Mapping rice paddy extent and intensification in the Vietnamese Mekong River Delta with dense time stacks of Landsat data

Caitlin Kontgis a,b,⁎, Annemarie Schneider a,b, Mutlu Ozdogan a,c

a Center for Sustainability and the Global Environment, Nelson Institute for Environmental Studies, University of Wisconsin-Madison, 1710 University Avenue, Madison, WI 53726, USA
b Department of Geography, University of Wisconsin-Madison, 550 N. Park Street, Madison, WI 53706, USA
c Department of Forest and Wildlife Ecology, University of Wisconsin-Madison, 1630 Linden Drive, Madison, WI 53726, USA

A R T I C L E   I N F O

Article history:
Received 7 February 2015
Received in revised form 29 July 2015
Accepted 6 August 2015
Available online xxxx

Keywords:
Landsat;
Rice paddy mapping;
Decision trees;
Remote sensing;
Area estimation;
Enhanced vegetation index;
Normalized difference water index

A B S T R A C T

Rice is a staple food crop for the majority of the world’s population, yet paddy fields are threatened by urban expansion, climate change, and degraded agricultural land. Vietnam, one of the largest exporters of rice globally, grows most of its rice in the Mekong River Delta, but this low-lying and heavily populated area is susceptible to major land cover changes. To properly monitor and manage rice crops in this region, remote sensing with satellite imagery has been particularly useful; however, most efforts to map regional paddy area utilize coarse resolution MODIS or AVHRR data due to their high temporal frequency. Because the average size of a rice paddy field in the region is smaller than a coarse resolution pixel, we map the Mekong study area using finer-scale Landsat data collected across multiple growing seasons. First, we exploit dense Landsat time stacks for circa 2000 and circa 2010 to map rice paddy extent using vegetation trajectories, then combine these pixel-based rice maps with image-based segments to generate a polygon-based rice map. The results show that this method can map rice paddies with over 90% overall accuracy (and errors of omission and commission ranging from 6 to 25%) at a finer spatial resolution than previous efforts. Next, we differentiate between single-, double-, and triple-cropped rice paddies in the delta using a supervised classification based on exemplars of these different cropping trends. From circa 2000 to circa 2010, we find that triple-cropped rice fields have nearly doubled in area from one-third to nearly two-thirds of paddy area. Our work also highlights the importance of scenes that capture flooded fields, and the utility of cloud-covered scenes within the dense time stacks of data, to achieve higher classification accuracies. Methods to map rice paddies are vital to understanding the sustainability of these agricultural systems, and the work presented here makes strides toward routine monitoring at a field-level resolution.

© 2015 Elsevier Inc. All rights reserved.

1. Introduction

Globally, rice (Oryza sativa) is one of the most widely harvested and nutritionally important food sources. It accounts for 20% of the world’s calorie supply, and has been called “the most important food crop for the poor,” since it is the staple food item for over 900 million people who subsist on less than 1.25 USD per day [Dawe, Pandey, & Nelson, 2011]. Among the five countries with the largest exports of rice (Thailand, Vietnam, India, United States, Pakistan) account for approximately four-fifths of trade [Dorosh & Wailes, 2010]. To this end, changes to rice output in any one of these five countries greatly alter the global rice market, which in turn has ramifications for vulnerable populations dependent on rice for sustenance. Further affecting rice-reliant populations, recent research demonstrates that rising temperatures may decrease rice yields (even when fertilization from higher levels of CO2 is considered), while rising seas could inundate low-lying rice fields [Long, Ainsworth, Leakey, Nössberger, & Ort, 2006; Peng et al., 2004; Wassmann et al., 2009]. Because rice is so susceptible to impending climate change, rapid and efficient monitoring of these crop systems is critical for (1) continued management of paddy fields; (2) understanding the efficacy of different policy levers; and (3) illuminating broader issues of food security, agricultural sustainability, and ongoing climate change.

Land use changes coupled with rapid economic development have affected rice production globally, but these impacts are best exemplified in Vietnam. As a result of economic reforms instituted in the late 1980s targeting the agricultural sector, rice production in Vietnam grew by >25% between 2000 and 2011, while the area devoted to rice paddies remained nearly constant [General Statistics Office of Vietnam, 2011]. Much of this change can be attributed to the Green Revolution including new rice varieties, irrigation, pesticides, and fertilizers, all of which were adopted by Vietnamese farmers after market liberalization [Cleaver, 1972; Hazell, 2010]. For example, while none of Vietnam’s rice area was planted with modern rice varieties during the Communist regime, now 89% of rice seeds are modern varieties that provide higher...
yields and shorter growing periods, allowing for multiple harvests per year (Dawe, 2002). Moreover, in 1975, only 16% of agricultural land was irrigated and 50.7 kg/ha of fertilizer was applied. By 1995, these numbers had reached nearly 30% for irrigation and 214 kg/ha for fertilizer application (Rosegrant & Hazell, 2000). These advancements have allowed the Mekong Delta to emerge as one of the most intensely cultivated regions in the world for rice, and one of only a handful of places to practice triple-cropping rice agriculture (Dawe et al., 2010).

While this vast transformation has pushed Vietnam to the world stage, the sustainability and effective management of rice production in the Mekong Delta requires accurate information on location and cultivation practices at the scale of individual fields. Remote sensing, either alone or in combination with annual surveys, offers the most practical means of monitoring rice production in the region. Synoptic coverage, repeated observations, and the archival nature of observations are some of the key attributes of remote sensing that allow rice monitoring over large areas. What remains unknown, however, is the timing and the number of observations required for satellite imagery to accurately capture the vegetation phenology associated with rice cultivation. This issue is particularly important in the Mekong Delta where the monsoonal climatic conditions lead to persistent cloud cover that obscures the land surface over a significant period of time.

With these issues in mind, the goal of this work is to develop a change detection strategy involving dense time stacks of Landsat (30-m) data to monitor changes in rice paddy area as well as shifts in the number of annual harvests for two time points (circa 2000, circa 2010) at fine spatial scales in the Mekong Delta. In the first step, we take advantage of object-oriented image classification algorithms to map the areal extent of rice paddy agriculture. Second, we develop a new approach to count the number of annual rice harvests using an incomplete – due to clouds and Scan Line Corrector (SLC)-off gaps – set of observations to determine rice-cropping intensity. When applied to the two time periods (2000, 2010), these methods allow us to determine how the areal extent of rice paddy has changed, and whether crop rotations intensified over the ten-year study period. We hypothesize that the number of cropping rotations has increased due to the adoption of Green Revolution technologies (e.g. rice cultivars with a short 100-day growing period), and the construction of dikes and sluices that protect paddy from flooding and saltwater intrusion.

This study builds upon methods developed for MODIS data that exploit spectral–temporal indices from satellite data to map rice paddy and cropping rotations (Xiao et al., 2005; Xiao, et al., 2006). Since the average area of rice fields in the Mekong Delta is only one half-hectare, it is critical to map this region at a finer resolution than past efforts with MODIS. Therefore, we expand and refine the MODIS approach in three ways: (1) we adapt the method for use with Landsat data by segmenting the image data into individual fields to provide greater spatial detail for mapping complex, highly heterogeneous landscapes; (2) we map multiple time points (2000, 2010), and test the ability of our approach to capture change through time, and most critically, (3) we distinguish whether fields are cultivated once, twice, or three times per year to provide valuable information on changes in cropping.

Fig. 1. The Mekong River Delta in Southern Vietnam. In this study we focus on the 12-province area within Landsat path 125, row 53.
rotations during the last decade. To the best of our knowledge, this study is the first to differentiate single-, double-, and triple-cropped rice paddy fields at a regional-level with Landsat spatial resolution. This work produces maps that are useful for informing science and policy researchers regarding food security and environmental degradation in the Mekong region. In addition, we conduct analyses to better understand how to more efficiently map rice at this resolution in the future by assessing which scenes are most effective for discriminating between single-, double-, and triple-cropped fields. Finally, our method relies solely on remote sensing data, so it can be widely applied to generate up-to-date data on rice paddy extent and cropping rotations in regions dominated by rice paddy agriculture even under persistent cloud cover.

2. Study area

The study area (Landsat path 125, row 53) bounds nearly the entire delta from Ho Chi Minh City in the northeast to the Ca Mau peninsula in the south (Fig. 1). Fourteen provinces are fully or partially encompassed within the study area. The study area abuts Cambodia to the west and the East Sea to the east, and includes the Mekong River as well as the Bassac River, the main distributary of the Mekong in the region. Weather patterns are driven by the East-Asian summer monsoon, creating a dry season from December through April, and a rainy season from May to November. Annual rainfall in the Mekong Delta averages nearly 2000 mm, while temperatures are warm all year (20 °C–35 °C, sometimes exceeding 40 °C). Soils in the region are high in sediments due to the annual flooding of the Mekong River, which begins at the end of September and lasts until late October or early November. The low-lying region has an average elevation of two meters above sea level, rendering it particularly susceptible to climate change-induced storm surge and sea level rise.

The Vietnamese Mekong Delta is home to >17 million people (20% of Vietnam’s population), and has a population density of 429 persons/km² (General Statistics Office of Vietnam, 2011b). Over 60% of the land area is devoted to agricultural activities, and the predominant crop is rice paddy (General Statistics Office of Vietnam, 2011a). More than half of domestic rice production and nearly 90% of rice for export are produced in the Mekong Delta, with 65% of the income share coming from farming activities, 43% of which is attributable to rice (Pandey, Paris, & Bhandari, 2010; Thanh & Singh, 2006; Wassman et al., 2010). Recent research indicates that yields have reached their maximum potential, and even with improved management techniques, rice paddy fields in the Vietnamese Mekong Delta are unlikely to attain greater yields (Licker et al., 2010). As a result, ongoing changes in the region will affect the livelihoods of farmers, the profitability of rice, and the sustainability of rice paddy systems.

3. Methodology

3.1. Overview

While there is a long history of mapping rice with satellite data (Frolking et al., 2002; Huke & Huke, 1997; Knox, Matthews, & Wassmann, 2000; Le Toan, Ribbes, Wang, Flourey, & Ding, 1997; Liew, Kam, & Tuong, 1998; Ribbes & Le Toan, 1999; Yang, Shen, Li, Le Toan, & He, 2008), optical remote sensing efforts have shown particular promise, especially those that extract spectral indices from MODIS data to differentiate rice from other crops (e.g., normalized difference vegetation index (NDVI), enhanced vegetation index (EVI), normalized difference water index (NDWI)) (Gumma, Nelson, Thenkabail, & Singh, 2011; Gumma, Thenkabail, Maunahan, Islam, Nelson, 2014; Sakamoto et al., 2005; Sakamoto, Van Nguyen, Ohno, Ishitsuka, & Yokozawa, 2006; Xiao et al., 2005; Xiao, et al., 2006). While higher spatial resolution Landsat imagery has also been used, these data have only been applied to relatively small study sites and locations where cloud cover is minimal (Okamoto & Fukuhara, 1996; Van Niel, McVicar, Fang, & Liang, 2003). Studies within the tropics that use Landsat data have required data fusion with RADAR to overcome missing observations due to cloud cover (Okamoto & Kawashima, 1999), or have also incorporated data with resolutions as fine as 1–4 m (Sakamoto, Van Nguyen, Ohno, Ishitsuka, Yokozawa, 2006). Here, we aim to rely solely on Landsat data for rice paddy mapping.

Unlike other crops, rice spends a portion of its phenologic cycle submerged in water (Wade, Fukai, Samson, Ali, & Mazid, 1999). Rice has three distinct growing phases: the sowing-transplanting period, the growing period, and the after-harvest period (Le Toan et al., 1997).
Table 1
The thresholds used to determine rice paddy extent. Rice paddies are alternately flooded and dry throughout the season, which is captured by a high NDWI standard deviation. Likewise, fields are bare just after rice is planted, then become bright green as the plant matures, which is captured by a high EVI standard deviation. A mid-range mean EVI separates rice paddy from areas that are vegetated all year long. The combination of these three variables delineates rice paddy area at the pixel-level.

<table>
<thead>
<tr>
<th>Spectral index variable</th>
<th>Circa 2000 image</th>
<th>Circa 2010 image</th>
</tr>
</thead>
<tbody>
<tr>
<td>NDWI standard deviation across Landsat stack2</td>
<td>0.12–0.437</td>
<td>0.12–0.425</td>
</tr>
<tr>
<td>NDWI standard deviation value range:</td>
<td>(0–0.437)</td>
<td>(0–0.425)</td>
</tr>
<tr>
<td>EVI standard deviation across Landsat stack6</td>
<td>0.21–0.527</td>
<td>0.18–0.794</td>
</tr>
<tr>
<td>EVI standard deviation value range:</td>
<td>(0–0.527)</td>
<td>(0–0.794)</td>
</tr>
<tr>
<td>EVI mean across Landsat stack2</td>
<td>0.1–0.6</td>
<td>0.2–0.75</td>
</tr>
<tr>
<td>EVI mean value range:</td>
<td>(−0.241–0.848)</td>
<td>(−0.287–1.058)</td>
</tr>
</tbody>
</table>

1 NDWI refers to the normalized difference water index; EVI refers to the enhanced vegetation index.
2 Separate time stacks were used for circa 2000 and circa 2010.

each of which has a distinct spectral signature that can be exploited for remote sensing analysis. During the sowing-transplanting phase, rice fields are flooded, and NDWI can be used to document this phenologic stage since vegetation strongly reflects near-infrared radiation ($P_{nir}$), while water absorbs both near-infrared and shortwave-infrared radiation ($P_{swir}$) (Gao, 1996):

$$NDWI = \frac{P_{nir} - P_{swir}}{P_{nir} + P_{swir}}$$

(1)

After sowing, rice matures into its vegetative stage, which can be mapped using NDWI in conjunction with EVI (Liu & Huete, 1995; Xiao et al., 2005; Xiao et al., 2006). EVI is computed as follows:

$$EVI = 2.5 \times \frac{P_{nir} - P_{red}}{P_{nir} + 6 \times P_{red} - 7.5 \times P_{blue} + 1}$$

(2)

where $P_{nir}$ again corresponds to near-infrared radiation, $P_{blue}$ to visible blue radiation, and $P_{red}$ to visible red radiation. Xiao et al. (2005), Xiao et al. (2006) showed that a pixel can be identified as flooded if NDWI + 0.05 ≥ EVI. In tropical regions like Southeast Asia, rice takes approximately 100 days to complete a full growth cycle, so three cycles per year are possible (Le Toan et al., 1997). Consequently, spectral indices can reveal the timing of rice phenology, which in turn, can be used to determine whether fields have single-, double-, or triple-cropped rice paddy.

3.2 Data

With data acquisition beginning in 1972 and continuing today (Williams, Goward, & Arvidson, 2006), Landsat data are ideal for assessment of long-term changes in cropped area and cropping rotations. Since 2003, all Landsat 7 (ETM+) images have data gaps due to a malfunction in the Scan Line Corrector (SLC) (Chander, Markham, & Helder, 2009), resulting in ~22% loss of scene area (Storey, Scaramuzza, Schmidt, & Basi, 2005). Despite these data gaps, Landsat 7 imagery is still widely used in land cover and land use change studies through a variety of gap-filling and dense time stack approaches (Castrence, Nong, Tran, Young, & Fox, 2014; Hilker et al., 2009; Huang et al., 2010; Schneider, 2012). In this work, we acquire data from Landsat 5, Landsat 7, and Landsat 8 for circa 2000 and circa 2010 (65 images total, Fig. 2). Prior to analysis, clouds and cloud shadows for all images were masked using Fmask (Zhu & Woodcock, 2012). The images were corrected to Top of Atmosphere (TOA) reflectance, which uses the Earth–Sun geometry to define the reflectance measured by a satellite without performing correction for atmospheric effects. A full atmospheric correction is not necessary in this work since we are comparing a time series of images across a single Landsat footprint, rather than analyzing each individually (Song, Woodcock, & Seto, 2001). Image correction to TOA is sufficient since this analysis relies on the temporal patterns of EVI and NDWI, rather than the absolute values of these indices. Because the variability in the temporal dynamics of rice is far greater than any variability introduced by atmospheric effects, a full atmospheric correction would be superfluous.

Rice in the region is planted in three-year cycles, with three growing seasons per cycle, totaling nine seasons. However, many farmers leave a field fallow during one or two growing seasons to avoid soil degradation. As a result, triple-cropped rice is planted in at least seven of nine growing seasons, double-cropped rice is planted in at least six growing seasons, and single-cropped rice is planted during three growing seasons in the three-year cycle. To capture these trends, we rely on three years of data for each time point (2000–2002, 2010–2012), taking additional images from ±1–2 years to fill months and seasons where no scene was available during each period of interest (Fig. 2). Using multiple years of imagery provides additional data points that help overcome missing...
observations due to clouds, low temporal resolution, and SLC-off errors. In this strategy, we assume that there was no major land use change in areas planted with rice during the images used for each time period. More supplemental scenes were used for circa 2010 than circa 2000 to compensate for missing data as a result of SLC-off gaps. We gathered all available Landsat 8 data (2013–2014) for this strategy, we assume that there was no major land use change in areas planted with rice during the images used for each time period. More supplemental scenes were used for circa 2010 than circa 2000 to compensate for missing data as a result of SLC-off gaps. We gathered all available Landsat 8 data (2013–2014) for

Fig. 4. After mapping rice vs. non-rice areas using spectral indices, we segmented the pixel-based map into individual fields. Based on the classification of the majority of pixels within the polygon, each field was assigned the label rice paddy agriculture, or non-rice paddy agriculture (non-rice agriculture, mangroves, water, or built-up land).

Fig. 5. To discriminate between single-, double-, and triple-cropped rice paddies, we tallied the number of EVI peaks within mapped paddy areas. Here, we visualize what an ideal, pixel-based EVI trajectory would look like over a three-year period, and what a single pixel of our Landsat data looks like over a three-year period for circa 2000 and circa 2010. Because the two time period image stacks contained different numbers of images, how cropping cycles were defined varied.

note: EVI refers to the enhanced vegetation index
the study area and used a ‘brute-force’ approach in which all data were provided to the classification algorithm to take advantage of as many observations as possible for each pixel (Schneider, 2012).

During a field visit to the Mekong Delta in February and March of 2014, we acquired ancillary data, including land use maps and expert knowledge for six Mekong provinces (Can Tho, Soc Trang, Anh Giang, Bac Lieu, Tra Vinh, and Vinh Long) (Fig. 1). We met with colleagues at Can Tho University and government officials who provided information on approximate planting dates for triple-, double-, and single-cropped rice. These data were used in conjunction with high-resolution Google Earth imagery and medium-resolution Landsat imagery to collect sample sites for input into a supervised classification algorithm (Section 3.3, Step 2). Local land use maps were also compared to our satellite-derived maps, to assist in both calibration and validation of the methods.

3.3. Technical approach

For this analysis, we mapped the study area in a hierarchical fashion: (1) we first discriminated rice extent from all other land cover types; and (2) within areas defined as rice, we differentiated between single-, double-, and triple-cropped fields. Each of these steps was performed for two time points: circa 2000 and circa 2010. In step (3), we assess the accuracy of the rice extent and crop cycle maps, and finally in step (4) we compare the maps across the two dates to understand ongoing changes in the region. We describe the detailed methodology for each step in the following sections.

3.3.1. Step one: spectral–temporal rice signatures for mapping paddy extent

We adapted the parameters defined by Xiao et al. (2005); Xiao, et al. (2006) to map the location of rice. For this step, we are simply interested in whether a field is planted with rice at any point during the growing cycle. Since fields are systematically flooded and vegetated throughout the year (Section 3.1), we captured this variability by selecting areas with a high standard deviation across all years for both EVI and NDWI (Table 1). Next, we delineated areas with a mid-range mean EVI (Table 1) to separate rice from areas that are (a) vegetated all year long (e.g., mangroves, orchards, household vegetable farms, etc.), which have a correspondingly high annual mean EVI; and (b) non-vegetated (e.g., salt flats, buildings, aquaculture, rivers, etc.), which exhibit a very low annual mean EVI. Thresholds for EVI and NDWI were chosen for circa 2000 and circa 2010 based on visual interpretation of the imagery for that period and expert knowledge acquired during fieldwork. For each time point, circa 2000 and circa 2010, we combined the three variables (high standard deviation EVI, high standard deviation NDWI, and mid-range mean EVI) to delineate areas as “rice paddy”, and all other areas were placed in a generalized “non-rice” class (Fig. 3).

Because of the relatively high level of spectral–temporal variability and predominance of missing data across the study area, the raw output from the NDWI/EVI thresholds contained noise such that fields were not fully mapped at the pixel level. To remedy this, we turned the pixel-level assessment of rice into a field-based (or object-based) assessment by combining the pixel-based rice map with a map segmented into individual fields (Fig. 4). The premise of this approach is to cluster spatially-adjacent pixels into objects based on their spectral similarity, so that classification can be performed on the object, rather than pixels (Huang & Zhang, 2008). Segmentation of the data into objects is relatively easy in our paddy-dominated study area since the extensive dyke irrigation systems have distinct boundaries readily identifiable at the 30-m Landsat scale (Fig. 4, raw Landsat panels). We segmented a single, cloud-free, dry-season image for each time point (12/10/2001; 12/09/2009) using the iterative, non-parametric mean-shift segmentation algorithm (Cheng, 1995; Fukunaga & Hostetler, 1975; Huang & Zhang, 2008). This algorithm detects maxima of local density functions, and then iteratively shifts data points to their spatial neighborhood average (Cheng, 1995; Fukunaga & Hostetler, 1975). For remote sensing applications, the pixels in each spatial neighborhood act as data points in feature space, and they are clustered into the same polygon if they converge on the same final position after the mean-shift iterations have completed, which occurs when all points have been clustered into an object (Cheng, 1995; Comaniciu & Meer, 2002; Friedman, Netanyahu, & Shoshany, 2003).

To implement the mean-shift segmentation algorithm, we utilized the open-source Orfeo Toolbox software, and designated a spatial radius...
From circa 2000 to circa 2010, triple-cropped rice fields expanded by nearly 30% of total paddy area, with the majority of newly converted areas coming from fields that were double-cropped in circa 2000. At both time points, single crop rice took up only a small percentage of the total paddy area (<5%).
of five pixels (French Space Agency (CNES), 2014). We constrained the minimum mapping unit to a half-hectare, or the average farm size in Vietnam, meaning that no polygon was composed of fewer than six Landsat pixels. It is safe to assume that the number of annual harvests for rice paddy is uniform within a field, so we combined the pixel-based rice map with the image-based segments to generate a polygon-based rice map using the majority rule (i.e., for each field, a majority vote of classes was taken, and the winning class was assigned to all pixels in the field). The result is a field-based classification with all polygons labeled as “rice” or “non-rice”.

3.3.2. Step two: rice phenology for mapping crop cycles

The distinct spectral signature (inundation, full vegetation, bare soil) associated with each growing stage of the rice phenological cycle allowed us to use EVI trajectories to map not only the location of rice paddies (Step 1), but also the number of annual harvests per field. To differentiate between single-, double-, and triple-cropped fields for each time period, we first chose example polygons of each cropping type using data from locations visited during field work, as well as information from land use maps and Landsat EVI trajectories. For these exemplars, we counted the number of EVI peaks with a value greater than 0.4; this threshold was chosen based on photo interpretation and by assessing EVI trajectories from the sample polygons. Ideally, the trajectory curves for triple-cropped rice over a three-year period would include nine EVI peaks, double-cropped rice would include six EVI peaks, and single-cropped rice would include three EVI peaks (Fig. 5, left). However, the variability in field planting dates, decisions by individual farmers to leave fields fallow, and available image data led to considerable variation in EVI peak values. Therefore, our definition of cropping rotations varied from the ideal trajectory as follows:

Fig. 8. Cropping cycle changes: circa 2000–circa 2010.
formation gain ratio (Quinlan, 1993, 1996). The ef-

computer classification algorithm recursively splits training data into increasingly similar

groups based on statistical testing of the exemplars. This extensively-

satellite data (Friedl & Brodley, 1997; Hansen, Dubayah, & Defries, 1996),

ulesr's accuracies are particularly adept at discriminating these types of complex patterns in

els that were known to be single-, double-, or triple-cropped rice

fields and circa 2010, it was applied to each corresponding image stack to pro-

Once the decision tree was constructed for each time point, circa 2000 and circa 2010, it was applied to each corresponding image stack to produce a final classified map of annual rice paddy harvests.

3.3.3. Step three: Accuracy assessment of rice paddy and crop cycle mapping

The accuracy assessments of the resulting maps from Step 1 and Step 2 were conducted separately, yet in a similar fashion. Since maps were created independently for each time point, separate accuracy assessments were also conducted for each time period (circa 2000 or circa 2010) within each step. For each time point of Step 1 and Step 2, we compared the maps to a set of randomly selected, independently labeled polygon-based 'truth sites,' and then quantified areas of agreement (Congalton, 1991). For Step 1, a proportional, random sample of 200 polygons were independently labeled as 'rice' or 'non-rice' by utilizing the initial Landsat images, high-resolution Google Earth imagery, and maps provided by officials in Vietnam. For Step 2, a similar approach was used to assess accuracy. A stratified random sample of 245 polygons was independently labeled by utilizing EVI trajectories from the Landsat image stacks, original Landsat images, and local knowledge gathered during fieldwork.

Once we completed the accuracy assessments for each map, we followed recommendations set forth by Olofsson et al. (2014) to adjust our areal estimates based on the accuracy results. The accuracy-adjusted methods provide greater transparency for estimates of area in land cover/use maps by using high-quality reference information to estimate or adjust the actual land cover class areas (Olofsson et al., 2014). The proportions of each class from the maps generated in Steps 1 and 2 are used as weights to standardize their respective user's accuracies. The standardized user’s accuracies are then summed for each reference class to obtain a value for the estimated proportion of samples (EPS) in the class (i.e., the areal proportions of each class according to the distribution of the reference data). This value is then multiplied by the total number of hectares in the classification to obtain the adjusted estimate for each class area.

Table 3

<table>
<thead>
<tr>
<th>Class</th>
<th>Triplet-cropped</th>
<th>Double-cropped</th>
<th>Single-cropped</th>
<th>Total</th>
<th>User’s accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. Circa 2000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Classified map of</td>
<td>59</td>
<td>19</td>
<td>1</td>
<td>79</td>
<td>74.7%</td>
</tr>
<tr>
<td>cropping cycles</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Double-cropped</td>
<td>35</td>
<td>101</td>
<td>1</td>
<td>137</td>
<td>73.7%</td>
</tr>
<tr>
<td>Single-cropped</td>
<td>9</td>
<td>6</td>
<td>14</td>
<td>29</td>
<td>48.3%</td>
</tr>
<tr>
<td>Total</td>
<td>103</td>
<td>126</td>
<td>16</td>
<td>245</td>
<td></td>
</tr>
<tr>
<td>Producer’s accuracy</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Triplet-cropped</td>
<td>57.3%</td>
<td>80.2%</td>
<td>87.5%</td>
<td>71.0%</td>
<td></td>
</tr>
<tr>
<td>Double-cropped</td>
<td>95</td>
<td>40</td>
<td>0</td>
<td>135</td>
<td>70.4%</td>
</tr>
<tr>
<td>Single-cropped</td>
<td>18</td>
<td>56</td>
<td>2</td>
<td>76</td>
<td>73.7%</td>
</tr>
<tr>
<td>Total</td>
<td>118</td>
<td>106</td>
<td>21</td>
<td>245</td>
<td></td>
</tr>
<tr>
<td>Producer’s accuracy</td>
<td>80.5%</td>
<td>52.8%</td>
<td>50.5%</td>
<td>69.4%</td>
<td></td>
</tr>
</tbody>
</table>

1 For circa 2000, single-cropped refers to ≤2 EVI peaks, double-cropped refers to 3–4 EVI peaks, and triple-cropped refers to 5+ EVI peaks.

2 For circa 2010, single-cropped refers to ≤4 EVI peaks, double-cropped refers to 5–7 EVI peaks, and triple-cropped refers to 8+ EVI peaks.

3.3.4. Step four: Mapping changes in rice extent and number of annual harvests

Once we determined the accuracies of the maps, we created a map of change to assess how rice area changed from circa 2000 to circa 2010, and we created a second map of change to illustrate how annual harvests were altered during the same period. To do so, we subtracted each set of maps to make a change map, a widely-used method that has been tested in a range of geographical environments (Singh, 1989). Because our accuracies are reasonable for each time point, we

Table 2

<table>
<thead>
<tr>
<th>Class</th>
<th>Rice</th>
<th>Non-rice</th>
<th>Total</th>
<th>User’s accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. Circa 2000</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Classified map of</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rice paddy extent</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rice</td>
<td>76</td>
<td>15</td>
<td>91</td>
<td>83.5%</td>
</tr>
<tr>
<td>Non-rice</td>
<td>24</td>
<td>85</td>
<td>109</td>
<td>78.0%</td>
</tr>
<tr>
<td>Total</td>
<td>100</td>
<td>100</td>
<td>200</td>
<td></td>
</tr>
<tr>
<td>Producer’s accuracy</td>
<td>76.0%</td>
<td>85.0%</td>
<td>80.5%</td>
<td></td>
</tr>
<tr>
<td>b. Circa 2010</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Classified map of</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>rice paddy extent</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rice</td>
<td>95</td>
<td>7</td>
<td>102</td>
<td>93.1%</td>
</tr>
<tr>
<td>Non-rice</td>
<td>5</td>
<td>93</td>
<td>98</td>
<td>94.9%</td>
</tr>
<tr>
<td>Total</td>
<td>100</td>
<td>100</td>
<td>200</td>
<td></td>
</tr>
<tr>
<td>Producer’s accuracy</td>
<td>95.0%</td>
<td>93.0%</td>
<td>94.0%</td>
<td></td>
</tr>
</tbody>
</table>

(a) Triple-cropped rice: polygons with five or more EVI peaks (circa 2000), or eight or more peaks (circa 2010);
(b) Double-cropped rice: polygons with three-four EVI peaks (circa 2000), or five-seven peaks (circa 2010); and,
(c) Single-cropped rice: polygons with fewer than two EVI peaks (circa 2000), or fewer than four peaks (circa 2010).

These definitions were established after visually analyzing EVI peaks for fields that were known to be single-, double-, or triple-cropped rice paddies. In this approach, each time period uses a different number of


estimating multiple classifiers while forcing the classifier to focus on difficult classes (Quinlan, 1996). The final classification is produced by an accuracy-weighted vote across all the classifications (Quinlan, 1996).
feel confident that the resulting change maps represent the patterns and amounts of change in the region.

3.4. Assessing feature importance in Landsat dense time stacks

This research aims to assess the efficacy of Landsat dense time stacks for mapping the location and annual harvests of rice in Vietnam. While there were greater than 25 images available in each time period, the number of high-quality observations for each pixel was highly variable. In locations with persistent cloud cover such as the Mekong Delta, missing observations are considered one of the biggest obstacles to mapping rice paddy agriculture with satellite data (Currey, Fraser, & Bardsley, 1987). We therefore evaluate the impact of feature selection – both the reduction of data points and the inclusion of specific features – by performing two sets of experiments in which different numbers and combinations of inputs were tested for their ability to discriminate single-, double-, and triple-cropped rice for the circa 2000 time point.

In the first experiment, we focused on the removal of features from the full, 25-scene dataset. To do this, we removed one Landsat scene (i.e., two features, EVI and NDWI for that date) at random, and tested the remaining scenes as input to the boosted decision tree classifier in a 20-fold cross-validation. We repeated this experiment sixteen times, each time removing an additional scene from the image stack (i.e., in the second round, two scenes/four features were removed at random, in the third round, three scenes/six features were removed at random, and so on). These results were then compared with the accuracies and errors of the maps generated with the full stack of features to determine whether a reduced image stack could achieve comparable accuracies.

In the second experiment, we evaluated the importance of specific features for mapping rice crop cycles using random forests (Breiman, 2001). The random forest algorithm draws multiple bootstrap samples from the training data to estimate many classification trees. Each tree is grown with a randomized set of predictors, a portion of which are randomly permuted. Here, the random permutation of a feature mimics the absence of that feature from the model. Random forests – an extension of the boosted decision trees used in this work – provides a measure of feature importance that is the difference between prediction accuracy (i.e., the number of observations correctly classified) before and after permuting the feature, averaged over all trees (Breiman, 2001). A high increase or decrease in accuracy denotes the relative importance of that feature. For this work, an ensemble of 1000 trees was grown by the random forest algorithm (implemented in Matlab, Jaintilal, 2009) using the 25-scene (50-feature) circa 2000 Landsat stack as input.

4. Results

4.1. Trends in rice expansion and intensification

The resulting maps of the Vietnamese Mekong River Delta illustrate that rice paddy agriculture has been and continues to be the dominant form of land cover in the region from 2000 to 2010 (Fig. 6). For both time points, rice paddies are the predominant land cover in the northwestern and western portions of the landscape, and they occur less frequently near the coastline and northeastern portions of the study area. This spatial distribution is expected, as the northeastern portion of the study area contains the southern extent of Ho Chi Minh City, the largest...
city in Vietnam. Recent growth in this burgeoning metropolitan area has likely affected soil and water quality, as well as land and water availability for agricultural activities. Salinity intrusion from seawater near the coastline renders this area much more amenable to aquaculture than rice agriculture, which helps explain the low density of rice paddy fields in the eastern portion of the study area. Overall, the absolute amount of rice paddy area changed from 1,184,384 ha to 1,280,175 ha from circa 2000 to circa 2010, an increase of approximately 100,000 ha. These values refer to fields planted with rice at least once during the growing cycle, rather than gross paddy area (total harvested amount summed over the three seasons of a growing cycle).

With respect to spatial patterns, several trends are apparent. First, rice paddy area expanded in coastal provinces from 2000 to 2010, particularly in Tra Vinh and Tien Giang (Fig. 6C, eastern portion). This result may seem counterintuitive given the vulnerability of these areas to salinity intrusion and storm surge that make rice expansion risky in dynamic coastal regions. However, sluice construction in the Mekong Delta has boomed since 2000, which may explain why rice is grown closer to the coast in circa 2010 (Vietnamese government officials, personal communication). The opposite spatial pattern is evident for the expansion of fields with increased annual harvests: triple-cropping became more widespread in provinces farthest from the coast, specifically An Giang, Dong Thap, and Long An (Fig. 7). Though intensive farming practices (i.e. multiple harvests per year) are widespread at both time points, triple-cropped fields are only the dominant land cover in circa 2010, consisting of 62% of paddies (compared with 35% in 2000). The transition to triple-cropped fields generally comes at the expense of double-cropped fields, which shrunk from 64% to 33% of paddy area (Figs. 7 and 8). This change from double- to triple-cropped fields is possible in the far western reaches of the delta thanks to sophisticated dike and canal systems operated by the Vietnamese government, which have been enhanced and augmented since a series of flood-related disasters in 2000 (Disaster Management Unit, 2004; Kazama, Muto, Nakatsuji, & Inoue, 2000; Sakamoto et al., 2006). At both time points, single-
cropped rice represents the smallest proportion of paddy fields, with the majority clustered in the northwestern corner of the study area (near or across the Cambodian border). These spatial patterns align with government-produced provincial land use maps of the region, as well as other studies that have used MODIS data to track the intensification of rice farming in the region (Chen, Son, & Chang, 2012).

4.2. Map accuracies

The circa 2000 map of rice paddy extent has an overall accuracy of greater than 80%, while the circa 2010 map has an overall accuracy of 94%, indicating that the object-oriented NDWI/EVI threshold method for delineating rice paddy agriculture produces strong agreement between the maps and ground reference information (Table 2). The circa 2010 rice paddy map corresponds well with the reference data (user’s and producer’s accuracies both greater than 93% for the rice class), while the circa 2000 map slightly underestimates rice paddy area (producer’s accuracy 76%).

Our method for discriminating cropping rotations performs reasonably well at each time point, as indicated by overall accuracies of 71% and 69% for the circa 2000 and circa 2010 maps, respectively (Table 3). For circa 2000, double- and triple-cropped rice fields have reasonable user’s accuracies (73–75%), indicating that the commission errors for these two classes are minimal. Meanwhile, the producer’s accuracies for single- and double-cropped rice fields are high (roughly 87% and 80%, respectively), suggesting that the method captures these areas well and omits very few fields from these classes. However, the map underestimates triple-cropped rice fields in circa 2000 (producer’s accuracy, 57%); much of what should be classified as triple-cropped rice is instead labeled as double-cropped rice. Although still a small portion of the study area, the map overestimates single-cropped rice (user’s accuracy, 48%), suggesting that several double- and triple-cropped rice fields are mislabeled as single-cropped rice. A similar trend is apparent for the circa 2010 time point: we overestimate single-cropped rice fields, with a user’s accuracy of 60%. This pattern of overestimating single-cropped rice is to be expected; if data were particularly limited, cloud cover or SLC-off gaps, rice phenology over the study period would be represented by only a few EVI peaks, and the classifier would naturally label these as single-cropped fields. The circa 2010 map also underestimates double-cropped rice fields, with greater than 80% of omitted fields labeled as triple-cropped rice on our map.

4.3. Accuracy-adjusted estimates of area

Given the trends revealed by the accuracy assessment, we present the accuracy-adjusted areal estimates for each class in Table 4. The adjusted values indicate what the class proportions should be for each map and each time point given the reference information. The adjusted results show that we underestimate rice paddy extent in both circa 2000 and circa 2010 maps, though by a much smaller margin in the latter time point (omission of approximately 700,000 ha in circa 2000 versus approximately 100,000 ha in circa 2010) (Table 4A, B). This table also reiterates the pattern of systematic overestimation of single-cropped rice fields at each time point that is described above. The trends for triple-cropped fields – once adjusted for the ‘true’ landscape proportions represented in the reference data – still show that many areas intensified to triple-cropped rice, with an increase of 164,712 ha (±6314 ha) between the time points, and a corresponding drop in double-cropped fields of 92,715 ha (±7371 ha) (Table 4C, D). These changes are especially striking in light of the accuracy-adjusted area results for paddy areal extent, which indicate that rice paddy area actually shrunk by over 500,000 ha between circa 2000 and circa 2010 (Table 4A, B).

4.4. Comparison to available census data

To further validate our method, we used ancillary data sources to determine how accurately we captured rice paddy extent (Section 3.3, Step 1). Vietnam provides much of its recent annual census data online, allowing us to compare our circa 2010 mapped paddy area totals to the census-reported paddy area totals (Fig. 6B) (General Statistics Office of Vietnam, 2011c). In our maps, we classified a field as ‘rice’ even if it only elicited one rice harvest per year. The census discloses province-level paddy area totals for each of the three planting seasons separately: winter, spring, and autumn. Because we are interested in whether a field is planted with rice at any point during 2010, we extracted the highest value regardless of season for each province to compare to our mapped values. Since we limited our study area to one Landsat scene (path 125, row 53), some provinces are not completely represented in the maps (e.g. only 45% of Kien Giang province and 43% of An Giang province are included in the study area). Because of this, the census values needed to be standardized by the proportional areas for each province. These standardized census data show that our satellite-derived estimates of total paddy area are analogous to the totals reported by the Vietnamese government, thereby corroborating our NDWI/EVI threshold method for mapping total rice area (Fig. 9).

4.5. Feature importance in Landsat dense time stacks

While our method discriminated between single-, double-, and triple-cropped fields well, it may be possible to achieve the same levels of accuracy with fewer images or features (Fig. 10). When
an increasing number of scenes are removed at random from the image stack (experiment one), the results show that the overall map accuracy decreases only slightly, dropping from 85 to 75% accuracy (Fig. 10). The variability in the results for the 20 classification iterations increases substantially, however, as additional scenes are removed. This result suggests that certain scenes and/or features may affect the classification more so than others, since omitting some data leads to a minor decline while dropping other data leads to a large decline in accuracy.

The errors of class commission and omission from experiment one (Fig. 11) reveal trends similar to those for overall accuracy (Fig. 10). Again, it is possible to obtain low commission and omission errors (below 25%) for all classes, even with up to 16 dates removed from the image stack, yet there is also increasing variability in error amounts as more dates are removed. Specifically, single-cropped rice paddy produces the greatest variability in commission errors as an increasing number of scenes are removed, indicating that this class would likely be over-classified with a reduced image stack. This makes sense given that the training data for each class relies on the number of EVI peaks, and with fewer dates there are fewer EVI peaks for all classes, rendering double- and triple-cropped fields more likely to be classified as single-cropped fields. Likewise, triple-cropped rice paddy produces the greatest amount of variability in omission errors across the scene removal experiments, indicating that it would likely be under-classified with a reduced image stack since there would be fewer EVI peaks with which to distinguish these fields.

To better understand the impacts of specific images and features on the classification results, we explored the importance values obtained from the random forest analysis (experiment 2, Figs. 12, 13). As described in the methods, unitless importance values are derived from random forests by assessing how randomly removing features (e.g. NDWI and EVI for a specific image date) from decision tree construction affects the classification. Higher importance values denote that the removed features have major impact on the classification, while lower importance values denote a comparatively minor impact on the classification (Breiman, 2001). This analysis provided insight into the type of data that is most beneficial when attempting to distinguish between single-, double-, and triple-cropped rice fields. Interestingly, the results of the random forest analysis indicate that scenes where fields are flooded are more important than scenes that are not flooded, even when mostly cloudy scenes are included (Fig. 12A). When EVI is compared to NDWI, the average importance values for EVI scenes where fields are flooded are over three times higher than the average importance values of EVI scenes without flooding and NDWI scenes with and without flooding (Fig. 12B). Again, the level of cloudiness in a given scene appears to have minimal effect; although cloud cover ranged from 30 to 70% in the ‘cloud-covered’ scenes, these still provided cloud-free pixels that aided the classification. To further illustrate these trends, the importance values for all of the dates in the image stack (EVI and NDWI combined) were averaged by month (Fig. 13). The highest average values are recorded for months in the latter half of the calendar year. With the rainy season occurring from May–October, fields are inundated with water during and after this time period, and are more likely to have increased cloud cover. This result further highlights the fact that flooded fields are more important than non-flooded fields when trying to distinguish between single-, double-, and triple-cropped fields.

5. Discussion

The Vietnamese Mekong River Delta is one of the few places in the world that has the climate, infrastructure, and labor force to harvest rice three times annually. And though rice paddy agriculture is the dominant land cover across the region (55% in circa 2000 and 39% in 2010), average field size is only one half-hectare, so individual field processes occur on a very fine scale. From a remote sensing point of view, this makes characterizing rice trends in space/time incredibly challenging, and points to the necessity of adapting coarse spatial resolution data methods for data with higher resolutions, despite the more limited temporal trajectory available from medium-resolution sources (Landsat, SPOT, and the upcoming Sentinel-2). This research accomplishes that goal by (1) taking advantage of information from multiple spectral indices (EVI, NDWI); (2) incorporating an object-oriented approach to reduce within-field variability for mapping extent; (3) using multiple years of Landsat data for each time point to ensure sufficient temporal information for mapping extent and number of crop rotations; and (4) using decision rules on single-, double-, and triple-cropped patterns developed from field work and expert knowledge of the area. The results provide new, detailed information on rice-farming practices that were previously unavailable from remote sensing: rice has clearly intensified in the region, with triple-cropped fields expanding from approximately 34% to 62% of rice paddy agriculture over the ten-year study period.

A remaining challenge in mapping changes to rice-dominated areas is the availability of detailed field-level data for map validation. Though we utilized land use maps obtained during field visits to the Mekong Delta as well as freely accessible census data, these ancillary datasets are collected and distributed at a coarser spatial resolution than the field-level results produced by our method. While we could compare absolute values for rice paddy area (Fig. 9), it was impossible for the ancillary data to verify whether specific fields were correctly classified since the finest aggregation of available census data is at the province-level. In addition, it proved
extremely difficult to use non-satellite data to validate our maps of cropping rotations since census data do not identify number of annual harvests per field, as our maps do. Further, there is a mismatch in how data are reported, as well as a lack of temporal detail in both census and remote sensing data that would be requisite for this type of comparison. These issues notwithstanding, we feel confident that the resulting maps reflect the ongoing trends in the region because we focused heavily on collection and synthesis of field data and expert knowledge in all phases of the methodology.

This research demonstrates that our new, fine-scale method of mapping rice paddy agriculture works well even in cloudy, tropical regions like the Vietnamese Mekong River Delta, where missing observations are common, and where significant within-pixel mixing occurs due to the incredibly heterogeneous and complex nature of the landscape over small areas. Missing data due to SLC-off gaps and clouds led to overestimation of single-cropped rice, since many fields had fewer EVI peaks than expected. While cloud cover clearly affected these results, additional analysis of scene variable importance using random forest revealed that, despite high levels of missing data, the cloud-covered scenes played a critical role in mapping crop cycles since even these provide some glimpses at cloud-free pixels. If the method was applied in regions with additional available data (e.g. temperate areas with more cloud-free data), and/or in locations where fields are slightly larger and more homogenous (e.g. the USA), we expect map accuracies to increase.

As new medium resolution imagery (e.g. Landsat 8) and diverse data sources (e.g. RADAR, LIDAR) become increasingly available (many freely distributed), the work here can provide a foundation for future efforts to characterize rice and crop rotation changes through time and space. The June 2015 launch of the Sentinel-2 satellite by the European Space Agency, for example, will provide more Landsat-type data to map rice and crop cycles. Because the combined data from Landsat 7, Landsat 8, and Sentinel-2 will substantially increase the temporal availability of 30-m resolution data, the work presented here takes important first steps to adapting MODIS-based methods for higher resolution data. As a result, rice can be mapped at the field-scale and at multiple time points, achievements that have not been made in past work. By investigating the types of Landsat data most important for this type of mapping (Figs. 12 – 14), our research also makes strides toward routine, automated monitoring of rice paddy extent and management practices. Specifically, our results suggest that future efforts should focus on scenes with fields that are nearly all flooded – even if these scenes are cloudy – since these images were particularly instrumental in increasing map accuracies.

6. Conclusions

This research demonstrates that Landsat 30-m spatial resolution data is capable of mapping changes to rice paddy area and annual cropping cycles over time in the Vietnamese Mekong River Delta. Using dense time stacks of satellite imagery in conjunction with object-oriented segmentation, we were able to overcome limitations due to missing data from clouds, low temporal resolution, and SLC-off errors that often plague tropical applications of Landsat data. Specifically, we thresholded spectral indices to characterize segmented polygons as ‘rice’ or ‘non-rice’, and found that in the Mekong River Delta, rice expanded only slightly between circa 2000 and circa 2010. However, a supervised classification of rice cropping cycles using exemplars of different cropping rotations illustrated that nearly one-third of existing farms in the region transitioned from double- to triple-cropped fields, suggesting the intensification of farming practices. Finally, we conducted a random forest analysis, which showed that flooded scenes are the most important type of imagery for discriminating between single-, double-, and triple-cropped rice fields, even if these scenes are over 50% cloudy. To date, this work represents the highest spatial resolution time series analysis at the regional scale for rice paddy mapping, and it is the first to demonstrate the relative importance of flooded scenes. It is still unclear whether this method will work in landscapes that are less engineered than the Mekong River Delta, but we are confident that coupling Sentinel-2 data with Landsat 7 and Landsat 8 imagery will aid this task.

Acknowledgments

We would like to extend our gratitude to Dr. Van Pham Dang Tri and his students at Can Tho University for their insight and assistance. This work was supported by the NASA Earth and Space Science Fellowship Program — Grant NNX13AN51H. Dr. Ozdognak acknowledges partial support from the NASA NNX12DA001N-MEASURES Program awarded through a USGS cooperative agreement.

References


