HHS Public Access

Author manuscript

*Remote Sens Environ*. Author manuscript; available in PMC 2017 November 01.

Published in final edited form as:

*Remote Sens Environ*. 2016 November ; 185: 142–154. doi:10.1016/j.rse.2016.02.016.

# Mapping paddy rice planting area in northeastern Asia with Landsat 8 images, phenology-based algorithm and Google Earth Engine

**Jinwei Dong**[a,b,*], **Xiangming Xiao**[a,b,c,*], **Michael A. Menarguez**[a,b], **Geli Zhang**[a,b], **Yuanwei Qin**[a,b], **David Thau**[d], **Chandrashekhar Biradar**[e], and **Berrien Moore III**[f]

[a]Department of Microbiology and Plant Biology, University of Oklahoma, Norman, OK 73019, USA

[b]Center for Spatial Analysis, University of Oklahoma, Norman, OK 73019, USA

[c]Institute of Biodiversity Science, Fudan University, Shanghai 200433, China

[d]Google, Mountain View, CA, USA

[e]International Center for Agricultural Research in Dry Areas, Amman 11195, Jordan

[f]College of Atmospheric and Geographic Sciences, University of Oklahoma, Norman, OK 73019, USA

## Abstract

Area and spatial distribution information of paddy rice are important for understanding of food security, water use, greenhouse gas emission, and disease transmission. Due to climatic warming and increasing food demand, paddy rice has been expanding rapidly in high latitude areas in the last decade, particularly in northeastern (NE) Asia. Current knowledge about paddy rice fields in these cold regions is limited. The phenology- and pixel-based paddy rice mapping (PPPM) algorithm, which identifies the flooding signals in the rice transplanting phase, has been effectively applied in tropical areas, but has not been tested at large scale of cold regions yet. Despite the effects from more snow/ice, paddy rice mapping in high latitude areas is assumed to be more encouraging due to less clouds, lower cropping intensity, and more observations from Landsat sidelaps. Moreover, the enhanced temporal and geographic coverage from Landsat 8 provides an opportunity to acquire phenology information and map paddy rice. This study evaluated the potential of Landsat 8 images on annual paddy rice mapping in NE Asia which was dominated by single cropping system, including Japan, North Korea, South Korea, and NE China. The cloud computing approach was used to process all the available Landsat 8 imagery in 2014 (143 path/rows, ~3290 scenes) with the Google Earth Engine (GEE) platform. The results indicated that the Landsat 8, GEE, and improved PPPM algorithm can effectively support the yearly mapping of paddy rice in NE Asia. The resultant paddy rice map has a high accuracy with the producer (user) accuracy of 73% (92%), based on the validation using very high resolution images and intensive field photos. Geographic characteristics of paddy rice distribution were

*Corresponding authors at: Department of Microbiology and Plant Biology, University of Oklahoma, 101 David L. Boren Blvd., Norman, OK 73019, USA, jinwei.dong@ou.edu (J. Dong), xiangming.xiao@ou.edu (X. Xiao). *URL's:* http://www.eomf.ou.edu (J. Dong), http://www.eomf.ou.edu (X. Xiao).

analyzed from aspects of country, elevation, latitude, and climate. The resultant 30-m paddy rice map is expected to provide unprecedented details about the area, spatial distribution, and landscape pattern of paddy rice fields in NE Asia, which will contribute to food security assessment, water resource management, estimation of greenhouse gas emissions, and disease control.

**Keywords**

Paddy rice; Landsat 8; Phenology- and pixel-based algorithm; Cloud computing; Google Earth Engine

## 1. Introduction

Paddy rice planting area information is important for a series of issues influencing human well-being, ranging from human food security, water use, climate change, to disease transmission. First, rice is a staple grain and food source for more than half of the global population and paddy rice accounts for more than 12% of global cropland area (Elert, 2014; FAOSTAT, 2010). Second, paddy rice is the largest water-consuming crop and the water use efficiency for rice planting is related to global water security in the future (Salas et al., 2007). Third, effects of paddy rice agriculture on methane emission and future climate change cannot be ignored because it accounts for more than 10% of the total methane flux in the atmosphere (Ehhalt et al., 2001; Sass & Cicerone, 2002). Finally, paddy rice agriculture is related to avian influenza virus transmission as paddy rice fields are important habitats for free-ranging ducks and wild waterfowls (Gilbert et al., 2008, 2014). The existing remote sensing based global or national land cover and land use datasets do have cropland layers, e.g., the MCD12Q1 from the Moderate-resolution Imaging Spectroradiometer (MODIS) (Friedl et al., 2002, 2010), the GlobCover from the MEdium Resolution Imaging Spectrometer (MERIS) (Arino et al., 2008), and the Finer Resolution Observation and Monitoring-Global Land Cover (FROM-GLC) from Landsat (Gong et al., 2013). However, the general cropland category in these products did not have paddy rice information. Despite the significance of paddy rice information, regional or global paddy rice maps for the scientific communities are rarely available, especially for high latitude regions in northeastern (NE) Asia with a large amount of newly reclaimed rice planting area.

The difficulty of paddy rice mapping arises largely from the similar spectral characteristics between paddy rice and other land covers as well as frequent clouds and cloud shadows in the rice planting regions. Existing efforts have used different data sources and approaches for paddy rice mapping and the methods can be generally divided into three categories. First, some studies mapped paddy rice using the images from certain stages (e.g., transplanting or early growing season) and the image statistics-based approaches, e.g., the supervised classifiers like Maximum Likelihood (Oguro et al., 2001), Support Vector Machine (Zhang, Wang, Wu, Qi, & Salas, 2009), and unsupervised classifiers like the Iterative Self-Organizing Data Analysis Technique (ISODATA) (Pan, Uchida, Liang, Hirano, & Sun, 2010; Thi, De Bie, Ali, Smaling, & Chu, 2012). This approach would have various outcomes when using varying images or training samples from different phases or regions. It is often

difficult to extend the classifier rules and parameters, as all these methods are region- and phase-dependent due to the spectral variability in periods and regions (Dong et al., 2015; Zhong, Gong, & Biging, 2014). The second approach used time series data for cropland classification by using different algorithms such as the threshold-based segmentation (Lobell & Asner, 2004; Pan et al., 2012; Thenkabail et al., 2009; Thenkaball, GangadharaRao, Biggs, Krishna, & Turral, 2007; Wardlow & Egbert, 2008; Wardlow, Egbert, & Kastens, 2007). For example, Thenkaball et al. (2007) and Thenkabail et al. (2009) presented a spectral matching techniques (SMT) to map land cover and irrigated areas.

The third approach identifies paddy rice based on the flooding signatures (a mix of water and rice plants) during the rice transplanting phase to early vegetative growth phase, which can be detected with the time series land surface water index (*LSWI*), normalized difference vegetation index (*NDVI*), and enhanced vegetation index (*EVI*) in a pixel (Xiao et al., 2002, 2005, 2006). This phenology- and pixel-based paddy rice mapping (PPPM) algorithm has been proved to have a good performance in spatially extensive application at large scale tropical regions using time series MODIS data. For example, it has been used for mapping paddy rice in South and Southeast Asia (Bridhikitti & Overcamp, 2012; Xiao et al., 2006), and China (Sun, Huang, Huete, Peng, & Zhang, 2009; Xiao et al., 2005). Recently researchers have conducted Landsat-based paddy rice change detection in a case region of the cold temperate NE China (Dong et al., 2015), in which the temperature-based thermal growing season was used to define the time window of the rice transplanting phase. Moreover, the study indicated that the Landsat data availability, especially in the flooding and transplanting phase, is critical for guaranteeing a high quality map of paddy rice planting area (Dong et al., 2015).

The reliance of phenology-based algorithms on high temporal frequency images determines the selection of the low spatial and high temporal resolution data (e.g., MODIS) in previous studies. Despite a 16-day revisit frequency, the complete observations from Landsat 5/7 are usually inaccessible due to missing data and SLC-off gaps. Due to the Landsat 5/7 data availability, paddy rice area was identified at a five-year epoch scale instead of an annual scale by assuming no changes in one epoch (Dong et al., 2015). The recently launched Landsat 8 provides a new opportunity for land cover mapping (Roy et al., 2014). In addition to improved performance in spectral and radiometric resolutions, an improved duty cycle allows collection of a significantly greater number of images per day (Roy et al., 2014), which provides more data sources for phenology-based land cover mapping. Can the enhanced temporal and geographic coverage of Landsat 8 provide an opportunity to realize the paddy rice mapping at annual scale? It would be a priority to demonstrate the feasibility of all the available Landsat 8 images on annual paddy rice mapping in the new frontier of rice expansion, NE Asia. This study would verify whether a single year of Landsat 8 data with improved temporal coverage can facilitate paddy rice map on the large scale area like NE Asia, which was proven difficult using Landsat 5/7 imagery in previous studies (Dong et al., 2015).

Since the free release of Landsat data by USGS/EROS (Wulder, Masek, Cohen, Loveland, & Woodcock, 2012), an unprecedented opportunity became available for us to map large scale paddy rice at 30-m resolution and longer temporal range (back to 1980s). Utilizing the open

and dense Landsat archive data for the land cover studies has been attracting lots of attention (DeVries, Verbesselt, Kooistra, & Herold, 2015; Dong et al., 2015; Schroeder et al., 2014; Schroeder, Wulder, Healey, & Moisen, 2011; Zhu, Woodcock, Holden, & Yang, 2015). For example, time series Landsat data have been used for separating cropland and pasture (Müller, Rufin, Griffiths, Barros Siqueira, & Hostert, 2015), forest disturbance (Masek et al., 2013; Zhu, Woodcock, & Olofsson, 2012), impervious surface dynamics (Zhang, Pan, Chen, Zhan, & Mao, 2013), glacier extent changes (Yavaşlı, Tucker, & Melocik, 2015), and paddy rice dynamics (Dong et al., 2015; Kontgis, Schneider, & Ozdogan, 2015). Although these efforts have greatly improved the previous land cover and land use change efforts from spatial and temporal domains, these studies all focus on local or limited study areas due to the intensive amount of data processing. Computing capability has been improved greatly in last few years. For example, Google Earth Engine (GEE), based on its millions of servers around the world and the start-of-the-art cloud-computing and storage capability, has archived a large catalog of earth observation data and enabled the scientific community to work on the trillions of images in an intrinsically-parallel processing way. For example, Hansen et al. (2013) have generated annual global forest change maps from 2000 to 2012, using over 650,000 Landsat 7 scenes and one million CPU hours in few days. Besides Google, the NASA Earth Exchange (NEX) provides a state-of-the-art supercomputing platform (Pleiades supercomputer) to realize the remote sensing data processing and large data analysis (Nemani, Votava, Michaelis, Melton, & Milesi, 2011). Amazon Web Service (AWS) also hosts petabytes of Landsat data now and can provide on-demand cloud-computing services for planetary scale analysis. Our question is whether we can use the emerging cloud-computing approach to extend our algorithm to subcontinental scale for mapping paddy rice planting area at 30-m resolution.

Based on the necessity of paddy rice information in high latitude regions and the existing challenges in large scale moderate resolution paddy rice mapping, the objective of this study is to generate an unprecedented 30-m paddy rice map in NE Asia, including NE China, South Korea, North Korea, and Japan, by using all the time series Landsat 8 imagery in 2014, the phenology- and pixel-based paddy rice mapping (PPPM) algorithm, and the cloud-computing technology with the GEE platform. Specifically, we would like to answer three questions: (1) Is the PPPM algorithm based on thermal phenology definitions reliable for large scale paddy rice mapping in NE Asia? (2) How well can the time series Landsat 8 archive data support paddy rice mapping at annual scale? and (3) whether can the cloud-computing approach effectively support thousands of Landsat image processing for paddy rice mapping? The resultant paddy rice map in 2014 would provide a detailed baseline map for the food security in these four countries and regions. This study also sheds light on large scale mapping of other land use types or global paddy rice mapping at 30-m spatial resolution.

## 2. Materials and methods

### 2.1. Study area

NE Asia is composed of NE China, North Korea, South Korea, and Japan (Fig. 1). NE China and the Korean peninsula have two major mountain ranges, one is the Lesser Khingan

Mountains, located on the northwestern part of Heilongjiang province; the other one is the Changbai Mountains which range from northeast to southwest and cover the eastern part of NE China and the eastern part of Korean peninsula. The central Japan is also mountainous with an average elevation over 600 m (Fig. 1). Besides these mountain ranges, all the other areas generally have an elevation lower than 400 m. The plains in NE China are mainly located in the northeastern (Sanjiang Plain), western (Songnen Plain), and southern parts (Liaohe Plain), while plains in the Korean peninsula are located in the western part, and the plains in Japan are distributed in the surrounding alluvial plains in the islands (Fig. 1). Abundant rivers are distributed in these alluvial plains with an average elevation less than 200 m and these plains are the major paddy rice production regions in the study area.

The climates in these four countries/regions are different with a transition from cold temperate monsoon climate in the northern parts of NE China to subtropical monsoon climate in southern Japan (Fig. 1). NE China has cold temperate and humid/sub-humid climates. The annual accumulated air temperatures above 0 and 10 °C range from 2000 – 4200 °C·day and 1600–3600 °C·day respectively. Annual precipitation varies approximately 500–800 mm which is concentrated in July and August. The number of frost-free days varies between 140 days and 170 days. Due to the temperature limit, generally single cropping dominates this region. Japan has different climates between the south and the north due to the large distance from the south to the north. The mean temperature in January (July) is −6 (17) °C in the north region and 16 (28) °C in the south region. The average annual precipitation ranges from 700 mm to 3500 mm. The Korean Peninsula belongs to the temperate monsoon climate with continental climate characteristics in the north like NE China and marine climate characteristics in the south like Japan.

### 2.2. Landsat data and pre-processing

All the available standard Level 1 Terrain-corrected (L1T) orthorectified images from Landsat 8 in 2014 were used for this study, which have been archived in the GEE platform as the image collection of United States Geological Survey (USGS) Landsat 8 Top of Atmosphere (TOA) Reflectance (Orthorectified). There are 143 path/rows to cover the whole study area, NE Asia (Fig. 1). We counted the observation numbers of individual pixels from Landsat 8 within a year and within the rice transplanting phase in 2014. On average, Landsat 8 had an average of 36.4 observations in the study area in 2014 and an average of 7.8 observations in the rice transplanting phase. Moreover, 66% of Landsat 8 pixels had over 30 observations in 2014 (Fig. 2a) and 62% of Landsat 8 pixels had over 5 observations in the rice transplanting phase (Fig. 2c). The higher intensity observations in the northern regions are due to the sidelaps of the high latitudes, which contribute to a higher temporal resolution of observations and provide more phenology information for paddy rice mapping.

The bad observations, including clouds, cirrus, and snow/ice, were removed according to the Landsat 8 quality assessment band (Roy et al., 2014; Scaramuzza, Bouchard, & Dwyer, 2012). For each category of bad observations, there are four levels of conditions indicating their confidences, including “not determined” (Algorithm did not determine the status of this condition), ‘no’ (0–33% confidence), ‘maybe’ (34–66% confidence), and ‘yes’ (67–100% confidence). We used the 67–100% confidence level to exclude all the potential bad

observation effects from clouds, cirrus, and snow/rice. The time series of Landsat TOA image collection was used to calculate three vegetation indices, including *NDVI* (Tucker, 1979), *EVI* (Huete et al., 2002; Huete, Liu, Batchily, & vanLeeuwen, 1997), and *LSWI* (Xiao et al., 2005). The spectral indices were calculated using following equations:

$$\mathrm{NDVI}=\frac{\rho_{\mathrm{NIR}}-\rho_{\mathrm{Red}}}{\rho_{\mathrm{NIR}}+\rho_{\mathrm{Red}}} \quad (1)$$

$$\mathrm{EVI}=2.5\times\frac{\rho_{\mathrm{NIR}}-\rho_{\mathrm{Red}}}{\rho_{\mathrm{NIR}}+6\times\rho_{\mathrm{Red}}-7.5\times\rho_{\mathrm{Blue}}+1} \quad (2)$$

$$\mathrm{LSWI}=\frac{\rho_{\mathrm{NIR}}-\rho_{\mathrm{SWIR}}}{\rho_{\mathrm{NIR}}+\rho_{\mathrm{SWIR}}} \quad (3)$$

where $\rho_{\mathrm{blue}}$, $\rho_{\mathrm{red}}$, $\rho_{\mathrm{NIR}}$, and $\rho_{\mathrm{SWIR}}$ are the surface reflectance values of the blue band (0.45–0.52 mm), red band (0.63–0.69 mm), near-infrared band (for short hereafter, 0.76–0.90 mm), and shortwave-infrared band (1.55–1.75 mm) in the Landsat OLI sensor.

All the Landsat data processing was conducted using the cloud-computing technology in the GEE platform (https://earthengine.google.org/), which enables parallel computing and big data processing feasible in the large study area. After excluding these bad observations, the resulting time series data were used for the following phenology-based classification.

### 2.3. Algorithms to identify inundation and paddy rice fields

A phenology- and pixel-based paddy rice mapping (PPPM) algorithm was used for paddy rice mapping (Fig. 3). The flooding and rice transplanting signals are key features to identify paddy rice fields as paddy rice is the only crop needing to be transplanted in a water–soil mixture environment (LeToan et al., 1997). The remote sensing recognition of the transplanting signals is critical to identify paddy rice. Previous studies have revealed that the relationship between *LSWI* and *NDVI* (*EVI*) can effectively discriminate flooding/transplanting signals (Boschetti, Nutini, Manfron, Brivio, & Nelson, 2014; Xiao et al., 2002). We conducted temporal profile analysis at random paddy rice sites from the four countries/regions and found that all the paddy rice agriculture systems in the study area have flooding and transplanting signals in the early growing season and can be captured using vegetation indices (Fig. 4); specifically, here we recognized flooding signals using the criteria $LSWI > EVI$ or $LSWI > NDVI$.

The transplanting phases are different among regions (Fig. 4). Recognition of paddy rice transplanting phase is critical for the identification of rice as it can help to avoid introducing occasional noises from other seasons, e.g., rainfall flooding in summer and snowmelt flooding in early spring. As temperature is a major limiting factor for vegetation growth and cropping in the cold temperate zone, previous studies have applied air temperature or land

surface temperature to simulate the start of the thermal growing season in temperate zones (Dong et al., 2015; Dong, Liu, Tao, Xu, & Wang, 2009; Dong et al., 2013b; Zhang et al., 2015). For example, time series MODIS LST data was successfully used to define the starting date of rice transplanting (SOT) using nighttime LST over 5 °C (Zhang et al., 2015). The use of nighttime LST is based on a truth that the daily temperature variation is more remarkable in the high latitudes and daily minimum temperature is a more reliable indicator for biophysical limitation (Fig. S1). Moreover, the flooding/transplanting signals can last for approximately two months after transplanting until rice canopy covers most of the surface area in tropical areas (Xiao et al., 2005, 2006). By contrasting the temporal profiles of nighttime LST and spectral indices (*NDVI*, *EVI* and *LSWI*), the start of thermal growing season above 5 °C was close to the SOT in NE China, North Korea, northern Japan, and northern South Korea (Fig. 5), but had time lags with the SOT in the southern Japan and southern South Korea, which belong to the oceanic climate types (Fig. 1). Considering the variability of transplanting timing in different fields (Fig. 4), we defined the length of transplanting phase to open canopy as 80 days in this study.

Therefore, we detected the flooding and transplanting signals using the vegetation indices within the rice transplanting phase. Specifically, the following equation is used,

$$\text{Flood}=\begin{cases} 1 \ (\text{LSWI}_{Ti}>\text{EVI or LSWI}_{Ti}>\text{NDVI}) \\ 0 \ (\text{LSWI}_{Ti} \le \text{EVI and LSWI}_{Ti} \le \text{NDVI}) \end{cases} \quad (\text{SOT}<T_i<\text{EOT}) \tag{4}$$

where *Flood* is the status of flooding/transplanting, the starting (*SOT*) and ending (*EOT*) of transplanting phase are defined using the method stated above, and $T_i$ is the period when the observations are acquired.

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

**2.4.1. Non-cropland masks—**Although the flooding signal in the paddy rice transplanting period is a unique characteristic for paddy rice, flooding signals from other land covers could also exist in the paddy rice transplanting phase, mainly including water bodies (e.g., rivers and lakes) and natural wetlands. Also, some noises could be caused by miscellaneous non-cropland land covers such as natural vegetation and built-up lands due to some unexpected weather effects. The generation of these non-cropland masks can help to reduce sporadic commission errors in the resultant paddy rice maps.

We generated several masks by refining the rules used in previous studies (Dong et al., 2015; Xiao et al., 2005, 2006; Zhang et al., 2015), including (1) Sparse vegetation mask, including saline and alkaline land, built-up land, water body, and other low-vegetated lands, which are covered by sparse vegetation or no vegetation covered. It uses the maximum *EVI* value during the thermal growing season above 5 °C smaller than 0.6 (see Fig. S2 for the justification); (2) Natural vegetation mask, including forests, natural wetlands, grass, etc. which grow earlier and are greener than paddy rice by the start of the thermal growing season above 10 °C (S_GS10). That uses the maximum *EVI* value before the S_GS10 larger than 0.4 (see Fig. S3 for the justification). Note that this rule only was used on the cold

regions with single cropping which were defined using climate zones (Fig. 1). Specifically, Cwa (temperate, dry winter and hot summer) and Cfa (temperate, without dry season and hot summer) were excluded as a second growth cycle exists there (Fig. 4); (3) Forest mask, from the fine resolution (25-m) forest map based on the Japan Aerospace Exploration Agency (JAXA)/Phased Array type L-band Synthetic Aperture Radar (PALSAR) (Dong et al., 2012; Qin et al., 2016); and (4) Sloping land mask, where slopes larger than 3° with the Shuttle Radar Topography Mission (SRTM) Digital Elevation Data Version 4 (Farr et al., 2007; Jarvis, Reuter, Nelson, & Guevara, 2008), as paddy rice cannot be planted in the slopping lands due to the flooding demand in the transplanting phase. These masks were all excluded before the phenology information was used to map paddy rice, to minimize commission error.

**2.4.2. Generation of paddy rice map—**After masking the above non-cropland masks, paddy rice can be identified according to its unique flooding characteristics in the transplanting phase. However, due to the variation of transplanting timing in different fields, the transplanting information from one specific image or one image composite could be biased. Here we used a frequency-based approach to clarify the flooding signals by accounting for all the flooding signals in the transplanting phase.

$$F=\frac{\sum N_{\text{flood}}}{\sum N_{\text{total}}-\sum N_{\text{bad}}} \quad (5)$$

where $F$ is the frequency of flooding signals among all of the good-quality observations in the transplanting phase, $N_{flood}$ is the number of flooding/transplanting signals, $N_{total}$ is the total observation number in the transplanting phase, and $N_{bad}$ is the number of bad observations (e.g., clouds, cloud shadows, snow, and cirrus). After a statistical analysis based on ground truth data, we used 10% as a threshold to generate the paddy rice layer (see Fig. S4 for the justification).

## 2.5. Validation and comparison

Accuracy assessment of our resultant paddy rice maps includes two aspects: 1) validation using the very high resolution (VHR) images and field photos, and 2) comparison with MODIS and Landsat 7 based paddy rice maps in a few case regions.

As our study area is huge, in order to realize a reasonable coverage of the validation samples, we collected the reference data from VHR images from the Google Earth (Fig. 6a) and the National Geospatial-Intelligence Agency (NGA), together with large amount (>20,000) of field photos (Fig. 6b) from the Global Geo-Referenced Field Photo Library (a Google Maps-based scientific field photos management platform, www.eomf.ou.edu/photos/) which are freely available to the public. Most of the photos in NE China were collected in ground surveys in 2013 summer while only around 300 photos were available in Japan, South Korea, and North Korea from the public contribution on the crowd-sourcing platform.

We used the stratified random sampling approach to acquire validation samples. First, according to the MODIS land cover products (Friedl et al., 2010), paddy rice map (Zhang et al., 2015) and the field photos, the study area was partitioned into several strata (paddy rice, upland crops, forest and other natural vegetation, water body, and built-up land). Second, random points were generated in each stratum and then we made areas of interest (AOIs) as square buffers of the points (100-m by 100-m) (see Fig. 6c for an AOI example). Third, we checked each AOI with the VHR images and selected and labeled the pure land cover AOIs by referring to both VHR images and the field photos (if any); the AOIs without clear land cover information were excluded. For the AOIs with field photos available, the photos were used to clarify land cover types; for the AOIs without field photos, the field shapes, sizes, and flooding features were considered in labelling land cover information (See Dong et al., 2015 for a detailed AOI generation description). Finally a total of 7400 AOIs (81,984 pixels, Fig. 6d) were generated for the validation of resultant paddy rice and non-paddy rice map (Table 1), including paddy rice (1710 AOIs/18,887 pixels), upland crops (590/6592), forest (4881/54,150), water body (61/680), and built-up land (154/1675). A confusion matrix of paddy rice map was calculated to evaluate the accuracy of the results. Individual confusion matrix for different countries was also examined. Following the previous studies (Card, 1982; Olofsson, Foody, Stehman, & Woodcock, 2013), we adjusted accuracies of the resultant rice/non-rice map by considering the area of each stratum, also, the rice areas were corrected based on the area-adjusted accuracies. Both accuracies and areas were presented with a 95% confidence interval.

Besides the validation, we also conducted comparisons between the Landsat 8 based result and the Landsat 7 based rice map as well as the existing paddy rice maps from the MODIS (Zhang et al., 2015) in a randomly selected case region. The aim of the comparison is to check whether the multi-temporal Landsat 8 images can effectively capture paddy rice phenology information comparing to higher temporal resolution of MODIS data and the previous generation of sensor (Landsat 7).

## 3. Results

### 3.1. Pattern of paddy rice transplanting phase in NE Asia

Paddy rice transplanting phase varied in different geographical regions due to different thermal conditions. Fig. 5 shows the spatial distribution of the starting dates of rice transplanting phase (SOT) and its distribution along different latitude and elevation gradients. The SOT delayed along with the increasing of latitudes from around 71 (±29) DOY to 165 (±5) DOY, with a long time range (~3 months) from early March to early June (Fig. 5a). The gradients were also indicated in the Fig. 4. In contrast to the significant variation in latitude-based SOT gradients (Fig. 5b), the trend in the SOT change along with the elevation gradients was not remarkable (Fig. 5c). The weak variation in elevation could be attributed to the offset effects of the wide latitude gradients.

### 3.2. Accuracy assessment and comparison with existing paddy rice maps

The validation based on the VHR images and field photos indicated that the produced paddy rice map in 2014 (Fig. 7) had a high accuracy, with the overall accuracy of 98% (Table 1).

The paddy rice category had a producer accuracy (PA) and user accuracy (UA) in the order of 73% and 92%, which were adjusted by the area of stratum. The accuracies of the paddy rice maps were not identical in different countries/regions, NE China possessed the highest accuracy of paddy rice map (PA 83% and UA 93%), followed by North Korea (PA 61% and UA 96%); while Japan (PA 52% and UA 91%) and South Korea (PA 50% and UA 89%) had the slightly lower accuracies (Table 1). The highest paddy rice accuracy in NE China could be related to the large rice parcels in the three large plains while paddy rice fields were mainly distributed in the mountainous valleys in other countries, so the omission errors were relatively high in Japan, South Korea, and North Korea. Another reason could be related to the estimation weights associated with each stratum and the low rice area proportion. For more uncertainty analysis please see the discussion Section 4.2.

The comparison between Landsat 8 based paddy rice map and the existing MODIS-based map (Zhang et al., 2015) in 2014 showed that both products had a high spatial consistency in a case region of the Sanjiang Plain (Fig. 8), while Landsat 8-based result had more details, in particular, it can exclude the roads between paddy rice fields. The Landsat 7 based rice map also had more details at the field scale; however, it suffered from the missing of observations in the southwestern part and some strips across the region (Fig. 8c), which would lead to noteworthy omission errors. The spatial consistency between the paddy rice maps from the hyper-temporal resolution of MODIS and 16-day revisited Landsat 8 indicates the temporal coverage of Landsat 8 data can reasonably support yearly paddy rice mapping in NE Asia. Moreover, the Landsat 8 has a higher capability in mapping paddy rice using phenology information compared to Landsat 7.

### 3.3. Geographical characteristics of paddy rice distribution in NE Asia

Despite the main staple for the study area, paddy rice was a minority of the land use types (10.4 ± 0.08 million ha), which accounted for only 7.6% of the land in 2014. Fig. 7 shows that the paddy rice fields were mainly distributed in the low alluvial plains, which was consistent with the terrain shown in Fig. 1. Here we have analyzed the geographical distribution characteristics from aspects of countries, latitudes, elevations, and climate zones.

In terms of distribution among countries, NE China had an area of paddy rice of 6.2 ± 0.05 million ha, which accounted for 60% paddy rice in the NE Asia; while Japan had a paddy rice area of 2.1 ± 0.05 million ha (20%); South Korea and North Korea had relatively less paddy rice areas and that accounted for 14% and 7% of the whole study area (Fig. 9a), which is consistent with the statistics from FAO (Fig. S5).

In terms of distribution within different climate zones, a half of paddy rice area was located in the Dwa (cold, dry winter and hot summer) zone, followed by Dwb (cold, dry winter and warm summer, 29%), Cfa (temperate, without dry season and hot summer, 9%) and Dfa (cold, without dry season and hot summer, 7%). Therefore, 90% paddy rice area was located in the cold climate zones (Fig. 9b).

In terms of distribution along the elevation gradients, 4.4 million ha (57%) of paddy rice area were located in the area with an average elevation less than 100 m asl; and 90% was

distributed on the area below 200 m asl (Fig. 9c). Paddy rice area decreased rapidly with the elevation over 200 m. That is because the paddy rice needs good irrigation conditions and water resources which are more suitable in plains. In terms of distribution along the latitude gradients, 62% paddy rice located in the range from 44 to 47°N (Fig. 9d), was generally contributed by Sanjiang Plain, China. The variation along the latitude gradients was generally controlled by the main paddy rice production plains. Note that these area statistics in different climate zones, elevation/latitude gradients were not adjusted using the area-adjusted accuracies as no specific confusion matrix was available for each climate zone and elevation/latitude gradient; however, it did show the general distribution of paddy rice in different geographical regions.

## 4. Discussion

### 4.1. Reliability of paddy rice mapping using OLI data, GEE and improved PPPM method

This study indicated the feasibility and reliability of mapping yearly 30-m paddy rice in cold NE Asia using Landsat 8 images and the improved PPPM method based on the pre-definition of transplanting time window. The GEE was used for the cloud computing given the huge data size and processing requirement. To our knowledge, this study was the first investigation to map paddy rice at 30-m spatial resolution on the subcontinental scale. The implementation of this study was attributed to several factors including improved data, platform, algorithm, and simple natural conditions.

First, the improved data was from the enhanced observation capability of Landsat 8, which provides high temporal frequency of observations (Roy et al., 2014) and can effectively meet the time series data requirement of phenology-based approaches. In addition, the Landsat sidelaps in the high latitude regions increased the observations by 59.6%, from a theoretically spatially-averaged 22.8 observations to 36.4 observations in the entire year of 2014 (Fig. 2a). Meanwhile, Landsat 7 had an average of 22.1 observations in the study area in 2014. Landsat 8 have much higher advantages than Landsat 7 in the data availability in both the yearly scale and the transplanting phase (Fig. 2). Landsat 5 data was unavailable in 2014 due to the sensor decommission in June 2013. Last but not least, we used all the good observations of individual pixels to increase the representation of phenology information (Dong et al., 2015), instead of the selection of the less cloud/shadow scenes.

Second, the GEE-based cloud computing effectively facilitated processing thousands of Landsat images for paddy rice mapping. Specifically, GEE synchronizes all the Landsat data from the Earth Resources Observation and Science (EROS) center of USGS, and also it provides different levels of processed products, including the raw data, TOA reflectance, and surface reflectance (SR) data at scene and composite formats (8-day, 16-day, 32-day, and annual). Its parallel processing platform is powerful by utilizing millions of servers over the world, which sped up the whole process time of this study to a considerable degree (<1 day). This study shed light on the encouraging capability of cloud computing technology in the emerging platforms (e.g., GEE and AWS) for land cover mapping.

Third, the PPPM approach (Xiao et al., 2005, 2006) based on the transplanting time window was promising in NE Asia. While temperature is the major limitation for agriculture in this

region, the land surface temperature based paddy rice planting calendar (i.e., transplanting) can be applied in the study area. The transplanting time window can help to remove the potential effects of spring flooding from freezing snow/ice on land or rivers. This study furthermore proved this approach reliable for the cold regions in Asia, which is consistent with previous pilot studies in single path/row (Dong et al., 2015; Qin et al., 2015).

Fourth, simpler natural conditions in NE Asia relative to tropical regions were another factor contributing to the successful implementation of paddy rice mapping. The longer growing season, lower cropping intensity (generally single cropping), and less cloud coverage in the study area provided opportunities to acquire more effective observations of paddy rice in the growing season, in contrast to the tropical regions (shorter growing season, multiple cropping, and more cloud coverage).

### 4.2. Uncertainty analysis and implications for extensive applications

This study indicates the advances of the improved Landsat 8 imaging systems in paddy rice mapping and characterization, in particular, the enhanced temporal and geographic coverage plays an important role. We used TOA reflectance data in this study as the SR data for Landsat 8 was not available in the GEE platform until after the study was complete and the parallel approach of the atmospheric correction (e.g., LEDAPS) was not accessible either. The use of TOA reflectance data could potentially impact the classification results; for example, the TOA reflectance based $EVI$ values were generally higher than SR-based one; however, the reasonably high accuracy of the result showed the effects of TOA-reflectance could be limited, as we used a set of robust strategies such as utilizing all the available images, flooding frequency based segmentation approach, and phenology-based approach. A comparison will be helpful in the future between TOA and SR reflectance to determine whether the SR based paddy rice mapping results are more reliable. In this study, the bad observations including clouds, snow/ice, and cirrus were masked according to the quality assessment band. Our previous study showed the good performance of Fmask-based clouds/cloud shadows (Dong et al., 2015) and that could be involved in the processing chain to improve the results when the Fmask package is available on the GEE.

Phenology-based approach and all the available Landsat imagery have been increasingly used for land cover and land use change studies. For example, Zhong et al. (2014) applied phenological metrics (e.g., starting date and ending date) to map soybean and corn; Dong et al. (2013a) identified deciduous rubber plantations by integrating the phenology information from time series Landsat images and the tree structure information from synthetic aperture radar data. However, the phenology-based land cover mapping relies on the high temporal frequency of remote sensing images in order to acquire the phenology characteristics. Fig. 10 shows the spatial distribution of the Landsat 8 good observations in the entire year of 2014 and in the transplanting phase defined in this study. We can find the spatially average good observation frequency of Landsat 8 (Landsat 7) were 20.3 (13.3) for the entire year and 4.1 (3.0) for the transplanting phase, which was generally sufficient to pick up paddy rice characteristics. However, Landsat 8 (Landsat 7) still had 12% (18%) pixels with only one to two good observation(s) in the transplanting phase which still had risks to lose the flooding

signals in paddy rice fields. This can explain that the paddy rice maps in Japan, North Korea, and South Korea had higher omission errors than commission errors (Table 1).

In addition, the method could not be applied in pre-Landsat 8 periods due to lower observation frequency. Our previous study has proved Landsat 7 images insufficient to map paddy rice on annual scale due to limited observation capability and SLC-off gaps (Dong et al., 2015). But the methodology can be used to map paddy rice in the multiple-year epoch level. Also, most recent studies have improved the capability of generating synthetic time series gap-free Landsat surface reflectance and predict reflectance in given time (Hermosilla, Wulder, White, Coops, & Hobart, 2015; Zhu et al., 2015), which shed light on the phenology-based paddy rice mapping as it could improve the availability of good observations in the transplanting phase.

Phenology-based approach has been proved effective in the paddy rice mapping using time series MODIS data (Xiao et al., 2005, 2006). This study clarified the feasibility of the PPPM and Landsat 8 data on annual paddy rice mapping in the cold Asian regions. However, we can expect more challenges when it is extended to tropical regions (e.g., Southeast and South Asia) with more clouds (Asner, 2001), more fragmented croplands, and higher cropping intensity (Gibbs et al., 2010). Increasing availability of Landsat-like moderate resolution multi-spectral sources, in particular, the new ESA/EU Sentinel-2, would provide seamless combination and lead to higher temporal resolution observation capability (Wulder, White, Masek, Dwyer, & Roy, 2011). The combination of Landsat 8 and the Sentinel-2 will improve the data revisit cycle from 16 days to <5 days (Wulder et al., 2011). So an extensive application of the methodology in tropical area is foreseen and encouraging.

## 5. Conclusions

Existing efforts on large scale paddy rice mapping are generally focused on the tropical or subtropical Asia using high temporal MODIS datasets, and our knowledge is still limited in the new frontier of paddy rice expansion–NE Asia. Based on the improved phenology-and pixel-based paddy rice (PPPM) algorithm and all the available Landsat 8 images in a single year, this study generated an unprecedented 30-m paddy rice map in NE Asia in 2014. The products will be freely available to the public through the EOMF data portal (http://maps.eomf.ou.edu) for users' assessments and applications. To our knowledge, this is the first application of Landsat imagery on subcontinental scale paddy rice mapping. Its effective implementation in large area with thousands of Landsat scenes was greatly attributed to the utility of cloud-computing technology, which showcased potentials and prospects of the emerging parallel/cloud-computing platforms (e.g., GEE, NASA NEX, and AWS) in global and longer temporal scale of land cover and land use change studies.

Due to less cloud coverage, longer growing season, and more sidelaps observations relative to the tropical regions, the improved PPPM algorithm has been proved robust in mapping paddy rice in cold regions. More importantly, Landsat 8 imagery provides an effective support, although our paddy rice mapping algorithm needs an intensive observations in the rice transplanting phase. This study demonstrates that Landsat 8 data can reasonably meet the high temporal observation requirement of the phenology-based paddy rice mapping in

temperate regions. That suggests Landsat 8 imagery is promising in the applications of more detailed land use mapping such as different crop types (e.g., rice), forest species (e.g., industrial plantation). The increasing availability of moderate resolution imagery from the additional Landsat-like sensors (e.g., Sentinel-2) (Wulder et al., 2011) would be expected to lead to more precise global land cover and land use maps with finer category information which will provide more applicable information for ecological, hydrological and other applications.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgments

This study was supported by the NASA Land Use and Land Cover Change program (NNX11AJ35G, NNX14AD78G), the National Institutes of Health (1R01AI101028-01-A2) and the Google Earth Engine Research Award to Prof. Xiangming Xiao at the University of Oklahoma. The high resolution images were provided by NASA through National Geospatial-Intelligence Agency (NGA) Commercial Data Archive (cad4nasa.gsfc.nasa.gov). We thank Dr. Noel Gorelick for his technical support in the application of the GEE platform, and Sarah Xiao at Yale University for the English editing, and Drs. Xiaopeng Song and Zhe Zhu for their suggestions on the manuscript. We thank the Guest Editors of the Landsat 8 Special Issue, Drs. Tom Loveland and Jim Irons, as well as three anonymous reviewers for their valuable comments.

## Appendix A. Supplementary data

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.rse.2016.02.016.

## References

Arino O, Bicheron P, Achard F, Latham J, Witt R, Weber JL. GLOBCOVER — The most detailed portrait of Earth. Esa Bulletin-European Space Agency. 2008:24–31.

Asner GP. Cloud cover in Landsat observations of the Brazilian Amazon. International Journal of Remote Sensing. 2001; 22:3855–3862.

Boschetti M, Nutini F, Manfron G, Brivio PA, Nelson A. Comparative analysis of normalised difference spectral indices derived from MODIS for detecting surface water in flooded rice cropping systems. PloS One. 2014; 9

Bridhikitti A, Overcamp TJ. Estimation of Southeast Asian rice paddy areas with different ecosystems from moderate-resolution satellite imagery. Agriculture Ecosystems and Environment. 2012; 146:113–120.

Card DH. Using known map category marginal frequencies to improve estimates of thematic map accuracy. Photogrammetric Engineering and Remote Sensing. 1982; 48:431–439.

DeVries B, Verbesselt J, Kooistra L, Herold M. Robust monitoring of small-scale forest disturbances in a tropical montane forest using Landsat time series. Remote Sensing of Environment. 2015; 161:107–121.

Dong JW, Liu JY, Tao FL, Xu XL, Wang JB. Spatio-temporal changes in annual accumulated temperature in China and the effects on cropping systems, 1980s to 2000. Climate Research. 2009; 40:37–48.

Dong JW, Xiao XM, Sheldon S, Biradar C, Duong ND, Hazarika M. A comparison of forest cover maps in mainland Southeast Asia from multiple sources: PALSAR, MERIS, MODIS and FRA. Remote Sensing of Environment. 2012; 127:60–73.

Dong J, Xiao X, Chen B, Torbick N, Jin C, Zhang G, Biradar C. Mapping deciduous rubber plantations through integration of PALSAR and multi-temporal Landsat imagery. Remote Sensing of Environment. 2013a; 134:392–402.

Dong JW, Liu JY, Zhang GL, Basara JB, Greene S, Xiao XM. Climate change affecting temperature and aridity zones: A case study in Eastern Inner Mongolia, China from 1960 to 2008. Theoretical and Applied Climatology. 2013b; 113:561–572.

Dong J, Xiao X, Kou W, Qin Y, Zhang G, Li L, Jin C, Zhou Y, Wang J, Biradar C, et al. Tracking the dynamics of paddy rice planting area in 1986–2010 through time series Landsat images and phenology-based algorithms. Remote Sensing of Environment. 2015; 160:99–113.

Ehhalt, D.; Prather, M.; Dentener, F.; Derwent, R.; Dlugokencky, EJ.; Holland, E.; Matson, P. Atmospheric chemistry and greenhouse gases. Richland, WA (US): Pacific Northwest National Laboratory (PNNL); 2001.

Elert E. Rice by the numbers: A good grain. Nature. 2014; 514:S50–S51. [PubMed: 25368886]

FAOSTAT. Statistical database of the food and agricultural organization of the United Nations. 2010 (In).

Farr TG, Rosen PA, Caro E, Crippen R, Duren R, Hensley S, Kobrick M, Paller M, Rodriguez E, Roth L, et al. The shuttle radar topography mission. Reviews of Geophysics. 2007; 45

Friedl MA, McIver DK, Hodges JCF, Zhang XY, Muchoney D, Strahler AH, Woodcock CE, Gopal S, Schneider A, Cooper A, et al. Global land cover mapping from MODIS: Algorithms and early results. Remote Sensing of Environment. 2002; 83:287–302.

Friedl MA, Sulla-Menashe D, Tan B, Schneider A, Ramankutty N, Sibley A, Huang XM. MODIS collection 5 global land cover: Algorithm refinements and characterization of new datasets. Remote Sensing of Environment. 2010; 114:168–182.

Gibbs HK, Ruesch AS, Achard F, Clayton MK, Holmgren P, Ramankutty N, Foley JA. Tropical forests were the primary sources of new agricultural land in the 1980s and 1990s. Proceedings of the National Academy of Sciences of the United States of America. 2010; 107:16732–16737. [PubMed: 20807750]

Gilbert M, Xiao XM, Pfeiffer DU, Epprecht M, Boles S, Czarnecki C, Chaitaweesub P, Kalpravidh W, Minh PQ, Otte MJ, et al. Mapping H5N1 highly pathogenic avian influenza risk in Southeast Asia. Proceedings of the National Academy of Sciences of the United States of America. 2008; 105:4769–4774. [PubMed: 18362346]

Gilbert M, Golding N, Zhou H, Wint GR, Robinson TP, Tatem AJ, Lai S, Zhou S, Jiang H, Guo D, et al. Predicting the risk of avian influenza A H7N9 infection in live-poultry markets across Asia. Nature Communications. 2014; 5:4116.

Gong P, Wang J, Yu L, Zhao Y, Zhao Y, Liang L, Niu Z, Huang X, Fu H, Liu S, et al. Finer resolution observation and monitoring of global land cover: First mapping results with Landsat TM and ETM + data. International Journal of Remote Sensing. 2013; 34:2607–2654.

Hansen MC, Potapov PV, Moore R, Hancher M, Turubanova SA, Tyukavina A, Thau D, Stehman SV, Goetz SJ, Loveland TR, et al. High-resolution global maps of 21st-century forest cover change. Science. 2013; 342:850–853. [PubMed: 24233722]

Hermosilla T, Wulder MA, White JC, Coops NC, Hobart GW. An integrated Landsat time series protocol for change detection and generation of annual gap-free surface reflectance composites. Remote Sensing of Environment. 2015; 158:220–234.

Huete AR, Liu HQ, Batchily K, vanLeeuwen W. A comparison of vegetation indices over a global set of TM images for EOS-MODIS. Remote Sensing of Environment. 1997; 59:440–451.

Huete A, Didan K, Miura T, Rodriguez EP, Gao X, Ferreira LG. Overview of the radiometric and biophysical performance of the MODIS vegetation indices. Remote Sensing of Environment. 2002; 83:195–213.

Jarvis, A.; Reuter, HI.; Nelson, A.; Guevara, E. Hole-filled SRTM for the globe version 4. 2008. (Available from the CGIAR-CSI SRTM 90 m Database: http://srtm.csi.cgiar.org)

Kontgis C, Schneider A, Ozdogan M. Mapping rice paddy extent and intensification in the Vietnamese Mekong River Delta with dense time stacks of Landsat data. Remote Sensing of Environment. 2015; 169:255–269.

LeToan T, Ribbes F, Wang LF, Floury N, Ding KH, Kong JA, Kurosu T. Rice crop mapping and monitoring using ERS-1 data based on experiment and modeling results. IEEE Transactions on Geoscience and Remote Sensing. 1997; 35:41–56.

Lobell DB, Asner GP. Cropland distributions from temporal unmixing of MODIS data. Remote Sensing of Environment. 2004; 93:412–422.

Masek JG, Goward SN, Kennedy RE, Cohen WB, Moisen GG, Schleeweis K, Huang CQ. United States forest disturbance trends observed using Landsat time series. Ecosystems. 2013; 16:1087–1104.

Müller H, Rufin P, Griffiths P, Barros Siqueira AJ, Hostert P. Mining dense Landsat time series for separating cropland and pasture in a heterogeneous Brazilian savanna landscape. Remote Sensing of Environment. 2015; 156:490–499.

Nemani R, Votava P, Michaelis A, Melton F, Milesi C. Collaborative supercomputing for global change science. Eos, Transactions American Geophysical Union. 2011; 92:109–110.

Oguro, Y.; Suga, Y.; Takeuchi, S.; Ogawa, M.; Konishi, T.; Tsuchiya, K. Comparison of SAR and optical sensor data for monitoring of rice plant around Hiroshima. In: Tsuchiya, K., editor. Calibration and characterization of satellite sensors and accuracy of derived physical parameters. 2001. p. 195-200.

Olofsson P, Foody GM, Stehman SV, Woodcock CE. Making better use of accuracy data in land change studies: Estimating accuracy and area and quantifying uncertainty using stratified estimation. Remote Sensing of Environment. 2013; 129:122–131.

Pan XZ, Uchida S, Liang Y, Hirano A, Sun B. Discriminating different landuse types by using multitemporal NDXI in a rice planting area. International Journal of Remote Sensing. 2010; 31:585–596.

Pan Y, Li L, Zhang J, Liang S, Zhu X, Sulla-Menashe D. Winter wheat area estimation from MODIS-*EVI* time series data using the crop proportion phenology index. Remote Sensing of Environment. 2012; 119:232–242.

Peel MC, Finlayson BL, McMahon TA. Updated world map of the koppengeiger climate classification. Hydrology and Earth System Sciences. 2007; 11:1633–1644.

Qin Y, Xiao X, Dong J, Zhou Y, Zhu Z, Zhang G, Du G, Jin C, Kou W, Wang J, et al. Mapping paddy rice planting area in cold temperate climate region through analysis of time series Landsat 8 (OLI), Landsat 7 (ETM+) and MODIS imagery. ISPRS Journal of Photogrammetry and Remote Sensing. 2015; 105:220–233. [PubMed: 27695195]

Qin Y, Xiao X, Dong J, Zhang G, Roy PS, Joshi PK, Gilani H, Murthy MSR, Jin C, Wang J, et al. Mapping forests in monsoon Asia with ALOS PALSAR 50-m mosaic images and MODIS imagery in 2010. Scientific Reports. 2016; 6:20880. [PubMed: 26864143]

Roy DP, Wulder MA, Loveland TR, Woodcock CE, Allen RG, Anderson MC, Helder D, Irons JR, Johnson DM, Kennedy R, et al. Landsat-8: Science and product vision for terrestrial global change research. Remote Sensing of Environment. 2014; 145:154–172.

Salas W, Boles S, Li C, Yeluripati JB, Xiao X, Frolking S, Green P. Mapping and modelling of greenhouse gas emissions from rice paddies with satellite radar observations and the DNDC biogeochemical model. Aquatic Conservation-Marine and Freshwater Ecosystems. 2007; 17:319–329.

Sass RL, Cicerone RJ. Photosynthate allocations in rice plants: Food production or atmospheric methane? Proceedings of the National Academy of Sciences of the United States of America. 2002; 99:11993–11995. [PubMed: 12221297]

Scaramuzza PL, Bouchard MA, Dwyer JL. Development of the Landsat data continuity mission cloud-cover assessment algorithms. IEEE Transactions on Geoscience and Remote Sensing. 2012; 50:1140–1154.

Schroeder TA, Wulder MA, Healey SP, Moisen GG. Mapping wildfire and clearcut harvest disturbances in boreal forests with Landsat time series data. Remote Sensing of Environment. 2011; 115:1421–1433.

Schroeder TA, Healey SP, Moisen GG, Frescino TS, Cohen WB, Huang C, Yang Z. Improving estimates of forest disturbance by combining observations from Landsat time series with U.S.

Forest Service Forest inventory and analysis data. Remote Sensing of Environment. 2014; 154:61–73.

Sun HS, Huang JF, Huete AR, Peng DL, Zhang F. Mapping paddy rice with multi-date moderate-resolution imaging spectroradiometer (MODIS) data in China. Journal of Zhejiang University-Science A. 2009; 10:1509–1522.

Thenkabail PS, Dheeravath V, Biradar CM, Gangalakunta ORP, Noojipady P, Gurappa C, Li Y. Irrigated area maps and statistics of India using remote sensing and national statistics. Remote Sensing. 2009; 1:50–67.

Thenkaball PS, GangadharaRao P, Biggs TW, Krishna M, Turral H. Spectral matching techniques to determine historical land-use/land-cover (LULC) and irrigated areas using time-series 0.1-degree AVHRR pathfinder datasets. Photogrammetric Engineering and Remote Sensing. 2007; 73:1029–1040.

Thi THN, De Bie C, Ali A, Smaling EMA, Chu TH. Mapping the irrigated rice cropping patterns of the Mekong delta, Vietnam, through hyper-temporal SPOT *NDVI* image analysis. International Journal of Remote Sensing. 2012; 33:415–434.

Tucker CJ. Red and photographic infrared linear combinations for monitoring vegetation. Remote Sensing of Environment. 1979; 8:127–150.

Wardlow BD, Egbert SL. Large-area crop mapping using time-series MODIS 250 m *NDVI* data: An assessment for the US Central Great Plains. Remote Sensing of Environment. 2008; 112:1096–1116.

Wardlow BD, Egbert SL, Kastens JH. Analysis of time-series MODIS 250 m vegetation index data for crop classification in the US Central Great Plains. Remote Sensing of Environment. 2007; 108:290–310.

Wulder MA, White JC, Masek JG, Dwyer J, Roy DP. Continuity of Landsat observations: Short term considerations. Remote Sensing of Environment. 2011; 115:747–751.

Wulder MA, Masek JG, Cohen WB, Loveland TR, Woodcock CE. Opening the archive: How free data has enabled the science and monitoring promise of Landsat. Remote Sensing of Environment. 2012; 122:2–10.

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

Xiao XM, Boles S, Liu JY, Zhuang DF, Frolking S, Li CS, Moore B. Mapping paddy rice agriculture in southern China using multi-temporal MODIS images. Remote Sensing of Environment. 2005; 95:480–492.

Xiao XM, Boles S, Frolking S, Li CS, Babu JY, Salas W, Moore B. Mapping paddy rice agriculture in South and Southeast Asia using multi-temporal MODIS images. Remote Sensing of Environment. 2006; 100:95–113.

Yavaşlı DD, Tucker CJ, Melocik KA. Change in the glacier extent in Turkey during the Landsat era. Remote Sensing of Environment. 2015

Zhang Y, Wang CZ, Wu JP, Qi JG, Salas WA. Mapping paddy rice with multitemporal ALOS/PALSAR imagery in southeast China. International Journal of Remote Sensing. 2009; 30:6301–6315.

Zhang XP, Pan DL, Chen JY, Zhan YZ, Mao ZH. Using long time series of Landsat data to monitor impervious surface dynamics: A case study in the Zhoushan Islands. Journal of Applied Remote Sensing. 2013; 7:073515.

Zhang G, Xiao X, Dong J, Kou W, Jin C, Qin Y, Biradar C. Mapping paddy rice planting areas through time series analysis of MODIS land surface temperature and vegetation index data. ISPRS Journal of Photogrammetry and Remote Sensing. 2015; 106:157–171. [PubMed: 27667901]

Zhong LH, Gong P, Biging GS. Efficient corn and soybean mapping with temporal extendability: A multi-year experiment using Landsat imagery. Remote Sensing of Environment. 2014; 140:1–13.

Zhu Z, Woodcock CE, Olofsson P. Continuous monitoring of forest disturbance using all available Landsat imagery. Remote Sensing of Environment. 2012; 122:75–91.

Zhu Z, Woodcock CE, Holden C, Yang Z. Generating synthetic Landsat images based on all available Landsat data: Predicting Landsat surface reflectance at any given time. Remote Sensing of Environment. 2015; 162:67–83.

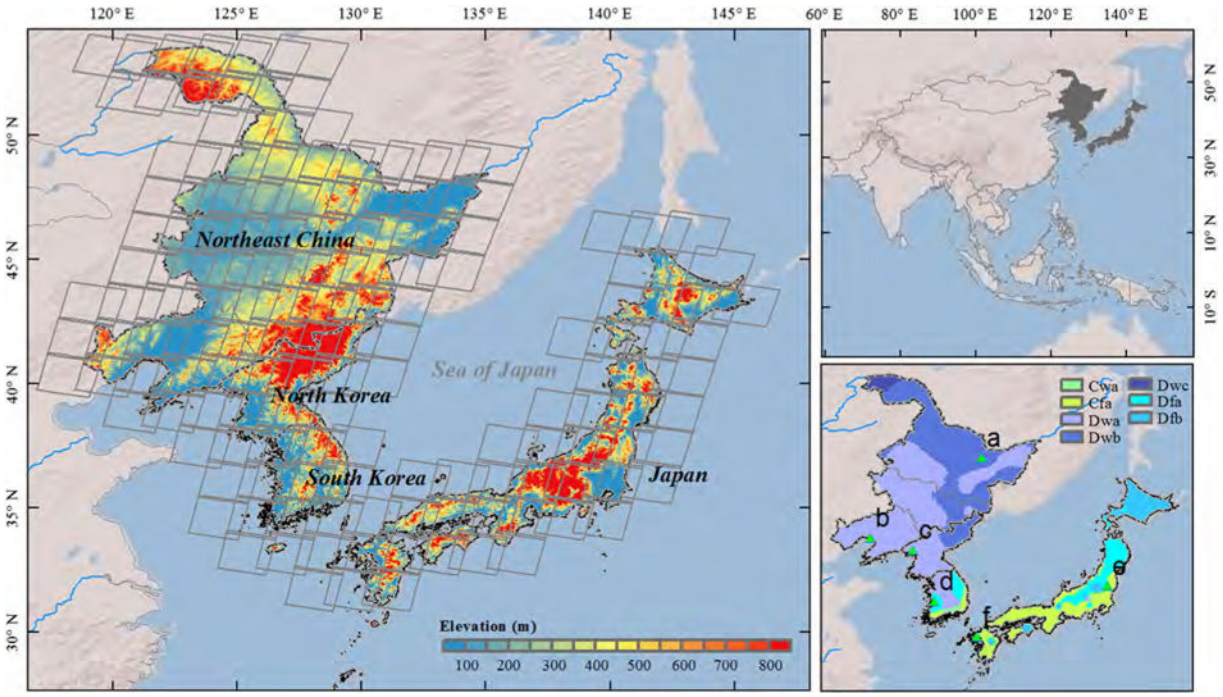


**Fig. 1.**

The topographical and climatic characteristics in the northeastern (NE) Asia, including three countries (Japan, North Korea, and South Korea) and one region (NE China) in China. The climate zone data is from Peel, Finlayson, and McMahon (2007), Cwa means temperate, dry winter and hot summer; Cfa means temperate, without dry season and hot summer; Dwa means cold, dry winter and hot summer; Dwb means cold, dry winter and warm summer; Dwc means cold, dry winter and cold summer; Dfa means cold, without dry season and hot summer; Dfb means cold, without dry season and warm summer. The boundary of Landsat path/rows is showed on the map to showcase the Landsat sidelaps.

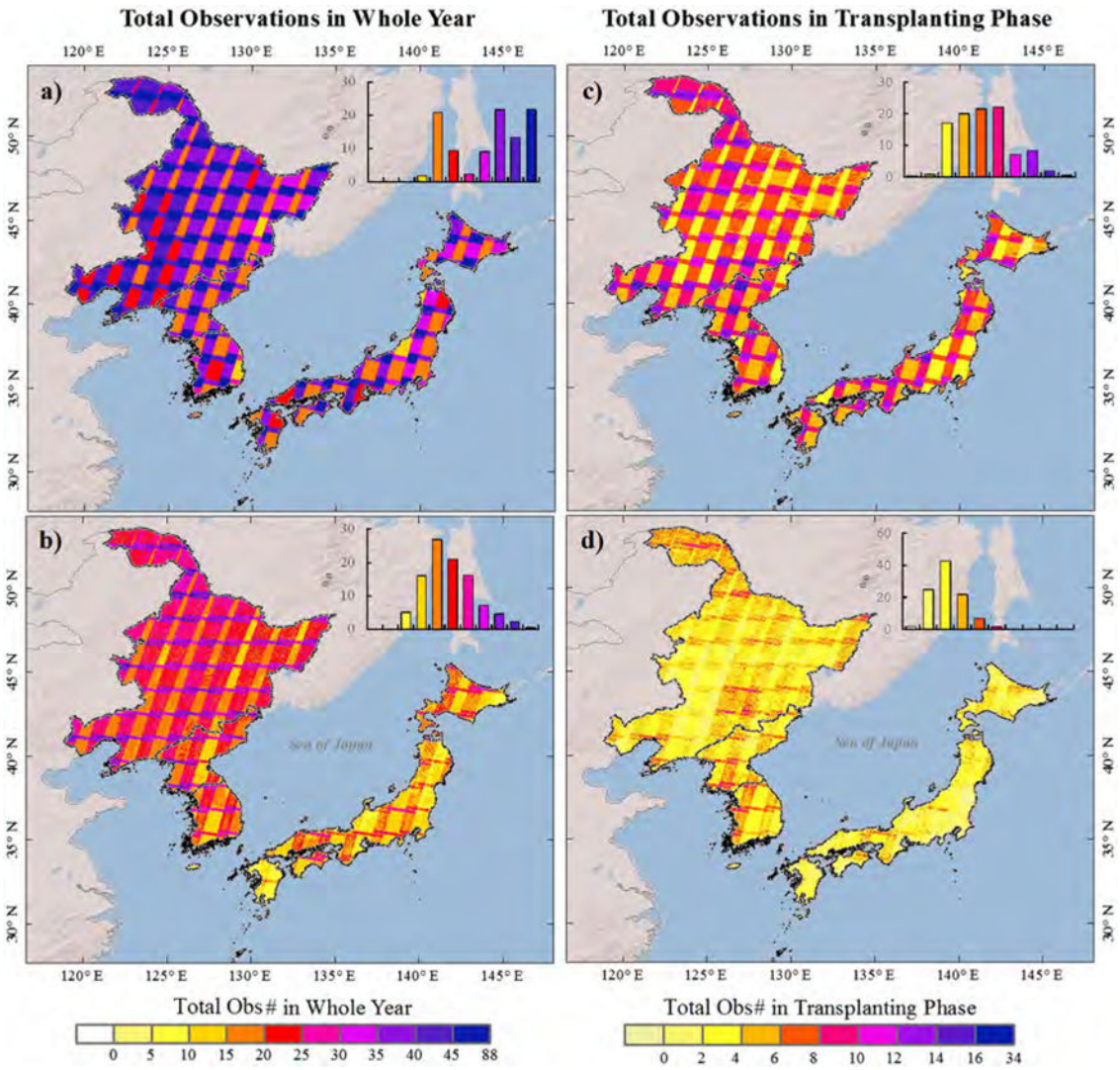


**Fig. 2.**

Availability of time series Landsat images in the study area for 2014. The total observation numbers in the whole year for a) Landsat 8 and b) Landsat 7; the total observation numbers in the transplanting phase for c) Landsat 8 and d) Landsat 7. The inset histograms show the distribution of Landsat imagery numbers with different intensities; the legends are consistent with the map legends.

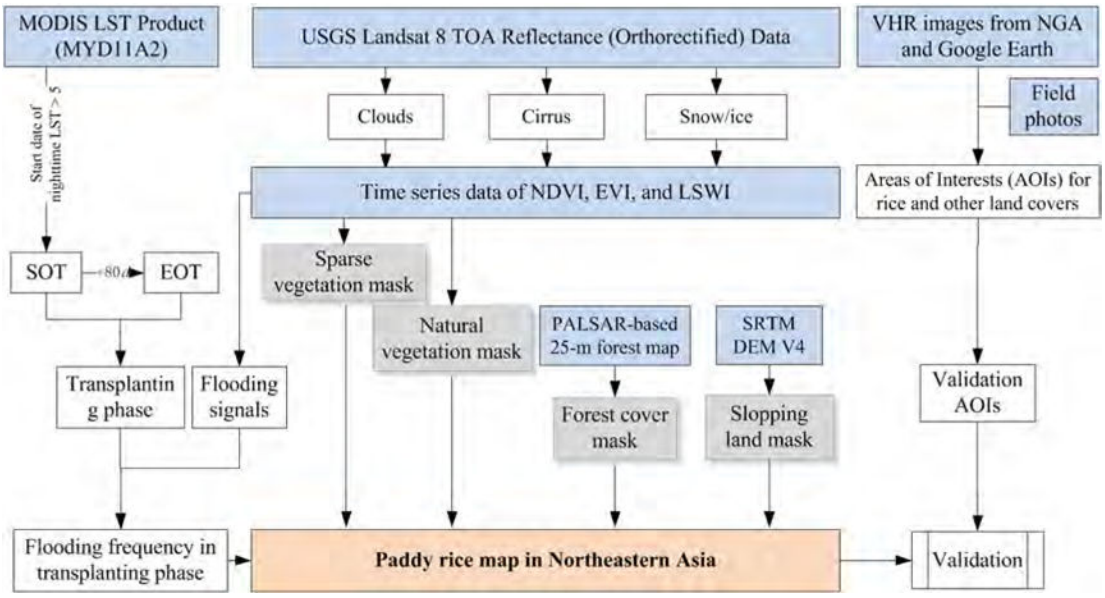


**Fig. 3.**

The workflow for the phenology and pixel-based paddy rice mapping in this study, major modules include time window determination of the rice transplanting phase (starting point: SOT, ending point: EOT), Landsat data preprocessing, phenology- and pixel-based mapping for non-cropland masks and paddy rice flooding, validation based on the areas of interest (AOIs) from very high resolution (VHR) images and field photos.

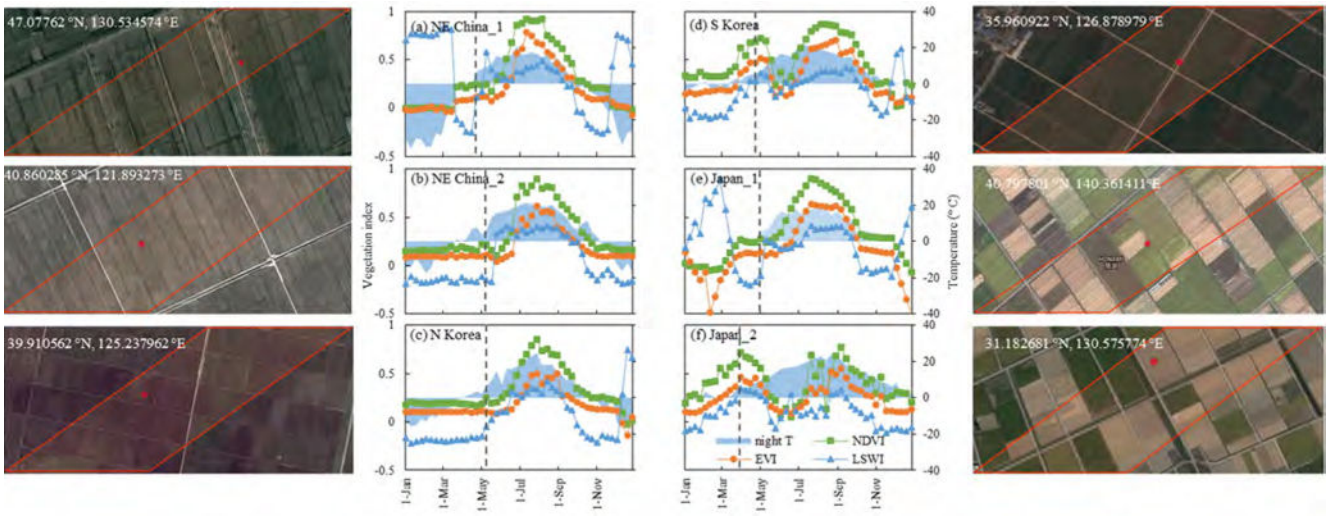


**Fig. 4.**

Temporal profile analysis of MODIS vegetation indices (*NDVI*, *EVI*, and *LSWI*) and nighttime land surface temperature (LST) at selected paddy rice sites in different biomes and climate types: a) Dwb and single cropping, b) Dwa and single cropping, c) Dwa and single cropping, d) Dfa and single cropping, e) Cfa and single cropping, and f) Cfa and double cropping (paddy rice plus others). See Fig. 1 for locations of selected sites and detailed descriptions about different climate types. Both the VIs and LST data have been gap-filled by using the linear interpolation approach. The surrounding snapshots show the landscapes of the corresponding paddy rice fields. The start points of LST over 5 °C are marked in the figures with gray dashed line.

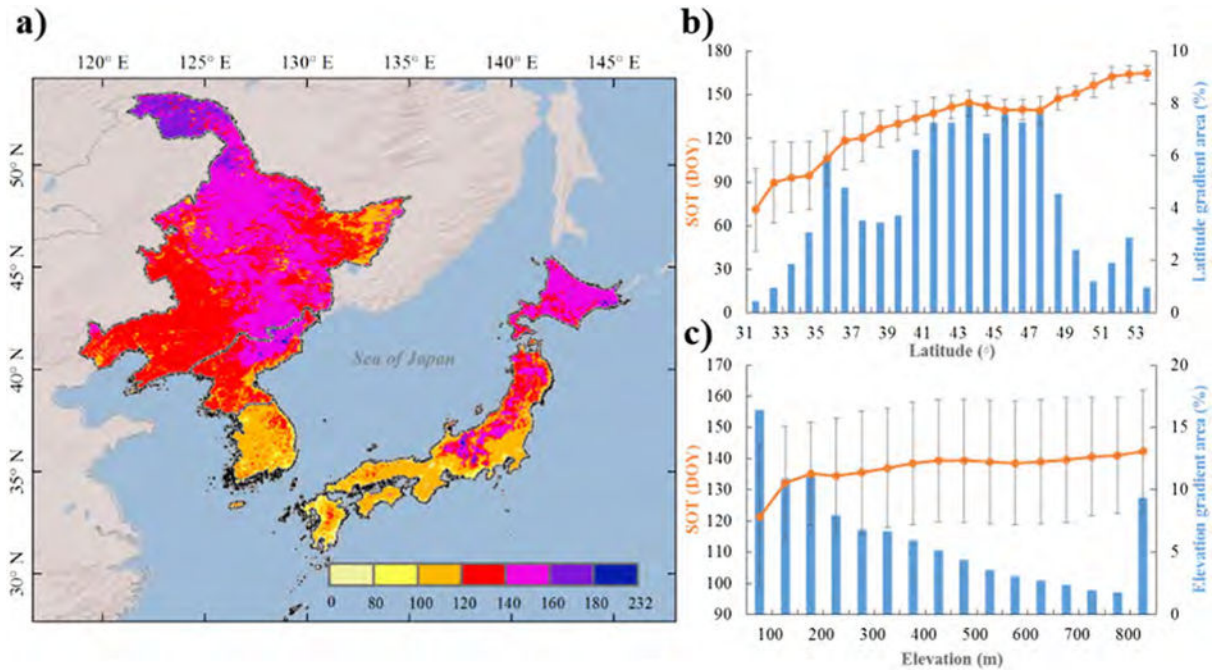


**Fig. 5.**

The start of the transplanting phase defined by using MODIS nighttime land surface temperature (LST) data. a) Spatial distribution of the starts of the transplanting phase (SOT) according to the LST over 5 °C in spring; b) variations of the SOT along with the latitude gradients; and c) variations of the SOT along with the elevation gradients.

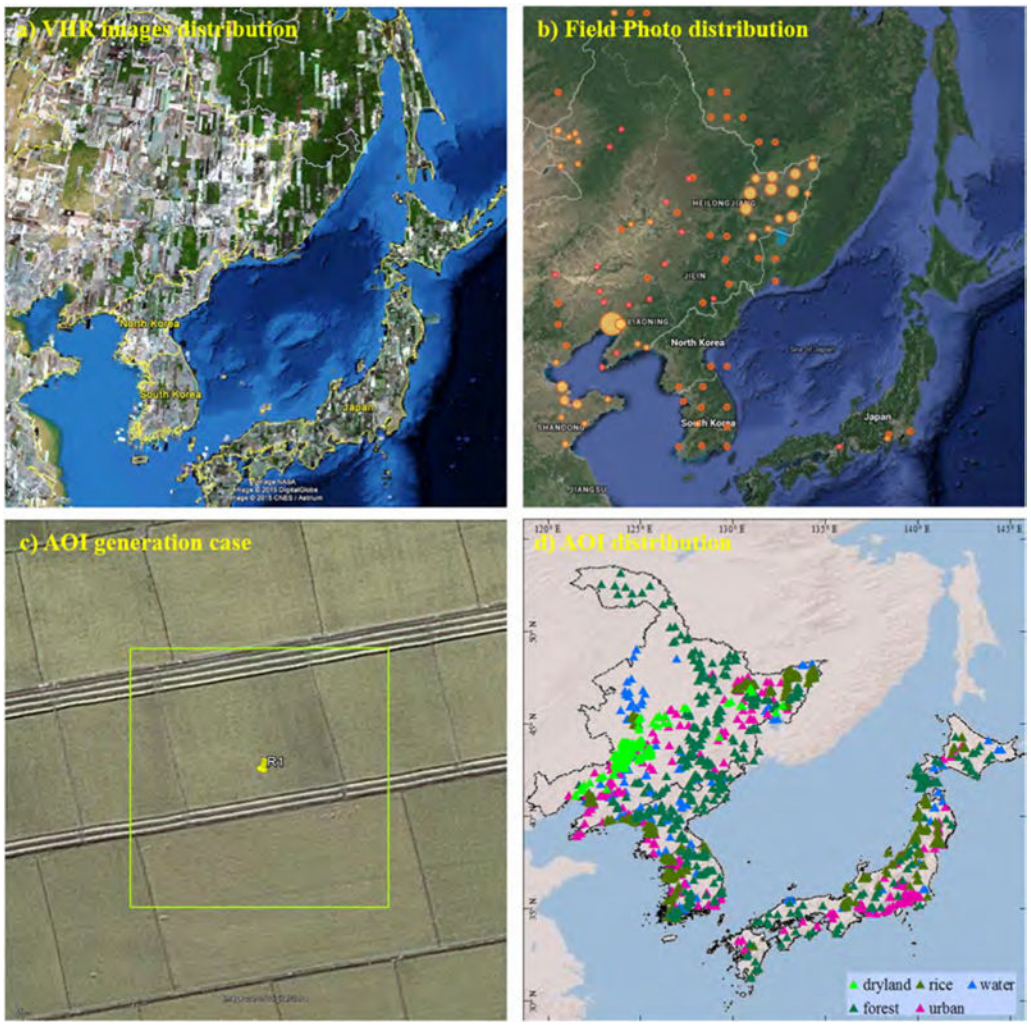


**Fig. 6.**

Extraction of the validation areas of interest (AOIs) for the resultant paddy rice map. a) The very high resolution (VHR) images from Google Earth around 2014 using historical imagery function; b) spatial distribution of field photos from the Global Field Photo Library (http://eomf.ou.edu/photos) at the University of Oklahoma, the cycle sizes stand for the photo intensity in the regions; c) a square AOI case of 100-m by 100-m; d) spatial distribution of the validation AOIs digitized referring to VHR images (a) and field photos (b).

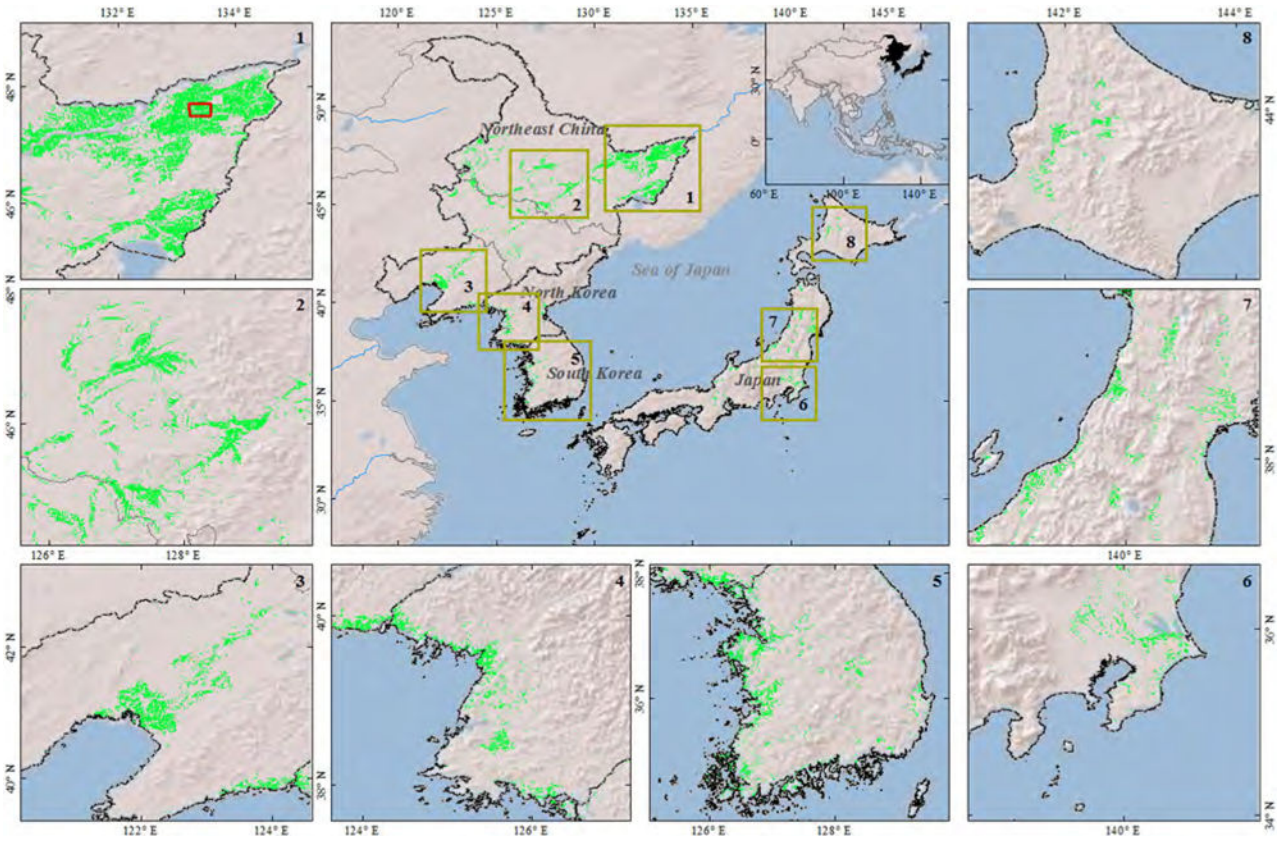


**Fig. 7.**

The resultant paddy rice map in the northeastern (NE) Asia in 2014 based on the phenology- and pixel-based paddy rice mapping (PPPM) algorithm, Landsat 8 imagery, and the Google Earth Engine (GEE) platform. The surrounding zoom-in maps show the local details in different countries and regions. The original 30-m map is resampled into 500-m in the central map and 100-m in the surrounded maps for visualization aim. Maps 1–8 show the spatial distributions of paddy rice in main production regions, including Sanjiang Plain, Songnen Plain, Liaohe Plain, western plains of North Korea, western plains of South Korea, Tokyo surrounding area, Middle Honshu, and Hokkaido, respectively. The red rectangle in Map 1 shows the extent of the case region in Fig. 8. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

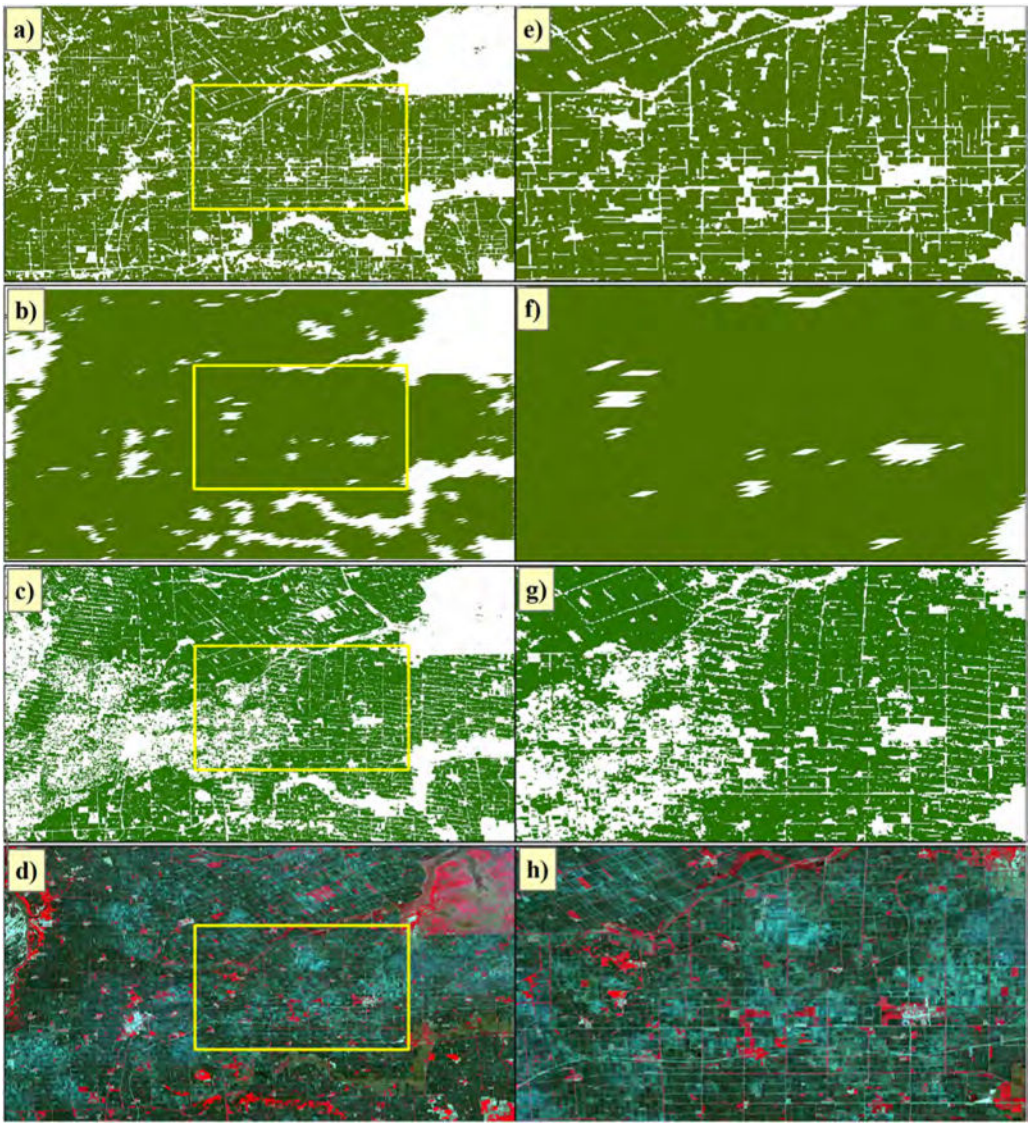


**Fig. 8.**

Comparison between Landsat 8-based 30-m paddy rice map, Landsat 7-based 30-m paddy rice map, and MODIS-based 500-m paddy rice map in 2014 in a typical paddy rice cropping region (the extent is marked in Fig. 7 Map 1) in Sanjiang Plain, NE China: a) 30-m Landsat 8-based result, b) 500-m MODIS-based result, c) 30-m Landsat 7-based result, and d) Landsat color composite image acquired on June 1, 2014. Maps e, f, g and h showcase the spatial details of a–d within the yellow box. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

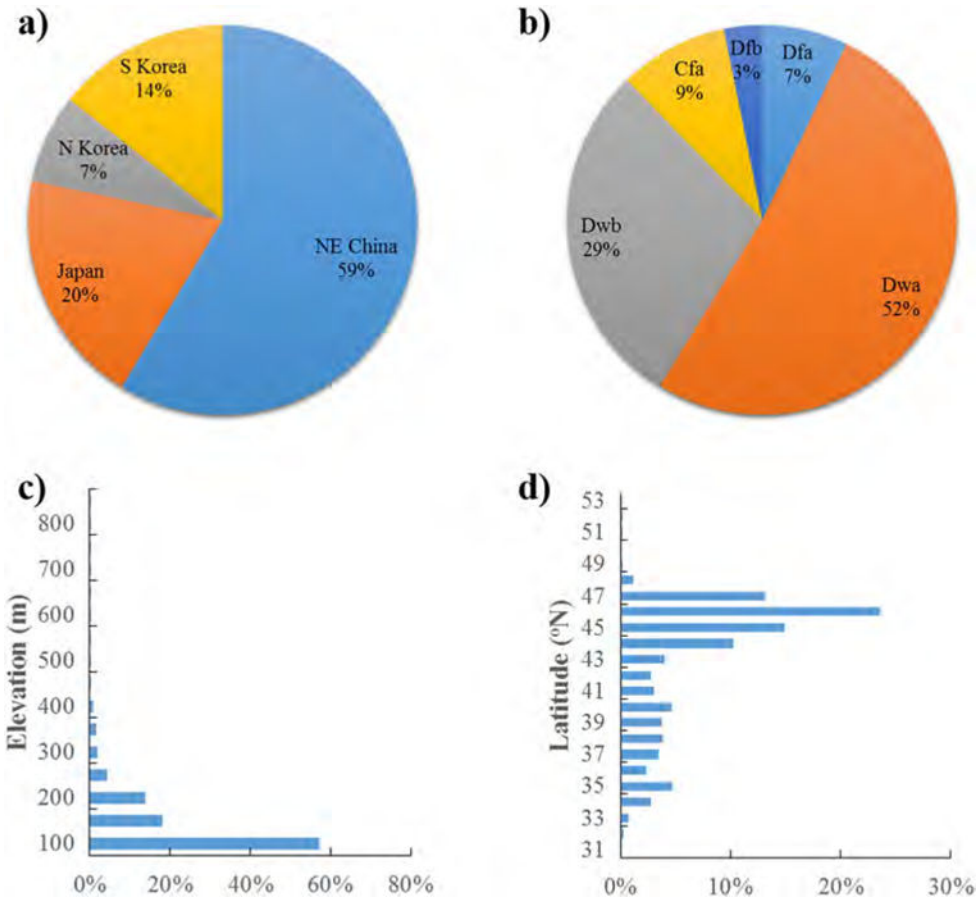


**Fig. 9.**

Statistical analysis of paddy rice area along different geographical regions: a) countries, b) climate zones, c) elevation gradients, and d) latitude zones. The abbreviations for the climate zones are described in Fig. 1 caption.

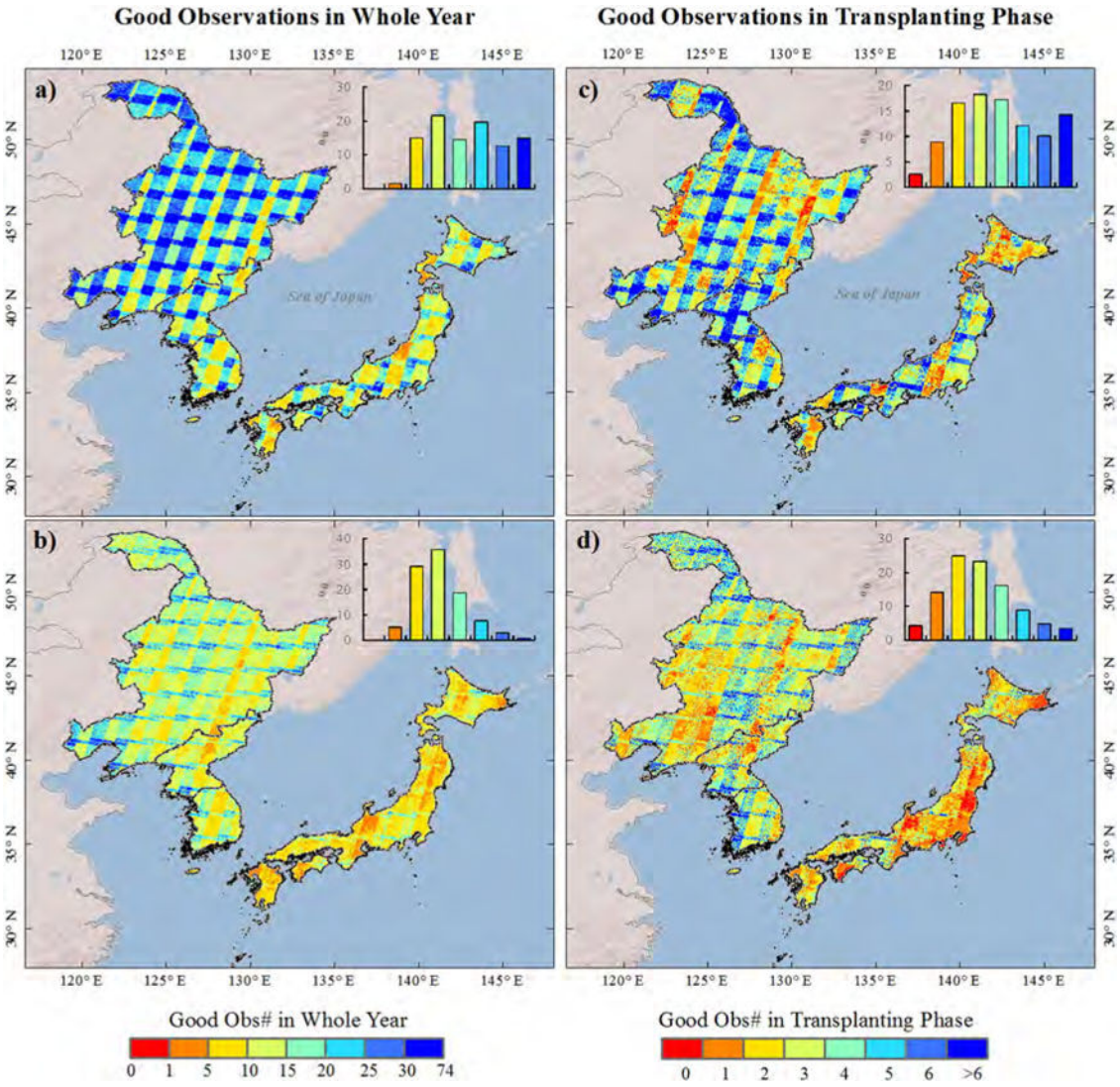


**Fig. 10.**

Spatial distribution of the good observation numbers of Landsat pixels for 2014 by excluding clouds, snow/ice and cirrus. The good observation numbers in the whole year for a) Landsat 8 and b) Landsat 7; the good observation numbers in the transplanting phase for c) Landsat 8 and d) Landsat 7. The inset histograms show the distribution of good observation numbers with different intensities; the legends are consistent with the map legends.

**Table 1**

Confusion matrix of land cover validation based on the areas of interest (AOIs) from very high resolution images of Google Earth and the National Geospatial-Intelligence Agency (NGA) Commercial Archive Data as well as the field photos from the Global Field Photo Library (eomf.ou.edu/photos). Accuracy measures are adjusted with the area of each stratum and presented with a 95% confidence interval (Card, 1982; Olofsson et al., 2013).

| Region | Class | Rice | Non-rice | Classif. total | User's acc. | Producer's acc. | Overall acc. |
|---|---|---|---|---|---|---|---|
| NE Asia | Rice | 17,536 | 1460 | 18,996 | 0.92 ± 0.004 | 0.73 ± 0.01 | 0.98 ± 0.001 |
| | Non-rice | 1351 | 61,637 | 62,988 | 0.98 ± 0.001 | 0.995 ± 0.0002 | |
| | Sample total | 18,887 | 63,097 | 81,984 | | | |
| Japan | Rice | 1877 | 191 | 2068 | 0.91 ± 0.01 | 0.52 ± 0.02 | 0.97 ± 0.003 |
| | Non-rice | 448 | 15,849 | 16,297 | 0.97 ± 0.003 | 0.997 ± 0.0004 | |
| | Sample total | 2325 | 16,040 | 18,365 | | | |
| NE China | Rice | 13,064 | 1064 | 14,128 | 0.93 ± 0.005 | 0.83 ± 0.01 | 0.98 ± 0.001 |
| | Non-rice | 427 | 29,434 | 29,861 | 0.99 ± 0.001 | 0.99 ± 0.0003 | |
| | Sample total | 13,491 | 30,498 | 43,989 | | | |
| N Korea | Rice | 1409 | 54 | 1463 | 0.96 ± 0.009 | 0.61 ± 0.03 | 0.97 ± 0.003 |
| | Non-rice | 207 | 7854 | 8061 | 0.97 ± 0.004 | 0.998 ± 0.0004 | |
| | Sample total | 1616 | 7908 | 9524 | | | |
| S Korea | Rice | 1186 | 151 | 1337 | 0.89 ± 0.02 | 0.50 ± 0.03 | 0.97 ± 0.004 |
| | Non-rice | 269 | 8500 | 8769 | 0.97 ± 0.004 | 0.996 ± 0.001 | |
| | Sample total | 1455 | 8651 | 10,106 | | | |