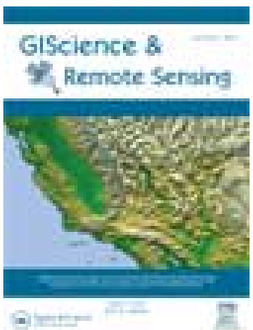

GIScience & Remote Sensing

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/tgrs20

Mapping rice area and yield in northeastern asia by incorporating a crop model with dense vegetation index profiles from a geostationary satellite

Jong-Min Yeom , Seungtaek Jeong , Ravinesh C Deo & Jonghan Ko

To cite this article: Jong-Min Yeom , Seungtaek Jeong , Ravinesh C Deo & Jonghan Ko (2021): Mapping rice area and yield in northeastern asia by incorporating a crop model with dense vegetation index profiles from a geostationary satellite, GIScience & Remote Sensing

To link to this article: https://doi.org/10.1080/15481603.2020.1853352

Published online: 11 Jan 2021.

GISCIENCE & REMOTE SENSING
https://doi.org/10.1080/15481603.2020.1853352

Taylor & Francis
Taylor & Francis Group

Mapping rice area and yield in northeastern asia by incorporating a crop model with dense vegetation index profiles from a geostationary satellite

Jong-Min Yeom[a], Seungtaek Jeong[a♪], Ravinesh C Deo[b] and Jonghan Ko[c]

[a]Satellite Information Center, Korea Aerospace Research Institute, Yuseong-gu, Daejeon, Republic of Korea; [b]School of Sciences, University of Southern Queensland, Springfield, QLD Australia; [c]Applied Plant Science, Chonnam National University, Buk-gu, Gwangju Republic of Korea

ABSTRACT

Acquiring accurate and timely information on the spatial distribution of paddy rice fields and the corresponding yield is an important first step in meeting the regional and global food security needs. In this study, using dense vegetation index profiles and meteorological parameters from the Communication, Ocean, and Meteorological Satellite (COMS) geostationary satellite, we estimated paddy areas and applied a novel approach based on a remote sensing-integrated crop model (RSCM) to simulate spatiotemporal variations in rice yield in Northeastern Asia. Estimated seasonal vegetation profiles of plant canopy from the Geostationary Ocean Color Imager (GOCI) were constructed to classify paddy fields as well as their productivity based on a bidirectional reflectance distribution function model (BRDF) and adjusted normalized difference vegetation indices (VIs). In the case of classification, the overall accuracy for detected paddy fields was 78.8% and the spatial distribution of the paddy area was well represented for each selected county based on synthetic applications of dense-time GOCI vegetation index and MODIS water index. For most of the Northeast Asian administrative districts investigated between 2011 and 2017, simulated rice mean yields for each study site agreed with the measured rice yields, with a root-mean-square error of 0.674 t ha^{-1}, a coefficient of determination of 0.823, a Nash-Sutcliffe efficiency of 0.524, and without significant differences (p-value = 0.235) according to a sample t-test (α = 0.05) for the entire study period. A well-calibrated RSCM, driven by GOCI images, can facilitate the development of novel approaches for the monitoring and management of crop productivity over classified paddy areas, thereby enhancing agricultural decision support systems.

ARTICLE HISTORY
Received 19 September 2020
Accepted 15 November 2020

KEYWORDS
crop model; rice classification; rice yield; geostationary satellite; northeast Asia

Introduction

Rice (*Oryza sativa* L.) is a staple crop for more than half of the world's population. It feeds approximately 557 million people in Asia, a continent that produces more than 90% of the world's rice (MacLean et al. 2013). The growing conditions required for rice cultivation are likely to deteriorate over the coming decades, as climate change is projected to have an increasingly adverse effect on rice crop yields, particularly in Asia (Bregaglio et al. 2017; IPCC 2013; Kim et al. 2013). Extreme rainfall and drought events in Northeast Asia are projected to occur more frequently (Kusunoki and Mizuta 2013), making it particularly challenging to secure the water resources required for rice cultivation in the region, especially with limited irrigation facilities. A significant increase in global food production will be necessary to sustain the projected (the year 2050) demands for human food and livestock feed (Barrett 2010; Tilman et al. 2011; Zhang et al. 2016). Accurate monitoring of spatiotemporal changes in rice productivity (area × yield) in key supplier regions (*e.g.*, Northeastern Asia) is an essential first step toward helping prepare a sustainable agricultural production environment and supporting informed agricultural decision-making (Yeom et al. 2018; Zhang et al. 2017).

Empirical crop modeling approaches were among the earliest of the broad range of methods employed to simulate spatial variations in rice productivity (Mosleh, Hassan, and Chowdhury 2015; Peng et al. 2014). However, such methods suffer from being largely dependent on field measurements of the target paddy and restricted to particular regions and time periods (Doraiswamy and Cook 1995; Reeves, Zhao, and Running 2005). An empirical crop model is constructed according to a particular principle using a procedure based on observation and experimentation. An empirical model can make the connection

CONTACT Jonghan Ko Email jonghan.ko@chonnam.ac.kr
♪Equal contribution goes to the first and second authors.

with experience in such a way that its derivation can be traced to a connection given-in-experience, e.g., describing relationships among variables and crop yields based on experimentation (Lai et al. 2018). Accordingly, it is difficult to predict time-dependent processes related to crop growth and yield – critical measures for real-time monitoring and forecasting of crop productivity – using empirical approaches.In contrast, process-based crop models can provide discrete estimates of crop growth and development through the simulation of crop biophysical processes (such as photosynthesis and evapotranspiration) in the soil-crop-atmospheric system (Hodson and White 2010; Ahuja et al. 2000). Such models, thus, have the advantage of being able to simulate crop growth and production variables based on independent agro-biophysical processes. However, crop models cannot provide detailed information on crop characteristics or estimate crop yields at a regional scale without the input of spatial biophysical and meteorological variables (Ko et al. 2005). By integrating a crop model and remote-sensing approaches in a complementary manner, it is possible to offset the weaknesses of each technique. Recent methods incorporating vegetation index from remote sensing into crop models have shown some success in estimating crop production at regional or continental scales (Huang et al. 2016; Lobell et al. 2015; Yeom et al. 2018).

In the present study, Communication, Ocean, and Meteorological Satellite (COMS) Meteorological Imager (MI) and Geostationary Ocean Color Imager (GOCI) products were incorporated into a remote sensing-integrated crop model (RSCM) to assess the spatiotemporal variation in paddy rice area and yield in Northeast Asia. The RSCM updated from a gramineous crop model that uses remote-sensing data, GRAMI (Maas 1992), can be employed for crop growth monitoring and yield mapping (Ko et al. 2015; Nguyen et al. 2019). RSCM can use simpler input parameters and variables than other process-based crop models owing to the integration with remote-sensing data. This feature allows RSCM to improve simulation accuracy by assimilating satellite images into the modeling procedure through the direct use of observed scene data. RSCM can carry out within-season simulation using operational satellite images obtained during the crop-growing season (Maas 1993b). In this process, simulated values are calibrated to agree with observed values based on the application of mathematical procedures using parameters that control crop growth. Approaches that improve model performance, the WOrld FOod STudies (WOFOST) (van Diepen et al. 1989) and the Simple and Universal CROp growth Simulator (SUCROS) models (Spitters, Van Keulen, and Van Kraalingen 1989), can be used to mathematically calibrate a model based on the integration of satellite images (Huang et al. 2016; Zhao, Chen, and Shen 2013; Launay and Guerif 2005). However, these approaches require the same number of input parameters and variables as most process-based crop models, which hampers the simulation of geographical projections of crop productivity, especially for data-sparse and inaccessible regions without scouting through satellites.

The process of obtaining spatial rice area and yield using satellite imagery can be divided into the mapping of paddy areas and the estimation of input parameters for modeling paddy rice productivity. To classify rice paddy areas, most optical satellite-based approaches use the unique flooding signatures generated from the transplanting phase to the early growing season under the specific water management strategies adopted in paddy rice fields (Xiao et al. 2006; Zhang et al. 2015). Classification methods include supervised procedures, such as the maximum likelihood (McCloy, Smith, and Robinson 1987) and support vector machine (Li et al. 2014), and unsupervised procedures, such as the Iterative Self-Organizing Data Analysis Technique (ISODATA; Nguyen et al. 2012). However, as a result of spectral variability based on time period and region, such approaches, in addition to their limitations in extending classifier rules and parameters, are region- and phase-dependent (Dong et al. 2016). Recently developed phenology-based methods using time-series vegetation indices (VIs) obtained from polar orbit satellites, such as Moderate Resolution Imaging Spectroradiometer (MODIS), Advanced Very High Resolution Radiometer (AVHRR) and Landsat, have been demonstrated to perform reliably in large-scale applications (Kontgis, Schneider, and Ozdogan 2015; Zhang and Zhang 2016; Zhang et al. 2017). Existing phenology-based methods can draw upon polar orbit satellite data obtained under clement-weather spring seasons to successfully capture the distinctive spectral signals of vegetation and water in paddy rice fields during the season's transplanting period. However, during

the monsoon summer growing season, when there are long periods of rain and numerous cloudy days, the instrument's relatively low temporal resolution results in polar orbit satellites increases the likelihood of discontinuous crop signal transitions, mainly due to extended cloudiness (Xiao et al. 2006; Yeom and Kim 2015).

From a rice yield perspective, it is also important to secure reliable and continuous rice vegetation index profiles, especially in the growing season, since the degree to which results extracted from indirect satellite imaging indicators (e.g., NDVI) accurately reflect real paddy rice crop phenology and productivity is closely correlated with the accuracy of national-scale productivity simulations (Kastens et al. 2005; Mkhabela et al. 2011; Wu et al. 2011; Zhang et al. 2017). In conclusion, to accurately classify paddy rice areas and simulate paddy rice productivity using a national-scale crop model, the degree to which indirect indicators extracted from satellite data (*e.g.*, NDVI) reflect real crop phenology remains an essential factor to consider (Kastens et al. 2005; Mkhabela et al. 2011; Wu et al. 2011; Zhang et al. 2017).

In the present study, high temporal resolution imagery from GOCI satellite was newly applied to obtained time-dense vegetation index profiles. Such data can produce higher temporal-resolution continuous NDVI-based phenology profiles for paddy rice than existing polar orbit satellite images, particularly during Northeast Asia's monsoon growing season (Yeom and Kim 2015). Accordingly, the present study's objectives were to classify paddy areas across Northeast Asia and to simulate the geospatial variability in rice productivity in the region using the RSCM, time-dense vegetation index profiles, and meteorological variables from a COMS geostationary satellite. The approach fills a significant knowledge gap, as there have been few efforts previously to generate annual maps of rice area and yield in Northeast Asia (including South and North Korea, Japan, and the northeastern part of China), or to perform a comparative analysis of the spatiotemporal patterns of paddy rice cultivation.

Study area and materials

Characteristics of the study area in northeast asia

The study region for rice cultivation and productivity was limited to the GOCI full-disk area, which fully covered the nations of South and North Korea, Japan, and the Manchuria region (northeast) of China (Figure 1). The latter region includes the important grain-growing areas of the Sanjiang, Songnen, and Liaohe plains (Dong et al. 2016). As a topographical feature, the average altitude of the study area is

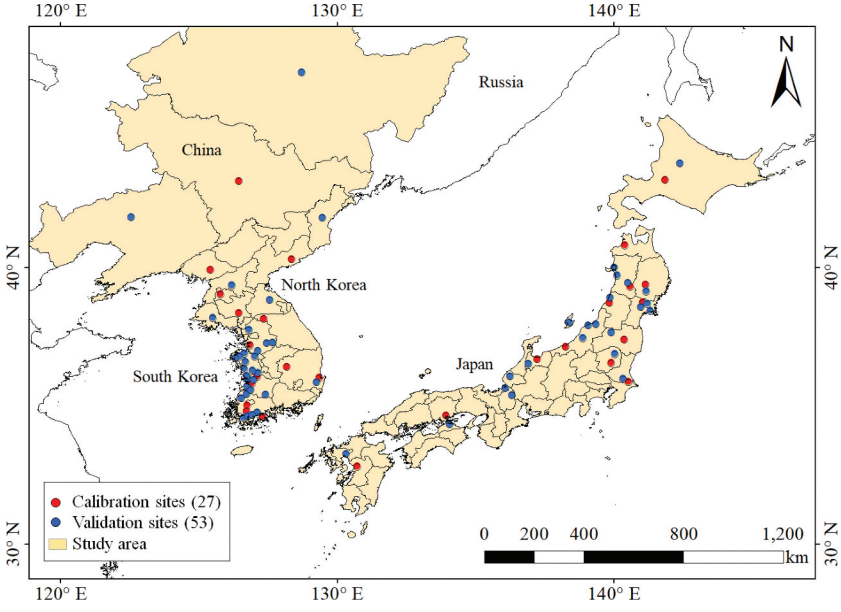

Figure 1. Study area for the geospatial simulation of rice yield. The red and blue circles indicate locations (80) for calibration (27) and validation (53), respectively, of the crop model in counties (or, in the case of North Korea and China, provinces) where rice is primarily cultivated.

approximately 900 m, and the variation in altitude within the area is large. The large rivers form a large-scale sedimentary plain downstream, creating a unique topographical feature in Northeast Asia. The surface slope in the study area ranges from 0 to 54 degrees, and the average slope is 3.94 degrees, indicating significant spatial variability (Park 2014). The study region's climate varies from the subtropical monsoon in southern Japan to the cold temperate monsoon in northern China (Peel, Finlayson, and McMahon 2007). The four countries are generally dominated by a single rice cropping system, as a result of their wide temperature ranges and temporal concentration of precipitation in the summer under a monsoon climate. The rainy monsoon season that uniquely characterizes East Asia is a phenomenon in which wet air forms a front between mid-June and the end of July, producing a large amount of precipitation as it moves north and south (Yoshikane, Kimura, and Emori 2001). During the rainy summer season, it is difficult to acquire polar orbit satellite imagery as a result of persistent clouds, making it reasonable to use high-temporal-resolution data obtained from the GOCI satellite observation area to directly or indirectly represent rice productivity in Northeast Asia.

RSCM evaluation sites for northeast asia

The parameterization and validation sites for the RSCM-rice run were selected from 34 counties in South Korea, 8 provinces in North Korea, 36 counties in Japan, and three provinces in northeast China, based on data availability (Figure 1). The parameterization and validation counties or provinces (80 sites) were carefully chosen to best represent each country's paddy growth environments. Rice was cultivated over an area of 778,734 ha in South Korea in 2016, according to the Korean Statistical Information Service (KOSIS). The current study used the rice yields reported by KOSIS to define the geospatial variability of yield throughout the study region. Table 1 presents information on the cultivated areas and production of rice in the study region.

COMS vegetation indices and solar radiation

The first geostationary satellite orbited by South Korea, COMS, was launched on 27 June 2010 by the Korean Aerospace Research Institute (KARI). The satellite has two payloads: GOCI, to monitor short-term biological phenomena and the ocean environment, and a meteorological imager (MI) to observe weather phenomena and atmospheric conditions (Table 2). In the present study, the vegetation indices and solar radiation data estimated from the GOCI and MI sensors, respectively, served as input variables to the RSCM.

GOCI obtains hourly top of atmosphere (TOA) reflectance values from 9 A.M. to 4 P.M. local time in eight multispectral bands from visible to near-infrared at a half-kilometer spatial resolution. An atmospheric correction for GOCI TOA reflectance values was performed to estimate surface reflectance using the second simulation of the satellite signal in the solar spectrum (6S) model. The GOCI surface reflectance data were further corrected using a semi-empirical bidirectional reflectance distribution function (BRDF) to correct for surface anisotropy effects. More details about the BRDF model are described in section 4.1 below. Four vegetation indices (VIs) were determined using the BRDF-adjusted GOCI reflectance data for use as inputs to the crop model: the NDVI (Rouse et al. 1974), the renormalized difference vegetation index (RDVI) (Rouse et al. 1974), the optimized soil-adjusted vegetation index (OSAVI) (Rondeaux, Steven,

Table 1. Information about the cultivated areas and productions of paddy rice in the study regions.

Region	*Area (ha)*	*Production (Tg [Mg ha^{-1}])*	*Data source*
South Korea	778,734	5.612 [7.23]	Korean Statistical Information Service (KOSIS, http://kosis.kr)
North Korea	468.677	2.536 [5.41]	United Nations Food and Agriculture Organization (FAO, http://www.fao.org), United States Department of Agriculture (USDA, https://www.fas.usda.gov)
Japan	1,478,511	8.043 [5.44]	Portal Site of Official Statistics of Japan (https://www.e-stat.go.jp)
NE China	4,454,000	32.975 [7.40]	China Statistical Yearbook (http://www.stats.gov.cn)

All the data obtained are based on the year 2016.

Table 2. Detailed characteristics of the COMS GOCI, COMS MI, and MODIS sensors for estimating crop productivity in Northeast Asia.

Satellite sensor	*Orbit type (Altitude)*	*Wavelength (μm)*	*Spatial resolution*
GOCI	Geo-synchronous (36,000 km)	B1: 0.402–0.422 B2: 0.433–0.453 B3: 0.480–0.500 B4: 0.545–0.565 B5: 0.650–0.670 B6: 0.675–0.685 B7: 0.735–0.755 B8: 0.845–0.885	500 m at nadir
MI	Geo-synchronous (36,000 km)	B1: 0.55–0.80 B2: 3.50–4.00 B3:6.50–7.00 B4:10.30–11.30 B5:11.50–12.50	1 km
MODIS	Sun-synchronous (≈ 705 km)	B1: 0.620–0.670 B2: 0.841–0.876 B3: 0.459–0.479 B4: 0.545–0.565 B5: 1.230–1.250 B6: 1.628–1.652 B7: 2.105–2.155	500 m at nadir for visible bands

and Baret 1996), and the modified triangular vegetation index (MTVI) (Haboudane et al. 2004). The detailed equations to determine the VIs are as follows:

$$NDVI = (R_{865}R_{660})/(R_{865}+R_{660}) \quad (1)$$

$$RDVI = (R_{865}-R_{660})/\sqrt{(R_{865}+R_{660})} \quad (2)$$

$$OSAVI = (R_{865}R_{660})/(R_{865}+R_{555}+0.16) \quad (3)$$

$$MTVI1 = 12.2 \cdot [12.2 \cdot (R_{865} - R_{555})2.5 \cdot (R_{660} - R_{555})] \quad (4)$$

where R_{865}, R_{660}, and R_{555} represent reflectance at 865, 660, and 555 nm from GOCI, respectively. We used all the four VIs to estimate LAI based on the empirical relationships between the VIs and the LAI. The fundamental idea is to adopt an ensemble approach based on the four relationships between the structural VIs and LAI.

Daily solar radiation with a 1-km spatial resolution was estimated by integrating hourly solar radiation of the COMS MI during the daytime using the Kawamura physical model (Kawamura, Tanahashi, and Takahashi 1998), as modified by Yeom et al. (2016). Given the complex physical characterization and time-consuming calculations involved in interpreting the radiative effects of atmospheric constituents and clouds from a hemispherical source, a grid-based physical model using instantaneous satellite observations and atmospheric information was employed. As the pass depth of clouds is more representative of the magnitude of radiance attenuation than to cloud temperature, the model applied an improved cloud factor based on the solar zenith angle and visible reflectance (Yeom, Han, and Kim 2012). Details of the methods of solar radiation estimation are described in Yeom et al. (2016). In addition, we resampled the 1-km resolution of the COMS MI solar radiation to a resolution of 500 m using the nearest-neighbor interpolation method.

MODIS reflectance product

In this study, we used MODIS 8-day composite near-infrared (NIR) and shortwave infrared (SWIR) spectral bands in the MYD09A1 product of the Aqua satellite (https://ladsweb.modaps.eosdis.nasa.gov/), as an additional dataset to calculate the land surface water index (LSWI), for the classification of the spatial distribution of cultivatable paddy fields, since GOCI only observes from the visible to the near-infrared (VNIR) spectral domain. The LSWI equation is as follows:

$$LSWI = (R_{865}-R_{2,130})/(R_{865}+R_{2,130}) \quad (5)$$

where R_{865} and $R_{2,130}$ represent surface reflectance at 865 and 2,130 nm from MODIS, respectively. The MYD09A1 data, at a spatial resolution of 500 m, were selected to reduce the effects of atmospheric water vapor and clouds over 8-day periods. To resolve the contamination issue during the crop-growing season, arising from the typically cloudy monsoonal conditions, we performed a simple linear interpolation to

average the indices obtained before and after poor pixels contaminated by clouds.

Numerical weather prediction reanalysis data and SRTM DEM data

The geographical variation in the RSCM's temperature data input was reproduced using the Regional Data Assimilation and Prediction System (RDAPS), based on a fifth-generation Mesoscale Model (Grell, Dudhia, and Stauffer 1994) provided by the Korea Meteorological Administration (KMA). The RDAPS reproduces temperature data as a reanalysis weather variable, with a 12-km three-dimensional resolution across 70 levels, from the ground surface to about 40 km. In addition, digital elevation model (DEM) data were generated by the Shuttle Radar Topography Mission (SRTM) model of the United States National Aeronautics and Space Administration (NASA). Drawing upon Interferometric Synthetic Aperture Radar data, the SRTM model is designed to provide a world DEM at a spatial resolution of 90 m (Rabus et al. 2003). The data were employed to compute the threshold values for surface elevation and slope for arable land conditions. Detailed information on the obtained and manipulated geospatial data are presented in Table 3. RDAPS and SRTM DEM datasets were resampled to 500 m using the nearest-neighbor interpolation method.

Methodology

Temporal profiles of NDVIs from the semi-empirical BRDF model

It was very important to secure a stable NDVI profile in both paddy rice classification and crop yield prediction to obtain input parameters for the crop model for the countries that are influenced by monsoon. Although both Terra and Aqua MODIS can provide frequent observations, it remains difficult to secure continuous NDVI profiles over the study area during the cloudy summer season. We assumed that measuring the crop area more frequently would provide a more reliable reflection of the vegetation's actual condition and used the output of the GOCI sensor aboard a geostationary satellite, rather than the MODIS products. A semi-empirical BRDF model was applied to correct the primarily solar-generated surface anisotropy effects occurring in geostationary satellite images as well as to estimate angularly independent crop phenology using GOCI NDVI profiles. Specifically, a linear combination of three basic scattering kernels from the Ross-Thick Li-Sparse Reciprocal (RTLSR) – volumetric, geometric, and isotropic – was used (Ross and Marshak 1988; Roujean, Leroy, and Deschamps 1992):

$$\rho(\theta_s, \theta_v, \emptyset) = f_{iso} + f_{geo}k_{geo}(\theta_s, \theta_v, \emptyset) + f_{vol}k_{vol}(\theta_s, \theta_v, \emptyset) \quad (6)$$

where k_{geo} is the LiSpare-Reciprocal geometric kernel, k_{vol} is the Ross-Thick volumetric kernel, f_{iso} is the Lambertian reflectance at the nadir direction, f_{geo} is the coefficient of k_{geo} (the LiSpare-Reciprocal geometric kernel), f_{vol} is the coefficient of k_{vol} (the Ross-Thick volumetric kernel), θ_v is the viewing zenith angle, θ_s is the solar zenith angle, and $\emptyset$ is the relative azimuth angle.

To estimate the adjustable surface reflectance ($\rho(\theta_s, \theta_v, \emptyset)$) using the BRDF model, the kernel coefficients of f_{iso}, f_{vol}, and f_{geo} were fitted based on multiple linear regression using entire sets of hourly surface reflectance data during 16-day composite periods; 128 samples of surface reflectance observations were used to normalize each of the data points. In addition, BRDF adjustable surface reflectance was simulated only when at least seven observations of clear surface reflectance were secured. The

Table 3. Information on collected and manipulated geospatial data for detection of the spatial distribution of paddy fields and transplanting dates and for the simulation of rice yield and production in Northeast Asia from 2011 to 2017.

Purpose	Data type	Produced data	Spatial resolution
Detection of paddies and transplanting dates	MODIS reflectance	Water indices	500 m
	SRTM DEM	Elevation and surface slope	90 m
	KME land cover	Paddy field map of South Korea	5 m
	JAXA land cover	Paddy field map of Japan	30 m
Simulation of rice yields and production	COMS GOCI	Vegetation indices	500 m
	COMS MI	Solar radiation	1,000 m
	LDAPS	Temperatures	12,000 m

determined coefficients by the BRDF model kernels were used to normalize the fixed view angle and mean solar zenith angle using Eq. (6) for each pixel grid (Lucht, Schaaf, and Strahler 2000; Schaaf et al. 2002, 2011). Thus, the BRDF model in Eq. (6) allowed for the estimation of the BRDF-adjusted reflectance (BAR) for GOCI (Yeom and Kim 2015).

Phenology-based paddy rice classification in northeast asia

Paddy rice (*Oryza sativa* ssp. japonica) grows in wetlands, making it unique among staple crops. Accordingly, this study focused on detecting irrigated paddy fields by using all available spectral indices from GOCI and MODIS data to classify paddy fields based on interrelationships between NDVIs obtained from GOCI data, the land surface water index (LSWI) obtained from MODIS, and the threshold (*T*) of the MODIS LSWI. This approach, proposed by Xiao et al. (2005); Xiao et al. (2002), is based on the constraint that LSWI + *T* ≥ NDVI (Figure 2). The *T* value is a threshold for enhancing the sensitivity of the LSWI in mixed pixels, which is connected with the land cover heterogeneity of the study area. Therefore, the land cover characteristics could have an influence on the determination of the *T* value. Using this methodology, it is not only possible to detect paddy fields, but also to determine transplanting dates, a factor vital for determining the initiation of paddy growth in most crop models.

In addition to the T value, we used several thresholds related to the characteristics of rice cultivation. Thresholds of altitude and surface slope from SRTM DEM data were used to reflect the topographical features of paddy fields and determine environments in which it was impossible to cultivate paddy rice because of high altitudes or severe slopes. The maximum NDVI attained during the reproductive stage was based on growing degree days (GDD) and the minimum NDVI on the transplanting date, based on NDVI and LSWI conditions. Finally, the NDVI increase rate was used as a threshold to reflect the distinctive growth rate of paddy rice relative to other plants (Figure 2). The equation is as follows:

$$NDVI\ increase\ rate = (NDVI_{max} - NDVI_{min})/(DOY_{NDVImax} - DOY_{NDVImin}) \times 1,000 \quad (7)$$

where $NDVI_{max}$ and $NDVI_{min}$ are the maximum and minimum NDVI from transplanting (about May to June) and heading periods (about July to August), respectively. $DOY_{NDVImax}$ and $DOY_{NDVImin}$ are the dates on which $NDVI_{max}$ and $NDVI_{min}$ were obtained. 1,000 is a constant for scaling the value. Additionally, pixels corresponding to days which satisfied the paddy detection condition but on which the air temperature was lower than 10°C were ignored, as transplanting was not possible under these conditions

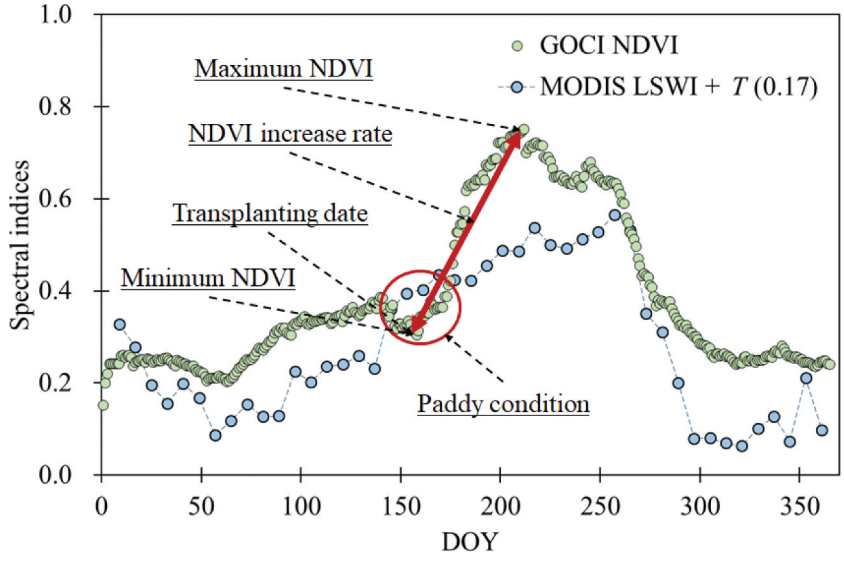

Figure 2. Seasonal variations in the normalized difference vegetation index (NDVI) derived from geostationary ocean color imager (GOCI), and land surface water index (LSWI) derived from moderate resolution imaging spectroradiometer (MODIS) imagery of a paddy field in South Korea. *T* (0.17) is the threshold value of the LSWI.

(Dingkuhn and Miezan 1995). According to the daily NDVI from GOCI, the irrigation period for satisfying detection conditions was defined to be at least one week.

The paddy field map was obtained from high-spatial-resolution land cover maps of South Korea (based on KOrea Multi-Purpose SATellite-2, KOMPSAT-2) and Japan (based on Advanced Land Observing Satellite, ALOS) produced in 2013 and 2014, respectively. The maps were provided by the Environmental Geographic Information Service (EGIS) of the Korean Ministry of the Environment (KME) (https://egis.me.go.kr) and by the Earth Observation Research Center (EORC) of the Japan Aerospace Exploration Agency (JAXA) (https://www.eorc.jaxa.jp). Given the inaccessibility of official national land cover data for the northeast region of China and North Korea, these nations' paddy areas were assumed to have the same distinguishing characteristics as South Korea and Japan because there is no officially reported spatial data. Paddy fields in all the study regions are mostly transplanted under the monocropping practice. Rice cultivation depends on flood irrigation during the summer growing season after transplanting.

RSCM for rice

The integration of remotely sensed data into the RSCM for rice (Figure 3) enabled the simulation and monitoring of crop growth and the capacity to make productivity estimates, based on the agreement between simulation and observation from remote-sensing information (Ko et al. 2015). The RSCM for rice simulates the daily growth of rice using four processes: (1) determining daily growing degree days, (2) absorption of incoming solar radiation by the crop canopy, (3) daily increase in above-ground dry mass, and (4) daily change in the LAI (increase or senescence), based on mathematical equations (Eqs. A1–A4). The RSCM fits a simulated LAI curve to the observed LAI curve using a Bayesian procedure as a part of the within-season calibration method. The LAI concept employed in this study is the green LAI, as the modeling regime is designed to absorb solar radiation through photosynthesis in the vegetation canopy. The LAI was indirectly estimated using the four VIs determined from GOCI imageries onboard the COMS geostationary satellite.

The RSCM was initially designed to accept remote-sensing data as an input variable to obtain a mathematical agreement between simulation and measurement based on a "within-season" calibration procedure (Maas 1993a). In this study, four parameters were employed by the RSCM to model rice crop growth processes: L_0, the initial LAI at transplanting; *a* and *b*, parameters of the leaf allocation function; and *c*, the parameter of the leaf senescence function. The Bayesian method for optimization of these parameters was applied to obtain these parameter values with prior distributions selected in relation to estimates from previous findings. The LAI is a three-dimensional measure of the volume of a crop canopy, while the reflectance of the crop canopy to solar

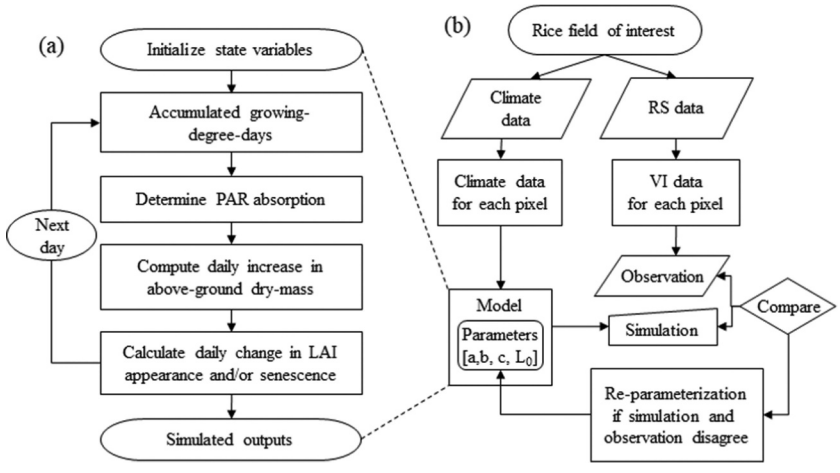

Figure 3. Schematic diagram of remote sensing-integrated crop modeling (RSCM) system used (a) to simulate crop growth and yield and (b) to project spatiotemporal crop production based on the crop modeling process and remote sensing (RS) data. (LAI: leaf area index; PAR: photosynthetically active radiation; VI: vegetation index).

radiation is a two-dimensional measure corresponding to the flat surface atop the canopy. Accordingly, the square root of reflectance should be proportional to the cube root of the LAI, and we presumed that the relationship between the two could be determined using a log-log regression regime with a slope of about two-thirds (McMahon 1973). Using this approach, we quantified the relationships between the LAI and the four VIs (NDVI, RDVI, OSAVI, and MTVI) with log-log linear regression models, which were adopted to obtain more robust LAI estimates using an ensemble approach based on the relationships between the four Vis and LAI. The current approach was developed earlier in the RSCM formulation process to achieve the best fit between remote-sensing data and LAI (Nguyen et al. 2019).

An empirical model was designed for each VI using the following equation.

$$\log(\mathrm{VI}_t) = \alpha_{\mathrm{VI}} + \beta_{\mathrm{VI}} \log(\mathrm{LAI}_t) + \in_t \tag{8}$$

,

where α_{VI}, β_{VI}, and $\in_t$ $(\sim N(0, \sigma_{VI}^2)$ represent the intercept, slope, and error of the linear regression model, respectively. The evolution of the LAI for each pixel was explained by the RSCM-rice regime, using four parameters $\theta =$ (L_0, a, b, and c). These parameters were assumed to be generated from the prior distribution$\psi \sim \mathrm{N}(\mu, \mathrm{D})$, where transformations were used to guarantee that all four parameters (L_0, a, b, and c) range between 0 and 1 as follows.

$$\begin{aligned}\psi &= (\psi_1, \psi_2, \psi_3, \psi_4) \\ &= \left(\log\frac{a}{1-a}, \log\frac{b}{1-b}, \log\frac{c}{1-c}, \log\frac{L_0}{1-L_0}\right)\end{aligned}$$

,

$$\theta = \theta(\psi) = \left(\frac{e^{\psi_1}}{1+e^{\psi_1}}, \frac{e^{\psi_2}}{1+e^{\psi_2}}, \frac{e^{\psi_3}}{1+e^{\psi_3}}, \frac{e^{\psi_4}}{1+e^{\psi_4}}\right) \tag{9}$$

We obtained both the regression coefficients $(\alpha_\ell, \beta_\ell, \sigma_\ell^2), \ell = 1, 2, 3, and 4$ and the hyper-parameters (μ, D) from the data collected in previous studies (Ko et al. 2015; Kim et al. 2017). These included both the VIs and the measured LAI values. The parameter μ was specified using the 'before-calibration' values (Table A1). Parameter D is a diagonal matrix with all diagonal elements equivalent to 0.5.

The following numerical procedure was adopted to obtain θ for each pixel.

Step 1: For each pixel, the set μ served as the initial guess ofψ.

Step 2: $\mathrm{LAI}_t = \mathfrak{G}(t;\psi) = G(t;\theta(\psi))$ was defined and the objective function below considered.

$$\sum_{\ell=1}^{5}\left\{\frac{1}{\sigma_\ell^2}\sum_{t=1}^{n}\left(\log VI_{\ell t} - \alpha_\ell - \beta_\ell \log \tilde{G}(t;\psi)\right)^2\right\} + (\psi-\mu)' D^{-1}(\psi-\mu) \tag{10}$$

Step 3: The simulated curve for each pixel was generated from the estimated ψ in Step 2.

Step 4: μandD were updated as the sample means and sample variances of the estimates in Step 2.

In this procedure, the parameter ψ was estimated by minimizing the above function, and the optimization was performed using the POWELL optimization routine (Press et al. 1992) for one-point simulation cases and the Quasi-Newton minimizer (Nash 1990) for 2-dimensional simulation cases. Previously established as an extended system of the RSCM-rice, the crop information delivery system uses remotely sensed images from various platforms to simulate grid-based crop growth and yield maps (Ko et al. 2015) (Figure 3). This system receives grid-based remote sensing and climate data as input variables to the RSCM-rice system. When available, climate data are extracted from one or more weather stations as grid data, and the RSCM-rice is then implemented to simulate crop growth in each grid using both types of input variables. Details of the modeling procedures and related equations are provided in Ko et al. (2015) and Kim et al. (2017). In the current study, we applied the same initial conditions and parameter values used in these previous studies to simulate geographical rice productivity in Northeast Asia using the RSCM-rice system (see also Table A1). As described earlier in the study area section (section 3), the modeling regime was calibrated for 27 locations and validated for 53 locations. These locations were chosen considering the representation of each country's paddy production conditions, randomly dividing the calibration and validation sites at a 3:7 ratio.

Statistical analysis

For the training and validation of paddy field classifications, paddy, and non-paddy pixels, in total 225,000 pixels, were randomly selected from the reference paddy field maps, out of which 75,000 were paddy pixels, and 150,000 were non-paddy pixels. The pixels were divided into a 150,000-pixel (50,000 paddy and 100,000 non-paddy pixels) training set and a 75,000-pixel (25,000 paddy and 50,000 non-paddy pixels) validation set. To evaluate the classified results, accuracy metrics, including overall, user, and producer accuracies, in addition to Cohen's Kappa Coefficient (Cohen 1968) were used.

The simulation performance of RSCM was evaluated using three statistical indices: the root mean squared error (RMSE), the coefficient of determination (r^2), and the Nash-Sutcliffe model efficiency (NSE) (Nash and Sutcliffe 1970) as well as paired sample *t*-tests in the R software (https://www.r-project.org/). NSE determines the relative magnitude of the residual variance of the simulated data in comparison with that of the observed data, designed to evaluate how well simulated versus measured data fit the 1:1 line. NSE values can vary from $-\infty$ to 1. The modeling regime is more dependable if the value is closer to 1. Simulated values are lower or larger than the observed data when NSE values are close to zero.

Results

Comparison of time-dense profiles of vegetation index from BRDF-adjusted GOCI and MODIS surface reflectance

Figure 4 presents an example of the annual variation in seasonal paddy rice dynamics obtained using BRDF-adjusted GOCI- and MODIS-based NDVI profiles from 2011 to 2017. Under clement weather conditions in the spring and fall, both GOCI (red circles) and MODIS (gray circles) effectively captured the NDVI profiles over the rice paddy areas. During the summer rainy season (light gray boxes), however, MODIS produced discontinuous crop signal transitions owing to its limited acquisition of observations during the period. The GOCI BAR NDVIs accurately captured the respective rice crop growth and development patterns corresponding to the sowing, transplanting, heading, and harvesting stages (Kim et al. 2016). In particular, the NDVI values for the rice-growing paddy season, based on their higher temporal resolution, exhibited a stable pattern despite the

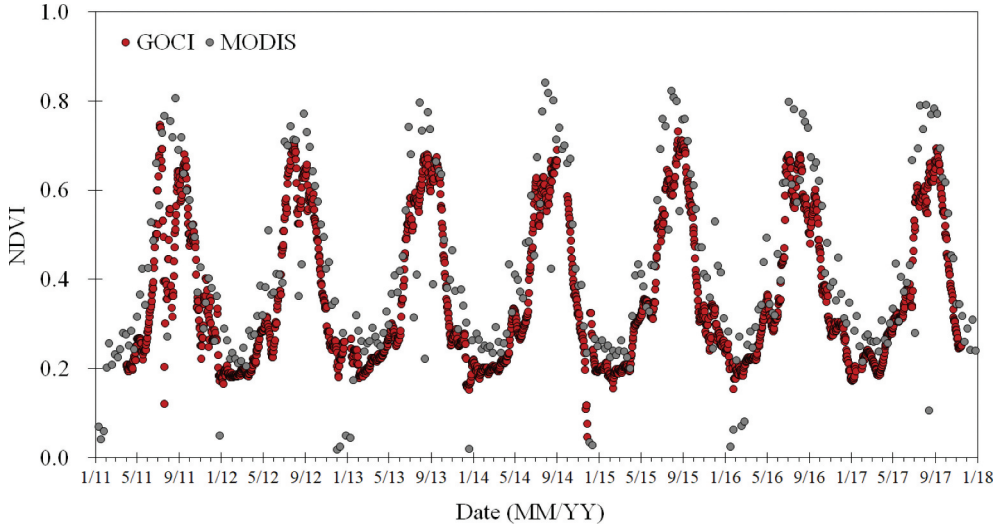

Figure 4. Annual variations in normalized difference vegetation index (NDVI) profiles between geostationary ocean color imager (GOCI) and moderate resolution imaging spectroradiometer (MODIS) in a sampled paddy pixel in the study area. MODIS NDVIs are 8-day composite data from the MYD09 A1 product (gray circles), and BRDF adjusted GOCI NDVIs (red circles) is daily data. when simulating BRDF model of GOCI imagery, 128 samples of surface reflectance observations during a 16-day composite period were used to normalize each of the data points. The light gray boxes are summer monsoon periods in the study area.

rainy summer weather conditions, which suggests that the continuous NDVI profiles produced by GOCI are potentially viable alternatives for addressing the previously discussed challenges experienced during the classification of paddy rice areas and estimation of rice productivity at national scales. For reference, the NDVI profiles of MODIS in Figure 4 are due to lack of BRDF correction; the present study used official products of NDVI profiles from MODIS 8-day composite surface reflectance (MYD09A1) to assess the capability of current polar orbit MODIS products. MODIS Nadir BRDF-adjusted Reflectance (NBAR) products (MYD43) or BRDF MODIS/Multi-Angle Implementation of Atmospheric Correction (MAIAC) products would be superior choices for minimizing BRDF effects. However, an inherent persisting limitation of polar orbit satellite is its inability to acquire continuous rice growth and development patterns in spite of BRDF correction (Yeom and Kim 2015).

Classification of rice area in northeast asia

Following the determination of the optimal values by combining and changing each threshold values for paddy rice field classifications, the following values were set: altitude, 800 m; surface slope, 7°; maximum and minimum NDVIs, 0.43 and 0.42; NDVI increase rate, 2.0; and *T* value, 0.17. Following the application of these thresholds to the validation dataset, the overall accuracy and Kappa coefficient for detected paddy fields were found to be 78.8 and 51.2%, respectively (Table 4). The user accuracies for paddy and non-paddy fields were 69.7 and 82.7%, respectively, while the producer accuracies were 64.1 and 86.1%, respectively. Figure 5 shows a representative map of classified paddy fields based on GOCI and MODIS for the study period. A representative map was estimated by compositing the overlapping areas of seven-rice classification maps from 2011 to 2017. The spatial distribution of the paddy rice area was well described for each selected country.

In the case of the Korean peninsula, most rice cultivation areas are located in the western plains (Jeong, Ko, and Yeom 2018), as shown in Figure 5. In the northeast part of China, paddy rice fields were mostly distributed in the Sanjiang Plain of Heilongjiang Province (the northeast region) and in a plain in Liaoning Province (the western region), which is consistent with the findings of previous studies (Clauss, Yan and Kuenzer 2016; Zhang et al. 2017). In Japan, paddy fields were mainly distributed along the central-eastern region and along the coastline (Inoue, Ito, and Yonezawa 2020). The classified spatial distribution patterns of paddy fields were well reflected in the actual fields. In addition, from 2011 to 2017, the paddy field areas showed decreasing trends. For reference,

Table 4. Error matrix to validate the spatial distribution of detected paddy fields from geostationary ocean color imager (GOCI) and moderated resolution imaging spectroradiometer (MODIS) data for South Korea in 2013 and Japan in 2014.

		Reference			
		Paddy	Non-paddy	Total	User's accuracy (%)
Classification	Paddy	16,021	6,953	22,974	69.7
	Non-paddy	8,979	43,047	52,026	82.7
	Total	25,000	50,000	75,000	
Producer's accuracy (%)		64.1	86.1		

Overall Accuracy = 78.8%, Kappa coefficient = 51.2%

Table 5. Root mean square errors (RMSE), coefficient of determination (r^2), nash-sutcliffe efficiencies (NSE), and p values of two-sample t-tests (α = 0.05) between observed and simulated rice yields for the model calibration in 27 and validation in 53 selected counties (or provinces) in northeast asia for the entire years (2011–2017). SD is standard deviation.

Site	Observed mean yield (SD)	Simulated mean yield (SD)	RMSE	r^2	NSE	t-test
-	– – – – – – – t ha^{-1} – – – – – –			-	-	– *p* –
Calibration	6.12 (± 0.99)	6.04 (± 1.07)	0.551	0.903	0.692	0.235
Validation	6.25 (± 0.98)	6.15 (± 1.14)	0.674	0.823	0.524	0.434

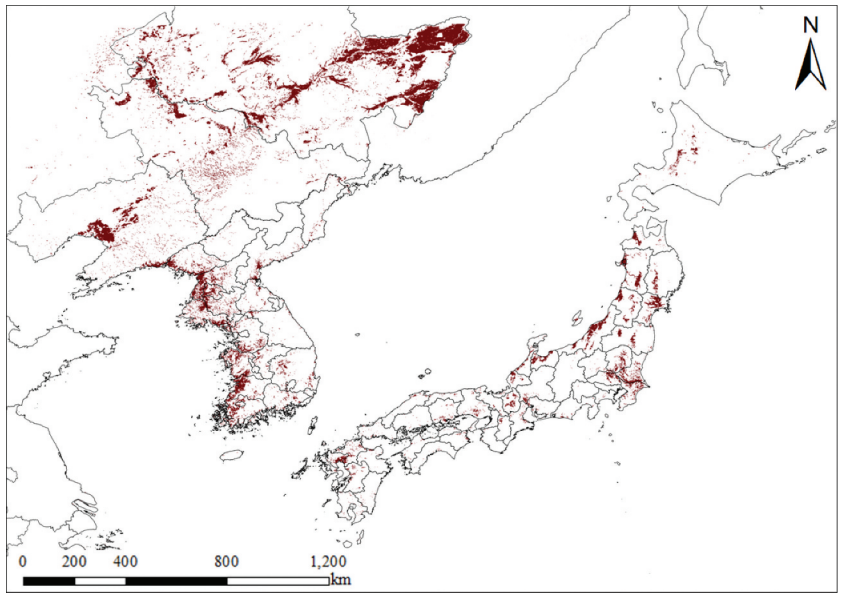

Figure 5. Spatial distribution of classified paddy fields based on geostationary ocean color imager (GOCI) and moderate resolution imaging spectroradiometer (MODIS) composited from 2011 to 2017 in northeast asia.

each of the classified paddy maps for the entire year for the period from 2011 to 2017 is presented in Figure 1 A.

Projection of crop productivity in northeast asia

In this study, paddy rice yield was geospatially simulated and analyzed in most of the Northeast Asian counties and provinces in China, North Korea, South Korea, and Japan from 2011 to 2017. We could successfully reproduce the geographical variation in rice yield according to the grid with a pixel resolution of 500 m (Figure 6 and A2) and the respective administrative county or province (Figure 7 and A3). In the case of Figures 6 and 7, each of the representative maps of paddy rice yields shows the simulated mean rice yield and standard deviations from 2011 to 2017. Each of the representative maps presented distinctive spatiotemporal rice yield trends in each of the countries, which could be due to differences in climate zones, rice cultivars, and agricultural infrastructure among the countries. In the case of North Korea, rice yield is relatively lower than in the other neighboring countries, which could be attributed to poor agricultural infrastructure (Yeom et al. 2018; Ryu et al. 2019). All rice yield maps on the grid-scale and the country/province scale from 2011 to 2017 are represented in Appendix (Figure A2 and A3) for more detailed information. These maps represented the temporal and spatial variations in rice yield in the Northeast Asian countries well; the spatiotemporal characteristics of rice yield differed across countries and were well illustrated in annual maps. Based on the spatiotemporal annual maps of paddy rice yield across the Northeastern Asia, we could perform further rice productivity analyses by considering various factors such as cultivars planted, environmental conditions, and agricultural infrastructure in each of the countries.

In the statistical error analysis for the calibration sites, the RMSE, r^2, and NSE were 0.551 t ha^{-1}, 0.903, 0.692, and 0.434, respectively (Figure 8a). For the validation sites, the RMSE, r^2, and NSE were 0.674 t ha^{-1}, 0.823, 0.524, and 0.235, respectively (Figure 8b). There were no significant differences between simulated and measured rice yields, where *p*-values were 0.434 and 0.235 for the calibration and validation sites according to a sample *t*-test ($\alpha = 0.05$). This indicates that the simulated mean rice yields for each county from 2011 to 2017 were in agreement with the observed rice yields within an acceptable statistical range in both calibration and validation sites. In a more detailed statistical analysis for the entire

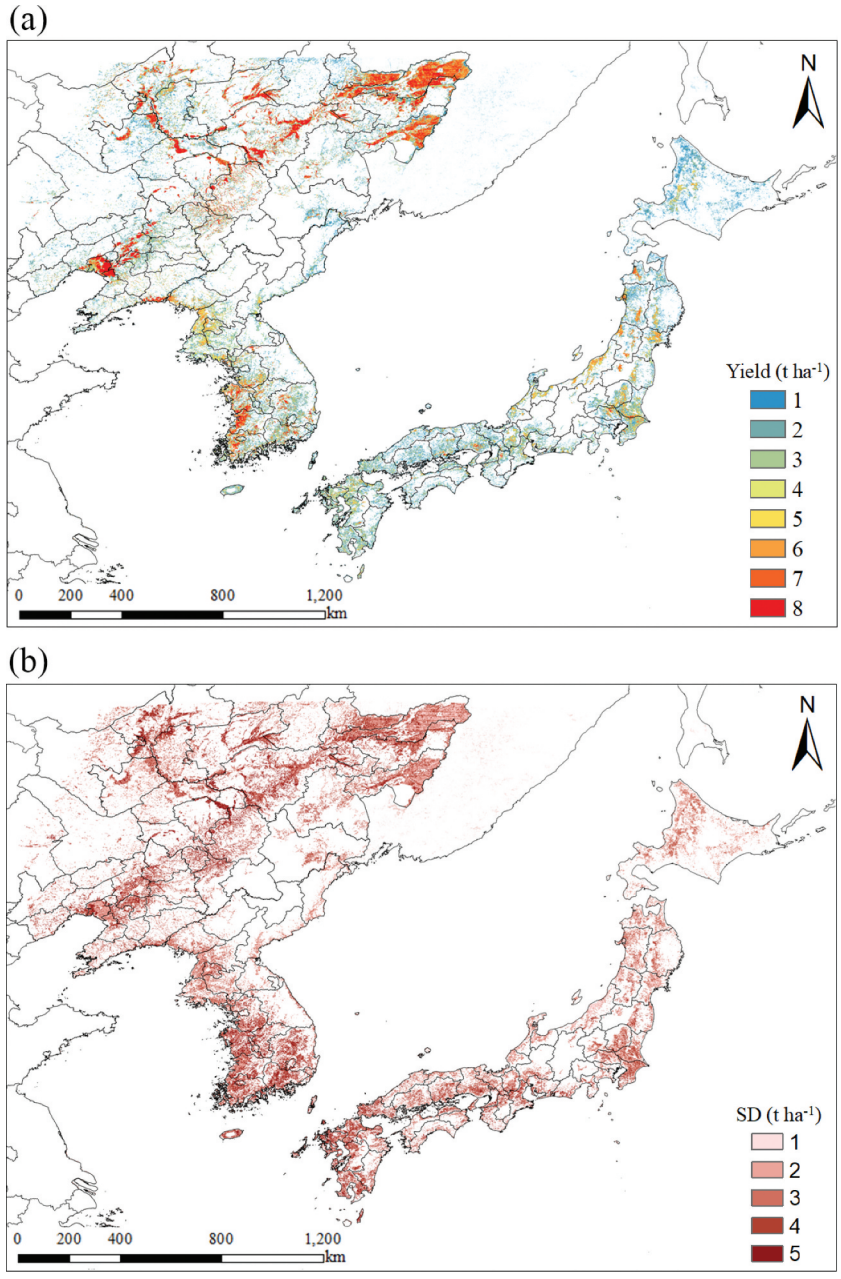

Figure 6. Spatial distribution of simulated (a) mean rice yield and (b) standard deviation (sd) in northeast asia from 2011 to 2017 based on communication, ocean, and meteorology satellite (COMS) images integrated into the RSCM for rice.

dataset, in the calibration, the RMSE, r^2, NSE, and *p* values ranged from 0.367 to 0.680 t ha^{-1}, 0.661 to 0.834, 0.581 to 0.802, and 0.571 to 0.951 from 2011 to 2017, respectively (Figure A4a and Table A2). Similarly, in the validation, the RMSE, r^2, NSE, and *p* values ranged from 0.620 to 0.676 t ha^{-1}, 0.462 to 0.782, 0.108 to 0.660, and 0.130 to 0.953, respectively (Figure A4b and Table A3). While there were small disagreements attributable to subpixel heterogeneity, overall, the simulated rice yields corresponded with the observed values (Figure 8).

Discussion

In most previous studies, rice cultivation area and yield at the national scale have been estimated using one type of satellite, especially polar orbit

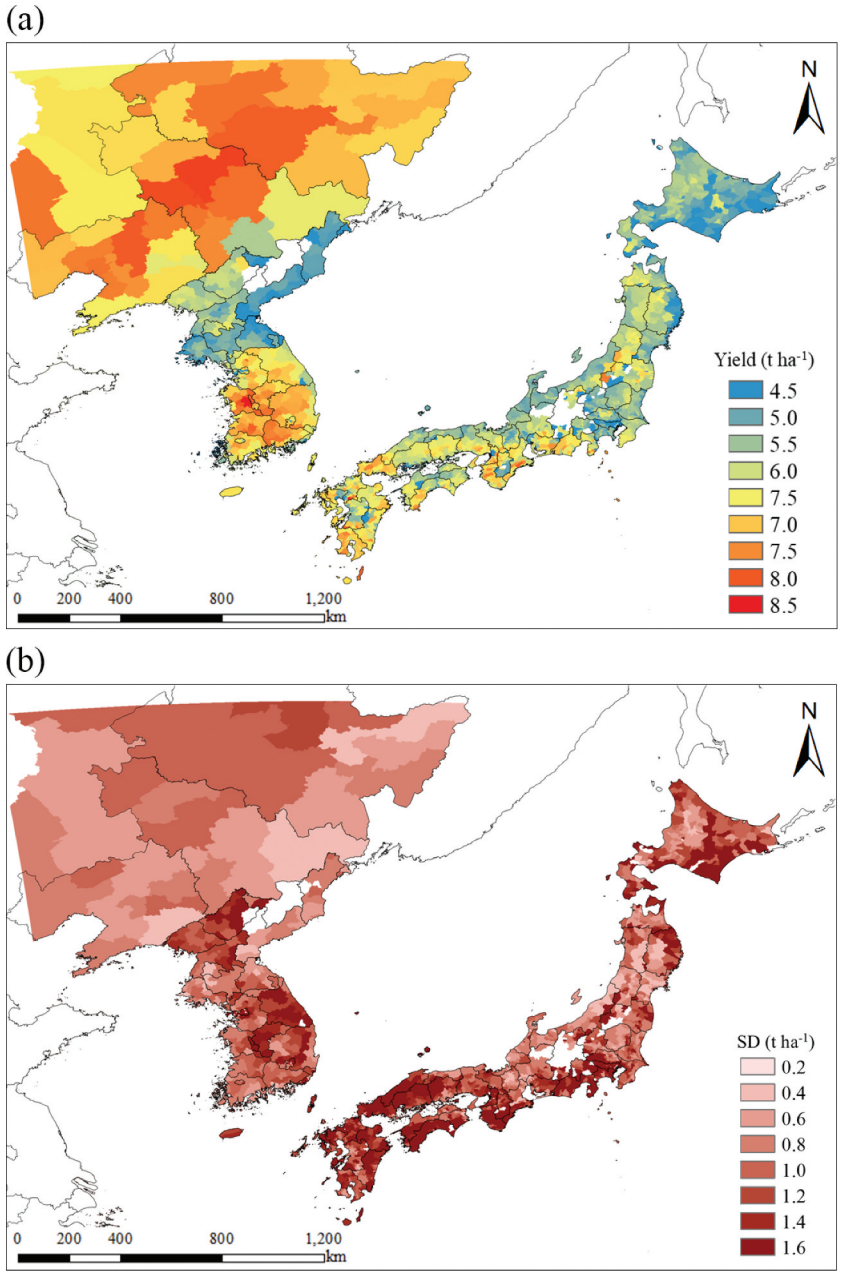

Figure 7. County- (or province-) wide spatial distribution of (a) simulated rice yields and (b) standard deviation (SD) in northeast asia from 2011 to 2017 based on communication, ocean, and meteorology satellite (COMS) images integrated into the RSCM for rice.

satellite (mostly MODIS) (Kontgis, Schneider, and Ozdogan 2015; Zhang and Zhang; Zhang et al. 2017). Although the previous studies have reported reliable results, they highlighted the systemic limitations of polar orbit satellites in terms of data acquisition. In particular, there are challenges with regard to the observation of time-dense vegetation indices of rice-growing periods, especially under the rainy monsoon season (Xiao et al. 2006; Yeom and Kim 2015). This study estimated rice area and yield in Northeast Asia after addressing the aforementioned problems, based on a newly adopted high temporal resolution GOCI geostationary satellite. With regard to the classification of rice area, discontinuous vegetation index profiles attributed to cloudiness were resolved by capturing distinctive spectral signals of rice growth patterns from the GOCI satellite. Furthermore, rice yield was estimated based on GOCI by stably

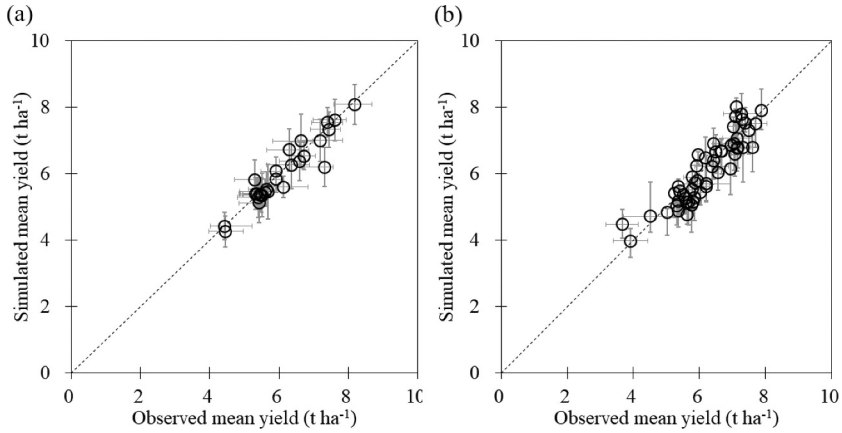

Figure 8. Comparisons of observed and simulated rice yields for (a) 27 calibration and (b) 54 validation counties or provinces (refer to Fig. 1) in northeast asia from 2011 to 2017. The dotted lines are 1:1 lines. RMSE, NSE, and *p* represent root mean square errors, nash-sutcliffe efficiencies, and p values of two-sample *t* tests (α = 0.05). refer to Table 5 for the detailed statistical analysis results of the (a) calibration and (b) validation, respectively.

observing the vegetation indices in the growing seasons, which is essential for the determination of rice final yield (Kastens et al. 2005; Mkhabela et al. 2011; Wu et al. 2011; Zhang et al. 2017).

The present study showed that the synthetic application of polar orbit and geostationary satellites could complement their limitations to facilitate the acquisition of more comprehensive information on crops in large-scale applications.

Additional considerations in rice area classification

The classification approach used in this study has been widely applied to detect the distribution of paddy fields and irrigated and flooded areas within optical satellite images (Sakamoto et al. 2007; Peng et al. 2011; Dong et al. 2016; Clauss, Yan and Kuenzer 2016) since it was first suggested by Xiao et al. (2002). The approach applies a simple concept and clear logic that reflect the spectral time series characteristics of paddy fields. Studies have shown that this approach produces reliable results, making it applicable to most remote-sensing data equipped with optical spectral bands, including the SWIR band. It can also simultaneously detect transplanting dates, a key indicator of the onset of rice growth, and is used as a key input to RSCM and other crop models. The advantages of the classification method used in this study, along with its potential errors and pitfalls, are discussed below.

The most important threshold in this classification approach is the *T* value of the LSWI, which serves to detect paddy fields among pixels mixed with other vegetation by increasing the sensitivity to irrigated water. The *T* value can vary depending not only on the land cover features of the study site, but also on the spectral responses of the satellite sensors. It was, therefore, important to define the appropriate threshold value for the approach used in this study. We believe that the *T* values determined in this study (between 0.05 and 0.20) were appropriate (Xiao et al. 2006; Sun et al. 2009; Teluguntla et al. 2015), given previous results. However, we did not consider the fact that in double cropping regions, the *T* values for double rice crops or single crops after a winter crop can differ significantly in late-season rice fields (Peng et al. 2011). Therefore, a more accurate classification would have been possible if the double-cropped regions had been distinguished from single-cropped regions, and a different threshold for single-rice crop regions had been applied.

In addition, it was necessary to apply an additional algorithm to remove land covers with temporal NDVI and LSWI patterns similar to those of paddy fields, such as wetlands, snowmelt cover during the spring season, or land following heavy rain. Several conditions (or thresholds) were therefore applied, namely: maximum and minimum NDVIs, periods with specific NDVI values, and a normalized difference snow index to remove snowmelt cover. In some studies, temporal windows to define the period available for cultivation based on the MODIS land surface temperature (LST) have been used (Dong et al. 2016; Zhang et al. 2015). In this study, air temperature instead of LST and the rate of increase in the NDVI were used to remove

misleading pixels. Because of tillage and irrigated water, initial NDVI values are lower in paddy fields than in other types of fields but increase rapidly to a maximum value within 2 months of transplanting. Wetlands and other water-related land covers show lower rates of NDVI increases because of their low maximum NDVIs. The detection error caused by heavy rainfall was resolved using the condition that irrigated fields had to remain in the data for more than one week, which was enabled by our use of GOCI data to provide a stable daily NDVI.

Finally, the use of satellite imagery with a coarse spatial resolution, such as GOCI or MODIS, results in errors caused by mixed-pixel effects, including, most significantly, a tendency toward underestimation in small paddy patches and overestimation in large paddy patches, particularly in areas with a large heterogeneity in land cover (Jeong et al. 2012; Peng et al. 2011; Sakamoto et al. 2007). This tendency is even more noticeable when making an assessment based on a comparison of paddy areas, because small patches are likely to be undetectable, whereas for large patches, each detected pixel is likely composed solely of paddy fields (Jeong, Ko, and Yeom 2018; Xiao et al. 2006; Ozdogan and Gutman 2008). Despite these errors, using coarse resolution imagery with high temporal resolution is necessary, not only for continuous and consecutive land cover classification, but also for the monitoring of critical crop growth information over broad areas, from national to global scales.

Spatial extensibility of the RSCM model and its limitations

Before applying remotely sensed satellite images to paddy productivity simulations in North Korea, we calibrated RSCM-rice using an experimental field site in Cheorwon, South Korea, and then validated the approach using geospatial data from two administrative districts bordering North Korea (data not shown), similar to the approach in the earlier study by Yeom et al. (2018). The spatial distribution of the simulated rice yields in Northeast Asia from 2011 to 2017 (Figures 6, 7) show a consistent classification of the actual rice-growing areas and trends in the four countries, both spatially and temporally. In North Korea, in particular, the inter-annual variation in rice yields was found to be higher than in the other countries because of the country's poor irrigation infrastructure. Our results were generally comparable to those of earlier studies conducted using the GRAMI model (Kim et al. 2017; Jeong, Ko, and Yeom 2018; Yeom et al. 2018). These studies demonstrated that the satellite image-integrated rice model could reproduce the regional and nation-wide rice yield and productivity. However, our study widened the application of RSCM to cover several countries, while the scope of previous studies was more limited and focused on the farm field or national scale. The accuracy of our results and those of previous studies suggest that integration of various types of remote-sensing data, including satellite imagery, into the RSCM system, can be useful for monitoring crop productivity and support decision making in crop management in a variety of fields and regions. We demonstrated earlier that the current hybrid modeling approach could be efficiently applied to retrieve and deliver information on rice area and yield for a data-sparse or inaccessible region (Yeom et al. 2018). The methodology can also be employed to monitor rice yield and production of a nation of interest to secure the current and future food crop production, as an essential agricultural resource (Jeong, Ko, and Yeom 2018).

The input parameters involved in RSCM represent a possible source of error in simulation. The parameters used to define the relationships between the LAI and specific VIs are critical to the RSCM system and can substantially affect its simulation results. According to Doraiswamy et al. (2005), the relationship between the LAI and NDVI used in the calibration procedure might vary by research application, as a result of changes in soil features, plant population densities, and other environmental aspects. If the relationship between the LAI and VI falls outside the optimal range of the preset parameters, the resulting error could be severe (Ko et al. 2015). To address this issue, we applied log-log-linear regression models, with correlations between the LAI and either canopy reflectance values or the VIs (e.g., MTVI, NDVI, RDVI, and

OSAVI) framed collectively. As discussed above, we assumed that the relationship between the LAI and reflectance or VI could be described using a log-log regression design with a slope of approximately two-thirds. A Bayesian method for parameter estimation was formulated in the RSCM regime to enable an agreement between simulation and observation according to the Quasi-Newton optimization method (Nash 1990) for pixel-based geospatial simulation cases. Our optimization approach was framed to integrate various remote-sensing data into RSCM, which allowed us to rely heavily on the LAI inputs estimated from remotely sensed information. This approach of assimilating remote-sensing data into RSCM can have several benefits. First, it allows the use of a modest input state, in which existing observations are introduced as a crucial driver for the characterization of environmental conditions. Second, it allows the RSCM system to advance simulation performance. Third, it allows RSCM to assimilate remote-sensing information obtained from a variety of platforms, including unmanned aerial systems (Jeong, Ko, Choi, Xue, and Yeom 2018) and operational optical satellite-based sensors, with varying spatial resolutions (Kim et al. 2017; Yeom et al. 2018). Finally, it renders RSCM applicable to any area of interest on the Earth's surface, assuming satellite imagery can be obtained. Yeom et al. (2018) demonstrated this aspect, reporting rice productivity and geographical variability in inaccessible North Korea.

At the same time, the optimization approach of RSCM has several limitations, such as including imperfect representations of remote-sensing information, along with the limited observations obtainable during the crop season. These restrictions can ultimately result in a degree of dissimilarity between simulations and observations, reproducing an inaccurate crop productivity. Nevertheless, the simple input parameters and variable requirements of RSCM have some important advantages, predominantly concerning unapproachable and data-sparse regions. Although the RSCM approach has some limitations in such regions, including a constant dependence on remote-sensing data, satellite observation-based crop modeling is generally more important; it would be nearly impossible to reproduce and monitor crop productivity without the use of satellite images.

Conclusions

In the present study, we successfully classified paddy areas and reproduced the national rice yields in North and South Korea, Japan, and the northeast region of China using an RSCM for rice in combination with geostationary satellite-based GOCI images. The overall classification accuracy of paddy field area was 78.8% based on synthetic applications of dense-time GOCI vegetation and MODIS water indices, indicating dependable representations of the spatial distributions of the paddy rice areas in each country. In the case of rice yield, there were no significant differences (p-value = 0.235) between simulated and measured rice yields according to a sample t-test ($\alpha = 0.05$) for the entire period (from 2011 to 2017), showing an agreement with RMSEs of 0.674 t ha^{-1}, r^2 of 0.823, and NSEs of 0.524. The overall accuracy metrics, comparing simulated and observed variables, were statistically acceptable, given the continental scale of our study, and in spite of the subpixel heterogeneity of the satellite images arising from their rough spatial resolution. The current study results are, therefore, potentially useful in the development of methods to provide a more accurate real-time understanding of regional crop cultivation areas and productivity, using the integration of GOCI images into crop models. Such information could be invaluable in helping to manage staple food production in the countries of Northeast Asia. We believe that our assimilation of GOCI imagery into the RSCM system could help provide diagnostic information for the monitoring of crop conditions and contribute to an agricultural decision delivery system for simulating crop growth and monitoring crop yield and production for various terrestrial divisions from a region to a nation. Further development would add value to the current integration scheme and its capability to simulate other major staple crops, and it would incorporate the fundamental ecophysiology of crop productivity (e.g., carbon assimilation and evapotranspiration).

Acknowledgements

This research was supported by the Basic Science Research Program through the National Research Foundation of Korea (NRF), funded by the Ministry of Education, Science, and Technology (NRF-2018R1D1A1B07042925).

Data availability (Updated upon acceptance)

The data that support the findings of this study are available in [repository name] at [URL/DOI], reference number [reference number]. These data were derived from the following resources available in the public domain: [list resources and URLs].

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This research was supported by the Basic Science Research Program through the National Research Foundation of Korea (NRF), funded by the Ministry of Education, Science, and Technology [NRF-2018R1D1A1B07042925].

Highlights

- Northeast Asian paddy field areas were classified based on Sun- and Geo-synchronous satellite images.
- Dense vegetation index profiles of a geostationary satellite were integrated into a crop model.
- Spatiotemporal maps of rice yields in Northeast Asia were simulated with reasonable accuracy.
- The method can be applied to monitor rice area and yield in large-scale applications.

References

Ahuja, L. R., K. W. Rojas, J. D. Hanson, M. J. Shaffer, and L. Ma. 2000. *Root Zone Water Quality Model: Modeling Management Effects on Water Quality and Crop Production*. Highland Ranch, CO, USA: Water Resources Publications, LLC.

Barrett, C. B. "Measuring Food Insecurity." *Science*,327 (5967) (2010).:825–828. doi: 10.1126/science.1182768.

Bregaglio, S., L. Hossard, G. Cappelli, R. Resmond, S. Bocchi, J.-M. Barbier, F. Ruget, and S. Delmotte. 2017. "Identifying Trends and Associated Uncertainties in Potential Rice Production under Climate Change in Mediterranean Areas." *Agricultural and Forest Meteorology*,237-238:219–232. doi: 10.1016/j.agrformet.2017.02.015.

Cohen, J. 1968. "Weighted Kappa: Nominal Scale Agreement Provision for Scaled Disagreement or Partial Credit". *Psychological Bulletin*, 70(4): 213.doi: 10.1037/h0026256.

Dingkuhn, M., and K. M. Miezan. 1995. "Climatic Determinants of Irrigated Rice Performance in the Sahel — II. Validation of Photothermal Constants and Characterization of Genotypes". *Agricultural Systems*, 48 (4):411–433. doi: 10.1016/0308-521X(94)00028-J.

Dong, J., X. Xiao, M. A. M. G. Zhang, Y. Qin, D. Thau, C. Biradar, and B. Moore. 2016. "Mapping Paddy Rice Planting Area in Northeastern Asia with Landsat 8 Images, Phenology-based Algorithm and Google Earth Engine". *Remote Sensing of Environment*,185:142–154. doi: 10.1016/j.rse.2016.02.016.

Doraiswamy, P. C., and P. W. Cook. 1995. "Spring Wheat Yield Assessment Using NOAA AVHRR Data". *Canadian Journal of Remote Sensing*, 21 (1):43–51. doi: 10.1080/07038992.1995.10874595.

Doraiswamy, P. C., T. R. Sinclair, S. Hollinger, B. Akhmedov, A. Stern, and J. Prueger. 2005. "Application of MODIS Derived Parameters for Regional Crop Yield Assessment" *Remote Sensing of Environment*, 97 (2):192–202.doi: 10.1016/j.rse.2005.03.015.

Grell, G. A., J. Dudhia, and D. Stauffer. 1994. "A Description of the Fifth Generation Pen State/NCAT Mesoscale Model (MM5)." In *NCAR Technical Note NCAR/TN-398+STR*

Haboudane, D., J. R. Miller, E. Pattey, P. J. Zarco-Tejada, and I. B. Strachan. 2004. "Hyperspectral Vegetation Indices and Novel Algorithms for Predicting Green LAI of Crop Canopies: Modeling and Validation in the Context of Precision Agriculture". *Remote Sensing of Environment*,90 (3):337–352. doi: 10.1016/j.rse.2003.12.013.

Hodson, D., and J. White. 2010. "GIS and Crop Simulation Modelling Applications in Climate Change Research." In *Climate Change and Crop Production*, edited by M. P. Reynolds, 245–262. Cambridge, MA CABI.

Huang, J., F. Sedano, Y. Huang, H. Ma, X. Li, S. Liang, L. Tian, X. Zhang, J. Fan, and W. Wu. 2016. "Assimilating a Synthetic Kalman Filter Leaf Area Index Series into the WOFOST Model to Improve Regional Winter Wheat Yield Estimation". *Agricultural and Forest Meteorology*,216:188–202. doi: 10.1016/j.agrformet.2015.10.013.

Inoue, S., A. Ito, and C. Yonezawa. 2020. "Mapping Paddy Fields in Japan by Using a Sentinel-1 SAR Time Series Supplemented by Sentinel-2 Images on Google Earth Engine". *Remote Sensing*, 12(10), 1622.doi: 10.3390/rs12101622.

IPCC. 2013. "Climate Change 2013: The Physical Science Basis." In*Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, edited by T. F. Stocker, D. Qin, G.-K. Plattner, M. Tignor, P. M. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex, and P. M. Midgley. Cambridge, United Kingdom and New York, NY, USA: Cambridge University Press

Jeong, S., S. Kang, K. Jang, H. Lee, S. Hong, and D. Ko. 2012. "Development of Variable Threshold Models for Detection of Irrigated Paddy Rice Fields and Irrigation Timing in Heterogeneous Land Cover". *Agricultural Water Management*,115:83–91.doi: 10.1016/j.agwat.2012.08.012.

Jeong, S., J. Ko, J. Choi, W. Xue, and J. M. Yeom. 2018. "Application of an Unmanned Aerial System for Monitoring

Paddy Productivity Using the GRAMI-rice Model." *International Journal of Remote Sensing*, 39 (8):2441–2462. 10.1080/01431161.2018.1425567

Jeong, S., J. Ko, and J. M. Yeom. 2018. "Nationwide Projection of Rice Yield Using A Crop Model Integrated with Geostationary Satellite Imagery: A Case Study in South Korea" *Remote Sensing*,10 (10):1665. 10.3390/rs10101665

Kastens, J. H., T. L. Kastens, D. L. A. Kastens, K. P. Price, E. A. Martinko, and R.-Y. Lee. 2005. "Image Masking for Crop Yield Forecasting Using AVHRR NDVI Time Series Imagery". *Remote Sensing of Environment*, 99 (3):341–356.doi: 10.1016/j.rse.2005.09.010.

Kawamura, H., S. Tanahashi, and T. Takahashi. 1998. "Estimation of Insolation over the Pacific Ocean off the Sanriku Coast." *Journal of Oceanography*, 54 (5):457–464. doi: 10.1007/bf02742448

Kim, H. Y., J. Ko, S. Kang, and J. Tenhunen. 2013. "Impacts of Climate Change on Paddy Rice Yield in a Temperate Climate" *Global Change Biology*, 19 (2):548–562. doi: 10.1111/gcb.12047.

Kim, M., J. Ko, S. Jeong, J. M. Yeom, and H. O. Kim. 2017. "Monitoring Canopy Growth and Grain Yield of Paddy Rice in South Korea by Using the GRAMI Model and High Spatial Resolution Imagery". *GIScience & Remote Sensing*,54 (4):534–551. 10.1080/15481603.2017.1291783

Kim, S. I., D. S. Ahn, K. S. Han, and J. M. Yeom. 2016. "Improved Vegetation Profiles with GOCI Imagery Using Optimized BRDF Composite."*Journal of Sensors* 2016: 7. doi:10.1155/2016/7165326.

Ko, J., S. J. Maas, R. J. Lascano, and D. Wanjura. 2005. "Modification of the GRAMI Model for Cotton." *Agronomy Journal*, 97:6. 5 10.2134/agronj2004.0267

Ko, J., S. Jeong, J. Yeom, H. Kim, J.-O. Ban, and H.-Y. Kim. 2015. "Simulation and Mapping of Rice Growth and Yield Based on Remote Sensing." *Journal of Applied Remote Sensing*, 9 (1):096067. 10.1117/1.JRS.9.096067

Kontgis, C., A. Schneider, and M. Ozdogan. 2015. "Mapping Rice Paddy Extent and Intensification in the Vietnamese Mekong River Delta with Dense Time Stacks of Landsat Data" *Remote Sensing of Environment*,169:255–269. 10.1016/j.rse.2015.08.004

Kusunoki, S., and R. Mizuta. 2013. "Changes in Precipitation Intensity over East Asia during the 20th and 21st Centuries Simulated by a Global Atmospheric Model with a 60 km Grid Size." *Journal of Geophysical Research: Atmospheres*, 118 (19):11,007–11,16. doi:10.1002/jgrd.50877.

Lai, Y. R., M. J. Pringle, P. M. Kopittke, N. W. Menzies, T. G. Orton, and Y. P. Dang. 2018. "An Empirical Model for Prediction of Wheat Yield, Using Time-integrated Landsat NDVI." *International Journal of Applied Earth Observation and Geoinformation*, 72:99–108.doi: 10.1016/j.jag.2018.07.013.

Launay, M., and M. Guerif. 2005. "Assimilating Remote Sensing Data into a Crop Model to Improve Predictive Performance for Spatial Application" *Agriculture, Ecosystems & Environment*,111 (1–4):321–339. doi: 10.1016/j.agee.2005.06.005.

Li, Q., H. Zhang, X. Du, N. Wen, and Q. Tao 2014. "County-level Rice Area Estimation in Southern China Using Remote Sensing Data." *Journal of Applied Remote Sensing* 8: 083657.

Lobell, D. B., D. Thau, C. Seifert, E. Engle, and B. Little. 2015. "A Scalable Satellite-based Crop Yield Mapper." *Remote Sensing of Environment*, 164:324–333.doi: 10.1016/j.rse.2015.04.021.

Lucht, W., C. B. Schaaf, and A. H. Strahler. 2000. "An Algorithm for the Retrieval of Albedo from Space Using Semiempirical BRDF Models" *IEEE Transactions on Geoscience and Remote Sensing*, 38 (2):977–998. doi: 10.1109/36.841980.

Maas, S. J. 1992. "GRAMI: A Crop Model Growth Model that Can Use Remotely Sensed Information." *In ARS-91*, 78. Washington, DC, USA: USDA-ARS.

Maas, S. J. 1993a. "Within-season Calibration of Modeled Wheat Growth Using Remote Sensing and Field Sampling." *Agron. J*, 85:669–672. 3 10.2134/agronj1993.00021962008500030028x

Maas, S. J. 1993b. "Parameterized Model of Gramineous Crop Growth: II. Within-season Simulation Calibration." *Agron. J*, 85:354–358.

MacLean, J. L., D. C. Dawe, B. Hardy, and G. P. Hettel. 2013. "*Rice Almanac: Sourcebook for the Most Important Economic Activity on Earth* Fourth Edition Ed". *International Rice Research Institute, Metro Manila, Philippines Global Rice Science Partnership*, 283. Metro Manila, Philippines: International Rice Research Institute.

McCloy, K. R., F. R. Smith, and M. R. Robinson. 1987. "Monitoring Rice Areas Using LANDSAT MSS Data." *International Journal of Remote Sensing*, 8:741–749. 5 10.1080/01431168708948685

McMahon, T. 1973. "Size and Shape in Biology". *Science*, 179 (4079):1201–1204. 10.1126/science.179.4079.1201

Mkhabela, M. S., P. Bullock, S. Raj, S. Wang, and Y. Yang. 2011. "Crop Yield Forecasting on the Canadian Prairies Using MODIS NDVI Data". *Agricultural and Forest Meteorology*, 151 (3):385–393. doi: 10.1016/j.agrformet.2010.11.012.

Mosleh, M., Q. Hassan, and E. Chowdhury. 2015. "Application of Remote Sensors in Mapping Rice Area and Forecasting Its Production: A Review". *Sensors*,15 (1):769. 10.3390/s150100769

Nash, J. C. 1990. *Compact Numerical Methods for Computers: Linear Algebra and Function Minimisation*. Adam Hilger, New York, USA: CRC press.

Nash, J. E., and J. V. Sutcliffe. 1970. "River Flow Forecasting through Conceptual Models Part I — A Discussion of Principles". *Journal of Hydrology*,10 (3):282–290. doi: 10.1016/0022-1694(70)90255-6.

Nguyen, T. T. H., C. A. J. M. De Bie, A. Ali, E. M. A. Smaling, and T. H. Chu. 2012. "Mapping the Irrigated Rice Cropping Patterns of the Mekong Delta, Vietnam, through Hyper-temporal SPOT NDVI Image Analysis". *International Journal of Remote Sensing*,33:415–434. 2 10.1080/01431161.2010.532826

Nguyen, V. C., S. Jeong, J. H. Ko, C. T. Ng, and J. M. Yeom. 2019. "Mathematical Integration of Remotely-sensed Information into a Crop Modelling Process for Mapping Crop Productivity" *Remote Sensing*,11 (2131). 18 10.3390/rs11182131

Ozdogan, M., and G. Gutman. 2008. "A New Methodology to Map Irrigated Areas Using Multi-temporal MODIS and Ancillary Data: An Application Example in the Continental US". *Remote Sensing of Environment*,112 (9):3520–3537. doi: 10.1016/j.rse.2008.04.010.

Park, S. J. 2014. "Generality and Specificity of Landforms of the Korean Peninsula, and Its Sustainability". *Journal of the Korean Geographical Society*, 49(5): 656–674.

Peel, M. C., B. L. Finlayson, and M. Thomas 2007. "Updated World Map of the Köppen-Geiger Climate Classification". *Hydrology and Earth System Sciences*, 11:1633–1644. 5 10.5194/hess-11-1633-2007

Peng, D., J. Huang, C. Li, L. Liu, W. Huang, F. Wang, and X. Yang. 2014. "Modelling Paddy Rice Yield Using MODIS Data." *Agricultural and Forest Meteorology*, 184:107–116.doi: 10.1016/j.agrformet.2013.09.006.

Peng, D., A. R. Huete, J. Huang, F. Wang, and H. Sun. 2011. "Detection and Estimation of Mixed Paddy Rice Cropping Patterns with MODIS Data". *International Journal of Applied Earth Observation and Geoinformation*,13 (1):13–23. doi: 10.1016/j.jag.2010.06.001.

Press, W. H., S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery. 1992. *Numerical Recipes: The Art of Scientific Computing*. New York: Cambridge University Press.

Rabus, B., M. Eineder, A. Roth, and R. Bamler. 2003. "The Shuttle Radar Topography Mission—a New Class of Digital Elevation Models Acquired by Spaceborne Radar". *ISPRS Journal of Photogrammetry and Remote Sensing*,57 (4):241–262. doi: 10.1016/S0924-2716(02)00124-7.

Reeves, M. C., M. Zhao, and S. W. Running. 2005. "Usefulness and Limits of MODIS GPP for Estimating Wheat Yield." *International Journal of Remote Sensing*,26 (7):1403–1421. doi: 10.1080/01431160512331326567.

Rondeaux, G., M. Steven, and F. Baret. 1996. "Optimization of Soil-adjusted Vegetation Indices". *Remote Sensing of Environment*, 55 (2):95–107.doi: 10.1016/0034-4257(95)00186-7.

Ross, J. K., and A. L. Marshak. 1988. "Calculation of Canopy Bidirectional Reflectance Using the Monte Carlo Method". *Remote Sensing of Environment*, 24 (2):213–225. doi: 10.1016/0034-4257(88)90026-0.

Roujean, J. L., M. Leroy, and P. Y. Deschamps. 1992. "A Bidirectional Reflectance Model of the Earth's Surface for the Correction of Remote Sensing Data". *Journal of Geophysical Research: Atmospheres*, 97 (D18):20455–20468. doi: 10.1029/92JD01411.

Rouse, J. W., Jr., R. H. Haas, J. A. Schell, and D. W. Deering. 1974. "Monitoring Vegetation Systems in the Great Plains with ERTS." *In NASA. Goddard Space Flight Center 3d ERTS-1 Symp*., 309–317. Washington, D.C., USA: NASA.

Ryu, J. H., K. S. Han, Y. W. Lee, N. W. Park, S. Hong, C. Y. Chung, and J. Cho. 2019. "Different Agricultural Responses to Extreme Drought Events in Neighboring Counties of South and North Korea." *Remote Sensing*,11:1773. 15 10.3390/rs11151773

Sakamoto, T., N. V. Nguyen, A. Kotera, H. Ohno, N. Ishitsuka, and M. Yokozawa. 2007. "Detecting Temporal Changes in the Extent of Annual Flooding within the Cambodia and the Vietnamese Mekong Delta from MODIS Time-series Imagery". *Remote Sensing of Environment*,109 (3):295–313. doi: 10.1016/j.rse.2007.01.011.

Schaaf, C. B., F. Gao, A. H. Strahler, W. Lucht, X. Li, T. Tsang, and N. C. Strugnell. 2002. "First Operational BRDF, Albedo Nadir Reflectance Products from MODIS". *Remote Sensing of Environment*,83 (1–2):135–148. doi: 10.1016/S0034-4257(02)00091-3.

Schaaf, C. B., J. Liu, F. Gao, and A. H. Strahler. 2011. "MODIS Albedo and Reflectance Anisotropy Products from Aqua and Terra."*In Land Remote Sensing and Global Environmental Change: NASA's Earth Observing System and the Science of ASTER and MODIS, Remote Sensing and Digital Image Processing Series*, edited by B. Ramachandran, C. Justice, and M. Abrams, 873. Berlin, Germany: Springer-Verlag.

Spitters, C. J. T., H. Van Keulen, and D. W. G. Van Kraalingen. 1989. "A Simple and Universal Crop Growth Simulator: SUCROS87." *In Simulation and Systems Management in Crop Protection*, 147–181. Wageningen, The Netherlands: Pudoc.

Sun, H., J. Huang, A. R. Huete, D. Peng, and F. Zhang. 2009. "Mapping Paddy Rice with Multi-date Moderate-resolution Imaging Spectroradiometer (MODIS) Data in China". *Journal of Zhejiang University-SCIENCE A*,10 (10):1509–1522. doi: 10.1631/jzus.A0820536.

Teluguntla, P., D. Ryu, B. George, J. P. Walker, and H. M. Malano. 2015. "Mapping Flooded Rice Paddies Using Time Series of MODIS Imagery in the Krishna River Basin, India". *Remote Sensing*,7 (7):8858–8882. 10.3390/rs70708858

Tilman, D., C. Balzer, J. Hill, and B. L. Befort. 2011. "Global Food Demand and the Sustainable Intensification of Agriculture." *Proceedings of the National Academy of Sciences*,108 (50):20260–20264. doi: 10.1073/pnas.1116437108.

van Diepen, C. A., J. Wolf, H. van Keulen, and C. Rappoldt. 1989. "WOFOST: A Simulation Model of Crop Production". *Soil Use and Management*,5 (1):16–24. doi: 10.1111/j.1475-2743.1989.tb00755.x.

Wu, F., C. Wang, H. Zhang, B. Zhang, and Y. Tang. 2011. "Rice Crop Monitoring in South China with RADARSAT-2 Quad-polarization SAR Data." *IEEE Geoscience and Remote Sensing Letters*,8 (2):196–200. doi: 10.1109/LGRS.2010.2055830.

Xiao, X., S. Boles, S. Frolking, W. A. Salas, B. Moore, C. Li, L. He, and R. Zhao. 2002. "Observation of Flooding and Rice Transplanting of Paddy Rice Fields at the Site to Landscape Scales in China Using VEGETATION Sensor Data". *International Journal of Remote Sensing*,23 (15):3009–3022. doi: 10.1080/01431160110107734.

Xiao, X., S. Boles, J. Liu, D. Zhuang, S. Frolking, C. Li, W. Salas, and B. Moore. 2005. "Mapping Paddy Rice Agriculture in Southern China Using Multi-temporal MODIS Images". *Remote Sensing of Environment*,95 (4):480–492. doi: 10.1016/j.rse.2004.12.009.

Xiao, X., S. B., . S. Frolking, C. Li, J. Y. Babu, W. Salas, and B. Moore. 2006. "Mapping Paddy Rice Agriculture in South and Southeast Asia Using Multi-temporal MODIS Images." *Remote Sensing of Environment*,100 (1):95–113. doi: 10.1016/j.rse.2005.10.004.

Yan, C. K. H., and C. Kuenzer. 2016. "Mapping Paddy Rice in China in 2002, 2005, 2010 and 2014 with MODIS Time Series". *Remote Sensing*,8 (5):434. 10.3390/rs8050434

Yeom, J. M., K. S. Han, and J. J. Kim. 2012. "Evaluation on Penetration Rate of Cloud for Incoming Solar Radiation Using Geostationary Satellite Data." *Asia-Pacific Journal of Atmospheric Sciences*, 48 (2):115–123. doi: 10.1007/s13143-012-0011-9.

Yeom, J. M., S. Jeong, G. Jeong, C. TimNg, R. C. Deo, and J. Ko. 2018. "Monitoring Paddy Productivity in North Korea Employing Geostationary Satellite Images Integrated with GRAMI-rice Model". *Scientific Reports*, 8 (1):16121. 10.1038/s41598-018-34550-0

Yeom, J. M., and H. O. Kim. 2015. "Comparison of NDVIs from GOCI and MODIS Data Towards Improved Assessment of Crop Temporal Dynamics in the Case of Paddy Rice." *Remote Sensing*, 7 (9):11326. 10.3390/rs70911326

Yeom, J. M., Y. K. Seo, D.-S. Kim, and K. S. Han. 2016. "Solar Radiation Received by Slopes Using COMS Imagery, a Physically Based Radiation Model, and GLOBE." *Journal of Sensors*, 2016:1–15. doi: 10.1155/2016/4834579.

Yoshikane, T., F. Kimura, and S. Emori. 2001. "Numerical Study on the Baiu Front Genesis by Heating Contrast between Land and Ocean". *Journal of the Meteorological Society of Japan*,79 (2):671–686. doi: 10.2151/jmsj.79.671.

Zhang, G., X. Xiao, C. M. Biradar, J. Dong, Y. Qin, M. A. Menarguez, and Y. Zhou. 2017. "Spatiotemporal Patterns of Paddy Rice Croplands in China and India from 2000 to 2015" *Science of the Total Environment*, 579:82–92.doi: 10.1016/j.scitotenv.2016.10.223.

Zhang, G., X. Xiao, J. Dong, W. Kou, C. Jin, Y. Qin, Y. Zhou, J. Wang, M. A. Menarguez, and C. Biradar. 2015. "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*, 106:157–171. doi: 10.1016/j.isprsjprs.2015.05.011.

Zhang, X., X. Xu, Y. Liu, J. Wang, and Z. Xiong. 2016. "Global Warming Potential and Greenhouse Gas Intensity in Rice Agriculture Driven by High Yields and Nitrogen Use Efficiency." *Biogeosciences*, 13:2701–2714. doi: 10.5194/bg-13-2701-2016. 9

Zhang, X., and Q. Zhang. 2016. "Monitoring Interannual Variation in Global Crop Yield Using Long-term AVHRR and MODIS Observations." *ISPRS Journal of Photogrammetry and Remote Sensing*, 114:191–205. Doi: 10.1016/j.isprsjprs.2016.02.010.

Zhao, Y., S. Chen, and S. Shen. 2013. "Assimilating Remote Sensing Information with Crop Model Using Ensemble Kalman Filter for Improving LAI Monitoring and Yield Estimation." *Ecological Modelling*,270:30–42.doi: 10.1016/j.ecolmodel.2013.08.016.

Appendix Equations

$$\Delta D = Max[T - T_b, 0] \quad \text{(A1)}$$

where ΔD is a daily change in growing degree days, T is the average daily air temperature (°C), and T_b is the crop-specific base temperature. Daily increase in the aboveground dry mass (ΔM) was calculated using the equation$\Delta G = P_2 \cdot \Delta M$, where ε is the crop-specific radiation use efficiency (RUE), and Q is the daily total photosynthetically active

$$Q = \beta \cdot R \cdot (1 - e^{-k \cdot LAI}) \quad \text{(A2)}$$

where Q is the absorption of photosynthetically active radiation, R is the incident daily total solar irradiance (MJ m^{-2}), β is the fraction of total solar irradiance that is PAR, and k is the crop-specific light extinction coefficient. Daily LAI increase with new leaf growth (ΔL) was obtained using the equation$\Delta L = \Delta M \cdot P_1 \cdot L_s$, where ΔM is the daily increase in AGDW, P_1 is the fraction of ΔM allocated to new leaves, and L_s is the specific leaf area.

$$P_1 = Max[1 - p_a \cdot e^{p_b \cdot D}, 0] \quad \text{(A3)}$$

where P_1 is *a* dimensionless leaf-allocation parameter, p_a and p_b are parameters that control the magnitude and shape of the function, and *D is* the cumulative GDD. The leaf senescence used in the model was formulated by assuming that the leaves would start to senesce after attaining the maximum LAI and that the senescence rate varies depending on plant genetic traits and environmental conditions. Daily increase in grain (ΔP) was calculated using the equation$\Delta G = P_2 \cdot \Delta M$, where P_2 is the fraction of ΔM partitioned to the grains and ΔM is the daily increase in AGDW.

$$P_2 = Max[1 - p_a \cdot e^{p_b \cdot fG_D}, 0] \quad \text{(A4)}$$

where P_2 is a dimensionless grain-partitioning parameter, p_a and p_b are parameters that control the magnitude and shape of the function, and fG_D *is* the grain partitioning factor based on the cumulative GDD.

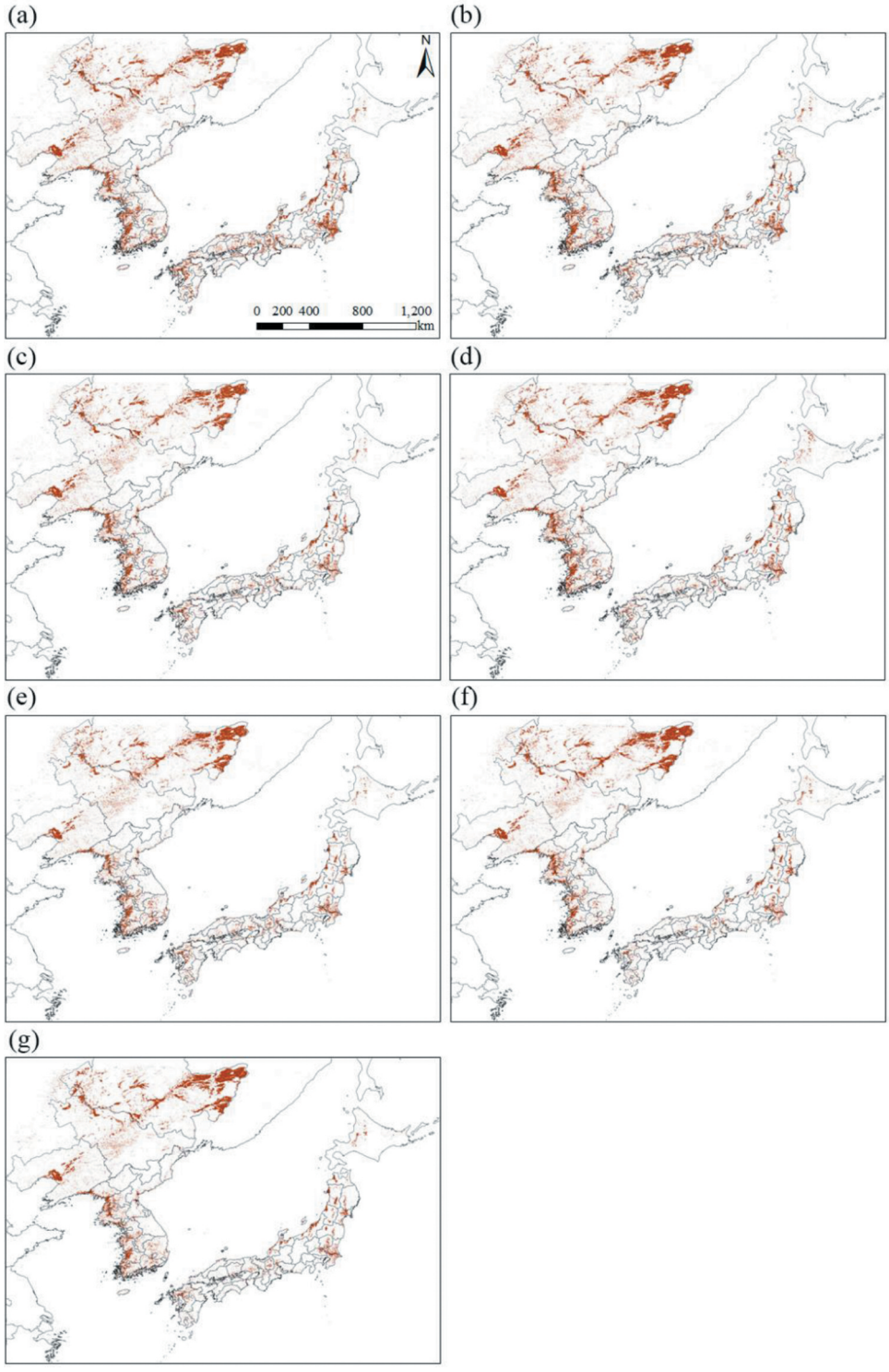

Figure. A1. Spatial distribution of classified paddy fields based on Geostationary Ocean Color Imager (GOCI) and Moderate Resolution Imaging Spectroradiometer (MODIS) from 2011 to 2017 (a–g, respectively) in Northeast Asia.

Appendix Figures

Appendix Tables

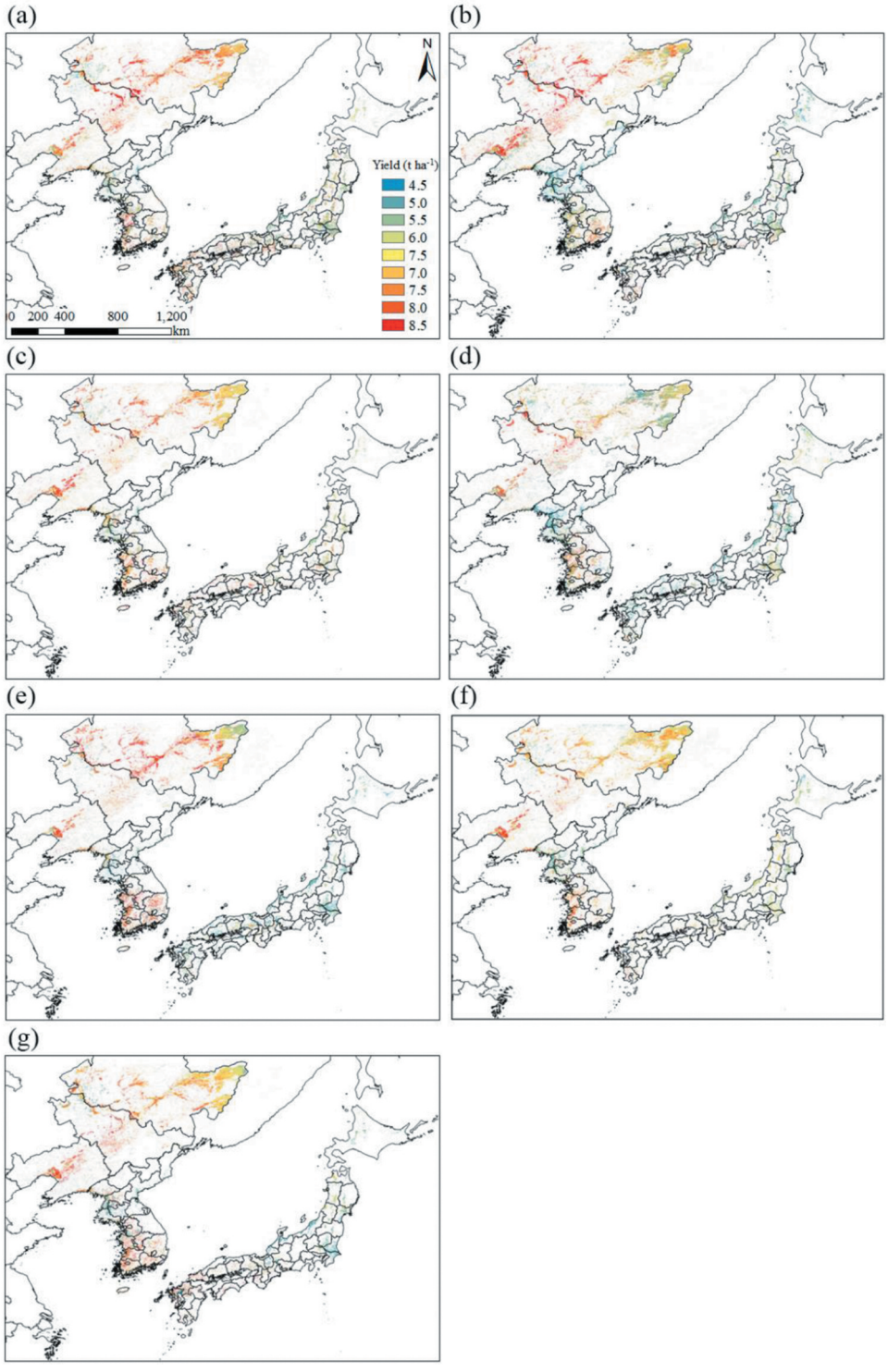

Figure. A2. Spatial distribution of simulated rice yields in Northeast Asia from 2011 to 2017 (a–g, respectively) based on Communication, Ocean, and Meteorology Satellite (COMS) images integrated into the RSCM for rice.

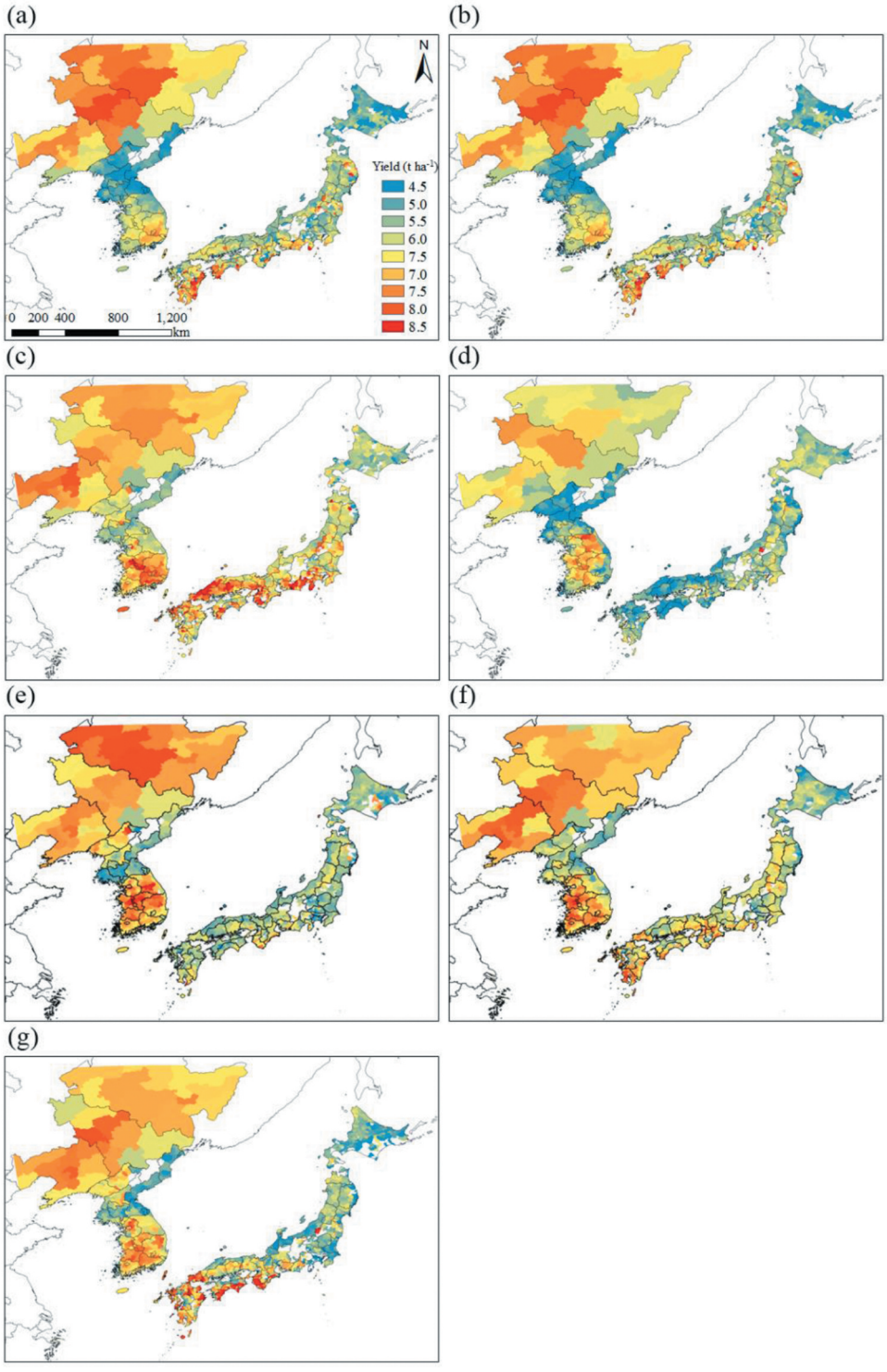

Figure. A3. County- (or province-) wide spatial distribution of simulated rice yields in northeast asia from 2011 to 2017 (a–g, respectively) based on Communication, Ocean, and Meteorology Satellite (COMS) images integrated into the RSCM for rice.

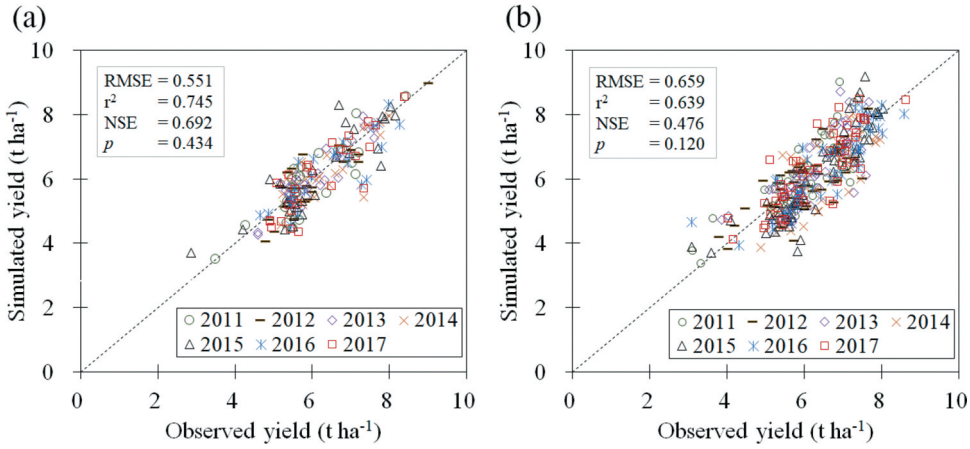

Figure. A4. Comparisons of observed and simulated rice yields for (a) 27 calibration and (b) 53 validation counties (or provinces) in Northeast Asia from 2011 to 2017. The dotted lines are 1:1 lines. RMSE, NSE, and *p* represent root mean square errors, Nash-Sutcliffe efficiencies, and p values of two-sample *t* tests (α = 0.05). Refer to Tables 5 and 6 for the detailed statistical analysis results of the (a) calibration and (b) validation.

Table A1. Constant and parameter values used for RSCM for rice.

Symbol	Description	Unit[a]	Value
ε	Radiation use efficiency	g MJ^{-1}	3.49
k	Light extinction coefficient	na	0.6
L_s	Specific leaf area	m^2 g^{-1}	0.016
T_b	Base temperature	°C	12.0
L_0	Leaf area index at transplant	m^2 m^{-2}	0.2
a	The parameter in the leaf allocation function	na	0.325
b	The parameter in the leaf allocation function	na	0.00125
c	The parameter in the leaf senescence function	na	0.00125

[a] 1 μmol mol^{-1} is equivalent to 1 ppm.

Table A2. Root mean square errors (RMSE), coefficient of determination (r^2), Nash-Sutcliffe efficiencies (NSE), and p values of two-sample t-tests (α = 0.05) between observed and simulated rice yields for the model calibration in 27 selected counties (or provinces) in Northeast Asia from 2011 to 2017.

Year	Observation	Simulation	RMSE	r^2	NSE	*t* test
-	- - - - - - - t ha^{-1} - - - - - - -			-	-	– *p* –
2011	5.97	5.95	0.535	0.758	0.699	0.951
2012	5.97	5.85	0.506	0.752	0.690	0.663
2013	6.06	6.09	0.367	0.834	0.802	0.901
2014	6.31	6.15	0.572	0.672	0.581	0.581
2015	6.14	6.08	0.680	0.749	0.705	0.863
2016	6.30	6.15	0.557	0.737	0.696	0.571
2017	6.14	6.02	0.592	0.661	0.621	0.676

Table A3. Root mean square errors (RMSE), coefficient of determination (r^2), Nash-Sutcliffe efficiencies (NSE), and p values of two-sample t-tests ($\alpha = 0.05$) between observed and simulated rice yields in 54 counties (or provinces) of interest in Northeast Asia from 2011 to 2017.

Year	Observation	Simulation	RMSE	r^2	NSE	*t* test
-		– – – – – – – t ha^{-1} – – – – – – –		-	-	– *p* –
2011	6.02	6.10	0.673	0.623	0.478	0.683
2012	5.98	5.82	0.684	0.462	0.286	0.314
2013	6.24	6.27	0.685	0.589	0.353	0.894
2014	6.45	6.16	0.767	0.535	0.108	0.130
2015	6.36	6.16	0.726	0.782	0.596	0.454
2016	6.50	6.28	0.620	0.751	0.660	0.299
2017	6.26	6.27	0.658	0.685	0.543	0.953