Detecting fluctuation of rice cultivated areas in semi-arid regions by combined use of MODIS and Landsat imageries

Takanori Nagano¹, Yumiko Ono², Akihiko Kotera³ and Ranvir Singh⁴

¹Graduate School of Agricultural Science, Kobe University, Japan
²Technical Section, Kitai Planning, Japan
³Research Department, Research Institute for Humanity and Nature, Japan
⁴Institute of Agriculture and Environment, Massey University, New Zealand

Abstract:
We developed an improved method to detect rice cultivated areas in semi-arid regions by combining the usage of Landsat imagery for detecting rice fields and MODIS for spatial and temporal upscaling of rice cultivation areas. We selected Haryana State in northwestern India as a case study area, where average farm plot size is small (~4,000 m²) and cultivated areas largely fluctuate from year to year due to water availability. Firstly, rice cultivated areas were detected by unsupervised classification of Landsat images of a specific year. Secondly, two conditional parameters for time-series EVI and LSWI were optimized by Powell’s method to best match the rice cultivated areas detected by MODIS to the areas detected by Landsat. Thirdly, calculated rice cultivated areas were compared with statistical data. Until results showed reasonable agreement, rice cultivated areas were reclassified in Landsat images and the following procedures were repeated. A good coefficient of determination ($R^2 = 0.82$) was obtained between the estimated rice area and the statistical data at the district (sub-state) level in the study area. This demonstrates the potential and increased accuracy of the developed methodology to detect and map harvested rice areas in water scarce arid and semi-arid regions.

KEYWORDS irrigated agriculture; water scarcity; remote sensing; EVI; LSWI

INTRODUCTION

Climate change may lead to changes in the frequency, intensity, spatial extent, duration, and timing of extreme weather and climate events, and it can result in unprecedented extreme weather and climate events (IPCC, 2012). In recent years, irrigated agricultural areas recognized as stable food production areas in semi-arid and arid regions are increasingly affected by such climatic events and gradual changes in climatic patterns and water availability. Adaptive capacity of irrigated agriculture against water resource fluctuations is difficult to assess because it comprises many technical and institutional factors. It is only visible in its past behavior against water crises of different magnitudes. Therefore it is important to collect and analyze data on cropping patterns, productivity, water resource availability and management, to help inform policy and planning for sustainable management of irrigated agriculture in increasingly water scarce arid and semi-arid regions.

Rice is the staple and dominant crop cultivated in large parts of Asian agriculture regions. Many researchers have attempted to detect changes in rice cultivated areas using remote sensing techniques. For example, Xiao et al. (2002) used Normalized Difference Vegetation Index (NDVI) and Normalized Difference Water Index (NDWI) observed by SPOT (Satellite Pour l’Observation de la Terre)-4 to detect characteristic features of rice fields in western China. Xiao et al. (2006) similarly used Enhanced Vegetation Index (EVI) and Land Surface Water Index (LSWI) derived from MODIS (MODerate resolution Imaging Spectrometer) imagery to detect rice fields in 13 of the South and the Southeast Asian countries. These methods showed that the accuracy of estimated total rice area was high at the state level, but low at the sub-state level. Sun et al. (2009) and Peng et al. (2011) followed the methodology of Xiao et al. (2006) with changed threshold values of EVI and LSWI for different regions and different years to improve the accuracy of rice area detection at sub-state level in China. Generally, unsupervised classification, which does not require ground truth data, has been used to detect and map rice cultivated areas on larger spatial and temporal scales. This however limits the accuracy achieved in detection and mapping of rice cultivated areas.

In this study, we propose combined use of Landsat and MODIS imageries and a step-wise reclassification method to improve the accuracy of rice area detection. We aim to achieve reasonable accuracy of rice area detection by effective use of freely available satellite imagery and commonly available statistical crop data, so that it can be implemented anywhere. Landsat is used to increase the accuracy of rice field detection where size of cultivation plots are small (~4000 m²) and land use is heterogeneous. We selected Haryana State of India as a case study area, and carried out the analysis to estimate and map rice cultivated area over 11 years from 2002 to 2012.

STUDY AREA

Haryana State is situated in the northwestern part of India (Figure 1). It has a semi-arid and arid climate with annual precipitation ranging between 344 mm in Fatehabad and 1,108 mm in Yamunanagar (CGWB, 2006). About 80% of the annual precipitation falls in the rainy season between...
Out of 4.42 million ha of the state’s territory, about 3.7 million ha or 84% is agricultural land. Crops are typically cultivated twice a year, mainly rice and cotton in the Kharif (monsoon, summer) season and wheat in the Rabi (post monsoon, winter) season.

In Haryana State, the average size of a farm plot is around 4,000 m². Farms are irrigated with water sourced from surface water (canal) and groundwater resources, where pumped groundwater accounts for 64%, canal water for 1% and conjunctive use of two (canal and groundwater) for 35% on the irrigated area basis (Erenstein, 2009). Electricity fees for electric pumps are heavily subsidized in Haryana State and groundwater use is critical, using more than 100% of estimated annual groundwater recharge in some parts of the district (CGWB, 2006). The key factors that affect irrigated agriculture in the state are low and erratic rainfall, delay and failure of monsoons, fluctuations in canal water supply, and rising and declining groundwater levels.

DATA AND METHODS

MODIS data

In this study, four sets of Terra/Aqua MODIS collection 5 BRDF-corrected-surface-reflectance 8-day composite data (MOD09Q1, MYD09Q1, MOD09A1 and MYD09A1) (Vermote et al., 2011) were acquired from Land Processes Distributed Active Archive Center, USGS (LP DAAC). To produce a 250-m resolution map, we resampled 500-m MOD09A1 and MYD09A1 data to 250-m resolution by applying the ratio of red band (band 1) in MOD09Q1 and MYD09Q1 to that of MOD09A1 and MYD09A1. EVI (Equation (1)) and LSWI (Equation (2)) were then calculated.

\[
EV\text{I} = \frac{2.5 \times (\rho_{\text{NIR}} - \rho_{\text{red}})}{\rho_{\text{NIR}} + 6 \times \rho_{\text{red}} - 7.5 \times \rho_{\text{swir}} + 1.0}
\]

\[
LSWI = \frac{(\rho_{\text{NIR}} - \rho_{\text{SWIR}})}{(\rho_{\text{NIR}} + \rho_{\text{SWIR}})}
\]

where \(\rho_x\) is the reflectance of spectral band \(x\), \text{NIR}: near infrared, \text{SWIR}: shortwave infrared.

To produce noise-free and smooth time-series, we applied the following procedure to \(EVI\) and \(LSWI\) time-series;
1. Eliminate cloud and sensor noise pixels by referring the image quality information.
2. Reconstruct daily time-series data by referring the actual date of observation for each pixel.
3. Synthesize daily time-series of Terra and Aqua (The Aqua data is available after the year 2002).
4. Interpolate eliminated pixel value on time-series axis by linear interpolation method.
5. Apply the Savitzky–Golay (SG) filter (Savitzky and Golay, 1964; Chen et al., 2004; Gu et al., 2009) to time-series data to produce smoothed time-series data.

Landsat ETM+

Because of frequent cloud interference, most of the Landsat archive data over Haryana State during rice cultivation (rainy) seasons between 2002 and 2012 were not useful except for three images on the path P147R40 acquired on 15 June, 19 September and 5 October, 2002. After converting irradiance to reflectance, atmospheric corrections were made using the Fast Line-of-sight Atmospheric Analysis of Hypercubes (FLAASH) algorithm of ENVI (Exelis) followed by calculation of the \(EVI\) and \(LSWI\). The FLAASH algorithm uses first-principles theory to correct wavelength in the visible through near-infrared and shortwave infrared regions. Unsupervised classification was carried out on combination of 6 bands and two indices of different images using K-means method to derive 6 classes. After checking consistency of classification using phenological pattern of different classes of MODIS EVI time series, 1) \(EVI\)-\(LSWI\) in the image of 15 June and 2) \(EVI\) and Bands 1–5 and 7 in the image of 19 September were found as useful vectors representing mixed land use of rice with other crops. The vector 1) discriminates state of inundation of paddy and the vectors 2) discriminate growth and maturity of rice near harvest. Therefore multiple combinations of 6 classes were used as the reference rice area to optimize rice field detection by MODIS images.

Statistical data of rice cultivation

District wise statistical data of rice cultivated area in the study area was obtained from the website of the Government of India (2014). The area sown for rice was used. Statistical data of total rice production in India was obtained from FAOSTAT database (FAOSTAT, 2014).

Determination of threshold parameters of MODIS indices

The process of rice detection is illustrated in Figure 2. To best-match rice area derived from Landsat ETM+ images, threshold parameters \(a\) and \(\beta\) were optimized using the Powell method for MODIS time series in Equation (3) and (4). This is a modification of the rice detection method by Xiao et al. (2006). Equation (3) detects the early stage of rice cultivation when \(EVI\) of transplanted rice generally becomes
close to or smaller than \( LSWI \) in inundated fields. Parameter \( \alpha \) was introduced to optimize the threshold for a specific region. Equation (4) discriminates rice from natural vegetation growing in inundated areas because \( EVI \) of rice generally becomes larger than that of natural vegetation in such areas. It also excludes rice which did not grow to a certain level.

\[
EV < LSWI + \alpha \quad \text{Peak EVI} > \beta
\]

Once \( \alpha \) and \( \beta \) were determined, their values were used to derive rice area over the entire state for 11 years from 2002 to 2012. The estimated rice areas at district level were then compared with the statistical data obtained. The agreement between the estimated and statistical rice area was low, \( \alpha \) and \( \beta \) parameter optimization was repeated to match different combination of classes in Landsat ETM+. This was repeated until rice areas reached a good agreement with the statistical data. Finally, \( \alpha \) and \( \beta \) were determined as 0.33 and 0.4038, respectively. Such optimization is possible only where sub-state level statistical data are accessible. Compared to the case where only state-level data are accessible, an increased number of samples contributes to the better optimization.

**Verification of estimation error**

The size of a single MODIS pixel (250 m × 250 m) is far larger than the average size of a single plot (4,000 m\(^2\)) in Haryana and therefore omission and commission errors of rice detection were inevitable. To evaluate these errors, histograms of peak EVI values of pixels classified as rice were drawn for each district in each year. Peak EVI values of real rice pixels in each district were assumed to be normally distributed. The assumed shape of the normal distribution was drawn using that part of the histogram that is larger than the median value of EVI and assuming the line symmetry shape for the part smaller than the median value (Figure 3). Then the actual histogram was compared with the assumed shape of the normal distribution.

**RESULTS AND DISCUSSION**

**Estimated rice cultivated area**

Figure 4 shows the relation between the statistical data of
district wise rice cultivated area (2002–2012) and the estimated area by MODIS. As a result of repeated optimization, a coefficient of determination \( R^2 = 0.82 \) was obtained between the estimated and statistical rice area in the state. The regression line, however, indicates the overall tendency of under-estimation of the rice cultivated area (Figure 4). This will be discussed in the following section.

Figure 5 reproduces the maps of estimated rice cultivated area from 2002 to 2012. The map of 2002 shows much smaller areas of rice cultivation compared to the other years. Although Landsat images of year 2002 were used to optimize threshold parameters \( \alpha \) and \( \beta \) MODIS platform Aqua was not in full operation at the time so cloud noise reduction of 8-day composites was not performed, unlike the later years. This resulted in relatively lower accuracy of rice detection in 2002 as compared to the other years. The intensity of rice fields seems higher in the northeastern part of the state and it generally decreases towards the southwest with the decreasing trend of precipitation. This is also potentially affected by the canal irrigation network in the state, which runs from the northeast to the southwest (Figure 7), and relatively higher prospects of shallow groundwater availability in the northeastern parts of the state.

**Omission and commission errors**

Table I shows the average of omission and commission errors occurring in each district for 10 years from 2003 to 2012 together with the standard deviations. Omission errors were generally larger than the commission errors with the exception of Panchkula District. Larger commission and omission errors occurred as intensity of rice fields decreased from the northeast to southwest, resulting in more occurrences of mixed land use in a single pixel of MODIS. This might have affected rice detection in two ways. As the ratio of inundated area in a pixel decreases, threshold parameter \( \alpha \) in Equation (3) must be increased so as not to omit rice fields included in the pixel. This resulted in larger commission errors. As the EVI value of a pixel with the mixed land use became lower than that of the rice only, the omission error would occur due to Equation (4). These omission and commission errors are inevitable with the analysis of images of coarse resolution. Use of fine resolution optical images, however, is not the immediate solution as return period of the platform (as in the case of Landsat) is long which would make cloud noise reduction difficult. We would need either increased number of observations by multi-platform high-resolution optical sensors or non-cloud-affected observations by high-resolution Synthetic Aperture Radars to overcome this issue.

**Changes in the rice cultivated area**

Figure 6 shows the comparison of estimated rice cultivated area in Haryana State with the statistical rice cultivated area (Government of India, 2014) and total rice production data in India (FAOSTAT, 2014). Over the years, the estimated rice cultivated area is largely fluctuating whereas the statistical rice cultivation area remains relatively stable. This difference might come from the definition of cultivated area. The methodology taken in this study detects the harvested rice area due to the condition of Equation (4). The statistical data, on the other hand, is generally based on the sown area collected by the government agencies for water use or subsidies. The difference seems to largely come from the fluctuation of estimated rice cultivated area in the southwestern part of the state (Figure 5). In Figure 6, the estimated rice area
showed a good agreement with the annual rice production of India, in terms of their fluctuations from year to year. Therefore the methodology taken in this study seems capable of estimating rice production changes at the state level. However, due to different degrees of commission and omission errors occurring in different districts, the methodology taken in this study seems incapable of estimating district-wise rice productions correctly.

Figure 7 shows the coefficient of variation of estimated rice cultivated area in each district of Haryana state over 10 years. The districts situated towards the southwestern part of the state showed a large fluctuation in the rice cultivated area. This is potentially due to low and erratic precipitation and relatively less canal water and shallow groundwater availability in the southwestern part of the state.

### CONCLUSIONS

The combined use of unsupervised classification of
Landsat images to optimize rice detection by EVI and LSWI of MODIS imagery improved accuracy of rice area detection at sub-state level in Haryana State, India, from 2002 to 2012 where average plot size is small (~4,000 m²). Higher accuracy was obtained by creating 8-day composite data of MODIS indices using observations of both Aqua and Terra. Omission and commission errors were evaluated assuming a normal distribution of peak EVI of rice fields in different districts of the state. Due to the coarse resolution of analysis and small plot size (~4,000 m²), mixed land use occurred which attributed to higher omission and commission errors in the districts where intensity of rice cultivation was relatively lower.

The proposed method, however, demonstrated its potential to estimate with reasonable accuracy the harvested rice area rather than the sown area in the state. It was capable of successfully capturing large spatial and temporal fluctuations in the cultivated rice area from year to year probably due to water availability. A higher coefficient of variation of rice cultivation area was observed in water scarce districts of the state. The proposed methodology can be applied to larger spatial and temporal extents with relative ease and therefore it is promising for measuring adaptive capacity of rice cultivation and irrigation in increasingly water scarce semi-arid and arid regions.

ACKNOWLEDGEMENTS

This research was supported by Core Research for Evolutional Science and Technology program of the Japan Science and Technology Agency (JST-CREST) and JSPS KAKENHI Grant Number 15H04047. The research was partly conducted under the Program for Risk Information on Climate Change supported by the Ministry of Education, Culture, Sports, Science, and Technology-Japan (MEXT).

REFERENCES

CGWB. 2006. Ground water year book 2005–2006, National Capital Territory, Delhi. Central Ground Water Board, New Delhi; 31.
Chen J, Jönsson P, Tamura M, Gu Z, Matsushima B, Eklundh L. 2004. A simple method for reconstructing a high-quality NDVI time-series data set based on the Savitzky–Golay filter. Remote Sensing of Environment 91: 352–344. DOI: 10.1016/j.rse.2004.03.014.
Directorate of Economics and Statistics, Department of Agriculture and Cooperation, Ministry of Agriculture, Government of India. 2014. Crop Production Statistics for the Year 1998–99 to 2012–13. http://eands.dacnet.nic.in/. Last access June 30, 2015.
Erenstein O. 2009. Comparing water management in rice–wheat production systems in Haryana, India and Punjab, Pakistan. Agricultural Water Management 96: 1799–1806. DOI: 10.1016/j.agwat.2009.07.018.
FAOSTAT. 2014. Area under principal crops. Food and Agriculture Organization of the United Nations. http://faostat3.fao.org/download/Q/QC/E. Last access June 30, 2015.
Gu J, Li X, Huang C, Okin GS. 2009. A simplified data assimilation method for reconstructing time-series MODIS NDVI data. Advances in Space Research 44: 501–509. DOI: 10.1016/j.ast.2009.05.009.
Intergovernmental Panel on Climate Change (IPCC). 2012. Managing the risks of extreme events and disasters to advance climate change adaptation. A special report of working groups I and II of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge, UK, and New York, NY, USA; 582.
Land Processes Distributed Active Archive Center (LP DAAC). USGS. http://e4ftl01.cr.usgs.gov. Last access June 24 2015.
Peng D, Huete AR, Huang J, Wang F, Sun H. 2011. Detection and estimation of mixed paddy rice cropping patterns with MODIS data. International Journal of Applied Earth Observation and Geoinformation 13: 13–23. DOI: 10.1016/j.jag.2010.06.001.
Savitzky A, Golay MJE. 1964. Smoothing and differentiation of data by simplified least squares procedures. Analytical Chemistry 36: 1627–1639. DOI: 10.1021/ac60214a047.
Sun H, Huang J, Huete AR, Peng D, Zhang F. 2009. Mapping paddy rice with multi-date moderate-resolution imaging spectroradiometer (MODIS) data in China. Journal of Zhejiang University SCIENCE A 10: 1509–1522. DOI: 10.1631/jzus.a0820536.
Vermote EF, Kotchenova SY, Ray JP. 2011. MODIS Surface Reflectance User’s Guide. http://modis-sr.ltdri.org/guide/MOD09_UserGuide_v1_3.pdf. Last access June 30, 2015.
Xiao X, Boles B, Froliking S, Salas W, Moore III B, Li C, He L, Zhao R. 2002. Landscape-scale characterization of cropland in China using Vegetation and Landsat TM images. International Journal of Remote Sensing 23: 3579–3594. DOI: 10.1080/01431160110106069.
Xiao X, Boles S, Froliking S, Li C, Babu JY, Salas W, Moore B. 2006. Mapping paddy rice agriculture in South and Southeast Asia using multi-temporal MODIS images. Remote Sensing of Environment 100: 95–113. DOI: 10.1016/j.rse.2005.10.004.