MODIS Satellite Data Based Rice Yield Forecasting Model for Sri Lanka: A Pilot Study in Kurunegala District

T. L. Dammalage[1*], P. M. T. S. Sirisena[1] and Junichi Susaki[2]

[1]*Department of Remote Sensing and GIS, Sabaragamuwa University of Sri Lanka, Sri Lanka*

[2]*Department of Civil and Earth Resources Engineering, Kyoto University, Japan*

Abstract

Rice is the staple food of the inhabitants of Sri Lanka. Hence, it is of paramount importance in proper strategic planning and decision making regarding the food security and facilitation of safe harvest storages, based on the vital statistics of accurate pre-harvest seasonal rice yield estimations. In Sri Lanka, the conventional method known as "crop cutting survey" is used to estimate the seasonal rice production. However, it fails to forecast rice yield before harvest as this experiment is conducted during the harvest. Several previous research studies have identified satellite Remote Sensing (RS) as a potential tool for estimating and forecasting crop yields. Therefore, this pilot study was focused to utilize free RS data for accurate pre-harvest seasonal rice yield estimations. This study used 250m spatial resolution satellite data from 2007 to 2016 over Kurunegala district of Sri Lanka, captured by Moderate Resolution Imaging Spectro-radiometer (MODIS) sensor onboard NASA EOS Terra satellite. The challenge of identifying the cultivated paddy lands for a considerable accuracy with the use of 250m resolution MODIS satellite data has been overcome by analyzing the unique temporal dynamics of paddy cultivation. In this study an average overall accuracy of approximately 77% was achieved for identifying cultivated paddy lands. According to the comparison of estimated yield based on the proposed method with national yield statistics, both NDVI and EVI2 based models provided more reliable estimations nearly 80 days after transplanting. However, the EVI2 based model shows more reliable results than NDVI model based estimations with an average overall accuracy of 92%. Therefore, this experiment has demonstrated that the seasonal rice yield can successfully be forecasted one month prior to harvest with considerably higher accuracy.

Key words: Rice yield forecasting, MODIS-(MOD09Q1) image product, NDVI, EVI2.

1. Introduction

Rice is significant to the livelihood of Sri Lanka both as a widely consumed food source and a primary agricultural product. The paddy crops are cultivated as a wetland crop in all the districts of Sri Lanka with total land devoted for paddy estimated to be approximately 708,000 hectares (DCS, 2017). There are two paddy cultivation seasons, namely *Maha* and *Yala*, which are synonymous with the two principal annual monsoons. However, the entire area devoted for paddy is not being cultivated due to a number of factors including shortage of water during the growing seasons and prevailing unsettled conditions in the ground (DCS, 2017). Average paddy yield in Sri Lanka is estimated by means of a sample survey. This practice, commonly known as "crop cutting survey", was introduced by the Food and Agriculture Organization of the United Nations (FAO) in 1951. At present, a sample of 3,000 villages for the main season (*Maha*) and 2,000 villages for the second season (*Yala*) are selected to conduct this survey (DCS, 2017). However, this method fails to forecast rice yield before the harvest as the sampling is conducted during the harvest. In addition, it is time-consuming, subjective, and prone to significant discrepancies because of limited ground observations, resulting in poor rice yield estimations.

The capability of estimating seasonal crop production before a substantial time to harvest is truly important to take necessary management decisions regarding seasonal crop production. In particular, it is of vital importance in strategic

* Corresponding author: thilantha@geo.sab.ac.lk

planning and decision making regarding the food security and facilitation of safe harvest storages. According to the Sri Lankan Department of Agriculture, in the 2015/2016 *Maha* and 2016 *Yala* seasons there were 2.5 and 1.3 million metric tons of rice produced, respectively. However, during these two seasons, it was recorded that farmers in many parts of the country waited days near the storage facilities to receive payment for their crop. In the end, they were disappointed over the failure to set up an adequate mechanism to purchase their harvest. On the other hand, in early 2017, the price of rice increased significantly due to the inadequate stores of rice within the country. These two recent incidents reflect the inadequate management and poor decision making of national rice production in Sri Lanka. The conditions of surplus in 2015/2016 and scarcity in 2017 were primarily due to the unavailability of a system to forecast the expected paddy yield by a considerable time before harvesting.

The previously mentioned two incidents could have been overcome by proper import or arranging additional storage facilities (in shortfall case) and export (in surplus case) policies, taken based on reliable yield estimations as suggested by Noureldin *et al*. (2013). Since there is a critical requirement for pre-harvest yield estimation system for Sri Lanka, this study is focused on utilizing modern science and technology in proposing an accurate, reliable and cost effective paddy yield forecasting model.

Ferencz *et al*. (2004) concluded with their study that Satellite Remote Sensing (RS) can be successfully applied for this particular research area as it is a powerful and effective tool for estimating and forecasting crop yields. However, accurate identification of crop cultivated lands using RS data is imperative before building yield forecasting models. In the case of paddy cultivation, some unique physical characteristics (paddy plants are grown on flooded or irrigable soil) and distinct temporal dynamics of paddy cultivation help to classify paddy from other crops. Moreover, the temporal dynamics of paddy cultivation can be characterized by three main periods; the flooding and rice transplanting, growing (vegetative growth, reproductive and ripening stages) and the period following harvest (Toan *et al*., 1997). Therefore, the time dependent relationship between Land Surface Water Index (LSWI) and Normalized Difference Vegetation Index (NDVI) or Enhanced Vegetation Index (EVI) at the period of rice flooding and transplanting can be used to identify flooded pixels (Xiao *et al*., 2005a). Accordingly, the paddy cultivated pixels can be isolated by applying a water and forest mask or by analyzing the changing pattern of vegetation index values (Xiao *et al*., 2002b, 2002c, 2005b and Lio *et al*., 2011). However, this method has posed a great challenge at large spatial scales (Xiao *et al*., 2005b) since individual farmers follow diverse flooding and transplanting schedules in their paddy rice fields. Moreover, the spatial & temporal resolutions and the cost of satellite imagery are also considerable challenges.

Accordingly, even after a relatively long time (more than 20 years), no routine yield estimation method has been developed for a wide range of operational applications (Ferencz *et al*., 2004). However, considerable developments in this field have been achieved at present day due to the open access of satellite data in higher temporal resolutions. Some studies have discovered a linear relationship between rice yield and vegetation indices calculated based on satellite data (Cheng *et al*., 2010) whereas some have showcased an exponential relationship (Nuarsa *et al*., 2011). Yet, the most important issue addressed by any yield forecasting experiment involves determining at which growth stage of the plant can most accurately forecast its expected yield. Various solutions to this question have been proposed through numerous studies. A study was conducted in Indonesia using Landsat ETM+ data which found that the best age of a paddy plant for yield prediction is 63 days after transplanting (Nuarsa *et al*., 2011). Another experiment conducted in Egypt has identified that the best age for yield prediction is 90 days after transplanting (Noureldin *et al*., 2013). Factors such as variation in soil attributes, climatic conditions, paddy varieties, and different agricultural practices from one area to another have led to these diverse conclusions. Thus, yield estimation models developed or tested locally are compatible solely for local use and not for regional or global use (Shresthan *et al*., 2003).

At present there are several earth observation satellites systems available which offer open-access imagery. One of these systems which is of particular interest is the Moderate Resolution Imaging Spectro-radiometer (MODIS) sensor. MODIS has a high temporal resolution of 12 hours, making it ideal for the type of temporal analysis in this study. Therefore, the objective of this study to explore the potential of applying remote sensing technology for rice yield estimation and to propose a unique rice yield forecasting model for Sri Lanka based on MODIS satellite data. In order to model with reliable accuracy, the study aims to propose a new method to identify paddy cultivated lands in each season (*Yala* and *Maha*) using MODIS satellite data, as well as to determine the best age of paddy plants for yield forecasting in Sri Lankan conditions.

2. Study Area and Materials Used

2.1 Study Area

For this initial experiment of formulating a yield forecasting model for Sri Lanka, the Kurunegala district in the North-Western province of Sri Lanka was selected as the study area. The district has a total area of 481,600 hectares and provides a significant contribution for the total rice production in the country. According to the Sri Lankan Census and Statistics Department's data, the district had an average cultivated paddy area of about 61,000 hectares and average rice yield of 3,775 kg per hectare from 2007 to 2016 (DCS, 2017). The climatic condition of the study area is tropical with major changes in weather occurring during the monsoons from May to August and October to January every

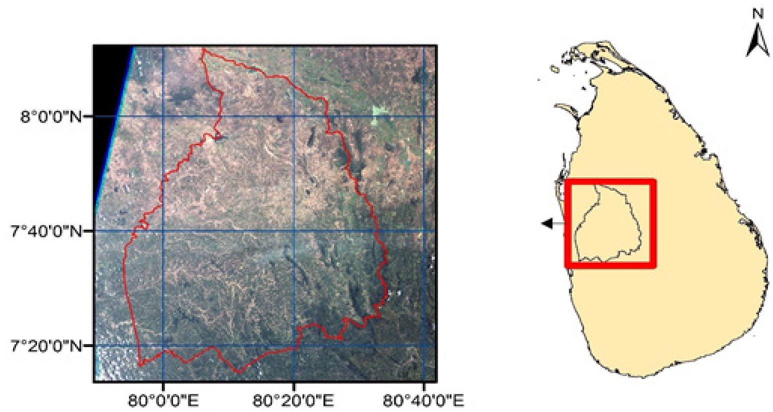


Figure 1. Spatial extent and location of the study area (Kurunegala district). Red color polygon marked on a Landsat 8 true color composite image, acquired in 2015

year. With the expected heavy rains during these two periods of the year, two seasonal rice crops are common in this region.

Kurunegala district was selected for this initial experiment, in particular, due to its availability of miscellaneous spatial size of paddy land plots, extending from 1or 2 acres to several hectares. Moreover, based on field investigations it was found that the farmers of the area adopt slightly different transplanting time intervals. Therefore, this initial experiment would also provide an applicability analysis of proposed methodology to be utilized for the whole country.

2.2 Satellite and Rice Yield Data Used

MOD09Q1 is the satellite image product used in this study. It is a composite product generated with the use of eight consecutive days of 250m spatial resolution MODIS satellite imagery. The MOD09Q1 image product has uniform spatial resolution and retains the "best" observation during each eight day period, for every cell in the image (Vermote *et al*., 2011). This has helps to reduce and eliminate clouds or cloud shadows from a satellite image (Vermote *et al*., 2015). This product contains three MODIS science data sets, including 250m surface reflectance band 1 (620-670 nm), 250m surface reflectance band 2 (841-876 nm) and 250m reflectance band quality data set. For this study, three-hundred ninety-one MOD09Q1 images from 2007 to 2016 were acquired through the USGS-EROS data center (http://edc.usgs.gov/). Rice yield data and other required local paddy statistics from 2007 to 2016 were obtained from the Sri Lankan Census and Statistics Department.

3. Methodology

In order to accomplish the main objective of deriving a unique rice yield forecasting model for Sri Lanka based on MODIS satellite data, the study first aimed to propose a method to identify paddy cultivated lands in each growing season using MODIS satellite data (Figure 2).

Paddy cultivation in Sri Lanka is based on two monsoons, known as "*Yala*" and "*Maha*" seasons. The beginning of the *Yala* season (flooding and transplanting) is variable and can take place from the end of March to mid-May (DCS, 2017). Similarly, *Maha* season can vary from end of September to mid-December (DCS, 2017). The study was conducted during the two paddy growing seasons and commenced during the same months of each year from 2007 to 2016. The time series analyses of selected vegetation indices were conducted for the identification of cultivated paddy lands throughout both seasons. Twenty four MOD09Q1 satellite images of 250 m spatial resolution were used for the time series analysis for both seasons.

3.1 Vegetation Indices Used

In this study, Vegetation Indices (VI) were used to minimize the effect of the vegetation canopy radiometric response by factors such as optical properties of the soil background, sun-illumination, view geometries and meteorological factors (wind, cloud, etc.) as discussed by Leblon (1997). Vegetation Indices are measures of vegetation "greenness" obtained by combining the results of surface reflectance measurements of the vegetation canopy in different spectral bands (Myneni *et al*., 1995; Guzinski, 2010). The indices work based on the principle that vegetation reflects different amounts of electromagnetic radiation in different spectral bands (Guzinski, 2010). However, vegetation indices are not indicative of one particular physical property of the vegetation, but they represent a combination of properties such as leaf chlorophyll content, canopy cover and architecture or leaf area (Jiang *et al*., 2008). There are a number of different vegetation indices improved by researchers. The most commonly used vegetation indices include ratios of single-band or linear combined reflectance.

The first ratio-based vegetation index used was the Reflectance Ratio or Ratio Vegetation Index (RVI). RVI is the ratio between near infrared and red bands (Pearson *et al*., 1972). Latter, Rouse *et al*. (1974) improved the ratio

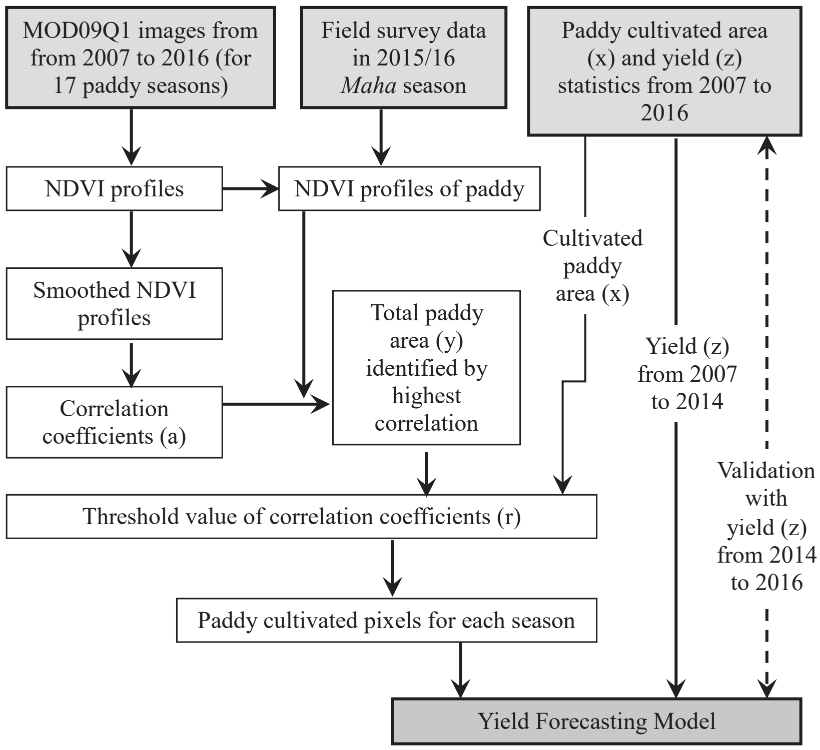


Figure 2. Proposed methodology.

vegetation index and derived the Normalized Difference Vegetation Index (NDVI) based on red (*RED*) and near infrared (*NIR*) spectral bands as illustrated in equation 1. Those two bands are used as green vegetation that displays strong absorption in the red wavelength region of the spectrum (reflectance of around 3-5%) and weak absorption in the NIR wavelength region (reflectance around 40 - 60%) (Gitelson, 2004).

$$NDVI = \frac{NIR - RED}{NIR + RED} \qquad (1)$$

Enhanced Vegetation Index (EVI) is formulated as a standard satellite vegetation product for the Terra and Aqua Moderate Resolution Imaging Spectro-radiometers (MODIS) (Huete *et al*., 1999). EVI provides improved sensitivity in high biomass regions while minimizing soil and atmospheric influence (Huete *et al*., 1999; Huete *et al*., 2002). The major limitation of EVI is that it utilizes the blue band in addition to the red and near-infrared bands. Jiang *et al*. (2008) developed and evaluated a two-band EVI (EVI2), without blue band, which has the best similarity with the three-band EVI, particularly when atmospheric effects are insignificant and data quality is good. EVI2 is defined by equation 2.

$$EVI2 = 2.5\,\frac{NIR - RED}{NIR + 2.4RED + 1} \qquad (2)$$

3.2 Cultivated Paddy Land Identification

One challenging objective of this study was to identify the cultivated paddy lands for a considerable accuracy with the use of 250m spatial resolution free satellite data. Such a classification is practically difficult due to the complexity of different paddy cultivation practices adopted by the farmers, even in this comparatively small geographical area. Furthermore, the amount of paddy cultivated lands changes from season to season; the beginning of flooding and transplanting time can also vary from location to location due to many different motivations. However, the unique temporal dynamics of paddy cultivation over other vegetation was used to overcome the challenge. Accordingly, the paddy pixels were distinguished from over other vegetation pixels by analyzing the correlation coefficients between pre-defined NDVI pattern/profile of paddy and NDVI time series obtained through satellite data.

Accordingly, NDVI values for each of the twenty-four, 8-day composite images (MOD09Q1 products) of each season were calculated using Equation 1. The resulting NDVI profiles were smoothed using the “moving average” method to calculate the correlation coefficients between NDVI changing patterns of normal paddy cultivation and pre-defined smoothed profiles. Subsequently, several correlation coefficients (a) were calculated for one smoothed NDVI temporal profile by the shifting matching pattern along the NDVI time series. By analyzing the resulting correlation coefficients, an appropriate threshold value (r) was selected to extract cultivated paddy pixels by equalizing the total paddy area (y) identified by this algorithm and available census data (x). The pixels that have higher correlation coefficient values than the identified threshold value were classified as paddy lands of that specific growing season. For the validation of this method of identifying cultivated paddy lands, several classified pixels were compared with ground truth data to determine whether those pixels are truly from paddy lands. The same procedure was repeated for all the other seasons to identify the cultivated paddy lands in the study area.

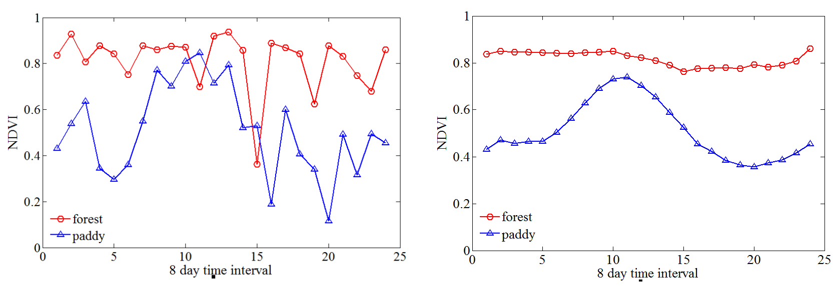
3.3 Rice Yield Forecasting Model

The rice yield forecasting models were derived based on the average NDVI and EVI2 values and maximum NDVI of each cultivated paddy pixel for seventeen (17) paddy growing seasons from 2007 to 2016. Then the relationship between the values of average vegetation indices and yield was modeled using the recorded average rice yield data from 2007 to 2014. Reflectance values, derived by satellite images, correspond to "greenness" of paddy plants and that is changing throughout the growing season according to the temporal change of paddy plants. Therefore, the vegetation indices, calculated based on reflectance values for a particular pixel on any growing season, are changing with the time. Therefore, the models were derived only for six different stages of paddy growth with a sixteen (16) day time interval.

The best model to represent the relationship between rice yield and vegetation indices was defined by comparing the results of the accuracy assessment obtained through linear and exponential models. For the accuracy validation of the derived models, the remaining average rice yield data from 2014 to 2016, which are not used for modeling, are to be used. In addition to the NDVI and EVI2 based models, another rice yield forecasting model was derived by taking cumulative NDVI (ΣNDVI) based on average NDVI values. The cumulative NDVI values were calculated by using both original NDVI temporal profile and smoothed NDVI profiles obtained through moving average technique.

4. Results and Discussions

4.1 Cultivated Paddy Area Identification



(a) Original NDVI pattern/profile | (b) Smoothed NDVI pattern/profile

Figure 3. Temporal change of NDVI values for a cultivated paddy and forest pixel

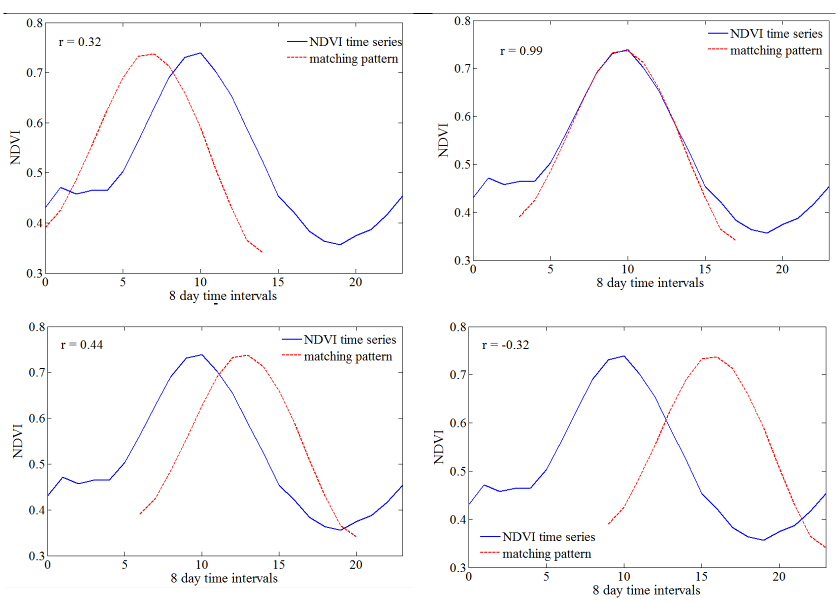


Figure 4. Shifting matching pattern along the NDVI time series derived for a specific paddy pixel on a growing season (correlation coefficient denoted by "r")

The cultivated paddy pixels were identified for each season by adopting the proposed methodology. Based on field investigations several comparatively larger (4 to 5 hectares) paddy cultivated lands in the 2015-2016 *Maha* season were located with GPS observations. NDVI values were calculated for each of those identified paddy pixels based on equation 1 for all 24 MOD09Q1 images (8 day) of the season to obtain the time series variation of NDVI. Figure 3(a) shows the original NDVI time series values for a paddy and a forest pixel. Figure 3(b) shows the same NDVI time series after the moving average filter technique was used to obtain the smooth values. For all the other tested paddy samples the same unique temporal change of the smoothed NDVI values was observed. Therefore, this was used as the pre-defined NDVI pattern/profile of cultivated paddy pixels and also as the shifting matching pattern of correlation calculation.

Several correlation coefficients were calculated by shifting the pre-defined NDVI matching pattern/profile along the NDVI time series derived based on MODIS 8 day composites data (MOD09Q1 images) for each season. The correlation coefficient (r) was taken the highest value at a point on its NDVI time series over cultivated paddy pixels. For instance, as illustrated in Figure 4, the calculated NDVI time series corresponded to a cultivated paddy pixel and it was correlated with the matching pattern showing maximum value of correlation coefficient of 0.99 at an only one specific time.

Based on the resulting correlation coefficients, an appropriate threshold value for each season was selected, by equalizing the total paddy area identified by this algorithm and available census data. The threshold values that were used to determine the cultivated paddy lands from other land uses are listed in Table 1, including the total cultivated paddy area identified and available in census data for each season. However, it is evident that for all seventeen (17) seasons, from 2007/2008 *Maha* to 2015/2016 *Maha*, the threshold value of correlation coefficients are not unique or common. Therefore, further investigation is required to identify a unique threshold value. It could be possible by varying the shape of the matching NDVI pattern/profile based on advanced smoothing and filtering techniques.

The pixels that have higher correlation coefficient values than the defined threshold value were then classified as cultivated paddy pixels for each season. By applying this method, the study area is classified into two classes: cultivated paddy and other lands. Accordingly, the spatial distribution of cultivated paddy lands from 2007/2008 *Maha* to 2015/2016 *Maha* seasons in Kurunegala district at 250m spatial resolution is illustrated in Figure 5(a) to 5(q) respectively. One of the significant observations made is that the cultivated paddy lands are sporadically distributed throughout the study area for each season. Moreover, it has also been identified that some of the paddy fields are not continuously cultivated in each season.

In the study area, most of the paddy lands are cultivated during *Maha* season with the southwest monsoon. In *Yala* season, it has been observed that only the paddy lands with

Table 1. The threshold value identified for each season to differentiate cultivated paddy pixels from other pixels.

Paddy cultivated Season	**Threshold value (r)**	**Total paddy area (hectares)**	
		Identified (y)	**Census data (x)**
2007/08 *Maha*	0.938	75050	75243
2008 *Yala*	0.889	67969	68188
2008/09 *Maha*	0.985	76206	76812
2009 *Yala*	0.875	46088	46516
2009/10 *Maha*	0.978	71381	68023
2010 *Yala*	0.943	55125	56386
2010/11 *Maha*	0.855	79494	80206
2011 *Yala*	0.943	66600	67458
2011/12 *Maha*	0.962	56600	56718
2012 *Yala*	0.965	14913	15153
2012/13 *Maha*	0.946	80581	80851
2013 *Yala*	0.920	51631	52300
2013/14 *Maha*	0.984	55156	55480
2014 *Yala*	0.942	51924	52587
2014/15 *Maha*	0.895	76385	77108
2015 *Yala*	0.921	62225	61319
2015/16 *Maha*	0.978	80119	80625

irrigated water are cultivated due to the lack of rainwater (DCS, 2017). Therefore, the main factor that controls the amount of cultivated lands is the water availability for paddy growth. In addition, this might also depend on the individual farmer interest. For instance, the spatial distribution of cultivated paddy lands in 2012-Yala season, a recorded dry period, are concentrated mostly around major water tanks, which only had the capability of supplying necessary water requirement of the paddy cultivation. Based on these observations, further investigations will be extended to perform spatial analyses for cultivated paddy area distribution and change for *Yala* and *Maha* seasons. It is expected, that this would provide valuable information for better water management practices to maintain a uniform cultivation for the majority of paddy lands in both seasons.

An accuracy assessment was performed to validate this classification (Figure 5) of cultivated paddy lands. Twenty-

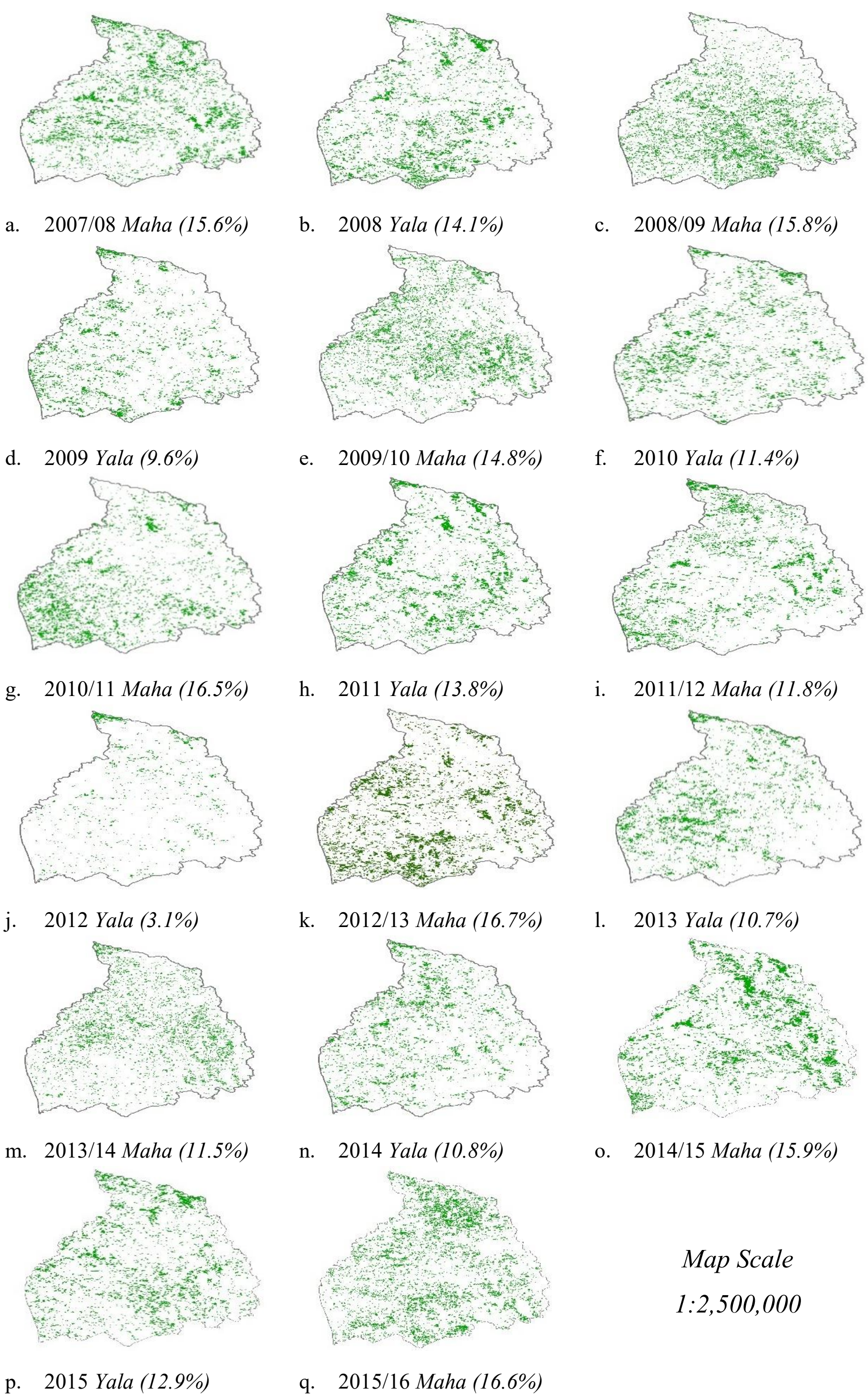


Figure 5. Spatial distribution and percentage of cultivated paddy fields (green) in Kurunegala district based on MODIS satellite data at 250m spatial resolution.

Table 2. Accuracy assessment of identified cultivated paddy lands.

Season	No of sample points (classified as paddy)	Actual Land Use		Accuracy (%)
		Paddy	Other	
2008 *Yala*	25	18	7	72
2009/10 *Maha*	25	19	6	76
2010/11 *Maha*	25	18	7	72
2012 *Yala*	25	21	4	84
2012/13 *Maha*	25	19	6	76
2014 *Yala*	25	20	5	80

Table 3. Rice yield forecasting models and their accuracies as a percentage of discrepancy with respect to national statistical data.

Forecaster	Model*	R^2 value	** Deviation %				
			2014 *Yala*	14/15 *Maha*	2015 *Yala*	15/16 *Maha*	Average
NDVI maximum	$y = 4672x + 123$	0.29	6.3	3.4	9.9	2.2	5.5
	$y = 1459e^{1.2097x}$	0.29	5.9	2.9	9.4	2.7	5.2
Cumulative NDVI	$y = 208x + 2673$	0.29	9.7	12.8	10.9	0.7	8.5
	$y = 2796e^{0.0556x}$	0.31	9.1	12.6	10.4	0.9	8.3
Cumulative NDVI (Smoothed)	$y = 408x + 1629$	0.40	7.4	7.2	5.7	1.5	5.5
	$y = 2143e^{0.1066x}$	0.41	6.9	6.6	5.1	1.9	5.1
NDVI after 80 days	$y = 3263x + 1746$	0.41	7.6	6.1	5.2	0.4	4.8
	$y = 2213e^{0.8498x}$	0.42	7.1	5.5	4.7	0.2	4.4
EVI2 after 80 days	$y = 8061x + 481$	0.43	10.4	1.8	11.4	0	5.9
	$y = 1603e^{2.0827x}$	0.44	9.7	1.4	10.8	0.4	5.6

*y - rice yield (kg per hectare)

*x - corresponding vegetation index

**Deviation % = ((model based yield - Census data)/ Census data)×100

five random points were selected to compare image classification results with ground truth data. The resulting accuracies for six different seasons are illustrated in Table 2 as a percentage of the tested sample points. Accordingly, the suggested cultivated paddy land identification method has demonstrated an average overall accuracy of 77% for identifying cultivated paddy lands.

Most of the sample points which are classified as “Other” in Table 2 are located near comparatively large paddy lands. One of the reasons for that could be the effect of mixed-pixels, due to the low spatial resolution (250m) of MODIS09Q1 data. Because of this, other land uses, which are located near comparatively large paddy lands, could represent as a paddy pixel in this classification. Similarly, the small paddy lands, which are surrounded by another land use, may not be represented as a paddy pixel. Using high spatial resolution imagery of cultivated paddy could possibly overcome pixels being classified as the other class. However, satellite imagery with high spatial and temporal resolutions are not freely available.

4.2 Rice Yield Forecasting Model

Linear and exponential relationships between rice yield and vegetation indices have been tested for average NDVI, maximum NDVI, average EVI2 and cumulative NDVI. They were calculated based on identified cultivated paddy pixels. The relationships, yield forecasting models, are derived for six different stages of paddy growth with a sixteen (16) day time interval. However, the model derived after approximately 80 days of transplanting has shown higher reliable results. The 80 days model is depicted in Table 3. Accordingly, the relationships based on NDVI and cumulative NDVI have shown comparatively lower determinant of coefficient (R^2) value of about 0.3. However, for smooth cumulative NDVI based relationship and EVI2 and NDVI relationships derived at 80 days after the transplanting have shown almost similar R^2 values of about 0.4. Therefore, all those listed six relationships, yield forecasting models, are showing almost similar correlation between the reference rice yield data. Hence, the listed relationships are reasonably accurate to use as rice yield forecasting models.

Rice yield is predicted for the remaining four paddy cultivation seasons, from 2014-*Yala* to 2015/2016- *Maha*, based on the derived models to evaluate the yield forecasting accuracies. The resulting accuracies are presented in Table 3 as percentage difference of model-based yield as well as census data based on the conventional “crop-cutting survey” method. Except for cumulative NDVI based models, all the other models have shown less than 6% discrepancy between model-based yield and census data. Out of those, the exponential model derived that is based on NDVI values at 80 days after the transplanting demonstrates the best accuracy with only 4.4% of discrepancy.

Hence, the exponential model based on NDVI values at 80 days after the transplanting is selected as the best model for rice yield forecasting using MODIS satellite imagery in this particular study area. Moreover, the proposed model has shown reasonable accuracy on rice yield forecasting approximately one month before harvesting (or about 80 days after transplanting). The model will allow authorities to make effective decisions over rice yield management without cost or field work. This study has successfully tested the proposed yield forecasting methodology.

5. Conclusions and Recommendations

Sri Lanka is facing many changes on rice yield management for many years now due to lack of yield statistics before harvesting. Therefore, the intent of this study was to check the potential of applying remote sensing technology for rice yield estimation and propose a unique rice yield forecasting model for Sri Lanka based on freely available satellite data. Three-hundred ninety-one MOD09Q1 images, from 2007 to 2016 over Kurunegala district, have been used for this initial experiment. Based on the analysis, it has identified that some of the paddy fields in this study area are not being cultivated continuously in each season. Therefore, the cultivated paddy lands are essential to be identified before accurate yield estimation. A new cultivated paddy land identification method was suggested that has demonstrated an average overall accuracy of 77% on identifying cultivated paddy lands by coarse spatial resolution MODIS satellite data (250m). The use of high spatial resolution images could be one possibility to improve this accuracy of cultivated paddy land identification. However, satellite images with comparatively high spatial and reasonably higher temporal resolution, are not freely available at present scenario. In addition, it is recommended to improve the cultivated paddy land identification method to have a unique threshold value of correlation coefficients to be used for any season.

Linear and exponential yield forecasting models derived based on average NDVI, maximum NDVI, average EVI2 and cumulative NDVI have shown comparatively lower determinant of coefficient (R^2) value than smooth cumulative NDVI, EVI2 and NDVI based models derived at 80 days after the transplanting. Moreover, the accuracy of the model-based yield forecasting is evaluated by comparing the resulting yield with census data that was based on the conventional “crop-cutting survey” method. However, the exponential model derived based on NDVI values at 80 days after the transplanting has shown the best accuracy with only about 4.4% of discrepancy with census data. The 80 days exponential model was selected as the best model for rice yield forecasting using MODIS satellite imagery in this particular study area.

The proposed yield forecasting model has shown reasonable accuracy on rice yield forecasting approximately one month before harvest. Based on these pre-harvest yield estimations, the authorities could make effective decisions over rice yield management without any cost or field work to overcome the present national issues. Before generalizing this model for a Sri Lankan nation-wide national statistics application, it is recommended to test and validate it by implementing the same method for other parts of the country as well.

References

Cheng, Q., & Wu, X., 2010. Mapping paddy rice yield in Zhejiang Province using MODIS spectral index. Turk J Agric For, 35 579-589. DOI:10.3906/tar-1003-826.

DCS, Agriculture and Environment Statistics Division of the Department of Census and Statistics., 2017. Retrieved from http://www.statistics.gov.lk/agriculture/index.htm

Ferencz, C., Bognár, P., Lichtenberger, J., Hamar, D., Tarcsai, G., Timár, G., &Ferencz-Árkos, I., 2004. Crop yield estimation by satellite remote sensing. International Journal of Remote Sensing, 25, 4113-4149.

Gitelson, A.A., 2004. Wide Dynamic Range Vegetation Index for Remote Quantification of Biophysical Characteristics of Vegetation. Journal of Plant Physiology, 161, 165–173.

Guzinski, R.,2010. Comparison of vegetation indices to determine their accuracy in predicting spring phenology of Swedish ecosystems. Retrieved January 11, 2015, from http://www.natgeo.lu.se/ex-jobb/exj_193.pdf

Huete, A., Justice, C., &Leeuwen, V.W., (1999).MODIS Vegetation Index (MOD13), Algorithm Theoretical Basis Document. Available at http://modis.gsfc.nasa.gov/data/atbd/atbd_mod13.pdf, Accessed 15 November 2015

Huete, A., Didan, K., Miura,T., Rodriguez, E.P., Gao, X., &Ferreira, L.G., (2002) Overview of the radiometric and biophysical performance of the MODIS vegetation indices, Remote Sensing of Environment 83(1–2), 195-213.

Jiang, Z., Huete,A.R., Didan,K., &Miura, T., 2008. Development of a two-band enhanced vegetation index without a blue band.Remote Sensing of Environment, 112(10), 3833–3845.

Leblon, B., 1997. RSCC - Volume 4 - Module 9. Retrieved February 10, 2015, from http://rscc.umn.edu/rscc/Volume4/Leblon/leblon.htm

Le Toan, T., Ribbes, F., Wang, L., Floury, N., Ding, K., Kong, J., *et al*. 1997. Rice crop mapping and monitoring using ERS-1 data based on experiment and modeling results. IEEE Transactions on Geoscience and Remote Sensing, 1, 41–56.

Lio, Y., Sha, H., & Liao, L., 2011. Estimation of Rice Yield in Southern Taiwan Using MODIS Data. Retrieved January 12, 2015, from http://www.isprs.org/proceedings/2011/Gi4DM/PDF/PP15.pdf

Myneni, R. B., F. G. Hall, P. J. Sellers, and A. L. Marshak (1995). The interpretation of spectral vegetation indexes, in *IEEE Transactions on Geoscience and Remote Sensing* 33(2), 481-486.

Noureldin, N., Aboelghar, M., Saudy, H., & Ali, A., 2013. Rice yield forecasting models using satellite imagery in Egypt. The Egyptian Journal of Remote Sensing and Space Sciences 13, 125-131.

Pearson, R.L., & Miller, L.D., 1972. Remote mapping of standing crop biomass for estimation of the productivity of the short-grass Prairie, Pawnee National Grasslands, Colorado in Proc. of the 8th International Symposium on Remote Sensing of Environment, ERIM, Ann Arbor, MI, 1357-1381.

Rouse, J. W., Haas, R. H., Shell, J. A., Deering, D.W., & Harlan, J.C., 1974. Monitoring the vernal advancement of retrogradation of natural vegetation. Final Report, Type III, NASA/GSFC, Greenbelt, MD, 371 pp.

Shresthan, R.P., & Naikaset, S., 2003. Agro-spectral models for estimating dry season rice yield in the Bangkok Plain of Thailand. Asian Journal of Geoinformatics 4, 11–19.

Vermote, E.F., Kotchenova, S.Y., &Ray, J.P., 2011.MODIS Surface Reflectance User's Guide, Version 1.3.

Vermote, E.F., (2015). MOD09Q1 MODIS/Terra Surface Reflectance 8-Day L3 Global 250m SIN Grid V006. NASA EOSDIS Land Processes DAAC. Available at https://doi.org/10.5067/MODIS/MOD09Q1.006, Accessed 20 December 2015.

WayanNuarsa, I., Nishio, F., &Hongo, C., 2011. Rice Yield Estimation Using Landsat ETM Data and Field Observation. Journal of Agricultural Science 4

DCS, 2017, "Paddy Statistics". Department of Census and Statistics, http://www.statistics.gov.lk/agriculture/Paddy%20Statistics/PaddyStats.htm

Xiao, X., Boles, S., Liu, J., Zhuang, D., Frolking, S., Li, C., Salas, W., &Moore, B., 2005a. Mapping paddy rice agriculture in southern China using multi-temporal MODIS images. Remote Sens. Environ. 95, 480–492.

Xiao, X., Boles, S., Frolking, S., Li, C., Babu, J., Salas, W., & Moore, B., 2005b. Mapping paddy rice agriculture in South and Southeast Asia using multi-temporal MODIS images. Remote Sensing of Environment, 95-113.

Xiao, X., Boles, S., Frolking, S., Salas,W., Moore, B., Li, C., *et al*. 2002b. 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, 3009– 3022.

Xiao, X., Boles, S., Liu, J., Zhuang, D., & Liu, M., 2002c. Characterization of forest types in Northeastern China, using multi-temporal SPOT-4 VEGETATION sensor data. Remote Sensing of Environment, 82, 335–348.