CHEMICAL ENGINEERING TRANSACTIONS

VOL. 106, 2023

A publication of AIDIC

The Italian Association of Chemical Engineering
Online at www.cetjournal.it

Guest Editors: Jeng Shiun Lim, Nor Alafiza Yunus, Peck Loo Kiew, Hon Huin Chin

ISBN 979-12-81206-05-2; **ISSN** 2283-9216

DOI: 10.3303/CET23106139

Rice Yield Estimations Based on Transformed Surface Reflectance from Orbital Hyperspectral Remote Sensing

Poonsak Miphokasap

Department of Geography, Faculty of Social Science, Kasetsart University. 50 Ngam Wong Wan Rd., Ladyaow, Chatuchak, Bangkok, Thailand
fsocpsm@ku.ac.th

The estimation of rice yield using remote sensing (RS) technology is a critical task for sustainable agriculture. RS has been widely applied to monitor the rice yield estimations in several weeks before harvesting. However, the spectral data from the common earth observation satellites with the low spectral data results in the loss of important spectral characterization data that affecting the accuracy of the evaluation of rice yields. The main objective of this research is to examine the potential of hyperspectral data to estimate the rice yield before harvesting. In addition, to remove the atmospheric noise, EO-1 Hyperion surface reflectance data was transformed into the First Derivative Spectrum (FDS). The results of the estimations of rice yield using high spectral resolution remote sensing data and Stepwise Multiple Linear Regression (SMLR) analysis revealed that FDS dataset is the best estimator variables, compared to using the surface reflectance data. The best model for estimating rice yield of specific varieties (RD-41 variety) had a coefficient of determination of 0.884 and Root Mean Square Error (RMSE) of 286.3 kg/ha. The selected FDS bands for the estimation model were centered at red edge and shortwave Infrared (SWIR) region of electromagnetic spectrum which showed the good correlation with the rice yields. Red edge is related to the amount of chlorophyll in the rice leaves and can be used to explain the variation in rice yield. The shortwave infrared is a wavelength region that can be used to analyse the health and water in plants because SWIR is defined as water absorption feature. It was found that the FDS reflectivity could explain the variability of rice yield more than the common surface reflectivity. The usefulness of transformed spectral information for the rice yield estimation provides a solution to improve the crop yield estimation method from hyperspectral remote sensing.

1. Introduction

Rice is one of the most important economic crops in several countries including Thailand. Rice is a staple food for approximately 50% of the World's population worldwide (Yang et al., 2008). Traditional agriculture without good field management is an important source of greenhouse gas emissions and leads to global warming (Sadenovaa et al., 2022). Global warming has a direct impact on food security. The estimation of rice yields before the pre-harvest time is very important, especially in areas with variations in weather conditions. Yield estimation model helps the government to plan production, manage the overall productivity and finally make decisions to reduce or control the cultivation activities that emit greenhouse gases, e.g., fertilizing, maintaining water levels in rice fields (Nuarsa et al., 2011). Based on the conventional techniques, estimation of rice yield focuses on the use of statistical data acquired through site visits and finally writing reports (Mostafa et al. 2015). Such reports are often subjective, costly, time-consuming.

Remote sensing (RS) is the process of measuring and tracking the reflectance and emissivity characteristics of any object on the surface using light intensity measuring equipment installed on satellites or aircraft. RS has been widely applied to monitor the growth of paddy rice and other crops, including the spread of diseases or pests and yield estimations. Evaluating rice yield using reflectance data from satellite images allows for accurate, reliable and time-saving yield estimation especially in the large farming areas with a non-destructive method.

However, the spectral data from the multispectral data satellites, e.g., Landsat, Spot usually captures a spectrum of light that is reflected from any object surface by dividing the light into several discrete bands in greater than 100 nm in width of electromagnetic spectrum. The restrictive properties mentioned above results

Paper Received: 28 September 2023; Revised: 24 October 2023; Accepted: 5 November 2023
Please cite this article as: Miphokasap P., 2023, Rice Yield Estimations Based on Transformed Surface Reflectance from Orbital Hyperspectral Remote Sensing, Chemical Engineering Transactions, 106, 829-834 DOI:10.3303/CET23106139

in the loss of important spectral characterization data that affecting the accuracy of the evaluation of crop characteristics and crop yields (Galvao et al. 2005, Miphokasap and Wannasiri, 2018). In contrast, hyperspectral data enables us to collect hundreds of reflectance data at continuous and narrow wavelengths usually less than 10 nm in width covering the range of electromagnetic spectrum. It has been proven that hyperspectral data are more useful than multispectral data in applications to evaluate physical characteristics and biochemistry of various crop types (Galvao et al. 2005, Miphokasap and Wannasiri, 2018).

This study aims to examine the potential of hyperspectral data to estimate the rice yield before harvesting. If the yield can be estimated accurately and in advance, it will help farmers in planning and managing their plantations, optimal management of fertilizer and water supply which leads to the sustainable use of natural resources and reduction in greenhouse gas emissions from the agricultural sector.

2. Study area

The study was conducted in the central lowlands of Thailand which covers some areas of Ayutthaya, Saraburi and Lopburi provinces. Topographical features It is a flood plain. Most of the area is rice fields. There are no mountains and forests. Most of the area is rice growing area. There are only three popular rice cultivars; RD41, RD57 and SuphanBuri-1 with a total planting area of 1,208.36 ha.

3. Research Methodology

Details of data collection, preparation and analysis to develop a rice yield estimation model were illustrated and explained in Figure 1.

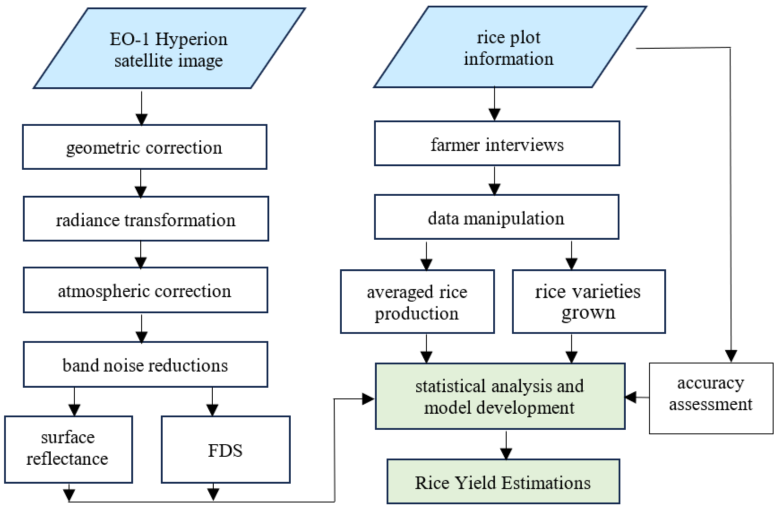

Figure 1: Research Methodology Framework

3.1 Data Collection

EO-1 Hyperion orbital image of 140/45 path/row recorded on 27 January 2017 was downloaded. This is the time when the rice is fully grown before the flowering stage. EO-1 Hyperion measured the reflectivity in 242 continuous bands covering 400–2,500 nm. The scene centre was at 14.3304° N and 100.5296° E. The relevant data on 1,750 rice cultivation plots within the study area that were grown during the image acquisition were requested from the government agencies. The data received includes the plot center coordinates, rice varieties and average rice yield. The topography of the study area is a flood plain and is a small study area. Based on the assumption that rice growing areas within the study area have similar physical environments, weather conditions, and access to water sources. Therefore, 225 samples of rice plot data were randomly selected for use in creating the training and testing data set.

3.2 Image Pre-Processing

To increase the accuracy of EO-1 Hyperion data, image rectification was performed using ten clearly visible Ground Control Points (GCP). The precision in XY coordinate must be less than 3 m. Subsequently, radiance transformation was performed to obtain the radiance values in the unit of W/m^2/µm/sr, the Hyperion calibration coefficients were applied to rescale the digital number (DN) value of EO-1 image particularly in the visible, near infrared and shortwave infrared region of electromagnetic spectrum. Radiance was calculated using constant value by divided DN with 40 for visible-near infrared (VNIR) and 80 for shortwave infrared (SWIR) wavelength. Afterward, the FLAASH algorithm was applied to reduce the aerosol effects, water content and gases. In this step, radiance measured at the atmosphere was converted to surface reflectance using MODTRAN4. Band noise reduction is the last step for image pre-processing. The total 44 bands of Hyperion raw data were identified as zero value bands due to the atmospheric noise. In addition, Datt et al. (2003) found that bands 77 and 78 in the shortwave infrared region have a higher level of spectral noise. Finally, there are only 196 bands used in this experiment.

3.3 Spectral Transformation

In this study, First-Derivative Spectrum (FDS) which is widely used to enhance absorption wavelengths that were interfered with the atmosphere effects (Dawson and Curran 1998). FDS was calculated and used as a independent variable to develop the yield estimation models. The usefulness of FDS technique is that it reduces the effects of light scattering due to the roughness and light geometry. However, FDS data are not calculated continuously throughout full wavelength (400-2,500 nm). This is because 46 bands of spectral data were removed due to interference in the atmosphere. The equation for FDS calculation was shown in Eq(1).

$$\mathrm{FDS} = (R_{\lambda(j+1)} - R_{\lambda(j)}) / \Delta\lambda \tag{1}$$

where FDS is the first-derivative transformation at a specific wavelength i midpoint between wavebands j and j + 1. $R_{\lambda(j)}$ is the reflectance at the j waveband, $R_{\lambda(j+1)}$ is the reflectance at the j + 1 waveband and $\Delta\lambda$ is the difference in wavelength between j and j + 1.

3.4 Statistical analysis and model validation

The relationship between actual rice yield (kg per ha) and reflectivity data were statistically analysed using the Stepwise Multiple Linear Regression (SMLR) method. SMLR is the step-by-step iterative construction of a regression model that involves the selection of independent variables to be used in a final model. It is an analysis of whether to add or remove independent variables from the proposed equation by testing statistical significance after each iteration. To verify the accuracy of the proposed models, the actual rice yield data per area unit data were compared with the yield data deriving from the estimation model as shown in RMSE and in scatter plot. In addition, coefficients of determination (R2) which is a measurement used to explain how much the variability of one factor is caused by its relationship to another factor will be calculated. This correlation is represented as a value between 0.0 and 1.0. The equation for RMSE calculation was shown in Eq(2).

$$\mathrm{RMSE} = \frac{\sqrt{\sum_{i=1}^{n}(y_i' - y_i)^{2}}}{n} \tag{2}$$

where $\hat{y}_i$ and y_i are the estimated and measured crop variables, respectively, and n is the number of samples.

4. Results and Discussion

The results of the study were presented in two parts including the Development of the rice yield estimation models for all varieties and model for each rice variety. The effectiveness of the independent variables is tested and evaluated in explaining the variance of the dependent variable (yield value).

4.1 The distribution of rice yield data

The rice yields obtained from the government agency ranged from 1,822.92 to 9,722.22 kg/ha with a mean of 4,624.41 kg/ha and median of 4,761.9 kg/ha. Testing for normality by the two tailed Kolmogorov-Smirnov test a p-value of 0.35 which confirms the hypothesis that the yield data is normally distributed, therefore parametric analysis techniques can be employed for further analysis without fulfilling any transformation requirements.

4.2 Development of the rice yield estimation models for all varieties

SMLR analysis and determination coefficient were performed based on the combination of all rice varieties. The proposed models using FDS dataset showed the precision higher than models using surface reflectance dataset illustrated in Table 1. According to lists of proposed estimation model, it can be observed that a distinctive relation between rice yield and spectral data centred in green, near infrared or shortwave infrared region of spectrum wavelength. In the case of using surface reflectance data as an independent variable, there were two bands selected for use in rice yield model when comparing with the case of using FDS as a variable, it was found that there were five bands selected. It is consistent with the research conclusion that using more than one variable for rice yield prediction increased the efficiency of the accuracy of the generated models due to enhancing R^2 values (Noureldin et al., 2013). However, the proposed model based on the two variable types shows the quite low accuracy. This is because of the effects of radiation scattering due to the difference in the structure of the rice canopy, which can be classified into planophile and erectophile. This is consistent with Galvao et al (2005) which summarizes the important findings that light incident and reflectance value from the plant canopy. Thus, it is necessary to develop separate models for estimating yield values for each rice variety. Figure 2 depicts the scatter plots of measured yield versus estimated yield for all rice varieties.

Table 1: Lists of proposed models for estimating rice yield including all varieties

variable	estimation model	R^2	RMSE (kg/ha)
surface reflectance	$y = -0.043x_1 + 0.043x_2 + 792.817$ [a]	0.224	494.53
first-derivative spectrum	$y = -4.156x_1 - 2.857x_2 + 9.626x_3 + 0.326x_4 + 7.414x_5 + 797.197$ b[b]	0.343	466.33

[a] referred to the *center wavelength* at band 164 (SWIR) *and* 134 (SWIR) for X_1 and X_2 respectively
[b] referred to the *center wavelength* at band 159 (SWIR), 138 (SWIR), 17 (GREEN), 79 (SWIR), and 47 (NIR) for X_1 and X_5 respectively

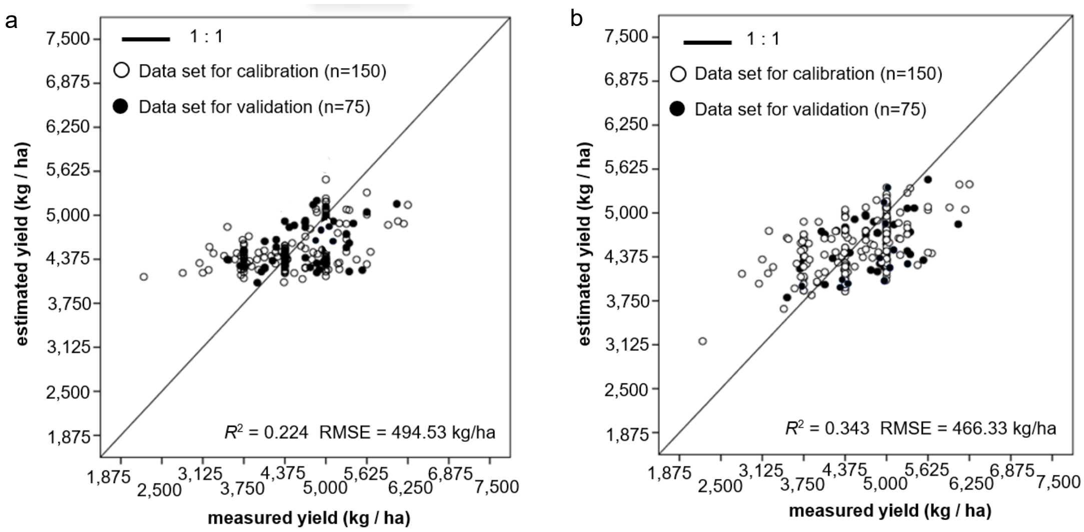

Figure 2: Measured versus estimated rice yield for all varieties; the model was developed (a) using surface reflectance as variables and (b) using first-derivative spectrum

4.3 Development of the rice yield estimation models for each rice variety

This is because the structure of the rice canopy is different in each variety. Models for estimating rice yields were developed separately for each rice variety. When analysing the relationship between the reflectance value and the average yield value for each rice variety. It was found that the proposed model could be able to significantly increase the accuracy of estimations illustrated in Table 2.
The top three of yield estimation model were the model of RD-41, RD-57 and SuphanBuri-1 with R^2 values of 0.884, 0.869 and 0.804 and with RMSE value of 286.3 kg/ha, 301.84 kg/ha and 317.23 kg/ha, respectively. The results show that the accuracy in estimating rice yield of the proposed models using surface reflectance data as a variable is lower than that using FDS data. This is because the surface reflectance data still contains the effects of radiation scattering due to geometry, surface roughness and atmosphere effects.
Most of the FDS wavelengths were selected as variables in the best estimation model for each rice variety were centered at the region of shortwave infrared and red edge. Reflectance from shortwave infrared could be used to detect moisture content or the water distribution within the plant leaf and canopy. Because water absorbs infrared light at a wavelength range from 1,450-1,500 nm. The amount of water in the rice canopy is therefore directly related to the growth rate, rice health and rice yields. Red edge region refers to the region of rapid change in reflectance of vegetation in the near infrared range of the electromagnetic spectrum usually covering 680-720 nm. Thus, red edge is related to the amount of chlorophyll in leaf and canopy. The transformation of the reflectivity data to FDS is a method that helps to emphasize the wavelength range and eliminate interference due to reflections.

Table 2: Lists of proposed models for estimating each rice variety

Rice variety	Variable	Estimation model	R^2	RMSE (kg/ha)
RD-41	surface reflectance	$y = -0.334x_1 + 0.141x_2 + 1204.433^{\mathbf{a}}$	0.537	398.39
	first-derivative spectrum	$y = -26.698x_1 + 1.745x_2 + 5.551x_3 + 0.751x_4 - 18.414x_5 + 1369.247^{\mathbf{b}}$	0.884	286.30
RD-57	surface reflectance	$y = -1.565x_1 - 0.019x_2 + 729.248^{\mathbf{c}}$	0.475	415.97
	first-derivative spectrum	$y = 22.928x_1 + 0.206x_2 + 5.485x_3 + 17.575x_4 + 7.760x_5 + 601.922^{\mathbf{d}}$	0.869	301.84
SuphanBuri-1	surface reflectance	$y = 0.482x_1 - 0.289x_2 + 958.336^{\mathbf{e}}$	0.424	515.33
	first-derivative spectrum	$y = 18.946x_1 - 11.152x_2 - 1.319x_3 - 1.044x_4 - 7.798x_5 + 791.207^{\mathbf{f}}$	0.804	317.23

[a] referred to the *center wavelength* at band 37 (Red edge) *and* 216 (SWIR) for X_1 and X_2 respectively

[b] referred to the *center wavelength* at band 33 (Red edge), 192 (SWIR), 104 (SWIR), 79 (SWIR), and 147 (SWIR) for X_1 and X_5 respectively

[c] referred to the *center wavelength* at band 15 (Blue) *and* 98 (SWIR) for X_1 and X_2 respectively

[d] referred to the *center wavelength* at band 154 (SWIR), 98 (SWIR), 104 (SWIR), 44 (NIR), and 135 (SWIR) for X_1 and X_5 respectively

[e] referred to the *center wavelength* at band 188 (SWIR) *and* 36 (NIR) for X_1 and X_2 respectively

[f] referred to the *center wavelength* at band 31 (Red), 85 (SWIR), 21 (Green), 18 (Green), and 206 (SWIR) for X_1 and X_5 respectively

Figure 3 depicts the scatter plots of measured yield versus estimated yield for separated variety. The estimation models were developed using surface reflectance as variables showed in (a) and using first-derivative spectrum in (b). It can be observed that the sample points of both were clustered and trends to close to the line of equality (1:1). It shows that the average rice yield per unit area has a strong relationship with the reflectance value above the rice canopy. Reflectance from shortwave infrared could be used to detect moisture content or the water distribution within the plant leaf and canopy. Because water absorbs infrared light at a wavelength range from 1,450-1,500 nm. Red edge region has been indicated to be more sensitive to water content, chlorophyll and nitrogen in leaf and canopy and can be used to explain the variation in the rice yields. Figure 4 illustrated the optimal estimation model method using FDS was applied to map the spatial distribution of rice yield for RD-41 in the cropland. The rice yield map obtained by applying the estimation equations to satellite image was displayed in colored shades ranging from 1,250 to 9,375 kg / ha. The boundaries of rice fields were used to exclude areas that were not rice fields.

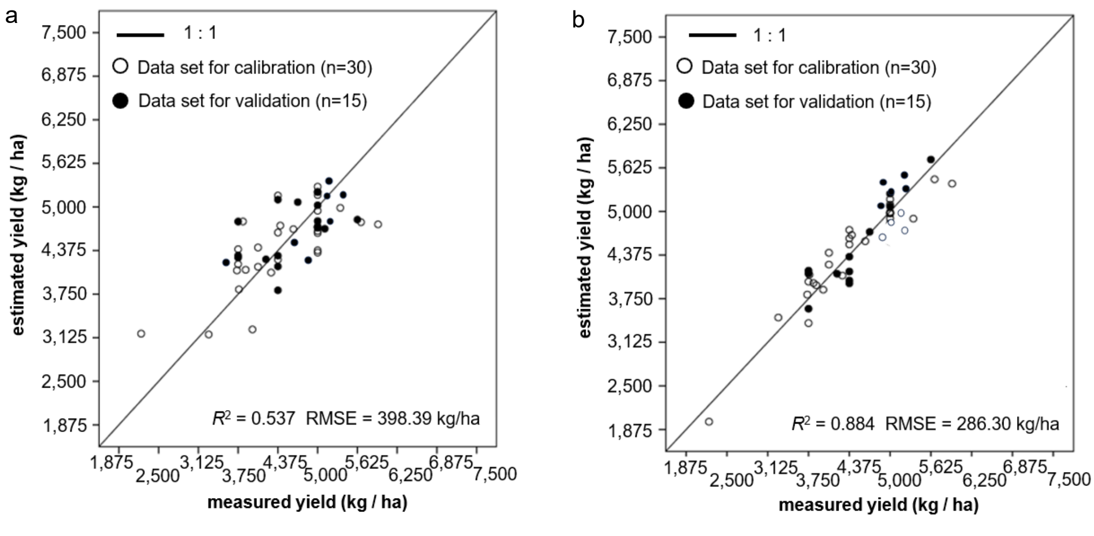

Figure 3: Measured versus estimated rice yield for RD-41variety; the model was developed (a) using surface reflectance as variables and (b) using first-derivative spectrum

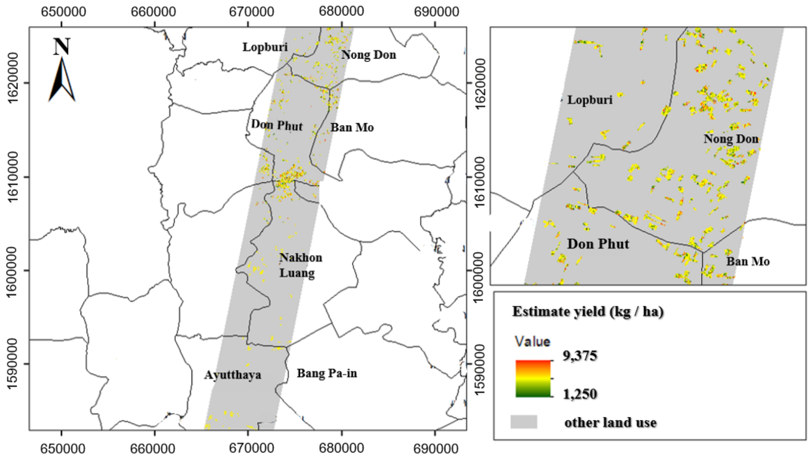

Figure 4: Spatial distribution of rice yield estimation (kg/ha) for RD-41 varieties using FDS dataset as variables

5. Conclusions

This experiment was an effort to estimate pre-harvest yield using orbital hyperspectral data. The benefits of hyperspectral data with the continuous and narrow wavelength of the electromagnetic spectrum allow researcher to develop the sustainable rice yield estimation model with the high accurate enough to be used in practice**.** It is clearly noticeable the use of FDS as variables in the model can significantly increase the accuracy of rice yield estimates, compared to the use of surface reflectance as variables.

However, it can be concluded that development of rice estimation models in the study area must be developed by separating each rice variety. This is because of the effects of radiation scattering due to the difference of rice canopy structure. Most of the selected wavelengths are centred in the shortwave infrared regions and longer portion of red edge (680–780 nm) in the electromagnetic spectrum. However, the yield estimation models derived from the study which is defined as empirical model, were needed further to calibrate and validate over other geographical location and different agro-climate zones. The usefulness of transformed spectral information to FDS for the rice yield estimation provides a solution to improve the crop yield estimation method from orbital hyperspectral imagery.

Acknowledgments

The authors would like to thank Department of Geography and Bachelor of Science Program in Geography (Special Program), Faculty of Social Science, Kasetsart university to support the budget for doing this research.

References

Datt B., McVicar T.R., Van Niel T.G., Jupp D.L.B., Pearlman J.S., 2003, Preprocessing EO-1 Hyperion hyperspectral data support the Application of agricultural indexes, IEEE Transactions on Geoscience and Remote Sensing, 41, 1246-1259.

Dawson T.P. and Curran P.J., 1998, A new technique for interpolating the reflectance red edge position, International Journal of Remote Sensing, 19, 2133-2139.

Galvao L.S., Formaggio A.R., Tisot D.A., 2005, Discrimination of Sugarcane Varieties in Southeastern Brazil with EO-1 Hyperion data, Remote Sensing of Environment, 94, 523–534.

Miphokasap P. and Wannasiri W., 2018, Estimations of Nitrogen Concentration in Sugarcane Using Hyperspectral Imagery, Sustainability, 10(4),1266.

Mostafa K.M., Quazi K.H. and Ehsan H.C., 2015, Application of remote sensor in mapping rice area and forecasting its production: A review, Sensors, 15, 769-791.

Noureldin J., Aboelghar M.A., Saudy H.S. and Ali A.M., 2013, Rice yield forecasting models using satellite imagery in Egypt, The Egyptian Journal of Remote Sensing and Space Sciences**,** 6(1), 125-131.

Nuarsa I.W., Nishio F., Hongo C., 2011, Spectral characteristics and mapping of rice plants using multi-temporal Landsat data, The Journal of Agricultural Science, 3, 54-67.

Sadenova M.A., Beisekenov N.A., Varbanov P.S., Kulenova N.A., Abitaev F., Kamenev Y., 2022, Digitalization of Crop Production for Transition to Climate-optimized Agriculture Using Spring Wheat in East Kazakhstan as an Example, Chemical Engineering Transactions, 96, 415-420.

Yang S., Shen S., Li B., Le Toan T., He W., 2008, Rice mapping and monitoring using ENVISAT ASAR data, IEEE Geoscience and Remote Sensing Letters, 5, 108-112.