Estimation of Leaf Area Index from UAV Multispectral Indices and Machine Learning Models

Asniyani Nur Haidar Abdullah (1), Farrah Melissa Muharam(1), Zed Diyana Zulkafli(2), Khairudin Nurulhuda (2), Muhamad Faiz Che Hashim(2) and Siti Najja Mohd Zad (2), Mohd Razi Ismail(3)

[1]Faculty of Agriculture, Universiti Putra Malaysia, 43400, UPM Serdang, Selangor Darul Ehsan, Malaysia
[2]Faculty of Engineering, Universiti Putra Malaysia, 43400, UPM Serdang, Selangor Darul Ehsan, Malaysia
[3]Institute of Tropical Agriculture and Food Security, Universiti Putra Malaysia, 43400, UPM Serdang, Selangor Darul Ehsan, Malaysia
Email:asniyaninh.ab@gmail.com, farrahm@edu.com.my, zeddiyana@upm.edu.my, k_nurulhuda@upm.edu.my, muhamadfaiz.ch@gmail.com, SitiNajja.MohdZad@gmail.com, razi@upm.edu.my

Keywords: germination stage, end of effective tillering, panicle initial stage

Abstract: Rice (Oryza sativa L) is the most important staple crop in Malaysia. Malaysian government has taken full concerns to maximize the quality of rice yield while minimizing resources utilization and adverse impact on the environment. Currently, rice producer and agencies still practice conventional management and monitoring tools such as field visit, interview, rice check guideline and ground based sensors. Ground data monitoring is extremely tedious, time-consuming, costly, inconsistent, labour intensive and prone to significant discrepancies. Consequently, delayed monitoring may postpone the next treatment to improve rice growth performance. Hence, this research aims to assess the potential of combining multispectral unmanned aerial vehicle (UAV) and data mining approach as a platform to estimate and monitor leaf area index (LAI). This study was conducted at a farmer's field in IADA KETARA, Lubuk Kawah, Jerteh Terengganu (5.717497oN, 102.492691oE) for the off-season from February to June 2018, involving five different rice cultivars i.e. MR269, MR297, UPUTRA, MR219 and MR220CL and four N rates 76 kg ha^{-1},108 kg ha^{-1} and 141 kg ha^{-1}. Physiological data presenting the growth performances of the crop i.e. LAI as well as multispectral images were acquired three times throughout the season; tillering, booting and milking stages. Red, blue, green, red edge and near infrared based indices were derived in order to determine the best vegetative index to represent growth performance. Consequently, two models i.e. Gradient Boosting Regressor (GBR) and Decision Tree (DT) were applied to explore the relationship between UAV images and LAI. The DT model yielded the highest stability and accuracy for all growth stages for LAI. Besides, combination of single NIR band performed best for estimating LAI at all stages. Given these results, the proposed approach could provide a new opportunity for the integration of UAV-multispectral and data mining in providing useful information for decision making on crop growth performance.

1. Introduction

Leaf area index (LAI) is an important biophysical variable in rice development. Specifically, LAI is a major indicator determining stable net primary crop production, nutrient and utilization water, improvement on photosynthesis and carbon balance (Zheng, 2009). Conventionally, scanners are the standard tools in monitoring LAI, whereby the leaf area is measured subsequently after the harvesting procedure. Despite the high accuracy of traditional data monitoring, this method is extremely tedious, time-consuming, costly, labour intensive and also prone to significant discrepancies (Hirooka et al., 2018). Hence, the delayed in monitoring may postpone the next agronomic practice such as fertilizer application which will reduce the rice growth performance.

As an alternative to the destructive method, remote sensing has been introduced as a tool to monitor and estimate LAI in simple, reliable, real-time and non-destructive manners (Shibayama et al., 2011; Yuan et al., 2017). Yao et al. (2017) had tested several vegetation indices such as modified triangular vegetation index (MTVI2), range vegetation index (RVI), green-blue normalized differential vegetation index (GBNDVI), renormalized difference vegetation index (RDVI), NDVI, and SAVI to estimate LAI of wheat using a UAV equipped with multispectral sensors. They found that the MTVI2 produced the highest coefficient of determination (R2) of 0.80 for predicting various value of LAI. In another study, Silva et al. (2018) examined two vegetative indices i.e. soil adjusted vegetation index (SAVI) and normalized differential vegetation index (NDVI) for predicting the LAI of mangrove tree from UAV and RapidEyes images. The authors reported that the LAI prediction using SAVI-UAV and SAVI-RapidEyes obtained a moderate R^2 of 0.58 and 0.70, respectively. A better relationship was observed for the RapidEyes given its lower spatial resolution compared to the UAV, and therefore the later had a greater influence of the background which influenced the R^2.

Accurate predictions remain as the main bottleneck in remote sensing application for predicting LAI based on multispectral reflectance bands. Models utilizing spectral information often adopts high non-linear modeling, yet the available methods are still unstable. Previous studies as conducted by Yuan et al. (2017) have compared the partial least square (PLS) regression, random forest (RF), artificial neural network (ANN), and support vector machine (SVM) in predicting LAI. The results from the study showed that the RF in combination with stratified sampling (STR) produced the highest R^2, and lowest standard deviation (SDR^2), V-RMSE and SDRMSE. Despite that, the RF model performed well to estimate LAI over the whole single growth rather than single growth stage, which was a contrast to the findings found by Upreti et al. (2019). The least squares linear regression (LSLR) and PLS regression were more appropriate to estimate LAI at single stage (R^2=0.78). However, previous studies presented the contrast results on the accuracy of the machine learning (ML) techniques.

In this light of reviews, this research aims to assess the potential of combining a multispectral sensor onboard a UAV and data mining approach as a platform to estimate LAI of rice.

2. Materials and Methods

2.1 Experimental Sites and Designs

This study was conducted at a farmer's field in IADA KETARA, Lubuk Kawah, Jerteh Terengganu (5.717497N, 102.492691E) from February to June 2018. N fertilizer rates were applied as the following: 76, 108, and 141 kg ha^{-1}, and henceforth will be referred to as N1, N2, and N3, respectively. The 108 kg ha^{-1} rate is the recommended rate from the Department of Agriculture, and 30% of the N input was reduced or added to categorize formulate the N1 and N3, respectively. Other nutrients were applied as standard agronomic practices. The applications of N fertilizer were made in three split application according to development of vegetative phase: 39% at the germination stage (18 days after seeding (DAT)), 42% at the end of effective tillering (39 DAT), and 19% at the panicle initiation stage (55 DAT). Five cultivars, MR216, MR220CL2, MR219, MR297, UPUTRA were seeded with 2 kg in each N plots between 17 to 21 February 2018. The water was kept constantly flooded between 5 to 10 cm until 15 days before harvesting.

2.2 Field Data Collection

2.2.1 Ground Data

LAI were measured at three different growth stages: stage 1 (tillering), stage 2 (booting) and stage 3 (milking) and limited to 30 samples per day. Two quadrants were used to select randomly a 0.5 m x 0.5 m square area and the coordinates of these quadrant were recorded using a Trimble R8 RTK. The plant above ground were uprooted manually and cut for each quadrant. Later, all set of samples were packed and stored in an iced cooler and transported to Universiti Putra Malaysia (UPM). Then, green leaf blades from all 30 set of samples were detached manually and immediately scanned for leaf area using LI-3100C Area Meter. The leaf area index (LAI) was calculated using Equation 1 as follows:

$$LAI=LA/Q \tag{1}$$

Where, LA is the leaf area and Q is the area of quadrant (cm2).

2.2.2 Spectral Data

Multispectral images were acquired simultaneously with the ground data collected at two hours before and after local solar noon. The flying altitude was set up approximately 50 m above the local terrain and resulted in a ground sample distance (GSD) of 3 cm. Side and front lap were approximately 75%. A Micasense Red-Edge multispectral camera acquiring images in five narrow bands with a gimbal was mounted on a Quadcopter drone. The Pix4DMapper Pro (version 4.0, Pix4D Lausanne, Switzerland) was used to align the five separated images. The pre-processing of images was conducted as follows; initial processing, point cloud and mesh, DSM and orthomosaic. Further, for radiometric correction, each band was calibrated with known reflectance values from the calibration panel as provided from MicaSense. The pixel values of the quadrants were finally extracted accordingly.

2.3 Data Pre-processing

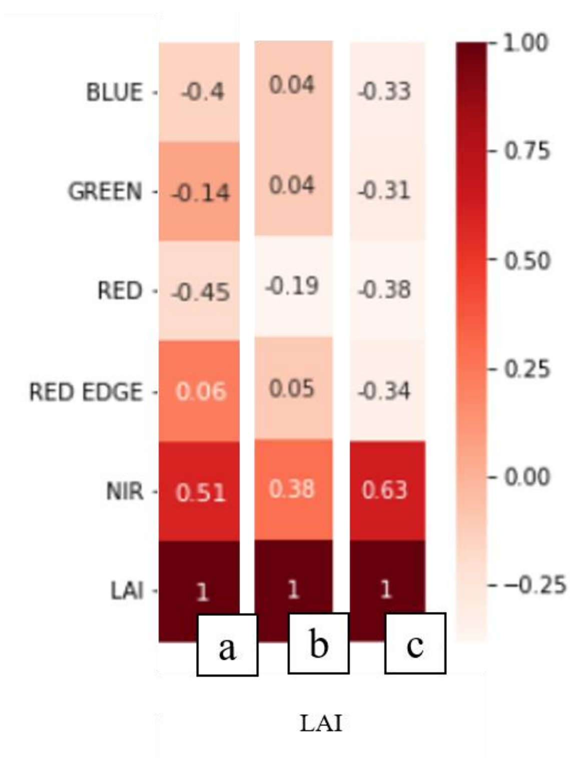
Scikit-learn packages in Python language was used to interpret and test the Pearson coefficient (*r*) between the reflectance at five different wavelengths and for each pair with LAI. Additionally, data were tested by individual sampling dates rather than entire season. Only bands with correlation values at *r*> 0.35 were used for further non-parametric regression analysis. Calibration set was set up 70% for training sample and 30% for validation set (Wang et al., 2018) with 120 data for every single growth period. Further, the DT and GBR models were applied to estimate LAI, also by using the Scikit-learn packages whereby with parameters of *max_depth*, *max_feature*, *min_sample_split* and *min_sample_leaf* were adjusted simultaneously using RandomizedSearchCV model selection (Yuan et al., 2017). The performance and stability for each model was measured by root mean squared error (RMSE) and coefficient of determination (R^2) (Govaerts et al., 1999).

3. Results and Discussion

3.1 Analysis of Spectral Parameter for Each Band

Pearson's correlation analysis showed both significant negative and positive correlation between five spectral parameters i.e. blue (B), green (G), red (R), red edge (RE) and near infrared (NIR) with LAI. At stage 1, a strong negative correlation was found between B and R with *r* = -0.45 and -0.4, respectively while NIR exhibited a strong positive correlation (0.51) with LAI. For stage 2, NIR spectral band depicted higher *r* value i.e. 0.38, as compared to the rest of the bands. B, G, R, and RE also demonstrated high correlation with LAI, however, *r* value of NIR was the highest at *r*= 0.63 for stage 3. Taken together, these results indicated that LAI has a higher association with certain spectral bands such as NIR compared to the rest of spectral bands.

Figure 1: Pearson Coefficient (*r*) of GL Biomass with Multispectral UAV Bands :(a) Stage 1, (b) Stage 2 and (c) Stage 3.



3.2 Appropriate Regression Model for LAI Growth Model

Table 1 illustrates the performance of spectral models in estimating LAI. Regardless of models, the single NIR model was the best predictor in estimating LAI of stage 1 (R^2 = 0.66-0.80) and 2 (R^2=0.61-0.80). Correspondingly, Hatfield et al. (2008) discovered that NIR bands had a good correlation in estimating LAI across the growing seasons of winter wheat. Contrary to stage 1 and 2, the appearance of panicle may lead to the changes of canopy reflectance causing R bands to obtain the highest value of R^2 ranging from 0.62-0.70. This is confirmed by findings reported by He et al. (2019), that the R band is the most affected by panicles, while the RE band are less affected in estimated LAI.

At stage 1, regardless of models, combine spectral model were generally better than single band. The R value of B+R+NIR resulted in the highest R2 value ranging from 0.72-0.77. Likewise, the combination of single band between NIR and R bands at stage 3 also depicted high values of R^2 with ranging from 0.57-0.72. Combination of bands was found to result in better relationships between spectral measurement and crop variables as reported by Zhou et al. (2007), Din et al. (2017) and Wang et al. (2018).

In comparing the performance of models, the DT model performed best at stage 1 and 3 with R^2 values ranging from 0.43-0.84 (RMSE = 0.07-0.12) and 0.53-0.70 (RMSE = 0.11-0.14), respectively, regardless of spectral bands. The DT model remarkably had more stability and precision with lower RMSE and consistently produced higher R^2 value for the calibration and validation dataset, which was a contrast for the GBR model. However, the GBR model was most appropriate for LAI estimation for a single band at stage 2.

For the combination of single band, the GBR model exhibited the best performance at stage 1 with R^2 and RMSE values of 0.58 -0.77 and 0.08-0.13, respectively. In the case of small sample size, the GBR model was more stable due to the use of sample feature dimensions to predict the results while the DT model was based on the splitting of the predictor space with respect to the target data. The stronger the interpretation of the sample, the stronger the generalization ability of the trained data (Yuan et al., 2017). In contrast, the DT model performance slightly improved at stage 3 with R^2 and RMSE values of 0.62-0.72 and 0.10-0.14. This is due to the model parameterization that distinctly improved the adjusted parameter towards stronger learning and more generalization of the training data in order to produce better accuracy during the validation phase.

Table 1: Different model performances spectral index for rice LAI estimation.

Stages	Spectral Index	Calibration (n=42)				Validation (n=18)			
		DT		GBR		DT		GBR	
		RMSE	R^2	RMSE	R^2	RMSE	R^2	RMSE	R^2
Stage 1	**B**	**0.09**	**0.84**	0.11	0.75	**0.10**	**0.60**	0.10	0.55
	R	**0.12**	**0.68**	0.13	0.63	**0.12**	**0.43**	0.12	0.41
	NIR	**0.11**	**0.72**	0.13	0.66	**0.07**	**0.80**	0.08	0.69
	B+R	0.12	0.71	**0.10**	**0.77**	**0.09**	**0.65**	0.10	0.58
	B+R+NIR	0.11	0.73	**0.11**	**0.76**	**0.08**	**0.72**	0.08	0.70
Stage 2	**NIR**	0.15	0.61	**0.08**	**0.80**	**0.11**	**0.75**	0.11	0.66
Stage 3	**R**	**0.13**	**0.70**	0.14	0.62	**0.11**	**0.67**	0.12	0.64
	NIR	**0.14**	**0.67**	0.14	0.63	**0.13**	**0.53**	0.14	0.48
	R+NIR	**0.14**	**0.62**	0.15	0.57	**0.10**	**0.72**	0.11	0.67

4. Conclusion

The aim of the present research was to assess the potential of combining multispectral images acquired using a UAV platform and data mining approach as a platform to estimate LAI of rice. This study found that the DT model in combination have the high potential to estimate these crop variables. Further research should be undertaken to investigate highly dispersed data and to utilize vegetation indices as spectral variables.

Acknowledgements

This research was funded by the Higher Institution Centers of Excellence (HICoE) research grant (HICoE-ITAFoS/2017/FC6, vote no 6569105).

References

Chen, J., L, Hui., H, C. & Fang, C, 2009. The relationship between the leaf area index (LAI) of rice and the C-band SAR vertical/horizontal (VV/ HH) polarization ratio. International Journal of Remote Sensing, pp. 30(8), 2149–2154.

Govaerts, Y.M., Verstraete, M.M., Pinty, B. & Gobron, N., 1999. Designing optimal spectral indices: A feasibility and proof of concept study. International Journal of Remote Sensing, 20(9), pp.1853-1873.

Hatfield, J.L., Gitelson, A.A., Schepers, J.S. & Walthall, C.L., 2008. Application of spectral remote sensing for agronomic decisions. Agronomy Journal, 100(Supplement_3), pp. S-117.

He, J., Zhang, N., Su, X., Lu, J., Yao, X., Cheng, T., Zhu, Y. & Cao, W., 2019. Estimating Leaf Area Index with a New Vegetation Index Considering the Influence of Rice Panicles. Remote Sensing, 11(15), pp.1809.

Hirooka, Y., Homma, K. & Shiraiwa, T., 2018. Parameterization of the vertical distribution of leaf area index (LAI) in rice (*Oryza sativa* L.) using a plant canopy analyser. Scientific Report, 8, pp. 6387.

Muharam, F. M., Delahunty, T. & Maas, S. J., 2018. Evaluation of nitrogen treatment effects on the reflectance of cotton at different spatial scales. International Journal of Remote Sensing, 39(23), pp. 8482-8504.

Shibayama, M., S., Sakamoto, T., Takada, E., Inoue, A., Morita, K., Takahashi, W. & Kimura, A., 2011. Estimating paddy rice leaf area index with fixed point continuous observation of near infrared reflectance using a calibrated digital camera. Plant Production Science, 14(1), pp. 30-46.

Shibayama, M., Sakamoto, T., Takada, E., Inoue, A., Morita, K., Takahashi, W. & Kimura, A., 2009. Continuous monitoring of visible and near-infrared band reflectance from a rice paddy for determining nitrogen uptake using digital cameras. Plant Production Science, 12(3), pp. 293-306.

Silva, E.D.O., Xavier, A.C., de Souza Tedesco, A.N., Neto, A.A.B., de Lima, L.E.M., Pezzopane, J.E.M. & Tognella, M.M.P., 2018. Estimates of the leaf area index using unmanned aerial vehicle images of an urban mangrove in the vitória bay, brazil.

Upreti, D., Huang, W., Kong, W., Pascucci, S., Pignatti, S., Zhou, X., Ye, H. & Casa, R., 2019. A Comparison of Hybrid Machine Learning Algorithms for the Retrieval of Wheat Biophysical Variables from Sentinel-2. Remote Sensing, 11(5), pp. 481.

Wang, L., Chang, Q., Yang, J., Zhang, X. & Li. F., 2018. Estimation of paddy rice leaf area index using machine learning methods based on hyperspectral data from multi-year experiments. PLoS ONE, 13(12), pp. 1-16.

Yao, X., Wang, N., Liu, Y., Cheng, T., Tian, Y., Chen, Q. & Zhu, Y., 2017. Estimation of wheat LAI at middle to high levels using unmanned aerial vehicle narrowband multispectral imagery. Remote Sensing, 9(12), pp. 1304.

Yuan, H., Yang, G., Li, C., Wang, Y., Liu, J., Yu, H., Feng, H., Xu, B., Zhao, X. & Yang, X., 2017. Retrieving soybean leaf area index from unmanned aerial vehicle hyperspectral remote sensing: Analysis of RF, ANN, and SVM regression models. Remote Sensing, 9(4), pp. 309.

Zhao, D., Huang, L., Li, J. & Qi, J., 2007. A comparative analysis of broadband and narrowband derived vegetation indices in predicting LAI and CCD of a cotton canopy. ISPRS Journal of Photogrammetry and Remote Sensing, 62(1), pp.25-33.

Zheng, H., Cheng, T., Li, D., Zhou, X., Yao, X., Tian, Y., & Zhu, Y., 2018. Evaluation of RGB, color-infrared and multispectral images acquired from unmanned aerial systems for the estimation of nitrogen accumulation in rice. Remote Sensing, 10(6), pp. 824.

Zheng, G., & Moskal, L. M., 2009. Retrieving leaf area index (LAI) using remote sensing: theories, methods and sensors. Sensors, 9(4), pp. 2719-2745.