

# A Research of Yield Estimation on Jiang Xi Early Rice Using MODIS Land Product

Liushui Yan, Xuehui Gao  
Nanchang Hang Kong University  
Nanchang, China

Shu-e Huang  
Meteorological Research Institute of Jiangxi Province  
Nanchang, China

**Abstract**— The conventional multiband data need a lot of processing before application in rainy area. MODIS land products synthesis algorithm reduces the influence of cloud, aerosol and other factors in the greatest degree. MODIS vegetation index product was directly used to estimate total yield production of early rice in Jiangxi Province. The summation of pixel value of Enhanced vegetation index in each administrative division was taken as the metric. Correlation analysis of the total production of early rice versus metrics shows significantly related relationship. 7 estimation models were built based on regression equations. One optimal model was selected with an integrated consideration of their root mean square error and relative error in prefecture-level estimation. The applicability of the model was validated by province-level estimation test using the data of 2001 to 2004. The optimal model indicated that MODIS land product can be used in yield estimation and the enhanced vegetation index during the tillering and jointing stage were most appropriate for the estimation of the total rice production in Jiangxi Province. Provincial production estimation result is better than that of the prefecture-level estimation.

**Index Terms**— MODIS data, Double early rice, Yield estimation.

## I. INTRODUCTION

Since the 1980s, satellite remote sensing has become an important means in crop monitoring research [1]. In the 1990s, China began to concern the processing and information extraction technology of the MODIS (Moderate Resolution Imaging Spectrometer) data. [2]. In recent years, research based on MODIS data has developed widely in many science fields, such as the water environment, atmosphere, agricultural monitoring and so on. Since it is provided with moderate spatial resolution and relatively high temporal frequency as well as many kinds of products which are freely available from the National Aeronautics and Space Administration website, MODIS data has a large advantage in crop growth monitoring and yield estimation.

Study of the relationship between the spectral characteristics with rice growth and yield components is the basis of rice yield estimation. Studies have shown that using the normalized difference vegetation index is an effective method to estimate crop yields [3-6]. However, with the rice

growing to heading stage the crop canopy is almost entirely covered by the rice. In this case, the normalized difference vegetation index is easy to reach saturation [7]. As for the introduction of the blue band, Enhanced vegetation index (EVI) has effectively reduced the impact of the air, soil, aerosols, and avoided high biomass saturation. Comparing with NDVI, EVI is better in the establishment of a linear relationship with leaf area index, biomass and photo-synthetically active radiation component and more effective in crop monitoring and yield estimation with a superior forecast accuracy [8]. As the rice yield of double-crop rice in the area of the Yangtze River basin is difficult to dynamically estimate accurately, the objective of this study is to carry out a research of early rice yield estimation in Jiangxi province with an application of the enhanced vegetation index (EVI) product, in the aim of finding an more macro, dynamically and accurate method.

## II. MATERIALS AND METHODS

### A. Study area and data resources

#### 1) Study area

Jiang xi province is located in the southeast of china. It lies in the southern bank of the Yangtze River's lower and middle sections. Jiang Xi has a sub-tropical rainy climate, warm and humid. Jiangxi has 11 administrative divisions. It's main crop is rice. The average grain output was 19.2827 million tons, and the average rice production was 18.3511 million tons in 2005 to 2010, during which rice production accounted for more than 94% of the total grain output. 8 administrative divisions were chosen as the study area including Ganzhou, Fuzhou, Ji'an, Nanchang, Shangrao, Xinyu, Yichun and Yingtan.

#### 2) Data resources

##### a) Remote sensing data

MOD13Q1 data is one product of MODIS land products which has 12 layers including Enhanced vegetation index (EVI) data. MOD13Q1 data used in the study was downloaded from the NASA (the National Aeronautics and Space Administration, NASA) website. MODIS level 1 data was from the Jiangxi Institute of Meteorological Sciences digital satellite broadcasting system (DVB-S).

##### b) Official statistics

Historical yield production data and sown area during 2002 to 2010 was downloaded from Statistic Bureau of Jiangxi website.

##### c) rice growth stage data

Early rice growth period data in 2005-2010 comes from 12 meteorological observation stations of Jiangxi province. The rice growth process phenophase include: the seedling stage, the three-leaf stage, transplanting period, turning green stage, tillering, jointing stage, booting stage, heading stage, milky stage and maturity. The early rice growth period start from late March to late July in Jiangxi province, the whole growth period lasts 110 days to 130 days.

### B. Data processing

#### 1) Remote sensing images interpretation

MODIS level 1 data from the DVB-S was used to extract rice cultivation information using the Supervised classification method. Before the interpretation, the MODIS level 1 data was processed into Albers Equal Area projection and WGS84 datum in .img file type.

#### 2) Metrics calculation

As for the transplanting injure, rice often need some time to restore , and therefore this study chose the tillering, jointing, booting, heading, flowering and milky growth stages which last from early May to early July as the research time. According to MOD13Q1 synthesis time, the study selected four remote sensing images for every year whose synthetic time cover tillering stage to the milk stage. (See table 1).

Table1 Rice growth periods and corresponding MOD13Q1data in Jiang Xi Province

	Image No.			
	1	2	3	4
Acquire time (Month/day)	5/9-5/24	5/25-6/9	6/10-6/25	6/26-7/13
Growth stage	Tillering	Tillering Jointing	Booting Heading	Heading Milky

EVI data was extracted from the MOD13Q1 layers with the MODIS Reprojection Tools. The sum of pixel value over each administrative divisions was calculated. The calculation result was taken as the metrics for yield estimation. As there were four images each year, there would be four groups data for every administrative division each year. We named them as X<sub>1</sub>, X<sub>2</sub>, X<sub>3</sub>, and X<sub>4</sub> according to the time sequence.

## III. RESULTS AND ANALYSIS

### A. Correlation analysis

The metrics data of 2005-2009 was used to build regression models to estimate the yield production of 2010 , so there were 40 samples in the six years. Accordingly, there were 40 yield production samples. Correlation analysis between the metrics X and yield production Y shows that the metrics is significantly related to the total yield production.

Table 2 The Relationship of the Metrics with Total production

Metric name	X <sub>1</sub>	X <sub>2</sub>	X <sub>3</sub>	X <sub>4</sub>
Correlation coefficients	0.957**	0.951**	0.935**	0.97**

\*\* Significant at 0.01 level

As shown in Table 2, the correlation coefficients range from 0.957 to 0.97, significant at 0.01 level. The correlation coefficient of Y versus X<sub>4</sub> is the highest among the four metrics. X<sub>4</sub> data was acquired from Issue 177 images and correspond with the heading and milky growth period. However, the biggest of correlation coefficient value doesn't indicate the best capacity in estimation of the early rice. The result of correlation analysis was used to develop regression equations of Y versus X. Instead of correlation coefficients, the study chose the root mean square error (Root Mean Square Error: RMSE) as well as the relative error as the indicator to select the optimal fitting model. The smaller the value of indicator is, the better the model is in estimation.

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (y_i - \tilde{y}_i)^2}{n}}$$

Where  $y_i$  is the statistical total production,  $\tilde{y}_i$  is the estimated total production, n indicates the number of samples

### B. Regression model

#### 1) Single variable models

The four group data derived form the four EVI imges cover different growth stage. As the influence of the four stages is different, single variable models were built to analysis their contribution to the final yield output. According to the R square, the sample data fitted well. The model 4 which established with the variable X<sub>4</sub> has the best fitting result but the highest value of RMSE. The RMSE value of model 1 is smaller than that of the other models. As the smaller value of RMSE indicates a better estimation effect, the model 1 is better in estimation than the other three models. Model 1 was established with X<sub>1</sub> which is derived from the EVI data in tillering stage. Tillering stage is the critical period for vegetative organs growing and decisive period to the panicles per unit area.

Table 3 Regression Equations of Single Variable Model

Model No	Regression equations	R <sup>2</sup>	RMSE
1	Y=2.48+0.009X <sub>1</sub>	0.909	6.13
2	Y=2.04+0.008X <sub>2</sub>	0.904	8.77
3	Y=5.52+0.008X <sub>3</sub>	0.875	8.80
4	Y=1.40+0.008X <sub>4</sub>	0.947	18.80

#### 2) Multivariate models

As different growth stage have different influence on the yield production, the combination of different stage may have a different result. The study combined the continuous growth stages to establish three multivariate models. The value of RMSE varied. Model 5 is better than the other two models

with a RMSE value of 5.60. Variables  $x_1$  and  $x_2$  were derived from the EVI images whose synthesis time is May 9 to June 9. According to the average growth stage time data of Jiangxi early rice, the period is tillering to jointing stage. The RMSE value of Model 5 is smaller than that of model 1 and model 2, meanwhile the RMSE value of model 6 is smaller than that of model 7, model 2 and model 3, which suggests that the combination improves the estimation effect. Variable  $x_4$  was derived at the milky stage during which rice leaves start to turn yellow and the leaf area index decreased rapidly, after the heading stage, photosynthesis declines and respiration strengthens. The importance of the metric derived in the period decreased, so model 4 and model 7 have a bigger RMSE value than the others. As the RMSE value of model 5 is the smallest one among the 7 models, it should be the optimal model which has a better estimation effect. It's relative error in prefecture-level was within the range of -5% to 5%. The relationship between the estimated value and statistics can be found in the Scatter plots whose points of distribution was close to the 1:1 reference line.

Table 4 Regression Equations of Multivariate Model

Model No.	Regression equations	R <sup>2</sup>	RMSE
5	$Y=1.411+0.005X_1+0.004X_2$	0.916	5.60
6	$Y=1.412+0.005X_1+0.004X_2+0.000018X_3$	0.916	7.99
7	$Y=1.728-0.002X_1+0.002X_2-0.002X_3+0.01X_4$	0.952	20.31

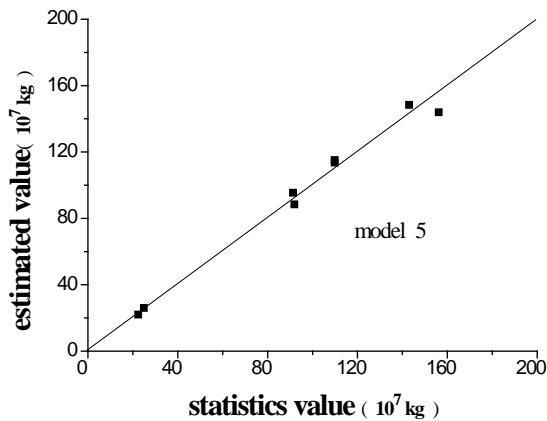


Figure 1 Relationship of Estimation Value with the Statistic Value

### 3) Model Validation

To vify the applicability of the modes, the study tested the models in province-level. The study took the data of 2001-2004 and 2010 to estimate their province yield. The relative error of model 5 is between -2% and 2% smaller than that of the other models. The province-level test show that model 5 is much better than other model. Optimal Model 5 is also valid for the provincial yield estimation. The result of provincial

yield estimation for the five years can be seen in table 5. It's relative error was within the range of 0.1% to 1.7%.

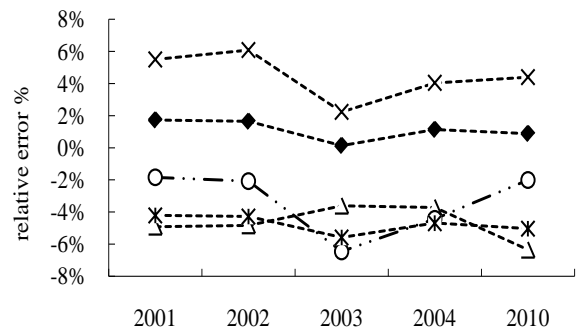


Figure 2 Relative Error for Provincial Estimation

---Δ--- model 1    —○— model 2    ---×--- model 3  
 ---◆--- model 5    ---\*--- model 6

## IV. CONCLUSION AND DISCUSSION

Comparing to the northern areas, the use of remote sensing technology for crop yield estimation is relatively less in the south areas. In one hand, the weather in Jiangxi is cloudy and rainy, optical data influenced by cloud can't be used directly for yield estimation. On the other hand, rice cultivation area distribution is very scattered in some area which makes remote estimation even harder. The study used the MOD13Q1 data directly to establish regression equations according the correlation analysis, which indicated that the sum of EVI pixel value was significantly correlated with the total yield production within the divisions. For better estimation, single rice growing stage and multiple growth stages regression equation were developed respectively. The applicability of the model was test in two ways in prefecture-level and province-level respectively. The results showed that the data in tillering stage and jointing stage were better fit for yield estimation. The combination of data in different growth stage could improve the capability of estimation and the MODIS data could be used directly for the yield estimation in Jiangxi province. The shortage is that this study did not consider the difference between the north and south areas of Jiangxi Province. Firstly, as rice cultivation area distribution scattered in some region, the MODIS data interpretation accuracy would be reduced, so we abandoned some prefecture-level cities; on the other hand, statistics based on administrative divisions units, so it is difficult to obtain historical statistics for other division. The related error is 0.8% for provincial estimation and within  $\pm 7\%$  for the prefecture-level the yield estimation of 2010 using the optimal model, which has certain guiding significance for the estimate of the total early rice production in Jiangxi Province.

#### ACKNOWLEDGMENT

This work is supported by the National Science Foundation of China (21165013, 51208248) and the Natural Science Foundation of Jiangxi Province, China (2010GZH0112; 20114BAB213015).

#### REFERENCES

- [1] Qingbo Zhou. Status and Tendency for Development in Remote Sensing of Agriculture Situation[J]. Journal of China Agricultural Resources Regional Planning, 2004, 25 (5):9 -14.
- [2] Yujie Liu, Zhongdong Yang. MODIS remote sensing information processing principles and algorithms [M].Beijing: Science Press, 2001.
- [3] Rui Deng, Jingfeng Huang, Fumin Wang, et al. Research on Yield Estimation of Rice Based on Remote Sensing Using Moderate Resolution Imaging Spectroradiometer (MODIS) Data: A Case Study of Jiangsu Province, China [J], China Journal of Rice Science. 2010(1): 87 - 92.
- [4] Haizhen Zhang, Zehong Ma, Zhiyue Zhou, et al. Research on the Estimation of Rice in Chengdu Based on MODIS data[J],Remote Sensing Information,2008,5:63 – 67.
- [5] Wenhuan Xie, Youzhi Zhang, and Li WU, Forecast Methods of Crop Production in Wangkui Country Based on the Normalized Difference Vegetation Index[J].Heilongjiang Agricultural Sciences, 2011(4): 113 – 115.
- [6] Junying Sun, Jinliang Huang, and Lihui Wang.Grain Yield Estimation for Hubei Province Using Sensing Data-with Medium Rice as an example[J],World Science Technology Research and Development, 2009,31(5): 897 – 900.
- [7] Zhengxing Wang, Chuang Liu, Huete A, et al. From AVHRR-NDVI to MODIS-EVI: Advances in vegetation index research [J] Acta Ecologica Sinica, 2003, 23(3): 979-987.
- [8] Wim J D, Alfredo R. MODIS Vegetation Index Composition Approach: A prototype with AVHRR Data [J]. Remote Sensing of Environment, 1999, 69: 264 – 280.