



# Rice Yield Estimation Using Parcel-Level Relative Spectral Variables From UAV-Based Hyperspectral Imagery

Feilong Wang<sup>1</sup>, Fumin Wang<sup>2,3\*</sup>, Yao Zhang<sup>2,3</sup>, Jinghui Hu<sup>2</sup>, Jingfeng Huang<sup>3,4</sup> and Jingkai Xie<sup>1</sup>

<sup>1</sup> Institute of Hydrology and Water Resources, Zhejiang University, Hangzhou, China, <sup>2</sup> Institute of Applied Remote Sensing & Information Technology, Zhejiang University, Hangzhou, China, <sup>3</sup> Key Laboratory of Agricultural Remote Sensing and Information System, Zhejiang University, Hangzhou, China, <sup>4</sup> Ministry of Education Key Laboratory of Environmental Remediation and Ecological Health, Zhejiang University, Hangzhou, China

Time-series Vegetation Indices (VIs) are usually used for estimating grain yield. However, multi-temporal VIs may be affected by different background, illumination, and atmospheric conditions, so the absolute differences among time-series VIs may include the effects induced from external conditions in addition to vegetation changes, which will pose a negative effect on the accuracy of crop yield estimation. Therefore, in this study, the parcel-based relative vegetation index ( $\Delta VI$ ) and the parcel-based relative yield are proposed and further used to estimate rice yield. Hyperspectral images at key growth stages, including tillering stage, jointing stage, booting stage, heading stage, filling stage, and ripening stage, as well as rice yield, were obtained with Rikola hyperspectral imager mounted on Unmanned Aerial Vehicle (UAV) in 2017 growing season. Three types of parcel-level relative vegetation indices, including Relative Normalized Difference Vegetation Index (RNDVI), Relative Ratio Vegetation Index (RRVI), and Relative Difference Vegetation Index (RDVI) are created by using all possible two-band combinations of discrete channels from 500 to 900 nm. The optimal VI type and its band combinations at different growth stages are identified for rice yield estimation. Furthermore, the optimal combinations of different growth stages for yield estimation are determined by F-test and validated using leave-one-out cross validation (LOOCV) method. The comparison results show that, for the single-growth-stage model, RNDVI<sub>[880,712]</sub> at booting stage has the best correlation with rice yield with a  $R^2$ -value of 0.75. For the multiple-growth-stage model, RNDVI<sub>[808,744]</sub> at jointing stage, RNDVI<sub>[880,712]</sub> at booting stage and RNDVI<sub>[808,744]</sub> at filling stage gain a higher  $R^2$ -value of 0.83 with the mean absolute percentage error of estimated rice yield of 3%. The study demonstrates that the proposed method with parcel-level relative vegetation indices and relative yield can achieve higher yield estimation accuracy because it can make full use of the advantage that remote sensing can monitor relative changes accurately. The new method will further enrich the technology system for crop yield estimation based on remotely sensed data.

Keywords: hyperspectral image, unmanned aerial vehicles, relative spectral variables, growth stages, rice yield estimation

#### OPEN ACCESS

#### Edited by:

Guijun Yang, Beijing Research Center for Information Technology in Agriculture, China

#### Reviewed by:

Changwei Tan, Yangzhou University, China Xia Yao, Nanjing Agricultural University, China

> \*Correspondence: Fumin Wang wfm@zju.edu.cn

#### Specialty section:

This article was submitted to Technical Advances in Plant Science, a section of the journal Frontiers in Plant Science

> Received: 29 October 2018 Accepted: 26 March 2019 Published: 10 April 2019

#### Citation:

Wang F, Wang F, Zhang Y, Hu J, Huang J and Xie J (2019) Rice Yield Estimation Using Parcel-Level Relative Spectral Variables From UAV-Based Hyperspectral Imagery. Front. Plant Sci. 10:453. doi: 10.3389/fpls.2019.00453

1

# INTRODUCTION

Remote sensing technology is an important measure for collecting data on the Earth and its changes, and it has been widely used in all kinds of subjects such as water resources (Schneider et al., 2018), geology (Govil et al., 2017), ecology (Echappé et al., 2018), and agriculture (Guo et al., 2018). Remote sensing for agriculture has the advantages of nondestructive, non-invasive, fast, and cost-efficient monitoring, well-correlated with agronomical and important physiological crop traits (Reynolds et al., 2015). It has usually been applied to monitor crop growth status (Poenaru et al., 2017), to map vegetation area (Xiao et al., 2005), to estimate crop yield (Peng et al., 2014), etc. In recent years, those issues such as the environmental degradation and water pollution have caused the reduction of arable land and grain production, grain security has become a tremendous challenge in many countries and regions. So accurate grain yield estimation is a significant means to ensure grain security and agriculture production.

Although satellite images have been widely used in crop yield estimation, there are still many problems that we need to address. For example, high spatial and temporal resolution satellite images usually cannot be obtained at same time, sometimes it is impossible to obtain valid data at the key growth stage due to poor weather conditions. Unmanned aerial vehicles (UAV) equipped with sensors can remedy those defects mentioned above (Zhou et al., 2017). UAV remote sensing is a low altitude remote sensing system, which can acquire high spatial-temporal resolution remotely sensed data on demand. It has been used for agriculture monitoring in sugarcane (Luna and Lobo, 2016), sunflower (Vega et al., 2015), soybean (Yu et al., 2016), and triticale (Noack, 2016), yield prediction in rice (Zhou et al., 2017), wheat (Du and Noguchi, 2017) and barley (Honkavaara et al., 2013). At present, remote sensing sensors mounted on UAV include RGB digital camera (Luna and Lobo, 2016; Du and Noguchi, 2017; Zhou et al., 2017), NIR camera (Nebiker et al., 2016) and multispectral camera (Vega et al., 2015; Yu et al., 2016; Noack, 2016; Zhou et al., 2017). In recent years, with the development of imaging hyperspectral technology, imaging hyperspectral cameras have gradually been equipped on UAV to acquire remotely sensed data combining image with spectra (Honkavaara et al., 2012, 2013; Yue et al., 2017). Imaging hyperspectral technology can obtain more spectral bands and precise spectral information, which is expected to further improve the monitoring accuracy. Hyperspectral images have been used for measuring individual parcel plots using ultra-high spatial resolution up to 1 cm per pixel (Turner et al., 2012), mapping high-precision leaf carotenoid concentration of vine in region scale (Zarcotejada et al., 2013), monitoring soybean LAI precisely by combining hyperspectral image with artificial neural network (ANN) (Yuan et al., 2017), early detection of olive verticillium using airborne hyperspectral and thermal imagery (Calderón et al., 2013).

Unmanned Aerial Vehicle platform equipped with remote sensing sensors can supply high spatial and temporal resolution images for precision agriculture, but there are still some issues that need be further investigated. For example, accurate radiometric correction for time-series images is a challenge for UAV remote sensing. The ideal condition for multi-period UAV data acquisition is that the weather is clear and the angle of the sun is the same for each flight. Actually, for each flight, the absorption and scattering of solar radiation by atmospheric molecules, dust and water vapor are different and the altitude angle of the sun is not the same due to the different flying time (Wang et al., 2003), which make the measured values of the sensor inconsistent with physical quantities, such as spectral reflectance or spectral radiance of the target (Wang et al., 2003; Roy et al., 2010). For example, the different compositions of water vapor, ozone, and aerosol in the atmosphere will affect the reflection of the Red band and the NIR near-infrared band (Teillet et al., 2000), which make radiometric corrections for time-series images more difficult. In addition, different soil backgrounds may also affect target reflections. Therefore, eliminating the interference of atmospheric and illumination conditions on remote sensing images and obtaining accurate canopy reflection data is the basis for predicting crop yield accurately. Radiometric calibration is an essential process for UAV remote sensing, in which the reflectance spectra are calculated by using reference white or gray board (Hall et al., 1991; Ding and Elvidge, 1996; Carvalho et al., 2013) and then vegetation indices at different growth stages are derived for yield estimation. However, due to the influence of atmospheric and illumination conditions in different dates, it is difficult to obtain absolutely accurate reflectance using Pseudo-Invariant Features (PIF) method (Carvalho et al., 2013) or standard reference calibration board method (Du et al., 2001), although these methods can mitigate these effects in some way. In fact, it is impossible to obtain absolutely accurate reflectance spectra, but it is easy to characterize the relative change of radiation variable for remote sensing technology. This is because the two variables used for calculating the relative changes are obtained under the same external conditions, such as atmospheric and illumination conditions, even background conditions, so the relative variable can eliminate the effects of these external conditions on target reflectance. Therefore, this study attempts to construct a parcel-level relative spectral value to weaken the negative effects from different external conditions, and estimate rice yield based parcel-level relative spectral values from multiple-growth-stage hyperspectral images. The purposes of this study are (1) to propose a crop yield estimation method based parcel-level relative spectral value, which can eliminate the effects of atmospheric and illumination conditions and rice field background on the target reflectance, (2) to make full use of the strength of hyperspectral image in band numbers to determine the optimal band combinations for rice yield estimation, (3) to determine the best combination of growth stages for rice yield estimation.

# MATERIALS AND METHODS

# **Field Experiment**

The field experiment was conducted in May–November 2017 in Deqing County, Zhejiang Province, China  $(120^{\circ}10'49.34''E, 30^{\circ}34'21.21''N)$  (**Figure 1**). The average annual precipitation





and temperature in study area were 1379 mm and  $15^{\circ}$ C, respectively (China Meteorological Data, 2018). Two rice cultivars Japonica Rice [Jia-58 (V1) and Zhe-99 (V2)] were sown on 20 May, transplanted on 22 June, and harvested on 24 November. Twenty-two parcel plots (1–22), with five different nitrogen fertilizer level including no fertilizer (N0), half of normal fertilization (N1), normal fertilization (N2), 1.5 times normal fertilization (N3), and two times normal fertilization (N4), were set (**Figure 2**). Field data were collected in July–October in 2017 for six times from tillering stage (28 July), jointing stage (23 August), booting stage (3 October), heading stage (19 September), filling stage (3 October), to ripening stage (24 October).



FIGURE 3 | UAV-based hyperspectral platform.

# UAV-Based Hyperspectral Images Acquisition and Data Processing

Rikola Hyperspectral Imager (Rikola Ltd., Oulu, Finland) mounted on the DJI Matrice 600 Pro UAV (**Figure 3**) are used to obtain hyperspectral images at different growth stages. The UAV maximum flight altitude is 2.5 km and payload weight is 6 kg. The Rikola hyperspectral imager can capture two-dimensional hyperspectral images ranging from 500 to 900 nm with 62 bands in total, and the full width at the half maximum of the spectral band is 8 nm. Hyperspectral sensor exposure times were set ranging from 10 to 15 ms depending on sunlight conditions. Hyperspectral images were obtained continuously and saved to memory cards when UAV was flying. For each band, a 1010 by 1010 pixels image with 12 bit (4096 DN) is created. All pixels are true image pixels, no interpolation is used. Flight altitude was set to 200 m and flight route was fixed for each UAV flight campaign with flying time of 10 min. UAV-based hyperspectral images were obtained at rice tillering stage, jointing stage, booting stage, heading stage, filling stage, and ripening stage from 28 July to 24 October (**Table 1**).

Every hyperspectral images were processed with vignetting correction, lens distortion correction, black level calibration and band to band registration. DN values of raw images were transformed to radiance values. Except for band to band registration, all data processing was carried out with Rikola Hyperspectral Imager software (Rikola Ltd., Oulu, Finland).

# Parcel-Level Relative VI and Relative Yield

As mentioned above, the use of relative vegetation index is expected to diminish the limitation of absolute differences that contain the uncertain information caused by different background, illumination and atmospheric conditions at different growth stages. The idea of parcel-level relative vegetation index is based on the hypothesis that solar radiation, solar altitude and atmospheric conditions are the same, and background are similar at one time of data acquisition. So, the one parcel of rice field is selected as reference parcel and others are seen as studied parcels. The parcel-level relative vegetation index and relative yield is calculated on the basis of vegetation index and yield of reference parcel. In our study, a well-grown parcel with normal fertilization is selected as a reference parcel from all experimental parcels. Relative vegetation index is calculated by vegetation index of studied parcel divided by vegetation index of the reference parcel [Eq. (1)], and relative yield is calculated by yield of studied parcel divided by yield of reference parcel [Eq. (2)]. Schematic plot of the calculation method is shown in **Figure 4**.

In our study, three types of relative vegetation indices are tested, including Relative Normalized Difference Vegetation Index (RNDVI), Relative Ratio Vegetation Index (RRVI), and Relative Difference Vegetation Index (RDVI). All possible two-band combinations of discrete channels from 500 to 900 nm are used to create RNDVI, RRVI, RDVI, as demonstrated by Le Maire et al. (2008) [28], and linear regressions between those combinations and relative yields are performed to determine the correlation coefficient (r). All the r-values are plotted in matrix plots. The advantages of the matrix plots are that they give a quick overview of



hundreds of wavelength combinations and make it possible to identify the optimal band combination for further analysis.

$$\Delta VI = \frac{VI}{VI_R}, (1) \tag{1}$$

Where  $\Delta VI$  is parcel-level relative vegetation index, VI is the vegetation index of a study parcel,  $VI_R$  is the vegetation index of reference parcel.

$$_{RY=\frac{Y}{Y_{D}}},(2) \tag{2}$$

Where RY is parcel-level relative yield, Y is the measured yield of a study parcel, and  $Y_R$  is the measured yield of reference parcel.

# **Yield Estimation Model**

Since the good linear relationship between VI and crop yield, like rice and wheat, have been found in the previous studies (Tucker et al., 1980; Siyal et al., 2015), multiple linear regression method will be used to construct the rice yield estimation models in this study. The independent variables are depended on analysis results of the optimal band combinations above, and dependent variable is relative yield.

The relative VIs at different growth stages are all related with relative yield, but which single growth stage and growth stage combinations that can more accurately estimate rice yield will be determined through this study. Therefore, firstly, the optimal single-growth-stage model can be directly obtained by comparing the fitted model with different single growth stage. Secondly, the multiple-growth-stage model can be acquired by comparing the fitted model with two, three, four, five, and all combinations of growth stages. Coefficient of determination ( $R^2$ ) and root mean square error (RMSE) between estimated and measured yield are calculated to quantify the performance of estimation models.

## RESULTS

# Band Selection for Relative VI and Determine the Optimal Relative VI Type

The illustrations of *r* between the vegetation indices of different growth stages and relative yield are shown in **Figure 5**. As seen for all growth stages, band combinations provided by wavelength 1 at red edge spectral region combined with wavelength 2 in near-infrared area almost occupy high *r*-values areas. The locations of wavelength 1 and wavelength 2 of the optimal band combinations with the highest *r*-value for three types index and for every growth stage are shown in **Table 1**. It can be seen that the red edge bands from 712 to 744 nm and the near-infrared regions from 808 to 888 nm occur in all three vegetation types. These results indicate the great potential of red edge band and near-infrared regions for the estimation of rice yield.

The comparison of the optimal band combinations at every growth stage for all three types index, i.e., RNDVI type index, RRVI type index, and RDVI type index, are shown in **Figure 5**. As illustrated, compared with RRVI and RDVI type index, RNDVI type index is the most efficient type (**Figure 6**). The optimal relative vegetation indices of every growth stage are all RNDVI type index. In addition, the correlation coefficients show an increased tendency from tillering stage to booting stage and then decrease to ripening stage. The highest correlation coefficients between optimal relative vegetation index and relative yield for all three type index are always obtained at booting stage, with *r* of 0.87 for RNDVI<sub>[880,712]</sub>, 0.85 for RRVI<sub>[776,744]</sub>, and 0.84 for RDVI<sub>[840,744]</sub>. It can be concluded that booting stage is the most efficient growth stage for rice yield estimation.

# Yield Estimation Model With Different Growth Stages Involved

Rice yield is related to the relative spectral variables at every growth stages. As concluded from above results, the estimation efficiency of RNDVI type index is the most effective compared with RRVI type index and RDVI type index, so in the process of model construction, only RNDVIs are considered. In order to identify the number of involved RNDVIs and the involved growth stages in the optimal yield estimation model, the univariate models with one single growth stage and the multivariate models including two, three, four, five, and all growth stages are tested using  $R^2$  and RMSE.

#### Yield Estimation Model With One Single Growth Stage Involved

Yield estimation model with one single growth stage involved is the simplest yield estimation model, which has the advantages of simplicity and easy to calculate. The  $R^2$  and RMSE of all estimation models with one single RNDVI at each growth stages are shown in **Table 2**.

As seen, compared with models using RNDVIs at other growth stages, the model using RNDVI at booting stage gains the highest  $R^2$  of 0.75 and the lowest RMSE of 228.04 kg/ha. So the booting stage is regarded as the optimal growth stage for rice yield estimation. The model can be expressed as:

 $Y_E = (-0.60 + 1.55 \times \text{RNDVI}_{[880,712] (Booting)}) \times Y_R$ , (3)

where  $Y_E$  is estimated Yield,  $Y_R$  is reference yield.

#### Yield Estimation Model With Two Growth Stage Involved

The models including two RNDVIs of any two growth stages are shown in **Table 3**.

As seen, the best estimation model with  $R^2$  of 0.80 and RMSE of 205.74 kg/ha, include booting stage (RNDVI<sub>[880,712]</sub>) and filling stage (RNDVI<sub>[808,744]</sub>). The expression of model is:

$$Y_E = (-1.45 + 2.65 \times \text{RNDVI}_{[880,712](\text{Booting})} - 0.26 \times \text{RNDVI}_{[808,744](\text{Filling})}) \times Y_R, \quad (4)$$

### Yield Estimation Model With Three Growth Stages Involved

The model expressions and  $R^2$  and RMSE of all three-variate models are shown in **Table 4**. As seen, the highest  $R^2$  of 0.83





**TABLE 2 |** The performances for the models with one single growth stage involved.

Grow stage	Model Expression	R <sup>2</sup>	RMSE(kg/ha)
Tillering stage (TS)	$y = -0.13 + 1.06x_1$	0.60	285.89
Jointing stage (JS)	$y = 0.45 + 0.50x_2$	0.70	250.35
Booting stage (BS)	$y = -0.60 + 1.55x_3$	0.75	228.04
Heading stage (HS)	$y = 0.19 + 0.77x_4$	0.66	270.02
Filling stage (FS)	$y = 0.67 + 0.30x_5$	0.56	304.10
Ripening stage (RS)	$y = 0.68 + 0.29x_6$	0.52	318.39

 $x_1$  is RNDVI<sub>[720,880]</sub> at tillering stage,  $x_2$  is RNDVI<sub>[744,808]</sub> at jointing stage,  $x_3$  is RNDVI<sub>[712,880]</sub> at booting stage,  $x_4$  is RNDVI<sub>[736,888]</sub> at heading stage,  $x_5$  is RNDVI<sub>[744,808]</sub> at filling stage,  $x_6$  is RNDVI<sub>[744,872]</sub> at ripening stage.

and the lowest RMSE of 189.13 kg/ha are obtained by the model including jointing stage (RNDVI<sub>[808,744]</sub>), booting stage (RNDVI<sub>[808,744]</sub>) and filling stage (RNDVI<sub>[808,744]</sub>). The model expressed as Eq. (5):

$$Y_E = (-1.06 + 0.20 \times \text{RNDVI}_{[808,744]}(\text{Jointing})$$

$$+2.04 \times \text{RNDVI}_{[880,712]}$$
 (Booting)

$$-0.23 \times \text{RNDVI}_{[808,744](\text{Filling})} \times Y_{\text{R}}$$
 (5)

# Yield Estimation Model With Four Growth Stages Involved

As shown in **Table 5**, among the models that include four RNDVIs, the highest  $R^2$  (0.84) and the lowest RMSE (185.84 kg/ha) is obtained by the model incorporating jointing stage (RNDVI<sub>[808,744]</sub>), booting stage (RNDVI<sub>[872,744]</sub>), filling stage (RNDVI<sub>[808,744]</sub>), and ripening stage (RNDVI<sub>[872,744]</sub>), the **TABLE 3** | Expressions and performance for the model with two

 growth stage involved.

Combination grow stage	Model expression	R <sup>2</sup>	RMSE(kg/ha)
TS, JS	$y = 0.18 + 0.39x_1 + 0.36x_2$	0.73	239.39
TS, BS	$y = -0.59 + 0.06x_1 + 1.48x_3$	0.76	227.84
TS, HS	$y = 0.0026 + 0.43x_1 + 0.50x_4$	0.68	261.48
TS, FS	$y = 0.087 + 0.74x_1 + 0.12x_5$	0.64	276.90
TS, RS	$y = -0.0051 + 0.87x_1 + 0.073x_6$	0.62	282.30
JS, BS	$y = -0.27 + 0.22x_2 + 0.99x_3$	0.79	208.02
JS, HS	$y = 0.29 + 0.33x_2 + 0.34x_4$	0.74	234.23
JS, FS	$y = 0.47 + 0.38x_2 + 0.10x_5$	0.73	238.51
JS, RS	$y = 0.46 + 0.44x_2 + 0.053x_6$	0.71	247.69
BS, HS	$y = -0.89 + 2.21x_3 - 0.36x_4$	0.77	223.15
BS, FS	$y = -1.45 + 2.65x_3 - 0.26x_5$	0.80	205.74
BS, RS	$y = -0.99 + 2.06x_3 - 0.12x_6$	0.77	218.87
HS, HS	$y = 0.046 + 1.00x_4 - 0.10x_5$	0.66	267.60
HS, RS	$y = -0.020 + 1.12x_4 - 0.16x_6$	0.67	263.20
FS, RS	$y = 0.67 + 0.25x_5 + 0.053x_6$	0.57	303.43

model expression is:

$$Y_E = (-1.02 + 0.23 \times \text{RNDVI}_{[808,744](\text{Jointing})} + 1.97$$

$$\times \text{RNDVI}_{[880,712](\text{Booting})} - 0.14 \times \text{RNDVI}_{[808,744](\text{Filling})}$$

$$-0.10 \times \text{RNDVI}_{[872,744](\text{ripening})}) \times Y_R$$
(6)

# Yield Estimation Model With Five Growth Stages Involved

As seen in **Table 6**, for multivariate models that include five RNDVIs, the best estimation model is the last one, with  $R^2$ 

TABLE 4 | Expressions and performance for the model with three growth stage involved.

Combination grow stage	Model expression	R <sup>2</sup>	RMSE(kg/ha)
TS, JS, BS	$y = -0.26 - 0.15x_1 + 0.24x_2 + 1.13x_3$	0.80	206.86
TS, JS, HS	$y = 0.21 + 0.16x_1 + 0.30x_2 + 0.27x_4$	0.74	233.02
TS, JS, FS	$y = 0.31 + 0.23x_1 + 0.34x_2 + 0.07x_5$	0.73	236.16
TS, JS, RS	$y = 0.16 + 0.42x_1 + 0.37x_2 - 0.014x_6$	0.73	239.26
TS, BS, HS	$y = -0.90 + 0.13x_1 + 2.11x_3 - 0.39x_4$	0.77	222.28
TS, BS, FS	$y = -1.45 + 0.13x_1 + 2.52x_3 - 0.27x_5$	0.80	204.76
TS, BS, RS	$y = -1.02 + 0.21x_1 + 1.90x_3 - 0.15x_6$	0.78	216.85
TS, HS, FS	$y = -0.17 + 0.46x_1 + 0.77x_4 - 0.13x_5$	0.69	257.95
TS, HS, RS	$y = -0.28 + 0.51x_1 + 0.89x_4 - 0.20x_6$	0.70	251.23
TS, HS, RS	$y = 0.063 + 0.77x_1 + 0.16x_5 - 0.051x_6$	0.64	276.27
JS, BS, HS	$y = -0.66 + 0.26x_2 + 1.90x_3 - 0.55x_4$	0.82	195.85
JS, BS, FS	$y = -1.06 + 0.20x_2 + 2.04x_3 - 0.23x_5$	0.83	189.13
JS, BS, RS	$y = -0.74 + 0.27x_2 + 1.59x_3 - 0.18x_6$	0.83	189.75
JS, HS, FS	$y = 0.26 + 0.32x_2 + 0.38x_4 - 0.019x_5$	0.74	234.14
JS, HS, RS	$y = 0.059 + 0.33x_2 + 0.73x_4 - 0.18x_6$	0.76	224.21
JS, FS, RS	$y = 0.45 + 0.43x_2 + 0.23x_5 - 0.15x_6$	0.75	232.17
BS, HS, FS	$y = -1.44 + 2.61x_3 + 0.037x_4 - 0.27x_5$	0.80	205.70
BS, HS, RS	$y = -1.01 + 2.14x_3 - 0.06x_4 - 0.12x_6$	0.77	218.78
BS, FS, RS	$y = -1.45 + 2.65x_3 - 0.26x_5 - 0.0028x_6$	0.80	205.74
HS, FS, RS	$y = -0.05 + 1.18x_4 - 0.040x_5 - 0.15x_6$	0.67	262.91

TABLE 5 | Expressions and performance for the model with four growth stage involved.

Combination grow stage	Model expression	R <sup>2</sup>	RMSE(kg/ha)
TS, JS, BS, HS	$y = -0.65 - 0.09x_1 + 0.27x_2 + 1.96x_3 - 0.54x_4$	0.82	195.47
TS, JS, BS, FS	$y = -1.05 - 0.06x_1 + 0.20x_2 + 2.08x_3 - 0.23x_5$	0.83	188.94
TS, JS, BS, RS	$y = -0.73 - 0.01x_1 + 0.27x_2 + 1.59x_3 - 0.18x_6$	0.83	189.74
TS, JS, HS, FS	$y = 0.16 + 0.18x_1 + 0.30x_2 + 0.34x_4 - 0.03x_5$	0.74	232.76
TS, JS, HS, RS	$y = -0.07 + 0.24x_1 + 0.30x_2 + 0.65x_4 - 0.20x_6$	0.77	221.50
TS, JS, FS, RS	$y = 0.25 + 0.28x_1 + 0.38x_2 + 0.19x_5 - 0.17x_6$	0.75	228.47
TS, BS, HS, FS	$y = -1.45 + 0.13x_1 + 2.52x_3 + 0.006x_4 - 0.27x_5$	0.80	204.76
TS, BS, HS, RS	$y = -1.04 + 0.21x_1 + 1.98x_3 - 0.07x_4 - 0.13x_6$	0.78	216.75
TS, BS, FS, RS	$y = -1.46 + 0.15x_1 + 2.52x_3 - 0.25x_5 - 0.02x_6$	0.80	204.63
TS, HS, FS, RS	$y = -0.32 + 0.52x_1 + 0.96x_4 - 0.05x_5 - 0.18x_6$	0.70	250.80
JS, BS, HS, FS	$y = -1.08 + 0.22x_2 + 2.23x_3 - 0.25x_4 - 0.19x_5$	0.83	187.37
JS, BS, HS, RS	$y = -0.79 + 0.27x_2 + 1.81x_3 - 0.21x_4 - 0.14x_6$	0.83	188.71
JS, BS, FS, RS	$y = -1.02 + 0.23x_2 + 1.97x_3 - 0.14x_5 - 0.10x_6$	0.84	185.84
JS, HS, FS, RS	$y = 0.13 + 0.35x_2 + 0.583x_4 + 0.09x_5 - 0.21x_6$	0.77	222.73
BS, HS, FS, RS	$y = -1.44 + 2.60x_3 + 0.06x_4 - 0.26x_5 - 0.01x_6$	0.80	205.67

TABLE 6 | Expressions and performance for the model with five growth stage involved.

Combination grow stage	Model expression	R <sup>2</sup>	RMSE(kg/ha)
TS, JS, BS, HS, FS	$y = -1.06 - 0.05x_1 + 0.23x_2 + 2.26x_3 - 0.24x_4 - 0.18x_5$	0.83	187.25
TS, JS, BS, HS, RS	$y = -0.79 - 0.01x_1 + 0.28x_2 + 1.82x_3 - 0.21x_4 - 0.14x_6$	0.83	188.70
TS, JS, BS, FS, RS	$y = -1.02 - 0.01x_1 + 0.23x_2 + 1.98x_3 - 0.14x_5 - 0.10x_6$	0.84	185.83
TS, JS, HS, FS, RS	$y = 0.004 + 0.22x_1 + 0.32x_2 + 0.54x_4 + 0.07x_5 - 0.22x_6$	0.77	220.48
TS, BS, HS, FS, RS	$y = -1.45 + 0.15x_1 + 2.45x_3 + 0.05x_4 - 0.25x_5 - 0.03x_6$	0.80	204.58
JS, BS, HS, FS, RS	$y = -1.04 + 0.24x_2 + 2.08x_3 - 0.13x_4 - 0.13x_5 - 0.08x_6$	0.84	185.46

of 0.84 and RMSE of 185.46. The included growth stages are jointing stage (RNDVI<sub>[808,744]</sub>), booting stage (RNDVI<sub>[880,712]</sub>), heading stage (RNDVI<sub>[888,736]</sub>), filling stage (RNDVI<sub>[808,744]</sub>), and ripening stage (RNDVI<sub>[872,744]</sub>). The model is expressed as Eq. (7):

 $Y_E = (-1.04 + 0.24 \times RNDVI_{[808,744](Jointing)} + 2.08 \times RNDVI_{[880,712](Booting)} - 0.13 \\ \times RNDVI_{[888,736](Heading)} - 0.13 \\ \times RNDVI_{[808,744](Filling)} - 0.08 \\ \times RNDVI_{[872,744](ripening)}) \times Y_R, \quad (7)$ 

#### Yield Estimation With All Grow Stage Combination

Yield estimation model including all optimal RNDVIs of every growth stage is expressed as Eq. (8),  $R^2 = 0.84$ , RMSE = 185.45 kg/ha:

 $Y_E = (-0.98 - 0.12 \times \text{RNDVI}_{[888,720](\text{Tillering})} + 0.28$  $\times \text{RNDVI}_{[808,744](\text{Jointing})} + 2.18$  $\times \text{RNDVI}_{[880,712](\text{Booting})} - 0.23$  $\times \text{RNDVI}_{[888,736](\text{Heading})} - 0.11$  $\times \text{RNDVI}_{[808,744](\text{Filling})} - 0.06$  $\times \text{RNDVI}_{[872,744](\text{ripening})}) \times Y_R, \quad (8)$ 

#### 3.3 F-Test for Optimal Model Selection

As can be seen in Table 7 and Figure 7, with the excessive number of RNDVIs including in the models,  $R^2$ -values do not increase significantly while RMSE values don't decrease significantly. Yield estimation model would be complex if too many RNDVIs are included in the estimation model, but a univariate model using one RNDVI would decrease the estimation accuracy. In order to identify the optimal yield estimation model with greatest simplicity and accuracy simultaneously, all optimal multivariate models were tested using F-test. The optimal estimation model would be selected from the models that with F-test reaching 1% significant level and simultaneously with lower RMSE and less RNDVIs included. As can be seen in Table 7, F-test of all models reach 1% significant level. Then, taking into consideration  $R^2$  and RMSE, the multivariate model including three RNDVIs, with  $R^2$  of 0.83 and RMSE of 189.13 kg/ha, is identified as the optimal estimation model. The equation of selected model is expressed in Eq. (5).

## **Model Validation**

The estimation model Eq. (5) is validated using leave-oneout cross validation (LOOCV) method. The agreement between estimated and measured yield is tested using RMSE. The scatter plot of estimated and measured yield are shown in **Figure 8**. Diagonal lines are the 1:1 line. Ideally, when points are located on the 1:1 line, it should be a perfect match. As can be seen, all the estimated yields are acceptable and a low average relative error of 3%, and an RMSE of 215.08 kg/ha are obtained.

### DISCUSSION

Grain yield estimations with UAV-based remotely sensed data have been carried out for many types of crops. Stroppiana et al. (2015) used UAV-derived NDVI at one single growth stage to construct rice yield estimation in northern Italy and with  $R^2$ of 0.5. Zhou et al. (2017) also constructed rice yield estimation models with UAV-derived NDVI at one single growth stage and NDVIs at two or more grows stages, with  $R^2$  of 0.75 for univariate model and 0.76 for multivariate model, respectively. Compared with those researches, our study presents the idea of relative vegetation index and relative yield to improve the rice yield estimation accuracy. As the results have shown, the univariate model using RNDVI at booting stage obtained a  $R^2$ of 0.75 and the multivariate model including three RNDVIs, i.e., RNDVI<sub>[808,744]</sub> at jointing stage, RNDVI<sub>[880,712]</sub> at booting stage and RNDVI<sub>[808,744]</sub> at filling stage, gained a  $R^2$  of 0.83. The results demonstrated that when using relative vegetation index and the relative yield, a high accuracy of yield estimation can be obtained.

In this study, the RNDVI at booting stage shows the highest correlation with rice yield. This conclusion is consistent with previous studies (Chang et al., 2005; Zhou et al., 2017), in which booting stage was regarded as the optimal growth stage for rice yield estimation. Booting stage is a transitional period from vegetative growth to reproductive growth and can reflect yield potential (Zhou et al., 2017). Combining booting stage with other growth stages can further improve model estimation accuracy. The great potential of booting stage used for rice yield estimation is also proved in present study. As seen in **Table 7**, the optimum univariate model and multivariate models all included the RNDVIs at booting stage.

The large number of narrow spectral bands results in a high inter-correlation between them (Darvishzadeh et al., 2008; Yue et al., 2017). In our study, the correlation between

TABLE 7 | Models performance and F-test

TABLE / Middels performance and / -test.				
Model Expression	R <sup>2</sup>	RMSE(kg/ha)	F	1% level of significance
$y = -0.60 + 1.55x_3$	0.75	228.04	58.51	YES
$y = -1.45 + 2.65x_3 - 0.26x_5$	0.80	205.74	36.11	YES
$y = -1.06 + 0.20x_2 + 2.04x_3 - 0.23x_5$	0.83	189.13	27.94	YES
$y = -1.02 + 0.23x_2 + 1.97x_3 - 0.14x_5 - 0.10x_6$	0.84	185.84	20.57	YES
$y = -1.04 + 0.24x_2 + 2.08x_3 - 0.13x_4 - 0.13x_5 - 0.08x_6$	0.84	185.46	15.50	YES
$y = -0.98 - 0.12x_1 + 0.28x_2 + 2.18x_3 - 0.23x_4 - 0.11x_5 - 0.06x_6$	0.84	185.45	12.06	YES



FIGURE 7 | Scatter plots of measured yield vs. estimated yield derived from the optimal estimation models with (A) one single growth stage involved; (B) two growth stages involved; (C) three growth stages involved; (D) four growth stages; (E) five growth stages involved; (F) six growth stages involved.



relative VI from all possible two-band combinations and relative yield are used to identify the optimal band combination to avoid the band autocorrelation and redundancy. The optimum relative vegetation indices are mostly composed by the bands located at red edge region (712, 720, 736, and 744 nm) and near infrared region (808, 872, and 888 nm). These

results are consistent with the previous study (Ren et al., 2011; Fu et al., 2014; Verger et al., 2014). The bands in the red edge region are sensitive to the vegetation changes while the bands in near infrared region are sensitive to plant grow and can indicate the health status of leaves.

In this study, we conducted a 1-year rice experiment with two rice varieties and five nitrogen levels at a specific region. The results obtained in this study may be limited by rice varieties and research regions. In addition, due to the limitation of data volume, no systematic comparison between the models with the relative spectral index and with the traditional spectral index was made. So, this method needs to be further validated by using multi-year, multi-region, and multi-variety data.

# CONCLUSION

The study proposed a rice yield estimation method with the parcel-level relative vegetation index as input variables to overcome the external effects, such as different background, illumination and atmospheric conditions, on the absolute differences of time-series vegetation indices. Three types of relative vegetation indices were constructed using all possible two-band combinations of discrete channels from 500 to 900 nm, including RRVI, RNDVI, and RDVI. The main conclusions drawn from this study are: (1) RNDVI is the optimal type of relative vegetation index for rice yield estimation compared with RRVI and RDVI, and the correlation between RNDVIs and rice yield are significantly higher than RRVIs vs. yield and RDVI vs. yield; (2) the optimal RNDVIs at different growth

# REFERENCES

- Calderón, R., Navas-Cortés, J. A., Lucena, C., and Zarco-Tejada, P. J. (2013). Highresolution airborne hyperspectral and thermal imagery for early detection of verticillium, wilt of olive using fluorescence, temperature and narrow-band spectral indices. *Remote Sens. Environ.* 139, 231–245. doi: 10.1016/j.rse.2013. 07.031
- Carvalho, O. A. D., Guimarães, R. F., Silva, N. C., Gillespie, A. R., Gomes, R. A. T., and Silva, C. R. (2013). Radiometric normalization of temporal images combining automatic detection of pseudo-invariant features from the distance and similarity spectral measures, density scatterplot analysis, and robust regression. *Remote Sens.* 5, 2763–2794. doi: 10.3390/rs5062763
- Chang, K. W., Shen, Y., and Lo, J. C. (2005). Predicting rice yield using canopy reflectance measured at booting stage. *Agron. J.* 97, 872–878. doi: 10.2134/agronj2004.0162
- China Meteorological Data (2018). *China Meteorological Data*. Available at: http://data.cma.cn/ (accessed on May 1, 2018).
- Darvishzadeh, R., Skidmore, A., Schlerf, M., Atzberger, C., Corsi, F., and Cho, M. (2008). Lai and chlorophyll estimation for a heterogeneous grassland using hyperspectral measurements. *ISPRS J. Photogramm. Remote Sens.* 63, 409–426. doi: 10.1016/j.isprsjprs.2008.01.001
- Ding, Y., and Elvidge, C. D. (1996). Comparison of relative radiometric normalization techniques. *ISPRS J Photogramm. Remote Sens.* 51, 117–126. doi: 10.1016/0924-2716(96)00018-4
- Du, M., and Noguchi, N. (2017). Monitoring of wheat growth status and mapping of wheat yield's within-field spatial variations using color images acquired from uav-camera system. *Remote Sens.* 9:289. doi: 10.3390/rs9030289
- Du, Y., Teillet, P. M., and Cihlar, J. (2001). Radiometric normalization of multitemporal high-resolution satellite images with quality control for land

stages are generally composed by bands in red edge and near infrared regions; (3) the booting stage is the optimum stage for yield estimation when estimation model constructed by RNDVI at one single growth stage; (4) the multiple-growth-stage model, including three RNDVIs, i.e., RNDVI<sub>[808,744]</sub> at jointing stage, RNDVI<sub>[880,712]</sub> at booting stage and RNDVI<sub>[808,744]</sub> at filling stage was determined as the optimum model among the multivariate models that include RNDVIs at different growth stages. The promising ability of parcel-level relative vegetation indices for rice yield estimation is proven in this study, however, the further applications of relative vegetation indices on rice or other crop yield estimation are still needed for a solid conclusion using more measurements. This will be carried out in our future work.

# **AUTHOR CONTRIBUTIONS**

FeW analyzed the data and wrote the manuscript. FuW conceived and designed the experiments, and revised the manuscript. YZ, JhH, and JX collected the UAV hyperspectral and yield data. JfH provided comments for the manuscript.

# FUNDING

This study was supported by "National Key Research and Development Plan of China" (2016YFD0300601), "Open Fund of Laboratory of Target Microwave Properties" (2018KF02), and "National Natural Science Foundation of China" (41871328).

cover change detection. *Remote Sens. Environ.* 82, 123–134. doi: 10.1016/S0034-4257(02)00029-9

- Echappé, C., Gernez, P., Méléder, V., Jesus, B., Cognie, B., and Decottignies, P. (2018). Satellite remote sensing reveals a positive impact of living oyster reefs on microalgal biofilm development. *Biogeosciences* 15, 1–30. doi: 10.5194/bg-15-905-2018
- Fu, Y., Yang, G., Wang, J., Song, X., and Feng, H. (2014). Winter wheat biomass estimation based on spectral indices, band depth analysis and partial least squares regression using hyperspectral measurements. *Comput. Electron. Agric.* 100, 51–59. doi: 10.1016/j.compag.2013.10.010
- Govil, H., Gill, N., Rajendran, S., Santosh, M., and Kumar, S. (2017). Identification of new base metal mineralization in kumaon himalaya, india, using hyperspectral remote sensing and hydrothermal alteration. *Ore Geol. Rev.* 92, 271–283. doi: 10.1016/j.oregeorev.2017.11.023
- Guo, C., Zhang, L., Zhou, X., Zhu, Y., Cao, W., Qiu, X., et al. (2018). Integrating remote sensing information with crop model to monitor wheat growth and yield based on simulation zone partitioning. *Precis. Agric.* 19, 55–78. doi: 10. 1007/s11119-017-9498-5
- Hall, F. G., Strebel, D. E., Nickeson, J. E., and Goetz, S. J. (1991). Radiometric rectification: toward a common radiometric response among multidate, multisensor images. *Remote Sens. Environ.* 35, 11–27. doi: 10.1016/0034-4257(91)90062-B
- Honkavaara, E., Kaivosoja, J., Mäkynen, J., Pellikka, I., Pesonen, L., Saari, H., et al. (2012). Hyperspectral reflectance signatures and point clouds for precision agriculture by light weight UAV imaging system. *ISPRS Ann. Photogramm. Remote Sens. Spat. Inform. Sci.* 7, 353–358. doi: 10.5194/isprsannals-I-7-353-2012
- Honkavaara, E., Saari, H., Kaivosoja, J., Pölönen, I., Hakala, T., and Litkey, P. (2013). Processing and assessment of spectrometric, stereoscopic imagery

collected using a lightweight uav spectral camera for precision agriculture. Remote Sens. 5, 5006–5039. doi: 10.3390/rs5105006

- Le Maire, G., François, C., Soudani, K., Berveiller, D., Pontailler, J. Y., and Bréda, N. (2008). Calibration and validation of hyperspectral indices for the estimation of broadleaved forest leaf chlorophyll content, leaf mass per area, leaf area index and leaf canopy biomass. *Remote Sens. Environ.* 112, 3846–3864. doi: 10.1016/j.rse.2008.06.005
- Luna, I., and Lobo, A. (2016). Mapping crop planting quality in sugarcane from uav imagery: a pilot study in nicaragua. J. Remote Sens. 8:500. doi: 10.3390/ rs8060500
- Nebiker, S., Lack, N., Abächerli, M., and Läderach, S. (2016). Light-weight multispectral UAV sensors and their capabilities for predicting grain yield and detecting plant diseases. *Int. Arch. Photogramm. Remote Sens. Spat. Inform. Sci.* 41, 963–970. doi: 10.5194/isprsarchives-XLI-B1-963-2016
- Noack, P. O. (2016). Estimating triticale dry matter yield in parcel plot trials from aerial and ground based spectral measurements. *IFAC-Papers Online* 49, 404–408. doi: 10.1016/j.ifacol.2016.10.074
- Peng, D., Huang, J., Li, C., Liu, L., Huang, W., and Wang, F. (2014). Modelling paddy rice yield using modis data. *Agric. Forest Meteorol.* 184, 107–116. doi: 10.1016/j.agrformet.2013.09.006
- Poenaru, V., Badea, A., Dana Negula, I., and Moise, C. (2017). Monitoring vegetation phenology in the Braila plain using sentinel-2 data. *Sci. Pap. Ser. E Land R.* 6, 175–180.
- Ren, H., Zhou, G., and Zhang, X. (2011). Estimation of green aboveground biomass of desert steppe in inner Mongolia based on red-edge reflectance curve area method. *Biosyst. Eng.* 109, 385–395. doi: 10.1016/j.biosystemseng.2011.05.004
- Reynolds, M., Tattaris, M., Cossani, C. M., Ellis, M., Yamaguchi-Shinozaki, K., and Pierre, C. S. (2015). "Exploring genetic resources to increase adaptation of wheat to climate change," in Advances in Wheat Genetics: from Genome to Field, eds Y. Ogihara, S. Takumi, and H. Handa (Tokyo: Springer), 355–368. doi: 10.1007/978-4-431-55675-6\_41
- Roy, D. P., Ju, J., Kline, K., Scaramuzza, P. L., Kovalskyy, V., Hansen, M., et al. (2010). Web-enabled landsat data (WELD): landsat ETM+ composited mosaics of the conterminous United States. *Remote Sens. Environ.* 114, 35–49. doi: 10.1016/j.rse.2009.08.011
- Schneider, R., Ridler, M. E., Godiksen, P. N., Madsen, H., and Bauer-Gottwein, P. (2018). A data assimilation system combining CryoSat-2 data and hydrodynamic river models. J. Hydrol. 557, 197–210. doi: 10.1016/j.jhydrol. 2017.11.052
- Siyal, A. A., Dempewolf, J., and Beckerreshef, I. (2015). Rice yield estimation using landsat etm plus data. J. Appl. Remote Sens. 9:095986. doi: 10.1117/1.JRS.9. 095986
- Stroppiana, D., Migliazzi, M., Chiarabini, V., Crema, A., Musanti, M., and Franchino, C. (2015). "Rice yield estimation using multispectral data from UAV: a preliminary experiment in northern Italy," in *Proceedings of the IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, (Milan: IEEE), 4664–4667. doi: 10.1109/IGARSS.2015.7326869
- Teillet, P. M., El Saleous, N., Hansen, M. C., Eidenshink, J. C., Justice, C. O., and Townshend, J. R. G. (2000). An evaluation of the global 1-km AVHRR land dataset. *Int. J. Remote Sens.* 21, 1987–2021. doi: 10.1080/01431160050 021259

- Tucker, C. J., Holben, B. N., Elgin, J. H. Jr., and McMurtrey, J. E. III (1980). Relationship of spectral data to grain yield variation. *Photo Eng. Remote Sens.* 46, 657–666.
- Turner, D., Lucieer, A., and Watson, C. (2012). An automated technique for generating georectified mosaics from ultra-high resolution unmanned aerial vehicle (uav) imagery, based on structure from motion (sfm) point clouds. *Remote Sens.* 4, 1392–1410. doi: 10.3390/rs4051392
- Vega, F. A., Ramírez, F. C., Saiz, M. P., and Rosúa, F. O. (2015). Multi-temporal imaging using an unmanned aerial vehicle for monitoring a sunflower crop. *Biosyst. Eng.* 132, 19–27. doi: 10.1016/j.biosystemseng.2015.01.008
- Verger, A., Vigneau, N., Chéron, C., Gilliot, J. M., Comar, A., and Baret, F. (2014). Green area index from an unmanned aerial system over wheat and rapeseed crops. *Remote Sens. Environ.* 152, 654–664. doi: 10.1016/j.rse.2014.06.006
- Wang, Z. X., Liu, C., and Alfredo, H. (2003). From AVHRR-NDVI to MODIS-EVI: advances in vegetation index research. *Acta Ecol. Sin.* 23, 979–987. doi: 10.3321/j.issn:1000-0933.2003.05.020
- Xiao, X., Boles, S., Liu, J., Zhuang, D., Frolking, S., Li, C., et al. (2005). Mapping paddy rice agriculture in southern china using multi-temporal modis images. *Remote Sens. Environ.* 95, 480–492. doi: 10.1016/j.rse.2004.12.009
- Yu, N., Li, L., Schmitz, N., Tian, L. F., Greenberg, J. A., and Diers, B. W. (2016). Development of methods to improve soybean yield estimation and predict plant maturity with an unmanned aerial vehicle based platform. *Remote Sens. Environ.* 187, 91–101. doi: 10.1016/j.rse.2016.10.005
- Yuan, H., Yang, G., Li, C., Wang, Y., Liu, J., and Yu, H. (2017). Retrieving soybean leaf area index from unmanned aerial vehicle hyperspectral remote sensing: analysis of rf, ann, and svm regression models. *Remote Sens.* 9:309. doi: 10.3390/ rs9040309
- Yue, J., Yang, G., Li, C., Li, Z., Wang, Y., and Feng, H. (2017). Estimation of winter wheat above-ground biomass using unmanned aerial vehicle-based snapshot hyperspectral sensor and crop height improved models. *Remote Sens.* 9:708. doi: 10.3390/rs9070708
- Zarcotejada, P. J., Guilléncliment, M. L., Hernándezclemente, R., Catalina, A., González, M. R., and Martín, P. (2013). Estimating leaf carotenoid content in vineyards using high resolution hyperspectral imagery acquired from an unmanned aerial vehicle (uav). *Agric. Forest Meteorol.* 171-172, 281–294. doi: 10.1016/j.agrformet.2012.12.013
- Zhou, X., Zheng, H. B., Xu, X. Q., He, J. Y., Ge, X. K., and Yao, X. (2017). Predicting grain yield in rice using multi-temporal vegetation indices from uav-based multispectral and digital imagery. *ISPRS J. Photogramm. Remote Sens.* 130, 246–255. doi: 10.1016/j.isprsjprs.2017.05.003

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2019 Wang, Wang, Zhang, Hu, Huang and Xie. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.