Uav Hyperspectral Imagery Mining to Identify New Spectral Indices for Predicting Field-Scale Yield of Spring Maize

Yue Zhang *, Yansong Wang , Hang Hao , Ziqi Li , Yumei Long , Xingyu Zhang , Chenzhen Xia

Posted Date: 12 August 2024

doi: 10.20944/preprints202408.0804.v1

Keywords: unmanned aerial vehicles; hyperspectral imagery; spectral indices; contour map; yield

 doi:10.20944/preprints202408.0804.v1

Article

UAV Hyperspectral Imagery Mining to Identify New Spectral Indices for Predicting Field-Scale Yield of Spring Maize

Yue Zhang [1,2,3,*], Yansong Wang [4], Hang Hao [1,2], Ziqi Li [1,2], Yumei Long [1,2], Xingyu Zhang [5] and Chenzhen Xia [1,2]

[1] College of Resources and Environment, Jilin Agricultural University, 130118 Changchun, China;
[2] Key Laboratory of Soil Resource Sustainable Utilization for Jilin Province Commodity Grain Bases, Jilin Agricultural University, 130118 Changchun, China;
[3] Key Laboratory of Straw Comprehensive Utilization and Black Soil Conservation, Ministry of Education, 130118 Changchun, China;
[4] Songliao Basin Soil and Water Conservation Monitoring Center of Songliao Water Resources Commission of the Ministry of Water Resources, 130021 Changchun, China;
[5] Jilin Emergency Warning Information Dissemination Center, 130062 Changchun, China.
* Correspondence: lisa_ling7892002@163.com

Abstract: Non-destructive, accurate, and timely approach for crop yield prediction at field scale is vital for precision agriculture. This study aimed to investigate the appropriate wavelengths and their combinations to explore the new SIs derived from UAV hyperspectral images in predicting yield during the growing season of spring maize. The best wavelengths and new SIs, including the difference spectral index, ratio spectral index, and normalized difference spectral index forms, were obtained by the contour maps constructed by the coefficient of determination (R^2) from the linear regression models between the yield and all possible SIs screening out from the 450-950 nm wavelengths. The results showed that the most sensitive wavelengths were 640-714 nm at WJQ, 450-650 nm and 750-950 nm at SKS, and 450-700 nm and 750-950 nm at FJJ. The new SIs established here were different across the three experimental fields, and their performance on maize yield prediction were generally better than that of the published SIs. In addition, the new SIs presented different response to various N fertilization levels. This study demonstrated the potential of exploring new spectral characteristics from remote sensing technology for predicting field-scale crop yield in spring maize cropping systems before harvest.

Keywords: unmanned aerial vehicles; hyperspectral imagery; spectral indices; contour map; yield

1. Introduction

Climate change, environmental degradation, and unsuitable farm practices impact the crop yield and food security all over the world. Maize (Zea mays L.) is one of the most important cereal crops, and it can be used as food and materials for humans and animals in many regions of the world. Maize yield data can help the farmers and decision-makers to determine potential yield reductions, food prices, and optimizing nutrient management [1,2]. However, obtaining accurate and timely yield prediction at field scale is also a big challenge due to limited ground meteorological observations and unadvanced technologies [3,4]. Therefore, reliable field-scale maize yield predicting before harvest is essential for many regions with great significance to food security.

The traditional methods for the determination of crop yield based on measurement surveys are labor intensive and time consuming [5]. Compared with the traditional methods, remote sensing (RS) provides a cost-effective and non-destructive way for timely monitoring crop yield at a range of spatial scales [6,7]. The optical properties of chlorophylls and proteins pave the way for crop yield

 doi:10.20944/preprints202408.0804.v1

assessment by using RS techniques [8,9]. While the spatial resolution, revisit frequency combined with cloud generally limits the use of satellite for agricultural applications, because it could not meet the requirement to investigate site-specific crop yield for precision agriculture [10].

With the rapid development of imaging spectrometers and unmanned aerial vehicles (UAV), a UAV-based spectral imaging system offers a very attractive alternative: it can be operated conveniently and provides the images with high spatial, temporal and spectral resolutions [11,12]. For these reasons, many studies were conducted to monitor the crop nitrogen (N)-related properties, such as LAI, chlorophyll content, nitrogen content, N nutrition index (NNI), biomass and yield through UAV aerial imagery [13–15]. Recently hyperspectral techniques developed rapidly, and the hyperspectral data generally concentrates on visible to near-infrared (VIS-NIR) spectroscopy, with high spectral resolution (<10 nm) [16]. It can increase the detection sensitivity of crop N status, by acquiring a large number of contiguous narrow bands for obtaining subtle variations in spectral reflectance of the plant canopy [17,18]. The existing spectral indices, such as normalized difference vegetation index (NDVI), enhanced vegetation index (EVI), soil adjusted vegetation index (SAVI), normalized difference red edge index (NDRE), red-edge chlorophyll index (CIre), etc., are the most sensitive predictors for crop yield during the growing season [19,20]. In summary, most spectral indices were constructed by the difference, ratio, normalization forms and their combinations based on the VIS-NIR bands to diagnose crop N status and predict crop yield. While each spectral index only represents part spectral information on crop nutrition status, and it can also be easily affected by leaf architecture, canopy structure, growth stage, fertilization amounts, and field environment.

As far as we know, most studies adopted the published SIs which were not designed for the yield assessment to predict crop yield, and the SIs they employed varied across geographic locations, climatic environment, crops, varieties, and canopy structures. For example, some research tested several vegetation indices including lnRE, MSAVI, GNDVI, NDRE, SCCCI via sub-orbital multispectral sensor for predicting soybean crop yield and got satisfactory results [21]. Some studies combined high spatial resolution data and vegetation indices (NDVI, RVI, EVI2, and WDRVI) to predict the yield of winter wheat, and achieved improved predictive accuracy [22]. However, revealing the response characteristics of spectral reflectance to crop yield is the basis for establishing a spectral index and providing reliable yield prediction accordingly. To date, few studies have been conducted to find the sensitive spectral bands and designed the new SIs with good performance for crop yield prediction during the crop growing season.

Due to the complex climate conditions, topography, field managements, smallholder farms are located in Northeastern China and with high heterogeneity existed in the farmland, even within a field. At present, crop yield prediction within field-scale is limited and challenging, due to the difficult accessibility of large numbers of observations both at spatial and temporal scale and the capacity of processing the big data such as hyperspectral imagery with a large number of spectral bands. Meanwhile, the most sensitive wavelengths and SIs forms for maize yield prediction were unclear at different growth stages during the growing season. Therefore, the objectives of this study are to (i) screen out the narrowband SIs in 450-950 nm spectrum for the maize yield prediction by using the linear regression method; (ii) compare the performance of new SIs in this study and the published SIs for the yield prediction across the key growth stages; (iii) investigate the sensitivity of SIs to the yield of spring maize under different N fertilizer treatments and soil types.

2. Materials and Methods

2.1. Study Area and Field Experiments Design

The field experiments in this study were conducted since 2009 at three sites with different soil types, alluvial soil for Wang-Jia-Qiao (WJQ), black soil for San-Ke-Shu (SKS) and aeolian sand soil for Fu-Jia-Jie (FJJ), in Lishu County of Jilin Province, Northeast China (Figure 1 & Table 1). Lishu County (123°45′-124°53′E, 43°02′-43°46′ N) is located in the center of Jilin Province. This region has cold temperatures and a humid/sub-humid climate. The average annual temperature is 6.5°C, with a sunlight time of approximately 2541.4 h and 155 frost-free days. Every experimental site was divided to 15 microplots assigned to 5 treatments with 3 replications for each treatment. The area of each

 doi:10.20944/preprints202408.0804.v1

microplot was 60 m^2 (10.0 m×6.0 m), and the furrow between two microplots was 0.6 m. Five N fertilizer application rates were 0, 168, 240, 270 and 312 kg N ha-1 (designated as N0, N168, N240, N270 and N312) and one spring maize type (Liangyu 99) was used. Half of the total N fertilizer was applied as base fertilizer together with phosphorus (72 kg P_2O_5 ha^{-1} as calcium superphosphate ($Ca(H_2PO_4)_2$) and potassium (90 kg K_2O ha-1 as K_2SO_4) before sowing, while the remaining N fertilizer was applied in the 0-10 cm soil layer at about 15 cm away from the rows at the 12 leaf collar stage (V12). Spring maize was sown in early May and harvested in middle October, and the planting density was 65000 plants ha^{-1}. The experimental fields were conducted using conventional tillage methods. All the microplots were not irrigated and no obvious pest and disease stress during the growing season of spring maize, thus N application rates was the only limiting factor for maize yield.

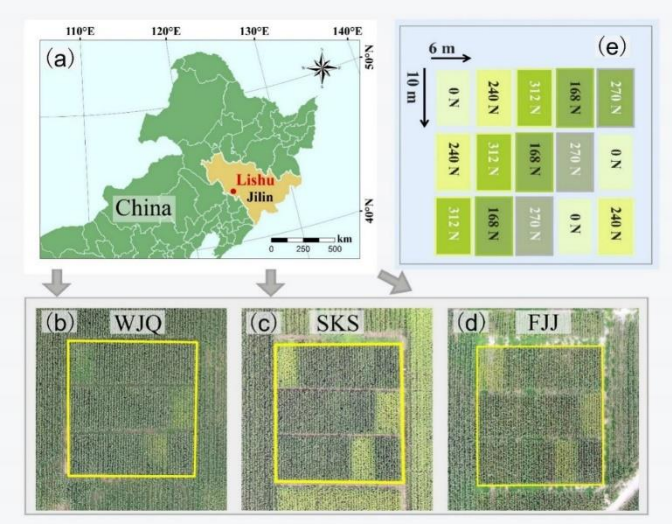


Figure 1. Location of the study area (a), UAV hyperspectral images (b)-(d) and the nitrogen application rates (e) of three experimental fields (WJQ, SKS, and FJJ).

Table 1. Soil properties of the three experimental fields.

Experimental fields	Sand (%)	Silt (%)	Clay (%)	pH	Organic matter (g/kg)	Total N (g/kg)	Alkali-hydrolyzable N (mg/kg)	Available P (mg/kg)	Available K (mg/kg)
WJQ	47.7	29.6	22.7	5.15	12.2	1.04	91.6	29.1	52
SKS	32.5	25.2	42.3	6.16	25.3	1.69	128.2	43.9	122
FJJ	73.6	9.6	10.8	6.71	13.2	1.01	58.8	13.6	43

2.2. Crop Yield Measurements

Yield surveys were conducted for the three experimental fields in 2019 and 2020. When maize reached fully maturity, we removed the border areas of each field and harvested the maize in netting bags for each microplot. Then, the seeds were dried and weighted, and maize yield was calculated in kilograms per hectare (kg ha^{-1}).

2.3. UAV Hyperspectral Imaging Acquisition and Preprocessing

UAV hyperspectral images were collected on June 23 (jointing), August 1 (silking) and September 28 (maturity) of 2019 and June 28 (jointing), August 15 (silking) and September 24 (maturity) of 2020, respectively. The UAV campaign was conducted by using a DJI S1000 UAV (SZ DJI Technology Co., Ltd., Guangdong, China) with six propellers as the hyperspectral sensor platform under cloudless conditions between 10:30 and 14:30. It has a flying speed about 8 m/s, and a takeoff weight of up to 6-11 kg. The hyperspectral images were acquired by a UHD 185 sensor (Cubert GmbH Company), which was fitted with fiber optics with a 27° field of view. The sensor was

 doi:10.20944/preprints202408.0804.v1

operated in 450-950 nm at a sampling interval of 4 nm and 125 channels of better imaging quality. The flight was designed to ensure 70% overlap both across- and along- track. A digital camera (Sony RX1R II, Sony, Tokyo, Japan, with a 4000×3000 pixel detector and a mass of 1000 g) was also employed along with the UHD 185 sensor. Finally, a total of 1506 hyperspectral images were acquired at an altitude of 100 m and a spatial resolution of 5 cm.

Then, these images were mosaicked using Agisoft Photoscan Pro software (Agisoft LLC, St. Petersburg, Russia) through the process of aligning photos, building dense point clouds, building mesh, building texture, and building othomosaic. After the othomosaic was acquired, it was conducted atmospheric correction using the Fast Line-of-Site Atmospheric Analysis of Spectral Hypercubes (FLAASH) module in ENVI 5.5 software (Exelis Visual Information Solutions, America). And then, geometric correction was applied using 15 ground control points in each field in ArcGIS 10.6 software (ESRI Inc., USA), and the root mean square error (RMSE) was smaller than 0.5 pixels for each point.

2.4. Spectral Indices Calculations

The spectral index is defined as the combination of some specific band reflectance, and the determination of the spectral bands refers to a certain physical basis and can help to improve the estimation sensitivity of target parameter (Zhao et al. 2018). In this paper, several spectral indices forms including difference spectral index (DSI), ratio spectral index (RSI), and normalized difference spectral index (NDSI), were selected to construct a variety of spectral indices within a spectral range of 450-950 nm to explore the sensitive indices for yield prediction of maize. Their formulars were as the following:

$$DSI = R_1 - R_2 \quad (1)$$

$$RSI = R_1 / R_2 \quad (2)$$

$$NDSI = (R_1 - R_2)/(R_1 + R_2) \quad (3)$$

where R_1 and R_2 were the spectral reflectance of random wavelengths from 450 to 950 nm. Then, all possible DSI, RSI, and NDSI based on any two random bands from 450 to 950 nm at an interval of 4 nm were regressed with the maize yield by using the linear regression model. The contour map of coefficients of determination (R^2) was plotted and the sensitive ranges of spectral bands were identified from this map. Meanwhile, the best SIs were selected based on the coefficient of determination (R^2) between the maize yield and different forms of SIs. The above procedures were conducted in R 4.3.2 software.

In addition, several widely used SIs for crop yield prediction were also considered here to test and compare with the new SIs for their performance on maize yield prediction across different growth stages (Table 2). These SIs were selected because they are related to various biophysical parameters including the pigment, canopy structure, and physiology, and they have the similar forms with the SIs constructed in this paper, except for DDi, CI_{re}, Maccioni, and mND705.

Table 2. The selected published narrowband SIs for maize yield prediction in this study.

Narrowband VIs	Calculation formula in this study	Reference
DDi (Desertification difference index)	$(R_{749} - R_{720}) - (R_{701} - R_{672})$	[23]
VOG (Vogelman red edge index)	R_{740}/R_{720}	[24]
CIre (Red-edge chlorophyll index)	$R_{850}/R_{730} - 1$	[25]
NDI (Normalized difference index)	$(R_{850} - R_{710})/(R_{850} + R_{680})$	[26]
NDSI (Normalized difference spectral index)	$(R_{860} - R_{720})/(R_{860} + R_{720})$	[27]
NDRE (Normalized difference red edge index)	$(R_{790} - R_{720})/(R_{790} + R_{720})$	[28]
Maccioni	$(R_{780} - R_{710})/(R_{780} - R_{680})$	[29]

 doi:10.20944/preprints202408.0804.v1

mND705(750,705,450) (Modified normalized difference 705)	$(R_{750} - R_{705})/(R_{750} + R_{705} - 2R_{450})$	[30]

2.5. Comparison between New SIs and Published SIs for Yield Prediction

The leave-one-out cross-validation process was employed to test the yield prediction performance. The prediction accuracy of the linear prediction models was assessed by using the coefficient of determination (R^2), and root mean square error (RMSE) [31]. The R^2 and RMSE values were calculated as follows:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(O_i - P_i)^2}{\sum_{i=1}^{n}(O_i - \overline{O})^2} \quad (4)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(O_i - P_i)^2}{n}} \quad (5)$$

3. Re

4. Su

where n is the number of samples, P_i is the estimated yield from the regression model, O_i is the observed yield, and $\overline{O}$ is the average of the observed yields. A 1:1 line was also drawn to test the model fit degree of the measured vs. predicted yield values.

Results

3.1. Description of Maize Yield Measurement

Table 3 shows the descriptive statistics of the maize yield including minimum, maximum, mean, median, standard deviation (SD), skewness, and coefficient of variation in 2019 and 2020. A total of 90 yield observations were collected across the three experimental fields in this study. The crop yield ranged from 447.86 to 12904.03 kg ha^{-1} with the CV of 44.48% in 2019, and from 148.70 to 12299.61 kg ha^{-1} with the CV of 58.69 % in 2020. The large data ranges and variations can be used to test the sensitivity and reliability of the new SIs established for maize yield prediction in the following steps.

Table 3. Descriptive statistics of the yield (kg ha^{-1}) of spring maize in 2019 and 2020.

Year	Min	Max	Mean	Median	SD	Skewness	CV (%)
2019	447.86	12904.03	8155.47	10100.83	3627.79	-0.45	44.48
2020	148.70	12299.61	7106.39	8833.99	4170.78	-0.49	58.69

* Abbreviations: SD, standard deviation; CV, coefficient of variation.

3.2. The Canopy Spectral Characteristics of Maize

Canopy spectral reflectance was affected by leaf nitrogen content, leaf area index (LAI), canopy architecture, and different growth stages. As shown in Figure 2, the canopy spectra were declined with increasing level of N treatments in the VIS wavebands (450-750 nm), and raised with increasing level of N treatments in the NIR wavebands (780-900 nm). The figure also presented that canopy spectra generally reached the maximum value at silking growth stage, compared with the reflectance values at the other two growth stages. In addition, the three experimental fields presented some difference in the spectral reflectance across different growth stages, while the spectra at silking stage in the three fields always had obvious heterogeneity across the five N application rates. These characteristics indicated that the maize canopy spectra had the capability to discriminate various N status and the yield of spring maize accordingly, across different N treatments, growing stages, and experimental fields.

 doi:10.20944/preprints202408.0804.v1

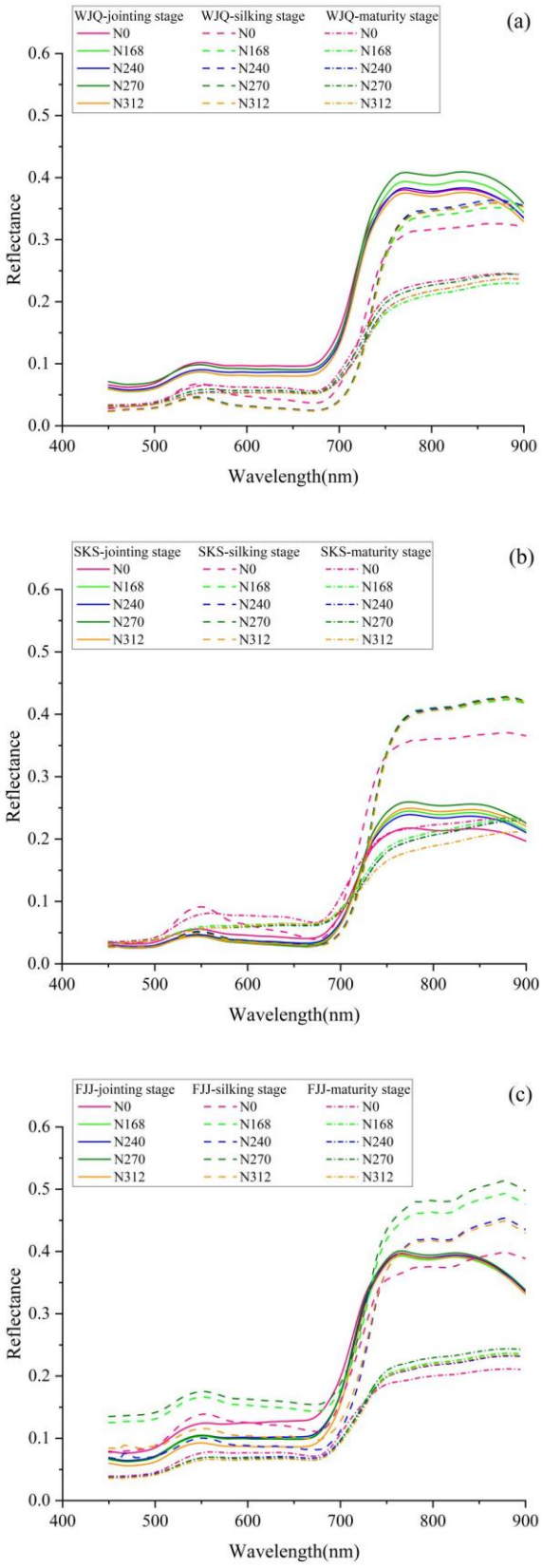


Figure 2. Mean canopy reflectance spectra curves of spring maize with different levels of N treatments across three growth stages in the three experimental fields. (a): WJQ, (b): SKS, (c): FJJ.

3.3. Determination Coefficients (R2) between DSI, RSI, and NDSI and Yield

 doi:10.20944/preprints202408.0804.v1

All possible spectral indices formed by difference, ratio, and normalization were developed to regressed with maize yield across the jointing, silking and maturity stages of spring maize in 2019-2020. The determination coefficient R^2 of the regression models were computed and shown in contour maps, as shown in Figure 3 (WJQ), Figure 4 (SKS), and Figure 5 (FJJ). These figures showed that the contour maps derived from the silking stage generated higher R^2 values and broader hot zones than the contour maps from the jointing and maturity stages, for all the three experimental fields. Furthermore, the areas of hot zones in the subplots of RSI and NDSI were much more than those in the subplots of DSI. RSI and NDSI showed the similar patterns in the contour maps across different growth stages and experimental fields. Compared with the WJQ experimental field, SKS and FJJ got more sensitive wavelength ranges at the silking stage for all the DSI, RSI, and NDSI forms. At the silking stage, the higher R^2 values derived from RSI and NDSI forms were generally ranged 0.50-0.70 in the hot zones of contour maps at WJQ, 0.60-0.90 in the hot zones of contour maps at SKS, and 0.50-0.70 in the hot zones of contour maps at FJJ. Accordingly, the wavelengths with the higher R^2 were concentrated in 640-714 nm at WJQ, 450-650 nm and 750-950 nm at SKS, and 450-700 nm and 750-950 nm at FJJ.

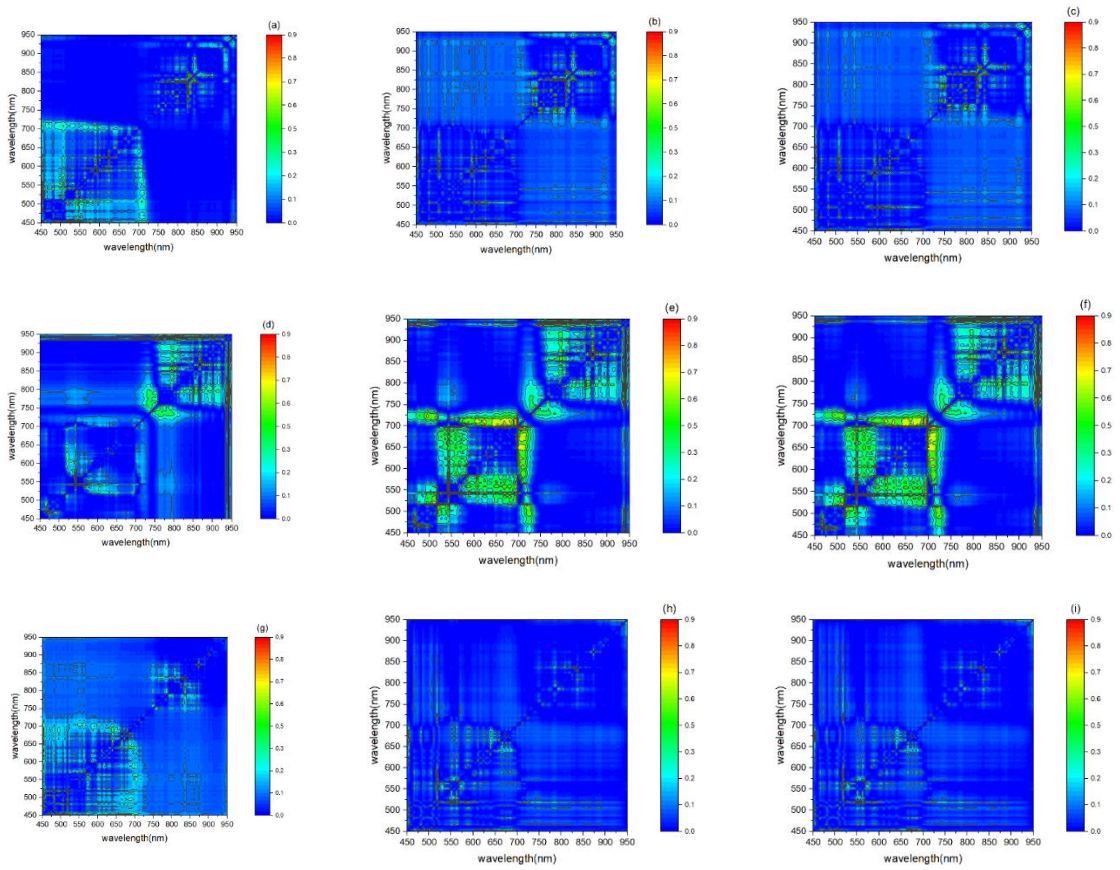


Figure 3. The contour maps for linear model between the difference spectral index (DSI), ratio spectral index (RSI), normalized difference spectral index (NDSI) and maize yield in WJQ experimental field. (a)-(c): DSI, RSI, and NDSI forms at the jointing stage; (d)-(f): DSI, RSI, and NDSI forms at the silking stage; (g)-(i): DSI, RSI, and NDSI forms at the maturity stage, respectively.

 doi:10.20944/preprints202408.0804.v1

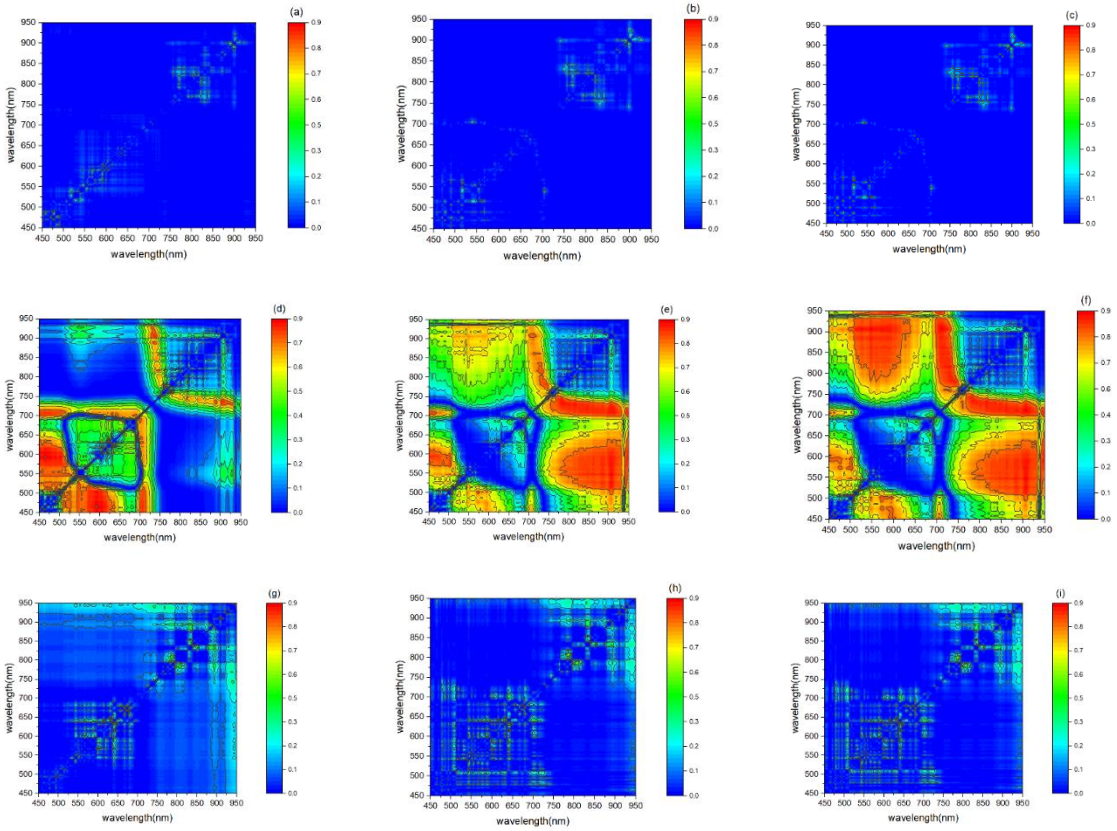
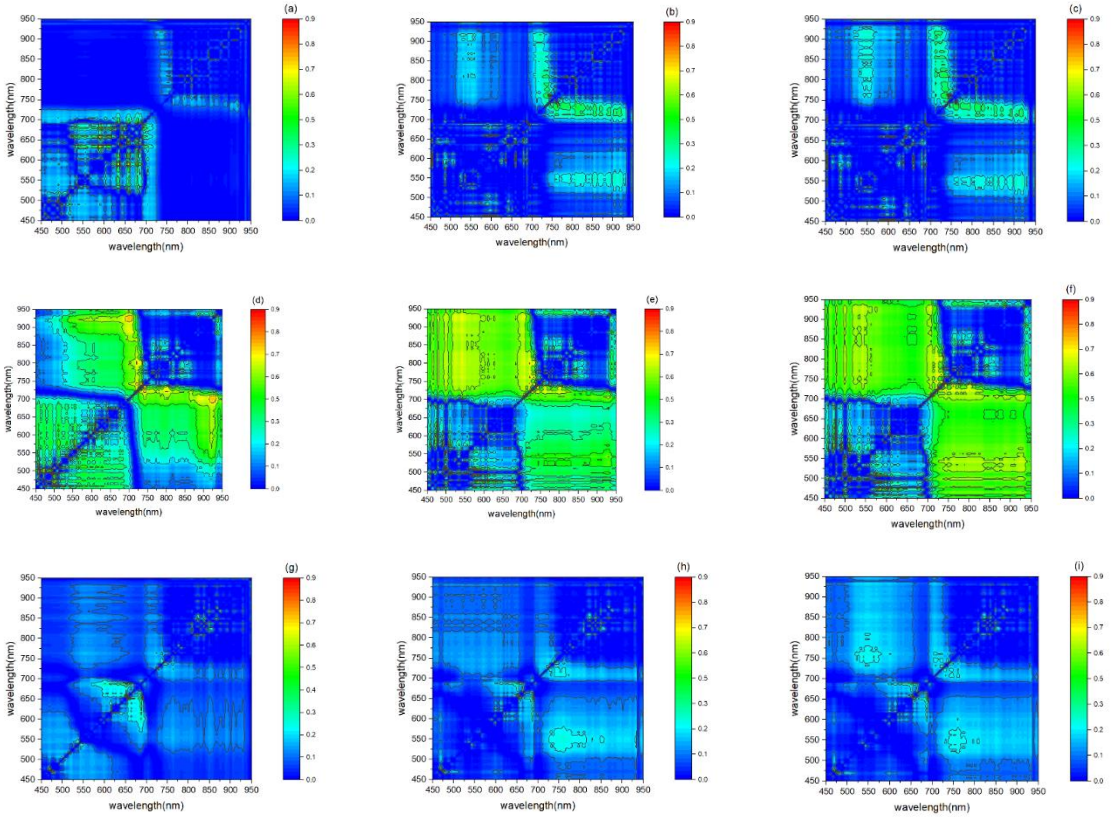


Figure 4. The contour maps for linear model between the difference spectral index (DSI), ratio spectral index (RSI), normalized difference spectral index (NDSI) and maize yield in SKS experimental field. (a)-(c): DSI, RSI, and NDSI forms at the jointing stage; (d)-(f): DSI, RSI, and NDSI forms at the silking stage; (g)-(i): DSI, RSI, and NDSI forms at the maturity stage, respectively.



 doi:10.20944/preprints202408.0804.v1

Figure 5. The contour maps for linear model between the difference spectral index (DSI), ratio spectral index (RSI), normalized difference spectral index (NDSI) and maize yield in FJJ experimental field. (a)-(c): DSI, RSI, and NDSI forms at the jointing stage; (d)-(f): DSI, RSI, and NDSI forms at the silking stage; (g)-(i): DSI, RSI, and NDSI forms at the maturity stage, respectively.

3.4. Comparison with the Published SIs for Maize Yield Prediction

The results in Section 3.3 indicated that the best period for maize yield prediction before harvest was the silking stage, so only the contour maps derived from this growth stage were considered to construct the new SIs for yield prediction across the three experimental fields. Table 4 showed the new SIs established in this paper and the comparison results with the published SIs for maize yield prediction in WJQ, SKS, and FJJ fields. The new SIs constructed here were different across the three experimental fields, such as NDSI (690,710), RSI(678,710), and DSI(858,794) in WJQ, NDSI (902,546), RSI(906,546), and DSI(590,462) in SKS, and NDSI (730,758), RSI(706,758), and DSI(698,922) in FJJ, respectively. The performance of the new SIs on yield prediction was achieved with R^2 ranged from 0.463 to 0.773 in WJQ, R^2 ranged from 0.851 to 0.892 in SKS, and R^2 ranged from 0.651 to 0.726 in FJJ, respectively. The performance of the existed SIs on yield prediction was achieved with R^2 ranged from 0.008 to 0.254 in WJQ, R^2 ranged from 0.629 to 0.862 in SKS, and R^2 ranged from 0.083 to 0.199 in FJJ, respectively. All the three SI forms (DSI, RSI, and NDSI) in SKS dataset obtained higher R^2 and lower RMSE than that from the dataset in other two fields, for both the dataset from the new SIs and published SIs.

Table 4. Accuracy of linear regression models of the new SIs and published SIs for maize yield prediction in the three experimental fields.

WJQ			SKS			FJJ		
SIs	**R^2**	**RMSE (kg ha^{-1})**	**SIs**	**R^2**	**RMSE (kg ha^{-1})**	**SIs**	**R^2**	**RMSE (kg ha^{-1})**
NDSI(690,710)	0.773	1842.563	NDSI (902,546)	0.890	1102.623	NDSI (730,758)	0.696	2224.780
NDSI(690,706)	0.738	2065.236	NDSI (902,550)	0.877	1169.254	NDSI (730,798)	0.668	2239.541
NDSI(690,714)	0.719	2154.235	NDSI (906,550)	0.871	1174.236	NDSI (730,762)	0.654	2298.251
RSI(678,710)	0.689	2203.587	RSI(906,546)	0.892	1095.245	RSI(706,758)	0.651	2230.547
DSI(858,794)	0.463	1326.741	DSI(590,462)	0.851	1503.252	DSI(698,922)	0.726	2169.541
DDi	0.008	2772.250	DDi	0.629	2389.804	DDi	0.083	2562.750
VOG	0.208	2475.841	VOG	0.862	1454.887	VOG	0.170	2438.515
CIre	0.074	2677.032	CIre	0.762	1913.550	CIre	0.116	2517.017
NDI	0.254	2403.923	NDI	0.827	1629.380	NDI	0.096	2544.796
NDSI	0.059	2699.700	NDSI	0.857	1481.413	NDSI	0.199	2395.427
NDRE	0.038	2729.652	NDRE	0.639	2355.802	NDRE	0.101	2538.020
Maccioni	0.086	2660.352	Maccioni	0.766	1902.562	Maccioni	0.198	2397.110
mND705	0.247	2415.400	mND705	0.791	1794.349	mND705	0.170	2437.842

Figure 6 presented the validation results of maize yield prediction model derived from the new SIs with the best performance in WJQ, SKS, and FJJ, respectively. The plots were evenly distributed near the 1:1 line, which pointed out the better performance of the new SIs for the yield prediction.

 doi:10.20944/preprints202408.0804.v1

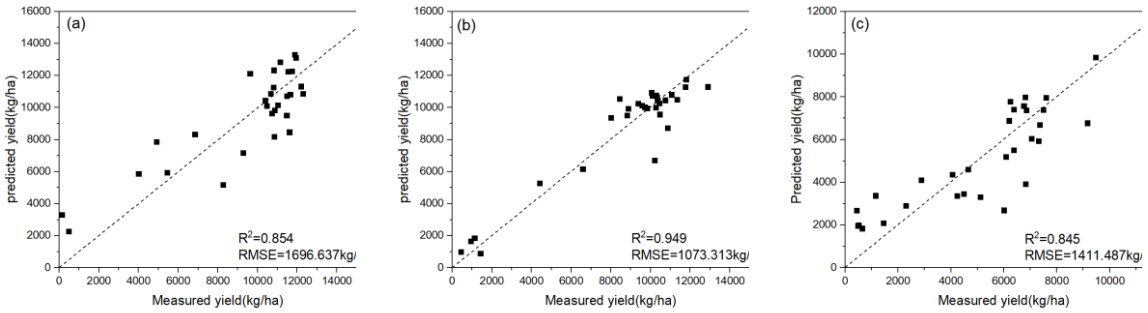


Figure 6. Scatter plots of the measured yield (kg/ha) versus predicted yield (kg/ha) by the new SIs: (a) NDSI (690,710) in WJQ, (b) RSI(906,546) in SKS, and (c) DSI(698,922) in FJJ, respectively.

3.5. New SIs Response to Different N Treatments

The response of maize yield to different N fertilization treatments across the three experimental fields was shown in Figure 7. The FJJ field with aeolian sand soil type had the lowest yield, while the WJQ field with alluvial soil type had the highest yield. When N application rates was less than 168 kg ha^{-1}, the yield improved with the increasing N application rates; when N application rates was more than 168 kg ha^{-1}, the yield did not increase with the increasing N application rates. This was consistent among the three experimental fields with different soil types.

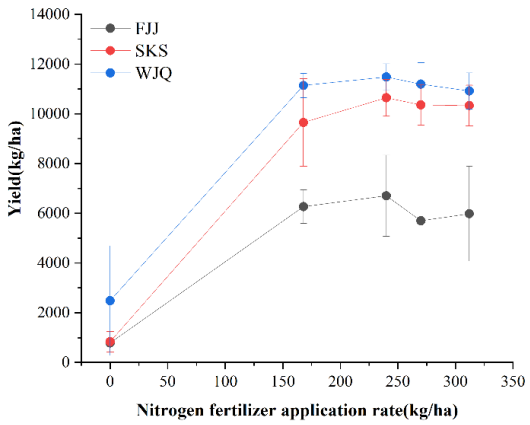


Figure 7. The response of maize yield to different N application rates at the three experimental fields.

At the silking stage, the response of the new SIs (DSI, RSI, and NDSI) to different N fertilizer treatments in the three experimental fields (WJQ, SKS, and FJJ) was presented in Figure 8. Since the interval range of N treatments was not the same and the optimal N fertilizer amounts was 168 kg/ha, the new SIs did not show the linear increased or decreased trend with the increase of N fertilization levels obviously. Compared with WJQ and FJJ fields, the new SIs in SKS showed distinct response to the five N fertilization levels obviously. For example, NDSI(902,546) reached the maximum value at the N168 level and the minimum value at the N0 level, and decreased slightly from N240 to N312 levels. RSI and DSI forms had better performance in discriminate different N fertilization levels compared with the NDSI form in the three experimental fields. For example, the medium fit line of the RSI(678,710) and DSI(858,794) in WJQ were between 0.41-0.46, and -40-210 respectively, while the NDSI(690,706) had the medium values all around -0.30 across the five N fertilization levels.

 doi:10.20944/preprints202408.0804.v1

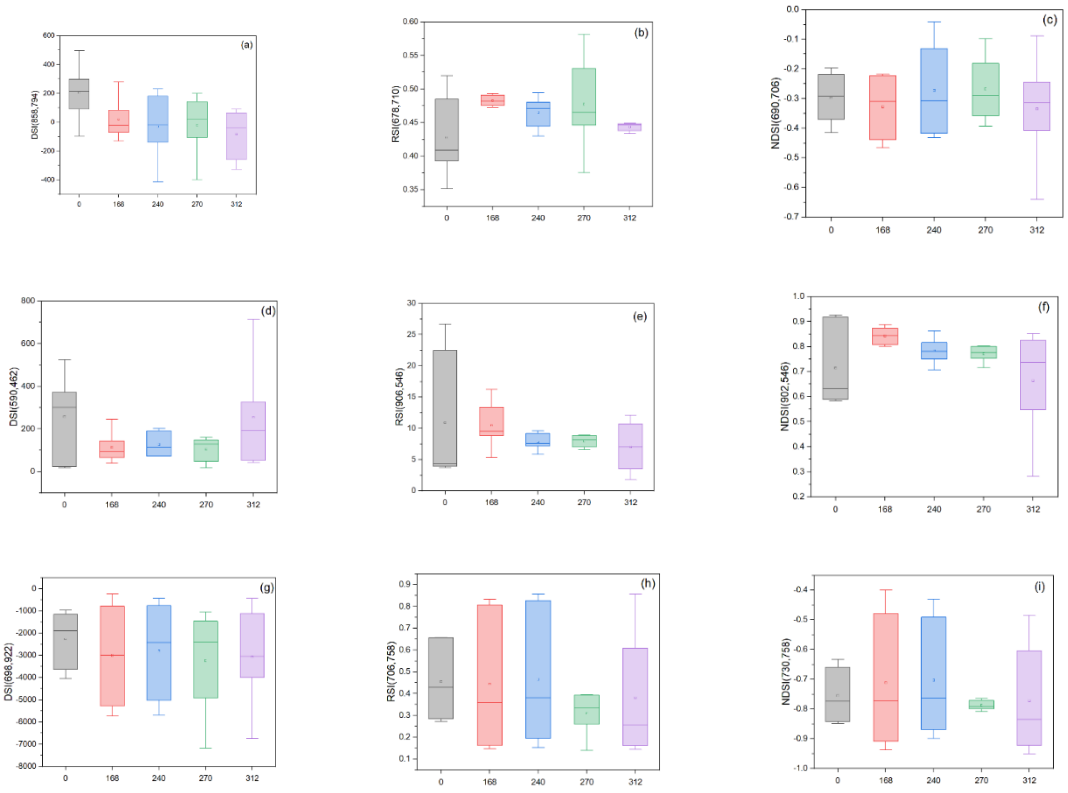


Figure 8. The response of new SIs to different N treatments in the three experimental fields. (a)-(c): DSI, RSI, and NDSI forms in WJQ, (d)-(f): DSI, RSI, and NDSI forms in SKS, and (g)-(i): DSI, RSI, and NDSI forms in FJJ, respectively.

4. Discussion

4.1. Screening out the Sensitive Wavelengths and SIs

Development of new technologies and devices has led to the inventions for various agricultural management by using remote sensing data. The new approaches, such as UAV, are capable of detecting crop growth status and predict the yield, especially at the field scale. However, the spectral information collected form UAV platform, particularly the canopy reflectance of crops, is impacted by many factors such as crop nutrition status, crop types, canopy structure, and growth stages. Based on this background, the sensitive wavelengths and SIs need to be investigated and designed for the accurate prediction of crop yield at field scale, and the relevant studies were limited to our knowledge. According to earlier and present studies [32], NIR bands (>780 nm) were more efficient in describing crop N status than visible light band generally, and this was also validated in this study (Figure 2). However, it was not always this situation in this study. The performance of visible light and NIR bands on N levels were different across the three fields with different soil types. For example, visible light band was more efficient in describing N levels than red edge and NIR bands in FJJ with aeolian sand soil type, especially at jointing and silking stage. It would also affect the selection of sensitive wavelengths and SIs for crop yield prediction in the three experimental fields in this study.

This study identified the most sensitive wavelengths and SIs to predict maize yield by screening out the determination coefficient (R^2) through the contour mapping approach. All the three types of SI forms (DSI, RSI, and NDSI) used here indicated that the sensitive bands found in the three experimental fields were different, such as 640-714 nm at WJQ, 450-650 nm and 750-950 nm at SKS, and 450-700 nm and 750-950 nm at FJJ. It revealed that the spectral characteristics of crop is also affected by soil types, even with the same crop type, crop variety and N fertilization levels. However, the red edge region was screened out in all the three fields, because the red edge was proved to be

 doi:10.20944/preprints202408.0804.v1

closed related to plant biomass and N levels in previous studies [33,34]. Among the above sensitive wavelength regions, new SIs for maize yield prediction were constructed by using the linear regression models between the yield and SIs with the best performance. It generally indicated that the normalized difference indices and ratio indices showed great potential for predicting maize yield compared with the difference indices in this study, except for the DVI(698,922) in FJJ field. The result was consistent with the previous studies (Wu et al. 2021), in which the NDSIs also outperformed other indices forms in crop yield predictions. It was because that NDSI and RSI can avoid the effects of atmospheric and soil background on the canopy spectral reflectance due to their formular forms.

4.2. Comparison between the New SIs vs. Published SIs

Most of the published spectral indices for yield prediction were not originally designed to estimate crop yield, the present study selected some SIs widely used for crop yield prediction from previous studies. By using the same experimental N fertilization treatments in the three fields respectively, the performance of the eight published SIs presented large differences in maize yield prediction, with the determination coefficient (R^2) ranged from 0.008 to 0.862. At the same time, the published SIs did not always achieve the ideal performance on the yield prediction across the three experimental fields. Moreover, the performance of the published SIs for yield prediction were different in the previous studies [3,7]. This is because these existed SIs were mostly pigment, structure, and physiology-related indicators, which were designed to estimate plant nitrogen content, crop N uptake and biomass, not for the crop yield. While the new SIs with the best performance for yield prediction were obtained from this study, with the validation results R^2=0.854 and RMSE=1696.637 kg/ha by NDSI(690,710) in WJQ, R2=0.949 and RMSE=1073.313 kg/ha by RSI(906,546) in SKS, and R^2=0.845 and RMSE=1141.487 kg/ha by DSI(698,922) in FJJ, respectively. Many researchers have contributed to find universal spectral indices to monitor crop N status, biomass, and yield, but it is so difficult to achieve it due to the complex interactions of crop biochemical characteristics, spectral information, soil properties, and farm management practices [9,15,34]. The results in this study proved that sensitive SIs need to be designed for crop yield prediction, rather than using the exited indices to get the poor performance and unreliable yield prediction results. In addition, the complex form of spectral indices would not always get good performance on crop yield prediction. Just as the existed SIs (e.g., NDI, Maccioni, mND705) selected in this paper also did not get ideal prediction results, although they had more wavelengths or more complex index structures than the difference, ratio, and normalization forms.

Crop yield is an important end-of-season trait, however, which needs to be predicted before harvest, and it combines the cumulative impact of crop growth status during the growing season. Which growth stage is the most optimal period to predict crop yield is vital for the farm management and food security. The best growing period determined in this study is the silking stage of spring maize across all the three experimental fields. Since maize canopy is usually sparse at jointing stage and during this period the variations in plant height, biomass, LAI may weak crop growth status reflected in canopy reflectance, so the jointing stage would not be the ideal prediction period. In addition, N is the vital element for the yield and it was stored in vegetative organs to promote the crop growing at vegetative stage, and then N is transported from these vegetative organs to reproductive organs during reproductive growth stage. The silking stage is just attributed to the later period of the vegetative stage and crop N status can also be monitored by the spectrum, while the maturity stage is attributed to the reproductive stage and most N had transported to the seeds, so the silking stage is the best period for maize yield prediction.

Although the new SIs showed good performance on the calibration and validation results, the index calculation methods here included the three basal forms (difference, ratio, and normalization) only, and other combination structures and more spectral bands were required to be involved for crop yield prediction in the next step. In addition, only the linear relationship between maize yield and SIs were investigated, future work would also be concentrated on the non-linear relationships between them. Due to the limitation of the validation data in this study, future work is planned to

 doi:10.20944/preprints202408.0804.v1

test the reliability of the new spectral indices for crop yield prediction across different climatic regions, crop types, and farm management practices.

4.3. New SIs Response to Different N Levels and Soil Types

The important effect of N fertilizer on the crop N content, NNI, biomass, and yield have been widely reported [35–37], and the relationship between maize yield and N fertilization levels here showed that the yield did not increase with the increasing N fertilization amounts when the amounts reached 168 kg ha^{-1}. It means that if N fertilizer amounts is lower than 168 kg ha^{-1} the yield would reach the lower values with N deficiency, while if N fertilizer amounts is higher than 168 kg ha^{-1} the yield would reach the higher values with optimized or excessive N supply. In this paper, the response of new SIs to different N fertilization levels were investigated aimed to analyze the sensitivity of new SIs to the deficient and excessive N fertilizer levels and the relevant yield accordingly. The yield for spring maize varied from 148.70 to 12904.03 kg ha^{-1}, which were significantly affected by the N fertilizer amounts, soil types, and the climatic conditions in 2019-2020. The above-mentioned linear regression models between the yield and new SIs showed a highly satisfactory performance at the silking stage of spring maize in the three experimental fields. It indicated that the new indices constructed in this paper have the ability to reflect the large variations in the yields.

The sensitivity of SIs in different N fertilizer amounts is also a key aspect to diagnose the performance of SIs for crop yield prediction. Figure 8 showed that all the three SI forms (DSI, RSI, and NDSI) had the ability to discriminate different N fertilization levels in WJQ and SKS field, although the new SIs in FJJ had the poor response. That is because that the soil types of WJQ and SKS is alluvial soil and black soil respectively, with higher soil fertility and quality than the aeolian sand soil in FJJ. Accordingly, the crops planted in fertile soil would present obvious response to different N fertilization amounts, while the crops planted in infertile soil would show unapparent response to different N fertilization levels. Therefore, the soil property is also required to be considered for the establishment of sensitive indices for crop yield prediction, more than the spectral response to crops to different fertilization conditions. Sometimes DSI and RSI had better performance and larger response ranges than NDSI in this study, and it was indicated that different index structures and forms need to be investigated to obtain the optimal and robust SIs for crop yield prediction.

5. Conclusions

This study investigated the potential of UAV canopy hyperspectral images and explore new spectral indices designed for predicting maize yield during the growing season in three experimental fields with different soil types. The most sensitive wavelengths and SIs were identified by screening out the determination coefficient (R^2) of linear model between the yield and all possible SIs with the wavelengths ranged from 450 to 950 nm. The best growth period for the yield prediction was the silking stage of spring maize. Different sensitive wavelengths and SIs were recommended as the best indicators to assess maize yield across the three experimental fields. It revealed that the best indices determined for the yield prediction were not only affected by the crop spectral characteristics, but also be influenced by soil properties of farm fields. From the eight published SIs and fifteen new indices, NDSI(690,710), RSI(906,546), and DSI(698,922) outperformed the existing indices and achieved the best performance for the yield prediction with the determination coefficients 0.773, 0.892, and 0.726 respectively. It can be used to establish the optimal spectral indices applicable to predict crop yield in other fields. The new SIs also presented distinct different response to various N fertilization levels across the three SI forms (DSI, RSI, and NDSI), except for FJJ field. This augmentation will subsequently improve the accuracy and reliability for predicting crop yield across various growth stages, N fertilization levels, and soil properties.

Author Contributions: Conceptualization, Zhang Y.; methodology, Wang Y.; software, Li Z.; validation, Long Y. and Zhang X.; formal analysis, Xia C.; investigation, Li Z.; resources, Wang Y.; data curation, Zhang Y.; writing—original draft preparation, Zhang Y.; writing—review and editing, Zhang Y. All authors have read and agreed to the published version of the manuscript.

 doi:10.20944/preprints202408.0804.v1

Funding: This research was funded by the National Natural Science Foundation of China, grant number 42301074; and the Natural Science Foundation of Jilin Province, China, grant number 20240101041JC; and the National Key Research and Development Program of China, grant number 2021YFD1500800.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. García-Lara, S.; Serna-Saldivar, S.O. *Chapter 1 - corn history and culture*. In: Serna-Saldivar, S.O. (Ed.), Corn, 3rd ed.; AACC International Press: Oxford, 2019; pp. 1-18.
2. Jaafar, H.; Mourad, R. GYMEE: A global field-scale crop yield and ET mapper in google earth engine based on Landsat, weather, and soil data. *Remote Sensing* **2021**, *13 (4)*, 773.
3. Zhang, L.; Zhang, Z.; Luo, Y.; Cao, J.; Xie, R.; Li, S. Integrating satellite-derived climatic and vegetation indices to predict smallholder maize yield using deep learning. *Agricultural and Forest Meteorology* **2021**, *311*, 108666.
4. Desloires, J.; Ienco, D.; Botrel, A. Out-of-year corn yield prediction at field-scale using Sentinel-2 satellite imagery and machine learning methods. *Computers and Electronics in Agriculture* **2023**, *209*, 107807.
5. Chlingaryan, A.; Sukkarieh, S.; Whelan, B. Machine earning approaches for crop yield prediction and nitrogen status estimation in precision agriculture: A review. *Computers and Electronics in Agriculture* **2018**, *151*, 61-69.
6. Liu, H.; Xiong, W.; Mottaleb, K.A.; Krupnik, T.J.; Burgueno, J.; Pequeno, D.N.L.; Wu, W. B. Contrasting contributions of five factors to wheat yield growth in China by process-based and statistical models. *European Journal of Agronomy* **2021**, *130*, 126370.
7. Xiao, G.; Zhang, X.; Niu, Q.; Li, X.; Li, X.; Zhong, L.; Huang, J. Winter wheat yield estimation at the field scale using sentinel-2 data and deep learning. *Computers and Electronics in Agriculture* **2024**, *216*, 108555.
8. Wang, L.A.; Zhou, X.; Zhu, X.; Dong, Z.; Guo, W. Estimation of biomass in wheat using random forest regression algorithm and remote sensing data. *Crop Journal* **2016**, *4*, 212-219.
9. Wu, S.; Yang, P.; Ren, J.; Chen, Z.; Li, H. Regional winter wheat yield estimation based on the WOFOST model and a novel VW-4DEnSRF assimilation algorithm. *Remote Sensing of Environment* **2021**, *255*, 112276.
10. Maimaitijiang, M.; Sagan, V.; Sidike, P.; Hartling, S.; Esposito, F.; Fritschi, F.B. Soybean yield prediction from UAV using multimodal data fusion and deep learning. *Remote Sensing of Environment* **2020**, *237*, 111599.
11. Sankaran, S.; Khot, L.R.; Espinoza, C.Z.; Jarolmasjed, S.; Sathuvalli, V.R.; Vandemark, G.J.; Miklas, P.N.; Carter, A.H.; Pumphrey, M.O., Knowles, R.R.N., Pavek, M.J. Low-altitude, high-resolution aerial imaging systems for row and field crop phenotyping: a review. *European Journal of Agronomy* **2015**, *70*, 112-123.
12. Van Der Meij, B.; Kooistra, L.; Suomalainen, J.; Barel, J.M.; De Deyn, G.B. Remote sensing of plant trait responses to field-based plant-soil feedback using UAV-based optical sensors. *Biogeosciences* **2016**, *14*, 733-749.
13. Zhou, X.; Zheng, H.B.; Xu, X.Q.; He, J.Y.; Ge, X.K.; Yao, X.; Cheng, T.; Zhu, Y.; Cao, W.X.; Tian, Y.C. Predicting grain yield in rice using multi-temporal vegetation indices from UAV-based multispectral and digital imagery. *ISPRS Journal of Photogrammetry Remote Sensing* **2017**, *130*, 246-255.
14. Jay, S.; Baret, F.; Dutartre, D.; Malatesta, G.; Héno, S.; Comar, A.; Weiss, M.; Maupas., F. Exploiting the centimetre resolution of UAV multispectral imagery to improve remote-sensing estimates of canopy structure and biochemistry in sugar beet crops. *Remote Sensing of Environment* **2018**, *231*, 110898.
15. Zhang, Y.; Xia, C.; Zhang, X.; Cheng, X.; Feng, G.; Wang, Y.; Gao, Q. Estimating the maize biomass by crop height and narrowband vegetation indices derived from UAV-based hyperspectral images. *Ecological Indicators* **2021**, *129*, 107985.
16. Corti, M.; Cavalli, D.; Cabassi, G.; Gallina, P.M.; Bechini, L. Does remote and proximal optical sensing successfully estimate maize variables? A review. *European Journal of Agronomy* **2018**, *99*, 37-50.
17. Homolová, L.; Malenovský, Z.; Clevers, J. G. P. W.; García-Santos, G.; Schaepman, M. E. Review of optical-based remote sensing for plant trait mapping. *Ecological Complexity* **2013**, *15*, 1-16.
18. Kerkech, M.; Hafiane, A.; Canals, R. Vine disease detection in UAV multispectral images using optimized image registration and deep learning segmentation approach. *Computers and Electronics in Agriculture* **2020**, *174*, 105446.
19. Foster, A.; Kakani, V.; Mosali, J. Estimation of bioenergy crop yield and N status by hyperspectral canopy reflectance and partial least square regression. *Precision Agriculture* **2017**, *18*, 192-209.

20. Su, X.; Wang, J.; Ding, L.; Lu, J.; Zhang, J.; Yao, X.; Cheng, T.; Zhu, Y.; Cao, W.; Tian, Y. Grain yield prediction using multi-temporal UAV-based multispectral vegetation indices and endmember abundance in rice. *Field Crops Research* **2023**, *299*, 108992.
21. Da Silva, E.E.; Baio, F.H.R.; Teodoro, L.P.R.; da Silva Junior, C.A.; Borges, R.S.; Teodoro, P.E. UAV-multispectral and vegetation indices in soybean grain yield prediction based on in situ observation. *Remote Sensing Applications: Society and Environment* **2020**, *18*, 100318.
22. Zhang, H.; Zhang, Y.; Liu, K.; Lan, S.; Gao, T.; Li, M. Winter wheat yield prediction using integrated Landsat 8 and Sentinel-2 vegetation index time-series data and machine learning algorithms. *Computers and Electronics in Agriculture* **2023**, 213, 108250.
23. le Maire, G.; Francois, C.; Dufrene, E. Towards universal broad leaf chlorophyll indices using PROSPECT simulated database and hyperspectral reflectance measurements. *Remote Sensing of Environment* **2004**, *89*, 1-28.
24. Vogelmann, J. E.; Rock, B. N.; Moss, D.M. Red edge spectral measurements from sugar maple leaves. *International Journal of Remote Sensing* **1993**, *14*, 1563-1575.
25. Gitelson, A.A.; Keydan, G.P.; Merzlyak, M.N. Three-band model for noninvasive estimation of chlorophyll, carotenoids, and anthocyanin contents in higher plant leaves. *Geophysical Research Letters* **2006**, *33*, 1-6.
26. Datt, B. A new reflectance index for remote sensing of chlorophyll content in higher plants: tests using eucalyptus leaves. *Journal of Plant Physiology* **1999**, *154 (1)*, 30-36.
27. Yao, X.; Zhu, Y.; Tian, Y.C.; Feng, W.; Cao, W.X. Exploring hyperspectral bands and estimation indices for leaf nitrogen accumulation in wheat. *International Journal of Remote Sensing* **2010**, *12 (2)*, 89-100.
28. Fitzgerald, G.; Rodriguez, D.; O'Leary, G. Measuring and predicting canopy nitrogen nutrition in wheat using a spectral index - the canopy chlorophyll content index (CCCI). *Field Crops Research* **2010**, *116*, 318-324.
29. Maccioni, A.; Agati, G.; Mazzinghi, P. New vegetation indices for remote measurement of chlorophylls based on leaf directional reflectance spectra. *Journal of Photochemistry and Photobiology B* **2001**, *61 (1-2)*, 52-61.
30. Sims, D.A.; Gamon, J. A. Relationships between leaf pigment content and spectral reflectance across a wide range of species, leaf structures and developmental stages. *Remote Sensing of Environment* **2002**, *81 (2-3)*, 337-354.
31. Bellon-Maurel, V.; Fernandez-Ahumada, E.; Palagos, B.; Roger, J.M.; McBratney, A. Critical review of chemometric indicators commonly used for assessing the quality of the prediction of soil attributes by NIR spectroscopy. *TrAC-Trends in Analytical Chemistry* **2010**, *29*, 1073-1081.
32. Piikki, K.; Söderström, M.; Stadig, H.; Wolters, S. Remote sensing and on-farm experiments for determining in-season nitrogen rates in winter wheat-Options for implementation, model accuracy and remaining challenges. *Field Crops Research* **2022**, *289*, 108742.
33. Li, F.; Miao, Y.; Feng, G.; Yuan, F.; Yue, S.; Gao, X.; Liu, Y.; Liu, B.; Ustin, S.L.; Chen, X.P. Improving estimation of summer maize nitrogen status with red edge-based spectral vegetation indices. *Field Crops Research* **2014**, *157*, 111-123.
34. Zhao, B.; Duan A.; Ata-Ul-Karim, S.T.; Liu, Z.; Chen, Z.; Gong, Z.; Zhang, J.; Xiao, J.; Liu, Z.; Qin, A.; Ning, D. Exploring new spectral bands and vegetation indices for estimating nitrogen nutrition index of summer maize. *European Journal of Agronomy* **2018**, *93*, 113-125.
35. Ata-Ul-Karim, S.T.; Yao, X.; Liu, X.; Cao, W.; Zhu, Y. Development of critical nitrogen dilution curve of Japonica rice in Yangtze River Reaches. *Field Crops Research* **2013**, *149*, 149-158.
36. Roobroeck, D.; Kimutai, G.; Kanampiu, F.; Ng'etich, W.; de Nowina, K.R.; Vanlauwe, B. Effective Striga control and yield intensification on maize farms in western Kenya with N fertilizer and herbicide-resistant variety. *Field Crops Research* **2023**, *296*, 108924.
37. Vijayakumar, S.; Ivany, H.P.; Bazrgar, A.B.; Deen, B.; Thimmanagari, M.; Schneider, K.; Lauzon, J.; Voroney, P.; Thevathasan, N. The effect of land-use conversion from agriculture to perennial biomass crops and nitrogen fertilizer on soil organic carbon stock in southern Ontario, Canada. *Geoderma Regional* **2023**, *35*, e00732.