



Soybean yield prediction from UAV using multimodal data fusion and deep learning



Maitiniyazi Maimaitijiang^a, Vasit Sagan^{a,*}, Paheding Sidike^{a,b}, Sean Hartling^a, Flavio Esposito^c, Felix B. Fritsch^d

^a Department of Earth and Atmospheric Sciences, Saint Louis University, St. Louis, MO 63108, USA

^b Department of Electrical and Computer Engineering, Purdue University Northwest, Hammond, IN 46323, USA

^c Department of Computer Science, Saint Louis University, St. Louis, MO 63108, USA

^d Division of Plant Sciences, University of Missouri, Columbia, MO 65211, USA

ARTICLE INFO

Edited by Marie Weiss

Keywords:

Remote sensing

Yield prediction

Multimodality

Data fusion

Deep learning

Phenotyping

Spatial autocorrelation

ABSTRACT

Preharvest crop yield prediction is critical for grain policy making and food security. Early estimation of yield at field or plot scale also contributes to high-throughput plant phenotyping and precision agriculture. New developments in Unmanned Aerial Vehicle (UAV) platforms and sensor technology facilitate cost-effective data collection through simultaneous multi-sensor/multimodal data collection at very high spatial and spectral resolutions. The objective of this study is to evaluate the power of UAV-based multimodal data fusion using RGB, multispectral and thermal sensors to estimate soybean (*Glycine max*) grain yield within the framework of Deep Neural Network (DNN). RGB, multispectral, and thermal images were collected using a low-cost multi-sensor UAV from a test site in Columbia, Missouri, USA. Multimodal information, such as canopy spectral, structure, thermal and texture features, was extracted and combined to predict crop grain yield using Partial Least Squares Regression (PLSR), Random Forest Regression (RFR), Support Vector Regression (SVR), input-level feature fusion based DNN (DNN-F1) and intermediate-level feature fusion based DNN (DNN-F2). The results can be summarized in three messages: (1) multimodal data fusion improves the yield prediction accuracy and is more adaptable to spatial variations; (2) DNN-based models improve yield prediction model accuracy: the highest accuracy was obtained by DNN-F2 with an R^2 of 0.720 and a relative root mean square error (RMSE%) of 15.9%; (3) DNN-based models were less prone to saturation effects, and exhibited more adaptive performance in predicting grain yields across the Dwight, Pana and AG3432 soybean genotypes in our study. Furthermore, DNN-based models demonstrated consistent performance over space with less spatial dependency and variations. This study indicates that multimodal data fusion using low-cost UAV within a DNN framework can provide a relatively accurate and robust estimation of crop yield, and deliver valuable insight for high-throughput phenotyping and crop field management with high spatial precision.

1. Introduction

Accurate and non-destructive prediction of crop yield over large areas at low cost is critical for grain policy making and food security (Pantazi et al., 2016; Wang et al., 2014), particularly in the context of population growth and climate change (Alexandratos and Bruinsma, 2012; Kang et al., 2009). Additionally, early estimation of yield at field/farm scales, in conjunction with mapping within-field spatial variations of yield, play an important role in crop management in terms of site-specific fertilization, irrigation and pesticide application (Mourtzinis et al., 2013; Schut et al., 2018; Yang et al., 2013). Yield prediction using early-season data can contribute to increasing crop production and

subsequent profit while reducing input resources and environmental pollution (Bausch and Duke, 1996; McBratney et al., 2005; Panda et al., 2010). Furthermore, non-destructive prediction of crop yield with high-accuracy would allow the identification of high-yielding genotypes rapidly and efficiently from a large number of promising genotypes (Elsayed et al., 2017; Yu et al., 2016).

A variety of approaches have been presented and applied for crop yield prediction, such as ground-based field surveys or farmers' expert knowledge (Allen et al., 2002; Geipel et al., 2014), crop-growth models (Batchelor et al., 2002; Thorp et al., 2008) and remote sensing based methods (Tucker et al., 1980; Wang et al., 2014), as well as crop growth models coupled with environmental factors (i.e., hydrological,

* Corresponding author.

E-mail address: vasit.sagan@slu.edu (V. Sagan).

meteorological, soil factors etc.) and remote sensing data (Betbeder et al., 2016; Horie et al., 1992).

With various spatial, temporal and spectral resolutions, satellite data have been broadly used for crop yield prediction over a wide range of geographic locations and scales (from local to national, continental and global) since the 1970s (Battude et al., 2016; Johnson, 2014; Mariotto et al., 2013; Wall et al., 2008). However, the application of satellite data for yield prediction is hampered by coarse spatial resolution and less optimal temporal sampling (Wang et al., 2010; Zaman-Allah et al., 2015). The development of sensor technologies in recent years has significantly promoted the application of Unmanned Aerial Vehicle (UAV) for data acquisition due to its improved spatial, spectral and temporal resolution compared to airborne and satellite platforms (Colomina and Molina, 2014; Sagan et al., 2019b; Sidiike et al., 2018b). UAV platforms have demonstrated their versatility and cost-benefits across a wide array of applications, particularly high-throughput plant phenotyping (Maimaitijiang et al., 2017; Sagan et al., 2019a) and precision agriculture (Honkavaara et al., 2013; Maresma et al., 2018).

Spectral, structural, thermal, and textural information (or features) extracted from UAV-acquired datasets have been utilized for estimation of various plant traits, respectively, throughout previous research. For example, canopy spectral information derived from UAV-based multispectral and hyperspectral imagery was employed for physiological and geometric properties of vegetation (Houborg and Boegh, 2008), as well as accurate estimation of leaf chlorophyll content (Uto et al., 2013), nitrogen concentration (Caturegli et al., 2016; Zheng et al., 2018c) and grain yield for a variety of crops (Du and Noguchi, 2017; Gong et al., 2018; Maresma et al., 2018). Canopy structure information such as crop height derived from Light Detection and Ranging (LiDAR) system or digital photogrammetry-based point clouds, and canopy coverage or vegetation fraction, have been employed for crop biomass (Bendig et al., 2014; Wang et al., 2017), LAI (Leaf Area Index) (Maimaitijiang et al., 2017) and grain yield prediction (Geipel et al., 2014; Yin et al., 2011; Yu et al., 2016). For soybean, positive correlations between canopy structure (i.e., canopy height) and yield have been shown (Cicek et al., 2006; Mansur et al., 1996; Wilcox and Sedyama, 1981), but the relationship between canopy height and yield can also be negative (Panthee et al., 2007). Canopy thermal information has been used for evaluating plant transpiration (an important gauge for crop water status) (Berni et al., 2009b; Ehrler, 1973; Virlet et al., 2014; Zarco-Tejada et al., 2012), monitoring of crop growth and yield (Guo et al., 2016), as well as identifying heat and drought-tolerant genotypes (Reynolds et al., 2009). Since transpiration cooling is in part a function of stomatal conductance, which also is critical for CO₂ movement into leaves and in turn photosynthesis, canopy temperature measurements and/or differences between canopy and air temperatures have long been of interest for crop yield prediction (Idso et al., 1977). Previous studies showed negative relationships between canopy temperature depression (CTD), the difference between canopy temperature and air temperature, and soybean yield (Harris et al., 1984; Hou et al., 2009). UAV imagery-based canopy texture information has also been employed for estimating plant traits such as crop biomass (Zheng et al., 2018b), LAI (Li et al., 2018), chlorophyll content (Lu et al., 2018), and nitrogen concentration (Zheng et al., 2018a).

Recent studies demonstrated that the combination of canopy spectral, structure, thermal and texture information contained in diverse sensor systems has the potential to improve plant trait estimations in a variety of agricultural applications (Maimaitijiang et al., 2017; Shi et al., 2016) over using features from a single sensor. This may be attributed to the fact that estimating plant traits, such as biomass or grain yield, from spectral information is hampered by asymptotic saturation observed from multi and hyperspectral optical sensors that does not account for three-dimensional (3D) structural information, especially among dense and heterogeneous canopies (Greaves et al., 2015; Rischbeck et al., 2016). Fusion of canopy structure and spectral information has been reported to improve plant trait estimations such as

crop biomass (Li et al., 2015; Maimaitijiang et al., 2019; Wang et al., 2017), LAI (Maimaitijiang et al., 2017) and grain yield (Bendig et al., 2015; Geipel et al., 2014). Additionally, combining canopy thermal information with spectral and structure features has demonstrated to increase the robustness of yield prediction under different weather conditions and development stages of crops (Rischbeck et al., 2016). This is because leaf temperature can affect photosynthetic ability while also influencing leaf nutrient status, such as nitrogen and chlorophyll concentration (da Luz and Crowley, 2010; Maimaitijiang et al., 2017; Sharkey, 2005), which are the critical factors that affect grain yield of cultivars (Elsayed et al., 2017; McKinney et al., 1989; Rischbeck et al., 2016). Compared to the use of spectral information from optical sensors, only a few studies have examined the application of thermal information for yield prediction. Especially fusion of canopy thermal information with spectral and structure features for yield prediction is less utilized (Du et al., 2011; Guo et al., 2016; Rischbeck et al., 2016).

Texture information, characterized by spatial variation in pixel intensities within an image, can highlight the structural and geometric features of plant canopies (De Grandi et al., 2009), and helps to suppress saturation of models that are not suitable for high spatial heterogeneity of the landscape (Mutanga and Skidmore, 2004; Sibanda et al., 2017). Although texture information has been widely used for image classification and segmentation in many previous studies (Mongus and Žalik, 2018; Murray et al., 2010; Sidiike et al., 2018a), there are few reported attempts at directly evaluating the potential of image texture in crop yield prediction (Shiu and Chuang, 2019); and investigating the inherent spatial information captured within texture features remains unexplored for yield prediction. Correspondingly, texture information has been applied to estimate crop nitrogen concentration, biomass and LAI (Zheng et al., 2018a; Zheng et al., 2018b), which are important indicators of crop grain yield (Prasertsak and Fukai, 1997; Rochester, 2007; Serrano et al., 2000). Therefore, texture information could potentially be applied to grain yield prediction.

Data fusion from multiple UAV-integrated sensors has been adopted for plant trait estimations in previous studies (Maimaitijiang et al., 2017). However, less is known on the contribution of combining canopy spectral, structure, thermal, and texture information from RGB, multispectral, and thermal sensors to crop grain yield prediction in the context of high-throughput plant phenotyping and precision agriculture.

Many statistical and machine learning (ML) based regression methods, such as Multiple Linear Regression (MLR) (Jin et al., 2017), Partial Least Squares Regression (PLSR) (Rischbeck et al., 2016), Artificial Neural Networks (ANN) (Fieuza et al., 2017), Random Forest Regression (RFR) (Aghighi et al., 2018), and Support Vector Regression (SVR) (Kuwata and Shibasaki, 2016) have been implemented to achieve accurate yield predictions for a variety of crops.

Deep Learning (DL), a subfield of ML, has gained much attention recently within the scientific community. Based on ANN theory, DL can automatically learn representations from data using a multi-layer architecture, which supports complex nonlinear functions that are learned from the hierarchical outputs of the previous layers. DL accuracies can be further improved given large amounts of data, which is further made possible with the advancement of hardware, technology, data optimization and collection. For these reasons, DL methods such as Deep Neural Network (DNN), Convolutional Neural Network (CNN), and Recurrent Neural Network (RNN), have produced definitively higher accuracies across various regression and classification tasks (Ball et al., 2017; Cai et al., 2018; LeCun et al., 2015; Schmidhuber, 2015; Sidiike et al., 2019). Within the context of data fusion applications, DL also demonstrates its capacity to learn features across multiple modalities (e.g. image, text or audio) to improve feature learning for a single modality (Ngiam et al., 2011; Srivastava and Salakhutdinov, 2012). Multimodal DL has widely been used in human activity recognition, medical applications, autonomous systems (i.e., autonomous navigation) (Novak and Riener, 2015; Poria et al., 2016; Ramachandram and

Taylor, 2017; Srivastava and Salakhutdinov, 2012), and remote sensing based applications such as classification (Benedetti et al., 2018), semantic mapping (Audebert et al., 2018) and image segmentation (Valada et al., 2018). However, few studies have employed DL for crop yield prediction (Kuwata and Shibasaki, 2016; You et al., 2017), and UAV-based crop yield prediction using multimodal data fusion within a deep neural network framework has yet to be attempted.

The objectives of this study were to (i) investigate the potential of low-cost multi-sensor UAV in crop yield prediction within the framework of multimodal data fusion and deep neural network, (ii) test the contributions of canopy spectral, structure, thermal, texture information and their combinations to yield prediction, and (iii) evaluate the robustness and adaptability of multimodal data fusion and deep neural network based yield prediction model across different soybean genotypes, and over space.

2. Material

2.1. Test site and experimental setup

The test site was setup at the University of Missouri Bradford Research Center (BRC) near Columbia, Missouri, USA (Fig. 1). Missouri experiences a humid continental climate characterized by four distinct seasons. A weather station at the BRC continuously records weather conditions which provided the weather data over the 2017 growing season for this study. High (116.4 mm) and low (20.2 mm) monthly growing season precipitation totals were recorded for July and August, respectively. The average monthly growing season temperatures peaked at 25.0 °C in July and were lowest at 13.8 °C in October.

The test site design was broken up into four main plot replications for three soybean cultivars – two obsolete, ‘Pana’ and ‘Dwight’, along with one modern, ‘AG3432’. These main plots were then split into irrigated and rainfed sub-plots (either 17 × 21 m or 21 × 27 m in size) according to a randomized block design. All cultivars were planted on May 19, 2017 at a depth of 2.5 cm and row spacing of 0.76 m at a density of 40 seeds/m². Pre-emergence herbicide, sulfentrazone (0.3 kg ha⁻¹), and post-emergence herbicide, sethoxydim (2.6 kg ha⁻¹), were applied to control weeds with manual weed extraction conducted as needed. No fertilizers were used during this study. Drip irrigation was only applied at the end of August, due to ample precipitation throughout the 2017 growing season.

2.2. Data acquisition

2.2.1. Field data collection

Soybean were harvested with a small-plot combine (ALMACO SPC40, ALMACO, Nevada, IA) in October 2017. To increase the spatial resolution of the yield data beyond the size of the water treatment sub-plots within each cultivar and replication, two-row plots of either 8.1 m or 9.1 m length were harvested and weighed separately, resulting in a total of 960 12.3 m² or 13.8 m² small yield plots within the entire field (Fig. 1). Note that, in a low portion of the field (south-central) where standing water soon after planting resulted in a poor soybean stand, about 10 sub-plots were excluded from yield data collection. The grain from each yield plot was weighed and yield was expressed as kg ha⁻¹ and normalized to a moisture content of 13.0%. The statistics of field harvested grain yield are shown in Table 1.

Grain yield differences across water treatments and genotypes were

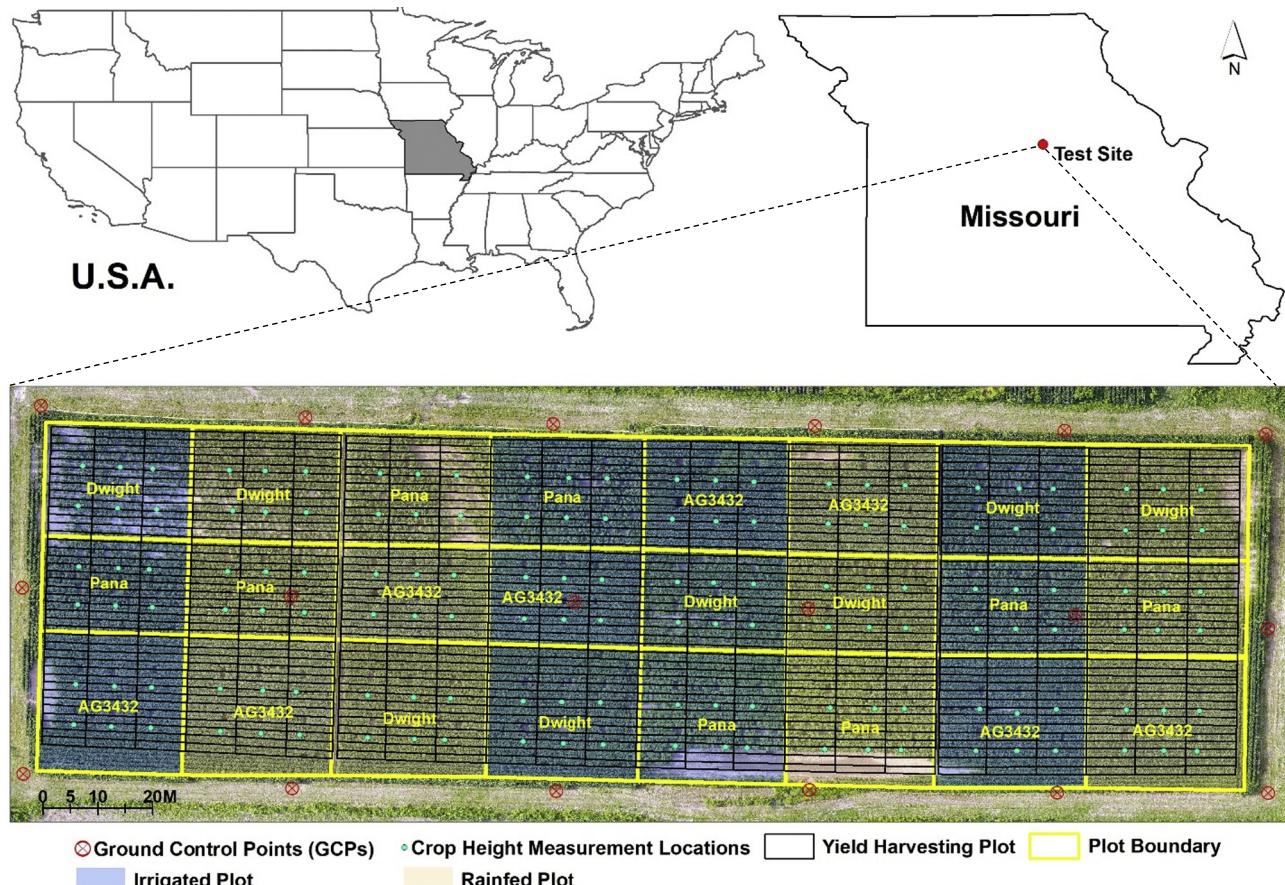


Fig. 1. Test site location. Cyan dots are sampling locations of canopy height; black rectangles represent grain yield harvest plots; red circles with x are locations of Real-Time Kinematic (RTK) GPS survey Ground Control Points (GCPs). Note: this figure was modified after Maimaitjiang et al. (2019). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Table 1

Descriptive statistics for field-measured parameters.

Parameters	No. of samples	Mean	Max.	Min.	SD	CV (%)
Grain yield (kg ha^{-1})	950	3064.7	5446.1	577.2	922.1	30.1
CH (m)	144	0.587	0.801	0.326	0.103	17.5

CH: canopy height; SD: standard deviation; CV: coefficient of variation.

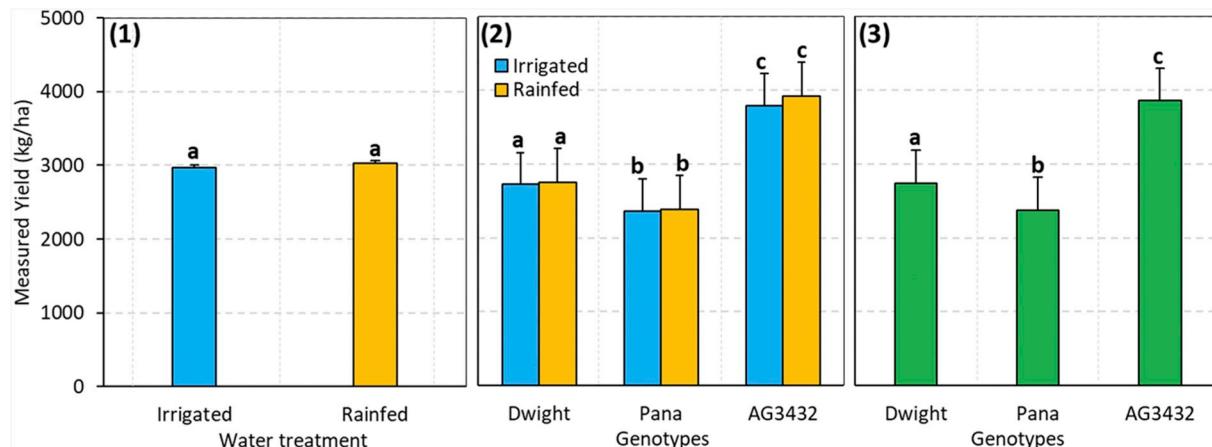


Fig. 2. Mean measured grain yield values under irrigated and rainfed conditions (1), mean measured grain yield values under irrigated and rainfed conditions of each genotype (2), mean measured grain yield values of each genotype (3). ANOVA was carried out, and different letters on the bars indicate significant differences according to the HSD Tukey's test ($\alpha = 0.001$).

assessed using ANOVA (Fig. 2). No significant differences were found in grain yield between irrigated and rainfed treatment within any of the three cultivars, but grain yield of AG3432 was greater than that of Dwight which was greater than that of Pana (Fig. 2(1)). The absence of an irrigation effect was likely due to timely rainfall throughout most of the season as well as due to the unexpected rainfall within hours of the single irrigation that was conducted.

Manual canopy height (CH) measurements were obtained via measuring sticks at 144 locations (Fig. 1) throughout the field (i.e., six CH samples in each plot) in order to validate UAV-derived CH extracted from photogrammetric 3D point clouds. Ground-based CH measurement statistics are outlined in Table 1.

2.2.2. Multi-sensor UAV development and image acquisition

Concurrent acquisition of UAV imagery from multiple sensors is integral to the applications of data fusion for crop monitoring (Maimaitjiang et al., 2017). A cost-effective UAV capable of collecting high-resolution stereo-RGB, multispectral and thermal imagery simultaneously was developed and employed for aerial data collection in this study (Fig. 3). A DJI S1000+ (DJI Technology Co., Shenzhen, China) octocopter frame was chosen as a base for the platform (Fig. 3a). A Pixhawk 2.1 (Hex Technology Limited, Xiamen, China) autopilot was incorporated as an automated flight control system allowing for user-defined waypoint operations and custom mission planning (Fig. 3b). Navigation and image geotagging was provided via Here GNSS GPS (Hex Technology Limited, Xiamen, China) offering concurrent reception of up to three Global Navigation Satellite Systems (GNSS). This system was equipped with an RFD900 long-range telemetry modem (RFDesign, Queensland, Australia) that extends data link connection to over 40 km. System setup, configuration, tuning and mission planning were performed using an open-source ground-station controller, Mission Planner (ArduPilot).

A custom multi-sensor payload tray was designed using a high-pressure fiberglass laminate G10 board and attached to the factory payload mounts beneath the DJI S1000+ (Fig. 3c). Mounting brackets were 3D printed using ABS filament to attach individual sensors to the payload tray. Two Mapir Survey2 RGB cameras were hard mounted to the payload tray at differing angles, nadir and 45 degrees, to facilitate stereoscopic

imaging and 3D modeling. The Mapir Survey2 RGB camera employed a 16-megapixel (MP) Exmore sensor and 23-mm focal length with an 82-degree field-of-view lens. A Mapir Survey-2 NDVI (Red + NIR) camera was also hard mounted to the payload tray, which utilizes the same 16 MP sensor and lens as the Survey2 RGB camera, and differs in its use of a dual-band filter that captures reflected red light (660 nm) in the RGB sensor's red channel and reflected near-infrared (NIR) light (850 nm) in the RGB sensor's blue channel. An additional high-resolution multispectral camera, a Parrot Sequoia, was included into the sensor array and hard mounted to the payload tray. The Sequoia camera incorporated its own integrated GPS/sunshine sensor for image geotagging and signal correction of downwelling radiance (Fig. 3b). In conjunction with a 16 MP rolling shutter RGB camera, the Sequoia featured four 1.2 MP global shutter narrowband spectral cameras that image in the green (550 nm), red (660 nm), red-edge (735 nm) and near-infrared (790 nm) spectral bands. To supplement the various spectral information gathered from the multiple VNIR sensors, a FLIR Vue Pro R 640 thermal imager was incorporated onto the payload tray using a 3D printed custom design 2-axis gimble (Fig. 3c). The FLIR Vue Pro R 640 model employed a 13 mm lens with a 30 Hz frame rate to record temperature measurements in the 7.5–13.5 μm spectrum range with a $\pm 5^\circ\text{C}$ measurement accuracy (Table 2).

The aerial survey was conducted with the multi-sensor UAV under low wind speed and clear skies conditions from 11:00 AM through 2:00 PM local time on July 20 of 2017. Soybean reached the beginning pod stage (R3 stage) (Fehr et al., 1971) on this date, and estimation of final yield based on crop conditions early during this reproductive growth is of great interest. The mission was planned at a flight height of 30 m and a flight speed of 6 m s^{-1} with a front overlap of 90% and a side overlap of 85% for the FLIR Vue Pro R 640 thermal sensor.

During the UAV flight, the surface temperature was measured with a handheld thermal spot imager FLIR TG167 (FLIR Systems, USA), over 8 custom-designed portable wooden boards and 6 black, white and gray-color panels for radiometric calibration of UAV thermal imagery (Fig. 3d). The FLIR TG167 thermal imager advertises a (\pm) 1.5°C measurement accuracy at a resolution of 80×60 pixels. Before surface temperature measurement of field calibration targets, the TG167 bias difference in temperature was measured against a calibrated black body



Fig. 3. UAV systems and integrated sensors. DJI S1000+ Octocopter platform (a), navigation, flight control and telemetry system (b), camera array integrated to the frame (c), handheld thermal spot imager and calibration board for radiometric temperature correction (d), and wooden target used as a ground control point (GCP) (e). Figs. 3 (c) and (d) were modified after [Sagan et al. \(2019a\)](#).

concentric disk radiator – an Everest Interscience Model 1000 calibration source (Everest Interscience Inc., USA) ([Sagan et al., 2019a](#)). The wooden calibration targets (Fig. 3d) were designed with an open squared center to facilitate the matching of ground-based and UAV-based temperature measurement of various ground surface types (e.g. wet/dry soil, grass, crop, etc.), thus eliminating the need for GPS-surveyed ground control points ([Torres-Rua, 2017](#)).

In order to geo-reference the imagery from multi-sensor UAV, 18 Ground Control Points (GCPs) made of portable and cross-patterned wooden targets (80×80 cm) (Fig. 3e) were set up in the field (Fig. 1), and the precise geographic coordinate of each target was measured with a Trimble R8 GNSS Rover that connected to the Real-Time GNSS Network of Missouri state.

2.2.3. Image preprocessing

The Pix4Dmapper software (Pix4D SA, Lausanne, Switzerland) was employed to orthorectify and mosaic the UAV RGB, MS and TIR images.

The geographic coordinates of the 18 GCPs were used during the photogrammetric workflow of Pix4Dmapper to improve the vertical and horizontal accuracy of the output orthomosaics. Point clouds with high density were generated using the SfM technique in Pix4Dmapper along with photogrammetric workflow.

The MS imagery was radiometrically calibrated using the irradiance values captured by the sun-shine sensor (Fig. 3b) during the flight ([Maimaitijiang et al., 2017](#)) along with lens distortion and vignetting issue correction performed within Pix4DMapper software ([Jhan et al., 2018](#)). The raw MS image values were converted to reflectance.

The raw image values captured by the thermal infrared camera (radiometric value) are often required to be converted to surface temperature in degrees Fahrenheit ($^{\circ}\text{F}$) or Celsius ($^{\circ}\text{C}$) by taking into account of object emissivity, atmospheric conditions and distance between the sensor and the target object ([Berni et al., 2009a; Kelly et al., 2019](#)). In addition, possible temperature drift during image collection, particularly for low-cost uncooled thermal sensors, is one of the

Table 2
Basic parameters for the sensors mounted on the UAV.

Sensor name	Sensor type	Spectral region (μm)	Resolution (pixels)	Focal length (mm)	Shutter type	FOV ($\text{H}^{\circ} \times \text{V}^{\circ}$)	Weight (g)
Parrot Sequoia	MS	0.550, 0.660, 0.735, 0.790	1280 × 960	3.98	Global	62.2 × 48.7	107.0
	RGB	N/A	4608 × 3456	4.88	Rolling	64.6 × 50.8	
Mapir Survey2	RGB	N/A	4608 × 3456	2.80	Global	82.0 × 64.5	64.0
Mapir Survey2	NDVI	0.850, 0.660	4608 × 3456	2.80	Global	82.0 × 64.5	64.0
FLIR Vue Pro R 640	TIR	7.50–13.5	640 × 512	13.0	Global	45.0 × 35.0	113.4

MS: multispectral, TIR: thermal infrared, FOV: field of view, H: horizontal, V: vertical.

important error sources that should be considered (Mesas-Carrascosa et al., 2018).

In this study, the procedures applied in our previous research (Sagan et al., 2019a) were employed for thermal image calibration. First, to reduce the influence of possible temperature drift, the thermal camera was powered and pre-heated for about 15 to 20 min on the field before the flight, and the automated NUC (non-uniformity correction) was enabled during the data collection as well (Kelly et al., 2019; Mesas-Carrascosa et al., 2018). Following the camera pre-heat procedure, one radiometric JPEG image of the Model 1000 blackbody (Everest Inter-sciene, Inc.) and the background scene was taken using the thermal camera, which was immediately followed by one 14-bit tiff format image from the same scene. The absolute measurement accuracy of the Model 1000 blackbody is 0.3 °C and thermal sensitivity is 0.1 °C. During post-processing, this radiometric JPEG image was loaded to FLIR Tools software program, which provides options to adjust environmental conditions and imaging targets related parameters, to conduct radiometric conversion. Air temperature and humidity recorded at an on-farm weather station were input, along with object emissivity and target distance (flight height) to extract temperature values, which were exported to a TXT file (Bergkamp et al., 2018; Zhang et al., 2018c). The bias between the image temperature and the blackbody was also determined. At the end, each pixel value (radiometric value) in the 14-bit TIFF image was correlated to the corresponding value in the TXT file, which resulted in a linear model being used as a radiometric conversion equation (Fig. 4(a)). It is worth noting that the FLIR Tools program only supports imagery in radiometric JPEG format and can only export to TXT format, with no batch processing option provided (Sagan et al., 2019a).

The conversion equation in Fig. 4(a) was applied to all 14-bit tiff images collected during the flight, the ground measured temperatures from the various surfaces using the FLIR TG167 thermal imager (Fig. 3(d)) were compared with the corresponding temperatures in the UAV thermal imagery to validate and further convert to a surface level temperature through a linear model (Fig. 4(b)) (Sagan et al., 2019a). Before the surface temperature measurement of each ground thermal target, the TG167 bias difference in temperature was measured against the Model 1000 blackbody calibration source (Sagan et al., 2019a).

3. Methods

3.1. Canopy spectral, structure, thermal and texture information extraction

3.1.1. Canopy spectral information

The raw bands (Green, Red, Red-edge and Near-infrared) from Parrot multispectral orthomosaics were used as canopy spectral features. Additionally, a set of vegetation indices (VIs), previously used for grain yield prediction, were selected (Table 3).

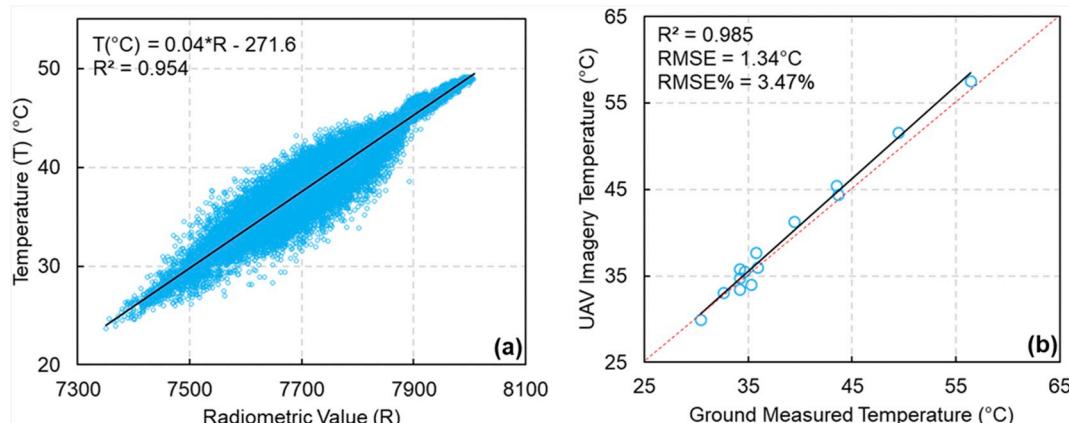


Fig. 4. Conversion of thermal radiometric value (or DN) to temperature (a), UAV imagery temperature validation and conversion (b).

3.1.2. Canopy structure information

Canopy height was extracted from photogrammetric point clouds and used as a canopy structure feature for yield prediction in this study. Prior to plant emergence on May 25, 2017, UAV-based high-resolution RGB imagery was collected in order to create a bare-earth Digital Elevation Model (DEM) using photogrammetric 3D point clouds. A separate 3D point cloud model was built from UAV-based RGB imagery collected on July 20, 2017, representing a Digital Surface Model (DSM) of all objects on the ground - earth and vegetation. From there, a Canopy Height Model (CHM) was obtained from a pixel-wise subtraction of DSM and DEM (Bendig et al., 2013; Zhao et al., 2018). Soil and weed pixels in CHM were excluded using a binary mask created from the SVM-based classification result of orthorectified MS imagery (see Section 3.1.5). The 144 ground-based CH measurements were used to evaluate the accuracy of CHM (Fig. 5c). The distribution of CH reflects different genotypes and heterogeneity in the soybean field (Fig. 5a&b).

Additionally, Vegetation Fraction (VF), which is the percentage of vegetation area per ground surface area, was derived and employed for yield prediction as a valuable index for crop density and structure (i.e. LAI) information (Schirrmann et al., 2016). As shown in Fig. 6, the soybean canopy area was extracted by excluding background soil and weeds from MS imagery using Support Vector Machine (SVM) based classifier following procedures similar to those outlined in Maimaitjiang et al. (2015) (see Section 3.1.5). Classified soybean pixels in each subplot were then divided the total number of pixels in that plot to calculate VF (Yu et al., 2016). The formula for VF calculation is as follows:

$$VF = \frac{\text{Number of crop pixels in the plot}}{\text{Total number of plot pixels}} \quad (1)$$

3.1.3. Canopy thermal information

The normalized relative canopy temperature (NRCT) (Elsayed et al., 2017; Elsayed et al., 2015) was calculated from UAV TIR imagery (Fig. 7) and then used as a thermal feature for yield prediction. The NRCT, which is similar to crop water stress index (CWSI) (Idso et al., 1981), was calculated using the actual canopy temperature, the lowest temperature measured in the whole field trial (T_{min} , lower baseline), and the highest temperature in the whole field trial (T_{max} , upper baseline) (Elsayed et al., 2015). The NRCT has been used to assess and predict water status and grain yield of wheat cultivars in different irrigation regimes (Elsayed et al., 2017; Elsayed et al., 2015; Rischbeck et al., 2016). The NRCT was calculated based on the following equation:

$$NRCT = \frac{T_i - T_{min}}{T_{max} - T_{min}} \quad (2)$$

where T_i represents the canopy temperature of the i th pixel, T_{min} is the lowest temperature measured in the whole field trial, and T_{max} the highest temperature measured in the whole field trial.

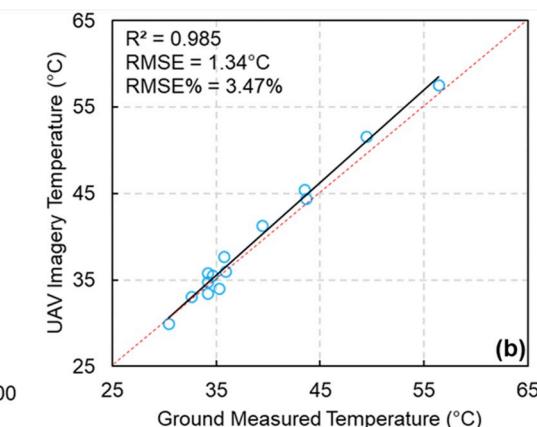


Table 3

Definitions of the features extracted from different sensors and imagery.

Sensor/info.	Features	Formulation	References
MS (Spec. Info.)	Green (G), Red (R), Red-edge (RE), Near-infrared (NIR) Ratio vegetation index Green chlorophyll index Red-edge chlorophyll index Normalized difference vegetation index Green normalized difference vegetation index Green-red vegetation index Normalized difference red-edge Normalized difference red-edge index Simplified canopy chlorophyll content index The enhanced vegetation index Two-band enhanced vegetation index Optimized soil adjusted vegetation index Modified chlorophyll absorption in reflectance index Transformed chlorophyll absorption in reflectance index MCARI/OSAVI TCARI/OSAVI Wide dynamic range vegetation index	The raw value of each band $RVI = NIR/R$ $GCI = (NIR/G) - 1$ $RECI = (NIR/RE) - 1$ $NDVI = (NIR - R) / (NIR + R)$ $GNDVI = (NIR - G) / (NIR + G)$ $GRVI = (G - R) / (G + R)$ $NDRE = (NIR - RE) / (NIR + RE)$ $NDREI = (RE - G) / (RE + G)$ $SCCCI = NDRE / NDVI$ $EVI = 2.5 * (NIR - R) / (1 + NIR - 2.4 * R)$ $EVI2 = 2.5 * (NIR - R) / (NIR + 2.4 * R + 1)$ $OSAVI = (NIR - R) / (NIR - R + L) (L = 0.16)$ $MCARI = [(RE - R) - 0.2 * (RE - G)] * (RE/R)$ $TCARI = 3 * [(RE - R) - 0.2 * (RE - G) * (RE/R)]$ $MCARI/OSAVI$ $TCARI/OSAVI$ $WDRVI = (a * NIR - R) / (a * NIR + R) (a = 0.12)$	/ (Tucker, 1979) (Gitelson et al., 2005) (Gitelson et al., 2005) (Rouse et al., 1974) (Gitelson et al., 2003) (Tucker, 1979) (Gitelson and Merzlyak, 1997) (Hassan et al., 2018) (Rapaport and Varco, 2015) (Huete et al., 2002) (Jiang et al., 2008) (Rondeaux et al., 1996) (Daughtry et al., 2000) (Haboudane et al., 2002) (Daughtry et al., 2000) (Haboudane et al., 2002) (Daughtry et al., 2000) (Haboudane et al., 2002) (Gitelson, 2004)
RGB (Struc. Info.)	Canopy Height (m)	$CH = DSM - DEM$	/
MS (Struc. Info.)	Vegetation fraction (%)	$VF = \frac{\text{Number of crop pixels in the plot}}{\text{Total number of the plot pixels}} * 100$	(Torres-Sánchez et al., 2014)
TIR (Therm. Info.)	Normalized relative canopy temperature index	$NRCT = \frac{T_i - T_{min}}{T_i - T_{max}}$	(Elsayed et al., 2015)
MS + RGB + TIR (Texture Info.)	Gray-level co-occurrence matrix (GLCM)	ME, VA, HO, CO, DI, EN, SE, CO	(Haralick and Shanmugam, 1973)

DSM: digital surface model; DEM: digital elevation model;

3.1.4. Canopy texture information

Texture features were extracted from MS-based Green, Red, Red-edge and NIR bands as well as RGB-based CH and TIR-based NRCT,

then used as input variables for yield prediction. Among several texture algorithms, the commonly used gray level co-occurrence matrix (GLCM) (Haralick and Shanmugam, 1973) was selected to test the

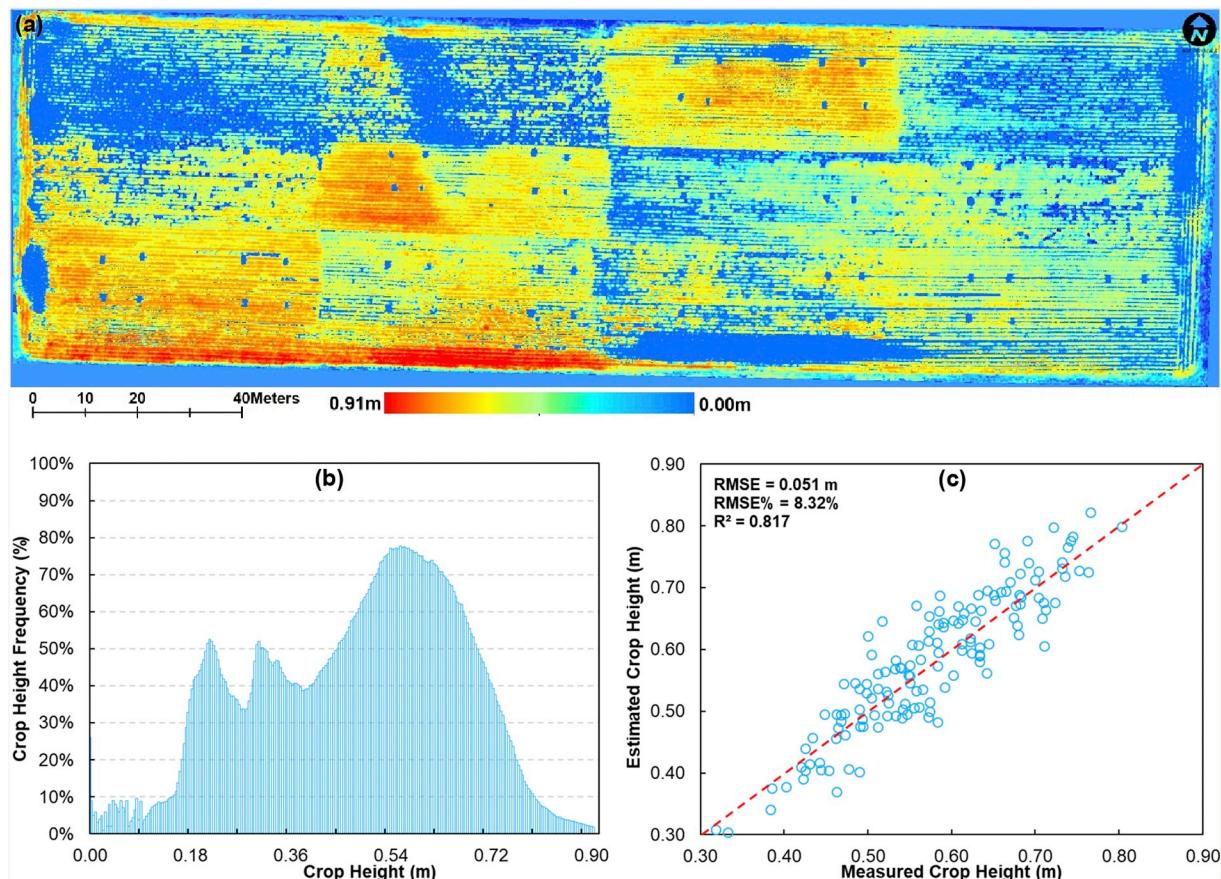


Fig. 5. Crop height estimation map showing the entire field (a), Histogram of crop height distribution (b), and scatter plot of field measured and estimated crop height (c).

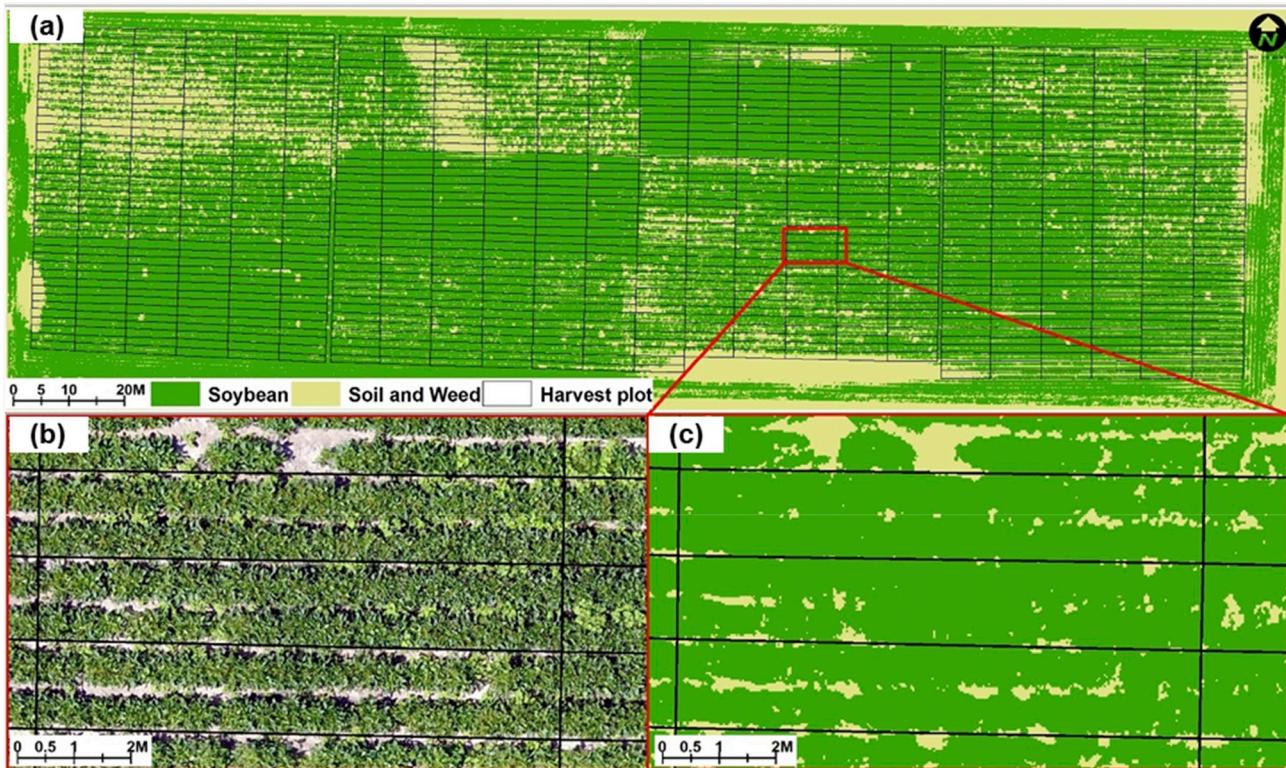


Fig. 6. Vegetation fraction and soil and shadow removal. (a) shows the entire field, (b) is a close-up RGB image, and (c) shows the corresponding vegetation and weeds-soil map of the close-up view.

potential of texture information from UAV imagery for grain yield prediction. The GLCM-based texture measurements including mean (ME), variance (VA), homogeneity (HO), contrast (CO), dissimilarity (DI), entropy (EN), second moment (SE) and correlation (CR), were computed using the ENVI 5.4.1 software. A detailed description of the eight texture metrics used in this study can be found in [Nichol and Sarker \(2011\)](#).

3.1.5. Soil and weed removal

The Support Vector Machine (SVM) ([Cortes and Vapnik, 1995](#)) classifier was employed to identify soybean, weeds and background soil using UAV MS orthomosaics. A binary mask layer was established to exclude weed and background soil pixels from all spectral, structure, thermal, and texture features extracted from UAV imagery ([Table 3](#)) for further processing. The performance of the SVM classifier was evaluated using the confusion matrix and accuracy statistics overall accuracy and Kappa coefficient based on randomly selected independent test samples. The SVM classification resulted in an overall accuracy of

97.4% and a Kappa coefficient of 0.960.

Average pixel values for each of the spectral, structure, thermal, and texture raster layers listed in [Table 3](#) were computed at small yield plot level using zonal statistics in order to relate them with the corresponding grain yield. The Arcpy library and Python 2.7 programming language were used to apply zonal statistics, removes soil/weeds, as well as automate and streamline the extraction of VIs, CH, VF, NRCT raster layers.

3.2. Yield prediction model calibration and validation

3.2.1. Modeling methods

Multimodal data fusion can be conducted at data-level (or early fusion), feature-level, and decision-level (or late fusion), and within a DNN framework, feature-level fusion can be carried out based on input-level features, as well as intermediate-level features ([Ramachandram and Taylor, 2017](#); [Williams et al., 2018](#)). This study employed both input-level and intermediate-level feature fusion within the framework

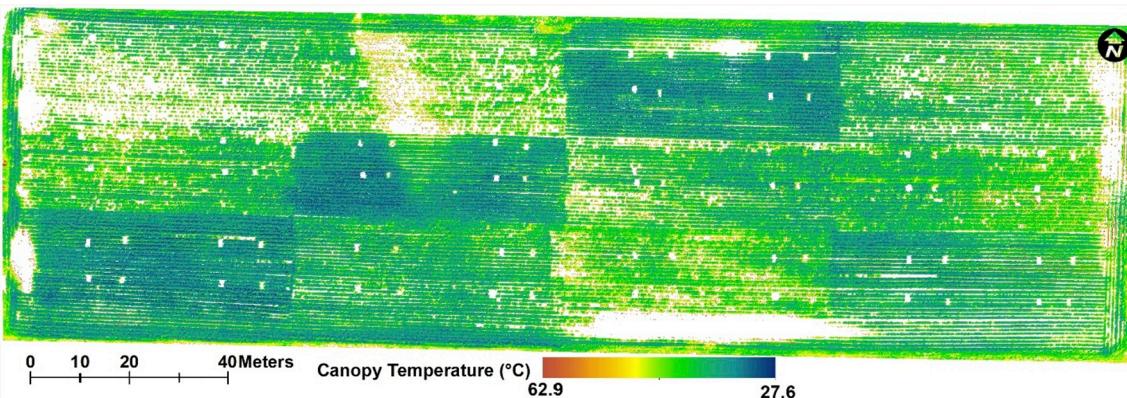


Fig. 7. Canopy temperature distribution map (removed soil/weeds pixels are represented by white color).

of fully-connected feedforward DNN (Biganzoli et al., 1998) to integrate information across different modalities for grain yield prediction. The fully-connected feedforward DNN is based on simply increasing the depth (number of layers) of ANNs (Langkvist et al., 2014), and it has been applied successfully to a range of machine learning problems, typically a variety of classification and regression problems (Cai et al., 2018; Langkvist et al., 2014; Zhang et al., 2018a).

The architecture of the input-level feature fusion DNN, denoted as DNN-F1, proposed in this study is shown in Fig. 8(a), wherein it concatenates multimodal information (i.e., canopy spectral, structure, thermal, and texture features) before feeding them into a DNN. Fig. 8(b) demonstrates the proposed architecture of intermediate-level feature fusion DNN, denoted as DNN-F2, which is composed of four-parallel-stream structure that contains spectral, structure, thermal, and texture feature/modality subnetworks, in which each modality trained using a unimodal DNN, then the four modalities are eventually combined into a joint representation by a concatenation layer.

In our case, the optimized configuration, in terms of the hidden layers and activation functions that achieve a stable and optimal performance is depicted in Fig. 8. The activation function selected in our DNN model was the Rectified Linear Unit (ReLU) with the form of $f(z) = \max(0, z)$ (LeCun et al., 2015). Meanwhile, dropout (Srivastava et al., 2014) was employed to reduce overfitting in the DNN. Furthermore, a batch-normalization (Ioffe and Szegedy, 2015) was also utilized to speed up the learning process by lessening covariance shift.

The DNN-F1 and DNN-F2 were implemented using the Python with TensorFlow (Abadi et al., 2016) and Keras (Keras, 2018) libraries. It is worth noting that, other machine-learning methods, such as PLSR, SVR, and RFR were also employed for yield prediction as baseline comparing with DNN. For a fair and uniform comparison of machine learning methods, 70% randomly selected input features and grain yield data were used as training samples, and the remaining 30% were used as unseen samples to test the performance of the prediction algorithms.

To evaluate the performance of yield prediction model, the coefficients of determination (R^2), root mean square error (RMSE), and relative RMSE (RMSE%) were computed, and can be expressed as follows:

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n-1}} \quad (3)$$

$$\text{RMSE\%} = \frac{\text{RMSE}}{\bar{y}} * 100 \quad (4)$$

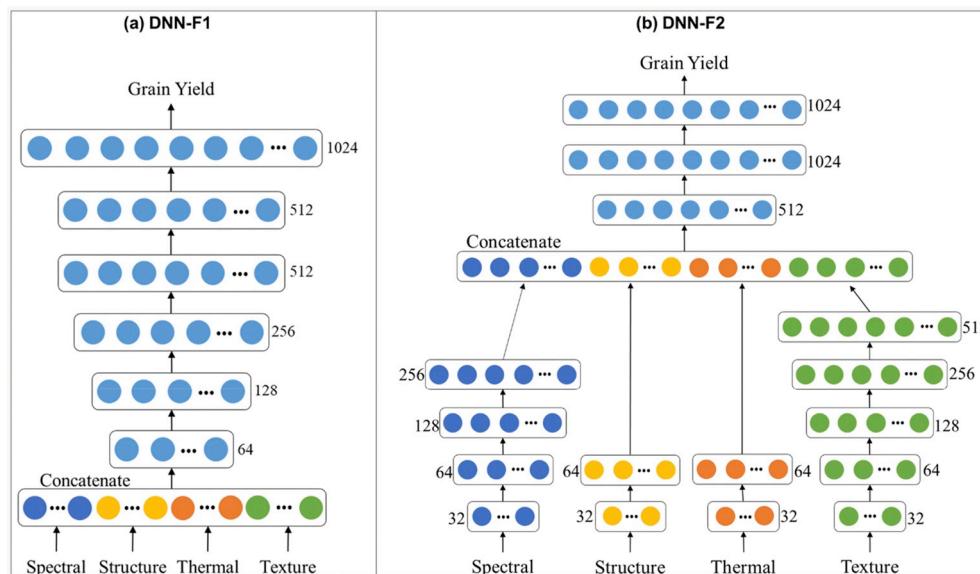


Fig. 8. A schematic illustration of the multimodal deep network architectures used for prediction in this work; input-level feature fusion DNN: DNN-F1 (a); intermediate-level feature fusion DNN: DNN-F2 (b).

where y_i and \hat{y}_i are the measured and the predicted grain yield, respectively. \bar{y} is the mean of measured grain yield, and n is the total number of samples in the testing set.

3.2.2. ANOVA test

To explore the performance and applicability of the models across different soybean genotypes, a one-way analysis of variance (ANOVA) using an honest significant difference (HSD) Tukey test ($\alpha = 0.001$), was carried out to test for significant differences of the estimated grain yield between the three cultivars (Pana, Dwight, and AG3432). The IBM SPSS software (version 24, IBM Corp., Armonk, NY, US) was used to perform the ANOVA test.

3.2.3. Spatial autocorrelation analysis

Spatial independence and location invariance were often assumed for yield prediction in many previous studies (Russello, 2018; You et al., 2017). However, the spatial pattern of grain yield variability is permanently influenced by spatial (terrain attributes, erosion classes and soil properties, irrigation conditions) and temporal (soil pathogens, diseases and production issues in planting the crop) factors (Kravchenko and Bullock, 2000; Peralta et al., 2016). Additionally, crop breeding suffers from intrinsic spatial dependence and has demonstrated non-random variance across a field (Haghigattalab et al., 2017). Therefore, spatial heterogeneity and dependency of crops can occur over many fields, or even within a single field. Without accounting for spatial variation or autocorrelation, prediction models may lead to inaccurate estimations, distorted variance or, more importantly, wrong conclusions (Peralta et al., 2016). For this reason, attention should be paid to the spatial correlation of regression residuals (i.e. prediction error) when examining yield data (Anselin et al., 2004).

To examine the ability of yield prediction models in coping with the influence of spatial variations due to different genotypes, irrigation conditions and other environmental factors over space, the Global Moran's I has been employed to examine the spatial distribution pattern of regression residuals from yield prediction models (Cai et al., 2014; Imran et al., 2015). Differences between measured and predicted grain yield of each plot were used as prediction error or regression residuals. Spatial autocorrelation, or degree of clustering of variables, can be identified and measured using Global Moran's I statistic (Anselin, 1995; Harries, 2006), and has been applied for regression residuals as a diagnostic measure to evaluate model performance over space in many

previous studies (Ghulam et al., 2015; Maimaitjiang et al., 2015; Zhang and Gove, 2005). Global Moran's I values range from -1 to 1, demonstrating negative to positive spatial autocorrelation, respectively, with 0 indicating randomness (Moran, 1950), where higher randomness often indicates better model performance (Anselin et al., 2004; Maimaitjiang et al., 2015). Global Moran's I was calculated for plot-wise prediction error from each yield prediction model using spatial statistics tools in ArcGIS 10.4 software package. The workflow from feature extraction to model building and evaluation was demonstrated in Fig. 9.

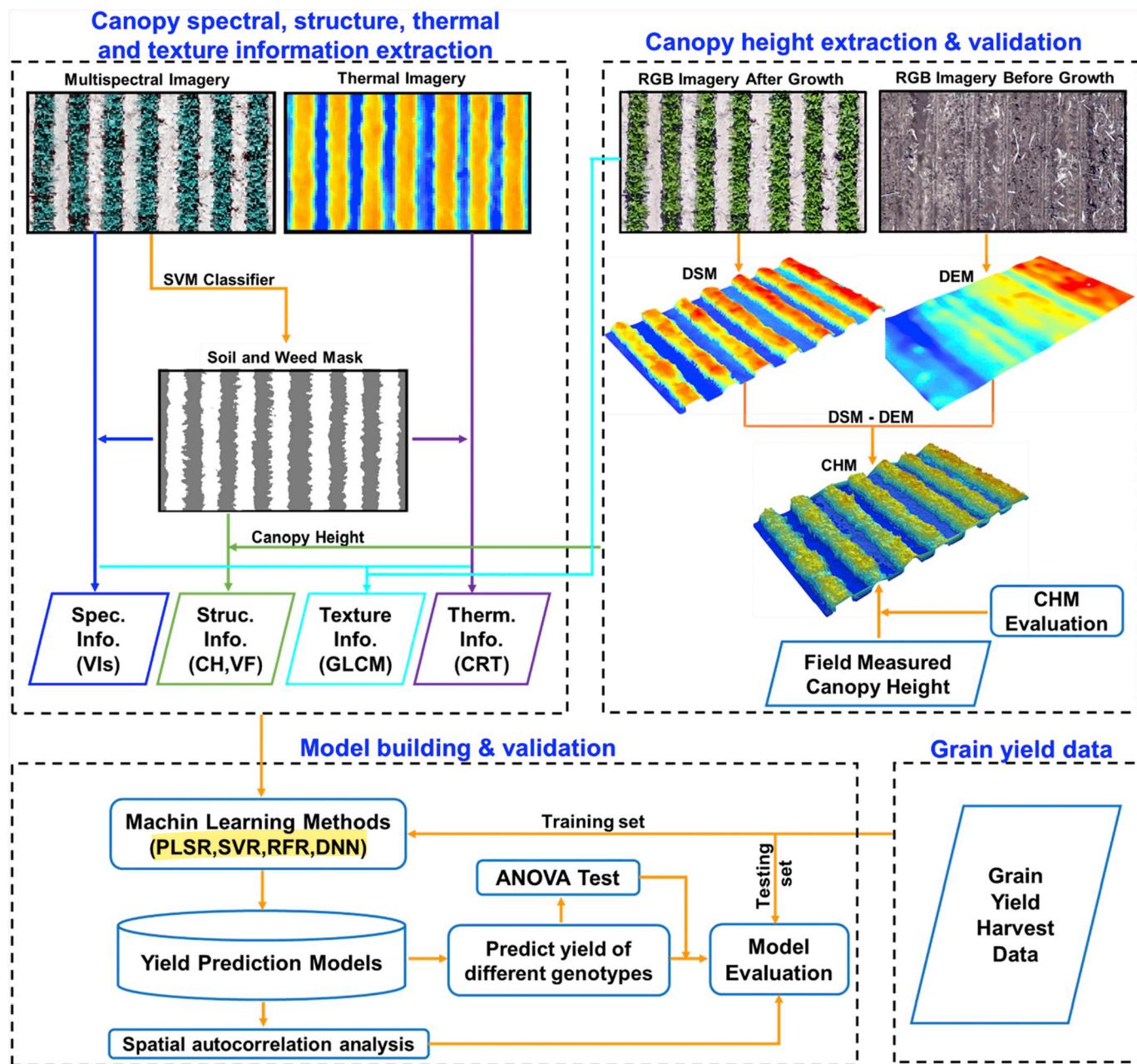


Fig. 9. A workflow diagram of data processing, feature extraction and modeling.

4. Results and discussion

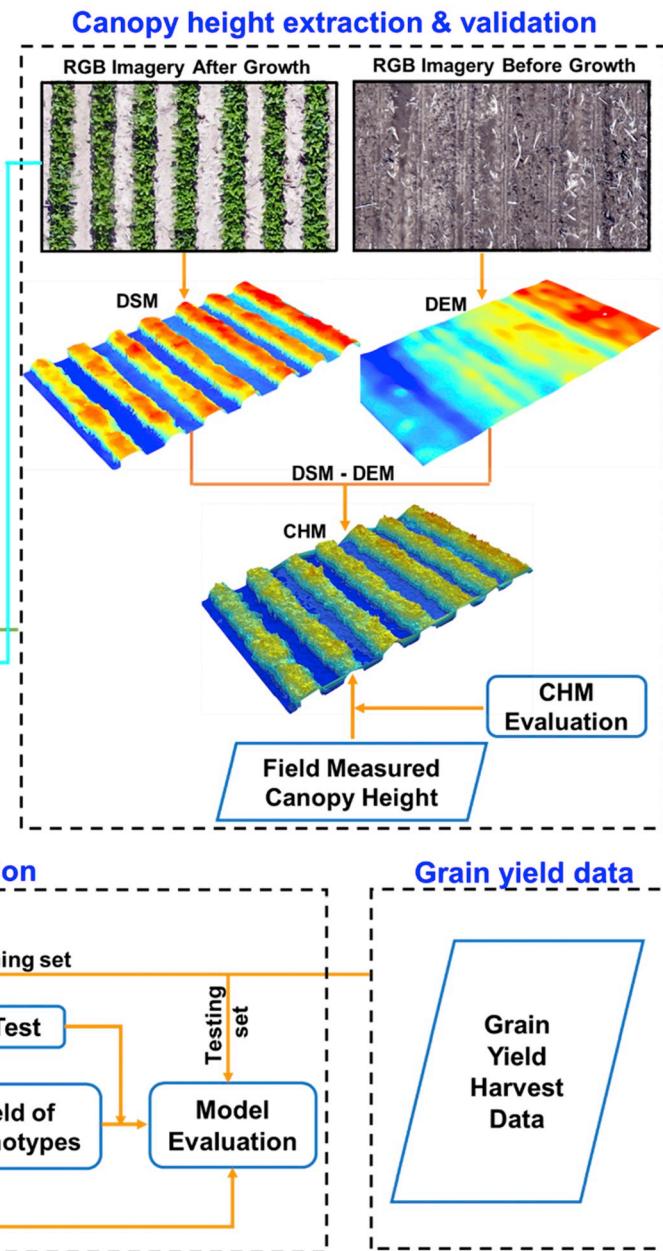
4.1. Modeling and validation of grain yield

Partial Least Squares Regression, SVR, RFR, and DNN-F1 were employed for soybean grain yield prediction using extracted canopy

spectral, structure, thermal, and texture information from either UAV RGB, MS, or TIR as well as from different combinations of sensors and modalities as shown in Table 4. On the other hand, DNN-F2 was only applied for the fusion of two or three sensor combinations.

4.1.1. Contributions of different information/modalities in crop yield prediction

Spectral information obtained from MS sensor resulted in the best estimation of crop yield consistently with all modeling methods with R^2 ranging from 0.464 to 0.515 in the case of using a single sensor



(Table 4). As reported in many previous works, spectral information such as VIs has become the mostly used remote sensing indicator for crop yield prediction due to its stable and superior performance (Ballester et al., 2017; Zhou et al., 2017). RGB sensor-derived canopy structure information (CH and VF) yielded slightly lesser, yet comparable, performance to the MS sensor-based spectral information,

indicating, to some extent, that canopy structure information is a promising alternative to commonly used VIs. Previous research has also demonstrated that UAV imagery-based canopy structure features, such as VF and row length, are good indicators for soybean grain yield prediction in the context of high-throughput phenotyping (Yu et al., 2016). TIR sensor-extracted thermal information presented the poorest performance in grain yield prediction, as compared to MS and RGB sensors with relatively low R^2 which varied from 0.167 to 0.261. Canopy thermal information is associated with leaf water content, pigment concentration and canopy structure properties (da Luz and Crowley, 2010; Elarab et al., 2015; Neinavaz et al., 2016), and is consequently related to crop grain yield (Grassini et al., 2011). However, a wide range of environmental conditions such as background soil, water availability and atmospheric conditions, have major effects on canopy temperature as well (Aubrech et al., 2016; Jones, 1992), which may degrade model performance. Nonetheless, canopy thermal information still exhibits potential in grain yield prediction in this study, which was also recognized in previous studies that examined canopy temperature as a tool for monitoring crop growth and grain yield (Becker and Schmidhalter, 2017; Guo et al., 2016). Systematic and in-depth research is needed to better understand the relationship between canopy thermal information and grain yield, specifically in terms of various factors such as plant species, development stages, environmental conditions, along with their interactions.

As shown in Table 4, the combination of MS-spectral and RGB-structure information led to a noticeable improvement in accuracy, compared to using only a single sensor. RGB-derived canopy structure features such as CH and VF contain independent information on canopy growth and architecture, not obtained from spectral features (Rischbeck et al., 2016; Stanton et al., 2017). In addition, canopy structure information, to some extent, is able to overcome the asymptotic saturation problems inherent to spectral features (Maimaitijiang et al., 2017; Wallace, 2013). Thus, a combination of both spectral and structure information would increase prediction accuracy. Many previous studies have validated the potential of coupled spectral and structure information in grain yield prediction (Bendig et al., 2015; Geipel et al., 2014; Stanton et al., 2017).

The inclusion of TIR-based canopy thermal information in MS-spectral information slightly increased the prediction accuracy with PLSR, RFR and DNN methods, but reduced accuracy when using SVR method. Canopy thermal information tends to add supplementary information to spectral features in grain yield prediction, but to a lesser extent than canopy structure information, indicating that canopy thermal and spectral information tends to suffer information overlap and further impact the predicting capability (Maimaitijiang et al., 2017; Rischbeck et al., 2016). Prediction accuracy was slightly improved across all methods when TIR-based canopy thermal information was added to RGB-derived structure features, demonstrating that canopy thermal information is relatively independent to canopy structure properties (Maimaitijiang et al., 2017), thus adding to the structure information in grain yield prediction.

The fusion of canopy spectral, structure, and thermal information from the MS, RGB and TIR sensors outperformed the combination of any two sensors/modalities regardless of modeling method, with R^2 varying from 0.520 to 0.625. Consistent with previous research, these results demonstrate that canopy spectral, structure, and thermal information provide unique and complementary information that contributes to the grain yield prediction (Rischbeck et al., 2016). Furthermore, the fusion of canopy texture features from all three sensors further boosts prediction accuracy. Canopy texture features could potentially offer additional information associated with spatial canopy architecture and subtle structure characteristics (Colombo et al., 2003). These results further explain, at least in part, the benefits of using texture information in grain yield prediction. Future research should focus on a comprehensive evaluation of the impact of canopy texture on grain yield prediction for different crop species considering different

development stages.

Inclusion of texture information to canopy spectral, structure, and thermal features further improved grain yield prediction accuracy, with R^2 ranging from 0.652 to 0.720. This is likely due to the aforementioned benefits of texture information, which can enhance the spectral, structure, and thermal signature required for accurate prediction of crop grain yield. Improved model performance was also achieved in previous remote sensing imagery classification studies (Mongus and Žalik, 2018; Murray et al., 2010) and crop biomass estimation (Dube and Mutanga, 2015; Eckert, 2012) when using coupled spectral and texture information. Be that as it may, it should be noted that the accuracy improvement was not substantial when combining all spectral, structure, thermal, and texture information as compared to using only multi-sensor based texture information, which is likely attributed to the information homogeneity and redundancy among canopy spectral, structure, thermal, and texture features (Maimaitijiang et al., 2017; Pelizari et al., 2018).

4.1.2. Performance of regression methods in soybean yield prediction

Regarding regression methods, PLSR, SVR, RFR, DNN-F1, and DNN-F2 generally exhibited very close performance in both single-sensor and multi-sensor cases (Table 4). SVR performed slightly better in each single-sensor based grain yield prediction case, while RFR performed best in the case of RGB and TIR fusion. Additionally, relatively higher R^2 , and lower RMSE% were achieved for DNN-based regression with the rest of the multimodal data fusion cases, compared to other methods (Table 4). The best performance for grain yield prediction was observed when using DNN-F2 method with all input features, achieving an R^2 of 0.720 and RMSE% of 15.9%.

As shown in Fig. 10 and Table 4, R^2 gradually increased and RMSE% decreased for all methods with an increasing number of input features, indicating that all tested regression methods, to some extent, are able to cope with an efficient fusion of multimodal input features. However, popular machine learning methods such as RFR and SVM outperformed DNN-based regression when fewer input variables were available, whereas DNN achieved superior performance to other methods when a larger number of input variables were present (e.g. input features > 21 as demonstrated in this study). This is likely due to the fact that DL often exceeds popular machine learning methods when dealing with larger sample size, complex, nonlinear and redundant datasets (Kang and Kang, 2017; LeCun et al., 2015; Schmidhuber, 2015; Zhang et al., 2018a). Furthermore, DNN-F2 consistently outperformed DNN-F1 (Table 4 and Fig. 10), which followed previous studies regarding classification and regression applications (Kang and Kang, 2017; Karpathy et al., 2014; Williams et al., 2018). DNN-F1 employs direct concatenation of low-level features from different modalities at the input-level, whereas DNN-F2 transforms the raw inputs to higher-level representations through multiple DNN streams (Fig. 8(b)). Each stream within DNN-F2 typically alternates linear and nonlinear operations that scale, shift, and skew its input, producing a new representation of the original data (Ramachandram and Taylor, 2017). For this reason, DNN-F2 tends to be less vulnerable to redundancy and dependency between multimodal variables, which consequently leads to better performance (Kang and Kang, 2017).

Although DNN-F1 and DNN-F2 yielded superior performance over PLSR, SVR, and RFR methods when dealing with a larger number of input features, the improvement in grain yield prediction accuracy was not substantial (Table 4). One reason could be that DNN-based methods generally need a large amount of data for training, and the 950 samples used in this study might still not be sufficient. This result agrees with the previous studies (Cai et al., 2018; Zeng et al., 2018; Zhang et al., 2018a) which demonstrated that DNN was only slightly superior to SVM and Random Forests in classification and estimation applications. Future work will examine the ability of more advanced deep learning architectures (e.g., CNN and its variants) to extract better multimodal information for grain yield prediction.

Table 4

Validation statistics of different models for grain yield prediction (The best result for each type of input variable is highlighted in boldface).

Sensor type	Feature type ^a	No. of features	Metrics	PLSR	SVR	RFR	DNN-F1	DNN-F2
MS	Sp	21	RMSE	662.8	630.5	658.1	651.2	–
			RMSE%	22.0	20.9	21.8	21.6	–
			R ²	0.464	0.515	0.472	0.480	–
RGB	St	2	RMSE	700.5	688.2	693.2	689.2	–
			RMSE%	23.2	22.8	23.0	22.8	–
			R ²	0.402	0.422	0.414	0.421	–
TIR	Th	1	RMSE	794.3	778.6	844.6	786.8	–
			RMSE%	26.3	25.8	28.0	26.1	–
			R ²	0.23	0.261	0.167	0.250	–
MS + RGB	SpSt	23	RMSE	629.9	607.4	618.3	597.9	575.3
			RMSE%	20.9	20.1	20.6	19.8	19.1
			R ²	0.516	0.550	0.533	0.567	0.596
MS + TIR	SpTh	22	RMSE	655.2	646.5	647.3	636.3	611.6
			RMSE%	21.7	21.4	21.4	20.8	20.3
			R ²	0.477	0.490	0.489	0.525	0.544
RGB + TIR	StTh	3	RMSE	698.8	680.1	650.2	674.7	671.0
			RMSE%	23.1	22.5	21.5	22.4	22.2
			R ²	0.404	0.436	0.484	0.442	0.451
MS + RGB + TIR	SpStTh	24	RMSE	627.6	601.8	616.6	576.7	554.9
			RMSE%	20.8	19.9	20.4	19.1	18.4
			R ²	0.520	0.558	0.536	0.591	0.625
MS + RGB + TIR	Te	48	RMSE	573.9	556.6	542.0	522.4	–
			RMSE%	19.0	18.4	18.0	17.2	–
			R ²	0.598	0.622	0.642	0.674	–
MS + RGB + TIR	SpStThTe	72	RMSE	534.1	521.0	526.4	506.9	478.9
			RMSE%	17.7	17.3	17.4	16.8	15.9
			R ²	0.652	0.669	0.662	0.691	0.720

^a Sp represents spectral features, St represents structure features, Th represents thermal features, Te represents texture features, SpStTh represents spectral, structure and thermal features, SpStThTe represents spectral, structure, thermal, texture features. The best result in terms of R², RMSE and RMSE% values through different sensors with various modeling methods are boldfaced.

The best estimates of grain yield achieved from each regression method were compared with corresponding observed values using scatterplots (Fig. 11). As expected from the comparable R² and RMSE% values, the distribution pattern of estimated against the observed grain yield for each regression method was very similar. It is worth noting that grain yield was underestimated by all regression methods for samples with grain yields exceeding ~4500 kg ha⁻¹, as demonstrated by the data points covered by the blue circles in Fig. 11. The samples with grain yields exceed ~4500 all belong to the cultivar AG3432, AG3432 showing higher density and canopy height comparing to other two cultivars Dwight and Pana, which could be observed in Fig. 5. This is likely attributed to a saturation issue, which is common to optical remote sensing, particularly at medium to high LAI and yield (Becker-Reshef et al., 2010; Vergara-Díaz et al., 2016). Furthermore, the amount of underestimation by DNN-F1 and DNN-F2 for samples with higher

grain yield was much less compared to other regression methods (Fig. 9), suggesting that DNN is more effective in utilizing multimodal information and less prone to optical saturation issues as it pertains to grain yield prediction.

4.2. Yield prediction for different varieties

The adaptability of yield prediction models to a variety or genotypes is significant, especially with respect to remote sensing based high-throughput plant phenotyping (Li et al., 2014). Therefore, the grain yield prediction models were applied to each variety in order to evaluate their reliability and adaptability. As shown in Table 5, DNN-F2 exhibited the highest prediction accuracy for all three soybean varieties with R² ranging from 0.820 to 0.848, followed by DNN-F1 with R² ranging from 0.780 to 0.838, with PLSR yielding the poorest

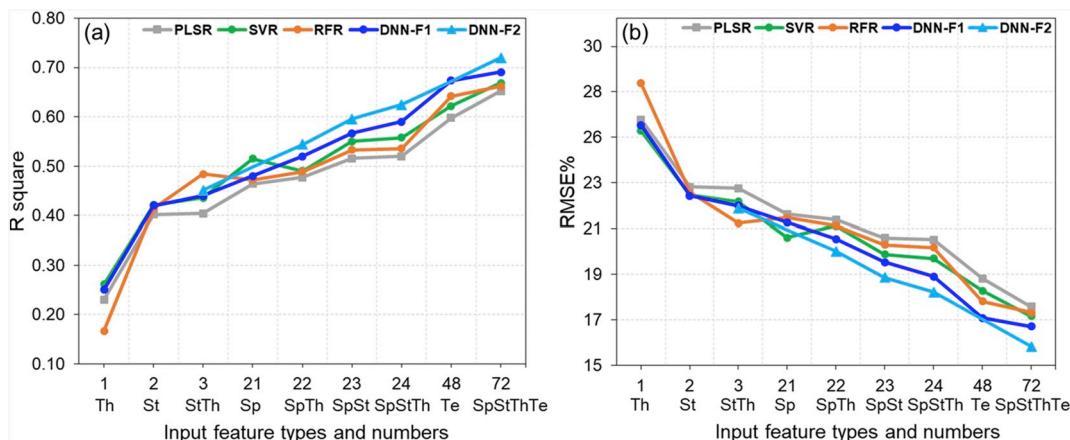


Fig. 10. Grain yield prediction performance of different models with various input feature types and numbers.

Sp represents spectral features, St represents structure features, Th represents thermal features, Te represents texture features, SpStTh represents spectral, structure and thermal features, SpStThTe represents spectral, structure, thermal, texture features.

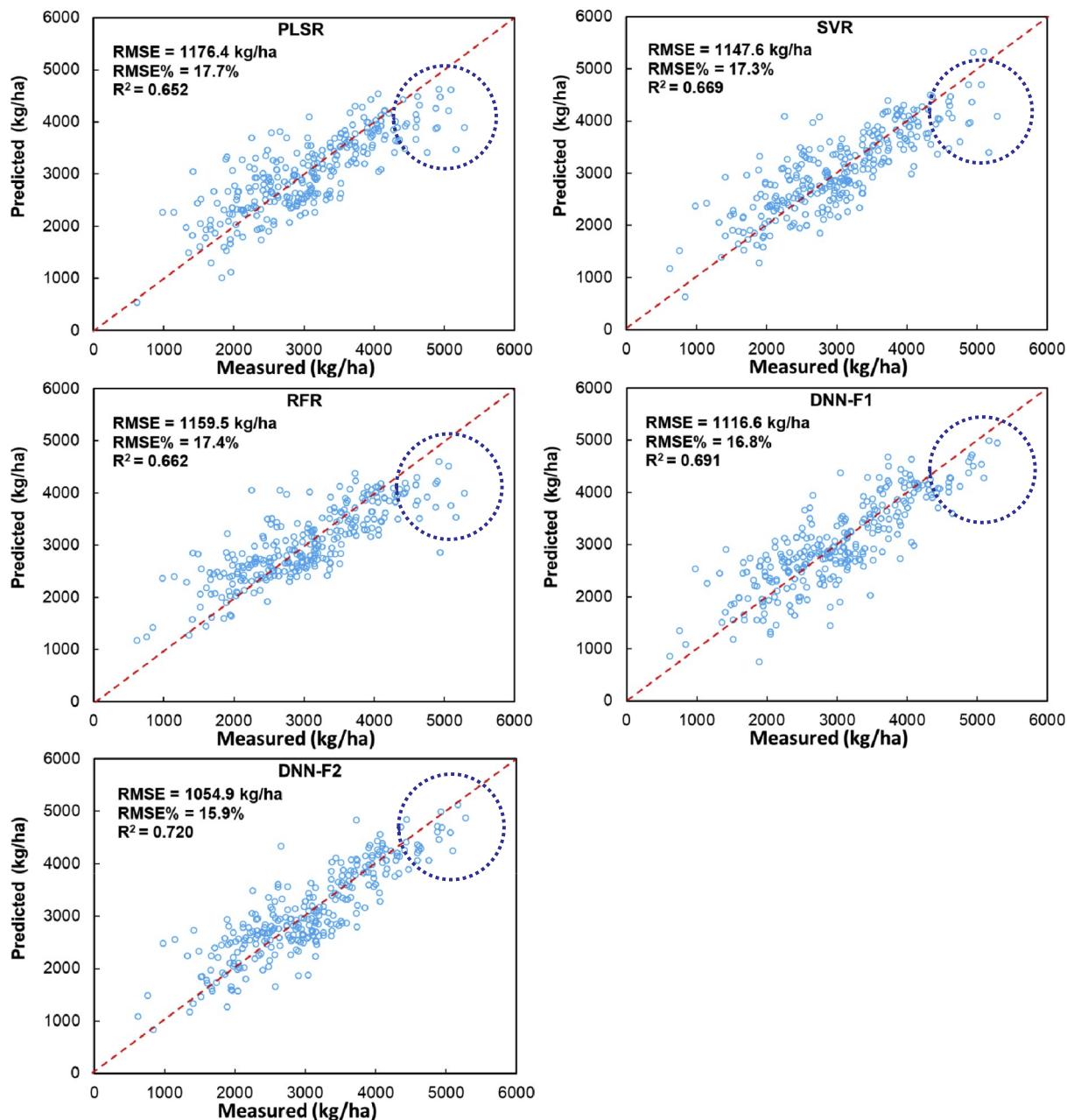


Fig. 11. The cross-validation scatter plots for measured versus prediction grain yield. The red dash lines indicate a 1:1 relationship. The blue dashed-line circles represent samples with grain yields exceeding $\sim 4500 \text{ kg ha}^{-1}$. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

performance with R^2 varying from 0.379 to 0.453. These results highlight that the DNN based models, especially DNN-F2, yielded more adaptive performance for different varieties, compared to the other machine learning methods. In contrast, although the PLSR technique could effectively solve multiple collinearity problems among independent variables, it only simulates the limited relationship between input variables/features and modeling targets (i.e., grain yield), and is unable to map highly non-linear and complex relationship between variables (Nawar et al., 2014; Wang et al., 2018).

The decent performance achieved by RFR for different genotypes with R^2 ranging from 0.693 to 0.786, may be attributed to the combination of multiple diverse individual decision trees. Additionally, RFR has shown to be robust to noise and overfitting (Belgiu and Dragut, 2016; Hu et al., 2018). The modest performance obtained by SVR with R^2 varying from 0.451 to 0.645, demonstrates its complex fitting

Table 5

Performance of each model in predicting grain yield of different genotypes (The method produced best prediction result for each genotype is highlighted in boldface).

Genotypes	Metrics	PLSR	SVR	RFR	DNN-F1	DNN-F2
Dwight	RMSE	463.9	373.7	290.1	271.6	266.3
	RMSE%	16.9	13.6	10.6	9.92	9.73
Pana	R^2	0.453	0.645	0.786	0.813	0.820
	RMSE	524.1	503.5	337.2	312.1	259.1
AG3432	RMSE%	22.1	21.2	14.2	13.2	10.9
	R^2	0.379	0.427	0.743	0.780	0.848
	RMSE	533.3	503.0	376.0	272.9	266.4
	RMSE%	13.9	13.1	9.79	7.11	6.91
	R^2	0.383	0.451	0.693	0.838	0.842

properties, even for non-linear data, through RBF kernel mapping (Liu et al., 2017).

Further research should be targeted towards the implementation and evaluation of multimodal DNN in grain yield prediction, using a larger number of genotypes to explore the suitability for high-throughput grain yield phenotyping.

The performance of each model to detect grain yield differences across genotypes was evaluated using ANOVA with HSD Tukey tests ($\alpha = 0.001$). The results show that predicted grain yields of the three varieties based on each regression model were significantly different (Fig. 12), which is consistent with the measured grain yield for the three genotypes (Fig. 12). These results demonstrate that all methods were effective in distinguishing grain yield differences among genotypes.

4.3. Yield prediction for different water treatment

In addition to investigating the yield prediction across different varieties, grain yield was predicted under irrigated and rainfed conditions as well. As shown in Table 6, DNN-F2 provided the best prediction accuracy with R^2 of 0.845 for irrigated soybean, and 0.824 for rainfed soybean; PLSR yielded the poorest performance with R^2 of 0.500 and 0.572 for irrigated and rainfed soybean.

The prediction accuracies of each model for soybean under irrigated and rainfed conditions are very close and comparable (Table 6). DNN-F2 and SVR models exhibited slightly better performance for irrigated soybean, while DNN-F1, RFR and PLSR yielded slightly higher accuracies for rainfed soybean. Additionally, the performance of each model to detect grain yield differences across water treatments was also evaluated using ANOVA with HSD Tukey tests ($\alpha = 0.001$). The results show that no significant differences were found for predicted grain yields under irrigated and rainfed conditions (Fig. 13), which is consistent with the measured grain yield for irrigated and rainfed soybean (Fig. 13).

4.4. Spatial analysis of yield prediction

In addition to being adaptable to different genotypes, within-field, multi-field, and regional applications require grain yield models to cope with variation and heterogeneity in space caused by differences in soil, irrigation, fertilization and other field conditions that affect plant growth (Laurent et al., 2015; Rischbeck et al., 2016). Therefore, to evaluate the adaptability of the grain yield models over space, Moran's I was calculated for the plot-wise prediction errors from each model to determine their spatial distribution patterns. Conceptualization of spatial relationships (CSR) of input variables such as prediction error is an important parameter for Moran's I calculation (Scott and Janikas, 2010), the commonly used 'inverse distance' and 'contiguity edges corners' methods were employed in this study. The contribution of multimodal data fusion and the performance of different regression methods to site-specific yield prediction were evaluated.

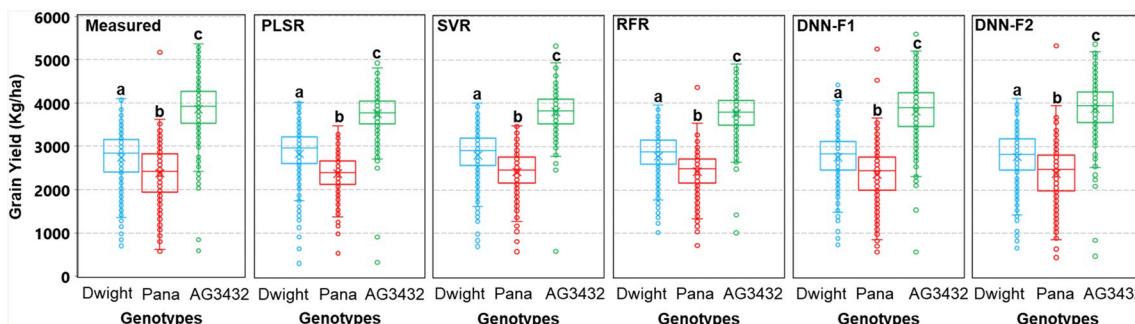


Fig. 12. Boxplot for predicted grain yield from each model and measured grain yield for different genotypes.

ANOVA was carried out, and different letters (a, b, c) on the bars indicate significant differences according to the HSD Tukey's test at the HSD Tukey's test ($\alpha = 0.001$).

Table 6

Performance of each model in predicting grain yield of different water treatments.

Genotypes	Metrics	PLSR	SVR	RFR	DNN-F1	DNN-F2
Irrigated	RMSE	763.7	593.1	457.8	424.9	357.8
	RMSE%	25.6	19.9	15.4	14.2	12.0
	R^2	0.500	0.645	0.760	0.788	0.845
Rainfed	RMSE	604.2	537.7	408.8	377.6	350.4
	RMSE%	20.0	17.8	13.5	12.5	11.6
	R^2	0.572	0.638	0.766	0.797	0.824

4.4.1. The contribution of multimodal data fusion to yield prediction over space

Positive and statistically significant Moran's I indicate the spatial dependent and clustering distribution patterns of variables (Anselin et al., 2004; Peralta et al., 2016). Generally, the distribution of prediction errors from most of the yield prediction models presented spatial clustering patterns at different levels under both CSR methods (Table 7, Table 8 and Fig. 14). Compared to single sensor based yield prediction, multi-sensor/multimodal data fusion resulted in less pronounced spatial clustering patterns with lower Moran's I value regardless of CSR method and regression model, while Moran's I decreased gradually with the increase in input feature numbers or modalities (Fig. 14). Fusion of all modalities led to the weakest clustering pattern of prediction errors from all regression models with Moran's I ranging from 0.028 to 0.144 when using the 'inverse distance' CSR method, and 0.045 to 0.153 when using the 'contiguity edges corners' CSR method (Table 7 and Table 8). This indicates that multimodal data fusion is able to improve model adaptability over space via the weakening of spatial clustering of prediction errors.

It should be noted that the development stage used for soybean yield prediction is a critical factor (Ma et al., 2001) as noted by a number of studies that showed the optimal time window for soybean yield prediction is from blooming to initial seed-filling stages (R2 to R5 stages) (Gao et al., 2018; Wu et al., 2013). Particularly, the initial seed-filling stage (R5) was found to be the best development stage for predicting soybean yield (Ma et al., 2001; Wiegand et al., 1986; Zhang et al., 2019). The remote sensing data in this study was collected at the beginning pod stage (R3 stage). Although this offered a unique opportunity to predict terminal yield at early growth stage, the performance and applicability of soybean yield prediction in the context of multimodal UAV data fusion and deep learning should be evaluated at different development stages, especially at the R5 stage.

4.4.2. Comparison of site-specific yield predictions from different methods

Regarding the performance of different regression methods for site-specific prediction, DNN-F2 outperformed other methods (lowest Moran's I), followed by DNN-F1, whereas PLSR and SVR exhibited diminished performance with higher Moran's I value (Table 7, Table 8

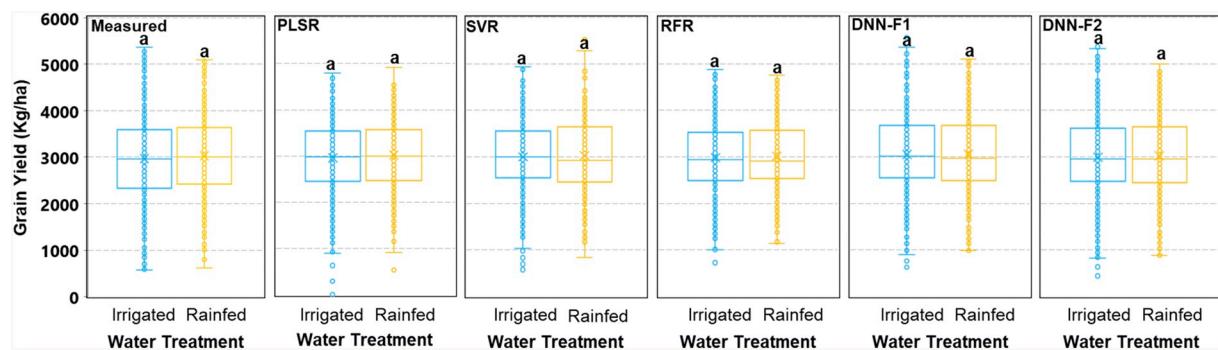


Fig. 13. Boxplot for predicted grain yield from each model and measured grain yield for different water treatments.

ANOVA was carried out, and the letter 'a' on the bars indicate no significant differences according to the HSD Tukey's test at the HSD Tukey's test ($\alpha = 0.001$).

Table 7

Comparison of Moran's I of the prediction errors from different models using various input features (CSR method: Inverse distance method; the lowest Moran's I value for each method is highlighted in boldface).

Modalities/sensors	No. of input features	PLSR		SVR		RFR		DNN-F1		DNN-F2	
		MI	P								
Th	1	0.536	0.000	0.497	0.000	0.276	0.000	0.503	0.000	–	–
St	2	0.479	0.000	0.421	0.000	0.241	0.000	0.363	0.000	–	–
StTh	3	0.462	0.000	0.421	0.000	0.251	0.000	0.180	0.000	0.173	0.000
Sp	21	0.301	0.000	0.294	0.000	0.226	0.000	0.130	0.000	–	–
SpTh	22	0.287	0.000	0.251	0.000	0.230	0.000	0.096	0.022	0.075	0.022
SpSt	23	0.251	0.000	0.275	0.000	0.219	0.000	0.113	0.000	0.076	0.020
SpStTh	24	0.245	0.000	0.238	0.000	0.174	0.000	0.101	0.002	0.053	0.107
Te	48	0.206	0.000	0.144	0.000	0.097	0.000	0.102	0.000	–	–
SpStThTe	72	0.144	0.000	0.095	0.048	0.086	0.000	0.081	0.014	0.028	0.391

MI: Moran's I, P: P-value. Sp represents spectral features, St represents structure features, Th represents thermal features, Te represents texture features, SpStTh represents spectral, structure and thermal features, SpStThTe represents spectral, structure, thermal, texture features.

Table 8

Comparison of Moran's I of the prediction errors from different models using various input features (CSR method: Contiguity edges corners; the lowest Moran's I value for each method is highlighted in boldface).

Modalities/sensors	No. of input features	PLSR		SVR		RFR		DNN-F1		DNN-F2	
		MI	P								
Th	1	0.507	0.000	0.474	0.000	0.319	0.000	0.468	0.000	–	–
St	2	0.479	0.000	0.436	0.000	0.261	0.000	0.372	0.000	–	–
StTh	3	0.460	0.000	0.442	0.000	0.262	0.000	0.17234	0.000	0.149	0.000
Sp	21	0.322	0.000	0.294	0.000	0.222	0.000	0.157	0.000	–	–
SpTh	22	0.3059	0.000	0.282	0.000	0.203	0.000	0.13534	0.000	0.099	0.000
SpSt	23	0.278	0.000	0.290	0.000	0.219	0.000	0.154	0.000	0.088	0.000
SpStTh	24	0.266	0.000	0.252	0.000	0.194	0.000	0.130	0.000	0.079	0.000
Te	48	0.212	0.000	0.153	0.000	0.144	0.000	0.122	0.000	–	–
SpStThTe	72	0.153	0.000	0.103	0.000	0.118	0.000	0.086	0.000	0.045	0.007

MI: Moran's I, P: P-value. Sp represents spectral features, St represents structure features, Th represents thermal features, Te represents texture features, SpStTh represents spectral, structure and thermal features, SpStThTe represents spectral, structure, thermal, texture features.

and Fig. 11). It is worth noting that, for three-modality fusion, prediction errors from the DNN-F2 model under the 'inverse distance' CSR method presented spatial random-distribution pattern, with higher P-value ($P = 0.107$ and $P = 0.391$) indicating statistical non-significance of the corresponding Moran's I (Table 7).

Albeit at a relatively weaker level, most models demonstrated spatial clustering pattern for yield prediction, indicating that the performance of these models was not consistent over space. This is likely due to the spatial independent and location invariant assumption, and the fact that spatial variation and autocorrelation structure was not taken into account in this study (Peralta et al., 2016). Nonetheless, the DNN-F2 model provided yield prediction with randomly distributed regression residuals when using the 'inverse distance' CSR method, and a much weaker

clustering pattern with the 'contiguity edges corners' method. This was in contrast to PLSR, SVR, RFR and DNN-F1 models, indicating the DNN-F2 model outperformed other models in accounting for spatial variations (Anselin et al., 2004; Maimaitjiang et al., 2015). Further investigation regarding the spatial adaptability of DNN-based model in yield prediction should be conducted for assorted crop species, over various field types and larger heterogeneous regions. It is worth noting that PLSR model exhibited a relatively stronger spatial clustering pattern for yield prediction with higher statistically significant Moran's I in both 'inverse distance' and 'contiguity edges corners' methods, indicating inferior performance over space compared to other models. This could be explained by the fact that PLSR simulates the potential linear relationship between input variables/features and modeling targets (i.e. grain yield),

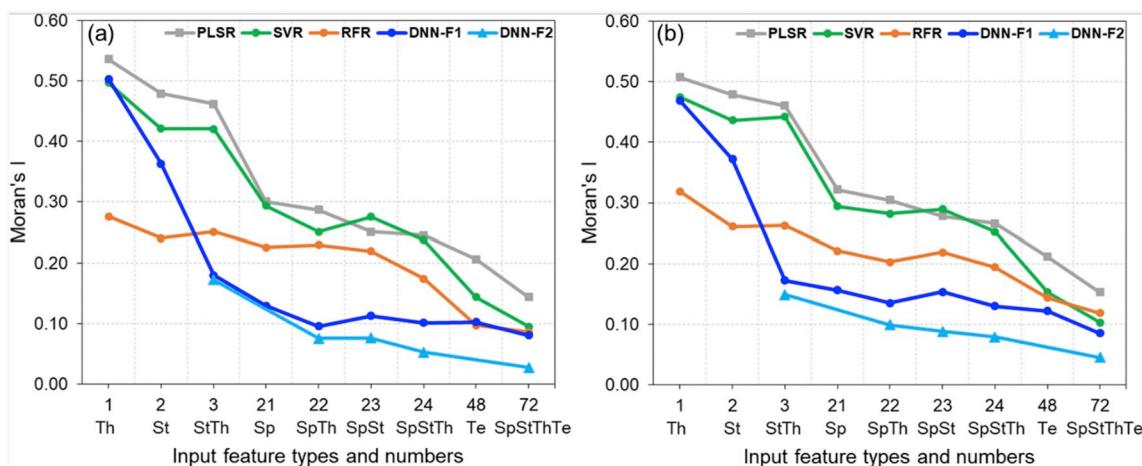


Fig. 14. Grain yield prediction performance of different models with various input feature types and numbers. Moran's value when using 'inverse distance' CSR method (a); Moran's value when using 'contiguity edges corners' CSR method (b). Sp represents spectral features, St represents structure features, Th represents thermal features, Te represents texture features, SpStTh represents spectral, structure and thermal features, SpStThTe represents spectral, structure, thermal, texture features.

but, to some extent, exhibits poor performance over non-linear and complex datasets (Nawar et al., 2014; Wang et al., 2018).

The clustering pattern of prediction errors over space from the PLSR model, and the random-distribution pattern of the DNN-F2 model are

visually evidenced in Fig. 15. Fig. 15(a) displays the spatial distribution of prediction errors from PLSR model, while Fig. 15(b) depicts the prediction error pattern from DNN-F2.

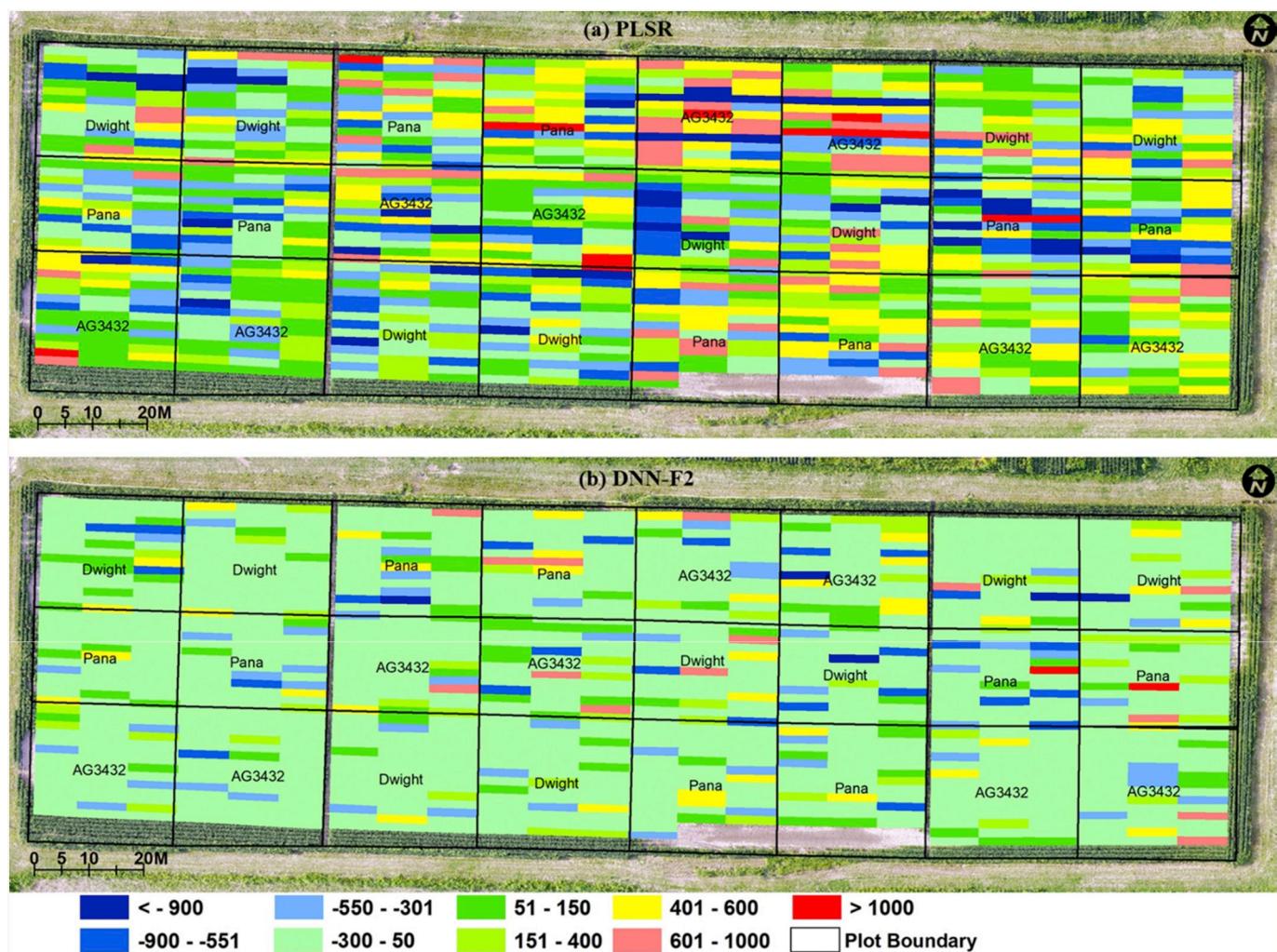


Fig. 15. Spatial distribution of prediction errors (regression residuals) from PLSR model (a), and DNN-F2 model (b); the unit of the values is kg ha⁻¹.

5. Conclusion

The potential of canopy spectral, structure, thermal and texture information derived from low-cost UAV-based multiple sensors for crop yield prediction within the framework of multimodal data fusion and deep learning were investigated. The main conclusions are:

1. Spectral information outperformed structure and thermal features in the case of single sensor/modality based yield prediction. However, RGB-based structure information proved to be a potential alternative to spectral features with comparable prediction accuracy. Although with relatively poorer performance, thermal information illustrated potential in crop yield prediction as well. Texture information from multiple sensors showed great potential for yield prediction with higher prediction accuracy achieved when using all spectral, structure, and thermal features.
2. Multimodal data fusion yielded superior performance for yield prediction over single sensor data, regardless of modeling methods. Additionally, multimodal data fusion contributed to model adaptivity over space with less spatial dependency and variation.
3. DNN-based model outperformed PLSR, SVR, and RFR methods when the number of input features was increased. Furthermore, DNN exhibited strong adaptability to different varieties with high prediction accuracy, along with robustness in accounting for spatial dependency and variation. Moreover, the DNN-based yield prediction model was less prone to saturation effects.
4. Intermediate-level feature fusion DNN (DNN-F2) outperformed input-level feature fusion DNN (DNN-F1) in terms of prediction accuracy, spatial adaptivity and robustness across different genotypes.

The results presented in this study demonstrate the tremendous potential of UAV based crop yield prediction when employing a multimodal data fusion and deep neural network approach. However, to further evaluate the robustness, this approach needs to be tested for different crop types and larger numbers of genotypes at different development stages and environmental conditions.

Author contributions

Vasit Sagan is the primary investigator who developed and framed the initial project and concept that is the foundation upon which this research is based. Maitiniyazi Maimaitijiang, Vasit Sagan and Paheding Sidike conceived the idea and developed and framed research questions. Maitiniyazi Maimaitijiang prepared the original draft. Sean Hartling contributed to the UAV system development and relevant descriptions. Flavio Esposito and Felix B. Fritsch contributed to manuscript review and editing. Felix B. Fritsch managed the Bradford soybean experiment and conducted yield data collection.

Declaration of competing interest

All authors have read and approved this version of the article, and due care has been taken to ensure the integrity of the work. This paper is original work and is currently not being considered for publication in any other peer-reviewed journal. No part of this paper has been published or submitted elsewhere. No conflict of interest exists in the submission of this manuscript.

Acknowledgment

This work was supported in part by the National Science Foundation (IAA-1355406 and IIA-1430427), in part by the National Science Foundation (CNS 1836906, CNS 1647084, and CNS 1908574) and in part by the National Aeronautics and Space Administration (NNX15AK03H). The authors thank Matthew Maimaitiyiming and Kyle

Peterson from Saint Louis University who assisted the UAV imagery collection phase; and the yield harvest team from University of Missouri in Columbia who spent long and strenuous hours collecting data. The authors also would like to thank the editor and the anonymous reviewers for their thoughtful review and constructive comments.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.rse.2019.111599>.

References

- Abadi, M., Barham, P., Chen, J., Chen, Z., Davis, A., Dean, J., Devin, M., Ghemawat, S., Irving, G., Isard, M., 2016. Tensorflow: a system for large-scale machine learning. In: OSDI, pp. 265–283.
- Aghighi, H., Azadbakht, M., Ashourloo, D., Shahrabi, H.S., Radiom, S., 2018. Machine learning regression techniques for the silage maize yield prediction using time-series images of Landsat 8 OLI. *Ieee Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 11, 4563–4577.
- Alexandratos, N., Bruinsma, J., 2012. World Agriculture Towards 2030/2050: The 2012 Revision. ESA Working paper FAO, Rome.
- Allen, R., Hanuschak, G., Craig, M., 2002. Limited Use of Remotely Sensed Data for Crop Condition Monitoring and Crop Yield Forecasting in NASS. US Department of Agriculture, Washington, DC, USA.
- Anselin, L., 1995. Local indicators of spatial association—LISA. *Geogr. Anal.* 27, 93–115.
- Anselin, L., Bongiovanni, R., Lowenberg-DeBoer, J., 2004. A spatial econometric approach to the economics of site-specific nitrogen management in corn production. *Am. J. Agric. Econ.* 86, 675–687.
- Aubrecht, D.M., Helliker, B.R., Goulden, M.L., Roberts, D.A., Still, C.J., Richardson, A.D., 2016. Continuous, long-term, high-frequency thermal imaging of vegetation: uncertainties and recommended best practices. *Agric. For. Meteorol.* 228, 315–326.
- Audebert, N., Le Saux, B., Lefèvre, S., 2018. Beyond RGB: very high resolution urban remote sensing with multimodal deep networks. *ISPRS J. Photogramm. Remote Sens.* 140, 20–32.
- Ball, J.E., Anderson, D.T., Chan, C.S., 2017. Comprehensive survey of deep learning in remote sensing theories, tools, and challenges for the community. *J. Appl. Remote. Sens.* 11, 042609.
- Ballester, C., Hornbuckle, J., Brinkhoff, J., Smith, J., Quayle, W., 2017. Assessment of in-season cotton nitrogen status and lint yield prediction from unmanned aerial system imagery. *Remote Sens.* 9, 1149.
- Batchelor, W.D., Basso, B., Paz, J.O., 2002. Examples of strategies to analyze spatial and temporal yield variability using crop models. *Eur. J. Agron.* 18, 141–158.
- Battude, M., Al Bitar, A., Morin, D., Cros, J., Huc, M., Sicre, C.M., Le Dantec, V., Demarez, V., 2016. Estimating maize biomass and yield over large areas using high spatial and temporal resolution Sentinel-2 like remote sensing data. *Remote Sens. Environ.* 184, 668–681.
- Bausch, W., Duke, H., 1996. Remote sensing of plant nitrogen status in corn. *Transactions of the Asae* 39, 1869–1875.
- Becker, E., Schmidhalter, U., 2017. Evaluation of yield and drought using active and passive spectral sensing systems at the reproductive stage in wheat. *Front. Plant Sci.* 8.
- Becker-Reshef, I., Vermote, E., Lindeman, M., Justice, C., 2010. A generalized regression-based model for forecasting winter wheat yields in Kansas and Ukraine using MODIS data. *Remote Sens. Environ.* 114, 1312–1323.
- Belgiu, M., Dragut, L., 2016. Random forest in remote sensing: a review of applications and future directions. *ISPRS J. Photogramm. Remote Sens.* 114, 24–31.
- Bendig, J., Bolten, A., Bareth, G., 2013. UAV-based imaging for multi-temporal, very high resolution crop surface models to monitor crop growth variability. In: *Photogrammetrie Fernerkundung Geoinformation*, pp. 551–562.
- Bendig, J., Bolten, A., Bennertz, S., Broscheit, J., Eichfuss, S., Bareth, G., 2014. Estimating biomass of barley using crop surface models (CSMs) derived from UAV-based RGB imaging. *Remote Sens.* 6, 10395–10412.
- Bendig, J., Yu, K., Asen, H., Bolten, A., Bennertz, S., Broscheit, J., Gnyp, M.L., Bareth, G., 2015. Combining UAV-based plant height from crop surface models, visible, and near infrared vegetation indices for biomass monitoring in barley. *Int. J. Appl. Earth Obs. Geoinf.* 39, 79–87.
- Benedetti, P., Ienco, D., Gaetano, R., Ose, K., Pensa, R.G., Dupuy, S., 2018. \$ M^3{text{fusion}} \$: a deep learning architecture for multiscale multimodal multitemporal satellite data fusion. *Ieee Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 11, 4939–4949.
- Bergkamp, B., Impa, S., Asebedo, A., Fritz, A., Jagadish, S.K., 2018. Prominent winter wheat varieties response to post-flowering heat stress under controlled chambers and field based heat tents. *Field Crop Res.* 222, 143–152.
- Berni, J.A., Zarco-Tejada, P.J., Suárez Barranco, M.D., Fereres Castiel, E., 2009a. Thermal and Narrow-band Multispectral Remote Sensing for Vegetation Monitoring from an Unmanned Aerial Vehicle. Institute of Electrical and Electronics Engineers.
- Berni, J.A., Zarco-Tejada, P.J., Suárez, L., Fereres, E., 2009b. Thermal and narrowband multispectral remote sensing for vegetation monitoring from an unmanned aerial vehicle. *IEEE Trans. Geosci. Remote Sens.* 47, 722–738.
- Betbeder, J., Fieuza, R., Baup, F., 2016. Assimilation of LAI and dry biomass data from optical and SAR images into an agro-meteorological model to estimate soybean yield.

- Ieee Journal of Selected Topics in Applied Earth Observations and Remote Sensing 9, 2540–2553.
- Biganzoli, E., Boracchi, P., Mariani, L., Marubini, E., 1998. Feed forward neural networks for the analysis of censored survival data: a partial logistic regression approach. *Stat. Med.* 17, 1169–1186.
- Cai, R.H., Yu, D.L., Oppenheimer, M., 2014. Estimating the spatially varying responses of corn yields to weather variations using geographically weighted panel regression. *J. Agric. Resour. Econ.* 39, 230–252.
- Cai, Y.P., Guan, K.Y., Peng, J., Wang, S.W., Seifert, C., Wardlow, B., Li, Z., 2018. A high-performance and in-season classification system of field-level crop types using time-series Landsat data and a machine learning approach. *Remote Sens. Environ.* 210, 35–47.
- Caturegli, L., Corniglia, M., Gaetani, M., Grossi, N., Magni, S., Migliazzi, M., Angelini, L., Mazzoncini, M., Silvestri, N., Fontanelli, M., Raffaelli, M., Peruzzi, A., Volterrani, M., 2016. Unmanned aerial vehicle to estimate nitrogen status of Turfgrasses. *PLoS One* 11.
- Cicek, M.S., Chen, P., Maroof, S., Buss, G.R., 2006. Interrelationships among agronomic and seed quality traits in an interspecific soybean recombinant inbred population. *Crop Sci.* 46, 1253–1259.
- Colombo, R., Bellingeri, D., Fasolini, D., Marino, C.M., 2003. Retrieval of leaf area index in different vegetation types using high resolution satellite data. *Remote Sens. Environ.* 86, 120–131.
- Colomina, I., Molina, P., 2014. Unmanned aerial systems for photogrammetry and remote sensing: a review. *ISPRS J. Photogramm. Remote Sens.* 92, 79–97.
- Cortes, C., Vapnik, V., 1995. Support-vector networks. *Mach. Learn.* 20, 273–297.
- Daughtry, C., Walthall, C., Kim, M., De Colstoun, E.B., McMurtrey III, J., 2000. Estimating corn leaf chlorophyll concentration from leaf and canopy reflectance. *Remote Sens. Environ.* 74, 229–239.
- De Grandi, G.D., Lucas, R.M., Kropacek, J., 2009. Analysis by wavelet frames of spatial statistics in SAR data for characterizing structural properties of forests. *IEEE Trans. Geosci. Remote Sens.* 47, 494–507.
- Du, M., 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.
- Du, W.Y., Zhang, L.D., Hu, Z.F., Shamaila, Z., Zeng, A.J., Song, J.L., Liu, Y.J., Wolfram, S., Joachim, M., He, X.K., 2011. Utilization of thermal infrared image for inversion of winter wheat yield and biomass. *Spectrosc. Spectr. Anal.* 31, 1476–1480.
- Dube, T., Mutanga, O., 2015. Investigating the robustness of the new Landsat-8 Operational Land Imager derived texture metrics in estimating plantation forest aboveground biomass in resource constrained areas. *ISPRS J. Photogramm. Remote Sens.* 108, 12–32.
- Eckert, S., 2012. Improved forest biomass and carbon estimations using texture measures from WorldView-2 satellite data. *Remote Sens.* 4, 810–829.
- Ehrler, W.L., 1973. Cotton leaf temperatures as related to soil water depletion and meteorological factors. *Agron. J.* 65, 404–409.
- Elarab, M., Ticlavilca, A.M., Torres-Rua, A.F., Maslova, I., McKee, M., 2015. Estimating chlorophyll with thermal and broadband multispectral high resolution imagery from an unmanned aerial system using relevance vector machines for precision agriculture. *Int. J. Appl. Earth Obs. Geoinf.* 43, 32–42.
- Elsayed, S., Rischbeck, P., Schmidhalter, U., 2015. Comparing the performance of active and passive reflectance sensors to assess the normalized relative canopy temperature and grain yield of drought-stressed barley cultivars. *Field Crop Res.* 177, 148–160.
- Elsayed, S., Elhoweity, M., Ibrahim, H.H., Dewir, Y.H., Migdadi, H.M., Schmidhalter, U., 2017. Thermal imaging and passive reflectance sensing to estimate the water status and grain yield of wheat under different irrigation regimes. *Agric. Water Manag.* 189, 98–110.
- Fehr, W., Caviness, C., Burmood, D., Pennington, J., 1971. Stage of development descriptions for soybeans, Glycine Max (L.) Merrill 1. *Crop Sci.* 11, 929–931.
- Fieuza, R., Sicre, C.M., Baup, F., 2017. Estimation of corn yield using multi-temporal optical and radar satellite data and artificial neural networks. *Int. J. Appl. Earth Obs. Geoinf.* 57, 14–23.
- Gao, F., Anderson, M., Daughtry, C., Johnson, D., 2018. Assessing the variability of corn and soybean yields in central Iowa using high spatiotemporal resolution multi-satellite imagery. *Remote Sens.* 10, 1489.
- Geipel, J., Link, J., Claupein, W., 2014. Combined spectral and spatial modeling of corn yield based on aerial images and crop surface models acquired with an unmanned aircraft system. *Remote Sens.* 6, 10335–10355.
- Ghulam, A., Ghulam, O., Maimaitjiang, M., Freeman, K., Porton, I., Maimaitiyiming, M., 2015. Remote sensing based spatial statistics to document tropical rainforest transition pathways. *Remote Sens.* 7, 6257–6279.
- Gitelson, A.A., 2004. Wide dynamic range vegetation index for remote quantification of biophysical characteristics of vegetation. *J. Plant Physiol.* 161, 165–173.
- Gitelson, A.A., Merlyak, M.N., 1997. Remote estimation of chlorophyll content in higher plant leaves. *Int. J. Remote Sens.* 18, 2691–2697.
- Gitelson, A.A., Gritz, Y., Merlyak, M.N., 2003. Relationships between leaf chlorophyll content and spectral reflectance and algorithms for non-destructive chlorophyll assessment in higher plant leaves. *J. Plant Physiol.* 160, 271–282.
- Gitelson, A.A., Vina, A., Ciganda, V., Rundquist, D.C., Arkebauer, T.J., 2005. Remote estimation of canopy chlorophyll content in crops. *Geophys. Res. Lett.* 32.
- Gong, Y., Duan, B., Fang, S.H., Zhu, R.S., Wu, X.T., Ma, Y., Peng, Y., 2018. Remote estimation of rapeseed yield with unmanned aerial vehicle (UAV) imaging and spectral mixture analysis. *Plant Methods* 14.
- Grassini, P., Thorburn, J., Burr, C., Cassman, K.G., 2011. High-yield irrigated maize in the Western US Corn Belt: I. On-farm yield, yield potential, and impact of agronomic practices. *Field Crop Res.* 120, 142–150.
- Greaves, H.E., Vierling, L.A., Eitel, J.U.H., Boelman, N.T., Magney, T.S., Prager, C.M., 2015. Estimating aboveground biomass and leaf area of low-stature Arctic shrubs with terrestrial LiDAR. *Remote Sens. Environ.* 164, 26–35.
- Guo, J.X., Tian, G.L., Zhou, Y., Wang, M., Ling, N., Shen, Q.R., Guo, S.W., 2016. Evaluation of the grain yield and nitrogen nutrient status of wheat (*Triticum aestivum* L.) using thermal imaging. *Field Crop Res.* 196, 463–472.
- Haboudane, D., Miller, J.R., Tremblay, N., Zarco-Tejada, P.J., Dextraze, L., 2002. Integrated narrow-band vegetation indices for prediction of crop chlorophyll content for application to precision agriculture. *Remote Sens. Environ.* 81, 416–426.
- Haghghatalab, A., Crain, J., Mondal, S., Rutkoski, J., Singh, R.P., Poland, J., 2017. Application of geographically weighted regression to improve grain yield prediction from unmanned aerial system imagery. *Crop Sci.* 57, 2478–2489.
- Haralick, R.M., Shanmugam, K., 1973. Textural features for image classification. *IEEE Transactions on systems, man, and cybernetics* 610–621.
- Harries, K., 2006. Extreme spatial variations in crime density in Baltimore County, MD. *Geoforum* 37, 404–416.
- Harris, D., Schapaugh, W., Kanemasu, E., 1984. Genetic diversity in soybeans for leaf canopy temperature and the association of leaf canopy temperature and yield 1. *Crop Sci.* 24, 839–842.
- Hassan, M.A., Yang, M., Rasheed, A., Jin, X., Xia, X., Xiao, Y., He, Z., 2018. Time-series multispectral indices from unmanned aerial vehicle imagery reveal senescence rate in bread wheat. *Remote Sens.* 10, 809.
- Honkavaara, E., Saari, H., Kaivosoja, J., Pöllönen, I., Hakala, T., Litkey, P., Mäkinen, J., Pesonen, L., 2013. Processing and assessment of spectrometric, stereoscopic imagery collected using a lightweight UAV spectral camera for precision agriculture. *Remote Sens.* 5, 5006–5039.
- Horie, T., Yajima, M., Nakagawa, H., 1992. Yield forecasting. *Agric. Syst.* 40, 211–236.
- Hou, A., Chen, P., Alloatti, J., Li, D., Mozzoni, L., Zhang, B., Shi, A., 2009. Genetic variability of seed sugar content in worldwide soybean germplasm collections. *Crop Sci.* 49, 903–912.
- Houborg, R., Boegh, E., 2008. Mapping leaf chlorophyll and leaf area index using inverse and forward canopy reflectance modeling and SPOT reflectance data. *Remote Sens. Environ.* 112, 186–202.
- Hu, S.B., Liu, H.Z., Zhao, W.J., Shi, T.Z., Hu, Z.W., Li, Q.Q., Wu, G.F., 2018. Comparison of machine learning techniques in inferring phytoplankton size classes. *Remote Sens.* 10.
- Huete, A., Didan, K., Miura, T., Rodriguez, E.P., Gao, X., Ferreira, L.G., 2002. Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Remote Sens. Environ.* 83, 195–213.
- Idso, S., Jackson, R., Pinter Jr., P., Reginato, R., Hatfield, J., 1981. Normalizing the stress-degree-day parameter for environmental variability. *Agric. Meteorol.* 24, 45–55.
- Idso, S.B., Jackson, R.D., Reginato, R.J., 1977. Remote-sensing of crop yields. *Science* 196, 19–25.
- Imran, M., Stein, A., Zurita-Milla, R., 2015. Using geographically weighted regression kriging for crop yield mapping in West Africa. *Int. J. Geogr. Inf. Sci.* 29, 234–257.
- Ioffe, S., Szegedy, C., 2015. Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift. arXiv preprint arXiv:1502.03167.
- Jhan, J.P., Rau, J.Y., Haala, N., 2018. Robust and adaptive band-to-band image transform of UAS miniature multi-lens multispectral camera. *ISPRS J. Photogramm. Remote Sens.* 137, 47–60.
- Jiang, Z., Huete, A.R., Didan, K., Miura, T., 2008. Development of a two-band enhanced vegetation index without a blue band. *Remote Sens. Environ.* 112, 3833–3845.
- Jin, Z., Azzari, G., Lobell, D.B., 2017. Improving the accuracy of satellite-based high-resolution yield estimation: a test of multiple scalable approaches. *Agric. For. Meteorol.* 247, 207–220.
- Johnson, D.M., 2014. An assessment of pre- and within-season remotely sensed variables for forecasting corn and soybean yields in the United States. *Remote Sens. Environ.* 141, 116–128.
- Jones, H., 1992. Plants and Microclimate, 2nd edn. Cambridge University Press, Cambridge, UK.
- Kang, H.-W., Kang, H.-B., 2017. Prediction of crime occurrence from multi-modal data using deep learning. *PLoS One* 12, e0176244.
- Kang, Y., Khan, S., Ma, X., 2009. Climate change impacts on crop yield, crop water productivity and food security—a review. *Prog. Nat. Sci.* 19, 1665–1674.
- Karpathy, A., Toderici, G., Shetty, S., Leung, T., Sukthankar, R., Fei-Fei, L., 2014. Large-scale video classification with convolutional neural networks. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 1725–1732.
- Kelly, J., Kljun, N., Olsson, P.-O., Mihai, L., Liljeblad, B., Weslien, P., Klemetsson, L., Eklundh, L., 2019. Challenges and best practices for deriving temperature data from an uncalibrated UAV thermal infrared camera. *Remote Sens.* 11, 567.
- Keras, 2018. Keras: The Python Deep Learning Library. Available online. <https://keras.io/>, Accessed date: 25 December 2018.
- Kravchenko, A.N., Bullock, D.G., 2000. Correlation of corn and soybean grain yield with topography and soil properties. *Agron. J.* 92, 75–83.
- Kuwata, K., Shibasaki, R., 2016. Estimating corn yield in the United States with modis evi and machine learning methods. *ISPRS Annals of Photogrammetry, Remote Sensing & Spatial Information Sciences* 3.
- Langkvist, M., Karlsson, L., Loutfi, A., 2014. A review of unsupervised feature learning and deep learning for time-series modeling. *Pattern Recogn. Lett.* 42, 11–24.
- Laurent, A., Loyce, C., Makowski, D., Pelzer, E., 2015. Using site-specific data to estimate energy crop yield. *Environ. Model Softw.* 74, 104–113.
- LeCun, Y., Bengio, Y., Hinton, G., 2015. Deep learning. *nature* 521, 436.
- Li, D., Gu, X.F., Pang, Y., Chen, B.W., Liu, L.X., 2018. Estimation of forest aboveground biomass and leaf area index based on digital aerial photograph data in Northeast China. *Forests* 9.
- Li, L., Zhang, Q., Huang, D.F., 2014. A review of imaging techniques for plant phenotyping. *Sensors* 14, 20078–20111.

- Li, W., Niu, Z., Wang, C., Huang, W.J., Chen, H.Y., Gao, S., Li, D., Muhammad, S., 2015. Combined use of airborne LiDAR and satellite GF-1 data to estimate leaf area index, height, and aboveground biomass of maize during peak growing season. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 8, 4489–4501.
- Liu, H.Z., Shi, T.Z., Chen, Y.Y., Wang, J.J., Fei, T., Wu, G.F., 2017. Improving spectral estimation of soil organic carbon content through semi-supervised regression. *Remote Sens.* 9.
- Lu, B., He, Y., Liu, H.H., 2018. Mapping vegetation biophysical and biochemical properties using unmanned aerial vehicles-acquired imagery. *Int. J. Remote Sens.* 39, 5265–5287.
- da Luz, B.R., Crowley, J.K., 2010. Identification of plant species by using high spatial and spectral resolution thermal infrared (8.0–13.5 μm) imagery. *Remote Sens. Environ.* 114, 404–413.
- Ma, B., Dwyer, L.M., Costa, C., Cober, E.R., Morrison, M.J., 2001. Early prediction of soybean yield from canopy reflectance measurements. *Agron. J.* 93, 1227–1234.
- Maimaitijiang, M., Ghulam, A., Sandoval, J.S.O., Maimaitiyiming, M., 2015. Drivers of land cover and land use changes in St. Louis metropolitan area over the past 40 years characterized by remote sensing and census population data. *Int. J. Appl. Earth Obs. Geoinf.* 35, 161–174.
- Maimaitijiang, M., Ghulam, A., Sidike, P., Hartling, S., Maimaitiyiming, M., Peterson, K., Shavers, E., Fishman, J., Peterson, J., Kadam, S., Burken, J., Fritsch, F., 2017. Unmanned Aerial System (UAS)-based phenotyping of soybean using multi-sensor data fusion and extreme learning machine. *ISPRS J. Photogramm. Remote Sens.* 134, 43–58.
- Maimaitijiang, M., Sagan, V., Sidike, P., Maimaitiyiming, M., Hartling, S., Peterson, K.T., Maw, M.J., Shakoor, N., Mockler, T., Fritsch, F.B., 2019. Vegetation Index Weighted Canopy Volume Model (CVMVI) for soybean biomass estimation from Unmanned Aerial System-based RGB imagery. *ISPRS J. Photogramm. Remote Sens.* 151, 27–41.
- Mansur, L., Orf, J.H., Chase, K., Jarvik, T., Cregan, P., Lark, K., 1996. Genetic mapping of agronomic traits using recombinant inbred lines of soybean. *Crop Sci.* 36, 1327–1336.
- Maresma, A., Ariza, M., Martinez, E., Lloveras, J., Martinez-Casanovas, J.A., 2018. Analysis of vegetation indices to determine nitrogen application and yield prediction in maize (*Zea mays* L.) from a standard UAV service (vol 9, 648, 2017). *Remote Sens.* 10.
- Mariotti, I., Thenkabail, P.S., Huete, A., Slonecker, E.T., Platonov, A., 2013. Hyperspectral versus multispectral crop-productivity modeling and type discrimination for the HypsPIRI mission. *Remote Sens. Environ.* 139, 291–305.
- McBratney, A., Whelan, B., Ancev, T., Bouma, J., 2005. Future directions of precision agriculture. *Precis. Agric.* 6, 7–23.
- McKinney, N., Schapaugh, W., Kanemasu, E., 1989. Canopy temperature, seed yield, and vapor pressure deficit relationship in soybean. *Crop Sci.* 29, 1038–1041.
- Mesas-Carrascosa, F.-J., Pérez-Porras, F., Meroño de Larriba, J., Mena Frau, C., Agüera-Vega, F., Carvajal-Ramírez, F., Martínez-Carricundo, P., García-Ferrer, A., 2018. Draft correction of lightweight microbolometer thermal sensors on-board unmanned aerial vehicles. *Remote Sens.* 10, 615.
- Mongus, D., Žalik, B., 2018. Segmentation schema for enhancing land cover identification: a case study using Sentinel 2 data. *Int. J. Appl. Earth Obs. Geoinf.* 66, 56–68.
- Moran, P.A., 1950. Notes on continuous stochastic phenomena. *Biometrika* 37, 17–23.
- Mourtzinis, S., Arriaga, F.J., Balkcom, K.S., Ortiz, B.V., 2013. Corn grain and Stover yield prediction at R1 growth stage. *Agron. J.* 105, 1045–1050.
- Murray, H., Lucieer, A., Williams, R., 2010. Texture-based classification of sub-Antarctic vegetation communities on Heard Island. *Int. J. Appl. Earth Obs. Geoinf.* 12, 138–149.
- Mutanga, O., Skidmore, A.K., 2004. Narrow band vegetation indices overcome the saturation problem in biomass estimation. *Int. J. Remote Sens.* 25, 3999–4014.
- Navar, S., Buddenbaum, H., Hill, J., Kozak, J., 2014. Modeling and mapping of soil salinity with reflectance spectroscopy and Landsat data using two quantitative methods (PLSR and MARS). *Remote Sens.* 6, 10813–10834.
- Neinavaz, E., Skidmore, A.K., Darvishzadeh, R., Groen, T.A., 2016. Retrieval of leaf area index in different plant species using thermal hyperspectral data. *ISPRS J. Photogramm. Remote Sens.* 119, 390–401.
- Ngiam, J., Khosla, A., Kim, M., Nam, J., Lee, H., Ng, A.Y., 2011. Multimodal deep learning. In: Proceedings of the 28th International Conference on Machine Learning (ICML-11), pp. 689–696.
- Nichol, J.E., Sarker, M.L.R., 2011. Improved biomass estimation using the texture parameters of two high-resolution optical sensors. *IEEE Trans. Geosci. Remote Sens.* 49, 930–948.
- Novak, D., Riener, R., 2015. A survey of sensor fusion methods in wearable robotics. *Robot. Auton. Syst.* 73, 155–170.
- Panda, S.S., Ames, D.P., Panigrahi, S., 2010. Application of vegetation indices for agricultural crop yield prediction using neural network techniques. *Remote Sens.* 2, 673–696.
- Pantazi, X.E., Moshou, D., Alexandridis, T., Whetton, R.L., Mouazen, A.M., 2016. Wheat yield prediction using machine learning and advanced sensing techniques. *Comput. Electron. Agric.* 121, 57–65.
- Panthee, D., Pantalone, V., Saxton, A., West, D., Sams, C., 2007. Quantitative trait loci for agronomic traits in soybean. *Plant Breed.* 126, 51–57.
- Pelizari, P.A., Sprohnle, K., Geiss, C., Schoepfer, E., Plank, S., Taubenbock, H., 2018. Multi-sensor feature fusion for very high spatial resolution built-up area extraction in temporary settlements. *Remote Sens. Environ.* 209, 793–807.
- Peralta, N.R., Assefa, Y., Du, J., Barden, C.J., Ciampitti, I.A., 2016. Mid-season high-resolution satellite imagery for forecasting site-specific corn yield. *Remote Sens.* 8.
- Poria, S., Cambria, E., Howard, N., Huang, G.B., Hussain, A., 2016. Fusing audio, visual and textual clues for sentiment analysis from multimodal content. *Neurocomputing* 174, 50–59.
- Prasertsak, A., Fukai, S., 1997. Nitrogen availability and water stress interaction on rice growth and yield. *Field Crop Res.* 52, 249–260.
- Ramachandram, D., Taylor, G.W., 2017. Deep multimodal learning: a survey on recent advances and trends. *IEEE Signal Process. Mag.* 34, 96–108.
- Raper, T.B., Varco, J.J., 2015. Canopy-scale wavelength and vegetative index sensitivities to cotton growth parameters and nitrogen status. *Precis. Agric.* 16, 62–76.
- Reynolds, M., Manes, Y., Izanloo, A., Langridge, P., 2009. Phenotyping approaches for physiological breeding and gene discovery in wheat. *Ann. Appl. Biol.* 155, 309–320.
- Rischbeck, P., Elsayed, S., Mistele, B., Barneier, G., Heil, K., Schmidhalter, U., 2016. Data fusion of spectral, thermal and canopy height parameters for improved yield prediction of drought stressed spring barley. *Eur. J. Agron.* 78, 44–59.
- Rochester, I.J., 2007. Nutrient uptake and export from an Australian cotton field. *Nutr. Cycl. Agroecosyst.* 77, 213–223.
- Rondeaux, G., Steven, M., Baret, F., 1996. Optimization of soil-adjusted vegetation indices. *Remote Sens. Environ.* 55, 95–107.
- Rouse Jr., J.W., Haas, R., Schell, J., Deering, D., 1974. Monitoring Vegetation Systems in the Great Plains with ERTS.
- Russello, H., 2018. Convolutional Neural Networks for Crop Yield Prediction Using Satellite Images.
- Sagan, V., Maimaitijiang, M., Sidike, P., Ebblimit, K., Peterson, K.T., Hartling, S., Esposito, F., Khanal, K., Newcomb, M., Pauli, D., 2019a. UAV-based high resolution thermal imaging for vegetation monitoring, and plant phenotyping using ICI 8640 P, FLIR Vue pro R 640, and thermoMap cameras. *Remote Sens.* 11, 330.
- Sagan, V., Maimaitijiang, M., Sidike, P., Maimaitiyiming, M., Erkbol, H., Hartling, S., Peterson, K., Peterson, J., Burken, J., Fritsch, F., 2019b. UAV/satellite multiscale data fusion for crop monitoring and early stress detection. *International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences* XLII-2/W13, 715–722.
- Schirrmann, M., Giebel, A., Gleiniger, F., Pflanz, M., Lentschke, J., Dammer, K.H., 2016. Monitoring agronomic parameters of winter wheat crops with low-cost UAV imagery. *Remote Sens.* 8.
- Schmidhuber, J., 2015. Deep learning in neural networks: an overview. *Neural Netw.* 61, 85–117.
- Schut, A.G.T., Traore, P.C.S., Blaes, X., de By, R.A., 2018. Assessing yield and fertilizer response in heterogeneous smallholder fields with UAVs and satellites. *Field Crop Res.* 221, 98–107.
- Scott, L.M., Janikas, M.V., 2010. Spatial statistics in ArcGIS. In: *Handbook of Applied Spatial Analysis*. Springer, pp. 27–41.
- Serrano, L., Filella, I., Penuelas, J., 2000. Remote sensing of biomass and yield of winter wheat under different nitrogen supplies. *Crop Sci.* 40, 723–731.
- Sharkey, T.D., 2005. Effects of moderate heat stress on photosynthesis: importance of thylakoid reactions, rubisco deactivation, reactive oxygen species, and thermo-tolerance provided by isoprene. *Plant Cell and Environment* 28, 269–277.
- Shi, Y.Y., Thomasson, J.A., Murray, S.C., Pugh, N.A., Rooney, W.L., Shahian, S., Rajan, N., Rouze, G., Morgan, C.L.S., Neely, H.L., Rana, A., Bagavathiannan, M.V., Henrickson, J., Bowden, E., Valasek, J., Olsenholter, J., Bishop, M.P., Sheridan, R., Putman, E.B., Popescu, S., Burks, T., Cope, D., Ibrahim, A., McCutchen, B.F., Baltensperger, D.D., Avant, R.V., Vidrine, M., Yang, C.H., 2016. Unmanned aerial vehicles for high-throughput phenotyping and agronomic research. *PLoS One* 11.
- Shiu, Y.-S., Chuang, Y.-C., 2019. Yield estimation of paddy rice based on satellite imagery: comparison of global and local regression models. *Remote Sens.* 11, 111.
- Sibanda, M., Mutanga, O., Rouget, M., Kumar, L., 2017. Estimating biomass of native grass grown under complex management treatments using worldview-3 spectral derivatives. *Remote Sens.* 9, 55.
- Sidike, P., Asari, V.K., Sagan, V., 2018a. Progressively Expanded Neural Network (PEN Net) for hyperspectral image classification: a new neural network paradigm for remote sensing image analysis. *ISPRS J. Photogramm. Remote Sens.* 146, 161–181.
- Sidike, P., Sagan, V., Quimsiyeh, M., Maimaitijiang, M., Essa, A., Asari, V., 2018b. Adaptive trigonometric transformation function with image contrast and color enhancement: application to unmanned aerial system imagery. *IEEE Geosci. Remote Sens. Lett.* 15, 404–408.
- Sidike, P., Sagan, V., Maimaitijiang, M., Maimaitiyiming, M., Shakoor, N., Burken, J., Mockler, T., Fritsch, F.B., 2019. dPEN: deep Progressively Expanded Network for mapping heterogeneous agricultural landscape using WorldView-3 satellite imagery. *Remote Sens. Environ.* 221, 756–772.
- Srivastava, N., Salakhutdinov, R.R., 2012. Multimodal learning with deep boltzmann machines. *Adv. Neural Inf. Proces. Syst.* 2222–2230.
- Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., Salakhutdinov, R., 2014. Dropout: a simple way to prevent neural networks from overfitting. *J. Mach. Learn. Res.* 15, 1929–1958.
- Stanton, C., Starek, M.J., Elliott, N., Brewer, M., Maeda, M.M., Chu, T.X., 2017. Unmanned aircraft system-derived crop height and normalized difference vegetation index metrics for sorghum yield and aphid stress assessment. *J. Appl. Remote. Sens.* 11.
- Thorp, K.R., DeJonge, K.C., Kaleita, A.L., Batchelor, W.D., Paz, J.O., 2008. Methodology for the use of DSSAT models for precision agriculture decision support. *Comput. Electron. Agric.* 64, 276–285.
- Torres-Rua, A., 2017. Vicarious calibration of sUAS microbolometer temperature imagery for estimation of radiometric land surface temperature. *Sensors* 17.
- Torres-Sánchez, J., Pena, J.M., de Castro, A.I., Lopez-Granados, F., 2014. Multi-temporal mapping of the vegetation fraction in early-season wheat fields using images from UAV. *Comput. Electron. Agric.* 103, 104–113.
- Tucker, C., Holben, B., Elgin Jr., J., & McMurtrey III, J. (1980). Relationship of spectral data to grain yield variation. *Photogramm. Eng. Remote. Sens.*, 46, 657–666.
- Tucker, C.J., 1979. Red and photographic infrared linear combinations for monitoring vegetation. *Remote Sens. Environ.* 8, 127–150.

- Uto, K., Seki, H., Saito, G., Kosugi, Y., 2013. Characterization of Rice paddies by a UAV-mounted miniature hyperspectral sensor system. *Ieee Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 6, 851–860.
- Valada, A., Mohan, R., Burgard, W., 2018. Self-Supervised Model Adaptation for Multimodal Semantic Segmentation. arXiv preprint [arXiv:1808.03833](https://arxiv.org/abs/1808.03833).
- Vergara-Díaz, O., Zaman-Allah, M.A., Masuka, B., Hornero, A., Zarco-Tejada, P., Prasanna, B.M., Cairns, J.E., Araus, J.L., 2016. A novel remote sensing approach for prediction of maize yield under different conditions of nitrogen fertilization. *Front. Plant Sci.* 7, 666.
- Virlet, N., Lebourgeois, V., Martinez, S., Costes, E., Labbe, S., Regnard, J.L., 2014. Stress indicators based on airborne thermal imagery for field phenotyping a heterogeneous tree population for response to water constraints. *J. Exp. Bot.* 65, 5429–5442.
- Wall, L., Larocque, D., Leger, P.M., 2008. The early explanatory power of NDVI in crop yield modelling. *Int. J. Remote Sens.* 29, 2211–2225.
- Wallace, L., 2013. Assessing the stability of canopy maps produced from UAV-LiDAR data. In: *Geoscience and Remote Sensing Symposium (IGARSS), 2013 IEEE International IEEE*, pp. 3879–3882.
- Wang, C., Nie, S., Xi, X.H., Luo, S.Z., Sun, X.F., 2017. Estimating the biomass of maize with hyperspectral and LiDAR data. *Remote Sens.* 9.
- Wang, J., Ding, J., Abulimiti, A., Cai, L., 2018. Quantitative estimation of soil salinity by means of different modeling methods and visible-near infrared (VIS-NIR) spectroscopy, Ebinur Lake Wetland, Northwest China. *PeerJ* 6, e4703.
- Wang, K., Franklin, S.E., Guo, X., Cattet, M., 2010. Remote sensing of ecology, biodiversity and conservation: a review from the perspective of remote sensing specialists. *Sensors* 10, 9647–9667.
- Wang, L., Tian, Y., Yao, X., Zhu, Y., Cao, W., 2014. Predicting grain yield and protein content in wheat by fusing multi-sensor and multi-temporal remote-sensing images. *Field Crop Res.* 164, 178–188.
- Wiegand, C.L., Richardson, A.J., Jackson, R.D., Pinter, P.J., Aase, J.K., Smika, D.E., Lautenschlager, L.F., McMurtrey, J., 1986. Development of agrometeorological crop model inputs from remotely sensed information. *IEEE Trans. Geosci. Remote Sens.* 90–98.
- Wilcox, J., Sediyyama, T., 1981. Interrelationships among height, lodging and yield in determinate and indeterminate soybeans. *Euphytica* 30, 323–326.
- Williams, J., Comanescu, R., Radu, O., Tian, L., 2018. DNN multimodal fusion techniques for predicting video sentiment. In: *Proceedings of Grand Challenge and Workshop on Human Multimodal Language (Challenge-HML)*, pp. 64–72.
- Wu, Q., Qi, B., Zhao, T., Yao, X., Zhu, Y., Gai, J., 2013. A tentative study on utilization of canopy hyperspectral reflectance to estimate canopy growth and seed yield in soybean. *Acta Agron. Sin.* 39, 309–318.
- Yang, C., Everitt, J.H., Du, Q., Luo, B., Chanussot, J., 2013. Using high-resolution airborne and satellite imagery to assess crop growth and yield variability for precision agriculture. *Proc. IEEE* 101, 582–592.
- Yin, X., McClure, M.A., Jaja, N., Tyler, D.D., Hayes, R.M., 2011. In-season prediction of corn yield using plant height under major production systems. *Agron. J.* 103, 923–929.
- You, J., Li, X., Low, M., Lobell, D., Ermon, S., 2017. Deep Gaussian Process for Crop Yield Prediction Based on Remote Sensing Data. *AAAI*, pp. 4559–4566.
- Yu, N., Li, L.J., Schmitz, N., Tiaz, L.F., Greenberg, J.A., 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.
- Zaman-Allah, M., Vergara, O., Araus, J., Tarekegne, A., Magorokosh, C., Zarco-Tejada, P., Hornero, A., Albà, A.H., Das, B., Craufurd, P., 2015. Unmanned aerial platform-based multi-spectral imaging for field phenotyping of maize. *Plant Methods* 11, 35.
- Zarco-Tejada, P.J., Gonzalez-Dugo, V., Berni, J.A.J., 2012. Fluorescence, temperature and narrow-band indices acquired from a UAV platform for water stress detection using a micro-hyperspectral imager and a thermal camera. *Remote Sens. Environ.* 117, 322–337.
- Zeng, W.Z., Zhang, D.Y., Fang, Y.H., Wu, J.W., Huang, J.S., 2018. Comparison of partial least square regression, support vector machine, and deep-learning techniques for estimating soil salinity from hyperspectral data. *J. Appl. Remote. Sens.* 12.
- Zhang, L.J., Gove, J.H., 2005. Spatial assessment of model errors from four regression techniques. *For. Sci.* 51, 334–346.
- Zhang, N., Rao, R.S.P., Salvato, F., Havelund, J.F., Möller, I.M., Thelen, J.J., Xu, D., 2018a. MU-LOC: a machine-learning method for predicting mitochondrially localized proteins in plants. *Front. Plant Sci.* 9.
- Zhang, X., Zhao, J., Yang, G., Liu, J., Cao, J., Li, C., Zhao, X., Gai, J., 2019. Establishment of plot-yield prediction models in soybean breeding programs using UAV-based hyperspectral remote sensing. *Remote Sens.* 11, 2752.
- Zhang, Y., Zhou, J., Meng, L., Li, M., Ding, L., Ma, J., 2018c. A method for deriving plant temperature from UAV TIR image. In: *2018 7th International Conference on Agro-Geoinformatics (Agro-Geoinformatics)*, pp. 1–5 (IEEE).
- Zhao, K.G., Suarez, J.C., Garcia, M., Hu, T.X., Wang, C., Londo, A., 2018. Utility of multitemporal lidar for forest and carbon monitoring: tree growth, biomass dynamics, and carbon flux. *Remote Sens. Environ.* 204, 883–897.
- Zheng, H., Cheng, T., Li, D., Yao, X., Tian, Y., Cao, W., Zhu, Y., 2018a. Combining unmanned aerial vehicle (UAV)-based multispectral imagery and ground-based hyperspectral data for plant nitrogen concentration estimation in rice. *Front. Plant Sci.* 9, 936.
- Zheng, H., Cheng, T., Zhou, M., Li, D., Yao, X., Tian, Y., Cao, W., Zhu, Y., 2018b. Improved estimation of rice aboveground biomass combining textural and spectral analysis of UAV imagery. *Precis. Agric.* 1–19.
- Zheng, H.B., Cheng, T., Li, D., Zhou, X., Yao, X., Tian, Y.C., Cao, W.X., Zhu, Y., 2018c. Evaluation of RGB, color-infrared and multispectral images acquired from unmanned aerial systems for the estimation of nitrogen accumulation in rice. *Remote Sens.* 10.
- Zhou, X., Zheng, H., Xu, X., He, J., Ge, X., Yao, X., Cheng, T., Zhu, Y., Cao, W., Tian, Y., 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.