Received: 21 March 2024 | Accepted: 10 October 2024

DOI: 10.1002/ppj2.70009

The Plant Phenome Journal OPEN ACCESS

ORIGINAL ARTICLE

Multi-sensor and multi-temporal high-throughput phenotyping for monitoring and early detection of water-limiting stress in soybean

Sarah E. Jones[1] | **Timilehin T. Ayanlade**[2] | **Benjamin Fallen**[3] | **Talukder Z. Jubery**[2] | **Arti Singh**[1] | **Baskar Ganapathysubramanian**[2] | **Soumik Sarkar**[2] | **Asheesh K. Singh**[1]

[1]Department of Agronomy, Iowa State University, Ames, Iowa, USA

[2]Department of Mechanical Engineering, Iowa State University, Ames, Iowa, USA

[3]USDA-ARS Soybean and Nitrogen Fixation Research Unit, Raleigh, North Carolina, USA

Correspondence
Asheesh K. Singh, Department of Agronomy, Iowa State University, Ames, IA, United States.
Email: singhak@iastate.edu
Soumik Sarkar, Department of Mechanical Engineering, Iowa State University, Ames, IA, United States.
Email: soumiks@iastate.edu

Funding information
National Institute of Food and Agriculture, Grant/Award Numbers: 2019-67021-29938, 2021-67021-35329; Iowa Soybean Association; Iowa State University Plant Sciences Institute; National Science Foundation, Grant/Award Number: S&CC 1952045; United Soybean Board; COALESCE: CPS Frontier, Grant/Award Number: 1954556; U.S. Department of Agriculture, Grant/Award Number: CRIS Project IOW04714; Raymond F. Baker Center for Plant Breeding

Abstract

Soybean (*Glycine max* [L.] Merr.) production is susceptible to biotic and abiotic stresses, exacerbated by extreme weather events. Water limiting stress, that is, drought, emerges as a significant risk for soybean production, underscoring the need for advancements in stress monitoring for crop breeding and production. This project combined multi-modal information to identify the most effective and efficient automated methods to study drought response. We investigated a set of diverse soybean accessions using multiple sensors in a time series high-throughput phenotyping manner to: (1) develop a pipeline for rapid classification of soybean drought stress symptoms, and (2) investigate methods for early detection of drought stress. We utilized high-throughput time-series phenotyping using unmanned aerial vehicles and sensors in conjunction with machine learning analytics, which offered a swift and efficient means of phenotyping. The visible bands were most effective in classifying the severity of canopy wilting stress after symptom emergence. Non-visual bands in the near-infrared region and short-wave infrared region contribute to the differentiation of susceptible and tolerant soybean accessions prior to visual symptom development. We report pre-visual detection of soybean wilting using a combination of different vegetation indices and spectral bands, especially in the red-edge. These results can contribute to early stress detection methodologies and rapid classification of drought responses for breeding and production applications.

1 | INTRODUCTION

Biotic and abiotic stresses, exacerbated by weather events, can lead to billions of dollars in U.S. crop insurance payments, economic loss for farmers, and increased consumer prices (Dice & Rodziewicz, 2020). The impact of drought on major grain crops can be severe; therefore, its patterns are traced globally in relation to yield loss (Santini et al., 2022). Across different species, yield loss attributed to drought has been investigated. These losses in soybean range from 28% to 74% depending on various factors such as the temperature and timing of drought (Jumrani & Bhatia, 2018). Furthermore, a recent study of the magnitude, frequency, duration, and timing of droughts has shown that North America is at high risk for reduced soybean (*Glycine max* [L.] Merr.) yield associated with drought (Santini et al., 2022).

In the 1980s, landrace PI 416937, from Japan, was observed to wilt more slowly in the field compared to modern cultivars and had a lower yield penalty under water deficit stress (Sloane et al., 1990). Further studies reported that increased root mass, volume, and density, larger leaf area, higher nodule count, maintained turgor pressure, aluminum tolerance, water conservation strategies, and limited transpiration rate, among other traits, could confer the slow wilting phenotype (Carter & Rufty, 1992; Fletcher et al., 2007; Hudak & Patterson, 1995; Pantalone et al., 1996; Valliyodan et al., 2017; Ye et al., 2020). Slow wilting lines have also shown lower yield reductions under drought stressed conditions (Pathan et al., 2014; Sloane et al., 1990; Ye et al., 2020). These lines have contributed to breeding progress, as their progenies showed higher-yield under drought conditions (Carter & Rufty, 1992; Zhou et al., 2020). Canopy wilting has become a measure of drought tolerance in soybean breeding efforts because of the association between the slow wilting phenotype and higher seed yield (Pathan et al., 2014; Ye et al., 2020). Furthermore, it has been found that soybean growth stage during drought stress affects the response to drought stress, making growth stage an important factor to consider in drought studies (Kpoghomou et al., 1990). Ideally, breeders could simultaneously evaluate cultivars and breeding lines for seed yield and growth stage under drought stressed and well-watered environments to fully understand drought response in tested lines. However, in situations where drought-prone land is limited, or in early generations when seed quantity is too limited to plant multi-environment trials, the use of proxy traits, such as canopy wilting, can be utilized for screening and selection of lines with potential drought tolerance until full-sized yield trials are possible later in the breeding pipeline. For example, Zhou et al. (2021) used canopy wilting, maturity group, and various traits extracted from RGB and multispectral imagery data to predict soybean yield of breeding lines under drought conditions.

Core Ideas

- Sensors, wavebands, and vegetation indices are evaluated for importance in phenotyping canopy wilting in soybean.
- Aerial versus ground based sensing study shows necessity of balancing speed of data collection and field of view.
- Random forest classification can be applied to support selection decisions in plant breeding program.
- Multispectral UAV data enables pre-visual early detection of canopy wilting drought stress in soybean.

Several challenges emerge in phenotyping for canopy wilting in breeding and crop production scenarios. Traditional methods require an expert to visually rate canopy wilting severity. Several methods and scales exist for this visual classification in soybean, including the commonly cited scale from 0 to 100 (0 = no wilting, 20 = slight wilting and leaf rolling at the top of the canopy, 40 = severe leaf rolling at the top of the canopy and moderate leaf wilting throughout the canopy and loss of petiole turgidity, 60 = severe wilting throughout the canopy and loss of petiole turgidity, 80 = severe petiole wilting and dead leaves scattered throughout the canopy, and 100 = plant death (Chamarthi et al., 2021; King et al., 2009; Kaler et al., 2017), and the equivalent 0 (no wilting) to 5 (plant death) scale (Charlson et al., 2009). Additional methods include a 1–5 scale (1 = no wilting, 2 = few top leaves showed wilting, 3 = half of the leaves showed wilting, 4 = severe wilting 75% of the leaves showed wilting, and 5 = severely wilted) followed by dividing scores into two categories of slow wilting = average wilting score $\leq$2.5 or fast wilting = average wilting score $\geq$2.5 used in cases rated on the 1–5 scale (Ye et al., 2020; Zhou et al., 2020). While integral to breeding programs and crop production, several limitations emerge with visual ratings. Visual ratings can be prone to inter- and intra-rater variation, provide no early warning for farmers equipped with irrigation systems, and are time-consuming to collect. Furthermore, breeding programs evaluate thousands of test lines at multiple locations each year (Vieira & Chen, 2021). Canopy wilting is affected by temporally variable environmental conditions such as increasing temperature and solar radiation, leading to increasingly severe wilting symptoms as the day progresses. Therefore, it is essential that rapid, automated methods for drought screening be implemented in breeding programs to facilitate selection with increased speed and accuracy (Singh et al., 2021a; Singh et al., 2021). The immense scale of breeding programs necessitates rapid phenotyping to decrease time and labor costs, and appropriate statistical analysis that can handle complex data.

Sensors on unmanned aerial vehicles (UAVs) offer improved speed and spectral resolution beyond human vision with high applicability in breeding programs (Guo et al., 2021; Herr et al., 2023). A comprehensive array of sensors, including RGB, multispectral, hyperspectral, and thermal, are readily accessible for integration with UAV flights tailored to many agricultural applications (Herr et al., 2023; Singh et al., 2021b). The combined speed of collection, spectral density, and often multicollinearity of high throughput phenotyping data can benefit from machine learning (ML) analysis strategies to detect feature hierarchy and recognize patterns in multi-modal data to provide data-driven solutions (Singh et al., 2016, 2018, 2021; Chan et al., 2022). Common ML objectives in plant stress phenotyping include identification, classification, quantification, and prediction (ICQP) of plant stress (Singh et al., 2016). Identification detects the causal stress, classification groups severity expressions into a finite number of classes, quantification determines the severity of stress expression at a finer level such as percent expression, and prediction combines multiple data sources such as environmental parameters and imagery to predict stress onset (Singh et al., 2016). The ICQP paradigm, via ML/ deep learning (DL) approaches, has been applied to solve crop production and breeding issues, including disease identification, severity classification (Ghosal et al., 2018; Rairdin et al., 2022), insect identification (Chiranjeevi et al., 2023), and has even be applied to below ground traits (Carley et al., 2023).

Limited research is available studying UAV-based methods for drought-induced canopy wilting assessment in the field, especially in earlier maturity group soybean. One study used multiple sensors, including RGB, infrared thermal, and multispectral cameras mounted on UAV to classify 116 soybean genotypes into two classes (slow vs. fast wilting soybeans) utilizing four imagery-based features inputted into a support vector machine (SVM) algorithm with an accuracy of 0.80 (Zhou et al., 2020). Similar to SVMs, the random forest (RF) model, an ensemble learning algorithm, also has been widely used (Singh et al., 2016) and has been successful in plant stress identification and classification (de Oliveira et al., 2023). Peanut (*Arachis hypogaea*) canopy wilting was classified into six wilt classes using 11 features extracted from aerial-based RGB imagery and logistic regression achieving an accuracy of 0.69 (Sarkar et al., 2021). The same data was used to classify wilting into binary categories of turgid vs. wilted canopy with an increased accuracy of 0.88. A few studies in soybean have approached drought phenotyping by predicting yield under drought stress. Nine features from multispectral and RGB-based UAV imagery of 116 soybean genotypes were input into a convolutional neural network to estimate yield under drought stress with an R^2 of 0.78 (Zhou et al., 2021). To test application in a breeding program, a follow-up study showed that 38 features from UAV-based multispectral imagery could be used to classify over 11,000 progeny rows and over 1,000 preliminary yield trial entries into select vs. non-select classes (Zhou et al., 2022). The model also selected over 60% of the same selections as a breeder, while increasing average yield of the selected class compared to breeder selections (Zhou et al., 2022). This agreement and improved yield show the promise of UAV data and model-based selection methods in the breeding pipeline. Both canopy wilting and yield prediction methods are valuable and offer unique information to plant breeders.

Likewise, early stress detection plays a pivotal role in crop production, giving farmers more time to act to alleviate stress. Early stress detection has shown to be promising in insect, herbicide, and disease stress studies utilizing wavelengths beyond human vision for pre-visual detection of stress. For example, a study of pine tree bark beetles showed that RGB data alone could not detect early infestations, however, red-edge bands played a significant role in early infection detection in a time series data set (Yu et al., 2022). In soybean, artificial neural networks have also been used in the early detection of herbicide injury, classifying 240 soybean genotypes into three classes of tolerant, moderate, and susceptible Dicamba injury responses (Vieira et al., 2022). Pre-visual detection and quantification of the fungal disease, charcoal rot (*Macrophomina phaseolina*), was possible with an accuracy of 0.97 when using six selected wavebands, including 475.56, 548.91, 652.14, 516.31, 720.05, and 915.64 nm (Nagasubramanian et al., 2018). In a UAV-based study, the early detection of charcoal rot was successful using Normalized Difference Red Edge (NDRE) calculated from multispectral data (Brodbeck et al., 2017). Early drought stress detection has not been widely studied in soybean. However in grapevine, one study found that vegetation indices in the red and near-infrared (NIR) range, such as Normalized Difference Vegetation Index (NDVI), Atmospherically Resistant Vegetation Index (ARVI), Enhanced Vegetation Index (EVI), Soil-Adjusted Vegetation Index (SAVI), and Optimized Soil-Adjusted Vegetation Index (OSAVI), had a higher correlation with leaf water potential compared to other indices, while chlorophyll based indices, especially NDRE, were also highly correlated with leaf water potential (Tang et al., 2022). The objectives of this study were to: (1) develop a pipeline for rapid classification of soybean drought symptoms in early maturity group soybeans and (2) investigate methods for early detection of drought stress.

2 | MATERIALS AND METHODS

2.1 | Field design

Field experiments were conducted in a rain-fed drought nursery at the Muscatine Island Research Farm in Fruitland, IA (41°21′ N, 91°08′ W) on Fruitfield coarse sand in 2022. Two-row, 1.52 m plots with 76 cm row spacing and 0.91 m alleys were planted with a seeding density of 33 seeds

per meter on May 28, 2022. The experimental design was a randomized complete block design with three replications of 450 MG 0-III lines. Test lines included a 31-member subset from the SoyNAM panel (Song et al., 2017; Diers et al., 2018), eight maturity and yield checks, and 411 PI lines from a mini core collection subset from the USDA Soybean Germplasm Core Collection representing the diversity of soybean that can be grown in Iowa. PI lines originate from 28 countries across North America, Europe, Asia, and Africa.

A subset of 12 lines were also planted under pivot irrigation about 470 m from the rain-fed nursery. Irrigated plots were planted in a randomized complete block design with three replications at the same plot size and planting density as in the rain-fed nursery. These 12 lines included seven lines that had low wilt score best linear unbiased predictors (BLUPs) from previous drought trial data from 2020 and 2021, three lines with high wilt score BLUPs, and two high yielding check lines.

2.2 | Data collection

Daily precipitation (mm), was obtained from Iowa State University's Iowa Environmental Mesonet NWS COOP station IA5837 (https://mesonet.agron.iastate.edu/request/coop/fe.phtml), and daily summaries of soil volumetric water content was obtained from the Iowa Environmental Mesonet Fruitland, IA station FRUI4 at (https://mesonet.agron.iastate.edu/agclimate/hist/daily.php). Data was collected across multiple single-day time points 46 days after planting (DAP), 64 DAP, and 81 DAP in 2022 to investigate stress progression. Entire replications were phenotyped in the shortest possible time and the on same day. Visual canopy wilting was recorded on the equivalent scale of 1–6, quantifying the level of wilting severity seen in the plant canopy (1 = no wilting, 2 = slight wilting and rolling in the top of the canopy, 3 = somewhat severe leaf rolling in the top canopy, moderate wilting of leaves throughout the canopy, some loss of petiole turgidity, 4 = severe wilting of leaves throughout the canopy with advanced loss of petiole turgidity 5 = petioles severely wilted and dead leaves throughout much of the canopy, 6 = plant death (Charlson et al., 2009; King et al., 2009). One individual collected visual wilt scores from both irrigated and rain-fed plots between 12:00 and 2:00 p.m. to prevent inter-rater variability within a replication. The average soybean growth stage per plot was recorded (Fehr & Caviness, 1977). All sensor and growth stages data were collected between 10:00 a.m. and 2:00 p.m as shown in Table 1.

2.3 | Data processing

Data processing included stitching and orthomosaic generation from UAV-based images, cropping plot boundaries, background removal, and feature generation from canopy pixels. UAV-based RGB, multispectral, and thermal flight data were stitched in Pix4D using GPS coordinates of ground control points in the field. From the resulting orthomosaics, plot boundaries were demarcated and snipped via Python code interfacing with ArcGIS (Carroll et al., 2024). Vegetation pixels were separated from the soil background, and mean reflectance was extracted from each sensor via sensor-dependent pipelines. For RGB images, we separated the vegetation in the foreground from the soil in the background using the Hue, Saturation, and Value (HSV) color space. RGB images were transformed to the HSV color space, and the pixels with HSV colors between the range of HSV values from (25, 20, 50) to (80, 255, 255) were kept while the other pixels were masked. This range, obtained through the inspection of the peak in the distribution of the HSV values on our data, thresholded the contours that contain the pixels defined as vegetation. A similar pipeline was employed for the RGB bands of the multispectral imagery data and the masks created were used to crop out the vegetation from each of the ten multispectral bands. The mean reflectance of all soybean plots for multispectral imagery and RGB imagery was extracted for further analysis. Figure 1 is a flowchart showing the pipeline developed for data processing and analysis. To characterize various vegetation characteristics, vegetation indices (VIs) were calculated from the multispectral and hyperspectral (ASD) sensors using equations provided in Tables 2 and S1. These indices are widely used to quantify plant health, greenness, and other physiological conditions. These VIs also include water-based indices and indices based on only the visible (red, green, and blue) bands.

2.4 | Data analysis

Data from each sensor was analyzed in two phases. Phase one targeted monitoring drought stress severity, and phase two targeted early detection of drought stress prior to visual symptom development. The traditional wilt score rating on a scale of 1–6 was re-grouped into classes that could be applied in a breeding program for advancement decisions. This led to a merger of classes 1, 2, and 3 into 'Select' and classes 4, 5, and 6 as 'Discard', as shown in Figure 2.

2.4.1 | Monitoring of drought stress

For monitoring, we focused on time point 2 (65 DAP) due to the higher severity of stress observed at this time point. Complete sensor data were available for 2 replications, or 900 of the 1350 plots. In this data set, the 'Select' category consisted of 431 plots and the 'Discard' class contained 469 plots. To explore the performance of wavebands falling within the visual range of 400–700 nm to mimic human vision as used by raters, the red, green, and blue bands were selected from

TABLE 1 List of sensors, sensor specifications, and overview of methodology and data points captured in this study.

Sensor	Sensor specification	Wavelength range (nm)	Spectral resolution	Altitude	Overlap	Genotypes	Replications	Time points
Canon T5i digital SLR camera[1]	EF-S 18-55 mm f/3.5-5.6 IS II	400 - 700	R,G,B	1.3 m	N/A			
ASD FieldSpec 4 Hi-Res[2]	hyperspectral reflectance	350 - 2500	1 nm	1 m above canopy	N/A			
DJI Phantom 4 Advanced UAV[3]	20 MP, 1-inch CMOS with a 24 mm focal length	400 - 700	R,G,B	30 m	85% front, 75% side			
Zenmuse X5S + DJI Inspire UAV[3]	45 mm	400 - 700	R,G,B	45 m	70% front, 80% side	450	3	3
Micasense RedEdge-Mx Dual camera system[4] + DJI Matrice 600 Pro UAV[3]	10 band multispectral	444 - 842	444(28), 475(32), 531(14), 560(27), 650(16), 668(14), 705(10), 717(12), 740(18), 842(57)	30 m	80% front, 80% side			

[1]Canon USA, Inc., Melville, NY.
[2]Malvern Panalytical, Malvern, United Kingdom.
[3]SZ DJI Technology Co., Ltd., Shenzhen, China.
[4]AgEagle Aerial Systems, Inc., Wichita, KS.
[5]Teledyne FLIR LLC, Wilsonville, OR.

the multispectral and ASD sensors and were stacked to create two additional visual range-based sensors.

In this study, random forest (Breiman, 2001) models were used to classify the drought susceptibility traits for the various sensors. After tuning, we set the number-of-estimator parameter of the RF models to 100 to learn the non-linear classifier. The number-of-estimator parameter was selected to enhance generalization by capturing diverse patterns across each sensor while accounting for computational constraints. The model performance was evaluated using a 10-fold cross-validation method, where the data is randomly split into 10-folds, and for each training iteration, the RF model was trained on 9-folds, while the remaining fold was used for model testing, repeated 10 times. Classification accuracy was calculated using Equation (1). We used cross-validation as a conservative estimate of the model accuracy (Joalland et al., 2018) as it is also valuable for models dealing with limited sample sizes (James et al., 2023).

$$\text{Accuracy} = \frac{\text{Number of samples classified correctly}}{\text{Total number of samples}} \quad (1)$$

To identify the most effective sensor for monitoring drought stress, we trained a RF model using various sets of sensor data across multiple levels of representation. First, we trained the model on a classification task aimed at distinguishing between 'Select' and 'Discard' classes using only the visual range of wavelengths used by human raters. Subsequently, we extended the training to include both visible and non-visible spectral information. And finally, we incorporated vegetation indices.

Additionally, we evaluated the benefits of integrating data from multiple sensors by analyzing the performance of combined sensor setups. Given the complexity introduced by the high dimensionality of the combined sensor data, we formulated a sensor selection problem by implementing a backward elimination technique at the sensor level. This allowed us to identify the most advantageous combinations of sensors for effective drought stress monitoring. Furthermore, we refined the selection process to the feature level to determine the top 10 features across all sensors that are critical for accurate monitoring of drought stress.

2.4.2 | Early detection

To address the early detection of drought stress, we utilized data from two of the three time points, framing it as a time series analysis problem. Since all plots at time point 1 were initially rated as a wilt score of 1, we reclassified these plots

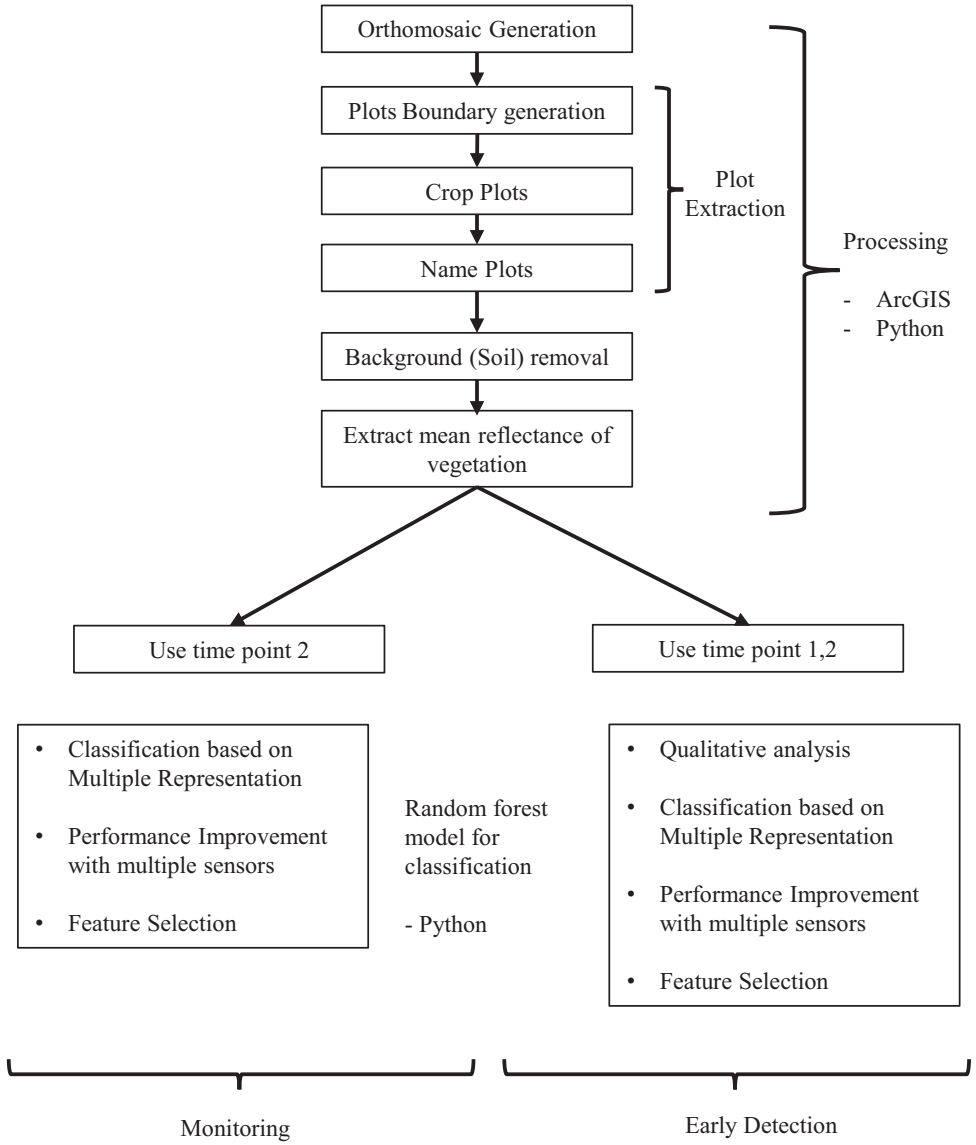

FIGURE 1 Flowchart of the methods pipeline and software utilized for data processing and analysis to address monitoring and early detection goals.

using the 'Select' and 'Discard' class labels from time point 2 which had the widest range of canopy wilting scores. This reclassification strategy aligns the early stage time point 1 sensor data with their subsequent time point 2 wilt score classifications, similar to (Yu et al., 2022). Complete sensor data were available for 2 replications at time point 1, or 900 of the 1350 plots. The class counts for the 'Select' and 'Discard' classes were 406 and 494, respectively.

As with the drought stress monitoring task, a RF model was employed to assess the class discriminative capability at various sensor representation levels using the newly reclassified dataset. Additionally, we extended our analysis to include multiple sensors, evaluating their collective performance in the context of early detection. This approach also involved identifying the most important sensor combinations and features, drawing parallels to the techniques used in the monitoring task for optimizing drought stress detection.

To further investigate performance of the multispectral data, a Welch's t-test was employed to examine the spectral variations in multispectral imagery bands between 'Select' and 'Discard' plots at each of the three time points. Multispectral imagery was available for all three replications at each time point resulting in 1350 plots of data analyzed at each time point. The analysis sought to to identify the time during the growing season when the classes exhibited significant differences in mean reflectance as well as the specific bands showing significant differences between categories that could be utilized for early detection of stress.

3 | RESULTS

From the week of May 24, 2022 to the week of July 5, 2022, the majority of Muscatine County, Iowa was under

TABLE 2 List of vegetation indices (VIs) utilized for drought monitoring and early detection.

Index	Name	Formula	Reference
NDVI	Normalized Difference Vegetation Index	$\frac{p780 - p670}{p780 + p670}$	(Rouse et al., 1973)
PRI	Photochemical Reflectance Index	$\frac{p531 - p570}{p531 + p570}$	(Gamon et al., 1997)
RARSa	Ratio Analysis of Reflectance Spectra A	$\frac{p675}{p700}$	(Chappelle et al., 1992)
RARSb	Ratio Analysis of Reflectance Spectra B	$\frac{p675}{p650 \cdot p700}$	(Chappelle et al., 1992)
RARSc	Ratio Analysis of Reflectance Spectra C	$\frac{p760}{p500}$	(Chappelle et al., 1992)
RDVI	Renormalized Difference Vegetation Index	$\frac{p800 - p670}{\sqrt{p800 + p670}}$	(Roujean & Breon, 1995)
EVI	Enhanced Vegetation Index	$2.5 \cdot \frac{(p800) - (p670)}{(p800) + 6 \cdot (p670) - 7.5 \cdot (pBLUE) + 1}$	(Huete et al., 2002)
GCI	Green Chlorophyll Index	$\frac{(p800)}{(p570)} - 1$	(Esri, 2024)
MSAVI	Modified Soil-Adjusted Vegetation Index	$\frac{2 \cdot (p800) + 1 - \sqrt{(2 \cdot (p800) + 1)^2 - 8 \cdot ((p800) - (p670))}}{2}$	(Qi et al., 1994)
NDRE	Normalized Difference Red Edge	$\frac{(p790 - p720)}{(p790 + p720)}$	(Sims & Gamon, 2002)
RECI	Red-Edge Chlorophyll Index	$\frac{(p800)}{(p740)} - 1$	(Esri, 2024)
REV	Red-Edge Vegetation Index	$\frac{(pp740)}{\sqrt{p670}}$	(Xie et al., 2018)
ARI	Anthocyanin Reflectance Index	$\frac{1}{p570} - \frac{1}{p740}$	(Gitelson et al., 2009)
NDLI	Normalized Lignin Index	$\log\left(\frac{1}{p1754}\right) - \log\left(\frac{1}{p1680}\right)$	(Serrano et al., 2002)
NMDI	Normalized Multi-band Drought Index	$\frac{p860 - (p1640 - p2130)}{p860 + (p1640 - p2130)}$	(Wang & Qu, 2007)
NWI	Normalized Water Index I*	$\frac{p970 - p900}{p970 + p900}$	(Prasad et al., 2007)
PSRI	Plant Senescence Reflectance Index	$\frac{p680 - P500}{p750}$	(Merzlyak et al., 1999)
VREI2	Vogelmanns Red Edge Index 2	$\frac{p734 - p747}{p715 + p726}$	(Vogelmann et al., 1993)
NDWI	Normalized Difference Water Index	$\frac{p860 - p1240}{p860 + p1240}$	(Gao, 1996)
WI	Water Index	$\frac{p900}{p970}$	(Peñuelas et al., 1993)

(Continues)

TABLE 2 (Continued)

Index	Name	Formula	Reference
DRI	Drought Response Index	$\frac{p850 - p2218}{p850 + p1928}$	(Datt, 1999)
SWIWSI	Short wave infrared water stress index	$\frac{p860 - p1640}{p860 + p1640}$	(Fensholt & Sandholt, 2003)
CWM*	Canopy Water Mass Index*	*	(Winterhalter et al., 2011)
Y*	Reflectance Ratios Y*	*	(Yu et al., 2000)

*These VIs have accompanying numbered variations which can be found in Table S1.

Visual	Visual Description	Wilt Score Rating Traditional	Multi-Class	Binary
	No wilting	1	Tolerant	Select
	Light wilting and rolling in the top of the canopy	2		
	Somewhat severe leaf rolling in top of canopy, moderate wilting of leaves throughout the canopy, some loss of petiole turgidity	3	Moderately Susceptible	
	Severe wilting of leaves throughout the canopy, with advanced loss of petiole turgidity	4	Susceptible	Discard
	Petioles severely wilted and dead leaves throughout much of the canopy	5		
	Plant death	6		

FIGURE 2 Class categorization of drought stress symptoms in soybean utilized in the analysis. Visual ground-truth data was collected on 450 diverse accessions in water-limited screening nurseries via a traditional wilt score rating scale of 1–6 adapted from previous studies utilizing the same visual description of classes of the 0–5 and 0–100 scales (Charlson et al., 2009; King et al., 2009). Visual plot scores of 1–6 were re-classified into a binary setting of 'Select' and 'Discard' that could be applied in a breeding program.

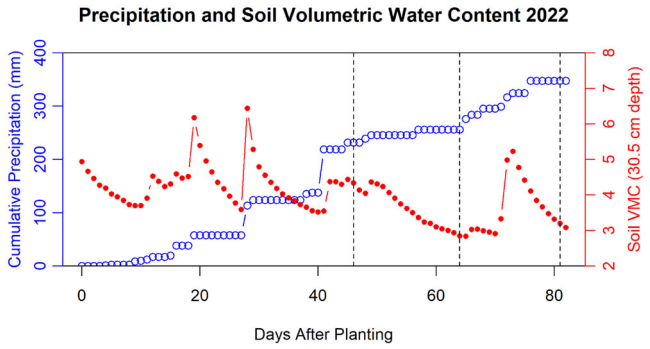

FIGURE 3 Cumulative precipitation (mm) as shown by blue open circle points and soil volumetric water content (%) from Iowa Environmental Mesonet as shown by solid red points beginning on the day of planting (May 28, 2022) and continuing to one day post final data collection. Time points 1, 2, and 3 are marked by vertical dashed lines at 46, 64, and 81 days after planting.

D0 abnormally dry conditions according to the United States Drought Monitor. In the week of July 12, 2022, the abnormally dry conditions moved southwest and the abnormally dry conditions abated in most of the county. However, as seen in Figure 3, soil volumetric water content (%) at 30.5 cm decreases rapidly after each rain event due to the rapidly draining Fruitfield Coarse Sand soil type. Volumetric water content was 4.3% at time point 1, 2.8% at time point 2, and 3.2% at time point 3. Lower soil volumetric water content, seen especially at time point 2, corresponds with more severe wilting scores in the soybean plots. In the rain-fed field, the average wilt score rating was 1 at time point 1, 3.5 at time point 2 and 2.8 at time point 3. Taken 7 days after rain, and after a period of only 10.2 mm of precipitation in the previous 15 days, the second time point showed the most severe wilting symptoms of the season and had a volumetric water content below the permanent wilting point for primarily sand textured soil (Ratliff et al., 1983). In the irrigated field, no wilting symptoms were visible at time points 1, 2, or 3. Due to the close proximity of the two fields, and the lack of disease or stress in the irrigated field, we concluded that wilting symptoms in the rain-fed field were due to moisture deficit.

Growth stage ranged from V3 - R2 at time point 1, from V9 - R5 at time point 2, and from R2 - R8 at time point 3. The wilt scores of plots at R8 were not collected due to final maturity and were, therefore, not included in the analysis.

3.1 | Monitoring of soybean wilt

3.1.1 | Classification based on multiple representation

Table 3 presents the performance outcomes of the RF models used for classifying canopy wilting severity across multiple sensor data representations. The evaluation, based on a 10-fold cross-validation, reveals distinct accuracies for different data forms. Considering only the RGB representations, which utilize the spectral wavelengths used by expert raters, the multispectral sensor, with a ground sampling distance of 2.0 cm/pixel, exhibited superior performance in detecting canopy wilting severity compared to other sensors. Additionally, in the spectral representation, where non-visible wavelengths augment the RGB data, the multispectral sensor surpassed the handheld ASD sensor with an accuracy of 71.8%. This trend of higher performance with the multispectral sensor persisted even in the VIs representation, though accuracy of multispectral VIs was slightly lower than accuracy of multispectral bands. Most important features selected via RF for monitoring within sensor representations can be found in Tables S2 and S3.

3.1.2 | Performance improvement with multiple sensors

The combinations of multiple sensor data were evaluated using a RF backward elimination algorithm with cross-validation at sensor level. The model using all nine sensor representations (handheld RGB, phantom RGB, inspire RGB, multispectral RGB, ASD RGB, multispectral, ASD, and vegetation indices from multispectral and ASD indices) as predictors (the full model) had an overall classification accuracy of 72.8%. The highest accuracy model via backward elimination used a subset of sensor representations including handheld RGB, phantom RGB, inspire RGB, multispectral RGB, multispectral, and vegetation indices from Multispectral and ASD indices with the ASD RGB, and ASD sensor representations dropped, reached an accuracy of 76.2%. The confusion matrix for this model is presented in Figure 4. Details on the step-wise performance and sensors removed can be found in Table S4.

With the complexity and high dimensionality of combined sensor data and its impact on performance, it is important to have an efficient analysis pipeline in place for wilt classification. Hence, we determined a set of the most effective features across the combined data to reduce band correlation, preserve spectral information, and lower the computational costs of working with multiple sensors. Using the RF backward elimination algorithm with cross-validation at a feature level, the performance of the selected combinations of bands was evaluated based on their classification accuracy. A classification using 10 selected features-the multispectral 'blue-444', 'green', 'green-531', 'red edge', and 'RARSa' features, the phantom 'red', 'green', and 'blue' features, and the inspire 'green', and 'blue' features, obtained a classification accuracy of 75%, with a 3.7% increment compared to using all sensor features. This selection predominantly utilizes visible

TABLE 3 Monitoring of soybean wilt—Performance of the random forest model for classifying two classes of canopy wilting severity ('Select' and 'Discard') in response to moisture deficit stress through different sensor representations. Results show mean and standard deviation of performance across a 10-fold classification, focused on monitoring stress when canopy wilting was visible in the field.

	Sensor				
	Handheld RGB	**Phantom**	**Inspire**	**Multispectral**	**Handheld ASD**
RGB[a]	0.591 ± 0.045	0.659 ± 0.040	0.621 ± 0.041	0.687 ± 0.043	0.520 ± 0.044
Spectral[b]				0.718 ± 0.042	0.601 ± 0.042
VIs[c]				0.710 ± 0.030	0.634 ± 0.049

[a]RGB: RGB bands were also extracted from the multispectral and handheld ASD sensors.
[b]Spectral: 10 multispectral UAV bands, 2151 proximal bands.
[c]VIs: 13 multispectral based VIs, 35 handheld ASD based VIs.

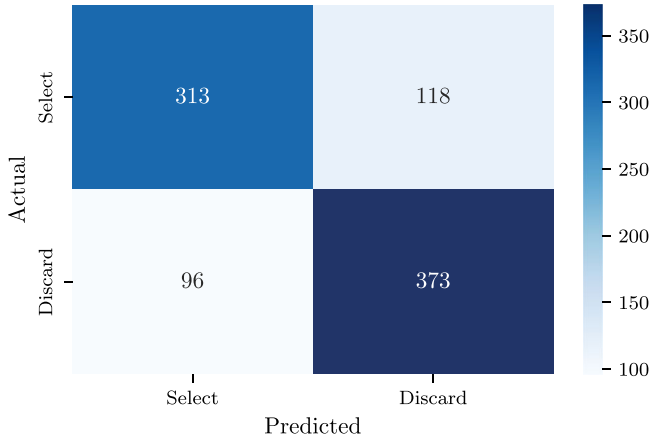

FIGURE 4 Confusion matrix for the highest accuracy model obtained via backward elimination. The model utilized a subset of sensor representations, including handheld RGB, phantom RGB, inspire RGB, multispectral RGB, multispectral, and vegetation indices from Multispectral and ASD indices, with the ASD RGB and ASD sensor representations excluded. The model achieved an accuracy of 76.2%, with class counts of 431 for 'Select' and 469 for 'Discard'.

wavelengths, with an inclusion of near-infrared wavelengths for calculating the RARSa vegetation index, showcasing a balanced approach between spectral types. To see expanded list of top 20 selected features see Table S5.

3.2 | Early detection of soybean wilt

3.2.1 | Qualitative analysis

In the 2022 growing season, multi-temporal data was collected. Field investigations within the first time point (46 DAP) observed no evident wilting symptoms in susceptible soybean plots. At the second time point (64 DAP), the soybean leaves and petioles began to showcase wilting symptoms in the susceptible plots, which were also evident in the third time point at 81 DAP. Therefore, the UAV-based data we explore contains pre- and post-visual canopy wilting development.

3.2.2 | Classification based on multiple representation

For the early detection of soybean wilting, RF classification models were utilized to differentiate between 'Select' and 'Discard' soybean plots before they were differentiable to the human eye. As shown in Table 4, the models were evaluated based on their mean reflectance performance across various sensor representations, using the same 10-fold cross-validation method employed in the monitoring studies. In the RGB spectrum, which mimics the wavelengths used by expert raters, the Inspire sensor, with a ground sampling distance of 0.33 cm/pixel, demonstrated superior performance in classifying the severity of canopy wilting compared to other sensors. In the spectral representation, which included non-visible wavelengths in addition to RGB, both the multispectral sensor and the handheld ASD sensor achieved a classification accuracy of approximately 0.58. Further improvement was observed when spectral data was converted into VIs, with the multispectral-based VIs achieving the highest accuracy of 0.60. Most important features selected via RF for early detection within sensor representations can be found in Tables S6 and S7.

3.2.3 | Performance improvement with multiple sensors

The integration of data from multiple sensors was explored using a RF backward elimination algorithm tailored to sensor-level cross-validation. The full model, which incorporated all nine sensor representations (handheld RGB, Phantom RGB, Inspire RGB, multispectral RGB, ASD RGB, multispectral, ASD, multispectral VIs and ASD VIs), achieved a classification accuracy of 59.6%. The reduced model, resulting from the backward elimination of the ASD sensor representation while retaining other sensor data, notably improved the accuracy to 64.1%. Details on the step-wise performance and sensors removed through backward elimination can be found in Table S8.

TABLE 4 Early detection of soybean wilt - Performance of the random forest model for early detection of canopy wilting severity ('Select' and 'Discard') through different sensor representations. Results show mean and standard deviation of performance across a 10-fold classification aimed at detecting drought stress before it becomes visually observable.

	Sensor				
	Handheld RGB	**Phantom**	**Inspire**	**Multispectral**	**Handheld ASD**
RGB[a]	0.563 ± 0.048	0.519 ± 0.043	0.572 ± 0.046	0.527 ± 0.042	0.500 ± 0.048
Spectral[b]				0.581 ± 0.057	0.580 ± 0.039
VIs[c]				0.596 ± 0.042	0.587 ± 0.054

[a]RGB: RGB Bands were extracted from the sensor multispectral and handheld ASD.
[b]Spectral: 10 multispectral UAV bands, 2151 proximal bands.
[c]VIs: 13 multispectral based VIs, 35 handheld ASD based VIs.

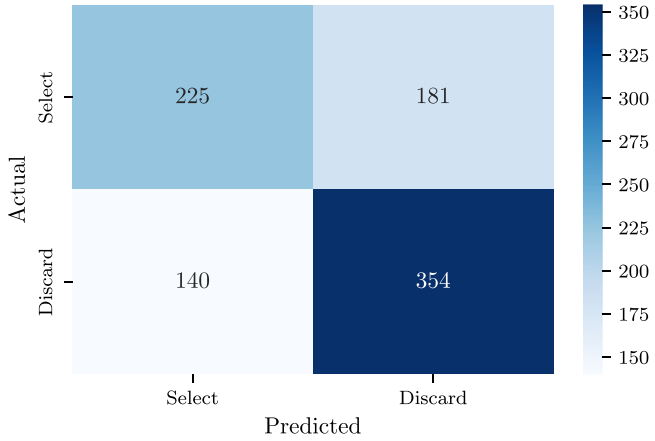

FIGURE 5 Confusion matrix for the highest accuracy model obtained via backward elimination. The model utilized a subset of features - multispectral 'red edge-705', 'EVI', Inspire and handheld 'blue', and ASD '1854nm', '1885nm', '1925nm', 'NMDI', 'WI', and 'NWI-4'. The model achieved an accuracy of 64.3%, with class counts of 406 for 'Select' and 494 for 'Discard'.

Given the complexity and high dimensionality of combined sensor data, establishing an efficient analysis pipeline is crucial for optimizing wilt classification. To this end, the most impactful features across the sensor data were identified, aiming to minimize band correlation, maintain spectral integrity, and reduce computational demands. Employing the RF backward elimination approach at the feature level, the selected combination of 10 features –including multispectral 'red edge-705', 'EVI', Inspire and handheld 'blue', and ASD '1854nm', '1885nm', '1925nm', 'NMDI', 'WI', and 'NWI-4'–led to a classification accuracy of 64.3%. This represents an improvement of 5.8% over models using the full set of sensor features, highlighting the efficacy of strategically selected features predominantly in the near-infrared and short-wave infrared region for early detection of stress. The confusion matrix for this model is presented in Figure 5. To see an expanded list of the top 20 selected features, see Table S9.

3.2.4 | Spectral-temporal changes

In time point 1, an independent Welch's two-sample t-test showed significant differences in the red edge band ($p < 0.1$) spectrum between 'Select' and 'Discard' soybean plots (Figure 6A), demonstrating the utility of this band in comparison to other bands for classifying soybean plots in the early stages of wilting. For the subsequent time points, once stress symptoms became visible, the other multispectral bands showed significant differences ($p < 0.05$) (Figures 6B and 6C). Reflectance intensity declined in both groups for the green wavebands and remained steady in the red to infrared wavebands.

4 | DISCUSSION

A two-pronged approach to drought phenotyping, including (1) classification of canopy wilting severity and (2) early detection of stress, is highly applicable in both breeding and production environments. In this experiment, three RGB, one multispectral, one hyperspectral reflectance ASD sensor, and vegetation indices derived from multispectral and ASD sensors were evaluated for their classification performance for two classes ('Select' and 'Discard') that grouped plots based on their severity level of canopy wilting. The efficiency and accuracy of phenotyping protocols are governed by several sensor specifications, including speed of collection, ground sampling distance (GSD), spectral range, and spectral resolution.

In the classification approach that examines data from the visible range only (400–700 nm) from each sensor modality, the aerial based sensors (Inspire, Phantom, and Multispectral) outperforms the ground based sensors (Handheld RGB and Handheld ASD). Though more work is necessary to confirm, two possibilities arise to explain this trend. The aerial sensors image the full plot area which the ground based sensors are unable to achieve. This larger field of view may provide a more representative view of the whole plot's performance resulting in improved accuracy. This trend is mirrored

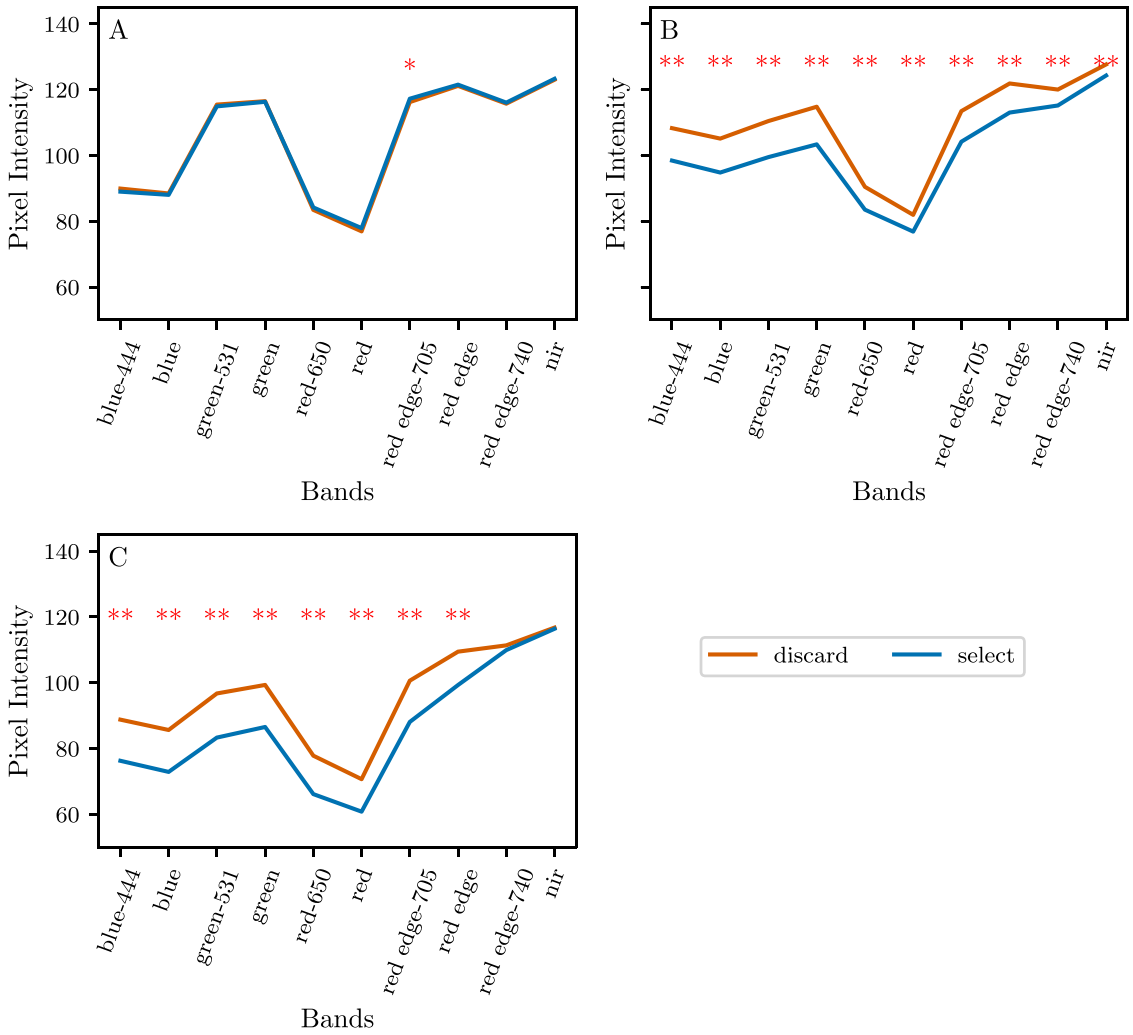

FIGURE 6 Mean spectra of 'Select' and 'Discard' soybean plots of time points (A) Time point 1: 46 DAP, (B) Time point 2: 64 DAP, (C) Time point 3: 81 DAP, at each band of the Micasense RedEdge MX Dual Camera. The symbol * indicates significant differences, with *, and ** indicating differences at $p < 0.1$, and $p < 0.05$, respectively.

in field plot testing in breeding programs, where traits examined on small plots tend to have lower accuracy (Singh et al., 2021c). Furthermore, handheld data collection requires 2.5–3 h for 1350 plots, while a UAV flight requires about 20 min depending on the sensor specifications. Due to the temporal variation of wilting in response to temperature and solar changes throughout the day, longer data collection time duration can introduce higher variance. In this study, we observe higher mean variances in the RGB bands of ground based RGB imagery compared to the RGB bands of the multispectral UAV data. The increase in variation in the handheld imagery could affect the accuracy of stress severity classification. In the comparison of sensor representations, the multispectral spectral bands and multispectral based vegetation indices showed the highest mean classification accuracy of 0.718 and 0.710, respectively. Without multispectral bands in the near-infrared and infrared region, classification accuracy decreases to 0.687 highlighting the importance of the red edge, near-infrared, and infrared regions in improving accuracy of stress classification. Therefore, for classification of drought induced wilting symptoms in this scenario, aerial collection as well as spectral data beyond the visual spectrum prove to be important.

In addition to stress severity classification, early detection is important for production scenarios, where farmers can mitigate drought stress through irrigation, and also in breeding programs to identify early symptom development prior to visual wilting in plant canopies. In this study of early detection of pre-visual soybean canopy wilting due to drought stress, extended wavelengths beyond human vision are shown to be highly important because multispectral and hyperspectral spectral bands and indices outperform all RGB-only representations. In early detection, the multispectral vegetation indices outperform all other sensor representations with a classification accuracy of 0.596. This again supports the usefulness of rapid and full plot phenotyping enabled by UAV

data collection. In time point 1, prior to visual symptom development, susceptible and tolerant canopies differ significantly in the red-edge multispectral band, showing the importance of multispectral imaging for early detection prior to visual symptom development. Implementation of early stress detection strategies would further benefit from the development early stress signatures of multiple stresses, such as disease or insect pressure that may be encountered in production and breeding fields so that correct mitigation strategies can be applied.

Individual features, selected through RF backward elimination, that are important in both severity monitoring as well as early detection include the blue waveband from inspire aerial RGB data. In this study, monitoring stress severity relies on RGB visual spectral information while early detection relies more on spectral data outside the visual range, especially in the short-wave infrared region (SWIR). This study primarily focused on RGB, NIR, and SWIR ranges, but other studies have found links between narrow bands for chlorosis measurement due to short-term water stress and use of wider bands for long-term water stress detection especially in combination with thermal and evapotranspiration data from canopy, soil, and air via the crop water stress index and water deficit index, among others (Moran et al., 1994, 2003). While this study did not investigate yield, several studies have reported various overlapping high-performing features in the prediction of soybean yield, including red wavelengths 665 nm, and 675 nm (Yoosefzadeh-Najafabadi et al., 2021), NDVI, canopy color (Zhou et al., 2021), and multiple features including green, red edge, MSAVI, and NDRE (Zhou et al., 2022). These studies further support the utility of RGB, red-edge, NIR bands while the current study expands on the utility of SWIR for early detection. These features mentioned above would be interesting to pursue in future studies investigating simultaneous phenotyping of canopy wilting and yield production under drought to tease out soybean physiological mechanisms contributing to high yield under drought.

It will also be valuable to investigate the role of soil type and weather related variables in drought response phenotyping, as it has shown usefulness in yield estimation and prediction (Chattopadhyay et al., 2023), though many challenges such as excessive precipitation and need for rain-out shelters in certain environments can can be costly and labor intensive. Drought studies on large genotypic panels can utilize deep learning based methodologies using transfer learning (Chiranjeevi et al., 2021) and image based phenotyping that has shown finer grained classification of diseases in soybean (Rairdin et al., 2022). With finer grained classifications, biotic and abiotic stresses can be combined for meta-analysis (Shook et al., 2021) along with traits such as nodulation and rooting depth (Zubrod, 2022; Falk et al., 2020) that could also contribute to alternative mechanisms of drought tolerance. This current study utilized high-density spectral and two-dimensional (2D) data to detect and classify drought symptoms. However, to further enhance our understanding, future research could leverage the concept of "canopy fingerprinting" by incorporating three-dimensional (3D) point cloud data. This innovative approach would enable a comprehensive examination of plants response to drought stress in both spectral and 3D changes such as leaf angle and altered volumetric distribution (Young et al., 2023; Chiranjeevi et al., 2021). By connecting these multi-modal data, researchers can lay the groundwork for developing sophisticated cyber-agricultural systems that seamlessly integrate sensing, modeling, and actuation processes, thereby enhancing our ability to monitor and respond to dynamic environmental conditions in agriculture (Sarkar et al., 2023). Early detection methods are particularly of use in cyber-agricultural systems as they promote timely actuation and response.

5 | CONCLUSION

Enhancing soybean diversity is imperative, given the current lack of variation in modern soybean cultivars. Introducing diverse genetic material and wild landraces into breeding populations relies on the ability to make binary selections and eliminations within a breeding program. Leveraging aerial imagery, particularly in the visible spectrum, can facilitate this process. Identifying drought stress at an early stage, even before visible symptoms manifest, relies heavily on the utilization of spectral bands in the red edge, NIR, and SWIR regions. This research develops a comprehensive pipeline for integrating data from multiple sensors to classify and detect canopy wilting in soybean. This study emphasizes the significance of high-speed data collection through UAV based sensing and the inclusion of both visual and non-visual bands in achieving accurate detection and severity rating.

ACKNOWLEDGMENTS

The authors would like to acknowledge Brian Scott, Jennifer Hicks, and Ryan Dunn for their effort in planting trials in Muscatine. The authors also thank the many graduate students who assisted in data collection, particularly, Ashlyn Rairdin, Liza Van der Laan, Sam Blair, and Joscif Raigne. The authors would also like to thank Dr. W.T. Schapaugh and Anirudha Powadi for reviewing the paper. The authors sincerely appreciate the funding support from the United Soybean Board (A.K.S., B.F.), Iowa Soybean Association (A.K.S.), USDA CRIS project IOW04714 (A.K.S., A.S.), AI Institute for Resilient Agriculture (USDA-NIFA #2021-67021-35329) (B.G., S.S., A.S., A.K.S.), COALESCE: COntext Aware LEarning for Sustainable CybEr-Agricultural Systems (CPS Frontier #1954556) (S.S., B.G., A.S., A.K.S.), USDA-NIFA FACTS (#2019-67021-29938) (A.S.), Smart Integrated Farm Network for Rural Agricultural Communities (SIRAC) (NSF S&CC #1952045) (A.K.S., S.S.), Raymond F. Baker

Center for Plant Breeding (A.K.S.), and Plant Sciences Institute (A.K.S., S.S., B.G.).

AUTHOR CONTRIBUTIONS

Sarah E. Jones: Conceptualization; data curation; formal analysis; investigation; methodology; writing - original draft. **Timilehin T. Ayanlade**: Conceptualization; formal analysis; investigation; methodology; writing - original draft. **Benjamin Fallen**: Methodology; writing - review & editing. **Talukder Z. Jubery**: Validation; writing - review & editing. **Arti Singh**: Funding acquisition; methodology; resources; supervision; writing - review & editing. **Baskar Ganapathysubramanian**: Funding acquisition; methodology; resources; supervision; writing - review & editing. **Soumik Sarkar**: Conceptualization; funding acquisition; methodology; resources; supervision; validation; writing - review & editing. **Asheesh K. Singh**: Conceptualization; funding acquisition; methodology; project administration; resources; supervision; validation; writing - review & editing. All authors contributed to the final manuscript production. Sarah E. Jones and Timilehin T. Ayanlade contributed equally to this work.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest with this work.

ORCID

Sarah E. Jones https://orcid.org/0000-0002-8780-6933
Benjamin Fallen https://orcid.org/0000-0001-7185-5493
Arti Singh https://orcid.org/0000-0001-6191-9238
Baskar Ganapathysubramanian https://orcid.org/0000-0002-8931-4852
Asheesh K. Singh https://orcid.org/0000-0002-7522-037X

REFERENCES

Breiman, L. (2001). Random forests. *Machine Learning*, *45*, 5–32. https://doi.org/10.1023/A:1010933404324

Brodbeck, C., Sikora, E., Delaney, D., Pate, G., & Johnson, J. (2017). Using unmanned aircraft systems for early detection of soybean diseases. *Advances in Animal Biosciences*, *8*(2), 802–806. https://doi.org/10.1017/S2040470017001315

Carley, C. N., Zubrod, M. J., Dutta, S., & Singh, A. K. (2023). Using machine learning enabled phenotyping to characterize nodulation in three early vegetative stages in soybean. *Crop Science*, *63*(1), 204–226. https://doi.org/10.1002/csc2.20861

Carroll, M., Rairdin, A., Van der Laan, L., Ferela, A., Hanumolu, S., Sarkar, S., Ganapathysubramanian, B., Singh, A., & Singh, A. K. (2024). *Time series gwas for iron deficiency chlorosis tolerance in soybean using aerial imagery*. Manuscript submitted for publication.

Carter, T. E., & Rufty, T. W. (1992). Soybean plant introductions exhibiting drought and aluminum tolerance. In *Adaptation of food crops to temperature and water stress* (pp. 335–346). Asian Vegetable Research and Development Center.

Chamarthi, S. K., Kaler, A. S., Abdel-Haleem, H., Fritschi, F. B., Gillman, J. D., Ray, J. D., Smith, J. R., Dhanapal, A. P., King, C., & Purcell, L. C. (2021). Identification and confirmation of loci associated with canopy wilting in soybean using genome-wide association mapping. *Frontiers in Plant Science*, *12*, 698116–698116. https://doi.org/10.3389/fpls.2021.698116

Chan, J. Y.-L., Leow, S. M. H., Bea, K. T., Cheng, W. K., Phoong, S. W., Hong, Z.-W., & Chen, Y.-L. (2022). Mitigating the multicollinearity problem and its machine learning approach: A review. *Mathematics*, *10*(8), 1283.

Chappelle, E. W., Kim, M. S., & McMurtrey III, J. E. (1992). Ratio analysis of reflectance spectra (RARS): An algorithm for the remote estimation of the concentrations of chlorophyll a, chlorophyll b, and carotenoids in soybean leaves. *Remote Sensing of Environment*, *39*(3), 239–247. https://doi.org/10.1016/0034-4257(92)90089-3

Charlson, D. V., Bhatnagar, S., King, C., Ray, J. D., Sneller, C. H., Carter, T. E., & Purcell, L. C. (2009). Polygenic inheritance of canopy wilting in soybean [Glycine max (L.) Merr.]. *Theoretical and Applied Genetics*, *119*, 587–594. https://doi.org/10.1007/s00122-009-1068-4

Chattopadhyay, S., Gupta, A., Carroll, M., Raigne, J., Ganapathysubramanian, B., Singh, A., & Sarkar, S. (2023). A comprehensive study on soybean yield prediction using soil and hyperspectral reflectance data. https://doi.org/10.20944/preprints202310.1232.v2

Chiranjeevi, S., Sadaati, M., Deng, Z. K., Koushik, J., Jubery, T. Z., Mueller, D., Neal, M. E. O., Merchant, N., Singh, A., Singh, A. K., Sarkar, S., Singh, A., & Ganapathysubramanian, B. (2023). Deep learning powered real-time identification of insects using citizen science data. https://doi.org/10.48550/arxiv.2306.02507

Chiranjeevi, S., Young, T., Jubery, T. Z., Nagasubramanian, K., Sarkar, S., Singh, A. K., Singh, A., & Ganapathysubramanian, B. (2021). Exploring the use of 3d point cloud data for improved plant stress rating. In *AI for Agriculture and Food Systems*. AAAI Press. https://openreview.net/forum?id=kTHfZzeIoH4

Datt, B. (1999). Remote sensing of water content in eucalyptus leaves. *Australian Journal of Botany*, *47*(6), 909–923. https://doi.org/10.1071/BT98042

de Oliveira, B. R., Zuffo, A. M., Steiner, F., Aguilera, J. G., & Gonzales, H. H. S. (2023). Classification of soybean genotypes during the seedling stage in controlled drought and salt stress environments using the decision tree algorithm. *Journal of Agronomy and Crop Science*, *209*(5), 724–733. https://doi.org/10.1111/jac.12654

Dice, J., & Rodziewicz, D. (2020). Drought risk to the agriculture sector. *Economic Review*, *105*(2), 61–86. https://doi.org/10.18651/ER/v105n2Rodziew

Diers, B. W., Specht, J., Rainey, K. M., Cregan, P., Song, Q., Ramasubramanian, V., Graef, G., Nelson, R., Schapaugh, W., Wang, D., Shannon, G., McHale, L., Kantartzi, S. K., Xavier, A., Mian, R., Stupar, R. M., Minchno, J.-M., An, Y.-Q. C., Goettel, W., … Beavis, W. D. (2018). Genetic architecture of soybean yield and agronomic traits. *G3: Genes, Genomes, Genetics*, *8*(10), 3367–3375. https://doi.org/10.1534/g3.118.200332

Esri. (2024). Band arithmetic raster function. Retrieved January 7, 2024. https://pro.arcgis.com/en/pro-app/latest/help/analysis/raster-functions/band-arithmetic-function.htm

Falk, K., Jubery, T., O'Rourke, J., Singh, A., Sarkar, S., Ganapathysubramanian, B., & Singh, A. (2020). Soybean root system architecture trait study through genotypic, phenotypic, and shape-based clusters. *Plant Phenomics*, *2020*, 1925495–1925495. https://doi.org/10.34133/2020/1925495

Fehr, W. R., & Caviness, C. E. (1977). *Stages of soybean development* (Special Report 80). Iowa Agricultural Experiment Station, Iowa Cooperative External Series. Iowa State University of Science and Technology.

Fensholt, R., & Sandholt, I. (2003). Derivation of a shortwave infrared water stress index from modis near-and shortwave infrared data in a semiarid environment. *Remote Sensing of Environment*, *87*(1), 111–121. https://doi.org/10.1016/j.rse.2003.07.002

Fletcher, A. L., Sinclair, T. R., & Allen, L. H. J. (2007). Transpiration responses to vapor pressure deficit in well watered 'slow-wilting' and commercial soybean. *Environmental and Experimental Botany*, *61*(2), 145–151. https://doi.org/10.1016/j.envexpbot.2007.05.004

Gamon, J., Serrano, L., & Surfus, J. (1997). The photochemical reflectance index: an optical indicator of photosynthetic radiation use efficiency across species, functional types, and nutrient levels. *Oecologia*, *112*(4), 492–501. https://doi.org/10.1007/s004420050337

Gao, B.-C. (1996). NDWI–a normalized difference water index for remote sensing of vegetation liquid water from space. *Remote Sensing of Environment*, *58*(3), 257–266. https://doi.org/10.1016/s0034-4257(96)00067-3

Ghosal, S., Blystone, D., Singh, A. K., Ganapathysubramanian, B., Singh, A., & Sarkar, S. (2018). An explainable deep machine vision framework for plant stress phenotyping. *Proceedings of the National Academy of Sciences*, *115*(18), 4613–4618. https://doi.org/10.1073/pnas.1716999115

Gitelson, A. A., Chivkunova, O. B., & Merzlyak, M. N. (2009). Nondestructive estimation of anthocyanins and chlorophylls in anthocyanic leaves. *American Journal of Botany*, *96*(10), 1861–1868. https://doi.org/10.3732/ajb.0800395

Guo, W., Carroll, M. E., Singh, A., Swetnam, T. L., Merchant, N., Sarkar, S., Singh, A. K., & Ganapathysubramanian, B. (2021). UAS-based plant phenotyping for research and breeding applications. *Plant Phenomics*, *2021*, 9840192–9840192. https://doi.org/10.34133/2021/9840192

Herr, A. W., Adak, A., Carroll, M. E., Elango, D., Kar, S., Li, C., Jones, S. E., Carter, A. H., Murray, S. C., Paterson, A., Sankaran, S., Singh, A., & Singh, A. K. (2023). Unoccupied aerial systems imagery for phenotyping in cotton, maize, soybean, and wheat breeding. *Crop Science*, *63*(4), 1722–1749. https://doi.org/10.1002/csc2.21028

Hudak, C., & Patterson, R. (1995). Vegetative growth analysis of a drought-resistant soybean plant introduction. *Crop Science*, *35*(2), 464–471. https://doi.org/10.2135/cropsci1995.0011183X003500020031x

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 Sensing of Environment*, *83*(1-2), 195–213. https://doi.org/10.1016/S0034-4257(02)00096-2

James, G., Witten, D., Hastie, T., Tibshirani, R., & Taylor, J. (2023). *An introduction to statistical learning: With applications in python*. Springer Nature. https://doi.org/10.1023/A:1010933404324

Joalland, S., Screpanti, C., Varella, H. V., Reuther, M., Schwind, M., Lang, C., Walter, A., & Liebisch, F. (2018). Aerial and ground based sensing of tolerance to beet cyst nematode in sugar beet. *Remote Sensing*, *10*(5), 787. https://doi.org/10.3390/rs10050787

Jumrani, K., & Bhatia, V. S. (2018). Impact of combined stress of high temperature and water deficit on growth and seed yield of soybean. *Physiology and Molecular biology of Plants*, *24*(1), 37–50. https://doi.org/10.1007/s12298-017-0480-5

Kaler, A. S., Ray, J. D., Schapaugh, W. T., King, C., & Purcell, L. C. (2017). Genome-wide association mapping of canopy wilting in diverse soybean genotypes. *Theoretical and Applied Genetics*, *130*, 2203–2217. https://doi.org/10.1007/s00122-017-2951-z

King, C., Purcell, L. C., & Brye, K. R. (2009). Differential wilting among soybean genotypes in response to water deficit. *Crop Science*, *49*(1), 290–298. https://doi.org/10.2135/cropsci2008.04.0219

Kpoghomou, B., Sapra, V., & Beyl, C. (1990). Sensitivity to drought stress of three soybean cultivars during different growth stages. *Journal of Agronomy and Crop Science*, *164*(2), 104–109. https://doi.org/10.1111/j.1439-037X.1990.tb00793.x

Merzlyak, M. N., Gitelson, A. A., Chivkunova, O. B., & Rakitin, V. Y. (1999). Non-destructive optical detection of pigment changes during leaf senescence and fruit ripening. *Physiologia Plantarum*, *106*(1), 135–141. https://doi.org/10.1034/j.1399-3054.1999.106119.x

Moran, M., Clarke, T., Inoue, Y., & Vidal, A. (1994). Estimating crop water deficit using the relation between surface-air temperature and spectral vegetation index. *Remote Sensing of Environment*, *49*(3), 246–263. https://doi.org/10.1016/0034-4257(94)90020-5

Moran, S., Zarco-Tejada, P., & Clarke, T. (2003). Crop water stress detection using remote sensing. In J. H. Lehr & J. Keeley (Eds.), *Water encyclopedia* (Vol. 3, pp. 719–724). Wiley. https://doi.org/10.1002/047147844X.aw428

Nagasubramanian, K., Jones, S., Sarkar, S., Singh, A. K., Singh, A., & Ganapathysubramanian, B. (2018). Hyperspectral band selection using genetic algorithm and support vector machines for early identification of charcoal rot disease in soybean stems. *Plant Methods*, *14*(1), 1–13. https://doi.org/10.1186/s13007-018-0349-9

Pantalone, V., Rebetzke, G., Burton, J., & Carter, T. J. (1996). Phenotypic evaluation of root traits in soybean and applicability to plant breeding. *Crop Science*, *36*(2), 456–459. https://doi.org/10.2135/cropsci1996.0011183X003600020039x

Pathan, S. M., Lee, J.-D., Sleper, D., Fritschi, F., Sharp, R., Carter Jr, T., Nelson, R., King, C., Schapaugh, W., Ellersieck, M., & Nguyen, H. T. (2014). Two soybean plant introductions display slow leaf wilting and reduced yield loss under drought. *Journal of Agronomy and Crop Science*, *200*(3), 231–236. https://doi.org/10.1111/jac.12053

Peñuelas, J., Filella, I., Biel, C., Serrano, L., & Save, R. (1993). The reflectance at the 950–970 nm region as an indicator of plant water status. *International Journal of Remote Sensing*, *14*(10), 1887–1905.

Prasad, B., Carver, B. F., Stone, M. L., Babar, M., Raun, W. R., & Klatt, A. R. (2007). Genetic analysis of indirect selection for winter wheat grain yield using spectral reflectance indices. *Crop Science*, *47*(4), 1416–1425. https://doi.org/10.2135/cropsci2006.08.0546

Qi, J., Chehbouni, A., Huete, A., Kerr, Y., & Sorooshian, S. (1994). A modified soil adjusted vegetation index. *Remote Sensing of Environment*, *48*(2), 119–126. https://doi.org/10.1016/0034-4257(94)90134-1

Rairdin, A., Fotouhi, F., Zhang, J., Mueller, D. S., Ganapathysubramanian, B., Singh, A. K., Dutta, S., Sarkar, S., & Singh, A. (2022). Deep learning-based phenotyping for genome wide association studies of sudden death syndrome in soybean. *Frontiers in Plant Science*, *13*, 966244–966244. https://doi.org/10.3389/fpls.2022.966244

Ratliff, L., Ritchie, J., & Cassel, D. (1983). Field-measured limits of soil water availability as related to laboratory-measured properties. *Soil Science Society of America Journal*, *47*(4), 770–775. https://doi.org/10.2136/sssaj1983.03615995004700040032x

Roujean, J.-L., & Breon, F.-M. (1995). Estimating par absorbed by vegetation from bidirectional reflectance measurements. *Remote Sensing of Environment*, *51*(3), 375–384. https://doi.org/10.1016/0034-4257(94)00114-3

Rouse, J. W., Haas, R. H., Schell, J. A., & Deering, D. W. (1973). Monitoring vegetation systems in the great plains with ERTS. In *Third earth resources technology satellite - 1 Symposium: Vol. I; Technical presentations section A* (p. 309). NASA.

Santini, M., Noce, S., Antonelli, M., & Caporaso, L. (2022). Complex drought patterns robustly explain global yield loss for major crops. *Scientific Reports*, *12*(1), 5792. https://doi.org/10.1038/s41598-022-09611-0

Sarkar, S., Ganapathysubramanian, B., Singh, A., Fotouhi Ardakani, F., Kar, S., Nagasubramanian, K., Chowdhary, G., Das, S. K., Kantor, G., Krishnamurthy, A., Merchant, N., & Singh, A. (2023). Cyber-agricultural systems for crop breeding and sustainable production. *Trends in Plant Science*, *29*(2), 130–149. https://doi.org/10.1016/j.tplants.2023.08.001

Sarkar, S., Ramsey, A. F., Cazenave, A.-B., & Balota, M. (2021). Peanut leaf wilting estimation from RGB color indices and logistic models. *Frontiers in Plant Science*, *12*, 658621–658621. https://doi.org/10.3389/fpls.2021.658621

Serrano, L., Penuelas, J., & Ustin, S. L. (2002). Remote sensing of nitrogen and lignin in mediterranean vegetation from aviris data: Decomposing biochemical from structural signals. *Remote Sensing of Environment*, *81*(2), 355–364. https://doi.org/10.1016/S0034-4257(02)00011-1

Shook, J. M., Zhang, J., Jones, S. E., Singh, A., Diers, B. W., & Singh, A. K. (2021). Meta-GWAS for quantitative trait loci identification in soybean. *G3: Genes - Genomes - Genetics*, *11*(7), jkab117. https://doi.org/10.1093/g3journal/jkab117

Sims, D. A., & Gamon, J. A. (2002). Relationships between leaf pigment content and spectral reflectance across a wide range of species, leaf structures and developmental stages. *Remote Sensing of Environment*, *81*(2), 337–354. https://doi.org/10.1016/S0034-4257(02)00010-X

Singh, A., Singh, A. K., Ganapathysubramanian, B., & Sarkar, S. (2016). Machine learning for high-throughput stress phenotyping in plants. *Trends in Plant Science*, *21*(2), 110–124. https://doi.org/10.1016/j.tplants.2015.10.015

Singh, A. K., Ganapathysubramanian, B., Sarkar, S., & Singh, A. (2018). Deep learning for plant stress phenotyping: Trends and future perspectives. *Trends in Plant Science*, *23*(10), 883–898. https://doi.org/10.1016/j.tplants.2018.07.004

Singh, A. K., Singh, A., Sarkar, S., Ganapathysubramanian, B., Schapaugh, W., Miguez, F. E., Carley, C. N., Carroll, M. E., Chiozza, M. V., Chiteri, K. O., Falk, K. G., Jones, S. E., Jubery, T. Z., Mirnezami, S. V., Nagasubramanian, K., Parmley, K. A., Rairdin, A. M., Shook, J. M., Van der Laan, L., … Zhang, J. (2021). High-throughput phenotyping in soybean. In J. Zhou & H. T. Nguyen (Eds.), *High-throughput crop phenotyping* (pp. 129–163). Springer International Publishing. https://doi.org/10.1007/978-3-030-73734-4_7

Singh, D. P., Singh, A. K., & Singh, A. (2021a). Chapter 21 - Breeding for resistance to abiotic stress. In D. P. Singh, A. K. Singh, & A. Singh (Eds.), *Plant breeding and cultivar development* (pp. 399–424). Academic Press. https://doi.org/10.1016/B978-0-12-817563-7.00011-8

Singh, D. P., Singh, A. K., & Singh, A. (2021b). Chapter 28 - Phenomics and machine learning in crop improvement. In D. P. Singh, A. K. Singh, & A. Singh (Eds.), *Plant breeding and cultivar development* (pp. 563–594). Academic Press. https://doi.org/10.1016/B978-0-12-817563-7.00011-8

Singh, D. P., Singh, A. K., & Singh, A. (2021c). Chapter 4 - Primer on population and quantitative genetics. In D. P. Singh, A. K. Singh, & A. Singh (Eds.), *Plant breeding and cultivar development* (pp. 77–127). Academic Press. https://doi.org/10.1016/B978-0-12-817563-7.00011-8

Sloane, R. J., Patterson, R. P., & Carter, T. E. J. (1990). Field drought tolerance of a soybean plant introduction. *Crop Science*, *30*(1), 118–123. https://doi.org/10.2135/cropsci1990.0011183X003000010027x

Song, Q., Yan, L., Quigley, C., Jordan, B. D., Fickus, E., Schroeder, S., Song, B.-H., Charles An, Y.-Q., Hyten, D., Nelson, R., Rainey, K., Beavis, W. D., Specht, J., Diers, B., & Cregan, P. (2017). Genetic characterization of the soybean nested association mapping population. *The Plant Genome*, *10*(2), 1–14. https://doi.org/10.3835/plantgenome2016.10.0109

Tang, Z., Jin, Y., Alsina, M. M., McElrone, A. J., Bambach, N., & Kustas, W. P. (2022). Vine water status mapping with multispectral uav imagery and machine learning. *Irrigation Science*, *40*(4-5), 715–730. https://doi.org/10.1007/s00271-022-00788-w

Valliyodan, B., Ye, H., Song, L., Murphy, M., Shannon, J. G., & Nguyen, H. T. (2017). Genetic diversity and genomic strategies for improving drought and waterlogging tolerance in soybeans. *Journal of Experimental Botany*, *68*(8), 1835–1849. https://doi.org/10.1093/jxb/erw433

Vieira, C. C., & Chen, P. (2021). The numbers game of soybean breeding in the United States. *Crop Breeding and Applied Biotechnology*, *21*(spe). https://doi.org/10.1590/1984-70332021v21Sa23

Vieira, C. C., Sarkar, S., Tian, F., Zhou, J., Jarquin, D., Nguyen, H. T., Zhou, J., & Chen, P. (2022). Differentiate soybean response to off-target dicamba damage based on uav imagery and machine learning. *Remote Sensing*, *14*(7), 1618. https://doi.org/10.3390/rs14071618

Vogelmann, J., Rock, B., & Moss, D. (1993). Red edge spectral measurements from sugar maple leaves. *International Journal of Remote Sensing*, *14*(8), 1563–1575. https://doi.org/10.1080/01431169308953986

Wang, L., & Qu, J. J. (2007). NMDI: A normalized multi-band drought index for monitoring soil and vegetation moisture with satellite remote sensing. *Geophysical Research Letters*, *34*(20), L20405. https://doi.org/10.1029/2007GL031021

Winterhalter, L., Mistele, B., Jampatong, S., & Schmidhalter, U. (2011). High throughput phenotyping of canopy water mass and canopy temperature in well-watered and drought stressed tropical maize hybrids in the vegetative stage. *European Journal of Agronomy*, *35*(1), 22–32.

Xie, Q., Dash, J., Huang, W., Peng, D., Qin, Q., Mortimer, H., Casa, R., Pignatti, S., Laneve, G., Pascucci, S., Dong, Y., & Ye, H. (2018). Vegetation indices combining the red and red-edge spectral information for leaf area index retrieval. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, *11*(5), 1482–1493. https://doi.org/10.1109/JSTARS.2018.2813281

Ye, H., Song, L., Schapaugh, W. T., Ali, M. L., Sinclair, T. R., Riar, M. K., Mutava, R. N., Li, Y., Vuong, T., Valliyodan, B., Pizolato Neto, A., Klepadlo, M., Song, Q., Shannon, J. G., Chen, P., & Nguyen, H. T. (2020). The importance of slow canopy wilting in drought tolerance in soybean. *Journal of Experimental Botany*, *71*(2), 642–652. https://doi.org/10.1093/jxb/erz150

Yoosefzadeh-Najafabadi, M., Earl, H. J., Tulpan, D., Sulik, J., & Eskandari, M. (2021). Application of machine learning algorithms in plant breeding: predicting yield from hyperspectral reflectance in

soybean. *Frontiers in Plant Science*, *11*, 624273–624273. https://doi.org/10.3389/fpls.2020.624273

Young, T. J., Jubery, T. Z., Carley, C. N., Carroll, M., Sarkar, S., Singh, A. K., Singh, A., & Ganapathysubramanian, B. (2023). "Canopy fingerprints" for characterizing three-dimensional point cloud data of soybean canopies. *Frontiers in Plant Science*, *14*, 1141153–1141153. https://doi.org/10.3389/fpls.2023.1141153

Yu, G.-R., Miwa, T., Nakayama, K., Matsuoka, N., & Kon, H. (2000). A proposal for universal formulas for estimating leaf water status of herbaceous and woody plants based on spectral reflectance properties. *Plant and Soil*, *227*(1/2), 47–58. https://doi.org/10.1023/a:1026556613082

Yu, R., Huo, L., Huang, H., Yuan, Y., Gao, B., Liu, Y., Yu, L., Li, H., Yang, L., Ren, L., & Luo, Y. (2022). Early detection of pine wilt disease tree candidates using time-series of spectral signatures. *Frontiers in Plant Science*, *13*, 1000093–1000093. https://doi.org/10.3389/fpls.2022.1000093

Zhou, J., Beche, E., Vieira, C. C., Yungbluth, D., Zhou, J., Scaboo, A., & Chen, P. (2022). Improve soybean variety selection accuracy using uav-based high-throughput phenotyping technology. *Frontiers in Plant Science*, *12*, 768742–768742. https://doi.org/10.3389/fpls.2021.768742

Zhou, J., Zhou, J., Ye, H., Ali, M. L., Chen, P., & Nguyen, H. T. (2021). Yield estimation of soybean breeding lines under drought stress using unmanned aerial vehicle-based imagery and convolutional neural network. *Biosystems Engineering*, *204*, 90–103. https://doi.org/10.1016/j.biosystemseng.2021.01.017

Zhou, J., Zhou, J., Ye, H., Ali, M. L., Nguyen, H. T., & Chen, P. (2020). Classification of soybean leaf wilting due to drought stress using uav-based imagery. *Computers and Electronics in Agriculture*, *175*, 105576. https://doi.org/10.1016/j.compag.2020.105576

Zubrod, M. (2022). Counting the nodules that count, relationships between seed nitrogen and root nodules. *Natural Sciences Education*, *51*(2), e20088. https://doi.org/10.1002/nse2.20088

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Jones, S. E., Ayanlade, T. T., Fallen, B., Jubery, T. Z., Singh, A., Ganapathysubramanian, B., Sarkar, S., & Singh, A. K. (2024). Multi-sensor and multi-temporal high-throughput phenotyping for monitoring and early detection of water-limiting stress in soybean. *The Plant Phenome Journal*, *7*, e70009. https://doi.org/10.1002/ppj2.70009