Utilizing spatial variability from hyperspectral imaging to assess variation in maize seedlings

Sara B. Tirado\textsuperscript{1,2} | Susan St. Dennis\textsuperscript{2} | Tara A. Enders\textsuperscript{3} | Nathan M. Springer\textsuperscript{2}

\textsuperscript{1} Dep. of Agronomy and Plant Genetics, Univ. of Minnesota, St. Paul, MN, 55108, USA
\textsuperscript{2} Dep. of Plant and Microbial Biology, Univ. of Minnesota, St. Paul, MN, 55108, USA
\textsuperscript{3} Dep. of Biology, Hofstra Univ., Hempstead, New York 11549, USA

Correspondence
Nathan Springer, Univ. of Minnesota, Dep. of Plant and Microbial Biology, 140 Gortner Laboratory, 1479 Gortner Avenue, St. Paul, MN 55108, USA.
Email: springer@umn.edu
Tara Enders, Hofstra Univ., Dep. of Biology, 114 Gittleson Hall, Hempstead, NY 11549, USA.
Email: tara.a.enders@hofstra.edu

Funding information
National Science Foundation Plant Genome Award, Grant/Award Number: IOS-1444456

Abstract
There is significant enthusiasm about the potential for hyperspectral imaging to document variation among plant species, genotypes, or growing conditions. However, in many cases the application of hyperspectral imaging is performed in highly controlled situations that focus on a flat portion of a leaf or side-views of plants that would be difficult to obtain in field settings. We were interested in assessing the potential for applying hyperspectral imaging from a top-down view to document variation in genotypes and abiotic stresses for maize (\textit{Zea mays} L.) seedlings grown in controlled environments. A top-down image of a maize seedling includes a view into the funnel-like whorl at the center of the plant with several leaves radiating outward. There is substantial variability in the reflectance profile of different portions of this plant. To deal with the variability in reflectance profiles that arises from this morphology we implemented a method that divides the longest leaf into 10 segments of equal length from the center to the leaf tip. We show that there is large variability in the hyperspectral profiles across leaf segments, which are masked when performing whole-plant averages as tend to be done when analyzing hyperspectral data. We found that using these segments provides improved ability to discriminate different genotypes (B73, Mo17, K11, MS71, PH207) and abiotic stress conditions (heat, cold, or salinity stress) for maize seedlings. This provides an approach that can be implemented to help classify genotype and environmental variation for maize seedlings from a top-down view such as that which would be collected in field settings.

1 | INTRODUCTION

Abiotic stresses cause major yield declines across many crops and can limit production by up to 70\% (Boyer, 1982; Majeed & Muhammad, 2019). Advances in molecular tools have greatly facilitated breeders in efficiently identifying and selecting germplasm with favorable traits such as tolerance to abiotic stresses; however, breeders still rely on obtaining high quality phenotypic data for developing and implementing these methods (Masuka et al., 2012). Phenotyping has become the main bottleneck in making breeding advances because current methods of phenotyping involve a large amount of time and labor. This limits their applications across breeding programs that typically consist of large populations comprised of thousands of lines grown in replicates across multiple environments (Myles et al., 2009). To effectively breed
for tolerance to abiotic stresses, quantifying the severity of the response to a particular stress across different genotypes and their ability to recover from the stress is crucial. This requires temporal measurements of phenotypes linked to the stress response, which increases the complexity in making progress in breeding for such traits.

The development of high throughput phenotyping tools has increased as scientists have sought to gather phenotypic data at larger scale and lower cost. Most of these methods rely on remote sensing techniques to capture images of plants and subsequently processing the images to extract meaningful traits. Sensors that measure different ranges of the electromagnetic spectrum have been applied in agriculture. Red-green-blue (RGB) imaging has been widely used to extract morphological traits linked to plant productivity such as plant height across different crop species in field and indoor settings (Enders et al., 2019; Feng et al., 2018; Varela et al., 2017; Watanabe et al., 2017). Thermal imaging and near infrared combined with visible imaging have been used to extract information of drought stress across grasses and legume crops (Benavente et al., 2013; Biju et al., 2018; Jin et al., 2017; Martynenko et al., 2016; Zhang et al., 2012).

More recently, with the advent of advanced machine and deep learning algorithms, multispectral and hyperspectral sensors that generate large amounts of data at very high spectral and spatial resolutions have been applied in four key areas in plant phenotyping: identification, classification, quantification, and prediction of a particular stress (Singh et al., 2016). Multispectral sensors can monitor several wavelengths not measured by RGB sensors whereas hyperspectral sensors scan a broad series of wavelengths. With hyperspectral imaging, the user can take advantage of hundreds of spectral channels and attempt to uncover materials and biochemical processes, such as the degradation of pigment molecules and changes in water content, within plant tissues that can differentiate and potentially quantify differences across species, genotypes, and stresses. The degradation of pigments such as chlorophyll alters the amount of reflected, absorbed, and transmitted radiation and can therefore be passively captured using spectral imaging (Blackburn, 2006).

Hyperspectral imaging has been applied for the identification and quantification of several bacterial and fungal infections including fusarium head blight and leaf rust in wheat (Triticum aestivum L.) (Alisaac et al., 2019; Ashourloo et al., 2014; Bauriegel et al., 2011; Mahlein et al., 2019; Qiu et al., 2019); powdery mildew in barley (Hordeum vulgare L.) (Thomas et al., 2018; Cercospora leaf spot, powdery mildew, and leaf rust in sugarbeet (Beta vulgaris L.) (Mahlein et al., 2012; Rumpf et al., 2010); and Sclerotinia sclerotorium in oilseed rape plants (Brassica napus L.) (Kong et al., 2018). Diseases caused by bacterial or fungal infections tend to have characteristic features such as bacterial pustules with neighboring chlorotic tissue or necrotic lesions that are picked up and easily distinguished using spectral imaging. However, identifying more subtle symptoms such as those caused by abiotic stresses can be challenging. Pandey et al. (2017) found hyperspectral imaging to be useful in quantifying plant leaf chemical properties that could aid in detecting water and nutrient deficiencies among crops and Obeidat et al. (2018) discovered that spectral indices correlated with chlorophyll content could help distinguish between genotypes and cold-stressed plants in indoor settings. Similarly, Behmann et al. (2014) found that hyperspectral imaging can be used to cluster barley plant pixels into different levels of drought-stress based on amount of chlorosis and senescence. Also, Römer et al. (2012) was able to detect drought stress early in development for cereal grains in both indoor and field settings based on a matrix factorization technique that allows for the computation of how similar a plant pixel is to the typical spectrum of a healthy plant. Another way to identify subtle abiotic stress signals using spectral imaging is by correlating reflectance data with other more laborious, time-consuming, or costly measurements correlated to the stress response. Feng et al (2019) found strong correlations between hyperspectral measurements of okra [Abelmoschus esculentus (L.) Moench] leaves with measurements linked to leaf chlorophyll content and fresh weight traditionally used to assess salt stress across crops.

A common problem when analyzing spectral data of plant surfaces is taking into account uneven light scattering that occurs upon the interaction between incident light and the plant surface being captured (Makdessi et al., 2017). Plant material possesses non-Lambertian reflectance properties and plants themselves contain a large amount of morphological variation causing differences in angle relative to the sensor across leaf segments. Many studies that have evaluated the use of hyperspectral imaging for assessing plant abiotic stresses have utilized indoor setups with uniform, nonreflective backgrounds and have dismissed the effects of plant morphology by securing the plant leaves on a flat background (Obeidat et al., 2018); however, this limits their application in natural settings. Other studies have proposed ways to account for plant architectural variation. Behman et al. (2015) proposed

Core Ideas
- It is important to account for variation along the leaf length for the analysis of hyperspectral images of maize seedlings.
- Leaf segmentation enhances the identification of genotypic responses to stress conditions.
- There are spectral differences across leaf segments that enable genotypic differentiation.
a method to account for differential light scattering by performing geometric calibration of hyperspectral cameras that connects a 3D model with a 2D image and Mohd Asaari et al. (2018) applied a standard normal variate normalization method to correct spectra for uneven illumination effects. Moghimi et al. (2018) circumvented plant architectural differences by identifying endmembers, or single inherent values representative of a given property, indicative of all plant pixels for a given inbred genotype in a given treatment and used these to identify salt stress across wheat lines. Feng et al. (2019), on the other hand, was able to develop an instance segmentation model using deep learning to segment individual okra plant leaves for further evaluation, which is a suitable approach for crops where leaves lie relatively flat horizontally with respect to the sensor.

Variation in plant architecture across different crops species can make finding a single approach for analyzing spectral data challenging; however, it can also be taken advantage of in the context of finding discriminatory patterns within an individual crop. Upon accounting for differences in light scattering of different plant surfaces, the large number of plant pixels representing single or multiple individuals are commonly reduced to a single value such as an endmember (Moghimi et al., 2018) or an average (Mohd Asaari et al., 2018). Looking at all plant pixels throughout the plant can elucidate biochemical processes in response to certain stimuli that vary spatially throughout a plant. This variation along the length of the leaf could be useful for identifying more subtle symptoms that may be masked when reducing the data to a single value per line or treatment or by normalizing the data to account for scattered light due to differences in plant morphology. Moreover, although multiple studies have identified indices that are useful for a particular stress, they have not described how these change due to plant morphology. This study aims to elucidate the effects of morphology and stress on the variation of reflectance values within different segments of plant leaves and compares the ability of reflectance data for different regions of the leaf to resolve genetic and environmental factors relative to reflectance data for the entire plant. Applying hyperspectral analysis to leaf segments could enlighten new mechanisms for identifying, classifying, quantifying, and predicting the onset and recovery of biotic and abiotic stresses where little variation is observed with the naked eye.

# METHODS

## 2.1 Plant growth

Two experiments were conducted, E1 and E2 (Supplementary Table S1). For all experiments, seeds were planted approximately 5 cm (2 inches) below the surface in 655.5 cm$^3$ (40 inch$^3$) D-40 DeePots (Stuewe and Sons, Inc.) containing a 1:1 mix of SunGro horticulture professional growing mix and autoclaved field soil. All plants were grown in Conviron growth chambers with a 16 h 30 °C and 8 h 20 °C day/night cycle and watered every other day.

In experiment E1, five maize genotypes (B73, Mo17, PH207, Ki11, and MS71) were subjected to four treatment conditions (control, cold, heat, salt). Plants for all treatments were grown in standard (non-stress) conditions until 11 days after sowing (DAS) when the stresses were applied for the cold, heat, and salt treatments. The cold-stress treatments were implemented using a Thermo Scientific refrigerated incubator programmed with a 16 h 6 °C and 8 h 2 °C day/night cycle and applied for 48 h. The heat-stress treatments were implemented using a Thermo Scientific refrigerated incubator programmed with a 16 h 39 °C and 8 h 29 °C day/night cycle and applied for 48 h. The salt stress was a single 50 ml 0.75 M NaCl treatment at Zeitgeber Time 2 (ZT2) at 11 DAS. Plants were imaged with our hyperspectral and RGB imaging systems at the end of the stress treatment at ZT2 at 13 DAS.

Three experimental replicates were grown each consisting of three plants per genotype per treatment condition, for a total of nine plants per genotype per treatment.

Experiment E2 consisted of two maize genotypes (Mo17 and PH207) subjected to four treatment conditions (control, low salt, medium salt, and high salt stress). The low, medium, and high salt stress treatments were implemented by a single 50 ml 0.5 M NaCl, 0.75 M NaCl, or 1 M NaCl treatment, respectively, at ZT2 at 11 DAS. Plants were imaged with our hyperspectral and RGB imaging systems before undergoing stress at ZT2 at 11 DAS, at the end of the stress treatment at ZT2 at 13 DAS and 2 d after the stress treatment at 15 DAS. The experiment consisted of 15 plants per genotype per treatment.

## 2.2 Hyperspectral image acquisition

To capture the hyperspectral images, a custom-built line-scanning system from Middleton Spectral Vision was utilized (Supplemental Figure S1). The system contains a Specim V10 spectrograph with a spectral range of 400–1,000 nm and approximately a 1 nm spectral resolution. The spectrograph was mounted on an Imperx IPX-2M30 camera with 1600 × 1200 pixel spatial resolution. When acquiring hyperspectral images, a spectral binning of 2x was applied to obtain an average spectral resolution of 2.3 nm. The system utilized halogen bulbs (Ushio 1003345, 12 V, 20 W) located 110 cm above soil level to provide illumination.
2.3 Hyperspectral data pre-processing and normalization

Black references were gathered by capturing and averaging 10 hyperspectral frames with the camera shutter closed and white references were taken from a white lambertian reference panel included in each image. Intensity values for each pixel at each wavelength were converted to reflectance values by subtracting the dark reference and dividing the result by the difference between the white and dark references (Yoon & Park, 2015). The resulting reflectance values were then normalized by dividing each spectrum by its L_2 norm, or the square root of the sum of the squares of that signature, following the equation:

\[ x_{\text{norm}} = \frac{x}{\sum_{i=1}^{n} \left(x_i^2\right)^{1/2}} \]

where \( x \) is the full vector of reflectance data in image, \( i \) is the response band, \( n \) is the total number of measured wavelengths, and \( x_i \) is the full vector of reflectance data in image for the response band \( i \).

2.4 Segmenting plant material and longest leaf into individual segments

All approaches for identifying, segmenting and extracting data from leaf segments were implemented utilizing custom MATLAB algorithms (MATLAB, 2018a) that are available at https://github.com/SBTirado/HS_LeafSegmentation.git. Plant material was segmented by calculating the NDVI value of each pixel and thresholding to a value of .35 or greater to create a binary plant segmentation mask. This threshold was found to balance maintaining the highest percentage of plant pixels while minimizing the number of mixed pixels in the extracted data. Each hyperspectral image contained three plants of the same genotype and treatment in a defined location. Individual plant objects were identified from the plant material mask using the `bwconncomp` function in MATLAB, which returns connected components with a connectivity of 8 (MATLAB, 2018b). Objects caused by background noise were then removed by only keeping objects that had a minimum of 1,000 pixels and a maximum of 3,000 pixels and allocated an ID (plant A, B, or C) based on their location in the image.

For each plant (plant A, B, and C) of each image, the approximate center of the leaf whorl was identified by manually identifying the \( x \)-\( y \) coordinates of the plant center from an NDVI grayscale representation. This represented the only manual input in the pipeline. Extrema of the plant object were then automatically identified using the `regionprops` function in MATLAB (MATLAB, 2018b). Using the \( x \)-\( y \) coordinates for each terminal extrema and whorl center, the extrema farthest away from the center was identified for each plant representing the tip of the longest leaf. The distance of the center to the longest leaf tip was then divided into 10 equidistant points in linear space and 10 concentric rings were generated utilizing the identified distance between points as the radius (Figure 1a). Each ring segment was used as a mask coupled with the plant material segmentation mask to extract leaf segments along the plant (Figure 1a). To ensure that only segments belonging to a single, constant leaf were kept, only segments that also overlapped with a straight line that extended from the center of the plant to the longest leaf tip were kept. Each of these 10 segments were used as a mask for the hyperspectral image cube, and reflectance data was extracted for wavelengths 420–1,000 nm after trimming off noisy wavelengths. The reflectance values of each pixel in the plant were then normalized by the L2-norm calculated on a noisy wavelengths. This normalized data was then exported for further analysis.

2.5 Outlier detection and removal

Individual leaf segments for each plant were visually assessed by manually assessing the leaf segment binary masks and all data from the given segment excluded in the analysis if the segment encompassed multiple leaves (which happened in cases where part of the given segment was close to the whorl before leaves separated or when leaves overlapped each other), if the leaf segment had less than 25 pixels, or if the segment was not on the primary longest leaf (which occurred in some cases where the leaf curled or overlapped another leaf). A total of 156 plants of the entire 540 had at least one leaf segment excluded (29%); however, the final number of leaf segments excluded was 187 out of the total 5,400 (3%). The majority of excluded segments were located adjacent to the whorl in cases where the leaves were short, and the second leaf segment encompassed multiple leaves.

2.6 Prediction model development and implementation

All cubic support vector machine (SVM) models were developed using the Classification Learner application in MATLAB with specified response and predictor variables as the input (MATLAB, 2018c). Cubic SVM models were selected after testing 13 different machine learning algorithms as these models consistently provided the highest prediction accuracy based on cross-validation results during model training across the different applications specified in this study. The models compared were based on a variety of classifiers including decision trees, discriminant analysis, logistic regression, SVM, nearest neighbor classifiers, ensemble
TIRADO ET AL.

FIGURE 1 Evaluating individual leaf segments. (a) Procedure for extracting hyperspectral data from 10 segments across the longest leaf of each plant. (b) Mean (dark line) and variance in reflectance of individual leaf segments, the entire longest leaf (all leaf segments combined) and the entire plant of a Mo17 control plant at 13 d after sowing. Pixel counts represented in each panel are indicated by the blue text. Leaf segment is indicated at the top of each panel in black text.

2.7 Statistical analyses of leaf segment comparisons

The average spectra for each leaf segment of each plant was compared with the average spectra of each other segment and the average spectra for the entire leaf and the entire plant to assess which leaf segments significantly differ from each other. The comparisons were made by performing a pairwise Wilcoxon Rank Sum Test between all pairwise comparisons of leaf segments using the pairwise.wilcox.test function from the R stats package (R Core Team, 2012; Figure 2d).

2.8 RGB trait data acquisition

Red–green–blue side-view images of each set of plants were collected immediately following hyperspectral data collection using the procedures specified in Enders et al. (2019).

3 RESULTS AND DISCUSSION

In most field settings, hyperspectral images of cereal crops are collected from above, resulting in a top-down view of the plants. We sought to develop approaches for the analysis of hyperspectral images for maize plants collected from above. We obtained raw intensity data using a top-down approach for
FIGURE 2  Reflectance of leaf segments across all Mo17 and PH207 control plants in experiment E2 at 15 d after sowing. (a) Principal component analysis (PCA) biplots of the mean reflectance across all wavelengths of all 10 leaf segments for individual plants. Each datapoint constitutes an individual leaf segment for a single plant. (b) Mean reflectance for the 10 leaf segments, the entire leaf, and the entire plant. All plants were utilized to calculate the mean. Vertical black lines indicate wavelengths at 498, 550, and 625 nm. (c) Distributions of the average reflectance per plant of leaf segments, the entire leaf, and the entire plant at 498, 550, and 625 nm. (d) Adjusted $p$ values from a pairwise Wilcoxon Rank Sum test between pairwise comparisons of leaf segment, entire leaf (L), and whole plant (WP) reflectance, at 498, 550, and 625 nm. Results were adjusted for multiple comparisons using the “holm” method. Black color indicates $p$ value < 0.001.

wavelengths ranging from approximately 400–1,000 nm for several controlled-condition experiments. The experiments contained 11 DAS maize seedlings of multiple inbred genotypes subjected to different environmental treatments. Plants were illuminated by halogen lights, which are oriented in two parallel rows of bulbs on either side of the camera and each image contained three plants (Supplemental Figure S1). Raw intensity values for the resulting images of plants were converted to reflectance using white and dark references and then normalized by their L2 norm (see the Methods section). For each plant, the NDVI values were utilized to identify pixels containing plant tissue and thresholded to generate a binary image mask to extract the reflectance values at each wavelength for entire plants. The longest leaf for each plant was then identified and broken down into 10 segments from the center whorl to the outer leaf tip and reflectance data for each leaf segment was extracted.

Entire plant and leaf segment hyperspectral data was obtained for plants across several different experiments that are summarized in Supplemental Table S1. We sought to address different themes in our analyses of this data. First, we evaluated the potential of top-down hyperspectral data to capture morphological differences by utilizing changes in reflectance for specific regions of a plant. In our analyses, we compared the ability of using whole plant data relative to using specific regions of leaves in the ability to resolve genetic or environmental factors. Second, hyperspectral imaging provides opportunities to identify genetic or environmental variation. We assessed the ability of using hyperspectral imaging data to accurately classify abiotic stress conditions for different genotypes using machine-learning approaches. To achieve these goals, we collected hyperspectral data using the described system for two experiments. The first experiment (E1) consisted of three replicates of five genotypes of maize seedlings under four treatment groups (control, cold, salt, and heat) imaged at a single time point. The chosen genotypes were previously demonstrated to differ in morphological responses to cold stress (Enders et al., 2019). The second experiment (E2) consists of two genotypes under four treatment groups (control and three severities of salt stress) imaged...
at three timepoints (immediately before the stress, 2 d following the stress treatment, and 4 d after following the stress treatment). A total of 540 plant images are represented in the dataset. We have used the dataset to address a series of questions about the ability of hyperspectral imaging to resolve the effects of growth stage, genotype, and various abiotic stresses across portions of maize leaves.

### 3.1 Spatial variation in top-down images of seedlings

The system and approach that was used to generate images of maize seedlings results in a relatively large number of pixels (1,908–13,831) for each plant. These pixels exhibit a range of reflectance values with substantial standard deviation (Supplemental Figure S2). A top-down image of a maize seedling consists of a central whorl from which leaves extend. The whorl has a funnel-like shape with each leaf extending in an arc (Figure 1a). Given the variation in reflectance based on the orientation of the plant surface relative to the lights and camera, there was substantial variability in reflectance from the central whorl to the leaf blade tip. In addition, there is biological variation in gene expression and physiological properties of leaves from the base to the tip (Li et al., 2010). We sought to compare the average normalized reflectance values from each plant within zones extracted along the length of a leaf. To classify relatively consistent zones of a plant leaf, we implemented an approach to divide the longest leaf into 10 sections and identify the plant pixels within each section (Figure 1, see the Methods section for details). This resulted in a set of 10 segments that were used as a mask for the hyperspectral image cube to extract normalized reflectance values by leaf zone.

The average reflectance profiles and variance were compared for the entire leaf relative to each of the 10 leaf sections. Substantial variability is observed in pixel counts across leaf segments due to differences in width across the length of the leaf (Figure 1). Although the overall pattern of reflectance values is generally similar among plant segments, there is substantial variability within and across plant segments for the magnitude of specific patterns. Principal component analysis (PCA) of the reflectance values for each segment of each plant suggests differences in the most outer segments of the leaves relative to more central regions of the leaf (Figure 2a). A comparison of the reflectance profiles across leaf segments focused only on the visible range of the spectrum revealed distinct reflectance profiles between leaf segments near the center or leaf tip relative to the middle portion of the leaf in both Mo17 and PH207 (Figure 2b) and the distribution of average reflectance values for all plants were assessed for each of the leaf segments and the entire leaf and the entire plant (Figure 2c). For several of the segments the reflectance values for these three wavelengths exhibited distributions of values that were significantly different from each other or from the entire plant; however, the whole-leaf and whole-plant averages closely aligned with each other (Figure 2c, 2d). In general, the patterns were similar for the two genotypes. Relative to the values observed for the entire leaf there were often significant differences in the distribution of values seen from the middle and the tip of the leaf. The tip of the leaf is often distinct from many other zones as well. These observations highlight the variability throughout a single leaf and suggest that using average reflectance values for a plant or leaf will likely obscure spatial variation that may occur due to developmental, genetic or environmental factors. Further experiments will be necessary to understand what factors drive this variation along the length of the leaf.

### 3.2 Stable patterns of hyperspectral signal for different stages of seedling growth

The differences in hyperspectral profiles were assessed for PH207 seedlings grown in control conditions that were 11, 13, and 15 DAS to document whether there are differences as seedlings mature and whether the differences among leaf segments are consistent over time (Figure 3). Principal component analysis revealed that differences across leaf segments account for most of the observable variation in reflectance intensity compared to differences observed between days (Figure 3a). Across the three time points, average reflectance values cluster into groups corresponding to leaf segments near the leaf tip, leaf segments toward the middle of the leaf, and leaf segments toward the center whorl. No clustering by date was observed even though plant size and morphology changes were observed based on trait data obtained from side-view RGB images (Supplemental Figure S3, see the Methods section). The profiles of reflectance values were slightly different on the three dates (Figure 3b) but the distribution patterns for the different leaf segments relative to each other remained consistent. Examination of the distribution of values at three wavelengths revealed similar distributions for plants at the three dates and similar trends among the different leaf segments (Figure 3c).

### 3.3 Ability to distinguish genotypes using hyperspectral imaging

Hyperspectral profiling has been widely used for separating different plant species such as weeds and tropical forest tree species (Pantazi et al., 2016; Cochrane, 2000) and
FIGURE 3 Reflectance of leaf segments across all PH207 control plants in experiment E2 across 11, 13, and 15 d after sowing. (a) Principal component analysis (PCA) biplots of the mean reflectance across all wavelengths of all 10 leaf segments for individual plants. Each datapoint constitutes an individual leaf segment for a single plant. (b) Mean reflectance for the 10 leaf segments, the entire leaf, and the entire plant. All plants were utilized to calculate the mean. Vertical black lines indicate wavelengths at 498, 550, and 625 nm. (c) Distributions of the average reflectance per plant of leaf segments, the entire leaf, and the entire plant at 498, 550, and 625 nm across the 3 d of imaging

Fewer studies have used hyperspectral profiling to separate different genotypes or lines of the same species. A study by Obeidat et al. (2018) showed that genotype main effects across short-season maize lines significantly contributed to variation in various spectral reflectance indices and spectral reflectance in the visible and near-infrared range when comparing hyperspectral scans of flat leaves. This variation, particularly in the spectral reflectance across the visible range of the spectrum, was likely due to chlorophyll and carotenoid differences (Obeidat et al., 2018). To assess variation in spectral reflectance among maize genotypes, we applied our leaf segmentation approach to compare the reflectance values across individual leaf segments among the different inbred lines grown in control conditions in experiment E1. Experiment E1 consisted of images from nine plants for five different genotypes. Although the overall average profiles for the entire leaf were relatively similar in the visible range, there were some leaf segments that showed more pronounced differences among genotypes (Figure 4a). In particular, MS71 shows higher reflectance for wavelengths near 550 nm for central segments of the leaf but lower reflectance compared with other genotypes for wavelengths near 500 and 675 nm for segments near the leaf tip. However, these differences are reduced in averages that include all pixels for the entire leaf (Figure 4a).

To further evaluate whether specific leaf segments can be more useful than whole leaves in distinguishing individual genotypes, a cubic SVM model trained on one-sixth of all the pixel values from all plants grown in control conditions in experiment E1 was developed to predict the corresponding genotype of the validation set comprising of the remaining five-sixths of pixels (Supplemental Figure S7). The predictor variables included normalized reflectance values for wavelengths in the visible range as well as the leaf segment of the corresponding pixel. After applying this model on the validation set to predict the genotype, the proportion of pixels per plant that were classified into each genotype for each of the zones of each maize line was determined (Figure 4b). The model was able to correctly identify MS71 pixels in most leaf segments, although the accuracy was somewhat lower in the whorl region (Figure 4b). The fourth leaf segment showed the highest accuracy with an average of 78% of pixels correctly identified as MS71, and the whorl showed the lowest accuracy with an average of 49%. The other genotypes are
Genotypic differences in hyperspectral profiles for all control plants of each genotype in Experiment E1. (a) Average reflectance values across all control plants of each genotype. (b) Average proportion of pixels per plant from all plants of a certain genotype (rows) for each leaf segment (columns) predicted to belong to a certain genotypic class. Error bars indicate the standard deviation in pixel values from the mean.

The B73 genotype was most accurately identified for the central leaf segments with an average prediction accuracy ranging from 47 to 51% across the fifth through seventh leaf segments but was often confused with Ki11 (Figure 4b). Genotypes PH207 and Mo17 generally exhibited relatively low correct prediction accuracies, which varied on average from 23 to 30% for Mo17 and 24 to 39% for PH207, throughout the leaf and were frequently mis-classified as B73 or Ki11 (Figure 4b).

3.4 Ability to distinguish and quantify abiotic stresses using hyperspectral imaging

Two different experiments were performed to investigate the potential to utilize hyperspectral profiling for documenting the effects of abiotic stress on maize seedlings. For experiment E2 we treated two genotypes, Mo17 and PH207, with three different concentrations of salt applied on Day 11 immediately after imaging (Figure 5a). These plants were then imaged 2 and 4 d after the stress application. Two cubic SVM models were trained using one-sixth of all pixels from the control and the medium salt stress (0.75 M NaCl) treatment groups with the goal of comparing the use of whole plants and individual leaf segments for salt stress quantification. In both models, the treatment (control or salt stressed) was the response variable. The first model was trained using pixels randomly selected throughout the entire plant and the normalized reflectance values for wavelengths in the visible range, the genotype, and the day of imaging represented as days after sowing were set as the predictor variables. This model was then used to predict all the pixels in the validation set comprised of the remaining five-sixths of pixels from entire plants in the E2 dataset from all treatment groups into belonging to the control or salt stressed class (Figure 5b). The second model only contained pixels randomly selected from the longest leaf and the corresponding leaf segment was also included as a predictor variable. This model was used to predict all pixels from the longest leaf of all plants (Figure 5c).

As expected, the proportion of pixels classified as salt-stressed were not significantly different for the control and the treatment groups at Day 11 for either model as these images were collected prior to the actual stress treatment application (Figure 5b, 5c). However, at Day 13 and Day 15, 4 and 4 d after application of salt, we observed substantial increases in the proportion of pixels classified as salt stressed (Figure 5b, 5c).
When looking at the classifications based on the entire plant using the first model, the proportion of pixels classified as salt stressed increases at higher concentrations of salt treatment and is higher at Day 15 than at Day 13 (Figure 5b). However, there are a high frequency of pixels in control plants classified as salt stressed in this analysis. A comparison of the predictive ability for different leaf segments revealed substantial variation across the leaf. At Days 13 and 15 the segments from the middle of the leaf had higher correct prediction accuracy than segments near the leaf tip or whorl (Figure 5c). At Day 13, the
The differences in hyperspectral profiles for the different treatments were more severe in some leaf segments compared with others (Figure 6a). A cubic SVM model was developed to predict the treatment using predictor variables of normalized reflectance values, genotype, and leaf segment. The proportion of pixels classified into each condition is shown for each segment of each actual treatment (Figure 6b, 6c). In general, the SVM model was able to differentiate abiotic stress conditions within individual genotypes. For Ki11 there is a high true prediction accuracy for cold and salt stress across all segments of the plants with an average prediction accuracy ranging from 53 to 81% across leaf segments; however, the prediction accuracy is further improved for segments near the middle of the leaf for cold stress, with an average prediction accuracy at the 6th leaf segment of 81%, and the tip of the leaf for salt stress, with an average prediction accuracy of 84% at the 9th leaf segment (Figure 6b). Heat stress is not predicted as accurately with an average prediction accuracy ranging from 26 to 52% across segments and with substantial confusion between control and heat stress (Figure 6b). This likely reflects minimal phenotypic response to heat stress for Ki11. Similar patterns of enhanced prediction accuracy utilizing middle segments of the leaf compared with the entire leaf across treatments are also observed for the other four genotypes (Supplemental Figures S5, S6).

If we compare the accuracies of a representative leaf segment in the center whorl, the middle portion of the leaf, the tip of the leaf, and the entire leaf in predicting the abiotic stress response across genotypes, we observe differences in utility of different leaf segments based on the stress being predicted as well as the genotype (Figure 6c). Leaf segments in the middle portion of the leaf as well as the leaf tip provided the highest true prediction accuracy overall across treatments and genotypes compared with the center whorl and utilizing all segments form the leaf. Leaf segments toward the middle portion of the leaf (Segment 6 in this case) provide a higher prediction accuracy across most genotypes for predicting cold stress and across some genotypes (Mo17, MS71, and PH207) for salt stress; however, the leaf tip was more informative for heat stress and across some genotypes such as Ki11 for salt stress (Figure 6c). Overall, most pixels that were misclassified across genotypes and treatments were predicted to belong to the control treatment group rather than alternate stress treatments. The variability in prediction of stress for different genotypes could reflect differences in genotype responses to abiotic stress. Genotypes that are more tolerant for a specific stress would be more likely to be misclassified as control. This shows the value of hyperspectral imaging in differentiating stress conditions compared to visual inspection or RGB imaging of plants, which result in similar responses across treatments.

There are several limitations to our approach for segmenting the longest leaf and making comparisons of specific leaf segments across dates, genotypes, and treatments. In this work, we performed manual detection of the center of the plant and the remainder of the leaf identification and segmentation process was automated. It is likely that spectral properties or plant shape could be used for automated detection of the center of the whorl. This would also be important for using this approach in field settings in which plants would likely overlap with neighboring plants. Focusing on automated identification of plant centers and leaf segments at a fixed distance from the whorl could be applied at large scale in field settings. Additionally, when comparing leaf segments for varying genotypes or growing conditions the length of the longest leaf may vary as plants exhibit different growth rates for different stress conditions or among genotypes. This results in differing numbers of pixels for the segments being compared. However, since we are segmenting into 10 equally sized regions the relative segmentation of the leaf should remain consistent. Another potential issue is the variation in the angle of the leaf tip. As a leaf emerges from the whorl it
FIGURE 6 Differences in hyperspectral profiles across treatments for all Ki11 and Mo17 plants in Experiment E1. (a) Average reflectance values for all plants in the cold, control, heat, and salt stress treatments of Ki11 and Mo17 in Experiment E1. (b) Average proportion of pixels from all Ki11 plants of a certain treatment (rows) for each leaf segment (columns) predicted to belong to a certain treatment class. Bars represent the mean proportion of all plants per category and error bars represent the standard deviation around the mean. (c) Average proportion of pixels from all plants of a certain treatment (rows) for genotypes across selected leaf segments predicted to belong to a certain treatment class. Bars represent the mean proportion of all plants per genotype and treatment class predicted to belong to a certain treatment.

has an upright angle. As the leaf extends the tip will shift from upright to horizontal to having a downward angle. There is biological variation among plants at the same developmental stage for the angle of the leaf tip. This may result in increased variance for segments near the leaf tip, as noted in our PCA plots (Figure 2a, 3a). However, many stress conditions have visible effects on the leaf margins near the leaf tip and this region provided the best classification for some stresses. One additional potential complication is the presence of mixed pixels that include some plant tissue as well as background. We implemented relatively strict cutoffs to minimize the number of mixed pixels obtained after plant segmentation but there are likely a small number of mixed pixels captured in our plant masks. These may be represented in uneven quantities across
leaf segments with more mixed pixels appearing in narrower segments such as the leaf tip relative to the base of leaf.

4 | CONCLUSIONS

Hyperspectral profiling provides new opportunities for optical analysis of trait variation in crops. Many studies have reported the ability to monitor physiological changes in plants using point profiles of reflectance from a single region of a leaf (Meacham-Hensold et al., 2019; Silva-Perez et al., 2018; Smith et al., 2004; Yendrek et al., 2017). However, there has been less analysis of the ability to separate effects of genotype or environmental conditions using whole plant images. In this study we highlight the potential for using hyperspectral imaging but also show that using averages of whole plants provides less resolution than focusing on specific regions of plants. The variation in spectral profiles observed from the base of the leaf to the tip of the leaf could represent a combination of physiological differences and variation in the plant shape/leaf angle resulting in differing reflectance. Although there have been studies that have shown variation in biochemical compounds across different sections of plant leaves and in response to different growing conditions, the relative ability of hyperspectral data to capture these changes throughout the plant in a non-destructive and high-throughput manner had not previously been looked at (Borsuk & Brodersen, 2019; Song et al., 2018).

In this study, we have not separated the specific physiological factors but instead have simply relied upon the segmentation of hyperspectral data across the leaf to reduce variance and improve stress discrimination. We have also evaluated which portions of the maize seedling leaves provide higher discriminatory accuracies across treatments and found that these vary by condition and genotype.

Our findings highlight the utility of plant segmentation for improved accuracy of genotype or environment predictions using hyperspectral data. It is worthwhile to note that there is not a single region of the leaf with the highest performance. Instead the most accurate regions varied for different stresses or genotypes. We found that there are distinct differences across some leaf segments among genotypes that enable genotypic differentiation. Performing such analysis on large genotypic panels would allow for the identification of groups of genotypes that are morphologically similar. Finding these groups would be important when evaluating environmental responses as these differences could mask the environmental effects if not taken into account when imaging large genotypic panels.

The use of wider panels of genotypes would likely also result in the classification of genotypic groups with similar stress responses, but in this study we focused on improving the methods for stress detection in a small set of variable genotypes. Our classification prediction rates vary substantially. This is likely due to variation among the genotypes. Some genotypes are more tolerant of certain abiotic stresses and we can observe a higher proportion of pixels misclassified into the control classes for these. In contrast, genotypes that are more sensitive to a certain stress exhibit a larger prediction accuracy for stress prediction for the particular treatment. Our experiments are also focused on a single developmental stage and classification of responses in adult plants likely would involve different spectral changes. Although these machine-learning models would not directly transfer to new genotypes, developmental stages, sensor platforms, or environmental conditions such as outdoor field sites, we show that the approach of segmenting leaves for hyperspectral image analysis provides higher discriminatory power for identifying genotypic responses to stress conditions that cannot be accurately assessed utilizing whole-plant averages of spectral profiles. Extracting and analyzing hyperspectral data across plant segments can likely be conducted for larger scale field experiments and may improve the utility of hyperspectral profiles for documenting genotype, environment, and genotype × environment effects.

CONFLICT OF INTEREST

The authors do not have any conflict of interest to declare.

ACKNOWLEDGMENTS

This work was supported by the National Science Foundation Plant Genome Award IOS-1444456. Sara B. Tirado was supported by the Bayer Graduate Student Fellowship. We thank Candice N. Hirsch, Cory D. Hirsch, and Zhikai Liang for providing feedback and insight about the work executed. We also thank Kjell Sandstrom, Shale Demuth, and Danielle Sorensen for helping collect RGB images and Amber Troesch for helping identify plant center coordinates.

AUTHOR CONTRIBUTIONS

S.B. Tirado: Conceptualization, Formal analysis, Investigation, Methodology, Validation, Visualization, Writing-original draft, Writing-review & editing. Susan St. Dennis: Data curation, Methodology, Project administration. T.A. Enders: Conceptualization, Formal analysis, Methodology, Writing-original draft, Writing-review & editing. N.M. Springer: Conceptualization, Visualization, Writing-original draft, Writing-review & editing.

DATA AVAILABILITY STATEMENT

Scripts utilized for processing the hyperspectral images as well as for leaf segmentation are available at https://github.com/SBTirado/HS_LeafSegmentation.git. The hyperspectral dataset for the E1 experiment is available on Cyverse at https://datacommons.cyverse.org/browse/
The hyperspectral dataset for the E2 experiment is available on Cyverse at https://doi.org/10.25739/2wrx-h405.

ORCID
Sara B. Tirado https://orcid.org/0000-0003-0432-091X
Nathan M. Springer https://orcid.org/0000-0002-7301-4759

REFERENCES
Alisaac, E., Behmann, J., Rathgeb, A., Karlovsky, P., Dehne, H., & Mahlein, A. (2019). Assessment of Fusarium infection and mycotoxin contamination of wheat kernels and flour using hyperspectral imaging. Toxins, 11(10). https://doi.org/10.3390/toxins11100556
Ashourloo, D., Mobasher, M. R., & Huete, A. (2014). Evaluating the effect of different wheat rust disease symptoms on vegetation indices using hyperspectral measurements. Remote Sensing, 6(6). https://doi.org/10.3390/rs6061070
Bauriegel, E., Giebel, A., Geyer, M., Schmidt, U., & Herppich, W. B. (2011). Early detection of Fusarium infection in wheat using hyperspectral imaging. Computers and Electronics in Agriculture, 75(2), 304–312. https://doi.org/10.1016/j.compag.2010.12.006
Behmann, J., Mahlein, A., Paulus, S., Kuhlmann, H., Oerke, C., & Plümer, L. (2015). Calibration of hyperspectral close-range push-broom cameras for plant phenotyping. ISPRS Journal of Photogrammetry and Remote Sensing, 106, 172–182. https://doi.org/10.1016/j.isprsjprs.2015.05.010
Behmann, J., Schmitter, P., Steinrücker, J., & Plümer, L. (2014). Ordinal classification for efficient plant stress prediction in hyperspectral data. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XL-7, 29–36. https://doi.org/10.5194/isprsarchives-XL-7-29-2014
Benavente, E., García-Toledano, L., Carrillo, J. M., & Quemada, M. (2013). Thermographic imaging: Assessment of drought and heat tolerance in Spanish germplasm of Brachypodium distachyon. Procedia Environmental Sciences, 19, 262–266. https://doi.org/10.1016/j.proenv.2013.06.030
Biju, S., Fuentes, S., & Gupta, D. (2018). The use of infrared thermal imaging as a non-destructive screening tool for identifying drought-tolerant lentil genotypes. Plant Physiology and Biochemistry, 127, 11–24. https://doi.org/10.1016/j.plaphy.2018.03.005
Blackburn, G. A. (2006). Hyperspectral remote sensing of plant pigments. Journal of Experimental Botany, 58(4), 855–867. https://doi.org/10.1093/jxbер/erl123
Borsuk, A. M., & Brodersen, C. R. (2019). The spatial distribution of chlorophyll in leaves. Plant Physiology, 180(3) 1406–1417. https://doi.org/10.1104/pp.19.00094
Boyer, J. S. (1982). Plant productivity and environment. Science, 218(4572), 443–448. https://doi.org/10.1126/science.218.4571.443
Cochrane, M. A. (2000). Using vegetation reflectance variability for species level classification of hyperspectral data. International Journal of Remote Sensing, 21(10), 2075–2087. https://doi.org/10.1080/01431160050021303
Enders, T. A., St. Dennis, S., Oakland, J., Callen, S. T., Gehan, M. A., Miller, N. D., Spalding, E. P., Springer, N. M., & Hirsch, C. D. (2019). Classifying cold-stress responses of inbred maize seedlings using RGB imaging. Plant Direct, 3(1). https://doi.org/10.1002/pld3.104
Meacham-Hensold, K., Montes, C. M., Wi, J., Guan, K., Fu, P., Ainsworth, E. A., Pederson, T., Moore, C. E., Brown, K. L., Raines, C., & Bernacchi, C. J. (2019). High-throughput field phenotyping using hyperspectral reflectance and partial least squares regression (PLSR) reveals genetic modifications to photosynthetic capacity. Remote Sensing of Environment, 231, 111176. https://doi.org/10.1016/j.rse.2019.04.029

Moghim, A., Yang, C., Miller, M. E., Kianian, S. F., & Marchetto, P. M. (2018). A novel approach to assess salt stress tolerance in wheat using hyperspectral imaging. Frontiers in Plant Science, 9, 1182. https://doi.org/10.3389/fpls.2018.01182

Mohd Asaari, M. S., Mishra, P., Mertens, S., Dhondt, S., Inzé, D., Wuys, N., & Scheuners, P. (2018). Close-range hyperspectral image analysis for the early detection of stress responses in individual plants in a high-throughput phenotyping platform. ISPRS Journal of Photogrammetry and Remote Sensing, 138, 121–138. https://doi.org/10.1016/j.isprsjprs.2018.02.003

Myles, S., Peiffer, J., Brown, P. J., Ersoz, E. E., Zhang, Z., Costich, D. E., & Bluckler, E. S. (2009). Association mapping: Critical considerations shift from genotyping to experimental design. The Plant Cell, 21(8), 2194–2202. https://doi.org/10.1105/tpc.109.068437

Obeidat, W., Avila, E., Earl, H., & Lukens, L. (2018). Leaf spectral reflectance of maize seedlings and its relationship to cold tolerance. Crop Science, 58, 2569–2580. https://doi.org/10.2135/cropsci2018.02.0115

Pandey, P., Ge, Y., Stoerger, V., & Schnable, J. C. (2017). High-throughput In vivo analysis of plant leaf chemical properties using hyperspectral imaging. Frontiers in Plant Science, 8, 1348. https://doi.org/10.3389/fpls.2017.01348

Pantazi, X., Moshou, D., & Bravo, C. (2016). Active learning system for weed species recognition based on hyperspectral sensing. Biosystems Engineering, 146, 193–202. https://doi.org/10.1016/jbiosystemseng.2016.01.014

Qiu, R., Yang, C., Moghim, A., Zhang, M., Stefferson, B., & Hirsch, C. (2019). Detection of Fusarium head blight in wheat using a deep neural network and color imaging. Remote Sensing, 11(22), 2658. https://doi.org/10.3390/rs11222658

R Core Team. (2012). R: A language and environment for statistical computing. R Foundation for Statistical Computing. http://www.R-project.org/

Römer, C., Wahabzada, M., Ballvora, A., Pinto, F., Rossini, M., Pani-gada, C., Behmann, J., Léon, J., Thurau, C., Baukhage, C., Kersting, K., Rascher, U., & Plümer, L. (2012). Early drought stress detection in cereals: Simplex volume maximisation for hyperspectral image analysis. Functional Plant Biology, 39(11), 878–890. https://doi.org/10.1071/FP12060

Rumpf, T., Mahlein, A., Steiner, U., Oerke, E., Dehne, H., & Plümer, L. (2010). Early detection and classification of plant diseases with Support Vector Machines based on hyperspectral reflectance. Computers and Electronics in Agriculture, 74(1), 91–99. https://doi.org/10.1016/j.compag.2010.06.009

Silva-Perez, V., Molero, G., Serbin, S. P., Condon, A. G., Reynolds, M. P., Furbank, R. T., & Evans, J. R. (2018). Hyperspectral reflectance as a tool to measure biochemical and physiological traits in wheat. Journal of Experimental Botany, 69(3), 483–496. https://doi.org/10.1093/jxb/erx421

SUPPORTING INFORMATION
Additional supporting information may be found online in the Supporting Information section at the end of the article.

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

Smith, K. L., Steven, M. D., & Colls, J. J. (2004). Use of hyperspectral derivative ratios in the red-edge region to identify plant stress responses to gas leaks. Remote Sensing of Environment, 92(2), 207–217. https://doi.org/10.1016/j.rse.2004.06.002

Song, H., Li, Y., Zhou, L., Xu, Z., & Zhou, G. (2018). Maize leaf functional responses to drought episode and rewatering. Agricultural and Forest Meteorology, 249, 57–70. https://doi.org/10.1016/j.agrformet.2017.11.023

Thomas, S., Behman, J., Steier, A., Kraska, T., Muller, O., Rascher, U., & Mahlein, A. (2018). Quantitative assessment of disease severity and rating of barley cultivars based on hyperspectral imaging in a non-invasive, automated phenotyping platform. Plant Methods, 14, 45. https://doi.org/10.1186/s13007-018-0313-8

Varela, S., Assf, Y., Prasad, P. V., Peralta, N. R., Griffin, T. W., Sharda, A., Ferguson, A., & Ciampitti, I. A. (2017). Spatio-temporal evaluation of plant height in corn via unmanned aerial systems. Journal of Applied Remote Sensing, 11(3). https://doi.org/10.1117/1.JRS.11.036013

Watanabe, K., Guo, W., Ari, K., Kakanashi, H., Kajiya-Kanegae, H., Kobayashi, M., Yan, K., Tokunaga, T., Fujiwara, T., Tsutsumi, N., & Iwata, H. (2017). High-throughput phenotyping of sorghum plant height using an unmanned aerial vehicle and its application to genomic prediction modeling. Frontiers of Plant Science, 8, 421. https://doi.org/10.3389/fpls.2017.00421

Yendrek, C. R., Tomaz, T., Montes, C. M., Cao, Y., Morse, A. M., Brown, P. J., McIntyre, L. M., Leakey, A. D. B., & Ainsworth, E. A. (2017). High-throughput phenotyping of maize leaf physiological and biochemical traits using hyperspectral reflectance. Plant Physiology, 173, 614–626. https://doi.org/10.1104/pp.16.01447

Yoon, S. C., & Park, B. (2015). Hyperspectral image processing methods. In B. Park & R. Lu (Eds.), Hyperspectral imaging technology in food and agriculture. Food Engineering Series. Springer.

Zhang, Q., Li, Q., & Zhang, G. (2012). Rapid determination of leaf water content using VIS/NIR spectroscopy analysis with wavelength selection. Journal of Spectroscopy, 27. https://doi.org/10.1155/2012/276795

How to cite this article: Tirado SB, St. Dennis S, Enders TA, Springer NM. Utilizing spatial variability from hyperspectral imaging to assess variation in maize seedlings. Plant Phenome J. 2021;4:e20013. https://doi.org/10.1002/ppj2.20013