

*Using spectral distance, spectral angle
and plant abundance derived from
hyperspectral imagery to characterize crop
yield variation*

Chenghai Yang & James H. Everitt

Precision Agriculture

An International Journal on Advances in
Precision Agriculture

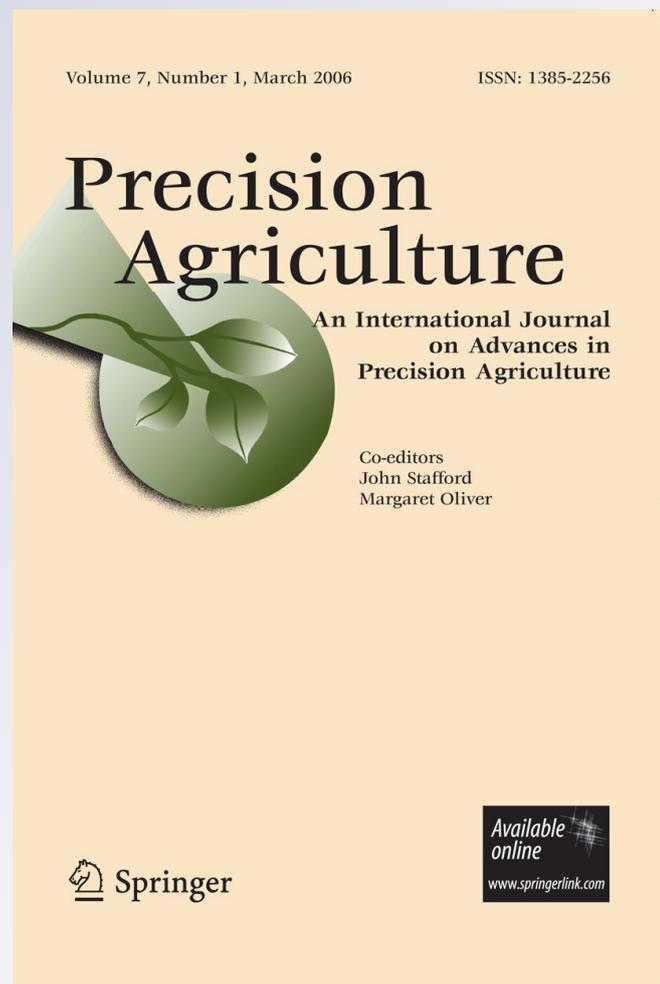
ISSN 1385-2256

Volume 13

Number 1

Precision Agric (2012) 13:62-75

DOI 10.1007/s11119-011-9248-z



Your article is protected by copyright and all rights are held exclusively by Springer Science+Business Media, LLC (outside the USA) . This e-offprint is for personal use only and shall not be self-archived in electronic repositories. If you wish to self-archive your work, please use the accepted author's version for posting to your own website or your institution's repository. You may further deposit the accepted author's version on a funder's repository at a funder's request, provided it is not made publicly available until 12 months after publication.

Using spectral distance, spectral angle and plant abundance derived from hyperspectral imagery to characterize crop yield variation

Chenghai Yang · James H. Everitt

Published online: 21 October 2011

© Springer Science+Business Media, LLC (outside the USA) 2011

Abstract Vegetation indices (VIs) derived from remote sensing imagery are commonly used to quantify crop growth and yield variations. As hyperspectral imagery is becoming more available, the number of possible VIs that can be calculated is overwhelmingly large. The objectives of this study were to examine spectral distance, spectral angle and plant abundance (crop fractional cover estimated with spectral unmixing) derived from all the bands in hyperspectral imagery and compare them with eight widely used two-band and three-band VIs based on selected wavelengths for quantifying crop yield variability. Airborne 102-band hyperspectral images acquired at the peak development stage and yield monitor data collected from two grain sorghum fields were used. A total of 64 VI images were generated based on the eight VIs and selected wavelengths for each field in this study. Two spectral distance images, two spectral angle images and two abundance images were also created based on a pair of pure plant and soil reference spectra for each field. Correlation analysis with yield showed that the eight VIs with the selected wavelengths had r values of 0.73–0.79 for field 1 and 0.82–0.86 for field 2. Although all VIs provided similar correlations with yield, the modified soil-adjusted vegetation index (MSAVI) produced more consistent r values (0.77–0.79 for field 1 and 0.85–0.86 for field 2) among the selected bands. Spectral distance, spectral angle and plant abundance produced similar r values (0.76–0.78 for field 1 and 0.83–0.85 for field 2) to the best VIs. The results from this study suggest that either a VI (MSAVI) image based on one near-infrared band (800 or 825 nm) and one visible band (550 or 670 nm) or a plant abundance image based on a pair of pure plant and soil spectra can be used to estimate relative yield variation from a hyperspectral image.

Keywords Hyperspectral imagery · Plant abundance · Spectral angle · Spectral distance · Spectral unmixing · Vegetation index · Yield

C. Yang (✉) · J. H. Everitt
USDA-ARS Kika de la Garza Subtropical Agricultural Research Center, 2413 E. Highway 83,
Weslaco, TX 78596, USA
e-mail: chenghai.yang@ars.usda.gov

Introduction

Hyperspectral imagery contains tens to hundreds of bands of spectral data and therefore provides much finer spectral information than multispectral imagery. Traditionally, broadband vegetation indices (VIs) derived from multispectral imagery are commonly used to characterize crop growing conditions and productivity such as leaf area index (LAI) (Baret and Guyot 1991), biomass (Moran et al. 1995) and yield (Wiegand et al. 1991; Yang and Anderson 1999). These VIs are typically a sum, difference, ratio or other combination of reflectance observations from two or more wavebands. The simple ratio index (SRI) (Jordan 1969) and the normalized difference vegetation index (NDVI) (Rouse et al. 1973) derived from the red band and near-infrared (NIR) band are two of the earliest and most widely used VIs. More recently, many new VIs have been developed to improve the linearity and sensitivity such as the modified simple ratio (MSR) (Chen 1996) and the renormalized difference vegetation index (RDVI) (Rougean and Breon 1995) and to compensate for the effect of soil background such as the soil-adjusted vegetation index (SAVI) (Huete 1988) and the modified SAVI (MSAVI) (Qi et al. 1994). In addition to the two-band VIs, several three-band VIs have also been developed, including the modified chlorophyll absorption in reflectance index (MCARI) (Daughtry et al. 2000) and the triangular vegetation index (TVI) (Broge and Leblanc 2000). MCARI and TVI were originally proposed for chlorophyll estimation and require 700-nm and 750-nm wavelengths in the red edge, respectively. Haboudane et al. (2004) proposed two modified versions of MCARI (MCARI1 and MCARI2) and two modified versions of TVI (MTVI1 and MTVI2) by replacing the red edge wavelengths with the 800-nm wavelength to lower the sensitivity to chlorophyll effects, increase the sensitivity to LAI changes and reduce soil and atmospheric effects. MCARI1 and MTVI1 are mathematically the same vegetation index. Similarly, the MCARI2 and MTVI2 indexes, obtained respectively from MCARI1 and MTVI1 by the introduction of a soil correction factor represent a single index.

For a multispectral image, which typically contains three bands to seven bands as in Landsat-7 ETM+, there are only one green band, one red band and one NIR band. The multispectral image can be easily converted to one single VI image based on the selected VI. However, a hyperspectral image contains dozens of blue, green, red or NIR narrow bands and the number of VIs that can be calculated is overwhelmingly large. For example, if a hyperspectral image has 40 red bands and 50 NIR bands, the number of SRIs or NDVIs can be as many as 2 000. It is not always practical to calculate and examine all the possible VIs (i.e., the 2 000 NDVI images in the example) to identify the best VI for a particular application. Therefore, the commonly used multispectral VIs have been applied to hyperspectral imagery based on selected narrow bands. For example, the 800 and 670 nm narrow bands extracted from airborne hyperspectral imagery were used as the NIR band and red band, respectively, in the broadband VIs for estimating crop LAI (Haboudane et al. 2004) and crop yield (Zarco-Tejada et al. 2005). Other combinations of narrow bands derived from hyperspectral imagery have also been used for estimating crop growth parameters (Ray et al. 2006; Wu et al. 2010).

Thenkabail et al. (2000) used ground reflectance data measured in 490 discrete narrow bands between 350 and 1 050 nm to characterize yield and other crop biophysical variables. They calculated narrow-band NDVI-type indices with all possible two-band combinations of the 490 bands and identified the best band centers and bandwidths for each crop variable. Based on the results from NDVIs and other hyperspectral indices, they recommended 12 hyperspectral bands (495, 525, 550, 568, 668, 682, 696, 720, 845, 920, 982 and 1 025 nm) for estimating agricultural crop biophysical information.

Yang et al. (2004) applied stepwise regression analysis to grain yield monitor data and 102-band airborne hyperspectral imagery and identified four optimum bands for one field and seven different bands for a second field for estimating yield. Clearly, the identified optimum bands were the best for the particular datasets from which they were derived and may not be the best for different datasets. To avoid the need for band selection and make use of all the bands in hyperspectral imagery, Yang et al. (2007) used linear spectral unmixing to generate plant and soil abundance images from airborne hyperspectral imagery for quantifying the variation in crop yield. Yang et al. (2008) also applied spectral angle mapper (SAM) to airborne hyperspectral imagery to derive spectral angle images for the same purpose. Both linear spectral unmixing and SAM have been commonly used in remote sensing for image classification (Bateson and Curtiss 1996; Dennison et al. 2004; Franke et al. 2009). Fractional abundance images determined by linear spectral unmixing may be preferred to NDVI as all the bands in the image are used (Bateson and Curtiss 1996). Yang et al. (2007, 2008) demonstrated that both plant abundance images and spectral angle images provided better r values with yield than most of the 5 151 possible narrow band NDVIs derived from the 102-band hyperspectral images.

Crop yield is perhaps the most important piece of information for crop management in precision agriculture. Despite the commercial availability and increased use of yield monitors, not all harvesters are equipped with them. Relative yield maps derived from remote sensing imagery can be used as an alternative for both within-season and post-season management. Relative yield maps can be derived using any of the two-band and three-band VIs and they can also be generated either from plant abundance images or from spectral angle images. Two or three center wavelengths (e.g., 800 nm for NIR, 670 nm for red and 550 nm for green) have to be selected to calculate the VIs, while plant and soil endmembers need to be defined to derive plant abundance and spectral angle. Although many VIs are available and different center wavelengths have been suggested (Thenkabail et al. 2000; Haboudane et al. 2004; Zarco-Tejada et al. 2005; Ray et al. 2006), it is still not clear which VIs and wavelengths should be used to estimate relative yield variation from a hyperspectral image. Therefore, the first objective of this study was to compare five two-band VIs (SRI, NDVI, RDVI, SAVI and MSAVI) based on the 800 and 670 nm center wavelengths and three three-band VIs (MCARI1, TVI and MTVI2) for yield estimation. These VIs have been found to be good indicators of LAI and/or crop yield (Haboudane et al. 2004; Zarco-Tejada et al. 2005). The second objective was to apply one NIR wavelength (825 nm) and eight visible wavelengths (495, 525, 550, 568, 668, 682, 696 and 720 nm) suggested by Thenkabail et al. (2000) to the five two-band VIs for yield estimation. The last objective was to relate yield to spectral distance, spectral angle and plant abundance derived from airborne hyperspectral imagery based on a pair of reference plant and soil spectra for each field and compare the correlations with those from the above VIs.

Methods

Hyperspectral imagery and yield data

The airborne imagery acquired at the boot to half-bloom stages (around peak canopy development) and yield data collected from two grain sorghum fields (19 and 14 ha in size) in south Texas were used for this study. The description of the study sites, image acquisition, rectification and calibration as well as yield data collection and processing is given in the article by Yang et al. (2008). The airborne imagery was acquired using a

hyperspectral imaging system described by Yang et al. (2003) and contained 102 usable bands with center wavelengths from 477.2 to 843.7 nm at 3.63 nm intervals. The swath of the imagery was 640 pixels and the radiometric resolution was 12 bits. The imagery was converted to a reflectance range of 0–1 based on three 8 × 8 m tarpaulins with reflectance values of 4, 32 and 48%. The images were resampled to 1-m pixel resolution using the nearest-neighbor algorithm during the rectification process. The root mean square (RMS) errors for the rectified hyperspectral images were 3.8 and 4.3 m for fields 1 and 2, respectively, based on first-order polynomial transformations. Yield data were collected using an Ag Leader Yield Monitor 2000 system (Ag Leader Technology, Ames, Iowa, USA). Both the image and yield data were aggregated to 9 m resolution (close to the effective cutting width of the harvester) and the number of aggregated samples was 2 265 for field 1 and 1 658 for field 2.

Hyperspectral vegetation indices (VIs)

Eight VIs listed in Table 1 were selected as the hyperspectral VIs to be calculated for this study based on their performance for the estimation of LAI and yield by other researchers. The five two-band hyperspectral VIs (SRI, NDVI, RDVI, SAVI and MSAVI) were first calculated based on the 800 and 670 nm center wavelengths suggested by Haboudane et al. (2004) and the three three-band hyperspectral VIs (MCARI1, TVI and MTVI2) were calculated based on the center wavelengths given in the formulas.

The second group of hyperspectral VIs was calculated based on the five two-band VIs using the center wavelengths suggested by Thenkabail et al. (2000). The 12 suggested center wavelengths include one blue band (495 nm), three green bands (525, 550 and 568 nm), three red bands (668, 682 and 696 nm), one red-edge band (720 nm) and four NIR bands (845, 920, 982 and 1 025 nm). Because of the narrow spectral range of the

Table 1 Vegetation indices calculated from hyperspectral imagery in this study

Vegetation index	Equation	Reference
Simple ratio index (SRI)	$SRI = R_{NIR}/R_{Red}$	Jordan (1969)
Normalized difference vegetation index (NDVI)	$NDVI = (R_{NIR} - R_{Red}) / (R_{NIR} + R_{Red})$	Rouse et al. (1973)
Renormalized difference vegetation index (RDVI)	$RDVI = (R_{NIR} - R_{Red}) / \sqrt{R_{NIR} + R_{Red}}$	Rougean and Breon (1995)
Soil-adjusted vegetation index (SAVI)	$SAVI = 1.5(R_{NIR} - R_{Red}) / (R_{NIR} + R_{Red} + 0.5)$	Huete (1988)
Modified SAVI (MSAVI)	$MSAVI = 0.5 \left[2R_{NIR} + 1 - \sqrt{(2R_{NIR} + 1)^2 - 8(R_{NIR} - R_{Red})} \right]$	Qi et al. (1994)
Modified chlorophyll absorption in reflectance index (MCARI1)	$MCARI1 = 1.2[2.5(R_{800} - R_{670}) - 1.3(R_{800} - R_{550})]$	Haboudane et al. (2004)
Triangular vegetation index (TVI)	$TVI = 0.5[120(R_{750} - R_{550}) - 200(R_{670} - R_{550})]$	Broge and Leblanc (2000)
Modified TVI (MTVI2)	$MTVI2 = \frac{1.5[1.2(R_{800} - R_{550}) - 2.5(R_{670} - R_{550})]}{\sqrt{(2R_{800} + 1)^2 - (6R_{800} - 5\sqrt{R_{670}}) - 0.5}}$	Haboudane et al. (2004)

hyperspectral data used in this study, there was no NIR center wavelength to match the four suggested NIR center wavelengths. Thenkabail et al. (2000) stated in the description of the suggested 845 nm wavelength that a broad band or a narrow band in the NIR shoulder (845 ± 35) will provide the same results due to the near-uniform reflectance throughout the NIR shoulder. In order to examine the sensitivity of NIR wavelengths on the results, the 810, 825 and 840 nm wavelengths were selected as the NIR band and the other eight visible bands as the red band in the NDVI formula to calculate the 24 possible NDVI-type indices as well as their correlations with yield for each field. The results showed that the three NIR wavelengths gave essentially the same results. Therefore, the 825 nm wavelength and the eight visible wavelengths were used to calculate hyperspectral indices based on the five two-band VIs.

Spectral distance, spectral angle and plant abundance

Spectral distance is a spectral measure commonly used in unsupervised classification and supervised minimum distance classification (Campbell 2002). The spectral distance between a pixel spectrum and a reference spectrum can be calculated by Euclidean distance as follows:

$$d = \sqrt{\sum_{i=1}^n (y_i - r_i)^2} \quad (1)$$

where d is the spectral distance, y_i is the reflectance in band i for a pixel, r_i is the reflectance in band i for a reference and n is the number of bands in the image.

Spectral distance has the potential to quantify the variation in crop growth and yield. For grain sorghum, the best phenological stage is around the peak vegetative development for yield estimation (Yang and Everitt 2002). For example, if a pure healthy crop canopy is selected as the reference, the spectral distance between healthy plants and the reference will be small, whereas the spectral distance between stressed plants and the reference will be large. Therefore, spectral distance can be used as an indirect indicator of plant health and cover abundance.

Spectral angle is a spectral measure used in spectral angle mapper (SAM), a spectral classification technique that assigns pixels to classes based on the spectral angles between image pixel spectra and reference spectra (Kruse et al. 1993). The spectral angle between a pixel spectrum and a reference spectrum can be calculated by the following formula:

$$\alpha = \cos^{-1} \frac{\sum_{i=1}^n y_i r_i}{\sqrt{\sum_{i=1}^n y_i^2} \sqrt{\sum_{i=1}^n r_i^2}} \quad (2)$$

where α is the spectral angle between a pixel spectrum and a reference spectrum measured in radians or degrees, y_i is the reflectance in band i for the pixel, r_i is the reflectance in band i for a reference and n is the number of bands in the image. Similar to spectral distance, spectral angle is also an indirect measure of plant health and abundance. When healthy crop canopy is selected as the reference, small angle values correspond to healthy plants and large values correspond to stressed plants.

Fractional plant cover or plant abundance within a pixel can be estimated using linear spectral unmixing. Linear spectral unmixing models each spectrum in a pixel as a linear combination of a finite number of spectrally pure spectra of the components in the image, weighted by their fractional abundances (Adams et al. 1986; Garcia-Haro et al. 1996).

If a component such as a healthy crop canopy or a bare soil surface occupies the whole pixel, then the pixel spectrum can be considered as the reference spectrum or endmember spectrum of the ground component. For agricultural crop fields, crop plants and bare soil can be selected as the two meaningful ground components or endmembers for spectral unmixing analysis (Yang et al. 2007). Thus a simple linear spectral unmixing model has the following form:

$$y_i = r_{i1}x_1 + r_{i2}x_2 + \varepsilon_i, \quad i = 1, 2, \dots, n \quad (3)$$

where y_i is the reflectance in band i for a pixel, r_{i1} and r_{i2} are the known reflectance in band i for pure crop plants and bare soil, respectively, x_1 and x_2 are the unknown fractional abundance for plants and soil, respectively, ε_i is the residual between actual and modeled reflectance for band i and n is the number of spectral bands. This model is referred to as the unconstrained linear spectral unmixing model. For constrained linear spectral unmixing, x_1 and x_2 should sum to unity.

To calculate plant abundance, a plant spectrum and a soil spectrum are needed. In comparison, only one reference spectrum is necessary to calculate spectral distance and spectral angle. Reference or endmember spectra can be obtained directly from the image or measured on the ground. In this study, healthy crop plants and bare soil were selected as the relevant endmembers. A pair of plant and soil spectra was extracted from each image to represent pure and healthy plants and bare soil for the respective field. To obtain pure spectra for crop plants, 50 pixels that had a bright red color on a color-infrared (CIR) image (corresponding to healthy plants with a LAI of about 3.5 and high yielding areas) were first identified from each image. Similarly, 50 pixels that contained pure bare soil were identified from each image (corresponding to non-vegetative and zero yielding areas). The endmember spectra for plants and soil for each image were obtained by averaging the spectra of the 50 respective training pixels from that image. Alternatively, computerized methods such as the pixel purity index and the n -dimensional visualizer in ENVI (Research Systems, Inc., Boulder, Colorado, USA) can be used to identify purest pixels for the endmembers. However, these automatic methods are not always reliable. For example, weed plants can be mixed with crop plants and atypical soil surface areas with too dark or too bright colors can be mis-identified as typical soil. Since there were only two endmembers in this particular application, the simple pixel selection approach was used. Although only one reference spectrum is needed to calculate spectral distance and spectral angle, both the plant and soil spectra were used. Thus two spectral distance images, two spectral angle images and two unconstrained abundance images were calculated from all 102 bands for each field based on the two reference spectra using ENVI.

Statistical analysis

For correlation analysis, the 64 images based on the eight VIs and the selected wavelengths and the six images based on the three hyperspectral measures (spectral distance, spectral angle and abundance) for each field were aggregated by a factor of 9 to match the 9-m yield data resolution. The digital value for each output cell was the mean of the 81 input cells that the 9 m \times 9 m output cell encompassed. Correlation coefficients (r) between yield and each of the 70 spectral indices were calculated. Linear regression equations between yield and selected spectral indices were also determined. SAS software (SAS Institute Inc., Cary, North Carolina, USA) was used for statistical analysis.

Results and discussion

Table 2 gives the correlation coefficients between grain yield and the eight narrow band VIs for the two fields. The center wavelengths used to calculate SRI, NDVI, RNV, SAVI and MSAVI were 800 nm for the NIR band and 670 nm for the red band. The r values ranged from 0.74 to 0.78 for field 1 and from 0.83 to 0.85 for field 2. The three three-band VIs (MCARI1, TVI and MTVI2) were not superior to the two-band VIs for yield estimation.

Table 3 gives the correlation coefficients between grain yield and NDVI based on the three NIR bands (810, 825, 840 nm) and the eight visible bands for the two fields. For any visible center wavelength, the r values were essentially the same among the three NIR center wavelengths, indicating that any of the NIR center wavelengths can be used. However, the r values ranged from 0.74 to 0.80 for field 1 and from 0.82 to 0.85 for field 2 among the eight visible center wavelengths, indicating that choices of visible bands affect the r values. Two green bands (525 and 550 nm) had the highest r value (0.80) for field 1 and the red-edge band (720 nm) had the best r value (0.85) for field 2.

Table 4 summarizes the correlation coefficients between yield and the five two-band VIs based on the 825 nm NIR wavelength and the eight visible wavelengths for the two fields. Among the 40 VIs, the r values varied from 0.73 to 0.80 for field 1 and 0.82 to 0.86 for field 2. RDVI, SAVI and MSAVI appeared to produce more consistent r values than SRI or NDVI among the eight visible bands. For example, MSAVI provided similar r values of 0.77–0.79 for field 1 and 0.85–0.86 for field 2 among the eight visible bands. The commonly used NIR and red combinations were not the best for the two-band VIs for estimating crop yield, whereas the green bands tended to do better than the red bands, especially for SRI and NDVI. This is partly due to the fact that NDVI tends to saturate when crop canopy reaches its maximum cover, while green NDVI may have an advantage over NDVI at the particular crop growth stages. Although the green NDVI was proposed by Gitelson et al. (1996) for chlorophyll concentration estimation, it has also been used for yield estimation. The result from this study agrees with the findings of Thenkabail et al. (2000) and Yang and Everitt (2002). The red-edge band (720 nm) provided higher r values than the red bands for field 1 and the highest r values for field 2. However, the reflectance

Table 2 Correlation coefficients (r) between grain yield and eight narrow band vegetation indices (VIs) derived from 102-band hyperspectral images for two grain sorghum fields

Vegetation index ^a	Field 1	Field 2
SRI	0.74 ^b	0.83
NDVI	0.75	0.83
RNV	0.77	0.85
SAVI	0.77	0.85
MSAVI	0.78	0.85
MCARI1	0.75	0.85
TVI	0.75	0.84
MTVI2	0.76	0.85

^a The eight VIs are defined in Table 1. The center wavelengths used to calculate SRI, NDVI, RNV, SAVI and MSAVI were 800 nm for the NIR band and 670 nm for the red band

^b All the r values were significant at the 0.0001 level. The number of samples was 2 265 for field 1 and 1 658 for field 2

Table 3 Correlation coefficients (*r*) between grain yield and narrow band NDVI based on three NIR bands and eight visible bands derived from 102-band hyperspectral images for two grain sorghum fields

Visible band center (nm)	NIR band center (Field 1)			NIR band center (Field 2)		
	810 nm	825 nm	840 nm	810 nm	825 nm	840 nm
495 ^a	0.79 ^b	0.79	0.79	0.82	0.82	0.82
525	0.80	0.80	0.80	0.82	0.82	0.82
550	0.80	0.80	0.80	0.83	0.83	0.83
568	0.79	0.78	0.79	0.83	0.83	0.83
668	0.75	0.75	0.75	0.83	0.83	0.83
682	0.74	0.74	0.74	0.83	0.83	0.83
696	0.75	0.75	0.75	0.83	0.83	0.83
720	0.79	0.78	0.78	0.85	0.85	0.85

^a The NDVI-type indices were calculated with three NIR bands and eight visible bands. The red band in NDVI defined in Table 1 was replaced by the eight visible bands

^b All the *r* values were significant at the 0.0001 level. The number of samples was 2 265 for field 1 and 1 658 for field 2

Table 4 Correlation coefficients (*r*) between grain yield and five narrow band vegetation indices (VIs) derived from 102-band hyperspectral images for two grain sorghum fields

Visible band center (nm)	Field 1					Field 2				
	SRI ^a	NDVI	RDVI	SAVI	MSAVI	SRI	NDVI	RDVI	SAVI	MSAVI
495	0.78 ^b	0.79	0.78	0.77	0.78	0.82	0.82	0.84	0.84	0.85
525	0.80	0.80	0.79	0.78	0.78	0.83	0.82	0.85	0.85	0.85
550	0.80	0.80	0.79	0.79	0.79	0.85	0.83	0.85	0.85	0.85
568	0.78	0.78	0.79	0.79	0.79	0.85	0.83	0.85	0.85	0.85
668	0.74	0.75	0.77	0.77	0.77	0.83	0.83	0.85	0.85	0.85
682	0.73	0.74	0.76	0.76	0.77	0.83	0.83	0.85	0.85	0.85
696	0.73	0.75	0.77	0.77	0.77	0.84	0.83	0.85	0.85	0.85
720	0.78	0.78	0.79	0.79	0.79	0.86	0.85	0.86	0.86	0.86

^a The narrow band indices were calculated with one NIR band (825 nm) and eight visible bands. The red band in the five VIs defined in Table 1 was replaced by the eight visible bands

^b All the *r* values were significant at the 0.0001 level. The number of samples was 2 265 for field 1 and 1 658 for field 2

around this wavelength can change significantly with wavelength due to the steep slope in the red-edge portion of the spectrum, so the *r* values may not be stable. Therefore, MSAVI based on one NIR band (e.g., 825 nm) and one green band (e.g., 550 nm) appears to be one of the best choices. As shown in Table 2, a MSAVI image based on the 800-nm wavelength and the 670-nm wavelength gave equal or similar results for yield estimation.

Table 5 gives the correlation coefficients between grain yield and the three hyperspectral measures (spectral distance, spectral angle and fractional abundance) derived from the 102-band hyperspectral image based on the plant reference spectrum and the soil reference spectrum for each field. Yield was negatively related to spectral distance and spectral angle and positively related to plant abundance based on the plant reference

Table 5 Correlation coefficients (r) between grain yield and three hyperspectral measures (spectral distance, spectral angle and fractional abundance) derived from 102-band hyperspectral images based on plant and soil reference spectra for two grain sorghum fields

Hyperspectral measure	Field 1		Field 2	
	Plant-based	Soil-based	Plant-based	Soil-based
Spectral distance ^a	-0.76 ^b	0.75	-0.85	0.84
Spectral angle	-0.77	0.77	-0.83	0.84
Abundance	0.78	-0.75	0.85	-0.82

^a A pure plant spectrum and a pure soil spectrum extracted from each image were used to calculate spectral distance, spectral angle and fractional abundance. There were one plant abundance image and one soil abundance image for each field

^b All the r values were significant at the 0.0001 level. The number of samples was 2 265 for field 1 and 1 658 for field 2

spectra. In contrast, yield was positively related to spectral distance and spectral angle and negatively related to soil abundance based on the soil reference spectra. The magnitude of the r values ranged from 0.75 to 0.78 for field 1 and from 0.82 to 0.85 for field 2. These r values for the hyperspectral measures were within the ranges of the r values for the two-band and three-band VIs examined (0.73–0.79 for field 1 and 0.82–0.86 for field 2). This finding indicates that the two-band and three-band VIs can be as effective as the all-band hyperspectral measures for yield estimation.

Yang et al. (2008) evaluated 10 different reference spectra for grain sorghum plants, soil, roads and water extracted from hyperspectral images and from ground measurements for calculating spectral angle images. They found that spectral angle images based on reference spectra derived from bare soil, highway surface or water provided similar or slightly better r values with grain yield than those derived from plants. When the soil spectra were used as the reference spectra for generating spectral angle images in this study, the r values with yield were 0.77 for field 1 and 0.84 for field 2, compared with -0.77 and -0.83 for the respective fields based on the plant spectra (Table 5). Similarly, when the same soil spectra were used as the reference spectra for generating spectral distance images in this study, the r values with yield were 0.75 for field 1 and 0.84 for field 2, which are similar to the r values of -0.76 and -0.85 for the respective fields based on the plant spectra (Table 5).

Yang et al. (2007) examined how variations in endmember spectra affect plant abundance and its correlations with yield using 15 very different plant and soil spectrum pairs. Although the selection of the plant and soil endmember spectra affected the magnitude of the plant abundance values, the correlation coefficients between yield and unconstrained plant abundance were only minimally affected. In their study, the r values varied from 0.62 to 0.64 for one field and 0.79 to 0.81 for a second field among the 15 plant and soil spectrum pairs. In this study, the r values between yield and unconstrained plant abundance were 0.78 for field 1 and 0.85 for field 2. Although the r values for soil abundance (-0.75 for field 1 and -0.82 for field 2) were similar to those for plant abundance, it is more meaningful to use a plant abundance image as a relative yield map because plant abundance is a direct indicator of plant canopy cover.

Figures 1 and 2 show the scatter plots and regression lines of grain yield with spectral distance, spectral angle, plant abundance and MSAVI derived from the 102-band airborne hyperspectral image for fields 1 and 2, respectively. Spectral distance and spectral angle

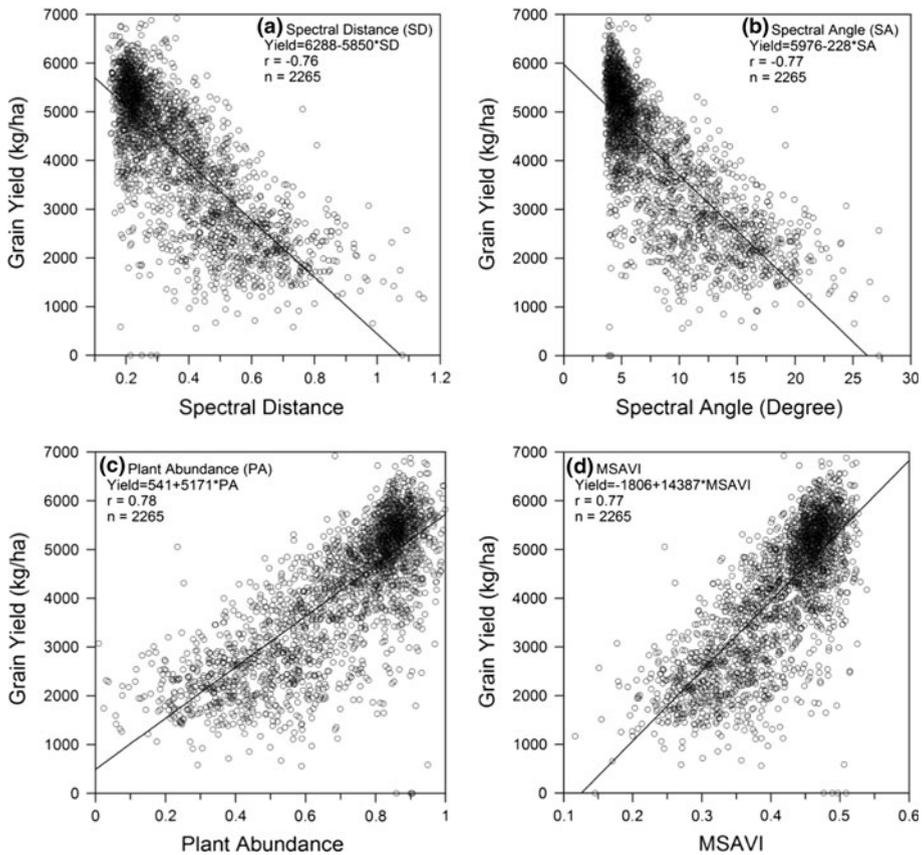


Fig. 1 Scatter plots and regression lines of grain yield with **a** spectral distance, **b** spectral angle, **c** plant abundance and **d** MSAVI (NIR = 825 nm and red = 668) derived from a 102-band airborne hyperspectral image for a grain sorghum field (field 1). Spectral distance and spectral angle were generated using a plant reference spectrum, while plant abundance was generated based on a pair of plant and soil reference spectra

were generated using the healthy plant reference spectra, while plant abundance was generated based on the healthy plant and soil reference spectra. Grain yield generally decreased with plant spectral distance and spectral angle and increased with plant abundance and MSAVI as expected. Although clear linear trends existed between yield and each of the four hyperspectral indices, there was also large variability in yield for any given value of each hyperspectral index. This is understandable because plant canopies with the same spectral index value may not always have the same yield value. Nevertheless, the general linear correlations with yield provide the basis for estimating relative yield variation from hyperspectral imagery using any one of the hyperspectral measures. Table 6 gives the correlation coefficients among the four hyperspectral indices for the two grain sorghum fields. The magnitude of the r values ranged from 0.96 to 0.99 for field 1 and from 0.98 to 1.00 for field 2. Clearly, these four hyperspectral indices were highly interrelated.

Figures 3 and 4 present the maps for the four hyperspectral indices as compared with the actual yield maps for fields 1 and 2, respectively. The spectral values for each map were arbitrarily grouped into 10–12 classes and different colors were assigned to the

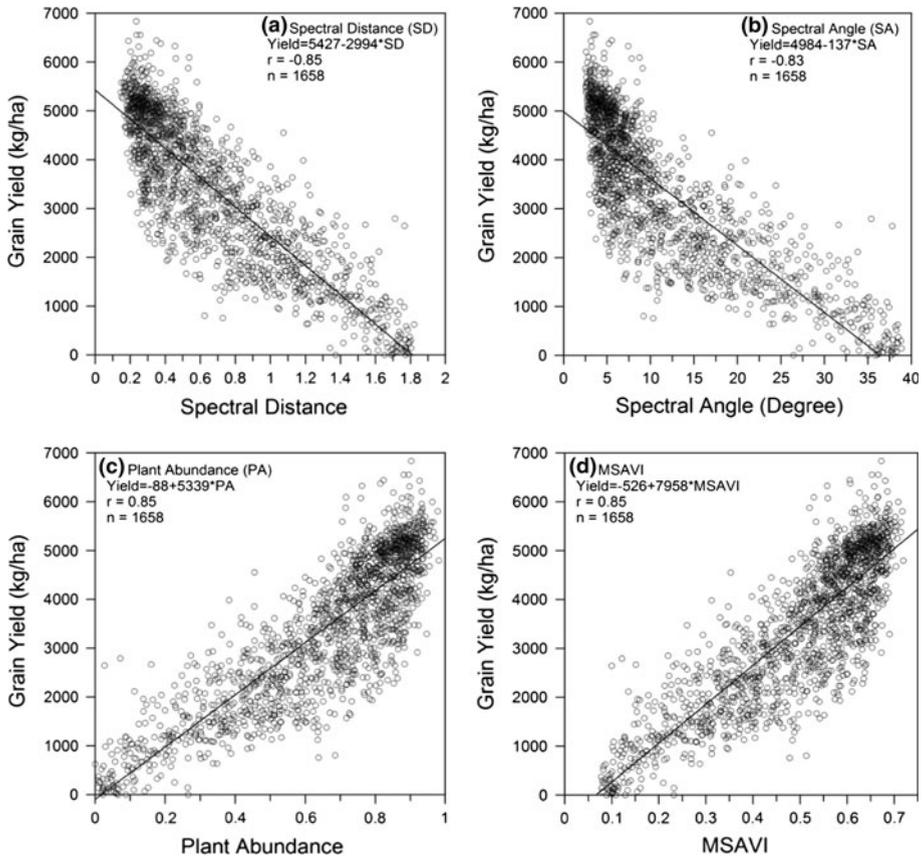


Fig. 2 Scatter plots and regression lines of grain yield with **a** spectral distance, **b** spectral angle, **c** plant abundance and **d** MSAVI (NIR = 825 nm and red = 668) derived from a 102-band airborne hyperspectral image based on a reference plant spectrum for a grain sorghum field (field 2)

Table 6 Correlation coefficients (*r*) among four hyperspectral indices (spectral distance, spectral angle, plant abundance and MSAVI) derived from 102-band hyperspectral images for two grain sorghum fields

Hyperspectral index	Field 1			Field 2		
	SD ^a	SA	PA	SD	SA	PA
SA	0.96 ^b			0.98		
PA	-0.98	-0.98		-1.00	-0.99	
MSAVI	-0.96	-0.97	0.99	-1.00	-0.98	1.00

^a *SD* spectral distance, *SA* spectral angle, *PA* plant abundance, *MSAVI* modified soil-adjusted vegetation index

^b All the *r* values were significant at the 0.0001 level. The number of samples was 2 265 for field 1 and 1 658 for field 2

classes with red showing low yielding areas and green showing high yielding areas. Although the four spectral maps reveal some differences, they generally show similar patterns of plant growth variability, which resemble the patterns on the actual yield maps.

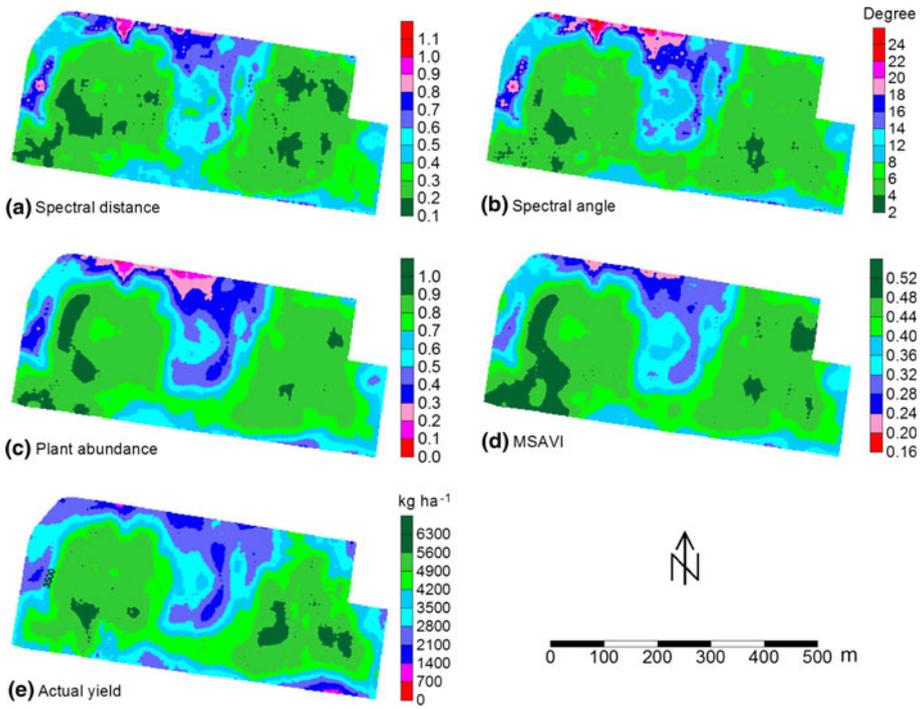


Fig. 3 Maps of **a** spectral distance, **b** spectral angle, **c** plant abundance, **d** MSAVI and **e** actual yield for a grain sorghum field (field 1). The four hyperspectral indices were derived from a 102-band airborne hyperspectral image of the field

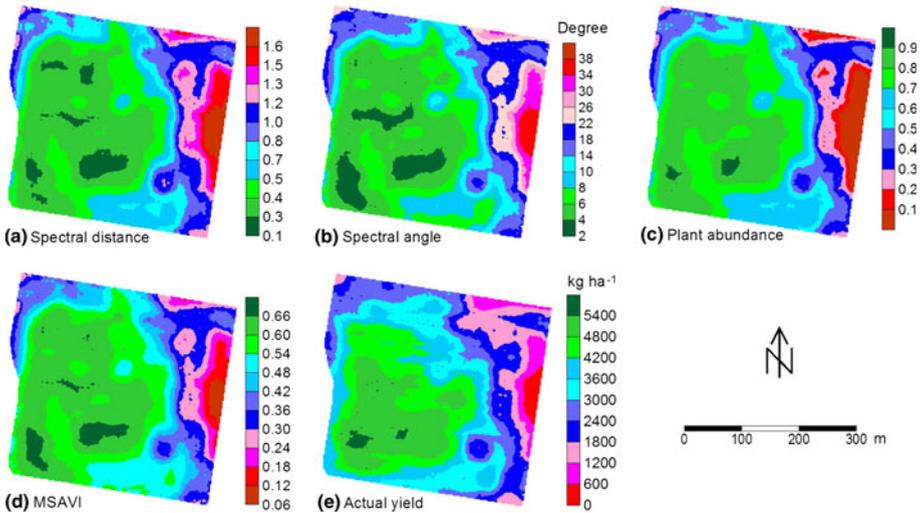


Fig. 4 Maps of **a** spectral distance, **b** spectral angle, **c** plant abundance, **d** MSAVI and **e** actual yield for a grain sorghum field (field 2). The four hyperspectral indices **a–d** were derived from a 102-band airborne hyperspectral image of the field

Conclusions

This study examined five two-band VIs (SRI, NDVI, RDVI, SAVI and MSAVI) and three three-band VIs (MCARI1, TVI and MTVI2) as well as three all-band hyperspectral measures (spectral distance, spectral angle and fractional abundance) for yield estimation. MSAVI produced more consistent and generally higher r values than the other VIs based on one NIR and eight visible wavelengths. The commonly used NIR and red combinations were not always the best for the two-band VIs and the green band used as an alternative to the red band in the two-band VIs tended to do as well as or better than the red band for yield estimation. In addition, the red-edge band (720 nm) as an alternative to the red band performed better as well. These findings indicate that VIs based on a NIR and green band combination or a NIR and red-edge band combination have the potential to do better than the traditional red and NIR combination for yield estimation.

The three all-band hyperspectral measures provided comparable results with the best VIs. In practice, the three all-band hyperspectral measures do not need band selection, but they require one or two reference spectra for the calculation. Compared with spectral distance and spectral angle, plant abundance derived from a hyperspectral image using spectral unmixing provides a direct measure of crop canopy cover. This feature makes plant abundance even more relevant than traditional VIs for crop cover and yield estimation. However, the traditional VIs only need two or three bands, which can also be obtained from less expensive multispectral imagery. Therefore, if crop fractional cover is not necessary, one can create a VI (e.g., MSAVI) image based on one NIR band (e.g., 800 or 825 nm) and one visible band (e.g., 550 or 670 nm) from multispectral or hyperspectral data to characterize crop yield variation. Otherwise, a plant abundance image can be derived from a hyperspectral image using linear spectral unmixing to estimate crop fractional cover and relative yield. More experiments are needed to validate these recommendations for other agricultural crops over diverse growing conditions and environments.

Acknowledgments The author thanks Rene Davis and Fred Gomez of USDA-ARS at Weslaco, TX for acquiring the imagery for this study and Jim Forward of USDA-ARS at Weslaco, TX for assistance in image rectification and calibration. Thanks are also extended to Rio Farms, Inc. at Monte Alto, TX for use of its fields and harvest equipment.

References

- Adams, J. B., Smith, M. O., & Johnson, P. E. (1986). Spectral mixture modeling: A new analysis of rock and soil types at the Viking Lander 1 site. *Journal of Geophysical Research*, *91*, 8098–8112.
- Baret, F., & Guyot, G. (1991). Potentials and limits of vegetation indices for LAI and APAR assessment. *Remote Sensing of Environment*, *35*, 61–173.
- Bateson, A., & Curtiss, B. (1996). A method for manual endmember and spectral unmixing selection. *Remote Sensing of Environment*, *55*, 229–243.
- Broge, N. H., & Leblanc, E. (2000). Comparing prediction power and stability of broadband and hyperspectral vegetation indices for estimation of green leaf area index and canopy chlorophyll density. *Remote Sensing of Environment*, *76*, 156–172.
- Campbell, J. B. (2002). *Introduction to remote sensing* (3rd ed.). New York: The Guilford Press.
- Chen, J. (1996). Evaluation of vegetation indices and modified simple ratio for boreal applications. *Canadian Journal of Remote Sensing*, *22*, 229–242.
- Daughtry, C. S. T., Walthall, C. L., Kim, M. S., Brown de Colstoun, E., & McMurtrey, J. E., I. I. I. (2000). Estimating corn leaf chlorophyll concentration from leaf and canopy reflectance. *Remote Sensing of Environment*, *74*, 229–239.

- Dennison, P. E., Halligan, K. Q., & Roberts, D. A. (2004). A comparison of error metrics and constraints for multiple endmember spectral mixture analysis and spectral angle mapper. *Remote Sensing of Environment*, 93, 359–367.
- Franke, J., Roberts, D. A., Halligan, K., & Menz, G. (2009). Hierarchical Multiple Endmember Spectral Mixture Analysis (MESMA) of hyperspectral imagery for urban environments. *Remote Sensing of Environment*, 113, 1712–1723.
- Garcia-Haro, F. J., Gilabert, M. A., & Melia, J. (1996). Linear spectral mixture modelling to estimate vegetation amount from optical spectral data. *International Journal of Remote Sensing*, 17(17), 3373–3400.
- Gitelson, A. A., Kaufman, Y., & Merzlyak, M. N. (1996). Use of green channel in remote sensing of global vegetation from EOS-MODIS. *Remote Sensing of Environment*, 58, 289–298.
- Haboudane, D., Miller, J. R., Pattey, E., Zarco-Tejada, P. J., & Strachan, I. (2004). Hyperspectral vegetation indices and novel algorithms for predicting green LAI of crop canopies: Modeling and validation in the context of precision agriculture. *Remote Sensing of Environment*, 90, 337–352.
- Huete, A. R. (1988). A soil adjusted vegetation index (SAVI). *Remote Sensing of Environment*, 25, 295–309.
- Jordan, C. F. (1969). Derivation of leaf area index from quality of light on the forest floor. *Ecology*, 50(4), 663–666.
- Kruse, F. A., Lefkoff, A. B., Boardman, J. W., Heidebrecht, K. B., Shapiro, A. T., Barloon, J. P., et al. (1993). The spectral image processing system (SIPS): Interactive visualization and analysis of imaging spectrometer data. *Remote Sensing of Environment*, 44, 145–163.
- Moran, M. S., Maas, S. J., & Pinter, P. J., Jr. (1995). Combining remote sensing and modeling for estimating surface evaporation and biomass production. *Remote Sensing of Environment*, 12, 335–353.
- Qi, J., Chehbouni, A., Huete, A. R., Keer, Y. H., & Sorooshian, S. (1994). A modified soil vegetation adjusted index. *Remote Sensing of Environment*, 48, 119–126.
- Ray, S. S., Das, G., Singh, J. P., & Panigrahy, S. (2006). Evaluation of hyperspectral indices for LAI estimation and discrimination of potato crop under different irrigation treatments. *International Journal of Remote Sensing*, 27(24), 5373–5387.
- Rougean, J.-L., & Breon, F. M. (1995). Estimating PAR absorbed by vegetation from bidirectional reflectance measurements. *Remote Sensing of Environment*, 51, 375–384.
- Rouse, J. W., Haas, R. H., Shell, J. A., Deering, D. W. (1973). Monitoring vegetation systems in the Great Plains with ERTS. In *Proceedings of 3rd ERTS symposium* (vol. 1, pp. 309–317). NASA SP-351. Washington, D.C.: U.S. Government Printing Office.
- Thenkabail, P. S., Smith, R. B., & Pauw, E. D. (2000). Hyperspectral vegetation indices and their relationships with agricultural crop characteristics. *Remote Sensing of Environment*, 71, 158–182.
- Wiegand, C. L., Richardson, A. J., Escobar, D. E., & Gerbermann, A. H. (1991). Vegetation indices in crop assessments. *Remote Sensing of Environment*, 35, 105–119.
- Wu, C., Han, X., Niu, Z., & Dong, J. (2010). An evaluation of EO-1 hyperspectral Hyperion data for chlorophyll content and leaf area index estimation. *International Journal of Remote Sensing*, 31(4), 1079–1086.
- Yang, C., & Anderson, G. L. (1999). Airborne videography to identify spatial plant growth variability for grain sorghum. *Precision Agriculture*, 1(1), 67–79.
- Yang, C., & Everitt, J. H. (2002). Relationships between yield monitor data and airborne multitemporal digital imagery for grain sorghum. *Precision Agriculture*, 3(4), 373–388.
- Yang, C., Everitt, J. H., & Bradford, J. M. (2004). Airborne hyperspectral imagery and yield monitor data for estimating grain sorghum yield variability. *Transactions of the ASAE*, 47(3), 915–924.
- Yang, C., Everitt, J. H., & Bradford, J. M. (2007). Airborne hyperspectral imagery and linear spectral unmixing for mapping variation in crop yield. *Precision Agriculture*, 8(6), 279–296.
- Yang, C., Everitt, J. H., & Bradford, J. M. (2008). Yield estimation from hyperspectral imagery using spectral angle mapper (SAM). *Transactions of the ASABE*, 51(2), 729–737.
- Yang, C., Everitt, J. H., Davis, M. R., & Mao, C. (2003). A CCD camera-based hyperspectral imaging system for stationary and airborne applications. *Geocarto International*, 18(2), 71–80.
- Zarco-Tejada, P. J., Ustin, S. L., & Whiting, M. L. (2005). Temporal and spatial relationships between within-field yield variability in cotton and high-spatial hyperspectral remote sensing imagery. *Agronomy Journal*, 97, 641–653.