Industrial Crops and Products 34 (2011) 1150–1159

Contents lists available at ScienceDirect

Industrial Crops and Products

journal homepage: www.elsevier.com/locate/indcrop

Color image segmentation approach to monitor flowering in lesquerella

K.R. Thorp[a,*], D.A. Dierig[b]

[a] USDA-ARS, U.S. Arid Land Agricultural Research Center, 21881 N Cardon Ln, Maricopa, AZ 85138, United States
[b] USDA-ARS, National Center for Genetic Resources Preservation, 1111 S Mason St, Fort Collins, CO 80521, United States

ARTICLE INFO

Article history:
Received 18 November 2010
Received in revised form 1 April 2011
Accepted 6 April 2011
Available online 14 May 2011

Keywords:
Agriculture
Anthesis
Arizona
Camera
Connected components
Digital
Flower
Hue
Intensity
Lesquerella
Management
Monte Carlo
Oilseed
Phenotyping
Saturation

ABSTRACT

Lesquerella (*Lesquerella fendleri* (Gray) Wats.) seed oil has been proposed as a petroleum alternative in the production of many industrial products, but several crop management and breeding challenges must be addressed before the crop will be grown commercially. Lesquerella canopies characteristically exhibit quite prominent and vibrant yellow flowers at anthesis, and remote detection of lesquerella flowering patterns can provide useful crop development information to aid management and breeding decisions. In the present study, we used a consumer-grade digital camera to collect images 2 m above lesquerella canopies throughout two growing seasons. Biomass samples within 0.125 m^2 areas were also regularly collected and processed to obtain flower numbers. Image processing algorithms were developed to extract information on lesquerella flower features from the images. Key features of the image processing approach included an image transformation to the hue, saturation, and intensity (HSI) color space and a Monte Carlo approach to address uncertainty in HSI parameters used for image segmentation. Flower numbers were estimated from image-based flower cover percentage with root mean squared errors that ranged from 159 to 194 flowers, which was better than the reported results for other studies with a similar objective. Attempts to resolve individual flowers were less successful due to the complexity of the flowering patterns within the image scenes. Digital imaging offers an inexpensive and quite practical means for remote monitoring of flowering patterns in lesquerella canopies.

Published by Elsevier B.V.

1. Introduction

Onset of flowering is a critical development stage for most agricultural crops. Flowering marks the transition from vegetative to reproductive development, at which time the plants begin to set the reproductive structures that may lead to mature grains or fruits. These reproductive processes are particularly sensitive to plant stresses from temperature extremes (Ferris et al., 1998; Ohnishi et al., 2010) and nutrient and water deficits (Nielsen and Nelson, 1998; Moser et al., 2006). Thus, knowledge of flowering onset is useful information in the development of optimized crop management plans that aim to correct nutrient deficiencies or adjust water management strategies during this critical time. The rate of crop development to anthesis is dependent on plant genetics and several environmental factors, such as photoperiod, temperature, and plant stress levels (Bonaparte, 1975; Johansen et al., 1985). Methods to rapidly quantify the time to flowering or the flower number are of interest to breeders who want to select plants with specific flowering characteristics. Such information is also important for current high-throughout phenotyping efforts that aim to better understand the link between plant genetics and phenotypic expression of traits, such as flowering (Montes et al., 2007).

Techniques for remote detection of flowering have been demonstrated in several common crops, particularly those with prominent and spectrally distinguishable flowers. For example, Mogensen et al. (1996) showed that spectral reflectance indices were sensitive to flowering time in oilseed rape (*Brassica napus* L.), which displays vibrant yellow flowers at anthesis. The results of Vina et al. (2004) demonstrated the use of a spectral index computed from several visible light bands to detect the emergence of tassels in maize (*Zea mays* L.), and Pimstein et al. (2009) developed a spectral index to monitor emergence of heads in wheat (*Triticium aestivum* L.). In another study by Kaleita et al. (2006), several hyperspectral data processing techniques were used to detect onset of tasseling and pollen shed in maize. They concluded that the reflectance of maize canopies was spectrally unique during the critical tasseling stage, at which time kernel set is established and grain fill is initiated.

Lesquerella (*Lesquerella fendleri* (Gray) Wats.) is an alternative oilseed crop native to the southwestern United States and northern Mexico. Hydroxylated fatty acids derived from lesquerella seed are

* Corresponding author.
E-mail address: kelly.thorp@ars.usda.gov (K.R. Thorp).

0926-6690/$ – see front matter. Published by Elsevier B.V.
doi:10.1016/j.indcrop.2011.04.002

being developed as a biorenewable diesel fuel additive (Geller and Goodrum, 2004; Moser et al., 2008) and as a petroleum substitute in the production of many industrial products, such as greases, lubricants, cosmetics, paints, inks, and coatings (Dierig et al., 1992). The indeterminate and vibrant yellow flowering patterns of lesquerella make the crop spectrally intriguing, and application of remote techniques to detect and monitor flowering patterns in this new crop should be quite feasible and practical. Recently, Thorp et al. (2011) measured the spectral reflectance of lesquerella canopies with a field spectroradiometer and developed a partial least squares regression model to estimate flower counts. The model estimated flower counts with root mean squared errors from 251 to 304 flowers. In an alternative approach, Adamsen et al. (2000) used a digital camera to photograph lesquerella canopies and developed an automatic image segmentation algorithm to estimate flower counts. Image segmentation is the process of defining regions of similar characteristics within a digital image. By regressing manual flower counts against the number of image pixels segmented as flowers, they were able to estimate flower counts with a root mean squared error of 235 flowers. This digital imaging approach was subsequently used to assess the response of lesquerella flowering to nitrogen fertilization and seeding rate (Adamsen et al., 2003) and to asses the response of oilseed rape flowering to planting date (Adamsen and Coffelt, 2005).

Application of digital cameras to monitor lesquerella flowering offers several advantages over more traditional remote sensing approaches with spectroradiometers. Digital images allow a combined spatial and spectral approach for analyzing crop scenes, whereas radiometers provide only a point-based spectral measurement. Inexpensive digital cameras are also readily available on the consumer market, whereas commercial remote sensing instruments are usually quite specialized and more expensive. Given these advantages, the objective of this study was to develop a novel approach for estimating flower counts from ground-based images of lesquerella canopies. The new approach refines the work of Adamsen et al. (2000) by incorporating an image transformation to the hue, saturation, and intensity (HSI) color space and by using a Monte Carlo algorithm to address uncertainty in defining the color space parameters for flower segmentation.

2. Materials and methods

2.1. Field experiments

Lesquerella was grown at the University of Arizona's Maricopa Agricultural Center (MAC) near Maricopa, Arizona (33.067547° N, 111.97146° W) over the winters of 2007–2008 and 2008–2009. The soil type at the site was a Casa Grande sandy loam, classified as fine-loamy, mixed, hyperthermic, Typic Natrargids. In both growing seasons, the field layout consisted of nine experimental plots, each 20 m × 180 m and hydrologically isolated with border dikes. Three planting date treatments were replicated three times over the nine plots (Table 1). In the 2007–2008 experiment, the first and second planting dates were September 28 and February 15, respectively. The third treatment was planted in March, but poor stand density prevented any useful data from being collected from this treatment. In the 2008–2009 experiment, planting dates were October 6, January 8, and February 6. All plots were broadcast planted at a rate of 12 kg ha^{-1} using a Brillion planter with a roller ring. Plots were flood irrigated by siphoning water from a canal along the southern edge of the field. After crop emergence, multiple locations within each plot were randomly selected and flagged for biomass sampling and recurrent collection of digital images. The marked areas were each 0.125 m^2. In 2007–2008, twenty sample areas were marked for the first planting while twelve areas were marked for the second planting. In 2008–2009, the total number of marked areas was twelve, eight, and eight for the first, second, and third plantings, respectively (Table 1).

Table 1
Summary of the 2007–2008 and 2008–2009 lesquerella experiments.

	Planting 1	Planting 2	Planting 3
2007–2008 Experiment			
Planting date	9/28/2007	2/15/2008	3/10/2008
Replications	3	3	3
Sample areas per plot	20	12	0
Biomass samples per plot	16	7	0
Image collection dates	20	10	0
2008–2009 Experiment			
Planting date	10/6/2008	1/8/2009	2/6/2009
Replications	3	3	3
Sample areas per plot	12	8	8
Biomass samples per plot	8	4	3
Image collection dates	10	6	5

2.2. Field measurements

Ground-based digital images were collected at each of the 0.125 m^2 sampling locations from emergence until biomass was destructively sampled at that location. Square frames of white PVC tubing were constructed and laid within the canopy to delineate each 0.125 m^2 sample area prior to image collection. Additional images (approximately 24 per plot) were collected while walking along a 180 m linear transect on the western edge of each plot. Images were typically obtained in the mid-morning hours. Image collection occurred on a weekly basis during the 2007–2008 experiment and at a two-week interval during the 2008–2009 experiment (Table 1).

A digital camera (EOS Digital Rebel XT, Canon U.S.A. Inc., Lake Success, New York) was used to collect the images of the lesquerella canopy. The camera was a high-end consumer instrument equipped with an 8.2 megapixel complimentary metal–oxide semiconductor (CMOS) detector and a red, green, and blue (RGB) color filtering system. Sixteen-bit digital images with resolution of 3456 × 2304 pixels were collected and saved to the onboard CompactFlash card in the RAW image format.

An L-shaped metal pole was fabricated from 2.54 cm square tubing to suspend the camera at nadir view angle over the crop canopy. Design of the pole allowed the camera height to be adjusted between 200 and 350 cm from the soil surface. Throughout each lesquerella growing season, the camera height was adjusted to maintain a distance of approximately 200 cm between the camera lens and the crop canopy. To collect images with the mounted camera, the remote control device provided by the manufacturer was used to trigger the shutter.

Biomass was destructively sampled at one of the 0.125 m^2 sampling locations in each plot on a weekly basis during the 2007–2008 experiment and every two weeks during the 2008–2009 experiment. In 2008, the first planting date treatment was sampled sixteen times per plot from January 15 to May 21, and the second planting date treatment was sampled seven times per plot from May 1 to June 19. In 2009, the first treatment was sampled eight times per plot from February 3 to May 12, and the second treatment was sampled four times per plot from April 30 to June 11. The third treatment in 2009 was sampled three times per plot from May 13 to June 12. Typically, we overestimated the number of sampling areas needed to document lesquerella growth and development over the growing season, and some of the predefined sample areas remained unsampled at crop maturity (Table 1). To collect the biomass samples, the square frames of white PVC tubing were again used to delineate each 0.125 m^2 sample area. Samples were typically collected in the early morning hours. Plant material

was then immediately processed in the laboratory to obtain estimates of several biophysical properties of the lesquerella canopy. Counts of lesquerella flowers within each 0.125 m^2 area was the key measurement used in the present study.

2.3. Image processing algorithms

After each image collection outing, proprietary software provided by the camera manufacturer was used to convert the images from the RAW format to the TIF format. This produced three-band images in the RGB color space with 16-bit color depth on each channel. Images were then transformed to the hue, saturation, and intensity (HSI) color space (Gonzalez and Woods, 1992). This transformation decoupled the intensity (or brightness) information from the color information in the RGB images, which allowed for thresholding on the more relevant color parameters of hue and saturation. Since others have found the HSI transformation to be useful for color image segmentation in the variable lighting conditions of outdoor agricultural scenes (Tang et al., 2000), we hypothesized that an HSI transformation could facilitate image segmentation of yellow flowers within lesquerella canopies. Further image processing procedures for extracting information on lesquerella flowers were conducted as shown in the flow diagram (Fig. 1).

2.3.1. HSI transformation

To convert between the RGB and HSI color spaces, RGB pixel values were first normalized to range [0, 1] by dividing by the bit depth of each channel. The normalized RGB data were then converted to HSI in the range [0, 1] using the following equations:

$$I = \frac{1}{3}(R+G+B) \tag{1}$$

$$S = 1 - \frac{3}{(R+G+B)}[\min(R,G,B)] \tag{2}$$

$$H = \arccos\left\{\frac{(1/2)[(R-G)+(R-B)]}{[(R-G)^2+(R-B)(G-B)]^{1/2}}\right\} \tag{3}$$

For cases where $(B/I) > (G/I)$, we computed $H = 2\pi - H$. Also, to normalize hue to the range [0, 1], we let $H = H/2\pi$. Transformations from RGB to HSI color space are ill-conditioned, because a singularity exists at $R = G = B = 0$ and hue is undefined along the intensity axis, $R = G = B$. We set $H = S = 0$ at their singularities. Further details regarding the RGB and HSI color spaces and transformations between them can be found in Gonzalez and Woods (1992) and Tang et al. (2000).

2.3.2. Image segmentation

Segmentation of lesquerella flowers was accomplished by thresholding the images in HSI color space. Boundary conditions were based on six parameters, including the maximum and minimum hue, maximum and minimum saturation, and maximum and minimum intensity. The image segmentation processing step produced a binary image with 1 indicating pixels that fell within the boundary criteria and 0 otherwise. Because the quality of image segmentation is highly dependent on the parameter values used to specify the boundary conditions (Tang et al., 2000), we implemented a Monte Carlo approach (discussed below) to reduce the subjectivity of boundary condition specification and improve the reliability of image segmentations.

2.3.3. Noise removal

As an optional step in our image processing algorithm, we included a 5×5 median filter convolution for reducing noise and smoothing the segmented binary image (Fig. 1). The specific type and size of filter was determined by experimenting with several

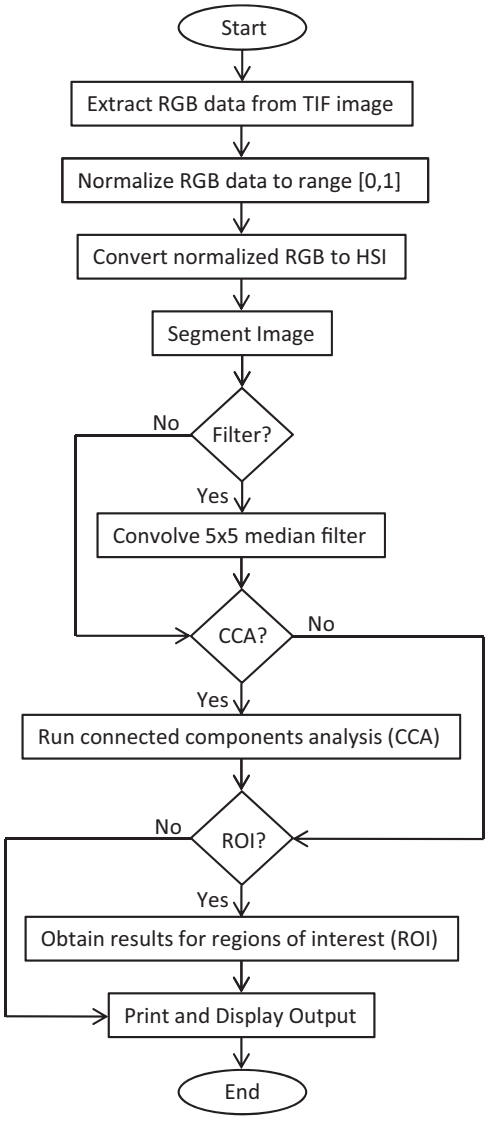

Fig. 1. Flow chart for the image processing procedure.

filtering schemes. Although not likely essential for our image analysis, we elected to include the median filter convolution for all image processing results reported herein. After this filtering step, we counted the number of pixels identified as lesquerella flowers, and this number served as the first image-based metric for estimation of lesquerella flower counts.

2.3.4. Connected components analysis

A second metric for estimation of flower count was obtained by performing a connected components analysis on the segmented, median-filtered, binary image. The purpose of this processing step was to identify the total number of disconnected features or 'blobs' in the segmented image. Gonzalez and Woods (1992) provide a basic methodology for this procedure.

2.4. Algorithm implementation

Image processing algorithms were developed using the Interactive Data Language (IDL) within the Environment for Visualizing Images (ENVI 4.5, ITT Visual Information Solutions, Boulder, CO, USA). An interactive color image segmentation tool with graphics user interface was created to facilitate user implementation of the image processing algorithms. The tool allows users to con-

duct supervised image segmentations, manage HSI segmentation parameters, and perform several batch-mode processes for automatic segmentation of multiple images. Several processing options can also be selected, including whether or not to perform median filter convolution, to conduct connected components analysis, or to compile segmentation results for user-defined regions of interest (ROIs) within the images. In the present study, native ENVI algorithms were used to manually draw ROIs corresponding to the image area delineated by the 0.125 m^2 PVC frame at biomass sampling locations. Further processing of these images focused only on the area within the image ROIs.

2.4.1. HSI parameter selection

To understand the HSI parameter ranges that were appropriate for segmenting lesquerella flowers, manual supervised segmentations were conducted on a subset of images using the interactive color image segmentation tool. Images were selected based on their nearness in time to the actual biomass sampling date at each sampling location. As a result, our manual segmentations were performed on the same images that were used to develop relationships between the image-based estimates of flower counts and actual flower counts. A total of 63 images were selected from those collected in the 2007–2008 growing season, and 42 were selected from those collected in 2008–2009. For each of these 105 images, the image segmentation tool was used to interactively select image pixels that represented lesquerella flowers, to perform full image segmentations based on the HSI ranges of selected pixels, to visually inspect the segmentation result, and to select additional flower pixels and reprocess until the overall segmentation of lesquerella flowers was deemed adequate by human inspection. This endeavor resulted in a set of 105 HSI parameter sets that defined lesquerella flowers in images collected over multiple plots on multiple dates throughout two growing seasons. All supervised segmentations were conducted by the same person to avoid bias in judgement of adequate segmentations.

2.4.2. Monte Carlo sampling

The set of 105 HSI parameter sets obtained through manual supervised segmentations was used to understand the uncertainty in the maximum and minimum hue, saturation, and intensity parameters required for adequate segmentation of lesquerella flowers. Calculation of the mean and standard deviation of these six parameter sets defined their normal probability distributions. All the images in the complete data set were then processed using a Monte Carlo sampling technique to vary the HSI parameters used for image segmentation. The HSI parameter values were iteratively and independently sampled from their respective normal distributions by implementing a random number generator to computationally randomize the selection. With each sampled parameter set, images were analyzed for total number of segmented flower pixels and total number of connected components. A total of 500 Monte Carlo parameter samples were used to process images collected over the 0.125 m^2 biomass sampling areas, and 250 Monte Carlo parameter samples were used to process images collected along the western transect of each plot. The number of iterations was selected using a Shapiro–Wilk normality test to determine the number of sampled parameters needed for normality at the 0.01 significance level. This demonstrated that the sampled parameters adequately represented the theoretical normal distributions from which they were drawn. This Monte Carlo sampling approach was used to process 974 images of biomass sampling areas and 1787 transect images from the 2007–2008 growing season and 356 images of biomass sampling areas and 1583 transect images from the 2008–2009 season. Results were averaged to obtain one value for flower pixel cover percentage and connected component number from each image. This approach provided image-based metrics for estimating lesquerella flower counts while considering the uncertainty in the HSI parameter values used for image segmentation.

2.5. Application

After processing all the photos, simple linear regression models were built to relate flower pixel number and connected component number to the flower count measurements from the biomass sample areas. Since the lesquerella canopy was sampled more frequently in 2007–2008, we used the information from this season to build the regression models. Data from the 2008–2009 season were then used to independently test the models.

Processing results from images collected along the 180 m transects were used to demonstrate temporal monitoring of lesquerella flowering patterns. The regression equations were applied to estimate flower count from the images collected at each location along the transect on each measurement date. Estimates were then averaged according to the measurement date and planting date treatment to generate time-series plots of flower number. These plots demonstrated the ability of the image processing procedure to temporally monitor lesquerella flowering patterns over each growing season.

3. Results and discussion

3.1. Visual example

The image processing result for an example image is given in Fig. 2, where Fig. 2a shows the original RGB image and Fig. 2b shows the image segmentation result. Pixels identified as flowers after HSI transformation and image segmentation are marked with a magenta color, and non-flower pixels are shown in their original RGB format. Within the ROI defined by the 0.125 m^2 PVC frame, flower cover was 35.5% for this example image. Fig. 2c demonstrates the image processing result after a 5×5 median filter convolution on the segmented binary image. This processing step tended to adjust a few scattered pixels in the soil or green vegetation parts of the image that were incorrectly segmented as flowers. It also tended to fill holes in the segmented flower features and produced a better representation of lesquerella flowers overall. A slight adjustment in flower cover percentage, from 35.5% to 34.5%, resulted from this image processing step. Fig. 2d demonstrates the result of the connected component analysis on the segmented, median-filtered binary image. Although 150 unique connected components were present within the image ROI, we are using only six primary colors to display them. By comparing Fig. 2a and d, it is evident that flowers must be well separated in space for connected component analysis to adequately delineate individual lesquerella flowers.

3.2. Image segmentation

The mean and standard deviation of the 105 HSI parameter sets obtained from manual segmentation of images is given in Table 2. Minimum and maximum hue were 0.12 and 0.18, respectively, corresponding as expected to yellow hue in the HSI color space. Hue parameters were also relatively stable with lower standard deviations than the other parameters. Maximum saturation was 1.0 for all the HSI parameter sets, indicating that a fully saturated yellow color was necessary for adequate segmentation of flowers in all the photos. Higher standard deviations for the remaining parameters, minimum saturation and maximum and minimum intensity, demonstrated higher uncertainty in defining their values for flower segmentations. Likely, the variability in these parameters is related to the variable outdoor lighting conditions in which the images

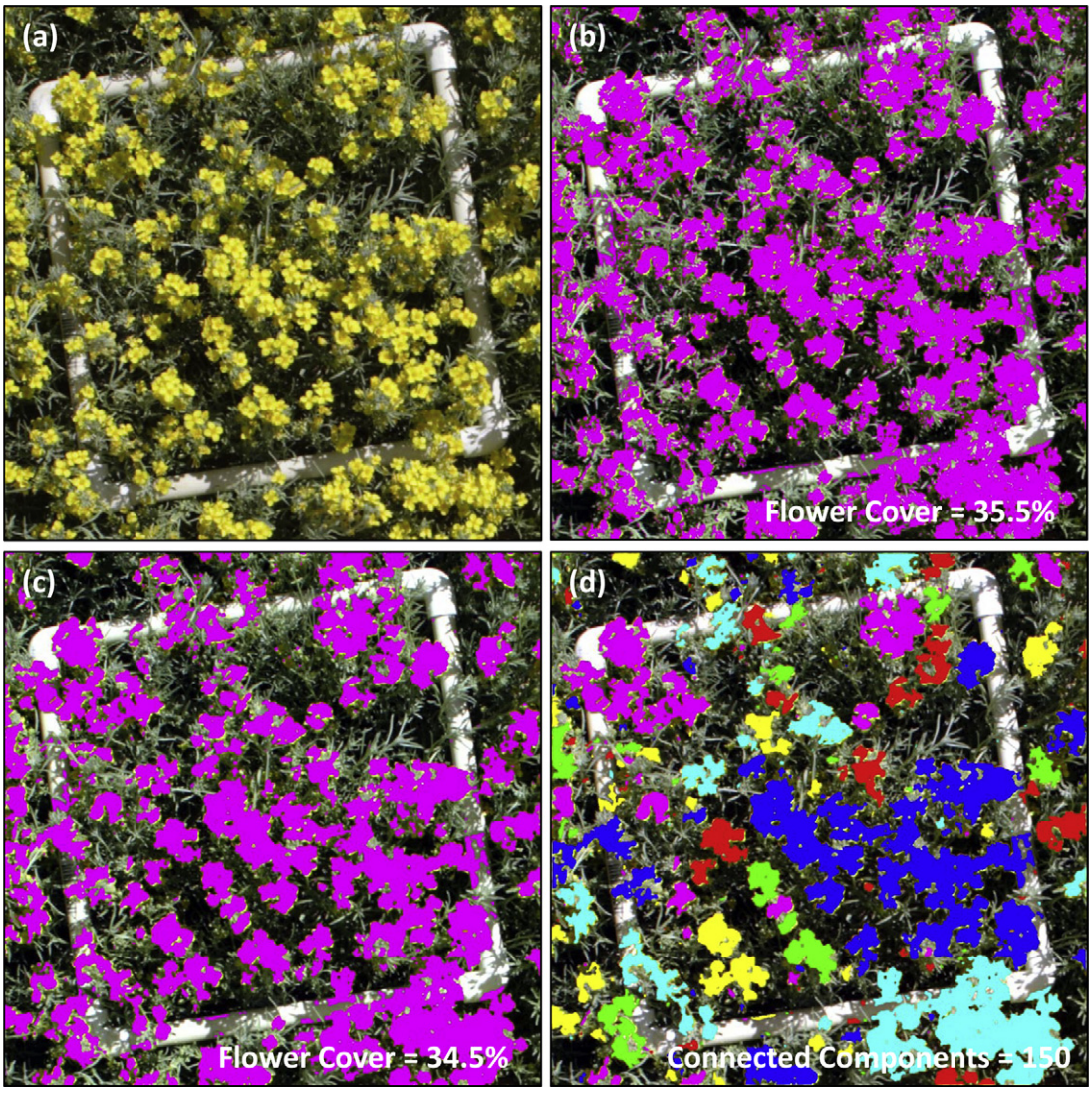

Fig. 2. Example image processing result: (a) original image, (b) segmented image, (c) median-filtered, segmented image, and (d) median-filtered, segmented image with unique connected components identified. Area of the PVC square was 0.125 m^2.

Table 2
Descriptive statistics of HSI parameters from manual image segmentations of lesquerella flowers. Coefficient of determination (r^2) is computed for each HSI parameter and the resulting percentage of segmented flower pixels in each image.

	Mean	St. Dev.	r^2
Min Hue	0.1216	0.0111	0.002
Max Hue	0.1768	0.0084	0.013
Min Sat	0.2902	0.0916	0.022
Max Sat	1.0000	0.0000	0.086
Min Int	0.3763	0.0584	0.277
Max Int	0.7513	0.0947	0.003

were collected. Variation in minimum saturation may be due to hot spot effects in the flower areas of the images. Hot spots on flowers were characteristically diluted with white light, thereby reducing the level of purely saturated color and requiring lower values for the minimum saturation parameter to adequately segment flowers. Maximum and minimum intensity parameters inherently depended on the overall brightness of the image, which was certainly related to the outdoor lighting conditions when images were collected. Histograms of the HSI parameter sets (not shown) demonstrated normal distributions, which justified the use of the Monte Carlo sampling approach to process the remaining images in the data set.

Coefficients of determination were calculated to assess the relationship between each of the six HSI parameters and total flower cover in the manually segmented images. For five of the six parameters, the coefficient of determination was less than 0.1, indicating relatively little correlation between these variables. This further justified the use of the Monte Carlo sampling approach to image processing, because manual selections for five of the six HSI parameters had no dependence on the amount of flowers present in the image. A slight correlation of 0.277 existed between the minimum intensity parameter and flower cover. Likely, this was the result of shadowing. As the canopy grew and increased its flower cover, there was greater likelihood that some flowers would be shaded by nearby parts of the plant. This shadowing effect lowered the intensity of the pixels defining the shaded flowers, and selection of shaded flower pixels during manual segmentations lowered the threshold for the minimum intensity parameter. This result somewhat complicated the use of the Monte Carlo approach for image processing, because the minimum intensity parameter and flower cover were not independent. To address the issue, we included code within the Monte Carlo algorithm to adjust the mean minimum

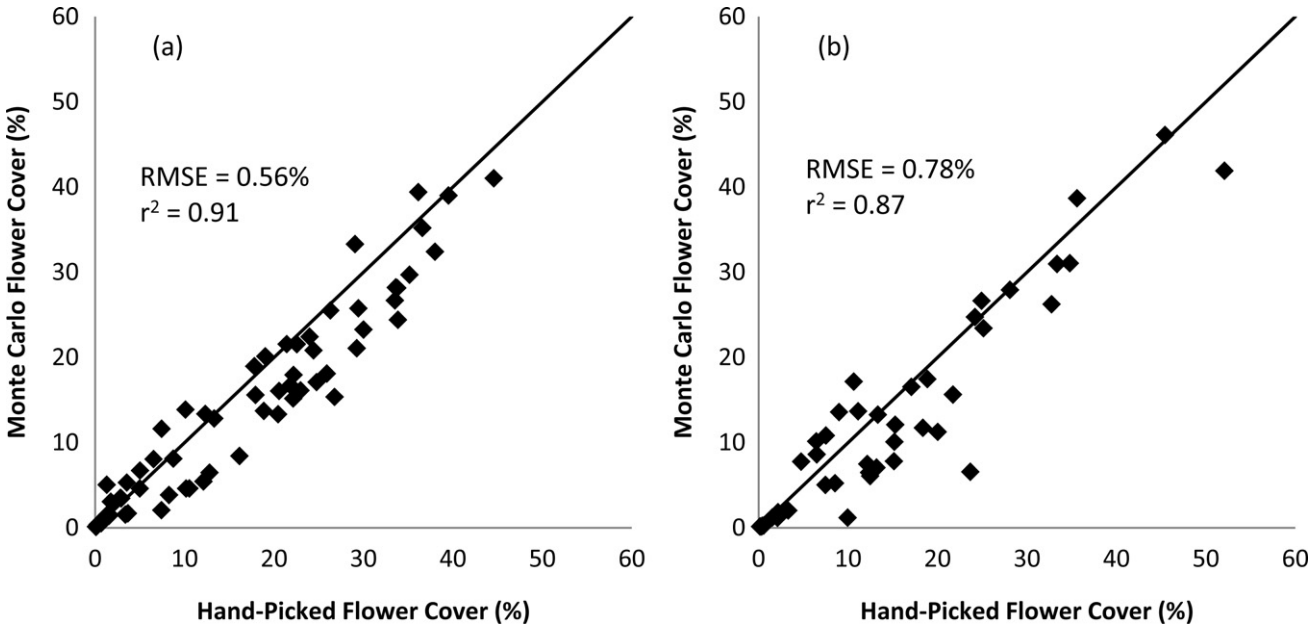

Fig. 3. Mean flower cover from the Monte Carlo image segmentation procedure versus flower cover from manual image segmentations for the (a) 2007–2008 and the (b) 2008–2009 growing seasons.

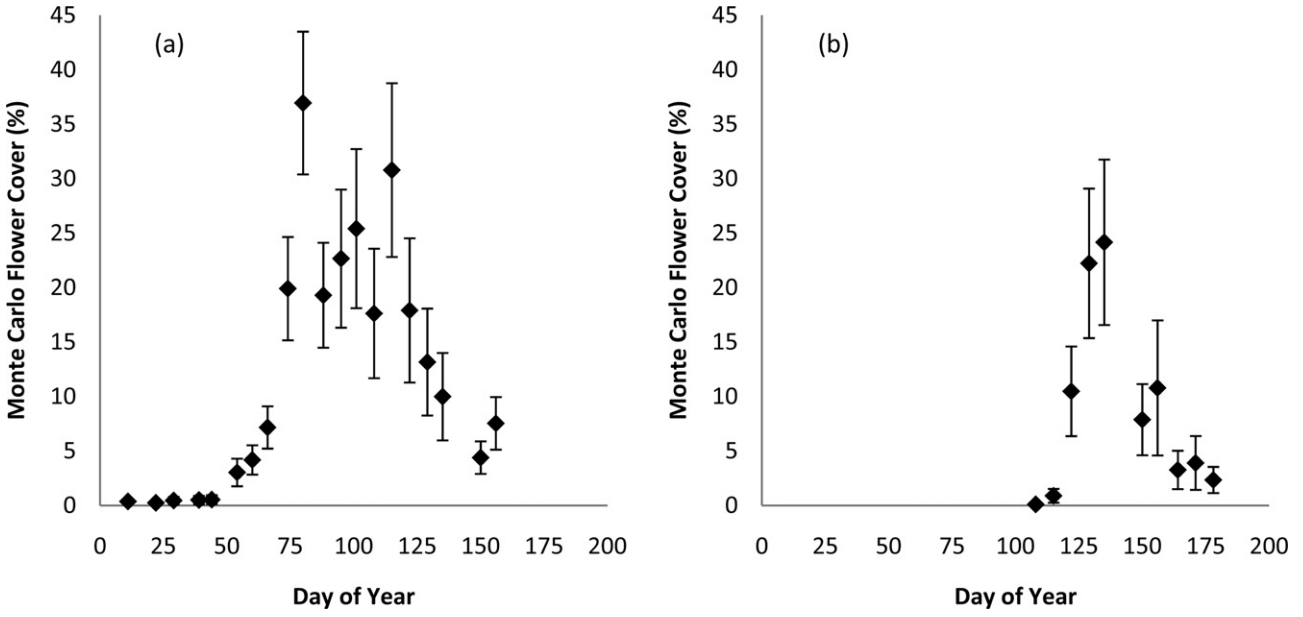

Fig. 4. Mean Monte Carlo flower cover for all images collected over the biomass sampling areas in the (a) first and (b) second plant date treatments in the 2007–2008 growing season. Error bars represent one standard deviation from the mean on each image collection date.

intensity parameter based on the mean flower cover calculated from previous Monte Carlo iterations. The adjustment to the minimum intensity parameter was based on the regression equation defining the relationship between minimum intensity and flower cover from the manual segmentations.

The standard deviation of flower cover resulting from 500 iterations of the Monte Carlo algorithm was up to 11% for some images. This demonstrates the high sensitivity of the flower cover estimates to the input HSI parameters for image segmentation. By computing the mean flower cover resulting from the Monte Carlo approach, we aimed to obtain a better estimate of flower cover than by relying on a segmentation result for a single set of HSI parameters. A comparison of the mean flower cover from the Monte Carlo procedure to the flower cover values from manual segmentations demonstrates the usefulness of the Monte Carlo procedure to generate reliable estimates of lesquerella flower cover (Fig. 3). For the images collected in the 2007–2008 and the 2008–2009 growing seasons, the root mean squared errors between flower cover values from Monte Carlo and manual segmentations were 0.56% and 0.78%, respectively. Coefficients of determination for these variables were 0.91 and 0.87, respectively. From this, we conclude that the Monte Carlo approach can provide adequate estimates of lesquerella flower cover, while also addressing parameter uncertainty associated with image segmentation in HSI color space.

3.3. Predicting flower count

Preliminary analysis of the image processing results for the 2007–2008 growing season provided guidance on how to build regression models to relate image-based metrics to field measure-

ments of flower count (Fig. 4). These results highlighted some unforeseen limitations in our field methods that should be corrected in future studies. Primarily, we conclude that lesquerella flowering is a very dynamic temporal process. In the first planting date treatment in 2007–2008, the mean flower cover among the plots progressed from 7% on day of year (DOY) 66–37% on DOY 80. In just two weeks, flower cover increased by a factor of five. Similarly for the second planting date treatment, the mean flower cover increased from 1% on DOY 115 to 22% on DOY 129. Progression of flowering to an initial peak occurred quite rapidly for both planting date treatments. Once the initial peak in flowering was reached, flowering patterns began to fluctuate, as demonstrated between DOY 80 and DOY 115 for the first planting date treatment. During this 35-day period, mean flower cover varied by up to 20% with no clear temporal pattern. Lesquerella flowering was also shown to be spatially variable. Error bars in Fig. 4 represent one standard deviation in flower cover percentage among the areas that were imaged each day. One standard deviation in flower cover percentage was up to 8% in some cases.

We took great care to coordinate the spatial location of imaging and biomass sampling endeavors in the field. However, due to issues with field labor scheduling, the time difference between imaging and biomass sampling could be up to several days in length. Given the dynamic temporal nature of lesquerella flowering patterns, this created some problems when attempting to relate flower cover percentage from images to the measured flower counts from biomass samples. To correct the problem, we implemented a procedure to interpolate the flower cover percentages between imaging dates, such that the interpolated flower cover percentage matched the date of biomass sampling. Results from the final image collected before biomass sampling was used as one point for interpolating. Since no images were collected after biomass sampling at a particular location, we used the mean and standard deviation of flower cover from all images collected on the next imaging date to estimate the second interpolation point. The interpolation procedure substantially improved our ability to relate image-based flower cover percentages to actual flower counts with meaningful coefficient of determination and low root mean squared error. Further improvements were obtained by eliminating from the regression analysis the information from DOY 88 through DOY 115 in the first planting date treatment. Since the lesquerella flowering patterns were quite sporadic during this time, it was uncertain if linear interpolation was an appropriate way to temporally correct this data. Regression model fitting then focused only on the earlier (DOY 11 to DOY 80) and later (DOY 122 to DOY 156) portions of the first planting date treatment when temporal correction by linear interpolation was more justified by the data. All temporally corrected flower cover percentages from the second planting date treatment were included in the regression exercise without problem. To render this interpolation step unnecessary, future studies should strive to coordinate crop canopy imaging and biomass sampling in both time and space.

Measured flower counts versus Monte Carlo flower cover for the data collected in the 2007–2008 growing season are shown in Fig. 5. A simple linear regression between the two variables resulted in a coefficient of determination of 0.75, which is slightly lower than that obtained by Adamsen et al. (2000). However, unlike Adamsen et al. (2000), we elected not to include an intercept, since it did not make physical sense to do so. Zero flower cover should equal zero flower count. From the regression equation, we found a multiplier of 30.291 to convert from Monte Carlo flower cover to flower count. Using this multiplier to convert the flower cover to flower count for the 2007–2008 growing season, we computed a root mean square error of 159 flowers. This error is substantially lower than previously reported attempts to estimate lesquerella flower count from both digital images (Adamsen et al., 2000) and hyperspectral

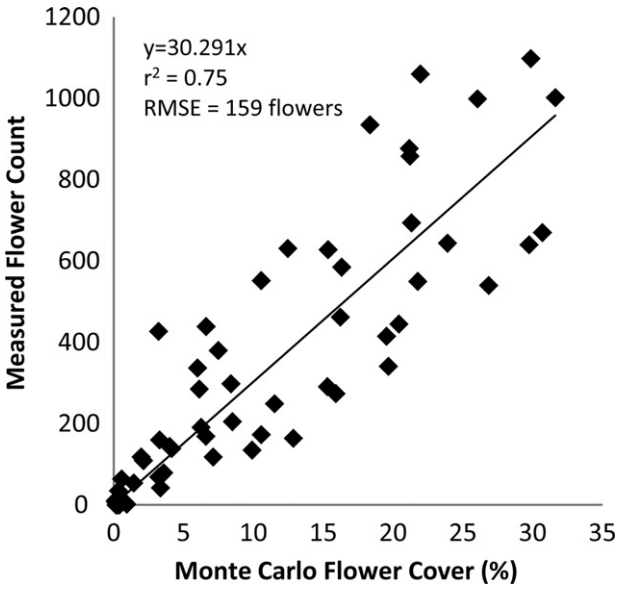

Fig. 5. Regression analysis of measured flower count versus Monte Carlo flower cover for the data collected in the 2007–2008 growing season.

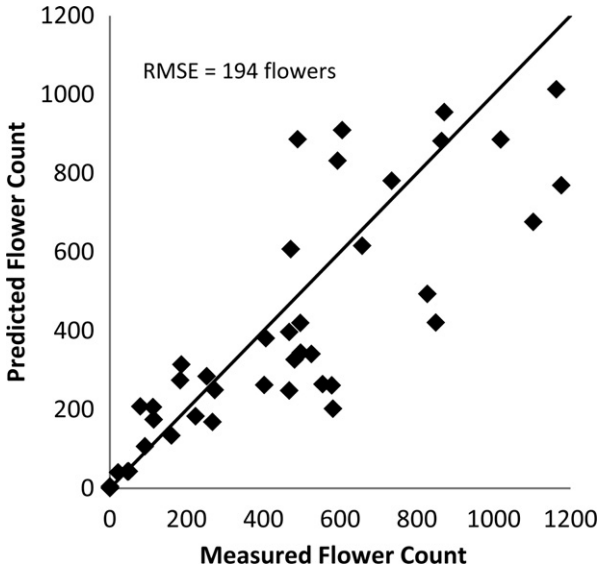

Fig. 6. Predicted flower count (obtained from Monte Carlo flower cover) versus the measured flower count for the data collected in the 2008–2009 growing season.

remote sensing (Thorp et al., 2011). Our results were also obtained over a wider range of flower cover, from 0% to 30%, as compared to Adamsen et al. (2000), whose data ranged from 0% to 10%.

To test the 30.291 multiplier on an independent dataset, Monte Carlo flower coverages from the 2008–2009 growing season were converted to flower count and plotted versus measured flower count (Fig. 6). The resulting root mean squared error between predicted and measured flower counts was 194 flowers. This error value is also lower than that previously reported by both Adamsen et al. (2000) and Thorp et al. (2011).

Results for estimating flower counts from the number of connected components identified in the images were less favorable than estimates from flower cover percentages (Fig. 7). Moderate correlations were achieved with coefficients of determination of 0.58 and 0.38 for the 2007–2008 and 2008–2009 growing seasons, respectively. However, for the approach to make physical sense, the regression multiplier should ideally be 1.0, and there should be

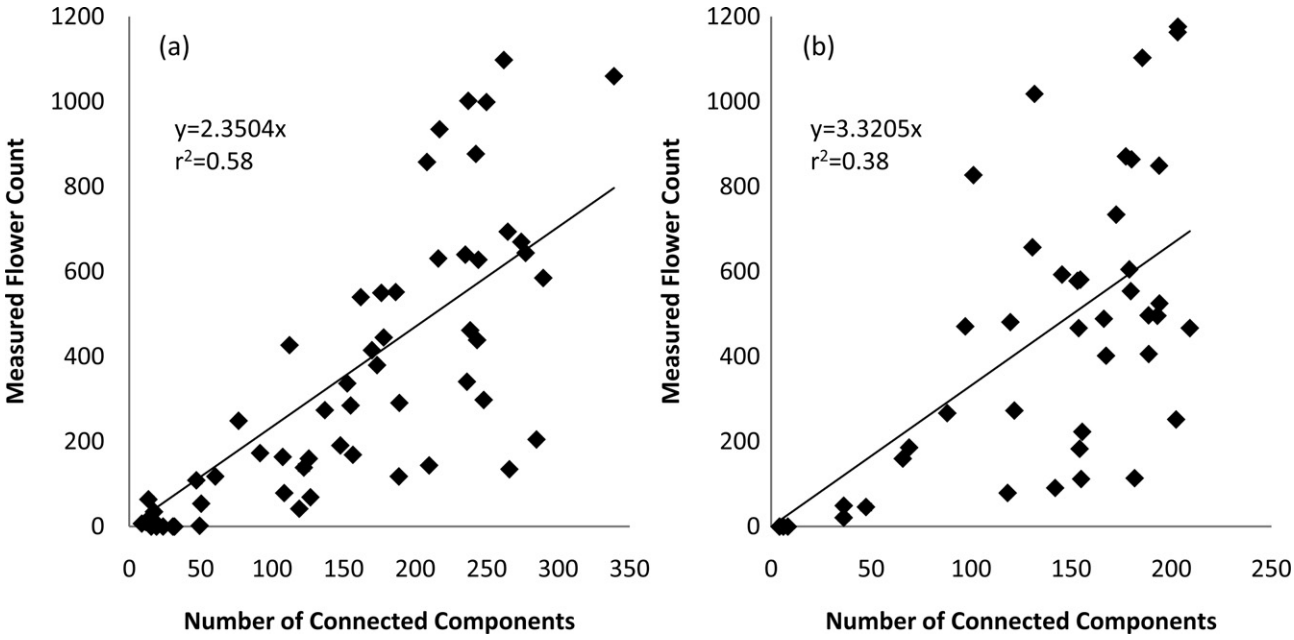

Fig. 7. Measured flower count versus the number of connected components delineated in images from the (a) 2007–2008 and (b) 2008–2009 growing seasons.

a direct one-to-one relationship between connected components and flower count. In reality, the number of connected components was consistently less than the measured flower count, and regression multipliers of 2.35 and 3.32 were required to relate these two variables in the 2007–2008 and 2008–2009 growing seasons, respectively. Likely, the underestimation of flower count by connected components is due to overlapping bunches of flowers in the images being identified as a single connected component and counted as a single flower (Fig. 2d). The problem becomes greater as flower numbers increase, and flowers become more dominant in the image scene. This idea is supported by the widening range of measured flower count as connected component number increases (Fig. 7). Adamsen et al. (2000) was able to achieve higher correlations between connected components and flower counts; however, flower cover was not more than 10% in their images. We have images with flowers covering more than 30% of the area. Adamsen et al. (2000) also did not demonstrate a one-to-one relationship between connected component number and flower count. Given the spatial complexity of lesquerella flowering, we favor an approach that uses total flower pixel number or flower cover percentage to estimate flower counts, rather than attempting to segment individual flowers. For our data, the former approach was more practical than the latter.

3.4. Monitoring flower count

Application of the image processing algorithms to the images collected along the 180 m transect in each plot demonstrated the ability of the procedure to monitor lesquerella flowering patterns for each of the planting date treatments. Of particular interest is the ability of the procedure to identify the day of year (DOY) or days after planting (DAP) that peak flowering occurred. For the 2007–2008 experiment, the average peak flowering date was March 20 (DOY 80; DAP 174) for the first planting (Fig. 8). Flowering patterns for this treatment then fluctuated for 49 days until May 9 (DOY 129; DAP 223) before permanently declining. A similar flowering pattern was observed from concurrent remote sensing measurements of the canopy, although the actual peak flowering date was observed on April 24 (DOY 115; DAP 209) with the spectral reflectance approach (Thorp et al., 2011). Peak flowering was May 14 (DOY 135; DAP 89) for the second planting in 2007–2008. The results demonstrated a notable reduction in the length of time available for flowering between fall-planted and winter-planted lesquerella. For the 2008–2009 experiment, the average peak flowering date was April 20 (DOY 110; DAP 196) for the first planting. In the second season, the fall-planted lesquerella did not experience the prolonged flowering period as was observed in the previous season. Rather, flowering peaked abruptly and then permanently declined. A similar result was observed using the canopy spectral reflectance approach (Thorp et al., 2011). Peak flowering for the second and third planting dates occurred on May 6 (DOY 126; DAP 118) and May 21 (DOY 141; DAP 104), respectively. These peak flower dates were a few weeks earlier than that obtained from the spectral reflectance approach; however, the general flowering patterns and peak mean flower counts estimated with the two methods were reasonably similar.

3.5. Relevance to past studies

In the present study and a companion study (Thorp et al., 2011), we have demonstrated two approaches for remotely monitoring lesquerella flowering patterns: the previous approach based on canopy spectral reflectance measurements and the present approach using digital images of the canopy. There are several trade-offs between the two approaches. To complete the previous study, we implemented a hand-held spectroradiometer valued at roughly $20,000. Likely, more inexpensive radiometers could be developed specifically for lesquerella flower detection if sufficient demand existed. However, this type of instrument does not yet exist. In the present study, we obtained better results using a consumer-grade digital camera valued at roughly $800. Likely, lesquerella growers or breeders already own a digital camera that could be used to sufficiently replicate the results of the present study. Several types of platforms could be developed for digital image collection, including hand-held systems, tractor-mounted systems, or field stationary platforms. Given the dynamic nature of lesquerella flowering patterns, we favor the latter approach with wireless transmission of images to a computer dedicated for image

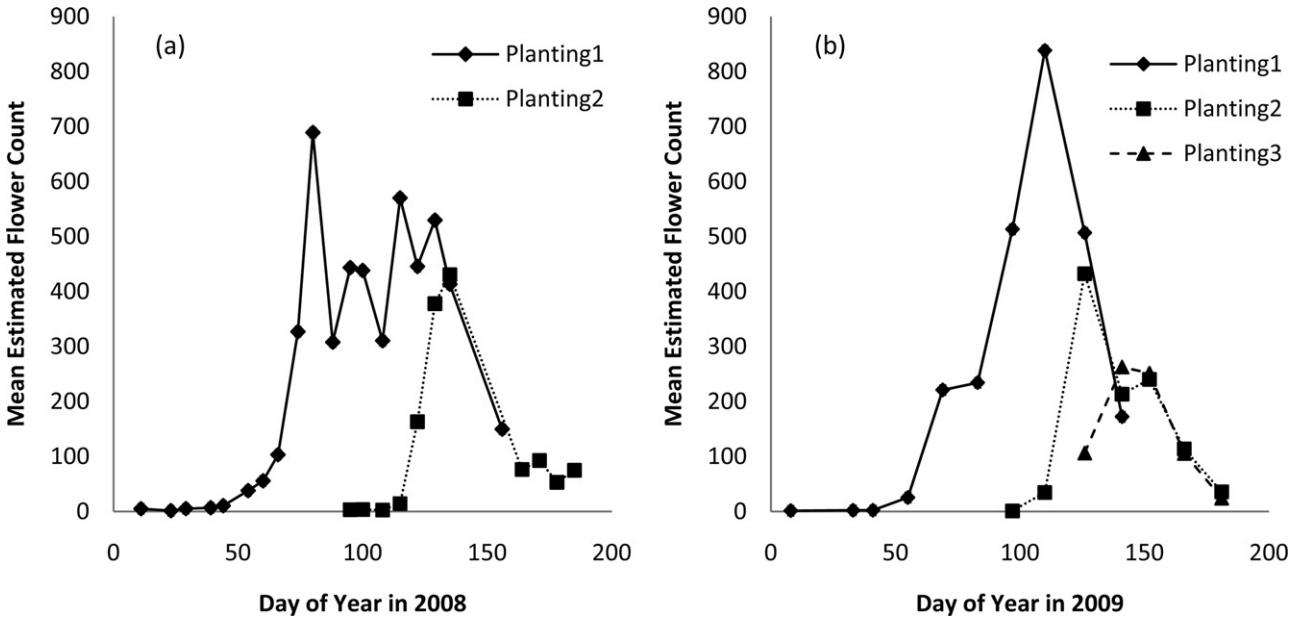

Fig. 8. Mean estimated flower count using the images collected along linear transects from January to July in the (a) 2008 and (b) 2009 growing seasons.

processing. In this way, lesquerella canopies could be monitored on a sub-hourly basis. One drawback to the digital imaging approach is that one camera can cover only a small spatial extent, on the order of a few meters. Vehicle-mounted cameras could be designed to map an entire field, but such a system would be inefficient with the time required to collect and process the images. The other option for wider spatial coverage is to collect images at a higher altitude from aerial or satellite platforms. However, the image processing routines developed in the present study would not apply to images collected from high altitude, and the spectral reflectance approaches of the previous study would be required.

Image processing algorithms developed in the present study refine the algorithms developed previously by Adamsen et al. (2000). The most fundamental difference between the present approach and that of Adamsen et al. (2000) is the inclusion of the image transformation to the HSI color space. This step decoupled the image intensity from the color information (hue and saturation) such that image thresholding could be more directly based on the color components of the image. The advantage of this is most effectively demonstrated by the relatively low variability in the hue parameter for images collected over a wide range of conditions (Table 2). Maximum and minimum hue and maximum saturation parameters for segmentation of yellow lesquerella flowers were quite narrowly defined and were quite stable in the HSI color space. Use of the Monte Carlo approach to address larger uncertainties in the remaining parameters resulted in a much more computationally intensive algorithm than that of Adamsen et al. (2000). However, refinement of the image collection protocols based on the results of our HSI transformation should reduce this parameter variation, which is likely attributable to variability in outdoor lighting conditions. For example, incorporation of an appropriate shading mechanism to block the direct solar beam or use of flash bulbs for consistent canopy illumination would likely provide images with reduced intensity variation. Appropriate shading of the image scene should also reduce hot spots that cause larger variation in the minimum saturation parameter (Table 2). These modifications to the image collection protocol may eliminate the need for Monte Carlo analysis and may allow for image thresholding based on a set of well-defined, well-controlled, and highly meaningful HSI parameters.

4. Conclusions

Vibrant yellow flowers are prominently displayed at the top of lesquerella canopies throughout anthesis. This characteristic permits the development of remote techniques for monitoring the lesquerella flowering process. Because lesquerella flowering is indeterminate in nature and also quite dynamic, digital image processing can be used to track flowering patterns and to identify key growth stages, such as peak flowering and flowering decline. Such information is useful for determining the optimum times for irrigation cut-off, desiccant application, crop harvest, or installation of bee boxes to maximize pollination. Since lesquerella in Arizona will likely be produced as a winter annual in rotation with cotton or a bioenergy crop, earlier lesquerella harvest dates allow more flexibility to plan field activities for the summer crop. The image processing techniques developed here could also be used in breeding programs to select cultivars that exhibit more desirable developmental characteristics for lesquerella, such as earlier peak flowering dates. Future work will provide techniques for predicting yield based on in-season lesquerella flowering patterns. We will also focus on improving our data collection protocol in the future. Improved ability to predict flower count from image-based flower cover is likely possible with better temporal correspondence between biomass sampling and image collection. Improved image segmentation is also likely with the development of appropriate shading mechanisms for better control of the intensity and saturation characteristics of the images. Our methods are quite simple and practical, requiring only a digital camera and the image processing algorithms. The algorithms are also quite generic and could be easily applied for many other digital image processing applications that require image segmentation based on color features.

Acknowledgements

The authors are grateful to the field technicians who conducted the experiments and collected and processed the digital images and biomass samples, including Ms. Suzette Maneely, Mr. Rafael Chavez-Alcorta, Ms. Gail Dahlquist, and Ms. Stephanie Johnson.

References

Adamsen, F.J., Coffelt, T.A., 2005. Planting date effects on flowering, seed yield, and oil content of rape and crambe cultivars. Ind. Crop. Prod. 21 (3), 293–307.

Adamsen, F.J., Coffelt, T.A., Nelson, J.M., 2003. Flowering and seed yield of lesquerella as affected by nitrogen fertilization and seeding rate. Ind. Crop. Prod. 18 (2), 125–131.

Adamsen, F.J., Coffelt, T.A., Nelson, J.M., Barnes, E.M., Rice, R.C., 2000. Method for using images from a color digital camera to estimate flower number. Crop Sci. 40 (3), 704–709.

Bonaparte, E.E.N.A., 1975. The effects of temperature, daylength, soil fertility and soil moisture on leaf number and duration to tassel emergence in *Zea mays* L. Ann. Bot. 39 (163), 853–861.

Dierig, D.A., Thompson, A.E., Nakayama, F.S., 1992. Lesquerella commercialization efforts in the United States. Ind. Crop. Prod. 1 (4), 289–293.

Ferris, R., Ellis, R.H., Wheeler, T.R., Hadley, P., 1998. Effect of high temperature stress at anthesis on grain yield and biomass of field-grown crops of wheat. Ann. Bot. 82 (5), 631–639.

Geller, D.P., Goodrum, J.W., 2004. Effects of specific fatty acid methyl esters on diesel fuel lubricity. Fuel 83 (17–18), 2351–2356.

Gonzalez, R.C., Woods, R.E., 1992. Digital Image Processing. Addison-Wesley Publishing Company, Inc.

Johansen, C., Waseque, M., Begum, S., 1985. Effect and interaction of photoperiod, temperature, water stress and nitrogen on flowering and growth in jute. Field Crop. Res. 12 (4), 397–406.

Kaleita, A.L., Steward, B.L., Ewing, R.P., Ashlock, D.A., Westgate, M.E., Hatfield, J.L., 2006. Novel analysis of hyperspectral reflectance data for detecting onset of pollen shed in maize. Trans. ASABE 49 (6), 1947–1954.

Mogensen, V.O., Jensen, C.R., Mortensen, G., Thage, J.H., Koribidis, J., Ahmed, A., 1996. Spectral reflectance index as an indicator of drought of field grown oilseed rape (*Brassica napus* L.). Eur. J. Agron. 5 (1–2), 125–135.

Montes, J.M., Melchinger, A.E., Reif, J.C., 2007. Novel throughput phenotyping platforms in plant genetic studies. Trends Plant Sci. 12 (10), 433–436.

Moser, B.R., Cermak, S.C., Isbell, T.A., 2008. Evaluation of castor and lesquerella oil derivatives as additives in biodiesel and ultralow sulfur diesel fuels. Energ. Fuel. 22 (2), 1349–1352.

Moser, S.B., Feil, B., Jampatong, S., Stamp, P., 2006. Effects of pre-anthesis drought, nitrogen fertilizer rate, and variety on grain yield, yield components, and harvest index of tropical maize. Agric. Water Manage. 81 (1–2), 41–58.

Nielsen, D.C., Nelson, N.O., 1998. Black bean sensitivity to water stress at various growth stages. Crop Sci. 38 (2), 422–427.

Ohnishi, S., Miyoshi, T., Shirai, S., 2010. Low temperature stress at different flower developmental stages affects pollen development, pollination, and pod set in soybean. Environ. Exp. Bot. 69 (1), 56–62.

Pimstein, A., Eitel, J.U.H., Long, D.S., Mufradi, I., Karnieli, A., Bonfil, D.J., 2009. A spectral index to monitor the head-emergence of wheat in semi-arid conditions. Field Crop. Res. 111 (3), 218–225.

Tang, L., Tian, L., Steward, B.L., 2000. Color image segmentation with genetic algorithm for in-field weed sensing. Trans. ASAE 43 (4), 1019–1027.

Thorp, K.R., Dierig, D.A., French, A.N., Hunsaker, D.J., 2011. Analysis of hyperspectral reflectance data for monitoring growth and development of lesquerella. Ind. Crop. Prod. 33 (2), 524–531.

Vina, A., Gitelson, A.A., Rundquist, D.C., Keydan, G., Leavitt, B., Schepers, J., 2004. Monitoring maize (*Zea mays* L.) phenology with remote sensing. Agron. J. 96 (4), 1139–1147.