ABSTRACT - Weed mapping is a useful tool for site-specific herbicide applications. The objectives of this study were (1) to determine the percentage of land area covered by weeds in no-till and conventionally tilled fields of common bean using digital image processing and geostatistics, and (2) to compare two types of cameras. Two digital cameras (color and infrared) and a differential GPS were affixed to a center pivot structure for image acquisition. Sample field images were acquired in a regular grid pattern, and the images were processed to estimate the percentage of weed cover. After calculating the georeferenced weed percentage values, maps were constructed using geostatistical techniques. Based on the results, color images are recommended for mapping the percentage of weed cover in no-till systems, while infrared images are recommended for weed mapping in conventional tillage systems.

Keywords: precision agriculture, geostatistic, common bean.

INTRODUCTION

In Brazil, common beans (Phaseolus vulgaris), which are cultivated in all regions, are a staple in the diet and an important source of protein for poor people. Common beans are typically grown by small producers, but several large producers have also grown the crop since the mechanical harvests availability. Generally, the larger producers use advanced technologies that allow them to achieve higher yields than the smaller producers.

According to Kozlowski et al. (2002), weed infestation can reduce common bean production by 15 to 80%. Herbicide use is the most common method used to control weeds because of its high efficiency and ease of application (Concenço et al., 2006). However, farmers tend to apply a uniform dose of herbicides over entire areas, regardless of the...
prevalence of weeds. To improve the efficiency of herbicides, farmers should apply the product only where necessary. Thus, weed mapping, which identifies infestations throughout a production area, has been pointed as one of the tools for the site-specific application of herbicides (Singh et al., 2011).

Site-specific herbicide applications not only enable a higher efficiency of weed control but also result in higher grain yields (Machado et al., 2006). This technique optimizes herbicide use and reduces the environmental impacts of the product. According to Shiratsuchi & Christoffoletti (2002), several experiments have demonstrated that site-specific herbicide applications are economically feasible.

According to Lamb & Brown (2001), although weed mapping can be performed with different techniques, one of the most promising is remote sensing. These authors state that one of the main advantages of remote sensing is the speed at which weed infestation maps can be generated. This process relies on the extraction of information from digital images instantaneously acquired by platforms attached to farm machinery.

Spectral bands or indices aid in plant mapping by minimizing variations in external factors and by enhancing the objects of interest. Moreover, with image processing and analysis of techniques, important information can be extracted to more easily discriminate between crops and weeds (Sartori et al., 2005).

Most of the systems proposed for acquiring spatial weed data have used two approaches, namely continuous and discrete area sampling (Rew & Cousens, 2001). In continuous weed sampling, data are collected over the entire field by using visual or remotely sensed imagery by an expert. Weed remote sensed images can be acquired by aircraft and satellite platforms (Bajwa & Tian, 2001; Casady et al., 2005; Cuneo et al., 2009) or tractor-based systems, known as machine vision systems (Lee et al., 1999; Lamm et al., 2002). In discrete weed sampling, data are collected only from pre-defined sites throughout the field and interpolation method is used to estimate weed infestation for the entire field (Expósito et al., 2003).

Machine vision systems have been usually proposed for real-time herbicide application by scanning the entire field. The premise of this research is that the farmer can use the center-pivot irrigation system as a machine vision platform for sensing the weed. However, designing a system for scanning the entire field under the center-pivot structure could be a complex task that could require an expensive solution. Sampling images in some points of the center-pivot field could be a feasible option. To create a weed infestation map from sample images, geostatistics can be used to calculate values for areas where information is not available.

Site-specific herbicide applications can help farmers optimize their production, improve the quality of their crops, and reduce environmental problems. Thus, the objectives of this work were (1) to map the percentage of weed cover using a machine vision system and (2) to test two types of cameras, color and near infrared.

**MATERIALS AND METHODS**

This study was conducted in a 0.8-hectare experimental field irrigated by a center pivot system and planted in common beans. The field was divided in half and equally cultivated under no-till and conventional agricultural practices.

**Image Acquisition System**

The image acquisition system consisted of two 1/3" CMOS cameras, one monochrome and one color, that were controlled by a portable microcomputer via a 10 m cable and an image acquisition board (PCMCIA IEEE 1394). The cameras used C-mount lenses with 2.8 mm focal lengths and captured two images of the same scene simultaneously. An infrared longpass filter with a cut-off wavelength of 695 nm was coupled to the monochrome camera to acquire near infrared images (NIR). The color camera acquired R (red), G (green), and B (blue) image bands. The images were saved in bitmap (BMP) format with dimensions of 480 x 640 pixels.

The system was attached to the center-pivot structure and pointed down to focus the
crop rows. While the center-pivot structure was running on its circular trajectory, the cameras were manually moved on the center-pivot structure to acquire sample images over the entire field. Using this setup, each image represented 4.9 x 6.5 m of the field for both types of images, color and NIR.

Wooden flag stakes, georeferenced with a Trimble® Pathfinder Pro XRS DGPS, were placed in the field in a regular grid pattern. The stakes were arranged in circular sections and the distance between stakes in the same section was approximately 6.5 m. The distance between sections was also 6.5 m. When a stake was observed near the center of the camera’s field of view, an image was acquired. In total, 75 and 82 georeferenced images were acquired in the conventionally tilled and no-till fields, respectively.

All images were taken under cloudy sky conditions at 25 days after emergence (DAE) of the plants.

**Image processing and analysis**

An image processing algorithm was developed in the NI LabView® environment to output the percentage of weed cover in each georeferenced image. This percentage was estimated for both color and NIR images of the same scene.

The main idea implemented in the algorithm was to assume that plants between two crop rows were weeds. Firstly, plants in the image were discriminated from the background. Then, the identified plants were classified as crop and weed by detecting the crop row location. Finally, once the crop rows were identified, the other segmented plant pixels were summed to estimate the weed cover in each image.

Color and NIR images were tested. In color image, the color differences between plant and other objects in the scene could favor the plant discrimination in images by enhancing the green information. In NIR images, the spectral response differences between plant and soil could favor plant discrimination using only one NIR band instead of three RGB bands used in color image.

To enhance the contrast between the plants and the background, image processing began with the transformation of color images (Figure 1A) into excess green index images (Figure 1B). No contrast transformation was performed on the NIR images.

The excess green index (ExG) was calculated according to Woebbecke et al. (1995) from equation 1 as follows:

\[
ExG = \frac{2G - R - B}{R + G + B}
\]

where R, G, and B are the values of each pixel of the color image in the red, green, and blue bands, respectively.

The NIR and ExG images were then automatically segmented by the iterative method proposed by Yang et al. (2001) in which a single-threshold classifies the pixels into plants (crops and weeds) and background (soil, straw, etc.). After segmentation, opening and closing morphological operations were applied for noise reduction (Figure 1C).

Next, the angle of the crop rows in the binary images was determined using the Hough Transform (Duda & Hart, 1972). Since the cameras were fixed on the center pivot structure, crop rows had different directions in different images. Once the crop row angle was established, the image was rotated until the crop rows were vertical (Figure 1D). A block of 320 (V) x 480 (H) pixels was cut in the center of the rotated image to ensure that the crop rows had the same length in all image blocks (Figure 1E). Summing each column of the image block, the algorithm identified the position and the width of the crop rows and removed them, leaving only plants identified between the crop rows (Figure 1F). Having image blocks with the same crop row length was essential for identifying the crop row position and width; on the other hand, the image column summing would be too different for each crop row. Finally, based on the number of pixels representing plants between crop rows, the percentage of weed cover was estimated for each georeferenced image.

Each image generated a georeferenced weed cover sampling. Thus, at the ending of the image processing and analysis, a grid
Figure 1 - Images processing steps for estimating the percentage cover of weeds between crop rows (A) the original color image, (B) the excess green index image, (C) the thresholded and filtered image, (D) the rotated image, (E) the image block, and (F) the plant between crop rows image.

Sample of the weed cover was obtained with 75 and 82 sampling points in the conventionally tilled and no-till fields, respectively.

Mapping the percentage cover of weeds

The GS+ software (version 9.0) was used for geostatistical analysis of the georeferenced data. The objective of this analysis was to determine if the data points were spatially dependent and to build a map by interpolation. Analyses were performed for both no-till and conventional tillage systems and for both color and NIR cameras.

Geostatistical analysis consisted of calculating summary statistics, fitting predefined semivariogram models, choosing
the semivariogram model and interpolating by kriging.

The GS+ uses nonlinear regression for fitting the predefined semivariogram models: spherical, Gaussian, linear, and exponential. According to Isaaks & Srivastava (1989), these are the most common semivariogram models and, in the fitting process, the following semivariogram parameters are evaluated: the nugget effect, the sill, and the range. Interpolation by ordinary kriging was performed using each of these models. The criterion for choosing the best semivariogram model was based on parameters obtained in the cross-validation analysis. These parameters were the correlation coefficient and standard error between measured and estimated values. Therefore, the model with correlation coefficient closer to 1 and standard error closer to 0 was selected.

Although the quality of the weed map generated by geostatistical techniques is very sensitive to the sampling strategy, this study did not aim to discuss this issue. Further details of the weed map quality generated by geostatistical interpolation can be found at Clay et al. (1999) and Cousens et al. (2002).

To validate the weed maps, images with dimensions of 2304 (V) x 3072 (H) pixels were captured at a height of 1.60 m with a digital camera (CASIO EX-Z75) and saved in JPEG (Joint Photographic Experts Group) format. Using this setup, each image represented 1.0 x 1.4 m of the field. The resolution of those images was much higher than the resolution of the images used by the proposed system since the camera had a smaller field of view and used a greater number of pixels.

The geographic coordinates of each validation image, acquired by the Trimble Pathfinder Pro XRS DGPS, were assumed to be in the center of the camera’s field of view.

The high-resolution validation images were transformed to ExG and segmented using the same algorithms described above. However, to obtain the percentage of weed cover, crop rows were manually removed in the computer. Using the geographic coordinates, the manual segmentation values were compared to the map values.

For the map validation process, 100 high-resolution images were randomly acquired at 26 DAE in the no-till and conventionally tilled fields.

RESULTS AND DISCUSSION

Summary statistics

The summary statistics obtained by analyzing the percentage of weed coverage between the rows are presented in Table 1. The values obtained with the color camera in no-till field were not significant at the 5% level in the Shapiro-Wilk normality test. According to Cressie (1993), when the normal distribution is rejected, as happened with the conventional tillage data, the non-normality must not compromise the continuity of analysis.

The kurtosis coefficient (Ck), which describes the degree of flattening of the frequency distribution curve, showed values greater than 0.0 for all cropping systems and cameras tested. This indicated that the distributions were leptokurtic. However,

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Median</th>
<th>Minimum</th>
<th>Maximum</th>
<th>CV</th>
<th>s</th>
<th>Cs</th>
<th>Ck</th>
<th>W</th>
</tr>
</thead>
<tbody>
<tr>
<td>CT-RGB</td>
<td>6.73</td>
<td>6.34</td>
<td>1.90</td>
<td>16.85</td>
<td>42.83</td>
<td>2.88</td>
<td>0.89</td>
<td>1.22</td>
<td>*</td>
</tr>
<tr>
<td>CT-NIR</td>
<td>8.49</td>
<td>8.24</td>
<td>3.11</td>
<td>19.92</td>
<td>35.54</td>
<td>3.02</td>
<td>0.79</td>
<td>1.83</td>
<td>*</td>
</tr>
<tr>
<td>NT-RGB</td>
<td>6.97</td>
<td>7.02</td>
<td>1.34</td>
<td>10.88</td>
<td>25.21</td>
<td>2.04</td>
<td>-0.47</td>
<td>0.15</td>
<td>ns</td>
</tr>
<tr>
<td>NT-NIR</td>
<td>8.33</td>
<td>8.40</td>
<td>1.87</td>
<td>12.76</td>
<td>28.02</td>
<td>2.33</td>
<td>-0.59</td>
<td>0.35</td>
<td>*</td>
</tr>
</tbody>
</table>

CV = coefficient of variation (%); s = standard deviation; Cs = skewness coefficient; Ck = kurtosis coefficient; W = Shapiro-Wilk normality test; * = non-normal distribution in the Shapiro-Wilk test at 5% probability; ns = normal distribution in the Shapiro-Wilk test at 5% probability.
values of Ck for the no-till system were close to zero, indicating a mesokurtic distribution.

In all cases, the mean and median values were similar, indicating a trend in the data toward a symmetrical distribution. This was also confirmed after analyzing the skewness coefficients, all of which were close to zero.

According to Cressie (1993), data normality is not a requirement for geostatistical analysis; the only requirement is that the distributions do not have very long tails, which could compromise the analysis. Thus, according to the summary statistics analysis, the shapes of the distribution curves of the collected data were assumed to have no effect on the geostatistical analysis.

**Geostatistical analysis and map building**

The semivariogram models, their parameters, and the cross-validation results are presented in Table 2. The range represents the distance above which the variables do not have spatial dependence between sampled values. All range values recorded did not surpass the maximum distance between sampling points and were not below the smallest sampling distance, thus validating the chosen sampling grid. According to Silva et al. (2010), points located in a circular area with a radius less than or equal to the range are spatially dependent and can be used to estimate values for locations that were not sampled.

The nugget effect (Co) represents the sampling errors and miscalculations of the variable, which causes discontinuity at the origin of the semivariogram. Therefore, Co is the component of spatial variability that cannot be related to a specific cause. Rather, it can be attributed to random variability, and smaller random variations are indicative of more accurate estimates (Vieira, 2000). The values of Co for the color camera in both tillage systems were the smallest ones and were close to zero.

In general, in the cross-validation analysis, all regression coefficient (RC) values were close to 1.0, especially the values from the color camera in the no-till system and from the NIR camera in the conventional tillage system. Both had RC values of 1.01. In all cases, the determination coefficient (R²) values were relatively low (less than 0.50). In contrast to the estimated values, the sampled values were more clustered and were closer to the ideal line in the conventional tillage system (Figures 2A and 2B) than in the no-till system (Figures 2C and 2D). This indicated that the estimated values in the conventional tillage system may have been more accurate than estimates in the no-till system.

Based on semivariogram models, values were interpolated using ordinary kriging to create a map of weed cover percentage between the common bean plant lines (Figures 3-6). Since the weed cover percentage ranged from 1.34 to 19.92%, these maps were arbitrarily created using five classes with intervals of 5% weed coverage for both planting systems and both types of cameras. By analyzing the four maps, the weed coverage

<table>
<thead>
<tr>
<th>Model</th>
<th>n</th>
<th>Nugget effect</th>
<th>Sill</th>
<th>Range</th>
<th>R²</th>
<th>SDI (%)</th>
<th>RSS</th>
<th>RC</th>
<th>R²</th>
<th>SE</th>
</tr>
</thead>
<tbody>
<tr>
<td>CT-RGB</td>
<td>75</td>
<td>1.70</td>
<td>7.62</td>
<td>20.80</td>
<td>0.96</td>
<td>78</td>
<td>0.30</td>
<td>1.10</td>
<td>0.49</td>
<td>0.13</td>
</tr>
<tr>
<td>CT-NIR</td>
<td>75</td>
<td>3.40</td>
<td>8.57</td>
<td>27.14</td>
<td>0.91</td>
<td>60</td>
<td>0.84</td>
<td>1.01</td>
<td>0.33</td>
<td>0.17</td>
</tr>
<tr>
<td>NT-RGB</td>
<td>82</td>
<td>0.95</td>
<td>3.99</td>
<td>22.50</td>
<td>0.75</td>
<td>76</td>
<td>0.54</td>
<td>1.01</td>
<td>0.23</td>
<td>0.21</td>
</tr>
<tr>
<td>NT-NIR</td>
<td>82</td>
<td>3.28</td>
<td>7.01</td>
<td>54.04</td>
<td>0.95</td>
<td>53</td>
<td>0.71</td>
<td>0.87</td>
<td>0.28</td>
<td>0.16</td>
</tr>
</tbody>
</table>

n = number of data points in geostatistical analysis; SDI = Spatial Dependence Index; RSS = residue sum of squares; RC = regression coefficient; SE = standard error.
Weed mapping using a machine vision system

was predominant in the 5-10% class in all maps.

The conventional tillage system maps had higher percentages of weed cover than the no-till system maps, as observed in the field. The no-till system generally had lower weed infestation than the conventional tillage system because the layer of straw in this system made weed seed germination difficult (Mateus et al., 2004). The weed cover in the conventional tillage system was more variable than that in the no-till system.

Validation of the proposed system

For the no-till field, the weed cover percentage was manually estimated in the validation images to be between 0.50 and 17.20%, with an average of 3.80% and a standard deviation of 3.01%. For the conventionally tilled field, values ranged from 1.26-23.68%, with an average of 10.07% and a standard deviation of 4.40%. Due to the presence of the straw layer on the soil surface, the no-till field tended to have less weed infestation than the conventionally tilled field.

For the conventionally tilled field, the average absolute errors between the manually estimated weed cover percentages in the validation images and those from the maps were 4.1 and 3.4% for color and NIR cameras, respectively. On the other hand, in the no-till field, the color camera was more efficient than the NIR camera, and the average absolute errors between the manually estimated weed cover percentages in the validation images and those from the maps were 4.0 and 5.4%, respectively. Thus, the excess green indexes from the color images were more efficiently discriminated between plant and straw than NIR radiation. Furthermore, the performance of the excess green index was similar in both tillage systems.

Figure 7 shows the weed cover values measured in the validation images versus those estimated in the maps on the same locations, for the conventionally tilled field. It was observed that the points tended to be above the red line
**Figure 3** - Map of percentage cover of weeds estimated using the color camera in the conventional tillage field.

**Figure 4** - Map of percentage cover of weeds estimated using the NIR camera in the conventional tillage field.

**Figure 5** - Map of percentage cover of weeds estimated using the color camera in the non-tillage field.
y=x, which means that the measured values tended to be greater than the estimated ones. Thus, for the conventionally tilled field, the proposed method underestimated the weed cover percentage. On the other hand, the method tended to overestimate the weed cover percentage in the no-till system (Figure 8). The contrast between soil and plants in the conventionally tilled field was greater with a higher resolution camera, and some small weeds were not captured in the images. This has led to an underestimation of weed infestation. In contrast, for lower resolution images of the no-till field, the low contrast between straw and weeds hampered the ability of the method to discriminate objects, and the method incorrectly classified some straw areas as weed areas.

The right determination of the crop row width was the main source of error in the proposed algorithm. The row width was defined by summing the pixel representing plants on each column of the binary image (Figure 1E). Thus, the row width was assumed uniform and equal to the distance between two image columns. Due to the fact that the shape of the crop row top-view is not uniform, two types of errors can occur: crop pixel was considered as weed pixel when the defined width was thinner than the actual width, or weed pixel was counted as part of the crop row when the defined width was wider than the actual width. These two types of errors can be seen in Figure 1F.

The thematic maps presented in this paper show the classes of weed cover percentage. This information alone is not useful for determining the necessary dosage or type of herbicide. Such decisions must be supported by additional information, including stage of maturity of crops and weeds, species of dominant weeds, and dangers posed by these species (Ferreira et al., 2007). These factors must be determined by field verification. These factors added to information from the maps could assist in decision making even if weed control methods are deemed unnecessary. Moreover, georeferenced information from the maps could reduce the randomness of weed scouting.

The weed mapping methods suggested in this paper may help monitoring weed growth after the application of pre-emergent herbicides. Thus, the effectiveness of herbicide applications could be reviewed and specific sites where weed control is ineffective could be identified.

Digital image processing and geostatistical techniques were used in this study to map the percentage of an area covered by weeds. Based on results obtained, color images were determined to be the most suitable for mapping weed infestations in no-till fields, while NIR images were determined to be the most suitable for conventionally tilled fields.
ACKNOWLEDGMENTS

The authors would like to thank Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (Capes), Fundação de Amparo a Pesquisa do Estado de Minas Gerais (FAPEMIG), and Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) for granting the scholarship and financing this study.

LITERATURE CITED


Figure 7 - Weed percentage cover measured in the validation images versus those estimated in the maps for the conventional tillage field using (A) color and (B) NIR cameras.

Figure 8 - Weed percentage cover measured in the validation images versus those estimated in the maps for the non-tillage field using (A) color and (B) NIR cameras.


