USE OF SATELLITE IMAGES TO OPTIMIZE REGIONAL MANAGEMENT STRATEGIES: ADAPTING A CLASSIFICATION PROCESS TO MAP COTTON FIELDS

P. J. Trichilo[1], L. T. Wilson[1], and R. K. Haldenby[2]
[1] Department of Entomology, Texas A&M University
College Station, TX
[2] Plains Cotton Growers, Inc.
Lubbock, TX

Abstract

Satellite Thematic Mapper (TM) images were used to quantify the size, shape, and location of cotton fields in 14 counties on the High Plains of Texas. These data were subsequently used to provide maps of cotton fields to Plains Cotton Growers, Inc. (PCG) for use in the High Plains Boll Weevil Diapause Control Program. The key to this study was in adapting a semi-automated system that could classify cotton and identify individual cotton fields. Several changes resulted in substantial improvements over last year, and facilitated a process that produced much more acceptable field maps. These modifications included the following. 1) Use of a satellite image collected in mid-August rather than mid-July. Only by mid-August had plants produced enough vegetation (enough ground cover) to be sufficiently recognizable in the image, especially during a dry year. 2) As mid-August is too late to be of value to the current year's program, an image from mid--August of the previous year was used. It was assumed that changes in field shape and location, as well as data lost to cloud cover would be corrected during land-based (vehicle) reconnaissance, otherwise known as ground-truthing. 3) A geographic information system (GIS) was used to produce buffer strips around digital line graphs (DLGs) of roads and highways, which delineated individual fields in the subsequent maps. Ground-truthing indicated that map accuracy was relatively high for irrigated fields, but less so for dryland cotton. Additionally, a method was developed to estimate susceptibility of cropland to spring weevil infestations, based on proximity to 1/2-mile buffers placed around overwintering habitats.

Introduction

Boll Weevil Management

The boll weevil, *Anthonomous grandis* Boheman, which is the most severe pest on cotton in Texas, overwinters as an adult in diapause. During an average winter in the Rolling and High Plains of Texas, where temperatures often drop below 0 °C, weevil overwintering survival is higher in habitats that produce abundant insulating leaf litter, than in habitats where leaf accumulation is reduced or absent. Weevils have little or no chance of survival in habitats that provide limited insulation (e.g., in a pasture or fallow field), whereas with the best types of insulation (e.g., shinnery oak), survival usually reaches 10 to 15% (Parajulee et al. unpublished). Weevil infestation in the spring is closely associated with winter severity and proximity of cotton fields to the best habitats, such as shinnery oak and various leafy hardwoods (Rummel & Adkisson 1970, Bottrell et al. 1972). Fields closer to habitats that foster higher weevil survival are at greater risk of infestation than fields more isolated from these types of vegetation. Given that the majority of weevils migrate one half mile (0.805 km) or less, in search of cotton after emerging from overwintering habitats (Rummel & Adkisson 1970), the ability to predict which fields are at risk would be of great economic benefit when formulating management strategies for mobile pests such as the boll weevil.

Each year Plains Cotton Growers, Inc. (PCG) implements the Boll Weevil Diapause Control Program. This program is based on knowing the location and size of all cotton fields within the program zone, an area that in 1995 included 19 counties and encompassed 2.7 million acres in the Rolling and High Plains of Texas. For the last three years the objective of our project has been to use satellite images to identify and map cotton fields within the Diapause Control Program zone. These maps are used by PCG in several ways including spatially identifying and quantifying the acreage of all cotton within the boundaries of the program, identifying the location of fields for the monitoring of boll weevils during late summer through early Fall (i.e., sampling weevil densities), and physically pinpointing the location and shape of cotton fields for applicator aircraft during the chemical control phase of the program. Maps are also useful in designating places having special sensitivity, where chemical applications should be avoided, such as areas with honey bees, horses, or exotic livestock (e.g., ostriches or emus).

During the first year of study, cotton fields were identified by visual inspection of the image, followed by digitizing the shapes of the fields directly from the computer screen (i.e., heads-up digitizing) using the geographic information system (GIS) ARC/INFO. To meet the PCG deadline of 25 August when maps are required for the weevil monitoring program, and accounting for the time necessary to acquire and process the data, a mid-July image was used. Identification of cotton fields was based upon a general understanding of where major cotton growing areas existed, and what was believed to be cotton (or at least agricultural land) reflectance in the satellite image. This process was especially difficult due to the immaturity of the plants at this time of the year, as cotton was often confused with surrounding soil and other vegetation. The process was further hindered by the drought conditions prevalent that year.

Reprinted from the *Proceedings of the Beltwide Cotton Conference*
Volume 2:1086-1090 (1996)
National Cotton Council, Memphis TN

The following year, image processing software, ERDAS Imagine®, was used in an attempt to identify cotton fields from a satellite image. As in the first year, a mid-July image was acquired for this procedure. A large amount of variability (i.e., a wide range of color patterns) was again found to be associated with cotton and was attributed to differences in vegetative growth, stemming from inconsistent rainfall and drought in many areas (Trichilo et al. 1995). Collecting signatures resulted in 32 different spectral patterns for cotton, which still left some cotton unclassified and caused some non-cotton areas (which shared similar patterns) to be classified. Spectral variability of cotton also made it difficult to visually separate individual fields from each other for delineation on a map. Individual fields often could not be clearly isolated in the original image and appeared to run together in the classified image. As time was a critical factor, this method was abandoned for that year, with a return to heads-up digitizing. The approximate locations of cotton fields had been acquired previously by field crews using global positioning system (GPS) devices, which record position in longitude/ latitude. These locations were displayed as points over the image on the screen (Trichilo, et al. 1995). This procedure allowed a rough estimate of field locations, but no specific identification of field shapes. Thus, cotton field recognition had to be based on the best estimate of the perceived color patterns of cotton as of mid-July. This effort resulted in six maps, one for each of six counties in the High Plains. Feedback from PCG personnel suggested that the sizes of most fields had been overestimated.

The objectives for this year's project were as follows. 1) To develop a method for identifying and classifying cotton from a satellite image. Mid-July was deemed too early to get a satisfactory image of cotton, as plants have not developed enough vegetation or sufficient ground cover to produce a unique signature. Moreover, acquisition of an image in mid-August, which shows better plant development than is available in mid-July, is too late to be of value to the current year's program. Therefore, an image from mid-August of the previous year was used. It was assumed that changes in field shape and location, as well as data lost to cloud cover would be corrected through land-based (vehicle) reconnaissance, otherwise known as ground truthing. 2) To identify individual cotton fields and produce maps of cotton for 14 counties in the High Plains Diapause Control Program, through an image classification process. Associated with the maps would be computerized data files providing coordinates in longitude and latitude, and area in acres of all mapped cotton fields. 3) To adapt a method to spectrally characterize vegetation associated with cotton-growing areas, and to quantify the relative susceptibility of all potential cotton acreage to infestation by weevils emerging in the spring from overwintering habitats. Traditional pest management strategies are based on the population dynamics of arthropods within a given field. A more appropriate strategy for mobile (flying) arthropods is one that addresses management on a landscape or regional level, focusing on multi-crop, multi-field ecosystems, and the influence of surrounding crops and native vegetation on arthropod population dynamics (Wilson et al. 1994).

Materials and Methods

Image Classification - Cotton Recognition

A Landsat 5, Thematic Mapper (TM) satellite image (185 x 175 km) with 28 m resolution recorded on 15 August 1994, was purchased from the EOSAT Company in June 1995. The scene, which encompassed a major portion of the Diapause Control Program Region of the High Plains, was downloaded to a SUN Spark workstation from 8-mm tape, using ERDAS Imagine® software. The image was then run through a histogram equalization process to redistribute pixel values equally throughout the available spectral range. Equalizing results in approximately the same number of pixels in each value within the pixel brightness range (i.e., 0-255).

The raster image, obtained in Universal Transverse Mercator (UTM) projection, was then exported from ERDAS into ARC/INFO, where it was georeferenced (fine tuned) to known locations of roads and towns in the form of digital line graph (DLG) vector files that were also in UTM projection. After georeferencing, the image was imported back into ERDAS for classifying. In ERDAS, a color-infrared (IR) image was displayed using bands 4, 5, and 3 (representing near IR, mid-IR, and red reflectance), as red, green, and blue colors, respectively. A signature file was prepared based on collecting signatures, called training samples, that represented cotton. Using the color-IR projected image (bands 4, 5, and 3) and techniques from the previous 2 years (i.e., knowledge of cotton growing areas and GPS-recorded points), cotton fields were identified as those areas having a very bright orange-red color pattern. All of the signatures were then merged into a single feature representing cotton.

Once the signature file was complete, a supervised classification produced an image of all pixels associated with the identified cotton reflectance. Pixel values in the image that were not associated with cotton were left unclassified (i.e., were left as part of the unidentified background). The classified image typically exhibited considerable spectral artifacts or noise that did not represent cotton. Consequently, the classified image was run through a majority filter to clump similar pixels and eliminate much of the spectral noise.

Delineation and Mapping of Cotton Fields

Classified filtered images were exported from ERDAS to ARC/INFO, and converted into vector polygon (outline) files, called coverages. Polygons representing areas smaller than 10 acres were eliminated from the coverage. In the real world, a field and a road cannot occupy the same space. However, much of the cotton coverage did not agree perfectly with the DLG road files, as parts of some fields were incorrectly represented to be located on roads. This inconsistency occurred because dirt roads that separate many of the fields do not always show up well in the image (due in part to the 28 m resolution), and because the majority filter process tends to clump similar pixels. This difficulty was resolved with a new file created by placing a 30 m buffer around all roads and highways. Whereas roads are represented by a line (having no area), road buffers are represented by an area bounded by two lines. All of the road buffer areas were then made part of the outer undefined space, and were thus also undefined. This undefined area was intersected with the cotton coverage to produce a new file in which all fields were separated from each other by the width of the road buffer, and were thus brought into agreement with the DLGs. Images were then clipped to the spatial dimensions of each of 14 counties in the High Plains, which covered as far Northeast as Briscoe County and as far Southwest as Andrews County. Maps were made by overlaying cotton field coverages with DLGs and adding appropriate documentation.

Vegetation Analysis

Identification of general vegetation was conducted in six counties of the Southern Rolling Plains of Texas, using a GPS device. The device allowed up to 250 coordinates to be collected at a given time, and as they accumulated, data were downloaded to a PC computer. Signature files (i.e., pixel samples) were constructed for all habitats based on GPS-recorded locations throughout the Southern Rolling Plains. Satellite images, which were obtained during either spring or late summer, were then classified based on appropriate signature files, in which all pixels were classified (i.e., no area was left unidentified). Images were then exported to ARC/INFO and converted into polygon vector files. For all of these coverages, habitat type was carried over from the original signature files. Thus, it was relatively easy to extract and quantify polygons associated with each type of vegetation. Locating and quantifying areas potentially subject to high infestations in the spring was accomplished by placing one half mile (0.805 km) buffers around the polygons of a given vegetation type. This approach is especially appropriate from a landscape or regional management perspective, because the focus is on regions of potential infestation rather than individual fields. Relative susceptibility to invading weevils within these buffer areas should be a function of the quality of the adjacent habitat (i.e., the degree to which the habitat insulates weevils against ambient winter temperatures), thereby influencing weevil winter survival and the numbers that emerge the following spring (Eq. 1).

$$S_b = e^{-\alpha Dh^2} \tag{1}$$

where:

S_b = susceptibility of a given buffer area to weevil infestation
D_h = habitat quality (cumulative negative °D), specific to each type of habitat)
α = a modifying coefficient (0.0035) that allows the function to approximate recorded values of proportional survival, averaged for data collected over a 15-year period

Because winter mortality only becomes a significant factor at temperatures below 0 °C (Parajulee et al. unpublished), habitat quality can be quantified as the total number of negative degree-days (°D < 0 °C) accumulating between 1 September and 31 May of the next year. In general, the better the insulating capacity of the habitat, the lower the number of cumulative negative °D, relative to ambient. Data indicate that leaf litter from oak and other broadleaf hardwoods provide the best insulation for weevils against severe cold. For example, the number of cumulative negative °D would be fewer in an oak habitat than in a mixed shrub, and fewer in a mixed shrub than in grassland or fallow, which should be close to ambient. Buffer areas will differ in susceptibility based on the type of habitats they represent (e.g., mixed shrub versus oak-broadleaf). In areas where habitat buffer areas overlap, susceptibility will be based on the buffer belonging to the best habitat (i.e., that which provides the best insulation, resulting in the highest relative survival). Algorithms have been developed to calculate Dh (negative °D) for each of the major habitats, given the ambient temperature (Parajulee et al. 1995).

Results and Discussion

Image Classification - Cotton Recognition

By mid-August, irrigated cotton was well developed and identified as a bright orange-red pattern in the color-IR image. Using bands 4, 5, and 3 as red, green, and blue, this color pattern is representative of high reflectance in the near and mid-IR part of the spectrum (Bands 4 & 5), and heavy absorption (low reflectance) in the red part of the spectrum (Band 3). Cotton reflectance of Band 4 (near-IR) for a section of Crosby County is illustrated in Fig. 1. After classification, all pixels sharing the identified signature were displayed as a single category for cotton. Using a majority filter allowed much of the spectral noise to be removed (Fig. 2). In this figure, the white background was left unidentified, thus rendering only two categories; classified cotton versus unclassified matrix. This scheme was very convenient for the next step of the process, exporting to ARC/INFO and production of a cotton polygon vector file.

Delineation and Mapping of Cotton Fields

In Fig. 2, a single polygon may appear to represent several fields merged together due to the inability of the image to distinguish many of the dirt roads between fields, and the subsequent filtering process that tends to clump similar pixels. This problem was substantially resolved by producing a new file from intersecting the field coverage with DLG road buffers (Fig. 3). In the figure, roads (from the original DLG) are displayed running between the newly bisected cotton fields. While not a perfect representation of all fields (some fields within road boundaries remain merged), given the savings of time and expense, this method appears to be extremely useful for identifying cotton fields. Moreover, minimal on-screen editing is necessary to produce a map-worthy product, especially when compared to the heads-up digitizing of the previous year. Maps of cotton fields and associated spatial data (e.g., location, area, field number, etc.) were then produced in ARC/INFO for each of 14 counties on the High Plains using cotton coverages and DLGs. Computerized data files of all coverages and DLGs were transferred directly into PCG data bases via FTP.

The accuracy of the cotton field maps has not yet been tested, although reports from field observations suggest that there was a high level of agreement between irrigated cotton and mapped fields. However, dryland cotton which had not produced enough vegetation to be recognized in the image, and fields behind clouds or in cloud shadows, could not be accurately mapped. These fields were mapped using GPS devices in land vehicles. With the satellite image-based field maps as references, this process of ground-truthing was much faster, more efficient, and required dramatically fewer worker hours than in past years. Presently, efforts are underway to summarize 1994 data from the Agricultural Stabilization & Conservation Service (ASCS) on the shapes and locations of cotton fields in a 25 square mile section in each of 5 counties on the High Plains. These data will facilitate comparisons with identical regions represented in the image-based maps, which will allow us to quantify map accuracy.

Vegetation Analysis

Over 450 locations, involving 12 habitat types, were collected throughout the Southern Rolling Plains of Texas with a GPS device during a 6-day period in the spring of 1995. These data points were used to construct habitat signature files on all habitats for several images corresponding to this region. Supervised classifications were conducted using ERDAS Imagine® in which all of the area within each image was classified. Classified images were then converted to polygon vector files in ARC/INFO. Each polygon in the vector coverage contained attribute information relating to the original habitats. This habitat identity therefore allowed polygons to be extracted as subgroups from the original coverage, establishing specific coverages for each type of habitat. Habitats included oak woodland, oak-mesquite, broadleaf-riparian, mesquite, mixed shrub, juniper, grassland, cropland, bare ground, water, roads, and urban.

Calculation of buffers representing 0.805 km (maximum distance for weevil spring migration) around each habitat enables identification of potential boll weevil emergence zones. These buffer zones can then be quantified as to the relative susceptibility to weevil infestations (Eq. 1). As an example, habitats such as oak woodland can be isolated and buffered to delineate areas of potential susceptibility to high weevil infestations in the spring, as shown in a representation of a section of Concho County (Fig. 4). Presently, vegetation coverages have been produced for 4 counties in the Southern Rolling Plains, with 3 more to be completed shortly. After extraction and buffering of all potential weevil habitats, maps of all areas can be produced and infestation probabilities quantified.

References

1. Bottrell, D. G., J. R. White, D. S. Moody & D. D. Hardee. 1972. Overwintering habitats of the boll weevil in the Rolling Plains of Texas. Environmental Entomology 1: 633-638.

2. Parajulee, M. N., L. T. Wilson, D. R. Rummel, S. C. Carroll, & P. J. Trichilo. 1995. Insulating values of different boll weevil overwintering habitats: Implications of using habitat versus ambient temperature in developing predictive models. Proceedings Beltwide Cotton Insect Research & Control Conference. Vol. 2: 1029-1031.

3. Rummel, D. R. & P. L. Adkisson. 1970. Distribution of boll weevil-infested cotton fields in relation to overwintering habitats in the High and rolling Plains of Texas. Journal Economic Entomology 63: 1906-1909.

4. Trichilo, P. J., L. T. Wilson, & R. K. Haldenby. 1995. Use of satellite images to optimize regional management strategies. Proceedings Beltwide Cotton Insect Research & Control Conference. Vol. 2: 804-808.

5. Wilson, L. T. P. J. Trichilo, S. Bozkurt, P. K Tapadiya, D. R. Rummel, S. C. Carroll, J. E. Slosser, T. W. Fuchs, & R. E. Frisbie. 1994. Impact of boll weevil migration on selection of eradication strategies for Texas. Proceedings Beltwide Cotton Insect Research & Control Conference. Vol. 2: 975-980.

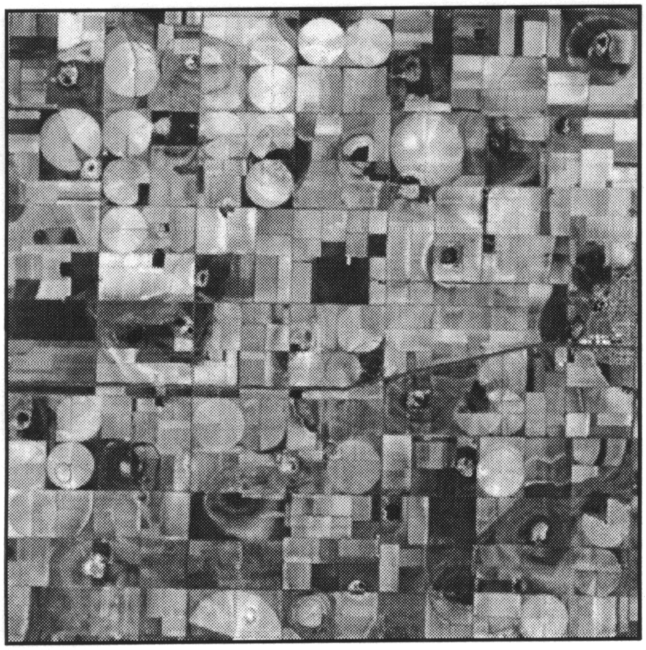

Fig. 1. Thematic Mapper (TM) satellite image showing Band 4 (Near-IR) reflectance of a section of Crosby County. Scale: 1/2" = 1 mile.

Fig. 2. Results of supervised classification of cotton reflectance using TM image of Crosby County represented in Fig. 1.

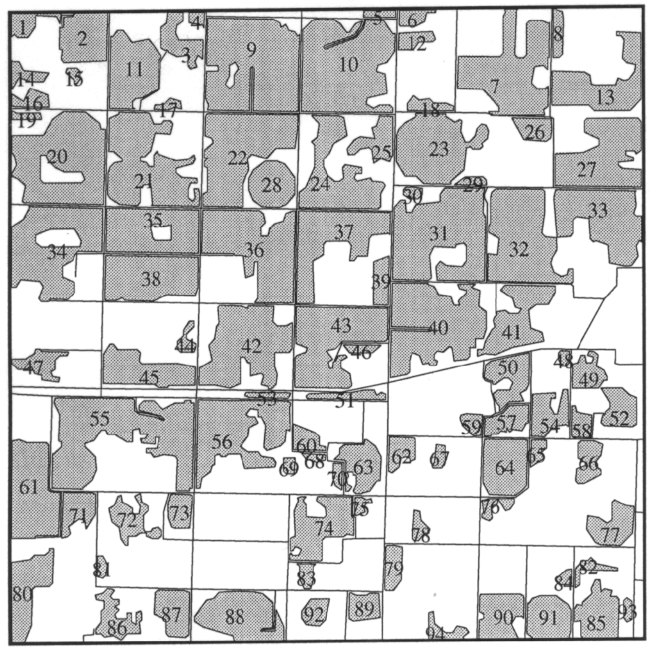

Fig. 3. Map of cotton fields bisected by roads in digital line graph (DLG) format for same section of Crosby County represented in Fig. 1.

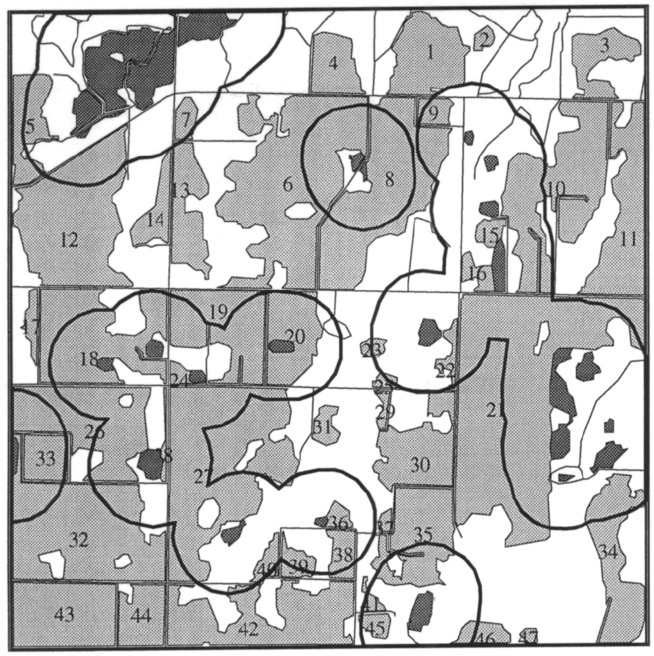

Fig. 4. Section of Concho County showing areas of cropland (light gray) with high susceptibility for spring boll weevil infestations, based on 1/2-mile buffers (black-lined circles) placed around oak habitats (dark gray).