

Kansas State University Libraries

**New Prairie Press**

---

Conference on Applied Statistics in Agriculture

2004 - 16th Annual Conference Proceedings

---

## INFORMATION TECHNOLOGIES AND THE DESIGN AND ANALYSIS OF SITE-SPECIFIC EXPERIMENTS WITHIN COMMERCIAL COTTON FIELDS

J. L. Willers

G. A. Milliken

C. G. O'Hara

J. N. Jenkins

*See next page for additional authors*

Follow this and additional works at: <https://newprairiepress.org/agstatconference>



Part of the [Agriculture Commons](#), and the [Applied Statistics Commons](#)



This work is licensed under a [Creative Commons Attribution-Noncommercial-No Derivative Works 4.0 License](#).

---

### Recommended Citation

Willers, J. L.; Milliken, G. A.; O'Hara, C. G.; and Jenkins, J. N. (2004). "INFORMATION TECHNOLOGIES AND THE DESIGN AND ANALYSIS OF SITE-SPECIFIC EXPERIMENTS WITHIN COMMERCIAL COTTON FIELDS," *Conference on Applied Statistics in Agriculture*. <https://doi.org/10.4148/2475-7772.1151>

This is brought to you for free and open access by the Conferences at New Prairie Press. It has been accepted for inclusion in Conference on Applied Statistics in Agriculture by an authorized administrator of New Prairie Press. For more information, please contact [cads@k-state.edu](mailto:cads@k-state.edu).

---

**Author Information**

J. L. Willers, G. A. Milliken, C. G. O'Hara, and J. N. Jenkins

# Information Technologies and the Design and Analysis of Site-Specific Experiments within Commercial Cotton Fields

J. L. Willers<sup>1†</sup>, G. A. Milliken<sup>2†</sup>, C. G. O'Hara<sup>3†</sup>, and J. N. Jenkins<sup>1</sup>

<sup>1</sup> Genetics and Precision Agriculture Research Unit  
USDA ARS, Mississippi State, MS

<sup>2</sup>Department of Statistics, Kansas State University

<sup>3</sup> GeoResources Institute, Mississippi State, MS

(<sup>†</sup>These authors shared equally in the work.)

**Abstract.** Information products derived from multi-spectral remote sensing images, LIDAR elevations, or data products from other sensor systems (soil electrical conductivity measurements, yield monitors, etc.) characterize potential crop productivity by mapping bio-physical aspects of cropland variability. These sensor systems provide spectral, spatial, and temporal measurements at resolutions and accuracies describing the variability of in-field, physical characteristic phenomena, including management practices from cropland preparation, selection of crop cultivars, and variable-rate applications of inputs. In addition, DGPS-equipped (differential, global positioning system) harvesters monitor yield response at closely spaced, geo-referenced points. Geographic information system and image processing techniques fuse diverse information sources to spatially characterize cropland, describe management practices, and quantify the variable yield response. Following fusion of information sources, effectiveness of spatially applied management practices may be evaluated by designed experiments assessing impacts on yield caused by geo-referenced relationships between (1) uncontrollable spatial components (the environment) and (2) controllable management practices (cultivar selection, fertility management, herbicide, insecticide, and plant growth regulator applications, etc.). These kinds of experiments can be designed because farming equipment can be computer controlled through DGPS giving farmers the ability to continuously change applied treatments for many farming operations. A mixed linear model involving both uncontrollable and controllable management attributes attached as spatial descriptors to yield monitor points evaluates effects of management practices on yield. An example based upon cotton production demonstrates the methodology. Additional strategies for designing studies in commercial cotton fields involving spatial information are discussed.

## I. Introduction.

Several years of attempts to build designed experiments to evaluate site-specific (precision agriculture) management (Dupont *et al.* 2000) of cotton in commercial fields have demonstrated the need for new analytical methods. Traditional experimental designs (CRD, RCB, Split-Plot, Lattices, etc.) do not perform well within large commercial fields because it is difficult for farm operators to impose planned treatments to randomized, replicated, small-sized

plots and complete other necessary farming operations. Traditional designs also generally fail to incorporate spatial characteristics of harvested sub-samples within and among (small) plots (experimental units). Defining representative units of replication (Mead 1988; Milliken and Johnson 1992) and their size in commercial fields is also difficult. If experimental units are arbitrarily defined, deciding how to assign site-specific (or even control) treatments is problematic due to crop growth being influenced by uncontrollable sources of spatial and temporal variability. Uncontrollable variability results from interactions among weather, soil types, and elevation causing variation in water utilization and assimilation of nutrients available for plant growth. Influences of past management practices may be difficult to prevent (e.g., choice of herbicides applied the previous year, land leveling operations, etc.). These sources of uncontrollable agricultural, edaphic, or environmental influences often mask planned treatment effects if a traditional, small plot based experiment is implemented in a large commercial field. Similarly, the spatial application of one crop input is often not coincident in time or space with other precision agriculture practices in large fields, complicating assessment of joint effects when using small plots. Further, responses of spatial management practices may differ for other fields (or farms) making it difficult to define the inferential space (Stroup 1989) of results from analysis of small plots. With experimental units (plots) of the same size and shape, traditional techniques of random assignment of treatments and increases in replications do not solve these problems in commercial fields.

Therefore, there are opportunities to develop new designs for experiments in commercial agricultural fields where precision agriculture is practiced. These 'topological experimental designs' (TED) geographically apply traditional treatments and site-specific practices to different zones in the field. These zones are not artificial, symmetrical arrangements of small plots. Topological relationships among various layers of spatial information (including uncontrollable and controllable sources) allow DGPS (Kennedy 1996), variable-rate (VR) farm equipment to assign different treatments to various polygons (zones) at geographic locations in fields. Several forms of regression models analyze these spatial designs. At least one approach is to build a mixed linear model of yield (omitting yield data from site-specific management zones) as a function of traditional management practices, including environmental factors (or other site characteristics) as covariates. Using covariates, this model predicts yield at the coordinates of yield monitor points within each site-specific zone as if the traditional management practice had been applied. Diverse covariate values of field topography, measured by remote sensing, are extracted by GIS analysis at the geographic coordinates of yield points. These predicted yields (assuming the traditional practice) are compared to the actual site-specific yields to compare effects of management. This general methodology can be extended to more than one new management practice.

The objective of this paper is to describe the integration of agricultural, GIS, remote sensing, and statistical disciplines by demonstrating analysis of an unreplicated cotton cultivar trial that includes one spatially variable application of a plant growth regulator (PGR). We illustrate spatial processing of different data types using geographic information system (GIS) and statistical techniques. Summary results illustrate further applications of the methodology. Additional strategies for designing studies in commercial fields utilizing spatial information are discussed.

## II. Methods.

Designing, completing, and analyzing a site-specific experiment in a commercial cotton field requires teamwork throughout three distinct phases. The first part is fieldwork, requiring a producer's cooperation and other farm resources. However, in order to apply site-specific management practices, spatial information has to be acquired and processed into 'farm ready' formats. Thus, the second part is participation of service providers, who acquire, process, and deliver spatial data products to the farm. This exchange occurs many times throughout the season and continues through harvest. Once harvesting is finished, the third part is analysis involving spatial and statistical processing, iterative analysis and model building procedures, and interpretation.

**The Field Study:** During the 2003 production season, the case study TED was established in the western part of a 160 acre (64.75 ha) field at Good's Longview Farm, Noxubee County, several miles east of Macon, MS. The row spacing was 30 in (0.762 m) to promote yield production while reducing soil erosion in a dryland production system. The selected study area was comprised of 17 cotton cultivars (Table 1) in a configuration of 8 rows by 0.5 mi (804 m) in length for each cultivar. A cultivar was planted using an 8-row planter where hopper bins were thoroughly emptied between changes to the next cultivar. Planting occurred within 2 days of 22 April 2003 with the planter set at a seeding rate 48,521 seeds/ac (120,101 seeds/ha) for all cultivars. Other planting requirements were the application of N-Sol<sup>®</sup> at 10 gal/ac (93.7 l/ha), Caporal<sup>®</sup> at 1 pt/ac (1.2 l/ha), Ammo<sup>®</sup> at 1.3 oz/ac (0.07 l/ha), GammoxonMax<sup>®</sup> at 0.7 pt/ac (0.8 l/ha), Temik<sup>®</sup> at 6 lbs/ac (2.72 kg/ha), and TSX<sup>®</sup> (fungicide) at 9.4 lbs/ac (4.26 kg/ha). The total area involved was approximately 22 acres (8.88 ha). The cultivar trial is an unreplicated strip-plot design (Milliken and Johnson, 1992) and was the first management factor of the TED.

The first application of PGR (mepiquat pentaborate) to these cultivars was a broadcast application (5 oz/ac (0.37 l/ha)) applied 26 June 2003, using a tank mix applied by a John Deere<sup>®</sup> Model 6800 sprayer. The second and last application was a variable-rate PGR prescription (the second management factor of the TED) applied 18 July 2003 using a CASE-IH<sup>®</sup> Model SPX 2130 sprayer equipped with a Mid-West Technologies<sup>®</sup> (Springfield, IL) 6600 variable-rate controller and injection system. The PGR was applied to each cultivar using a boom width of 20' (6.09 m). The water carrier channel was set to apply 30 gal. of water/ac (281 l/ha). The four rates of PGR applied were 0, 4, 6, and 8 oz/ac (0, 0.29, 0.43, and 0.59 l/ha) according to a prescription map built from a multi-band image composite acquired 9 June 2003.

Harvesting took place 16 September 2003, using a John Deere<sup>®</sup> Model 9965 4-row cotton picker equipped with an Ag-Leader<sup>®</sup> (Ag Leader Technologies, Ames, IA) 3000 Pro Yield Monitor. Differential, global positioning information was supplied to the yield monitor using the OmniSTAR<sup>®</sup> DGPS subscription service ([www.omnistar.com](http://www.omnistar.com)). The yield monitor supplies several attributes of each yield point, including load\_id, time and location. Yield data were collected once every two seconds and (as harvester speed varied) the distance between logged points slightly differed (3.30 – 14.60 ft (1.00 – 4.45m)).

The cooperating farm conducted all field operations (including harvest) except for the second variable-rate application of PGR in July, accomplished in cooperation with the

Department of Weed Science, Mississippi State University. All other inputs were blanket applications applied to all cultivars and were not included in the analysis.

**Imagery and LIDAR Acquisitions:** Endogenous traits of the field can be measured by remote sensing (RS) and LIDAR (LIght Detection And Ranging) systems on aircraft (or ground based equipment). Imagery and LIDAR data were acquired in cooperation with the GeoSpatial Resources Institute (GRI) at Mississippi State University.

Multi-spectral images were acquired by the GeoVantage camera system ([www.geovantage.com](http://www.geovantage.com)) flown by GeoData Airborne and Mapping, Inc. (Weir, MS). Multi-spectral imagery was obtained at pixel resolutions of 0.5 m<sup>2</sup>. The center bands of the imagery were: 450 nm (blue), 550 nm (green), 650 nm (red), and 850 nm (near infrared (NIR)). Three image dates utilized were bare ground imagery from 15 January 2002, and two in-season images acquired on 9 June 2003, and 17 July 2003. Image data products were delivered as geo-referenced mosaics in the Universal Transverse Mercator (UTM) projection, Zone 16, using the North American 1983 (NAD83) datum.

LIDAR data were acquired 11-12 May 2003 (0.5 m<sup>2</sup> horizontal resolution and 8.1-10.7 cm RMSE (vertical resolution)) by Earthdata, Inc. (Frederick, MD) using the ALS50 Airborne Laser Scanner developed by Leica Geosystems ([www.gis.leica-geosystems.com](http://www.gis.leica-geosystems.com)). LIDAR data were delivered in point theme format to GRI, Mississippi State University, but were not available for the 2003 growing season. The initial three-dimensional coordinates were compiled in a mass point file of x, y, z on the UTM projection. Ellipsoidal heights were converted to NAVD88 using Geoid 99. These data were converted to a digital surface model (DSM) with 1 m<sup>2</sup> ground resolution using specially developed scripts in the ArcInfo<sup>®</sup> macro language. The coordinate system of the DSM was the UTM projection, Zone 16, using the NAD83 datum.

**Geographic Information System Analyses of Data:** Using ERDAS Imagine<sup>®</sup> 8.7 or ArcInfo<sup>®</sup> Workstation, multispectral imagery was converted from its native bitmap format into raster grids. Information in the imagery was processed using unsupervised classification procedures (Pouncey *et al.* 1999; Richards and Jia 1999) emphasizing the normalized vegetation index (NDVI). Generally, the NDVI varies over the growing season, particularly before peak bloom, due to interactions among soil type, weather, and management practices. Focal processing (Anonymous 1997) prepared the variable-rate PGR prescription file (using the NDVI values of 9 June) and is an operation involving a grid where the output value of the currently processed cell depends upon the values of its neighboring cells in the defined focal area. In summary, the in-season GIS/RS processing steps include (1) evaluating the NDVI value for each cell in the grid form of the image, (2) focal smoothing and re-grouping of the NDVI values into three classes, (3) re-sampling of the original image spatial resolution (1.64 ft (0.5 m)) to a spatial resolution matching the boom width (20 ft (6.096 m)) of the sprayer, (4) assignment of a PGR rate (according to the class codes established in step 2) to each cell in the sprayer grid, (5) imbedding any 'overlap plots' to serve as controls (0 PGR) for future analysis, and (6) conversion of the prescription grid into a polygon coverage. This coverage was converted to a shapefile (latitude and longitude in decimal degrees in the NAD83 datum) for proper translation by the sprayer controller while applying the variable-rate PGR to the cotton.

Once harvest was completed, it was necessary to establish the data structure (a dBase IV table derived from a shapefile) characterizing the TED. The data structure is assembled by one or more GIS procedures applied to the yield file obtained by the cotton yield monitor, where records are initially sorted by load number (two loads per cotton cultivar) and location/time stamp. Harvesting of cultivars occurred from west to east. These techniques (see Anonymous 1997; Theobald 2003 for details) generally involve converting yield point and prescription shapefile themes into arc coverage formats, converting raster images into raster grids, and geographic sub-setting (masking) of spatial information to only the area of interest (AOI). Focal processing of yield data within a cultivar strip was necessary in order to align (east-west) adjacent picker passes, because the two picker passes differed in direction of travel (north and south) while harvesting a strip. Additional focal processing converted the yield point theme into a surface (or grid) to depict yield response across all cultivars. Yield points are originally in a spherical format (latitude and longitude as decimal degrees in the WGS84 datum); therefore, GIS processing projected these data to the UTM, Zone 16, NAD83 planar coordinate system (Bugayevskiy and Snyder 1995).

GIS processing derived several topographic (environmental) variables from raster images and LIDAR DSM to be covariates in the regression model. One new feature derived from the images represented relative NDVI change between January 2002 and June 2003. This grid of NDVI change captured those regions of the field where the cotton crop first emerged after planting. These plants had the most developed leaf canopy at the time of the June acquisition. Another covariate was the NDVI status on 17 July 2003. (While doing these GIS operations in preparation for analysis, it was found necessary to geographically adjust (ca. 3-5 m south to north) the June and July mosaics using ERDAS Imagine<sup>®</sup> and ESRI ArcInfo<sup>®</sup>. The 15 January 2002 image did not need geometric corrections.) Additional grids of hydrology covariates were derived from the DSM grid.

Once all data layers were prepared, the final, key GIS processing step bilinearly interpolated an estimate of the PGR rate and other covariate values at the UTM coordinates (easting, northing) of each yield point. Intersecting yield point coordinates with each vector or grid theme extracted a location specific 'spot' parameter value attached to every yield record. Variables (see Appendix 1 for acronyms) in the yield attribute table prior to regression analysis were: fieldname, load\_id, yield monitor serial number, crop type (cotton), record number, track degree heading of the harvester, swath width, distance (ft) to next point, time duration between data logging of points, differential status of the GPS unit, date, time, cotton flow (analog output), seed cotton yield (lbs/ac), PGR rate as a 'grid\_code' (0 = 0 oz, 1 = 4 oz, 2 = 6 oz, and 3 = 8 oz), the x-coordinate (easting), y-coordinate (northing), slope (percent), aspect (a categorical number corresponding to compass direction), flow accumulation (m<sup>2</sup>), concavity/convexity (negative to positive real numbers), Euclidean distance from a synthetic stream (m) network, and elevation (m). The attribute table also included the following NDVI values at different times of the year: NDVI\_1 = 15 January 2002, NDVI\_2 = 9 June 2003, NDVI\_717 = 17 July 2003, and NDVI\_Dif is the grid subtraction of NDVI\_2 - NDVI\_1. Plant *et al.* (2001) and Willers *et al.* (1999) discuss other NDVI applications in cotton production.

**Statistical Analyses of Data:** The treatment structure of the case study TED was the 17 cotton cultivars and the 4 levels of the variable-rate PGR application. These factors represent the

controllable management attributes. Additional field information affecting data preparation and analysis is (1) a buffer strip of 40 rows was planted on the west border of the trial, (2) an extra load was logged on the east side outside of the study area, and (3) the west to east order of the cultivar strips (ascending Plot ID, Table 1) was established by (grower and research) opinions. These decisions determined the 9 stacked gene cultivars (Plots 5-13, Bollgard<sup>®</sup> and Round-up Ready<sup>®</sup>) would be planted in one 'block' and the 8 single gene cultivars (Plots 14-21, Round-up Ready<sup>®</sup> only) would be planted in another 'block'. The PGR application in July also progressed one cultivar strip at a time from west to east. The design structure of the TED was comprised of 17 strips of 8 rows harvested in 2 loads from 2 passes per cultivar by a 4-row cotton picker and the various spatial management zones for the variable-rate PGR application.

Once the data structure of the TED was available in shapefile format, the attribute table was imported into SAS<sup>®</sup>. A mixed regression model (Gotway and Stroup 1997; Littell *et al.* 1996), or mixed analysis of covariance model (Milliken and Johnson 2001) described seed cotton yield as a function of cultivar, PGR rate (Rate), and the site characteristic variables. Additional variables computed in a SAS<sup>®</sup> data step were cultivar name, load\_id, pass (or direction 'north' or 'south' traveled by the harvester within a cultivar strip),  $\log_{\text{fac}} = \log(\text{fac}+1)$ ,  $\log_{\text{euc}} = \log(\text{euc})$ ,  $\text{one\_cvx} = 1/\exp(\text{cvx})$ ,  $\text{one\_ndvi} = 1/\exp(\text{ndvi\_dif})$ , and  $\text{index} = \log(\text{fac}+1)/(\text{slo})$ . The UTM coordinates of each yield point were also adjusted to a new origin (0 m, 0 m) in the southwest corner of the field, where new\_y is the northing value minus 3666170.70 m (a range from 0 to 800 m) and new\_x is the easting value minus 361084.77 m (a range from 0 to 135 m). A variable called RateClass was generated to denote areas of the field where each level of PGR was applied. The intersection of PGR rate class boundaries and cultivar strips defines experimental units for the treatment structure.

The initial model included explanatory variables as well as interactions of explanatory variables with cultivar. Rate and Rate<sup>2</sup> as well as DSM and DSM<sup>2</sup> were used in the modeling and were included in interactions with cultivar and with other explanatory variables. Load\_id within cultivar was used as a random effect. The complete data set initially consisted of 8745 data points and included data from 10 load\_ids not belonging to the study AOI. These data from load\_id = L1-L9 and L55 were deleted, as well as observations with cvx < -900 and slo < -900 (missing values occurring during GIS pre-processing), leaving a data set with 7465 observations. There were some outlying yield monitor data points occurring at the ends of the cotton rows as well as during times when the harvester went through a water way or ditch. The first and last 30 feet of each harvest load were deleted from the data set since the yield monitor often produced extreme readings in these portions of the load. These values were caused by changes in speed, etc., as the harvester reached the end of a load, turned around, and started the next load for a cultivar strip. The data set now consisted of 6550 data points. The yield monitor data are quite variable from one time point to the next because of the process of measuring yield, thus a Loess filter smoothed the data to lessen extreme values. The focal and loess processing steps jointly adjusted seed cotton yield values to create a dependent variable named 'say' (or smoothed adjusted yields). The correlation between focal-smoothed (orig\_say) and twice-smoothed predicted values (say) was 0.97039. Thus, little information was lost by the loess smoothing process, but several very extreme values were shrunken.



The first analysis step was to fit the whole model to this filtered data set and compute residuals. Observations that had outliers greater than 1000 lbs. seed cotton in magnitude were removed from the data set, leaving 6307 observations. Next, a stepwise deletion process simplified the model where variables and interactions were removed, one at a time, until all remaining variables were significant at 0.05 or less. The exception was for variables involved with significant interactions with cultivar, as deleting them from the model would not improve the fit. Before the model evaluated cultivar effects, residuals were examined one more time and observations with residuals greater in magnitude of 800 lbs. were deleted, leaving the final data set with 6301 observations. The distribution of residuals was very symmetric, but the distribution tails were quite long, even after deleting observations with large residuals. An even more aggressive filtering process may provide residuals that have tails more like a normal distribution. The following SAS<sup>®</sup> statements fit the filtered data for this case study TED:

```
proc mixed data=predres1; where -800<res1<800;
title 'analysis with site characteristics';
class pass cultivar load_id;
model say = cultivar rate rate*rate cultivar*rate
cultivar*rate*rate Slo LogFac Cvx logEuc Dsm
Ndvi_dif index one_cvx ndvi_717 logfac*cvx Dsm*cultivar
Dsm*rate Cvx*cultivar*rate logEuc*cultivar*rate
ndvi_717*cultivar*rate LogFac*cultivar*rate*rate
logEuc*cultivar*rate*rate Dsm*cultivar*rate*rate
ndvi_717*cultivar*rate*rate LogFac*rate
cvx*cultivar*rate*rate Ndvi_dif*cultivar*rate*rate
Slo*cultivar*rate*rate index*cultivar*rate
Ndvi_dif*cultivar*rate Slo*rate logfac*dsm slo*dsm
logfac*ndvi_dif one_ndvi dsm*dsm*cultivar
/outp=preds solution;**say is smoothed yield;
random load_id/subject=cultivar;
repeated /type=sp(GAU)(new_y)
subject=load_id(rateclass*cultivar) local;
lsmeans cultivar/at means diff;
lsmeans cultivar/at rate=0 diff;
lsmeans cultivar/at rate=2 diff;
lsmeans cultivar/at rate=4 diff;
lsmeans cultivar/at rate=6 diff;
lsmeans cultivar/at rate=8 diff;
```

The `random load_id/subject=cultivar` statement specifies that `load_id`'s within a cultivar are random effects. The `'repeated / type = sp(GAU)(new_y) subject = load_id(rateclass*cultivar) local;'` statement specified that residuals within a harvest pass of a cultivar intersected with a rate class are spatially correlated. The Gaussian spatial correlation model provided the smallest AIC value over all other spatial correlation structures. The `'lsmeans cultivar/at means diff;'` statement requests that adjusted cultivar means evaluated at average values of the response variables be computed. These means can be thought of as being computed by predicting the response of a cultivar at each yield monitor point in the entire field and taking the mean of those points. That is, these least square

means are predictions of the mean yield of the cultivars as if each had been grown on the complete field. The other lsmeans statements provide predictions as if each cultivar would have been grown on the complete field using the specified blanket rate of PGR. The 'outp=preds' part of the model statement provides a data set with a predicted value for each observation.

For comparison purposes, another analysis of the final data set was completed without site characteristics, but used the PGR rate as a continuous variable entered as both a linear effect and a quadratic effect. The following Proc Mixed code fit the data set:

```
proc mixed data=predres1;where -800<res1<800;
title 'model without site characteristics';
class load_id cultivar;
model say= cultivar rate cultivar*rate
cultivar*rate*rate/ddfm=satterth;
random int /subject=load_id(cultivar) ;
repeated /type=sp(sph)(scaled_new_y) subject =
load_id(rateclass*cultivar) local;
lsmeans cultivar/at means diff;
lsmeans cultivar/at rate=0 diff;
lsmeans cultivar/at rate=2 diff;
lsmeans cultivar/at rate=4 diff;
lsmeans cultivar/at rate=6 diff;
lsmeans cultivar/at rate=8 diff;
```

The 'repeated /type=sp(sph)(scaled\_new\_y) subject = load\_id (rateclass\*cultivar) local;' statement provides a spatial correlation (using the spherical option) among yield monitor measurements within a load\_id of a cultivar intersected with the PGR rate class.

Tables of LSD values for both types of models were also determined. Both models estimate effects of a 2 oz PGR rate which was not applied as an actual rate in the field.

### III. Results.

**Geographic Information System/Remote Sensing Data Products:** Images and LIDAR data were acquired for the focus field and various data products were derived. A site-specific application was made, field observations were collected (not discussed here), and the crop harvested with equipment monitoring continuous yield rates as discrete points. These operations provided a great deal of information that must be extensively processed. Therefore, the GIS objectives in the analysis of this TED were:

- ➊ Extract information from remote sensing data to support site-specific application of a variable-rate application of PGR.
- ➋ Implement 'controls' (named as overlap plots) in the field and embed them in the planned PGR application to automate their consideration; thereby removing the need to manually turn off the sprayer when 'flagged control areas' are encountered.

- Create highly accurate data products from high-resolution imagery and DSM data to explore interactions of topography, hydrology, and other topology relationships on cotton growth and yield.
- Process and clean data products to address uncertainty, normalize data, and remove or minimize unwanted noise.
- Utilize coordinates of monitored yield points to extract thematic information of field topography with a sample density sufficient to examine micro-variability effects on cultivar performance and PGR application effectiveness.

Summarized in Figure 1 are intermediate data products (extracted from the raster image of 9 June 2003) used to build the variable-rate PGR prescription. Focal processing resolved the continuous NDVI values into an intermediate grid of three comprehensible rate classes at the same spatial resolution of the original multispectral image. Focal processing also ‘rasterized’ the information into another grid (Figure 1, far right) whose cell geometry matched the boom width of the sprayer that applied the PGR to each cultivar, 8 crop rows at a time along its travel path.

The first data product (Figure 2) derived from the LIDAR digital surface model (DSM) was a computer-generated hillshade model of elevation providing perspective of the shadows caused by the sun at a particular angle and compass direction. The hillshade surface is a visualization tool (Theobald 2003) providing analysts with graphical information about field topography and was not part of the statistical modeling process. The DSM also provided several hydrological data products (Figure 2). The concavity/convexity grid is a measure of whether water would pond or runoff at a particular cell in the grid and was highly relevant to the statistical modeling work.

Yield data points have spatial coordinates and provide measurements about crop yield at that location. These point data provide the best resolution for studying relationships among topographic variables influencing growth, the spatially variable management practice (PGR rate), and cotton cultivar effects on yield. (These point data ‘on the ground’ are actually small areas defined by the swath width and 2 sec recording time of the picker.) The yield data were highly variable and noisy between north and south picker loads within a cultivar. This problem was first addressed through focal processing of the point coverage within the cultivar strips (see example for loads 30 and 31, Figure 3). Next the yield point data (Figure 4, left) were transformed to a grid (Figure 4, middle) and spatially averaged over the area of study using a second focal processing step to create a final surface grid (Figure 4, right) of seed cotton yield.

GIS functions applied at locations of yield data coordinates extracted attributes from various thematic layers to create an output file (one record per yield point) that captured spatial relationships among management practices and topography. Intersecting yield data points with various thematic layers at different spatial scales is illustrated in Figure 5.

**Statistical Analyses:** The general, linear mixed model analysis of the TED data table was insightful. To complete the first part of the analysis, site characteristics (topography themes) and variable-rate PGR application (*i.e.*, grid\_code) classes were used to model each cultivar’s yield response at each yield point coordinate within every ‘load\_id’. For example, for cultivar 215, Figures 6a and 6b compares focal data (orig\_say), focal-loess smoothed data (say), and predicted

data for picker loads 10 and 11. These data are displayed using adjusted Northing coordinate information from south (0 m) to north (ca. 800 m). These graphs show that modeling yield response using site characteristics of each different yield point provides a good fit. The statistical model predicted seed cotton yield for the other 16 cultivars (data not shown) with equally comparable results. Table 2 summarizes these models by showing the correlation coefficients between predicted values and either loess-and-focally smoothed (say) or focally smoothed (orig\_say) yield monitor data by cultivar. The correlation coefficients range from 0.614 to 0.797 for twice-smoothed data and from 0.592 to 0.775 for only focally smoothed data. Using random and repeated statements with a Gaussian spatial structure, the load\_id(cultivar) variance component was estimated as 13.96, sill estimate was 32,465, range estimate was 8.0767, and nugget effect estimate was 0.0146.

Statistical models built on the basis of topographical site characteristics and PGR prescription rates obtained seed cotton yield predictions at the coordinates of other cultivar yield points as if a selected cultivar were grown for the entire study area (and not only the strip where it was actually grown). The SAS<sup>®</sup> output of this modeling process produced additional tables keyed to the original data table through the UTM coordinates of the yield points. These predicted 'whole' field means compared performance among cultivars. The results of these lsmean statements are displayed in Table 3 where predicted yields of different PGR rates are named rate\_0, rate\_2, rate\_4, rate\_6, and rate\_8. The mean PGR rate for the field was 6.17 oz/acre providing results very similar to those for 6 oz/acre; thus, results for the mean PGR rate are not included in the table. Cultivar ranks within a PGR rate are also included (for general comparison to the cultivar ranks in Table 1, if these predicted seed cotton weights are adjusted (not shown) to lint weights).

Once a model of cultivar response at different sites was possible, other models examining cultivar responses to PGR rates could be built. It is possible to answer questions about the effect of different PGR rates on each cultivar with different combinations of site characteristics. To our surprise, analyses of spatial PGR application by cultivar were quite productive and rich (> 600 pages of tables and graphs). Differences of predicted means at each rate were computed and the mean, minimum, and maximum standard errors of the difference were determined (Table 4). The standard errors were multiplied by a t-value to provide a 0.05 LSD value for the three standard error values for each rate. Since the mixed model is a regression model, the estimated standard errors are larger for 0 and 8 oz/per acre and smaller values for 6 oz/acre, a value close to the mean rate.

In general, there are several interaction terms involving cultivar, indicating that yield models for each cultivar are different functions of the independent variables. The rate of PGR also occurs in several interaction terms (both Rate and the square of Rate), indicating effects of PGR rate on a cultivar's yield depends on both site and cultivar characteristics at the coordinates of the yield points. The analysis of variance table of fixed effects (Table 5) indicates that several characteristics interact with the cultivar yield responses, including Rate, Rate<sup>2</sup>, DSM, DSM<sup>2</sup>, Rate\*Cvx, Rate<sup>2</sup>\*Cvx, Rate\*Index, etc. These results suggest one can apply too high of a PGR rate to vigorous growing cotton residing in regions of fields that collect water and cause decreases in yield. Many of these interactions spatially correspond to the middle right region displayed in Figure 4 (any panel) showing yellow hues for seed cotton yield. The agricultural

application of simultaneous site and cultivar interactions with PGR rate classes is that spatial prescription maps cannot be generated by an elementary process. This implies the processing (Figure 1) building our PGR application map was too simplistic.

Figure 7 is a summary display of predicted values of seed cotton yield for each cultivar at 5 levels of PGR based upon a mixed model using site characteristics. Most cultivars perform better at lower levels of PGR than at higher levels of PGR. This evidence suggests that the initial 5 oz/acre blanket rate in June has cumulative effects with the spatial application of PGR rates in July. These exploratory results (total PGR amount applied at any location  $\leq 13$  oz/acre) suggest that current recommendations typically averaging about 16 oz/acre per season may be too high for managing growth of many cotton cultivars. Other optimization experiments on total PGR amounts topographically applied throughout a season need to be planned.

The above analysis employing site characteristics can be compared to a second analysis without yield point site characteristics. Using random and repeated statements with a spherical spatial structure, the estimate of the  $\text{load\_id}(\text{cultivar})$  variance component was 321.25, sill estimate was 96,110, range was 0.4678, and nugget effect was 0.9114. These estimates of covariance parameters are a lot larger for the model without site characteristics than for the model with site characteristics. The predicted yields for each cultivar evaluated at 5 levels of PGR without site characteristics are displayed in Table 6. The ranks of cultivars within a level of PGR are also included. Table 7 contains the mean, minimum, and maximum estimated standard errors of the difference between pairs of cultivar means and corresponding 0.05 LSD values. Again, estimated standard errors for 0 and 8 oz/acre are larger than those close to the mean PGR rate of 6.17 oz/acre. Standard errors of the differences are functions of the covariance parameters, thus estimated standard errors and corresponding LSD values are much smaller for analysis that incorporates site characteristics than for analysis without site characteristics. Table 8 presents the analysis of variance table of the model without site characteristics. Figure 8 summarizes predicted seed cotton yields for cultivars at five levels of PGR and similarly indicates most cultivars have depressed yields at higher rates of PGR. A few cultivars have a larger response at both 0 and 8 oz/acre than at the middle rate ranges. Such a response may be due to the 'luck of the draw' for yield points at these locations within these cultivar strips.

Two different types of general linear mixed models analyzed the case study TED and provided diverse results evaluating both cultivar and PGR performance. Of the two approaches, analyses utilizing site characteristics provided models with greater precision. The VR controller, yield monitor, and other sensors installed on aircraft generated the required spatial information needed to develop these analyses.

#### **IV. Discussion.**

The capability to apply different rates of crop inputs to different areas of fields raises the issue of how to design and analyze experiments to demonstrate treatment effects. Many authors (Conquest 2000; Hurlbert 1984; McDonald *et al.* 2000; Murtaugh 2000; Scheiner and Gurevitch 2001) discuss various difficulties in analyzing landscape level experiments from an ecological perspective. A common theme is how to establish controls for comparisons and obtain replications of treatments (see also Mead 1988). However, around the mid-1990's, yield

monitors (Pierce *et al.* 1997) and high resolution, RS capabilities (Jensen 2000; Moran *et al.* 1997; Pinter *et al.* 2003; Willers *et al.* 1999) began to appear upon the market. As result, it is now possible to address the problem of experimental design in more detail than described by Gotway *et al.* (1997) or Doerge and Gardner (1999). The major change is thinking differently about the characteristics and arrangements of plots. The ability of sensor systems to map several uncontrollable, spatially variable factors and the capabilities of VR controllers to apply (and map) spatial inputs of agrichemicals allows one to geographically establish adaptable experimental units (zones) and treatment and design structures needed to answer questions. In a TED, geo-referenced experimental units do not need to be symmetrically arranged, similar in size, or in close proximity to one another, and should not be randomly assigned to treatments. Experimental units are established by geometric relationships between travel paths of the largest sized toolbar (or boom) on equipment applying spatial inputs and the producer's intent to differently manage diverse geographic zones in the field. Therefore, in a TED, treatments are assigned to experimental units by producer's (or researcher's) judgment and not through a process of randomization, because randomization results in irrelevant assignments of treatments to some experimental units. Independent treatments are achieved by changes in state of the VR controller as it spatially applies different recommendations along paths of travel.

High data density provided by high resolution imagery, LIDAR, and yield monitors (Birrell *et al.* 1996) provide valuable covariate information to better test hypotheses. Topological relationships of field topography provided by these sensors best allow one to define experimental units useful for replication and to have independence between control and site-specific treatments. Our analysis method builds upon several assumptions and is dependent (as noted above) upon capabilities of VR equipped machinery. First, the VR path of the sprayer or implement is only possible in long strips (polygons). The strips are straight lines or curves following contours or property boundaries. Second, precision agricultural practice prescriptions are formulated to be applied to polygons of interest (management zones) and are spatially varied along the travel path of the farm equipment. Third, demographic characteristics of the polygons are available at the same spatial resolution (or greater) provided by the harvester. To create an analysis approach, spatial demographics of various thematic layers within fields create large zones where 'recommended' (traditional or precision) agricultural practices are applied. Other 'experimental' (traditional or precision) agricultural practices can be evaluated against recommended practices by using smaller 'overlap plots (or zones)' imbedded by analysts into paths of travel of VR farm equipment. Overlap zone locations and recommendations are both built into the prescription file (following recommendations by field personnel) and the VR controller directs the farm equipment to apply the spatial management practices. The key for obtaining data to evaluate (traditional or precision) agricultural practices is to establish overlap zones of alternate management practices within larger zones where recommended practices are applied. This overlapping process provides the necessary control data. The various agricultural practices are evaluated using analysis of covariance to obtain regression effects describing both the large (i.e., recommended or traditional) and overlap (i.e., site-specific or alternative) management zone demographics. For this analysis, treatments applied to experimental units (zones) are fixed effects, while the locations of experimental units are random effects.

A TED must be easily implemented in a commercial field. If the experiment is difficult to accomplish by the operator, then definite changes in design are necessary. At times, some

agricultural inputs are applied to all areas of the field, and here the creation of overlap zones is straightforward. Other times, management practices are applied at different times and locations and require one or more styles of overlap zones to evaluate their effectiveness. Often, spatial applications of one agrichemical cannot be tank mixed with blanket applications of other products applied to the field, unless 'off/on', or changes in rate, criteria are appropriate. Multiple travel paths by equipment to apply several different inputs in various spatial arrangements can be inconvenient to the farm's schedule of events, thus some questions may be too expensive and time consuming. (However, multiple channel, multi-product injection systems now available on some equipment solve this type of limitation.)

Questions of interest addressed by a TED are of at least three general types. The first question evaluates only a single management tactic at one point in time. The second type compares one or more management tactics at one point in time. The third, and more complex question, evaluates one, two or more management approaches throughout a season. The long-term goal of the project is to develop several styles of site-specific experiments to appropriately assess (1) agronomic and entomological research in cotton and (2) provide producers and consultants the ability to implement and analyze TEDs on production farms.

It would be best with a TED to state *a priori* planned comparisons of interest; however, unplanned comparisons are to be expected because one cannot predict in advance how seasonal (weather) and/or other agronomic conditions (market prices, pest outbreaks, etc.) exactly play out in commercial fields. In these cases, the fact that any changes in management practices are topologically registered information will be of great advantage. Flexibility in planning is possible and the TED can keep the best interests of the cooperating farm in mind.

The methods developed to implement and analyze the case study TED provided excellent insight on effects of topography and site-specific management (the PGR application) on yield of several cotton cultivars. Preliminary conclusions from the analysis provide insight for developing additional directions of agricultural research. The inference space of results will require more investigation. We expect the integrated methodology to apply to other difficult to replicate experiments established in commercial fields.

#### **IV. Summary.**

To implement a TED at the field level requires close cooperation among the producer(s), his consultant(s), farm staff, research team, extension agent, and one or more commercial companies (or research organizations) in order to acquire, process and apply various information layers. The analysis also requires close cooperation among the skills of various analysts and intensive applications of diverse types of spatial information.

#### **V. References.**

Anonymous, 1997. *Understanding GIS. The ARC/INFO Method*. ESRI Press, Redlands, CA.

Birrell, S. J., K. A. Suddeth, and S. C. Borgelt. 1996. Comparisons of sensors and techniques for crop yield mapping. *Comp. Elec. Agric.* 14. 215-233.

Bugayevskiy, L. M. and J. P. Snyder. 1995. *Map Projections. A Reference Manual*. Taylor and Francis, Philadelphia, PA.

Conquest, L. L. 2000. Analysis and interpretation of ecological field data using BACI designs: Discussion. *J. Agric. Biol. Env. Stat.* 5: 293-296.

Doerge, T. A and D. L. Gardner. 1999. On-farm testing using the adjacent strip comparison method. In *Precision Agriculture* (P. C. Robert, R. H. Rust, and W. E. Larson (eds)). Am. Soc. Agronomy, Crop Science Soc. America, and Soil Science Soc. America, Publishers. Madison, WI.

Dupont, J. K., R. Campenella, M.R. Seal, J. L. Willers and K. B. Hood. 2000. Spatially variable insecticide applications through remote sensing. *Proc. Beltwide Cotton Conf.*, San Antonio, TX. Vol. 1: 426- 429.

Gotway, C. A., D. G. Bullock, F. J. Pierce, W. W. Stroup, G. W. Hergert, and K. M. Eskridge. 1997. Experimental design issues and statistical evaluation techniques for site-specific management. In *The state of site-specific management for agriculture* (F. J. Pierce and E. J. Sadler (eds)). Am. Soc. Agronomy, Crop Science Soc. America, and Soil Science Soc. America, Publishers. Madison, WI.

Gotway, C. A. and W. W. Stroup. 1997. A generalized linear model approach to spatial data analysis and prediction. *J. Agric. Biol. Env. Stat.* 2: 157-178.

Hurlbert, S. H. 1984. Pseudoreplication and the design of ecological field experiments. *Ecol. Monog.* 54: 187-211.

Jensen, J.R. 2000. *Remote sensing of the environment: An earth resource perspective*. Prentice-Hall, Upper Saddle River, NJ.

Kennedy, M. 1996. *The Global Positioning System: An Introduction*. Ann Arbor Press, Chelsea, MI, USA. 268 pp.

Littell, R. C., G. A. Milliken, W. W. Stroup and R. D. Wolfinger. 1996. *SAS<sup>®</sup> System for Mixed Models*. SAS Institute Inc., Cary, NC.

McDonald, T. L., W. P. Erickson and L. L. McDonald. 2000. Analysis of count data from before-after control-impact studies. *J. Agric. Biol. Env. Stat.* 5: 262-279.

Mead, R. 1988. *The design of experiments: Statistical principles for practical application*. Cambridge University Press, Cambridge.

Milliken, G. A. and D. E. Johnson. 1992. *Analysis of Messy Data, Vol. 1. Designed Experiments*. Chapman and Hall/CRC, New York.



- Milliken, G. A. and D. E. Johnson. 2001. *Analysis of Messy Data, Vol. 3. Analysis of Covariance*. Chapman and Hall/CRC, New York.
- Moran, M.S., Y. Inoue, and E.M. Barnes. 1997. Opportunities and limitations for image-based remote sensing in precision crop management. *Remote Sens. Environ.* 61:319-346.
- Murtaugh, P. A. 2000. Paired intervention analysis in ecology. *J. Agric. Biol. Env. Stat.* 5: 280-292.
- Pierce, F. J., N. W. Anderson, T. S. Colvin, J. K. Schueller, D. S. Humburg, and N. B. McLaughlin. 1997. Yield mapping. In *The state of site-specific management for agriculture* (F. J. Pierce and E. J. Sadler (eds)). Am. Soc. Agronomy, Crop Science Soc. America, and Soil Science Soc. America, Publishers. Madison, WI.
- Pinter, P. J., J. L. Hatfield, J. S. Scheppers, E. M. Barnes, M. S. Moran, C. S. T. Daughtry, and D. R. Upchurch. 2003. Remote sensing for Crop Management. *Photogramm. Engr. & Rem. Sens.* 69: 647-664.
- Plant, R. E., D. S. Munk, B. R. Roberts, R. N. Vargas, R. L. Travis, D. W. Rains and R. B. Hutmacher. 2001. Application of remote sensing to strategic questions in cotton management and research. *J. Cot. Sci.* 5: 30-41.
- Pouncey, R., K. Swanson and K. Hart (ed.) 1999. *ERDAS Field Guide*, 5<sup>th</sup> ed. ERDAS, Atlanta, GA.
- Richards, J. A. and X. Jia. 1999. *Remote Sensing Digital Image Analysis. An Introduction*, 3<sup>rd</sup> ed. Springer, Berlin.
- Scheiner, S. M. and J. Gurevitch, eds. 2001. *Design and Analysis of Ecological Experiments*, 2<sup>nd</sup> ed. Oxford University Press, Oxford.
- Stroup, W. W. 1989. Predictable functions and prediction space in the mixed model procedure. *Applications of Mixed Models in Agriculture and Related Disciplines*, Southern Coop. Ser. Bull. No. 343. Louisiana Agricultural Experiment Station, Baton Rouge.
- Theobald, D. M. 2003. *GIS Concepts and ArcGIS® Methods*. Conservation Planning Technologies, Fort Collins, CO.
- Willers, J. L., M. R. Seal, and R. G. Luttrell. 1999. Remote sensing, line-intercept sampling for tarnished plant bugs (Heteroptera: Miridae) in Mid-south cotton. [Online]. *J. Cotton Sci.* 3(4): 160-170. Available at <http://www.jcotsci.org> (verified 10 Mar. 2005).

**Acknowledgements.** The July 2003 spatial PGR prescription was applied in cooperation with Dr. Dan Reynolds, Mr. Matt Kirkpatrick and Mr. Tom Barber, Department of Weed Science, Mississippi State University. Appreciation is expressed to Mr. Paul Good and Mr. Dale Weaver of Good's Longview Farm for cooperation in this research. The cultivar trial was planted in

cooperation with Good's Longview Farm, and sponsored by Mr. Dave Roberts, Delta Pine and Land Seed Company, Scott, MS and Dr. Will McCarty, State Agronomist, Mississippi State University Extension Service. Research was supported in part by grant funds from the GeoResources Institute, Mississippi State University and the USDA-ARS, Special Project, Area-Wide Pest Management of the Tarnished Plant Bug.

**Disclaimer.** Mention of trade names or commercial products in this publication is solely for the purpose of providing specific information and does not imply recommendation or endorsement by the U.S. Department of Agriculture.

**Table 1.** Ranked (high to low) yield responses (lint (lbs/acre) of 17 cotton cultivars and their percent turnouts. These yields were obtained using traditional methods without any adjustment for effects of topography and the variable PGR prescription. (Courtesy of Dennis Reginelli, Mississippi Cooperative Extension Service, Noxubee County, MS)

<b>Cultivar</b>	<b>Plot ID</b>	<b>Lint Yield (Lbs/Acre)</b>	<b>Percent Turnout</b>
<b>SG 501 BR</b>	<b>6</b>	<b>951</b>	<b>37.9</b>
<b>SG 215 BR</b>	<b>5</b>	<b>940</b>	<b>39.3</b>
<b>FM 991 BR</b>	<b>13</b>	<b>933</b>	<b>36.6</b>
<b>DP 494 R</b>	<b>21</b>	<b>897</b>	<b>38.8</b>
<b>ST 5599 BR</b>	<b>10</b>	<b>890</b>	<b>36.3</b>
<b>FM 960 BR</b>	<b>12</b>	<b>887</b>	<b>35.1</b>
<b>ST 5303 R</b>	<b>20</b>	<b>861</b>	<b>37.4</b>
<b>DP 5415 R</b>	<b>14</b>	<b>856</b>	<b>39.0</b>
<b>FM 989 R</b>	<b>19</b>	<b>832</b>	<b>36.0</b>
<b>DP655 BR</b>	<b>11</b>	<b>830</b>	<b>37.7</b>
<b>ST 4892 BR</b>	<b>9</b>	<b>829</b>	<b>37.2</b>
<b>SG 521 R</b>	<b>17</b>	<b>828</b>	<b>38.3</b>
<b>ST 4793 R</b>	<b>15</b>	<b>824</b>	<b>38.2</b>
<b>DP 451 BR</b>	<b>7</b>	<b>817</b>	<b>34.8</b>
<b>FM 989 BR</b>	<b>8</b>	<b>806</b>	<b>36.3</b>
<b>DP 436 R</b>	<b>16</b>	<b>767</b>	<b>34.9</b>
<b>DP 5690 R</b>	<b>18</b>	<b>764</b>	<b>35.4</b>

**Table 2.** Pearson correlation coefficients between predicted values from mixed model using site characteristics and focal-loess smoothed data points (say) and focal smoothed data points (orig\_say) for each cultivar.

Cultivar	say	orig_say
501 BR	0.75540	0.73281
215 BR	0.61408	0.59210
991 BR	0.74590	0.68756
494 R	0.72194	0.66983
5599 BR	0.75436	0.72852
960 BR	0.75265	0.68892
5303 R	0.63507	0.57841
5415 R	0.78542	0.74210
989 R	0.62923	0.56943
655 BR	0.70470	0.67695
4892 BR	0.79670	0.77531
521 R	0.69374	0.56814
4793 R	0.74602	0.72055
451 BR	0.72781	0.70838
989 BR	0.74890	0.73175
436 R	0.69458	0.65572
5690 R	0.61695	0.53376

**Table 3.** Predicted seed cotton yield means for each of the cultivars for a second blanket application (above the base rate of 5 oz) of PGR at 0, 2, 4, 6, and 8 (oz/acre) denoted by rate\_0,...rate\_8 with the ranks of the mean within a rate (rank\_0,...,rank\_8) from the model using site characteristics. Cultivars are listed as presented in Table 1 to facilitate comparisons (after adjustments to lint yield using the gin turnout values in Table 1).

Cultivar	rate_0	rank_0	rate_2	rank_2	rate_4	rank_4	rate_6	rank_6	rate_8	rank_8
501 BR	2715	1	2604	3	2542	4	2530	1	2569	1
215 BR	2702	2	2660	2	2571	2	2436	5	2253	5
991 BR	2652	3	2666	1	2603	1	2463	2	2246	6
494 R	2162	17	2102	17	2008	17	1879	17	1715	17
5599 BR	2295	14	2434	5	2484	5	2445	4	2317	3
960 BR	2488	5	2554	4	2543	3	2454	3	2287	4
5303 R	2416	8	2310	11	2204	11	2098	12	1992	12
5415 R	2471	6	2432	7	2336	8	2183	10	1973	13
989 R	2384	9	2281	13	2194	13	2123	11	2066	11
655 BR	2302	13	2434	6	2461	6	2384	6	2203	9
4892 BR	2236	16	2383	9	2430	7	2377	7	2224	8
521 R	2306	12	2212	16	2125	16	2045	16	1972	14
4793 R	2352	11	2294	12	2201	12	2072	15	1908	16
451 BR	2570	4	2397	8	2331	10	2372	8	2519	2
989 BR	2433	7	2383	10	2335	9	2289	9	2244	7
436 R	2292	15	2263	14	2194	14	2084	14	1933	15
5690 R	2357	10	2226	15	2139	15	2096	13	2096	10

**Table 4.** Standard deviation information for pairwise comparisons among the cultivars and 0.05 LSD values computed using mean standard deviation (mn\_std), minimum standard deviation (min\_std) and maximum standard deviation (max\_std) for each of the rates of PGR from a regression model using site characteristics.

rate	mn_std	LSD_mn	min_std	LSD_min	max_std	LSD_max
0.00	92	194	86	180	104	218
2.00	64	134	59	125	72	152
4.00	55	115	46	98	66	138
6.00	38	80	34	72	43	90
6.17	37	77	34	71	41	85
8.00	58	122	39	81	98	206

**Table 5.** Tests of the fixed effects for the model with site characteristics (continued next page).  
 See Appendix 1 for label descriptions for effects.

<b>Type 3 Tests of Fixed Effects</b>				
<b>Effect</b>	<b>Num DF</b>	<b>Den DF</b>	<b>F Value</b>	<b>Pr &gt; F</b>
<b>cultivar</b>	<b>16</b>	<b>17</b>	<b>10.28</b>	<b>&lt;.0001</b>
<b>rate</b>	<b>1</b>	<b>5979</b>	<b>2.82</b>	<b>0.0933</b>
<b>rate*rate</b>	<b>1</b>	<b>5979</b>	<b>4.45</b>	<b>0.0349</b>
<b>rate*cultivar</b>	<b>16</b>	<b>5979</b>	<b>6.18</b>	<b>&lt;.0001</b>
<b>rate*rate*cultivar</b>	<b>16</b>	<b>5979</b>	<b>7.41</b>	<b>&lt;.0001</b>
<b>Slo</b>	<b>1</b>	<b>5979</b>	<b>3.00</b>	<b>0.0834</b>
<b>logfac</b>	<b>1</b>	<b>5979</b>	<b>21.60</b>	<b>&lt;.0001</b>
<b>Cvx</b>	<b>1</b>	<b>5979</b>	<b>544.65</b>	<b>&lt;.0001</b>
<b>logeuc</b>	<b>1</b>	<b>5979</b>	<b>3.92</b>	<b>0.0477</b>
<b>Dsm</b>	<b>1</b>	<b>5979</b>	<b>527.19</b>	<b>&lt;.0001</b>
<b>Ndvi_dif</b>	<b>1</b>	<b>5979</b>	<b>3.09</b>	<b>0.0788</b>
<b>index</b>	<b>1</b>	<b>5979</b>	<b>0.05</b>	<b>0.8207</b>
<b>one_cvx</b>	<b>1</b>	<b>5979</b>	<b>678.02</b>	<b>&lt;.0001</b>
<b>Ndvi_717</b>	<b>1</b>	<b>5979</b>	<b>2.54</b>	<b>0.1112</b>
<b>logfac*Cvx</b>	<b>1</b>	<b>5979</b>	<b>290.96</b>	<b>&lt;.0001</b>
<b>Dsm*cultivar</b>	<b>16</b>	<b>5979</b>	<b>10.35</b>	<b>&lt;.0001</b>
<b>rate*Dsm</b>	<b>1</b>	<b>5979</b>	<b>2.80</b>	<b>0.0944</b>
<b>rate*Cvx*cultivar</b>	<b>17</b>	<b>5979</b>	<b>9.77</b>	<b>&lt;.0001</b>
<b>rate*logeuc*cultivar</b>	<b>17</b>	<b>5979</b>	<b>1.91</b>	<b>0.0135</b>
<b>rate*Ndvi_71*cultivar</b>	<b>17</b>	<b>5979</b>	<b>5.94</b>	<b>&lt;.0001</b>
<b>rate*rate*logf*cultivar</b>	<b>17</b>	<b>5979</b>	<b>8.29</b>	<b>&lt;.0001</b>
<b>rate*rate*loge*cultivar</b>	<b>17</b>	<b>5979</b>	<b>1.89</b>	<b>0.0145</b>
<b>rate*rate*Dsm*cultivar</b>	<b>17</b>	<b>5979</b>	<b>7.40</b>	<b>&lt;.0001</b>
<b>rate*rate*Ndvi*cultivar</b>	<b>17</b>	<b>5979</b>	<b>4.85</b>	<b>&lt;.0001</b>
<b>rate*logfac</b>	<b>1</b>	<b>5979</b>	<b>1.52</b>	<b>0.2176</b>
<b>rate*rate*Cvx*cultivar</b>	<b>17</b>	<b>5979</b>	<b>9.62</b>	<b>&lt;.0001</b>
<b>rate*rate*Ndvi*cultivar</b>	<b>17</b>	<b>5979</b>	<b>3.18</b>	<b>&lt;.0001</b>

<b>Type 3 Tests of Fixed Effects</b>				
<b>Effect</b>	<b>Num DF</b>	<b>Den DF</b>	<b>F Value</b>	<b>Pr &gt; F</b>
<b>rate*rate*Slo*cultivar</b>	<b>17</b>	<b>5979</b>	<b>64.09</b>	<b>&lt;.0001</b>
<b>rate*index*cultivar</b>	<b>17</b>	<b>5979</b>	<b>8.76</b>	<b>&lt;.0001</b>
<b>rate*Ndvi_di*cultivar</b>	<b>17</b>	<b>5979</b>	<b>3.85</b>	<b>&lt;.0001</b>
<b>rate*Slo</b>	<b>1</b>	<b>5979</b>	<b>0.79</b>	<b>0.3748</b>
<b>logfac*Dsm</b>	<b>1</b>	<b>5979</b>	<b>22.91</b>	<b>&lt;.0001</b>
<b>Slo*Dsm</b>	<b>1</b>	<b>5979</b>	<b>2.89</b>	<b>0.0894</b>
<b>logfac*Ndvi_dif</b>	<b>1</b>	<b>5979</b>	<b>51.67</b>	<b>&lt;.0001</b>
<b>one_ndvi</b>	<b>1</b>	<b>5979</b>	<b>2.79</b>	<b>0.0947</b>
<b>Dsm*Dsm*cultivar</b>	<b>17</b>	<b>5979</b>	<b>55.82</b>	<b>&lt;.0001</b>



**Table 6.** Predicted seed cotton yield means for each of the cultivars for blanket application (above a base of 5 oz) of PGR at 0, 2, 4, 6, and 8 (oz/acre) denoted by rate\_0,...rate\_8 with the ranks of the mean within a rate (rank\_0,...,rank\_8) from a regression model without site characteristics. Cultivars are listed as presented in Table 1 to facilitate comparisons.

Cultivar	rate_0	rank_0	rate_2	rank_2	rate_4	rank_4	rate_6	rank_6	rate_8	rank_8
501 BR	3093	1	2538	6	2348	10	2522	5	3061	1
215 BR	2967	3	2548	5	2371	8	2436	7	2743	4
991 BR	2791	5	2728	3	2658	3	2580	3	2494	9
494 R	2422	16	2454	11	2354	9	2122	17	1759	17
5599 BR	2725	7	2769	2	2783	2	2768	1	2723	5
960 BR	2795	4	2834	1	2800	1	2694	2	2515	8
5303 R	2585	11	2476	7	2372	7	2275	12	2184	16
5415 R	2557	12	2460	9	2375	6	2305	10	2247	11
989 R	2602	10	2449	12	2347	11	2295	11	2293	10
655 BR	2694	8	2645	4	2603	4	2570	4	2546	7
4892 BR	2673	9	2467	8	2412	5	2510	6	2759	3
521 R	2458	15	2312	15	2226	15	2199	13	2233	13
4793 R	2474	14	2341	14	2250	14	2199	14	2189	15
451 BR	2973	2	2458	10	2270	13	2410	8	2877	2
989 BR	2766	6	2414	13	2288	12	2388	9	2714	6
436 R	2398	17	2276	17	2202	16	2174	16	2194	14
5690 R	2493	13	2304	16	2200	17	2181	15	2247	12

**Table 7.** Standard deviation information for pairwise comparisons among the cultivars and 0.05 LSD values computed using mean standard deviation (mn\_std), minimum standard deviation (min\_std) and maximum standard deviation (max\_std) for each of the rates of PGR from model without using site characteristics.

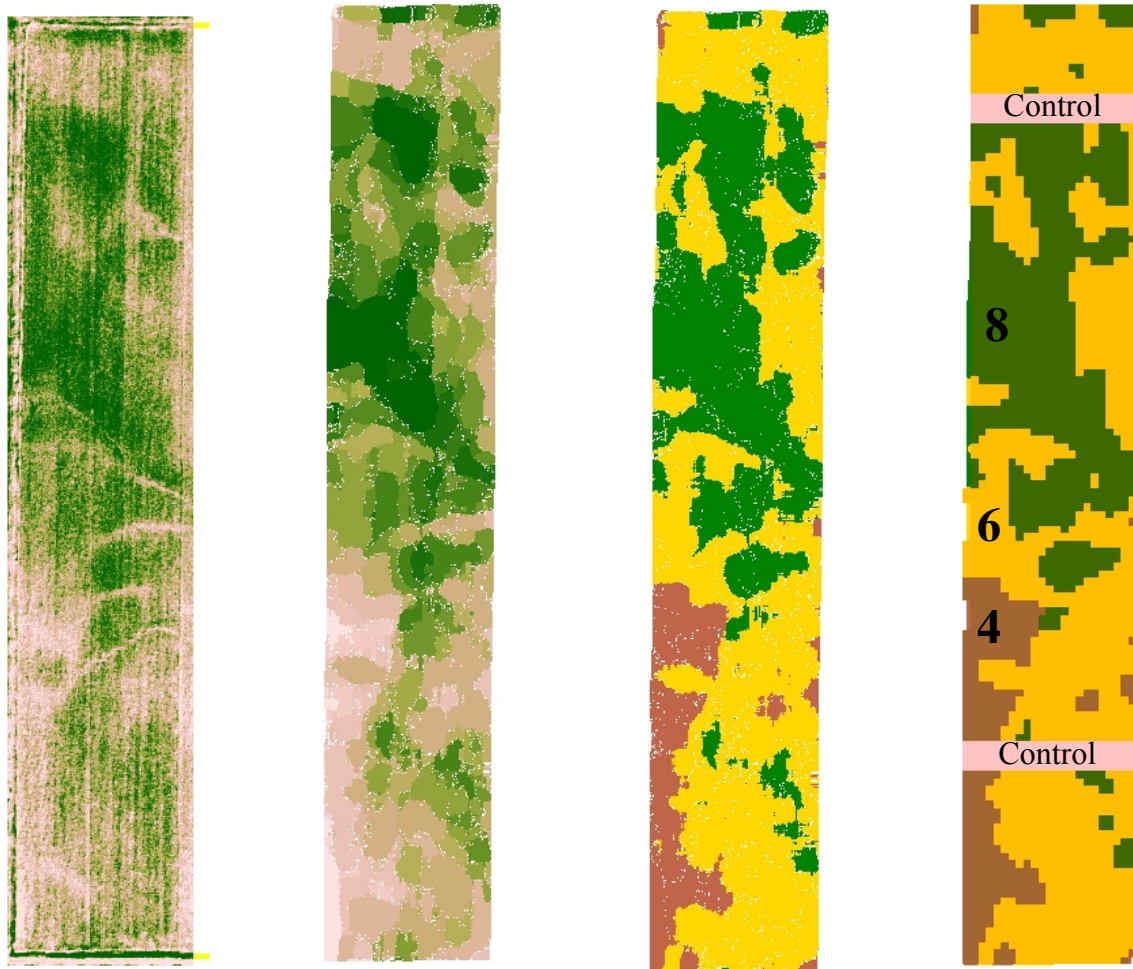
rate	mn_std	LSD_mn	min_std	LSD_min	max_std	LSD_max
<b>0.00</b>	<b>216</b>	<b>455</b>	<b>214</b>	<b>449</b>	<b>232</b>	<b>487</b>
<b>2.00</b>	<b>190</b>	<b>398</b>	<b>163</b>	<b>342</b>	<b>223</b>	<b>468</b>
<b>4.00</b>	<b>203</b>	<b>427</b>	<b>168</b>	<b>353</b>	<b>248</b>	<b>520</b>
<b>6.00</b>	<b>144</b>	<b>303</b>	<b>125</b>	<b>263</b>	<b>169</b>	<b>354</b>
<b>6.17</b>	<b>138</b>	<b>290</b>	<b>122</b>	<b>256</b>	<b>159</b>	<b>335</b>
<b>8.00</b>	<b>190</b>	<b>399</b>	<b>178</b>	<b>373</b>	<b>208</b>	<b>436</b>

**Table 8.** Tests of the fixed effects for the model without site characteristics.

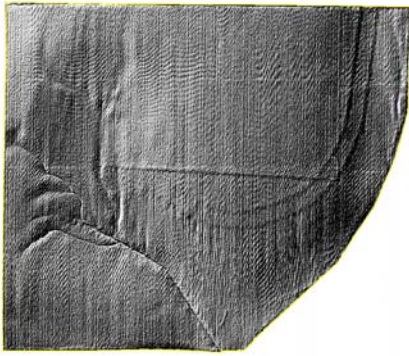
<b>Type 3 Tests of Fixed Effects</b>				
<b>Effect</b>	<b>Num DF</b>	<b>Den DF</b>	<b>F Value</b>	<b>Pr &gt; F</b>
<b>cultivar</b>	<b>16</b>	<b>3139</b>	<b>1.80</b>	<b>0.0257</b>
<b>rate</b>	<b>1</b>	<b>6250</b>	<b>23.36</b>	<b>&lt;.0001</b>
<b>rate*cultivar</b>	<b>16</b>	<b>6250</b>	<b>2.05</b>	<b>0.0082</b>
<b>rate*rate*cultivar</b>	<b>17</b>	<b>3440</b>	<b>3.43</b>	<b>&lt;.0001</b>

Num DF = Numerator Degrees of Freedom;

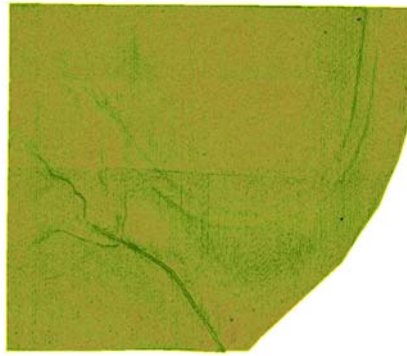
Den DF = Denominator Degrees of Freedom



**Figure 1.** NDVI data products are generated to apply spatially variable-rates of a plant growth regulator (PGR) to control cotton growth. From left to right, the NDVI index is classified (left), focally clarified (second from left), and re-classed into three codes (third from left). The patterns are refined and filled to match the geometry of the sprayer boom (right). Imbedded are two control strips (no PGR applied) for statistical analysis of the variable-rate PGR application after harvest. Overlaid numbers (4, 6, 8) correspond to different PGR rates applied to the underlying three color classes during July 2003.



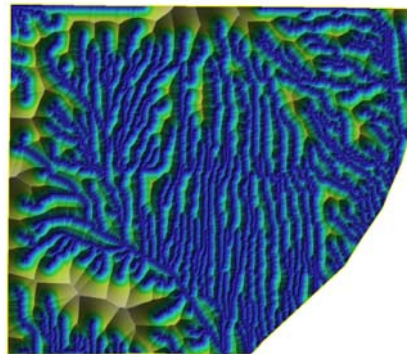
Hillshade



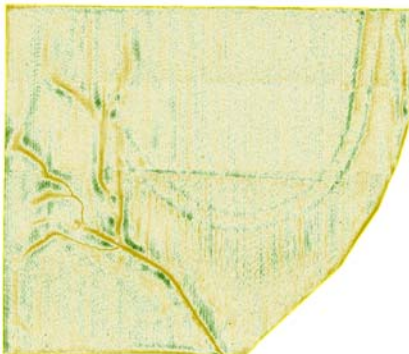
Slope



Aspect

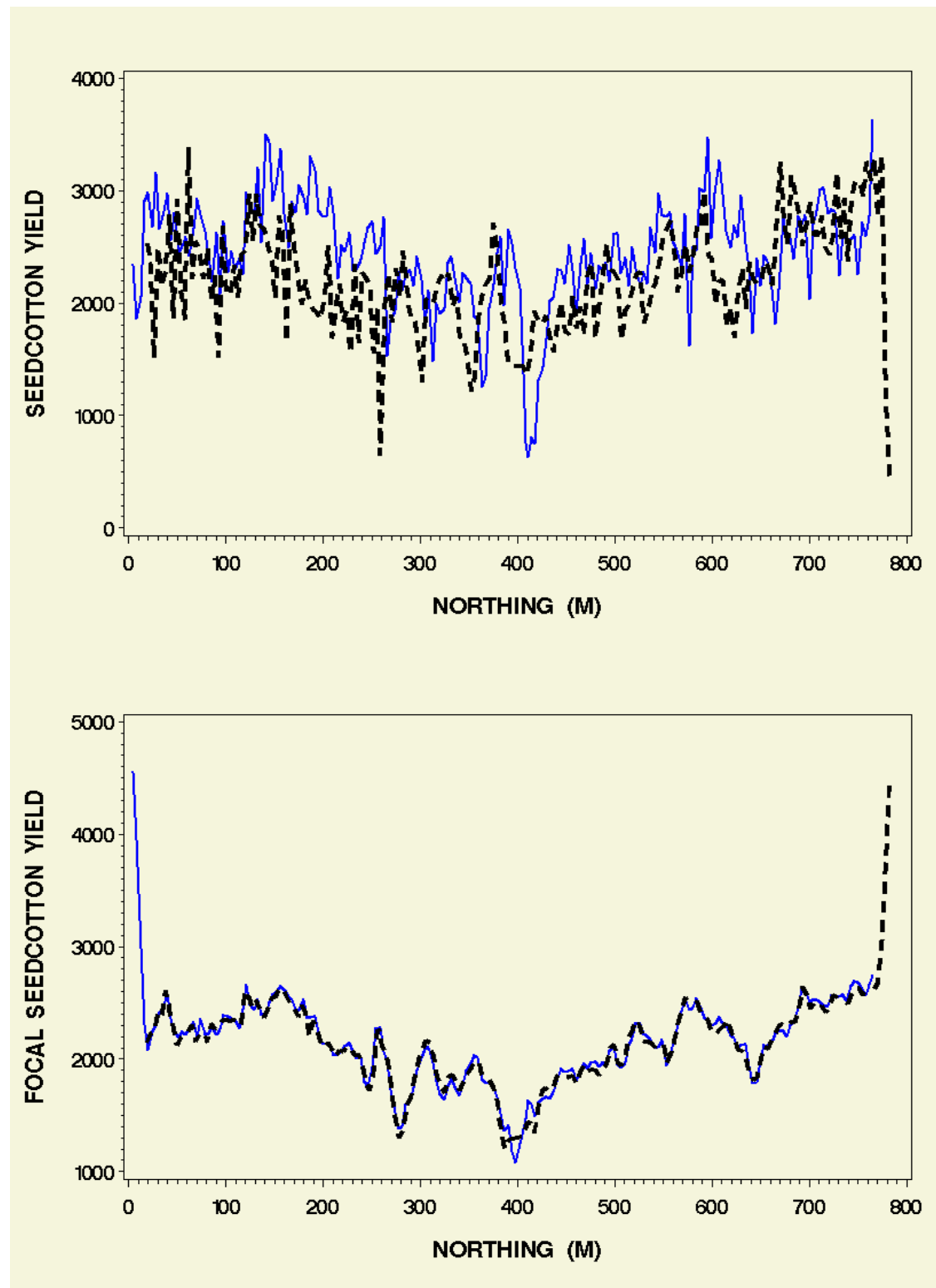


Distance

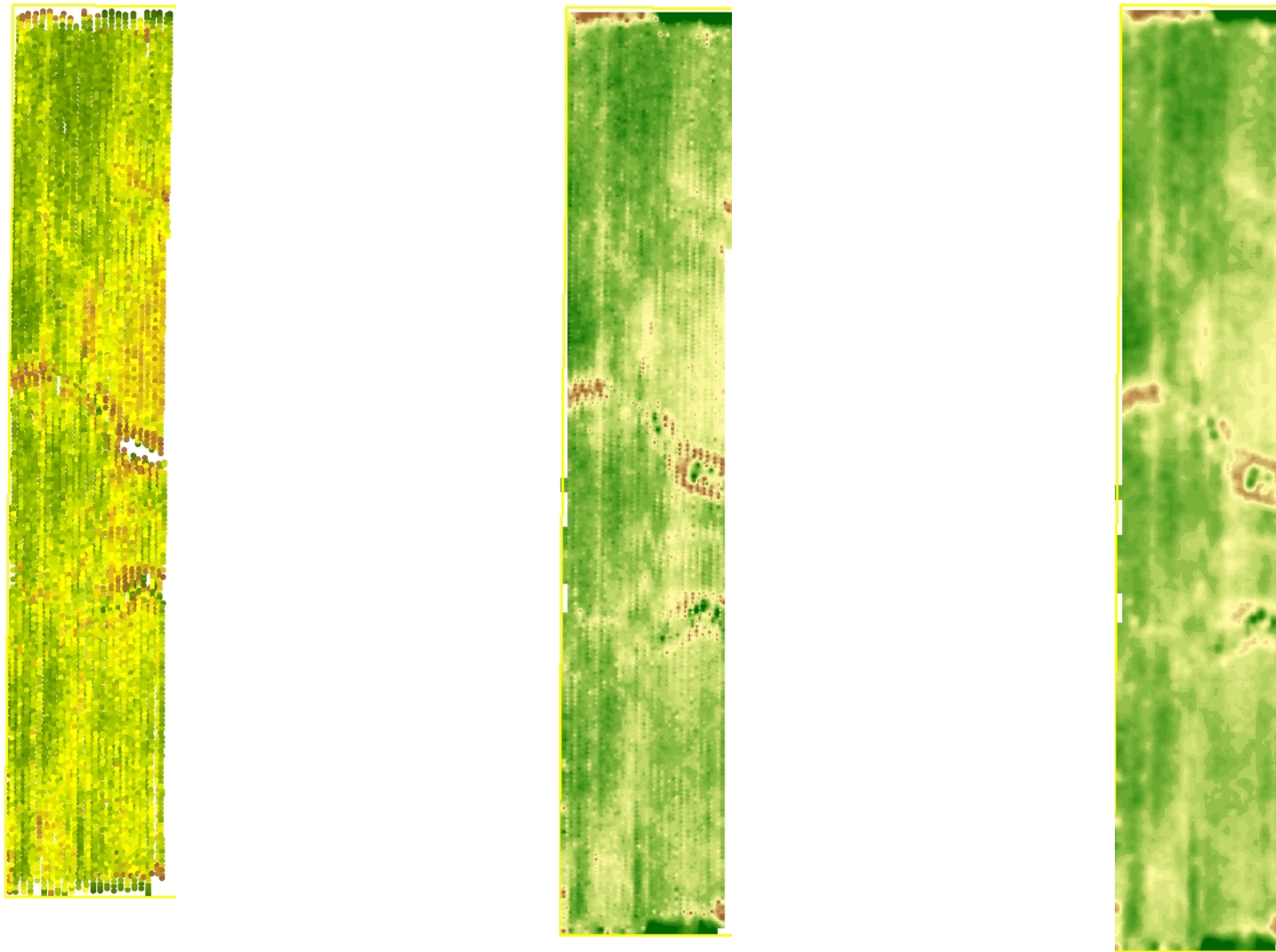


Concavity

**Figure 2.** LIDAR technology can generate highly detailed, accurate terrain data. These are pictorial examples of at least five data products (grids) derived from the digital surface model (DSM). Shown is all of Field 160 where the western edge (Figure 1) was utilized for the 2003 cultivar trial.

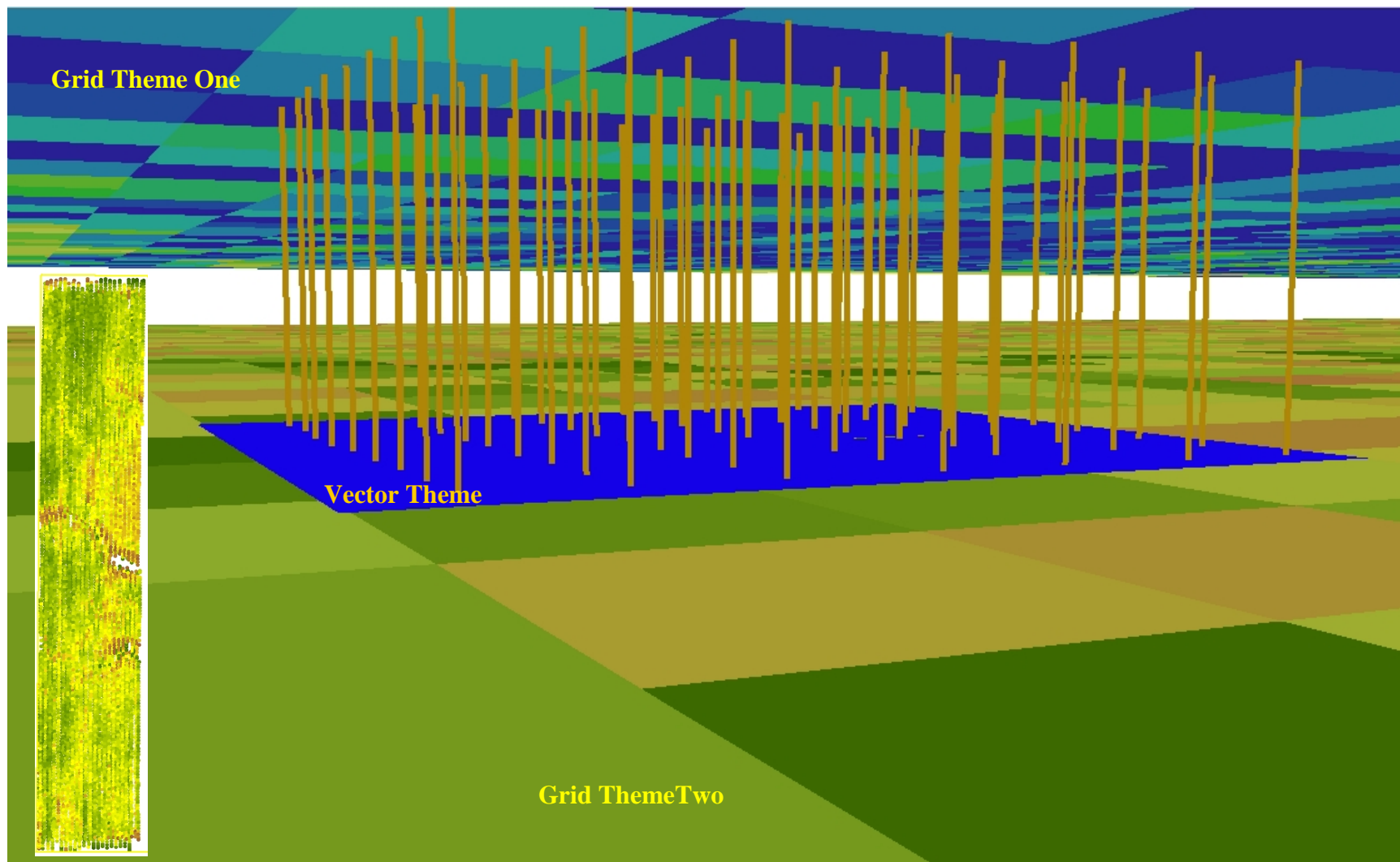


**Figure 3.** For a given cultivar, the north- and south-bound picker loads exhibited noisy (top) yield response characteristics addressed through focal processing (bottom). Shown here are loads 30 (blue line) and 31 (black line).

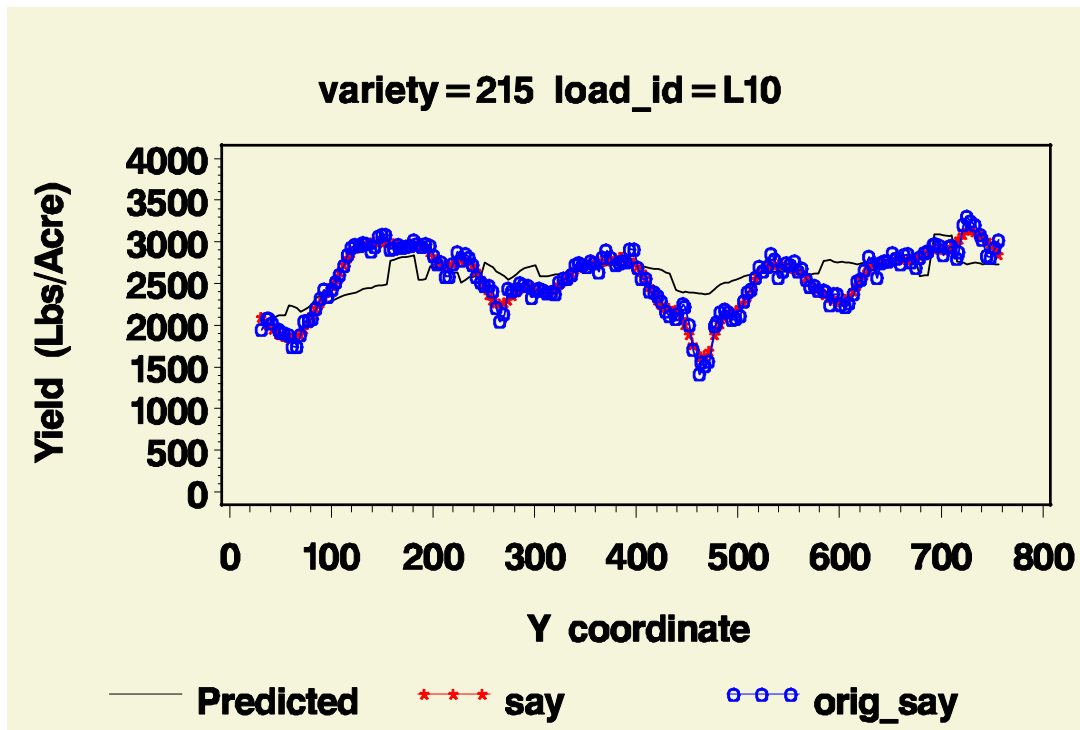


**Figure 4.** The yield point data (left panel) were transformed to a raster grid (center panel) and spatially averaged (right panel) using focal geometry (see Fig. 3) spanning the two picker passes (4 rows each) within each cultivar strip of 8 rows. Focally adjusted yield data (orig\_say) were extracted at their coordinate locations for statistical analysis.

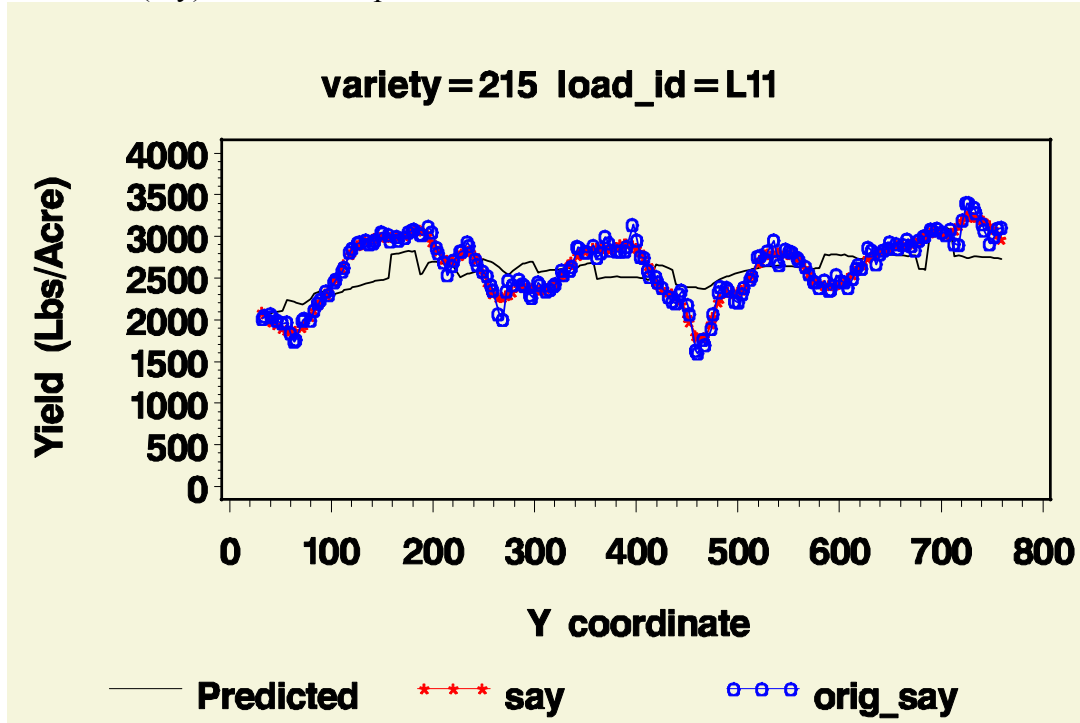




**Figure 5.** Illustration showing yield data point coordinates (left panel inset) intersected (represented by the brown vertical rods) with two grid themes (at different spatial scales) and a vector theme using GIS functions.

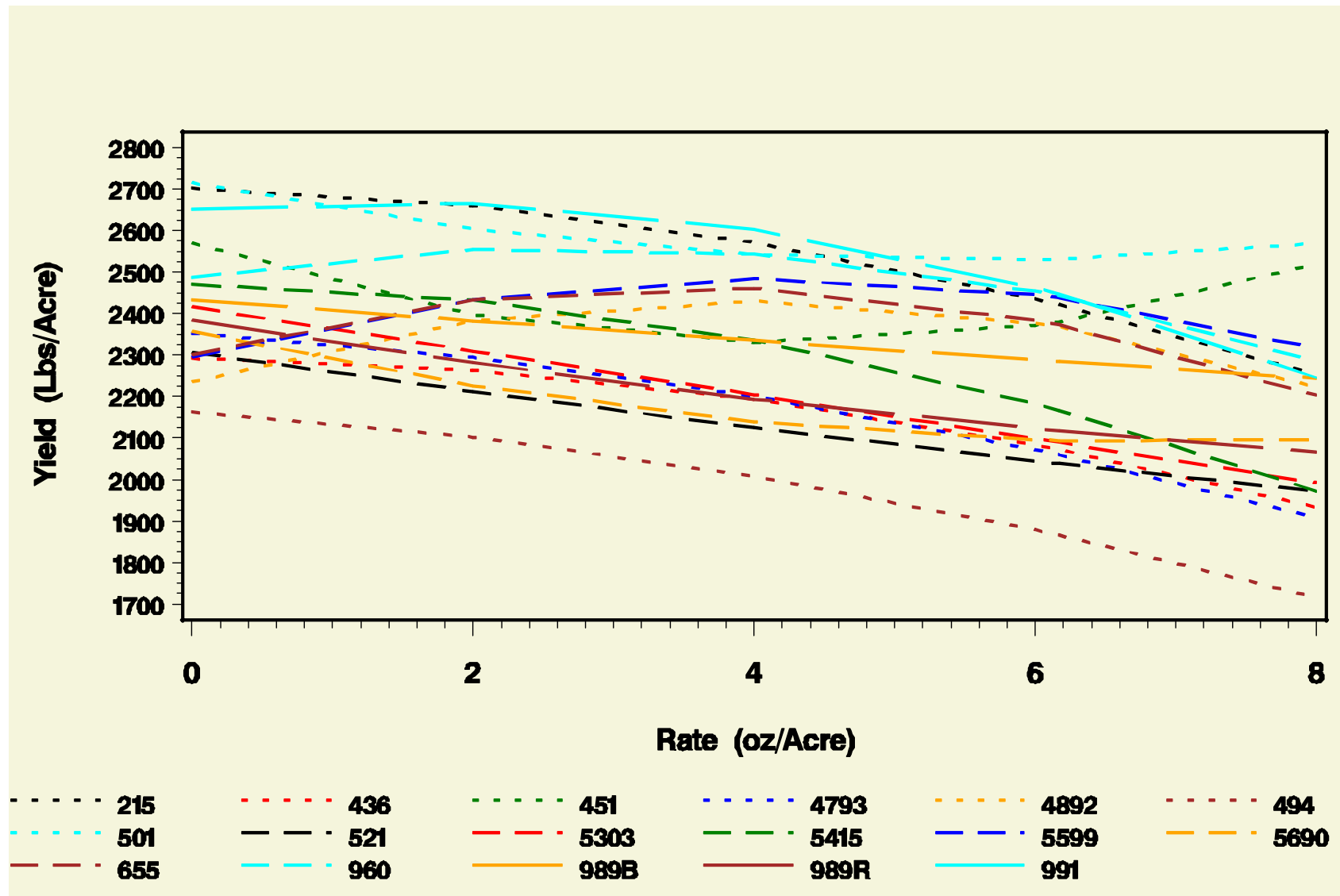


**Figure 6a.** Graph of the focal smoothed seed cotton yield data (orig\_say), the focal-loess smoothed (say) data and the predictions from the model for cultivar 215 and load L10.

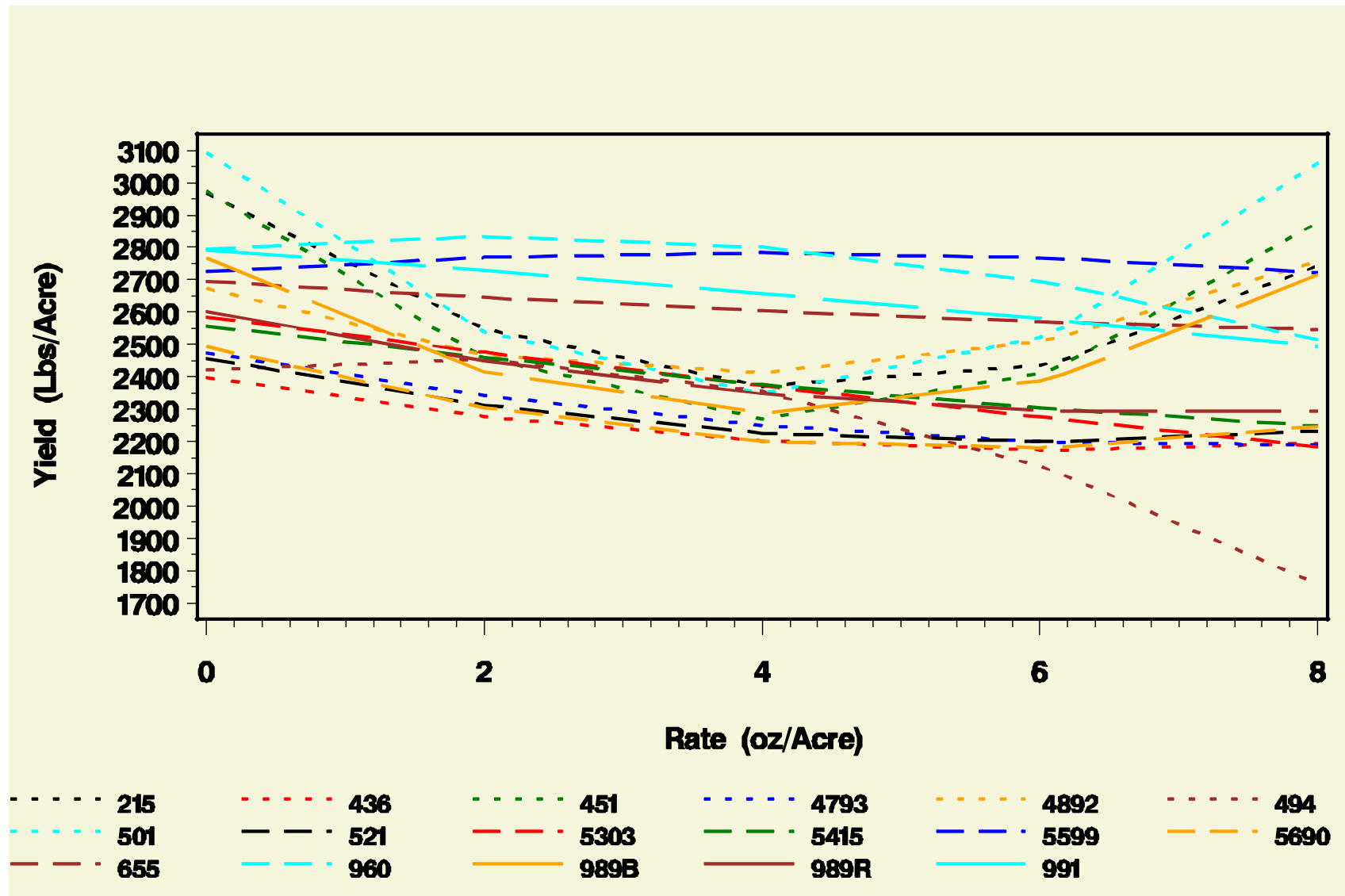


**Figure 6b.** Graph of the focal smoothed seed cotton yield data (orig\_say), the focal-loess smoothed (say) data and the predictions from the model for cultivar 215 and load L11.





**Figure 7.** Graph of the predicted responses (seed cotton yield) from the model using site characteristics for each of the cultivars for five rates of PGR (applied above the base rate of 5 oz).



**Figure 8.** Predicted responses (seed cotton yield) from the model without site characteristics for each of the cultivars at five PGR rates (applied above the base rate of 5 oz).

## Appendix 1

### Variable Labels and Acronymns

(Most abbreviations in general order of appearance as presented in Table 5)

1. **Cultivar**-The factor for one of 17 cotton cultivars used in the study. See Table 1 for the cultivar labels.
2. **Rate**-The factor for one of the plant growth regulator (PGR) rates applied to the cotton cultivars in July according to a variable-rate prescription.
3. **Slo**-The slope of the land at the location of the yield point coordinate.
4. **logfac (=logf)**-The log of the flow accumulation (fac) to a point in the field.
5. **Cvx**-The convexity/convexity value at a yield point coordinate.
6. **Logeuc (=loge)**-The log of the Euclidean distance of a yield point coordinate from a synthetic stream network.
7. **Ndvi (=NDVI)**-The normalized difference vegetation index on 9 June 2003, which is a ratio of the (near infrared band minus the red band) divided by the (near infrared band plus the red band).
8. **Ndvi\_dif (=Ndvi\_di)**-The raster difference between the Ndvi value on 9 June 2003 and the Ndvi value on 15 January 2002.
9. **Ndvi\_717**-The Ndvi value on the 17 July 2003.
10. **Index**-The logfac divided by the tangent of the slope. The index indicates 'bogginess' for a grid cell in the digital surface model (DSM).
11. **Dsm (=DSM)**-The elevation (m) of a yield point coordinate provided by the LIDAR digital surface model.
12. **one\_cvx**- The reciprocal of the Cvx value.
13. **PAT**-The point attribute table of a point theme.
14. **VR**-Variable-rate.
15. **RS**-Remote sensing.
16. **GIS**-Geographic Information System.
17. **DGPS**-Differential Global Positioning System.
18. **AOI**-Area of interest; the geographic boundaries of a study.