United States
Department of
Agriculture

National
Agricultural
Statistics
Service

Research and
Development
Division

RDD Research Report
Number RDD-02-05

April 2002

Evaluation of Yield Monitor Data Importance to NASS

Paul W. Cook

United States Department of Agriculture

National Agricultural Statistics Service

1400 Independence Avenue, SW
Washington, DC
20250-2000

April 11, 2002

TO: Administrator
Associate Administrator
Deputy Administrators
Division Directors
Branch Chiefs
Section Heads
State Statisticians

FROM: Carol House
Director
Research and Development Division

SUBJECT: Distribution of Research Report, "Evaluation of Yield Monitor Data Importance to NASS"

The enclosed report, "Evaluation of Yield Monitor Data Importance to NASS" documents NASS's first experience with yield monitor data. The report is authored by Paul Cook.

Valuable lessons were learned as yield monitor data proved to be quite complex and subject to many potential sources of measurement error. This small pilot level project was conducted with Purdue University as a research partner. At the time the study was initiated, NASS was under pressure to consider becoming the "keeper of the data" for yield monitor data for the nation.

For several fields on Purdue U.'s Davis Research Farm, yield monitor data was collected at harvest for corn and soybeans, as well as weigh wagon data. In addition, thanks to the Indiana SSO, nearly 60 objective yield plots per field had both field and lab data collected. The Illinois SSO processed the lab data.

The general conclusion was that yield monitor data was still subject to too many potential sources of measurement error to be useful to NASS in its yield forecasting and estimation program, and the data volume is also quite cumbersome to deal with. The report distribution is designated for limited distribution to the research community.

If you have any questions, or comments or would like additional copies of the report, please contact Paul Cook at 703- 235-5218, Extension 122.

EVALUATION OF YIELD MONITOR DATA IMPORTANCE TO NASS, by Paul W. Cook, Spatial Analysis Research Section, Geospatial Information Branch, Research and Development Division, National Agricultural Statistics Service, U.S. Department of Agriculture, Washington, D.C. 20250-2000, October 2001, NASS Research Report RDD-02-05.

ABSTRACT

The National Agricultural Statistics Service (NASS) currently uses farmer-reported (list-frame) based procedures to make state and county level crop yield estimates. NASS also collects objective yield data on small field sites as a complement to the farmer provided information for state and national level yield estimation. Two crops of particular interest for crop yield determination are corn and soybeans. In conjunction with Purdue University, NASS collected objective yield data plots chosen within two selected fields of corn and soybeans at the Purdue University Davis Research Farm in east-central Indiana in order to compare with yield monitor data and weigh wagon data that the university staff collected for this project.

This study analyzes the within field relationships of corn and soybean yields from the NASS objective yield data with the yield monitor data. Crop yields for corn and soybeans varied substantially within the fields for both the objective yield and the yield monitor data. The yield monitor yield levels and weigh wagon crop yields were generally in agreement with the objective yield estimated crop yields at the field level for soybeans, but not for corn. However, the spatial relationships within the two fields were substantially different for the two types of data.

The study concluded that yield monitor data is not of importance for direct use by NASS. The Economic Research Service (ERS) of USDA has recommended incorporating questions about yield monitors into future surveys to monitor adoption rates.

KEYWORDS

Geographic Information Systems(GIS); Global Positioning System (GPS); Objective Yield; Precision Farming; Yield Monitor; Soil Variables.

ACKNOWLEDGMENTS

My special thanks go to George Hanuschak, Branch Chief of NASS's Geospatial Information Branch for his efforts to arrange the agreement with Purdue University, oversee many aspects of the project as needed, travel to help with laying out the NASS samples, and provide needed training. My thanks go as well to Ralph Gann, State Statistician, Indiana for his strong support of the research with Purdue University and leading the NASS objective yield data collection for the two study fields at Purdue Davis Farms. For technical assistance, my thanks go to Martin Ozga, Research and Development Division, for FLDHARV, the program that made preliminary corrections to the yield monitor data, Jim Burt, Research, for preliminary studies of the yield monitor data, Keith Morris, Purdue University, for obtaining GPS locations for the NASS samples, Tim Keller, Sampling and Estimation Research Section, for providing information on objective yield procedures and current research efforts, and Gail Wade and Jeff Johnson for suggestions to make better use of ArcView 3.2, an ESRI Geographic Information System (GIS) computer program. The Illinois State Office did the two fields' laboratory work on the NASS objective yield sample data. Bob Hale, Head, Spatial Analysis Research Section, and Charles Day have provided a careful review and editorial recommendations that have helped improve my paper.

TABLE OF CONTENTS

LIST OF TABLES

TABLE OF FIGURES

SUMMARY

This report documents some results from a small pilot study by NASS staff to become familiar with yield monitor data. The USDA Under Secretary of Research, Education, and Economics (REE) authorized NASS to conduct this study. This Summary Section and the Discussion Section (pages 11- 13) provide a complete executive summary. The remaining sections of the report provide an analysis of one year of corn and soybean data for two fields within the Purdue University Davis Research Farm. The limited quantity of data prevents any broad statistical inferences regarding yield monitor data characteristics outside the two fields studied.

Crop yield estimation relies primarily on farmer-reported yields for both county-and state-level crop yield estimation. NASS objective yield procedures use a sample of small, hand-harvested plots within crop fields to provide an additional indication of state-level yield estimates.

Precision Farming Researchers develop new technologies to monitor and reduce within-field crop yield variability to assist in reducing the costs of farming while improving crop yields. Developers of precision farming procedures use the latest in available technology in four areas:

1) Development of yield monitors, often with positional capabilities (Global Positioning System (GPS)) to locate their observations and include crop moisture levels, protein, and water stress (AgLeader Technology, 1997),

2) Evaluation of improved variable rate technology to deliver water, seeds, nutrients or chemical sprays only where needed within the field thereby reducing waste and improving efficiency (Rawlins, 1996),

3) Production of Geographic Information Systems (GIS) maps displaying the field's crop yield variability to develop a plan for variable rate technology devices (Blackmore, 1999), and

4) Production of more detailed soil maps to provide soil type and nutrient differences within a field.

This paper provides information regarding the basics of precision farming research and its importance to NASS in setting crop yields. This goal requires examining an example of precision farming using yield monitor data (category one above) in greater detail. A thorough evaluation will provide a better understanding of the strengths and weaknesses of yield monitor data for evaluating crop yields. Specifically, this report analyzes the yield monitor data collected by Purdue University in 1998 for two fields at the Purdue University Davis Research Farm. An analysis of NASS objective yield sample data for each of the two fields is an integral part of the analysis, as well.

Three general conclusions regarding the yield monitor data and its importance to NASS are evident from these analyses. First, the large quantity of data required for analysis combined with the modest rate of adoption of yield monitors would argue that NASS will currently not find yield

monitor data of importance in setting either U.S. or state-level crop yields. Second, the degree of calibration accuracy exhibited by the various available yield monitors, potential analysis errors, and complexity of yield monitor data prevent its use in operational crop yield surveys. Finally, yield monitor data cannot substitute for weigh wagon, crop marketing (sales tickets), or on-farm storage volume information because its development is not sufficiently advanced. Indeed, yield monitor data is not a sufficiently error-free data source to be considered a standard when compared with weigh wagon data.

1. BACKGROUND

This project examines Purdue University yield monitor data with NASS objective yield data. NASS currently has an operational program to use the objective yield procedures in combination with list frame farmer reported data to estimate state-level crop yields. Specifically, NASS harvests a small portion of grain from two plots, typically for 200-300 fields within a major crop producing state, to provide early indications of state-level crop yields. Also, NASS mails questionnaires to a list frame sample of cooperating farmers to confirm a state's crop yields. Use of these surveys provides NASS a method of making both forecasts, early to mid-season and estimates of crop yields near the time of harvest.

Although precision farming is a recently developed field of study, university researchers and commercial interests have already published many journal articles. Many authors praise precision farming procedures for improving crop yields with reduced farming costs. Others discuss potential problems in using the data from precision farming. This report focuses on the latter.

Yield monitors are electronic devices measuring crop harvested weights from harvesters at intervals of one to five seconds during harvest. GPS receivers give associated latitude and longitude coordinates from satellites on the move at harvest, with some accuracy constraints, such as the number of satellites and signal interference from trees, and so forth. Additional instruments make other crop measurements such as moisture percentage of the crops at harvest. Since the yield monitor requires a calibration of its sensors for accurate readings, the yield monitor data can exhibit irregularities. Typically, yield monitors require a calibration with weigh wagon records taken for corresponding sub-field areas (Ag Leader Technology, 1997). The literature documents many problems with using yield monitor data. For example, *Precision Agriculture, Proceedings of the 3rd International Conference* (1996) contains articles regarding corrections necessary to make yield monitor data of value. Of course, the yield monitor data can help the farmer focus on sub-field areas where crop yields need improvement. One particular Midwestern benefit has been the identification of soil drainage problems due to broken tiles that require replacement.

NASS initiated an agreement in 1996 with Purdue University to examine the capabilities and potential value of precision farming data in our work. The primary focus of the studies was on yield monitor data. Five professors with the Agriculture Engineering Department of Purdue University were the primary researchers cooperating in the research. Dr. Dan Ess was the lead researcher during the agreement's early years. Dr. Sam Parsons assumed that role in early 1998. The remaining three

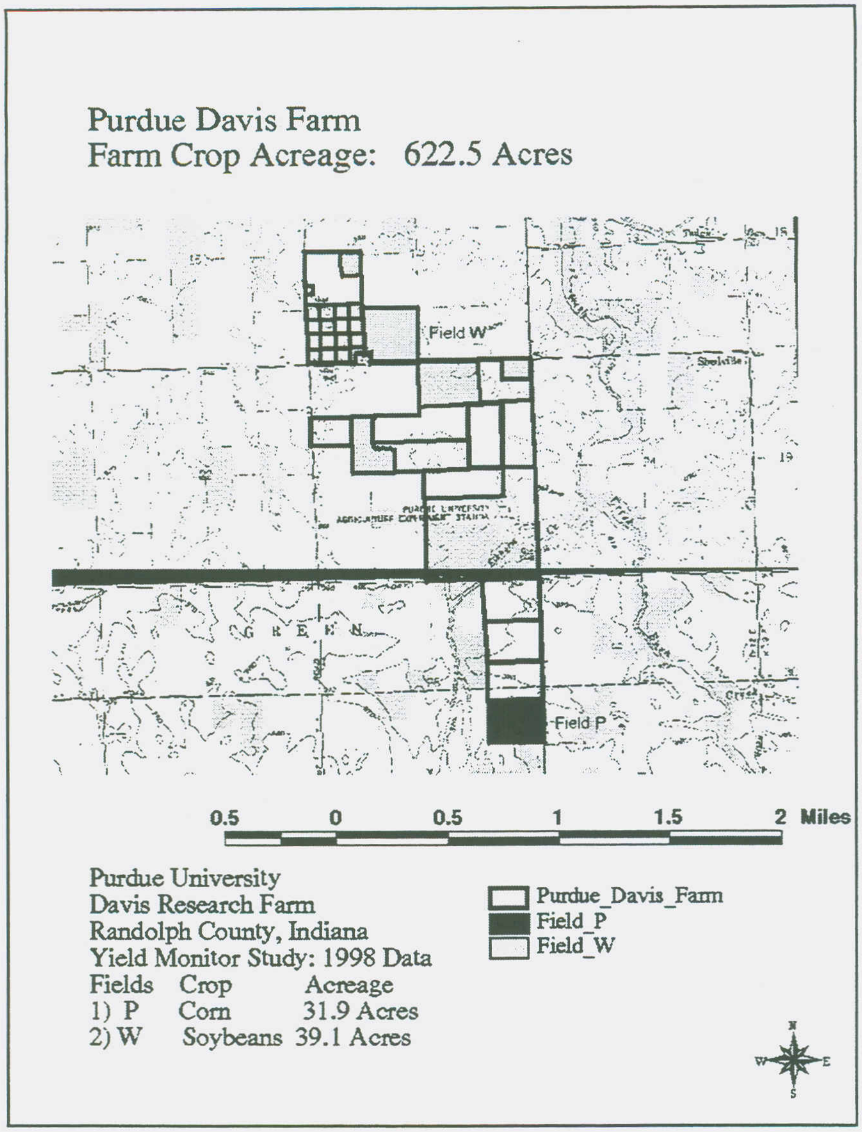

Figure 1. Purdue University Davis Research Farm with Study Fields P and W shown.

professors were Dr. John Trott, Director, Purdue Agricultural Centers, Dr. Chris Johannsen, Director, Laboratory for Applications of Remote Sensing, and Dr. Gaines Miles, Assistant Professor of Agricultural Engineering. A group of Purdue graduate students worked with the project as well. Two students, Patrick Willis and Montie O'Neal, wrote papers (Willis, 1999 and O'Neal, 2000a and 2000b).

Purdue University chose the Davis-Purdue Research Farm (DPAC) in Randolph County, Indiana (near Muncie, Indiana) as the study's site. The farm contains 622.5 acres of land with fields of corn and soybeans (Figure 1). Although many fields were the focus of research, Purdue University shared two years of yield monitor data for fields' P and W (Figure 1). NASS collected objective yield samples from both fields during 1997 and 1998. Purdue University obtained yield monitor data during 1997 and 1998 as well. However, the 1997 yield monitor data had additional flaws that made it unsuitable for this analysis.

Purdue University collected data from other sources including weigh wagon, remote sensing, and weather data. Patrick Willis (Willis, 1999) analyzed the available remote sensing data and yield monitor data for both fields. Montie O'Neal (O'Neal, 2000a) evaluated weather effects on crop yields for many fields within the Purdue Davis Research Farm. His second paper examined potential yield monitor errors (O'Neal, 2000b). This paper evaluates the relationships among the 1998 yield monitor data and the corresponding NASS objective yield samples for fields P and W.

2. DATA COLLECTION

This report focused on an analysis of the Purdue 1998 yield monitor (Figure 1) and NASS objective yield associated data sets for two Davis fields, P and W. Purdue University collected one major set of data associated with this project: yield monitor data from the Ag Leader yield monitor (Appendix A). University researchers also assisted with collection of the NASS objective yield site data (Appendix B) by providing GPS locations of the selected locations within the fields and with laying out the samples.

The yield monitor data for each recorded geographic location contains data from a larger area. Indeed, each individual reading contains portions of grain from many areas within the harvesting path. The harvester must fill the hopper before taking the first reading (lag), then the readings remain low as the hopper fills (ramp-up). As a pass ends, the yield monitor readings exhibit decreasing readings (ramp-down). Unequal grain movement can prevent associating areas from two consecutive GPS readings. High accuracy differential GPS using a stationary GPS unit can provide locations accurate to within five meters (Blackmore, 1999). Movement of the tractor at varying speeds as the operator adjusts for changes in the field conditions decreases positional accuracy. Improved procedures continue to increase the positional accuracy (Rawlins, 1996). Further automation of the harvester and other agricultural equipment will aid in these efforts.

Additional positional errors with the collected data can create inaccuracies in the calculated yields. These errors include the following: flyers (stray points from incorrect GPS readings), an overlay of consecutive points (from incorrect GPS antenna placement) and the equipment's inability to measure blank spots. Montie O'Neal (2000a) has recently catalogued many errors inherent in the yield monitor data. Each data set can exhibit all possible errors, but both the frequency and severity of errors can vary. The header not remaining full (often difficult to detect) can cause fluctuations in the measured weight of grain during harvest.

The resulting reduction in the accuracy of the yield monitor data from such errors does not invalidate the estimates of yields obtained from the yield monitor data. However, effective use of the yield monitor data requires a careful calibration with weigh wagon data. Of particular interest is the mapping scale at which the yield monitor data is of most value. Some researchers have evaluated the data at the one-meter level (Willis, 1999). However, additional characteristics of the data that suggest aggregating individual values to a larger area for evaluation are the following:

1) Inaccuracies from GPS errors (Nolan, 1996) and analysis of the Purdue data,

2) Possible mis-calibrations of the yield monitor sensors preventing accurate measurement of grain weight and moisture (Missotten, 1996),

3) A partially full header with no objective data to verify that it is full, as with Purdue, and

4) Considerable variations both in location and grain weight data (as with fields P and W).

Not all authors agree on an ideal sized area to aggregate the yield monitor data. For example, O'Neal (2000b) chose 9.13 meter grid cells corresponding to two passes of the harvester with a fifteen-foot header. One author (Dunn, 1998) established how rapidly the yield monitor would need to adjust to measuring field areas containing zero yield, normal yields, and 1.5 times normal yield for corn. He concluded that a minimum sized area would be a 100-foot grid cell that is equivalent to a 30-meter grid cell. Another author (Nolan, 1996) also stated larger areas provide greater accuracy. His estimate was that calculated yields will be within 5 per cent for a 400 square meter area (that is, a 20-meter grid cell). A 30-meter grid cell should provide an acceptable accuracy within that range,

Data collection accuracy limitations make an aggregate of the collected yield monitor data to a larger area for analysis essential. One goal was to avoid making corrections to the data locations beyond what could be easily evaluated. A second goal was to make the evaluation of the aggregated data more meaningful. Since LANDSAT TM data with a 30-meter pixel size has been successful in crop acreage estimation, a 30-meter aggregation area for yield monitor data would seem reasonable. Many yield monitor data sets would be needed to evaluate fully the strengths and weaknesses of other grid cell sizes.

2.1 NASS OBJECTIVE YIELD DATA COLLECTION: FIELDS P AND W

The number of objective yield plots varied somewhat for the two fields of interest. Laboratory counts are for only a portion of the sampled areas, though NASS takes field counts such as number of corn ears or soybean pods for the entire sampled area. Both sending samples for laboratory evaluation and the physical demands from collecting the data limited the number of objective yield sites to 58 in field P and to 64 objective yield sites in field W.

Ralph Gann, State Statistician, Indiana SSO, led 12 NASS enumerators in obtaining the USDA/NASS objective yield data. These objective yield data sets for the two fields provided a source of comparisons with the 1998 yield monitor data. An overlay with USGS digital maps at 1:24,000 scale to ensure accurate location of the data sets provided confirmation of the correct location and overlay of the data sets. The objective yield data nearly mimicked, with minor adjustments, the operational objective yield procedures as described in Reference 21.

3. PREPARATION OF YIELD MONITOR DATA FOR ANALYSIS

Before examining the available data sets, each data set was overlaid on a 7.5 minute map Digital Orthophoto Quadrangle (DOQ) obtained from USGS. This overlay process was necessary to ensure the following: the data sets were in the same projection; each data set properly overlaid all others with the map; and basic distances and areas were correct for each field. Creation of the grid data sets required all the original data sets to be in decimal degrees, with the ArcView (ESRI, GIS software) projection being the same as the USGS map. Specifically, the map projection chosen was the following: Universal Transverse Mercator (UTM) 1927, Clarke 1866 spheroid, and NAD27 Datum.

3.1 CORRECTION OF MAJOR POSITIONAL AND RECORDING ERRORS IN THE YIELD MONITOR DATA

Martin Ozga, Research and Development Division, developed the FLDHARV program (Appendix C) in Delphi 3. He translated algorithms into code intended to correct for lag, ramp-up, ramp-down, and zero data points. Preparing the data for use in ArcView 3.2 with the Spatial Analyst add-in (ESRI, 1996) remains a laborious process. However, FLDHARV did make possible consistent corrections of many data errors for fields P and W (Appendix C). Purdue did not ensure that harvest paths were straight nor prevent overlap of passes. Recent developments in associated precision agriculture (such as a light bar tracking system to align the harvester passes) may aid in improving the data positioning over what was observed in these data sets. No data was available to allow us to make corrections of this kind.

The FLDHARV program summed small data points corresponding to the ramp-up and ramp-down, eliminated zero data points corresponding to the lag, and adjusted the data values to find the starting values for each row more accurately. Elimination of many zero values and more proper placement of the observed values improved the appearance of the data. However, additional errors remained, since rows varied in their spacing width along the field's length.

After correcting major positional and recording errors in the yield monitor data, the first step was to map the initial values. The readings obtained were in pounds of crop. The usual way of printing out these preliminary values is to convert them to Bushels/Acre in a GIS. Since FLDHARV did not provide accurate calculations of the new distances between recorded points, accurate yield estimates were not possible on a location by location basis. Examining the yield monitor data revealed many cases of overlapping data, breaks in the data, open areas with no data values, and values outside the mapped field boundaries. Another component of the yield monitor data that was evident were changes in recorded weights throughout the fields and within each pass. Few adjacent recorded weights were the same so that calculated yields varied. Correcting every error in the data would require assumptions that are difficult to justify. Use of potential mapping corrected many of these errors and permitted analysis of the resulting grid data.

3.2 EVALUATION OF THE FLDHARV SUPPLEMENTARY OUTPUTS

The corrections made by FLDHARV showed considerable variability within each data file. Running each file individually within FLDHARV helped by permitting evaluation of data in the order recorded. The program output provided the following information:

1) Calculation of the number of points within a pass,

2) A count of zero weight values before a positive weight value (lags),

3) The number of points with positive values before more stable values start (ramp- up),

4) Evaluation of the reduced weights at the end of a pass (ramp-down), and

5) A count of the corresponding zero values with header-up as the pass ends (Appendix C).

3.3 POTENTIAL MAPPING OF THE YIELD MONITOR DATA

Although FLDHARV did not correct all errors in the data, potential mapping, (Blackmore and Marshal, 1996) made the data made useful. Blackmore and Marshal explained that potential mapping sums all available yield monitor weights within grid cells of a certain size. A 30-meter square (900 square meters) was the starting point for this analysis. Other researchers have considered areas as small as one-meter (Willis, 1999). However, this report focused on reducing overall errors in the yield monitor data with only small position corrections and crop weight corrections. Smaller areas would require more work with the data to correct it. Potential mapping did especially well in correcting for a partially full header, zero weights, and variations in positional accuracy. However, adjustments for lag, ramp-up and ramp-down were necessary to make potential mapping procedures work correctly.

Potential mapping required adding up the total yield within each 30-meter grid cell. Smaller grid cells require greater confidence in the accuracy of the crop weights measured, the GPS readings, the various corrections necessary to position the readings accurately, and a full header. The width of the header (15 feet) means that six passes of the combine are usually within a 30-meter grid cell. Since passes were not always in a straight line, using this size grid helped reduce the effects of various uncorrected data errors on the final analyses. Therefore, the evaluation of the data could focus on larger blocks of data rather than on the more variable individual readings.

4. EVALUATION OF THE POTENTIAL MAPPED YIELD MONITOR DATA

The process of potential mapping in this study was to create 30-meter square grid cells from the yield monitor data by aggregating the observed yield monitor data within the grid area. One concern with potential mapping for the yield monitor data in both fields was the presence of low data values and many zero weight values within the end row areas. The lower yield values in the end-rows should create an edge-effect for the eastern and western portions of field P. Similarly, the end-rows should give lower yields for the northern and southern portions of field W. Strong evidence that the yields were lower in the end-rows of field W comes from two NASS data points in the northern edge of the fields that were 134.7 and 0.2 Bu/Acre, respectively: both values were below the field average yield for the NASS data points. Indeed, one value was virtually zero with virtually no corn present. Potential mapping at 30-meter grid cell sizes reduced the number of observations within the edge grid cells for both fields since these grid cells could be outside the field.

4.1 EXAMINATION OF FIELD P POTENTIAL MAPPED YIELD MONITOR DATA

After correction of the original data set using FLDHARV, the remaining 19,419 weights show a total of 92,743 pounds of soybeans was obtained from field P. Without moisture adjustment, the yield would be 48.5 Bushels/Acre for the yield monitor without regard to the planned potential mapping of the data. As a comparison to this data, weigh wagon data provided a value of 45.5 Bushels/Acre while the mean yield of the NASS objective yield plots was 47.8 Bushels/Acre. An agreement within three Bushels/Acre is certainly acceptable at this scale since this difference would mean only 100 bushels difference at the field level.

Creating the grid cells for field P required summing the yield monitor weights (in pounds) on a 30-meter grid to obtain total weights. Field P (planted with soybeans for 1998) with 31.9 acres produced an initial grid with 11 rows and 14 columns for a total of 154 grid cells. After using the average moisture values and adjusting to a standard moisture of 12.5 per cent, the total weights for each grid were converted to acres (a factor of 4.49636 Grid Cells/Acre) and divided by 60 pounds per bushel within ArcView's Spatial Analyst. These calculations produced estimated soybean bushels per acre yields for each grid cell.

After creating the grid data set, ArcView's Spatial Analyst was used to create contours using the full set of grid cells at a five (5) Bushel/Acre resolution. These contour maps would help in evaluating the accuracy of this first attempt. An edge effect on the eastern edge of field P is especially apparent in the closely spaced contours that ramp-up quickly from a minimum of five (5) Bushels/Acre to 35 Bushels/Acre (Figure 6, Appendix D). The remainder of the field shows less variation except one section in the northeastern part of the field that had a much lower yielding grid surrounded by larger yielding grid cells (Ibid.).

4.2 EXAMINATION OF FIELD W POTENTIAL MAPPED YIELD MONITOR DATA

After correction of the original data set using FLDHARV, the remaining 22,569 recorded weights totaled 299,275 pounds of corn for field W. Without moisture correction, the yield would be 134.1 Bushels/Acre without considering the planned potential mapping of the data. The weigh wagon estimate was 139.9 Bushels/Acre (Willis, 1999) whereas the mean yield for the objective yield sample lots was 153.3 Bushels/Acre.

The mean of objective yield plot estimates for field W was 13.4 Bushel/Acre greater than the weigh wagon data and its 95% confidence interval of 150.9 to 161.2 Bushels/Acre excluded the weigh wagon estimate as well. This lack of agreement is somewhat disconcerting and unexpected since the number of sample plots should have been sufficient to estimate the field corn yield with sufficient precision. Only the one end-row value appeared to be an extreme outlier, but that was exluded from analysis .

Creating the grid cells for field W required summing the yield monitor weights (in pounds) on a 30-meter grid to obtain total weights by grid cell as for field P. Field W (planted to corn for 1998) had 39.1 acres that produced an initial grid containing 14 rows and 14 columns for a total of 196 grid cells. After using the average moisture values and adjusting to a standard moisture of 15.5 per cent, the total weights for each grid were expanded to acres (a factor of 4.49636 Grid Cells/Acre) and divided by 56 pounds per bushel. These calculations produced estimated corn bushels per acre yields for each grid cell.

After creation of the grid data set, contours were generated from the full set of grid cells at ten (10) Bushels/Acre resolution to help in evaluating the accuracy of this first attempt. The edge effects for all contours near the field W boundaries are evident from the closely spaced contours that ramp-up quickly from minimums of 40 Bushels/Acre in the southern portion of the field, 60 Bushels/Acre in the western and northern sections, and 90 to the east. The central sections of the field show contours that generally range from 120 to 170 Bushels/Acre (Figure 7, Appendix D).

Two causes of the reduced yields in the southern and northern edges of the field are apparent. The first is that these edges of the field contain end-rows which are typically lower yielding than the remainder of the field. The second cause of reduced yields for these grid cells is that the number of observations in these grid cells is less than for the remainder of the field.

5. FIELD P COMPARISON OF NASS OBJECTIVE SOYBEAN YIELDS WITH YIELD MONITOR YIELDS

For the purpose of making comparisons between the NASS objective yield site yield estimates and the grid potential mapped yield monitor yield estimates, the 30-meter yield monitor grid cell containing the NASS objective yield value was selected. Although other methods are possible, this procedure should be the simplest and provide comparable yield estimates. Considerations of the yield monitor data regarding the number of observations that went into the yield calculation are possible. However, the first check would be with all available data from the NASS objective yield sites compared with the corresponding yield monitor 30-meter grid cells.

Figure 2, Appendix D provides the NASS objective yield locations with the NASS objective yield values. Figure 4, Appendix D gives a plot of the yield monitor yields versus the NASS Objective yield calculated yields. Table E.1, Appendix E. provides summaries for the two data sets of all 56 NASS objective yield sites after deletion of the two end-row observations.

Next the analysis continued with creating a numbered row and column grid for the data values starting in the south-west corner of the field. The purpose of this process was to permit analysis of the data according to rows and columns to determine if there were consistent differences between different regions of the field in terms of observed yield values. Significant differences across the field would provide evidence of spatial dependence of the yield estimates and permit evaluation of causes for those differences. Multivariate Analysis of Variance (MANOVA) is one method to evaluate the correlated yield values according to the grid of row and column values.

Two MANOVA analyses provided the means to evaluate the spatial relationships among the yield estimates within the field. The first MANOVA relates both the YIELD and YMYLD variables to the row factors. The row factors' MANOVA evaluated the degree of change in the North-South direction within the field while the column factors' MANOVA evaluated the East-West variation.

The row factors MANOVA presented in Table E.2.A of Appendix E was significant at the p = 0.009 level, thereby indicating a strong presence of North-South yield variation among the defined rows. However, Table E.2.B. of Appendix E provides the same MANOVA analysis for the column factors that shows a non-significant p-value of 0.313. Therefore, there appears to be a North-South component of change in yield estimates for the yield monitor 30-meter grid within the field, but not a corresponding East-West component of change.

The next question is which variable(s) show(s) a difference among rows. A row factors Analysis of Variance (ANOVA) provided a means of evaluating these differences. Table E.3. of Appendix E provides the two ANOVAs for row and column factors, respectively. The row effects' ANOVA for YIELD is not significant with a p-0.32 level test value. However, the row effects' ANOVA for YMYLD is significant at the p = 0.001 level indicating a significant trend in rows North-South for the yield monitor 30-meter grid cells at the NASS objective yield sites.

Continuing this rows effects' ANOVA evaluation for YMYLD, a multiple-comparisons evaluation of the row means showed the following: Row eight differs significantly from all rows except row two. A probable explanation of these results is evident, since the density command in Arc-View's Spatial Analyst showed that the number of points in some grid cells on the northern edge of field P had fewer yield monitor values on average than did other grid cells within the field. For this reason, calculated yield monitor 30-meter grid cell yields for row eight (northern edge of the field) have lower yields. Consequently, the row eight yield monitor yield estimates' mean differs statistically from the other rows, except row two (2).

6. FIELD W COMPARISON OF NASS OBJECTIVE YIELD CORN YIELDS WITH YIELD MONITOR YIELDS

Analyses of Field W proceeded in a similar manner as that in Field P. Two sample locations, locations 63 and 64, were in the end rows of the northern part of the field and were deleted to exclude end-row effects. A row and column number for the remaining sample locations provided the means for an MANOVA to evaluate the relationship between the yield monitor 30-meter grid corn yield estimates and the objective yield corn yield estimates. Table E.4. of Appendix E shows that the means are134.2 for the yield monitor estimates and 156.1 for the objective yield estimates. The weigh wagon had an estimate of 139.9 Bushels/Acre. The yield monitor 95% confidence interval includes the weigh wagon estimate. However, the NASS objective yield locations create a 95% confidence interval of (150.9, 161.2) Bushels/Acre with a mean of 156.1 Bushels\Acre that does not include the weigh wagon estimate. A correlation of 0.29 showed that the two sets of yield estimates were not highly correlated as indicated in the graph of yield monitor at 30-meter grid corn yields versus the objective yield estimates (Figure 5, Appendix D).

The MANOVA for the objective yield corn yields with the yield monitor corn yields at 30-meter grid cells versus rows is not significant at p = 0.17. However, the columns' MANOVA is significant at p = 0.0004 (Table E.5, Appendix E). Supplementary ANOVA's for the individual objective yield estimates and yield monitor 30-meter grid yield estimates show non-significant p values except for the column ANOVA of yield monitor yield estimates with a p-value of 0.0007(Table E.6, Appendix E). A subsequent multi-comparison procedure showed that column one differed with columns six and seven while column two differed with column eight. Although these results are an indication of East-West differences, the relationships are not as strong as in the case of field P. There is evidence, however, of there being fewer observations than the field average within the 30-meter grid cells on the western edge of the field.

7. DISCUSSION

This report has examined relationships among three sets of data: 1) yield monitor data, 2) weigh wagon data, and 3) NASS objective yield sample locations. These sources of data were available for two fields, P and W, on Purdue University's Davis Research Farm for 1998. The purpose of this study was to provide an analysis of relationships among the three sets of data for each field and assist in showing how yield monitor data might aid in providing a better understanding of soybean and corn yield variability within a field.

7.1 OPINIONS DERIVED FROM THIS STUDY

At this time, NASS cannot conclude that yield monitor data will be an important data set in determination of U.S. or State yields. Variability in the recorded data observations, potential errors in analysis, and overall complexity of the yield monitor data would argue against its use in any surveys obtaining crop yields. Yield monitor data cannot at this time be a substitute for weigh wagon information, crop marketing (sales tickets) data, or on-farm storage volumes. Nor can it be a useful input into remote sensing studies without overcoming its multiple shortcomings. While these multiple limitations of the yield monitor yields continue to be significant, NASS data collection of farmers' reported data from yield monitors is not recommended. Monitoring the degree of adoption of yield monitor data and its importance to farmers should be, at this time, the only potential interests that NASS should maintain.

The results of this study show quite different relationships between the NASS objective yield 30-meter interpolated grid cells and Purdue yield monitor 30-meter potential mapped grid cells for the two fields in the study. The objective yield grid cells for the NASS soybean objective yield locations have a similar mean as the Purdue soybean yield monitor 30-meter grid cells. However, the NASS objective yield values have a greater range and low correlation of -0.12 with the Purdue yield monitor 30-meter potential mapped soybean grid cells at the objective yield sample locations. These facts would argue that the two procedures are estimating different processes for field P. Conversely, field W NASS estimated objective yield values at the objective yield sample locations have a larger mean than do the Purdue corn yield monitor 30-meter potential mapped grid cells. They also have a positive correlation of 0.29 and a 95 per cent confidence interval for the mean yield with nearly the same range (not overlapping)as for the Purdue yield monitor corn 30-meter potential mapped grid cells at the objective yield sample locations. These results seem to suggest that the two processes are similar, but that the NASS objective yield method of estimating the number of plants or ears per acre produced a higher corn yield for field W.

The yield monitor data does not appear to represent field W in the same way as does the NASS objective yield sample. Although this degree of difference may not always occur, this field provides a contrary example as follows: The yield monitor will not measure all fields in a way that will generate estimates close to those obtained from objective yield measurement sites. Complexity of the yield monitor data and the need to make frequent calibrations would argue that this might be a frequent occurrence.

Factors that reduce the variability of crop yields across the field will help the effectiveness of the NASS objective yield methodology. Any detrimental soil or drainage conditions within the field can increase the crop yield variability and thereby reduce the effectiveness of the NASS objective yield methodology. How well farmers follow good farming practices in the management of their fields by proper field preparation and the provision of required crop nutrients will also play a role in the range of crop yields observed within the field.

One remark regarding the range of crop yields observed with a field seems evident from the NASS objective yield sample data and the yield monitor data. Specifically, the possibility exists for rather large degrees of crop yield variability within a field. End rows, in particular, tend to have particularly large variations in crop yields, due to stress from weather conditions and poor soil conditions. The NASS objective yield does not address the possibility of the potentially large range of crop yields both at the sample level and the field level that are observed here. Indeed, evidence from these data sets shows that wide ranging yield values can be present in the fields. The Purdue field W corn data, for example, shows the difficulty of estimating the field level corn yield using objective yield data. At the very least, the within field yield variability will increase the confidence interval range of the estimate determined by the NASS objective yield sample.

The author can catalog just a few of the possible difficulties that might cause the NASS objective yield methodology challenges, over time, as follows:

1) Widespread Planting of Numerous Crop Varieties within a State:
New crop varieties may change germination rates as well as variations in the number of the crop plants throughout the field in response to soil variables, weather, and planting rate,

2) Modern Farming Practices that Emphasize Denser Planting Rates:
Changes in the planting rate within the field can have profound effects on the resulting density of plant stands and thereby the resulting crop yields throughout the field, and

3) Maintaining Field Drainage by Repair of Field Drainage Tiles:
Improved drainage throughout the field with no broken drainage tiles should be a factor in reducing yield variability throughout the field as, for example, broken tiles had caused decreased yields in portions of Purdue field P (Willis, 1999).

Any of the above-enumerated possibilities can affect the accuracy of the NASS objective yield. Additional studies would be necessary to evaluate how consistent the results of this study would be under different cropping conditions.. The small area of the objective yield plot sizes and sample sizes that are necessary due to cost constraints involved in making objective yield data collections, will likely continue to be a limitation for NASS in the foreseeable future. However, improved farming practices should aid in all four areas listed above and thereby reduce the range of observed field yields.

The most important component of yield monitor information for NASS to examine further at this time would be a possible joint follow-up survey with ERS involvement, every three years or so, to obtain the following information:

1) Survey farmers to determine those who have adopted use of yield monitors,

2) Question those farmers using yield monitors as to how valuable they have found the use of yield monitor data to be for them,

3) Question these same farmers regarding any difficulties that they might be having in using information obtained from yield monitors, and

4) Finally, question the degree of agreement that the farmers have found that their yield monitor yields have had to weigh wagon data, crop sales tickets, or on-farm storage volume.

9. REFERENCES

1. Ag Leader Technology, "Yield Monitor 2000, Operators Manual," Ames, Iowa, June 1997.

2. A and L Laboratories, "Soils Report on Field P and W," Fort Wayne, Indiana, August 1998.

3. Birrell, S. J. and K. A. Sudduth, "Nutrient Mapping Implications of Short-Range Variability," Precision Agriculture, Proceedings of the 3rd International Conference, American Society of Agronomy, Crop Science of America, and Soils Science Society of America, Minneapolis, Minnesota, p. 207, June 23-26, 1996.

4. Blackmore, S. and M. Moore, "Remedial Correction of Yield Map Data," in Precision Agriculture, Pierre Robert (Editor-In-Chief), Volume 1, Number 1, p.53, January 1999.

5. Brown, J. R., Editor, "Recommended Chemical Soil Test Procedures for the North-Central Region," Missouri Agricultural Experiment Station, January 1998.

6. Cressie, N. A. C., "Statistics for Spatial Data," Revised Edition, New York, John Wiley and Sons, 1993.

7. Doucette, P. and Beard, K., "Exploring the Capability of Some GIS Surface Interpolators for DEM Gap Fill," Photogrametric Engineering & Remote Sensing, ASPRS, vol. 66, no. 7, p. 881, July 2000.

8. Dunn, R. F., "Where Oh Where Does the Grain Go?", in Precision Ag Illustrated, Clear Window Multimedia, St. Louis, MO, December 1998.

9. Environmental Systems Research Institute, Inc., ArcView Spatial Analyst, Applied Spatial Analysis Using Raster and Vector Data, Redlands, California, 1996.

10. Huddleston, H. "Sampling Techniques for Measuring and Forecasting Crop Yields," Economics, Statistics, and Cooperative Service, USDA, 193 pages, 1978.

11. Mallarino, A., E. Oyarzabal, and Hinz, P., "Interpreting Within-Field Relationships Between Crop Yields and Soils and Plant Variables Using Factor Analysis," in Precision Agriculture, Pierre Robert (Editor-In-Chief), Volume1, Number 1, p. 15, 1999.

12. Missotten, B., Strubbe, G., de Baedemaeker, J., "Accuracy of grain and straw yield mapping," Precision Agriculture, Proceedings of the 3rd International Conference, American Society of Agronomy, Crop Science of America, and Soils Science Society of America, Minneapolis, Minnesota, p. 713, June 23-26, 1996.

13. Nolan, S., Haverland, G., Goddard, T., Green, M., Penney, D., Henriksen, J., Lachapelle, G., "Building a Yield Map from Geo-referenced Harvest Measurements," Precision Agriculture, Proceedings of the 3rd International Conference, American Society of Agronomy, Crop Science of America, and Soils Science Society of America, Minneapolis, Minnesota, p. 885, June 23-26, 1996.

14. O'Neal, M., Frankenberger, J., Parson, S., Ess, D., Crisler, M., Strickland, R., "Correcting yield monitor data for improved yield mapping," Presented at the 2000 ASAE Annual International Meeting, Paper No. 001088, ASAE, St. Joseph, MI USA.

15. O'Neal, M., Frankenberger, J., Ess, D., Grant, R., "Spatial Variability of Precipitation as a Factor of Yield Variability," Presented at the 2000 ASAE Annual International Meeting, Paper No. 991144, ASAE, St. Joseph, MI USA.

16. Personal Communication, Mark Schleusener, NASS.

17. Pocknee, S., and others, "Directed Soil Sampling," Precision Agriculture, Proceedings of the 3rd International Conference, American Society of Agronomy, Crop Science of America, and Soils Science Society of America, Minneapolis, Minnesota, p. 159, June 23-26, 1996.

18. Precision Agriculture, Proceedings of the 3rd International Conference, American Society of Agronomy, Crop Science of America, and Soils Science Society of America, Minneapolis, Minnesota, June 23-26, 1996.

19. Rawlins, S., "Moving from Precision to Prescription Farming: The Next Plateau," Precision Agriculture, Proceedings of the 3rd International Conference, American Society of Agronomy, Crop Science of America, and Soils Science Society of America, Minneapolis, Minnesota, p. 283, June 23-26, 1996.

20. SISA, Simple Interactive Statistical Analysis, http:www.home.clara.net/sisa/bonhlp.htm.

21. Statistical Methods Branch, "The Yield Forecasting and Estimating Program of NASS," Washington, D.C., June 1998. NASS Staff Report Number SMB 98-01.

22. Top-Soil Testing Service, "Soil Test Report," Frankfort, Illinois, Spring, 1994.

23. Willis, P., "Evaluation of Crop Yield Estimation Using Yield Monitor Data, Remotely Sensed Imagery, and Hand-Harvest Data," M.S. Thesis, Purdue University, 1999.

APPENDIX A: YIELD MONITOR DATA CHARACTERISTICS AND PREPARATION

Yield monitor data has many potential errors that make its use challenging. However, careful preparation of the data can reduce some errors. The simplest errors to correct are those of the lag (zero weight values at the start of a pass), ramp-up (increasing weight values as the harvester's hopper fills), and ramp-down (decreasing weights as the hopper empties at the end of a pass). These errors are most evident from evaluating the data. The most efficient way of adjusting these errors was to develop a computer program.

FLDHARV, a program written in Delphi 3 by Martin Ozga, corrects only these most troublesome characteristics of the yield monitor data. The program does not correct the following: a partially full header, stray points from loss of the GPS signal, and incorrect distances between the recorded points. Potential mapping, aggregation of the adjacent data points, reduces the overall measurement errors. A 30-meter grid size increased this averaging effect, so we did not evaluate smaller grid cell sizes. This grid cell size would permit comparison with Landsat Thematic Mapper data that also has a 30-meter pixel size.

Purdue University collected yield monitor data with an AgLeader 2000 yield monitor (Willis, 1999). The latitude and longitude data came from a Vision System Omnistar 4000 DGPS with sub-meter accuracy. They used a Case 1460 harvester with a 15-foot wide (4.572 meters)or a six-row header to harvest the crop. Table A.1 provides an example of the collected data. Although other data is in the file, the analysis of the data concentrated on the flow (weight of grain). Of course, GPS provided latitude and longitude coordinates for each recorded point. The flow is the harvested crop's weight measured by a sensor on the harvester for the corresponding area. Moisture readings provided the percentage of crop moisture. The header records provide information on the position of the harvester's header for being up (not harvesting) or down (harvesting).

Field Id provides the location of the harvester as within the end-rows (End) or within the primary part of the field (Bulk). The time shows each location being taken one-second after the preceding location. Cycles confirm that each recording was at one-second intervals. The distance provides a calculated distance between recorded locations. Pass remains one throughout the field and so did not describe the travel of the harvester well. Swath stays 15 to represent the intended width in feet of the harvester header. The serial number, zero (0), was the serial number of the yield monitor. Load ID remained one. Grain signified soybeans for field P and corn for field W. DGPS gave a quality value for the GPS signal while the PDOP gave a measure of the GPS signal strength. The values for DGPS and PDOP remained good for both fields P and W the GPS signal did not contribute to inaccuracies in recording locations. Finally, the elevation values provided the altitude of each location. Since Field W was very level, the evaluation values changed little while field P had more variability in readings. Since altitude readings from GPS devices tend to be inaccurate, no attempt was made to analyze these measurements.

Table A.1. Uncorrected Yield Monitor Data Example: Purdue Davis Farm Field P, 1998

Longitude	Latitude	Flow	Time	Cycles	Distance	Swath	Moisture	Header	Pass	Serial Num	Field ID	Load ID	Grain	DGPS	PDOP	Elevation
-85.152095	40.241835	0	148885	1	0	15	12.7	Up	1	0	Bulk	1	Beans	2	2.3	298
		Forty Observations with zero Flow Readings														
-85.151743	40.241840	2.86	148931	1	61	15	12.8	Down	1	0	Bulk	1	Beans	2	2.3	298
-85.151707	40.241838	2.84	148933	1	61	15	12.7	Down	1	0	Bulk	1	Beans	2	2.3	298
-85.151688	40.241838	2.19	148934	1	60	15	12.7	Down	1	0	Bulk	1	Beans	2	2.3	298
-85.151688	40.241838	2.67	148934	1	59	15	12.6	Down	1	0	Bulk	1	Beans	2	2.3	298
-85.151653	40.241838	4.41	148936	1	60	15	12.7	Down	1	0	Bulk	1	Beans	2	2.3	298
-85.151637	40.241838	3.44	148937	1	60	15	12.7	Down	1	0	Bulk	1	Beans	2	2.3	299
-85.151618	40.241838	3.12	148938	1	59	15	12.7	Down	1	0	Bulk	1	Beans	2	2.3	298
-85.151600	40.241840	3.45	148939	1	60	15	12.9	Down	1	0	Bulk	1	Beans	2	2.3	299
-85.151600	40.241840	2.87	148939	1	60	15	12.8	Down	1	0	Bulk	1	Beans	2	2.3	299
-85.151565	40.241840	2.59	148941	1	60	15	12.8	Down	1	0	Bulk	1	Beans	2	2.3	299
-85.151550	40.241840	4.54	148942	1	60	15	12.9	Down	1	0	Bulk	1	Beans	2	2.3	299
-85.151550	40.241840	4.11	148942	1	61	15	12.7	Down	1	0	Bulk	1	Beans	2	2.3	299
-85.151513	40.241840	2.87	148944	1	61	15	12.9	Down	1	0	Bulk	1	Beans	2	2.3	299
-85.151495	40.241840	4.41	148945	1	61	15	12.9	Down	1	0	Bulk	1	Beans	2	2.3	299
-85.151495	40.241840	4.83	148945	1	61	15	12.7	Down	1	0	Bulk	1	Beans	2	2.3	299
-85.151460	40.241842	4.29	148947	1	60	15	12.8	Down	1	0	Bulk	1	Beans	2	2.3	299
-85.151442	40.241842	3.11	148948	1	61	15	12.9	Down	1	0	Bulk	1	Beans	2	2.3	299
-85.151423	40.241843	2.89	148949	1	61	15	12.9	Down	1	0	Bulk	1	Beans	2	2.3	299
-85.151405	40.241843	2.42	148950	1	61	15	13.0	Down	1	0	Bulk	1	Beans	2	2.3	299
-85.151388	40.241843	2.36	148951	1	61	15	12.9	Down	1	0	Bulk	1	Beans	2	2.3	299
-85.151372	40.241843	1.96	148952	1	61	15	13.0	Down	1	0	Bulk	1	Beans	2	2.3	299
-85.151353	40.241845	2.36	148953	1	60	15	13.1	Down	1	0	Bulk	1	Beans	2	2.3	299
-85.151353	40.241845	2.52	148953	1	61	15	12.9	Down	1	0	Bulk	1	Beans	2	2.3	299
-85.151317	40.241845	2.54	148955	1	60	15	13.0	Down	1	0	Bulk	1	Beans	2	2.3	298
-85.151300	40.241845	2.27	148956	1	60	15	12.8	Down	1	0	Bulk	1	Beans	2	2.3	299
-85.151300	40.241845	2.49	148956	1	60	15	12.9	Down	1	0	Bulk	1	Beans	2	2.3	299
-85.151265	40.241847	2.55	148958	1	58	15	12.7	Down	1	0	Bulk	1	Beans	2	2.3	298
-85.151250	40.241845	2.66	148959	1	57	15	12.8	Down	1	0	Bulk	1	Beans	2	2.3	299
-85.151250	40.241845	2.56	148959	1	57	15	12.8	Down	1	0	Bulk	1	Beans	2	2.3	299
-85.151215	40.241847	2.60	148961	1	55	15	12.7	Down	1	0	Bulk	1	Beans	2	2.3	299
-85.151200	40.241845	2.63	148962	1	56	15	12.6	Down	1	0	Bulk	1	Beans	2	2.3	299
-85.151182	40.241845	3.15	148963	1	57	15	12.8	Down	1	0	Bulk	1	Beans	2	2.3	299
-85.151165	40.241845	2.80	148964	1	57	15	12.8	Down	1	0	Bulk	1	Beans	2	2.3	299
-85.151165	40.241845	4.16	148964	1	57	15	12.7	Down	1	0	Bulk	1	Beans	2	2.3	299
-85.151133	40.241845	4.71	148966	1	56	15	12.5	Down	1	0	Bulk	1	Beans	2	2.3	299
-85.151117	40.241847	5.36	148967	1	56	15	12.6	Down	1	0	Bulk	1	Beans	2	2.3	299
-85.151117	40.241847	5.79	148967	1	55	15	12.6	Down	1	0	Bulk	1	Beans	2	2.3	299
-85.151083	40.241848	5.64	148969	1	56	15	12.6	Down	1	0	Bulk	1	Beans	2	2.3	299
-85.151068	40.241847	4.74	148970	1	56	15	12.7	Down	1	0	Bulk	1	Beans	2	2.3	299
-85.151052	40.241847	4.01	148971	1	56	15	12.8	Down	1	0	Bulk	1	Beans	2	2.3	299
-85.151035	40.241845	4.64	148972	1	56	15	12.7	Down	1	0	Bulk	1	Beans	2	2.3	299
-85.151020	40.241845	5.61	148973	1	56	15	12.8	Down	1	0	Bulk	1	Beans	2	2.3	300
-85.151020	40.241845	5.71	148973	1	56	15	12.7	Down	1	0	Bulk	1	Beans	2	2.3	300

*** AgLeader 2000 Yield Monitor soybeans' data for Purdue University field P for the 1998 crop (AgLeader, 1997).**

APPENDIX B: NASS OBJECTIVE YIELD SAMPLES: FIELDS P AND W

Collection of the NASS objective yield samples for fields P and W required considerable effort by a group of twelve NASS enumerators led by Ralph Gann. The Illinois State Statistical Office (SSO) evaluated the data to determine grain counts on the soybeans pods (field P), the corn ears (field W), and moisture measurements for both. The enumerators counted the plants of soybeans and stalks of corn in the field. Table B1 provides descriptive statistics of soybean data for field P and Table B2 provides field W statistics for corn.

Table B.1. Field P 1998 Soybeans NASS Variables Descriptive Statistics

	Variables[1]										
	RA	RB	RC	RD.	PA	PB	APd	BPd	PdD	PdU	Mst	AYld
Min	0	0	1	2	34,848	34,848	0	0	21	0	7.8	2.7
Max	21	17	17	15	302,016	302,016	155.5	147.5	447	4	8.2	101.1
Mean	7.3	8.2	8.6	8.2	187,959	187,960	72.1	76.2	184.0	0.74	8.0	47.8
SD	4.5	3.9	3.4	3.1	54,805	54,806	36.1	33.6	86.9	1.0	0.12	22.2
LCL	6.2	7.2	7.7	7.4	173,548	173,549	62.6	67.2	161.2	0.48	8.0	41.9
UCL	6.5	9.3	9.5	9.0	202,369	202,370	81.6	85.1	206.9	1.00	8.0	53.6

[1] RA = Plants in Row A, RB = Plants in Row B, RC = Plants in Row C, RD. = Plants in Row D, PA = Population Row A, PB = Population Row B, APd = Row A Pods, BPd = Row B Pods, PdD = Pods Rows A and B Summed, PdU = Ratio of Pods with Beans to Total Fruit, Mst = Pod Moisture, AYld = Calculated Soybeans Yield Bushels\Acre after deletion of 2.02 Bushels\Acre Harvest Loss Average.

Table B.2. Field W 1998 Corn NASS Variables Descriptive Statistics

	Variables[1]											
	Stlk1	Stlk2	Er1	Er2	ErW1	ErWt	GrnWt	Moist	EarsA	WtEar	Fract	lbEar	AYld
Min	16	9	3.0	3	0.2	54	37	11.7	3484.8	0.067	0.69	0.47	0.2
Max	33	27	26.0	27	11.3	549.7	454.9	15.8	30,201.6	0.491	0.85	0.43	201.4
Mean	23.7	23.3	23.2	22.8	9.0	326.7	268.9	13.4	26,698.7	0.385	0.82	0.32	153.3
SD	2.0	2.8	2.9	3.5	1.5	82.6	70.4	0.89	3455.4	0.06	0.029	0.011	28.1
LCL	23.2	22.6	22.4	21.9	8.6	306.1	251.3	13.1	25,835.5	0.37	0.81	0.31	146.3
UCL	24.2	24.0	23.9	23.7	9.3	347.3	286.5	13.6	27,561.8	0.40	0.83	0.34	160.3

[1] Stlk1 = Stalks in Row 1, Stlk2 = Stalks in Row 2, Er1 = Number of Ears in Row 1, Er2 = Number of Ears in Row 2, ErW1 = Ears Weight in Row 1, Er Wt = Ear Weight Total for Rows 1 and 2, GrnWt = Grain Weight from Ears in Row 1 plus Row 2, Moist = Moisture of Grain, EarsA = Number of Estimated Ears\Acre (Calculated), WtEar = Weight per Ear, Fract = Shelling Fraction (Calculated), lbEar = Estimated weight of shelled corn from one ear (pounds), Ayld = Calculated Corn Yield Bushels\Acre after deletion of 2.7 Bushels\Acre Harvest Loss Average.

APPENDIX C: FLDHARV: A PROGRAM TO CORRECT LAG, RAMP-UP, AND RAMP-DOWN OF YIELD MONITOR DATA

The purpose for FLDHARV was to correct the data locations for the lag, ramp-up, and ramp-down errors that are evident in all uncorrected yield monitor data-sets. Although many programs are commercially available to do this preprocessing, such programs usually make assumptions about the lag by using a constant value. Ramp-up and ramp-down variations are not usually considered for correction.

Yield monitor use continues to be operator dependent for the saving of data. Ten files of uncorrected data for field P corresponded to the following: (1) two files for the end-rows, (2) six files for the main section of the field, and (3) two files with only zero flow readings. Similarly, field W had 14 files with the following: (1) three files for the end-rows, (2) nine files for the main section of the field, and (3) two files with all zero flow readings. Because the data within each file was not always adjacent, the processing of the data went from one file at a time. Of course, the category (3) files with zero flow readings were not used.

Development of FLDHARV came from some basic considerations about the data from field W. First, the definition of a row was set as starting when a header down occurred (no matter what flow value that observation had - including zero). Next, the program counted the number of downs with zero flow readings, the number of ups, and the number of ups with zero flow readings. Algorithms to adjust for the ramp-up and ramp-down effects had three cases:

1) Zero < = Number of header down observations with zero flows Minus Number of header up observations with positive flow < = 5,

2) -5 < = Number of header down observations with zero flow Minus Number of header up observations with positive flows < = -1, and

3) If (1) and (2) are both false, then provide no correction to the recorded data.

For case (1), FLDHARV added together flow values for beginning and ending flow observations to adjust the number of observations with both header down and zero flow at the beginning of a pass to equal the number of header ups with zero flows at the end of a pass. The purpose was to prevent corrections for the location of lag at the beginning of the pass from displacing the locations of points at the end of the pass Case (2) required adding together more values to adjust for the ramp-up and fewer values for the ramp-down. Short passes were the most frequent cause for case (3) where no corrections are necessary.

FLDHARV worked better in field W. However, field P had greater variability in the number of header downs and so had more variable results. Since evaluation of the data within Arcview showed that the corrections were satisfactory for input into potential mapping, no changes to FLDHARV were necessary for field P data. The great amount of variability within the individual files for the number of zero downs (lags) within individual passes shows that an assumption of a standard lag adjustment would often be incorrect and thereby provide reduced accuracy.

Output from FLDHARV consists of two files. The first file contains simple counts of the output files of corrected data (Table C.1.). The second file contained the following statistics: minimum, maximum and mean number of output header downs, the input number of header downs with zero flows, the input number of header ups, and the input number of header ups with zero flows for each row (Table C.2).

Table C.1. Purdue Field P: An example of count data output from the FLDHARV program[1, 2]

File	Passes	H.Downs	Downs.0	H.Ups	Ups.0
2	1	259	15	17	11
2	2	27	12	15	11
2	3	179	15	14	10
2	4	11	11	12	6
2	5	210	12	11	9
2	6	13	12	12	11
2	7	10	10	11	10
.....	......	....	.....	....	....
2	8	208	13	13	11
3	1	199	16	1	1
3	2	31	0	5	5
9	24	207	12	14	12
9	25	32	12	1	1
9	26	161	0	19	14
Totals		Note: 2			
1038	3008	19671	1616	1664	1366

[1] Files = The File Number for Field P with Yield Monitor Data, Passes = Pass Defined by FLDHARV as defined above, H.Downs = Output Header Downs for the designated Pass, Downs.0 = Input Weights with Harvester Header in Down Position with Zero Flows for the designated Pass, H.Ups = Input Weights with Harvester Header in Up Position for the designated Pass, and Ups.0 = Input Weights Equal Zero (0) with Harvester Header in Up for the designated Pass.

2. The reduced number of weights (19,419) mentioned in the text was the result of deletion of 252 zero weights missed by FLDHARV.

Some smaller rows (less than 100 observations) are the result of the harvester stopping and\or lifting the header possibly to avoid obstacles within the field. Adjustments to such passes within a file that had less than one-half of the median number for that file were not corrected since there would be the uncertainty of stopping and starting points for the pass and its relationship to other passes. Similarly, many observations exceeding three times the median within that file were not corrected either. There were few such passes within a field.

Table C2. Purdue Field P: An example of statistics for the File Output Information from the FLDHARV program[1]

File	Passes	Statistics	H.Downs	Downs.0	H.Ups	Ups.0
2	8	Min	10	10	11	6
NA	NA	Max	259	15	17	11
NA	NA	Mean	114.6	12.5	13.1	9.9
3	3	Min	31	0	1	1
NA	NA	Max	199	16	11	9
NA	NA	Mean	87	6	5.7	5
......	......	......	......	.....	......	....
8	68	Min	13	0	1	1
NA	NA	Max	283	16	37	
NA	NA	Mean	119.7	8.6	10.5	8.6
9	26	Min	7	0	0	0
NA	NA	Max	248	39	19	15
NA	NA	Mean	146.2	14.2	9.6	7.9

[1] Files = The File Number for Field P with Yield Monitor Data, Passes = Pass Defined by FLDHARV as defined above, H.Downs = Output Header Downs for the designated Pass, Downs.0 = Input Weights with Harvester Header in Down Position with Zero Flows for the designated Pass, H.Ups = Input Weights with Harvester Header in Up Position for the designated Pass, and Ups.0 = Input Weights Equal Zero (0) with Harvester Header in Up for the designated Pass.

APPENDIX D: NASS OBJECTIVE YIELD FIGURES FOR FIELDS P AND W

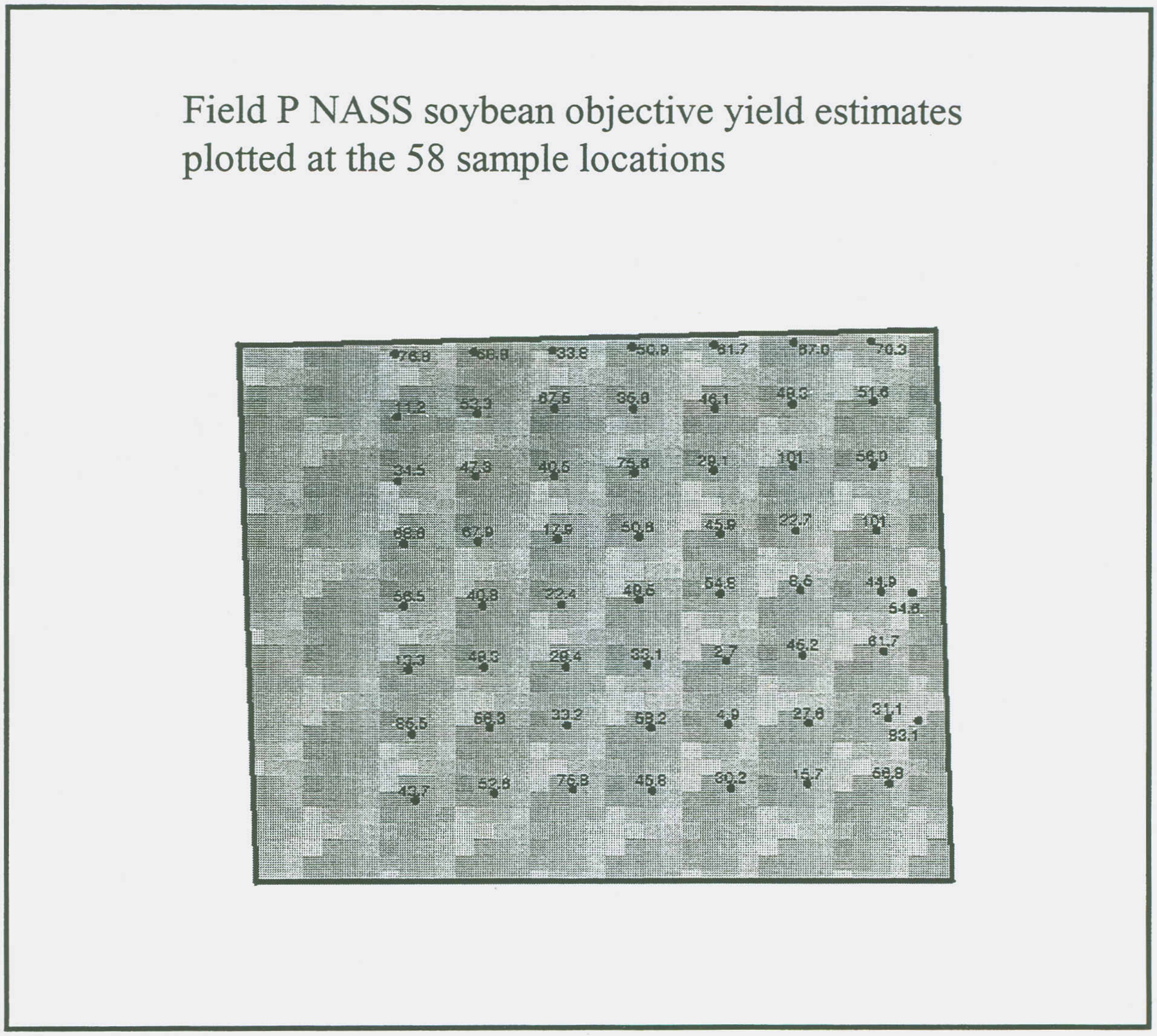

Figure 2. Field P NASS soybean objective yield estimates (Bushels\Acre) plotted at the NASS objective yield locations.

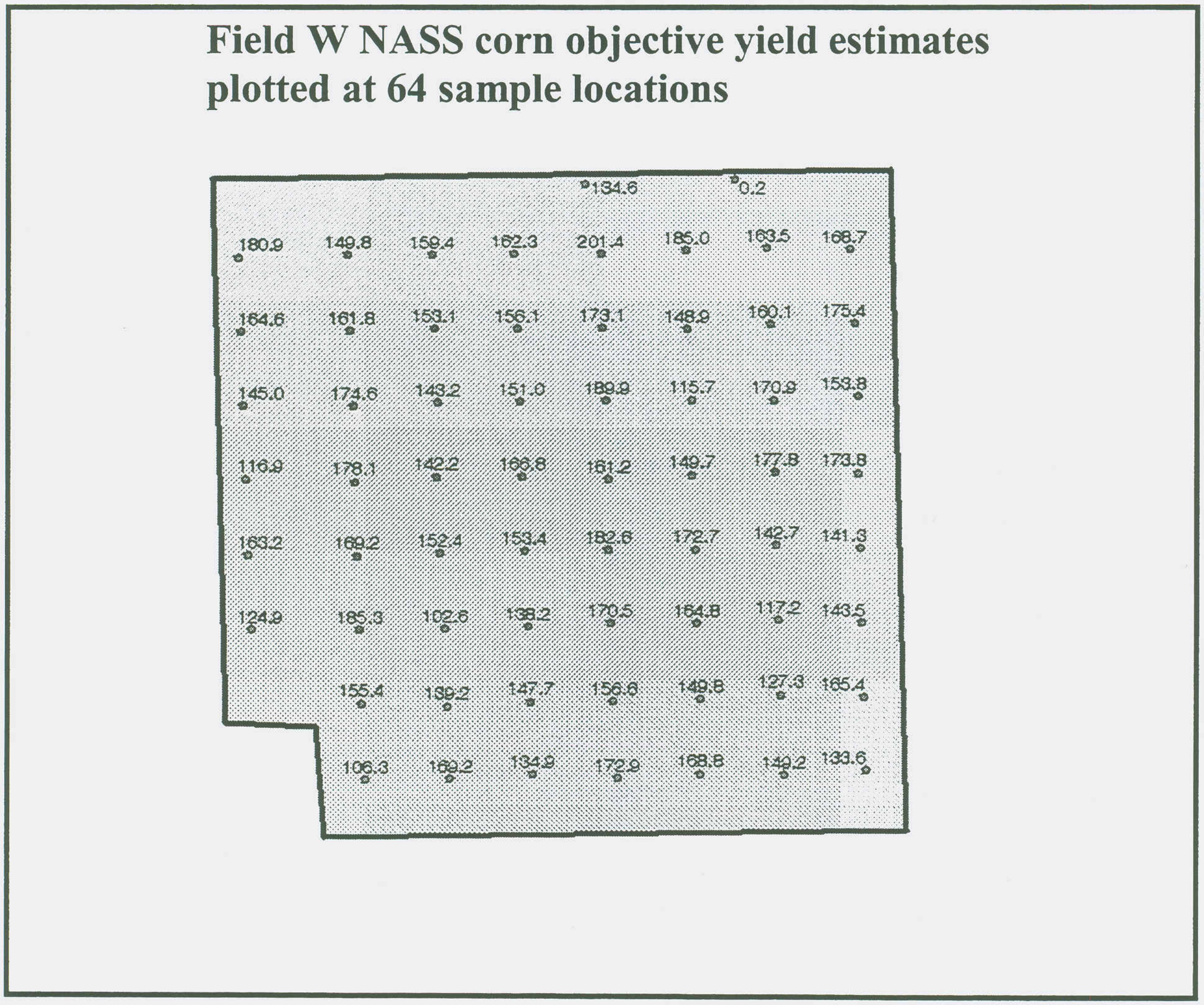

Figure 3. Field W NASS corn objective yield estimates (Bushels/Acre) plotted at sample locations.

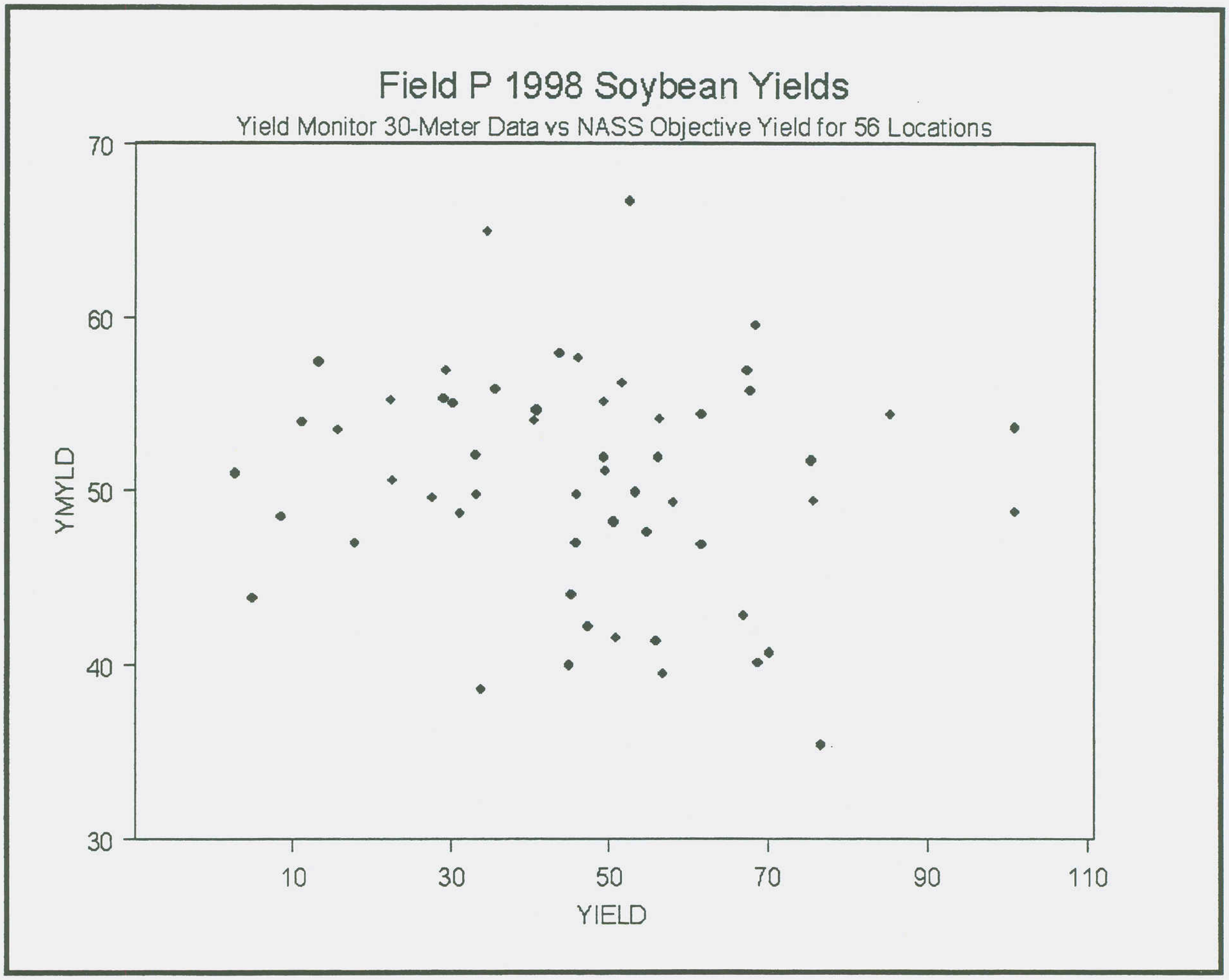

Figure 4. Field P 1998 soybean yields for NASS objective yield site estimates (YIELD) versus corresponding yield monitor 30-meter aggregated yield estimates (YMYLD).

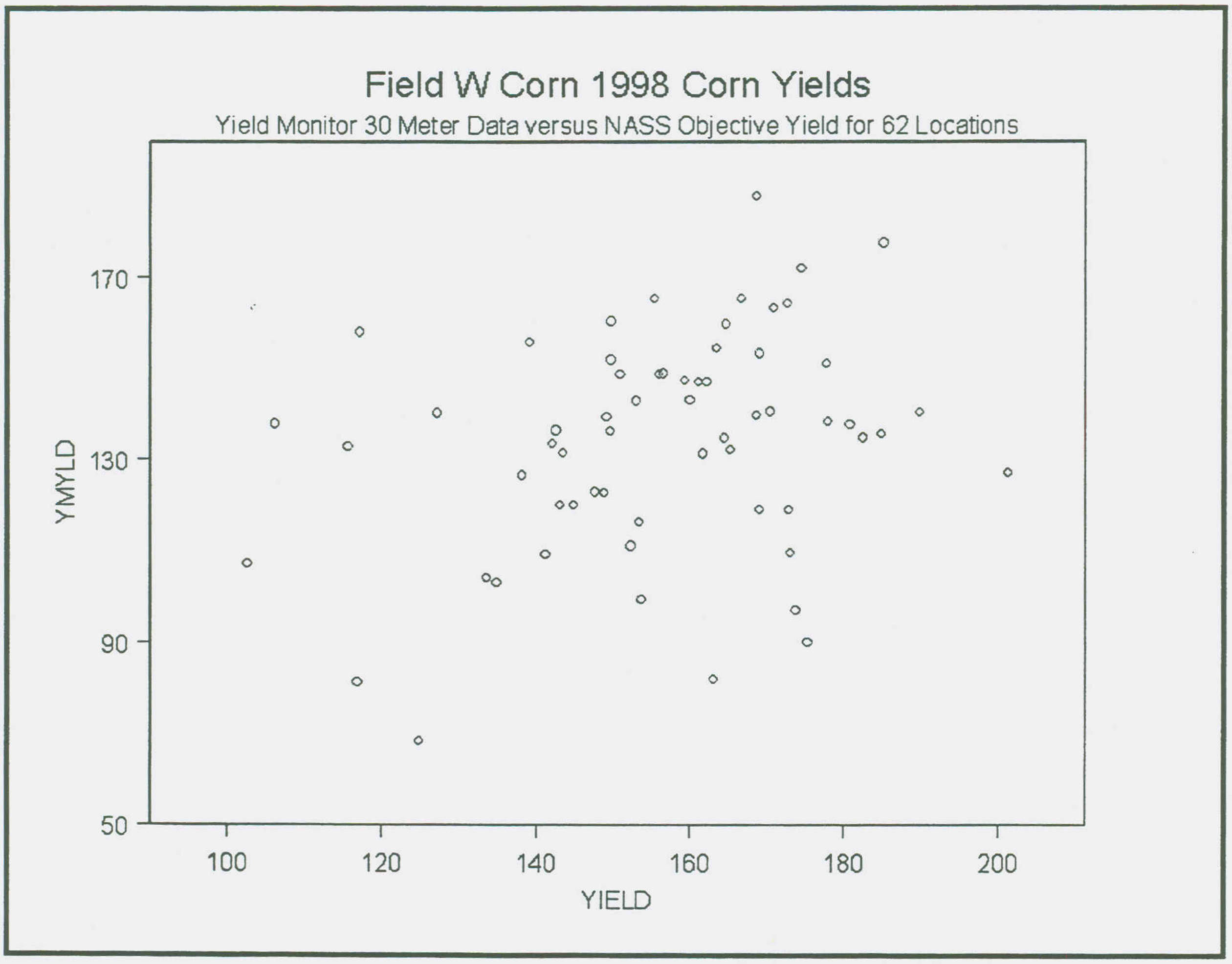

Figure 5. Field W 1998 NASS corn objective yield site estimates (YIELD) versus corresponding yield monitor 30-meter aggregated yield estimates (YMYLD).

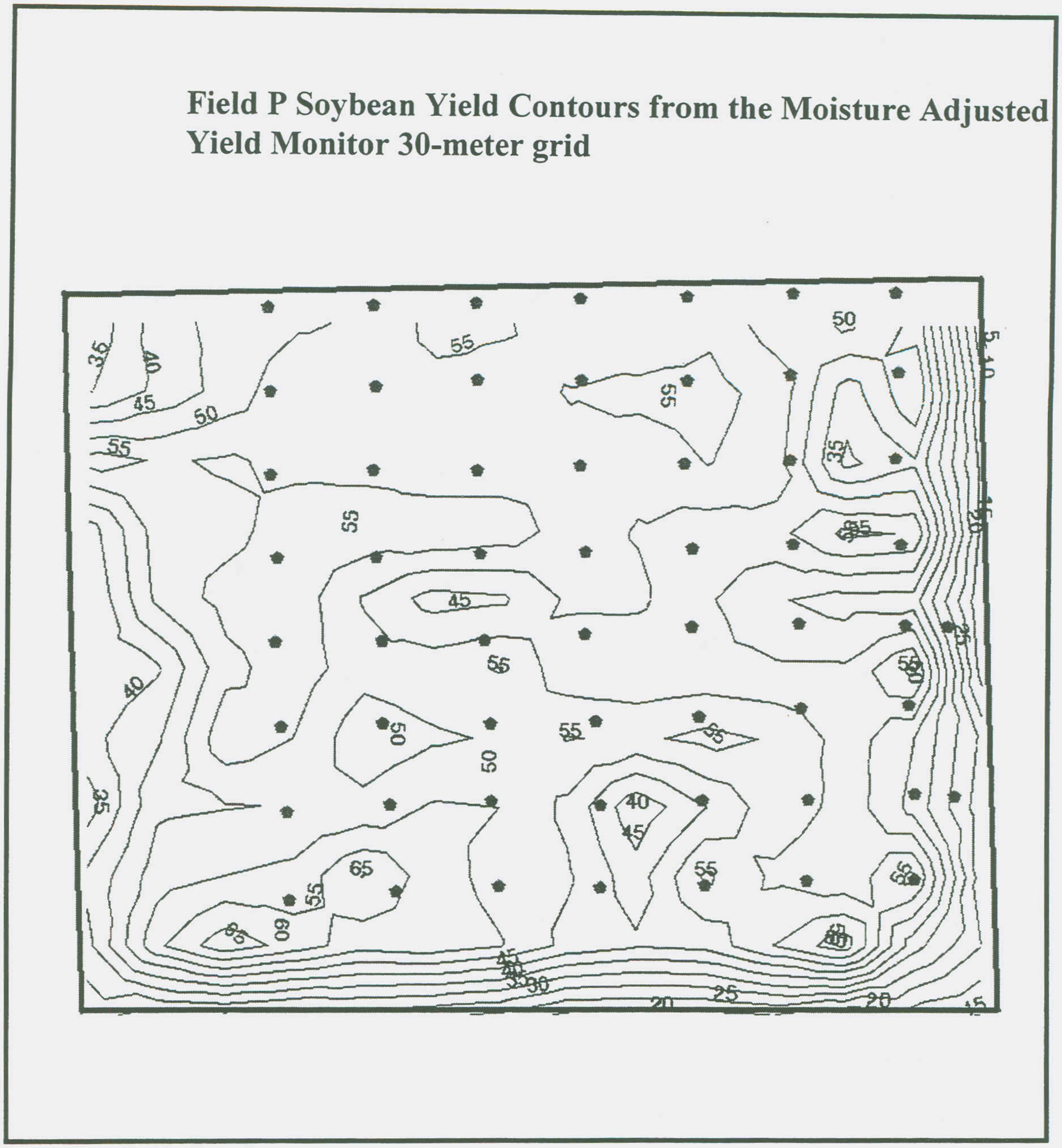

Figure 6. Field P soybean yield contours at five (5) bushels per acre intervals created from the yield monitor adjusted 30-meter soybean yield estimates.

Table E.3. ANOVA Analysis results for Field P Soybean YIELD and YMYIELD Variables for 1998

A. One-Way Analysis of Variance Formula		YIELD~ROW		
Terms:	ROW	Residuals		
Sum of Squares	4004.89	22841.09		
Degrees of Freedom	7	48		
Residual Standard Error:		21.81412		
Estimated Effects:	Balanced			

	Df	Sum of Squares	Mean Square	F Value	Pr(F)
ROW	7	4004.89	572.1278	1.202313	0.320
Residuals	48	22841.09	475.8560		

B. One-Way Analysis of Variance Formula		YMYLD~ROW		
Terms:	ROW	Residuals		
Sum of Squares	883.256	1435.709		
Degrees of Freedom	7	48		
Residual Standard Error:		5.469058		
Estimated Effects:	Balanced			

	Df	Sum of Squares	Mean Square	F Value	Pr(F)
ROW	7	883.256	126.17985	4.218555	0.001
Residuals	48	1435.709	29.9106		

Table E.4. Descriptive Statistics for the Field W, 1998 Variables: NASS Corn Yields and Yield Monitor Corn 30-Meter grid Yields for the corresponding 62 locations with Weigh Wagon data

	Variables[1]		
Statistics	Yield	YMYLD	Weigh Wagon
Bu/A	Bu/A	Bu/A	Bu/A
Minimum	102.6	68.4	
Maximum	187.9	201.4	
Mean	156.1	134.2	139.9
Std Dev.	20.4	24.1	
LCL Mean	150.9	128.1	
UCL Mean	161.2	140.4	
Number of Locations	62	62	
Correlation	0.29	0.29	

[1] Yield = Sample NASS Yields for 62 locations, YMYLD = Yield Monitor 30-M Grid cell Yields containing NASS Objective Yield Sample Locations for 62 locations after deletion of endrows..

Table E.5. MANOVA Analysis results for Field W NASS Objective Yield Corn yields and YMYIELD Variables for 1998.

A.	MANOVA	Formula =	cbind(YIELD,YMYLD)~ROW			
		ROW	Residuals			
Degrees of	Freedom	7	54			
Estimated	Effects:	UnBalanced				
Analysis of Variance Table						
	Degrees of Freedom	Pillai Trace	approx. F	df	Density df	P-value
ROW	7	0.3037	1.3812	14	108	0.1747

B.	MANOVA	Formula =	cbind(YIELD,YMYLD)~COLUMN			
		ROW	Residuals			
Degrees of Freedom		7	54			
Estimated	Effects:	UnBalanced				
Analysis of Variance Table						
	Degrees of Freedom	Pillai Trace	approx. F	df	Density df	P-value
COLUMN	7	0.5745	3.1088	14	108	0.0004

Table E.6. ANOVA Analysis results for Field W Corn Yield and YMYIELD Variables for 1998

A. One-Way Analysis of Variance Formula Yield~Column					
Terms:		Column	Residuals		
Sum of Squares		4936.48	20541.83		
Degrees of Freedom		7	54		
Residual Standard Error:			19.50395		
Estimated Effects:		UnBalanced			
	Df	Sum of Squares	Mean Square	F Value	Pr(F)
Column	7	4936.48	705.2118	1.853848	0.096
Residuals	54				
	6				

B. One-Way Analysis of Variance Formula YMYLD~Column					
Terms:		Column	Residuals		
Sum of Squares		12829.99	22706.16		
Degrees of Freedom		7			
Residual Standard Error:			20.50572		
Estimated Effects:		UnBalanced			
	Df	Sum of Squares	Mean Square	F Value	Pr(F)
Column	7	12829.99	1832.856	4.358915	0.0007
Residuals	54	22706.16	420.484		