EXPANDING GIS APPLICATIONS THROUGH PARTNERING

Invited paper submitted by the National Agricultural Statistics Service,
United States of America*

Introduction

1. This paper consists of three sections. The first outlines a number of present or possible applications of Geographic Information Systems (GIS) techniques to agricultural statistics. The second describes practical and legal concerns that National Statistical organizations must deal with in adopting and implementing GIS applications. The third section outlines the approach the National Agricultural Statistics Service (NASS) used to greatly expand its capability for creating GIS and remote sensing products in spite of personnel and budget limitations.

2. The paper is illustrative instead of exhaustive in coverage of examples and opportunities. Most section one and two material is written from a general standpoint, with a few specific United States examples. The third section is a case study totally focused on NASS efforts.

GIS Applications for Agricultural Statistics

3. GIS techniques have already had an effect on the products and operations of many agricultural statistics organizations around the world. Organizations are able to present more attractive and informative data products for their users; new research and operational Government programs can be based on spatial relationships; and organizations are often able to improve sampling, editing, and analysis by utilizing the advanced technology.

4. As stated in the theme for this sesion, GIS applications can make agricultural information more user friendly, particularly in the display and interpretation of data relationships. Users may miss significant relationships if they can only view static tables of area, yield, production, and inventory data, which are probably printed in an alphabetic order. Placing the same data in a GIS map format and presenting alternative views will clarify where the peak areas of production and inventory are located. By changing scales, and presentation colors, it is possible to compensate for varying sizes of political areas and illustrate the areas which have the greatest intensity of production or inventory. If crop progress data are being presented during a growing season, it is enlightening to the audience to see variation across the total area being mapped. Users who have GIS capabilities can take the agricultural statistics information and relate it to soil types, marketing areas, transportation grids, etc.

5. One of the most significant products or services that an agricultural statistics organization can provide for customers is to apply GIS techniques to create time series information. There is great interest and concern in changes over time in production and inventories. On one hand,

---

*Prepared by Mr. Rich Allen, National Agricultural Statistics Service.

there are concerns about losses of agricultural land to commercial development, roads, and new

housing. Other individuals and organizations are concerned about increases in agricultural production, particularly in the case of intensive livestock production operations. GIS data sets and displays can shed light on these changing patterns, without agricultural statistics organizations becoming embroiled in political debates. GIS capabilities can also be used for displaying nontraditional farm related data such as total government payments, average payments per farm, etc. by small areas within a country or sub-country divisions.

6. Government organizations can not predict the uses that customers will make of well designed and presented GIS products. As mentioned above, companies will want to resummarize and display data by their own marketing or transportation areas. They will also compile such things as chemical usage data related to total crop area and to soil types from other data sources.

7. One growing use of GIS technology has been in investigative and regulatory activities. Some States in the United States, and some other countries, are licensing certain types of farming operations (such as swine production units). In many cases. this means that the actual production facilities have been accurately geographically located and regulatory GIS data bases have been created. Discussion of the implications of such files for statistical organizations is included in section 2 below.

8. One important use of GIS technology which might be overlooked is the capability to improve statistics for all customers through internal GIS applications. Organizations can examine sample allocations in relationship to population distributions to determine adjustments that might be needed in weighting reported data. They can also review reported data by area during the analysis stage to better understand current data relationships and changes over time. Response patterns can be evaluated in order to target improvement efforts. By combining population and sample location data with transportation grids, organizations might be able to bring in new data collection efficiencies for surveys involving face-to-face interview.

GIS Concerns for Statistical Organizations

9. Most country level government agricultural statistical organizations operate under specific regulations which govern the handling of collected data. In most cases, organizations must protect confidentiality regardless of whether data were collected on mandatory or voluntary surveys. This requirement presents a number of challenges for organizations in expanding the use of GIS products for customers. The confidentiality rules used for determining which printed data aggregates are publishable must be applied to any GIS displays. If GIS products are being created from census type data sets it is possible to create products for smaller geographic areas but the statistical organization must guard against “custom” data products which might inadvertently reveal some confidential data when compared against other small area products which have been created. Organizations often may need to expand their confidentiality review and clearance procedures in order to properly meet these new concerns.

10. Legal arrangements vary by country, but many government statistical organizations are prohibited from creating proprietary products. Having actual production locations, such as concentrated feedlots, in a GIS data base means that statistical organizations could create some

desirable aggregations which would meet confidentiality restrictions but would provide marketing advantages if they went only to one customer. The approach NASS has always used, and which has been expanded to GIS products, is to release all products for public consumption. In the past, a number of special reports have been created and released on a regularly scheduled basis because of funding from an industry group which needed improved data. NASS now has an expanded capability to create special data tabulations because of the Census of Agriculture data files. Many of the requested tabulations are not of interest beyond the requesting organization, even if they result in a GIS display, but all such tabulations are tracked and a description of each is available on the agency web site.

11. An extremely sensitive GIS concern is the handling of specific location data for concentrated livestock production. In some cases, an agricultural statistics organization might be required to be involved in the creation or maintenance of such location data because of its data handling capabilities and its obligations to the broader government agricultural organization. This might provide an opportunity for the organization to utilize improved quality administrative data for sampling and estimation even though the organization can not release or reveal the location data. The statistical organization might even need to be involved in creating GIS products from the location data but those products must be issued by the regulatory organization with the legal authority, not from the statistical organization.

12. NASS has had to operate carefully and judiciously in the past year because of concerns about possible Foot and Mouth outbreaks. Almost every State department of agriculture wanted to set up contingency plans for how information, notices, and quarantines would be handled in the case of a suspected outbreak. Many States wanted to hold practice alerts or to create names and address files, along with physical locations and livestock numbers so they would be prepared. Since all NASS files are protected under strong confidentiality provisions, the Agency can not provide such detailed information but needed to keep in touch and be supportive of the efforts of the other organizations. In many cases, State Statisticians have offered to compile new tabulations and maps, which could be released, to show livestock concentrations from the Agency sampling frames. In the States where regulatory GIS files were available those files could be provided –but they often contain information for only selected species.

13. Creating new GIS products, particularly maps, can lead to new confidentiality questions and decisions. NASS utilizes field boundary and crop identification data from its annual area frame survey as the training set for remote sensing clustering and crop specific classification of large areas. These classification results from many images are merged together in a mosaic to form a very valuable GIS product referred to as the cropland data layer. While actual field boundaries are used as inputs to the clustering algorithms, the final product contains only classification results for every data pixel within a satellite scene and no specific original data remain. Thus, NASS is able to fully release all of the classified products which have been created. However, NASS releases all products with a data warning that accuracy of any specific data point is not guaranteed.

14. The fact that many organizations are developing GIS capabilities can also lead to new proprietary concerns. One recent, unprecedented request was for NASS to use its area frame data to evaluate the accuracy of the land use classification that another organization had done.

This request was rejected as being a proprietary benefit. However, it also would have been rejected for two additional reasons: 1. the survey respondents in the area frame were not informed that their information might be used for such a purpose (they are only informed that NASS will use their data for statistical purposes) and 2. the area frame data in this instance had not been subject to the extra scrutiny that NASS uses for States in the cropland data layer studies. (If this had been a State for which NASS had created a current cropland data layer product which was publically available, the organization could have manipulated the two files and formed their own conclusions.)

Resource Sharing for GIS Information

15. NASS has been a leader in the fields of using remote sensing for crop identification, creating satellite based data products, and creating GIS applications. However, NASS efforts had been restricted to research and demonstration levels because of budget . Budget constraints also meant that only a small trained staff was available for remote sensing/GIS work. Because of the potential for creating GIS products that could broadly benefit agriculture, NASS initiated a new effort in 1998 to identify new partners who could benefit and who might be willing to make staffing and equipment investments.

16. Since NASS had operated for more than 80 years through cooperative agreements with State governments, public universities, and other Federal agencies, a ready vehicle was available for establishing partnerships. The NASS State Statisticians in every State were asked to pursue their public sector contacts to identify individuals and organizations interested in such an approach. A total of 5 States were identified for work in 1999: 3 were new partnerships and 2 were States in which NASS had been concentrating its recent developmental work. Two more full States were added in 2000 plus a separate pilot effort in a third State. In 2001, two additional pilot areas were added to the Program.

17. The NASS basic offer was to provide the software (for both remote sensing classification and GIS applications), ground data, satellite imagery (through an agreement with two other U. S. Department of Agriculture agencies), and the necessary training. The new partners were to provide an analyst for training and the actual product creation work, plus a properly configured workstation and printer.

18. The approach has been extremely successful but not without needing additional efforts and compromises. For instance, the analysts' backgrounds varied tremendously and some required additional training on processing and remote sensing concepts to bring them up to the level expected for the main classification training. It was originally planned that each project would first focus on crop area estimation and then on creating a cropland data layer. However, the pilot State in 2000 was directed more towards training student analysts in a minority university for future employment, so the emphasis there changed to having those individuals also involved in ground data collection and concentrating on land cover mapping rather than area estimation.

19. In every new effort there are lessons to be learned. The partnerships involved universities in some cases and State or Federal agencies in others. The State and Federal arrangements have been the most productive in the first two years since new university analysts were added the

second year and needed training. There will also be re-training needs for the other partners since the analysts are in "one-of-a-kind" positions and there is no continuity if they move on to other opportunities.

20. There has been great interest and excitement in the cropland data layer products that have been created and which are available as CD-ROM products. Some business locations decisions have been made in one State based on the new information. In another State, other land cover information was inputted into the classification and a full Land Cover data layer was created which will have broad use for planning purposes.

21. The experience to date has been extremely encouraging since: 1. customers are receiving a useful product and 2. NASS has been able to more broadly serve agriculture without an increase in staffing or budget. However, the effort will not reach its potential unless more partners can be recruited for multi-year commitments so large areas of the major agricultural producing areas can be covered consistently.

References

Allen, Rich (1995). "Guidelines and Principles for Public Use Geographic Data." *USDA/NASS Official Memorandum to Donald Bay, Administrator*, Washington, D.C., USA, February 1995.

Allen, Rich, Hanuschak, G. (1998) "Policy Issues Associated with the Utilization of Geographic Information systems in the U.S. National Agricultural Statistics Service." *Working Paper No. 25, UN/ECE Work Session on Geographical Information Systems,* Ottawa, Canada, October 1998.

Craig, Mike. (2001). "A Resource Sharing Approach to Crop Identification and Estimation". *2001 ASPRS/ACSM Annual Conventions Technical Papers*, available on the CD, *Proceedings of the ASPRS 2001 Conference*, Bethesda, MD, USA.

Griffith, Daniel A. (1999) "A Methodology for Small Area Estimation with Special Reference to a One-Number Agricultural Census and Confidentiality: Results for Selected Major Crops and States." *USDA/NASS, Research Division, Research Report Number RD-99-04*, Washington, D.C., USA, August 1999.

Hanuschak, George, Hale R., Craig, M., Mueller , R., and Hart, G. (2001). "The New Economics of Remote Sensing for Agricultural Statistics in the United States." *Proceedings of the CAESAR* Conference, *Italian Statistics Agency (ISTAT), Rome Italy, June 2001.*