THE SCIENTIFIC AND TECHNICAL ISSUES IN INTEGRATING REMOTELY SENSED IMAGERY WITH GEOCODED DATA BASES

W. Murray Strome and Berne Grush
Perceptron Computing Inc.
Downsview, Ontario, Canada M3H 5S8

ABSTRACT

Since the launch of LANDSAT-1 in 1972, major advances have occurred in computer technology, which have in turn led to improvements in our ability to more effectively analyze digital remotely sensed data. The decreasing cost of computer equipment, especially both on-line and off-line memory, has led to an explosion in the use of digital mapping techniques.

For several years, many have recognized that the full potential benefits of satellite remote sensing would not be realized until digital analysis of satellite image data and automated cartography/geocoded information systems become more compatible and integrated.

Some of the difficulties associated with this integration will be explored, as well as some of the efforts being made to overcome them.

KEYWORDS: Geographic Information Systems, remote sensing, automated cartography, LANDSAT

INTRODUCTION

Workers in many disciplines associated with resource management, environmental monitoring and land use planning are required to collect, store, analyze and display spatial data collected from several sources. For decades, the main sources of such data have been maps, ground surveys and aerial photography. Until relatively recently, the spatial data processing was performed using paper maps, transparent overlays, human interpretation and manual drafting.

Over the past decade or so, new technological tools have become available. Computer systems have permitted a more automated approach to map making as well as to the storage and manipulation of geographic information. The launch of LANDSAT-1 in 1972 ushered in a new era for the collection of data concerning the earth's surface. For the first time, synoptic data in digital form became available regularly and repetitively for all land areas of the earth. The high quality of the data and its abundance, coupled with advances in computer technology have led to considerable progress in developing digital processing techniques for analyzing the remotely sensed data from satellites. These advances have increased our ability to determine more accurately and objectively the material present on or near the earth's surface and its condition.

For some time, it has been recognized that the key to realization of the full potential benefits of satellite remote sensing for land applications lies in the successful integration of the data from this new source with geographic data bases. Much success has been achieved using manual techniques such as overlaying satellite images and maps. However, the task of integrating satellite data into computer based Geographic Information Systems has proven difficult.

GEOGRAPHIC INFORMATION SYSTEMS

A large number of digital Geographic Information Systems have been developed over the years. The basis for computer processing of spatial data lies in the creation of consistent digital data files which can be manipulated in a manner to extract the particular information desired for any specific purpose. The files are usually generated initially by digitization of existing maps (hopefully, ones which are accurate). The systems are intended to permit the user to manipulate and integrate data files in order to bring together information from several spatial data sets into a composite for either visual display or analytic modelling. To bring these data sets into correspondence is known as overlaying, which demands a reference to a common geographic coordinate system.

The typical products of a Geographic Information System (GIS) are documents in tabular or graphical form which are directly usable by a professional planner, resource manager or decision maker without further manipulation or modification (Calkins and Tomlinson, 1977). The utility of a GIS cannot be measured in terms of its complexity, but rather, its ability to provide the right data to decision makers in the proper manner and at the right time. A GIS may be manual, computer based or a combination of the two.

A typical series of questions which might be posed to a GIS are: in a given geographical region (e.g. country),

a) how much land was converted from agriculatural to urban use in the past five years? Ten years?

b) how much land is suitable for future urban development, and where is it?

c) how much of this land is not prime agricultural land, and where is it?

The replies to these questions could take the form of printed information as well as maps showing the location of all land suitable for development in the region as well as that portion which is not prime agricultural land.

A GIS generally deals with three data characteristics - the actual phenomenon, classification or measurement; its geographical location; and the time at which the data were measured. It is important to note that an effective spatial data management system must permit location and nonlocation (attributes) data to vary independently, as the attributes may change character without changing spatial location or vice versa. For example, a field planted to wheat one year may be left fallow the next, or a river may change its course over a period of time.

In attempting to make the transition from the paper map-based manual GIS to the automated computer based system, two fundamentally different approaches have been taken. The first approach is based upon the manner by which the spatial location of objects is depicted on maps, using points, lines and polygons. These are most commonly defined on maps using x, y cartesian coordinates such as latitude and longitude. In a computer based GIS, the location of points are recorded. A single point will have one set of coordinates while both lines and polygons will be defined by a sequence of coordinates. Once the location of a geographic entity is defined, it must be identified or assigned attributes. There are several approaches which may be taken for the storage of the spatial information and the attributes (e.g. a line may represent a river, a specific highway etc.; a polygon, a city, county, wheat field, etc.; and a point an aircraft beacon, lighthouse, etc.). The manner in which the data are stored has an important impact on the way analyses may be performed. An excellent description of the options available has been given by Dangermond (1983).

A second approach to the computer based GIS is to represent the geographic data by grid cell encoding. The area of interest is subdivided into a large number of grid cells and each cell is assigned the attributes corresponding to its location from a map. Usually the grid cells are square, although some interesting coding efficiencies can be achieved using hexagonal cells as is done with a system called AGIS/GRAM.

The conversion of map data to digital form using each of these approaches is illustrated in Figure 1. In the first instance, geographic entities are defined by a linked list of coordinate pairs with a reference to an attribute list. Attributes may be assigned to an object or in the case of a line representing a boundary, to the object on either side of the line. Grid cells usually form an array with a label assigned to each cell. To conserve space, the data are often run-length encoded.

POLYGON DATA BASES

There are several hundred different polygon or line segment based Geographic Information Systems in existence. In the U.S. Geological Survey alone, it has been reported that there are over fifty such systems in use. Since no standards have been developed, each system uses a different internal data structure, making it very difficult to combine the data from different systems.

Two factors resulted in this form of GIS. Because the original data source was usually on printed or hand drawn maps, it was natural to digitize these by tracing the lines on the maps to define points or lines and polygons to define areas. Perhaps more importantly, this approach leads to efficient use of computer memory. When the original systems were devised, computers were equipped with relatively small amounts of internal memory which was very expensive. Even external memory was at a premium. The key to a useful GIS is its ability to manipulate the data by overlaying the information from different data sets. This is a complex problem, especially when the data are stored in the form of a linked list of coordinate points defining lines and polygons. The complexity of the computing required to effectively manipulate the data was deemed worth the effort in order to minimize the cost of storage.

Today, the cost of memory has fallen drastically so that space efficiency is no longer a primary consideration. However, the line segment or polygon representation of data

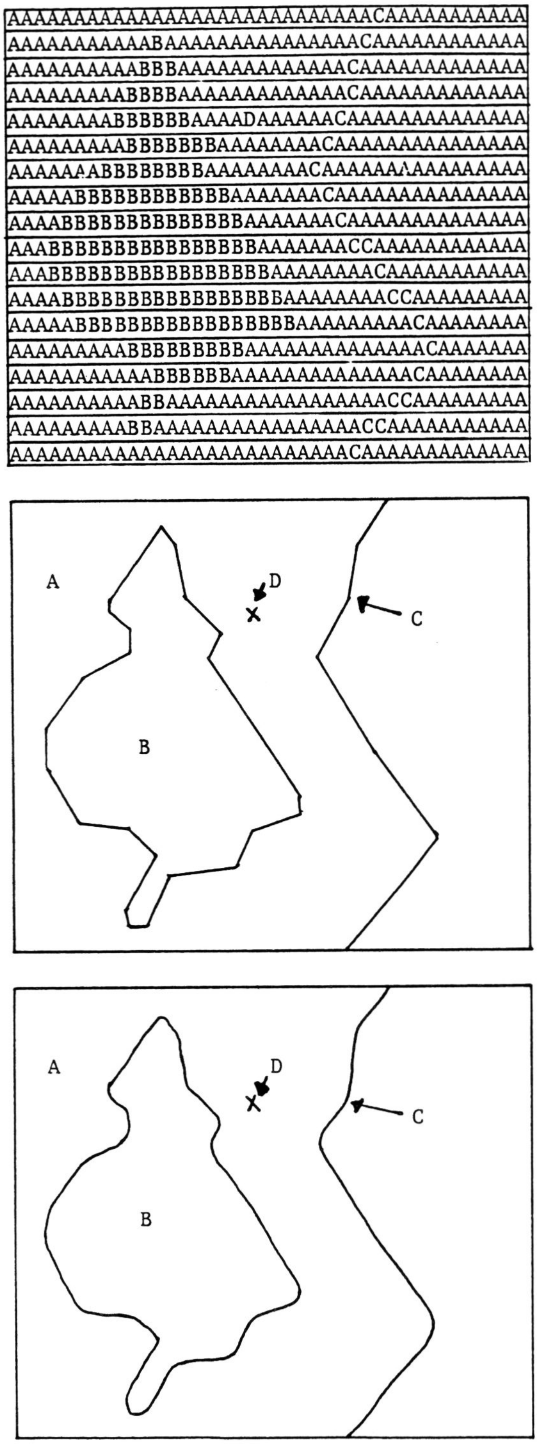

Figure 1. Representation of a map (bottom) in polygon (centre) and grid cell (top) formats.

still has two major advantages. First is the fact that a tremendous amount of geographic data has been collected and stored in this format. Secondly, and perhaps more importantly, this representation most accurately corresponds to the hard-copy format of maps, which are the major tools of resource managers and land use planners.

The Canada Geographic Information System (CGIS) developed in the mid-1960's was one of the most ambitious projects for collecting, storing and manipulating spatial data. The highly innovative design work on this, one of the largest and most complex GIS in operation, created a system which has been able to migrate through several generations of computers and which is still one of the most cost-effective in existence (Calkins and Tomlinson, 1977) after two decades of operation. One of the most innovative aspects of the CGIS was the method chosen for data entry. Most such systems to that time (and even today) captured the line data from maps by manually tracing the lines using a point digitizer. Later systems have used automated line-following digitizers. The CGIS system uses a high resolution drum scanner to digitize maps into a raster format. Complex software then converts the raster data into the usual linked list of point coordinates to represent points, lines and polygons. The results are plotted using a graphical plotter or are displayed on a graphics terminal where they are edited with manual assistance.

GRID CELL DATA BASES

The principal attraction of grid cell geographic data bases is their conceptual simplicity. Overlaying data from different sources is a relatively simply task. Since the data are laid out in computer memory in the same way for each layer of data, it is easy to perform logical and arithmetic operations among data sets. Manual data entry from maps is easily performed by placing a gridded sheet of mylar or tracing paper over the map to be coded. Identification labels are then entered into each cell. The data may then be entered into a computer. Expensive peripheral equipment, such as high precision plotters are not required to display the results, as they may be printed using a standard line printer.

The original major objection to the simple grid cell approach was the vast amount of computer storage required. In the very simple example shown in Figure 1, the polygon, line and point shown could be represented by about 30 sets of x,y coordinates. Including attributes, this could be represented by less than 70 data elements. The grid cell representation, however, requires 720 data points, an order of magnitude greater. For very large data bases with complex maps requiring very small grid cell sizes, the difference can be several orders of magnitude.

Today, with the cost of storage much less important than it was in the past, the grid cell approach still has important limitations. Unless the grid cells are extremely small, the output products "look different" than the maps. Depending upon the particular users and applications, this may or may not be important.

One of the better known grid cell systems is the Maryland Automated Geographic Information (MAGI) System (Maryland, 1981). While the system itself is independent of scale or grid cell size, the State of Maryland uses a system based upon its 1:63,360 scale maps and the Maryland State Coordinate Grid Lines which are separated by 10,000 feet in each direction. Each 10,000 x 10,000 foot square is subdivided into 25 2,000 x 2,000 foot 91.8 acre grid cells, which may be further subdivided into twenty 400 x 500 foot 4.57 acre grid cells. The system is easy to use, is being constantly updated and is very comprehensive with respect to the wide variety and completeness of resource, environmental, demographic and cultural data contained.

The MAP Analysis Package (MAP) (Tomlin, 1983) is particularly useful as a teaching aid. It is a set of computer programs that provide for the input, output and transformation of cartographic data stored in a grid cell format. Because of its low cost availability and ease of use, it is a popular tool for demonstrating the fundamental GIS concepts to students. The system has been used to implement at least one small scale practical geocoded data base (Tomlin, 1982).

REMOTELY SENSED IMAGE DATA

Image data acquired by remote sensing satellites are provided in two basic formats: photographs and digitally recorded computer tapes. The photographs may be used in a manner similar to that used for aerial photographs. They can be interpreted and the data transferred to base maps for later digitization and entry into a GIS.

The digital data may be handled in a similar manner: after analysis and classification, perhaps using digital techniques, the data may be converted to hard copy using a line printer, plotter or photographic output. The analysed data may then be transferred to base maps using standard manual or semi-automated methods and then entered into the GIS.

The vast majority of satellite remote sensing data now available has been acquired by the LANDSAT series launched by the National Aeronautics and Space Administration (NASA) and later the National Oceanographic and Atmospheric Administration (NOAA) in the United States. The early satellites carried Multi Spectral Scanners (MSS) which provides digital data for four spectral bands in a two dimensional array covering an area on the ground of about 185 x 185 km. The individual points or picture elements (pixels) represent an area of about 57 x 80 metres in the original data. The good radiometric fidelity of the data has helped to foster significant advances in computer assisted analysis and classification of the digital data. Data from the Thematic Mapper (TM) carried on satellites launched after 1982 are even better quality radiometrically and provide better spatial resolution (30 x 30 metres).

It has long been recognized that the acceptance of LANDSAT data by the resource management agencies was somewhat hampered by the difficulty in incorporating the data into existing geographic information systems. While the geometric fidelity of the LANDSAT data is excellent in comparison to aerial photography or that from other satellites, the coordinate system is non standard and the errors in position are too great for direct input to a GIS. In most countries, even the geometrically corrected LANDSAT data are in a coordinate system which is based on orbital geometry rather than on standard mapping coordinates.

Canada is fortunate to have high quality base maps for the entire country based on a single standard projection system: Universal Transverse Mercator (UTM). Furthermore, through the Canadian Advisory Committee for Remote Sensing and the National Remote Sensing Program administered by the Canada Centre for Remote Sensing, an excellent dialogue is maintained between the user community and the technology developers. This resulted in what has proven to be a significant decision with regard to the format of digitally corrected satellite data distributed in Canada. It was decided that MSS data should be made available to Canadian users in a format which exactly corresponds to standard UTM map sheets. The resampled pixel size is 50 x 50 metres and the orientation is along the map coordinates (Northings and Eastings). Thus, after the data are digitally analyzed, no further corrections are required in order for the data to match digital data bases which were keyed to the national mapping system. This philosophy is being carried over to the Thematic Mapper data, where the resampled pixel size is 25 x 25 metres and to SPOT (to be launched by France in 1987) where the pixel sizes will be 12.5 x 12.5 metres for the 20 metre multispectral instrument and 6.25 x 6.25 metres for the panchromatic sensor.

INTEGRATION OF REMOTE SENSING AND GIS

The integration of remotely sensed data analysis automated cartography and digital GIS offers several benefits. If the data can yield the correct information, it provides the most economical method for repetitive data collection over large areas for updating geographic data bases. On the other side of the coin, the best analysis of satellite data requires the input of all available ancilliary information. Much of this could be obtained from existing geographic information systems.

The data from LANDSAT have been used to update and even to create georeferenced data bases, and information from such data bases has been used to assist in the analysis of satellite data. However, most work to date has used the two systems more or less independently. Part of the reason for this lies in the differences in the training and background of the scientists and engineers responsible for the technological developments in the two areas. The developers of Geographic Information Systems had strong cartographic and photogrammetric influences. The remote sensing equipment developers were influenced by the environmental sciences, physics (in particular, spectroscopy) and signal processing. The differing backgrounds and influences resulted in systems whose data did not mesh naturally.

In the United States, where there are a fairly large number of grid cell based geographic data bases, the geometry of corrected LANDSAT digital products is incompatible (for MSS, 57 x 57 metre pixels aligned along the axes of the Space Oblique Mercator projection, which was devised to mimimize the corrections required for data acquired by satellite). In Canada, the bulk of geocoded data is stored in polygon format.

With the advent of LANDSAT, there was considerable pressure (much if it from the remote sensing community) to move from polygon based systems to grid cells. The arguments used were that it is much easier to manipulate data stored by grid cells (which is true) and that the cost of computer memory was no longer an impediment to this approach. This format also happens to correspond, in concept at least, to the raster format of satellite data. Thus, it was felt that it would be easier to encourage resource managers and planners to adopt remote sensing data and analysis technology if the data bases they used had a similar internal format. Many have predicted the demise of the polygon data base. This is unlikely for the forseeable future, simply because the format is inherently too useful! Moreover, while it is difficult to manipulate the data (e.g. to create overlays) in the sense that this requires complex, sophisticated software, the problems have been solved, more than once!

Some of the impediments to the successful integration of automated cartography or digital GIS technology and remote sensing are simply symptomatic of the difficulty experienced in integrating data from one data base into another. In particular, there are difficult problems in moving data between polygon and grid cell data bases and vice versa. Indeed, the difficulties occur in attempting to merge the data from two different data bases of the same type. The fundamental source of the problem is that each data set is a (mis)-representation of the real world, and is subject to errors in measurement. Figure 2 illustrates the difference between the map (which was probably different than the world it represented) and each of the two representations, line segment and grid cell. The important point to be made here is that these errors are different. If two different people had digitized the lines or encoded the grid cells, the results would have been different again.

It is possible to design a GIS system which is based upon a vector approach for digitization of maps for data entry. The data may be converted to grid cell format for ease of internal manipulation. The resulting output products cannot exactly match the original maps, unless the grid cells are extremely small. If maps are an unimportant byproduct, or if the user can be educated to accept the differences, this approach is quite acceptable.

Remotely sensed data is provided in what is inherently a grid cell format. In Canada, where the cells are oriented along the standard mapping coordinates, the data could be quite easily integrated into a grid cell based GIS. Since most of these systems in Canada are vector or polygon oriented, successful integration required conversion between grid cell and polygon representations. In addition to the differences in the representation of the earth's surface using these two approaches, there are other difficulties to be overcome, especially in any attempt to automatically convert between the formats.

When a map is converted to a grip cell format, it is possible to assign an attribute to a cell to indicate whether it corresponds to a part of an areas, a part of a line or an isolated point. However, when the data are obtained from satellite, ambiguities arise. Does a single isolated pixel represent a small area, a point or part of a line which is not connected because of errors in classification? Do two or three adjacent pixels represent part of a line, or a small elongated area? These are problems which cannot be solved on the basis of the satellite imagery alone. Moreover, with today's state of the art, it is not practical to define the location of a boundary to subpixel accuracy.

B.C. MINISTRY OF FORESTS SOLUTION

Provincial legislation passed in 1978 placed a requirement on the B.C. Ministry of Forests to report annually on the current status of the forest resources in that province, which covers 52 million hectares. The task would be virtually impossible without automatic cartography and remote sensing technology. Because the current state of the art, integration of these technologies has not advanced very far. Instead, a pragmatic and highly practical approach has been taken (Hegyi and Quenet, 1983). Complete LANDSAT coverage of the province was used to update the entire forest inventory in 1981-82. New LANDSAT data are geometrically registered to previous data, which were registered to the forestry base maps (which may be in error). Changes due to selective logging and other disturbances are detected automatically and displayed as a (grid) overlay to the existing (vector) digital map base. A human operator verifies that changes indicated by the satellite data analysis are valid. On the basis of general knowledge of the area as well as field data and aerial photography, the polygon boundaries may be modified in the data base. At this stage, no attempt is made to automatically convert between the grid and polygon format, nor to automatically update the data base using the satellite data. The changes in the forest are flagged for human operators who make the final decision.

The computer is used to analyze vast amounts of satellite data in order to identify those areas where changes seem to have occurred. Again, the computer is used to store and manipulate similar amounts of geographic data for the province. However, the human operator is much more efficient at using

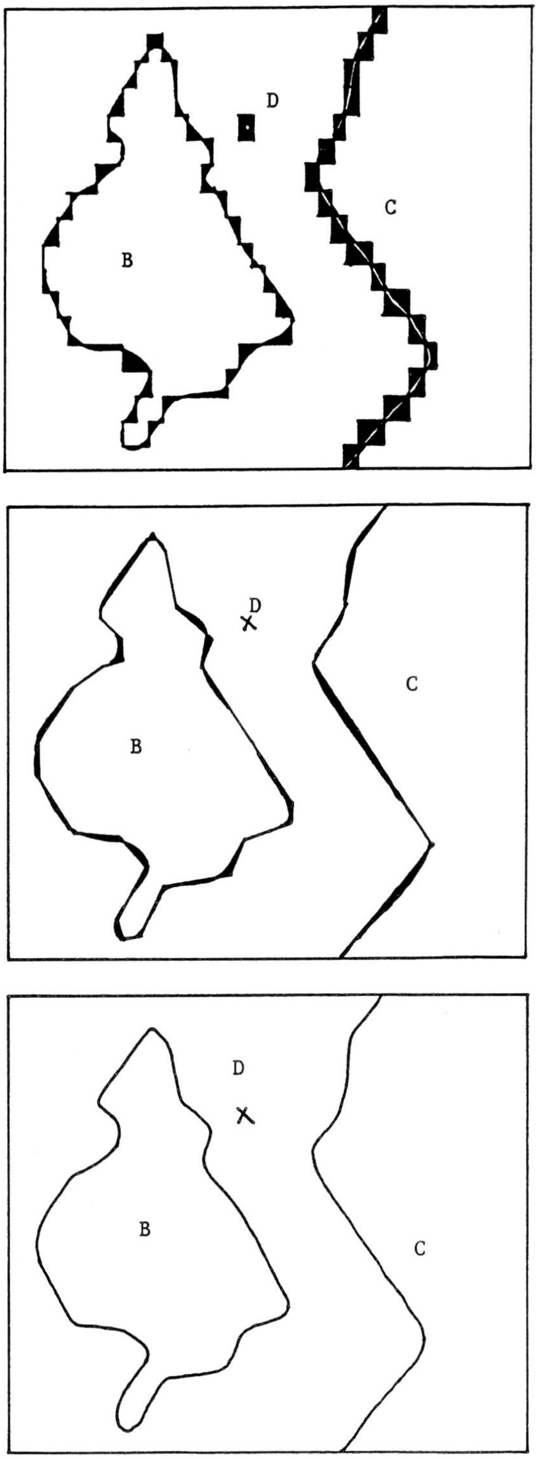

Figure 2. Differences between original map and GIS grid cell (top) and polygon (centre) representations.

contextual information (some of which may not even ben in the data base) to determine if a change really has occurred, and if so, exactly where with respect to the original map.

THE FUTURE

One attempt to automate the process of integrating remote sensing digital image and automated cartography technologies is the development of an Expert System (Goldberg, et al., 1985). In this system, expert knowledge is encoded in symbolic form to attempt to emulate the human activities involved in selecting appropriate image processing and analysis techniques, recognizing valid changes in the forest conditions and making intelligent decisions regarding map boundary changes. The system is based upon the forest inventory update procedures now performed by the B.C. Ministry of Forests in a less automated manner.

The system design includes the incorporation of a large number of existing image analysis programs developed for the LANDSAT Data Image Analysis System (LDIAS) at the Canada Centre for Remote Sensing and is based upon a hierarchical organization for the forestry map updating problem. The overall design is now complete, and a number of the "experts" have been coded. To acquire new knowledge to incorporate into the system, a number of maps of different regions in British Columbia will be used, together with multidate LANDSAT MSS and TM images to detect logging. As the system becomes capable of handling change detection in forests in British Columbia, it will be applied to other regions and expanded to include sensor inputs from the SPOT satellite, airborne synthetic aperture radar and eventually, RADARSAT.

REFERENCES

Calkins, H.W. and R.F. Tomlinson 1977. Geographic Information Systems, Methods and Equipment for Land Use Planning, 1977. Reston VA: U.S. Geological Survey

Dangermond, J. 1983. A Classification of Software Components Commonly Used in Geographic Information Systems. A Chapter in "Design and Implementation of Computer-Based Geographic Information Systems", Amherst, NY: IGU Commission on Geographical Data Sensing and Processing, 1983.

Goldberg, M., David G. Goodenough, Mayer Alvo and Gerald Karam 1985. A Hierarchical Expert System for Updating Forestry Maps with LANDSAT Data. To be published.

Hegyi, F. and R.V. Quenet 1983. Integration of Remote Sensing and Computer Assisted Mapping Technology in Forestry. Canadian Journal of Remote Sensing, Vol. 1 No. 2, December 1983.

Maryland, 1981. Maryland Automated Geographic Information System. Maryland Department of State Planning Publication #81-13, Revised April, 1981.

Tomlin, C. Dana 1983. An Introduction to the MAP Analysis Package. 1983 National Conference on Resource Management Applications: Energy and Environment, CERMA International Conference Series, Aug. 22-26, 1983. San Francisco, CA.

Tomlin, S.M. 1982. Market analysis for the Mount Tom ski area of Holyoke, Mass. Staff paper 82-3, Department of Agricultural Economics and Rural Sociology. University of Connecticut College of Agriculture and Natural Resources.