Report to the

Legislative Commission on Minnesota Resources

WATER ALLOCATION AND MANAGEMENT

VOLUME II

DEVELOPING A GEOGRAPHIC INFORMATION SYSTEM FOR MINNESOTA WATER RESOURCES

AUGUST, 1987

University of Minnesota

Water Resources Research Center

Cooperating Departments:

Geography
Soil Science
Agricultural Engineering
Geology

Report to the

Legislative Commission on Minnesota Resources

WATER ALLOCATION AND MANAGEMENT

VOLUME II

DEVELOPING A GEOGRAPHIC INFORMATION SYSTEM FOR MINNESOTA WATER RESOURCES

AUGUST, 1987

WATER RESOURCES RESEARCH CENTER
UNIVERSITY OF MINNESOTA
ST. PAUL, MN 55108

Project Coordinator

Dwight A. Brown

Principal Investigators

Dr. Donald G. Baker, Soil Science
Dr. Dwight A. Brown, Geography
Dr. Philip J. Gersmehl, Geography
Dr. Curtis Larson, Agricultural Engineering
Dr. John L. Nieber, Agricultural Engineering
Dr. Hans-Olaf Pfannkuch, Geology
Dr. Richard Skaggs, Geography

PREFACE

This research was performed as part of a Water Resources Research Center multidisciplinary project supported by contract with the Minnesota Department of Natural Resources. Funds provided by the Legislative Commission on Minnesota Resources.

Separate reports detailing each phase of this project have been published. Interested parties are encouraged to contact the WRRC to obtain copies of these reports.

Water Resources Research Center
The Graduate School
The University of Minnesota
866 Biological Sciences Center
Saint Paul, Minnesota 55108

Patrick L. Brezonik, Director
Elizabeth Espointour, Secretary

DEVELOPING A GEOGRAPHIC INFORMATION SYSTEM FOR MINNESOTA WATER RESOURCES

FINAL REPORT SUMMARY

Water Resources Research Center
The University of Minnesota
Dwight Brown
Project Coordinator

BACKGROUND

The development of Minnesota's first generation of geographic information system (GIS) in the late 1960s and early '70s was a pioneering effort that has served the state very well. However, as the system has matured, federal, state, and local governing bodies are asking it to address new types of questions that the data and system are not designed to handle. In addition, a strong push toward regional and local control over resource decisions has resulted in increasing use of the state's GIS facilities by those not familiar with the potential for serious conceptual error. These users generally lack the funds to pay for the appropriate data and for the expert advice that they need. Too often they fail to realize their own need for expert help and treat the GIS as an expert system, which it emphatically is not.

The Planning Information Center has historically lacked resources for GIS system development, modernization of data collection, and expert support for the new and numerous nonexpert users.

This project address those facets of the problem outlined above that affect water resources planning and management. We recognize the potential for misuse and have guided our recommendations toward minimizing that potential. In most respects the problems of data collection and management in a GIS are general and apply to a much wider array of natural resources planning data needs than just water resources.

OBJECTIVES

1) To develop design criteria for the next-generation geographic information systems that will support water resources simulation procedures. These criteria include the way the next generation of resource data are collected and the way data should be stored and manipulated in a second-generation GIS.
2) To demonstrate the use of simulation procedures as a tool to aid planning and resource management decisions.
3) To conduct an analysis of the effect of drought on state-wide grain production for use in drought scenarios in the study of water value conducted by economists working through NRRI at Duluth.

These objectives focus on helping the State meet the 6 responsibilities outlined by the Water Planning Board in 1983.

INSTITUTIONAL ARRANGEMENT

The Water Resources Research Center, with Dwight Brown as project coordinator and one of the principal investigators, entered into a contract with the Department of Natural Resources to address the development of GIS data and file design criteria, the needs of water simulation procedures, and the analysis of the frequency and impact of drought on grain production in Minnesota.

PROJECT PRODUCTS

Relevant project reports are listed in parentheses.

1. Develop design criteria of a collection, storage and manipulation of precipitation data for use in GIS. (See WRRC Special Report 10, Chapters 2, through 5)
2. Develop design criteria of a collection, storage and manipulation of temperature data for use in GIS. (Temperature was determined to have less spatial variability than precipitation and the precipitation recommendations are sufficient. See WRRC Special Report 10, Chapters 2, 3 and 4)
3. Summarize precipitation parameters for NRRI's IPASS analyses. (See WRRC Special Report 10, Chapters 3 and 4)
4. Examine Minnesota's precipitation probabilities and methods of inclusion in GIS file structure of PIC. (See WRRC Special Report 10, Chapters 2, 3 and 4; WRRC Special Report 15)
5. Analyses of the spatial and serial relationships of drought severity for Minnesota. (See WRRC Special Report 15)
6. Summarize drought severity for use in NRRI's IPASS analyses. (See WRRC Special Report 9)
7. Design of soil water file structure for PIC. (See WRRC Special Report 10, Chapters 8 and 10)
8. Create soil water storage and infiltration-rate change files for a demonstration basin where sufficient soils mapping data exist. (See WRRC Special Report 14)
9. Design and create links between soil water files and surface water systems produced from existing PIC files. (These files proved to be inadequate for such analysis and as a result we have recommended how appropriate files should be developed based on detailed precipitation, topographic and soil survey data. See WRRC Special Report 10, Chapters 8, 10, 13, and 14.)
10. Design and create links between the soil water files and atmospheric data files. (See WRRC Special Report 10, Chapters 2, 3, 4, 8 and 10)
11. Validate and demonstrate surface water simulation applications based on appropriate GIS data with stream flow data provided by DNR. (See WRRC Special Report 11, 12, and 13)
12. Link models of land cover and cultural practices that affect soil water loss, recharge rates, and storage capacity to the way we collect and classify land cover and land use data. (See WRRC Special Report 10, Chapters 6 and 7)
13. Create a structural design for development/population geography models that link NRRI's IPASS outputs to Minnesota's regional geographic situation and resource context. (Early project discussions concluded that the inputs to the IPASS model were insufficient for running it for spatial units at the county or smaller level. We therefore expanded our efforts in other areas including groundwater and soil erosion modeling GIS file requirements. We were asked originally to cut these from the proposal but their importance to major water resource issues related to water quality and availability and the unity of the hydrologic cycle dictated that we give them some consideration. See WRRC Special Report 10, Chapters 9, 13, and 14.)
14. Support the development of microcomputer version of EPPL (EPPL7) by PIC. This was done with $10,000 provided to the WRRC project by DNR for use of PIC facilities. After discussion with Les Maki to determine its feasibility, Bill Becker agreed that this would satisfy many of the mutual objectives of WRRC, PIC, and DNR.
15. Program the computer to convert Landsat data files that are rectified in an ERDAS image analysis system to the EPPL7 system developed by PIC. (See WRRC Special Report 16.)

OTHER PRODUCTS

Work on this project provided financial support and research experience in water resources and geographic information systems for 16 graduate and undergraduate students from 5 departments at the University.

Nine oral and poster presentations based on this research were made to various national and regional professional organizations in an effort to obtain criticisms and suggestions from peers.

COOPERATING FACULTY

Dr. Donald Baker (atmosphere models, soil water balance models)
Dr. Dwight Brown (surface water systems, geog. info. systems)
Dr. Phil Gersmehl (soil water budget, geog. info. systems)
Dr. Curtis Larson (soil water submodels)
Dr. John Nieber (soil water submodels)
Dr. Olaf Pfannkuch (groundwater submodels)
Dr. Richard Skaggs (Atmosphere models; stochastic modeling)

REPORTS

D. Brown, C. Gersmehl, J. Drake, and R. Skaggs, Crop Production Response to Moisture Supply in Minnesota, Water Resources Research Center, Special Report 9.

D. Brown and P. Gersmehl, Editors. 1987. File structure design and data specifications for water resources geographic information systems, Water Resources Research Center, Special Report 10. (with 15 separately authored chapters)

1. D. Brown and P. Gersmehl, Introduction: toward water resources analysis with geographic information systems.
2. P. Gersmehl and D. Brown, File structure and cell size considerations for a water resources GIS.
3. R. Swerman and D. Baker, Precipitation network density requirements for short-term analysis.
4. J. Drake and R. Skaggs, Climatic network density analysis.
5. R. Swerman, D. Rushy, and D. Baker, Precipitation Data for a Water Resources GIS.
6. P. Gersmehl, K. Anderson, R. Greene, N. Dunning, C. Gersmehl, and D. Brown, Hydrologic classification of land cover.
7. C. Gersmehl, Land cover data for a water resources GIS.
8. J. Nieber and I. Lopez Bakovic, Soil water systems analysis and modeling.
9. T. Beach, A review of soil erosion modeling.
10. P. Gersmehl, J. Corbett, and R. Greene, Soil data for a water resources GIS.
11. J. Corbett and P. Gersmehl, Terrain data for a water resources GIS.
12. D. Brown, K. Anderson, and P. Gersmehl, Hydrographic data for a water resources GIS.
13. H. Pfannkuch, P. Jones, and L. Guo, Groundwater systems analysis and modeling.
14. S. Beach, Groundwater data for a water resources GIS.
15. P. Gersmehl and D. Brown, Conclusions and summary of recommendations for a water resources geographic information system.

R. Skaggs, and D. Brown, Relationship Between Climate and the Mean Annual Flow of the Mississippi River at Saint Paul, Minnesota, Water Resources Research Center, Special Report 11.

K. Anderson, J. Corbett, N. Dunning, C. Gersmehl, R. Greene, P. Gersmehl, and D. Brown, Twin Cities Surface Water Simulation Modeling Demonstration, Water Resources Research Center, Special Report 12.

K. Anderson, Bear Creek Surface Water Simulation Modeling Demonstration, Water Resources Research Center, Special Report 13.

I. Lopez Bakovic, Modeling Soil Water Variability, Water Resources Research Center, Special Report 14.

R. Swerman, D. Baker, and R. Skaggs, Minnesota Drought, Water Resources Research Center, Special Report 15.

K. Anderson and B. Scheer, A Computer Program to Exchange ERDAS and EPPL7 Data Files, Water Resources Research Center, Special Report 16.

SUMMARY OF FINDINGS AND RECOMMENDATIONS

WATER RESOURCES RESEARCH CENTER

FILE STRUCTURE DESIGN AND DATA SPECIFICATIONS FOR WATER RESOURCES GEOGRAPHIC INFORMATION SYSTEMS

Special Report 10.

D. Brown and P. Gersmehl, Editors.

INTRODUCTION: TOWARD WATER RESOURCE ANALYSIS WITH GIS

Chapter 1 by Brown and Gersmehl, WRRC, University of Minnesota

ISSUES

Managers of water resources face uncertainties about the quantity, quality, and variability of both supply and demand; these uncertainties are confounded by the impacts of human actions, both planned and unanticipated. The historical record, by itself, is not a reliable tool for predicting unless all environmental conditions remain constant. In this context, "constant" includes many changes that are part of the normal environment (e.g. weather fluctuations, climatic cycles, geologic erosion, and the gradual infilling of lake basins to become marshes).

To predict consequences of unprecedented changes or human activity, we must find another tool to supplement the historical record. Hydrologic simulation based on physical principles is one powerful alternative. Necessary data can come from geographic information systems (GIS), if file structures, cell sizes and variables are matched to the simulation requirements. The utility of geographical data bases can be compromised if we begin with less-than-optimal choices about scale of analysis, classification of data, and relational structure of the data files.

FINDINGS

The workings of the hydrologic cycle impose a fundamental unity on the water resource picture of the state, but water budgets in different parts of the state are significantly different.

Control over local water resources should rest with the people in the local area, responsible for determining appropriate use.

Analysis should precede action when decisions affect resources; the GIS should therefore be legally sensitive to such matters as property rights, accessibility, and privacy.

Local governments may eventually be connected to a statewide data base, but near-future support will come from microcomputers.

With these four ideas in mind, project personnel began a two-year investigation of hydrologic simulation models and their data requirements. The overriding concern was to design a structure in which data of all kinds could be stored for easy retrieval and analysis. Various chapters of this report will discuss each major sub-part of the hydrologic system, together with its linkages to other aspects of the system. The focus is on regional scales of analysis (100-10,000 square kilometers).

RECOMMENDATIONS

After extensive examination of alternative data-handling methods, we recommend a point-sampling procedure that can store and relate the following variables: precipitation, streamflow, ground-water flow, stream networks, artificial drainage systems, water quality, lake basins and wetlands, bedrock aquifers and aquicludes, surficial materials, soil traits, topography, land cover, and land use. Much of this information already exists in various places throughout the state, but the data should be translated into a form that allows different kinds of information to be combined and applied to specific water problems of a region.

FILE STRUCTURE AND CELL SIZE CONSIDERATIONS FOR A WATER RESOURCES GIS

Chapter 2 by Gersmehl, Brown, and Anderson, WRRC, University of Minnesota

ISSUES

As a basis for predicting water use, runoff potential, groundwater recharge, and other hydrologic phenomena, a resource manager needs good information about the soil, climate, land cover, terrain, and other environmental conditions at a place. To translate existing data into computer-usable form, and to gather future data in an efficient manner, the designers of a water-resources Geographic Information System must deal with three fundamental issues:

1) Types of hydrologic simulations. Every analytical tool has its own theoretical foundation, treatment of area and time, and scale of analysis, which affect the type of input data necessary and the validity and form of the output. Users must be aware of these differences in making their choice of a hydrologic simulation.
2) Differences in data quality. Each source of data has its own spatial resolution, temporal scope, and measurement precision. Users of a GIS must be aware of data limitations in order to obtain appropriate answers for the questions they have formed.
3) Output from a geographic information system or a computer simulation has a "mystique" that makes it extremely easy to imply much greater precision than can be justified on the basis of the structure of the simulation or the quality of input data. Users of a GIS must be aware of its limitations and adjust its output accordingly, in order to avoid product-liability problems.

FINDINGS

The goal is to build a water-resources GIS that could describe areas of moderate size with reasonable precision, make regional inventories with reasonable accuracy, facilitate relational use with other data, and allow addition of better data when they become available. To obtain a basis for recommendations, we examined many data-handling strategies. Our findings include the following:

There is a striking lack of analytical models capable of simulating groundwater, surface runoff, and soil moisture as an integrated system at the "regional" scale (100 to 10,000 square kilometers).

A planner or resource manager in Minnesota has access to a rich array of water-resources data, although many data files have features that limit their usefulness for hydrologic simulation.

A water-resources GIS must relate all data to a single spatial framework in order to provide input for a hydrologic simulation (such as the Universal Soil Loss Equation or Peak Flood Model). A simulation cannot produce valid results unless the input data are spatially related (i.e. come from exactly the same locations).

The U.S. Public Land Survey has some severe locational problems and statistical effects that render it undesirable as a spatial framework for a water-resources GIS.

A uniform metric coordinate base for all GIS files will minimize spatial bias in coding, allow easier merging of data, and be electronically efficient for data storage and manipulation.

The point-counting (inventory) and the area-tagging (classificatory) approaches to data collection are fundamentally incompatible. Each approach has its own specific purposes and methods, and data collected in one way should not be used to answer questions that require data of the other type.

The EPPL/MLMIS and Arc/Info systems at the Planning Information Center lack "pedigree" records, which would allow a user to trace the origin or examine the input resolution of a data file, in order to determine its applicability for a particular use.

RECOMMENDATIONS

On the basis of these findings and the results of many other investigations at various scales, we make the following recommendations concerning the file structure for data in a water-resources geographic information system:

1) Fund development of a regional-scale, hydrologically-integrated, physically-based simulation model which can use GIS data in order to solve emerging problems of water quantity and quality.
2) Develop a count-based GIS to supplement current EPPL and Arc/ Info data bases. This system would concentrate on handling detailed, point-relational data of the kind needed for hydrologic simulation. In the meantime, EPPL can serve as an adequate display system for inventory data (see Anderson et al. 1987).
3) Adopt the UTM coordinate base and a one-kilometer resolution as the base file for environmental analysis. Relating the corners of 40-acre Public Land Survey parcels to metric coordinates is not sufficient, given the inherent problems with PLS.
4) Modify EPPL (and other state GIS packages) to include a better identification system for every data file. Header records should note the file's "pedigree" and scale. EPPL7 currently keeps track of vertical resolution as an explicit header item; entries on horizontal resolution, measurement precision, temporal validity, and data source should be mandatory in the header of all files.
5) Develop a system that uses the resolution information in a file header to control display of data. The implied precision of data output should not exceed the precision (spatial, temporal, or mathematical) of the least precise input variable. This rule is especially important in the case of sample data, which (because they cannot be both locationally and statistically valid) should be displayed on a map at a resolution that is at least a full order of magnitude less detailed than the sample density.

PRECIPITATION NETWORK DENSITY REQUIREMENTS FOR SHORT-TERM ANALYSIS.

Chapter 3 by Swerman and Baker, WRRC, University of Minnesota

ISSUES

As a basis for predicting water use, runoff potential, groundwater recharge, and other hydrologic phenomena, a resource manager needs good information about the soil, climate, land cover, terrain, and other environmental conditions at a place. To translate existing data into computer-usable form, and to gather future data in an efficient manner, the designers of a water-resources Geographic Information System must deal with two fundamental issues:

1) The spatial variability of precipitation. A stringent level of accuracy of precipitation measurement is needed to validate (and use) hydrologic simulation models (both event and continuous-synthesis) for areas of a few tens to a few thousand square miles. The density of gauges needed to make an accurate record of the spatial pattern of rainfall is an unknown but usually large number that depends on the characteristic patterns of these events and how accurately they must be recorded. Our recommendations must accommodate the tradeoff between cost and accuracy in precipitation measurement.
2) The costs of measurement error in urban areas. The impact of error in precipitation measurement on the outcome of the hydrologic model is particularly acute in urban areas for two reasons. First, the abundance of impermeable surfaces greatly accentuates the response of watersheds to small differences in rainfall. Second, the cost of flood damage and service interruption is high because of capital improvements and population density.

FINDINGS

The goal is to build a water-resources GIS that could describe the precipitation pattern in areas of moderate size with reasonable precision, calculate regional totals with reasonable accuracy, facilitate relational use with other data, and allow addition of better data when they become available. To obtain a basis for recommendations, we examined the efficiency of several different networks of gauges as measures of the size and pattern of storm events in the Twin Cities area. Our findings include the following:

The ability to make an accurate record of the volume and pattern of precipitation events differs with complexity of the storm. In general, winter storms have broader patterns that require fewer gauges for accurate measurement than summer storms.

The accuracy of measurement of average rainfall for an 1100 mi^2 segment of the Twin Cities metropolitan area improves rapidly from 1 to about 25 gauges; after approximately 70 gauges are sampled the results improve very slowly.

About 50 gauges are needed to report the volume of rainfall from a single event for an area of 1100 mi^2 to within 90% correct at least 95% of the time.

An attempt to calibrate a runoff simulation or to model the general pattern of soil water and groundwater recharge on an event basis with fewer than 15 gauges for an area of 1100 mi^2 will add substantial error to the analysis. Modeling soil water on a site scale will require a much large number of gauges.

RECOMMENDATIONS

On the basis of these findings and the results of other investigations at various scales, we make the following recommendations concerning the methods for gathering precipitation data for a water-resources geographic information system:

1) Maintain a network of at least 70 precipitation gauges in the Twin Cities metropolitan area to insure that there are a sufficient numbers of good records for any individual storm event.
2) In important groundwater recharge areas, where it is possible to obtain good information of land cover, land use, soils, and topography, the precipitation network may be the factor that is limiting the accuracy of hydrologic analysis. Especially in these areas, the number of precipitation gauges should be increased to or maintained at a level of one gauge per 40 mi^2 (25 gages per 1000 mi^2) to allow monitoring of the effect of surface changes on the budget of important aquifers.

CLIMATIC NETWORK DENSITY ANALYSIS

Chapter 4 by Drake and Skaggs, WRRC, University of Minnesota

ISSUES

As a basis for predicting crop yields, water use, and runoff potential, a resource manager needs good information about the climate in an area. To translate existing information into computer-usable form, and to gather future information in an efficient manner, the designers of a water-resources Geographic Information System must deal with three fundamental issues:

1) Intrinsic variability of precipitation. The input of precipitation into a local hydrologic system varies greatly, both spatially and temporally, over any scale of analysis. Our recommendations must make allowances for that variability.
2) Uneven data sources. Data from the National Weather Service (NWS) Cooperative Observer Network are readily available, but they vary in spacing and duration of record. In general, these data are not adequate for hydrologic simulation on an hourly or daily basis. Several specialized precipitation networks have been established in recent years, but their frequency, methods, and seasonality of observations often do not match those of the NWS. Moreover, these data are not always available to other users. Our recommendations must make allowances for the lack of strict comparability in daily or historic records.
3) Necessity for relational structure in the GIS. In order to provide data for a predictive simulation (such as the Universal Soil Loss Equation or the Peak Flood Model), a GIS must relate all data to a common coordinate system. When that is done, the simulation computer program can get information on climate, soil, slope, and land cover at exactly the same point and thus reach a valid conclusion about erosion or runoff.

FINDINGS

The goal is to build a data base that can describe the climate of Minnesota with reasonable precision, facilitate relational use with other data, and allow easy addition of data as they become available. To obtain a basis for recommendations, we made a series of empirical investigations in two study areas. Our findings include the following:

Correlation-fields analysis of annual precipitation reveals some sharp gradients in the southeastern part of the state, especially near the Mississippi River, and a greater degree of areal consistency in the southwestern part of the state.

The existing network of observation stations is generally adequate for annual measurements in the southwest, but a slightly greater density would be desirable in the southeastern part of the state.

A kriging method, with a reduced search radius to accommodate spatial gradients of precipitation, is the preferred method of interpolation between observation stations.

Interpolation error for annual precipitation is only about 3 percent in the southwest and between 6 and 16 percent in the southeast.

Interpolation error for daily precipitation exceeded 300 percent in both study areas, indicating that the NWS network is simply not closely spaced enough for hydrologic simulation on a short-term basis.

RECOMMENDATIONS

On the basis of these findings and the results of other investigations, we make the following recommendations:

1) Climatic data should not be stored in grid-cell form in a water resources geographic information system. Rather, the climatic information should be obtained for each application from point-data files that are relevant for the question being investigated.
2) The State Climatologist should establish a comprehensive data base from all observation networks in the state; in time, data from earlier years should be added to it.
3) The State Climatologist should provide electronic access to this data base; logical access should be by user-specified space coordinates, time coordinates, and desired climatic elements (e.g. minimum temperature, monthly precipitation, etc.).
4) A package of programs to prepare climatic data for use with a GIS should be available in the user interface with the State Climatologist's data base; this package should include, at the minimum, programs to grid data and programs to assign values to GIS data cells using simple (unweighted) averaging, Thiessen polygons, and isopleth interpolation.

PRECIPITATION DATA FOR A WATER RESOURCES GIS

Chapter 5 by Swerman, Ruschy, and Baker, WRRC, University of Minnesota

ISSUES

As a basis for predicting crop yields, water use, and runoff potential, a resource manager needs good information about the precipitation input to a local hydrologic cycle. To translate existing information into computer-usable form, and to gather future information in an efficient manner, the designers of a water-resources Geographic Information System must deal with three fundamental issues:

1) Intrinsic variability of precipitation. The input of precipitation into a local hydrologic system varies greatly, both spatially and temporally, over any scale of analysis. Our recommendations must make allowances for that variability.
2) Data requirements for hydrologic simulation. Computer programs have become fairly effective in predicting floods from small watersheds, but they require input data at a fine spatial and temporal resolution.
3) Necessity for relational structure in the GIS. In order to provide data for a predictive simulation (such as the Universal Soil Loss Equation or the Peak Flood Model), a GIS must relate all data to a common coordinate system. When that is done, the simulation can get information on climate, soil, slope, and land cover at exactly the same point and thus reach a valid conclusion about erosion or runoff.

FINDINGS

The goal is to build a data base that can describe the precipitation patterns of Minnesota with reasonable precision, facilitate relational use with other data, and allow easy addition of data as they become available. To obtain a basis for recommendations, we made a detailed evaluation of the existing observation networks. Our findings include the following:

Data from the National Weather Service (NWS) Cooperative Observer Network are readily available, but the stations are unevenly spaced and their records are of unequal duration. In general, these data are not adequate for hydrologic simulation on a short-term basis.

Several specialized precipitation networks have been established in recent years, but they vary more widely than the NWS in time and methods of observation; some of these networks function only in summer.

The greatest density of precipitation observation stations is in the Twin Cities metropolitan area.

The existing network of observation stations is generally adequate for annual measurements in the southwest, but a greater density would be desirable in the southeastern part of the state.

Differences in observation time, variations in observation practice, and occasional observer error make it all but impossible to analyze individual precipitation events unless they are separated by several days.

The number of stations making hourly precipitation observations is extremely small and their spatial distribution is very uneven. Even so, the mass of data generated by these stations is difficult to use. The lack of accessible data for particular locations is one of the major limitations on hydrologic simulation in Minnesota.

For the most part, the computer data base of the State Climatologist contains only measurements that were made since 1972.

RECOMMENDATIONS

On the basis of these findings and the results of other investigations, we make the following recommendations:

1) The data base of the State Climatologist should be kept up to date, and data from earlier years should be added to it.
2) The State Climatologist should provide electronic access to this data base; logical access should be by user-specified space coordinates, time coordinates, and desired climatic elements [see Chapter 4 for details].
3) Programs to prepare precipitation data for use with a GIS should be available in the user interface; these programs should include routines to grid data and several options for assigning values to GIS data cells, including simple (unweighted) averaging, Thiessen polygons, and isopleth interpolation.
4) Establishment of a "floating" network of hourly precipitation gauges would be desirable if we would like to refine our flood-forecasting ability.

HYDROLOGIC CLASSIFICATION OF LAND COVER

Chapter 6 by Gersmehl, Anderson, Greene, Dunning, Gersmehl, and Brown, WRRC, U of Minn

ISSUES

As a basis for predicting runoff potential, water use by plants, and crop yields, a resource manager needs good information about the land cover in a watershed. To translate existing information into computer-usable form, and to gather data efficiently in the future, the designers of a water-resources Geographic Information System must deal with four fundamental issues:

1) The distinction between land cover and land use. Land use is an economic concept; land cover is a physical idea. Typical maps of land use categories can hide information that is extremely important for hydrologic simulation.
2) he temporal changeability of land cover. The hydrologic effects of bare plowed ground differ from those of a mature crop or a harvested field. Our recommendations must make allowances for seasonal cycles, annual changes (e.g. crop rotation), and long-term trends (e.g. forest growth, land abandonment, or urbanization).
3) The incompatibility of the tag and count perspectives in mapping. A tag (classificatory) map of land cover is an attempt to delimit areas that are reasonably homogeneous. To map large areas, one must generalize, usually by omitting parcels that are too small to be mapped separately. The inevitable result is an underestimate of the areal extent of some features (and those omitted may have great hydrologic significance). By contrast, a count procedure uses point sampling to provide a better estimate of area covered by particular features, but it should not claim to show locations of individual features. A set of point samples will "hit" only a few cases of a widely dispersed but individually small kind of cover. Mapping the locations of sample points will give an erroneous picture of the spatial pattern of that kind of cover. Our recommendations must recognize that a GIS cannot classify individual parcels accurately (especially at small scale) and still do a statistically valid inventory of land cover for a large area.
4) The necessity for relational structure in the GIS. In order to provide data for a simulation such as the Agricultural Non-Point Source Pollution Model or the Universal Soil Loss Equation, a GIS must relate all data to a common coordinate system. When that is done, the simulation can get information on land cover, soil, and slope for the same places and thus reach a valid conclusion about erosion or runoff. Our recommendations must make allowances for the nature of the data files that will be used with the land cover file in solving a resource problem.

FINDINGS

The goal is to build a land-cover data base that can describe tracts of land with reasonable precision, make regional inventories with acceptable accuracy, facilitate relational use with other data, and allow easy addition of better data when they become available. To obtain a basis for recommendations, we examined many different data-handling strategies. Our findings include the following:

Differences in previous crops and present tillage methods can cause erosion to be as much as thirty times greater from one field than from another with the same crop, climate, slope, and soil type.

Existing land-use maps published by the Metropolitan Council and the USGS are not valid as sole input for hydrologic simulations. Differences in proportion of impervious surface and connectivity of drainage systems can cause more than a tenfold variation in flood potential from areas all mapped as "residential land use."

Land-cover classes that occupy as little as ten percent of the area of a small watershed can produce more than half of the total runoff.

Reader errors and seasonal inconsistencies in aerial photo interpretation are often high enough to limit the accuracy of hydrologic simulation, yet they rarely are reported by planning agencies.

Recording complete data at sample points should have higher priority than refining maps of arbitrarily defined land-cover classes.

RECOMMENDATIONS

On the basis of these findings and the results of many other investigations at various scales, we make the following recommendations concerning the file structure for land-cover data in a water-resources geographic information system:

1) Form a Task Force of people from State agencies to refine and/or modify the framework land-cover classification proposed herein.
2) Use a relational point-sampling system, as described in Chapter 2, to enter land-cover and land-use data into the GIS. Recording details of surface condition and drainage connectivity is more important for hydrologic studies than improving map resolution.
3) Base the system on Universal Transverse Mercator coordinates. The tendency for land uses to be aligned with survey lines can introduce big statistical aberrations in a section-based system.
4) Once the GIS is operational, use the climate, terrain, and soil data to evaluate and improve the land-cover files. For example, ambiguity in Landsat images of marshland can be reduced by noting poorly-drained soils in the soil file and low areas in the terrain file. Inclusion of building permits, plat maps, and zoning data in the GIS may clarify patterns in urban-fringe areas.
5) Pay special attention to methods used to make maps from a point-inventory system, in order to avoid misinterpretation. Point sampling allows us to estimate the extent of a given land-cover type quite accurately, but it does not permit precise description of small areas. For that reason, the system should simply not be permitted to portray data at individual sample points; instead, it should display only the percentages of larger areas that fall into particular categories. As a rule of thumb, the output map should display data at a resolution that is a full order of magnitude less detailed than the sample data. If more detailed information is needed, field surveys are necessary.

LAND COVER DATA FOR A WATER RESOURCES GIS

Chapter 7 by C. Gersmehl, WRRC, University of Minnesota

ISSUES

To gather land-cover information efficiently, the designers of a water-resources Geographic Information System must deal with four fundamental issues:

1) Temporal changeability of land cover. Land cover can change markedly from year to year, season to season, even day to day. Satellite imagery has relatively poor spatial resolution but better temporal resolution than aerial photographs or ground surveys. Our recommendations should consider the tradeoffs between map accuracy and timeliness in making maps of land cover.
2) The weather vulnerability of remote sensing. Benign atmospheric conditions are important to the success of most kinds of civilian remote sensing. Our recommendations should make allowance for the fact that place-to-place variations in air transparency and cloud cover can disrupt plans to use Landsat tapes (or other satellite imagery) as a source of data on land cover.
3) Automated image analysis. Modern software makes it possible to "train" a computer to recognize the reflectance traits of areas of known land cover. With an adequate selection of training fields, a computer can classify most of the individual elements (pixels) of a Landsat scene into spectral categories that are reasonably homogeneous. Our recommendations should assess the utility of automated image-processing in land-cover analysis.
4) The necessity for relational structure in the GIS. When data are related to a common coordinate system, a simulation can get data on land cover, soil, and slope at the same place and thus reach a valid conclusion about erosion or runoff. Our recommendations must make allowances for the nature of the data files that will be used with the land cover file in solving a resource problem.

FINDINGS

The goal is to build a land-cover data base that can describe tracts of land with reasonable precision, make regional inventories with acceptable accuracy, and facilitate relational use with other data. To obtain a basis for recommendations, we investigated the methods, accuracy, and efficiency of automated classification of Landsat imagery. Our findings include the following:

Before we can get them, Landsat images have undergone analog-to-digital translations and several radiometric and geometric corrections that inevitably introduce some locational imprecision.

A rectified image can be related directly to other data files on UTM coordinates, but locations on such an image are only approximate and may deviate by as much as 100 meters from "true" positions.

Satellite sensors "see" energy reflected from a large area (almost two acres per data "point"), and therefore small features like houses are just part of a a spatially averaged signal.

Spectral data can overlap in complex ways, making it difficult to separate buildings from gravel pits or marshes from corn fields.

Despite these limitations, we can achieve an accuracy of 60 percent in classifying 100-meter pixels into level II land cover types.

Lumping 51 individual spectral "signatures" into 16 composite groups for training did not reduce accuracy significantly, but a large number of signatures is important in order to span the range of spectral characteristics within each major land-cover class.

Fine spatial resolution is expensive. 100-meter pixels need one-fourth the storage space of 50-meter pixels, but seem to provide almost as much accuracy in identifying broad land-cover categories.

Healthy vegetation produces a strong reflectance in the infrared bands on multi-spectral scanner data tapes; this fact may allow a fairly direct gauge of evapotranspiration in a watershed.

RECOMMENDATIONS

On the basis of these findings and the results of other investigations, we make four recommendations concerning the use of Landsat imagery as a source for land-cover data in a water-resources geographic information system:

1) Use a supervised classification and obtain at least three or four training areas for each category, so that the composite signature spans a range of reflectances from examples of each land-cover type with demonstrated hydrologic significance.
2) Store data in grid form, with 100 pixels per square-kilometer data cell, registered to Universal Transverse Mercator coordinates. The GIS should not output the data at that resolution, however, nor try to match Landsat values one-on-one with other data at that fine a resolution, because the satellite imagery has an unavoidable, inconsistent, but significant positional error.
3) Once the GIS is operational, use the climate, terrain, and soil data files to evaluate and improve the land-cover files. For example, ambiguity in Landsat images of marshlands can be reduced by noting poorly-drained soils in the soil file and low areas in the terrain file. Inclusion of building permits, plat maps, and zoning data in the GIS may clarify patterns in urban-fringe areas.
4) Pay special attention to methods used in making maps from a Landsat classification, in order to avoid misinterpretation. The automated classification procedure allows us to estimate the areal extent of a given land-cover type quite accurately, but it does not permit precise description of small areas. For that reason, the system should simply not be permitted to portray data for individual pixels; instead, it should display only percentages of larger areas that fall into particular categories. As a rule of thumb, the output map should display data at a resolution that is a full order of magnitude less detailed than the sample data. In the case of Landsat classifications, that means that a square-kilometer data cell is the minimum size of area that should be shown on an output map.

SOIL WATER SYSTEMS ANALYSIS AND MODELING

Chapter 8 by Nieber and Lopez Bakovic, WRRC, University of Minnesota

ISSUES

As a basis for predicting infiltration, evapotranspiration, deep percolation, and groundwater recharge at a particular place, a resource manager must have good information about the water status of the soils there. To translate existing data into computer-usable form, and to gather future information in an efficient manner, the designers of a water-resources geographic information system must deal with three fundamental issues:

1) The temporal and spatial variability of soil moisture. The amount of moisture in the soil at a point can change dramatically from day to day, and even more so from season to season or year to year. Moreover, soil moisture varies spatially between points, even over very short distances, which makes it risky to try to extrapolate from existing measurements. Our recommendations must include ways of dealing with the intrinsic variability of soil moisture.
2) Processes of water movement in soils. Principles of soil physics can help explain the quantity and rate of vertical and lateral flow of water in soils. To simulate flows of moisture, it is necessary to incorporate these physical principles, or simplifications of these principles, into mathematical models of the flow processes. Our recommendations must accommodate the data requirements of soil moisture simulations.
3) Principles of interpolation between measurement points. To analyze the spatial distribution of soil water, one can make measurements of soil water status at numerous discrete points, or one can use an interpolation model to predict soil water status for the areas between measurement points. Interpolation requires knowledge of the spatial distributions of soil water properties (field capacity, wilting point, hydraulic conductivity, etc.). Our recommendations must take into account the feedback loops that exist between field measurements and computer simulation in soil moisture modeling.

FINDINGS

Our efforts concentrated on analyzing soil water spatial variability, developing a Soil Water Balance Model (SWBM), and examining the effects of scale. Scale is important in the process of selecting data appropriate to run hydrologic models and in extending soil water analyses from points to areas.

A Soil Water Balance Model (SWBM) was developed for one-dimensional analysis of hydrologic processes. The model uses simplified mathematical equations of the vertical flow of water in the soil profile, and the output from the model is sensitive to inputs such as soil properties and weather data. The model can characterize the temporal variability of soil water status at discrete points.

Geostatistical methods can characterize the spatial variability of soil water. These methods help us analyze the spatial structure of point measurements and predictions (such as those by SWBM) of soil water status. Geostatistical tools require some theoretical understanding, but the results are somewhat easier to visualize than some of the traditional statistical methods.

Modeling soil water systems is a complex task. The objectives and scale of analysis determine not only the hydrologic models and other tools that may be used, but also the type of variability involved and the nature of the output. The scale dependence of analytical methods and data-gathering strategies should warn against using large-scale analyses to make inferences about small-scale problems, or vice versa.

RECOMMENDATIONS

On the basis of these findings and the results of other investigations, we make the following recommendations concerning the file structure for soils data in a water-resources geographic information system:

1) Use known relationships and physical principles to augment field measurements of soil moisture. A single point measurement (or prediction) is not sufficient for characterizing the soil water status of an area, because of the potential spatial variations of soil moisture. Point measurements (or predictions) should be supplemented by studies to quantify the spatial distribution of soil moisture. The tools of geostatistical methods (e.g. semi-variogram analysis) can quantify the variability of soil water status in an area and identify those factors which most influence its distribution. In addition, Kriging can be used as a means of interpolating between measurement points and selecting sites for additional field studies.
2) Clearly differentiate between primary and derived data in a water-resources GIS. A point measurement has validity only at the place and time it was obtained. An interpolation between point measurements is one step farther removed from the real world, and should be clearly identified as such in subsequent analyses. Despite that caveat, interpolated data are essential for most hydrologic simulations, because the cost of gathering field measurements at the density needed for effective modeling is too great for most practical purposes.

A REVIEW OF SOIL EROSION MODELING

Chapter 9 by T. Beach, WRRC, University of Minnesota

ISSUES

To predict soil erosion, an analyst can choose among a number of simulation models. Each of these models has its own purposes, data requirements, strengths, and weaknesses. A geographic information system could serve as a source of data for many of these simulation models. To facilitate its use in this way, the designer of a water-resources GIS should consider three main issues:

1) Data requirements of sophisticated simulations. Some erosion models use a small number of variables to obtain a general estimate of erosion rates in an area. Others aim for greater precision by including more variables that can affect soil erosion. Simulation complexity, however, cannot substitute for inaccurate or generalized data. Our recommendations must take into account the tradeoff between faithfulness to real-world processes and ease of gathering data for a simulation.
2) Regional differences in erosional processes. Simulation models can deal with different aspects of soil erosion, including detachment, sheetflow, rill erosion, gullying, channel scour, deposition, and/or remobilization. These processes vary in absolute and relative importance in different environments. Our recommendations must consider the ability of particular erosion models to simulate the processes that are important in different parts of Minnesota.
3) Types of soil erosion models. Some models deal with single runoff events, whereas others simulate long-term averages. Physically based models try to mimic the underlying processes of erosion, whereas empirically based models predict erosion on the basis of relationships observed in similar settings in the past. Some models deal with erosion alone, whereas other erosion equations are parts of multi-component models that can simulate a variety of ecosystem processes. Our recommendations must accommodate the dissimilar structures and data requirements of different kinds of simulations.

FINDINGS

This chapter reviews a number of soil erosion simulation models, in order to identify their overall strengths, weaknesses, and data requirements. Many soil erosion models are still in a developmental stage and need further testing before use with a GIS. No individual model is best for each application. Each model has strengths and weaknesses. Our findings include the following list of basic characteristics of soil erosion models:

The Universal Soil Loss Equation (USLE) is perhaps the most widely used method of predicting soil loss in the world. It is an empirical equation with a relatively simple structure and limited data requirements. It works reasonably well at a field scale, but use at regional planning scales can pose problems.

People have modified the USLE in many ways to enable it to work in specific environments or to meet certain prediction goals. Some modifications include estimates of sediment yield and deposition; others add a physical base to make the simulation more sound.

The Water Erosion Prediction Project (WEPP) is a current attempt to replace the USLE. This model will deal with more aspects of erosion than USLE and will be physically based. The target date for completion of this project is August 1988.

AGNPS I and II are multi-component watershed models that predict erosion and sediment transport from single runoff events. These models apply to areas ranging from 2.5 to 23,000 acres. They are primarily process-based models that require more data than the USLE; the use of regional averages for some key inpu data can limit the precision of the output.

ANSWERS is a multi-component watershed simulation that predicts erosion and deposition from single runoff events. It is physically based and works with areas up to about 25,000 acres, which partly explains its very large data requirements.

CREAMS is another physically based multi-component model with large data requirements. It simulates erosion and deposition for field-sized areas and can handle either single runoff events or long-term sequences.

EPIC is a multi-component model applied to field-sized areas that simulates erosional effects on soil productivity. It uses a physically based model presented i 1975, but the whole model currently needs further validation. EPIC's erosion component requires relatively easily gathered data.

RECOMMENDATIONS

On the basis of this review of soil erosion modeling, we make the following recommendations concerning the selection of an erosion model for a water-resources GIS:

1) Select a soil erosion model whose data requirements can be met with available resources. The data requirements of some models are prohibitive. Additionally, even small errors with numerous variables can accumulate into a large overall error.
2) Select a soil erosion model that simulates the important processes in Minnesota and estimates the appropriate measures. For instance, if an estimate of sediment yield is needed, choose the MUSLE rather than the USLE for the GIS.
3) Use primary data as much as possible. The manuals for using these predictive models usually have guidelines on how to estimate data from secondary sources, but simulations are more reliable when based on relational field data [see chapters 2 and 10]. At the very least, primary data from a few selected sample points can supplement information derived from generalized maps and other indirect sources.

SOIL DATA FOR A WATER RESOURCES GIS

Chapter 10 by Gersmehl, Corbett, and Greene, WRRC, University of Minnesota

ISSUES

As a basis for predicting crop yields, water use, and runoff potential, a resource manager needs good information about the soils in a watershed. To translate existing information into computer-usable form, and to gather future information in an efficient manner, the designers of a water-resources Geographic Information System must deal with three fundamental issues:

1) Uneven quality of input data. In Minnesota, computerized soil maps are available for about 15 counties; another forty counties have modern soil surveys; and the remaining areas have no up-to-date soil data beyond the generalized Soil Atlas. Our recommendations for providing soil information to the GIS must make allowances for the wide range of input data quality.
2) Classificatory nature of a soil survey. A soil survey is an attempt to describe soil in each delimited area on a map accurately. To that end, soil surveyors have several conventions to deal with the soil variability that can occur within even a short distance. At the risk of oversimplifying, they follow a four-part rule: delimit soils that cover at least an acre, ignore inclusions of radically dissimilar soils if they cover less than ten percent of an acre, describe inclusions in the text if they are of moderate extent, and list dissimilar soils in the map legend as part of a soil complex when they cover more than about a third of the area mapped as having a given kind of soil. This procedure ensures a consistency that allows a trained map reader to infer actual soil patterns, but it makes calculating the areal extent of various soils difficult. Our recommendations must make allowances for the fact that a map cannot tag (classify) individual parcels of land accurately and still provide a precise count (inventory) of soils for a larger area.
3) Necessity for relational structure in the GIS. In order to provide data for a predictive simulation (such as the Universal Soil Loss Equation or the Peak Flood Model), a GIS must relate all data to a common coordinate system. When that is done, the simulation can get information on soil, slope, and land cover at exactly the same point and thus reach a valid conclusion about erosion or runoff. Our recommendations must accommodate the data that will be used with the soil file in solving a resource problem.

FINDINGS

The goal is to build a soil data base that would describe areas of moderate size with reasonable precision, make regional inventories with reasonable accuracy, facilitate relational use with other data, and allow easy addition of better data when they become available. To obtain a basis for recommendations, we examined many different data-handling strategies. Our findings include the following:

Point-counting methods typically have one-third to one-tenth as much error as area-tagging methods of the same spatial resolution.

The Minnesota Soil Atlas and computer files developed from it under-represent poorly drained soils by a factor of from two to ten.

Soils that occupy less than a fourth of the area can contribute more than half of the storm runoff from a watershed.

Statewide productivity ratings often err by more than 25 percent in describing yields for soils in individual counties.

Environmental models can achieve accuracies of 75 percent or better in "predicting" general soil patterns in unsurveyed counties.

RECOMMENDATIONS

On the basis of these findings and the results of many other investigations at various scales, we make the following recommendations concerning the file structure for soils data in a water-resources geographic information system:

1) Strictly avoid use of Soil Atlas data (as stored in the Planning Information Center computer) to determine the areal extent of peat bogs, wetlands, or other soil features. This tag-based file is designed to categorize broad areas accurately, which it does well, but it is simply not appropriate for answering inventory questions.
2) Use a relational point-sampling system, as described in Chapter 2, to enter data from soil surveys into the GIS. A sampling intensity of 4 to 9 data points per square kilometer is adequate for hydrologic simulations at the scale of a township. As many as 25 data points per square kilometer are needed in order to make a map of the spatial pattern of soils within a watershed.
3) Once the GIS is established, use an interpreted combination of climate, terrain, and land-cover data files to generate a file of "predicted" soil patterns in unsurveyed parts of the state.
4) Pay special attention to the methods for making maps from a point-inventory system, in order to avoid misinterpretation. Point sampling allows us to estimate the areal extent of a given kind of soil quite accurately, but it does not permit precise description of small areas. It is difficult to emphasize this fact too much -- descriptive errors on some of the soil maps currently used in Minnesota can exceed 60 percent. In the field, a good soil surveyor can "predict" soil traits between sample points, because related information about micro-terrain and land cover are right there to see; in the office, by contrast, it is extremely dangerous to interpolate between sample points. For that reason, the system that produces maps that are derived from sample data should not be permitted to portray data at individual sample points. Instead, it should display only the percentages of larger areas that fall into particular categories. As a rule of thumb, the output map should display data at a resolution that is a full order of magnitude less detailed than the sample data. If more detailed information is needed, field surveys are necessary.

TERRAIN DATA FOR A WATER RESOURCES GIS

Chapter 11 by Corbett and Gersmehl, WRRC, University of Minnesota

ISSUES

As a basis for predicting crop yields, water use, and runoff potential, a resource manager needs good information about the topography in a watershed. To translate existing information into computer-usable form, and to gather future information in an efficient manner, the designers of a water-resources Geographic Information System must deal with three fundamental issues:

1) The mathematical relationship between slope and elevation. The absolute elevation of a point is usually of little significance in hydrologic simulation, but its position with respect to nearby points is very important. One can compute slope from elevation measurements, but the result may be wildly inaccurate, because the terrain between two sample points may be a uniform slope, a series of steps, a concave chute, a high hill, or any one of an infinite variety of slope forms. Our recommendations must therefore deal with the complexity of real-world topography.
2) Tradeoffs between accuracy of terrain description and cost of data storage. Topographic information can be stored in a GIS in several ways: as digitized contours, spot elevations, Delauney triangles, profiles, or slope measurements at sample points. In each case, closely spaced data can reveal more intricacy of the topography, but halving the sample interval will increase costs of data storage by a factor of four or more. Our recommendations must consider the cost of data storage as well as the accuracy needed for the intended uses of the system.
3) Necessity for relational structure in the GIS. In order to provide data for a predictive simulation (such as the Universal Soil Loss Equation or the Peak Flood Model), a GIS must relate all data to a common coordinate system. When that is done, the simulation can get information on slope, soil, and land cover at exactly the same point and thus reach a valid conclusion about erosion or runoff. Our recommendations must accommodate the data that will be used with the terrain file in solving a resource problem.

FINDINGS

The goal is to build a terrain data base that would describe land forms with reasonable precision, facilitate relational use with other kinds of data, and allow easy addition of better data when they become available. To obtain a basis for recommendations, we examined many different data-handling strategies. Our findings include the following:

Contour digitizing and terrain modeling are costly and provide data in a form that existing hydrologic simulations cannot use.

The USGS Digital Elevation Model is well suited for general maps, but erosion predictions based on it have unacceptably high error.

Tabulating slope classes from soil maps is statistically valid, but most hydrologic simulations also require measurements of slope length.

Intersubjective error in interpolating elevations from contour maps is relatively low; slope angle measurements are likewise fairly reliable, but slope length estimates are more error-prone.

Interpolating elevations at 100-meter intervals from 1:24,000 topographic maps can supply adequate terrain data at reasonable cost.

RECOMMENDATIONS

On the basis of these findings and the results of many other investigations at various scales, we make the following recommendations concerning the file structure for terrain data in a water-resources geographic information system:

1) Avoid attempts to derive slope angle and slope length from the USGS Digital Elevation Model; this data file is suitable only for highly generalized maps.
2) Obtain slope data by direct measurement from 1:24,000 topographic maps, either at 100-meter intervals for local profiles, or at a density of 1-4 samples per square kilometer data cell if only regional averages are needed (e.g. for a soil erosion simulation).
3) Record slope classes while entering data from soil survey maps. The data are inexpensive, reasonably accurate, and can serve as a check on the validity of the terrain data file.
4) Use data from several data files to improve the accuracy of terrain data. For example, the presence of marsh vegetation in a land-cover file, or an area of poor drainage in a soil file, may clarify a potentially ambiguous situation in the topography file.
5) Pay special attention to the methods for making maps from a point-inventory system, in order to avoid misinterpretation. Point sampling allows us to say, with some confidence, that "the average slope in area A is B" or "C percent of area D is in slope class E," but we cannot have the same confidence in trying to describe the slope at a point that did not happen to be one of the sample points. It is difficult to emphasize this too much -- the slope halfway between two sample points may indeed be equal to the arithmetic average of the slopes measured at them, but it also may be much greater or less than either measured slope. For that reason, it is dangerous to interpolate between sample points; the system that produces maps derived from sample data should simply not be permitted to portray data at individual sample points. Rather, it should display only the percentages of larger areas that fall into particular categories. As a rule of thumb, the output map should display data at a resolution that is at least a full order of magnitude less detailed than the sample data. If more detailed information is needed, field surveys are necessary.

HYDROGRAPHIC DATA FOR A WATER RESOURCES GIS

Chapter 12 by Brown, Anderson, and Gersmehl, WRRC, University of Minnesota

ISSUES

As a basis for predicting runoff potential, a resource manager needs good information about the streams and artificial drainage channels in a watershed. To translate existing data into computer-usable form, and to gather future information in an efficient manner, the designers of a water-resources Geographic Information System must deal with five fundamental issues:

1) Linear nature of stream channels. Unlike most other hydrologic phenomena, hydrographic data are linear and not areal in nature. On maps or aerial photographs, most streams appear as thin lines. Reservoirs and lakes on rivers are an exception to this linear view -- water may enter and leave at endpoints, but a reservoir must have substantial width in order to serve as an important storage place. Our recommendations must make allowance for the fact that data on stream channels are qualitatively different from terrain, land-cover, and soils data.
2) Hierarchical nature of stream systems. All water that enters stream channels, rivers, and gullies will flow downhill through a system that is hierarchical in design. Most streams have tributaries that add water from upstream. Each in turn serves as a tributary to another body of water. The surrounding terrain is an important factor in the formation of stream networks, and the topology of each system is unique. Our recommendations must make allowances for the difficulty of generalization about drainage systems.
3) Time- and place-bound nature of stream data. Streams have length on maps, but it is the only continuous hydrographic information we have -- other stream data comes from samples at individual points. To make inferences about traits of streams between sampling points, we usually look at factors such as topography, soil, geologic formations, and vegetation cover. The adequacy of these inferences can vary, depending on landscape complexity, frequency of observations, and spacing of sampling points.
4) Temporal changeability of stream networks and channels. Any swale or depression in the terrain (natural or man-made) can serve as channel to move storm runoff to the nearest mapped channel. The parts of a hydrographic network that actually carry water at a given time can vary, depending on rainfall intensity, channel geometry, and the traits of the surrounding land. Moreover, the topology of a drainage system and the dimensions of the channels can change even during a single storm. Our recommendations must make allowances for the variations in channel capacity and flow velocity that accompany these changes in the system.
5) Necessity for relational structure in a GIS. In order to provide data for a predictive simulation (such as the Universal Soil Loss Equation or the Precipitation-Runoff Modelling System), a GIS must relate all data to a common coordinate system. When that is done, the simulation can get information on land cover, soil, and slope at exactly the same place and thus reach a valid conclusion about erosion or runoff. Our recommendations must make allowances for the nature of the data files that will be used with the hydrographic file in solving a resource problem.

FINDINGS

Our goal is to build a hydrographic data base that can describe the drainage system in areas of moderate size with reasonable accuracy, facilitate relational use with other kinds of data, and allow easy addition of better data when they become available. To obtain a basis for a recommendation, we reviewed current data sets collected by various agencies in Minnesota, including information on:

- the state-wide Common Stream and Watershed Numbering System,

- the River Kilometer and Mile Index files,
- the MLMIS40 watershed boundary files,
- the MLMIS40 and Arc/Info digitized stream network and lake files,
- the SWIM Lake Summary File, and
- the U.S. Geological Survey stream discharge and water quality data.

RECOMMENDATIONS

On the basis of our review of these existing data sets, we can make the following recommendations:

1) Maintain a separate vector-encoded file of stream channel data. This file should include data on location, shape, length, and place within the drainage system of the State. Do not summarize the information by larger units of area such as 40-acre parcels, square-mile sections, townships, or counties except for actual analyses. When larger data cells are used, maintain an inventory of the number and characteristics of stream segments at each level of the system hierarchy within the area.
2) Consolidate all stream, lake, and watershed boundary information into a single file with the same spatial resolution and stream identification system. Currently this information is maintained within several incompatible files of differing scales and formats. It therefore requires too much time and processing cost to extract complete hydrographic data for a limited geographic area such as a county or watershed.
3) Begin a program of systematic field measurement of channel dimensions, slope, and roughness at key sample points. These measurements should be geocoded by UTM coordinates and time-stamped, because data of this kind is time- and place-specific and can become obsolete very quickly.

GROUNDWATER SYSTEMS ANALYSIS AND MODELING

Chapter 13 by Pfannkuch, Jones, and Guo, WRRC, University of Minnesota

ISSUES

As a basis for predicting water availability, a resource manager needs good information about the water that occurs below the ground, in surficial and bedrock aquifers. To translate existing information into computer-usable form, and to gather future information in an efficient manner, the designers of a water-resources Geographic Information System must deal with three fundamental issues:

1) Uneven distribution of groundwater aquifers. Groundwater is the most abundant and, in some places, the only reliable source of water in Minnesota. Our recommendations must deal with the geographic variations in the quantity and quality of the groundwater resource.
2) Uneven quality of groundwater data. Evaluating a groundwater resource is more difficult than for water in other parts of the environment, because we cannot see it. We must infer its quantity, quality and flow from observation wells and pumping data at irregularly and often quite widely spaced points. Our recommendations must take into account the difficulty of obtaining groundwater measurements where we might want them.
3) Slow "turnover" of groundwater reservoirs. Flow in most aquifers is very slow, and the resource is therefore subject to long-term damage if misused or contaminated. Our recommendations must recognize the need to develop more creative ways to evaluate and monitor groundwater than we have used in the past.

FINDINGS

The goal is to build a groundwater data base that would describe areas of moderate size with reasonable precision, make regional inventories with reasonable accuracy, facilitate relational use with other data, and allow easy addition of better data when they become available. To obtain a basis for recommendations, we examined many aspects of groundwater. Our findings include the following:

The flow of streams that drain groundwater basins can be used to define the status of the aquifer; this fact seems to offer a useful tool for analyzing the behavior of regional groundwater systems.

In hummocky Minnesota terrain, the groundwater table is a subdued replica of the topographic surface. These local surfaces and the general regional slope drive the flow of groundwater in ways that can produce separate budgets, with different residence times and flow rates for contaminants.

Temperature profile data can be used as an inexpensive tool to define the separation boundary between local and regional groundwater flow systems.

We developed a theoretical model to separate local from regional flow systems in shallow groundwater. Use of this model to define the local and regional groundwater flow boundary may be a good way of assessing contamination and evaluating water-level problems with lakes that are connected to aquifers.

Groundwater simulation models on microcomputers can determine probable well interference problems. This technology would allow access to better data at critical times in the decision- making process, as long as the complexity of the problem does not exceed the capabilities of the program and computer.

To work most effectively, these groundwater simulation models should be linked, via a deep percolation component, with soil water simulation models.

RECOMMENDATIONS

On the basis of these findings and the results of many other investigations at various scales, we make the following recommendations concerning the file structure for groundwater data in a water-resources GIS:

1) Maintain a groundwater information system that is compatible with the states GIS and regularly update it with well and pumping permits and reports (see chapter 14).
2) Use baseflow-analysis techniques developed here as a way to determine the status and behavior of major aquifers. Establish a network of monitoring wells in a series of benchmark basins that have stream gauge records and are representative of different geological and hydrological conditions in the state.
3) Develop a statewide groundwater temperature survey for shallow aquifers as an inexpensive means of delineating local and regional flow systems.
4) Establish criteria for use of groundwater flow simulation to define boundary conditions, complexity, and other characteristics of the system being monitored, so that users can determine if these tools are appropriate for their problem. This should be done with simple rules or menus in the microcomputer programs themselves.
5) Use the groundwater data base to simulate flow and detect possible well interference prior to permitting.
6) Define important groundwater recharge areas through the use of baseflow analysis, regional and local water systems definition, and linked simulation of soil water budgets and groundwater aquifers.

7) **Use the local/regional flow model developed in this project to examine aquifer traits, especially where there are concerns with pollution or the effects of groundwater on lake levels.**

GROUNDWATER DATA FOR A WATER RESOURCES GIS

Chapter 14 by S. Beach, WRRC, University of Minnesota

ISSUES

The usefulness of a Geographic Information System is ultimately linked to the quality of the data it stores and manipulates. Data-collection strategies must be designed carefully to ensure statistical validity and appropriateness for the questions being asked. Recognizing our increasing dependence on groundwater supplies, the designers of a water-resources GIS must deal with three issues:

1) The "hidden" nature of groundwater reservoirs. The kinds of rocks and surficial materials that can store and transmit water are not distributed uniformly throughout the state, and their presence is not always obvious to an observer on the surface. Much of the information that we have came from logs of wells drilled for private use or other purposes. Our recommendations must make allowances for inadequacies in background data.
2) The long-term implications of pollution. Groundwater reservoirs usually have long detention times and slow turnover rates. Our recommendations must make allowances for the fact that a spill or other source of contamination may take a long time to affect an aquifer. Usually, cleanup takes even longer.
3) The possibility of unanticipated pollution in unforeseen places. The timetable for detecting and monitoring contamination is extremely short in the case of events such as vehicle accidents, tank car spills, or pipeline breaks. These events often occur in places where available groundwater data are inadequate. Our recommendations must accommodate the need for quick retrieval of data in emergencies.

FINDINGS

The goal is to build a groundwater data base that can store aquifer traits, flow characteristics, and water quality. This data base should facilitate relational use with other data bases, allow easy addition of better data when they become available, and permit rapid retrieval of data in case of emergency. To obtain a basis for recommendations, we evaluated groundwater monitoring strategies and methods of simulating groundwater movement. Our findings include the following:

Public consciousness of the potential threat to groundwater supplies is quite high; support for proper monitoring is therefore strong.

A spatial-statistical monitoring network offers an efficient way of gathering background data to evaluate contamination hazards.

Existing record-keeping systems are being upgraded as PCA develops its Integrated Ground-water Information System (IGWIS) and the DNR/MGS expands its Ground Water Data Base (GWDB).

A USGS ground-water simulation, the McDonald/Harbaugh Three-Dimensional Finite Difference Ground Water Flow Model, may be useful for "predicting" regional trends in groundwater quality.

RECOMMENDATIONS

1) Explore ways of combining data from GWDB and IGWIS with other GIS information (e.g. land cover, terrain, soils, bedrock geology) in order to detect patterns and clarify ambiguities in the files.

2) Use statistically valid methods to investigate regional patterns of background water quality. This study must consider trends and seasonal cycles as well as spatial traits of aquifers.

3) Target intensive patterns of monitoring wells around underground storage tanks, landfills, or other potential sources of aquifer contamination. A few upgradient observation wells can establish background water quality. More wells are needed down gradient to monitor water levels, determine flow direction and velocity, and detect lateral migration of possible contaminants.

4) Formulate contingency plans and build a regional data base to improve reaction time in cases of unanticipated contamination. A survey phase consists of three wells drilled in a triangular pattern around the site, to gather hydrogeologic information, to trace the movement of contaminants, and to aid in designing the monitoring network. The monitoring phase includes testing water from private wells to determine local background water quality and the areal extent of the contaminant plume. Information gaps are then filled by drilling more (perhaps 15-30) boreholes, each one systematically placed on the basis of information from two or three prior boreholes.

5) Pay special attention to the methods used for making maps from a point-sampling system, in order to avoid misinterpretation. Point sampling allows us to make a reasonably accurate description of water quality and the dimensions of groundwater aquifers at a particular place, but it does not permit precise description of areas between wells. It is difficult to emphasize this fact too much -- interpolating between sample points can be extremely dangerous. For that reason, the system that produces maps from sample data should not be permitted to use the same symbols for the areas between individual sample points. As a rule of thumb, the output map should display data at a resolution that is a full order of magnitude less detailed than the sample data. If more detailed information is needed, field surveys are necessary.

CONCLUSIONS AND SUMMARY OF RECOMMENDATIONS FOR A WATER RESOURCES GEOGRAPHIC INFORMATION SYSTEM.

Chapter 15 by P. Gersmehl and D. Brown, WRRC, University of Minnesota.

ISSUES

Planners and managers of water resources in the State of Minnesota have expressed a need for access to data, which could be stored and manipulated in a water resources geographic information system (GIS). These data are of two two qualitatively different kinds:

The first set of files consists of records of the location, quality, and quantity of water resources - streams, lakes, aquifers, wetlands, etc. In general, the existence of these data in on-line GIS is spotty, although some are in the process of being compiled and digitized by various agencies.

The second set of files should contain data on those aspects of the surrounding environment that can affect the quantity or quality of water. This GIS could provide a baseline against which we measure the hydrologic effects of environmental changes, both natural or human-induced. It could also serve as a source of input data for hydrologic simulations of the potential impacts of environmental changes on water resources.

Our research project was to examine the data requirements of hydrologic simulations that appear to be of value in analyzing water resources in Minnesota; on the basis of that examination, we made a series of recommendations concerning the file structure, coordinate system, cell size, data precision, analytical methods, and output procedures for a water resources GIS that could provide environmental information needed by planners and managers. Some of these recommendations could be carried out with existing primary data sources. Others may require development of interim techniques to make existing data more useful.

RECOMMENDATIONS

In this concluding chapter, we will summarize the most important of those recommendations:

1) A water-resources GIS for Minnesota should store data in a two-dimensional grid registered to the Universal Transverse Mercator coordinate system. This method of geographic location provides a uniform size for data cells and an unambiguous referencing language. An alternative coordinate grid based on the Public Land Survey would make data files easier to relate to ownership patterns. However, this apparent "advantage" has some serious legal implications, especially when a sampling method is used to maximize the statistical validity of the files. Moreover, the sections in the Public Land Survey are of uneven size and shape, and the alignment of many land uses and associated features with survey lines can introduce some very serious statistical abberrations in water-resource data files.
2) A water-resources GIS for Minnesota should facilitate input of data obtained by point sampling at specific locations within the one-square-kilometer data cells. Although a point-sampling method does sacrifice some precision in locational display, this apparent drawback is more offset by the gain in statistical accuracy in providing inventory data for hydrologic simulations. Providing different kinds of data -- soils, terrain, land cover, drainage, etc. -- at exactly the same location is essential if the GIS is to maintain relational accuracy when it combines separate data files in order to answer specific questions.
3) A water-resources GIS for Minnesota should include a set of algorithms for translating climatic information into data values for specific GIS cells. The

main source of climatic information is a network of National Weather Service recording stations, which are too widely spaced for short-term hydrologic studies. Indeed, their spacing is barely adequate for describing annual totals. Interpolation between those stations is risky at best; statistical accuracy is maximized if one uses a kriging method to derive data that are optimized for the areal extent and temporal frame of each specific study.

4) A water-resources GIS for Minnesota should clearly differentiate the concepts of *land cover* and *land use*. A "land cover type" is a physical concept, observable on (and mappable with) aerial photographs and satellite imagery. A "land use type" is an economic category, which can (and usually does) have a very wide range of physical traits. One square mile of "multi-family residential land use" can produce a flood that is twenty times as big as that from another section with the same land-use category and a different physical arrangement of impervious areas and sewers. For this reason, a map or computer file of land *cover* is a preferable source of data for hydrologic simulation.

5) A water-resources GIS for Minnesota should be able to use current LANDSAT imagery to provide a picture of the broad patterns of land cover for a study area. Spectral classification with a microcomputer can achieve accuracies of 60 percent or better at relatively low cost, especially when compared with obtaining and interpreting large-scale aerial photographs. Both data sources have the drawback of being unable to perceive details of surface condition, drainage system connectivity, and channel dimensions, which can cause order-of-magnitude variations in the hydrologic response of a watershed. For this reason, we recommend placing a high priority on rigorous point sampling within defined landcover areas, and a much lower priority on improving the accuracy of placement of boundaries on land-cover maps.

6) A water-resources GIS for Minnesota could provide a method of augmenting the existing file of soil data in the Planning Information Center. This data file is based on a *tagged-area* map, which tried to minimize error in describing individual parcels of land. A data file with a tag perspective will systematically underestimate the areal extent of hydrologically important soils, often by a factor of five or more. A single relational record of soil in each square kilometer data cell would be a significant improvement; eight point samples per data cell would achieve a statistical accuracy surpassing that of virtually every other kind of data in the GIS.

7) A water-resources GIS for Minnesota should flag certain files for restricted use. For example, the USGS Digital Elevation Model (DEM) should not be used for any purpose except to provide a general picture of the terrain in an area of county size or greater. This file does not "capture" the complexity of typical Minnesota topography, nor can it do a good job of defining drainage basin boundaries for small study areas. Calculating slope on the basis of this DEM can lead to estimates of soil erosion that range from one tenth to twice as much as is actually likely to occur. For many hydrologic simulations at a regional scale, fully relational measurements of slope angle and length at only one site per square kilometer would be preferable to a hundred elevation measurements in the same area.

8) A water-resources GIS for Minnesota should include two separate water files: a vector-encoded file of natural and artificial drainage systems and a point-data file of all groundwater monitoring wells. These two files should be registered to the UTM coordinate system, so that their information can be related to the data in the rest of the geographic information system. The square-kilometer inventory files in the GIS should include data on the presence, direction, and departing elevation of major streams, along with

field measurements of channel dimensions, which are critically important for flood-routing algorithms.

9) A water-resources GIS for Minnesota should provide tools for using the contents of related files to improve the quality of data files in the GIS. For example, the presence of low spots in the terrain file and poorly drained soils in the soil file can help clarify some ambiguities in the land-cover file. Moreover, when strict relationality is maintained, it is possible to extrapolate from one area and create "synthetic" data files for adjacent areas that may lack certain data sources, such as detailed soil surveys or satellite imagery for a particular day. These extrapolations (and indeed all files derived by merging primary data files) should have a strict "sunset" date, beyond which they cannot be used without recompiling with up-to-date data.

10) A water-resources GIS for Minnesota should not produce maps directly from point-sample data. Statistical sampling at discrete points allows us to relate data files to each other and to estimate the areal extent of features quite accurately, but it is dangerous to extrapolate from sample points for data that are not spatially continuous. For that reason, the system to produce maps derived from point-sample data should not be able to portray data at individual sample points. Instead, it should display only the proportions of larger areas that belong to certain categories. As a rule of thumb, the output map should display at a resolution that is a full order of magnitude less detailed than the sample data. For more detailed informatin, field surveys are necessary.

11) A water-resources GIS for Minnesota should require creation and attachment of a detailed "pedigree" to each file entered into or retrieved from the GIS. The output program should be able to sense the measurement precision, spatial resolution, and temporal attributes of each input file, and it should adjust the output specifications so that the printed or displayed results do not imply any more precision than can be justified statistically. This is extremely important, in view of the recent trend to extend the concepts of misrepresentation and produce liability to apply also to the "informatin sector" of the economy.

12) A water-resources GIS for Minnesota should have an established mechanism for the systematic evaluation of data files in order to set priorities for data upgrading. In most cases, the automatic monitoring procedure that was recommended in point 11 will cause the output precision to be limited to that of the least precise input variable. A method of identifying and recording this variable after each use of the system will enable the users of the GIS to target variables for further investigation.

These twelve principles would provide the basis for a water-resources GIS that would be scientifically sound, economically feasible, and legally defensible. The recommendations are specifically for water resources; they are not intended as blanket suggestions for other GIS applications. Implicit in the framework is a recognition of three conceptual levels of geographic variation: in the hydrologic characteristics of the place we are examining, in the quality of information available for that place, and in the quality of information needed there. The first level says that a "tested principle" in one place may not always work in another, and therefore we need access to good data about different places. The second level is an admission that our understanding of hydrologic systems (or our arsenal of available data) may not be adequate to allow us to use the same hydrologic methods in all places. And the third level implies that we should not wastefully do more than is really necessary to solve a problem in a given place. One data point per sqaure kilometer may be adequate to provide the information needed for a particular application in one area, where it is not hard to insure against the consequences of failure. By contrast, a more dense sampling network may be desirable for someone trying to solve the same kind of land puts more "at risk" and makes the cost of failure higher. In the last analysis, that is the justification for a water-resources GIS -- to provide a sound basis of information for people making decisions that could have an adverse impact on the way other Minnesotans use our resources. For a preventive medicine to work, it must be both palatable and effective.

SUMMARY OF DEMONSTRATION AND MISCELLANEOUS REPORT FINDINGS

CROP PRODUCTION RESPONSE TO MOISTURE SUPPLY IN MINNESOTA

SPECIAL REPORT 9.

WATER RESOURCES RESEARCH CENTER

D. Brown, P. Gersmehl, J. Drake and K. Scaggs

CROP PRODUCTION RESPONSE TO MOISTURE SUPPLY IN MINNESOTA

ABSTRACT

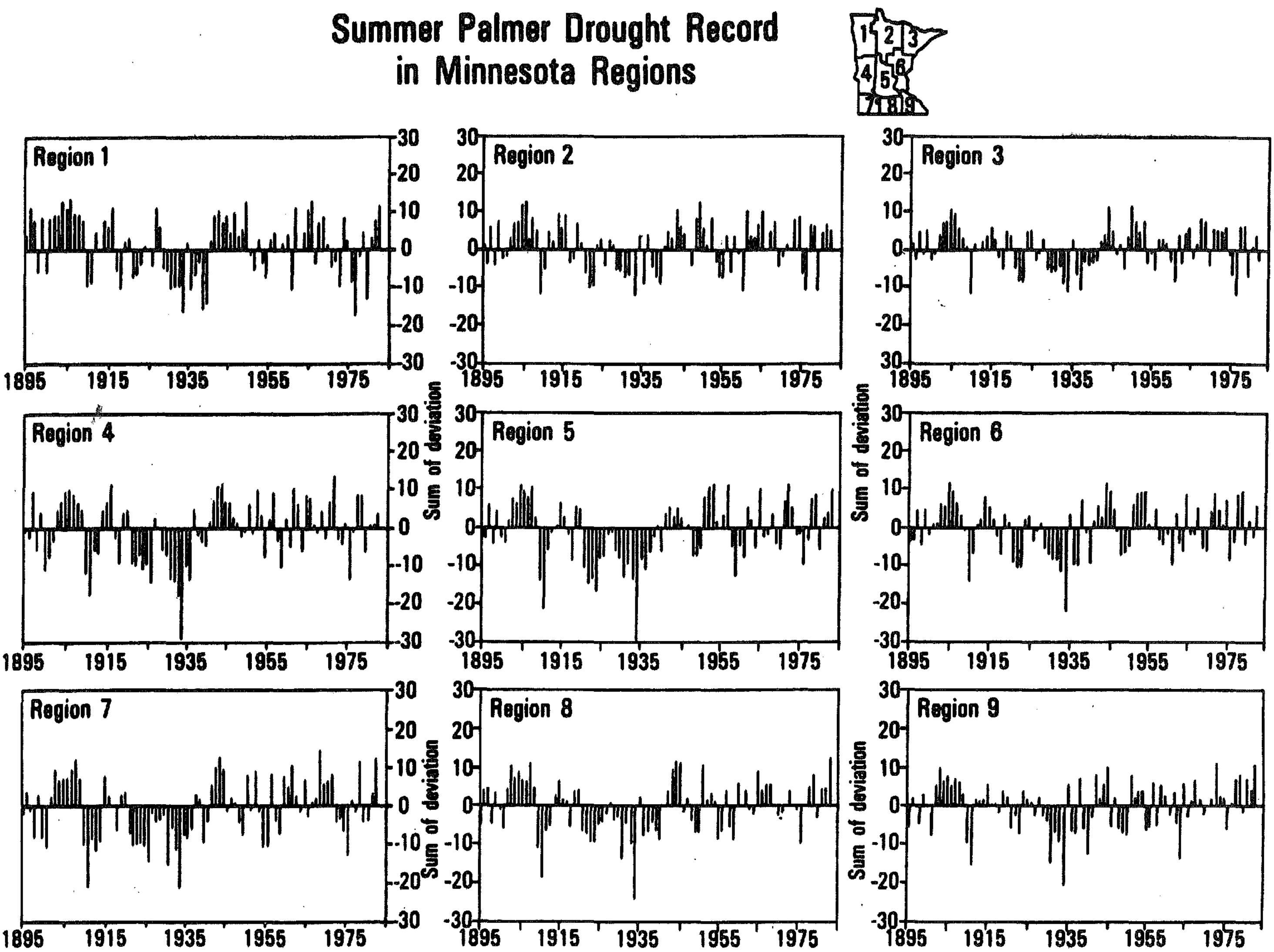
This report defines how variations in moisture affect state-wide crop production. To accomplish this goal we controlled geographic variations in the response of crops to changing technology, and regional differences in the severity and timing of wet and dry periods. The regional differences in the moisture supply and crop response are treated by examining them as deviations from local norms or expected values. Technology changes are controlled by subtracting the general trend in yield from the actual yield history.

Yields of major grains are used here with the summer Palmer Drought Index record to examine the role of atmospheric moisture supply on state-wide grain production. The northeast contributes so little grain production that we excluded this region from total state production figures.

For oats, wet years are more damaging to total production than drought in the east and both wet and dry years result in slightly reduced production in the southwest. Corn production is slightly poorer during wet years in the east and north, and slightly reduced by both extremely wet and extremely dry years in most other areas of the state. Only in west central Minnesota does soybean production show a weak tendency to be limited by drought events. In other regions the soybean response to moisture is very poorly defined or not statistically significant.

Several factors limit the interpretations of these results. These include: 1) the statistical significance of the relationship between the Palmer Drought Index and the yields, which is not always present in Minnesota; 2) the assumptions of resource homogeneity within the regions, which is a known fact; 3) the fact that these results are not intended for use as a forecast tool, because they are limited to the time period and place of the derivation.

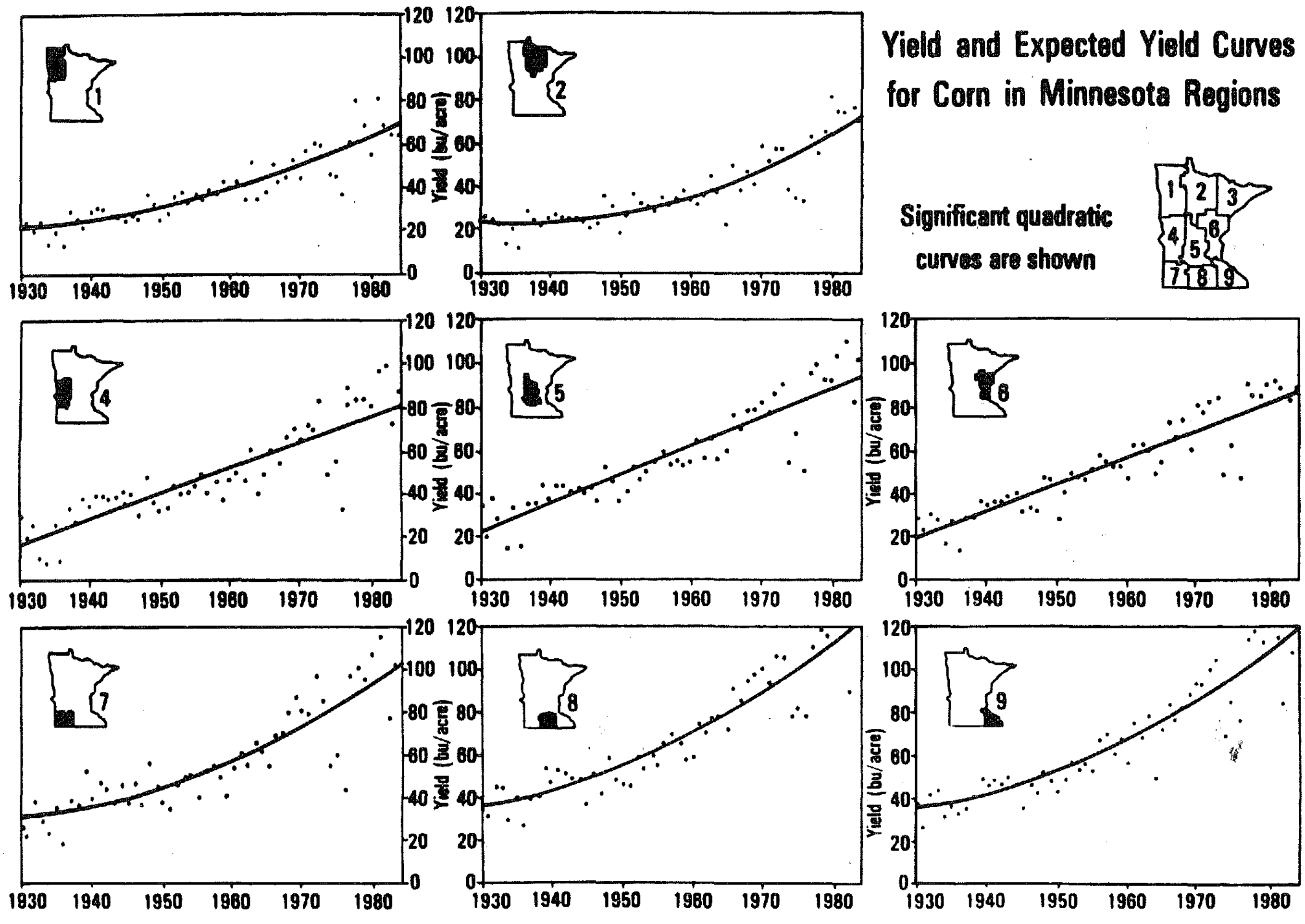
Pronounced relationships between moisture supply and yield are lacking for a number of crops and regions, with factors other than moisture emerging as prominent in the total state grain production. The large area averaging effect and the variability caused by other factors combine to be so great that drought responses do not appear to overwhelm state aggregate production much more than one year in ten. One should, however, expect local production to be much more responsive to drought than the entire state which is influenced by large-area averaging.



Cartography Laboratory, Department of Geography, The University of Minnesota.

Figure 1: Values represent the sum of June, July, and August PDI values for the years 1895 to 1983. Negative values represent dryer than normal years and positive values represent years with above normal moisture conditions. The patterns of historic moisture variability are not uniform among the 9 Minnesota climatic subdivisions.

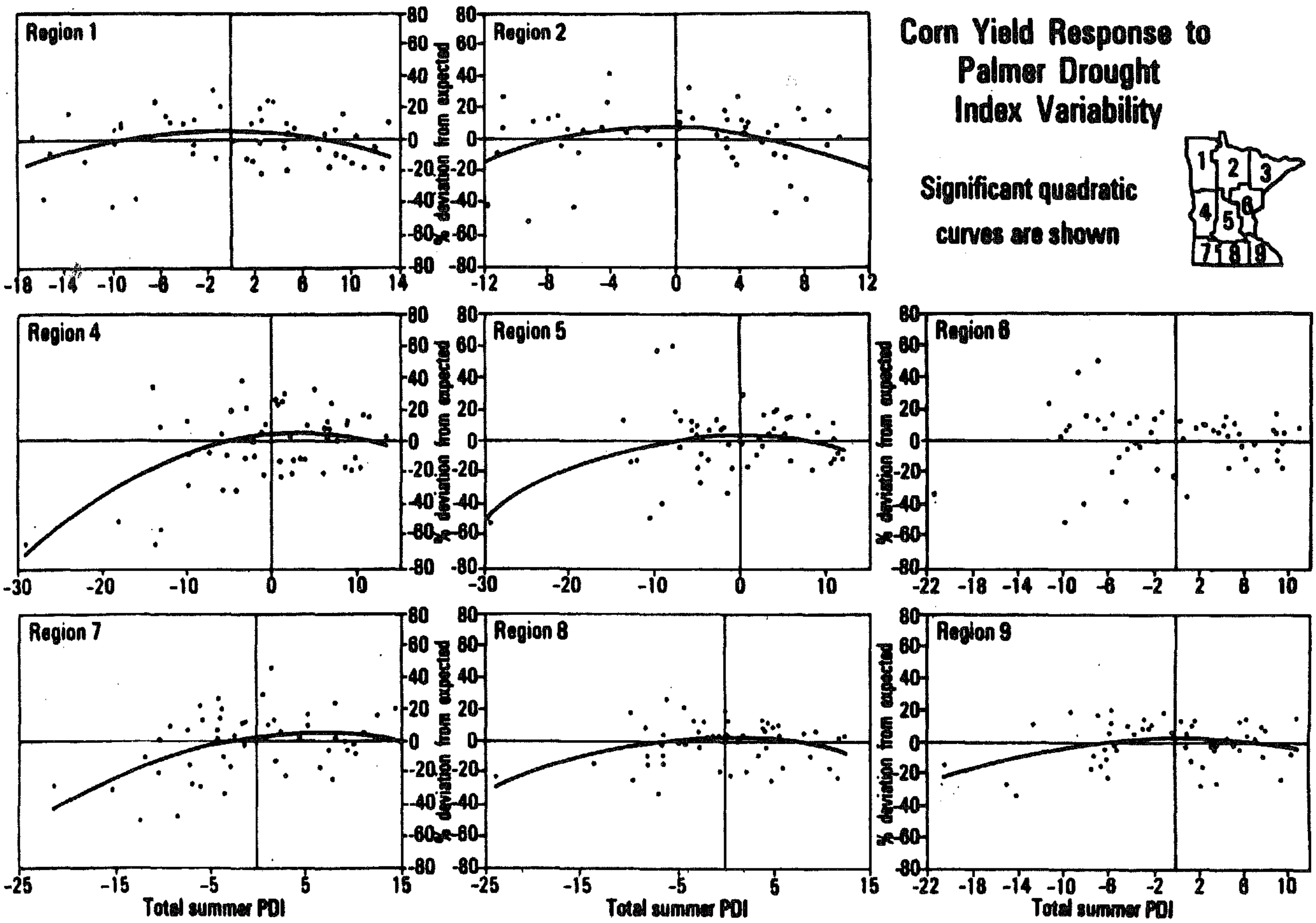
Data from the National Weather Service.



Cartography Laboratory, Department of Geography, The University of Minnesota.

Figure 2: Historic trend of corn yield records for 8 Minnesota Climatic subdivisions.

Data from the Minnesota Department of Agriculture.



Cartography Laboratory, Department of Geography, The University of Minnesota.

Figure 3: Corn yield response to summer Palmer Drought Index values for 8 Minnesota climatic subdivisions.

Data from the Minnesota Department of Agriculture and the National Weather Service.

Table 1: Estimated impact of moisture variability on state-wide production of major Minnesota grains for selected moisture exceedance years. The response of corn (C) is fairly strong. Oat (O), and soybean (S) production are more strongly affected by other factors than by their response to summer moisture differences. Production figures are based on Minnesota acreage planted to each crop in 1982 using the technology of 1982. Bushel figures are in millions.

Year	Percent Time Exceeded	Normal Expected (Bushels)			Difference Predicted (Bushels)			Difference Predicted (Percent)		
1978	24	645	94	135	119	-30	17	18	-32	13
1982	30	645	94	135	0	-2	0	0	-2	0
1963	50	645	94	135	67	-11	-23	10	-12	-17
1959	74	645	94	135	-61	-23	-77	-9	15	18
1977	81	645	94	135	104	14	24	16	15	18
1976	93	645	94	135	-209	-56	-178	-32	-60	-131

SUMMARY OF DEMONSTRATION AND MISCELLANEOUS REPORT FINDINGS

RELATIONSHIP BETWEEN CLIMATE AND THE MEAN ANNUAL FLOW OF THE MISSISSIPPI RIVER AT SAINT PAUL, MINNESOTA

SPECIAL REPORT 11.

WATER RESOURCES RESEARCH CENTER

R. Skaggs, and D. Brown

RELATIONSHIP BETWEEN CLIMATE AND THE MEAN ANNUAL FLOW OF THE MISSISSIPPI RIVER AT SAINT PAUL, MINNESOTA

ABSTRACT

This study examines the statistical relationship between the mean annual flow of the Mississippi River at St. Paul and the water balance surpluses (water not used in evapotranspiration and soil moisture storage) for the six climatological divisions of Minnesota that contribute to the drainage of the Mississippi and Minnesota Rivers in Minnesota. The water surpluses are calculated by the Thornthwaite method of estimating the water balance.

About 70 percent of the mean annual flow of the Mississippi River at St. Paul is statistically explained by variations in the surpluses computed for these two divisions.

Although the statistical model was developed on the first half of an eighty three year record of mean annual flows (1900-1982) it was quite capable of predicting mean annual flows of the second half of the record. From these results we conclude that it possible to statistical predict large scale water resources in Minnesota if the data are collected very rapidly (near real time) and used in a water budget calculation (such as the Thornthwaite method) to determine the water surpluses.

From the excellent performance of the statistical equations in the second half of the record we conclude that a substantial component of the increase in mean annual flow of the Mississippi in the past 40 years results from greater precipitation since 1940 in both the eastern and western portions of Minnesota and not lower temperatures.

We also examined the possible effects of a temperature increase as a result of the "Greenhouse Effect" (increasing carbon dioxide in the atmosphere). We used a postulated 3 degree Celsius temperature rise and recomputed the water surpluses for the west central and east central climatological divisions. The results suggest a major decrease in the amount of surplus in both divisions if the temperature rise occurs and a corresponding reduction in the mean flow of the Mississippi River at St. Paul.

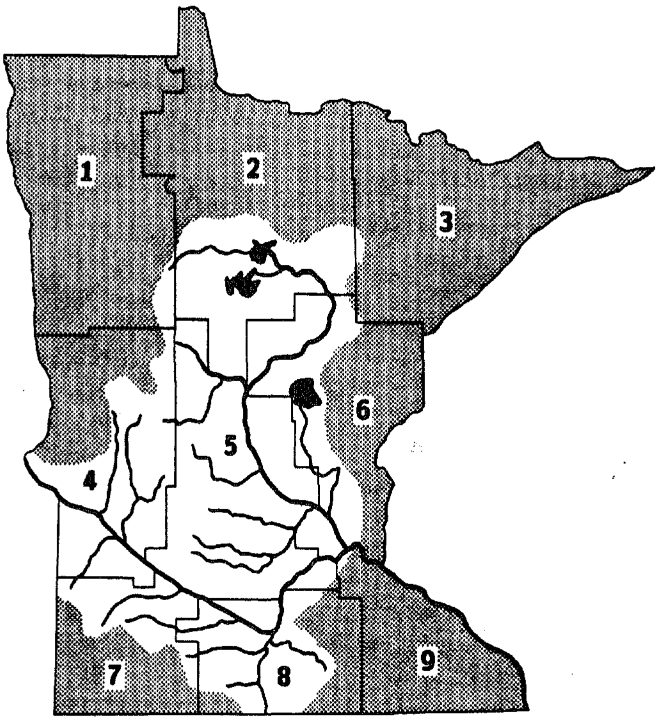


Figure 1: Climatic subdivisions of Minnesota and the Mississippi River Basin above Saint Paul, Minnesota.

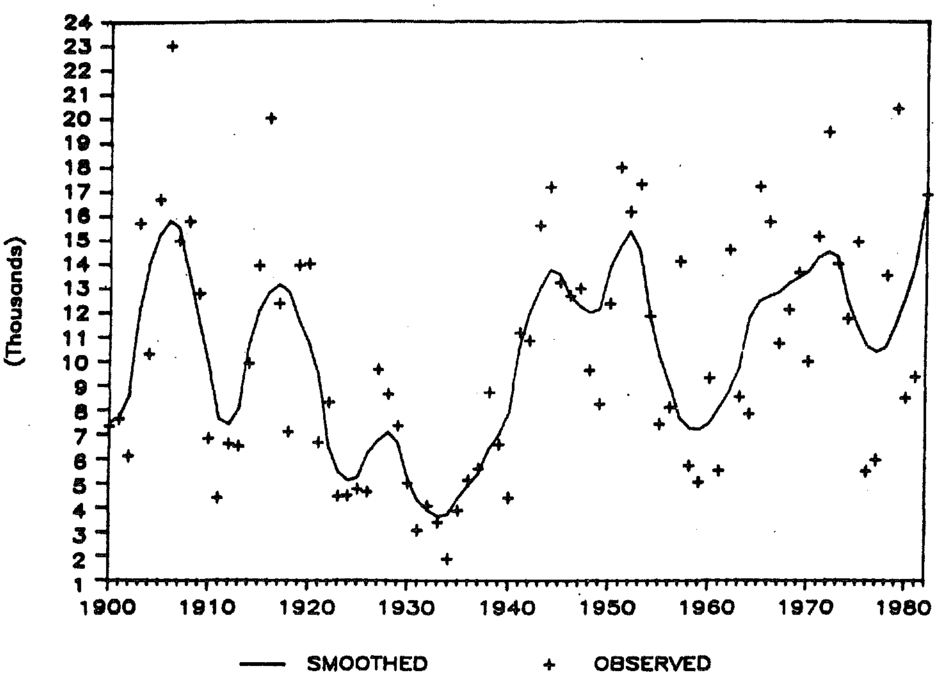


Figure 2: Mean annual flow record of the Mississippi River at Saint Paul, Minnesota.

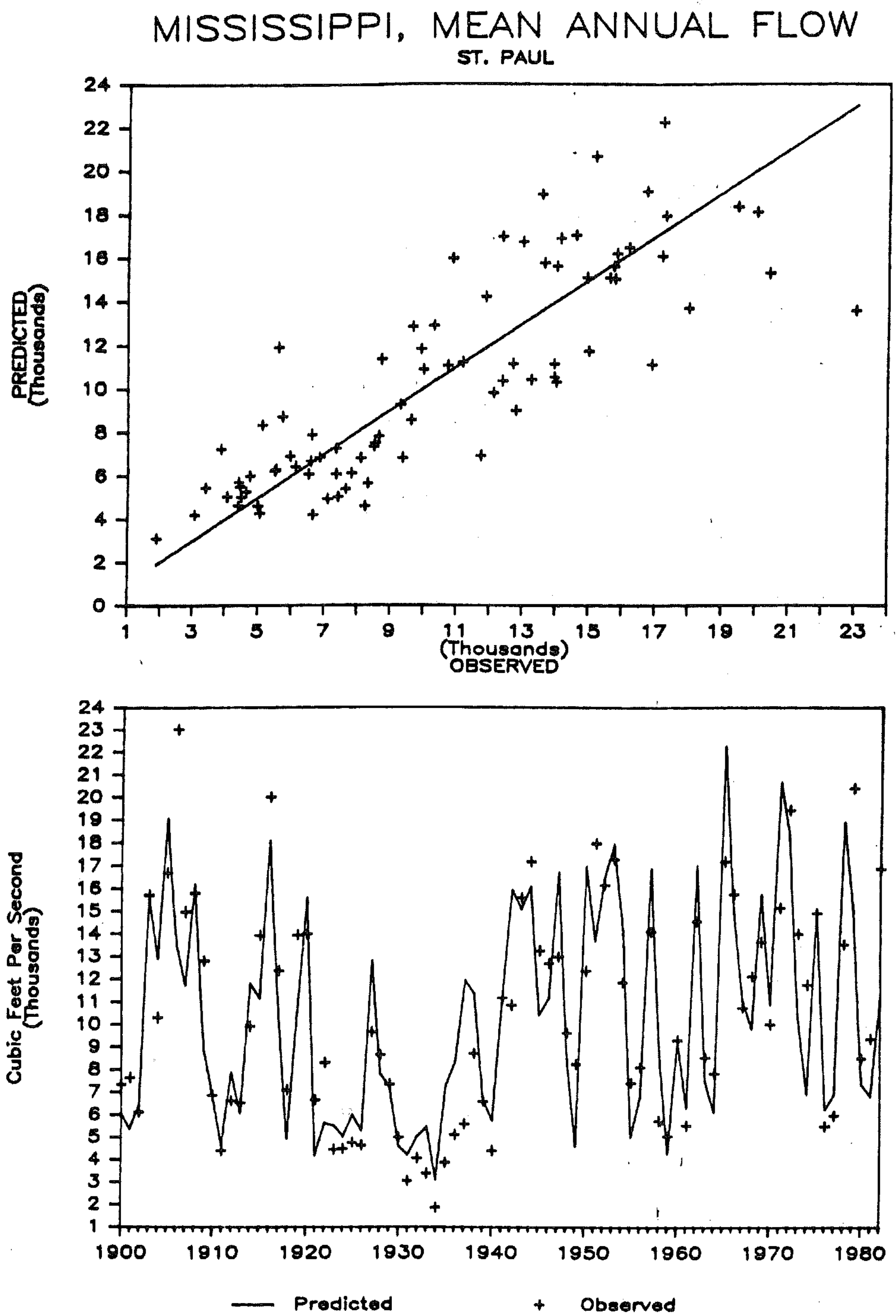


Figure 3: Predicted and observed plots of mean annual flow of the Mississippi River at Saint Paul, Minnesota. The upper graph is a plot of the predicted values versus the observed. The lower graph is a plot of both the observed and predicted flows versus time.

Table 1: Effect of temperature change on computed water balance surplus.

Year	Surplus* with Observed Temperature		Surplus* with Increased Temperature		Percent Change**	
	Div 4	Div 6	Div 4	Div 6	Div 4	Div 6
1971	178	343	41	167	-67	-51
1972	173	251	136	187	-21	-25
1973	50	198	0	64	--	-69
1974	20	120	0	69	--	-42
1975	103	280	53	172	-49	-39
1976	11	112	0	92	--	-18
1977	31	92	20	52	-35	-43
1978	181	257	89	118	-51	-54
1979	122	242	56	170	-54	-30
1979	122	242	56	170	-54	-30
1980	42	84	19	48	-55	-43
1981	0	166	0	17	--	-90
1982	74	174	18	146	-76	-18
1983	85	317	37	160	-56	-50
Average	82	203	43	112	-56	-45

*Values in millimeters per year.

**Observed Temp. Surplus/Increased Temp. Surplus*100

SUMMARY OF DEMONSTRATION AND MISCELLANEOUS REPORT FINDINGS

TWIN CITIES SURFACE WATER SIMULATION MODELING DEMONSTRATION.

SPECIAL REPORT 12.

WATER RESOURCES RESEARCH CENTER

K. Anderson, J. Corbett, N. Dunning, C. Gersmehl, R. Greene, P. Gersmehl, and D. Brown

TWIN CITIES SURFACE WATER SIMULATION MODELING DEMONSTRATION.

ABSTRACT

This report looks at how a water-resources geographic information system (GIS) and computer program can simulate runoff from storm events. We introduce a land-cover classification designed to improved on the way current land-use classifications deal with impervious (paved) surfaces. Our proposed classification also accounts for seasonal changes in vegetative growth and evapotranspiration. The file structure of our GIS uses point observations tabulated within square-kilometer areas. This inventory-based file structure is specifically designed to relate soils and land cover measurements taken from identical places, as well as to identify the range of phenomena within larger areas.

Our study areas consist of two medium-sized watersheds located within the seven-county Twin Cities Metropolitan Area. The Elm Creek basin in northern Hennepin County drains 220 km2 of urban and suburban developments mixed with agricultural land uses and hobby farms. The Vermillion River basin of central Dakota and eastern Scott Counties (285 km2) consists primarily of agricultural land, with a number of free-standing communities around which new urban growth is taking place. Both watersheds include a number of land cover types on a diverse mixture of land forms; both also provide an adequate record of stream discharge out of the basin to test simulation results.

The GIS used in this study is the cell-based EPPL7 system that runs on 16-bit personal computers (such as IBM PCs and their compatibles). EPPL7 is a product of the Planning Information Center (PIC), State Planning Agency. The point-counting inventory consists of a series of files, one file for each category in the GIS (e.g. "medium-density residential" in the land-cover data), with a frequency of occurrence recorded for each cell. We derive land-cover information from Landsat MSS imagery rectified to 100 meter areas on a UTM coordinate base. Soils data are point samples from published county soil surveys. The Appendix shows a method of implementing a point-counting inventories in EPPL7.

Runoff is calculated using the well-known Soil Conservation Service "Curve Number" approach. We produced a series of maps of runoff estimates based on current land-cover from two separate Landsat images for each basin. Estimates are also calculated for an "urbanized" Elm Creek basin to test the applicability of the simulation approach to future growth scenarios.

Three major conclusions can be drawn from this study: (1) useful hydrologic analysis can be performed with a simple GIS and simulation models; (2) the point-counting method is the best approach for coding environmental data; and (3) point-relational data sets are a must for accurate simulation results. Square-kilometer cells are more than adequate for analysis of surface runoff in the Elm Creek and Vermillion River basins. If necessary, results can actually be improved more by better measurements at specific points in an area than by using data at a finer spatial resolution in the GIS.

Table 1: Categories used with LANDSAT classification of Twin Cities land cover

Level I	Level II	Category
1		Hard Surfaces*
2		Pervious Earth Materials
	21	Gravel Pits; extractive
3		Surface Water
	30	Surface Water (Undifferentiated)
4		Persistent Vegetation
	41	Forested; woods; trees
	42	Grassland (Pasture, open-space recreation)
5		Wetlands
	50	Wetlands (Undifferentiated)
6		Temporarily Vegetated Areas
	61	Cover Crop (Hay, Alfalfa)
	62	Small Grains
	63	Row Crops (Corn, Beans)
7		Developed Areas
	71	Commercial; Industrial (large structures, +85% impervious)
	72	High Density Residential (HDR) (<1/8 acre, small to medium size structures, roughly 65% impervious)
	73	Medium Density Residential (MDR) (1/8 to 1/2 acre, small structures, roughly 35% impervious)
	74	Low Density Residential (LDR) (1/2 to 1 acre, small structures, roughly 25% impervious)
	75	Very Low Density Residential (VDR) (>1 acre, small structures, roughly 20% impervious)

* "Hard Surfaces" as a category was not distinguishable from commercial, industrial, or other high density land covers using LANDSAT at 100 meter resolution; in most cases, it was lumped in with Category 71, "Commercial; Industrial".

Table 2: Percentages of sample observations with various combinations of land cover and soil hydrologic group: Elm Creek Basin on the June 2, 1986, classified LANDSAT image*

	Land Cover Category	Soil Hydrologic Group A	B	C	D
21	Extractive (gravel pits, dirt roads)	<1	<1	0	<1
30	Surface Water **	0	<1	<1	<1
41	Forested	0	2	<1	<1
42	Grasslands	0	1	<1	<1
50	Wetlands	<1	4	1	4
61	Cover Crop	<1	13	4	8
63	Row Crop	0	12	3	6
71	Commercial/Industrial	0	1	<1	<1
72	High Density Residential	0	2	1	2
73	Medium Density Residential	<1	3	1	2
74	Low Density Residential	0	4	1	3
75	Very Low Density Residential	<1	6	1	2

* Out of 3712 sample points, 306 (8 percent) were unusable due to uncertain or undefined soil hydrologic group classification.
** Not included in this table are an additional 61 soil records (2 percent) that were recorded as "water" in the soil file, but classified as "non-water" on the land cover file. These areas were assumed to be correctly identified as "water" by the soil survey.

Table 3: Curve numbers used with the LANDSAT land cover data.

	Land Cover Category	Soil Hydrologic Group A	B	C	D
21	Extractive (gravel pits, dirt roads)	72	82	87	89
30	Surface Water *	100	100	100	100
41	Forested	25	55	70	77
42	Grasslands	39	61	74	80
50	Wetlands	85	85	85	85
61	Cover Crop	58	72	81	85
63	Row Crop	67	78	85	89
71	Commercial/Industrial	89	92	94	95
72	High Density Residential	77	85	90	92
73	Medium Density Residential	61	75	83	87
74	Low Density Residential	54	70	80	85
75	Very Low Density Residential	51	68	79	84

* Curve numbers for surface water present a big theoretical problem. A curve number of 100 (what SCS assigns to this land cover) implies "maximum runoff: -- all rainfall is assumed to be automatically available for runoff. As long as a lake doesn't dry up, this is probably true, especially for those water bodies with outlets. In Minnesota, however, lakes are also associated with hydrologic "sinks" -- water is held in the depression until it infiltrated into the groundwater system. A zero curve-number may be more appropriate. Here is one area where the relational power of a GIS might come in handy -- a network file will tell us how connected a water body is to the rest of the surface hydrologic system.

(Source: After Soil Conservation Service 1972 and 1986; Young *et al.* 1985

SUMMARY OF DEMONSTRATION AND MISCELLANEOUS REPORT FINDINGS

BEAR CREEK SURFACE WATER SIMULATION MODELING DEMONSTRATION

SPECIAL REPORT 13.

WATER RESOURCES RESEARCH CENTER

K. Anderson

BEAR CREEK SURFACE WATER SIMULATION MODELING DEMONSTRATION

ABSTRACT

The role people play in changing the character of the Earth's surface has a profound impact on water resources. Reasonable predictions of the results of human actions would be of enormous benefit to planners. Computer simulations can be used to provide these predictions. As with any analysis technique, these simulations require specific input data and make assumptions about reality that limit their application to watersheds of specific sizes and geographic locations.

This report summarizes a test of the interface between existing computer models and water-resource data in Minnesota's current geographic information system. The study area for this demonstration is the Bear Creek basin of Olmsted County, eighty square miles of rolling farmland near Rochester. We used two "off-the-shelf" computer programs, AGNPS and USDAHL, to estimate runoff for the Bear Creek basin. AGNPS is a cell-based, distributed model that uses the Soil Conservation Service's "Curve Number" method with the Universal Soil Loss Equation to estimate surface runoff and soil erosion for storm events. USDAHL is a non-cellular, fitted model that is calibrated to a basin to estimate runoff over a continuous period of time.

Results from the AGNPS and USDAHL models lead us to several conclusions:

- our current GIS lacks good data on antecedent basin conditions, particularly soil moisture, which is necessary information for storm-event models such as AGNPS;
- a point-count (inventory) approach to data collection is necessary to identify the range of basin characteristics. Describing each area in terms of a single soil, land-cover type, or slope (the current practice in the state's GIS) produces poor results. For example, soils or land covers that occupy only a small fraction of a data cell can still produce the majority of the runoff; and
- the cellular approach to watershed subdivision captures the diversity of hydrologic responses and lends itself to use with a GIS better than a polygon approach.

At present, the state has no model that incorporates all of these features at a scale appropriate to analysis of medium-to-large-sized watersheds. It is this very scale in Minnesota that can benefit from policy decisions and planning based on the use of simulation with a GIS.

SUMMARY OF
DEMONSTRATION
AND
MISCELLANEOUS REPORT FINDINGS

MODELING SOIL WATER VARIABILITY

SPECIAL REPORT 14.

WATER RESOURCES RESEARCH CENTER

I. Lopez Bakovic

MODELING SOIL WATER VARIABILITY.

ABSTRACT

This report demonstrates the process of simulating the temporal and spatial variability of soil water. We use a highly instrumented catchment in Texas for this demonstration. We use a one-dimensional water budget model, based on equations from the SWRRB model to demonstrate time variability of soil moisture at a point. We also demonstrate methods of interpolating the spatial patterns of soil water with various semi-variogram and kriging techniques. These methods help us analyze the spatial structure of point measurements and predictions (such as those by SWBM) of soil water status. In combination, these methods are demonstrated to be useful tools for examining the soil water under various environmental conditions and therefore could be used to study the effects of land-cover and land-use changes on water resources.

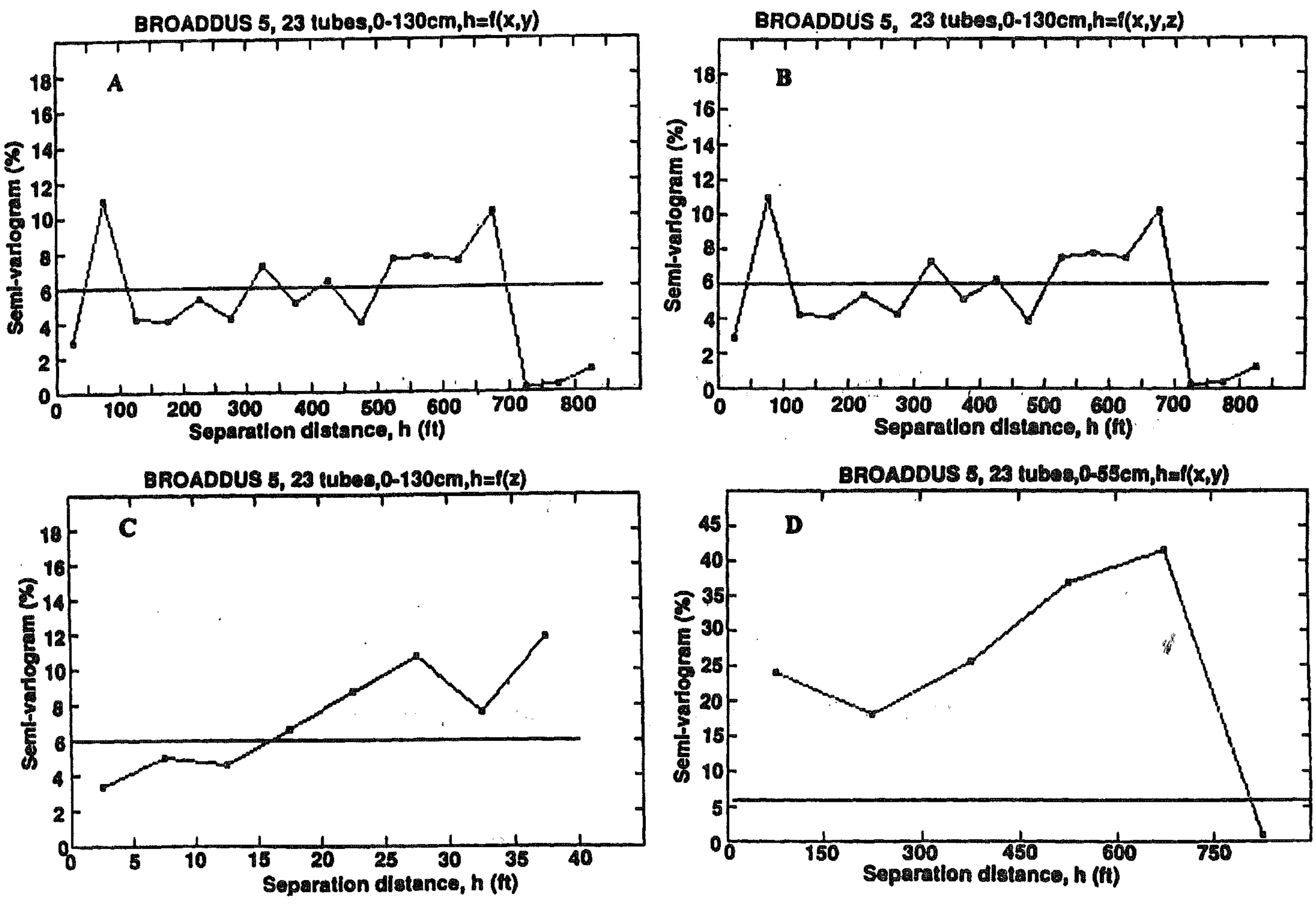


Figure 1: Semi-variograms for soil water (down to a 130 cm soil depth) of a forested catchment in Texas showed that the water content was not significantly correlated to separation distances from the place of measurement in the plane or in a three-dimensional soil surface (Figures 1a and b). Analysis using elevation as separation distances showed significant correlation, with lower variance at higher water content (Figure 1c). The correlation in the x,y-plane increased for shallower soil water content down to a depth of 55 cm (Figure 1d).

SUMMARY OF DEMONSTRATION AND MISCELLANEOUS REPORT FINDINGS

MINNESOTA DROUGHT

SPECIAL REPORT 15.

WATER RESOURCES RESEARCH CENTER

R. Swerman, D. Baker, and R. Skaggs

MINNESOTA DROUGHT

ABSTRACT

Drought is an ordinary and expected part of the climate of any location. However, there are few measures of drought and often it is difficult to recognize when a drought has begun and when it has ended. In the United States, the Palmer Drought Severity Index (PDSI) is the most commonly employed measure of drought. Examination of the time and space averages and variability of the PDSI allows some conclusions about the climatology of drought over Minnesota to be drawn.

There is a consistent gradient in the duration and severity of drought occurrence from southwest to northeast across the state. The droughts in the southwest are more intense and have the longest duration. Toward the north and east, the droughts become much less severe and are much shorter in average duration. In the north and east there is a tendency for droughts that are more frequent, shorter, and milder.

The persistence is an outstanding characteristic of drought. Once a drought has become established, it tends to persist for several weeks to several months. The persistence is much stronger in the southern and western portions of the state. In these areas a drought established by the beginning of the growing season has a likelihood of 50 percent or greater of continuing through the end of the growing season in August. Thus, it is important to have a near real time monitoring system for drought during the late spring and early summer in order to anticipate the effects.

The frequency, severity, and duration of drought is not constant in time. In the early part of this century, from the early 1920's through the 1930's, much of Minnesota became progressively drier as measured by the PDSI. After 1940, precipitation increased substantially and drought was much less frequent and persistent. Much of the famous 1950's drought on the Great Plains did not affect Minnesota. The past five years or so have been among the wettest on record. However, it cannot be concluded that favorable moisture conditions and infrequent drought are likely in the future. A return to the drier conditions of the early part of this century with more persistent droughts should be expected sometime in the future. The question, which cannot be answered, is when these more prolonged drought conditions will reoccur.

Table 1: Maxima and minima of the Palmer Drought Severity Index.

Climatic Division	Maximum	Mo/Yr	Mo/Yr	Mo/Yr	Minimum	Mo/Yr
1	5.9	10/03			-6.0	9/34
2	5.6	11/05			-7.0	3-4/11
3	5.7	1/69			-7.8	2/77
4	5.2*	9/65	5/72		-9.9	7-8/34
5	5.2*	9/65			-9.9	7/34
6	4.9*	11/05	12/65	1/69	-8.0	4/11
7	6.5*	2/69			-7.2	7/11
8	5.8*	1/69			-8.0	8/34
9	4.8*	10/03			-7.1	6/34
*4)	7.0	9/86				
*5)	7.0	9/68				
*6)	6.3	9/86				
*7)	8.1	9/86				
*8)	6.1	10/86				
*9)	4.9	9/86				

* indicates PDSI values for 1986 that were not part of the analysis exceeded the maximum values during the analysis period.

Table 2: Annual probability of drought by severity

Frequency of months with occurrences exceeding the indicated values:

	Climatic Division								
Severity	1	2	3	4	5	6	7	8	9
-1.0	33.0	32.3	31.9	38.5	42.9	34.4	43.7	37.6	31.3
-2.0	22.1	23.1	17.6	25.6	29.9	23.7	31.1	20.9	20.4
-3.0	12.5	12.4	9.1	14.5	15.4	13.2	17.3	12.1	12.4
-4.0	6.5	5.2	3.8	8.7	9.4	6.2	7.2	6.5	5.4

SUMMARY OF
DEMONSTRATION
AND
MISCELLANEOUS REPORT FINDINGS

A COMPUTER PROGRAM TO EXCHANGE ERDAS AND EPPL7 DATA FILES.

SPECIAL REPORT 16.

WATER RESOURCES RESEARCH CENTER

K. Anderson and B. Scheer

A COMPUTER PROGRAM TO EXCHANGE ERDAS AND EPPL7 DATA FILES.

ABSTRACT

This report is the documentation and user manual for a FORTRAN-77 computer program to convert geographic information system (GIS) files and satellite between the ERDAS image analysis system and the EPPL7 GIS formats. This expands the data-capture ability of the EPPL7 GIS and also allows users to take advantage of the analytical techniques of both systems when using an IBM-AT or compatible computer equipped with Enhanced or Professional Graphics Adapter and appropriate ERDAS hardware and software.