APPLICATION OF DIGITAL MAPPING TECHNOLOGY TO THE DISPLAY OF HYDROLOGIC INFORMATION

A Proof-of-Concept Test in the Fox-Wolf River Basin, Wisconsin

By G. K. Moore[1], L. G. Batten[2],
G. J. Allord[3], and C. J. Robinove[3]

U.S. GEOLOGICAL SURVEY

Water Resources Investigations Report 83-4142

1 National Mapping Division, U.S. Geological Survey.
2 Technicolor Government Services, Inc. Work under U.S. Geological Survey Contract no. 14-08-0001-20129.
3 Water Resources Division, U.S. Geological Survey.

Sioux Falls, South Dakota

1983

UNITED STATES DEPARTMENT OF INTERIOR

JAMES G. WATT, Secretary

GEOLOGICAL SURVEY

Dallas L. Peck, Director

For additional information
write to:

Chief, Applications Branch
EROS Data Center
Sioux Falls, South Dakota
57198
(Telephone (605) 594-6114)

Copies of this report can be
purchased from:

Open-File Services Section
Western Distribution Branch
U.S. Geological Survey
Box 25425, Federal Center
Denver, Colorado 80225
(Telephone: (303) 234-5888)

CONTENTS

ILLUSTRATIONS

TABLES

CONVERSION FACTORS

For use of readers who prefer to use metric units, conversion factors for terms in this report are listed below:

Multiply	By	To Obtain
foot (ft)	0.3048	meter (m)
cubic foot per second per square mile [(ft^3/s)/mi^2]	0.01093	cubic meter per second per square kilometer [(m^3/s)/km^2]
cubic yard (yd^3)	0.7646	cubic meter (m^3)
gallon per minute (gal/min)	0.0630	liter per second (L/s)
million gallons per year (Mgal/yr)	3,785	cubic meter per year (m^3/a)
inch (in.)	25.40	millimeter (mm)
mile (mi)	1.609	kilometer (km)
square mile (mi^2)	2.590	square kilometer (km^2)

National Geodetic Vertical Datum of 1929 (NGVD of 1929): A geodetic datum derived from a general adjustment of the first-order level nets of both the United States and Canada, called NGVD of 1929, is referred to as sea level in this report.

ABBREVIATIONS

IDIMS - Interactive Digital Image Manipulation System.

VAX - Virtual Address eXtension.

WATSTORE - Water Data Storage and Retrieval System.

WRD - Water Resources Division, U.S. Geological Survey.

APPLICATION OF DIGITAL MAPPING TECHNOLOGY TO
THE DISPLAY OF HYDROLOGIC INFORMATION

A Proof-of-Concept Test in the Fox-Wolf River Basin, Wisconsin

---

By G. K. Moore, L. G. Batten, G. J. Allord, and C. J. Robinove

---

ABSTRACT

The Fox-Wolf River basin in east-central Wisconsin was selected to test concepts for a water resources information system, using digital mapping technology. A large amount of recently updated hydrologic information is available in this 16,800 km^2 basin. Several characteristics of the basin are also similar to those in many areas of the country: four surface water cataloging units and three important aquifers; neither great shortages nor great surpluses of water; and local water problems but not problems of extreme severity. The principles exemplified by information system processing in the Fox-Wolf basin are the same as those that can be used in other areas or nationwide.

Fifty data sets were included in the Fox-Wolf information system, not counting different forms of the same data set. Many data sets were digitized in vector form from 1:500,000 scale maps and overlays. All data were geometrically transformed into a Lambert Conformal Conic map projection and converted to a raster format with a 1-km grid cell resolution for further processing. In this format, the Fox-Wolf basin is an array with maximum dimensions of 160 by 250 grid cells.

Some thematic data were acquired in digital form from the WATSTORE, Water Use, Basin Characteristics, and Ground-Water Site Inventory files. A digital file of city and town populations was also prepared. The inclusion of these data sets required reading the source tape and correctly positioning data features within the raster array. Some data sets also required mosaicking, sorting, and selecting among closely spaced data points. Several of the digital data sets represent point samples of a continuous surface. These points and a minimum curvature algorithm (which approximates the steady state results of a heat-flow equation) were used to generate the surfaces. The result of digitizing and preliminary processing is a group of spatially registered data sets.

Parameter evaluation, areal stratification, data merging, and data integration were used to achieve the processing objectives and to obtain analysis results for the Fox-Wolf basin. Parameter evaluation includes the visual interpretation of individual data sets and digital processing to obtain new derived data sets. In the areal stratification stage of data

processing, masks were used to extract from one data set all features that are within a selected area on another data set. Most processing results were obtained by data merging. Merging is the combination of two or more data sets into a composite product, in which the contribution of each original data set is apparent and can be extracted from the composite. One processing result was also obtained by data integration. Integration is the combination of two or more data sets into a single new product, from which the original data cannot be separated or calculated.

There are eight main benefits of a digital water resources information system:

1. More complete and more objective resource information can be obtained by data analysis. Human errors of omission are minimized, and any human bias has less effect on results. All results are reproducible.

2. The data sets are permanently stored in a form that is easily updated and revised.

3. Digital data sets can be provided to other agencies in a readily usable form for other types of resource assessment or management.

4. There are many available options for formatting, data processing, and display; these options are easily and cheaply implemented.

5. Measures that would be impractical with conventional procedures can be derived from digital data sets (neighborhood slope and relief, for example). Other measures can be obtained faster than with conventional procedures (area measurements, for example).

6. Lower grade personnel can be used for digitizing and simple data processing. Skilled personnel can concentrate on planning, data analysis, and evaluation of results.

7. Both the spatial distribution and amplitude values (or labels) of features in the data sets can be used to obtain useful information. Some information can be obtained from single data sets, but other information is obtained by areal stratification, data merging, data integration, and modeling of multiple data sets.

8. Registered digital data sets offer an improved capability for the study of complex hydrologic problems; the data are in a format that is suitable for most distributed parameter models.

The storage and processing of digital maps cannot accomplish all objectives of a water resources summary. Some factors, relationships, and problems are best described in narrative form. Also, some data types, relationships, and totals are more understandable when shown in graphic, tabular, or statistical form. Spatial processing and display mainly offer a highly useful, objective method of showing facts, relationships, and results in readily understandable maplike forms.

The system requirements to produce national data sets and national information products for a water resources summary include hardware and software for (1) digitizing maps; (2) reading and entering digital point-data files, vector data sets, and raster data sets; (3) geometric registration to a selected map projection, conversion to a common raster or vector format, and resampling to a common resolution; (4) surface generation; (5) on-line storage of at least several complete data sets; (6) a considerable variety of data analysis options; (7) display and evaluation of data sets and processing results; and (8) generation of maplike products on film or paper.

INTRODUCTION

Over the past 50 years, more than 20 commissions, committees, and Federal agencies have studied the Nation's water problems and policies. The most recent results of these efforts were the "Second national water assessment" (U.S. Water Resources Council, 1979) and "Soil, water, and related resources in the United States, conditions and trends" (U.S. Department of Agriculture, 1981). It has proven difficult to adequately address local resources and problems in these studies because of limited funds and time. Each study has also required a new, complex, and expensive national compilation of all available water information. The previous 20 studies have demonstrated the need for a national water information system, in which the latest and best water resources data are compiled and maintained.

One of the Department of Interior goals for 1983 is "to establish and implement sound management concepts and practices." An objective within this goal is "to develop and make available information regarding lands and resources that permits the public and their congressional delegates to understand and act in an informed manner on critical natural resources issues." The subobjective for the Geological Survey is "to design and initiate (jointly with the Interior Department Office of Water Policy), by September 30, 1984, a continuing process for a water resources summary that will characterize the condition of the Nation's water resources." One task assigned to the Water Resources Division (WRD), for completion by June 30, 1983, consists of "designing a water resources information system to support water resources summaries."

For the purpose of designing a water resources information system, WRD has defined a summary as a process for continually evaluating the condition of the Nation's water resources, with the objective of identifying and understanding the problems and opportunities associated with man's activities and the use of water in various hydrologic settings. Timely water resources summaries thus present water conditions as a set of accounts, trends, and problem analyses from which inferences may be drawn about the capability to meet future demands for water. This presentation indicates the regional diversity of water supplies, water uses, and water problems. A national summary may not define individual local problems and may not be useful for water planning at a site, but it should reduce uncertainties about regional water conditions and the existence of water crises. Summaries should also contribute some of the information needed to evaluate the effects of governmental policies and programs in water resources development.

As presently envisioned by WRD, a water resources summary should include the following types of information:

1. Summaries of past and current water resources conditions, including availability, quantity, quality, and use.

2. Descriptions of water balances and trends in water conditions.

3. Identification and description of the severity and extent of major water problems.

4. Hydrologic and political boundaries.

A summary may also include statistical measures of water availability or quality by river basin or state, illustrations of water-supply and water-quality trends over time, descriptions of alternative solutions to water problems, and the predicted hydrologic effects of man's activities.

A water resources information system should be capable of description, prescription, and prediction; it should be a versatile and efficient method of compiling, revising, processing, and summarizing water information; and it should be capable of displaying the relationships of man's activities to water resources over space and time (see, for example, Calkins and Tomlinson, 1977, Estes, 1982, and Loveland and Johnson, 1982). WRD is evaluating various computer-based hardware and software systems that process and analyze digital data sets in map formats to accomplish these objectives.

A water resources information system (fig. 1) can be conceptually defined as a multilayered, georeferenced digital data base that can be used for inventory, description, update, problem solving, and prediction in an area. It consists of geographically registered digital data sets, associated files, and various data processing functions. The keys to this definition are (1) all data sets consist of a two-dimensional array of points, lines, or areas; some data sets also have attribute values that show magnitude or change over time, (2) all data arrays are registered and geographically referenced to a selected map projection, and (3) the data base can be used for a variety of objectives. In conventional terms, a water resources information system can be thought of as a group of registered map overlays plus comparison procedures which are determined by study objectives.

The storage and processing of digital hydrologic data in map forms cannot accomplish all objectives of a national water resources summary. For example, the relationships between components in hydrologic systems, the social and economic factors in water use, and some water problems are best described in narrative form. Other data types, relationships, comparisons, or totals are more understandable when shown in graphic, tabular, or statistical form. Information system processing and display mainly offer a highly useful, objective method of showing facts, relationships, and modeling results in readily understandable maplike forms. Digital data sets are also easily revised and updated.

The Fox-Wolf River basin in Wisconsin (fig. 2) was selected to test the concepts for a national water resources information system and a national summary. This basin drains 16,800 km^2 (6,500 mi^2), and thus is large enough to evaluate the data-set resolution that is required by a national summary. A large amount of hydrologic information is available in this basin (Olcott, 1968), and this information has recently been updated. The glacial geology of the area is similar to much of the northeastern United States. Water resources development in the vicinity of Green Bay has also produced a number of typical urban problems, including surface water pollution and water-level declines in wells.

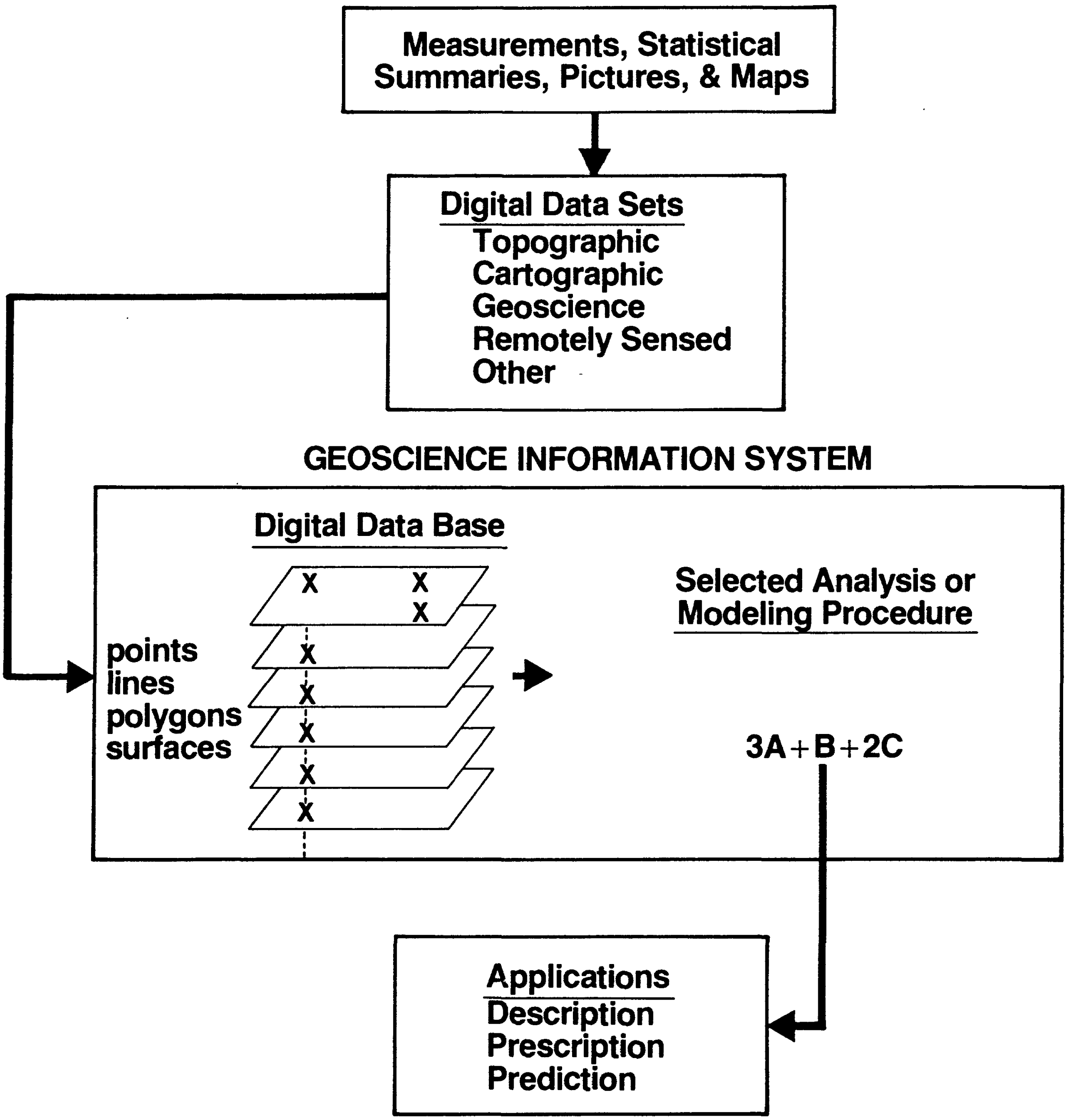
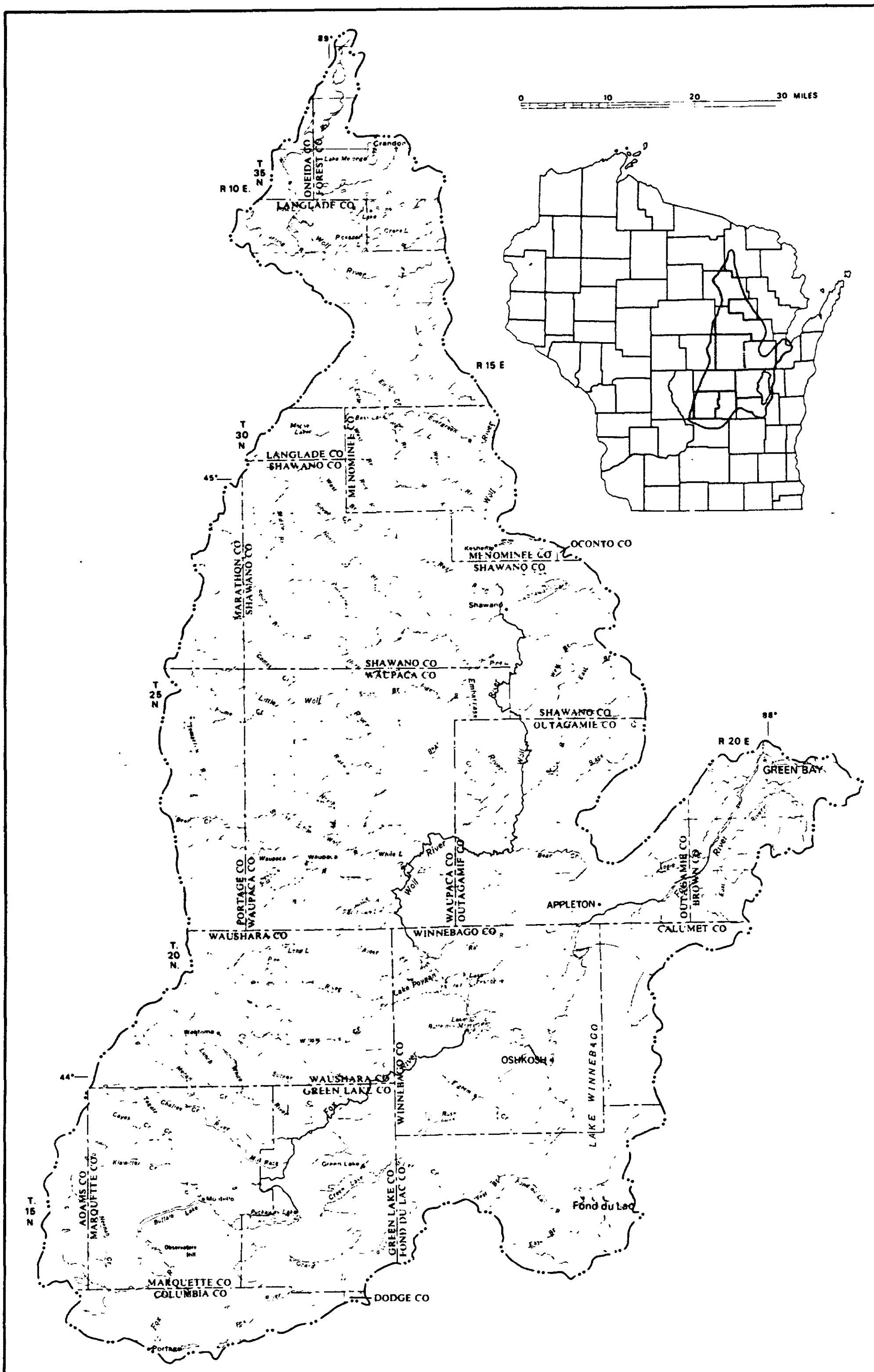


Figure 1.--A water resources information system is a multilayered, georeferenced digital data base that can be used for description, problem solving, and prediction in an area. It includes various data processing options in the available software.



Base from U.S. Geological Survey
State base map, 1968

Figure 2.--The Fox-Wolf River basin is in eastern Wisconsin and includes the city of Green Bay.

It is important to make a distinction between data requirements for a national summary and those for small project studies. A small project may require all types of data, and results may be limited only by the availability, accuracy, and reliability of these data. A typical objective in this case is to obtain site-specific information that can be used for water-supply planning and development. In contrast, a national summary is intended to provide information that can be used for regional policy and management decisions. It need not include all available data, because it is not meant to provide site-specific information. Instead, a national summary should show regionally important characteristics of the water resources for purposes of evaluation by stream reach, river basin, aquifer, water resources planning unit, and political unit. Categories and classes of water resources are generalized in a national summary, but local variability may be indicated by a range of attribute values or by a mean and standard deviation.

It is difficult to simulate a national water resources summary within a single river basin, because (1) the reader anticipates a large variety of data types, detailed data configurations, and a complete assessment of water resources in this size of study area, (2) the characteristics and problems of water resources in a single basin are limited in type, range, and number of classes, (3) results cannot be compared with those in adjacent basins or elsewhere in the country, (4) enlargement of map scale to show the Fox-Wolf basin at a visually acceptable size makes the effects of a low data resolution very apparent, and (5) some data sets, such as water-table elevations, are not representative of the accuracy that is possible at the map scale that is used. It should be kept in mind that the objective of this prototype study was to develop methods and procedures for the national summary.

The implementation of a water resources information system for the Fox-Wolf basin was done at the EROS Data Center in Sioux Falls, South Dakota by personnel of the Earth Resources Observation Systems (EROS) Office of the National Mapping Division, in accordance with design criteria specified by WRD. Cartographic and hydrologic data were supplied by Wisconsin District personnel of WRD. Data processing and modeling objectives were determined by WRD, in consultation with other project personnel.

This report describes the design and implementation of a prototype water resources information system for the Fox-Wolf River basin. The design decisions are discussed as are the constraints imposed by these decisions and the implications for a national water resources summary. This report also describes the use of the Fox-Wolf information system for an assessment of selected hydrologic conditions in the basin. Finally, results in the Fox-Wolf basin are evaluated in terms of objectives for the national summary.

A number of people made significant contributions to the successful completion of this prototype water summary. The authors especially appreciate the efforts of Bruce K. Quirk and Sue A. Mattson, Technicolor Government Services, at the EROS Data Center. Dr. Quirk processed the digital data sets from WATSTORE and similar files, generated the surfaces from these data, and solved two software problems, which were constraints on the data processing. Ms. Mattson digitized all maps and overlays of the

Fox-Wolf basin. Special thanks are due the WRD personnel in the Wisconsin District, who contributed good ideas and considerable time to the project; they include Ralph D. Cotter, Warren A. Gebert, Gerald L. Goddard, Timothy J. McElhone, and Wendy J. Danchuk. The authors also gratefully acknowledge the project guidance and direction provided by Joseph S. Cragwall, Jr., David W. Moody, Wesley L. Bradford, Edward S. Davidson, and H. C. Riggs in WRD Headquarters at Reston, Virginia.

INFORMATION SYSTEM COMPILATION AND ANALYSIS

The two phases in the development and use of a water resources information system are compilation and analysis. Compilation includes digitizing, formatting (vector or raster), geometric registration, positioning the data in digital files, resampling, surface generation, classification, and feature coding. Analysis includes parameter evaluation, areal stratification, data merging, data integration, and modeling.

Digitizing

Many water resources data sets are available only as thematic maps that show points, lines, polygons (areas), and surfaces (contours). The significant information on these maps must be digitized by raster or vector methods. Both methods utilize the geometry of maps or overlays to digitally record the relative locations and attributes of all significant features. Raster methods use a light-beam scanning system or a television camera for automatic recording of these features in a grid cell array. Vector methods use a cursor (an electronic pointer) for manual or semiautomatic recording of separate points and series of linked points (lines and polygon boundaries).

Raster methods are used mainly for the digitizing of aerial photographs and other images. Maps and overlays have been digitized with these methods, but the time spent in classifying, editing, enhancing, and encoding the digital products commonly results in costs that are higher than those of vector methods. The resolution of data produced by raster methods may also be lower than those of vector methods. Most television cameras are limited to data arrays of about 520 by 520 grid cells. The vector digitizing system used for maps of the Fox-Wolf basin automatically scales any size of study area into 64,000 units in the longest dimension; a square study area thus contains a data array of 64,000 by 64,000 points.

Vector digitizing methods mark individual points, the end points of straight lines, and groups of points along curved lines and boundaries. The recorded data include the end points and intersection points of all lines. A group of 40-50 control points (with known latitude and longitude locations) is also digitized in the first data set of a new study area. Fewer control points are required for subsequent data sets in the same map projection. The control points are used for a geometric transformation of the digitized data to a selected map projection.

Points, lines, and polygons are generally assigned to classes during the vector digitizing process. The digitizing system permits a maximum of 94 classes in a single data set. This limit has never been a problem, but a larger number of classes could be handled, if necessary. The digitized records include data that can be used for determining number of lines and arcs, number and locations of intersections, line length by class (including perimeter length), and total line or boundary length.

Vector digitizing proved suitable for all maps in the Fox-Wolf River basin and should also be appropriate for a national water resources summary. Modern digitizing systems have adequate flexibility, accuracy, and capacity to handle the anticipated requirements for national data sets.

Geometric Transformation and Formatting

After digitizing, the data sets must be geometrically transformed into the selected map projection. The Lambert Conformal Conic map projection is one of 20 available map projections (table 1) and was selected for the Fox-Wolf River basin. In Wisconsin, this map projection is also used as the State Plane Coordinate system. A Lambert Conformal projection is commonly used for areas less than a hemisphere in size and for all but area measurement purposes. The conformal characteristic means that angles and shapes of all features are true within small areas.

The known latitudes, longitudes, and x,y positions of the control points, which were selected during digitizing, are used to determine polynomial coefficients for the geometric transformation. These coefficients are calculated from a polynomial fit of the control points to their correct positions in the map projection. One software function then fits the data points between the control points by interpolation. A second software function fits all data points into position by a third order polynomial transformation, which is based on the coefficients of the control points. The latter procedure was used for the Fox-Wolf data sets, but either procedure should give satisfactory results for national data sets.

After geometric transformation, all digitized data sets in the Fox-Wolf basin were converted to a raster format. There are advantages to raster processing (table 2) that may be necessary for the analysis and modeling of national data sets.

The biggest advantage of a raster format is the capability for analysis and display of surface data sets. Surfaces are important data sets for almost any geoscience objective. In the raster domain, a wide variety of image analysis and display options are available for surfaces, and the data are handled as numerically continuous arrays. In the vector domain, surfaces must be level sliced, displayed as contours or as a 3-D mesh, and handled as polygons. Raster processing of surface data sets produces numerically continuous results for modeling purposes, whereas the results of vector processing are discontinuous.

The main disadvantage of a raster format is the inaccuracy that may occur in measurements of line length. This problem can be solved by measuring line length in the vector domain as a part of the digitizing process. Various complications arise, however, when the source of a data set is digital vector or raster data. Vector processing is also more efficient for some types of overlay analysis, and computer-driven plotter displays are faster and neater in vector form.

Table 1.--Map projections for information system processing.

1. Albers Conic Equal Area
2. Azimuthal Equidistant
3. Equidistant Conic
4. Equirectangular
5. General Vertical Near-Side Perspective
6. Gnomonic
7. (Hotine) Oblique Mercator
8. Lambert Azimuthal Equal Area
9. Lambert Conformal Conic
10. Mercator
11. Miller Cylindrical
12. Orthographic
13. Polar Stereographic
14. Polyconic
15. Sinusoidal
16. State Plane Coordinates
17. Stereographic
18. Transverse Mercator
19. Universal Transverse Mercator
20. Van der Grinten

Table 2.--The advantages of a raster format outweigh the disadvantages for many purposes.

ADVANTAGES OF RASTER FORMATS

1. Conceptually easy to understand.
2. Effects of resolution decisions are readily apparent.
3. Inclusion of interpolative data sets (surfaces).
4. Image processing and display options.
5. Simple procedures for data merging, data integration, and cartographic modeling; numerically continuous results.
6. Suitable format for finite element model parameters.

DISADVANTAGES OF RASTER FORMATS

1. Lines and edges are encoded and displayed as stairsteps.
2. Lines are not a uniform width.
3. Line and perimeter length is difficult to measure accurately.
4. Fine detail in some locales or some data sets may be thrown away.
5. Results of overlay analysis may be more difficult to evaluate.
6. Adjacency and connectivity measures may be more difficult to evaluate.
7. Multiple attribute data may take more time to encode and process.
8. Cell bounds have nonuniform latitude and longitude positions and spacings.

The ideal software system would handle both raster and vector data formats and would automatically convert from one format to the other for different types of data processing and display. All presently available software and hardware systems have limits, however, and it is necessary to obtain processing results within these constraints. Data format restrictions are less of a problem for information system objectives than are limits to the size of data arrays, types of data processing, and options for display of results.

The resolution (grid cell size) selected for the Fox-Wolf information system is 1 km by 1 km. The irregularly shaped 16,800 km^2 basin is thus within an array of 160 by 250 grid cells. In considering a national water resources summary, some hardware and software systems are limited to a data array of about 5,000 by 5,000 grid cells. At a 1 km resolution, national data sets would fit within this limit. The Fox-Wolf data sets are thus representative of the detail that can be used and displayed for national data sets.

Other Data Sources

A number of thematic data sets are available as digital arrays in vector or raster form. Other digital data are referenced to point locations (generally latitude and longitude) and are available from WATSTORE and similar data files. The inclusion of any type of digital data in an information system requires reading the source tape and correctly positioning data features within the array. Other processing steps that may be necessary are: (1) mosaicking blocks of data to cover the study area, (2) extracting only data that are located within the study area, (3) reducing noise or selecting among closely spaced data points, (4) resampling raster data to a selected grid cell size, and (5) converting from vector to raster format.

The Fox-Wolf basin information system includes digital topography data in raster format and point-located attribute data from various digital data files in WRD. The digital topography data were from the Digital Elevation Model (DEM) formatted, 1:250,000 scale data base, which is available from the National Cartographic Information Center (NCIC).

The 1° by 1° blocks of digital topography data were first mosaicked to fill the Fox-Wolf study area, geometrically transformed to a Lambert Conformal map projection, and resampled from a 3-arc-second to a 100-m grid cell resolution. Next, software algorithms were used to produce a shaded relief data set. Finally, neighborhood functions were used to reduce the grid cell resolution from 100 m to 1 km, while preserving as much topographic information as possible. These functions were used to determine mean, variance, and relief within 10 by 10 grid cell windows in the elevation data set. Thus the original topographic data set was used to produce four (including shaded relief) different data sets in the information system.

Digital data were used from WATSTORE, Water Use, Basin Characteristics, and Ground-Water Site Inventory files. A digital file of town populations was also prepared. The other information in these files is described later. All digital data consisted of point locations (referenced by latitude and longitude) and attribute values for these locations. A separate file in the IDIMS (Interactive Digital Image Manipulation System)[1/] software system was used for each attribute in each data file. Points in the data sets were then correctly positioned in raster format, and a separate layer in the information system was produced for each attribute.

The use of a raster format may cause a problem in the positioning of points from a digital data file; closely spaced points may fall into the same grid cell. In the case of a 1 km resolution, this situation may occur frequently, but it cannot be predicted. Cell bounds have unknown and nonuniform latitude and longitude positions and spacings. If a computer is instructed to position the points, any later entry in a grid cell will automatically delete an earlier entry in the same cell. This situation is generally considered to be unacceptable, because some data values may be more accurate or more reliable than others.

The partial solution that was used for the problem of closely spaced points in the Fox-Wolf basin was to first position the data points in raster space with a 500-m resolution. Fewer data points fall into the same grid cells at this resolution. After any required processing (such as surface generation) the resolution was reduced to 1 km by use of an averaging algorithm and a 2 by 2 grid-cell window. This solution will not be possible for creation of national digital data sets. Instead, the data points could first be positioned in a raster array of 30 by 30 arc second grid cells. An array of this type has a variable cell size (about 0.9 km by 0.7 km at 40° N latitude) and does not match a map projection. However, cell bounds are known; attribute values can be examined by a hydrologist in order to delete the overlapping data points that are less accurate or less reliable. After surface generation, the data sets can be geometrically transformed and registered by use of a resampling procedure.

Surface Generation

Most of the data sets that were acquired in digital form represent point samples of a continuous surface--elevations in wells of the top of the Cambro-Ordovician sandstone aquifer, for example. These points were used to generate the surfaces. For this purpose, the data points are assumed to be representative samples that can be used to accurately reproduce the surface.

Surfaces can be generated by any of three available procedures: neighborhood, Kriging, and minimum curvature functions. To interpolate a value at a point, the neighborhood functions look at the amplitudes of a

---

[1/] The use of proprietary hardware and software names is for identification purposes and does not constitute endorsement by the U.S. Geological Survey.

selected number (generally 3-6) of nearest known points and either weight each amplitude by the inverse of distance (or distance squared or cubed) to the known point or weight the amplitudes by the slopes of planes fitted to the points. Kriging assumes that the surface can be represented by determining the relationship between the amplitudes at the known points and the distances between points (Davis, 1973, p. 383). The Kriging method first calculates variograms that represent plots of semivariance (variance/2) against vector distance in three different directions (one at a 45° angle to the other two). These results are then used to find equations that describe the surface between known points. Minimum curvature is an iterative procedure to approximate the steady state results of a heat-flow equation, thereby interpolating a smooth surface between known data points and producing a smooth surface with minimum relief in areas of sparse control.

All surface generation functions tend to honor the amplitude values of the original data points, but all methods are subject to the introduction of unlikely local configurations (artifacts) near data points. Tests have shown that all three algorithm types produce similar results, and that nearly all results are reasonable and acceptable. These tests have also shown that contours on any of the computer-generated surfaces are more nearly identical than when contours of the original data points are drawn by several scientists.

The minimum curvature procedure was originally developed for automated computer contouring and is the same algorithm that is used for the contouring of geophysical data. When used for surface generation, it runs about four times faster than Kriging on a Vax 11-780 computer and requires less trial-and-error testing by an operator. Contours produced by the minimum curvature algorithm are slightly smoother and more realistic than those produced by the neighborhood functions. The minimum curvature function was used to generate surface data sets for the Fox-Wolf River basin; this function should be equally well suited for use with national data sets.

Other Preliminary Processing

Other types of processing may be necessary to prepare digital data sets for data analysis and modeling. In the Fox-Wolf project, this processing included boundary delineation, classification, and feature coding. It is sometimes desirable to display polygons as boundary lines rather than as digitally coded areas. Polygon boundaries are a by-product of the digitizing and vector-to-raster conversion procedures in the IDIMS software system. If desired, these boundaries can be retained in the data base as a separate digital data set.

Classification is a subjective selection of data set categories for future processing and must be done carefully to meet the objectives of the study. Examples are the groupings of counties and drainage basins into classes that represent ranges of water use or streamflow. Some classification of data sets in the Fox-Wolf basin was done prior to digitizing. Several data sets, which are inherent surfaces, were digitized as polygon amplitude ranges; examples are mean annual precipitation and the well yields from aquifers. These decisions were made to simplify and speed-up the data compilation process. The assumption was that detailed local configurations in these data sets would not be required for future processing. It should be kept in mind, however, that the polygon bounds represent contours on the original surface--some of the surface information is retained as class boundaries.

Feature coding is necessary because classes generally are given labels (digital values) of convenience during the digitizing process. Thus, coding is necessary to change these labels to the actual physical values. Alternatively, the data set features can be coded as byte data (within the 0-255 range for 8-bit digital data) for television or film displays. Coding can also be used to minimize digitizing requirements. For example, some of the same basins in the Fox-Wolf study area are shown in separate data sets that represent mean annual runoff and an index of low streamflow; the basin boundaries were digitized once, and two separate data sets were then created by copy and coding procedures.

Data Analysis

The result of digitizing and preliminary processing operations is a large group of registered data sets. A logical approach to the analysis of these data sets consists of parameter evaluation (including the generation of derived data sets), areal stratification of multiple data sets, data merging, data integration, model derivation, and operational model use. Only parameter evaluation, areal stratification, data merging, and data integration were required to achieve the objectives of the Fox-Wolf project.

Parameter evaluation includes the display and visual interpretation of individual data sets; it may also include digital processing to obtain new, derived data sets. The processing options at this stage are partly determined by the types of data sets in the information system (table 3) and by the inherent characteristics (table 4) of the data types. Many inherent characteristics of the data types, however, represent measures or features that can be extracted to produce new data sets (table 5) of the same or a different type. Data set conversions can be important or necessary to achieve project objectives. The processing options for parameter evaluation (table 6) typically include the obtaining of various measures, the location of internal features, and the characterization, enhancement, extraction, or reclassification of features, neighborhoods, and data sets. The processing options are constrained by the availability of software algorithms, but a wide variety of these algorithms are available in various software systems.

The Fox-Wolf data sets required only a small amount of additional processing at the parameter evaluation stage. Points are represented by a single grid cell in a raster data set and are barely visible on displays; a spread or proximity algorithm was used to incorporate a number of grid cells around the points in order to improve this visibility. The same software function was used to delineate 1 km zones of influence adjacent to the

Table 3.--Various types of data sets may be included in a water resources information system.

DISCRETE - DISTINCT AND NONCONTINUOUS [1]

- POINTS
- LINES
- POLYGONS (AREAS)

INTERPOLATIVE - NUMERIC AND CONTINUOUS [2]

- SURFACES

---

[1] Also termed categorical, dichotomous, and chloroplethic. May represent a nominal, ordinal, or numeric attribute.

[2] Also termed continuous and isoplethic. Numeric and continuous within the area of occurrence.

Table 4.--Each data type has inherent characteristics that determine processing flexibility and the options for derived data sets. All data sets have an identifying label and a resolution. Discrete data may be nominal, ordinal, or numeric (numbers may represent physical values or labels of convenience). Interpolative data are numeric and continuous. Distance from any feature in any data set may be an important parameter.

DISCRETE DATA

Points	Lines	Polygons
Number	Number	Number
Distribution	Distribution	Distribution
Spatial density	Spatial density	Perimeter nodes or arcs (number)
	Length	Perimeter length
	Straightness	Shape
	Direction	Area
	Pattern	Centroid location
	Intersection location	

INTERPOLATIVE DATA

Surfaces

Local mean amplitude, range, and variance
Slope (1st derivative)
Rate of change of slope (2nd derivative)
Aspect (relationship to a point or direction)
Regional and local configuration:
  Locations of peaks and pits
  Axes of ridges and valleys
  Directional trends
Image attributes:
  Texture
  Pattern
  Feature enhancement
  Feature identification

Table 5.--The characteristics of a data set represent measures or features that can be extracted to produce a new data set of the same or a different type.

DATA TYPE CONVERSIONS

Surface to surface - Derive neighborhood characteristics (relief, variance, or derivatives); derive residuals by surface trend analysis; factor analysis or principal components transformation of multiple data sets.

Surface to polygon - Define classes and slice, rank, or weight amplitude ranges; locate class boundaries by slope or amplitude change; label pattern or configurations; classification of multiple data sets by minimum variance clustering (unsupervised training) and discriminant analysis (supervised training).

Surface to line - Delineate axes of amplitude high, lows, or gradients; delineate other directional trends.

Surface to point - Extract points of high or low amplitude (peaks and pits).

Point, line, or polygon to surface - Measure and interpolate spatial frequency; measure distance from features or centroids; interpolate between data points that represent presence by 1 and absence by 0.

Polygon to polygon - Reclassify based on local or neighborhood mode; rank or weight (linear or nonlinear) polygon classes; aggregate classes; overlay and combine data sets.

Other - Determined by inherent characteristics of features, surroundings, and intersections.

Table 6.--Data processing begins with parameter evaluation and areal stratification. There are a variety of processing and graphic options for these procedures.

PROCESSING OPTIONS

1. Distance from points, lines, or polygons.
2. Feature count. Boundaries that separate presence and absence of features.
3. Line length, straightness, and directionality.
4. Polygon perimeter, area, and shape index.
5. Adjacency analysis by autocorrelation (subjective or numeric evaluation of connections between features and their surroundings).
6. Neighborhood characterization:
   A. Slope (first derivative).
   B. Rate of change of slope (second derivative).
   C. Aspect: relationship to an interior or exterior point (orientation and compass direction).
   D. Spatial frequency of occurrence.
   E. Other characteristics: mean, mode, variance, texture, and relief.
7. Reclassification of data sets or selected features within data sets:
   A. Rescaling (linear or nonlinear).
   B. Ranking or weighting of features.
   C. Grid cell interpolation or aggregation.
   D. Level slicing: anomaly definition.
   E. Replacement of amplitude values with neighborhood mean or mode.
8. Enhancement of features in surface data sets, especially high spatial frequency enhancement.
9. Ridge and trough axes in surface data sets. Extraction of edge and line segments.
10. Local correlation coefficients and other statistical measures for two or more data sets.
11. Surface trend analysis.

GRAPHICS OPTIONS

Graphics	Maps
Histograms of occurrence	Point Maps
Cumulative histograms	Line maps
Surface profiles along selected lines	Polygon maps
	Interpolative (surface) maps or images
	Shaded relief maps
	Isometric projections
	Linear perspective views

streams and lakes; the assumption is that activities-of-man within these zones will have more effect on surface water quality than will the same activities farther away.

A part of parameter evaluation during the Fox-Wolf study was the documentation of selected individual data sets (Appendix) for later visual review and project planning purposes. Film images of each data set were made on both Polaroid print film using a Matrix camera, and black-and-white negative film using a laser-beam film recorder. The stairsteps in a raster data format are less obvious and objectionable on a film output product than on a map drawn by a plotter. A film product is also faster and easier to produce, because special command files are required to drive a plotter. These files must be created for each different data set.

In the stage of work called areal stratification, small areas of two or more data sets are extracted and combined to evaluate local relationships. Measures and features can also be obtained from one data set under constraints imposed by other data. These constraints can be numerical, or they can be in the form of a mask, to extract from one data set all features that are within a selected area on a second data set (a cookie-cutter type of overlay analysis). This type of extraction is accomplished by assigning a value of 1 to grid cells in the desired area and a value of 0 to the grid cells in the surrounding area. The first data set is then multiplied by the mask. The operation and graphic options for areal stratification are generally the same as those for single data sets (table 6).

The cookie-cutter type of areal stratification was used to obtain several processing results in the Fox-Wolf basin. This operation was used, for example, to isolate and show waste disposal sites that are near streams or within flood-prone areas.

In the next stage of work, information is obtained by the merging and integration of data sets. Merging consists of combining two or more data sets into a composite product. The contribution of each original data set is apparent and can be extracted from the composite. The purpose of data merging is to compare positions and identify features, anomalies, and trends in the composite. Operations and intrinsic graphics options (table 7) include pattern, color, and stereoscopic composites.

Integration is the combination of two or more data sets into a single new product from which the original data cannot be separated or calculated. The purpose of data integration is to enhance and extract features which are obscure in the original data sets. The occurrence of such features in the integrated data indicates a partial correlation of the original data sets. However, the integrated data have a physical significance that is different from any of the original data sets. Integration operations (table 7) include arithmetic combinations, classification, and various logic functions and decision rules. The graphics options for integrated data sets are generally the same as those for single data sets (table 6). Other options include plots, graphs, and statistical measures of data set relationships.

Table 7.--The purposes and options for data merging and data integration are different.

Merging operations and graphics:

1. Patterned map.
2. Combined map and image.
3. Color composite map or image.
4. Steroscopic map or image.

Integration operations:

1. Addition, subtraction, multiplication, or division.

2. Maximum or minimum--highest or lowest value in two or more data sets.

3. Binary logic functions: >, <, =, AND, OR, TRUE, and FALSE--used to select and combine significant levels and ranges in multiple data sets.

4. Principal components transformation--creates uncorrelated data sets from original, partly correlated data.

5. Albedo calculation or color space transformation--for multispectral, remotely sensed data; reduces the number of data sets.

6. Multiband classification: parallelepiped, minimum distance, maximum likelihood, or other algorithm--for multiband data of the same type (such as multispectral remotely sensed data); extracts similar areas from the data sets.

7. Multiset classification--class assignments result from multiple decision rules for amplitude and spatial measurements in grid cell neighborhoods.

These additional options are feasible products of an information system, but they do not require the spatial relationships of data in an information system.

Nearly all processing objectives for the Fox-Wolf basin (figs. 3-14) were achieved by data merging. Various significant data sets were combined by overlay analysis procedures and were then color coded for display and for film recording of results. The only need for data integration was the combining of well yields from various aquifers to show the maximum well yield from any aquifer.

It is anticipated that the processing requirements for a national water resources summary will be similar to those for the Fox-Wolf basin. The data sets that are available on maps and in WATSTORE and similar files generally have some inherent significance, or they represent the results of manual analysis and interpretation procedures--the data have already been processed to the point where they have intrinsic importance for a national summary. Thus, a national summary can probably be achieved with relatively simple forms of parameter evaluation, areal stratification, and data merging and integration. It is possible, however, that some objectives will require more complex models of the data sets.

A wide variety of models have been used for various purposes in geoscience studies. In recent years, most models have been objective types that are implemented by computer, but subjective interpretation and expert opinion have been important for a few models (table 8). All the models in table 8 are at least partly dependent on the spatial attributes of available data. Some models (distributed parameter types) are entirely cartographic, whereas others (lumped parameter types) are non-cartographic.

The registered data sets in a water resources information system can be used in a variety of ways. Statistical measurements and summaries can be made of single data sets, and regression analysis can be used to determine the relationships between data sets. Similarly, the data sets can be processed and input to both lumped parameter and distributed parameter models. The most effective use and maximum benefit from an information system are obtained by exploitation of both the spatial and amplitude relationships of multiple data sets.

Data sets in an information system can be thought of as a group of registered map overlays. The answer to some hydrologic questions in any area requires the preparation of only a single map overlay. This type of question can also be answered by a single data set in an information system. However, an information system is expensive to prepare and compile. Questions and processing objectives that require only one or two data sets, by themselves, are not cost-effective uses of the system.

The answer to another type of hydrologic question is obtained from the relationships between data sets. The techniques of areal stratification, data merging, and data integration are designed for the detection and evaluation of spatial correlations. On the other hand, many procedures exploit only the amplitude relationships at points in the data sets or

Table 8.--A wide variety of models have been used in geoscience studies. Adapted and modified from Singer and Mosier, 1981, p. 1008.

GEOSCIENCE MODELS

1. Subjective evaluation based on knowledge, experience, and logic; evaluation based on genetic model and inference network.

2. Extension of directional trends, as shown by lines, polygons, and surfaces.

3. Anomaly definition.

4. Addition (weighted or unweighted), ratio, or other arithmetic combination of data.

5. Determination of probability, based on geometric attributes or concepts.

6. Adjacency analysis, based on autocorrelation.

7. Bayesian statistical inference--the use of prior results to determine probability in a new data set.

8. Analysis of ranges and frequencies within distances from local features.

9. Frequency analysis based on a different, well-known and representative area.

10. Surface trend analysis--determination of regional configuration and calculation and interpretation of residuals.

11. Minimum variance clustering (unsupervised training).

12. Multivariant discriminant analysis (supervised training).

13. Factor analysis or principal components transformation.

14. Time series analysis.

15. Multiple regression analysis.

16. Mathematical simulation.

determine trends through time; these procedures do not consider cartographic positions or the spacings between points. Conventional digital files, such as those in WATSTORE can be used for the latter procedures, because geometric registration is unnecessary. By itself, this is not a cost-effective use of an information system.

A water resources information system need not be compiled for every problem of a cartographic type, because the human mind is an efficient and effective tool for the evaluation of small numbers of map overlays. Manual analysis of these overlays, however, requires more personnel expertise than does computer analysis of digital data sets. In the manual analysis of a group of map overlays, for example, the mind is used to visualize implications and to evaluate three-dimensional conceptual models of occurrences, trends, gradients, and relationships. As amounts and complexities of the data increase, this procedure becomes progressively more time consuming and subject to errors of omission. Manual analysis techniques are further limited by the mathematical, statistical, and artistic skills of the scientist. Computer analysis techniques, on the other hand, are further limited only by imagination and software availability.

Once a water resources information system has been compiled, it can become a logical first choice for one-stop information shopping. Single data sets might be extracted from an information system faster than they could be located in original source materials. Similarly, it might be easier to obtain values for a lumped parameter model or a multiple regression analysis from the layers in an information system than from several digital data files. Nevertheless, the cost-effectiveness of an information system is best shown by answers to questions that require multiple data sets and a variety of analysis procedures.

DATA SETS FOR THE FOX-WOLF BASIN

The sources for data sets in the Fox-Wolf information system were maps and various digital data files. A total of 50 different data sets were included in the information system, not counting different types of the same data set (table 9). Numerous other data sets are available and could have been included. The 30 data sets in the Appendix include (1) data sets used for the processing objectives, (2) very significant single data sets, and (3) other representative examples of the types of data sets. Each data set is shown in image form, along with an explanation of its source, initial form, method of entry, and any problems associated with data entry. All maps were digitized at a scale of 1:500,000.

Point data sets show the locations of features that have a small areal extent at the resolution of the data base. Points can be digitized as a single class, but generally are given identifying labels or physical values. A labeled data set is a more flexible format, because additional data sets can be created by copy and coding procedures. In this manner, liquid waste disposal sites were coded (labeled) to show the type of disposal. Once points have been digitized as a single class, they can only be reclassified by overlay analysis, using polygons in other data sets as cookie cutters.

Table 9.--Digital data sets of many types were included in the Fox-Wolf water resources information system.

Geographic

1. Mean land surface elevations in 1 km cells - surface.
2. Variance of land surface elevations in 1 km cells - surface.
3. Relief of land surface elevations in 1 km cells - surface.
4. Shaded relief map - surface.
5. County lines - line and polygons.
6. Populations of cities and towns - points.

General Hydrology

1. Surface-water divide for basin - line and polygon.
2. Ground-water divide for basin - line and polygon.
3. Hydrologic cataloging units - lines and polygons.
4. Minor drainage basins - lines and polygons.
5. Drainage lines - lines.
6. Lakes and double-line streams - polygons.
7. Mean annual precipitation - lines and polygons.
8. Mean annual evapotranspiration - lines and polygons.
9. Trout streams - lines
10. Soil infiltration classes - polygons.
11. Wet soil areas - polygons.
12. Hydrologic unit map - lines and polygons.

Surface Water

1. Gaging stations and partial-record stations - points.
2. Mean annual runoff - polygons, points, and surface.
3. Centroids for basins with mean annual runoff values - points.
4. Low flow index; 7-day average, 10-year recurrence low flows - polygons.
5. Suspended sediment yields - polygons.
6. Water temperatures at base flow - points.
7. Specific conductance of water at base flow - points.

Table 9.--Continued.

Ground Water

1. Extent of the Cambro-Ordovician sandstone aquifer - polygons.
2. Well yields in the sandstone aquifer - polygons.
3. Elevations of the top of the sandstone aquifer - points and surface.
4. Average elevation of water levels in wells, 1945-54 - points and surface.
5. Average elevation of water levels in wells, 1975-83 - points and surface.
6. Dissolved solids in water in the sandstone aquifer - polygons.
7. Extent of the Silurian dolomite aquifer - polygons.
8. Well yields in the Silurian dolomite aquifer - polygons.
9. Well yields in the shallow glacial sand and gravel aquifer - polygons.

Water Problems

1. Landfills - points.
2. Liquid waste disposal sites - points.
3. Polluted lakes and stream reaches - polygons and lines.
4. Flood-prone areas - polygons.
5. Water-level declines in the Cambro-Ordovician sandstone aquifer - surface, polygons, and lines.

Water Use

1. Power plants - points.
2. Major points of ground-water use - points.
3. Major points of surface-water use - points.
4. Total ground-water use by county - polygons.
5. Industrial and commercial ground-water use by county - polygons.
6. Irrigation and stock ground-water use by county - polygons.
7. Residential ground-water use by county - polygons.
8. Total surface-water use by county - polygons.
9. Industrial and commercial surface-water use by county - polygons.
10. Irrigation and stock surface-water use by county - polygons.
11. Residential surface-water use by county - polygons.

Several data sets from the various digital data files represent points with numeric attributes: town population, site water use, and specific conductance and temperature of base flows at gaging stations. The points were digitally positioned and then numerically coded with the attribute values.

Several data sets from the digital data files represent point samples of a surface: elevations in wells on the top of the sandstone aquifer and elevations of water levels in wells, 1945-54 and 1975-83. Each set of points was digitally positioned, and the surface was interpolated.

Drainage lines are the only data set that is an inherent line type in the Fox-Wolf information system, because lakes and double-blue-line streams were digitized as polygons. As a result, the drainage system cannot be shown as a network unless the separate line and polygon data sets are combined. It may be desirable to have only line data in a national drainage network. Another decision consisted of digitizing drainage lines in only three classes: streams in the Fox River basin, streams in the Wolf River basin, and streams below the confluence of the Fox and Wolf Rivers. Maps showing reaches designated as trout streams or as polluted segments were digitized separately. As a result, these separate significant stream reaches do not exactly overlie drainage lines in the original data set.

Line data sets offer more flexibility for reclassification at the digitizing stage than do point data. Each digitized line is controlled by points at each end and at junctions with other lines. Any line segment can be automatically deleted between any two points. After deletion, new junction points can be added for the beginning and ending of important reaches (such as a trout stream or a polluted reach), and the short deleted segments can then be redigitized. This approach can be used to solve the problem of an inexact overlay (registration) of drainage lines.

Most data sets (table 9 and Appendix) in the Fox-Wolf information system were digitized as polygons. The procedure consists of digitizing boundary lines and of labeling the areas between bounds. Two data sets are produced, and either or both can be incorporated in the information system. One shows all boundary lines in the data set as a single class; the other data set shows all polygons as contiguous areas. Some contour maps were included in these data sets (mean annual precipitation and evapotranspiration, well yields in the sandstone and glacial aquifers, and dissolved solids content of water in the sandstone aquifer). The procedure in this case consisted of (1) digitizing the contours as polygon boundaries, (2) applying labels of convenience to the resulting polygons, and (3) describing the numeric range for each label in a separate file.

Polygon data sets offer a considerable flexibility for reclassification, either at the digitizing stage or during later processing. Since each polygon is labeled, it can be coded with any physical value, or it can be deleted at the data processing stage. The deletion of polygons does not change boundary lines in the second data set, however. For this purpose or that of revising the data set by adding new polygons, it is necessary to return to the digitizing stage. The bounds between junction points are deleted and redigitized, as was described for line data.

The other data sets in the Fox-Wolf information system are surfaces showing, for example, elevation of the land surface and various derivative products. The source, characteristics, and processing of these data were described previously. Surface data sets offer the greatest flexibility for reclassification and further processing. Any number of classes can be selected and displayed as polygons by level slicing into numeric increments.

The most critical task in preparation of an information system is the selection of data types, levels of detail, and numbers of classes. Decisions must be made about the compilation of many data sets; wrong decisions constrain processing options or affect processing results. Several wrong decisions in the Fox-Wolf project were described previously; another was the digitizing of too much detail in the drainage network and basin boundaries data sets. These data sets are crowded on displays, and processing results may be difficult to interpret.

Many other decisions were necessary in the preparation of data sets for the Fox-Wolf basin. One example, described previously, was the digitizing of several surface data sets as polygons. Another example was the reduction in number of polygon classes in some data sets. The original map showing chemical quality of water in the Cambro-Ordovician sandstone aquifer included eight classes of total dissolved solids; for summary purposes, these were grouped into three classes:

Original classes, in mg/L	Summarized classes, in mg/L
100 - 200	Less than 500
200 - 300	
300 - 400	
400 - 600	500 - 1,000
600 - 800	
800 - 1,000	
1,000 - 2,000	More than 1,000
over 2,000	

Decisions of this type must be based on experience, knowledge of the hydrologic significance of the data sets, understanding of data reliability and precision, data processing goals and objectives, knowledge of the significance of class bounds and intervals, and understanding of the effects of these decisions on processing procedures and results. Each data set that is entered in a water resources information system must be screened to ensure that all decisions make good hydrologic sense as well as good sense for stated or anticipated processing requirements. These decisions are especially important for national data sets, where there may be different data sources, different class intervals, and differing amounts of detail in the source materials. In this case, questions may also arise as to whether or not data sets from different parts of the country are equivalent -- whether they should be combined into a single data layer or included as separate layers.

USES OF THE FOX-WOLF INFORMATION SYSTEM

The questions and processing objectives (table 10) for the Fox-Wolf information system were selected to (1) describe several aspects of the water resources in this area, (2) show the relationship of man's activities to various water problems, (3) demonstrate answers to questions that require multiple data sets, (4) illustrate the analysis of several data types from a variety of sources, (5) demonstrate various data processing procedures, and (6) show several different types of processing results.

Six categories of questions have been asked of the information system. These categories are surface-water supply, surface-water quality, ground-water supply, ground-water quality, water use, and water problems. Many other questions could be answered by the data layers in the information system. The selected questions and the processing results (figs. 3-14) are representative of the capabilities, benefits, and limitations of a water resources information system. In answering each question, below, an explanation is given as to how the data sets were produced, handled, analyzed, and displayed.

How Is the Drainage Network of the Fox-Wolf River Basin Related to the Physiography?

The data sets (Appendix) chosen to respond to this question are (1) a shaded relief map of topography in the region, (2) a polygon of the surface-water divide for the Fox-Wolf basin, (3) a polygon of the ground-water divide for the basin, (4) polygons of the four hydrologic cataloging units in the basin, (5) the streams, (6) the lakes, and (7) locations of streamflow measurement sites. The source for the shaded relief map was digital topography data from 1:250,000 scale maps. All other data sets were digitized from 1:500,000 scale maps in vector form and were then converted to raster form. Streamflow measurement sites were digitized as points, the streams were digitized as lines, and the other data were digitized as polygons.

In merging the separate data sets to produce a colored information product (fig. 3), the polygons for the surface-water and ground-water divides were combined and used to produce a lighter color inside the basin on the shaded relief base. The boundary lines for these polygons (by-products of the polygon digitizing process) are also shown to outline the basin. The polygons for the four hydrologic cataloging units are not shown on the composite illustration, but these polygons were used to separately mask, classify, and color code all streams and lakes within the area of each unit. Lakes are shown as polygons on the composite map. The 22 regular gaging stations were extracted from all points where streamflow information is available; this procedure was possible, because the points were labeled

Table 10.--Six questions about water resources in the Fox-Wolf basin result in 12 objectives for data analysis.

I. QUESTION - HOW IS THE DRAINAGE NETWORK OF THE FOX-WOLF BASIN RELATED TO THE PHYSIOGRAPHY?

1. Objective - Show the drainage network in the Fox-Wolf basin, color coded by hydrologic cataloging unit, on a shaded relief map of the physiography; also show the regular gaging stations. Data sets - Shaded relief map of the region, surface-water divide, ground-water divide, hydrologic cataloging units, streams, lakes, and streamflow measurement sites. Analysis - Code, classify and combine data sets; color code features.
2. Objective - Show the drainage system in a way that emphasizes the many lakes in the basin and de-emphasizes the hydrologic cataloging units; also show all sites where some streamflow information is available. Data sets -Surface-water divide, ground-water divide, streams, lakes, hydrologic cataloging units, and all streamflow data collection points. Analysis - Combine data sets and color code features.

II. QUESTION - WHAT IS THE AVAILABILITY OF SURFACE WATER IN THE BASIN?

3. Objective - Show the basic water budget and centers of water use. Data sets -Surface-water divide, ground-water divide, county lines, mean annual precipitation, mean annual runoff, and populations of cities and towns. Analysis - Code and classify data sets; proximity analysis to show population classes; combine data sets and color code features.
4. Objective - Show an index of low streamflow and the specific conductance of surface water at base flows. Data sets - Surface-water divide, ground-water divide, streams, lakes, index of low streamflow, and specific conductance of base flows at gaging stations. Analysis - Code, classify, and combine data sets; color code features.

III. QUESTION - WHAT IS THE NATURAL QUALITY OF SURFACE WATER AND HOW IS IT AFFECTED BY MAN?

5. Objective - Show the relationship of an index of low streamflow to designated trout streams and to polluted streams and lakes. Data sets - Surface-water divide, ground-water divide, county lines, index of low streamflow, trout streams, polluted streams, and polluted lakes. Analysis - Code, classify, and combine data sets; color code features.
6. Objective - Show the landfills that are within 1 km of a stream or lake and the landfills that are farther away. Data sets - surface-water divide, ground-water divide, streams, lakes, county lines, and landfill sites. Analysis - Proximity analysis and overlay analysis; classify and combine data sets; color code features.
7. Objective - Show the landfills and liquid waste disposal sites that are within flood-prone areas. Data sets - Surface-water divide, ground-water divide, county lines, landfills, liquid waste disposal sites, and flood-prone areas. Analysis - Overlay analysis; classify and combine data sets; color code features.

Table 10.--Continued.

IV. QUESTION - WHAT IS THE AVAILABILITY OF GROUND WATER IN THE BASIN?

8. Objective - Show the maximum well yield from the best aquifer in any area. Data sets - Surface-water divide, ground-water divide, county lines, well yields and extent of the Cambro-Ordovician sandstone aquifer, well yield and extent of the Silurian dolomite aquifer, and well yields in the glacial sand and gravel aquifer. Analysis - Maximum; combine data sets; color code features.
9. Objective - Show the maximum amount of water that can be pumped from a well at any point and identify the aquifer or combination of aquifers. Data sets - Surface-water divide, ground-water divide, county lines, well yields in the glacial sand and gravel aquifer, well yield and extent of the Silurian dolomite aquifer, and well yields and extent of the Cambro-Ordovician aquifer. Analysis - Code and add; combine data sets; color code features.

V. QUESTION - WHAT ARE THE WATER PROBLEMS IN THE SANDSTONE AQUIFER?

10. Objective - Show the water-level drawdown in wells tapping the sandstone aquifer and total ground-water pumpage by county for industrial and commercial use. Data sets - Surface-water divide, ground-water divide, county lines, total industrial and commercial ground-water use by county, and water-level drawdown in wells from the 1945-54 period to the 1975-83 period. Analysis - Code, classify, and combine data sets; color code features.
11. Objective - Show areas where water from the sandstone aquifer exceeds 500 mg/L and 1,000 mg/L of total dissolved solids. Data sets - Surface-water divide, ground-water divide, county lines, and content of total dissolved solids in water. Analysis - Classify and combine data sets; color code features.

VI. QUESTION - IN WHICH AREA MAY GROUND - WATER POLLUTION HAVE A SERIOUS IMPACT?

12. Objective - Show the areas where the glacial sand and gravel aquifer consists of coarse-grained materials and yields large amount of water to wells; also show the distribution of landfills in the basin. Data sets - Surface-water divide, ground-water divide, county lines, well yields in the glacial sand and gravel aquifer, and landfills. Analysis - Code and combine data sets; color code features.

Figure 3.--The Fox-Wolf basin consists of a rolling topography, as shown by the shaded relief background. The drainage network includes four hydrologic cataloging units, as shown in different shades of blue. The red squares show the locations of the 22 regular gaging stations.

(classified) during the digitizing process. The final step was to increase the visibility of the gaging station points by showing these points as a 3 by 3 array of grid cells and by color coding.

The intent of an illustration like figure 3 is to create an impression in the viewer's mind of the physiography, the drainage, and the relationships between these features in the river basin. Information on the hydrologic cataloging units is included so that a viewer can compare the physiography and the drainage features in the four units. Gaging stations are included to show the locations in the basin where streamflow information is collected regularly.

In compiling the original data sets and producing the merged composite, several decisions that affected the composition of the illustration were made. The first decision was that a 1 km resolution was generally adequate to show the types of topography in the basin and that the topography can be visualized best on a shaded relief map. The second decision was that the density of drainage shown by 1:500,000 scale maps was appropriate for the resolution of the data sets. The third decision was the selection of data sets for the composite product. There were also other decisions concerning data set format, data set precedence, and color selection; these are discussed later. It is important to understand, however, that some trial-and-error experimentation may be necessary to produce the best combination of data sets in a colored composite, in which all features contribute to the intended visual impression.

An alternative illustration (fig. 4) showing most of the same information in a different way leaves a different impression of drainage characteristics in the mind of the viewer. A shaded relief base is not included in the colored information product; streams are shown as a single class and a single color; lakes are colored differently than streams; and 91 steamflow measurement points (regular gaging stations and partial record sites) are included as a single class to show all points where some streamflow information is available. In addition, polygon boundary lines are used to show the surface-water and ground-water divides for the Fox-Wolf basin and to show the drainage divides for the hydrologic cataloging units. These few changes in data set selection and handling emphasize the many lakes in the basin and the many points where some streamflow information is available. The changes also omit the physiographic relationships of the drainage system and de-emphasize the importance of the hydrologic cataloging units.

What Is the Availability of Surface Water in the Basin?

Two illustrations are used to address this question. The first (fig. 5) shows mean annual precipitation and water runoff for the Fox-Wolf basin. The second (fig. 6) shows an index of low streamflow, which is important in determining perennial water availability without storage facilities. The second illustration includes information on the specific conductance of surface water at base flow and can also be used to address questions about the chemical quality of surface water.

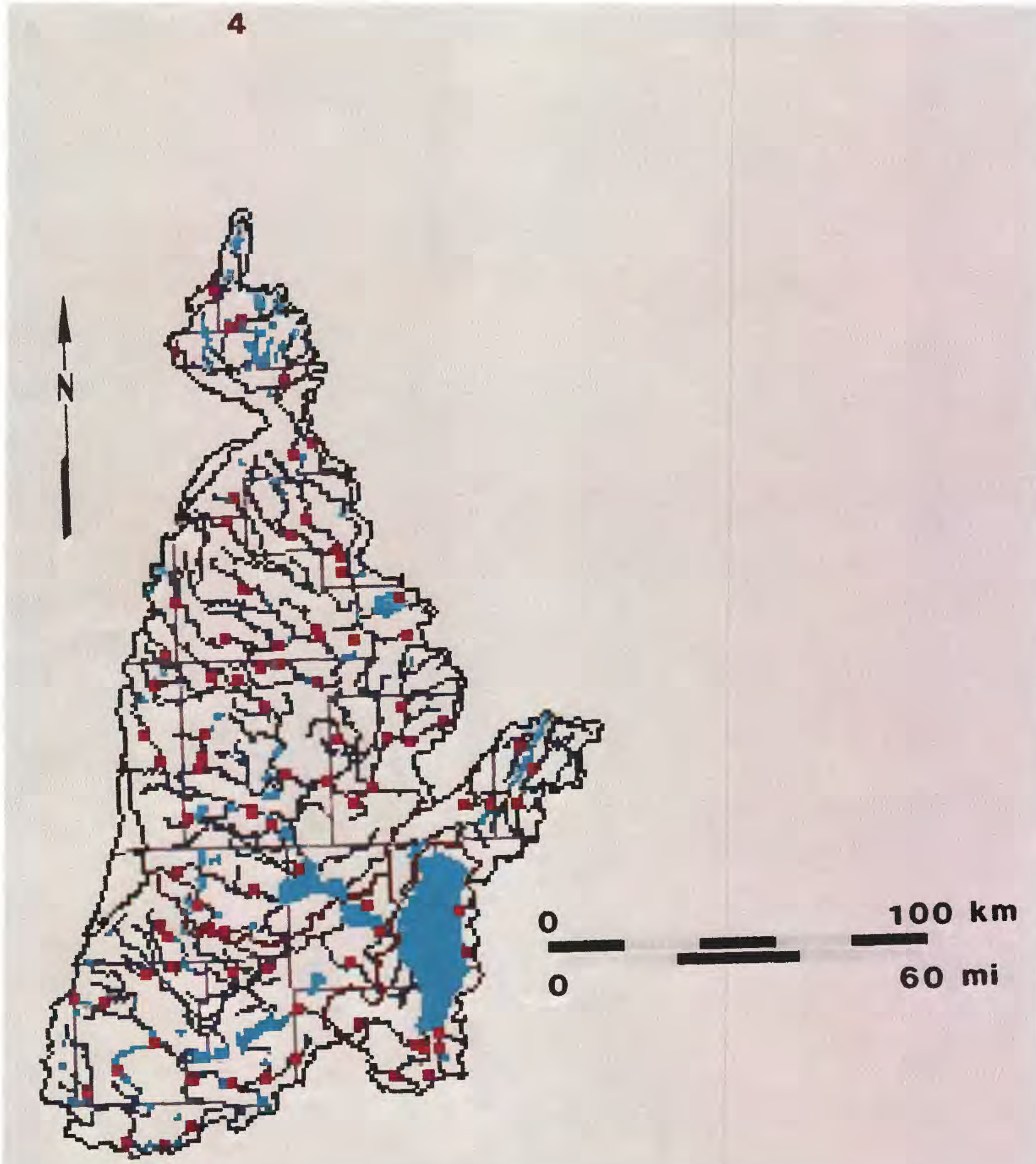


Figure 4.--Different shades of blue for streams and lakes emphasize the many lakes in the basin. Red lines show the boundaries of the four hydrologic cataloging areas. The small squares show all 91 points where some streamflow information is available.

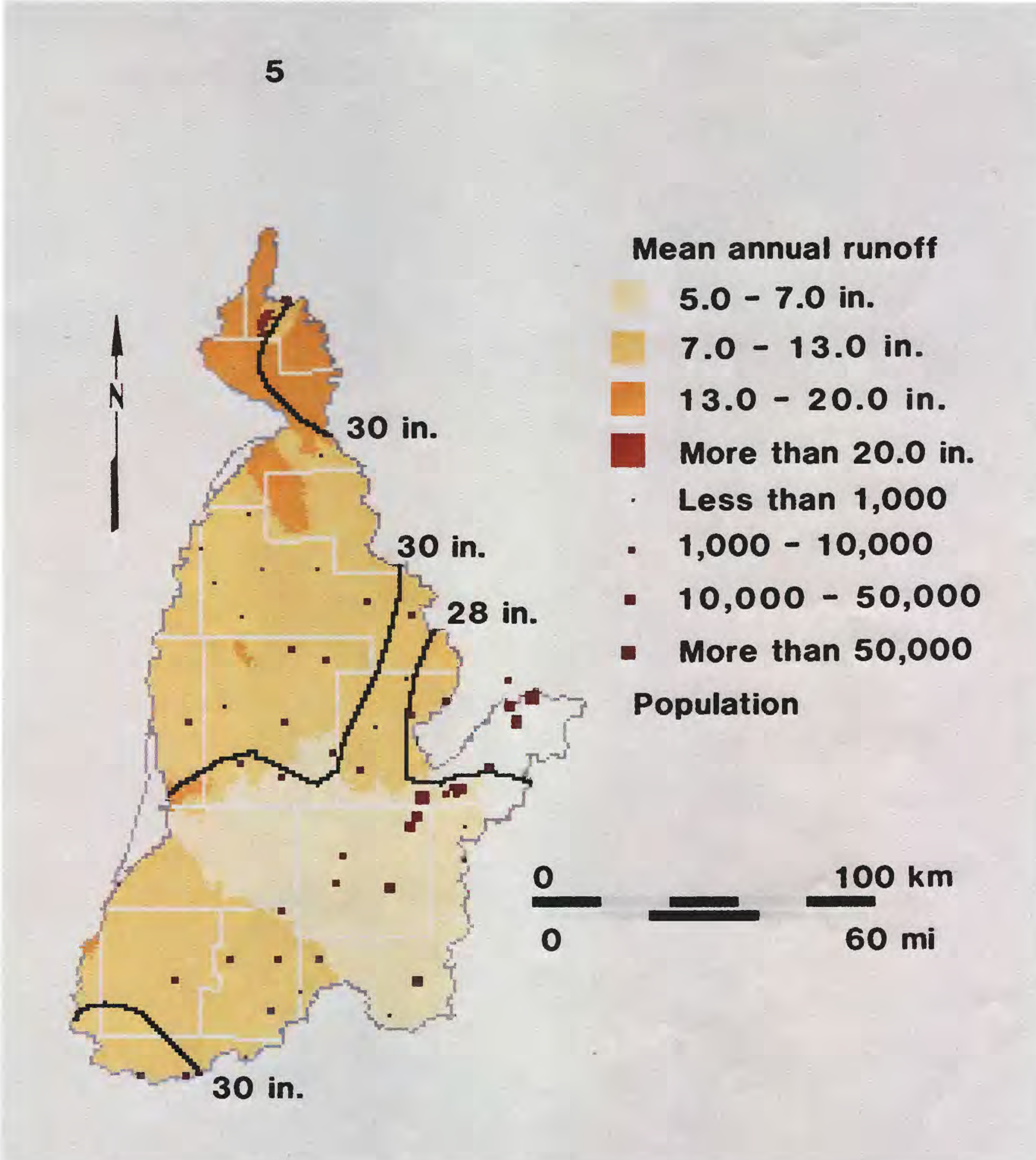


Figure 5.--Water availability is addressed by contours of mean annual precipitation, color coded basins showing mean annual runoff, and squares that are sized to show the population classes of cities and towns. Runoff information is not available in the easternmost part of the study area.

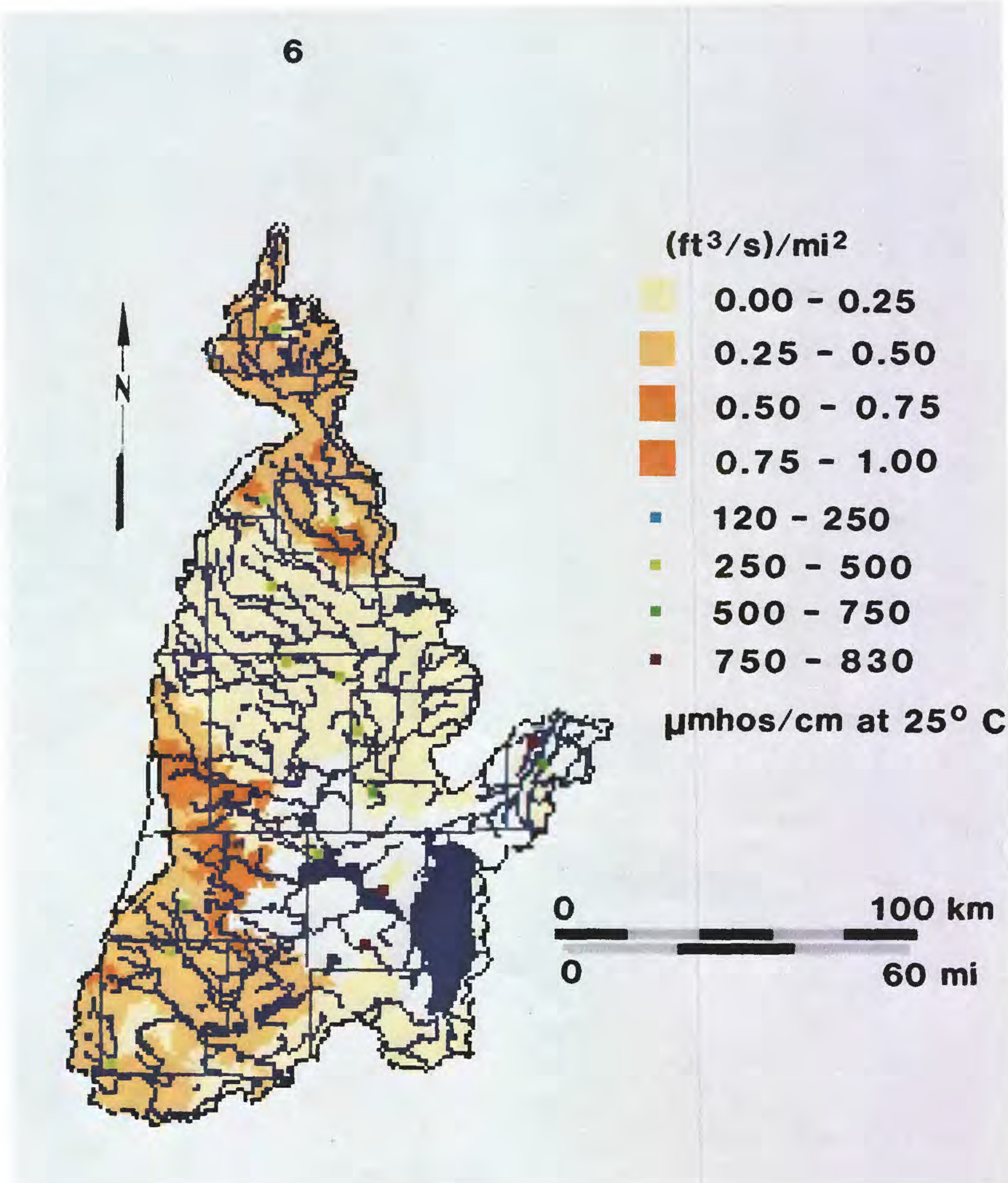


Figure 6.--Additional surface-water information is shown by color coded polygons for the 7-day, 10-year low flows and by color coded points representing the specific conductance of streams at base flows.

The data sets for the first illustration (fig. 5) are surface-water divide, ground-water divide, county lines, mean annual precipitation, mean annual runoff, and populations of cities and towns. The basin divides and county lines are included as base-map information for viewer location and orientation purposes. Mean annual precipitation shows total water income to the basin. Mean annual runoff shows the total amount of water available for man's use after evapotranspiration. The locations and populations of cities and towns are included, because these are centers of water use and water demand.

The locations and populations of cities and towns were acquired as a digital file; a point data set, coded by population, was created from this file. A decision was made to show four classes of population, and the points were sized to represent these classes (fig. 5). Boundary lines from the basin divide and county polygons are shown on the illustration. Polygon bounds are also shown for mean annual precipitation; in this case, the bounds represent and are labeled as contours. The basins for which mean annual runoff information is available in the Fox-Wolf drainage system are shown as color coded polygons. All basins for which any streamflow information is available were digitized once, and each basin was given a unique label. The separate data sets for figures 5 and 6 were then created by digital copy and coding procedures. Mean annual runoff values were acquired as a digital file, referenced by downstream gaging station number and by latitude and longitude location. A decision was next made to show mean annual runoff in four classes on the illustration. The final step was to code each basin with the digital value for its class and to select appropriate colors for display of the classes.

A number of decisions must be made for the display of several types of information on a single illustration. An illustration must be thoroughly planned or obtained by considerable trial-and-error experimentation. On figure 5, for example, there are five polygon data sets and multiple classes for three of these data sets; if an attempt were made to show all of these data sets as polygons, there would be at least hundreds of unique classes in the composite; the illustration would be nearly impossible to interpret. This is the reason for the decision to show only mean annual runoff classes as color coded polygons.

Other decisions were required for the determination of precedence for each data set in figure 5. Some of these decisions are obvious: for maximum visibility of all data, points take precedence over lines, and lines take precedence over polygons. Thus, precipitation contours are shown within the mean annual runoff polygons. The other decisions of this type generally are based on the relative importance of each data set. Thus precipitation contours are continuous, whereas county lines are broken at intersections of these features. The final decision was the selection of colors for all features on the illustration. Bright colors were generally selected for the more important features, especially those that are less obvious on the composite illustration. A decision also was made to show higher amplitude values by darker colors on the composite illustrations but by lighter gray tones on the single data sets (Appendix).

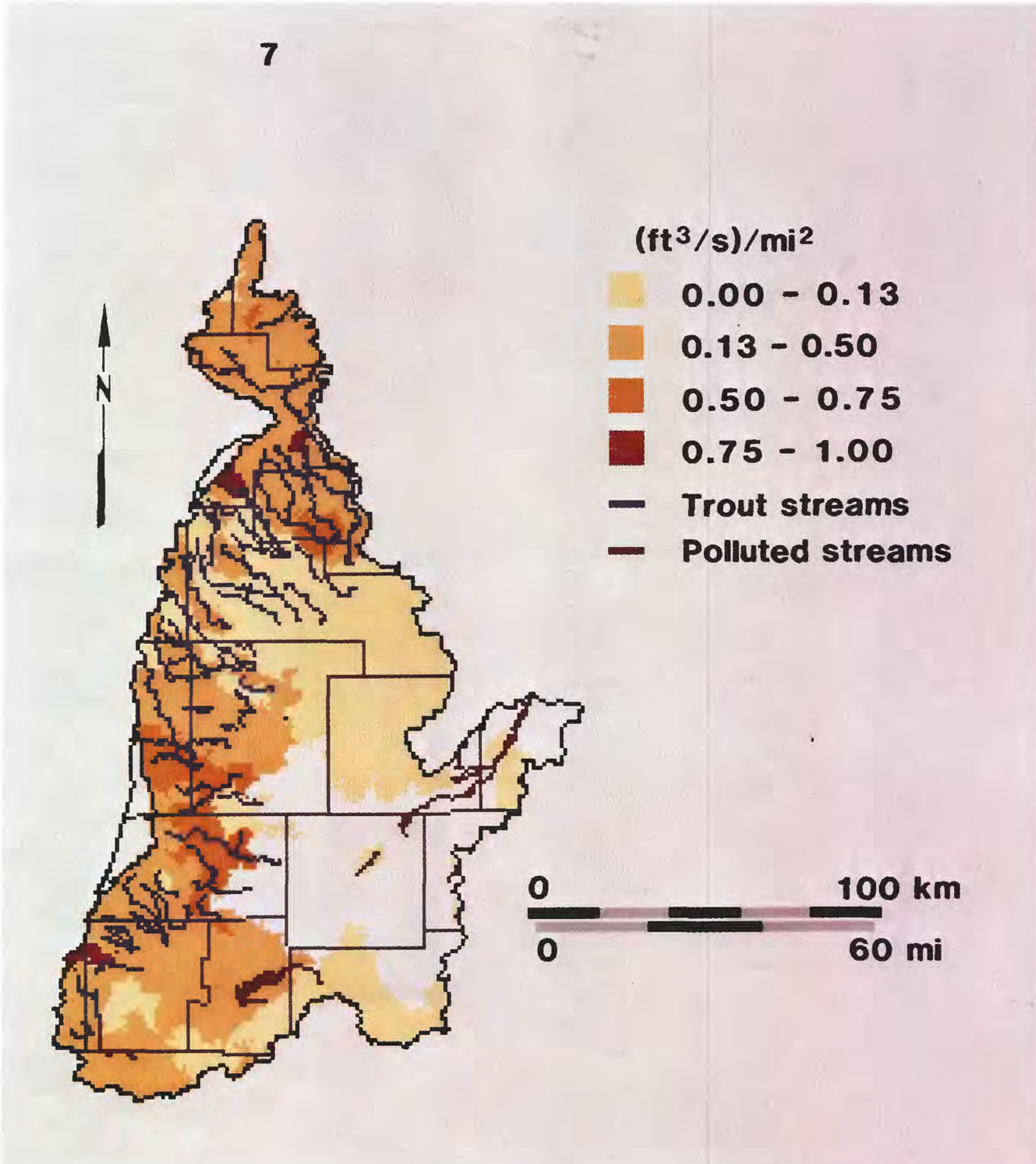


Figure 7.--Most trout steams are in basins that have relatively large low flows. Most polluted streams and lakes are in the downstream part of the study area.

The second illustration (fig. 6) that addresses the question consists of six data sets: surface-water divide, ground-water divide, streams, lakes, low streamflows, and specific conductance of base flows at gaging stations. The basin divide and the drainage network are included for viewer location and orientation purposes. Both low flow classes and classes of the specific conductance of base flows are shown so that an index of the chemical quality of surface water can be compared with an index of the amount of available surface water.

Values for the 7-day, 10-year low flows were obtained for 91 basins from a map overlay supplied by the Wisconsin District, WRD. A decision was made to show these values in four classes on the illustration. Then the polygon for each basin was coded with the digital value for its class. Specific conductance values were acquired as a digital file, referenced to a gaging station number and to a latitude and longitude location. Four classes of specific conductance were selected, and each gaging station point was coded with the digital value for its class.

Both low flow index and specific conductance (fig. 6) represent values measured at the gaging stations. It would have been possible to display classes of both values at the points by showing, for example, a size range for low flow classes and different colors for specific conductance classes. Instead, a decision was made that it would be informative to show the basin areas that produce the low flow classes; the basin polygons were color coded for this purpose. Gaging station points were given precedence over basin polygons and all other data sets for display on the composite illustration. Both lakes and the basins are polygons, but the lakes are smaller and were given precedence on the illustration.

What Is the Natural Quality of Surface Water and How Is It Affected by Man?

Four illustrations were prepared to respond to this question. One illustration (fig. 6) shows the specific conductance of base flows at gaging stations and has been discussed previously. Another (fig. 7) shows the relationships of low streamflow index classes to designated trout steams and to polluted streams and lakes. The third illustration (fig. 8) shows landfills that are within 1 km of a stream, and the last illustration (fig. 9) shows the landfills and liquid waste disposal sites that are within flood-prone areas.

The illustration that compares low streamflows with significant lakes and stream reaches (fig. 7) consists of seven data sets, including the surface-water divide, ground-water divide, county lines, low flow index classes, trout streams, polluted streams, and polluted lakes. The basin divide and county lines are included as base map information. Low flow index classes are shown because trout generally are found in streams that have relatively large rates of base flow. Designations for the trout streams and polluted streams and lakes were assigned by state or local government agencies (Wisconsin Department of Natural Resources, 1982).

Most of the data sets on figure 7 have been discussed previously. The trout streams (lines), polluted streams (lines), and polluted lakes

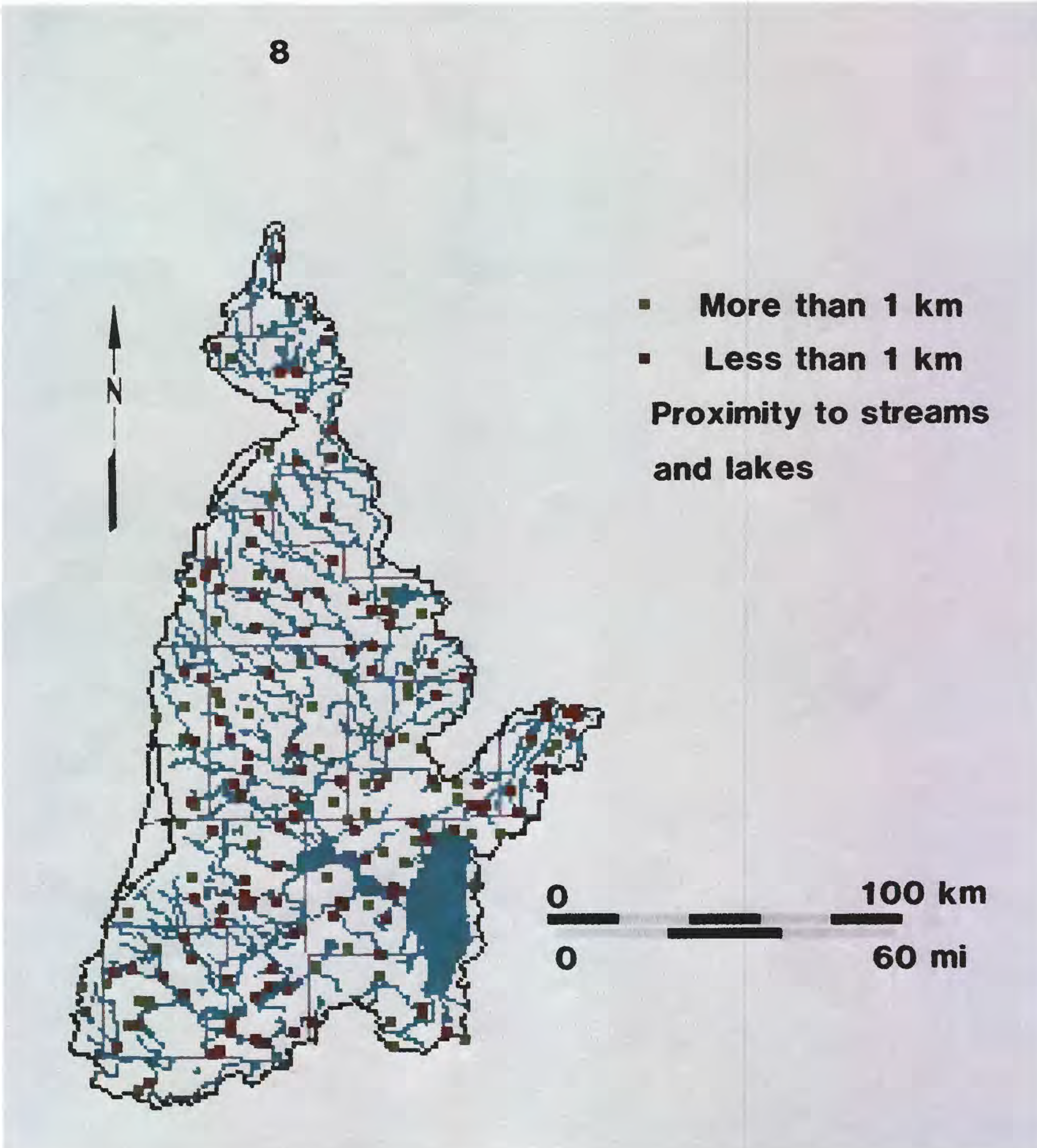


Figure 8.--Landfills within 1 km of a stream or lake are color coded red, whereas landfills farther away are shown in green. Landfills near streams and lakes are more likely to affect the chemical quality of surface waters than landfills that are farther away.

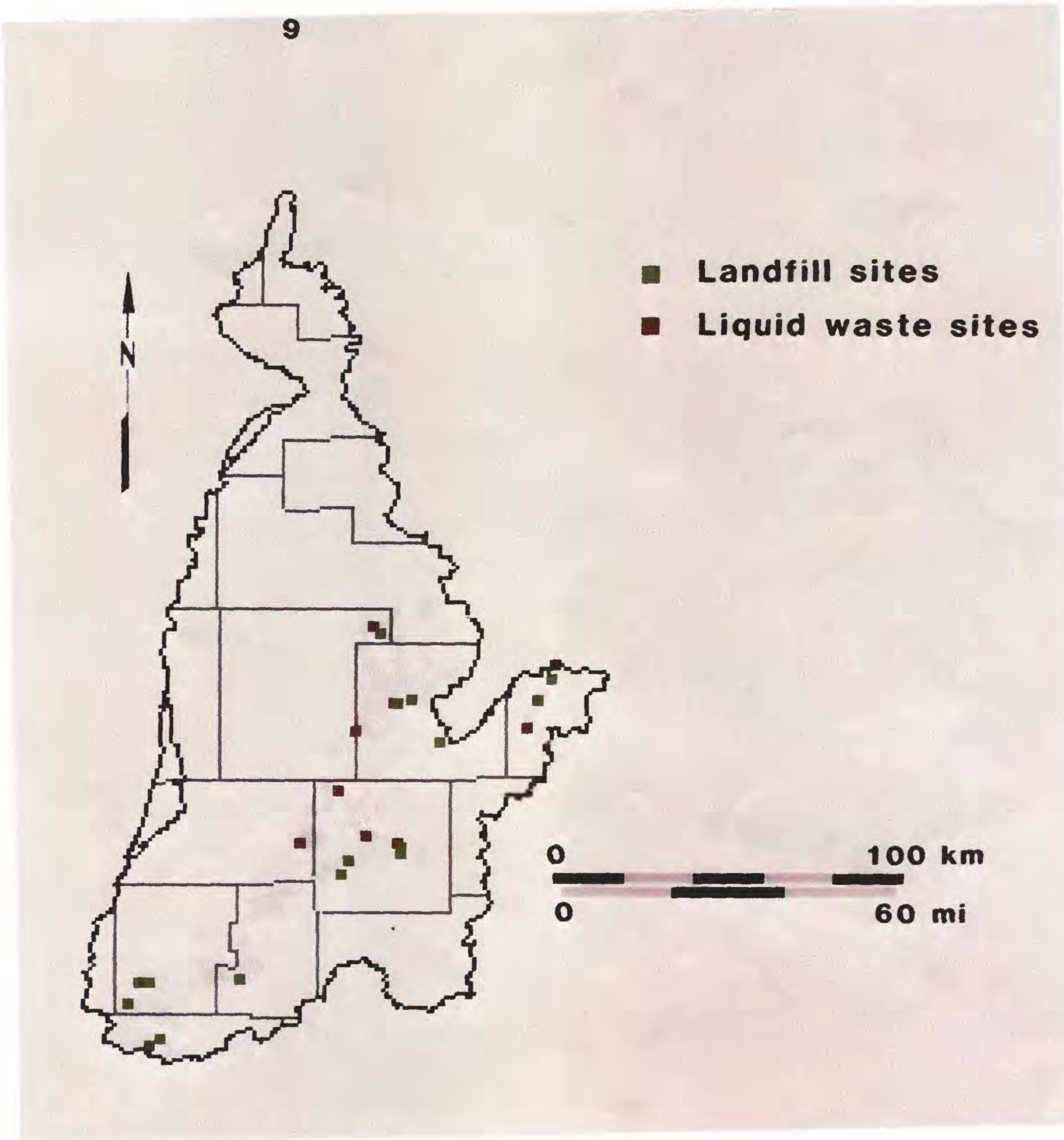


Figure 9.--Some of the landfills (green) and liquid waste disposal sites (red) occur within flood-prone areas. If waste-disposal sites are flooded, surface-water pollution may occur.

(polygons) were digitized as separate data sets from 1:500,000 scale map overlays. Different bounds were used to show four classes of 7-day, 10-year low flows on figure 7; the low flow index classes on figure 7 have a more obvious relationship to the location of trout streams than do the classes on figure 6. The highest precedence in the composite illustration was given to trout streams and polluted streams and lakes; county lines have a higher precedence than the low flow index classes.

The next illustration (fig. 8) shows a classification of landfills into sites within 1 km of a stream and sites that are farther away. The data sets consist of surface-water divide, ground-water divide, streams, lakes, county lines, and landfill sites. The landfills were digitized as points and were labeled by class of disposal volume. For the purpose of this illustration, landfills were treated as points with a single class label.

Streams and lakes are shown on the composite illustration (fig. 8) and were also used with a proximity function to create a polygon that includes all areas within 1 km of these features. A decision was made that landfills within this zone are more likely to affect the chemical quality of surface water than are landfills farther away. This polygon was used to classify and separately color code the landfills within and outside of the 1 km distance.

The composite illustration (fig. 8) is affected by both the accuracy of the original maps and by the resolution of the data sets. All streams and lakes are not shown on 1:500,000 scale maps, and the positions of these features may be somewhat generalized. Similarly, landfill locations may not be shown accurately on maps at this scale. In addition, all landfills are represented by grid cells that are 1 km by 1 km in size in the digital data set, and stream channels are represented as having a minimum width of 1 km in a second data set. The computer classification of landfills produced an accurate result but one that may not be correct in all cases. Nevertheless, an information product of this type can be used to indicate sites that need further study or checking in the field.

The other illustration (fig. 9) shows all landfills and liquid waste disposal sites that are within flood-prone areas. If waste disposal sites are flooded, surface-water pollution may occur. The data sets consist of surface-water divide, ground-water divide, county lines, landfills, liquid waste disposal sites, and flood-prone areas. Flood-prone areas were digitized as polygons from 1:500,000 scale maps. Liquid waste disposal sites were digitized as points and labeled by type and status of the disposal site. For the purpose of this illustration, all liquid waste disposal sites were treated as points with a single class label.

The results of data processing in figure 9 are affected by the accuracy of the 1:500,000 scale source materials and the accuracy with which locations and boundaries are shown by 1-km grid cells in the digital data sets. Most of the waste disposal sites are close to boundary lines of the flood-prone areas, and conclusions about the susceptibility of these sites to flooding should not be based entirely on the results of digital processing of this type.

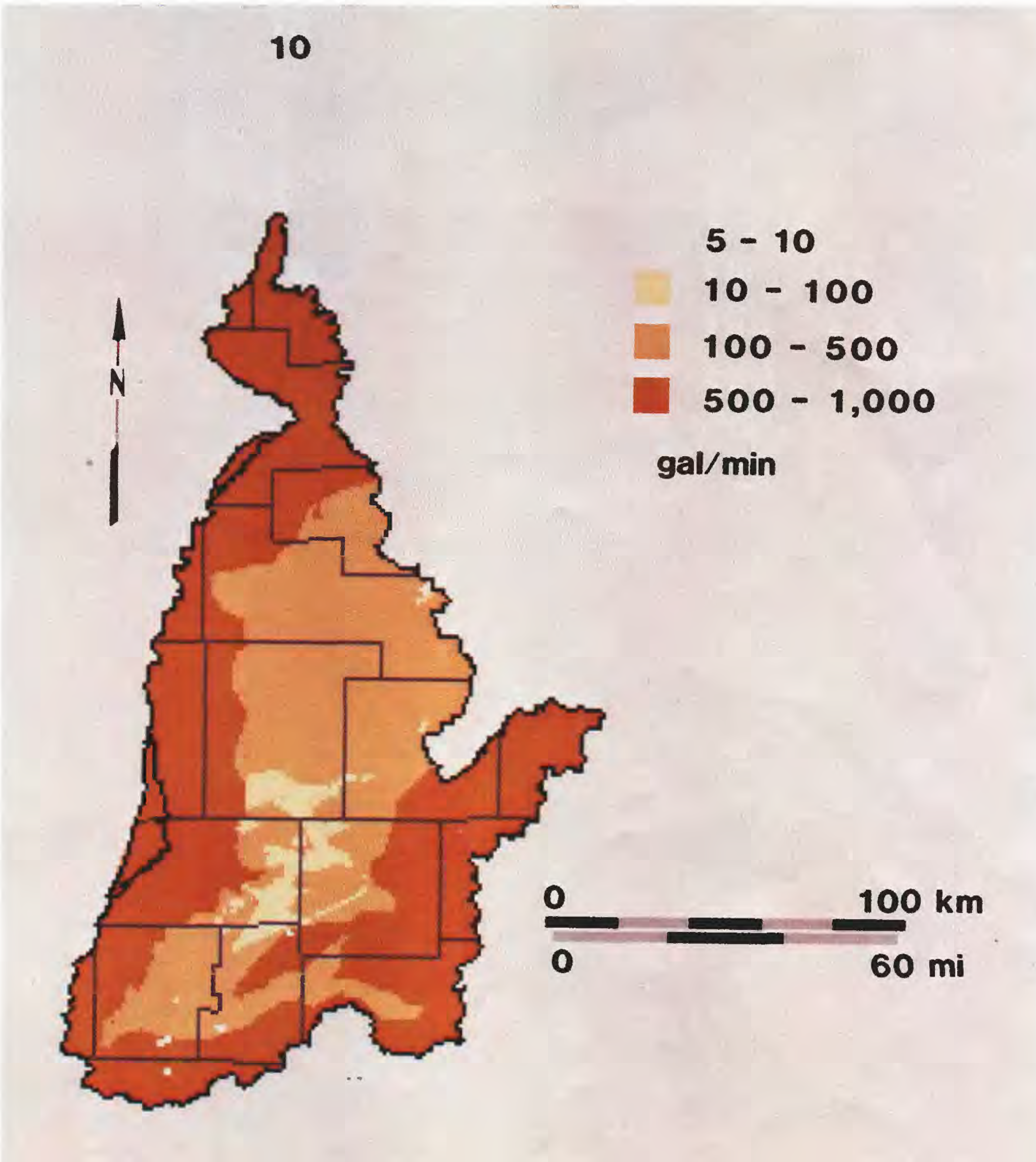


Figure 10.--The maximum well yield from the best aquifer in any area generally is from 10 gal/min to 1,000 gal/min. Wells in the white areas of the map generally yield less than 10 gal/min.

What Is the Availability of Ground Water in the Basin?

Ground-water availability is shown by two illustrations. The first (fig. 10) shows the maximum amount of water that a well may produce from any one of the three aquifers in the Fox-Wolf basin. The Cambro-Ordovician sandstone aquifer occurs only in the southern half of the basin; the Silurian dolomite aquifer underlies only the southeastern edge of the basin; and the glacial sand and gravel aquifer occurs in the entire basin. The maximum well yield from the best aquifer in any area is shown in figure 10, but the aquifer is not identified. Also, some wells in these areas have lower yields.

The six data sets used for the illustration (fig. 10) are surface-water divide, ground-water divide, county lines, well yields and extent of the Cambro-Ordovician sandstone aquifer, well yield and extent of the Silurian dolomite aquifer, and well yields in the glacial sand and gravel aquifer. Maps showing well-yield ranges and extents of the three aquifers were supplied by the Wisconsin District, WRD. These maps were digitized as polygons. A maximum algorithm (which finds the maximum value in two or more data sets) was then used to produce the color coded polygons on the illustration.

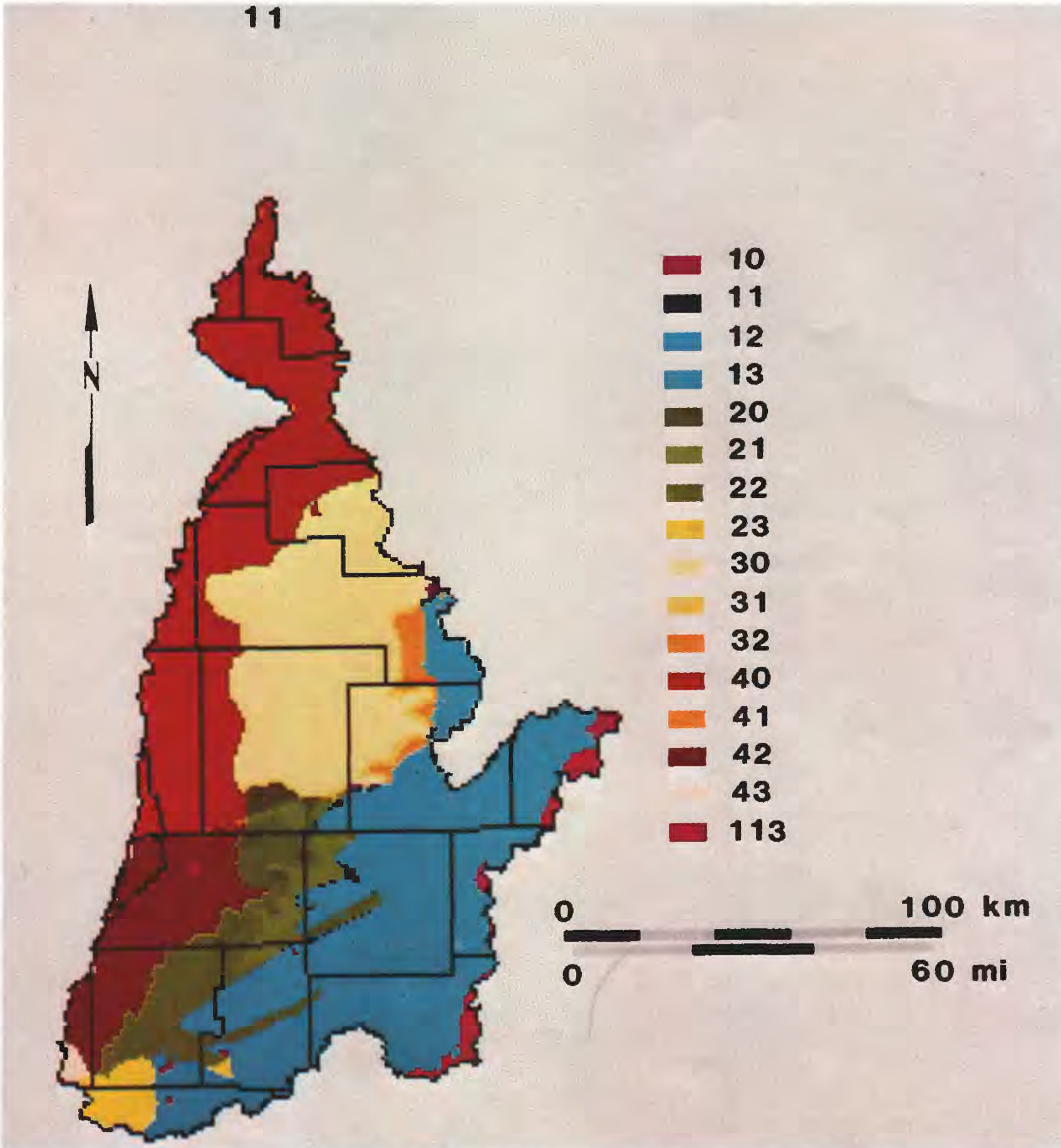
The second illustration (fig. 11) is the most complicated in this report; it shows the maximum amount of water that can be pumped from a well at any point, and it identifies the aquifer or combination of aquifers. The data sets are the same as those in figure 10, but a different algorithm was used to produce the composite illustration. The well yield polygons for each aquifer were first digitally coded. Well yield polygons for the sandstone aquifer were coded by units (ones), polygons for the sand and gravel aquifer were coded by tens, and the dolomite aquifer was coded by hundreds. An additive algorithm was then used to total the polygons in these three data sets. The final step was the selection of colors for the sums in the resulting polygons.

The interpretation of figure 11 requires referencing a color on the illustration to a number in the explanation and then finding the key to this number in the table. For example, violet on the illustration is 113 in the explanation; the table then shows that this number (100 plus 10 plus 3) represents well yields of 10-100 gal/min. (gallons per minute) from the dolomite, 5-10 gal/min. from the sand and gravel, and 500-1,000 gal/min. from the sandstone.

The responses (figs. 10-11) to a question on ground-water availability show (1) the same data sets can be used to show different amounts of information on a composite illustration, and (2) interpretation complexity increases with the amount of information on a single illustration.

What Are the Water Problems in the Sandstone Aquifer?

Two illustrations are used to address this question. The first (fig. 12) shows water-level drawdowns in wells, caused mainly by industrial water use near Green Bay. The second (fig. 13) shows classes of dissolved



	Well Yield, in Gal/Min			
	5-10	10-100	100-500	500-1,000
Sandstone	n.a.	1	2	3
Sand and Gravel	10	20	30	40
Dolomite	n.a.	100	n.a.	n.a.

Figure 11.--The maximum well yield from any aquifer at any point is shown by the colored polygons on the map, the number for each color in the explanation, and the key to the numbers in the table, above. A description of the interpretation procedure and an interpretation example are in the text. The symbol n.a. in the table means not applicable.

solids content of water from the aquifer. Both illustrations are relatively simple, and they could have been combined. There are at least three different options for showing the values or classes of two different data sets as polygons: (1) the combined polygons can be separately color coded (fig. 11, for example); (2) colors can be used to code one data set and shading or patterns to code the other; and (3) values or classes in the second data set can be shown in the third dimension (as a stereo pair or as a pedestal diagram). The decision to show figures 12 and 13 separately was based mainly on the fact that water-level drawdowns and water quality are two separate, unrelated problems in this aquifer. Also, a combined illustration would have been more difficult to interpret, and the additional complexity did not seem to be warranted. Many decisions of this type will be necessary during the processing of national data sets.

The data sets selected to show the effects of industrial pumpage are surface-water divide, ground-water divide, county lines, total industrial and commercial ground-water use by county, and water-level drawdown in wells. The water-level drawdowns were digitized as polygons from a 1:500,000 scale map overlay. The county water-use data set was produced by digital copy and coding procedures from county polygons that were digitized previously. Water-use polygons are color coded on the composite illustration, and boundaries of the drawdown polygons are shown as contours.

Class intervals for all polygon data in figure 12 were selected by the Wisconsin District, WRD, and the polygons could not be reclassified later by digital processing. Also there were not separate values for ground-water pumpage from only the sandstone aquifer. These data selections and the procedure seemed logical at the time, but the composite illustration is not optimum for showing the relationships between industrial water use and water-level drawdown near Green Bay. This result demonstrates that processing objectives must be known to ensure the best possible decisions on selection and handling of data sets.

The second illustration (fig. 13) shows areas where water from the sandstone aquifer exceeds 500 mg/L and 1,000 mg/L of total dissolved solids. The data sets are surface-water divide, ground-water divide, county lines, and three classes of total dissolved solids. The latter data set was digitized as polygons from a 1:500,000 scale map overlay.

In Which Area May Ground-Water Pollution Have a Serious Impact?

Pollution of a surficial aquifer is often caused by the leaching of undesirable chemical constituents from landfills. These leachates are more likely to affect the chemical quality of ground water in areas of coarse-grained materials; these are the same areas that have the highest yields of water from shallow wells. The data sets selected to address this question are surface-water divide, ground-water divide, county lines, well-yield classes from the glacial sand and gravel aquifer, and landfills. Well-yield classes were digitized as polygons from 1:500,000 scale map

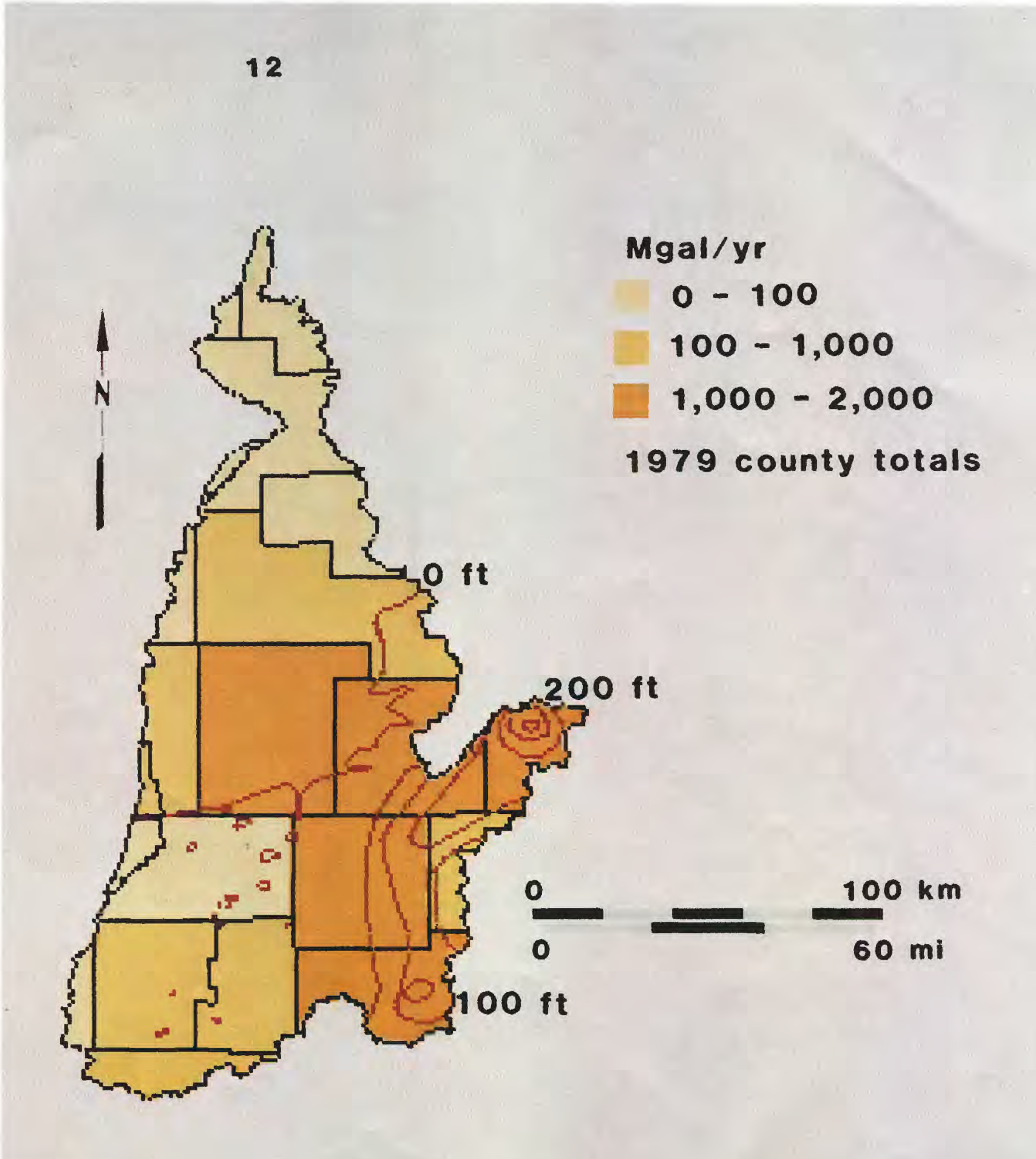


Figure 12.--Colored polygons show total ground-water pumpage for industrial and commercial use by county. The contours show the drawdown (0-300 feet, in 50-foot intervals) of water levels in wells tapping the sandstone aquifer; the drawdown is caused mainly by industrial water use near the city of Green Bay.

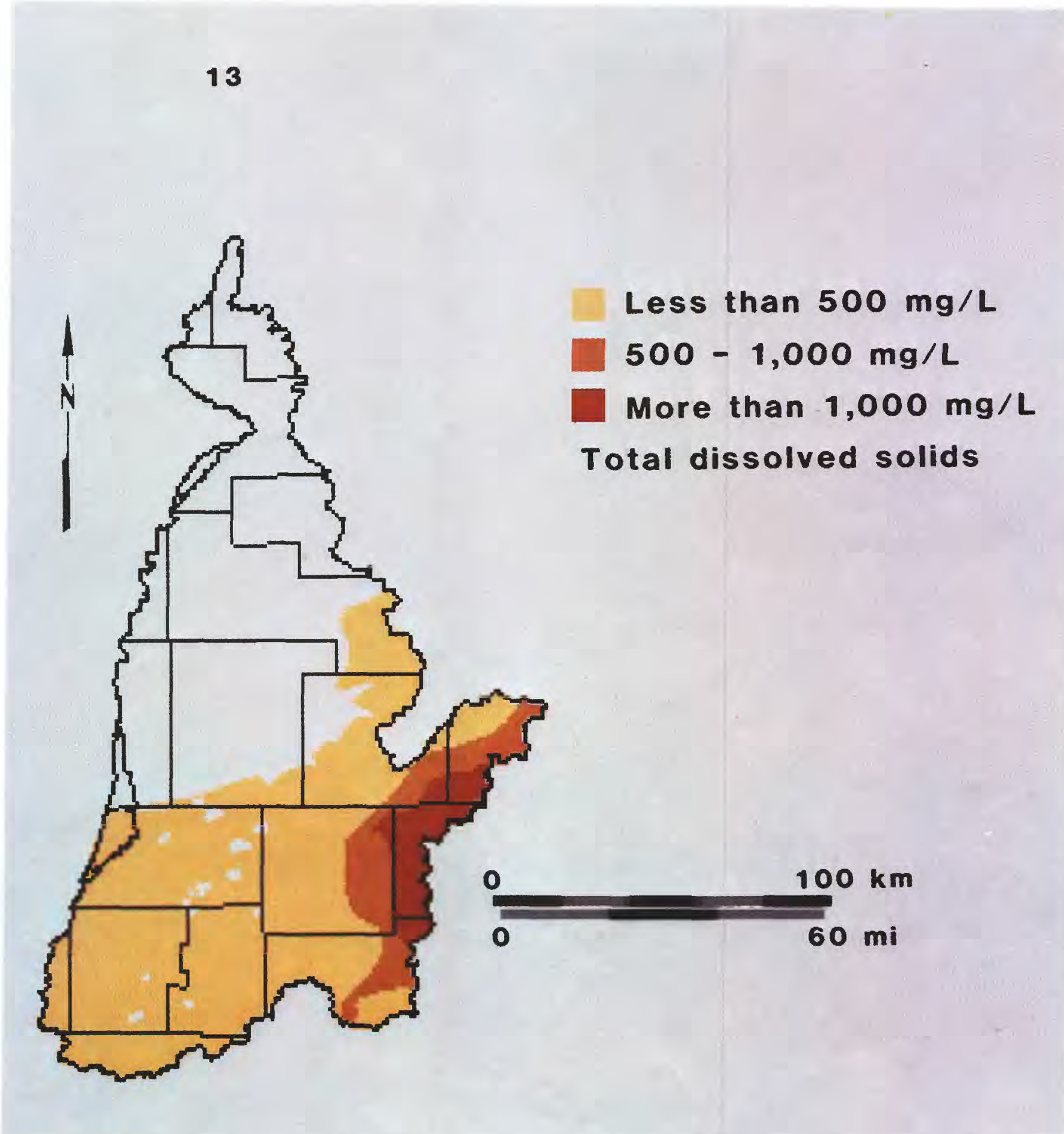


Figure 13.--Three color coded polygons show the overall chemical quality of water from the sandstone aquifer. The aquifer does not occur in the white areas on the map.

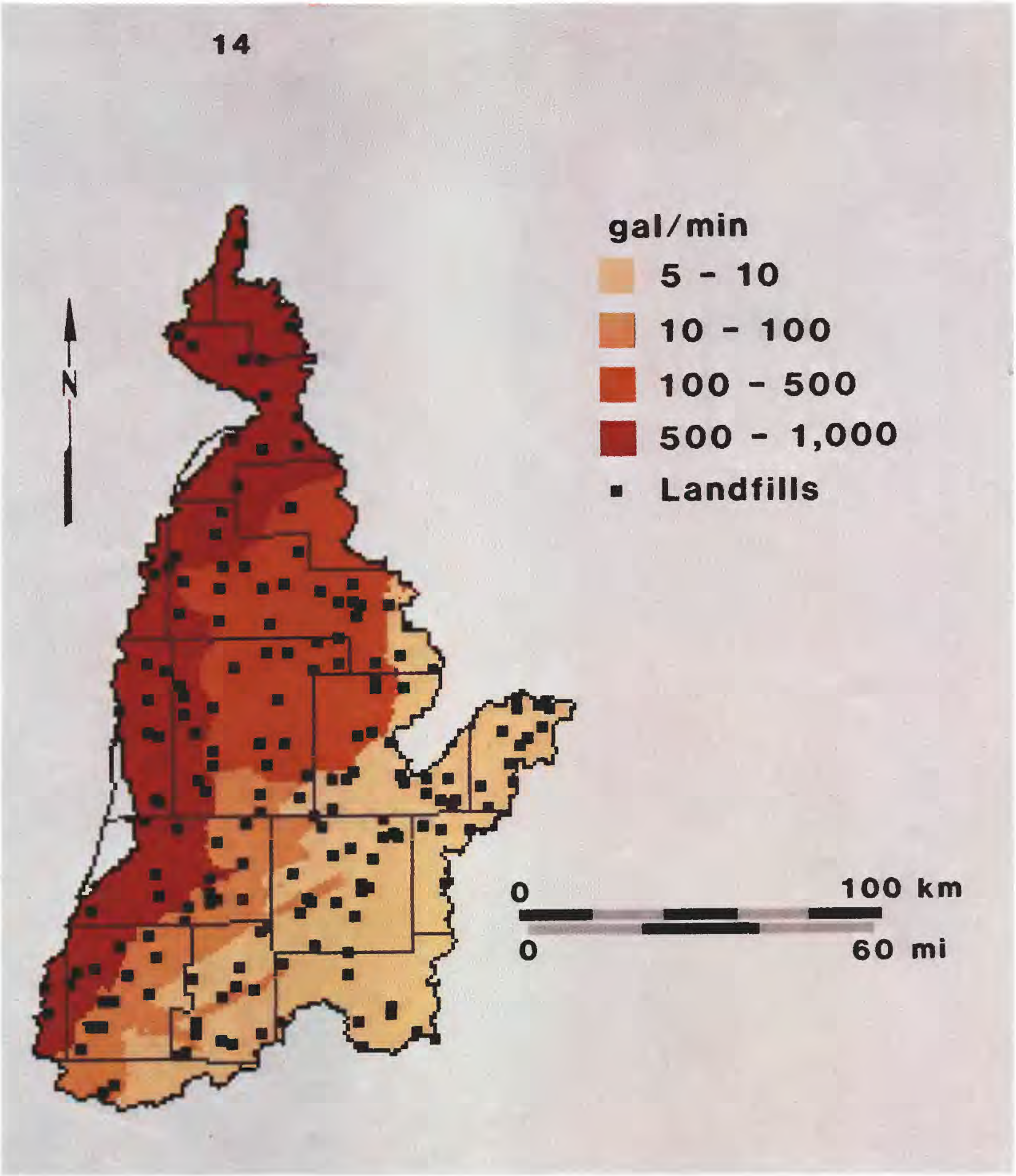


Figure 14.--Landfills are fairly evenly distributed in the basin. Leachates from these waste disposal sites are more likely to affect the chemical quality of ground water in areas of coarse-grained materials; these are the same areas that have the highest yields of water from shallow wells.

overlays and are shown as color coded polygons on the composite illustration (fig. 14). Landfills were enlarged for visibility by using a proximity algorithm.

CONCLUSIONS

The data sets for the Fox-Wolf drainage basin were digitized in vector form, but all data processing and product generation represent a raster (grid cell) format at a resolution of 1 km by 1 km. The same specifications and procedures are adequate for national water information needs. The very fine detail that is possible with vector data storage is unnecessary for national data products, because only generalized maplike products are possible at scales of 1:2,000,000 to 1:25,000,000.

All data sets were compiled digitally before processing objectives were selected; this proved to be inefficient. Only 25 of the 50 data sets were used for data analysis. Some early subjective decisions on data selection, format, and classification also later proved to constrain the results that could be obtained by data analysis. Nevertheless, the results of this study are typical of those that can be obtained by digital mapping technology. Specifically, a relatively few digital processing techniques can be used to create many distinctive maplike products. The results of a national summary can be optimized, and costs can be minimized, by prior decisions on the number, type, and content of output products.

Most of the problems encountered in the Fox-Wolf study were minor and easily solved. However, some decisions on the selection and digitizing of data sets were irreversible; they could only have been changed by returning to the digitizing stage of work. Line data sets of the same feature may not exactly overlie, for example, unless they are digitized as separate classes of the same feature. Similarly, polygon bounds constitute a single class that may be difficult or impossible to reclassify (separate into different classes) at a later stage of processing. Finally, class bounds generally cannot be changed after digitizing in order to include more detail.

It is anticipated that the data analysis requirements for a national water resources summary will be similar to those for the Fox-Wolf basin, because the available data will have already been processed to the point where they have intrinsic importance for a summary of this type. Thus, a national summary can probably be accomplished with relatively simple forms of parameter evaluation, areal stratification, and data merging and integration.

A digital information system in map form is expensive to prepare and compile. The most effective uses and maximum benefits from an information system are obtained by exploitation of both the spatial and amplitude relationships among multiple data sets. Questions and processing objectives that require only one or two data sets are, by themselves, not cost-effective uses of the system. Neither, by themselves, are procedures that require graphs, equations, or statistics, and that exploit only the amplitude relationships at points in the data sets.

There are eight main benefits of a digital water resources information system:

1. More complete and more objective resource information can be obtained by data analysis. Human errors of omission are minimized, and any human bias has less effect on results. All results are reproducible.

2. The data sets are permanently stored in a form that is easily updated and revised.

3. Digital data sets can be provided to other agencies in a readily usable form for other types of resource assessment or management.

4. There are many available options for formatting, data processing, and display; these options are easily implemented.

5. Measures that would be impractical with conventional procedures can be derived from digital data sets (neighborhood slope and relief, for example). Other measures can be obtained faster than with conventional procedures (area measurements, for example).

6. Lower grade personnel can be used for digitizing and simple data processing. Skilled personnel can concentrate on planning, data analysis, and evaluation of results.

7. Both the spatial distribution and amplitude values (or labels) of features in the data sets can be used to obtain useful information. Some information can be obtained from single data sets, but other information is obtained by areal stratification, data merging, data integration, and modeling of multiple data sets.

8. Registered digital data sets offer an improved capability for the study of complex hydrologic problems; the data are in a format that is suitable for most distributed parameter models.

The storage and processing of digital maps cannot accomplish all objectives of a national water resources summary. Some factors, relationships, and problems are best described in narrative form. Also, some data types, relationships, and totals are more understandable when shown in graphic, tabular, or statistical form. Spatial processing and display mainly offer a highly useful, objective method of showing facts, relationships, and results in readily understandable maplike forms.

SYSTEM DESIGN CONSIDERATIONS

A national water resources summary should serve several purposes and should provide water information at several scales and levels of detail. A national information system must be useful for large regional assessments and for policy decisions at the federal level. In addition, national data sets should be useful for the study of any area that is the size of a county or larger. It is not feasible today to produce a national water resources information system that meets all of these requirements. Countrywide coverage of digital cartographic and topographic data are not yet available at the desired scale and resolution (level of detail). Also, the software necessary for efficient storage and processing of very high resolution data is not yet perfected. However, a national water summary can be designed and produced to meet information needs at the national level, as well as some needs at the regional level. Planning and compilation can also begin in some areas to meet other water information needs at the river-basin, state, and county levels.

Since the immediate needs for national water information represent different data and system requirements than those of counties, the procedures for compiling each type of information system are considered separately. The system requirements to produce national data sets and national information products include hardware and software for (1) digitizing maps in vector format; (2) reading and entering digital point-data files, vector data sets, and raster data sets; (3) geometric registration to a selected map projection, conversion of all data sets to a raster format, and resampling to a common grid-cell size; (4) surface generation; (5) on-line storage of at least several complete data sets; (6) a variety of data processing and derivation options (similar to those in table 5-8); (7) display and evaluation of data sets and processing results; and (8) generation of maplike products on film or paper.

Hardware or software limitations can constrain the compilation and analysis of national data sets at any stage of system use. One laser-beam film recorder, for example, uses a 38 µm spot size to print a maximum of 28 million grid cells (in a 5,300 by 5,300 array) on film. This fact restricts national data sets to a resolution of 1 km by 1 km, thereby producing a maplike image that is about 4,700 by 2,800 grid cells in size. Similarly, color television displays, which are used for interactive evaluation of data sets and processing results, are generally limited to an array of 512 by 512 grid cells at one time. This means that only a small part of each data set can be examined at full resolution, and that repetitive displays may be necessary for editing and evaluation of data sets. If a national water resources information system is compiled in the near future, all plans, work, and products will have to fit these or similar system constraints. The only other known problems are the efficiency of some software algorithms and some limits on the size of data arrays; these problems should be easily solved by software programming on an as-needed basis. All available hardware and software systems have limitations, and it is important to understand the resulting constraints on data processing.

Some presently available digital data are satisfactory for national water information needs and can be used to begin the compilation of a national water resources information system. Digital cartographic data from the Geological Survey's 1:2,000,000 scale digital line graph data base include data sets that show political boundaries, federal land ownership, drainage network, transportation network, and lakes and reservoirs. Selected data sets from this group can be used as cartographic base maps. Digital topographic data in a 30-arc-second raster format are available from the U.S. National Geophysical Data Center in Boulder, Colorado; these data can be easily resampled to a 1 km resolution. Much of the necessary hydrologic information is available as digital point data in the WATSTORE, Basin Characteristics, Water Use, and Ground-Water Site Inventory files. Other hydrologic information is available on small-scale maps, which can be digitized. The digitizing of large-scale hydrologic maps could be a bottleneck in the compilation process, however; only the most important of these data sets should be selected for the national information system.

During the preparation of an information system to meet national water information needs, consideration should be given to plans for a more detailed data base to satisfy information needs at river basin, state, and county levels. A data resolution of 25 m to 100 m is required for these purposes. In raster format, a 100 m resolution would require a grid-cell array of about 47,000 by 28,000 for the conterminous 48 states. State-of-the-art hardware systems and many software procedures are adequate for the handling and analysis of this size of data base. Vector data storage is more efficient than raster storage, and plans could be made for the compilation, storage, and processing of all data in vector format. The limitations of this approach are (1) blocks of vector data can be difficult to mosaic and merge, (2) surface data sets cannot be stored and processed at full resolution in vector form, and (3) vector data must be converted to raster for output as maplike images; this is a faster and more efficient method than is the use of a plotter to draw line maps.

An alternative to vector storage and processing is the use of data compression to minimize storage requirements for raster data. Some compression is obtained by minimizing bit-size requirements (and number of classes) in the data sets. Other methods reduce the number of cells by omitting most cells in groups with identical values. If there were no more than 63 classes of information in any single data set, for example, 6-bit digital data could be used for class coding; a data set of this type should fit easily on a 600-megabyte hard disk (for on-line processing) or on a computer compatible tape (for off-line storage).

Maplike image products can be easily output from very large data bases. One large-format film writer has the capability to record a data array slightly larger than 40,000 by 40,000 grid cells. This means that an output product could be generated at full resolution (with 100 m by 100 m grid cells) for any area a little smaller than the conterminous 48 states.

A near ideal amount of detail for digital base maps at 25 m to 100 m resolution is represented by the cartography and topography on maps in the Geological Survey's new 1:100,000 scale series. Unfortunately, very few of these maps have been completed, and there are no plans at present to offer this information in digital form. Digital topographic, but not cartographic data are available from the 1:250,000 scale map series. Both digital cartographic and topographic data are available for some maps in the 1:24,000 scale series, but there is not nationwide coverage; these data represent about 30 m resolution. It may be necessary to digitize base-map data from either the 1:250,000 scale map series or from 1:500,000 scale state base maps for a more detailed national information system.

In summary, it is feasible today to produce a water resources information system to meet national needs for water information; a data set resolution of 1 km is adequate for this purpose. State-of-the-art hardware and software systems are adequate or near adequate today for national data sets with 100 m resolution, but cartographic base maps would have to be digitized. A data set resolution of 25 m may be feasible in 5-10 years, with anticipated advances in hardware and software technology. Vector format storage and processing are more efficient than raster format for very high resolution data, but raster output as maplike images is faster and cheaper than plotter-drawn line maps.

RECOMMENDATIONS

The following recommendations are based on results obtained in the Fox-Wolf prototype study and on experience gained by the authors in the compilation and use of digital information systems. Some recommendations are direct conclusions of the prototype study, but others are beyond the scope of this study and are not documented.

1. A digital water resources information system should be developed by the Geological Survey, because a system of this type can form a useful part of a national water summary. Some water information is best analyzed and displayed in map form, and digital maps can be easily updated or revised.

2. Facilities for the design and compilation of a national water resources information system could be located anywhere. However, the Headquarters Office of WRD has prime responsibility for national water information requests. For this reason, a single, centralized facility at the Headquarters Office should include the required hardware, software, display devices, and output product devices, as required for analysis of the data and for documentation of all results.

3. Cartographic base map data should be obtained from U.S. GeoData files, digital line graphs from 1:2,000,000 scale maps, produced by National Mapping Division of the Geological Survey. These data sets should be supplemented with digitized and registered data showing hydrologic units (to the cataloging unit level) and water resources planning subregions. All additional water resources data should be geographically referenced and registered to this data base.

4. A search should be made for all national and regional maps of water resources and directly related data that meet criteria for a current summary of national water resources. Selected maps should be digitized and registered for inclusion in the information system.

5. Digitized maps and other data should be in as basic (primitive) a form as possible, because this form offers maximum flexibility for digital data processing, as required by any selected objectives.

6. All individual digital data sets and composite digital products (obtained as a result of processing) should be made available for public distribution in a manner similar to that of Geological Survey publicatons.

The development, handling, processing, and output-product capabilities of digital information systems are changing rapidly. Present methods are adequate for many purposes of a water resources summarization, but hardware, software, and display limitations impose constraints on many methods. Further research is needed to automate some data handling procedures and to address data analysis capabilities for time series and trends. There are two specific recommendations for further research.

1. Information system software presently does not have a good capability to describe and characterize lines, such as stream reaches. Research is needed to improve characterization, analysis, and display capabilities for lines and line segments.

2. Automated techniques should be developed as an interface between the digital data files, such as Watstore, of the Geological Survey and the spatially positioned and registered data sets in a water resources information system. An interface of this type should allow an easy interchange of data from either format to the other.

Digital storage of basic water data is operational and is being improved at present. However, much water resources information is compiled and archived as maps; some of this information is not readily accessible at regional and state levels, and it is expensive to update or revise. The final two recommendations address the relationships between a water resources information system and source data.

1. Comprehensive and flexible sorting procedures should be developed to permit extraction of any basic data, in any form, in any area, from the digital data files of the Geological Survey.

2. Future national, regional, and state maps of water features and conditions should be prepared in digital form, in accordance with the specifications of National Mapping Division, for digital line graph data. If prepared by manual methods in a field office, maps should be in as basic a form as possible, so that they can be digitized inexpensively, and so that the data sets will permit maximum flexibility in future digital processing. The latter method may require the manual preparation of very basic overlays (map separates) that are specially prepared for raster scanning systems.

REFERENCES

Calkins, H. W., and Tomlinson, R. F., 1977, Geographic information systems, methods, and equipment for land use planning, 1977: Ottawa, International Geographical Union Commission on Geographical Data Sensing and Processing, 394 p.

Davis, J. C., 1973, Statistics and data analysis in geology: New York, John Wiley, 550 p.

Estes, J. E., 1982, Remote sensing and geographic information systems coming of age in the eighties *in* Pecora Symposium, 7th, Sioux Falls, S. Dak., 1981, Proceedings: Falls Church, Virginia, American Society of Photogrammetry, p. 23-40.

Loveland, T. R., and Johnson, G. E., 1982, The role of remotely sensed and other spatial data for predictive modeling--the Umatilla, Oregon, example *in* Pecora Symposium, 7th, Sioux Falls, S. Dak., 1981, Proceedings: Falls Church, Virginia, American Society of Photogrammetry, p. 442-454.

Olcott, P. G., 1968, Water resources of Wisconsin, Fox-Wolf River basin: U.S. Geological Survey Hydrologic Atlas HA-321, 4 sheets.

Singer, D. A., and Mosier, D. L., 1981, A review of regional mineral resource assessment methods: Economic Geology, v. 76, p. 1006-1015.

U.S. Department of Agriculture, 1981, Soil, water, and related resources in the United States, 1980 appraisal, pt. 1 - Status, condition, and trends, and pt. 2 - Analysis of resource trends: U.S. Department of Agriculture, I - 328 p. and II - 296 p.

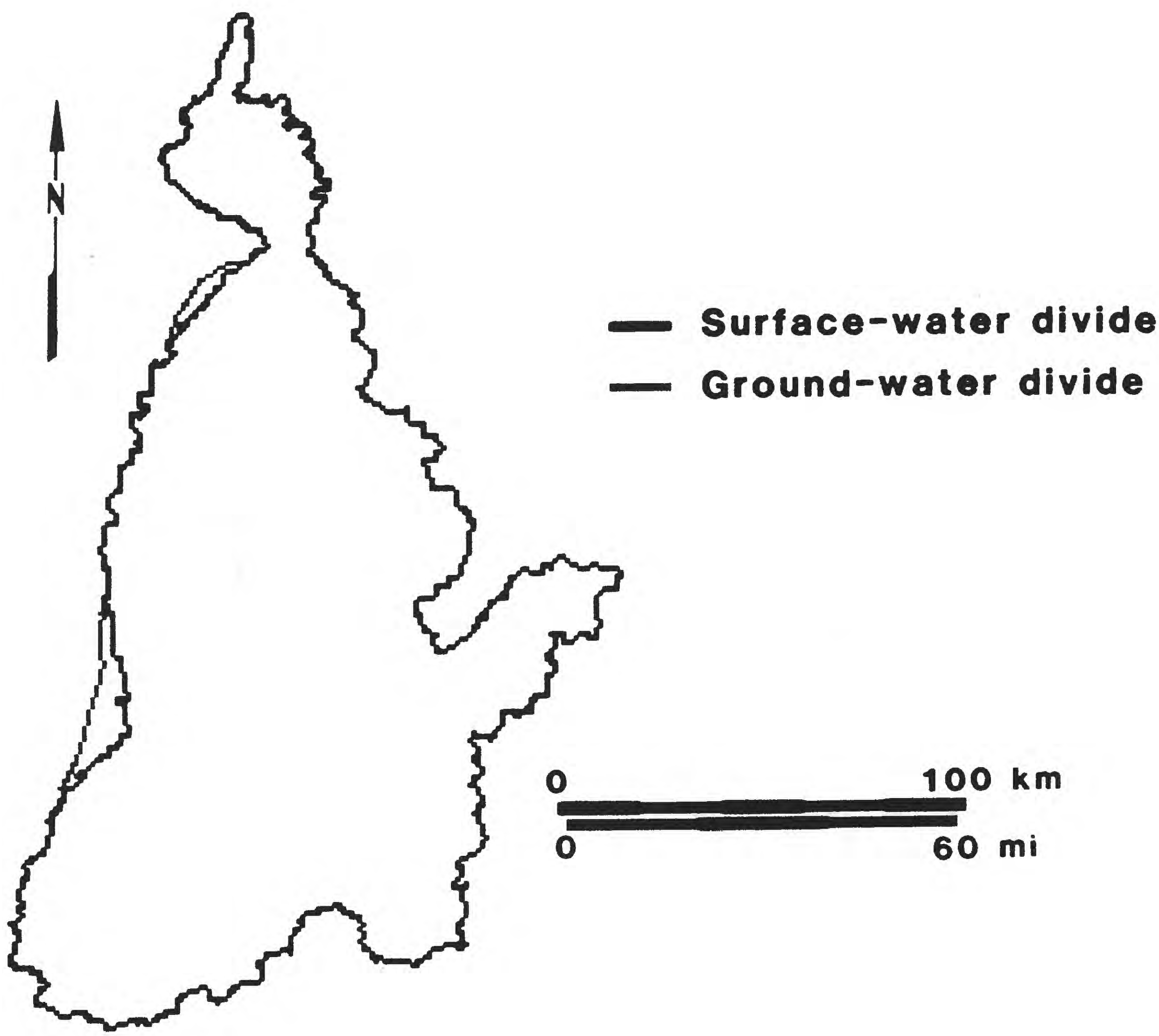
U.S. Water Resources Council, 1978, The Nation's water resources--1975-2000, Second national water assessment, v. 1 - Summary, v. 2 - Water quantity, quality, and related land considerations, v. 3 - Analytical data summary, and v. 4 - Water resources regional reports (21 repts.): Washington, D.C., U.S. Government Printing Office, v. 1 - 83 p., v. 3 - 89 p., other paging by section.

Wisconsin Department of Natural Resources, 1982, Wisconsin nonpoint source water pollution abatement program--A report to the Governor and Legislature: Wisconsin Department of Natural Resources, Bureau of Water Quality Management, 53 p.

APPENDIX

This appendix consists of 30 illustrations, including data sets that were used for information processing (figs. 3-14), very significant single data sets, and other representative examples of the types of data sets. Each illustration shows a single data set in a basic or primitive form. A primitive data set is defined as two or more features in the form of points, lines, polygons, or a surface. It is also defined as the form in which the data set was digitized, or otherwise acquired, and entered as a layer in the information system. There are four exceptions to this general definition of a primitive data set. First, all data sets are shown inside an outline of the Fox-Wolf basin for location purposes. Second, all points have been enlarged for visibility on the illustrations. Third, all numerical values were classified into intervals by level slicing and are shown as distinctive gray tones on the illustrations. Finally, several data sets, which are inherent surfaces, were acquired as points, which represent samples of the surfaces. In these cases, the surfaces were generated and then level sliced for the illustrations. Only the shaded relief map is shown as a true surface (with continuous gray tones) in these illustrations.

Each data set may have more than one form in the information system. In general, the most primitive form is shown in this appendix; other forms are indicated in table 9. Some data sets were acquired as digital attribute files. These files are not shown in the appendix, because they are tabular in nature. However, a number of the digital files are shown as point, polygon, and level-sliced surface data sets. Each illustration is accompanied by a description of the data set, including its origin and the processing required for its inclusion as one or more layers in the Fox-Wolf information system.



1

Data Set:

SURFACE-WATER AND GROUND-WATER DIVIDES FOR THE FOX-WOLF BASIN

Source:

1:500,000 scale map overlays; complete polygon for surface-water divide; deviations from this polygon for ground-water divide.

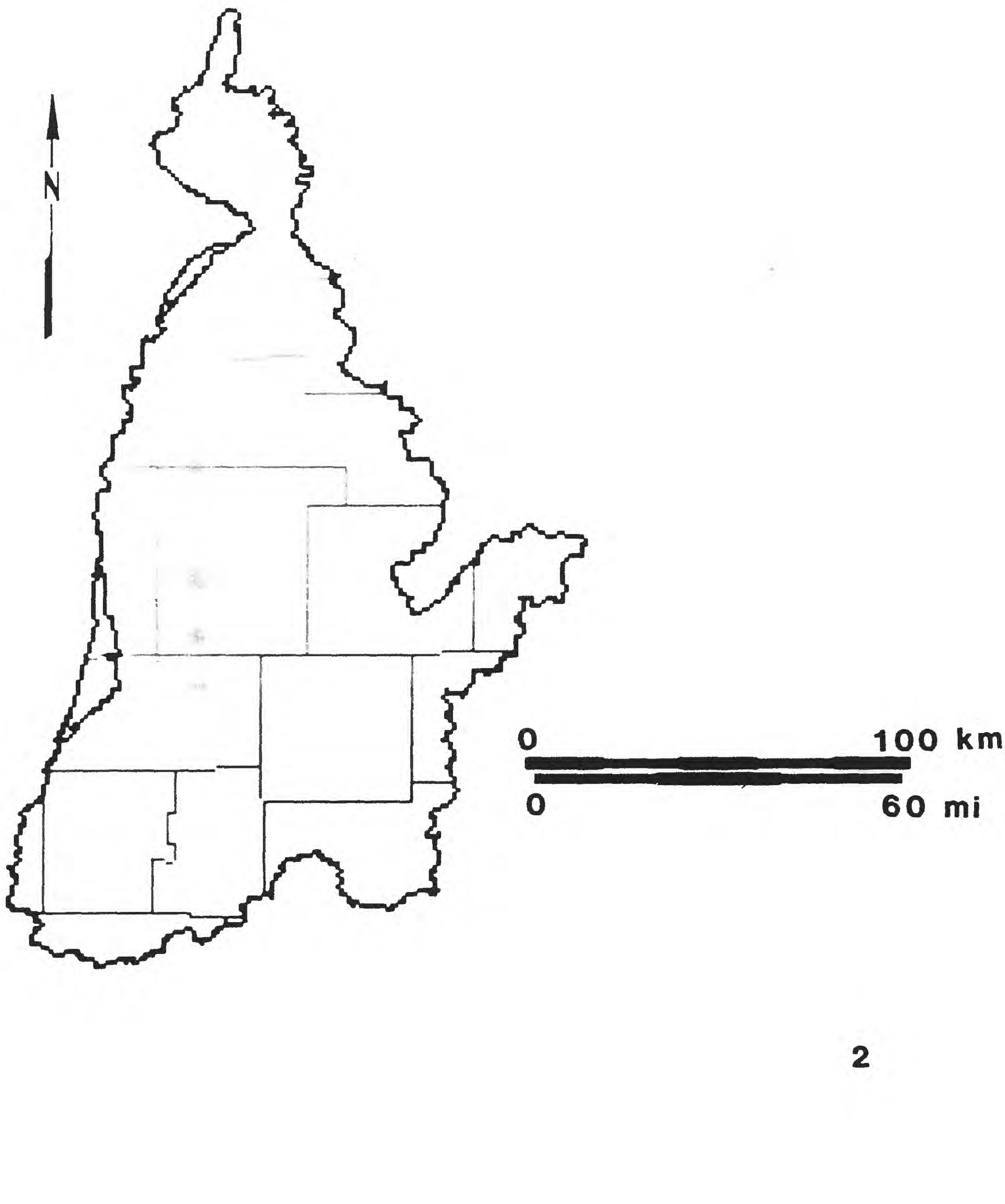
Processing:

1. Digitize polygons in vector form, convert to raster form, and register.
2. Make separate data sets of the boundary lines.
3. Add boundaries to make a single composite data set.

Coding:

Class 0. Background.

1. Surface-water divide.
2. Ground-water divide.



2

Data Set:

COUNTY LINES

Source:

1:500,000 scale State map.

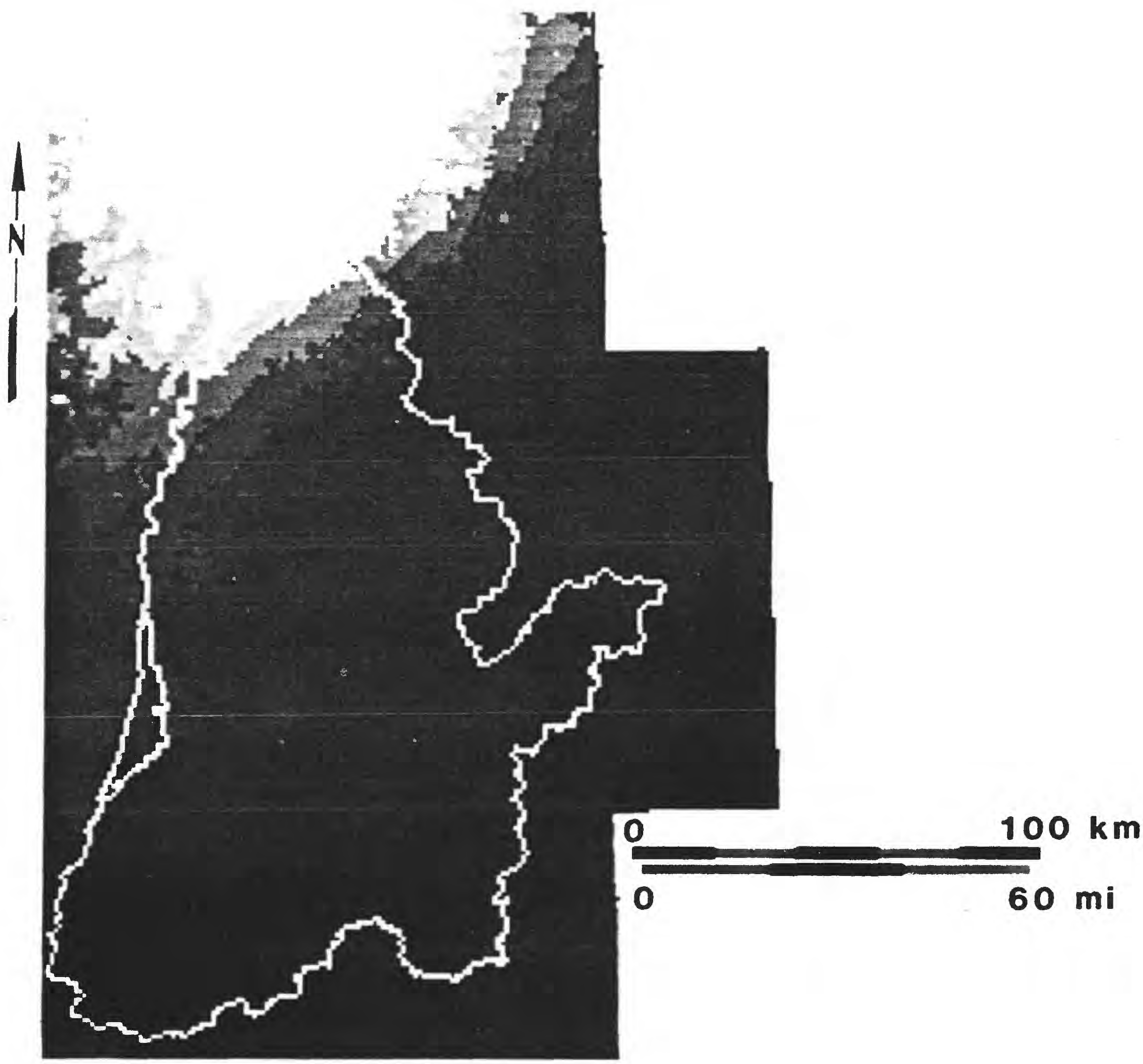
Processing:

1. Digitize county polygons in vector form, convert to raster form, and register.
2. Use the ground-water divide polygon as a mask to delete counties outside the basin.
3. Make a separate data set of all county lines.

Coding:

Class 0. Background.

1. County lines.



3

Data Set:

MEAN SURFACE ELEVATIONS

Source:

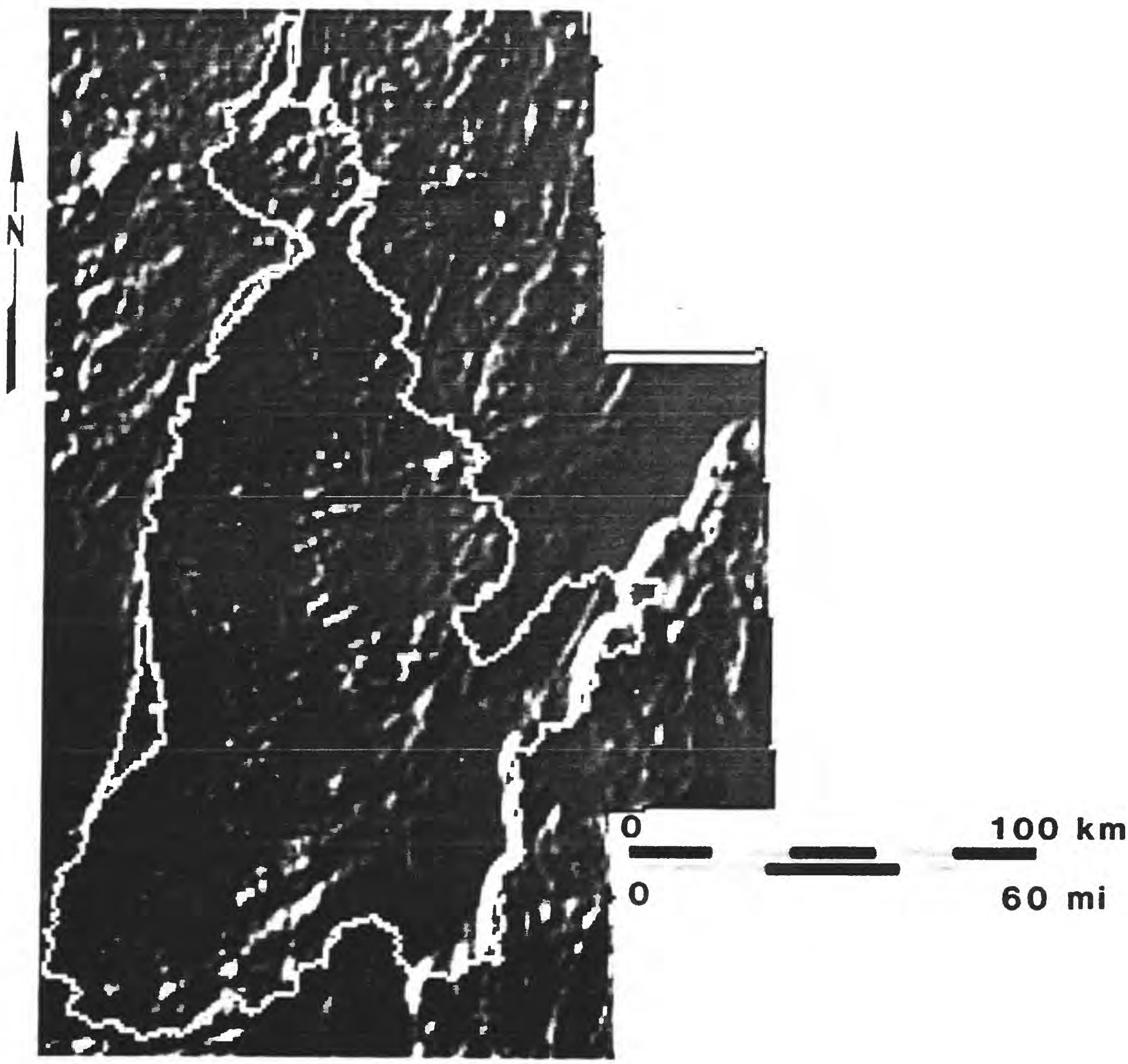
Digital Elevation Model data from the 1:250,000 scale map series.

Processing:

1. Mosaic seven 1° by 1° blocks of digital data.
2. Resample from 3-arc-second to 100-m grid cells and register.
3. Use maximum and minimum algorithms to determine these values in 1-km grid cells.
4. Average the maximum and minimum data sets to obtain mean elevations.
5. Level slice the elevation surface for display.

Coding:

Code each 100-ft. elevation interval for display as a separate gray tone. The lightest tone has an elevation, of 1,900-2,000 feet above sea level (NGVD of 1929). The darkest tone has an elevation of 500-600 feet above sea level.



4

Data Set:

SHADED RELIEF MAP

Source:

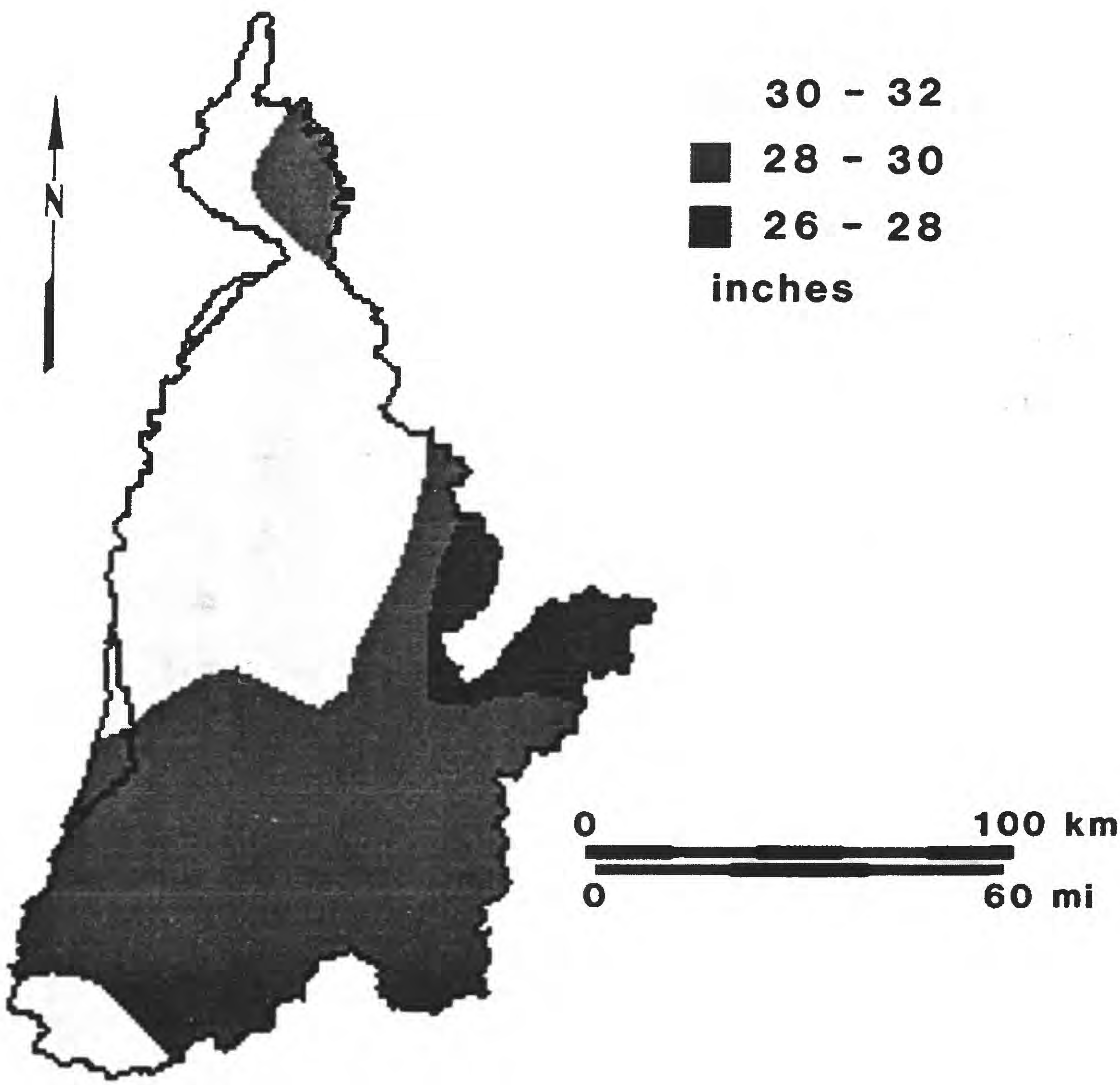
Digital data set showing mean surface elevations.

Processing:

Use algorithm that calculates relative brightness and areas of shadow for a simulated sun. Position the sun at N. 60° W. and at an elevation angle of 10°. Use parameters that result in a 20X vertical exageration.

Coding:

As calculated by the shaded relief algorithm.



5

Data Set:

MEAN ANNUAL PRECIPITATION

Source:

1:500,000 scale map overlay.

Processing:

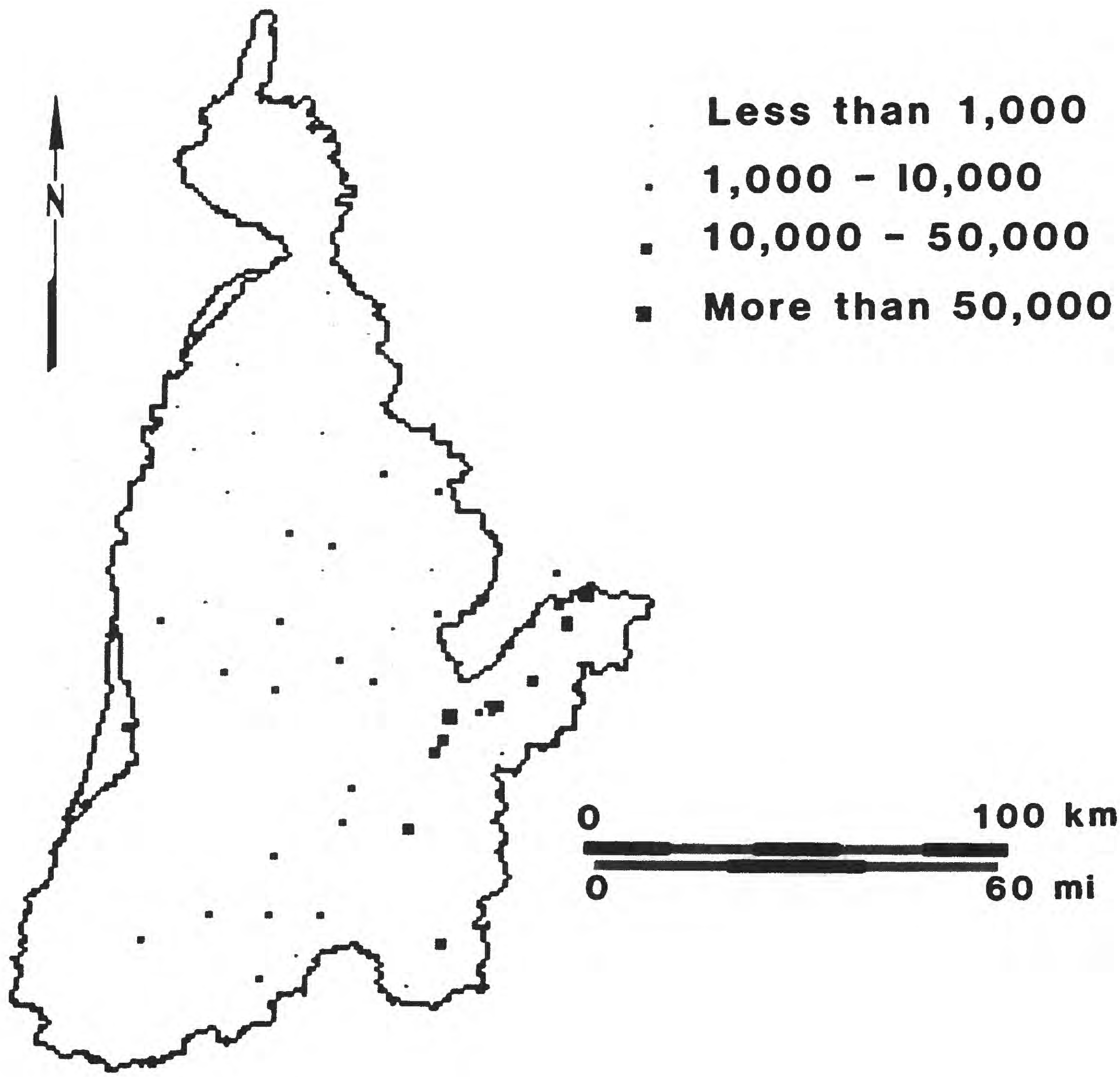
Digitize intervals of mean annual precipitation as polygons in vector form; convert to raster form and register.

Coding:

Class 1. 26-28 inches.

2. 28-30 inches.

3. 30-32 inches.



Less than 1,000
1,000 - 10,000
10,000 - 50,000
More than 50,000
N
0
100 km
0
60 mi

6

Data Set:

POPULATIONS OF CITIES AND TOWNS

Source:

Digital file containing latitude and longitude locations of points and the associated populations.

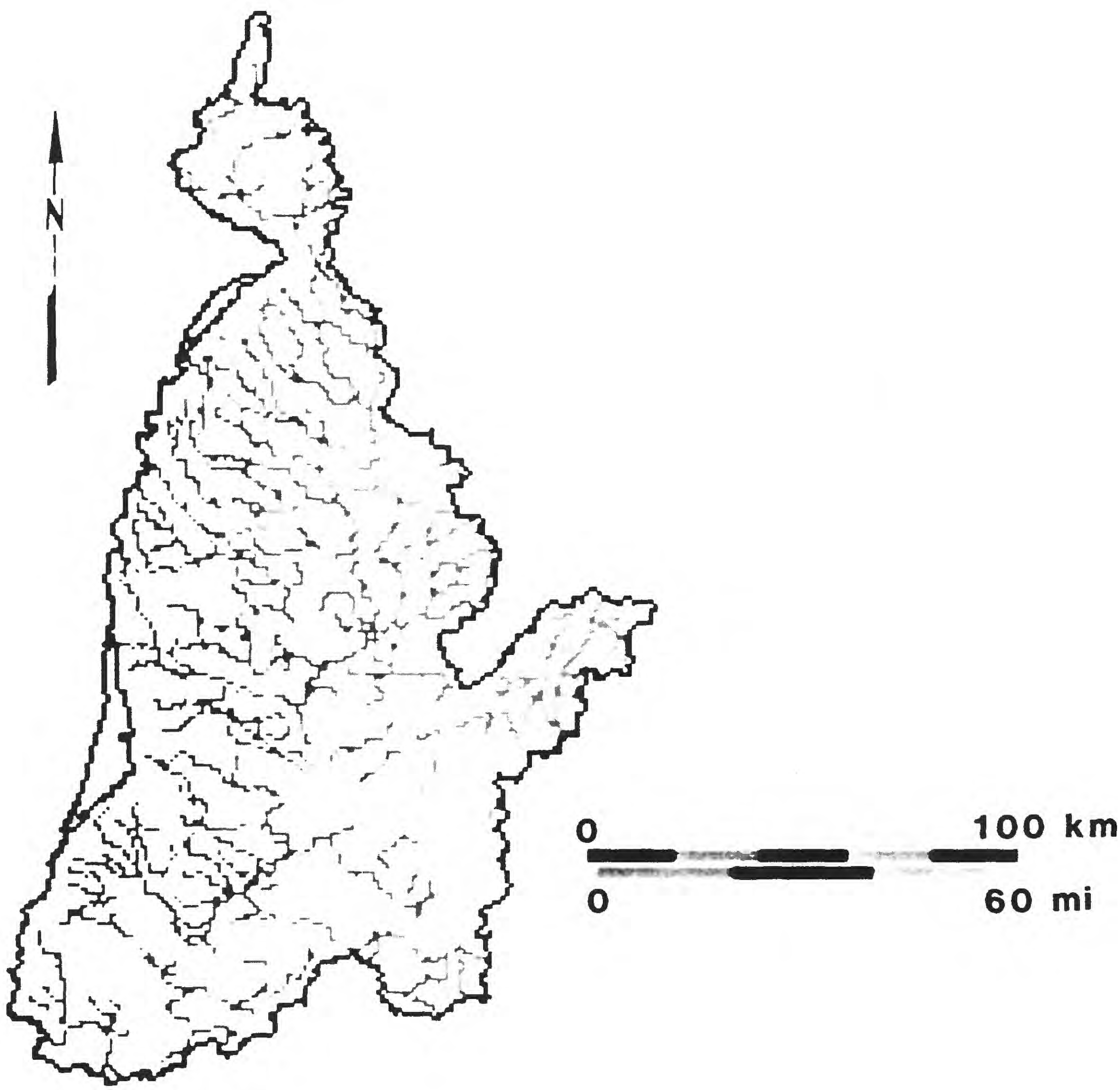
Processing:

1. Enter digital file as a text file.
2. Convert latitude and longitude values to line and sample coordinates in a Lambert Conformal Conic map projection.
3. Code the grid cell for each city or town with the population number.
4. Group populations into four classes.
5. Use proximity algorithm to show population classes by size of grid-cell array.

Coding:

Class 0. Background.

1. Less than 1,000 population; 1 by 1.
2. 1,000-10,000 population; 2 by 2.
3. 10,000-50,000 population; 3 by 3.
4. More than 50,000 population; 4 by 4.



7

Data Set:

STREAMS

Source:

1:500,000 scale State map.

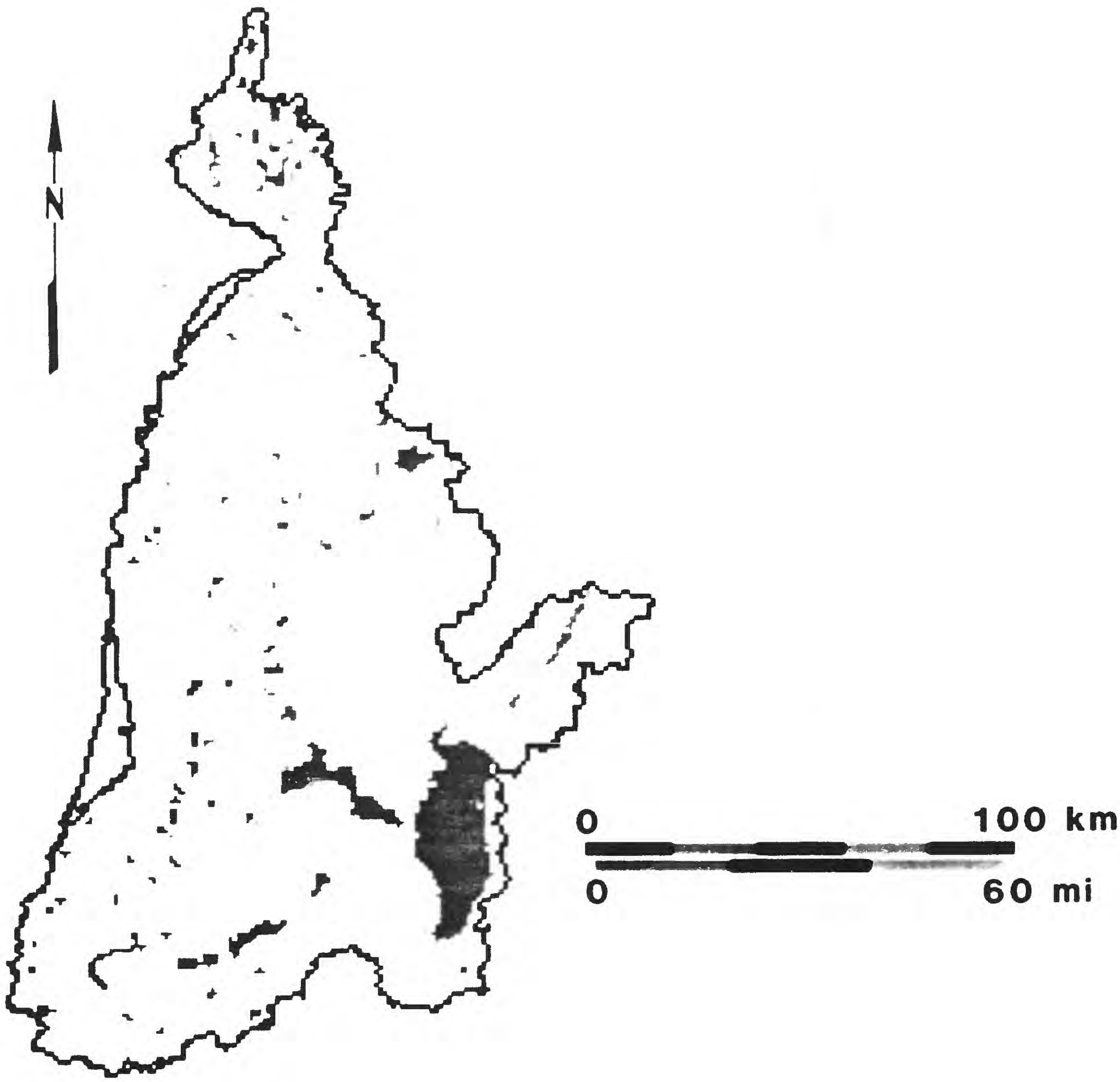
Processing:

Digitize lines in vector form, convert to raster form and register.

Coding:

Class 0. Background.

1. Streams.



8

Data Set:

LAKES

Source:

1:500,000 scale State map.

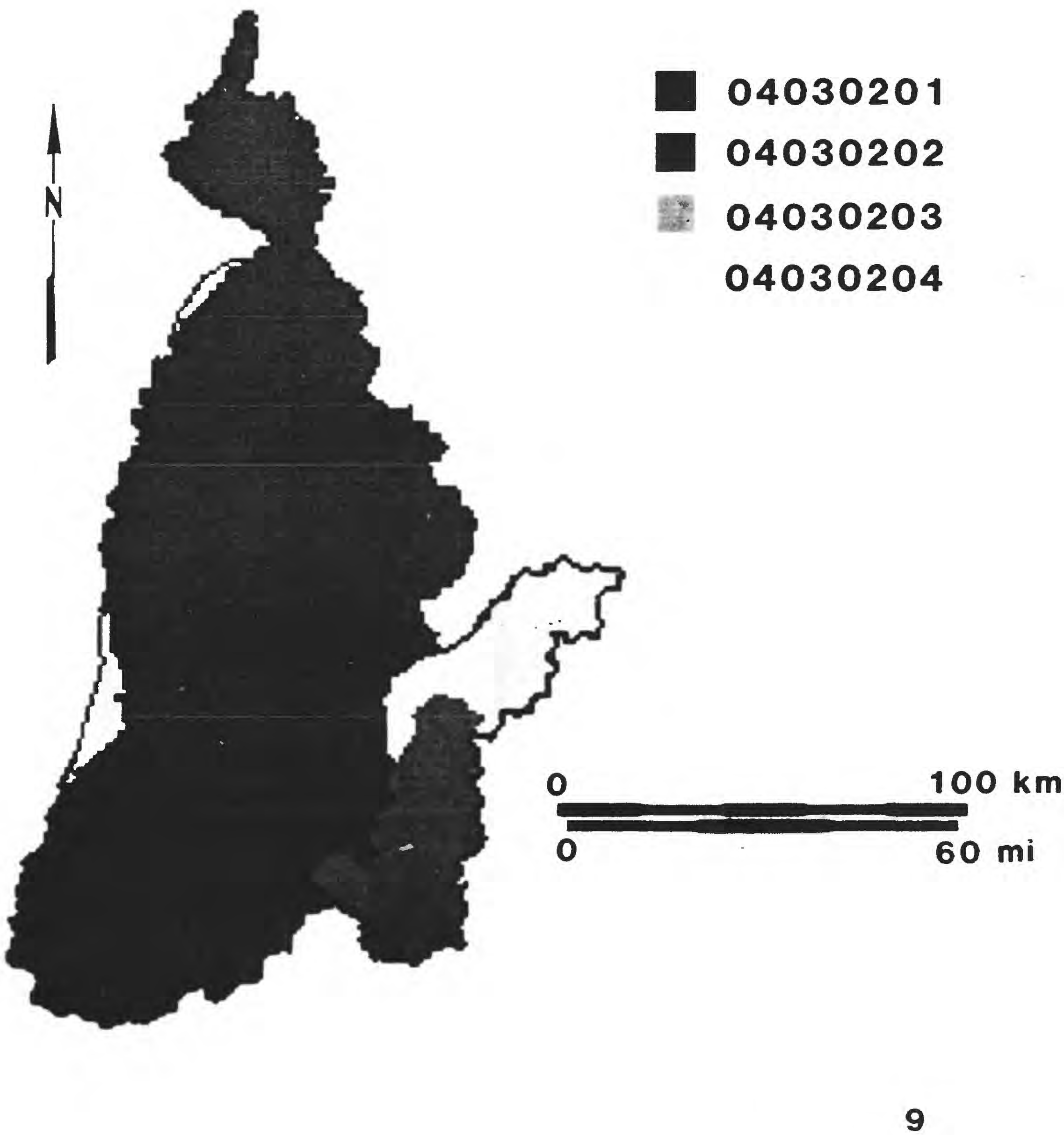
Processing:

Digitize all lakes and double-line streams as polygons in vector form, convert to raster form, and register.

Coding:

Class 0. Background.

1. Lakes.


N
04030201
04030202
04030203
04030204
0
100 km
0
60 mi
9

Data Set:

HYDROLOGIC CATALOGING UNITS

Source:

1:500,000 scale map overlay.

Processing:

Digitize polygons in vector form, convert to raster form, and register.

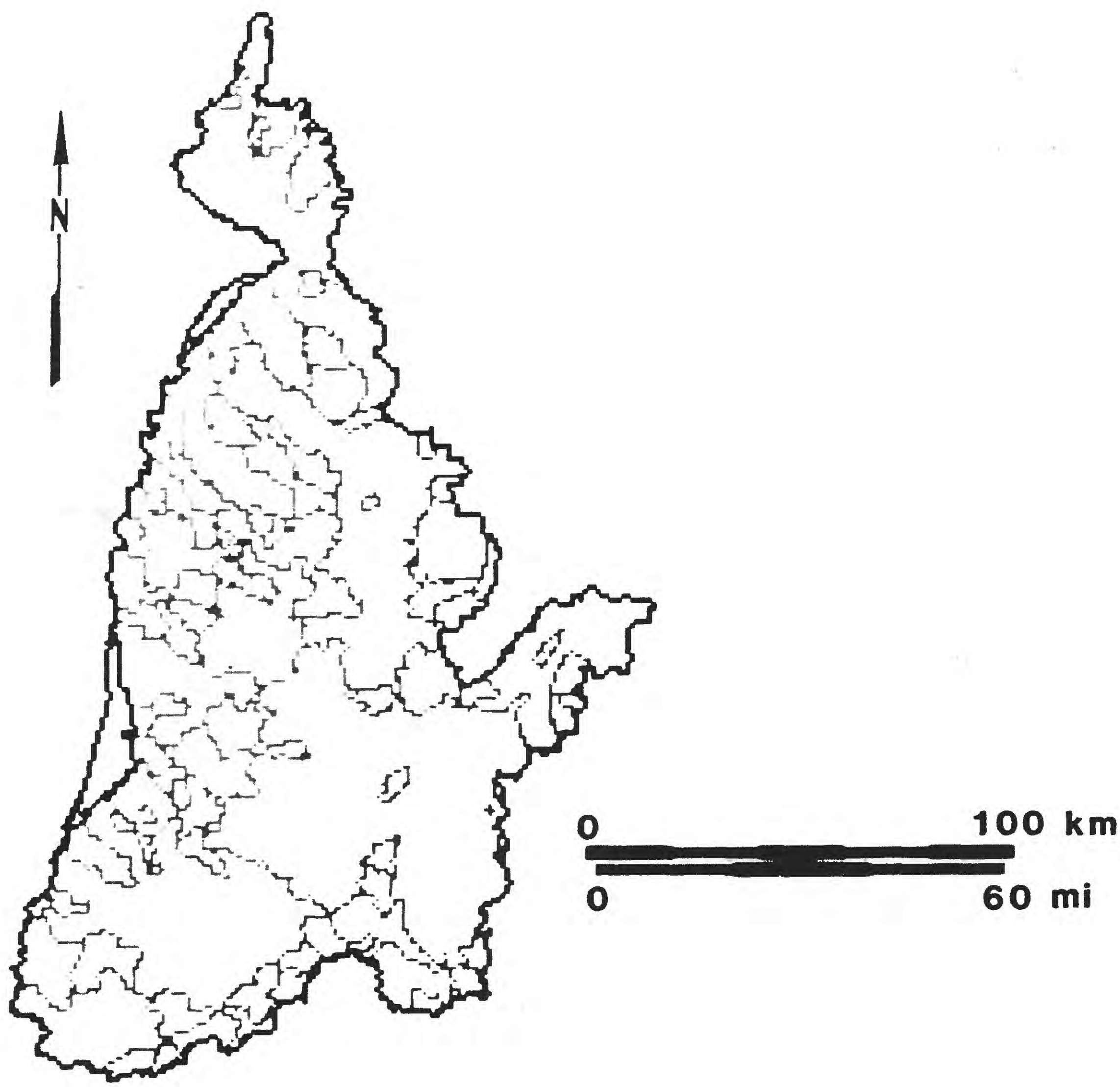
Coding:

Class 1. Fox River basin, 04030201.

2. Wolf River basin, 04030202.

3. Lake Winnebago basin, 04030203.

4. Fox-Wolf River basin, 04030204.



10

Data Set:

BASIN BOUNDARIES

Source:

1:500,000 scale map overlays (3).

Processing:

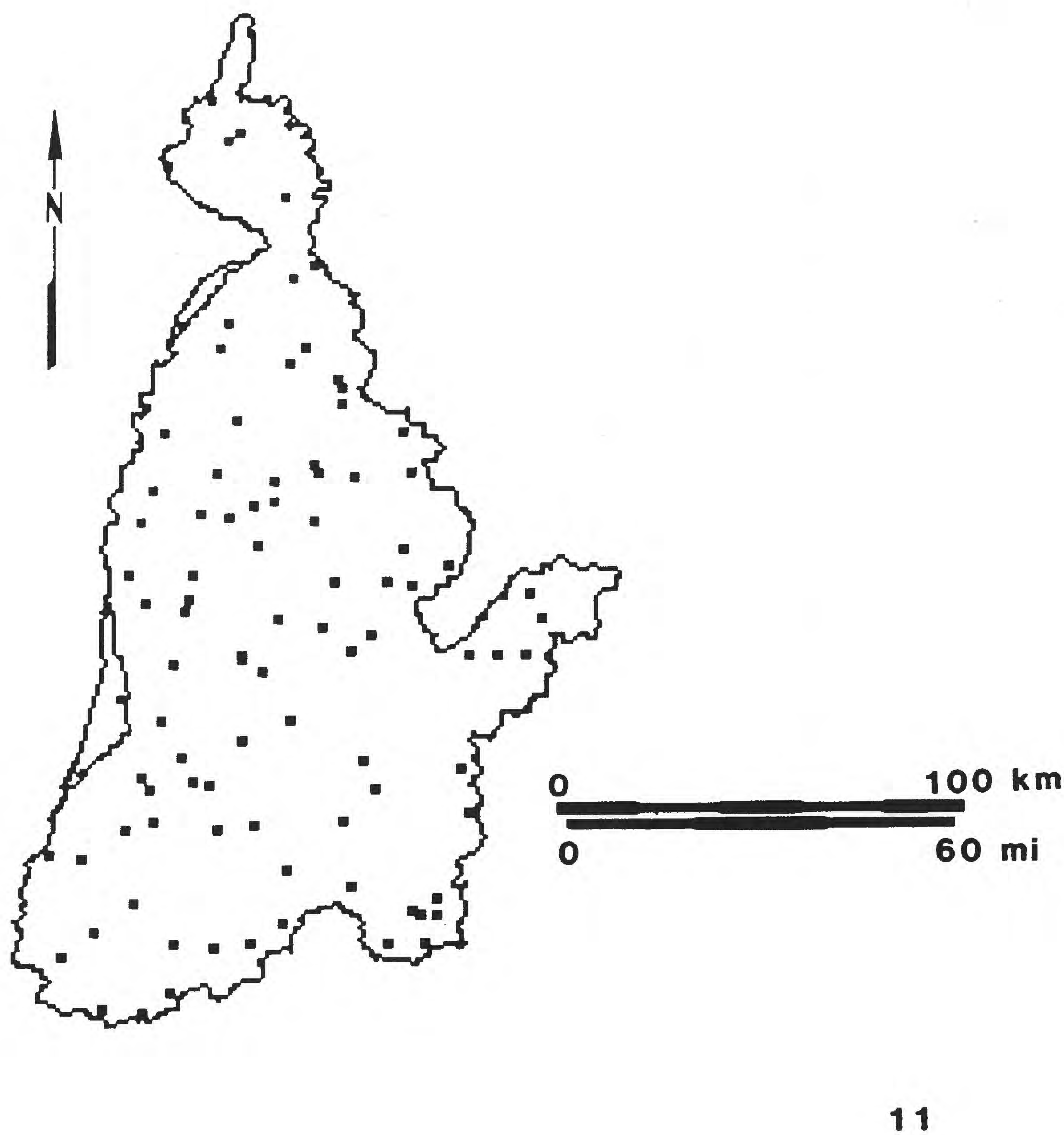
1. Digitize polygons in vector form, convert to raster form, and register.

2. Make a separate data set of the boundary lines.

Coding:

Class 0. Background.

1. Minor basin boundaries.


N
0
100 km
0
60 mi
11

Data Set:

CONTINUOUS AND PARTIAL RECORD STREAMFLOW GAGING STATIONS

Source:

1:500,000 scale map overlays.

Processing:

1. Digitize points in vector form, convert to raster form and register.

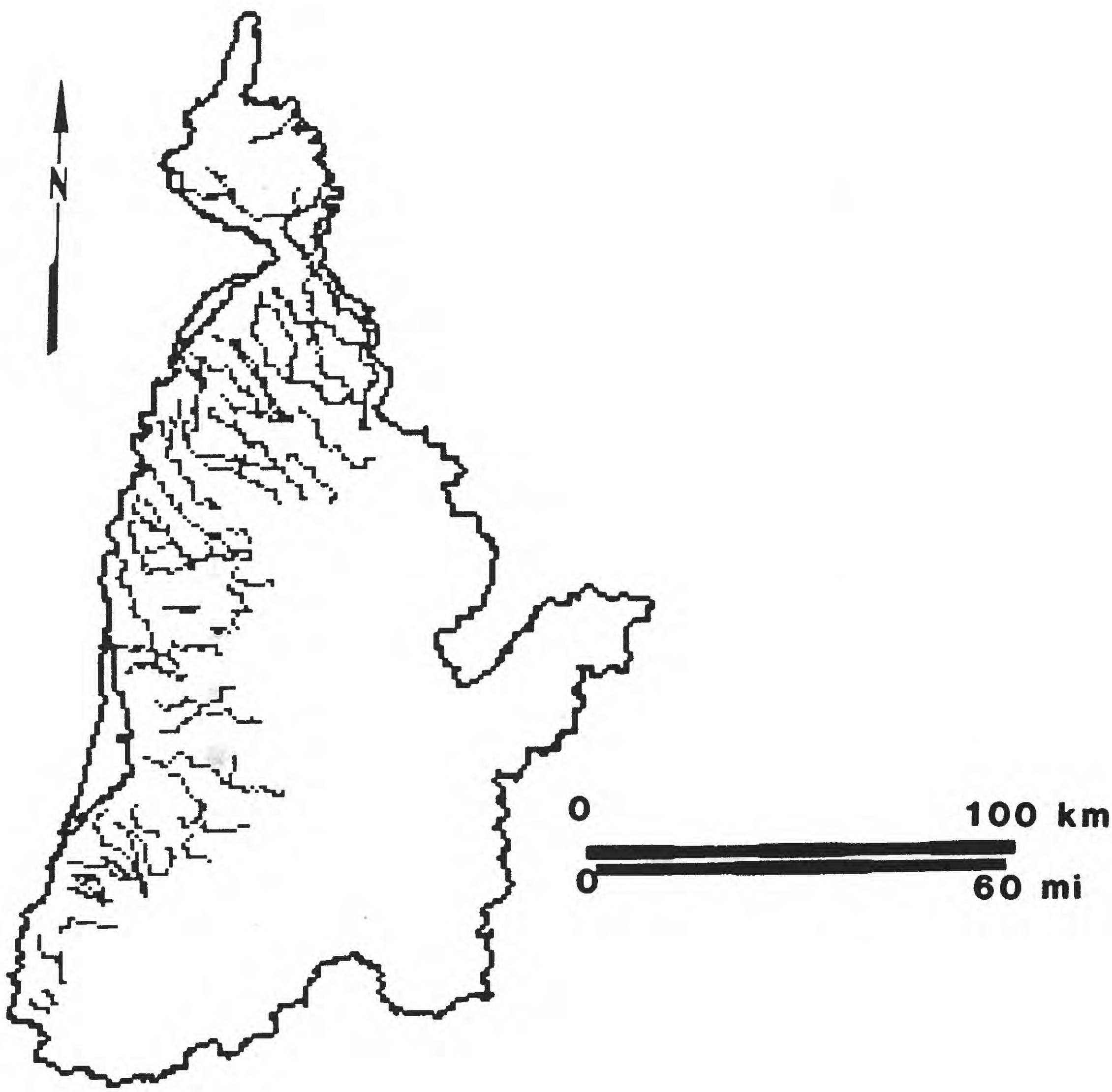
2. Code all points as a single class.

3. Use proximity algorithm to enlarge points for display.

Coding:

Class 0. Background.

1. Stream gaging station or partial record station.



12

Data Set:

TROUT STREAMS

Source:

1:500,000 scale Wisconsin Land Resources Analysis Program map.

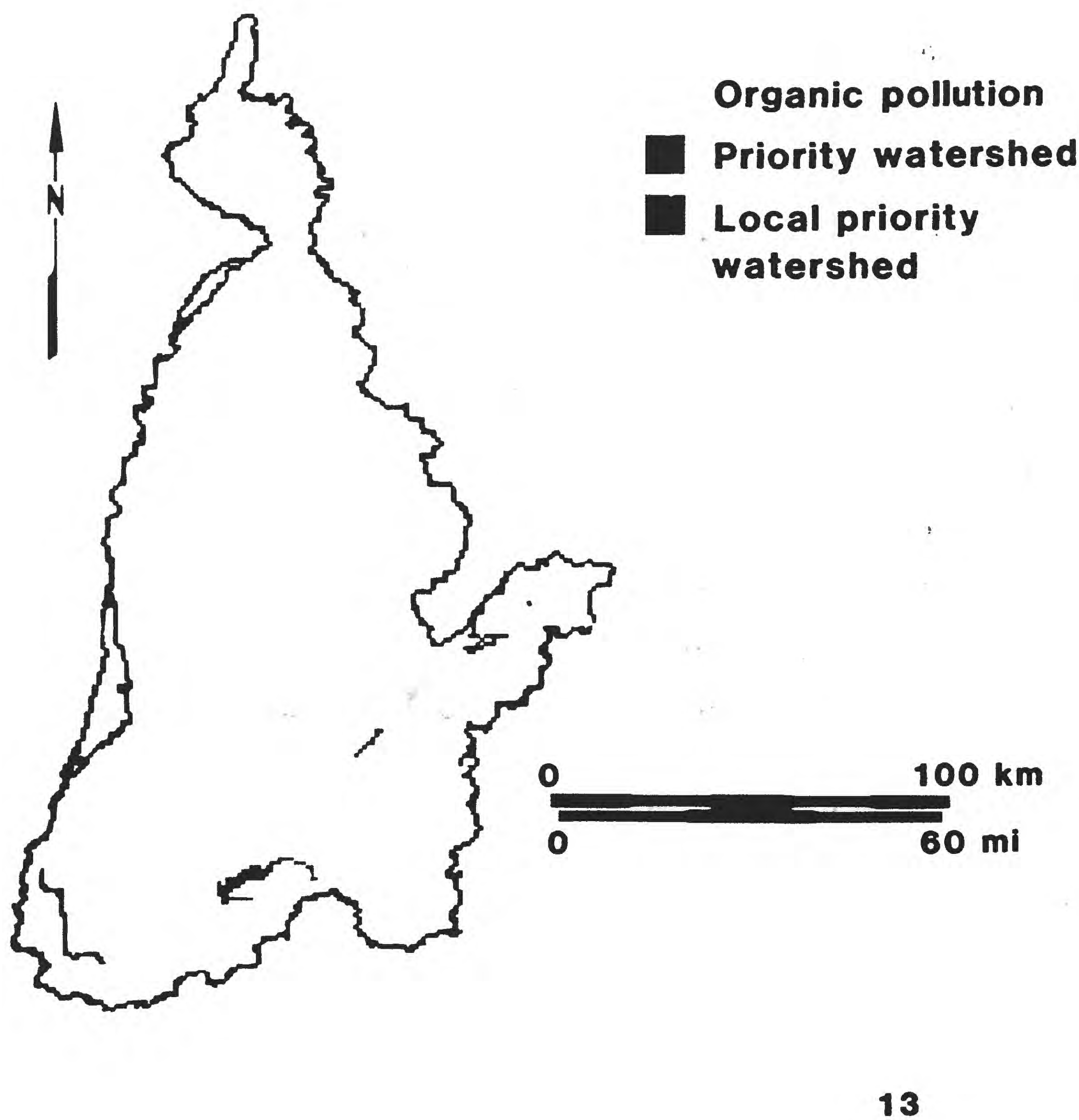
Processing:

Digitize lines in vector form, convert to raster form, and register.

Coding:

Class 0. Background.

1. Trout streams.


N
Organic pollution
Priority watershed
Local priority watershed
0
100 km
0
60 mi
13

Data Set:

POLLUTED STREAMS AND LAKES

Source:

1:500,000 scale map overlay (from Wisconsin Department of Natural Resources, 1982) and lakes data set.

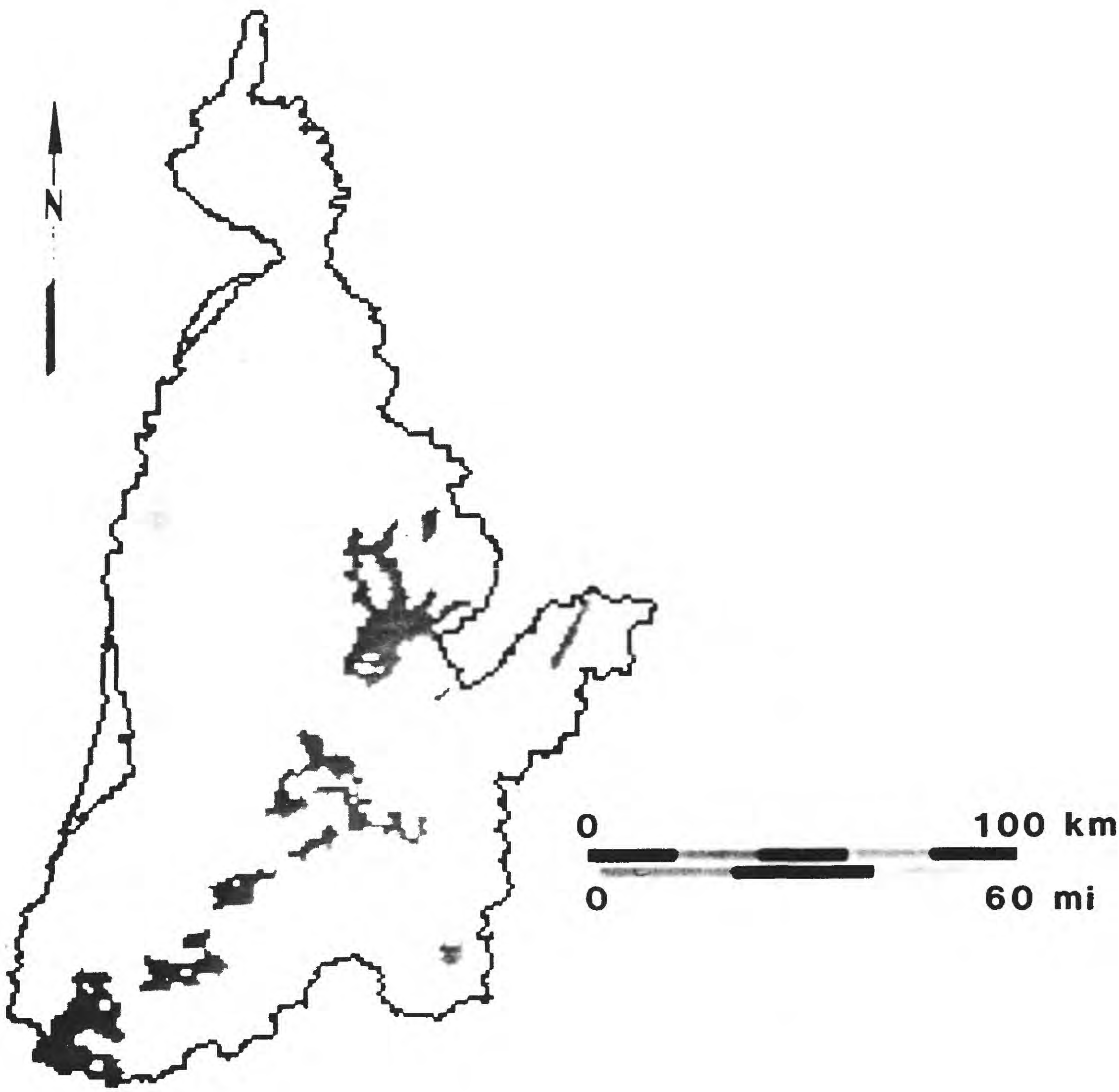
Processing:

1. Digitize stream lines in vector form, convert to raster form, and register.
2. Extract polluted lakes from the lakes data set.
3. Add polluted streams and lakes to make a single composite data set.

Coding:

Class 0. Background.

1. Organic pollution.
2. Water feature in a priority watershed.
3. Water feature in a local priority watershed.



14

14

Data Set:

FLOOD-PRONE AREAS

Source:

1:500,000 scale Wisconsin Land Resources Analysis Program map.

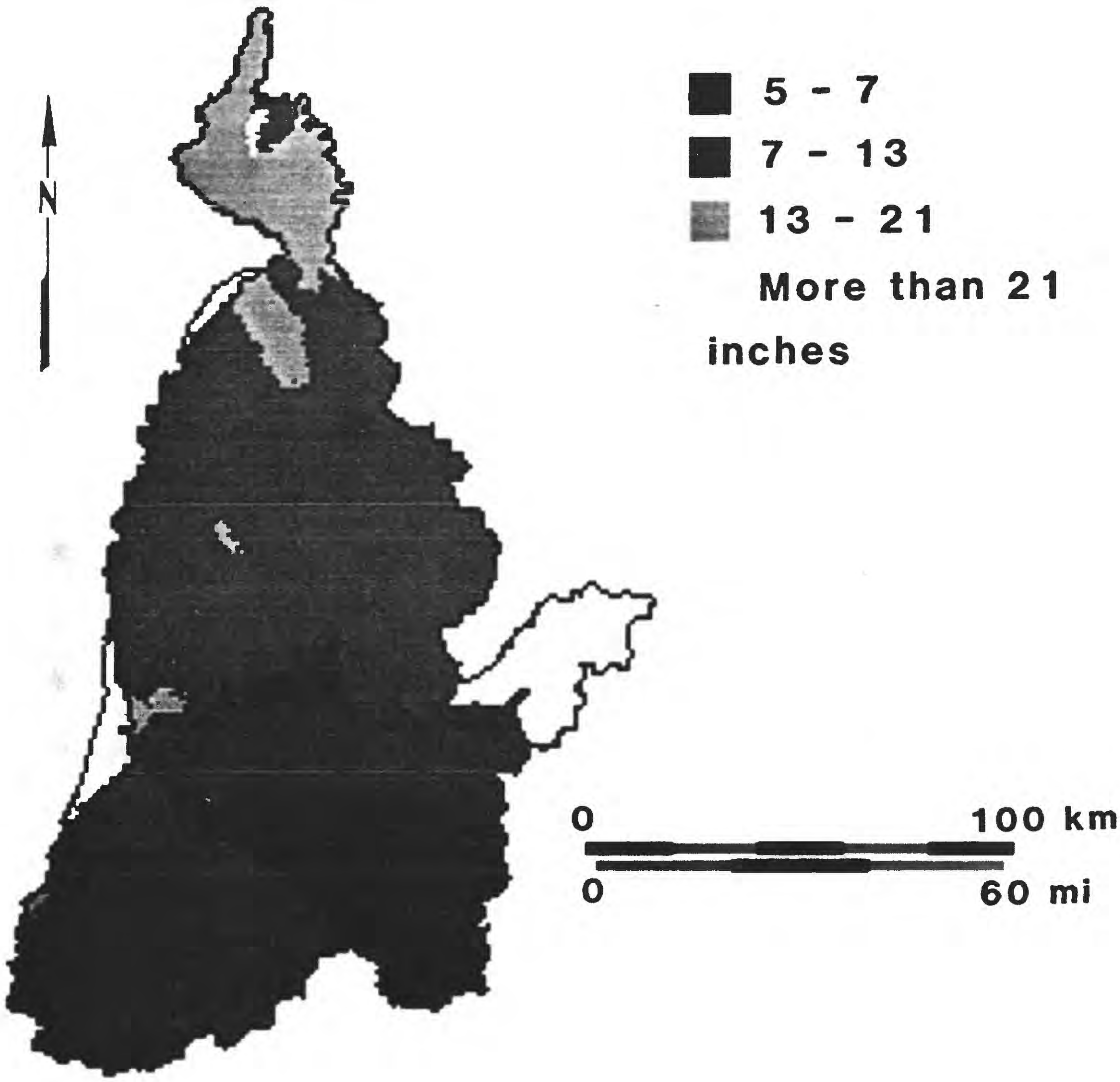
Processing:

Digitize polygons in vector form, convert to raster form, and register.

Coding:

Class 0. Background.

1. Flood-prone areas.



15

Data Set:

MEAN ANNUAL RUNOFF

Source:

Polygons from minor basin boundaries data set and digital file of mean annual runoff values.

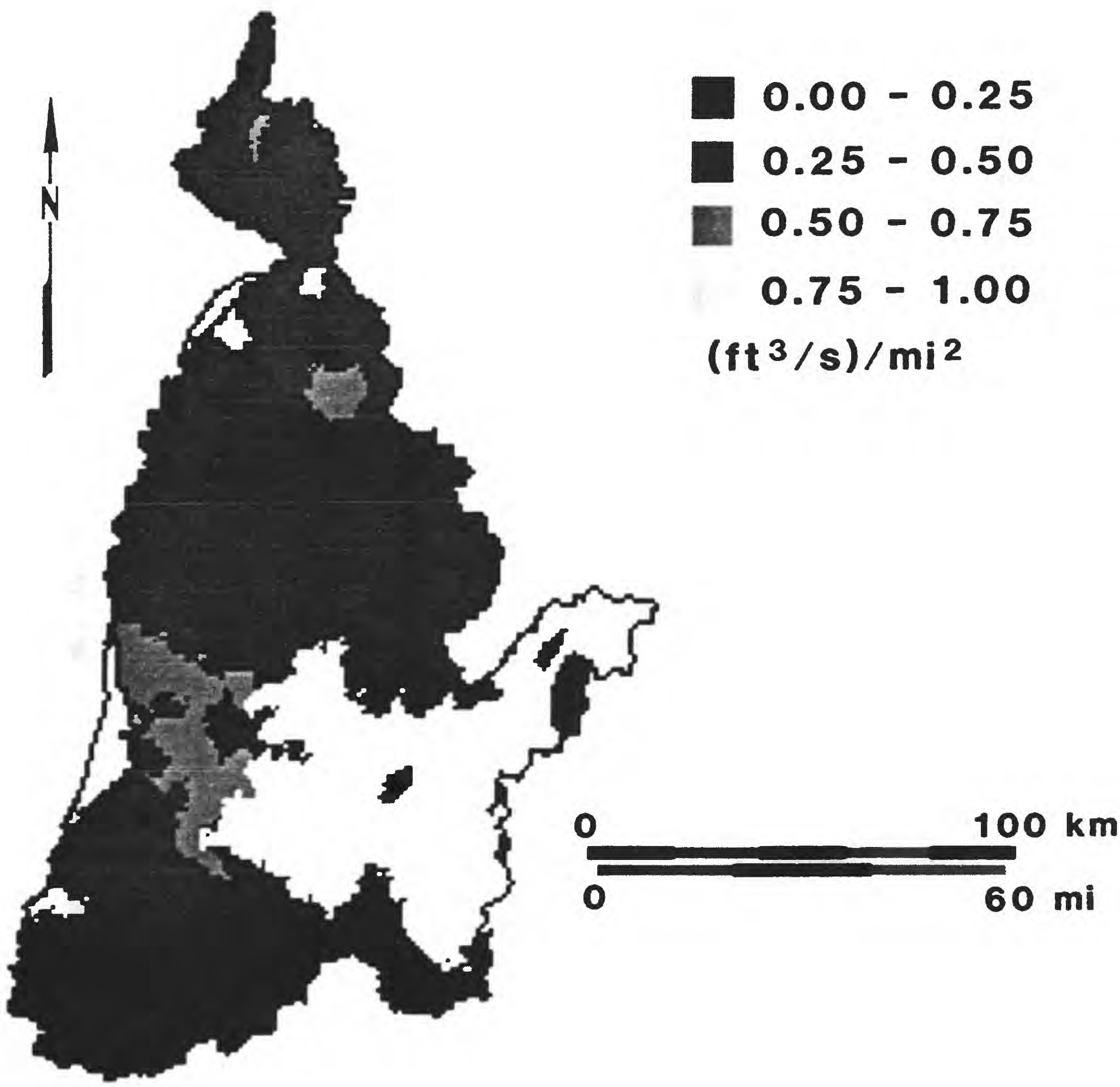
Processing:

1. Extract polygons for which mean annual runoff values are available from the minor basin boundaries data set.
2. Recode the extracted polygons with values of mean annual runoff.
3. Group mean annual runoff values into four classes.

Coding:

Class 0. Background; mean annual runoff value not available.

1. 5-7 inches.
2. 7-13 inches.
3. 13-21 inches.
4. More than 21 inches.



16

Data Set:

SEVEN-DAY, 10-YEAR LOW FLOW INDEX

Source:

Polygons from minor basin boundaries data set and 1:500,000 scale map overlay showing 7-day, 10-year low flow values.

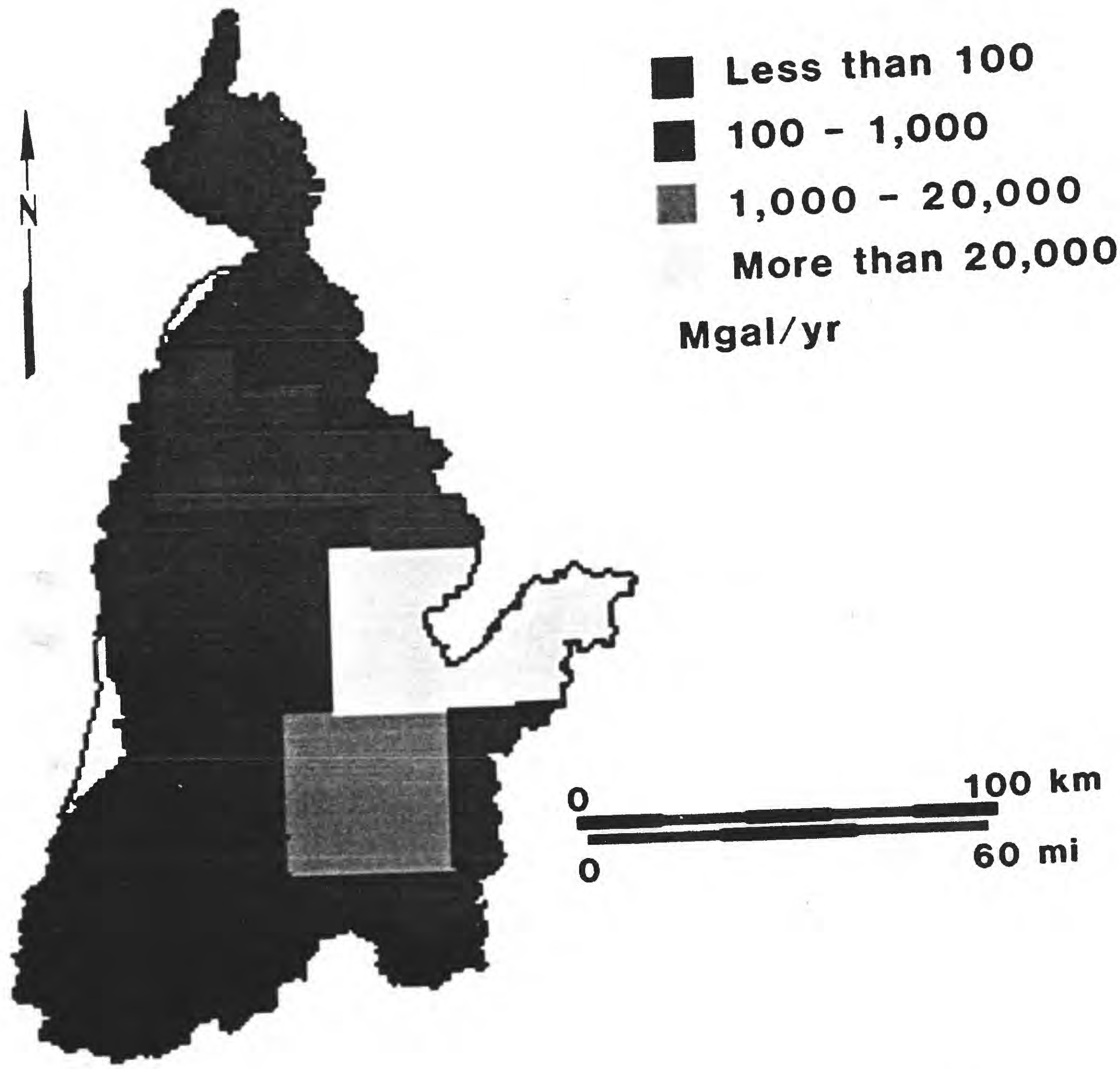
Processing:

1. Extract polygons for which 7-day, 10-year low flow values are available from the minor basin boundaries data set.
2. Recode the extracted polygons with the 7-day, 10-year low flow values.
3. Group the low flow values into four classes.

Coding:

Class 0. Background; 7-day, 10-year low flow value not available.

1. 0.00-0.25 $(ft^3/s)/mi^2$.
2. 0.25-0.50 $(ft^3/s)/mi^2$.
3. 0.50-0.75 $(ft^3/s)/mi^2$.
4. 0.75-1.00 $(ft^3/s)/mi^2$.


Less than 100
100 - 1,000
1,000 - 20,000
More than 20,000
Mgal/yr
N
0
100 km
0
60 mi

17

Data Set:

TOTAL SURFACE-WATER USE BY COUNTY

Source:

Polygons from county line data set and 1:500,000 scale map overlay showing total surface-water use in each county.

Processing:

Code county polygons to indicate water use intervals (classes of water use.)

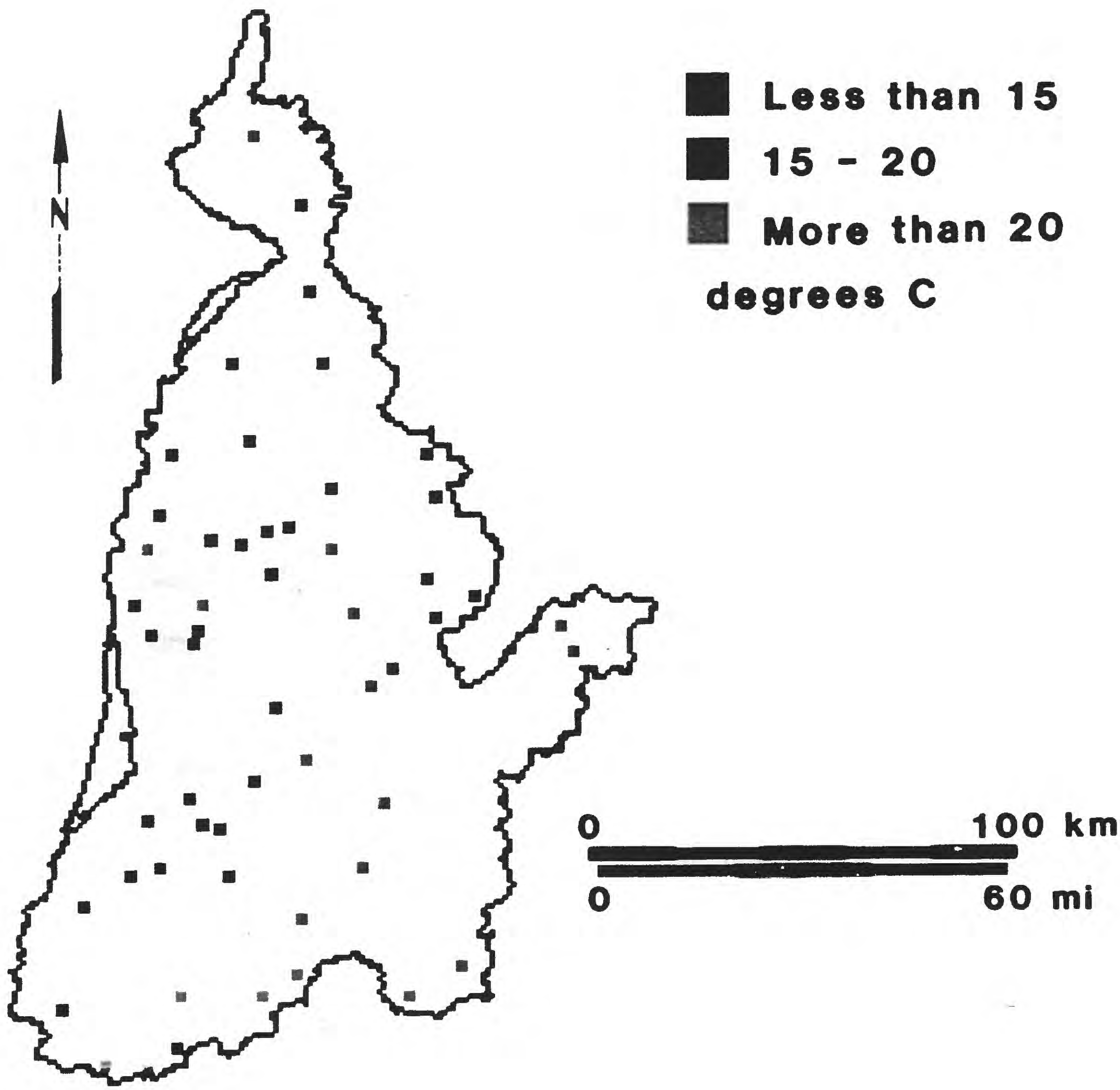
Coding:

Class 1. Less than 100 Mgal/yr.

2. 100-1,000 Mgal/yr.

3. 1,000-20,000 Mgal/yr.

4. More than 20,000 Mgal/yr.



18

Data Set:

STREAM TEMPERATURE AT BASE FLOW

Source:

Digital file containing latitude and longitude locations of points and the associated water temperature.

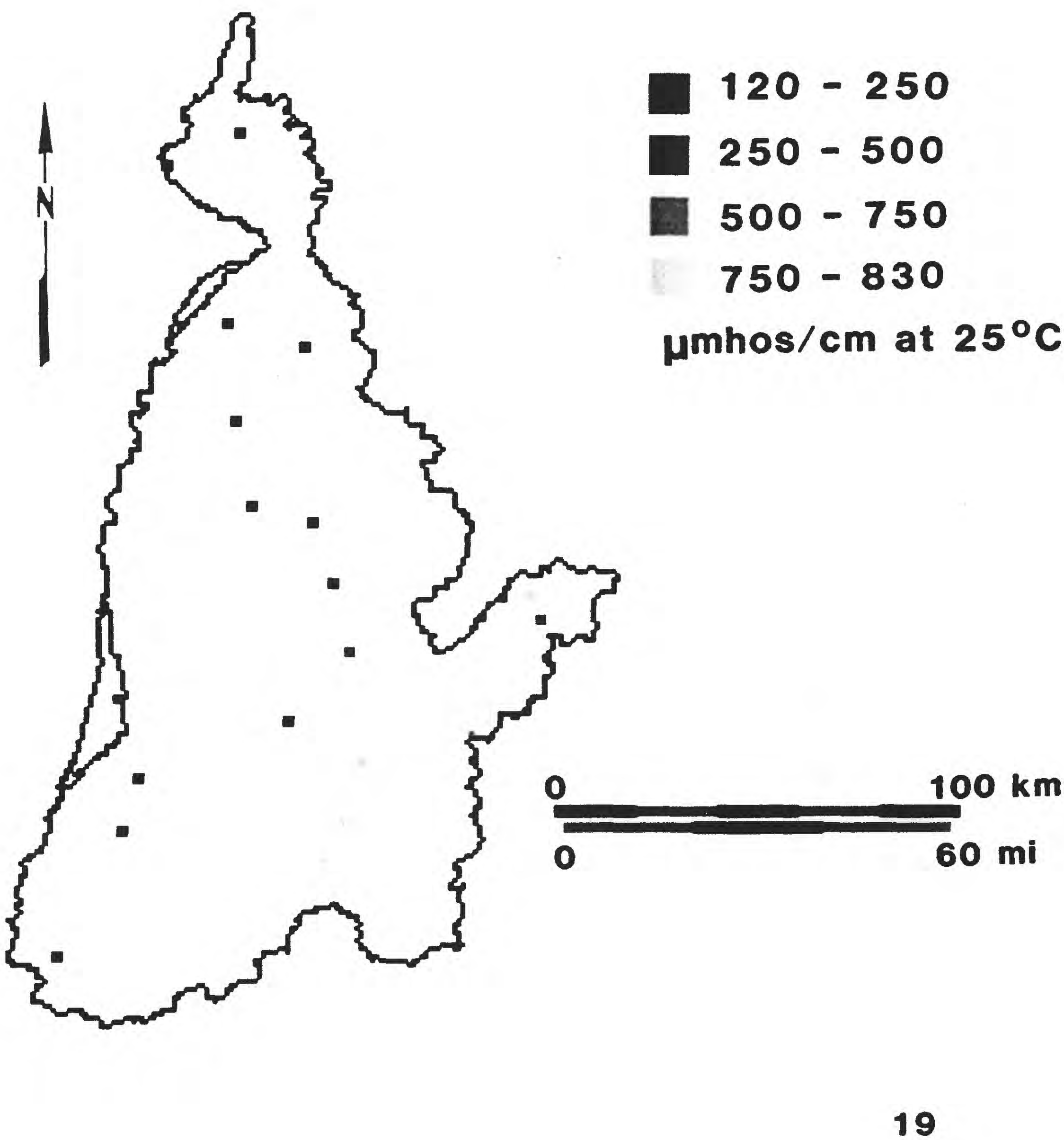
Processing:

1. Enter digital file as a text file.
2. Convert latitude and longitude values to line and sample coordinates in a Lambert Conformal Conic map projection.
3. Code the grid cell at each sampling location with the water temperature.
4. Group temperature values into four classes.
5. Use proximity algorithm to enlarge points for display.

Coding:

Class 0. Background.

1. Less than 15°C.
2. 15-20°C.
3. More than 20°C.



N
120 - 250
250 - 500
500 - 750
750 - 830
µmhos/cm at 25°C
0
100 km
0
60 mi
19

Data Set:

SPECIFIC CONDUCTANCE OF STREAMS AT BASE FLOW

Source:

Digital file containing latitude and longitude locations of points and the associated specific conductance of water.

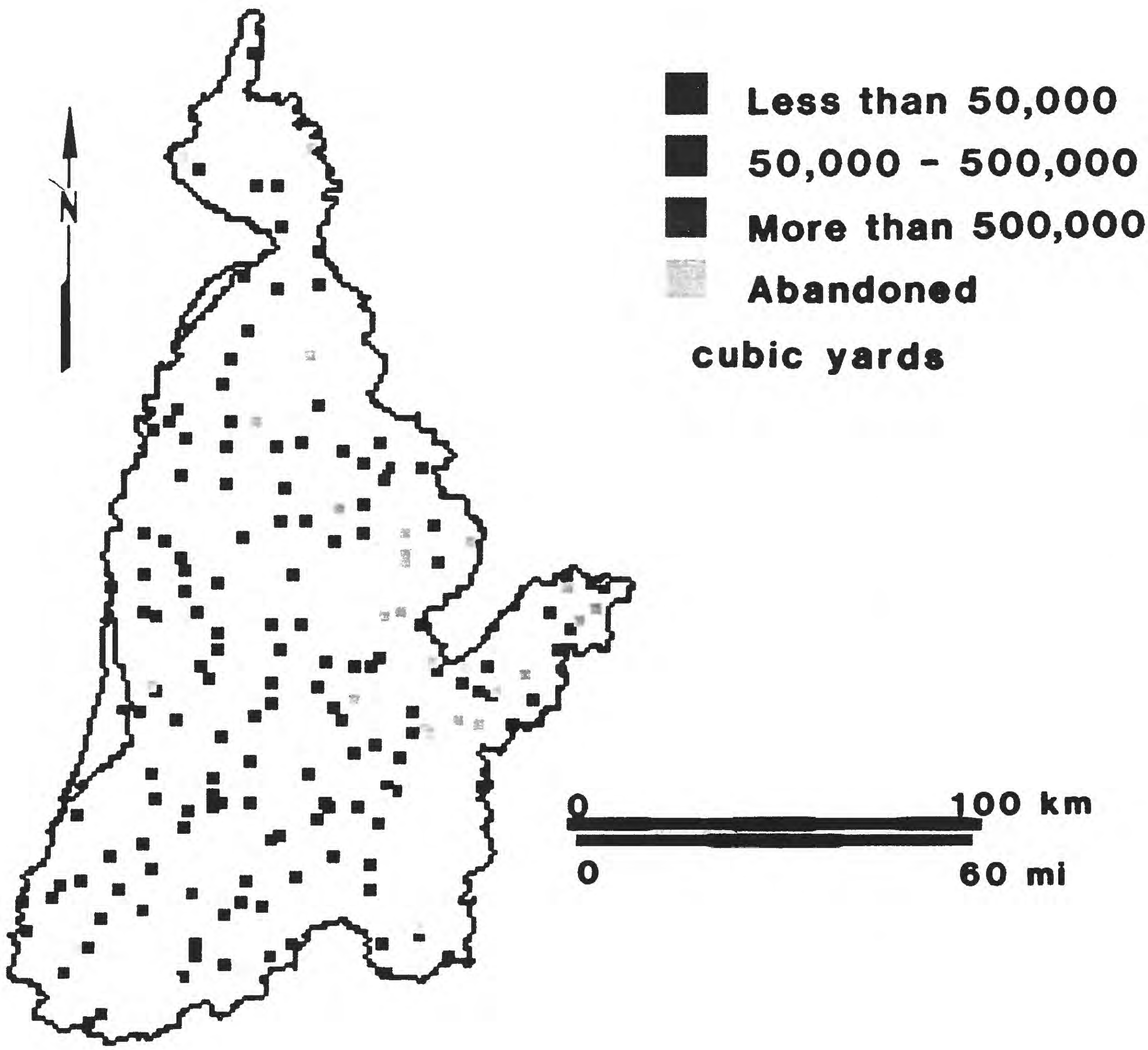
Processing:

1. Enter digital file as a text file.
2. Convert latitude and longitude values to line and sample coordinates in a Lambert Conformal Conic map projection.
3. Code the grid cell at each sampling location with the specific conductance value.
4. Group specific conductance values into four classes.
5. Use proximity algorithm to enlarge points for display.

Coding:

Class 0. Background.

1. 120–250 umhos/cm at 25°C.
2. 250–500 umhos/cm at 25°C.
3. 500–750 umhos/cm at 25°C.
4. 750–830 umhos/cm at 25°C.


N
Less than 50,000
50,000 - 500,000
More than 500,000
Abandoned
cubic yards
0
100 km
0
60 mi

20

Data Set:

LANDFILLS

Source:

1:500,000 scale Wisconsin Land Resources Analysis Program map.

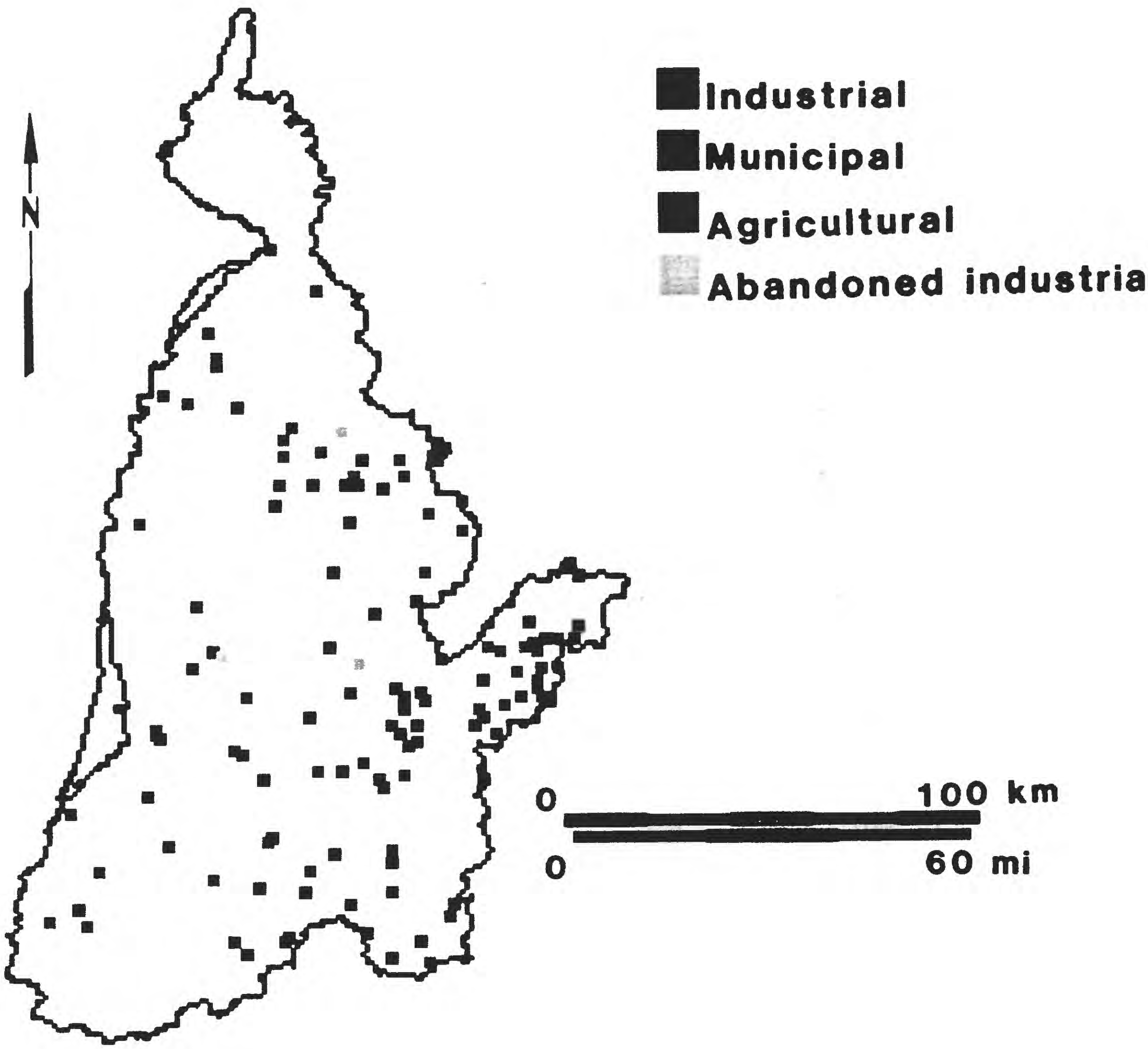
Processing:

1. Digitize points in vector form, convert to raster form, and register.
2. Use proximity algorithm to enlarge points for display.

Coding:

Class 0. Background.

1. Less than 50,000 yd^3.
2. 50,000-500,000 yd^3.
3. More than 500,000 yd^3.
4. Abandoned.



N
Industrial
Municipal
Agricultural
Abandoned industrial
0
100 km
0
60 mi

21

Data Set:

LIQUID WASTE DISPOSAL SITES

Source:

1:500,000 scale Wisconsin Land Resources Analysis Program map.

Processing:

1. Digitize points in vector form, convert to raster form, and register.

2. Use proximity algorithm to enlarge points for display.

Coding:

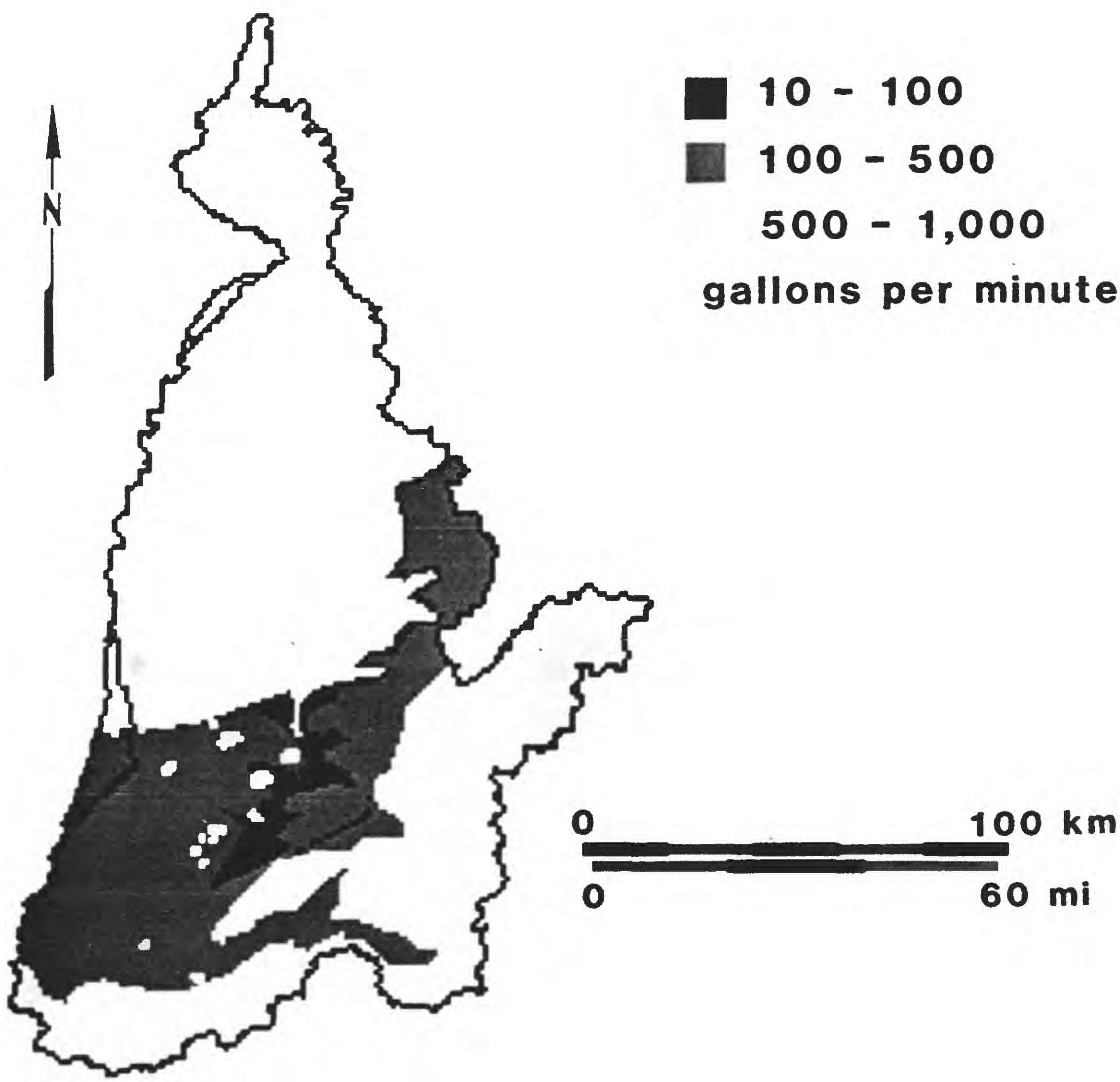
Class 0. Background.

1. Industrial.

2. Municipal.

3. Agricultural.

4. Abandoned industrial.



22

Data Set:

WELL YIELDS IN THE CAMBRO-ORDOVICIAN SANDSTONE AQUIFER

Source:

1:500,000 scale map overlay.

Processing:

Digitize polygons in vector form, convert to raster form, and register.

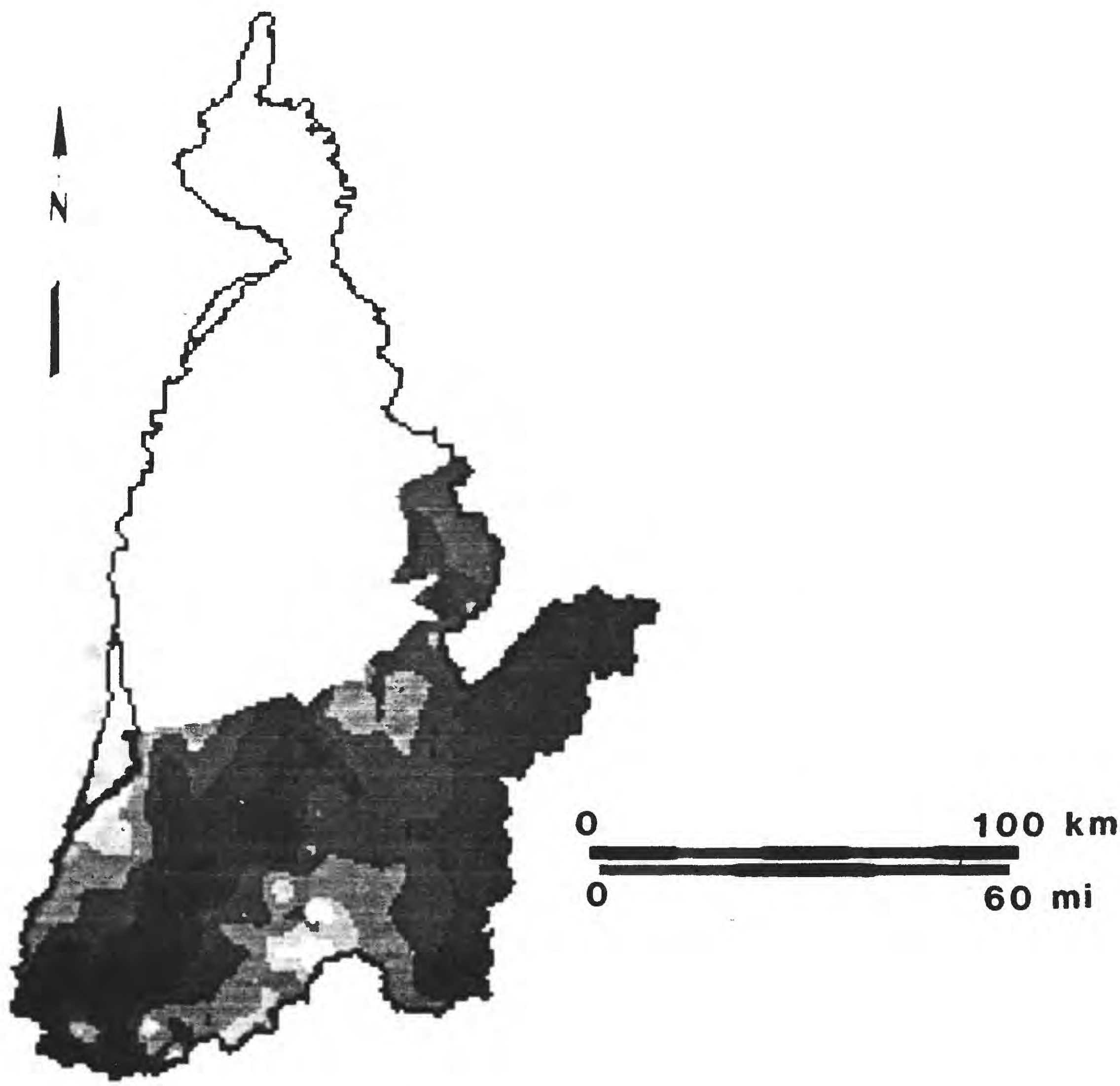
Coding:

Class 0. Background; aquifer absent.

1. 10-100 gal/min.

2. 100-500 gal/min.

3. 500-1,000 gal/min.



23

Data Set:

CONFIGURATION OF THE TOP OF THE CAMBRO-ORDOVICIAN SANDSTONE AQUIFER

Source:

Digital file containing latitude and longitude locations of wells and the associated elevations of the top of the sandstone aquifer.

Processing:

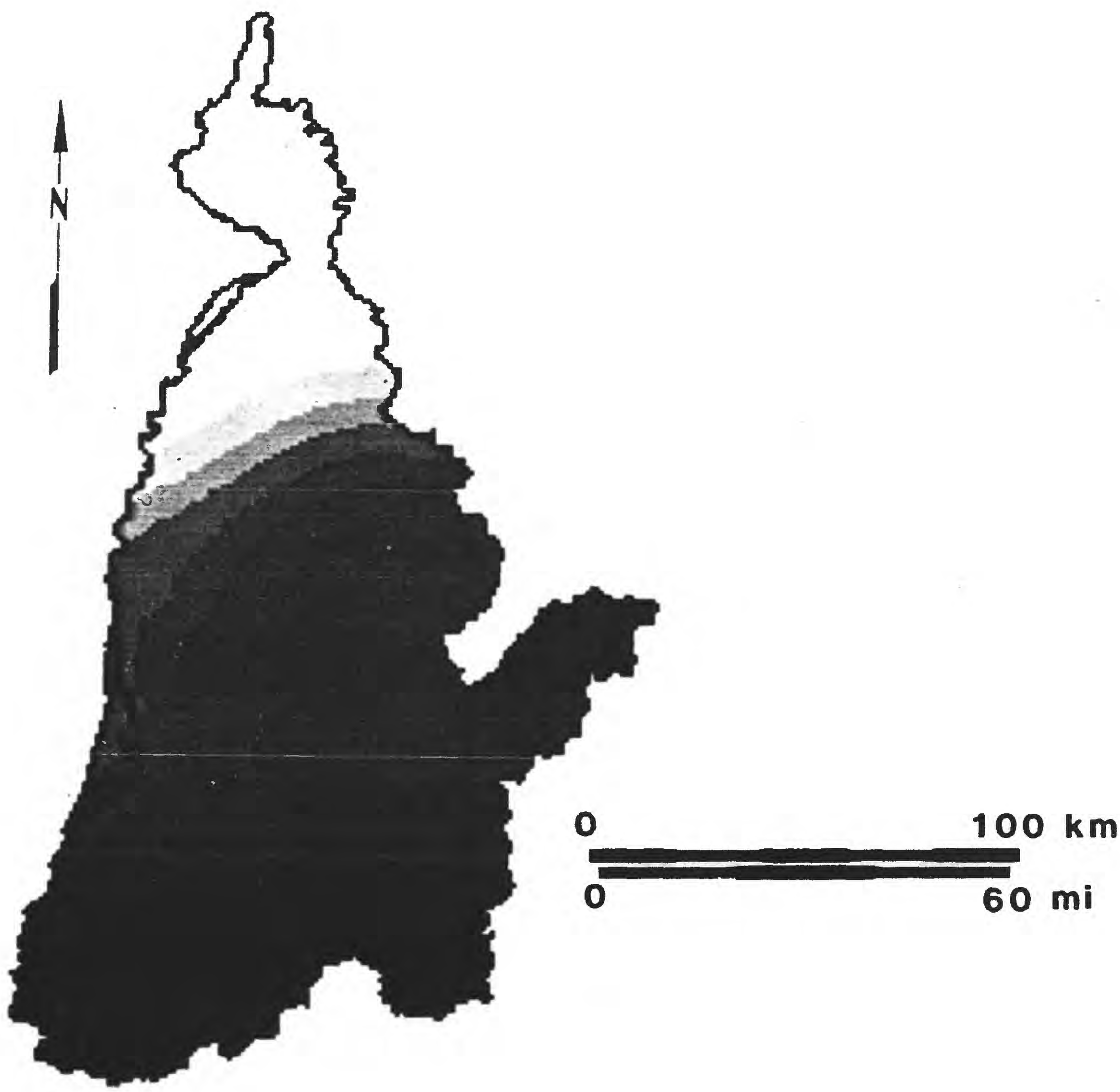
1. Enter digital file as a text file.
2. Convert latitude and longitude values to line and sample coordinates in a Lambert Conformal Conic map projection.
3. Code the grid cell at each well location with the associated elevation; select for use the first of any multiple wells at the same location.
4. Use a minimum curvature algorithm to generate a surface.
5. Use the ground-water-divide polygon as a mask to delete areas outside of the basin; use polygons from the well yield data set to delete grid cells outside the area of aquifer occurrence.
6. Level slice the elevation surface for display.

Coding:

Code each 100-ft elevation interval for display as a separate gray tone. The aquifer is absent in the white area. The lightest gray tone has an elevation of 1,600-1,700 feet above sea level (NGVD of 1929). The darkest tone has an elevation of 100-200 feet above sea level.

Remarks:

Data points are from the Ground-Water Site Inventory digital file and include reported values. The surface may not be correct in all areas.



24

Data Set:

AVERAGE ELEVATIONS OF WATER LEVELS IN WELLS, 1945-54

Source:

Digital file containing latitude and longitude locations of wells and the associated elevations of water levels.

Processing:

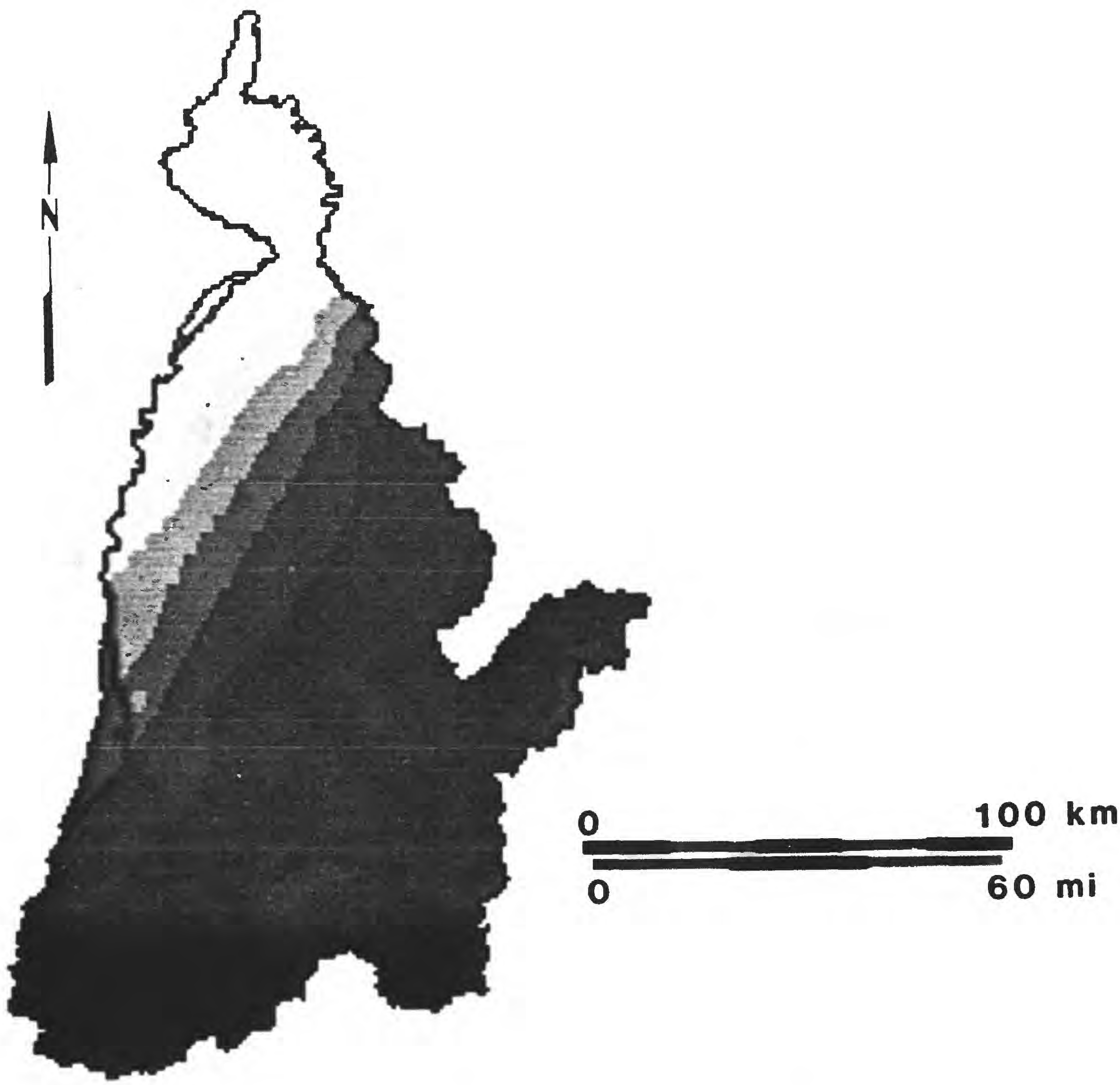
1. Enter digital file as a text file.
2. Convert latitude and longitude values to line and sample coordinates in a Lambert Conformal Conic map projection.
3. Code the grid cell at each well location with the associated elevation; select for use the first of any multiple wells at the same location.
4. Use a minimum curvature algorithm to generate a surface.
5. Use a 2 by 2 averaging algorithm to reduce data set resolution from 500 m to 1 km; use the ground-water-divide polygon as a mask to delete areas outside of the basin.
6. Level slice the elevation surface for display.

Coding:

Code each 100-ft elevation interval for display as a separate gray tone. The lightest gray tone has an elevation of 1,500-1,600 feet above sea level (NGVD of 1929). The darkest tone has an elevation of 300-400 feet above sea level.

Remarks:

Data points are from the Ground-Water Site Inventory digital file and include reported values. The surface may not be correct in all areas.



25

Data Set:

AVERAGE ELEVATIONS OF WATER LEVELS IN WELLS, 1975-83

Source:

Digital file containing latitude and longitude locations of wells and the associated elevations of water levels.

Processing:

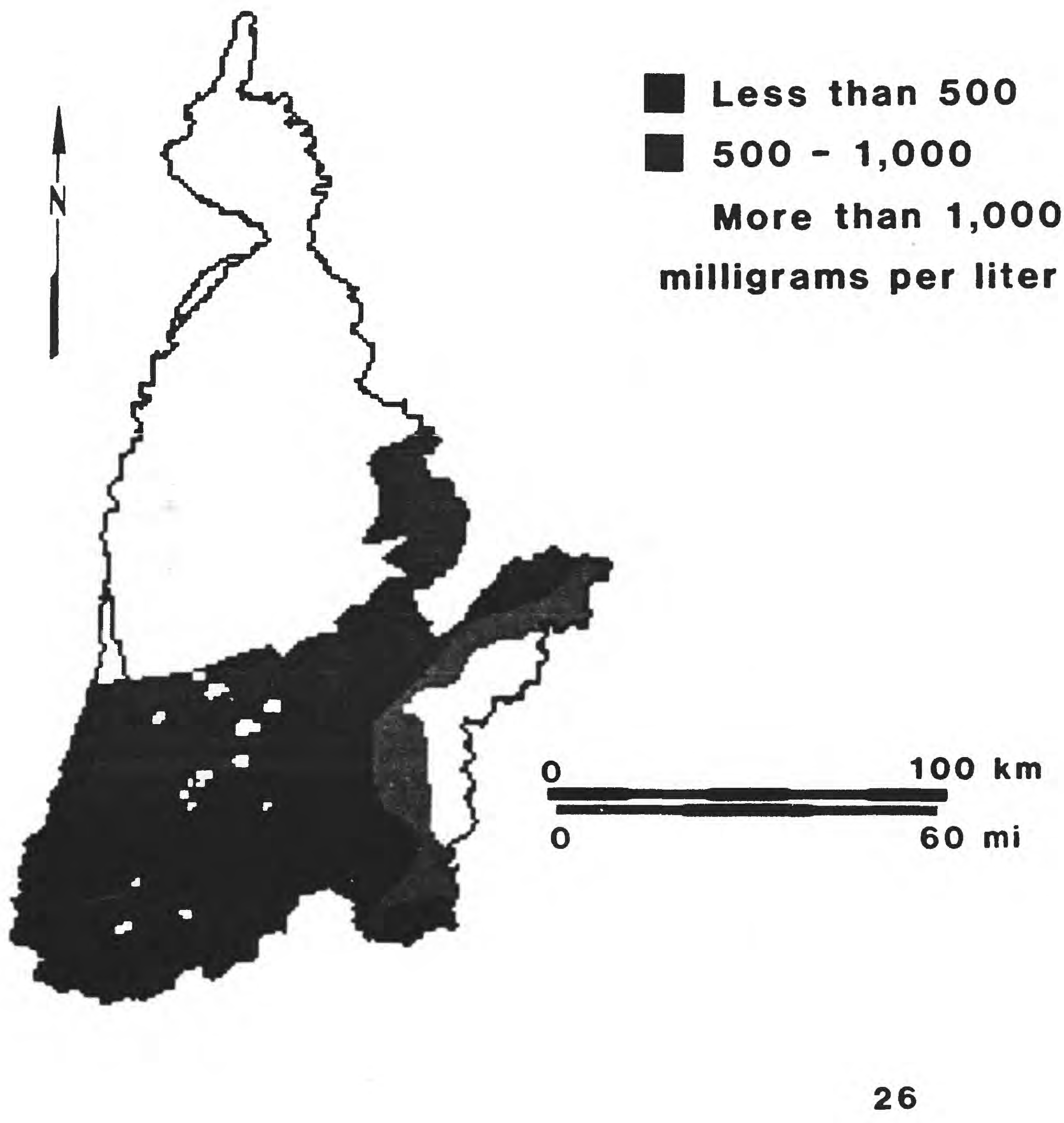
1. Enter digital file as a text file.
2. Convert latitude and longitude values to line and sample coordinates in a Lambert Conformal Conic map projection.
3. Code the grid cell at each well location with the associated elevation; select for use the first of any multiple wells at the same location.
4. Use a minimum curvature algorithm to generate a surface.
5. Use a 2 by 2 averaging algorithm to reduce data set resolution from 500 m to 1 km; use the ground-water-divide polygon as a mask to delete ares outside of the basin.
6. Level slice the elevation surface for display.

Coding:

Code each 100-ft elevation interval for display as a separate gray tone. The lightest gray tone has an elevation of 1,600-1,700 feet above sea level (NGVD of 1929). The darkest tone has an elevation of 300-400 feet above sea level.

Remarks:

Data points are from the Ground-Water Site Inventory digital file and include reported values. The surface may not be correct in all areas.


N
Less than 500
500 - 1,000
More than 1,000
milligrams per liter
0
100 km
0
60 mi
26

Data Set:

TOTAL DISSOLVED SOLIDS IN WATER FROM THE CAMBRO-ORDOVICIAN SANDSTONE AQUIFER

Source:

1:500,000 scale map overlay.

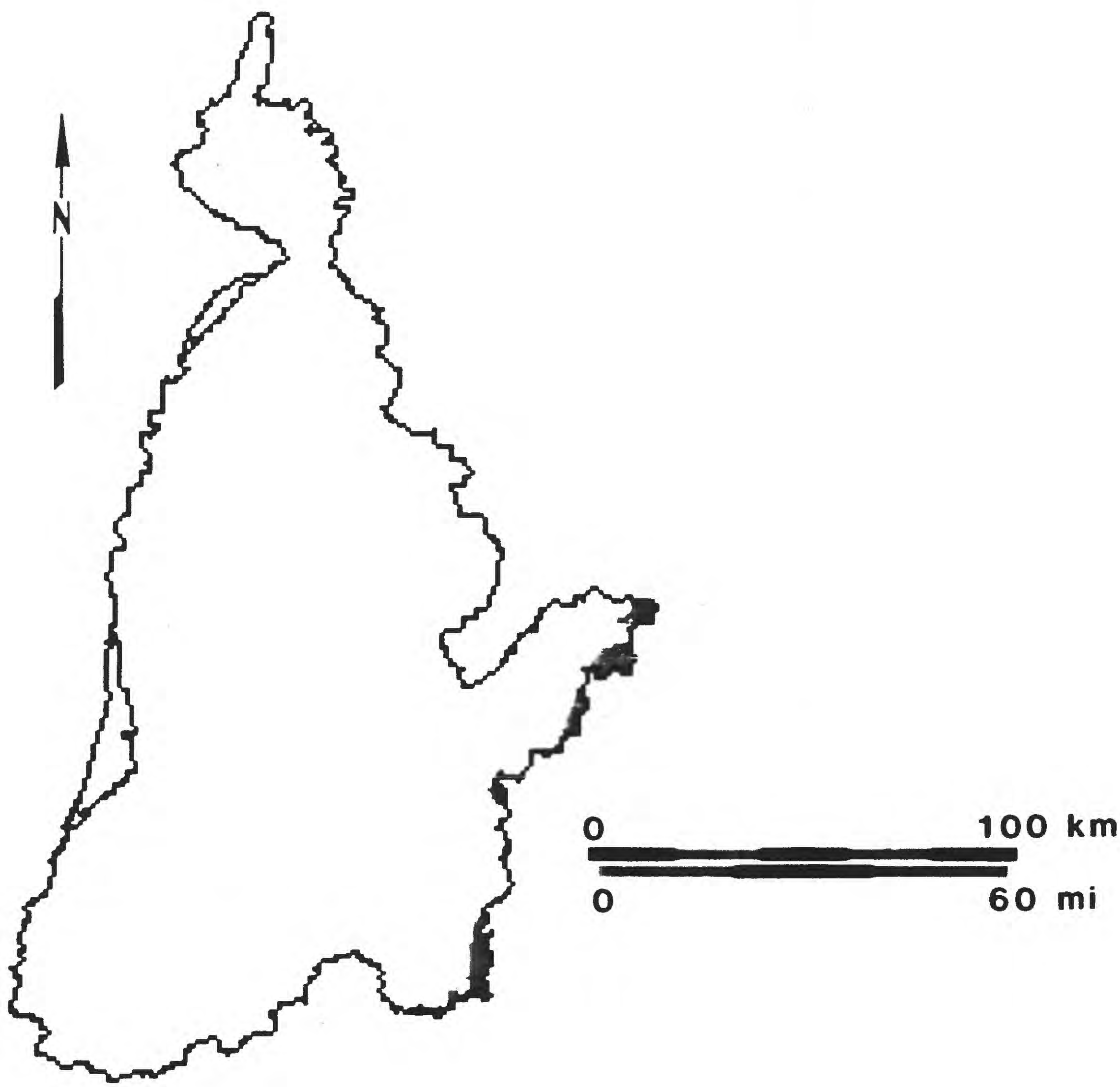
Processing:

Digitize polygons in vector form, convert to raster form, and register.

Coding:

Class 0. Background; aquifer absent.

1. Less than 500 mg/L.
2. 500-1,000 mg/L.
3. More than 1,000 mg/L.



27

Data Set:

WELL YIELD IN THE SILURIAN DOLOMITE AQUIFER

Source:

1:500,000 scale map overlay.

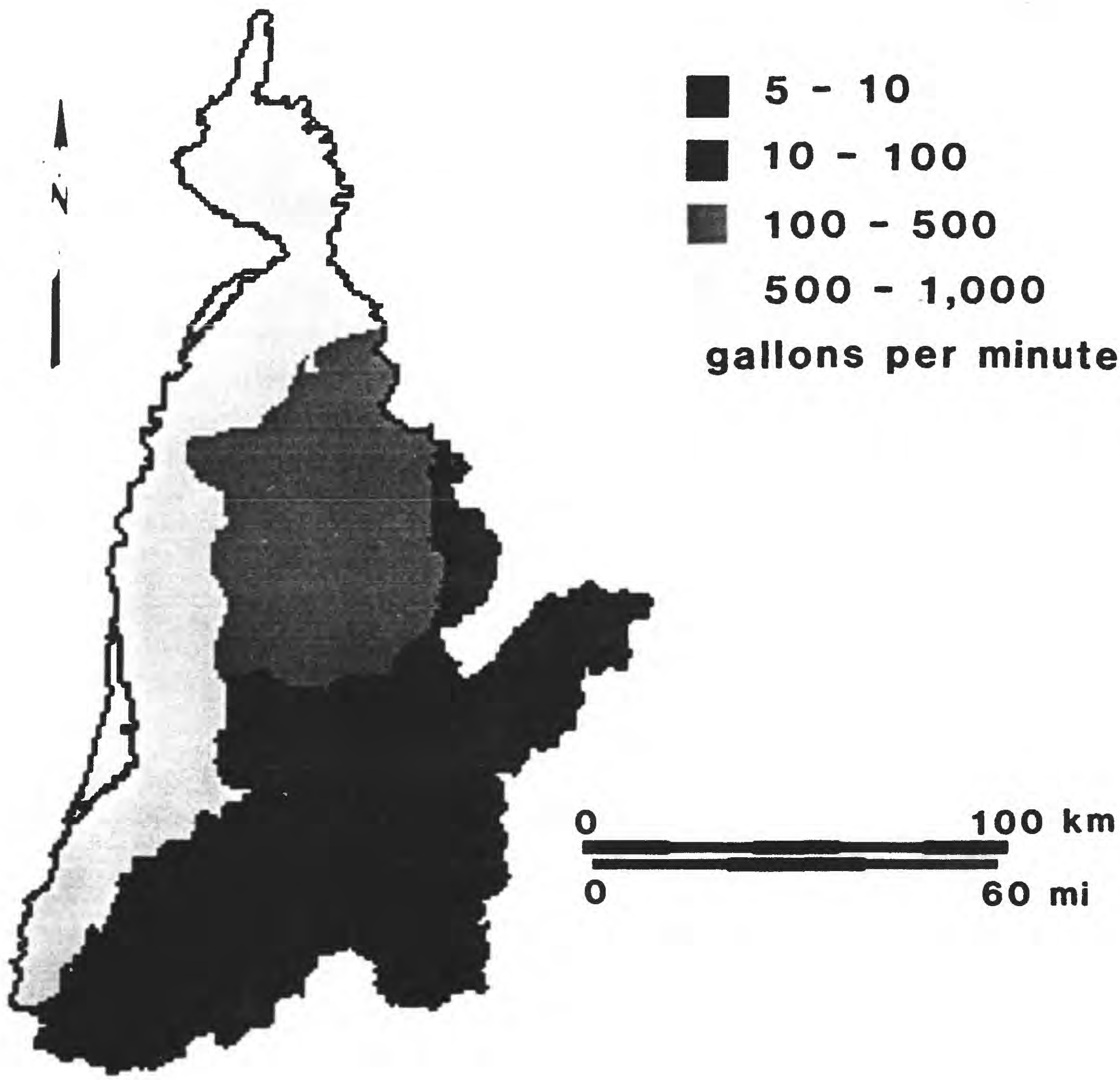
Processing:

1. Digitize polygons in vector form, convert to raster form, and register.
2. Use the ground-water-divide polygon as a mask to delete areas outside of the basin.
3. Code the polygons with the well yield value.

Coding:

Class 0. Background; aquifer absent.

1. 10-100 gal/min.



28

Data Set:

WELL YIELDS IN THE GLACIAL SAND AND GRAVEL AQUIFER

Source:

1:500,000 scale map overlay.

Processing:

Digitize polygons in vector form, convert to raster form, and register.

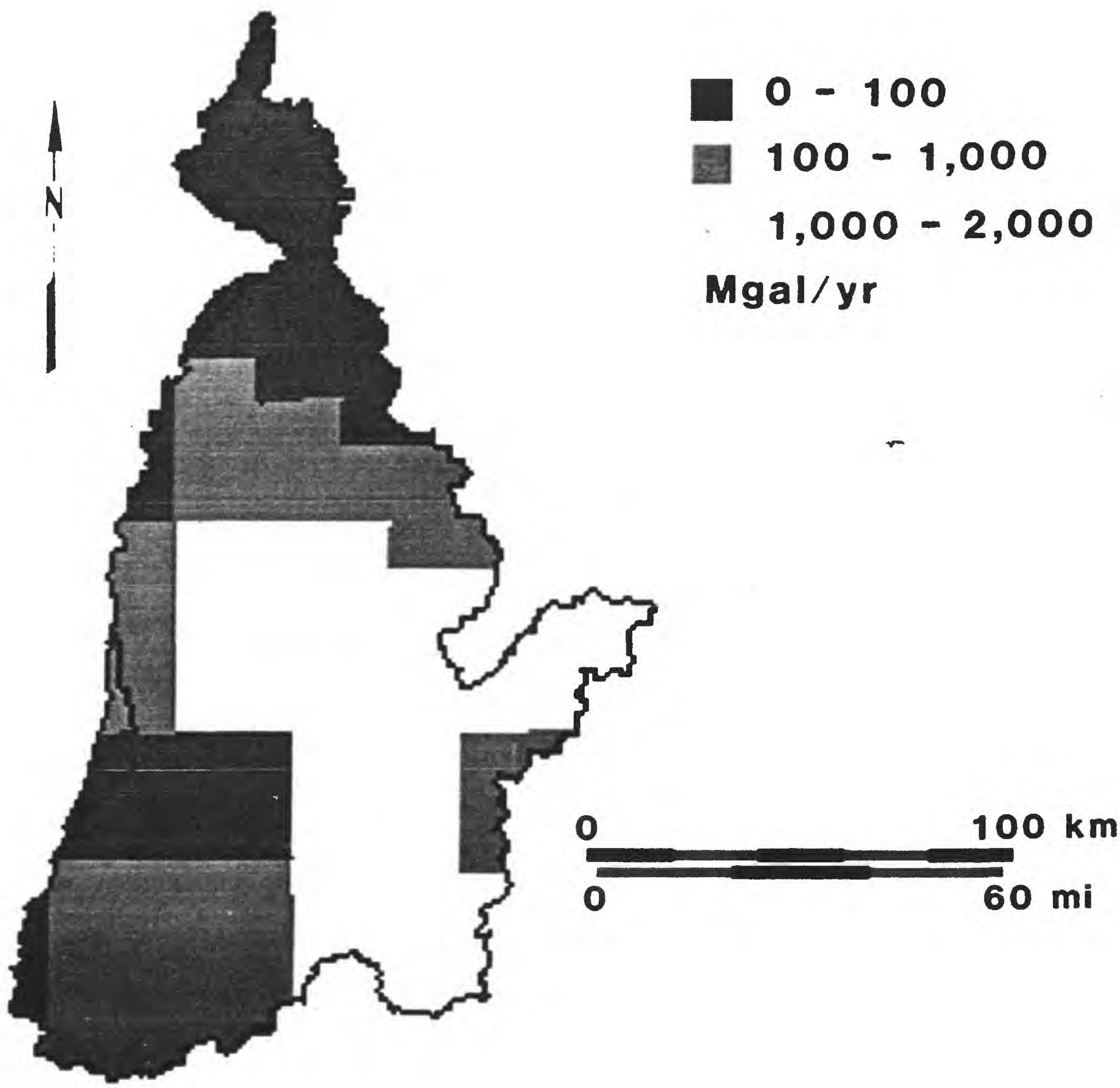
Coding:

Class 1. 5-10 gal/min.

2. 10-100 gal/min.

3. 100-500 gal/min.

4. 500-1,000 gal/min.



29

Data Set:

INDUSTRIAL AND COMMERCIAL GROUND-WATER USE BY COUNTY

Source:

Polygons from county line data set and 1:500,000 scale map overlay showing industrial and commercial ground-water use in each county.

Processing:

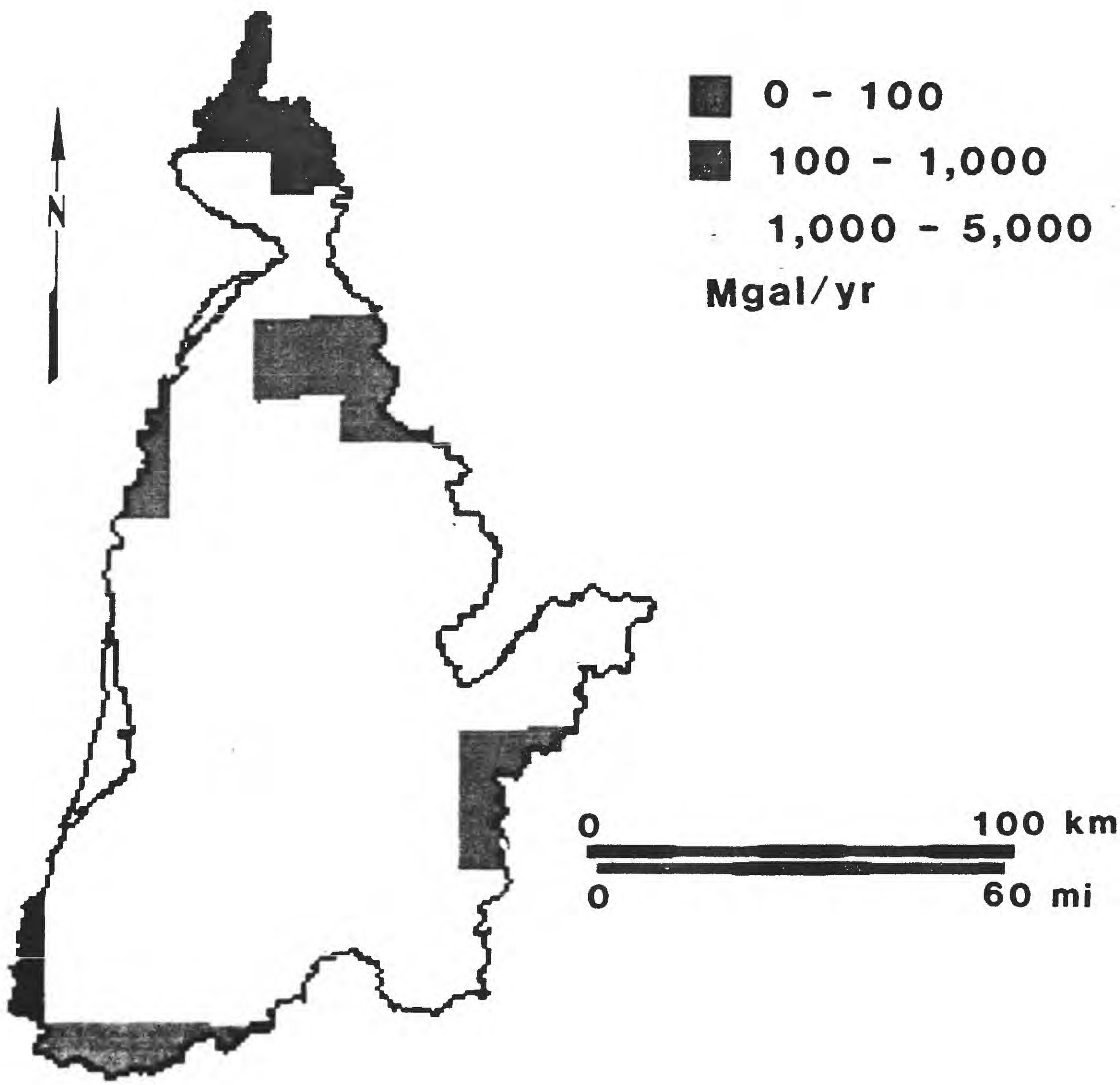
Code county polygons to indicate water use intervals (classes of water use.)

Coding:

Class 1. 0-100 Mgal/yr.

2. 100-1,000 Mgal/yr.

3. 1,000-2,000 Mgal/yr.



30

Data Set:

TOTAL GROUND-WATER USE BY COUNTY

Source:

Polygons from county line data set and 1:500,000 scale map overlays showing total ground-water use in each county.

Processing:

Code county polygons to indicate water use intervals (classes of water use.)

Coding:

Class 1. 0-100 Mgal/yr.

2. 100-1,000 Mgal/yr.

3. 1,000-5,000 Mgal/yr.