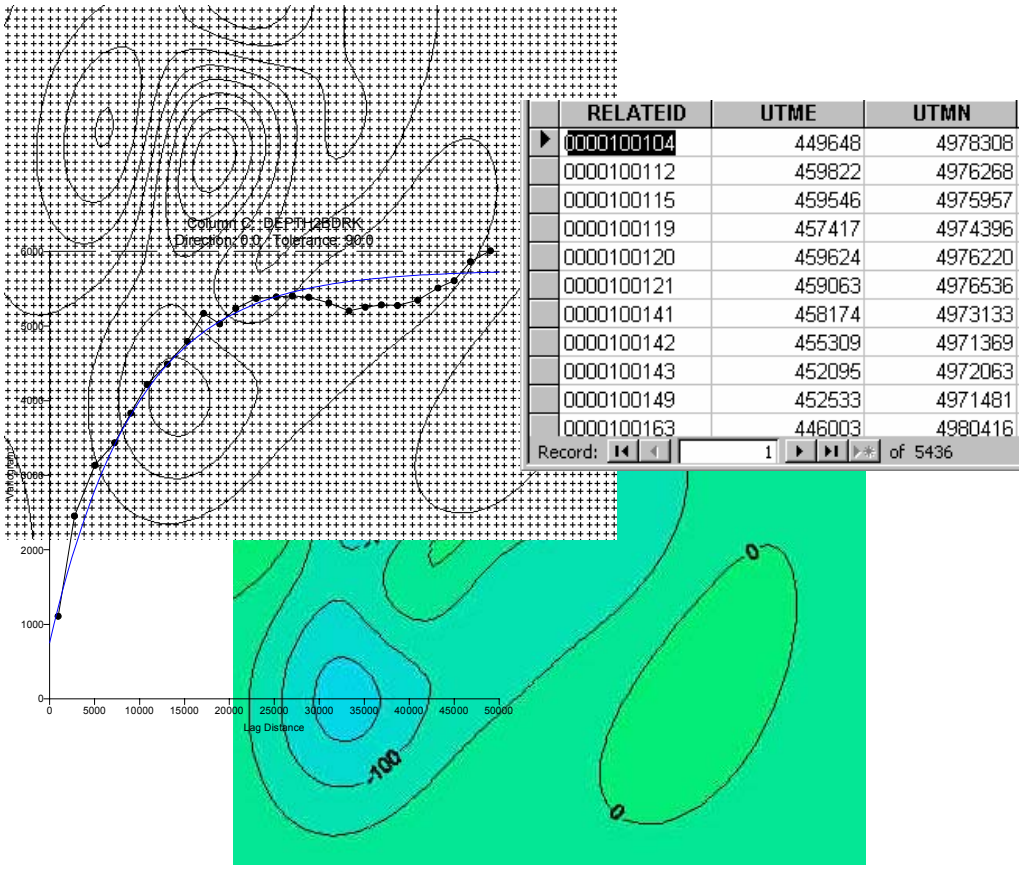
Metropolitan Area Groundwater Model Project

Preparation of Supporting Databases for the Metropolitan Area Groundwater Model

Version 1.00, August 2003

Andrew R. Streitz



RELATEID	UTME	UTMN
0000100104	449648	4978308
0000100112	459822	4976268
0000100115	459546	4975957
0000100119	457417	4974396
0000100120	459624	4976220
0000100121	459063	4976536
0000100141	458174	4973133
0000100142	455309	4971369
0000100143	452095	4972063
0000100149	452533	4971481
0000100163	446003	4980416

Record: 1 of 5436

Figures (continued)

Figure 9. Location of Wells Installed in 1987, Prairie du Chien-Jordan Aquifer, Hennepin County 19
Figure 10. Groundwater Surface in 1987 for the Prairie du Chien- Jordan Aquifer 20
Figure 11. Location of Groundwater Site in Hennepin County 21
Figure 12. Groundwater Surface as Viewed from the South, Edge-on 21
Figure 13. The Groundwater Surface of the Prairie du Chien-Jordan Aquifer in 3D 22
Figure 14. Number of Well Logs with Quaternary Entries by Elevation 25
Figure 15. Averaged Sand Content for Northwest Province 27
Figure 16. Averaged Sand Content for Northwest Province, by Intervals 28
Figure 17. A Comparison of Interpolation Techniques; Northwest Province, Interval 3 30
Figure 18. Study Area 31
Figure 19. Well Locations 31
Figure 20. Query Design in ACCESS 32
Figure 21. Variogram Analysis of Elevation Information 33
Figure 22. Surface Elevation 34
Figure 23. Thickness of Glacial Drift 35
Figure 24. Sand Content Values- Individual Borings 36
Figure 25. Sand Content Values- Interpolated to a Grid 37
Figure 26. Focused Analysis Area 37
Figure 27. Results of Focused Analysis 38
Figure 28. Ranked Residuals Outliers 40

Metropolitan Area Groundwater Model Project

Preparation of Supporting Databases For the Metropolitan Area Groundwater Model

Version 1.00, July 2003

Andrew R. Streitz

Introduction

Objective

This report presents an overview of the many databases that support the Twin Cities Metropolitan Groundwater Model (Metro Model), a computer model that simulates regional groundwater flow in the seven-county Twin Cities metropolitan area (Figure 1). The modeled area includes Anoka, Carver, Dakota, Hennepin, Ramsey, Scott, and Washington Counties.

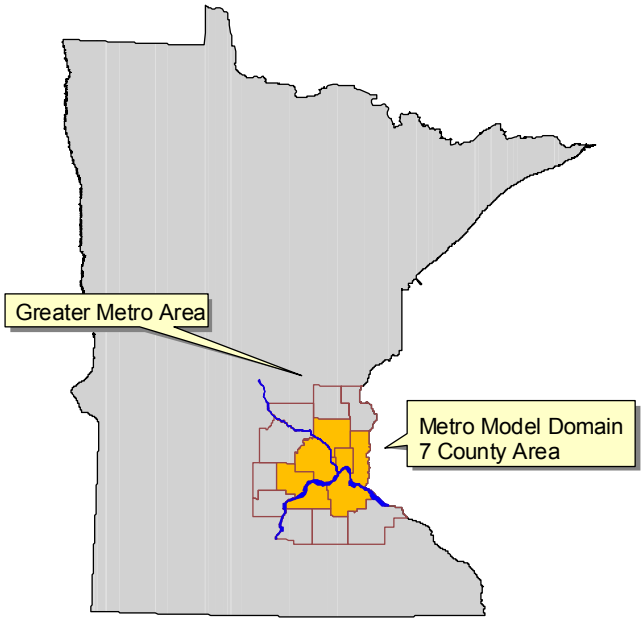


Figure 1. Database Coverage vs. Metro Model Domain

The Introduction presents background information regarding the Metro Model as well as the supporting data and information. It is followed by two main sections. The first presents descriptions of how the databases were created, and the second discusses the databases themselves.

Database Primer

Geographic Information Systems and Other Software Tools

Much of the work of the Metro Model project was handled in a Geographic Information Systems (GIS) environment. Specifically, the Metro Model team used ArcView, a proprietary GIS software package that is widely used among parties in Minnesota and the environmental field. Data and information that can be displayed within a GIS environment include bedrock geology, sand content of glacial drift, bedrock topography, thicknesses and surface elevations of the tops of selected bedrock layers, and model outputs including head calibration plots and piezometric surfaces. The GIS environment allows ready comparison of different location-oriented databases and coverages. For example, well and pumping test locations with hydraulic conductivity values may be superimposed on displays of geology or piezometric surfaces.

This software package was one of many used in the analysis of data for this report. ArcView can be used separately without interaction with other programs, but it was most often used to review datasets that were produced with some of the programs listed below. Most of the datasets presented in this report are the product of analysis by a full suite of proprietary programs used in sequence.

Proprietary software programs used for data handling, manipulation, and analysis are listed below, along with a brief description of how each was applied to the project. The mention of specific software packages in this report does not constitute an endorsement of a commercial product by the State of Minnesota or its employees.

Spreadsheet- EXCEL

Sorting and converting files between different formats, and graphing results. Able to produce a descriptive statistics analysis. Modifies datasets to facilitate transfer of data between analysis programs.

Database- ACCESS and ORACLE 8i

Dataset management, filtering, sorting, merging , and table generation. Able to produce specific filtered interpretations that can be exported directly into GIS.

Spatial Analysis- SURFER, VARIOWIN and GEOEAS

Variogram analysis, cross validation, dataset interpolation to a grid, contouring, and DXF graphic file generation. A quasi-GIS function, and a sophisticated spatial analysis set of tools for finding patterns in large datasets.

Three-dimensional Analysis- GMS and Savi3D

Dataset presentation and three-dimensional visualization.

The use of these software tools in developing the supporting databases will be described later in the report.

Central Role of Databases in Metro Model

Geologic and environmental databases played a critical role in the development of the Metro Model. A technically sound model can only be developed after a proper base of information is collected. One way to visualize this is to imagine a groundwater model as the apex of a pyramid made up of layers of supporting databases and assumptions (Figure 2). Each step up the pyramid is possible only because of the support provided from the step below, and each step up requires more sophisticated interpretations of the data.

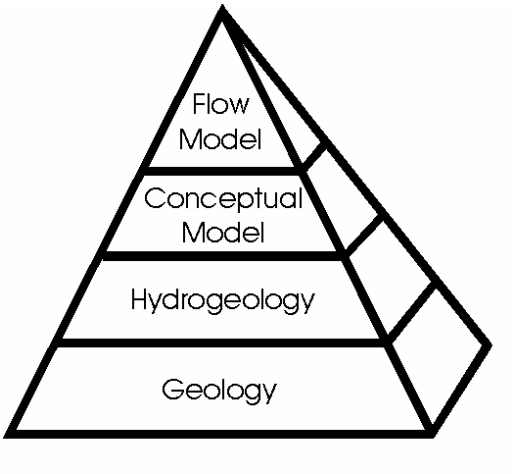


Figure 2. Groundwater Model Pyramid

Construction of a computer groundwater flow model, shown at the top of the pyramid, is based on a conceptual model of groundwater flow. A conceptual model may be described as a set of assumptions regarding groundwater flow expressed in words (Bear and Verruijt, 1990). As described in the Overview Report for the Metro Model (Seaberg, 2000):

> These assumptions include identification of the system's geometry, boundary conditions, type of flow, composition of the system, aquifer recharge and discharge zones, and hydraulic properties of the media. These assumptions represent a simplified perception of the hydrogeologic system intended to meet the objectives of the modeling effort by including only the features that are relevant to the questions being answered. Data and information used in the development of the conceptual model are approached from two fronts: 1) the hydrogeology is evaluated to identify features and processes likely to have a significant impact on the groundwater flow system; and 2) hydraulic head data are evaluated to ascertain indirectly the nature of interaction between hydrostratigraphic units and to help identify "hidden" hydrogeologic features that impact flow.

The geologic map and upper surface elevation maps of selected hydrostratigraphic units were used by Metro Model staff in the development of the conceptual model. Additionally, the upper surface elevation maps were used to construct isopach maps of selected hydrostratigraphic units.

Department of Natural Resources, Water Resources Data

Obwell Network

Since 1944, the Department of Natural Resources (DNR) Division of Waters has managed a statewide network of water level observation wells, more commonly known as the Obwell Network. Data from these wells are used to assess groundwater resources, determine long term trends, interpret impacts of pumping and climate, plan for water conservation, evaluate water conflicts, and otherwise manage the water resource. Soil and Water Conservation Districts under contract with DNR Waters measure the wells monthly and report the readings to DNR Division of Waters. The U.S. Geological Survey also monitors some wells using continuous recorders, and readings are also obtained from volunteers at other locations. Currently, about 700 wells are being monitored in the observation well network. The DNR Obwell Network can be accessed at:

http://www.dnr.state.mn.us/waters/groundwater_section/obwell/waterleveldata.html

These high-quality data gave Metro Model staff a clear picture of piezometric conditions in certain areas, and helped them to look at past and current conditions of aquifer water levels.

Department of Natural Resources, State Water Use System

SWUDS

The State Water Use Data System (SWUDS) is a database of high volume wells that are permitted through the DNR Division of Waters. This database may be explored further at:

http://www.dnr.state.mn.us/waters/watermgmt_section/appropriations/index.html

Metro Model staff used the SWUDS database to develop model datasets of high-capacity pumping wells.

Data Analysis Techniques

Overview

This section describes the data analysis steps that were taken to convert the source databases into the Metro Model supporting databases. Some databases were used without additional processing, such as the Minnesota Geological Survey (MGS) bedrock maps. However, most of the source databases required modification because they were incomplete or because there was a need to identify biases and spatial patterns.

Biased and Incomplete Data

The CWI database is a large, spatially well-distributed database, but it is in many ways incomplete, at least for the purposes of building a groundwater model. This is because the database is based on well logs filled out and submitted by water well drillers, who are hired generally in response to new domestic and commercial construction. Though this ensures that hundreds of new logs are entered each month this source of information brings with it several limitations.

The first problem is location bias. The vast majority of wells in CWI are drilled for new residential homes, which in Minnesota means they are drilled within the rapidly expanding Twin Cities metropolitan area that stretches from Rochester to St. Cloud. Thus the spatial coverage of new data is denser in the metropolitan region than in Greater Minnesota. Even within this region development does not occur evenly, but is focused in specific townships and cities depending on growth-limiting factors such as highway expansions and changes in the Metropolitan Council's Metropolitan Urban Service Area (MUSA). Consequently, a large number of records might have been submitted from the west metropolitan area in 1978, while in 1999 a larger number might come from the east metro. For example, analyzing CWI by date of entry might provide a valuable review of residential development in the area, as first one county and then another receives hundreds of new homes. This time and location bias is not fatal to the use of the data, but is something that must be understood by potential users.

A second problem is location error. Wells are commonly reported with an incorrect location because the forms are submitted in the confusing noncartesian Township-Range-Section (TRS) system. As budgets allow, the MGS physically field locates individual wells, using Global Positioning System (GPS) technology to generate a UTM location coordinate. One principal advantage of the UTM system over the TRS is that the former is GIS ready (see previous section on coordinate systems). Because the Metro Model supporting databases are based on the field-located data and not the larger TRS dominated data, this location error does not apply. Outside of the Metro region however, the error associated with TRS-collected data would remain a problem. This of course remains a problem even if TRS coordinate locations are converted to the UTM coordinate system via a program such as SECTIC or MNCon.

The terms “many” and “well distributed” are not precise, and can only be determined in context. Generally, the sample number should be greater than 50 and the data points should fill the area of interest without big gaps. Finally, the points should be close enough together so as to produce a variogram (defined starting on p.13) with values that can be correlated. An example of a large, well-distributed dataset is displayed in Figure 3. Dakota County (outlined in red) shows good coverage for bedrock wells (almost all from wells in the Prairie du Chien-Jordan aquifer) in the display on the left, but inadequate data coverage for a countywide analysis of glacial drift wells in the display on the right. In the same way that a dataset mean gains power from a larger sample size, so too does an analysis based on the variogram gain from a spatially well distributed dataset.

As an illustration of the differences between statistical measures that use and don’t make use of location, consider the difference between determining the height of children in a school versus the height of plants in a field. For the children it is reasonable to disregard the location of individuals when calculating the average height of students in a classroom since the location of students within the classroom is certainly not fixed, and is therefore irrelevant to the measurement. But if you were analyzing the height of plants growing in a field then location could be critical to understanding the processes that led to different growth rates. (Even location can be brought to bear on the height of children. A correlation between height and home address that is random within a city may be meaningful across a country or between countries in the world.)

None of the shortcomings of the CWI database discussed previously render the data useless. On the contrary its size and relatively even spatial distribution over most of the state make it an ideal candidate for trend and spatial analyses. Earlier in this report specific examples were presented of different kinds of errors that exist in CWI. The same list will now be discussed in light of the use of data analysis techniques that can provide greater confidence in use of the data.

The first error was the location bias that stems from a spatially uneven collection of new data points. Specifically this could cause problems for analyses if there were significant areas within the region that were not sampled. Figure 3 shows that CWI data has accumulated in a spatially well distributed fashion over the last 30 years. Based on an analysis of CWI water elevations, trend analyses of the change in groundwater head over time for the unconsolidated Quaternary material (Layer 1) and the Prairie du Chien-Jordan aquifer (Layer 3) show no discernable patterns significant enough to affect the Metro Model. Within these aquifers there are local areas where groundwater levels are rising or falling.

If there is any area under-represented in the database it is in the highly urbanized center of the metropolitan area. This is an area that is both dependant on wells drilled before geologic logs were routinely recorded and incorporated into CWI, and areas served predominantly by public supply systems, not individual domestic wells. Estimates of parameters of interest, such as head measurements, can be made in this area, but where the data are sparse the confidence in the estimates is low.

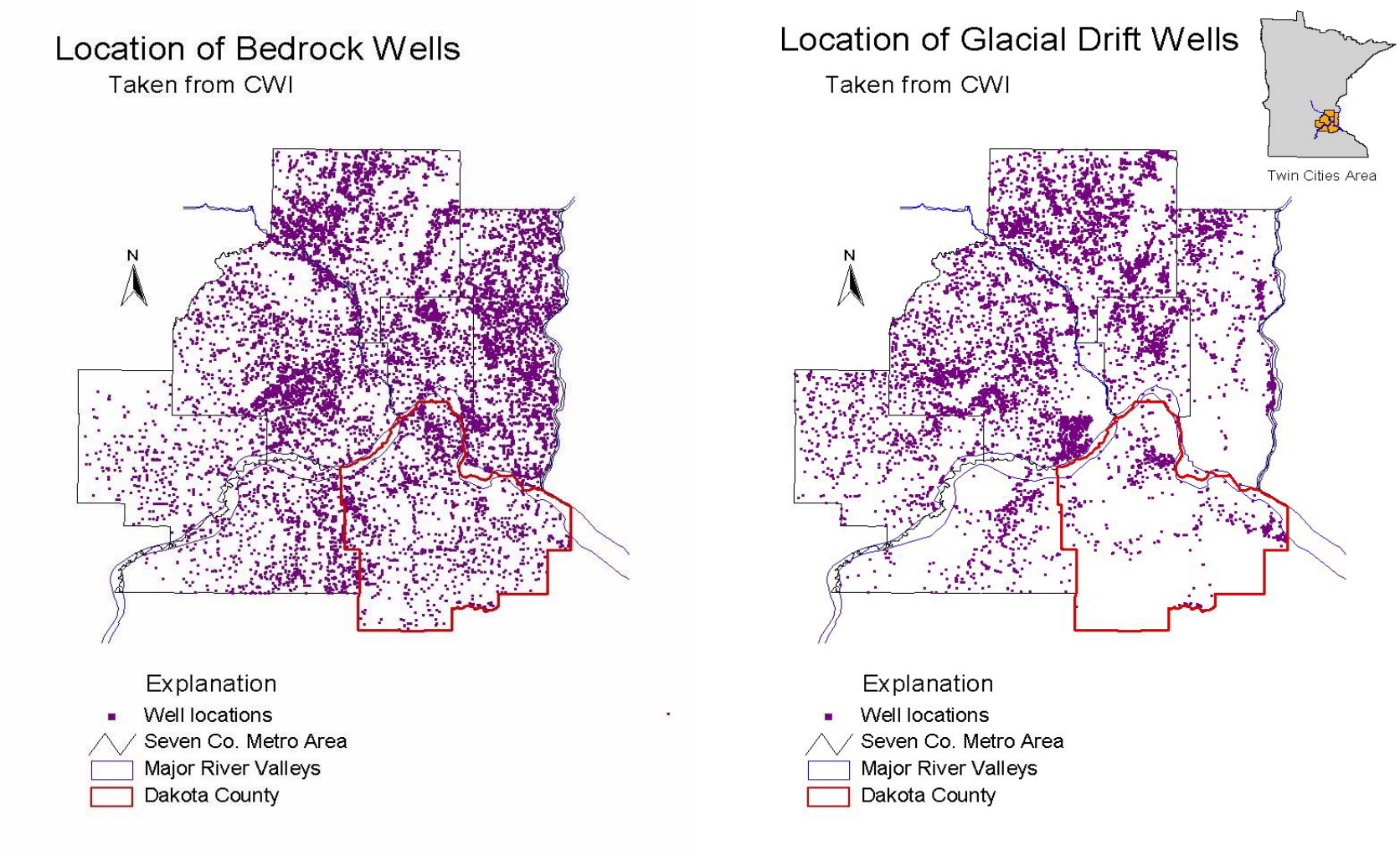


Figure 3. Location of Wells, Completed in the Bedrock and the Glacial Drift, Respectively