ADVANCED WATER DISTRIBUTION MODELING AND MANAGEMENT

Authors
Thomas M. Walski
Donald V. Chase
Dragan A. Savic
Walter Grayman
Stephen Beckwith
Edmundo Koelle

Contributing Authors
Scott Cattran, Rick Hammond, Kevin Laptops, Steven G. Lowry,
Robert F. Mankowski, Stan Plante, John Przybyla, Barbara Schmitz

Peer Review Board
Lee Cesario (Denver Water), Robert M. Clark (U.S. EPA),
Jack Dangermond (ESRI), Allen L. Davis (CH2M Hill),
Paul DeBarry (Borton-Lawson), Frank DeFazio (Franklin G. DeFazio Corp.),
Kevin Finnan (Bristol Babcock), Wayne Hartell (Bentley Systems),
Brian Hoefer (ESRI), Bassam Kassab (Santa Clara Valley Water District),
James W. Male (University of Portland), William M. Richards
(WMR Engineering), Zheng Wu (Bentley Systems ),
and E. Benjamin Wylie (University of Michigan)
Integrating GIS and Hydraulic Modeling

A geographic information system (GIS) is a powerful configuration of computer hardware and software used for compiling, storing, managing, manipulating, analyzing, and mapping (displaying) spatially-referenced information. It integrates database operations such as data storage, query, and statistical analysis with visual and geographic analysis functions enabled by spatial data. A GIS can serve as an integral part of any project that requires management of large volumes of digital data and the application of special analytical tools.

GIS is becoming an increasingly valuable tool for the water distribution modeler both as a source for modeling data and as a decision-support tool. Anderson, Lowry, and Thomte (2001) reported that although approximately 15 percent of water utilities currently use GIS in their modeling, almost 80 percent plan to use GIS in the future.

In the past, most models were developed with a batch-run philosophy in which a text file (or card deck) containing the model data was required to drive the model. Gradually, models gained the ability to interact with databases that were either internal to the model or more general and commercial. Finally, because a GIS is essentially a spatially aware database, ongoing development has led to models that are highly integrated with GIS. Shamsi (2001) describes the evolution of model/GIS integration as a three-step process:

1. **Interchange**: Data are exchanged through an intermediary file, which may be an ASCII text file or a spreadsheet. Data is written to this intermediary file, where it is reformatted for the model, if necessary, and then read into the model. The model and GIS are run independently.

2. **Interface**: Links are built between the model and GIS. These links are used to synchronize the model and GIS. The data are duplicated on each side of the link.
and the model and GIS are run independently. One common approach is the use of shapefiles, which can pass data between the model and the GIS and optionally update either based on data contained in the other.

3. **Integration**: A single repository for the data is used. The model can be run from the GIS and vice versa.

This integration of the hydraulic model and the GIS leads to the following benefits:

- Time-savings in constructing models
- Ability to integrate disparate land use, demographic, and monitoring data using GIS analysis tools to more accurately predict future system demands
- Visual, map-based quality control of model inputs
- Map-based display and analysis of model outputs in combination with other GIS layers

The most powerful feature of a GIS, from a planner's perspective, is probably the ability of the GIS to integrate, through their spatial relationships, databases that would be difficult or impossible to integrate outside of a GIS environment. For example, a GIS can overlay soil data, repair data, and hydraulic modeling output to automatically assign a condition rating to pipes.

This chapter contains background information on GIS development and uses for both water model creation and other application areas. The experienced GIS practitioner with a specific interest in applications of GIS technology to model development may want to begin with Section 12.3, “Model Construction.” The experienced modeler with access to GIS, but with no GIS experience, may want to begin with Section 12.1, “GIS Fundamentals,” and then skip ahead to Section 12.4, “GIS Analysis and Visualization.”

## 12.1 GIS FUNDAMENTALS

An easy way to think of GIS is to imagine it as a set of transparencies that are layered in such a manner that any point in one layer would appear at the same location in any other layer, as shown in Figure 12.1. In an actual GIS graphical user interface (GUI), these layers appear together, and the user can manipulate the order in which they appear.

Within a GIS, *features* (objects on a map) are not simply points and lines; they have *attributes* (information about the feature) associated with them. In a water distribution system, facilities such as pipes, tanks, and pumps are features possessing attributes. For instance, a pipe is represented in a GIS as a feature, and the diameter of the pipe is an attribute of that feature.

As shown in Figure 12.1, maps may contain more than one type of feature, each of which is displayed as a layer on the GIS map. By selecting which layers are displayed, the order in which layers are displayed, and the *symbology* (size, shape, and color of symbols), the user can control the appearance of the resulting map. Figure 12.2 shows a GIS map.
In addition to being used for map-making, a GIS can be used to perform system analysis, answering questions about:

- Location (using proximity, buffer, or overlay analysis)
- Condition
- Temporal and spatial patterns (trends)
- What-if scenarios (in modeling)
Two primary and opposing data management paradigms are in use today: centralized data management and decentralized data management. The mainframe computing environment is an example of centralized data management, and the PC environment is an example of decentralized data management. In the mainframe environment, all applications and data reside on a central server. This data management approach is very practical, but the hardware and software are typically very expensive to develop and maintain. Within the PC environment, special-purpose applications and databases are less expensive than their mainframe counterparts, but data and applications can reside on different networked PCs. The decentralized management philosophy is therefore very practical due to the economics involved; however, it does lead to the creation of “data islands.”

The data developed within these “islands” are often generated and maintained redundantly. For example, the diameter (6 in./150 mm), length (250 ft/76 m), and material (ductile iron) for a pipe can be entered into the hydraulic modeling application, asset management system, and maintenance management system, and the pipe can be drawn and annotated on a map. Very few links to a master database or other data islands are ever developed. As a result, the utility is unable to tap the knowledge and efficiency that can be gained by analyzing and acting upon this information centrally. Unfortunately, the data management situation of many water utilities around the world is characterized by data islands.

The computer industry has created technology such as SQL and ODBC to assist in centralizing these data islands. However, because they were designed and developed
independently, they frequently do not have the key identifiers needed to form meaningful data relationships. For example, the billing system may use account number as its primary identifier, and the hydraulic modeling software may use pipe and node numbers. It is therefore difficult to relate the billing information to the hydraulic model for use in customer service and system planning.

What do the systems and databases for billing, customer service, asset management, work management, inspections/permits, water quality testing, facility mapping, hydraulic modeling, and document management have in common? The answer is geography. All of these systems have information about items (for example, permits, work orders, test reports) that can be tied to a geographic location such as parcel number, address, or facility number. Geography is the essence of GIS-centric data management, a compromise between centralized and decentralized data management; features that cannot be linked explicitly through database table-to-table relationships can be associated geographically by determining the proximity (connectivity, distance, closeness) to each other. This aspect of GIS uniquely qualifies it as the preferred integrating technology and unifying information resource within an organization.

The compromise that GIS attains between centralized and decentralized data management is this: not all data needs to be centralized; only the GIS layers need to be centralized. Therefore, the only limitation to the integration of specialized systems into the GIS-centric model is that they must have some reliable means of geographic referencing, such as facility ID, parcel number, or street address. Virtually any type of system or database can be linked to a GIS layer, provided that a geographical association and consistency and quality of data are present in both systems. Usually, expanding the utility of a GIS system is primarily a data development and quality control exercise.
**Geographic Data Models**

Three representations have emerged to handle most geographic data: raster, vector, and TIN. The following list describes these representations, which are shown in Figure 12.4.

- **Raster:** This representation stores data as discrete grids in which a single value for each attribute is associated with each grid cell. Each grid cell has an attribute value and location coordinate. Because the data is stored in a matrix, the coordinates for each cell do not need to be stored explicitly (ESRI, 2001). Rather, they can be determined “on the fly” because the origin of the raster model, the grid cell size, and the rotation are known.

- **Vector:** This representation stores discrete features as points, lines, or polygons. Each feature has attribute values and a set of $x$-$y$ coordinates (and possibly a $z$ coordinate) associated with it. In this way, vector data differ from raster data, which have coordinates intrinsically associated with the cells (ESRI, 2001).

- **Triangulated irregular networks (TIN):** TINs divide space into a set of contiguous (non-overlapping) triangular faces. The triangular faces are derived from irregularly spaced sample points, breaklines, and polygon features, and each sample point has coordinate $(x, y)$ and attribute ($z$ coordinate or other attribute to be modeled) information associated with it. Attribute values at other locations (along breaklines and polygon features, or even on the triangular faces themselves) are calculated using interpolation within the TIN. TINs are sometimes considered a special case of a vector model, but Zeiler (1999) presents them as a separate data model.
Most data used in hydraulic modeling are vector data. For example, junction nodes are points, pipes are lines, and node service areas are polygons. Although most modeling data is made up of vector features, modelers also use raster and TIN data for tasks such as extracting elevation data or using an aerial photo as a background for the model.

12.2 DEVELOPING AND MAINTAINING AN ENTERPRISE GIS

Many publications describe the process of developing a GIS, including Orne, Hammond, and Cattran (2001) and Przybyla (2002). The pace of technologic advancement and the diversity of commercial implementations outstrip the ability of any general text to provide detailed guidelines on the configuration and management of a specific GIS-based modeling system. Consequently, this section provides a general discussion of the GIS development process and outlines the major steps involved in developing a GIS.

GIS is generally implemented at one of four levels:

1. **Project**: Supporting a single project objective
2. **Departmental**: Supporting the needs of one department
3. **Enterprise**: Inter-departmental sharing of data that meets the needs of many departments
4. **Interagency**: Sharing of application and data with external agencies

Building a GIS on a project basis for the sole purpose of using it with a hydraulic model is quite rare because an organization derives a variety of benefits from a GIS, including:

- Elimination of redundant data maintenance activities
- Streamlining of workflow processes with GIS functions
- Improvements in access to quality information
- Ability to use spatial analysis to solve problems

For these reasons, a GIS is often implemented at the departmental or enterprise level. The subsections that follow discuss some of the key steps that should be taken to ensure a successful enterprise-level GIS implementation. For more information on planning, developing, and maintaining a GIS for modeling and other applications, see the references listed at the end of this chapter or the GIS section of this book's bibliography (see page 727).

**Keys to Successful Implementation**

A community or water utility that commits to developing a GIS must consider several key factors:
- Development of a GIS that is capable of supporting a hydraulic model requires a high level of data quality, accuracy, and detail.

- Creation of a GIS necessitates a review of existing hardware and software and does not imply that legacy systems will go away. Systems such as a customer information system (CIS), CMMS, and SCADA are usually fixtures on the landscape.

- Enterprise GIS development calls for a high level of interdepartmental cooperation. All departments involved in enterprise GIS development should share the same vision for the system. The "people" aspect of GIS development can be more challenging than the technological aspects.

- Enterprise GIS development demands a GIS leader — someone to champion the effort. The GIS leader must build communication bridges between departments so that people talk, cooperate, and share.

- The largest expense of the GIS project is usually the data development effort, which often requires expensive conversion from paper map sources. If an organization fails to consider all of the cartographic, asset database, and hydraulic modeling needs during the database development, then its GIS system will either (a) fail to meet future application demands or (b) necessitate an expensive data upgrade effort in the future. For example, the GIS must be able reproduce the map sources used in its creation, or all that will have been accomplished is the creation of a new, redundant dataset to maintain.

**Needs Assessment**

Critical to a successful GIS implementation is a detailed understanding of the business processes and operational and management needs of the organization. An understanding of the specific GIS functions required by the individual users is also crucial. This information is gathered through a needs assessment.

A needs assessment has three components:

1. **User needs assessment**: The purpose of this assessment is to answer the following questions: who are the people that are going to use the system; what roles do they fulfill; what task or functions do they need to accomplish with the GIS; what skill levels do they possess; where (physically) are they; and how often will they use the GIS applications.

2. **Data source assessment**: The data source assessment determines what data sources are available to support the GIS data development, including their formats (e.g., electronic, paper), geographic extents, spatial \((x, y, z)\) and attribute accuracies, update frequencies, and dates last updated.

3. **System design assessment**: The purpose of this assessment is to determine the types, locations, and characteristics of existing servers, individual workstations, and computer networking components. This includes operating system platforms, current applications being served, and current levels of utilization.
The needs assessment is an essential part of creating any IT system and is usually accomplished by conducting detailed surveys and interviews with all potential GIS users (including hydraulic engineers) and the appropriate organization decision-makers. It can also be accomplished as part of a business process workflow analysis (recommended).

A thorough needs assessment is crucial for developing a GIS that will provide adequate service now and in the future. It should reveal specific problems or constraints associated with present systems and identify project implementation requirements. A properly conducted needs assessment will culminate in an integrated GIS with the following benefits and advantages:

- Increased operational and management efficiency and staff productivity
- Better sharing of data
- Quicker access to quality and timely information
- Full leveraging of the capabilities of the system
- Support for organizational operations that reflects the organization's mission and priorities
- Staff endorsement and regular use
- Immediate value to the organization
- Functionality that supports all current and future needs

**Design**

The second phase of the GIS development process is the design, which potentially includes the following tasks:

1. **Application design**: Describes the commercial software to be deployed and the custom programs that will be combined to create the applications required to support user needs.

2. **Database design**: Describes the format for the layers, individual features, and their attributes that will comprise the new GIS database.

3. **Data development plan**: Describes the techniques, methods, and procedures that will be used to convert the data sources into the desired GIS database.

4. **System design**: Describes the hardware and software to be installed on new servers, workstations, and network components, as well as the reorganization and redeployment of existing hardware and software components. (Beyond the scope of this book.)

5. **Implementation plan and schedule**: Describes the tasks and provides a schedule for developing the GIS. (Beyond the scope of this book.)

**Application Design.** The value of digital data in general, and in a GIS in particular, is the ability to create data one time and use it over and over for many purposes without the need to manually handle the data. This recycling of data enables the data user to work with much larger datasets than would be possible with manual data entry.
and manipulation. Once developed, a GIS can serve many applications — beyond hydraulic modeling — in a community or water utility.

In evaluating the responsibilities and workflow within a department, certain tasks that can be performed more efficiently or effectively in a GIS will be identified. These tasks will form the basis of GIS applications, and application descriptions prepared as part of the needs assessment will document these tasks.

Popular GIS interfaces include:

- Interface to hydraulic and hydrologic modeling
- Interface to customer service and maintenance management systems
- Interface to customer information (billing) system
- Interface to laboratory information system
- Interface to SCADA (Supervisory Control and Data Acquisition) system
- Interface to Document Management/Workflow

Popular GIS-based applications (often utilizing/integrating the above system interfaces) include:

- Facility mapping (GIS data maintenance)
- Service request tracking/work management
- Asset management/GASB 34 reporting
- Crew dispatch/vehicle routing
• Field data collection/inspections
• Leak detection (compare master meter to individual account data)
• Link to as-built/intersection drawings (CAD or images)
• Isolation tracing/customer notification
• Demand projections/demographic data
• CIP planning/construction monitoring
• One call/underground service alert response
• New connection processing
• Cross-connection (backflow) test tracking
• Well monitoring/water resources analysis

**Database Design.** The database design for a GIS developed for a water utility should strive to accomplish three fundamental goals that will enable the GIS to become a strategic asset for the organization:

1. **Cartographically represent the water distribution facilities (assets).** This representation can be used to create map products.

2. **Inventory the network.** The GIS is often the primary record for geographically-distributed assets (that is, assets outside the plant).

3. **Model the network.** The GIS should be able to model the flow of water in the system and support the integration of hydraulic modeling software.

Assuming that hydraulic modeling will be one of the activities supported by the GIS, the GIS analyst should identify the entire range of related hydraulic applications that potentially will be used. The analyst can then determine the types of data required for these applications (for example, SCADA data and as-built drawings) and how the various types of data relate to one another. This information is necessary if the database design is to meet all functional and interrelational requirements.

An important step in the database design process is the compilation of information about the dataset, which is called *metadata*. Metadata provides the user with:

• Source of the data
• Data reliability, quality, and quality confidence levels
• Methods used in collecting and associating the data
• QA/QC (quality assurance/quality control) and validation procedures
• Other applications and software systems that the data might interact with

Metadata is becoming more important as the ability to share data between organizations over the Internet using eXtensible Markup Language (XML), a programming language for structured information, becomes commonplace. Portal sites use metadata so that users can search for data layers and determine whether the datasets listed meet their accuracy and spatial extent needs.
As GIS applications mature, the major GIS vendors have advanced their software from relational-relational architectures (application/database) to object-relational architectures. In the older relational-relational environment, connectivity, attribute domain validation, and relationships between features were implemented as relationships between tables (executed by the DBMS). In the new architecture, relationships between features, connectivity, and attribute domain validation are implemented as object-oriented components (executed by client-side software) with the raw objects being stored in the RDBMS. The principal advantage of the new object-relational structure is the compatibility of the GIS applications and data to other object-oriented software applications.

This evolution provides new opportunities and poses some new challenges for integration with modeling. Objects can be more complex, with more sophisticated connectivity, than with older GIS data types (points, lines, polygons). Because the data modeling effort in an object world is so much more involved (Zeiler, 1999), the hydraulic modeler has to be even more involved at the GIS database design stage when using an object-relational GIS.

As an example, consider a GIS that has been developed for a water distribution system using the older relational-relational system. In this model, each valve, hydrant, water service, and service shutoff throughout the entire water distribution system is required to be a node in the network in order to maintain connectivity. As a result, the GIS might easily contain hundreds of thousands of short pipe segments (for example, fitting to valve, valve to valve, and valve to fitting). To the hydraulic modeler, this level of detail is unnecessary and even problematic.

Object-relational systems provide powerful new opportunities to the hydraulic modeler. For instance, the connectivity between features is now controlled by the software according to rules supplied by the user. For instance, valves and services can be part of the network without creating nodes in the pipe segments (that is, pipes can be fitting-to-fitting). In addition, connectivity rules can be established that enhance the integrity of the database, such as “a pipe can only connect to another pipe through a fitting, PRV, or pump station” or “only a customer meter or backflow device can be at the end of a service line.” The result can be a GIS database that more closely meets the needs of the hydraulic modeler.

In older data models, the modeler would likely need to develop code to perform the tasks of discarding unnecessary point features and merging small pipe segments into the larger segments that would form the link-node system in the model. This code would rely not only on a programmer's expertise to properly process the myriad conditions which exist, but on a high level of data quality to avoid unintentional errors during the translation. With object-oriented data storage, the processing of data to form new, merged features can be manifested through behaviors introduced at the database design stage. Further, rules can be applied during dataset development to ensure that, for example, a 2-inch pipe is not connected to a 16-inch pipe without some type of required transition element. By designing these rules, properties, and behaviors up front, extracting the features needed to model a system can become a basic function of the GIS. The challenge in this system is that the technology is new, and a significant investment in design time and funding will be required in the coming years to turn the promise into reality.
Although it is relatively easy to define those characteristics of the buried piping system that are needed to support hydraulic modeling, modeling the behavior of a complex pump station within the GIS is more difficult and may not make sense. In many cases, some of the information needed by the modeler will not be stored in the GIS and will have to be acquired from other sources, such as as-builts or design documents. Ideally, all necessary input data for the model would exist in the GIS, and the data format and content would be capable of supporting all modeling goals and applications identified in the needs assessment. In practice, however, all the data needed by modelers is sometimes not in a readily available format or is not economical to develop (for example, field data) during the initial GIS implementation. Therefore, many hydraulic model interface implementations are staged to address the most urgent or time-saving modeling priorities first. The GIS should be capable of generating, in a repeatable manner, a high percentage of the data required for model development.

It is important to realize that GIS is often part of a wider information management program that may include maintenance management systems, SCADA, facility automation, CAD, flow monitoring databases, a water asset database, as-built maps and drawings, and other elements. Developing a database design that supports the multiple needs of an organization is sometimes difficult to accomplish, but the end result should not be compromised.

**Data Development Plan.** As previously mentioned, development of the data asset is often the most important and expensive step in the GIS development process. The following subsections describe the main components and issues of the data development plan.

*Land Base.* A water utility GIS must use some type of land base layer as a spatial reference. Some design issues that must be considered in the base map are the accuracy (+/- x ft or m), the scale, the projection (latitude/longitude, state plane coordinates, UTM [Universal Transverse Mercator], etc.), the vertical datum (NAD [North American Datum] 27, NAD83, etc.), the methods used to create the land base, the frequency of updates, and the timeliness of the data.

The details of these decisions are beyond the scope of this book, but users must realize that they cannot simply make a quick decision about using a United States Geological Survey (USGS) quadrangle map, aerial photo, or commercially produced map without giving serious consideration to the long-term potential implications of the land base choice.

Developing the land base is expensive and must be done correctly and accurately in the initial stages of the project; otherwise, costs escalate when problems arise that must be addressed after the fact. The lack of an accurate land base is the most common problem in developing an accurate GIS because multiple datasets developed without a common land base will almost always be inconsistent with each other and therefore not suitable to be used together.

For example, if spatially accurate water facilities (such as those located with a differential global positioning system [GPS]) are used in combination with inaccurate base map layers, a good map will not be produced (for example, pipes may not fall on the
Correct side of the street centerline or property line). Not being able to produce reliable maps is more than a nuisance; it could be a legal liability.

If an inaccurate street-centerline file is used as the spatial reference for placing water facilities (say, for putting pipes four feet to the right of the street centerline), and a new, more-accurate data source for the street centerline is provided in the future, all of the facilities will have to be moved in the GIS to match the new spatial reference — an expensive proposition.

**Data Conversion.** GIS database design must be matched to the specific needs of the applications that the GIS is to serve, such as those of hydraulic models. Through data conversion, data is made to conform to a uniform format that supports all functional
requirements. If necessary, GIS database design is modified to effectively integrate it with the hydraulic modeling software, information management systems, as-built drawing records, GPS, and CAD.

In supporting the data development process, the GIS analyst must develop and maintain data accuracy standards and implement QA/QC procedures to ensure data integrity. The higher the accuracy standard, the higher the cost of the data conversion.

In some cases, the information necessary to support current and future hydraulic modeling applications is not available in digital format. It will therefore be necessary to convert paper maps to digital format or use GPS to collect data in the field. If multiple data sources exist, the analyst should identify the sources that are best-suited for the intended use, as well as the method for entering this data into the GIS.

Raster conversion, or *rasterization*, is the conversion of vector data (that is, points, lines, and polygons) to cell or pixel data. Vector conversion, or *vectorization*, is the conversion of cell or pixel data into points, lines, and polygons. Scanning of paper maps produces raster files that must be converted to vector files to be useful for model creation.

Rasterization and vectorization are only used when the quality of the data sources supports it, which means that the information to be captured is easily distinguishable by the vectorization software. Dealing with the errors of the vectorization process can be onerous for the majority of GIS development projects. Heads-up digitizing is more efficient, produces excellent results, and is rapidly becoming the industry standard.

Converting map data from CAD files into the desired GIS format can also pose a number of challenges. For example, a single line in CAD file may actually represent several pipe segments in the GIS. When this element is converted to GIS, the line must be divided into multiple lines representing individual pipe segments. As an additional hurdle, map annotations such as pipe diameter and valve size in a CAD drawing often do not have any data records and are not linked to the CAD features being annotated. To be turned into attribute information in the GIS, these floating graphics must be associated with the nearest pipe, and the text strings must be filtered to produce the correct attribute. For instance, "1987 - 3 ½" may need to be converted into "1987" for the year installed attribute and "3.50" for the pipe diameter attribute.

**Pilot Study**

After design, the next phase of the GIS development is usually to perform a pilot study, which can include the following activities:

1. Create a pilot database following the data development plan.
2. Develop prototypes of high-priority applications following the application design.
3. Provide core software training to key staff.
4. Test the applications and data during several pilot review sessions with end-users and management.
5. Finalize the database design, data development plan, and system design documents as appropriate, incorporating what was learned from the pilot review sessions.

**Production**

The next phase is the Production phase, which can include the following tasks:

1. Finalize the QA/QC software and techniques that will be used during the entire-service-area data conversion.
2. Perform the service-area-wide data conversion following the data development plan.
3. Procure new hardware and software.
4. Finalize the applications.
5. Develop end-user and system maintenance documentation.
6. Begin user training and rollout of high-priority applications (such as facility mapping).

**Rollout**

The final phase is a rollout phase, which can include the following tasks:

1. Installation of full complement of operational hardware and software.
2. User training and system maintenance training, which will likely be a combination of core GIS software courses and application training.
3. Acceptance testing (formal testing to determine whether the system satisfies the acceptance criteria and thus whether the customer should accept the system).
4. Roll-out, which may include transition from any legacy systems being retired.

**12.3 MODEL CONSTRUCTION**

Constructing a water model and maintaining it over time can be one of the most time-consuming, costly, and error-prone steps of a hydraulic modeling project. Prior to widespread integration of GIS and modeling, building a water model was a specialized activity, separate from an organization's routine business procedures and workflows. Engineers created model input files by gathering, combining, and digitizing data from a variety of hard-copy source documents, such as water system maps, topographic maps, and census maps, among others. If CAD data was available, features required for modeling had to be extracted for use with the hydraulic modeling software. The process was manual and required great attention to detail and many engineering judgments along the way. Once a model was developed, calibrated, and run, the modeler generated the required outputs and what-if scenarios and typically produced a master plan so that a community or utility could begin making the required capital improvements.
Despite rapid advances in hydraulic modeling software throughout the last decade, which have included tools for the automatic translation of CAD data into modeling data and for linking the model to external data sources, many communities and water utilities have found it difficult to build, update, and maintain anything but highly skeletonized models. Because of the impracticality of using manual methods to gather and manage large volumes of data, many communities and water utilities have not performed routine modeling and have had no mechanism for water model maintenance. Although organizations may have intended to keep models updated, time constraints or business-process issues have often interfered. Even if CAD layers have been updated or as-builts red-lined, model maintenance has usually been ignored because the model input data has had to be maintained separately from system drawings. Thus, in communities experiencing rapid development, water models could quickly become outdated and have often had to be re-created from scratch when a current system model was required.

A GIS professional can use a GIS to create a model more efficiently, more accurately, and more cost-effectively than an engineer creating a model input file from scratch inside a traditional modeling environment. Consider the following:

- Because GIS tools can automate the process, model building can be faster and more efficient, especially for large models.
- Because GIS can manage large volumes of data, the model can incorporate more detail.
- Both hydraulic modeling software and GIS have advanced editing tools. The user needs to look at each task and decide if it is better done in the model or the GIS.
- In the ideal case, where GIS data-entry has a consistent spatial reference and a high level of quality control applied, the integrated model should contain better data and should therefore be easier to calibrate and potentially lead to better decision-making.
- Data collected and stored in the GIS primarily for other applications can be extracted and incorporated into the model input file, if needed.
- As long as the GIS is maintained routinely, GIS data needed for reconstructing the model input file will be maintained routinely.
- Contour interpolations and digital elevation models (DEMs) in the GIS can be used to assign elevations to model nodes automatically.
- Digital orthophotos available in the GIS can be overlaid with the model to provide a base map reference.
- Georeferenced customer billing records can be used to generate and allocate water demands for the model.
- If modeling results are returned to the GIS, further analyses can be run, and other users such as planners and developers can manipulate modeling data in conjunction with other GIS data.
Model Sustainability and Maintenance

By working with GIS professionals to build a water model from a properly constructed GIS, the engineer can spend time evaluating the water system and making engineering decisions rather than constructing — and, over time, reconstructing — the water model. If the GIS is used as the foundation for building the water model, and assuming the GIS is maintained in the long term, the water model can be rebuilt easily in the future using up-to-date GIS data (see Figure 12.5).

For instance, a model may be constructed by using the GIS to select and extract all pipes that are 8 in. (200 mm) and larger, and by manually selecting additional pipes that have smaller diameters but are necessary to close important loops or to service large water users. All pipes and other network elements included in the model must be marked appropriately in the GIS. Thus, as pipes, nodes, and associated features and attributes in the GIS are updated over time, these elements can easily be selected again and used to reconstruct the model with current data. Alternatively, the full system can be imported into the model and then reduced by skeletonization.

The GIS professional and the modeler should determine where certain features required for the model will reside — in the GIS or in the model. Ideally, the required datasets should be detailed from the outset so that they can be stored and maintained in the GIS, which helps to ensure that the model-building process can be repeated and minimizes adjustments that the modeler must make to the data delivered by the GIS professional.

Making good decisions about whether new pipes added to the distribution system should be marked for inclusion in the model requires engineering judgment (based on pipe size or other criteria). The GIS professional can use GIS tools to do much of the sorting and selecting required for skeletonization, but the modeler must review the entire network carefully to ensure that the model reflects reality.

Communication Between GIS and Modeling Staff

In some small systems, the modeler and the GIS professional are the same person, but in most instances, the modeler and GIS professional are two individuals who are in
different departments or companies. To be successful, model development should be a coordinated effort between the modeler and the GIS professional.

During the model-building process, the GIS professional must confer regularly with the modeler to ensure that the model is developed efficiently. Both professionals should acknowledge their differing perspectives to help ensure effective communication. The GIS professional understands GIS technology, what it can do, and how to manipulate data from various systems and databases, and the modeler understands how the model works, what data are required by the model, and how to ensure that the model generates meaningful results.

Because the modeler is the user or consumer of the data to be developed or managed by the GIS professional, he or she should take time to explain the data needs of the model. Similarly, the GIS professional should explain GIS capabilities and limitations, and what is technologically feasible.

Sufficient time spent in model planning and design ensures a smooth model-development process. Specifically, the GIS professional and the modeler should discuss the following issues at the outset of the project:

- **Modeling basics:** The modeler should share modeling basics with the GIS professional (such as flow in or out of the system occurs only at nodes; closed pipes do not have flow; the model predicts pressures at nodes; and so on). Before using the GIS to develop the model input file, the GIS professional must have a fairly complete understanding of input data requirements, types of model calculations, and how the model operates. In most instances, the GIS professional can benefit from receiving some modeling training. By understanding how the model works and having a vision of the intended outcome, the GIS professional can determine which GIS tools should best serve the modeler's needs. Merely listing required features and attributes will not suffice; the modeler must discuss why these elements are crucial.

- **Modeling terminology:** For successful model development, each professional must have a basic understanding of the relevant terminology familiar to the other. The GIS professional is comfortable discussing fields, tables, and technology, and the modeler is comfortable discussing pipes and roughness coefficients. Both professionals must discuss common modeling terms, their meanings, and how they are applied. For example, they need to discuss terms such as *demand* (the GIS professional may use demand to mean an individual customer billing record, but the modeler may think of aggregate demand at a modeling node) and *junction* (the GIS professional may use junction synonymously with any node, and the modeler may think of junction as a location where two or more pipes meet).

- **Standard units of measure:** The modeler must define the standard units of measure required and the calculations needed to make unit conversions. Although the modeler understands intuitively how certain data must be represented in the model, the GIS professional may not. For example, if the modeler needs average demand in gallons per minute, and the demand in the billing system is in gallons per day, then the units must be converted either...
by the model (preferred) or the GIS professional before the GIS data is useful in the model. The modeler must therefore communicate very clearly how the GIS professional should prepare the data for model input.

**Using an Existing GIS for Modeling**

A common misconception in the industry is that if a utility has a water system GIS dataset, using the GIS to generate data needed for the model will be easy, if not automatic. If hydraulic modeling was identified as a potential application of the GIS during the needs assessment, this should be true. However, if the GIS was designed only for purposes such as hard-copy mapping, maintenance management, or capital improvement planning, it might not be easy to use the GIS to generate a hydraulic model. This limitation is most pronounced when the GIS is used primarily for hard-copy mapping and its “users” never interact directly with the underlying GIS database. The GIS data are not necessarily bad, but they may not be suitable for hydraulic modeling purposes, even though they satisfy the existing GIS needs.

Some typical challenges often faced by modelers using GIS data include the following.

Database design incompatibilities or omissions:

- The GIS incorporates identification numbers for system elements that differ from what can be used by the modeler.
- The GIS lacks critical valve information.
- The GIS lacks pump performance curve information.
- The GIS may have many short pipe segments that were introduced to provide full topology, but that would unnecessarily add to the complexity of the model.

Data errors:

- The GIS may not include connections that occur in the real system (because components are not “snapped”), and may include connections that do not occur in the real system (for example, at locations where pipes cross but do not connect).
- The GIS may have artificial gaps between pipes that were introduced because the data was created in a tiled system and pipes were not connected across sheet boundaries.

**Network Components**

Even if hydraulic modeling was considered during the development of the GIS, unless the GIS was created *solely* to support modeling, it is likely to possess a much greater level of detail than what is needed by the model. This excess is especially true with regard to the number of piping elements. It is not uncommon for the GIS to include every service line and hydrant lateral. Such information is not needed for most modeling applications and should be removed to improve model runtime, reduce file size, and save costs.
In addition to the extraneous service lines and hydrant laterals, a GIS may begin a new pipe element at every isolation valve or fitting — an unnecessary level of detail for most hydraulic modeling applications. Conversion of the GIS to the model therefore involves combining GIS elements to form a smaller number of model elements, as shown in Figure 12.6.

Two steps exist at which the GIS data can be cleaned up for model use — importing and skeletonization. In importing data from the model to the GIS, the extraneous GIS features such as air release valves, in-line meters, open gate valves, and locations where new pipe segments were installed in response to a break must be associated with model elements. This association usually requires that mapping be established between the GIS and model. The primary criterion for handling these elements is whether or not they have appreciable head loss associated with them. Open gate valves and pipe bends often have negligible head loss and can be represented as pressure junctions, which can be eliminated during skeletonization. Those features that have significant head loss either need to be assigned to an adjacent pipe or treated as a control valve element or a general-purpose valve. Typical relationships are shown in Table 12.1.

Trade-offs exist with each situation. For example, modeling an open gate valve or water meter as a throttling control valve or general-purpose valve makes it easy to assign minor losses to the element. However, the three elements (upstream pipe, valve, and downstream pipe) could also be modeled as a single pipe element with a minor loss, thereby reducing model size.
<table>
<thead>
<tr>
<th>GIS Feature</th>
<th>Model Elements</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bend or other fitting with negligible loss</td>
<td>Pressure junction to be skeletonized out</td>
</tr>
<tr>
<td>Bend or other fitting with significant loss</td>
<td>Valve or minor loss on adjacent pipe</td>
</tr>
<tr>
<td>Isolating valve that is always open with negligible loss</td>
<td>Pressure junction to be skeletonized out</td>
</tr>
<tr>
<td>Isolating valve that is always open with significant loss</td>
<td>Valve or minor loss on adjacent pipe</td>
</tr>
<tr>
<td>Isolating valve that is normally closed</td>
<td>Valve or a pipe segment with two nodes</td>
</tr>
<tr>
<td>Isolating valve that is always closed</td>
<td>Valve or nodes connected by closed pipe</td>
</tr>
<tr>
<td>Air release, blowoff, or surge relief valve</td>
<td>Pressure junction to be skeletonized out</td>
</tr>
<tr>
<td>Customer or hydrant lateral</td>
<td>Pressure junction and pipe to be skeletonized out unless individual customer is to be modeled</td>
</tr>
<tr>
<td>Check valve that is in-line</td>
<td>Property of adjacent pipe</td>
</tr>
<tr>
<td>Check valve that is at a pump</td>
<td>Usually automatically included in pump, pipe on either side should be combined</td>
</tr>
<tr>
<td>System water meter</td>
<td>Valve or minor loss on adjacent pipe</td>
</tr>
<tr>
<td>Control valve</td>
<td>PRV, PSV, TCV, or GPV, depending on function of valve</td>
</tr>
<tr>
<td>Pump control valve</td>
<td>Pressure junction to be skeletonized out</td>
</tr>
<tr>
<td>Reducer</td>
<td>Pressure junction with different diameter pipe on either side</td>
</tr>
<tr>
<td>Change in pipe material</td>
<td>Pressure junction which may be skeletonized out depending on difference in hydraulic properties</td>
</tr>
</tbody>
</table>

After the data are imported into a model, the number of elements can be further reduced by skeletonization. Section 3.11 provides an overview of the skeletonization process.
Some would argue that with the increasing power of models and the ease with which models can share data with GIS, skeletonization has become less important and the state-of-the-art is evolving toward more “all-pipe” models. Different levels of skeletonization are still appropriate, however, depending on how the model will be used.

The GIS professional and the engineer (or modeler) should discuss specific skeletonization criteria in detail, including the importance of network connectivity. The GIS professional must understand that a fully connected network is required for modeling.

**Retrieval of Water Use Data**

Utilities collect and compute water usage data by means of several possible methods, ranging from the highly accurate to the more generalized. Three common techniques are:

- Storing individual customer meter records for each billing period in a customer information system
- Aggregating usage data for larger areas such as meter routes or pressure zones
- Computing water usage estimates based on land-use or population

A GIS is not essential for loading the hydraulic model with water usage data, but it can be used to effectively address each of these use cases and streamline the demand allocation process. Much of the early work in using GIS for modeling focused on accurately placing demands (Basford and Sevier, 1995; Buyens, Bizier and Conbee, 1996; Davis and Braun, 2000).

**Node Service Polygons.** Junction nodes are point features, but some demand allocation methods require that the nodes have a polygon service area associated with them. These polygons can be constructed manually, but automatic techniques, such as construction of Thiessen polygons, exist. *Thiessen polygons* define the individual tributary areas for each node. The space is divided such that any point within a particular Thiessen polygon is nearer to that polygon's node than to any other node. A Thiessen polygon for a node is created by connecting the perpendicular bisectors of lines drawn between it and all adjacent nodes. The polygons for nodes along the outer edge of the model will have no outer boundary, so it is essential to specify some method for closing the boundary to calculate areas. The boundary can be based on a buffering distance, but it is usually best to draw the outer boundary manually. Also, areas having no customers may exist within the model area (for example lakes, parks, landfills). Figure 12.7 shows a typical system with Thiessen polygons around each node.

**Customer Meter Data.** When water usage data is available for individual customer meters, the GIS can be used to automatically geocode the customer location. *Geocoding* is the process of matching an address data field or equivalent spatial reference, in this case a customer service address, against (usually) a street centerline file or a parcel file that also contains address information. The resulting file is a set of points that coincide with parcel centroids or the interpolated length along a line segment, depending on the source file used to geocode. It is important for the modeler to
understand how the meter location was geocoded. Using the underlying point coordinates and the coordinates of the nodes in the model, the demand can be assigned to a node, usually based on which node is nearest to the customer meter (see Figure 12.8). When the actual service line connection point is stored in the GIS, then the demand can be placed at one of the nodes for that main or proportioned based on distance from the end nodes. Usually, the difference in model results due to different methods of assigning demands along the pipe is negligible.

**Figure 12.7**
Thiessen polygons for distribution system nodes

Geocoding can be a very effective method of demand placement, but it relies on source datasets (GIS and billing) containing address data that are seldom standardized and validated for geocoding purposes. Therefore, it is important to understand that the results of a geocoding effort may inaccurately attribute the load to one polygon versus another due to geocoding inaccuracies. The geocoding results should be checked for accuracy, and changes to the load assignments may need to be performed manually.

Another problem in using customer meter data is that CISs are set up to capture total volume used for billing purposes, not flow rates that are required for models. Consider Table 12.2. The CIS is likely to contain the first four columns, but the modeler needs data from the final two columns (that is, flow rate). The conversion from volume billed to flow rate must be done either in the CIS software, the GIS, or by some code written specifically for this calculation.
### Table 12.2 Typical billing information from a Customer Information System (CIS)

<table>
<thead>
<tr>
<th>Reading (100 ft³)</th>
<th>Read Date</th>
<th>Billing Cycle (days)</th>
<th>Usage (100 ft³)</th>
<th>Flow (gpd)</th>
<th>Flow (gpm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>6,754.83</td>
<td>3/12/2002</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6,770.25</td>
<td>4/9/2002</td>
<td>28</td>
<td>15.42</td>
<td>412</td>
<td>0.29</td>
</tr>
<tr>
<td>6,786.72</td>
<td>5/11/2002</td>
<td>32</td>
<td>16.47</td>
<td>385</td>
<td>0.27</td>
</tr>
<tr>
<td>6,805.99</td>
<td>6/11/2002</td>
<td>31</td>
<td>19.27</td>
<td>465</td>
<td>0.32</td>
</tr>
<tr>
<td>6,826.93</td>
<td>7/11/2002</td>
<td>30</td>
<td>20.94</td>
<td>522</td>
<td>0.36</td>
</tr>
<tr>
<td>6,850.74</td>
<td>8/9/2002</td>
<td>29</td>
<td>23.80</td>
<td>614</td>
<td>0.43</td>
</tr>
<tr>
<td>6,879.08</td>
<td>9/9/2002</td>
<td>31</td>
<td>28.35</td>
<td>684</td>
<td>0.48</td>
</tr>
<tr>
<td>6,900.10</td>
<td>10/9/2002</td>
<td>30</td>
<td>21.02</td>
<td>524</td>
<td>0.36</td>
</tr>
<tr>
<td>Cumulative</td>
<td>211</td>
<td>145.27</td>
<td>515</td>
<td></td>
<td>0.36</td>
</tr>
</tbody>
</table>
The modeler must also decide on whether the value to be loaded into the model is the value from the most recent billing period or an average over a longer period. This decision depends on whether the modeler is interested in loading the model with average annual flows or flows from a particular period for, say, a calibration exercise.

Other issues include the fact that not all customer meters are read on the same day. For instance, “July water use” for one customer may be calculated from July 1 to July 31, but for another customer, it may be July 11 through August 12. In addition, some utilities may use different units for commercial customers and residential customers. Meters may have become stopped during the year or been replaced during the year such that simply subtracting the last reading from the first reading will not give the correct volume used. Sometimes corrections or adjustments are made to billed amounts. The modeler must decide whether to use the raw value or corrected value. Most of these cases must be dealt with manually. For the modeler, the key to working with customer billing data is talking with the individual who understands the data to determine its true meaning.

When nodal demands are determined from customer metering, the modeler must remember that unaccounted-for water, by definition, is not included in the flow rates. The modeler must determine how to assign unaccounted-for water to nodes. This assignment may be done by simply dividing the unaccounted-for demand evenly among the nodes. For greater accuracy in EPS runs, the modeler may want to set up a
different demand pattern for unaccounted-for water demands, although sufficient data are not usually available for this.

**Area Flow Data.** When demand data are available for large areas such as pressure zones and meter routes, these areas should be incorporated into the GIS as a system meter polygon layer. Using overlay analysis, the water use within each polygon can be equally distributed among the model nodes that fall within the polygon by using *point-in-polygon analysis*. For customer demand data that have been placed using a geocoding process, use of the meter route identifier to place points that could not be geocoded is a common backup procedure.

Alternatively, if the individual model nodes have service area polygons associated with them, the total demand from the pressure zone or meter route polygon can be proportionally assigned to the service area polygons (and then to the model nodes) based on the percentage of the larger polygon area taken up by the service area polygon (see Figure 12.9). See page 549 (“Node Service Polygons”) for more information about this method.

![Figure 12.9](image)

*Proportional flow distribution*

<table>
<thead>
<tr>
<th>Area</th>
<th>Q&lt;sub&gt;a&lt;/sub&gt; (gpd)</th>
<th>Q&lt;sub&gt;b&lt;/sub&gt; (gpd)</th>
<th>Q&lt;sub&gt;total&lt;/sub&gt; (gpd)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Area 1</td>
<td>162</td>
<td>0</td>
<td>162</td>
</tr>
<tr>
<td>Area 2</td>
<td>250</td>
<td>25</td>
<td>275</td>
</tr>
<tr>
<td>Area 3</td>
<td>0</td>
<td>175</td>
<td>175</td>
</tr>
<tr>
<td>Area 4</td>
<td>0</td>
<td>210</td>
<td>210</td>
</tr>
<tr>
<td>Area 5</td>
<td>440</td>
<td>35</td>
<td>475</td>
</tr>
<tr>
<td>Area 6</td>
<td>148</td>
<td>55</td>
<td>203</td>
</tr>
</tbody>
</table>

Flow A: Total Flow = 1000 gpd
Flow B: Total Flow = 500 gpd
**Land-Use/Population Data.** In some cases, the attribute data in a database may be sufficient to get the bills out but may be very poor for geocoding purposes. In other cases, demand data is only available as land-use or population information based on census tracts, traffic analysis zones, or other similar areas. When data are stored in the GIS in terms of the area of a given land use, population, number of structures, or population density, the modeler needs to develop the corresponding rates for water usage. For example, if considering land use, rates need to be established for water demand per unit of land use (for example, gal/day/ac or l/day/ha). If the number of structures is being considered, then rates need to be established for water demand for each type of dwelling unit. With this data, the aggregate demands can be automatically computed within the GIS.

Figure 12.10 and Table 12.3 show how water usage can be determined based on land use.
### Table 12.3 Computing water use based on land-use area

<table>
<thead>
<tr>
<th>Node</th>
<th>Total Node Area (ha)</th>
<th>Land Use Type</th>
<th>Land Use Area (ha)</th>
<th>Unit Demand (l/day/ha)</th>
<th>Demand (l/day)</th>
<th>Node Total (l/day)</th>
</tr>
</thead>
<tbody>
<tr>
<td>J-1</td>
<td>6.88</td>
<td>Industrial</td>
<td>6.88</td>
<td>11,200</td>
<td>77,100</td>
<td>77,100</td>
</tr>
<tr>
<td>J-2</td>
<td>7.69</td>
<td>Industrial</td>
<td>1.38</td>
<td>11,200</td>
<td>15,500</td>
<td>60,200</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Commercial</td>
<td>0.92</td>
<td>4,700</td>
<td>4,300</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Residential</td>
<td>5.38</td>
<td>7,500</td>
<td>40,400</td>
<td></td>
</tr>
<tr>
<td>J-3</td>
<td>7.69</td>
<td>Commercial</td>
<td>1.31</td>
<td>4,700</td>
<td>6,100</td>
<td>44,800</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Residential</td>
<td>5.15</td>
<td>7,500</td>
<td>38,600</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Undeveloped</td>
<td>1.23</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>J-4</td>
<td>8.50</td>
<td>Industrial</td>
<td>0.17</td>
<td>11,200</td>
<td>1,900</td>
<td>20,800</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Commercial</td>
<td>0.10</td>
<td>4,700</td>
<td>470</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Residential</td>
<td>2.45</td>
<td>7,500</td>
<td>18,400</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Undeveloped</td>
<td>5.78</td>
<td>0</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>J-5</td>
<td>8.09</td>
<td>Industrial</td>
<td>6.48</td>
<td>11,200</td>
<td>72,500</td>
<td>80,100</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Commercial</td>
<td>1.62</td>
<td>4,700</td>
<td>7,600</td>
<td></td>
</tr>
<tr>
<td>J-6</td>
<td>4.86</td>
<td>Industrial</td>
<td>0.20</td>
<td>11,200</td>
<td>2,200</td>
<td>33,400</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Commercial</td>
<td>1.36</td>
<td>4,700</td>
<td>6,400</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Residential</td>
<td>3.30</td>
<td>7,500</td>
<td>24,800</td>
<td></td>
</tr>
</tbody>
</table>

GIS can also be used to calculate demands for future conditions based on population or land-use projections supplied to the modeler. Custom GIS operations enable modelers to compute future water usage rates by overlaying data such as population projection and future land-use polygon layers with a modeling node layer.

**Retrieval of Elevation Data**

Many GISs contain elevation data, and numerous sources of elevation data in digital form exist, such as published digital elevation models (DEMs). These models are often used to develop the contours found in base mapping, but they can also be used to derive the elevation for any other GIS feature, or for hydraulic model nodes. It is typically a very simple and fast process to drape, or overlay, the model nodes onto a surface model and compute the elevation at each point. As with most GIS layers, depending on the means used to develop it, a surface model can be very detailed or quite coarse.

Elevation data can be stored in a GIS in several ways, two of which are TINs and raster DEMs. Figure 12.11 shows a shaded TIN of ground elevation data with the hydraulic network draped over it, and Figure 12.12 shows a raster DEM with a hydraulic model superimposed on it.
**Figure 12.11**
Network draped over a TIN

**Figure 12.12**
Network superimposed on a DEM
For hydraulic modeling purposes, elevations are needed at specific points, including junction nodes, pumps, valves, and tanks. To determine the elevation at these points, the x-y coordinates of the node are passed to the GIS, and the GIS software uses the elevations of DEM grid points or TIN vertices surrounding the node to determine the node elevation, which can be passed back to the model. Use of DEMs is described in more detail in Miller (1999), Price (1999), and Walski et al. (2001).

DEMs have been prepared for much of the United States by the U.S. Geological Survey and are available with grid spacing of 30 m (although 10-m grid spacing is available for an increasing number of locations). These maps are based on the contour intervals found in the 7.5-minute quad sheets of that area (for example, 20 ft and 10 ft). Interpolating between these lines may only be accurate to several feet. In some cases, it may be necessary to find sources of elevation data with higher accuracy in order to achieve the model accuracy desired. It may be worthwhile to use free or low cost data initially, and then determine if it is accurate enough. If it is not, more accurate DEMs may need to be obtained.

Software that converts the data from a DEM to elevation attributes in a model usually requires that the DEM data be in a specific format. Some type of function is usually necessary to convert the raw DEM data into the required raster grid format and project the data to the coordinate system being used in the GIS (most USGS data sources are provided in Universal Tranverse Mercator [UTM] coordinates, which are not commonly used in municipal GIS applications).

Most models cover more than a single USGS quad sheet. Although elevation data import software can work with one sheet at a time, it is usually easier to mosaic the raster grid files together into a single file using a GIS function rather than deal with numerous files. In recent years, improvements in LIDAR (Light Detection and Ranging) technology have proven to be a source for higher resolution and accuracy in elevation data. LIDAR data can be represented in a number of file formats, including DEM and TIN.

**Modeling GIS Versus Enterprise GIS**

Few would doubt the value of using data from a GIS to create and update a model. One issue that arises in modeling with GIS, however, is whether model results should be carried back to the source GIS. Typically, the modeling data is maintained in a separate GIS layer created specifically for modeling.

Hydraulic modeling results in the enterprise GIS can be used for a number of purposes, including:

- Pressure mapping (used by both engineering and customer service)
- Establishment of water main replacement priorities (when combined with other GIS layers such as soils and repair data)
- Connection permit processing (available capacity can be reviewed and future demands reserved)
- Contaminant isolation/remediation (contaminant is introduced accidentally or intentionally)
**Extract, Transform, Load.** Modelers will leverage many data sources within an enterprise. An enterprise GIS houses the centrally managed data that is shared by individuals within the organization. Modelers typically envision the GIS as the primary data source or hub from which they will derive their models. Modelers also need access to other important data sources managed outside of the GIS. Often, this information will not reside in the same database or physical server that contains the GIS. It may be maintained using a blend of database technologies and/or proprietary file formats. The data sources may be distributed across the enterprise and hosted on various servers and client workstations. Figure 12.13 presents a generalization of such an enterprise-based GIS modeling system.

The figure shows the data pathways between the distributed data sources and the modeling GIS. The key to successful modeling in the enterprise is to use automation tools to accomplish the flow of data between these pathways whenever possible, avoiding manual intervention or transcription of data between the GIS and the original data sources. The modeler can accomplish this automation in several ways:

- By using general utilities to extract the data from one source, transform it as required for the target source, and then load it into the target source. Such utilities, called extract/transform/load utilities (ETLs), are readily available and are extremely valuable to modelers.
- By using programming and script Application Programming Interfaces (APIs) provided by the GIS and the databases to develop custom extensions of the standard GIS commands for accomplishing the ETL steps.
- By using commercially available technologies that are specialized for modeling within GIS. These technologies usually focus on some of the intensive and key data transformation services (for example, automation of the protocols for skeletonization, demand loading, and terrain extraction).

**Modeling Features.** Several aspects of hydraulic modeling must be taken into account when using a GIS for this purpose, and these considerations often lead to the separation of the enterprise GIS layers from the modeling GIS layers. These aspects of modeling are:

- Network granularity
- Scenarios
- Time-series data
- Ownership

*Network granularity.* For many hydraulic modeling applications, the model network does not need to contain every pipe in the actual system to obtain accurate results. For example, a skeletonized version of the system is often sufficient to make informed planning decisions and is often desirable to improve the efficiency of the hydraulic modeling software. Providing fields in the enterprise GIS layers to manage the hydraulic results would be wasteful because many of the GIS features are eliminated during the skeletonization process. However, providing a single field for the modeling identifier is possible, and this identifier allows the extraction or skeletonization software to make the link between the GIS feature and model feature. This link then provides a trail back to the GIS feature for the analysis of hydraulic results.
Scenarios. In water system planning, the modeler is most often dealing with what-if conditions, not as-built or in-service conditions. These conditions may include future demands, proposed pipes and system facilities, or facility outages for emergency response planning.
The GIS can be designed to incorporate the various what-if conditions and phases of a water system facility. For example, a given pipe goes through many stages:

1. Alternative pipe in the model
2. Proposed pipe in a planning study or budget
3. A pipe under design, bid, and then construction
4. An installed pipe that has not been tested or placed in service
5. A pipe placed in service

Ideally, the enterprise GIS will be designed to manage this whole life cycle, because many users need to see proposed pipes with other GIS data. A tremendous amount of activity within a water utility revolves around planned or proposed pipes. The sharing of data on planned or proposed pipes is one area where inter-agency GIS needs are high. However, papers on modeling (for example, Deagle and Ancel, 2002) typically describe how models use GIS, but few describe incorporating model information in the GIS. An exception to the situations described in most papers is the Indianapolis Water Company (Schatzlein and Dieterlein, 2002), which has a separate area in its GIS for proposed projects.

**Time-series data.** Hydraulic models and water quality models are dynamic, meaning they can be used to predict the response of the water system over an extended period of time. The modeler needs to visually analyze this time-series data in an efficient manner. Although an enterprise GIS can be designed to handle time-series data, they typically lack the tools for working with time-series data efficiently because most attributes for GIS features contain a single value (such as node elevation). With time-series data, it is not uncommon to have 48 hourly values for each hydraulic parameter for dozens of scenarios. The modeling software is set up to handle these large numbers of values, but the enterprise-wide GIS is usually not the best repository.

**Ownership.** In an enterprise-level GIS, the majority of the data is typically not “owned” by the hydraulic modelers, and changes to this data are often outside the modeler's direct control. Several issues arise from this lack of control. First, the enterprise GIS will often have data inaccuracies or omissions that are insignificant from a maintenance management or asset management perspective, but that are important from a hydraulic modeling viewpoint. The “owners” of the GIS may not be able to make these corrections as quickly as the hydraulic modeler requires, so the modeler is forced to make these corrections directly in the modeling GIS layers. An opposite but equally problematic issue is that the hydraulic modeler often does not want changes to the GIS to be immediately reflected in the model. For example, when calibrating against fire hydrant flow test data, it is important that the model reflect the “as-built” conditions at the time of the test. The GIS owners may be updating for recent water main installations, and this could have a significant impact on the model's results. The bottom line is that the modeler needs to be in control of the data used for modeling.
12.4 GIS ANALYSIS AND VISUALIZATION

This section illustrates ways that GIS might support a water utility in the hydraulic model application cycle, from model development to capital planning, decision support, and operations support. It is important to note that a significant amount of planning and data conversion may be required to develop the datasets to empower a GIS to perform all of these operations. Section 12.3 covers GIS development to support water utilities, concentrating on the hydraulic modeling aspects.

Using Attributes to Create Thematic Maps

A classic and very simple use of GIS is to change the appearance of features in a dataset based on an underlying attribute (attributes are the data that have been collected and associated with each feature in a dataset). In Figure 12.14, this capability illustrates an apparent error in GIS data that will be imported into a model, where a 6-in. (150-mm) line still exists in the middle of all new 8-in (200-mm) piping. In this

---

### Case Study: Germantown, Tennessee

The City of Germantown created a GIS to get a complete and accurate digital mapping inventory of its water network. The city also wanted to use the GIS to create a skeletonized water network suitable for hydraulic modeling. Germantown's approach to GIS development was to develop a basic infrastructure GIS for a single application—modeling—and then consistently enhance and build upon the GIS for additional applications over time.

In the past, Germantown had maintained its water data on a 1 in. = 500 ft scale, hard-copy map. The city used digital orthophotos to compile planimetric data so a spatially accurate GIS could be created and used as the basis for the water model. Valve and hydrant locations from the digital orthophotos and planimetric mapping were used with the 1 in. = 500 ft scale map as the basis for connecting water pipes. The result was a highly accurate GIS dataset that could be used immediately for modeling applications, inventory, and map maintenance, and in the future, for building additional applications.

GIS tools were used to extract a skeletonized water network for hydraulic modeling. The model was skeletonized to include water lines larger than 6 in., plus all other lines critical for looping. Additional criteria were used to perform batch queries and skeletonize the network. For example, because Germantown did not need to perform unidirectional flushing, all hydrant laterals were queried and removed from the model dataset. Once the skeletonization was complete, any critical pipe spans omitted were added back into the skeletonized network.

The GIS was used to assign customer demand information, taken from the city's billing database and stored in parcel centroids, to model nodes. Each billing data address was matched with a parcel centroid address, and demands were then aggregated and assigned to modeling nodes. The procedure resulted in very precise demand assignments, which would not have been possible without GIS. Elevations were also extracted from Germantown's DEM and automatically assigned to model nodes. The model was then calibrated and run to determine whether the system could meet the present-day demand, and where future capital improvements were needed.

Based on the success of the GIS, Germantown is now considering a GIS expansion that would produce an integrated information system featuring an online parcel application and a document management system for managing the city's as-built drawings.
figure, water pipes 4-in. (100-mm) and smaller are represented by dashed lines, 6-in. (150-mm) pipes by thin lines, and 8-in. (200-mm) and larger pipes by thick lines. Used in this manner, GIS can serve as an excellent quality-control tool for all manner of data related to the water distribution system piping prior to model construction (such as pipe length, C-value, diameter, and so on). Similarly, thematic maps can be used to show which customers are late on payments or distinguish between different types of water quality complaints (or any underlying attribute of any dataset).

**Figure 12.14**
Basic thematic mapping — the thin line between thick lines indicates a diameter error in the GIS.

*Note:* Most water distribution modeling software can perform basic thematic mapping, but the number of available settings to distinguish between unique attribute values is usually higher in a GIS environment.

A variation of basic thematic mapping for quality control or map production involves color-coding data by underlying attribute ranges. For example, water distribution nodes at elevations below 500 ft could be color-coded dark blue, 501-600 ft in light blue, 601-700 ft in green, and so on. These types of thematic maps can also be effective in quality control and debugging operations.

Communicating the results of modeling to managers, regulators, the media, and the general public is often difficult. A good map can convey the information much more clearly than long oral explanations or text. Publications such as ESRI’s Map Book series contain hundreds of examples of maps drawn using GIS (ESRI, series).

Some examples of thematic mapping are:

- Color-coding pressure zones and overlaying the model to show which undeveloped land parcels belong in which pressure zones.
• Illustrating the areal extent of deficient fire flow capacity before and after the proposed improvements are constructed
• Displaying concentrations of a contaminant in the system in conjunction with epidemiological data to aid in correlating illness with water quality
• Showing a three-dimensional view of the hydraulic grade line overlaying a system

Most GIS packages provide many alternatives within the framework of basic thematic mapping that can be used to produce innumerable effects.

**Using the Spatial Coincidence of Features to Assign New Data**

GIS software can analyze features on different layers to determine which coincide. In this way, data from one set of features can be transferred into a second set of features. A typical example in water model construction involves the overlay of model nodes against a layer that contains elevation data. The process is more fully explored in Section 12.3, but essentially, the elevation associated with the point at each node can be incorporated as an attribute of the node based on this coincidence. Nodes that lie outside the extent of the elevation layer will not be assigned an elevation in this case. Other examples of this type of analysis include

• Overlaying water distribution model nodes on meter routes to assign the meter route number to each node
• Overlaying water distribution model pipes and nodes on pressure zone boundaries to add the zone code to each model feature, enabling zone-specific models to be created (see Figure 12.15)
• Overlay a land-use map on a node service zone map to determine the land-use in each node service zone

**Using Spatial Relationships Between Features to Select Certain Elements and Assign New Data**

In addition to using the coincidence of features to assign new data, many GIS packages can select or isolate features based on their proximity to other features. Classic cases include finding the closest feature in another dataset and finding all features within a specified distance of a selected feature. Examples of this type of analysis include

• Finding the water distribution node closest to a parcel centroid, so that the billed water demand may be assigned to that node.
• Finding all parcels within 300 feet (91 m) of a set of water distribution pipes, so that the owners can be notified that a construction project will take place from October to November. In Figure 12.16, the dark pipes in the center of the figure were selected and buffered by 300 ft (91 m) to select the parcels to be notified.
Using Relationships to Trace Networks

When structured properly, datasets such as those that make up a hydraulic model (pipes and nodes), can be subjected to a process known as network tracing. In network tracing, GIS software uses information on which links are connected to which
nodes to allow a system to be traversed. These functions are commonly applied to street networks and utility networks. Examples of this type of analysis include:

- Generating shortest-path driving directions from point A to point B.
- Indicating the location of a hydrant flushing into a catch basin and identifying the location at which the chlorinated water will enter a stream.
- Indicating the location of a water pipe break and identifying the valves that must be closed to isolate the break (see Figure 12.17).
- Tracing the network to identify segments that are disconnected, either due to inaccurate data or inadvertently closed valves. This analysis can be a great aid in GIS and model input data quality control.

![Figure 12.17](image)

Using trace analysis to simulate the isolation of a main break

**Using Combinations of GIS Capabilities to Perform Complex Analyses**

Although many advanced types of GIS analyses can be performed, the examples listed in the preceding sections illustrate some of the most common uses available through most commercial GIS packages and can be easily mastered. These simple capabilities can also be used in combination and series to perform more complex analyses. Examples pertinent to the water distribution industry include:
• **Developing or optimizing pressure zone boundaries:** A polygon map of the static hydraulic grade line in each pressure zone can be draped on an elevation model to calculate theoretical static pressures across an entire service area. The resulting pressures can then be color-coded and shown together with piping and valve locations to indicate where the boundaries might be adjusted for optimum service and to determine how undeveloped areas should be incorporated into the zones (see Figure 12.18).

![Figure 12.18](image)

**Figure 12.18**
Pressure zone topographic map (darker shading corresponds to higher pressure zone)

• **Locating potential sites for facilities:** The GIS can be used to identify good locations for water system facility sites. In this case, the GIS is not used as a source of data for modeling but as a way to present alternatives to decision-makers. As an example, Figure 12.19 illustrates the results of an analysis to identify the top five ranked sites for the location of a new water storage tank. About 13,000 parcels were ranked using various criteria, including parcel size, land ownership, distance to a large water main, and parcel elevation. The resulting scores were filtered to select the top five parcels. Better parcels are shown in darker colors. Areas shown in white were eliminated from consideration because of size or elevation constraints.

• **Locating potential sites for monitoring equipment:** Another example described by Walski (2002) shows how model results can be imported into a GIS and used with other data, such as property ownership and locations of utility-owned buildings, to determine good locations for water quality or pressure monitoring equipment. For example, siting a pressure monitor or chlorine sensor in an area where the pressure or chlorine concentration is almost always constant would not provide a great deal of information to the system operators. Model runs can provide data on locations in the system with widely fluctuating pressures or chlorine concentrations. These locations can be displayed in the GIS by using color-coding and can be overlaid with other layers to show desirability of monitoring locations in terms of utility land ownership, location of power, and existence of pump stations and vaults.
12.5 THE FUTURE OF GIS AND HYDRAULIC MODELING

Both GIS software and hydraulic models are embracing open computing standards promulgated by the International Standards Organization (ISO), the Open GIS Consortium (OGC), and others. These standards enable the once disparate systems of models and GIS to share databases and objects, allowing users to quickly make more informed decisions with less risk of using outdated or inaccurate information.

Users can look forward to even tighter integration between systems in the next few years as vendors begin to incorporate object technologies and object-oriented programming methods into their products. Ultimately, this technology will bring models and GIS so close to one another as to be indistinguishable in certain applications.
REFERENCES


DISCUSSION TOPICS AND PROBLEMS

Read the chapter and complete the problems. Submit your work to Haestad Methods and earn up to 11.0 CEUs. See Continuing Education Units on page xxix for more information, or visit www.haestad.com/awdm-ceus/

12.1 Match the definition with the GIS term on the left. Place the letter in the blank.

1) ___ Polygon a) Data structure in which features are represented by set of coordinates
2) ___ TIN b) A closed two-dimensional figure
3) ___ Raster c) Assign x-y coordinates to a location such as an address
4) ___ DEM d) Data structure made up of contiguous non-overlapping triangles
5) ___ Vector e) Type of projection of earth surface to Cartesian coordinates
6) ___ Analytical paradigm f) Assigns unique attribute value to even sized cells
7) ___ Geocode g) Polygons generated around points
8) ___ UTM h) File used for storing elevation data
9) ___ Thiessen i) Using GIS as a place to store mapping data

12.2 Consider a small pressure zone with known flow rates. The demands are to be assigned to the nodes based on the area of each node. A set of Thiessen polygons was created and their areas are listed in the table below. Given that the long term average demand is 85 gpm and the peak is 215 gpm, find the average and peak demand at each node as would be done by a GIS-based tool.

Average and peak demands

<table>
<thead>
<tr>
<th>Node</th>
<th>Area (acres)</th>
<th>Average Demand (gpm)</th>
<th>Peak Demand (gpm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>J-151</td>
<td>15</td>
<td></td>
<td></td>
</tr>
<tr>
<td>J-152</td>
<td>25</td>
<td></td>
<td></td>
</tr>
<tr>
<td>J-153</td>
<td>23</td>
<td></td>
<td></td>
</tr>
<tr>
<td>J-154</td>
<td>41</td>
<td></td>
<td></td>
</tr>
<tr>
<td>J-155</td>
<td>12</td>
<td></td>
<td></td>
</tr>
<tr>
<td>J-156</td>
<td>15</td>
<td></td>
<td></td>
</tr>
<tr>
<td>J-157</td>
<td>11</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

When is it logical to expect water use to be proportional to the service area of the node?
12.3 The interpolation for elevation used in a GIS is based on two-dimensional TINs using three points. In the simple problem below, interpolate the elevation at point \( x \), given the elevation at points \( x_1 \) and \( x_2 \).

Extracting elevation data from points is based on the assumption that the points are close enough such that interpolation gives accurate results. What would the elevation at \( x \) be if, between the points, there is a steep embankment as shown by the dashed gray line?

12.4 Using the nearest node method based on customer meters, the demands assigned to a given node are 16 gpm. What demand should be placed on the node if unaccounted-for water is 20 percent of production?

12.5 Given the demand (already corrected for unaccounted-for water) for the land-uses in the following table:

<table>
<thead>
<tr>
<th>Demand Type</th>
<th>Unit Demand (L/day/hectare)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Single Family Residential</td>
<td>1,400</td>
</tr>
<tr>
<td>Multi-family Residential</td>
<td>1,800</td>
</tr>
<tr>
<td>Commercial</td>
<td>1,200</td>
</tr>
<tr>
<td>Light Industrial</td>
<td>2,500</td>
</tr>
<tr>
<td>Open Space</td>
<td>100</td>
</tr>
</tbody>
</table>
And the following land uses for each node:

<table>
<thead>
<tr>
<th>Node</th>
<th>Area</th>
<th>Land Use</th>
<th>% in Land Use</th>
</tr>
</thead>
<tbody>
<tr>
<td>J-201</td>
<td>5</td>
<td>Single Family Residential</td>
<td>100</td>
</tr>
<tr>
<td>J-202</td>
<td>8</td>
<td>Commercial</td>
<td>100</td>
</tr>
<tr>
<td>J-203</td>
<td>12</td>
<td>Single Family Residential</td>
<td>65</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Commercial</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Light Industrial</td>
<td>15</td>
</tr>
<tr>
<td>J-204</td>
<td>15</td>
<td>Single Family Residential</td>
<td>40</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Multi-family Residential</td>
<td>60</td>
</tr>
<tr>
<td>J-205</td>
<td>9</td>
<td>Light Industrial</td>
<td>75</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Open Space</td>
<td>25</td>
</tr>
</tbody>
</table>

Determine the demand at each node in Liters/day.

To perform the above calculation in a GIS, it is necessary to have two different polygon layers. Describe them?

12.6 In building a GIS for water distribution system modeling, indicate whether the following items should be vectors (also indicate if they should be point, line, or polygon), rasters, or TINs. There may be more than one correct answer for each depending on the system.

<table>
<thead>
<tr>
<th>Item</th>
<th>GIS representation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Junction Node</td>
<td>Vector (point)</td>
</tr>
<tr>
<td>DEM</td>
<td></td>
</tr>
<tr>
<td>Tank</td>
<td></td>
</tr>
<tr>
<td>Pipe</td>
<td></td>
</tr>
<tr>
<td>Aerial Photo Background</td>
<td></td>
</tr>
<tr>
<td>DXF File Background</td>
<td></td>
</tr>
<tr>
<td>Raw Water Reservoir</td>
<td></td>
</tr>
<tr>
<td>Pump</td>
<td></td>
</tr>
<tr>
<td>Node Service Area</td>
<td></td>
</tr>
</tbody>
</table>