# UPDATING THE NATIONAL HYDROGRAPHY DATA FOR THE TWIN CITIES METROPOLITAN AREA WITH LOCAL SUBSURFACE DRAINAGE INFORMATION

**Steven M. Kloiber**
**Jennifer Hinz**

Metropolitan Council
390 North Robert Street
St. Paul, MN 55101
Telephone: (651) 602-1056
Fax: (651) 602-1130
E-mail: steve.kloiber@metc.state.mn.us

## TABLE OF CONTENTS

## ABSTRACT

The National Hydrography Dataset (NHD) is a digital spatial dataset that describes the location, extent, and relationships of surface water features for the entire United States. The NHD is a key federal and state GIS dataset that is used for making maps, reporting data, and supporting environmental modeling and analysis. However, the NHD is lacking some important information on drainage for urban areas. Most of the GIS drainage features in NHD were derived from aerial photo interpretation. Therefore, subsurface features such as storm sewer pipelines are largely missing. In heavily urbanized areas, storm sewers can be the dominant drainage pathway. The deficiency was addressed for the Twin Cities Metropolitan Area by aggregating, selecting, and standardizing locally generated GIS data for urban storm water systems.

Local drainage data was acquired for 38 different organizations. The combined dataset included over 200,000 individual pipe features with a total length of almost 10,000 km. Locally generated data exhibited a wide array of data issues including incomplete data, missing features, lack of connectivity, lack of directionality, inconsistent attributes, and lack of metadata documentation. These data were cleaned, processed, and analyzed to identify drainage features with a high degree of hydrologic influence. Selected drainage features were merged into 1,078 features with a total length of 588 km and added to the NHD data for two subbasins: Mississippi River - Twin Cities (07010206) and Lower Minnesota River (07020012). Recommendations for future improvements to the process include: 1) address data quality and inconsistency issues of locally generated data by developing and promoting a simplified GIS data standard and tools for local data generators, and 2) direct additional resources toward updating and improving the documentation for the NHD editing tools.

## INTRODUCTION

The National Hydrography Dataset (NHD) is a digital spatial dataset that describes the location, extent, and relationships of surface water features for the entire United States. The NHD is comprised of several inter-related geographic information system (GIS) feature classes including points (e.g. gaging stations), lines (e.g. bridges and dams), areas, flow lines, and waterbodies (USGS 2007). Not only are these data important for making maps, but they also support a wide array of other water management activities. For example, NHD includes a spatially unique reach code system that all state and tribal agencies must use when reporting data to the USEPA to meet the reporting requirements of the Clean Water Act (USGS and USEPA 2000). Using a common spatial addressing system like the reach code system solves an important technical problem for data sharing. NHD has also been incorporated into hydrologic and water quality models such as the Soil and Water Assessment Tool (SWAT) and Better Assessment Science Integrating Point and Nonpoint Sources (BASINS) to aid in watershed delineation (Di Luzio et al. 2002). The capabilities of NHD have also been extended by incorporating it into the ArcHydro data model (Maidment 2002) which includes additional information on terrain and watershed boundaries.

Like all GIS data, the NHD is an imperfect representation of real world features; partly because the data are an abstraction of the real world, and partly because not all features of interest are represented or some of the representations are inaccurate or dated. However, the federal agencies leading the development of NHD have created a system for maintaining and improving the data that relies on a shared development network of federal, state, and local entities. This paper describes a pilot study to improve the flow line features of the NHD for the Twin Cities Metropolitan Area (TCMA) of Minneapolis and St. Paul, Minnesota by incorporating additional local drainage information. Information gained through this effort should benefit future efforts to update the NHD for other metropolitan areas.

This project aims to correct a deficiency in the density of drainage features for the Twin Cities Metropolitan Area (Figure 1). The flow line feature class is used to represent the hydrologic drainage systems as a single-line, flow network. The flow line features for NHD, including streams, ditches, canals, and pipes, were principally derived from the drainage features of the USGS topographic maps. These maps were created using photo interpretation techniques, so, in general, only features visible on aerial photographs were mapped. In most urban areas, a major component of the drainage system is below the surface in the form of storm sewer pipes (Figure 2). Therefore, while NHD has a design that supports various water management applications, it lacks important local details for urban areas. On the other hand, many local governments have digital data on storm sewers, but these data are frequently developed without standardization or in a manner that doesn't easily support geospatial analyses. The goal of this project is to update the NHD for the TCMA to achieve a feature density in the urban area that is consistent with the surrounding rural areas by incorporating drainage features captured from local data sources.

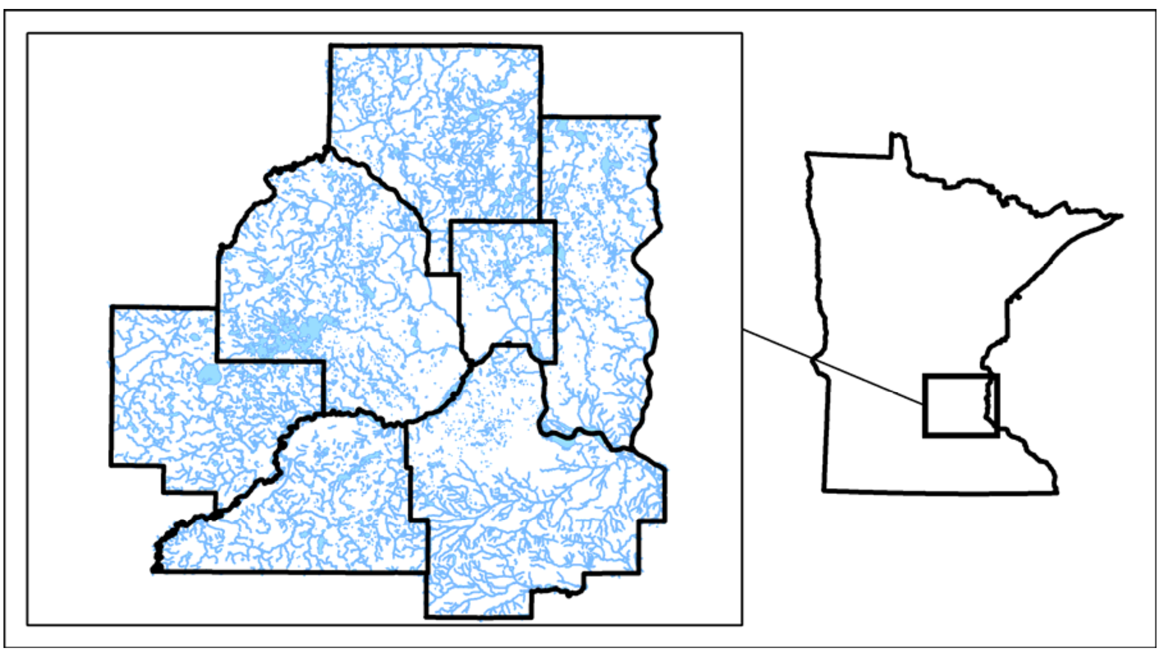

Figure 1: The project area is centered on the seven-county metropolitan area of the Twin Cities (Minneapolis and St. Paul, Minnesota).

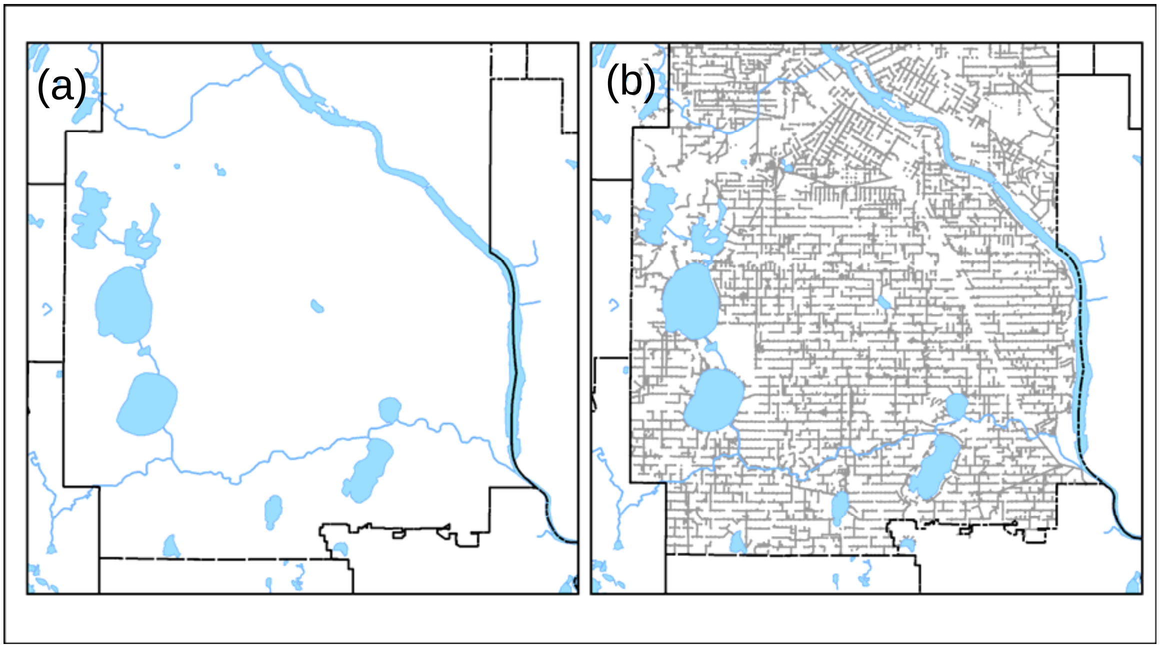


Figure 2: (a) The existing hydrologic drainage data contained in the NHD flow line feature class (in light blue) is missing a considerable amount of local drainage information. (b) The city of Minneapolis has over 60,000 linear drainage features in their storm water system geodatabase. The southern half of Minneapolis is shown here.

## Project Area Description

The TCMA includes a seven-county region centered on the cities of Minneapolis and St. Paul, MN. The region has a population of 2.85 million people (Metropolitan Council 2007). This region includes three major rivers: the Mississippi, the Minnesota, and St Croix Rivers. The project area focused on updating and enhancing the drainage information of the urban and suburban core of the TCMA (Figure 3) so that it will have a feature density similar to the exurban and rural portion of the seven-county area. The project area is encompassed by a rectangle 44km from east to west and 35 km from north to south. Updates were made to two subbasins: Mississippi River - Twin Cities (07010206) and Lower Minnesota River (07020012).

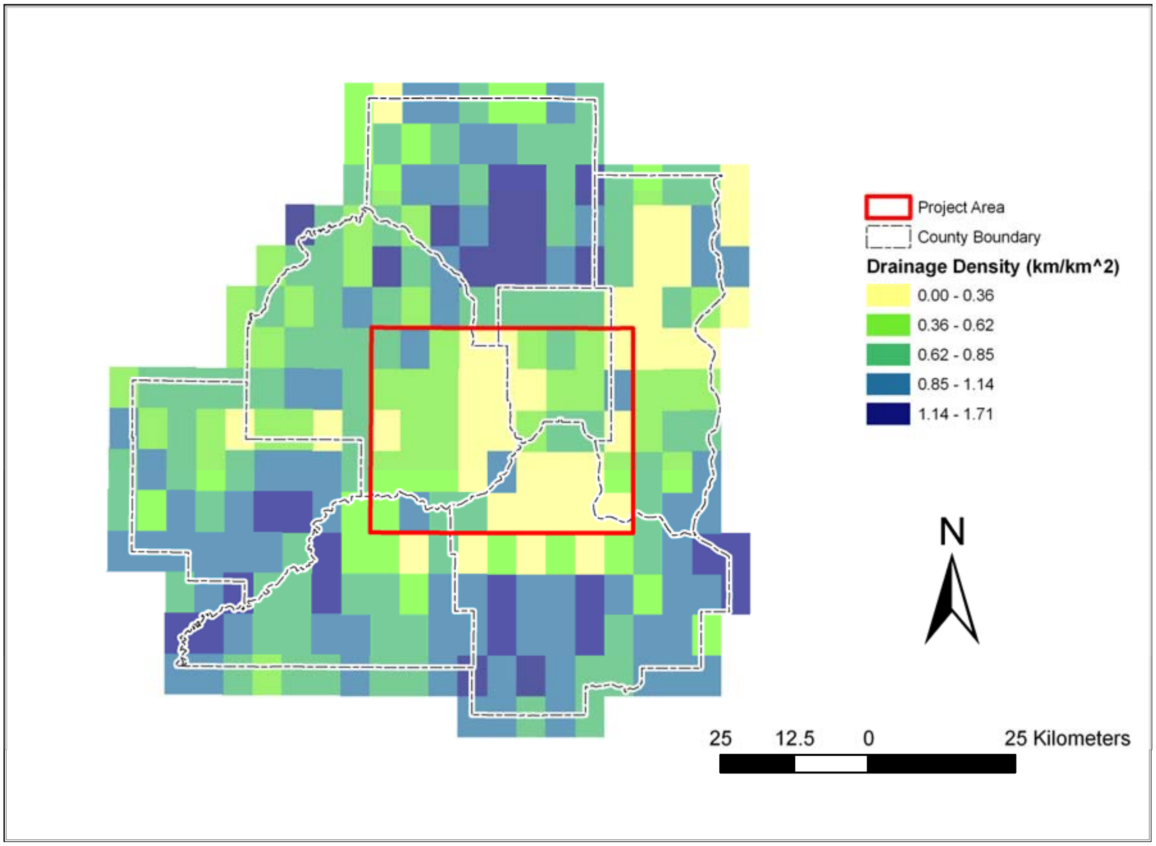


Figure 3: The density of existing NHD 1:24,000 scale drainage features for the TCMA is shown in the color gradient scale. Yellow and light green areas have the lowest drainage density. This project focuses on enhancing the drainage density for the urban core area within the red rectangle.

## METHODS

For this effort, we compiled local digital data on drainage with a particular focus on storm sewers. The local datasets were extracted from various native formats, transformed into a standardized format, georeferenced to the common coordinate system (Universal Transverse Mercator, Zone 15, North American Datum 1983) and loaded into a common geodatabase. These data were analyzed and ranked at the feature level according to their potential hydrologic influence. Features that drain larger contributing areas are given higher priority, as are features that connect to water features that are already mapped in NHD.

Hydrologic influence was determined by evaluating a number of characteristics in a semi-automated procedure. Pipe diameter, where available, was used as a surrogate for drainage area. We assumed a positive correlation between pipe size and drainage area because standard engineering practices dictate larger pipes for larger drainage areas. The total path length of a set of connected drainage features was also considered; however, data quality issues required a manual approach for this part of the assessment. A subset of the data with relatively high hydrologic influence was extracted from the storm sewer geodatabase and imported into the NHD.

NHD data was checked-out for editing and downloaded from the NHD website. The data were edited using version 3.2 of the NHD editing tool following the documentation provided with the software. The cleaned local data was loaded into the NHD using the import linework function. After importing, the data go through several quality control checks. These checks include topology checks for various types of geometry errors such as dangles and overlaps. A network analysis is performed to ensure correct connectivity and directionality. Attributes are checked to ensure that there are no duplicates in the unique ID field (ComID) and that all reach codes and ComIDs are assigned as positive integer values. And procedures are run to ensure that artificial flow paths do not occur outside of waterbodies and stream/rivers do not occur inside waterbodies. All of these quality control checks were performed independently by both the Metropolitan Council staff and the staff at the Minnesota Land Management Information Center (LMIC) prior to submitting the final data to the USGS.

## DATA ASSESSMENT

### Assessment of Existing NHD

The existing 1:24,000 scale NHD data for the TCMA has a considerably lower density of drainage features in the urban core than in the surrounding rural areas (Figure 3). The average drainage feature density for the urban area is 0.44 kilometers of linear drainage feature per square kilometer of land area. The average drainage density for the surrounding area nearly doubles this at 0.80 $km/km^2$.

This lack of drainage feature density hinders the use of NHD for various uses including assessing hydrologic connectivity. For example, the NHD contains 3129 waterbodies for the project area, mostly of the lake/pond feature type, but also quite a few of the swamp/marsh feature type. Of these, only about 20% (647 waterbodies) fall within 200 meters of a NHD flow line feature. Therefore, the NHD provides little or no information on hydrologic connectivity for the vast majority of these waterbodies. While some of these features may be landlocked, most of them do have a defined outlet, but that outlet has not been mapped in NHD.

## Assessment of Local Drainage Data

We successfully acquired local drainage data covering 38 municipalities in the project area (Table 1). From these data, over 200,000 linear drainage features (mostly pipes) were extracted to a geodatabase. About one-third of the data was delivered as computer-aided drawing (CAD) files which we had to import into GIS. The remaining files were delivered as either GIS shapefiles or geodatabases, but it was readily apparent that some of these had been transformed from CAD to GIS prior to delivery.

Several data issues were encountered with local drainage data. These issues and their resolution are briefly described here.

### *Unknown geographic coordinate system*

Most of the local data sets came without metadata describing the geographic coordinate system. Some of the datasets even appeared to have arbitrary coordinate systems with a coordinate origin (0, 0) located in the lower left-hand extent of the data. The result of this is that many datasets were offset from their true geographic locations by significant margins (Figure 4) and incorrectly scaled. ArcGIS spatial adjustment tool was used to georeference these data prior to integration. This process required locating pairs of ground control points (GCP) on the imported data and on a georeferenced target dataset. A second or third order polynomial transformation was generated using these GCP pairs and then applied to the imported data.

**Table1:** Inventory of local drainage data acquired for the NHD update of the Twin Cities Metropolitan Area.

| City/Organization | Feature Count | Feature Length (m) | Data Delivery Format |
|---|---|---|---|
| Arden Hills | 2,228 | 93,502 | CAD |
| Brooklyn Park | 5,613 | 187,461 | CAD |
| Cottage Grove | 4,958 | 156,786 | CAD |
| Edina | 4,947 | 123,014 | CAD |
| Fridley | 5,399 | 188,013 | CAD |
| Hopkins | 884 | 46,556 | CAD |
| Little Canada | 1,372 | 53,171 | CAD |
| Maplewood | 6,922 | 284,504 | CAD |
| Mounds View | 1,318 | 44,566 | CAD |
| New Hope | 2,275 | 81,656 | CAD |
| Robbinsdale | 3,167 | 63,805 | CAD |
| Shoreview | 2,674 | 94,311 | CAD |
| Saint Anthony | 189 | 23,404 | CAD |
| White Bear Lake | 3,176 | 96,362 | CAD |
| Blaine | 7,384 | 233,520 | GIS |
| Bloomington | 12,479 | 556,077 | GIS |
| Brooklyn Center | 3,371 | 128,433 | GIS |
| Burnsville | 12,087 | 346,427 | GIS |
| Columbia Heights | 2,424 | 71,448 | GIS |
| Crystal | 2,846 | 95,239 | GIS |
| Eagan | 14,443 | 464,178 | GIS |
| Eden Prairie | 3,936 | 326,834 | GIS |
| Falcon Heights | 544 | 15,013 | GIS |
| Golden Valley | 5,945 | 168,272 | GIS |
| Inver Grove Heights | 5,264 | 152,232 | GIS |
| Mendota Heights | 2,435 | 82,252 | GIS |
| Minneapolis | 60,691 | 3,856,447 | GIS |
| Minnesota Dept. of Transportation | 4,510* | 70,759 | GIS |
| New Brighton | 2,837 | 96,258 | GIS |
| North St. Paul | 2,157 | 66,510 | GIS |
| Plymouth | 8,143 | 252,333 | GIS |
| Richfield | 6,127 | 143,758 | GIS |
| Rosemount | 4,811 | 124,465 | GIS |
| Roseville | 7,323 | 235,788 | GIS |
| Saint Paul | 6,170 | 316,760 | GIS |
| Shakopee | 4,161 | 134,259 | GIS |
| Vadnais Heights | 970 | 41,137 | GIS |
| Woodbury | 13,697 | 381,378 | GIS |

* Many of the features from MNDOT were included in other local data.

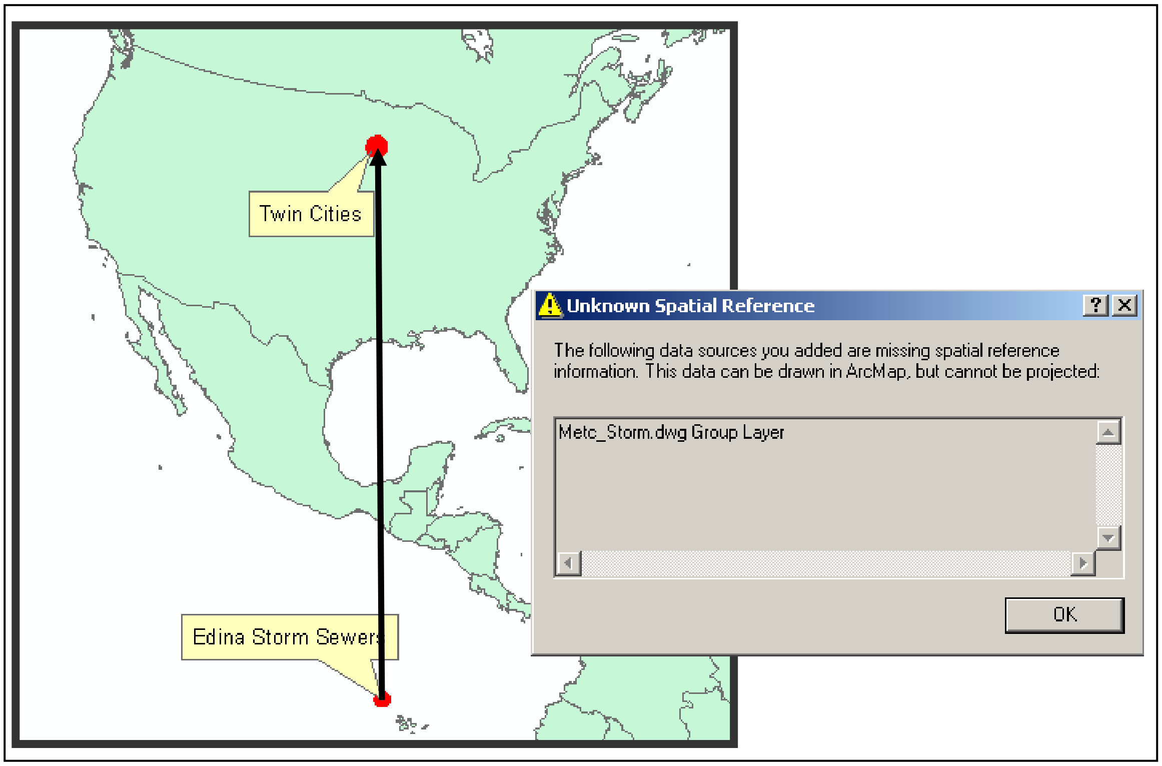


Figure 4: Much of the local data had undocumented geographic projections or in some cases arbitrary coordinate systems that created alignment issues. In this figure, the storm sewers for the City of Edina initially appeared to be located in the Pacific Ocean near the equator. The spatial adjustment tool in ArcGIS was used to correct these alignment issues.

The local GIS community in the Twin Cities has made major strides in improving GIS data sharing by establishing a regional organization, MetroGIS, to promote standards for data and metadata documentation. However, the development of a GIS data standard for local drainage data was only recently begun and it will probably be sometime before any of the local governments have adopted and implemented a data standard.

***Mixed feature classes***

Some local drainage data included mixed feature types such as manholes, catch basin inlets, and parcel boundaries (Figure 5). Parcel boundaries were often helpful in the georeferencing process because we could easily find paired GCPs at the intersections of parcel boundary lines, but our overall process required that these features be removed from the data before integration.

Strategies for removing unnecessary feature classes varied from data set to data set. In some cases, fairly obvious data attributes could be queried to select and remove these features. In other cases, a small set of features targeted for removal would be visually identified and selected on-screen. The attributes of these features could then be reviewed to find some unique characteristic. For example, the data for Edina, Minnesota includes catch basins and manholes (Figure 5). These features are polyline features with a circular shape. It turns out that all of the catch basins have a length of exactly 7.65 meters while the manholes have a length of 9.56 meters. We could then select all such features by querying the data on the feature length attribute.

### *Lack of Connectivity*

There are two main classes of problems that lead to a lack of connectivity in the local drainage data sets; missing features and undershoots. The lack of connectivity is a significant issue because one of the principal design elements of NHD is a connected hydrologic network.

#### Missing Features

Gaps in local drainage data is a significant issue affecting the update process. These data gaps arose for several reasons, but two of the more common reasons include a local data model that is restricted to pipe features or the local data model only includes features that are owned by the jurisdiction.

In the first case, pipe features are included in the local data set, but other important hydrologic connections are omitted. For example, a pipe that discharges to a ditch might be included, but the ditch itself is not. Another common occurrence is for pipes entering and leaving a pond or wetland to be captured, but there is no flow path connection through the pond or wetland (Figure 6). Very few of the local data sets were created with an eye toward creating a connected hydrologic network like NHD. Therefore, we

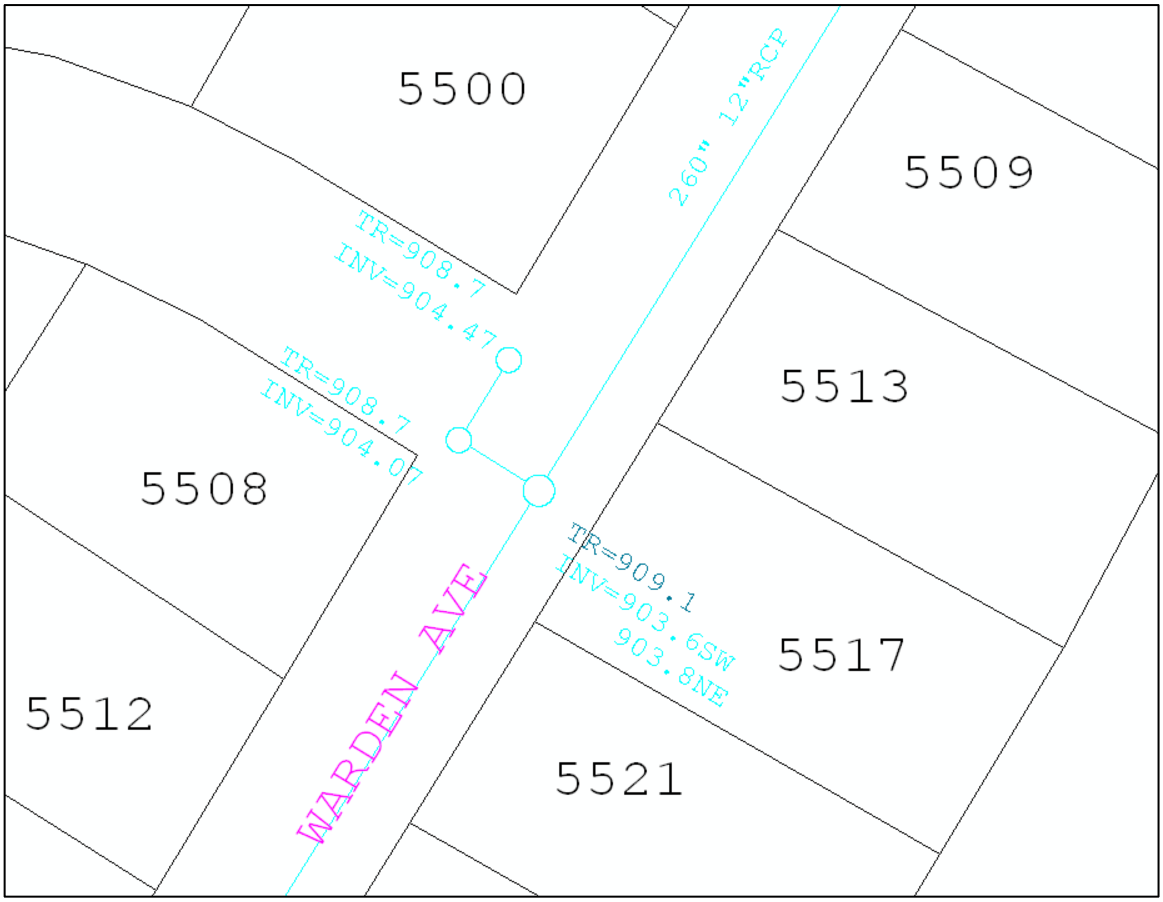

Figure 5: Local drainage datasets frequently included mixed features. This example shows the storm sewer pipe lines in light blue (cyan), property boundaries are shown in black. The data also include stormwater catch basins (small light blue circles and manholes (slightly larger circles). Annotation is included for pipes, manholes, catch basins, parcels, and streets.

frequently had to interpolate connecting features to fill in the gaps for these local data sets. Our general approach to address this issue was to overlay the georeferenced local data on high resolution aerial photography. We then manually added pond and wetland connectors and drainage ditches that were missing from the existing NHD and the local drainage data (Figure 6).

The second case leading to data gaps was a result of jurisdictional issues. The typical case occurs when a city only mapped pipe features that it owned regardless of whether the pipes were within its municipal boundaries. For example, if a county road or state highway went through a city, the city may show storm sewer that end in the middle of a road, but not the county or state-owned storm sewer that continues from that point (Figure 7). Our approach to address this issue was to try to obtain as much local drainage data from as many sources as possible, including overlapping jurisdictions such as cities, counties, and the state. However, this led to a problem of duplicate features in cases where the local data model did include features from other jurisdictions.

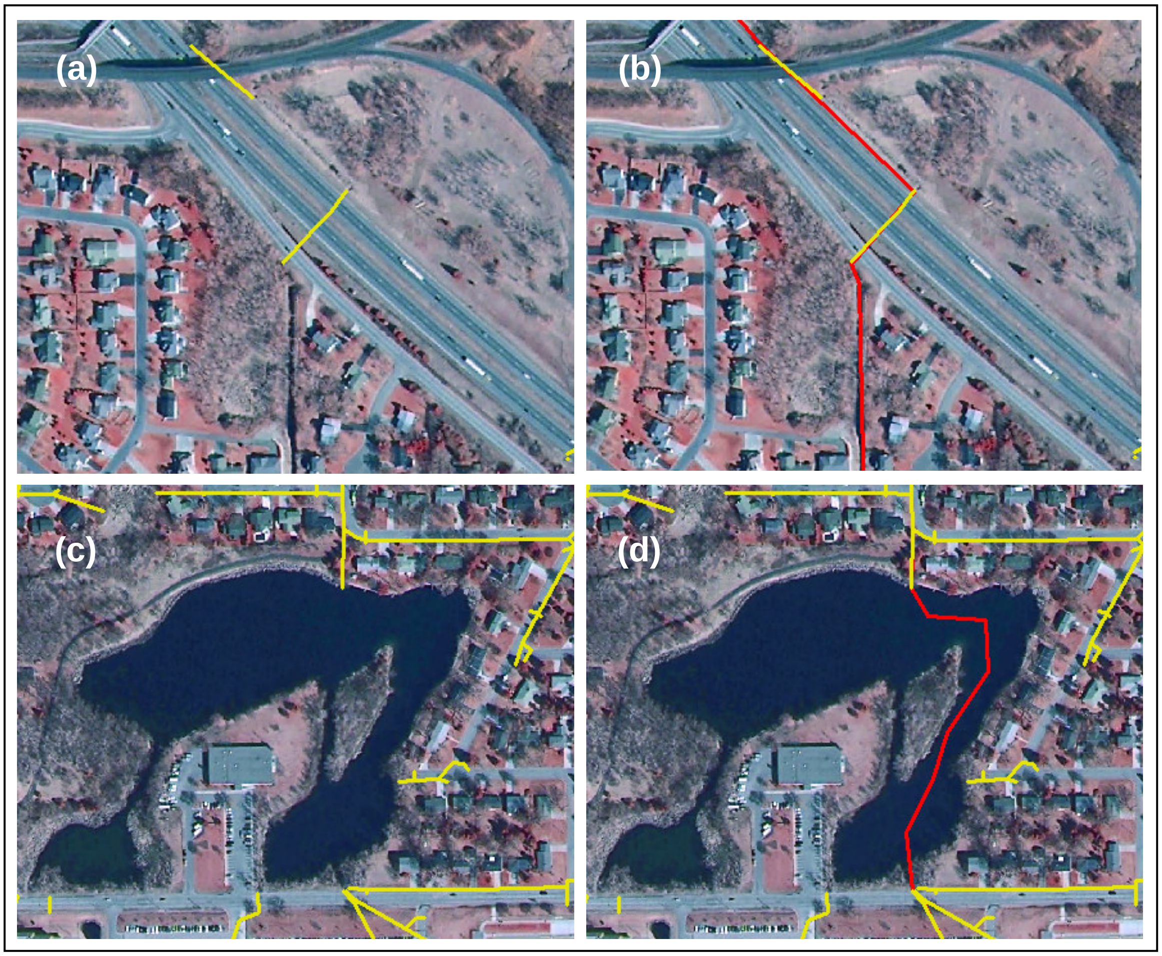


Figure 6: Local drainage datasets frequently exclude surface water connections such as ditches and artificial path connectors through lakes, ponds, and wetlands. Figure 6 (a) and (c) show pipe features in yellow from the local drainage data. Figure 6 (b) and (d) show drainage ditch features and an artificial flow path connector through a lake that were not in either the local data or the existing NHD data.

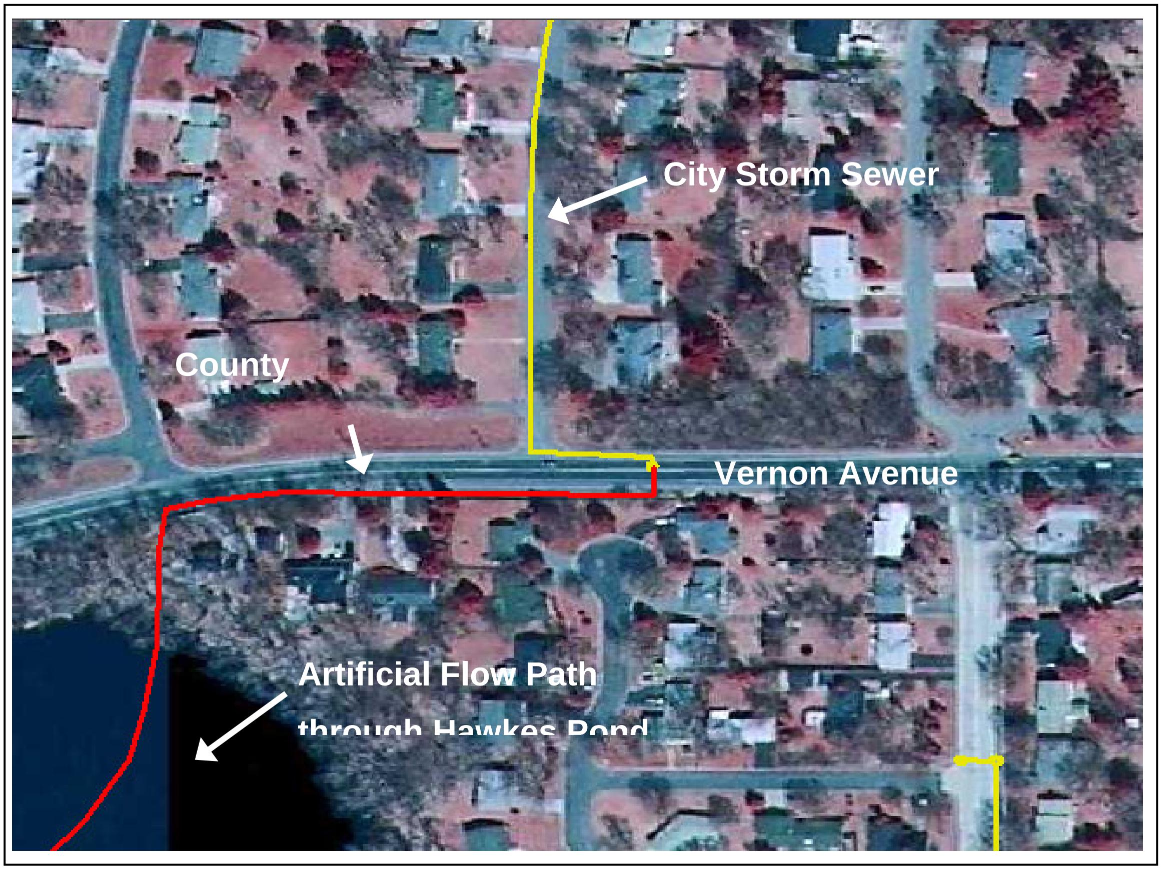


Figure 7: Local drainage data for some cities only includes features that the city owns. In this figure, a city-owned pipe carrying stormwater from Highland Park Pond terminates in the middle of Vernon Avenue. However, there is a county owned pipe under Vernon Avenue that connects to the city's pipe and carries stormwater to Hawkes Pond.

**Undershooting**

Undershooting, a type of dangling node, is another major connectivity issue. An undershoot occurs when a line falls short of another line that it should intersect. In this case, there isn't necessarily a missing feature or incomplete data model as much as the features that were digitized were not properly snapped in an end-to-end fashion. Undershooting tends to exhibit as very narrow gaps between line segments that are supposed to connect (Figure 8). Many of these undershoots occurred in data sets where the pipe features were snapped to a representation of a manhole or a catch basin inlet instead of snapping the linear pipe features end-to-end. Our approach to reconciling these

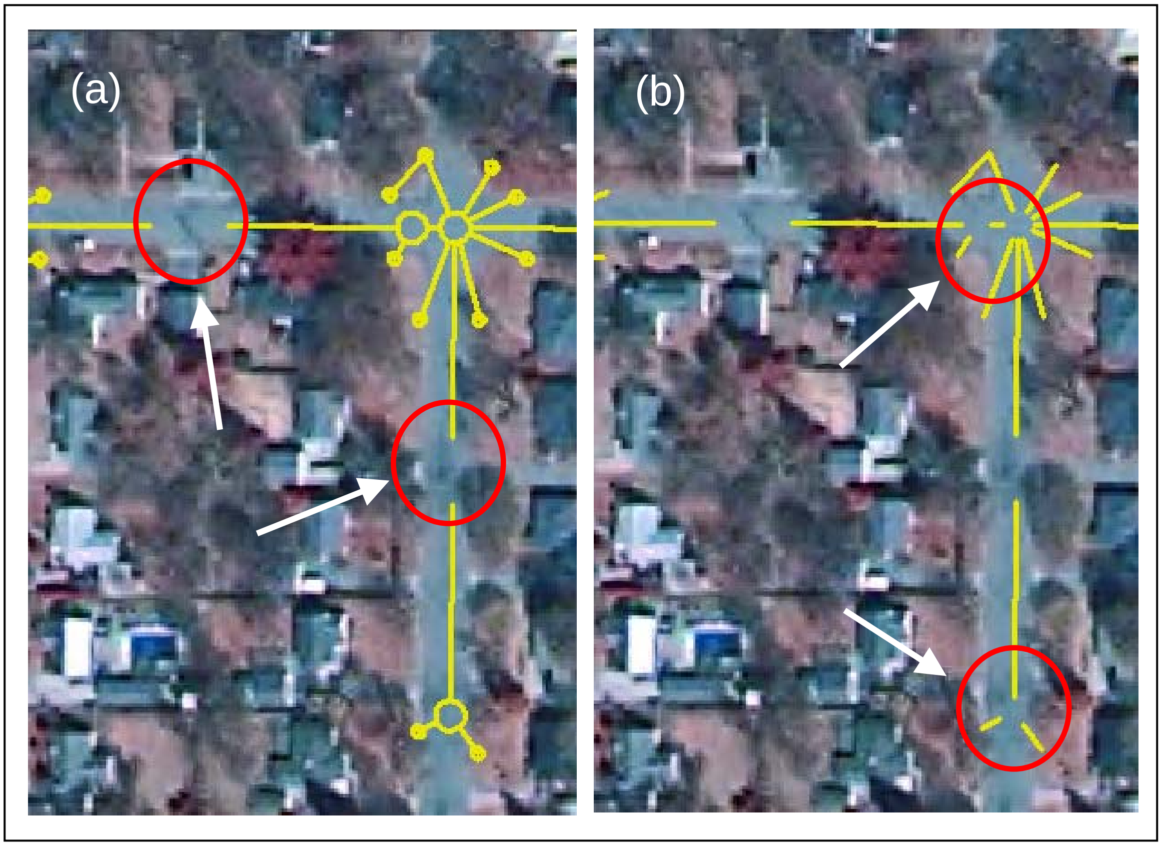


Figure 8: (a) Some local data exhibit short unexplained breaks. (b) Other small breaks are a result of line work being snapped to manholes and catch basins represented as circles, rather than creating a pipe network. The gaps between pipe features become readily apparent once the catch basins and manholes are removed from the data.

gaps was to manually edit the data with a snapping environment set to ensure coordinate coincidence for our selected features. We only fixed undershoots for the features we selected to include in the NHD. We chose not to use an automated procedure to snap flow line features because initial testing suggested that this created additional errors. Overshoots may also be a problem in some datasets, but these are not nearly as prevalent and undershoots.

***Lack of directionality***

In addition to connectivity, directionality is a key property of hydrologic networks like the NHD because it lets us resolve important relationships such as determining what is upstream and downstream of a given point. In GIS data, directionality is implied by the order of the vertices that make up an arc. The first vertex is the upstream end of the arc and the last vertex is the downstream end. Some of the local data was digitized following flow direction (Figure 9), but much of it was not aligned with flow direction.

We corrected flow direction for only the features selected for incorporation into NHD. Where necessary, we manually flipped arcs to align them with flow direction. Flow direction was verified by symbolizing the features with an arrow at the end with the last vertex. Since the vast majority of storm sewers are drained by gravity flow, there is a correlation between pipe flow direction and surface elevation over longer distances. Some pipes may cut against the surface topography for short distances, but eventually all the water ends up at one of the two major river valleys in the project area. Directionality and connectivity were also tested at various points in the NHD editing process using the network build and trace functions.

Figure 9: The storm water drainage data for Cottage Grove, MN was connected and directional even though it was created in a CAD system rather than GIS. The arrows are symbolically rendered by the GIS software and indicate the direction of flow. This particular data set also appeared to include some surface water features such as the backyard swale/ditch on the right hand side of the image.

***Incomplete and inconsistent attributes***

Another issue that impacted the update process was the inconsistent or incomplete attributes. This issue affected two parts of the process. The lack of a standardized attribute field for feature type made it difficult to select and separate the features of interest from data sets with mixed feature classes. The other issue with data attributes was the incompleteness of the pipe size information. We used pipe size as a surrogate measure for hydrologic influence, but the source data only had pipe size information for about half of the features.

Our approach to address filtering out mixed features classes was described earlier in this paper. Our approach to evaluating hydrologic significance for the purposes of selecting a subset of features for inclusion in NHD was to first use the pipe size information that we did have. Afterwards, we visually examined the combined local drainage data with reference to the existing NHD. We identified areas that were lacking drainage information in NHD. Within these areas we looked for local drainage data that met two main criteria: 1) the local data indicated a long network branch and 2) the local data could be readily connected to the existing NHD network.

## DATA INTEGRATION

### <u>NHD Editing Process</u>

As discussed in the method section, we loaded the cleaned local data into NHD using the import linework function of the NHD edit tool. The import linework function allows the user to bring in features from an outside source, but it only supports the importation of a single feature type at a time. Our local data geodatabase contained several feature types including pipelines, canals/ditches, connectors, and artificial paths. We tested two options for importing the data: 1) importing all the data as a single feature type and then correcting the feature type in NHD, and 2) importing the linework in subsets based on feature type. Using the first method, considerable time and effort was spent finding and correcting all the feature type (FTYPE) attributes. In addition, the process of modifying the attribute resulted in a second attribute error. When the FTYPE field was modified the software reset the unique identifier field (ComID) to zero. Further effort was then required to resolve these errors by repopulating this field with the original ComID value. The second method of importing linework avoided the issue of attribute errors, but under some circumstances it created geometry errors. This process was sensitive not only to the parameters set for the snapping environment, but also to the order in which the data were imported. The import seemed to have fewer snapping problems if we started with the feature type that was most common in the source data. Of these two methods, importing the data in subsets based on feature type was the preferred method because it was less labor intensive.

Other issues we encountered during the editing process included a problem assigning reach codes and a problem building the network. Part of the problem with assigning reach codes was due to the security settings on the Metropolitan Council's firewall. In addition, even after the firewall issue was resolved, the global process "assign reach codes" only seemed to work for a single feature at a time; however, the process "add permanent reach codes - selected" did allow us to assign reach codes for all selected features. Our work-around solution was to select all the features we added and use this second function to assign reach codes. The process to build the geometric network using the NetBuild tool was not successful and we were never able to resolve the problem. We did find that the USGS's FlowCheck tool provided similar functionality and was able to successfully build the geometric network. However, the network it built appeared to be a complex network allowing feature connections at unsplit edges. Our work-around solution for this was to build a simple geometric network using the Geometric Network tools in ArcCatalog. Specifying a simple network to be built, the unsplit edges would not be included in the network and could be identified using the Utility Network Analyst extension in ArcGIS. These edges were then manually split.

**<u>Summary of NHD Updates</u>**

In total, we identified and added 1078 linear drainage features (pipelines, ditches, connectors and artificial paths) with a total length of 588 kilometers (figure 10). These features correspond to approximately 6,000 individual features from the local data sources, but the local pipe data had to be merged into continuous pipeline features to be consistent with the NHD data model. In addition, we connected another 363 NHD waterbodies that were previously disconnected from the NHD drainage network.

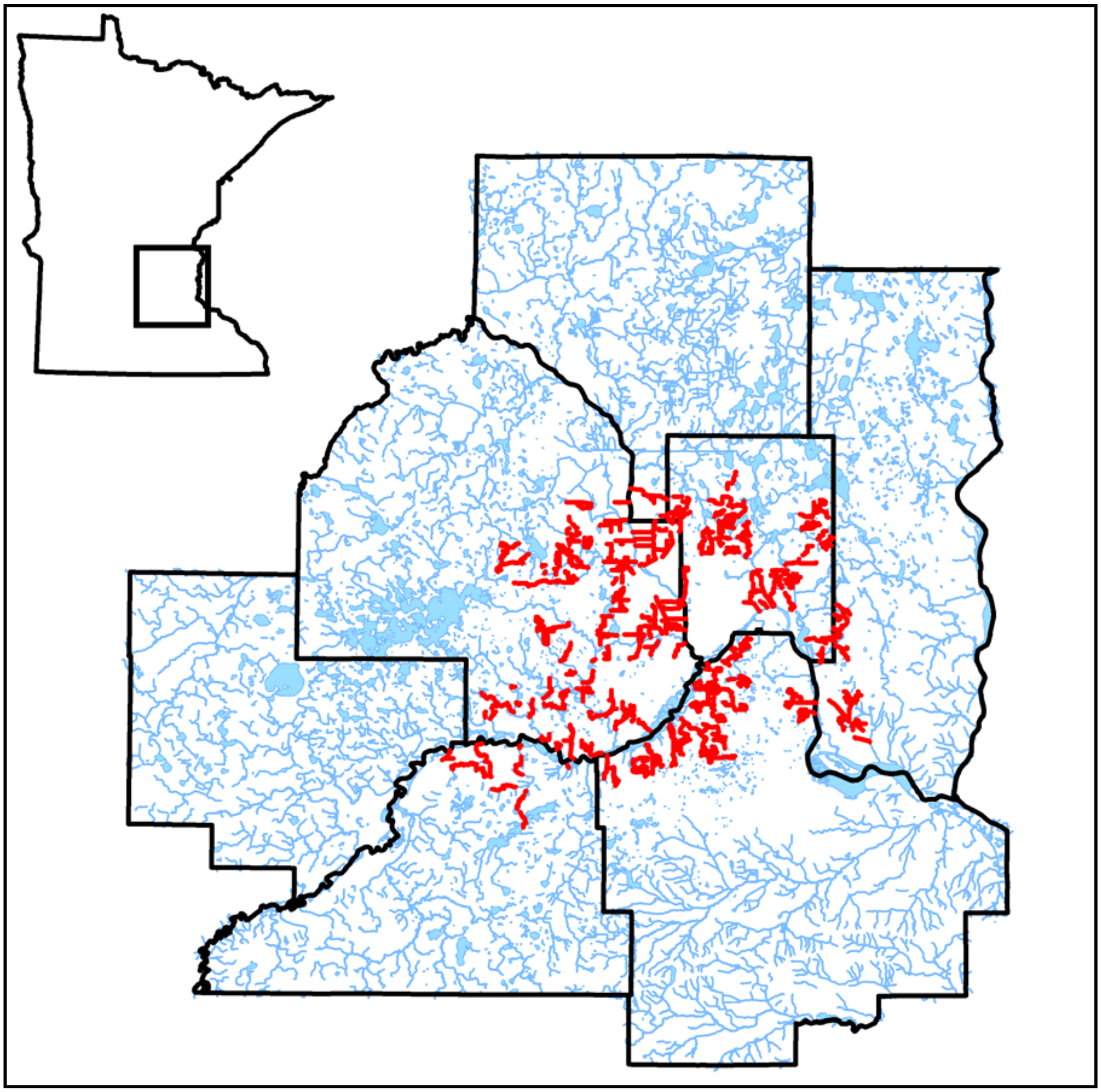

Figure 10: Over one thousand features were added to the NHD through this project. The new features are depicted here in red.

## DISCUSSION

### Data Quality

The NHD editing process includes several quality control checks including checks for connectivity, directionality, dangles, overlaps, and attribute accuracy. The local data that was incorporated into NHD was edited to ensure that it was topologically correct. In addition, the attribute accuracy was verified for all features that were added or modified during this process.

Some of the local data we received may have positional errors. Some cities and agencies verify at least some of their infrastructure locations with global positioning systems, but others do not. We did not perform an independent quantitative analysis of the position accuracy of these features. However, we did perform a qualitative evaluation of each of the data sets that we georeferenced by comparing them against other data sets with a known accuracy such as our regional parcel data set and a set of orthorectified imagery. Data that did not align properly with other infrastructure such as ponds and streets was reprojected. Given that standard residential street width are 35 feet wide, it would be a reasonable estimate to say that most (90% or more) of the local drainage features are within ±35 feet of their true location. The positional accuracy standard for rivers and streams in the high-resolution (1:24,000 scale) NHD is ±40 feet (USEPA and USGS 1999).

### Comparison to Similar Efforts

A review of the literature for this project found only one documented effort to evaluate and update the national hydrography dataset by incorporating local storm sewer drainage information. Sheng et al. (2007) evaluated the high-resolution NHD for regional watershed assessment in southern California. The most common problem found in the NHD for their project area was dangling streams, many of which were the result of missing connecting features.  Although some of these dangling streams end in areas of low relief and may be a function of the semi-arid climate. The NHD for the TCMA had relatively few dangling streams, but the local drainage data contained many of these errors.

Sheng et al (2007) used several approaches to address network connectivity issues. One method they used was to captured selected storm sewer features from a GIS dataset of Los Angeles County storm sewers. Using this data they added 168 new drainage routes with a length totaling 277 km.

Other types of errors found by Sheng et al. (2007) included flow divergences, attribute errors, and duplicate stream segments. Of these types of errors, the NHD for the Twin Cities was only really affected by errors in the feature type field (FTYPE). This was likely due to difficulties discriminating between streams and drainage ditches. All the attributes on features that were added through our update effort were verified for accuracy. Attribute errors in the NHD for features that were not directly impacted by our efforts were not addressed.

## Considerations for Future Efforts

### *Data Standards for Storm Water Systems*

There will continue to be an increasing need to find a solution to integrating local drainage data into a comprehensive, inter-jurisdictional hydrologic analysis system. Integrating local drainage data into a hydrologic network improves our capability for addressing important environmental and public health and safety issues such as flooding, chemical spill response, and water quality. Whether the solution to this problem is integration with NHD or some other solution, it is clear that the current highly variable state of local drainage data will make any comprehensive effort quite costly. The scale of such an effort can be gauged against the fact that this update of NHD took nine months to complete and we only added a fraction of the available local drainage data.

To address issues of storm sewer data inconsistency and incompatibility, a collaborative effort has been initiated to develop and adopt GIS data standards for the state of Minnesota. A multi-participant workgroup has been established to collect information, solicit input from stakeholders, review various user needs, and develop a data standard.

This effort is patterned after earlier successful efforts to develop GIS data standards and integrate local data sets (Johnson and Arbeit 2001) such as parcel boundaries.

In this model, the municipally separate storm sewer system operators (MS4) are the primary producers of the data. They will be encouraged to implement the adopted data standards through several mechanisms. Also, a state or regional agency would serve as the area integrator, responsible for assembling data from primary producers. As such, data standards become a key part of ensuring an efficient process for data integration.

***Data Sharing / Security Issues***

Developing a shared hydrographic network data set that incorporates detailed local information on storm sewers has significant and sometimes complicated implications for security issues. For example, during the course of this effort, very few organizations expressed much reluctance to share their data. However, we were only able to acquire about half of the storm sewer data for the city of St. Paul. Getting access to the remainder of the storm sewer data was affected by planning and security concerns surrounding the Republican National Convention. A few other organizations, also expressed some general concerns about the security risks involved, but they ultimately agreed to share their data as well.

Looking at the security issue from a different perspective though, Amstutz et al (2008) point out that there are clear risks associated with not sharing this information with affected organizations. The Nation’s water infrastructure system incorporates natural hydrologic systems such as river and stream networks, as well as water distribution systems, wastewater collection systems, and storm sewer systems. Managing these systems is a task that is complicated by the connectivity and feedback loops within and between these systems. Network analysis of these systems is an important component of planning, monitoring, and response strategies to address potential accidental or deliberate toxic contamination.

## CONCLUSIONS

A considerable amount of local drainage information is available for urban areas. This data is frequently in electronic format and can be incorporated into the National Hydrography Dataset to help provide additional important drainage information in metropolitan regions such as the Twin Cities. However, the current state of much of the data for the Twin Cities is such that considerable effort is required to clean up the data prior to incorporating it into the NHD or any other integrated network dataset.

The NHD is a good choice for a state or regional hydrologic network model because it has a well developed and tested network data model and because it is supported by the U.S. federal government. However, one potential drawback to NHD as a data model is that the tools for the collaborative maintenance of this dataset are relatively new and quite complex. Even the installation of the NHD editing tools is a non-trivial task. This complexity of the tools and maintenance process may limit the participation in direct updates to a few dedicated data stewards in each state. Investing in improved documentation and testing will address some of this problem. In addition, it may be helpful to develop an extension of the NHD stewardship model that includes a simpler system and data model that can be used by local data generators. Then data from the local data generators can be more easily aggregated by state or regional data stewards and incorporated into the NHD.

## ACKNOWLEDGMENTS

Funding for this project was provided by the U.S. Geological Survey, thanks in large part to the efforts of Paul Kimsey. Thanks also go to Ron Wencl who served as our project liaison with the U.S. Geological Survey. We gratefully acknowledge the generous expert advice and assistance received from the staff at the Minnesota Land Management Information Center, including Sandi Kuitunen, Susanne Maeder, and Jim Krumrie. We would also like to thank the many organizations that shared their data with us for this project.

## REFERENCES


Amstutz, D.E., Bahadur, R., Pickus, J.M., Monteith, M.C., Johnson, J.P., Ziemniak, C.J. Chehata, M., Herrera, M.E., and Samuels, W.B. 2008. The integration of network-based models for spill response and homeland security. American Journal of Environmental Science. 4:6:544-550.

Di Luzio, M., Srinivasan, R., and Arnold, J.G. 2002. Integration of watershed tools and SWAT model into BASINS. Journal of the American Water Resources Association. 38:2:1127-1141.

Johnson, R. and Arbeit, D. 2001. MetroGIS. In Lessons from Practice: A Guidebook for Organizing and Sustaining Geodata Collaboratives. Johnson, R., Nedovic-Budic, Z. and Covert, K. eds. GeoData Alliance, Reston Virginia. 29-42.

Maidment, D.R. 2002. ArcHydro: GIS for Water Resources. ESRI Press, Redlands, California.

Metropolitan Council. 2007. Twin Cities Region – Population and Household Estimates, 2007. Metropolitan Council. St. Paul, MN. 5 pages.

Sheng, J., Wilson, J.P., Chen, N., Devinny, J.S., and Sayre, J.M. 2007. Evaluating the quality of the national hydrography dataset for watershed assessments in metropolitan regions. GIScience and Remote Sensing. 44:3:1-22.

USEPA and USGS. 1999. Standards for National Hydrography Dataset. U.S. Environmental Protection Agency and U.S. Geological Survey. 131 pages.

USGS. 2007. National Hydrography Dataset (NHD) Model (v1.06). U.S. Geological Survey. National Geospatial Technical Operation Center – II, Denver, Colorado.

USGS and USEPA. 2000. The National Hydrography Dataset: Concepts and Content. U.S. Geological Survey and U.S. Environmental Protection Agency.