City of Santa Ana: GIS Integrated for Drainage Facility Inventory and Analysis

Jonathan Posner, GISP – CDM, GIS Project Manager
Joe Parco, P.E.– City of Santa Ana, Senior Civil Engineer

Background

The City of Santa Ana (City), with a population of nearly 340,000 residents and 27 square miles (mi^2) of commercial, industrial and residential areas, maintains approximately 1,600 storm drain inlets and 34,000 linear feet of open channels that transport urban runoff generated from nonpoint sources within the City. Runoff transported by these drainage facilities discharges to the Lower Santa Ana River, Newport Bay, and Bolsa Chica water bodies. Major pollutants typically found in runoff from urban areas such as the City of Santa Ana include sediment, nutrients, oxygen-demanding substances, heavy metals, petroleum hydrocarbons, pathogenic, and bacteria.

CDM collaborated with the City to write the Proposition 13 Nonpoint Source Control grant proposal, which outlined a program to decrease the pollutants reaching sensitive water sources. As a result of the grant proposal, the City was awarded a grant from the State Water Resources Control Board (SWRCB) under Agreement Number (II-63)-03-142-558-0. The grant extends through May 2006. Geographic Information Systems (GIS) was integral to this project as it provided the City the means to enhance their existing data assets.

Abstract

This paper describes the integration of GIS in the creation of the City's drainage facility inventory and analysis program. It will describe the GIS planning process, assembling drainage facility and industrial site information, incorporating inspection and global positioning system (GPS) location information into the GIS database, developing a drain inspection and cleaning component, and analyzing drain inspection and cleaning data and its use in supporting a targeted source reduction program.

Introduction

The implementation of the drainage inlet inventory and analysis centered on the use of GIS. The City recognized that the use of GIS to perform the tasks for the project would provide the best results with its capability for data integration, organized and centralized data storage, analytical capabilities, and ability to produce various outputs (i.e., statistics, maps, and reports). The ultimate goal for the City is to establish a long term program that would decrease the debris entering their storm drain inlets and show measurable results over the long term.

A description of the efforts involved in carrying out the GIS-related tasks is presented in this paper. These included:

- **Task A** - Collection of existing base and drainage facility data into a centralized GIS database

- **Task B** - Performance of GPS based location of drainage facilities
- **Task C** – Integration of Cleaning/Inspection data and Industrial/Commercial Sites
- **Task D** - Analysis

As with any project, there was a great deal of project management that required a measured amount of client interaction and team work. However, it was CDM's priority to not burden the City with a lot of extra work and to meet timelines for deliverables. Timelines were critical as they were prescribed by the Regional Water Quality Control Board (RWQCB) in the grant proposal.

Getting Started: Task A – Collection of Existing Base and Drainage Facility Data Into a Centralized GIS Database

The first task was to collect and integrate existing data. The collection effort served as an excellent springboard to develop location maps and other mapping products that helped explain issues that would arise throughout the project. It also provided the means to conduct the GPS field collection task in an organized and efficient manner as described below.

Source Data Collection

The data associated with the City's drainage facilities came in several formats and from varying sources. Formats included Microstation DGN files, shapefiles and coverages as exports from GeoMedia, native shapefiles and coverages, MrSID imagery, and spreadsheets. Data sources included the City, Thomas Bros. for street data, County of Orange agencies, and other City contractors responsible for various inputs to the program. Table 1 below outlines a portion of the data required for the project, their formats and sources:

Table 1 – Listing of Some Core Datasets, Their Source and Format

Dataset	Source	Format
Parcels	City via County Assessor	Export Shapefile from GeoMedia
City Boundary	City	Export Shapefile from GeoMedia
Parks	City	Export Shapefile from GeoMedia
Schools	City	Export Shapefile from GeoMedia
Streets	Thomas Bros.	Coverages
Land Use	City	Parcels joined to Access table
Existing Drainage Inlets	City	Microstation DGN
Old Map Grid	City	Microstation DGN
Open Channels	City	Shapefile
Industrial Site Locations	City via outside Consultant	MS Excel
Cleaning Inspection Forms	City via outside Contractors	MS Excel and Hardcopy
Imagery	City	MrSID

This data was integrated in ArcGIS using ArcView, ArcEditor, and ArcInfo as the need arose for certain functionality. Some of it needed simple conversion, while in other cases additional processing was required. For example, the Industrial Site Locations had to be geocoded (located geographically by their address information). The City provided this information in MS Excel format for approximately 1,700 locations. There are several steps that must be carried out in order to achieve a high percentage match. Geocoding is also very dependent on the quality of the address ranges of the street centerline data used. Once completed with the processing, we achieved an approximate 90 percent match rate with a score of 80 or better.

Data Structure

For any GIS related data development effort it is important to establish how the data will be structured. Initially we discussed the ability to integrate with the existing Orange County Watershed & Coastal Resources Division (OCWCRD) data model standard for their Stormwater Program. At the time we initiated work on this program (June 2004), the data model from OCWCRD was not finalized. Therefore, we made a decision to revisit this option later in the project.

A practical approach was agreed upon to deliver the results as shapefiles. The data would be developed very specific to the needs of this grant project and later we would review the need to integrate with a more robust structure as developed by OCWCRD.

How Many Drainage Inlets? : Task B - Performance of GPS Based Location of Drainage Facilities

The grant proposal required that we determine the location of the drainage inlets. We would also determine exactly how many drainage inlets the City owns and maintains. The initial state of the data was a DGN file containing locations that were determined to be inaccurate.

Drainage Inlet Identification

To provide the City with the greatest benefits from our effort, we proposed an updated identification scheme for the drainage inlets. This proposal tied in nicely to the City's broader GIS data structure direction. The City was in the process of establishing a new grid system that all new maps would be based upon. In hindsight, it made practical sense to do this since many of the drainage inlets did not correlate between what was found on the City's old maps and the field investigation.

Planning Efforts

We collected the drainage inlet information with handheld GPS devices—the GeoXT with TerraSync software from Trimble. Working with the City, we determined the attributes and values to collect for the drainage inlets. These included XY coordinates, elevation, drainage inlet opening length, shape of opening to drainage inlet, and whether it was an existing facility based on the City's old maps or not. After establishing the attributes and domain values to be collected, we entered those parameters into TerraSync on the GeoXT.

Based on the estimated number of features in the existing source data, it was estimated that the effort would take approximately eight to ten weeks to collect all the data. With effective team coordination, scheduling, and data management, we finished the GPS field collection in just six weeks. The planning carried out prior to the collection effort was attributed to the near 40 percent efficiency gain.

Initially, we conducted user training with the GeoXT and thoroughly covered the process for collecting the drainage facility data. A review of the source maps and a coding system was discussed to indicate which features were to be located on a daily basis. These source maps were an important asset in the review of the data collected and in determining data quality issues. The training also covered the layout of the City and the strategy for covering the City as efficiently as possible.

We organized two teams to collect field data in a designated territory. The City was divided into two halves with the dividing line being north-south running Bristol Street. Each team took with them the needed materials to collect the data: a Health and Safety Plan, measuring tape, an official letter stating we were conducting work for the City, the GeoXT GPS unit, and a copy of the old source maps of the drainage system.

Daily the teams collected field data in a designated portion of the City. Nightly, the teams downloaded their data from the GeoXT and uploaded it to the GIS. For quality purposes and to make sure the data was backed up on a server, it was important to offload data each night.

Each team indicated on hardcopy maps the location of drainage inlets. This information was then generalized and transferred to an index map. This procedure provided multiple places to evaluate each team's performance and to keep track what portions of the City were completed on a daily basis.

Quality Methods

As part of quality assurance, we used the software accompanying the GeoXT, called Pathfinder Office, and ArcGIS. The Pathfinder Office software was geared toward the assurance in the accuracy of the coordinate and elevation data collected at each location. ArcGIS was used to visually correlate locations collected with the historical data to determine validity of the captured feature, to perform counts in terms of the verification of the total number of drainage inlets collected, and to track team performance.

At the end of each collection day, data was downloaded and post-processed using Pathfinder Office. This included a process that corrected the data by comparing the coordinate and elevation data collected against a base station that is located in Palmdale, California. Post-processed correction was not always required as the GeoXT unit was capable of collecting data to within +/- 1 meter horizontal accuracy on-the-fly, as long as certain protocols were followed. This included assuring that the unit could receive signals from at least four satellites for several seconds. In any case, we ran the post-processing correction on each dataset to reassure the accuracy standards required were met with each data set.

The final process used in Pathfinder Office was converting the proprietary Trimble GPS data to traditional shapefiles that could be brought into the GIS. Each dataset had to have their projection defined to State Plane NAD 83 California Zone 406. Afterwards, the data was brought into the GIS and visual inspection of the data was conducted. It was during this routine visual inspection that any errors in the data were noted. This included duplicate drainage inlets, drainage inlets located with no shape attribute indicated, improper identification of the drainage inlet, and review of the comments entered by the field teams. All fatal flaws were sent back to the teams for recapture or were handled within the GIS in conjunction with further insight from team members and the source maps.

Recognizing that some drainage inlets would be hard to reach due to traffic conditions, the comments entered by the field teams became invaluable pieces of information to more accurately place them. The comments read something like: 'feature 30 feet north of GPS point'. Therefore, we performed this final step in the GIS to ensure a more accurate placement of the drainage inlets, and indicated such manual adjustments in the metadata of the deliverable shapefile.

Final Count of Drainage Inlets

After finalizing all the QA/QC, including some field verification, we had collected 1657 drainage inlets. This was approximately 10 percent more than the estimated 1500 drainage inlets suggested by the City.

Parallel Processes: Task C – Integration of Cleaning/Inspection Data and Industrial/Commercial Sites

Integrating the cleaning and inspection data and the industrial/commercial site data into the GIS was carried out while the effort to collect all the drainage inlets was occurring.

Cleaning/Inspection Data

Prior to CDM's involvement on this grant project, we had collaborated with the City to create a form for gathering relevant information during drainage inlet cleaning. The City recognized the importance of keeping watch on water quality issues and was also under regulatory guidelines to track this information. These forms became the basis for the structure we used to develop our data deliverables to the City. The forms contained the following information: date of inspection, old facility ID, cross street information, indication of a filter being present, type of inlet, comments regarding the dimensions of the drainage inlet, whether the debris was removed, and other comments.

The City contracted with two companies to carry out the drainage inlet inspection and cleaning work in late 2003 and early 2004. Each contractor was given approximately half the City to inspect and clean. The contractor made minor modifications to the hard copy forms for to record the drainage inlet and inspection and cleaning data. This data was integrated with the drainage inlet data. The data integration required creating some basic forms in MS Access for data entry of the hard copy forms. The data was then merged with the data captured in the MS Excel spreadsheets. The City made sure that the contractors identified each form with the drainage inlet ID that was on their old maps. We in turn used this ID to link the cleaning and inspection data to the GPS location of the drainage inlets.

Industrial Sites Locations

The industrial site locations were an important aspect of this grant program, and will continue to be as the program continues. The City had contracted with another consultant to analyze all the industrial sites within the City and prioritize them according to the 2003 Orange County Drainage Area Management Plan (DAMP). The database contained: business name, contact information, address, SIC code, General Industrial Permit No., vicinity to Area of Special Biological Significance (ASBS), and priority classification. Approximately, 102 industrial/commercial facilities were designated as "high priority."

Under the scope of this program, we determined if some sites may be impacting the debris volumes and the need to implement some best management practices to lessen their impacts.

The industrial sites were provided in the form of a basic Microsoft Excel spreadsheet. Using the street address for each site and the geocoding functionality in ArcGIS, we integrated these locations into the GIS. The City had also provided Thomas Bros. street centerline data, which is one of the more accurate dataset sources for accomplishing this process.

Geocoding is the process of spatially locating a point feature with an address to its geographical location. Essentially the street address location is matched to a series of address ranges assigned to lines in a centerline dataset and then tied to that line based on some embedded algorithms. These calculate the address of the feature and determine where that address falls on a particular line. For example if you had a site at 40 Main Street and the address range on the line was 10 to 100 Main Street that feature should be located approximately 40 - 50 percent of the way down the line.

In ArcCatalog you need to establish a geocoding service. This is a series of dialogue boxes that establishes the parameters you will use to perform the match. We found this to be a very iterative process to find the best parameters for the highest and most accurate matching. The results of the geocoding service can be revised in ArcGIS. Using the GeoCode Plus toolbar in ArcGIS you can step through your results and make manual adjustments and decisions regarding some of the more difficult locations.

As mentioned above, we achieved a very good match after stepping through the process a few times. We achieved a 90 percent match with a score of 80 (out of 100) or better.

Pilot Results: Task D - Analysis

The work during the first year of the program centered on three pilot areas. Data collected in subsequent years will be used to broaden our analysis to the entire city. The analysis focused on the evaluation of high accumulation areas throughout the City. What this meant was determining which drainage inlets were found to contain the highest debris volumes, and then to understand the spatial relationship of, and potential impacts by, high priority industrial/commercial sites. Another aspect of the grant program was to determine potential indicators and to help implement educational, and other best management practices for preventing debris from entering drainage inlets.

The cleaning and inspection contractors collected volume data for 220 drainage inlets that fell into the three pilot areas. The first area was located in the Northwest corner of the City and is bounded by Westminster Avenue to the North, Bristol Street to the East, Ward Street to the West, and 1st Street to the South. The second area is located in the Northeast corner of the City and is bounded by 17th Street to the North, Grand Avenue to the East, Santiago Avenue to the West, and 6th Street to the South. The third area is located in the Southwest corner of the City and is bounded by Segerstrom Avenue to the North, Bristol Street to the East, Hyland Avenue to the West, and Sunflower Avenue to the South.

Drainage Inlet Debris Volume Review

The volumes calculated at these drainage inlets ranged from 0.1 ft^3 to 80.0 ft^3 of debris (Note: 80ft^3 equates to about sixteen 33-gallon hefty bags). The City decided that in order to categorize the drainage inlets three equal ranges would be created with the maximum set to 80 ft^3. These ranges were used just for the pilot phase. Once data is collected in 2005 for the entire city, the ranges may be adjusted. Range 1 signifies the lowest volume range, Range 2 the middle, and Range 3 signifies the highest volume range. A color and symbol scheme was applied to provide easy visualization within the GIS environment and on mapping output.

The symbols and ranges from low to high were as follows:

0.01 – 26.67 ft3	–	Range 1
26.68 – 53.33 ft3	-	Range 2
53.34 – 80.00 ft3	-	Range 3

The breakdown of the number of drainage inlets falling within the defined ranges is as follows:

- Range 1 (0.01 – 26.67 ft^3) – 212 Drainage Inlets
- Range 2 (26.68 – 53.33 ft^3) – 6 Drainage Inlets
- Range 3 (53.34 – 80.00 ft^3) – 2 Drainage Inlets

Drainage Inlet Heavy Accumulation Areas

Upon review of the data, Pilot Area 1 contained drainage inlets with the highest debris volume relative to the other Pilot Areas, and is therefore considered the heaviest accumulation area. There are two drainage inlets from Range 3 and four out of the six drainage inlets in Range 2 were located in this area.

There were seven areas considered high accumulations in the three Pilot Areas. Pilot Area 1 contained five high accumulation locations. Pilot Area 2 contained two high accumulation locations. Pilot Area 3 did not contain any.

Potential Linkage to Nearby Industrial/Commercial Zones

The City implemented a color and symbol schema for the industrial site locations. Each industrial site was symbolized as either a large red triangle for high priority, a medium sized orange triangle for medium priority, or a smaller green triangle for low priority sites.

Using the symbolized drainage inlets and industrial site locations in the GIS environment proved an invaluable tool to determine spatial correlation. As an example, for the heaviest accumulation area of all the Pilot areas, it was apparent that, while there was no high priority industrial sites, there were several medium priority industrial/commercial sites that were within 2500 feet of the drainage inlet.

Although we can hypothesize that there may be a link between these industrial/commercial sites and the high debris volumes in the drainage inlets, it was not a 100 percent certain. Spatial proximity (in this case being located near something else) indicates that a relationship exists, but does not indicate the strength of that relationship or the specific variables involved. Strengthening these hypotheses will require further correlation with data to be collected in mid-2005 and in 2006.

The other intersections of high accumulation areas listed above show similar results. Table 2 indicates the number of and approximate maximum distance of high and medium industrial/commercial facilities for the listed high accumulation area locations (Distances are straight line distances.):

Table 2: Heavy Accumulation Areas and Priority Site Proximity

Locations	Number	Distance (ft)	High Priority Site	Medium Priority Site
Pilot Area 1 Location 1	17	2800	0	17
Pilot Area 1 Location 2	6	1600	0	6
Pilot Area 1 Location 3	21	2000	1	20
Pilot Area 1 Location 4	21	2400	1	20
Pilot Area 1 Location 5	21	2700	1	20
Pilot Area 2 Location 1	11	2800	2	9
Pilot Area 2 Location 2	2	1850	2	0

Methods Used

Two GIS analysis methods were used in developing strategies for this program during the pilot phase. They included proximity analysis and overlay analysis. Proximity analysis means the determination of which features are near to/far from others. As was indicated above we were able to determine that some medium priority industrial/commercial sites were located near the highest accumulation areas. In addition, we could see that there were some other high priority sites located near drainage inlets with a low ranging volume. By flagging these locations and using proximity analysis in forthcoming months as more data comes into the GIS database, it will help us see if this relationship turns into a high accumulation area that should be earmarked for further scrutiny. This basic analysis provides a very effective means of understanding these potential spatial relationships.

We used overlay analysis to drape one layer of GIS data over another and to distinguish relationships from a birds-eye point of view. A common example is relating points to polygonal features. In this context, it understood the relationship of heavy accumulation areas and land use designation. Based on the initial data collected and the determination of the heavy accumulation areas, we can now take it a step further by understanding what types of land uses occurred nearby.

As an example, if we were to see that there is a drainage inlet with a significant debris volume and it sits within a low density residential area, one could assume that the cause of the high debris volume may be from yard clippings and other similar vegetative matter. Another example is if a high-ranging volume drainage inlet is located in the middle of the urban center, we could extrapolate that the cause may be attributed to an area containing a school and fast food restaurants.

Table 3 outlines the most abundant land use in each area High Accumulation Area.

Table 3: Heavy Accumulation Areas and Predominant Land Use

Locations	Land Use
Pilot Area 1 Location 1	It is contiguous with General Commercial but Low and Low-Medium Residential predominates.
Pilot Area 1 Location 2	It is near a parcel of Institutional, but it is not a school, and Low and Low-Medium Residential predominates.
Pilot Area 1 Location 3	Low Density Residential
Pilot Area 1 Location 4	Low Density Residential
Pilot Area 1 Location 5	Industrial
Pilot Area 2 Location 1	Industrial and Professional & Administrative Office
Pilot Area 2 Location 2	Medium Density Residential

Final Remarks

GIS has been central to the efforts in establishing the drainage inlet inspection and cleaning program for the City of Santa Ana under the auspices of this grant from the SWRCB. This paper described the planning of the GIS process, assembling drainage facility and industrial/commercial site information, developing a drain inspection and debris measurement component, incorporating inspection and GPS location information into a GIS database, the analysis of drain cleaning and inspection data, and its use in the development of a targeted source reduction program.

GIS has been harnessed as the main tool in the planning and implementation phases of this program as the first year unfolded. Incorporating the functionality of GIS with some core data management strategies has provided the City with a very strong foundation to integrate the data and perform the analyses required. The use of GIS will continue as the program enters its second and third years. Map products for the program have been a substantial by-product of its use and the City has found them to be extremely great references.

Establishing this program required excellent communication and planning amongst the various people involved from both the City and CDM. By leveraging this technology, we were able to reinforce and strengthen this area tremendously by having up to date data, references, and processes at our fingertips. As we move into the next year of the program, we will build upon this structure.

Acknowledgements

The contributions of the project team comprised of CDM and City personnel made the project a profound success. These included Joe Parco, Souri Amarani, and Trevor Burgan from the City; Melinda McCoy, Luis Leon, Thomas Lo from CDM, and others who contributed in a collaborative manner to assure the best results for this project.