4 METHODOLOGY

As discussed in section 1.4, the methodology followed in this study is partitioned into eight major tasks: (1) Establishment of a digital database, (2) digital modeling of the watershed, (3) definition of a rainfall/streamflow relationship, (4) linking expected mean concentration of pollutants to land use, (5) calculation of pollutant loadings in the watershed, (6) predicting the aerial distribution of pollutant concentrations, (7) simulation of point sources, and (8) estimating EMC values.

Chapter 3 discussed the establishment and preparation of digital data sets for the nonpoint source pollution assessment. In the discussion of the remaining tasks, this chapter is similarly formatted to provide a descriptive narrative of the steps performed along with the actual Arc/Info and UNIX commands executed. This format provides the reader insight into the specific steps performed and describes the theoretical bases for each procedure. As in Chapter 3, automated Arc Macro Language (AML) scripts are referenced where appropriate.

4.1 Grid-Based Watershed Modeling Using Digital Elevation Data

The process of digitally simulating a watershed starts with the digital elevation model of the basin. The fine mesh of 1 hectare cells laid out over the basin is simply represented by a rectangular array, or grid. For the San Antonio-Nueces region, the total number of cells in this array is approximately 1.87 million. Processing of this digital basin relies heavily on the Arc/Info version 7.0 GRID module.

Establishing a Digital Stream Network

Before digitally simulated stream networks and subwatersheds can be created, the raw USGS digital elevation model accessed from the Internet must be corrected for data errors that exist in the original data file or are introduced as a result of reprojection to a different coordinate system. In particular, raw digital elevation models downloaded from Internet may contain many sinks. Sinks are single grid cells or groups of cells surrounded by cells of higher elevation. In order to create a

"hydrologic DEM" (Reed and Maidment, 1995), all of the sinks in the digital elevation model must be removed. This is accomplished through use of the Fill command. The Fill command redefines the elevations of each of the sink points to be equal to that of its lowest elevation neighbor. This smoothing process should always be used on a digital elevation model after reprojection because the data resampling that occurs during reprojection often creates artificial holes, or sinks, in the grid.

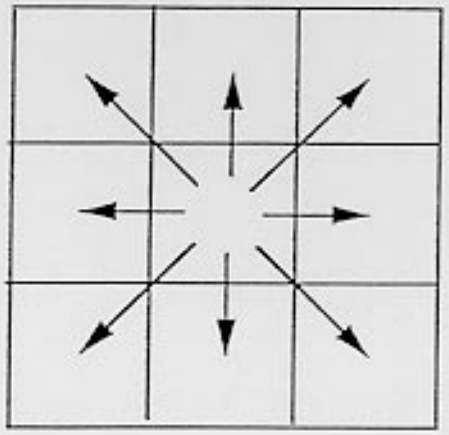
Grid: **fill sndemalb sanfil SINK**

Once the filled hydrologic digital elevation model has been created, it can be processed to determine the direction of the flow of water from cell to cell and to determine, for each cell in the grid, the number of cells that are upstream. The Flowdirection and Flowaccumulation commands are used for these purposes. The conceptual basis for this process relies on the 8-direction pour point model (Figure 4.1a). This model represents a cell surrounded by its eight neighbors. Drainage passes from each cell to only one of its neighbors in the direction of steepest descent, as defined by the filled digital elevation model (Figure 4.1b). By tracing these cell to cell drainage connections downstream, a flow direction network for a complete basin is established (Figure 4.1c). By counting the number of cells that occur upstream of each particular cell, a flow accumulation grid (Figure 4.1d) is established (Maidment, 1993).

Grid: **sanfdr = flowdirection(sanfil)**
Grid: **sanfac = flowaccumulation(sanfdr)**

A digital representation of the stream network in the basin is established by acknowledging that, just as surface runoff accumulates in creeks and streams, flow accumulation values along the digital streams should be greatest. The Conditional (Con) function is used to extract the flow accumulation cells that have value greater than a certain threshold (in this case, 1000). The resulting grid (str1) and equivalent coverage (covstr) actually reflect strings of cells whose flow accumulation values are greater than 1000.

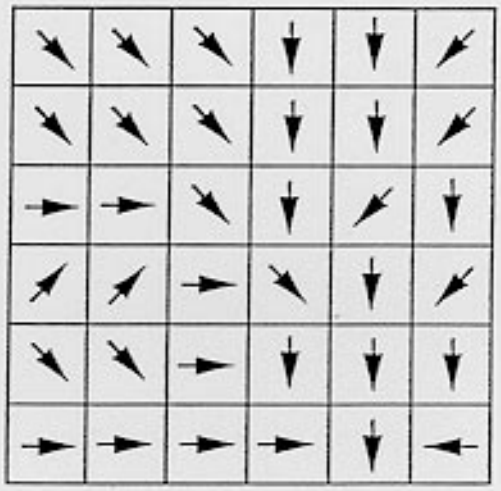
Grid: **str1 = con(sanfac > 1000,1)**
Grid: **covstr = gridline(str1)**



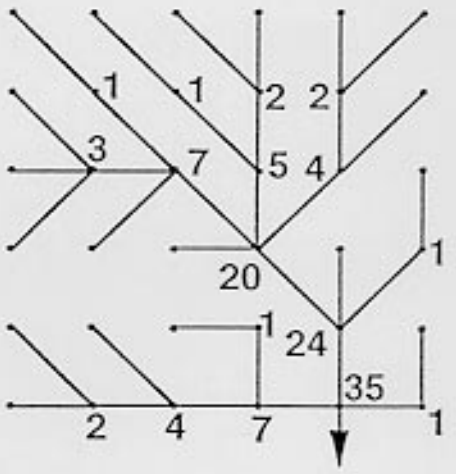
(a)

78	72	69	71	58	49
74	67	56	49	46	50
69	53	44	37	38	48
64	58	55	22	31	24
68	61	47	21	16	19
74	53	34	12	11	12

(b)



(c)



(d)

Figure 4.1 : Processing of Digital Elevation Data.
(a) the 8-direction pour point model; (b) a grid of elevation values; (c) flow direction grid; (d) flow accumulation grid. (Source: Maidment, 1993)

Figure 4.2 shows a comparison of the digitally delineated stream network (covstr) and the 1:100,000-scale hydrography digital line graph representation of the basin streams (Saunders and Maidment, 1995). As can be seen in the figure, the delineated streams in the inland portions of the basin match quite closely with the digital line graphs. However, closer to the coast, the differences between the Grid-delineated and digital line graph streams are much more apparent. This is expected, as slopes in this region of the San Antonio-Nueces coastal basin are generally flat. Elevations in this region do not change as significantly (or at all) from cell to cell and flow directions must be determined over larger areas of equal elevation.

Burning Digital Line Graph Streams into the Digital Elevation Model

The digital stream network established in the above procedure is derived using pure elevation data. However, the poor match that exists with the digital line graphs in the near-shore portions of the watershed is of concern. These digital line graphs are the result of manual digitizations of USGS 1:100,000-scale maps of the region and are considered to be fairly accurate. A review of the digital line graph coverage indicates many straight constructed channels in the region. Elevations of these channel beds may not be accounted for in the digital elevation model. In order to correct for this inconsistency, and to ensure that all digitally derived drainage paths adhere to the accepted stream networks reflected in the digital line graphs, a process of "burning" the digital line graphs into the digital elevation model is performed (Maidment and Saunders, 1996).

As can be seen from Figure 3.3, the hydrography digital line graphs of the San Antonio-Nueces coastal basin include lakes, in-stream lakes, coastlines, and "disappearing" streams in addition to the streams that flow to the bay network. The first step in preparing the digital line graph coverage for the "burn-in" process is to remove all of the features that do not contribute to providing contiguous drainage paths throughout the basin. The Arc/Info ArcEdit module is used for this purpose. In ArcEdit, each stand-alone lake and "disappearing" stream is removed. All in-stream lakes are replaced with arc segments that would otherwise bisect the lakes. Additionally, in the deltas of the Nueces and San Antonio Rivers, where the braiding effects of bifurcating and distributary streams occur, a main channel is identified

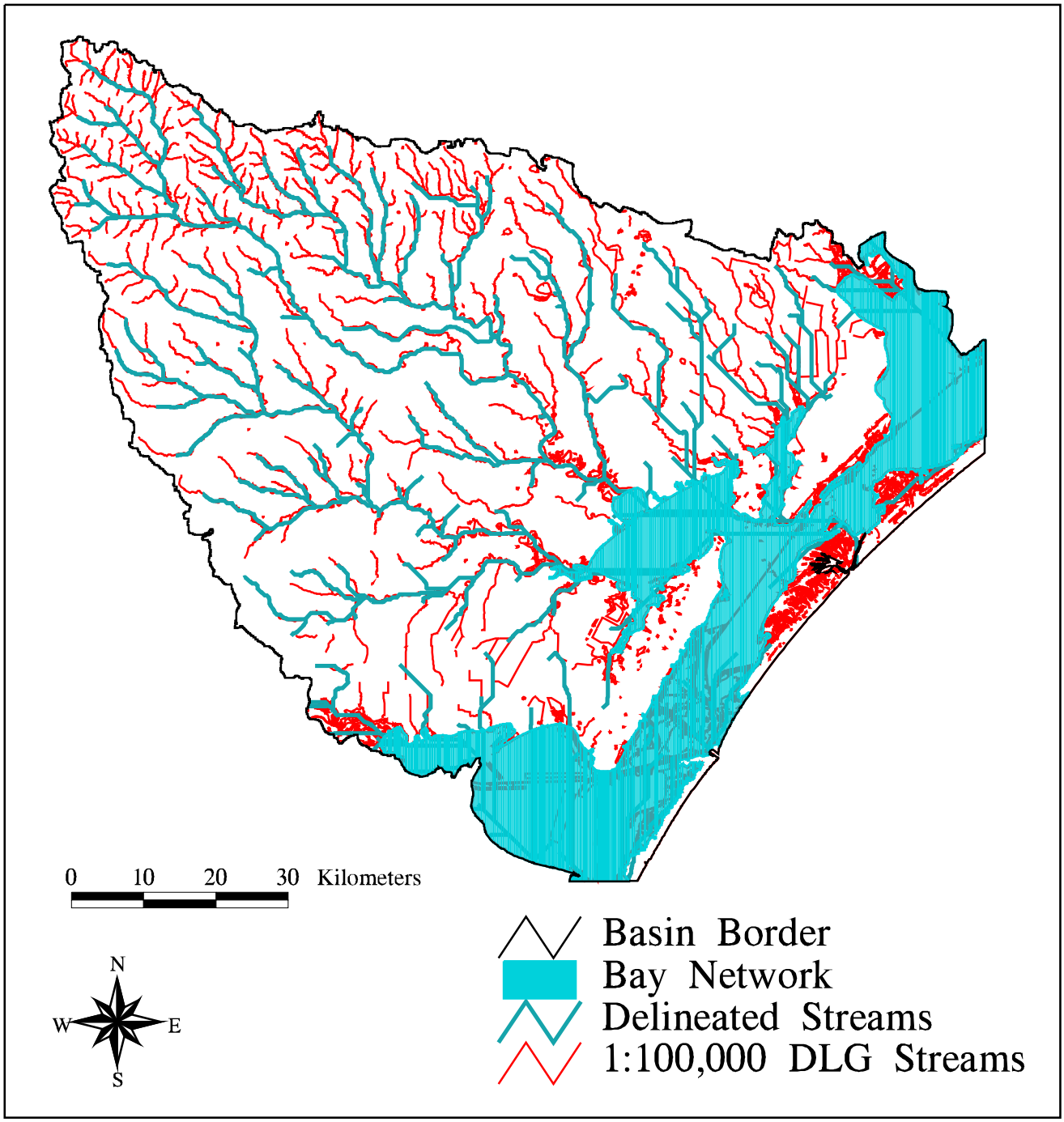


Figure 4.2 : Comparison of GRID-Delineated Streams with 1:100,000-Scale Hydrography Digital Line Graphs

through the delta and all other split channels and sinuous side channels are removed. This maintains one and only one drainage path for each upstream cell.

Other editing performed on the digital line graph coverage includes the removal of marsh channels throughout the barrier islands, removal of pipelines, shipping lanes, and islands within the Intracoastal Waterway, and the addition of arc segments to bound the Intracoastal Waterway between Corpus Christi Bay and San Antonio Bay. The final edited coverage, defined as sanrivs4, is shown in Figure 4.3.

Polygons are established from this line coverage by using the Arc/Info Clean command to create the sanpolys coverage. When all of the edits have been implemented correctly, the only polygons produced are those of the Intracoastal Waterway and the barrier islands. Unique polygon coverages of the Intracoastal Waterway and barrier islands are created by displaying sanpolys in ArcView 2.0, selecting the appropriate polygons, and converting them into shape files (bays.shp and barriers.shp). The Arc/Info Shapearc command is then used to build coverages from these shape files:

```
Arc:  clean sanrivs4 sanpolys
Arc:  shapearc bays bays
Arc:  build bays poly
Arc:  shapearc barriers barriers
Arc:  build barriers poly
```

The bays coverage is buffered by 100 meters (one cell width) to create an approximate bay network coverage that can be used to remove coastlines from the edited digital line graph coverage. First a rectangular coverage spanning the extent of the filled digital elevation model is created through use of the Con and Gridpoly commands. The buffered bay coverage is then combined with this rectangle through the Arc Union command. The resulting coverage is converted back into 100 m grid cell format, using Polygrid:

```
Arc:  buffer bays baybuff # # 100 # poly
Grid:  sqgrid = int(con(sanfil,1,1))
Grid:  sqcov = gridpoly(sqgrid)
Arc:  union sqcov baybuff baycov
Grid:  baygrid = polygrid(baycov,#,#,#,100)
```

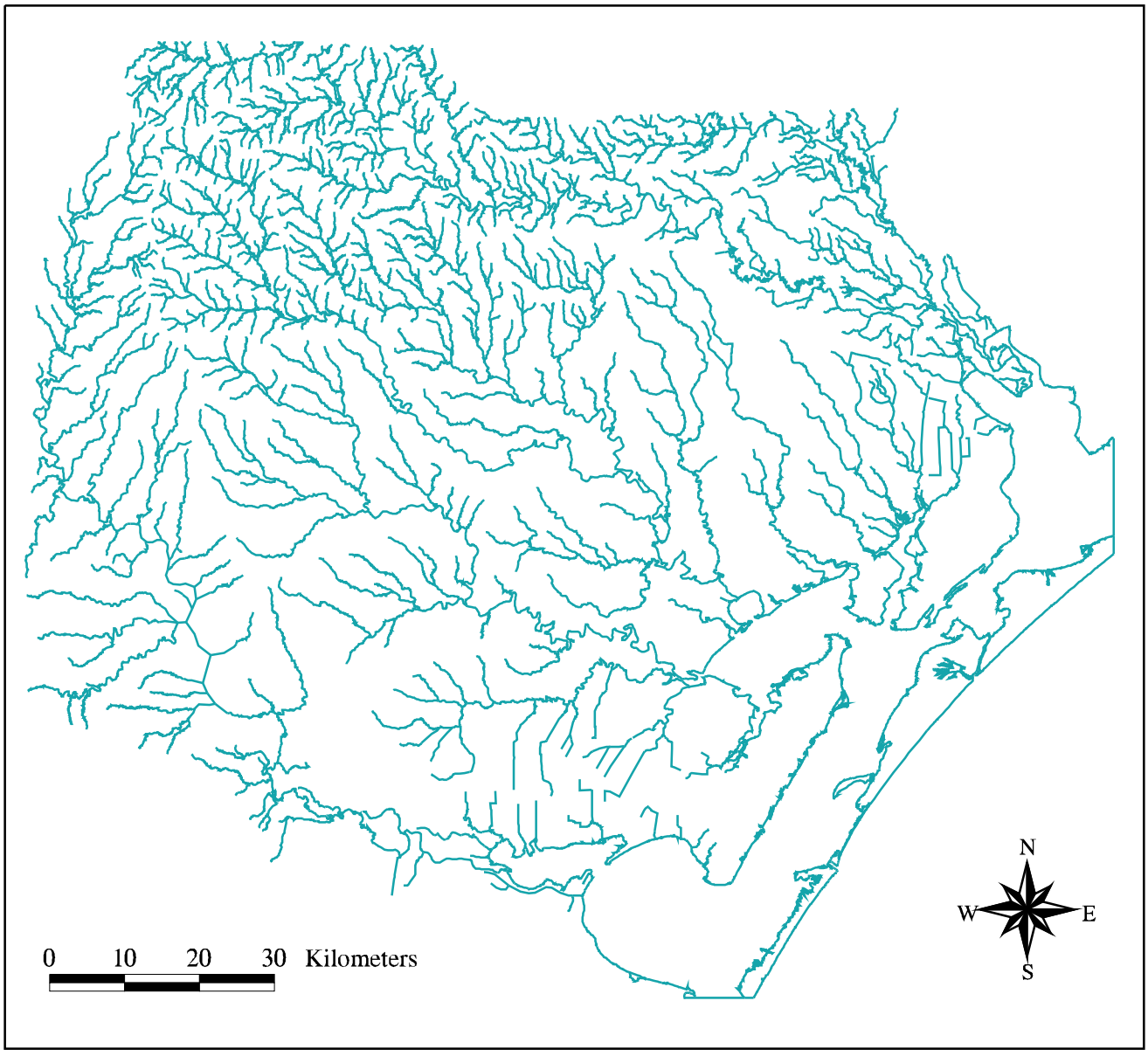


Figure 4.3 : Final Edited Digital Line Graph Stream Coverage for the San Antonio-Nueces Coastal Basin

The grid analysis window is then set to the size of the digital elevation model. An equivalent grid of the edited stream coverage is created, using the Linegrid command. The coastlines of the stream grid are removed with the Con statement, by selecting only the cells that correspond to the mainland portion of baygrid (i.e. baygrid cell value = 2). In effect, this step reduces all subsequent analyses to the mainland portion of the basin, as all other grid cells (bay network and barrier islands) are represented by NODATA, or null values.

Grid: **setwindow sanfil**
Grid: **strgrid = linegrid(sanrivs4,#,#,#,100,zero)**
Grid: **strmgrid = con(baygrid == 2,strgrid)**

Strmgrid is "burned" into the digital elevation model with the Con statement by artificially raising the elevation of all off-stream grid cells by five meters while holding the in-stream grid cells to a value of zero elevation. This creates a new digital elevation model with which to restart the digital stream delineation process.

Grid: **ditstrm = con(strmgrid > 0,0,sanfil + 5)**

After the new digital elevation model is filled, the bay network region is redefined with values of zero elevation in place of the NODATA values, using baygrid and the Con statement. This is required in order to avoid erroneous flow direction computations in the subsequent steps. A flow direction grid is established from the updated bayfil grid, and then NODATA values are reinserted into the bay network, so that subsequent analyses will be specific to the mainland region, only. This last step is accomplished by using baygrid and the Con statement to isolate the flow direction cells specific to the mainland:

Grid: **fill ditstrm ditfil SINK**
Grid: **bayfil = con(baygrid == 2,ditfil,0)**
Grid: **ditfdr = flowdirection(bayfil)**
Grid: **clipfdr = con(baygrid == 2,ditfdr)**

A flow accumulation grid is created and, as before, flow accumulation cells with a value greater than 1000 are extracted to define the locations of the digitally simulated streams:

Grid: **ditfac = flowaccumulation(clipfdr)**

Grid: **ditstr1 = con(ditfac > 1000,1)**
Grid: **covstr1 = gridline(ditstr1)**

Figure 4.4 shows the new digital streams, as burned into the digital elevation model and superimposed over the 1:100,000-Scale hydrography digital line graph files of the basin.

Digital Delineation of Subwatershed Drainage Areas from USGS Flow Gauges

In order to provide a more quantitative check on the accuracy of the digitally derived basin, drainage areas from the existing USGS flow gauges in the basin are determined from the flow accumulation grid, using an overlay of the sangages coverage created in section 3.2. These digitally delineated subwatershed drainage areas are then compared with values provided through the USGS-Texas Internet site identified in Table 3.2.

In order to digitally delineate drainage areas, outlet cells for each particular area must first be established. This is accomplished through the Arc/Info Grid module, by displaying the flow accumulation grid, overlaying the sangages coverage, and selecting each gauge location along a flow accumulation string. The fact that each of the stream gauges in the coverage fall exactly on the flow accumulation network is a testament to the accuracy of the "burn-in" process used above. The Selectpoint command allows the user to interactively define each outlet point. Once the outlet cell grid is defined, the Watershed function uses it, along with the flow direction grid, to define the area draining to the selected cell. An equivalent coverage of the drainage area is then created using the Gridpoly command. This process is performed for all five USGS gauges in the coastal basin. For example, the commands for delineating drainage area to the Aransas River gauge are:

Grid: **drainpt1 = selectpoint(ditfac,*)**
Grid: **aranarea = watershed(clipfdr,drainpt1)**
Grid: **arancov = gridpoly(aranarea)**
Grid: **list aranarea.vat**

Record	Value	Count
1	56	63291

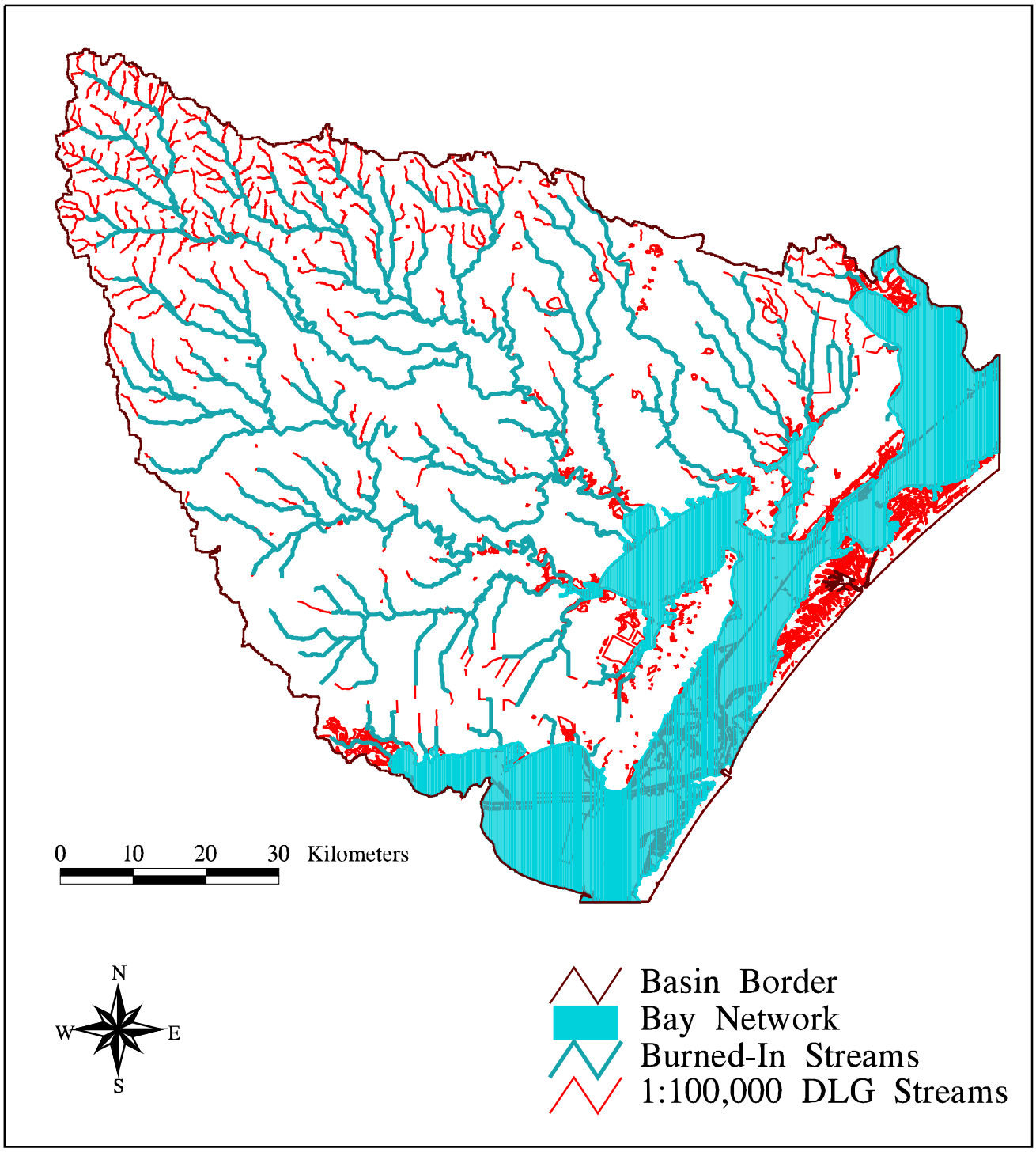


Figure 4.4 : Comparison of Burned-In Streams with 1:100,000-Scale Hydrography Digital Line Graphs

By displaying the value attribute tables (vat's) for each of the five drainage area grids, a count of the number of cells simulating each drainage area is obtained. Since it is known that each cell has area of 1 hectare = 10,000 m^2, the area in square kilometers is established by dividing the number of cells by 100. These areas, converted to square miles, are then compared with the USGS drainage areas obtained from the Internet site. Table 4.1 shows the comparison of the digitally delineated drainage areas with USGS drainage areas and Figure 4.5 shows the digital drainage areas as they exist within the basin.

Percent errors from Table 4.1 indicate that the digitally delineated drainage areas match the USGS areas fairly accurately. The largest errors, 8.89% for the Copano Creek drainage and 2.85% for the Chiltipin Creek drainage, occur in the flattest portions of the basin, which are also closest to the coast. The smallest error, 0.32% for the Medio Creek drainage, occurs for the furthest inland area.

Figure 4.6 shows a close-up of the Copano Creek drainage area and one potential contributing factor to the errors occurring in the digital delineation. The sinuous nature of the digital subwatershed boundary results when using the "burn-in" process for establishing the digital elevation model. While the actual cause of this anomaly is unknown, it is suspected that the flow direction grid is affected by the sharp drops in elevation to the burned-in streams. Even with these boundary anomalies, the percent errors for the delineated drainage errors are considered to be acceptable.

USGS GAGE #	STREAM	# CELLS	DELINEATED DRAINAGE AREA (km^2)	DELINEATED DRAINAGE AREA (mi^2)	ACTUAL USGS DRAINAGE (mi^2)	% ERROR
08189200	COPANO	20,782	207.82	80.2	88	8.89
08189300	MEDIO	52,708	527.08	203.3	204	0.32
08189500	MISSION	176,619	1766.19	681.4	690	1.25
08189700	ARANSAS	63,291	632.91	244.2	247	1.15
08189800	CHILTIPIN	32,233	322.33	124.4	128	2.85

Table 4.1 : Comparison of Digitally Delineated and USGS Drainage Areas

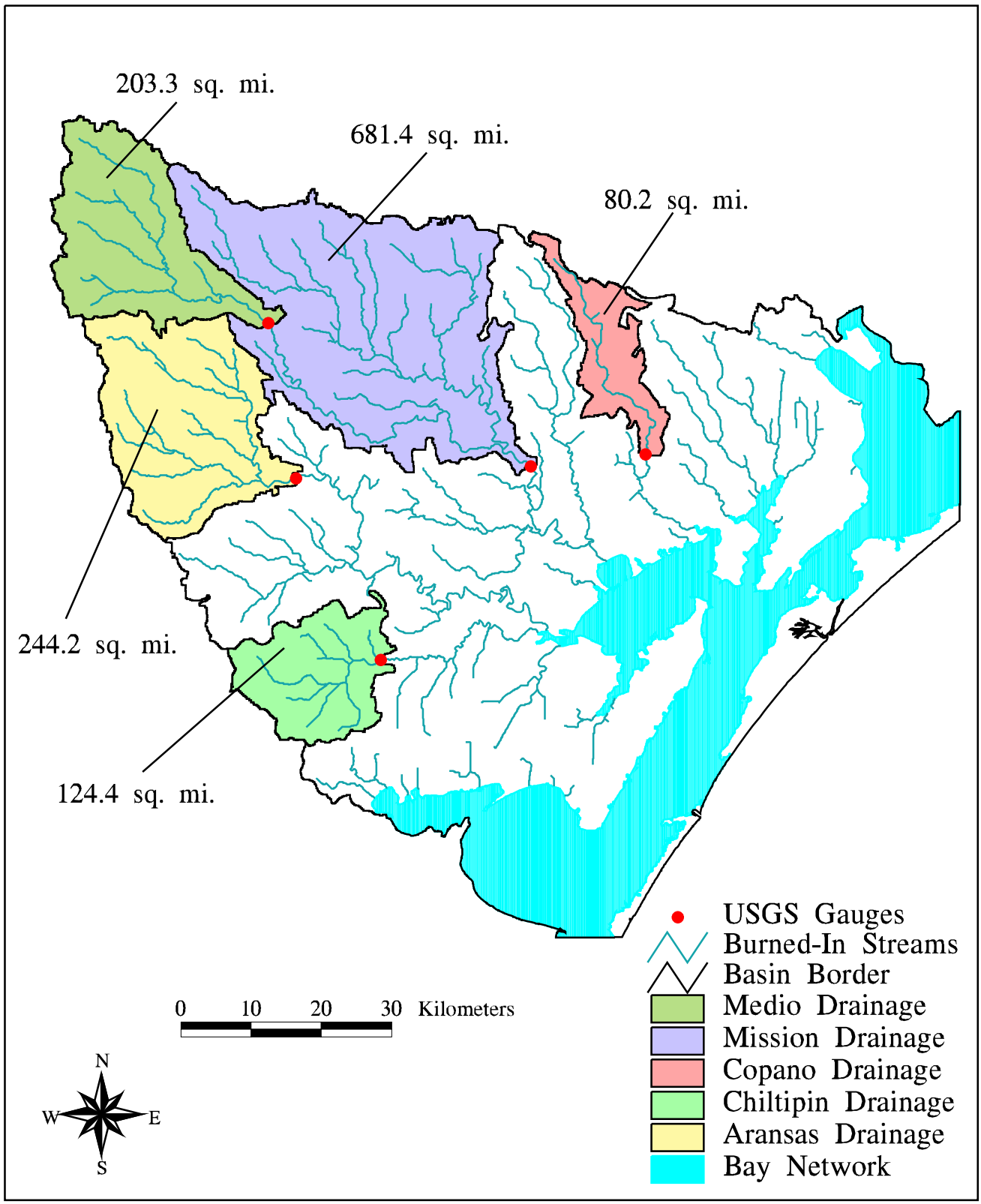


Figure 4.5 : Digital Drainage Areas Delineated from USGS Streamflow Gauges

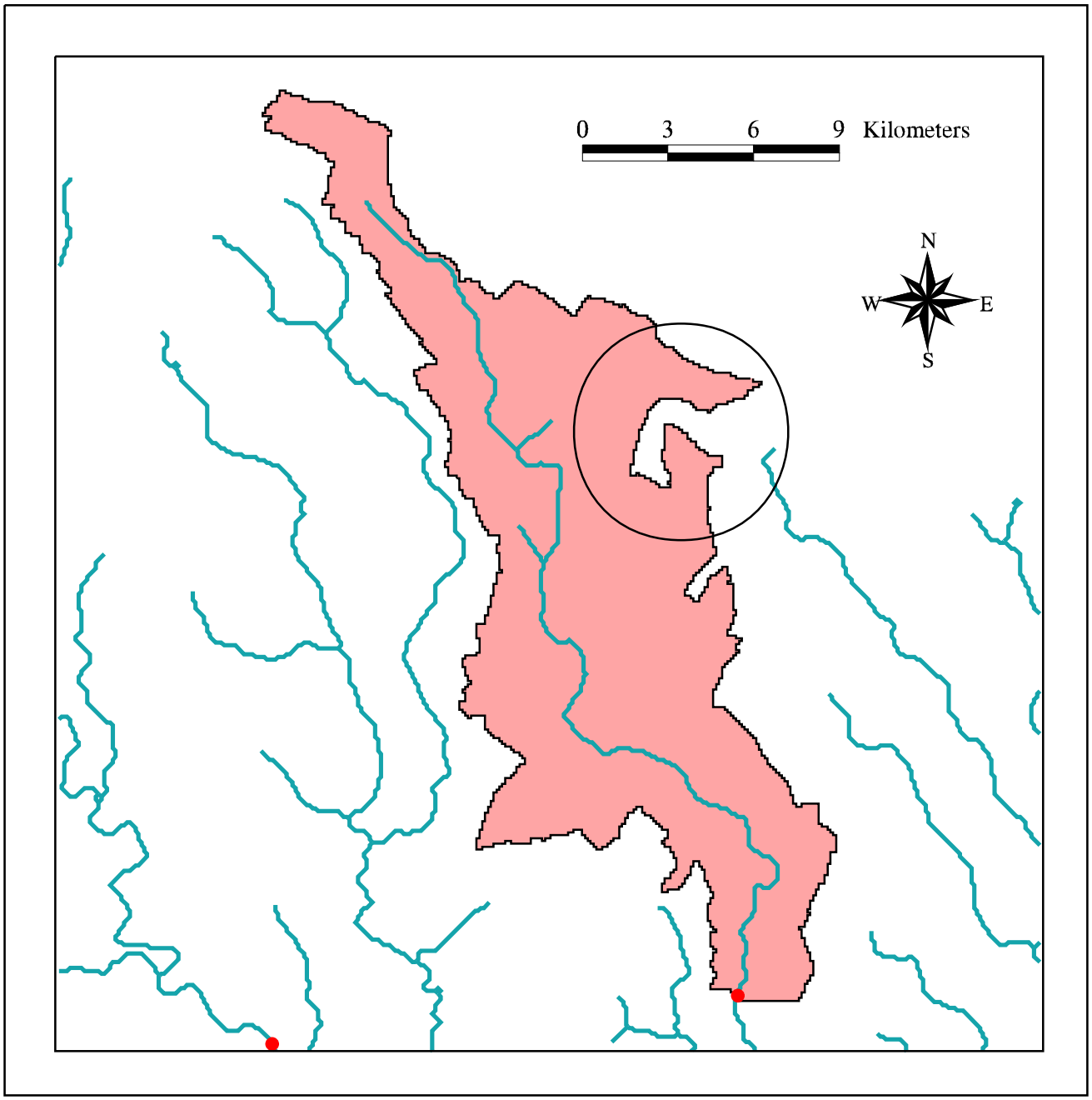


Figure 4.6 : Drainage Area Delineated from the Copano Creek USGS Stream Gauge

Defining the Coastal Basin Boundary

For many of the figures in section 3.2, a coverage of the San Antonio-Nueces coastal basin boundary is used to clip out the particular features of the display. This boundary is created to facilitate watershed-level analyses of the respective spatial parameters. Both Arc/Info version 7.0 and ArcView 2.0 are used in the establishment of this border.

The Arc/Info Grid module is first employed to delineate subwatersheds within the complete basin. A threshold value (i.e. number of cells) defining the size of subwatersheds to be delineated is specified. This threshold value should be chosen to ensure that the total number of subwatersheds delineated is manageable. The total area of the San Antonio-Nueces coastal basin is known to be approximately 7000 km^2. In order to keep the number of subwatersheds in the basin under 100, a threshold value of 8000 cells (i.e. 80 km^2) is chosen. The Con statement is used to identify all flow accumulation cells in the basin with value greater than the threshold. As discussed previously, this results in strings of grid cells that represent a stream grid of the basin.

The Streamlink command is used to identify specific stream reaches, based on the stream grid and flow direction grids. The Zonalmax command then produces a grid of accumulation zones, using the grid of stream reaches along with the flow accumulation grid. This command stores the maximum value of each of the stream reaches into all cells of the corresponding accumulation zones.

Next, using the Con statement, the outlet cells of each accumulation zone are defined as those cells with identical flow accumulation and accumulation zone grid values. The Watershed function is then used, as before, to delineate the drainage areas to each zonal outlet cell. Finally, an equivalent coverage of the delineated subwatersheds is created through the Gridpoly command:

Grid: **ditstr8 = con(ditfac > 8000,1)**
Grid: **ditlnk8 = streamlink(ditstr8,clipfdr)**
Grid: **ditacc8 = zonalmax(ditlnk8,ditfac)**
Grid: **ditout8 = con(ditacc8 == ditfac, ditlnk8)**
Grid: **ditshd8 = watershed(clipfdr,ditout8)**
Grid: **shed8cov = gridpoly(ditshd8)**

Figure 4.7 shows the digitally delineated subwatersheds of the San Antonio-Nueces coastal basin overlaid with the USGS Hydrologic Unit Codes to provide an estimate of which subwatersheds fall within the basin and which are associated with the Nueces and San Antonio River basins.

Using ArcView 2.0, the subwatersheds coverage (shed8cov) is displayed and each of the polygons that fall within the San Antonio-Nueces basin are selected. Once selected, these polygons are converted into the shapefile, subsheds.shp. As can be seen from Figure 4.7, the complete San Antonio-Nueces basin is not accounted for by the polygons of shed8cov. This occurs because the San Antonio-Nueces basin is a coastal basin and not a river basin. River basins have a single outlet point, but coastal basins drain to the ocean in a more diffuse manner. Since many of the actual drainage areas along the coast are smaller than 80 km^2, they are not included in the subwatersheds coverage.

This problem is resolved by selecting shed8cov polygons that, along with the baybuff coverage and the subsheds shapefile, completely enclose the basin area not accounted for in shed8cov. Only three additional polygons are selected for this purpose and converted into the shapefile, trimshed.shp. Figure 4.8 shows the shapefiles subsheds.shp and trimshed.shp displayed with the baybuff coverage to completely enclose the undelineated area of the coastal basin.

The subsheds and trimshed shapefiles are converted to coverages using the Arc/Info Shapearc command. The coverages are then cleaned to construct polygon topology. This process creates the coverages covsheds and covtrim. The Append command is used to merge the covsheds, covtrim, baybuff, and barriers coverages into one large coverage blanketing the entire coastal basin.

Arc: **shapearc subsheds subsheds**
Arc: **shapearc trimshed trimshed**
Arc: **clean subsheds covsheds**
Arc: **clean trimshed covtrim**
Arc: **append basin**
Enter the 1st coverage: **covsheds**
Enter the 2nd coverage: **covtrim**
Enter the 3rd coverage: **baybuff**
Enter the 4th coverage: **barriers**
Enter the 5th coverage: **~** <return>
Done entering coverage names (Y/N)? **y**
Do you wish to use the above coverages (Y/N)? **y**

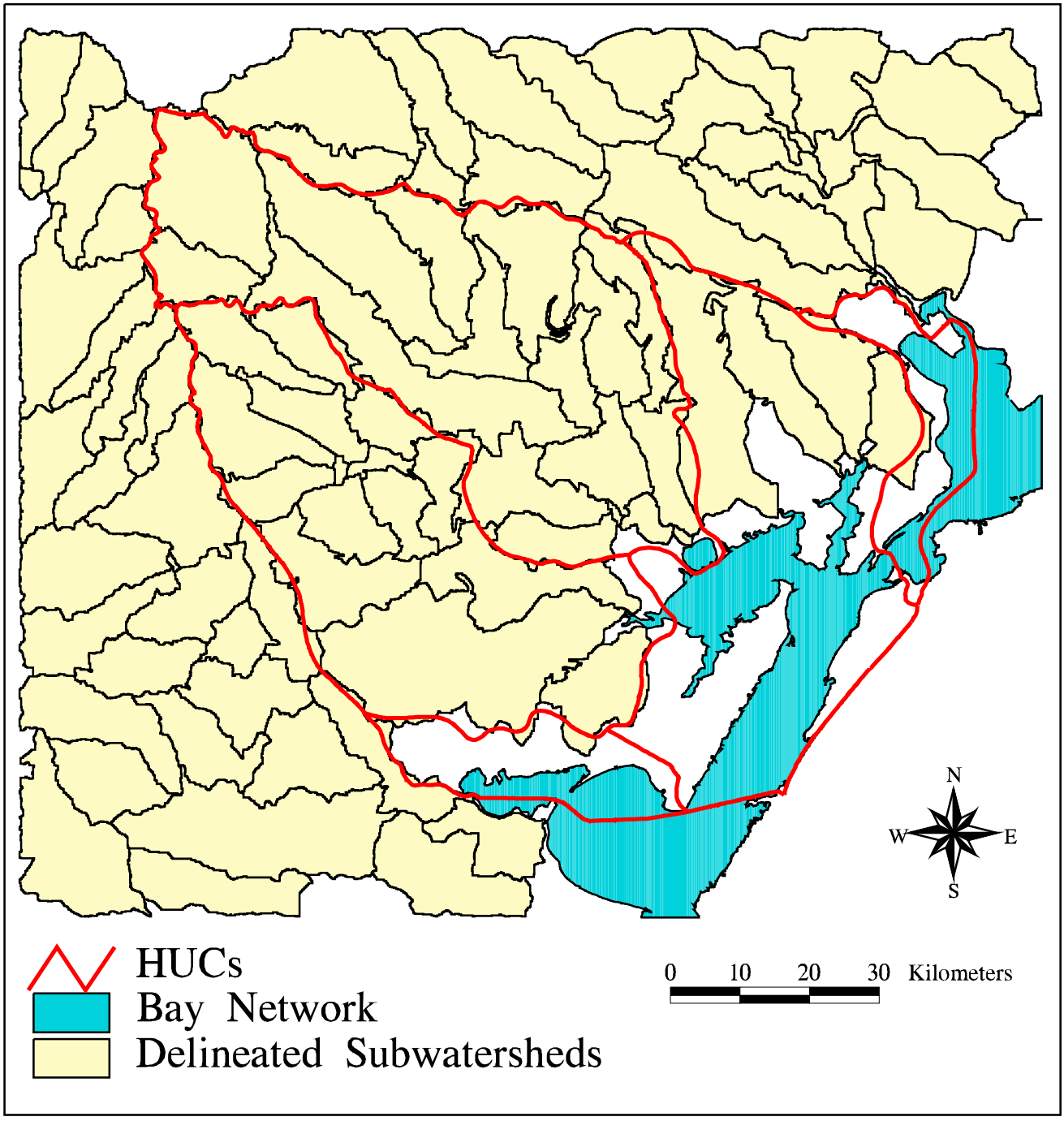


Figure 4.7: Digitally Delineated Subwatersheds of the San Antonio-Nueces Basin Compared with USGS Hydrologic Unit Codes

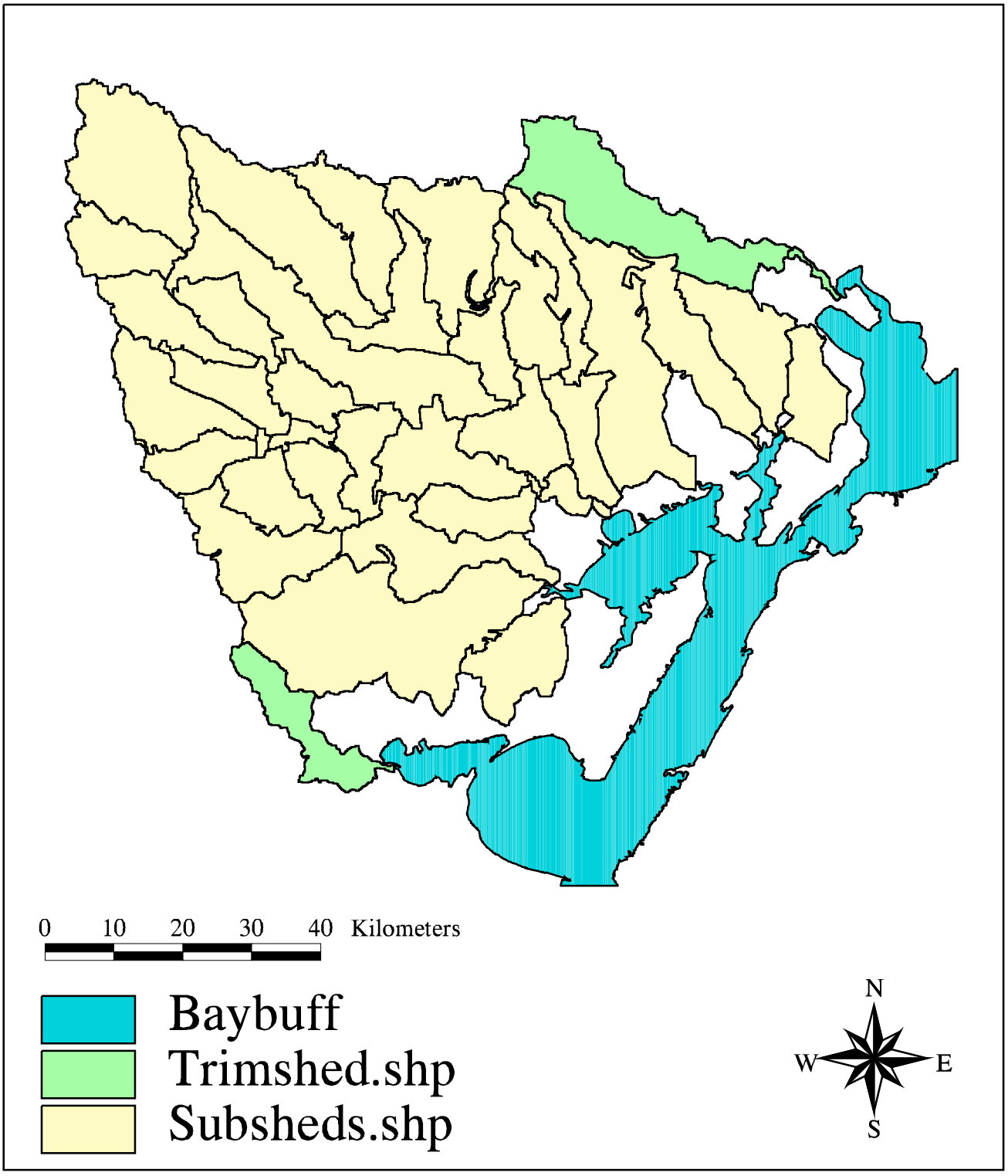


Figure 4.8 : Creating the Basin Boundary through use of ArcView Shapefiles

```
  Appending coverages....
Arc:  clean basin sanbasin
```

The final cleaned basin coverage, sanbasin, actually contains the three polygons from the trimshed shapefile. By displaying the sanbasin coverage in ArcView 2.0, all sanbasin polygons except for those from trimshed are selected and converted to the shapefile, bord.shp. Once again, the Shapearc and Clean commands are used to create a border coverage. Finally, the Reselect command is used to select the exterior polygon of the coverage. This has the effect of removing all of the interior subwatershed boundaries and leaving only the outline of the basin.

```
Arc:  shapearc bord bord
Arc:  clean bord border
Arc:  reselect bord sanbord
>:  res bord# = 1
>:  ~
Do you wish to re-enter expression (Y/N)?  n
Do you wish to enter another expression (Y/N)?  n
  1 features out of 60 selected
```

The final sanbord coverage is used throughout this project to define the boundary of the San Antonio-Nueces Coastal Basin. For aesthetics, the complete bodies of both Corpus Christi Bay and San Antonio Bay are included in the coverage. The Clip command is used, along with this basin border, to select data specific to the basin from the data sets described in section 3.2.

```
Arc:  clip sanhydro sanbord sanhyd line
Arc:  clip sanlus sanbord sanlu poly
Arc:  clip rainbfcv sanbord snrainyr poly
```

4.2 Determination of a Rainfall/Runoff Relationship

In order to assess the transport of pollutant loads in a region, an understanding of the means by which the loads migrate is first required. Nonpoint source pollutants are carried over land and into the stream networks of a region by direct runoff. This runoff is largely the result of precipitation over the area, although some runoff may also be generated by over-irrigation in agricultural areas. For this study, the volume of

runoff from a grid-cell is completely attributed to precipitation over the cell. By comparing average annual stream flows at each of the USGS flow gauges with the average annual precipitation that occurs upstream of those gauges, a mathematical relationship between rainfall and runoff is established.

Determining Average Rainfall for each Delineated Drainage Area

The Parameter-elevation Regressions on Independent Slopes Model (PRISM) discussed in section 3.2 provides the precipitation data used for this study. This data is provided as total annual depth of precipitation (mm) averaged over the 30-year period from 1961 to 1990. Two methods of determining average rainfall for each drainage area are performed and compared in this analysis.

The first method for calculating average rainfall for each drainage area makes use of a process called a weighted flow accumulation. This is an extension of the regular Flowaccumulation command. However, instead of counting the number of cells that occur upstream of each particular grid cell, the weighted Flowaccumulation command uses a second grid, called a weight grid, and sums the weight grid values of the cells that occur upstream. Using the buffered precipitation grid as the weight grid, a grid representing total annual potential runoff is generated:

Grid: **weighfac = flowaccumulation(clipfdr,rainbuff) * 10**

The factor of ten is used in this command to convert from the rainbuff units of depth (mm) to units of volume (m^3), using the knowledge that each cell is equal to 10,000 m^2, or

Volume = Depth (mm) * Area (#cells) * 10,000 m^2/cell * .001 m/mm. (4-1)

Once the weighted flow accumulation grid is established and displayed, the USGS stream gauge coverage is overlaid and each of the gauge points are queried, using the Cellvalue command, to determine the potential runoff that would occur at each gauge. By dividing these potential runoff values by the delineated drainage areas associated with each gauge (from Table 4.1), the average depth of precipitation is established for each drainage area:

Grid: **gridpaint weighfac value linear nowrap gray**
Grid: **points sangages**
Grid: **cellvalue weighfac ***
<9 to END>
The cell containing point (1233178.620,682331.934) has value 510618944.000
The cell containing point (1266688.298,684048.117) has value 1487741184.000
The cell containing point (1282992.941,685733.054) has value 192068960.000
The cell containing point (1229206.739,704427.678) has value 412713952.000
The cell containing point (1245272.269,656404.121) has value 273848544.000

A second method of determining average precipitation at each gauge is to create separate precipitation grids corresponding to each subwatershed grid, using the Con statement. Once the localized precipitation grids are created, the Describe command provides the mean value of all cells in the grid as a statistic. Using the Aransas drainage area as an example, this process is performed as:

Grid: **aranrain = con(aranarea,rainbuff)**
Grid: **describe aranrain**

Description of Grid ARANRAIN

Cell Size =	100.000	Data Type:	Integer
Number of Rows =	1325	Number of Values =	37
Number of Columns =	1520	Attribute Data (bytes) =	8

BOUNDARY		STATISTICS	
Xmin =	1180828.125	Minimum Value =	761.000
Xmax =	1332828.125	Maximum Value =	860.000
Ymin =	612183.250	Mean =	806.792
Ymax =	744683.250	Standard Deviation =	15.708

Table 4.2 shows the average annual precipitation values determined by both methods for each gauge. As can be seen from the table, results are consistent for both methods.

A precipitation grid that adheres to the watershed boundary is established by first creating an equivalent grid from the sanbord coverage established in section 4.1. Then, using that grid with the Con statement, the precipitation cells particular to the basin are selected.

Grid: **bordgrid = polygrid(sanbord,#,#,#,100)**
Grid: **sanpyr = con(bordgrid,rainbuff)**

	Method #1			Method #2
Drainage Subwatershed	Potential Runoff (m^3)	Drainage Area (km^2)	Precip Depth (mm)	Precip Depth (mm)
Mission	1,487,741,184	1766.19	842.34	842.326
Aransas	510,618,944	632.91	806.78	806.792
Copano	192,068,960	207.82	924.21	924.252
Chiltipin	273,848,544	322.33	849.59	849.618
Medio	412,713,952	527.08	783.02	783.033

Table 4.2 : Comparison of Methods for Determining Average Annual Precipitation for each Gauged San Antonio-Nueces Drainage Area

Determining Average Depth of Runoff at each USGS Gauge

The montflow.f FORTRAN algorithm (Appendix B) calculates values for total monthly, annual, and average annual streamflow volume, given average daily streamflow in cubic feet per second (cfs). Table 4.3 shows the output from this algorithm for each USGS streamflow gauge in the San Antonio-Nueces Coastal Basin, given the raw input data for the years 1961-1990. Table 4.4 shows the equivalent depths of streamflow for those volumes, calculated by dividing each value by the delineated drainage area of the particular gauge (from Table 4.1). Figure 4.9 shows how annual depths of streamflow have varied from the average annual depths at each gauge for the period 1961-1990.

One may note from Tables 4.3 and 4.4 that, of the five USGS gauges in the basin, only the Mission River gauge has recorded streamflow values for the total period of applicable precipitation data. Ideally, for the establishment of a rainfall/runoff relationship, rainfall and streamflow data from the same periods of record should be used. To that end, projected 30-year average annual streamflows at each gauge, Q_g, are estimated using the average annual 1961-1990 streamflow at the Mission gauge, Q_m. These estimates are established by multiplying Q_m by the ratio of q_g / q_m, where q_g is the average annual streamflow at the gauge and q_m is the average annual streamflow at the Mission gauge over the same time period, or

Year	Mission	Aransas	Chiltipin	Copano	Medio
1961	57,685,664				
1962	40,983,796				
1963	5,693,702				3,769,459
1964	10,694,530				3,144,356
1965	47,063,808	12,482,179			11,207,823
1966	106,309,680	23,827,040			1,432,233
1967	632,705,728	184,715,696			163,328,112
1968	131,968,248	18,562,584			12,908,496
1969	74,330,552	14,724,674			2,892,822
1970	65,834,276	14,914,258			7,378,471
1971	379,032,896	115,493,312	117,657,808	97,337,648	11,217,219
1972	177,693,296	34,983,532	36,046,596	58,093,640	6,511,890
1973	356,130,304	70,796,616	82,647,592	76,333,720	10,388,754
1974	106,735,128	52,987,968	12,367,189	21,977,854	745,549
1975	35,551,872	4,430,039	11,762,097	1,716,429	557,798
1976	253,111,616	30,784,200	59,696,076	42,789,296	18,338,360
1977	117,446,048	16,581,756	26,458,148	14,502,448	
1978	61,703,216	6,657,413	15,928,468	57,803,472	
1979	123,047,520	16,923,788	55,162,504	47,387,740	
1980	114,900,872	21,109,020	57,560,848	10,808,809	
1981	347,880,480	55,757,024	43,350,032	134,456,512	
1982	113,334,800	11,405,166	25,378,954	21,914,878	
1983	164,663,248	26,732,898	46,031,200	84,999,136	
1984	26,053,482	7,954,423	41,102,256	7,781,302	
1985	70,610,344	19,403,550	51,825,828	14,094,454	
1986	39,910,080	3,505,644	775,226	11,878,824	
1987	90,450,640	26,621,798		14,231,760	
1988	8,253,274	9,077,310	3,634,653	0	
1989	1,103,216	2,086,059	419,566	467,225	
1990	179,311,024	50,048,796	1,853,683	32,815,878	
Avg Annual =	131,339,778	32,791,029	42,734,426*	37,569,551	18,130,096

*calculated for 1971-1986 due to break in service in 1987

Table 4.3 : Annual Volume (m^3) of Recorded Streamflow (1961-1990) for the Five USGS Gauges in the San Antonio-Nueces Coastal Basin

Year	Mission	Aransas	Chiltipin	Copano	Medio
1961	32.7				
1962	23.2				
1963	3.2				7.2
1964	6.1				6.0
1965	26.6	19.7			21.3
1966	60.2	37.6			2.7
1967	358.2	291.9			309.9
1968	74.7	29.3			24.5
1969	42.1	23.3			5.5
1970	37.3	23.6			14.0
1971	214.6	182.5	365.0	468.4	21.3
1972	100.6	55.3	111.8	279.5	12.4
1973	201.6	111.9	256.4	367.3	19.7
1974	60.4	83.7	38.4	105.8	1.4
1975	20.1	7.0	36.5	8.3	1.1
1976	143.3	48.6	185.2	205.9	34.8
1977	66.5	26.2	82.1	69.8	
1978	34.9	10.5	49.4	278.1	
1979	69.7	26.7	171.1	228.0	
1980	65.1	33.4	178.6	52.0	
1981	197.0	88.1	134.5	647.0	
1982	64.2	18.0	78.7	105.5	
1983	93.2	42.2	142.8	409.0	
1984	14.8	12.6	127.5	37.4	
1985	40.0	30.7	160.8	67.8	
1986	22.6	5.5	2.4	57.2	
1987	51.2	42.1	-------	68.5	
1988	4.7	14.3	11.3	0.0	
1989	0.6	3.3	1.3	2.2	
1990	101.5	79.1	5.8	157.9	
Avg Annual =	74.4	51.8	132.6*	180.8	34.4

*calculated for 1971-1986 due to break in service in 1987

Table 4.4 : Equivalent Depth (mm) of Recorded Streamflow (1961-1990) for the Five USGS Gauges in the San Antonio-Nueces Coastal Basin

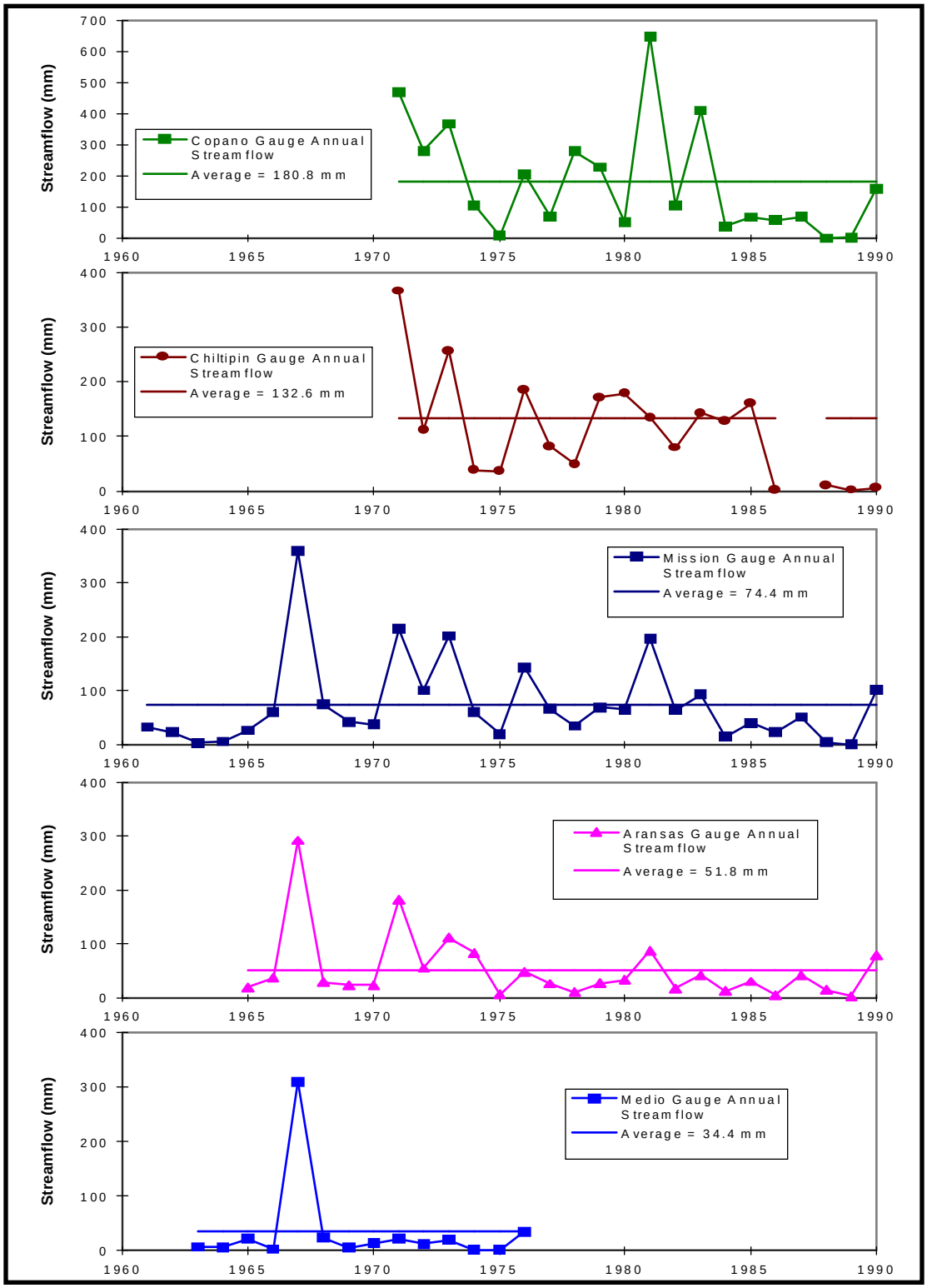


Figure 4.9 : USGS Recorded Annual Streamflows for Five Gauges in the San Antonio-Nueces Basin

$$Q_g = Q_m * (q_g/q_m). \quad (4\text{-}2)$$

This approach is legitimate for temporally averaged estimates in a region, where variations from year to year generally conform to similar trends. Figure 4.9 illustrates these regional trends with coincident occurrences of local maximum and minimum streamflow values. Table 4.5 shows the projected 30-year average annual depths of streamflow for each of the five USGS gauges.

Establishing a Mathematical Relationship Between Rainfall and Runoff

Using the five values for average annual precipitation along with the five values for projected 30-year average annual depth of streamflow, the Microsoft Excel 5.0 Regression tool is employed to determine the best fit curve between the two data sets. Assessments of the best linear, best quadratic, and best exponential fits show that the linear relationship most accurately reflects runoff in the San Antonio-Nueces coastal basin. Figure 4.10 shows the Microsoft Excel output of the regression for the linear case. This regression run produces a squared multiple correlation coefficient (r^2) value of 0.964, which indicates that the best fit line approximates the actual data well.

Based on the regression output, the linear relationship that best approximates the rainfall/runoff relationship in the San Antonio-Nueces Coastal Basin is

$$Q\ (mm) = 1.0527 * P\ (mm) - 799.37, \quad (4\text{-}3)$$

where Q represents depth of streamflow and P represents precipitation.

In order to create an Arc/Info grid of runoff, this relationship would be applied to every cell in the precipitation grid. However, since the precipitation grid has an effective range of values between 739 mm and 985 mm, it is noted that there is a small range of cells (739 - 759 mm) for which the relationship produces negative numbers. In order to avert this irregularity, the rainfall/runoff relationship of equation 4-3 is only applied to precipitation cells with value greater than 759 mm. In other words, the adjusted rainfall/runoff relationship becomes

$$\begin{aligned} Q\ (mm) &= 1.0527 * P\ (mm) - 799.37, \quad P > 759\ mm \\ Q\ (mm) &= 0, \quad P \leq 759\ mm. \end{aligned} \quad (4\text{-}4)$$

USGS Streamflow Gauge	Average Depth of Streamflow (mm)	Years of Continuous Operation	Avg Mission Depth for those Years (mm)	Projected 30-Year Avg (61-90) (mm)
Mission	74.4	1961-90	74.4	74.4
Aransas	51.8	1965-90	83.3	46.3
Copano	180.8	1971-90	78.3	171.6
Chiltipin	132.6	1971-86	88.1	112.0
Medio	34.4	1963-76	96.7	26.5

Table 4.5 : Projected 30-Year Average Annual Depth of Streamflow for the Five USGS Gauges in the San Antonio-Nueces Basin

SUMMARY OUTPUT

Regression Statistics	
Multiple R	0.9818
R Square	0.9640
Adjusted R Square	0.9519
Standard Error	12.6196
Observations	5.0000

ANOVA

	df	*SS*	*MS*	*F*	*Significance F*
Regression	1	12779.4255	12779.4255	80.2450	0.0029
Residual	3	477.7651	159.2550		
Total	4	13257.1906			

	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*
Intercept	-799.3698	99.0143	-8.0733	0.0040	-1114.4778	-484.2617
X Variable 1	1.0527	0.1175	8.9580	0.0029	0.6787	1.4267

Figure 4.10 : Regression Tool Output for Best Linear Fit Relationship Between Average Annual Precipitation and Depth of Streamflow

The fact that this equation produces values of Q = 0 for precipitation values less than 759 mm is a limitation of the linear modeling function. However, since the region of the San Antonio-Nueces basin that annually receives less than 759 mm of rain is limited to a 78 square kilometer area in the northwest corner of the watershed (approximately one percent of the basin's area), the adjusted linear rainfall/runoff relationship is considered acceptable for the basin. However, it should be stressed that the equation is specific to the San Antonio-Nueces coastal basin and should not be applied outside the watershed. A plot of this adjusted rainfall/runoff relationship is shown in Figure 4.11. The five points denoted on the graph represent the average precipitation and 30-year projected depth of streamflow for each gauge.

While equation 4-4 provides reasonable estimates of runoff for portions of the San Antonio-Nueces basin that drain to gauged locations, a more comprehensive relationship for the basin might be established by considering runoff data from gauges in adjacent basins which receive greater and less precipitation. Consideration of this additional runoff data would extend the range of application of the rainfall/runoff relation and a mathematical form of the relationship could be estimated more accurately.

Using the rainfall/runoff relationship of equation 4-4 in conjunction with the precipitation grid and the Con statement, a grid of runoff is produced. So that subsequent flow accumulations may be performed on this grid without encountering cells of NODATA (null) value, the Isnull command is used with a second Con statement to zero fill all of the null cells resulting from application of the rainfall/runoff relationship. Finally, an equivalent coverage of runoff is created through use of the Gridpoly command. Figure 4.12 shows this runoff coverage, with annual runoff amounts depicted in intervals of 50 mm.

Grid: **runoffeq = con(sanpyr > 759, 1.0527 * sanpyr - 799.37, 0)**
Grid: **runoff = con(isnull(runoffeq),0,runoffeq)**
Grid: **runoffcv = gridpoly(int(runoff))**

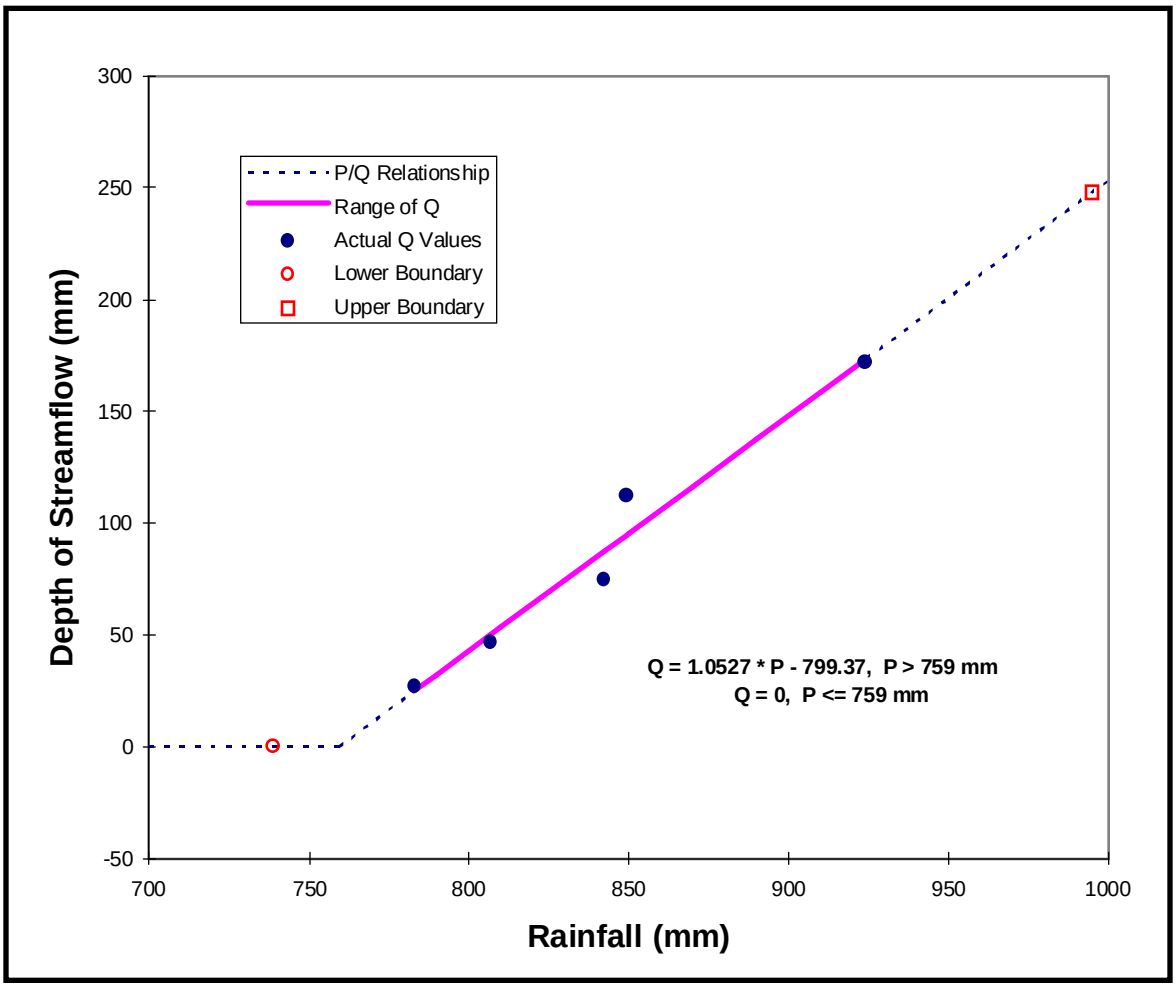


Figure 4.11 : Realtionship Between Rainfall and Streamflow in the San Antonio-Nueces Coastal Basin (Linear)

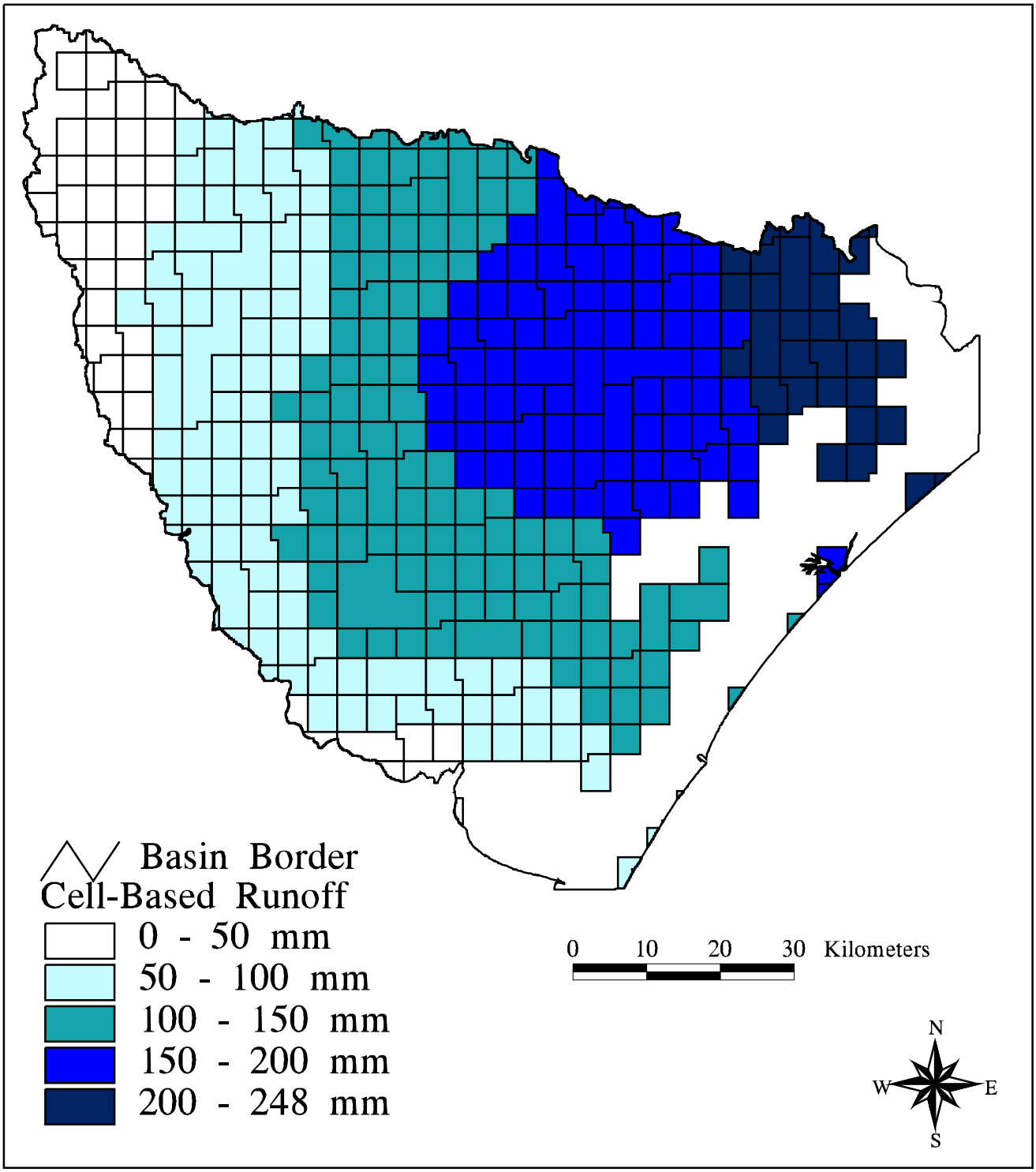


Figure 4.12 : Estimated Runoff in the San Antonio-Nueces Coastal Basin

4.3 Linking Expected Mean Concentration of Pollutants to Land Use

The measure of pollutant level that occurs during a runoff event is the expected mean concentration, or EMC, defined as the mass of pollutant transported per volume of runoff. For this study, it is assumed that expected mean concentrations of various pollutants are directly related to land uses in the drainage areas. In order to associate pollutant expected mean concentrations with land use, the land use coverage shown in Figure 3.5 is used along with the expected mean concentration data from Table 3.6.

Establishing a Link Attribute

A review of the data in Table 3.6 shows that, while expected mean concentration values are included for each subcategory of urban land use, only one value is included for the agricultural, range, and barren land use categories. However, all polygons in the land use coverage are delineated by subcategory. In order to facilitate the assignment of expected mean concentrations to land uses in the region, an additional attribute is first created in the polygon attribute table (pat) of the land use coverage. This new attribute, called lusecat, identifies the unique land use categories to which the expected mean concentrations are assigned.

The Arc/Info Tables tool is used to create the lusecat attribute. The attribute, defined as an integer, is first added to the polygon attribute table, using the Additem command. All land use subcategory polygons for which no unique expected mean concentrations exist are then reselected and the lusecat attribute for these polygons is defined as the truncated lanuse-id field, rounded to the lowest multiple of ten. This has the effect of redefining all agriculture land use subcategories, for example, to one value of land use category. For those land use subcategory polygons which do have corresponding unique expected mean concentrations (i.e. urban land uses), the lusecat attribute is defined as the value of the lanuse-id field. Finally, the Arc/Info Dissolve command is used to create a land use map with distinct category, versus subcategory, polygons.

Arc: **tables**
Enter Command: **additem sanlu.pat lusecat 8 8 i**
Enter Command: **sel sanlu.pat**

Enter Command: **reselect lanuse-id > 19**
Enter Command: **calc lusecat = lanuse-id / 10**
Enter Command: **sel**
File SANLU.PAT is now closed.
Enter Command: **sel sanlu.pat**
Enter Command: **calc lusecat = lusecat * 10**
Enter Command: **sel**
File SANLU.PAT is now closed.
Enter Command: **sel sanlu.pat**
Enter Command: **reselect lanuse-id < 19**
Enter Command: **calc lusecat = lanuse-id**
Enter Command: **quit**
Arc: **dissolve sanlu sanluse lusecat poly**
Arc: **kill sanlu all**
Arc: **rename sanluse sanlu**

Attaching the Expected Mean Concentration Data to Land Use

In order to attach the Expected Mean Concentration data from Table 3.6 to the land use coverage, a separate data table with each of the values listed by land use category must first be created. This data table, called emc3a.dat, is shown in Figure 4.13. Note that land use category appears as the first item in each row of the data and that expected mean concentration values for each pollutant are listed horizontally, in order of their appearance in Table 3.6, for each land use category. It should also be noted that expected mean concentration values for water, wetlands, tundra, and snowfield land uses are assumed to be zero for all pollutants and that the concentration values for range land uses are also applied to forest land uses in the basin. For the creation of this data file, special care must be taken to ensure that items in the file are delimited by single spaces and that the data is followed by an 'end' statement.

Once the raw expected mean concentration data file is created, it is used to fill a formatted data file, called attrib.dat, that is subsequently attached to the polygon attribute table of the land use coverage. Construction of the formatted data file is done with the Tables tool. A field for land use category is defined and then fields for each pollutant expected mean concentration value are defined. This process of defining the formatted data table is cumbersome and the potential for error in data input is significant. The process is more efficiently performed through use of an AML.

```
0 0.0 0.0 0.0 0.0 0.0 0 0.0 0.0 0 0.0 0.0 0.0 0.0 0.0 0.0 0 0
11 1.82 1.5 0.23 0.57 0.48 41.0 134 9.0 15.0 80 0.75 2.1 5.0 25.5 49.5 1.7 20000 56000
12 1.34 1.1 0.26 0.32 0.11 55.5 185 13.0 14.5 180 0.96 10.0 11.8 23.0 116.0 9.0 6900 18000
13 1.26 1.0 0.3 0.28 0.22 60.5 116 15.0 15.0 245 2.0 7.0 8.3 14.0 45.5 3.0 9700 6100
14 1.86 1.5 0.56 0.22 0.1 73.5 194 11.0 11.0 60 0.5 3.0 4.0 6.4 59.0 0.4 53000 26000
15 1.30 1.05 0.28 0.3 0.17 58.0 151 14.0 14.8 207 1.48 8.5 10.1 18.5 81.0 6.0 8300 12050
16 1.57 1.25 0.34 0.35 0.23 57.9 157 12.0 13.9 141 1.05 5.5 7.3 17.2 67.5 3.5 22400 26525
17 1.57 1.25 0.34 0.35 0.23 57.9 157 12.0 13.9 141 1.05 5.5 7.3 17.2 67.5 3.5 22400 26525
20 4.4 1.7 1.6 1.3 0.0 107.0 1225 1.5 1.5 16 1.0 5.0 0.0 4.0 0.0 0.0 0 0
30 0.7 0.2 0.4 0.0 0.0 1.0 245 5.0 5.0 6 0.5 7.5 0.0 0.5 0.0 0.0 200 0
40 0.7 0.2 0.4 0.0 0.0 1.0 245 5.0 5.0 6 0.5 7.5 0.0 0.5 0.0 0.0 200 0
50 0.0 0.0 0.0 0.0 0.0 0.0 0 0.0 0.0 0 0.0 0.0 0.0 0.0 0.0 0.0 0 0
60 0.0 0.0 0.0 0.0 0.0 0.0 0 0.0 0.0 0 0.0 0.0 0.0 0.0 0.0 0.0 0 0
70 1.5 0.96 0.54 0.12 0.03 70.0 0 1.52 0.0 0 0.0 0.0 0.0 0.0 40.0 0.0 0 0
80 0.0 0.0 0.0 0.0 0.0 0.0 0 0.0 0.0 0 0.0 0.0 0.0 0.0 0.0 0.0 0 0
90 0.0 0.0 0.0 0.0 0.0 0.0 0 0.0 0.0 0 0.0 0.0 0.0 0.0 0.0 0.0 0 0
end
```

Figure 4.13 : Conversion of Tabulated Expected Mean Concentration Values to an Arc/Info Data File

Appendix B includes the attrib.aml file, which is used to define item formats in the attrib.dat file and then fill the formatted file with raw data from the emc3a.dat file.

Finally, the expected mean concentration data is attached to the land use polygon attribute table through use of the Joinitem command, using the lusecat field as the linking item between both tables:

Arc: **joinitem sanlu.pat attrib.dat sanlu.pat lusecat lusecat**

The resulting land use coverage includes 18 new fields identifying pollutant expected mean concentrations for each land use category within the basin. The land use coverage can be used to show how expected mean concentrations for a particular pollutant vary throughout the land use polygons of a particular region. For instance, Figure 4.14 shows expected mean concentrations for total phosphorus, based on the land use polygons within the San Antonio-Nueces Coastal Basin. As expected, the highest concentrations of total phosphorus are identified in the regions where agricultural land uses are predominant.

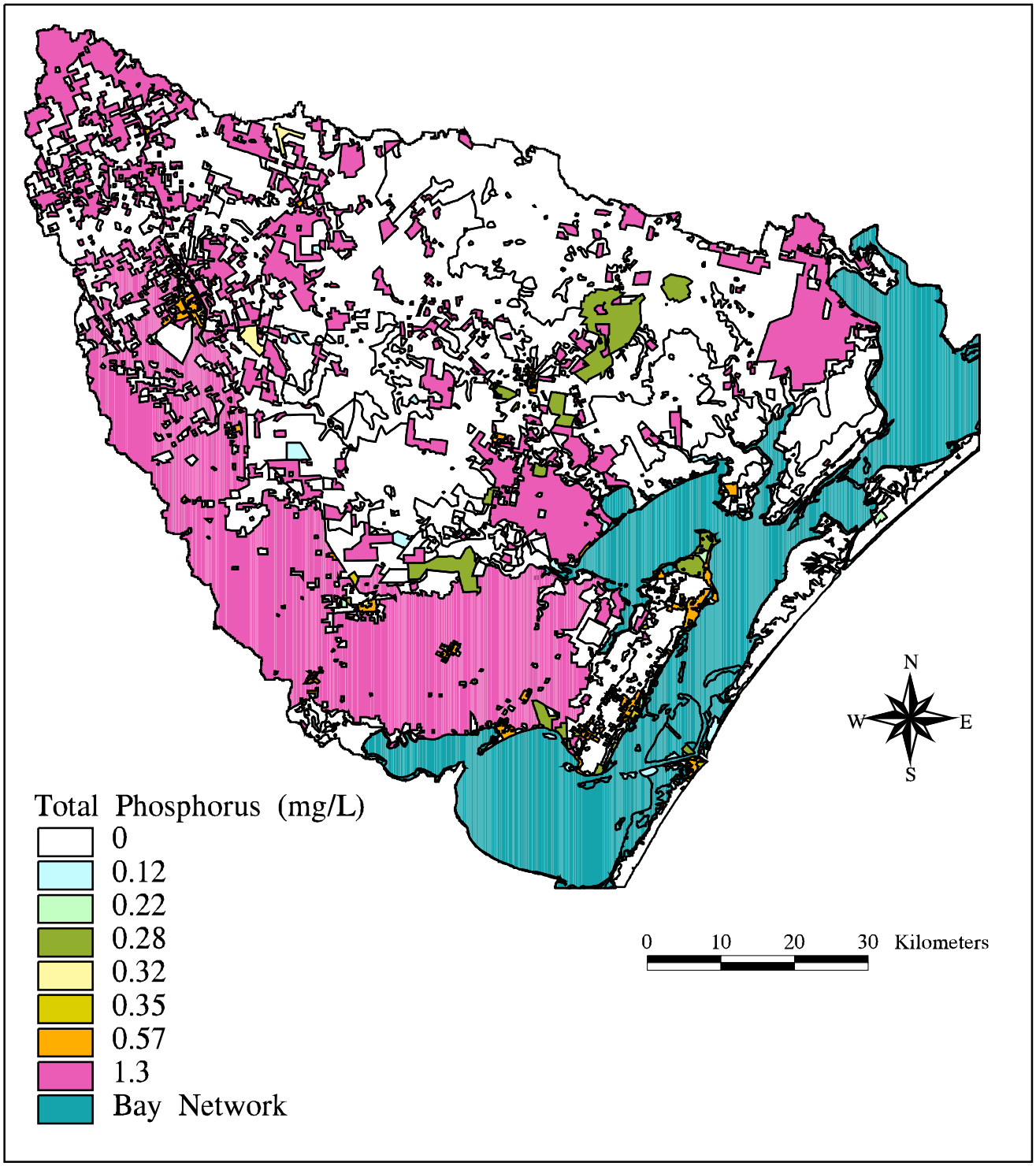


Figure 4.14 : Expected Mean Concentrations for Total Phosphorus in the San Antonio-Nueces Coastal Basin

4.4 Estimating Annual Loadings Throughout the Watershed

The pollutant mass contribution that each cell makes to downstream pollutant loading is calculated by taking the product of the expected mean concentration and runoff associated with the cell, or

$$\text{Load (mass/time)} = \text{EMC (mass/volume)} * \text{Q (volume/time)}. \qquad (4\text{-}5)$$

For load computations in this study, equation 4-5 becomes

$$L = K * Q * EMC * A, \qquad (4\text{-}6)$$

where Q is given in units of mm/year, EMC is given in units of mg/Liter, A is the area of one grid cell (10,000 m^2), and K is a constant to make the units consistent, i.e. K = 10^{-6} kg-m-L/mg-mm-m^3, so that L is determined in units of kg/year. This approach to representation of loadings assumes that the downstream transport process is conservative, i.e. no pollutant decay occurs along the flow paths. This assumption is considered appropriate for the pollutants in Table 3.6 along the short flow paths of the San Antonio-Nueces Basin. Another important point about this relationship is that it applies expected mean concentration, which is typically associated with single runoff events, to mean annual runoff, which generally includes stream base flow as well as runoff from storm events.

Pollutant loadings associated with each grid cell are determined by first converting the expected mean concentration map coverage to a grid, through use of the Polygrid command. For the creation of this grid, cell values are determined from the appropriate concentration attribute of the land use coverage. For the case of total phosphorus, the tp field is specified as the item from which to extract cell values. Once the expected mean concentration grid is created, a cell-based loading grid is established as the product of this grid and the runoff grid.

Grid: **phosgrid = polygrid(sanlu,tp,#,#,100)**
Grid: **phosrnof = phosgrid * runoff**

Before a cumulative annual loading grid is created, it is noted that, for display purposes, a representation of cumulative loads in the stream networks is desired. One

way to accomplish this is through the conversion of grid cell strings to an equivalent arc coverage, using the Streamline command. However, arcs created using Streamline start at the geographic center of the endpoint cell, rather than including the full width of the cell. The result of this idiosyncrasy is that the equivalent arc of a gridded stream falls one-half cell short of its expected outlet point.

To correct for this anomaly, the mainland portion of the baycov coverage, created in section 4.1, is isolated using the Reselect command. The new mainland coverage is then buffered by 100 meters and the buffered coverage is converted to an equivalent grid, using Polygrid. Finally, a flow direction grid specific to the buffered mainland coverage is created with the Con statement. This procedure has the effect of creating a flow direction grid that covers the mainland plus a 100-meter boundary extending out into the bay network.

```
Arc:  reselect baycov mainland
>:  res baycov-id = 1
>:  ~
Do you wish to re-enter expression (Y/N)?  n
Do you wish to enter another expression (Y/N)?  n
  1 features out of 30 selected
Arc:  buffer mainland main # # 100
Arc:  grid
Grid:  maingrid = polygrid(main,#,#,#,100)
Grid:  mainfdr = con(maingrid,ditfdr)
```

Cumulative annual loading in the basin is determined by performing a weighted flow accumulation, using the cell-based loading grid as the weight grid and the new buffered mainland flow direction grid. Division by 100 is introduced into this command, as per equation 4-6, to provide the result in units of kg/year.

```
Grid:  phosld = flowaccumulation(mainfdr,phosrnof) / 100
```

In order to facilitate the conversion of the cumulative loading grid to a coverage, an integer grid of cumulative load is first created. Then the Con statement is used with the Streamline command to effectively reselect all grid cells with value greater than or equal to a threshold of 1000. Selection of this threshold value reduces the number of cells to be converted to those that occur at in-stream locations, where accumulated loads are greatest. The specific threshold value is not arbitrary, but

should be selected so as to reflect as much of the known stream network as possible. Finally, the cumulative loadings coverage is clipped with the mainland template, so that the endpoints of the streams occur exactly at the bay network borders.

Grid: **phosload = int(phosld)**
Grid: **tpline = streamline(con(phosload >= 1000,phosload),mainfdr,grid-code)**
Arc: **clip tpline mainland tpload line**

By performing a Describe command on the annual cumulative loading grid (phosload), the maximum value (i.e. load) in the grid can be identified. Also, by querying the various outlet cells to the bay network with the Cellvalue command, annual cumulative loads from each subwatershed in the basin can be established.

Grid: **describe phosload**

Description of Grid PHOSLOAD

Cell Size =	100.000	Data Type:	Integer
Number of Rows =	1325	Number of Values =	4884
Number of Columns =	1520	Attribute Data (bytes) =	8

BOUNDARY		STATISTICS	
Xmin =	1180828.125	Minimum Value =	0.000
Xmax =	1332828.125	Maximum Value =	60900.000
Ymin =	612183.250	Mean =	74.213
Ymax =	744683.250	Standard Deviation =	1553.429

Grid: **gridpaint phosload value linear nowrap gray**
Grid: **cellvalue phosload ***
The cell containing point (1267701.191,660318.274) has value 60900

Figure 4.15 shows annual cumulative loads of total phosphorus in the San Antonio-Nueces basin, using the grid-code attribute of the tpload coverage to display aerial distributed values of load greater than thresholds of 1000 kg/yr, 5000 kg/yr, 10,000 kg/yr, and 50,000 kg/yr. Specific load values at five bay network outlet points are identified on the figure. It should be noted that the largest contributions of phosphorus load are seen to be from the agricultural part of the basin in the Aransas subwatershed.

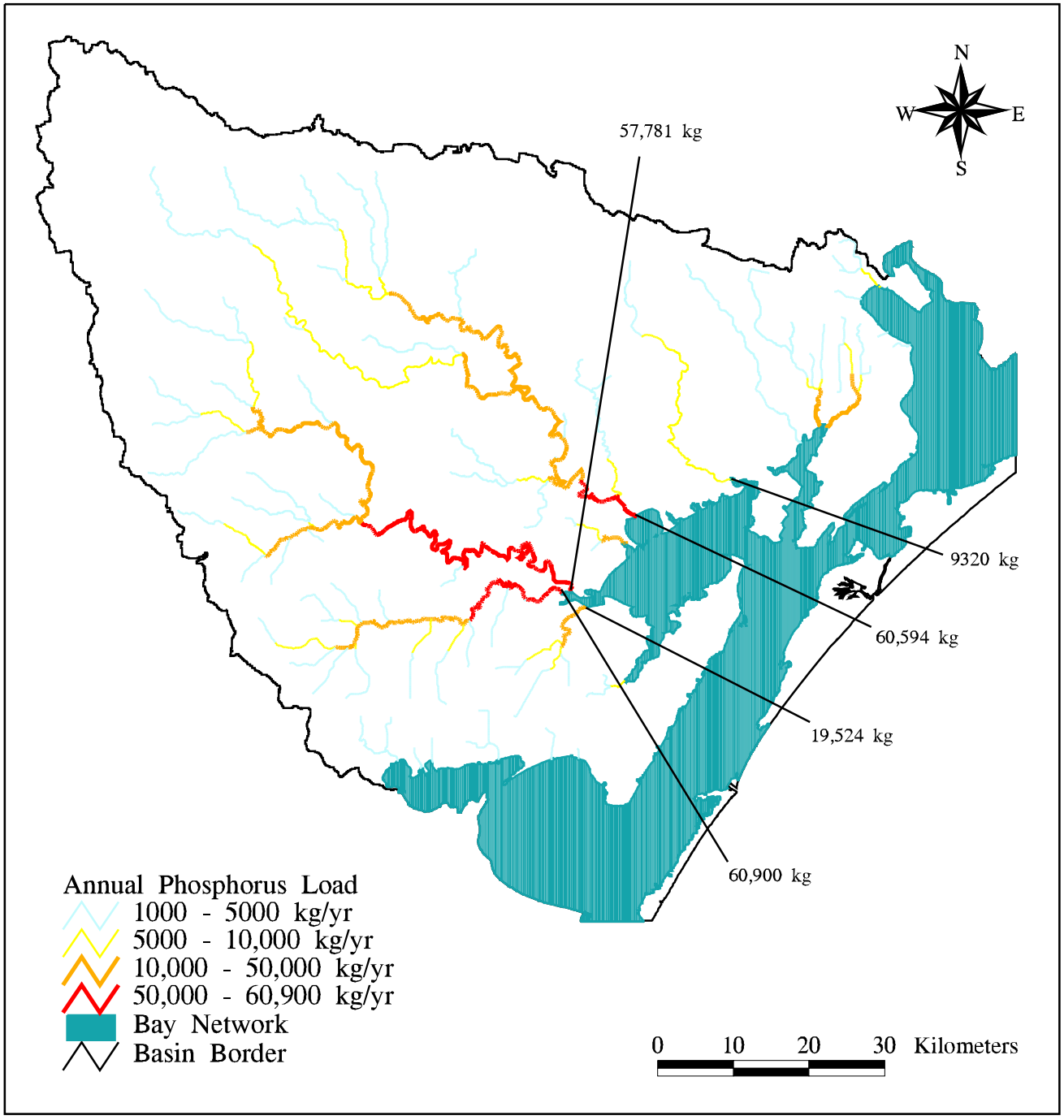


Figure 4.15 : Average Annual Total Phosphorus Loads in the San Antonio-Nueces Coastal Basin

4.5 Predicting Downstream Pollutant Concentrations in Watershed Stream Networks

Pollutant concentrations that are sampled at various in-stream locations result from the mixing of all pollutant-laden flows draining from upstream of the particular location. For a digitally discretized grid model, this mixing process is approximated by dividing the accumulated load at each cell by the accumulated runoff that also occurs there. Mathematically, this is represented by

$$C_a = L_a / Q_a, \tag{4-7}$$

where L_a is the annual cumulative loading, Q_a is the annual cumulative runoff, and C_a is the average concentration expected at the location.

These predicted concentration values can be compared with measured data from a sampling program in order to assess the accuracy of the predicted values. For this study the water quality measurement data described in section 3.2 are used for comparison. For each sampling location in the data set, the assumption is made that the expected observed concentration is simply the average of all the measurements made there, or

$$C_o = (1/n) * \sum_{i=1}^{n} C_i , \tag{4-8}$$

where C_i is each concentration value measured at a particular sampling location, n is the total number of samples made at that location, and C_o is the average observed concentration.

Estimating Average Concentrations

Before estimated concentrations can be calculated, grids of annual cumulative loading and annual cumulative runoff need to be established. Grids of annual cumulative loading are created as per the procedure in section 4.4. Annual cumulative runoff is created by performing a weighted flow accumulation, using the runoff grid as

the weight grid. The result of the weighted flow accumulation is multiplied by 10 to convert from runoff units of mm/yr to accumulated units of m^3/yr, as in equation 4-1.

Grid: **runoffac = flowaccumulation(mainfdr,runoff) * 10**
Grid: **describe runoffac**

By performing a Describe command on the cumulative runoff grid, the maximum value of the grid is determined as more than 290 million m^3/yr. This is the value at the outlet of the Mission River to Mission Bay. The equivalent annual cumulative runoff grid, in units of cubic feet per second (cfs), is calculated by multiplying the runoffac grid by the number of cubic feet per cubic meter and dividing by the number of seconds per year. In these units of measure, the annual cumulative runoff is represented as an average stream flow and is more easily compared with recorded USGS stream flow values. For display purposes, an equivalent coverage of the accumulated runoff grid is created by first converting the real number grid to an integer grid. Then the Streamline command is used, along with the Con statement, to create arcs for all cells having value greater than or equal to a certain threshold value, specified so that only in-stream cells are converted. For this conversion, the threshold value is chosen to be 1 cfs. The cumulative runoff coverage is then clipped with the mainland coverage to create cumulative runoff arcs that end exactly at the boundaries of the bay network. Figure 4.16 shows average stream flows in units of cubic feet per second.

Grid: **rofaccfs = runoffac * 35.2875 / 31557600**
Grid: **introfac = int(rofaccfs)**
Grid: **rofaclin = streamline(con(introfac >= 1,introfac),mainfdr,grid-code)**
Arc: **clip rofaclin mainland rofaccov line**

Once the annual cumulative runoff grid is created, a grid of predicted pollutant concentration can be created as per equation 4-7. Using total phosphorus as an example pollutant, a grid of predicted concentrations is produced by dividing the annual total phosphorus cumulative load grid by the annual cumulative runoff (m^3/yr) grid. Multiplication of this result by 1000 produces a concentration grid in units of mg/L as per the equation

$$C\ (mg/L) = L\ (kg/yr) / Q\ (m^3/yr) * 10^6\ mg/kg * .001\ m^3/L. \qquad (4\text{-}9)$$

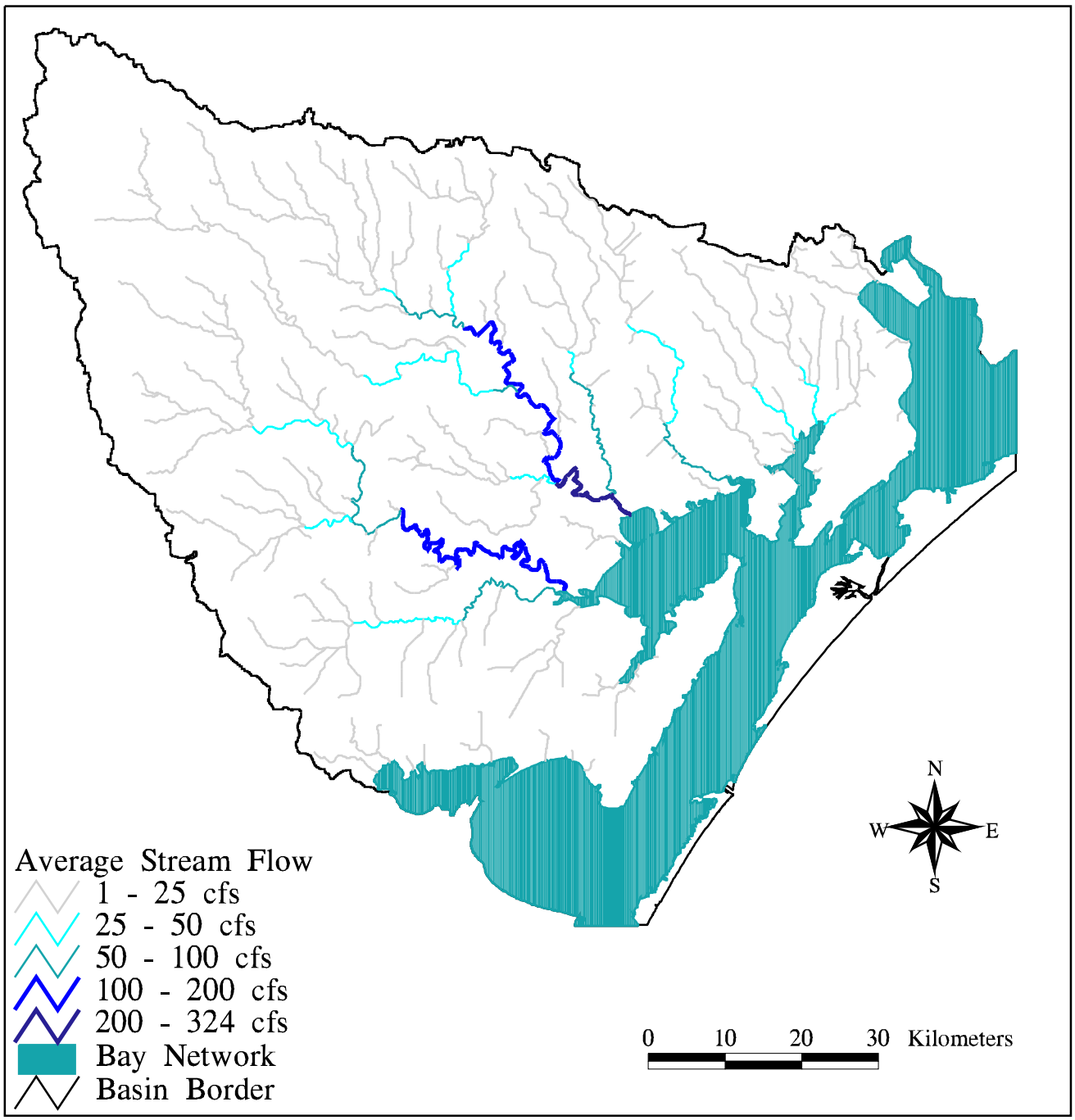


Figure 4.16 : Average Annual Stream Flows in the San Antonio-Nueces Coastal Basin

A grid of concentration values specific to the basin stream network is established using the Con statement with the introfac grid created above. Values from the predicted total phosphorus concentration grid are filled into those cells that correspond to locations along the stream networks. Since arc coverages may only be converted from integer value grids, the stream concentration grid is multiplied by 1000 to retain significant figures, the product is truncated to create the integer grid, and the resulting grid is converted to a coverage, using the Streamline command. Finally, the phosphorus concentrations arc coverage is clipped so that the concentration arcs end exactly at the shores of the bay network.

Grid: **phosconc = phosload / runoffac * 1000**
Grid: **phconstr = con(introfac >= 1,phosconc)**
Grid: **phline = streamline(int(phconstr * 1000),mainfdr,grid-code)**
Arc: **clip phline mainland phcon line**

Figure 4.17 shows the predicted concentrations for total phosphorus in the San Antonio-Nueces coastal basin. These predicted concentrations represent the levels of pollution that are attributed to nonpoint source runoff, only. Additional point source pollutant loadings are considered in section 4.6.

Attaching Observed Concentration Data to Measurement Locations

The Surface Water Quality Monitoring (SWQM) data described in section 3.2 are used for comparison with the predicted concentration values. With the data linked in ArcView 2.0 as shown in Figure 3.11, the average measured value of a particular pollutant constituent is established through use of the Summary Statistics tool. First, a pollutant is selected in the storet.dbf table. Then, with the station_id field selected in the value.dbf table, the Summary Statistics tool is invoked. This tool allows the user to sort and manipulate data from the selected table, using the previously selected field to sort by. Using the tool, the Value field is specified as the data to manipulate and the Summary Statistics Averaging function is performed on the data. This process creates a new database file (.dbf) that includes three fields: (1) all station-id's reporting data for the particular pollutant, (2) a field called count that represents the total number of measurements of the pollutant at that station, and (3) a field called ave_value that

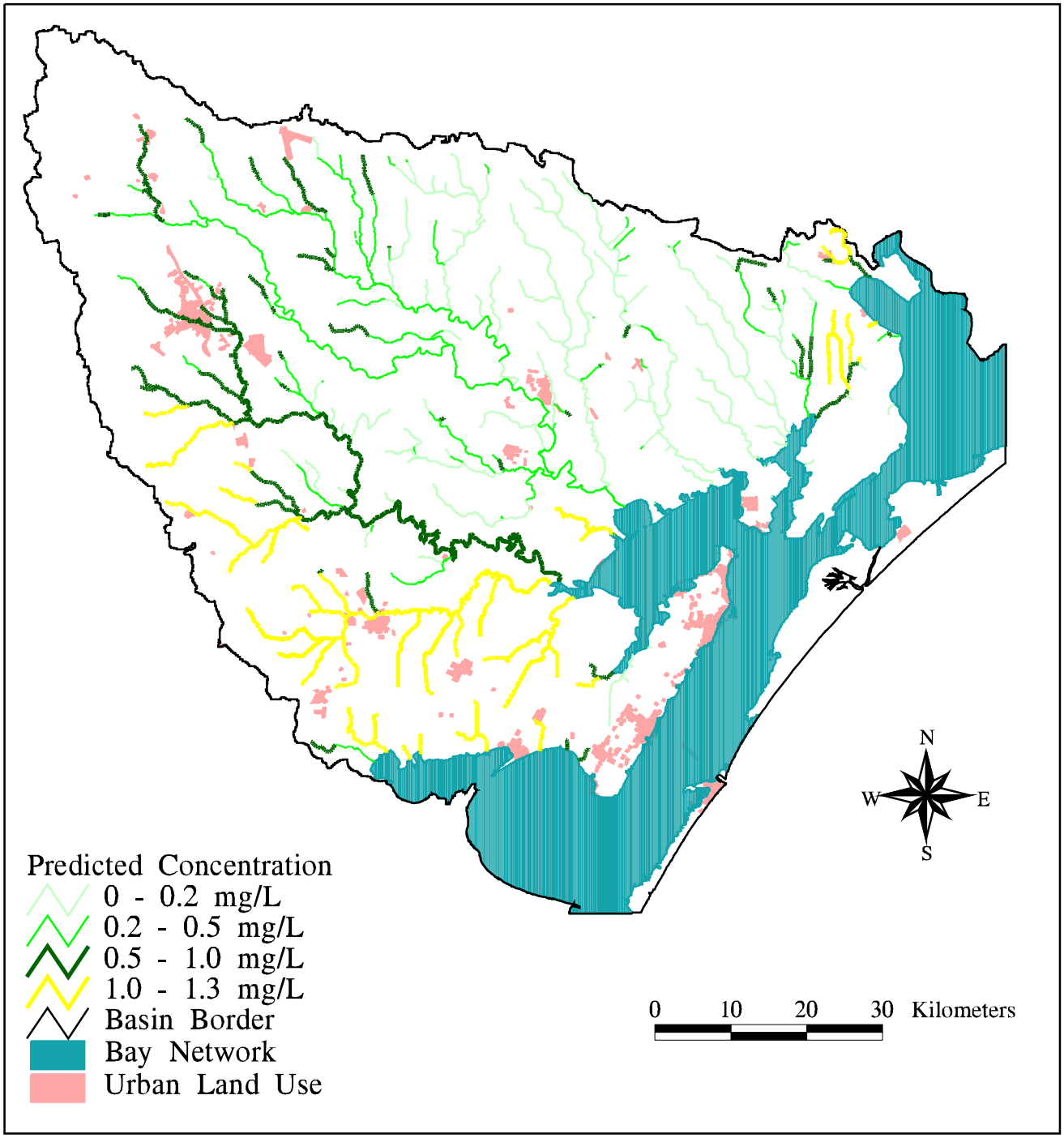


Figure 4.17 : Estimated Total Phosphorus Concentrations in the San Antonio-Nueces Coastal Basin

represents the mean value of the specified measurements. Table 4.6 shows a portion of the tp.dbf file identifying all measurement locations where total phosphorus is measured, the number of measurements at each location, and the average concentrations at each location.

The tp.dbf file is attached to the water quality measurement stations point coverage in Arc/Info. First, the file is converted to an Arc/Info Information file (.dat) using the Dbaseinfo command. The new tp.dat file is then attached to the sanwq point attribute table using the Joinitem command with the station_id field specified as the link item. Using the Arc/Info Tables module, the new count and ave_value fields of the sanwq point attribute table are altered to have the more definitive tp_cnt and tp_avg field names.

```
Arc:  dbaseinfo tp.dbf tp.dat
Arc:  joinitem sanwq.pat tp.dat sanwq.pat station_id station_id
Arc:  tables
Enter Command:  sel sanwq.pat
  105 Records selected
Enter Command: alter
Enter item name: count
COLUMN   ITEM NAME      WIDTH OUTPUT  TYPE N.DEC  ALTERNATE NAME
  22                COUNT             8       11         F         0
Item name: tp_cnt
Item output width: 11
Item type: f
Item decimal places: 0
Alternate item name: ~
COLUMN   ITEM NAME      WIDTH OUTPUT  TYPE N.DEC  ALTERNATE NAME
  22                TP_CNT            8       11         F         0
Enter item name: ave_value
COLUMN   ITEM NAME      WIDTH OUTPUT  TYPE N.DEC  ALTERNATE NAME
  30           AVE_VALUE          8       16         F         2
Item name: tp_avg
Item output width: 16
Item type: f
Item decimal places: 2
Alternate item name: ~
COLUMN   ITEM NAME      WIDTH OUTPUT  TYPE N.DEC  ALTERNATE NAME
  30                TP_AVG            8       16         F         2
Enter item name:  ~
Enter Command:  quit
```

STATION_ID	COUNT	AVE_VALUE
12932	2	0.61
12933	5	6.60
12934	1	7.36
12935	6	6.28
12937	2	6.61
12938	2	5.94
12939	2	4.26
12940	2	4.22
12941	1	0.25
12942	1	0.16
12943	27	0.15
12944	75	0.06
12945	27	0.14
12946	1	0.28
12947	2	0.50
12948	39	1.09
12949	2	1.73
12950	1	2.19
12951	2	2.91
12952	3	4.47
12953	1	3.01
13030	1	0.14
:	:	:
:	:	:

Table 4.6 : Summary Statistics for Total Phosphorus

The procedure of using the ArcView Summary Statistics tool and attaching average concentration values to the sanwq point attribute table is repeated for each pollutant constituent of interest (i.e. those pollutants identified in Table 3.6). Nitrogen, however, is not sampled and reported as total nitrogen in the Surface Water Quality Monitoring data set. Instead, total kjeldahl nitrogen (organic plus ammonia nitrogen), nitrate nitrogen, and nitrite nitrogen are reported separately. These are the components that total nitrogen is comprised of (American Public Health Association, American Water Works Association and Water Environment Federation, 1992). Each of the three nitrogen components is summarized, averaged, and attached to the sanwq point attribute table along with the other pollutant constituents from Table 3.6. Then two additional fields, tn_cnt and tn_avg, are added to the point attribute table using the Joinitem command. In the Tables module, the number of effective total nitrogen measurements at each location is determined as the average of the number of measurements for each component. The average value for total nitrogen concentration at each location is determined as the sum of the average values for each component. Finally, X- and Y-coordinate values are added to each record in the sanwq point attribute table through use of the Addxy command:

```
Arc:  additem sanwq.pat sanwq.pat tn_cnt 8 8 f 0
Arc:  additem sanwq.pat sanwq.pat tn_avg 8 8 f 2
Arc:  tables
Enter Command:  sel sanwq.pat
  105 Records selected
Enter Command:  calc tn_cnt = ( tkn_cnt + no2_cnt + no3_cnt ) / 3
Enter Command:  calc tn_avg = tkn_avg + no2_avg  + no3_avg
Enter Command:  quit
Arc:  addxy sanwq
```

Analyses of the Surface Water Quality Monitoring data at specific locations and for specific pollutants reveal some interesting points. Figure 4.18 shows all of the total phosphorus measurements taken at station #12948 along the Aransas River about 15 kilometers upstream of Copano Bay. By plotting these concentration levels against the sampling dates, the variations in concentration magnitude are plainly seen. A plot of the average concentration overlaid on the data shows the effect of a few elevated concentration measurements on the average value and suggests that consideration and

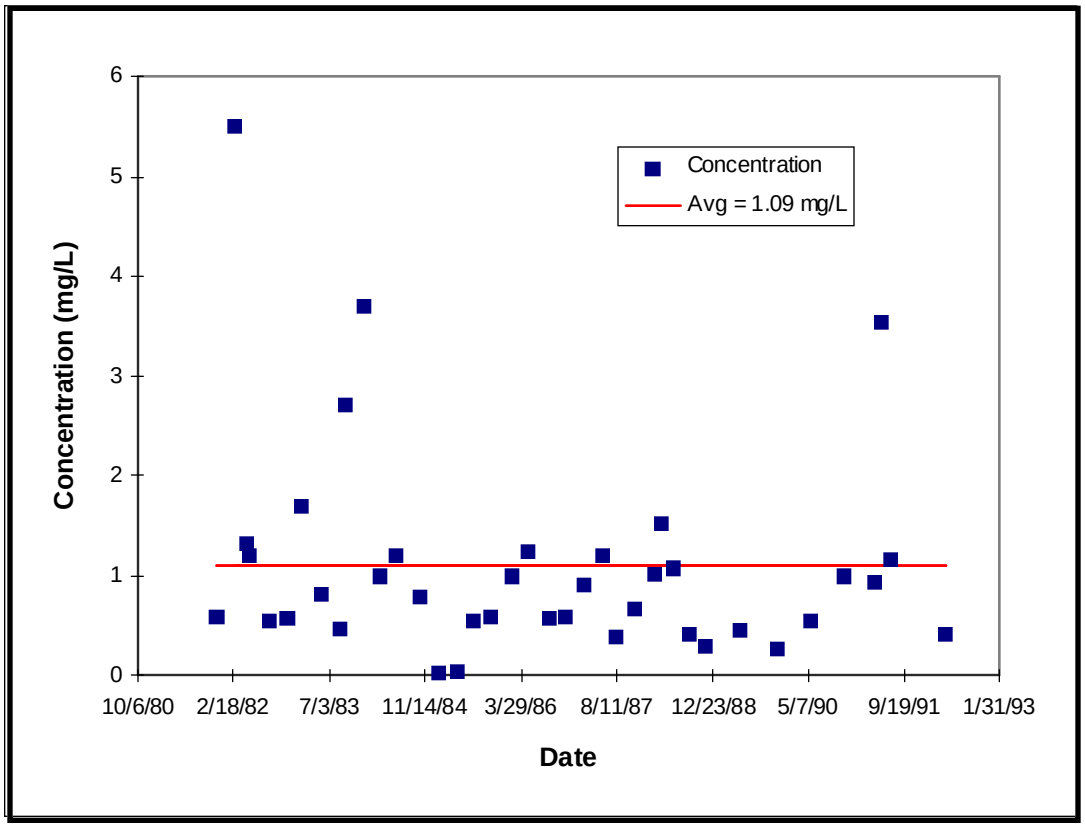


Figure 4.18 : Total Phosphorus Concentration Measurements at TNRCC SWQM Station # 12948 (Aransas River)

possible removal of outlying data points may be appropriate for determination of a revised average.

Figures 4.19 and 4.20 respectively show the nitrogen component measurements made at the Aransas station and at station #12944 along the Mission River about 10 kilometers upstream of Mission Bay. Each of these plots also shows the value for total nitrogen, calculated as the sum of the average total kjeldahl, total nitrate, and total nitrite levels. Values for total kjeldahl and total nitrite nitrogen generally fall into fairly well-bounded ranges, but nitrate nitrogen concentration values, particularly at the Mission River station, show an occasional tendency to vary significantly from the normal range. These atypical measurements have a significant effect on the calculated average total nitrate concentration which, in turn, affects the calculation of average total nitrogen concentration. In fact, the single outlying total nitrate concentration data point observed at the Mission station (Figure 4.20) affects the calculated average total nitrate concentration by almost 200%, increasing it from about 0.077 mg/L to 0.22 mg/L. As a result, average total nitrogen calculated for the station is 18% higher than it would be without inclusion of the anomalous data point. This point emphasizes that outlying data points should be considered when establishing averaged values for pollutant concentration at a particular location.

A second point of interest regarding the Surface Water Quality Measurement nitrogen data is illustrated in Figure 4.21, which shows the percentile distributions, for both the Aransas and Mission stations, of the three components that contribute to the calculated average total nitrogen concentrations. The charts in this figure have been determined using all data points from each of the stations, i.e. without consideration and removal of outlying data points. The chart shows that, for both locations, most of the total nitrogen observed is of an organic nature. The oxidized forms of nitrogen account for only 25-30% of the total observed (before consideration of outlying points). Organic and ammonia nitrogen is typically associated with agricultural land uses and the fact that kjeldahl nitrogen accounts for over 70% of the total nitrogen measured in the two main streams of the basin indicates a significant contribution from the local agricultural lands.

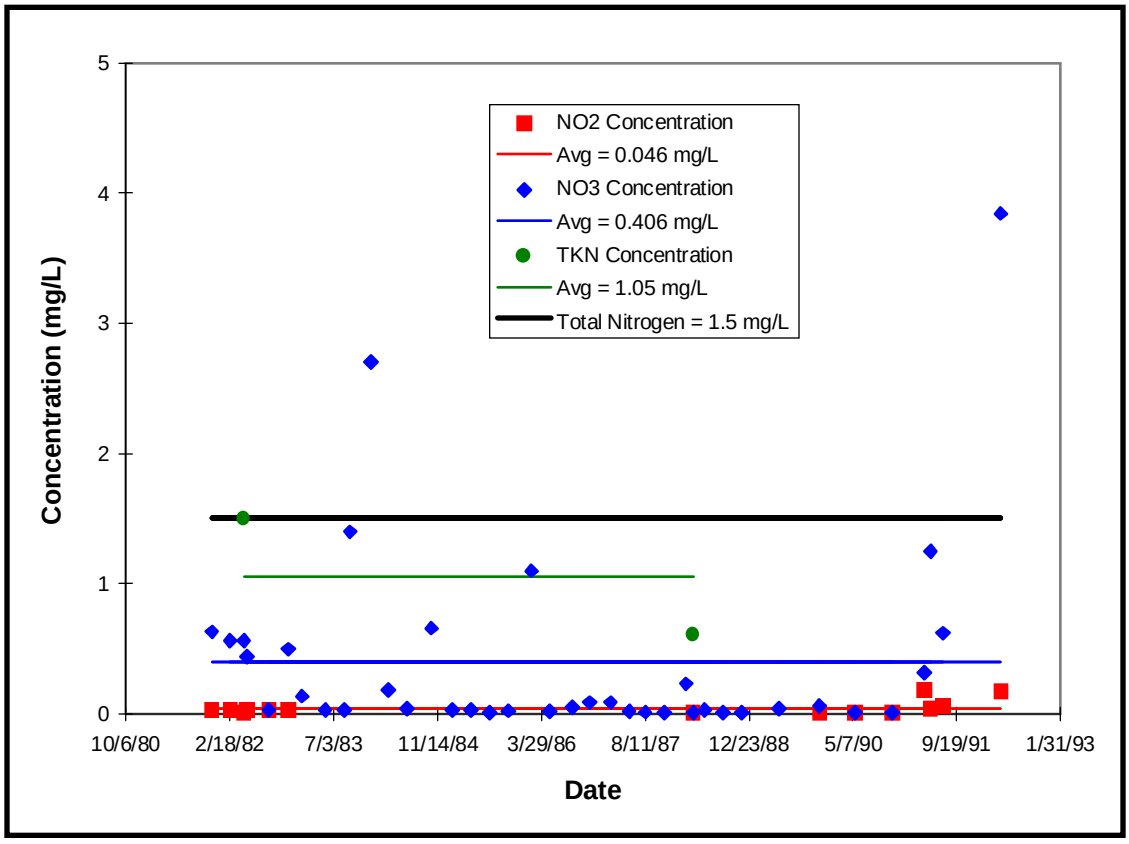


Figure 4.19 : Total Nitrogen Component Concentration Measurements at TNRCC SWQM Station # 12948 (Aransas River)

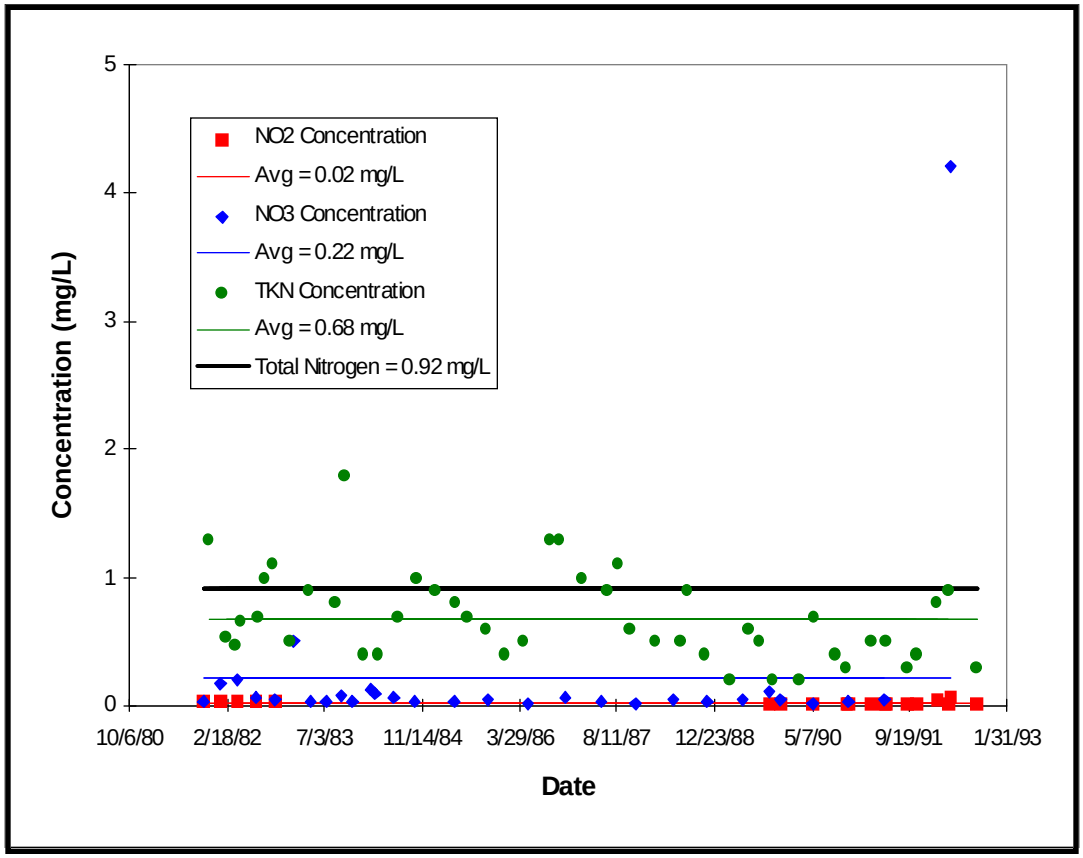


Figure 4.20 : Total Nitrogen Component Concentration Measurements at TNRCC SWQM Station # 12944 (Mission River)

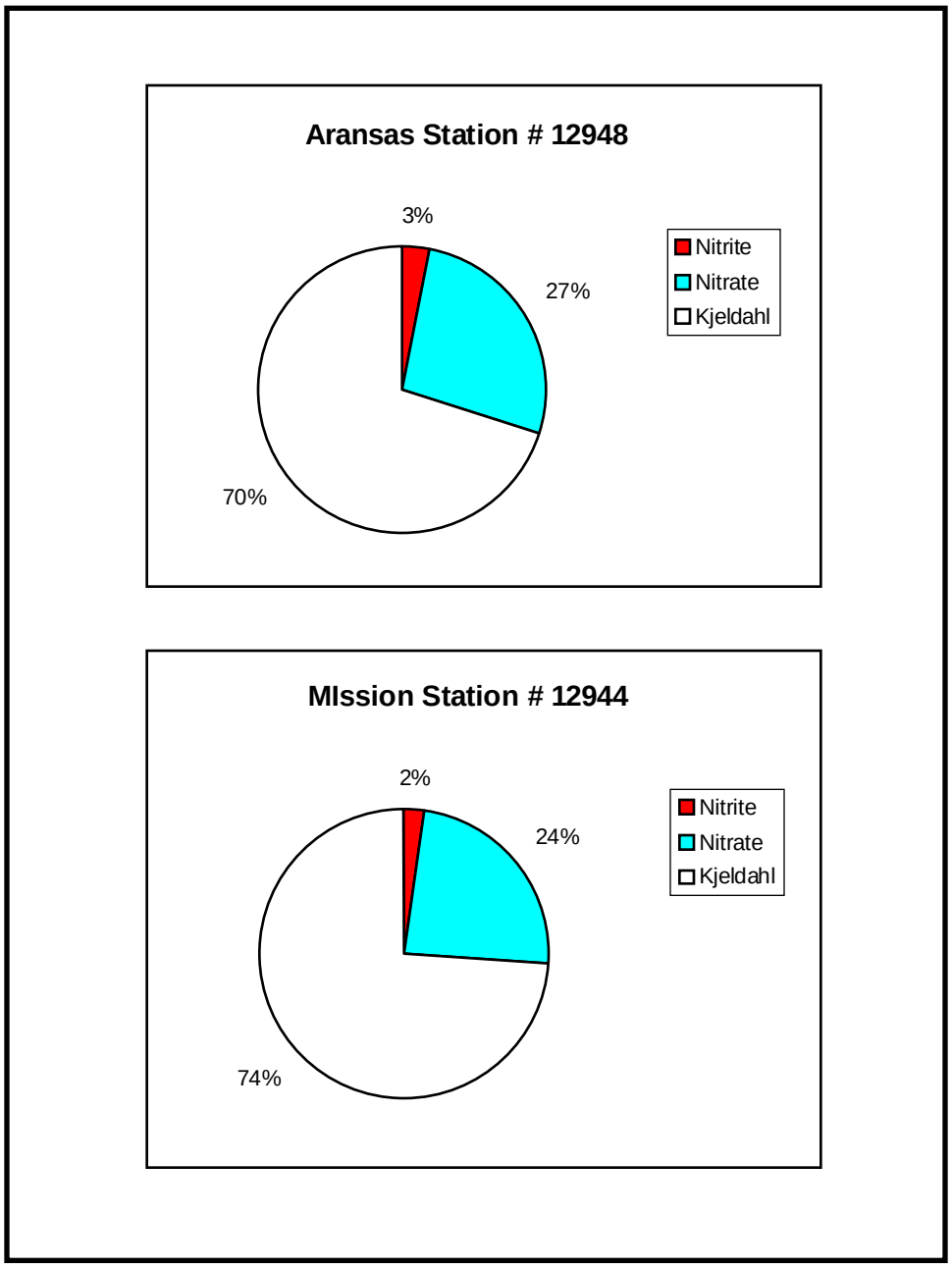


Figure 4.21 : Percentile Distribution of Total Nitrogen Components Measured at Two Locations in the San Antonio-Nueces Basin

Graphically Depicting Variations in the Frequency of Concentration Sampling

The average concentrations that are attached to the water quality measurement points are calculated by averaging various numbers of measurements. In fact, for total phosphorus, Table 4.6 shows one average concentration derived from 75 measurements while a number of locations have only one measurement defining average concentration. One would be correct in placing more statistical validity in those averages derived from larger numbers of measurements.

A method of depicting this variation in the number of concentration measurements is established by converting the water quality measurement point coverage into a polygon coverage of circles, where each circle is centered about the measurement location coordinates and each circle's area is approximately proportional to the number of measurements made at the station. This is done by (1) adding a radius field to each record in the sanwq point attribute table, (2) calculating values for radius based on the number of measurements for the pollutant constituent of interest, (3) creating a text-delimited data file from the station-id, x-coordinate, y-coordinate, and radius fields, (4) generating a polygon coverage from the data file, and (5) attaching the pollutant measurement data to the new polygon coverage.

The first three of these steps are performed in ArcView 2.0: For the case of total phosphorus measurements, the sanwq point attribute table is displayed and the Properties feature in the Table menu is used to deselect all fields except for station_id, x-coord, y-coord, and tp_cnt. The Table menu is used once again to Start Editing of the table. The Add Field feature from the Edit menu is then invoked and the Radius field is defined as an 8-character numeric item.

The Calculate feature of the Field menu is used to specify that values in the Radius field are determined as the truncated square root of the tp_cnt field multiplied by 200 meters, or

$$\text{Radius} = \text{tp\_cnt.sqrt.truncate} * 200. \qquad (4\text{-}10)$$

The value of 200 meters is selected, by trial and error, as the smallest radius that produces a discernible circle for single measurement stations, while maintaining a reasonably sized circle for locations with many measurements. By taking the square

root of the number of pollutant measurements, the area of the circle (π * radius^2) is made proportional to the number of measurements. Once the values for the Radius field are filled, the Stop Editing feature is selected from the Table menu.

The Properties feature in the Table menu is used to deselect the tp_cnt field from the sanwq point attribute table, leaving only the station-id, x-coord, y-coord, and radius fields displayed, in that order. The Export feature from the File menu is then invoked to create a text-delimited file containing the values of these four fields. A portion of this text-delimited file, called rad.txt, is shown in Figure 4.22.

A raw data file (rad.dat) is created from this text-delimited file by removing the column labels in the header and appending the bottom of the file with an END statement. This raw data file is then used in conjunction with the Arc/Info Generate command to create a coverage of circles at each measurement location. Polygon topology is created through use of the Clean command:

```
Arc:  generate phospts
Generate:  input rad.dat
Generate:  circles
  Creating Circles with coordinates loaded from rad.dat
Generate:  quit
  Externalling BND and TIC.....
Arc:  clean phospts phopts
```

Finally, water quality measurement data is attached to the phopts coverage by adding an integer field called station_id to the phopts polygon attribute table, filling those fields with the values from the phopts-id field, altering the station_id field to character type, and performing a Joinitem command with the tp.dat file, using the station_id field to join the two files.

```
Arc:  additem phopts.pat phopts.pat station_id 5 5 i
Arc:  tables
Enter Command:  sel phopts.pat
    24 Records selected
Enter Command:  calc station_id = phopts-id
Enter Command: alter
Enter item name: station_id
COLUMN   ITEM NAME       WIDTH OUTPUT  TYPE N.DEC  ALTERNATE NAME
  17          STATION_ID          5         5         I         -
Item name: station_id
Item output width: 5
Item type: c
```

```
"Sanwq-id","X-coord","Y-coord","Radius"
12931,1253025.250,696013.875,0
12932,1222723.500,694795.125,200
13399,1321153.000,694116.750,0
12933,1223832.250,693739.500,400
12934,1224645.625,693725.188,200
12936,1223313.875,693697.688,0
12935,1225518.375,693651.250,400
12937,1226820.250,692784.062,200
12939,1226039.375,690734.125,200
12938,1226287.000,690616.125,200
13660,1282946.625,685779.625,800
12942,1225275.500,685540.438,200
12944,1266646.625,684073.938,1600
12940,1227986.625,682454.125,200
12952,1233187.250,682256.125,200
12953,1230096.875,681881.750,200
12941,1226212.125,681493.000,200
13398,1315794.250,678917.375,0
12951,1242619.250,676083.625,200
13401,1309669.875,673810.625,0
13406,1299136.750,672437.562,0
12943,1273454.375,672223.500,1000
13400,1307925.750,670925.375,0
12950,1242018.500,670924.062,200
12949,1249132.375,668866.188,200
13404,1292647.500,666251.562,0
12948,1252749.000,665714.812,1200
   :          :          :          :
   :          :          :          :
   :          :          :          :
```

Figure 4.22 : Text-Delimited File of Water Quality Measurement Radii

```
Alternate item name: ~
COLUMN   ITEM NAME       WIDTH OUTPUT  TYPE N.DEC  ALTERNATE NAME
  17          STATION_ID              5        5         C        -
Enter item name: ~
Enter Command:  quit
Arc:  joinitem phopts.pat tp.dat phopts.pat station_id station_id
```

This procedure is performed for each pollutant constituent of interest. However, since no .dat file exists for total nitrogen, the polygon attribute table for that coverage of circles is joined with the sanwq point attribute table, which contains the average values for all pollutant constituents of interest. The sanwq point attribute table is actually an alternative source of average concentration data for all of the circle coverages.

Figure 4.23 shows the predicted total phosphorus concentration data overlaid with the phopts polygons. For display purposes, these circles are provided with a label of the average concentration at the location concatenated with the number of total phosphorus measurements. This label is created in ArcView 2.0 by adding a new character field and, using ArcView's internal Avenue programming language, defining the contents of the character string as

```
[pho_tag] = [tp_avg].SetFormat("d.dd").AsString ++
            "(" ++ [tp_cnt].AsString ++ ")",                    (4-11)
```

where .AsString converts the value of the preceding variable to a character string and .SetFormat("d.dd") specifies a floating point numeric format for the preceding variable.

Figure 4.23 also shows interesting trends in the comparison of predicted and average observed values for total phosphorus concentration. Using the same color coding scheme to represent predicted and observed concentrations, it can be seen that, within the Mission and Copano subwatersheds, estimated concentrations generally match the minimal levels that have historically been recorded there, between 0.1 and 0.3 mg/L. However, in the Aransas subwatershed, observed concentrations significantly exceed predicted levels. In particular, observed concentrations just downstream from the city of Beeville (Figure 4.24) are seen to reach above 7 mg/L, whereas predicted concentrations in the same reaches of the river are less than 1 mg/L.

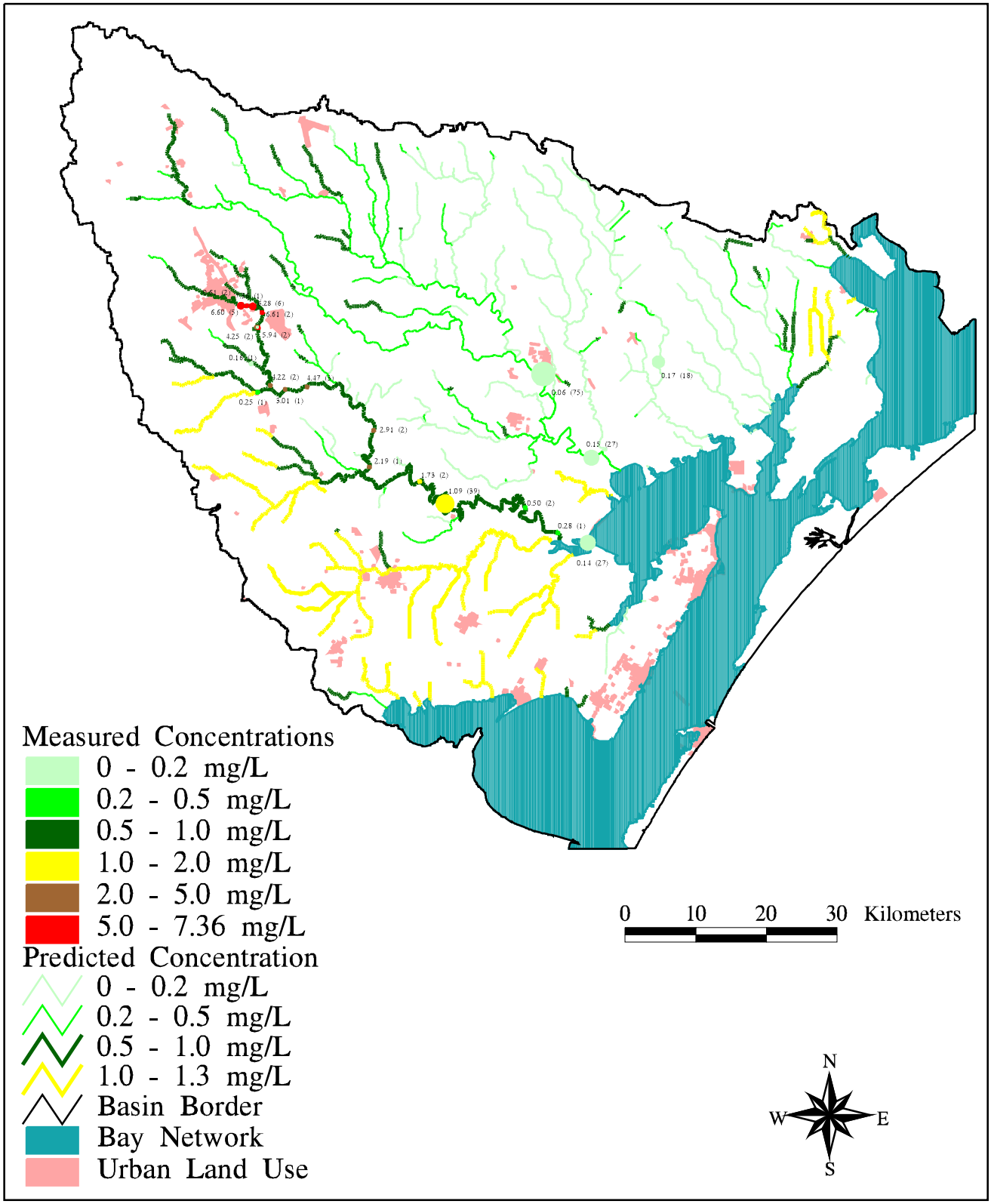


Figure 4.23 : Comparison of Estimated and Average Observed Total Phosphorus Concentrations in the San Antonio-Nueces Coastal Basin

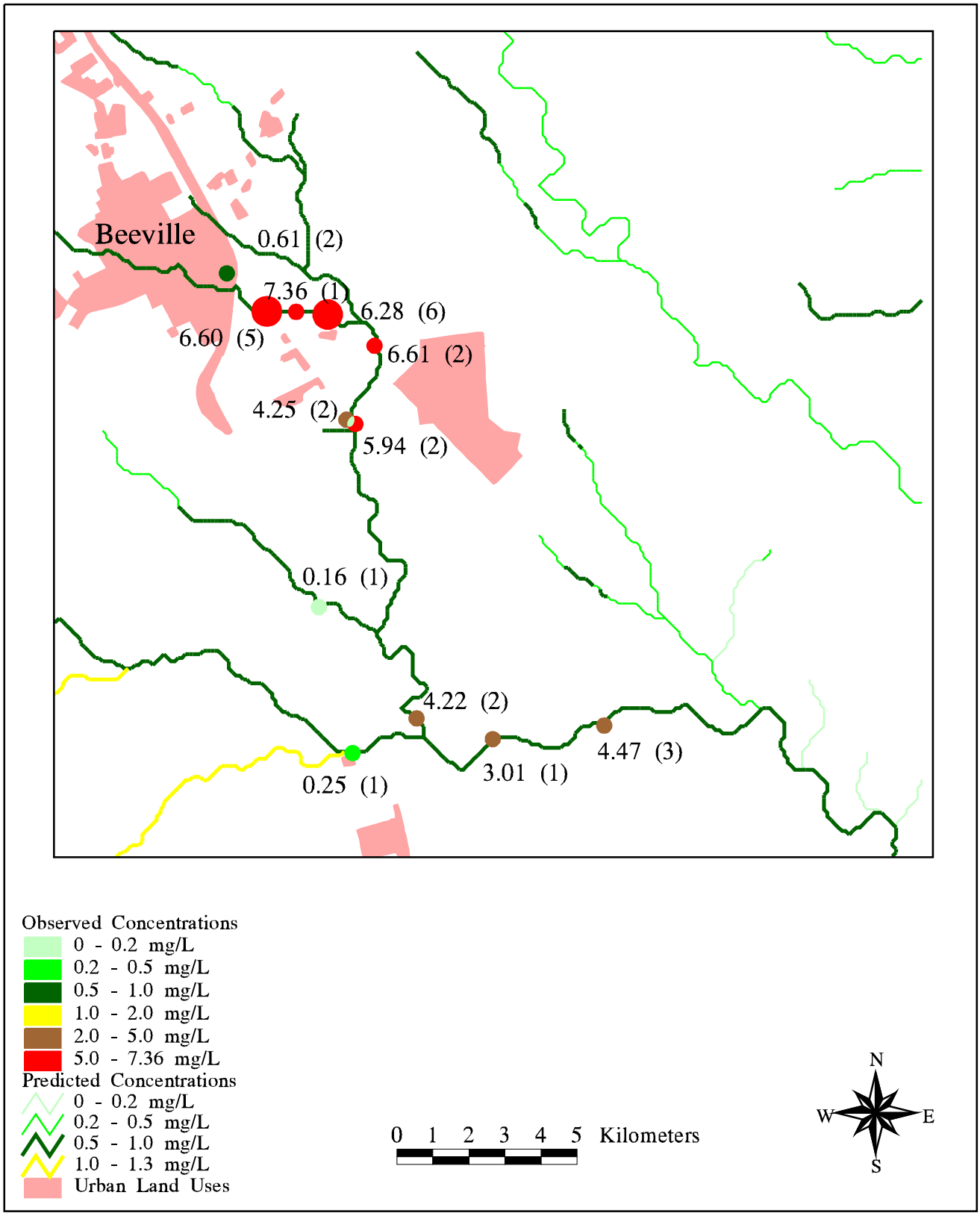


Figure 4.24 : Estimated vs. Observed Total Phosphorus Concentrations Just Downstream of Beeville, TX

These discrepancies would tend to indicate a significant point source in the area contributing to total phosphorus loads. Consultation with TNRCC personnel have identified that the data points in question were sampled to investigate suspected effluent problems from a wastewater treatment plant in Beeville. However, it should also be noted that most of these measurements were made within a short period in the early 1980's and it is not known whether total phosphorus at the sampling locations has remained at these elevated levels.

4.6 Considering and Simulating Point Sources

As can be seen from section 4.5, the characterization of nonpoint source pollution for a particular region may not provide a complete representation of the pollutant levels in that area. Point sources along stream networks can contribute significantly to the measured pollutant levels. Pollutant level data for point sources in the San Antonio-Nueces Coastal Basin were unavailable at the time of this study. However, a method of simulating point sources is investigated by considering the difference between predicted nonpoint source pollution concentration levels and observed concentration levels at a specific location, and then accounting for the difference with a single point load at the location. The point source pollutant load is then included in every downstream location in the digital basin.

Estimating an Annual Point Load

Figure 4.24 shows a number of measurement points just downstream of Beeville, TX where observed total phosphorus concentrations significantly exceed the values expected from nonpoint sources alone. Assuming that the Beeville wastewater treatment plant effluent enters the Aransas River at the furthest upstream location where a significant concentration discrepancy exists, a point source phosphorus contribution for that location is estimated to account for the discrepancy.

To establish the exact value of estimated nonpoint source total phosphorus concentration at the location, the phosconc grid is displayed in the Grid tool, overlaid with the phopts coverage, and queried at the suspected point source location, using

the Cellvalue command. Similarly, the annual cumulative runoff grid is displayed and queried to determine cumulative runoff at the point source location. By multiplying the cumulative runoff by the difference between observed and estimated concentrations, the amount of observed annual phosphorus load attributable to the point source is calculated.

Grid: **gridpaint phosconc value linear nowrap gray**
Grid: **polygonshades phopts 2**
Grid: **cellvalue phosconc ***
The cell containing point (1223830.414,693729.621) has value 0.621
Grid: **gridpaint runoffac value linear nowrap gray**
Grid: **polygonshades phopts 2**
Grid: **cellvalue runoffac ***
The cell containing point (1223830.414,693729.621) has value 5467914

Noting that the average observed total phosphorus concentration at the point source location is 6.6 mg/L, the amount of this concentration attributed to the point source effluent is calculated as 6.6 mg/L - 0.621 mg/L = 5.979 mg/L. By multiplying this value by the cumulative runoff at the point source, the total annual estimated cumulative phosphorus point load is determined as

$$5.979 \text{ mg/L} * 5{,}467{,}914 \text{ m}^3\text{/yr} * 1000 \text{ L/m}^3 * 10^{-6} \text{ kg/mg} = 32{,}694 \text{ kg/yr}. \quad (4\text{-}12)$$

This value for estimated load is compared with an algorithm from Thomann and Mueller (1987), where load is calculated as the product of daily per capita municipal flow, population of the municipality, and typical effluent concentration. For Beeville, using the population data from Table 1.1, and Thomann and Mueller's typical average values for per capita flow (125 gallons/capita-day) and total phosphorus municipal effluent concentration (7 mg/L), this algorithm results in an estimate of

$$125 \text{ gcd} * 13547 \text{ pop.} * 365 \text{ d/yr} * 3.785 \text{ L/gal} * 7 \text{ mg/L} * 10^{-6} \text{ kg/mg} = 16{,}376 \text{ kg/yr}. \quad (4\text{-}13)$$

According to the Beeville wastewater treatment plant chief operator, daily flow at the facility, averaged over the year, is approximately 2,000,000 gallons per day (Barrera, 1996). Using this value for flow, instead of Thomann and Mueller's typical daily per capita flow value, estimated total phosphorus load is calculated as

$$2{,}000{,}000 \text{ gal/d} * 365 \text{ d/yr} * 3.785 \text{ L/gal} * 7 \text{ mg/L} * 10^{-6} \text{ kg/mg} = 19{,}341 \text{ kg/yr}. \quad (4\text{-}14)$$

This value represents 58% of the value calculated in equation 4-12. The fact that these other estimates are within the same order of magnitude show that this method of estimating point loads has some validity. However, the other estimates also indicate that the additional phosphorus loads contributing to the measured concentrations at the Beeville location are probably not from the wastewater treatment plant alone.

Considering Point and Nonpoint Sources Together

In order to combine the point source load from equation 4-12 with the nonpoint source load, the point source load value is added to the cell where the observed concentration discrepancy exists. First, the flow accumulation grid is displayed and overlaid with the phosphorus measurement location point coverage. Through visual identification of the discrepant Beeville measurement location and use of the Selectpoint command, a single-cell grid representing the location is established. This grid has values of NODATA in all other cells. So that map algebra may be performed with this grid, the NODATA cells are converted to zero-value cells through use of the Isnull command and the Con statement. The annual point load value is simultaneously stored into the selected cell.

Grid: **gridpaint ditfac value linear nowrap gray**
Grid: **points phopts**
Grid: **beepoint = selectpoint(ditfac,*)**
Grid: **beeload = con(isnull(beepoint),0,32694)**

A new cell-based loading grid is established by adding the existing nonpoint source cell-based load grid (phosrnof) and the Beeville point load grid. However, since the Beeville point load grid is in units of kg/yr, it must first be converted to the aerial mg-mm/L-yr units of phornof. As shown in equation 4-15, this is accomplished by multiplying the point load grid by 100.

$$Q * \text{EMC (mg-mm/L-yr)} = \text{kg/yr} * 10^{6} \text{ mg/kg} * .0001 \text{ cells/m}^2 * .001 \text{ m}^3\text{/L} * 1000 \text{ mm/m} \quad (4\text{-}15)$$

A new total phosphorus load grid is created as the weighted flow accumulation of the new cell-based loading grid divided by 100, as per equation 4-6. The

phosphorus concentration grid is then recalculated as the new total phosphorus load grid divided by the accumulated runoff grid. A factor of 1000 included in this product produces concentration in units of mg/L, as per equation 4-9. As in section 4.5, a grid of concentration values specific to the basin stream network is established using the Con statement with the introfac grid. The stream concentration grid is multiplied by 1000 to retain significant figures, the product is truncated to create the integer grid, and the resulting grid is converted to a coverage, using the Streamline command. The mainland coverage is then used to clip the concentration coverage so that concentration streams end exactly at the shores of the bay network.

```
Grid:  beernof = phosrnof + (beeload * 100)
Grid:  totpload = flowaccumulation(mainfdr,beernof) / 100
Grid:  totpconc = totpload / runoffac * 1000
Grid:  tophostr = con(introfac >= 1,totpconc)
Grid:  topholin = streamline(int(tophostr * 1000),mainfdr,grid-code)
Arc:  clip topholin mainland tophocon line
```

Since the beeload point source pollutant grid only affects load values along the Aransas River, the only differences between this new concentration coverage and the one created in section 4.5 occur along the Aransas. Figure 4.25a shows the Beeville portion of the newly calculated concentration coverage with the observed concentration circles overlaid. Likewise, figures 4.25b and 4.25c show portions of the Aransas River between the Beeville area and the Copano Bay outlet. A review of the newly calculated concentrations in these three figures shows better agreement with the average observed concentrations along the length of the Aransas River. However, it should be re-emphasized that this new concentration coverage is derived with the assumption that the Beeville wastewater treatment plant effluent accounts for the difference between observed concentrations and estimated nonpoint source concentrations. In fact, there may be a number of point sources along the Aransas River that contribute to the total phosphorus concentration profile there.

For more accuracy, this method of simulating point sources should be implemented with values of reported annual loads or permitted average concentrations for all of the permitted point source effluents in the basin.

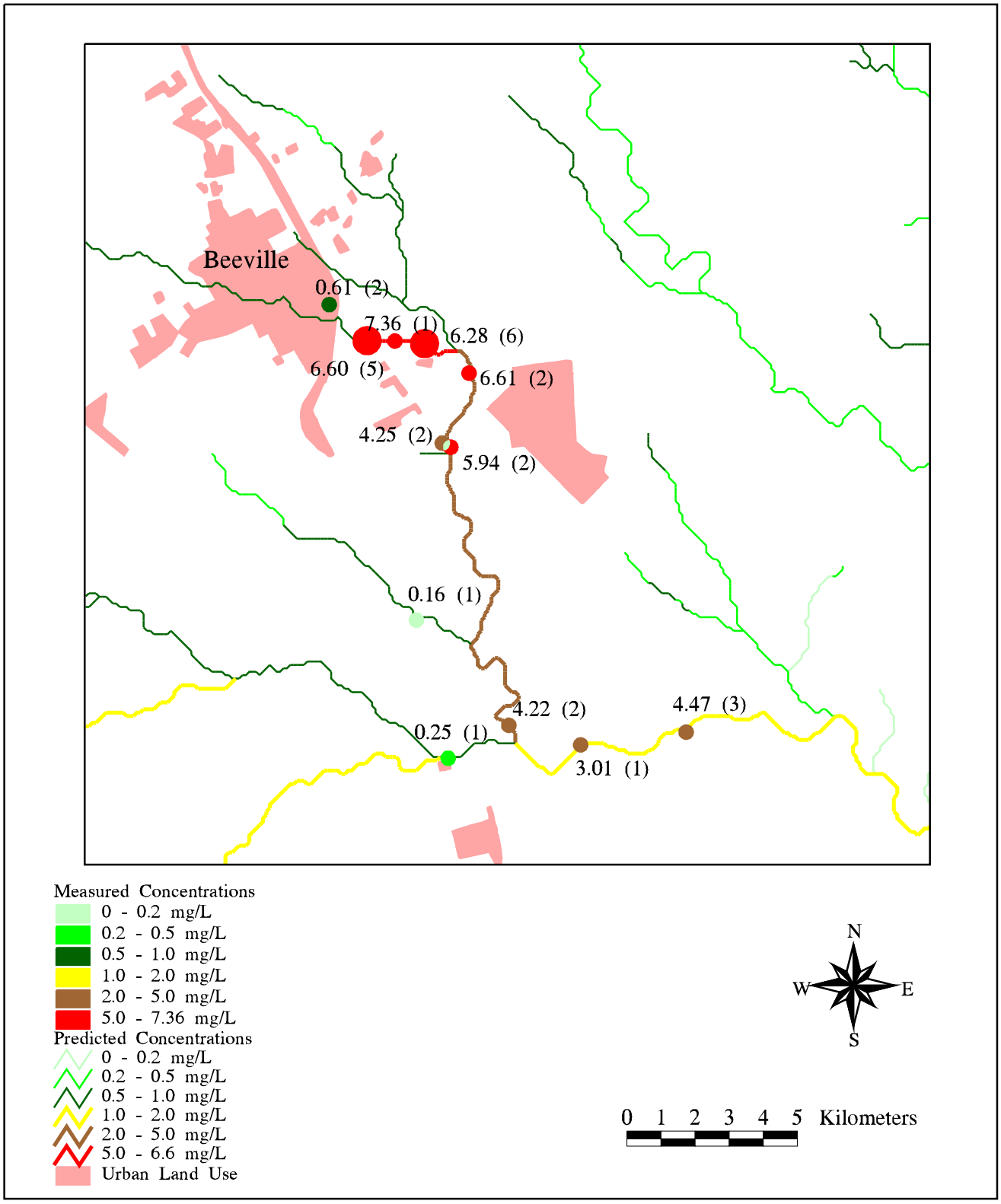


Figure 4.25a : Predicted vs. Measured Total Phosphorus Concentrations (Beeville Point Source Included) Just Downstream of Beeville, TX

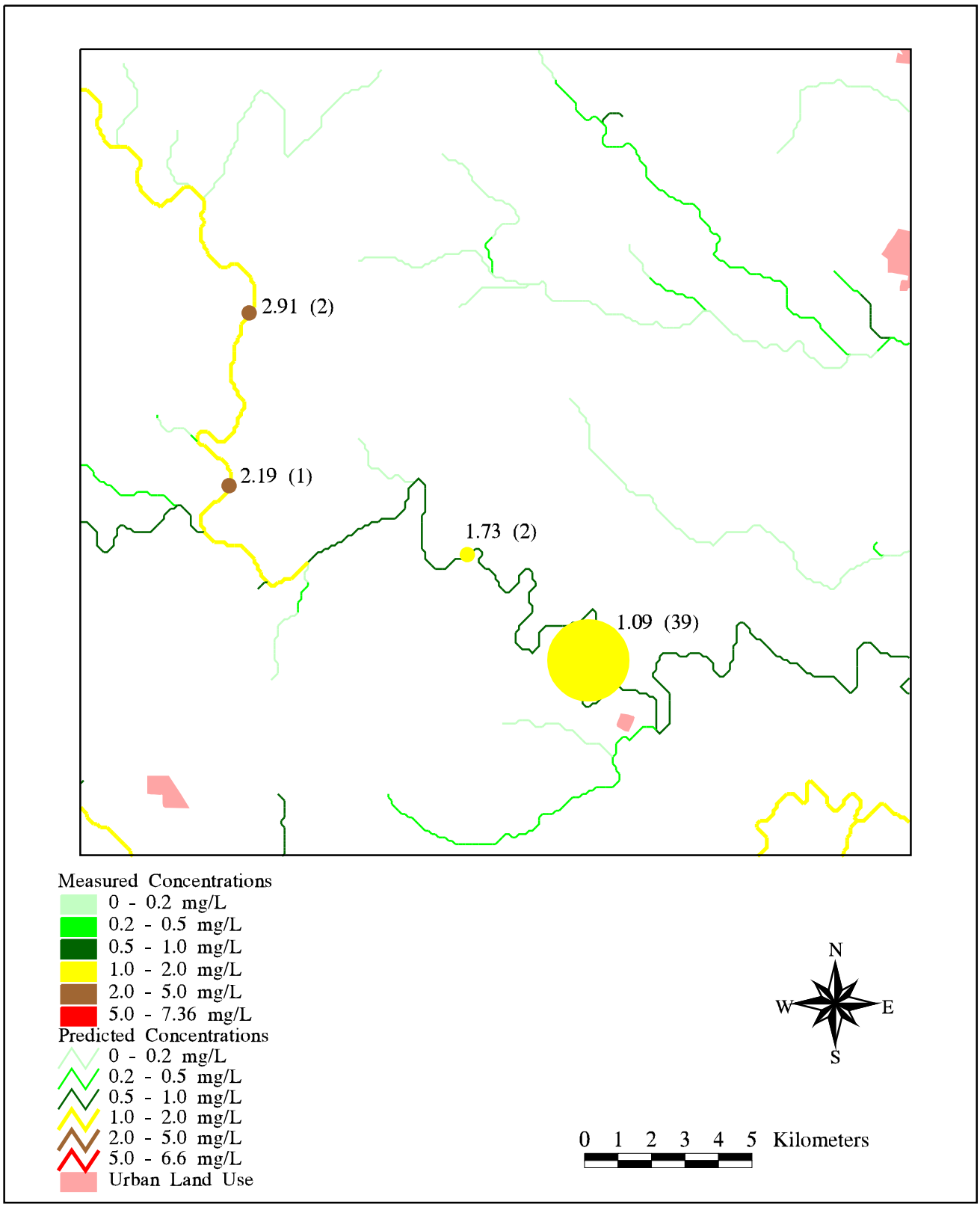


Figure 4.25b : Predicted vs. Measured Total Phosphorus Concentrations (Beeville Point Source Included) for Middle Aransas River

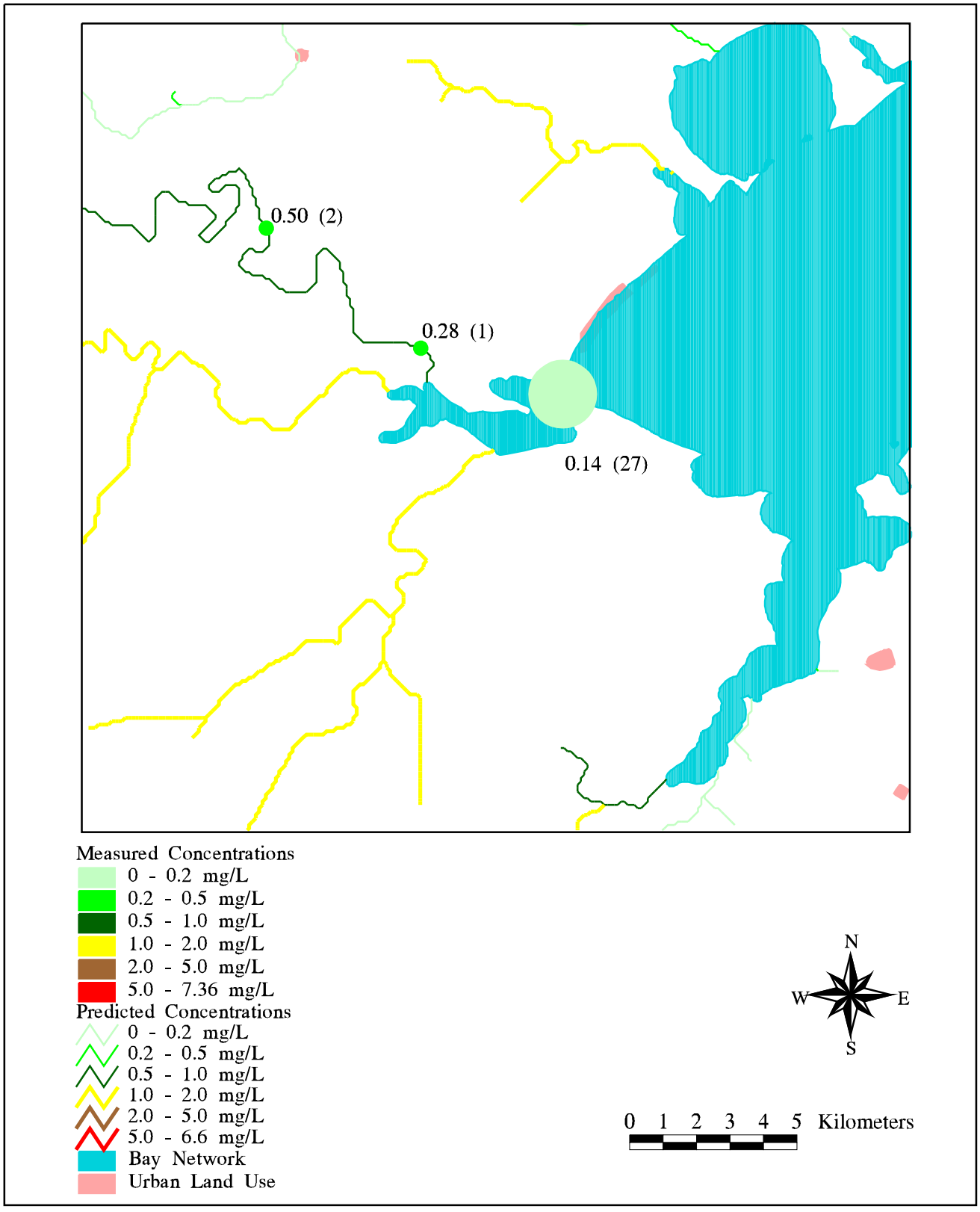


Figure 4.25c : Predicted vs. Measured Total Phosphorus Concentrations (Beeville Point Source Included) for Lower Aransas River

4.7 Using an Optimization Routine to Provide Estimates of EMC Values

The land use expected mean concentration values included in Table 3.6 are integral to this assessment of nonpoint source pollution. As outlined in section 3.2, these data are literature-based values used and published in a previous study (Baird, et al., 1996). Even though the agriculture and rangeland expected mean concentrations in this study were established empirically from measurements made near the San Antonio-Nueces coastal basin, it is desirable to establish a full set of expected mean concentration data that fits local conditions in the basin and does not necessarily rely on literature-based values.

One alternative method of determining expected mean concentration values for each land uses involves the use of a computer-based optimization routine. The input data required for this routine are (1) average observed pollutant concentrations at significant sampling locations, (2) all upstream pollutant point loads, (3) total annual cumulative runoff at the sampling locations, and (4) the annual cumulative runoff occurring from each land use upstream of each sampling location.

Determination of Optimization Routine Inputs

Average observed pollutant concentrations are established from the methods discussed in section 4.5 and upstream point load data should be acquired from reported or permitted values, as identified in section 4.6. However, for this analysis, the total phosphorus point load data estimated in section 4.6 is used.

Total annual cumulative runoff and land use-based cumulative runoff are established for the TNRCC sampling sites where significant numbers (more than 15) of historical phosphorus measurements exist. There are five such locations in the San Antonio-Nueces coastal basin; two along the Aransas River, two on the Mission River, and one on Copano Creek. Upon further review, one of these sampling locations, in Copano Bay a few kilometers east of the Aransas River outlet, is rejected since pollutant transport to the location does not follow a strict linear path along the stream network and is assumed to have a significant dispersion component.

Determination of total annual cumulative runoff is accomplished by displaying the cumulative runoff grid of the basin, overlaying the phosphorus sampling locations,

and querying the locations of significant phosphorus measurements. These steps are performed using the Gridpaint, Points, and Cellvalue commands. For a sampling site along the Aransas River, the procedure is as follows:

Grid: **gridpaint runoffac value linear nowrap gray**
Grid: **points phopts**
Grid: **cellvalue runoffac ***
The cell containing point (1252520.808,665484.913) has value 94664336.000

The cumulative runoff values for each land use upstream of a sampling location are determined by first delineating a subwatershed from the sampling site, using the Gridpaint, Points, Selectpoint, and Watershed commands along with the basin flow accumulation grid, flow direction grid, and sampling sites coverage. An equivalent polygon coverage of the subwatershed grid is created, using Gridpoly. The polygon coverage is then used to clip the basin land use coverage, so that only those land uses occurring upstream of the sampling location are retained.

Grid: **gridpaint ditfac value linear nowrap gray**
Grid: **points phopts**
Grid: **aranpt = selectpoint(ditfac,*)**
Grid: **arptarea = watershed(clipfdr,aranpt)**
Grid: **araptcov = gridpoly(arptarea)**
Arc: **clip sanlu araptcov aranlu poly**

The clipped land use coverage is converted back to a grid, using Polygrid. Cells in the land use grid are filled with land use category values (lusecat). Finally, cumulative runoff from each land use is established by using the Zonalsum command with the land use grid and the cell-based runoff grid. This command sums the grid cell values from a target grid (runoff) based on regions of equal value defined in a zone grid (land use category). The result of this Zonalsum is multiplied by 10, as per equation 4-1, in order to convert cumulative runoff to units of m^3/yr. The product is then converted to an integer grid, so that a value attribute table may be subsequently created for the grid.

Grid: **arlugrid = polygrid(aranlu,lusecat,#,#,100)**
Grid: **arrunoff = int(zonalsum(arlugrid,runoff) * 10)**

By listing the value attribute tables (vat) of the land use grid and the cumulative runoff grid, cumulative runoff values from each land use category in the subwatershed are established by matching the values from the two tables, based on the count of cells in each grid.

Grid: **list arlugrid.vat**

Record	VALUE	COUNT
1	11	1312
2	12	1229
3	13	6
4	14	437
5	16	25
6	17	30
7	20	65400
8	30	25711
9	40	35419
10	50	19
11	60	97
12	70	866

Grid: **list arrunoff.vat**

Record	VALUE	COUNT
1	6730	6
2	14420	19
3	15100	25
4	18450	30
5	117060	97
6	256980	437
7	752240	866
8	785360	1312
9	906050	1229
10	25114850	25711
11	25536140	35419
12	41141650	65400

Once this procedure is performed for each of the four significant sampling locations in the basin, mass balance equations are set up for each subwatershed. These mass balances equate the total measured load (total cumulative runoff at the sampling location multiplied by the observed concentration) with the sum of the loads from each land use and point source. The loads from each particular land use are denoted by taking the product of the cumulative runoff from that land use and an expected mean concentration variable associated with the land use. Known point sources upstream of

the sampling location are also included in the sum. Mathematically, the mass balance equation for each subwatershed is written as

$$C_o * Q_a = \sum_{i=1}^{n} (C_i * Q_i) + \sum_{j=1}^{m} Pt_j \qquad (4\text{-}16)$$

where C_o is the average observed concentration at the sampling location, Q_a is the total cumulative runoff at the sampling location, n is the number of subwatershed land uses, C_i is the expected mean concentration for each land use, Q_i is the cumulative runoff from each land use, m is the number of subwatershed point sources, and Pt_j is the load from each point source.

Execution of the Optimization Routine

The four mass balance equations are entered into the Microsoft Excel Solver optimization routine and solved simultaneously to establish the best fit values for the land use-based expected mean concentration variables. Initially, the optimization routine does not converge to a solution since, for the four subwatershed mass balance equations, a total of 12 expected mean concentration variables exist. In order to solve for 12 variables in four equations, additional constraints on the variables are introduced. These constraints are derived from observations about the literature-based event mean concentration data in Table 3.6 and are outlined below:

- All phosphorus EMC's are limited to within +/-50% of their initially entered value.
- No pollutant contribution is expected from water and wetland land uses (i.e. phosphorus EMC's for those land uses are set to 0)
- Phosphorus EMC's for mixed urban and other urban land uses are assumed to be equal to the linear average of the phosphorus EMC's for residential, commercial, industrial, and transportation land uses.

The constraints do provide some bounds for the solution of the 12 variables, but still do not amount to 12 unique equations. However, the solution is further constrained by entering the total phosphorus data from Table 3.6 as the initial set of values for the expected mean concentration variables. Unfortunately, this limits the

function of the routine to that of an adjustment algorithm, rather than an independent method of establishing expected mean concentration values.

In order to run the optimization routine, all terms from equation 4-16 are placed on one side of the equation and are divided by total cumulative runoff at the sampling location, Q_a. Mathematically, this manipulation appears as

$$C_o - \left[\sum_{i=1}^{n} (C_i * Q_i) + \sum_{j=1}^{m} Pt_j \right] / Q_a = CB, \qquad (4\text{-}17)$$

where CB is the concentration balance, which should equal zero when the appropriate values for the land use-based expected mean concentrations are entered.

The concentration balances for each subwatershed are established and optimized solutions for the land use expected mean concentrations are calculated in two different ways. First, the sum of the absolute values of the concentration balances for each subwatershed is minimized. This optimization produces the expected mean concentration values shown in the fourth column of Table 4.7. A second optimization of the land use expected mean concentrations is performed by minimizing the maximum absolute value of the concentration balances for each subwatershed. This optimization method results in the recalculated expected mean concentration values shown in the fifth column of Table 4.7. Both of these methods have the effect of minimizing each of the individual subwatershed concentration balance values.

Land Use	Land Use Code	EMC Values (mg/L)		
		From Table 3.6	Minimized Conc Bal SUM	Minimized Conc Bal MAX
Urban Residential	11	0.57	0.332	0.609
Urban Commercial	12	0.32	0.228	0.327
Urban Industrial	13	0.28	0.14	0.269
Urban Transportation	14	0.22	0.33	0.226
Mixed Urban	16	0.35	0.257	0.358
Other Urban	17	0.35	0.257	0.358
Agricultural	20	1.3	1.424	1.306
Range Land	30	0.005	0.0025	0.0047
Forest Land	40	0.005	0.0036	0.0035
Water	50	0	0	0
Wetlands	60	0	0	0
Barren Lands	70	0.12	0.18	0.123

Table 4.7 : Expected Mean Concentration Values Calculated Using the Microsoft Excel Solver Optimization Routine