3 DATA DESCRIPTION

This study uses raster and vector data sets that are publicly available from a variety of sources. Raster data sets have values stored in a uniform rectangular array and are typically referred to as grids. A digital elevation model is an example of a raster data set. Vector data sets include points, lines, and/or polygons and are typically referred to as coverages. A point coverage includes data represented by single coordinate values, such as locations of streamflow gauges. Line coverages, such as stream networks, are defined by series of points, with nodes specified as the starting and ending points of each line. Polygon coverages, such as watershed boundaries, are made up of connected sequences of lines. Vector data sets also have associated tables of values that describe the geographic features they represent. (Environmental Systems Research Institute, 1990).

Vector data layers can be converted into raster data layers (and vice versa) by using the conventions that a point may be represented as a single grid cell, a line may be represented as a string of grid cells, and a polygon may be represented as a zone of cells. The Arc/Info GIS supports the transformations between these raster and vector data sets.

3.1 Map Projection

A standard map projection is needed for any study where the superposition and spatial analysis of geographic data from different sources is performed. Spatial data sets are typically available at various map scales and in different coordinate systems. Arc/Info GIS allows for the successful adjoining of spatial data, even if the data are of different spatial scales, as long as that data have common datum and map projections. Arc/Info also allows for conversion from one map projection to another.

The Texas State Mapping System, which is sometimes referred to as the Shackleford State Mapping System, is defined using a Lambert conformal conic projection, which preserves shapes on a map. For this study, a map projection that preserves area, such as the Albers equal-area conic projection, is preferred to the Lambert projection because it simplifies computations of water and mass balances over a region (Snyder, 1987). Thus, a hybrid map projection is used for the study, called the Texas State Mapping System-Albers (TSMS-Albers) projection. A list of the TSMS-Albers projection parameters is shown in Table 3.1.

The North American Datum of 1983 (NAD83) uses the Geodetic Reference System of 1980 (GRS80) ellipsoid as a reference ellipsoid defining orientation relative to the geoid of Earth. The Texas State Mapping System projection uses this datum instead of the North American Datum of 1927 (NAD27), which uses the older Clarke (1866) ellipsoid as a reference (Snyder, 1987).

3.2 Establishing a Digital Database

The establishment of a watershed digital description involves the assembly of the data that is ultimately used for each of the subsequent steps of the assessment. Table 3.2 summarizes the data sources used in this study and provides Internet addresses for obtaining the data. Procedures for accessing this data can be obtained from the University of Texas at Austin GIS Hydrologic Modeling World Wide Web site at http://civil.ce.utexas.edu/prof/maidment/gishydro/.

This section describes each of the data sets and provides a discussion of how they are managed in order to extract the data specific to the San Antonio-Nueces Coastal Basin. A running narrative of the steps performed is provided along with the

Projection:	Albers
Datum:	NAD83
Units:	meters
Spheroid:	GRS1980
1st Standard Parallel:	27 25 0.00
2nd Standard Parallel:	34 55 0.00
Central Meridian	-100 0 0.00
Latitude of Origin:	31 10 0.00
False Easting (m):	1,000,000
False Northing (m):	1,000,000

Table 3.1 : Texas State Mapping System-Albers Projection Parameters

DATA SOURCE	INTERNET ADDRESS
Digital Elevation Models (DEMs)	http://sun1.cr.usgs.gov/eros-home.html
Hydrography Digital Line Graphs	http://sun1.cr.usgs.gov/eros-home.html
Hydrologic Unit Codes (HUCs)	http://h2o.er.usgs.gov/nsdi/wais/water/huc250.HTML
Land Use/Land Cover (LULC) Files	http://www.epa.gov/epahome/search.html
USGS Daily Discharge Values	http://txwww.cr.usgs.gov/cgi-bin/nwis1_server
USGS Stream Gauge Locations	http://txwww.cr.usgs.gov/cgi-bin/nwis1_server
Precipitation Grids	fsl.orst.edu (anonymous ftp site)
Expected Mean Concentration values	CCBNEP (not available via Internet)
Water Quality Measurement Data	tnris.twdb.state.tx.us (anonymous ftp site)

Table 3.2 : Internet Addresses for Data Sources

actual Arc/Info and UNIX commands. This format provides the reader insight into the specific steps performed and describes the theoretical bases for each procedure. In addition, some of the steps in this chapter are more efficiently performed via automated Arc Macro Language (AML) scripts. Where appropriate, these AMLs are referenced and included in Appendix B.

Hydrologic Unit Codes (HUCs)

Watersheds typically define the boundaries of a hydrologic study. Reasonable approximations of the drainage basin boundaries in the United States are available through the USGS 1:250,000-scale Hydrologic Unit Codes (HUCs). This data was created through digitization of a combination of 1:250,000-, 1:100,000-, and 1:2,000,000-scale Hydrologic Unit Maps, which divide the United States into 21 major hydrologic regions and further subdivide the regions into subregions, accounting units, and cataloging units. Each of these subdivisions are uniquely identified by two-digit fields contained within an eight-digit attribute code referred to as the Hydrologic Unit Code. The first two digits in the code identify water resources region; the first four digits identify subregion; the first six digits identify accounting unit; and the whole eight-digit code identifies the cataloging unit (Steeves and Nebert, 1994).

The Hydrologic Unit Codes are available on Internet from the USGS in an Albers equal area conical projection (see Table 3.2 for address). The San Antonio-Nueces Coastal Basin HUCs are not specifically required data for the nonpoint source pollution assessment, but they do provide a useful frame of reference for comparison with the digitally delineated versions of the basin and subwatersheds (see Figure 3.1).

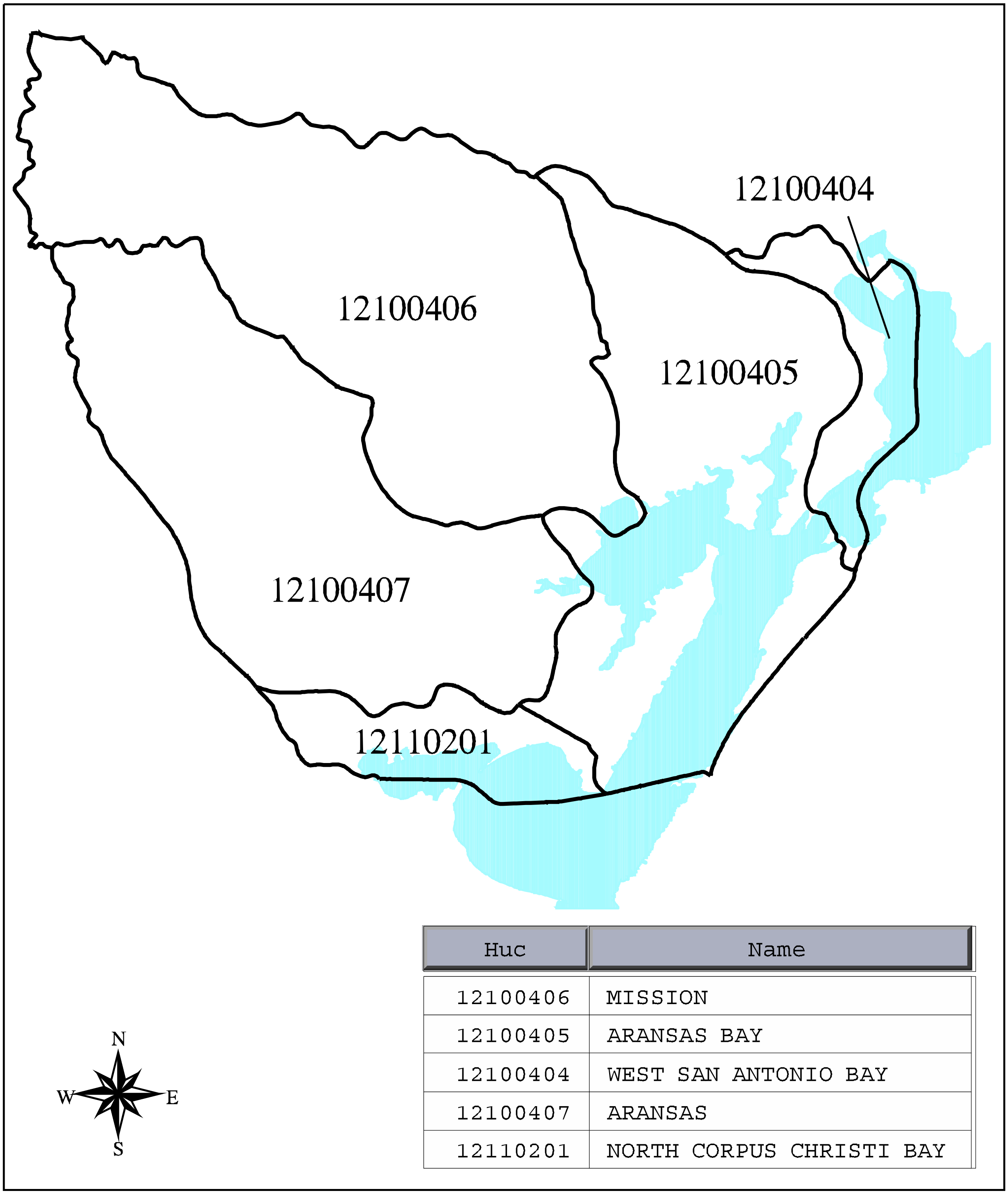
The Hydrologic Unit Codes start as Arc/Info interchange files (denoted by a file extension of .e00). A coverage is created from the interchange file through use of the Arc/Info Import command:

Arc: **import cover huc250.e00 huc250**

The huc250 coverage is displayed in the ArcView 2.0 program and the regional location of the San Antonio-Nueces coastal basin is magnified. Through use of the ArcView query builder, five polygons that approximate the basin are identified. Using ArcView Tables, the eight-digit Hydrologic Unit Code for each of the polygons is determined and recorded. Table 3.3 lists the five Hydrologic Unit Codes that approximate the San Antonio-Nueces coastal basin.

To create a new Hydrologic Unit Code coverage including only the five San Antonio-Nueces polygons, the Arc/Info Reselect command is invoked. Through use of the re-select and add-select features of the command, the HUCs with values between 12100404 and 12100407 are chosen and then appended with the code of 12110201. The new coverage is then converted into the desired TSMS-Albers projection and polygon topology is restored with the Arc/Info Clean command:

Arc: **reselect huc250 hucs**
>: **res huc >= 12100404**
>: **~** <return>
Do you wish to re-enter expression?(Y/N): **n**
Do you wish to enter another expression? (Y/N): **y**
>: **res huc >= 12100407**
>: **~**
Do you wish to re-enter expression?(Y/N): **n**
Do you wish to enter another expression? (Y/N): **y**
>: **asel huc = 12110201**
>: **~**
Do you wish to re-enter expression?(Y/N): **n**
Do you wish to enter another expression? (Y/N): **n**
5 features out of 2157 selected
Arc: **project cover hucs hucsan alb-tsms.prj**
Arc: **clean hucsan sanhucs**



Huc	Name
12100406	MISSION
12100405	ARANSAS BAY
12100404	WEST SAN ANTONIO BAY
12100407	ARANSAS
12110201	NORTH CORPUS CHRISTI BAY

Figure 3.1 : Hydrologic Unit Codes of the
San Antonio-Nueces Coastal Basin

Hydrologic Unit Code	Name
12100404	West San Antonio Bay
12100405	Aransas Bay
12100406	Mission River
12100407	Aransas River
12110201	North Corpus Christi Bay

Table 3.3 : Hydrologic Unit Codes Approximating the San Antonio-Nueces Coastal Basin

The above Project command makes use of a projection file (alb-tsms.prj) that specifies conversion from the national Albers projection to TSMS-Albers parameters. This projection file is included in Appendix B. The sanhucs coverage provides an initial approximation of the San Antonio-Nueces coastal basin boundaries. The sanhucs polygons and corresponding Hydrologic Unit Code values are displayed in Figure 3.1.

Hydrography Digital Line Graphs (DLGs)

The 1:100,000-scale Hydrography Digital Line Graph (DLG) data files are derived from USGS 30 x 60 minute quadrangle topographic maps and include stream networks, standing water, and coastlines as hydrographic features. These graphs are distributed in groups of files that cover a 30 x 30 minute area (the east or west half of the 1:100,000-scale source map). Typically, each 30-minute area is represented by four 15-minute files. Thus, each 30 x 60 minute quadrangle is represented by eight 15-minute files (USGS, 1989).

The 1:100,000 digital line graphs are available in either standard or optional format. The standard format has a larger logical record length (144 bytes) than the optional format (80 bytes), but is projected in an internal file coordinate system (thousandths of a map inch) that is not as easy to work with as the Universal Transverse Mercator (UTM) projection of the optional format (USGS, 1989). For this reason, the optional format hydrography digital line graphs are used in the San Antonio-Nueces Coastal Basin study.

In addition to hydrography, the USGS distributes 1:100,000-scale digital line graphs for roads, rail lines, and pipelines. These are all available publicly via the Internet address in Table 3.2. Alternatively, digital line graph files for the United States are available (in optional format) from the USGS Earth Science Information Center in a 14-volume Compact Disc-Read Only Memory (CD-ROM) set. For this analysis, the required hydrography 15-minute files were accessed and downloaded from Volume 8 (Texas and Oklahoma) of the CD-ROM series (USGS, 1993).

The Hydrography files for Texas are located in the 100k_dlg directory of the USGS 1:100,000-Scale Digital Line Graph Data CD-ROM (USGS, 1993). This directory contains separate subdirectories for each of the 1:100,000-scale USGS mapsheets (60' x 30') in Texas and Oklahoma. By cross-referencing the 1:100,000-Scale Digital Line Graph Index Map at the USGS EROS Data Center Internet World Wide Web site (Table 3.2) with a map of delineated watershed boundaries in Texas (USGS, 1985), five 1:100,000-scale mapsheets that completely overlay the watershed are identified (Figure 3.2). These mapsheets are: Beeville, Goliad, San Antonio Bay, Corpus Christi, and Allyns' Bight.

The hydrography files from each of the five mapsheet subdirectories are copied from the CD-ROM into a local UNIX workspace:

```
$:  cp /cdrom/100k_dlg/beeville/be3hydro.zip ./
$:  cp /cdrom/100k_dlg/sananbay/be4hydro.zip ./
$:  cp /cdrom/100k_dlg/goliad/be1hydro.zip ./
$:  cp /cdrom/100k_dlg/corpus_c/cc1hydro.zip ./
$:  cp /cdrom/100k_dlg/allyns/cc2hydro.zip ./
```

Each of these files exist in a compressed (zipped) format. Uncompressing them creates eight 15-minute (1:62,500-scale) coverages, arranged in a 2-row by 4-column format. For example:

```
$:  unzip be3hydro.zip
Exploding:  be3hyf01
Exploding:  be3hyf02
        :          :
Exploding:  be3hyf08
```

Once all five hydrography files are unzipped, forty separate 15-minute map coverages exist. Through consultation with an atlas (USGS, 1970), the number of

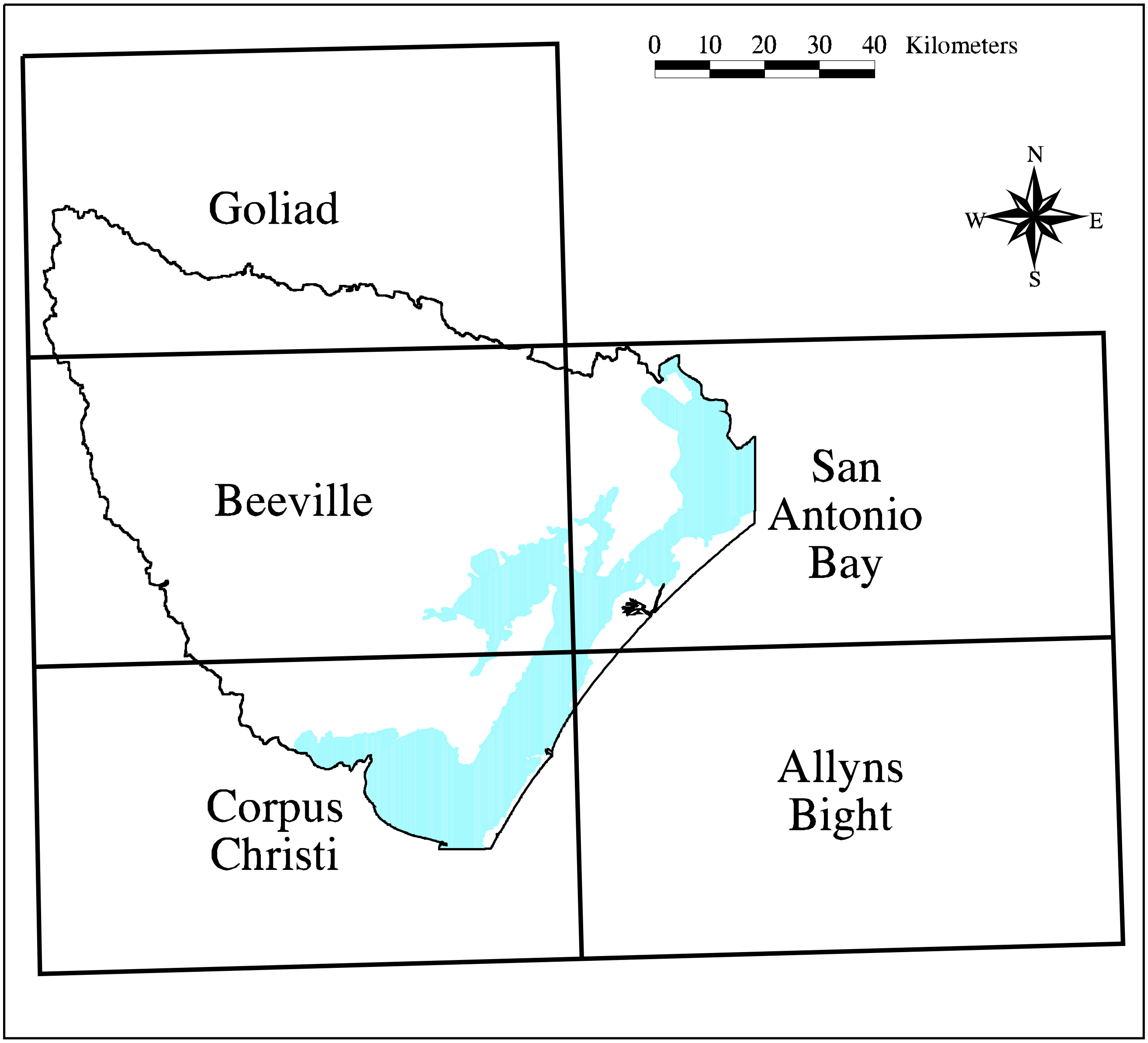


Figure 3.2 : USGS 1:100,000-Scale Mapsheets Covering the San Antonio-Nueces Coastal Basin

these coverages required to completely overlay the San Antonio-Nueces Coastal Basin is determined as eighteen. These map coverages include all 8 associated with the Beeville mapsheet, maps 5-8 from the Goliad mapsheet, maps 2-4 from the Corpus Christi mapsheet, maps 1 & 5 from San Antonio Bay, and map 1 from Allyns' Bight.

Before manipulation of the hydrography coverages can be performed, each of the 18 maps must be converted from its digital line graph format to an Arc/Info format. The Dlgarc command, with the optional format argument specified, is used for this purpose. Once converted, line topology is restored to each new Arc/Info coverage through application of the Build command. For example, conversion of the first Beeville 15-minute coverage is performed as:

Arc: **dlgarc optional be3hyf01 beef01**
Arc: **build beef01 line**

Each of the 15-minute hydrography coverages contain lines representing the streams, lakes, and coastlines associated with a particular map. A border around each coverage, representing 15-minute meridians and parallels, is also included. If all of these maps were merged together into a single coverage of the basin hydrography, the 15-minute meridians and parallels would be included. Alternatively, these border lines may be removed. This is performed by acknowledging that all arcs in a line coverage have a left polygon number and right polygon number field associated with them and that the value of the exterior polygon in a coverage is always defined as one. Using this information, the meridians and parallels can be trimmed away from each coverage through use of the Reselect command. Using the first Beeville 15-minute coverage as an example:

Arc: **reselect beef01 bee1 line # line**
>: **res rpoly# > 1**
>: **~**
Do you wish to re-enter expression?(Y/N): **n**
Do you wish to enter another expression? (Y/N): **y**
>: **res lpoly# > 1**
>: **~**
Do you wish to re-enter expression?(Y/N): **n**
Do you wish to enter another expression? (Y/N): **n**
187 features out of 240 selected

Once the meridian/parallel removal process is performed on all 18 hydrography coverages, they can be joined together using the Append command. Line topology is added with the Build command and the appended coverage is then converted from its initial Universal Transverse Mercator projection to TSMS-Albers parameters using the projection file, utmtsms.prj (included in Appendix B).

```
Arc:  append sanutm
Enter the 1st coverage:  bee1
Enter the 2nd coverage:  bee2
            :           :
Enter the 18th coverage:  allyn1
Enter the 19th coverage:   ~
Done entering coverage names (Y/N)?  y
Do you wish to use the above coverages (Y/N)?  y

Appending coverages.....
Arc:  build sanutm line
Arc:  project cover sanutm sanhydro utmtsms.prj
```

This procedure is much more efficiently performed using an AML. Dlgmerge.aml, is used to convert individual files from the 30' x 60' mapsheet subdirectories into a single coverage and is inlcuded in Appendix B. Figure 3.3 shows the final hydrography coverage, sanhydro, as clipped by a coverage of the basin boundary, which is created as per discussion in Chapter 4.

Digital Elevation Models (DEMs)

Three-arc second (3") digital elevation models (DEMs) are created by the Defense Mapping Agency by first digitizing cartographic maps ranging in scale from 1:24,000 to 1:250,000, and then processing elevation data from these digitized maps into a rectangular matrix format. The USGS distributes digital elevation models (via the Internet site noted in Table 3.2) in 1º x 1º blocks that correspond to either the eastern or western half of a USGS 1:250,000-scale map sheet. The models contain elevation data points at 3" intervals, or 20 elevation data points per minute. With 60 minutes per degree, each digital elevation model contains 1201 rows and 1201 columns of data (including the data points on the whole degree latitudes and longitudes, which are repeated in adjacent 1º x 1º grids) (USGS, 1990).

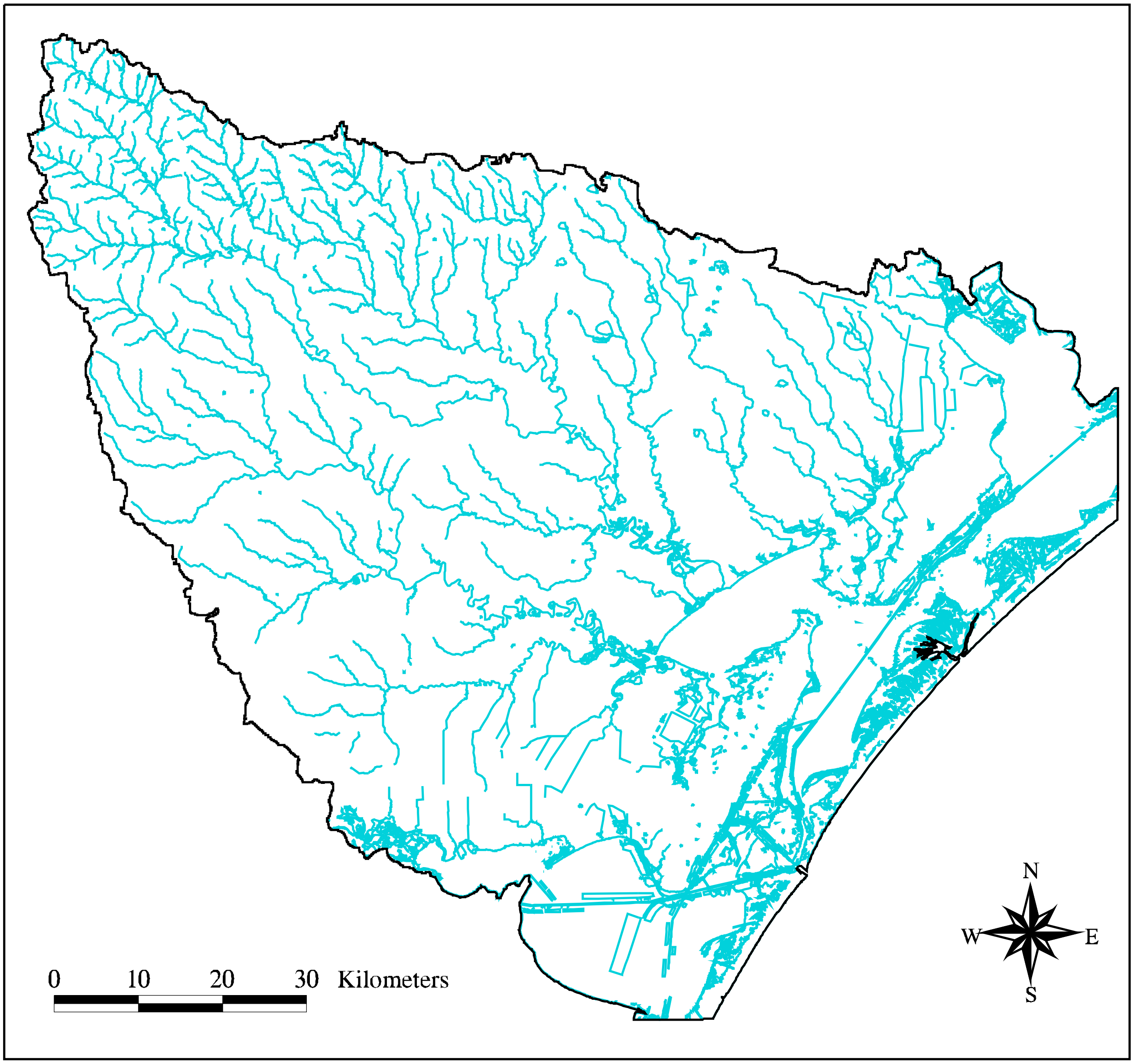


Figure 3.3 : 1-100,000-Scale Hydrography Digital Line Graphs of the San Antonio-Nueces Coastal Basin

Because the meridians of longitude converge at the poles, the latitudinal distance between 3” data points decreases as one moves north or south away from the equator. The distance along the surface of the earth at a specific radian of latitude (L_λ) can be calculated as $L_\lambda = R\cos\phi$, where R (Earth's radius) = 6371.2 km and ϕ = latitude. The distance between 3” elevation points at that latitude can then be calculated as $(L_\lambda * \pi/180^o)/1200$ (Reed and Maidment, 1995). For the San Antonio-Nueces Coastal Basin, which is bisected by the 28º North parallel, the latitudinal distance between elevation points is

$$[6371.2 \text{ m} * \cos(28^o) * \pi/180^o]/1200 = 81.8 \text{ meters} \quad (3\text{-}1)$$

and the longitudinal distance between points is

$$(6371.2 \text{ m} * \pi/180^o) /1200 = 92.67 \text{ meters.} \quad (3\text{-}2)$$

For use in a hydrologic analysis, digital elevation model data is first reprojected from geographic coordinates to a flat map coordinate system, in which horizontal dimensions can be measured in units of length and slopes can then be calculated by comparison with elevation values, also in units of length. When the digital elevation model is reprojected, a new grid is created by resampling the data at uniform intervals in the new projection. For example, a 3” x 3” geographic grid cell size is typically converted into a 100 m x 100 m flat map grid cell size.

Three arc-second (3”) digital elevation models are available via the US Geodata section of the USGS EROS Data Center Internet World Wide Web site specified in Table 3.2. Each 1° x 1° model is identified by the east or west half of a 1:250,000-scale Index Map. For the San Antonio-Nueces basin, four digital elevation models (Beeville East/West and Corpus Christi East/West) are required to completely cover the watershed.

When accessing compressed versions of the digital elevation models, the local UNIX file extension should be defined to show that the file is compressed (.gz). Compressed files can be uncompressed using the UNIX gunzip utility. These files must then have their record lengths modified to a format that Arc/Info can recognize. The UNIX dd command adds a carriage return at the end of every 1024 bytes. For example, these steps, as performed on the Beeville East digital elevation model appear as:

```
$:  gunzip beevillee.dem.gz
$:  dd if=beevillee.dem of=beeve cbs=1024 conv=unblock
```

where if = input file name, of = output file name, cbs = conversion buffer size, and "conv=unblock" specifies to allow for variable sized record lengths. Once these commands are performed for all four digital elevation models, the unblocked files can be converted into Arc/Info grids by using the Arc/Info Demlattice command:

```
Arc:  demlattice beeve beedeme usgs
```

This creates a grid called beedeme from the input digital elevation model beeve, which is specified as existing in a standard USGS format.

After the four four grids are created, they are combined into one large digital elevation model using the Arc/Info Grid Merge function. The large digital elevation model is then converted from its initial geographic projection into the desired TSMS-Albers using the projection file al72tsms.prj (included in Appendix B), and specifying a grid cell size of 100 m.

A smaller digital elevation model that contains just the area corresponding to the San Antonio-Nueces Coastal Basin is created by using the previously created sanhucs coverage. A five-kilometer buffer is first established around the sanhucs boundary through use of the Arc/Info Buffer command. Then the Grid Setwindow command is used to reduce the analysis window to the mapextent of the buffered sanhucs coverage. Once this analysis window has been reduced, a new digital elevation model (sndemalb) is defined that contains the values of the larger model within the analysis window.

```
Arc:  grid
Grid:  bcdem = merge(beedeme,beedemw,corpdeme,corpdemw)
Grid:  bcdemalb = project(bcdem,al72tsms.prj,#,100)
Grid:  quit
Arc:  buffer sanhucs hucbuff # # 5000
Arc:  grid
Grid:  setwindow hucbuff bcdemalb
Grid:  sndemalb = bcdemalb
```

Figure 3.4 shows the gray-shaded digital elevation model overlayed with the USGS Hydrologic Unit Codes, the major streams from the 1:100,000-scale hydrography digital line graphs, and a coverage of the intracoastal waterway features near the basin.

Land Use/Land Cover (LULC) Files

The 1:250,000-scale Land Use/Land Cover (LULC) data files are GIS polygon coverages and were created by the USGS through manual interpretation of aerial photographs acquired from NASA high-altitude missions in the late 1970's. Digitization of the land use maps resulted in the creation of the Geographic Information Retrieval Analysis System (GIRAS) (USGS, 1986). The land use files are available electronically from the USGS (conforming to an Universal Transverse Mercator projection) and the EPA (conforming to an Albers equal area projection). For this study, the land use files are downloaded from the EPA Internet World Wide Web site. Procedures for accessing this data can be obtained from the University of Texas at Austin GIS Hydrologic Modeling World Wide Web site at http://civil.ce.utexas.edu/prof/maidment/gishydro/.

The land use files employ the Anderson Land Use Classification System, which identifies two-digit subcategories within the categories of urban, agricultural, range, forest, water, wetland, barren, tundra, and snowfield land uses (Anderson et al., 1976). While widely available and frequently used, this data set is significantly dated and is considered out of date by many municipalities conducting urban assessments. However, this data set is still considered to be fairly accurate for the San Antonio-Nueces Coastal Basin, which is largely rural.

The land use files are organized and accessible by their associated 1:250,000-scale USGS mapsheet name. Starting at the EPA Internet site identified in Table 3.2, the user performs a query on "land use". This query results in the display of the EPA WAIS Gateway page, where the user selects the EPA EPAGIRAS (HTML) link. Finally, at the EPAGIRAS Data Sets page, the user performs queries on the 1:250,000-scale mapsheet names of interest. Only two land use files (corresponding to the Beeville and Corpus Christi mapsheets) are required to cover the San Antonio-Nueces Coastal Basin. These files are downloaded as compressed Arc/Info interchange files and have extensions of .e00.gz.

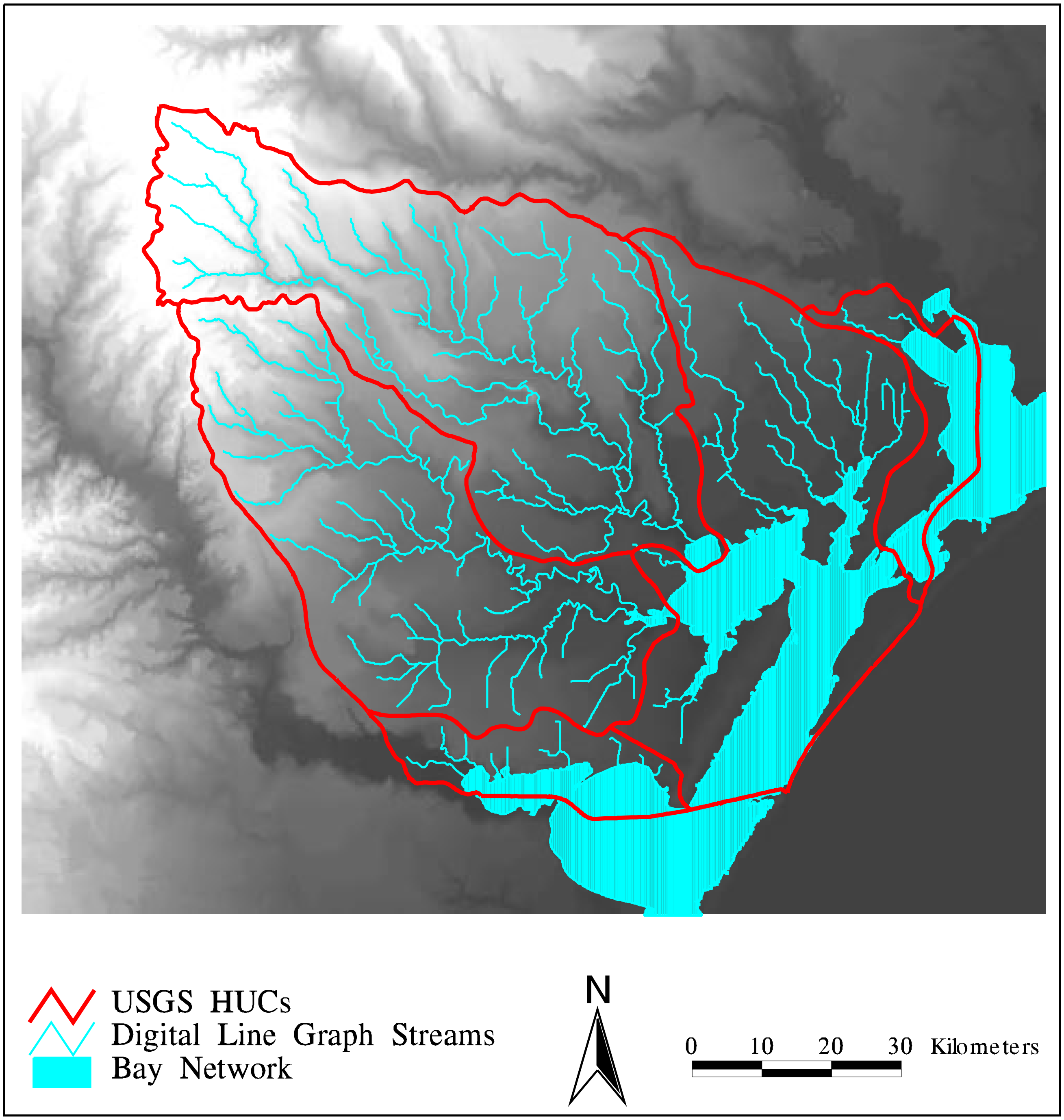


Figure 3.4 : Digital Elevation Model of the
San Antonio-Nueces Coastal Basin

The land use files are first uncompressed, imported, and cleaned as per the previous discussions. Using the Beeville land use file (lbe28096.e00.gz) as an example:

$: **gunzip lbe28096.e00.gz**
Arc: **import cover lbe28096.e00 lbe28096**
Arc: **clean lbe28096 beelu**

Once both land use coverages have been created, they are appended together and converted into the TSMS-Albers projection using the alb-tsms.prj file. The parallel line between the two mapsheets is removed using the Arc/Info Dissolve command. This command eliminates arcs between polygons that have the same value for a specified attribute, or "dissolve item". The attribute lanuse-id contains the value of the Anderson land use code for each polygon. By selecting lanuse-id as the dissolve item, any arcs between polygons of the same land use are eliminated.

Arc: **mapjoin landuse**
Enter the 1st coverage: **beelu**
Enter the 2nd coverage: **cclu**
Enter the 3rd coverage: **~**
Done entering coverage names (Y/N)? **y**
Do you wish to use the above coverages (Y/N)? **y**

Appending coverages.....
Arc: **project cover landuse lanuse alb-tsms.prj**
Arc: **dissolve lanuse luse lanuse-id poly**

Using ArcView 2.0 to inspect the luse coverage and selecting lanuse-id as the field through which to display shows that most of the polygons have values reflective of the Anderson land use codes. However, one polygon has a lanuse-id value of 200000. Upon further inspection in ArcView, this anomaly is identified as the lanuse-id for the Gulf of Mexico. By performing a Reselect on the luse coverage, the anomalous polygon is removed:

Arc: **reselect luse sanlus**
>: **res lanuse-id < 100**
>: **~**
Do you wish to re-enter expression?(Y/N): **n**
Do you wish to enter another expression? (Y/N): **n**
6513 features out of 6514 selected

Figure 3.5 shows the final land use coverage, sanlus, as clipped by a coverage of the basin boundary, which is created as per discussion in Chapter 4.

USGS Daily Discharge Values

Daily average discharge values (in units of cubic feet per second) are available for all active and inactive USGS streamflow gauges in Texas from the Texas Surface Water Database section of the USGS-Austin, TX World Wide Web site listed in Table 3.2. For the San Antonio-Nueces Coastal Basin, five streamflow gauges (three active, two inactive) exist. Table 3.4 identifies the periods of record for each gauge.

The discharge values recorded by each USGS gauge represent average streamflow at the gauge for that particular day. Daily, monthly, and annual streamflow volumes are calculated by processing the raw discharge data through the FORTRAN algorithm montflow.f (included in Appendix B).

USGS Stream Gauge Locations

Geographic locations (in degrees, minutes, and seconds) of the USGS streamflow gauges cited above are available from the same section of the USGS-Austin, TX World Wide Web site. Table 3.4 shows the latitudes and longitudes for each of the five San Antonio-Nueces coastal basin streamflow gauges.

In order to create a GIS coverage of these stations, the latitudes and longitudes are first converted into decimal degrees via the relationship,

$$DD = D + MIN/60 + SEC/3600 \qquad (3\text{-}3)$$

where DD = decimal degrees, D = degrees, MIN = minutes, and SEC = seconds. A raw data file of the digital coordinates (longitude listed first) is then built in a UNIX text editor window and named lonlat.dat. A copy of this raw data file, constructed by increasing USGS gauge number, is shown in Figure 3.6. Note that West longitude is treated as negative in decimal degrees.

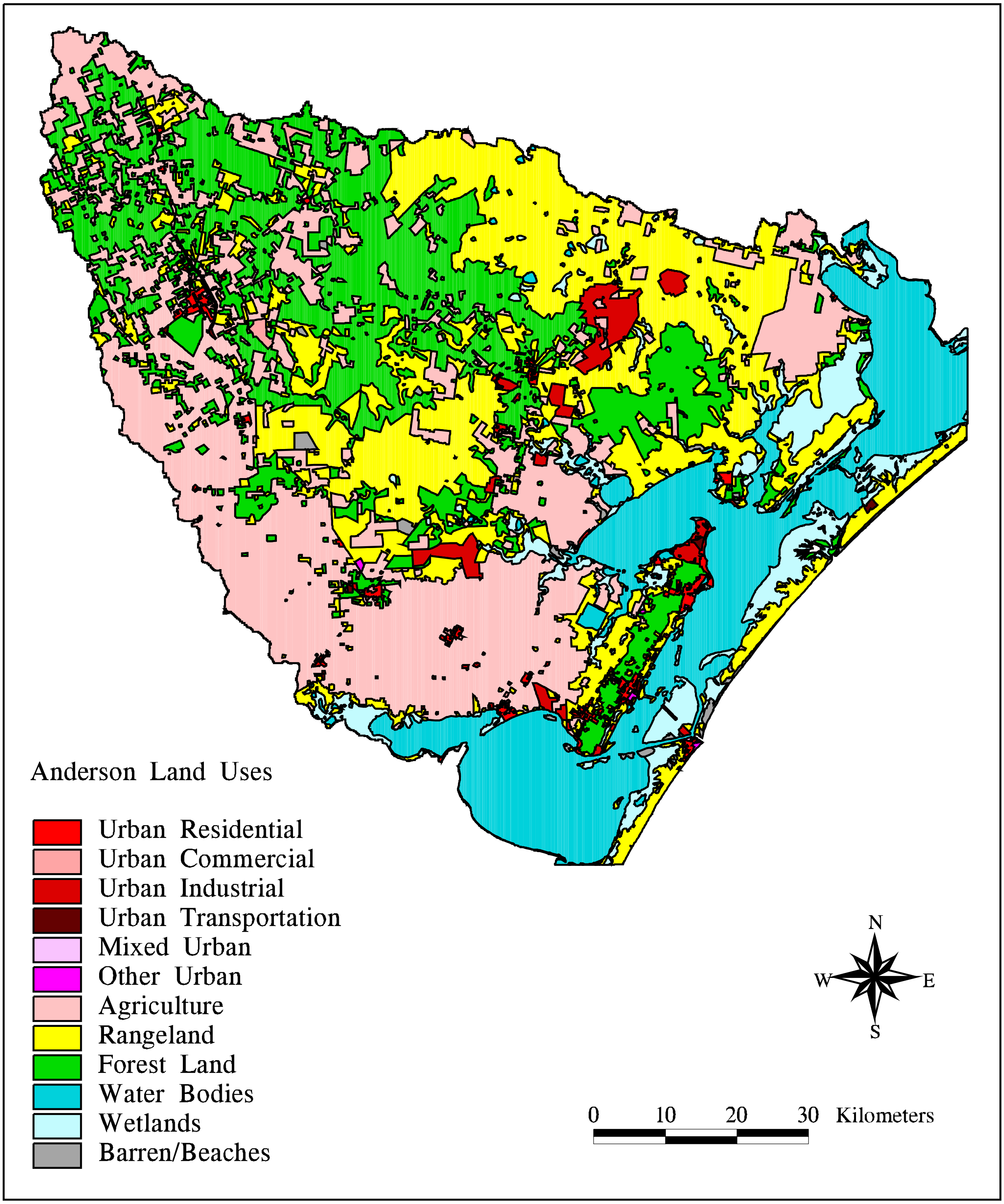


Figure 3.5 : Land Uses in the San Antonio-Nueces Coastal Basin

USGS Gauge	Gauge Description	Period of Operation	Latitude (N)	Longitude (W)
08189200	Copano Creek near Refugio, TX	6/17/1970 - present	28º 18' 12"	97º 06' 44"
08189300	Medio Creek near Beeville, TX	3/1/1962 - 10/17/1977	28º 28' 58"	97º 39' 23"
08189500	Mission River at Refugio, TX	7/1/1939 - present	28º 17' 30"	97º 16' 44"
08189700	Aransas River near Skidmore, TX	4/1/1964 - present	28º 16' 56"	97º 37' 14"
08189800	Chiltipin Creek at Sinton, TX	7/23/1970 - 4/6/1987, 8/4/1987 - 9/30/1991	28º 02' 48"	97º 30' 13"

Table 3.4 : USGS Streamflow Gauge Information

A point coverage of this digital coordinate data is built using the Arc/Info Generate command, specifying the lonlat.dat file as input and points as the geographic feature type. Once the coverage is created, point topology is established through the Build command and the digital coordinate values are added as attributes to each point by using the Addxy command:

Arc: **generate stations**
Generate: **input lonlat.dat**
Generate: **points**
    Creating points with coordinates loaded from lonlat.dat
Generate: **quit**
    Externalling BND and TIC...
Arc: **build stations points**
    Building points...
Arc: **addxy stations**

```
1    -97.1122    28.3033
2    -97.6564    28.4828
3    -97.2789    28.2917
4    -97.6206    28.2822
5    -97.5036    28.0467
end
```

Figure 3.6 : Digital Coordinate Data File for San Antonio-Nueces Stream Gauges

1	08189200	Copano
2	08189300	Medio
3	08189500	Mission
4	08189700	Aransas
5	08189800	Chiltipin
end		

Figure 3.7 : Gauge Number and Name Data File for San Antonio-Nueces Stream Gauges

A second data file, called statname.dat, is then created as per Figure 3.7. This file includes the gauge-id's and names listed in order. The shell of an attribute data file, called attribut.dat, is then built through use of the Arc/Info Tables function. Attribute field names and formats are defined for each of the items in the statname.dat file, making sure to define the first item, stations-id, to be in the same format as the stations-id field in the stations coverage. The data from statname.dat is used to fill in the formatted attribut.dat file, using the Tables "add from" command. The attribute data is then appended to the stations point attribute table (pat) through use of the Arc/Info Joinitem command. This command links data from two tables through the use of a common relate item. In this case, the station-id field is used as the relate item. Finally, the stream gauge coverage is converted from geographic to the required TSMS-Albers projection, using the geotsms.prj file:

Arc: **tables**
Enter Command: **define attribut.dat**
        1
Item Name: **stations-id**
Item Width: **4**
Item Output Width: **4**
Item Type: **i**
        5
Item Name: **stat-num**
Item Width: **10**
Item Output Width: **10**
Item Type: **c**
        15
Item Name: **stat-name**
Item Width: **15**

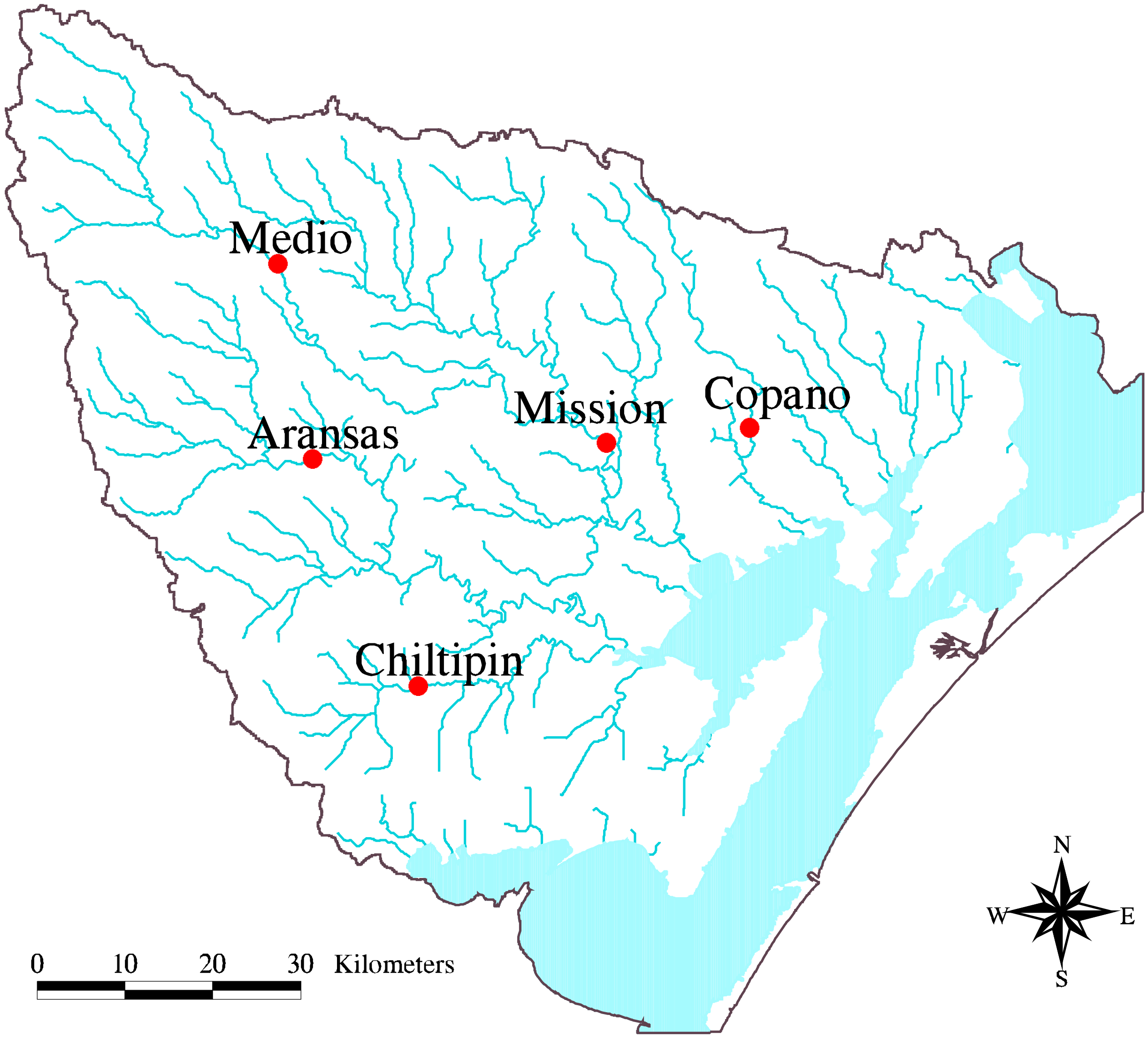
Item Output Width: **15**
Item Type: **c**
Item Name: ~
Enter Command: **add from statname.dat**
Enter Command: **quit**
Arc: **joinitem stations.pat attribut.dat stations.pat stations-id stations-id**
Arc: **project cover stations sangages geotsms.prj**

The resultant sangages coverage, shown in Figure 3.8, identifies the locations of each USGS stream gauge in the San Antonio-Nueces Coastal Basin and is used to define outlet points from which subwatersheds can be delineated for hydrologic analysis.

Precipitation Grids

Rainfall data typically provide a prime input to any nonpoint source pollution model. Much has been written about the importance of establishing definitive rainfall inputs for nonpoint source pollution load estimation. Collins and Dickey (1989) employed a stepwise least squares optimization procedure in the development of a stochastic model for simulating individual rainfall-runoff events and performing nonpoint source pollutant load assessments. Rudra et al. (1993) have identified that, for some nonpoint source pollution models that accept non-steady state rainfall inputs, variations in the selected rainfall time step interval can significantly affect estimates of runoff, sediment yield, and erosion characteristics.

This study considers precipitation as a steady state quantity averaged over an extended (30 year) time period. As a result, nonpoint source loads are also estimated as static quantities and concerns about temporal variations in rainfall inputs are somewhat mitigated. Precipitation data for the San Antonio-Nueces coastal basin is extracted from a set of grids developed at the Oregon State University Forestry Sciences Laboratory. These grids are part of the Parameter-elevation Regressions on Independent Slopes Model (PRISM) and cover the conterminous United States. PRISM is an analytical model that uses precipitation data measured at over 7000 National Weather Service and cooperator stations, 500 SNOTEL stations, and some selected State network stations (Daly et al., 1994).



Sangages-id	Stat-num	Stat-name	X-coord	Y-coord
1	08189200	Copano	-97.112	28.303
2	08189300	Medio	-97.656	28.483
3	08189500	Mission	-97.279	28.292
4	08189700	Aransas	-97.621	28.282
5	08189800	Chiltipin	-97.504	28.047

Figure 3.8 : USGS Stream Gauges in the San Antonio-Nueces Coastal Basin

GRASS Format		Arc/Info Format	
north:	50:01:15 N	ncols	1465
south:	24:03:45 N	nrows	623
east:	64:58:45 W	xllcorner	-126.020833333
west:	126:01:15 W	yllcorner	24.0625
rows:	623	cellsize	0.041666667
cols:	1465	nodata_value	-9

Table 3.5 : ASCII Header Formats for PRISM files in GRASS and Arc/Info

Estimated precipitation values are established for intermediate grid-cells through the use of a regression function, considering the measured precipitation point data along with digital elevation model data to account for orographic effects (Daly et al., 1994). The result of this process is a completely gridded surface of average precipitation across the nation. Average monthly (January-December) and average annual precipitation grids for the period between 1961 and 1990 are available.

The PRISM grids exist as compressed Geographical Resource Analysis Support System (GRASS) ASCII files at the ftp site noted in Table 3.2. For this study, only average annual precipitation data is required and is downloaded from the ftp site as the prism_us.ann.Z ASCII file. In order to uncompress the file, the file extension is changed from .Z to .gz and the gunzip utility is invoked:

```
$: mv prism_us.ann.Z  prism_us.ann.gz
$: gunzip prism_us.ann.gz
```

GRASS is a different GIS than Arc/Info, and there are some file format differences. The prism_us.ann ASCII file is compatible for immediate conversion to a GRASS GIS grid, but must have some modification to its' header before conversion to an Arc/Info grid. Table 3.5 shows the ASCII header formats that both GRASS and Arc/Info recognize. To create Arc/Info header information, (1) the nrows and ncols fields are directly transferrable from the GRASS rows and cols fields. (2) The xllcorner and yllcorner fields are just digital degree representations of the GRASS west and south fields. (3) Cellsize is calculated as the decimal degree difference between the GRASS east and west coordinates, divided by the number of columns. (4) Finally, nodata_value is specified as the value that GRASS uses to represent NODATA cells, -9 in this case.

Once the ASCII header information is modified from the GRASS format, the Arc/Info Asciigrid command is used to convert the ASCII file into an Arc/Info grid:

Arc: **asciigrid prism_us.ann p_ann**
Arc: **describe p_ann**

Description of Grid P_ANN

Cell Size =	0.042	Data Type:	Integer
Number of Rows =	623	Number of Values =	3470
Number of Columns =	1465	Attribute Data (bytes) =	8

BOUNDARY		STATISTICS	
Xmin =	-126.021	Minimum Value =	36.000
Xmax =	-64.979	Maximum Value =	6539.000
Ymin =	24.063	Mean =	771.181
Ymax =	50.021	Standard Deviation =	441.307

NO COORDINATE SYSTEM DEFINED

The Arc/Info Describe command is used to obtain projection and statistical information about the p_ann grid. This description shows that, while no coordinate system is defined for the grid, the X and Y boundary values are digital representations of the original GRASS coordinates, indicating that the grid is in a geographic projection with decimal degrees specified as the units of measure. For projection definition purposes, this information can be used, along with the datum and spheroid information (NAD83, GRS1980) of the TSMS-Albers projection.

In order to select the portion of the precipitation grid applicable to the San Antonio-Nueces Coastal Basin, a copy of the buffered Hydrologic Unit Code coverage (hucbuff) is first reprojected from TSMS-Albers to a Geographic coordinate system, using the tsmsgeo.prj file, included in Appendix B. The Grid Setwindow command is then used to reduce the analysis window to the mapextent of the new geobuff coverage. Once this analysis window has been reduced, a smaller precipitation grid (p_ann2) is defined that contains the values of p_ann within the analysis window. The smaller precipitation grid is then projected to the TSMS-Albers projection using the geotsms.prj file and specifying a grid cell size of 100 meters:

```
Arc:  project cover hucbuff geobuff tsmsgeo.prj
Arc:  grid
Grid:  setwindow geobuff  p_ann
Grid:  p_ann2 = p_ann
Grid:  rainbuff = project(p_ann2,geotsms.prj,#,100)
Grid:  rainbfcv = gridpoly(rainbuff)
```

A vector representation of the rainbuff grid is created using the Arc/Info Gridpoly command. When this command is invoked, each feature of the resulting coverage is assigned an attribute field called Grid-Code that contains the value of the corresponding grid cell. Figure 3.9 shows this precipitation coverage, as clipped by a coverage of the basin boundary, which is created as per discussion in Chapter 4.

Expected Mean Concentration Values

In order to calculate loadings of pollutants from each grid cell in the San Antonio-Nueces basin, pollutant concentration values need to be associated with the cells. Using literature-based expected mean concentration (EMC) values associated with land use is one way to spatially assign average pollutant concentrations. For this study, a set of expected mean concentration values used in a previous Corpus Christi Bay National Estuary Program analysis (Baird et al., 1996) was applied to the land uses in the basin. These expected mean concentrations were developed from water quality analyses performed at the Oso Creek and Seco Creek USGS stream gauges in south Texas. The Oso Creek stream gauge is located just west of Corpus Christi and represents the outlet of a predominantly agricultural subwatershed. The Seco Creek gauges are northwest of Hondo, Texas and represent drainage of rangeland (Baird et al., 1996). Expected mean concentration values for eighteen pollutants were used during this study and are included in Table 3.6.

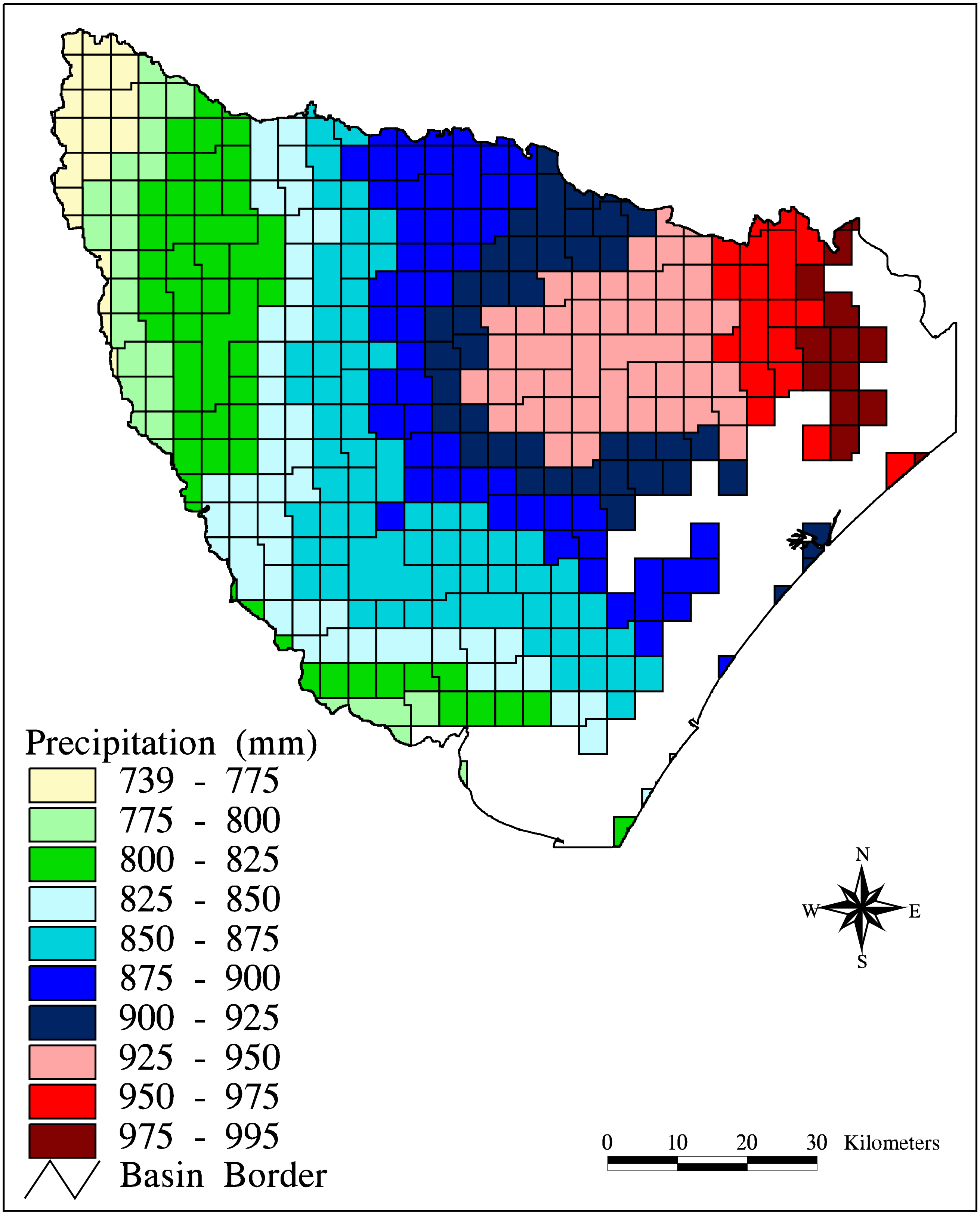


Figure 3.9 : Average Annual Precipitation in the San Antonio-Nueces Coastal Basin

Constituent	Urban Res 11	Urban Comm 12	Urban Ind 13	Urban Trans 14	Urban Mixed 16/17#	Agr 2*	Range 3*	Undev/ Open 7*
Total Nitrogen (mg/L)	1.82	1.34	1.26	1.86	1.57	4.4	0.7	1.5
Total Kjeldahl N. (mg/L)	1.5	1.1	1	1.5	1.25	1.7	0.2	0.96
Nitrate + Nitrite (mg/L as N)	0.23	0.26	0.3	0.56	0.34	1.6	0.4	0.54
Total Phosphorus (mg/L)	0.57	0.32	0.28	0.22	0.35	1.3	<0.01	0.12
Dissolved Phos (mg/L)	0.48	0.11	0.22	0.1	0.23			0.03
Suspended Solids (mg/L)	41	55.5	60.5	73.5	57.9	107	1	70
Dissolved Solids (mg/L)	134	185	116	194	157	1225	245	
Total Lead (ug/L)	9	13	15	11	12	1.5	5	1.52
Total Copper (ug/L)	15	14.5	15	11	13.9	1.5	<10	
Total Zinc (ug/L)	80	180	245	60	141	16	6	
Total Cadmium (ug/L)	0.75	0.96	2	<1	1.05	1	<1	
Total Chromium (ug/L)	2.1	10	7	3	5.5	<10	7.5	
Total Nickel (ug/L)	<10	11.8	8.3	4	7.3			
BOD (mg/L)	25.5	23	14	6.4	17.2	4	0.5	
COD (mg/L)	49.5	116	45.5	59	67.5			40
Oil and Grease (mg/L)**	1.7	9	3	0.4	3.5			
Fec Coliform (col./100 ml)**	20,000	6,900	9,700	53,000	22,400		200	
Fecal Strep (col./100 ml)**	56,000	18,000	6,100	26,000	26,525			

calculated as avg of land uses 11-14

* applied to all subcategories within the land use type

**average concentrations base on instantaneous rather than flow-averaged samples

Table 3.6 : Relationship Between Land Use and Expected Pollutant Concentrations

Water Quality Measurement Data

Once estimated average pollutant loads and concentrations have been established, they need to be compared with sampled data to validate the analysis. In support of this, a ten-year period (1982-1992) of water quality data measured in the region is used. This data set was previously used for the screening analysis portion of the 1994 Regional Assessment of Water Quality in the Nueces Coastal Basins (TNRCC, 1994) and was made available by the Texas Surface Water Quality Monitoring (SWQM) Program, managed by the Watershed Management Division of the TNRCC.

The Surface Water Quality Monitoring data available for the Nueces Coastal Basins (both San Antonio-Nueces and Nueces-Rio Grande basins) include 37 fixed monitoring stations measuring various combinations of 107 different water quality parameters. The parameters typically fall into three classes: (1) conventional parameters, such as pH, dissolved oxygen, and temperature, (2) nutrients (e.g. nitrogen and phosphorus), and (3) toxics (e.g. metals and pesticides). As the coordinating agency, TNRCC oversees and collects sampling data from other various Federal, State, and local agencies that perform the sampling (TNRCC, 1994).

The water quality data is provided, via the TNRCC ftp site identified in Table 3.2, as one compressed GIS point coverage identifying the sampling locations and two database (.dbf) files: one specifying each of the available water quality parameters in the EPA's standard STORET code format, and the other providing the actual time-tagged measurement values. Once the three files are accessed from the ftp site, the station location point coverage is imported and reprojected using the wqtsms.prj file in Appendix B. The .dbf files are converted to INFO files using the Dbaseinfo command;

Arc: **import cover snwqsites.e00 wqsites**
Arc: **project cover wqsites sanwq wqtsms.prj**
Arc: **build sanwq points**
Arc: **dbaseinfo value.dbf value**
Arc: **dbaseinfo storet.dbf storet**

Figure 3.10 shows the TNRCC water quality measurement points in the San Antonio-Nueces Coastal Basin.

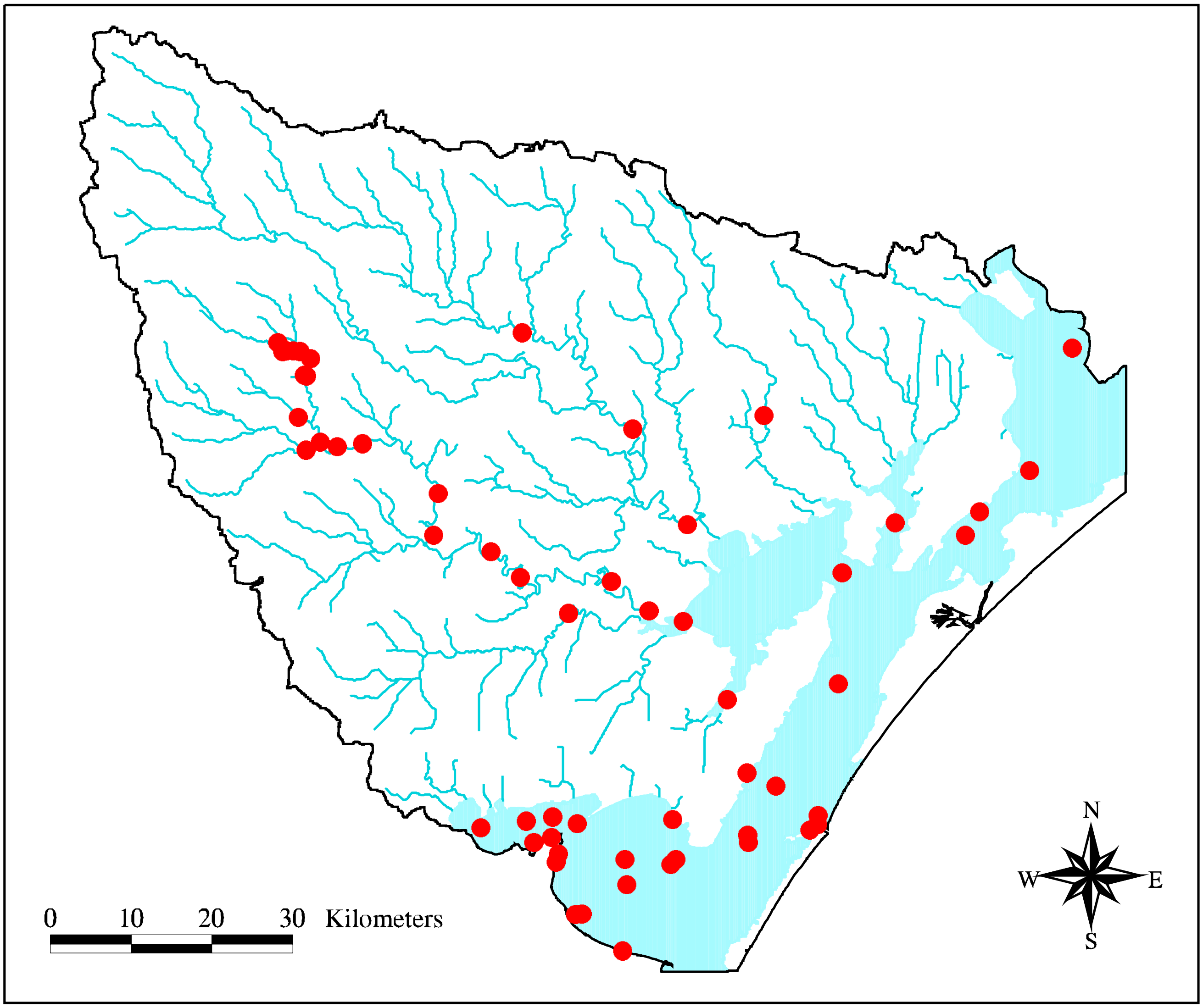


Figure 3.10 : TNRCC Water Quality Measurement Locations in the San Antonio-Nueces Coastal Basin

In order to link specific concentration values from the value table to stations in the sanwq coverage, a common linkage item must be identifed between the value table and the point attribute table (pat) of the coverage. A review of the two tables shows that the sanwq-id field in the pat contains the same data as the station-id field in the value table. However, the two fields are in different formats and must be in a common format in order to be linkable. This problem is resolved by adding a station-id field to the pat of the coverage, filling in the field with values from the sanwq-id field, and then changing the format of the new station-id field from integer to character type, using the Arc/Info Tables Alter feature:

```
Arc:  additem sanwq.pat sanwq.pat station-id 5 5 i
          Adding station-id to sanwq.pat to produce sanwq.pat.
Arc:  tables
Enter Command:  sel sanwq.pat
          105 Records Selected
Enter Command:  calc station-id = wqsites-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
Alternate item name:  ~
COLUMN  ITEM NAME   WIDTH  OUTPUT  TYPE  N.DEC  ALTERNATE NAME
    17     STATION-ID       5           5          C        -
Enter item name:  ~
```

Using ArcView 2.0, the sanwq point attribute table and the value table are linked through their station-id fields and the storet table is linked to the value table through their respective param-id and storetcode fields. Figure 3.11 shows portions of the three linked tables and demonstrates how selection of a pollutant constituent in the storet table identifies the sanwq locations where that pollutant is measured and the values of those concentration measurements in the value table.

Param_id	Long_desc	Short1_des	Short2_des
01034	CHROMIUM, TOTAL (UG/L AS CR)	CHROMIUM	CR,TOT
01040	COPPER, DISSOLVED (UG/L AS CU)	COPPER	CU,DISS
01042	COPPER, TOTAL (UG/L AS CU)	COPPER	CU,TOT
01043	COPPER IN BOTTOM DEPOSITS (MG/KG AS CU DRY WGT	CU MUD	DRY WGT
01045	IRON, TOTAL (UG/L AS FE)	IRON	TOTAL
01046	IRON, DISSOLVED (UG/L)	IRON	FE,DISS
01049	LEAD, DISSOLVED (UG/L AS PB)	LEAD	PB,DISS

storet.dbf

Station_i	On_seg_fl	Seg_id	Enddate	Tag	Storetcode	Gtlt	Value
12943	1	2001.0000	08/24/1983	Z195629	01042		16.0000
12944	1	2002.0000	02/10/1982	U009065	01042		10.0000
12944	1	2002.0000	05/04/1982	U009067	01042		3.0000
12944	1	2002.0000	08/31/1982	U009069	01042	<	1.0000
12944	1	2002.0000	11/19/1981	U009064	01042		5.0000
12945	1	2003.0000	08/24/1983	Z195890	01042	<	10.0000
12948	1	2004.0000	08/24/1983	Z196056	01042		13.0000
13074	1	2202.0000	05/31/1989	Z203336	01042		11.0000
13074	1	2202.0000	07/14/1986	Z203318	01042		4.0000
13074	1	2202.0000	07/20/1987	Z203325	01042		5.0000
13074	1	2202.0000	07/28/1988	Z203332	01042	<	20.0000
13074	1	2202.0000	08/07/1985	Z203309	01042	<	20.0000
13091	1	2203.0000	08/20/1982	Z203549	01042		11.0000
13093	1	2204.0000	08/19/1982	Z203616	01042		11.0000
13094	1	2204.0000	08/19/1982	Z203625	01042		18.0000
13096	1	2204.0000	08/19/1982	Z203642	01042		47.0000
13098	1	2204.0000	08/19/1982	Z203646	01042	<	10.0000
12931	0	2004.0000	01/04/1982	Z194514	00011		67.0000
12931	0	2004.0000	01/04/1982	Z194514	00300		7.4000
12931	0	2004.0000	01/04/1982	Z194514	00400		7.7000
12931	0	2004.0000	01/09/1989	Z194589	00011		61.0000

value.dbf

sanwq.pat

Shape	Area	Perimeter	Sanwq#	Sanwq-id	Station_i	Wqsites#
Point	0.000	0.000	14	12943	12943	14
Point	0.000	0.000	15	12944	12944	15
Point	0.000	0.000	16	12945	12945	16
Point	0.000	0.000	19	12948	12948	19
Point	0.000	0.000	35	13091	13091	35
Point	0.000	0.000	37	13093	13093	37
Point	0.000	0.000	38	13094	13094	38
Point	0.000	0.000	40	13096	13096	40
Point	0.000	0.000	42	13098	13098	42
Point	0.000	0.000	1	12930	12930	1
Point	0.000	0.000	2	12931	12931	2
Point	0.000	0.000	3	12932	12932	3
Point	0.000	0.000	4	12933	12933	4
Point	0.000	0.000	5	12934	12934	5
Point	0.000	0.000	6	12935	12935	6
Point	0.000	0.000	7	12936	12936	7
Point	0.000	0.000	8	12937	12937	8

Figure 3.11 : Linking Water Quality Measurement Data Files

3.3 Scales of Analysis

For this study, there are four spatial scales at which hydrologic and loadings analysis can be performed: (1) the 100 m digital elevation model grid cell (0.01 km^2 in area), (2) the PRISM 20 km^2 rainfall grid cell, (3) the subwatersheds defined by drainage area to the USGS streamflow gauges (average area = 650 km^2), and (4) the coastal basin (7235 km^2) taken as a whole. Figure 3.12 demonstrates the relationships between these scales of analysis.

Processes in this study are performed using the 100 m x 100 m (1 hectare) digital elevation model grid cell as the analysis unit. This is the only reasonable scale to use for the watershed modeling step, since an accurate replica of the stream network in the basin is required. Even at this scale, the resultant digital streams are all of a uniform 100 m width (or 141 m when flowing to diagonally adjacent cells).

For calculations performed using the PRISM rainfall data, each 20 km^2 cell is discretized into approximately 2000 grid cells corresponding to the digital elevation model cells. One may note, from Figure 3.12, that a number of the rainfall cells are irregular in shape. This is the result of (1) the reprojection of the grid from its initial geographic map projection and (2) the discretization process performed on each rainfall cell.

While the digital elevation model grid cell is used as the analysis unit for determination of loadings from each subwatershed, these loadings are also accumulated and reported on a subwatershed basis. Finally, the coastal basin scale is not used at all for this study. Coastal basins differ from river basins in that there are multiple outlets versus just one. For river basins, characteristic parameters such as runoff or load that are determined on a subwatershed basis can be lumped into single values associated with the outlet point of the basin. To perform the same accumulations for a coastal basin would leave the false impression that these quantities might be measurable at a specific point. For this reason, analysis on the coastal basin scale is avoided.

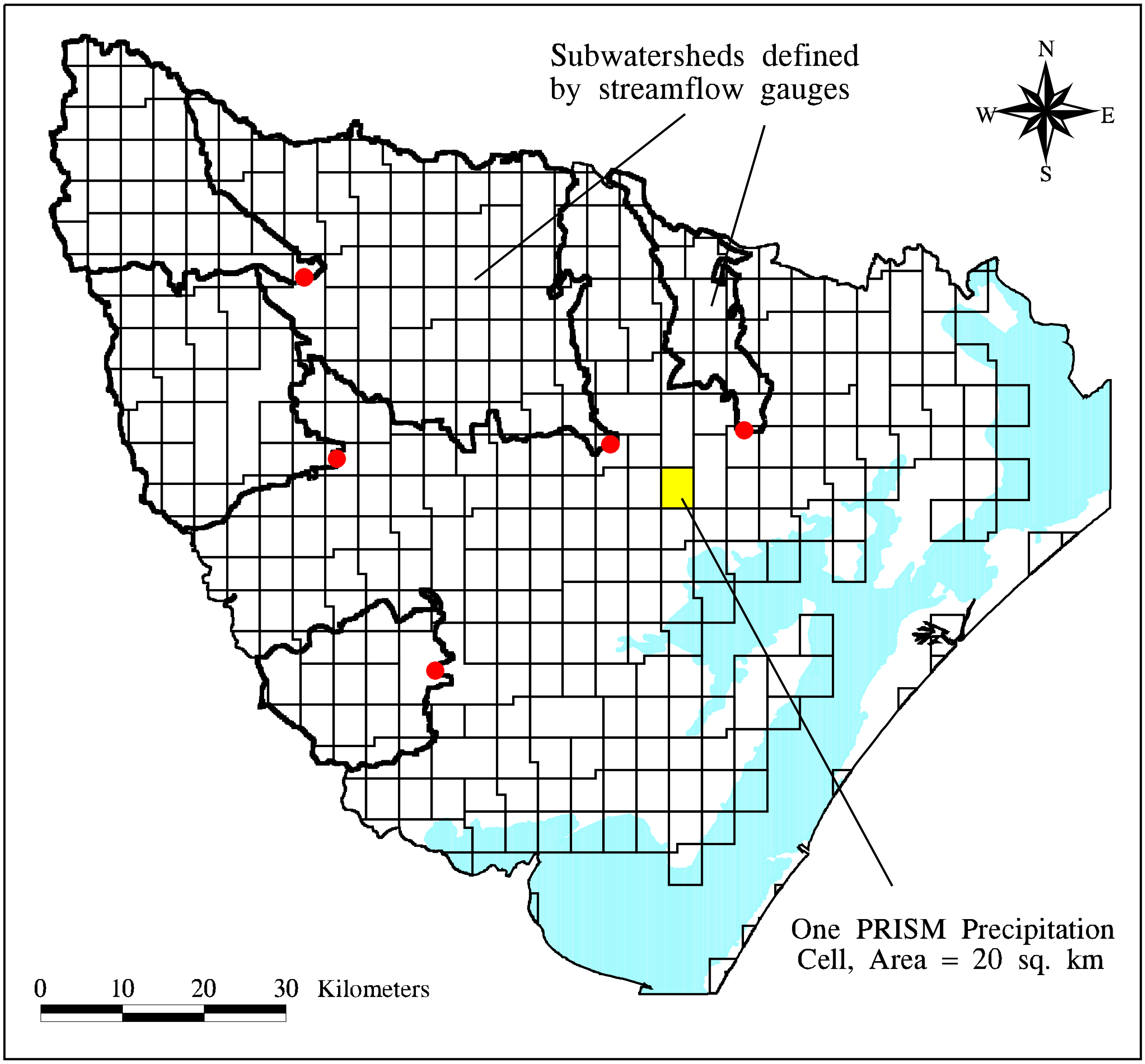


Figure 3.12 : Scales of Analysis