In cooperation with the U.S. Army Corps of Engineers

Design and Compilation of a Geodatabase of Existing Salinity Information for the Rio Grande Basin, from the Rio Arriba-Sandoval County Line, New Mexico, to Presidio, Texas, 2010

Data Series 499

U.S. Department of the Interior
U.S. Geological Survey

Cover: Upstream view of the Rio Grande near Ohara Road, Anthony, New Mexico, May 2010 (photograph by Julian D. Maltby).

Design and Compilation of a Geodatabase of Existing Salinity Information for the Rio Grande Basin, from the Rio Arriba-Sandoval County Line, New Mexico, to Presidio, Texas, 2010

By Sachin D. Shah and David R. Maltby II

In cooperation with the U.S. Army Corps of Engineers

Data Series 499

U.S. Department of the Interior
U.S. Geological Survey

U.S. Department of the Interior
KEN SALAZAR, Secretary

U.S. Geological Survey
Marcia K. McNutt, Director

U.S. Geological Survey, Reston, Virginia: 2010

This and other USGS information products are available at http://store.usgs.gov/
U.S. Geological Survey
Box 25286, Denver Federal Center
Denver, CO 80225

To learn about the USGS and its information products visit http://www.usgs.gov/
1-888-ASK-USGS

Suggested citation:
Shah, S.D., and Maltby, D.R., II, 2010, Design and compilation of a geodatabase of existing salinity information for the Rio Grande Basin, from the Rio Arriba-Sandoval County line, New Mexico, to Presidio, Texas, 2010: U.S. Geological Survey Data Series 499, 24 p.

Acknowledgments

The authors would like to thank Dale Doremus and Gregory Lewis of the New Mexico Interstate Stream Commission for providing valuable data and insight on the Rio Grande Basin; James Hogan of the University of Arizona for providing data; and members of the Rio Grande Salinity Management Coalition for their interest and involvement in developing the work plan for this project.

Contents

Figures

Tables

Conversion Factors, Datums, and Water-Quality Unit

Inch/Pound to SI

Multiply	By	To obtain
	Length	
mile (mi)	1.609	kilometer (km)
	Area	
square mile (mi^2)	2.590	square kilometer (km^2)

SI to Inch/Pound

Multiply	By	To obtain
	Length	
kilometer (km)	0.6214	mile (mi)
meter (m)	3.281	foot (ft)

Datums

Vertical coordinate information is referenced to North American Vertical Datum of 1988 (NAVD 88).

Horizontal coordinate information is referenced to North American Datum of 1983 (NAD 83).

Water-Quality Unit

Concentrations of chemical constituents in water are given in milligrams per liter (mg/L).

Design and Compilation of a Geodatabase of Existing Salinity Information for the Rio Grande Basin, from the Rio Arriba-Sandoval County Line, New Mexico, to Presidio, Texas, 2010

By Sachin D. Shah and David R. Maltby II

Abstract

The U.S. Geological Survey, in cooperation with the U.S. Army Corps of Engineers, compiled salinity-related water-quality data and information in a geodatabase containing more than 6,000 sampling sites. The geodatabase was designed as a tool for water-resource management and includes readily available digital data sources from the U.S. Geological Survey, U.S. Environmental Protection Agency, New Mexico Interstate Stream Commission, Sustainability of semi-Arid Hydrology and Riparian Areas, Paso del Norte Watershed Council, numerous other State and local databases, and selected databases maintained by the University of Arizona and New Mexico State University. Salinity information was compiled for an approximately 26,000-square-mile area of the Rio Grande Basin from the Rio Arriba-Sandoval County line, New Mexico, to Presidio, Texas. The geodatabase relates the spatial location of sampling sites with salinity-related water-quality data reported by multiple agencies. The sampling sites are stored in a geodatabase feature class; each site is linked by a relationship class to the corresponding sample and results stored in data tables.

Introduction

The Rio Grande Basin from the Rio Arriba-Sandoval County line, N. Mex., to Presidio, Tex. (fig. 1), an area of about 26,000 square miles, is characterized as arid or semi-arid and has numerous natural sources of dissolved solids (Benke and Cushing, 2005). High concentrations of dissolved solids (also expressed as salinity) in this part of the Rio Grande Basin have been noted for almost 100 years (Stabler, 1911). The problems associated with high salinity are of growing concern for water-resource managers as rapid urban growth in cities along the Rio Grande in the United States and Mexico causes increased water demand and changes urban, agricultural, and environmental conditions and water uses. For example, the quality of water in the Rio Grande is becoming increasingly important as more surface water is proposed for diversion from the river for potable and non-potable uses (Langman, 2009). Historically, high concentrations of salinity in this reach have been attributed to (1) reservoir evaporation, which increases the concentration of dissolved solids in the remaining water in storage; (2) displacement of shallow saline groundwater during irrigation, which subsequently contributes to irrigation return flow; (3) erosion and dissolution of natural mineral deposits containing high concentrations of dissolved solids ; and (4) inflow of deep saline or geothermal groundwater (Pillsbury, 1981; Allison and others, 1990, Moore and Anderholm, 2002; Philips and others, 2003).

The U.S. Geological Survey (USGS), in cooperation with the U.S. Army Corps of Engineers (USACE), compiled salinity-related water-quality information in a geodatabase for use in the development of a comprehensive baseline salinity budget for a part of the Rio Grande included in the overall study area of the geodatabase, the reach from San Acacia, N. Mex., to Fort Quitman, Tex. The geodatabase was designed as a tool for water-resource management and includes readily available digital data sources from the USGS, U.S. Environmental Protection Agency (USEPA), New Mexico Interstate Stream Commission (NMISC), Sustainability of semi-Arid Hydrology and Riparian Areas (SAHRA), Paso del Norte Watershed Council, numerous other State and local databases, and selected databases maintained by the University of Arizona and New Mexico State University. The geodatabase was designed so that any salinity-related water-quality data collected by an agency or university can be integrated into a single repository of salinity-associated data for the Rio Grande Basin. Despite many studies by numerous Federal, State, and local agencies and universities investigating salinity sources and possible mitigation strategies, water-quality data and the ancillary information associated with data collected during the course of these studies had never been compiled into a regional, comprehensive geodatabase for the Rio Grande Basin from the Rio Arriba-Sandoval County line, N. Mex., to Presidio, Tex.

Figure 1. Spatial extent of existing data for the Rio Grande Basin, from the Rio Arriba-Sandoval County line, New Mexico, to Presidio, Texas, 2010.

Purpose and Scope

The purpose of this report is to provide information on the design and compilation of a salinity geodatabase for the Rio Grande Basin from the Rio Arriba-Sandoval County line, N. Mex., to Presidio, Tex. Salinity-related data and ancillary information for more than 6,000 sampling sites were compiled from readily available sources in the study area. Following a brief description of the study area, the characteristics and function of the geodatabase are described, the methodology used to compile the water-quality components of the geodatabase is presented, and adiscussion of the compiled data is provided. The geodatabase is intended to provide detailed information regarding the primary sampling locations (referred to as sites) and the associated salinity data.

Description of the Study Area

Geospatial data were compiled for the part of the Rio Grande Basin extending from the Rio Arriba-Sandoval County line, N. Mex., to Presidio, Tex. In addition to compiling readily available surface-water-quality data, groundwater data were compiled for selected shallow alluvial and basin-fill aquifers located within the Rio Grande Rift Basin along an approximately 300-mile reach of the Rio Grande in the study area, from about 5 miles north of San Acacia, N. Mex., to about 6 miles south of Fort Quitman, Tex. (fig. 1).

The Rio Grande overlies six alluvial-fill basins in the study area—Socorro, San Marcial, Engle, Palomas, Mesilla, and Hueco (Wilkins, 1998). Detailed descriptions of the geologic structure of the alluvial-fill basins are given in Chapin (1971), Hawley (1978), Riecker (1979), Hawley and Kennedy (2004), Hawley and others (2005), and Hutchison (2006).

In the study area, the Rio Grande is generally a gaining stream; groundwater inflow contributes to streamflow throughout Rincon and Mesilla Valleys. The Rio Grande is the ultimate point of discharge for the regional flow system which includes both the shallow alluvial aquifer of Rincon and Mesilla Valleys and the deeper regional flow system (Wilson and others, 1981; Bexfield and Anderholm, 1997; Hibbs and others, 2003). However, during periods of drought, some reaches become losing streams (Wilson and others, 1981; Nickerson, 1995; Anderholm, 2002) when increased groundwater pumping causes drawdown and reversed gradients (Conover, 1954; Frenzel and others, 1992).

According to Phillips and others (2003), salinity concentrations are highest during the winter low-flow period when the majority of the observed flow in the Rio Grande consists of groundwater and municipal discharge. Low-flow conditions with high concentrations of salinity-related constituents limit the use of Rio Grande water for municipal supply and adversely affect agricultural and environmental uses. Phillips and others (2003) reported that salinity increased in a series of distinct "steps" from about 40 to about 2,000 milligrams per liter (mg/L) in a 1,200-kilometer stretch of the Rio Grande between its headwaters in southern Colorado and the United States-Mexico border region; some distinct steps of large increases in salinity were measured at San Acacia, N. Mex., Selden Canyon upstream from Las Cruces, N. Mex., and El Paso Narrows upstream from El Paso, Tex. (fig. 1).

Methodology

A geodatabase is a spatially enabled database that contains spatial and non-spatial information; it is an extension of tabular data that allows users to correlate data with physical and spatial components. With a geodatabase, geographically referenced data can be manipulated using a geographic information system (GIS) to produce maps, interactive queries, and various types of spatial analyses. A geodatabase provides a framework and an interactive tool to aid in understanding spatial trends in water-quality analysis. The various types of data are separated into relational tables in the geodatabase on the basis of how they interact and correspond with the spatial feature class. The relational tables represent a collection of features and the relations between them.

The geodatabase was built using ESRI ArcGIS and Microsoft© Access software. Multiple agencies and universities (listed in the "Introduction" section) maintain salinity-related data that were used to construct the geodatabase in that part of the Rio Grande Basin defined by the study area. The primary steps involved in developing the geospatial database were compiling data, entering data into the geodatabase, ensuring data quality, and documenting the associated metadata.

Geodatabase Design

The geodatabase was designed to be a single, comprehensive repository for salinity-related data and ancillary information for the study area maintained by numerous Federal, State, and local agencies and universities. Wellborn and Moreo (2007, p. 4) describes the elements of an efficient geodatabase design, elements used in the design of this geodatabase: "a thematic approach [is used] to create layers of feature data within a geographic information system (GIS). The various features are stored as relational tables in the geodatabase on the basis of how the features interact and correspond to one another. These features, tables, and relations represent real-world spatial, temporal, and descriptive attribute interactions (Zeiler, 1999)."

For this geodatabase, point feature classes represent sampling locations in the study area. Attribute information tables are used to store detailed information needed to link information from related tables of data. The geodatabase contains one point feature class, five attribute information tables, and two relationship classes (table 1). The feature class stores the spatial location of primary sampling sites. The information tables store records linked to the feature class and include the site table, sample table, result table,

Table 1. Description and definition of data compiled and entered into the geodatabase of water-quality information for the Rio Grande Basin, from the Rio Arriba-Sandoval County line, New Mexico, to Presidio, Texas, 2010.

[id, identifier]

Feature class or table name	Data type	Definition
site	Point feature class	Spatial locations of sampling sites. Attributes include site id, site description, type code original, latitude, longitude, horizontal datum, altitude, vertical datum, site source, comments, subtype code, and site alias.
sample	Attribute information table	Non-spatial site information from collecting agencies. Attributes include site id, date/time, top of sample, bottom of sample, aquifer code, and collecting agency.
result	Attribute information table	Non-spatial, salinity data. Attributes include sample id, distributing data source, parameter code, result value, and remark code, result description, units, and result comments. Nondetections are reported as the laboratory reporting level, when available, with a remark code value of less than (<).
param_code	Attribute information table	Non-spatial site information compiled from reported data. Used as a domain to describe the parameter code stored in the result table. Attributes include parameter code, description, common name, and parameter group.
collecting_agency	Attribute information table	Non-spatial data source information. Used as a domain to describe the data collector and provide metadata in the sample table. Attributes include data collecting agency and description.
data_source	Attribute information table	Non-spatial data source information. Used as a domain to describe the data source and provide metadata in the result table. Attributes include data source code and description.
site_sample	Relationship class	Connection between the primary site feature class and the site information table. Allows duplicate representations of a single site to be connected for display and analysis. The site id field in the primary site feature class is linked to the primary site id field in the sample information table.
sample_result	Relationship class	Connection between the sample information table and the result table. The sample id field in the sample information table is linked to the sample id field in the result information table.

param_codes (parameter list) table (hereinafter, parameter refers to water-quality physical properties and constituents), collecting-agency table, and data-source table (fig. 2, table 1). To connect data from the information tables to the geographic features contained in the feature class, relationship classes are created that use identical, unique identifiers in each data type to establish a relation between records in the geodatabase. The site table contains all site information compiled during sample collection as well as a primary site identification code so site locations that are common among the reporting agencies are coded to a single feature. The result table contains salinity-related water-quality data and documents the reporting data source at the results level. The parameter list table contains descriptions of all water-quality parameters, unique parameter codes, commonly used parameter names, and parameter groups. A parameter stored in the result table is automatically interpreted using a domain (Zeiler, 1999) based on the parameter list table, so that the parameter description is visible in tables and in spatial queries when viewed in ArcGIS. A domain definition includes the precise description of the set of values (objects) constituting the domain (McLeod, 1976). The use of a parameter code rather than the full parameter description in the result table increases storage, retrieval, and updating efficiency. The collecting-agency table and data-source table contain non-spatial data that provide detailed information about the agencies and universities that collected and provided the data. This information is also used as domains and provides record-level metadata in the result and sample tables.

Geodatabase Compilation

As a first step, environmental data were compiled from existing digital databases maintained by Federal, State, and local agencies and universities (table 2). The USGS, USEPA, NMISC, and SAHRA maintain extensive environmental databases that include salinity-related water-quality data; these databases provided the majority of the data compiled. USGS water-quality data were obtained from the National Water Information System (NWIS). USGS data include

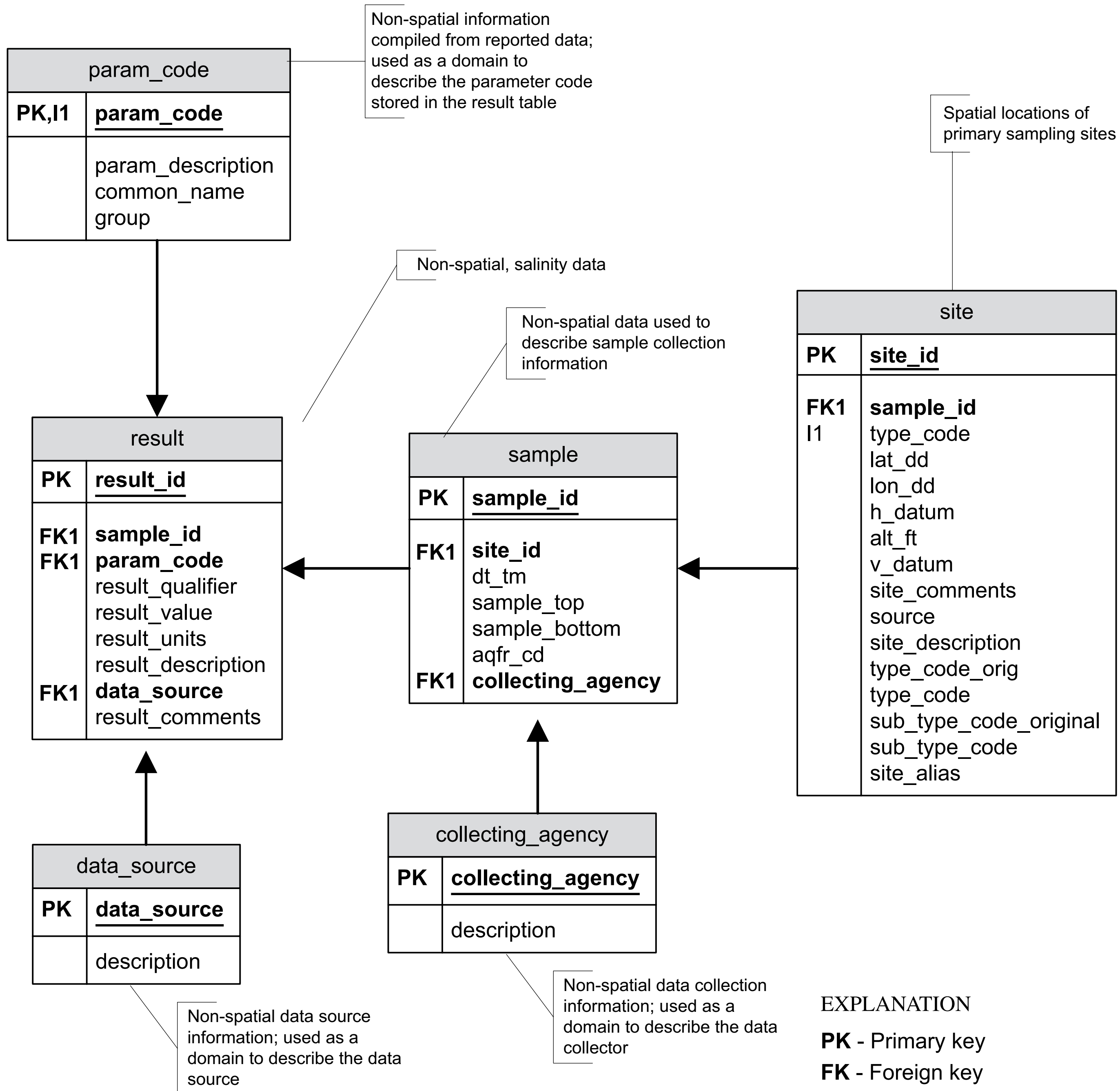

Figure 2. Diagram showing relational tables in the geodatabase of water-quality information for the Rio Grande Basin, from the Rio Arriba-Sandoval County line, New Mexico, to Presidio, Texas, 2010.

water-quality measurements taken as part of routine sampling and project-specific sampling in the New Mexico and Texas Water Science Centers (U.S. Geological Survey, 2009). USEPA data were obtained from the Modernized and Legacy Storage and Retrieval Repository (STORET) and include data supplied by State and local agencies (U.S. Environmental Protection Agency, 2009). NMISC data were obtained from and compiled by S.S. Papadopulos and Associates (written commun., 2003) from masters' theses, historical database compilations, reports compiled by the New Mexico Office of the State Engineer, and data downloaded from USEPA and USGS. SAHRA data were obtained as a database and include routine sampling data. Additional data sources (appendix 1) were included in the final database and are listed in table 2. Data compilation notes are listed in appendix 2.

Table 2. Sources of data compiled and entered into the geodatabase of water-quality information for the Rio Grande Basin, from the Rio Arriba-Sandoval County line, New Mexico, to Presidio, Texas.

Supplying entity or report	Sample count	Sample years
U.S. Geological Survey (USGS), National Water Information System		
New Mexico Water Science Center	6,416	1901 to 2009
Texas Water Science Center	1,321	1970 to 2008
U.S. Environmental Protection Agency (EPA), Modernized and Legacy Storage and Retrieval Repository (STORET)		
EPA STORET Modern data download	1,099	1937 to 2009
EPA STORET Legacy data download	8,299	1937 to 1998
New Mexico Interstate Stream Commission (NMISC), S.S. Papadopulos and Associates (SSPA) (written commun., 2003)		
Data compilation for Rio Grande Project Conveyance System 1996 (Boyle Engineering Corp./ Parsons Engineering Science, Inc., written commun., 2009)	5,693	1938 to 1995
State Engineers Office 1986 database; CEP – EPEC	20	1951 to 1972
State Engineers Office 1986 database; CEP – LNE – TX	1	1951
State Engineers Office 1986 database; City of El Paso, El Paso Water Utilities	89	1951 to 1980
State Engineers Office 1986 database; CEP, EPTL	3	1971
State Engineers Office 1986 database; CEP, USGS	31	1936 to 1984
State Engineers Office 1986 database; City of El Paso	9	1936 to 1972
State Engineers Office 1986 database; El Paso Water Utilities	102	1979 to 1984
State Engineers Office 1986 database; CEP–CURTIS	4	1953 to 1960
State Engineers Office 1986 database; CEP–US ARMY	1	1931
State Engineers Office 1986 database; EP CHEM-TEC	1	1981
Lower Rio Grande (LRG) compendium (12/2008)		
Data compilation for Rio Grande Project Conveyance System 1996 (Boyle Engineering Corp./ Parsons Engineering Science, Inc., written commun., 2009)	4,118	1964 to 1965
U.S. Geological Survey and Lower Rio Grande (USGS/LRG)	2,019	1970 to 1984
New Mexico Environment Department (NMED)	5,817	2003 to 2004
New Mexico Interstate Stream Commission (NMISC) Lower Rio Grande (LRG) Quarterly Monitoring	344	2007
New Mexico State University, M.S. Thesis, Jerry H. Williams		
"Salt Balance in the Rio Grande Project from San Marcial, New Mexico to Fort Quitman, Texas" (Williams, 2001)	8,213	1934 to 1963
University of Arizona, Department of Hydrology and Water Resources		
Sustainability of semi-Arid Hydrology and Riparian Areas (*SAHRA)* database *(http://www.sahra.arizona.edu/*)	1,082	2000 to 2006
New Mexico Environment Department (NMED)		
New Mexico Environment Department (NMED) and New Mexico Interstate Stream Commission (NMISC) Lower Rio Grande (LRG) Quarterly Monitoring	816	2005 to 2008
Wilson, Orr, White, and Roybal		
"Water Resources of the Rincon and Mesilla Valleys and adjacent areas, New Mexico" (Wilson and others, 1981)	660	1936 to 1977
Texas Commission on Environmental Quality (TCEQ)		
TCEQ data download (*http://www.tceq.state.tx.us/cgi-bin/compliance/monops/water_daily_summary.pl?cams=719*)	3,543	1968 to 1992
Texas Water Development Board (TWDB)		
TWDB data download (*http://www.twdb.state.tx.us/ publications/reports/GroundWaterReports/GWDatabaseReports/GWdatabaserpt.htm*)	3,928	1800? to 2007
Daniel B. Stephens and Associates, Inc. (DBSA)		
Williams (2001) data from TetraTech report to NMISC	360	1934 to 1963
Daniel B. Stephens and Associates, Inc. (written commun., 2005)	29	1934 to 2005

The next step after the environmental data were compiled was to identify and isolate salinity-specific data from the rest of the data received. A representative from the USGS, New Mexico Water Science Center (Doug Moyer) was consulted to develop a list of potential terms that might have been used to describe salinity-related constituents (referred to as parameters in this geodatabase) as part of a salinity-related water-quality investigation. A search of all parameters used in each agency or university database was done and the resulting salinity-specific data were then reviewed by NMISC and USACE. A table of all parameters used to describe results reported by source agencies was generated and used for data querying and quality-assurance measures.

Data Input

Comparable fields within each agency's data structure were identified. Structured query language (SQL) code automated the creation of tables for site, sample, parameter, and result data (appendixes 3 and 4). The SQL code isolates appropriate data, then maps comparable fields and appends them into a single table. Once the tables were created, they were loaded into a geodatabase framework that supports linked tables, domains to interpret coded values, and spatial locations. This automation documents the comparable fields, selects appropriate data, and facilitates updates to minimize human error.

Data Quality Assurance

Duplicate data can occur when multiple agencies are reporting the same data in one or more databases used in a compilation effort. Elimination of duplicate data is essential to ensure the integrity of data being queried and ensures that the database does not contain information that could bias subsequent analyses.

Identification of duplicate data was performed both spatially (location on the earth) and temporally (date and time). Database queries analyzed site name and sample dates whereas spatial analysis compared site locations. Sample information queries were written to identify potential duplicate data on the basis of location and date. Samples were considered duplicates if the site identifier, sample date, and parameter result values were identical. If an entire dataset was duplicated elsewhere, the original dataset was used and the duplicate excluded.

To examine duplicates based on proximity or spatial location, all sites within 30 meters of another site were visually inspected using ESRI ArcGIS™ software. If a site was within the 30-meter buffer or very close to a USGS site (well or gage), then the USGS site location was given priority and input into the geodatabase. If no USGS site locations were found within the 30-meter buffer, then criteria based on originating agency, site period of record, and data richness of the site location were used to establish the site identifier, which was captured in the geodatabase. Additional methods such as examining well locations with digital orthophoto quarter quadrangles and USGS topographic maps were used to determine the accuracy of the spatial coordinates assigned by an agency or individual.

Metadata

Metadata that comply with Federal Geographic Data Committee (FGDC) standards were created for each spatial component. The metadata record documents the basic characteristics of the data or information resource in the study area. Metadata components include information such as the title, abstract, and publication date of source documents; geographic elements such as geographic extent and projection information; and database elements such as attribute label definitions and attribute domain values. The metadata record for the sites feature class is contained in appendix 5.

References

Allison, G.B., Cook, P.G., Barnett, S.R., Walker, G.R., Jolly, I.D., and Hughes, M.W., 1990, Land clearance and river salinization in the western Murray Basin, Australia: Journal of Hydrology, v. 119, p. 1–20.

Anderholm, S.K., 2002, Water-quality assessment of the Rio Grande Valley, Colorado, New Mexico, and Texas—Surface-water quality, shallow ground-water quality, and factors affecting water quality in the Rincon Valley, south-central New Mexico, 1994–95: U.S. Geological Survey Water-Resources Investigations Report 02–4188, 117 p.

Benke, A.C., and Cushing, C.E., eds., 2005, Rivers of North America: Academic Press, 1,144 p.

Bexfield, L.M., and Anderholm, S.K., 1997, Water-quality assessment of the Rio Grande Valley, Colorado, New Mexico, and Texas—Ground-water quality in the Rio Grande flood plain, Cochiti Lake, New Mexico, to El Paso, Texas, 1995: U.S. Geological Survey Water-Resources Investigations Report 96–4249, 93 p.

Chapin, C.E., 1971, The Rio Grande Rift, part 1—Modifications and additions, *in* James, H.L., ed., Guidebook of the San Luis Basin, Colorado: New Mexico Geological Society, Twenty-Second Field Conference, p. 191–201.

Conover, C.S., 1954, Ground-water conditions in the Rincon and Mesilla Valleys and adjacent areas in New Mexico: U.S. Geological Survey Water-Supply Paper 1230, 200 p.

Frenzel, P.F., Kaehler, C.A., and Anderholm, S.K., 1992, Geohydrology and simulation of ground-water flow in the Mesilla Basin, Doña Ana County, New Mexico, and El Paso

County, Texas: U.S. Geological Survey Professional Paper 1407–C, 105 p.

Hawley, J.W., 1978, Guidebook to the Rio Grande Rift in New Mexico and Colorado: New Mexico Bureau of Mines and Mineral Resources Circular 163, 241 p.

Hawley, J.W., and Kennedy, J.F., 2004, Creation of a digital hydrogeologic framework model of the Mesilla Basin and southern Jornada del Muerto Basin: Las Cruces, N. Mex., New Mexico Water Resources Research Institute of New Mexico State University, Technical Completion Report 332, 105 p.

Hawley, J.W., Kennedy, J.F., Ortiz, Marquita, and Carrasco, Sean, 2005, Digital hydrogeologic framework model of the Rincon Valley and adjacent areas of Doña Ana, Sierra and Luna Counties, NM: Las Cruces, N. Mex., New Mexico Water Resources Research Institute of New Mexico State University, addendum to Technical Completion Report 332, accessed October 22, 2009, at *http://wrri.nmsu.edu/publish/techrpt/tr332/cdrom/addendum.pdf*.

Hibbs, Barry, Phillips, Fred, Hogan, James, Eastoe, Chris, Hawley, John, Granados, Alfred, and Hutchison, Bill, 2003, Hydrogeologic and isotopic study of the groundwater resources of the Hueco Bolson aquifer, El Paso, Texas/Juarez, Mexico area: Hydrological Science and Technology, v. 19, no. 1–4, p. 109–119.

Hutchison, W.R., 2006, Groundwater management in El Paso, Texas: University of Texas at El Paso, Ph.D. dissertation, 329 p.

Langman, J.B., 2009, Traveltime of the Rio Grande in the Middle Rio Grande Basin, water years 2003–05: U.S. Geological Survey Scientific Investigations Report 2007–5292, 34 p.

McLeod, D.J., 1976, High level domain definition in a relational data base system, *in* Conference on Data—Abstraction, Definition and Structure, Proceedings: Association for Computing Machinery Portal, v. 8, no. 2, p. 45–57.

Mills, S.K., 2003, Quantifying salinization of the Rio Grande using environmental tracers: Socorro, N. Mex., New Mexico Institute of Mining and Technology, M.S. thesis, 397 p., available at *http://dspace.nmt.edu/dspace/handle/10136/437?mode=full&submit_simple=Show+full+item+record*.

Moore, S.J., and Anderholm, S.K., 2002, Spatial and temporal variations in streamflow, dissolved solids, nutrients, and suspended sediment in the Rio Grande Valley study unit, Colorado, New Mexico, and Texas, 1993–95: U.S. Geological Survey Water-Resources Investigations Report 02–4224, 52 p.

Nickerson, E.L., 1995, Selected hydrologic data for the Mesilla ground-water basin, 1987 through 1992 water years, Doña Ana County, New Mexico, and El Paso County, Texas: U.S. Geological Survey Open-File Report 95–111, 123 p.

Phillips, F.M., Hogan, J.F., Mills, S.K., and Hendrickx, J.M.H., 2003, Environmental tracers applied to quantifying causes of salinity in arid-region rivers—Preliminary results from the Rio Grande, southwestern USA, *in* Alsharhan, A.S., and Wood, W.W., eds., Water resources perspectives—Evaluation, management, and policy: Amsterdam, Elsevier, Developments in Water Science, v. 50, p. 327–334.

Pillsbury, A.F., 1981, The salinity of rivers: Scientific American, v. 245, no. 1, p. 54–65.

Riecker, R.R., ed., 1979, Rio Grande Rift—Tectonics and magmatism: American Geophysical Union, 438 p.

Stabler, Herman, 1911, Some stream waters of the western United States, *with chapters on* Sediment carried by the Rio Grande and the industrial application of water analyses: U.S. Geological Survey Water-Supply Paper 274, 188 p.

U.S. Environmental Protection Agency, 2009, Modernized Storage and Retrieval repository: accessed January 5, 2009, at *http://www.epa.gov/storet/about.html*.

U.S. Geological Survey, 2009, National Water Information System (NWIS Web): accessed January 4, 2009, at *http://waterdata.usgs.gov/nwis*.

Wellborn, T.L., and Moreo, M.T., 2007, Irrigated acreage within the Basin and Range carbonate-rock aquifer system, White Pine County, Nevada, and adjacent areas in Nevada and Utah: U.S. Geological Survey Data Series 273, 17 p., accessed online January 21, 2010, at *http://pubs.usgs.gov/ds/2007/273/pdf/ds273.pdf*.

Wilkins, D.W., 1998, Summary of the southwest alluvial basins regional aquifer-system analysis in parts of Colorado, New Mexico, and Texas: U.S. Geological Survey Professional Paper 1407–A, 49 p.

Williams, J.H., 2001, Salt balance in the Rio Grande project from San Marcial, New Mexico to Fort Quitman, Texas: Master's thesis, New Mexico State University, 80 p.

Wilson, C.A., White, R.R., Orr, B.R., and Roybal, G.R., 1981, Water resources of the Rincon and Mesilla Valleys and adjacent areas, New Mexico: New Mexico State Engineer Technical Report 43, 514 p.

Zeiler, Michael, 1999, Modeling our world—The ESRI guide to geodatabase design: Redlands, Calif., ESRI Press, 216 p.

Appendixes

Blank Page

Abbreviations

DBSA, Daniel B. Stephens and Associates, Inc.

EPA, U.S. Environmental Protection Agency

FGDC, Federal Geographic Data Committee

ID, identifier

LRG, Lower Rio Grande

NMED, New Mexico Environmental Department

NWIS, National Water Information System

SAHRA, Sustainability of semi-Arid Hydrology and Riparian Areas

SSPA, S.S. Papadopulos and Associates

STORET, Storage and Retrieval Repository

TCEQ, Texas Commission on Environmental Quality

TWDB, Texas Water Development Board

USGS, U.S. Geological Survey

WILSON, Wilson and others (1981)

Appendix 1—Data Sources

USGS-New Mexico

Source: Internal data access direct from NWIS.
Processing: Created unique parameters and sample IDs by prefixing with 'USGSNM'.
Date Range: 1/1/1901–7/28/2009

USGS-Texas

Source: Internal data access direct from NWIS.
Processing: Created unique parameters and sample IDs by prefixing with 'USGSTX'.
Date Range: 1/1/1901–7/28/2009

SAHRA

Source: *SAHRA Data*
Processing: Created sample table. Prefixed parameters and sample IDs with 'SAHRA'. Data from Mills (2003) included.
Date Range: 1/28/2000–8/10/2006

EPA STORET Modern

Source: *EPA Storet Modern*
Processing: Created sample table. Filtered out data collected by 21TXWQB (Texas Water Commission, now TCEQ). Created unique parameters and sample IDs by prefixing with 'EPA'.
Date Range: 10/11/1937–5/3/2029*

EPA STORET Legacy

Source: *EPA Storet Legacy*
Processing: Created sample table. Filtered out data collected by 21TXWQB (Texas Water Commission, now TCEQ). Created unique parameters and sample IDs by prefixing with 'EPAL'.
Date Range: 10/11/1937–5/3/2029*

TCEQ

Source: *Texas Commission on Environmental Quality*
Processing: Imported data by basin (Rio Grande, Rio Grande-Nueces) for sites and results. Created sample table by site and sample date. Created unique parameters and sample IDs by prefixing with 'TCEQ'.
Date Range: 9/18/1968–8/26/1992

TWDB

Source: *Texas Water Development Board*
Processing: Downloaded geodatabase. Appended water-quality data in a new table, changing horizontal data to vertical. Filtered out USGS data. Created unique parameters and sample IDs by prefixing with 'TWDB'.
Date Range: 1/1/1800*–9/15/2007

SSPA

Source: *v3 WaterQuality*
Processing: Created unique parameters and sample IDs by prefixing with 'SSPA'.
Date Range: 2/11/1931–12/15/1995

Wilson, Orr, White, and Roybal

Source: *Wilson Data*
Processing: Created unique parameters and sample IDs by prefixing with 'W'.
Date Range: 4/22/1936–9/18/1977

DBSA

Source: *DBSA Water Quality*
Processing: Excluded all data except Williams (2001) data from Tetra Tech and Daniel B. Stephens and Associates, Inc. (written commun., 2005) data. Mills (2003) data are included in SAHRA database. Created unique parameters and sample IDs by prefixing with 'DBSA'.
Date Range: 1/15/1934–1/27/2005

NMED

Source: *NMED LRG Monitoring*
Processing: Created unique parameters and sample IDs by prefixing with 'NMED'.
Date Range: 5/23/20057*–2/14/2008

* Date not reliable

Appendix 2—Database Compilation Notes

Data Preparation

To compile data from disparate sources, it is necessary to analyze, prepare, and standardize the data to create a comprehensive dataset. Data preparation included:
Identifying key data sources
Creating list of unique water-quality parameters
Creating table of unique samples
Analyzing data for duplicate records
Determining site locations (latitude and longitude)
Resolving duplicate sites
Once data were processed, *UNION* queries for site, sample, result, and parameter values put the data together in four tables.

Duplicate Resolution

Samples and Results

Sample and result data were imported into the salinity staging database and cross-checked for duplicates by sample date/time and parameter results. If a sample and result pair was duplicated in another dataset, the data closest to the original (native) data source was used. Example: USGS staff sampled a well on July 9, 2007, and uploaded results into NWIS. TWDB received the sample and result data and uploaded the results into their system. Because both systems contain the same information, the original data would be used, in this case NWIS.

Example SQL Query for Locating Duplicate Samples

SELECT Sample.site_id, Sample.dt_tm, Sample.collecting_agency, Sample.sample_id **FROM** Sample **WHERE** (((Sample.site_id) **IN** (**SELECT** [site_id] **FROM** [Sample] **AS** Tmp **GROUP BY** [site_id],[dt_tm] **HAVING Count**(*)>1 **AND** [dt_tm] = [Sample].[dt_tm])) **ORDER BY** Sample.site_id, Sample.dt_tm;

Sites

The following process was used to determine duplicate sites:
Remove duplicate sites with same name, longitude, and latitude (use original data)
Create 30-meter buffer around all sites
Consider duplicates to be within the overlap of buffers
All sites that are duplicates of USGS sites were flagged with USGS site name as primary site
All sites within the buffer and with different names were left in the database

Example SQL Query for Locating Duplicate Sites

SELECT Sites.site_id, **Count**(Sites.site_id) **AS** CountOfsite_id **FROM** Sites **GROUP BY** Sites.site_id **HAVING** (((**Count**(Sites.site_id))>1));

Appendix 3—Union Queries

A union query combines the result sets of several similar select queries. For example, suppose that you have one table that stores information about customers, another table that stores information about suppliers, and no relation exists between the two tables. Suppose that both tables have fields that store contact information, and you want to look at all of the contact information from both tables at the same time. You could create a select query (select query: A query that asks a question about the data stored in your tables and returns a result set in the form of a datasheet, without changing the data) for each table to retrieve only those fields that contain contact information, but the information that is returned would still be in two separate places. To combine the results of two or more select queries into one result set, you can use a union query.

SQL Code for Site Union Query

SELECT s.site_no as site_id, s.site_tp_cd **AS** type_code, s.dec_lat_va **AS** lat_dd, s.dec_long_va **AS** lon_dd, s.coord_datum_cd **AS** h_datum, s.alt_va **AS** alt_ft, s.alt_datum_cd **AS** v_datum, "" **AS** site_comments, "NWIS NM" **AS** source
FROM [data\USGS_NM\update_20090819\nwisdb_gdb_01.mdb].sitefile **AS** s
UNION
SELECT s.site_no as site_id, s.site_tp_cd **AS** type_code, s.dec_lat_va **AS** lat_dd, s.dec_long_va **AS** lon_dd, s.coord_datum_cd **AS** h_datum, s.alt_va **AS** alt_ft, s.alt_datum_cd **AS** v_datum, "" **AS** site_comments, "NWIS TX" **AS** source
FROM [data\USGS_TX\update_20090819\nwisdb_gdb_01.mdb].sitefile **AS** s
UNION
SELECT s.[SiteCode] **AS** site_id, "SW" **AS** type_code, s.[Latitude] **AS** lat_dd, s.[Longitude] **AS** lon_dd, "" **AS** h_datum, s.[Elevation_m] **AS** alt_ft, s.[VerticalDatum] **AS** v_datum, "" **AS** site_comments, "SAHRA" **AS** source
FROM [data\SAHRA_20081210\RioGrandeWQ_ODM_working_v3_NoMisVal.mdb].Sites **AS** s
UNION
SELECT s.Site_ID, s.[site_type] **AS** type_code, s.[Latitude] **AS** lat_dd, s.[Longitude] **AS** lon_dd, s.[Datum] **AS** h_datum, 0 **AS** alt_ft, "" **AS** v_datum, "" **AS** site_comments, "STORET MODERN" **AS** source
FROM [data\EPA\Modernized\20090728\EPAModern.mdb].Sites **AS** s
UNION
SELECT s.site_id, s.[TYPE] **AS** type_code, s.lat_dd, s.lon_dd, s.h_datum_dd, s.alt_ft, s.v_datum, "" **AS** site_comments, "STORET LEGACY" **AS** source
FROM [data\EPA\Legacy\EPA_Legacy2.mdb].TBL_LOC **AS** s
UNION
SELECT s.[Station_ID] **AS** site_id, s.[Stream_Station_Type_1] **AS** type_code, s.Latitude, s.Longitude, "" **AS** h_datum, "" **AS** alt_ft, "" **AS** v_datum, "" **AS** site_comments, "TCEQ" **AS** source
FROM [data\TCEQ\TCEQ.mdb].TCEQ_Stations **AS** s
UNION
SELECT w.[state_well_number] **AS** site_id, "WELL" **AS** type_code, w.[lat_dec] **AS** lat_dd, w.[long_dec] **AS** lon_dd, "" **AS** h_datum, "" **AS** alt_ft, "" **AS** v_datum, "" **AS** site_comments, "TWDB" **AS** source
FROM [data\TWDB\TWDB.mdb].m**AS**ter_sample **AS** s **INNER JOIN** [data\TWDB\TWDB.mdb].welldata **AS** w ON s.state_well_number = w.state_well_number **WHERE** (((w.county_code)=141 **Or** (w.county_code)=229 **Or** (w.county_code)=377))
UNION
SELECT c.[Site ID] **AS** site_id, s.[Water Type Code] **AS** type_code, c.[LAT (DD)] **AS** lat_dd, c.[LONG (DD)] **AS** lon_dd, c.[DD Datum] **AS** h_datum, c.[ALTITUDE (FT)] **AS** alt_ft, c.[ALT Datum] **AS** v_datum, "" **AS** site_comments, "SSPA" **AS** source
FROM [data\SSPA_V2_2004_Data_compendium\WaterQuality\v3WaterQuality.mdb].[TBL WQ Site Info] **AS** s **INNER JOIN** [data\SSPA_V2_2004_Data_compendium\WaterQuality\v3WaterQuality.mdb].v2_SiteID_CoordinateData **AS** c ON s.[Site ID] = c.[Site ID]
UNION
SELECT l.Site_ID, "" **AS** type_code, l.[Latitude] **AS** lat_dd, l.[Longitude] **AS** lon_dd, "NAD83" **AS** h_datum, l.[GE] **AS** alt_ft, "" **AS** v_datum, "" **AS** site_comments, "WILSON Calc" **AS** source
FROM [data\Wilson\WilsonData.mdb].WilsonWells **AS** l
UNION
SELECT l.[Site ID] **AS** Loc_id, s.[Water Type Code] **AS** type_code, l.[Latitude DD] **AS** lat_dd, l.[Longitude DD] **AS** lon_dd, l.[Horiz Datum DD] **AS** h_datum, l.[Elevation (ft)] **AS** alt_ft, l.[Vert Datum] **AS** v_datum, "" **AS** site_comments, "DBSA" **AS** source

FROM [data\NMISC_20081120\SurfaceWater\Sources\DBSA\DBSAWaterQuality.mdb].[TBL WQ Location] **AS** l **INNER JOIN** [data\NMISC_20081120\SurfaceWater\Sources\DBSA\DBSAWaterQuality.mdb].[TBL_Sample] **AS** s ON l.[Site ID] = s.[Site ID]
UNION
SELECT l.Site_ID, l.[Type] **AS** type_code, l.[Latitude] **AS** lat_dd, l.[Longitude] **AS** lon_dd, l.[HorizontalDatum] **AS** h_datum, l.[Elevation] **AS** alt_ft, l.[VerticalDatum] **AS** v_datum, "" **AS** site_comments, "NMED" **AS** source
FROM [data\NMISC_Compendium_20090624\Compendium_1208\SurfaceWater\Sources\NMED\NMED_LRG_Monitoring_1208.mdb].TBL_LOC **AS** l;

SQL Code for Sample Union Query

SELECT 'USGSNM' & s.[record_no] **AS** sample_id, s.site_no **AS** site_id, s.sample_start_dt **AS** dt_tm, 0 **AS** sample_top, 0 **AS** sample_bottom, s.aqfr_cd, "USGS_" & s.coll_ent_cd **AS** collecting_agency **FROM** [data\USGS_NM\update_20090819\nwisdb_gdb_01.mdb].qw_sample **AS** s
UNION
SELECT 'USGSTX' & s.[record_no] **AS** sample_id, s.site_no **AS** site_id, s.sample_start_dt **AS** dt_tm, 0 **AS** sample_top, 0 **AS** sample_bottom, s.aqfr_cd, "USGS_" & s.coll_ent_cd **AS** collecting_agency
FROM [data\USGS_TX\update_20090819\nwisdb_gdb_01.mdb].qw_sample **AS** s
UNION
SELECT s.sample_id, s.site_ID, s.dt_tm, s.Sample_top, s.Sample_bottom, "" **AS** aqfr_cd, "SAHRA" **AS** collecting_agency
FROM [data\SAHRA_20081210\RioGrandeWQ_ODM_working_v3_NoMisVal.mdb].Sample2 **AS** s
UNION
SELECT s.sample_id, s.Site_ID, s.dt_tm, s.Sample_Top, s.Sample_Bottom, s.aqfr_cd, "EPA_" & s.collecting_agency
FROM [data\EPA\Modernized\20090728\EPAModern.mdb].Samples **AS** s
UNION
SELECT s.sample_id, s.Site_ID, s.dt_tm, s.Sample_Top, s.Sample_Bottom, s.aqfr_cd, "EPA_" & s.collecting_agency
FROM [data\EPA\Legacy\EPA_Legacy2.mdb].Sample **AS** s
UNION
SELECT "TCEQ" & s.[Tag_ID] **AS** sample_id, s.Station_ID **AS** site_id, Format(s.[enddate] & " " & s.[endtime],"General Date") **AS** dt_tm, s.StartDepth **AS** sample_top, s.EndDepth **AS** smaple_bottom, "" **AS** aqfr_cd, "TCEQ_" & s.CollectingEntity **AS** collecting_agency
FROM [data\TCEQ\TCEQ.mdb].Sample **AS** s
WHERE (((Sample.SubmittingEntity)<>"GS"))
UNION
SELECT "TWDB" & s.[sample_idx] **AS** sample_id, s.state_well_number **AS** site_id, s.date_tm **AS** sample_dt, s.top_s_interval **AS** sample_top, s.bottom_s_interval **AS** sample_bottom, s.samp_int_aqcode **AS** aqfr_cd, "TWDB_" & entity_codes.entity **AS** collecting_agency
FROM [data\TWDB\TWDB.mdb].master_sample **AS** s **INNER JOIN** [data\TWDB\TWDB.mdb].entity_codes ON s.collecting_agency = entity_codes.entity_code
WHERE (((entity_codes.entity_code)<>"03")) **AND** ((welldata.county_code)=141 **OR** (welldata.county_code)=229 **OR** (welldata.county_code)=377))
UNION
SELECT s.sample_id, s.Site_ID, s.dt_tm **AS** sample_dt, s.Sample_Top, s.Sample_Bottom, "SSPA_" & s.coll_agency **AS** collecting_agency, s.aqfr_cd
FROM [data\SSPA_V2_2004_Data_compendium\WaterQuality\v3WaterQuality.mdb].TBL_Sample **AS** s
WHERE (((s.coll_agency) In (1,4,5,8,9,10,11,13,14,15,16,17,18,19,21,26)))
UNION
SELECT s.Sample_ID, s.Site_ID, s.Sample_Dt, s.Sample_Top, s.Sample_Bottom, "WILSON" **AS** collecting_agency, w.GU **AS** aqfr_cd
FROM [data\Wilson\WilsonData.mdb].WilsonWells **AS** w **INNER JOIN** [data\Wilson\WilsonData.mdb].Wilson_Sample **AS** s ON w.Site_ID = s.Site_ID
UNION
SELECT s.sample_id, s.Site_ID, s.sample_date **AS** sample_dt, s.Sample_Top, s.Sample_Bottom, "DBSA_" & s.collecting_agency, s.aqfr_cd
FROM [data\NMISC_20081120\SurfaceWater\Sources\DBSA\DBSAWaterQuality.mdb].TBL_Sample **AS** s

WHERE (((s.collecting_agency) In (30,32)))
UNION
SELECT s.Sample_ID, s.SiteID **AS** site_id, s.dt_tm **AS** sample_dt, s.Sample_Top, s.Sample_Bottom, "NMED" **AS** Collecting_agency, "" **AS** aqfr_cd
FROM [data\NMISC_Compendium_20090624\Compendium_1208\SurfaceWater\Sources\NMED\NMED_LRG_Monitoring_1208.mdb].Sample **AS** s;

SQL Code for Result Union Query

SELECT "USGSNM" & r.record_no **AS** sample_id, "USGS" & r.parm_cd **AS** param_code, r.remark_cd **AS** result_qualifier, r.result_va **AS** result_value, p.units **AS** result_units, "" **AS** result_description, "USGS" **AS** distributing_agency, "" **AS** [result_comments]
FROM tbl_search **AS** t **INNER JOIN** ([\\Igskiacwgsnas\gis_proj\b8653_dp6\data\USGS_NM\update_20090819\nwisdb_gdb_01.mdb].qw_sample **AS** s **INNER JOIN** ([\\Igskiacwgsnas\gis_proj\b8653_dp6\data\USGS_NM\update_20090819\nwisdb_gdb_01.mdb].qw_result **AS** r **INNER JOIN** [\\Igskiacwgsnas\gis_proj\b8653_dp6\data\USGS_NM\update_20090819\nwisdb_gdb_01.mdb].Pmcodes **AS** p **ON** r.parm_cd=p.parameter_cd) **ON** s.record_no=r.record_no) **ON** p.parameter_nm like t.Term **WHERE** (((s.sample_start_dt) Is Not Null))
UNION
SELECT "USGSTX" & r.record_no **AS** sample_id, "USGS" & r.parm_cd **AS** param_code, r.remark_cd **AS** result_qualifier, r.result_va **AS** result_value, p.units **AS** result_units, "" **AS** result_description, "USGS" **AS** distributing_agency, "" **AS** [result_comments]
FROM tbl_search **AS** t **INNER JOIN** ([\\Igskiacwgsnas\gis_proj\b8653_dp6\data\USGS_TX\update_20090819\nwisdb_gdb_01.mdb].qw_sample **AS** s **INNER JOIN** ([\\Igskiacwgsnas\gis_proj\b8653_dp6\data\USGS_TX\update_20090819\nwisdb_gdb_01.mdb].qw_result **AS** r **INNER JOIN** [\\Igskiacwgsnas\gis_proj\b8653_dp6\data\USGS_TX\update_20090819\nwisdb_gdb_01.mdb].Pmcodes **AS** p **ON** r.parm_cd=p.parameter_cd) **ON** s.record_no=r.record_no) **ON** p.parameter_nm like t.Term **WHERE** (((s.sample_start_dt) Is Not Null))
UNION
SELECT sa.sample_id, "SAHRA" & Format([v].[VariableID]) **AS** param_code, cv.symbol **AS** result_qualifier, dv.DataValue **AS** result_value, UCase([UnitsAbbreviation]) **AS** result_units, cv.Definition **AS** result_description, "SAHRA" **AS** distributing_agency, "" **AS** result_comments
FROM [\\Igskiacwgsnas\gis_proj\b8653_dp6\data\SAHRA_20081210\RioGrandeWQ_ODM_working_v3_NoMisVal.mdb].Units **INNER JOIN** ([\\Igskiacwgsnas\gis_proj\b8653_dp6\data\SAHRA_20081210\RioGrandeWQ_ODM_working_v3_NoMisVal.mdb].Sample2 **AS** sa **INNER JOIN** ([\\Igskiacwgsnas\gis_proj\b8653_dp6\data\SAHRA_20081210\RioGrandeWQ_ODM_working_v3_NoMisVal.mdb].CensorCodeCV **AS** cv **INNER JOIN** ([\\Igskiacwgsnas\gis_proj\b8653_dp6\data\SAHRA_20081210\RioGrandeWQ_ODM_working_v3_NoMisVal.mdb].Variables **AS** v **INNER JOIN** [\\Igskiacwgsnas\gis_proj\b8653_dp6\data\SAHRA_20081210\RioGrandeWQ_ODM_working_v3_NoMisVal.mdb].DataValues **AS** dv **ON** v.VariableID = dv.VariableID) **ON** cv.Term = dv.CensorCode) **ON** sa.sample_id = dv.sample_id) **ON** Units.UnitsID = v.VariableUnitsID
UNION
SELECT r.Sample_ID, "EPA" & r.PCODE **AS** param_code, r.result_qualifier, r.result_value, UCase(r.[result_units]) **AS** Expr1, r.result_description, "EPA" **AS** distributing_agency, "" **AS** result_comments
FROM [\\Igskiacwgsnas\gis_proj\b8653_dp6\data\EPA\Modernized\20090728\EPAModern.mdb].PCODE **AS** p **INNER JOIN** [\\Igskiacwgsnas\gis_proj\b8653_dp6\data\EPA\Modernized\20090728\EPAModern.mdb].QW_Res **AS** r **ON** p.PCODE = r.PCODE
WHERE (((p.SALINITY)="Y"))
UNION
SELECT r.sample_id, "EPA" & r.Param **AS** param_code, r.R **AS** result_qualifier, r.Result_Value, UCase(rc.[Name]) **AS** result_units, "" **AS** result_description, "EPA" **AS** distributing_agency, "" **AS** result_comments
FROM ([\\Igskiacwgsnas\gis_proj\b8653_dp6\data\EPA\Legacy\EPA_Legacy2.mdb].Parameter **AS** p **INNER JOIN** [\\Igskiacwgsnas\gis_proj\b8653_dp6\data\EPA\Legacy\EPA_Legacy2.mdb].Reporting_codes **AS** rc **ON** p.Reporting_Units = rc.Code) **INNER JOIN** [\\Igskiacwgsnas\gis_proj\b8653_dp6\data\EPA\Legacy\EPA_Legacy2.mdb].Results **AS** r **ON** p.Param = r.Param
WHERE (((p.selected)="Y"))
UNION
SELECT "TCEQ" & [r].[Tag_ID] **AS** sample_id, "TCEQ" & r.Parameter_Code **AS** param_code, r.GreaterThan_LessThan **AS** result_qualifier, r.ResValue **AS** result_value, p.[Units of Measurement] **AS** result_units, "VERIFIED: " & r.[VerifyFlag] **AS**

```
result_description, "TCEQ" AS distributing_agency, "" AS result_comments
FROM [\\Igskiacwgsnas\gis_proj\b8653_dp6\data\TCEQ\TCEQ.mdb].Sw_parm AS p INNER JOIN [\\Igskiacwgsnas\gis_proj\b8653_dp6\data\TCEQ\TCEQ.mdb].Results AS r ON p.Parameter_Code = r.Parameter_Code
WHERE (((p.selected)="Y"))
UNION
SELECT "TWDB" & [r].[sample_idx] AS sample_id, r.parm_cd AS param_code, r.remark_cd AS result_qualifier, r.result_va AS result_value, sc.units_of_measure AS result_units, "" AS result_description, "TWDB" AS distributing_agency, "" AS result_comments
FROM tbl_search INNER JOIN ([\\Igskiacwgsnas\gis_proj\b8653_dp6\data\TWDB\TWDB.mdb].welldata INNER JOIN ([\\Igskiacwgsnas\gis_proj\b8653_dp6\data\TWDB\TWDB.mdb].master_sample AS s INNER JOIN ([\\Igskiacwgsnas\gis_proj\b8653_dp6\data\TWDB\TWDB.mdb].storet_code AS sc INNER JOIN [\\Igskiacwgsnas\gis_proj\b8653_dp6\data\TWDB\TWDB.mdb].master_result AS r ON sc.storet_code = r.parm_cd) ON s.sample_idx = r.sample_idx) ON welldata.state_well_number = s.state_well_number) ON sc.long_description LIKE tbl_search.Term
WHERE (((s.collecting_agency)<>"03") AND ((welldata.county_code)=141 OR (welldata.county_code)=229 OR (welldata.county_code)=377))
UNION
SELECT p.sample_id, "SSPA" & p.[Parameter Code] AS param_code, p.Qualifier AS result_qualifier, p.Value AS result_value, UCase([Unit]) AS result_units, c.CommentDescrip AS result_description, "SSPA" AS distributing_agency, "" AS result_comments
FROM tbl_search INNER JOIN ([\\Igskiacwgsnas\gis_proj\b8653_dp6\data\SSPA_V2_2004_Data_compendium\WaterQuality\v3WaterQuality.mdb].QRY_Sample AS s INNER JOIN ([\\Igskiacwgsnas\gis_proj\b8653_dp6\data\SSPA_V2_2004_Data_compendium\WaterQuality\v3WaterQuality.mdb].[REF WQ Comment Code] AS c INNER JOIN [\\Igskiacwgsnas\gis_proj\b8653_dp6\data\SSPA_V2_2004_Data_compendium\WaterQuality\v3WaterQuality.mdb].[TBL WQ Parameter] AS p ON c.[Comment Code] = p.[Comment Code]) ON s.sample_id = p.sample_id) ON p.Parameter like tbl_search.Term
UNION
SELECT r.Sample_ID, r.Param AS param_code, "" AS result_qualifier, r.Result AS result_value, r.Units AS result_units, "" AS result_description, "WILSON" AS distributing_agency, "" AS result_comments
FROM [\\Igskiacwgsnas\gis_proj\b8653_dp6\data\Wilson\WilsonData.mdb].Wilson_WQ_Results AS r
UNION
SELECT r.sample_id, "DBSA" & r.[Parameter Code] AS param_code, r.Qualifier AS result_qualifier, r.Value AS result_value, UCase([Unit]) AS result_units, r.[Comment Code] AS result_description, "DBSA" AS distributing_agency, "" AS result_comments
FROM [\\Igskiacwgsnas\gis_proj\b8653_dp6\data\NMISC_20081120\SurfaceWater\Sources\DBSA\DBSAWaterQuality.mdb].QRY_Sample AS s INNER JOIN [\\Igskiacwgsnas\gis_proj\b8653_dp6\data\NMISC_20081120\SurfaceWater\Sources\DBSA\DBSAWaterQuality.mdb].[TBL WQ Parameter] AS r ON s.sample_id = r.sample_id
UNION
SELECT r.Sample_ID, "NMED" & r.Pcode AS param_code, r.Qualification AS result_qualifier, r.Value_ AS Result_Value, UCase(r.[Units]) AS result_units, r.Comment AS result_description, "NMED" AS distributing_agency, "" AS result_comments
FROM [\\Igskiacwgsnas\gis_proj\b8653_dp6\data\NMISC_Compendium_20090624\Compendium_1208\SurfaceWater\Sources\NMED\NMED_LRG_Monitoring_1208.mdb].SW_GW_Results AS r;
```

SQL Code for Parameter Union Query

```
SELECT "USGS" & Pmcodes.parameter_cd AS param_code, Pmcodes.parameter_nm AS param_description
FROM [data\USGS_TX\nwisdb_gdb_01.mdb].tbl_search INNER JOIN [data\USGS_TX\nwisdb_gdb_01.mdb].Pmcodes ON Pmcodes.parameter_nm LIKE tbl_search.Term
UNION
SELECT "SAHRA" & [VariableID] AS param_code, Variables.VariableName AS param_description
FROM [data\SAHRA_20081210\RioGrandeWQ_ODM_working_v3_NoMisVal.mdb].Variables
UNION
SELECT "EPA" & [PCODE] AS param_code, PCODE.Characteristic_Name AS param_description
FROM [data\EPA\Modernized\20090728\EPAModern.mdb].PCODE WHERE (((PCODE.SALINITY)="Y"))
UNION
SELECT "EPA" & [parameter_code] AS param_code, STORET_PRM.Parameter AS param_description
FROM [data\EPA\Legacy\EPA_Legacy2.mdb].STORET_PRM WHERE (((STORET_PRM.selected)="Y"))
```

```
UNION
SELECT "TCEQ" & [Parameter_Code] AS param_code, Sw_parm.[Parameter Description] AS param_description
FROM [data\TCEQ\TCEQ.mdb].Sw_parm INNER JOIN [data\TCEQ\TCEQ.mdb].tbl_search ON Sw_parm.[Parameter Description] LIKE tbl_search.Term
UNION
SELECT "TWDB" & TWDBstoret.storet_code AS param_code, TWDBstoret.long_description AS param_description
FROM [data\TWDB\TWDB.mdb].tbl_search INNER JOIN [data\TWDB\TWDB.mdb].TWDBstoret ON TWDBstoret.long_description LIKE tbl_search.Term
UNION
SELECT "SSPA" & [REF WQ Parameter Code].[Parameter Code] AS param_code, [REF WQ Parameter Code].[ParameterDesc] AS param_description
FROM [data\SSPA_V2_2004_Data_compendium\WaterQuality\v3WaterQuality.mdb].tbl_search INNER JOIN [data\SSPA_V2_2004_Data_compendium\WaterQuality\v3WaterQuality.mdb].[REF WQ Parameter Code] ON [REF WQ Parameter Code].ParameterDesc LIKE tbl_search.Term
UNION
SELECT Wilson_USGS_Param.Param AS param_code, Wilson_USGS_Param.Description AS param_description
FROM [data\Wilson\WilsonData.mdb].Wilson_USGS_Param
UNION
SELECT "DBSA" & [REF WQ Parameter Code].[Parameter Code] AS param_code, [REF WQ Parameter Code].ParameterDesc AS param_description
FROM [data\NMISC_20081120\SurfaceWater\Sources\DBSA\DBSAWaterQuality.mdb].tbl_search INNER JOIN [data\NMISC_20081120\SurfaceWater\Sources\DBSA\DBSAWaterQuality.mdb].[REF WQ Parameter Code] ON [REF WQ Parameter Code].ParameterDesc LIKE tbl_search.Term
UNION
SELECT "NMED" & [PCODE] AS param_code, SW_Parameters.[Parameter] AS param_description
FROM [data\NMISC_Compendium_20090624\Compendium_1208\SurfaceWater\Sources\NMED\NMED_LRG_Monitoring_1208.mdb].SW_Parameters;
```

Appendix 4—Data Query Examples

Query Results by Parameter Example

SELECT sites.site_id, sample.dt_tm, param_codes.common_Name, result.result_qualifier, result.result_value, result.result_units **FROM** ((sites **INNER JOIN** sample **ON** sites.site_id = sample.site_id) **INNER JOIN** result **ON** sample.sample_id = result.sample_id) **INNER JOIN** param_codes **ON** result.param_code = param_codes.param_code **WHERE** (((param_codes.common_Name)=[Enter Parameter to Search]));

Query Results by Site Example

SELECT sites.site_id, sample.dt_tm, param_codes.common_Name, result.result_qualifier, result.result_value, result.result_units **FROM** ((sites **INNER JOIN** sample **ON** sites.site_id = sample.site_id) **INNER JOIN** result **ON** sample.sample_id = result.sample_id) **INNER JOIN** param_codes **ON** result.param_code = param_codes.param_code **WHERE** (((sites.site_id)=[Enter Site ID to Search]));

Query Results by Site and Parameter Example

SELECT sites.site_id, sample.dt_tm, param_codes.common_Name, result.result_qualifier, result.result_value, result.result_units **FROM** ((sites **INNER JOIN** sample **ON** sites.site_id = sample.site_id) **INNER JOIN** result **ON** sample.sample_id = result.sample_id) **INNER JOIN** param_codes **ON** result.param_code = param_codes.param_code **WHERE** (((sites.site_id)=[Enter Site ID to Search]) **AND** ((param_codes.common_Name)=[Enter Parameter to Search]));

Query Results by Parameter and Minimum Detection[1] Value Example

SELECT sites.site_id, sample.dt_tm, param_codes.common_Name, result.result_qualifier, result.result_value, result.result_units **FROM** ((sites **INNER JOIN** sample **ON** sites.site_id = sample.site_id) **INNER JOIN** result **ON** sample.sample_id = result.sample_id) **INNER JOIN** param_codes **ON** result.param_code = param_codes.param_code **WHERE** (((param_codes.common_Name)=[Enter Parameter to Search]) **AND** ((result.result_value)>[Enter Minimum Value]));

[1] Laboratory reporting level.

Appendix 5—FGDC-Compliant Metadata Record

Identification_Information:
 Citation:
  Citation_Information:
   Originator: Various sources.
   Publication_Date: 20090915
   Title: sites
   Geospatial_Data_Presentation_Form: vector digital data
   Online_Linkage: \\Igskiacwgsnas\gis_proj\b8653_dp6\data\Salinityv3.mdb
 Description:
  Abstract: The U.S. Geological Survey, in cooperation with the U.S. Army Corps of Engineers, compiled salinity-related water-quality data and information in a geodatabase containing more than 6,000 sampling sites. The geodatabase was designed as a tool for water-resource management and includes readily available digital data sources from the U.S. Geological Survey, U.S. Environmental Protection Agency, New Mexico Interstate Stream Commission, Sustainability of semi-Arid Hydrology and Riparian Areas, Paso del Norte Watershed Council, numerous other State and local databases, and selected databases maintained by the University of Arizona and New Mexico State University. Salinity information was compiled for an approximately 26,000-square-mile area of the Rio Grande Basin from the Rio Arriba-Sandoval County line, New Mexico, to Presidio, Texas. The geodatabase relates the spatial location of sampling sites with salinity-related water-quality data reported by multiple agencies. The sampling sites are stored in a geodatabase feature class; each site is linked by a relationship class to the corresponding sample and results stored in data tables.
  Purpose: These data are for informational purposes only. These data have not received Bureau approval and as such are provisional and subject to revision. The data are released on the condition that neither the U.S. Geological Survey, its cooperators, nor the U.S. Government may be held liable for any damages resulting from its authorized or unauthorized use. Although these data have been processed successfully on a computer system at the U.S. Geological Survey, no warranty expressed or implied is made regarding the accuracy or utility of the data on any other system or for general or scientific purposes, nor shall the act of distribution constitute any such warranty.
 Time_Period_of_Content:
  Time_Period_Information:
   Single_Date/Time:
    Calendar_Date: 2009
  Currentness_Reference: 1901–2009
 Status:
  Progress: On-going
  Maintenance_and_Update_Frequency: Unknown
 Spatial_Domain:
  Bounding_Coordinates:
   West_Bounding_Coordinate: -107.729762
   East_Bounding_Coordinate: -104.246110
   North_Bounding_Coordinate: 36.039700
   South_Bounding_Coordinate: 29.519630
 Keywords:
  Theme:

Theme_Keyword: salinity
Place:
Place_Keyword: Rio Grande Valley
Use_Constraints: These data are for informational purposes only. These data have not received Bureau approval and as such are provisional and subject to revision. The data are released on the condition that neither the U.S. Geological Survey, its cooperators, nor the U.S. Government may be held liable for any damages resulting from its authorized or unauthorized use. Although these data have been processed successfully on a computer system at the U.S. Geological Survey, no warranty expressed or implied is made regarding the accuracy or utility of the data on any other system or for general or scientific purposes, nor shall the act of distribution constitute any such warranty.
Native_Data_Set_Environment: Microsoft Windows XP Version 5.1 (Build 2600) Service Pack 3; ESRI ArcCatalog 9.3.1.3000
Data_Quality_Information:
Lineage:
Process_Step:
Process_Description: Sample information queries were written to identify potential duplicate data based on location and date. Samples were considered duplicate if the site identifier and sample date were the same. If an entire dataset was duplicated elsewhere, the original dataset was used and the duplicate excluded.
Process_Date: Unknown
Process_Step:
Process_Description: To examine duplicates based on proximity, all sites within 30 meters of another site were inspected. If a site was within the 30-meter buffer and was on top of or very close to a USGS site (well or gage), then the USGS site location was given priority and input into the geodatabase. Other methods such as examining well locations with digital orthophoto quarter quadrangles and USGS topographic maps were used to determine the accuracy of the spatial coordinates assigned by an agency or individual.
Spatial_Data_Organization_Information:
Direct_Spatial_Reference_Method: Vector
Point_and_Vector_Object_Information:
SDTS_Terms_Description:
SDTS_Point_and_Vector_Object_Type: Entity point
Point_and_Vector_Object_Count: 6397
Spatial_Reference_Information:
Horizontal_Coordinate_System_Definition:
Geographic:
Latitude_Resolution: 0.000000
Longitude_Resolution: 0.000000
Geographic_Coordinate_Units: Decimal degrees
Geodetic_Model:
Horizontal_Datum_Name: North American Datum of 1983
Ellipsoid_Name: Geodetic Reference System 80
Semi-major_Axis: 6378137.000000
Denominator_of_Flattening_Ratio: 298.257222
Entity_and_Attribute_Information:
Detailed_Description:
Entity_Type:

Entity_Type_Label: sites
Attribute:
Attribute_Label: OBJECTID
Attribute_Definition: Internal feature number.
Attribute_Definition_Source: ESRI
Attribute_Domain_Values:
Unrepresentable_Domain: Sequential unique whole numbers that are automatically generated.
Attribute:
Attribute_Label: SHAPE
Attribute_Definition: Feature geometry.
Attribute_Definition_Source: ESRI
Attribute_Domain_Values:
Unrepresentable_Domain: Coordinates defining the features.
Attribute:
Attribute_Label: site_id
Attribute_Definition: Site name.
Attribute:
Attribute_Label: site_description
Attribute:
Attribute_Label: type_code
Attribute_Definition: Site type.
Attribute:
Attribute_Label: lat_dd
Attribute_Definition: Reported site latitude in decimal degrees.
Attribute:
Attribute_Label: lon_dd
Attribute_Definition: Reported site longitude in decimal degrees.
Attribute:
Attribute_Label: h_datum
Attribute_Definition: Reported site horizontal datum.
Attribute:
Attribute_Label: alt_ft
Attribute_Definition: Reported site altitude (in feet).
Attribute:
Attribute_Label: v_datum
Attribute_Definition: Reported vertical datum.
Attribute:
Attribute_Label: site_comments
Attribute_Definition: Comments about the site.
Attribute:
Attribute_Label: source
Attribute_Definition: Source of the site data.
Attribute:
Attribute_Label: primary_site
Attribute_Definition: Primary site name.
Detailed_Description:
Entity_Type:

Entity_Type_Label: site2sample
Distribution_Information:
Resource_Description: Downloadable Data
Metadata_Reference_Information:
Metadata_Date: 20090929
Metadata_Contact:
Contact_Information:
Contact_Organization_Primary:
Contact_Organization: U.S. Geological Survey
Contact_Person: Public Information Officer
Contact_Address:
Address_Type: mailing and physical address
Address: 1505 Ferguson Lane
City: Austin
State_or_Province: Texas
Postal_Code: 78754
Country: USA
Contact_Voice_Telephone: 512-927-3500
Contact_Facsimile_Telephone: 512-927-3590
Contact_Electronic_Mail_Address: gs-w-txpublic-info@usgs.gov
Metadata_Standard_Name: FGDC Content Standards for Digital Geospatial Metadata
Metadata_Standard_Version: FGDC-STD-001-1998
Metadata_Time_Convention: local time
Metadata_Extensions:
Online_Linkage: http://www.esri.com/metadata/esriprof80.html
Profile_Name: ESRI Metadata Profile

Publishing support provided by
Lafayette Publishing Service Center

Information regarding water resources in Texas is available at
http://tx.usgs.gov/

USGS

Shah and Maltby—Design and Compilation of a Geodatabase of Existing Salinity Information for the Rio Grande Basin—DS 499

ISBN 978-1-4113-2853-2

Printed on recycled paper