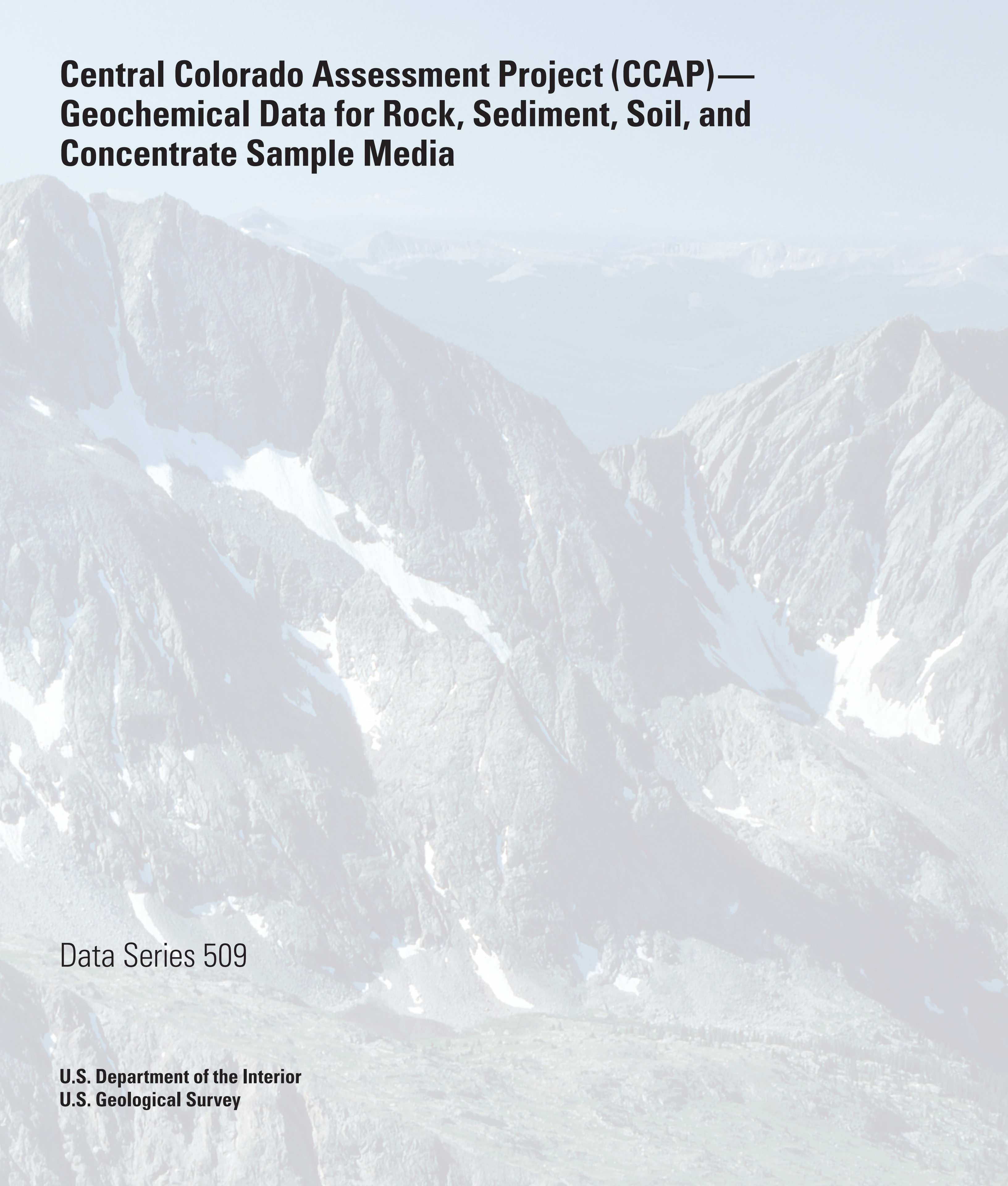
Central Colorado Assessment Project (CCAP)—Geochemical Data for Rock, Sediment, Soil, and Concentrate Sample Media



Data Series 509

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

Cover Photograph: View due south from Huron Peak to Ice Mountain (13,845-foot altitude), the central peak of the Three Apostles in the Collegiate Range. Pyramidal peak on far left skyline is Mount Aetna (13,648-foot altitude), north of Garfield. Pyramidal peak on skyline between the two summits of the Three Apostles is Emma Burr Mountain (13,533-foot altitude), east of Tincup. Taylor Park is north of the tree-covered area below the right skyline. Photograph by Ed DeWitt, 1984.

Central Colorado Assessment Project (CCAP)—Geochemical Data for Rock, Sediment, Soil, and Concentrate Sample Media

By Matthew Granitto, Ed H. DeWitt, and Terry L. Klein

Data Series 509

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

Suggested citation:
Granitto, Matthew, Klein, T.L., and DeWitt, E.H., 2010, Central Colorado Assessment Project (CCAP)—Geochemical data for rock, sediment, soil, and concentrate sample media: U.S. Geological Survey Data Series 509, 29 p. Available online at *http://pubs.usgs.gov/ds/2010/509*.

Contents

Abstract........1
Introduction........1
Geographic Setting........1
Methods of Study........2
    Sample Media and Collection........2
    Analytical Techniques........3
    Quality Assurance/Quality Control........3
Characteristics of the Relational Database........3
    Contents........3
    Structure........3
    Other Data Formats........6
    Database Query Examples........6
    Data Enhancement, Correction, and Processing........7
Acknowledgments........7
References Cited........7
    Database References........7
    References Cited........8
Appendix 1. Analytical Methods........11
Appendix 2. Database Query Examples........13
Appendix 3. Data Enhancement, Correction, and Processing........19

Figures

1. Geographic area covered by CCAP Geochemical Database........2
2. Table relationships in the CCAP Geochemical Database........5

Appendix Figures

A2–1. Summary query qsumSampleType in Query Design View........13
A2–2. Summary query qsumAnalyticMethod in Query Design View........14
A2–3. Select query qselNURE_orig_As in Query Design View........15
A2–4. Select query qselNURE_reanal_As in Query Design View........15
A2–5. Select query qselNURE_orig_VS_reanal_As in Query Design View........17
A2–6. Crosstab query qctabIgnRx_Geol_10MajorsChem in Query Design View........18

Tables

1. CCAP Geochemical Database tables........4
2. List of spreadsheets containing CCAP data........7

Appendix Table

A1–1. Analytical methods........11

Central Colorado Assessment Project (CCAP)—Geochemical Data for Rock, Sediment, Soil, and Concentrate Sample Media

By Matthew Granitto, Terry L. Klein, and Ed H. DeWitt

Abstract

This database was initiated, designed, and populated to collect and integrate geochemical data from central Colorado in order to facilitate geologic mapping, petrologic studies, mineral resource assessment, definition of geochemical baseline values and statistics, environmental impact assessment, and medical geology. The Microsoft Access database serves as a geochemical data warehouse in support of the Central Colorado Assessment Project (CCAP) and contains data tables describing historical and new quantitative and qualitative geochemical analyses determined by 70 analytical laboratory and field methods for 47,478 rock, sediment, soil, and heavy-mineral concentrate samples. Most samples were collected by U.S. Geological Survey (USGS) personnel and analyzed either in the analytical laboratories of the USGS or by contract with commercial analytical laboratories. These data represent analyses of samples collected as part of various USGS programs and projects. In addition, geochemical data from 7,470 sediment and soil samples collected and analyzed under the Atomic Energy Commission National Uranium Resource Evaluation (NURE) Hydrogeochemical and Stream Sediment Reconnaissance (HSSR) program (henceforth called NURE) have been included in this database. In addition to data from 2,377 samples collected and analyzed under CCAP, this dataset includes archived geochemical data originally entered into the in-house Rock Analysis Storage System (RASS) database (used by the USGS from the mid-1960s through the late 1980s) and the in-house PLUTO database (used by the USGS from the mid-1970s through the mid-1990s). All of these data are maintained in the Oracle-based National Geochemical Database (NGDB). Retrievals from the NGDB and from the NURE database were used to generate most of this dataset. In addition, USGS data that have been excluded previously from the NGDB because the data predate earliest USGS geochemical databases, or were once excluded for programmatic reasons, have been included in the CCAP Geochemical Database and are planned to be added to the NGDB.

Introduction

The U.S. Geological Survey (USGS) conducted the Central Colorado Assessment Project (CCAP) from 2003 to 2008 to provide Federal, State, and local land-management entities with geoscience information to address issues related to urban development in the Colorado Front Range Urban Corridor. Rapid population growth along the Front Range and in the CCAP region, which encompasses about 20,800 mi^2 (53,800 km^2) of central Colorado between the New Mexico and Wyoming borders, and from west of Denver to east of Vail Colorado, has caused tremendous demand for natural resources and has created challenging land-management issues related to the interface between wilderness and urban expansion. Regional CCAP data, including geologic, geochemical, geophysical, and remote-sensing data, were compiled to assess the availability of mineral resources, the geoenvironmental effects of historical mining and wildfires, geologic controls on groundwater resources, and hazards such as landslides and flooding.

The CCAP Geochemical Database described herein is the project's data warehouse for new and historical geochemical analyses of rock, sediment, soil, and concentrate samples. The database is useful to delineate geochemical baselines and the geochemical signatures of different types of mineral deposits and for exploration in previously mined areas. These data represent analyses of earth material samples collected in support of various USGS programs and projects, in support of the National Uranium Resource Evaluation (NURE) program, and in support of work investigations in a few postgraduate theses.

Geographic Setting

The geographic boundaries of the CCAP Geochemical Database enclose an area shaped like three overlapping rectangles. These boundaries are 41.12501 degrees latitude on the north, 36.87499 degrees latitude on the south, 104.74999 degrees longitude on the east, and 106.62501 degrees longitude on the west (fig. 1). In addition to central Colorado, small strips of northern New Mexico and southern Wyoming have been included for continuity in mapping.

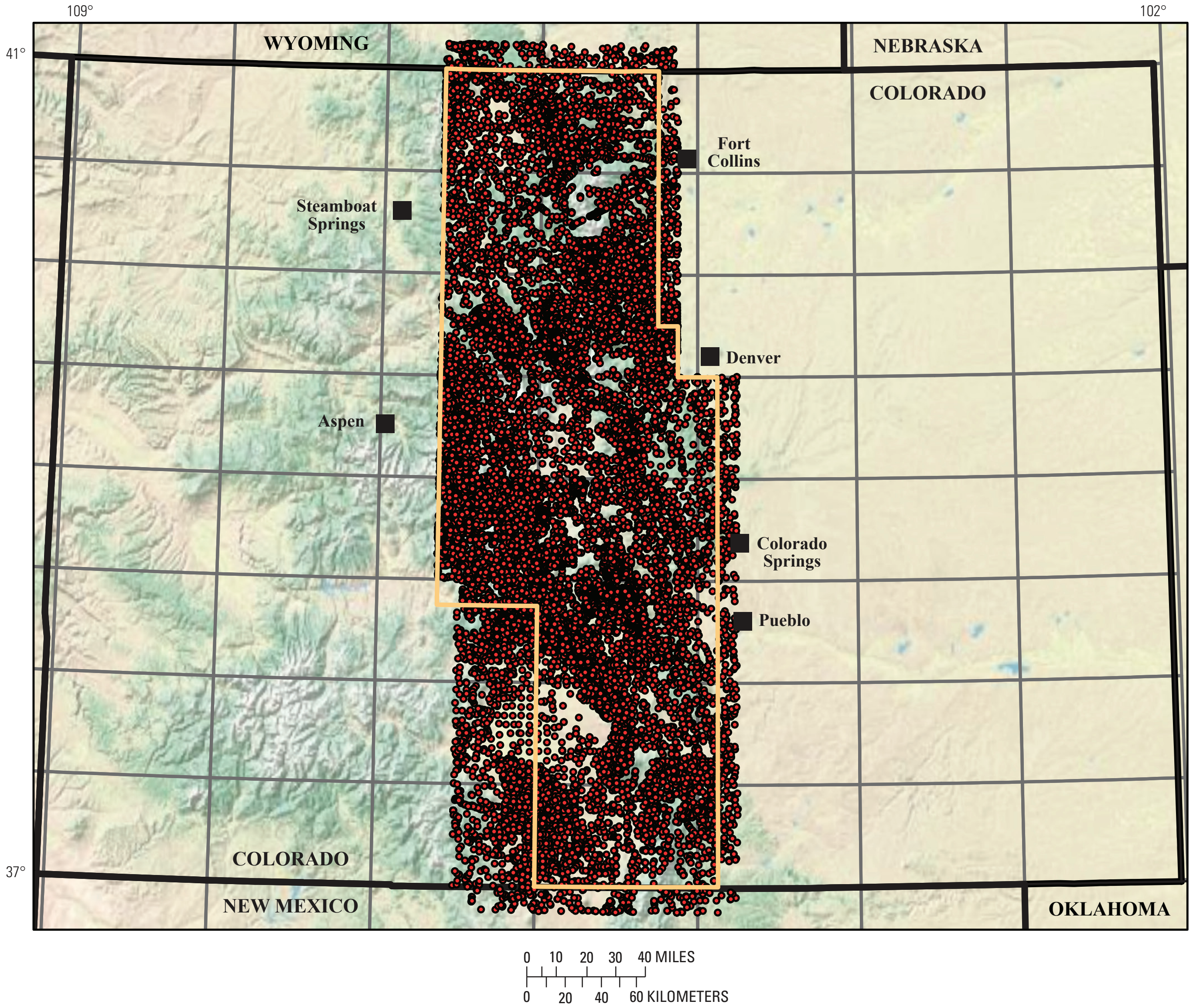


Figure 1. Geographic area covered by CCAP Geochemical Database. Sampled sites are red circles. Black circles represent the overlap and high density of sampled sites. Gray rectangles are the 1:100,000 topographic quadrangle outlines. The yellow line is the CCAP boundary. The approximate locations of major cities are shown by black squares. State boundaries are heavy black lines.

Methods of Study

Sample Media and Collection

In order to facilitate the use of geochemical data by the various geologic tasks within CCAP, analyses of 22,125 rock, 17,002 sediment, 4,637 soil, and 3,714 heavy mineral concentrate samples have been incorporated into the CCAP Geochemical Database presented here. The database includes new analyses of 694 rock and stream sediment samples that were collected and analyzed during CCAP and 979 NURE samples that were reanalyzed during CCAP (Caine and others, 2006; Klein and others, 2008; Van Gosen, 2008; Caine and Bove, 2010; and C.A. San Juan, D.L. Fey, T.S. Schmidt, T.L. Klein, and Ed DeWitt, written commun., 2010). Chemical data of collected and analyzed water samples and insect material samples that were separately studied under CCAP have been released in other publications (Fey and others, 2007; C.A. San Juan, D.L. Fey, T.S. Schmidt, T.L. Klein, and Ed DeWitt, written commun., 2010) and are not repeated in this database. Coal, humus, and peat samples have also been excluded from this database. Samples were collected between 1880 and 2007 and prepared according to a variety of USGS standard methods (variously described in Arbogast, 1990, 1996; Hillebrand, 1900, 1907, 1919; Miesch, 1976; Taggart, 2002) or by NURE methods (described in Smith, 1997).

Analytical Techniques

Seventy analytical field and laboratory methods were used to describe sample geochemistry. The various analytical methods reflect the evolution of analytical chemistry over time (from 1880 to 2007) and include methods such as bench-top wet chemistry techniques, atomic absorption, semiquantitative emission spectrography, inductively coupled plasma (ICP) atomic emission spectrometry, neutron activation, X-ray fluorescence, and ICP mass spectrometry. Refer to Appendix 1 for a complete list of the analytic methods and descriptive information. Refer to the **AnalyticMethod** table in the CCAP Geochemical Database for more detailed information about techniques and citations for analytic methods.

Quality Assurance/Quality Control

Data on quality assurance and quality control (QA/QC) are not incorporated into the CCAP Geochemical Database. Field sample-site duplicates and laboratory analytical replicates were left in the database. The USGS and contract laboratories use constituent standards and blanks for in-house QA/QC (Arbogast, 1990; Taggart, 2002). Data for these reference samples are not included in the database.

Characteristics of the Relational Database

Because of the scope and complexity of geologic materials data collected as part of the CCAP, a relational database structure was designed for data storage. The CCAP Geochemical Database (hereinafter called the database) was constructed in Microsoft Access 2003 as a tool to be used for data synthesis and analysis and as an archive of data collected during the study. The database structure and format are a modification of that used by the National Geochemical Database (NGDB) because more than 75 percent of the data was retrieved from the NGDB (Smith and others, 2003). This tabular relational database contains field site and sample measurements and observations and laboratory analyses of samples collected at point locations.

Contents

The database contains 14 tables, which are described in table 1. Two core tables consist of quantitative results, sample data, field site information, and geologic data. From these two relational datasets, nine analytical output data tables were created for various sections of data. Analytic method information and analytic method bibliography core-lookup tables provide needed reference for quantitative results, and a reference table of field name definitions was included to assist the user in understanding database field names and contents. In this report, names of tables and queries cited are in boldface; cited field names of tables and queries are italicized.

Structure

Data are contained in two core tables, **Geol_Data** and **Chem_Data**, and relationships are defined to link the tables (fig. 2). This structure provides efficient storage of information and provides for built-in data verification checks. For example, all valid results must have corresponding sample information. Relationships between these tables are depicted as lines in figure 2. **Geol_Data** is linked to **Chem_Data** by including a common field (*LAB_ID*) in both tables. Therefore, a chemical value cannot exist without having a corresponding sample in **Geol_Data**. The symbols “**1**” and “∞” at the ends of the relationship line indicate a one-to-many relationship; that is, a single sample may have many results. Data may be extracted from the database to meet specific user needs by constructing user-defined queries. Example queries are given in Appendix 2.

Relationships between **Geol_Data** and other tables in the database are shown in figure 2. **Geol_Data** contains 40 fields with information about the sample material collected at each site. Each analyzed sample has a unique *LAB_ID*, as well as a *FIELD_ID* that was provided by the sample collector. *LAB_ID* is a unique identifier assigned to each submitted sample by the analytical laboratory that received the sample. It is a key field that links the sample to its chemical and physical data found in **Chem_Data**. The dates of sample submission and collection are stored in the *DATE_SUBMITTED* and *DATE_COLLECT* fields; however, less than 10 percent of all samples have a collection date recorded. *LATITUDE* and *LONGITUDE* contain the geographic coordinate data, whose precision is set at five digits to the right of the decimal separator. Associated *SPHEROID* and *DATUM* information is also provided (see Appendix 3). *PRIMARY_CLASS* defines the sample material type, while *SECONDARY_CLASS* and *SPECIFIC_NAME* provide more detailed information about the sample medium. Media type should be carefully noted when assessing data so that data from different sample types are not mistakenly equated. For example, the database contains analyses for copper found in four different subsample media types (described in *SPECIFIC_NAME)* that were derived from one sediment sample site. Information regarding the collection and preparation of the sample may be found in *METHOD_COLLECTED*, *PREP*, and *MESH_PORE_SIZE*. Most of the entries in *LAB_ID* represent samples that were entered in the USGS laboratory information management system and whose data have been archived in the National Geochemical Database (NGDB) (Smith and others, 2003). Thus, the CCAP database can be linked to data within the NGDB by using the *LAB_ID* field.

Table 1. CCAP Geochemical Database tables.

Table name	Type	Description	Primary key field	Fields	Records
Geol_Data	Core	Spatial, geologic and descriptive attributes for samples whose chemical data were as compiled for CCAP	LAB_ID	41	47,478
Chem_Data[1]	Core	All chemical data compiled for CCAP	CHEM_ID	13	1,891,760
AnalyticMethod	Core-lookup	Analytic methods used to obtain chemical and physical data	ANALYTIC_METHOD	4	70
AnalyticMethodBiblio	Core-lookup	References for analytic methods used to obtain chemical data	ANALYTIC_METHOD_PUB_ID	9	702
Chem_HMC	Output	Chemical data for heavy mineral concentrate samples	LAB_ID	111	3,714
Chem_Rx_Majors	Output	Chemical “whole rock” data for rock samples	LAB_ID	122	21,157
Chem_Rx_Traces_Ag-Gd	Output	Trace-element data—silver through gadolinium—for rock samples	LAB_ID	125	20,205
Chem_Rx_Traces_Ge-Sb	Output	Trace-element data—germanium through antimony—for rock samples	LAB_ID	125	20,499
Chem_Rx_Traces_Sc-Zr	Output	Trace-element data—scandium through zirconium—for rock samples	LAB_ID	127	21,179
Chem_Sed_Ag-Mg	Output	Chemical data—silver through magnesium—for sediment samples	LAB_ID	179	16,713
Chem_Sed_Mn-Zr	Output	Chemical data—manganese through zirconium—for sediment samples	LAB_ID	183	16,815
Chem_Soil_Ag-Mg	Output	Chemical data—silver through magnesium—for soil samples	LAB_ID	170	4,545
Chem_Soil_Mn-Zr	Output	Chemical data—manganese through zirconium—for soil samples	LAB_ID	168	4,620
FieldNameDictionary	Reference	Field name descriptions for all tables in the CCAP database	FIELD_NAME	5	658

[1] Table not included in Excel spreadsheet files.

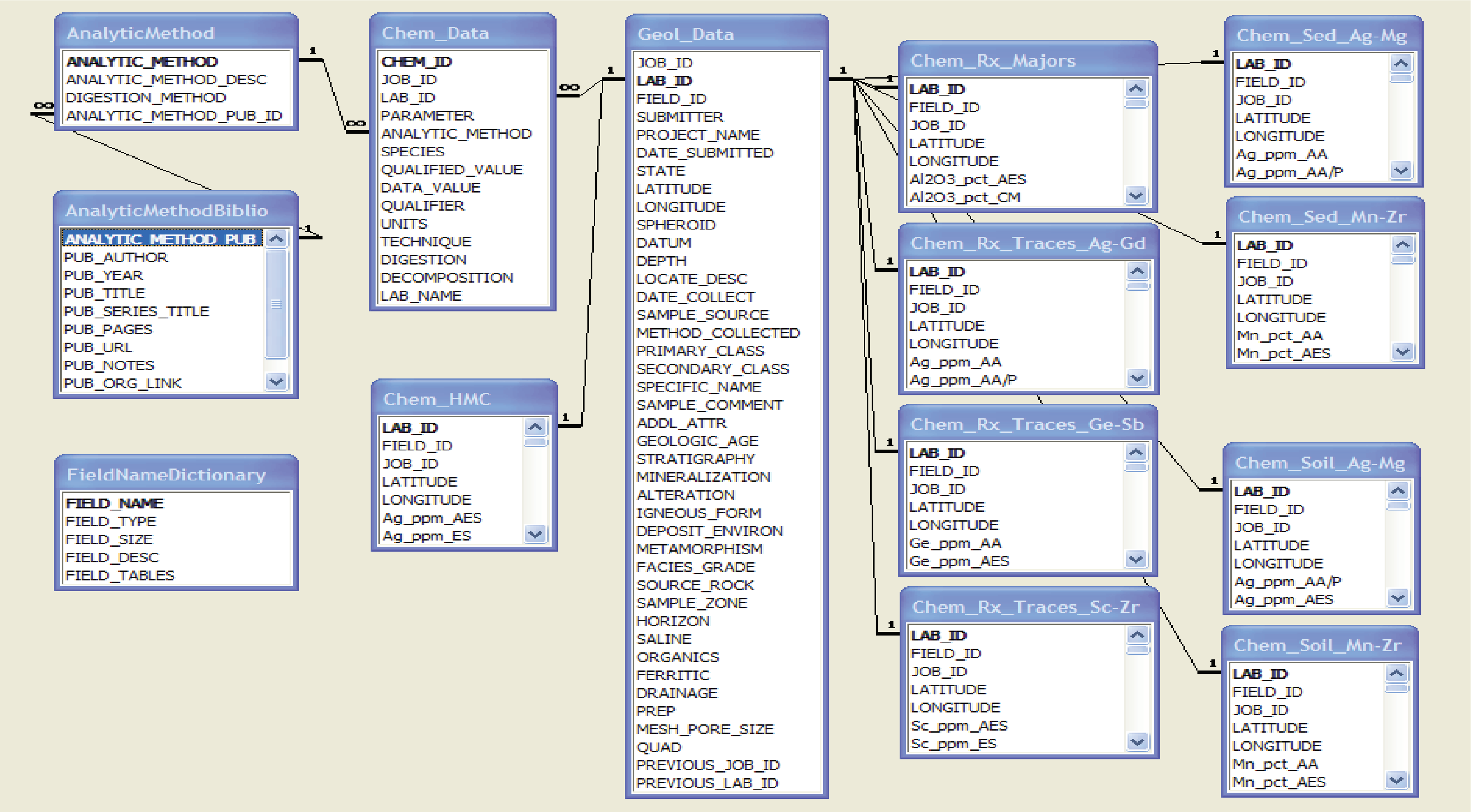


Figure 2. Table relationships in the CCAP Geochemical Database.

Chem_Data contains 13 fields with laboratory and field measurements, expressed as numeric values. *CHEM_ID* is a unique identifier assigned to each measurement in the table and is a key field of software-assigned integers. Measurements in **Chem_Data** consist of a numeric *DATA_VALUE* and an optional *QUALIFIER*, which is used to describe results such as nondetectable or estimates based on limits of instrumental detection (for example, "less than" values, such as <2). *QUALIFIER* entries are "<" or "N," meaning that the element was not detected at concentrations above the lower limit of determination for the method; "L," meaning that the element was detected, but at concentrations below the lower limit of determination for the method; and ">" or "G," meaning that the element was measured at a concentration greater than the upper limit of determination for the method. *QUALIFIED_VALUE* was populated by combining the data in *DATA_VALUE* with its complement in *QUALIFIER*, according to the following conventions: *DATA_VALUE* entries that are accompanied by "<," "N," or "L" entries in *QUALIFIER* are represented in *QUALIFIED_VALUE* as negative numbers (for example, "–2"); and *DATA_VALUE* entries that are accompanied by ">" or "G" entries in *QUALIFIER* are represented in *QUALIFIED_VALUE* as values with 0.00999, 0.09999 or 0.99999 added to them (for example, >0.25 becomes 0.25999, >0.5 becomes 0.59999, and >10 becomes 10.99999). Further, measured characteristics such as units and techniques are identified using a *PARAMETER* code, which is a concatenation of data from the fields *SPECIES*, *UNITS*, *TECHNIQUE*, and *DIGESTION*. There are 559 unique parameters in the CCAP database. For example, the parameter "Sb_ppm_AA/P" represents the concentration of antimony, expressed in parts per million, as detected by atomic absorption spectrometry after a partial digestion. *PARAMETER* is a succinct 35-character-length field that can be used as a column name in a data report or spreadsheet. *PARAMETER* entries are used as the field and column names of the output tables and are described in the **FieldNameDictionary** table. Information regarding the method of analysis or measurement used to obtain data is found in *ANALYTIC_METHOD*, an abbreviated label in the **Chem_Data** table linked to the **AnalyticMethod** table—a look-up table that provides additional information on the 70 field and laboratory techniques used for analysis of the various geologic materials. The table includes a description of the analytic methods and relevant published references to them and is linked by *ANALYTIC_METHOD_PUB_ID* to bibliographic reference information in **AnalyticMethodBiblio**. Further information regarding method of analysis is found in *DECOMPOSITION*. *LAB_NAME* provides information regarding the laboratory or work group responsible for the analysis. Relationships between **Chem_Data** and other tables in the database are shown in figure 2.

To facilitate ease of use, nine "ChemData" output tables have been created from the database. Each table contains a subset of analytical results for the analysis of a specific sample media. For example, **Chem_Rx_Majors** contains chemical and physical data from rock samples and major elements expressed as oxide concentrations for the "whole rock." In other output tables, major elements are not presented as oxide concentrations. The **Chem_Data** table contains the original major-element data as received from analytical laboratories. Due to the 256-field limit of Access and the 256-column limit of Excel, three of these datasets have been further subdivided. For example, the trace-element data for rock samples are found in three output tables that separate the analyzed elements alphabetically (Ag to Zr). For sediment, soil, and heavy metal concentrate samples, the major elements are not expressed as oxides but as elemental concentrations. The concentration values in all output tables are in the *QUALIFIED_VALUE* format described previously. These output tables have results for each parameter in their own columns (crosstab format) for ease of analysis in Excel. Relationships between the output tables and **Geol_Data** are shown in figure 2. The table **FieldNameDictionary** contains the field name, size, definition, and general data type of the 658 fields that are used in the tables of the database, as well as the table or tables in which these fields appear. This is of particular importance for the nondatabase user as it also contains the descriptions of field names of the nine output chemical data tables.

Other Data Formats

All of the Access tables in the database except **Chem_Data** were exported into Excel as 13 spreadsheets for use by the nondatabase user. **Chem_Data** was excluded because all of its data are presented in the nine analytical ChemData output tables. Table 2 lists the spreadsheet files included in the data release.

The Access tables also are exported to tab-delimited ASCII flat-file form and may be accessed using any text editor. The Access tables, however, are best used by loading each flat file into a relational database and reestablishing the links as shown in the accompanying relationship diagram (fig. 2).

Relational databases can be implemented using a variety of proprietary or nonproprietary software packages. The database is attached to this report in a proprietary (Microsoft Office Access 2003) and nonproprietary (ASCII tab-delimited) format. The spreadsheets are presented in Microsoft Office Excel 2003 format.

Database Query Examples

Within relational database software packages, queries may be constructed and saved to retrieve data using user-defined criteria. This database contains several examples of Access queries that aid the user in viewing and extracting selected datasets. The graphical Query Design Views of these

Table 2. List of spreadsheets containing CCAP data.

Spreadsheet name	Information contained in Spreadsheet
Geol_Data.xls	Spatial, geologic and descriptive attributes for samples whose chemical data were compiled for CCAP
AnalyticMethod.xls	Analytic methods used to obtain chemical and physical data
AnalyticMethodBiblio.xls	References for analytic methods used to obtain chemical data
Chem_HMC.xls	Chemical data for heavy mineral concentrate samples
Chem_Rx_Majors.xls	Chemical “whole rock” data for rock samples
Chem_Rx_Traces_Ag-Gd.xls	Trace-element data—silver through gadolinium—for rock samples
Chem_Rx_Traces_Ge-Sb.xls	Trace-element data—germanium through antimony—for rock samples
Chem_Rx_Traces_Sc-Zr.xls	Trace-element data—scandium through zirconium—for rock samples
Chem_Sed_Ag-Mg.xls	Chemical data—silver through magnesium—for sediment samples
Chem_Sed_Mn-Zr.xls	Chemical data—manganese through zirconium—for sediment samples
Chem_Soil_Ag-Mg.xls	Chemical data—silver through magnesium—for soil samples
Chem_Soil_Mn-Zr.xls	Chemical data—manganese through zirconium—for soil samples
FieldNameDictionary.xls	Explanation of field names used for all tables in the CCAP database

queries are translated into Structured Query Language (SQL) statements that are displayed within the database by clicking on the SQL View of the View box of the Query. Examples of three query types—summary, select, and crosstab queries—are presented in Appendix 2.

Data Enhancement, Correction, and Processing

Data retrieved from the NGDB were generated by the analytical laboratories of the USGS over a number of decades, beginning in the early 1960s. Upon completion of the chemical analyses, the data were stored in the RASS database (1963 to 1987), the PLUTO database (1979 to 1997), or the specific commercial laboratory-information-management software (LIMS) used by USGS analytical laboratories (1996 to present). Beginning in 2002, data from all three sources were combined, reformatted, and standardized into the Oracle-based NGDB. The NGDB is composed of three data tables: **GEO** consisting of sample site, collection, and description information; **CHEM** consisting of sample analysis information; **JOB** consisting of sample submitter information. All three tables were queried to produce most of the CCAP database.

To create the CCAP database, rock, sediment, soil, and heavy mineral concentrate sample data were retrieved from the NGDB by using the following criteria: (1) each sample must have a valid and unique laboratory identification number (lab ID); (2) each sample must have latitude and longitude coordinates; and (3) each sample must be identified as a one of these four geologic materials. This dataset was then examined to remove any samples that could be identified as a processed derivative of these media. Single minerals, mineral separates, rock coatings, insoluble residues, partial digestions, leachates, experimental or artificial samples, and some misidentified samples were thereby eliminated. An effort was made to fix incorrect or incomplete attributes. Several standardized sample descriptive fields were more completely populated using information previously found only in the comment field of PLUTO, the free coding fields of RASS, or the paper sample submittal forms. The geologic material sample-site location and descriptive data were repackaged into a single table called **Geol_Data**. Through this process, 4,806 samples in the CCAP area have been added to the NGDB, and thousands more have better coordinates as a result of the CCAP data cleanup effort. A detailed explanation of the data enhancement, correction, and processing plan that was used to produce the CCAP database is presented in Appendix 3.

Acknowledgments

The authors thank the following USGS geoscientists for their assistance in providing data from the NGDB and NURE databases: Dorothy M. Trujillo, Steven M. Smith, Robert G. Eppinger, and Richard B. Wanty. Many thanks to CCAP team member and GIS expert Carma A. San Juan for her encouragement and thoughtful discussion while this publication was in process. The authors acknowledge Richard M. O’Leary and Robert G. Eppinger for their technical review of the manuscript and database; their expertise and assistance were invaluable in improving the work.

References Cited

Database References

Within the database, references are cited for analytic methods that were used to determine elemental concentrations in the **Chem_Data** table. Information regarding the method of analysis or measurement used to obtain data is found in **AnalyticMethodBiblio** (table 1). Refer to the *ANALYTIC_METHOD_PUB_ID*—in most cases the USGS Library call number—when researching the analytic method in question.

References Cited

Arbogast, B.F., 1990, Quality assurance manual for the Branch of Geochemistry, U.S. Geological Survey: U.S. Geological Survey Open-File Report 90–668, 184 p.

Arbogast, B.F., ed., 1996, Analytical methods manual for the Mineral Resource Surveys Program, U.S. Geological Survey: U.S. Geological Survey Open-File Report 96–525, 248 p.

Caine, J.S., and Bove, D.J., 2010, Rock geochemistry and mineralogy from fault zones and polymetallic fault veins of the central Front Range, Colorado: U.S. Geological Survey Data Series 492, 6 p.

Caine, J.S., Manning, A.H., Verplanck, P.L., Bove, D.J., Kahn, K.G., and Ge, S., 2006, Well construction information, lithologic logs, water level data, and overview of research in the Handcart Gulch, Colorado—An alpine watershed affected by metalliferous hydrothermal alteration: U.S. Geological Survey Open-File Report 2006–1189, v. 1.0, accessed November 2, 2009, at *http://pubs.usgs.gov/of/2006/1189/*.

Fey, D.L., Church, S.E., Schmidt, T.S., Wanty, R.B., Verplanck, P.L., Lamothe, P.J., Adams, M.G., and Anthony, M.W., 2007, Selected field parameters from streams and preliminary analytical data from water and macroinvertebrate samples, Central Colorado Assessment Project, environmental assessment task, 2004 and 2005: U.S. Geological Survey Open-File Report 2007–1044, 41 p., v. 1.0, accessed November 2, 2009, at *http://pubs.usgs.gov/of/2007/1044/*.

Hillebrand, W.F., 1900, Some principles and methods of rock analysis: U.S. Geological Survey Bulletin 174, 114 p.

Hillebrand, W.F., 1907, The analysis of silicate and carbonate rocks: U.S. Geological Survey Bulletin 305, 200 p.

Hillebrand, W.F., 1919, The analysis of silicate and carbonate rocks: U.S. Geological Survey Bulletin 700, 285 p.

Klein, T.L., Church, S.E., Caine, J.S., Schmidt, T.S., and DeWitt, Ed, 2008, Central Colorado Assessment Project—Application of integrated geologic, geochemical, biologic, and mineral resource studies: U.S. Geological Survey Fact Sheet 2008–3084, 4 p., v 1.0, accessed November 2, 2009, at *http://pubs.usgs.gov/fs/2008/3084/pdf/FS08-3084_508.pdf*.

Miesch, A.T., 1976, Sampling designs for geochemical surveys—Syllabus for a short course: U.S. Geological Survey Open-File Report 76–772, 128 p.

Neuendorf, K.K.E., Mehl, J.P., Jr., and Jackson, J.A., eds., 2005, Glossary of geology (5th ed.): Alexandria, Va., American Geological Institute, 779 p.

Smith, S.M., 1997, National Geochemical Database reformatted data from the National Uranium Resource Evaluation (NURE) Hydrogeochemical and Stream Sediment Reconnaissance (HSSR) Program: U.S. Geological Survey Open-File Report 97–492, v. 1.4, accessed November 2, 2009, at *http://pubs.usgs.gov/of/1997/ofr-97-0492/*.

Smith, S.M., Smith, D.B., and Bailey, E.A., 2003, The U.S. Geological Survey's National Geochemical Database [abs.]; Geological Society of America Abstracts with Program, v. 35, no. 7, p. 366.

Streckeisen, A.L., 1967, Classification and nomenclature of igneous rocks (final report of an inquiry): Neues Jahrbuch für Mineralogie, Abhandlungen, v. 107, no. 2, p. 144–214.

Streckeisen, A.L., 1976, To each plutonic rock its proper name: Earth Science Review, v. 12, p. 1–33.

Taggart, J.E., ed., 2002, Analytical methods for chemical analysis of geologic and other materials, U.S. Geological Survey: U.S. Geological Survey Open-File Report 2002–0223, v. 5.0, accessed November 2, 2009, at *http://pubs.usgs.gov/of/2002/ofr-02-0223/*.

U.S. Geological Survey, 2004, The National Geochemical Survey, database and documentation: U.S. Geological Survey Open-File Report 2004–1001, v. 5.0, accessed November 2, 2009, at *http://pubs.usgs.gov/of/2004/1001/*.

Van Gosen, B.S., 2008, Geochemistry of rock samples collected from the Iron Hill Carbonatite Complex, Gunnison County, Colorado: U.S. Geological Survey Open-File Report 2008–1119, 27 p., v. 1.0, accessed November 2, 2009, at *http://pubs.usgs.gov/of/2008/1119/*.

Appendixes

Appendix 1. Analytical Methods

Appendix 1 contains a table of analytical method names and descriptions of the analytical techniques that provided the chemical data of the CCAP Geochemical database.

Table A1–1. Analytical method.

Analytic method	Description
AA/P_fusion-MIBK	molybdenum and antimony by flame atomic absorption spectrometry after potassium pyrosulfate fusion, partial acid digestion, and methyl butyl isoketone extraction.
AA/P_H2O	calcium, magnesium, manganese, and arsenic in saturation paste of soil by flame atomic absorption spectrometry after water extraction.
AA/P_HCl	copper and manganese by flame atomic absorption spectrometry after partial digestion with hydrochloric acid.
AA/P_HCl-MIBK	silver, arsenic, bismuth, cadmium, copper, lead, antimony, tin, and zinc by flame atomic absorption spectrometry after digestion with hydrochloric acid and selective organic extraction with 336- methyl butyl isoketone.
AA/P_HNO3	silver, cadmium, copper, lead, and zinc by flame atomic absorption spectrometry after partial digestion with hot nitric acid.
AA_CVAA	mercury by cold-vapor atomic absorption spectrometry after multiacid digestion and solution.
AA_FAA_fusion	major and minor elements by flame atomic absorption spectrometry after fusion digestion.
AA_FAA_HBr	silver, gold, and tellurium by flame atomic absorption spectrometry after hydrogen bromide/bromine digestion and methyl butyl isoketone extraction.
AA_FAA_HF	major and minor elements by flame atomic absorption spectrometry after multiacid digestion with hydrofluoric acid.
AA_GFAA_HBr	gold and tellurium by graphite furnace atomic absorption spectrometry after hydrogen bromide/bromine digestion and methyl butyl isoketone extraction.
AA_GFAA_HF	arsenic, gold, bismuth and tellurium by graphite furnace atomic absorption spectrometry after multiacid digestion with hydrofluoric acid and extraction.
AA_HGAA_fusion	arsenic, antimony, and selenium by flow injection or continuous flow hydride generation–atomic absorption spectrometry after fusion digestion.
AA_HGAA_HF	arsenic, antimony, selenium, and tellurium by flow injection or continuous flow hydride generation–atomic absorption spectrometry after multiacid digestion with hydrofluoric acid.
AA_TR	mercury by thermal release and atomic absorption spectrometry after multiacid digestion (Vaughn-McCarthy method).
AES/P_AR	major and minor elements by inductively coupled plasma–atomic emission spectrometry after partial digestion with aqua regia.
AES/P_A-Z	silver, arsenic, gold, bismuth, cadmium, copper, molybdenum, lead, antimony, and zinc by inductively coupled plasma–atomic emission spectrometry after partial digestion with hydrochloric acid/hydrogen peroxide.
AES/P_H2O	boron by inductively coupled plasma–atomic emission spectrometry after hot water extraction.
AES/P_leach	major and minor elements by inductively coupled plasma–atomic emission spectrometry after weak multiacid leach.
AES_acid	major and minor elements by inductively coupled plasma–atomic emission spectrometry after digestion with hydrofluoric acid/hydrochloric acid/nitric acid/perchlorate.
AES_fusion	major and minor elements by inductively coupled plasma–atomic emission spectrometry after fusion digestion.
AES_sinter	major and minor elements by inductively coupled plasma–atomic emission spectrometry after sinter digestion.
CB	ash or loss on ignition by weight loss after heating at 900°C.
CB_IRC	carbon and sulfur by infrared detection after combustion.
CB_TC	total carbon and organic carbon by thermal conductivity detection after combustion.
CB_TT	sulfur by iodometric titration after combustion.
CD	specific conductance by standard method conductivity electrode.
CM/P_As	arsenic by modified Gutzeit apparatus confined-spot method colorimetry after partial digestion in potassium hydroxide/hydrochloric acid and chemical separation.
CM/P_CX	copper by colorimetry after partial extraction in cold hydrochloric acid.
CM/P_fusion	molybdenum and antimony by colorimetry after fusion (Mo) or potassium hydroxide fusion-hydrochloric acid digestion (Sb, rhodamine B).
CM/P_HM	heavy metal elements by colorimetry after partial extraction in aqueous ammonium citrate solution.
CM/P_HNO3	copper, lead, and zinc by colorimetry after partial digestion with nitric acid.
CM/P_PC	uranium by paper chromatography after partial digestion with nitric acid.
CM/P_W	tungsten by colorimetry after carbonate sinter digestion.
CM_Cl	chloride by colorimetric spectrophotometry after sodium carbonate/zinc oxide sinter digestion.

Table A1–1. Analytical method.—Continued

Analytic method	Description
CM_F	fluorine by colorimetric spectrophotometry after acid digestion and chemical separation.
CM_fusion	major and minor elements by colorimetric spectrophotometry after fusion digestion.
CM_HF	major and minor elements by colorimetric spectrophotometry after multiacid digestion with hydrofluoric acid.
CP	organic carbon, carbonate carbon, and totals by computation.
ES_Q	major and minor elements by quantitative emission spectrography.
ES_SQ	major and minor elements by semiquantitative emission spectrography.
ES_SWR	major and minor elements by short-wavelength-region-response quantitative emission spectrography.
FA_AA	gold, silver and platinum group elements by graphite furnace atomic absorption spectrometry after lead oxide fire assay chemical separation.
FA_DCP	gold by direct current plasma–atomic emission spectroscopy or atomic absorption spectrophotometry after lead oxide fire assay.
FA_ES	gold and platinum group elements by direct-current arc quantitative emission spectrography after lead oxide fire assay chemical separation.
FA_MS	platinum group elements by inductively coupled plasma–mass spectrometry after nickel sulfide fire assay.
FL	beryllium, selenium, tin, and uranium by fluorometry after multiacid digestion with hydrofluoric acid.
GRC	uranium by gamma counting.
GV	density, moisture, and weight by gravimetry.
GV_acid	major and minor elements by gravimetry after acid digestion.
GV_CR	major and minor elements by gravimetry for Classical Rock Analysis after unknown digestion method.
GV_flux	moisture, bound water, and total water by heating and weight loss with flux.
GV_fusion	major and minor elements by gravimetry after fusion digestion.
IC	chloride, fluoride, nitrate, sulfate, and phosphate by ion chromatography.
IC/P	chloride, fluoride, nitrate, sulfate, and phosphate by ion chromatography after solution extraction.
ISE_fusion	chloride, fluoride, and iodide by ion-specific electrode after fusion digestion.
ISE_pH	pH by standard method combination pH electrode.
MS_acid	major and minor elements by inductively coupled plasma–mass spectrometry after hydrofluoric acid/hydrochloric acid/nitric acid/perchlorate digestion.
MS_sinter	major and minor elements by inductively coupled plasma–mass spectrometry after sodium peroxide sinter digestion.
NA_DN	uranium and thorium by delayed neutron activation counting.
NA_INAA	major and minor elements by instrumental neutron activation analysis.
TB_AR	acid-soluble sulfate, sulfur, and sulfide by turbidimetry after aqua regia digestion.
TB_NP	thorium by nephelometric turbidimetry; digestion unknown but could be hydrofluoric acid.
TT_flux	total water by Karl Fischer coulometric titration with flux after combustion.
TT_fusion	ferric oxide by fusion after decomposition and precipitation.
TT_HClO4	carbonate carbon and carbon dioxide (acid-soluble carbon) by coulometric titration after perchlorate digestion and extraction.
TT_HF	ferrous oxide by colorimetric or potentiometric titration after hydrofluoric acid/sulfuric acid digestion.
VOL	carbon dioxide or carbonate carbon by evolution after acid decomposition; also known as “gasometric” or “manometric.”
XRF_ED	minor elements by energy-dispersive X-ray fluorescence spectrometry.
XRF_WD_fusion	major and minor elements by wavelength-dispersive X-ray fluorescence spectrometry after lithium borate fusion digestion.
XRF_WD_raw	chlorine, iodine, and bromine by wavelength-dispersive X-ray fluorescence spectrometry on raw sample.

Appendix 2. Database Query Examples

Summary Queries

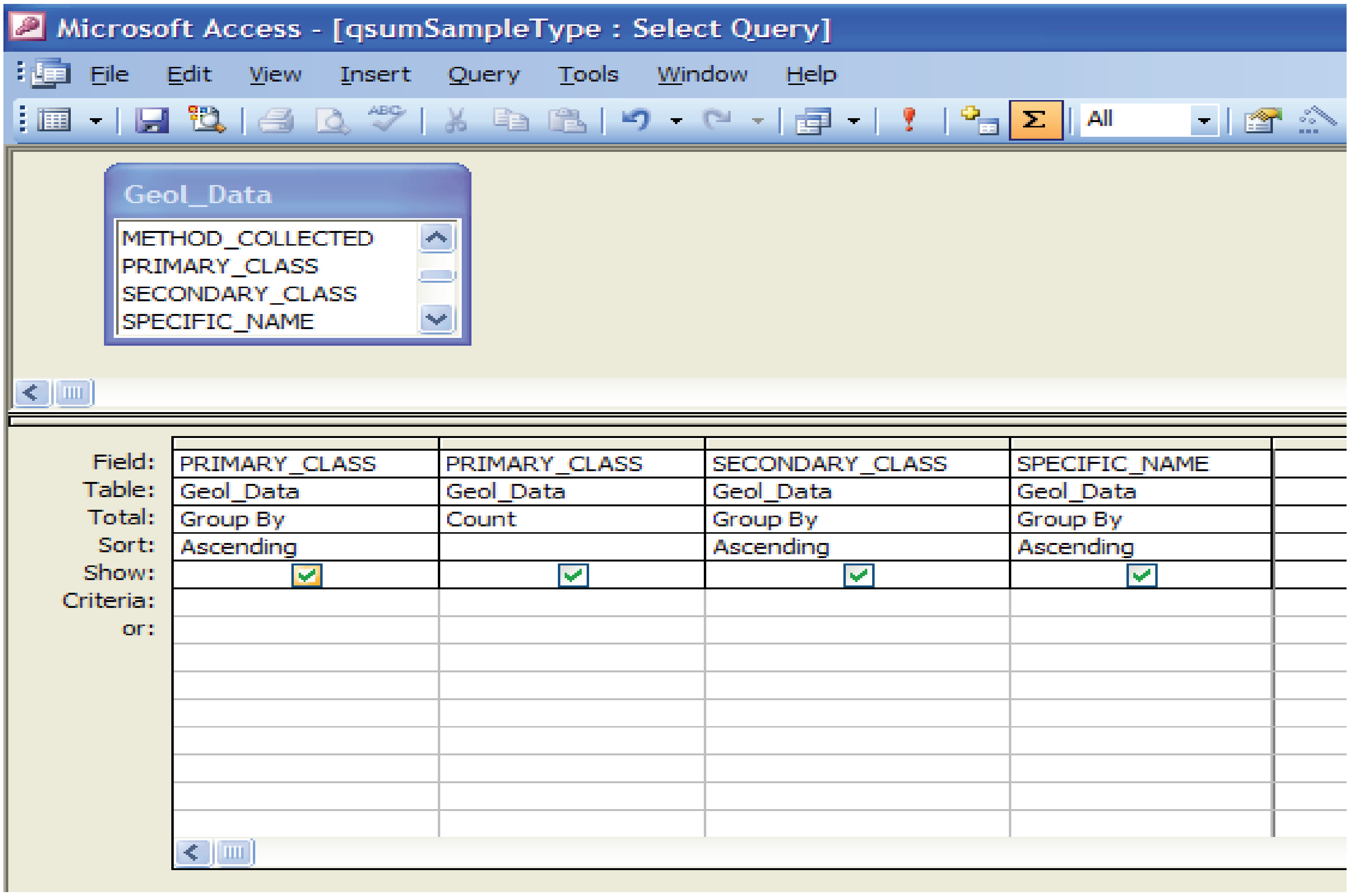
The query **qsumSampleType** was created to display the various types and subtypes of sample media that were collected and analyzed. This query displays data from the *PRIMARY_CLASS*, *SECONDARY_CLASS*, and *SPECIFIC_NAME* fields of the **Geol_Data** table, while providing a count of the number of samples in each Specific Name subtype (fig. A2–1). These data are sorted by fields in the order listed above. When the query is run, the Datasheet View shows 245 different *PRIMARY_CLASS-SECONDARY_CLASS-SPECIFIC_NAME* combinations—8 for the *PRIMARY_CLASS* category “concentrate,” 132 for the *SECONDARY_CLASS* category “igneous rock,” 73 for the *SECONDARY_CLASS* category “metamorphic rock,” and 30 for the *SECONDARY_CLASS* category “sedimentary rock.”

The query **qsumAnalyticMethod** was created to provide a breakdown of all chemical parameters and their analytic methods that are in the CCAP database. This query displays data from *ANALYTIC_METHOD* and *ANALYTIC_METHOD_DESC* in **AnalyticMethod** and *PARAMETER* in **Chem_Data** and provides a count of the number of chemical determinations for each parameter (fig. A2–2). These data are sorted first by *ANALYTIC*_METHOD and then by *PARAMETER*. When the query is run, the Datasheet View shows 844 unique *ANALYTIC_METHOD-PARAMETER* combinations.

Select Queries

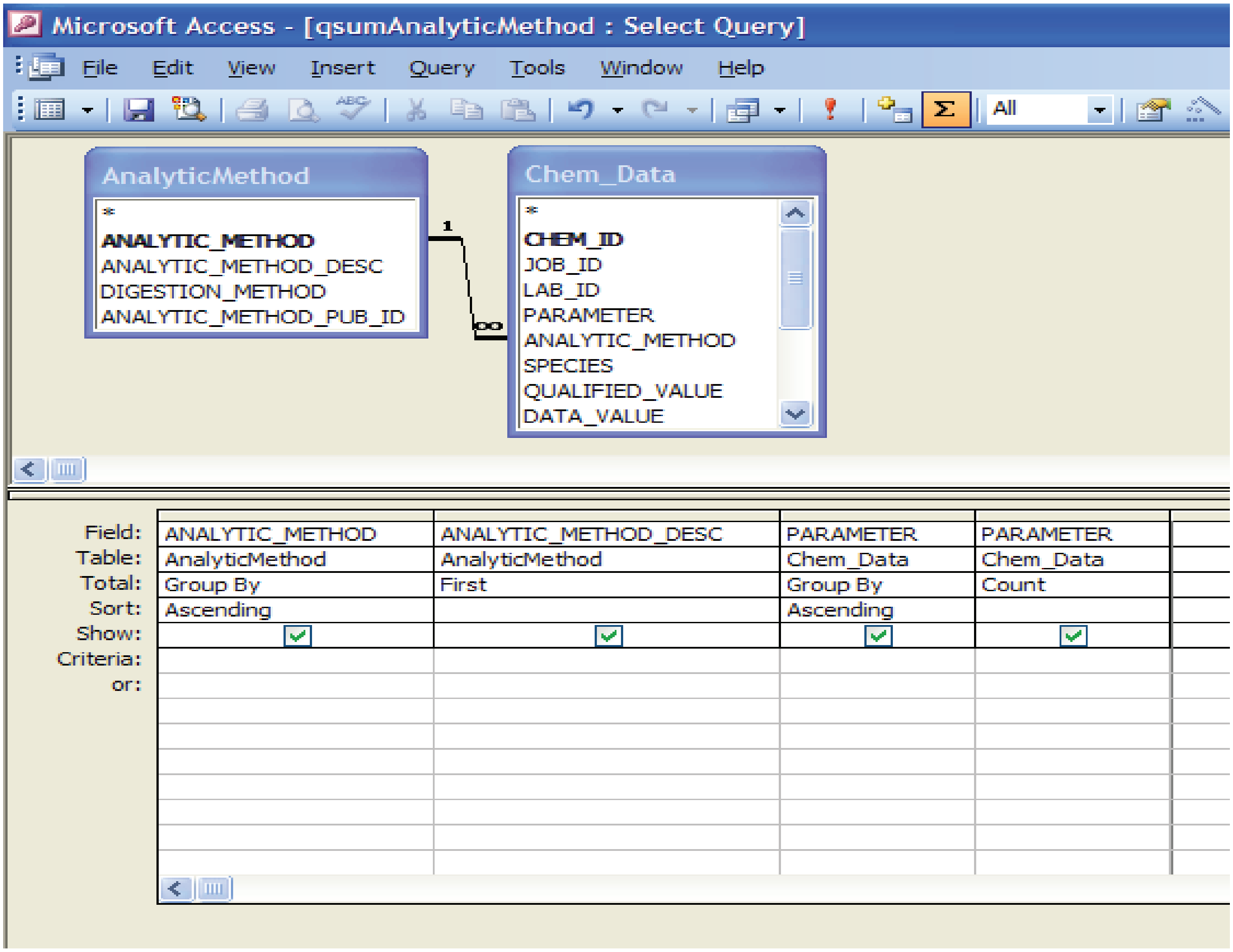
A series of queries was constructed to select specific data for sediment samples in the database. The select query **qselNURE_orig_As** was created to show all of the NURE samples that were originally analyzed for arsenic (As), among other elements (fig. A2–3).

The query combines data from two tables: **Geol_Data** (containing *LAB_ID*, *FIELD_ID*, *PROJECT_NAME*, *LATITUDE*, *LONGITUDE*, *DATE_SUBMITTED*, and *PRIMARY_CLASS* data) and **Chem_Data** (containing



Field:	PRIMARY_CLASS	PRIMARY_CLASS	SECONDARY_CLASS	SPECIFIC_NAME
Table:	Geol_Data	Geol_Data	Geol_Data	Geol_Data
Total:	Group By	Count	Group By	Group By
Sort:	Ascending		Ascending	Ascending
Show:	☑	☑	☑	☑
Criteria:				
or:				

Figure A2–1. Summary query qsumSampleType in Query Design View.



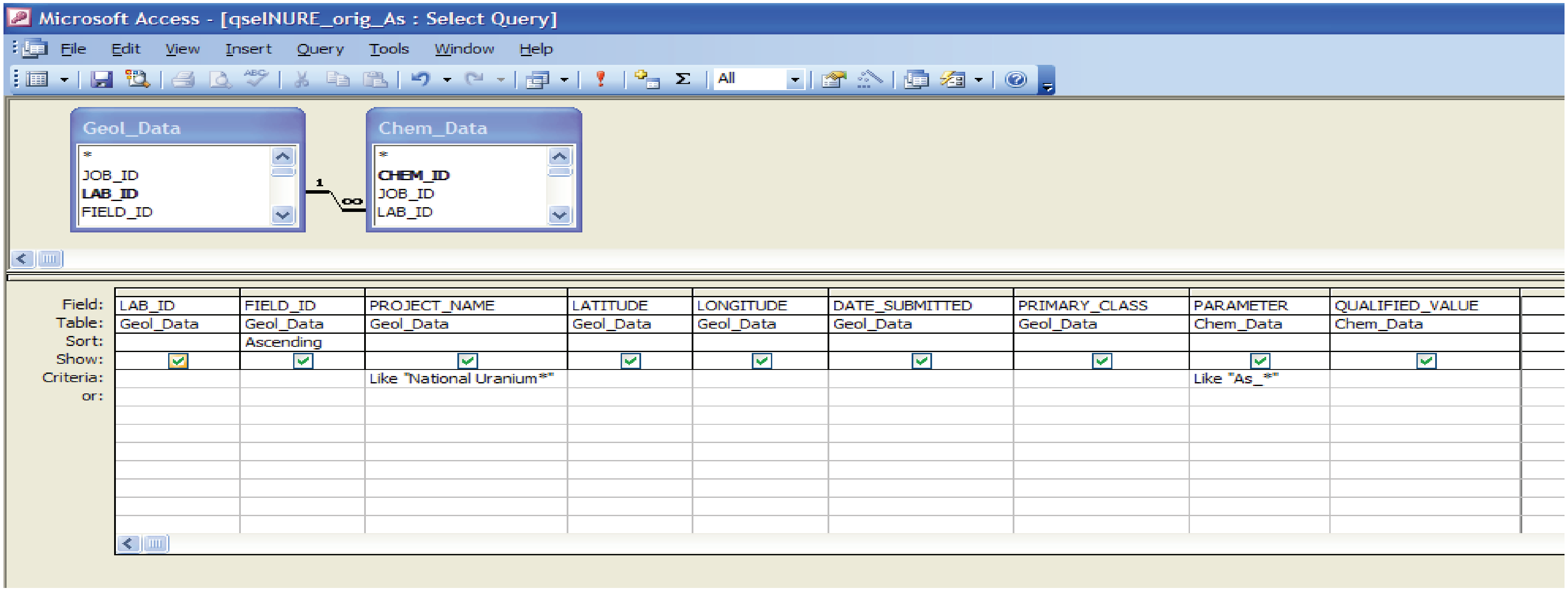
Field:	ANALYTIC_METHOD	ANALYTIC_METHOD_DESC	PARAMETER	PARAMETER
Table:	AnalyticMethod	AnalyticMethod	Chem_Data	Chem_Data
Total:	Group By	First	Group By	Count
Sort:	Ascending		Ascending	
Show:	☑	☑	☑	☑
Criteria:				
or:				

Figure A2–2. Summary query qsumAnalyticMethod in Query Design View.

PARAMETER and *QUALIFIED_VALUE* data). The tables are linked by the common field *LAB_ID*. A close look at the “Criteria” line of the Query Design View shows the conditions placed on this query: the phrase “National Uranium*” for *PROJECT_NAME* of **Geol_Data** identifies NURE samples; the phrase “As_*” for *PARAMETER* of **Chem_Data** identifies arsenic determinations. The data will be sorted by *FIELD_ID*. When the query is run, the Datasheet View shows 2,928 sediment samples that were analyzed for arsenic by X-ray fluorescence. From another select query of **Geol_Data**, we can learn that there are 7,470 NURE samples in the CCAP database, so less than one-third of these samples have arsenic data.

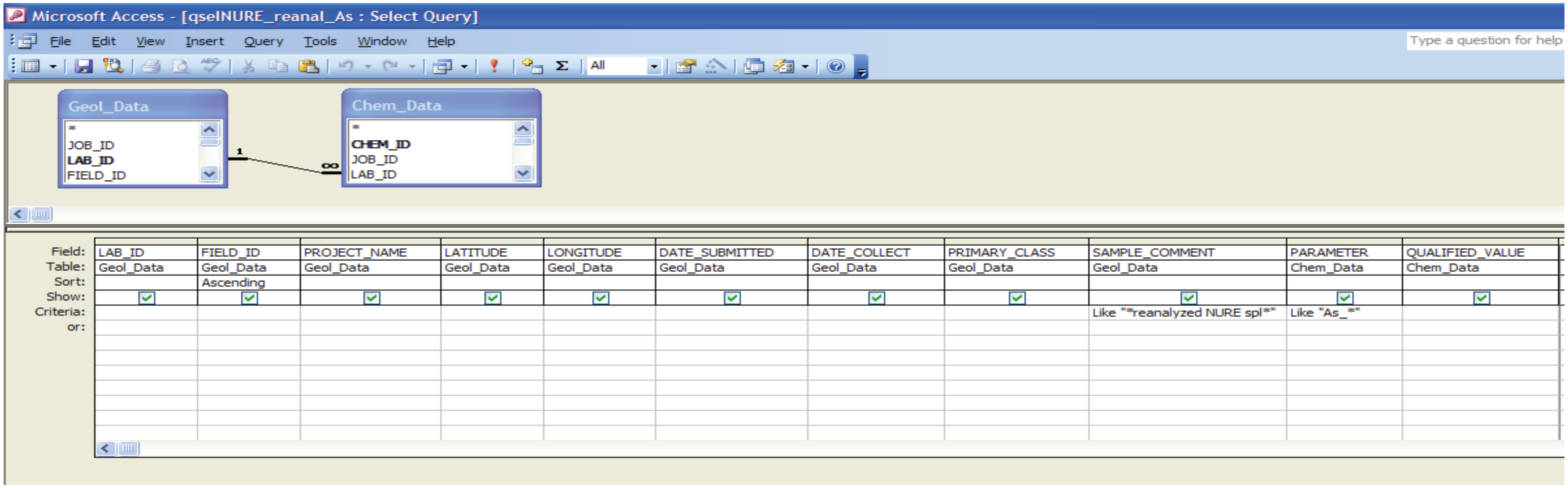
In much the same manner, the select **qselNURE_reanal_As** was created to show all of the NURE samples that were reanalyzed for arsenic among other elements (fig. A2–4).

The query also combines data from **Geol_Data** and **Chem_Data** and includes the additional fields *DATE_COLLECT* and *SAMPLE_COMMENT* added from **Geol_Data.** The tables are again linked by the common field *LAB_ID*. Search criteria for this query: the phrase “*reanalyzed NURE spl*” for *SAMPLE_COMMENT* identifies reanalyzed NURE samples; the phrase “As_*” for *PARAMETER* identifies arsenic determinations. When the query is run, the Datasheet View shows 1,525 arsenic values for reanalyzed NURE sediment samples. Summary query **qsumNURE_reanal_As_FieldID** was created and queries *FIELD_ID* of **qselNURE_reanal_As** and shows that 1,203 NURE samples were reanalyzed for arsenic—273 of them by more than one method. Six field IDs retrieved appear to have been analyzed four times, and they represent six samples and six sample duplicates—each analyzed twice for arsenic. Summary query **qsumNURE_reanal_As_Parameter** was created that queries *PARAMETER* of **qselNURE_reanal_As** and shows that three analytic methods were used to reanalyze these NURE sediments for arsenic. Because atomic absorption (AA) and inductively coupled plasma–mass spectrometry (MS) are the methods that have the lowest detection limits for arsenic, values gained from these methods would be preferred if two or more values for arsenic are found.



Field:	LAB_ID	FIELD_ID	PROJECT_NAME	LATITUDE	LONGITUDE	DATE_SUBMITTED	PRIMARY_CLASS	PARAMETER	QUALIFIED_VALUE
Table:	Geol_Data	Geol_Data	Geol_Data	Geol_Data	Geol_Data	Geol_Data	Geol_Data	Chem_Data	Chem_Data
Sort:		Ascending							
Show:	☑	☑	☑	☑	☑	☑	☑	☑	☑
Criteria:			Like "National Uranium*"					Like "As_*"	
or:									

Figure A2–3. Select query qselNURE_orig_As in Query Design View.



Field:	LAB_ID	FIELD_ID	PROJECT_NAME	LATITUDE	LONGITUDE	DATE_SUBMITTED	DATE_COLLECT	PRIMARY_CLASS	SAMPLE_COMMENT	PARAMETER	QUALIFIED_VALUE
Table:	Geol_Data	Geol_Data	Geol_Data	Geol_Data	Geol_Data	Geol_Data	Geol_Data	Geol_Data	Geol_Data	Chem_Data	Chem_Data
Sort:		Ascending									
Show:	☑	☑	☑	☑	☑	☑	☑	☑	☑	☑	☑
Criteria:									Like "*reanalyzed NURE spl*"	Like "As_*"	
or:											

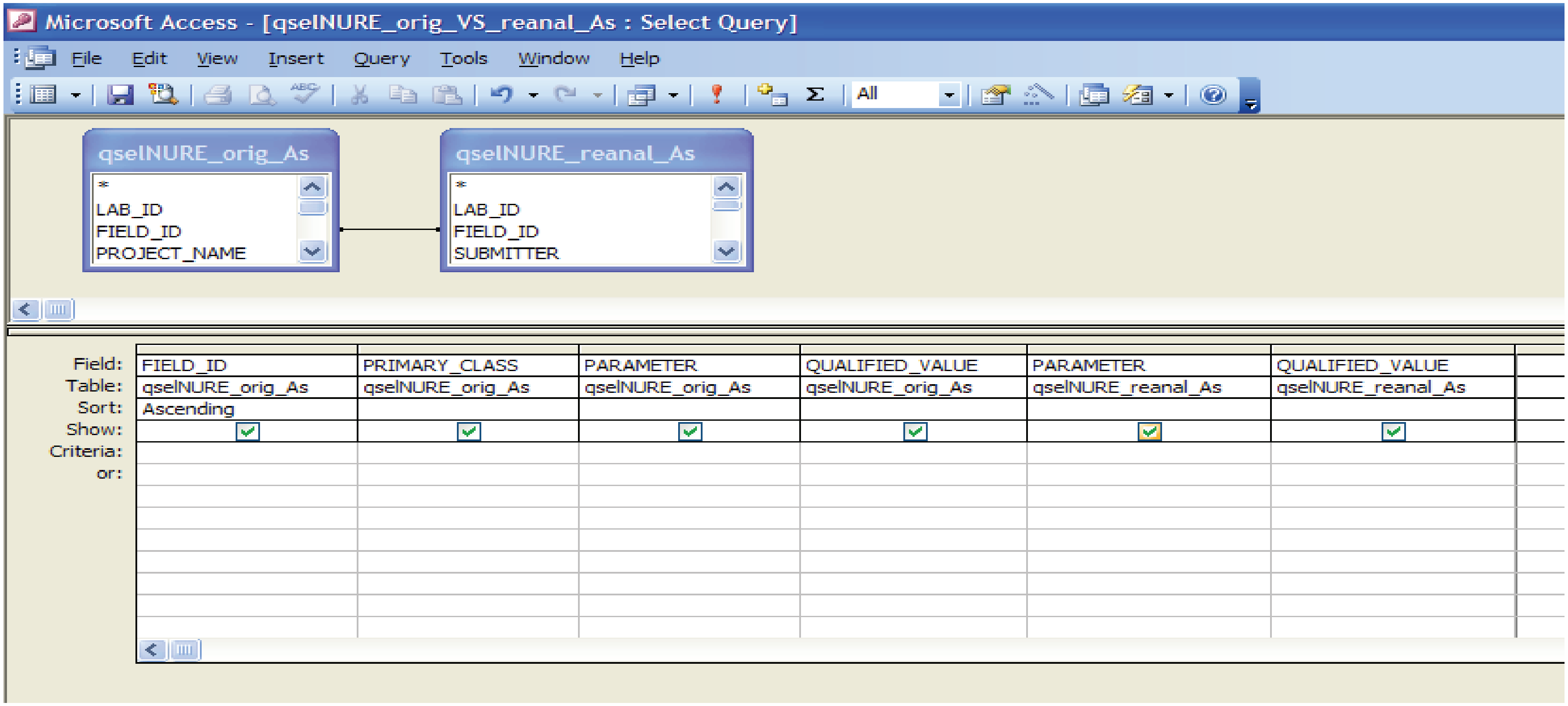
Figure A2–4. Select query qselNURE_reanal_As in Query Design View.

Select query **qselNURE_orig_VS_reanal_As** was created to compare the original NURE arsenic data for sediment samples with the reanalyzed NURE sample arsenic data from the same samples (fig. A2–5). Select queries **qselNURE_orig_As** and **qselNURE_reanal_As** are themselves queried, linked by the common field *FIELD_ID*, which combines data from *PRIMARY_CLASS*, *PARAMETER*, and *QUALIFIED_VALUE* in **qselNURE_orig_As**, and *PARAMETER* and *QUALIFIED_VALUE* from **qselNURE_reanal_As**. Running the query shows the Datasheet View with 483 arsenic values of original NURE data compared with data from reanalyzed NURE sediment samples. Summary query **qsumNURE_orig_VS_reanal_As** was created that queries *FIELD_ID* of **qselNURE_reanal_As** and shows that 400 NURE samples were originally analyzed for arsenic, and 73 reanalyzed samples had two or more values for arsenic.

Crosstab Query

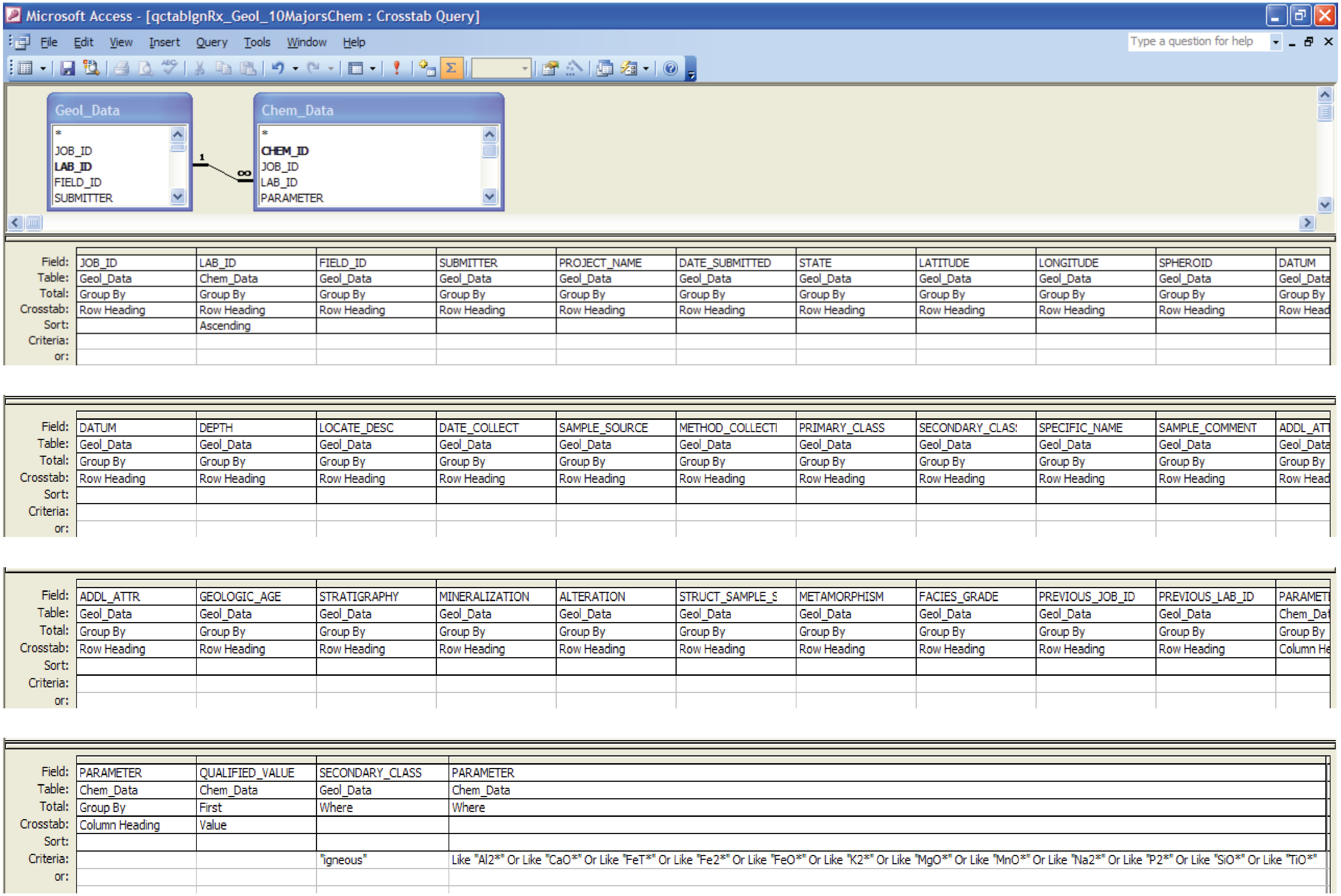
Note that in the select query **qselNURE_orig_As**, *QUALIFIED_VALUE* was used, which combines the determined value in *DATA_VALUE* and its qualifier in *QUALIFIER*. Likewise, *PARAMETER* was used rather than *SPECIES* and *UNITS* so that analytic method information could be included with the information regarding determined species and units of expression for all results. These two fields are critical in the construction of the crosstab query **qctabIgnRx_Geol_10MajorsChem**, which further aids the user by displaying the data in a flat file or spreadsheet view (fig. A2–6).

This crosstab query was constructed using **Chem_Data** so that *LAB_ID* is the key row heading, the unique entries in *PARAMETER* become the column headings, and the cells in each column are filled by the entries in *QUALIFIED_VALUE*. Fields providing sample identification, submittal, location, collection, and description were also added as row headings to the query from **Geol_Data** so that relevant descriptive information would be available in one Datasheet View. The "Criteria" line of the Query Design View shows the conditions placed on this query: the phrase "igneous" for *SECONDARY_CLASS* of **Geol_Data** identifies igneous rock samples; the phrase "" Like "Al2*" Or Like "CaO*" Or Like "FeT*" Or Like "Fe2*" Or Like "FeO*" Or Like "K2*" Or Like "MgO*" Or Like "MnO*" Or Like "Na2*" Or Like "P2*" Or Like "SiO*" Or Like "TiO*"" for *PARAMETER* of **Chem_Data** identifies major-element oxide data. The data retrieved will be sorted by *LAB_ID*. Running **qctabIgnRx_Geol_10MajorsChem** produces a Datasheet View containing all available major-element data, expressed as oxide percentage, as well as all geologic and geographic data for 3,204 igneous rock samples. Like the nine ChemData tables provided in the database, crosstab queries can be constructed to create unique datasets containing analytical data gathered from a specific sample media collected within specific geographic locales and determined by certain analytic methods.



Field:	FIELD_ID	PRIMARY_CLASS	PARAMETER	QUALIFIED_VALUE	PARAMETER	QUALIFIED_VALUE
Table:	qselNURE_orig_As	qselNURE_orig_As	qselNURE_orig_As	qselNURE_orig_As	qselNURE_reanal_As	qselNURE_reanal_As
Sort:	Ascending					
Show:	☑	☑	☑	☑	☑	☑
Criteria:						
or:						

Figure A2–5. Select query qselNURE_orig_VS_reanal_As in Query Design View.



Field:	JOB_ID	LAB_ID	FIELD_ID	SUBMITTER	PROJECT_NAME	DATE_SUBMITTED	STATE	LATITUDE	LONGITUDE	SPHEROID	DATUM
Table:	Geol_Data	Chem_Data	Geol_Data	Geol_Data	Geol_Data	Geol_Data	Geol_Data	Geol_Data	Geol_Data	Geol_Data	Geol_Data
Total:	Group By	Group By	Group By	Group By	Group By	Group By	Group By	Group By	Group By	Group By	Group By
Crosstab:	Row Heading	Row Heading	Row Heading	Row Heading	Row Heading	Row Heading	Row Heading	Row Heading	Row Heading	Row Heading	Row Head
Sort:		Ascending									
Criteria:											
or:											

Field:	DATUM	DEPTH	LOCATE_DESC	DATE_COLLECT	SAMPLE_SOURCE	METHOD_COLLECTI	PRIMARY_CLASS	SECONDARY_CLAS	SPECIFIC_NAME	SAMPLE_COMMENT	ADDL_ATT
Table:	Geol_Data	Geol_Data	Geol_Data	Geol_Data	Geol_Data	Geol_Data	Geol_Data	Geol_Data	Geol_Data	Geol_Data	Geol_Data
Total:	Group By	Group By	Group By	Group By	Group By	Group By	Group By	Group By	Group By	Group By	Group By
Crosstab:	Row Heading	Row Heading	Row Heading	Row Heading	Row Heading	Row Heading	Row Heading	Row Heading	Row Heading	Row Heading	Row Head
Sort:											
Criteria:											
or:											

Field:	ADDL_ATTR	GEOLOGIC_AGE	STRATIGRAPHY	MINERALIZATION	ALTERATION	STRUCT_SAMPLE_S	METAMORPHISM	FACIES_GRADE	PREVIOUS_JOB_ID	PREVIOUS_LAB_ID	PARAMET
Table:	Geol_Data	Geol_Data	Geol_Data	Geol_Data	Geol_Data	Geol_Data	Geol_Data	Geol_Data	Geol_Data	Geol_Data	Chem_Dat
Total:	Group By	Group By	Group By	Group By	Group By	Group By	Group By	Group By	Group By	Group By	Group By
Crosstab:	Row Heading	Row Heading	Row Heading	Row Heading	Row Heading	Row Heading	Row Heading	Row Heading	Row Heading	Row Heading	Column He
Sort:											
Criteria:											
or:											

Field:	PARAMETER	QUALIFIED_VALUE	SECONDARY_CLASS	PARAMETER
Table:	Chem_Data	Chem_Data	Geol_Data	Chem_Data
Total:	Group By	First	Where	Where
Crosstab:	Column Heading	Value		
Sort:				
Criteria:			"igneous"	Like "Al2*" Or Like "CaO*" Or Like "FeT*" Or Like "Fe2*" Or Like "FeO*" Or Like "K2*" Or Like "MgO*" Or Like "MnO*" Or Like "Na2*" Or Like "P2*" Or Like "SiO*" Or Like "TiO*"
or:				

Figure A2–6. Crosstab query qctablgnRx_Geol_10MajorsChem in Query Design View. The query involves many fields and is shown as four panels read left to right and top down.

Appendix 3. Data Enhancement, Correction, and Processing

Sample Identification

LAB_ID in the NGDB is the unique key field of the **GEO** table, despite the fact that the USGS has had a number of analytical chemistry laboratories, sample control groups, and computer data-processing projects associated with it over the years. *LAB_ID* is also a field in the NGDB **CHEM** table, grouping unique chemical determinations by the sample analyzed. There is potential duplication of *LAB_IDs* for pre-1965 samples because they were submitted for analysis at USGS laboratories in Washington, D.C., Denver, Colo., and Menlo Park, Calif., at a time when there was no perceived need for concern for *LAB_ID* uniqueness because a unified USGS geochemical database had not been developed. Special attention to this fact is required when samples from these older datasets are to be included in the NGDB. Very rarely, *LAB_ID* needed to be corrected (by the NGDB manager) because of archival data-entry error. This identification stamp was automatically computer generated in the PLUTO database and continues today through the Sample Submittal Form (SSF) that provides internal data entry for the NGDB and LIMS. The SSF has undergone a series of revisions to improve efficiency and user facility and currently is being redesigned. As the current (2010) version of the SSF provides accurate data to the NGDB, it will not be discussed in detail here. In addition, a new *LAB_ID* series—YL######—has been created in the NGDB to facilitate the inclusion of relevant datasets that were not issued standard *LAB_IDs* by Sample Control work groups at any work center. There are 1,139 of these added samples in the CCAP database that have these new *LAB_IDs*.

JOB_ID is the laboratory batch identifier assigned by the analytical laboratory that received the samples as a batch. In the NGDB, *JOB_ID* is the unique key field of the **JOB** table, linked to the *JOB_ID* fields in the **GEO** and **CHEM** tables. A given job may have between one and several hundred samples. There is duplication of entries in *JOB_ID* between the job number series of RASS and PLUTO archival databases that has been remedied by the NGDB database manager, enforcing uniqueness in this data field. The following examples illustrate this: old Branch of Analytical Laboratories-Washington, DC PLUTO *JOBNUMBER* “5534” containing Colorado samples is now called *JOB_ID* “W5534” in the NGDB, in order to avoid confusion with old Branch of Exploration Research-Denver RASS *JOB_ID* “5534”, now called *JOB_ID* “HM5534” in the NGDB, which contains Alaska samples; old Branch of Analytical Laboratories-Washington, DC PLUTO *JOBNUMBER* “11403” containing Colorado samples is now called *JOB_ID* “X1403” in the NGDB, in order to avoid confusion with old Branch of Exploration Research-Denver RASS *JOB_ID* “11403”, which contains Alaska samples; old Branch of Analytical Laboratories-Denver PLUTO *JOBNUMBER* “1136” containing Colorado samples is now called *JOB_ID* “Z1136” in the NGDB in order to avoid confusion with old Branch of Exploration Research-Denver RASS *JOB_ID* “1136” (now called *JOB_ID* “HM1136” in the NGDB) containing Georgia samples, or with old Branch of Analytical Laboratories-Washington, DC PLUTO *JOBNUMBER* “1136” (now called *JOB_ID* “W1136” in the NGDB) containing Idaho samples; old Branch of Analytical Laboratories-Denver PLUTO *JOBNUMBER* “M287” containing Colorado samples is not to be confused with old Branch of Analytical Laboratories-Menlo Park PLUTO *JOBNUMBER* “M287”, now called *JOB_ID* “M0287” in the NGDB, containing Alaska samples. Rarely, *JOB_ID* needed to be corrected (by the NGDB manager) because of archival data-entry error. Job number was automatically computer generated in the PLUTO database and continues today through the LIMS for the NGDB database. In addition, a new *JOB_ID* series—YJ#####—has been created in the NGDB to facilitate the inclusion of relevant datasets that were not issued standard *JOB_IDs* by Sample Control.

Unique identifiers have been assigned for 7,470 samples that have been retrieved from the NURE database and added to the CCAP database. The prefix “NURE” has been added to the sample’s *Prime_ID* (Primary Laboratory Sample Identification Number) of the NURE database. *Prime_ID* contains the final reported sample identification number that was assigned to each sample by the responsible Department of Energy (DOE) laboratory. The NURE *Prime_ID* then becomes the sample’s CCAP *FIELD_ID*. In addition, 1,285 of these NURE samples that have been reanalyzed by the USGS are included in the CCAP database. They have been assigned standard *LAB_ID* and *JOB_ID* identifiers and their reanalysis data stored in the NGDB. Their *FIELD_ID* entries have been corrected so that the NURE *Prime_ID* is now the sample’s *FIELD_ID*, allowing for easy comparison between original analyses and reanalyses of these NURE samples. This also facilitates the linking of the NGDB to the NURE database because the NURE *Prime_ID* is always populated in the NGDB *FIELD_ID* for all reanalyzed NURE samples, so *Prime_ID* can be linked to *FIELD_ID* in an Access query.

FIELD_ID is a field identifier assigned by the sample collector of the sample submitted for analysis and commonly has been modified by the data renovator due to the truncation of the original data entry. *FIELDNO* in the RASS database was an 8-character field, and many submitters did not follow this field size rule when creating their field numbers, so many field numbers were truncated by data-entry personnel. At times, this truncation created many identical field numbers in RASS that were in fact different from one another. For example, field numbers “87ABD001S” and “87ABD001C3” truncated to eight characters would be “87ABD001,” which is actually the field site number. A bulk stream sediment sample (suffix “S”) and a C3 fraction concentrate from a stream sediment sample (suffix “C3”) were taken at this site, but truncation obscures this fact. These sorts of field number errors have been corrected so that *FIELD_ID* entries match the field numbers listed in the Request for Analysis (RFA) paperwork that is on file at the NGDB Hard Copy Archives.

Various data-entry formats were used by RASS and PLUTO data-entry groups. A format has been applied to the CCAP *FIELD_ID* data whereby dashes or spaces between alpha and numeric characters, all slashes (/), and all underscores (_) have been removed. Spaces between two numeric characters or between two alpha characters have been replaced with a dash (-). Occasionally, a series of field numbers from a single field party were edited so that they could be usefully sorted. For example, "87ABD1" would become "87ABD001" so that it would appear in an ascending sort before "87ABD011" or "87ABD101." By implementing standard format and sorting criteria, all samples collected at one field site are grouped and, where one sample has accurate coordinates, provide an effective way of assigning coordinates to samples that have none or that have corner coordinates of topographic quadrangle maps instead of true coordinates. Alternate field numbers provided on RFA coding sheets have been entered in *SAMPLE_COMMENT*. Many NURE sediment and soil samples have been resubmitted for reanalysis by the USGS. These samples have been entered into the RASS, PLUTO, and NGDB databases, but there has been no standard data entry for these resubmittals, and as a result, there is no consistency for the entry of *FIELD_ID*. These samples have been corrected so that *FIELD_ID* now contains the *PRIME_ID* from the NURE database, and any alternate IDs have been entered in *SAMPLE_COMMENT*. This correction creates a link between the NGDB and NURE databases on the fields *FIELD_ID* and *PRIME_ID*, respectively. RASS and PLUTO samples that have been resubmitted into RASS, PLUTO, or the NGDB for further analysis are commonly assigned new *LAB_IDs*. In these cases, the original *LAB_ID* and *FIELD_ID* information may have been incorrectly entered. This has been remedied so that the original *LAB_ID* has been entered in *PREVIOUS_LAB_ID*, and the field number has been entered in *FIELD_ID* so that *FIELD_ID* remains a field containing data that are generated by the sample submitter.

The field *PREVIOUS_LAB_ID* contains the original *LAB_ID* of a resubmitted sample that has been given a new lab number upon resubmittal for further analysis. The *PREVIOUS_LAB_ID* is also entered into the NGDB by way of the SSF, but it is not a required field for data entry, which creates problems in the NGDB as one sample site may now have multiple samples, some of which are actually the same sample. RASS and PLUTO samples that have been resubmitted into RASS, PLUTO, or the NGDB for further analysis are commonly assigned new *LAB_IDs*. In these cases, the original *LAB_ID* and *FIELD_ID* information may have been incorrectly entered, if at all. This has been remedied in the CCAP database so that the original *LAB_ID* has been entered in the field *PREVIOUS_LAB_ID*, and the field number has been entered in *FIELD_ID*. *PREVIOUS_LAB_IDs* can be generated and populated by creating an Access query where the *FIELD_ID* in one table is linked to *LAB_ID* in a copy of the same table, thereby exposing the original *LAB_ID* and the resubmittal *LAB_ID*.

The field *PREVIOUS_JOB_ID* contains the original batch number (*JOB_ID*) of a resubmitted sample that has been given a new batch number upon resubmittal for further analysis. *PREVIOUS_JOB_ID* is entered into the NGDB on the SSF but is not a required field for data entry. It can be generated and populated by creating an Access query where the *LAB_ID* in one table is linked to *PREVIOUS_LAB_ID* in a copy of the same table, thereby exposing the resubmittal *JOB_ID* and the original *JOB_ID*, which is then entered in *PREVIOUS_JOB_ID*.

Sample Submittal

The field *SUBMITTER* contains the name(s) of the scientist(s) who submitted the sample in a batch to the laboratory for analysis. The sample submitter is not necessarily the sample collector. The standard format for *SUBMITTER* has been taken from the USGS Library Catalog format and is as follows: "Flintstone, Fred J."; or "Flintstone, Fred J., and Rubble, Barney H." if there are multiple submitters. In the past, the NGDB has had multiple spellings or formats for one name. This has been corrected so that one person has only one name in the NGDB. The name of the chemistry project chief or the chemist performing the analyses is commonly used as the *SUBMITTER*, though they may have had nothing to do with the sample collection or subsequent geochemical data interpretation. Because of this, the names of secondary sample submitters noted on RFAs have been added to *SUBMITTER* to better reflect the project aims and personnel and to allow for a more efficient linking of *SUBMITTER* with "name" fields in other databases, such as the Library Catalog at *http://library.usgs.gov/*.

The field *PROJECT_NAME* contains the project name of the work group that allocated funds for the collection and analysis of submitted samples. *PROJECT_NAME* entries can be used to group samples by the project that concerns them, as well as to indicate the scientific intent behind the samples and their data. In this manner, mineral-resource assessment data can be kept separate from environmental assessment data. Where data are absent, *PROJECT_NAME* can be derived from its match with the Project Number (project account number) that is present on the RFA. RFAs commonly have more than one Project Number as the project account that is funding the project might change during the duration of the project. *PROJECT_NAME* was not a required field in early versions of the SSF. *PROJECT_ACCOUNT* has always been required in the SSF but cannot be easily used to populate *PROJECT_NAME*.

The field *DATE_SUBMITTED* contains the date that the sample was submitted to Sample Control for initial database processing before laboratory sample prep and analysis. *DATE_SUBMITTED* is also useful as it places the samples in time with respect to the analytical methods available when the samples were submitted. Resubmitted samples may not always have a new *DATE_SUBMITTED* entry, however, as further analysis could be requested for samples already in

RASS or PLUTO without requiring the assignment of a new *LAB_ID* for the sample in question. This can result in situations where a *DATE_SUBMITTED* entry is matched with chemical data that were derived from an analytical method that was not available until 10 years later. The NGDB does not include "date of analysis" information, but it could be added retroactively by gleaning it from RFAs and laboratory reports.

Sample Location

The field *STATE* contains the name of the State from which a sample was collected. *STATE* entries do not always match the geographic coordinate entries given for the same sample, and these samples need to have their locations checked. RASS had no State field, but *STATE* in PLUTO entered a 2-character code that commonly contained errors (for example, "AL" was sometimes erroneously entered for Alaska), as well as 2-character *Country* codes entered in the wrong field (for example, the "CO" country code erroneously entered in the *State* field, resulting in Colombia records erroneously identified as derived from Colorado). An accurate *STATE* field can be GIS-derived once geographic coordinates for samples are found to be accurate.

The field *QUAD* contains the name of 1:250,000-scale quadrangle (1°×2° or 1°×3°) in which the sample was collected. For some USGS geochemical reconnaissance programs, the 1:250,000 quadrangle name, also known as a 2-degree (2°) quadrangle, is of much importance. In Alaska, geotechnical data of all kinds are routinely sorted and retrieved by quadrangle; less so in the conterminous States. As needed, accurate quadrangle information can be GIS-derived for samples with accurate coordinates.

The field *SPHEROID* contains the name of the reference spheroid or ellipsoid, when recorded, for the latitude and longitude coordinates of the sample site, and the field *DATUM* contains the reference datum for the site. Spheroid and datum information was not entered in RASS or PLUTO, so most NGDB sample records have no location reference information other than coordinates. In the numerous instances where datum and spheroid information is missing for samples listing latitude and longitude values, these older (commonly pre-2000) coordinates were most commonly obtained by digitizing or calculating locations based on USGS topographic maps. Since these topographic maps are most commonly projected using the NAD27 datum and Clarke 1866 spheroid, these are the recommended datum and spheroid to use in these instances.

The field *LATITUDE* contains the latitude coordinate of the sample site, and *LONGITUDE* contains the longitude coordinate of the sample site—both reported in decimal degrees and usually with NAD27 datum and Clarke 1866 spheroid for samples submitted prior to year 2000 (see preceding paragraph). Through the SSF, the submitter enters positive (Northern Hemisphere) and negative (Southern Hemisphere) decimal numbers for latitude and positive (Eastern Hemisphere) and negative (Western Hemisphere) decimal numbers for longitude. Both RASS and PLUTO entered coordinates in the format degrees-minutes-seconds, in varying degrees of precision. There is a logical sequence for obtaining missing geographic coordinates, and a discussion of these processes is detailed below.

The first step in obtaining accurate geographic coordinates for records that do not have them is to check RFAs against inaccurate data entry. Second, missing coordinates may be found in USGS Open-File Reports (OFRs) in cases where coordinates were published but were never furnished for database entry. Third, there are projects that have sample locations where several different types of sample media have been collected at one site. If one type of media from a site has good coordinates, the same coordinates apply to all other media type samples from this site. Also, these projects may have submitted sample splits to separate laboratories for different analyses. Sorting samples by *FIELD_ID* (see "Sample Identification" section concerning *FIELD_ID*) will group samples by field sites, which may yield accurate coordinates for *LAB_IDs* that are lacking.

A different sort of coordinate problem are "corner coordinates," defined as coordinates that plot on the corner—usually the southeast corner—of a 7.5′ or 15′ sample-site location topographic base map. In these cases, the actual sample sites are hand-plotted accurately on the field map, but the coordinates of the map corner were submitted to Sample Control for data entry as the location of the samples. This problem is most prevalent in the PLUTO database because the RFA coding sheets actually requested coordinates for the "SE corner of the 7 ½′ quadrangle." Samples entered in the PLUTO database may also have been entered in the RASS database and have accurate coordinates there, so sorting samples by *FIELD_ID* will retrieve these matching pairs. Although many corner coordinates have been corrected and updated, there remain many samples in the CCAP dataset that have corner coordinates that need to be resolved.

Finally, a search for archived or published sample-site location base maps is executed in an effort to provide coordinates for samples that are lacking them. Many of these maps reside in the geological data archives in the Field Records section of Denver's USGS Library and in the USGS Alaska Technical Data Unit. These maps are then geographically registered, the sample-site locations are identified and digitized, the resulting coordinates are converted to geographic Cartesian coordinates, and the data are entered in *LATITUDE* and *LONGITUDE*, with appropriate datum and spheroid information added.

Scientists who submitted samples to be entered in the RASS database had the option of submitting coordinates in Universal Transverse Mercator (UTM) projection instead of the latitude/longitude of geographic Cartesian projection. UTM Northing values, UTM Easting values, and the UTM Grid Zone number were required for data entry. UTM coordinates in the NGDB have been converted to geographic coordinates by the use of coordinate translation programs. Inadvertent truncation of the 7-character UTM Northing field to six characters during initial data entry created glaring location errors, but this has since been corrected as well. In this CCAP database, all UTM coordinates have been converted to latitude/longitude.

Scientists who submitted samples to be entered in the RASS database also had the option of submitting coordinates in X–Y projections instead of the latitude/longitude of geographic Cartesian projection. The Y component corresponds to Latitude or Northing, and the X component corresponds to Longitude or Easting. Some X–Y coordinate projections are actually accepted projections such as a State plane projection, but this fact may not be indicated on RFAs or in publications. The values may be coordinates of points located on a grid that is superimposed on the sample-site location base map, or they might be the coordinates of points that are measurements on the map from a point of origin (commonly the southeast corner of the map). X–Y coordinates always require the known location of the origin as well as the scale of distance used to create the grid (1″=1,000′, that is) on the sample-site location base map. Without a scale or an origin, the sample-site locations would need to be digitized. In this CCAP database, all X–Y coordinates had been converted to latitude/longitude before initiation of the CCAP.

By international convention, the latitude values of sample-site locations that are in the Northern Hemisphere are positive numbers, and negative values place sample-site locations in the Southern Hemisphere. *LATDIR* entries in RASS for samples in this dataset are “N,” placing them in the Northern Hemisphere, and latitude entries in PLUTO for CCAP samples are positive numbers. By similar convention, the longitude values of sample-site locations that are in the Western Hemisphere are negative numbers, and positive values place sample-site locations in the Eastern Hemisphere, based on the Greenwich prime meridian line. *LONGDIR* entries in RASS for samples in this dataset are “W,” placing them in the Western Hemisphere, and longitude entries in PLUTO for CCAP samples are negative numbers. Many samples that were collected in the Western or Northern hemispheres had site locations in the NGDB that plotted in the Eastern or Southern hemispheres because longitude or latitude was not entered as a negative number. These errors have been corrected in the NGDB. All samples in the CCAP database have positive LATITUDE (Northern Hemisphere) and negative LONGITUDE (Western Hemisphere) coordinates.

The field *DEPTH* contains information regarding the depth of the sample from the land surface if cored or trenched, accompanied by units of measure. *DEPTH* may refer to one measurement or a measured interval from which the sample was collected. “Surface” is a valid entry for *DEPTH*. Sample depth can be obtained from RFA Sample Comments, from the PLUTO field *DESCRIPT1*, or can be derived from *FIELD_ID* if it is known that the field number contains information that refers to sample depth (for example, DH1A10-20′ could imply that the sample comes from the interval between 10 and 20 feet in drill hole 1A). The SSF allows the submitter to enter depth value or range, with units, but earlier versions permitted the entry of “Surface (0–25 cm),” “1′–2′,” or any other entry desired by the submitter. The depth field *Z_COORD* in RASS was very sparsely populated, did not have units of measure, and could refer to measured distance above land surface or sea level (elevation). *DEPTH* is not to be confused with elevation above or below sea level.

The field *LOCATE_DESC* contains geographic information relating to the location of the sample site. Location description information was commonly recorded in the RASS Sample Comment field that was never digitally entered, or in the *DESCRIPT1* field of PLUTO. This field should not contain State or 1:250,000 quadrangle information as there are already fields for that information. Data for *LOCATE_DESC* can be derived from GIS layers. It can also be derived from fields in the NURE database when NURE samples are being resubmitted for further analysis.

Sample Collection

The field *DATE_COLLECT* contains the date, and time if applicable, when the sample was collected. Collection date information is a relatively new entity in the NGDB and is especially important for samples collected for environmental geoscientific study. It is also important in the case of NURE samples resubmitted for further analysis because reanalysis can take place 30 years after the samples were collected. For these resubmittals, *DATE_SUBMITTED* would be entered as the current date, but *DATE_COLLECT* usually is not entered at all, giving the false impression that the samples were recently collected and creating confusion when NGDB reanalyses are compared with the original NURE data. *DATE_COLLECT* can be derived from the *SAMPDAT* field in the NURE database by creating an Access query that links *Prime_ID* in NURE to *FIELD_ID* in the NGDB. This date will contrast with the date that the sample was submitted for reanalysis, which is found in the *DATE_SUBMITTED* field (“Sample Submittal” section). In the CCAP database *DATE_COLLECT* is in the format mm/dd/yyyy.

The field *SAMPLE_SOURCE* contains information regarding the source of sample or sample site type as indicated by the sample submitter. Examples of entries for *SAMPLE_SOURCE* are outcrop, stream, mine dump, and prospect pit. It is very important to know of cases when samples have been collected at sites that are directly related to mining, such as underground mines, mine dumps, prospect pits, and mill tailings piles, for these samples tend to be mineralized and may have high concentrations of metals. *SAMPLE_SOURCE* may also indicate the process or mode of transport that caused the sample to reach its source. Initially, many stream sediments have had a *SAMPLE_SOURCE* of “outcrop” but are actually stream deposits and have been changed to reflect this in the CCAP database. Much data that once resided in the PLUTO field *DEPOSITN* were migrated to the NGDB’s *SAMPLE_SOURCE* in order to describe the environment of deposition at the sample site. Samples from drilled holes (core, cuttings, sludge, and well fluids) should be checked for sample depth information that could be used to populate the *DEPTH* field. For soil and plant samples, data that once resided in the PLUTO field *BIOTIC* and *BIOTICSET* were migrated to *SAMPLE_SOURCE* to describe the biotic setting at the sample site. *SAMPLE_SOURCE* is not concerned with the sample preparation processes of segregation, concentration, or extraction that were used to create mineral separates, heavy mineral concentrates, or leachates but is concerned with the source of the parent material from which these sample types were derived. The field *METHOD_COLLECTED* refers to the field sample collection method—usually single grab, composite, or channel—not to further laboratory preparation. The field *SMPLCHAR* in PLUTO was used to describe whether the sample was typical or atypical of a larger body or population at the sample site, as well as a description of the collection method. *METHODCOL* in RASS was solely concerned with collection method. Thus, sample typicality is not noted in *METHOD_COLLECTED* of the NGDB.

Sample Description

The field *PRIMARY_CLASS* contains the primary classification of media, and its categories in the CCAP database are rock, sediment, soil, and concentrate. The field *SECONDARY_CLASS* is the secondary classification or subclass of sample media. It is a descriptor used to further define the basic nature or type of material collected as a sample and is an attribute of *PRIMARY_CLASS*. The SSF generates *SECONDARY_CLASS* entries for the media rock, organic, and sediment, and possibly for leachate and water. The field *SPECIFIC_NAME* further defines the basic nature or type of material collected as a sample and is an attribute used to modify *PRIMARY_CLASS* and(or) *SECONDARY_CLASS*. The SSF generates *SPECIFIC_NAME* entries for the media mineral, rock, organic, sediment, and miscellaneous, and possibly for leachate and water. A discussion of *PRIMARY_CLASS*, *SECONDARY_CLASS*, and *SPECIFIC_NAME* categories, their attributes, and their interrelationships is provided below.

Rock is defined as an aggregate of one or more minerals, for example, granite, shale, marble; or a body of undifferentiated mineral matter, for example, obsidian; or of solid organic material, for example, coal (Neuendorf and others, 2005). This definition includes ore samples, jasperoid, gossan samples, fault breccia (mylonite or gouge), tektites, and nodules or concretions. In the NGDB, coal is considered an "organic" sample rather than a "rock" sample. This point would certainly be disputed by many geologists but is moot for the CCAP database as coal samples and their data have been excluded. The degree of lithification may separate rocks from sediments and is determined by the sample collector or submitter. Saprolite samples are considered rocks. Samples that are derived from rock by processes of mineral separation, particle concentration, or chemical extraction are called minerals, concentrates, or leachates respectively. The *SECONDARY_CLASS* attributes of the class "rock" are "igneous," "metamorphic," "sedimentary," and "unspecified." There is commonly a fine line between the first three terms. For example, a water-laid tuff may be considered a type of igneous or sedimentary rock. For the CCAP database, the contents of all attribute fields for the rock sample concerned are considered when checking *SECONDARY_CLASS* entries, and consistency is maintained. For instance, where "igneous," "metamorphic," or "sedimentary" rock type could not be discerned the rock is classed as "unspecified."

The entries for *SPECIFIC_NAME* under "igneous" rocks are primarily petrologic terms based on classification according to Streckeisen (1967, 1976). Data from earlier RASS and PLUTO classifications, submitter's comments on RFAs, and other information at hand are considered when entering data in *SPECIFIC_NAME*. If nothing is known of the petrology or chemistry of the sample, its intrusive or extrusive nature is noted. If nothing is known of the form of the igneous body from which the sample was collected, *SPECIFIC_NAME* is considered "unspecified." The SSF provides a significantly abbreviated list of igneous rock names from which to choose. The field *IGNEOUS_FORM* contains information regarding the form or structure of the igneous body from which the sample was collected, whether the rock is considered igneous or metaigneous. The migration of RASS and PLUTO data to the NGDB has caused the data in *IGNEOUS_FORM* to become less definitive than it was originally. Many igneous structural categories in *ROCKFORM* of PLUTO were not represented in *IGNEOUS_FORM* of RASS, and others were merged to create one more general category. For the CCAP database, the data have been returned to their original entry. This field has been further populated using information derived from *SAMPLE_COMMENT* that pertains to the igneous structural type found at the site of the collected sample. As *IGNEOUS_FORM* is an attribute in the NGDB for igneous rocks only, it is not possible to define the structures of metaigneous rocks, such as metamorphosed dikes of metadiabase, through the use of this field. This is not the case in the CCAP database. For igneous rock samples, *ADDL_ATTR* contains code-derived information regarding the igneous texture and structure of the sample (from *TEXTURE* in PLUTO), the sample's quartz-feldspar relationship (from *QUAR_FELD* in PLUTO), as well as other attributes that came from *MODIFIER-IGNEOUS* in RASS. None of this information was migrated from PLUTO or RASS to the NGDB and must be entered in *ADDL_ATTR*. These data also can be used to further define *SPECIFIC_NAME*.

The entries for *SPECIFIC_NAME* under "metamorphic" rocks are derived from rock names from RASS, petrologic composition code data from PLUTO, metamorphic facies code information, and from submitter's comments on RFAs. If nothing is known regarding the petrology or chemistry or source of the rock from which the sample was collected, *SPECIFIC_NAME* is considered "unspecified." The SSF provides a significantly abbreviated list of metamorphic rock names from which to choose. The field *METAMORPHISM* contains code-derived information regarding the type of metamorphism present at the sample site, with the two entries being “regional” and “contact.” Though the data in *METAMORPHISM* was migrated from PLUTO, it is possible to further populate this field by using the “greisen” and “contact metamorphic” entries found in the field *MINERALDP* of RASS, as well as using information derived from *SAMPLE_COMMENT* that pertains to the type of metamorphism present at the site of the collected sample. The field *FACIES_GRADE* contains code-derived information regarding the metamorphic facies and grade classification of the rock from which the sample was collected. Though the data in *FACIES_GRADE* were migrated from RASS and PLUTO, it is possible to further populate this field by using information derived from *SAMPLE_COMMENT* that pertains to the type of metamorphic facies or grade present at the site of the collected sample. The field *SOURCE_ROCK* describes the type of rock that was metamorphosed to create the metamorphic rock that was sampled. This field only contains the entries “igneous” and “sedimentary” and is very sparsely populated in the NGDB. It could be further populated by using relevant information found in the field *SAMPLE_COMMENT*. It is not certain whether the data in the PLUTO field *METATYPE* were migrated to the NGDB, but if so, the data are most likely found in *ADDL_ATTR*. It is possible to further populate this field using information derived from *SAMPLE_COMMENT* as well as from the metaigneous and metasedimentary entries that were once in *METATYPE*. *SOURCE_ROCK* data are currently migrated to the NGDB by way of the SSF program. For metamorphic rocks, the field *ADDL_ATTR* contains code-derived information regarding the source of the rock metamorphosed (from *METATYPE* in PLUTO), the degree of metasomatism (from *METASOMAT* in PLUTO), the metamorphic texture (from *MTEXTURE* in PLUTO), the mineralogy and chemical composition (from *COMPOSIT* in PLUTO), as well as other attributes that came from *MODIFIER-METAMORPHIC* in RASS. Some of this information was not migrated from PLUTO or RASS to the NGDB and must be entered into *ADDL_ATTR*. The information also can be used to further define *SPECIFIC_NAME*.

The entries for *SPECIFIC_NAME* under “sedimentary” rocks are derived from earlier RASS and PLUTO classifications, submitter’s comments on RFAs, and other information at hand. The SSF provides a significantly abbreviated list of sedimentary rock names from which to choose. The field *DEPOSIT_ENVIRON* contains the categories “marine,” “continental,” and “transitional” that define the environment of deposition for the rock from which the sample was collected. As this is much generalized information and was entered only in the very sparsely populated *SDEPOSITN* of PLUTO, this field is of little use in the NGDB. For sedimentary rocks, *ADDL_ATTR* contains code-derived information regarding the chemistry, mineralogy, and texture of the sedimentary rock from which the sample was obtained. These data were usually migrated from RASS (from *MODIFIER-SEDIMENTARY*) and PLUTO (from *MODIFIER1* and *MODIFIER2*) to the NGDB and also can be used to help further define the entries in *SPECIFIC_NAME*.

Sediment is defined as solid fragmental material that originates from weathering of rocks and is transported or deposited by air, water, or ice, or that accumulates by other natural agents, such as chemical precipitation from solution or secretion by organisms. Sediment forms in layers on the Earth’s surface at ordinary temperatures in a loose, unconsolidated form; for example, sand, gravel, silt, mud, till, loess, alluvium (Neuendorf and others, 2005). There can be a fine line of definition between sediment and soil, especially if the soil in question has no true horizon development or the sediment is colluvially derived. Samples that are derived from sediment by processes of mineral separation, particle concentration, or chemical extraction are called minerals, concentrates, or leachates, respectively. In the NGDB, *SECONDARY_CLASS* attributes of “sediment” are currently “panned” and “panned only,” which refer to concentrated samples. Likewise, *SPECIFIC_NAME* entries for sediments in the NGDB all refer to concentrate samples. “Concentrate” is not yet a *PRIMARY_CLASS* category in the NGDB; but in the CCAP database, concentrate samples that have been derived from sediment or soil have been classified as concentrate. Concentrate terms are further described below. There are three possible places in the SSF that might generate *SECONDARY_CLASS* attributes for sediment: presubmittal sample concentration (preparation provided by submitter), postsubmittal sample concentration (preparation provided by analytical laboratory), and *SPECIFICSAMPLETYPE*, which seems the most likely of the three. For sediment samples, *ADDL_ATTR* contains code-derived information regarding the chemistry, mineralogy, texture, and grain size of the sediment that was collected. Some of these data were migrated from RASS (from *DESCRIPT1*) and PLUTO (from *MODIFIER1* and *MODIFIER2*) to the NGDB. Entries from *DEPOSITN* in PLUTO may provide data for *ADDL_ATTR* as well as for *SAMPLE_SOURCE*.

Soil is a term used in soil science for the unconsolidated mineral or organic matter on the surface of the Earth that serves as a natural medium for the growth of land plants and has been subjected to and shows effect of genetic and environmental factors of climate (including water and temperature effects) and macro- and microorganisms, conditioned by relief, acting on parent material over a period of time. A product-soil differs from the material from which it is derived in many physical, chemical, biological, and morphological properties and characteristics. Engineering geologists refer to soil as all unconsolidated material above bedrock (Neuendorf and others, 2005). These are obviously two very different definitions for soil, showing that there can be a vague boundary between sediment and soil, especially if the soil in question has no true horizon development or the sediment is colluvially derived. Samples that are derived from soil by processes of mineral separation, particle concentration, or chemical extraction are called minerals, concentrates, or leachates, respectively. There are two possible places in the SSF that might generate *SECONDARY_CLASS* attributes for soil: presubmittal sample concentration (preparation provided by submitter) and postsubmittal sample concentration (preparation provided by analytical laboratory). Another field in the NGDB that describes soil samples is *SAMPLE_ZONE*, which describes the soil zone or horizon from which the sample was collected. Samples from mixed zones cannot be coded accurately in *SAMPLE_ZONE* of the NGDB. The field *HORIZON* describes whether these horizons are well marked or poorly defined. The field *SALINE* describes whether a soil is saline or nonsaline in character. The field *FERRITIC* describes whether a soil is ferritic or nonferritic. The field *ORGANICS* describes whether a soil is organic rich or organically poor to nonorganic in nature. The field *DRAINAGE* describes whether a soil is well drained or poorly drained. These fields do not define the relative degree of these characteristics. It is possible to further populate these fields by using information derived from *SAMPLE_COMMENT*. For soil samples, *ADDL_ATTR* contains code-derived information regarding the chemistry, mineralogy, texture, and grain size of the sediment that was collected. These data were migrated from RASS (from *DESCRIPT1*) to the NGDB.

In the NGDB, concentrate is not yet a category of *PRIMARY_CLASS*, but in the CCAP database it has been given a class of its own. Concentrate, as a sample submitted for analysis, is the fraction of solid material remaining after one or a series of physical or mechanical separation or segregation processes have been completed. The sources of concentrated material, which may be rock (after crushing), sediment, or soil, are noted. The concentration procedure used is noted as well. These procedures can involve separation by specific gravity (density), magnetism, and particle size. In the NGDB, separation by particle size (grain size) alone does not constitute a concentrate, but when accompanied by one or more other separation processes, it is classified as such. A sample of mono-mineral separates is entered as “mineral” rather than “concentrate” for samples representing a single mineral.

Currently in the NGDB, only samples with a *PRIMARY_CLASS* entry of “sediment” can be coded as concentrates, but in the CCAP database, all sediment and soil concentrate samples have been coded as concentrates. Concentrates from rocks are not present in the CCAP database. The *SECONDARY_CLASS* attributes of sediment are “panned” and “panned only,” which refer to concentrated samples in the NGDB and in this database. Another similar mineral concentration can be achieved using a Wilfley table, but this method was not encountered in the CCAP dataset. If samples are “panned only,” *SPECIFIC_NAME* is a bulk-panned concentrate or heavy mineral concentrate (called “HMC”). If they are called “panned” in *SECONDARY_CLASS*, it is understood that other separation processes were used as well. Panned sample fractions in *SPECIFIC_NAME* are C1 (highly magnetic fraction of heavy mineral concentrate), C2 (less magnetic fraction of heavy mineral concentrate), C3 (nonmagnetic fraction of heavy mineral concentrate), C4 (combination of fractions C2+C3), C5 (combination of fractions C1+C2), and “lights” (nonheavy mineral concentrate with magnetite removed).

The separation parameters for these concentrates define the nature of the concentrate type entered. These data are entered in the *PREP* field of the NGDB. A typical entry might be “–35 hmc, bromo @ 2.85 SP, modified Frantz @ 0.25 & 1.75 amp ~ 0.2 & 0.6 amp w 15° fwd slope & 10° side tilt,” which means that the sample analyzed was a panned heavy mineral concentrate, sieved to –35 mesh, with a specific gravity greater than 2.85 after undergoing a density separation by bromoform, split into three magnetic fractions using a “modified Frantz Isodynamic Separator” method with electromagnetic settings of 0.25 and 1.75 amp, which is equivalent to Frantz settings of 0.2 and 0.6 amp with a 15° forward slope and a 10° side tilt. The fractions yielded by this process of separation are $\leq$0.25 amp fraction (C1), $\geq$0.25 to $\leq$1.75 amp fraction (C2), and $\geq$1.75 amp fraction (C3). The analyzed sample is most commonly one of these fractions, with the C3 fraction being the most common. Fractions were sometimes combined or their parameters of separation modified upon direction of the sample submitter.

In the NGDB, the category “miscellaneous” represents sample types that are more difficult to classify in the above categories. Some “miscellaneous” samples are precipitates, coatings, synthetic solutions or products, fossils, mill tailings, meteorites, dust, and so forth. Strong arguments can be made to place some of these terms with the sample types listed above, and the NGDB design group is currently investigating and discussing this issue. The *SPECIFIC_NAME* options for miscellaneous samples in the NGDB by way of the SSF log-in process are “peat,” “gossan,” “limonite,” “caliche,” and “other,” whereby one word is entered to describe this other type of sample. As previously mentioned, peat is best classed as a soil type. Gossans are iron-bearing weathered products overlying sulfide deposits and are usually rocks, though they can also be classed as sediment (Neuendorf and others, 2005). In the CCAP database, gossan, limonite, and caliche are classed as rock, sediment, or soil, depending on other information that would be used to determine the *PRIMARY_CLASS* choice that is accurate.

The field *SAMPLE_COMMENT* does not exist in the NGDB but has been created for the CCAP database. This field contains attributes used to modify *PRIMARY_CLASS*, *SECONDARY_CLASS*, or *SPECIFIC_NAME* and can be used to further describe the sample site, criteria, or any information about the sample that is useful. The contents of *SAMPLE_COMMENT* are not code-derived data but represent comments that have been written or typed on RFAs by sample submitters, data collected by the database renovator that are derived from field notebooks and OFRs, or further explanation of the sample that has been added by the database renovator. There is potential redundancy between the contents of *SAMPLE_COMMENT* and the code-derived data in other database fields, especially in *ADDL_ATTR*. This redundancy is removed from the CCAP database.

Petrologic and mineralogic information about rock samples may be found in *SAMPLE_COMMENT*. These data can be used to better define the contents of the *SECONDARY_CLASS* and *SPECIFIC_NAME* fields and can assist in the population of other fields in the NGDB. Detailed alteration and mineralization information also is found in *SAMPLE_COMMENT*, and these data may more clearly define the contents of the *ALTERATION* and *MINERALIZATION* fields. *SAMPLE_COMMENT* may contain structure and fabric information about rock samples, as well as comments regarding the relationship of the sample to larger features at the sample site. In the case of igneous rock samples, these data may assist in the population of *IGNEOUS_FORM*. Information regarding the composition of sediment and soil samples also may be found in *SAMPLE_COMMENT*. *SAMPLE_COMMENT* can contain the name of another agency that collected the sample or a note that the sample was originally collected under the NURE program. *SAMPLE_COMMENT* may also contain an alternate field number or a non-USGS laboratory identification number that is attached to the sample. In the event that the sample was previously submitted for analysis and received an earlier *LAB_ID* within an earlier *JOB_ID*, this lab number and job number have been transferred to the fields *PREVIOUS_LAB_ID* and *PREVIOUS_JOB_ID*, respectively. *SAMPLE_COMMENT* may contain detailed sample-collection and preparation information. Some of these data should be transferred to the *PREP*, *STABILIZATION*, and *MESH_PORE_SIZE* fields if relevant.

The field *ADDL_ATTR* contains additional attributes used to modify *PRIMARY_CLASS*, *SECONDARY_CLASS*, or *SPECIFIC_NAME* and in the CCAP database are solely derived from sample codes in fields of RASS and PLUTO that do not have current equivalent fields in the NGDB. In the NGDB, *ADDL_ATTR* also contains all data that would be found in *SAMPLE_COMMENT*, which may put a strain on the 255-character limit of this field. The only code-derived data from RASS that were migrated to *ADDL_ATTR* come from the RASS field *DESCRIPT1*, but only in the cases where *PRIMARY_CLASS* entries are “sediment” (“unconsolidated sediment” in the field *SMPLTYPE* in RASS) but *DESCRIPT1* is not a concentrate (“C,” “C1,” “C2,” or “C3”), or in cases where *PRIMARY_CLASS* entries are “soil.” Data entered in the following RASS fields were not migrated to *ADDL_ATTR*—or to any field—in the NGDB, but have been added to the *ADDL_ATTR* field of the CCAP database: *STRUCTURL* refers to the structural setting of the sample (“fracture/joint,” “shear or fault”); *MATRIX* refers to the matrix material of the sample—usually a rock (“silica,” “Fe/Mn,” “carbonate,” “clay”); *OXIDATION* refers to the degree of oxidation of a sample—usually a rock (“oxidized,” “partially oxidized,” “unoxidized”); *ORE/MINAL* refers to the ore mineral group in the sample (“base metals,” “precious metals,” “mixed base and precious metals,” “radioactive,” “rare earths”); and *MODIFIER* fields contain data that further define *SPECIFIC_NAME* for rock samples. Code-derived data from PLUTO that were migrated to *ADDL_ATTR* come from fields *MODIFIER1* and *MODIFIER2* if the entry in field *CATEGORY* is “sedimentary rock” or “unconsolidated sediment”—primary and secondary modifier data of sedimentary rock or sediment samples. Other code-derived data from PLUTO that were migrated to *ADDL_ATTR* come from *MATERIAL*, if *CATEGORY* is “unconsolidated sediment,” which refers to the primary material that makes up the sediment sample. Data entered in the following PLUTO fields were not migrated to the field *ADDL_ATTR*—or to any field—in the NGDB but have been added to the *ADDL_ATTR* field of the CCAP database: *TEXTURE* refers to the texture and structure of an igneous rock sample; *QUAR_FELD* refers to the relationship between quartz and feldspar of an igneous rock sample; *METASOMAT* refers to the degree of metasomatism exhibited in a metamorphic rock sample (“clearly,” “may be,” “no evidence”); *MTEXTURE* refers to the texture exhibited in a metamorphic rock sample (“schistose,” “gneissic,” “massive,” “grain size”); and *COMPOSIT* refers to the mineral composition of a metamorphic rock sample.

ADDL_ATTR in the NGDB is used as the destination for sample comments from the SSF, as well as for the migrated data from the sample comment field *DESCRIPT1* (PLUTO), and from the RASS and PLUTO fields previously listed. Sample site information (for example, adit names, prospect pit numbers, names of and distances from geographic features) that is stored in *ADDL_ATTR* in the NGDB has been transferred to the *LOCATE_DESC* field in the CCAP database. As the USGS created the NGDB from the migration of RASS and PLUTO data, some RASS and PLUTO fields were determined to be poorly populated or to contain information that might be considered irrelevant, confusing, or redundant. The CCAP database puts all non-code-derived data in *SAMPLE_COMMENT* and all code-derived data from fields with no RASS or PLUTO equivalent field, whether migrated to the NGDB or not, in *ADDL_ATTR*. This would make *ADDL_ATTR* a field containing RASS and PLUTO data only. In creating this database, the data from some RASS and PLUTO fields excluded from migration have been recovered and reentered into *ADDL_ATTR*, and the data from other excluded fields have been omitted. The data from a few included fields have been deleted from *ADDL_ATTR* as it has been found to be relatively useless. The contents from the RASS fields *STRUCTURL*, *MATRIX*, *OXIDATION*, *ORE/MINAL*, and *MODIFIER* were not migrated to the *ADDL_ATTR* in the NGDB but have been added in this database. Though these RASS fields have no equivalent PLUTO fields, the data within them have been deemed critical to the interpretation of information regarding the samples, especially rock samples. It is not certain whether any data from the RASS field *DESCRIPT1* were migrated to *ADDL_ATTR* as it appears that all were migrated to *SPECIFIC_NAME*. In the case of rock samples, this method works; but in the case of sediment, soil, and organic samples, problems are created. The only *DESCRIPT1* entries for sediment or soil samples that belong in *SPECIFIC_NAME* are “C1,” “C2,” “C3,” and “concentrate.” All others have been transferred from *SPECIFIC_NAME*, or added, to *ADDL_ATTR* except for “stream sediment,” which is redundant if “stream” is the *SAMPLE_SOURCE* entry. The contents from the PLUTO fields *TEXTURE*, *QUAR_FELD*, *METASOMAT*, *MTEXTURE*, and *COMPOSIT* were not migrated to the *ADDL_ATTR* in the NGDB but have been added in this database. Though these PLUTO fields have no equivalent RASS fields, the data within them have been deemed critical to the interpretation of information regarding the rock samples with which they were submitted.

A number of data-entry points in the SSF create potential data problems for migration into the NGDB. The fields "Oxidation" (of rock samples), "Modifier" (of sedimentary rock samples), and "Plant Part" are obvious ones that contain code-derived data. There are other SSF fields whose data may or may not be migrated to ADDL_ATTR in the NGDB: "Sample Comment"; "NURE Resubmitted Sample"; "RASS/PLUTO Resubmitted Sample"; "Reference Sample"; submitter analyzed parameters (pH, alkalinity, conductivity, dissolved oxygen, and temperature); specific sample type data; and numerous fields containing sample preparation information.

Four other fields in the CCAP database describe geologic media. *GEOLOGIC_AGE* contains information regarding the geologic age of the sample or the age of the material at the sample collection site. *GEOLOGIC_AGE* contains the names of the geologic time units of Era, Period, and in the case of Tertiary and Quaternary Periods, Epoch. It does not refer to geochronologic age. Data migrated from PLUTO may contain ranges of geologic age (for example, Cambrian–Devonian). This field has had many misspellings and variations of possible ranges of geologic age that need to be standardized. In the case of some sediment or soil samples, *GEOLOGIC_AGE* in the NGDB may refer to the age of the bedrock underlying the sample source or the estimated geologic age of rock outcrops and float samples in the vicinity rather than the age of the sediment or soil samples themselves. The entry of geologic age data in the SSF does not allow for the entry of ranges of geologic time. For samples that have accurate coordinates, this field could be populated using a GIS layer from small-scale USGS geologic maps, as has been done on CCAP geologic mapping tasks. The field *STRATIGRAPHY* contains stratigraphic unit data, or comments regarding the sample or the sample site. *STRATIGRAPHY* usually contains formal, accepted names of stratigraphic units, igneous bodies, metamorphic zones, and so forth. In these cases, names have been standardized and spelled correctly in the CCAP database. In other cases, stratigraphic comments and details have been migrated to the NGDB, representing information that the submitter felt was important to the interpretation of the geochemical significance of the sample.

In the SSF, the submitter is allowed 255 characters to describe stratigraphy.

The field *MINERALIZATION* contains information regarding the mineralization type evident in the sample or at the site where the sample was collected. The scope of *MINERALIZATION* is limited in the NGDB due to lack of population and, in the case of data migrated from PLUTO, lack of concise definition. This field has been further populated using information derived from *SAMPLE_COMMENT* that pertains to the mineralization of the collected sample. If the sample is known to be mineralized but the actual type of mineralization is not noted, the entry "mineralized" has been entered in this field. In the SSF, mineralization is an attribute of rock samples only, but in reality, samples coded as "soil," "sediment," and "miscellaneous" may exhibit types of mineralization as well. The field *ALTERATION* contains data regarding the alteration type evident in the sample or at the sample site. The scope of the field *ALTERATION* is limited in the NGDB due to lack of population and, in the case of data migrated from PLUTO, lack of concise definition. This field has been further populated using information derived from *SAMPLE_COMMENT* that pertains to the alteration of the collected sample. If the sample is known to be altered but the actual type of alteration is not noted, the entry "altered" has been entered in this field. In the SSF, alteration is an attribute of rock samples only, but in reality, samples coded as "soil," "sediment," and "miscellaneous" may exhibit types of alteration as well.

Sample Preparation

The field *PREP* contains sample preparation description or comments. Sample preparation information may be available on the cover sheet of the RFA. It also may be found if linking the sample to a *PROJECT_NAME*, and the project and(or) submitter to an OFR that contains fairly detailed explanation of the sample preparation procedures that were implemented for the various media types collected for that project. Some of this information has been used to populate the field *MESH_PORE_SIZE,* but all the rest populates *PREP*, especially when describing the preparation procedure used for concentrate samples. Also, sample preparation information residing in *ADDL_ATTR* or *SAMPLE_COMMENT* has been moved to *PREP*. *PREP* is poorly populated because there were no preparation fields in RASS or PLUTO, but most sample preparation procedures used by the USGS have changed very little over the years. Generally, USGS samples of earth material were prepared as follows: rock—crushed, pulverized to pass a 100-mesh screen, and mixed; sediment—dried, most organic material removed, sieved to pass an 80-mesh screen; the –80 mesh fraction is ground to pass a 100-mesh screen, and mixed; soil—dried, juiced, sieved to pass an 80-mesh screen, the –80 mesh fraction is ground to pass a 100-mesh screen, and mixed (however, the analysis of the –10 mesh fraction of soil samples is also quite common); heavy mineral concentrate—may vary from submitter to submitter, preparation information taken from RFA and OFR has been entered in *PREP*. In the SSF entry process, there are a number of fields that are concerned with the submitter's preparation procedures and the laboratory's preparation procedures. Some samples are submitted for analysis fully or partially prepared, but the majority of others will require some sort of laboratory preparation prior to analysis. These fields appear to produce data that will be migrated to *ADDL_ATTR* or *PREP* in the NGDB: "Ash/Ashed," "Blend/Blended," "Concentrate/Concentrated," "Grind/Ground," "Sieve/Sieved," "Rock/Sediment/Soil Request for Prep," "Request for Prep," "Completed Treatment (by Submitter)," "Other Completed Treatment (by

Submitter).” The field *MESH_PORE_SIZE* contains the mesh size through which a sediment or soil sample was sieved, or pore size through which a water sample was filtered, prior to analysis. Sieve mesh size is of great importance to sediment and soil samples as sieving promotes homogeneity. It is best to know and enter the actual parameters of a sieved fraction (for example, “–10 +80 mesh” rather than “+80” mesh). Sieve size of solid material is expressed using US standard sieve mesh number rather than in millimeters or in micrometers. Filter pore size is expressed in micrometers.

Chemical Analysis

The unique key field of the NGDB’s **CHEM** table is *UNIQ_SEQ_NUM*, software-assigned integers that are created as chemical analyses that are migrated into the NGDB from LIMS or from outside sources. *UNIQ_SEQ_NUM* was not used as the key field of **CHEM_DATA** in the CCAP database because the table contains many analyses that are not currently in the NGDB, such as the original NURE analyses found in the CCAP database. These data do not belong in the NGDB, but all other analyses will be migrated to the NGDB. All analytic records in **CHEM** are accompanied by associated LAB_IDs. At times, samples have been resubmitted for further analysis and have been assigned new LAB_IDs. These samples in the **CHEM** table cannot easily be linked to data from the original samples without linking to the **GEO** table on *LAB_ID*, or on *PREVIOUS_LAB_ID* if it has been fully populated for these resubmittals. Many entries for *JOB_ID* in **CHEM** are unpopulated, but those data also can be obtained by linking **CHEM** to **GEO** using *LAB_ID*. *JOB_ID* in **CHEM** has not been populated in this manner because these chemical analysis records may actually belong to a resubmittal job number that is not in *JOB_ID* of **GEO**.

The measured characteristic *SPECIES* is a chemical attribute of element, ion, or oxide concentration that has an associated data value. Species fields in RASS and PLUTO were very similar, and the NGDB has used PLUTO species names wherever chemical data were present in that database. Very few corrections were required in *SPECIES*, but some species names have been changed in the CCAP database to better support sorting. For example, “Carbonate C” and “Organic C” in the NGDB are “C-CO3” and “C-org” in the CCAP database. For carbon and sulfur, “Total C” and “Total S” in the NGDB are “C” and “S” in the CCAP database because total concentration is assumed for all species unless the sample digestion is known to be partial. In another example, “CV Th” in the NGDB—the coefficient of variance for thorium—is “Th-CV” in the CCAP database. The entries of trace elements expressed as oxides have been changed so that *SPECIES* names are elements; the *DATA_VALUE* entries were converted from oxide concentrations to element concentrations, which helps to reduce the number of possible fields in Access or columns in Excel that are attributes of given species. In the NGDB, samples that have been ashed prior to analysis have been assigned species such as “ash Mg.” The CCAP database has changed this attribution so that *SPECIES* would be “Mg,” and “after ashing” has been entered in *DECOMPOSITION*. Many entries have been shortened to assist the creation of *PARAMETER*, which is the concatenation of *SPECIES*, *UNITS*, *TECHNIQUE*, and *DIGESTION* (for example, “Loss on Ignition” changed to “LOI”). The fields *NID* (numerical ID) and *DESCRT* (description of the numerical ID) in PLUTO and their RASS equivalents were used to define *SPECIES* and *UNITS* in the NGDB and were migrated to the NGDB as the fields *NID* and *NID_DESC*, and to *STATUS* in part, but were not included in the CCAP database.

The field *UNITS* contains the units of concentration or measurement in which the *DATA_VALUE* is expressed in both the NGDB and the CCAP database. The units for trace elements expressed as weight percent have been changed so that *UNITS* are parts per million and the *DATA_VALUE* entries multiplied by 10,000 for these elements. The units for most major elements and major-element oxides expressed as parts per million have been changed so that *UNITS* are weight percent; the *DATA_VALUE* entries were divided by 10,000. In the same manner, units of parts per billion have been converted to parts per million and the *DATA_VALUE* entries divided by 1,000. These changes significantly help to reduce the number of possible fields in Access or columns in Excel that are attributes of given species. The entry “percent” in the NGDB has been shortened to “pct” in the CCAP database to assist in the creation of *PARAMETER*.

As a result of researching analytic methods used for chemical analysis by the USGS since its inception, the fields *TECHNIQUE*, *DIGESTION*, and *DECOMPOSITION* have undergone significant change. A search for publications that describe these analytic methods has yielded the 753 titles listed in **AnalyticMethodBiblio** of the CCAP database, and the 70 methods that were used to detect the data values in **Chem_Data** are listed in **AnalyticMethod** that details the scope, digestion, decomposition, and detection used. Senior USGS chemists and emeriti were approached with many questions over the past 9 years and their answers recorded and archived; this information became a critical part of analytic method description. The association of certain analysts with certain methods, the comparison of dates-of-sample submittal with the general date for the advent of new analytic methods, and the grouping of certain methods with specific laboratories all assisted in accurately defining the analytic methods of detection. Most of the newer data required little editing, but the older data benefited much from the research of analytic methods. Results of the process are, at the least, a good educated guess. This research was used to create the CCAP database fields *ANALYTIC_METHOD* and *PARAMETER*. The PLUTO fields *TOA* (code for type of analysis), *DESCCT* (description of the analytic method), and their RASS equivalents were helpful in defining some analytic methods. These data were migrated to *TOA* in the NGDB but were not included in the CCAP database.

The field *TECHNIQUE* is the abbreviation of the analytic method used to analyze samples. Nineteen *TECHNIQUE* entries are found in **Chem_Data** of the CCAP database to accurately reflect the analytic method used. The following CCAP entries have been added to the NGDB: “CB” (combustion), “CD” (conductance), “FL” (fluorometry), “GRC” (gamma ray counting), “GV” (gravimetry), “TB” (turbidimetry), and “VOL” (volumetric analysis). Many NGDB entries have been shortened to assist in the creation of *PARAMETER* (for example, “DC-ARC SPEC” changed to “ES” for emission spectrography). The PLUTO fields *LLOA* (combination of the 2-character abbreviation of the laboratory location with the 2-character mnemonic of the analytic method) and *LLOADS* (the description of the *LLOA*) help to define *TECHNIQUE* and *LAB_NAME*. The data in *LLOA* were migrated to the NGDB fields *CENTER* and *PROJECT* but are not included in the CCAP database.

The field *DIGESTION* is the abbreviation of the degree of sample digestion—total (T) or partial (P)—required by the *TECHNIQUE* used to analyze the sample for a specific species. Some methods of sample digestion that are called “total” are less “total” than others but still are considered “total” in the NGDB and in this database.

In the NGDB, the field *DECOMPOSITION* contains a brief description of the decomposition method used for a given *TECHNIQUE* in the analysis of the sample and has been expanded in the CCAP database also to include comments that further describe this *TECHNIQUE*. There are 81 different types of *DECOMPOSITION* in this database. In many cases, multiple entries that infer the same basic information have been reduced to a general but accurate description that works for all. In two examples, *DECOMPOSITION* contains information that makes the distinction between quantitative emission spectrography and the semiquantitative method, and between instrumental neutron activation analysis and delayed neutron counting. *DECOMPOSITION* contains key information used in the creation of the entries in *ANALYTIC_METHOD* of **Chem_Data** in the CCAP database.

DATA_VALUE and *QUALIFIER* are the fields in the NGDB and the CCAP database that contain the numeric results and, when present, their qualifying modifiers. This data-storage scheme also was used in RASS and PLUTO, so chemical data migration to the NGDB went smoothly, considering the amount of data at hand. Very few values or qualifiers needed to be changed except for the many cases that “0” (zero) had been entered in *DATA_VALUE* of the NGDB, accompanied by the “less than” qualifiers “N” or “L” (“Structure” section). These values usually represent the lower limits of detection (LLD) for emission spectrography data from PLUTO. Data coming from RASS were not a problem because the LLD was always entered in that database. One of the results of the research of analytic methods used by the USGS is a table of LLDs for nearly all species, specific to the method of analysis used, and related to the time that the method was employed. This study was used to correct all “zero” values in the CCAP database, as well as hundreds of thousands of values in the NGDB. *DATA_VALUE* was also changed in the CCAP database if *SPECIES* or *UNITS* was changed in a way that required a conversion of the value. The key drawback of working with *DATA_VALUE* and *QUALIFIER* together is having two required fields to create one value, which doubles the number of result fields needed if Access tables are created or Excel flat files are exported that contain these data. Also, crosstab queries only can work with one value field at a time, not two. This limitation is overcome by the creation of *QUALIFIED_VALUE* in the CCAP database, which presents chemical values in two different formats (“Structure” section).

The field *LAB_NAME* was created for the NGDB; though there was some useful laboratory information stored in the PLUTO field *LLOA* (for example, “LAXR” shows the laboratory location “LA” for Lakewood, Colo., and the laboratory mnemonics “XR” for the X-ray spectroscopy laboratory). *LAB_NAME* is more accurately defined in the CCAP database to clearly link the chemical data to the analytical laboratory.

Publishing support provided by:
Denver Publishing Service Center, Denver, Colorado
Manuscript approved for publication, December 14, 2009

For more information concerning this publication, contact:
Center Director, USGS Central Mineral and Environmental Resources
Science Center
Box 25046, Mail Stop 973
Denver, CO 80225
(303) 236-1562

Or visit the Central Mineral and Environmental Resources Science Center
Team Web site at:
http://minerals.cr.usgs.gov/

This publication is available online at:
http://pubs.usgs.gov/ds/509/

Granitto and others—**Central Colorado Assessment Project (CCAP)—Geochemical Data for Rock, Sediment, Soil, and Concentrate Sample Media**—USGS/Data Series 509