FAQs for rollout of *Geochemical and mineralogical data for soils of the conterminous United States*:

1. Why this, why now, why the USGS?

Why this? The rationale for a national-scale soil geochemical survey was eloquently articulated by Dr. Mary Lou Zoback in her Presidential address to the Geological Society of America in 2001:

> *"Documenting and understanding natural variability is a vexing topic in almost every environmental problem: How do we recognize and understand changes in natural systems if we don't understand the range of baseline levels?"*

In specific regard to soils, how can we speak knowledgably about soil pollution from human activities without a thorough understanding of the natural variability of the potentially toxic elements of interest? Our current understanding of the national-scale natural variability of chemical elements in soil is based on a set of 1,300 soil samples collected in the 1960s and 1970s by the USGS and analyzed for 50 elements by methods available at that time through the USGS analytical laboratories. Thirty of these 50 elements, including many of the environmentally important elements (e.g., Be, Cd, Cr, Cu, Ni, Pb, and V), were analyzed by semiquantitative techniques that would no longer be considered acceptable for any environmental study. Although this small data set has been remarkably useful over the past 30 years, it was definitely time to bring our soil geochemistry knowledge into the 21st century with establishing a new and improved data set for the nation with a higher sample density and vastly improved analytical methods.

Why now? Understanding the abundance and spatial distribution of chemical elements in the Earth's near-surface environment is critical for fields such as risk-based assessment of contaminated land, agriculture, animal and human health, water quality, land-use planning, mineral exploration, industrial pollution, and environmental regulation. The need for improved soil geochemical data has been noted by all these stakeholder groups. Some states (e.g., Pennsylvania, California, Ohio, Michigan, Washington, New Hampshire, New Jersey, Mississippi, Illinois, and Florida) have recognized this need and have conducted state-scale soil geochemical surveys. However, these state-scale surveys were not conducted to standardized sample collection and analytical protocols and, as a result, consist of data generated on soils collected from a variety of depths and horizons and analyzed by different analytical methodologies. Such data sets cannot be easily merged to identify and understand regional- and national-scale geochemical patterns. Given the wide-spread interest and need for a better national-scale soil geochemical database, it was obviously an appropriate time to initiate such an effort.

Why the USGS? The USGS has a long history of conducting broad-scale geochemical surveys for mineral resource assessments and environmental investigations using soil as the primary sample medium and has research staff recognized internationally as experts in this field. The USGS also has a mandate to conduct geological, geochemical, and geophysical investigations at a national scale. As such, we are probably the Federal agency best suited to

carry out such a national-scale project within a reasonable timeframe and interpret the observed geochemical patterns in terms of the processes which cause them.

2. When did this project begin?

In 2001, the Directors of the U.S. Geological Survey, the Geological Survey of Canada (GSC), and the Servicio Geológico Mexicano (Mexican Geological Survey, or SGM) jointly recognized the need to establish a soil geochemical database for North America for the reasons cited in #1 above. Based on the Directors' decision, a trinational project was established in 2002 to address this issue. An organizational meeting was held in October 2002 and attended by representatives of all three North American geological surveys and also by representatives of the Natural Resources Conservation Service (NRCS), Agriculture and Agri-Food Canada, and Health Canada. At that meeting, it was decided the first step should be to obtain input from stakeholders through a Soil Geochemical Workshop held in Denver, Colorado in March 2003.

3. What sort of stakeholder input was obtained prior to the start of the project?

In 2003, a Soil Geochemistry Workshop was held in Denver, Colorado with 112 attendees representing more than 40 stakeholder entities (see Table 1 below). At this workshop, the attendees were asked to provide input on (1) how a national-scale soil geochemical survey should be designed, (2) what the sampling protocols should be, and (3) what analytical methods should be used. This input has guided the project throughout the past 10 years.

Table 1

List of agencies/organizations represented at 2003 Soil Geochemistry Workshop:

Agency for Toxic Substances and Disease Registry
Agriculture and Agri-Food Canada
Bureau of Land Management
Bureau of Reclamation
Camp, Dresser, & McKee
Centers for Disease Control and Prevention, National Center for Environmental Health
Colorado School of Mines
Colorado State University, Pueblo
Earth Tech, Inc.
ENSR International
EnviroStat, Inc.
Florida Department of Environmental Protection
Florida Department of Health
Geological Survey of Canada
Geological Survey of Norway
Gradient Corporation
Health Canada
Lawrence Berkeley National Laboratory
Maryland Department of the Environment
Massachusetts Department of Environmental Protection

Minnesota Department of Transportation
National Jewish Medical and Research Center
National Park Service
Natural Resources Conservation Service
Nebraska Conservation and Survey Division
North Dakota Geological Survey
Servicio Geológico Mexicano (Mexican Geological Survey)
U.S. Army Corps of Engineers
U.S. Dept. of Agriculture; Plant, Soil & Nutrition Laboratory
U.S. Environmental Protection Agency
U.S. Forest Service
U.S. Geological Survey
University of California, Davis
University of Colorado Health Sciences Center
University of Essex (UK)
University of Florida
University of Maryland, Baltimore County
University of Michigan
University of Montreal
University of Nebraska
University of Wisconsin, Madison
Woodard and Curran
Xavier University of Louisiana

---

4. How were the soil sampling sites selected for the national-scale survey?

The target sampling sites were selected by a Generalized Random Tessellation Stratified design at a nominal density of 1 site per 1,600 square kilometers (1 site per 617 square miles). This sample design was recommended by a panel of experts from the USGS, GSC, NRCS, U.S. Environmental Protection Agency, Centers for Disease Control and Prevention, Minnesota Geological Survey, Savannah River Ecology Laboratory, Agriculture and Agri-Food Canada, and Environment Canada at a workshop held in 2006. The GRTS design ensures that the target sites are selected in a random manner while also ensuring that there is a relatively uniform distribution of sites throughout the nation. The sampling density was chosen to be consistent with recommendations from the International Union of Geological Sciences Task Group on Global Geochemical Mapping. The field crews were given a great deal of flexibility to move the sample site if the original target site was not accessible for any reason. They were only required to keep the alternate site as close as possible to the target site and to try to sample from the same general landscape setting as found at the target site.

5. Does the geochemical data generated by this project represent "background" concentrations of elements in soil?

Given atmospheric deposition of elements such as mercury and lead from industrial processes and historic automobile emissions, it can be argued that it is impossible to collect a soil sample anywhere in the conterminous U.S. whose geochemistry has not been affected by human

activities. However, our sampling protocols called for avoiding areas of obvious contamination and the samples are as close as possible in the early 21st century to representing natural background concentrations of the elements. It is most precise to say these samples represent a geochemical baseline for the period 2007-2010 against which future changes from natural processes or human activities may be measured.

6. What types of soil samples were collected at each site?

The sampling protocols called for collection of three samples at each site: (1) a sample from a depth of 0 to 5 centimeters (0 to 2 inches); (2) a composite sample of the soil A horizon (the uppermost mineral soil); and (3) a sample from the soil C horizon (generally the partly weathered soil parent material), or if the top of the C horizon was at a depth greater than 1 meter (3.3 feet), a sample over about a 20-centimeter (8-inch) interval at a depth of about 1 meter (3.3 feet).

7. Who are the primary customers for the resulting data sets?

The primary customers are those agencies or organizations needing information on "background" concentrations of elements in soil. These customers are primarily (1) environmental regulators, (2) risk assessors dealing with risk-based assessments of contaminated land, and (3) public health specialists dealing with soil pathways for human exposure to potentially toxic elements. Because soil is the source of most biologically active trace elements that reach humans through the food chain, specialists in agriculture and food safety will also find the data useful. The current issue of arsenic in rice has illustrated the need to know the elemental content of soils on which food crops are grown.

8. Why was "total or near total" elemental concentration determined as opposed to "total recoverable" concentration as done by some environmental studies?

This was another recommendation to come out of the 2003 workshop. The workshop attendees felt that the methodology for determining the "total or near-total" elemental content by a four-acid acid digestion was a more robust method than the weaker (partial) extractions (e.g., aqua regia) that are used to determine the "total recoverable" elemental content. These "partial" chemical extractions were felt to be more dependent on procedural detail and, likely, operator technique. Thus, these partial extraction data may not stand the test of time.

9. How many people were involved in collecting the samples?

From 2007 – 2010, approximately 40 people participated in collecting soil samples throughout the conterminous U.S. These included 9 USGS scientists, an unknown number of staff from the NRCS in North Dakota and South Dakota, an unknown number of staff from the state geological surveys of Minnesota, Nebraska, and Pennsylvania, and 19 students representing 12 universities as follows:

University of Missouri
University of Nebraska

North Carolina State University
Utah State University
University of Wisconsin, Platteville
University of Wisconsin, River Falls
University of Minnesota
Virginia Tech
North Dakota State University
Delaware Valley College
University of Arizona
University of Southern Maine

10. How much did it cost to collect soil samples at more than 4,800 sites throughout the conterminous United States?

Just the sample collection activities (salaries of the sample collectors, travel expenses, vehicle cost, supplies, etc.) cost approximately $1 million.

11. The USGS Mineral Resources Program funded a project called the "National Geochemical Survey" from about 1997 to 2012. This project analyzed over 70,000 samples from throughout the United States. How does the National Geochemical Survey differ from the Soil Geochemical Landscapes of the Conterminous United States Project, which conducted the soil sampling whose data are currently being published?

The National Geochemical Survey (NGS) was initiated in 1997 with the stated purpose of providing complete national-scale geochemical coverage of the U.S. using the sample archives of the National Uranium Resource Evaluation–Hydrogeochemical and Stream Sediment Reconnaissance (NURE-HSSR) Program as a starting point. The NURE Program, conducted by the U.S. Department of Energy from 1975 to 1984, had the purpose of delineating areas of the U.S. with potential for undiscovered uranium resources. The HSSR component of this program was initiated with the intent of determining the geochemical composition of sediments and soils from approximately 650,000 sites throughout the U.S. to aid in the search for undiscovered uranium resources. However, sampling ceased in 1980 with only about 65% of the U.S. covered. The sample medium most commonly collected by NURE-HSSR was stream sediments, followed by soils in the Great Basin states and a significant number of lake sediments in Alaska. The goal of the NGS was to obtain national coverage, and thus complete the effort begun by NURE-HSSR. The NGS sample design was based on a 17 by 17-kilometer (10.6 x 10.6 mile) sampling grid resulting in a planned sample density of 1 site per 289 square kilometers (112 square miles). To obtain national coverage, the NGS used data from the following sources: (1) NURE-HSSR samples that had already been reanalyzed by other USGS projects using appropriate methods; (2) A subset of the existing NURE-HSSR stream sediments, soils, and lake sediments that were reanalyzed by the NGS analytical methods; (3) Data from other USGS projects involving regional-scale geochemical sampling for stream sediments or soils, if the collection and analytical methods were compatible with those of the NGS; (4) When possible, samples taken from the archives of stream sediments collected by other USGS projects that were reanalyzed by

the NGS to fill in any unsampled areas; and (5) New sample collection and chemical analyses that were conducted by the NGS in areas where no coverage was available from either the NURE-HSSR Program or previous USGS projects. Data from all these sources representing at least three broad categories of sample media (stream sediments, soils, and lake sediments) were then combined into a single national database.

Stream sediment was the primary sample medium for NURE-HSSR and NGS and was collected wherever possible. However, good stream-sediment samples are not available in all locations. In the NURE-HSSR Program, stream sediments were collected at about 80% of the sites, soils at about 12% of the sites, and lake or pond sediments at about 8% of the sites. In those parts of the U.S. where the NGS collected new samples, soils were substituted for stream sediments only where necessary. This included areas of low relief and poor drainage and farm lands where local streams had been largely channelized and diverted for agricultural purposes.

Upon completion of the NGS, soils were only collected for about 30% of the country. Therefore, the NGS data set is not useful to those agencies and organizations needing national-scale soil geochemical data. These include the customers listed in FAQ7, above. The Soil Geochemical Landscapes of the Conterminous United States Project had the goal of establishing national-scale soil geochemical data set that would meet the needs of these customers as well as the research needs of the USGS.

12. What are the primary factors that control the abundance and spatial distribution of chemical elements in soils of the conterminous United States?

The primary controlling factor is the chemical composition of the underlying soil parent material. This parent material may be bedrock or unconsolidated glacial or eolian material. Climate is also an important factor in controlling the concentration of some elements (e.g., calcium). Human influences, including widespread atmospheric deposition illustrated by the distribution of mercury and lead, and agricultural practices, such as application of phosphate fertilizers and shown by the distribution of phosphorus and selenium, can also be detected in the data.

13. Will the data generated by this project for the conterminous United States be comparable with other national-scale soil geochemical surveys being conducted in Europe, Australia, and China?

Yes. We exchanged standard reference materials with all of these other projects to ensure comparability for a large number of elements.

14. The project was initiated in 2001 as a collaborative effort among the USGS, the Geological Survey of Canada and the Mexican Geological Survey to conduct a soil geochemical survey of North America. How are the other two countries proceeding?

Mexico completed its sampling in 2012 and is currently conducting chemical and mineralogical analyses. These are anticipated to be completed in late 2013 or early 2014. Unfortunately, Canada dropped out of the project in 2010 after completing sampling in New Brunswick, Nova Scotia, Prince Edward Island, and small portions of other provinces. With a new government

administration and new leadership of the Geological Survey of Canada, science priorities shifted and the soil geochemical survey was not one of their high priorities.

15. Was the Natural Resources Conservation Service (NRCS) consulted during the design of this project?

Yes. NRCS was a participant from day one of the planning phase. They co-hosted the 2003 Soil Geochemistry Workshop and had representatives in attendance at every follow-up workshop. They decided that they could not be a full partner in the project because of other priorities. However, NRCS leadership encouraged state offices to participate as time and resources allowed. NRCS personnel in North Dakota and South Dakota collected all the samples for this project in their states and this assistance was greatly appreciated. USGS has continually kept NRCS advised as to progress of the project.

16. How does this national-scale data set benefit my state or congressional district?

It must be emphasized that the goals of this project are to delineate ***national-scale*** geochemical patterns and establish ***national-scale*** variations in elemental composition for soils. This was not a project designed to answer definitively the question: "What is in my backyard?" The sample design established sample sites based on 1 site per 1,600 square kilometers (617 square miles). At this density, there are 4,857 sites in the conterminous U.S. However, at this density, the smallest states (Rhode Island and Delaware) only have 2 sites each; therefore, not a lot can be said about the geochemical variation in these individual states. The primary benefit for individual states, congressional districts, etc. is that the current data set establishes a national-scale context for consideration of soil geochemistry in these individual entities. The data set enables us to state that, for example, soils of Nebraska congressional district 3 (covering the western 75% of the state) have lower arsenic concentrations than soils in Nebraska congressional districts 1 and 2 (covering the eastern 25% of the state). The reason for this is the composition of the soil parent material. In the eastern part of Nebraska (Congressional districts 1 and 2), there are marine shales and sandstones of Pennsylvanian, Permian, and Cretaceous age. Most of the rest of the state (Congressional district 3) is underlain by much younger non-marine sandy and limey deposits. The older marine sediments in districts 1 and 2 are almost always elevated in trace element content compared to the younger non-marine rocks and this is reflected in the soils.

17. How is such a low-density, national-scale soil geochemical data set used by agencies conducting a risk-based assessment of, for example, a superfund site that may only cover an area of a few square kilometers?

Dr. Teresa Bowers of the Gradient Corporation outlined at our 2003 Soil Geochemistry Workshop how a national-scale soil geochemical data set can aid in the local-scale risk assessment process. She stated such a data set can:

1. Be useful as a screen for where background may exceed risk thresholds;
2. Provide a guide for when a local background study may be useful;

3. Provide an indication of when local background may be anthropogenic rather than natural; and
4. Be useful for narrowing or expanding the contaminant list for a local background study.

18. All the elements in the data set (with the exception of organic carbon) are inorganic. Did the project consider analyzing for organic compounds such as pesticides, polycyclic aromatic hydrocarbons (PAHs), etc.?

The recommendations from the 2003 Soil Geochemistry Workshop included analysis of organic compounds to address two issues: (1) long-range transport of organic compounds and (2) the geographic distribution of major pesticides and their transformation products. In the pilot phase of the project (2004–2007), we collected soil samples for analysis of organic compounds using sample collection protocols established by organic chemists. However, it quickly became evident that the cost of analyzing for organic compounds far exceeded our anticipated budget for the project. One of the decisions reluctantly reached on the basis of our pilot studies was to omit analysis of organic compounds.

19. How was collecting soil samples on private land handled?

Each field crew was responsible for contacting the landowner and obtaining permission to collect samples on private land. This was done once the crew was in the field; no attempt was made to obtain permission prior to going to the field. If a landowner could not be found, then the sampling site was moved to a location where the landowner could be contacted. No samples were collected on private land without first obtaining permission. It should be noted that the vast majority of landowners were happy to cooperate with our project and the project could not have been successfully completed without this cooperation.

20. Although not a part of the current publication, the project included determination of soil pathogens. How did this come about and will the pathogen data be published?

One of the recommendations from the 2003 Soil Geochemistry Workshop was to include determination of selected soil pathogens in the national-scale study. During the pilot phase, we determined 4 soil pathogens in the soil A horizon: (1) *Bacillus anthracis* (anthrax), (2) *Brucella abortus* (brucellosis), (3) *E. coli*, and (4) *Cryptosporidium parvum*. Based on these pilot studies, it was decided to only determine *Bacillus anthracis* in soil samples collected from a depth of 0 to 5 centimeters for the national-scale survey. In 2010, USEPA's Homeland Security Research Center in Cincinnati, Ohio asked if it would be possible for two additional pathogens to be determined in the samples collected during the final year of sampling (2010). These two pathogens are *Yersinia pestis* (plague) and *Francisella tularensis* (tularemia or rabbit fever). Through an Interagency Agreement, Dale Griffin (USGS microbiologist in St. Petersburg, FL) will be determining these new pathogens in addition to *Bacillus anthracis* and EPA provides the funding. Publication of the pathogen data will be handled by Dale and EPA colleagues.

21. What is the most important point learned from this project from the point of view of the primary stakeholders (i.e., environmental regulators, risk assessors, public health specialists)?

The most important point is the recognition that the concentration of chemical elements in soil can vary by a factor of 2-3 orders of magnitude, or in some cases more than this. The potentially toxic elements in soil analyzed by this project that vary in concentration by more than a factor of 1,000 (3 orders of magnitude) include arsenic, cobalt, chromium, copper, nickel, lead, uranium, and zinc. Those that vary between a factor of 100 and 1,000 include beryllium, cadmium, selenium, and thallium. Input from human activities (pollution) is superimposed on this variable, natural geochemical background. It is not possible to recognize and quantify human input without a thorough understanding of the natural variability. This project provides a more complete understanding of this natural variability for the conterminous U.S. than has ever been available.

22. The data also includes quantitative determinations of soil mineralogy. Why was this done and what methods were used?

Quantitative determinations of soil mineralogy were tested during a pilot phase of this project and found to be useful in interpreting the causes of geochemical variability. We decided to add such analyses to the full survey. Determinations were made using an x-ray diffraction technique that provides accurate percentages for the major minerals in soils. This data is useful in its own right by providing information on variables such as acid neutralizing capacity of soil and predicting a variety of physical and chemical responses of soils largely controlled by clay content. In addition, the mineralogic data allows predictions of the behavior of various elements of concern because we can now determine their mineralogic residence and contrast areas where elements are tightly bound in stable minerals versus areas where those elements are more mobile and bioavailable.

23. Were the approximately 14,400 soil samples collected during this study archived so they can be made available for future research?

Yes, splits of all the samples are archived at the USGS in Denver and can be made available to other investigators. Anyone interested should contact David B. Smith (dsmith@usgs.gov).

24. The sample archives, along with the geochemical and mineralogical data sets, should provide numerous opportunities for collaboration with research scientists in both government and academia. Have any collaborative efforts been established to date?

Even though the data sets are just being published, there are already research scientists using our sample archives from this project. In addition, we have an Interagency Agreement with the U.S. Environmental Protection Agency allowing them access to our unpublished data. Collaborative research projects currently underway include:

1. Johannes Lehmann, Associate Professor; Verena Sabine-Jauss, Ph.D. student; Cornell University, Department of Crop and Soil Sciences, Ithaca, NY. Dr. Lehmann and Ms.

Sabine-Jauss are using a subset of the soil samples for the determination of regional- and national-scale distribution of black carbon in soil.

2. Kang Xia, Associate Professor, Virgina Tech, Department of Crop and Soil Environmental Sciences, Blacksburg, VA. Dr. Xia is using a subset of the soil samples in her research on soil organic carbon and nitrogen speciation.

3. Li Ma, Ph.D. Student, Virgina Tech, Department of Crop and Soil Environmental Sciences, Blacksburg, VA. Ms. Ma is using a subset of the soil samples in her studies of bioavailable amino acids in soils of the United States.

4. David Brown, Associate Professor; and Ross Bricklemyer, Ph.D. student, Department of Crop and Soil Sciences, Washington State University, Pullman, WA. Dr. Brown and Mr. Bricklemyer are using a subset of the soil samples to develop a method using proximal sensing techniques to determine forms of carbon in soil.

5. Jose Almirall, Professor and Director of the International Forensic Research Institute & Trace Evidence Analysis Facility; and Sarah Jantzi, Ph.D. student, Department of Chemistry and Biochemistry; Florida International University, Miami, FL. Dr. Almirall and Ms. Jantzi are using a subset of the soil samples to investigate forensic analysis of soil evidence.

6. Mark Waldrop, USGS, Menlo Park, CA. The project collected a duplicate set of samples for Dr. Waldrop at 10% of the sites throughout the conterminous U.S. These samples are being used for soil metagenome sequencing.

7. Dale Griffin (USGS, St. Petersburg, FL; Tonya Nichols and Sarah Perkins, U.S. Environmental Protection Agency, National Homeland Security Research Center, Washington, DC (Nichols) and Cincinnati, OH (Perkins). The project has an Interagency Agreement with EPA to determine three soil pathogens in soils collected from the conterminous U.S. (see FAQ 20).

8. Randy Waite, U.S. Environmental Protection Agency, Research Triangle Park, NC; Jennifer Phelan, RTI International. Research Triangle Park, NC. Mr. Waite and Dr. Phelan are using our geochemical and mineralogical data and our land use information at each site in their pilot studies to develop improved terrestrial acidification critical loads for analysis of the impacts of air deposition of oxides of nitrogen and sulfur.

9. George Gray, Professor and Director of the Center for Risk Science and Public Health; Adam Johnston, Ph.D. student; George Washington University, Department of Environmental and Occupational Health, Washington, DC. Dr. Gray and Mr. Johnston are using our samples and land use information in New England as part of Mr. Johnston's PhD research on arsenic in agricultural soils.

10. Clemans Reimann (Geological Survey of Norway) and Patrice de Caritat (Geoscience Australia). The project is collaborating with Drs. Reimann and de Caritat to

establish improved global soil reference values based on three continental-scale geochemical surveys (Europe, Australia, U.S.).

11. Charles Partridge (USEPA, Denver, CO). Dr. Partridge and colleagues at EPA are using our lead data as the best soil lead background data set for the United States. See their web site at: http://www.epa.gov/superfund/lead/background.htm.

12. Rufus Chaney (US Dept. of Agriculture; Beltsville, MD) and Timothy Taylor (USEPA, Washington, DC). Dr. Chaney and Mr. Taylor are using our data in their study of the appropriate use of spent foundry sand in soils.

25. Do the geochemical and mineralogical data sets contain location information for each site?

Yes, the location of each site is identified by latitude and longitude coordinates.

26. Can the geochemical and mineralogical data be used to generate map representations of the data for each element and mineral component?

Yes, geochemical and mineralogical maps have already been prepared for each determined chemical element and mineral component, respectively. These maps provide a visual representation of the national-scale geochemical and mineralogical variation in soils of the conterminous U.S. The maps are published as USGS Open-File Report 2014-1082 (http://dx.doi.org/10.3133/ofr20141082).