Background Concentrations of Arsenic in Ohio Soils: Sources and Influencing Factors

NATE WANNER[1], Certified Professional Geologist, College of Earth and Mineral Sciences, Pennsylvania State University, University Park, PA, USA and Cox-Colvin & Associates Inc., Plain City, OH, USA.

ABSTRACT. Arsenic concentrations—representative of natural background conditions in Ohio soils—were compiled from the USGS National Geochemical Survey database, Ohio EPA studies, and private environmental site investigations. The resulting dataset of 2,783 samples collected from 1,116 locations was then correlated with soil data from USDA-NRCS SSURGO, glacial and bedrock geology, and analyses of other metals at these specific locations to identify factors most likely to affect arsenic concentrations in soil. Bedrock geology, particularly Devonian-aged materials and black shales, was found to significantly correlate to arsenic concentrations in soil. However, this correlation is complicated by both glaciation and post-glaciation erosion. Approximately 70% of Ohio's bedrock is covered in glacial materials, such as till and outwash. As glaciers advanced across Ohio, bedrock materials were eroded and deposited farther south. Arsenic concentrations in Ohio soils tend to be highest where Devonian bedrock materials were deposited by glaciation. Following glacial deposition of materials, arsenic can be eroded or leached from the deposits. Deposits of wind-blown loess and materials deposited by water (as opposed to ice) had lower arsenic concentrations than glacial tills and outwash plains that underwent minimal erosion following glacial deposition.

INTRODUCTION

Arsenic is an element that has been used for a variety of purposes in agriculture and industry. Arsenic in agriculture has been used as a herbicide and a pesticide, particularly in orchards (ATSDR 2007a). In industry, arsenic has been used in many applications: a wood preservative to resist decay; an additive to glass to prevent bubbles; and as an alloying metal in ammunition, solders, bearings, and batteries (ATSDR 2007b). Arsenic is also commonly found in coal ash and cinders (EPA 2016). Its use in residential lumber has been phased out due to its toxicity to humans and the hazards associated with sawing, sanding, or burning wood treated with arsenical preservatives. Its use in orchards has been phased out due to the possible ingestion of pesticide residues on produce (ATSDR 2007b).

Arsenic occurs naturally in soils and groundwater, often at concentrations higher than action levels set by regulatory agencies such as the Ohio Environmental Protection Agency's (Ohio EPA) Voluntary Action Program (VAP) residential direct contact soil standard of 12 mg kg⁻¹ (Ohio EPA 2016a). Concentration levels above the direct contact soil standard can complicate site assessments and remediation related to possible contamination. If an industry's activities resulted in elevated arsenic levels, they will likely need to perform corrective actions. However, if arsenic is naturally occurring, corrective actions are not required, and may not be practicable.

The influence of arsenic on investigations of possible contamination from human activities is particularly prevalent in Ohio. The author frequently performs and reviews evaluations of natural background concentrations of arsenic during remedial investigations of industrial properties in Ohio, and has had personal correspondence with numerous other environmental professionals performing similar evaluations. As seen in Fig. 1, background concentrations of arsenic in Ohio are higher than other areas of the continental United States. Furthermore, Ohio has a greater population density and industrial history than other areas of the United States with elevated arsenic concentrations. If relationships can be identified between surficial geomorphology and arsenic concentrations, then there may be less need to routinely perform extensive site-specific evaluations. Ideally, using existing knowledge of surficial geology and soils, scientists could demonstrate that site arsenic concentrations do not exceed expected background levels and no corrective actions are needed. Conversely, they may be able to predict that site

[1]Address correspondence to Mr. Nate Wanner, Cox-Colvin & Associates Inc., 7750 Corporate Blvd., Plain City, OH 43064. Email: NateWanner@gmail.com
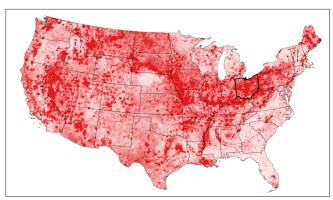
FIGURE 1. Generalized map of arsenic concentrations in soil and sediment of the continental United States created by kriging USGS National Geochemical Survey database data. Dark red areas have mean arsenic concentrations above 12 mg kg^{-1}.

concentrations exceed expected background levels and avoid the time and costs necessary to perform a site-specific study. Furthermore, site operations often cover entire properties and there is no access to unaffected soils. In these situations, it may be possible to estimate natural, pre-industry background concentrations that would otherwise be unknown.

Previous efforts have been made to understand what geomorphic processes result in the elevated arsenic concentrations in Ohio soils. The Ohio EPA has conducted county-level studies to determine background metal concentrations in soils (Ohio EPA 2015a). To date, the Ohio EPA has studied 6 of Ohio's 88 counties, and is currently studying an additional 4. While the detailed, county-level studies provide a valuable source of well-documented local data, it is impractical to perform this level of investigation for all counties, and the Ohio EPA may not continue their investigation beyond the initial 10 counties. A larger dataset is necessary to understand potential causes of elevated arsenic concentrations.

A 2014 paper by the Ohio Department of Natural Resources (ODNR) and The Ohio State University (Venteris 2014) used the United States Geological Survey (USGS) National Geochemical Survey database (NGSD) (USGS 2015b) and Gaussian simulation to compare arsenic concentrations of soil and sediment to bedrock geology. The paper concluded that bedrock, particularly Devonian shale, is likely to influence arsenic concentrations in Ohio soils. However, the relationships were not evident across the state, and areas of northwest Ohio (where Devonian bedrock was present but arsenic concentrations were not high) raised questions regarding the influence of glaciation or soil characteristics on arsenic concentrations. While Venteris (2014) sought to identify regional areas of

elevated arsenic concentrations using interpolation techniques, followed by a comparison to regional geology, the present study directly correlates individual sampling locations with soil and geologic characteristics.

Vosnakis et al. (2010) provided analysis of arsenic concentrations in soil and bedrock geology in Ohio and 6 other states south and east of Ohio. The paper concluded that background arsenic concentrations are commonly elevated above risk-based screening levels established by environmental agencies. However, consideration of geologic influences was limited to establishing that different geologic groupings showed distinct differences. Minimal discussion was given to possible causes of these differences, and there was no evaluation of glacial geology.

In the current study, a more robust evaluation of potential causes of, and correlations to, arsenic concentrations in Ohio soils was performed by combining multiple datasets in a spatially-enabled database. In this way, soil and geologic characteristics were assigned to soil samples, even when they were not recorded during sample collection. Arsenic concentrations in soil were obtained from the NGSD (a comparable dataset to Venteris (2014)), site investigations in the US EPA records (a comparable dataset to that used by Vosnakis et al. (2010)), Ohio EPA background studies, background soil data compiled from site investigations in Ohio EPA files by Cox-Colvin & Associates Inc. (Cox and Colvin 1996), and data from individual site investigations performed by private environmental consulting firms. To allow evaluation beyond the chemical and limited geomorphic data within these datasets, Quaternary geology obtained from the ODNR, bedrock geology obtained from the USGS, and soil data from the USDA-Natural Resources Conservation Service's (NRCS) Soil Survey Geographic (SSURGO) database were also loaded into the database. These geologic and soil datasets were matched to arsenic sampling locations using geographic coordinates and sampling depths.

METHODS AND MATERIALS

Three primary types of data were used in this project: geologic data, soil data, and chemical data. To effectively manage approximately 3 GB of data from multiple disparate datasets, data were compiled into a PostgreSQL (v. 9.5.4) database, with the PostGIS (v. 2.2.2) extension enabled for spatial operations. A variety of methods were used to manipulate and import the data, primarily involving Python and SQL

scripting. The study was performed using open source tools including PostgreSQL, PostGIS, QGIS, Python, R, and SAGA GIS.

Sources of Data

Soil data were obtained from the SSURGO database (Soil Survey Staff 2016). Both spatial and tabular data from 90 datasets (each representing a soil survey area) intersecting Ohio were imported into the relational PostgreSQL database using Python scripting. Importing the full set of data tables provided access to a variety of information including source materials, geomorphic settings, grain size analyses, hydrology, soil classifications, soil horizons, erosion, and other properties

Quaternary geology was obtained from the ODNR in Esri shapefile format (Ohio DNR 2015). The dataset includes information on the types of glacial geology in Ohio, such as outwash plains, glacial tills, and geologic ages of surficial glacial materials.

Bedrock geology was obtained from the US Geological Survey (USGS 2015a). The dataset included rock types and geologic ages for the uppermost bedrock units across Ohio.

Chemical data were obtained from the NGSD (USGS 2015b). The NGSD comprises over 85 different datasets, 5 of which contain data from samples collected in Ohio. The NGSD includes data used by Venteris (2014). Many of the samples in this database had analyses performed using multiple analytical methods, sometimes resulting in more than one concentration value being reported for the same analyte and sample. Different analytical methods have different levels of precision and accuracy. For instance, X-ray fluorescence (XRF) is often used to produce screening-level data without performing adjustment for moisture content, while atomic absorption spectrometry performed in an analytical laboratory typically provides more accurate data. Metadata for some datasets included in the NGSD have notations regarding the general reliability of the data. NGSD data were coded in the database so that the most reliable analysis for each analyte in each sample was used for evaluation. Furthermore, screening-level data, and data found to be otherwise unreliable, was excluded from evaluation. Although both soil and sediment data were imported from the NGSD, only soil data were used for further evaluation. Using only soil data eliminates uncertainties because the sediment may have been transported varying distances and soil chemistry may have been altered in the stream through dissolution, sorting, changes in redox potential, etc. Previous studies, including the

Venteris (2014) study of arsenic in Ohio, have not found significant correlation between stream sediments and upstream soils. An exception to using only soil data was made in generating Fig. 1, for which both soil and sediment data were used because of an absence of soil data in several states.

The Ohio EPA has undertaken county-level studies of background metal concentrations in soils. Data from these studies was obtained directly from the Ohio EPA. Where possible, original laboratory electronic data deliverables (EDDs) were obtained to facilitate database import and lessen the potential for transcription errors. Otherwise, data were converted or transcribed from report tables. The Ohio EPA has published soil background study reports for Cuyahoga County (Ohio EPA 2013a), Franklin County (Ohio EPA 2013b), Hamilton County (Ohio EPA 2015b), Lucas County (Ohio EPA 2015c), Montgomery County (Ohio EPA 2014), and Summit County (Ohio EPA 2015d). Unpublished data for Mahoning County (Ohio EPA 2016b) and Stark County (Ohio EPA 2016c) was also obtained. Additional unpublished data from private investigations ("Bratenahl project" in Cleveland, "Forest Hills Park" in East Cleveland, "Former Grimes Aerospace Facility" in Columbus, "Herron Ave. project" in Cincinnati, "MidTown Technology Center" in Cleveland, "Slavic Village" in Cleveland, and "Tiffin University" in Tiffin) was obtained from the Ohio EPA Background Study Group (Ohio EPA 2009).

Under contract to the Ohio EPA, Cox-Colvin & Associates Inc. (Cox-Colvin) previously performed an evaluation of background metal concentrations in Ohio soils (Cox and Colvin 1996). Their dataset was compiled from a variety of site-specific environmental site investigations that were selected by the Ohio EPA for inclusion in the study. Because coordinates for individual samples were not provided in this dataset, the sample locations were assigned the coordinates of the facility where the investigation was performed, as reported in the US EPA's Facility Registration System (EPA 2015). To ensure background conditions during the site-specific studies, actual sample locations were likely near or slightly outside facility boundaries, perhaps a few hundred to a few thousand feet from the facility coordinates. Data from the 1996 investigation was obtained directly from Cox-Colvin's archived database files.

Columbia Gas Transmission Corporation (CGTC) performed a number of characterization reports at compressor stations in Ohio and other states as part of a US EPA Superfund Administrative Order on Consent.

These investigations were conducted to determine if sites were contaminated, and frequently included background samples for comparison to samples collected within facilities. Although investigations largely focused on petroleum, analyses for metals were performed at the majority of sites. The CGTC data were included in a subsequent study by AECOM, which stated that the CGTC data were analyzed by US EPA-approved laboratories and that 10% of the data underwent detailed data validation (Vosnakis et al. 2010). Data from the investigations conducted in Ohio was obtained from US EPA Region 3 through a Freedom of Information Act (FOIA) request (Columbia Gas Transmission Corp. 2016). Analytical data were extracted from tables in provided report excerpts. If a sample was likely influenced by human activities, or it was uncertain whether a sample may have been so influenced, it was not included in the present study.

The resulting dataset of arsenic concentrations in soil consisted of 2,783 samples from 1,116 locations within Ohio (Fig. 2). Many of these locations were clustered so that metropolitan areas had higher sampling density than rural areas. By combining the datasets, at least a few samples were available within each Ohio county. In all datasets, laboratory analyses determined total arsenic concentrations, without speciation of arsenite (As^{3+}) versus less soluble arsenate (As^{5+}). Correlating general soil and geologic mapping units to sample locations is expected to be less precise than using lithologic observations specific to each sample. However, the variety of lithologic and geologic characteristics recorded with a chemical sample is frequently less than the variety of characteristics available from SSURGO and geologic datasets. Matching of analytical datasets to geologic and soil datasets thus greatly increases the amount of data available for evaluation. The loss of accuracy and precision in not having sample-specific data is expected to be more than offset by reductions in uncertainty when a large quantity of samples can be evaluated.

Samples included in the evaluation datasets were collected with the intent of establishing a baseline of natural background conditions in soil. Samples from the Ohio EPA and the NGSD were collected in locations thought to have no or minimal human influence. For other study datasets, where potential human sources were being investigated, the individual samples included in the evaluation dataset were collected with the intent of determining a local background concentration. Samples collected to investigate potential releases were not included in the dataset. While it is understood that essentially no area of Ohio is entirely devoid of human influence, potential influences were minimized by initial sampling location selections.
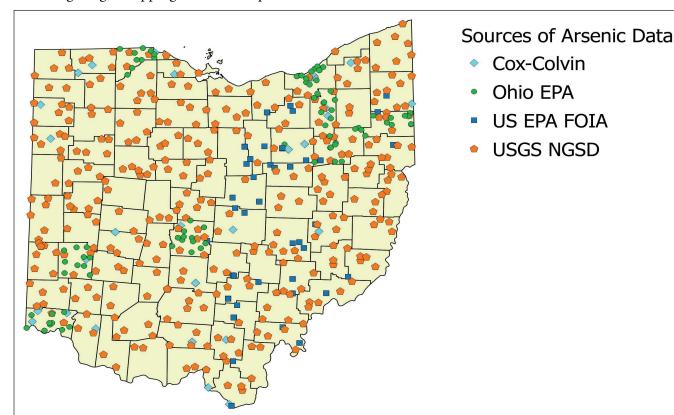

FIGURE 2. Sources and sampling locations of soil arsenic concentration data used in study. One icon may represent multiple locations.

Evaluation of Data

Following compilation of data into a single database, nearly 400 unique fields (or columns) describing soil characteristics, geology, laboratory data, and other information were available to evaluate. Some of these fields, such as arsenic concentrations, were scalar, while others were categorical. Each of the categorical fields had multiple options. For instance, geologic age could be represented as Late Wisconsinan, Holocene, Early-Middle Pleistocene, Cenozoic, or Illinoian. Many more subcategories were present in soil tables representing data on the various horizons and individual mapping units. The net result was a dataset including over 2,700 possible factors that could be evaluated for their influence on arsenic concentrations. Some of these factors and sub-datasets, such as mapping units with only one or a few samples, did not have sufficient data to provide significant insight. Python scripts were written to identify potentially significant relationships in all possible factors. These scripts compared the various factors to arsenic concentrations in soil to determine if a statistically significant correlation was present, as well as the practical significance and magnitude of correlations.

A Python script was written to retrieve data from the PostgreSQL database, manipulate it into an appropriate format, use the "rpy2" library (Gautier 2016) to perform analysis in R (R Development Core Team 2017), and then record results back in the PostgreSQL database. When evaluating data, representative soil horizon depths were compared to sample depths. A sample was considered to be from a particular horizon when greater than 50% of the sampling interval was within the representative soil horizon depth values from the SSURGO database.

Despite the careful selection of sampling locations, some samples may have been collected in areas with unrecognized human influence resulting in outlier concentrations; a potentially significant influence on parametric statistical methods that use mean concentrations. The use of non-parametric statistical methods—and comparison of median, rather than mean, concentrations—lessened the potential effect from unidentified outliers. Using non-parametric methods also precluded the need to demonstrate normal distributions, or distributions that can be transformed to a normal distribution, if parametric methods were performed on each of the evaluated factors. Because the combined datasets provided a relatively large number of samples for comparison,

the loss of statistical power in non-parametric tests is not expected to significantly affect findings.

When possible correlating factors could each be quantified, testing was performed using the non-parametric Kendall tau test. Because data were obtained from multiple studies, a variety of reporting limits were present. The Kendall tau test is easily modified for use with data that is censored at multiple reporting limits. For instance, the method can recognize that <5 mg kg^{-1} is lower than 10 mg kg^{-1}, but not necessarily lower than <10 mg kg^{-1}. In contrast, many statistical methods used in environmental analysis ignore censored data or make simple substitution of censored values to allow numerical evaluation. When datasets have a large proportion of censored data at multiple reporting limits, such substitution can cause misleading results. The "cenken" function of the R "NADA" package (Lee 2013) was used to perform the Kendall tau test, and was developed based on methods described by Helsel (2012).

When possible correlating factors were categorical in nature, survival analysis was used to compare observations from each category of a variable to observations from all other samples for which similar data were available. Tests were run individually for each group. For example, within the group of samples with descriptions of glacial geology, one test was conducted using the 298 observations in areas of alluvium as the "group" (or "factor") observations and the remaining 2,459 observations (beach ridges, end moraine, outwash, etc.) as the "nongroup" observations. The same test approach was then repeated for beach ridges, then end moraine, then outwash, etc. Running tests in this manner allowed evaluation of whether a statistically significant difference was observed for each individual factor, versus the variable as a whole (glacial geology). The "cendiff" function of the R "NADA" package was used to perform the evaluation. This function is equivalent to the Peto & Peto modification of the Gehan-Wilcoxon test, and provides a non-parametric score test which accounts for datasets with left-censored results at multiple reporting limits (Helsel 2012; Lee 2013). Means and medians for each evaluation group were also determined using the Kaplan-Meier method, as implemented in the "cenfit" function of the R "NADA" package.

Some soil samples did not have reported values for all variables. While performing evaluations for each variable, a sub-dataset was created of samples for which data on that variable was available. Comparisons were

then made within the sub-dataset. For instance, the number of samples for which water-ponding frequency information was available was less than the entire dataset. While evaluating whether areas of frequent ponding were significantly different than other areas, only those samples for which ponding values were reported were included in the evaluation. The median of these sub-datasets is typically different than the median of the entire dataset. If statistically significant differences were identified for a particular factor within a variable (for instance, sandstone within bedrock types) the *difference* of medians within the sub-dataset is presented in the following discussion rather than the calculated median concentrations. Not discussing the actual median values is intended to reduce bias when sub-datasets have different median values than that of the total dataset.

RESULTS

Results from Kendall tau and survival analysis tests were reviewed for potential correlations. In the following text, "significant" refers to correlations where probability values (p-values) for the respective tests were less than a significance level (α) of 0.05.

The median concentration of arsenic for all results in Ohio soils is 8.8 mg kg^{-1}, while the mean is 10.6 mg kg^{-1}. Ninety-nine percent of arsenic concentrations in the soil dataset do not exceed 34 mg kg^{-1}. A histogram of these concentrations is provided in Fig. 3. Because the dataset does not follow a distribution that could be normalized, formal outlier tests were not performed. To lessen the potential effect of unidentified outliers, *median* arsenic concentrations are considered in the following discussion rather than mean concentrations.
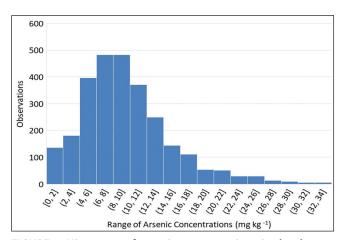

FIGURE 3. Histogram of arsenic concentrations in the dataset that do not exceed 34 mg kg^{-1} (99% of the dataset)

Bedrock Type

Comparison was made between arsenic concentrations and bedrock materials directly beneath the sample location, without consideration of depth to bedrock or possible material transport via glaciation or other mechanisms. Significant correlations were observed between arsenic concentrations and the bedrock types of black shale, sandstone, dolostone (dolomite), and limestone. Correlations to shale (other than black shale), siltstone, and mudstone were not significant. Samples collected above black shale had a median arsenic concentration 3.9 mg kg^{-1} higher than samples not collected above black shale. Soil samples collected over sandstone had a median arsenic concentration 2.4 mg kg^{-1} higher than samples not collected above sandstone. Median arsenic concentrations of soils above dolostone (dolomite) bedrock were 0.8 mg kg^{-1} less than samples not collected above dolostone, and median arsenic concentrations of soils above limestone were 3.0 mg kg^{-1} lower than samples not collected above limestone.

Age of Bedrock

Samples collected above Devonian-aged bedrock had a median arsenic concentration 3.8 mg kg^{-1} higher than samples not collected over Devonian-aged bedrock, and samples collected above Mississippian bedrock had a median concentration 2.6 mg kg^{-1} higher than samples not collected over Mississippian bedrock. Samples collected over Ordovician bedrock had median arsenic concentrations 4.0 mg kg^{-1} lower than samples not collected above Ordovician bedrock. Samples above Silurian bedrock had median arsenic concentrations 0.54 mg kg^{-1} lower than samples not collected over Silurian bedrock. Samples collected above Pennsylvanian bedrock were not significantly different from samples collected elsewhere. A box plot showing both bedrock type and age is provided in Fig. 4.

Age of Glaciation

Approximately two thirds of Ohio was covered by glacial ice masses. Glacial deposits are either Late Wisconsinan (13,000 to 24,000 years ago) or Illinoian (130,000 to 191,000 years ago) in age. Tests, performed to evaluate whether the age of glaciation correlated with arsenic concentrations in soil, indicated a significant difference. Median arsenic concentrations in areas of Late Wisconsinan glaciation were 2.2 mg kg^{-1} higher than samples not collected in areas of Late Wisconsinan glaciation, whereas samples collected in areas of
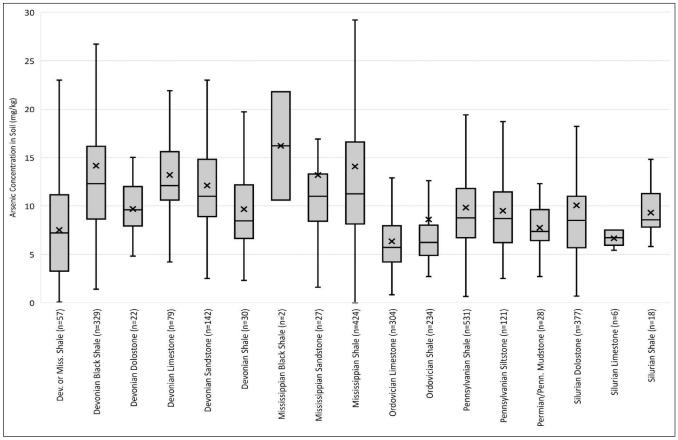
FIGURE 4. Box plot showing variation in soil arsenic concentrations overlying different bedrock. The number of samples is indicated by n=#. Results beyond the limits of the whiskers (greater than or less than 1.5 times the interquartile range) are not shown. Median concentrations are shown with a horizontal line, and mean concentrations are indicated with an "X".

Illinoian glaciation (which is older than Wisconsinan glaciation) had median arsenic concentrations 2.6 mg kg^{-1} lower than samples not collected in areas of Illinoian glaciation. Areas designated as Cenozoic in the ODNR Quaternary geology tables represent non-glaciated portions of Ohio (although the bedrock beneath these areas is older than the Cenozoic Era),
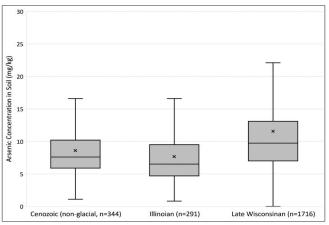

FIGURE 5. Box plot showing variation in soil arsenic concentrations within different ages of glacial materials. The number of samples is indicated by n=#. Results beyond the limits of the whiskers (greater than or less than 1.5 times the interquartile range) are not shown. Median concentrations are shown with a horizontal line, and mean concentrations are indicated with an "X".

and have a median arsenic concentration 1.5 mg kg^{-1} lower than samples not collected in areas designated as Cenozoic. A box-plot of these results is provided in Fig. 5.

Drift Thickness

A USGS study performed in Licking County, Ohio, (Thomas 2016) hypothesized that lower oxygen levels in aquifers underlying glacial drift create reducing conditions that allow for mobilization of arsenic, and thus higher groundwater concentrations. In the present study, arsenic concentrations in shallow soils were plotted against drift thickness to determine whether a similar correlation may exist with soil, but no visually discernible correlation was observed.

Post-Deposition Erosion

Following glacial deposition or erosion of parent bedrock material, soils may undergo additional erosion via stream transport, deposition in lakes, or eolian (wind) processes. Correlation between the degree of post-deposition soil erosion and arsenic concentrations was evident in multiple variables. It is important to note that soil erosion, in this case, refers to that observed

Table 1
Difference between median arsenic concentrations in samples from glacial materials and all other samples[a]

Difference in median arsenic concentration[b]	Glacial geology
2.6	Hummocky moraines
1.7	Ground moraines
0.4	Kames
-1.0	Alluvium
-1.5	Unglaciated areas
-1.6	Lacustrine deposits
-2.1	Lacustrine clays
-4.7	Valley train outwash

[a] Only statistically significant results calculated from a minimum of 50 observations are shown.
[b] mg kg^{-1}

in the geologic descriptions—not to the more recent conditions such as ditches or gullies present at the time of sample collection.

Correlations to the description field of the ODNR Quaternary geology landforms are shown in Table 1. Soil samples collected from glacial till (moraines) had higher median arsenic concentrations than samples not collected above glacial till. These materials would have undergone little erosion following glacial deposition. Kames are sand and gravel deposits that formed in depressions on glacial ice, these deposits then remained in place after the ice melted away. Kame materials would have undergone more erosion than glacial till, but less erosion than the stream and lake deposits that have lower arsenic concentrations.

Similar correlations were seen in the component parent material group ("pmgroupname") of the SSURGO database component parent material ("copm") table, as shown in Table 2. Residuum, weathered rock that has undergone minimal erosion after weathering, contains higher concentrations of arsenic than samples not collected over residuum; concentrations depending on the parent material rock type. Till materials that are closely associated with residuum, where rock materials were primarily eroded and transported by ice rather than water, also had elevated arsenic concentrations compared to samples not collected above till associated with residuum . More heavily eroded deposits, such as windblown loess and

sandy materials, have lower arsenic concentrations than samples not collected above heavily eroded deposits.

Similarly, in the parent material kind ("pmkind") field of the same table, till materials had median arsenic concentrations 1.4 mg kg^{-1} higher than samples not collected above till, while loess had median arsenic concentrations 1.6 mg kg^{-1} lower than samples not collected above loess material.

Other Metals in Soil

Kendall tau testing was used to compare arsenic concentrations to concentrations of other metals in soil. The Kendall tau algorithm is a non-parametric test used to evaluate whether higher arsenic concentrations consistently correlated to a higher (concordant) or lower (discordant) concentration of other analytes. Results for each of the analytes being compared (for instance arsenic and iron) were paired within each soil sample, and then compared to pairs of results from every other soil sample containing both analytes.

The most significant correlation was between arsenic and iron, shown in Fig. 6. This is consistent with other research where arsenic was found to be associated with iron (Welch et al. 2000; Thomas 2016). Correlations were also observed between arsenic and copper and arsenic and zinc, although these correlations were considerably lower. Arsenic in literature is commonly associated with sulfides, but the datasets used in this study did not include sulfide data.

Depth and Soil Horizons

Previous studies have suggested that arsenic concentrations in soil may increase with depth (Vosnakis et al. 2010; Ohio EPA 2013b). Most sampling locations in the present study's dataset had only one sample collected, making correlation to depth
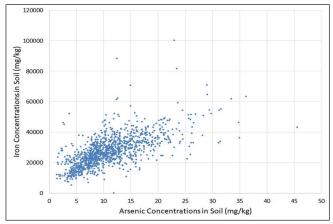

FIGURE 6. Arsenic concentrations in soil versus iron concentrations in soil (n=883)

difficult. Because there is great variety in soil types, soil horizon thicknesses, infiltration rates, and other factors, it is not appropriate to perform a direct comparison of sampling depth to arsenic concentrations. Where samples were collected from multiple depths at a single location, a simple comparison was made to determine if arsenic concentrations in soil increased or decreased

with depth for each sample at each location. This comparison found 131 samples having greater arsenic concentration with increasing depth, 93 samples having reduced arsenic concentrations with increasing depth, and 9 samples had no change in arsenic concentration with depth. Although the data were not statistically evaluated, the data were likely not sufficient to

Table 2
Difference between median arsenic concentrations in samples
from various parent materials and all other samples[a]

Difference in median arsenic concentration[b]	Parent material group
8.5	Residuum weathered from shale
6.8	Till over residuum weathered from sandstone and siltstone
5.5	Silty and clayey till over residuum
5.1	Till derived from shale
3.8	Wisconsinan loamy till derived from sandstone and shale
3.4	Loamy alluvium
3.3	Loamy till
3.1	Residuum weathered from shale and/or till
2.9	Silty material or loess over loamy till
2.9	Till
2.7	Wisconsinan till derived from limestone and shale
2.7	Silty and clayey outwash over sandy and gravelly outwash
2.3	Outwash
2.3	Loess over residuum weathered from limestone and shale
1.5	Sandy outwash and/or loamy outwash
1.0	Fine-loamy alluvium derived from sedimentary rock
-0.6	Glaciolacustrine deposits
-1.2	Alluvium
-1.4	Loess over alluvium over lacustrine deposits
-1.5	Colluvium over residuum
-1.5	Loamy alluvium over sandy and gravelly alluvium
-1.7	Loess over residuum
-3.3	Loess over Illinoian till
-3.3	Loess over pedisediment over till
-3.9	Clayey glaciolacustrine deposits over clayey till
-5.1	Loamy outwash over sandy and gravelly outwash
-7.2	Sandy glaciolacustrine deposits over loamy glaciolacustrine deposits

[a] Only statistically significant results calculated from a minimum of 20 observations are shown.
[b] mg kg^{-1}

demonstrate a statistically valid correlation between sampling depth and arsenic concentrations. However, depth may factor into multivariate predictions of arsenic concentrations, with some soil types perhaps showing stronger correlation to depth than others.

Evaluation of soil horizons was similarly inconclusive. The A horizons tended to have slightly higher median arsenic concentrations than either the underlying B horizons or overlying organic (O and H) horizons. However, these differences were not demonstrated to be statistically significant; correlations were weak, too few samples were available, or both. It is also possible that increasing natural arsenic concentrations with depth are masked at some locations by historical agricultural pesticide application.

The reader is referred to a 2008 USGS study of the relationship between solid phase and dissolved arsenic in groundwater underlying Preble County, Ohio, for additional discussion of how arsenic concentrations can change with depth (Thomas et al. 2008).

Soil Texture/Class

Significant correlations were observed between soil classification and arsenic concentrations. Silts and clays, as classified in the Unified Soil Classification System (USCS) as CL-ML, had median arsenic concentrations 3.0 mg kg^{-1} higher than soils other than silts and clays. Sands (designated with a SC-SM designation under USCS) had median arsenic concentrations 3.7 mg kg^{-1} lower than soils other than sands. Fine sand (as classified under the USDA classification system—USCS does not differentiate fine and coarse sands) had a median arsenic concentration 7.7 mg kg^{-1} lower than soils other than fine sand.

Hillslopes

Statistically significant correlations were observed between arsenic concentrations and the location of the samples with respect to hills ("geomposhill" field of "cogeomordesc" table in SSURGO database). Samples from nose slopes and base slopes showed median arsenic concentrations 1.9 mg kg^{-1} and 0.6 mg kg^{-1} higher than samples not from nose slopes and base slopes, respectively. Conversely, samples from side slopes and head slopes had median arsenic concentrations 0.8 mg kg^{-1} and 2.1 mg kg^{-1} lower than samples not from side slopes and head slopes, respectively. Correlations to hill crests and interfluves were not significant. It is possible that observed correlations could be representative of the amount of water flowing across soils prior to

reaching a stream channel, although additional data and research is recommended before any conclusions can be reached.

Ponding

A USGS study performed in Licking County, Ohio, found that elevated arsenic concentrations in groundwater are associated with strongly reducing conditions that promote dissolution of arsenic (Thomas 2016). Although redox levels are not reported in SSURGO, the database does include some data regarding the frequency and duration of ponding. Anticipating that reducing conditions associated with ponding could potentially leach arsenic from flooded soils, arsenic concentrations in soil were compared to ponding frequency and duration.

The SSURGO database's "comonth" table contains fields regarding ponding duration ("ponddurcl") and ponding frequency ("pondfreqcl"). Areas of brief ponding duration (2 to 7 days) had median arsenic concentrations 2.76 mg kg^{-1} higher than areas that did not exhibit brief ponding, while areas of long ponding duration (7 to 30 days) had median arsenic concentrations 1.8 mg kg^{-1} lower than areas that did not exhibit long ponding. Areas of very-long ponding duration (more than 30 days) had median arsenic concentrations 2.64 mg kg^{-1} lower than areas that did not exhibit very-long ponding. Similarly, areas of no ponding frequency had median arsenic concentrations 0.8 mg kg^{-1} higher than areas that experienced ponding, while areas of frequent ponding had median arsenic concentrations 0.8 mg kg^{-1} lower than areas without frequent ponding.

DISCUSSION
Bedrock

Statistical evaluation shows a significant correlation between bedrock type and soil arsenic concentrations. Soil materials above Devonian-aged bedrock, particularly black shale, tend to have higher arsenic concentrations than soils over other bedrock types. However, as shown in Fig. 7, when arsenic concentrations (interpolated by kriging) are mapped with Devonian-aged bedrock, there are a number of areas where elevated arsenic concentrations are expected, but not observed. In particular, Venteris (2014) noted that the northwestern portion of the state was underlain by Devonian-aged bedrock, but had low arsenic concentrations. Venteris hypothesized that arsenic may have been leached from soils, or an
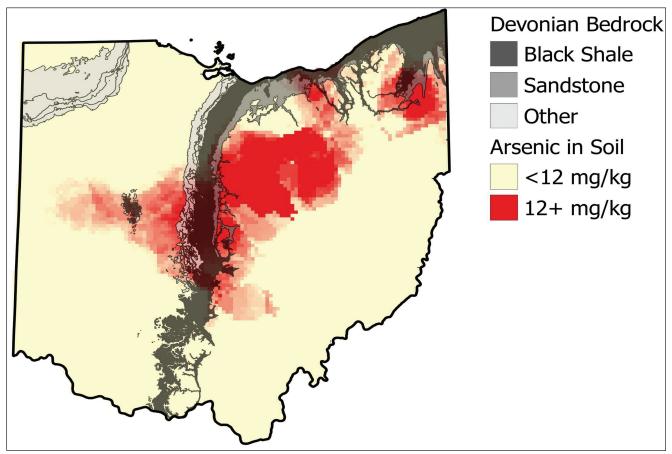
FIGURE 7. Spatial overlay of Devonian-aged bedrock materials and areas of Ohio with mean arsenic concentrations in soil greater than 12 mg kg⁻¹

unidentified soil characteristic may be influencing concentrations. Closer review shows that Devonian-aged bedrock in the northwestern part of the state consists of limestone, dolostone, and shale. This is in contrast to the black shale and sandstone found in the central and northeastern portions of the state. These rock types are formed in different depositional environments, suggesting an additional variable that may partly explain the lower arsenic concentrations in northwest Ohio. The lower concentrations may also result from arsenic being leached from soils by reducing conditions of the historical Great Black Swamp.

The south-central area of Ohio also contains Devonian-aged bedrock, but does not exhibit elevated arsenic concentrations in soil. This is not presently understood, but may be a result of few samples being collected directly from bedrock residuum. In northern parts of the state, glacial activities distributed Devonian-aged bedrock fragments over large areas. However, glacial activities did not extend to the south-central Ohio border, thus glacial distribution of Devonian-aged bedrock fragments would not have taken place in southern Ohio. Eroded bedrock fragments would have instead been concentrated into stream channels.

As a result, the limited number of samples collected (Fig. 2) were less likely to come from areas influenced by the narrow band of Devonian-aged bedrock (Fig. 7), and the dataset of bedrock-influenced samples may have been too small to demonstrate statistically significant correlations. Additionally, sampling in the southern and eastern portions of Ohio may be biased toward floodplains between hills, where access to soils along roadways is simpler. As discussed below, erosion and transport mechanisms (which result in floodplain deposits) are likely to remove arsenic from the source materials prior to deposition.

Glaciation

In the north-central and northeastern parts of Ohio are areas of elevated arsenic concentrations that are not overlying Devonian-aged bedrock. Figure 8 shows areas of the state where surficial materials were deposited from ice (till, ground moraine, etc.), as opposed to water. The areas of elevated arsenic concentrations in the north-central and northeastern parts of the state are generally found near the terminus of glacial activities that advanced across areas of Devonian black shales. Furthermore, it can be seen that surficial materials
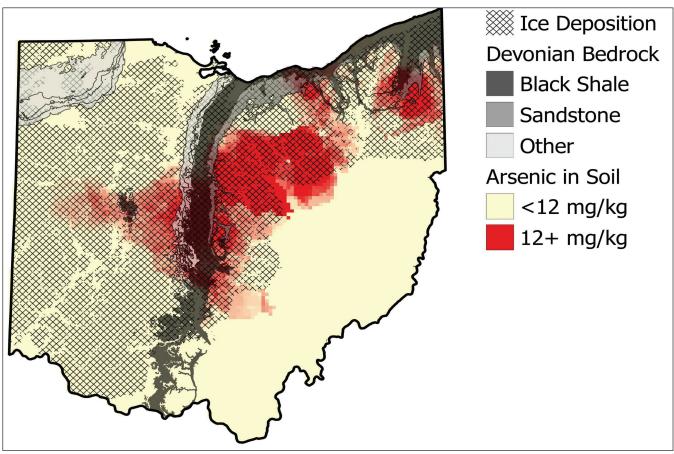
FIGURE 8. Spatial overlay of areas of surficial deposits from ice sheets, Devonian-aged bedrock materials, and mean arsenic concentrations in soil greater than 12 mg kg^{-1}
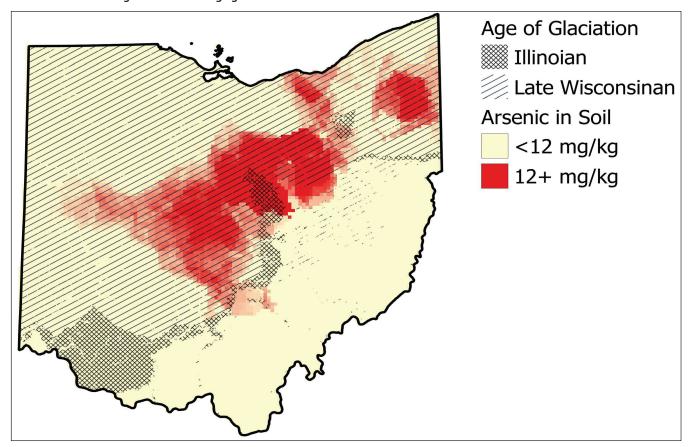

FIGURE 9. Spatial overlay of glacial advances and mean arsenic concentrations in soil

overlying Devonian bedrock in the south-central and northwestern areas of the state are less influenced by deposition from ice than are areas of Devonian bedrock in the central and northeastern areas of the state. In the south-central area of the state, the Scioto River flows south across the Devonian black shale in what was formerly the Teays River Valley. Thick deposits of sand and gravel are present along this river valley. The northwestern area of the state is influenced by lacustrine deposits.

Survival analysis tests indicated that the ages of glaciation may influence arsenic concentrations. Fig. 9 shows the areas of the state where deposits from glacial advances are at the surface. The Late Wisconsinan glaciation advanced over the Illinoian, but did not extend as far south. This resulted in an area in the southwestern part of the state, east of Cincinnati, where surficial materials are Illinoian in age. A narrow band of Illinoian deposits, not covered by Late Wisconsinan deposits, extends northeast from this area into central Ohio.

Two possible explanations for the differences in arsenic soil concentrations between Illinoian and Late Wisconsinan deposits are presented. First, glaciation in the western portion of the state seldom advanced across Devonian black shales. The dataset used in this study contained a greater number of samples from the larger exposure of Illinoian materials in the southwestern portion of the state as opposed to the narrow band extending northeast into central Ohio. With proportionately fewer samples collected from Illinoian deposits in the central portion of the state (where Devonian shales are located) arsenic concentrations representing Illinoian glaciation may have been biased low. If additional samples were to be collected from Illinoian deposits in central Ohio, they may show that the bedrock materials crossed by ice sheets have greater influence on arsenic concentrations than the time period during which glaciation occurred. A second possible explanation is that arsenic has been removed from the older Illinoian deposits through leaching. The Illinoian glaciation occurred between 130,000 and 191,000 years ago, while the Late Wisconsinan glaciation occurred only 13,000 to 24,000 years ago. As a result, materials deposited during Illinoian glaciation have been subjected to leaching for considerably longer than Wisconsinan deposits.

Although no correlation was found between drift thickness and arsenic concentrations in soils, this is not in conflict with the Thomas (2016) study conducted in Licking County. Thomas examined groundwater concentrations and hypothesized that thick layers of glacial till reduced oxygen concentrations in groundwater, resulting in reducing conditions more favorable for arsenic dissolution in groundwater. Because the majority of soil samples evaluated in the present study were surficial, and Thomas examined deeper groundwater, the processes influencing arsenic concentrations in each media are disparate.

Post-Deposition Erosion
While Devonian black shales are a primary source of elevated arsenic concentrations in Ohio, transport of the materials by ice sheets—and subsequent erosion of materials following glacial deposition—significantly influences arsenic concentrations in soil. Glacial till deposited from ice, with little or no post-depositional erosion by water, tends to have higher concentrations of arsenic. As materials are eroded and deposited in streams and lakes, arsenic concentrations decrease. Windblown deposits of loess, which tend to be highly-eroded fine-grained sands and silts, have very low arsenic concentrations. Areas of frequent ponding, such as wetlands, also have lower arsenic concentrations, possibly as a result of arsenic having been leached from soils in reducing conditions.

CONCLUSION

Elevated concentrations of naturally occurring arsenic in Ohio soils appear to be associated with bedrock materials. However, this correlation is complicated by both glaciation and post-glaciation erosion. As glaciers advanced across Ohio, bedrock materials were eroded and deposited farther south. Arsenic concentrations in Ohio soils tend to be highest where Devonian-aged bedrock materials, particularly black shales, were deposited directly from ice sheets. Following glacial deposition of materials, arsenic can be eroded or leached from the deposits. Deposits of wind-blown loess and materials deposited by water (as opposed to ice) had lower arsenic concentrations than glacial tills and outwash plains that underwent minimal erosion following glacial deposition. Because depositional environments represented by soils often vary over relatively short distances, heterogeneity can be expected within regional arsenic concentrations.

University, for guidance and technical review during the conduct of this research. Glen Lapsley of the US EPA Region III was also a tremendous help in locating background data for areas of Ohio with sparse coverage.

LITERATURE CITED

ATSDR. 2007a. Public Health Statement: Arsenic. Agency for Toxic Substances and Disease Registry. (Accessed 8 Apr 2017 https://www.atsdr.cdc.gov/ToxProfiles/tp2-c1-b.pdf)

ATSDR. 2007b. ToxGuide™ for Arsenic. Agency for Toxic Substances and Disease Registry. (Accessed 22 Jun 2016 http://www.atsdr.cdc.gov/toxguides/toxguide-2.pdf)

Columbia Gas Transmission Corp. 2016. US EPA Freedom of Information Request EPA-R3-2016-006345. Excerpts of Characterization Reports for 45 Sites in Ohio.

Cox CA, Colvin GH. 1996. Evaluation of Background Metal Concentrations in Ohio Soils. Cox-Colvin & Associates, Inc. (Accessed 17 Nov 2015 http://www.coxcolvin.com/documents/OhioBackgroundMetals.pdf)

EPA. 2015. Facility Registry System State Download Files. (Accessed 31 Oct 2015 https://www.epa.gov/enviro/epa-state-combined-csv-download-files)

EPA. 2016. Coal Ash Basics. United States Environmental Protection Agency. (Accessed 8 Apr 2017 https://www.epa.gov/coalash/coal-ash-basics)

Gautier L. 2016. rpy2 version 2.8.5. (Accessed 21 Dec 2016 https://rpy2.readthedocs.io/en/version_2.8.x/index.html)

Helsel DR. 2012. Statistics for Censored Environmental Data Using Minitab® and R. Hoboken, New Jersey: John Wiley & Sons, Inc. https://doi.org/10.1002/9781118162729

Lee L. 2013. Package "NADA" for R: Nondetects and Data Analysis for Environmental Data. https://cran.r-project.org/web/packages/NADA/NADA.pdf

Ohio DNR. 2015. Ohio Quaternary Map. Ohio Geology Store. (Accessed 17 Nov 2015 http://www.ohiogeologystore.com/browse.cfm/cd-rom-ohio-quaternary-map/4,59.html)

Ohio EPA. 2009. Unpublished data compiled by Ohio EPA Background Study Group.

Ohio EPA. 2013a. Evaluation of Background Metal Soil Concentrations in Cuyahoga County - Cleveland Area: Summary Report. Columbus, Ohio: Division of Environmental Response and Revitalization, Voluntary Action Program. (Accessed 31 Oct 2015 http://epa.ohio.gov/portals/30/vap/docs/Cleveland%20Background%20Summary%20Report%20%282%29.pdf)

Ohio EPA. 2013b. Evaluation of Background Metal Soil Concentrations in Franklin County - Columbus Area: Summary Report. Columbus, Ohio: Voluntary Action Program. (Accessed 31 Oct 2015 http://epa.ohio.gov/portals/30/vap/docs/Columbus%20Background%20Summary%20Report.pdf)

Ohio EPA. 2014. Evaluation of Background Metal Soil Concentrations in Montgomery County - Dayton Area: Summary Report. Columbus, Ohio: Voluntary Action Program. (Accessed 31 Oct 2015 http://epa.ohio.gov/portals/30/vap/docs/Dayton%20Background%20Summary%20Report.pdf)

Ohio EPA. 2015a. Evaluation of Background Metal Soil Concentrations. Ohio EPA. (Accessed 17 Nov 2015 http://epa.ohio.gov/portals/30/vap/docs/Evaluation%20of%20Background%20Metal%20Soil%20Concentrations.pdf)

Ohio EPA. 2015b. Evaluation of Background Metal Soil Concentrations in Hamilton County - Cincinnati Area: Summary Report. Ohio EPA: Ohio Voluntary Action Program. (Accessed 24 Dec 2016 http://epa.ohio.gov/portals/30/vap/docs/Hamilton%20Background%20Summary%20Report.pdf)

Ohio EPA. 2015c. Evaluation of Background Metal Soil Concentrations in Lucas County - Toledo Area: Summary Report. Columbus, Ohio: Voluntary Action Program. (Accessed 26 Dec 2016 http://epa.ohio.gov/portals/30/vap/docs/Toledo%20Background%20Summary%20Report_FINAL.pdf)

Ohio EPA. 2015d. Evaluation of Background Metal Soil Concentrations in Summit County - Akron Area: Summary Report. Columbus, Ohio: Voluntary Action Program. (Accessed 26 Dec 2016 http://epa.ohio.gov/portals/30/vap/docs/Summit%20Background%20Summary%20Report.pdf)

Ohio EPA. 2016a. Ohio Administrative Code (OAC) 3745-300-08, Rule 8 Appendix, Table 1, Effective 05/26/2016. (Accessed 12 Mar 2017 http://epa.ohio.gov/portals/30/Rules/2016/Appendix%20-%2008.pdf)

Ohio EPA. 2016b. Evaluation of Background Metal Soil Concentrations in Mahoning County. Columbus, Ohio: Unpublished - personal correspondence with Sydney Poole on 3 Jun 2016.

Ohio EPA. 2016c. Evaluation of Background Metal Soil Concentrations in Stark County. Columbus, Ohio: Unpublished - personal correspondence with Sydney Poole on 3 Jun 2016.

R Development Core Team. 2017. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. (Accessed 28 Jan 2017 http://www.R-project.org)

Soil Survey Staff. 2016. Soil Survey Geographic (SSURGO) Database. Natural Resources Conservation Service, United States Department of Agriculture. https://doi.org/10.15482/USDA.ADC/1242479

Thomas MA. 2016. Arsenic in Groundwater of Licking County, Ohio, 2012—Occurrence and Relation to Hydrogeology. U.S. Geological Survey Scientific Investigations Report 2015-5148, 38 p. https://doi.org/10.3133/sir20155148

Thomas MA, Diehl SF, Pletsch BA, Schumann TL, Pavey RR, Swinford EM. 2008. Relation between solid-phase and arsenic in the ground-water system underlying northern Preble County, Ohio. US Geological Survey Scientific Investigations Report 2008-5205.

USGS. 2015a. Ohio Geologic Map Data. USGS Mineral Resources On-Line Spatial Data. (Accessed 17 Nov 2015 http://mrdata.usgs.gov/geology/state/state.php?state=OH)

USGS. 2015b. National Geochemical Survey Database. USGS Mineral Resources On-line Spatial Data. (Accessed 17 Nov 2015 http://mrdata.usgs.gov/geochem/)

Venteris ER. 2014. Modeling Spatial Patterns in Soil Arsenic to Estimate Natural Baseline Concentrations. Journal of Environmental Quality. 43(3):936-946. https://doi.org/10.2134/jeq2013.11.0459

Vosnakis KA, Perry E, Madsen K, Bradley LJ. 2010. Background Versus Risk-Based Screening Levels - An Examination Of Arsenic Background Soil Concentrations in Seven States. Proceedings of the Annual International Conference on Soils, Sediments, Water and Energy. Vol. 14 (Article 10):97-117. (https://scholarworks.umass.edu/cgi/viewcontent.cgi?article=1093&context=soilsproceedings)

Welch AH, Westjohn DB, Helsel DR, Wanty RB. 2000. Arsenic in ground water of the United States: occurrence and geochemistry. Ground Water. 38(4):589–604. https://doi.org/10.1111/j.1745-6584.2000.tb00251.x