Jonathan L. Bradshaw is with the Department of Civil and Environmental Engineering, Princeton University (www.princeton.edu), Princeton, N.J. Arun Kumar is with the Department of Civil and Environmental Engineering, Michigan State University (www.msu.edu), East Lansing, Mich. Patrick L. Gurian and David E. Breen are with the Department of Civil, Architectural, and Environmental Engineering and Department of Computer Science, respectively, Drexel University (www.drexel.edu), Philadelphia.

Spatial Trends Shed Light on Arsenic Concentrations

Identifying and understanding where arsenic occurs in groundwater can be difficult. Geospatially referenced visualization data could provide a solution. JONATHAN L. BRADSHAW, PATRICK L. GURIAN, ARUN KUMAR, AND DAVID E. BREEN

TO BETTER UNDERSTAND geographic trends in arsenic occurrence, researchers recently used an interactive visualization technique to link geospatially referenced arsenic concentration information from a water quality database with data contained in digital terrain elevation data (DTED) files. DTED files allow users to develop 3-D plots of arsenic concentration and topography. For example, the plots show high arsenic is

- associated with topographic transition from plains to piedmont in New Jersey.
- found in Oklahoma in a region of transition to lower elevation.
- apparent in the southeastern portion of California's Central Valley.

These results are consistent with how arsenic moves in organic-rich sediments of river valleys. However, further statistical analysis is required to confirm the significance of this association.

A UNIQUE APPROACH

Abundant data on arsenic in groundwater are available but are often presented in forms difficult to interpret, such as in tables or spreadsheet databases. Visualization can present data in a way that facilitates interpretation of complex information and allows the data to be used to visually identify water contaminant hotspots and correlations among water quality parameters. Hypotheses can be formulated based on these visual, intuitive exploratory analyses and then tested rigorously using statistical methods. In addition, visualization can clarify where data gaps exist and identify areas in which more data collection is needed.

This project sought to develop a simple, user-friendly, interactive visualization technique that communicates complex information from otherwise incomprehensible data. The visualization approach was used to better understand which topographic features are associated with elevated arsenic concentrations.

METHODS

The study used the US Geological Survey arsenic point database, which contains water quality data for the United States and its territories. Each row in the database represents a different well. Organized by column, the database includes USGS station identification number; Federal Information Processing Standard state and county codes; latitude and longitude; depth; water usage code, such as private, industrial, municipal, etc.; site-use code; aquifer code; and 31 constituent values, such as arsenic, uranium, chloride concentrations, and pH.

Researchers used technical computing software that allows users to specify subset columns within a graphical user interface. Various constituents may be listed in the database but can consist only of numbers or blank cells. The software converts latitude and longitude from degrees–minutes–seconds to decimal degrees. However, with only minor programming modifications, any kind of spatial coordinate system is acceptable.

DTED files provide land elevation data that are useful for creating a surface plot and estimating well elevations, facilitating a user's recognition of a geographic area and permitting a user to view a well's depth adjusted to local elevation. DTED files are available for all land masses across the world.

The contaminant database is loaded directly into the computing software's programming environment from a spreadsheet file (Figure 1). Each column a user selects determines the plotting criteria and is stored as an individual vector variable. All constituent rank calculations, which determine color, are made after

The Dare County (N.C.) Water Department uses six arsenic removal vessels, arranged in two sets of three, filled with 40 in. of granular ferric oxide. The arsenic vessels were put on-line in November 2004 to meet the 10 µg/L federal maximum contaminant level for arsenic that went into effect in January 2005. The vessels continue to remove arsenic to below 1 µg/L.

PHOTOGRAPH: DARE COUNTY (N.C.) WATER DEPARTMENT

well and constituent selection. The computing software's mapping toolbox determines which DTED files are required to approximate the selected wells' elevations, and the appropriate DTED file is assigned to each well. Another mapping toolbox function translates DTED *.dt0 files into an elevation grid and reference vector. Each well's elevation is approximated by locating the well's position within the elevation grid and calculating the mean elevation of the nine closest points. A mapping toolbox function uses the elevation grid and reference vector produced from the DTED file to plot 3-D terrain.

For detailed visualization of county-level or smaller geographic areas, wells are indicated by icons that illustrate

- well depth, indicated by height.
- well type, indicated by the color of a cylindrical shell and outer circle.
- constituent concentrations, indicated by the darkness of colors of different sectors within the cylindrical shell of the icon.

Wells are represented on the plot by vertical lines parallel to the altitude axis with dot markers on each end. A line's length is equal to the well's depth. A well's location is determined by its latitude, longitude, and approximate elevation. Instead of incorporating and presenting all of a well's constituent and type information at once, less detailed plots visually present a single constituent with color and size used in the detailed plots.

Figure 1. Arsenic Point Database

The USGS arsenic point database consists of nearly 800,000 data points.

	A	B	C	D	E	F	G	H	I	AG	AH	AI	AJ
1	staid	States	FIPS Cour	Latitude	Longitude	Well Deptl	Water Use	Site Use	Aquifer Cod	p22703	p70301	p71888	p81366
2	15S	2S	3S	9N	9N	9N	1S	1S	8S	9N	9N	9N	9N
3	295737090051201	LA	22.071	295737	900512	811	A	W	112GZNO		534		
4	300054090012601	LA	22.071	300054	900126	658	A	W	112GZNO		574		
5	353026097274801	OK	40.109	353026	972748	164	A	W	318GRBR	3.05	711		
6	393931074482101	NJ	34.001	393932	744820	165	A	W	121CKKD	0.01			1.2
20254	484304099511301	ND	38.079	484304	995113	35	n/a	O	112SLVL		743		
20255	484416099523901	ND	38.079	484416	995239	25	n/a	T	112SLVL		392		
20256	484632099513001	ND	38.079	484632	995130	54	n/a	A	n/a		503		
20257	484731099504104	ND	38.079	484731	995041	38	n/a	O	112SLVL		493		
20258	485316122322601	WA	53.073	485316	1223226	20	n/a	W	112SUMS		155		
20259	485848122272701	WA	53.073	485848	1222727		n/a	W	112SUMS		90		

EPA

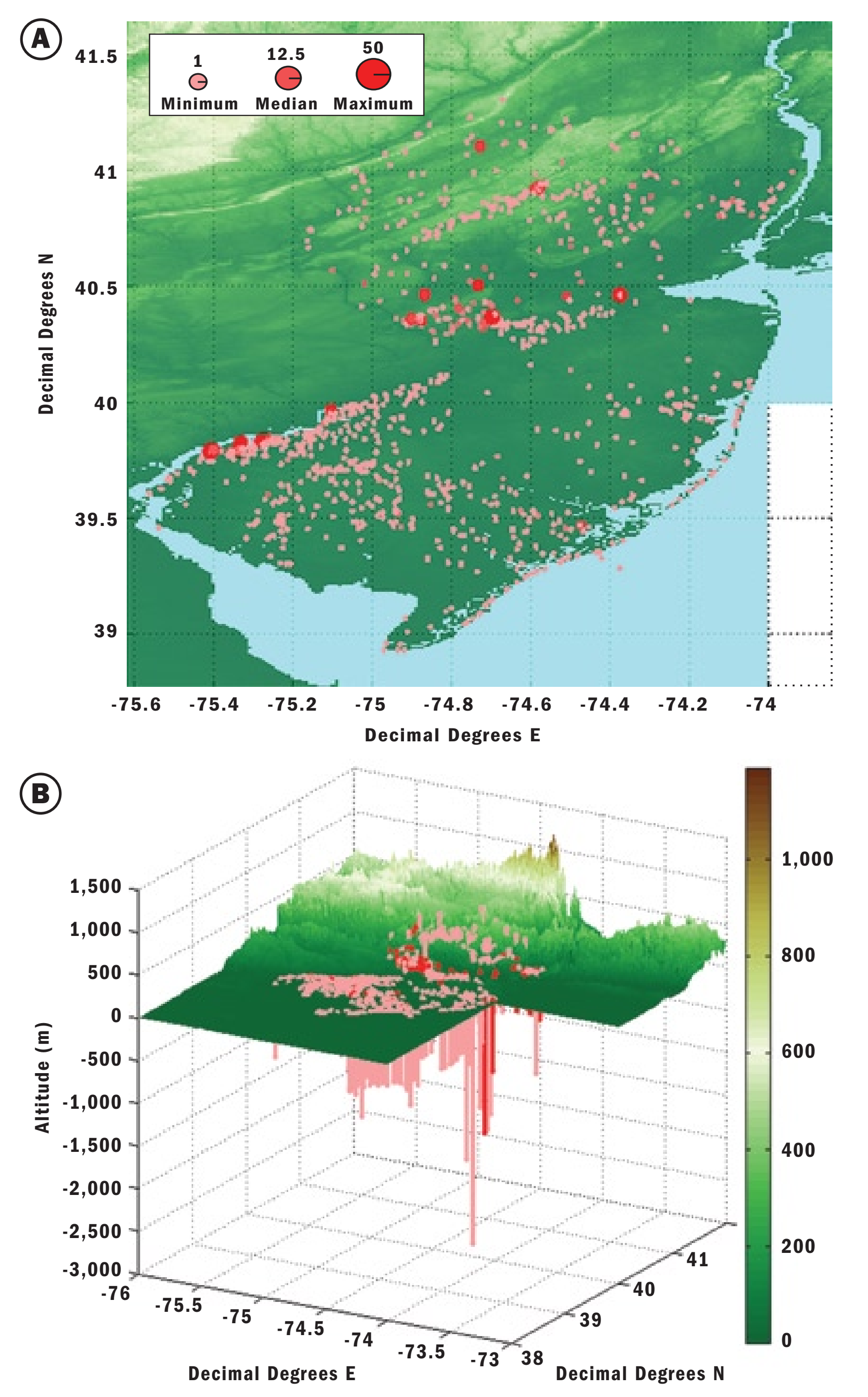
Ready

RESULTS

The visualization system's effectiveness was put to the test in evaluating arsenic occurrence in New Jersey, Oklahoma, and California. Previous studies identified these states as areas with elevated arsenic levels and, during exploratory analysis, demonstrated spatial trends of interest.

Figure 2. Arsenic in New Jersey

Visualization of arsenic distribution (in µg/L) in New Jersey reveals (A) horizontal spatial distribution and (B) vertical distribution of arsenic. An increase in arsenic concentration is indicated by an increase of darker circles.



New Jersey. Figure 2 shows a horizontal and vertical plot of New Jersey's arsenic distributions. Wells are shown in circles, and changes in arsenic concentrations are shown by variations in circle colors. For example, a color change from pink to red indicates an increase in arsenic concentration. Arsenic concentrations in the coastal plain are mostly low. Higher arsenic concentrations frequently aren't found until the coastal plain transitions into the piedmont. Wells with high arsenic levels are relatively shallow and located in a line approximately perpendicular to the elevation gradient, suggesting the wells may tap the same geological stratum.

Oklahoma. As shown in Figure 3A, the spatial distribution plot of arsenic for Oklahoma reveals serious gaps in available data for the state's north-central and southwest portions. The large number of wells in the center of the plot represents the Oklahoma City area, with the eastern side having mostly low-arsenic groundwater and the western side having noticeably more wells with high arsenic concentrations. A string of high-arsenic wells extends in a north–south direction roughly perpendicular to the elevation gradient, an indication that these wells tap the same geological stratum. This area of high arsenic corresponds to a transition from lower elevations at the state's center to higher elevations in the west.

California. As shown in Figure 3B, elevated arsenic concentrations are found in the southeastern portion of California's Central Valley region. However, the concentrations are somewhat lower elsewhere in the valley, an indication that within the same valley local geology influences arsenic occurrence trends. With few observations in the valley's northwestern portion, the plot reveals a gap in available data.

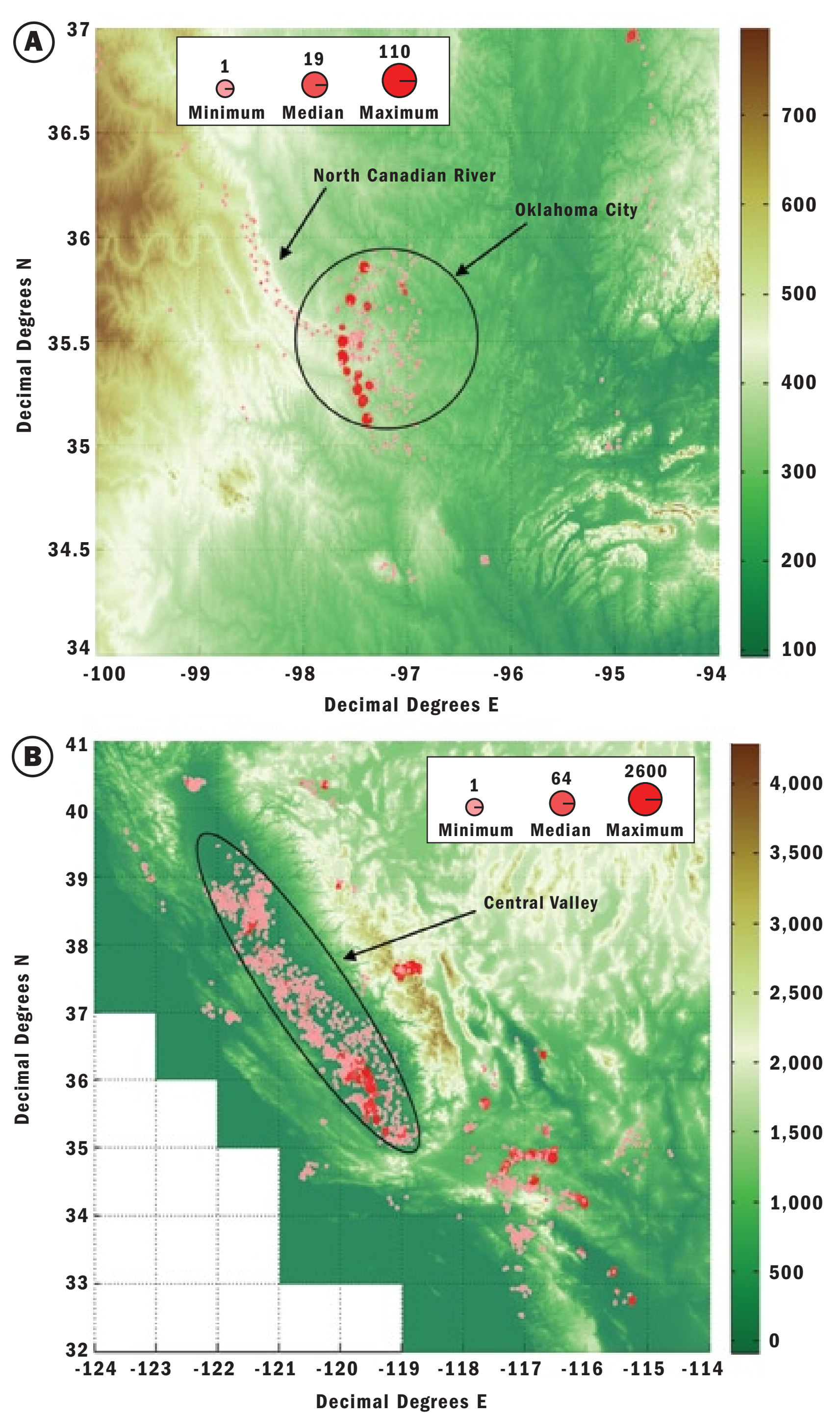
A PROMISING FUTURE

Visualization of arsenic concentration data revealed numerous localized hotspots and identified regions in which data gaps exist. This information may be useful to

Visualization of arsenic concentration data revealed numerous localized hotspots and identified regions in which data gaps exist.

Figure 3. Arsenic in Oklahoma (A) and California (B)

Visualization of spatial distribution of arsenic (in µg/L) reveals an increase in arsenic concentration, indicated by an increase in darker circles.



alert public health agencies to potential water quality problems or indicate areas where additional sampling is needed. Elevated arsenic levels can be associated with portions of particular river valleys or other changes in topography. However, not all river valleys or topographic transitions have elevated arsenic, and not all arsenic hotspots are associated with evident topographic features. Further statistical analysis is needed to assess if there's a significant association between particular topographic features and groundwater arsenic concentrations. Although this visualization approach was specifically developed for use with the USGS arsenic point database, it can support a wide variety of databases consisting of different water quality characteristics and geological information.

Authors' Note: *This research was funded by the Engineering Cities REU Program through NSF Grant No. EEC 0552792 [Engineering Cities]. For a copy of the software used in this analysis, contact jon.l.bradshaw@gmail.com. All DTED level 0 files are courtesy of the National Geospatial–Intelligence Agency and are available for free download (http://geoengine.nga.mil/muse-cgi-bin/rast_roam.cgi).*

RESOURCES

- Amy, Gary et al., 2005. *Adsorbent Treatment Technologies for Arsenic Removal*. AWWA and Water Research Foundation.
- Chang, Yu-Jung et al., 2006. *Advanced Processes for Simultaneous Arsenic and Manganese Removal*. AWWA, Water Research Foundation, and IWA Publishing.
- Chang, Yu-Jung et al., 2005. *Demonstration of Emerging Technologies for Arsenic Removal, Volume 2: Pilot Testing*. AWWA and Water Research Foundation.
- Samanta, Gautam and Clifford, Dennis A., 2006. *Preservation of Arsenic Species*. AWWA, Water Research Foundation, and IWA Publishing.