Spatial Fingerprinting of Biogenic and Anthropogenic Volatile Organic Compounds in an Arid Unsaturated Zone

Christopher T. Green,* Wentai Luo, Christopher H. Conaway, Karl B. Haase, Ronald J. Baker, and Brian J. Andraski

Subsurface volatile organic compounds (VOCs) can pose risks to human and environmental health and mediate biological processes. Volatile organic compounds have both anthropogenic and biogenic origins, but the relative importance of these sources has not been explored in subsurface environments. This study synthesized 17 yr of VOC data from the Amargosa Desert Research Site in Nevada with the goal of improving understanding of spatial and temporal variations that distinguish sources of VOCs from a landfill and from ambient sources including biogenic VOCs (bVOCs). Gas samples were collected from 1999 to 2016 from an array of shallow sample points (0.5- and 1.5-m depth) and from vertical profiles at three deep boreholes: two (109 m deep) near the border of a waste facility (33 and 100 m distant), and one (29 m deep) in a remote area 3 km to the south. Samples were analyzed for target VOCs and a subset was analyzed for non-target VOCs to enumerate a greater variety of potential bVOCs. Principal components analysis of the target and non-target VOCs provided an assessment of the spatial variability of VOCs originating from the landfill site and from ambient sources. Ambient VOCs occurred at all sample sites across a range of depths, and most were consistent with biogenic origins, indicating for the first time the presence of bVOCs in the deep unsaturated zone. Because some VOCs have both anthropogenic and biogenic sources, discrimination of sources can be important for estimating the extent and migration of anthropogenic plumes in arid unsaturated zones.

Abbreviations: ADRS, Amargosa Desert Research Site; bgs, below ground surface; bVOC, biogenic volatile organic compound; CFC, chlorofluorocarbon; LLRW, low-level radioactive waste; PC, principal component; PCA, principal component analysis; VOC, volatile organic compound.

The origin and fate of volatile organic compounds (VOCs) in the subsurface are key considerations for human and environmental health. The regional mass budgets of anthropogenic and biogenic VOCs (bVOCs) affect outdoor air quality as well as the formation of ozone and aerosols that affect atmospheric chemistry and climate (Jimenez et al., 2009). Subsurface plumes from spills and landfills can lead to the intrusion of VOCs into indoor spaces and the risk of chronic and acute health effects (Mitchell et al., 2007). While VOCs can pose risks to human and environmental health, they also serve important ecosystem functions—bVOCs in the atmosphere and subsurface mediate biological processes of plants and microbes and serve as signaling agents within and among species (Peñuelas et al., 2014). At the global scale, emissions of bVOCs surpass anthropogenic VOCs by about an order of magnitude (Geron et al., 2006), and bVOCs are of particular concern as contributors to hazardous air pollution such as ozone (Jardine et al., 2010). Despite the pervasive effects of VOCs in the biosphere, there are substantial knowledge gaps related to their subsurface sources, transport, and roles in ecological functions.

There are several challenges to identifying sources of VOCs, including uncertainty about the relative contributions of biological and anthropogenic sources. During previous decades, studies have found increasing numbers of biogenic sources. Starting in the 1950s and 1960s, plant sources were identified for isoprene (Sanadze, 2004) and terpene (Went, 1960) (both of which also have anthropogenic sources). Many studies have since elucidated the origins, biological pathways, and functions in the environment of isoprenoids.
and terpenoids (Kesselmeier and Staudt, 1999; Lalothaworkitkul et al., 2009). Additional classes of bVOCs have been identified that were generally assumed to be anthropogenic, such as benzenoids (Heiden et al., 1999; White et al., 2009, Jardine et al., 2010). Recent experimental and field observations have indicated that benzenoid emissions from vegetation rival those from fossil fuels (Miszta et al., 2015). Misztal et al. (2015) noted that the current lack of understanding of biological functions of some benzenoid compounds “mirrors the state of knowledge concerning many other important compounds of biogenic origin, including isoprene, two decades ago.” The ongoing discoveries related to sources of VOCs point to uncertainty in source classifications at contaminated sites. Commingling of anthropogenic and biogenic VOCs is recognized in regional atmospheric inventories (e.g., Pouliot and Pierce, 2009) but has not been addressed, to our knowledge, in subsurface investigations of contaminated sites.

The unsaturated zone has unique relevance to the study of VOC sources and fate. The transport characteristics of gases in the unsaturated zone can preserve a record of VOC emissions at a site at annual to multidecadal time scales, including potential information about VOC sources and reactivity. Unsaturated zone gas transport commonly occurs via diffusion (Altevogt et al., 2003), with diffusion coefficients that are several orders of magnitude greater than in the liquid phase. Transport can be further influenced by advection and density effects (Altevogt et al., 2003) as well as interactions with complex geology (Walvoord et al., 2014; Green et al., 2015). These gas transport processes are particularly relevant in arid unsaturated zones, where deep water tables allow greater distances and times of subsurface gas transport and low moisture contents provide increased pore space for transport. Such arid sites have been favored for use as waste repositories, such as at the Amargosa Desert Research Site (ADRS) and at Los Alamos National Laboratory (Stauffer et al., 2005; Behar et al., 2019). Because arid lands cover 35% of the world’s land area and are home to 20% of the world’s population (Tchakerian and Pease, 2015), VOC distribution in arid unsaturated zones has major importance to human and environmental health. Relatively little is known about bVOC cycling in arid soils (Peñuelas et al., 2014), and we know of no studies of bVOC occurrence in the subsoil unsaturated zone. The lack of information about VOC sources in arid unsaturated zones presents an opportunity to improve the understanding of the numbers and quantities of commingling VOCs at the scales of individual sites and regions and in the atmosphere.

Recent studies of desert ecosystems simplify the attribution of bVOCs in those settings. A plant community of *Larrea tridentata* (DC.) Coville (also known as creosote bush) and *Ambrosia dumosa* (A. Gray) Payne (also known as burro weed, burrobush, or white bursage) predominate across regions of the Mojave, Sonoran, and Chihuahuan deserts, which together encompass an area of 750,000 km² in the southwestern United States and northern Mexico. Allelopathic-root interactions between these two plant species have been attributed to biochemical inhibitors (Mahall and Callaway, 1991). Studies of bVOC emissions from shoot systems of desert plant species such as *L. tridentata* (Jardine et al., 2010) and *A. dumosa* (Geron et al., 2006), as well as from associated subsurface biota such as endophytic and pathogenic fungi (Strobel et al., 2011), have identified specific bVOCs that are likely to occur in those habitats. Roots are also well-known sources of bVOCs and extend to the maximum depth of water infiltration during the year in this region (e.g., 1 m, Andraski et al., 2005). Prior information about bVOCs around desert plants and soils may be combined with analysis of VOCs in the deep unsaturated zone to identify likely arid unsaturated zone bVOCs for consideration in regional mass balances and as background sources at contaminated sites.

The purpose of this study was to characterize the temporal and spatial variations of deep unsaturated zone VOCs and to develop spatial fingerprinting methods to distinguish biogenic from anthropogenic sources. Our overarching objectives were to improve delineation of anthropogenic plumes and to evaluate whether bVOCs are present in the deep unsaturated zone, below the root zone. The ADRS presents an opportunity to evaluate bVOCs in the subsurface beneath a desert ecosystem in undisturbed areas and disturbed areas in proximity to a low-level radioactive waste and hazardous waste facility (Fig. 1) that is a source of substantial subsurface anthropogenic VOCs. The site has produced a wealth of VOC data (Baker et al., 2012) and characterizations of transport processes (e.g., Striegl et al., 1996; Walvoord et al., 2004, 2005, 2008, 2014; Andraski et al., 2005; Mapers et al., 2005; Mapses et al., 2013; Green et al., 2015; Seltzer et al., 2017) that provide a context for unsaturated VOC occurrence. Baker et al. (2012) presented the spatial and temporal distributions of chlorofluorocarbon (CFC) and non-CFC classes of VOCs from 1999 to 2009 outside of the waste facility. In the current study, the dataset has been extended to 2016, with shallow samples from inside the facility and with non-target VOC analyses for a subset of samples in support of an expanded characterization of spatial and temporal variations. Data from the ADRS are combined with previously published characterizations of bVOCs emitted from southwestern-desert vegetation to identify likely bVOCs and to characterize their vertical extents in the unsaturated zone.

Methods: Field Site, Sampling, Analysis, and Statistical Approach

Field Site and Sample Locations

The ADRS is located 17 km south of Beatty, NV, and 2 km northeast of the intermittent Amargosa River (Fig. 1A). The area has an aridic soil moisture regime and thermic soil temperature regime. Field measurements and numerical models indicate that unsaturated zone water fluxes below the root zone (and recharge) are negligible under undisturbed vegetation (Andraski, 1997; Walvoord et al., 2004). Root-zone soil-water content varies with time from 0.02 to 0.14 m³m⁻³, with precipitation primarily during winter months (January–March); a gravelly layer at a depth of 1 m has a relatively constant soil water content of approximately 0.05 m³m⁻³, and water content below this depth shows
little temporal variability (Andraski, 1997; Andraski et al., 2005; Kauble et al., 2018). Rooting depths are limited to 1 m or less, which is the typical maximum depth of infiltration during the year (Andraski et al., 2005). Precipitation averaged 0.10 m yr$^{-1}$ from 1981 to 2016 (Kauble et al., 2018). Vegetation consists primarily of *L. tridentata* and *A. dumosa*. Shrub species cover about 6% of the land area, and *L. tridentata* constitutes about 83% of the vegetative cover (Garcia et al., 2009). Unsaturated zone sediments include unconsolidated debris-flow, fluvial, and alluvial-fan deposits.

The southwestern portion of the waste facility (Fig. 1B) was a low-level radioactive waste (LLRW) facility, active from 1962 to 1992. This area received mixed waste including LLRW, VOCs, acids, bases, and other hazardous chemicals; regulations before 1985 did not require separation of LLRW from other waste (Maples et al., 2013). Site operators excavated sediments, deposited waste, and backfilled the trenches. The unlined trenches received containers of solid waste, solidified liquid waste (in Portland cement), and some direct discharge of liquids (Striegl et al., 1996; Maples et al., 2013). The waste trenches ranged from 2 to 15 m deep. During operations, the trench covers were cleared of vegetation. After closure in 1992, operators installed a ≥2-m thick layer of stockpiled sediments over the waste trenches. The facility currently disposes of hazardous chemical waste in lined trenches north of the LLRW area.

Studies at the ADRS have characterized VOC occurrence and gas transport properties. Baker et al. (2012) reported the types of VOCs detected and their vertical and horizontal distributions outside of the boundaries of the waste facility. The most abundant VOCs were CFCs, common refrigerants since the 1960s. Concentrations of CFCs were as high as 37,000 ppbv (parts per billion by volume, equivalent to 1 cm$^3$ of analyte per 10$^9$ cm$^3$ of air) at 35.1-m depth in Borehole UZB-3 (Fig. 1B). The CFCs and xylenes were the most commonly detected VOCs. A 1-yr controlled tracer test at the ADRS using SF$_6$ showed that gas diffusion was a primary transport mechanism at scales of tens of meters, and pronounced density effects were not observed (Walvoord et al., 2014). Studies of longer term transport of plumes at ADRS have indicated greater extents of contamination than expected from pure diffusion (Striegl et al., 1996; Mayers et al., 2005; Walvoord et al., 2008; Maples et al., 2013), possibly stemming from additional transport processes such as interactions of barometric pumping (advection) with geological layering (Green et al., 2015). The molecular mass of the gases influences the steady-state vertical distribution of stable isotopes at the site (Seltzer et al., 2017). Potential effects of macroscopic gas density on vertical distributions during transport have not been characterized at the ADRS.

The conceptual model of VOC sources in the unsaturated zone at the ADRS includes multiple potential pathways (Fig. 2). Primary ambient sources of interest include local biological emissions and regional atmospheric transport from remote biogenic and anthropogenic sources. The landfill source includes the trenches in the waste facility. Biological sources may emit VOCs in the subsurface directly, or aboveground emissions may enter the subsurface by diffusion and by barometric pumping (Auer et al.,
1996; Choi et al., 2002; Stauffer et al., 2019). Elevated CO₂ near the water table at the ADRS indicated potential georespiration (Walvoord et al., 2005; Conaway et al., 2018), in which microbes respire relict carbon. Global emissions have led to accumulation in the atmosphere of refractory VOCs such as the CFCs, which has led to their use as tracers of atmospheric exchange with the unsaturated zone (Weeks et al., 1982) and as tracers of the time of recharge of groundwater (Busenberg and Plummer, 1992). For deep samples from UZB2, UZB3, and JFDB below 2 m, emissions from polymer plastic tubes used to access the deep unsaturated profile (as opposed to steel tubes used for samples <2 m) are a potential source, as discussed below.

Deep and Shallow Unsaturated Zone and Atmospheric Air VOC Sample Collection and Analysis

Samples from the deep unsaturated zone were collected from three boreholes at the UZB2, UZB3, and JFDB sites. The UZB2 and UZB3 boreholes are in the buffer zone outside the waste facility and are 160 and 100 m south, respectively, of the nearest waste trench (Fig. 1B). The JFDB site is 3 km south of the facility (Fig. 1A) and serves as a background site beyond the influence of the waste facility. These boreholes were previously installed for unsaturated zone gas research and have been used in numerous fate and transport studies. Each borehole is instrumented with multiple, vertically nested sampling points. Boreholes UZB2 and UZB3 extend to depths of >100 m (Prudic et al., 1997) and use 30-cm-long, stainless steel screened intervals connected to 6-mm outer diameter nylon tubing. The deep sample point (28.7 m below ground surface [bgs]) at the JFDB site uses a 60-cm-long polyvinyl chloride screen and 6-mm outer diameter polyethylene tubing (Baker et al., 2012). Shallow 0.5- and 1.5-m sample points were constructed of hand-driven, black steel pipe (10.29-mm outer diameter, 5.47-mm inner diameter) (Andraski et al., 2003). Installation details were provided by Prudic et al. (1997) and Andraski and Scanlon (2002).

Samples at depths of 0.5 m (in the rooting zone) and 1.5 m (below the typical rooting depth) were collected from outside (1999–2016) and inside (2012–2016) the LLRW area (Fig. 1). Samples from the deep unsaturated zone, atmospheric-air samples were collected in 2001 and 2008 and were analyzed for target VOCs. For the 2001 samples, the air-tube intake was set to 0.5 m above the land surface within the L. tridentata canopy at locations A2 (X = 501, Y = 374; Fig. 1; Green et al., 2019) and A7 (X = 500, Y = 251). In 2008, an air sample was collected with the intake tube at 1 m above the land surface (not within the canopy of a plant) near location BB2 (X = 549, Y = 375).

Unsaturated zone samples were collected by first purging the tubes (see below) and then pumping a volume (2 L in 1999, 1 L in other years) through VOC-adsorbent resin cartridges at 40 mL min⁻¹. Information about the laboratory analysis is provided in the Supplemental Material. The purge volumes before VOC sampling varied due to other gas sampling activities at the same locations and unexpected circumstances such as clogging of cartridges. Target purge volumes were one to a few volumes of the sample tube but ranged in practice from zero to hundreds of sample-tube volumes (Green et al., 2019). The relationships between VOC concentrations and the variation of purge volumes was used to screen for potential effects of tube-emitted VOCs in the UZB borehole samples below 1.5 m (0.5- and 1.5-m depth samplers did not use plastic tubing). Because the correlation between purge volume and concentration can also stem from concentration gradients in the formation, and the analysis treats any correlation as stemming entirely from tubing emissions, the analysis is intended to provide a minimum lower bound on concentrations rather than a maximum likelihood estimate. As discussed below, results indicated that potential effects were typically minor (e.g., less than a factor of 2) at sample depths shallower than 50 m (including all the sample locations used for principal component analysis), and the presence and absence of VOCs was consistent among the steel tube samples and the UZB samples. Further information about this procedure is provided in the Supplemental Material.

To characterize a broader range of potential bVOCs than available in the target analyses, 12 unsaturated-zone samples were selected for analysis of non-target VOCs. The 12 samples were selected to provide comparable numbers of samples from in, around, and distant from the waste facility at different depths. Six samples were selected from two sampling dates (2013 and 2014) at three depths (0.5, 1.5, and 28.7 m bgs) at JFDB to provide multiple samples distal to the source location and to check for possible temporal variability of non-target analytes. The remaining six samples were from in and around the LLRW area, including three depths at UZB3 (0.5, 1.5 and 23.8 m bgs) in 2014 and three additional shallow samples—one outside the facility (Site X300Y650), one at the edge (X400Y650), and one above a waste trench (X488Y480). Information about quantification of
non-target VOC concentrations is provided in the Supplemental Material. All target (including commonly analyzed VOCs for landfill sites) and non-target VOC data are available as a data release (Green et al., 2019).

**Statistical Methods and Spatial Characterization of VOCs for Source Discrimination**

Principal components analysis (PCA) was used to characterize differences in the spatial distributions of VOCs and to identify likely VOC sources (Fig. 2). Principal components analysis has long been used for exploratory data analysis to characterize the variance of datasets. Principal components are calculated based on eigenvalues of the covariance matrix to transform the original variables into linearly uncorrelated variables known as principal components (PCs). The PCs are linear combinations of the original variables such that PC1 captures the greatest variation of the data, PC2 the second greatest variation, and so on. Omitting the less influential PCs allows the problem to be viewed in reduced dimensions and helps to discern patterns in the underlying structure of the data. Principal components analysis results are often presented as the loadings, which quantify the extent to which the original variables (e.g., sample locations) contribute to the transformed variables, and the scores, which are the transformed observations. In a spatial analysis such as in the current study with multiple observations (e.g., VOCs) at each location, the locations can be assigned as the variables so that the analysis shows the relevance of individual locations (via loadings) to the data variability as well as the primary trends and groupings of the observations (via scores). Principal components analysis has been used for this type of fingerprinting in a variety of applications, for example analyzing facial features for recognition of individuals (Turk and Pentland, 1991) or in environmental studies to characterize petroleum biomarkers (Christensen et al., 2005). The approach of the current study is analogous, with assignment of sample locations as variables and VOC types as observations to analyze VOC spatial characteristics that can serve as “fingerprints” of source types.

The main VOC sources of interest in the current study were ambient sources (e.g., biogenic and atmospheric) and the landfill emissions from the LLRW facility. Principal components analysis was applied first to the non-target compounds, which included a greater number of potential ambient VOCs than the target compounds, to identify potential differences in spatial distributions observable in the PCA scores. Non-target locations included six samples in and around the landfill and six samples from the remote JFDB site beyond the influence of the landfill. Principal components analysis fingerprints were estimated by examining the scores plot of all VOCs and comparing the trend of known BVOCs from *L. tridentata* and other Mojave Desert vegetation (Jardine et al., 2010; Geron et al., 2006) with that of the most abundant landfill contaminants, including CFCs and chlorinated solvents (Baker et al., 2012). The fingerprints were then used to characterize the sources of other VOCs from the non-target and target databases for further analysis of the types and abundances of anthropogenic and biogenic VOCs. Correlations among the sample locations’ log-transformed and centered data are shown in the Supplemental Material (Supplemental Fig. S2), along with additional details of the data processing and PCA procedures.

**Results and Discussion**

**Horizontal, Vertical, and Temporal Variations of Target VOCs**

Plots of the shallow, horizontal distribution of VOCs around the ADRS show spatial patterns indicative of various types and locations of sources, including potential biogenic origins. Figure 3 includes contours of the average VOC concentrations based on all available years during the period of the study at the 1.5-m depth (steel sample tubes) for CFC-12, tetrachloroethene (PCE), CCl₄, and *m,p*-xylene. The hotspots of contours for CFC-12 and PCE indicate source areas in the waste trenches. Patterns indicate likely sources of CFC-12 in the western waste trenches and of PCE in the eastern and western waste trenches. In contrast, CCl₄ concentrations are greatest near the extreme eastern sample points, outside of the LLRW facility and nearest to the hazardous waste area (Fig. 1). Concentrations in that area are one to two orders of magnitude greater than other concentrations near and in the LLRW area. These results confirm previous postulation of multiple landfill source locations (with CCl₄ having a source to the east) based on VOC samples from outside the LLRW facility (Baker et al., 2012) and CO₂ and CH₄ concentration profiles (Conaway et al., 2018). The contours of *m,p*-xylene show a dispersed, patchy pattern with peaks separated from any known anthropogenic sources. Atmospheric concentrations of *m,p*-xylene were measured as 0.25 ppbv in the plant canopy (average of two samples in 2001) and <0.07 ppbv above the canopy (one sample in 2008), too low to account for the observed hotspot concentrations shown in Fig. 3D. The greater *m,p*-xylene concentrations in soil than in air and the absence of a coherent source location indicate possible biogenic sources in the soil. A biogenic source is also indicated by the concentrations at the shallow depth of the remote (3 km south) JFDB location, which were as high as 1.5 ppbv, similar to those in many shallow probes near the LLRW area. The tendency for greater concentrations outside of the LLRW area than inside may relate to the disturbed soils and relative lack of vegetation on the LLRW landfill cover, as *L. tridentata* is a source of xlenes (Jardine et al., 2010).

Vertical and temporal variations in VOCs around the ADRS also indicate distinct sources, including possible biogenic origins, and shed light on the processes of contaminant plume expansion. Figure 4 shows vertical profiles of concentrations at all depths at UZB3 during the entire period of record, with depth on the y-axis and date on the x-axis. High concentrations of CFCs are commonly associated with LLRW (Baker et al., 2012), and Fig. 4A shows peak concentrations of CFC-12 at depths of 20 to 50 m emanating from the mixed-waste trenches in the LLRW area, several orders of magnitude greater than peak atmospheric concentrations (Busenberg
and Plummer 1992). Studies of elemental Hg (Walvoord et al., 2008), tritium (Mayers et al., 2005), and VOCs (Baker et al., 2012) all identified this depth interval as a zone of peak concentrations of landfill contaminants. Waste emplacement was shallower than 15 m bgs, so the greater depths of the peak concentrations may indicate downward transport of nonaqueous liquid or dissolved CFCs below the trenches or sinking of denser-than-air CFCs during transport. The color profile patterns and contours in Fig. 4A also show a downward shift of the peak of concentrations with time. Tetrachloroethene (PCE) maximum concentrations occurred close to the 23.8-m-depth sampling port, with increasing concentrations with time in deeper sampling ports. Carbon tetrachloride concentrations before 2010 show indistinct peaks and no apparent trend of concentrations with time at greater depths. After 2010, however, CCl₄ concentrations increase markedly at 23.8 m, possibly related to late arrival from a separate source area in the hazardous waste facility, as discussed above. The concentrations of m,p-xylene occur in a spottier pattern throughout the profile, with peaks at the surface and at depths below 23.8 m. The relatively uniform profiles of m,p-xylene compared with the known landfill contaminants (Mayers et al., 2005; Walvoord et al., 2008; Baker et al., 2012) indicate a lack of association with the landfill sources, as observed also in the horizontal distributions. The biogenic and anthropogenic origins and their relations to spatial variations are explored in greater detail for all sampled VOCs below.

To better track the potential vertical movements of LLRW landfill VOCs (Fig. 4A and 4B), we calculated the centers of mass along the UZB3 vertical profile. For each VOC that showed a persistent distinct peak at 23.8 m bgs for all years with >10 samples...
across the depth of the UZB3 profile (2000, 2002, 2004, 2006, 2008–2011), the center of mass, $\tau_j$, was estimated based on

$$
\tau_j = \frac{\sum_{i=1}^{N} \Delta z_i C_{i,j} z_i}{\sum_{i=1}^{N} \Delta z_i C_{i,j}}
$$

where $i$ is a depth interval from the midpoint above a sampling depth to the midpoint below, $\Delta z_i$ is the length of the interval, $C_{i,j}$ is the concentration (applied uniformly) across the $i$th interval of the $j$th VOC, and $z_i$ is the depth of the sample point in the interval.

For VOCs that showed a distinct peak at 23.8 m, the $\tau_j$ values were plotted against time, and linear regression was used to estimate the slope of $\tau_j$ vs. time, $\partial \tau_j / \partial t$, which indicates the vertical velocity of the center of mass.

A scatterplot of VOC density vs. the vertical velocity of the center of mass, $\partial \tau_j / \partial t$, (Fig. 5) indicates that (i) there is downward movement of most VOCs emanating from the LLRW facility and (ii) the vertical velocities correlate with the density of the pure-phase gas ($r^2 = 0.38$, $p = 0.024$), indicating possible
The first three PCs (eigenvalues of 6.9, 2.2, and 0.8) were retained for further analysis based on multiple criteria, including the Kaiser criterion and the cumulative variance explained. Based on the loadings of the three PCs, the variance of the non-target dataset was primarily determined by ambient sources in distal and intermediate locations around the waste facility (PC1), landfill sources in the waste facility (PC2), and local occurrences at X488Y480 (PC3). The loadings of the 12 sample locations on these three PCs are shown in Fig. 6A. The vectors indicate the loading values for PC1 and PC2, and the colors of the vectors indicate the value of PC3. The vector labels indicate the year (2014 or 2015), location, and depth (YEAR_LOCATION_DEPTH [m bgs]). The loadings are the coefficients that convert original values (non-standardized, log-transformed VOC concentrations) to the transformed, linearly uncorrelated variables (PCA scores), and so the loadings for each sample indicate its influence on the data variability. Loadings are similar for different dates and years at JFDB, indicating relative uniformity of VOCs with time and depth at that undisturbed, background location. For PC1, all loadings are positive, with the lowest values (least influence) at UZB3_23.8 and X488Y480_1.5 locations, which are associated with emissions from the LLRW area (Conaway et al., 2018). In other words, high scores for PC1 indicate elevated concentrations in all samples, especially at distal and intermediate locations beyond the direct influence of the LLRW sources. For PC2, the loadings decrease from near-source to distal locations, with values of 0.5 at UZB3_23.8, from 0.1 to 0.3 at intermediate locations, and −0.25 at the distal JFDB location, indicating that PC2 is associated with proximity to, and influence by, the landfill VOCs from the LLRW facility. For PC3, loadings were greatest at X488Y480 (1.5-m depth) and appeared to reflect unique aspects of the VOC mixtures at this shallow location inside the LLRW, possibly due to the influence of local buried-waste materials.

The PCA scores of the VOCs show two distinct trends associated with the loadings of landfill VOCs and ambient VOCs, which correspond to anthropogenic and biogenic origins, respectively. The scores of the non-target VOCs are shown in Fig. 6B and are listed in Supplemental Table S1. Figure 6B also shows the loadings labeled based on relative proximity to, and influence by, the LLRW sources. The VOCs occurring at low concentrations near the detection limit are clustered at the left side of the plot at PC1 ≈ −2, PC2 ≈ 0 (negative PC1 values result from the use of logarithmic concentrations). From the cluster of low-concentration VOCs, a band of points extends up and to the right (greater scores tending to indicate greater concentrations), oriented with the vectors of loadings for near-landfill and intermediate sample locations. Another band of points extends to the right along PC1, aligned with the vectors of loadings for distal and intermediate samples. Based on the PCA scores of known bVOCs from L. tridentata (Jardine et al., 2010) and other common plant species of the Mojave and Sonoran deserts (Geron et al., 2006), shown as labeled points in Fig. 6C, the lower band of ambient VOCs corresponds to biogenic origins (Fig. 6B). The envelope of known bVOCs includes VOCs not previously identified by Jardine et al. (2010) or Geron et al. (2006) as bVOCs, as expected due to differences in analytical techniques, sampling locations (subsurface vs. canopy), or biological sources not present in the previous studies.

**Spatial Fingerprinting of VOC Sources**

Principal components analysis of the 12 non-target VOC samples found spatial variations indicative of distinct sources. The PCA of the log-transformed and centered data identified three PCs that explained 83% of the variance. The individual contributions (Supplemental Fig. S3) of the PCs decrease from 58 to 18 to 7% and then to <5% for the remaining PCs. The cumulative variance explained shows breaks in slope (or "elbows") at PC2 and PC3. The first three PCs (eigenvalues of 6.9, 2.2, and 0.8) were retained for further analysis based on multiple criteria, including the Kaiser criterion and the cumulative variance explained. Buoyancy effects during gas transport. All but one of the VOCs (diethyl ether) exhibited downward movement (positive changes in depth vs. time) of the center of mass from 2000 to 2011, with tetrachloroethene having the maximum rate of 2.2 m yr⁻¹. The best-fit line passes close to the point for no migration of the center of mass when the relative density = 1, which is the theoretical intercept for a neutrally buoyant gas without upper or lower boundary effects. The lack of vertical migration for neutrally buoyant VOCs indicates that the downward movement of the gaseous VOC plume probably does not stem from vertical migration of liquids under the waste trenches. Because the total VOC concentrations at UZB3 are too low (<100 ppm by volume) to create substantial macroscopic density gradients, the downward transport probably occurs closer to the source locations in pore-phase or high-concentration zones. Density-driven advection of high-vapor-concentration gas has been observed for TCE (Conant et al., 1996) and CFC-113 (Christophersen et al., 2005) on time scales of days within a few meters of the source in shallow unsaturated zones. The results from ADRS show that such effects can persist at distances of >100 m and after decades of transport.
The association of bVOCs with PC1 indicates a tendency for lower concentrations at UZB3_23.8 and X488Y480_1.5 because of the smaller PC1 loadings of those samples. Similar results were reported above, with relatively low \( m,p \)-xylene concentrations inside the LLRW facility (where X488Y480_1.5 is located) and at the UZB3_23.8 location (Fig. 3D). Lower bVOC concentrations at these locations may reflect the relative scarcity of plants on the LLRW landfill cover, greater microbial respiration of bVOCs near these locations as indicated by elevated CO\(_2\) and relatively low \( ^{13}\)C-CO\(_2\) (Conaway et al., 2018), or displacement of bVOCs by increased gas advection in the LLRW facility. The VOCs plotting along the trend extending up and to the right consist of CFCs, chlorinated ethenes and ethanes, and cyclohexane. These VOCs were detected at elevated concentrations in samples from UZB3_23.8 and other near-source locations associated with anthropogenic VOCs and are generally absent from the remote JFDB location.

The PCA transformation from the non-target analyses was applied to the target VOC concentrations from the same samples to characterize sources of the target VOCs. The scores of the target VOCs are shown in Fig. 7 (color points), overlain on the scores of the non-target VOCs (gray points). Values of PCA scores for target VOCs are listed in Supplemental Table S2. The target VOCs distribute along the same two primary trends as the non-targets, with ambient VOCs extending to the right along PC1 and landfill VOCs extending up and to the right. Target PC1 scores tended to be greater than those from non-target analyses (median difference of 3.1) due in part to differences in nondetects between target and non-target analyses, but PC2 scores were similar (median difference of −0.4) and the variations in VOCs along the PC2 axis remained adequate to separate landfill VOCs from ambient VOCs. The labeled VOCs in Fig. 7 are those with PC2 > 0.9 and include chlorinated ethenes and ethanes and CFCs typically associated with LLRW source area contaminants and observed at elevated concentrations in UZB3_23.8. Points between the primary trends for landfill VOCs and ambient VOCs may reflect multiple sources for some VOCs. Low concentration detections of carbon tetrachloride and CFCs occurred at JFBD, probably due to global or regional atmospheric transport. Carbon tetrachloride, CFC-12, and CFC-113 occur in the atmosphere worldwide (Blake, 2013), and all three were detected in the 2001 and 2008 ADRS atmospheric samples at concentrations comparable to those at the remote JFDB unsaturated zone site. Chloroform occurs in the landfill VOC plume and is also known to be produced by numerous natural sources, including soil fungi (Hoekstra et al., 1998). In the 2001 plant canopy air samples, detections of chloroform (0.34 ppbv average of two samples) and tetrachloroethylene (0.65 ppbv average) were more than an order of magnitude greater than global atmospheric concentrations (Blake, 2013) and are comparable to concentrations at the distal sample locations, indicating possible regional sources.

**Abundance and Spatial Trends of Biogenic VOCs in the Unsaturated Zone**

A comparison of concentrations in the subsurface, atmosphere, and plant emissions indicates that the unsaturated zone is probably...
a substantial source of bVOCs at the ADRS. At the remote JFDB site, 3 km south of the waste facility, concentrations indicate upward diffusive gradients; unsaturated zone concentrations of ambient VOCs (PC1 > 0 and PC2 > 0) frequently exceeded plant-canopy measurements in the current study and were comparable to or greater than previous above-canopy observations of bVOCs by Jardine et al. (2010). Figure 8 shows ADRS unsaturated-zone target and non-target ambient VOC concentrations (thin bars, 5th to 95th percentiles) along with ADRS plant-canopy concentrations (diagonal striped bars, maximum and minimum), and noontime ambient above-canopy (4 m above ground) concentrations from Jardine et al. (2010) (empty bars, maximum and minimum). The range of unsaturated-zone 5th to 95th percentile of target VOCs exceeded the range of ADRS plant-canopy values in seven cases (tetrahydrofuran, benzene, 2-ethyltoluene, and three tri- and one tetramethylbenzene), were greater than or in the range of ADRS canopy values in 12 cases, and bounded the ADRS canopy values in five cases (acetone, 2-butanol [methyl ethyl ketone], toluene, 2-hexanone [MBK], and chlorobenzene). Amargosa Desert Research Site plant-canopy observations are not available for the non-target VOCs. Among the ADRS target and non-target unsaturated zone ambient VOCs, 13 were also quantified by Jardine et al. (2010) in above-canopy air—three ADRS unsaturated-zone concentrations were greater than the canopy values (benzene, 

Fig. 7. Scores for target volatile organic compounds (VOCs) (colored points) and non-target VOCs (small gray points). Circled gray points are biogenic VOCs previously reported as biogenic above-ground emissions from US Southwest desert vegetation by Geron et al. (2006) and Jardine et al. (2010). The VOC names are given for Principal Component 2 > 0.9, which are primarily associated with the waste disposal facility sources. Number labels in parentheses after VOC names and by points correspond to the ID number for target VOCs in Supplemental Table S2.

Fig. 8. Concentration ranges of ambient volatile organic compounds (VOCs) in the unsaturated zone (0.5-, 1.5-, and 28.7-m depths) at the remote JFDB site with Principal Component (PC)1 > 0 and PC2 < 0 from target and non-target analyses. Also shown are ambient plant-canopy (Larrea tridentata) air concentrations measured for Amargosa Desert Research Site (ADRS) target VOCs (two samples in 2001: 25 and 150 m from the low-level radioactive waste area) and above-canopy (4 m above ground) values reported by Jardine et al. (2010) for target and non-target VOCs. Bars for JFDB unsaturated-zone samples show the span of the 5th to 95th percentiles based on regression on order statistics (see Supplemental Material) for 17 samples from 2000 to 2016. Bars for ADRS canopy concentration show the minimum and maximum, with bars extended down to 0.001 ppbv if one or more canopy sample was a nondetect (i.e., isopropylbenzene and naphthalene) and up to the detection limit if both canopy samples were nondetects (i.e., tert butyl alcohol, tetrahydrofuran, benzene). Bars for Jardine et al. (2010) above-canopy air span the minimum and maximum reported values.
phenol, and acetic acid) and the remaining 10 unsaturated-zone concentrations bounded the above-canopy values. The most abundant unsaturated-zone ambient VOC at the remote JFDB site was 2-ethyl-1-hexanol, which was identified as a bVOC in southwestern US desert habitats (Geron et al., 2006) but was not quantified by Jardine et al. (2010).

Deep concentration profiles at ADRS are consistent with long-term, shallow-subsurface bVOC production with equilibrium transport between the soil and the water table. Profiles were plotted for the deep boreholes UZB2 and UZB3 at 12 and 13 sample depths, respectively. Figures 9 and 10 show vertical profiles for all target ambient VOCs (n = 24) as well as a representative landfill VOC (CFC-12, lower right). The lower-bound concentration profiles (red lines on Fig. 9 and 10) account for potential effects of tube emissions on the measured VOC concentrations (black lines) and typically show uniform (though noisy) concentrations between the surface and groundwater. The uncertainty related to potential tube emissions (horizontal distance between

Fig. 9. Profiles of average concentrations of the ambient volatile organic compounds (VOCs) at UZB2, 160 m from the nearest waste trench. Black lines show measured data and red lines show purge-adjusted values. The pink horizontal line shows the typical depth of peaks for VOCs originating from the waste trenches, and the blue horizontal line shows the approximate depth of the groundwater. Landfill-source dichlorodifluoromethane (CFC-12) is included at bottom right for comparison.
red and black lines) is typically minor between 0 and 50 m and increases at depths >50 m. The increased uncertainty of concentrations at depths below 50 m points to the need for future work to deploy inert sampling tubes into this zone. The persistence of most VOCs across the full vertical profile is consistent with continuous production of bVOCs in the root zone and conservative diffusive transport for a sufficiently long time to propagate the source concentrations down to the water table. Peaks at depth may relate to variations of geology, production rates, or reactivity. For example, the majority of ambient VOCs exhibit peaks at approximately 100 m in both UZB2 and UZB3, which may relate to georespiration, consistent with elevated CO₂ near the water table (Walvoord et al., 2005; Conaway et al., 2018; Keller and Bacon, 1998), and potential secondary effects of equilibrium gravitational settling (Seltzer et al., 2017). Greater concentrations at UZB2 (160 m from the nearest waste trench) than at UZB3 (100 m from the

Fig. 10. Profiles of average concentrations of the ambient volatile organic compounds (VOCs) at UZB3, 100 m from the nearest waste trench. Black lines show measured data and red lines show purge-adjusted values. The pink horizontal line shows the typical depth of peaks for VOCs originating from the waste trenches, and the blue horizontal line shows the approximate depth of the groundwater. Landfill-source dichlorodifluoromethane (CFC-12) is included at bottom right for comparison.
nearest waste trench) indicate a negative association of ambient VOC sources with the waste facility, consistent with results of the PCA analysis and the distributed shallow-sample observations for m,p-xylene (Fig. 3D). Few ambient VOCs showed peak concentrations in the zones of landfill VOC contamination, with the notable exceptions of benzene and chlorobenzene at UZB3, indicating that these contaminants commingle from ambient and landfill sources.

At JFDB, depth trends are more difficult to ascertain because of the sparse vertical resolution, with two steel tubes at 0.5 and 1.5 m and a single plastic sampling tube at 28.7 m. The limited number of samples and variability of purge volumes from the 28.7-m depth were insufficient to conduct the purge-volume analysis, but we screened for potential purge effects by checking for vertical gradients in the ambient VOC concentrations between the shallow steel tubes at 0.5 and 1.5 m and the deeper sample point with plastic tubing. Depth-correlation p values were estimated for the slope of the linear least-squares fit for each ambient VOC vs. depth. Vertical gradients of concentrations were not significant (P > 0.05) except for increasing downward concentrations for 2-butanol and tetrahydrofuran and upward increasing concentrations for 2-ethyl-1-hexanol. The elevated concentrations of 2-butanol and tetrahydrofuran at depth, together with the results from UZB2 and UZB3, suggest potential tube-emission effects in samples from JFDB at 28.7 m.

The presence of shallow subsurface sources is further supported by correlation of the shallow and above-surface target ambient VOCs across the study site. The geometric-mean ambient VOC concentrations among all shallow (1.5 m bgs) sites were correlated with, and generally greater than, the concentrations measured in the plant canopy (Fig. 11). Because gas diffusion is relatively rapid from high- to low-concentration zones (Altevogt et al., 2003), elevated concentrations in the shallow subsurface probably indicate nearby sources such as roots and soil microbes (Peñuelas et al., 2014). The correlation between subsurface and above-surface concentrations may result from transport of the VOCs from the shallow unsaturated zone to the atmosphere due to diffusion and barometric pumping. As target analytes, these VOCs are commonly attributed to contaminant sources, but all have potential biogenic origins (Supplemental Table S3). For ambient VOCs with substantial commingling sources in the waste trenches (e.g., benzene and chlorobenzene, Fig. 10), subsurface inputs from the landfill may also elevate the average 1.5-m depth concentrations.

The vertical gradients of ambient VOCs shed new light on potential uses of these VOCs as indicators of biological sources, reactivity, and temporal changes. Subsurface biological activity is typically greatest in soils but may be elevated at depth if moisture and carbon are available (Tokunaga et al., 2016). Reactions of bVOCs may occur due to biological or abiotic activity, with daughter products tending to occur at greater distance from the source. The absence of such trends at ADRS indicates that reactions may be suppressed due to low microbial activity in deep, dry sediments and the lack of photooxidation. Because thick unsaturated zone profiles can equilibrate over centuries and millennia (Walvoord et al., 2004) and because vertical gas transport may be impeded by horizontal layering (Green et al., 2015), arid unsaturated zones potentially preserve chemical records of changes in bVOC emissions in response to climatic stresses or long-term ecological transitions, such as the millennial shift in the Mojave Desert from Juniperus steppe woodlands to L. tridentata dominated habitat (Spaulding, 1990; Koehler et al., 2005). With additional sampling and analysis of unsaturated zone profiles, future studies may better characterize the processes that determine the profiles of bVOCs in the unsaturated zone as well as the implications for environmental processes regulated by bVOCs.

**Implications for Contaminated-Site Characterizations**

The PCA fingerprinting indicated that non-target analyses directly complement target analyses in characterizing the number and quantity of landfill VOCs. The target analyses included 11 landfill VOCs (PC1 > 0, PC2 > 0.9), compared with four in the non-target analyses, including chlorotrifluoroethylene, cyclohexene, dichlorofluoromethane, and 1,2-dichloro-1,1,2-trifluoroethane. These non-target landfill VOCs were detected at concentrations ranging from 461 to 1075 ppbv at UZB3, 23.8-m depth, which is on the order of concentrations for the most abundant contaminants at the ADRS. In addition, there were six...
non-target VOCs tentatively identified as halogenated VOCs, including likely perfluorinated two- to eight-carbon compounds, that were not present at other locations and were not included in the PCA fingerprinting. The peak sizes of these VOCs ranged between those of CCl₄ and dichlorodifluoromethane, indicating concentrations on the order of 100 to 3000 ppbv. The reservoir and fluxes of landfill VOCs in the unsaturated zone are substantially greater than previously characterized based on the target analysis alone. The use of non-target analyses can improve the accuracy of estimates of total VOC mass emissions from landfills, as well as the variety of VOC contaminants present, compared with target analyses alone with limited numbers of VOCs (Robertson and Dunbar, 2005). Toxicological profiles of non-target VOCs are often less complete than those of the target VOCs, making it difficult to evaluate the implications of non-target VOCs for human and environmental health. However, non-target VOCs may reveal the presence of contaminants of emerging concern, such as perfluorooctane sulfonate (PFOS) and perfluorooctanoic acid (PFOA), as well as establishing a basis to screen for other, as yet unknown, VOC risk factors. Identifying the presence of contaminants of emerging concern at sites such as the ADRS opens the door to toxicological studies of long-term exposures near the sources.

The commingling of VOCs from multiple sources affects the delineation of unsaturated zone plume boundaries and necessitates careful discrimination of sources. At the ADRS, for example, the non-target analyses identified 152 VOCs compared with 60 from the target analyses, with numerous detections at the remote sampling location. The PCA fingerprinting indicates that most of the abundant VOCs (PC1 > 0) in the unsaturated zone are from ambient sources (74 ambient VOCs compared with 15 landfill VOCs) and are mainly biogenic, including many, such as the aromatic compounds, that are commonly thought to be indicative of anthropogenic sources. For VOCs with both anthropogenic and biogenic sources, such as benzene and chlorobenzene at ADRS, exact delineation of plume boundaries and characterization of transport may be difficult based on measured concentrations alone. Additional techniques, such as compound-specific carbon isotope analysis (Zwank et al., 2003), may be needed to identify plume boundaries and to estimate time scales of transport.

**Summary and Conclusions**

A combination of non-target analysis and principal components analysis based fingerprinting of landfill VOCs and ambient VOCs identified key aspects of anthropogenic and biogenic VOC occurrence in the unsaturated zone around a LLRW and hazardous waste facility. The non-target analysis provided concentrations of 152 VOCs compared with 60 from the target analyses. Principal components analysis was used to characterize the spatial distribution and potential sources of the non-target compounds. The PCA fingerprinting revealed 74 ambient VOCs that occur throughout the depth of the arid unsaturated zone and commingle with 15 landfill VOCs from the waste facility. Comparison of ambient VOCs with known bVOCs for *L. tridentata* and other Mojave Desert plants indicated biogenic emissions as a likely source of the ambient VOCs and was consistent with observations of vertical, horizontal, and temporal variations in the target-VOC dataset. This provides a proof-of-concept for discriminating sources using PCA fingerprinting with known bVOCs and a relatively small number (e.g., 12) of non-target samples.

The results showed, for the first time, that the deep arid unsaturated zone is a reservoir of bVOCs that are relatively uniform in concentration between the surface and the water table. Concentrations of bVOCs were correlated with, but often greater than, concentrations measured in the plant canopy, indicating subsurface sources of bVOCs. The bVOCs tended to be less strongly associated with locations near the LLRW source trenches, possibly due to soil disturbance and a relative lack of plants on the LLRW landfill cover, greater microbial respiration of VOCs in undisturbed areas outside of the facility, and/or displacement of bVOCs by gases emitting from the waste trenches. While peak concentrations of the bVOCs are lower than those of landfill VOCs near the waste facility, the overall inventory of unsaturated zone bVOCs probably occurs across an extensive region—e.g., the 750,000 km² of the Mojave, Sonoran, and Chihuahuan deserts—and may constitute a substantial potential reservoir of VOCs for consideration in regional inventories, e.g., 10⁶ kg of phenol at 10 ppbv distributed across 100 m of unsaturated zone with a porosity of 0.3. The bVOCs are important to air quality, atmospheric chemistry, and the functioning of biological systems; future studies will be needed to address the extent to which the substantial reservoir of bVOCs in the unsaturated zone influences these global, regional, and local issues.

The discovery of deep unsaturated zone bVOCs also opens the door to their use as tracers of gas transport. Methods of analysis of natural or modified unsaturated zone profiles have been established through numerous studies using, for example, water vapor (Walvoord et al., 2004), CO₂ (Conaway et al., 2018), atmospheric tracers (Busenberg and Plummer, 1992), and numerical models of disposed or injected gases (Stauffer et al., 2005; Green et al., 2015). The same methods can apply to profiles of bVOCs, with the added potential to link time scales of transport (e.g., between the ground surface and water table) to the ecological communities and conditions associated with particular bVOCs or assemblages of bVOCs (Peñuelas et al., 2014). In other words, future studies may better characterize not only the types and variety of deep unsaturated zone bVOCs and the transport processes that emplaced or displaced them but also the historical, environmental conditions that led to their production.

Our results indicate that the use of PCA fingerprinting to discriminate sources can improve contaminant characterizations. The commingling of anthropogenic and biogenic VOCs in the deep unsaturated zone has not been recognized in existing studies, to our knowledge, but can affect plume delineation. Many of the target analytes that were frequently detected at low or moderate concentrations, such as *m,p*-xylene, appear to originate primarily...
from biogenic sources, while others, such as benzene and chloroform, appear to originate from both biogenic and anthropogenic sources—drawing into question the accuracy of characterizing the zone of influence of a site based solely on concentrations. Resolution of this problem may be aided by using the PCA fingerprinting analysis developed in this study as well as complementary methods such as isotope analysis.

Observations of anthropogenic VOC transport at the ADRS provide a rare view of VOC transport on the scale of hundreds of meters and decades of time in a thick, sedimentary unsaturated zone. The current study showed that downward migration of landfill VOCs was greater for denser VOCs, indicating probable density effects during transport in pure-phase or high-concentration zones near the sources. Similar effects are likely to occur at other landfills where the waste stream includes concentrated VOCs. At the ADRS, the use of non-target analysis identified landfill VOCs that were similar in number and total mass to those identified in the target analyses, as expected (but not widely characterized) at most landfills as a result of the diversity of VOCs in waste streams. Estimates of fluxes and abundance based on target analysis alone will therefore tend to represent lower limits of the mass of VOCs around landfills.

Acknowledgments

This work has support from the USGS Toxic Substances Hydrology Program and the USGS Water Mission Area. Field support was provided by Jonathan Arthur, Michael Barrenchea, Amanda Garcia, Mike Johnson, Ron Kauble, Stephen Maples, Justin Mayers, Michelle Walvoord, and other members of the USGS ADRS team. Michelle Walvoord and James Thordsen provided supporting data. We gratefully acknowledge the constructive review comments of Phil Stauffer, Fred Tillman, and four anonymous reviewers. We thank the State of Nevada, the Bureau of Land Management, and US Ecology, Inc., for their cooperation in long-term research at the site. Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the US Government. All data generated or analyzed during this study are included in the main text of this publication, associated supplemental material, and as a data release (Green et al., 2019).