# Integrating Geophysics and Soil Sampling for Site Characterization: A Kernel Approach

Gabriel Fabien-Ouellet,[1] Erwan Gloaguen[1] and Gaël Plassart[2]
[1]INRS-ETE, Quebec City, Canada
Email: Gabriel.Fabien-Ouellet@ete.inrs.ca
[2]Envisol, France

## ABSTRACT

Reconstructing the variation of contaminant concentration with a limited number of soil samples is more or less the norm, even though it fails more often than not for problems of even moderate complexity. To overcome the limits inherent to discrete measurements, we propose to integrate soil sampling with continuous surface geophysical measurements in a geostatistical framework. We present this integrated analysis for a PAH contaminated site in France. For the study site, two 3D surveys were acquired: an electrical resistivity tomography survey and a seismic travel time tomography survey. Those two surveys permitted us to infer two spatially continuous physical properties on the whole volume, namely the electrical resistivity and P-wave velocity. The probability density function relating the velocity-resistivity pairs with each of the 75 lab measurements of PAH concentration was modeled using a Gaussian kernel. This probability density function combined with the 3D volumes of resistivity and P-wave velocity provided a means to translate the latter into a 3D map of PAH concentration. This 3D map of concentration was then used as a secondary variable in a cokriging simulation of the 75 lab samples, thus reintroducing the spatial correlation of the initial dataset. Comparing this final 3D PAH concentration model with the simple kriging of the PAH samples, the geophysical integrated model reproduce much better the distribution of measured concentration, shows a much more realistic spatial pattern of the contamination, and lowers the estimated contaminated volume.

## Introduction

Geophysical data are useful to bridge the gap between discrete measurements and a continuous 3D representation of the subsurface (Ruggeri *et al*., 2013). Indeed, geophysical surveys bring information on spatial connectivity that cannot be inferred from discrete measurements. Applied to soil characterization, geophysical properties can be affected directly by the presence of contaminants (Ajo-Franklin *et al*., 2006; Geller and Myer, 1995; Rucker *et al*., 2009), or indirectly by preferential migration pathways (Coscia *et al*., 2012). In this work, we show the integration of a 3D geophysical survey with direct soil sampling to obtain a more precise estimation of contaminated soil volumes in a complex, heterogeneous geology.

## Study Site

The study site is located in an industrial environment. The main contaminating activity is the creosote treatment of wood, which began in the 1950s. Creosote contamination of soil and groundwater is of concern in the area, as local inhabitants rely on groundwater for drinking water. This contamination takes mainly the form of polycyclic aromatic hydrocarbons (PAH). Our study focuses on a small area of the property, measuring 30 m by 200 m. In this area, two sources of PAH contamination have been identified: a tailings pond and a sump, which were used to treat wastewater on site. The goal of this work is to characterize the extent of those sources, particularly in the unsaturated sediments.

The natural sediments in place are comprised of a relatively homogeneous coarse alluvial sands deposit. However, due to the industrial nature of the site, backfill materials of different sources and composition have been put in place at different periods. Hence, the first several meters of soils are heterogeneous. For this reason, mapping the extent of the contaminated zone through direct soil sampling is challenging.

 DOI: 10.2113/JEEG22.3.305

## Field Measurements

We base our characterization effort on a recent soil and water sampling campaign. The following analysis is based on the total PAH concentrations in soil measured on 75 of those samples. Many measurements exceeded 70 mg/kg, the limit after which contaminants may migrate and contaminate water. The highest measured concentration reached 7,000 mg/kg. Hence, soil contamination is known to be present, but contaminated volumes are difficult to predict due to the backfill heterogeneity.

Soil sampling was complemented by surface geophysical measurements. Two methods were chosen on site because of their ability to infer physical properties correlated with the lithology: electrical resistivity tomography (ERT) and seismic refraction tomography. In total, nine parallel ERT lines were acquired along the east-west direction and 10 parallel lines were acquired along the northwest-southeast direction. Each line consisted of 48 electrodes with a spacing of 1.5 m, in a dipole-dipole configuration. The resulting electrical resistivity model is shown in Fig. 1(a). The inverted lines were interpolated with an inverse distance weighting algorithm to obtain a 3D continuous model. Resistivity values vary strongly laterally and vertically. Surface backfill materials are usually resistive, but can vary between 100 and 1,000 ohm-m. This high level of soil heterogeneity is caused by different backfills that were put in place at different times.

A total of five parallel seismic lines were acquired in the east-west direction. Geophones were planted at a 2 m interval and sources were positioned at every 6 m. The seismic lines were also inverted in 2D and interpolated with an inverse distance weighting to obtain a 3D continuous model. The quality of the inversion is very good with a root mean square error between 0.7 and 1.8 % for each line. The resulting velocity 3D model is shown in Fig. 1(b). Lateral variations are much more subtle than that shown for ERT. However, three layers are visible, with the first layer comprising a loose dry soil layer with a velocity around 300 m/s. The next layer is a saturated sediment with a velocity near 1,400 m/s. The lower layer represents the bedrock, around 2,500 m/s. Small velocity variations are present within each layer.

## Data Integration

The difficulty of using direct correlations between geophysical measurements and soil contaminant concentrations is well illustrated in this study. Insets within Figs. 1(a)–(b) show crossplots between the measured total PAH concentration in soil samples and electrical resistivity and P-wave velocity, respectively. A very poor correlation coefficient of -0.28 is obtained for ERT, and a better correlation of 0.6 for seismic velocities. This is to be expected as organic contaminants are far from being the only factor influencing the electrical resistivity or P-wave velocity of soils. On their own, very little quantitative information can be obtained from these correlations.

Based on the observation that other factors may influence either electrical resistivity or P-wave velocity, we seek to build a multi-parameter relationship that suppresses common variance due to unknown factors and amplify the influence of PAH concentrations. To do so, we use the kernel density estimation approach of (Silverman, 1981), with which we build a density function taking as input the electrical resistivity and the P-wave velocity and outputs the PAH concentration. The density function is obtained by interpolating known sample triplets with a Gaussian kernel. The results are shown in Fig. 1(c). In the resistivity-velocity plane, PAH samples are gathered in families of high and low levels of concentration. This shows that contaminated zones take specific pairs of P-wave and electrical resistivity. Application of the methodology allows a recovery of the mean of the PAH conditional distribution given any value of measured resistivity and P-wave velocity. These families cannot be obtained from one-to-one linear correlations alone.

Using this density function, we can convert the resistivity and velocity 3D models into a 3D PAH pseudo-concentration map, shown in Fig. 1(d). The model does not reproduce exactly the measured values of PAH concentration at the samples location, as shown in the inset of Fig. 1(d). This is to be expected, as the kernel averages the PAH values around each velocity-resistivity pair. However, the correlation between measured and predicted concentration is very good, at 0.95. In order to honor exactly the measured PAH values, we used collocated cokriging with the PAH model built with the kernel as a secondary variable (Fig. 1(f)). As shown in the crossplot, the final cokriged model reproduces exactly the measured concentrations. Comparing this model with the result of simple kriging that does not include information from geophysical measurements (Fig. 1(e)), we see that simple kriging produces overestimated values of concentration. However, when all of the geophysical information was included, it allowed the geostatistical estimation of PAH to be performed locally in zones of similar resistivity-velocity values.

## Conclusion

With appropriate geophysical measurements, a multiparameter function with geophysical properties as

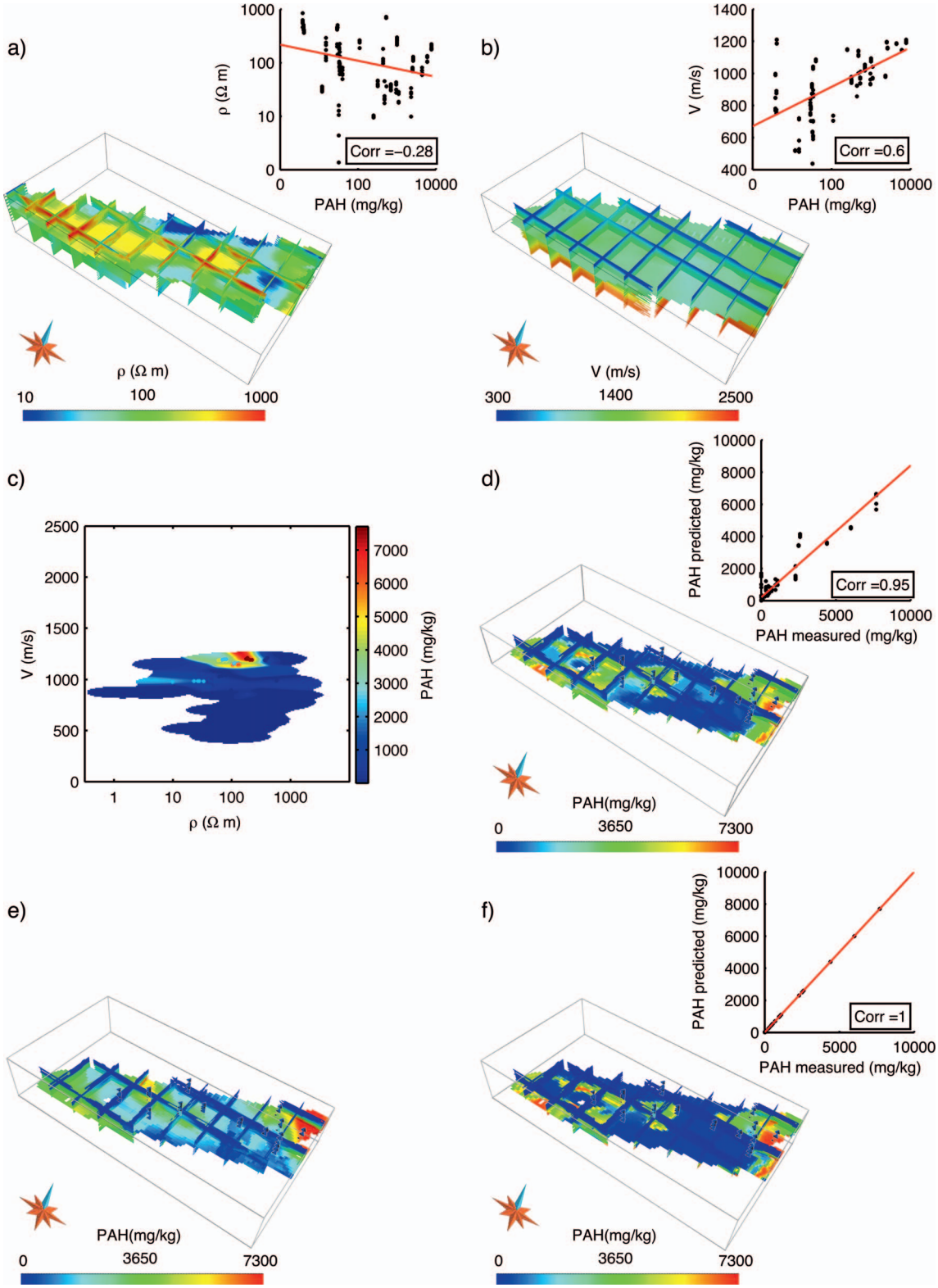


**Figure 1. a) 3D interpolated resistivity, b) 3D interpolated P-wave velocity, c) density function, d) PAH concentration given by the density function, e) PAH concentration obtained by simple kriging, and f) PAH concentration obtained by cokriging. The coloured triangles show the locations of PAH measurements. The insets show the correlation of either the measured value or the estimated PAH concentration from each respective method with laboratory measurements.**

input and concentration value as output can be built with a kernel estimation. Using this function, we built an intermediate concentration model that was used as a secondary variable for cokriging of the soil-sampled PAH values. The final PAH concentration model obtained with this methodology allowed to better delineate contaminated zones compared to the simple kriging of discrete PAH measurements. We do not claim that this workflow is applicable to all contamination types or all geological settings, but in some cases, as shown in this study, it can be used successfully to integrate geophysical information at low cost.

## References

Ajo-Franklin, J.B., J.T. Geller and J.M. Harris, 2006, A survey of the geophysical properties of chlorinated DNAPLs: Journal of Applied Geophysics, **59**, 177–189.

Coscia, I., Linde, N., Greenhalgh, S., Vogt, T., and Green, A., 2012, Estimating traveltimes and groundwater flow patterns using 3D time-lapse crosshole ERT imaging of electrical resistivity fluctuations induced by infiltrating river water: Geophysics, **77**, E239–E250.

Geller, J.T., and Myer, L.R., 1995, Ultrasonic imaging of organic liquid contaminants in unconsolidated porous media: Journal of Contaminant Hydrology, **19**, 85–104.

Rucker, D.F., Glaser, D.R., Osborne, T., and Maehl, W.C., 2009, Electrical resistivity characterization of a reclaimed gold mine to delineate acid rock drainage pathways: Mine Water and the Environment, **28**, 146–157.

Ruggeri, P., Irving, J., Gloaguen, E., and Holliger, K., 2013, Regional-scale integration of multiresolution hydrological and geophysical data using a two-step Bayesian sequential simulation approach: Geophysical Journal International, **194**, 289–303.

Silverman, B.W., 1981, Using kernel density estimates to investigate multimodality: Journal of the Royal Statistical Society. Series B (Methodological), 97–99.