

# STATISTICS INFORMS ENVIRONMENTAL CLEAN-UP

# STATISTICAL SIGNIFICANCE

Environmental contamination seems ubiquitous, with ever-increasing cleanup costs. Litigation stemming from claims of contamination increases the expenditures further. Statistics plays a critical role in the decisions about who should pay what portion of these costs. Working with soil scientists, geologists, chemists, ecologists, and toxicologists, statisticians conduct statistical analyses to inform decisions about who is responsible for contamination and the cost of cleanup.

## Contamination in the Environment

### FINGERPRINTING THE CULPRIT

PCBs (polychlorinated biphenyls) and dioxins/furans (polychlorinated dibenzodioxins and furans) represent a class of chemicals ideal for identifying the source of contamination because they consist of multiple chemicals (called congeners) that provide a chemical signature or “fingerprint” of a source. The chemical fingerprint indicates a source in the same manner an actual fingerprint provides evidence in a criminal investigation.

These longevity chemicals persist in the environment for decades after an initial source of contamination which provides the necessary chemical fingerprint for statistical analysis much later in time.



**PCBs** were a savior to electricians because they reduced the risk of fire around electrical work. Despite a ban on production in the 1970s, PCBs persist in the environment today.

**Dioxins/furans** arise from a variety of sources, including bleaching paper during manufacture, treating lumber, burning rubbish, and forest fires.

PCBs and dioxins/furans are attributable to cancer and reproductive and developmental problems.

### HELP FROM STATISTICIANS

Environmental investigators collect samples of soil, sediment, and/or other media, from a contaminated area for laboratory analyses to determine the concentrations of chemicals present.

Statisticians use these data in an analysis technique called principal component analysis (PCA) to evaluate the chemical compositions between samples potentially influenced by multiple sources of contamination. Results from this analysis indicate if the compositions of samples are similar to each other by grouping like samples closer together. The more similar the composition the closer the samples are to each other in the plots produced from the results.

Because of the variety of investigations and types of contamination, there is no standard for how this analysis is conducted. The input data, and any transformation of that data, can change the results, thereby altering the determination of responsibility and allocation of costs.

### COMPLEXITY OF ASSESSMENTS

When the chemical fingerprints, or compositions, of sources are distinct, mixing gradients can be easily identified and the contributions from each source quantified. The situation is more complex when sources become diluted or multiple sources have similar compositions.

Assessment of a site containing chemical contaminants generally includes evaluation of the level of contamination (how high are the concentrations) and consideration of the potential sources that may have contributed to the contamination. While significant efforts (involving statisticians) are continually applied to develop limits associated with specific hazard risks, analyses to distinguish potential sources are virtually unregulated. As a result, there is little guidance for how these statistical analyses can be conducted.

### DECISIONS AND COSTS

Decisions related to environmental cleanup, including apportionment and allocation of costs, frequently rely on the results of statistical analyses. These decisions can result in multi-million dollar expenditures, or more, by some of the largest companies in the U.S. and the world.

Correct and appropriate statistical analyses to address the question of interest are critical to the best, science-based decisions.

Melanie Edwards and Mihai Aldea

Exponent