Identification of Background Levels from Soil Constituent Data with a High Proportion of Non-Detects, at the USDOE Savannah River Site in South Carolina

by

Vikram M. Vyas*, Brian Hough**, Christopher Romanek**,
Panos G. Georgopoulos**, and David S. Kosson***

9 December, 2003

Presented at the 2003 Annual Meeting of the Society of Risk Analysis, Baltimore, MD

* EOHSI, a Joint Institute of UMDNJ – RW Johnson Medical School, and Rutgers University, Piscataway, NJ
** Savannah River Ecological Laboratory, Augusta, GA
*** Vanderbilt University, Nashville, TN

Role of Background Levels in Site Remediation

- Background levels are due to naturally occurring concentrations and also contamination from non-site related diffuse sources or activities
- Background levels are used to determine whether the threat to human health or environment posed by existing or potential release falls below the risk posed by naturally occurring substances in the impacted area, to set remediation goals, and to communicate cumulative risks associated with contaminated sites
- Background samples are those collected at or near the hazardous waste site in areas not influenced by the Superfund site contamination or other nearby Superfund sites
- References: USEPA, 2001, Guidance for Characterizing Background Chemicals in Soil at Superfund Sites, EPA 540-R-01-003, Office of Solid Waste and Emergency Response, Washington, D.C., June 2001

Issues Regarding Background Level Identification at SRS

Background soils data are used extensively in three processes in the environmental restoration program at Savannah River Site (SRS):

- Identification of unit specific constituents as a result of remedial investigation (RI) characterization activities;
- Reduction of Constituents of Potential Concern (COPCs) to constituents of concern (COC) in the baseline risk assessment (BRA) as a component of uncertainty analyses; and
- In establishing remedial goals and verifying if those remedial goals are met

Operational Issues in Identifying Background Levels

- High proportion of data below detection limit (non-detects) for certain COPCs
- Presence of multiple Method Detection Limits (MDLs) within data for each COPC
- Use of statistical methods to adjust distributions of COPCs with a high proportion of detects
- Relationship between background levels and health/ecological risk based screening criteria

Tasks Implemented for this Study

- Creation of SRS site-wide soils database for As, Cd, and Hg (to be extended to other COPCs)
  - Extraction and assimilation of data
  - Unit and spatial coordinate conversions
- Preprocessing of raw data
  - Screening of replicates
  - Soil type classification
  - Non-detect adjustment
  - Determination of “Top” interval
- Adjustment of distribution parameters for non-detects
- Computation of data percentiles

Creation of SRS Site-Wide Soils Database

- A “master” dataset was created combining the data found in the Looney (1990), Ross (1996), Dixon (1997), and Ross (1999) reports
- Unit conversions were necessary to express all analytical results as mg/kg in the new database
- Sample locations expressed in the SRS coordinate system were converted to Universal Transverse Mercator (UTM)

Preprocessing of Raw Data

- Screening of replicates
  - By using explicit replicate identification
  - Based on spatial location and depth of sample
- Non-detect adjustment
  - Using analysis qualifiers (“U”, “IU”, “JU”, “UJ” => non-detect) for Ross (96, 99) datasets
  - Using “LT” qualifier for Looney dataset
  - Using “<“ symbol for Dixon dataset
- Determination of “Top” interval
  - The “Top” interval is defined as the uppermost soil sample taken at a site, to avoid inconsistencies between different studies
  - Data from the top soil interval were used for further analyses

Number of Detects for Each COPC

COPC	Number of Records	Number of Detects
As	176	61
Cd	176	33
Hg	175	73

Survey of Methods for Statistical Treatment of Non-Detects

- Simple substitution methods will lead to unacceptable distortion of statistical distribution of data when high proportion of non-detects exist
  - Filtering on non-detects
  - Substituting non-detect values by MDL/2 or zero
- Maximum Likelihood based methods work best for normally or lognormally distributed data
- Helsel's Robust Method has been reported as the optimum method for cases where data are from highly skewed distributions
- References:

  - Helsel, D.R., 1990, Less than Obvious – Statistical treatment of data below the detection limit, Environmental Science and Technology, 24 (12), 1766-1774
  - Daniels, W.M., and N.A. Higgins, 2001, Environmental Distributions and the Practical Utilisation of Detection Limited Environment Measurement Data, NRPB, Oxon, UK. ISBN 0 85951 484 6
  - Newman, M.C, K.D. Greene, and P.M. Dixon, 1995, UNCENSOR Version 4.0 Users' Manual, Savannah River Ecological Laboratory, SC

Implementation of Helsel's Robust Method (HRM) for Multiple MDLs

- Log-transform all reported values
- Use the reported values to sort data and assign ranks. Data are sorted from lowest value to highest value. Ranks are assigned as 1 for the lowest value and N for the highest value, where N is the total number of data (detects as well as non-detects)
- Use the ranks to compute the normal scores for the data
- Do a linear regression between the reported detected values and their normal scores. The slope of the regression line will give the adjusted standard deviation for the data (detects as well as non-detects), and the Y-intercept will give the adjusted mean value of the data (detects as well as non-detects)
- The adjusted values for each non-detect are estimated from the regression equation

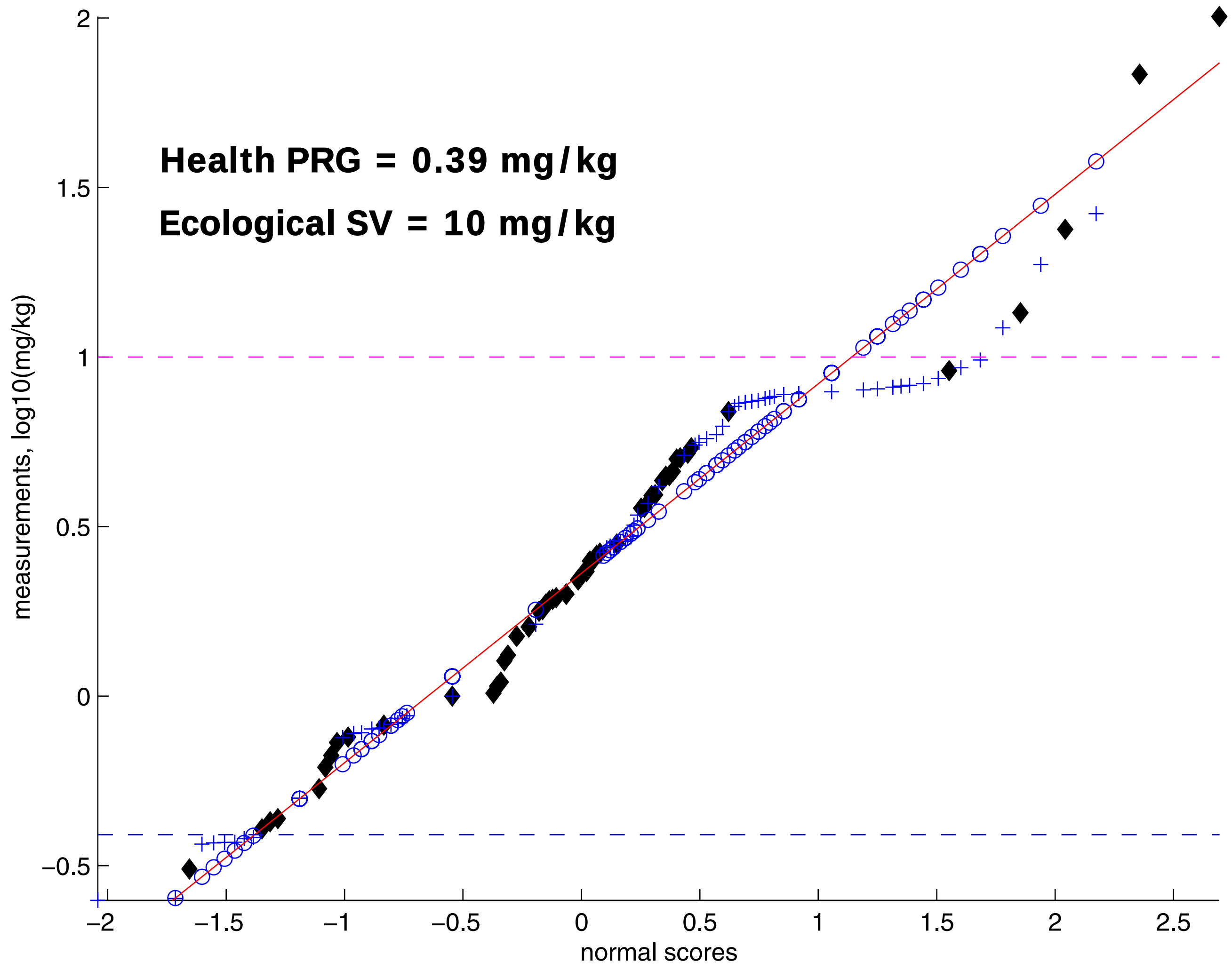
Evaluation of Implemented HRM by Comparison to MLE Method from UNCENSOR

- 10 synthetic realizations of 200 values each, generated from a normal distribution with mean of 200 and a standard deviation of 30
- Data artificially censored at 20th, 40th, 60th, and 80th percentiles
- Our implementation of HRM and MLE as implemented in UNCENSOR version 5.1 used to estimate adjusted mean for the censored data
- Bias and RMS error between actual and estimated means used to evaluate performance of HRM

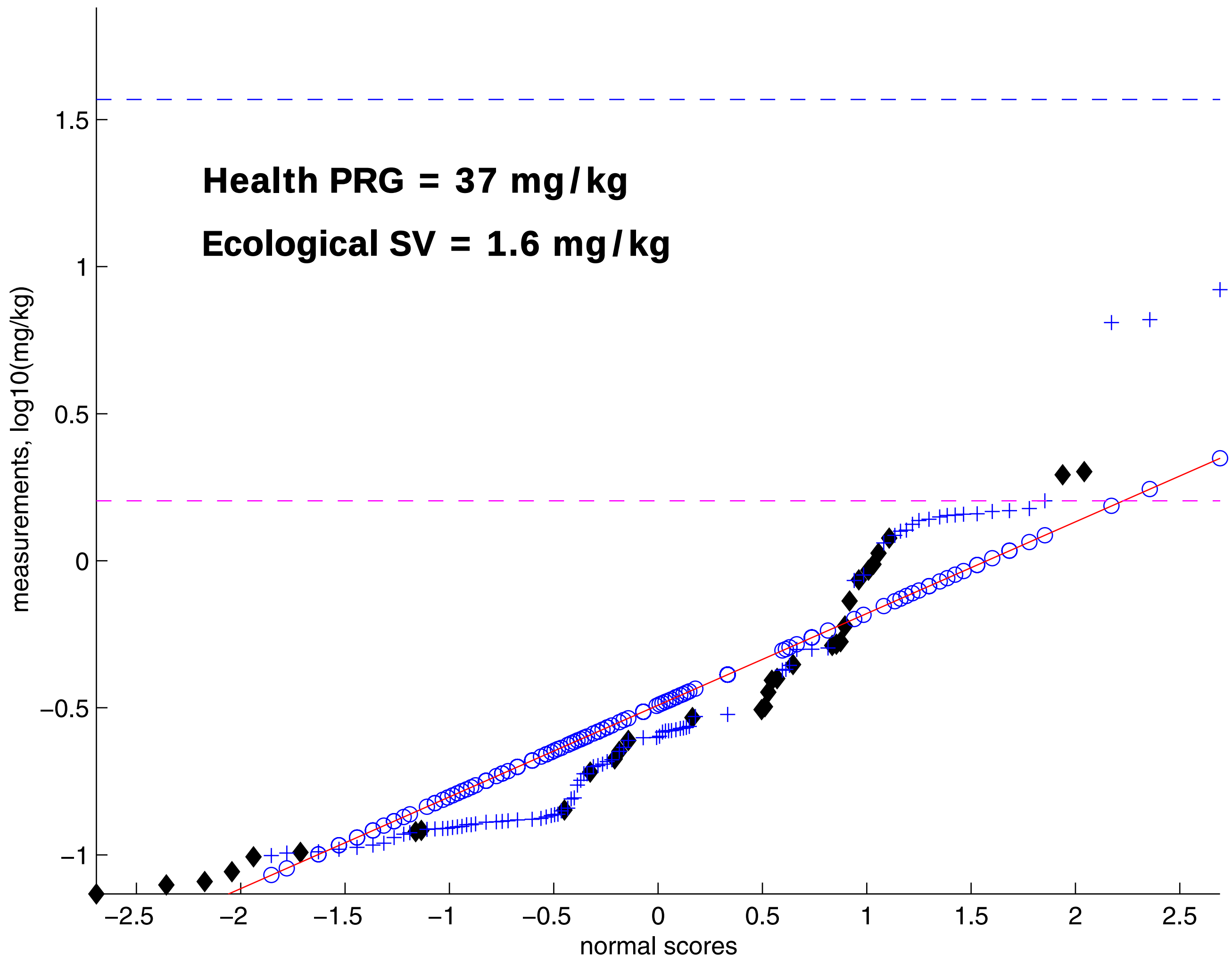
Results of Evaluation of HRM

		Censoring Percentile			
Method	**Metric**	**20th**	**40th**	**60th**	**80th**
MLE	Bias	7.55E-06	3.68E-05	0.000108	0.000271
HRM	Bias	1.47E-05	2.56E-05	0.00012	0.001029
MLE	RMSE	0.002749	0.006069	0.010394	0.016458
HRM	RMSE	0.003828	0.005061	0.010934	0.032077

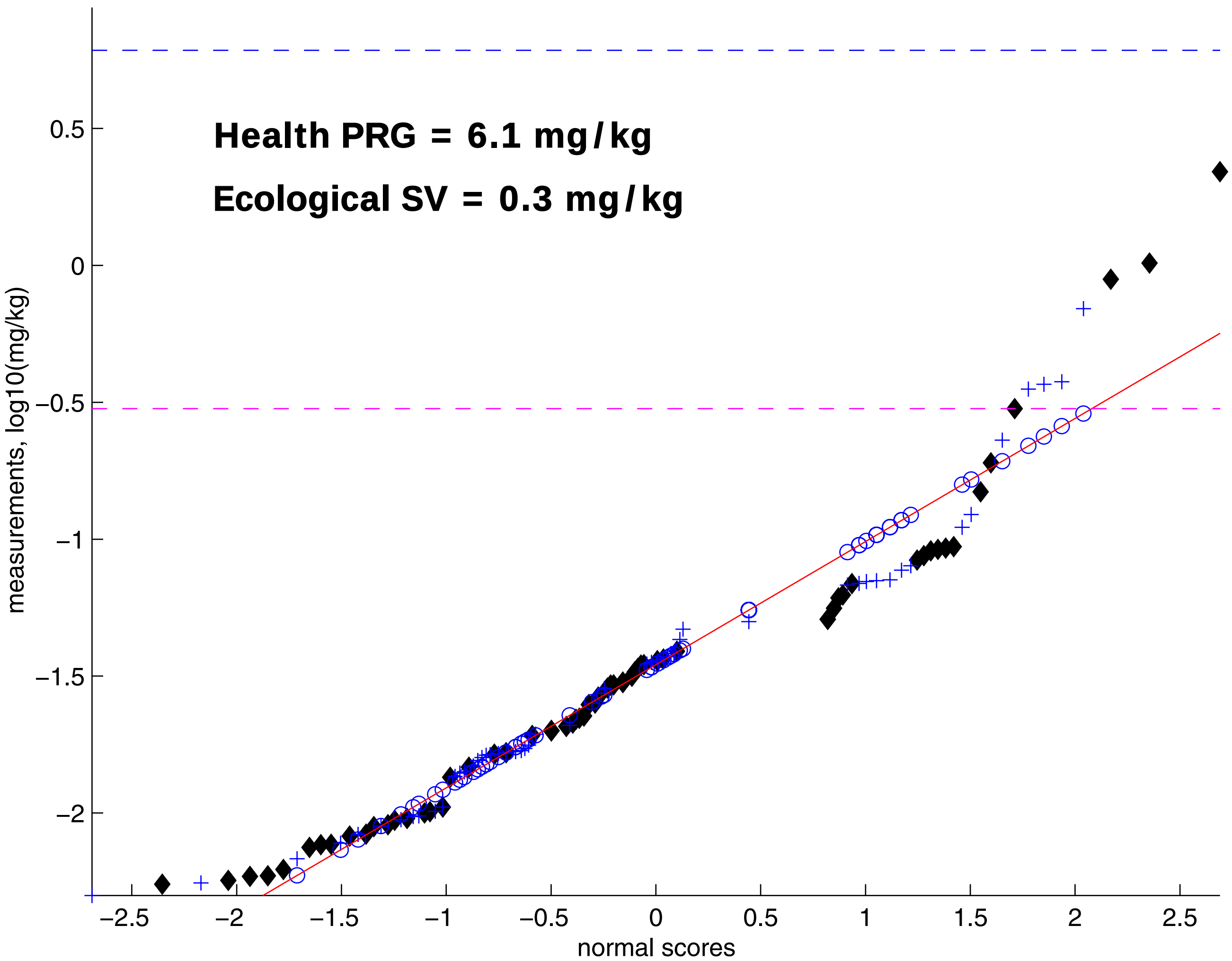
Adjustment of Distribution Parameters for Arsenic



Adjustment of Distribution Parameters for Cadmium



Adjustment of Distribution Parameters for Mercury



Percentiles for COPCs Before and After Adjustment for Non-Detects

		Percentiles (mg/kg)							Percentile of SV	
Type	**COPC**	**5th**	**10th**	**25th**	**50th**	**75th**	**90th**	**95th**	**Health**	**Ecological**
Before	As	0.33	0.44	1.00	2.25	7.33	8.05	9.65	0.08	0.96
After	As	0.30	0.44	1.09	2.25	5.61	11.51	17.47	0.08	0.89
Before	Cd	0.10	0.11	0.13	0.25	0.50	1.38	1.48	1.00	0.97
After	Cd	0.10	0.12	0.20	0.31	0.52	0.89	1.08	1.00	0.98
Before	Hg	0.01	0.01	0.02	0.04	0.05	0.09	0.22	1.00	0.96
After	Hg	0.01	0.01	0.02	0.04	0.06	0.11	0.19	1.00	0.98

Conclusions

- As, Cd, and Hg data contained a very high proportion of non-detects
- In order to make the statistical treatment generalizable, Helsel's Robust Method (HRM) was selected since it is considered to be the optimum method for non-normal distributions
- Presence of multiple MDLs for each COPC precluded use of existing implementation of HRM in UNCENSOR V5.1
- HRM for multiple MDLs was developed and tested for synthetic data, and it performed satisfactorily
- HRM was successfully applied to obtain adjusted distribution parameters for As, Cd, and Hg in SRS soils
- Analysis results indicate that background levels are higher than health and ecological risk based screening values for As

Acknowledgements

This work has been supported by a grant to the Institute for Responsible Management, Consortium for Risk Evaluation with Stakeholder Participation, from the US Department of Energy, Instrument DE-FG26-00NT 40938

The viewpoints expressed in this report are solely the responsibility of the authors and do not necessarily reflect the views of the US Department of Energy or its contractors