STATE OF NEVADA
Department of Conservation & Natural Resources
DIVISION OF ENVIRONMENTAL PROTECTION
NEVADA DIVISION OF ENVIRONMENTAL PROTECTION
protecting the future for generations
Jim Gibbons, Governor
Allen Biaggi, Director
Leo M. Drozdoff, P.E., Administrator

November 14, 2008

Mr. Mark Paris
Basic Remediation Company
875 West Warm Springs Road
Henderson, NV 89011

Ms. Susan Crowley
Tronox LLC
PO Box 55
Henderson, NV 89009

Mr. Curt Richards
Olin Corporation
3855 North Ocoee Street, Suite 200,
Cleveland, TN 37312

Mr. Joe Kelly
Montrose Chemical Corp of CA
600 Ericksen Ave NE, Suite 380
Bainbridge Island, WA 98110

Mr. Brian Spiller
Stauffer Management Co LLC
1800 Concord Pike
Wilmington, DE 19850-6438

Mr. Craig Wilkinson
Titanium Metals Corporation
PO Box 2128
Henderson, NV 89009

Re. **BMI Plant Sites and Common Areas Projects, Henderson, Nevada**
Statistical Analysis Recommendations for Field Duplicates and Field Splits

Dear Sirs and Madam:

Guidance on the treatment and use of field duplicates and field splits is provided in Attachment A. Please contact me with any questions (tel: 702-486-2850 x247; e-mail: brakvica@ndep.nv.gov).

Sincerely,

Brian A Rakvica, P.E.
Supervisor, Special Projects Branch
Bureau of Corrective Actions

BAR:s

CC: Jim Najima, NDEP, BCA, Carson City
Marysia Skorska, NDEP, BCA, Las Vegas
Shannon Harbour, NDEP, BCA, Las Vegas
Todd Croft, NDEP, BCA, Las Vegas
Greg Lovato, NDEP, BCA, Carson City
Barry Conaty, Holland & Hart LLP, 975 F Street, N.W., Suite 900,Washington, D.C. 20004
Brenda Pohlmann, City of Henderson, PO Box 95050, Henderson, NV 89009
Mitch Kaplan, U.S. Environmental Protection Agency, Region 9, mail code: WST-5, 75 Hawthorne Street, San Francisco, CA 94105-3901
Ebrahim Juma, Clark County DAQEM, PO Box 551741, Las Vegas, NV, 89155-1741
Ranajit Sahu, BRC, 311 North Story Place, Alhambra, CA 91801

Rick Kellogg, BRC, 875 West Warm Springs, Henderson, NV 89011
Kirk Stowers, Broadbent & Associates, 8 West Pacific Avenue, Henderson, Nevada 89015
George Crouse, Syngenta Crop Protection, Inc., 410 Swing Road, Greensboro, NC 27409
Nicholas Pogoncheff, PES Environmental, Inc., 1682 Novato Blvd., Suite 100, Novato, CA 94947-7021
Lee Erickson, Stauffer Management Company LLC, P.O. Box 18890 Golden, CO 80402
Keith Bailey, Environmental Answers, 3229 Persimmon Creek Drive, Edmond, OK 73013
Susan Crowley, Crowley Environmental LLC, 366 Esquina Dr., Henderson, NV 89014
Susan Crowley, Tronox LLC, PO Box 55, Henderson, Nevada 89009
Mike Skromyda, Tronox LLC, PO Box 55, Henderson, Nevada 89009
Jeff Gibson, AMPAC, 3770 Howard Hughes Parkway, Suite 300, Las Vegas, Nevada 89109
Sally Bilodeau, ENSR, 1220 Avenida Acaso, Camarillo, CA 93012-8727
Cindi Byrns, Olin Chlor Alkali, PO Box 86, Henderson, Nevada 89009
Paul Sundberg, Montrose Chemical Corporation, 3846 Estate Drive, Stockton, California 95209
Joe Kelly, Montrose Chemical Corporation of CA, 600 Ericksen Avenue NE, Suite 380, Bainbridge Island, WA 98110
Deni Chambers, Northgate Environmental Management, Inc., 300 Frank H. Ogawa Plaza, Suite 510, Oakland, CA 94612
Robert Infelise, Cox Castle Nicholson, 555 California Street, 10th Floor, San Francisco, CA 94104-1513
Michael Ford, Bryan Cave, One Renaissance Square, Two North Central Avenue, Suite 2200, Phoenix, AZ 85004
Dave Gratson, Neptune and Company, 1505 15th Street, Suite B, Los Alamos, NM 87544
Paul Black, Neptune and Company, Inc., 8550 West 14th Street, Suite 100, Lakewood, CO 80215
Teri Copeland, 5737 Kanan Rd., #182, Agoura Hills, CA 91301
Paul Hackenberry, Hackenberry Associates, 550 West Plumb Lane, B425, Reno, NV, 89509

Attachment A

Introduction

Duplicates and splits are usually included as part of data collection studies to assess environmental contamination or chemical concentrations. This document describes options for using data from duplicates and splits in statistical analyses of environmental data, where the statistical analyses might support site characterization, background comparisons, risk assessment or other environmental decisions. For the purpose of this document, duplicates and splits are defined as follows, with generalization to multiple duplicates and splits assumed:

- Field duplicate: A distinct sample collected from the same point in time and space as the first sample, or as near to the same time and place as possible. The field duplicate is analyzed separately from the first sample, and is assigned the label of field duplicate.

- Field split: A separate sample taken from a sample homogenized in the field. That is, a field sample is homogenized and then split into two samples. The second sample is analyzed separately, with the second one assigned the label of field split.

- Laboratory replicate: In the laboratory, a sample is sub-sampled for analysis. This process is repeated for a laboratory replicate. Both samples are analyzed, with the second one assigned the label of laboratory replicate.

- Laboratory split: A sample is homogenized and split into two samples. These samples are then analyzed, with the second one assigned the label of laboratory split.

The United States Environmental Protection Agency (EPA) describes the use of duplicate and split samples in the context of quality control (QC), with specific attention to the role these QC samples can play in the evaluation of measurement error (EPA G-5, Guidance for Quality Assurance Project Plans, 2002). The EPA guidance also suggests how many QC samples are appropriate to support data collection for chemical concentrations in environmental samples. In G-5, EPA describes how field duplicates and field splits, and laboratory replicates and laboratory splits, provide the information needed to break down the overall variance structure of chemical concentration data from environmental samples. If collected solely for the purpose of QC then this concept works. However, most field duplicates and field splits are collected as part of a site characterization program, which causes the duplicates and splits to be considered in that context as well. Laboratory replicates and splits, by contrast, are more often used in the context of QC. Consequently, the recommendations in this document pertain primarily to field duplicates and field splits, although the options described below can also be applied to laboratory replicates and laboratory splits.

In practice, field duplicates and splits are not used to serve the QC purpose as described in EPA guidance. Instead, field duplicates and splits are collected primarily because the EPA guidance prescribes their collection. They are usually used to support data validation, but often superficially from a decision making perspective. They are rarely, if ever, used to break down components of variability. Instead, and unfortunately, these QC samples serve to obfuscate statistical analysis performed to support site characterization, background comparisons, risk assessment, and other important environmental decisions.

When field duplicates and splits are collected some choices need to be made regarding how to process or statistically analyze their data to support nature and extent or risk assessment decisions. These choices are surprisingly difficult. The purpose of this guidance is to provide some clasrification and options on how to use the field duplicate and split sample results in the statistical analysis of chemical concentration data to support nature and extent, background, risk assessment, or other relevant environmental decisions.

Options

Options that can be considered for the inclusion of these QC samples in statistical analyses depend on the variance of the QC samples and the variance of the site samples. There is plenty of evidence to suggest that the variability in field duplicates for many chemicals under many environmental conditions is no different (statistically) than the variability between site samples. This depends on the effect of small-scale spatial correlation. If there is no small-scale spatial correlation then the field duplicate results are similar to other site samples. If there is small-scale spatial correlation, then field duplicates will show lower variance than site samples. The variability for field splits, however, is expected to be less than variability for site samples because of the within sample homogeneity caused by the physical mixing process prior to splitting the sample. The same is the case for laboratory replicates and laboratory splits because homogenization has occurred prior to creating separate samples. The options presented below are focused on field QC, but the treatment of laboratory replicates and splits should be similar to the treatment for field splits.

To simplify the discussion of statistical analysis options, assume that there is a site sample and one associated field duplicate, or a site sample and one associated field split. Extension of this guidance to accommodate situations involving more than 2 sample results is assumed. There are 3 basic options, each of which has different consequences depending on the data.

1. Use the first sample and ignore the second sample.

2. Treat the sample and its field duplicate as independent samples (field duplicates only).

3. Average the sample results.

Each of these options is discussed in more detail below:

1. The simplest solution is to use the site sample and ignore the duplicate or split. The only general argument against doing this concerns loss of information/data that has been collected. Specific arguments based on the actual data could also arise. For example, if the site sample's concentration for an analyte is much less than the duplicate or split concentration, or the site sample is a non-detect and the duplicate or split is a detect, then this approach might not be considered adequately protective from a risk perspective. The data should be checked for potential issues prior to choosing this method.

2. [Field duplicates only.] EPA in their Data Quality Assessment guidance "suggests" treating the site sample and field duplicate as independent samples (EPA G-9, Data Quality Assessment Guidance, 2006). This is reasonable if the variability of the site samples is similar to the variability of field duplicates. Some judgment can be made about relative variability by summarizing the data. For metals this is often a reasonable approach, but it might not be as reasonable for organic chemical spills or, more generally, if

there is small-scale spatial correlation. EPA does not force this approach in G-9, but acknowledges that this approach avoids statistical issues that can arise with averaging (Option 3). Averaging violates some basic statistical assumptions of the types of tests and estimation procedures used with environmental data. This Option also solves potential problems described in Option 1 with using the site sample only. Again the data should be reviewed prior to choosing this Option, including attention to the variability of field duplicates and the variability of site samples, and to the conceptual site model.

3. The third Option is to average the sample and the field duplicate or split sample. This option patently violates statistical assumptions that underlie hypothesis testing and estimation of confidence limits. Essentially, more information is included in the averaged result, because it represents greater sample volume. Statisticians term the relevant assumptions *iid*. That is, each sample must be independent and identically distributed. The statistical language could be described as "each sample must come from the same statistical distribution or population". When field duplicates are averaged, the result has a lower variance, and hence violates these basic assumptions.

Averaging might, nevertheless, be a reasonable option. If averaging is performed the violations of the statistical assumptions can usually be ignored. Averaging is straightforward when both the sample and its duplicate or split are detects. However, averaging is more complicated when non-detects are involved. Consider a simple case for which the sample results are 1 and 9 in some concentration units, and assume averaging is the selected option. If both results are non-detects then they should be averaged to produce a value of "less than 5". This reasonably describes the possibilities. Both results could be as low as zero, or they could be as high as 1 and 9 respectively. Their average ranges from 0 – 5, and can reasonably be used in statistical analysis as a value of < 5. The term "less than" can be interpreted statistically similarly to the way in which non-detects are interpreted.

If one of the values is a detect and one is a non-detect, then it is reasonable to assume that the chemical has been detected. For example, if the value of 1 is a non-detect and the value of 9 is a detect, then the average is between 4.5 and 5. An averaged value of "less than 5" could be used, but it is not fully descriptive. Instead, the substitution rule of ½ detection limit should be applied before averaging to generate a result of 4.75. The same approach should be used if the value of 1 is a detect and the value of 9 is a non-detect, a situation that is unfortunately common with environmental data. Now the average could be between 0.5 and 5. A value of "less than 5" could again be reported, however, if the same substitution rule is used, then a value of 2.75 would be generated. Given the options, the preferred averaging approach is to report the average as a detect and use ½ the detection limit for the non-detect before averaging.

Discussion

The 3 options presented above describe methods by which field duplicates and field splits can be managed for statistical analysis of environmental chemical data. It is important to defend the approach that is taken in light of the data, the conceptual site model, and the purpose of the statistical analysis. The reason for defending the approach is so that the risk results or background comparisons or other environmental statistical analyses are suitably documented and justified.

Summary

NDEP requires justification for the approach that is used, and has the following preference ordering for field duplicates and field splits.

For field duplicates, NDEP's preference is to use Option 2 unless the field duplicates exhibit obvious lower variance than the site data. If the latter is the case, then Option 1 is the preferred method. If there are obvious issues with Option 1 (see description above), then Option 3 can be used instead.

For field splits, Option 2 will rarely be appropriate. The preferred approach is to use Option 1, unless the data indicate potential issues with doing so (see description above). Otherwise, Option 3 can be used if necessary.

If laboratory replicates or laboratory splits are involved in the statistical analysis, then the preference order should follow that of field splits, because in all of these cases the variance of the QC samples is expected to be less than the variance of the site samples.

Although the 3 Options are described above, NDEP recognizes that in rare cases another approach might be more appropriate. NDEP expects that alternative approaches will be justified based on the data and the conceptual site model. For example, a sample might indicate a high concentration, whereas the QC sample indicates a low concentration or a non-detect. Options 1 and 2 are clearly inappropriate, and Option 3 might also be inappropriate if the initial high concentration is unexpected, does not match the remaining data, and does not fit the conceptual site model. Under these circumstances an argument might be made that the QC sample should be used in lieu of the site sample.

The Options presented herein describe approaches that can be taken for inclusion of these QC samples in statistical analysis of site data, and NDEP provides a preference order for these options for field duplicates, field splits, laboratory replicates and laboratory splits. However, NDEP will accept any of these Options or another approach provided sufficient justification is provided in the context of the site data and the conceptual site model.