Dealing With Beryllium Data: What About Those Limits?

Charles B Davis, PhD
Principal Statistician
ENVIROSTAT

!!!!!! Disclaimer !!!!!!

The views and opinions expressed here are those of the authors, and should not be taken to necessarily represent positions, policies, or views of the United States Government, the Department of Energy, or any of its contractors or subcontractors.

Any reference to products, companies, persons, or organizations is for information purposes only and does not represent any form of endorsement or criticism.

Plan for Today

Introduction

- Overview of Chapter 6 of the Beryllium Practitioners Guide
  - "Data Evaluation & Related Topics"
- 3 types of Limits
  - Established
    - Occupational Exposure Limit (OEL)
    - DOE Release Criterion (RC)
  - Data reporting and interpretation
    - "Detection", "Quantitation", "Reporting"
  - Limits to be established from your local background soil data
    - Main focus for today
    - More questions than answers

Introduction

Overview of Chapter 6

- "Data Evaluation & Related Topics" main sections
  - Data Quality Objectives and Data Quality Analyses
  - Sampling plans
  - The nature of actual (uncensored) beryllium data
  - "Detection" and Reporting limits
  - Data reporting and recording
  - Data evaluation
  - Supplement: More on "Detection" and Reporting Limits

Introduction

Limits - 1

- Established limits
  - OELs and RCs
  - RTELs: "Rarely-To-Exceed-Limits"
    Based on IH practice
  - Compare data with RTELs via upper tolerance limits
    95%-95% UTL is upper 95% confidence limit for 95th percentile
    Implements statistical hypothesis test with H_0: 95th %ile ≥ RTEL
    H_A: 95th %ile < RTEL
    Presume "dirty" unless you have enough evidence to show "clean"
    Who gets the null hypothesis?
    "Getting enough evidence" guides number of observations
- Discussed in Chapter 6

Introduction

Limits - 2

- "Detection" and Reporting Limits
  - Continuation of Chapter 8 of the Beryllium Book
    "Data Use, Quality, Reporting, and Communication", with Nancy Grams
    in *Beryllium: Environmental Analysis and Monitoring*
    MJ Brisson & AA Ekechukwu, eds, 2009
  - EPA recently released revisions to Method Detection Limit
    Some improvements but problems remain
  - Differences in policy between EPA and AIHA regarding RLs
  - Factsheet from EPA Region III Quality Assurance, 2006
    IDL-MDL-PQL
    What the "L" is going on?
    What does all this alphabet soup really mean?
- Discussed in Chapter 6

Introduction

Limits - 3

- Limits to be established from your background data
  - For comparing in-facility data with "local soils"
    - Need to be "not greater than", maybe actually "lower than"?
    - Who gets the null hypothesis?
  - Comparing a dataset with a dataset is one situation
    - Usual statistical tests in many cases
    - But there can be issues
  - But if we have a fixed background dataset and repeatedly compare data with it, we may wish to establish some sort of limit from background data
    - Can use to repeatedly compare data with local soil background; don't need to repeat statistical analysis

Introduction

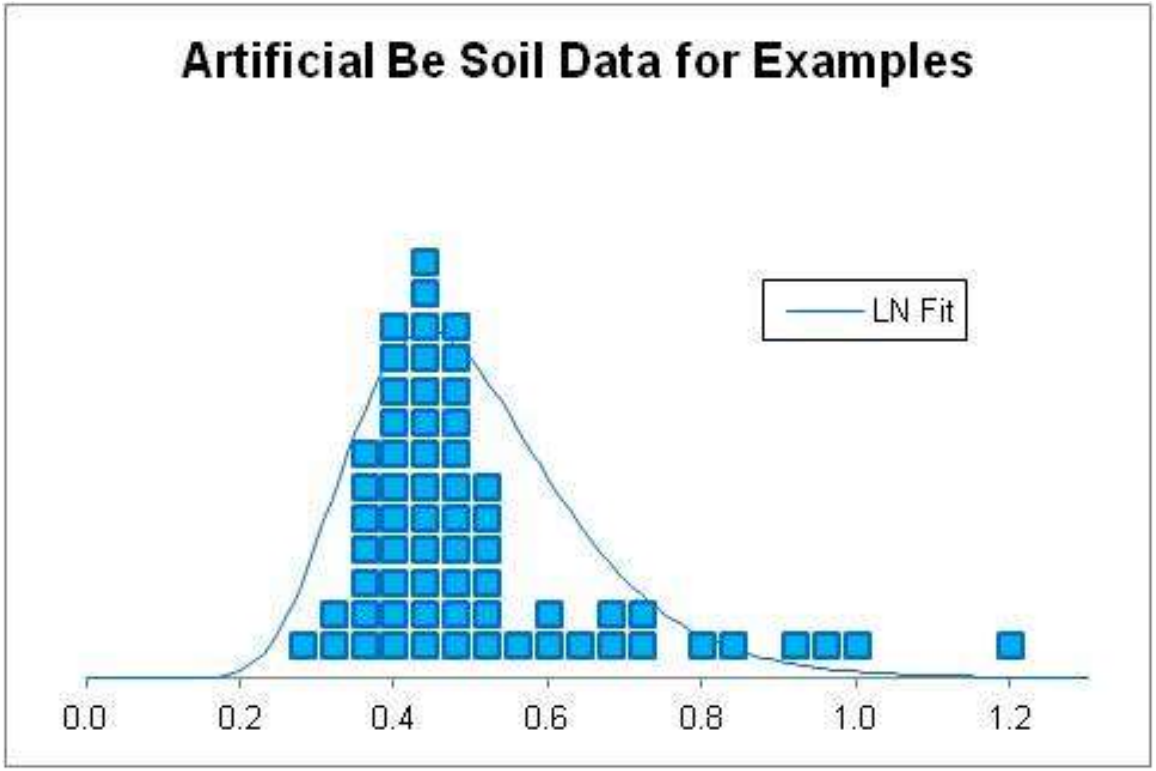
BHSC SAS Data Reporting Task Force ~ 2007

- **Attempted to mimic RTEL idea (assume dirty until shown clean)**
  - **Use BG soil mean plus allowance for variability**
    - **Allowance would be perhaps 110% of BG mean**
    - **Or maybe LCL for BG mean plus 1.0 µg/g**
    - **What to be compared with that was not stated (hopefully not all values!)**
  - **Saw that idea used once**
    - **Very dusty forward area facility**
    - **Limit established that way was 400% of BG mean**
    - **= BG mean + 1.5 µg/g**
    - **Actual decision was made by standard nonparametric tests with up-shifted BG data; areas of concern not found greater than BG**

Data for Examples

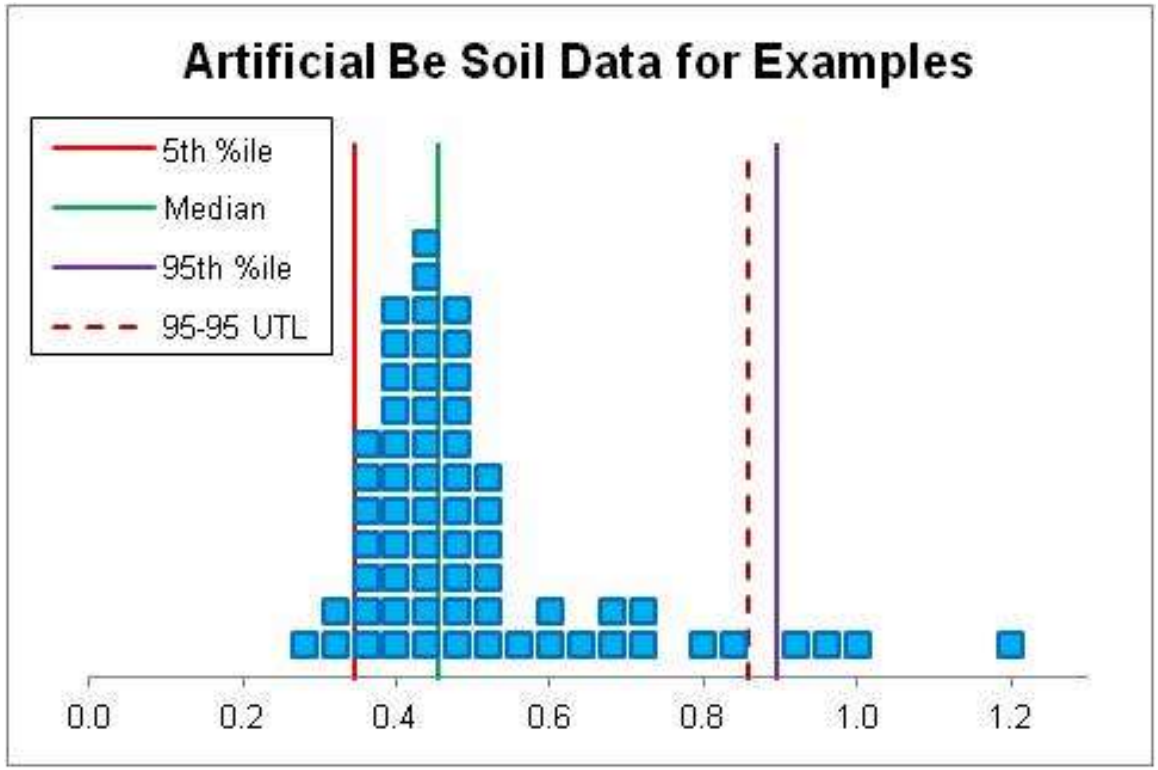
- **Artificial but realistic Be surface soil data**
  - **Fitted lognormal distribution**

    GM = 0.48, GSD = 1.33



Data for Examples

- **Some metrics of the example data**
  - **5th %ile = 0.35**
  - **Median = 0.45**
  - **95th %ile = 0.89 ≈ 95-95 Upper Tolerance Limit = 0.86**



What Limit to Use?

What limit do we compute for comparisons with future sample values?

- **Maximum?**
  - **Of how many BG measurements?**
  - **Obtained how and where?**
- **95th %ile or 95-95 UTL or ?**
- **Mean or GM?**
  - **Perhaps with fudge factor added to allow for sampling and analytical variability**

How do we compare future sample values with a limit?

- **What metric of future sample do we use?**
- **Who gets the null hypothesis?**

THESE ARE INTRINSICALLY RELATED!

What Limit to Use?

How do we compare future sample values with Limit?

- (1) Compare every data value with Limit
  - Should take Multiple Comparisons (MC) into account
    - How many comparisons are you going to make (this week/year)?
    - Will worst case create regulatory activity or headlines?
  - But draft DOE Reg wants in-facility bulks lower than local soils
    - " … employer must ensure that … [removable contamination level does not exceed DOE RC] … ; or [10 CFR 850.31(a)(2)] Concentration of beryllium in bulk material on the surface is lower than the concentration in soil at the point of release; or …"
  - Idea: bulk samples in many facilities are soils diluted with "fluff"; should Limit reflect this? If so, how?
    - (Motivation for Metal Ratio Method)
    - Can one incorporate nature of facility into such a Limit, perhaps using MRM ideas?

What Limit to Use?

How do we compare future sample values with Limit?

- (1) Compare every data value with Limit
  - Should take MC into account
  - What is decision rule?
    - Who gets the null hypothesis?
  - If we assume "clean" until an exceedence says "dirty", how many samples to take?
    - May be an arbitrary dictate – cf groundwater monitoring
  - Under "clean until shown dirty", have statistical tests
    - MC with control ≡ prediction limits for all future values (PLall)
    - Statistical power may be less than desirable
  - Can (should) one downshift PLall to allow for "fluff"?
    - Some technical issues come up regarding statistical assumptions such as equality of variance and similar distribution shape, particularly with censored data

What Limit to Use?

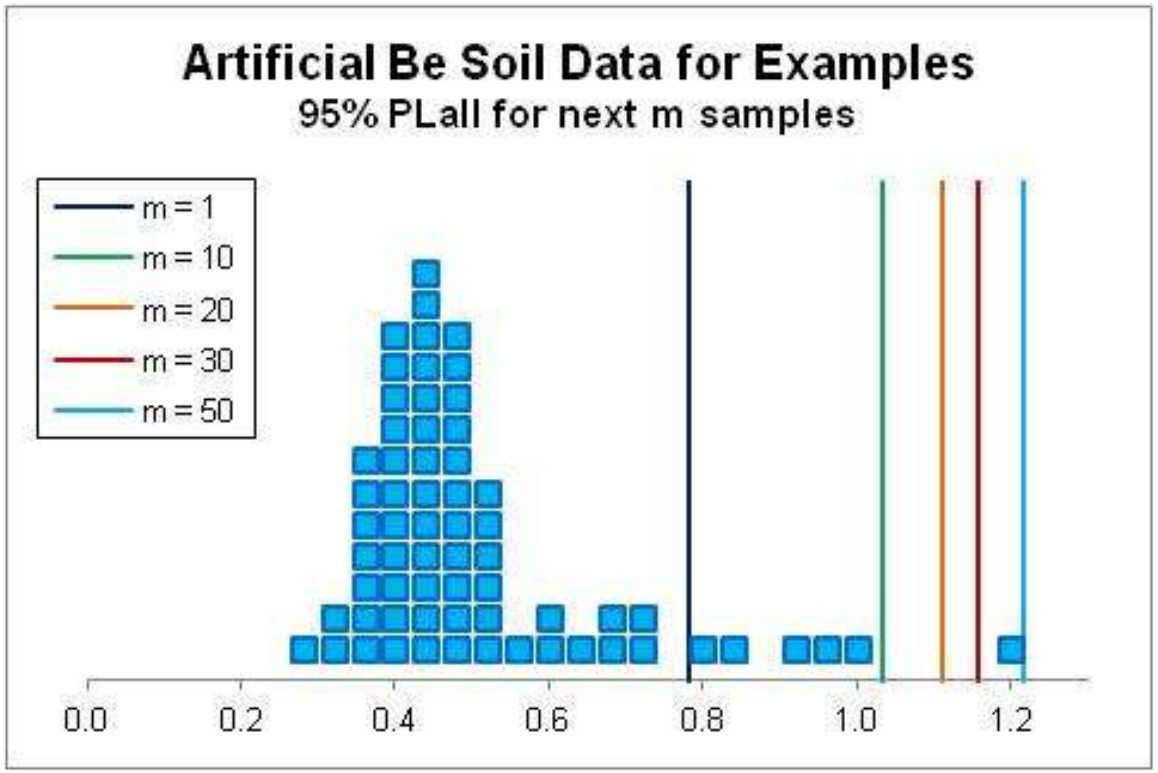
How do we compare future sample values with Limit?

- (1) Compare every data value with Limit
  - Can we come up with an RTEL for this situation?
    Chapter 6 comparisons with RTEL are for datasets, not individual values
  - *Caveat*: If Limit is high enough to take care of MC issue, there's lots of room for lower level contamination to sneak in

What Limit to Use?

How do we compare future sample values with Limit?

- (1) Compare every data value with Limit
  - Example of PLall (LN computations using log scale data)



What Limit to Use?

How do we compare future sample values with Limit?

- (2) Compare dataset means with Limit
  - Limit might include fudge factor, as in example in slide 8
  - Could use GMs, equivalent to using log-scale computation
  - MC issue is now how many datasets are compared with Limit
  - Comparing datasets is fairly standard procedure
    - Decide Who gets null hypothesis, frame statistical test accordingly
    - Have seen null hypothesis simply "distributions the same"
    - Can express usual *t*-test in terms of UCL or LCL comparison with an appropriate Limit (LCL or UCL) based on background data
    - These are UCLs or LCLs for mean (or GM)
    - Would be more challenging to do so with nonparametric test
  - One always has to deal with nondetects
  - This Limit would NOT be the same as that for option (1)!

What Limit to Use?

How do we compare future sample values with Limit?

- (2) Compare dataset means with Limit
  - How to deal with the “fluff” issue?
  - One might deflate Limit somehow to compensate for “fluff”
    - Might create issues with assumptions for statistical tests
    - Are distributions comparable? (What about nondetects)
    - One could compare QNP UTLs with downshifted limits
  - How to deflate? Maybe introduce Metal Ratio ideas

Quick Review

Agenda

- Overview of Chapter 6 of the Beryllium Practitioners Guide
  - "Data Evaluation & Related Topics"
- Established Limits (1): OELs and RCs => RTELs
- (2) "Detection" and Reporting Limits
  - See Chapter 6 for these (1) and (2)
- (3) Limits from <u>your</u> local background soil data: issues
  - What is to be compared with the Limit
    - All data values or mean or something else
  - How often (remember that MC issue)
  - Who gets the null hypothesis?
  - If "clean until shown dirty", how many samples?
  - If "dirty until shown clean", how to set up like an RTEL

Discussion?