CLASSIFICATION	DOCUMENT IDENTIFICATION NO. PNL-SA-6695
Battelle Pacific Northwest Laboratories Richland, Washington 99352	COPY AND SERIES NO.				
	DATE SEPTEMBER, 1977				
TITLE AND AUTHOR TRAN-STAT, ISSUE #1, SEPTEMBER. TOPICS DISCUSSED: QUANTITATIVE ASPECTS OF <u>PU</u> FIELD STUDIES. R. O. GILBERT	<table border="1" style="width:100%; border-collapse: collapse;"> <tr> <td style="width:50%;"> CONTRACT ☐ - 1830 ☐ - 1831 </td> <td style="width:50%;"> PROJECT NO. </td> </tr> <tr> <td colspan="2" style="text-align: center;"> RESERVED FOR TECH. INFO. USE </td> </tr> </table>	CONTRACT ☐ - 1830 ☐ - 1831	PROJECT NO.	RESERVED FOR TECH. INFO. USE	
CONTRACT ☐ - 1830 ☐ - 1831	PROJECT NO.				
RESERVED FOR TECH. INFO. USE					

D I S T R I B U T I O N

NAME	COMPANY	LOCATION	NAME	COMPANY	LOCATION

NOTICE

This report was prepared as an account of work sponsored by the United States Government. Neither the United States nor the United States Department of Energy, nor any of their employees, nor any of their contractors, subcontractors, or their employees, make any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness or usefulness of any information, apparatus, product or process disclosed, or represents that its use would not infringe privately owned rights.

MASTER

DISTRIBUTION OF THIS DOCUMENT IS UNLIMITED

ROUTE TO	PAYROLL NO.	COMPANY	LOCATION	FILES ROUTE DATE	SIGNATURE AND DATE

DISCLAIMER

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency Thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

DISCLAIMER

Portions of this document may be illegible in electronic image products. Images are produced from the best available original document.

TRAN-STAT

Statistics for Environmental Transuranic Studies

September 1977
Number 1

Battelle Memorial Institute
Pacific Northwest Laboratory
Richland, Washington 99352

Coordinator: Richard O. Gilbert, Energy Systems Department,
Statistics Section, (FTS) 444-7511, ext. 946-2378

This is the first issue of TRAN-STAT which we hope to issue on a fairly regular schedule. The purpose of TRAN-STAT is to stimulate your thinking about how you can make better use of statistics in the design and analysis of environmental transuranic studies. In this first issue we will be talking about some of our own current statistical work for DBER under the title "Quantitative Aspects of Plutonium Field Studies," some of our thoughts on appropriate data summaries for transuranic studies, and some new statistical references of interest. Future issues might discuss topics such as regression, ratios, nonparametric statistics, estimating concentration contours and/or inventories, the analysis of spatial data, and statistical approaches to sampling and analysis for deciding whether a contaminated area requires cleanup. We hope, however, that you will forward us your ideas on what you would like to see in future issues. The format of TRAN-STAT should be whatever is best suited to meeting your statistical design and analyses needs. TRAN-STAT will definitely not take the place of a statistician working closely with you on a day to day basis. Hopefully TRAN-STAT will stimulate you to consider more fully the statistical aspects of environmental transuranic studies and to encourage you to seek professional statistical help early in the design phases of your study.

QUANTITATIVE ASPECTS OF PLUTONIUM FIELD STUDIES

We would like to tell you something about our current activities under the above titled DBER funded program. The purpose of this program is to provide quantitative and statistical support and methodology for the design and analysis of plutonium and other transuranic ecosystem and field studies. Our approach has been to (1) become more aware of statistical design and analysis problems in current field studies through site visits, meetings, and phone conversations, (2) write papers on statistical topics applied to field studies to help meet these needs, (3) increase awareness of the role that statistics can play in field studies (through TRAN-STAT, e.g.), and (4) engage in the development of new statistical techniques tailored to transuranic field studies.

We have visited several transuranic field study sites during the past two years (Los Alamos, Oak Ridge, Savannah River, Mound Laboratory, Hanford, and, of course the Nevada Test site where we have provided statistical support for the Nevada

Applied Ecology Group (NAEG) since 1971). We are also in the process of revising two papers, one dealing with statistical analysis of ratio data ($^{238}\text{Pu}/^{239-240}\text{Pu}$, $^{239-240}\text{Pu}/^{241}\text{Am}$, etc.) and the other with effective methods for summarizing and reporting transuranic field data. Our work on the development of new statistical techniques is currently centered on experimenting with the use of Kriging techniques for estimating Pu concentration contours and inventories for transuranics in soil. Kriging attempts to make use of whatever correlation may exist between concentrations at various distances apart to obtain improved estimates of contours and inventories. Whether or not the technique is really applicable to Pu field studies remains to be seen. Kriging has been programmed and is currently available for ERDA funded projects on ERDA's computer in Las Vegas. The title of the program is BLUEPACK which is used in conjunction with a companion program called GAMMA. These programs were purchased by ERDA from the Center for Geostatistics, School of Mines, Fontainebleau, France. We are currently accessing these programs over a high speed terminal from Richland. Our research on Kriging is jointly funded by the NAEG. The data we are currently using are from NAEG study sites on the Nevada Test Site. Dr. Pam G. Doctor (FTS 444-7511, extension 946-2444), a statistician here at Battelle-Northwest, is in charge of our Kriging effort. Those of you interested in more information can give her a call.

DATA SUMMARIES

Transuranic data from a field study site should be summarized and displayed in appropriate ways to communicate effectively what the data have to tell us about the site from which samples were taken. The particular statistical methods used will depend on a number of factors including the number of data points, variability of the data, spatial orientation of the sampling locations, as well as the purpose for which the data were originally collected. Hence, there is not one best method that works in all situations.

In general, we should look at the data in a variety of ways so that it becomes thoroughly familiar. This implies that we not be content with computing only the arithmetic mean and standard deviation. These are sufficient only if the data being described are normally (gaussian) distributed, which is rarely the case for transuranic concentrations. The common occurrence of skewed distributions of data implies we should also consider computing the median (value above which and below which half the observations lie) and interquartile range (values between which lie the mid 50% of the data). If, say, 15 or more observations are available we should consider histograms or "stem-and-leaf" displays to describe the distributions shape. Gilbert, et al. (1977) discuss these descriptive measures relative to transuranic studies. Tukey (1977) gives a more complete account of stem-and-leaf displays.

The primary purpose of the sampling study influences the type of summarization used. If an estimate of the total amount of Pu in, say, surface soil is desired, then the arithmetic mean is usually required. However, if one is sampling to determine areas requiring cleanup to specified levels, then the arithmetic mean, median, proportion of observations about the cleanup criterion, largest concentration obtained, the spatial orientation of the samples, as well as other measures may all be required.

Note that no mention of statistical tests was made above. Statistical tests help us to make inferences about the universe from which we have obtained tiny bits of information (from our samples). Such tests certainly have a place in transuranic field studies, but they should be used with care and only after the data have been effectively summarized and displayed. Our discussion here has centered on the types of things you might do with the data before tests are done. In many, if not most cases, effective displays will make visually obvious any statistically significant differences that exist. Moreover, effective displays may point out features of the data that may not have been foreseen and which may suggest further research.

Since field studies involve collecting samples over geographical space, we should try to display our data to highlight the variability in the data over space. Many times it is the pattern of variation over space that is most informative, and this information may be completely masked by lumping data together and reporting only a median, arithmetic mean, and standard error.

Those of you interested in reading further on methods for getting more information out of your data might be interested in two new publications: Tukey (1977) and Mosteller and Tukey (1977). These are rather unique books at this point in time. They concentrate on exploring data sets and basic statistical analyses using new techniques (such as stem-and-leaf displays) to find out what the data seem to be saying to us. They are written in a style much more informal than the usual statistics text.

Another recent statistical reference of interest is by Link and Koch (1975). They deal with the question: How bad off are you if you use statistical methods appropriate for lognormally distributed data when in fact the data are not really lognormal, but some other skewed distribution? This reference is directed toward non-statisticians in the field of geology, but should be of interest to those engaged in the analysis of transuranic data.

FINAL COMMENT

Let us have your ideas on what topics you would like to see discussed in TRAN-STAT.

REFERENCES

- Gilbert, R. O., L. L. Eberhardt, and D. D. Smith. 1977. "An Initial Synthesis of Area 13 ^{239}Pu Data and Other Statistical Analyses." In NAEG Progress Report NVO-171, USERDA, Las Vegas. (In press)
- Link, R. R. and G. S. Koch, Jr. 1975. "Some Consequences of Applying Lognormal Theory to Pseudolognormal Distributions." Mathematical Geology 7:117-128.
- Mosteller, F. and J. W. Tukey. 1977. Data Analysis and Regression: A Second Course in Statistics. Addison-Wesley, Reading, Massachusetts.
- Tukey, J. W. 1977. Exploratory Data Analysis. Addison-Wesley Publishing Company, Reading, Massachusetts.