# STATISTICAL CONSIDERATIONS IN THE DESIGN OF A BASELINE RADIOLOGICAL SURVEILLANCE PROGRAM AT THE WASTE ISOLATION PILOT PLANT

N. Timothy Fischer
International Technology Corporation
Waste Isolation Pilot Plant
P.O. Box 2078
Carlsbad, New Mexico 88221

ABSTRACT

Two contradictory factors must be considered in developing a sampling strategy for a preoperational environmental surveillance program. First is the stochastic nature of radiological data, which dictates large sample sizes to ensure validity and statistical confidence, especially when a large geographic area is being characterized. Second is an upper limit to the total sampling effort, imposed by practical constraints such as available time, manpower, and funding. These contradictory factors create a dilemma in program design with respect to the allocation of sampling effort among the various components of the program.

The Radiological Baseline Program (RBP) was initiated at the Waste Isolation Pilot Plant (WIPP) in July 1985 to establish a preoperational database for the WIPP Operational Radiological Monitoring Program (ORMP). In the RBP, a partial resolution to this dilemma was achieved through the following design features:

(1) Sampling was limited to those sites which would be used as sample locations during the OKMP, or which might be used during operations under special circumstances. These sites were strategically located to maximize the probability of detecting a release from the WIPP and to characterize radiation at sites of special concern. Thus, each sampling site in the ORMP will have its own preoperational database.

(2) Sampling locations were classified as either primary or secondary sites based upon their importance to the goals of the ORMP. The intensity with which each will be sampled during the RBP will be adjusted according to this classification.

(3) Existing radiological data records from the area were used to predict the amount of spatial and and temporal variability which can be expected for equivalent parameters being measured in the RBP and ORMP. These data enable predictions to be made of the number of samples required to achieve a designated degree of statistical confidence.

(4) Where possible, existing data records were used to predict the type of statistical model which would most likely describe the behavior of equivalent parameters in the RBP. This information is useful in determining sample size and in prescribing methods of data transformation to facilitate statistical analysis.

## INTRODUCTION

The Waste Isolation Pilot Plant (WIPP), currently under construction in southeastern New Mexico, is a research and development facility to demonstrate the safe disposal of defense generated transuranic (TRU) waste in subsurface salt beds. Beginning in 1988, TRU waste will be transported to the facility and deposited in mined-out rooms in the Salado Formation, 670 meters below ground surface. Normal waste handling operations at the WIPP will not result in releases of radioactive elements to the environment.

The WIPP Operational Radiological Monitoring Program (ORMP) is scheduled to begin with the first arrival of waste. In this program, the environment around the WIPP and in adjacent communities will be carefully monitored for evidence of radionuclide releases. Radioactivity levels and concentrations of selected radionuclides in a variety of media (air, water, soil, biotic tissue) will be measured either continuously or at regular intervals throughout the operational life of the facility.

The effective evaluation of these data will require an estimate of the amount of radioactivity in the environment which is not accountable to WIPP operations. The WIPP Radiological Baseline Program (RBP), initiated in July 1985 is a preoperational radiological surveillance program designed to establish baseline values for the parameters being measured in the ORMP. The validity of these baseline values depends strongly on statistical considerations in the design of the program, such as sample size, sampling scheme, and the random representation of the various sources of variability.

Sample size is of particular importance in the design of a preoperational radiological surveillance program. Two contradictory factors which influence sample size must be considered in developing a sampling strategy: (1) the stochastic nature of the parameters, and (2) physical or logistical limitations imposed on the program.

The stochastic nature of environmental parameters is the result of a number of sources of random variability. These include such factors as the location from which the sample was taken, the time of collection, and various random events occurring during collection, transport, and analysis. The variation, or error, accrued in the measurement from all of these sources adds an element of uncertainty to an individual measurement. Certain statistical analyses provide ways to account for these sources of variability. However, these analyses require that repetitive samples be taken such that all known sources of variability are represented in the data. This means that very large sample sizes are required to achieve statistical confidence.

However, physical and logistical limitations may cause these sample sizes to be prohibitive. For example, a preoperational program is restricted in time to that period between the initial planning stage and the operational stage, thus making an extended temporal database impossible. Finite resouces and operating budgets also put ceilings on the extent of the sampling program.

These limitations create a dilemma in the design of a preoperational environmental surveillance program; that is, how should a finite sampling effort be allocated among the various subcomponents of the program while maintaining the statistical validity of each?

In designing the WIPP RBP, this dilemma was addressed in a number of ways. Four of the most important of these are: (1) the use of a strategic sampling scheme which mimics that planned for the ORMP, (2) the use of a hierarchical classification of sampling locations to determine sampling intensity, (3) the use of existing data records to predict adequate sample sizes, and (4) the use of existing records to predict the type of statistical model which best represent the measured parameters.

## STRATEGIC SAMPLING SCHEMES

A pervasive problem in designing an environmental surveillance program is the often large geographic area which must be covered by the program. Most environmental parameters are difficult to characterize on a large geographic scale since, as a general rule, the larger the area included in the sampling, the greater and more complex the variability in the data due to geography. Simple random sampling schemes, which are strongly dependent upon the magnitude of the variance for determining sample size, quickly reach unattainable sample sizes as area increases. This type of sampling scheme would be used where a general characterization of the parameter over a whole region is required, without reference to controlling factors or predictive models.

Environmental surveillance programs are often more specific in their objectives or scope than a general characterization of parameters over a region, *making* a simple random sampling scheme extravagant. Most operational radiological surveillance programs, for example, are designed in accordance with models of radionuclide release and dispersal patterns. These models predict the locations which are most likely to be affected by a release. These locations should be emphasized in the sampling program. Strategic sampling is a sampling scheme whereby sample locations are selected nonrandomly, in accordance with predetermined criteria[1].

For the WIPP ORMP, three principal criteria were used in the selection of sampling sites: (1) maximization of the probability of intercepting radionuclides being transported from the facility by either surface or subsurface media, (2) the inclusion of areas which have a high probability of exhibiting elevated levels of radioactivity not related to WIPP operations, and (3) the inclusion of media which, if contaminated, would bring radionuclides into potential contact with man. Figure 1 shows the distribution of sampling locations which will be used in the WIPP ORMP.

In this program, the WIPP facility is surrounded by two concentric rings of thermoluminescent dosimeters (TLD's), one around the security fence in the eight principal directions from the waste handling area and the second with a 5-mile radius around the site at each of 16 compass directions. The placement of these stations is consistent with the first of the three criteria listed above.

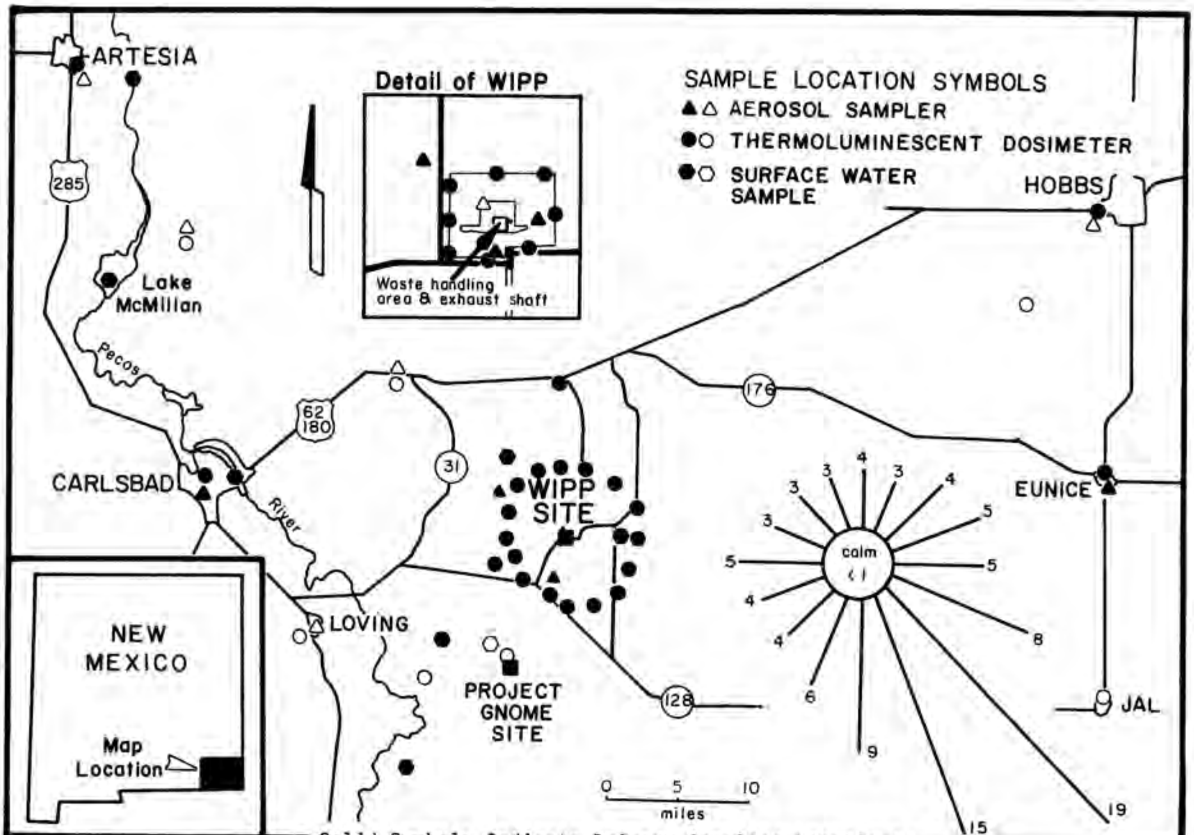
An example of TLD stations satisfying the second criteria are those established in the vicinity of the old Project Gnome site, approximately 9 miles southwest of the WIPP, where an underground nuclear test was performed by the Atomic Energy Commission in 1961. Contaminated particulates were released during this project and were generally carried to the northwest. Although an extensive decontamination program was performed[2], the WIPP environmental programs will continue to monitor this area to protect itself from possible allegations of being the source of the elevated levels of radioactivity at the Gnome site.

Additional TLD stations have been established in each of the principal communities around the site. Besides the data on ionizing radiation levels measured by the TLD's, soil samples will also be taken at each of the TLD locations. These will be analyzed for gross alpha and beta activity as well as concentrations of selected radionuclides.

Three airborne particulate samplers were located near the site in the major downwind directions from the waste handling area and the exhaust shaft of the repository. Nine other aerosol sampler locations were selected which also emphasized the prevailing wind direction from the site as well as nearby ranches and population centers. Samples of vegetation, rabbit, and quail tissues will also be collected from a location in the prevailing wind direction from the site.

Surface water and bottom sediment samples are taken from water bodies, particularly those in the downwind direction from the site, down-drainage from the Project Gnome site, and along the Pecos River, which flows through several population centers. Fish tissues will be collected periodically from the Pecos. Groundwater samples (locations not shown in Fig. 1) will emphasize the probable down-gradient direction (southwest) from the site.

The use of a strategic sampling scheme in the design of an environmental surveillance program allows the monitoring of a large geographic region with a minimum of sampling locations. However, the limitations of the database resulting from this type of scheme must be recognized. Due to the absence of randomization, it would be inappropriate to make estimates of a regional mean or variance. Each location should be treated independently of others.




Solid Symbols Indicate Primary Sampling Locations.
Hollow Symbols Indicate Secondary Locations.

Fig. 1. Map of the WIPP RBP strategic sampling scheme with respect to annual wind patterns. The numbers on the arms of the wind-rose at lower right are percentages of time the wind is from the indicated direction. Groundwater and biotic tissue sampling locations are not shown.

Baseline data collected from one should not be applied to or combined with any other location. Therefore, preoperational sampling performed at each location should be adequate to attain the desired degree of statistical confidence.

The locations selected for sampling in the preoperational program should, therefore, be the same as those which will be used in the operational program. This means that all potential sampling locations for the operational program must be identified prior to the start of the preoperational program. Information on local demography, physiography, and meteorology is usually necessary in making valid decisions on sample locations.

## SAMPLE SITE CLASSIFICATION

Since each site in a strategic sampling scheme will be analyzed independently, its relative importance to the objectives of the operational monitoring program should be carefully evaluated. Since some locations will have a lower potential importance, achieving the same degree of statistical confidence in all databases may not be a desirable goal. The classification of sampling locations according to their relative importance can provide a guide for determining sampling frequency and sample sizes.

In the WIPP RBP, all sampling locations were classified as either primary or secondary (Fig. 1). The primary sites are those within a 5 mile radius of the WIPP, those along the most probable routes of radionuclide transport (i.e., in the downwind or downflow directions), and those in nearby communities and residences. These locations will be monitored on a regular basis during the operational life of the facility. The secondary locations are those whose inclusion in the program is for verification purposes in the ORMP. These include the Project Gnome site, communities which have a low probability of being effected by an accidental release at WIPP, and other miscellaneous locations. These locations will not be monitored on a regular basis during operations, but sampling may be required at some point to verify an absence of WIPP-related effects.

All sites will be sampled at least once during the preoperational program. Primary sites will be sampled continuously or at regular intervals throughout the preoperational sampling period in order to account for as much temporal variability as possible during the three-year program. Supplementary sampling at the secondary sites will be done less regularly, less frequently, or not at all.

## USE OF EXISTING RECORDS

The temporal variability in an environmental parameter is difficult to measure in a preoperational program due to the finite time period within which data can be collected. Often it must be assumed that the time between the start of the data collection and the beginning of the operational phase is adequate to account for long-term temporal variance in the parameters. Verification of this can sometimes be achieved through the analysis of existing records of some of the parameters for the region in question.

The lifespan of the WIPP RBP will be slightly longer than 3 years. This time period should be adequate to give a good representation of the temporal variability in most radiological parameters. To verify this, a careful search was made for existing radiological data records for the area[3]. Although these data records cannot be used to directly supplement the RBP records due to differences in sampling and analytical techniques, useful information can be gained from them about the nature of the parameters being measured in the RBP.

The history of radiological surveillance in the area around the WIPP begins in 1961 with the Project Gnome predetonation measurements of gross beta activity in airborne particulates, tritium levels in water, and selected radioisotope concentrations in cattle tissues. Radiological sampling around the Gnome site was continued as part of the decontamination/decommissioning program following the detonation.

In addition to these records, the Public Health Service, and later the Environmental Protection Agency, has periodically measured radioactivity in air and groundwater samples from selected locations in southeastern New Mexico since 1966 and 1972, respectively. Gross beta activity in aerosol samples and ambient ionizing radiation were measured by Sandia National Laboratories in the vicinity of the WIPP as part of the site validation program. Four years of intermittent high volume aerosol sample data, beginning in 1976, and quarterly TLD data from 1976 to the present have been collected as part of this program.

The existing records were evaluated for their suitability in predicting the behavior of RBP parameters[4]. The length of the record, the age of the record, and the regularity of sampling were used as criteria for this evaluation. Table I shows the evaluation of five historical data records for airborne particulate samples from the vicinity of the WIPP.

An important use of existing data records is the validation of sampling design. For example, sample size predictions can be made for existing data records which meet the criteria for length and regularity. Existing records can also be used to evaluate the geographic variability in certain parameters and to locate possible areas of elevated radioactivity levels which were otherwise unpredicted, but may merit further examination.

## PREDICTION OF STATISTICAL MODELS

When historical data records are of good quality, i.e., they are of considerable length (three or more years), are well documented, and are consistent in sampling intervals and methodologies, more sophisticated statistical analyses can be performed than those described in the preceding section. For example, a prediction of a probability model might be made which will more accurately describe the behavior of the parameter than would be possible with the limited database allowed in the preoperational program.

Several types, or families, of probability distributions are potentially applicable as models of radiological environmental parameters and mathematical methods exist for determining which of these families will best fit a particular data set[5,6,7]. One of these, the method of moments[5], was applied to the existing data records for the region around WIPP. The results are summarized in Table II.

The method of moments predicted a Beta distribution for most of the records analyzed. This family of probability distributions can take a wide variety of shapes, such as a horizontal line (i.e., a uniform distribution), a slanted line, a symmetrical, bell-shaped curve, or an asymmetrical curve (Fig. 2). An important characteristic of the Beta family of probability distributions is that the random variable being modeled must be bounded by an upper and lower

TABLE I

Evaluation of Five Historical Radiological Data Records of Airborne Particulate Samples From Southeastern New Mexico For Use in the WIPP RBP

| Source of record | Period Sampled | Number of Locations | Number of Time Points | Accept/ Reject | Comments |
|---|---|---|---|---|---|
| Project Gnome (predetonation) | 12/2/61 - 12/10/61 | 5 | 5 - 9 | R | Pre-atmospheric test ban, not equivalent to present radiological conditions. Record very short. |
| Project Gnome (postdetonation) | 11/1/62 - 11/30/62 | 6 | 27 - 30 | R | Pre-atmospheric test ban, high variability probably not representative of present conditions. |
| Project Gnome (postdetonation) | 12/1/63 - 12/31/63 | 6 | 19 - 31 | A | Record predicts 18 days/month be sampled to estimate monthly gross beta activity within 0.05 pCi/m$^3$ with 95% statistical confidence. |
| Public Health Service/EPA | 5/21/66 - 12/21/84 | 1 | 231 | R | Gamma scans, not equivalent to analyses used in WIPP programs. |
| WIPP Site Validation Program | 2/76 - 8/79 | 1 | 200 | A | Record predicts 70 weekly samples/4 years required to estimate gross beta activity within 0.05 pCi/m$^3$ with 95% statistical confidence. |

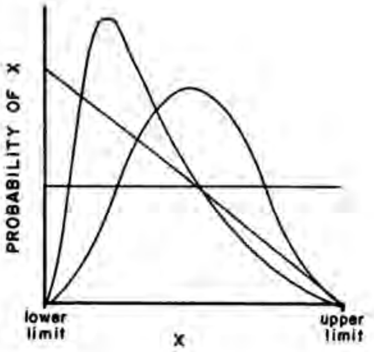



Fig. 2. Four examples of the Beta probability distribution models

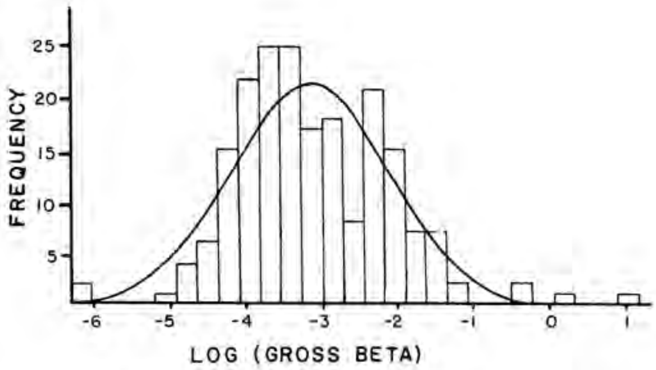



Fig. 3. Fit of log-transformed weekly airborne-particulate data (gross beta activity) to the normal distribution

limit. A likely reason for the frequency with which this model was predicted is that many of the data records analyzed are not long enough to include extreme values, causing the data behaves as if there is an upper limit.

The largest historical data set analyzed is the 4 years (200 weeks) of gross beta activity in airborne particulate samples performed by Sandia National Laboratories from 1976 through 1979. These data exhibited a best fit to the Lognormal distribution, a family of probability distributions which is commonly applied to environmental radiological data[8].

Predicting the probability model for a parameter in the preoperational radiological surveillance program has several advantages. Perhaps most important of these is in providing a basis for data transformation which will allow for more valid application of analytical procedures which specify a particular probability distribution, particularly the Normal distribution. For example, the identification of the Lognormal distribution as a model for gross beta activity in aerosol samples indicates that a simple logarithmic transformation of these data will result in Normally distributed data (Fig. 3).

## SUMMARY

The design of an effective preoperational radiological surveillance program demands careful consideration of the statistical validity of the database. The control of sample size without loss of validity requires a delicate arbitration between the mathematical requirements for validity and the capacity of the program to meet these requirements. This paper has presented and discussed four approaches used in designing the Radiological Baseline Program for the Waste Isolation Pilot Plant by which a partial resolution to this problem has been achieved. These are: (1) the use of a stategic sampling design, (2) the hierarchical classification of sampling locations, (3) the use of existing records to predict sample sizes, and (4) the use of existing records to predict statistical models of some parameters.

TABLE II

Probability Model Predictions From Historical Radiological

Data Records From Southeastern New Mexico

| Medium | Parameter | Number of Records | Percentage of Records Predicting Model | | | | | |
|---|---|---|---|---|---|---|---|---|
| | | | Normal | Lognormal | Beta | Gamma | Other | No model |
| Airborne particulates | beta activity | 26 | 0.0 | 7.7 | 46.2 | 11.5 | 0.0 | 34.6 |
| Air | gamma (HPIC) | 8 | 25.0 | 0.0 | 50.0 | 0.0 | 25.0 | 0.0 |
| Air | gamma (TLD) | 7 | 0.0 | 0.0 | 100.0 | 0.0 | 0.0 | 0.0 |
| Soil | Radionuclide concentrations | 6 | 0.0 | 0.0 | 100.0 | 0.0 | 0.0 | 0.0 |
| Water | Tritium concentrations | 2 | 0.0 | 0.0 | 100.0 | 0.0 | 0.0 | 0.0 |
| Plant tissues | $^{40}K$ concentrations | 5 | 0.0 | 0.0 | 100.0 | 0.0 | 0.0 | 0.0 |
| Bird tissues | $^{40}K$ concentrations | 2 | 0.0 | 0.0 | 100.0 | 0.0 | 0.0 | 0.0 |

These approaches, however, are not applicable to all situations. They are largely dependent upon the existence of an appreciable amount of preliminary data. An adequate database in regional demography, physiography, and meteorology is essential to the identification of strategic sampling locations. These data might not be readily available prior to the initiation of the preoperational program. Furthermore, adequate radiological data records are often rare, difficult to locate, or poorly documented. Deficiencies in these records may necessitate an expanded sampling program to fill in areas of uncertainty.

## REFERENCES


1. J. W. Johnston, "Index of Adequacy and Some Other Statistical Considerations for an Environmental Monitoring System," BNWL-B-250, Battelle, Pacific Northwest Laboratories, Richland, Washington (1973).

2. H. A. Berry, "Gnome Site Decontamination and Decommissioning Project: Radiation Contamination Clearance Report March 28, 1979 - September 23, 1979," DOE/NV/00410-59, Reynolds Electric and Engineering Co., Inc. and EG&G, Inc., Las Vegas, Nevada (1981).

3. E. T. Louderbough, "Compilation and Summary of Preliminary Radiological Data in the Vicinty of the WIPP Site," Draft report, International Technology Corporaton, Albuquerque, New Mexico (1985).

4. N. T. Fischer, "Validation of the Sampling Schedule in the Radiological Baseline Program for the WIPP Project: a Statistical Argument," Draft report, International Technology Corporation, Carlsbad, New Mexico (1986).

5. J. K. Ord, Families of Frequency Distributions, Griffin's Statistical Monographs and Courses, No. 30, Charles Griffin and Co., London (1972).

6. S. S. Shapiro and A. J. Gross, Statistical Modeling Techniques, Marcel Dekker, Inc., New York (1981).

7. R. A. Tapia and J. A. Thompson, Nonparametric Probability Density Estimations, Johns Hopkins University Press, Baltimore, Maryland (1978).

8. D. R. Speer and D. A. White, Statistical Distributions as Applied to Environmental Surveillance Data," BNW-SA-5482, Battelle, Pacific Northwest Laboratories, Richland, Washington (1975).



Work supported by the U.S. Department of Energy Assistant Secretary for Defense Programs, under DOE Contract No. DE-AC04-86AL31950.