

Statistical Graphics in Environmental Regulatory Studies

Larry R. Holden, Monsanto Company

ABSTRACT

Agricultural companies are required to conduct large numbers of environmental studies to evaluate the potential impact of their products on the environment. The results of these studies are typically reported in tabular form. Rarely are graphical or statistical analyses used to examine the data beyond the regulatory requirements. During the past few years, Monsanto has increased the use of graphical materials both in these submitted reports and in the conduct of the studies themselves. These efforts have resulted in an increased understanding of the data both by the researchers and the regulatory agencies. Most success has been found with simple scatterplot and dotplot variants. For display of mean comparisons, the significance interval technique of Andrews, et al. (1980) has proven very useful. All of these products can be created using commonly available graphics software.

INTRODUCTION

The Agricultural Group of Monsanto Company produces a diverse array of chemical and biological products aimed at the agricultural industry. These products include herbicides and fungicides as well as crops that have been improved through genetic engineering. The Environmental Protection Agency (EPA), the Food and Drug Administration (FDA), and the U.S. Department of Agriculture (USDA) each have some degree of responsibility for the approval and regulation of these agricultural products. Before granting approval for a new agricultural product, these regulatory agencies require many studies on product efficacy, safety, and environmental fate. On a case-by-case basis, specialized environmental studies may even be required for products currently on the market. The collection of reports summarizing these study results is called the submission or registration package. The agencies review this package and, if it supports the safety of the product, approves the product for commercial use. Any inadequacies in the research or in the reports themselves result in either a request for additional information or an outright rejection. Similar processes hold for evaluating new studies on existing products. Here rejection of a study can mean the removal of the product from the marketplace.

Since most of these environmental regulatory studies must follow published guidelines, the resulting reports have a standardized appearance. In addition, because a report rejection for incomplete data can severely delay approval of the new product, regulatory reports emphasize documentation over interpretation. Identifying characteristics of such documents are a minimal amount of interpretive text, a large number of tables, and few or no

graphical summaries. Regulatory agencies must find it unpleasant to review such tabular behemoths. Since it would be naive to expect a trend towards decreasing the amount of data submitted, there is a need for more summarizing and clarifying of the study results. During the last few years, Monsanto has attempted to meet this need by incorporating more graphical summaries into these reports and the entire regulatory research process itself. In a nutshell, this often means replacing, as much as possible, summary tables by figures.

In the environmental regulatory area, I have found no single graphical method that is optimal in every situation. The best approach to visual summarization depends upon how the display is going to be used. Is the figure to be in a report or in a presentation? In general, a graph that can be studied at leisure can tolerate a greater load of detail than can one being flashed on the screen for a limited time. Who is the intended audience? For internal use the purpose of the graph may be to show the present state of affairs—emphasizing both good and bad. For external usage the figure often serves to emphasize a point of view or a body of evidence pointing to a conclusion. Finally, is the graphical information intended for personal use? A graph used only by the statistician or researcher may have a tremendous complexity load and still be valuable. An audience member or report reviewer, however, may not have the knowledge base or the patience to deal with a high degree of visual complexity.

In this paper, I will cover three categories of graphical use that have proved very useful in our research and regulatory environment: (1) displaying distributions of replicate data values; (2) displaying structured data values, and (3) graphical displays for comparing means. In most cases, these graphs are designed for the regulatory reviewer or for a general audience, not the statistician. They can be used internally or externally and are suitable for either reports or presentations.

REPLICATE VALUES

The most useful type of plot I have found for displaying replicate measurements is the dot diagram or "dotplot" popularized by Box, Hunter, and Hunter (1978). In the dotplot (illustrated in Figure 1), each value is indicated by a single dot along a horizontal scale. Values that would normally overlap are represented by stacked dots. While similar to the histogram, the dotplot has a single important advantage—each data point is distinct. The dotplot provides a clear and immediate visual image of the number of data points represented. A histogram may provide a better summary of the relative distribution of a very large number of values, but for the sample sizes more common

in environmental work, the dotplot is preferable. It has been my experience, that researchers can rapidly grasp the point of a dotplot, whereas some orientation time is usually required for the same data displayed as a histogram.

The dotplot is especially useful for displaying censored values, a common feature of environmental studies. Frequently, chemical concentrations in study samples are too small to distinguish from random noise in the measurement device. Such environmental samples are classified as "not detected" by most chemists and are treated differently from those giving actual numbers. Non-detected "values" are easily represented on a separate position to the left of the dotplot's numerical scale. Figure 2 uses this dotplot format to display herbicide concentrations obtained from an analysis of soil samples.

To my knowledge, MINITAB is the only software that can directly generate this type of dotplot from a set of data. Although adequate for exploratory data analysis, the dotplot in MINITAB is in "fine printer" format and is a far cry from presentation quality. Hopefully, the SAS Institute and other vendors will add this capability to future releases of their graphic products. In the interim, however, dotplots can be constructed using any graphics package that can produce a simple X-Y scatterplot. To create a dotplot, first sort the replicate data in ascending order. Each value is then rounded to some interval, R . R determines how close two distinct values need to be before their dots are stacked. These rounded numbers become the X values on the scatterplot. For unique X values, the Y value is always 1. For multiple X values, assign the first $Y=1$, the second $Y=2$, and so on. Use any favorite graphics program to create the X-Y plot. The dotplot is completed by eliminating the Y-axis line and altering both axis lengths, as necessary, to move adjacent dots together. These steps can be programmed using the SAS® software macro facility or, more conveniently, use any of the common spreadsheet programs such as Microsoft EXCEL or LOTUS 1-2-3. I have found that the DeltaGraph Professional scatterplot feature can be used to generate high quality dotplots. For many applications, the charts created by a spreadsheet program itself may be of sufficient quality to eliminate the need for any further graphics step. When faced with generating a large number of similar dotplots, however, it may be worthwhile to consider using the SAS macro language and SAS/GRAPH® software.

While the dotplot is useful for a single batch of data, its value is multiplied when two or more batches must be compared. The stacked dotplot is probably the most productive visual tool I have found for this purpose. Shown in Figure 3 is a stacked plot contrasting the effects of different herbicide treatments on the amount of herbicide residue found in corn forage. Clearly, higher residues are associated with greater herbicide application amounts and later application stages. Stacked dotplots can also be very powerful tools for demonstrating the effects of treatment groups compared with a control,

especially if the treatment alters the variance of the data. Figure 3 also illustrates the use of a logarithmically scaled dotplot. The procedure for constructing a log-scale dotplot is straightforward: The data are first converted to logs, rounded, stacked, and then back-transformed to obtain the X-values on the original scale. The graphics package is then used to plot the data on a logarithmic X-axis. If the graphics software does not have a logarithmic axis feature, then the back-transformation step can be eliminated.

DATA WITH STRUCTURE

There are many environmental studies that do not result in a set of independent replicated samples. A commonly encountered situation is the existence of multiple treatments in the same experimental unit or "block". In environment regulatory studies, it is the location of the experiment that most often represents a block. The between-location variation imparted by the dotplot cannot be used as a reference to assess treatment differences. If only two treatments need to be contrasted, a simple solution is to display treatment differences (or ratios) within each location in the dotplot. For more than two treatments, however, other graphical approaches must be used. For simple impact, a 3-dimensional block chart showing an environmental variable plotted against both location (X) and treatment (Y) can be a very powerful visual device. Another common method for blocked data is the simple lineplot. Here, the blocks (e.g. locations) are arranged along the X-axis. The environmental measurement for each treatment is plotted on the Y-axis. To more easily identify treatment effects, points for the same treatment are connected across blocks by line segments. Such a lineplot is shown in Figure 4 illustrating the analysis of carbohydrate content in several genetically modified plant lines from nine different locations. Although there is a great deal of variation among locations, it is clear from this plot that Line A has a smaller carbohydrate content than Line C. The carbohydrate content in Line B is intermediate between A and C.

The only weakness of the lineplot is that it becomes very chaotic if the number of treatments, and hence line segments, gets too large. A simple alternative which I have found useful in some instances is the repeated scatterplot. Figure 5 presents the same carbohydrate data in a repeated scatterplot. Here, the same treatments in different blocks are not superimposed as in the line plot but are repeated along the X-axis. Now any repetitive pattern among the treatments is easier to discern. For visual impact, the order of the treatments within each location was arranged so that the mean increases from left to right. Such plots reduce the ability to make detailed comparisons of individual values separated along the x-axis. But if detection of pattern is the primary goal, which it often is in environmental regulatory research, then the loss is minor. As also shown in Figure 5, the average over blocks can be easily added to the plot as an extra "block".

Repeated scatterplots have also proved valuable as research tools in Monsanto's environmental studies. For example, in a nation-wide survey of rural domestic wells (Holden, et al., 1992), Monsanto wished to detect very small amounts of herbicides in well water. A repeated scatterplot was used to examine each batch of well water chemical results to help determine if a chemical was present. A sample plot of the type used in this study is shown in Figure 6. Each pair of dots represents the mean responses obtained for each well. The closed dots indicate results for one method of measuring the concentration in well water and the open dots for another method. A real herbicide effect would show up as a response in both methods. The vertical lines represent the range from analysis of two different aliquots of the same well water. A true contaminant in the well water might be displayed as shown in wells 1 and 6. The majority of the wells and all control samples will have no analyte present and appear as shown for wells 2-4. Occasionally, the measurement technique experiences interferences or other problems. For example, wells 5 and 7 illustrate false "positive" results that are quickly identified on the repeated scatterplot. Use of this graphical summarization method enabled Monsanto to effectively and quickly identify potential wells with the presence of herbicides. The EPA has cited this and similar graphical approaches for their ability to identify very low levels of contamination with no increases in the number of false positives.

DISPLAYING MEANS

Displays of individual data points, no matter how creative, cannot be used in every situation. Often there is a need to display and compare average values. Plots of means typically include "error bars" (representing standard errors or confidence intervals) as vertical lines above and below the symbol for each mean. A very large number of people mistakenly believe that if the bars from different means overlap, the means are not statistically different. At the conventional 5% level, using the overlap of standard error bars as a significance test gives too many significant results and the overlap of confidence intervals gives too few. One obvious solution is to incorporate some idea of the statistical significance of means on the visual graphic. For the case of a simple one-way experimental structure, the JMP® software uses a graphical device called comparison circles to show significant differences between means. Comparison circles, when aligned vertically with their centers at the mean value will intersect with an outside angle less than 90° when the means are statistically significant (Sall, 1992). When the comparison circles fail to overlap, statistical significance is clear. When they do overlap, however, it is very difficult for humans to judge a 90° angle between intersecting circles. Also when there are many means displayed, the circles become very jumbled and impossible to distinguish. In the JMP software, the user has the ability to manually select a circle and obtain a color indication of significance. This brings some order to the chaos on the

computer monitor, but is not very appealing for publication or presentation quality displays. A more visually appealing approach is the use of significance intervals proposed by (Andrews, et al., 1980). A significance interval (SI) is the interval defined by $\text{mean} \pm H$. Means whose SI overlap are not statistically significant. For samples of equal size n , the half-length of the SI, H , for any given mean is given by

$$H = (t/2)S\sqrt{(2/n)}$$

where t is the critical value for any statistical test and S is the pooled within-batch standard deviation. If the critical Student's t -value is used, then the SI is a graphical version of the common Least Significant Difference (LSD) procedure (e.g. Steele and Torrie, 1980). Mean quality scores established for fruit from five genetically modified tomato lines are compared in Figure 7 using the SI plot. The display of each SI as a rectangular "window" is preferred by Andrews et al. as a mechanism to focus the viewer's attention on the interval rather than the mean. Conveniently, SAS/GRAPH software has an option on the SYMBOL statement that will generate these windows, whereas most other programs do not. Most graphics packages can, however, display the SI as "error bars".

The SI concept works for most types of experimental structure. For example, if the treatments were used in a randomized block structure, the only adjustment would be to employ the correct residual error and the proper degrees of freedom. The real difficulty with SI plots comes when the samples sizes are unequal. (This is not a problem with the comparison circles in JMP software—they are exact for any combination of sample sizes.) Clearly, in the unbalanced case, the intervals must have different lengths to adjust for different samples sizes. While the length of H can be derived easily for any pair of means, it will differ with each particular pair of means being compared. Any solution to this problem will be only approximate. Fortunately, the approximation usually seems to be adequate within the visual resolution of the plot. In general, H , is constructed by approximating the standard error of a difference between pairs of means, $D_{ij} = (M_i - M_j)$, by summing a function of each individual standard error. That is:

$$se(D_{ij}) = F[se(M_i)] + F[se(M_j)]$$

Then the SI for mean i , $M_i \pm H_i$, can be constructed from:

$$H_i = t F[se(M_i)]$$

A good solution for F seems to be a quadratic function:

$$F(x) = A + Bx + Cx^2$$

Then A , B , and C can be found from a fit to the equation:

$$se(D_{ij}) = A + B se(M_i) + C se(M_i)^2 + A + B se(M_j) + C se(M_j)^2$$

Rewriting this as:

$$se(D_{ij}) = 2A + B[se(M_i) + se(M_j)] \\ + C[se(M_i) + se(M_j)]^2$$

gives the form of a simple quadratic regression of all pairs of means on the sum of standard errors, $se(M_i) + se(M_j)$. Using the SAS software, a macro can be developed to find A, B, and C. For simple designs, I have also found it easy to perform the calculations on a spreadsheet. For more complex designs, the SAS GLM procedure can be used to generate the $se(M_i)$ and $se(D_{ij})$ values necessary to compute each SI. An unequal sample size SI plot comparing effects of six herbicide treatments on corn grain residues is shown in Figure 8. This plot also makes it easy to explain how two treatments (EP@0.5 and EP@2) may be significantly different while those farther apart (EP@0.5 and LP@0.5) might not. SI plots are extremely useful for generating visual summaries of environmental science data. As important, they are very intuitive since many people already see intervals as significance tests.

OTHER GRAPHICS

There are many innovative methods for displaying visual information that can improve the quality and acceptability of environmental studies. For instance, simple diagrams of a corn plant at various stages of its development could be placed beneath the time axis of a concentration versus time plot. This gives the viewer an immediate feel for the physiological stage at which the data values were obtained. At Monsanto, the use of a plot as a background display has also proven an effective technique. Figure 9, for example, is a simple plot of some common pesticides on two environmental characteristics. As discussed by Gustafson (1989), these two characteristics, persistence and mobility of the chemical in the soil, are considered critical factors affecting the potential to contaminate groundwater (i.e. "leach"). The pesticides on this plot are given different symbols depending upon their classification as leacher or non-leacher by the California Department of Food and Agriculture. This plot can now serve as a reference background for comparing new or other compounds. The diagonally lined rectangle in Figure 9 shows the possible ranges of persistence and mobility for a new compound. Plots like this can be very powerful in a presentation.

CONCLUSIONS

I have illustrated how some rather simple types of graphical displays can be used to improve the presentation of results from environmental research in the regulatory arena. For single or multiple batches of replicated data, dotplots and stacked dotplots are very flexible and immediately interpretable. They can be modified easily to include censored data. For experiments with more structure, the use of lineplots and repeated

scatterplots to display blocks of results is useful. These plots are excellent for showing consistent patterns across structures. The use of significance interval plots is recommended for situations in which comparison of means is the primary purpose. These intervals can be developed for most experimental designs and can be used to compare means with differing sample sizes. Greater and more effective use of such graphical techniques will vastly improve the clarity and acceptability of environmental studies performed for regulatory approval.

While all the plots discussed here can be produced with available software, I would encourage graphics vendors such as SAS Institute to add dotplots to their graphical arsenal. I feel the only limitation to their widespread usage is the lack of a convenient way to generate them in a publication quality format.

REFERENCES

- Andrews, H. P., R. D. Snee, and M. H. Sarnier (1980), "Graphical Display of Means," *The American Statistician*, 34:4, 195-199.
- Box, G. E. P., W. G. Hunter, and J. Stuart Hunter (1978), *Statistics for Experimenters*, New York: Wiley.
- Gustafson, D. I. (1989), "Groundwater Ubiquity Score: A Simple Method for Assessing Pesticide Leachability," *Environmental Toxicology and Chemistry*, 8, 339-357.
- Holden, L. R., J. A. Graham, R. W. Whitmore, W. J. Alexander, R. W. Pratt, S. K. Liddle, and L. L. Piper (1992), "Results of the National Alachlor Well Water Survey," *Environmental Science and Technology*, 26:5, 935-943.
- Sall, J. (1992), "Graphical Comparison of Means," *Statistical Computing and Statistical Graphics Newsletter*, 3:1, 27-32.
- Steel, R. G. D., and J. H. Torrie (1980), *Principles and Procedures of Statistics*, New York: McGraw-Hill.

For Further Information Contact:

Larry R. Holden
Monsanto Company
700 Chesterfield Parkway North
St. Louis, MO 63198
Phone: (314) 537-6885
E-Mail: LRHOLD@CCMAIL.MONSANTO.COM

SAS, SAS/GRAPH, and JMP are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. ® indicates USA registration.

Other brand and product names are registered trademarks or trademarks of their respective companies.



Figure 1. Dotplot showing the distribution of a chemical in a set of soil samples.

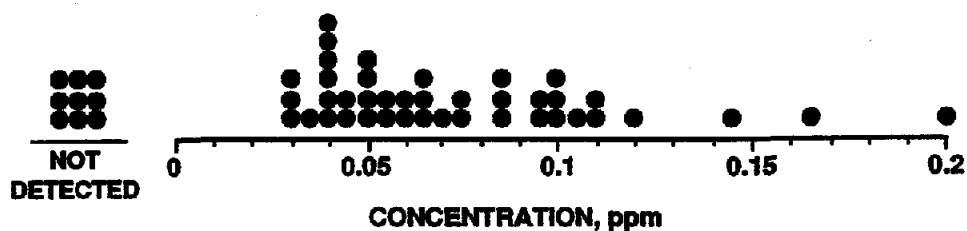


Figure 2. A dotplot of chemical concentrations found in soil samples. A separate category has been added indicating samples in which the chemical was not detected.

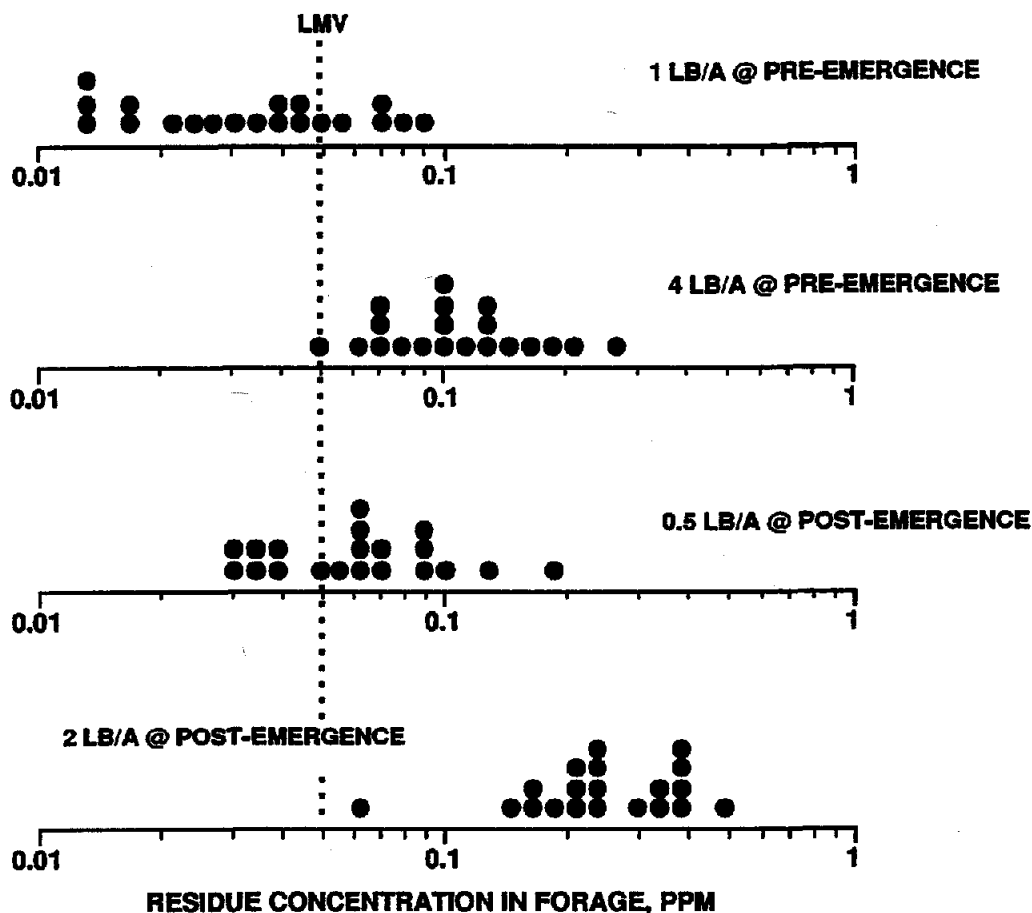


Figure 3. A stacked dotplot comparing the chemical residue found in corn forage resulting from application at different stages and rates. Each dot represents the results found at a single location. The LMV is the smallest concentration for which the analysis method has been validated.

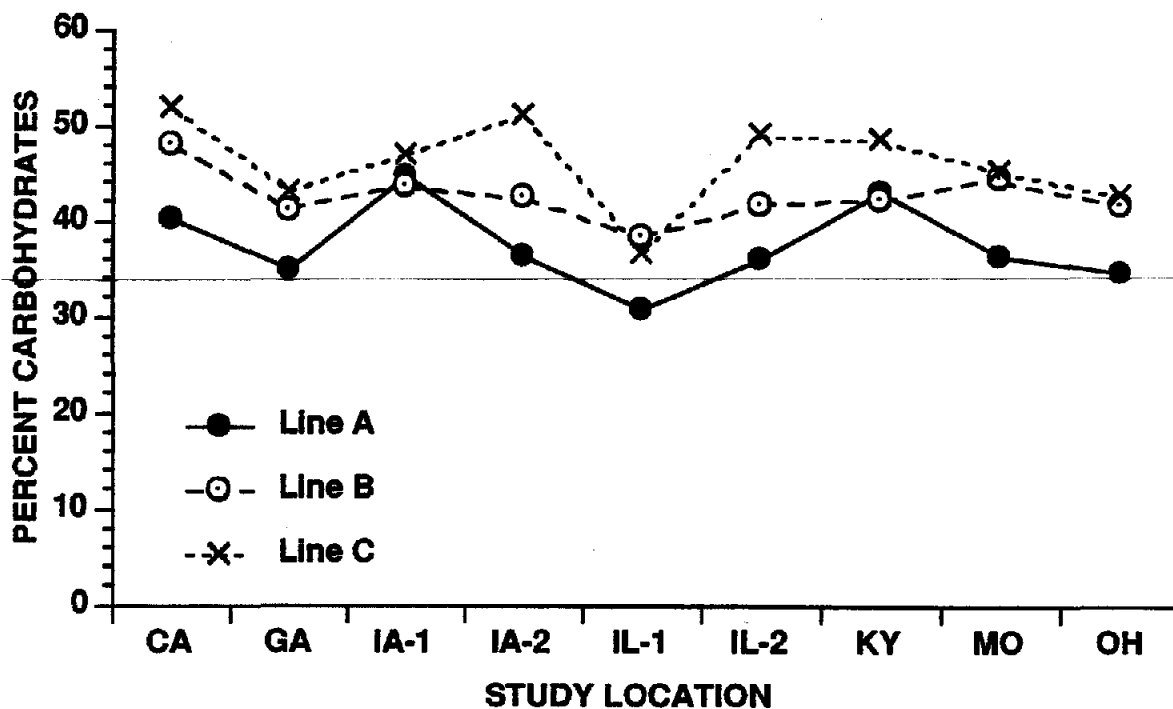


Figure 4. A lineplot showing the carbohydrate content in seed from three genetically modified plant lines from nine different field studies.

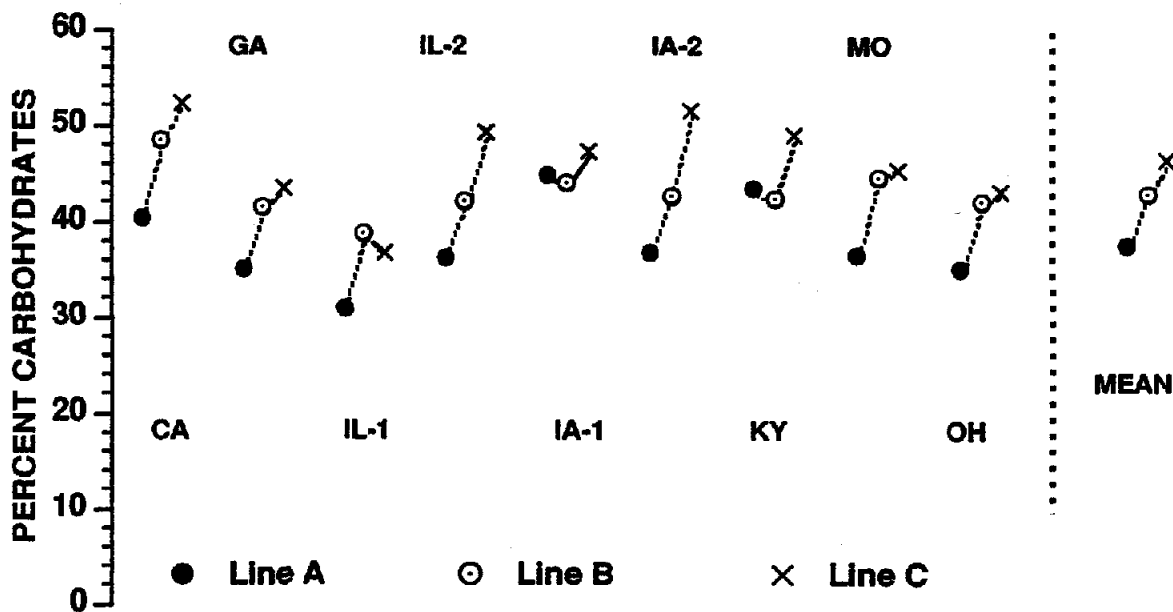


Figure 5. A repeated scatterplot of the same data displayed in Figure 4.

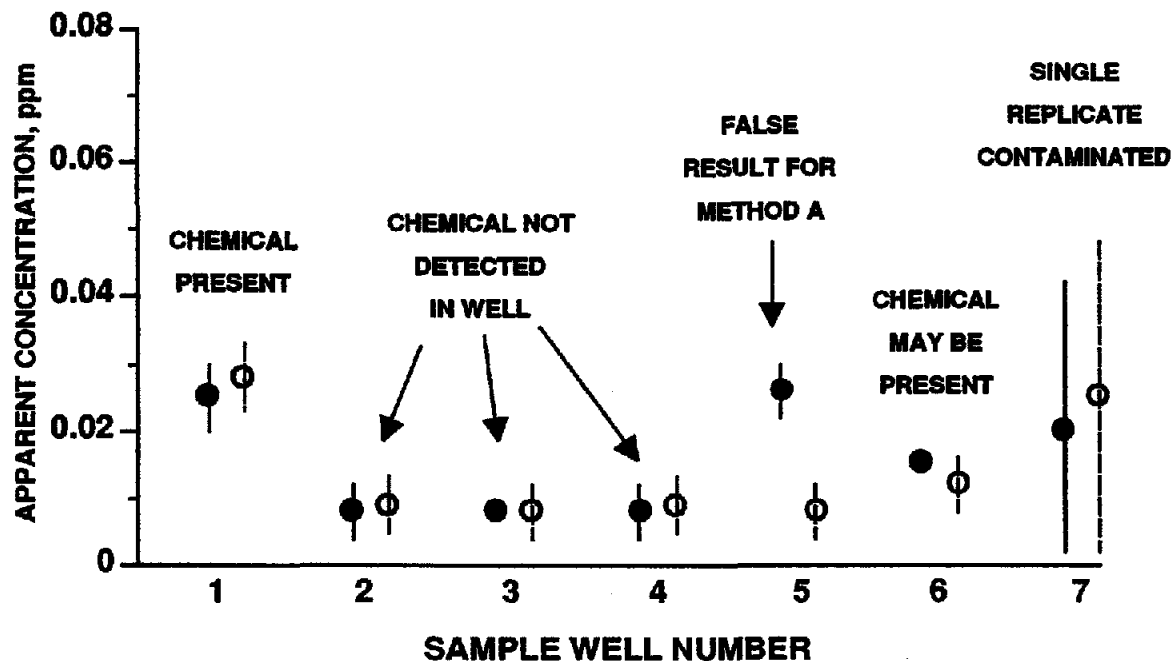


Figure 6. A more complicated repeated scatterplot used to display results from a chemical analysis of drinking water wells. The two dot symbols represent different ways to measure the herbicide. Vertical lines show the range obtained for two portions of the same well water.

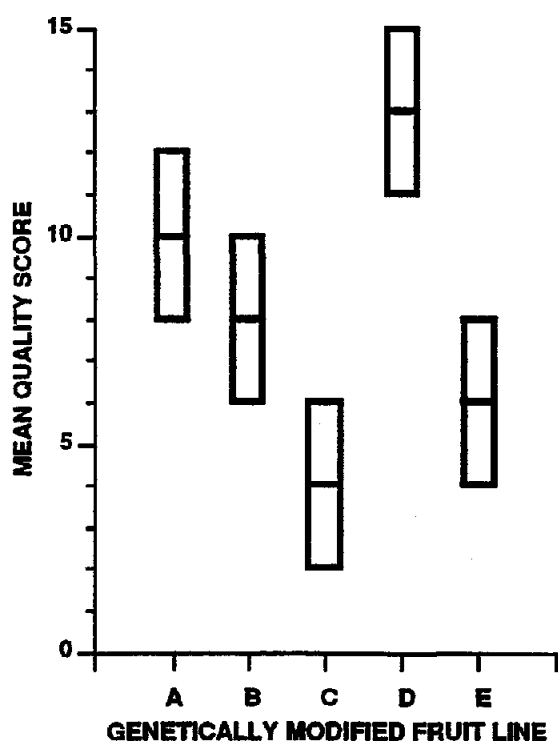


Figure 7. Significance interval plot used to compare the mean quality score obtained for tests of genetically modified fruit.

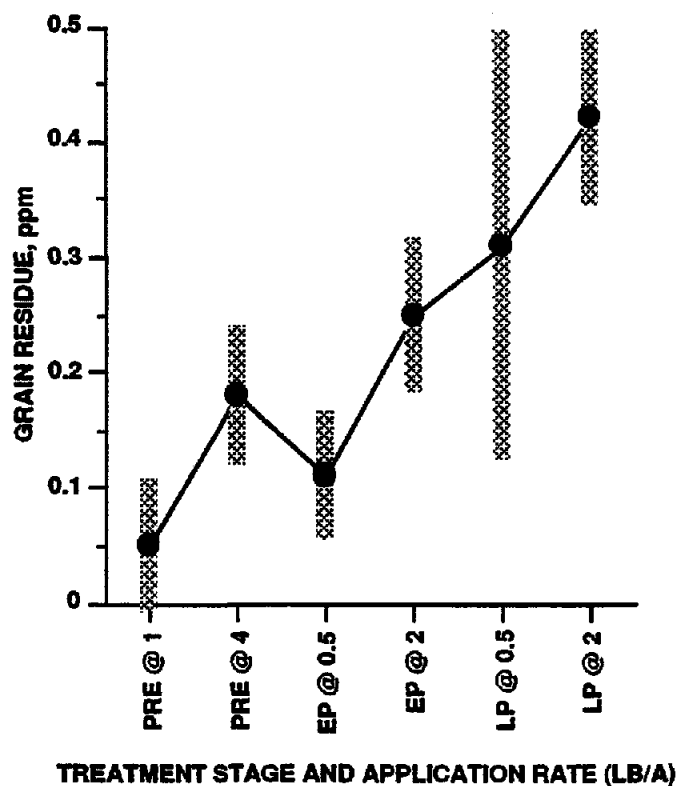
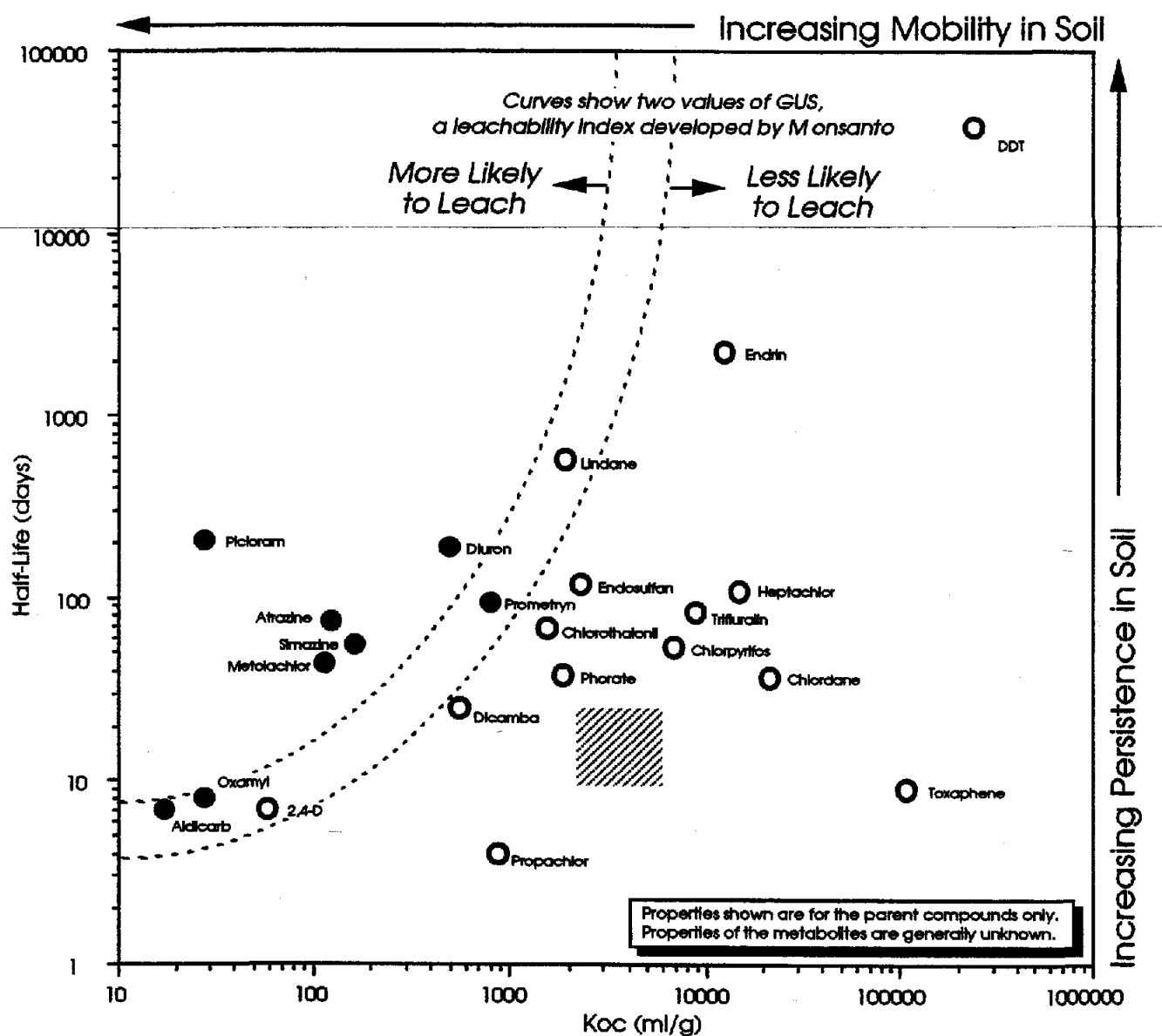


Figure 8. Significance interval plot comparing mean chemical residues in corn grain following different chemical treatments.

Relative Leachability of: X



- Known leachers according to California Dept. of Food and Ag. (CDFA)
- Known non-leachers according to California Dept. of Food and Ag. (CDFA)

 Size of rectangle reflects uncertainty in the values for the two properties, with position based on the available environmental fate data.

Figure 9. Example of a graph used as a reference background. The diagonally lined rectangle indicates the possible position of a chemical of interest relative to that of other compounds that have been classified as to their occurrence in ground water. In this example the compound does not appear to have the potential to contaminate groundwater.