Editors: Lloyd Treinish and Deborah Silver

Worm Plots

Geoffrey Matthews

Western Washington University

Mike Roze

Applied Software Technology

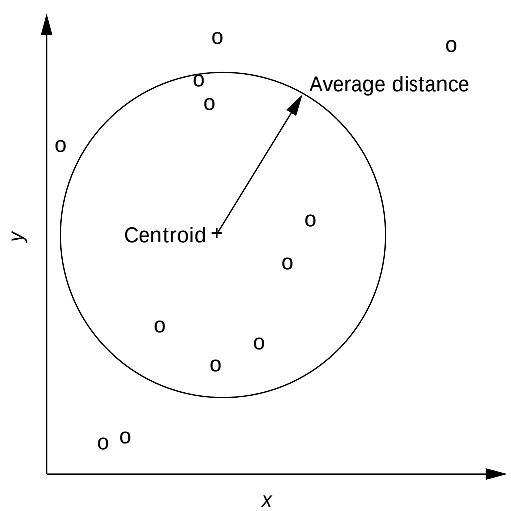
Scientists sometimes spend hours conducting experiments only to find that the resulting data proves difficult to analyze with traditional visualization techniques. A typical laboratory experiment, for example, will set up several systems in each of two or more groups—a control group and various treatment groups.

The investigator will then measure various parameters for each system over time (sometimes dozens to hundreds of parameters). The investigator tries to answer this question: What is different between the control and treatment groups? Field monitoring of polluted and unpolluted sites within a region result in the same kinds of data and the same difficulties in visualization. Plotting multiple lines on a single 2D plot quickly gets confusing. Plus, the dynamic interactions between parameters can be hard to see. 3D visualizations help researchers make qualitative insights about data more easily.

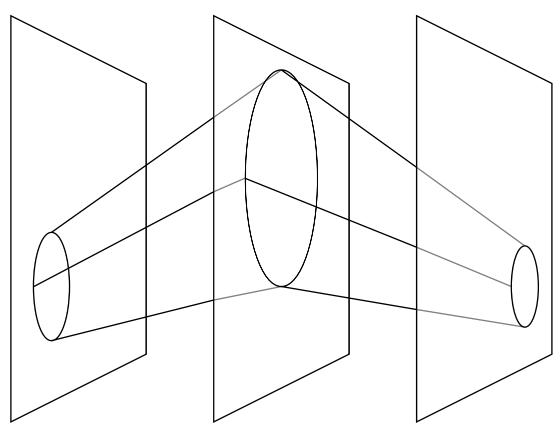
We developed a 3D plotting technique, called "worm plots," for visualizing these kinds of data. To see how they work, take a look at the 2D plot in Figure 1, which plots 12 points on two axes. We can summarize the positions of all 12 points with a 2D shape, in this case a circle. The center of the circle is located at the data's centroid. The radius of the circle equals the average distance of the points from the centroid. In this way, we summarize the central tendency of a group of points and their variability. If there were two or three groups of data points plotted—as when two or three treatment groups must be analyzed—we would plot a different circle for each group of points.

Our visualization tool examines how groups of points change over time. If we plot circles for each time slice, then these circles will change their relative sizes and positions as time goes by. To visualize these evolving dynamics, we connect these time slices with conic sections, as shown in Figure 2. This gives us spacetime worms that can be displayed graphically.

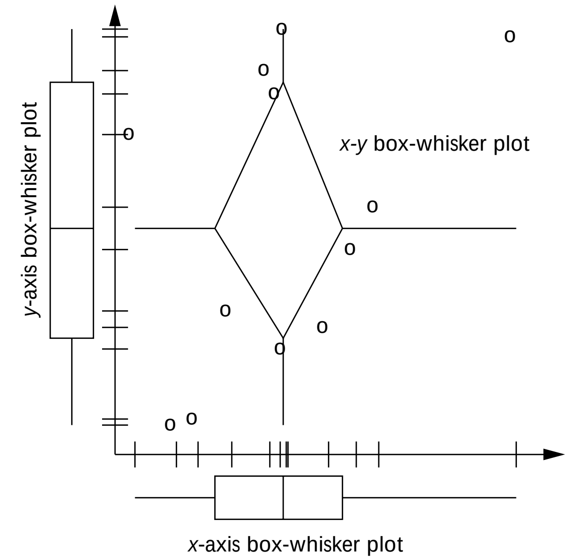
In our experience, these simple circular cross-sections work best when exploring data. They can be rendered quickly and give an easily intuited summary of each group. Further enhancements tend to make the image busy and difficult to interpret. However, other kinds of cross-sections can provide more detailed information. For example, we generalized Tukey's popular "box-whisker" plot for one dimension to two dimensions. Tukey recommended summarizing the interquartile range (the middle 50 percent) of a group of points with a box, having a division at the median, and adding whiskers to the box that extend the data's full range. Figure 3 (next page) illustrates two box-whisker plots for the same data as in Figure 1. Each point projects onto each axis, and the two one-dimensional data sets are summarized in the adjacent



1 A 2D scatterplot of 12 points at a single moment in time. The position and variability of the group of points is summarized by a circle centered at the group mean and with a radius equal to the group's average distance from the center.



2 The circular cross-sections summarizing a group's response are ordered by sampling period and connected with conic sections to form a spacetime worm.



3 A different kind of cross-section for worms. Two 1D box-whisker plots at right angles combine into a diamond-whisker shape. Note that only one point ends up inside the diamond.

box-whisker plot. Two such box-whisker plots at right angles combine into a single "diamond-whisker" plot. Note that in this example only one of the data points actually falls within the diamond. Unlike one-dimensional box-whisker plots (in which the box is guaranteed to contain 50 percent of the points), we can't guarantee that the diamond will contain any of the points.

The simplest plots, on the other hand, do not attempt any group summarization at all. If a plot doesn't contain too many points, each point can be treated as its own group and plotted as its own thin worm.

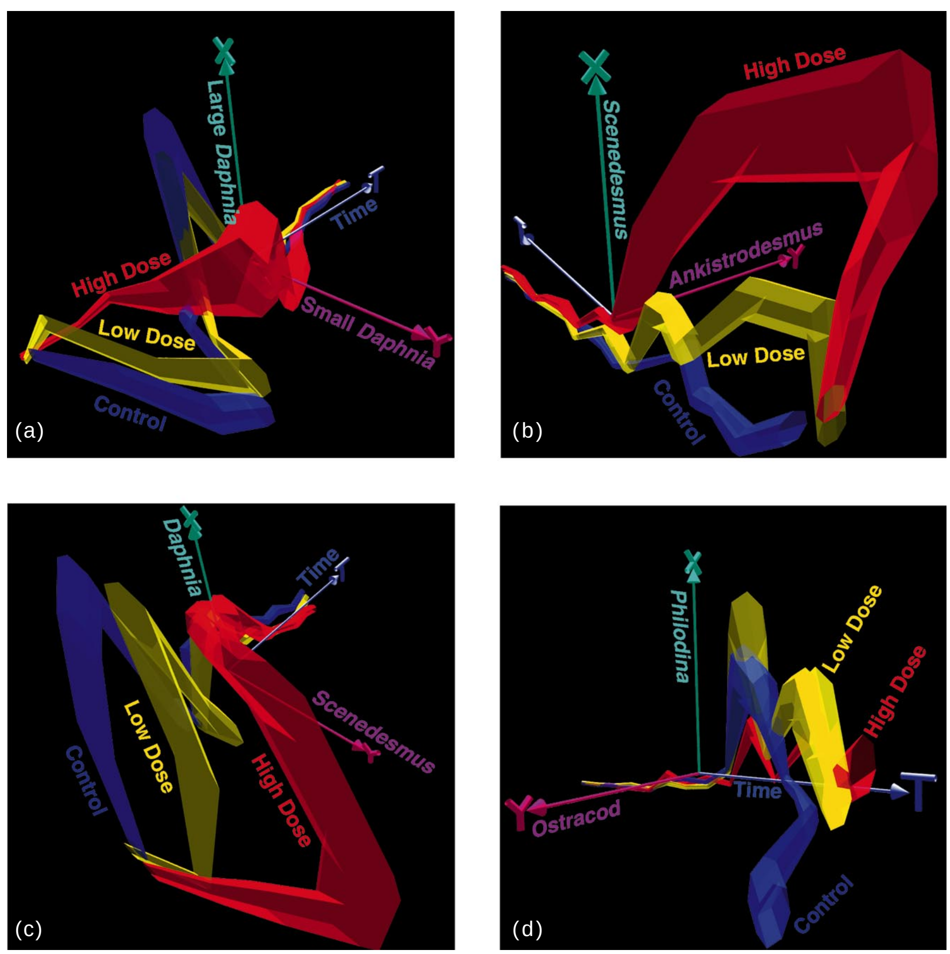
Aquatic microcosm experiments

Figures 4 and 5 illustrate worm plots of data from standardized aquatic microcosm experiments. These toxicology experiments involved several pickle-jar-sized containers of water that were innoculated with species of algae and other microorganisms, dosed in groups, and monitored for several months. The microcosm provided an excellent middle ground for toxicology testing, avoiding the huge expense of field testing and the simplistic approach of single-species testing. The small communities exhibited many of the characteristics of larger communities, but were reasonably inexpensive to maintain in a laboratory setting.

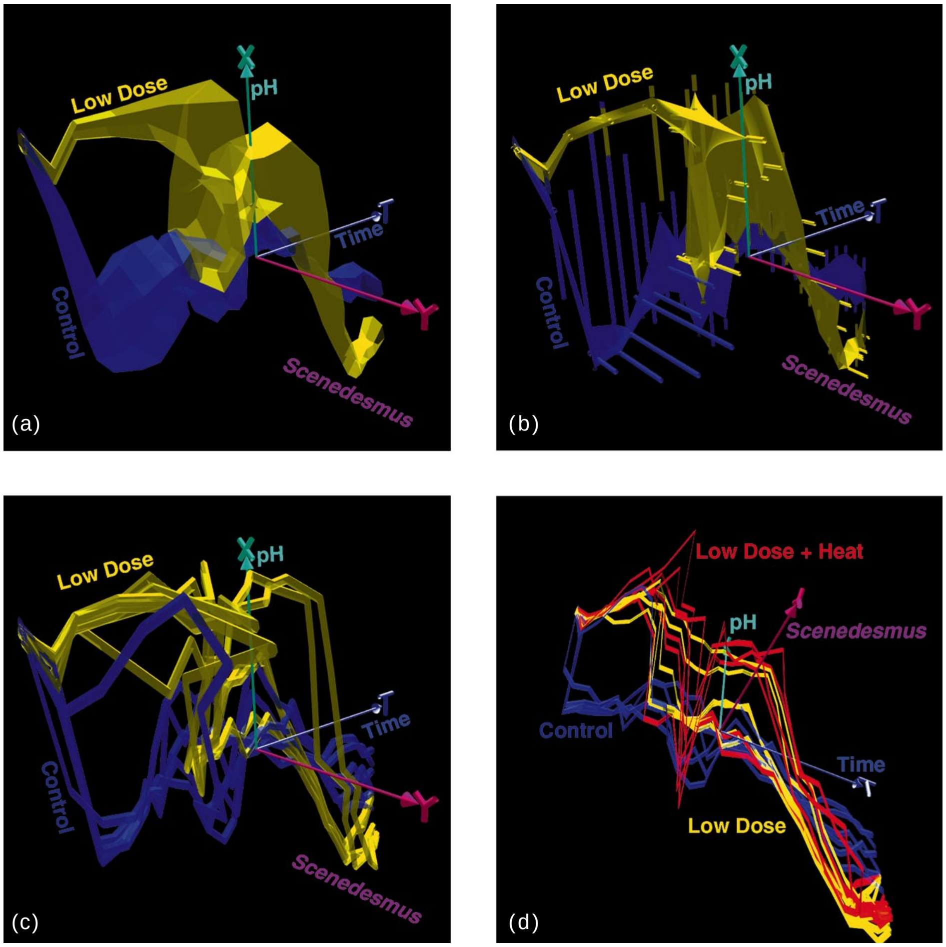
Effects of Jet-A fuel

Figure 4 illustrates some of the effects of the jet fuel Jet-A in zero, low, and high doses (blue, yellow, and red, respectively) in microcosms over a time period of about two months. Figure 4a shows a plot of the species *Daphnia*, the predominant algae-consumer in the microcosms.

The *x*-axis contains plotted numbers of large *Daphnia*, and the *y*-axis has plotted numbers of small *Daphnia*. In the control (blue) group, you can clearly see a shift in



4 Worms from a standardized aquatic microcosm experiment testing the effects of Jet-A jet fuel on aquatic organisms. Control group is in blue, medium dose of jet fuel in yellow, and high dose in red. (a) Large *Daphnia* versus small *Daphnia*, (b) *Scenedesmus* versus *Ankistrodesmus*, (c) small *Daphnia* versus *Scenedesmus*, and (d) *Philodina* versus *Ostracod*. To see an animation of Figure 4c, visit http://computer.org/cga/cg1997/extras/g60174c.mpg or g6017.avi.



5 **Worms from an extended microcosm experiment testing the combined effects of JP-8 jet fuel and heat, with pH on the *x*-axis and *Scenedesmus* on the *y*-axis. Control group is in blue, dosed group in yellow, dosed and heated group in red. (a) Using circular cross-sections, (b) using diamond-whisker cross-sections, (c) points shown individually without summarizing groups, and (d) individual points with dosed-heated (red) group added. To see an animation of Figure 5d, visit http://computer.org/cga/cg1997/extras/g60175d.mpg or g60175d.avi.**

population from small to large, occurring about one-third of the way through the experiment. In the high-dose (red) group, you can see that the population does not grow as well, and that a mixed group of small and large organisms emerges. The low-dose (yellow) group reveals an intermediate reaction. Although there was an early suppression of *Daphnia* growth, in the experiment's later stages the numbers of *Daphnia* essentially mimicked those in the control group. An interesting phenomenon occurs when the control and medium-dose groups pass through the high-dose group. Statistically, we would expect no differences among groups at this point in time, but the worms reveal that the static numbers are not nearly as important as the dynamics of the populations.

Figure 4b shows two species of algae—*Scenedesmus* (*x*-axis) versus *Ankistrodesmus* (*y*-axis)—from the same experiment. In the control (blue) group, algae populations remain low, due to the high numbers of *Daphnia* present. The high-dose (red) group shows the rapid growth of algae, but also shows a dynamic transition from the faster growing species (*Ankistrodesmus*) to the slower growing species (*Scenedesmus*).

Figure 4c shows small *Daphnia* (an algae-consumer) on the *x*-axis and *Scenedesmus* (an algae) on the *y*-axis. The spiral shape here is due to a classic predator-prey cycle. In the control (blue) group, the *Daphnia* show a boom-and-bust pattern. They quickly boom to large populations, consuming all the available algae. Then a population crash occurs, followed by another (smaller) boom, and so on. In the high-dose (red) group, the *Daphnia* have been decimated by the toxin, while the algae population booms. Some time later, the hardy *Daphnia* that survived the toxic insult begin to grow and simultaneously reduce the size of the algae population. This interaction between primary and secondary effects of the toxin creates the spiral in the red worm. The medium-dose (yellow) group shows a response intermediate between the control and high doses.

In Figures 4a through 4c, you can see another pattern for all the species plotted: they die off to small numbers by the end of the experiment. Traditionally this has been taken as evidence of system recovery after the toxic insult. However, Figure 4d plots two species from a different ecological niche. *Philodina* (*x*-axis) and *Ostracods* (*y*-axis) are detritivore species—organisms that survive by consuming dead and decaying matter. In a microcosm, little or no detritus exists at the beginning of the experiment. By the end of two months a sizable amount has accumulated and the detritivore species populations can begin to grow. Interestingly, one of them (*Ostracod*) grows better in the control (blue) group, while the other (*Philodina*) grows better in the low- and high-dosed groups. Evidently, the different population dynamics during the first month, seen in Figures 4a through 4c, have conditioned the detritus to favor different consumers in the second month. Here we can see the toxin's tertiary effects. We also have a clear demonstration that the dosed systems have not recovered at all, but that the microscopic community remains sensitive to its entire history.

Effects of JP-8 jet fuel

Figure 5 shows a microcosm experiment with a different design, which was run for four months and illustrates the effect of combined jet fuel JP-8 and heat shock, with control (blue), jet fuel alone (yellow), and both jet fuel and heat (red). In Figures 5a through 5d, pH is plotted on the *x*-axis and the algae *Scenedesmus* on the *y*-axis. These figures illustrate some of the more subtle investigations that can be pursued using Worm Plots. For example, the large amount of variance in the middle of the experiment calls for a closer look.

Using diamond-whisker plots in Figure 5b, we can see that the variance in the control (blue) group is due to outliers because the diamond is small relative to the whiskers. Looking at the individual points in Figure 5c gives us more information, because we can see that a single point from the blue group accounts for most of its variance in both the *x*- and *y*-axes, with tight replicability for the other points. Further, we can see that the outlying blue worm seems to follow a path similar to the yellow worms and not simply a randomly different path.

Looking farther along the temporal axis in Figure 5c reveals that the high variability in the dosed (yellow) group is due to two outlier worms, which diverge markedly from their own group and the control group. Rotating the image to get a better look at this part of the experiment and adding the dosed-heated (red) group gives us Figure 5d. For the middle part of the experiment, the dosed (yellow) and the dosed-heated (red) groups follow two different tracks: either a "low road," similar to the control (blue) group, or a "high road," which includes two of the dosed worms and four of the dosed-heated worms.

These graphics suggest that two paths can be followed through the middle part of the experiment. In addition, dosing and/or heating the systems has a stochastic tendency to push a system onto the wrong track, with a slightly bigger push coming from the combined dose-and-heat insults. This is suggestive, but requires more experimentation to determine if it is a real discovery or an artifact of the small number of systems used in the standard experimental design.

Conclusion

These examples reveal some of the qualitative insights that can be obtained when viewing multivariate time series of grouped data with Worm Plots. We believe that our visualization prototype solves some particularly thorny problems with these data and can play an important part in many kinds of laboratory experiments as well as field monitoring of the environment. ■

Contact Matthews at the Computer Science Department, Western Washington University, Bellingham, WA 98225, e-mail matthews@cs.wwu.edu.