The Water Line

Newsletter for the Lakes of Missouri Volunteer Program

Volume 2 Number 1 January 1998

THANK YOU VOLUNTEERS!!!!

Thank you to all of our volunteers for a successful 1997 sampling season. All of the lab analysis has been completed and we are in the process of writing the 1997 Data Report. We are excited about the new format of this year's Data Report. It will contain data from all of the public lakes in the LMVP so that you can compare your lake to other lakes around the state. There will also be a educational section that will provide you with information about the parameters, statistical explanations, and general information about Missouri watersheds.

We hope to begin scheduling our data review sessions in February. We will be contacting you to let you know about times and places for the review sessions.

VOLUNTEERS NEEDED!!

We are looking to fill some old sites and possibly expand the program to some new lakes. If you are interested or know someone who is interested in participating, please have them call and leave a message on our voice mail at 1-800-895-2260 or write us at - LMVP, 112 Stephens Hall, UMC Columbia, MO 65211.

We welcome any responses, but would especially like to fill these specific sites:
A) Lake of the Ozarks: Grand Glaise Arm (preferably 2 miles from Main Lake), Niangua Arm (preferably 9 miles from the main lake), main lake between the 32 and 36 mile markers.
B) Table Rock Lake - Kings River and main lake channel near the Kings River. Other lakes we would like to include in the program are Thomas Hill Reservoir, Long Branch Lake, Clearwater Lake, Wappapello Lake, and Lincoln Lake. We would appreciate any help we can get in filling these spots. Thank You!!

Hysteria Over Pfiesteria - A Pfreak of Nature?

Flesh-eating organisms that can change forms in a matter of hours, that are not classified as plants or animals, can produce toxins that cause massive fish kills and have even been linked to human health hazards? Is this the stuff of science fiction novels or a current environmental concern? The answer to both questions is yes. Some of you have probably heard about this very strange microbe called *Pfiesteria piscicida* (pronounced feast-er-ia pis-ki-seed-a) that is causing quite a stir all along the east coast from North Carolina to the Chesapeake Bay. A novel titled *And the Waters Turned to Blood* has been written about this organism and many scientists are now working to unravel its mystery.

Pfiesteria is a type of dinoflagellate that can receive energy by preying on algae, zooplankton, bacteria, and fish. It can also generate energy by photosynthesizing using chlorophyll from the algae they have eaten. Most dinoflagellates are a harmless part of the plankton that float in the ocean. Some dinoflagellates are toxic and large concentrations of them is also known as the "red tide". Red tide dinoflagellates produce toxins as a defense mechanism to prevent fish from eating (*continue Page 2)*

Statistics!

If this word makes you fog over, read Dan's article on page 3 of this newsletter. This article will provide you with information that can improve your understanding of the upcoming Data Review sessions.

them. Pfiesteria are different in that they produce the toxin so that they can actually attack and eat fish. This behavior has not been previously recorded in dinoflagellates.

Further studies revealed 24 different life stages for this one organism

In the late 1980's Pfiesteria was discovered at the North Carolina State Veterinary School. One of their aquariums was filled with brackish water from the Pamlico River and all of the fish in the tank died. Dr. Joanne Burkholder, an aquatic botanist at NC State, was called in to help. Along with the help of other scientists, she found that the cause of the fish kill was an unidentified organism with a very complex life cycle . Further studies revealed 24 different life stages for this one organism. One of the stages is a cyst that can remain dormant until environmental conditions are optimal for its success. Several of these forms are known to be toxic to fish. There have been individual cases reported of human health problems in areas where Pfiesteria has been located. Hazards in the natural environment have not been scientifically proven to be linked to this organism. At least thirteen people who have worked with Pfiesteria in a lab setting have had serious adverse health impacts. Some symptoms demonstrated were skin sores in the area of direct contact of samples containing toxic cultures, reddening of the eyes, headaches, kidney and liver dysfunction, acute short-term memory loss and severe cognitive impairment. Most of the acute symptoms reported reversed over time when the person was not directly exposed to Pfiesteria. One of the results of these health problems are that Pfiesteria studies must be conducted in isolated, quarantined, limited-access facilities. Studies are continuing to understand more about these potential human hazards.

Pfiesteria has probably been around for quite a long time spending most of its time as a nontoxic consumer of bacteria, algae and other small animals. But lab and field studies suggest that nutrient loading from septic systems, sewage treatment plants, runoff from feed lots and fertilizer application have caused the environment to change. These changes create an environment in which Pfiesteria spends more time as a predator of fish. Pfiesteria outbreaks have been found along the coast mostly in warm salt waters with low oxygen levels and high nutrient loads but so far no definite data have been produced showing exact correlations supporting these links.

Informational hot lines have been set up along the East Coast to keep the public informed about this problem. Scientist have asked citizens to report fish kills and have provided instructions to volunteers on collection of water samples from fish kill sites. Increased public awareness and education can provide a valuable link in trying to understand this organism and control further problems

To date there have been no reported outbreaks of this organism in fresh water. There are still many questions to be answered before we will completely understand this creature.

Survey Responses

Thank you to everyone who returned their surveys. We had 42 returned. The results will be presented at the Seventh International Symposium on Society and Resource Management at UMC this spring. Thank you again for your participation.

SIMPLIFYING STATISTICS

It was once said that "There are three kinds of lies: lies, damned lies and statistics*." While this may be an exaggeration, statistics can be confusing and misleading. The key to understanding statistics is knowing; 1) the terminology, 2) how the statistics were calculated, and 3) potential shortcomings of the statistics presented. The following article will hopefully provide you with some information about basic statistics (with an emphasis on some of the statistical terms used by the Lakes of Missouri Volunteer Program in the upcoming data report) and explain some of the ways that statistics can be misleading.

A lot of the information given in the data report is what we call descriptive statistics. Instead of giving you the 8 phosphorus values for a certain lake we use statistics to describe or summarize the values. The strength of descriptive statistics is that it allows for easier comparisons between lakes. For example, let's say the following is phosphorus data (in micrograms per liter) collected by volunteers from two lakes during the April-September sampling season.

Lake 1	20	32	41	24	36	35	24	28
Lake 2	25	42	30	41	25	21	33	23

A quick comparison of the numbers indicates that phosphorus concentrations in these two lakes were comparable but it is hard to tell just how similar they were. By calculating descriptive data such as an average (also known as the mean) for the two lakes we are able to determine just how similar the lakes were. The average for both lakes is 30. Notice that the phosphorus concentrations were never the same during any given sample yet we are able to say that the concentration is the same over the sample season. Because the concentrations of all of the parameters we measure in the Lakes of Missouri Volunteer Program vary over the course of the year we are better off looking at the average for the sample season instead of focusing on any one single value.

One potential problem with average values is that they do not reflect how the data were distributed. When we speak of distribution of data we are referring to the range (the difference between the minimum and maximum values) and how the individual values were grouped in that range. Let's look at another example to see how the average value can be misleading.

Lake 3	21	24	25	27	28
Lake 4	14	16	18	23	54

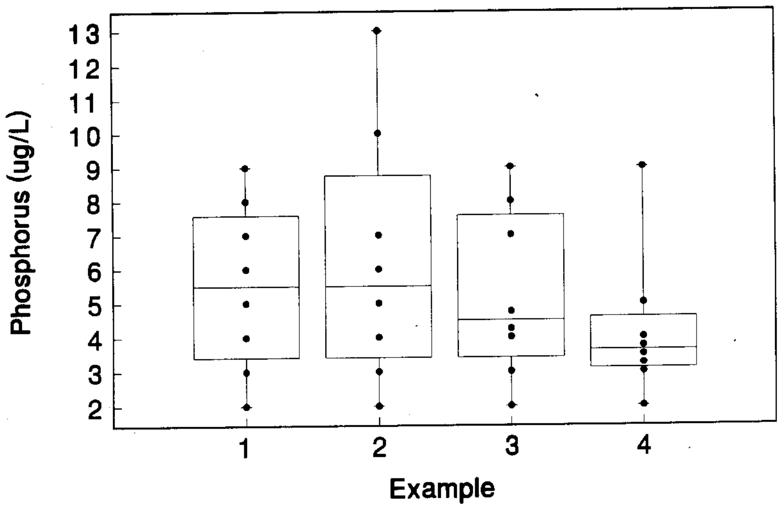
The average phosphorus concentration for both lakes is 25. When we look at the individual data we see that the average of 25 does a good job of describing the data for Lake 3. All of the values were similar to this average. The average of 25 does a poor job of describing conditions in Lake 4 as four out of five values are below this average. That means 80% of the time concentrations were less than the average. In the case of Lake 4 one extreme high value (54) influences the data in such a way that the average does a poor job of describing the "normal condition."

A term that is often confused with the average is the median value. Simply put, this number is the middle value if all of the values were placed in order lowest to highest (if you have an even number of values the median is the average of the two middle values). The median value is used much like the average, to gauge the general condition. It is useful because it is not influenced by extreme values like the average is. The median values for the Lakes 3 and 4 were 25 and 18 respectively. We can see that these median values do a good job of describing "normal conditions" in these two lakes.

Other descriptive data used by the LMVP include minimum and maximum values. The minimum and maximum define the range of values measured for a given parameter which aids in describing the distribution of the data. Maximum values are also important because they represent the water quality at its worst.

A common method for presenting data is with box plots (Figure 1). At first glance these plots can be intimidating but once you know what each aspect of the box plot represents you can tell a lot about the data and how it was distributed. These plots consist of three components: 1) the box, which encompasses the middle 50% of the data, 2) the horizontal line within the box, which represents the median value, and 3) the vertical lines extending above and below the box, which indicate maximum and minimum values respectively. To determine how the data were distributed use the location of the median line within the box as well as the location of the box relative to the minimum and maximum values. We will now look at four examples presented in Figure 1 and describe the average, median, minimum, and maximum values and investigate how the data influences the box plot.

FIgure 1



Example 1 is what is known as an even distribution. The individual values are;

2 3 4 5 6 7 8 9.

The average of these values is 5.5 which also happens to be the median value (remember when dealing with an even number of values the median equals the average of the two middle values, in this case 5 and 6). The minimum and maximum values are 2 and 9 respectively giving us a range of 7. This is considered an even distribution because the mean and median are located in the middle of the range. In other words the minimum and maximum value are equal distances from the median value. The box plot for this data has the median line in the middle of the box and the minimum and maximum lines extend equal distances from the box.

Example 2 contains the following values; 2 3 4 5 6 7 10 13. The average equals 6.25 and 5.5 is the median. Minimum and maximum values are 2 and 13, giving us a range of 11. Having an average that is larger than the median suggest that the data is not balanced around the median.

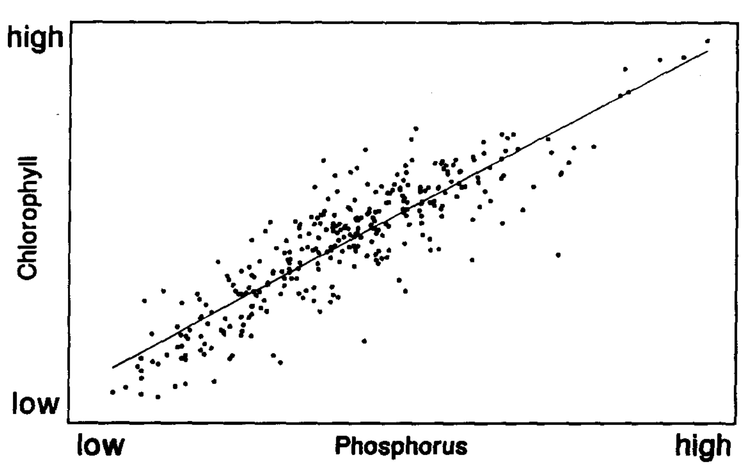
Review of the data shows that the two high values (10 and 13) deviate from the median more than the two low values (2 and 3). This leads to data that is referred to as skewed. When we compare the box plot for Example 2 to the box plot for Example 1 we see two indications that the data was skewed: 1) the vertical line identifying the maximum is longer than the minimum line, and 2) the box extends higher above the median line than it does below.

In Example 3 the values are; 2 3 4 4.25 4.75 7 8 9. The average is 5.25 and the median is 4.5. The range is 7 with minimum and maximum values of 2 and 9 respectively. We again have skewed data as the median and average are not equal. This time instead of extreme high values, the skewing is caused by the values of 4, 4.25 and 4.75 being clumped close together. Comparison of this box plot to Example 1 shows that the plots are the same with the exception of the median line. In Example 3 we are tipped off to the skewness in the data by the fact that the area of box above the median line is larger than the area below.

Example 4 consist of the following values; 2 3 3.25 3.50 3.75 4 5 9. The average for this data equals 4.19 and the median is 3.375. The minimum, maximum and range is the same as seen in Examples 1 and 3. Again we have skewed data caused by a clumping of low values (similar to Example 3). This time the number of values clumped together is greater, causing the box to be smaller. Tip off to the fact that the data is skewed are the median line is not in the center of the box and the maximum line extends farther away from the box than the minimum line.

Besides describing the data, statistics can provide us with information about how two parameters relate to each other. This is usually done with a regression plot (Figure 2). This plot uses what is known as independent and dependant variables. The regression is set up so that the independent factor is across the horizontal axis (also known as the x-axis). The dependant factor is scaled on the vertical or y-axis. Each data point in the regression represents the values for the two parameters. These values can be from a single sample, a yearly average or an average taken over a number of years. The basic idea is that the graph will show us how changes in the independent variable will lead to changes in the dependent variable. Regression plots have a regression line which is often described as the "best fit line" for the data. In other words the line can be thought of as describing the average relationship between the two parameters. The strength of a regression plot is that it tells not only what the relationship between the two parameters is but how strong of a relationship it is. The regression line can also be used to make predictions. If the independent variable changes, we can use the regression line to estimate what the expected change in the dependant variable will be.

Figure 2



Another way the LMVP uses statistics is to develop predictions concerning water quality. Since the parameters we are interested are dynamic and fluctuate one year to the next it is important to determine how much fluctuation is normal. Once we have determined the natural variability is we can better monitor for changes in water quality associated with changes in human influences. In order to determine natural variability we need data from a number of years (the number varies lake to lake and depends on how much influence weather patterns have on water quality in a given lake). After we have collected data over a period of years we use the average values from each year to predict a range of values that can be expected. Theoretically if human influences do not change then the following year the parameters should fall within the predicted ranges. If the values are not within the predicted ranges this may signal changes in water quality.

A Final Word Concerning Statistics.

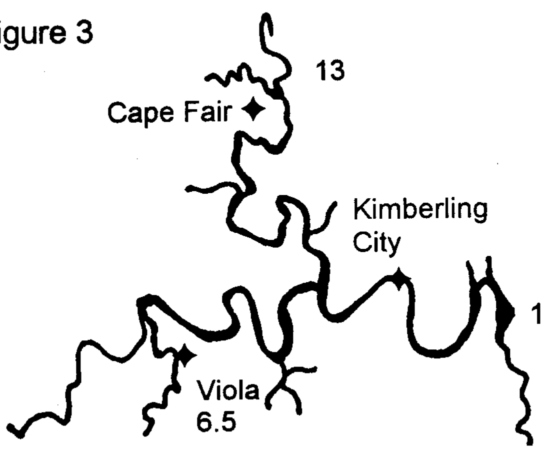
Statistics don't lie but they can mislead. If I collect one sample from a lake and report back that the chlorophyll concentration for this lake for the period April through September was 12.5 micrograms per liter I would be misleading you. Since there are 183 days during this period my sampling on only one of them represents a very small sample size. When we have a small sample size our chance of it not representing the true average condition increases. As sample size increases the probability of correctly describing conditions also increases. This is why we ask volunteers to collect 8 samples during the season. We feel this is an adequate number of samples to describe water quality in our lakes. Be wary anytime you are presented with statistical values and there is no mention of how the data were collected or how many "samples" were taken.

Another potential problem with statistics involves what is known as temporal and spatial variation (temporal refers to time and spatial with space). Inspection of volunteer collected data shows that we can expect differences in data over the course of a sample season as well as from year to year. The temporal variation that occurs during the course of the sample season relates to springtime storms and lake overturn (see September's Water Line). This temporal variation is why we sample from April to September. Spatial variability can be demonstrated by looking at Table Rock Lake data (Figure 3). We see that sites located on different parts of the lake can have extremely different water quality. On our smaller lakes this spatial variation is not as much of a concern. When presented with statistics relating to environmental issues consideration for natural variation due to time and space should be made.

Statistics can be a powerful tool in describing and managing environmental problems but like all tools it is only effective when used properly.

* This quote appears in Mark Twain's autobiography and is credited to Disraeli.

Figure 3



Average values from 1997 sampling season.

Site	Phosphorus (μg/L)	Chlorophyll (μg/L)	Secchi (inches)
1	11	4.5	156
6.5	26	11.3	57
13	128	55.7	35

Professor Limno's Fun Facts

Compiled by Steven McComas - in *Lakeline - A North American Lake Management Society Publication.*

*******How many different aquatic plants are there? The theory is that plants evolved on land and then adapted to the water in lakes and oceans. Of the roughly 350,000 land plants, species that produce seeds, about 2 percent of that number have gone to the wet side. It is estimated there are about 7,000 aquatic species.

*******Too much fertilizer in a lake can cause nasty algal blooms. Often phosphorus concentrations as low as 50 parts per billion (equivalent to five cans of Surge mixed in with 100,000,000 cans of Coke) can produce these blooms. How much phosphorus is 50 parts per billion in a lake? A pond the size of a city block (about two acres) and 6-feet deep with 50 parts per billion of phosphorus in the water column, has less than two pounds of phosphorus suspended in the water.

For comparison, if we followed directions on the fertilizer bag for lawns, we would apply about six pounds of phosphorus on the yards in a city block.

That's part of the challenge for keeping lakes clear. We add fertilizer to yards to make grass grow, but want to keep fertilizer out of the lakes to keep algae from growing too abundantly.

*******How far down can you see into lakes? Lowering a Secchi disk until it disappears from sight has been a standard of measurement technique for water clarity for over a hundred years. The world record for clarity appears to be 217 feet (66 meters) read in the Sargasso Sea, reported in 1972. For lakes, the record to beat is 144 feet (44 meters) recorded for Crater Lake, and reported by Doug Larson in 1973.

LAKE FORMATIONS

1. **Glacial Lakes** - Lakes formed by the various actions of glaciers, from scoured out holes from past glaciers to lakes that exist at the bottom of active glaciers. Examples: Cedar Bog Lake in Minnesota, Fremont Lake in Wyoming.
2. **Tectonic Lakes** - Tectonism refers to the warping, buckling and movement of the earth's shell. These movements cause depressions that can hold water. Examples: Lake Tahoe in Nevada, Lake Baikal in Siberia.
3. **Landslide Lakes** - Lakes that owe their existence to the impoundment of stream valleys by rock slides, mud flows, or other mass movements of soil or rock. Examples: Mountain Lake in Virginia, Lake Tali Karng in Australia.
4. **Volcanic Lakes** - Lakes created by volcanic eruptions or from a collapsed volcanic crater or from lava impoundments. Examples: Crater Lake in Oregon, Zuni Salt Lake in New Mexico.
5. **Solution Lakes** - Lakes created in basins where the dissolution or removal of materials may hold water. These are areas with a karst topography where sink holes are common. Example: Silver Springs in Florida. Salt collapse lakes are formed by the sapping of deep-lying salt pockets by groundwater Example: Montezuma Well in Arizona.
6. **Piping (false karst lakes)** - Lakes formed by the depression from piping. Piping is the process of subsurface tubular drainage that removes looses sands or soils away to be dumped down mountain slopes. The depressions left from these tunnels form lakes. Example: Dead Man Lake in Arizona.

See our next issue of Water Line for more lake types.