# ADVANTAGES AND HAZARDS OF USING MICROSOFT EXCEL™ TO ORGANIZE AND DISPLAY WATER QUALITY DATA

Robert C. Fuller

AUTHOR: North Georgia College & State University, Dahlonega, Georgia, 30597
REFERENCE: *Proceedings of the 2011 Georgia Water Resources Conference*, held April 11–13, 2011, at the University of Georgia.

**Abstract.** I have been collecting water quality data on nine major tributaries to Lake Sidney Lanier, as well as from a bridge over the lake, and from the lake's discharge below Buford Dam, since 1998. Dr. Mac Callaham (North Georgia College & State University) began this project in 1987 and continued it until 1998. One of the most significant changes that I made to the project was to organize my own data and all earlier data into spreadsheet form using Microsoft Excel™. This software provides tremendous flexibility and automation in organization, visualization, and analysis of the data. However, there are hazards to this methodology that can lead to serious misinterpretation and misunderstanding of the data if the user is unaware of certain subtleties in the software. This paper presents methodologies for storage, manipulation, and display of water quality data with Microsoft Excel™, discusses certain common, yet difficult to detect sources of error that can be interjected into large datasets with this software, and suggests data checking methodologies to guard against these unintended consequences.

## INTRODUCTION

The Upper Chattahoochee Basin Group, comprising City of Gainesville, Forsyth County, Gwinnett County, Hall County, and Upper Chattahoochee River Keeper (as a non-voting member), funds a long-term trend water quality monitoring project of the major tributaries to Lake Sidney Lanier, one site within the reservoir, and the discharge from the reservoir. Eleven sites (Figure 1) are sampled twice monthly. One site is chosen at random as a duplicate for each bimonthly sampling event. Duplicate tests are run and duplicate samples are collected at this site for subsequent lab analyses.

All sampling and testing is carried out in accordance with a Sampling and Quality Assurance Plan produced by the author and approved by the Environmental Protection Division (EPD) of the Georgia Department of Natural Resources in 2008. At each site, nine items of information are collected *in situ*: date, time, weather condition, air temperature, water temperature, pH, conductivity, specific conductance, and dissolved oxygen. Preserved samples are then transported back to the North Georgia College & State University Water Lab, where seven additional tests are conducted: total suspended solids, turbidity, alkalinity, hardness, biochemical oxygen demand ($BOD_5$), chemical oxygen demand (COD), and fecal coliform count. A portion of the samples are sent to the Institute of Ecology laboratory in Athens, Georgia, where four additional tests are run once per month, i.e., every other sampling event: nitrite plus nitrate, ammonium, total nitrogen, and total phosphorus.

.

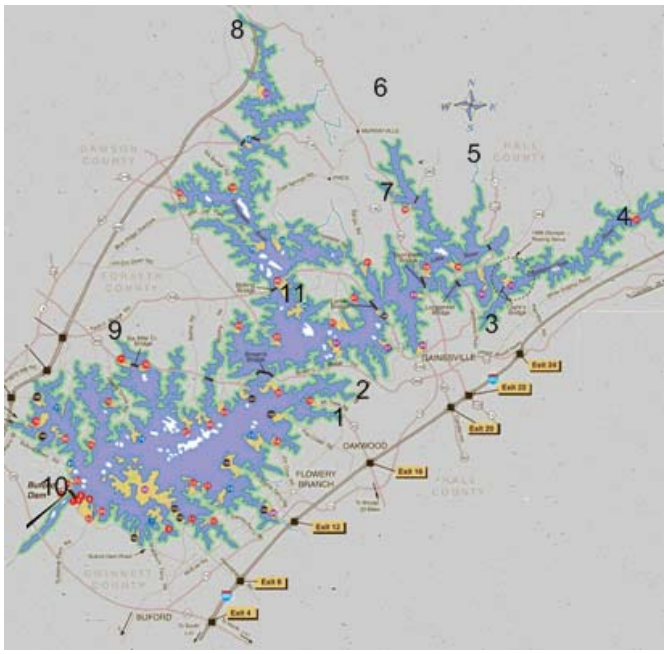



**Figure 1. Lake Lanier sampling sites 1 through 11. Base map: U.S. Army Corps of Engineers (2010).**

Data collected in the field and in the lab aggregate to 240 items (including duplicates) once per month and 192 items on alternate sampling dates, when nitrogen and phosphorus series are not being run. Of these, 220 data entries are stored, manipulated, and displayed using Microsoft Excel™ spreadsheet software once each month, and 176 are processed for the month's other sampling event. The 20 data from duplicate samples are kept on file but not entered into the spreadsheet. These duplicate samples are checked against other samples in accordance with approved Sampling and Quality Assurance Plan sole-

ly for the purpose of monitoring the quality of sampling and testing procedures. EPD stipulated no specific requirements for the use of duplicate samples, but we use them to determine the repeatability of test procedures. Annually, this adds up to 4,752 individual data entries. Using Excel™ to organize the data is a tremendous advantage, but it also creates an opportunity for insidious errors to be introduced into the dataset, errors which can be subtle in the manner of their introduction, pervasive in their impact on the dataset, and difficult to ferret out, as further explained in this paper.

## DATA STORAGE AND MANIPULATION

All results from the project prior to 1998 had been calculated manually using a hand-held calculator and entered into a static, i.e. non-electronic, spreadsheet. The use of a spreadsheet to organize and display the data became the accepted norm by the funding agency. This history, coupled with my own comfort level with Microsoft Excel™, led me to develop what began as a rather simple, 24-worksheet spreadsheet to store and organize the data. As time passed, though, I explored increasingly involved additional worksheets to summarize, reorganize, and display the data. As of 2010, I had expanded the spreadsheet from 24 to 194 worksheets. These additional worksheets were developed to provide ancillary information, summarize the data by quarter and year, reorganize the data by sampling site rather than date, and display the data both as scatter plots by constituent for the year and as quarterly and annual means for each constituent (except as geometric means for fecal coliform counts).

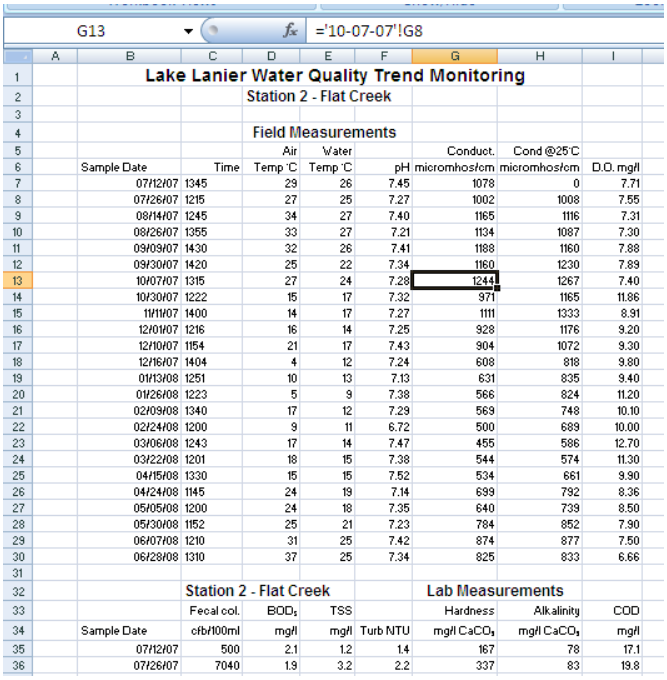
The 220 data for a single monitoring event are stored on one worksheet of a much larger spreadsheet (Fuller, 2010). Over a year's time, the spreadsheet eventually expands to comprise 194 individual worksheets. The first three worksheets are informational about the project and the spreadsheet organization. Worksheets 4 through 27 are the worksheets associated with the year's 24 sampling events and are where raw data are entered. Figure 2 provides a sample of the data entry worksheets' organization.

| | A | B | C | D | E | F | G | H | I | J |
|---|---|---|---|---|---|---|---|---|---|---|
| 1 | Lake Lanier Water Quality Trend Monitoring | | | | | | | | | |
| 2 | Samples taken: | | October 7, 2007 | | | | | | | |
| 3 | | | | | | | | | | |
| 4 | Field Measurements | | | | | | | | | |
| 5 | | | | Air | Water | | Conduct. | Cond @25°C | | |
| 6 | Station | Name | Time | Temp °C | Temp °C | pH | micromhos/cm | micromhos/cm | D.O. mg/l | Comments |
| 7 | 1 | Balus Cr. | 1200 | 26 | 19 | 7.39 | 106 | 118 | 8.2 | p. cloudy |
| 8 | 2 | Flat Cr. | 1315 | 27 | 24 | 7.28 | 1244 | 1267 | 7.4 | p. cloudy |
| 9 | 3 | Limestone Cr. | 1130 | 25 | 20 | 7.16 | 123 | 138 | 8.5 | p. cloudy |
| 10 | 4 | Chatt. R. | 1100 | 24 | 21 | 7.11 | 48 | 50 | 7.5 | p. cloudy |
| 11 | 5 | Little R. | 1040 | 24 | 19 | 7.22 | 60 | 67 | 7.1 | clear |
| 12 | 6 | Wahoo Cr. | 0945 | 20 | 18 | 7.12 | 60 | 70 | 7.0 | clear |
| 13 | 7 | Squirrel Cr. | 1005 | 23 | 20 | 7.08 | 73 | 82 | 8.5 | clear |
| 14 | 8 | Chestatee R. | 0920 | 19 | 20 | 7.24 | 41 | 45 | 7.9 | p. cloudy |
| 15 | 9 | Six Mile Cr. | 1405 | 28 | 20 | 6.96 | 189 | 207 | 7.9 | p. cloudy |
| 16 | 10 | Buford Dam Splw | 1440 | 29 | 10 | 6.42 | 36 | 49 | 4.5 | p. cloudy |
| 17 | 11 | Bolling Bridge | 1345 | 27 | 24 | 7.27 | 47 | 47 | 7.9 | p. cloudy |
| 18 | | | | | | | | | | |
| 19 | | | | | | | | | | |
| 20 | Lab Measurements | | | | | | | | | |
| 21 | | | Fecal | $BOD_5$ | TSS | | Hardness | Alkalinity | COD | |
| 22 | Station | Name | cfb/100ml | mg/l | mg/l | Turb NTU | mg/l $CaCO_3$ | mg/l $CaCO_3$ | mg/l | |
| 23 | 1 | Balus Cr. | 880 | 1.9 | 0.6 | 2.2 | 44 | 43 | 3.4 | |
| 24 | 2 | Flat Cr. | 80 | 1.9 | 0.6 | 0.8 | 217 | 54 | 12.3 | |
| 25 | 3 | Limestone Cr. | 100 | 2.0 | 1.2 | 3.3 | 54 | 54 | 7.9 | |
| 26 | 4 | Chatt. R. | 60 | 2.1 | 14.8 | 12.5 | 14 | 15 | 6.9 | |
| 27 | 5 | Little R. | 300 | 1.9 | 11.4 | 12.5 | 17 | 23 | 5.9 | |
| 28 | 6 | Wahoo Cr. | 1270 | 1.9 | 9.2 | 16.0 | 20 | 26 | 8.4 | |
| 29 | 7 | Squirrel Cr. | 870 | 2.0 | 11.2 | 5.8 | 27 | 33 | 7.4 | |
| 30 | 8 | Chestatee R. | 190 | 1.7 | 3.0 | 5.0 | 13 | 15 | 6.4 | |
| 31 | 9 | Six Mile Cr. | 1400 | 1.7 | 1.8 | 2.7 | 47 | 19 | 2.0 | |
| 32 | 10 | Buford Dam Splw | 8 | 1.7 | 1.8 | 4.7 | 14 | 15 | 2.5 | |
| 33 | 11 | Bolling Bridge | 0 | 1.5 | 2.2 | 2.5 | 13 | 16 | 3.9 | |
| 34 | | | | | | | | | | |
| 35 | | | $NO_2$+$NO_3$ | $NH_4$ | Tot N | Tot P | | | | |
| 36 | Station | Name | mg/l | mg/l | mg/l | mg/l | | | | |
| 37 | 1 | Balus Cr. | 0.6634 | 0.0099 | 1.1524 | 0.0041 | | | | |
| 38 | 2 | Flat Cr. | 17.0169 | 0.0222 | 23.9789 | 0.0263 | | | | |
| 39 | 3 | Limestone Cr. | 0.4982 | 0.0169 | 23.3754 | 0.0071 | | | | |
| 40 | 4 | Chatt. R. | 0.4082 | 0.0438 | 10.3025 | 0.0207 | | | | |
| 41 | 5 | Little R. | 0.7740 | 0.0283 | 5.5969 | 0.0115 | | | | |
| 42 | 6 | Wahoo Cr. | 0.2170 | 0.0423 | 1.9598 | 0.0489 | | | | |
| 43 | 7 | Squirrel Cr. | 0.2525 | 0.0642 | 5.2055 | 0.0717 | | | | |
| 44 | 8 | Chestatee R. | 0.1755 | 0.0159 | 1.9598 | 0.0153 | | | | |
| 45 | 9 | Six Mile Cr. | 8.3309 | 0.0178 | 18.9063 | 0.0151 | | | | |
| 46 | 10 | Buford Dam Splw | 0.2991 | 0.0629 | 5.9394 | 0.0017 | | | | |
| 47 | 11 | Bolling Bridge | 0.0147 | 0.0074 | 1.7477 | 0.0067 | | | | |

9-09-07 | 9-30-07 | **10-07-07** | 10-30-07 | 11-11-07 | 12-01-07 | 12-10-07

**Figure 2. Data entry worksheet.**

Data from these 24 worksheets are then referenced in Five sheets that calculate quarterly and annual mean values for the various water quality parameters and eleven worksheets that reorganize the data into annual listings of measurements by sampling site, one worksheet for each site. Figure 3 presents part of the annual measurements for one of the sampling sites. The entire worksheet is not shown. Take note of how the formula bar shows the reference for a value back in one of the data entry worksheets (Figure 2). This is typical of all cells in the sampling site worksheets. The worksheets are all arranged in such a way that only a single formula was required to be typed for each constituent. Remaining cells were filled by copying and pasting, taking advantage of relative addressing. Readers unfamiliar with this may wish to consult any of the many instructional manuals available on Microsoft Excel™ or the help files that are provided with the software. Only one worksheet had to be prepared in this manner. The other ten were copied from the original and modified using Find and Replace in order to reference the correct parts of the data entry worksheets.



G13 ='10-07-07'!G8

| | A | B | C | D | E | F | G | H | I |
|---|---|---|---|---|---|---|---|---|---|
| 1 | | Lake Lanier Water Quality Trend Monitoring | | | | | | | |
| 2 | | Station 2 - Flat Creek | | | | | | | |
| 3 | | | | | | | | | |
| 4 | | Field Measurements | | | | | | | |
| 5 | | | | Air | Water | | Conduct. | Cond @25°C | |
| 6 | | Sample Date | Time | Temp °C | Temp °C | pH | micromhos/cm | micromhos/cm | D.O. mg/l |
| 7 | | 07/12/07 | 1345 | 29 | 26 | 7.45 | 1078 | 0 | 7.71 |
| 8 | | 07/26/07 | 1215 | 27 | 25 | 7.27 | 1002 | 1008 | 7.55 |
| 9 | | 08/14/07 | 1245 | 34 | 27 | 7.40 | 1165 | 1116 | 7.31 |
| 10 | | 08/26/07 | 1355 | 33 | 27 | 7.21 | 1134 | 1087 | 7.30 |
| 11 | | 09/09/07 | 1430 | 32 | 26 | 7.41 | 1188 | 1160 | 7.88 |
| 12 | | 09/30/07 | 1420 | 25 | 22 | 7.34 | 1160 | 1230 | 7.89 |
| 13 | | 10/07/07 | 1315 | 27 | 24 | 7.28 | 1244 | 1267 | 7.40 |
| 14 | | 10/30/07 | 1222 | 15 | 17 | 7.32 | 971 | 1165 | 11.86 |
| 15 | | 11/11/07 | 1400 | 14 | 17 | 7.27 | 1111 | 1333 | 8.91 |
| 16 | | 12/01/07 | 1216 | 16 | 14 | 7.25 | 928 | 1176 | 9.20 |
| 17 | | 12/10/07 | 1154 | 21 | 17 | 7.43 | 904 | 1072 | 9.30 |
| 18 | | 12/16/07 | 1404 | 4 | 12 | 7.24 | 608 | 818 | 9.80 |
| 19 | | 01/13/08 | 1251 | 10 | 13 | 7.13 | 631 | 835 | 9.40 |
| 20 | | 01/26/08 | 1223 | 5 | 9 | 7.38 | 566 | 824 | 11.20 |
| 21 | | 02/09/08 | 1340 | 17 | 12 | 7.29 | 569 | 748 | 10.10 |
| 22 | | 02/24/08 | 1200 | 9 | 11 | 6.72 | 500 | 689 | 10.00 |
| 23 | | 03/06/08 | 1243 | 17 | 14 | 7.47 | 455 | 586 | 12.70 |
| 24 | | 03/22/08 | 1201 | 18 | 15 | 7.38 | 544 | 574 | 11.30 |
| 25 | | 04/15/08 | 1330 | 15 | 15 | 7.52 | 534 | 661 | 9.90 |
| 26 | | 04/24/08 | 1145 | 24 | 19 | 7.14 | 699 | 792 | 8.36 |
| 27 | | 05/05/08 | 1200 | 24 | 18 | 7.35 | 640 | 739 | 8.50 |
| 28 | | 05/30/08 | 1152 | 25 | 21 | 7.23 | 784 | 852 | 7.90 |
| 29 | | 06/07/08 | 1210 | 31 | 25 | 7.42 | 874 | 877 | 7.50 |
| 30 | | 06/28/08 | 1310 | 37 | 25 | 7.34 | 825 | 833 | 6.66 |
| 31 | | | | | | | | | |
| 32 | | | Station 2 - Flat Creek | | | | Lab Measurements | | |
| 33 | | | Fecal col. | $BOD_5$ | TSS | | Hardness | Alkalinity | COD |
| 34 | | Sample Date | cfb/100ml | mg/l | mg/l | Turb NTU | mg/l $CaCO_3$ | mg/l $CaCO_3$ | mg/l |
| 35 | | 07/12/07 | 500 | 2.1 | 1.2 | 1.4 | 167 | 78 | 17.1 |
| 36 | | 07/26/07 | 7040 | 1.9 | 3.2 | 2.2 | 337 | 83 | 19.8 |

**Figure 3. Partial annual data for one sampling site.**

Thirty-five charts are then built from the data rearranged by sampling site. Another forty-four worksheets disaggregate the data into four sets of quarterly summary sheets, by sampling site, and sixty-eight worksheets present quarterly mean values for the water quality parameters in the form of bar charts. A small number of place-holder worksheets, used simply to visually separate the tabbed worksheets into groups, make up the 194 worksheets.

## DATA ENTRY ERRORS

Inevitably, in the course of entering data into the 24 data entry worksheets, errors occur. The most common errors involve entering incorrect numbers, losing one's place in a paper data sheet containing field or lab results, leaving out a number, or typing a number twice (Goldberg et al., 2010). These errors become more frequent as the data entry technician becomes more tired or distracted by other, competing obligations (Jenks, 1981). We have experienced a fifth type of error that can be created by the spreadsheet software as a result of correcting other data entry errors and, while the fact that an error exists may be easy to detect, tracking down the source of the error may become quite difficult in large, complex spreadsheets.

**Typing errors.** Simple typing errors, often involving hitting an adjacent number on the keyboard or misreading data from lab data sheets, can be some of the most difficult of errors to detect, though once detected, they are quite easy to correct. Suppose that a fecal coliform count of 1460 is to be entered on a data entry worksheet, but the technician inadvertently types 1450 or 1640. The first error might result from a slip of the finger on the keyboard, while the second might result from an error in perception. Both types of errors have been fairly common in our project. In either case, the result is an incorrect value in the dataset, yet one that is close enough to the measured value to appear plausible. Scatter plots may appear quite reasonable. Calculating means with these incorrect values will further disguise the error, and bar graphs of the calculated means will appear reasonable.

Preventing these errors from entering or remaining in the dataset can generally be done using one or more of four strategies: 1) working slowly and carefully during data entry, 2) data checking by the data entry technician following each small set of data, and 3) read-aloud proofing with another person, or 4) double entry and validation by a second technician (Kawado, et al., 2003). In the Lake Lanier long term trend water quality project, all methods except the fourth are used.

Two other typing errors have also cropped up in this project, but they have been easily detected and corrected. The first involves typing an impossible value for the data type. The most common example of this is to type "O" where "0" was intended. These types of errors will generate error messages in subsequent data analysis, and can be traced back to their origin rather easily. The second of these involves inadvertently hitting a key twice. In this study, these errors are easily detected for nearly all data types by screening analyzed results for outliers. The exception to this is fecal coliform values, which may range from 0 to 10,000 or more. An incorrect value of 220, where the correct value was 20, may easily go undetected by screening for outliers. While we do screen analyzed results for outliers in this project, we have also had occasions when errors of this type went undetected by outlier screening, only to be found by double entry and validation.

**Losing one's place.** We have had several instances in which the person entering data lost his or her place and typed in an entire column of incorrect data. This type of error is normally caught by the data entry technician when the next column of data or the one after that contains values totally inappropriate for the next data type. This may be because of the precision to which we measure and report various data types or the more obvious example of entering text in a field requiring numerical data. Occasionally, though, a column of data may be entered twice, once in the correct location and once in the incorrect location, or a portion of a column of data may be interjected into the wrong place in the dataset. This has not happened often, but the strategies mentioned for

correcting other typing errors are effective in catching and correcting them.

**Missed or multiple data entries.** Missing a number or typing a number twice in a column of data both become obvious when the person entering the data reaches the end of the column. There will be one too few or one too many numbers for the column. The results are obvious and the method of correction is simple. In this simplicity, though, lies a hidden danger to the uninitiated.

The surest way to correct this type of error is to go back and retype the entries. It is tempting, though, to cut and paste, moving data up to correct a double entry or down with an added entry to correct for a skipped datum. This is where the problems begin. In its standard configuration, Microsoft Excel™ adjusts all references to a cell when the contents of the cell are moved. In other words, if cell D8 is referenced in formulas in five other cells, whether in the same worksheet or in other worksheets, and the contents of cell D8 are cut and pasted to cell D9, then all five formulas that referenced D8 will be changed to reference D9. If other cells contain references to any of these five cells, then the error propagates through the dataset. The following example illustrates the process.

Figure 4 presents a simple dataset with three columns of numbers and a scatter plot containing two line graphs. The first column contains numbers that are just what they appear to be, numbers typed directly into the spreadsheet. The second column, column C, contains formulas that reference adjacent values in column B. Similarly, column D contains formulas that reference adjacent numbers in column C and multiply those numbers by 2.

Suppose, though, that an error occurred that resulted in an extra digit being typed in column B. The result might look like Figure 5. It is important to note that the scatter plot does not alert the analyst to the error (unless, of course, she knew that the data should plot as a straight line). The data entry worksheet, however, clearly indicates an error, provided that the data entry technician pays attention. If the technician notices the error and simply retypes the data correctly, deleting the extraneous entry (in cell B in this example), then all will be well. If, however, the technician chooses to cut cells B4 through B8 and paste them into cells B3 through B7, the results will be as shown in Figure 6.

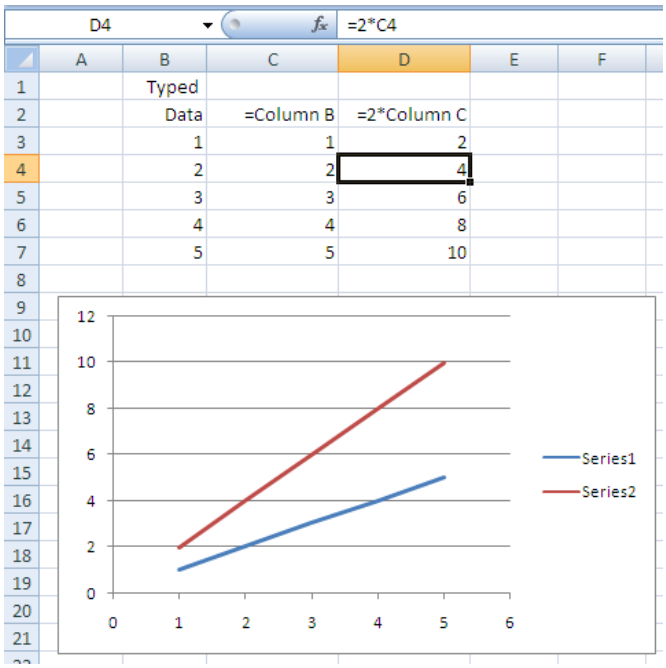



**Figure 4. Example worksheet.**

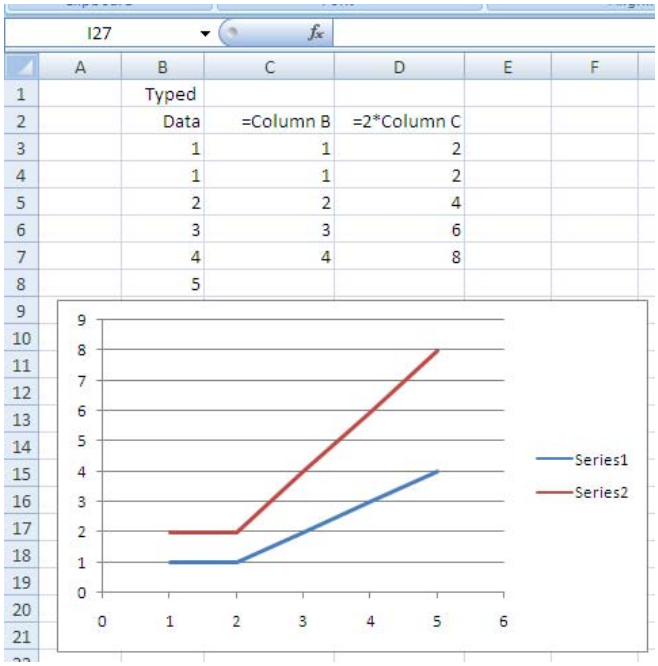



**Figure 5. Example worksheet with extra data entry.**

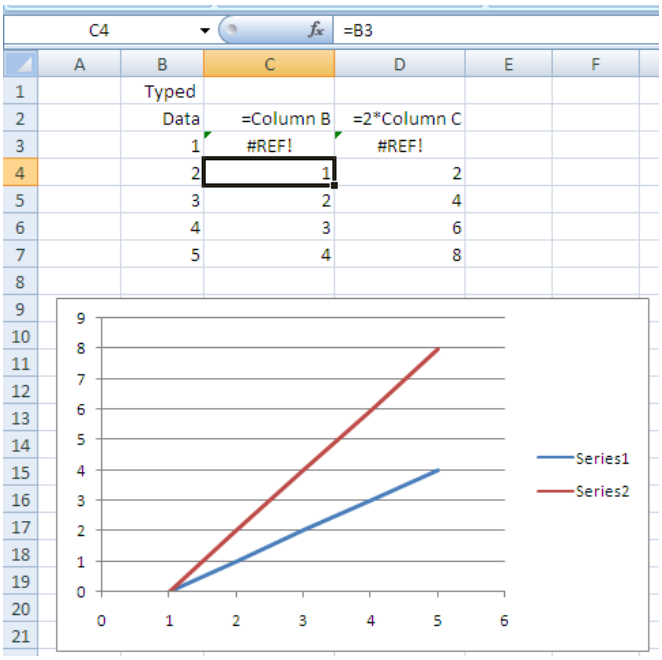



**Figure 6. Example worksheet after ill-advised cut and paste operation.**

In the example given in Figure 6, it is easy to detect the error. Notice, though, that observation of the entered data gives no hint that an error has crept into the dataset, nor does screening for outliers in the analyzed data, the scatter plot in this case, alert the analyst to the problem. It takes observation of the analyzed data that reference the cut and pasted data and from which the scatter plot is created in order to see that something is wrong. In our study, numerous worksheets reference the 24 data entry worksheets, and other worksheets reference these, including scatter plots and bar charts. With experience, we have learned which sheets to check for these types of errors, and we do so routinely, but before we became aware of the problem, we wasted dozens of person-hours trying to figure out what had gone wrong.

The first few times that this occurred, we also made the mistake of retyping the cut and pasted data, and then working through the dataset looking for all instances of reference to the moved data. In a large spreadsheet, this can be quite onerous. We have found that a better method is to undo the cut and paste operation if that is still an option. With newer versions of Excel™, formulas may have been automatically inserted into unwanted locations when the data series was inadvertently typed beyond the intended range. Care should be taken with Microsoft Excel™ 2007 and later to check for this. If undo is not an option, then it is better to cut and paste the data on the data entry worksheet back to its incorrect position, as originally entered, retyping it, then locating all references to the first cell, and correcting the cell with a REF# error. The other cells should have had their cell references corrected by the cut and paste back to the original entry position.

Of course, all persons who work on the Lake Lanier long-term trend water quality monitoring project dataset are now aware of this danger, but mistakes still happen, particularly if the person doing the data entry has an exam coming up, a term paper due, or some similar distraction, so we remain vigilant for this and other possible sources of data errors.

## CONCLUSIONS

There are a multitude of other methods for organizing and analyzing water quality data (Goodall et al., 2008; U.S. EPA, 2003; Whiteaker, 2008), many of which may be superior to the methods described in this paper. However, that is not the point of the current paper. The point of this paper is to describe what has been used in this study and to describe and discuss some of the difficulties that have arisen as a result. The lessons learned by the author include how easily a simple approach to organizing and storing data can expand into a monumentally complex system and how data entry errors can quickly become a major concern.

I am currently working to develop long term trend statistical analyses with the datasets described in this paper, including such tests as covariance and Pearson's R tests on various combinations of variables. This work has required combining the annual spreadsheets described in this paper for the years 1987 through the most recent sampling year. While many of the summary worksheets and graphs have been omitted, the resultant spreadsheet contains 24 times the data contained in an annual spreadsheet described in this paper. Clearly, the opportunities for errors have also increased with this new investigation. One of the most important issues to be addressed in the new work will be that of data quality control. Other storage and analysis tools will be investigated for this project and compared with Microsoft Excel™ for ease of use, acceptance by the funding agency and other users, data quality control options, and statistical analysis capabilities.

## REFERENCES

Fuller, R. C., 2010. Lake Lanier water quality trend monitoring data files. http://radar.northgeorgia.edu/~rcfuller/Research/Lake_Lanier/Data_Files.htm, Dec. 14, 2010.

Goldberg, S., A. Niemierko, M. Shubina, and A. Turchin, 2010. "Summary Page": a novel tool that reduces omitted data

in research databases. *BMC Medical Research Methodology*, Vol. 10, pp. 91-97.

Goodall, J. L.; J. Horsburgh, T. Whiteaker, D. Maidment, and I. Zaslavsky, "A first approach to web services for the National Water Information System", *Environmental Modelling & Software*, Apr2008, Vol. 23 Issue 4, p404-411.

Jenks, G. F., 1981. Lines, computers, and human frailty, *Annals of the Association of Human Geographers*, Vol 71, No. 1, pp. 1-10.

Kawado, M., S. Hinotsu, Y. Matsuyama, T. Yamaguchi, S. Hashimoto, and Y. Ohashi, 2003. A comparison of error detection rates between the reading aloud method and the double data entry method. *Controlled Clinical Trials*, Oct2003, Vol. 24, Issue 5, pp. 560-569.

U.S. Army Corps of Engineers, 2010. http://lanier.sam.usace.army.mil/maps/LakeRecMap.gif. December 12, 2010.

U.S. Environmental Protection Agency. 2003. "U.S. EPA Releases New Version of STORET **Water** Quality **Data Storage** Program", *Water Environment & Technology*; Vol. 15 Issue 6, p6.

Whiteaker, T and E. Ernest. 2008. "CUAHSI Web Services for Ground Water Data Retrieval", *Ground Water*, Jan2008, Vol. 46 Issue 1, p6-9.