**Spreadsheet Analysis of Logs**

E. R. (Ross) Crain, P.Eng.
Spectrum 2000 Mindware Ltd.
Calgary Alberta Canada Ltd
 *Presented at the 10th Formation Evaluation Symposium, Canadian Well Logging Society, September, 1985. Published in Symposium Transactions. This electronic version created May 2002.*
*Author’s Note: For a more modern approach, see “***META/LOG - A Knowledge Based Spreadsheet System To Reduce Complexity In Log Analysis, E. R. Crain**, P.Eng., **Trans. 12th Formation Evaluation Symposium, CWLS, Calgary, Sept 1989” for up-to-date information.** *ERC Jan 2005.*

**Introduction**The tools of the log analysis trade have evolved rapidly over the last 20 years, from charts and nomographs, to slide rules, to programmable calculators, to desktop microcomputers. Each method has its drawbacks. Charts are not very accurate or repeatable. Slide rules are difficult for some people to use and don't add or subtract very well. Let's face it, these methods are as obsolete as the ES log.

Programmable calculators are easy to use but not too easy to program. Desktop computers (or terminals to larger computers) may or may not be easy to use, but are almost invariably programmed by others, so custom analysis is difficult.

Another solution has been available for about seven years, but only recently have computers and software been powerful enough, large enough, fast enough and cheap enough to do log analysis. This solution is the electronic spreadsheet on modern microcomputers, typified by products like Visicalc, Supercalc, Context MBA, and Lotus 1-2-3.

Early versions of Visicalc could do log analysis, but memory and speed limitations made it impractical. Today, however, these limitations have been removed with the advent of professional desktop computers like the IBM-PC/XT, the HP 150 and HP 200 series, and many comparable machines.

**What Is An Electronic Spreadsheet?**An electronic spreadsheet is a computerized replacement for the pencil and columnar pad so familiar to log analysts who do hand calculations at the well site or in the office. One virtue of the electronic spreadsheet is that it can be much larger than a sheet of paper, often allowing up to 256 columns of data (with up to 72 characters per column) by 2000 rows long. The screen or monitor of the computer is a window on this large array of data.

Each row is assigned a row number and each column is headed by a column letter. The intersection of a row and column is termed a cell. Cells are identified by their row and column designation (e.g. C24).

Movement around the portion of spreadsheet seen on the screen is performed by moving a cursor using the keyboard, lightpen or mouse attached to the computer.

Cells can contain four kinds of data:

1. text or labels
2. numbers (raw data or answers)
3. formulae or algorithms
4. spreadsheet functions

Mathematical formulae in spreadsheets show relationships between cells (e.g. C1=+A1+B1. If the relationship of many cells is similar (e.g. C1=+A1+B1, C2=+A2+B2, etc.) then each individual relationship must be described. There are semi-automatic methods for doing this in most spreadsheets. There are no general purpose array operations such as add column A to column B to get column C.

Spreadsheet functions are shortcuts which help the user perform common mathematical computations, such as SUM, AVERAGE, MINIMUM, MAXIMUM, IF...THEN, and PRESENT VALUE. Some of these are not even available in conventional programming languages but must be coded uniquely each time they are needed.

The act of building a log analysis model with a spreadsheet defines all the mathematical relationships in the model, as well as the raw data and analysis parameters. You can edit formulae, insert rows or columns, fix typographical errors, enter new data, and recompute results without changing a Basic or Fortran program. The spreadsheet is the program!

You can also play "what if?." You could vary the water resistivity, matrix or fluid parameters, or shale values and see the change in the log analysis results in 10 to 30 seconds. Computations can be revised immediately after entering a change, or after entering all changes, as decided by the user.

The original spreadsheet, after testing and debugging, can be used immediately, or saved as a standard analysis package. Standard spreadsheets can be customized to fit individual problems when they are used. Managers or professionals may create spreadsheets for their own use, or for use by other professional or technical staff in other departments.

Spreadsheet software is the most widely used application for microcomputers. Over 500,000 copies of Visicalc have been sold, and there are more than 60 competing products, many with implementations on five or more microcomputers. Table 1 and 2 indicate some of these products and their suppliers.

**What To Consider In A Spreadsheet Package**A warning note should be issued. Spreadsheet software is no panacea. It is handy for small jobs needing quick turnaround, or those with a lot of "what if?" situations - the same kind of jobs you would do with charts or a calculator. The spreadsheet will not replace dedicated software on microcomputers or mainframes. One of the reasons for this is the difficulty in plotting results in conventional log analysis format (3 track or 4 track presentations).

Many spreadsheet packages do have integrated graphics, and are capable of illustrating most common crossplots (on linear axes) and can do a limited form of depth plot. A few packages have integrated word processing, so final reports can be constructed after the analysis has been completed.

Not all spreadsheets will work on all computers, nor are all computers large enough or fast enough to be useful. The minimum configuration is usually a computer with 256,000 bytes of memory, a dual disc drive with one to five megabyte storage, a CRT with at least 19 lines and 80 characters, an 80 character dot matrix printer, and a plotter (optional). All these components must be supported by the computer's operating system, as well as the spreadsheet software. A digitizer could be added, but custom software, probably in machine language or assembler language, will be required.

The minimum cost for a practical system would be in the order of: computer C$ 6,000, dual disc drive 2,000, printer 700l plotter 2,300 for a total of $11,000.

The spreadsheet software is often bundled with the hardware, but would cost less than $700 in any case.

Each spreadsheet package utilizes its own syntax (just as the variants of Basic or Fortran do), and interchange of models from one system to another may require some editing. The examples in this paper use the language of Lotus 1-2-3, one of the more sophisticated packages. However, the common commands are similar to Visicalc and conversion should be relatively easy.

One important consideration when choosing a spreadsheet package is whether or not the package supports forward referencing. The older packages, or cheaper packages, do not permit calculations of a cell if it refers to a cell not yet calculated. These programs start at the top left corner and work either horizontally or vertically.

Those that permit forward referencing search the spreadsheet for the key cell and work out from there. This allows the spreadsheet layout to be more flexible and more attractive to the user. Little used columns could be way off screen to allow more valuable columns, such as answers, to be closer to the window. This form of calculation is termed natural computation.

Another necessary feature involves security. There are two forms and both are desirable - hidden cells and protected cells. Hidden cells have a zero width, so formula or data in them cannot be seen. They can still be edited or even made visible, but it is unlikely that an accidental entry to such a cell will be made. Protected cells may be visible or hidden, but cannot be entered by the user to accidentally or erroneously change a key formula, or even lose it permanently. The originator of the spreadsheet can obviously unprotect a cell and modify the contents.

Another area of interest is the ability to use "macros." Macros are miniature programs which can be invoked by a single key stroke. They are stored with the spreadsheet and loaded when the spreadsheet is loaded. They could be used, for example, to print only the desired columns of a larger spreadsheet.

**Advantages Of Spreadsheet Analysis**The advantages of spreadsheet analysis can be stated simply - speed, cost, ease of use, familiarity, and limited programming skills required.

A typical data set of ten depth points and five data curves can be entered in about two minutes by even a poor typist. The calculations take about ten seconds. Three or four crossplots and results can be printed in less than two minutes. Total elapsed time is five minutes. A large data array - say fifty depth points - can be entered, computed and printed in about 20 minutes.

Progressive engineers and geologists familiar with microcomputers and spreadsheets, can learn to use such a package in less than an hour of practice. Modifying a spreadsheet or creating new ones for specialized analysis should take only one day's practice, and thereafter a few minutes to an hour may be needed to tune the spreadsheet to a particular problem.

The spreadsheet example illustrated later took about 60 hours to write, test and debug, but modifications to this basic structure take only minutes. The user must be familiar with the spreadsheet commands, however.

Because the screen layout, printed results and data structure are one and the same, the spreadsheet contents become familiar quickly. The data sheet is always available for viewing, compared to log analysis packages in which the data structure is invisible to the user.

Other factors, such as built in data management, file storing and retrieving, graphics, simple and friendly keystroke sequences, make spreadsheets more attractive than writing Basic or Fortran programs.

The software is usually available free of charge; that is, it is part of an existing micro computer system already in the work place. Even if it is not, it is a low cost item when compared to stand alone packages priced from $10,000 to $100,000.

**Limitations of Spreadsheet Software (1985)**The concept of spreadsheet software was developed for business and financial calculations. As a result, scientific use is not perfect. The following lists some of the problems encountered in developing the present program.

A cell can contain only one statement or formula. Therefore, these statements can be extremely complicated and unreadable if they are to be powerful enough to do the task of a multiline subroutine. For example, to select the desired porosity option from the several which may be computed requires a statement over 100 characters long and with 13 closing brackets at the end. Similarly, the statement to find the minimum shale volume (and not crash when no shale data is available) is 205 characters long. Lotus 1-2-3 allows statements to be 240 characters long, but previous programming experience helps, as does the ability to build complicated statements piece by piece.

Cells can be referenced by absolute locations or by relative addresses. Interpretation parameters are logically stored in absolute addresses, so you always know where they are. However, relative locations for log data curves can create unforeseen problems. For example, it is tempting to use the editing feature of spreadsheets to move a column of data up or down (depth shifting). However, any references to this data in equations will also be shifted, so that data from two different depths end up being used for a single depth. This can be solved by more absolute addresses in the equations, creating a new problem. Adding lines of data to allow for a larger zone becomes impossible unless one uses named data ranges - but it is starting to get unwieldy and complicated at this stage.

Other concepts of editing also fail. For example, re-scaling a curve (for sonde error or a units conversion) is simple. Just copy a formula into each cell and the new value is automatically computed. But if you leave the formula in the cell, it will recompute each time the job is run - not what you wanted. You might consider putting the rescaling equation into a different cell and feed the answer back to the original cell, then delete the rescaling cell. This doesn't work either, since the original cell still contains references to the deleted cell - it never contains the answer you want unless you type the number in yourself.

Arithmetic errors propagate rather badly. If a divide by zero (or similar) error occurs, the term ERR will show in the answer, as well as in every answer that depended on the original error. This is shown by the term ERR appearing in each formula that was affected by the error - thus losing the contents of totally innocent formulae. This might happen to 400 or 500 cells due to some very trivial data error. The situation will rectify itself, and all the ERRs will disappear if you fix the originating error. However, since so many equations show ERR, it is not always possible to find the offending equation on the first pass. If you don't, all is lost, and a lot of retyping will ensue.

The graphics programs are general purpose business oriented, and two dimensional, They create reasonable 2-D crossplots, but cannot handle Z-plots or 4-D plots. Depth plots cannot be made in log analysis format or to a specific depth scale. Even if asked to make a +0.45 to -0.15 scale for density porosity, the program insists on making a -0.20 to +0.50 scale. Depth plots were labelled in scientific notation (2.050E00 for 2050 meters) even when a fixed format was requested. These esthetic problems could not be overcome, and you can't phone your friendly local software supplier for a quick fix over the phone.

Another serious limitation is the inability to use text strings in a formula, or to generate text strings from an equation, or even from a lookup table. It would be nice, for example, to output lithology codes as words or to use the log type or log units in words to switch logic in the program, as Basic and Fortran allow. There is no way to see text on the screen unless you type it there, or use a macro to type it for you.

Spreadsheet size is also a problem. The examples presented here have the usual shale, porosity, saturation, permeability, and lithology models, as well as six crossplots and a hydrocarbon summary. For ten lines of data, the spreadsheet takes up about 70,000 bytes. A set of fifty data points takes a bit more than 260,000 bytes. Conventional Basic or Fortran data files would take much less space. Program "code" is stored with each data file, instead of only once for conventional programs.

The lack of a true DO loop is frustrating and accentuates the space problem. In order to obtain results for three sets of cutoffs, for example, you could run the job three times and print the answers three times. Or all three sets could be coded into the spreadsheet. This takes considerable space, since each cutoff set requires five columns of equations, each the full length of the data set. A DO loop would eliminate the need to code each iteration through the loop. Lotus 1-2-3 does have an iteration mode, but only one variable can be used as the iteration test - not enough to do cutoffs on four or five variables. Indirect addressing of cells would help reduce this problem also, but there doesn't seem to be any way to use the result in one cell to address another cell.

**Examples of Spreadsheet Analysis**The examples shown in the illustrations are from one well in which the Halfway sand was analyzed (same data as Appendix Two). The Halfway is characterized by anhydrite and dolomite in a shaly sand environment. Analysis requires a complex lithology model. The example was treated first as a hand entry job with eleven lines of data of unequal thickness. A second presentation was made using equally spaced data from a digitizer. Interpretation parameters, raw data, results, net pay, crossplots, and depth plots were printed from the spreadsheet. Little difference in hydrocarbon volume is evident between the two approaches.

The analysis model uses slightly simplified versions of the algorithms presented in the The Log Analysis Handbook (Pennwell Books). Plots are graphic dumps of the screen contents to the printer. More attractive, but slower, plots can be prepared using the utility programs supplied by Lotus 1-2-3 to drive a pen style colour plotter or colour printer.



*Figure 1: Detailed listing of raw data (top) and calculated log analysis results (bottom), from Log/Mate spreadsheet.*


*Figure 2: Summary Report for a detailed log analysis, generated automatically by the Log/Mate spreadsheet.*

This spreadsheet program has most of the features of LOG/MATE except for the length of data files, processing speed (five times slower), quality depth plots, 3 and 4-D crossplots, and many of the unique data editing features. Not bad for a $700 program.

The spreadsheet program is available by contacting the author.

**ConclusionS**Spreadsheet analysis of logs is a viable approach to reducing the burden of quantitative log analysis. It provides sophisticated analysis at low cost, is friendly and easy to use, and can be custom tailored to suit the needs of individual analysts or problems. Sufficient limitations on existing spreadsheet capabilities exist to suggest that this form of computer aided analysis will not replace standalone special purpose programs. But spreadsheets are certainly better than nearly all programmable calculator methods at only a modest increase in cost.

Microcomputers and spreadsheet software are becoming ubiquitous in engineering offices, so it is likely that more and more log analysis will be done by this method. The examples presented here should provide a good starting point for any one interested in pursuing this line of analysis.

**TABLE 1: THE TEN MOST POPULAR SPREADSHEET PROGRAMS - 1985**

Name Integrated Packages / Supplier

VisiCalc VisiFile Visicorp VisiCalc Advanced VisiTerm VisiTrend VisiWord VisiPlot San Jose CA 95134 408-946-9000

SuperCalc SuperChart SuperCalc 2 SuperWriter Sorcim Corp San Jose CA 95131 408-942-1727

CalcStar WordStar SpellStar MicroPro Corp San Raphael CA 94903 MailMerge 415-499-1200

Multiplan Nil Microsoft Corp Bellevue WA 98004 206-828-8080

PerfectCalc PerfectFiler PerfectSpeller Perfect Software Berkeley CA 94710 PerfectWriter 415-527-2628

ProCalc ProGraph ProOp Software Products San Diego CA 92121 LogiQuest 714-450-1526

Context MBA word processor data comm Context Mgmt Syst Torrance CA 90505 graphics 213-378-8277 data base

Lotus 1-2-3 graphics data base Lotus Corp Cambridge MA 02138 617-492-7171

**TABLE 2 SOME OTHER SPREADSHEET SUPPLIERS**

Name Supplier

BudgetPlan Dangen Corp

Cope-PC Plenary Systems Inc.

Desktop Plan Visicorp

Easy Calc Norell Data Systems

Easy Planner Information Unlimited Software

Electronic Spreadsheet American Planning Corp.

Ferox Modeler Ferox Microsystems

Finar Finar Research

Graph'n'Calc Desktop Computer Software

LogiCalc Software Products International

Master Planner Comshare Target Software

Microplan Chang Laboratories Inc.

MiniModel Westico Inc.

Plan 80 Business Planning Systems Inc.

Scratch Pad Super Soft

The Thinker Texa Soft

Knowledgeman Micro Data Base Systems

Vision Calc Visicorp

Graph Plan Chang Laboratories Inc.

The Planner Hayden Book Corp.

Peach Colc Peachtree Software

Super Comp-Twenty Access Technology, Inc.

Lisa Calc Apple Computer

Magi Calc Artsci Inc.

Senior Analyst Business Solutions Inc.

Basic Plan Cado Systems

Ultra Calc CIE Systems

Calc Result Computer Marketing Service

Planner Calc Comshare Target Software

Target Planner Comshare Target Software

Superscreen Creative Software Concept

Plan Master Cromemco Inc.

Business Planner Duosoft Corp.

Execu/Model Executic Software

Master Calc Four-M Marketing

Fusion/3-36 Fusion Products

Calc/1 H&A Computer Systems

Horizon Horizon Software

Novacalc Hourglass Systems

Colorcalc Intelligent Systems

Omnicalc ISA Software

Uni Calc Lifeboat Associates

Zen Calc Software Toolworks

Magic Worksheet Structured Systems Group

T/Maker T/Maker Co.

TMP/Calc United Software