

## DATA VALIDATION - REQUIREMENTS OF AN EGM EDITOR

Class 3200.1

R. Michael Squyres

President & CEO

Flow-Cal Inc.

2525 Bay Area Blvd #290

Houston, Texas 77058

The natural gas industry has adopted EGM as a means of increasing the speed and accuracy with which measurement information is obtained. This has created the need for an electronic data management system. These systems, if not properly designed and implemented, could potentially render the entire process useless. Therefore, it is essential that the system add functionality that complements the power of the hardware. With proper implementation, such a system will not only facilitate operations in today's fast paced, post-FERC 636 environment, but also will establish a foundation for meeting tomorrow's measurement challenges.

A good EGM data editing software package should provide a suite of tools to facilitate accurate and timely data processing. It should do this in a structured, feature rich, well designed environment. The program should include functions to do the following: import the data, recognize, review, and correct anomalies report, export, as well as provide advanced ad hoc query capabilities. Other considerations should include the developer's commitment, resources, and long term strategy, vis-à-vis electronic gas measurement, as well as industry's overall acceptance of the package.

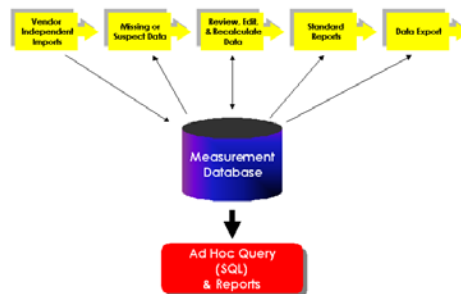
### INTRODUCTION

The natural gas industry is currently racing to update its downstream data handling capabilities to keep up with the technological advances of field automation. One area of concentration is electronic gas measurement (EGM) editing, recalculation, and reporting programs. Fundamental to all measurement systems is the ability to accurately measure, review, correct, and report the data. Any weakness in this chain undermines the speed and accuracy of the system, the primary reasons for automation. In this paper, I attempt to identify the ideal attributes necessary to streamline this process.

One common thread among all successful programs is a good basic design that is easy to understand and use.

This intuitive design, when combined with good documentation, helps to minimize training requirements and reduce problems. The documentation should include hard copy, as well as on-line, context-sensitive help. This can reduce mistakes and technical support requirements, fostering a more productive environment.

For installations with numerous flow computers, the ability to process in batches provides faster processing.



EGM Event Flow Diagram

This requirement exists in each phase, starting with the import process and ending with the report and export functions. Without batch capabilities, the process becomes too labor intensive and reduces the system's effectiveness.

## DATA IMPORT

The first step in operating an effective data management system is to import the data. The editor must have the ability to import the data from a variety of meters and manufacturers. This allows the user to combine all the data into a single system, and affords more options when replacing or acquiring additional field hardware. Importing is a critical step since any mistake made here ensures an incorrect output, which is why clean imports that include all pertinent data are an absolute necessity.

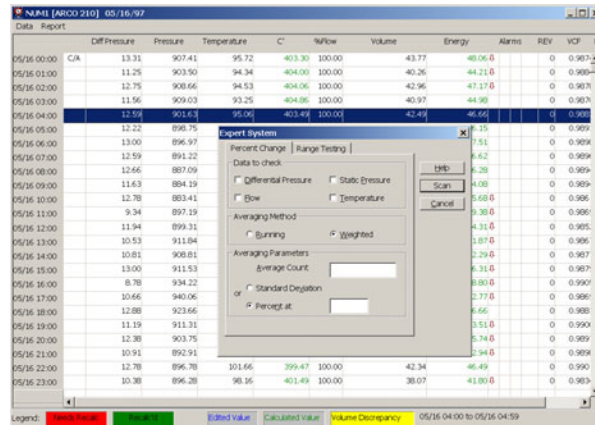


FIGURE 1. Statistical Expert System Data Scan

The import files should be binary or encrypted, as opposed to ASCII, to provide a more secure audit trail. ASCII files can be easily and undetectably altered, using any standard word processor. Deceptively altering a binary or encrypted file is much more difficult, since it involves a time-consuming process of reverse engineering.

Ideally, all import routines should be written and supported directly by the meter manufacturers. The data import requirements include not only the flowing parameters and meter station characteristics, but also the data points used in the evaluation process. These include such items as the alarm setpoints, and the calibration and transmitter ranges.

I will not attempt to list all of these data items; however, the five data categories include: characteristics, gas quality, event and alarm data, as well as the flowing parameters. Additionally, the program should facilitate the import of on-line gas chromatographic data, when applicable. In fact, every variable and value used in the computation of the volume and MMBTU, including calculation methodologies, should be imported to ensure data integrity.

## IDENTIFICATION OF POTENTIAL PROBLEMS

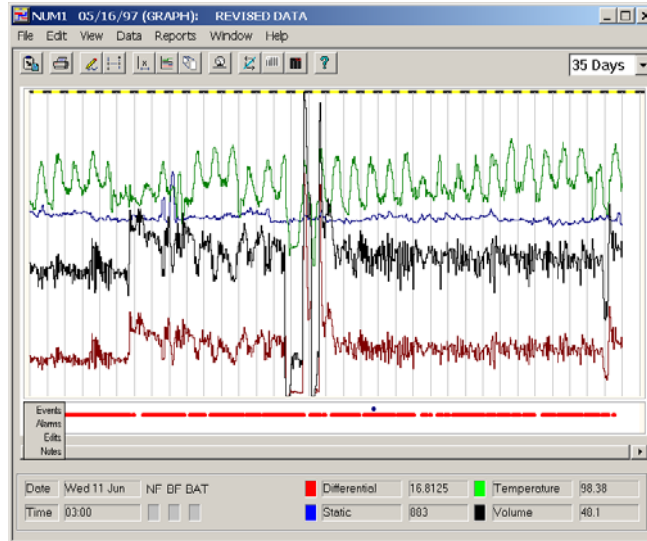
The second step is a means of identifying the missing or suspect data. Locating missing data involves not only a data scan to find unaccounted for flowing parameters, but also includes an event/characteristic audit to identify any discrepancies between a characteristic file constructed from logged events, and the actual characteristic file.

Identifying the suspect data is somewhat more complicated and should include a number of configurable parameters that are used in a statistical data scan. These parameters should be retained for

each meter individually, reflecting each meter's unique flow characteristics and history (see Figure 1). This process is an electronic equivalent to chart censoring.

With conventional meters, when a field technician notices erroneous or missing data he makes a note on the chart and possibly includes a field estimate of the needed correction. This procedure requires an electronic counterpart in the data management system, since erroneous data can easily go undetected by normal statistical methods.

The next step in electronic data management is review and editing of the data, as required. The review process should allow the user to readily review every variable and calculation methodology used in the volume and MMBTU determinations. In addition, each alarm and its associated setpoints should be available for data validation. This data should be presented in a tabular environment, and also graphically for pattern recognition and visual identification of cycles or trends (see Figure 2).



**FIGURE 2. Data Graph**

When the data requires editing, the user must be able to make changes to an individual record or to a date span, in either the graphical or tabular environment. Any changes to the data must be recorded and flagged, along with an associated reason for the change. However, the original data should always be maintained, thus allowing the system to produce reports based on the modified data, while preserving the original data for audit trail purposes.

The user should have the option of multiple editing techniques, depending upon the type of error. Among the available methods are: a cut and paste, point to point, freeform, and some means of creating shifts or offsets for instances of calibration error.

The technique employed should be determined by the user, based on the type of error, information available, and flow history.

Part of the editing process involves the recalculation of the volume and associated MMBTU. The editing program should allow for recalculation using the 1985 and 1992 AGA-3 algorithms, with super-compressibility (Fpv) options of NX-19, AGA-8 GROSS 1, AGA-8 GROSS 2, or AGA-8 DETAIL methods. Data that has not been edited should never be recalculated, since the original calculations are always more accurate.

The flow computer performs its calculations every second, based on live values, and accrues the volume and MMBTU; however, the editor must recalculate based on hourly averages. For this reason, the data manager's recalculation routines should employ some type of skewing technique to more closely approximate what the flow computer would have calculated had the error not existed in the original data.

## **REPORTING**

The data management system should provide for a number of standard reports. Each of these reports should include the ability to process meters individually or in batch. Probably the most basic of all reports is the daily volume statement. The user should have the option to produce the original or a revised version. Some of the additional standard reports include: volume summary, event report, alarm report, missing data, suspect data, and characteristics.

Although the program provides most standard reports, there are always special ad hoc reports required. For this reason, the data manager should provide some means of allowing the user to extract specific data from the database using some standard means, such as Structured Query Language (SQL). This allows a user to create unique reports based on some very specific criteria. [For example: generate a report showing all meters exceeding 4 ppm H<sub>2</sub>S or greater than 3% CO<sub>2</sub> for the last year, sorted in descending order, based on volume.] By allowing this type of access by the users, custom programming for unique applications is not required. This means end users can remain as autonomous as necessary or desired. It also means that no source code is required for modifications, should the editor be a third-party product.

## **DATA EXPORT**

Once all data has been corrected and validated, some means of providing the pertinent data (electronically) to the gas accounting group and to third parties must exist. In today's environment, it is no longer acceptable

to manually re-enter existing data. Therefore, the data management system must provide comprehensive, flexible data export capabilities.

The ability to provide audit data that is compliant with AGA Chapter 21 is mandatory, especially for custody transfer applications. The package should readily support an automated means of bundling the data in an electronic format and/or producing it as hard copy. By automating this process, the user can substantially reduce the time required to individually identify and produce the necessary reports and files.

Some of the more generic attributes needed for the ideal system is network support and an archiving tool for file size management. The system should have an easy means of archiving and retrieving the old data since the quantity of data for EGM meters can quickly exceed storage space. Network support is an important consideration, since most companies already operate on LANs/WANs and most others probably will in the near future.

## **SUPPORT**

The last and perhaps most important consideration for any mission critical data management system is technical support, with regular software updates. When entering closeout, it is essential to have access to a technical support team to assist with any unforeseen problems. The support team must have the tools to remotely diagnose and quickly resolve problems. Routine software updates are necessary, not only to correct minor problems, but to add additional functionality and enhancements. Ongoing development ensures against software obsolescence by continuing to meet any new AGA/API standards and requirements.

## **CONCLUSION**

A good EGM editor should facilitate quick and accurate data processing with a suite of tools for recognizing and correcting anomalies. It should have complete reporting and exporting functions, with advanced ad hoc querying capabilities. In addition to a well designed, feature rich program, one should feel comfortable with the developer's commitment, resources, and long term strategy.

