ATHENA is an interactive SAS®-PC application developed by Quantum Consulting. ATHENA uses advanced macro and windowing techniques that integrate a complex series of independent SAS programs into a coherent, user-friendly analysis environment. We offer this Interactive Posters session as an example of our experience integrating a number of SAS-PC products in the solution of large data handling and analysis problems. Although this application was developed for electric utility demand-side programs such as load monitoring, marketing, and load research programs, SAS-PC users will find that the software techniques used in ATHENA can be readily applied to their own large data editing, validation and analysis problems. Traditionally, solutions to such problems have been implemented in a mainframe computing environment. However, ATHENA takes advantage of the growing power of PC workstations to allow such problems to be solved on a cost-effective desktop system.

ATHENA can be run on any DOS-based microcomputer capable of running SAS and supporting expanded memory. In order to take fullest advantage of the software's capabilities, ATHENA should be run on a powerful 80386-based workstation. One such workstation, on which ATHENA was developed, is the Load Data Analysis Workstation (LDAW). Designed and built for the Electric Power Research Institute (EPRI) by Quantum Consulting under research project 2342-8, the LDAW is a 25-megahertz 80386-based microcomputer with a 300-megabyte fast-access hard disk, 9-track tape drive, 80387 math coprocessor and VGA monitor. Figure 1 illustrates the LDAW's configuration. ATHENA's performance on this machine is quite impressive.

The purpose of a utility load research program is to gather information about how different customers of an electric utility use electricity. A growing number of programs are gathering information about how residential and commercial customers use particular appliances and pieces of equipment (end-uses). These data are gathered through the use of load monitors that record energy consumption for a particular appliance end-use for a particular customer in a given market sector. Such load monitors may have several channels, allowing multiple end-uses at a particular site (home, premise) to be monitored concurrently. Usually, data are downloaded once a week from the load monitor to a computer system, via a built-in modem. Large load research programs utilize hundreds of load monitors in order to collect enough data to support statistically significant results for each end-use in each market sector. A load research database, therefore, can be quite large, as it can consist of from 24 hourly observations to 288 5-minute observations per day from up to several hundred premises over the course of a year or more. Database sizes of several hundred megabytes are not uncommon.

The goal of load research program analyses is to interpret these gigantic raw databases into clear and reliable analysis results. An important part of this goal concerns the production of accurate end-use load shapes that characterize electric demand for that end-use on a particular daytype across the entire utility service territory. End-use load shapes are used by utilities to support a variety of different activities, including:

- rate design;
- load and energy forecasting;
- evaluation of demand-side management and conservation programs;
- market research and planning; and

suggestions about using the SUGI version of ATHENA, please turn to the appendix at the end of this paper.
Quantum Consulting designed the user-friendly data editing, validation and analysis environment called ATHENA to help utility staff produce reliable end-use analysis results from a raw load research database. This process begins with pre-analysis data validation and editing routines that assure accuracy and high data quality. The validated data can then be analyzed and displayed using a sophisticated assortment of custom load research software tools.

Because of the variety and complexity of possible end-use load data formats, ATHENA incorporates a general, flexible scheme for data input and management. In addition to supporting a wide range of existing utility formats for load datasets, ATHENA provides a program for creating user-defined data formats. These formats can be used to read new end-use load data from an ASCII file or to translate mainframe SAS datasets into SAS-PC datasets. Additionally, the data formats are used throughout ATHENA to read, write, and manipulate new and existing SAS-PC datasets. This flexible data management scheme allows ATHENA to use load, weather, customer, and engineering databases. 'To create a new format, the user enters the "System Setup" menu and chooses the "Create A Format" option. This program prompts the user for several important characteristics of the dataset to be created. The user then gives the format a name, and ATHENA stores that format permanently for future use.

ATHENA allows the user to perform a series of data validation and editing procedures to enhance data reliability. The data validation and editing routines allow a user to identify and eliminate a host of problems associated with the load data collection process. These can include:

- **Load Monitor Installation Problems.** A load monitor can be improperly installed or inaccurately calibrated, yielding bad data. Channels can also be improperly identified at installation time, leading to major data inaccuracies.
- **Load Monitor Operation Problems.** A load monitor, in constant use over long periods, can fail. In addition, a customer can unplug an end-use from a monitor channel, resulting in zero electricity usage measured for that channel.
- **Data Transmission and System Problems.** A source of error exists in the data transmission link between a load monitor and the utility computer system. The computer itself can crash, with the result that all data are lost during that time period. Finally, electricity surges on power lines are recorded by load monitors as electricity consumption on a channel, resulting in further data distortion.

Described below are two of ATHENA's software tools, VEWare and DCWare.

ATHENA's VEWare incorporates the following pre-analysis data validation and editing routines to test for data integrity and to help an analyst identify the problems outlined above:

- **Spike Detector.** This routine identifies and flags power surges.

Using these and other routines, a load research analyst can identify and remove a high percentage of bad data. The validated data set can then be used as input to ATHENA's DCWare load research analysis tools. These tools include:

- **Appliance Consecutive Use Pattern Analyzer (ACUPA).** This routine analyzes how an appliance is used and compares the calculated usage pattern to a predefined range of valid values.
- **Data Cross-Referencing.** This validation test allows the user to cross reference files and compare data from a load monitor to customer billing data. Additionally, the sum of the end-use channels can be cross referenced against a whole household channel.

ATHENA is implemented as a menu-driven front-end that consists of a series of SAS/AF® windows. Each window corresponds to a user selection from the main menu and allows the user to specify information relevant to the chosen function. This information is passed as a set of parameters to a SAS module (SAS macros and programs) called in an SCL submit block. This modular approach allows for easy expansion of ATHENA's analysis and display options.

ATHENA's interface allows the user to make selections by positioning the cursor on the desired option and pressing <Enter>. The user may position the cursor with the cursor control keys (or "arrow keys"), with the Tab key, or using a mouse. Using the mouse provides a very quick, intuitive means of making a selection. To implement the mouse interface, we adapted the method used in SAS/ASSIST™. The mouse is used to position the cursor on the first word of the desired option. The left mouse button is used to select that option. The right mouse button can be used to cancel the most recent action.

Among the most useful features of ATHENA are two utility programs built in to increase the flexibility of the program. The "Change Directory" function allows the user to change both the current drive and current directory in which ATHENA expects to find datasets on which it operates. The "List Datasets" function will list all existing SAS datasets in the current directory. These two functions can
be accessed from every menu screen of ATHENA, so that the user can look and move around the DOS file structure without exiting to the operating system.

"Change Directory" takes as its default value a current directory defined in the initialization block, which runs when the program is started. The utility makes use of the "Substring" and macro "Concatenation" functions to define a new directory.

"List Datasets" calls the SAS/AF function "Dirlist", which displays the resulting list in a new window. The "List Datasets" utility has been placed in its own program, which makes it possible to call that program from any other menu or program. To accomplish this, an additional window was defined from which the utility runs. By using the "Zoom" function, the window created by "Dirlist" occupies the entire screen, covering up the empty new window.

The load research tools outlined above create a powerful analysis environment. ATHENA also incorporates presentation graphics, so that analysis results can be easily viewed. The presentation graphics allow the user to:

- specify a graphing device;
- graph a connected load distribution;
- graph a set of load profiles; and
- graph a duty cycle distribution.

Finally, ATHENA provides easy support for other environments. A user can create Lotus 1-2-3 files, create ASCII "flat files," or exit to MS-Windows. One MS-Windows application for which ATHENA provides direct support is Quantum Consulting's Load Data Display Tool (LDDT). The LDDT provides a quick, convenient means of viewing a large number of end-use load profiles or duty cycle distributions. Although the LDDT can make use of unformatted data files, its full capabilities can only be realized with a "scripted" input file, which provides a legend, title, and overlay instructions to the application. A "scripted" file can be created using the "Create Load Data Display Tool" program in the "Display Results" menu. Once such a file has been created, the user can exit to the LDDT to view the newly created file. When the user exits from the LDDT (and MS-Windows), control is then automatically returned to ATHENA.

This transfer of control is accomplished through the use of a DOS batch file. ATHENA is run from a batch file which checks the system environment variable "errorlevel." When the user chooses to "Exit to LDDT," ATHENA issues an "abort return errorlevel" command to SAS in order to exit with a specific value for "errorlevel." Since the value of this variable is known to the operating system, the batch file can conditionally assign control to different programs based on the value of "errorlevel." The section of the batch file that starts the LDDT is followed by a command that restarts ATHENA.

ATHENA was initially designed using the window definitions in the SAS Macro Language. However, this approach led to execution time bottlenecks, as several hundred lines of code were interpreted before each execution. To reduce this overload, Quantum Consulting rewrote ATHENA in SAS/AF, calling relevant modules from an SCL submit block. SAS/AF is fast and is easily integrated with other SAS products. It constitutes an ideal choice for a front-end facility for this type of data analysis and display application.

ATHENA has been successfully applied to analysis of data from a number of utility end-use load research projects. Quantum Consulting continues to expand the capabilities of ATHENA and other SAS-based interactive analysis environments. We look forward to discussing with other SAS users opportunities to apply the techniques we have demonstrated in ATHENA to a broad variety of analysis and display tasks.

For more information about ATHENA, please contact Burtt Blodgett or John Powers at:
Quantum Consulting Inc.,
2030 Addison Street, Suite 410
Berkeley, CA 94704
415/540-7200

Appendix

USING THE SUGI VERSION OF ATHENA

ATHENA is entered through the SUGI 14 PC Poster Session shell. The title screen appears while the program initializes macro variables and warms up. Pressing <Enter> brings up an initial help screen.

To enter the main menu, press <F10>. The main menu contains options that will branch to the appropriate screen. To move around within the menu system, tab or position the cursor to an option and press <Enter>. This is very similar to the menu system in SAS/ASSIST, from which it has been adapted.

For the purposes of the SUGI conference, many of the computationally intensive tasks supported by ATHENA have been replaced by screens explaining their functions. We now describe the use of the remaining functions.

Most of the interesting features can be found in the Display Results window. The Display Load Profile option will plot a SAS/GRAPH® of mean air conditioner demand vs. time. The Graph Duty Cycle Distribution option will send a duty cycle histogram to the monitor. A duty cycle represents the percentage of time an appliance was on over a given time interval. The graph portrays the distribution of air conditioner duty cycles over a population at a given time. The height of each bar represents the percentage of the population that had an air conditioner duty cycle within a given range. For example, at 18:00, roughly 55 percent of the population had a duty cycle between 0.9 and 1.0, meaning that these air conditioners were in constant use.

The user can graph duty cycle distributions between 18:00 and 19:00 at 15-minute intervals by changing the time parameter. In addition, the height of the vertical axis can be changed by entering an integer between 0 and 100. The user can also enter a title for the duty cycle distribution.

The View Connected Load Distribution option displays a histogram of air conditioner connected load (in kW) over the appliance population. The height of each bar corresponds to the percentage of air conditioners that had a connected load within the given range. The connected load of an appliance is the maximum load drawn when the appliance is in full operation.
Other interesting features include the Analyze Impacts from Cycling option, accessed from the DCWare menu. This screen allows the user to analyze the load impact of an appliance cycling strategy. As default input parameters have been provided, simply press <Enter> to start this program. When the program finishes (this may take a minute or two), use <F5> to move to the PGM Output window to view the results. The results displayed in the Output window give the average load reduction achieved during the given time interval through a 7.5/15 cycling strategy. To return to ATHENA, use <F5> to move back to the Analyze Impacts from Cycling window.

Additional operating features include the Set Directory and List Datasets options, which allow the user to move around within the DOS file structure. If you plan to use the default datasets supplied with ATHENA, make sure to leave the current directory as it is initially set.

*SAS, SAS/AF, SAS/ASSIST, and SAS/GRAPH are registered trademarks of SAS Institute Inc., Cary, NC, USA.