A computerized analysis and reporting system was developed for personnel involved in effluent monitoring operations at the Oak Ridge Y-12 Plant managed by Martin Marietta Energy Systems, Inc., for the United States Department of Energy under contract DE-ACO5-84OR21400. SAS/ACCESS® interface to Rdb/VMS®, base SAS software, SAS/GRAPH®, SAS® BATCH, the macro facility, and some aspects of Digital Command Language (DCL®) were used in the development of the reporting and analysis system. This system essentially contains three modules, each of which has its own function. The basic functions include the creation of monthly reports for easy perusal for excursions; run charts, and the trend charts with the associated trend analyses. Figure 1 is a depiction of the process layout.

DCL (Digital Control Language) command procedures were used to facilitate the end user interface and to control the execution of the SAS programs that make up the system. DCL command procedures were chosen due to the overhead and time involved in learning SAS/AF® or SAS® PMENUs facility.

Methods

The initial step involves the update of a permanent SAS data set. The target data of interest resides in a public derived Rdb data base on a DEC VAX computer that is part of a larger DEC VAX cluster. This derived data base is updated every evening. SAS/ACCESS was used to create a view to the target data base. Within the view descriptor, a WHERE clause is used so that subsetting on a particular variable can take place to select only those records that are needed. A DCL command procedure controls execution of a daily SAS program that reads, sorts, and then merges a temporary copy of the data from the target data base with a permanent SAS data set excluding any duplicates that occur. This SAS data set is then transferred to another DEC VAX computer to allow other end users access to the data. This execution takes place early in the morning to facilitate convenient access to the data by the time the end user reports to work.

Once the permanent SAS data set has been updated, one of three modules can be invoked by the use of simple DCL commands. The first module, called REPORT, once invoked will prompt users to specify a year and month for request a particular report. The command REPORT is supplied at the DCL prompt, and the user will depress the return key. Several informational messages follow along with prompts for the end user to supply the year and month of the report that is to be produced. Next, the end users with access to different computers might have easy access to the data. This would reduce the size of the paper trail. The initial step in the process begins with an update of a permanent SAS data set from a public derived data base located on a DEC VAX computer. Once completed, one of three modules can be invoked via simple commands to compile monthly summary reports, run charts, and the trend charts with the associated trend analyses. Figure 1 is a depiction of the process layout.

DCL (Digital Control Language) command procedures were used to facilitate the end user interface and to control the execution of the SAS programs that make up the system. DCL command procedures were chosen due to the overhead and time involved in learning SAS/AF® or SAS® PMENUs facility.

Methods

The initial step involves the update of a permanent SAS data set. The target data of interest resides in a public derived Rdb data base on a DEC VAX computer that is part of a larger DEC VAX cluster. This derived data base is updated every evening. SAS/ACCESS was used to create a view to the target data base. Within the view descriptor, a WHERE clause is used so that subsetting on a particular variable can take place to select only those records that are needed. A DCL command procedure controls execution of a daily SAS program that reads, sorts, and then merges a temporary copy of the data from the target data base with a permanent SAS data set excluding any duplicates that occur. This SAS data set is then transferred to another DEC VAX computer to allow other end users access to the data. This execution takes place early in the morning to facilitate convenient access to the data by the time the end user reports to work.

Once the permanent SAS data set has been updated, one of three modules can be invoked by the use of simple DCL commands. The first module, called REPORT, once invoked will prompt users to specify a year and month for a specified monthly report. However, interested end users do not have to wait until the end of the month to produce monthly reports. The report module can be invoked at any time although the user must recognize that the monthly report is only preliminary and may not contain all of the monthly results. The execution of the SAS programs that make up this module is entirely controlled by a DCL command procedure. Figure 2 is an example of the typical steps a user might follow to request a particular report. The command REPORT is supplied at the DCL prompt, and the user will depress the return key. Several informational messages follow along with prompts for the end user to supply the year and month of the report that is to be produced. Next, the end users with access to different computers might have easy access to the data. This would reduce the size of the paper trail. The initial step in the process begins with an update of a permanent SAS data set from a public derived data base located on a DEC VAX computer. Once completed, one of three modules can be invoked via simple commands to compile monthly summary reports, run charts, and the trend charts with the associated trend analyses. Figure 1 is a depiction of the process layout.

DCL (Digital Control Language) command procedures were used to facilitate the end user interface and to control the execution of the SAS programs that make up the system. DCL command procedures were chosen due to the overhead and time involved in learning SAS/AF® or SAS® PMENUs facility.

Methods
user has the option of submitting a batch job or of running the program interactively. Once the entries have been supplied by the end user, the parameters are passed and pulled into a SAS program with the aid of the GETSYM function \( \text{GETSYM function} \) (i.e., \text{variable = input (getsym('P1'),informat)}). If batch execution is chosen, the program will be submitted as a batch job using the SASBATCH command. If interactive execution is chosen, then the end user must wait until the report is generated but will have the option of observing intermediate values of all parameters on the terminal screen or printing them on an attached printer. In either case, the report will be generated and placed in the end user's default directory on the DEC VAX. Figure 3 depicts a typical generic report layout. The locations, parameters of interest, the monthly average and maximum, as well as the NPDES limits, highlight the report.

The second module, called GRAPH, once invoked will first prompt the user to decide whether or not a help file containing a listing of the available codes is to be displayed on the terminal screen. Next, the end user has the option of displaying the data and the associated run graph on the terminal or of creating the data files and PostScript files to print them on an attached printer. As with the REPORT module, the execution of the SAS programs that make up the GRAPH module are controlled by a DCL command procedure. Once a mode of display has been chosen, a beginning date, an ending date, a particular parameter of interest, and a particular sampling point of interest are entered by the end user as prompted by the DCL command procedure. These parameters are passed to a SAS program by using the GETSYM function, which will create a run graph of the data for the specified time of interest. A listing of the data is also available for on-screen perusal or for printing on an attached printer. This module can be repeatedly invoked when the current session is completed. Figure 4, created by the GRAPH module, is a typical run chart that was printed on a QMS* PS 410 printer using the PSLL device driver. The NPDES limits were placed in a footnote at the bottom of the graph by using macro variables. A DCL command procedure was run to put the correct escape sequences into the PostScript file to allow the end user to simply type "TYPE filename" at the DCL $ prompt. Essentially, this puts the terminal into the controller mode, which transfers output to printer rather than the terminal. This procedure also works well with Hewlett-Packard Series II and Series III Laser Printers with an available PostScript cartridge.

The final module, called TREND, once invoked, will prompt the end user to supply specific information to start the trend procedure. As with the other two modules, DCL command procedures are used to control the execution of the SAS programs that make up the TREND module. The trend analysis in this module uses the Mann-Kendall (Kendall) nonparametric test for trend because the data are gathered at irregular intervals. References to this test may be found in Hollander and Wolfe (1973) and Gilbert (1987). The program that calculates the Mann-Kendall statistic uses some array processing, with the dimensions being passed as macro variables. Because equally spaced observations are needed for this test, the test is applied to the medians calculated on a monthly basis from the individual results from each sampled effluent. Medians were chosen for their robust properties, in particular their lack of sensitivity to "extreme" values. A SAS program is submitted using the SASBATCH command, which calculates and updates a smaller data set that contains the summary values for each month. The end user has no control over this program because it is executed automatically at regular intervals.

In the first step of this module, the end user is prompted to supply information that includes the year, month, parameter, and location of interest in which the trend analysis is to be performed. The end user also has the option of specifying terminal display of the data and associated analyses or of creating the text and graphic files so they may be printed by a local printer. Figure 5 is an example of a plot of a set of medians for a specified period of time. The associated Mann-Kendall test statistic is included in the footnote. The labels in the footnote were placed there by using macro variables. Due to the ever-present problem with ties in the data, the Mann-Kendall statistic is only approximate, and the user must refer to Table A.21 of the Kendall statistic found in Hollander and Wolfe (1973) or table A.18 of the Mann-Kendall statistic in Gilbert (1987). Interpretation of the results is left up to the end user in consultation with management and/or statistical consultant. The Mann-Kendall statistic computed from the data from Figure 5 is -47. The negative sign is indicative of the type of trend. Checking the result in the table leads to the conclusion of a significant negative trend.

Findings

During user testing conditions, it was pointed out that certain situations arise when more than one trend analysis or run graph is desired. Under these circumstances, the SAS system must be repeatedly invoked. To avert this unnecessary cost, the SAS macro facility, namely macro windows (%window), will be used where repeated invocation presently occurs. The current modification of the analysis system is still run by DCL command programs, but windows requesting user input are now displayed. Once finished with an analysis, the user is prompted for the next analysis without invoking the SAS system again.
Several users also pointed out that plots generated on a printer with PostScript capability took longer than anticipated. It was discovered that by using default fonts a plot, such as Figure 5, took an average of 1 minute and 16 seconds to be generated. Using a hardware font, the time to generate the same plot was cut to an average of 9.7 seconds. Figure 5 was generated by using the Times-Bold font available with the PSLL driver. Use of this font significantly reduces the time the user must wait for a plot to be generated.

Conclusions

Effluent monitoring is required to identify/determine where excursions in the data have occurred. Timeliness in reporting this information is important. The SAS System provides a tool whereby data can be quickly extracted from a difficult-to-use derived data base. This, in turn, provides end users with a system capable of simple analysis and reporting of the data in a timely fashion.

I would like to thank Sybil A. Adams, David G. Ball, Ron D. Graves, and Lou Ann Ladd for their comments and suggestions during the developmental and continual improvement phase of this application.

References


Acknowledgements

Hewlett-Packard, LaserJet series II and III are registered trademarks of Hewlett-Packard, Inc.

PostScript is a registered trademark of Adobe Systems, Incorporated.

OMS-PS 410 is a trademark of OMS, Inc.

SAS, SAS/ACCESS, SAS/BASICS, and SAS/GRAPH, are registered trademarks of SAS Institute Inc., Cary, NC, USA.

VAX, DCL, and Rdb are registered trademarks of Digital Equipment Corporation.
This program will run a SAS job that will report summarized outfall results for a requested month's end. You must be logged onto system xxx.

Note: Be sure that your CAP LOCK key is on.

Enter a year and month in the form 1991 and JAN when asked or hit return to quit.

(Hit return after entering the year) Year = 1991
(Hit return after entering the month) Month = JUL

Enter SRO to create a summary report only. This will be submitted as a batch job and free up the terminal. Enter SRL to look at a list of the individual data values and create a summary report. With this option, your terminal will be tied up while you wait for the prompt to view or print the individual results.

Enter Here ===> Y

Job REPORT (queue xxxxx, entry xxx) started on xxxxx

This will take a few minutes, but when done, the report will be located in your directory and be called REPORT01.DAT.....Have a nice wait!

(DCL Prompt) $
### Monthly Effluent Reports
for Outfall xxx

<table>
<thead>
<tr>
<th>Year</th>
<th>Month</th>
<th>Parameter</th>
<th>Monthly Average</th>
<th>Monthly Max</th>
<th>NPDES Average</th>
<th>NPDES Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>1991</td>
<td>JUL</td>
<td>xxxxxxxx</td>
<td>xxxxx</td>
<td>xxxxx</td>
<td>xxxxx</td>
<td>xxxxx</td>
</tr>
<tr>
<td>1991</td>
<td>JUL</td>
<td>yyyyyyy</td>
<td>yyyy</td>
<td>yyyy</td>
<td>yyyy</td>
<td>yyyy</td>
</tr>
<tr>
<td>1991</td>
<td>JUL</td>
<td>zzzzzzzz</td>
<td>zzzzz</td>
<td>zzzz</td>
<td>zzzz</td>
<td>zzzz</td>
</tr>
</tbody>
</table>

(etc)

**Figure 3. Example of generic monthly report.**

### Effluent Data
**Parameter - SULFATE**

![Graph showing sulfate levels from August 31, 1990, to January 13, 1992.](image)

- **NPDES Average - na**
- **NPDES Max - na**

indicates a value below detection limits

**Figure 4. Example of a parameter run chart.**
The Mann-Kendall statistic is -47 for the time period starting SEP 1990. The number of months in the sample is 11.

Figure 5. Example of the trend analysis.