ABSTRACT

Flexibility is always an important part of any database system and SAS® has provided the needed tools. In the past, forecasts were loaded and analyzed in a non-standard format using spreadsheets. These spreadsheets were cumbersome, time consuming, and resulted in the forecaster spending the majority of his/her time on mechanical data processes rather than concentrating on modeling/analysis functions. Using SAS/AF®, SAS/FS®, SAS/STAT®, SAS/IML®, SAS/ETS® and a host of SAS procedures, the Usage Forecasting System (UFS) was developed to provide the ability to load, evaluate, validate, edit, and forecast over 480 items.

This system is divided into four basic modules consisting of ACTUALS, MODELS, FORECASTS, and REPORTS. Actuals are loaded through ASCII files or direct data set editing. Multiple models are allowed for each of the 480 items forecasted. The multiple models for a particular combination are weighted based on past performance and then combined into a single forecast. The forecasted units are merged with a "disaggregation data set" to forecast additional items that have little variability over time, yet provide additional detail that is sometimes required. The REPORT module provides the systems output file and report generation capability.

INTRODUCTION

This paper describes the GTE Telephone Operations Usage Forecasting System (UFS) developed using SAS/AF, SAS/FSP, SAS/STAT, SAS/IML, SAS/ETS and extensive SAS Macro techniques under MS-DOS. GTE provides service to over 15 million access lines in 40 states. These lines generate calls which are measured in minutes of use (MOUs). These MOUs are the basis for GTE's access revenues and a part of the overall budget planning process. This usage is recorded and classified into four basic product categories and three jurisdictions for forecasting purposes. The input MOU data is processed and forecasted for 480 different state/product/jurisdiction combinations using a multitude of SAS procedures.

Several different econometric and time-series modeling techniques are applied to each specific state/product/jurisdiction forecast item, and these results are combined using a combinatorial forecast methodology. The current data processing for this system is done on a dedicated MS-DOS PC at the GTE Telops headquarters. Some users are remotely located in other states, therefore authorized remote access has been provided to these users.

APPLICATION REQUIREMENTS

The design goals for the UFS system required extensive preliminary discussion with both the potential users of the system as well as with the groups that would be passed the final forecasted output. We wanted a system that had a reasonably fast response time, modular system design, incorporated sophisticated econometric modeling and graphics, included flexible reporting and was easy to use. Mainframes were considered since the company has several; however, these machines are heavily used around the clock for customer bill processing. Because of the heavy workloads, SAS is limited to batch processing on the mainframe, which gives very poor response time for end users. Therefore, it was decided to use a dedicated personal computer. Since the input usage data was sourced from may different systems, ASCII files were chosen as a common starting point. In many cases, users wanted to "adjust" the data in spreadsheets before handing it off to the system. Therefore, standard input formats and naming conventions were established for all ASCII data that would be converted into SAS data sets. Also, standard directory structures were chosen for loading input data and for output data on the PC. The modular design concept would allow each section of the system to be fine tuned independently and allow for flexibility when adding future enhancements. Data security was also a consideration both on site and with the ability for remote access. Password protection for access to certain files with SAS was a needed feature. The system needed the capability to forecast 480 items directly and apply additional factors to explode the base forecasts into over 20,000 elements.

SYSTEM OVERVIEW

UFS is divided into 4 main modules (Figure 1).

The first module is classified as ACTUALS. This unit consists of loading and verifying ACTUALS (historical data) such as MOUs, customer lines, prices, and economic and demographic data from Data Resources Inc. (DRI) for each state, jurisdiction, and product. Also loaded are the drivers (forecasted lines, prices and DRI variables) used in certain models. Several SAS/AF sub
menus are provided in this module. The choices require limited user input and provide authorized users (password protected) with the capability to load and update the existing monthly MOU SAS database using the UPDATE statement. The input data is standardized with users being required to edit the data set directly, provide a new SAS data set, or load a print file containing specific variables. Also, the data series can be expanded from monthly to quarterly for modeling using PROC EXPAND. PROC X11 is used to seasonally adjust the data, fill missing values, and smooth extreme outliers. Reports which aid in the evaluation of actuals can be obtained through the main reports menu under the ACTUALS MENU selection.

The second module is called MODELS, where both econometric and judgmental models are estimated, evaluated, and used to generate forecasts. This module generates forecasts using the system canned models and/or the judgmental forecasts. A set of standard models have been developed for each specific state/product/jurisdiction combination. These models may be of the type classified as Time-series Cross-sectional Pooled, OLS, Mixed Estimation, Winters, ARIMA and Polynomial Distributed Lag. The first part of the process involves running the models to estimate the parameters. The parameters and their statistics are reviewed until they are determined to be within acceptable bounds. The parameters are then applied to determine the forecasted volumes and are again examined. Finally, the judgmental forecast is loaded into the system.

The FORECAST module contains several sub menus allowing the system further flexibility. Once the individual forecasts are complete and all factors are finalized, we then select a job to estimate the combination of forecasts weights (ESTIMATE WEIGHTS option) for each model. These weights are based on the model's past performance (there can be up to 7 different models for each state, jurisdiction and product, including the judgmental model). We then select the number of historical periods to be used in the weight estimation process. Using the parameters above, the system then produces a forecast for each model that excludes the time period specified earlier. In effect, the system forecasts the actuals and compares them to the historical numbers for the excluded period to determine the weight of each model for a specific state/product/jurisdiction combination. For example, California, Interstate, Message Toll Service might use Pooled, OLS with autocorrelation correction, Mixed Estimation, Winters, Polynomial Distributed Lag, and Judgmental models. The performance of these models would be weighted on how well they can predict the actuals during the omitted time period.

Next, the system allows us to calibrate the models to a particular point in time, if required.

Then, we select the menu option GENERATE AGGREGATE FORECAST, which initiates code that weights the individual forecasts, resulting in one forecast series for each particular state, jurisdiction, and product. Thus, using the weights determined in the system and using our example above, the six models for California, Interstate, Message Toll service would be combined into one forecasted number.

Although only 480 items are forecasted directly, additional forecasted detail is often requested on specific products within these 480 items. These more detailed products have very little variability within a particular year, but may vary over several years. Therefore, to provide a practical solution to produce this level of detail with the limited resources available, a "FACTOR DATA SET" is used which contains annual factors that are applied to the primary forecasted MOUs. Users within this module can select an option to edit/update factors in using SAS/FSEDIT screens that are applied to the final forecast to generate additional forecasted detail (20,000 plus items). The factors are expressed as percents of the total volume for each state/product/jurisdiction combination (Figure 3).

Several data validation checks are used within the FSEDIT code to ensure that data integrity has not been violated (Figure 4). In addition, we have allowed for a manual overlay of the data in certain periods for special circumstances.

The last step in this module is to initiate the GENERATE DETAILED FORECAST option which combines the final forecasted series, detailed factors, and data overlays to produce 5 years of annual and 3 years of monthly forecasts in detail (20,000 plus records).

The REPORTS module is divided into three major divisions. The ACTUALS report menu provides a means to validate and evaluate the raw MOU data. Several different product based report combinations at both the summary and detailed level are available. You would enter basically the year for which you want the summary and statistic report to apply.

The MODELS REPORT menu provides access to annual, monthly, and user specified period reports. There is the flexibility of providing formatted report output to a printer or print file type output which can be readily imported into LOTUS and is stored in the standard output directory on the PC hard disk. SAS "PUT" and "FILE"DRIVE:DIRECTORY FILENAME" statements are used for all print file creations. Forecast comparisons are also an important part of the report process in evaluating the model output.

The FORECAST REPORT menu provides you with a series of reports to track the forecast performance during the
forecasted period.

CONCLUSION

The UFS system has indeed met and exceeded its design goals and is constantly evolving. Further enhancements are being considered including improving the man-machine interface design, more sophisticated modelling, and greater use of graphical output for forecast analysis and result reporting. There has also been consideration given to splitting the system into an area structure, where the United States is divided into four geography serving areas. Each area would have a duplicate copy of the UFS system loaded on a PC at the area headquarters containing data that applies to the states within that particular area. Each area could run the system independently with the final output being sent to the main PC database at the headquarters for a series of summarization processes.

REFERENCES

Tom Fischetti
Tel. # (805) 372-6808
FAX (805) 373-1156

Steve Heathcote
Tel. # (805) 372-8957
FAX (805) 373-1156

Don Perry
Tel. # (206) 261-6677
FAX (206) 258-5444

SAS, SAS/AF, SAS/FSP, SAS/STAT, SAS/IML, and SAS/ETS are registered trademarks of SAS Institute, Inc. in the USA and other countries. * indicates USA registration.

Other brand and product names are registered trademarks or trademarks of their respective companies.
Figure 1

USAGE FORECAST SYSTEM
Menu Layout

MAIN MENU

ACTUALS
DATA LOAD
SEASONAL ADJUSTMENT
QUARTERLY UPDATE

MODELS
ESTIMATE FORECAST
JUDGEMENT

FORECAST
FACTORS MENU
T/JO RATIOS
MONTHLY FACTORS
OVERLAYS
AVG. HAUL & TT
ESTIMATE WEIGHTS
GEN.AGG.FORECAST
GEN.DET.FORECAST

REPORTS
ACTUALS MENU
ACTUALS MENU
MODELS MENU
FORECAST MENU
Enter the Jurisdiction and Traffic Type Below:

Jurisdiction  
1 = Interstate
2 = Interstate/IntraLATA
3 = Intrastate/InterLATA
4 = Intrastate/IntraLATA
5 = Local

Traffic Type  
1 = Orig. MTS
2 = Term. MTS
3 = Orig. 800
4 = 800 Readyline
5 = Term. 800
6 = OutWATS

Hit the PgDn key to continue

Place an 'X' next to the model(s) you wish to estimate

- Pooled
- State Specific (PROC AUTOREG)
- Mixed Estimation (Svc Ratio)
- Mixed Estimation (Koyck Lag)
- Winter's
- PDL

Hit the PgDn key to continue

Do you want to save the parameters you estimate (Y,N) _

NOTE: DO NOT CHOOSE TO SAVE THE PARAMETERS UNLESS YOU ARE SURE AND YOU HAVE REVIEWED THEM THOROUGHLY!!

Enter password: 

Hit F10 to process or F3 to cancel
## Figure 3

<table>
<thead>
<tr>
<th>Combination 1</th>
<th>ST</th>
<th>Jurisdiction</th>
<th>Product</th>
<th>Rate</th>
<th>Factor</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>CA</td>
<td>Interstate Interlude</td>
<td>Message Toll</td>
<td>Non-Premium</td>
<td>.30</td>
</tr>
<tr>
<td></td>
<td>CA</td>
<td>Interstate Interlude</td>
<td>Message Toll</td>
<td>EOS 1 Premium</td>
<td>.10</td>
</tr>
<tr>
<td></td>
<td>CA</td>
<td>Interstate Interlude</td>
<td>Message Toll</td>
<td>EOS 2 Premium</td>
<td>.60</td>
</tr>
</tbody>
</table>

100%

<table>
<thead>
<tr>
<th>Combination 2</th>
<th>ST</th>
<th>Jurisdiction</th>
<th>Product</th>
<th>Rate</th>
<th>Factor</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>CA</td>
<td>Intrastate Interlude</td>
<td>Message Toll</td>
<td>Non-Premium</td>
<td>.50</td>
</tr>
<tr>
<td></td>
<td>CA</td>
<td>Intrastate Interlude</td>
<td>Message Toll</td>
<td>EOS 1 Premium</td>
<td>.20</td>
</tr>
<tr>
<td></td>
<td>CA</td>
<td>Intrastate Interlude</td>
<td>Message Toll</td>
<td>EOS 2 Premium</td>
<td>.30</td>
</tr>
</tbody>
</table>

100%

## Figure 4

### EDIT WORK.TESTX

Command ===>

<table>
<thead>
<tr>
<th>ENTITY: California STATE: CA TT: Orig MTS JURIS: IS/IL</th>
</tr>
</thead>
<tbody>
<tr>
<td>TOTALS 1.000000 1.000000 1.000000 1.000000 1.000000 1.000000 1.000000 1.000000</td>
</tr>
<tr>
<td>%AT&amp;T (IS/IL) 0.854717 0.854717 0.854717 0.854717 0.854717 0.854717 0.854717 0.854717</td>
</tr>
<tr>
<td>%OCC (IS/IL) 0.145283 0.145283 0.145283 0.145283 0.145283 0.145283 0.145283 0.145283</td>
</tr>
<tr>
<td>%PREM (IS/IL) 1.000000 1.000000 1.000000 1.000000 1.000000 1.000000 1.000000</td>
</tr>
<tr>
<td>%N-PR (IS/IL) 0.000000 0.000000 0.000000 0.000000 0.000000 0.000000 0.000000 0.000000</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>OMTS</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>OFX</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>OMTS</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>OFX</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>OMTS</td>
<td>0.854717</td>
<td>0.854717</td>
<td>0.854717</td>
<td>0.854717</td>
<td>0.854717</td>
<td>0.854717</td>
<td>0.854717</td>
</tr>
<tr>
<td>OMTS</td>
<td>0.145283</td>
<td>0.145283</td>
<td>0.145283</td>
<td>0.145283</td>
<td>0.145283</td>
<td>0.145283</td>
<td>0.145283</td>
</tr>
</tbody>
</table>