Paper 153-30

USING SAS® TO DETECT ERRORS IN X12-ARIMA SEASONAL ADJUSTMENT

Nicole Brooks, U.S. Bureau of Labor Statistics, Washington, DC

ABSTRACT

This paper features Base SAS and the X12-ARIMA Seasonal Adjustment program developed by the U.S. Bureau of the Census to seasonally adjust monthly or quarterly time series. X12-ARIMA commonly encounters problems with the time series used as input to the program and generates errors when problems occur. This paper illustrates how SAS can be used to detect errors generated by X12-ARIMA and to notify the user when these errors occur. In a production situation, X12-ARIMA is used to process a high volume of time series in a limited amount of time. This paper focuses on using the INDEX function in the DATA statement along with the SAS macro facility to reduce the amount of time taken to detect errors when processing a large number of time series with X12-ARIMA.

INTRODUCTION

X12-ARIMA is a seasonal adjustment program developed by the U.S. Bureau of Census to seasonally adjust monthly or quarterly time series. X12-ARIMA is available as an executable program for PC microcomputers running DOS, Sun 4 UNIX workstations, and VAX/VMS computers. To apply X12-ARIMA to a particular time series, a main input file, called a specification file, must be created. This ASCII (or "text") file contains a set of specifications or specs that X12-ARIMA reads to obtain the information it needs about the time series data, the time series model to be used, the analysis to be performed, and the output desired. The SAS System is often used to create spec files, as well as, execute X12-ARIMA in a production environment. The details of these procedures are beyond the scope of this paper.

In a production situation, it is essential to run more than one series in a given X12-ARIMA run. Before X12-ARIMA can run multiple series, a *metafile* must be created. This is an ASCII file which contains the names of the files to be processed by X12-ARIMA. If an error occurs in one of the spec files in a metafile run, the program will print the appropriate error message to another ASCII file called the error file. Execution will stop for that series and the program will continue processing the remaining spec files. A listing of all the spec files with errors is given in another ASCII file, called the X12-ARIMA log file. This paper illustrates how to use the SAS System to detect errors generated by X12-ARIMA in an automated production environment.

DETECTING ERRORS IN X12-ARIMA

When X12-ARIMA encounters a problem with the input data or discovers an error in the input specs, it reports these errors as they are discovered in the error file. The error file contains specific information about any problem X12-ARIMA encounters. In a production situation where you need to run X12-ARIMA on a large quantity of series, reviewing each individual error file for errors is very time consuming. The user needs to review the error file only when X12-ARIMA finds an error; otherwise, the error file is blank. The X12-ARIMA log file is a more useful source for the initial detection of errors.

Every time X12-ARIMA is run, a log file is produced where a summary of modeling and seasonal adjustment diagnostics, as well as error messages are stored for each series or spec file processed. The error messages stored in the X12-ARIMA log file are general messages that direct the user to the appropriate error file for more details on the problem. The X12-ARIMA log file is stored as an ASCII file and the error messages stored in it follow this format:

```
Error messages for the input files defined in \path\filename.mta
Input or runtime errors were found in the following files:
\path\filename.spc  (Error messages stored in \path\filename.err)
```

In a production situation, it is possible to have numerous X12-ARIMA log files since it is necessary to run multiple time series in various groups according to input specifications. In this situation, reviewing each X12-ARIMA log file could also be very time consuming. So, the next section illustrates how to use the SAS System to reduce the amount of time necessary to detect these errors.

USING THE SAS SYSTEM TO DETECT X12-ARIMA ERRORS

When X12-ARIMA discovers an error, an error message with the text string 'Error messages for the input files' is printed to the X12-ARIMA log file. X12-ARIMA prints a separate error message beginning with this text string on each line for the individual time series that have errors. The X12-ARIMA log file can be read with the INPUT statement in the DATA step. The INDEX function in the DATA step can then be used to search a character expression for a string of characters. The INDEX function is used to search for the text string 'Error messages for the input files':

```
data errormsg;
```

```
        infile "\path\filename.log" length=linelen;
        input @1 txtln $varying150. linelen @;
        errormsg = index(txtln,'Error messages stored');
        if errormsg > 0;
    run;
```

If the specified text string is found, INDEX returns the position of the string's first character. If the string is not found, INDEX returns a value of 0. Once an error message in the X12-ARIMA log file has been found by the INDEX function, the line containing the error message for each series is stored in a SAS data set using the sub setting IF statement. The SUMMARY procedure is then used to find the frequency that the error messages occur. PROC SUMMARY creates the automatic variable _FREQ_, which stores a count of the observations summarized in a data set. Once the frequency of the errors has been found the CALL SYMPUT routine can be used to assign the frequency to a macro variable for use in the next step:

```
    proc summary data=errormsg nway;
        output out=no;
      run;
    data _null_;
      set no;
      call symput ("freq",_freq_);
    run;
```

Next, iterative DO loop processing and the CALL SYMPUT routine is used to assign each individual error message stored in the SAS data set to a macro variable. These macro variables can then be used as input to a variety of notification options:

```
    data _null_;
      set errormsg;
      %do i=1 %to &freq;
        if _N_=&i then do;
          call symput ("errormsg&i",left(trim(txtln)));
        end;
        put &errormsg&i;
      %end;
    run;
```

One method of X12 error notification using SAS is to use the PUT function in the DATA step, which can be used to display the X12-ARIMA error message in the SAS log file. This method is shown in the DATA step above. This method requires the user to have access to the SAS log file. In a production situation, the SAS System may be executed remotely and the user may not have immediate access to the SAS log file. In this case another means of notification is necessary.

Another method of notification is the FILENAME, EMAIL (SMTP) access method, which allows you to send electronic mail programmatically from SAS using the SMTP (Simple Mail Transfer Protocol) e-mail interface. This method is used to send the X12-ARIMA error messages stored in the macro variables to the user through e-mail:

```
    filename outbox email "Brooks.Nicole@bls.gov"
      subject="X12-ARIMA Seasonal Adjustment Error";
    data _null_;
      file outbox;
        %do i=1 %to &freq;
          put "&&errormsg&i";
        %end;
    run;
```

EXAMPLE: X12-ARIMA ERROR DETECTION USING SAS

Suppose that you have multiple X12-ARIMA log files to review after running the X12-ARIMA Seasonal Adjustment program. Each of these X12-ARIMA log files contains seasonal adjustment diagnostics and error messages for errors found in the input data or spec files for multiple time series. In this case, you would need to use the SAS macro facility to conditionally generate the SAS code repetitively to review multiple X12-ARIMA log files with an iterative %DO loop. In this example, we have (20) X12-ARIMA log files to review for errors:

```
    %macro errorx12;
      %do n=1 %to 20;
      data errormsg;
        infile "\path\filename&n.log" length=linelen;
        input @1 txtln $varying150. linelen @;
        errormsg = index(txtln,'Error messages stored');
        if errormsg > 0;
      run;
```

2

```
      proc summary data=errormsg nway;
        output out=no;
      run;
      data _null_;
        set no;
        call symput ("freq",_freq_);
      run;
      %if &freq ne 0 %then %do;
        data _null_;
          set errormsg;
          %do i=1 %to &freq;
            if _N_=&i then do;
              call symput ("errormsg&i",left(trim(txtln)));
            end;
          %end;
        run;
        filename outbox email "Brooks.Nicole@bls.gov"
          subject="X12-ARIMA Seasonal Adjustment Error";
        data _null_;
          file outbox;
            %do i=1 %to &freq;
                put "&&errormsg&i";
            %end;
        run;
      %end;
    %end;
    %mend;
    %errorx12
```

Once, the source code has been included in either a local or remote SAS session, the user is notified by e-mail of any error messages generated in the X12-ARIMA log files.

CONCLUSION

In a production situation, errors need to be detected in a timely manner. SAS can be used to reduce the amount of time necessary to detect errors when processing a large number of time series with X12-ARIMA. This paper recommends using three components of SAS to detect errors in X12-ARIMA: 1) The INDEX function in the DATA step is useful for detecting standardized text strings, such as the error messages generated by X12-ARIMA. 2) The macro facility in SAS is useful for reviewing numerous X12-ARIMA log files with multiple error messages in a short amount of time. 3) If X12-ARIMA errors are detected, SAS provides various means of notification to the user. In an automated production environment, electronic notification through the SAS SMTP e-mail interface is useful to inform the user of an X12-ARIMA error.

REFERENCES

U.S. Bureau of the Census (1996), *X-12-ARIMA Reference Manual*, *Beta Version 1.0*, Washington, DC

SAS Institute Inc. (1999), *SAS® Language Reference: Dictionary, Version 8*, Cary, NC: SAS Institute Inc.

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author at:
> Nicole Brooks
> U.S. Bureau of Labor Statistics
> 2 Massachusetts Avenue, NE, Room 4860
> Washington, DC 20212
> Work Phone: 202-691-5097
> E-mail: Brooks.Nicole@bls.gov

SAS and all other SAS Institute Inc. product or service names are registered trademarks or trademarks of SAS Institute Inc. in the USA and other countries. ® indicates USA registration.

Other brand and product names are trademarks of their respective companies.