High Performance Forecasting with ITRM 3.2 and Subsequent Releases

Author: Patrick Curren

Introduction 3
Documentation 3
Case Study Example 3
Input Data 4
Enterprise Guide Process Flow 4
HPFENGINE Icon 5
Line Plots of Forecasts 11
Line Plot Icon 11
Output from the Line Plot Icon 13
HTML-HPFENGINE Icon 14
Conclusion 15
Appendix 1 (HPFENGINE code) 16

Introduction

Beginning with SAS IT Resource Management 3.2, SAS High-Performance Forecasting (HPF) is included in the solution. SAS High-Performance Forecasting software provides a large-scale system that automatically selects time series models for forecasting time-series data.

This paper will demonstrate use of the HPFENGINE procedure. The HPFENGINE procedure provides an automatic way to generate forecasts for many time series or transactional data in one step. The procedure can automatically choose the best forecast model from a user-defined model list or a default model list.

SAS Forecast Server, a robust, large-scale forecasting technology package that includes SAS Forecast Studio, an interactive graphical interface, SAS/ETS®, SAS High-Performance Forecasting, SAS/GRAPH, and SAS Integration Technologies is available from SAS and can be used to more easily perform the type of analysis discussed in this technical paper.

Documentation

The HPENGINE procedure is documented in the SAS High-Performance Forecasting 3.1 User's Guide. Information for some parameters of this procedure will be found in SAS/ETS 9.2 User's Guide. At the present time, Enterprise Guide does not have GUI pull-down menus to support HPF. You must create your solution using SAS code.

Case Study Example

We will demonstrate the features of HPF with an example of forecasting MIPS Used by LPAR, Shift and Week. A weekly forecast will be created for each LPAR/Shift combination. We have 16 LPARs that we will report on in this example. We will create a forecast for 12 weeks into the future.

Input Data

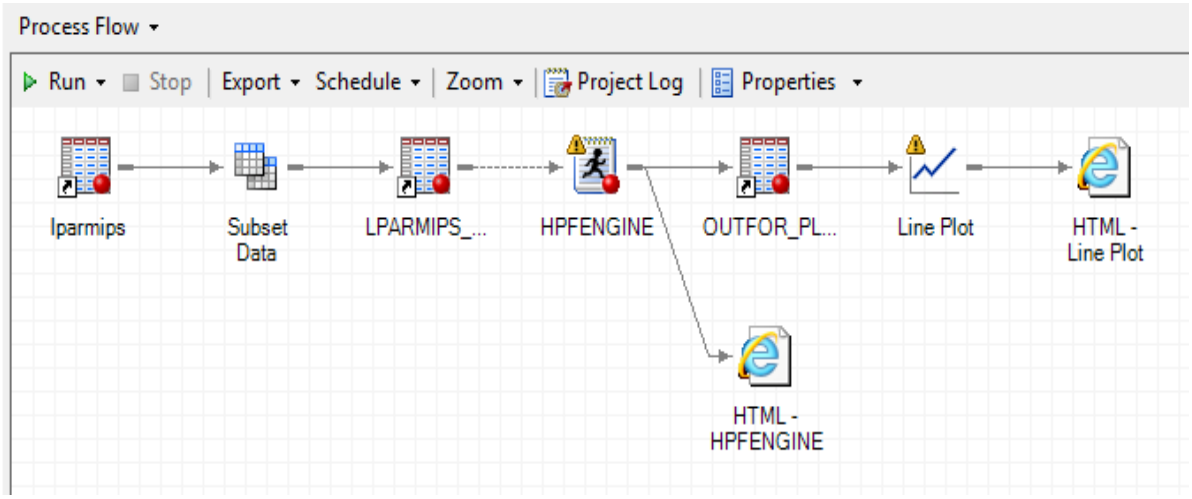
A snippet of the input data is shown here.

	DATE	SHIFT	MIPSUSED	LPARNAME
623	04SEP2011	1	1687.12	LPAR05
624	11SEP2011	1	1752.97	LPAR05
625	18SEP2011	1	1711.97	LPAR05
626	25SEP2011	1	1826.89	LPAR05
627	02OCT2011	1	1672.28	LPAR05
628	09OCT2011	1	1824.60	LPAR05
629	16OCT2011	1	1769.96	LPAR05
630	06FEB2011	2	1759.15	LPAR05
631	13FEB2011	2	1852.31	LPAR05
632	20FEB2011	2	1886.60	LPAR05
633	27FEB2011	2	1811.94	LPAR05
634	06MAR2011	2	1779.23	LPAR05
635	13MAR2011	2	1692.43	LPAR05

The data is sorted by LPARNAME-SHIFT-DATE

Enterprise Guide Process Flow

The Process Flow for this forecasting project is shown below.



Icon Descriptions :

- lparmips – input data
- Subset Data - Used to remove the Physical LPAR entries and sort the data
- LPARMIPS_... - Output from the Subset Data step
- HPFENGINE - SAS code to use the HPFENGINE procedure

- OUTFOR_PL … - Output from the HPFENGINE SAS code
- Line Plot – Line Plots of Forecasts
- HTML-HPFENGINE – Report output from the HPFENGINE procedure
- HTML – Line Plot - Out Forecast Plot Report

HPFENGINE Icon

This Icon represents the SAS Code to create the forecast data. The complete SAS code is shown in Appendix 1 of this document.

Forecast Code Part 1

```
options nofmterr;
ods graphics on;

%global forecastdate;
proc sql noprint;
  select max(DATE) into :forecastdate
  from WORK.LPARMIPS_DATA;
quit;
%put forecastdate = %sysfunc(putn(&forecastdate,date9.));

proc sort data=WORK.LPARMIPS_DATA out=WORK.LPARMIPS_DATA_SORT;
  by LPARNAME SHIFT DATE;
  run;
```

This section of the code obtains the last date from the input data and sets this to our forecastdate variable to be used later in the process. In our case study the SAS log shows “forecastdate = 16OCT2011”. Since the HPFENGINE procedure expects data to be presented in a time series format, the next step in our method is to sort our data accordingly.

Forecast Code Part 2

```
proc hpfengine data=WORK.LPARMIPS_DATA_SORT
      lead=12
      globalselection=tsfsselect
      task=select(criterion=MAPE alpha=0.10 override)
      print=(select summary)
      plot=forecasts
      outsum=work.outsum
      outfor=work.outfor
      out=work.out;
   by LPARNAME SHIFT;
   id DATE interval=week;
   forecast MIPSUSED;
run;
```

This is the most important piece of code for this project. This procedure does all the the forecasting work for you. An explanation of each parameter for our case study follows.

data=WORK.LPARMIPS_DATA_SORT

The input data to be used for the forecast.

lead=12

Specifies the number of periods ahead to forecast (forecast lead or horizon). In our case we want to forecast 12 weeks into the future. We thought projecting 12 weeks in the future would be enough lead time to plan for changes to the configuration.

globalselection=tsfsselect

Many models can be tried for each time series in the data. The globalselection parameter specifies the name of a catalog entry that contains the model selection list. Tsfsselect is a rich collection of models provided by the HPF system.

task=select(criterion=MAPE alpha=0.10 override)

- The criterion option specifies the model selection criterion (statistic of fit) to be used to select the best model from several candidate models. MAPE or mean absolute percent error is one of the most commonly used selection criterion.

- The alpha option specifies the significance level to use in computing the confidence limits of the forecast. In our case the alpha=0.10 produces 90% confidence intervals.
- The override option forces the use of any options listed.

print=(select summary)

Specifies the printed output desired.

- SELECT prints the label and fit statistics for each model in the selection list.
- SUMMARY prints the forecast summary.

plot=forecasts

Specifies the graphical output desired. By default, the HPFENGINE procedure produces no graphical output.

- FORECASTS plots forecast graphics.

outsum=work.outsum

The OUTSUM= data set records the summary statistics for each variable specified in a FORECAST statement. Variable _LEAD12_ is the forecasted MIPS Used 12 weeks in the future.

	LPARNAME	SHIFT	_LEAD12_	_NAME_
1	LPAR01	1	99.876890407	MIPSUSED
2	LPAR01	2	130.10288486	MIPSUSED
3	LPAR01	3	84.261743195	MIPSUSED
4	LPAR01	4	71.349683582	MIPSUSED
5	LPAR02	1	159.57893861	MIPSUSED
6	LPAR02	2	122.21368912	MIPSUSED
7	LPAR02	3	239.36152807	MIPSUSED
8	LPAR02	4	111.09621886	MIPSUSED

outfor=work.outfor

Names the output data set to contain the forecast time series components (actual, predicted, lower confidence limit, upper confidence limit, prediction error, and prediction standard error). The OUTFOR= data set is particularly useful for displaying the forecasts in tabular or graphical form. We will use this dataset to create the MIPS Forecast by LPAR-SHIFT line plot charts. Below are a few entries from this dataset.

	LPARNAME	SHIFT	_NAME_	DATE	ACTUAL	PREDICT	LOWER	UPPER
31	LPAR01	1	MIPSUSED	Sun, 4 Sep 2011	95.311189855	87.27684463	.	.
32	LPAR01	1	MIPSUSED	Sun, 11 Sep 2011	95.903633677	90.992469025	.	.
33	LPAR01	1	MIPSUSED	Sun, 18 Sep 2011	102.58735255	93.22949181	.	.
34	LPAR01	1	MIPSUSED	Sun, 25 Sep 2011	113.43662711	97.579458566	.	.
35	LPAR01	1	MIPSUSED	Sun, 2 Oct 2011	97.523034667	105.01874576	.	.
36	LPAR01	1	MIPSUSED	Sun, 9 Oct 2011	99.987912655	101.38115534	.	.
37	LPAR01	1	MIPSUSED	Sun, 16 Oct 2011	101.15362275	100.63653483	90.909021115	110.36404855
38	LPAR01	1	MIPSUSED	Sun, 23 Oct 2011	.	100.7979882	91.07047449	110.52550192
39	LPAR01	1	MIPSUSED	Sun, 30 Oct 2011	.	100.71425204	89.946589958	111.48191412
40	LPAR01	1	MIPSUSED	Sun, 6 Nov 2011	.	100.63051588	88.912871698	112.34816006
41	LPAR01	1	MIPSUSED	Sun, 13 Nov 2011	.	100.54677971	87.948903672	113.14465576
42	LPAR01	1	MIPSUSED	Sun, 20 Nov 2011	.	100.46304355	87.040954433	113.88513267
43	LPAR01	1	MIPSUSED	Sun, 27 Nov 2011	.	100.37930739	86.179264635	114.57935014
44	LPAR01	1	MIPSUSED	Sun, 4 Dec 2011	.	100.29557122	85.356604159	115.23453829
45	LPAR01	1	MIPSUSED	Sun, 11 Dec 2011	.	100.21183506	84.567440305	115.85622982
46	LPAR01	1	MIPSUSED	Sun, 18 Dec 2011	.	100.1280989	83.807427722	116.44877007
47	LPAR01	1	MIPSUSED	Sun, 25 Dec 2011	.	100.04436273	83.073079915	117.01564555
48	LPAR01	1	MIPSUSED	Sun, 1 Jan 2012	.	99.96062657	82.361549078	117.55970406
49	LPAR01	1	MIPSUSED	Sun, 8 Jan 2012	.	99.876890407	81.670473562	118.08330725
50	LPAR01	2	MIPSUSED	Sun, 6 Feb 2011	100.36870667	115.48125445	.	.

out=work.out

Names the output data set to contain the forecasts of the variables specified in the subsequent FORECAST statements. If an ID variable is specified, it will also be included in the OUT= data set. This dataset is similar to the OUTFOR dataset but different in that the MIPSUSED variable is a combination of the ACTUAL and PREDICT variables from the OUTFOR dataset. You can compare the screen shot below to the previous screen shot and notice that the values from 04Sep2011 through 16Oct2011 are coming from ACTUAL and the values from 23Oct2011 on are coming from PREDICT.

	LPARNAME	SHIFT	DATE	MIPSUSED
31	LPAR01	1	Sun, 4 Sep 2011	95.31
32	LPAR01	1	Sun, 11 Sep 2011	95.90
33	LPAR01	1	Sun, 18 Sep 2011	102.59
34	LPAR01	1	Sun, 25 Sep 2011	113.44
35	LPAR01	1	Sun, 2 Oct 2011	97.52
36	LPAR01	1	Sun, 9 Oct 2011	99.99
37	LPAR01	1	Sun, 16 Oct 2011	101.15
38	LPAR01	1	Sun, 23 Oct 2011	100.80
39	LPAR01	1	Sun, 30 Oct 2011	100.71
40	LPAR01	1	Sun, 6 Nov 2011	100.63
41	LPAR01	1	Sun, 13 Nov 2011	100.55
42	LPAR01	1	Sun, 20 Nov 2011	100.46
43	LPAR01	1	Sun, 27 Nov 2011	100.38
44	LPAR01	1	Sun, 4 Dec 2011	100.30
45	LPAR01	1	Sun, 11 Dec 2011	100.21
46	LPAR01	1	Sun, 18 Dec 2011	100.13
47	LPAR01	1	Sun, 25 Dec 2011	100.04
48	LPAR01	1	Sun, 1 Jan 2012	99.96
49	LPAR01	1	Sun, 8 Jan 2012	99.88

by LPARNAME SHIFT;

We will create a forecast for each LPARNAME/SHIFT combination

id DATE interval=week;

- The ID statement options also specify how the observations are accumulated and how the time ID values are aligned to form the actual time series.
- The interval statement specifies the frequency of the input time series.

forecast MIPSUSED;

The FORECAST statement lists the numeric variables in the DATA= data set whose accumulated values represent time series to be modeled and forecast.

Forecast Code Part 3

```
data work.outfor_plot;
   set work.outfor;
   if DATE < &forecastdate then
      do;
         lower=.;
         upper=.;
      end;
```

We are cleaning up the OUTFOR dataset for our plot charts later. We are setting the lower and upper to null for dates less than the forecastdate which we know is 16OCT2011.

Forecast Code Part 4

```
%global lastforecastdate;
proc sql noprint;
   select max(DATE) into :lastforecastdate
   from WORK.outfor;
quit;
%put lastforecastdate = %sysfunc(putn(&lastforecastdate,date9.));
```

This section of the code is getting the last date from our OUTFOR table and setting this to our lastforecastdate variable to be used later in the process. In our case study the SAS log is showing “lastforecastdate = 08JAN2012”.

SAS Code Part 5

```
proc datasets nolist nowarn library=WORK;
   delete LPARMIPS_DATA_SORT;
   delete OUTFOR;
run;
```

Cleaning up the work datasets.

SAS Code Part 6

```
%global forecastdays;
%let forecastdays = %eval(&lastforecastdate - &forecastdate);
%put forecastdays=&forecastdays;
run;

%put _user_;
run;
```

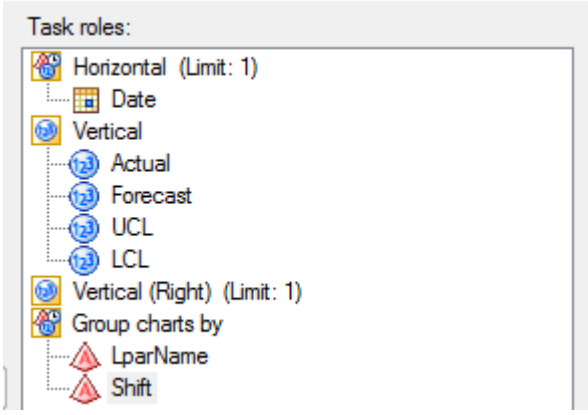
Here we are calculating the forecastdays for later use. In our case study the SAS log shows "forecastdays=84". This is seven days times 12 weeks forecast lead time.

Line Plots of Forecasts

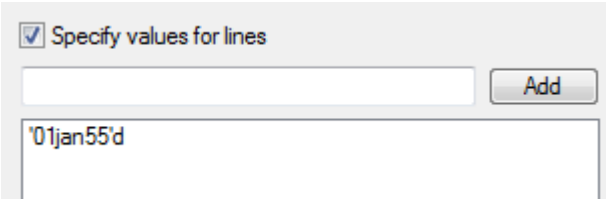
Here we will cover creating the Line Plots Forecasts for each LPAR and Shift.

Line Plot Icon

Here are the variables for the Line Plot task.

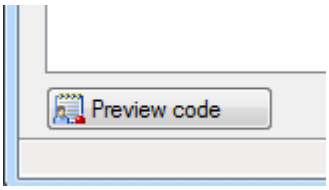


In the chart task we inserted a dummy HREF '01jan55'd

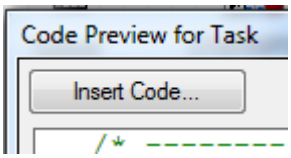


Then we insert code to override this with this statement:
HREF="%SYSFUNC(PUTN(%EVAL(&forecastdate.),date9.))"D

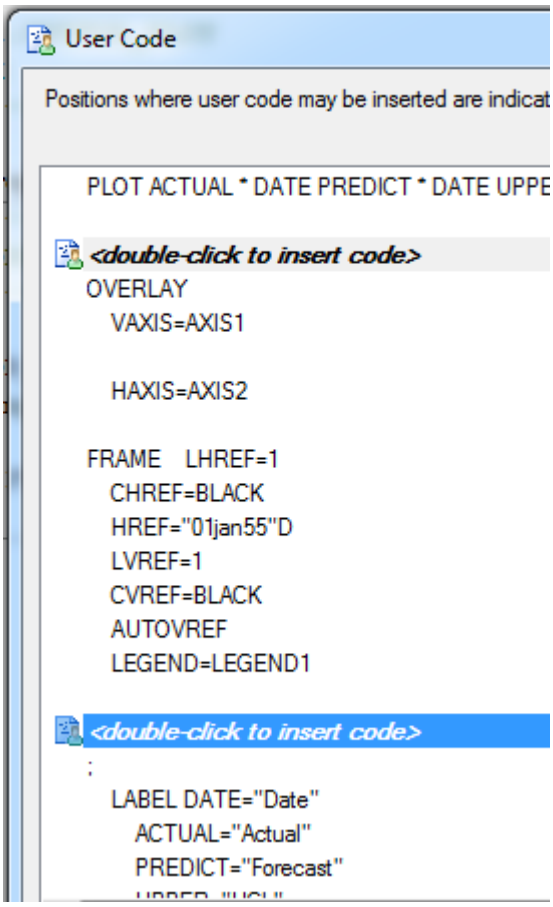
To do this click on the “Preview Code” button.



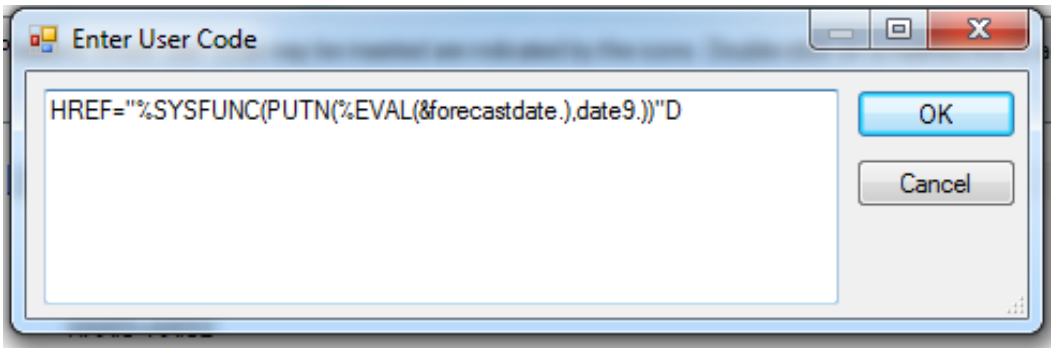
Then click the “Insert Code” button.



Double click the blue section below “<double-click to insert code>”

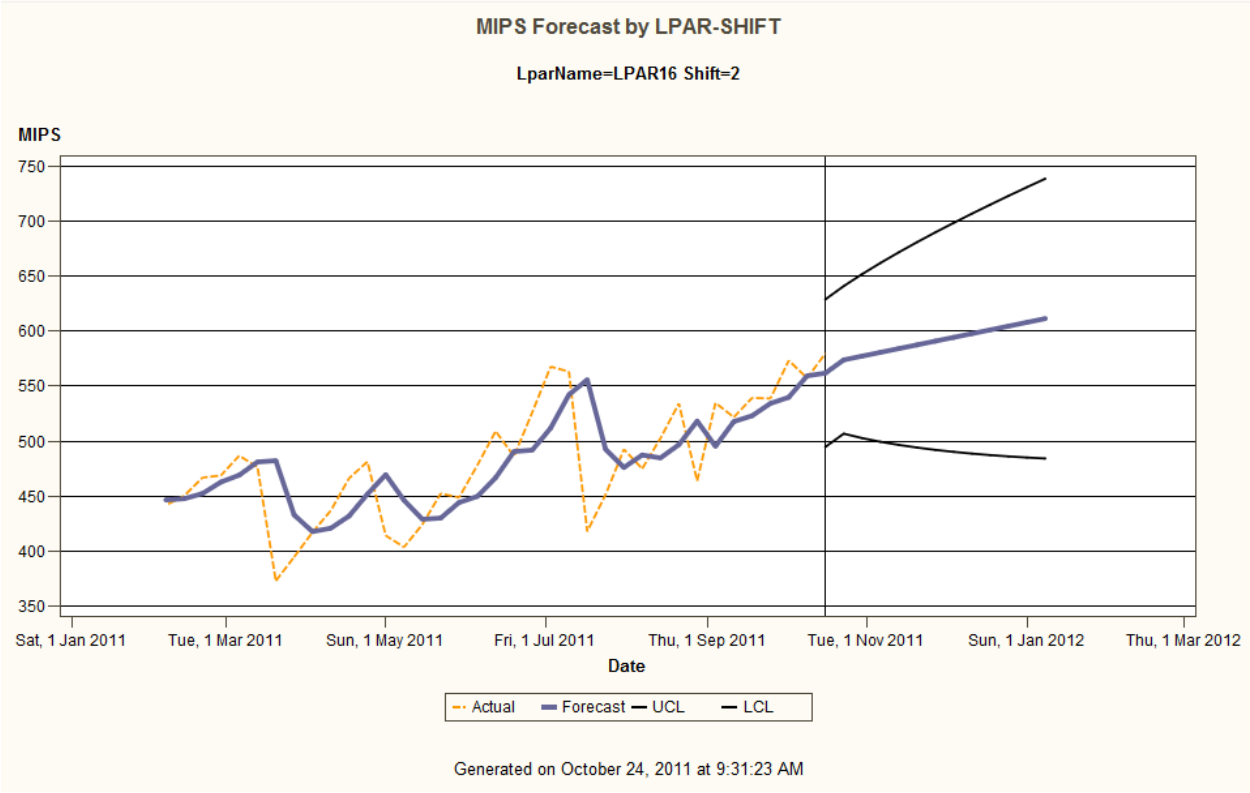


Paste the SAS code as follows:



Output from the Line Plot Icon

Below is one chart from the many charts in the output from the Line Plot task.



HTML-HPFENGINE Icon

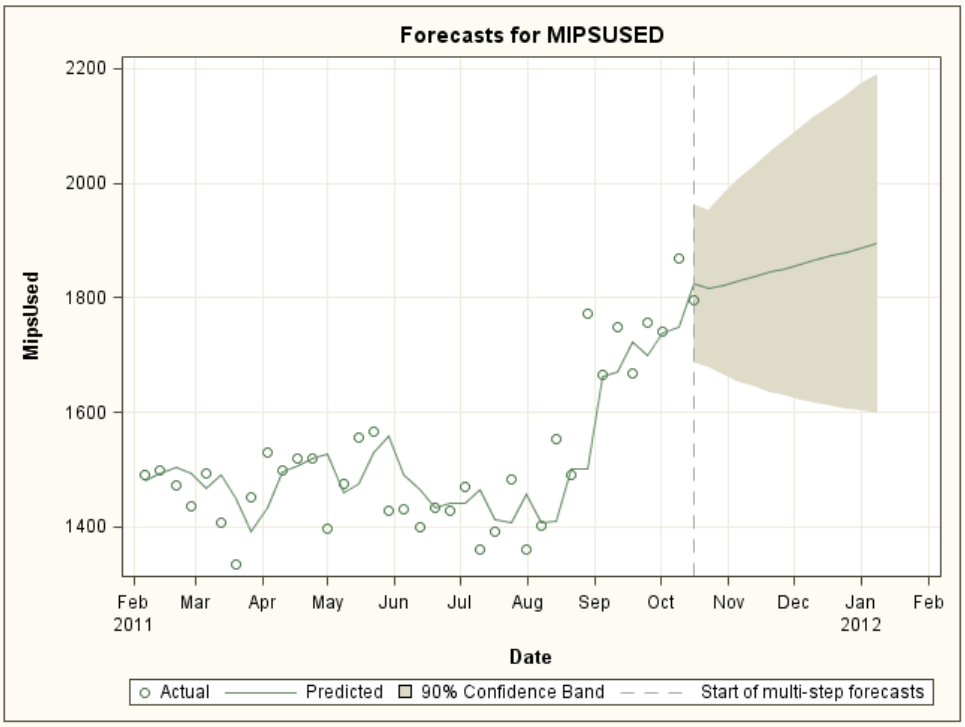
Opening this item reveals the output report from the HPFENGINE procedure. This report is used to show the programmer the various models used for each forecasted LparName/Shift combination.

A snippet of the model selection section is shown here.

Model Selection Criterion = MAPE			
Model	**Statistic**	**Selected**	**Label**
MEAN	5.8526917	No	ARIMA: Y ~ CONST
LINEARTREND	5.7056530	No	ARIMA: Y ~ CONST + INPUT: _LINEAR_
LINEARTRENDAR	.	Removed	ARIMA: Y ~ P = ((1,2,3)(1)s) + INPUT: _LINEAR_
LINEARSEASONALDUMMIES	.	Removed	ARIMA: AIR ~ NOINT + INPUT1: _LINEAR_ + INPUT2: _SEASONAL_
SEASONALDUMMIES	.	Removed	ARIMA: AIR ~ NOINT + INPUT: _SEASONAL_
SIMPLE	4.7161846	No	Simple Exponential Smoothing
LINEAR	4.6892891	Yes	Linear Exponential Smoothing
DAMPTREND	4.7161998	No	Damped-Trend Exponential Smoothing

You can see the procedure is removing some models and rejecting other models. The model selected for this data is the Linear Model.

This report also shows another version a forecast chart.



Conclusion

The steps presented in this paper will enable you to adapt this method to any data source needing forecast projections. The SAS High-Performance Forecasting software is a very powerful toolset that is included free with SAS ITRM 3.2.

Appendix 1 (HPFENGINE code)

```
options nofmterr;
ods graphics on;

%global forecastdate;
proc sql noprint;
  select max(DATE) into :forecastdate
  from WORK.LPARMIPS_DATA;
quit;
%put forecastdate = %sysfunc(putn(&forecastdate,date9.));

proc sort data=WORK.LPARMIPS_DATA out=WORK.LPARMIPS_DATA_SORT;
  by LPARNAME SHIFT DATE;
  run;

proc hpfengine data=WORK.LPARMIPS_DATA_SORT
     lead=12
        globalselection=tsfsselect
     task=select(criterion=MAPE alpha=0.10 override)
     print=(select summary)
     plot=forecasts
     outsum=work.outsum
     outfor=work.outfor
     out=work.out;
  by LPARNAME SHIFT;
  id DATE interval=week;
  forecast MIPSUSED;
run;

data work.outfor_plot;
  set work.outfor;
  if DATE < &forecastdate then
     do;
        lower=.;
        upper=.;
     end;

%global lastforecastdate;
proc sql noprint;
  select max(DATE) into :lastforecastdate
  from WORK.outfor;
quit;
%put lastforecastdate = %sysfunc(putn(&lastforecastdate,date9.));

proc datasets nolist nowarn library=WORK;
  delete LPARMIPS_DATA_SORT;
  delete OUTFOR;
run;

%global forecastdays;
%let forecastdays = %eval(&lastforecastdate - &forecastdate);
%put forecastdays=&forecastdays;
run;

%put _user_;
run;
```

DISCLAMIER

EXCEPT WHERE EXPRESSLY PROVIDED OTHERWISE IN AN AGREEMENT BETWEEN YOU AND SAS, ALL INFORMATION, SOFTWARE, PRODUCTS AND SERVICES ARE PROVIDED "AS IS" WITHOUT WARRANTY OF ANY KIND INCLUDING WARRANTIES OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE AND NON-INFRINGEMENT.

IN NO EVENT SHALL SAS BE LIABLE FOR ANY DIRECT, INDIRECT, INCIDENTAL, SPECIAL OR CONSEQUENTIAL DAMAGES, OR DAMAGES FOR LOSS OF PROFITS, REVENUE, DATA OR USE, INCURRED BY YOU OR ANY THIRD PARTY, WHETHER IN AN ACTION IN CONTRACT OR TORT, ARISING FROM YOUR ACCESS TO, OR USE OF, THIS WEB SITE, THE SERVICES PROVIDED ON THIS WEB SITE, OR ANY OTHER HYPERLINKED WEB SITE.

IF ANY PROVISION OF SAS' TERMS OF USE IS HELD TO BE UNLAWFUL, VOID, OR UNENFORCEABLE BY ANY COURT HAVING COMPETENT JURISDICTION, THEN SUCH PROVISION SHALL BE SEVERABLE WITHOUT AFFECTING THE ENFORCEABILITY OF ALL REMAINING PROVISIONS, WHICH SHALL REMAIN IN FULL FORCE AND EFFECT.