Paper SAS3061-2016

Dynamic Forecast Aggregation

Michael Leonard, Alex Chien, Pu Wang, and Yue Li, SAS Institute Inc.

ABSTRACT

Many organizations need to report forecasts of large numbers of time series at various levels of aggregation. Numerous model-based forecasts that are statistically generated at the lowest level of aggregation need to be combined to form an aggregate forecast that is not required to follow a fixed hierarchy. The forecasts need to be dynamically aggregated according to any subset of the time series, such as from a query. This paper proposes a technique for large-scale automatic forecast aggregation and uses SAS® Forecast Server and SAS/ETS® software to demonstrate this technique.

INTRODUCTION

Many traditional techniques are available for generating statistical forecasts for a particular time series, automatically generating and selecting time series models that can then be used to create forecasts, and automatically generating statistical forecasts for numerous time series that are arranged in a particular hierarchy. In addition to these traditional techniques, organizations that have numerous previously generated statistical forecasts want to dynamically view the aggregates of these forecasts to make better business decisions.

Aggregating forecasts can be difficult because forecasts are not simple numbers. The forecast for each time period is a distribution (which has prediction standard errors and confidence limits). So in order to aggregate forecasts, a means of preserving the basic distributional properties must be respected (or estimated).

This paper proposes a technique that dynamically aggregates numerous statistical time series forecasts and attempts to preserve their basic distributional properties. The proposed technique consists of the following steps, which are described in greater detail in the next section:

1. Generate the disaggregate forecasts.
2. Subset the disaggregate forecasts.
3. Aggregate the actual time series data and the predicted time series.
   Note: You cannot linearly aggregate the prediction standard errors or the confidence limits.
4. Forecast the actual time series.
5. Reconcile the aggregate forecasts.

This paper uses proven traditional time series analysis techniques along with standard forecast reconciliation techniques to provide dynamic forecast aggregation forecasts for the typical business user. Although this paper focuses only on extrapolation techniques, the concepts generalize to more complicated time series models.

DYNAMIC AGGREGATION TECHNIQUE

The steps in the proposed technique are illustrated in Figure 1 and described in greater detail in the following subsections.

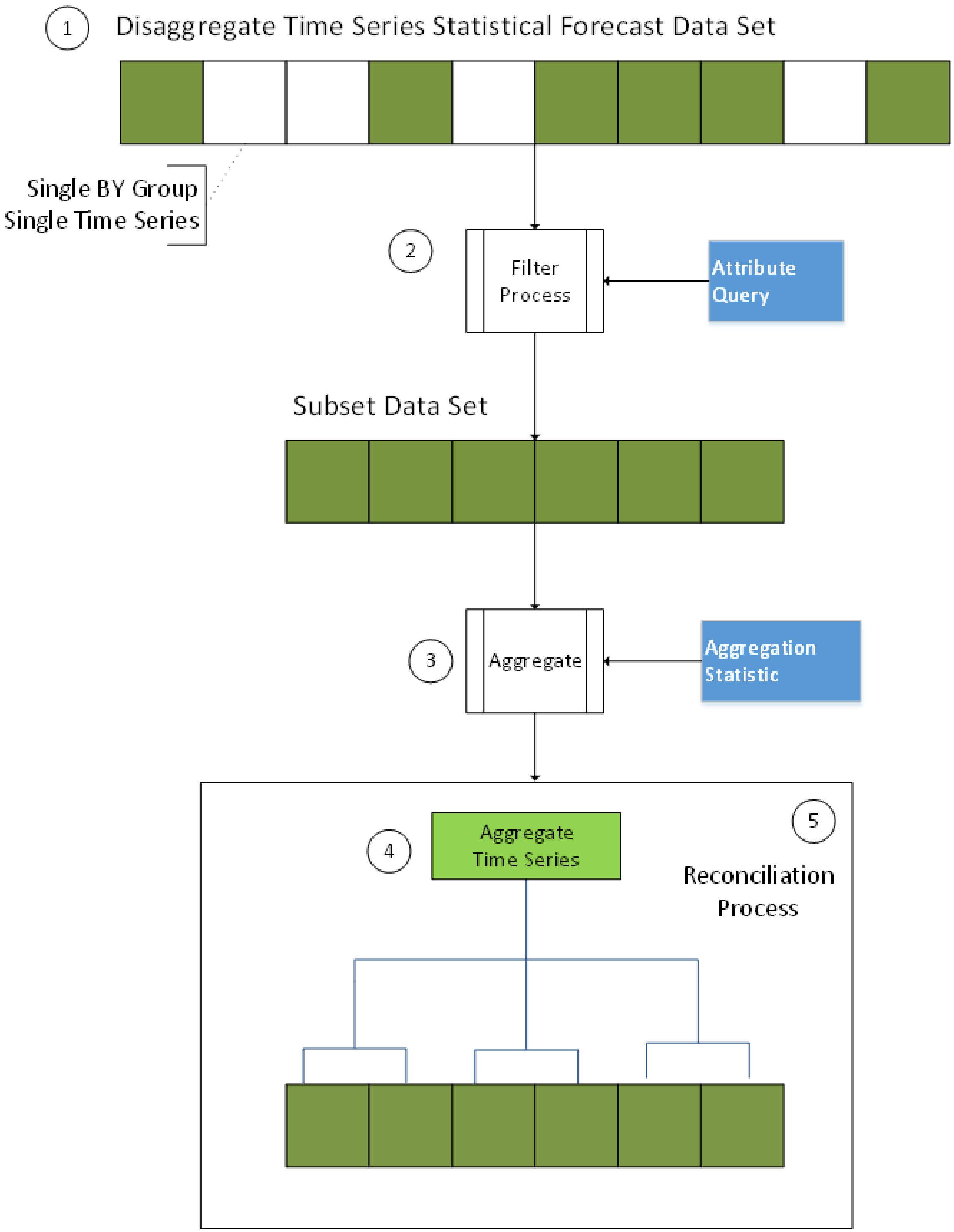


Figure 1. Dynamic Forecast Aggregation Conceptual Flow

STEP 1: GENERATE THE DISAGGREGATES

For the lowest level of aggregation (called disaggregation), you must provide statistical forecasts (called disaggregates) for each time series. There are many ways to generate these disaggregates.

The following time indices are used in time series analysis and forecasting:

- A (discrete) time index is represented by t = 1, …, T, where T represents the length of the historical time series.
- The index of future time periods (which are predicted by forecasting) is represented by l = 1, …, L, where L represents the forecast horizon (also called the lead).

The following series indices are used:

- A series index is represented by i = 1, …, N, where N represents the number of the historical time series.
- A subset of the time series is represented by $I \subseteq \{1, \ldots, N\}$, where N_I represents the number of series in the subset.

A time series value for series index i at time index t is represented by $y_{i,t}$. The *dependent series* (the historical time series vector) that you want to model and forecast is represented by $Y_{i,T} = \{y_{i,t}\}_{t=1}^{T}$.

The historical time series data can also contain *independent series*, such as inputs and calendar events that help model and forecast the dependent series. The historical and future predictor series vectors are represented by $\vec{X}_{i,T} = \{\vec{x}_{i,t}\}_{t=1}^{T+L}$.

Figure 1 shows an array of multicolored boxes that represent 10 disaggregate time series forecasts. Each box of the array represents a single time series.

STEP 2: SUBSET THE DISAGGREGATE FORECASTS

Subset the disaggregate forecasts by some means, such as manual selection, a query (filter) for various attributes of the individual times series, or some other means.

The colored boxes in Figure 1 represent the six disaggregates that you want to aggregate, and the white boxes represent the four aggregates you want to exclude. To subset this data set, you must apply a *filter*. The filter (query) represents attributes of each time series that help you select the six disaggregates.

STEP 3: AGGREGATE THE SUBSET

After the subset is created, the aggregate forecasts can be computed based on this subset. To aggregate the subset forecasts, you must specify an *aggregation statistic* (TOTAL, AVERAGE, or other statistic). Aggregating the actual time series (ACTUAL) and the predicted means (PREDICT) is usually fairly straightforward because the mean is a linear operator.

When you have numerous time series vectors, $Y_{i,T} = \{y_{i,t}\}_{t=1}^{T}$ for i = 1, …, N, you can aggregate the time series by means of an aggregation statistic, such as TOTAL or AVERAGE (and others).

$$\text{TOTAL} = y_{*,t} = \sum_{i=1}^{N} y_{i,t} \qquad \text{or} \qquad \text{AVERAGE} = y_{*,t} = \frac{1}{N}\sum_{i=1}^{N} y_{i,t}$$

The aggregation can occur over all the series indices, i = 1, …, N, or over a subset of the series indices, $I \subseteq \{1, \ldots, N\}$. Either way, the analysis that is described in the rest of these steps is the same.

Note: You cannot linearly aggregate the prediction standard errors (STD) or the confidence limits (LOWER and UPPER).

STEP 4: FORECAST THE AGGREGATE TIME SERIES

At this point, statistical forecasts are needed for the disaggregate time series, $Y_{i,T} = \{y_{i,t}\}_{t=1}^{T}$, and for the aggregate time series, $Y_{*,T} = \{y_{*,t}\}_{t=1}^{T}$. These forecasts can be provided in many ways. For ways of automatically producing statistical forecasts, see Leonard (2002), Leonard (2004), Leonard and Elsheimer (2015), and many others.

The disaggregate forecasts are represented by $\hat{Y}_{i,T+L} = \{\hat{y}_{i,t}\}_{t=1}^{T+L}$, where *L* is the forecast horizon. Likewise, the aggregate forecasts are represented by $\hat{Y}_{*,T+L} = \{\hat{y}_{*,t}\}_{t=1}^{T+L}$.

STEP 5: RECONCILE THE AGGREGATE FORECASTS

Aggregating the predicted standard errors (STD), lower confidence limit (LOWER), and upper confidence limit (UPPER) for each disaggregate time series might require information (reconciliation) from both the disaggregate forecasts and the forecasts that are directly generated from the aggregate time series (ACTUAL).

You can reconcile the aggregate forecasts and the numerous individual subset forecasts (subsets of the disaggregates) by using the usual hierarchal forecast reconciliation techniques (bottom-up). This reconciliation involves only two levels: a single aggregate and numerous disaggregates. Reconciliation enables you to better preserve the distributional properties, albeit not exactly.

At this point, you have forecasts for the numerous disaggregate time series, $\hat{Y}_{i,T+L} = \{\hat{y}_{i,t}\}_{t=1}^{T+L}$, and forecasts for the aggregate time series, $\hat{Y}_{*,T+L} = \{\hat{y}_{*,t}\}_{t=1}^{T+L}$.

However, the aggregate of the numerous disaggregate time series forecasts is not necessarily the same as the single aggregate time series forecasts:

$$\hat{y}_{*,t} \neq \sum_{i=1}^{N} \hat{y}_{i,t} \qquad \text{and} \qquad \hat{y}_{*,t} \neq \frac{1}{N}\sum_{i=1}^{N} \hat{y}_{i,t}$$

In order to reconcile these differences, reconciliation techniques are required. Consider a simple example, a simple two-series aggregation that uses TOTAL as an aggregation statistic:

$$y_{*,t} = y_{1,t} + y_{2,t}$$

Top-down (proportional) reconciliation results in the reconciled predictions that are shown in the first line of the following equations, and bottom-up (proportional) reconciliation results in the reconciled predictions that are shown in the second line of the following equations:

$$\hat{y}^{R}_{*_t} = \hat{y}_{*,t} \qquad \hat{y}^{R}_{1,t} = \hat{y}_{1,t}\left(\frac{\hat{y}_{*,t}}{\hat{y}_{1,t} + \hat{y}_{2,t}}\right) \qquad \hat{y}^{R}_{2,t} = \hat{y}_{2,t}\left(\frac{\hat{y}_{*,t}}{\hat{y}_{1,t} + \hat{y}_{2,t}}\right)$$

$$\hat{y}^{R}_{*_t} = \hat{y}_{1,t} + \hat{y}_{2,t} \qquad \hat{y}^{R}_{1,t} = \hat{y}_{1,t} \qquad \hat{y}^{R}_{2,t} = \hat{y}_{2,t}$$

Other forms of hierarchical forecast reconciliation are possible. For more information about hierarchical reconciliation see Trovero, Joshi, and Leonard (2007).

If the disaggregate forecasts are considered more reliable, bottom-up reconciliation is preferred. Bottom-up reconciliation is considered more reliable when hierarchical time series techniques are used to generate the forecasts for the disaggregate time series.

If the aggregate forecasts are more reliable, then there is no need to use reconciliation at all. Simply aggregate the numerous disaggregate time series and forecast the resulting time series. It is that simple.

Because bottom-up reconciliation is easier and more useful and because top-down reconciliation needs to be treated more carefully and is far more computationally intensive, this paper focuses on bottom-up reconciliation.

From the bottom-up equation, it seems that you just need to aggregate the forecasts:

$$\hat{y}^R_{*_t} = \hat{y}_{1,t} + \hat{y}_{2,t}$$

However, this equation represents the sum of two random variables—not the sum of two numbers. For example, you cannot always sum variances and you can never sum confidence limits.

The predictions (the mean of the forecasts) are the expected values. They can easily be aggregated:

$$E\left[\hat{y}^R_{*_t}\right] = E\left[\hat{y}_{1,t}\right] + E\left[\hat{y}_{2,t}\right]$$

Normally the reconciled prediction variances are calculated as follows:

$$Var\left[\hat{y}^R_{*_t}\right] = Var\left[\hat{y}_{1,t}\right] + Var\left[\hat{y}_{2,t}\right] + 2Cov\left[\hat{y}_{1,t}, \hat{y}_{2,t}\right]$$

Because the covariance is usually not available, the reconciled prediction variances are calculated in one of three possible ways:

- The reconciled prediction variances are the *same* as the aggregate series prediction variances:

$$Var\left[\hat{y}^R_{*_t}\right] = Var\left[\hat{y}_{*,t}\right]$$

- The reconciled prediction variances are *proportional* to the aggregate series prediction variances:

$$Var\left[\hat{y}^R_{*_t}\right] = \left[\frac{\hat{y}^R{}_{*,t}}{\hat{y}_{*,t}}\right]^2 Var\left[\hat{y}_{*,t}\right]$$

- The reconciled prediction variances are the *sum* of the numerous disaggregate series prediction variances:

$$Var\left[\hat{y}^R_{*_t}\right] = Var\left[\hat{y}_{1,t}\right] + Var\left[\hat{y}_{2,t}\right]$$

assumes that the two series are uncorrelated

The prediction standard errors are always the square root of the prediction variances, regardless of the method used to calculate them:

$$Std\left[\hat{y}^R_{*_t}\right] = \sqrt{Var\left[\hat{y}^R_{*_t}\right]}$$

The reconciled aggregate series confidence limits are calculated in one of two possible ways:

- Simply *shift* the confidence limits by using the difference between the reconciled aggregate series predictions and the aggregate series predictions:

$$Lower\left[\hat{y}^R_{*_t}\right] = Lower\left[\hat{y}_{*,t}\right] + \left(\hat{y}^R_{*_t} - \hat{y}_{*,t}\right)$$

$$Upper\left[\hat{y}^R_{*_t}\right] = Upper\left[\hat{y}_{*,t}\right] + \left(\hat{y}^R_{*_t} - \hat{y}_{*,t}\right)$$

- Compute the confidence limits by using the reconciled aggregate series prediction standard errors (assuming a Gaussian distribution with a confidence limit size of α):

$$Lower\left[\hat{y}^R_{*_t}\right] = \hat{y}^R{}_{*,t} + Z_{(\alpha/2)} Std\left[\hat{y}^R_{*_t}\right]$$

$$Upper\left[\hat{y}^R_{*_t}\right] = \hat{y}^R{}_{*,t} + Z_{(1-\alpha/2)} Std\left[\hat{y}^R_{*_t}\right]$$

The easiest way to aggregate the forecasts is to simply aggregate the predictions (PREDICT), copy the prediction standard errors (STD), and shift the confidence limits (LOWER and UPPER).

SAS IMPLEMENTATION

For this example, the Sashelp.PriceData data set contains monthly time series data for product sales by region. The variable Date records the time ID values; the variable Sale records the time series values for the number of units sold; and the variables RegionName, ProductLine, and ProductName are BY variables that categorize the 17 BY groups.

Figure 2 illustrates the 17 time series on a common scale. Figure 3 illustrates the 17 time series in stacked time series plot (no scale). Figure 4 illustrates the 3 time series on a common scale, where RegionName='Region1'. Figure 5 illustrates the 3 time series in stacked time series plot (no scale), where RegionName='Region1'.

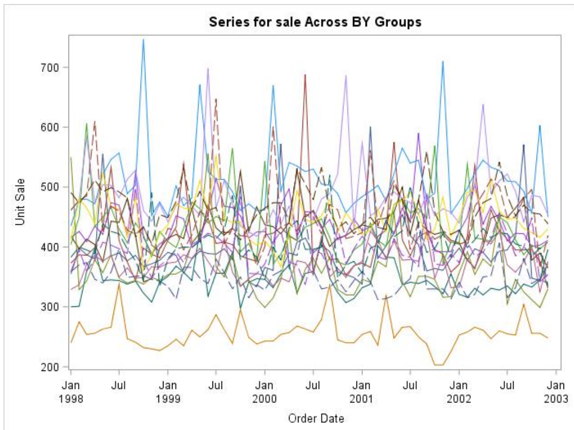


Figure 2. Time Series Plot

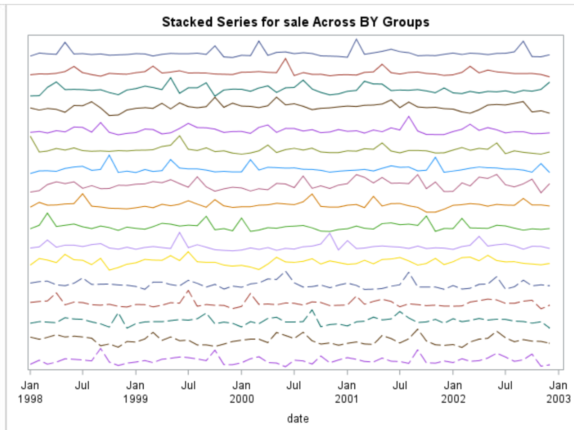


Figure 3. Stacked Time Series Plot

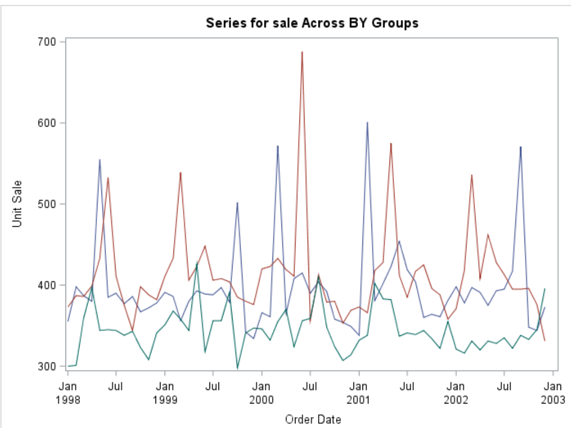


Figure 4. Time Series Plot (Region1 Only)

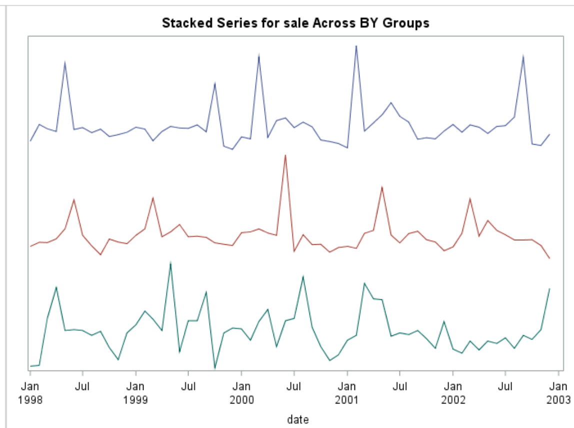


Figure 5. Stacked Time Series Plot (Region1 Only)

The following sections show how you can use SAS Forecast Server and SAS/ETS procedures to implement the steps of the dynamic aggregation technique described in this paper.

Step 1: Automatically forecast the disaggregates.

There are many ways to generate forecasts for the disaggregate time series. This paper uses the simplest for illustration.

The following SAS Forecast Server statements generate forecasts for the 17 time series in the Sashelp.PriceData data set:

```
proc hpf data=sashelp.pricedata out=_NULL_ outfor=disaggregates;
    by RegionName ProductLine ProductName;
     id Date interval=month;
     forecast Sale;
run;
```

The DATA= option specifies the time series data set to be analyzed. The OUTFOR= option specifies the output data set (Disaggregates) to contain the forecasts. The BY statement specifies the BY variables (RegionName, ProductLine, and ProductName). The ID statement specifies the time ID variable (Date), and the INTERVAL= option specifies the time interval (Month). The FORECAST statement specifies the time series variable to forecast (Sale).

At this point, you have forecasts for all the disaggregate time series. These forecasts can now be filtered and aggregated dynamically in a repeatable fashion for many filters.

Steps 2 through 5 can be accomplished by using the SAS/ETS TIMEDATA procedure, as follows:

```
proc timedata
    data=disaggregates(drop=STD Lower Upper where=(RegionName='Region1'))
    outarray=Forecasts(keep=Date Actual Predict STD Lower Upper) out=_NULL_;
    id Date interval=month accumulate=total notsorted;
    var Actual Predict;
    outarray AGGPREDICT STD LOWER UPPER;
    register tsm;

  declare object esm(tsm);
  rc = esm.initialize();
  rc = esm.sety(actual);
  rc = esm.run();
  rc = esm.getForecast('predict',aggpredict);
  rc = esm.getForecast('stderr',std);
  rc = esm.getForecast('lower',lower);
  rc = esm.getForecast('upper',upper);

  do t=1 to _LENGTH_;
     lower[t] = lower[t] + (predict[t] - aggpredict[t]);
     upper[t] = upper[t] + (predict[t] - aggpredict[t]);
  end;
run;
```

Step 2: Subset the disaggregates.

The DATA= option specifies the data set (Disaggregates) that contains the disaggregate time series and forecasts. The filter is specified by the WHERE clause, WHERE=(RegionName='Region1'). The ID statement specifies the time ID variable (Date), and the INTERVAL= option specifies the time interval (Month).The VAR statement specifies that the ACTUAL and PREDICT variables be used in the remaining code.

Step 3: Aggregate the disaggregate time series actuals and predictions.

The ID Statement ACCUMULATE=TOTAL specifies total aggregation.

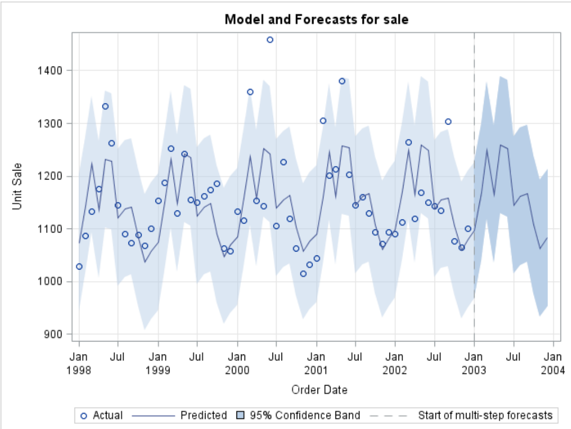
Step 4: Automatically forecast the aggregate time series.

The DECLARE statement creates an exponential smoothing model (ESM) object. The statements that follow forecast the aggregate time series.

Step 5: Reconcile the aggregate forecasts.

The DO loop performs the bottom-up reconciliation.

Figure 6 illustrates the time series forecasts of the aggregate time series *without* reconciliation. The blue circles represent the actual time series values. The blue line represents the predictions, and the blue shading represents the 95% confidence region. The vertical line separates the in-sample and out-sample regions. Figure 7 illustrates the time series forecasts of the aggregate time series *with* reconciliation.



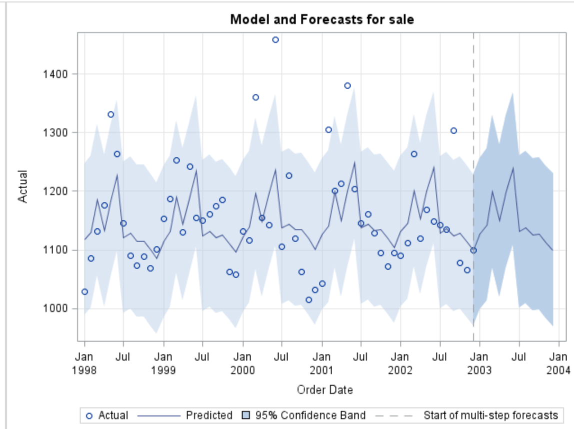


Figure 6. Aggregate Forecasts (No Reconciliation) Figure 7. Stacked Time Series Plot (Reconciliation)

As expected, the predicted values (means of the forecasts) differ when reconciliation is used. However, the reconciled predictions properly aggregate the disaggregate forecasts whereas the nonreconciled predicted values do not. Also note that the confidence limits are reasonable for the reconciled forecasts.

CONCLUSION

This paper proposes a technique for dynamically aggregating numerous previously statistically generated forecasts for use by decision makers. Because forecasts are not simple numbers but random variables, this technique attempts to provide ways to respect the distributional properties of the aggregate forecasts. This paper uses SAS/ETS and SAS Forecast Server software to demonstrate this technique.

REFERENCES

Leonard, M. J. 2002. “Large-Scale Automatic Forecasting: Millions of Forecasts.” International Symposium of Forecasting. Dublin.

Leonard, M. J. 2004. “Large-Scale Automatic Forecasting with Calendar Events and Inputs.” International Symposium of Forecasting. Sydney.

Leonard, M. J., and Elsheimer, B. M. 2015. “Count Series Forecasting.” *Proceedings of the SAS Global Forum 2015 Conference*. Cary, NC. SAS Institute Inc. Available http://support.sas.com/resources/papers/proceedings15/SAS1754-2015.pdf.

Trovero, M. A., Joshi, M. V., and Leonard, M. J. 2007. “Efficient Reconciliation of a Hierarchy of Forecasts in Presence of Constraints.” *Proceedings of the SAS Global Forum 2007 Conference*. Cary, NC. SAS Institute Inc. Available http://www2.sas.com/proceedings/forum2007/277-2007.pdf.

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the authors at:

Michael Leonard
SAS Campus Drive
Cary, NC 27513
919-531-6967
Michael.Leonard@sas.com

Alex Chien
SAS Campus Drive
Cary, NC 27513
919-531-9236
Alex.Chien@sas.com

Pu Wang
SAS Campus Drive
Cary, NC 27513
980-938-1811
Pu.Wang@sas.com

Yue Li
SAS Campus Drive
Cary, NC 27513
949-517-9342
Yue.Li@sas.com