› White Paper

Looking Inside SAS® Forecast Studio

sas
THE POWER TO KNOW®

Contents

Abstract

Organizations depend on statistical forecasting to provide a solid foundation for many important planning and decision-making processes. SAS® Forecast Server, which comprises a graphical user interface (SAS Forecast Studio) and the statistical forecasting engine (SAS Forecast Server Procedures), facilitates and speeds the forecasting process by providing a convenient, user-friendly interface to the large-scale automatic forecasting, model building and time series exploration capabilities available in SAS.

SAS Forecast Server also includes the SAS Time Series Studio graphical user interface, providing extensive time series exploration capabilities. The new SAS Forecast Server Client (available in release 14.1) makes certain functions of SAS Forecast Studio and SAS Time Series Studio available from anywhere through a web browser. Find more information about SAS Forecast Server Client in the white paper "Looking Inside SAS® Forecast Server Client" (forthcoming).

Through SAS Forecast Studio, SAS Forecast Server addresses the needs of novice forecasters by being largely automated, yet still meets the needs of more experienced analysts by providing layers of sophistication that can be accessed as needed.

SAS Forecast Studio enables users to set up forecasting projects, perform large-scale automatic forecasting, identify exceptions, override forecasts and construct their own models if desired. Given the scale of many forecasting problems, manually customizing many statistical models may not be feasible. SAS Forecast Studio provides an automated system that selects appropriate models and intelligently chooses influential variables that improve the model. Forecasts that violate business rules can be flagged for further attention. The system supports hierarchical forecasting processes by providing top-down, middle-out and bottom-up forecast reconciliation and generates SAS code that can be run in a batch environment.

Introduction

Generating forecasts is a crucial step in many strategic and tactical planning and decision-making processes. For example, future demand for products and services is forecast to support production planning, marketing activities, resource scheduling and financial planning. Much time and effort is spent in the planning process, and thus improving the reliability of statistical forecasts that feed these processes can result in huge rewards such as greater operational efficiency, reduced expenses and increased profits.

Although forecasting is a key business function, many organizations rely on a small forecasting team to generate large numbers of statistical forecasts. Therefore, a large degree of automation is often required to complete the forecasting process in the time available during each forecasting and planning cycle. Forecasting process tasks typically include generating forecasts based on recent data, reconciling forecasts in a hierarchical manner, identifying and correcting problematic forecasts, adding overrides to the forecasts based on business knowledge or scenario analysis and publishing the forecasts to other systems or as reports.

SAS Forecast Server, which consists of a graphical user interface (SAS Forecast Studio) as well as the statistical forecasting engine SAS Forecast Server Procedures, facilitates and speeds the forecasting process by providing a convenient, user-friendly interface to the large-scale automatic forecasting, model building and time series exploration capabilities available in SAS. In addition, the SAS Time Series Studio graphical user interface provides extensive time series exploration capabilities. And SAS Forecast Server Client makes certain functions of SAS Forecast Studio and SAS Time Series Studio available from anywhere through a web browser.

The wizard-driven graphical user interface of SAS Forecast Studio allows the novice forecaster to move quickly through the statistical forecasting process to generate forecasts using state-of-the-art forecasting methods. Forecasters do not need to know any SAS programming or how to build time series models to use SAS Forecast Server – the system can be entirely automated.

Although the SAS Forecast Studio user interface is streamlined to enable less-experienced forecasters to easily generate forecasts, SAS Forecast Studio has all the power and sophistication that more advanced analysts expect. SAS Forecast Studio provides a wide array of diagnostic and model-building tools, enabling more experienced analysts to benefit from the productivity of the analyst interface and to utilize even more of the SAS Forecast Server Procedures computational engine.

Overview of SAS® Forecast Studio

As the graphical user interface for SAS Forecast Server, SAS Forecast Studio has been designed to speed the work of both novice and advanced analysts:

- A New Project Wizard walks the analyst through the steps of setting up and running forecast projects.
- An Events dialog box simplifies the addition of events – for example, holidays and promotions – that might influence the forecast.
- A Model Builder dialog box guides the analyst through the development of custom forecasting models, built from scratch or based on models in the model repository.
- Forecast, Series and Model views give the analyst multiple ways to work with the forecast results.
- A Scenario Analysis view allows the analyst to create what-if scenarios based on the models generated by SAS Forecast Studio.
- Default and custom reports are readily generated in SAS Forecast Studio.
- Forecasts created with SAS Forecast Studio can be exported to a variety of other applications.

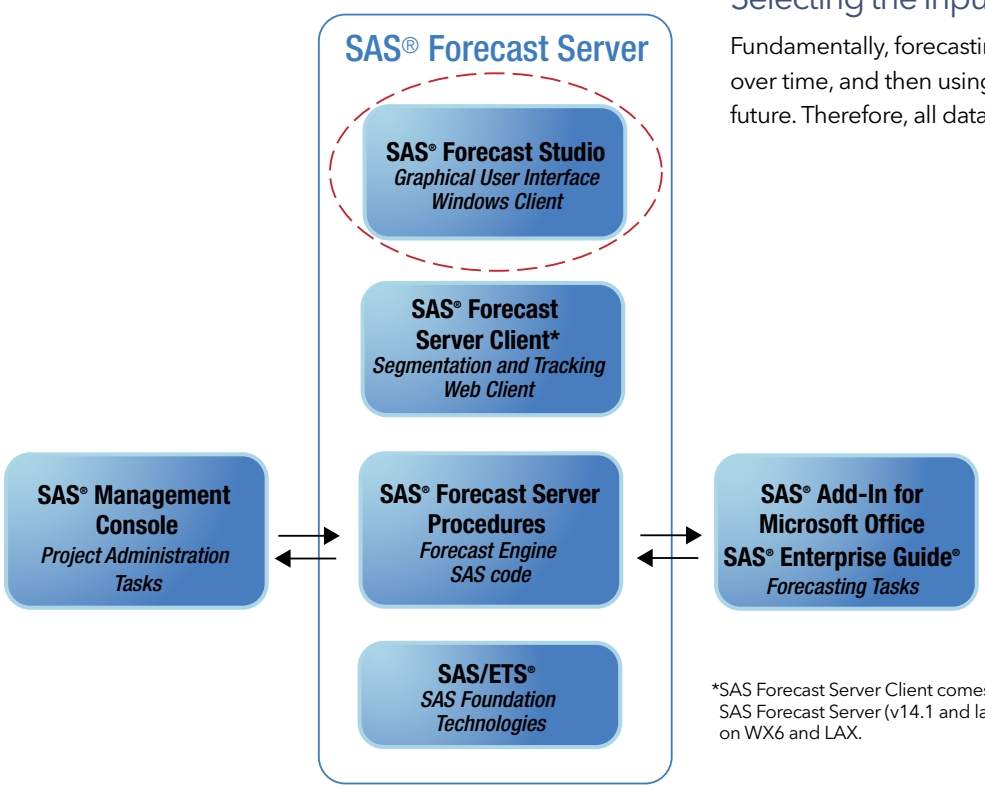


Figure 1: An overview of the components of SAS® Forecast Server.

Creating Forecast Projects

SAS Forecast Server organizes forecasts as projects. A forecasting project may be small, or it may contain thousands of time series. Organization of forecasts and all the information supporting them in projects has many advantages:

- Each member of a forecasting team may have his or her own projects, allowing him or her to work independently.
- Forecasters may maintain separate projects for the different accounts or product lines for which they are responsible.
- Forecasters may maintain separate projects for each forecasting cycle – the Q1 forecast, the Q2 forecast and so on.

The SAS Forecast Studio user interface includes an eight-step Project Wizard to guide analysts quickly through setting up a new project or working with existing projects.

Project folders – consisting mostly of SAS programs, SAS data sets and XML files – contain information about the input time series, event data, analyst-specified preferences, statistical models, parameter estimates, scenarios, SAS forecasting code and ultimately, the forecasted values. The input data is not stored within the project itself and does not need to be moved or copied to where projects are stored.

Selecting the Input Data Set

Fundamentally, forecasting is the process of detecting patterns over time, and then using those patterns to project into the future. Therefore, all data sets must contain a time variable, such

as month or day. For example, the data set may contain data on units sold per week, or electricity used per hour. If the data is recorded at regular intervals such as month or day, then it is time series data and is ready to be forecast. On the other hand, if data is collected at irregular intervals, such as raw transaction data from a retail store or call center, then it must be aggregated by time period to create a proper time series. If needed, SAS Forecast Studio performs this aggregation automatically.

The Project Wizard begins by prompting the analyst to navigate through SAS libraries and select the input data set. There is no predefined limit on the size of the data set or the number of series to be forecast in a project. In practice, the limits on project size will depend on the complexity of the data to be analyzed and the amount of history available for each series to be forecast.

Defining the Forecast Hierarchy

Many forecasts are hierarchical; for example, the forecast for units sold might include the number of individual products (SKUs) at the bottom level, and then aggregated upward by product line, region and finally, an overall forecast. In this case, the forecast hierarchy has four levels: product, product line, region and overall. "Product" is the fundamental unit to be forecast; "Product Line" and "Region" are SAS BY variables.

SAS Forecast Studio produces separate statistical forecasts at each level of a user-specified hierarchy, based on the history at each level. To generate the forecasts at upper levels of the hierarchy, the historical values from the lowest level are summed up to create the required input data time series.

Since each level of a hierarchy is forecast independently, it is highly unlikely that the statistical forecast at the second level of the hierarchy will equal the sum of the statistical forecasts at the first level, and so on. In this case, the forecasts must be reconciled within the hierarchy, so that the forecasts across the hierarchy levels are consistent (i.e., lowest level forecasts add up to the next level, and so on). The Project Wizard guides the user through the selection of an appropriate reconciliation technique: top down, bottom up or middle out.

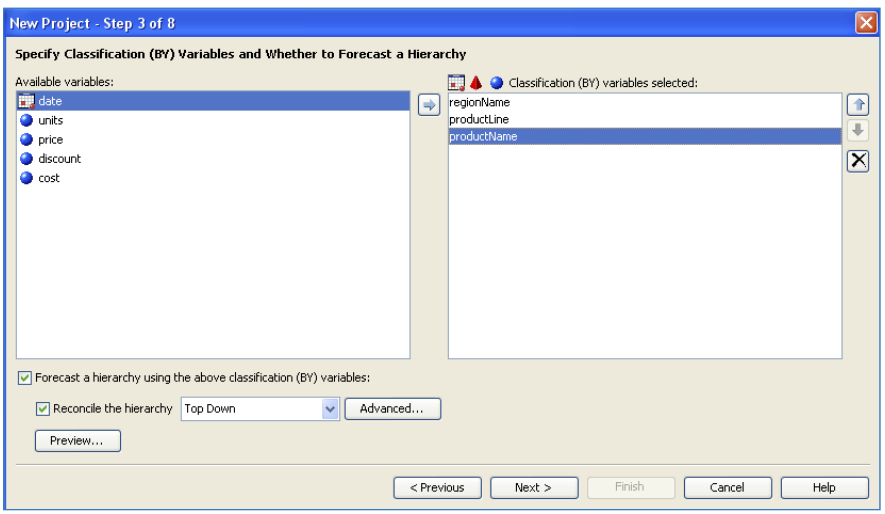


Figure 2: Defining a forecasting hierarchy.

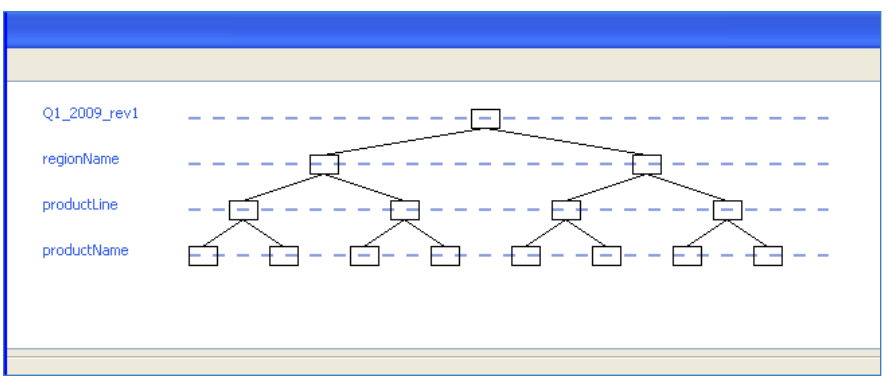


Figure 3: Example of a forecasting hierarchy.

Assigning Variable Roles

Variable roles determine how input variables are used within a project. The required time variable is assigned a role of "Time ID." As noted above, other input variables may be assigned the BY variable role and will determine the forecast's hierarchical structure.

The Project Wizard steps the analyst through the assignment of variable roles. The variable to be forecast – "units" in the preceding example – is assigned the role of dependent variable. In hierarchical forecasts, there can be only one dependent variable. Other possible variable roles include:

- Independent variables – used as potential explanatory inputs in the forecast models.
- Reporting variables – used in presenting the results of the forecast.
- Adjustment variables – used to transform variables for forecasting (for example, conversion to a common currency).

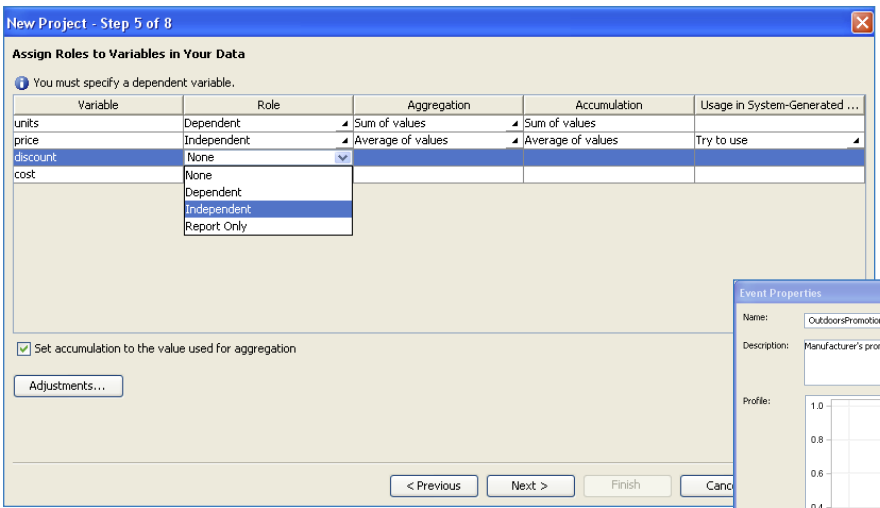


Figure 4: Defining variable roles.

Defining Events

Some models may be improved by the inclusion of events, which are occurrences out of the ordinary that may disturb the underlying time series. Some events are unplanned (such as strikes or storms); their impact should be isolated so that it is not propagated in future forecasts. Other events are planned (holidays, promotions, price changes) and may be recurring; their impact should be assessed and included in future forecasts if appropriate.

Events can be added to regression, ARIMAX and unobserved components models (UCM) as event variables: variables that indicate when something out of the ordinary occurred in the past or will occur again in the future. Once specified, SAS Forecast Server statistically estimates the impact of the event in the past and uses the estimated impact to calculate future forecasts where the event recurs.

The Project Wizard's Event Manager guides the user through selecting, creating, editing, combining or deleting events. SAS Forecast Studio includes a large collection of predefined events (mostly public holidays). In addition, externally generated tables of events can be imported easily into SAS Forecast Studio. To define a custom event within SAS Forecast Studio, the user specifies when the event occurred or will occur, the frequency of recurrence and the appropriate shape. For example, an event may be very short-lived (such as a power outage or the Olympics) and modeled as a pulse of limited duration, or an event may represent an ongoing change (such as a new law or regulation, or permanent price change) and thus modeled as a permanent level shift. Once defined, events may be saved externally for use by other forecasters or in other forecast projects.

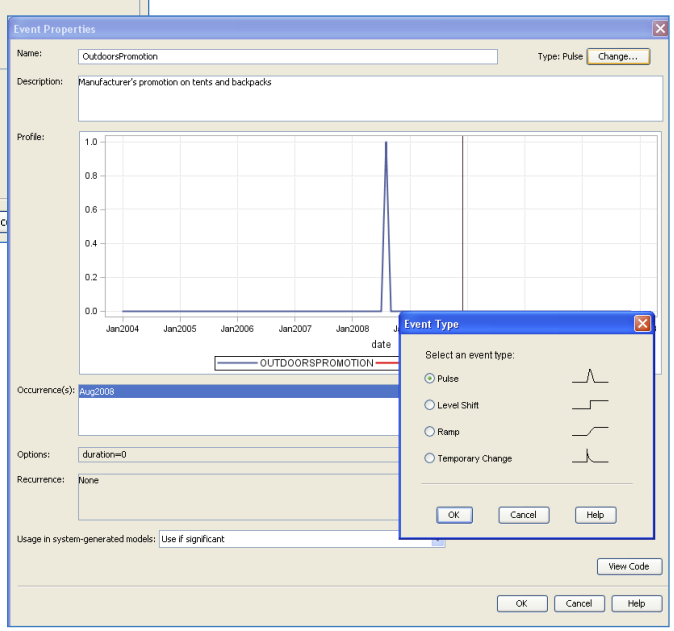


Figure 5: Defining events.

Setting Additional Forecast Options

When the basic components of a forecasting project are in place – the hierarchy defined, variable roles assigned and events created – the Project Wizard guides the analyst through the final, optional process of setting additional forecast parameters, including:

- How missing values in the data are to be interpreted.
- Which fit statistic should be used for model selection.
- How sensitive automatic outlier detection and other diagnostic tests should be.
- How leading and trailing zeros in the time series are to be treated.
- Whether or not forecasts are allowed to go below zero.
- Whether or not to use a holdout sample, and if so, its length definition.

SAS Forecast Studio offers standard default settings for all of these options, simplifying the forecast preparation task for novice analysts and streamlining the process for advanced forecasters.

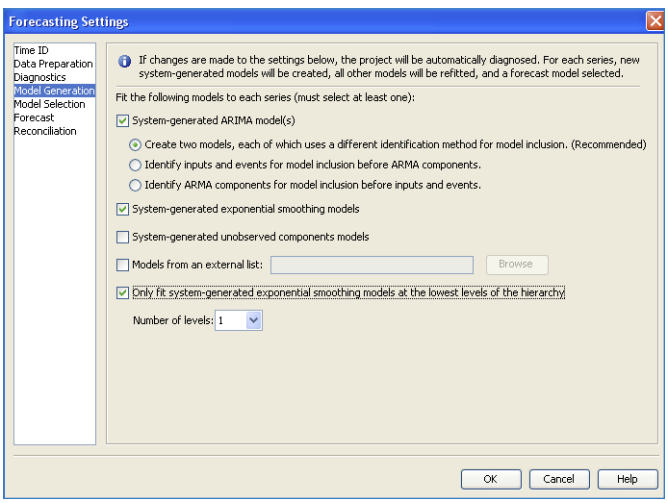


Figure 6: Project settings can be easily customized.

In addition, many businesses develop custom forecasting models that are used by many analysts throughout the company. The Project Wizard allows analysts to specify custom models to import into the project for use in the forecasting process, thus providing a means of further customizing the forecasting environment to fit the needs of both the forecast analyst and the business.

Forecast Generation

Upon closing the Project Wizard, SAS Forecast Studio automatically generates statistical forecasts for all series in the project. If a hierarchy has been specified, the system will reconcile the statistical forecasts according to the options specified by the user in the Project Set-Up wizard.

During processing, SAS Forecast Studio will:

- Diagnose the characteristics of each series to be forecast, at each level of the hierarchy.
- Estimate parameters for appropriate candidate models.
- Evaluate the performance of the candidate models.
- Select the best-performing model for each series, as well as alternative runner-up models.
- Produce forecasts.
- Reconcile the forecasts in the hierarchy.

Time Series Diagnosis and Model Construction

When generating forecasts, SAS Forecast Server must first determine an appropriate set of models for each time series being forecast. Several diagnostic procedures are automatically run for each time series in the forecast project. For example, in addition to the standard time series decomposition analysis, SAS Forecast Server tests for intermittency of the dependent variable, as well as the presence of outliers and structural shifts. If a time series is determined to be intermittent – that is, very few nonzero observations in the data – SAS Forecast Server will attempt to fit a specialized intermittent demand model. Otherwise, SAS Forecast Server constructs candidate exponential smoothing, ARIMAX and UCM models.

In addition to testing for intermittent dependent variables, SAS Forecast Server also examines each time series for outliers and shifts, and takes outliers and level shifts into consideration when constructing the forecast models.

SAS Forecast Studio performs many of the same diagnostic tests on independent and event variables as are performed on dependent variables. An important feature of SAS Forecast Studio is the ability to determine whether any of the independent or events variables are dynamically related to the dependent variable, and to specify the model accordingly. For example, when forecasting the demand for a product that exhibits a relationship with both product price and special promotional campaigns, SAS Forecast Server would detect and specify these separate dynamic relationships in an ARIMAX or UCM forecasting model.

Model Selection

When the properties of the time series have been evaluated, SAS Forecast Studio next evaluates a wide range of models, including unobserved components models (UCM), exponential smoothing models (ESM) and autoregressive integrated moving average models (ARIMA and ARIMAX), to determine which model best fits the time series data.

SAS Forecast Studio features more than four dozen model-fit statistics, including scale dependent measures such as MSE (Mean Square Error), measures based on percentage error such as MAPE (Mean Absolute Percentage Error), conventional statistical measures such as AIC (Akaike's Information Criterion) and many others.

Holdout samples can be specified so that forecasting models are selected not only by how well they fit the past data, but by how well they are likely to predict the future. If a holdout sample is specified – the preferred practice when enough data is available – the model will be selected based on the goodness-of-fit statistic in the holdout region (out-of-sample fit); otherwise in-sample fit is used to select the model. If a holdout sample was specified, the model parameters are estimated using the full range of data, including the holdout sample.

When selecting the best-fitting model, SAS Forecast Studio automatically tests candidate independent variables and events – identified during the forecast setup process – and determines how they should be used in the forecast models. In addition to examining the contemporaneous relationships between independent and dependent variables, lagged and dynamic relationships are explored. If appropriate, variable transformations, lags and transfer function definitions are calculated.

Forecast Summary

When the forecast process is complete, the Forecast Summary window provides an overview of important properties of the forecasting project by hierarchy level. The number of series forecast at each level of the hierarchy is given, as is the number of forecast failures (series that could not be statistically forecast, usually because the series contained too few observations). A bar chart displaying the distribution of fit statistics is shown, as are bar charts giving the distribution of model family and model characteristics (seasonal model, input variables present and outliers detected.)

Forecasting View

The detailed output from a SAS Forecast Studio project is accessed through a set of tabbed views. The first of these, the Forecasting view, displays the results of each time series forecast in a set of panels. The forecast graph panel shows the statistical forecast with 95 percent prediction intervals for each forecast; historical values are plotted as are the fitted values for the historical time period. The data table shows the historical data, the output of the statistical forecast and the fit statistic. If the model is part of a reconciled hierarchy, the reconciled (adjusted) forecast and fit statistic are displayed. Manual forecast overrides are also shown in the data table. The model details – type of model as well as parameter estimates – are revealed in a pop-up window.

Navigating the Forecast Hierarchy

Output is generated for every time series in the forecast project, at every hierarchy level. To facilitate navigating through the forecast hierarchy, SAS Forecast Studio provides two navigation modes: tree and table. In the Tree view, an expandable view of the forecast hierarchy is show in the navigation panel; the highlighted series is "active" and the corresponding output is shown in the graph and data table. Clicking up and down the hierarchy tree changes the active series and the displayed output.

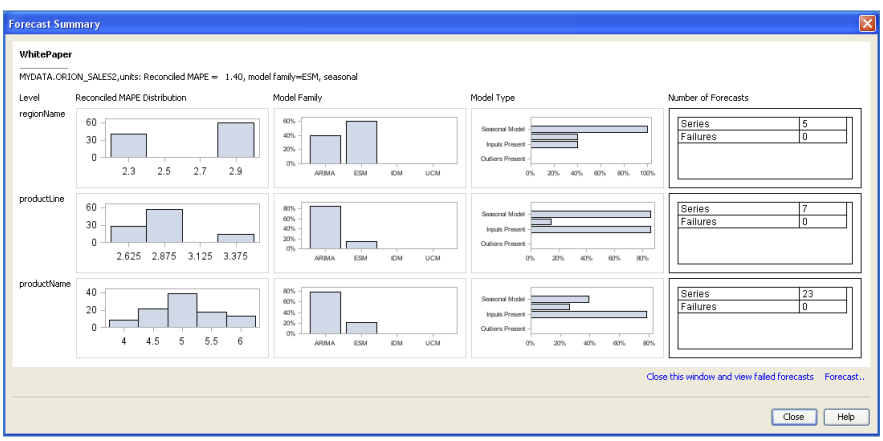


Figure 7: Forecast model statistics.

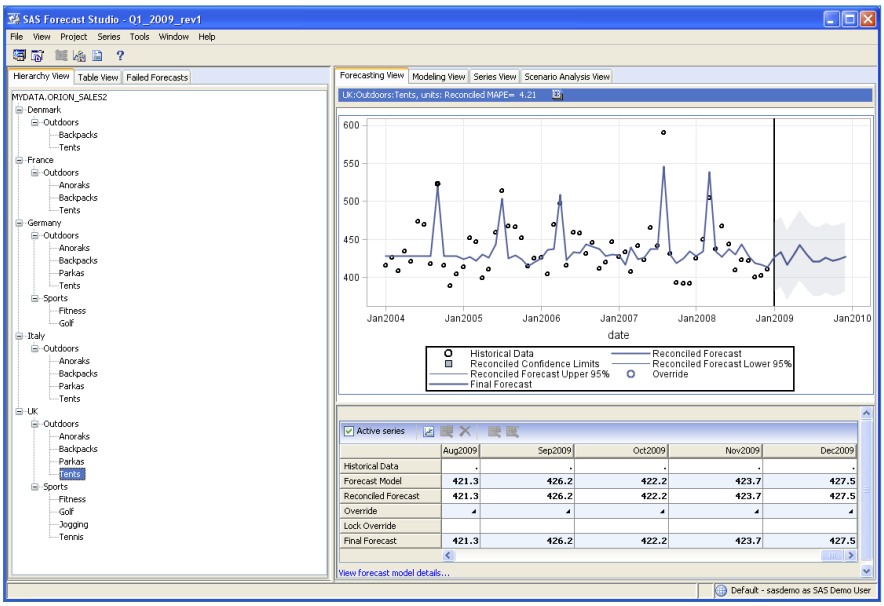


Figure 8: Forecasting View: Tree form.

In the Table Navigation view, the project series are listed in tabular form; columns for each BY variable are shown, as are the statistical and reconciled fit statistics. The table can be sorted by clicking on the column headers. In addition, a drop-down filter selector can be used to subset the displayed forecasts; filters are automatically generated for each BY variable in the forecast hierarchy.

Creating Forecast Filters

Filters are automatically created for each hierarchy level and a Filter dialog box guides the user through creating additional custom filters. For example, a filter can be created that isolates all the forecasts with a fit statistic value above a specified level, such as MAPE values greater than 5 percent. Filters are the basis of exception-based forecasting. Using filters, the forecast analyst can quickly identify the forecasts that are performing poorly and focus additional modeling attention on those time series. Other types of filters include filters on model property (for example, model family or inclusion of independent variables), series properties (for example, number of missing values, maximum value or date of first and last observation) and if applicable, reconciled fit statistic.

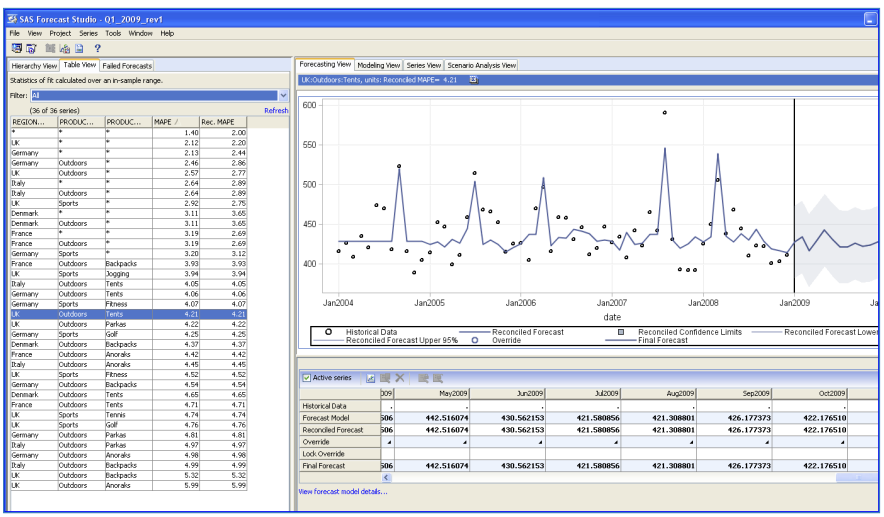


Figure 9: Forecasting View: Table form.

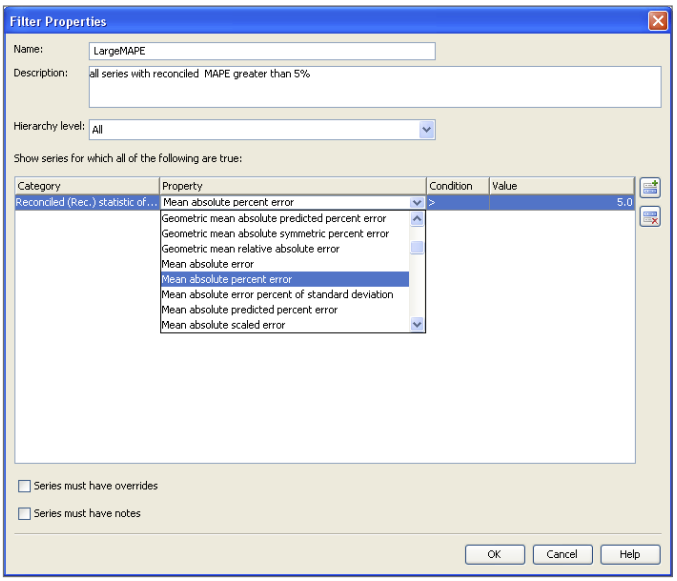


Figure 10: Forecast filtering.

Entering Forecast Overrides

Statistical forecast values can be manually overridden by entering new values in the data table; the numbers can be typed in, or the override calculator can be used to calculate new values by adjusting the statistical value by a fixed amount or a percentage. The override values are shown in the forecast graph, and, after forecast reconciliation, the impact of the overrides is shown at each level of the forecast hierarchy.

Series and Modeling Views

By using forecast filters to identify forecasts warranting further investigation because, perhaps, they exceed a forecast fit criterion threshold, the analyst can focus additional diagnostic and modeling efforts where the value to the forecasting process is greatest. Two SAS Forecast Studio views – the Series view and the Modeling view – help the analyst delve deeper into the forecasts generated by SAS Forecast Server.

Evaluating Series Properties

Opening the Series view in SAS Forecast Studio reveals a number of plots displaying properties of the active time series, including:

- Seasonal decomposition plots.
- Component plots.
- Autocorrelation and white noise probability plots.
- Cross-series plots.

In addition, in the Series view the analyst can explore the properties of both dependent and independent variables by experimenting with variable transforms using drop-down selection boxes. As in the Forecasting view, navigation among the time series is done with the Tree view or Table view in the navigation panel.

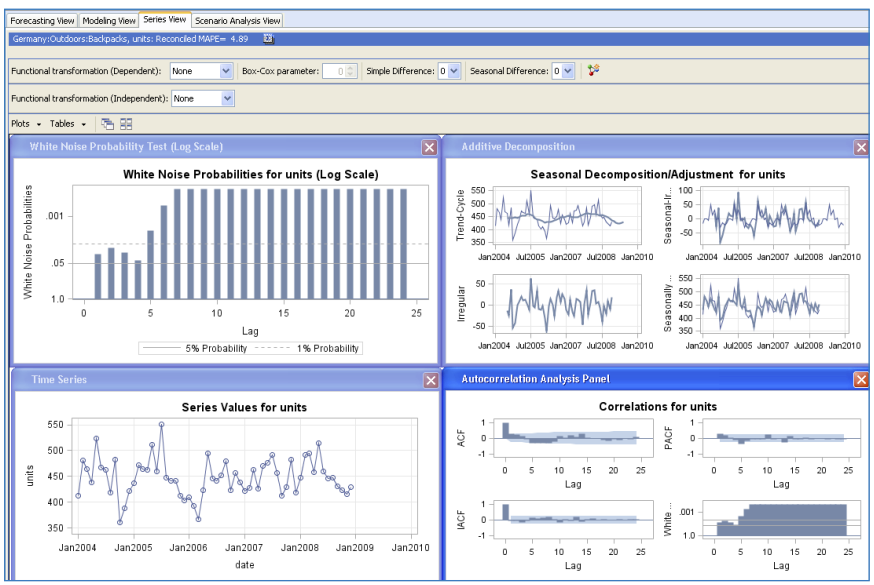


Figure 11: Series View.

Analyzing and Comparing Models

The heart of SAS Forecast Studio's advanced modeling capabilities lies in the Modeling view.

For the active forecast series, a list of the three most appropriate candidate models, based on model diagnostics and the chosen fit statistic, is shown. Clicking on a model activates it and displays details of the model as well as several diagnostic plots, including plots of:

- Prediction errors.
- Prediction error distribution.
- Autocorrelations of the prediction errors (and standardized autocorrelations).
- Partial autocorrelations of the prediction errors (and standardized partial autocorrelations).
- Inverse autocorrelations of the prediction errors (and standardized inverse autocorrelations).
- White noise probabilities for the prediction errors (also on log scale).
- Components of the series (smoothed trend, season and level states).

In addition, tables of model parameter estimates and forecast values are readily accessible. The Modeling view also provides tools for comparing multiple models, including tables showing the values of a large number of fit statistics as well as model comparison graphs.

Editing and Creating Models

Within the Modeling view, analysts can copy and edit models that were automatically generated by SAS Forecast Studio, or develop new models from scratch using the Modeling dialog box. From within the Modeling dialog box, users can select a model family and edit model parameters, add independent variables and events to a model and so on. By using SAS Forecast Studio's powerful Modeling dialog box, users create virtually any form of ARIMA, subset and factored ARIMA, unobserved components, exponential smoothing, multiple linear regression or intermittent demand model. The options offered in the Modeling dialog box vary dynamically with the type of model chosen.

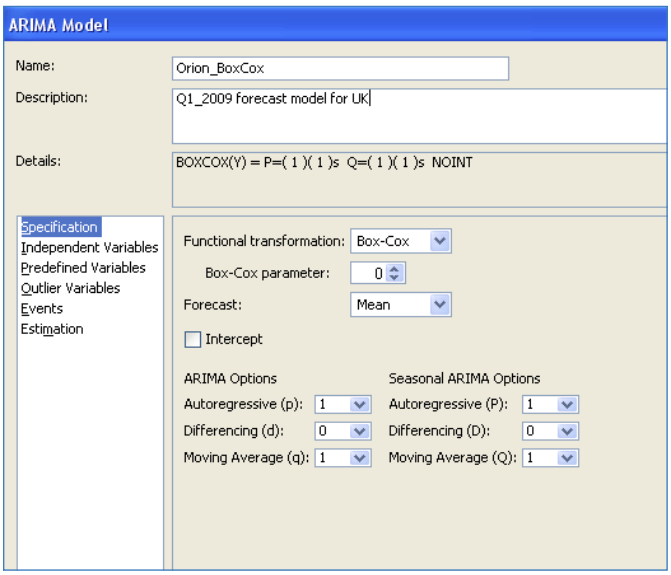


Figure 13: Sample modeling dialog box for ARIMA model.

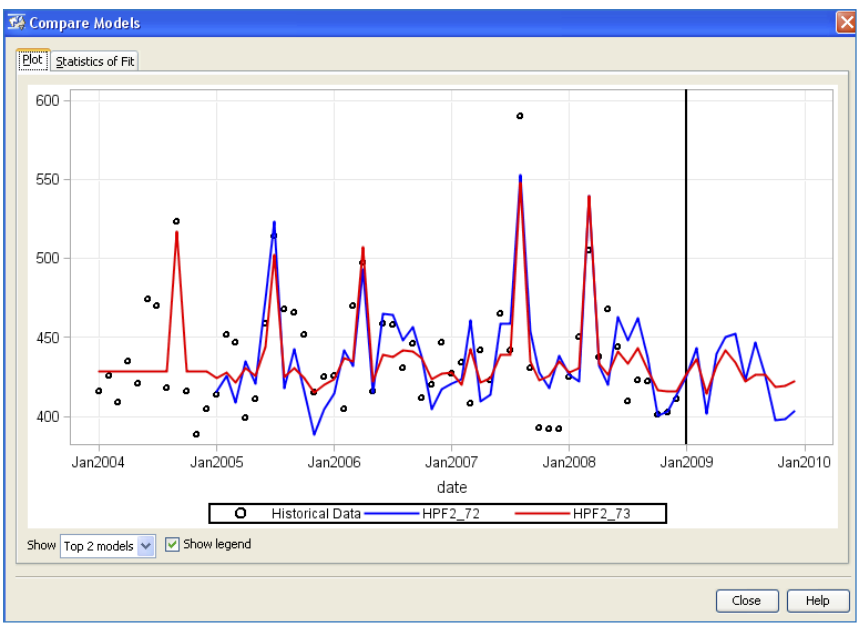


Figure 12: Model comparison.

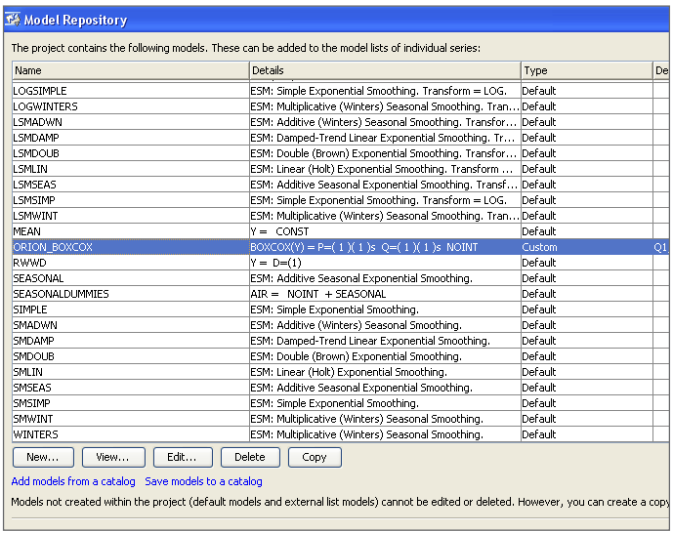


Figure 14: Model repository.

When the model specification is complete, the parameters of the new model are estimated and the model is added to the comparison table of models.

When a model is edited or created, the plots in the Modeling view are automatically updated, enabling users to easily evaluate the new model and compare it to existing models. In addition, models constructed or edited by the user can be added to the project's custom model repository for future use. In this way, the model repository can become increasingly tailored to the business or forecasting problem over time.

Scenario Analysis View

Forecasters and business analysts are frequently asked to conjecture about the future: "How many more units will be sold if the price is lowered by 5 percent?" or "What would the impact on production be of another late-winter blizzard?" The answers to questions like these can have a direct impact on the final forecast. One approach is to simply guess at the forecast impact by manually entering overrides into the forecast data table in the Forecasting view. A more sophisticated approach, however, is used in SAS Forecast Studio's Scenario Analysis view, in which the dynamic relationship between input variables and the forecast values, diagnosed and modeled by SAS Forecast Server, forms the crux of the scenario analysis.

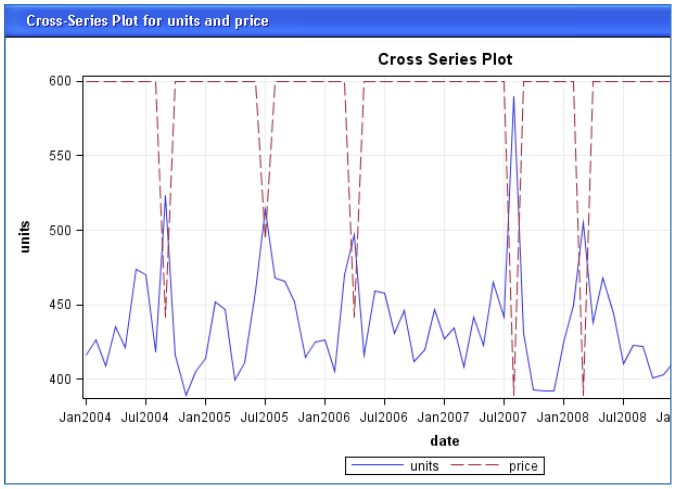


Figure 15: Cross-series plot.

SAS Forecast Studio scenarios are based on models that use independent variables, such as regression, ARIMAX and unobserved components models. Using the Scenario Analysis dialog box, a scenario is built based on a relationship (model) already estimated by the forecasting engine. New values for the independent variables are entered and the model is rerun, generating new forecast values. Multiple scenarios can be run and saved and the results compared. In addition, a scenario's results can be used as forecast overrides and incorporated in the final reconciliation of the forecasting project.

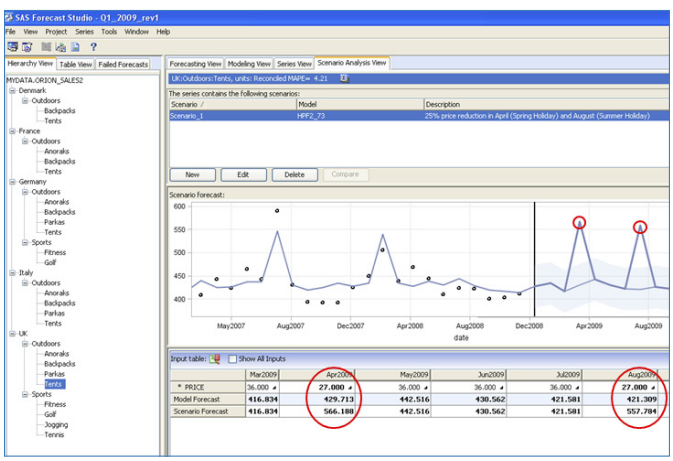
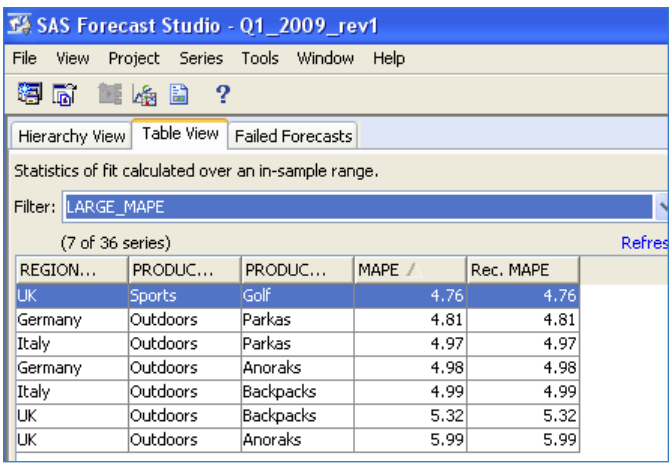


Figure 16: Overriding forecast based on scenario analysis.

Exception-Based Forecasting

When faced with forecasts for thousands of time series, manually reviewing each forecast is simply not feasible – the effort and time required would not be commensurate with the reward. Instead, forecasters and planners turn to exception-based forecasting, in which business rules and forecasting criteria are established, forecasts that meet all the rules and criteria are automatically passed on, and the analysts focus their attention on those forecasts that fail to meet a rule or pass a forecasting criterion. These are the forecast exceptions.



REGION...	PRODUC...	PRODUC...	MAPE	Rec. MAPE
UK	Sports	Golf	4.76	4.76
Germany	Outdoors	Parkas	4.81	4.81
Italy	Outdoors	Parkas	4.97	4.97
Germany	Outdoors	Anoraks	4.98	4.98
Italy	Outdoors	Backpacks	4.99	4.99
UK	Outdoors	Backpacks	5.32	5.32
UK	Outdoors	Anoraks	5.99	5.99

Figure 17: Forecast exceptions.

All the tools needed to set up business rules and flag potentially problematic forecasts are readily available in SAS Forecast Studio:

- Business rules and forecasting criteria can easily be set up using the Filter Generation dialog boxes. For example, the analyst may be happy with all forecasts with a MAPE less than 5 percent but may want to examine forecasts with a MAPE value that exceeds this threshold.
- Using the filter drop-down box in the List view, the list of forecasts requiring additional scrutiny can be generated quickly.
- Using the tools in the Forecasting view, forecasts can be manually adjusted using the forecast override function.
- In the Series view, the properties of the time series underlying a forecast exception can be evaluated, giving insight into the behavior of the forecast.
- Tools in the Model view allow the analyst to create or edit forecast models to generate, if possible, forecasts that comply with business rules and forecast exceptions.

By following the principles of exception-based forecasting, an analyst can quickly identify the forecasts which are performing poorly and focus additional attention on those time series. The net result of the focused effort stemming from exception-based forecasting is a significant increase in efficiency of the entire forecast process and a shortening of forecast-cycle time.

Forecast Reports

In addition to the powerful forecasting and scenario analysis capabilities of SAS Forecast Studio, several reporting tools are available. Graphs can easily be copied and pasted into presentations, and users can choose among several predefined graph styles as well as control the format of the data table.

The predefined reports in SAS Forecast Studio are based on SAS Stored Processes, which comprise an open and extensible system. Users can write custom reporting stored processes and add them to the reports catalog. Predefined reports include reports of the final forecast (forecast values and plots) as well as detailed reports of series and model diagnostics and hierarchy reconciliation.

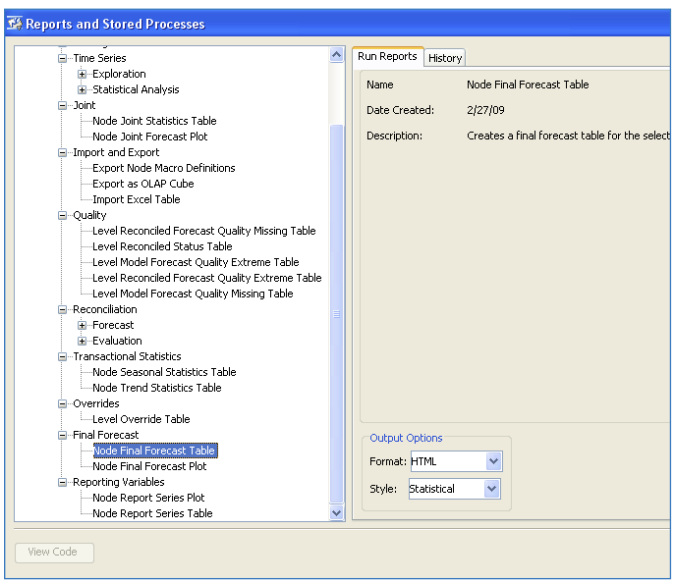


Figure 18: Available reports and stored processes.

Conclusion

SAS Forecast Studio delivers the sophisticated time series exploration, model building and large-scale automatic forecasting capabilities of SAS to novice forecasters through the SAS Forecast Studio interface. Advanced users can access SAS Forecast Server Procedures (writing SAS code) if they so desire. Thus, as an extensible and flexible system, SAS Forecast Studio meets the needs of both novice forecasters who need to move through the production forecasting process quickly as well as more experienced forecast analysts who wish to delve deeply into the forecast model building process.

SAS Forecast Studio supports exception-based forecasting; forecasts that violate business rules or forecasting criteria can be easily identified, allowing the analysts to focus their time and energy where it is most needed.

The scenario analysis capability in SAS Forecast Studio provides tools for what-if modeling, using the models and parameter estimates generated by the forecasting system or provided by the user. This means that SAS Forecast Studio scenarios leverage the underlying, potentially complex, dynamic relationship between the forecast variable and the independent variables, adding significantly to their value in the business planning process.

References

SAS White Papers

Turbo-Charging Spreadsheets: Accessing SAS® Forecast Server from Microsoft Excel.

Large-Scale Automatic Forecasting Using Inputs and Calendar Events.

Structured, Large-scale Statistical Forecasting Using SAS® Forecast Server

Recommended Reading

Box, G. E. P, Jenkins, G. M., and Reinsel, G. C. (1994), *Time Series Analysis: Forecasting and Control*, Englewood Cliffs, NJ: Prentice Hall Inc.

Brocklebank, John C. and Dickey, David A. (2003), *SAS® for Forecasting Time Series*, Second Edition.

Chatfield, C. (2000). *Time Series Models*, Boca Raton, FL: Chapman & Hall/CRC.

Gilliland, M. (2010). *The Business Forecasting Deal*, Hoboken, NJ: John Wiley & Sons Inc.

Makridakis, S. G., Wheelwright, S. C., and Hyndman, R. J. (1997), *Forecasting: Methods and Applications*, New York: John Wiley & Sons Inc.

To contact your local SAS office, please visit: sas.com/offices

102418_S144585.0915