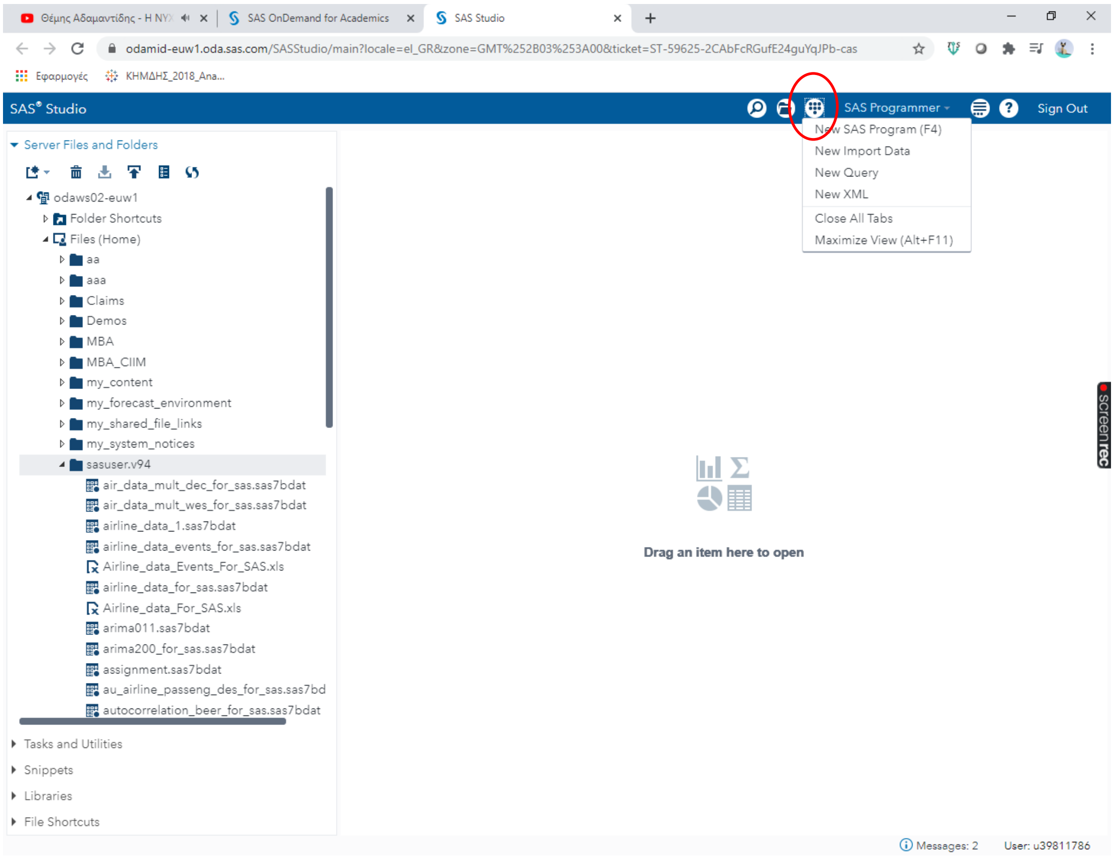
The purpose of the demonstration in this section is to show how a custom specified model can be included in SAS Forecast Studio 4.1 for refining a high-value and problematic series to improve precision of the forecasts. A thorough knowledge of the time series data and econometrics modeling could help to develop custom specific models.

1) Open google chrome
2) Go to odamid.oda.sas.com
3) Select Europe in the drop down menu and press sign in
4) Insert your credentials
5) In the dashboard select SAS Studio
6) On the left hand side select Server Files and Folders and then sasuser.v94
7) Press the upload button (the one with the arrow heading upwards).
8) Select the file “ssales4_31.sas7bdat”
9) Press Upload
10) Press the options button as shown below and select New SAS Program



11) Ctrl + V the following code:

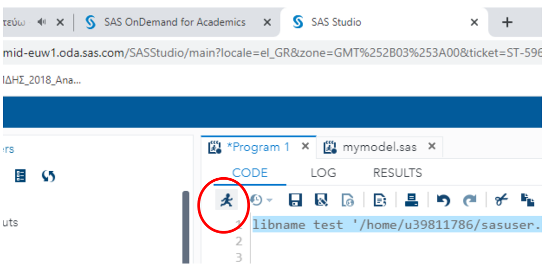
```
libname test '                    ';
proc hpfarimaspec repository=test.mymodels
        specname=myarima_1;
        forecast symbol=units p=(1 2 3 52);
        input symbol=coupon num=(1 2);
        input symbol=instore;
run;

proc hpfselect repository=test.mymodels
        name=mymsl;
        spec myarima_1;
run;
```

12) On the left hand side select Server Files and Folders and then sasuser.v94

13) Right click and properties.

14) Copy the location (Ctrl + C) and paste it in the first line of the code in step 7 between the quotes ‘’ e.g. libname test '/home/u39811786/sasuser.v94';

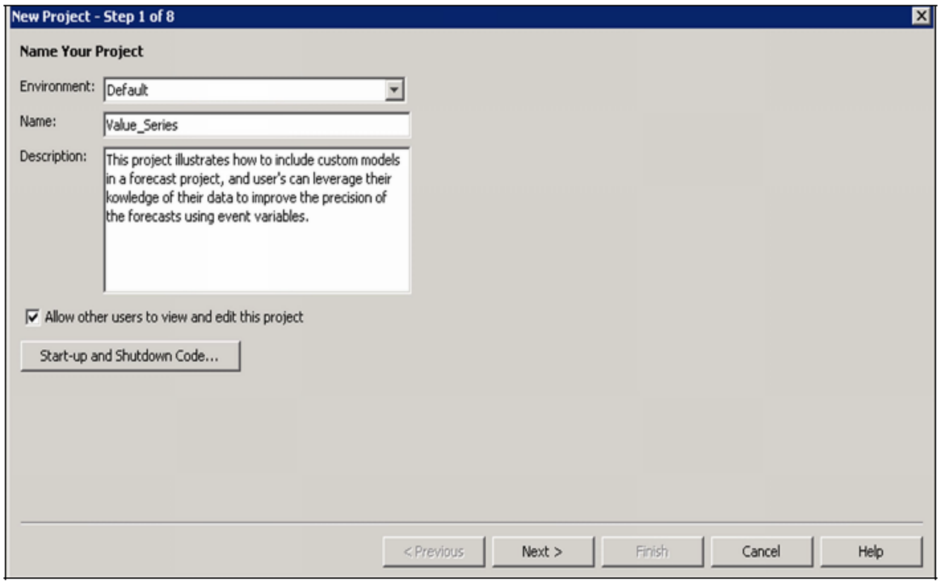
15) Press the Run button



16) Go to odamid.oda.sas.com
17) Select Europe in the drop down menu and press sign in
18) Insert your credentials
19) In the dashboard select SAS Forecast Studio
20) In the bottom of the browser press keep (Διατήρηση)
21) In the bottom of the browser click on the main.jnlp file
22) When prompted press "Run".
23) In the projects window press New.

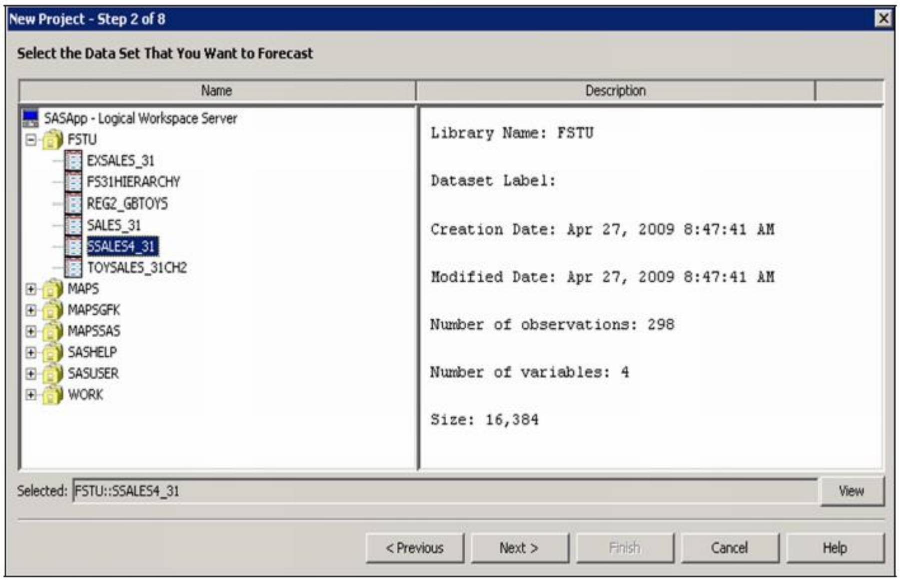
Setting Project Options

1. Select **File** ⇨ **New Project**.



2. Name the project **Value_Series** and provide a brief description.

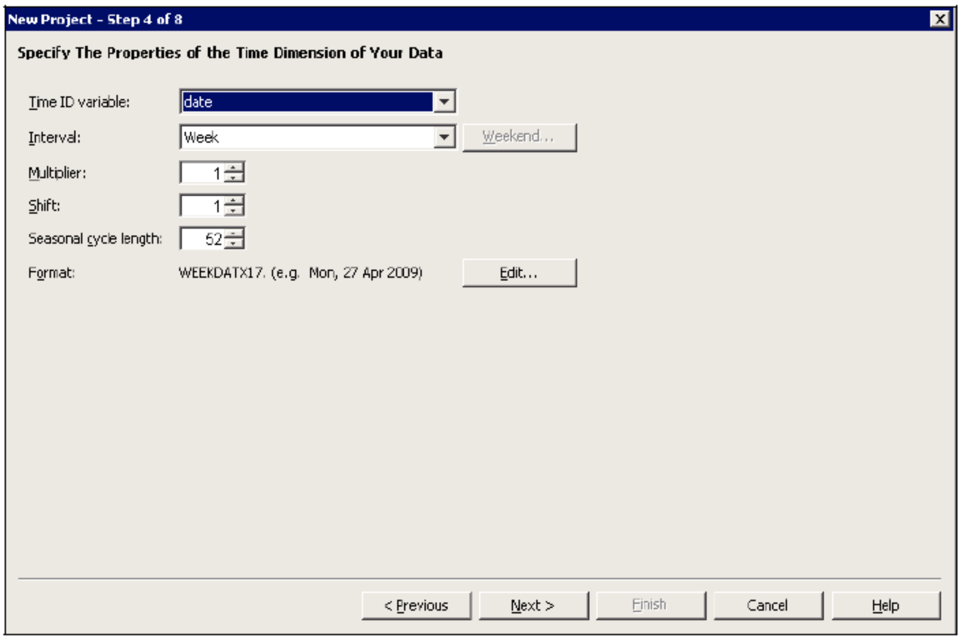
3. The data set for this demonstration is **SSALES4_31**.



This is a high-value series. The focus of this chapter is on investigating system tools for refining high-value and problematic series. The main purpose of the tools discussed is to help you leverage your knowledge of the data to improve precision of the forecasts.

This is a single series, and there is no associated hierarchical structure to the data; skip to step 4 in the wizard.

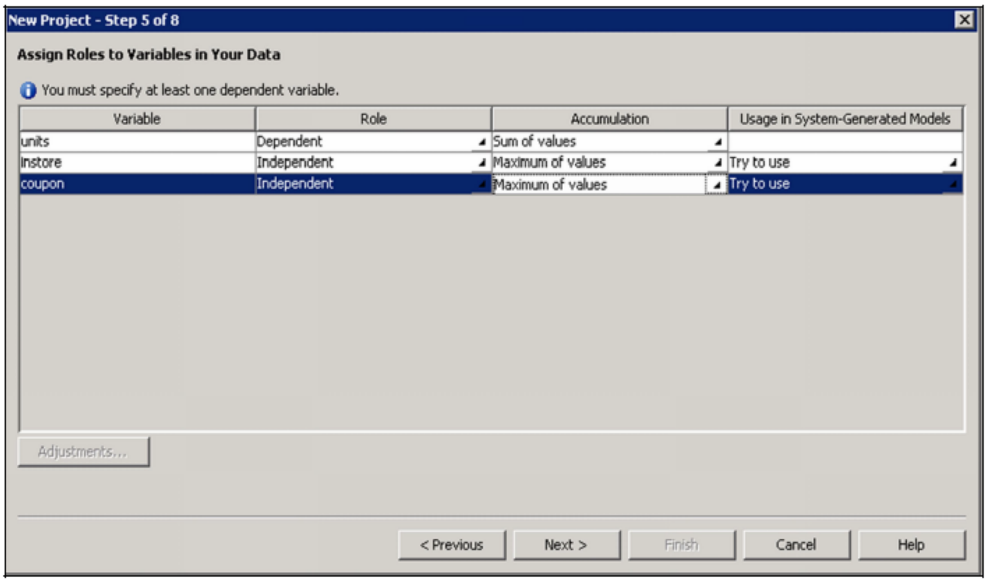
4. The time ID variable is **date**. The interval of the data is **week**.



5. The variable **units** is the units of sales within a store, and it is the dependent variable in the project. The default accumulation method, Sum of Values, is appropriate for this variable.

The variable **instore** is a binary promotion flag that represents promotional discounts for the product. Max is used as the accumulation method to preserve the zero-one coding of the variable.

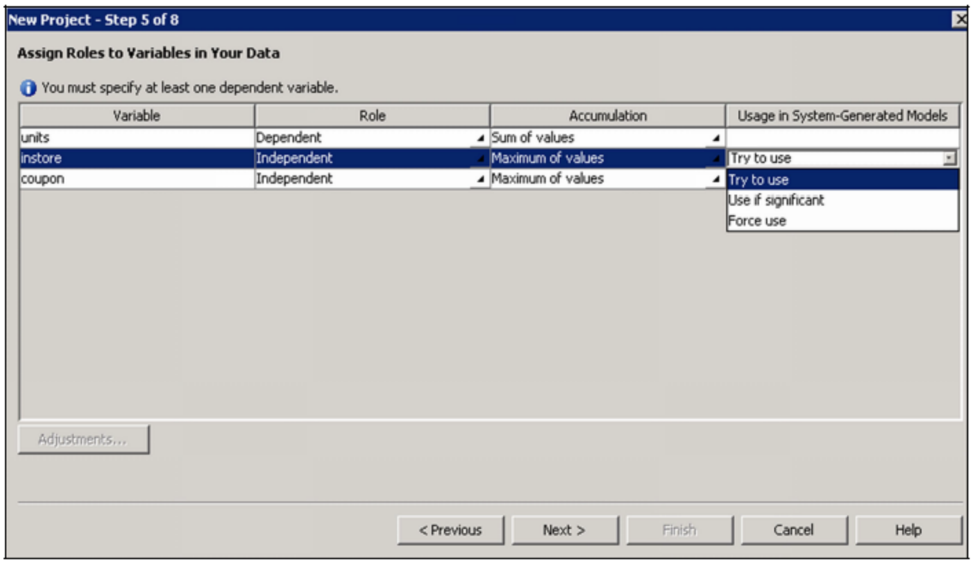
The variable **coupon** is a binary promotion flag that represents mailings and handouts of discount coupons for the product.



6. **Try to Use**: If inclusion of the independent variable significantly improves the overall fit of the model, it is included in the model (default).

   Use if Significant: If the estimated slope associated with the independent variable is significantly different from zero, it is included in the model.

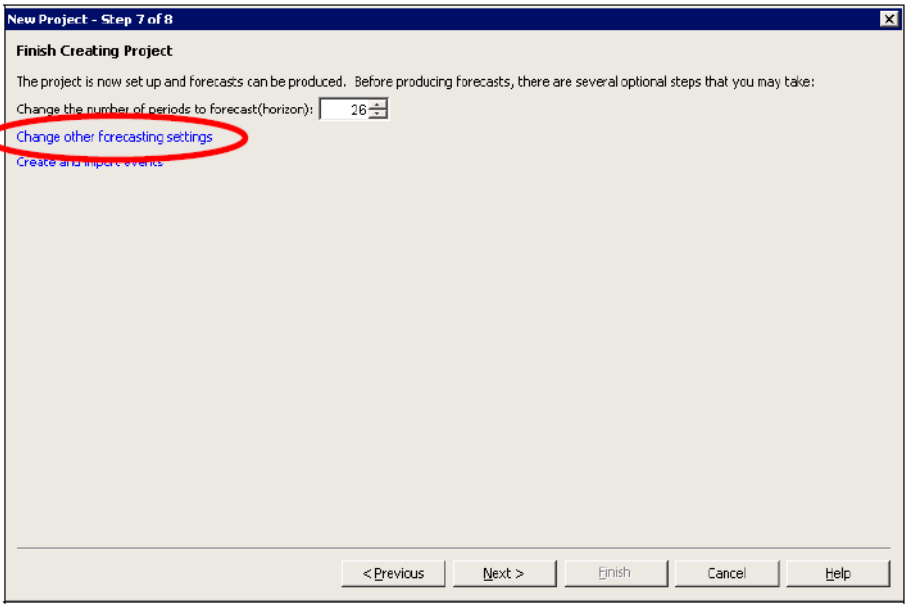
   Force Use: The independent variable is included in any system-generated model that accommodates independent variables and that does not fail.



If more than one candidate independent variable is specified, the collinearity between independent variables plays a role in the selection process. By default, ***the non-collinear combination of independent variables that most improves the overall fit of the model is chosen.***

If no data preparation is necessary, skip to step 7.

7. Forecast approximately six months ahead; change the lead forecast horizon to 26 weeks.

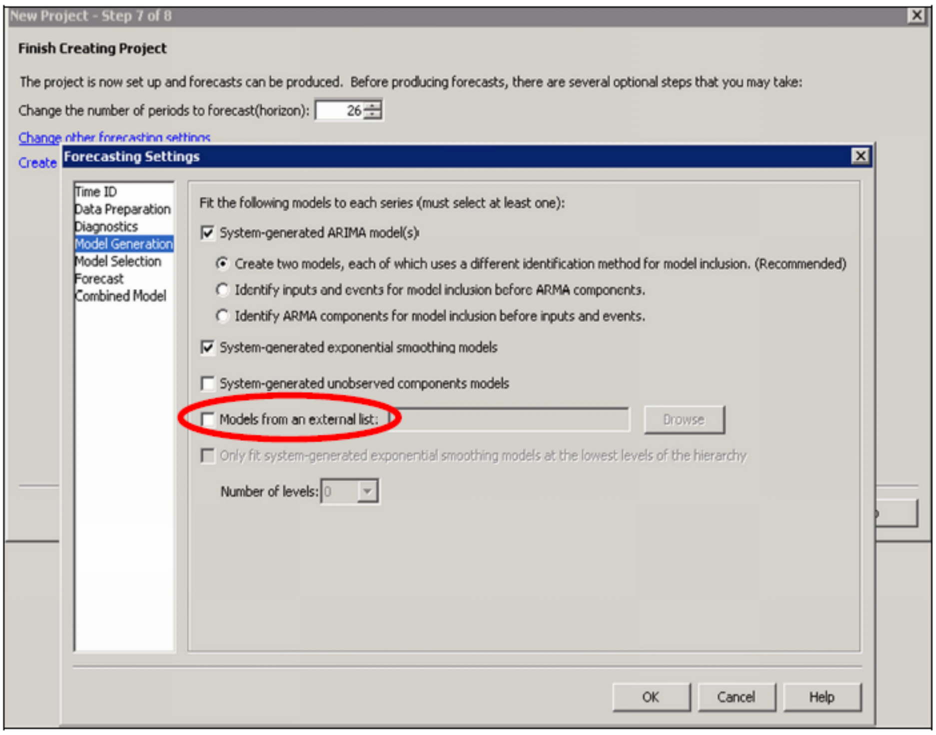


Custom, user-defined models can be included in the model selection process.

8. Select **Change other forecasting settings**.

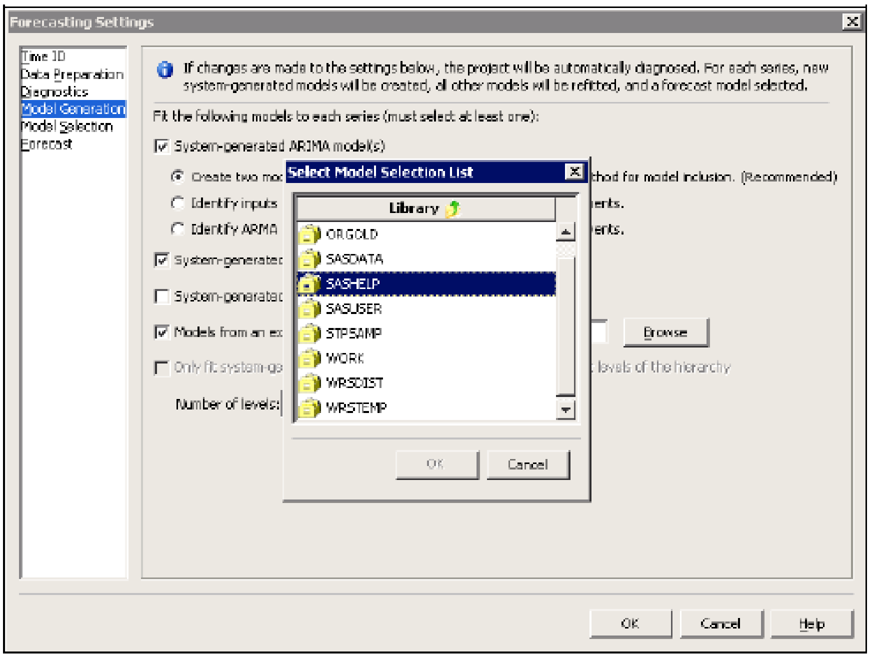
Options to include custom, user-specified models are under Model Generation. These models live on external Model Selection Lists.

9. Select the check box next to **Models from an external list**.

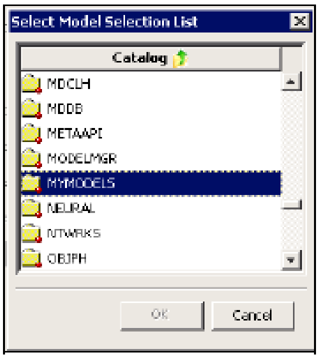


10. Click **Browse**.

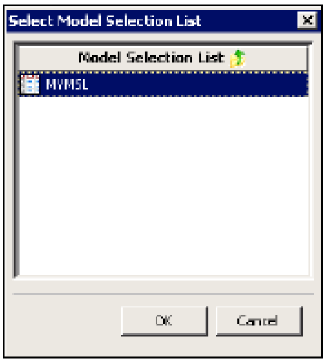
11. Scroll down and double-click the **SASHELP** library.



12. Select the **MYMODELS** catalog.

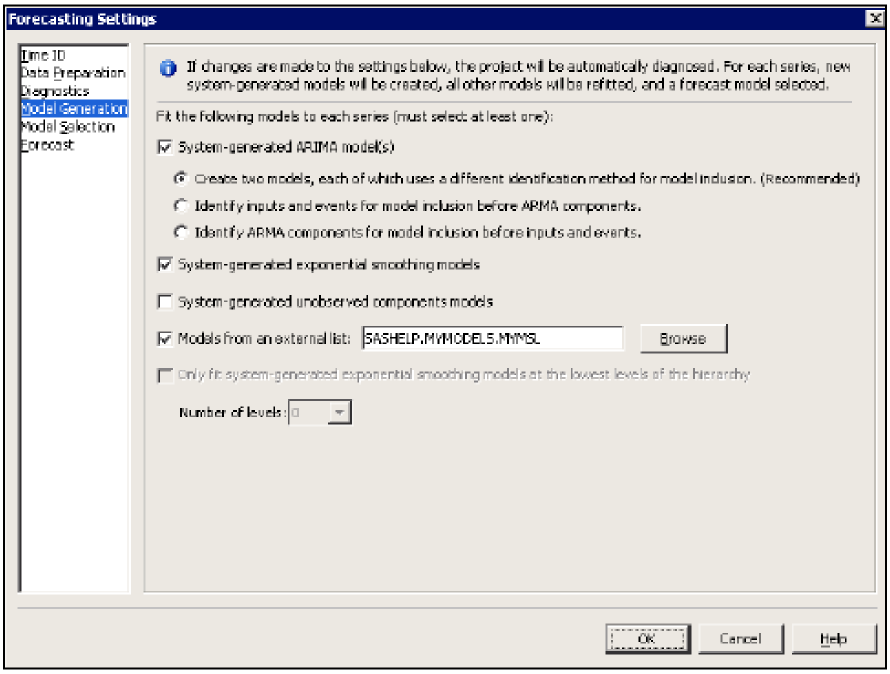


13. Select the **MYMSL** model selection list.



In SAS Forecast Server, SAS catalogs live in SAS libraries. Model selection lists live in SAS catalogs. Models can be system generated or prespecified (see above).

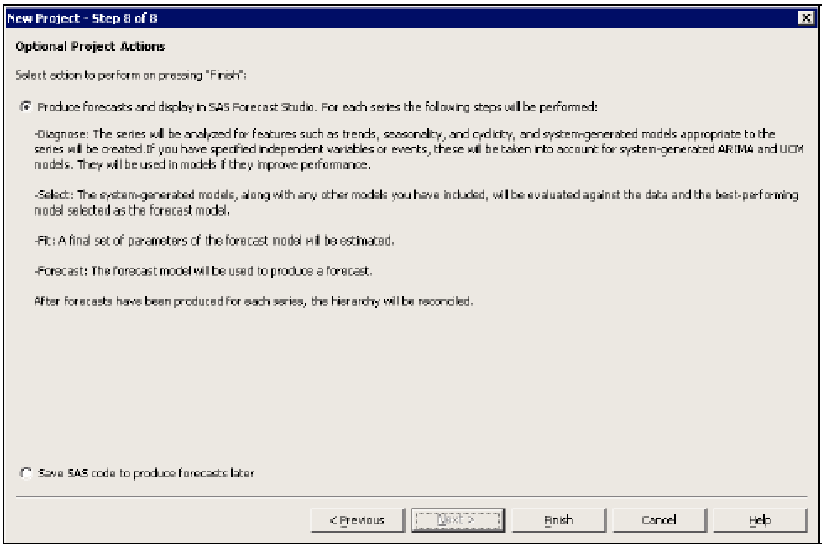
The model generation settings for this project are shown below.



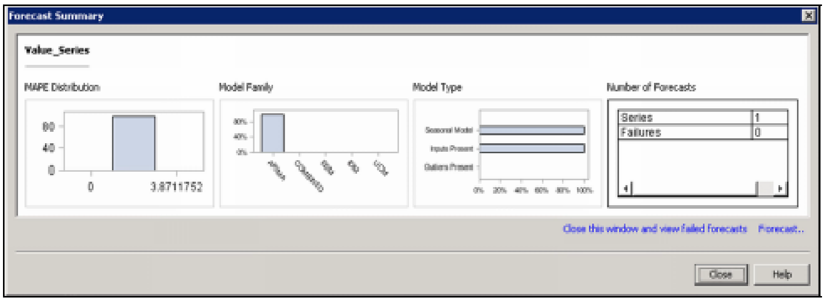
14. Click **OK** and then click **Next**.

All model specifications on the MYMSL model selection list will be fit to every series in the project and evaluated in the process of automatic model selection.

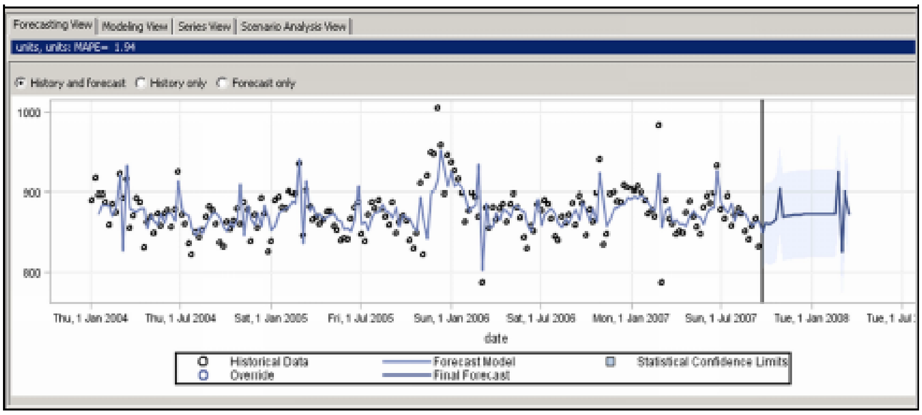
15. Click **Finish**.

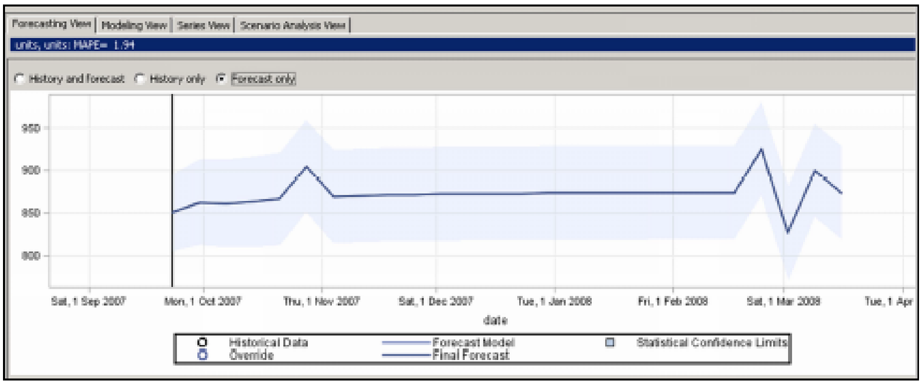


The selected forecast model is an ARIMA type that accommodates at least one of the candidate input variables. ***Seasonal variation is accommodated***. The MAPE associated with the forecast model seems quite low.



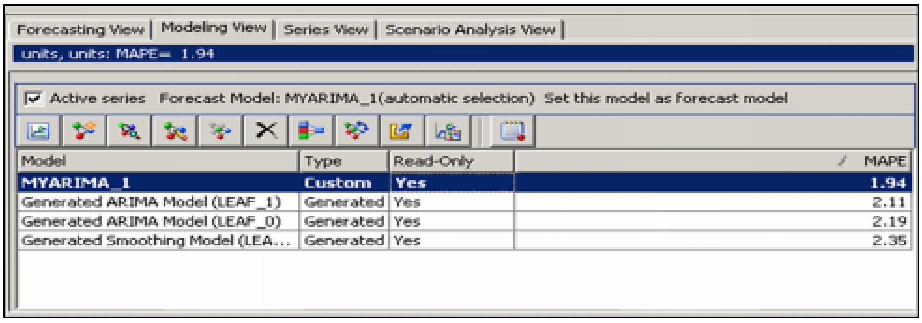
MAPE can be affected by the *scale of the series*. For example, a 10-unit difference between forecast and actual values in a series with a mean of 100 results in a difference in the neighborhood of 10%. The same 10-unit difference in a series with a mean of 1000 results in a difference in the neighborhood of 1%.





At least one of the promotional input variables seems to be scheduled to run in the lead forecast horizon.

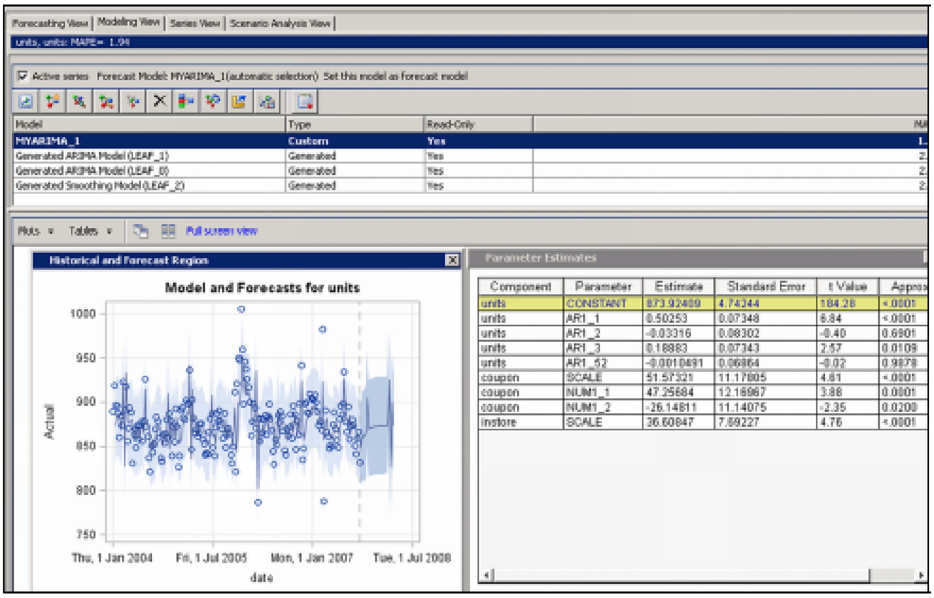
16. Switch to the Modeling view.



Model	Type	Read-Only	MAPE
MYARIMA_1	Custom	Yes	1.94
Generated ARIMA Model (LEAF_1)	Generated	Yes	2.11
Generated ARIMA Model (LEAF_0)	Generated	Yes	2.19
Generated Smoothing Model (LEA…	Generated	Yes	2.35

The custom model, **MYARIMA_1**, is selected as the forecast model, **MAPE=1.94**.

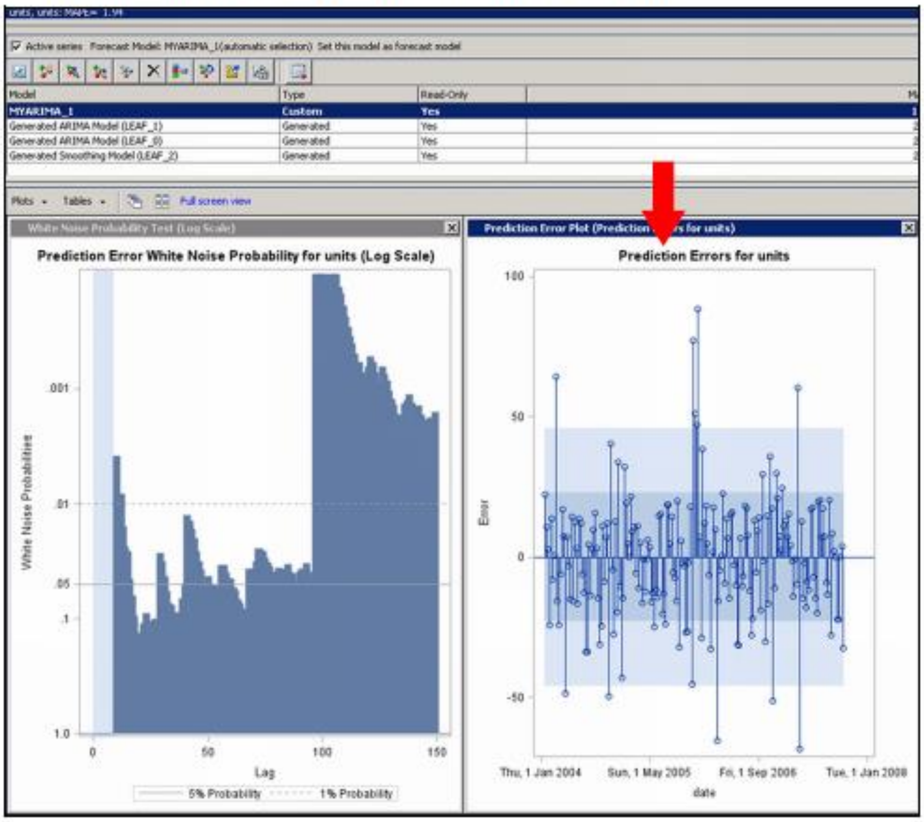
17. Select **Tables ⇨ Model Parameter Estimates**.

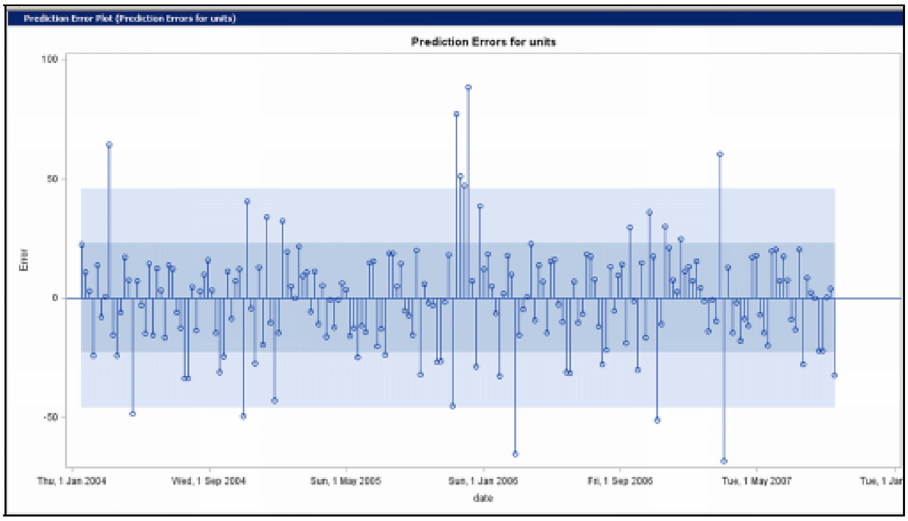


Both candidate input variables are included in the custom model.

While the generated MAPE for the forecast model seems quite low, a relevant question at this point is this: Should we try to improve the precision of the model? That is, can we make this model better? If the answer is yes, then what tools do we have to improve it?

One diagnostic to use in deciding whether to pursue further refinement of models is the plot of prediction errors. This plot is automatically generated in the Modeling view.

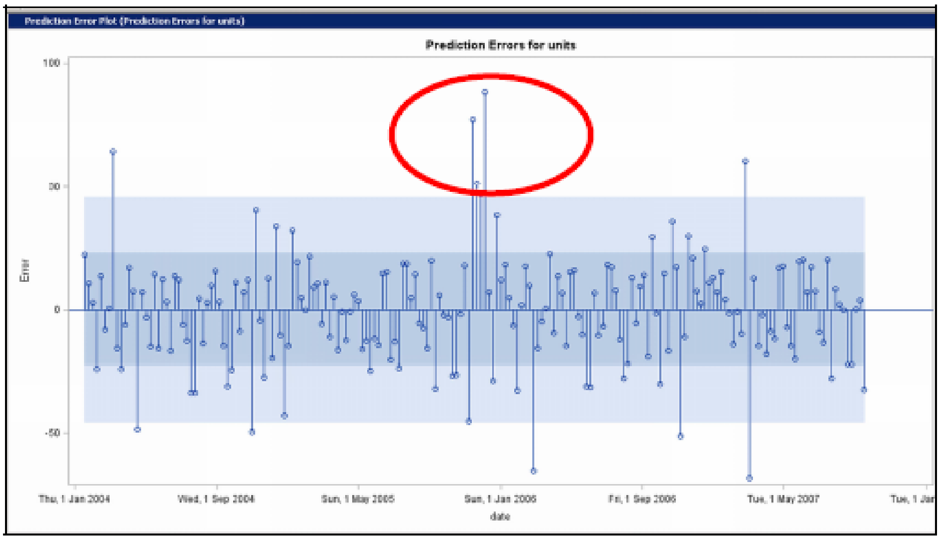




There appear to be several large errors in residuals of the model.

The dates of occurrence for large errors can be discerned by using the cursor to hover over different points in the plot.

The plot of prediction errors is the plot of residuals of the selected model in the Modeling View. The dark and light blue bands represent one and two standard errors. Large residuals, those falling outside the two standard error bands, might indicate undocumented promotions, competitor sales, and other demand drivers not captured in the model. They can also indicate random noise in the process generating the data. Systematic patterns in large residuals can be evidence of model inadequacy.

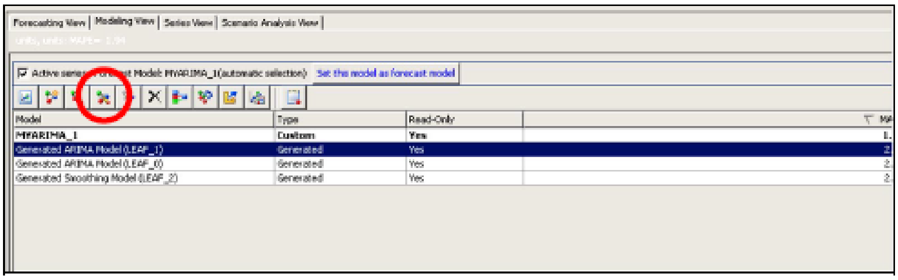


The marketing manager for this store is consulted. She describes the cluster of large residuals, circled above, as an undocumented, 20% price reduction aimed at reducing excess inventory of the product in the final quarter of 2005. The price reduction began on November 13, 2005, and continued until inventories reached targeted levels.

Event variables are very useful for incorporating this type of described variation into the model.

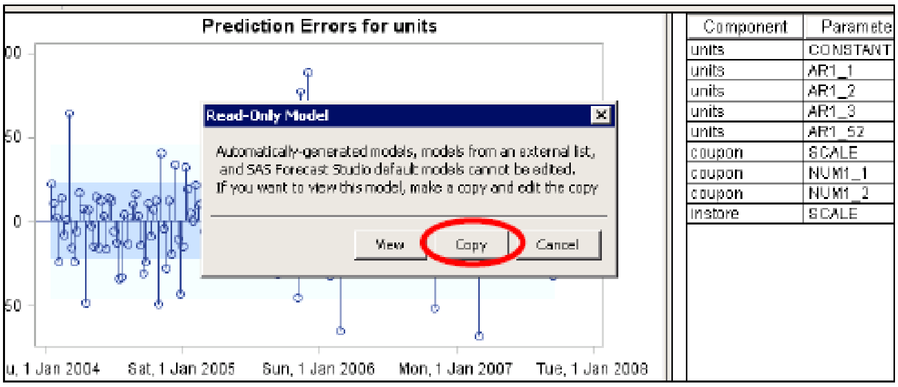
Creating Event Variables to Improve Forecast Precision

1. Highlight the runner-up model, **ARIMA, LEAF_1**, and click the **Edit Model** button.

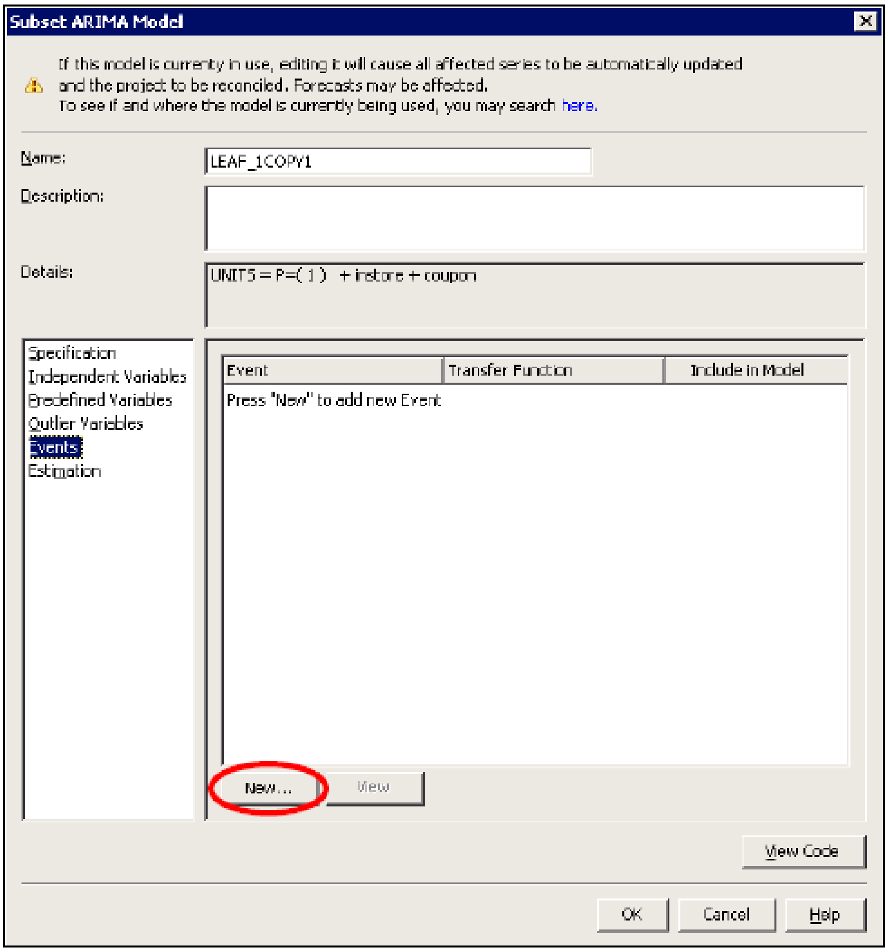


While the user-created forecast model could be edited, it has a complex characterization of the association between inputs and the target variable that could become confounded with new variables. It also has some problems with estimated parameter significance, noted above. The simpler, generated runner-up model will be used as a basis for accommodating the undocumented events described above.

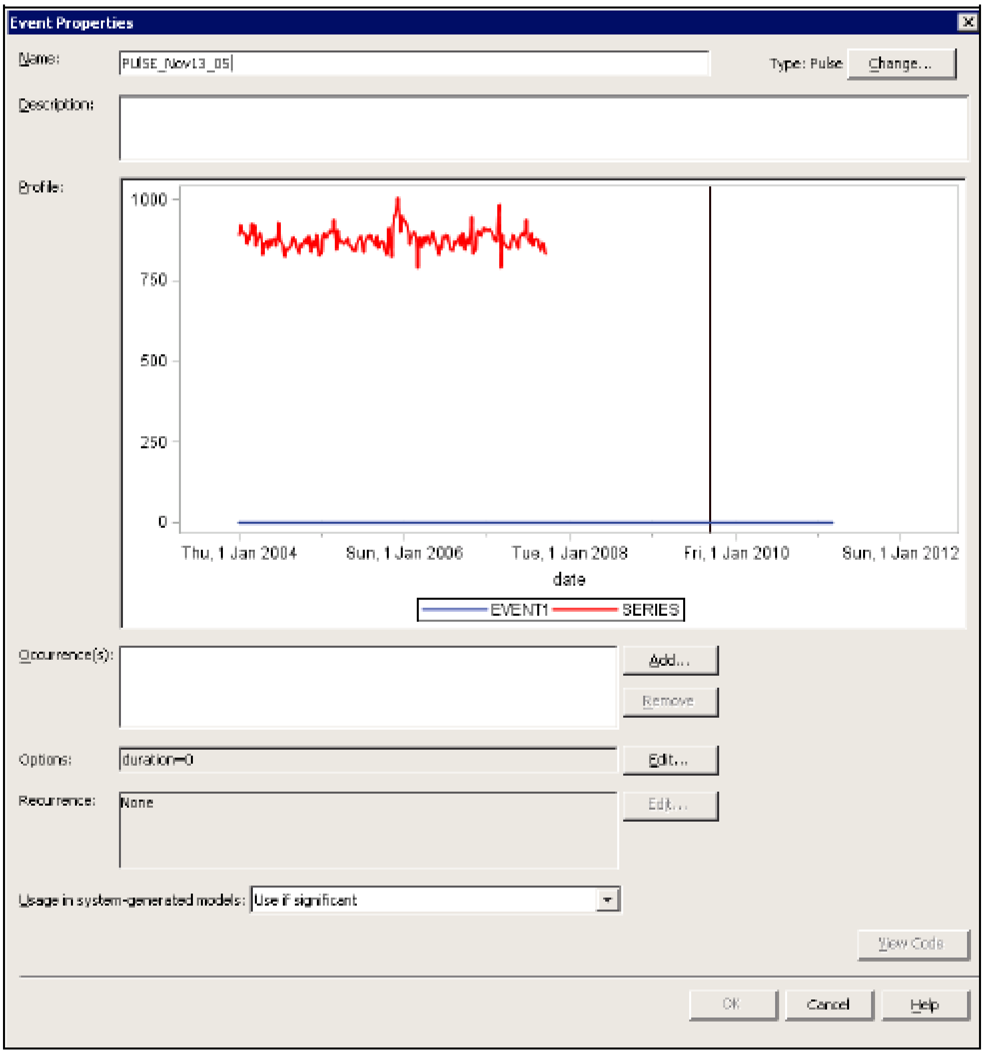
2. Automatically generated models and models from external lists cannot be edited directly. Select **Copy**.



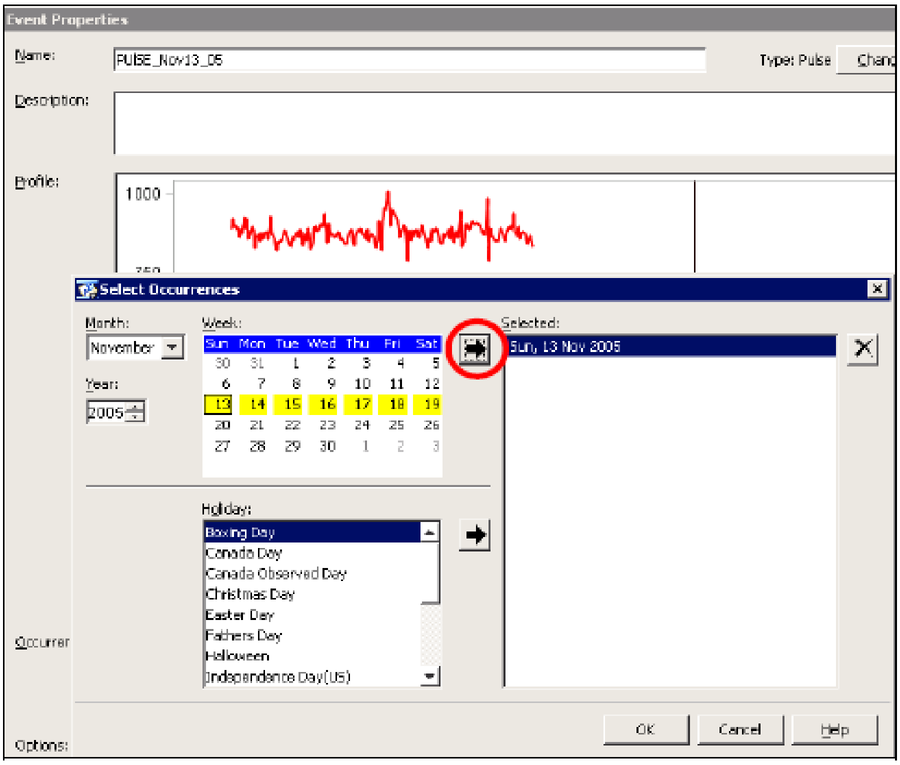
3. Select Events ⇨ New.



4. Name the new event variable **PULSE_Nov13_05**. The default type is p.
5. To add the occurrence date, click the **Add** button next to the Occurrence(s) box.

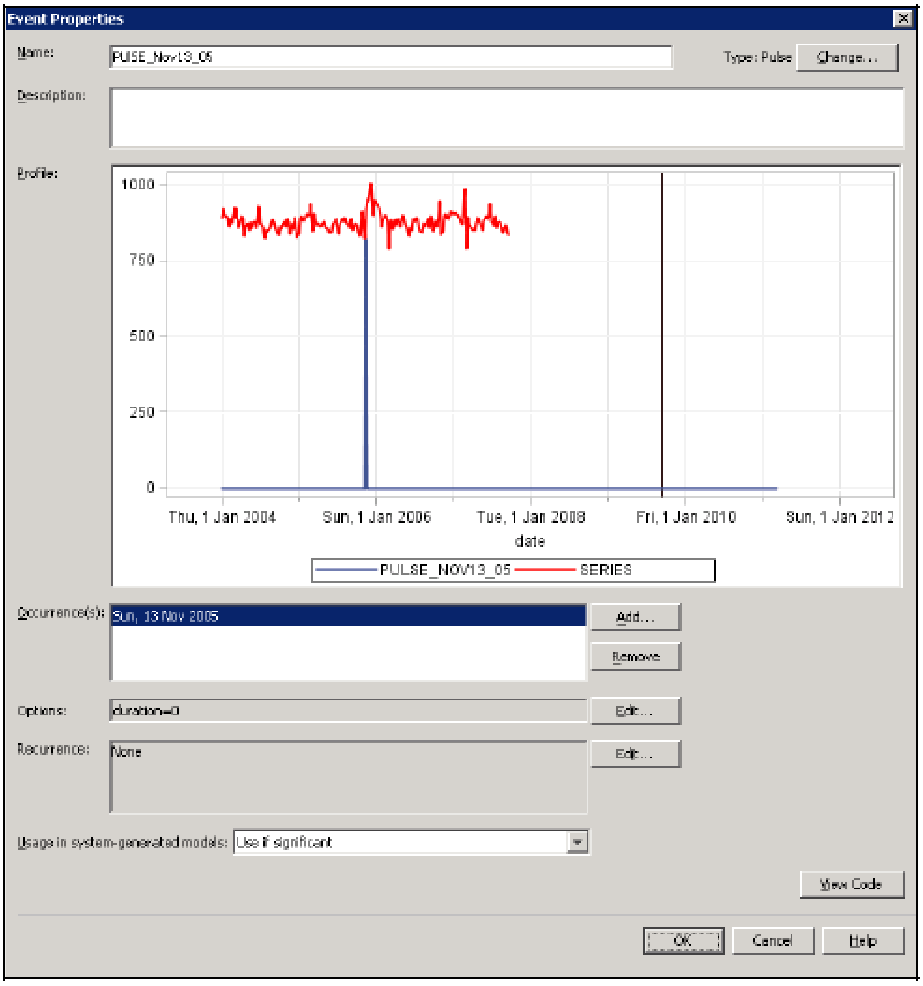


6. Set **2005** as the year and **November** as the month. Select the week beginning Sunday the 13th.
7. Choose the selector arrow so that the date is displayed in the Selected box.
8. Click **OK**.



The new event variable, **PULSE_Nov13_05**, is displayed below.

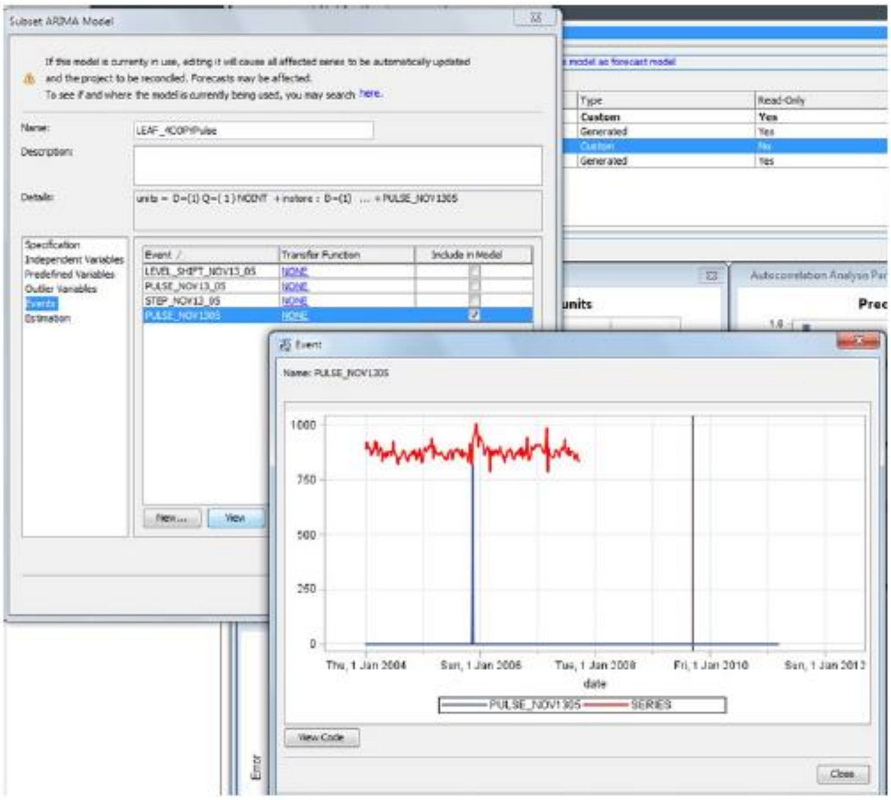
9. Click **OK**.



This is an abstract representation of the effect and characterization of the event.

The modified model specification with the included event variable is shown.

10. Click **OK**.

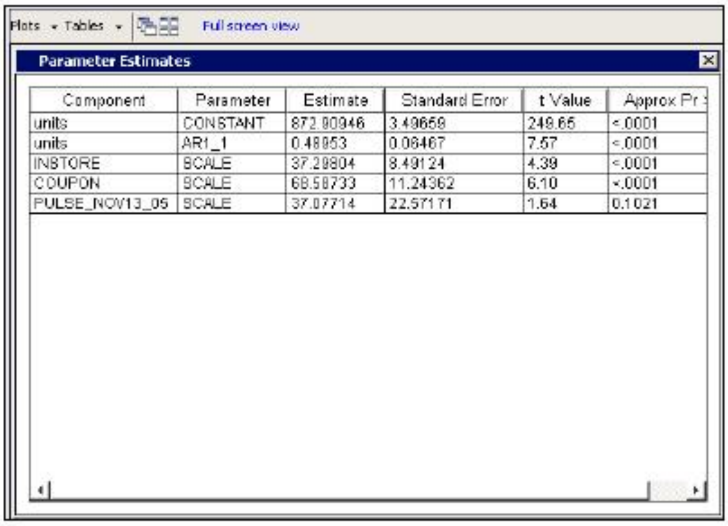


The modified forecast model specification shows very slight improvement.

Model	Type	Read-Only	MAPE
MYARIMA_1	**Custom**	**Yes**	**1.94**
LEAF_1COPY1	Custom	No	2.10
Generated ARIMA Model (LEAF_5)	Generated	Yes	2.11
Generated ARIMA Model (LEAF_4)	Generated	Yes	2.19
Generated Smoothing Model (LEAF_6)	Generated	Yes	2.35

The model type has been changed to Custom on **LEAF_1COPY1**.

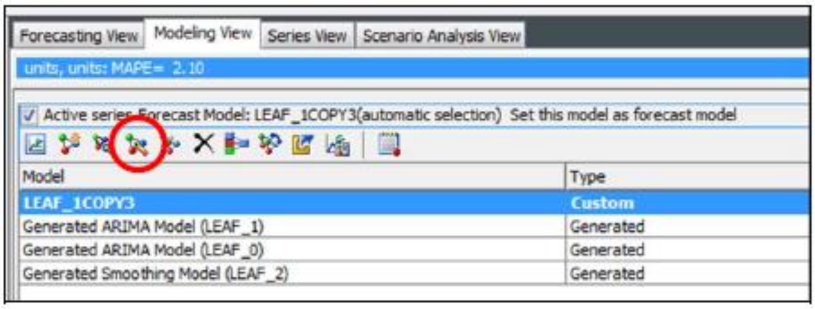
The estimated effect has the expected sign, but the estimated value of the event is not significantly different from zero.



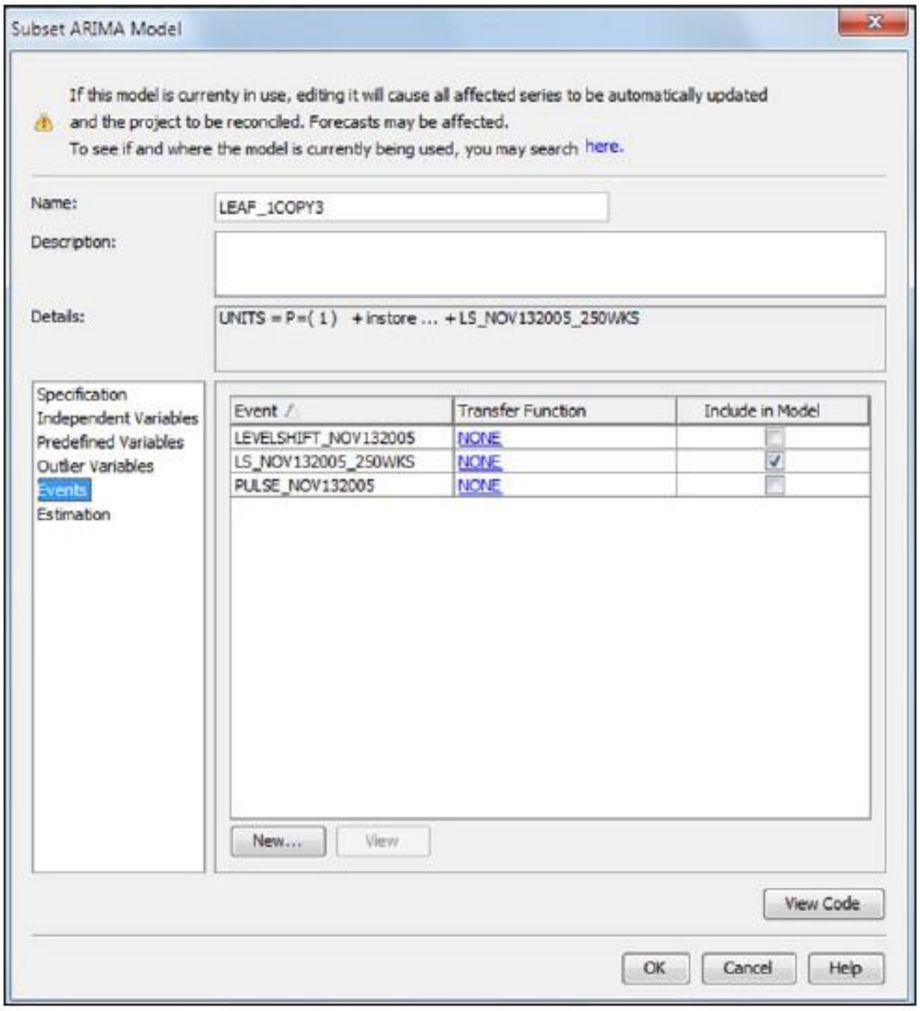
Component	Parameter	Estimate	Standard Error	t Value	Approx Pr
units	CONSTANT	872.80946	3.49659	249.65	<.0001
units	AR1_1	0.48853	0.06467	7.57	<.0001
INSTORE	SCALE	37.28804	8.49124	4.39	<.0001
COUPON	SCALE	68.58733	11.24362	6.10	<.0001
PULSE_NOV13_05	SCALE	37.07714	22.57171	1.64	0.1021

Another characterization of the event occurrence will be tried.

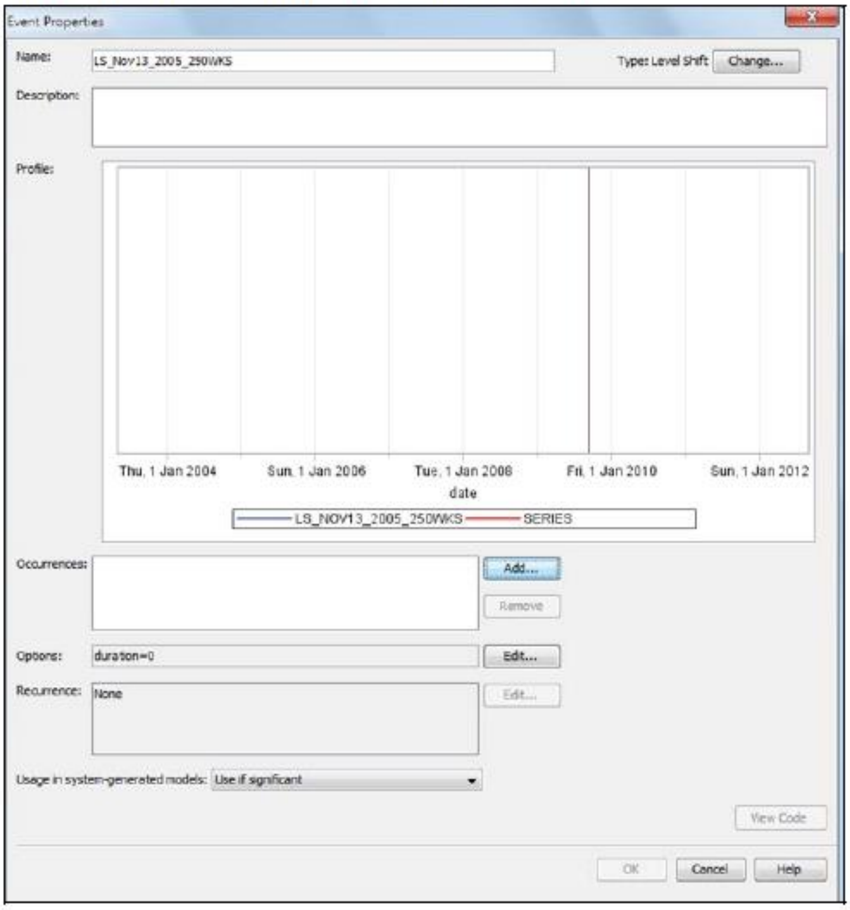
11. Select the **Edit Model** icon.



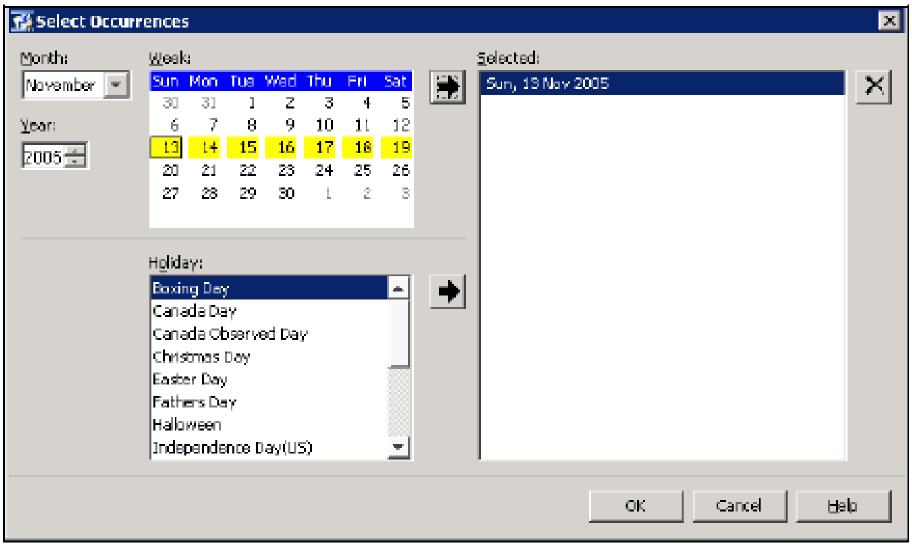
12. Select **Events** ⇨ **New** to create another characterization of the event effect.



13. Name the new event variable **LS_Nov13_2005_250WKS**.
14. Click the **Type** button and change the event type to **Level Shift** (Shift).
15. Click the **Add** button next to Occurrences to assign the date of occurrence.

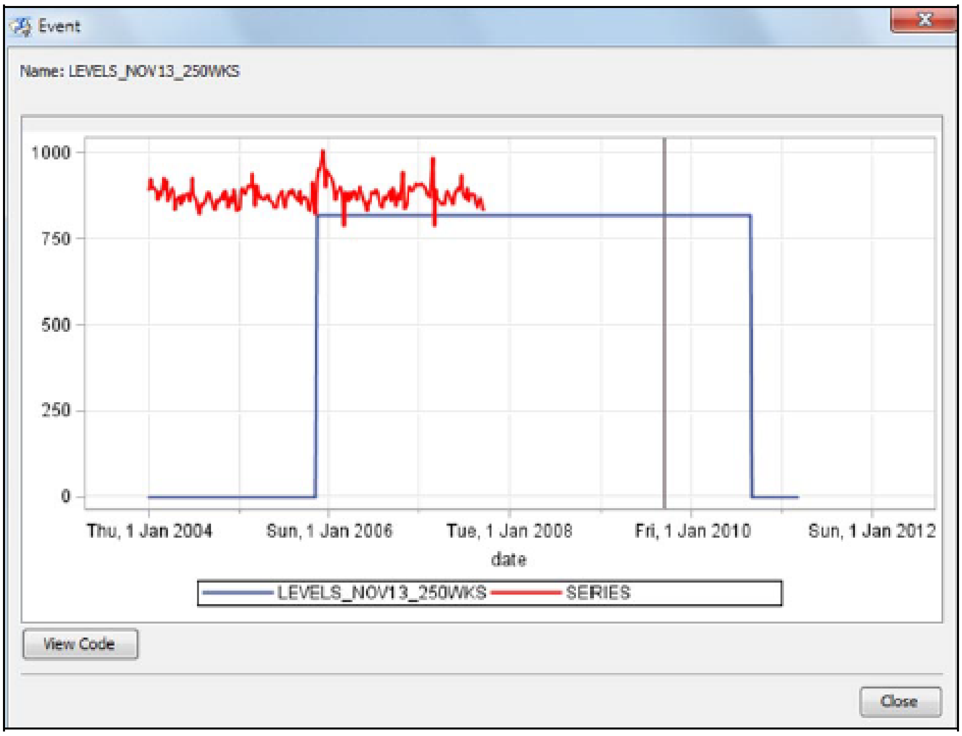


16. Set the date for the start of the Level Shift as 13 November, 2005.
17. Click **OK**.



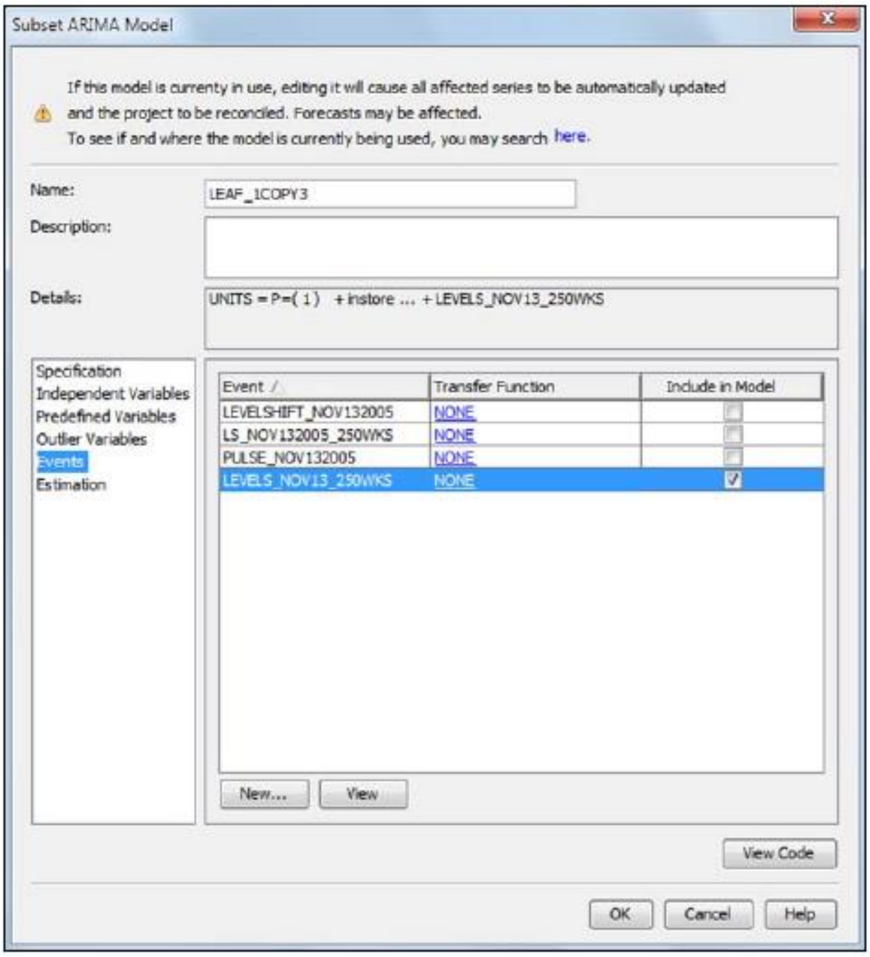
The plot indicates where the event variable begins flagging the shift in the series.

18. Select **OK**.

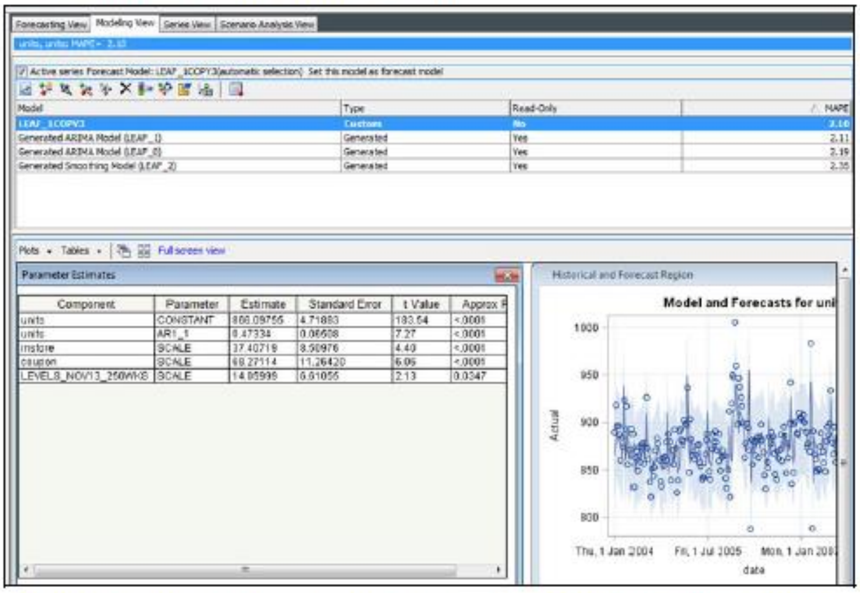


19. To focus on the relevancy of the level shift characterization of the 20% price reduction event, select the **LEVELS_NOV13_250WKS** event variable.

20. Click **OK**.

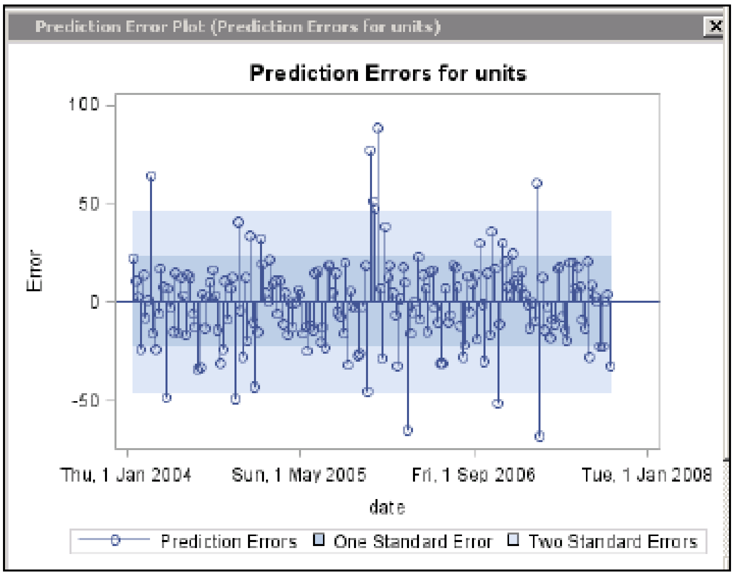


Again, there is only a very slight improvement in the fit of the model. The model parameter estimates associated with the step event variable are shown below.

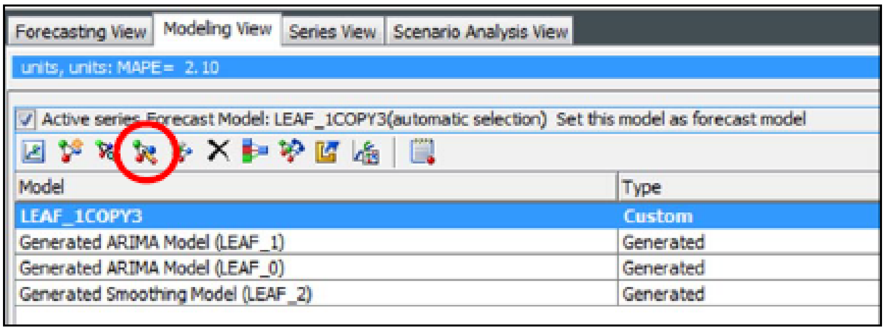


The estimated parameter has the expected sign, but the scale seems smaller than expected, and it is only marginally significant.

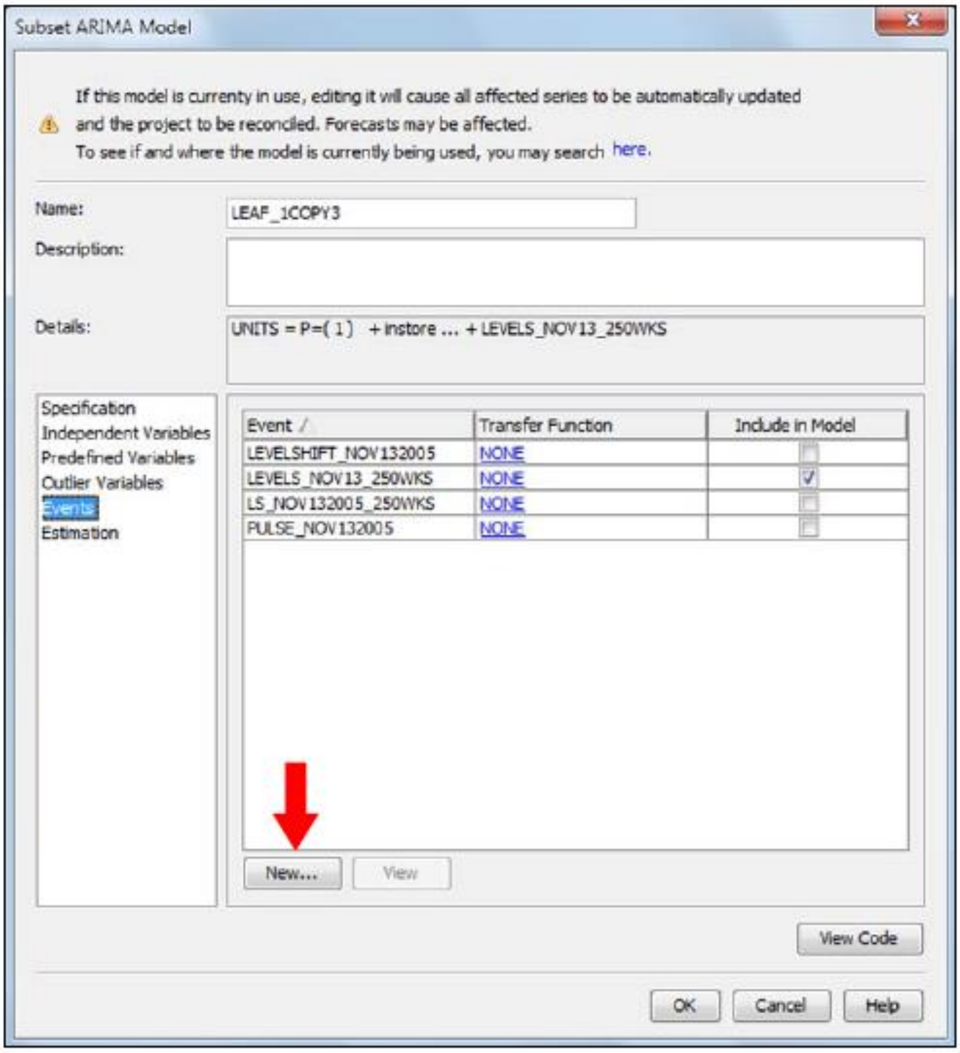
Take another look at the event. It begins on 13 Nov, 2005, and persists for a while, and then the discount ends and the effect vanishes. How can we characterize (qualify) the event variable to better capture this pattern?



21. Select the **Edit Model** icon again.



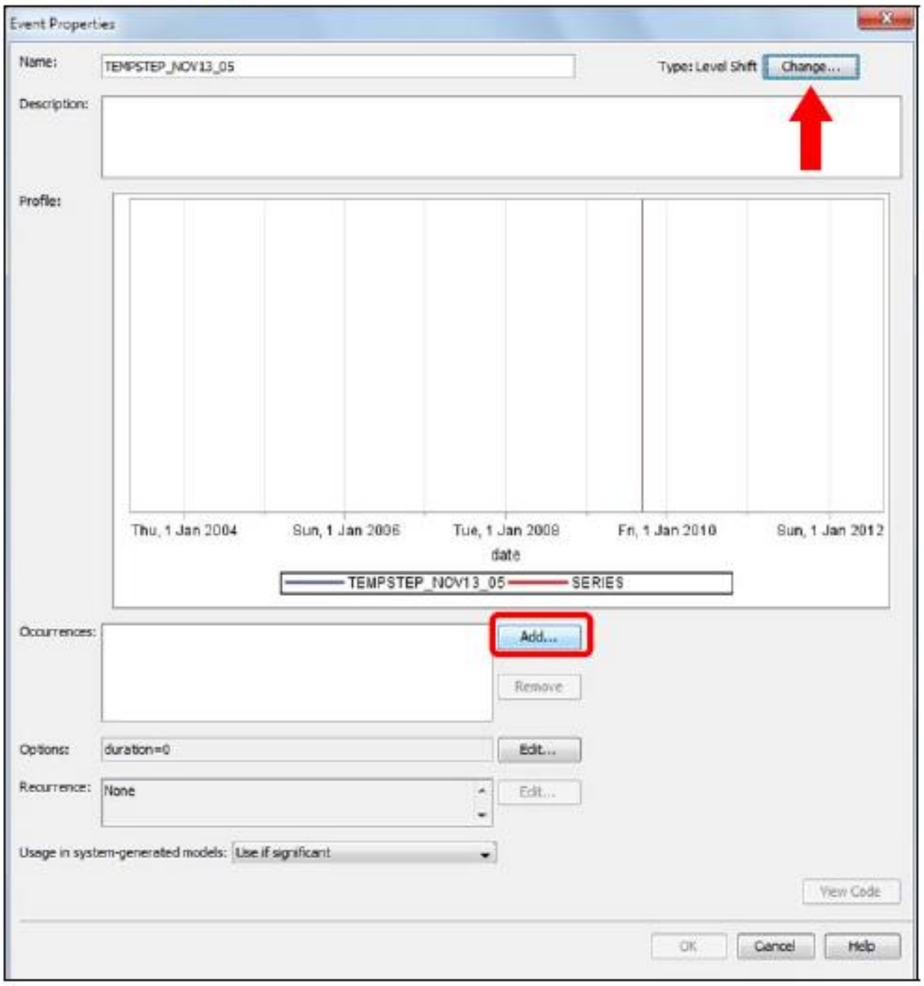
22. Select **Events** ⇨ **New**.



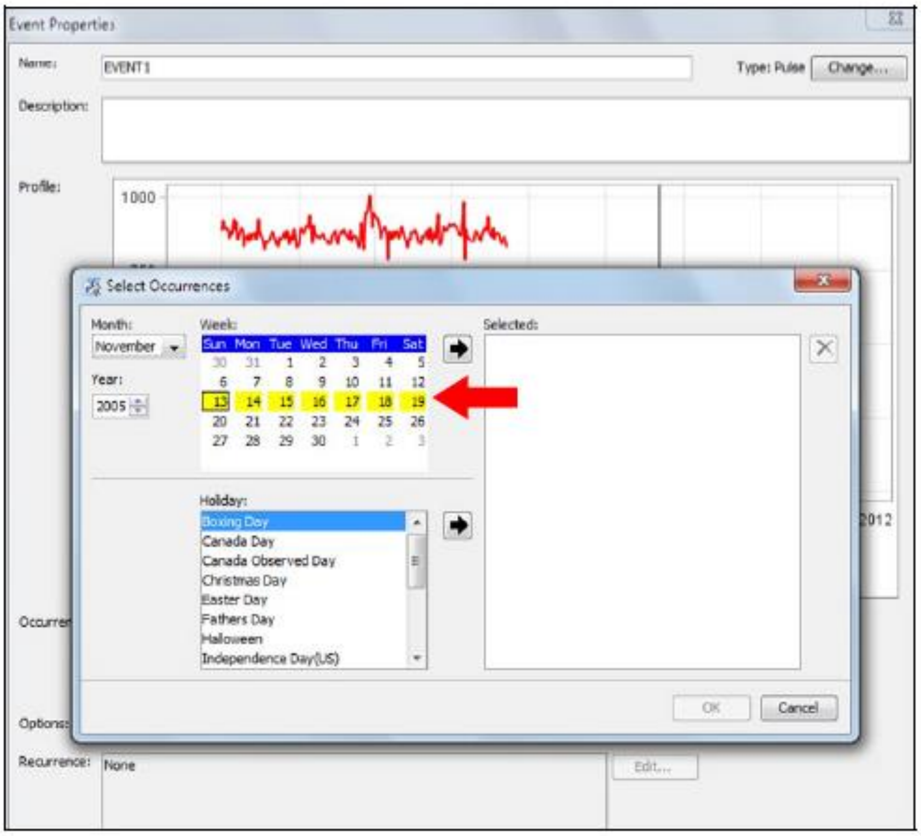
23. Name the new event variable **TEMPSTEP_NOV13_05**.

24. Change the Type to **Shift**.

25. Click **Add** to flag the first date of occurrence.

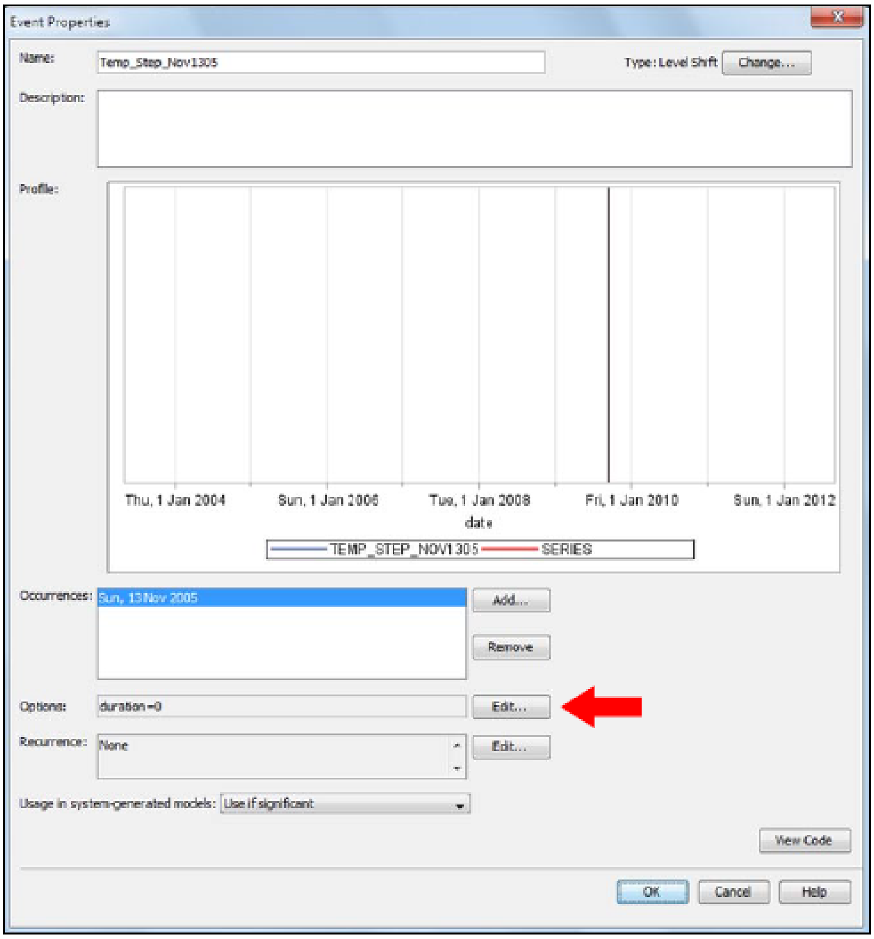


26. Set the beginning date of occurrence as before.
27. Click **OK**.



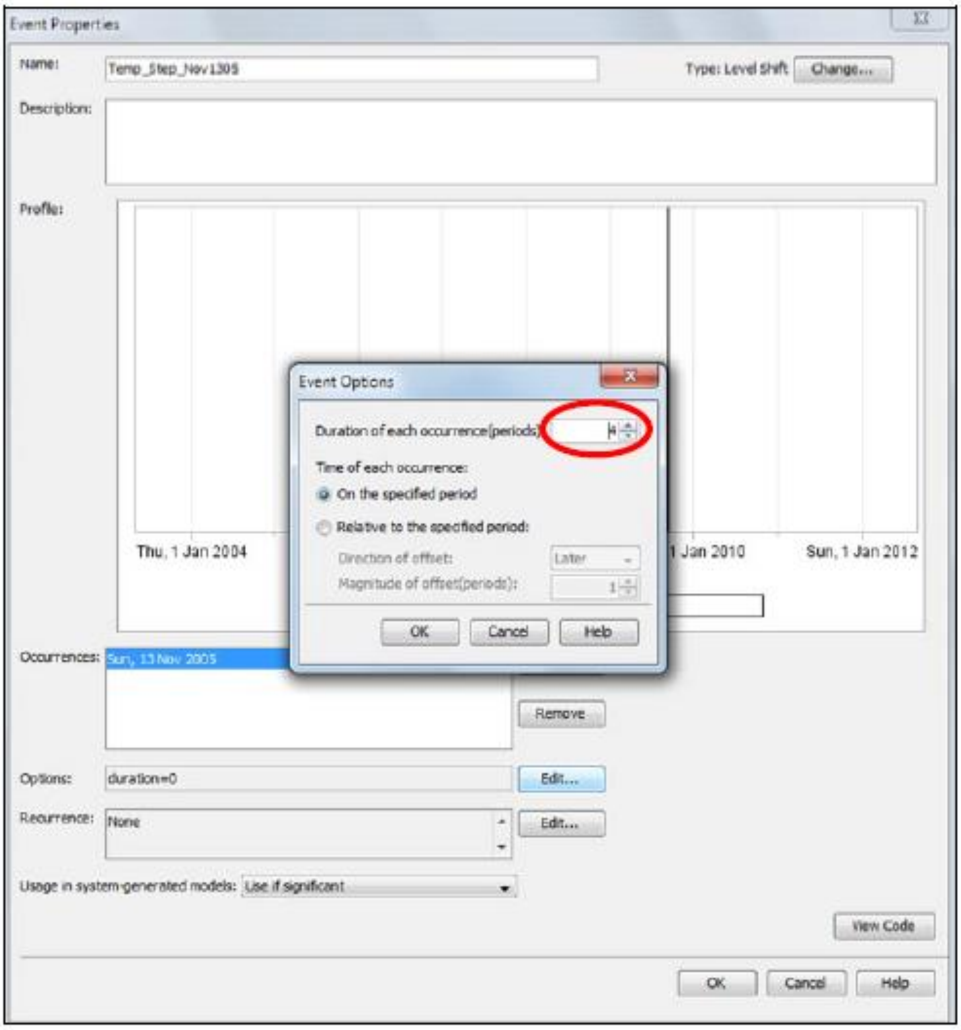
28. The aim here is to truncate the Shift event variable to better match the pattern of variation in the discount event.

29. Click the **Edit** button at the end of the **Options** field.



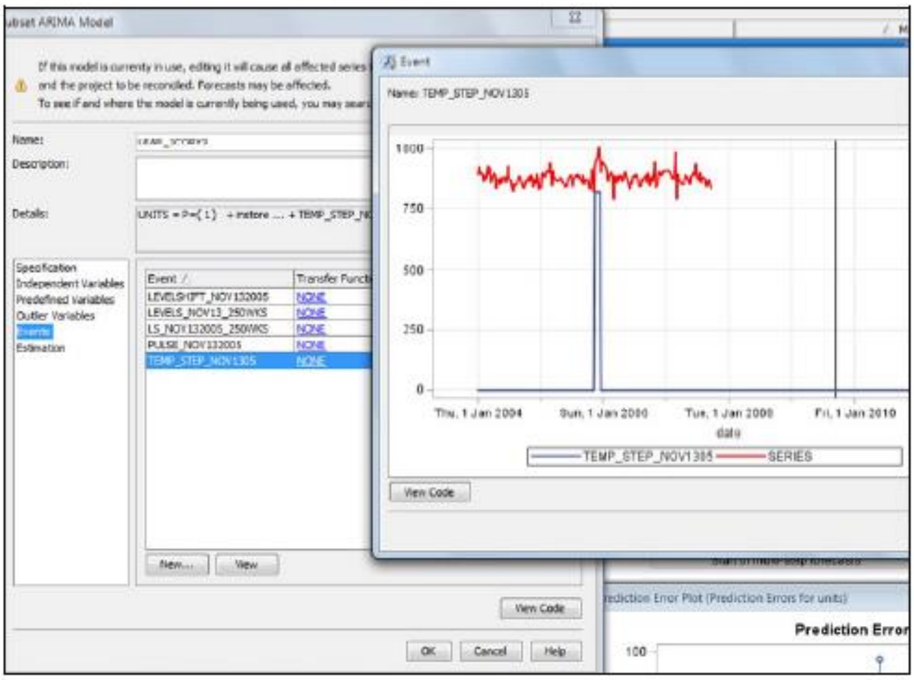
30. Change the Duration of each occurrence (periods) option from 0 to 4.

31. Click OK.



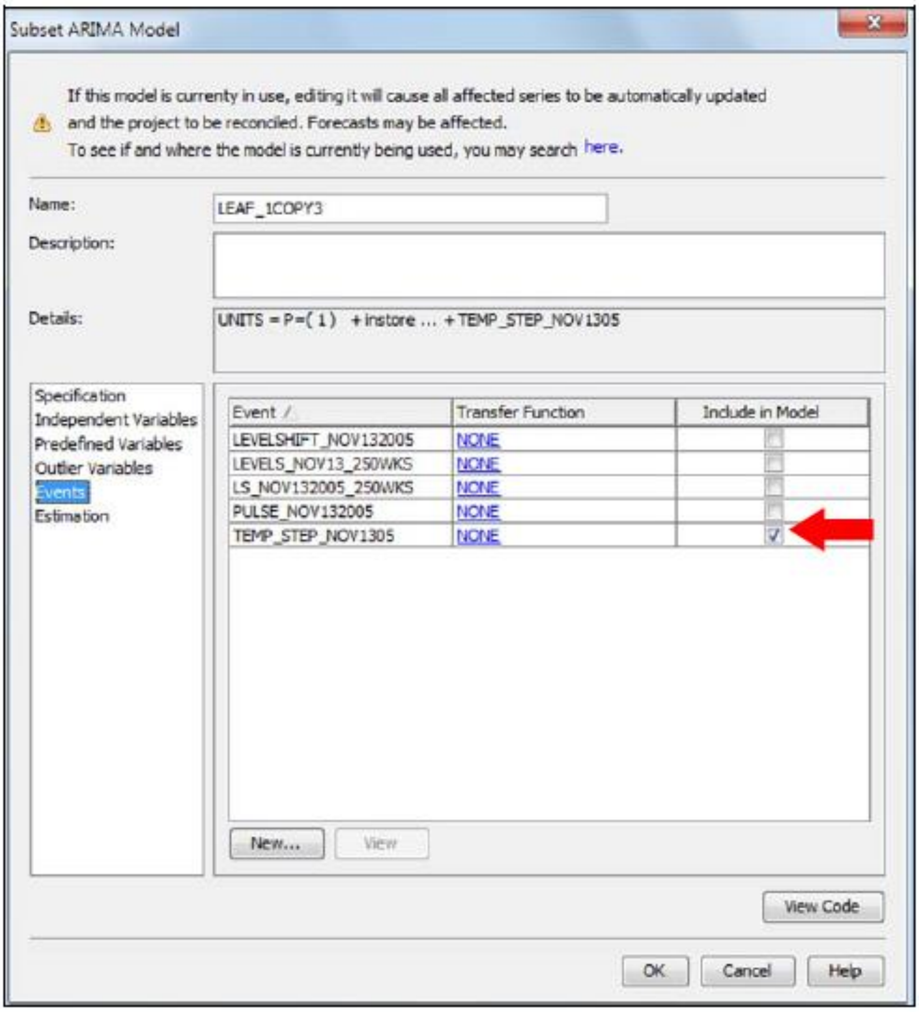
An abstract characterization of the event variable is shown.

32. Click **OK**.



33. Allow inclusion of only the truncated step event variable in the model.

34. Click **OK**.



The LEAF_1COPY1 model's fit is somewhat better with the *truncated-shift characterization* of the event.

The estimated scale seems more appropriate, and the estimated parameter is highly significant.