2018

Autobox 6.0 User's Guide Interactive Version

For Use by Only the Smartest Forecasters

How to get data in, How to run and How Understand output

AUTOMATIC FORECASTING SYSTEMS

Preface

About AFS Inc.

- AFS has been in business since 1976 and has many business and academic customers.
- AFS was given the award in the textbook "Principles of Forecasting" as the 'BEST DEDICATED FORECASTING PACKAGE".
- Autobox stands alone at the top of Automated Forecasting Software in the "Daily Data" 2008 International Society of Forecasters Forecasting Competition **(NN5)**
- We are dedicated to giving you a better class of service than other forecasting providers.

Technical support

Customers may contact technical Support for assistance in using Autobox, or even for installation help.
To reach Technical Support, see the AFS Inc. web site at *http://www.autobox.com*

Customer Service

If you have any questions concerning your shipment or account, contact your local office, listed on the Web site at *http://www.autobox.com*. Please have your serial number ready for identification.

Training Seminars

AFS is willing to provide training seminars, through monitor sharing, skype, or even onsite training. All seminars feature hands-on workshops. Seminars will be offered at an as-needed basis, so make sure to contact us if you need help.

Day 1

Introduction to Business Forecasting, Introduction to Time Series, Simple Averages, Moving Averages and Exponential Smoothing

Regression Models for Forecasting, Forecasting Accuracy, Putting it all Together – The Forecasting Process

Day 2

The Model Build Process, Daily Data modeled, Statistical Transformation, Verification of Assumptions, Model Augmentation (ie Stepup/Stepdown)

Day 3

Examples using Customer Data, Complete Walk through of Modeling Process, Verification of Model with Data, Examples using Autobox Data (if time permits)

Table of Contents

Part One: User's Guide

Chapter 6:

Chapter 7:

Part 2: Reference Guide

Chapter 1

Introduction

Thank you for selecting Autobox. If you are having any questions or issues call us at 215-675-0652.

In order to unlock the potential of Autobox, the program must be installed as an “administrator” on Vista/Windows 7 machines. Download Autobox and save it to disk. Right click on the downloaded and choose “run as administrator” to install. Then, right click on the icon to run as administrator whenever you go to run Autobox.

Autobox can be used with any “time series” data in 5 broad ways:

- **Data Cleansing** (i.e. Identify outliers and correct the data for its errors)
- **Modeling past behavior** (i.e. Did the Promotion Coupon work?)
- **Forecasting** (i.e. Extrapolate a series of numbers into the future)
- **Exception Reporting** (i.e. Which series are out of control? What time period has the most outliers across my different SKUs?)
- **What-If/Scenario Analysis** (i.e. What would happen if I lowered the price down to $xx?)

Autobox uses automatic modeling heuristics (not pick best) with intervention detection. It tailors the forecast model to the problem at hand including selecting the best lead and lag structures for each input series. It corrects for omitted variables (e.g., holidays or price changes that have affected the historical data that the system has no knowledge of) by identifying pulses, seasonal pulses, level shifts and local time trends, and then adding the needed structure through surrogate variables.

Autobox delivers not just a forecast, but a variety of information from equations to graphs that help present complex statistical information in a way that is easy and clear at every stage of the forecasting process and even a simple narrative description of the final model. Graphs of autocorrelation, partial-autocorrelation and cross-correlation functions are all available.

Included in the software package are over 700 example data sets taken from statistics textbooks that are available upon installation (look in the folders in the installation directory). This includes the classic “Airline Passenger series” which Box-Jenkins studied and is the most studied time series. It is located in the “Box-J” folder.

When modeling, you can include causal variables, retain future observations for error analysis, provide future values of the causal variables or tweak the modeling process that Autobox uses. Note that if you have causals you need to forecast using Autobox, you should do them by themselves and then post in the forecasts as future values. There are serious complications to bringing in interventions, etc. into the math behind the process that makes it too complicated and best done as the two step process as highlighted above.

Autobox will **automatically** aid the modeling process for weekly, daily, hourly and semi-hourly data. With weekly, daily, and hourly data, Autobox will add 51 dummy variables for the different weeks of the year. At least 1 1/2 year of historical data should be available for this to happen. For daily, hourly or semi-hourly data Autobox will add 6 dummy variables for days of the week. If it is hourly data Autobox will add 23 dummy variables for hour of the day. For semi-hourly data Autobox will add 47 dummy variables for each half-hour of the day. Autobox literally fills in the gaps to create a more complete forecast.

For daily data that covers all 7 days (Monday to Sunday), Autobox will use different modeling approaches. These approaches are entirely user selectable. You trigger Autobox to do this by providing a series name like this "__040106Y11". To tell Autobox to look for these daily effects, just add two "_" before the date and the name of the series where 040106 represents April 1, 2006 and the series(SKU) name is "Y11", for example.

- If a holiday lands on a weekend, Autobox will look for a "Friday before" and "Monday after" effect automatically.
- Search for a day of the month effect.
- Search for an "End of the Month" effect when the month ends on a Friday, Saturday and Sunday.
- Automatically add in U.S. holidays. Note that you can always create variables like this yourself and add them in as a causal variable.
- Monthly Fixed Effects for daily data given that there 78(52*1.5) weeks of data

If the prefix of the series is the date (ie __010107) AND then number of forecasts match to a month end then Autobox will report the probability of making a given month end number in a file named *.CAN. This will allow you to identify SKU's that are likely not to make the month end goal.

Autobox will look for Weeks of the Year when the series name does not have an underscore and date "__010108" and 78 weeks of data.

If you have a time series that is not annual, quarterly, daily or monthly, then Autobox will search for **interactions** between "fixed effects" automatically.

Autobox is capable of modeling unusual time series (ones that lack seasonality like monthly, quarterly or hourly). For example, if a business collected hourly data that only covers 20 hours out of the day, Autobox will still function. All that the user needs to do is to visit the installation directory that only covers 20 out of the 24 hours of the day, you can create a special file named '2season.afs" in the installation directory. Inside of the text file place a '20' in it so that Autobox can model it.

Our competitors will NOT let you pre-evaluate their software. They often will only let you see a slideshow. If they do let you "see" their software you will have to buy it with a 30 day refund policy. We would rather let you try before buying.

Note: If you have data that is very different in scale, we **strongly** recommend scaling your data(by dividing or multiplying) when you have small values and large values. If your Y is 10,000,000 and your causal is .075 then you should scale. You should keep a gap of 6 digits or smaller between the size of the variables (ie 1,000 in sales and causal variable .07 is ok). This is not a "quirk" of Autobox, but rather a common issue for everyone trying to estimate.

You should not have missing data points. If you have 6 years of monthly data, you should provide 72 observations. If you have a missing month or blank data, you should "zero fill" that period and post it. your data. If you have leading zeroes, Autobox will ignore the leading zeroes and begin the analysis at the first non-zero observation.

Contact us for any questions:

Afs Inc.
P.o. Box 563
Hatboro PA 19040
sales@autobox.com
Phone 215-675-0652 Fax 215-672-2534

The Menu System

File

New-Allows you create a new dataset by entering or pasting the data into a spreadsheet

Export-Save the current spreadsheet into Excel

Import-Import series from Excel worksheet.

Open-Allows you to directly import a dataset from an existing .TXT or .ASC file

Print-Allows you to print spreadsheets and graphs

Save-Save newly created dataset for permanent storage

Stock Retriever-Allows you to retrieve Stock Quotes to Forecast

Recent files-Allows you to open a recently used series

Exit-Exit Autobox

View

Show DataPoint Information-This gives the user the option to show/hide this information on graphs

Preferences

Allows you to change the number of decimals used during the process

Process

Run Autobox - Opens the "Model And Rules Wizard" that helps the user through the selection of execution options, model and rules required to process the data. Allows user to select automatic modeling, user specified model or special processing.

In **automatic modeling (OPTION 1)** the program will determine the model based on the data supplied; and the user can choose to process with default rules held in memory or provide their own.

In **user specified modeling (OPTION 2,3,4)** the user can select from a list of offered models with default rules held in memory or choose to provide a model with rules provided, or with specified default rules. The user can select to save changed rules, final model or reports, where applicable.

In **identification only (OPTION 5)** Autobox will provide the ACF and MPA tables.

RunWhatIf-Generates forecast data only based on user adjusted future values

Add

Generated Series-Allows user to select intervention series to be included

Level Intervention – Adds a variable with a series of 1's to account for a temporary change in the mean

Pulse Intervention – Adds a variable with one 1 to account for an outlier

Seasonal Pulse Intervention – Adds a variable with a 1 at regular intervals to account for a seasonal event

Trend Intervention – Adds a variable with 1,2,3,4, etc. to account for a trend in the data

Holidays-Allows user to select holiday series to be included in the dataset as dummy variables

User Defined-Allows user to include his specified type of series

Delete-Allows user to delete any input (Independent) series

Hide columns-Allows user to "hide" rather than delete an input series; and process the remaining dataset (you **can't** change series information in this mode)

Series Information – Allows you to change the data type (0,1,2,3) of a series

Select New Output Series- Allows user to select an input series to replace the original output series and process the dataset to determine effect

Restore Columns-restore any columns that were hidden

Restore Original Output Series – Reverses the Hide Columns process

Help

Contents

Search for Help on-Allows you use Index or Find options

User Guide-Open's this document

Autobox.com-Sends you to our website

About Professional-Sends you to website to see overview, model description and pricing

About Autobox -Tells you the version of the program

These Are the Tabs Seen In the Program

There are menu items which we have already reviewed and there are folder "tabs" that allow you to view the data, graphs, reports, etc.

Historical Data - Shows you the past of the series to be forecasted. You can right click on the spreadsheet which generates a pop-up menu with Copy, Paste, Hide Column, and Restore Columns options. The Copy and Paste options work just like a clipboard. The Hide Column and Restore Columns are unique features that allows the user to hide a column(s) and thus excluding it from the process without having to delete it. Any Input (independent) series can be hidden and restored; however, the Output (dependent) Series cannot. To hide a column(s) left click on the header(name)cell while holding down the "Ctrl" key; and, when you have finished selecting columns, right click on the spreadsheet and select the Hide Column option. To restore the columns, merely right click on the spreadsheet and select the Restore Columns option. Please note that if any of the hidden columns have related columns in the Future Values and Retained Data spreadsheets, they will automatically be hidden/restored.

Future Values-Shows the anticipated future values for the specified series. You can manually enter in or paste future values.

Retained Data-Shows the data retained from the end of the specified series. If you haven't already loaded retained data via an .ASC file, you need to reduce the number of observations in the series properties box to generate them. This is for testing "out of sample error" such as MAPE comparisons using different lead time periods.

Auxiliaries - Shows the Actual/Fit data, Actual/Forecast data and the Forecasts if generated for the series that was modeled.

Graph- As you open each graph in Autobox, it will then be generated and saved in the installation directory as a jpg file. Note that as the vertical bar passes over a data point the information is indicated. When an outlier is indicated, the text describes the type of outlier and the magnitude of its effect. We use a "P" for a pulse which is a one-time intervention which Autobox identifies and corrects for in the model itself (i.e. dummy variable). An "S" indicates a pulse that occurs every season. An "L" indicates that there has been a level change in the mean of the series a "SP" indicates a seasonal pulse.

Please note: We have added a new graph called "Adjustable Forecasts" which allows the user to "tweak" the forecasts for presentation purposes based on judgment outside the modeling process. Just left click on a data point and while holding down the mouse button move the point up or down as desired. When finished,

select the "Save Changes" button and the changes will be made to the Forecasts spreadsheet and the other graphs. You must click on each of the graphs to save the new graphs to a jpg file. Click the Restore Original Forecasts button and the Forecasts spreadsheet and graphs will be reverted. Again you would have to click on each graph to save them to a jpg file. If you select to use this graph and change the forecasts, you will be precluded from using the WhatIf function of the program.

Reports

Details.htm – Tracks the steps and decisions to create the model

Intrvent.htm – Lists all of the detailed information on interventions

Rhside.txt – Lists how each of the variables numerically contribute to the forecast.

Safety.txt - Lists out Safety Stock for 90% service level(put nosim.afs in the autobox folder to get this)

Adjusted.txt - Shows historical data free of outliers

Stat.htm – Shows the statistical fitting statistics (RMSE, AIC, etc.) and model with P-values

Verbal.txt – The model explained in "English"

Ab50pro.123 – The forecasts with upper and lower limits

Forecast.csv – The forecasts in Excel format

Equation.txt – Reporting only the equation

Equation.csv – The equation in Excel format

Fitted.csv – The fitted values in Excel format

*.**PMM** – When a file called "weights.afs" exists in the directory, a 4.1 PMML file is generated with the model built by Autobox ready to be imported into a post processor. See more at www.dmg.org Copy all pmm files to pmml after a run using the DOS copy command (ie copy *.pmm *.pmml)

*.**PMD** – When a file called "weights.afs" exists in the directory, a 4.1 PMML file is generated that is the dataset accompanying the pmml file ready to be imported into a post processor. See more at www.dmg.org

*.**TRA** – Ready to be imported to BI Tools like Tableau, Spotfire, etc. with name of series, actual, fit, forecasts, confidence limits, flag indicating if seasonality found.

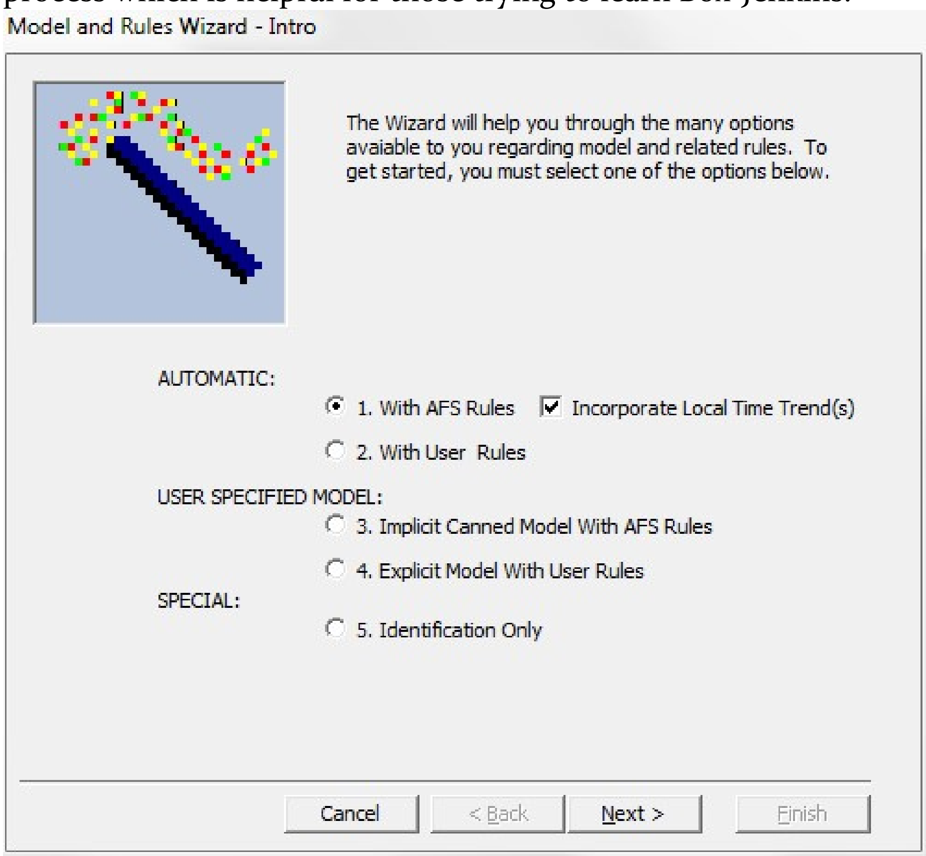
*.**OUT** – Reports all of the outliers sorted by size.

*.**PEK** – Shows the largest and and smallest values in the forecast. It's called "Peaks and Valleys" to show what the highs and the lows might be in the future.

Model And Rules Wizard

As we indicated above, select “Process\Run Autobox” and the wizard will appear as follows with the default set to option 1. This is because option ‘1’ uses our heuristics. Option ‘2’ is also automatic, but you can make many adjustments to the process.

Option ‘3’ lets you pick a typical model (i.e. holt winters, etc.) and run, but we use this more for more benchmarking than for analysis. Option ‘4’ allows you to specify your model for the dataset. This again could be a good comparison to what Autobox would report using option ‘1’. Option ‘5’ allows you to run the initial step of the modeling process which is helpful for those trying to learn Box-Jenkins.

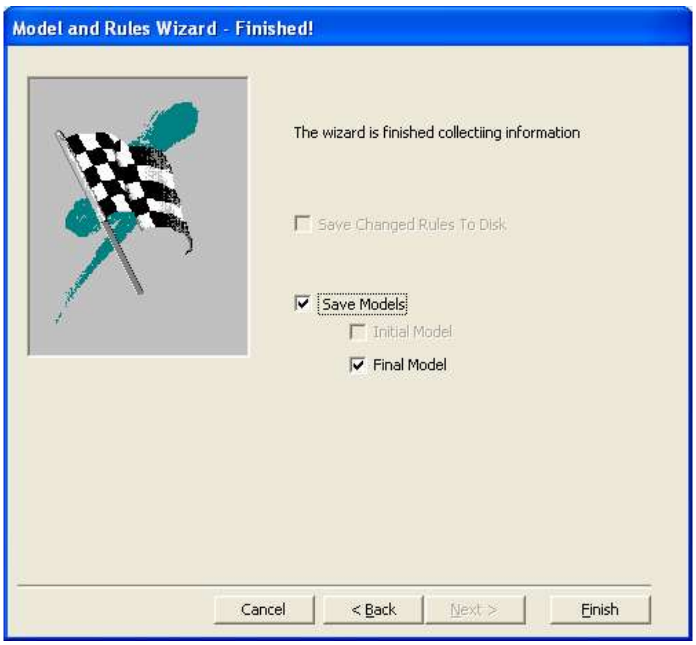


As the wizard indicates, your first need to choose what you want to do.

Option 1 – Run using proprietary AFS rules

Option 2 – Run using User's overrides of AFS rules
Option 3 – Choose a specific model like Holt Winters – it's only there for comparison in our opinion!
Option 4 – Choose your own model – this is for experts…be careful here
Option 5 – Runs only the identification process is good to run with as you can for yourself try and review the Autocorrelation Function and Partial Autocorrelation Function for lag relationships. If you select optional and press "Next", you will have the option to save the final model as indicated on the following:

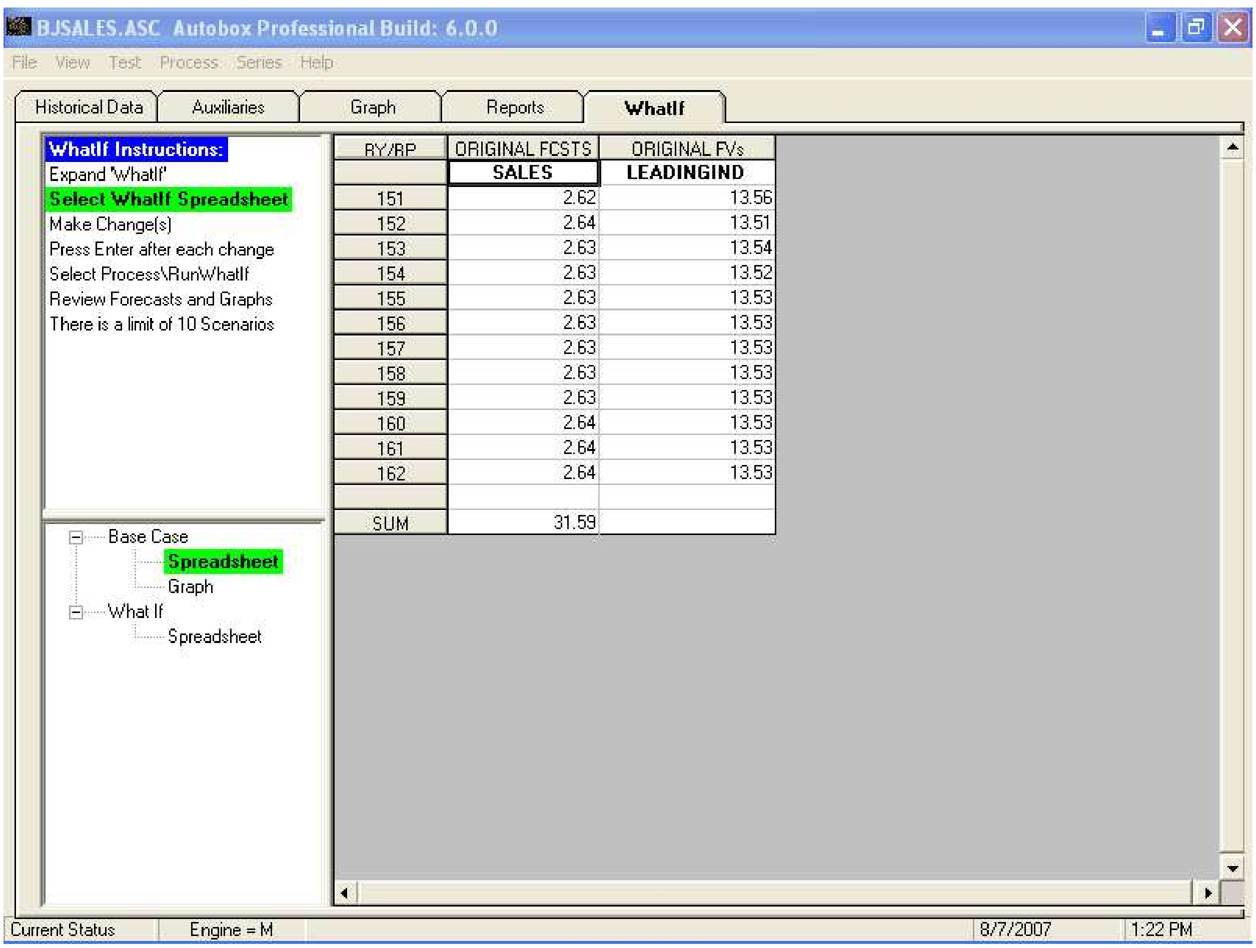
You can choose to save your rules and models for later use. No model will be saved when the Tsay test for changes in variance is used as well for any "demo" users.



Click "Finish" and your dataset will be processed automatically.

WhatIf Tab

This Tab appears for Causal Models only. It allows the user to change values to determine their effect. When the Tab is selected, the Base Case spreadsheet is shown which indicates the original forecasts generated and the original future values for each input series included in the model. The upper left pane includes the basic instructions and, as you proceed, the next step is highlighted to assist you. The lower left pane includes the options to be selected. The options expand as the process continues.



To begin the WhatIf process, you must select the WhatIf spreadsheet, click on the cell you wish to edit, make your change and press Enter. When you finished making desired changes, select Process/RunWhatIf from the main menu. When processing is completed, the Forecasts spreadsheet is shown indicating the original forecasts and the forecasts for the current scenario. A graph of the current WhatIf forecasts is also available.

At this time you have the option to select the WhatIf spreadsheet again and create another scenario or select Restore Values to revert to the original future values. Please note that there is a limit of 10 scenarios.

After each scenario is processed, the forecasts, graph data and specified values(changed values) are retained; and you can review the combined scenarios by expanding the Review All Scenarios option.

NOTE: If you have a series that has had a change in variance then no "What-if" option is displayed. It is not available in this release.

Chapter 2

A Quick Tour through Autobox

In this section, we will use sample data to familiarize you with the landmarks in our software. Click on "File/Open" and choose the file "Pizza.asc"

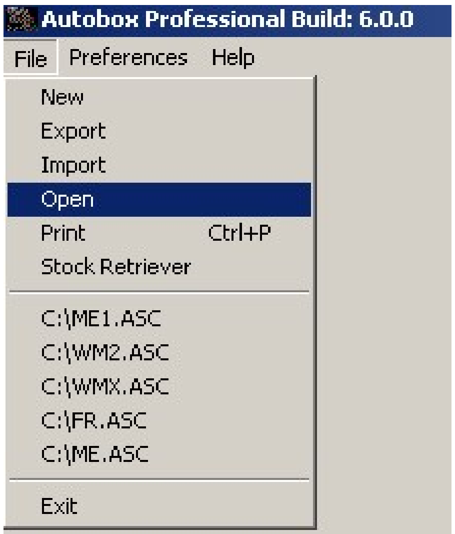


Fig 2-1 - Accessing Files

Double click on PIZZA.asc

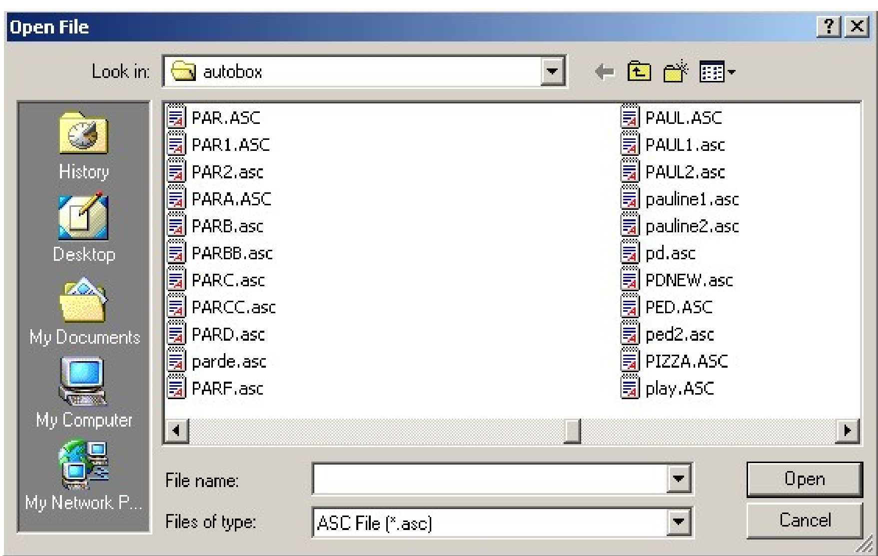


Fig 2-2 - File Selection

You can change the length of the forecast or the number of observations by editing the box on the right then clicking "Apply".

Autobox Toolbar

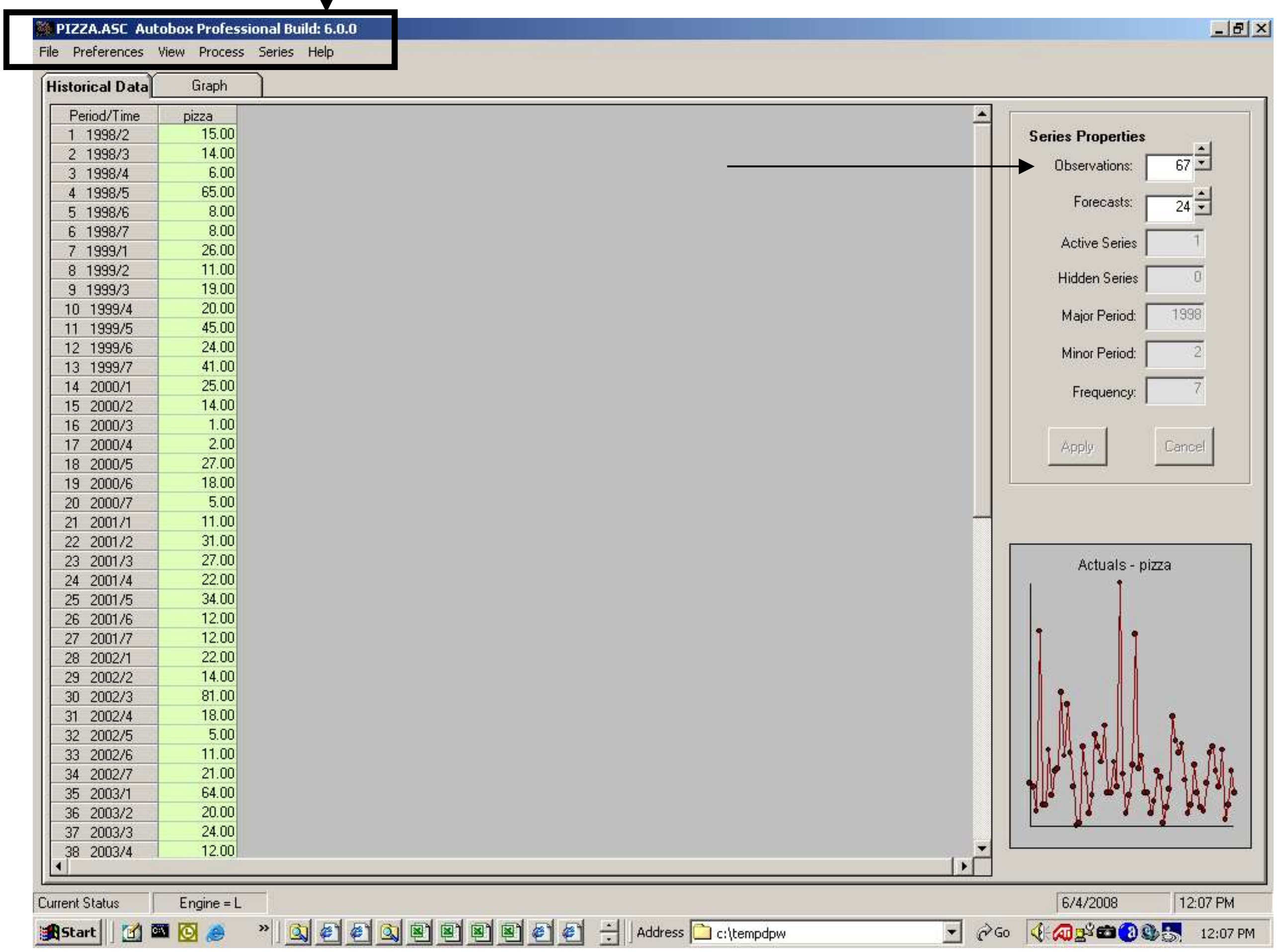


Fig 2-3 - Pizza Example

To run, click Process from the Autobox toolbar, then select “Run Autobox” from the dropdown menu. You will be prompted with the Model and Rules Wizard seen below. For this example, we are going to run using option “1”. Click “Next” to run using the “AFS Rules” or otherwise known as the “Expert system”. Option ‘1’ has **exclusive** conditions that are not disclosed.

We will walk through options ‘2’ through ‘6’ later in the “Models and Rules Wizard” section.

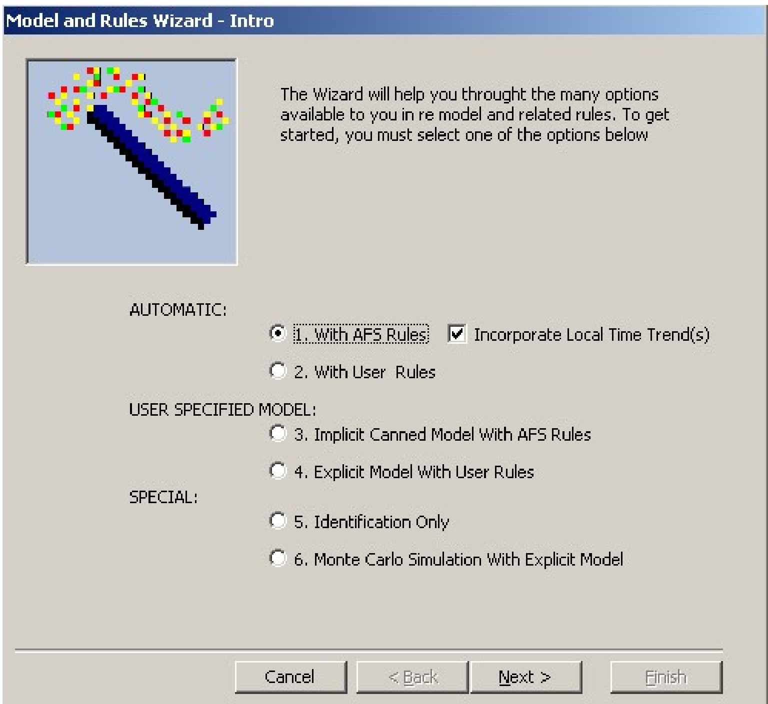


Fig 2-4 - Model and Rules Wizard

Click “Finish” to run. Notice that the option “saved changed rules to disk” is greyed out. This is because no adjustments were made. You can save the final model, but not the initial model as there were no initial model specified which is reserved for option ‘4’ where the user can specify the model.

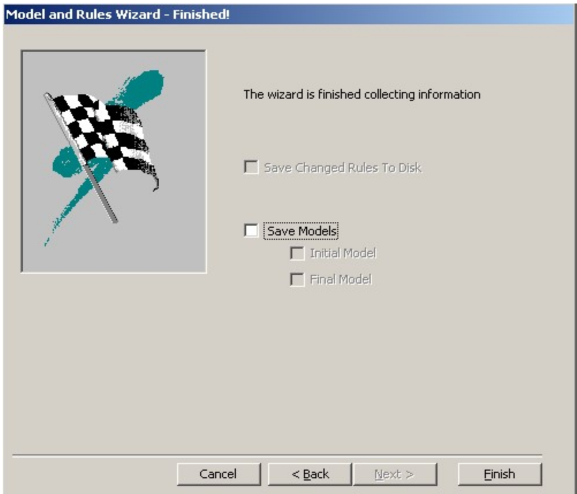


Fig. 2-5 - Rules have been set

Upon completion, the “Actuals and Forecasts” graph will automatically appear.   Note that when the vertical bar passes over a data point, information is displayed in the upper right-hand corner. When an outlier is indicated, the text box at the top right describes the type of outlier and magnitude of its effect.

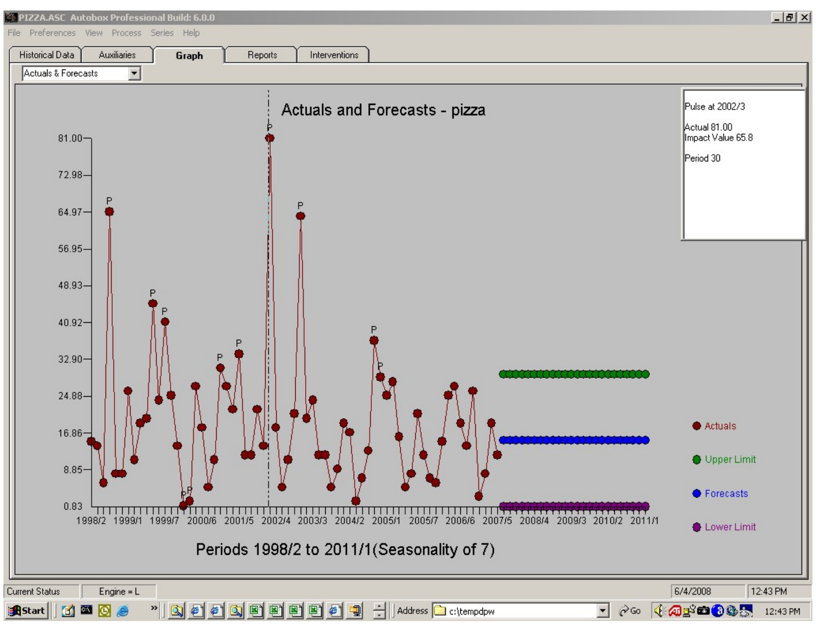


Fig. 2-5 - Actuals and Forecasts created

Click on “**REPORTS**” and you will see a list of reports to choose from on the left hand side of the screen. Note that all reports are written out to the Autobox installation directory so you can archive them. Graphs are generated in the installation directory only AFTER you have clicked on one to view them within Autobox.

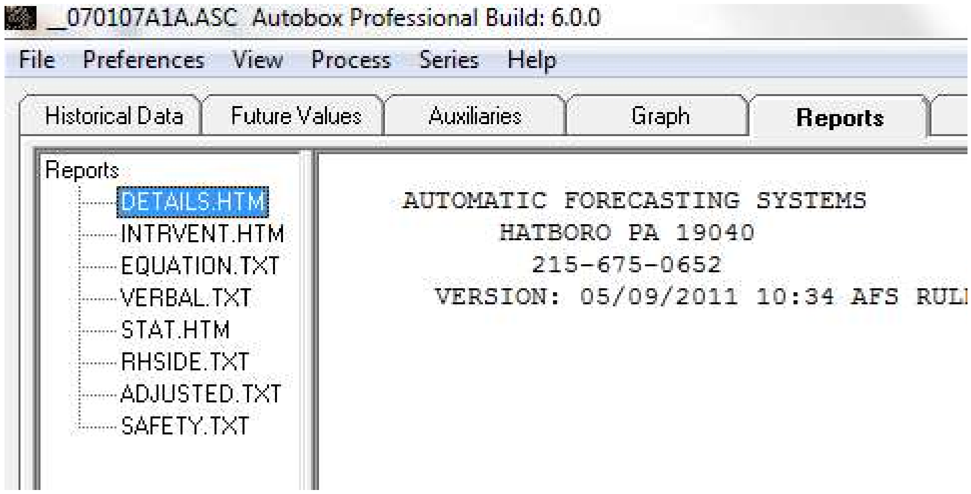


Fig. 2-6 - Reports are generated

Reports - There is a report generated that shows how the model was built and the forecast generated (DETAILS.HTM), a report showing(shown here)at what periods interventions were identified (INTRVENT.HTM), a report that shows the equation used to forecast(EQUATION.TXT), a report that provides of an executive summary of what was identified during the modeling process(VERBAL.HTM) plus breakdown of what each numerical part of the model contributed to the forecasts to get a full picture of the dynamics, a report showing the summary statistics (STAT.HTM), a report showing the equation as a pure regression equation(RHSIDE.TXT) which is suppressed when variance changes is detected as the "scale" is different (ie weighted least squares), a file cleansed of outliers (ADJUSTED.TXT) and a file with safety stock for a 90% service level(SAFETY.TXT). For example, if you click on "INTRVENT.HTM" you will see a list of all the interventions with the type of intervention (i.e. P – Pulse, S – Seasonal Pulse, L – Level), the date they took place and their magnitude.

Note: These reports exist in the installation directory if you want to open them in Excel, etc. or archive them for reference.

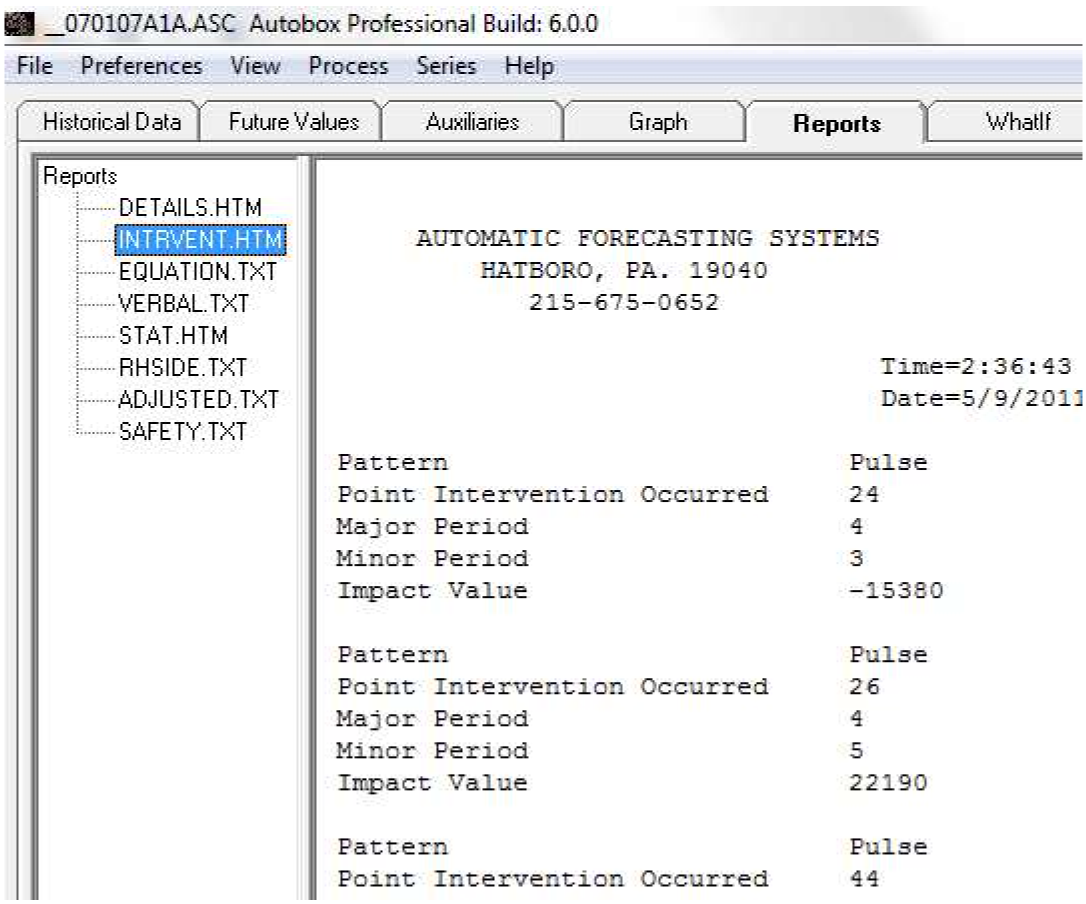
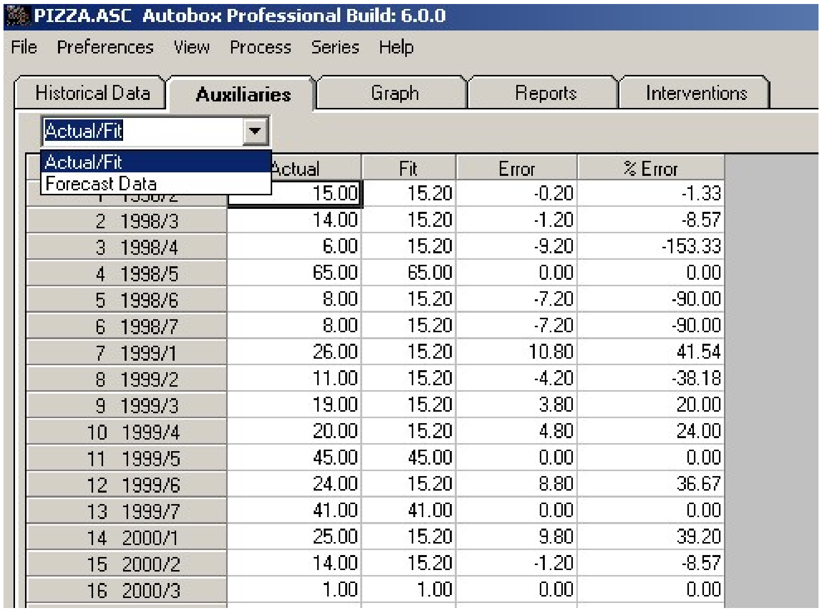


Fig. 2-7a - Browsing through reports to see interventions

Auxiliaries - If you choose the “Auxiliaries” TAB you will have the option of viewing the forecast, actuals and fit in a spreadsheets. Here we show the forecast and fit.



		Actual	Fit	Error	% Error
1	[illegible]	15.00	15.20	-0.20	-1.33
2	1998/3	14.00	15.20	-1.20	-8.57
3	1998/4	6.00	15.20	-9.20	-153.33
4	1998/5	65.00	65.00	0.00	0.00
5	1998/6	8.00	15.20	-7.20	-90.00
6	1998/7	8.00	15.20	-7.20	-90.00
7	1999/1	26.00	15.20	10.80	41.54
8	1999/2	11.00	15.20	-4.20	-38.18
9	1999/3	19.00	15.20	3.80	20.00
10	1999/4	20.00	15.20	4.80	24.00
11	1999/5	45.00	45.00	0.00	0.00
12	1999/6	24.00	15.20	8.80	36.67
13	1999/7	41.00	41.00	0.00	0.00
14	2000/1	25.00	15.20	9.80	39.20
15	2000/2	14.00	15.20	-1.20	-8.57
16	2000/3	1.00	1.00	0.00	0.00

Fig. 2-7b - Auxiliaries Tab

Graphs - You can click on the tab "Graphs" and see a variety of graphs. When you click to view a graph you will have a wide selection of the type of graph you would like to see: Actuals/Forecast, Fitted/Forecast, Actuals/Fit/Forecast, Actuals/Cleansed, Residuals, Actuals/Residuals, Forecasts, Actuals, ACF Originals, ACF Residuals

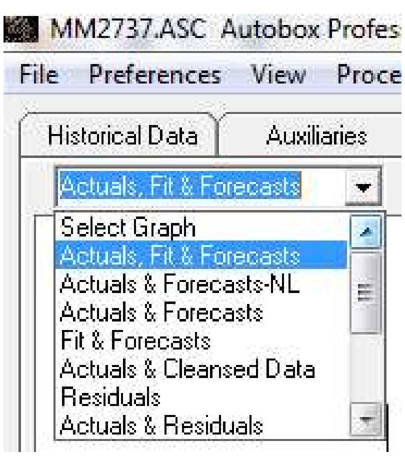


Fig. 2-7c

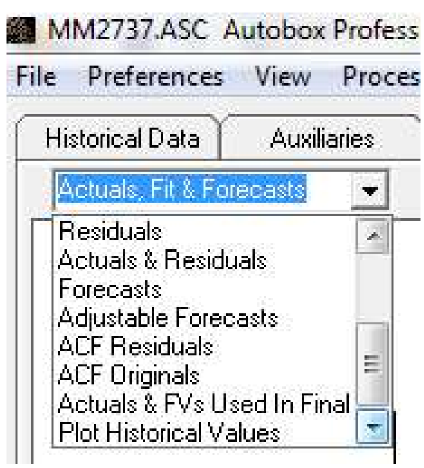


Fig. 2-7d

If you decide that you want to "cut and paste" in your data then you would click on "File/New" from the menu.

Fig. 2-8 – File Dropdown

You then need to fill in the properties of the series. When the spreadsheet appears, you can enter your data by cutting and pasting from Excel or entering it by hand, and then click "File/Save" to store the file for the future.

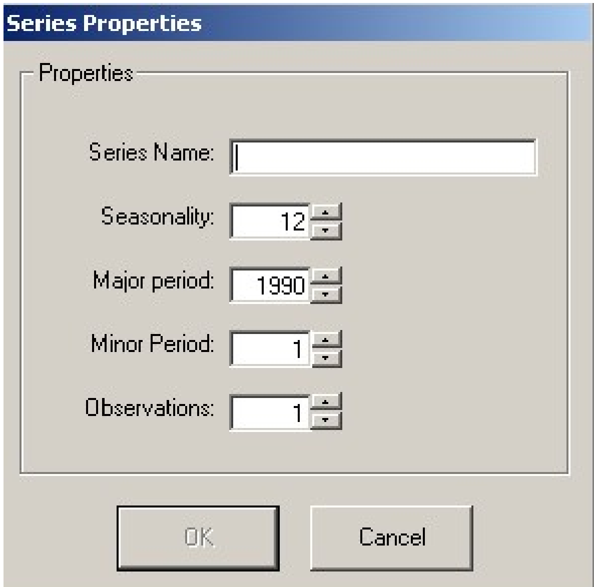


Fig. 2-9 - Properties of a New Series

Here are some examples on the type of **seasonality** you could enter – Choose 1 for annual, 4 for quarterly, 12 for monthly, 52 for weekly, 7 for daily(7 days in a week), 5 for daily(5 days in a week) and 24 for hourly.

Here are some examples on the **Major/Minor Period** –For a monthly data problem let's assume the starting month is February and the year is 2006. The Minor would be a '2' and the Major would be '2006'. For a quarterly data problem let's assume the starting quarter is 3 and the year 1974. The Minor would be a '3' and the Major '1974'. For a daily data problem let's assume the starting week 35 and day 5. The Minor would be '5' and the major '35' . For a weekly problem, let's assume the starting week was the 34th of the year in year 1992. The minor would be '34' and the major '1992'.

The Major and Minor are for helpful for reporting purposes only and not for the actual analysis. When we run daily data on test sets, we just put in a '1' and a '1' for each because it takes too much time to figure out the actual week and day so we don't bother with it.

Chapter 3

The Three Ways of Getting Data into Autobox

1)Import series from Excel

Select File/Import. The data ***MUST*** start in either row 1 or row 2 (row 2 if there is a header). You will be prompted for the series header information (i.e. what is the seasonality?)

2)Cut and paste into the data sheet or Data entry.

Select File\New. You will be prompted to enter the series name and properties. You can then enter the data by hand or cut and paste into the Historical Data spreadsheet.

3) Create your own Autobox ASC file.

You need to create a txt file with an *.ASC extension. You store the data and series header information in a specified format in a column. Select File\Open and the ASC file with the data. The properties and data will be automatically inserted into the proper spreadsheet(s). We have many ASC files that were created in the installation directory which can be used as reference.

Importing a Series from Excel

Our Excel import process only needs the data and any column headings. You will specify the date when do the import. ***<u>You must have the data beginning in row 1 or 2!</u>***

Let's say you have an XLS file with no causal variables and just the history to analyze. You will see that there are 3 time series and the third is missing values at the beginning. Autobox will ignore leading when running the analysis zeroes so this is 'ok'. However, in a causal you shouldn't have zeroes for one causal and the others populated as it will negatively impact your analysis!

Choose File/Import in Autobox and select the XLS file and hit "Next" five times.

	A	B	C
1	a	b	c
2	422	119	
3	182	49	
4	622	231	
5	54	16	
6	843	390	
7	242	36	
8	142	45	
9	629	183	
10	1615	415	
11	502	140	
12	1403	380	
13	868	227	
14	365	102	
15	650	145	
16	702	127	
17	446	108	5.418966
18	1836	383	19.20762
19	1246	257	12.88221
20	1450	263	13.17635
21	2953	407	20.38057
22	1161	149	7.457457
23	2924	434	21.71086
24	1729	169	8.45
25	1888	275	13.74313
26	919	178	8.891109
27	865	142	7.089366
28	1543	131	6.536926
29	2887	336	16.7581
30	1484	108	5.383848
31	1592	75	3.736921
32	1920	53	2.639442
33	1477	40	1.99104

Fig. 3-0 – Historical Data in Excel

In case you were importing a causal problem, we can specify what is the output (or Y) variable. We can select any field for the first box. Make sure to specify the **Seasonality** and **number of Forecasts**. **Major/Minor** can be left at the default as it is just used for some reports and not used at all when the model is generated. Click "Next" and then "Finish". A window will ask you were you should save the file. Specify the file and click "Save". This file will be "**causal ready**" and there will be 3 other files created assuming you only wanted to run them alone (ie Univariate). Open one up to see what it looks like using "notepad" for example.

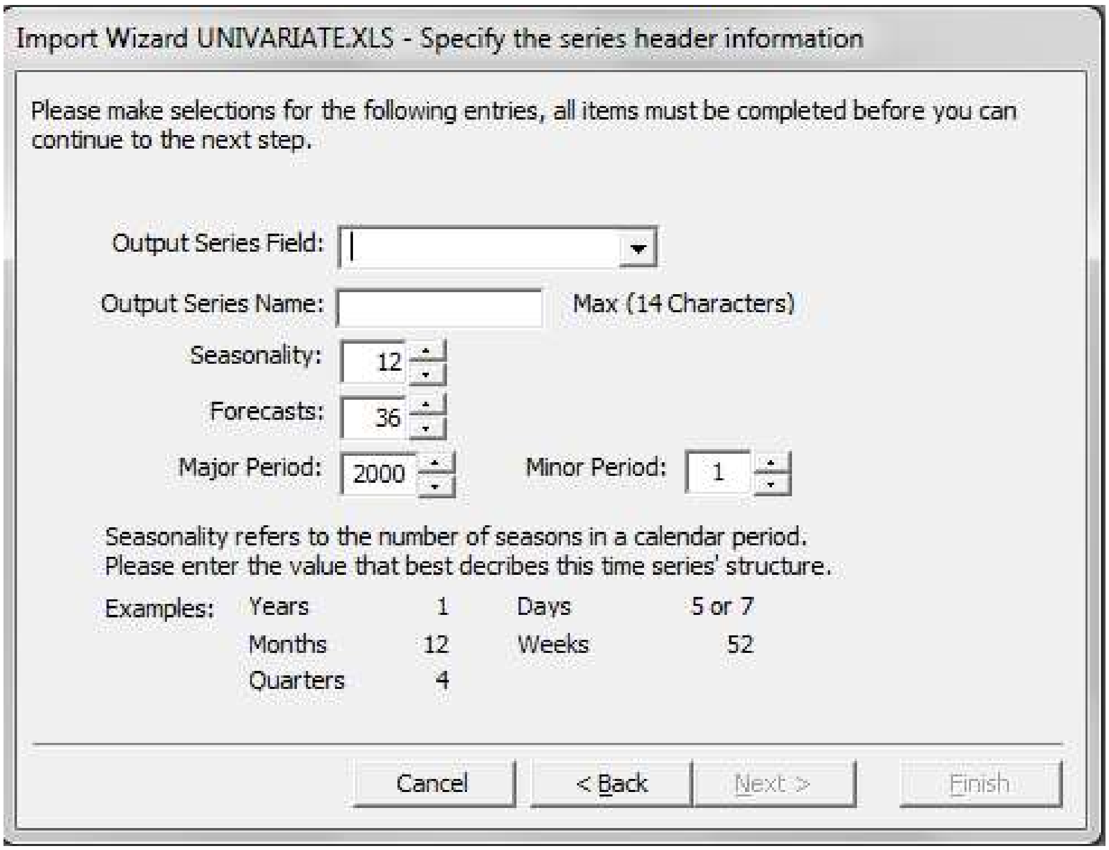


Fig. 3-1– Defining the Output, Properties

The following explains the individual properties

Output Series Field	Use this field to select the series you would like to forecast in this case it would be QTY
Output Series Name	This is just a name for the series that can be any alphanumeric characters we will use SKU1000
Seasonality	How often the observations are measured(ie 12)
Forecasts	The number of forecasts you would like Autobox to generate
Major Period	In this case it would be the year 2006
Minor Period	Enter the period the first observation had taken place, this number can 't exceed the seasonality

Table 3-2 – Properties of Data

Known issues and limitations:

- The spreadsheet must be closed in order for the import to work correctly. This means it cannot be open in Excel while you are trying to import.
- Horizontal data cannot be imported at this time, this means series that are listed in rows 1 – X and observations listed in columns A – Z.... will not import.
- Series names are limited to 14 characters.
- XLSX files will not import

Let’s try importing a causal problem now.
It is important to note that the causal variables will have a **data type** (a definition for Autobox on how the future values are created and used– there is more on this subject on page 40). The important thing to note is that if the series has no future values then the data type is defaulted to ‘0’.
If the causal variable does have future values then the data type is defaulted to ‘2’.
You can change these by clicking on “Series/Series Information” in the toolbar. Note that if you have a promotion with daily data that goes on for many periods you would want to use a data type ‘1’; as you can’t really look for lead/lag from multiple time periods.

demoBook.xls

	A	B	C	D
1	SKU	Qty	Adv	Date
2	1000	95623	0	3/31/2006
3	1000	83721	0	6/30/2006
4	1000	104953	1	9/31/2006
5	1000	123721	0	12/31/2006
6	1001	95623	0	3/31/2006
7	1001	83721	0	6/30/2006
8	1001	104953	1	9/31/2006
9	1001	123721	0	12/31/2006
10	1002	95623	0	3/31/2006
11	1002	83721	0	6/30/2006
12	1002	104953	1	9/31/2006
13	1002	123721	0	12/31/2006
14	1004	95623	0	3/31/2006
15	1004	83721	0	6/30/2006
16	1004	104953	1	9/31/2006
17	1004	123721	0	12/31/2006

Fig. 3-3 – Data with Causal Variables, Ready for Autobox

Shown in **figure 2.3** is the Import Wizard. Notice how each column has a heading identifying what is contained in the column and that each row contains an observation extending through time. When taking a close look at the first column, notice that the same number is repeated several times indicating observations for a particular SKU and then repeating with different SKU numbers. When importing a spreadsheet like this many SKU numbers as one series won't make much sense to Autobox, so we need some way to filter out just the items that we need. This can be accomplished by "filtering" the data.

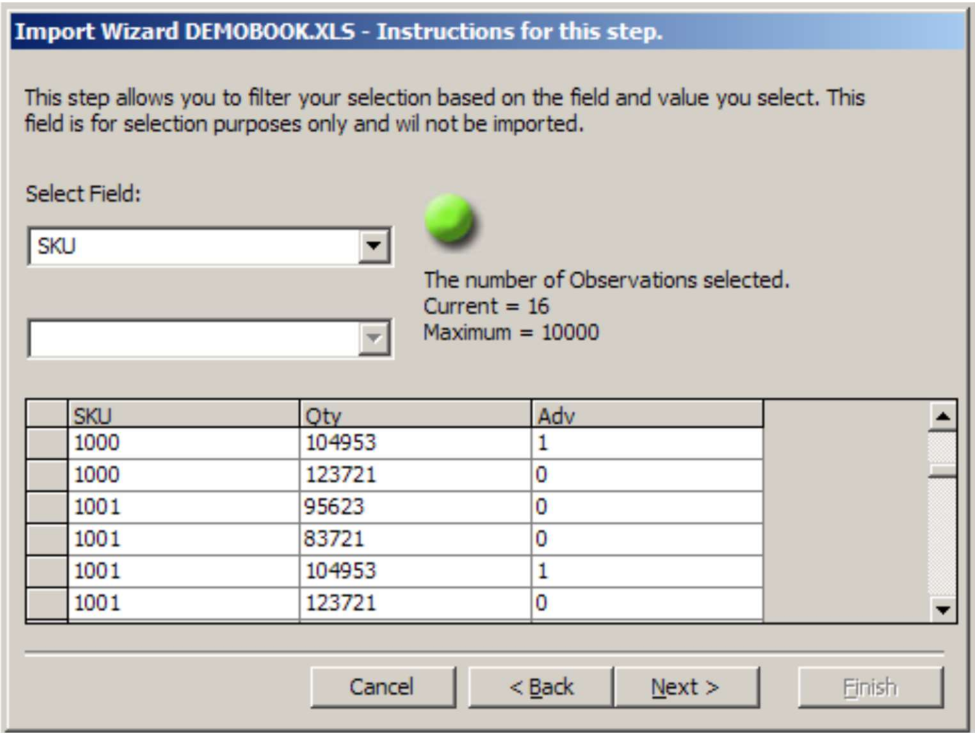


Fig. 3-4 – Import Wizard

To filter, select the field pull down menu to choose the field you would like use when filtering your data. Please note that only fields in the current data set will be available for selection. For this example we will be using the SKU field to filter our data. After selecting the SKU field, you can then select the value you would like to use as the filter. This is displayed in **Fig. 2.4** below.

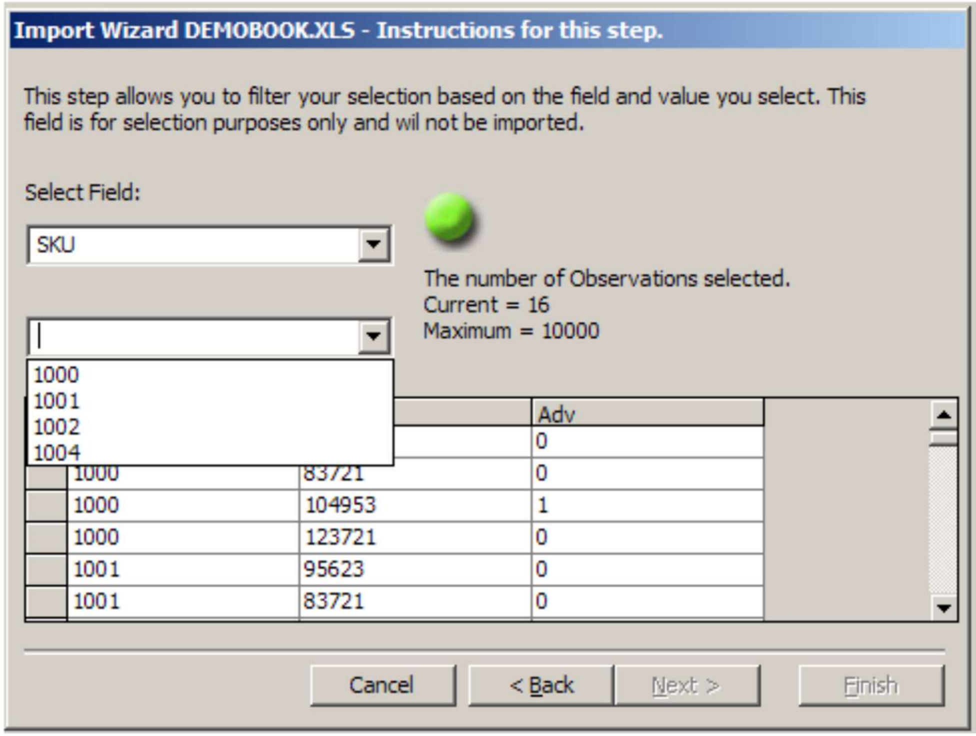


Fig. 3-5 – Data Filtering

When the pull down for the value is clicked, only the unique values that pertain to the SKU field are available for selection. By choosing the value 1000 we have selected the 1000 SKU as our filter; and only data that pertain to that SKU is available for import. An example of the data is displayed in **Fig. 2.5**.

NOTE: Please keep in mind when using the filter function that the field that you use to select your data will not be imported. The reason for this is that part numbers or SKU's are not used in modeling----just the data is needed.

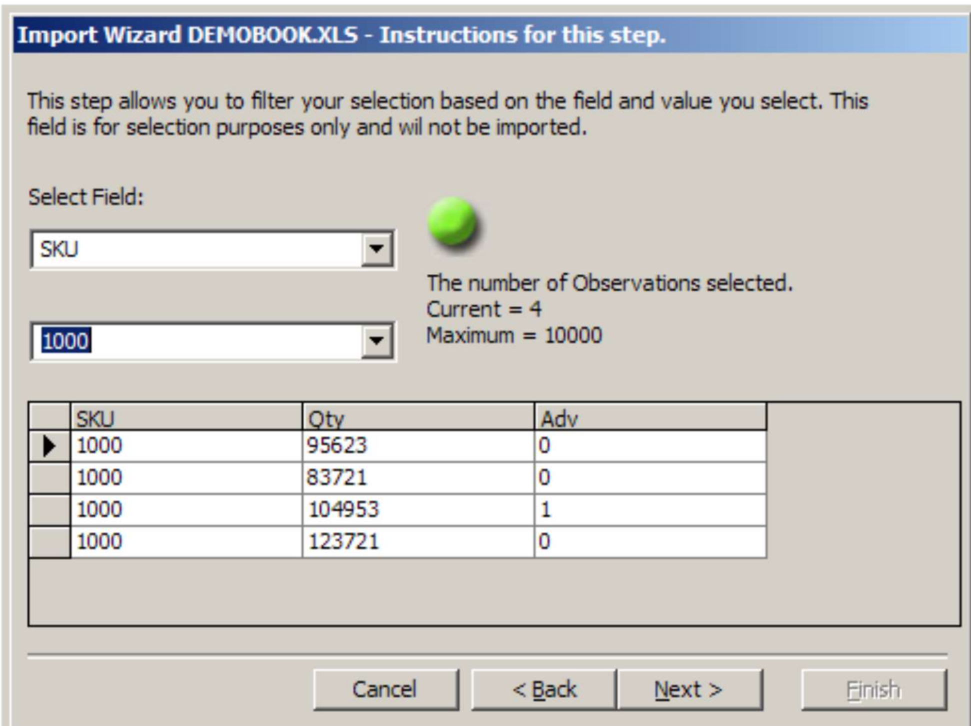


Fig. 3-6 – Data Filtering

The next step is to setup the properties for the Series you are importing. **Fig. 2.6** shows the Series properties step of the import wizard. If your data is not a causal problem, just pick any field as the output series field. Autobox will save the ASC file(s) into the folder where it imported the XLS file from. Autobox will also automatically generate an ASC file with a column for causal and an ASC file for each of the columns in anticipation of the data NOT being a causal dataset.

Note: International Users might need to change their Regional Options to "English" in Control Panel as some locations use comma's (,) instead of periods (.) which presents a problem for Autobox when decimals exist.

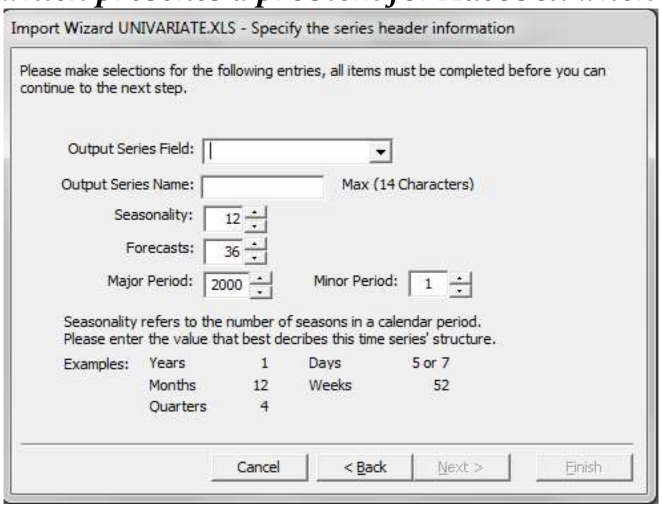


Fig. 3-7 International Users

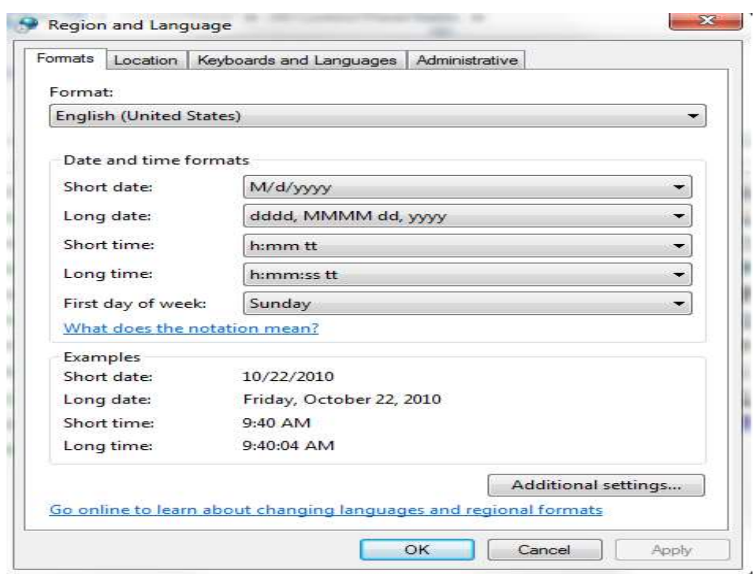


Fig. 3-8 - Language Selection

Cut and Paste or "Data Enter" to get Data into Autobox

For these examples, we provide you with the Excel file and a tutorial on how to take that data and get it into Autobox in order to get started.

Daily Sales

To access a **daily data** example, import a file named "daily.xls", which is installed when Autobox is installed. Here is a partial view of the data.

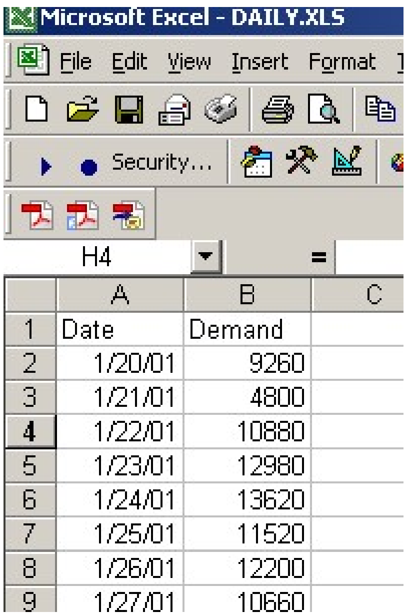


Fig. 3-9 – Data in Excel

Now, we enter Autobox and choose "File/New" and then add the series properties. We choose 7 for the seasonality as the data is daily observations. The major period of '3' is week 3 of the year and the minor period of '7' is the 7th day of that week.

NOTE: If you have the date and the series name embedded into the Series Name box, i.e. __010111Sales then you can just use 1 for major and 1 for minor as the date will automatically be reported.

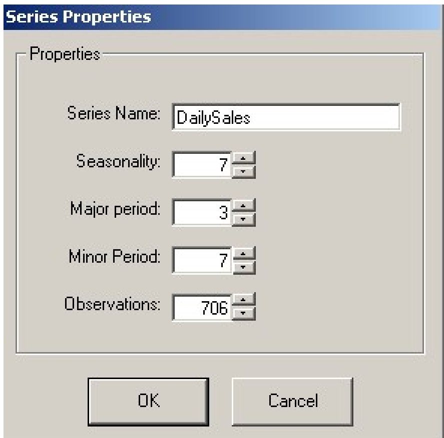


Fig. 3-10- Series Properties, Daily Data

After you click "OK", you can go into Excel and highlight your data (only the data, no need to highlight any text or dates!), copy it (CTRL-C or Edit/Paste) and paste it into Autobox's Historical Data spreadsheet.

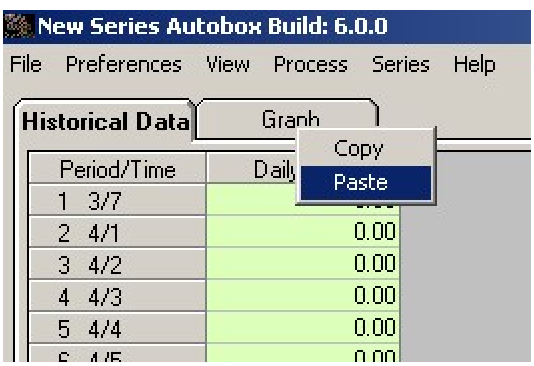


Fig. 3-11 – Pasting into Historical Data

Chapter 4

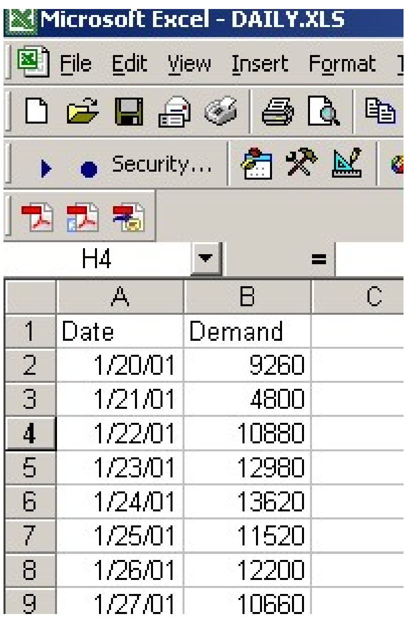
Some Examples to Show You How to Get Started

We will **repeat** some of the steps reading in the data **on the next two pages** as we did before and then introduce some new concepts.

For these examples, we provide you with the Excel file and a tutorial on how to take that data and get it into Autobox in order to get started. Again, a reminder that you must have the data in row 1 or row 2 in the Excel file.

Daily Sales

This example is daily data and our data is in a file named “daily.xls” which is installed when Autobox is installed. Here is a partial view of the data.



	A	B	C
1	Date	Demand	
2	1/20/01	9260	
3	1/21/01	4800	
4	1/22/01	10880	
5	1/23/01	12980	
6	1/24/01	13620	
7	1/25/01	11520	
8	1/26/01	12200	
9	1/27/01	10660	

Fig. 4 – Daily Data

We enter Autobox and choose “File/New” and then add the series properties. We choose 7 as the seasonality as the data is daily observations. The major period of ‘3’ is week 3 and minor period of ‘7’ is the 7th day of the week.

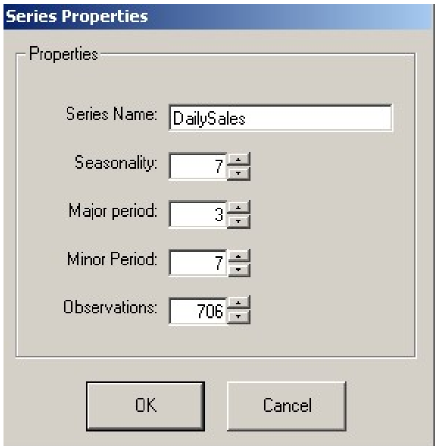


Fig. 4-1 – Daily Properties

After you click “OK”, you can go into Excel and highlight your data (only the data, no need to highlight any text or dates!), copy it (CTRL-V or Edit/Paste) and paste it into Autobox’s Historical Data spreadsheet.

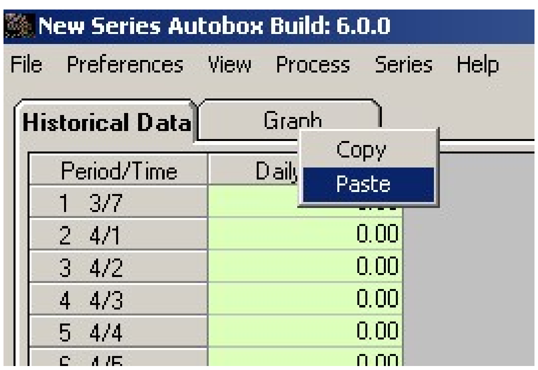


Fig. 4-2 – Pasting Historical Data

Change the number of forecasts to 365 as seen in the series properties on the right hand side of the screen and the click "APPLY". Note that the period & frequency can't be changed at all.

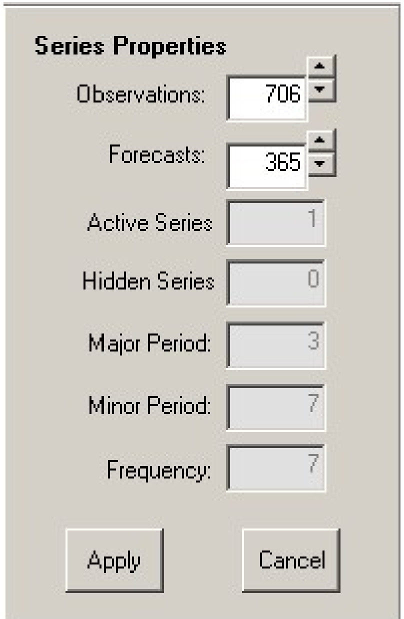


Fig. 4-3 – Setting Daily Properties

Choose Series/Add/Holidays.

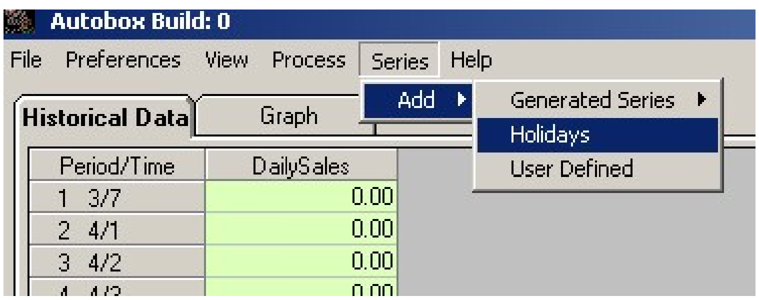


Fig. 4-4 – Adding a Holiday

Change the date to the start of the series 01/20/2001 and "then choose "Select All" and then click "OK'. This will let Autobox evaluate all of the different holidays in the year for any unusual behavior and account for it in the model/forecast. In addition, Autobox has created 365 future observations of these holiday variables so if they are significant they will be used to forecast the series.

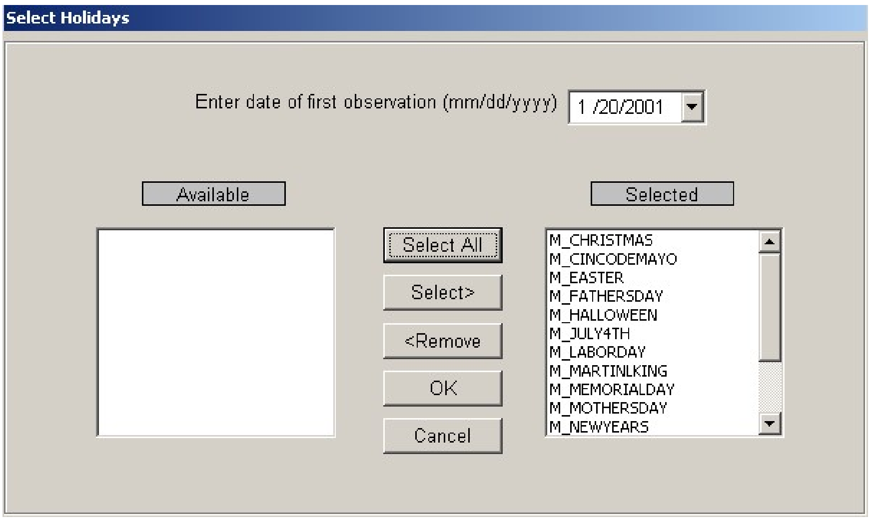


Fig. 4-5 – Selecting Holidays for Forecasting

The output series was renamed to include the first date and various holiday variables have been generated. This can be seen below in the column headers.

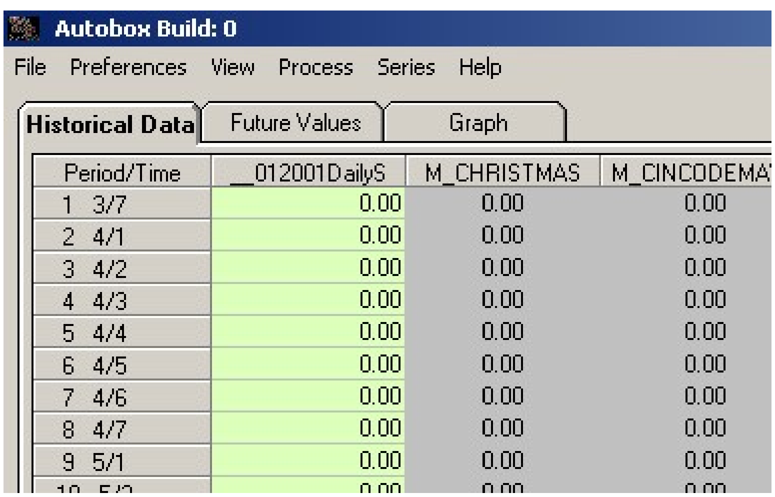


Fig. 4-6 – Data Ready to be Processed

You can then choose “Process/Run” and get your results. You can save your example by choosing “File/Save” so that you will have your example saved.

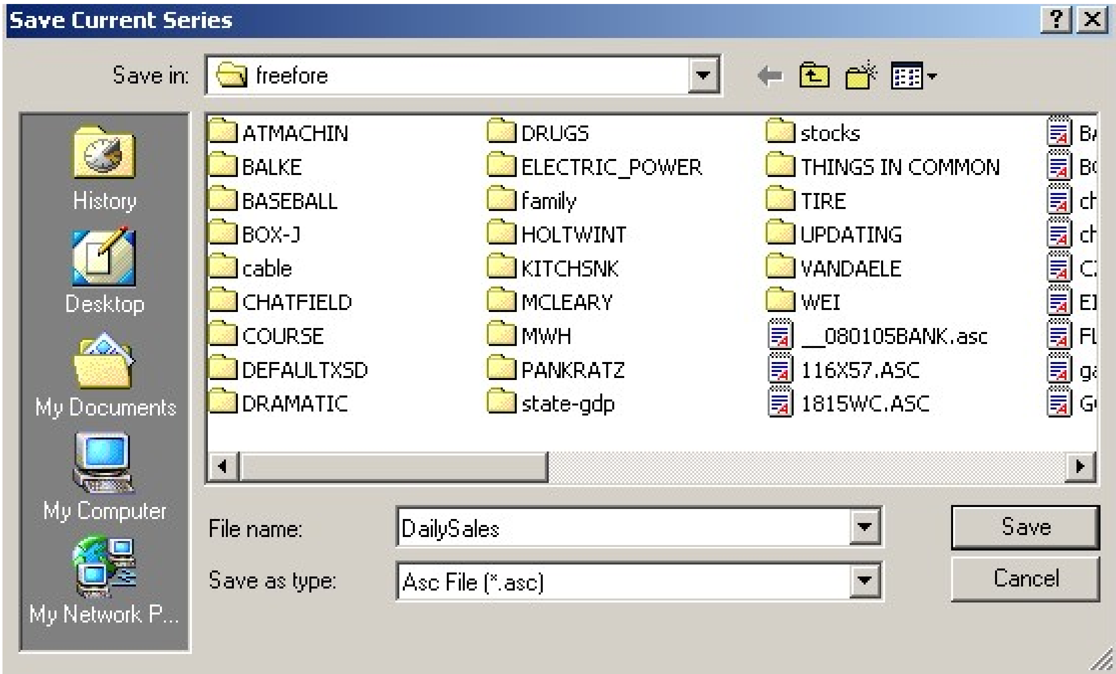


Fig. 4-7 – Saving the Example

Univariate example

We have one dataset that we are looking to forecast. The series, which contains sales in Bolivia, starts on 1/1/1991, and has a seasonality of 4. We have 44 observations that we are looking to forecast 4 periods out. Our data is in Excel file named Ex2.xls.

	A	B	C
1	Date	Sales	
2	1/1/1991	15602	
3	4/1/1991	17115	
4	7/1/1991	18795	
5	10/1/1991	19002	
6	1/1/1992	17851	
7	4/1/1992	18296	
8	7/1/1992	19980	
9	10/1/1992	20097	
10	1/1/1993	19381	
11	4/1/1993	19902	
12	7/1/1993	21030	
13	10/1/1993	22590	

Fig. 4-8 – Ex2.xls

We enter Autobox and choose “File/New” and then add the series name and properties.

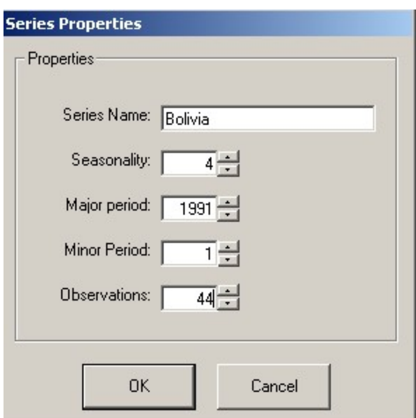


Fig. 4-9 – Univariate Properties

After you click “OK”, you can go into Excel and highlight your data and paste it into Autobox. Then change the number of forecasts to 4. Choose “Process/Run” to get your results. You can save your example by choosing “File/Save”.

Lead Effect example

There is going to be a promotion on 9/1996 and 7/1997. We would like to account for that in the modeling process. This example will detect an effect BEFORE the promotion as people may not end up buying the product before the sale and buy more when the sale is in effect. Autobox will look up to 4 periods before an event for a detectable trend. This function can be overridden using engine.afs (search for engine.afs in the manual for more on how to edit engine.afs). Also, creating causal variables with a data type of '3' within 4 periods of one another should be avoided, as this can conflict the search for lead/lag. Autobox will search for these instances and correct for this situation.

We enter Autobox and choose "File/New" and then add these series properties.

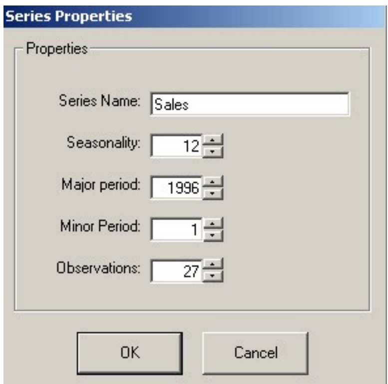


Fig. 4-10 – Lead Properties

Open the Excel file named "Ex1.xls" and cut and paste the data into Autobox. Then change the number of forecasts to 36.

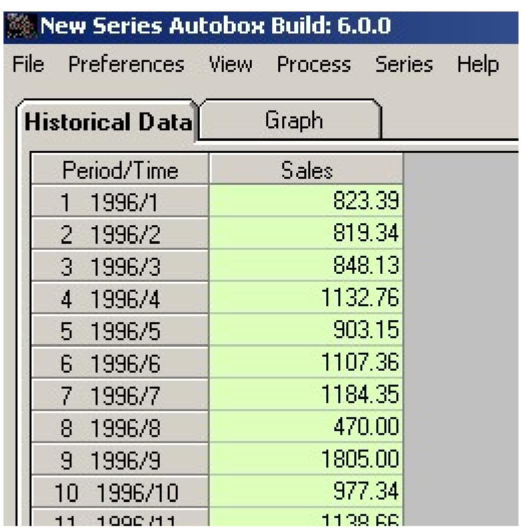


Fig. 4-11 – Ex1.xls

Let's add the promotion variable by choosing "Series/Add/User Defined"

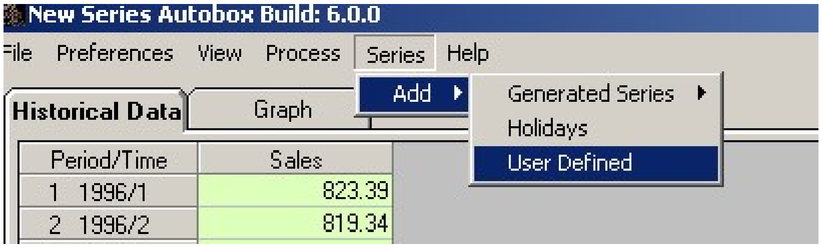


Fig. 4-12– Adding a User Defined Promotion

Enter "Promo" for the name and choose "data types"

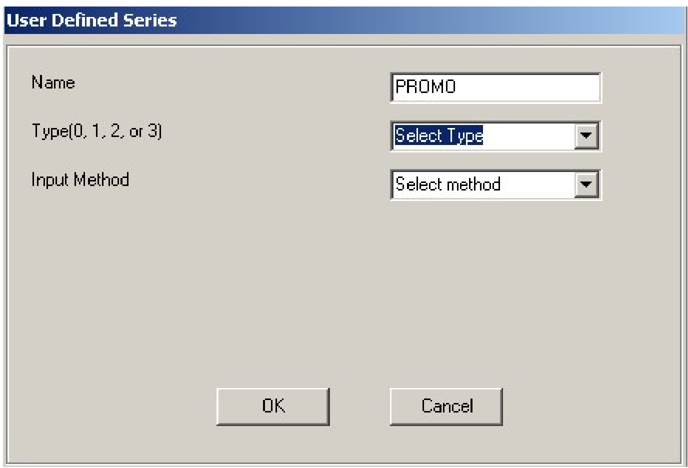


Fig. 4-13 – User Defined Promotion, Data Type

Let's discuss the ***data types*** as it is very important how Autobox treats your causals. If you *know* there is no way for the causal variables to have a lag then you wouldn't want to choose option '2' or '3'. Bad choices here can negatively affect your results.

Data Type Selection:

0 – Autobox will forecast the future values of the causal series and try and identify current and lag effects during the analysis

The next 3 have the future values defined by the user AND

1 - current time period effects (use this for something like “seasonal dummies” or you know there is no lag/lead effect) are analyzed

2 - current time period and lag effects (same as data type ‘0’ except the user is specifying the future values) are analyzed. Series name must be 8 or less characters

3 – current time period and lag an d effects are analyzed. Series name must be 8 or less characters

We choose 3 for this example. If you think an input series could have an effect “In advance” such as unemployment in the next quarter having an effect on sales revenue this quarter then you might specify unemployment to be a type “3”.

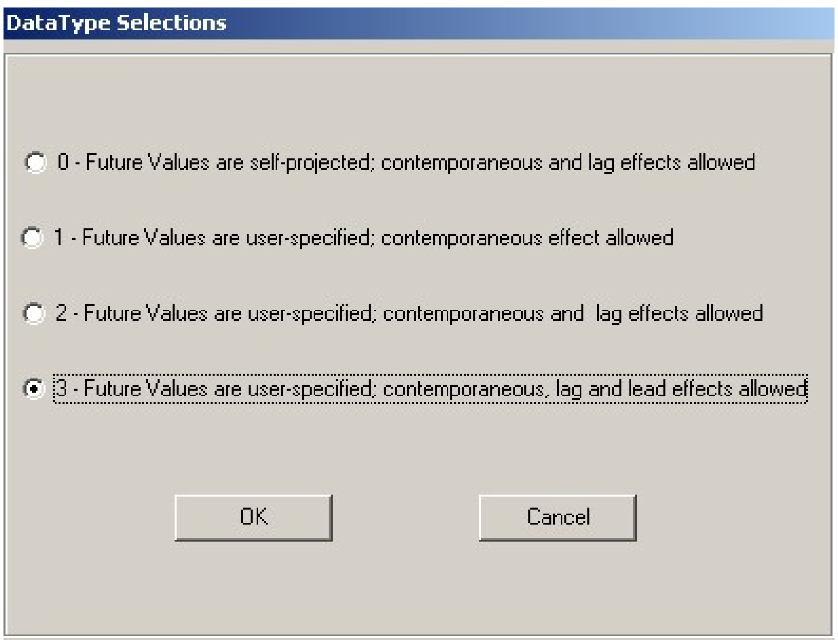


Fig. 4-14 – Data Type Selection Screen

Choose "Manual" as the input method.

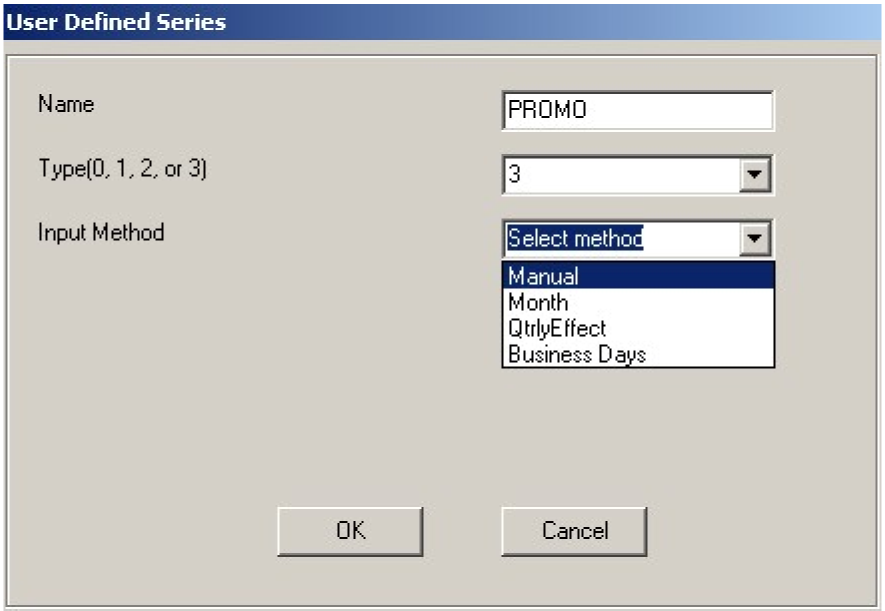
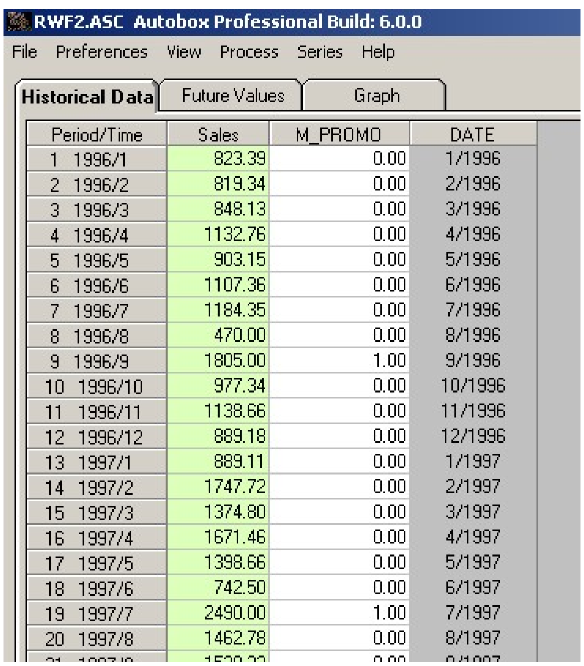


Fig. 4-15 – User Defined Promotion, Input Method List

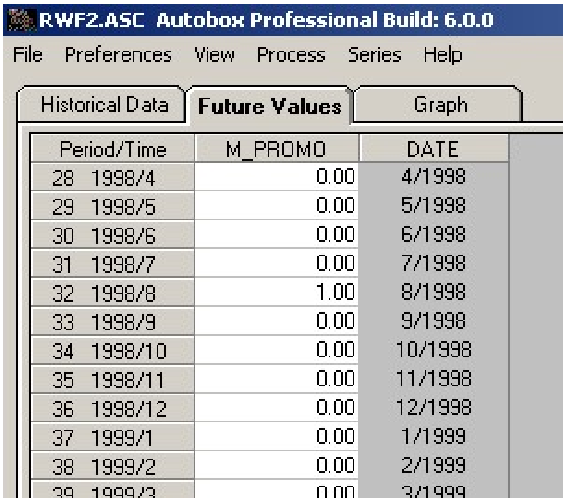
So, now we are enter the historical values of promotion placing a "1" for 9/96 and 7/97.



Period/Time	Sales	M_PROMO	DATE
1 1996/1	823.39	0.00	1/1996
2 1996/2	819.34	0.00	2/1996
3 1996/3	848.13	0.00	3/1996
4 1996/4	1132.76	0.00	4/1996
5 1996/5	903.15	0.00	5/1996
6 1996/6	1107.36	0.00	6/1996
7 1996/7	1184.35	0.00	7/1996
8 1996/8	470.00	0.00	8/1996
9 1996/9	1805.00	1.00	9/1996
10 1996/10	977.34	0.00	10/1996
11 1996/11	1138.66	0.00	11/1996
12 1996/12	889.18	0.00	12/1996
13 1997/1	889.11	0.00	1/1997
14 1997/2	1747.72	0.00	2/1997
15 1997/3	1374.80	0.00	3/1997
16 1997/4	1671.46	0.00	4/1997
17 1997/5	1398.66	0.00	5/1997
18 1997/6	742.50	0.00	6/1997
19 1997/7	2490.00	1.00	7/1997
20 1997/8	1462.78	0.00	8/1997

Fig. 4-16 – Historical Promotion Column Created

We then click on "future values" and enter our future promotion for 8/98.


RWF2.ASC Autobox Professional Build: 6.0.0

File Preferences View Process Series Help

Historical Data | Future Values | Graph

Period/Time	M_PROMO	DATE
28 1998/4	0.00	4/1998
29 1998/5	0.00	5/1998
30 1998/6	0.00	6/1998
31 1998/7	0.00	7/1998
32 1998/8	1.00	8/1998
33 1998/9	0.00	9/1998
34 1998/10	0.00	10/1998
35 1998/11	0.00	11/1998
36 1998/12	0.00	12/1998
37 1999/1	0.00	1/1999
38 1999/2	0.00	2/1999
39 1999/3	0.00	3/1999

Fig. 4-17 – Adding a Future Promotion

Now back to our example, you can enter a "1" at 9/1996 and 7/1997 to show that there was a promotion.

You can then choose "Process/Run" and get your results. You can see that the promotion did have an effect on the sales and the forecast reflects that. You will notice that before the promotion there is a reduction in demand and on the promotion it is increased. You can save your example by choosing "File/Save".

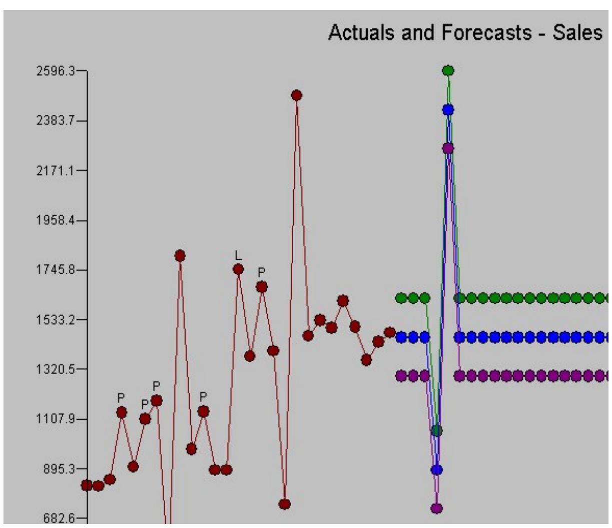
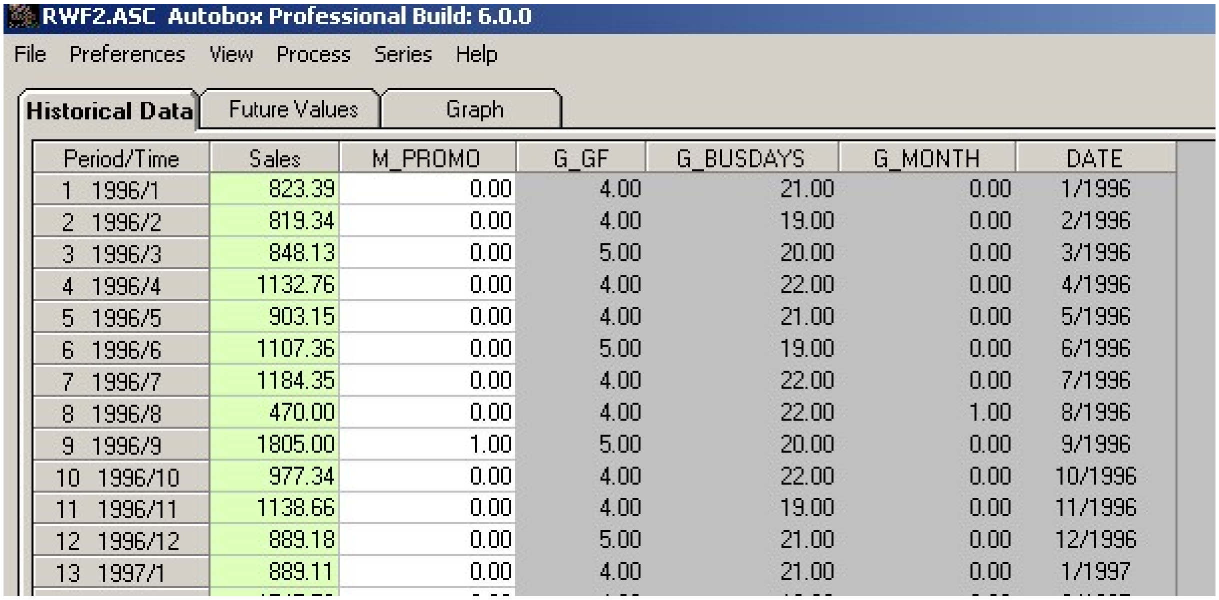


Fig. 4-18 – Promotion Graph

Here are some other examples. if you choose the other “Input Methods” you assign a monthly effect using “Monthly” or a business days effect “Business Days” reflecting the number of trading days in a month or a quarterly effect.



Period/Time	Sales	M_PROMO	G_GF	G_BUSDAYS	G_MONTH	DATE
1 1996/1	823.39	0.00	4.00	21.00	0.00	1/1996
2 1996/2	819.34	0.00	4.00	19.00	0.00	2/1996
3 1996/3	848.13	0.00	5.00	20.00	0.00	3/1996
4 1996/4	1132.76	0.00	4.00	22.00	0.00	4/1996
5 1996/5	903.15	0.00	4.00	21.00	0.00	5/1996
6 1996/6	1107.36	0.00	5.00	19.00	0.00	6/1996
7 1996/7	1184.35	0.00	4.00	22.00	0.00	7/1996
8 1996/8	470.00	0.00	4.00	22.00	1.00	8/1996
9 1996/9	1805.00	1.00	5.00	20.00	0.00	9/1996
10 1996/10	977.34	0.00	4.00	22.00	0.00	10/1996
11 1996/11	1138.66	0.00	4.00	19.00	0.00	11/1996
12 1996/12	889.18	0.00	5.00	21.00	0.00	12/1996
13 1997/1	889.11	0.00	4.00	21.00	0.00	1/1997

Fig. 4-19 – Using Input Methods to Forecast

And one more follow up on this topic.

When you have daily data with a periodicity of 5 (i.e. Mon-Fri) or 7 (Mon-Sun) you must select the date of the beginning observation.

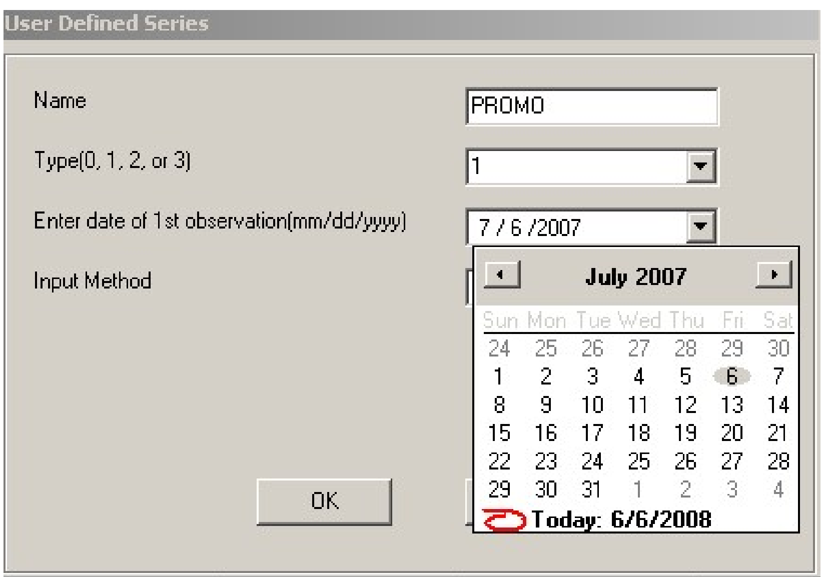


Fig. 4-20 – Selecting First Observation (Due to Data Type)

Also, you can choose some "day of the month" effects like 1st day of the month, 15th of the month or 1st and 15th.

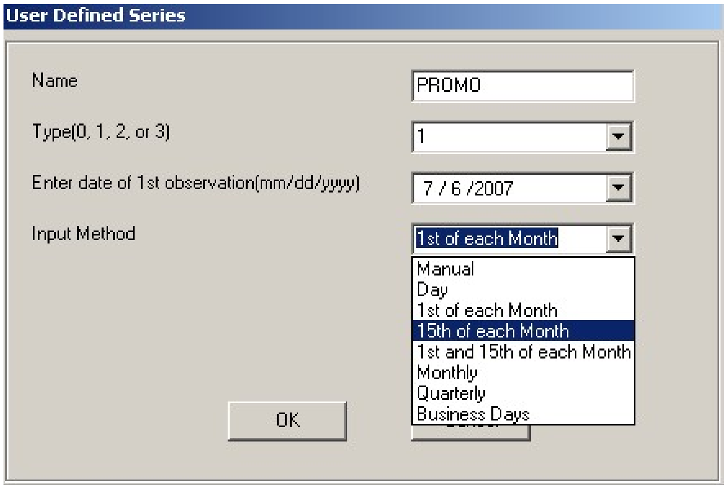


Fig. 4-21 – Daily Effects for Input

Using Causal Variables for Daily Prediction

We will create a new series, by choosing “File/New”. In this example, we will include not only holiday variables, but other variables that will help predict the series.

Enter in these properties

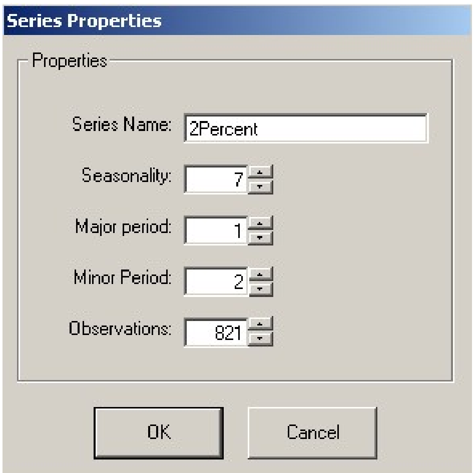


Fig. 4-22 – Daily Causal Properties

Since the actual start was January 1, 2001, a Monday, that would be a minor period of 2 since Sunday is considered the beginning of the week. Before we do anything else, we will need to set the number of forecasts on the right hand side to 90.

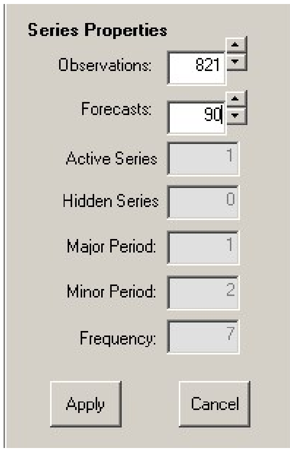


Fig. 4-23 – Setting Properties

We then open Ex3.XLS and paste the data from column "A" into Autobox.

Choose "Series/Add/User Defined" with the following properties:

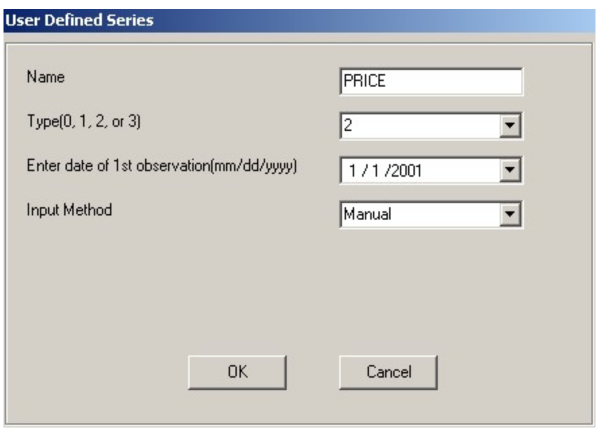
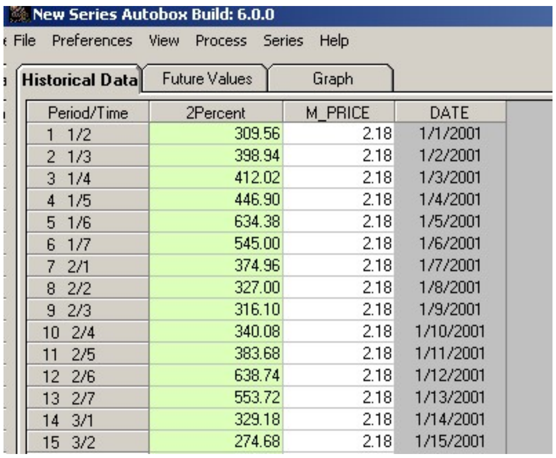


Fig. 4-24 – Ex3.XLS Causal Example (Daily)

Daily data like this will take MUCH longer to run. Series with a periodicity of 5 will be evaluated for "day of the week" and "week of the year" effects. In addition to those two checks, if you have a periodicity of 7 then a "day of the month" effect will also be evaluated (i.e. Social Security checks always come on the 3rd of the month). You can see how complicated actions like this could take longer.

Column B of Ex3.XLS is the price of 2% milk and it contains 821 historical data and 90 future values. Cut and paste the 821 observations into the Price column in the Historical Data spreadsheet.



Period/Time	2Percent	M_PRICE	DATE
1 1/2	309.56	2.18	1/1/2001
2 1/3	398.94	2.18	1/2/2001
3 1/4	412.02	2.18	1/3/2001
4 1/5	446.90	2.18	1/4/2001
5 1/6	634.38	2.18	1/5/2001
6 1/7	545.00	2.18	1/6/2001
7 2/1	374.96	2.18	1/7/2001
8 2/2	327.00	2.18	1/8/2001
9 2/3	316.10	2.18	1/9/2001
10 2/4	340.08	2.18	1/10/2001
11 2/5	383.68	2.18	1/11/2001
12 2/6	638.74	2.18	1/12/2001
13 2/7	553.72	2.18	1/13/2001
14 3/1	329.18	2.18	1/14/2001
15 3/2	274.68	2.18	1/15/2001

Fig. 4-25 – Ex3.XLS Inserting Data

Now, you need to click on the "Future Values" tab so that you can cut and paste in the 90 future values of the price variable into the Price column. You need to go to row 822 in the Excel file to find the future values to paste into Autobox.

Choose "Series/Add/Holidays". Select the date of 1/1/2001 and "select all" holidays and choose "ok".

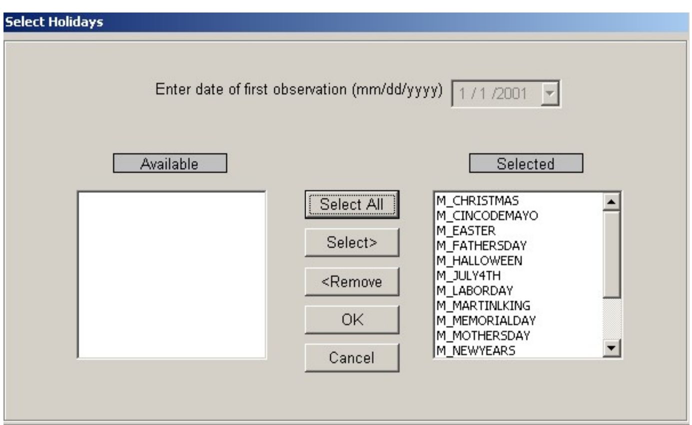
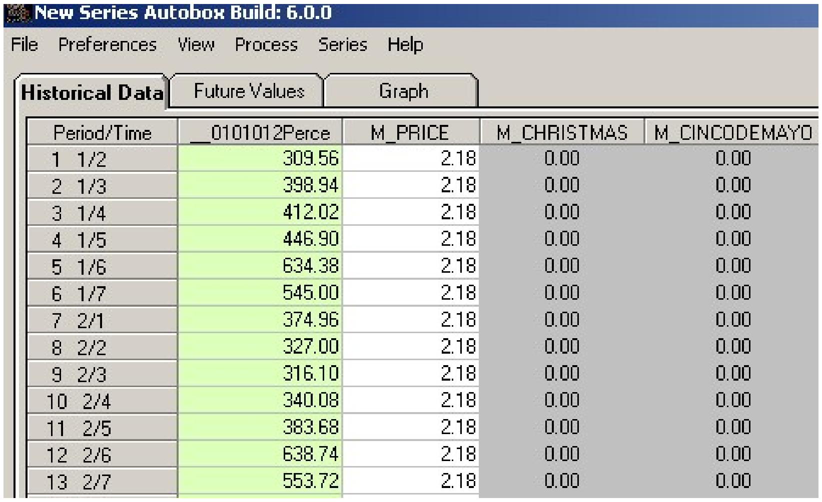


Fig. 4-26 – Adding in Holidays to Forecast

Press OK and Autobox will create dummy variables for each of the holidays… along with future values for the prediction of these events.



New Series Autobox Build: 6.0.0

File Preferences View Process Series Help

Historical Data | Future Values | Graph

Period/Time	__0101012Perce	M_PRICE	M_CHRISTMAS	M_CINCODEMAYO
1 1/2	309.56	2.18	0.00	0.00
2 1/3	398.94	2.18	0.00	0.00
3 1/4	412.02	2.18	0.00	0.00
4 1/5	446.90	2.18	0.00	0.00
5 1/6	634.38	2.18	0.00	0.00
6 1/7	545.00	2.18	0.00	0.00
7 2/1	374.96	2.18	0.00	0.00
8 2/2	327.00	2.18	0.00	0.00
9 2/3	316.10	2.18	0.00	0.00
10 2/4	340.08	2.18	0.00	0.00
11 2/5	383.68	2.18	0.00	0.00
12 2/6	638.74	2.18	0.00	0.00
13 2/7	553.72	2.18	0.00	0.00

Fig. 4-27 – Historical Values with Causal Variables

You can then choose "Process/Run" and get your results.

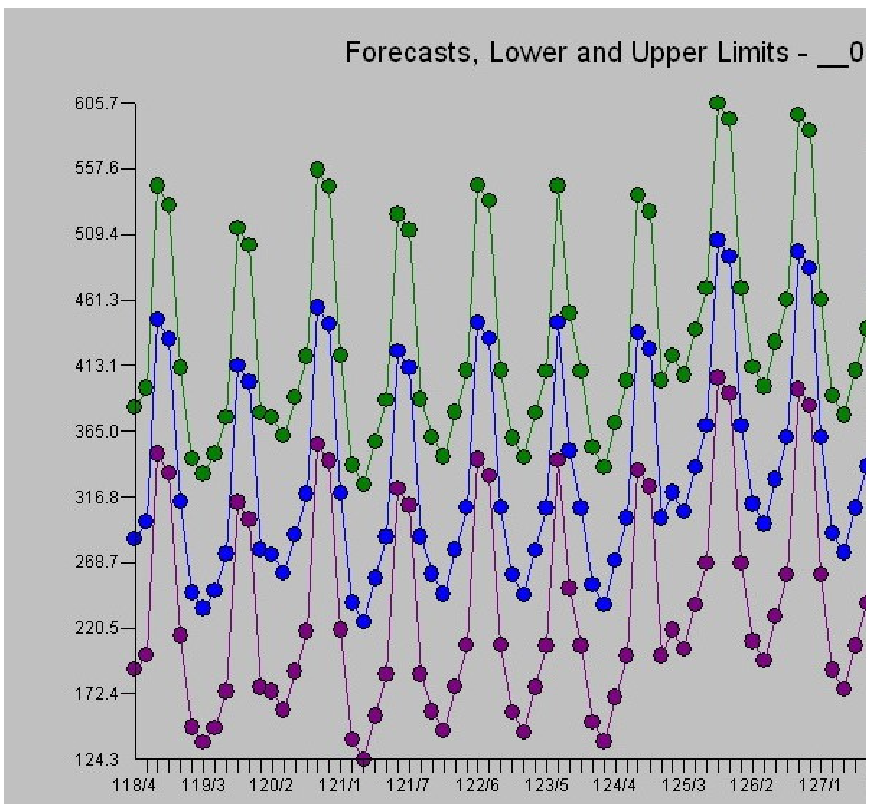


Fig. 4-28 - Results

More on Building Dummy Variables

Let's talk about how to build your own dummy variable a little bit more. If you had a promotion starting in February and it ended in April, you would want to put a '1' for those three months to indicate the event.

Historical Data | Future Values | Graph

	Period/Time	3	M_PROMO
1	2008/1	4186.00	0.00
2	2008/2	3462.00	0.00
3	2008/3	2552.00	0.00
4	2008/4	3045.00	0.00
5	2008/5	1287.00	0.00
6	2008/6	1395.00	0.00
7	2008/7	1290.00	0.00
8	2008/8	1681.00	0.00
9	2008/9	1435.00	0.00
10	2008/10	931.00	0.00
11	2008/11	401.00	0.00
12	2008/12	588.00	0.00
13	2009/1	2225.00	0.00
14	2009/2	15399.00	1.00
15	2009/3	9840.00	1.00
16	2009/4	22822.00	1.00
17	2009/5	6356.00	0.00
18	2009/6	3070.00	0.00
19	2009/7	2709.00	0.00

Fig. 4-29 – Indicating a Promotion

The question now becomes what type of "Data Type" is it? If you think that there is a "lead effect" to this promotion where demand is shifted before the event then this would suggest that the type should be set to '3', but it would be not be logical to have three dummy 1's consecutively when you are looking for a lead and lag relationship. There should just be 1 dummy set in March and the February and April should be a '0'. Autobox looks for lead effects up to 6 periods before the '1' dummy is specified. As a protective measure, Autobox will change your type from a '3' to a '1' if you have 1's that are within 6 periods of one another so be careful to follow these previous words or call us to discuss.

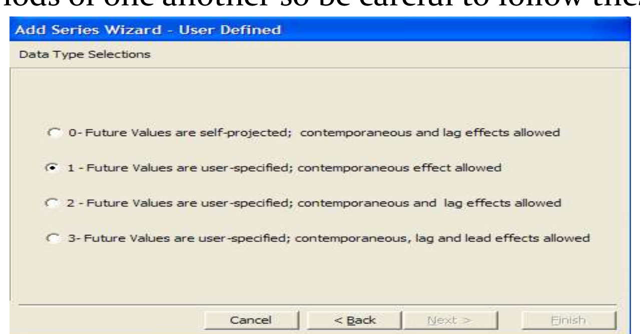


Fig. 4-30 – Defining the Data Type for the Data Set

Chapter 5

Overriding the Expert System

Why would you need this? When you run Autobox using option "1" it uses specific default conditions within the algorithm. If you want to override the defaults, we let you do that. You need to create trigger files in the installation directory and there is existence will cause Autobox to react. Some overrides require data to be in the file which are simple text files. The trigger files need to be named "*.AFS".

Here are the trigger files in alphabetical order:

2season.afs - THIS IS NOT JUST A TRIGGER FILE, BUT ALSO HAS CONTENT. If you have an unusual time series in that it doesn't have standard seasonality (ie monthly, quarterly, hourly) you can still model it using Autobox. If you have an hourly data that only covers 20 out of the 24 hours of the day, you can create a special file named '2season.afs' in the installation directory. Inside of the text file place a '20' in it so that Autobox can model it.

AFS2SIP.afs – This is called "Simulated Forecasting" or bootstrapping via Monte Carlo methods. This comes with the Enterprise + version of Autobox. Autobox generates Sipmath 2.0 ready XML files(see probabilitymanagement.org for more on their free and very useful simulation tool). Put "1000" in this file and Autobox will generate 1,000 simulations for each of the forecasting periods(see image below). The second line can have a 0 or 1. This determines if outliers will be allowed to exist in the future or not. We recommend you put a 1 in the second line so that outliers that were identified and adjusted for in the history will be allowed to occur in the forecast and in effect widen the confidence limits providing more realistic version of uncertainty! If you put a 0 on the second line then then no outliers are used in the forecast(more for educational use to show you the old vs the new). We recommend using 1,000 so that there are enough samples/trials are drawn to provide good results. If you are forecasting out 36 months then you will have a total of 36,000 forecasts along with the average, standard deviation, min and max. After running, look for a file in the Autobox folder with an "XML" file extension (ie.Sales.XML). The sampling needs a "seed value" which we use the clock to drive from. If you run an example, you will not get the same results as the time of day changes. If you want to reproduce, you can create a file called "seed.afs" and put a number in it so that the results don't change. We report the seed at the bottom of details.htm that was used in the simulation which you can post to seed.afs if you want to use that same outcome. You need to delete nosim.afs from the Autobox folder as there is some additional compute time to do this step. 500 would work fine and 1,000 gives similar results. 1,000 is actually testing the limits(think possible crash?) of Autobox with all of the memory that it uses with complicated algorithms.

Autobox generates *.FUM files(forecast univariate model) for each of the causal variables. Make sure to delete the previous *.FUM files if you made a change to the data.


```
<SLURP name="Example" coherent="True" about="Example SLURP of forecasts" origin="Autobox">
<SIP name="Forecast00001" Base="   58324.   " SipMean="   58296.   " SipStd="   596.99   " SipMin="   57140.   " SipMax="   59326.   " type="CSV"
 count=" 1000" csvr="02" about="Example SIP">
   58127.00,  58221.00,  58576.00,  58565.00,  57532.46,  58324.18,  59326.46,  58216.92,  57947.00,  58127.00,  58221.00,  57532.46,  58979.00,  57436.
</SIP>
<SIP name="Forecast00002" Base="   55810.   " SipMean="   55783.   " SipStd="   599.30   " SipMin="   54626.   " SipMax="   56812.   " type="CSV"
 count=" 1000" csvr="02" about="Example SIP">
   55702.92,  55433.00,  55613.00,  55707.00,  55018.46,  56465.00,  54922.00,  55925.00,  56170.50,  55810.18,  54922.00,  56074.92,  55810.18,  55810.
</SIP>
<SIP name="Forecast00003" Base="   58451.   " SipMean="   58437.   " SipStd="   602.39   " SipMin="   57267.   " SipMax="   59453.   " type="CSV"
 count=" 1000" csvr="02" about="Example SIP">
   58566.00,  58811.50,  58451.18,  57563.00,  58715.92,  58451.18,  58451.18,  57358.00,  57659.46,  59426.50,  58074.00,  59106.00,  59106.00,  58625.
</SIP>
<SIP name="Forecast00004" Base="   59824.   " SipMean="   59807.   " SipStd="   603.81   " SipMin="   58640.   " SipMax="   60826.   " type="CSV"
```


If retained data is in effect then simulation is disabled
If # of forecasts is greater than the maximum allowed # of series then simulation is disabled.
If intermittent demand is in effect then simulation is disabled (note nointerm.afs could control for this)
If non-constant variance is detected (tsay test) then simulation is disabled (note notsay.afs could control for this)
If power transform is in effect then simulation is disabled (note novarcon.afs could control for this)
If maximum # of allowed series is less than 150 (enterprise or Pro versions) then then causal simulation is disabled
Causal simulation is only available for type “0” series or type “2” series
Be careful with causals in that if you run and Autobox

This suppresses the breakin.txt and breakout.txt reports which detail how the fit and forecast are generated. To suppress simulation create an override called nosim.afs. Nogood reason toever do this.

Delphi.afs – You can specify forecasts with probability and ranges to be used with the simulated forecasting in this file.(see section on Simulation with Delphi method)

Foreconf.afs – THIS IS NOT JUST A TRIGGER FILE, BUT ALSO HAS CONTENT. This determines the confidence level for the confidence limits (For 80% confidence limits put ‘80.0’)
Integer.afs – Converts forecasts to integers

Interp.afs – Uses Interpolation to correct missing data.
Makealli.afs – Keeps all causal variables in the model.

Noarima.afs – Suppresses the use of ARIMA modeling. With daily data you don't want to overreact to the most recent data too much as the level shifts will do that for you. Daily data relies on monthly and daily dummies (ie deterministic) behavior.

Nofixday – For daily data, If you don't have special days of the month.
Noholdum.afs - If a user specifies their own holidays this suppresses Autobox from automatically generating holiday variables (ie only matters when series name is "__010108Y11" with daily data)
Nointer.afs – Stops testing for interactions. For use with unusual seasonalities only.
Nointerm.afs – Stops looking for intermittent demand. Useful with daily data
Nolevel.afs – Stops testing for level shifts
Noparcon.afs – Stops the testing & adjustment for constancy of parameters (ie. Removes older data that has a different pattern than the more recent data)
Nopulse.afs - Useful when you want to believe the outliers

Norule.afs – To stop these two overrides use this file. For daily data, Autobox will check if all weekends are zero and force the forecast weekends to be zero and for holidays the day of the week is not included to avoid double counting.

Noseasp.afs - Stops testing for seasonal pulse
Nostepdn.afs - Stops removal of causals.
Notrend.afs - Stops testing for local time trends.
Novarcon.afs – Stops the testing & adjustment for constancy of variance (ie. weight the observations based on their variance)

Noweekly.afs – When you want to force monthly dummies for short daily data. FYI: Autobox flips between monthly dummies and weekly dummies based on the length of the series.

Numbfore.afs – THIS IS NOT JUST A TRIGGER FILE, BUT ALSO HAS CONTENT. This defines the number of forecasts (For 24 forecasts put a '24' in this file)
Parpharm.afs - Put a '1' for a mean model. Refer to the Appendix "Creating your ASC file" and you will see that "Objective 1" has a list of different prespecified models that can be used to override Autobox.
Positive.afs – Converts forecasts to positive values
Sparse.afs – Specify two digits (ie .25 for that govern intermittent demand. Specify the % of zeroes in a dataset that trigger. Default is 25% in Autobox.
Stepupde.afs – THIS IS NOT JUST A TRIGGER FILE, BUT ALSO HAS CONTENT. This defines the number of interventions to be used in the model. (For 5 interventions put a '5' in this file) for daily data we recommend 80 outliers for 3 years of data.

Weekinmo.afs – Creates weekly dummy effects – Good for daily data that has a big month end, for example.

Here's a trick when you have long lagged response to a marketing event or "Adstock" type of data. For example, drilling done at an Oil Well will also have long lagged responses to injection of water and chemicals amongst similar type of data. If you specify the causal series name with the words "DYNA14" and place a "1" when the event occurred and leave all of the other data as a "0", Autobox will react by searching for a decay impact over the next 14 periods. Remember that the version of Autobox you are running might be constrained to 6/30/150 variables and you might blow out the array of the program if you are careless here. Note: You can only have names as long as 6 characters where the first 3 are "DYN" and the fourth is a letter or a number and 5th and 6th is the lag structure.

You will also need to put a unique character for each DYN variable. Note that we put an "A" after DYN to make the name unique. The max lag is 50. If you have a lag of 4 you need to put 04(ie DYNG04)

Note that the "data type" MUST be '1' for this to work.

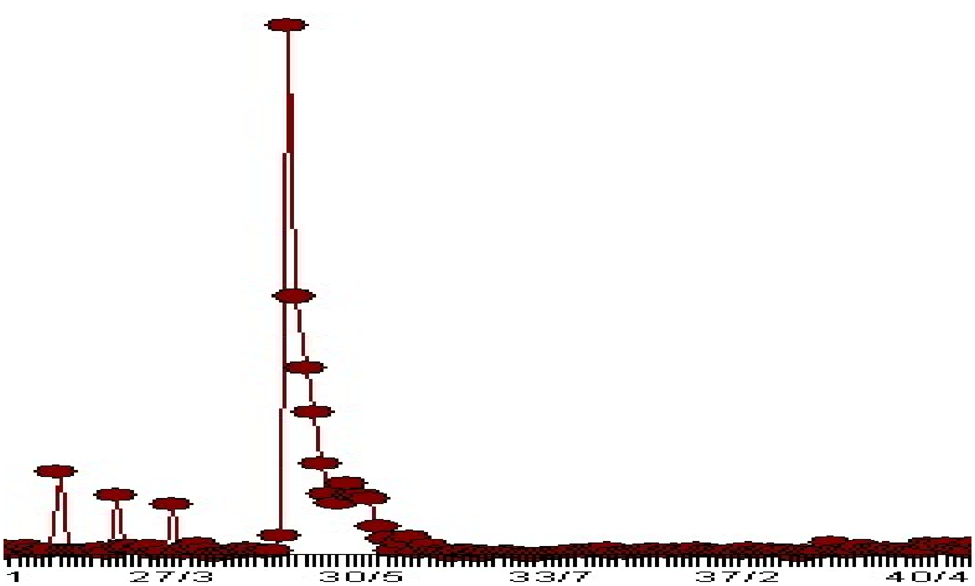


Fig. 5-0 – Example of a Long lagged impact

A More In-Depth look at the Model and Rules Wizard

As the wizard indicates, your first need to choose what you want to do.

Option 1 – Run using proprietary AFS rules
Option 2 – Run using User's overrides of AFS rules
Option 3 – Choose a specific model like Holt Winters – it's only there for comparison in our opinion!
Option 4 – Choose your own model – this is for experts…be careful here
Option 5 – Runs only the identification process

If you select option '2' and press "Next", you will be given the option to use the "Edit/Use Present or Existing Rules" (press the down arrow to be shown what is available). This will use the current rules or you select rules you had saved previously. If you want to reset the conditions back to the original factory installed defaults used choose "Restore Default Rules" and click "Next".

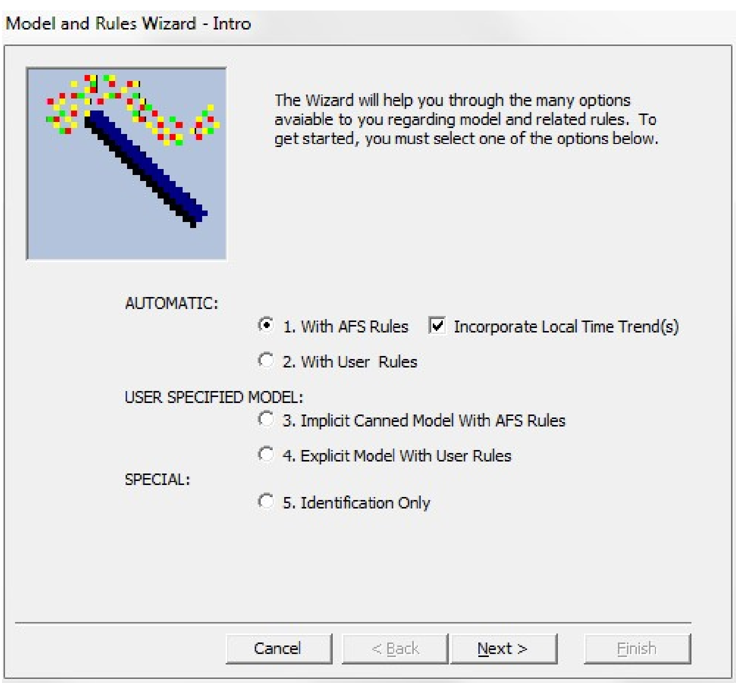


Fig. 5-1 – Rules Wizard

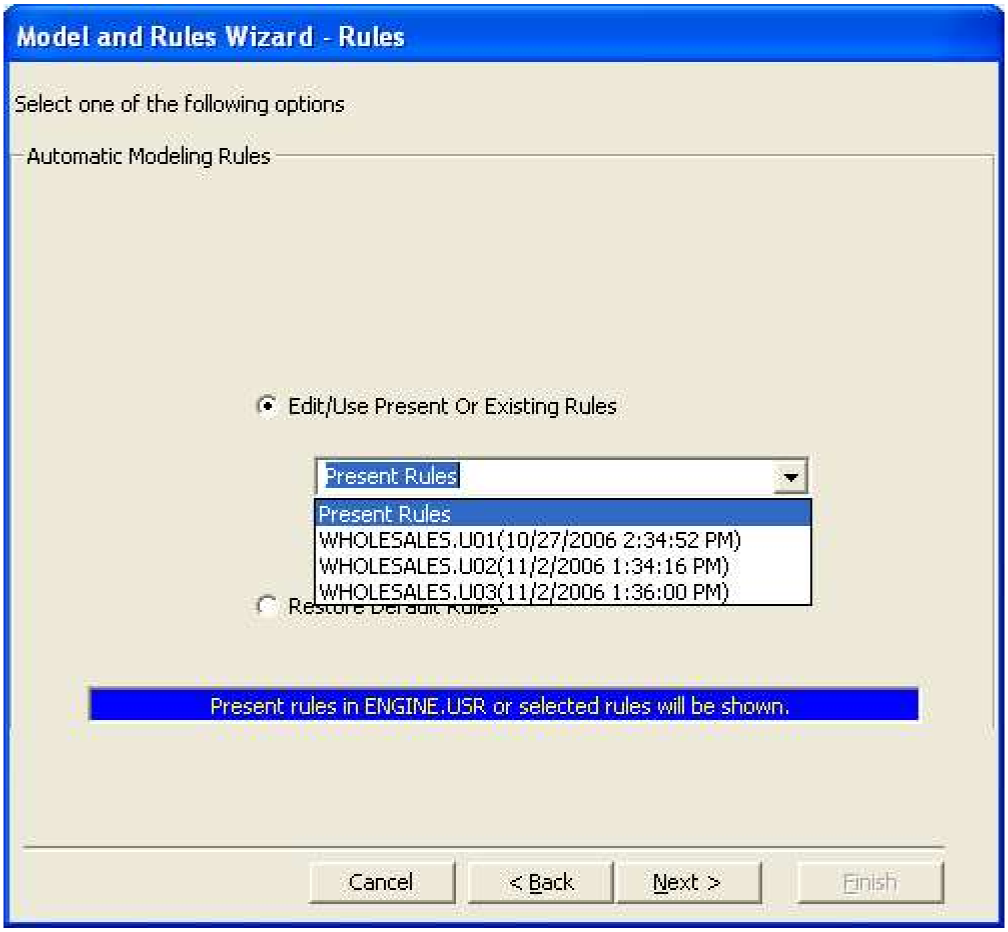


Fig. 5-2 – Option 2, Selecting Present Rules

Intermittent Demand

When you select the rules and press “Next”, they will be shown in the following order for your review and/or editing (please note that the rules vary between Noncausal(single) series and Causal(dependent series with independent series)series –we will show the causal screens after this section.

If you select this option, Autobox will check to see if more than 25% of the observations are zero and then run the following scheme to predict the intermittent demand.

Two new time series will be created: the interval and the rate. The interval series will be the number of periods with zero demand between periods where there is demand (i.e. If there is demand at every time period then the interval would be 1). The rate is the interval divided by the demand. We find this as a better alternative to Croston’s method.

Autobox approaches the case of Intermittent or Sparse Demand by making the problem into a causal problem. Autobox tracks the count of how many periods were zero and creates an “interval” time series. It takes the volume and divides it by the interval to get a “rate”. It projects out the interval and rate into the future and divides the two to get a forecast. Autobox will do all of this while still searching for level shifts. You need to have >25% zeroes in the series and 4 intervals to trigger intermittent demand.

You can override the % by creating a file in the installation directory named sparse.afs with .70 in the file to reflect a threshold of >70%.

You can also suppress intermittent demand by creating a file called “nointerm.afs” in the Autobox folder. We recommend creating an override file called “nopulse.afs” to NOT search for pulse outliers when dealing with intermittent data. We allow you decide which way to run it.

Fig. 5-3 – Option 2, Determining the Model

We will skip "Multiple Groups/Panels for now. It is discussed later. If you choose to change the "convergence criteria", you really won't need to make a change to the defaults. You can change the way Autobox iteratively goes about identifying the model. The % change in error sum of squares is set to 0 so you can't be more precise. The % change in parameter values are set to .1% and you can increase the iterations to a maximum of 200, but it doesn't make much of a difference.

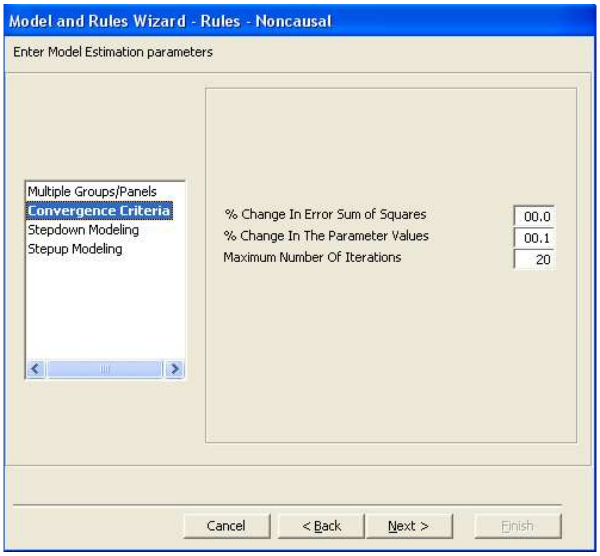


Fig. 5-4 – Option 2, Entering Model Parameters

For step down modeling (where variables are removed or weighted), you can either adjust the confidence level for the option or "uncheck" the box so that the test is disabled. If these are checked then:

Necessity - This will drop all variables that are introduced by Autobox (i.e. outliers). (The opposite case is that it keeps all variables)

Test for Constancy of Parameters - Autobox will determine if you have too much historical data that does not match the more recent data and delete older data.

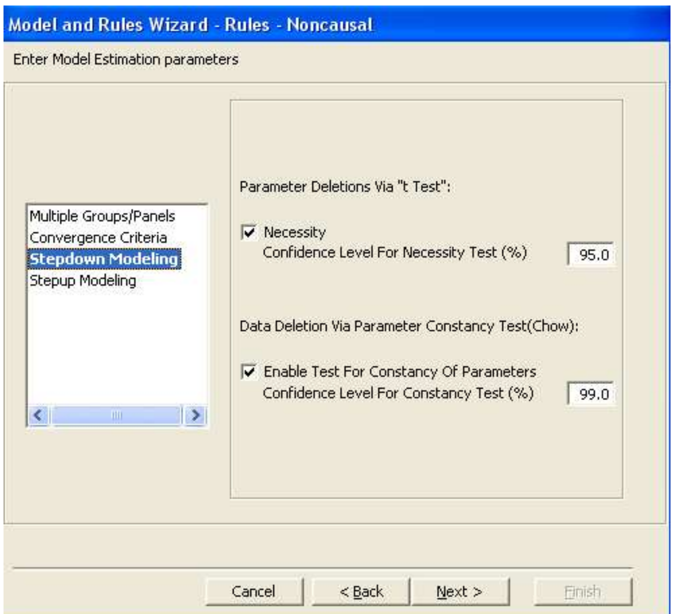


Fig. 5-5 – Stepdown Modeling Parameters

For stepup modeling (where variables are added into the model by Autobox), you can either adjust the confidence level for the option, “check” or “uncheck” the box, or modify the counts.

Sufficiency Test for Stochastic Structure- This will add in ARIMA variables if significant.
Sufficiency Test for Determinstic Structure – This will add in Intervention variables.
Maximum No. Of Outliers To Be Identified – This will limit the number of outliers that can be introduced.
Include Pulse Variables – This will allow one-time unusual variables to be accounted for.
Include Step Variables – This will allow sudden changes in mean to be accounted for.
Min. no. Of Observations In Group – This determines how many observations you need before you can create a step variable.
Include Local Trends – This looks for changes in intercept.
Include Seasonal Pulse Variables – This looks for pulses that occur every season.
Enable Auto Fix Up For Fixed Effects – For example, for monthly data, Autobox would include 11 dummies for monthly data to account for fixed effects.

<u>Discrete Change Test For Variance</u> – This will search for changes in variance and weight the history based on the observations.
<u>Minimum number of Residuals to Pool</u> – There are two groups created to compare the variance and this determines the smallest size of one of the groups.
<u>Enter Lamda Values</u> – click "Yes" to search for optimal variance stabilization

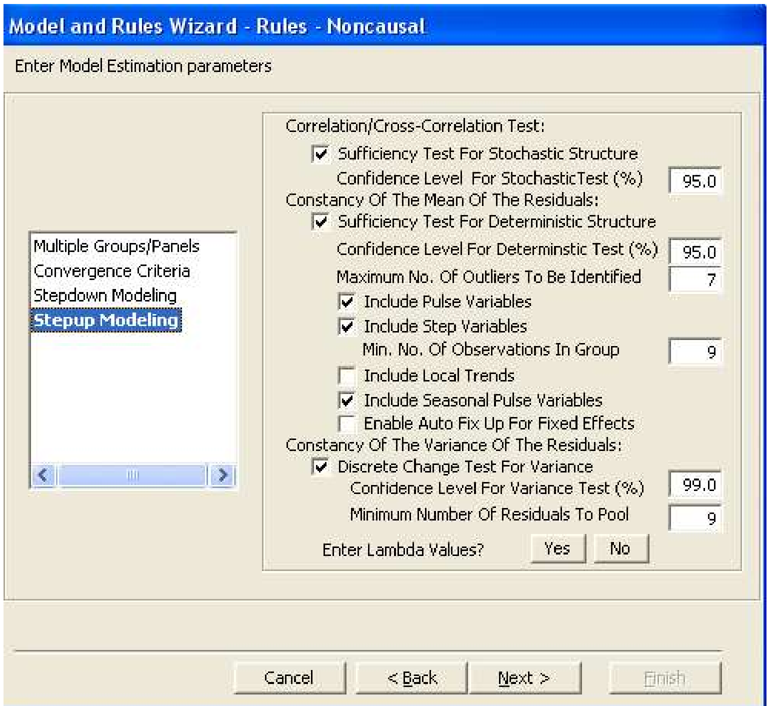


Fig. 5-6 – StepUp Modeling Parameters

Several common used variance-stabilizing transformations

Relationship of σ^2 to $E(y)$	Transformation
$\sigma^2 \propto$ constant	$y' = y$ (no transformation)
$\sigma^2 \propto E(y)$	$y' = \sqrt{y}$ (square root: Poisson data)
$\sigma^2 \propto E(y)(1-E(y))$	$y' = \sin^{-1}(\sqrt{y})$ (arcsin)
$\sigma^2 \propto (E(y))^2$	$y' = \log y$
$\sigma^2 \propto (E(y))^3$	$y' = y^{-1/2}$ (reciprocal square root)
$\sigma^2 \propto (E(y))^4$	$y' = y^{-1}$ (reciprocal)

Enter 0 for logs in row 2 and -1 for reciprocals in row 3, for example

The bottom option is the “Enter Lambda Values”. This allows you to determine the form of the Box-Cox Transformation. The first observation is always 1 as a rule.

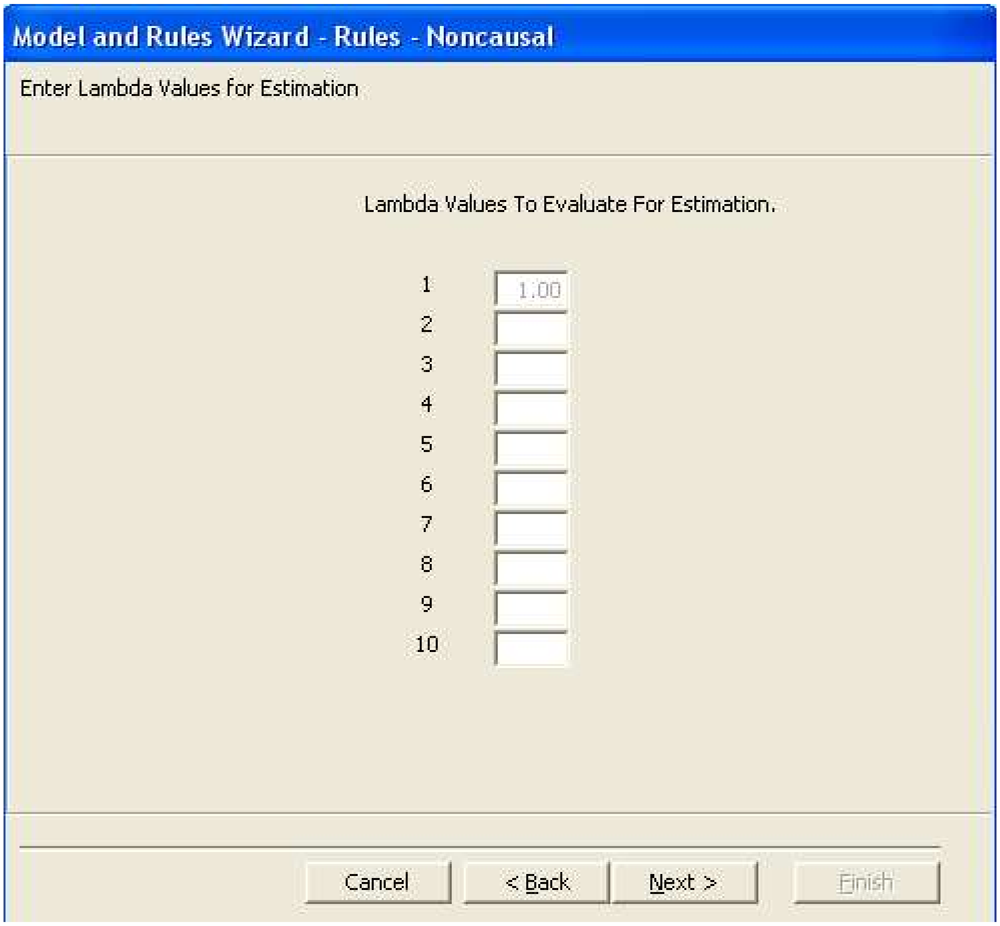


Fig. 5-7 - Setting Lambda Values

Click “Next” and the “Model Forecasting” parameters will be shown as follows(Please note that to get an explanation/definition of a particular checkbox or textbox merely click on it and press F1):

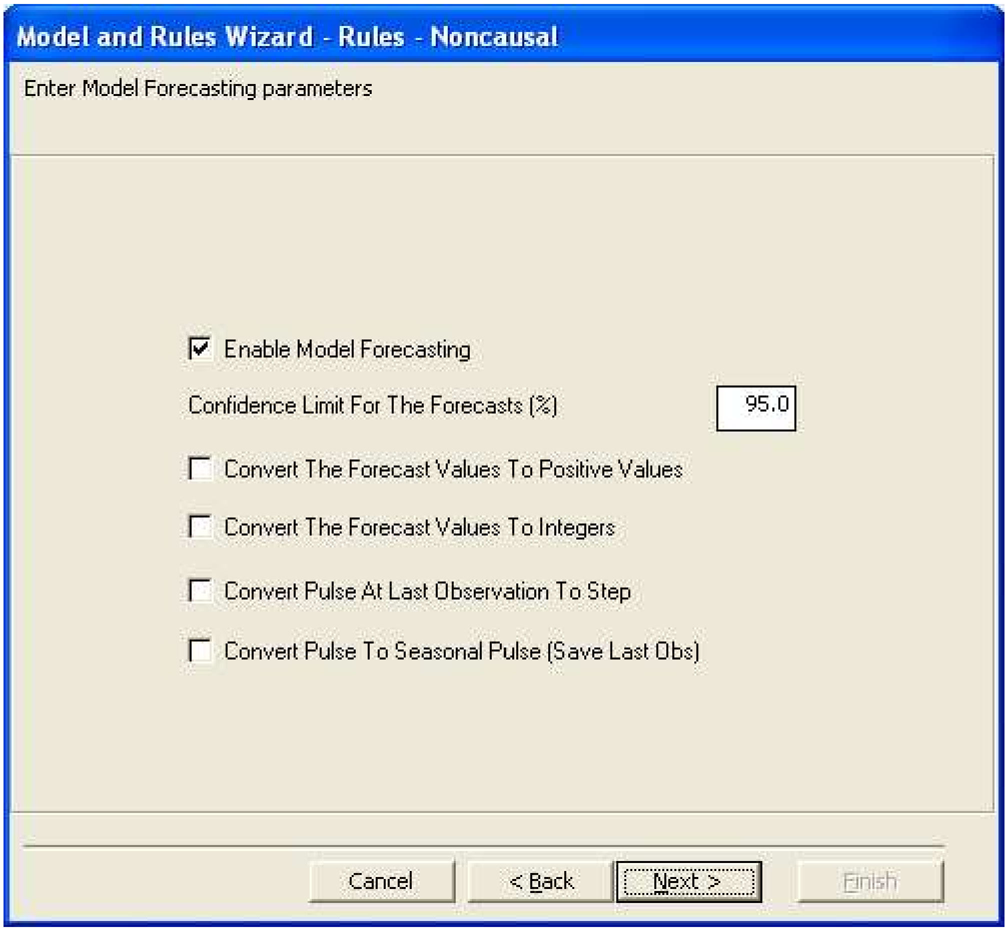


Fig. 5-8 – Model Forecasting Parameters

This is the process of using the model that best describes the past observations (to extrapolate the pattern of the data) to predict Confidence Levels into the future.

If this option is not enabled, then no forecasts will be made.

Press “Next” and you will be given the option to save changed rule and/or the final model as indicated in the following:

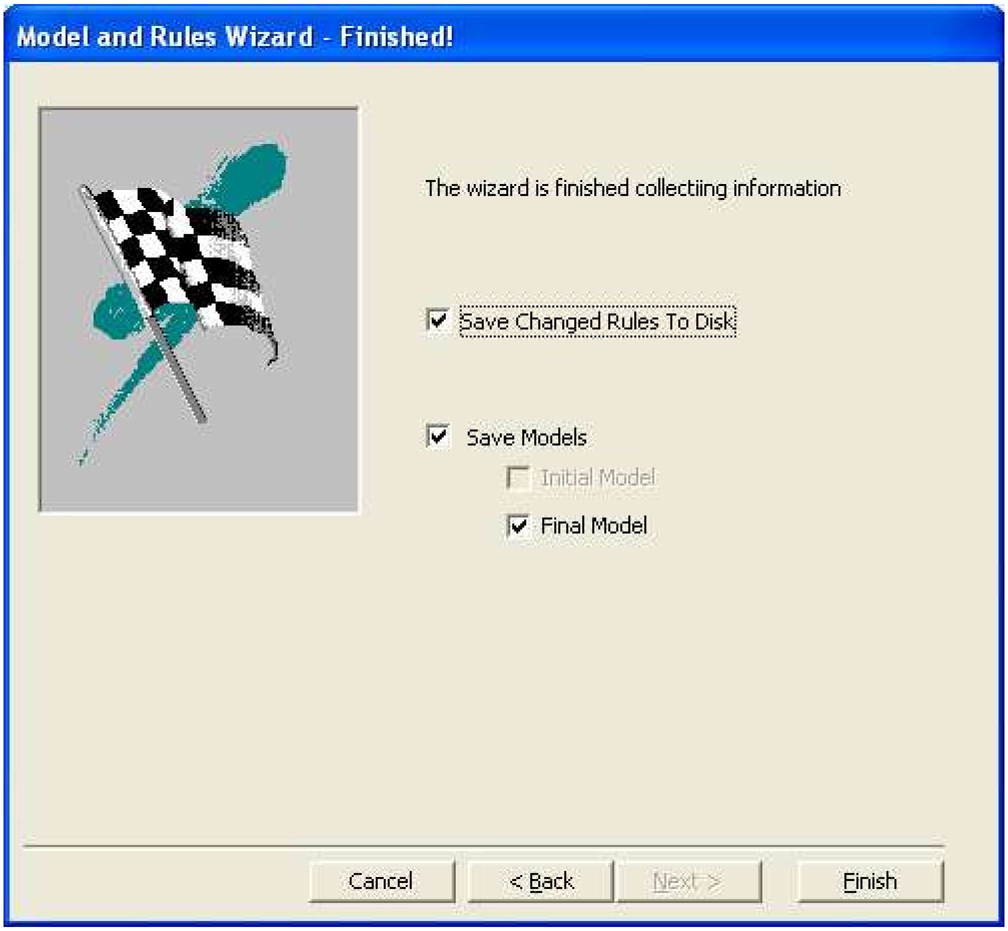


Fig. 5-9 – Model and Rules are Set, Prepared to Save

Again, press “Finish” and the dataset will be processed automatically.

If you selected option '2' and had causal variables then this additional screen would have been presented to you at the beginning.

Use Difference Factors From ARIMA Model In Causal Model – This might be helpful in identification and problematic in estimation.

Constrain All User Causal Coefficients in Model – This keeps all user causal variables in model regardless of their significance.

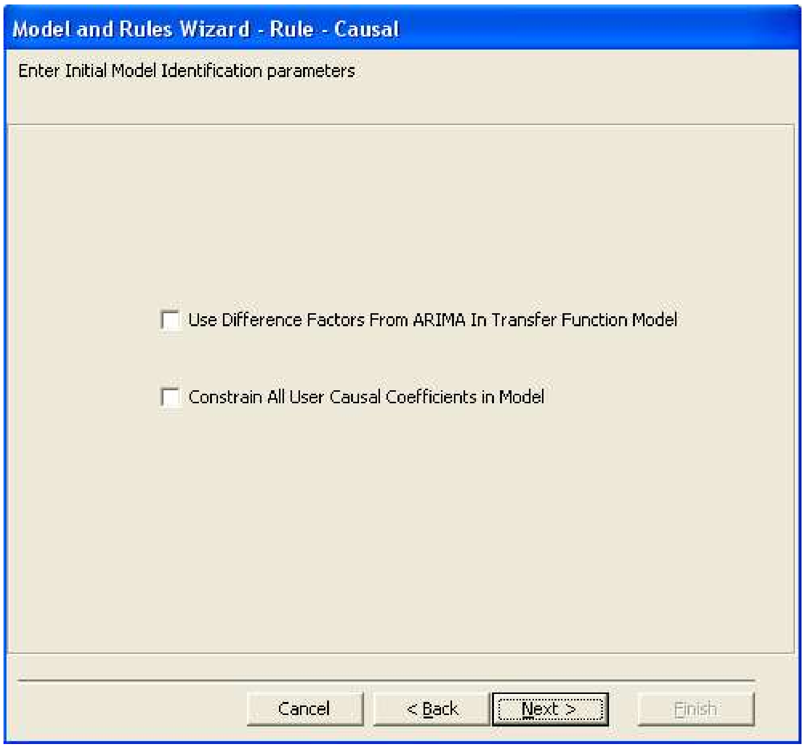


Fig. 5-10 – Causal Rules

If you select option '3' and press "Next", you must select one of the offered models as indicated in the following:

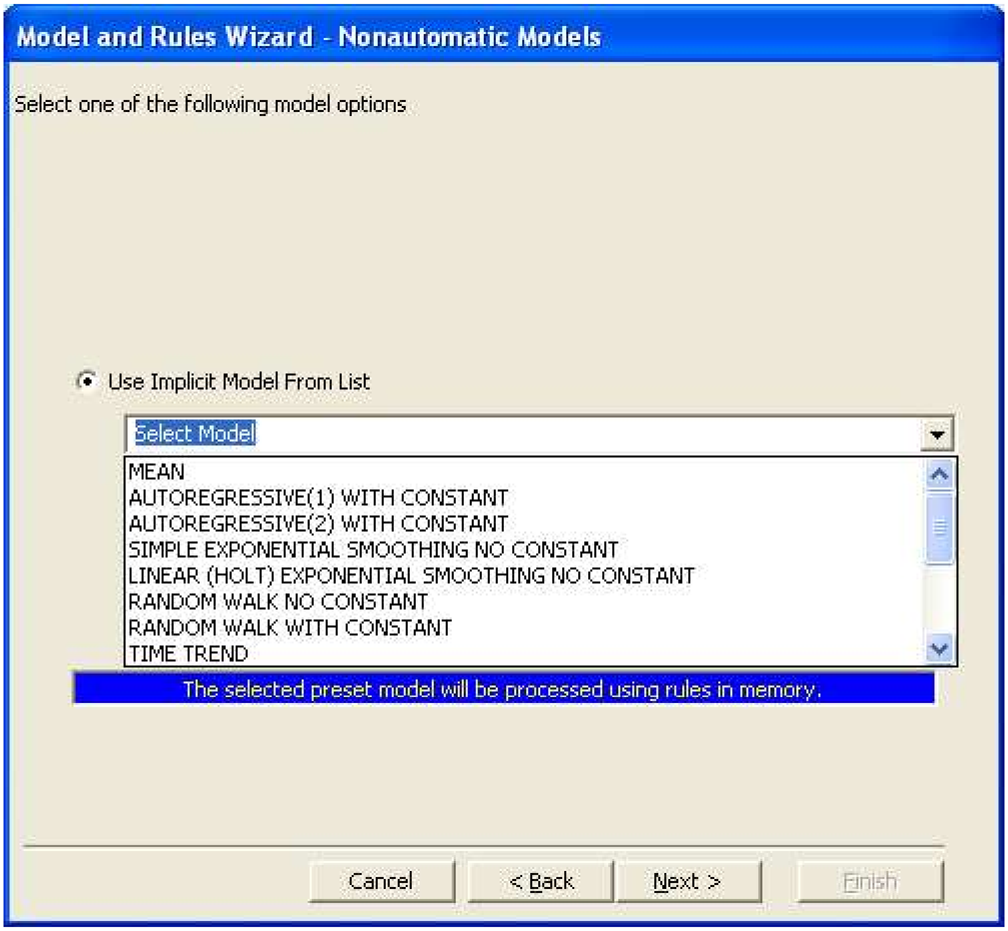


Fig. 5-11 – Option 3, Models

If you click the box “Allow Modifications” then Autobox will use all of its “expert system” approaches to try and create a better model given you picked a starting model like “Time Trend” as seen here. Note that we believe this is dangerous territory as all of the models in this area are typically a bad place to start.

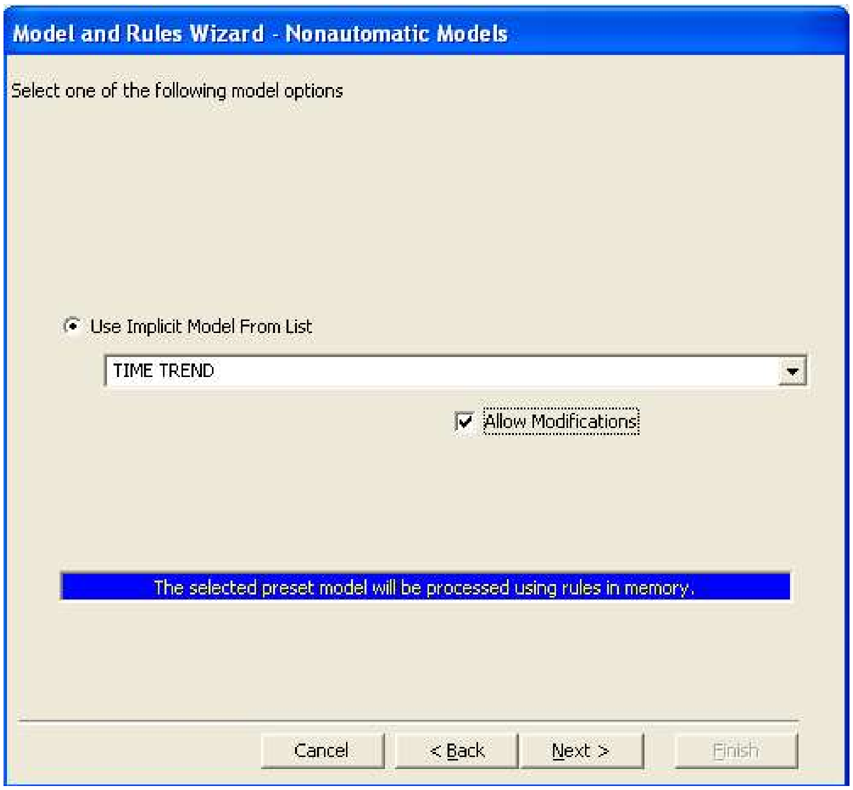


Fig. 5-12 – Option 3, Selecting a Preset Model

Press “Next” and you will have the option to save the final rules and model, then press “Finish” to process the dataset.

If you have causals then you have two “pre-selected” models to choose from along with the ability to choose “Allow Modifications”. Again, this is dangerous territory as opposed to using option ‘1’.

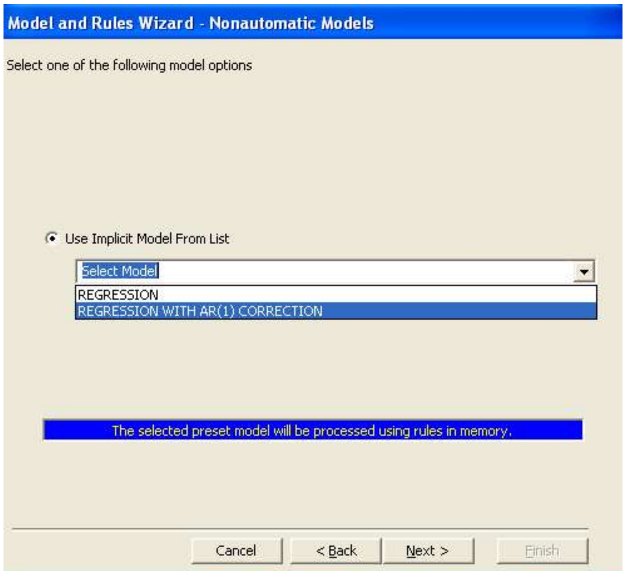


Fig. 5-13 – Option 3, Modifications

If you select option ‘4’ and press “Next”, you can choose to make a “New” model or use one of the saved models indicated in the following. Click “Next”

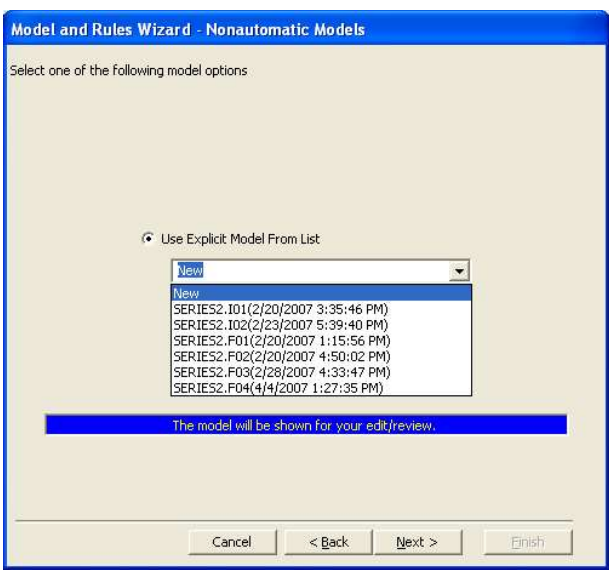


Fig. 5-14 – Option 4, Set Model

If you only have one series in your dataset you see the words “Noncausal” twice at the top of the screen.

If you only have multiple series(not shown) in your dataset you see the words “Causal” at the top and the word “noise” here. The noise refers to the ARIMA structure in the transfer function model. We will show you a causal example on the next screen.

This model would be of the form ARIMA(2,1,1)(1,1,0)12 . You can put any coefficient as a starting point.

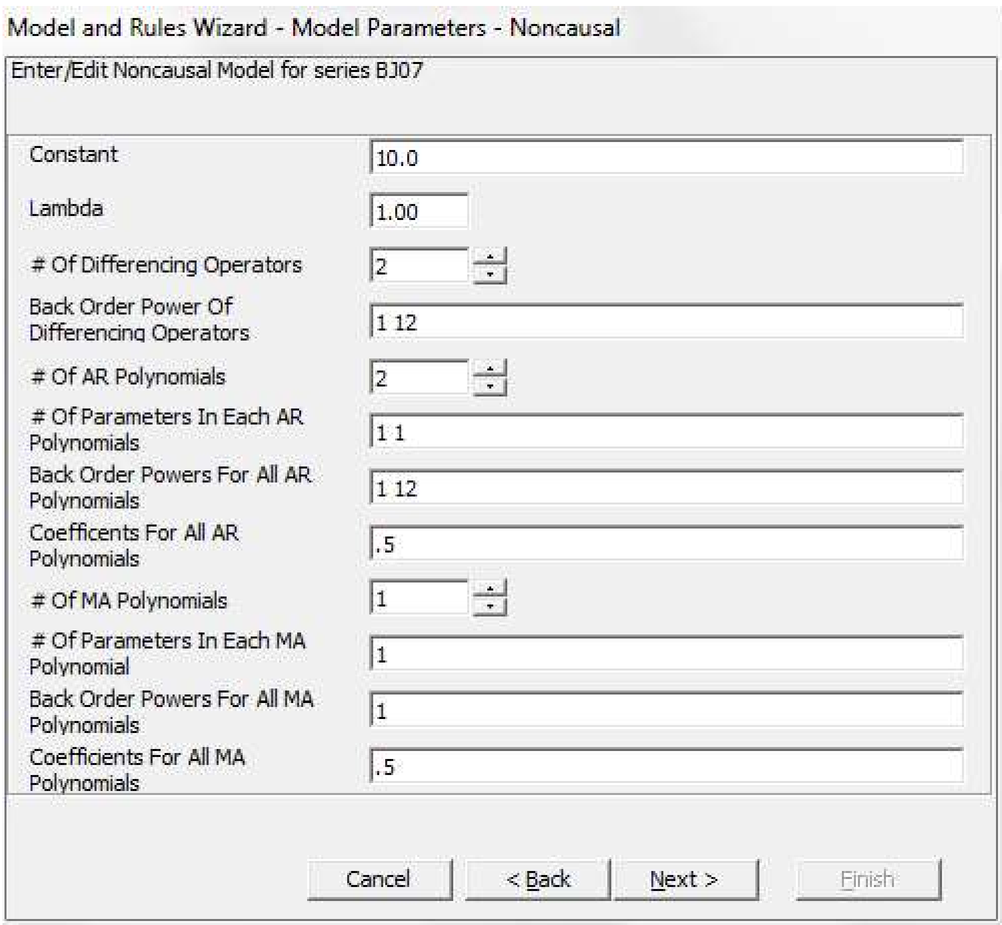


Fig. 5-15 – Option 4, Noncausal Model Creation

Here we specify the model for aom73, but we click “On Copy Parameters To” and then click on all of the other causal series so that they have the same model applied.

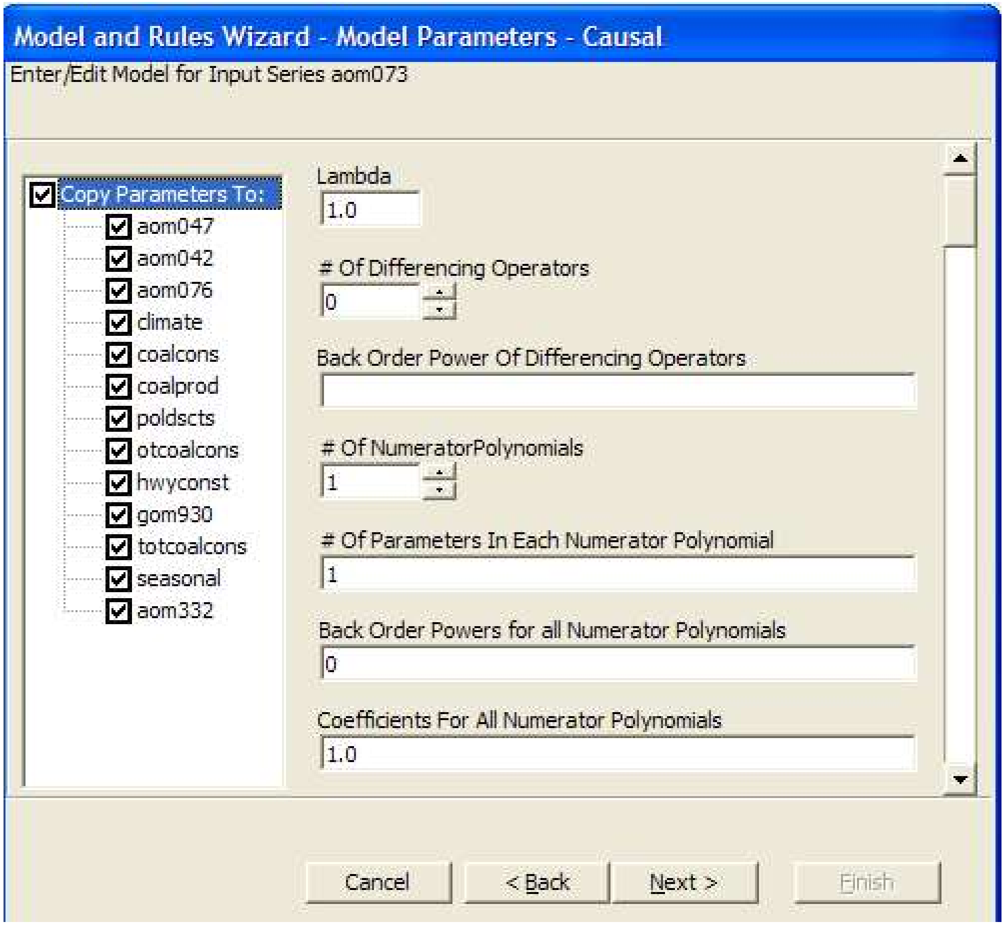


Fig. 5-16 – Option 4, Copy Parameters on All Series

Press “Next” and you will be given the option to use the “Edit/Use Present Or Existing Rules” (press the down arrow to be shown what is available) or “Restore Default Rules” and as follows:

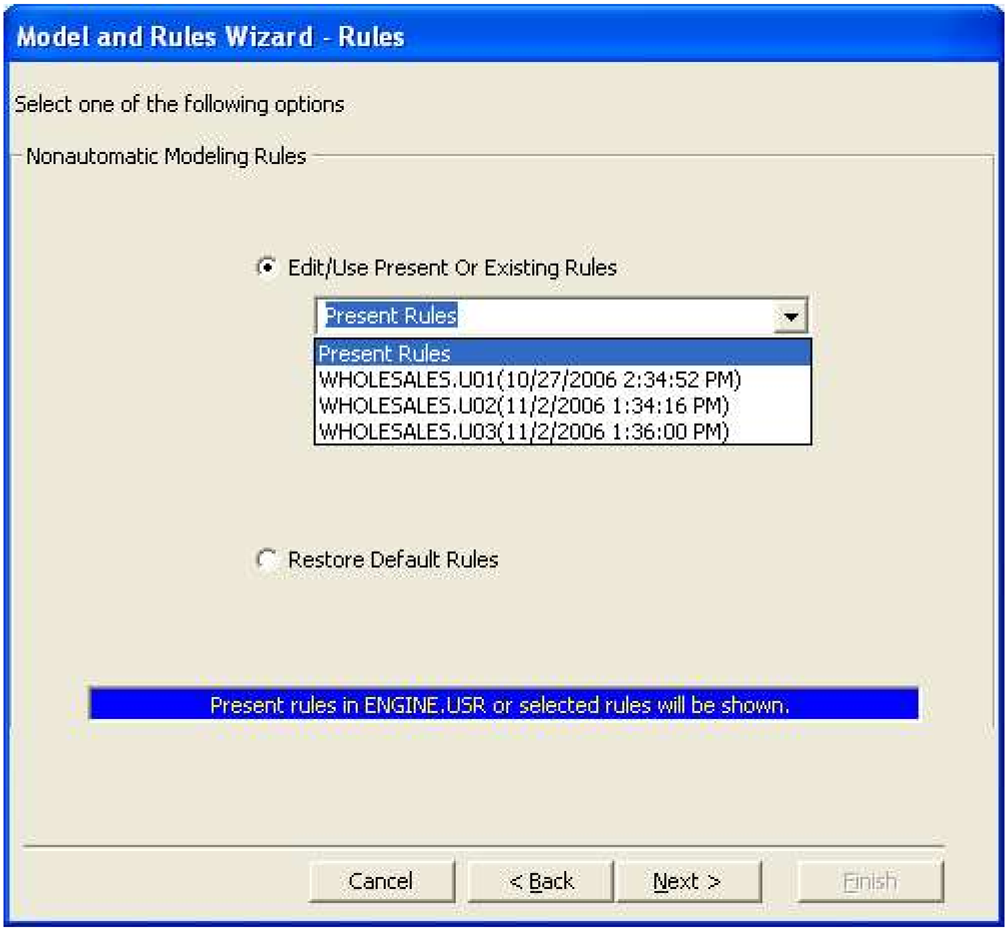


Fig. 5-17 – Option 4, Setting Rules

Press “Next” and the rules will be shown as follows for your review and/or edit:

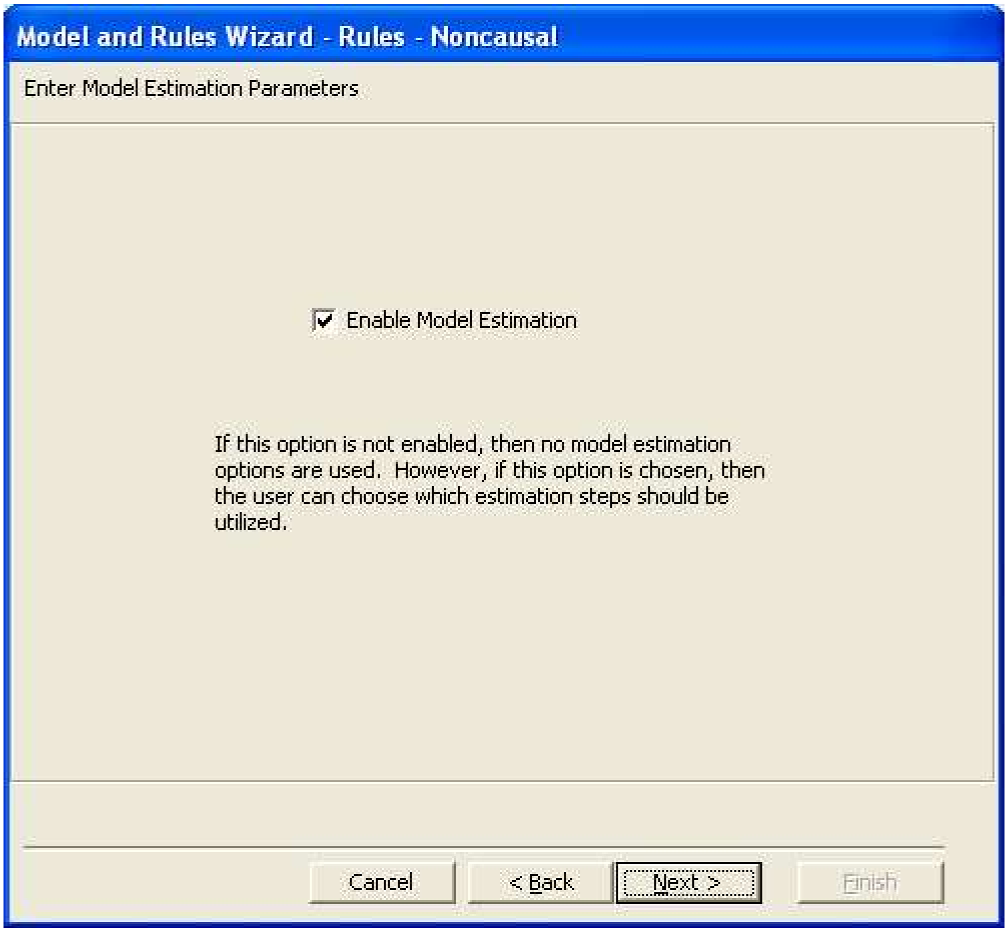


Fig. 5-18 – Option 4, Enabling Model Estimation

Estimation and diagnostic checking represent the second phase of the Box-Jenkins modeling procedure. The estimation option computes the model coefficients via nonlinear least squares and generates residuals. Various statistics are computed for both the estimated parameters and the residuals from the model.

Press “Next” and the “Model Forecasting” parameters will be shown as follows:

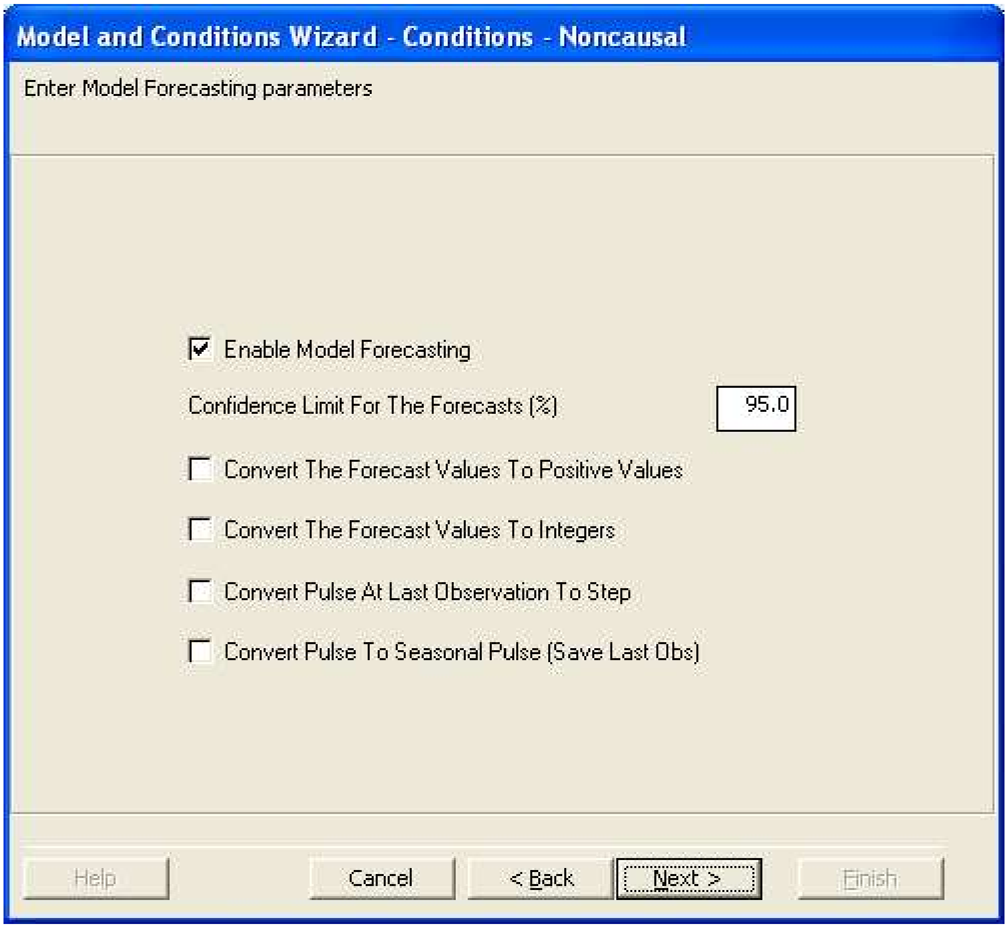


Fig. 5-19 – Option 4, Model Forecasting Parameters

This is the process of using the model that best describes the past observations (to extrapolate the pattern of the data) to predict Confidence Levels into the future.

If this option is not enabled, then no forecasts will be made.

Press "Next" and you may elect to save changed conditions and/or the "Initial Model" and/or the

"Final Model" as shown in the following"

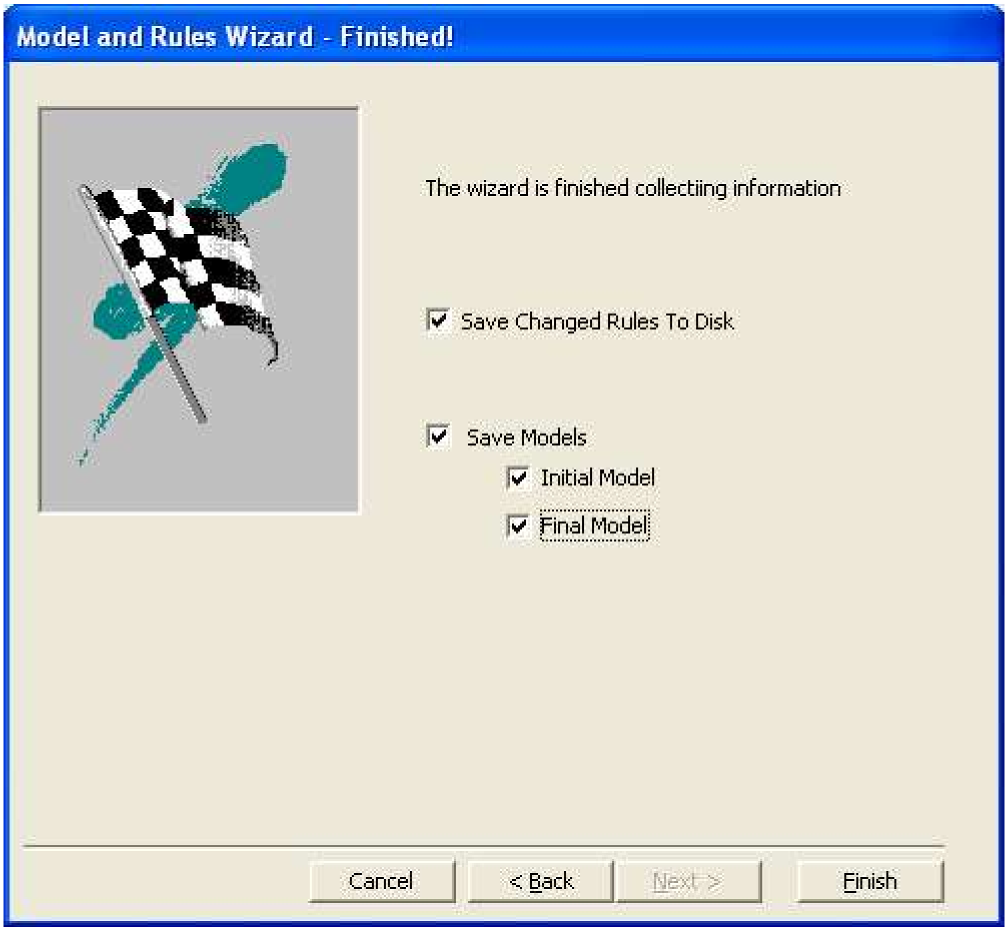


Fig. 5-20 – Option 4, Saving the Model

Again, press "Finish" to process the dataset.

If you select option5, press "Next" then "Finish" to process the dataset. Autobox will only perform identification. This is good for beginners just trying to learn Box-Jenkins to see if they can identify the model on their own.

Chapter 6

Case Study with Causals– Some Issues that May Arise from User Error

We have experienced clients who make mistakes when trying to perform analysis. The goal of this exercise is: to show you how to import the data and to get Autobox to properly do what you want it to do.

For example, we get clients who send us data that have no data (see the variable X1 on the next page) for the first half of the dataset as they did not have this data and then last half is real data. This is incorrect as if you use all of the data then you are sure to have a variable that won't be found to be significant as your data preparation was messy. So, let's walk you through our example to show you how to do analysis correctly.

Let's import the file the client sent us (with errors and all!- AFS20.XLS…you can find it in your installation directory if you want to follow along).

We will first discuss the errors and then import it into Autobox and show you how to fix them!

Note that the last 4 observations have the Y withheld to test the accuracy of Autobox.

A	B	C	D	E	F	G	H	I	J	K	L	M
	Y	X1	X2	X3	X4	X5	X6	M1	M2	M3	M4	M5
Y1M01	184616		6.3	110.32	0	121.69	163.47	1	0	0	0	0
Y1M02	181868		8.89	175.95	0	131.56	223.04	0	1	0	0	0
Y1M03	171800		11.26	17.2	0	129.11	65.83	0	0	1	0	0
Y1M04	152178		8.58	36.21	0	142.39	27.38	0	0	0	1	0
Y1M05	167492		5.15	48.08	0	138.8	31.32	0	0	0	0	1
Y1M06	120123		8.87	42.67	0	131.85	15.69	0	0	0	0	0
Y1M07	164718		6.3	82.65	0	154.07	11.31	0	0	0	0	0
Y1M08	156417		8.89	46.21	0	179.84	14.18	0	0	0	0	0
Y1M09	108093		11.26	0	0	202.45	0	0	0	0	0	0
Y1M10	152192		3.62	118.28	0	193.68	91.4	0	0	0	0	0
Y1M11	121089		0	161.32	0	190.07	115.48	0	0	0	0	0
Y1M12	156689		0	131.01	0	191.12	68.68	0	0	0	0	0
Y2M01	157451		0	27.38		220.38	110.32	1	0	0	0	0
Y2M02	162644		0	31.32		165.84	175.95	0	1	0	0	0
Y2M03	160602		0	15.69		173.97	17.2	0	0	1	0	0
Y2M04	249942		4.09	11.31		153.23	36.21	0	0	0	1	0
Y2M05	154513		12.37	14.18		158.57	48.08	0	0	0	0	1
Y2M06	159142		7.47	0		150.93	42.67	0	0	0	0	0
Y2M07	159728		6.63	91.4		145.34	82.65	0	0	0	0	0
Y2M08	184616		13.59	115.48		161.17	46.21	0	0	0	0	0
Y2M09	181868		8.48	68.68		175.33	0	0	0	0	0	0
Y2M10	171800		10.06	110.32		167.17	118.28	0	0	0	0	0
Y2M11	152178		14.46	175.95		191.29	161.32	0	0	0	0	0
Y2M12	167492		9.8	17.2		194.38	131.01	0	0	0	0	0
Y3M01	120123		14.24	36.21		186.17	146.54	1	0	0	0	0
Y3M02	164718		9.27	48.08		184.23	202	0	1	0	0	0
Y3M03	156417		9.17	42.67		201.48	53.94	0	0	1	0	0
Y3M04	108093		8.93	82.65		165.84	44	0	0	0	1	0
Y3M05	152192		10.13	46.21		173.97	0	0	0	0	0	1
Y3M06	121089		8.44	0		153.23	8.27	0	0	0	0	0
Y3M07	156689		0	118.28		158.57	11.2	0	0	0	0	0
Y3M08	157451		0	161.32		150.93	8.22	0	0	0	0	0
Y3M09	162644		0	131.01		145.34	7.25	0	0	0	0	0
Y3M10	160602		0	146.54		161.17	0	0	0	0	0	0
Y3M11	249942		0	202		165.84	8.2	0	0	0	0	0
Y3M12	154513		0	53.94		173.97	0	0	0	0	0	0
Y4M01	159142	5.99	0	32.55		153.23	10.37	1	0	0	0	0
Y4M02	159728	6.06	0.58	8.63		158.57	12.67	0	1	0	0	0
Y4M03	136379	6.25	0.72	10.18		150.93	6.41	0	0	1	0	0

Fig. 6-0 – An Example of Faulty Data

Problems:

There are 80 months of data (we show only an excerpt here). The causal variable X1 has missing data. All of the historical data from Y1M01 to Y3M12 should be dropped.

Seasonal dummies (11 of them for each month of the year) are delivered (see series M1 through M11). If you run Autobox and only a few seasonal dummies are in the final model, then maybe you shouldn't have used seasonal dummies. This is an example of "MSB" Model Specification Bias where the user makes assumptions that may be incorrect. So, if you don't include seasonal dummies then Autobox might include an AR12 variable. The point is that you should run it with and without and compare the "Adjusted Variance" reported by Autobox to determine which is the better model.

X4 is a linearly redundant variable as it is all zeros and should be removed.

Y is large and might need to be scaled by dividing by 100,000. (We tested this out and did not show it, but it after scaling returned the same results). This can be important as estimation of the model is based on % change and large numbers return small % changes.

Solutions:

Let's import the file. Just click "Next".

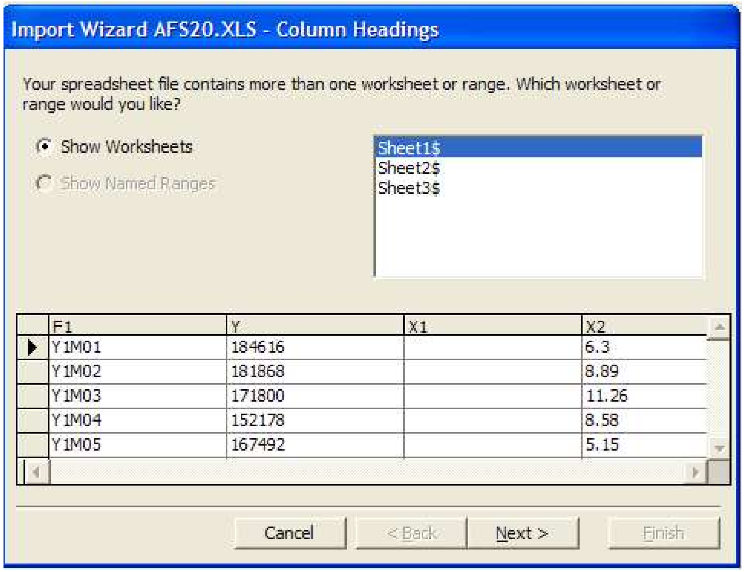


Fig. 6-1

The first row does have column headings so just click "Next".

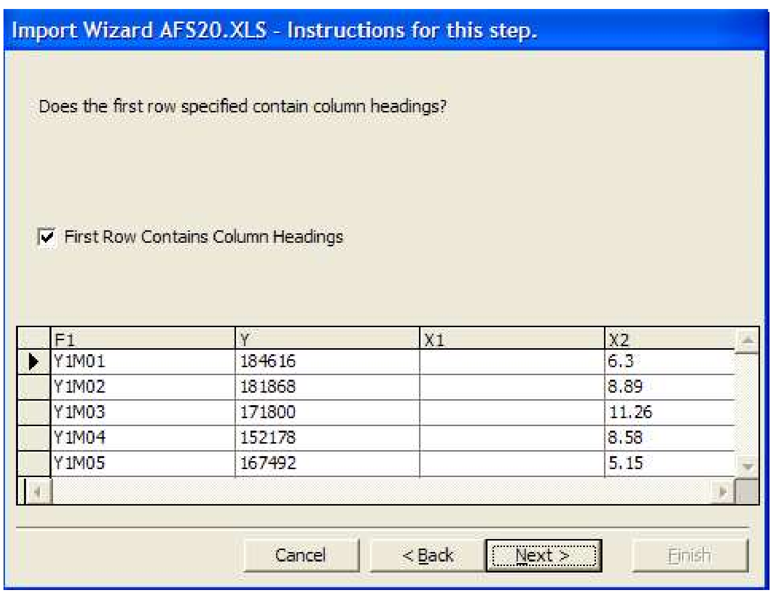


Fig. 6-2

The first column was identified by Autobox as a non-number field and highlighted it. We also clicked on X4(as it is all zeros) and clicked on "Do not import" and then "Next".

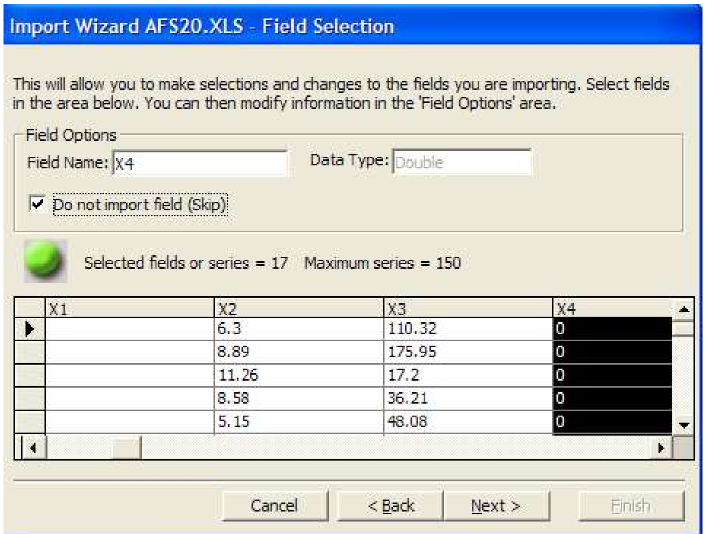


Fig. 6-2 – Do Not Import Empty Field

This shows you the proof that certain fields will not be imported. Click "Next" and also "Next" on the following screen (not shown here)

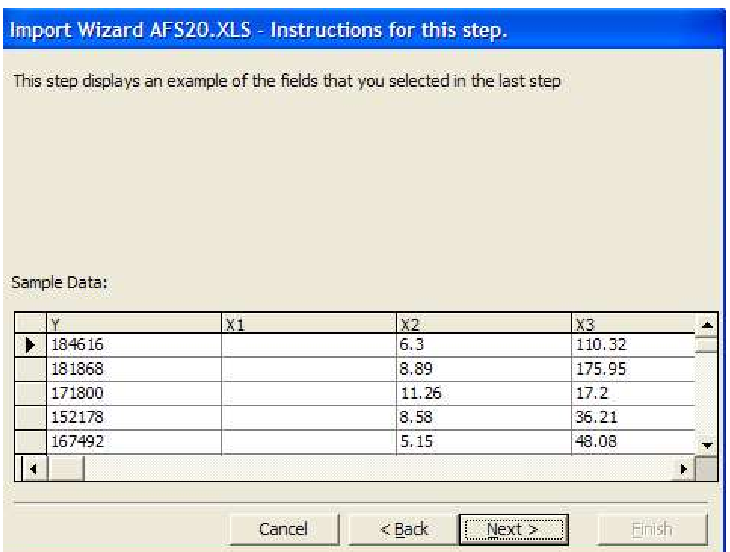


Fig. 6-3 – Submitting Fields

We select Y as the output series to be forecasted and change the number of forecasts to be 4 and click “Next”.

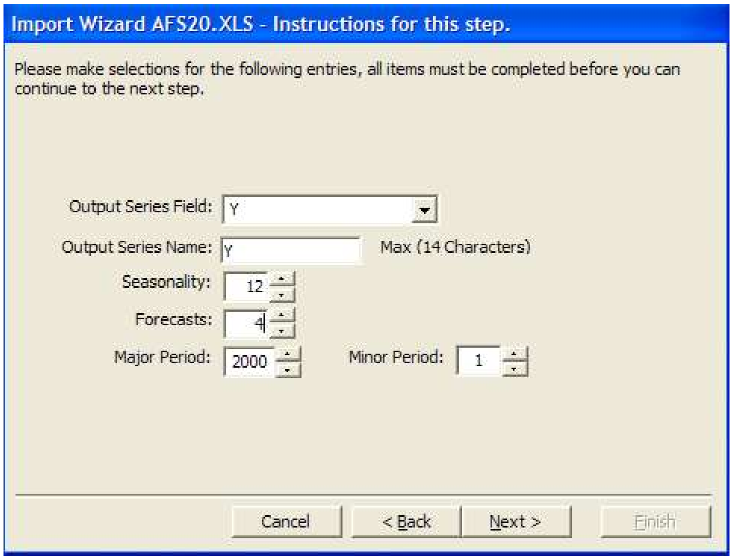


Fig. 6-4 – Updating Properties

Here is a summary showing what has been imported. Click “Finish”.

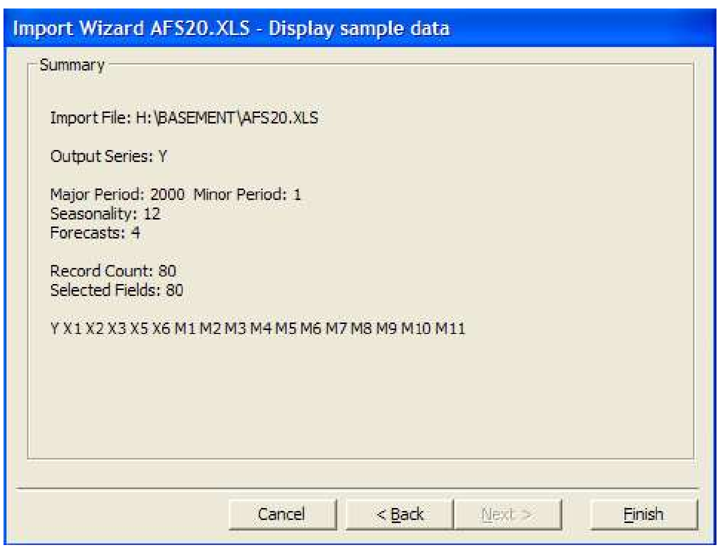


Fig. 6-5 – Summary of Import

Save the file in anywhere on your computer, but it is best to keep in the "Autobox" installation directory or a subfolder. We named it "readin.asc" and click "Save".

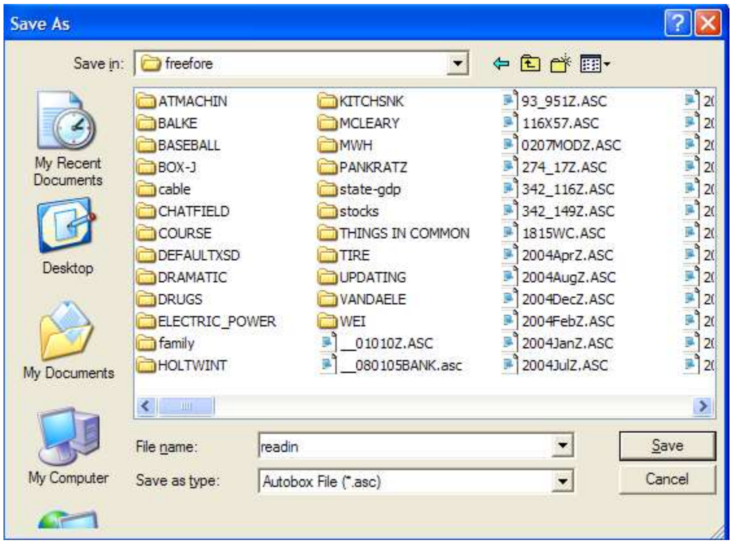


Fig. 6-6 – Save the File

We can change the observations to 44 to get rid of the first 36 observations due to the problem X1 had with missing data. (We could have deleted the column and ran the analysis without that causal variable which would be a thorough thing to do and we recommend doing just that on your own time!) We will change 80 to 44 so do that and click "Apply" and click on "Delete series from beginning of the series"(not shown below) which will shrink our dataset to 44 observations.

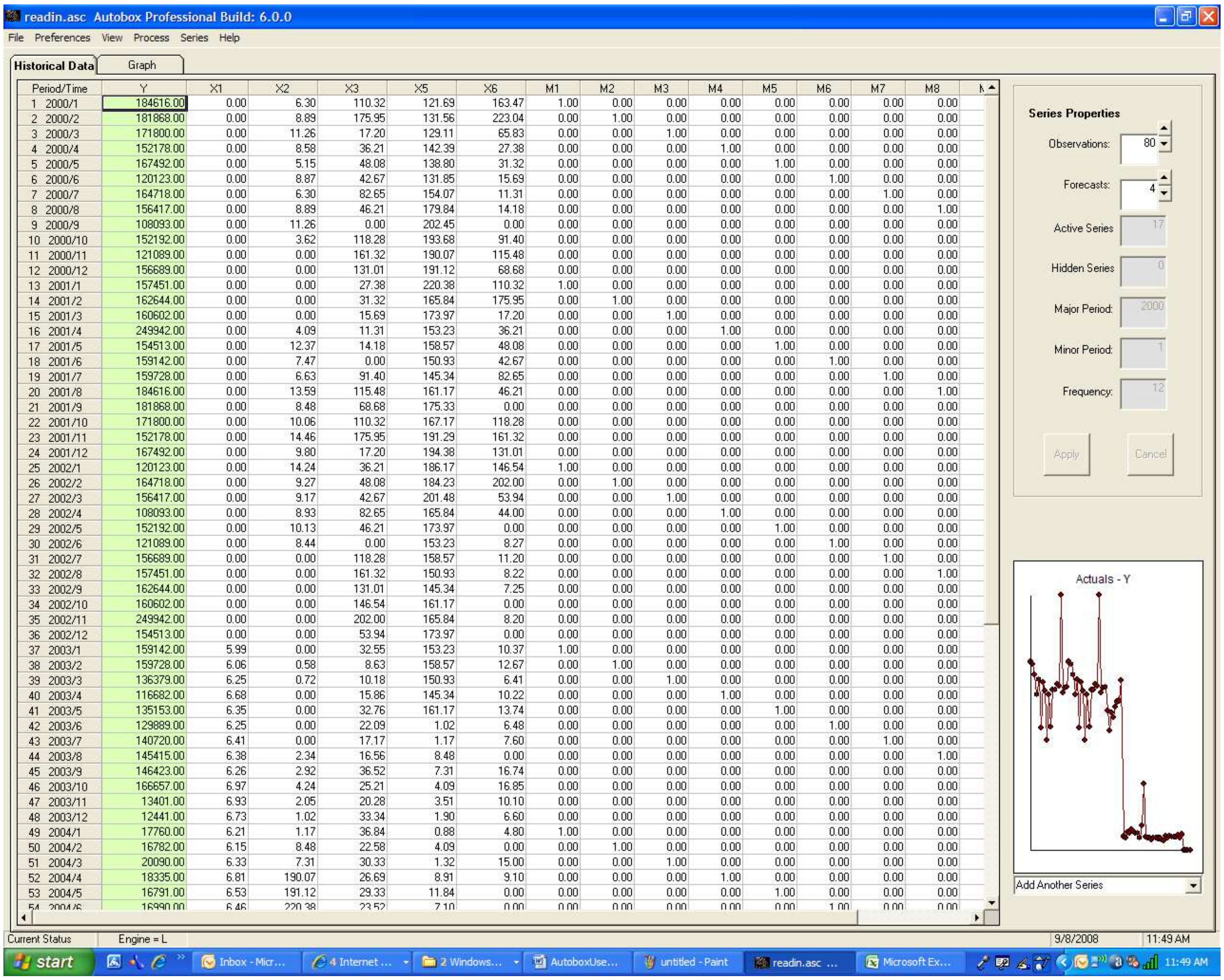


Fig. 6-7 – Data Set

Change the observations now from 44 to 40 so that you can withhold 4 observations for evaluation. Click "apply" and then choose "Delete observations at the end of the series" and click "Next" and then choose "Use the Deleted Observations as "Retained Data". The Number of Forecasts will be changed to Equal the number of retained data." and click "Next". Click "Ok" and then select all of the series to use for retained. Choose "select all" and then "ok".

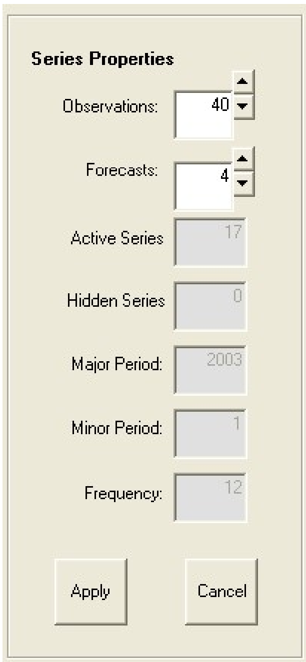


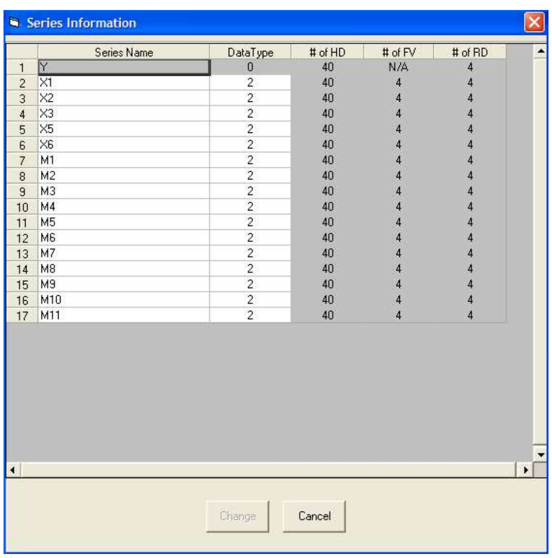
Fig. 6-8

In-Depth Look at Data Types

When you import data from excel you are in essence asking Autobox to determine the data type wheras when you create an Autobox ASC file you are telling Autobox how your data should be treated up front. Autobox needs to default the data type to something when it goes through the import. This default may not be the best case for your situation. If you have no future values, the default data type is set to ‘0’ which means Autobox will project the causal variable into the future to be used in the forecast process. If you have future values for causals then the default is set to ‘2’.

Let’s explain why this is important as there are two cases where you would want to override the default. If you have future values that you would like to provide **<u>or</u>** if you have a causal variable like 11 seasonal dummies (for the 12 months) it makes no sense for the data type to be a ‘2’ as with a ‘2’ Autobox will look for a lag relationship and there would be no lag relationship that could possibly exist for a given seasonal dummy as the next month’s dummy variable would control for that. For a situation where you have future values of the causal variable, you might know that there logically could be a lead effect so you would want to change it to ‘3’.

To change the default you need to have retained data. In this example, we have 4 future values of the causal variables. We created them on the previous page which now allows us to **now change them by choosing “ Series/Series Information” and change all of the seasonal dummies from a ‘2’ to a ‘1’.**



	Series Name	DataType	# of HD	# of FV	# of RD
1	Y	0	40	N/A	4
2	X1	2	40	4	4
3	X2	2	40	4	4
4	X3	2	40	4	4
5	X5	2	40	4	4
6	X6	2	40	4	4
7	M1	2	40	4	4
8	M2	2	40	4	4
9	M3	2	40	4	4
10	M4	2	40	4	4
11	M5	2	40	4	4
12	M6	2	40	4	4
13	M7	2	40	4	4
14	M8	2	40	4	4
15	M9	2	40	4	4
16	M10	2	40	4	4
17	M11	2	40	4	4

Fig. 6-9 - Change Series In formation

Now let's run Autobox and see what happens.

How interesting. Some of the seasonal dummies were not significant. The adjusted variance (otherwise known as the Mean Square Error) found in the stat.htm report was .001881.

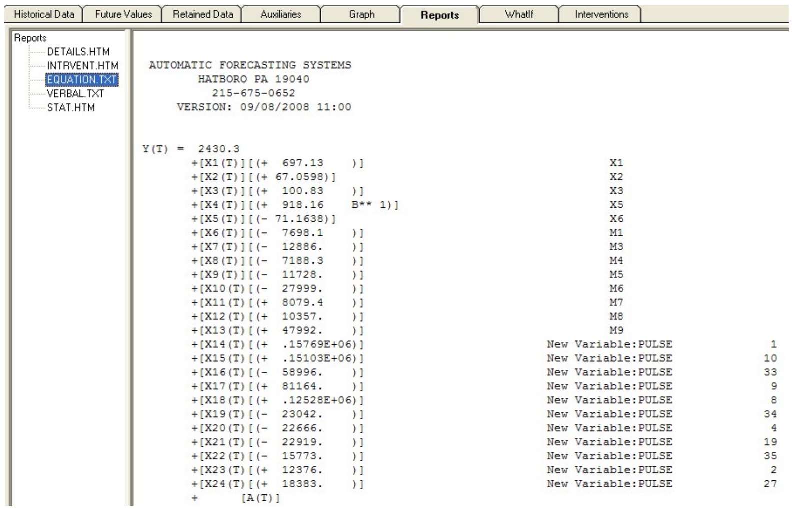


Fig. 6-10 – Checking for Trends

If we run the same problem, but remove the seasonal dummies we only get one indication of any seasonal bump at period 9 and onward. The causal variables are all now **not** significant (except X5), but we do have an autoregressive relationship of 1 period lag. The adjusted variance is .001623 which is better by 14% than the seasonal dummy model.

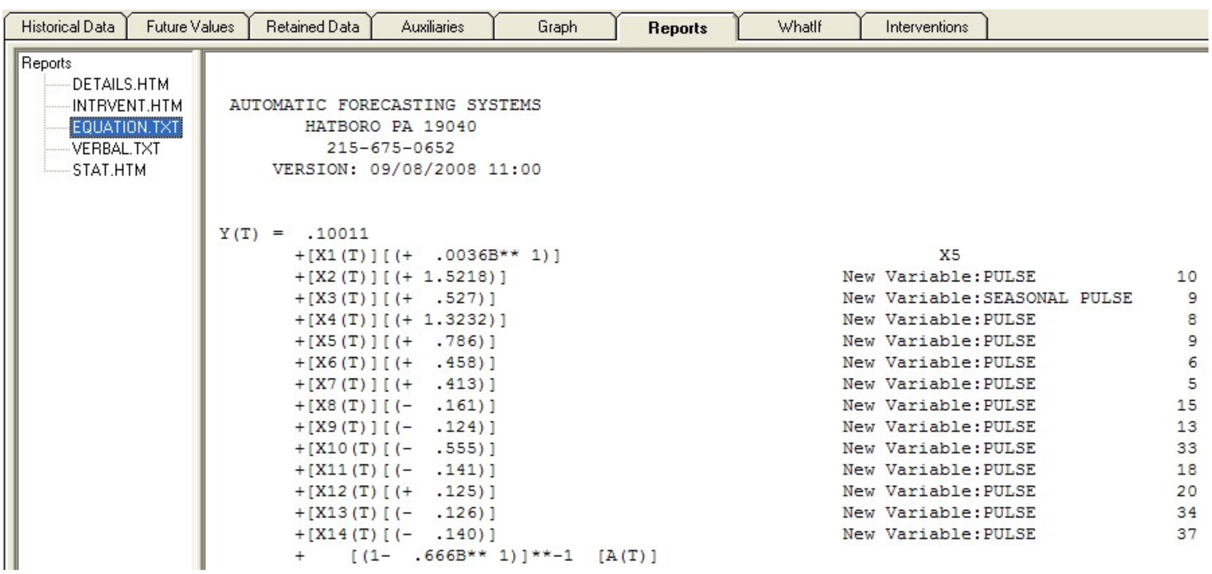


Fig. 6-11 – Checking for Trends without Dummies

Chapter 7

Pooled Time Series Cross-Sectional with Autobox

This is for those interested in comparing groups of data to see if the subsets are different than the whole. You may find that the states are different than the national level. Let's assume we have 10 states and the national level data in our example. We want to know which states have higher mortality, but first we need to first follow these steps:

Summary of Steps:

1)Create the file cleansed.afs in the installation directory

2)Run each of the states using option '1'. Autobox saves cleansed datasets free of outliers with the name "original name"-CLE.asc

4)Delete cleansed.afs and now create stepupde.afs with a zero in that file. So no outliers will be identified to keep this process workable.

5)Run all ten states using option '1' using the cleansed datasets

6)Delete stepupde.afs

7)Review each of the ten models built to try and identify common variables among the different groups

8)Run all 10 cleansed states using option '4' disabling all stepup and all stepdown options, but now only using the same model for all ten states

9)Create 1 giant series(ie asc file) using the cleansed the history and "stacking" the ten datasets. See pooled.asc in the installation directory and the estimate same model that was just run in step 8 by running using option '4' disabling all stepup and all stepdown options (see Fig 7.0 with the stepup/stepdown menus)

10)Compare Sum of Squares to get an F-test. Open the output file(s) "Stat.htm" and enter the "Sum of Squares" for each of the state models into Excel. Sum all the Sum of Squares and let's call that SOSA.

You then run Autobox using option '4' and specify your own model. You need to review the 10 models and establish a common model which is a superset of the 10 models (no interventions in the model). You then choose "Yes" on the "Pooled Cross-Sectional Time Series" option on the "Multiple Groups/Panels" screen.

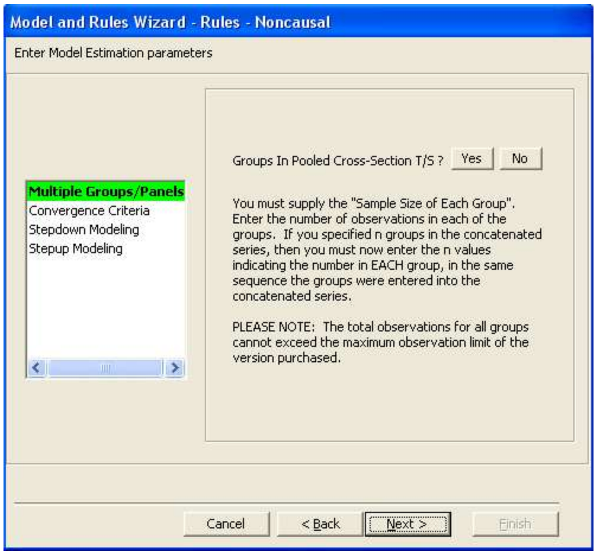


Fig. 7-0 – Cross-Section Series

You would specify the 5 states by typing a '1' here in each row.

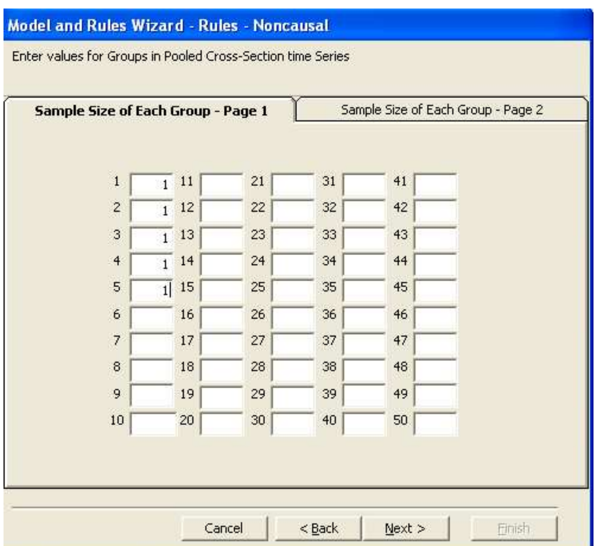


Fig. 7-1 – Sample Size

You would run Autobox option '1' time on an ASC file that has all of the states concatenated in one ASC file. Make sure to disable all 'step up' conditions as want to keep our model intact. Note the Error Sum of Squares for the national model. Let's call this SOSB.

Take the difference of SOSA and SOSB to get SOSC. Divide SOSC by the number of groups(5 in our case) to get MS1 (Mean Square Error). MS2 is calculated by taking SOSA divided by N*K – K*NUMPAR to get MS2 where N = observations in a group and K = # of groups NUMPAR = number of parameters in a model). The f value is then calculated from MS1/MS2.

Intermittent Demand

Autobox approaches the case of Intermittent or Sparse Demand by making the problem into a causal problem. Autobox tracks the count of how many periods were zero and creates an "interval" time series. It takes the volume and divides it by the interval to get a "rate". It projects out the interval and rate into the future and divides the two to get a forecast. Autobox will do all of this while still searching for pulses and level shifts. You need to have >25% zeroes in the series and 4 intervals to trigger intermittent demand. With causals, you need only 25% as the demand could be caused by a promotion and you wouldn't want to ignore the causals.

You can override the % by creating a file in the installation directory named sparse.afs with .70 in the file to reflect a threshold of >70%.

You can also suppress intermittent demand by creating a file called "nointerm.afs" in the Autobox folder.

What are the Odds of Making the Month End Number?

Everyone wants to know as quickly as possible if the month end number is in jeopardy. If you have daily data and the number of forecasts is equal to the remaining days in the month, Autobox will report a file with a "CAN" extension that gives various probabilities to assess the likelihood of making various month end sums. You can easily interpolate to locate your goal and locate the probability. Make sure you use afs2sip.afs when using this analysis(search on afs2sip.afs).

Since days of the week vary, you can't use a simple ratio estimate to forecast the month end. You need to use sophisticated tools like Autobox to account for trend, day of the week variations, etc to get a robust forecast. The file canwe.asc comes with the installation will create canwe.csv:

PROBABILITY OF	EXCEEDING TARGET __020910a	
100	7617.881	__020910a
100	7657.477	__020910a
100	7697.073	__020910a
100	7736.669	__020910a
100	7776.265	__020910a
100	7815.862	__020910a
100	7855.458	__020910a
99.999	7895.054	__020910a
99.999	7934.65	__020910a
99.997	7974.247	__020910a
99.994	8013.843	__020910a
99.987	8053.439	__020910a
99.974	8093.035	__020910a
99.95	8132.631	__020910a
99.905	8172.228	__020910a
99.825	8211.824	__020910a
99.687	8251.42	__020910a
99.455	8291.016	__020910a
99.077	8330.613	__020910a
98.482	8370.209	__020910a
97.574	8409.805	__020910a
96.239	8449.401	__020910a
94.345	8488.997	__020910a
91.752	8528.594	__020910a
88.336	8568.19	__020910a
84	8607.786	__020910a
78.709	8647.382	__020910a
72.499	8686.979	__020910a
65.495	8726.575	__020910a
57.903	8766.171	__020910a
57.903	8766.171	__020910a
50	8805.767	__020910a
42.097	8845.364	__020910a
34.505	8884.96	__020910a
27.501	8924.556	__020910a
21.291	8964.152	__020910a
16	9003.748	__020910a
11.664	9043.345	__020910a
8.248	9082.941	__020910a
5.655	9122.537	__020910a
3.761	9162.133	__020910a
2.426	9201.73	020910a

Fig. 7-2 – canwe.csv

The 50th percentile is all of the forecasts for the rest of the month summed up from the Autobox generated forecast.

You can locate your month end number probability on the table by using interpolation.

Let's say that the month end number is 7,770 then you can say that there is a 100% chance of making the number (barring any major intervention!)

Safety Stock

Autobox provides a safety stock calculation at a 90% service level out for as many periods as you are forecasting. Note that if there are not enough historical observations and given the seasonality and length of the forecasts this report won't be produced. The file created is called "SAFETY.TXT".

For an 8 week forecast, we compute the mean square error of the fit and then we generate forecast variances for the next 8 weeks. We then compute the covariance of the forecast variances and use them to create a forecast of the sum for each of the 8 weeks.

SERVICE LEVEL : 90 %

CYCLE	VARIANCE OF THE SUM	STD DEV OF THE SUM	SAFETY STOCK
1	0	1	1
2	1	1	1
3	4	2	3
4	5	2	3
5	5	2	3
6	9	3	4
7	10	3	4
8	13	4	5
9	13	4	5
10	19	4	5
11	20	4	5
12	20	4	5
13	23	5	6
14	23	5	6
15	24	5	6
16	26	5	6
17	28	5	6
18	29	5	6
19	29	5	6
20	34	6	8
21	34	6	8
22	36	6	8
23	36	6	8
24	37	6	8
25	40	6	8
26	40	6	8
27	44	7	9
28	45	7	9
29	48	7	9
30	48	7	9
31	55	7	9
32	55	7	9
33	56	7	9
34	59	8	10
35	59	8	10
36	60	8	10

Fig. 7-3 – Safety Stock Table

Simulation with the Delphi Method

You can use the Delphi forecasting method by following these steps:

This is a two-step process which can easily be scripted/concatenated into a single step. We talk about two steps for purposes to show you clearly what is happening. The first step is to generate 1,000 forecasts (ie a “family of forecasts”) for demand and store them in the disk as demand.fum using info in a file called delph.dat which contains uncertain expectations for the next period. The second step is to actually form a causal model and use the family of forecasts for demand (from demand.fum which was generated in the first step) to generate an xml file containing a family of forecasts for the output/target series for the next period.

The expected forecasts/probability ranges for the predictor series from your customer could be seen in the file Delphi.dat where, for example, 6 ranges are provided with the probability of that range in the first column and the ranges in column 2 and 3. We will copy this to Delphi.afs as that is the special name of a file that causes Autobox to generate the simulated forecasts for this “delphi” approach in the first step.

Delphi.dat

```
6
.05 35 40
.05 40 45
.20 45 50
.20 50 55
.40 55 60
.10 60 65
```

Template.dat might look like this (ASC file format with the historical data below the header information).

```
0
0
0
1
1
1
0
10
1
```

0
1
DELPHI
0
35.83
36.08
37.00
39.89
38.93
35.08
36.97
32.87
37.43
51.0816

FIRST STEP:

The first step is to use a template to get the process going in order to develop simulated forecasts for the casual variable named “Demand”. We first need to copy the file template.dat to template.asc and then copy Delphi.dat to Delphi.afs to trigger the creation and the storage of pseudo-forecasts for demand in demand.fum. When we go to run Autobox, the file AFS2SIP.AFS(explained already above) will tell Autobox to run 1,000 simulations with the second row telling Autobox to ultimately let outliers play in the forecasts for the target series in step 2 and not be ignored. Now it’s time to run Autobox. 1)Run Autobox using template.asc creating demand.fum.

A file called Delphi.fum will be generated(you could create this output with a different application too!). These values will have a similar distribution as the contents of Delphi.dat as that is what was used to generate the 1,000 simulations.

We now delete the trigger file Delphi.afs now as we are further along the process and already have our “FUM” file which will now be used.

SECOND STEP:

If you look at Delphi.asc(header information at the top) you will see that there is a causal called “Demand” with a 1 period out forecast. It has 60 months of historical data with a known future value of the causal named “Demand”.

Excerpt of Delphi.asc

0
0
0

```
2
1
1
0
        60
1
0
1
DELPHI
DEMAND
1
0
66.16
63.57
72.25
.
.
.
.
```

Use the Delphi.asc as an example file and note that the future value of the causal variable is just a placeholder and will not be used as the "FUM" file. Run Autobox again using Delphi.asc

You will now have a details.htm and an XML file generated with simulated forecasts of the Y variable(ie demand).

As you might know, statistical software will use the clock to get a random number. You can create a file called "seed.afs" and put any number in it if you want to reproduce the same exact results which is really just for needing to reproduce results (ie for a journal article for example).

FYI: If you have predictor series that are stochastic(data type 0) the system will automatically generate fum files and use them.

Test Data That Comes With Autobox

We provide some classic data sets found and analyzed time and time over. We also provide some datasets that we have found to have caused us to rethink the way we approach time series analysis. Each of these sets of data has its own Directory under the Autobox installation directory. We did not describe every folder and problem set as there are >700 series included. Here is a sample:

\BOX-J

Box, G.E.P. and Jenkins, G.M. (1976). Time Series Analysis:

Forecasting and Control, 2nd ed., San Francisco: Holden Day.

----

bj01 Chemical Process Concentration

bj02 IBM Stock Prices 5/17/62-11/2/62

bj03 IBM Stock Prices 6/29/59-6/30/60

bj04 Chemical Process Temperatures

bj05 Chemical Process Viscosity

bj06 Wolfer Sunspot Numbers

bj07 Batch Chemical Process

bj08 International Airline Passengers

bj09 Methane Gas Input Feed

bj10 Carbon Dioxide Gas Output

bj11 Coded Dynamic Output

bj12 Coded Input

bj13 Coded Gas

bj14 Leading Indicator

bj15 Sales

\Mcleary

McLeary, R. and Hay, R. (1980). Applied Time Series Analysis for the

Social Sciences, Los Angeles: Sage.

Troubleshooting

If you get an error or are very dissatisfied with the forecast, model, etc.   We would like to receive an email with a zip file with the following:

Data file ("*.ASC")
*.AFS files
Autobox.exe, freefore.dll, freel.dll
 ENGINE.*,
Create a file named 'snoop.afs' and 'apush.afs' and run Autobox and save all files that were created in the Autobox directory during the process of running the ASC file (You can use Windows Explorer and sort the directory to see these newly created files). Open a Word document and take print screens of each of your selections up so we can see exactly what you did. Upon receipt of your email, we will review and respond. Make sure you delete snoop.afs and apush.afs as they are diagnostic files!

IF AUTOBOX STOPS WORKING (YOU CAN TELL THIS BY VIEWING DETAILS.HTM ON YOUR FREEFORE DIRECTORY THEN AUTOBOX IS NOT FUNCTIONING PROPERLY) AND YOU NEED TO STOP IT, DON'T TURN OFF YOUR COMPUTER AND DO FOLLOW THESE DIRECTIONS

HIT CTRL-ALT-DEL AT THE SAME TIME
CHOOSE "TASK MANAGER"
CHOOSE "PROCESSES"
SORT ON IMAGE NAME BY CLICKING ON IMAGE NAME
SELECT AFSENGINE.EXE AND CLICK ON "END PROCESS"
SELECT AFSLITE.EXE AND CLICK ON "END PROCESS"
CHOOSE "APPLICATIONS"
SELECT "AUTOBOX" AND CHOOSE "END TASK"

Contact us for any questions:

Afs Inc.

P.o. Box 563

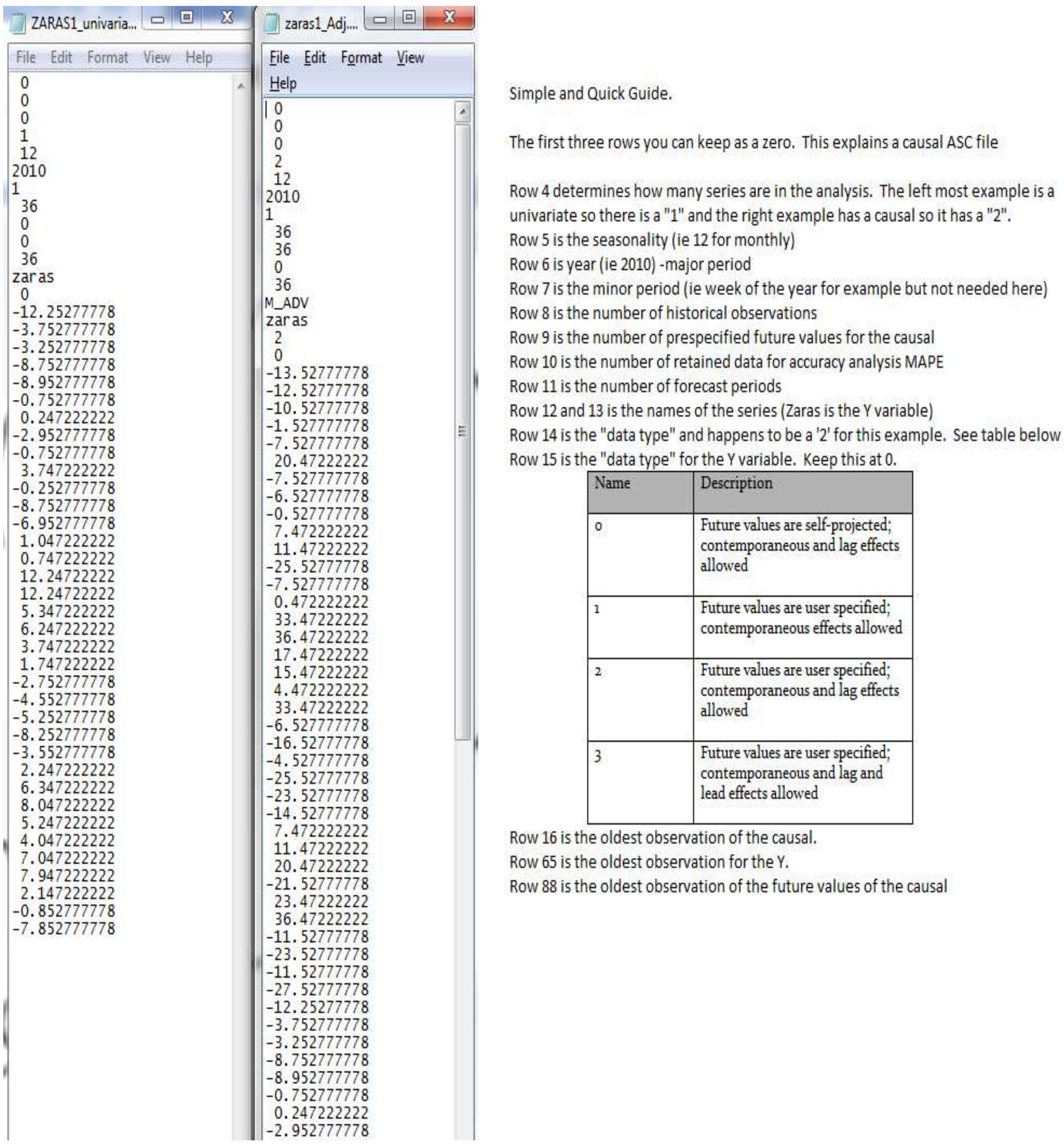
Hatboro PA 19040

sales@autobox.com

Phone 215-675-0652   Fax 215-672-2534

Create Your Own .ASC File

Instead of entering your own data or cutting and pasting you can create a file that can be read directly into Autobox. You can open any ASC file and you will quickly see that the first ~13 rows are "header information" that Autobox needs to know before running. This is really an easy process, but the information and data must be entered in a text file in a very specific order in a single column. Here's a quick way to figure out how to do this!


ZARAS1_univaria...

File Edit Format View Help

```
0
0
0
1
12
2010
1
36
0
0
36
zaras
0
-12.25277778
-3.752777778
-3.252777778
-8.752777778
-8.952777778
-0.752777778
0.247222222
-2.952777778
-0.752777778
3.747222222
-0.252777778
-8.752777778
-6.952777778
1.047222222
0.747222222
12.24722222
12.24722222
5.347222222
6.247222222
3.747222222
1.747222222
-2.752777778
-4.552777778
-5.252777778
-8.252777778
-3.552777778
2.247222222
6.347222222
8.047222222
5.247222222
4.047222222
7.047222222
7.947222222
2.147222222
-0.852777778
-7.852777778
```

zaras1_Adj....

File Edit Format View Help

```
0
0
0
2
12
2010
1
36
36
0
36
M_ADV
zaras
2
0
-13.52777778
-12.52777778
-10.52777778
-1.527777778
-7.527777778
20.47222222
-7.527777778
-6.527777778
-0.527777778
7.472222222
11.47222222
-25.52777778
-7.527777778
0.472222222
33.47222222
36.47222222
17.47222222
15.47222222
4.472222222
33.47222222
-6.527777778
-16.52777778
-4.527777778
-25.52777778
-23.52777778
-14.52777778
7.472222222
11.47222222
20.47222222
-21.52777778
23.47222222
36.47222222
-11.52777778
-23.52777778
-11.52777778
-27.52777778
-12.25277778
-3.752777778
-3.252777778
-8.752777778
-8.952777778
-0.752777778
0.247222222
-2.952777778
```

Simple and Quick Guide.

The first three rows you can keep as a zero. This explains a causal ASC file

Row 4 determines how many series are in the analysis. The left most example is a univariate so there is a "1" and the right example has a causal so it has a "2".
Row 5 is the seasonality (ie 12 for monthly)
Row 6 is year (ie 2010) -major period
Row 7 is the minor period (ie week of the year for example but not needed here)
Row 8 is the number of historical observations
Row 9 is the number of prespecified future values for the causal
Row 10 is the number of retained data for accuracy analysis MAPE
Row 11 is the number of forecast periods
Row 12 and 13 is the names of the series (Zaras is the Y variable)
Row 14 is the "data type" and happens to be a '2' for this example. See table below
Row 15 is the "data type" for the Y variable. Keep this at 0.

Name	Description
0	Future values are self-projected; contemporaneous and lag effects allowed
1	Future values are user specified; contemporaneous effects allowed
2	Future values are user specified; contemporaneous and lag effects allowed
3	Future values are user specified; contemporaneous and lag and lead effects allowed

Row 16 is the oldest observation of the causal.
Row 65 is the oldest observation for the Y.
Row 88 is the oldest observation of the future values of the causal

Header Ordering

- Objectives (all are required)
- Data properties (all are required)
- Data names (in the order of 1st input series to nth input series, if any; and then the output series)
- Data type (in the same order as the data names)
- Historical Data (in the same order as the data names)
- Future Values (for all input series which have a data type of 1, 2, or 3, if any, in the same order as data names)
- Retained Data (If any, for all series in the same order as the data names)

The following structure tables indicate the parameters and/or limitations for each of the above categories.

DATAPROP Structure:

Name	Description
OBJECTIVE(1)	Sets forth the model conditions as indicated by the following: 0 = Totally Automatic NONCAUSAL MODELS IN AUTOBOX MEMORY 1 = MEAN 2 = AUTOREGRESSIVE (1) WITH CONSTANT 3 = AUTOREGRESSIVE(2) WITH CONSTANT 4 = SIMPLE EXPONENTIAL SMOOTHING NO CONSTANT 5 = LINEAR (HOLT) EXPONENTIAL SMOOTHING NO CONSTANT 6 = RANDOM WALK NO CONSTANT 7 = RANDOM WALK WITH CONSTANT 8 = TIME TREND 9 = TIME TREND PLUS AR(1) CORRECTION 10 = FOURIER

	11 = HOLT LINEAR TREND PLUS ADDITIVE SEASONAL FACTORS (TREND FORM) 12 = DAMPED TREND LINEAR EXPONENTIAL SMOOTHING NO CONSTANT 13= SEASONAL EXPONENTIAL SMOOTHING NO CONSTANT 14 = HOLT LINEAR TREND PLUS ADDITIVE SEASONAL FACTORS (ARIMA FORM) 15 = AIRLINE MODEL 16= FOUR PERIOD MOVING AVG 17= DAMPING MODEL AR(1) AND MA(1) NO CONSTANT 18= DAMPING MODEL AR(1) AND MA(1) WITH CONSTANT CAUSAL MODELS IN AUTOBOX MEMORY 51 = REGRESSION 52 = REGRESSION WITH AR(1) CORRECTION COMMON 97 = IDENTIFICATION ONLY 99 = STARTMOD.123 199 = STARTMOD.123 + SIM 200 = Totally Automatic + ABOXLITE model is developed
OBJECTIVE (2)	Sets source of conditions for processing: 0 = Use default conditions in memory
OBJECTIVE (3)	0 = save reports

Table 3-12 – List of Objectives

DATAPROP Structure:

Name	Description
DATAPROP(1)	Number of series in the problem
DATAPROP(2)	Seasonality. How often the data was sampled. (i.e. 52 for weekly data) Please note that all series in the model must have the same seasonality
DATAPROP(3)	Beginning year. The year or major number identifying the starting point of the data. Please note that all series in the model must have the same Beginning Year. If you wish to use series whose original Beginning Year are different, you must determine the common matrix for the series and use that starting point as the Beginning Year.
DATAPROP(4)	Beginning period. the starting point of the data. (i.e. 1 for the 1st week in the year) Please note that all series in the model must have the same Beginning Period. If you wish to use series whose original Beginning Period are different, you must determine the common matrix for the series and use that starting point as the Beginning Period.
DATAPROP(5)	Number of historical values in each of the time series in the model
DATAPROP(6)	Number of future values to be included for each applicable input series.If a causal model (includes a dependent and independent series) and the DATATYPE of any the input(independent) series is 1 , 2, or 3, enter DATAPROP(7) + Number of Future Values (this must equal the number of forecasts to be calculated) to be supplied by the user. If DATATYPE of all input series is 0, or if a noncausal model, this must show a 0.
DATAPROP(7)	The number of values retained from the end of the series to be used to evaluate prior forecasts (enter 0 if none)
DATAPROP(8)	Number of forecast values to be calculated

Table 3-13 - Data Properties

There should be a row for each series with the name.

DATANAME Structure:

Name	Description
DATANAME	Actual name of each series in model in the order 1st Input series, 2d input series, ...N input series, output series These names must be limited to 22 characters for Input series and 14 characters for the output series; and they cannot contain space(s), period(.), exclamation point(!), backquote(‘), brackets([]), wild card characters such as * or ?, and control characters(ASCII values 0 through 31).

Table 3-14 - DATANAME Structure

There should be a row for each series with a 0,1,2,3 or 4.

DATATYPE Structure:

Name	Description
0	Future values are self-projected; contemporaneous and lag effects allowed
1	Future values are user specified; contemporaneous effects allowed
2	Future values are user specified; contemporaneous and lag effects allowed
3	Future values are user specified; contemporaneous and lag and lead effects allowed

Table 3-15 - DATATYPE Structure

Header Ordering

The following is an example of an .ASC file for a noncausal (single) series{annotations are not included in the file}:

0	(objective(1) indicates totally automatic modeling)
0	(objective(2) indicates use default rules in memory)
0	(objective(3) indicates full output)
1	(DataProp(1) number of series in the problem set)
52	(DataProp(2) seasonality of the series)
1998	(DataProp(3) beginning year or major period)
2	(DataProp(4) beginnng or minor period)
67	(Dataprop(5) number of historical data in series)
0	(DataProp(6) number of future values)
0	(DataProp(7) number of retained data
24	(DataProp(8) number of forecasts to be calculated
pizza	(output series name – causal variables would be added before this line, Also, by specifying the date (January 1, 2006) with the name in this format "__010106Y11" and line 5 has a '7' then U.S. holidays are automatically generated, day of the month effect is analyzed and the effect of "Friday before" and "Monday after" a weekend holiday is analyzed, search for "end of the month effect")
0	(Data type)
15	Historical Data
14	Historical Data
6	Historical Data
.	Historical Data
12	Historical Data

We recommend that you open an ASC file that already exists (such as 116x573.asc in \Kitchsnk directory) as a good example of how a file should look if it is to be a casual model.

Future values would be below the historical data.

Retained future values would below that as well.

We highly recommend that you download/install our interactive version (http://www.Autobox.com/abox.exe) as a way of Quality Control on how you have built the ASC file. You can quickly establish if there are any problems by opening the ASC file that you have built. Here is an example of a Univariate ASC file

```
0
0
0
1
12
1949
1
144
0
0
12
BJ07
    0
    112.000000
    118.000000
    132.000000
    129.000000
    121.000000
    135.000000
    148.000000
    148.000000
    136.000000
    119.000000
    104.000000
    118.000000
```

These are the three "Objectives". We prefer to leave these options all at zero, but you can change these as you wish. They determine automatic modeling, how much output to show

These are the 8 "Data Properties". They define the characteristics of the series like the seasonality, beginning year, beginning period, number of observations.

This is where the name of the series is specified. This field is "variable" in that if you have a causal problem where you have many series then you would list all of the names of the series HERE BEFORE the data. YOU should not USE THE SAME SERIES NAME FOR A DEPENDENT SERIES IN TWO SEPARATE ASC FILES as the output from a modeling run will overwrite each other. Also, by specifying the date (January 1, 2006) with the name in this format "__010106Y11" and line 5 has a '7' then day of the week effect is analyzed

This is the famous Box-Jenkins airline series (abridged) with the data shown here. If you have a causal problem then each of the series are placed end-to-end downward in a Blocked rectangular historical array. The last series would be the dependent series.

BUT If you have future values, you place this future values series below the dependent series

BUT If you have retained values, you place this series below the future values series

Output series

Future

Retained values

Measuring Accuracy

This utility is for evaluating forecasting accuracy from many withheld periods. The number of forecasts in the ASC file will determine how this will all behave.

Create a file in the Autobox folder called "Rolling.afs" using "Notepad". There will be 2 rows in the file. The first row tells the system how many period back you want to withhold. The 2nd row defines the periods out between each withheld amount. If you have 23 as the number of forecasts and have a 1 in each of the two rows, one snapshot will be produced from 23 periods back.

You would typically want the second row to be a "1" so you could evaluate each period getting a "rolling accuracy" from many origins, but if you wanted to get only every other period you could put a "2" for example, but this isn't usual to do. The # of forecasts in the ASC file needs to match the second row. In this example, we used 20 and a 1. If we have 144 observations and we have 12 forecast periods, the first observation you can withhold is 132 as you can see from the file created "summape.csv" as shown below. The max is 52 for row 1 as the maximum seasonality is 52 IN Autobox so 52 is the furthest out you can run a forecast when doing this.

20

1

ITEM	NOB	NF	MAPE	ACTUAL	FORECAST
BJ07	132	12	2.75	5714	5797
BJ07	131	12	2.74	5687	5649
BJ07	130	12	3.27	5659	5540
BJ07	129	12	3.34	5605	5526
BJ07	128	12	3.6	5560	5462
BJ07	127	12	3.41	5513	5427
BJ07	126	12	4.73	5439	5218
BJ07	125	12	3.61	5376	5274
BJ07	124	12	3.29	5324	5285
BJ07	123	12	3.73	5259	5101
BJ07	122	12	5.18	5246	4974

Autobox will run many different analysis from different origins and generate a CSV file to show accuracies. You will need to delete this output file (ie summape.csv) after your analysis as it will just grow and grow. The HTM files generated on the disk are the analysis files if you want to review. This can't be used with Simulated forecasting. This can be done using causal problems, but with no searching for day of the month impacts (ie nofixday.afs) due to complications.

Chapter 7

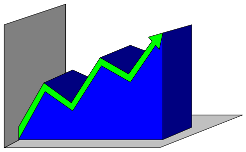
Daily Data – Using Holidays Outside of the U.S.

Autobox incorporates U.S. Holidays into the modeling and forecasting process. If you have data from outside of the U.S., create a file called “noholdum.afs” to suppress the search for U.S. in the Autobox folder.

If you want to specify a holiday on your own, make sure to start the series name with a “M_” as Autobox has a useful check of the history for all years having a zero to override the forecast to be a 0.

The TOP GDP producing countries were selected and there are 16 Countries that we have set up which are ready to be used in Autobox from 2001 out to the year 2027. If your forecast is only out to 2020, there is no need to adjust. To apply India holidays, create a file called “country.afs” and put the code as seen below (ie 5 for India) into a file called “country.afs” and save it to the Autobox folder. Autobox will use the Holiday variables from that Country. We will explain how you can customize Autobox for a different country not listed or even perhaps a local area within a country which follows a more specific holiday list than what we have included.

Code	Country
1	USA
2	BRAZIL
3	FRANCE
4	GERMANY
5	INDIA
6	ISRAEL
7	ITALY
8	JAPAN
9	MEXICO
10	RUSSIA
11	RWA
12	SOUTH AFRICA
13	SOUTH KOREA
14	SPAIN
15	SWITZERLAND
16	AUSTRALIA
17	UNITED KINGDOM



If you want to créate your own holiday file for a specific country you can use another country's holiday file as a template. For example, open the file "ROS01.HLD" found in your Autobox folder (which is the USA) and make changes to it and save it as ROS18.HLD. You can then créate a file called "country.afs" with an 18 and Autobox will use that to bring in the custom holidays.

Let's discuss the file format. We will not show you the whole file as it is too large.

The 18 in the first row tells Autobox how many holidays exist. If your country has only 12 then change this number to 12 and then save this file with a new name with a number higher than the ones listed on the page before(ie ros17.hld). Row two will have the first holiday's name and a count of 27. The 27 represents that there are 27 years of this holiday in the file. The date in the next row, 37250, would be the 37250th day since 1/1/2001(our internal date reference point). The numbers below 27250 are increments of 365 and sometimes 366. Set row 3's date to match the holiday that you want to add and set the dates below it to match. Note: You don't need to set this up like we did to go out 27 years like we did. Just remember that you would need to change 27 down to a different number.

```
AXFBJR(  1)( 1:  ) ='18                    '
AXFBJR(  2)( 1:  ) ='M_CHRISTMAS           27'
AXFBJR(  3)( 1:  ) ='37250                 '
AXFBJR(  4)( 1:  ) ='37615                 '
AXFBJR(  5)( 1:  ) ='37980                 '
AXFBJR(  6)( 1:  ) ='38346                 '
AXFBJR(  7)( 1:  ) ='38711                 '
AXFBJR(  8)( 1:  ) ='39076                 '
AXFBJR(  9)( 1:  ) ='39441                 '
AXFBJR( 10)( 1:  ) ='39807                 '
AXFBJR( 11)( 1:  ) ='40172                 '
AXFBJR( 12)( 1:  ) ='40537                 '
```

```
AXFBJR( 13)( 1:  ) ='40902                    '
AXFBJR( 14)( 1:  ) ='41268                    '
AXFBJR( 15)( 1:  ) ='41633                    '
AXFBJR( 16)( 1:  ) ='41998                    '
AXFBJR( 17)( 1:  ) ='42363                    '
AXFBJR( 18)( 1:  ) ='42729                    '
AXFBJR( 19)( 1:  ) ='43094                    '
AXFBJR( 20)( 1:  ) ='43459                    '
AXFBJR( 21)( 1:  ) ='43824                    '
AXFBJR( 22)( 1:  ) ='44190                    '
AXFBJR( 23)( 1:  ) ='44555                    '
AXFBJR( 24)( 1:  ) ='44920                    '
AXFBJR( 25)( 1:  ) ='45285                    '
AXFBJR( 26)( 1:  ) ='45651                    '
AXFBJR( 27)( 1:  ) ='46016                    '
AXFBJR( 28)( 1:  ) ='46381                    '
AXFBJR( 29)( 1:  ) ='46746                    '
AXFBJR( 30)( 1:  ) ='M_CINCODEMAYO            27'
AXFBJR( 31)( 1:  ) ='37016                    '
```

Many rows removed to avoid redundancy

```
              '

AXFBJR(503)( 1:  ) ='45972                    '
AXFBJR(504)( 1:  ) ='46337                    '
AXFBJR(505)( 1:  ) ='46702                    '
```

Change or make sure the numbers on the right are the counters for the first instance of a holiday. Note that Cinco De Mayo up above has a counter of 30 and 30 is in the second row shown below. Once you have done this, all you need to do is save the file and change the country.afs file and save it.

```
IAAA( 1)=  2
IAAA( 2)= 30
IAAA( 3)= 58
IAAA( 4)= 86
IAAA( 5)=114
IAAA( 6)=142
IAAA( 7)=170
IAAA( 8)=198
IAAA( 9)=226
IAAA(10)=254
IAAA(11)=282
IAAA(12)=310
IAAA(13)=338
IAAA(14)=366
IAAA(15)=394
IAAA(16)=422
IAAA(17)=450
IAAA(18)=478
CALL AASUB
RETURN
END'
```

Autobox

Reference Guide

Contents

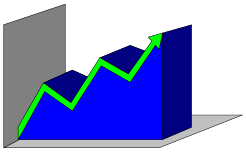
Overview

1. General Remarks

Autobox from Automatic Forecasting Systems (AFS) is an expert system which can be used to model and forecast both univariate and multivariate time series. Its methodology is based on the Box-Jenkin's ground breaking body of work. Outliers or intervention variables can be detected and incorporated into both the non-causal (univariate) and the causal (transfer functions) models. In addition, variance constancy and parameter constancy can be tested and remedies developed. The Autobox system has been under development (see references) since 1976 by AFS.

Autobox is attractive to both the expert and the non-expert. The expert can apply his knowledge and control the flow of modeling and estimation. The non-expert can use it as a "statistician-in-the-box" or as a "productivity aid" as it delivers a cost-effective solution. Autobox offers sophisticated analytical tools which can be recommended to teachers, academic researchers and forecasters who usually analyze a handful of time series with great care. Autobox offers a total automation of the modeling and forecasting tasks. Therefore, it can also be recommended to practitioners who must process large amounts of data automatically.

AFS was the first company to automate the Box-Jenkins model building process. Our approach is to program the model identification, estimation and diagnostic feedback loop as originally described by Box and Jenkins. This is implemented for both ARIMA (univariate) modeling and transfer function (multivariate or regression) modeling. What this means is that the user from novice to expert can feed Autobox any number of series and the program's powerful modeling heuristic can do the work for you. This option is implemented in such a way that it can be turned on at *any* stage of the modeling process. There is complete control over the statistical sensitivities for the inclusion/exclusion of model parameters and structures. These features allow the user complete control over the modeling process. The user can let Autobox do as much or as little of the model building process as you or the complexity of the problem dictates.



2. Modeling Overview

Autobox is a forecasting engine built on the central modeling steps of the Box-Jenkins paradigm. This core is extended by several useful modules: intervention detection and numerous others, all of which expedite the forecasting practitioner's tasks.

Autobox allows a single endogenous equation incorporating either user specified candidate causal series or empirically identified dummy series. The set of user specified candidate causal series can be either stochastic or deterministic (dummy) in form. During the search for the most appropriate model form and the optimal set of parameters the program can either be:

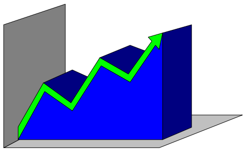
1. Purely empirical or
2. A starting model could be used.

A final model may require one or more of the following structures:

1. Power transformation like Log, Square Root, Reciprocal, etc.
2. Variance stabilization due to deterministic changes in the background error variance.
3. Data segmentation or splitting as evidenced by statistically significant changes in either model form or parameters.

En route to its tour de force, Autobox will evaluate numerous possible models/parameters that have been suggested by the data itself. In practice, a realistic limit is set on the maximum number of model form iterations. The exact specifics of each tentative model are not pre-set thus the power of Autobox emerges. The kind and form of the tentative models may never before been tried. Each data set speaks for itself and suggests the iterative process. The final model could be as simple as:

1. A simple trend model or a simple ordinary least squares model.
2. An exponential smoothing model.
3. A simple weighted average where the weights are either equal or unequal.
4. A Cochrane-Orcutt or ordinary least squares with a first order AR fixup.
5. A simple ordinary least squares model in differences containing some needed lags.



6. A spline-like set of local trends superimposed with an arbitrary ARIMA model with possible pulses.

The number of possible final models that Autobox could find is infinite and only discoverable via a true expert system.

A final model may require one or more of the following seasonal structures:

1. Seasonal ARIMA structure where the prediction depends on some previous reading S periods ago.

2. Seasonal structure via a ***complete*** set of seasonal dummy's reflecting a fixed response based upon the particular period.

3. Seasonal structure via a ***partial*** set of seasonal dummy's reflecting a fixed response based upon the particular period.

The final model will satisfy both:

1. Necessity tests that guarantee the estimated coefficients are statistically significant.

2. Sufficiency tests that guarantee that the error process is:

-unpredictable on itself

-not predictable from the set of causals or their lags (leads).

-has a constant mean of zero.

The final model will contain one or more of the following structures:

1. CAUSAL with correct lead/lag specification.

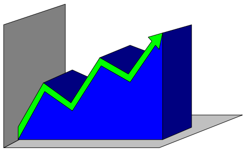
2. MEMORY with correct "autoregressive model components".

3. DUMMY with correct pulses, seasonal pulses, level shifts or spline time trends.

3. Statistical Features

Autobox considers the same model class, namely the seasonal autoregressive integrated moving average model with dynamic regressors (transfer components), abbreviated as SARIMAX(p,d,q)(P,D,Q)s[X]. The identification, estimation and outlier diagnosis issues are attacked in a rather elegant way via maximum likelihood estimation.

The user is responsible for selecting the dependent and independent series, and the time range so that Autobox will produce automatic forecasts. The way this is done is controlled by 155 switches specified by the user in various sub-menus.



Switches control the extent of automatic identification steps such as the use of Box-Cox transformation for variance stabilization, setting the confidence levels, the way how to handle differences to achieve stationarity and so on.

The results of the different modeling alternatives can be reviewed by setting all 45 switches governing "output report options" to a "YES". This yields huge amounts of output, but reveals at least some part of the inherent modeling wisdom of the program. Autobox uses several advanced and clever techniques to identify suitable forecast models, including necessity and sufficiency tests for the parameters, a stability check of the ARMA polynomials (stationarity and invertibility). In keeping with the latest methodologies, we have applied a recently developed test by Franses to discriminate between seasonal unit roots and seasonal dummy model components.

Outlier handling in Autobox recognizes several other types of outliers which **cannot** be handled by the other programs. One of these is the seasonal pulse, which can be very useful to model seasonal patterns which occur only in a few months (like Christmas effects in December). Another is the step or level shift variable which characterizes a permanent change in the level of the series. A recent advancement is the incorporation of trend detection where the model can

have a number of trends each with their own slope.

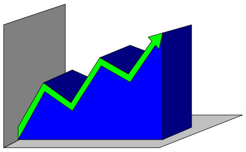
Autobox can detect dynamic patterns such as transient changes in the data (i.e. level shifts in the data can be detected in both systems as additive outliers in the first differences). It is also able to detect deterministic trend changes which are increasingly used by econometricians.

Different types of exponential smoothing schemes also can be estimated. Moreover, the parameters of the smoothing procedures are automatically optimized. More significantly for ARIMA users, Autobox extends its automatic modeling to causal modeling.

It offers two causal model identification options. The first technique is the well known prewhitening method developed by Box and Jenkins. As an alternative, the user can select the common filter/least squares technique which is useful when the input causal series are cross-correlated.

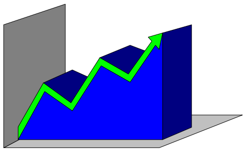
Autobox uses the theoretical covariances for alternative candidates as the scheme for both ARIMA and TRANSFER FUNCTION identification. This rather sophisticated pattern recognition selects the numerator/denominator polynomial orders for the rational transfer function weights.

Autobox permits the user to easily evaluate the forecast accuracy of the identified model.



First, the user can withhold some observations for forecast evaluation using an out-of-sample analysis. Second, different forecast evaluation measures including the mean error (bias) and mean absolute percentage error (MAPE) can be computed for different forecast horizons without any additional programming.

Autobox comes with a complete set of identification and modeling tools for use in the BJ framework. This means that you have the ability to transform or prewhiten the chosen series for identification purposes. Autobox handles both ARIMA (univariate) modeling and Transfer Function (multivariate) modeling allowing for the inclusion of interventions (see below for more information). Tests for interventions, need for transformations, need to add or delete model parameters are all available. Autocorrelation, partial autocorrelation and cross-correlation functions and their respective tests of significance are calculated as needed. Model fit statistics, including R^2, SSE, variance of errors, adjusted variance of errors all reported. Information criteria statistics for alternate model identification approaches are provided.



Intervention Detection

One of the most powerful features of Autobox is the inclusion of Automatic Intervention detection capabilities in both ARIMA and Transfer Function models. Almost all forecasting packages allow for interventions to be included in a regression model. What these packages don't tell you is how sensitive **all** forecasting methodologies are to the impact of interventions or missing variables. These packages don't tell you if your series may be influenced by missing variables or changes that are outside the current model.

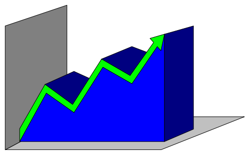
If a data series is impacted by changes in the underlying process at discrete points in time, both ARIMA models and Transfer Function models will produce poor results. For example, a competitors price change changes the level of demand for your product. Without a variable to account for this change your forecast model might perform poorly. Autobox implements ground breaking techniques which quickly and accurately identify potential interventions (level shifts, season pulses, single point outliers and changes in the variance of the series). These variables can then be included in your model *at your discretion*. The result is more robust models and greater forecast accuracy.

Graphical Analysis Tools

Autobox has a set of graphing tools that help present complex statistical information in a way that is easy and clear at every stage of the forecasting process. For example, graphs of autocorrelation, partial-autocorrelation and cross-correlation functions are all available. Even more incredibly these can be compared to theoretical values for various models forms.

Forecasting and Diagnostics

All forecast packages allow for you to produce forecasts using the models you have constructed. Autobox presents the critical information you need to determine if those forecasts are acceptable. Autobox has options that allow you to analyze the stability and forecasting ability of your forecast model. This is achieved through a series of ex-post forecast analyses. You can *automatically* withhold any number of observations, re-estimate the model form and forecast. Observations are then added back one at a time and the model is re-estimated and reforecast. Forecast accuracy statistics, including Mean Absolute Percent Error (MAPE) and Bias, are calculated at each forecast end point. Thus the stability of the model and its ability to forecast from various end points can be analyzed. Finally, you can optionally allow Autobox to actually re-identify the model form at each level of withheld data to see if the *model form* is unduly influenced by recent observations.



4. Conclusions

Academic researchers who usually analyze a handful of time series with great care and in great detail will be able to accomplish their tasks readily. For this group, the expert system component should be part of a data analysis support tool as it can be a “productivity aid”. Quite possibly, they might even find it a replacement (at times) for some of the repetitive tasks and allow them to focus on the “difficult series”.

Due to the high integration of the different time series analysis tools into an integrated framework, Autobox can be strongly recommended to practitioners who must handle many data sets. Practitioners with only little knowledge about Box-Jenkins models can use the program as a pure black-box tool, assuming some knowledgeable user with experience in forecasting methods has configured the system to match the specific features of the data and the specific needs of the end user.

Without the support of an experienced forecaster, a novice may unwittingly run into pitfalls. This is probably true for any sophisticated forecasting program not just this one.

2

Conceptual Overview

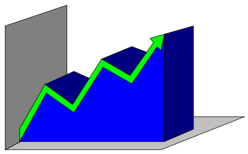
Autoregressive Integrated Moving Average (ARIMA) is a process designed to identify a weighted moving-average model specifically tailored to the individual dataset by using time series data to identify a suitable model. It is a rear-window approach that does use user-specified helping variables. It uses correlations within the history to identify patterns that can be statistically tested and then used to forecast. Often we are limited to using only the history and no causals whereas the general class of Box-Jenkins models can efficiently incorporate causal/exogenous variables (Transfer Functions). Thus ARIMA should be more accurately called Univariate Box-Jenkins rather than Box-Jenkins.

The modeling process includes an identification phase. This tutorial will introduce concepts and the steps to identify the model, estimate the model, and perform diagnostic checking to revise the model. We will also list the assumptions and how to incorporate remedies when faced with potential violations.

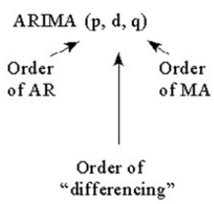
The Background

Our understanding of ARIMA has grown since it was introduced in (1). Properly formed ARIMA models are a general class that includes all well-known models save some state space and multiplicative Holt-Winters models. While classical ARIMA modeling attempted to capture stochastic structure in the data little was done about incorporating deterministic structure other than a possible constant or identifying change points in parameters or error variance.

In the Exceptions Section we will highlight procedures relevant to suggested augmentation strategies as they were not part of the original ARIMA approach suggested in (1) but are now standard. This step is often ignored as it is necessary that the mean of the residuals is invariant over time and that the variance of the final model errors are constant.



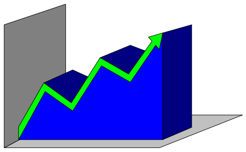
The idea of modeling is to characterize the pattern in the data and the goal is to identify an underlying model that is generating that pattern. The model that you will build should match the history which can then be extrapolated into the future. The actual minus the fitted values are called the residuals. The residuals should be random around zero (i.e. Gaussian) signifying that the pattern has been extracted and captured in the model. In primary school we learned how to determine the general term of both an arithmetic and geometric series, For example the series 1,3,5,7,9,11 would be easily characterized. A series 1,9,1,9,1,9 can also be easily characterized as a process. A series like 1,9,1,9,1,9,5,9,1,9 containing a pulse requires a more sophisticated characterizer as there is deterministic effect at period 7. In general ARIMA modeling is the search for the general term for a given data series.



- For example, a AR model for monthly data may contain information from lag 12, lag 24, etc.
  - i.e. $Y_t = A_1 Y_{t-12} + A_2 Y_{t-24} + a_t$
  - This is referred to as an ARIMA(0,0,0)x(2,0,0)$_{12}$ model
  -
- General form is ARIMA(p,d,q)x(ps,ds,qs)$_s$

While seasonal autoregressive structure might be intuitively appealing, early versions of ARIMA premised no seasonal fixed effects such as a February effect or a June effect. Modern ARIMA modeling can suggest the nature of these omitted fixed effects (pulses, seasonal pulses, level/step shifts, local time trends).

The Tools



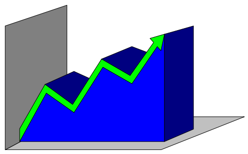
The ARIMA process uses regression/correlation statistics to identify the stochastic patterns in the data. Regressions are run to find correlations based on different lags in the data. The correlation between successive months would be the lag 1 correlation or in ARIMA terms, the ACF of lag 1. Is this month related to one year ago at this time would be the lag 12 correlation or in ARIMA terms, the ACF of lag 12. By studying the autocorrelations in the history, we can determine if there are any relationships and then take action by adding parameters to the model to account for that relationship. The different autocorrelations for the different lags are arranged together in what is known as a correlogram and are often presented using a plot. They are sometimes presented as a bar chart. We present it as a line chart showing 95% confidence limits around 0.0 . The autocorrelation is referred to as the autocorrelation function (ACF).

- The key statistic in time series analysis is the autocorrelation coefficient (the correlation of the time series with itself, lagged 1, 2, or more periods.)
- Recall the autocorrelation formula:

$$r_k = \frac{\sum_{t=k+1}^{n} (y_t - \bar{y})(y_{t-k} - \bar{y})}{\sum_{t=1}^{n} (y_t - \bar{y})^2}$$

Let's consider the Partial Autocorrelation Function (PACF). The PACF of lag 12 for example is a regression using a lag of 12, but also uses all of the lags from 1 to 11 as well, hence the name partial. The ACF could have been called the unconditional regression and the PACF the conditional regression as it measures the conditional impact of adding one more lag. Given this, you should then be able to realize that the PACF of lag 1 is identical to ACF of lag 1. The PACF is ordinarily only needed when performing initial model identification.

- Partial autocorrelations measures the degree of association between y_t and y_{t-k}, when the effects of other time lags 1, 2, 3, …, k-1 are removed.
- The partial autocorrelation coefficient of order k is evaluated by regressing y_t against $y_{t-1}, \ldots y_{t-k}$:



$$y_t = b_0 + b_1 y_{t-1} + b_2 y_{t-2} + \cdots + b_k y_{t-k}$$

Now that we have explained the ACF and the PACF, let's discuss the components of ARIMA. There are three pieces to the model. The "I" means Integrated, but for simplicity sake it really means that you took differencing on the Y variable during the modeling process. The "AR" means that you have a model parameter that explicitly uses the history of the series. The "MA" means that you have a model parameter that explicitly uses the previous forecast errors. Not all models have all parts of the ARIMA model. All models can be re-expressed as pure AR models or pure MA models. The reason we attempt to mix and match has to do with attempting to use as few parameters as possible. For example the somewhat curious model of single exponential smoothing uses 1 coefficient (an MA (1) coefficient) in lieu of an infinite set of AR weights. This model is curious because it expressly restricts the sample space of the so-called smoothing coefficient to ½ the possible space.

Identifying the order of differencing starts with the following initial assumptions, which are ultimately verified:

1) The sequence of errors (a's) are assumed to have a constant mean of 0 and a constant variance for all sub-intervals of time. Note that the inclusion of a constant in the model guarantees that the mean of all the a?s will be 0, but it doesn't guarantee this for each and every sub-interval. Power transformations should never be based upon the properties of

the original series but based on the distribution of the estimated a's as that is where the statistical requirements are in effect.

2) The sequence of errors (a's) are assumed to be normally distributed where the a's are independent of each other. Finally the model parameters and error variance are assumed to be fixed over all sub-intervals.

We study the ACF and PACF and identify an initial model. If this initial model is a good model then the residuals will be free of structure and we are done. If not we identify that structure and add that structure to the current model until a subsequent set of residuals is free of structure. This is akin to what happens when your eye doctor creates a new prescription/model for you. One could consider this iterative approach as moving structure in the errors to the model until there is no structure to relocate.

Following are some simplified guidelines to follow when identifying an appropriate ARIMA model with the following assumptions:

* Guideline 1: If the series has a large number of positive autocorrelations then differencing should be introduced. The order of the differencing is suggested by the significant spikes in the PACF based upon the standard deviation of the differenced series.

* Guideline 2: Include a constant if your model has no differencing; include a constant elsewhere if it is statistically significant.

* Guideline 3: Domination of the ACF over the PACF suggests an AR model while the reverse suggests an MA model. The order of the model is suggested by the number of significant values in the subordinate.

* Guideline 4: Parsimony: Keep the model as simple as you can but not too simple as overpopulation often leads to redundant structure.

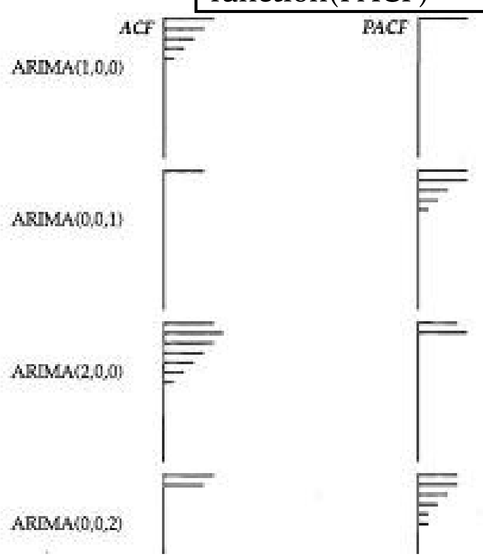
* Guideline 5: Evaluate the statistical properties of the residual (At) series and identify the additional structure (step-forward) required. Note that the ACF is the ratio of the covariance to the instantaneous variance which means that if there are violations in the At's such as pulse outliers, level shifts , seasonal pulses and/or local time trends this test

can be seriously biased downwards leading to the conclusion that there is no additional stochastic structure required. The variance is over estimated due to the pulses/level shifts/local time trends thus the ACF is understated. This has been referred to as “The Alice in Wonderland Effect”.

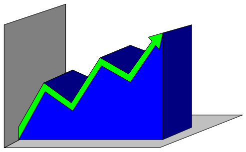
* Guideline 6: Reduce the model via step-down procedures to end up with a minimally sufficient model that has effectively deconstructed the original series to signal and noise. Over-differencing leads to unnecessary MA structure while under-differencing leads to overly complicated AR structure.

Guidelines for model form

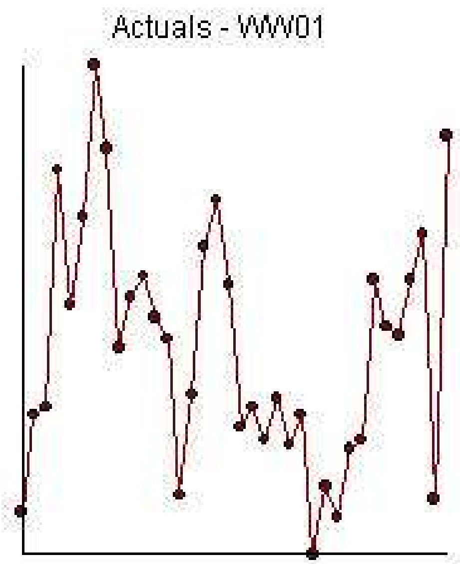
Process	MA(q)	AR(p)	ARMA(p,q)
Autocorrelation function(ACF)	Cuts off	Doesn't cut off	Doesn't cut off
Partial Autocorrelation function(PACF)	Doesn't cut off	Cuts off	Doesn't cut off



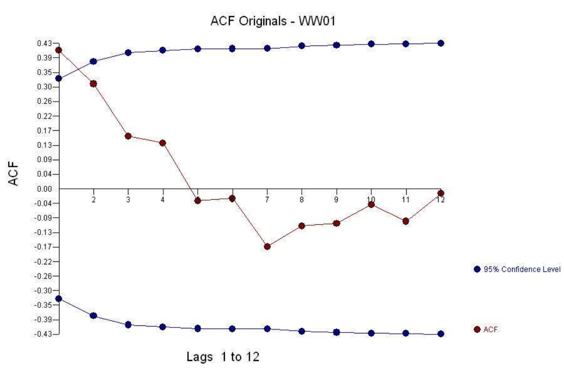
One could consider this a custom-fitting exercise, being cautious not to over fit or under fit. The key question is after estimation, is the current model sufficient? If not we identify the remedial action required via the ACF and PACF of the current (tentative) model's residuals and incorporate that structure into the model. Note that only statistically significant coefficients are used in the final model. If the series did not require differencing then you must have a constant in the model. If the model includes differencing then the constant is optional. In either case we still have to test that our assumptions regarding the errors are correct or at least can't be rejected.

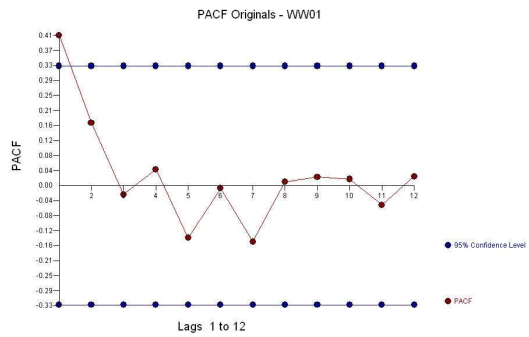


Example 1) We have 36 annual values.



The ACF and the PACF suggest an AR(1) model (1,0,0)(0,0,0).

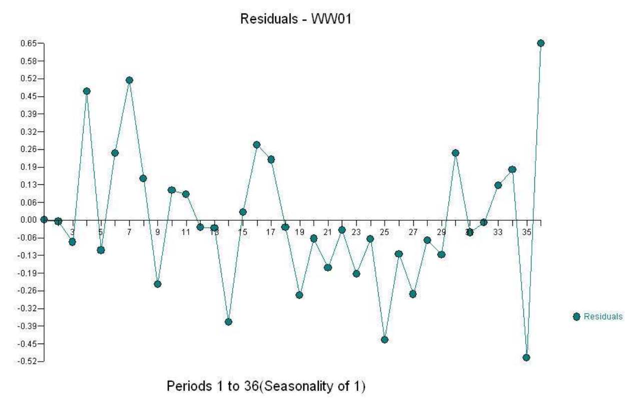




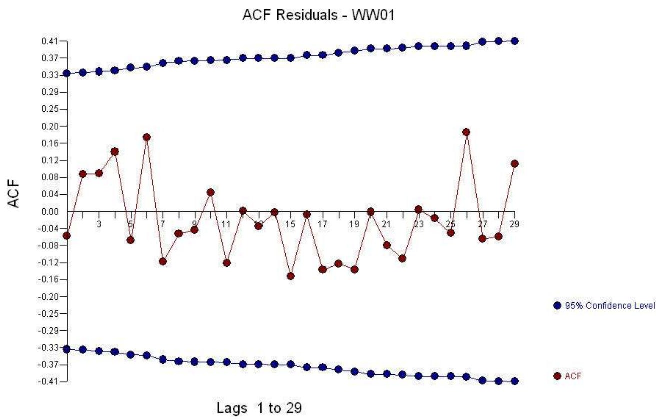
Leading to an estimated model (1,0,0)(0,0,0).

```
Y(T) =  .60123      +     [(1-  .454B** 1)]**-1  [A(T)]
```

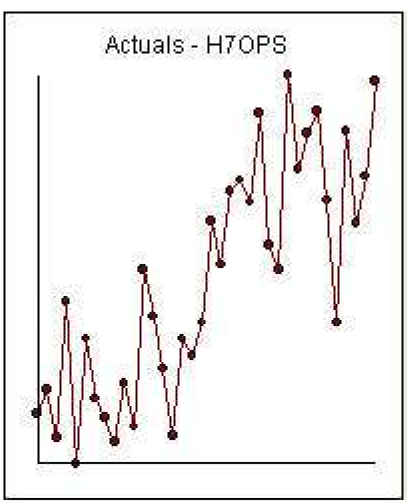
With the following residual plot, suggesting some "unusual values".



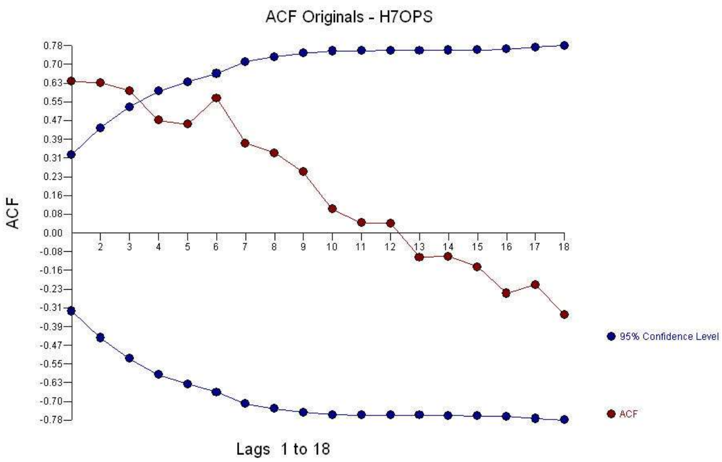
The ACF and PACF of the residuals suggests no stochastic structure as the anomalies effectively downward bias the results.

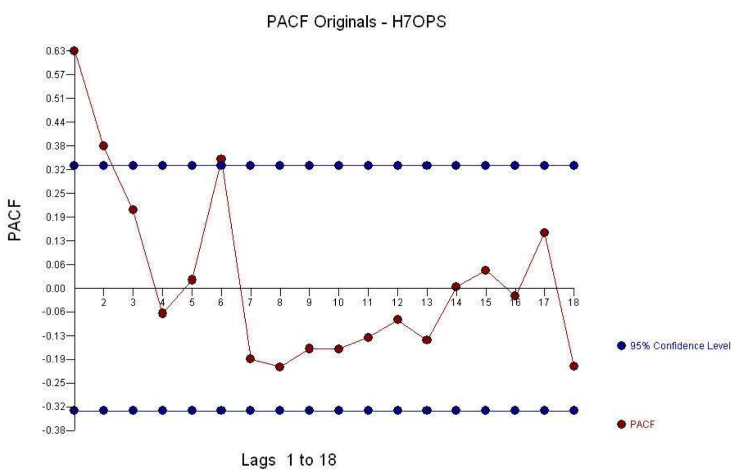


Example 2 has 36 monthly observations:



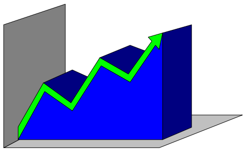
With ACF and PACF :



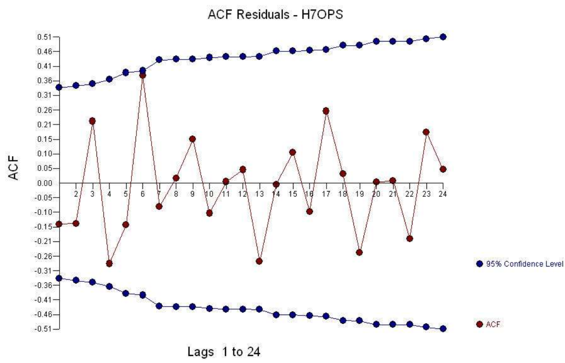


Leading to an estimated model: AR(2) (2,0,0)(0,0,0)

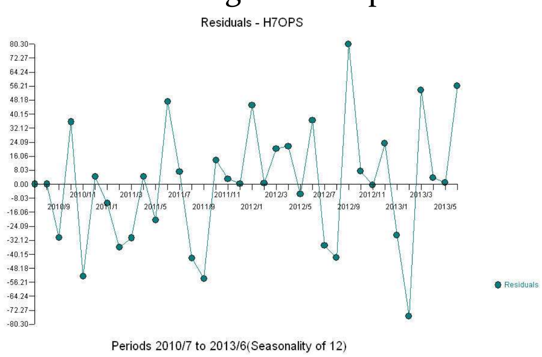
```
Y(T) =  381.96      +     [(1-  .378B** 1-  .462B** 2)]**-1  [A(T)]
```



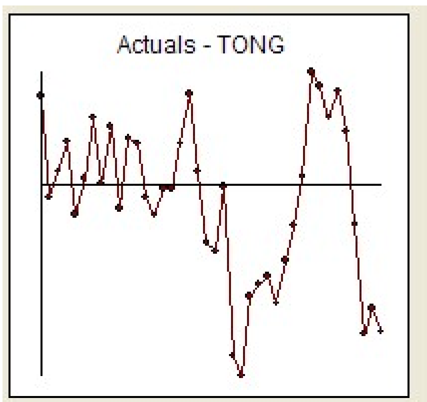
And with ACF of the residuals:



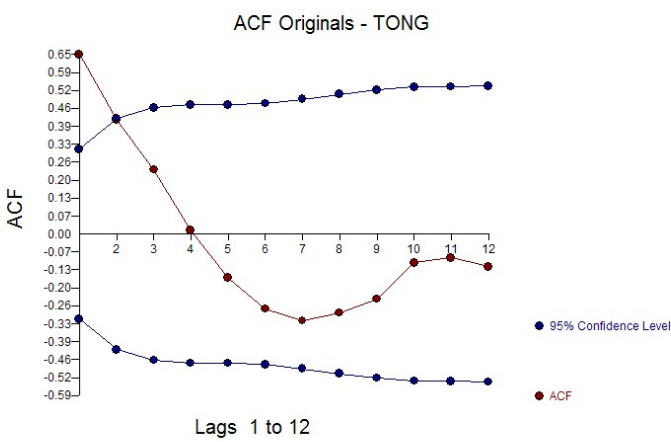
With the following residual plot:

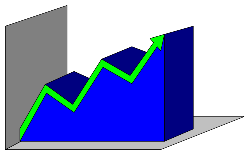


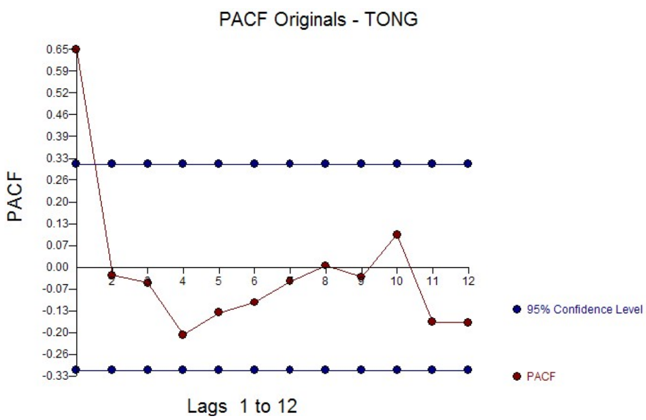
Example 3 has 40 annual values



The ACF and PACF of the original series are;



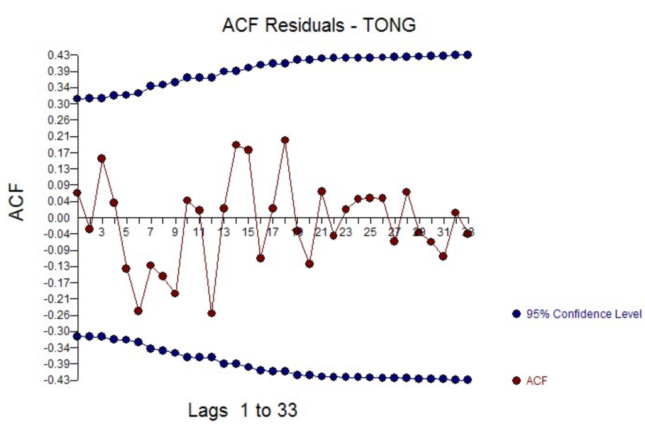


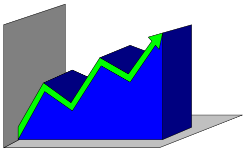


Suggesting a model (1,0,0,(0,0,0)

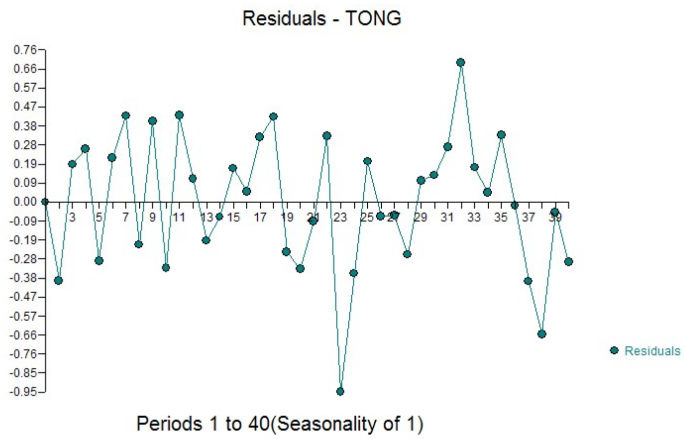
```
Y(T) = -.19797      +     [(1-   .706B** 1)]**-1   [A(T)]
```

With a residual ACF of





And residual plot



Suggesting a change in distribution of residuals during the second half.

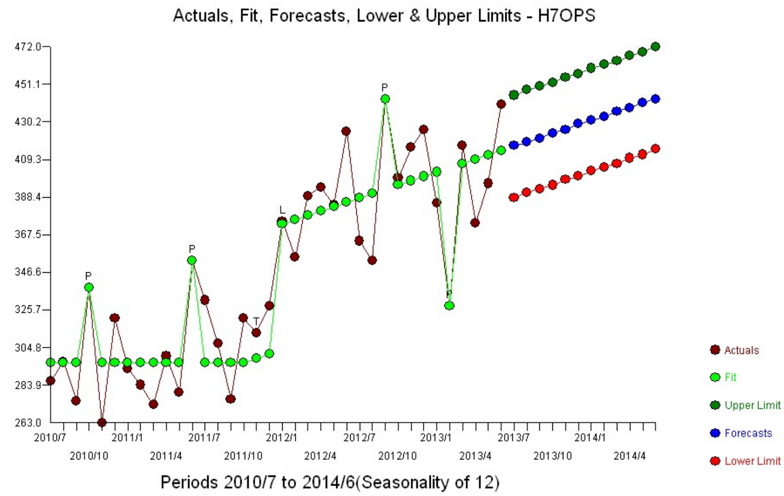
Exceptions: If a tentative model exhibits errors that have variance change this can be remedied in a number of ways ; 1) Identify the need to validate that the errors have constant mean via Intervention Detection (2,3) yielding pulse, seasonal pulse/level shift/ /local time trends 2) confirming that the parameters of the model are constant over time 3) confirming that the error variance has had no deterministic change points or stochastic change points. We will focus here on category 1. The tool to identify omitted deterministic structure is fully explained in references 2 and 3. The idea Step 1)is to use the model to generate residuals and then Step 2)identify the intervention variable needed following the procedure defined in reference 3. Step 3). Re-estimate the residuals incorporating the effect into the model and then go back to Step 1 until no additional interventions are found.

We now revisit Example 1 and augment the model with pulse indicators providing a more robust estimate of the ARIMA coefficients.

```
Y(T) =  .75857                                    WW01
       +[X1(T)][(+  .280)]                       :PULSE           7
       +[X2(T)][(+  .441)]                       :PULSE           4
       +[X3(T)][(-  .746)]                       :PULSE          35
       +[X4(T)][(-  .254)]                       :PULSE          25
       +[X5(T)][(-  .302)]                       :PULSE          14
       +[X6(T)][(-  .300)]                       :PULSE           9
       +[X7(T)][(-  .198)]                       :PULSE          19
       +[X8(T)][(+  .246)]                       :PULSE          30
      +     [(1-  .859B** 1)]**-1  [A(T)]
```

Example 2 is a series that is better modeled with a step/level shift.
The plot of the residuals suggests a mean shift. Empowering Intervention Detection leads to an augmented model incorporating a level shift and a local time trend with and 4 pulses and a level shift. This model is as follows:

```
Y(T) =  296.30                               H7OPS
       +[X1(T)][(+ 2.4008)]    :TIME TREND     17
       +[X2(T)][(- 76.6519)]   :PULSE          32
       +[X3(T)][(+ 69.9367)]   :LEVEL SHIFT    19
       +[X4(T)][(+ 56.6959)]   :PULSE          12
       +[X5(T)][(+ 50.3521)]   :PULSE          27
       +[X6(T)][(+ 41.6959)]   :PULSE           4
      +       [A(T)]
```



Example 3

When the parameters were tested for constancy. over time using the Chow Test (5) , a significant difference was detected at period 21.

.....

F TEST TO VERIFY CONSTANCY OF PARAMETERS

CANDIDATE	BREAKPOINT	F VALUE	P VALUE
18	18	6.0905	.0054861917
19	19	5.6979	.0073431038
20	20	6.2987	.0047097652
21	21	5.5804	.0080202803*

As period 1-20 was

#	MODEL COMPONENT		LAG (BOP)	COEFF	STANDARD ERROR	P VALUE	T VALUE
1	CONSTANT			.963E-01	.587E-01	.1183	1.64
2	Autoregressive-Factor #	1	1	-.122	.229	.6021	-.53

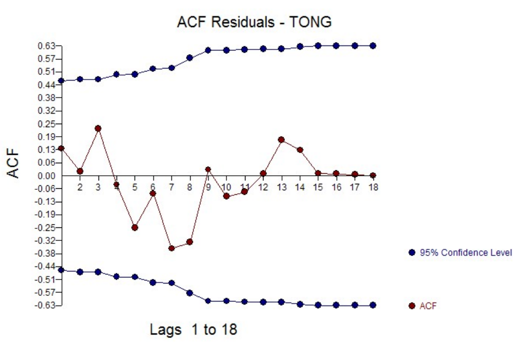
period 21-40

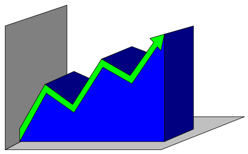
#	MODEL COMPONENT		LAG (BOP)	COEFF	STANDARD ERROR	P VALUE	T VALUE
1	CONSTANT			-.668E-01	.805E-01	.4175	-.83
2	Autoregressive-Factor #	1	1	.861	.128	.0000	6.72

A final model using the last 20 values was

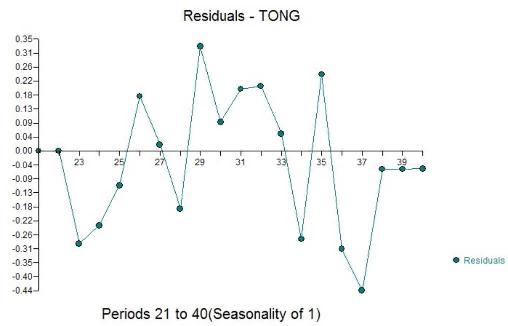
```
[(1-B**1)]Y(T)  = +[X1(T)][(1-B**1)][(-   .886)]          :PULSE             23
                  +[X2(T)][(1-B**1)][(+   .286)]          :PULSE             32
                  +[X3(T)][(1-B**1)][(-   .345)]          :PULSE             38
                  +[X4(T)][(1-B**1)][(-   .715)]          :PULSE             24
                 +      [(1-   .453B** 1)]**-1   [A(T)]
```

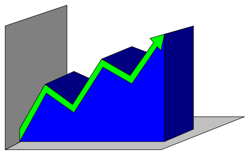
With residual ACF of





And residual plot of





3

References

Anderson, O.D. (1975). Time Series Analysis and Forecasting. London: Butterworth & Co.

Bell, W. (1983). "A Computer Program for Detecting Outliers in Time Series," in American Statistical Association 1983 Proceedings of the Business Economic Statistics Section, Toronto, pp. 624-639.

Box, G.E.P., and Cox, D.R. (1964). "An Analysis of Transformations (with discussion)", Journal of the Royal Statistical Society, B26, pp. 212-252.

Box, G.E.P., and Jenkins, G.M. (1976). Time Series Analysis: Forecasting and Control, 2nd ed. San Francisco: Holden Day.

Box, G.E.P., and Tiao, G. (1975). "Intervention Analysis with Applications to Economic and Environmental Problems," Journal of the American Statistical Association, Vol 70,70-79

Chang, I., and Tiao, G.C. (1983). "Estimation of Time Series Parameters in the Presence of Outliers," Technical Report #8, Statistics Research Center, Graduate School of Business, University of Chicago, Chicago.

Downing, D.J., and McLaughlin, S.B. (1986). "Detecting Shifts in Radical Growth Rates by the Use of Intervention Detection," Engineering Physics and Mathematics Division, Oak Ridge National Laboratory, Oak Ridge.

Franses, P., (1991) "Seasonality, non-stationarity and forecasting of monthly time series", International Journal of Forecasting", (7), pp 192-208.

Liu, L.M., and Hanssens, D.M. (1982). "Identification of Multiple-Input Transfer Function Models," Communications in Statistics - Theory and Methods, 11(3), pp. 297-314.

Mabert, V.A. (1975). "An Introduction to Short Term Forecasting Using the Box-Jenkins Methodology", Publication #2 AII.E. Monograph Series, Production Planning and Control Division, Norcross, Georgia.
Makridakis, S., Wheelright, S. and McGee, V. (1983). Forecasting: Methods and Applications. New York: Wiley.

McCleary, R., and Hay, R. (1980). Applied Time Series Analysis for the Social Sciences. Los Angeles: Sage.

Nelson, C.R. (1973). Applied Time Series Analysis for Managerial Forecasting. San Francisco: Holden Day.

Pack, D.J. (1978). "Forecasting Time Series Affected by Identified Isolated Events," College of Administrative Science, The Ohio State University, Ohio.

Pack, D.J. (1977). "Revealing Time Series Interrelationships," Decision Sciences, 8, pp. 377-402.

Quenouille, M.H. (1957). The Analysis of Multiple Time Series. London: Charles Griffin & Company.

Reilly, D.P. (1980). "Experiences with an Automatic Box-Jenkins Modeling Algorithm," in Time Series Analysis, ed. O.D. Anderson. (Amsterdam: North-Holland), pp. 494-508.

Reilly, D.P. (1987). "Experiences with an Automatic Transfer Function Algorithm," in Computer Science and Statistics Proceedings of the 19th Symposium on the Interface, ed. R.M. Heiberger, (Alexandria, VI: American Statistical Association), pp. 128-135.

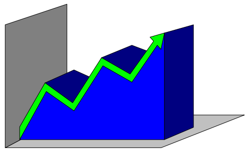
Shumway, R.H. (1988). Applied Statistical Time Series Analysis. Englewood Cliffs: Prentice Ha Hall.

Tiao, G.C., and Box, G.E.P. (1981). "Modeling Multiple Time Series with Applications," Journal of the American Statistical Association, Vol. 76, pp. 802-816.

Tsay, R.S. (1986). "Time Series Model Specification in the Presence of Outliers," Journal of the American Statistical Society, Vol. 81, pp. 132-141.

Tsay, R.S., and Tiao, G.C. (1983). "Consistent Estimates of Autoregressive Parameters and Extended Sample Autocorrelation Function for Stationary and Nonstationary ARMA Models," Technical Report #2, Statistics Research Center, Graduate School of Business, University of Chicago, Chicago.

Vandaele, W. (1983). Applied Time Series and Box-Jenkins Models. Orlando: Academic Press.



4

Reviews

Tim C. Ireland and Ramesh Sharda, A Test of Box-Jenkins Forecasting Expert Systems for Microcomputers, Department of Management, Oklahoma State University, College of Business Administration, Stillwater, OK 74078, Phone: (405) 744-8642, (405) 744-8638.

John C. Picket (1991), Autobox 3.0, International Journal of Forecasting, 7:395-398.

John C. Picket (1991), Autobox 3.0, OR/MS Today, pp. 32-35, April 1991.

David P. Reilly (1980) "Experiences with an Automatic Box-Jenkins Modeling Algorithm", in Time Series Analysis, ed. O.D. Anderson. (Amsterdam: North-Holland). Pp. 494-508

David P. Reilly (1987) "Experiences with an Automatic Transfer Function Algorithm", in Computer Science and Statistics Proceedings of the 19th Symposium on the interface, ed. R.M. Heiberger, (Alexandria, Va.: American Statistical Association), pp. 128-135.

J. Keith Ord (1986), Autobox - Software Review, International Journal of Forecasting, 2:512-513.

R.H. Shumway (1986), Autobox (Version 1.02), The American Statistician, 40(4):299-300.

Leonard J. Tashman and Michael L. Leach (1991), Automatic forecast software: A survey and evaluation, International Journal of Forecasting, 7:209-230.

Pamela A. Texter and J. Keith Ord, (1989), Forecasting using automatic identification procedures: A comparative analysis, International Journal of Forecasting, 5:209-215.

Pamela Texter-Geriner and J. Keith Ord, (1991), Automatic forecasting using explanatory variables: A comparative study, International Journal of Forecasting, 7:127-140.