Autobox 6.0
User's Guide
Batch Version

This manual is published by

Automatic Forecasting Systems©, Inc.

6/7/2010

Table of Contents

Introduction

Thank you for purchasing our software. Call us at 215-675-0652 or email us at sales@autobox.com, if you have any questions.

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

- Data Cleansing (i.e. Identify outliers and correct the data for it's 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?)
- Simulation/Scenario Analysis (i.e. What would happen if I lowered the price down to $xx?)

To use the Batch process the process follows like this: 1)The series to be forecast (and any causal variables) are created in separate flat files via importing from Excel or you the user create the ASC file(s). 2)The user chooses how he wants to run the system (using our conditions or making some modifications for your specific purposes). We definitely would recommend customizing Autobox to fit your needs! 3)Run Autobox 4)Use the Forecast files (again flat files) that are created which you can then use for your next step.

You might consider this strategy: 1)If you have data with different frequencies (i.e. monthly and quarterly), you should group the series that are alike into DIFFERENT Autobox Project files(XML files). This is due to the fact that **sometimes** different modeling approaches are used for different data. 2)You also can't mix Univariate and Multivariate problems into the same Project file.

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.

Let's define what a Causal Variable means first. It is BOTH of the following two lines:
Causal Variables – GNP, Price, Promotion, etc.
Dummy Variables – Christmas, Trend, Month of the year effect

Autobox will **automatically** aid the modeling process for weekly, daily, hourly and semi-hourly data. If you have weekly, daily or hourly data, Autobox will add 51 dummy variables for the different weeks of the year. You need at least 1 1/2 year of historical data for this to happen. If you have daily, hourly or semi-hourly data Autobox will add 6 dummy variables for day of the week. If you have hourly data Autobox will add 23 dummy variables for hour of the day. If you have semi-hourly data Autobox will add 47 dummy variables for each half-hour of the day.

For daily data that covers all 7 days(Monday to Sunday), Autobox will different modeling approaches. 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

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.
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 this text file place a '20' in it so that Autobox can model it.

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.

AFS has been in business since 1975 and has many business and academic customers. AFS was also selected in the textbook "Principles of Forecasting" as the 'BEST DEDICATED FORECASTING PACKAGE' See page 671 on a your favorite book web site.

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

Process Flow (For Interactive Batch Version – Univariate and Multivariate)

Interactive Batch Version – Imports XLS files or reads in Autobox "ASC" files and writes forecasts to a flat file.

1)Import your data into Autobox from Excel (XLS) or create an Autobox data file (ASC file) – one for each time series in a specific format that Autobox batch requires(more on that later) –
2)Create an Autobox Project (XML file) - This defines the location of the data, the series to run and, the output to be stored and where to store the XML file
3)Add Intelligence - Create Dummy variables based on your knowledge of historical events + holiday variables + special variables (trading days, monthly effects, etc.)
4)Choose which series will use the changes made in the previous step
5)Choose Batch Run – Runs Autobox
6)Forecasts Output is generated – Forecasts, Graphs, and detailed reports on the modeling process
7)Choose whether to save the asc files you have adjusted in step 5.

Process Flow for Running Against XLS Files – Univariate only)

Reads in an XLS file and appends forecasts to the same XLS file.

1)Create XLS data file – Series are in rows with the SKU in column 1 and the most recent data on the far right see Heinekin.xls as an example in the directory
2)Choose which graphs you want (ie "Actual/fir/Forecasts) and the seasonality of the data (ie 12 for monthly)
3)Choose "File/Run From Excel" and select the XLS file to run "HEINEKIN.XLS" for example.
4)Forecasts are appended into the XLS file. There are JPGs of different graphs and HTM files with reports from the process written to the directory that you might want to review

Limitations:
1)<u>Only uses the history(no causals, holidays, etc.) If you are ok with this limitation which might be ok for your situation then use this option.</u>
2)This isn't much of a limitation, but we thought just to mention it. Excel has a maximum of 230(16,384 for XLSX) columns so that limits the size of the problem to the number of historical observations plus the number of forecasts. There are 65,536(1,048,576 for XLSX) rows in Excel which is also a limitation.

We will not discuss this option any more in the manual as it is a basic way to run Autobox, but doesn't include causals just yet. You can run this way if you want and we will leave it at that.

Installation

A directory under the Autobox directory is created during installation named "OUTPUT". This is where the batch will save all of the reports from the run as a default. We include many data files for you to play with along with some XLS files to import as examples. We suggest you open our example XLS files in Excel and take a quick peek to see what the data looks like before importing. ***Click on aboxb.exe to start your Batch session for Interactive or click on go.bat(after editing it for your data first) to run the command version.***

Step 1
Getting your data ready

You can either run your data directly from Excel, or import your data from Excel or create an Autobox "ASC" DOS text file. Which way you go is based on what data you have to use. An ASC file is a flat file that lists some specifics for Autobox like (# of forecasts, starting year, starting period, frequency of the data, data type, etc.)

If you have 1,000 series in Excel to forecast and no causal variables then importing using Excel will be your best option so using the "command version" makes sense.

If you have 1,000 series to forecast and causal variables that are the same for every series(GNP, Interest rates, etc.) then importing using Excel will be your best option. Autobox will take your 1,000 series and create 1,000 ASC files. You can then use the "Add Generated Series/Add a User Defined Series" process to bring in your causals to all of the series and Autobox will adjust all of your ASC files automatically.

If you have 1,000 series to forecast and different causal variables that are different for each of the 1,000 series (Price for a given SKU, etc.) then creating an Autobox ASC file will be your best choice. This will require formatting your data using an Excel Macro, Access, Oracle, or SAS to get the data ready for Autobox. See Page 14 "Creating your own ASC file".

There is an example XLS file in the installation directory named AFS20.XLS which you can review and try importing. Your data should be in columns. There are two ways for you to let Autobox use your data. You can either import the data into Autobox using XLS or create your own Autobox ASC file. An XLS file that is included in the installation (AFS20.XLS) for you to review. Note that the last 4 rows are future values.

If you don't have any causal variables then think harder. Is there anything that drives the business that you could include in the modeling? After you identify causals, include the future values of your causals. Autobox can project your causals in the future which is a unique feature in the marketplace, but if you have knowledge of the future you should include that knowledge into the future estimates! Also, if you have dummy variables that you are including in the modeling process like a seasonal dummy model, you do know what those dummy variables would be in the future so include their future values as well.

Import your data from XLS

Open the XLS file that is included in the installation (AFS20.XLS). Note that the last 4 rows are future values. It is important to note that the causal variables will have a data type (which defines for Autobox how the future values are created and used in Autobox – more on that on the next page). The important thing to note is that if the series have no future values then the data type is defaulted to '0' which means Autobox will forecast the future values which may be what you want, but depending on your variable may not be what you want. If the causal variable does have future values then the data type is defaulted to '2' which means you supply the future values. If you want a '1' or '3' data type you can't use the XLS import facility and are forced to create your own ASC file as you can't change "en masse" data types. It just isn't possible to manage all of that data as Autobox is not a database. Note that if you have a promotion that if you have daily data and a promotion that goes on for many periods you would want to use a data type of '1' as you can't really look for lead/lag from multiple time periods.

Let's define the data types now so you will understand how important they are:

Data Types:

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

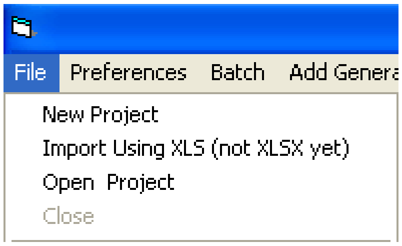
2 – current time period and lag effects (same as data type '0' except the user is specifying the future values) are analyzed

3 – current time period and lag and lead effects are analyzed

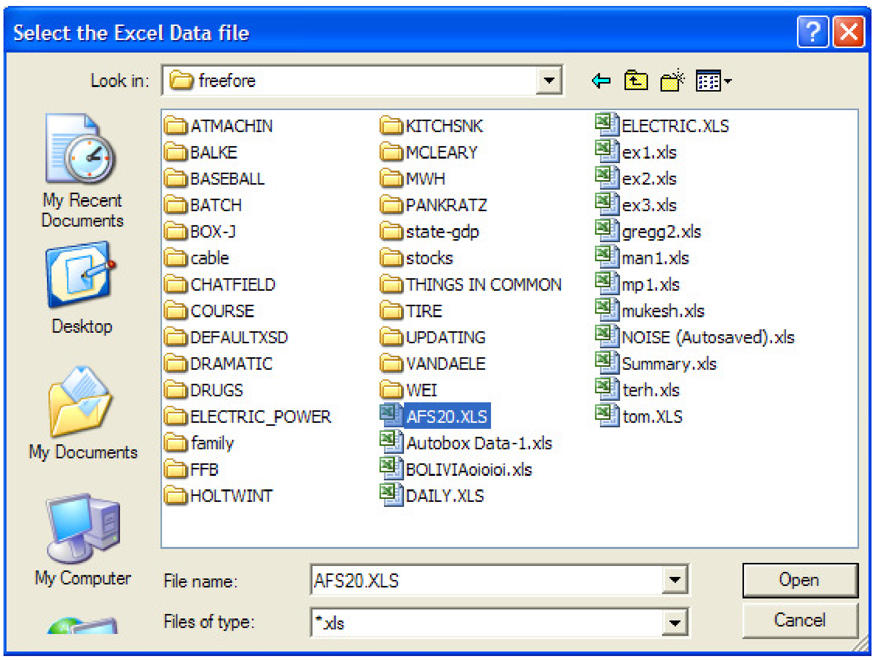
In order to change the data type default from a '2' after importing, you will need to import future values of causal variables (which would logically be at the very bottom of the dataset see the last 4 rows of AFS20.XLS as an example). Choose "Series/Series Information" to modify the data type.

Autobox will read in your XLS file and ask for you to tell it some specifics like when your first data point begins, the frequency and the number of forecasts you are looking for. Our Excel import is used for importing data from spreadsheets that are vertically oriented. This means that series are in columns and observations are in rows extending down. ***You must have the data beginning in row 1 or 2!***

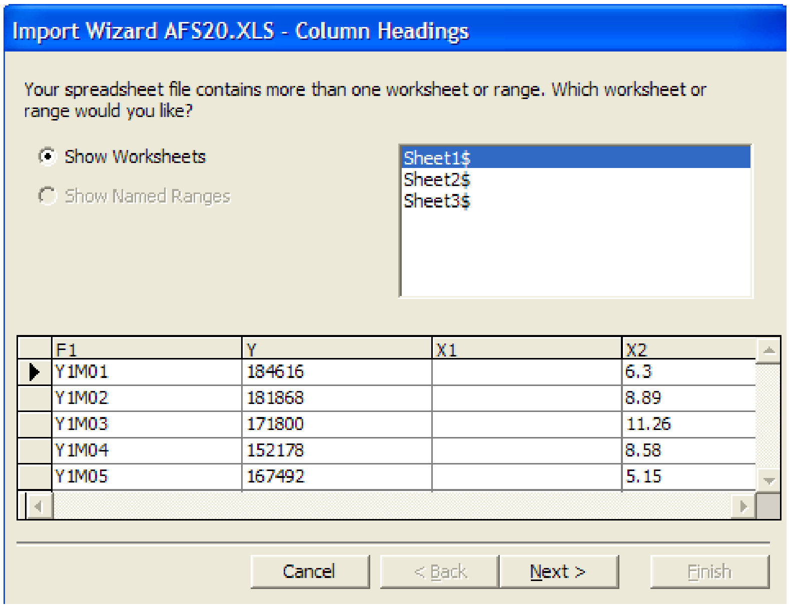
Let's Choose File/Import.



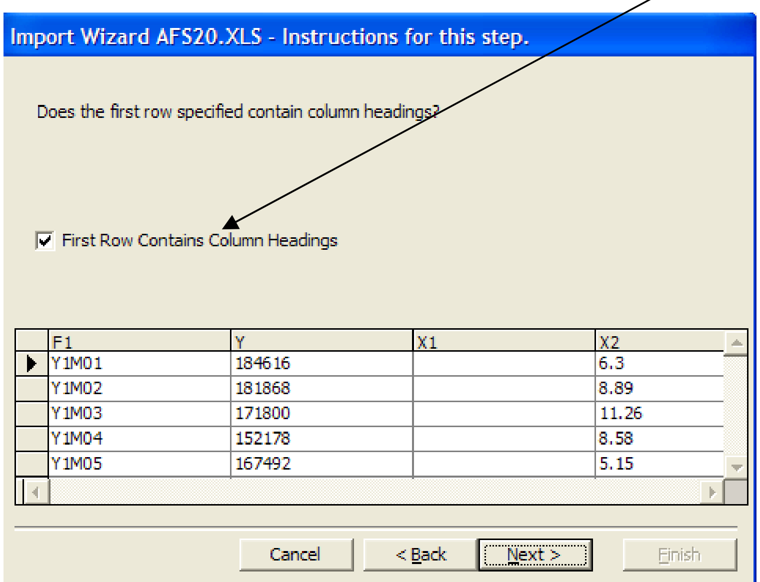
Let's import AFS20.xls(included in the installation) and click "open".



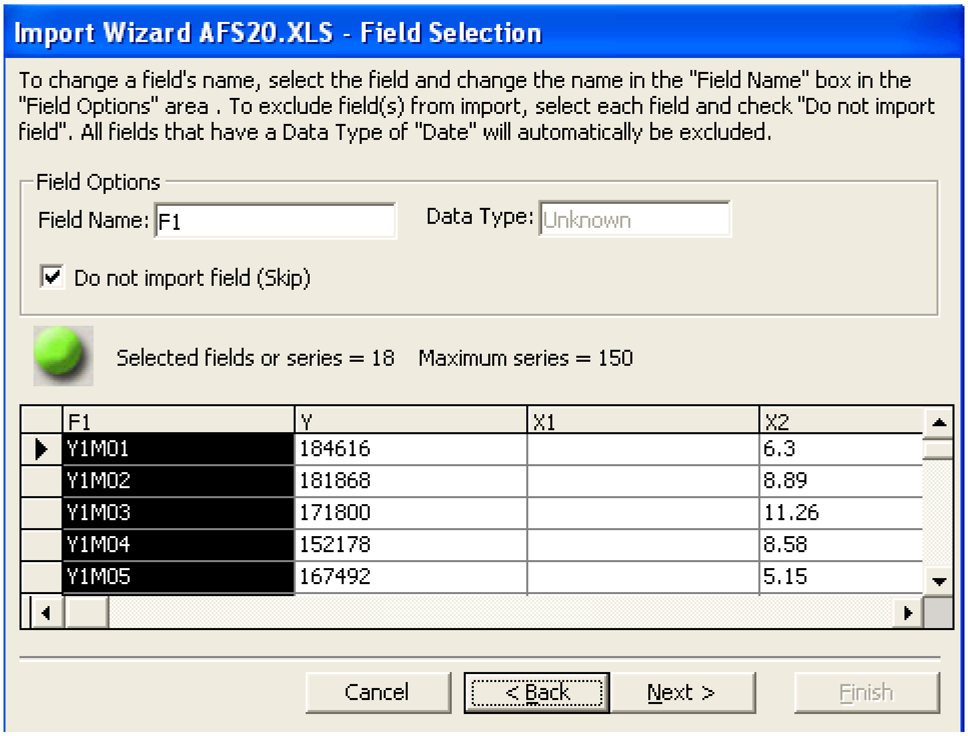
We need to tell which sheet to get the data from and our data is in the first sheet so we just click "next".



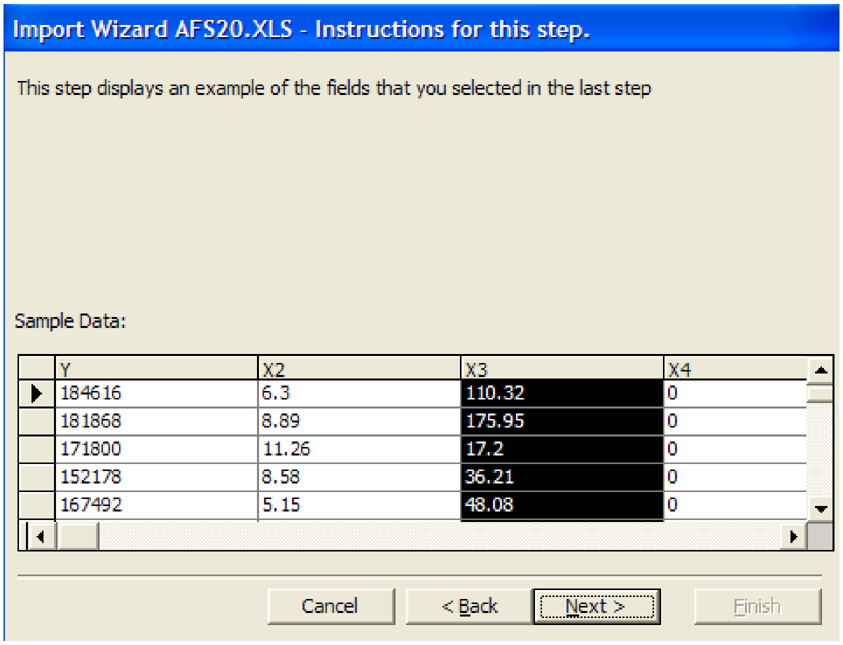
We have titles in the first row so we leave the box checked and click "next".



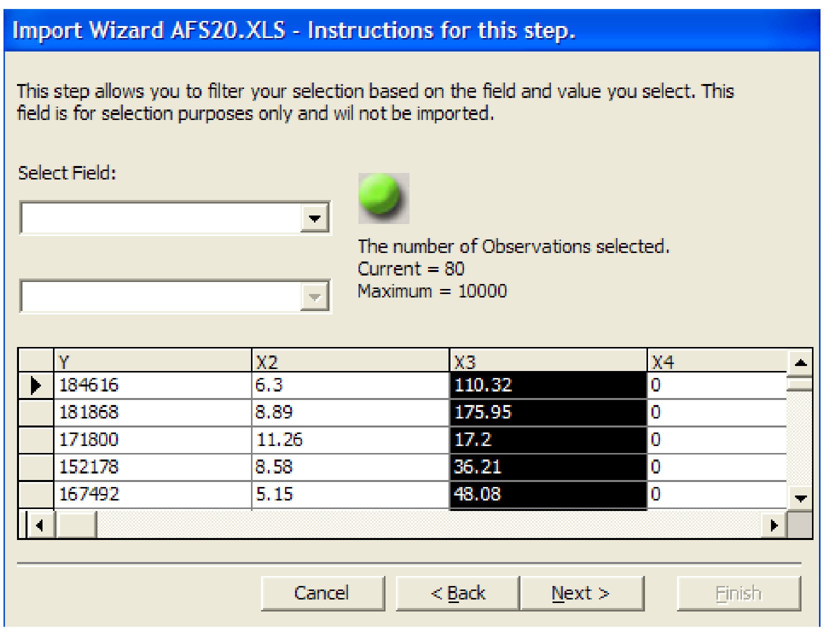
We will skip the first field as dates are not in the needed and defined in the next step so we will click on the box to remove F1(default name assigned by Autobox as it did not have a tile). Not shown here, but we also choose to remove X1 as only half of the observations were included. You click anywhere on the "X1" column to highlight it and then again click on the "Do not import field (Skip)". You might notice that there is only partial data in X1. There is another way to approach this and that is to remove the older data that did not have X1 values if you felt was so important in describing Y. You can run Autobox both ways to see the differences in approaches. Click “next”.



Notice that fields F1 and X1 are now dropped. Click “next”.

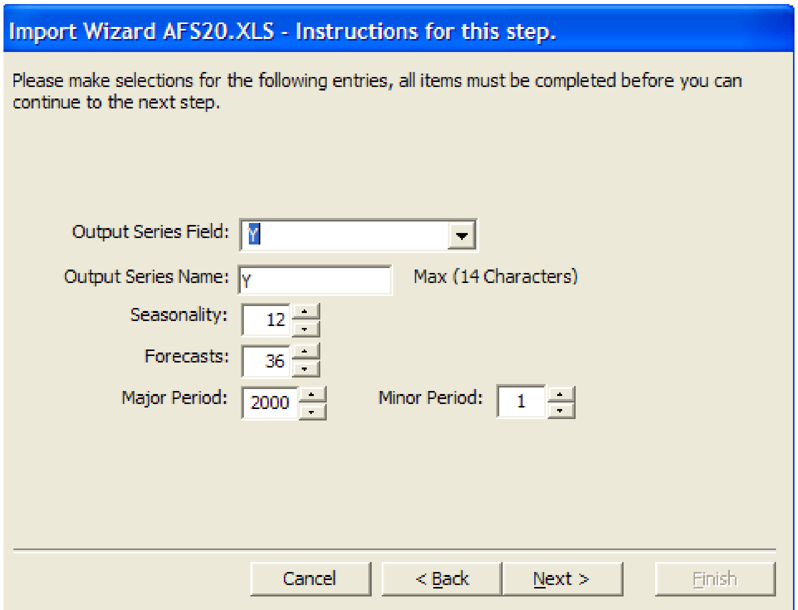


You have the option to filter the rows being read (like Excels' filter feature), but we don't have a need for this here. Click "next".



Click on the arrow and choose Y as the output variable you are trying to predict. Click "next".

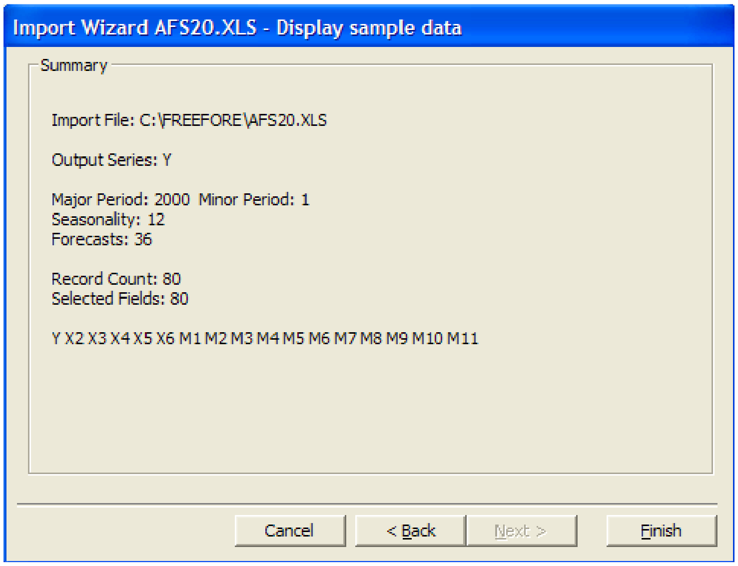
Your Excel spreadsheet may be a causal problem or it just might be many different problems. If your data is not a causal problem and each column is a different problem, just pick any field as the output series field as it won't make a difference. Autobox will save the ASC file(s) into the folder where it imported the XLS file from. Autobox will also automatically generate an ASC file with all columns for causal analysis and also an ASC file for each and every one of the columns in anticipation of the data NOT being a causal dataset.



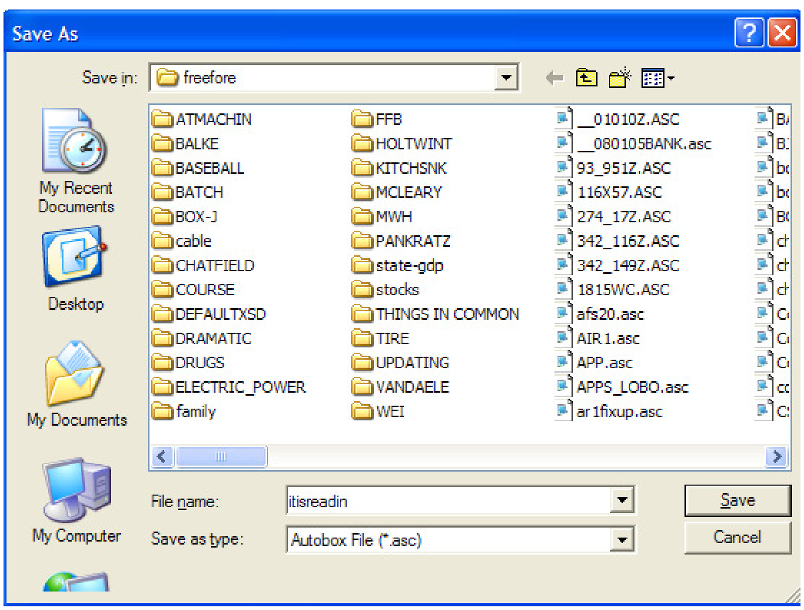
The following explains the individual properties
Output Series Field – For Causal problems, choose the series to be predicted, otherwise choose any field.
Output Series Name – This is prefilled when you import from Excel, but you can change it if you want to.
Seasonality – How often the observations are taken. This is a good time to layout the different seasonalities for Autobox – 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.
Forecasts – The number of forecasts you would like Autobox to generate
Major Period/Minor Period - Let's give a few examples of Major/Minor Period to help you see – For February 1990 monthly data, choose 1990 for the major period and 2 for the minor period. For January 2, 2006 daily data, choose 1 for the major period and 2 for the minor period.
Known issues and limitations:

- The spreadsheet must be closed in order for the import to work correctly. This means it can't be open in Excel while you are trying to import.
- Data must be in columns.
- Series names are limited to 14 characters.
- Vista XLSX files will not import.

Here is an audit of what was read in. Note that the last 4 observations are future values of the causals to help with forecasting. Click “finish”.

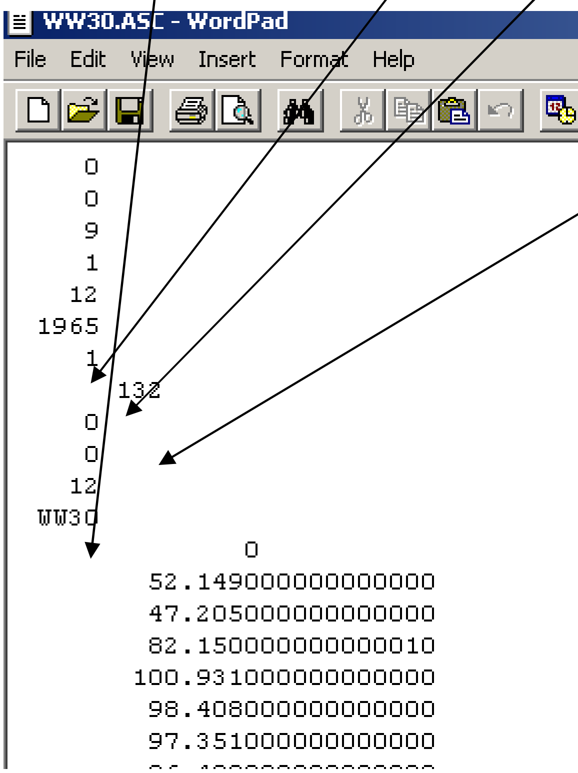


We saved the file as “itisreadin.asc”. Click “save” and a message informs you that it was saved.



Create your own ASC file

We include datasets to show you how to prepare data for Autobox. The files must have a name of “ASC”. Here is an example of “WW30.ASC” file (which you should use as a reference). Information like the # of series, the length of the series, # of forecasts, starting month, starting year are defined here. We have a full explanation in the ***appendix*** on how to create an ASC file. This data preparation takes time so it would make sense to get that process underway!!!



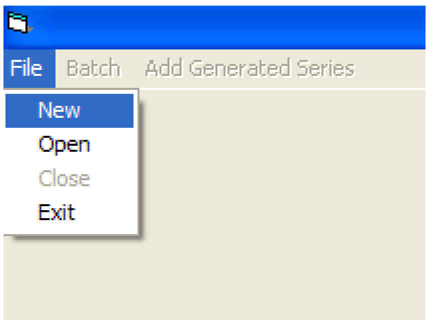
You will need to convert all of your data from SAS, Oracle, Excel, etc into an “ASCII” or “Flat file” format. Each dependent time series will need to be saved into a separate file for each series to be modeled. Note that if you have supporting series e.g. price, promotions, events these also need to be stored in each of the files. The file needs to have an extension of ASC (i.e. ww30.asc). There are some header records needed at the top of the series to define for Autobox batch details of the series (i.e. name, frequency, etc.). Future values of casuals are put at the very bottom of the file.

There are ASC files that came with the installation that you should examine and use to help test things out. Note: There is one very special Autobox trick that we want to discuss up front. For example, if you have a “dynamic” promotion over a period of two weeks (and you had daily data) that causes demand to shoot way up and it slowly ramps down back to the mean (decays). If you specify the causal series name with the words “DYN14” for example, Autobox will react by modeling the promotion to decay over the next 14 periods the promotion was running.

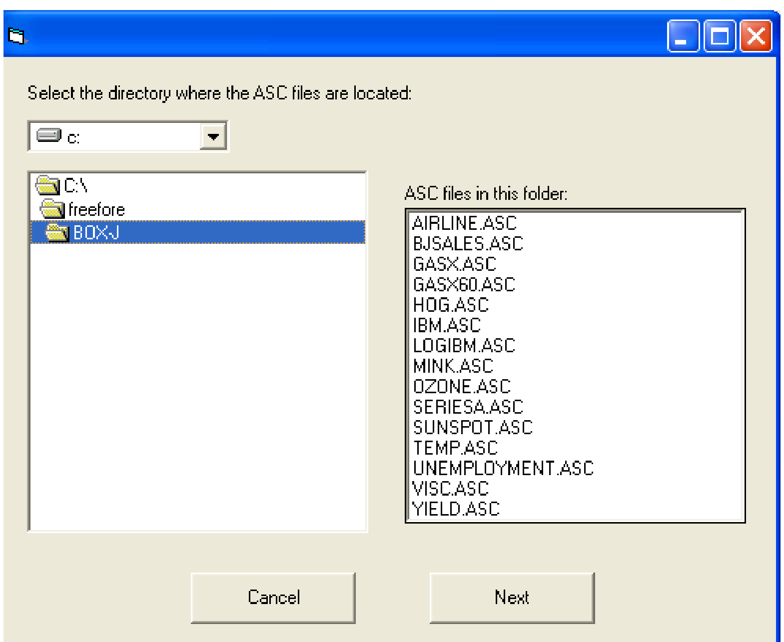
Step 2
Create an Autobox Project

You will need to define where the data is located and the series to be analyzed and where to store the results. This is called an Autobox project file and has an ".XML" file name.

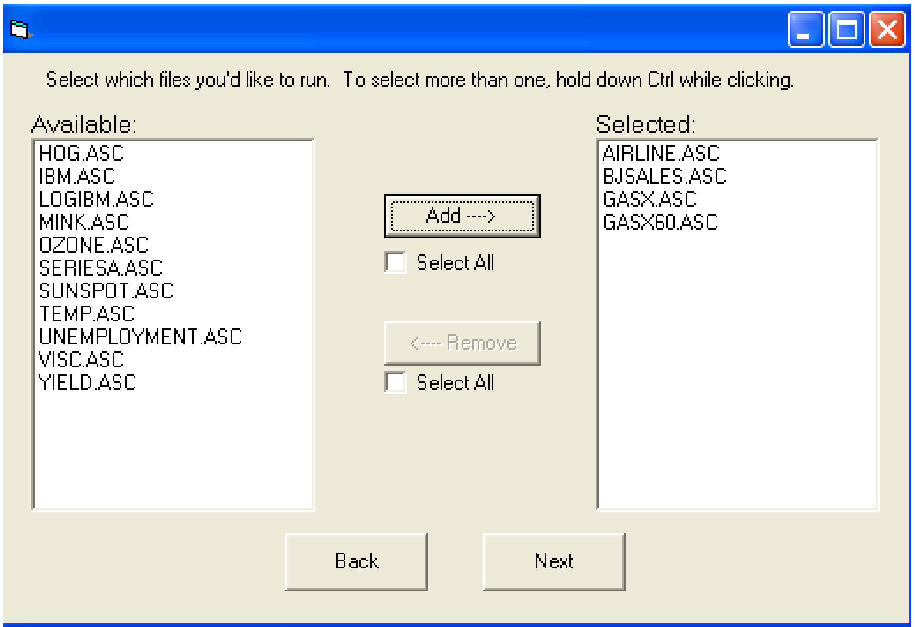
You choose "File/New" to begin.



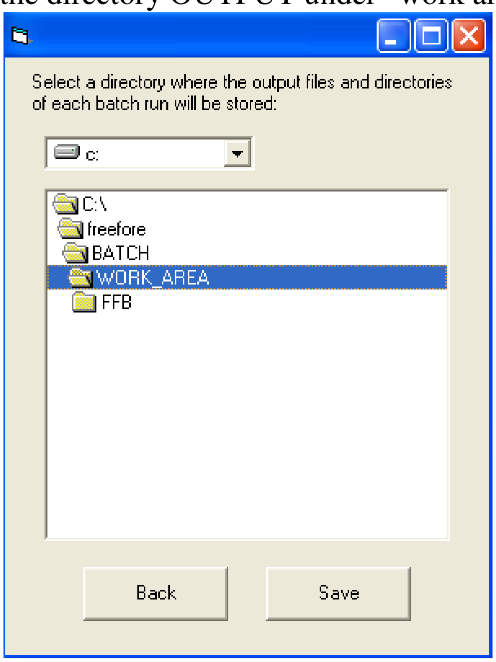
We need to tell Autobox which directory to find the data and we point to all of the Box Jenkins time series from the classic text book. Click 'next'.



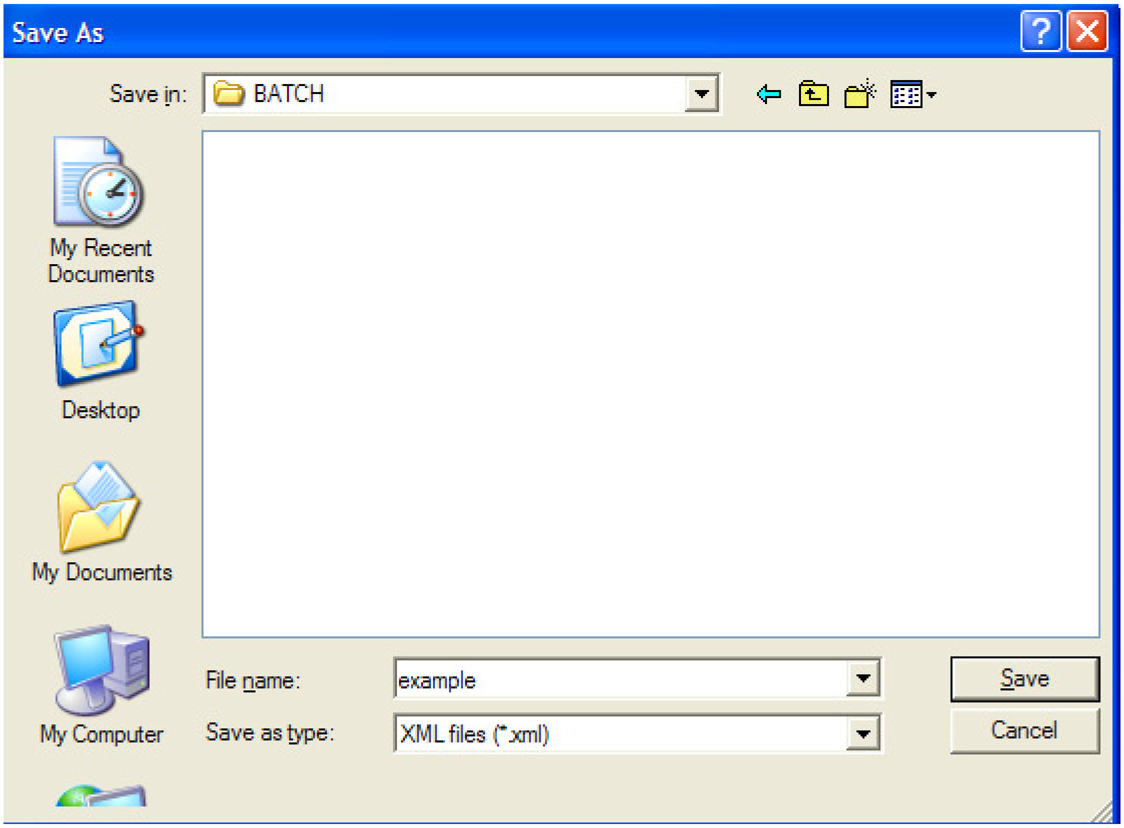
We only wanted to run the first 4 series so we highlighted them and clicked 'add' which moved them to the right and clicked 'next'.



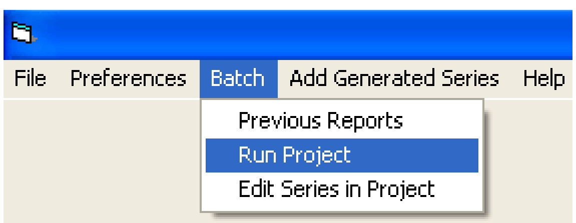
We recommend you save the work files Autobox creates(which aren't really need by you) to a working directory that was created during installation named "work_area". Click "save". Note the directory OUTPUT under "work area". This is where the output will be saved.



You should keep all of your Autobox projects in one spot for organization sake in the batch directory. Click 'save' and a "file successfully message" is generated (not shown below). Click "ok"



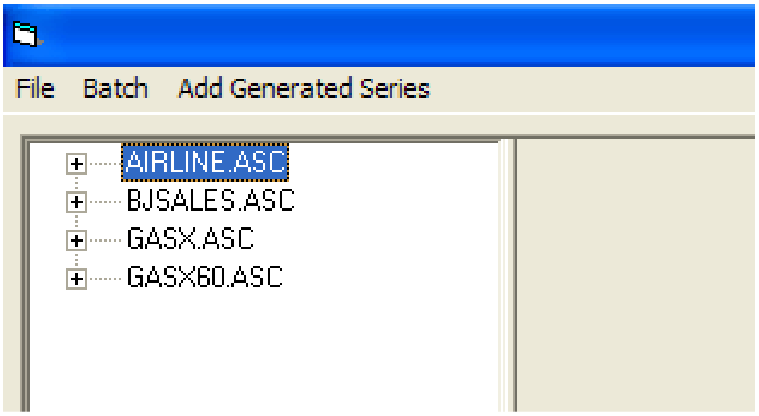
Now we run Autobox by choosing “batch/run”.



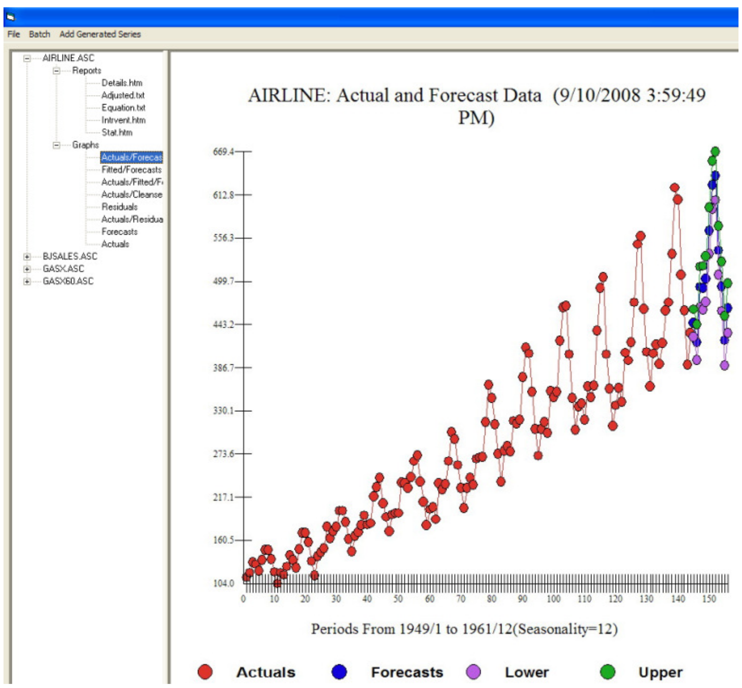
And we have results. The results are written to a flat file on your computer. The forecasts are all in one file (stored in the directory that contains the output) that has the name of the XML file but with a “PRO” file name. So, if the name of the XML file was “example.xml” the file would be named “example.pro”.. Various reports and graphs are also created for your review.

As for the reports: Details.htm has an audit of the modeling steps. Adjust.txt shows the historical cleansed of outliers. Equation.txt shows the final equation used to forecast. Intrvent.htm reports that intervention and the impact. Stat.htm reports the statistics for the model fit and t tests for the variables.

The graphs are self-evident.



This is the graph for the Airlines series.

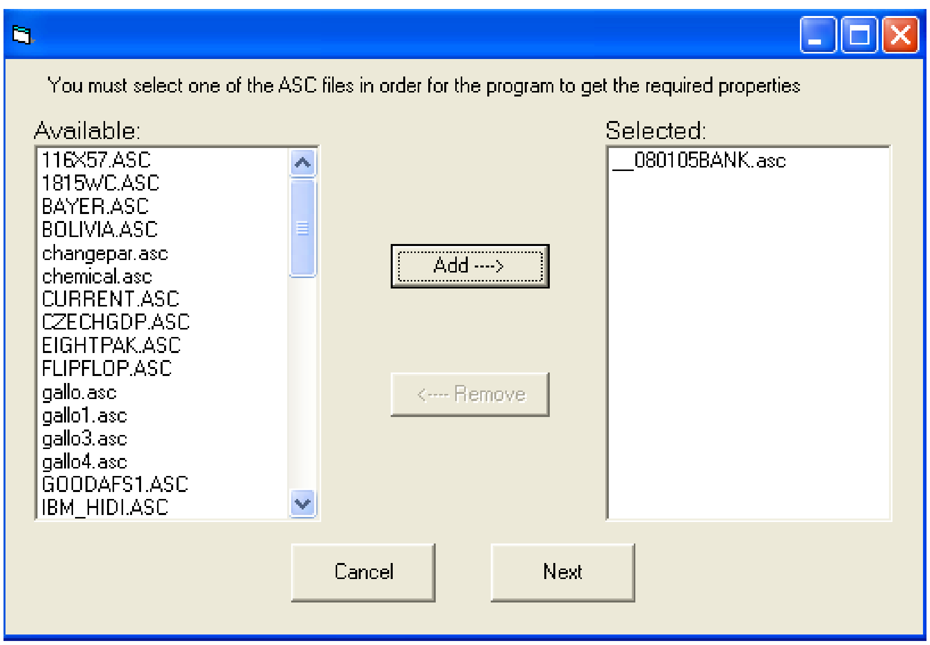


Step 3
Add Causal Variables

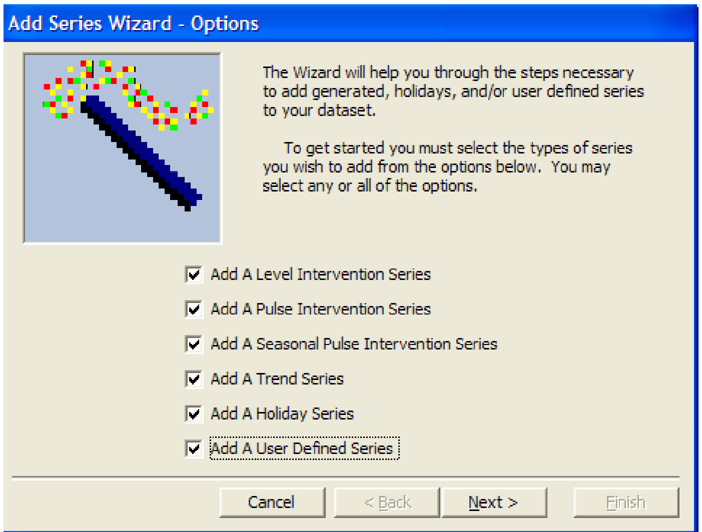
You can use Autobox to add important causal variables to be included in the modeling process. Choose "Add Generated Series" and you will be asked to select a series so that Autobox can find out what the properties of the series are and to be specific the seasonality (i.e. monthly) of the series. Once this is established then Autobox will offer causal variables that are appropriate to your problem at hand. It does not mean that any of these variables will be included in the model just that they will be considered to be in the model. It can be very powerful if you have "apriori" information like known interventions in the history or even in the future and include it to aid the model in the process. Note that there is a directory titled "ASCBACKUP" that contains the ASC files before you added the causal variable in case you are not liking the results WITH the causal.

You should note that for level, seasonal pulse and local time trends you can either allow the event to occur into the future or you can define the point that it has ended.

Here we select a daily series.



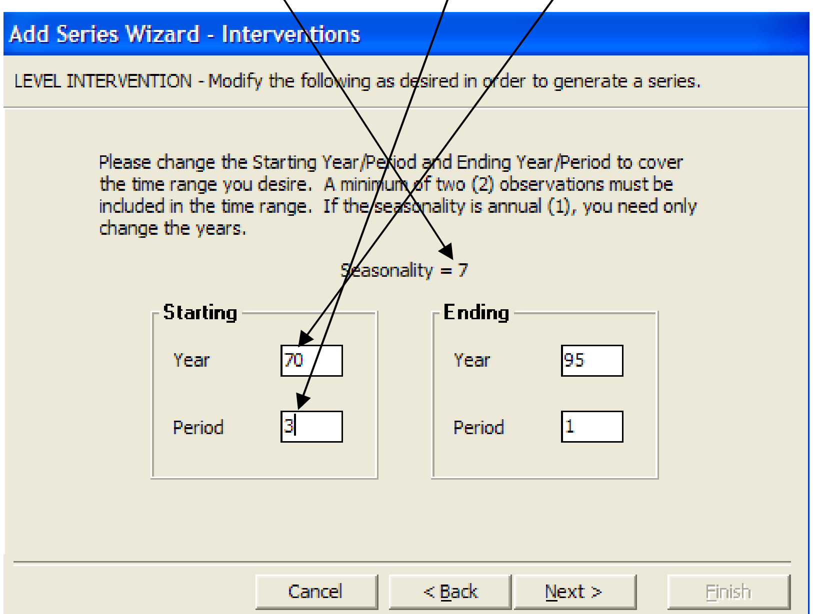
If you had annual data then you would not be allowed to add a seasonal pulse variable into the analysis as it makes no analytic sense to do so as the data has no seasonal fingerprint to it. Here we have all options selected for **illustrative** purposes only. Click “next”.



A level intervention is a sudden increase/decrease in level of the mean that continues into the future. A dummy variable is included in the dataset that looks like this 0,0,0,0,0,0,0,1,1,1,1,1,1,1.

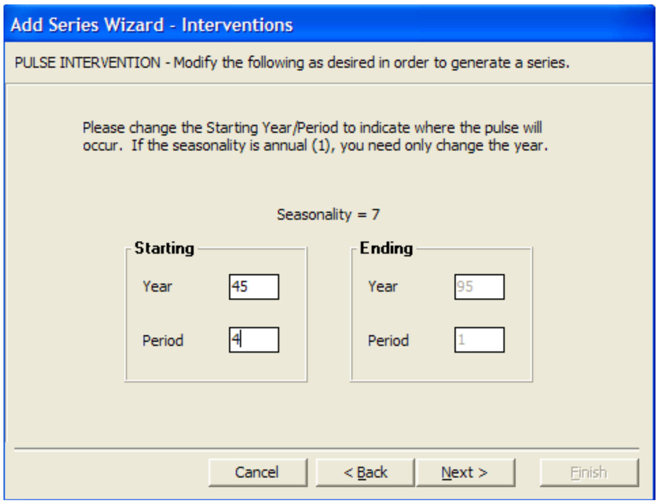
Now, since we are running in batch mode we are suggesting that every series in the batch run had a level shift at the 70th period. This might be the case as a competitor may have gone bankrupt causing your sales to have a one time increase in the mean which would apply to all of your product lines. On the flipside, you might have an outlier for just one series so you might want to move that series into its own project and run it by itself with the level shift variable. If you do not and bring a level shift variable for series that didn't experience a level shift you might change the model created so we recommend that you don't to this. This applies to all interventions (level shifts, pulses, seasonal pulses, local time trends). You can test each way out for yourself and give us feedback from your experience for our curiosity.

Here we are telling Autobox that at the 3rd day in the 70th week had a level shift (Note that the seasonality is 7) and that it continued to the end of the dataset as we did not adjust the ending dates which default to the very last observation. Let's assume we have a shirt factory and we added an embedded electric light in our collars and people like it and end rushing to the store and buying our shirts and continue to show up each day to buy these shirts.



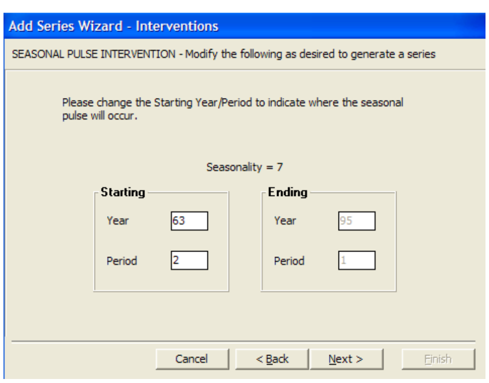
A pulse intervention is a one time sudden increase/decrease which occurs at only one point in time. A dummy variable is included in the dataset that looks something like this 0,0,0,0,0,0,0,1,0,0,0,0,0,0,0.

Here we create a pulse variable at the 45th week and 4th day. Here the factory lost power and no shirts were made at the factory and everyone had a siesta.



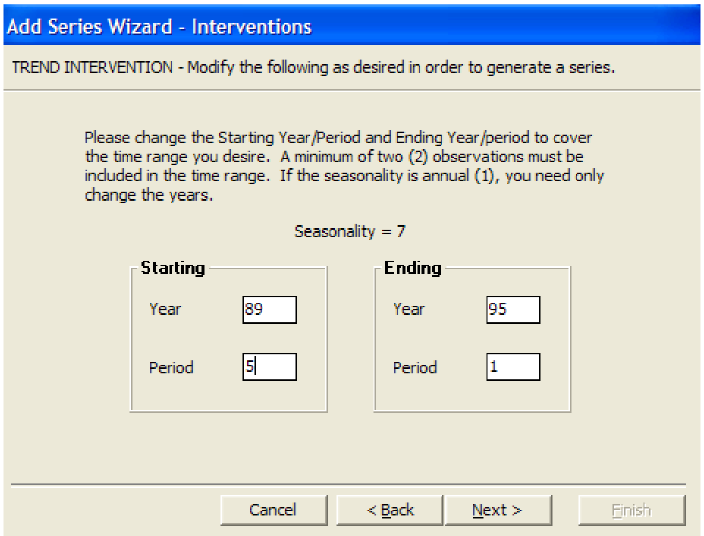
A seasonal pulse intervention is a cyclical increase/decrease which occurs at only one point in time and reoccurs. A dummy variable is included in the dataset that looks something like this 0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,1,0,0,0,0.

Here we create a seasonal pulse variable at the 63th week and 2nd day. Sales are now low on Tuesdays because the local church decided to start a bingo game in the town next door which is hurting our sales.



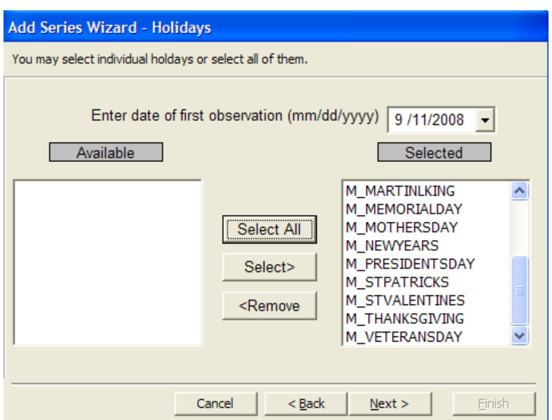
A trend intervention is a sudden change either increasing or decreasing in the slope which occurs at one point in time and into the future. A dummy variable is included in the dataset that looks something like this 0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,1,2,3,4,5,6,7,8,9

Here we create a trend variable at the 89th week and 5th day. Sales are now increasing on steady stream downwards as war has broken out in our county and people are afraid to come out of their houses to buy shirts.



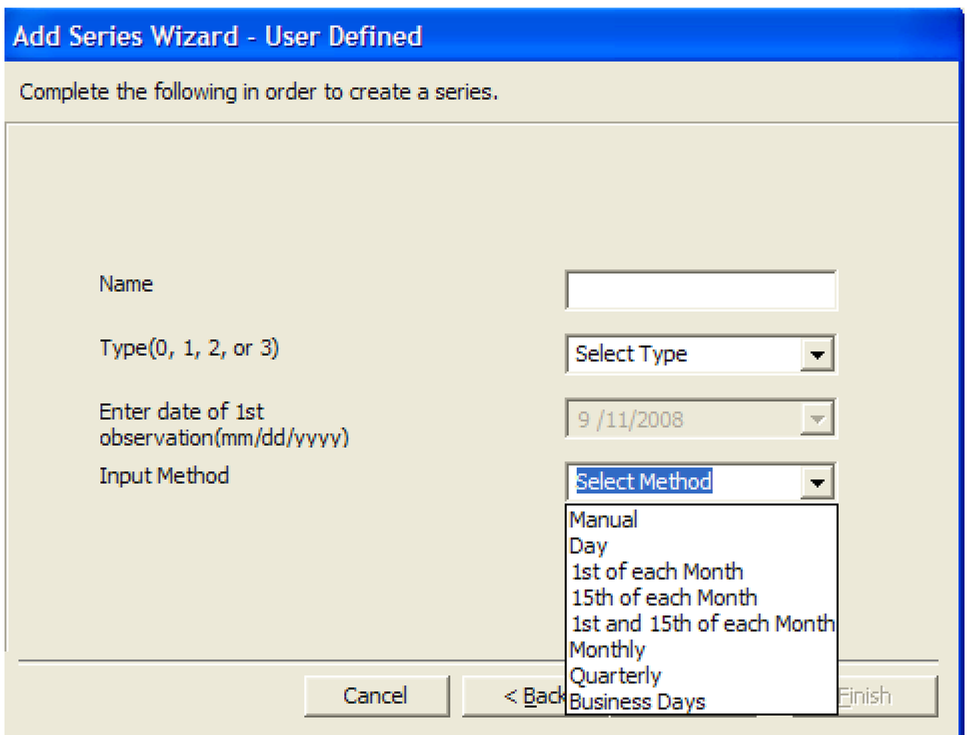
Holidays often have big impacts on many lines of business so we recommend including them to see what effects they may have on our business.

Our example is daily data and we have selected all 16 holiday variables. FYI: For monthly data, you only have the option of selecting 12 holidays as you are limited to their being only 12 months in a year.

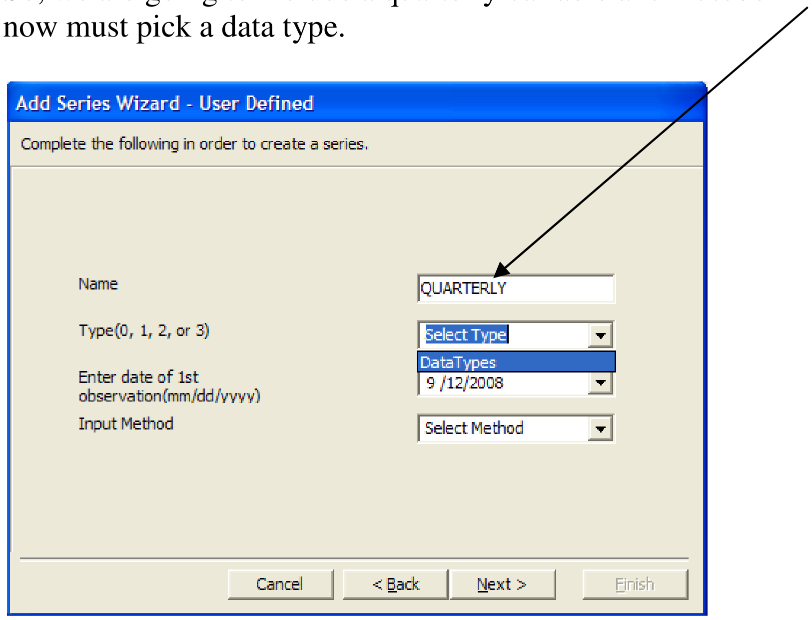


Last but not least is the "User Defined" variable which depending on the seasonality of your data you may have different options of including different variables. For example, while we have daily data there may be some quarterly variation that corresponds to the season due to weather that we would want to account for in the modeling process. So, we would select the quarterly variable.

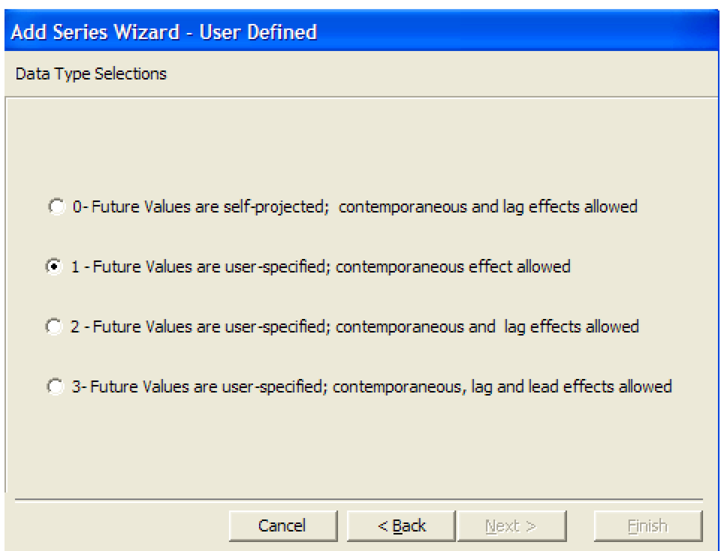
We must **<u>caution</u>** you that there are two options in this example that do not make sense to choose for daily data. There are two on this list you would not want to select as logic would have it. Day and Business Days make no sense to use for daily data. Business days is really only meant for monthly data. A day of the week effect is already included when Autobox analyzes daily data(go back and read page 3 if you are lost here!) so there is no need to include another day variable. Autobox would drop yours if you brought in a duplicate variable anyway as a self-protection scheme.



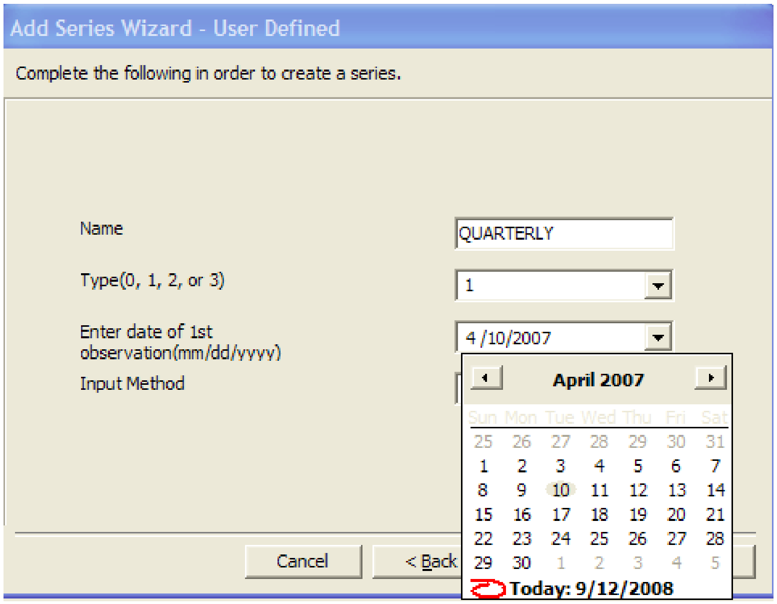
So, we are going to include a quarterly variable and Autobox fills in the name automatically. We now must pick a data type.



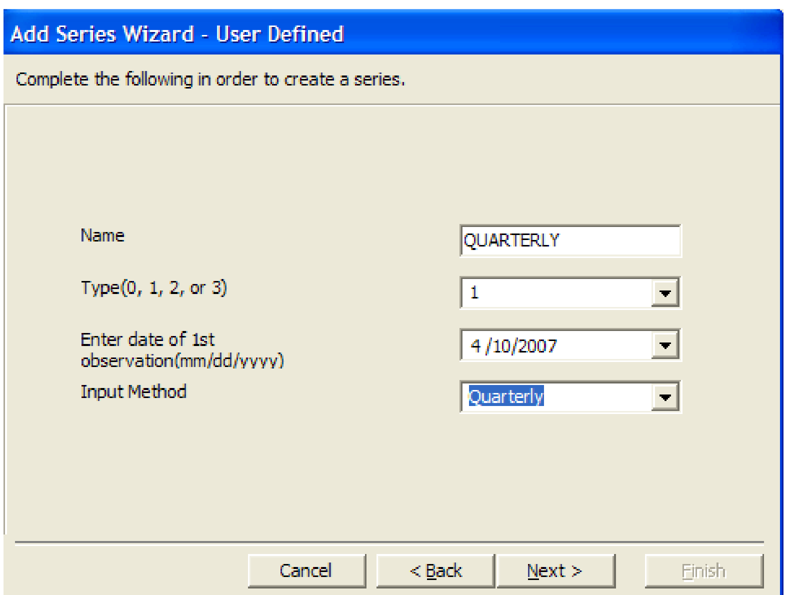
As discussed earlier, it would not make sense to choose 2 or 3 as there are variables generated for each quarter so there is no need to analyze if a lead or lag effect is necessary. Option 0 is not logical as you would not want to forecast a dummy variable. Choose 1 and click "back" to continue (strange as it sounds to click "back", but just do it).



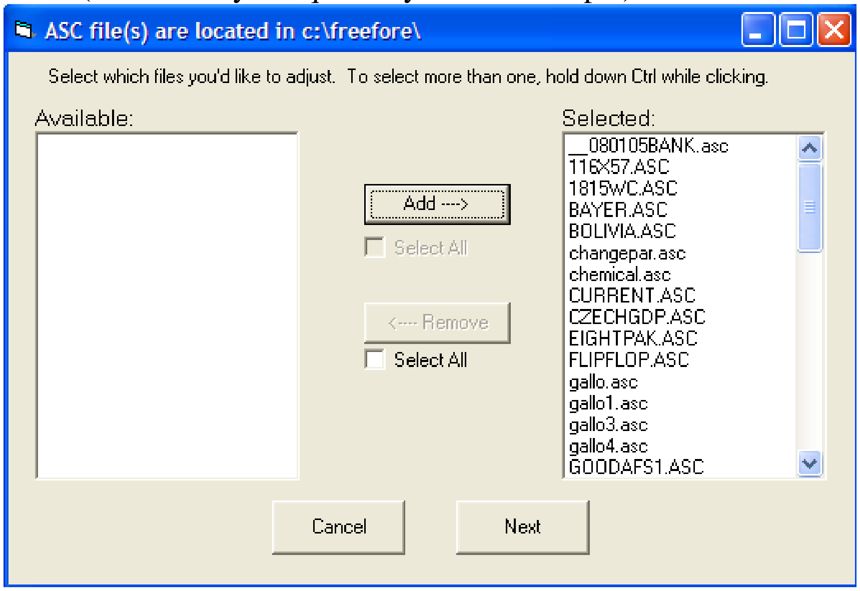
We need to tell Autobox the actual date when the data starts so that the dummy variables can be created properly using the calendar. We click on 4/10/2007



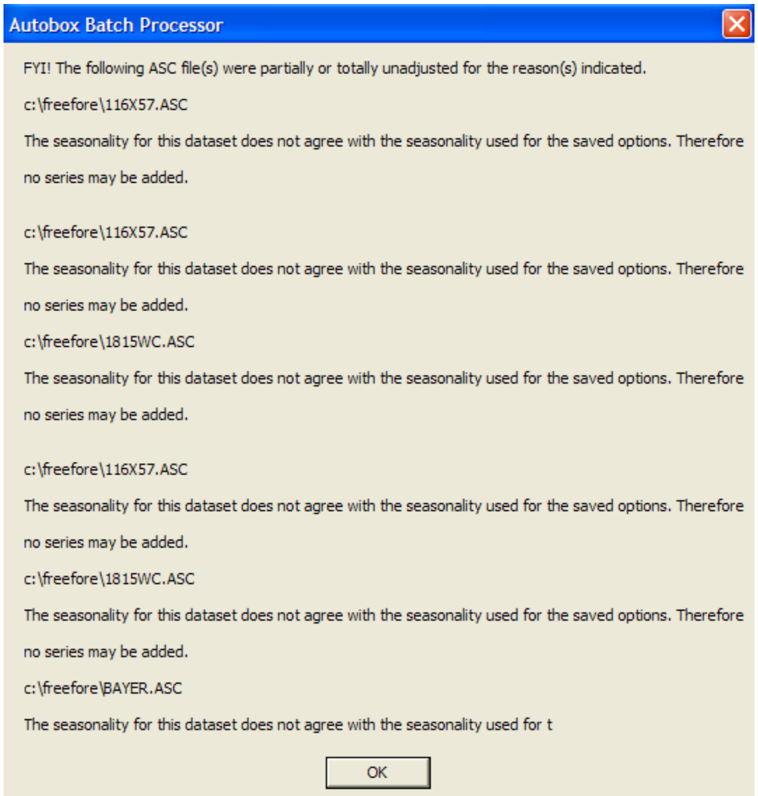
We click "next" and then "finish" (not shown).



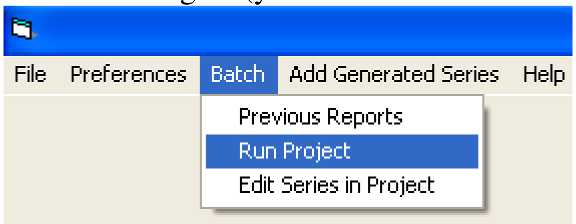
We are asked which series we want to use this new causal variable with and we select all the series (assume they are quarterly in this example) and click “next”.



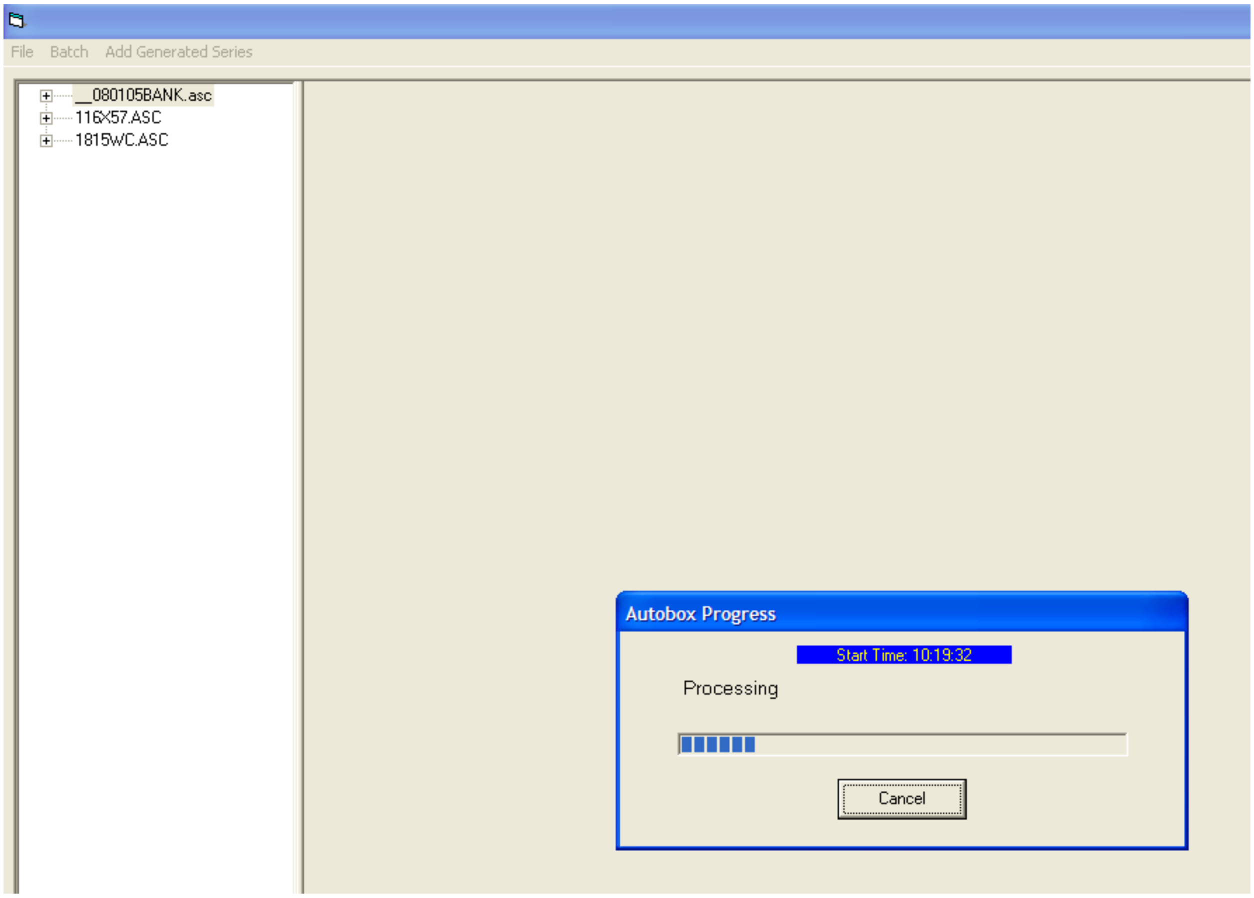
Now, if you didn’t read page 3 and you have series with different seasonality and you just plowed ahead and tried to go through this process without grouping your series into different projects by seasonality you will get this error message. You can still go on to run Autobox, but the reality is that you are not optimally modeling some of the series in this potpourri of different series and should create different project files (XML) for each type of seasonality.


Autobox Batch Processor
FYI! The following ASC file(s) were partially or totally unadjusted for the reason(s) indicated.
c:\freefore\116X57.ASC
The seasonality for this dataset does not agree with the seasonality used for the saved options. Therefore no series may be added.
c:\freefore\116X57.ASC
The seasonality for this dataset does not agree with the seasonality used for the saved options. Therefore no series may be added.
c:\freefore\1815WC.ASC
The seasonality for this dataset does not agree with the seasonality used for the saved options. Therefore no series may be added.
c:\freefore\116X57.ASC
The seasonality for this dataset does not agree with the seasonality used for the saved options. Therefore no series may be added.
c:\freefore\1815WC.ASC
The seasonality for this dataset does not agree with the seasonality used for the saved options. Therefore no series may be added.
c:\freefore\BAYER.ASC
The seasonality for this dataset does not agree with the seasonality used for t
OK

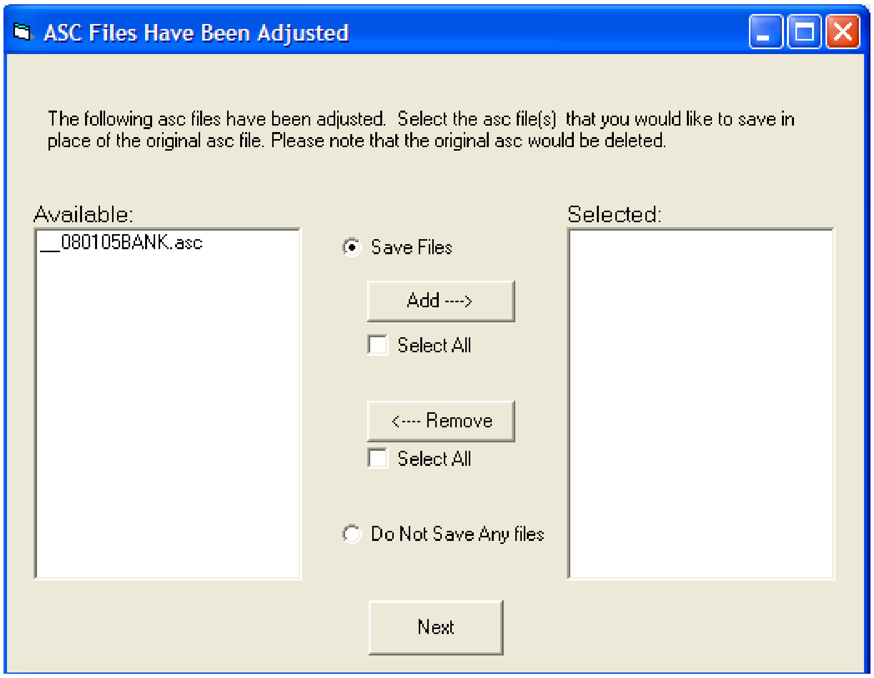
Now we run again (you can cancel the batch run by clicking on 'cancel').



You can see the progress as each series is completed, but you can't review any of the series modeled until after the batch run is completed.



We added casual variables to the process and Autobox asks for confirmation that you would like the ASC files to be adjusted to include these variables. Select the series and click “add” and then “next”.   A message will inform you that ASC files were modified.

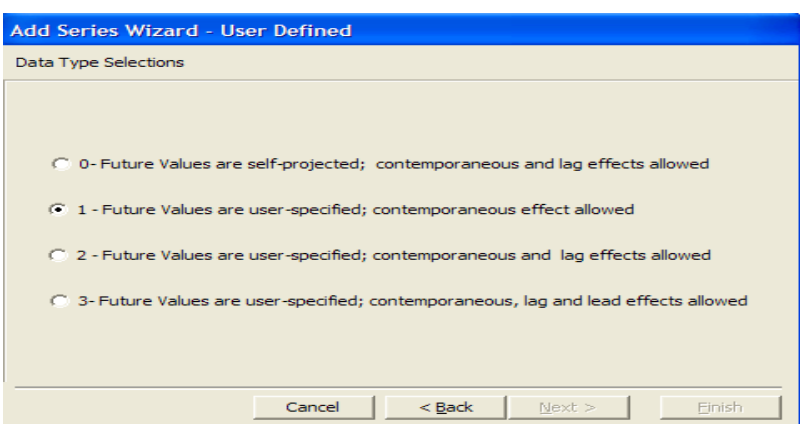


Let's talk about how to build your own dummy variable a little bit more. If you had a promotion starting in February and it ended in April, you would want to put a '1' for those three months to indicate the event.

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

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.



Step 4
Choose what modeling options to run with – Control the statistical options

There are two ways to override the options in Autobox.

1)Engine.XXX – These options are ***NOT*** the default options Autobox uses in automatic modeling (i.e. expert system) process. We would recommend that you use option '2' where you use Autobox's expert system options and choose to override ***some*** of the options using the "AFS" trigger files.

You modify ENGINE.XXX (refer to the section on "Engine.afs" in the Appendix for a complete description of each of the options) and save it as ENGINE.AFS for your overrides to be used. Engine.AFS will be deleted after every run so have a backup of that file.

2)Trigger files - Create "AFS" files that override the Autobox options. Note that you will need to have these special AFS files located in the installation directory and where your data is!!!

Copy specific "XXX" files to "AFS" to make them active.

You can also utilize by copying them to *.afs to have them active (i.e. copy integer.xxx to integer.afs).

Noparcon.xxx – Stops the testing for constancy of parameters (see line 35 page 40 of this manual or line 35 of engine.xxx)

Novarcon.xxx – Stops the testing for constancy of variance (see line 33 page 40 of this manual or line 33 of engine.xxx)

Integer.xxx – Converts forecasts to integers (see line 43 page 42 of this manual or line 43 of engine.xxx)

Positive.xxx – Converts forecasts to positive values (see line 42 page 42 of this manual or line 42 of engine.xxx)

Stepupde.xxx – THIS IS NOT JUST A TRIGGER FILE, BUT ALSO HAS CONTENT. This defines the number of interventions (see line 23 page 38 of this manual or line 23 of engine.xxx For 5 interventions put a '5' in this file)

Numbfore.xxx – THIS IS NOT JUST A TRIGGER FILE, BUT ALSO HAS CONTENT. This defines the number of forecasts (see line 40 page 40 of this manual or line 40 of engine.xxx For 24 forecasts put a '24' in this file)

Nointer.xxx – Stops Autobox from searching for Interactions

Foreconf.xxx – THIS IS NOT JUST A TRIGGER FILE, BUT ALSO HAS CONTENT. This determines the confidence level for the confidence limits (see line 42 page 40 of this manual or line 40 of engine.xxx For 80% confidence limits put a '80.0')

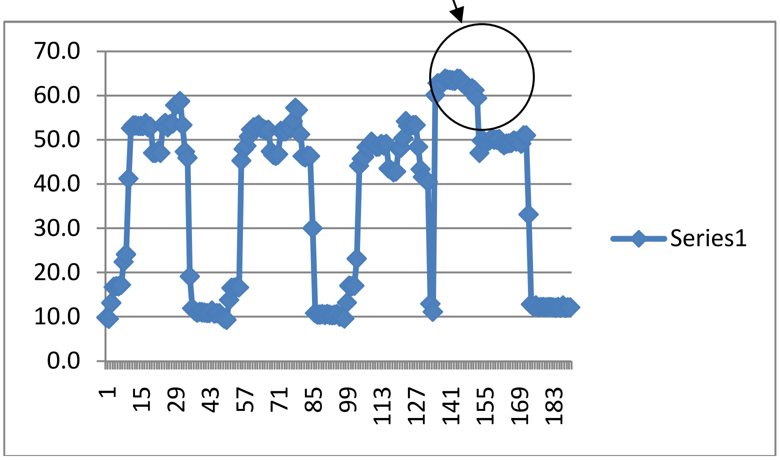
ALSO,

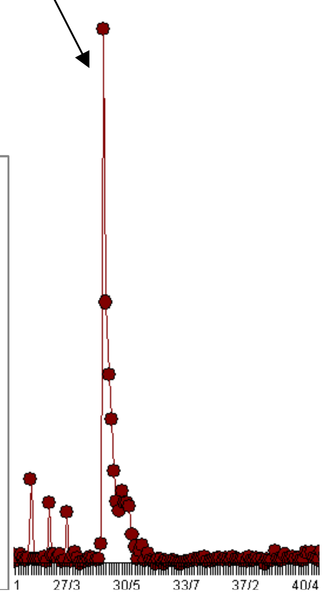
Allin.afs - keeps all causals in the model

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 - turns off looking for interactions
Noseason.afs - Stops testing for seasonal pulse
Notrend.afs - Stops testing for local time trends.
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 this text file place a '20' in it so that Autobox can model it.

Create the file ‘parpharm.afs’ in your Autobox directory and put a ‘1’ for a mean model (Refer to page 37 (objective 1) for the list of override models)

There is one trigger based on the name of the series: There is one very special Autobox trick that we want to discuss up front. For example, if you have a “dynamic” promotion over a period of two weeks (and you had daily data) that causes demand to shoot way up and it slowly ramps down back to the mean (decays). If you specify the causal series name with the words “DYN14” for example, Autobox will react by modeling the promotion to decay over the next 14 periods the promotion was running. Note that the data type(SEE THE SECTION ON CREATING YOUR OWN ASC FILE FOR MORE ON DATA TYPE) MUST be ‘3’ for this to work. Also, if you have a “patch of outliers” that are in the same range (ie all zeroes) then you can use a ‘1’ indicator during that patch. If you have a “patch of outliers” that varies wildly (ie high, low, etc.) then use “DYN” and the length and again data type must be equal to ’3’.





Step 5 Output

When Autobox has finished running, it automatically brings up the reports and graphs generated during the processing and tells you where this data is saved to disk.(They are stored in a "working" directory named "OUTPUT" which is automatically created wherever you have saved your XML file). You will need to delete the files in "OUTPUT" once in a while as they will just keep accumulating!!!! A file with all of the forecasts for the batch run is created and saved to disk with a "PRO" file name. The combined forecast file is in the "OUTPUT" directory and has the same filename as the xml file. So, if the xml file is "test.xml" then the forecast file will be "test.pro".

Now if you ran 1,000 series and you then go to "edit series in project" and choose lets say 5 SKUs you wanted to try a different methodology on and you rerun. NOTE: The old 1,000 reports will not be shown in Autobox, but they will still exist in the "OUTPUT" directory.

Reports

Detail.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.

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

*.pro – The forecasts for all series

Verbal.txt – The model explained in "English"

Pulserpt.csv – see next page

Graphs

JPGS - Actuals/Forecast, Fitted/Forecast, Actuals/Fit/Forecast, Actuals/Cleansed, Residuals, Actuals/Residuals, Forecasts, Actuals, ACF Originals,ACF Residuals

QC Autobox's output by reviewing the residuals graphs and the graphs of the actual fit and forecast. We use Google's Picasa to go through a slideshow(you can also use Windows Slideshow). Just copy "*actfitfore.jpg" and copy those files to the "My Pictures" folder and use Windows to do a slide show to QC the graphs.

Step 6 Exception Reports

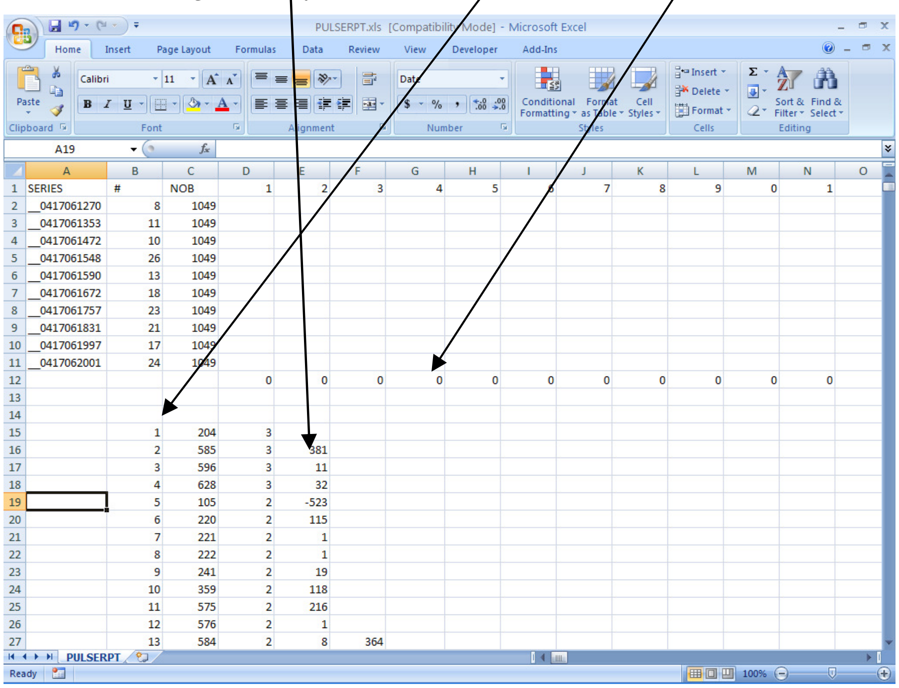
Autobox creates a number of reports to help you understand your data and track the quality of the forecasts being generated so that you can review them for accuracy.

- Outlier Exception Reports - Macro view of where outliers occurred which can suggest that "maybe something happened in the history". It may also trigger you to realize that it was due to a marketing campaign and then realize to bring these in as causal variables when modeling the data into the future
- Forecast Monitoring Report - Quick check to make sure the Autobox forecast is reasonable vs. a simple averaging method of your choosing (ie 4 period equally weighted average) as a baseline for comparison
- Forecasting Fit Exception Report - Compare Autobox fit vs 2 simple methods

Outlier Exception Reports

Pulserpt.csv, Trendrpt.csv, Levelrpt.csv – Log file showing a Table of a pulse(or trend or level) outlier at different time periods(see the first 11 rows to see what it looks like in the picture below). If you have 200 series and you find that 150 have an outlier at time period 02 then it might cause you to think about what happened at this point in time that you failed to include as a variable in the model from the beginning for these 150 series(possibly for all 200 series?). In a couple of steps you can find if these occurrences also occurred annually suggesting that it was a holiday that was omitted in the modeling process. Open the file in Excel and sum each of the periods. Copy and transpose that row to a column. Create the counting numbers next to this column (1,2,3,4, etc). Sort the two columns by the count largest to smallest. Now you have the count of the time period with the most outliers at the top. Below is an example with 10 SKU's with 1,049 daily observations. We did some investigation by subtracting different time periods to identify a missing holiday variable but we didn't find any differentials of 365 so given that we conclude that these are just interventions and not a systematic pattern since 3 out of 10 could randomly occur at a given time period by chance. Note that the series need to all start at the same time period so that the data is aligned! Also, this file is a log file so you need to delete it once in a while to keep the size down.

The level and trend don't show the sign of the coefficient making their report more ambiguous. You can turn off these reports by creating a file named "nocsv.afs" in the installation directory. This file is a log file so you can delete it once in a while to start fresh.



	A	B	C	D	E	F	G	H	I	J	K	L	M	N
1	SERIES	#	NOB	1	2	3	4	5	6	7	8	9	0	1
2	__0417061270	8	1049											
3	__0417061353	11	1049											
4	__0417061472	10	1049											
5	__0417061548	26	1049											
6	__0417061590	13	1049											
7	__0417061672	18	1049											
8	__0417061757	23	1049											
9	__0417061831	21	1049											
10	__0417061997	17	1049											
11	__0417062001	24	1049											
12				0	0	0	0	0	0	0	0	0	0	0
13														
14														
15		1	204	3										
16		2	585	3	381									
17		3	596	3	11									
18		4	628	3	32									
19		5	105	2	-523									
20		6	220	2	115									
21		7	221	2	1									
22		8	222	2	1									
23		9	241	2	19									
24		10	359	2	118									
25		11	575	2	216									
26		12	576	2	1									
27		13	584	2	8	364								

Early Warning System Report

The report "earlysig.txt" is created to help find out if the last observation is "out of control". The report tells you the name of the series, the last observations number, the probability of out being out of control, the observation, and what the observation was expected to be. There is one record added to this file every time (up to 20 series and then the file is purged to avoid a large file—the batch version will continue to write out to this file so if you have 50,000 series this file will have 50,000 records). Note: This report is saved to your installation directory NOT the OUTPUT directory.

You can bring this file into Excel and sort on probability (ascending) to find the series that seem to be "out of control". I ran the series inlier and there was nothing found to be "out of control in the last observation" as you can see here. However, I went and I changed the last observation from a 9 to 5,555 and then reran Autobox. The second row shows a low p-value to show that there is something wrong. It prints out what the value should have been here. This file is a log file so you can delete it once in a while to start fresh.

EARLYSIG.TXT

ITEM	NOB	PROBABILITY	ACTUAL	EXPECTED
inlier			9.0000000000	
inlier	10	.0000	5555.00000	5.0000000000

Forecast Monitoring Report

This report gives you a way to check that the forecast from Autobox compared to a simple method to locate if there are any forecasts that are very low or very high. The report takes the ratio of the two forecasts so that you can locate very low or very high ratios to inspect. This report has many false positives as a simple method is in fact that, simple.

You can create a file named "fore-mon.afs" in the installation directory and if you wanted your simple forecast benchmark to be a weighted average of the last 4 periods you would have a total of 5 rows in the file:

4

.25

.25

.25

.25

You would save this file and run Autobox and a report would be created named "fore-mon.csv". You can sort on the field "Autobox/base" to identify low and high differences in baseline forecasts. This file is a log file so you can delete it once in a while to start fresh.

ITEM	NOB		BASE	AUTOBOX	DIFFERENCE	AUTOBOX/BASE
REGION1	147		1025	500	-525	0.49
REGION10	147		0	341		
REGION11	147		225	58	-167	0.26
REGION12	147		12516	11477	-1039	0.92
REGION13	147		7642	7072	-570	0.93
REGION14	147		2342	4739	2397	2.02
REGION15	40		0	19		
REGION16	142		150	0	-150	0
REGION17	147		7213	6319	-894	0.88
REGION18	131		424	611	187	1.44
REGION19	91		0	0		

Forecasting Fit Exception Report

This report gives you a comparison of the fitted MAPE between the Autobox model, the naive model and a mean model. You create a file in the installation directory named "mapecomp.afs" and a report will be generated named "mapecomp.csv". The Autobox model column could be sorted and used to compare to the simple methods to provide a check that Autobox is doing better than the simple methods. Of course, like always you can have false positives. This file is a log file so you can delete it once in a while to start fresh.

ITEM	NOB	AUTOBOX MODEL	NAIVE	MEAN
REGION1	147	0.374	0.911	0.724
REGION10	147	0.225	0.376	0.162
REGION15	40	0.046	0.224	0.228

Measuring Accuracy

We have a utility that generates a report called "summape.txt". It will provide an accuracy report based on withholding observations and seeing how Autobox would have performed if it didn't have that data point. The steps to create this report are as follows:

1)Create a file named makemape.afs in your installation directory

2)Run Autobox for the series you want to review accuracy

3)You will see that there are now many files that have been created that used the name of the output series (ie Ted) and generated series that begin with "Ted". If you had 192 monthly observations, Autobox will withhold 12 months of data and create a file that is ready to forecast using 180 observations forecasting out for 12 periods. It will also have a file that has 179 observations forecasting out 12 periods. There are also files created from 178 back to 12 observations all forecasting out 12 periods. A similar set of files are generated that are forecasting out 11 periods starting from month 181 back to month 11 forecasting out 11 observations. A similar set files is created so that forecasts for 10 months to 1 month out are also created. So, this process will take a while to run!

To summarize, these files each have had a successive historical observation removed from the "actual" and placed into the withhold area of the asc file. Autobox won't use this data to build the model, but use it to evaluate the accuracy.

NOB	MAPE	SUM OF ACTUAL	SUM OF FORECAST
548	0.6%	48473.8	48746
549	0.6%	48470.5	48768.2
550	0.7%	48466	48810.4
551	0.7%	48461.5	48773
552	0.7%	48456.3	48810.8
553	0.9%	48450.3	48861.4
554	0.7%	48445	48775.5
555	0.9%	48438.3	48847.6
556	0.7%	48433.3	48785.9
557	0.7%	48428	48776.9
558	0.7%	48422.5	48778.7
559	0.8%	48418.3	48787
Overall	0.7%		

Discussion of Saving and Reusing models

If you need to reduce the time spent on remodeling when new data comes in, you can tailor the process to NOT to remodel and just reuse the previous model and use it with the latest data. The ***downside*** is that you are not using the latest information to make decisions which we will caution you only once on.

How can you reuse models in order to cut down the time expense of remodeling? In order to do so, you need to model and save it for later use.

To save your model(s):

- Create a file named savemod.afs in the installation directory
- Run Autobox batch and the models will be created and saved as *.MOD
- Create a zip file of all of the *.mod files for later use

To reuse your model using an existing model three things are necessary:

- We will assume it has been some time (1 month?) since you created the model so you should unzip the MOD files to the directory.
- To tell Autobox batch that there is a model saved and should be used:
    a. Edit Engine.xxx and change line 37 to a '1' and save as Engine.go or
    b. create a trigger file called USEMOD.AFS in the installation directory
- The actual model files (*.mod where * is the series name) must exist
- If you want to override AFS' expert system, edit Engine.xxx and modify any of the conditions as you see proper (see Appendix for more on Engine.afs) and save the file as "Engine.afs". Also, see predict.xxx for an example where we have restricted the analytics to just forecasting (note that estimation is turned off by setting it to '0'). Also, note that the structure is hierarchial. If an item is set to zero than any item that is indented underneath is also not going to be executed.

Here is how to reuse models for the "Comprehensive and Multiple batch modes"

In order to let Autobox batch know if you would like to save a model you need to place a 0,1,2,3 after the name of the zip file. A '0' runs the entire process and saves the model only. If we wanted to save our model we could put a '0' as seen here.

Issue the command at the DOS prompt:

OUTPUTATCH1 WEI3 0 (see next section on changing default modeling conditions).

When a '0' is used to run OUTPUTATCH1 a file is saved with the model. The name of the file is the series name with a "MOD" as the extension (i.e..We30.MOD). The MOD files are copied to MOZ files for backup purposes as the MOD files are deleted after being used via the model reuse (type 1,2,3). The contents of the saved model file (MOD) is discussed in Appendix 1 in the section "How to specify your own model".

The three options to put at the end of the line which use an existing model

Option/Goal	File Used
1. SIMPLY FORECAST	PREDICT.XXX
2. TUNE COEFFICIENTS AND THEN FORECAST	TUNE.XXX
3. TUNE COEFFICIENTS & POSSIBLY REMODEL THEN FORECAST	REMODEL.XXX

If we run Autobox batch with a '1' to run the R-squared is NOT CALCULATED as estimation is bypassed via PREDICT.XXX.

We now have a file named ALLCAST.USE which contains all the forecasts. The amount of time eliminated to develop the forecast versus just modeling and forecasting can be as much as 30:1 reduction, but in most cases it will be 20:1, a significant time saving feature especially when the modeling process is complex (i.e.. regression equations, daily time series, etc.).

Similarly, if you wish to have Autobox batch optimize the model coefficients and then forecast put a '2' at the end of the line:

OUTPUTATCH1 WEI3 2

If you wish to both TUNE and RECONFIGURE the model then put a '3' at the end of the line:

OUTPUTATCH WEI3 3

Reporting Errors

If you are experiencing an issue that you wish us to investigate, we will need to get some details from you. Call us if you are confused how to do any of the procedures. PHONE 215-675-0652

Create a dummy file in your directory named “SNOOP.AFS” and “PRESNOOP.AFS” to tell Autobox that you want to run some diagnostics and then run the process again.

1)Take a picture of the error using the Print-Screen key of the AFS message (if any) (Hit the “print screen” key and paste into a Word document and attach in an email)

2)The ASC file(s) that had been in the batch run. The ?.lst file that catalogues the series being analyzed.

3)Any and all files created in the directory after snoop.afs was created. You can sort the directory by time and zip these files using Windows Explorer.

4)All “AFS” files in the directory

5)Engine.go (if it existed)

6)A directory of all files - Go to the DOS prompt by Clicking “Start/Run” and type ‘cmd’ or ‘command’ to get the dos prompt and then type “dir *.* > direct”) and send the file named “direct”

7)Attach OUTPUTATCH.DLL

Make sure you delete the file called SNOOP.AFS when you wish to resume normal processing.

APPENDIX

Creating your own .ASC file

The ASC file contains some header records and the time series. This is really an easy procedure; but the information and data must be entered in a text file in a very specific order in a single column. We would prefer to first show you an example of an ASC file. There is some text below this example that explains the options for each line in great detail. Call us if you are the slightest bit confused.

<u>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.</u>

```
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
```

(Lines 1–3) 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

(Lines 4–11) These are the 8 "Data Properties". They define the characteristics of the series like the seasonality, beginning year, beginning period, number of observations.

(BJ07) 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

(Data) 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

This is a summary of the ASC files major sections:

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, in the same order as the data names)

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

Name	Description
OBJECTIVE(1)	Sets forth the model conditions as indicated by the following: 0 = Totally Automatic NONCAUSAL MODELS IN AUTOBOX BATCH 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) 91 = FIND HIDDDEN SEASONALITY THEN SET TO “BEST” 97 = IDENTIFICATION ONLY 98 = HOLT-WINTERS TREND PLUS MULTIPLICATIVE SEASONAL FACTORS (TREND FORM) CAUSAL MODELS IN AUTOBOX BATCH MEMORY 51 = REGRESSION 52 = REGRESSION WITH AR(1) CORRECTION 53 = STEPDOWN REGRESSION 54 = STEPDOWN REGRESSION WITH AR(1) CORRECTION STARTING MODEL SUPPLIED: 99 = STARTMOD.123 199 = STARTMOD.123 + SIM IF STARTING MODEL SUPPLIED = 99 or 199, STARTMOD.123 is required. 200 = TOTALLY AUTOMATIC + ABOXLITE model is developed
OBJECTIVE	No longer used leave as a ‘0’

(2)	
OBJECTIVE (3)	No longer used leave as a '0'

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

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). **Also, by specifying the date (January 1, 2006) with the name in this format “__010106Y11” and line 5 has a ‘5’ or a ‘7’ then daily effects such as “day of the week”, “week of the year”, “day of the month” are analyzed. Note:Day of the month only applies for ‘7’.**

DATATYPE Structure:

Name	Description
DATATYPE	Integer value for series type. This can be 1 of 4 values 0 = Future Values are self-projected; contemporaneous and lag effects allowed. All output series must be 0. 1 = Future Values are user specified; contemporaneous effect allowed. 2 = Future values are user specified; contemporaneous and lag effects allowed . 3 = Future values are user specified; contemporaneous, lag and lead effects allowed. A ‘0’ tells Autobox you want it to forecast future values of the causal. Types ‘1’,’2’, and ‘3’ are user supplied future values. In terms of effects on run-time Type ‘3’ would take the longest, Type ‘1’ takes the shortest and ‘2’ and ‘3’ would be similar.

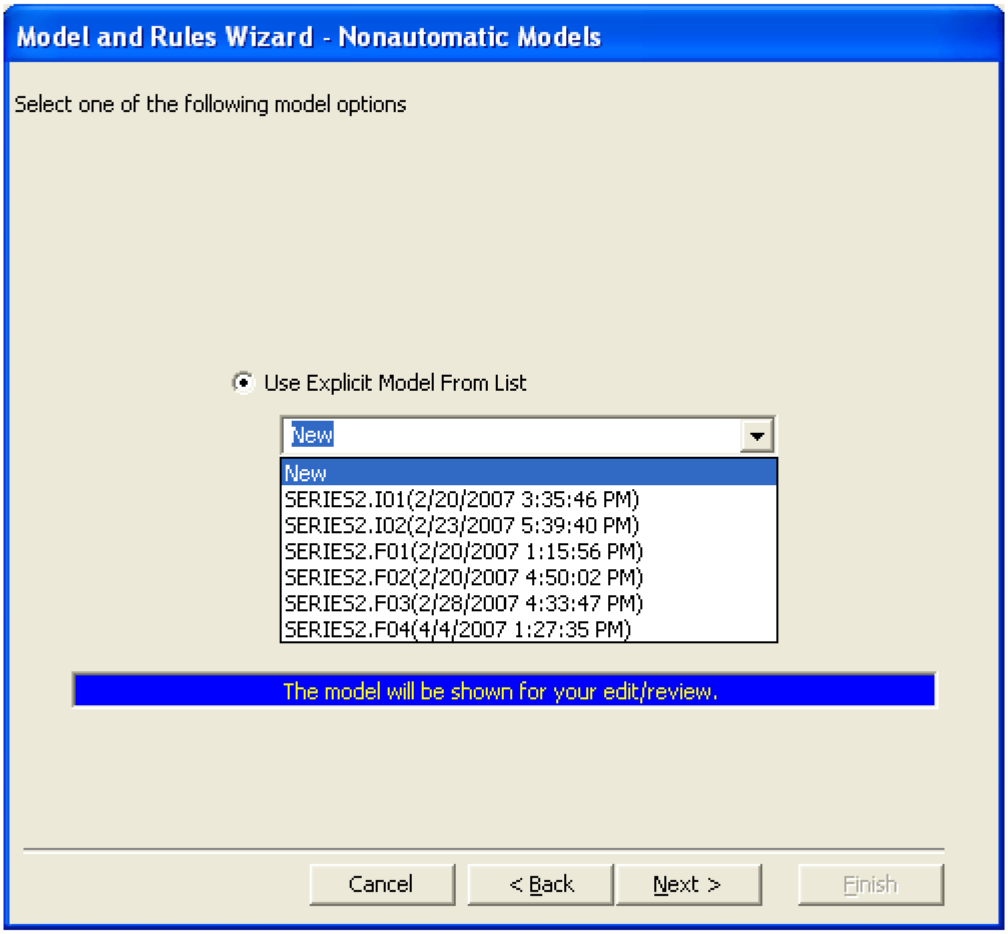
The following is an example of an .ASC file for a noncausal (single) series{annotations are not included in the file}:

0	(objective(1) indicates totally automatic modeling)
0	(objective(2) indicates use default conditions 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)
0	(data type)
15	(historical data – 67 observations)
14	
6	
.	
.	
12	

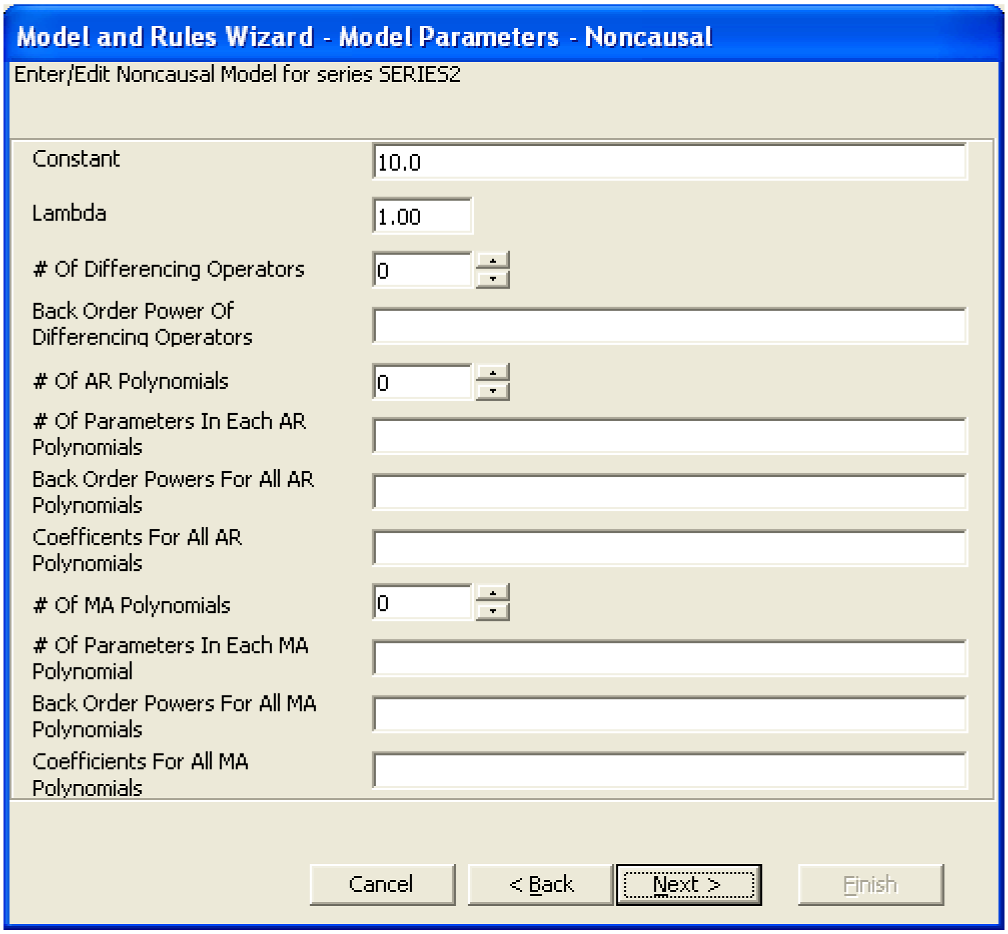
How to Specify Your Own Model

This is where an expert can use his expertise.

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:



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



Please note that for Causal series we will show model parameters for each of the input (independent) series as well as the output(dependent) series. You can copy that same model to the other input variables.

ENGINE.AFS

Listing of ENGINE.XXX which is copied to ENGINE.AFS to micro-manage the dll

```
NAME OF FILE CONTAINING INITIAL MODEL              |STARTMOD.123
 MONTE CARLO SIMULATION ENABLED (0=NO 1=YES)       |   0
  VARIANCE OF THE NOISE SERIES                     |01.0
  SEED VALUE TO START ( 0 FOR CLOCK )              |   0
NON-CAUSAL:                                        |
 EVALUATE THE NEED FOR INTERMITTENT DEMAND MODEL   |   0
CAUSAL:                                            |
 MAXIMUM LEAD FOR SERIES WITH POSSIBLE LEAD EFFECT|   0
 USE DIFFERENCE FACTORS FROM ARIMA IN TF MODEL     |   0
 CONSTRAIN ALL USER CAUSAL COEFFICIENTS IN MODEL   |   0
ENABLE MODEL ESTIMATION                            |   1
 # OF GROUPS IN POOLED-CROSS SECT. T/S (IF ANY)    |   0
  SAMPLE SIZE IN EACH OF THE GROUPS                |
 % CHANGE IN ERROR SUM OF SQUARES                  |00.0
 % CHANGE IN THE PARAMETER VALUES                  |00.1
 MAXIMUM # OF ITERATIONS                           |  20
 NECESSITY TEST:                                   |   1
  CONFIDENCE LEVEL FOR NECESSITY                   |95.0
 SUFFICIENCY TEST:(STOCHASTIC STRUCTURE )          |   1
  CONFIDENCE LEVEL FOR SUFFICIENCY (SS)            |95.0
 SUFFICIENCY TEST:(DETERMINISTIC STRUCTURE )       |   1
  CONFIDENCE LEVEL FOR SUFFICIENCY (DS)            |95.0
  MAXIMUM NUMBER OF OUTLIERS TO BE IDENTIFIED      |   7
  INCLUDE PULSE VARIABLES                          |   1
  INCLUDE STEP VARIABLES                           |   1
   MINIMUM NUMBER OF OBSERVATIONS IN GROUP         |   9
  INCLUDE SEASONAL PULSE VARIABLES                 |   1
  INCLUDE LOCAL TRENDS                             |   0
  ENABLE AUTOMATIC FIXUP FOR FIXED EFFECTS         |   0
  NUMBER OF LAMBDA VALUES TO EVALUATE IN EST       |   1
   LAMBDA VALUES TO EVALUATE (4F4.0)               |1.00
  DISCRETE CHANGE TEST FOR VARIANCE                |   1
   CONFIDENCE LEVEL FOR VARIANCE TEST              |99.0
   MINIMUM NUMBER OF OBSERVATIONS IN GROUP         |   5
 CONSTANCY OF PARAMETERS:                          |   1
  CONFIDENCE LEVEL FOR CONSTANCY TEST              |99.0
 STORE MODEL FORM (YES/NO)                         |   0
 DISPLAY MANAGEMENT ANALYSIS                       |   0
ENABLE MODEL FORECASTING                           |   1
 # OF FORECAST VALUES TO COMPUTE                   |  13
 CONFIDENCE LIMIT FOR THE FORECASTS (%)            |95.0
 CONVERT THE FORECAST VALUES TO POSITIVE VALUES    |   0
 CONVERT THE FORECAST VALUES TO INTEGERS           |   0
 CONVERT PULSE AT LAST OBSERVATION TO STEP         |   0
 CONVERT PULSE TO SEASONAL PULSE                   |   0
```

```
DISPLAY IDENTIFICATION INFORMATION                |   1
 DISPLAY ACF TABLE DURING IDENTIFICATION          |   1
 DISPLAY ACF GRAPH DURING IDENTIFICATION          |   1
 DISPLAY CCF TABLE DURING IDENTIFICATION          |   1
 DISPLAY CCF GRAPH DURING IDENTIFICATION          |   1
 DISPLAY PREWHITENING MODEL(S)                    |   1
 DISPLAY PREWHITENING MODEL(S) EQUATION           |   1
 DISPLAY IDENTIFIED MODEL                         |   1
 DISPLAY IDENTIFIED MODEL EQUATION                |   1
DISPLAY ESTIMATION INFORMATION                    |   1
 DISPLAY ESTIMATED MODEL PARAMETERS               |   1
 DISPLAY ESTIMATED MODEL EQUATION                 |   1
 DISPLAY FIT VS ACTUAL TABLE                      |   1
DISPLAY DIAGNOSTIC CHECKING INFORMATION           |   1
 DISPLAY RESIDUAL ACF TABLE                       |   1
 DISPLAY RESIDUAL ACF GRAPH                       |   1
 DISPLAY RESIDUAL CCF TABLE                       |   1
 DISPLAY RESIDUAL CCF GRAPH                       |   1
 DISPLAY NECESSITY TEST RESULTS                   |   1
 DISPLAY SUFFICIENCY TEST RESULTS                 |   1
 DISPLAY VARIANCE STABILITY TEST RESULTS          |   1
 DISPLAY OUTLIER TEST RESULTS                     |   1
 DISPLAY CONSTANCY TEST                           |   1
DISPLAY FORECASTING INFORMATION                   |   1
 DISPLAY FORECAST MODEL PARAMETERS                |   1
 DISPLAY FORECAST MODEL EQUATION                  |   1
 DISPLAY TABLE OF TRANSFORMED FORECAST VALUES     |   1
 DISPLAY TABLE OF FORECAST VALUES                 |   1
```

INTRODUCTION TO THE ENGINE.XXX FILE

HOW THE ENGINE VERSION WORKS

The Autobox batch version uses internal choices to model. If you would like to customize some of the choices, you can do so here by Editing a file named “Engine.xxx” and saving it as “Engine.go”. After the batch process runs it renames it as “Engine.afs” and will continue to use those overrides as long as “Engine.afs” exists. You can rename it to “Engine.old” if you want to stop using your overrides. The information seen to the right of the "|" character is where Autobox batch searches during an execution. A ‘0’ means not used and a ‘1’ means it is used.

Line 1 – Do not change this line

```
NAME OF FILE CONTAINING INITIAL MODEL            |STARTMOD.123
```

This points Autobox batch to the file use as the starting model. This is all handled in the “BAT” file process for you.

Line 2

```
MONTE CARLO SIMULATION ENABLED (0=NO 1=YES)      |0
```

Choose "1" to have Autobox batch create a time series based on a user specified model.

Line 3

```
VARIANCE OF THE NOISE SERIES                     |1.0
```

If you choose to simulate in line 2 then you will need to specify the amount of variance. The more variance the more randomness in the simulated data. The standard is 1.0.

Line 4

```
SEED VALUE TO START ( 0 FOR CLOCK )              |0
```

To start the simulation process, the program needs a starting point to iterate from. If you use the same seed value, then you can replicate your data. If you want persistent random simulated data, then use the "0" to use the time of day which is a pretty good random approximation.

Line 5 - Title only

NON-CAUSAL:

Line 6

EVALUATION THE NEED FOR INTERMITTENT DEMAND |0

If a 0 exists then the series will not be considered for intermittent demand modeling. If a 1 exists then if >50% of the data has a zero then intermittent demand modeling will proceed.

Line 7 - Title only

CAUSAL:

Line 8

```
MAXIMUM LEAD FOR SERIES WITH POSSIBLE LEAD EFFECT  |1
```

Enter the number of periods to search for a lead variable. This feature, if enabled, will evaluate possible leads for all input series that have names beginning with the string "MOVE". For example sales might arise the week before a holiday. To detect this temporal structure this option has to be enabled and the potential indicator series must have a name like MOVEX1 or MOVEX2.

Line 9

```
USE DIFFERENCE FACTORS FROM ARIMA IN TF MODEL     |1
```

A number or researchers have found that while differencing factors are important to Transfer Function Identification, they may be counter-productive when included in the actual estimated model. Since this version of AUTOBOX BATCH is rich in model augmentation procedures(step-up.... sufficiency), it may be possible to simplify the initial structure and then evolve via model augmentation procedures to the final model. The user has the choice of including the ARIMA model differences in the initially identified Automatic model. This feature only effects Automatic Transfer Function initial model identification.

Line 10

```
CONSTRAIN ALL USER CAUSAL COEFFICIENTS IN MODEL    |0
```

This option allows the user to constrain the final model such that all coefficients for user input series or model developed series, such as interventions, will be kept REGARDLESS of level of significance. Thus those coefficients that AUTOBOX BATCH would have considered insignificant and would have replaced with a zero instead become part of the model. This could be particularly interesting to the user who would like to see the affect of his causative input series instead of having them ignored be- cause of their perceived insignificance. This is then a CONSTRAINED REGRESSION option where certain coefficients are a permanent part of the model. Note that this does not constrain the actual values of the parameters.

Line 11

```
ENABLE MODEL ESTIMATION                            |1
```

Estimation and diagnostic checking represent the second phase of the B-J modeling procedure. The estimation option computes the model coefficients and the residual statistics via non-linear least squares.

Line 12

```
# OF GROUPS IN POOLED-CROSS SECT. T/S (IF ANY)   |   1
```

POOLED-CROSS SECTIONAL TIME SERIES

Consider the case where you have n distinct time series (max of 3) and you wish to test the hypothesis that the individual ARIMA models are equal to each other vs. the alternative that at least 1 model differs from the rest. This requires that 1 model be specified for all n and parameter estimation be done locally and compared to a global or generic set of coefficients. A STARTING MODEL MUST EXIST as this will be used. If AUTOMATIC MODELING IS DISABLED and this answer is greater than one (1) the program will: 1. disable all model modification options (sufficiency, necessity etc) 2. expect the time series to be a concatenated series of the n distinct time series and will estimate parameters without using the last set of group i to predict the start of i+1 , where i goes from 1 to 2 (max 3 groups). Hypothesis testing is done by summing the error sum of squares from the n local estimations (done separately) and divide by the total degrees of freedom to obtain a denominator mean square error. The numerator mean square error is the differential error sum of squares (composite estimation less the sum of the locals, divided by the number of groups see JOHNSTON : ECONOMETRIC METHODS 1963 Page 137)

Line 13

```
SAMPLE SIZE IN EACH OF THE GROUPS                |
```

Enter the number of observations in each of the groups. If you specified n groups in the concatenated series then you must now enter the n values indicating the number in EACH group, in the same sequence the groups were entered into the concatenated series.

Line 14

```
% CHANGE IN ERROR SUM OF SQUARES                 |00.1
```

Parameter estimation is an iterative process that stops when one of three conditions is met. If the relative change in the residual sum of squares is less than the value specified here, then the parameter estimation will stop.

Line 15

```
% CHANGE IN THE PARAMETER VALUES                 |00.1
```

Parameter estimation is an iterative process that stops when one of three conditions is met.. If the relative change in each individual parameter is less than the value specified here, then the parameter estimation will stop.

Line 16

```
MAXIMUM # OF ITERATIONS                          |20
```

Parameter estimation is an iterative process that stops when one of three conditions is met.. If the number of iterations in the estimation process exceeds the value specified here, then the parameter estimation will stop.

Line 17

```
NECESSITY TEST:                                   |    1
```

Choose "1" to enable diagnostic checking. One phase of diagnostic checking entails deleting unnecessary parameters from the model. This normally requires you to re-specify the model form, and then to estimate this model. With this option on, the program automatically deletes the non-significant parameters (one at a time) and re-estimates the model. The test for necessity is performed by examining the T-ratios for the individual parameter estimates. Parameters with nonsignificant coefficients will be deleted from the model.

Line 18

```
CONFIDENCE LEVEL FOR NECESSITY                   |95.0
```

If you elected to turn the parameter deletion option on, then you have the option of specifying the confidence level value that will be used to determine the significance of a parameter. For example, 95% indicates that the program should delete all parameters that are not significant at the 95% level.

Line 19

```
SUFFICIENCY TEST(STOCHASTIC STRUCTURE)           |    1
```

The diagnostic checking phase requires the analyst to make sure that the residuals can not be predicted from themselves (ACF) and in the case of multivariate models the pre-whitened input series (CCF). In the latter case, the test also has to be reversed, i.e. the residuals can not predict the pre-whitened input, otherwise the condition of feedback is identified. The residuals are tested for white noise in much the same way as model identification is performed. If there are patterns in the residual autocorrelations and partial autocorrelations, then the analyst may need to add parameters to the model. One follows the pattern recognition rules described above when adding parameters to the model. A "yes" will request the program to guide these adjustments.

Line 20

```
CONFIDENCE LEVEL FOR SUFFICIENCY (SS)            |95.0
```

If you elect to augment an estimated model with additional ARIMA structure as evidenced by the sample ACF and PACF of the residuals, you must indicate the significance of a parameter. For example, 95%

Line 21

```
SUFFICIENCY TEST(DETERMINISTIC STRUCTURE)        |    1
```

Outliers can occur in many ways. They may be the result of a gross error, for example, a recording or transcript error. They may also occur by the effect of some exogenous intervention. These can be described by two different, but related, generating models discussed by Chang and Tiao (1983) and by Tsay (1986).They are termed the innovational outlier (IO) and additive outlier (AO) models. AUTOBOX BATCH uses the AO approach due to estimation considerations. ARIMA modeling may be deficient when the series has been intervened with. This program will test the residuals from the ARIMA model for possible outlier (intervention) variables. We suggest that you modify either your model or your time series for any outlier variables that may be found. The automatic intervention detection option automatically determines the need for intervention variables using the residuals from an estimated model and automatically introduces them into the model.

Line 22

```
CONFIDENCE LEVEL FOR SUFFICIENCY (DS)          |90.0
```

If you select the outlier detection option, then you must specify the confidence limit to be used for detecting possible outlier variables. For example, .80 indicates that the program should identify all outliers that are significant at the 80% level.

Line 23

```
MAXIMUM NUMBER OF OUTLIERS TO BE IDENTIFIED    |   5
```

You may elect to limit AUTOBOX BATCH to a certain number of empirically identified outliers. As delivered, the standard product is limited to a maximum of 5 input series in a transfer function thus this integer can’t exceed that limit. AFS sells larger versions which allow up to 150 input series. This feature allows the user to control the incorporation of potentially spurious interventions leading to numerical instability.

Line 24

```
INCLUDE PULSE VARIABLES                        |   1
```

Select "1" to include pulse interventions.

Line 25

```
INCLUDE STEP VARIABLES                         |   1
```

Choose "1" to include step interventions.

Line 26

```
MINIMUM NUMBER OF OBSERVATIONS IN GROUP        |   2
```

The number entered determines how many successive values that are on a different level, before Autobox batch will consider there to be a level shift.

Line 27

```
INCLUDE SEASONAL PULSE VARIABLES               |   1
```

Choose "1" to include seasonal pulse interventions.

Line 28

```
INCLUDE LOCAL TRENDS                                |0
```

Choose "1" for Autobox batch to identify multiple trends.

Line 29

```
ENABLE AUTOMATIC FIXUP FOR SEASONAL DUMMIES         |0
```

Choose "1" to enable this option to test for the presence of a SEASONAL DETERMINISTIC VARIABLE which has a zero/one pattern according to the following:

a "1" in the corresponding period and a "0" in other periods

The formal test is outlined in Franses paper in the International Journal Of Forecasting, July 1991, pp 199-208 (see the help for the associated Confidence value indicates that the program should add all parameters that appear to be needed at the 95% level.

Line 30

```
NUMBER OF LAMBDA VALUES TO EVALUATE in EST          |0
```

Enter the number of values to be included in line 59. If you indicate 3, you must supply three values in the next prompt.

Line 31

```
LAMBDA VALUES TO EVALUATE                           |1.0 0 -.5
```

INITIAL MODEL IDENTIFICATION

The lambda value is the transformation parameter. In other words, the value that you specify here will be the exponent in the power transformation. Each data point in the time series is raised to the power lambda. The acceptable range of lambda values is from 1.0 to -1.0. For example, a lambda of 1.0 indicates that the original series is to be analyzed, a value of 0.0 indicates that the natural log of the series is to be analyzed, and a lambda of -1.0 indicates that the inverse of the series is to be analyzed. It represents the power transformation that is to the observed series in order to induce variance stationarity. Note however that this should only be applied when the non-constant variance is caused by a correlation between level and variability.

Line 32

```
DISCRETE CHANGE TEST FOR VARIANCE                   |   0
```

The residuals from a model may not have constant variance and consequently the standard estimation may be deficient. One form of non-constant variance is treated by the Box-Cox or lambda transformations. However, a different kind of non-constancy can occur if a series is affected by a period of unusual volatility. Consider the case where an upward trending series has a residual variance of say 10 for the first half and a variance of the residuals of 20 for the second half. It would be totally incorrect to either ignore the change in variance or to use the power transform procedures of Box-Cox. The suggested procedure is to simply identify a model and compute a vector of residuals. By breaking the residuals into consecutive but non-overlapping sections one can perform the standard F test for variance change. The time period with the greatest F value is then a potential point of variance change. There are cases in which the user has an "a priori" knowledge of the weights and wishes to estimate a user-specified model or to automatically build one using these weights or pre-assigned "degrees of believability". For example consider, the actuarial economist who has chronological data where the reading at time period t is based upon "n" samples. Thus a reading with a large "n" is more credible than one with small "n". The user can enter the weights in a disk file Weights.In . These weights will then be used and will be potentially modified if this test is enabled. The resultant weights are stored in Weights.Out, if the I/O option is specified.

Line 33

```
CONFIDENCE LEVEL FOR VARIANCE TEST             |90.0
```

If you select the variance stability test, then you must specify the confidence limit to be used for detecting possible change points. For example, .95 indicates that the program should identify all time periods that are significant at the 95% level. The interval for comparing variances (Variance Stability) is based on the number of forecasts. If you specify a 3 period forecast, then testing for variances (if enabled) will be done until a group of residuals is less than 3. If the number of forecasts is 0, the program defaults to a minimum of 10 residuals in a group.

Line 34

```
MINIMUM NUMBER OF RESIDUALS TO POOL            |    5
```

This entry controls the length of the interval for comparing variances . For example if you specify a 10 , then a minimum of 10 residuals will be pooled and compared against the remaining residuals.

Line 35

```
CONSTANCY OF PARAMETERS:                        |    0
```

Choose "1" to enable. Chow suggested a test to assess the statistical significance between two sets of regression coefficients. We have extended this to ARIMA/TF models and furthermore SEARCH for the point of maximum contrast thus identifying the local cluster of homogenous data. This is equivalent to pooled cross-sectional time series where the number of consecutive values in each of the two groups is unknown and to be determined.

Line 36

```
CONFIDENCE LEVEL FOR CONSTANCY TEST              |
```

If you select the variance stability test, then you must specify the confidence limit to be used for detecting possible change points. For example, .95 indicates that the program should identify all time periods that are significant at the 95% level. The interval for comparing variances (Variance Stability) is based on the number of forecasts. If you specify a 3 period forecast, then testing for variances (if enabled) will be done until a group of residuals is less than 3. If the number of forecasts is 0, the program defaults to a minimum of 10 residuals in a group.

Line 37

```
STORE MODEL FORM (YES/NO)                         |1
```

Choose ‘1’ if you have a model saved and want to reuse it now.

Line 38

```
DISPLAY MANAGEMENT ANALYSIS                        |   0
```

```
If you want a report that tries to summarize "in english" information about
the time series from the model used to fit the data.
```

Line 39

```
ENABLE MODEL FORECASTING                           |   1
```

The forecasting program generates the forecast values for each time series. This option should be selected upon successful identification of the final Box-Jenkins model form.

Line 40

```
# OF FORECAST VALUES TO COMPUTE                   |  10
```

This entry indicates how many forecast values you want the program to compute. The accuracy of the forecasts can be assessed by the resulting errors. Since this entry measures the length of the interesting interval, it is also used to control the interval for comparing variances (Variance Stability). If you specify a 3 period forecast, then testing for variances (if enabled) will be done until a group of residuals is less than 3. If the entry here is 0, the program defaults to a minimum of 10 residuals in a group.

The maximum is the seasonality multiplied by 3. So if the maximum seasonality of 60 is used then the maximum number of forecasts to compute is 180.

Line 41

```
CONFIDENCE LIMIT FOR THE FORECASTS (%)            |80.0
```

The reliability of a forecast is measured in terms of its uncertainty. This program will compute the individual confidence limit of each forecast, given the information available at the forecast origin. You can specify whatever percent confidence limit you want the program to use.

Line 42

```
CONVERT THE FORECAST VALUES TO POSITIVE VALUES    |    0
```

If your data arises only in positive values then you might wish to constrain forecasts to the set of positive real numbers. A "YES" will convert the forecasts and confidence limits. All error reports are presented in terms of these rounded forecasts. Note that the aggregated sum is rounded after the aggregation thus the sum of the forecasts may not be equal to the aggregated sum.

Line 43

```
CONVERT THE FORECAST VALUES TO INTEGERS          |    0
```

If your data arises only in integer form, then it is known as DARMA or a discrete ARIMA problem. One can approximate a DARMA model by estimating as if the data were continuous and then integerizing the forecasts. This is an approximation and the user should be guided by the results. A 'yes' will convert the forecasts and confidence limits to the nearest integer. Thus the forecasts will be rounded off rather than truncated. All error reports are presented in terms of these rounded forecasts. Note that the aggregated sum is rounded after the aggregation thus the sum of the forecasts may not be equal to the aggregate of the forecasts. ARIMA models are also an approximation to a process that is continuous and is sampled at fixed intervals. Again the ARIMA model is an approximation and the user should be guided by the results.

Line 44

```
CONVERT PULSE AT LAST OBSERVATION TO STEP        |    0
```

Allows the user to apply his knowledge that the last observation is not a pulse but a permanent step that must be considered.

Line 45

```
CONVERT PULSE TO SEASONAL PULSE (SAVE LAST OBS   |    0
```

You may elect to convert an identified pulse at a particular time period to a SEASONAL PULSE. Consider where insufficient data exists to confirm a SEASONAL PATTERN. This feature allows the user to enforce the rule that all pulses, save a pulse at the last observation will be treated as the first point in a repetitive pattern.

Line 46

```
DISPLAY IDENTIFICATION INFORMATION               |    1
```

If no detail is required in the initial identification process send a "no". Autocorrelation is a measure of the unconditional dependence that exists between observations in a time series that are separated by a particular time interval, called lag. The value of the autocorrelation lies between +1 and -1. The closer the autocorrelation is to +1 and -1, the more highly correlated are the observations separated by the particular lag being considered. In summary, the autocorrelation measures the unconditional relationship between lags. Partial Autocorrelation is a measure of the conditional dependence that exists between observations in a time series that are separated by a particular time interval. The value of the partial autocorrelation lies between +1 and -1 and is evaluated just like the ACF. In summary, the partial autocorrelation measures the conditional correlation between lags. Cross Correlation is a measure of the dependence that exists between observations in two time series that are separated by a particular time interval, called lag. The value of the cross correlation lies between +1 and -1. The closer the cross correlation is to +1 and -1, the more highly correlated are the observations separated by the particular lag being considered. If the correlation is closer to +1, a positive correlation is indicated; if it is closer to -1, a negative correlation exists. In summary, the cross correlation measures the strength of the relationship between the lags of two time series.

Line 47

```
DISPLAY ACF TABLE DURING IDENTIFICATION         |    1
```

Choose "1" to see a table of the correlations at the initial identification stage. The table displays rows of correlations and their standard errors.

Line 48

```
DISPLAY ACF GRAPH DURING IDENTIFICATION         |    1
```

Choose "1" to see a plot of the correlations at the INITIAL identification stage.

Line 49

```
DISPLAY CCF TABLE DURING IDENTIFICATION         |    1
```

Choose "1" to see the cross-correlations between the prewhitened input and the prewhitened output series. This information is a statistical 'tool' used to identify the form of a transfer model. This option allows you to control whether or not they get reported in a table.

Line 50

```
DISPLAY CCF GRAPH DURING IDENTIFICATION         |    1
```

Choose "1" to see the cross-correlations between the prewhitened input and the prewhitened output series. This information is a statistical 'tool' used to identify the form of a transfer model. This option allows you to control whether or not they get reported in a plot.

Line 51

```
DISPLAY PREWHITENING MODEL(S)                   |    1
```

Choose "1" to see the model form as a table.

Line 52

```
DISPLAY PREWHITENING MODEL(S) EQUATION          |    1
```

Choose "1" if you want the program to display the model(s) in the form of an equation.

Line 53

```
DISPLAY IDENTIFIED MODEL                        |    1
```

Choose "1" if you wish to see the model form as a table.

Line 54

```
DISPLAY IDENTIFIED MODEL EQUATION              |    1
```

Choose "1" if you want the program to display the model(s) in the form of an equation.

Line 55

```
DISPLAY ESTIMATION INFORMATION                 |    1
```

If this is set to "0" then no estimation information will be reported. This means that lines 123-128 would be skipped.

Line 56

```
DISPLAY ESTIMATED MODEL PARAMETERS             |    1
```

Choose "1" if you wish to see the model form as a table.

Line 57

```
DISPLAY ESTIMATED MODEL EQUATION               |    1
```

Choose "1" if you want the program to display the model(s) in the form of an equation.

Line 58

```
DISPLAY UNTRANSFORMED FIT VS ACTUAL TABLE      |    1
```

Choose "1" so that the program displays a chart which shows the fit values, the residual values and the actual values from the estimated model.

Line 59

```
DISPLAY DIAGNOSTIC CHECKING INFORMATION         |    1
```

Choose "1" to show detail regarding the diagnostic checking process.

Line 60

```
DISPLAY RESIDUAL ACF TABLE                     |    1
```

Choose "1" to see a table of the residual correlations each time that they are computed, the entry here should be a 'yes'. The table displays rows of correlations and their standard errors.

Line 61

```
DISPLAY RESIDUAL ACF GRAPH                    |    1
```

Choose "1" to see a plot of the residual correlations each time that they are computed, the entry here should be a 'yes'.

Line 62

```
DISPLAY RESIDUAL CCF TABLE                    |    1
```

Choose "1" to have Autobox batch display the cross-correlations between the prewhitened input and the residuals from the current model are the statistical 'tool' used to identify the form of a fixup required to the transfer model. This option allows you to control whether or not they get reported in a table.

Line 63

```
DISPLAY RESIDUAL CCF GRAPH                    |    1
```

The cross-correlations between the prewhitened input and the residuals from the current model are the statistical 'tool' used to identify the form of a fixup required to the transfer model. This option allows you to control whether or not they get reported in a plot.

Line 64

```
DISPLAY NECESSITY TEST RESULTS                |    1
```

Choose "1" to see the necessity test results. ARIMA modeling may be deficient when the model has too many coefficients. It is important to discard or delete unnecessary structure as it inflates forecast variances, among other things.

Line 65

```
DISPLAY SUFFICIENCY TEST RESULTS              |    1
```

Choose "1" to see the sufficiency test results. ARIMA modeling may be deficient when the MODEL does not have enough structure. The omitted structure can be identified by studying the sample ACF AND PACF of the residuals. In this way we move structure from the residuals to the model.

Line 66

```
DISPLAY VARIANCE STABILITY TEST RESULTS       |    0
```

Choose "1" to see the variance stability results. ARIMA modeling may be deficient when the series has a non-constant variance. The program will test the residuals from the ARIMA model for possible change points. Essentially interventions are changes in the mean level of the errors while variance stability measures changes in the variance.

Line 67

```
DISPLAY OUTLIER TEST RESULTS                        |    1
```

Choose "1" to enable this function. ARIMA modeling may be deficient when the series has been intervened with. This program will test the residuals from the ARIMA model for possible outlier (intervention) variables. We suggest that you modify either your model or your time series for any outlier variables that may be found. If you have enabled automatic fixup FOR outliers, in the choose analysis options section then these modifications will be done for you automatically. A "yes" shows the details of this process.

Line 68

```
DISPLAY CONSTANCY TEST                              |    0
```

Choose "1" to see the constancy test results. You get a table showing the observations and those values with a significant change in the reliability of the model parameters.

Line 69

```
DISPLAY FORECASTING INFORMATION                      |    1
```

If this is set to "0" then no forecasting information will be reported. This means that lines 143-150 would be skipped.

Line 70

```
DISPLAY FORECAST MODEL PARAMETERS                   |    1
```

Choose "1" to see the model form as a table.

Line 71

```
DISPLAY FORECAST MODEL EQUATION                     |    1
```

Choose "1" if you want the program to display the model(s) in the form of an equation.

Line 72

```
DISPLAY TABLE OF TRANSFORMED FORECAST VALUES        |    1
```

Choose "1" to see the forecast in transformed units. Your forecasting model may contain a transformation parameter (lambda). If so, then the program generates forecast values for both the original data and the transformed data.

Line 73

```
DISPLAY TABLE OF FORECAST VALUES                    |    1
```

Choose "1" to see the forecasts with their confidence bounds. Your forecasting model may contain a transformation parameter (lambda). If so, then the program generates forecast values for both the original data and the transformed data.

DISCUSSION OF FFLITE

Assuming that the series 0438.asc had been analyzed**,** the compressed model saved in 0438.abl would look as follows**:**

```
     147/   2  745  13   7  40   7   2     1.000
HIGH                 0  13   0
MOVE_PRICE       0   17  4
0438                 0
  1
  0           0.0084611576
  1
  4          -0.4498213812
  1
  0           0.0000000000
*****BASELINE****
          3.1871258281          0.7146756892
          3.0262497639          0.7146756892
          3.4692965832          0.7146756892
          3.3413501075          0.7146756892
          3.4936723981          0.7146756892
          3.1557451419          0.7146756892
          3.1244239210          0.7146756892
          3.1871258281          0.7146756892
          4.0408596480          0.7146756892
          3.4692965832          0.7146756892
          3.3413501075          0.7146756892
          4.7164280232          0.7146756892
          3.1557451419          0.7146756892
```

What we know after reading line 1

```
147/   2  745  13   7  40   7   2     1.000
```

The date of the LAST OBSERVATION is 147/2 based upon 745 historical values .

A forecast of 13 periods was made with seasonality = 7. The first data point was at 40/7 and there are 2 DYNAMIC VARIABLES(DY) in the model and the transformation parameter = 1.000 meaning no transformation is required

PLEASE NOTE: YOU SHOULD ARCHIVE THE SUM OF THE ORIGINAL NUMBER OF HISTORICAL VALUES AND THE ORIGINAL NUMBER OF FORECAST PERIODS(758=745+13) AS A CONTROL NUMBER FOR UPDATING THE FORECASTS. IN UPDATING, THE SUM OF ANY NEW NUMBER OF HISTORICAL VALUES AND/OR NEW NUMBER OF FORECAST PERIODS CANNOT EXCEED THE CONTROL NUMBER. IF YOU DESIRED TO EXCEED THAT CONTROL, YOU WOULD HAVE TO PROCESS THE SERIES WITH THE UPDATED DATA BY GOING THROUGH THE ENTIRE MODELING PROCESS IN TOTAL.

The next number of DY lines indicate the series name(position 1-22), the number of historical values required(position 25-28), the number of future values required(position 29-32); and the number of required lead values(position 33-36).

The next line indicates the output series name(position 1-22) and the number of historical values required(position 25-28)

The next lines indicate, for each series, a) the number of coefficients, b) the backorder power and the related coefficient. **These are not to be changed**

```
1
0      0.0084611576
1
4     -0.4498213812
1
0      0.0000000000
```

The *****BASELINE**** data reflects the original baseline forecasts in the left column and their related degree of accuracy in the right column. Baseline forecasts were generated in AUTOBOX using all the factors of the full model except the Dynamic series. **These values cannot be changed.**

RULES TO DETERMINE # OF HISTORICAL VALUES NEEDED BY THE MODEL

We will read the next DY+1 records which in this case is 3.and examine the 25-28th position to determine the number of historical values required for each of the DY+1 series.

E.G. the next 3 records or lines are:

```
HIGH             0  13  0
MOVE_PRICE       0  17  4
0438             0
```

0 + 0 + 0 gives us 0 and we have no need for historical values.

If, for example, the 3 records were:

```
HIGH            5  13   0
MOVE_PRICE      3   17   4
0438            7
```

then we would need 5+3+7 or 15 values in the 0438.his file, the supporting file containing the required historical values.

If historical values were needed, then the file would be saved with the number of historical values required for each of the DYNAMIC SERIES in the order indicated, followed by the number of historical values required for the OUTPUT SERIES. The values for each series would be entered from oldest to latest (that is: most recent data).

RULES TO DETERMINE # OF FUTURE VALUES NEEDED BY THE MODEL

The number of future values for each dynamic series is the number of forecasts in line 1, position 21 – 24 plus the number of leads indicated in position 33-36 in the respective DY lines. For HIGH there are 0 leads, so the number of future values in the 29-32th position will be 13 (13 + 0). For MOVE_PRICE the number of future values in the 29-32th position will be 17 (13 + 4).

Thus 13+17 gives 30 and we need 30 values for the 0413.fut file, the companion file containing the required future values. The first 13 values reflect "guesstimates" or planned values for the first DYNAMIC SERIES named "HIGH".

The next 17 values reflect "guesstimates" or planned values for the second DYNAMIC SERIES named "MOVE_PRICE".

Note, there are never any future values or leads indicated for the output series

Procedures:

TO OBTAIN REVISED FORECASTS AS NEW SALES DATA BECOMES AVAILABLE.

Consider that 3 periods have gone by and we have just observed OBSERVATION 748 at time period 147/5; and we wish to forecast out the next 9 days.

The new historical values are as follows:

#	date	sales	HIGH	MOVE_PRICE
746	147/3	3	5	6.10
747	147/4	2	7	6.05
748	147/5	1	3	6.15

We have to modify the three pieces of information ..

Model

Historical values

Future values

Model

> We would change the first record to the following reflecting the NEW YEAR OF THE LAST OBSERVATION(position 7-10), NEW PERIOD OF THE LAST OBSERVATION(position 12-15) THE NEW TOTAL OF HISTORICAL VALUES(position 16-20) and the NEW NUMBER OF FORECASTS(position 21-24):

```
147/   5  748  9   7  40  7   2
```

Please note: as indicated in the Explanations above, the sum of any new number of historical values and/or new number of forecast periods cannot exceed the sum of the original number of historical values and the original number of forecast periods

The next set of records up to and including "BASELINE" would be unchanged except for the new future values required. Again as explained in the Rules above, the number of future values for each dynamic series is the number of forecasts in line 1, position 21 – 24 plus the number of leads indicated in position 33-36 in the respective DY lines. For HIGH there are 0 leads so the number of future values in the 29-32th position will be 9 (9 + 0). For MOVE_PRICE the number of future values in the 29-32th position will be 13 (9 + 4).

.

```
HIGH              0   9   0
MOVE_PRICE        0  13   4
0438              0
  1
  0          0.0084611576
  1
  4         -0.4498213812
  1
  0          0.0000000000
*****BASELINE****
```

The next 9 records or lines would be taken from the original baseline data starting with the 4th line, since we now have real data for the first three periods.

```
3.3413501075          0.7146756892
3.4936723981          0.7146756892
3.1557451419          0.7146756892
3.1244239210          0.7146756892
3.1871258281          0.7146756892
4.0408596480          0.7146756892
3.4692965832          0.7146756892
3.3413501075          0.7146756892
4.7164280232          0.7146756892
```

Historical Values

If historical values were needed, as indicated in position 33-36 in the respective DY lines and the OUTPUT SERIES line, then the array would be loaded using the number of historical values required for each of the DYNAMIC SERIES in the order indicated, followed by the number of historical values required for the OUTPUT SERIES; however they would now incorporate the latest historical values. The values for each series would be entered from oldest to latest ..most recent.

In our example case there are no required historical so nothing would be added to the 0413.his file.

Future Values

Examining the required number of future values (position 29-32) tells us we need provide 22 future values, 9 for the first DYNAMIC SERIES named HIGH and 13 for the second DYNAMIC SERIES.