

Using JMP® Version 4 for Time Series Analysis

Bill Gjertsen, SAS Institute Inc., Cary, NC

Abstract

Three examples of time series will be illustrated. One is the classical airline passenger demand data with definite seasonal variation. The other two will be daily closing price data from a steady growth series (CSCO) and an accelerated growth series (CIEN).

Introduction

Among the many new enhancements to JMP in the version 4 release is the ability to do time series forecasting. The procedures include smoothing models, a seasonal smoothing method (Winter's method), and ARIMA (Autoregressive Integrated Moving Average) modeling. In this paper we will arrive at final forecast from:

- highly seasonal airline passenger data
- a steady growth stock (CSCO)
- an accelerated growth stock (CIEN)

Data Sources and Data Cleansing

A good free source of daily close stock data for CSCO and CIEN is provided via the Internet at:

<http://moneycentral.msn.com/investor/charts/charting.asp?Symbol=cscoc> and
<http://moneycentral.msn.com/investor/charts/charting.asp?Symbol=cienc>.

You can export a year's worth of daily data directly into Microsoft® Excel by choosing "Export Data". Once in Excel, strip off the first 4 rows and Save as with the default .csv file extension. Then within JMP choose Import from the File platform and the text file can be directly into JMP. Once within JMP we resort with the earliest dates appearing first and eliminate the holidays that have missing stock data. We will use only the daily close value, CLOSE, and not use HIGH, LOW or VOLUME. For the airline passenger data we use the Seriesg.jmp dataset supplied in the Time Series folder found in the Sample Data that comes with JMP4.

Smoothing and ARIMA Methods

Simple forecasting using smoothing models is based on the premise that reliable forecasts can be realized by modeling patterns in the data that are visible in a time series and/or spectral density plot. We in fact will use this approach to try our first forecasts of the Seriesg data. ARIMA modeling uses additional pattern information including differencing, autocorrelation and partial autocorrelation, and plots of AR coefficients to help identify the model. The advantage of ARIMA over smoothing methods is that

they are more flexible in fitting the data. However the iterative approach of fitting several ARIMA models, checking adequacy, and comparing results is more demanding. This approach is alleviated somewhat by using JMP's Model Comparison Table. This table summarizes the fit statistics for each model under consideration. Included in the table output are three criteria useful for model selection. They are:

- AIC (Akaike's Information Criterion)
- SBC (Schwarz's Bayesian Criterion)
- -2LogLikelihood

Smaller values of each of these criteria indicate better fits.

Example 1: Forecasting Airline Passenger Volume

An analyst at a major international airline wants to forecast passenger demand for the next 15 months. Excessive demand in any of these months could impact the flight schedule and/or the number of flights offered. Seriesg has 12 years of data on a monthly basis and wants to forecast monthly demand as input to projecting seating capacity and number and type of flights. The last 15 months will be contrasted with the forecasted values to understand incremental changes forecasted by month. (This actually is a classic seasonal time series from Box and Jenkins (1) page 531). The time series variable is international airline passengers in thousands, Passengers, and the time ID is Time in months. The last two years of the Seriesg passenger time series data is depicted in Figure 1a, and default graphs of the complete time series together with acf (autocorrelation function) and pacf (partial autocorrelation function) charts are depicted in Figure 1b. To assess that periodicity=12 we produce a Spectral Density plot in Figure 1c and as a "screening" forecast we attempt a seasonal Winter's Method forecast in Figure 1d. After realizing that the residuals are increasing in time from the residual plot we continue our iterative forecasting with Log Passengers as the dependent time series variable. We note also that Winter's method produces a non invertible forecast.

Results:

On $\text{Log}_{10}(\text{Passengers})$ we iteratively fit a seasonal Winters, a seasonal ARIMA(1,1,1)(1,1,1)12, and two ARIMA(0,1,1)(0,1,1)12 models. Finally a reduced model with the lowest AIC and SBC emerged as best: Seasonal ARIMA(0,1,1)(0,1,1)12 No Intercept. Look at the Model Comparison table in Figure 1e. By saving results and creating a subset comparing the next 15 months predicted with the last 15 months of actual passengers we find there are two months, March and August of the first year, where we expect passenger demand in excess of 60,000 in Figure 1f.

Example 2: Forecasting Stock Prices for a steady grower (CSCO)

An investor buys 100 shares of CSCO stock on 5/25/00 and decides to sell a short term call option that expires 15 trading days in the future on 6/16/00. What call option should he write to be reasonably certain that the stock will not be called away? For this problem, which is characteristic to short term options traders, what is needed is a short term forecast with a confidence interval. We could either do this with regression or time series methods. For purposes of this paper we will only look at time series estimates. Using the same techniques as in Example 2 we find from the residuals and using the acf plot and the AR coefficients plot that Log(Log(CSCO) (Log = Log₁₀) modeled with a simple AR1 process gives us a stable and invertible series with excellent fit RSquare= .99 and smallest values for AIC and SBC with several model contenders (Figure 2a). Figure 2b shows the short term forecast, 95% confidence limits and a good looking residual pattern. Figure 2c indicates what call option we would have sold (the June 70) if we wanted to 95% confident of not being exercised. By saving the results and generating predicted values for the upper 95% confidence limit and comparing it with actual close values we see the stock ended up at 67.813 and we were close but not in jeopardy of having the call exercised.

Example 3: Forecasting Stock Prices for a volatile stock (CIEN)

We repeat the same scenario as Example 2 but with an even more volatile response, Log(Log(CIEN). Again an AR1 model fits best and even though the stock rises 40% within the 15 day period we do not come close to exercising either a June 165 or June 170 option!

Summary

We have shown several instances of forecasting using time series techniques. In the passenger demand example using various tools and graphs led us to a seasonal ARIMA model. Here we introduced the useful Model Comparison report. And iteratively

used reports about stability and invertibility and residuals to converge on a "best" candidate.

Whereas the Spectral Density plot confirmed the annual cyclical nature of passenger demand the acf plot confirmed the autoregressive nature of the response for the Log(Log) of the two stocks

Notation

(1) and (2) are good references for a discussion of ARIMA modeling

B is the lag operator $By_t = y_t - y_{t-1}$

p is the order of the Autoregressive polynomial $\Phi(B)$

d is the differencing order

q is the order of the Moving Ave. polynomial $\Theta(B)$

where: $\Phi(B)(w_t - \mu) = \Theta(B)a_t$

and where $w_t = (1 - B)^d y_t$ is the response series after differencing and a_t are a sequence of random shocks.

The ARIMA model is described as ARIMA(p,d,q) and the seasonal as ARIMA(p,d,q) (P,D,Q) s

where: s is the number of periods per season.

References

1. G. E. P Box and G. M, Jenkins (1976), *Time Series Forecasting and Control*, Holden-Day Inc., San Francisco
2. W. Vandaele (1983) *Applied Time Series and Box-Jenkins Models*, Academic Press Inc, New York

Author

William Gjertsen, SAS Institute Inc., SAS Campus Drive, Cary, NC, 27513. Phone: (919) 677-8000 ext 7236
Fax: (919) 677-4444. bill.gjertsen@sas.com

JMP is a registered trademark of SAS Institute Inc. in the USA and in other countries. ® indicates USA registration.

Other brand and product names are registered trademarks or trademarks of their respective companies.

Seriesg						
Seriesg						
Notes from Box and Jenkins						
Time Series						
TS 1 Exploratory						
TS 2 Exploratory						
Columns (40)						
Passengers						
Time						
Season						
Log Passengers						
Rows						
All Rows	144					
Selected	0					
Excluded	0					
Hidden	0					
Labelled	0					
Passengers	Time	Season	Log Passengers			
121	360	121	1	2.5563025		
122	342	122	2	2.53402611		
123	406	123	3	2.60852603		
124	396	124	4	2.59769519		
125	420	125	5	2.62324929		
126	472	126	6	2.673942		
127	548	127	7	2.73878056		
128	559	128	8	2.74741181		
129	463	129	9	2.66558099		
130	407	130	10	2.60959441		
131	362	131	11	2.55870857		
132	405	132	12	2.60745302		
133	417	133	1	2.62013605		
134	391	134	2	2.59217676		
135	419	135	3	2.62221402		
136	461	136	4	2.66370093		
137	472	137	5	2.673942		
138	535	138	6	2.72835378		
139	622	139	7	2.79379038		
140	606	140	8	2.78247262		
141	508	141	9	2.70586371		
142	461	142	10	2.66370093		
143	390	143	11	2.59106461		
144	432	144	12	2.63548375		

Figure 1a: In this table Passengers is the number of international airline passengers (in thousands) who flew each month. Time is the cumulative month number (only the last 24 months depicted), Season is month if the year and Log Passengers is $\log_{10}(\text{Passengers})$. For many time series a log transformation together with differencing will produce stationary series with sensible forecasts.

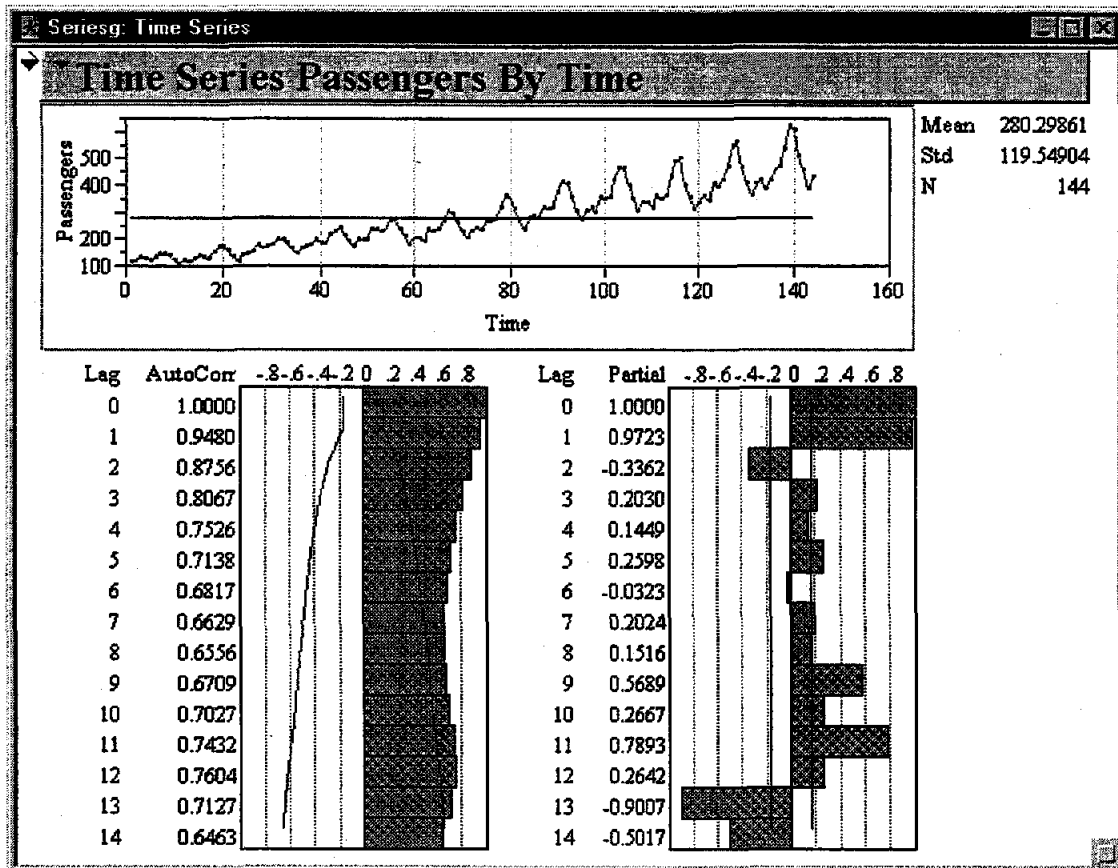


Figure 1b: Default plot of the time series and the acf and pacf functions. Even without a model and residual plot we see the series in increasing in variation as we increase in time.

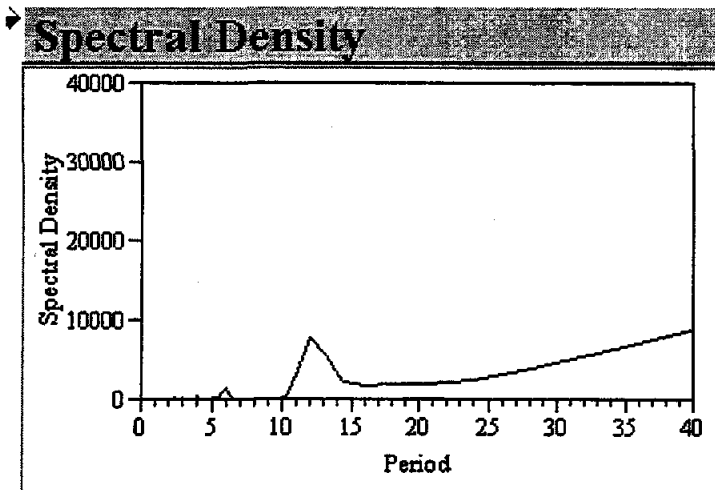


Figure 1c: There is a definite periodicity of 12.

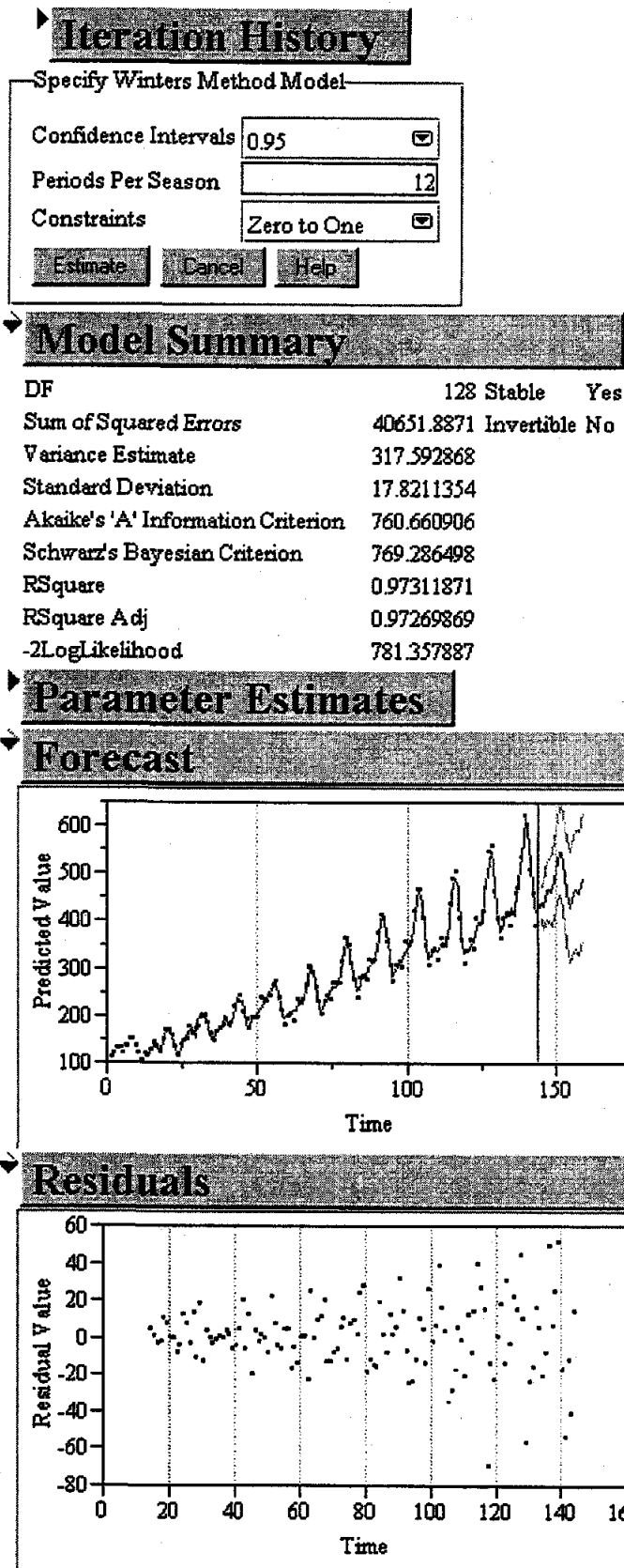


Figure 1d: Generating a Winter's method seasonal forecast produces a non-invertible result (unreliable forecast) and a residual pattern calling for a log transformation for the response

Model Comparison

Model	DF	Variance	AIC	SBC	RSquare	-2LogLH
Winters Method (Additive)	128	0.0002843	-1063.69	-1055.065	0.9900	-1066.414
Seasonal ARIMA(1, 1, 1)(1, 1, 1)12	126	0.0002619	-1070.446	-1056.07	0.9911	-1080.623
Seasonal ARIMA(0, 1, 1)(0, 1, 1)12	128	0.0002601	-1075.312	-1066.686	0.9910	-1079.699
Seasonal ARIMA(0, 1, 1)(0, 1, 1)12 No Intercept	129	0.0002582	-1078.288	-1072.538	0.9910	-1079.671

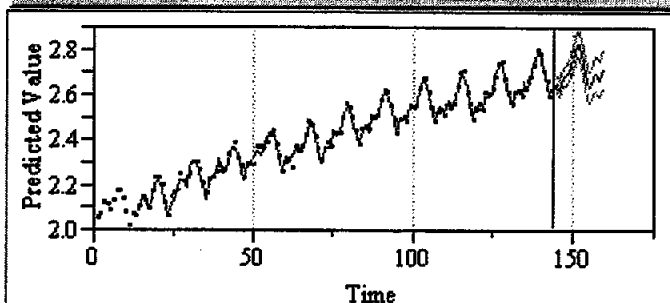
Model Summary

DF	129	Stable	Yes
Sum of Squared Errors	0.033309	Invertible	Yes
Variance Estimate	0.00025821		
Standard Deviation	0.01606889		
Akaike's 'A' Information Criterion	-1078.288		
Schwarz's Bayesian Criterion	-1072.5376		
RSquare	0.99097293		
RSquare Adj	0.99090295		
-2LogLikelihood	-1079.6714		

Parameter Estimates

Term	Factor	Lag	Estimate	Std Error	t	Prob> t
MA1,1	1	1	0.40182266	0.0896223	4.48	<.0001
MA2,12	2	12	0.55694078	0.0731031	7.62	<.0001

Forecast



Residuals

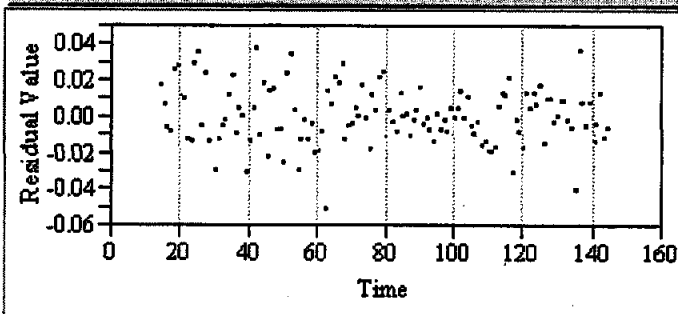


Figure 1e: Final model of Seasonal ARIMA(0, 1, 1)(0, 1, 1)12 No Intercept . It has the lowest AIC and SBC, is stable and invertible, and has good looking forecasts and residuals

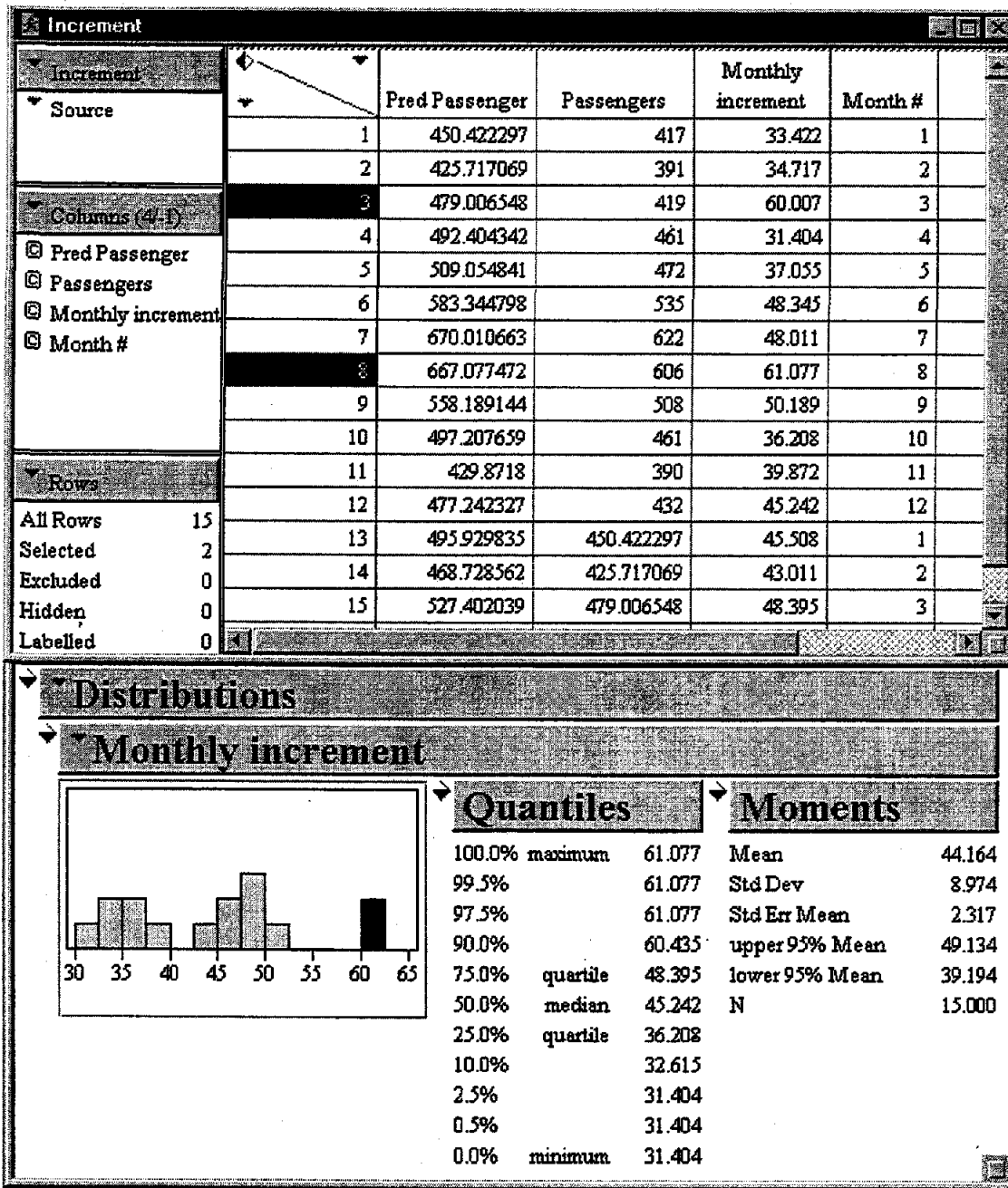


Figure 1f: Our airline analyst concludes that there are two months, March and August of the first year where we expect passenger demand in excess of 60,000

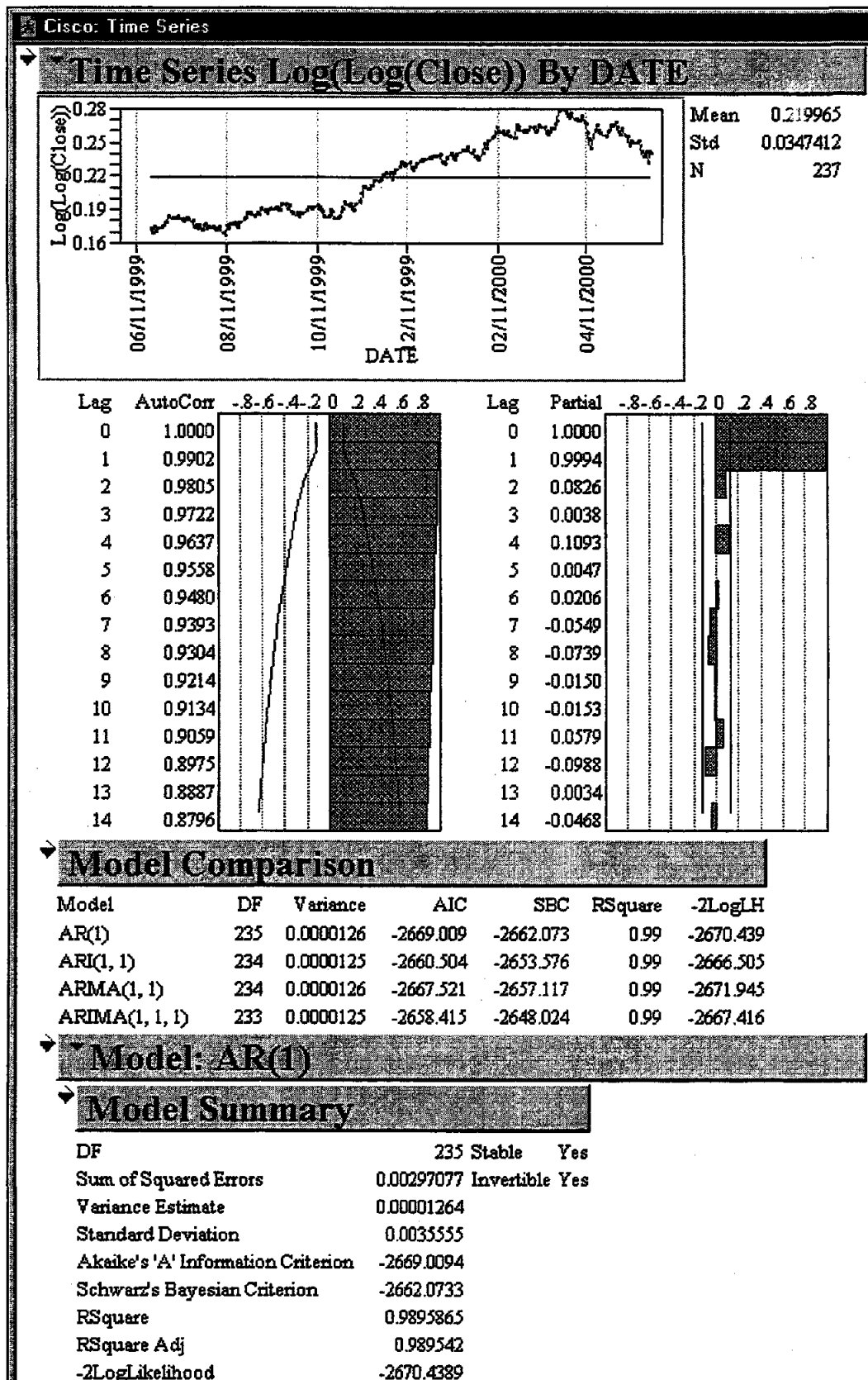


Figure2a: Showing a simple AR1 (ARIMA (1,0,0) process as being best from the Model Comparison Table and invertible and stable from the Model Summary.

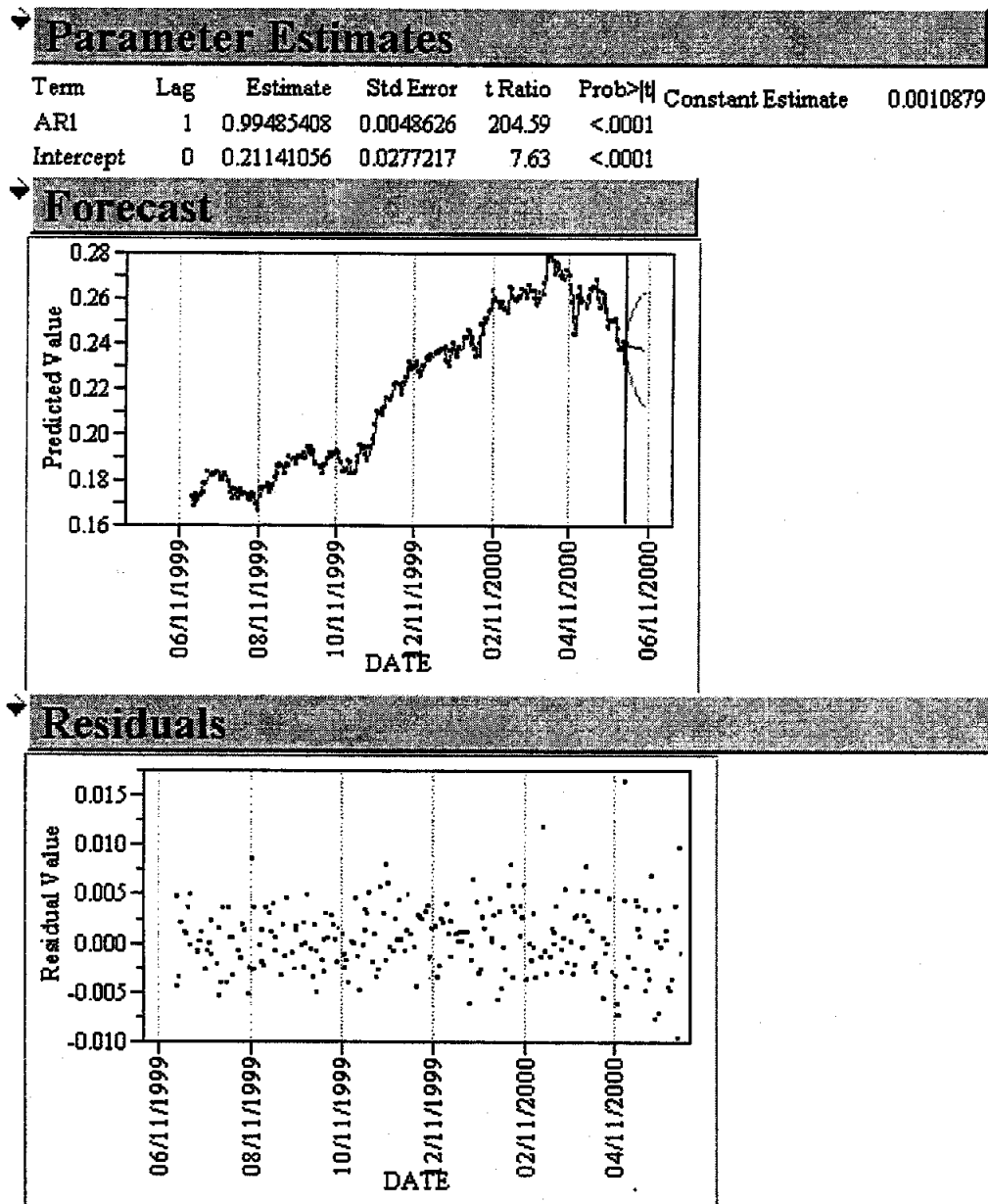


Figure 2b: The final AR1 forecast is shown together with its residuals.

Forecast for Option price of Cisco				
Forecast for Option price		DATE	CLOSE	Pred UpperCL(Close)
Source		1	05/25/2000	54.5
		2	05/26/2000	54.938
		3	05/30/2000	59.875
Columns (3/3)		4	05/31/2000	56.938
		5	06/01/2000	60.938
		6	06/02/2000	64.375
☒ DATE ☒ CLOSE ☒ Pred UpperCL(Close)		7	06/05/2000	63.25
		8	06/06/2000	61.313
		9	06/07/2000	62.875
Rows		10	06/08/2000	63.688
		11	06/09/2000	64.375
		12	06/12/2000	62.125
All Rows	16	13	06/13/2000	65
Selected	0	14	06/14/2000	65.188
Excluded	0	15	06/15/2000	66.5
Hidden	0	16	06/16/2000	67.813
Labelled	0			68.3612371

Figure 2c: The choice of selling the June 70 call for CSCO based on the predicted confidence limit ending at 68.36 resulted in profit since the call would not have been exercised.

Forecast for Option Price of Ciena				
Forecast for Option		DATE	CLOSE	pred UCL
Source		1	05/25/2000	104.5
		2	05/26/2000	99.688
		3	05/30/2000	121.5
Columns (3/3)		4	05/31/2000	119.688
		5	06/01/2000	130.188
		6	06/02/2000	138.313
☒ DATE ☒ CLOSE ☒ pred UCL		7	06/05/2000	133.625
		8	06/06/2000	137.875
		9	06/07/2000	131.438
Rows		10	06/08/2000	135.125
		11	06/09/2000	139.875
		12	06/12/2000	139.375
All Rows	16	13	06/13/2000	144.938
Selected	0	14	06/14/2000	144.438
Excluded	0	15	06/15/2000	145.5
Hidden	0	16	06/16/2000	145.25
Labelled	0			170.18386

Figure 3: Similarly, the June 170 call for CIEN was not exercised.