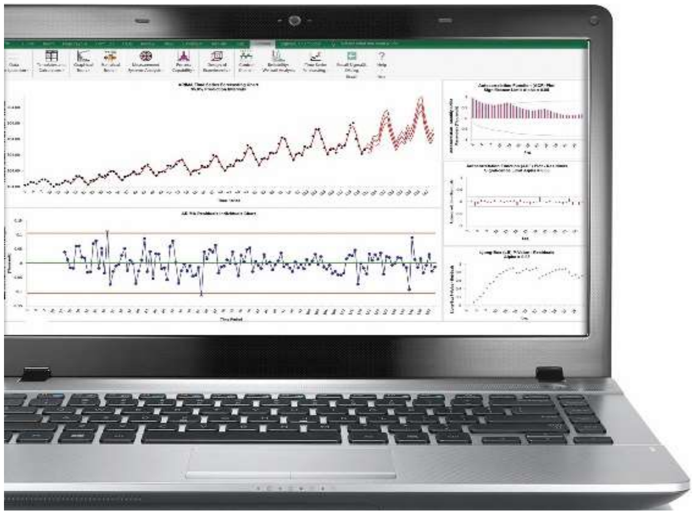
Lean Six Sigma Statistical Tools, Templates & Monte Carlo Simulation in Excel

What's New in SigmaXL® Version 9

Part 2 of 3: Time Series Forecasting



John Noguera
CTO & Co-founder
SigmaXL, Inc.
www.SigmaXL.com

Webinar November 12, 2020

SigmaXL V9: Time Series Forecasting

- Introduction
- Autocorrelation
- Example 1: Chemical Process Concentration
- Simple Exponential Smoothing
- Information Criteria
- Forecast Accuracy

SigmaXL V9: Time Series Forecasting

- Example 2: Monthly Airline Passengers
- Seasonal Trend Decomposition Plots
- Spectral Density Plots
- Error, Trend, Seasonal (ETS) Exponential Smoothing models

SigmaXL V9: Time Series Forecasting

- Autoregressive Integrated Moving Average (ARIMA) models
- Partial Autocorrelation
- ARIMA with Predictors
- Example 3: Electricity Demand with Temperature and Work Day Predictors
- References/Questions/Appendix

Introduction

- A time series is a series of data points indexed (or listed or graphed) in time order. Most commonly, a time series is a sequence taken at successive equally spaced points in time.
- Time series analysis comprises methods for analyzing time series data in order to extract meaningful statistics and other characteristics of the data. Time series forecasting is the use of a model to predict future values based on previously observed values.

From https://en.wikipedia.org/wiki/Time_series

Introduction

SigmaXL provides the following tools for exploratory data analysis of time series data:

- Run Chart
- Autocorrelation Function (ACF)/Partial Autocorrelation (PACF) Plots
- Cross Correlation (CCF) Plots with Pre-Whiten Data option
- Seasonal Trend Decomposition Plots
- Spectral Density Plot with Detection of Seasonal Frequency

Introduction

SigmaXL provides the following methods for time series analysis and forecasting:

- Exponential Smoothing
- Exponential Smoothing – Multiple Seasonal Decomposition (MSD)
- ARIMA – Box-Jenkins Autoregressive Integrated Moving Average
- ARIMA with Predictors
- ARIMA – MSD

Introduction

- Typically, either Exponential Smoothing or ARIMA may be used. It may be useful to try both to see which one gives a better model or use the average of the forecast from both methods.
- If the data has negative autocorrelation, ARIMA is recommended.
- If the data includes continuous or categorical predictors, use ARIMA with Predictors.

Introduction

- If the data are seasonal (i.e., influenced by seasonal factors), SigmaXL requires that the seasonal frequency be specified.
- Frequency is the number of observations per “cycle” unit of time, so monthly sales would be specified as seasonal frequency = 12 (observations per year). Quarterly revenue would be specified as seasonal frequency = 4. Hourly data would be 24 (observations per day).

Introduction

- Exponential Smoothing is limited to a maximum seasonal frequency of 24. For higher frequencies use Exponential Smoothing – Multiple Seasonal Decomposition (MSD).
- In MSD the seasonal component is first removed through decomposition, a nonseasonal exponential smooth model fitted to the remainder (+trend), and then the seasonal component is added back in.
- As the name implies, Multiple Seasonal Decomposition (MSD) also accommodates multiple seasonality.

Introduction

- ARIMA does not have a theoretical frequency limit, but for computational efficiency and to minimize the potential loss of observations through differencing, we recommend using ARIMA – MSD for seasonal frequency greater than 52 (or with multiple frequencies).
- Note, ARIMA with Predictors – MSD is not available.

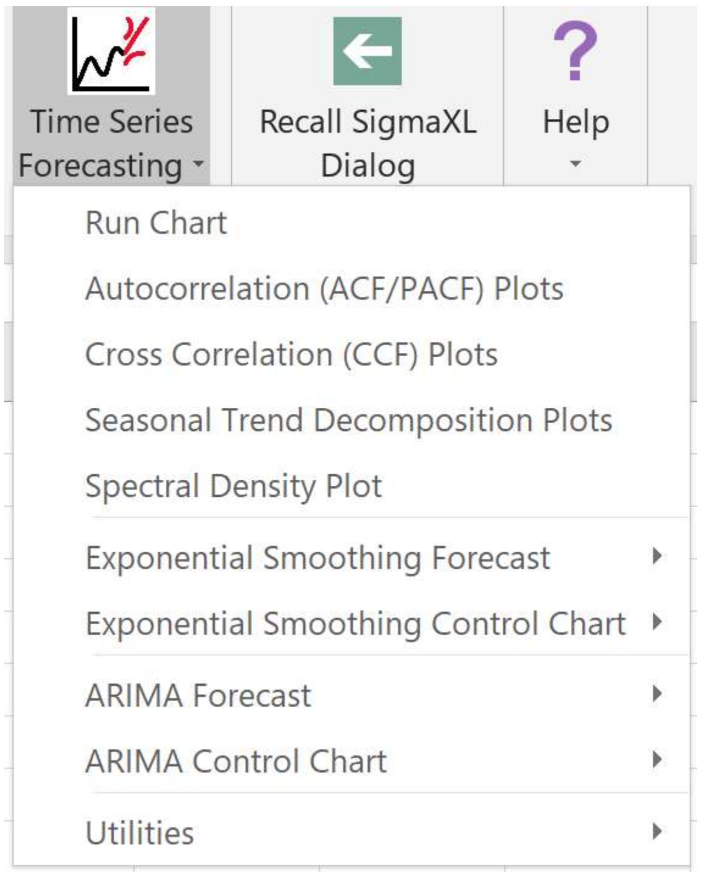
Introduction

- ARIMA assumes that the time series is stationary, i.e., it has the property that the mean, variance and autocorrelation structure do not change over time.
- If a time series mean is not stationary (e.g. trending), this can be corrected by differencing, computing the differences between consecutive observations for nonseasonal and between consecutive periods for seasonal data (e.g., Jan 2019 – Jan 2018, etc.).

Introduction

- If the variance changes over time, a Box-Cox transformation may be applied to achieve constant variance.
- Exponential Smoothing does not require stationarity.

SigmaXL Version 9
Time Series Forecasting Menu



Autocorrelation

- Just as correlation measures the extent of a linear relationship between two variables, autocorrelation (AC) measures the linear relationship between lagged values of data.
- A plot of the data vs. the same data at lag k will show a positive or negative trend. If the slope is positive, the AC is positive; if there is a negative slope, the AC is negative.
- The Autocorrelation Function (ACF) formula is:

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

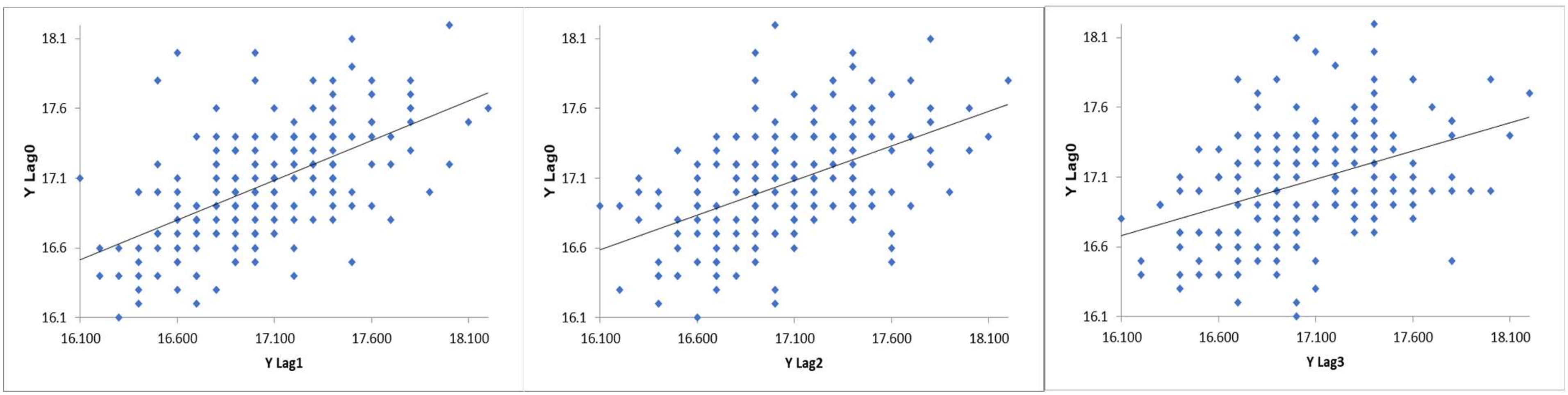
where T is length of the time series [4].

Autocorrelation

Y Lag0	Y Lag1	Y Lag2	Y Lag3
17			
16.6	17		
16.3	16.6	17	
16.1	16.3	16.6	17
17.1	16.1	16.3	16.6
16.9	17.1	16.1	16.3
16.8	16.9	17.1	16.1
17.4	16.8	16.9	17.1
17.1	17.4	16.8	16.9

Pearson Correlations	Y Lag1	Y Lag2	Y Lag3
Y Lag0	0.571	0.498	0.407

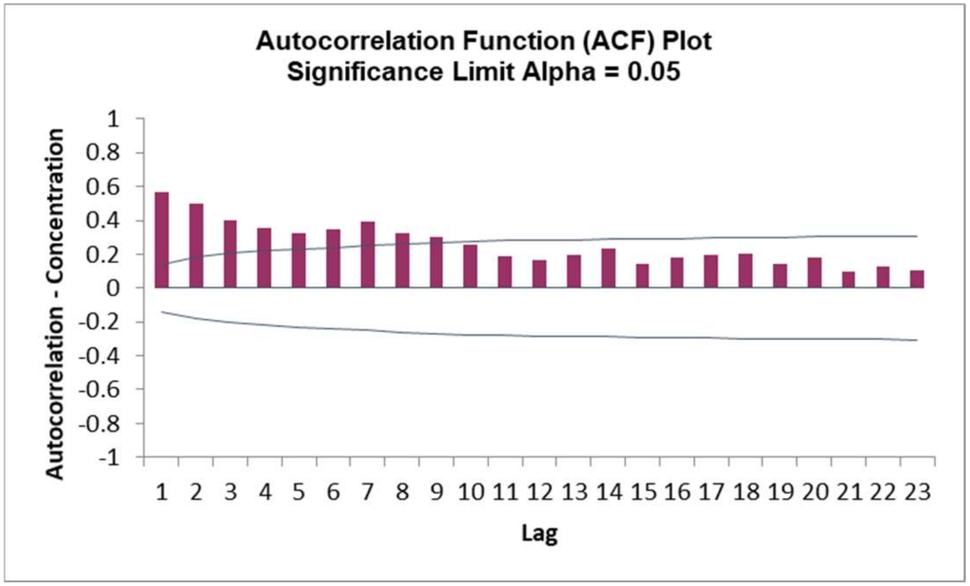
Pearson correlations are used here for demonstration purposes. They are approximately equal to the ACF correlation values.



Any statistically significant correlation ($r_k > 2/\sqrt{N}$) will adversely affect the performance of a Shewhart control chart.

The Ljung-Box test is used to determine if a group of autocorrelations are significant (see formula in Appendix).

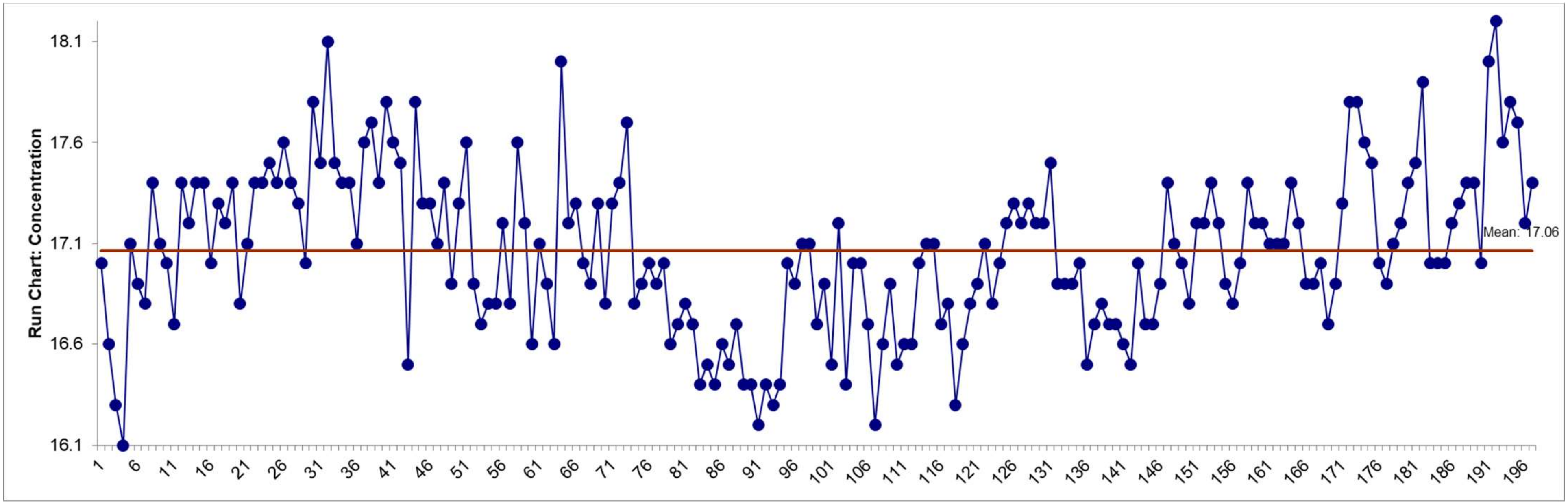
Example 1a: Box-Jenkins Series A - Chemical Process Concentration - Autocorrelation Function (ACF) Plot



SigmaXL > Time Series Forecasting > Autocorrelation (ACF/PACF) Plots

Example 1: Chemical Process Concentration - Series A.xlsx - Concentration

Example 1a: Box-Jenkins Series A - Chemical Process Concentration - Run Chart



SigmaXL > Time Series Forecasting > Run Chart

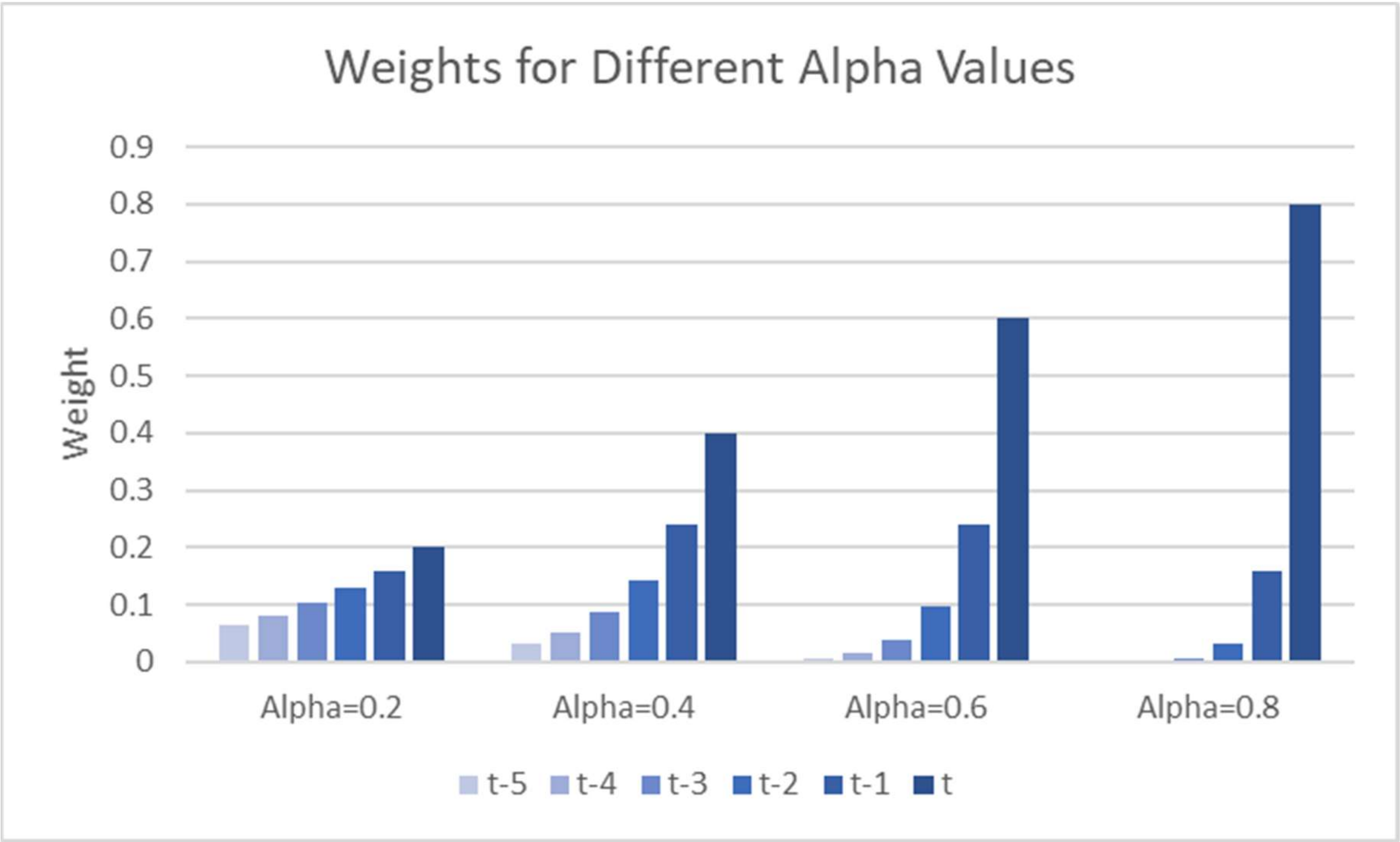
Simple (Single) Exponential Smoothing
Exponentially Weighted Moving Average (EWMA)

Forecasts are calculated using weighted averages, where the weights decrease exponentially as observations come from further in the past with the smallest weights associated with the oldest observations:

$$\hat{y}_{t+1} = \alpha\ y_t + \alpha(1-\alpha)\ y_{t-1} + \alpha(1-\alpha)^2\ y_{t-2} + \cdots$$

where $0 \leq \alpha \leq 1$ is the level smoothing parameter [4].

Simple (Single) Exponential Smoothing
Exponentially Weighted Moving Average (EWMA)



Simple (Single) Exponential Smoothing
Exponentially Weighted Moving Average (EWMA)

- An equivalent formulation for simple exponential smoothing is:

$$\hat{y}_{t+1} = \alpha y_t + (1 - \alpha)\hat{y}_t$$

with the starting forecast value (initial level) $\hat{y}_1$ typically estimated as y_1.

Simple (Single) Exponential Smoothing
Exponentially Weighted Moving Average (EWMA)

- The smoothing parameter and initial level are determined by minimizing the sum-of-square forecast errors (residuals):

$$\text{SSE} = \sum_{t=1}^{T} (y_t - \hat{y}_t)^2 = \sum_{t=1}^{T} e_t^2 .$$

- As usual for any statistical model, the residuals should be normal, independent and identically distributed.
- In SigmaXL, parameters are estimated by maximizing the Log-Likelihood function (which is similar to minimizing the residual sum-of-squares).

Model Selection and Information Criterion

– Akaike's Information Criterion

$$\text{AIC} = -2\log(L) + 2k,$$

where L is the likelihood of the model and k is the total number of parameters and initial states that have been estimated.

– The AIC corrected for small sample bias (AICc) is defined as:

$$\text{AIC}_\text{C} = \text{AIC} + \frac{k(k+1)}{T-k-1},$$

– The Bayesian Information Criterion (BIC) is:

$$\text{BIC} = \text{AIC} + k[\log(T) - 2]$$

Model Selection and Information Criterion

- Given a set of candidate models for the data, the preferred model is the one with the minimum Information Criteria value:
  - The Information Criteria rewards goodness of fit (as assessed by the likelihood function), but it also includes a penalty that is an increasing function of the number of estimated parameters.
  - The penalty discourages overfitting, because increasing the number of parameters in the model almost always improves the goodness of the fit.

Reference:
https://en.wikipedia.org/wiki/Akaike_information_criterion

Assess Forecast Accuracy

- Common forecast accuracy measures include:

  Root mean squared error: RMSE $= \sqrt{\text{mean}(e_t^2)}$

  Mean absolute error: MAE $= \text{mean}(|e_t|)$

  Mean absolute percentage error: MAPE

  $= \text{mean}\left(\left|\frac{100e_t}{y_t}\right|\right)$

  Mean absolute scaled error: MASE $= \text{mean}(|e_t|)/scale$

  - Scale is the MAE of the in-sample naïve or seasonal naïve forecast (set all forecasts to be the value of the last observation/period)
  - A scaled error is less than one if it arises from a better forecast than the average naïve/seasonal naïve forecast. Conversely, it is greater than one if the forecast is worse than the average naïve forecast [4].

Assess Forecast Accuracy

- Types of forecast error:
  - In-Sample One-Step-Ahead Forecast. This is less useful because the model may be over-fitted.
  - Out-of-Sample (Withhold) One-Step-Ahead. Model parameter estimates do not use any withhold data, but the forecast updates with every new withhold observation.
  - Out-of-Sample (Withhold) Full Period Forecast. This is important if one is assessing forecast accuracy over a horizon. This is used in forecast competitions.

Example 1b Demo of Simple Exponential Smoothing

Obs. No.	Concentration (y)	One-step-ahead forecast (y-hat)	Residuals (e)
1	17	17	0
2	16.6	17	-0.4
3	16.3	16.96	-0.66
4	16.1	16.894	-0.794
5	17.1	16.8146	0.2854
6	16.9	16.84314	0.05686
7	16.8	16.848826	-0.048826
8	17.4	16.8439434	0.5560566
9	17.1	16.89954906	0.20045094
10	17	16.91959415	0.080405846
11	16.7	16.92763474	-0.227634739
12	17.4	16.90487126	0.495128735
13	17.2	16.95438414	0.245615862
14	17.4	16.97894572	0.421054276
15	17.4	17.02105115	0.378948848
16	17	17.05894604	-0.058946037
17	17.3	17.05305143	0.246948567
18	17.2	17.07774629	0.12225371
19	17.4	17.08997166	0.310028339
20	16.8	17.12097449	-0.320974495
21	17.1	17.08887705	0.011122955
22	17.4	17.08998934	0.310010659
23	17.4	17.12099041	0.279009593
24	17.5	17.14889137	0.351108634
25	17.4	17.18400223	0.215997771
26	17.6	17.20560201	0.394397994
27	17.4	17.24504181	0.154958194

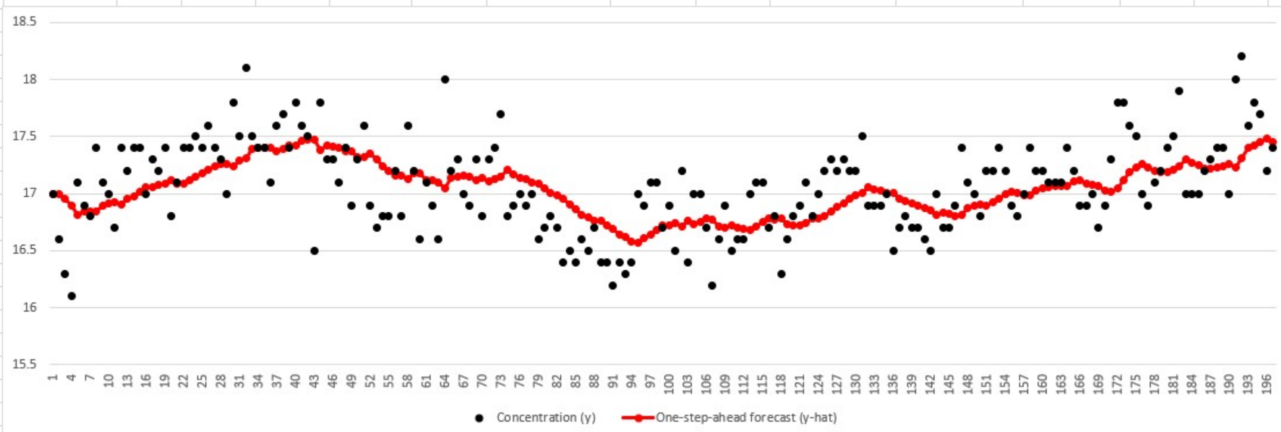
Exponential Smoothing Model Parameters	
alpha	0.1
initial level	17

Use Solver to obtain parameter values that minimize SSE:

SSE	21.8341

Exponential Smoothing Model Statistics	
Count	197
k terms	2
Log-Likelihood	-303.72228
AICc	613.56892
AIC	613.44457
BIC	623.29418

In-Sample Forecast Accuracy Metrics	
RMSE	0.332916231
MAE	0.271930445
MAPE	1.592107249
Scale (MR-bar)	0.275510204
MASE	0.9870068



Demo of Simple Exponential Smoothing - Concentration.xlsx

Example 1b Demo of Simple Exponential Smoothing

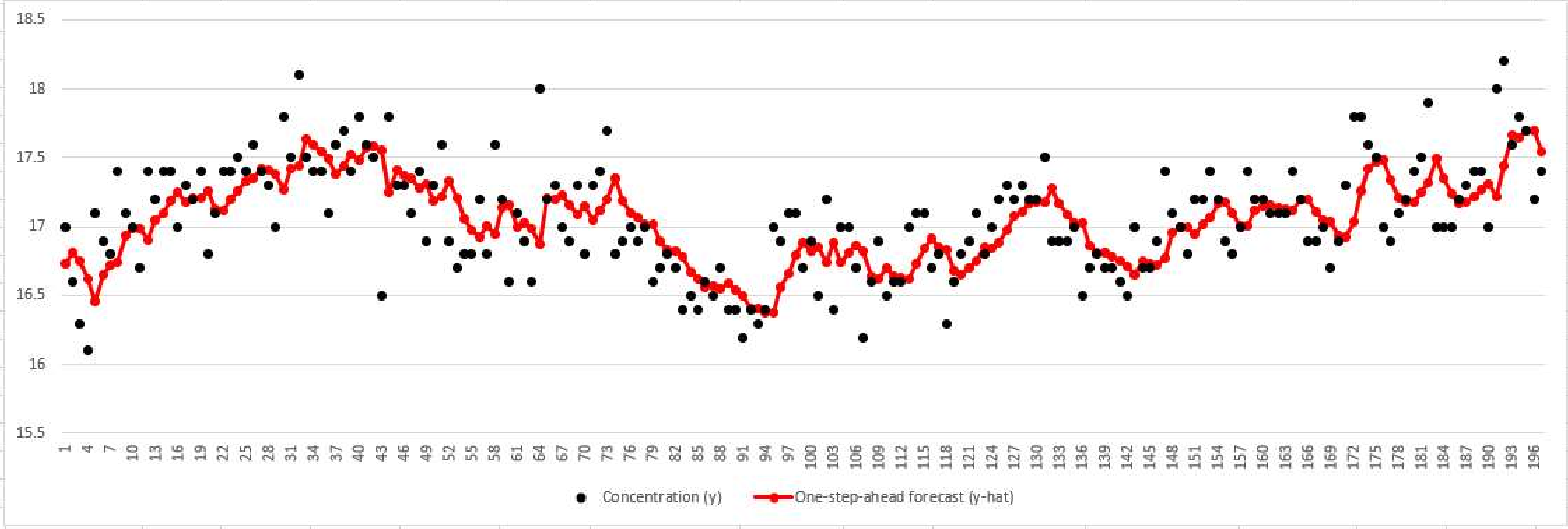
Exponential Smoothing Model Parameters	
alpha	0.29491
initial level	16.7319

Use Solver to obtain parameter values that minimize SSE:

SSE	19.7426

Exponential Smoothing Model Statistics	
Count	197
k terms	2
Log-Likelihood	-293.8036
AICc	593.73155
AIC	593.6072
BIC	603.45681

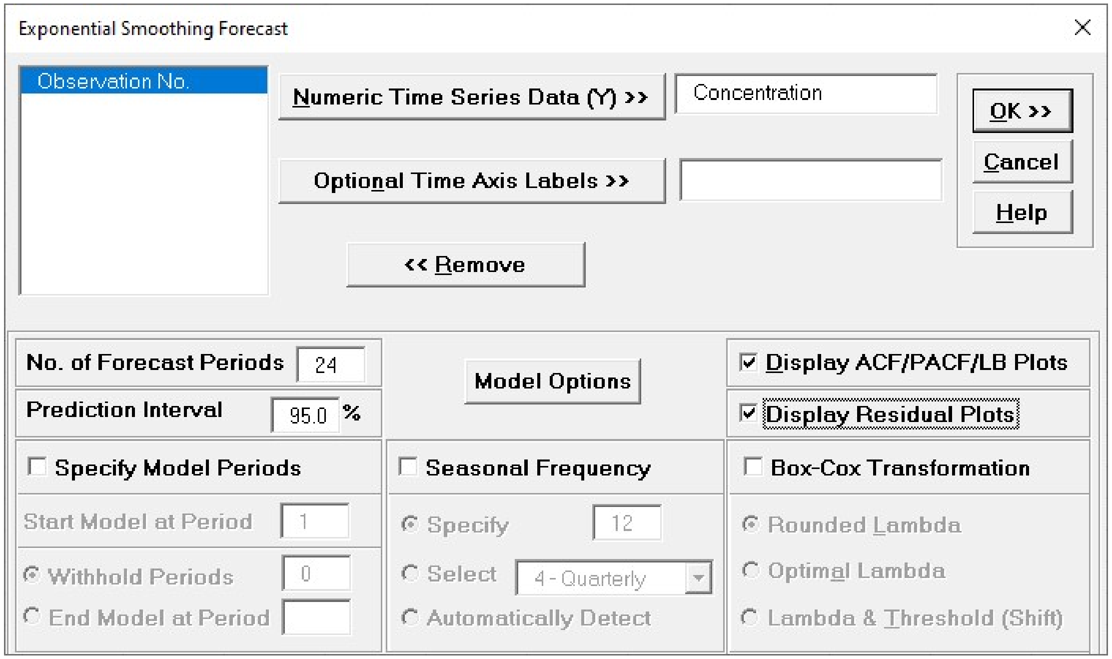
In-Sample Forecast Accuracy Metrics	
RMSE	0.316569323
MAE	0.247320195
MAPE	1.446469239
Scale (MR-bar)	0.275510204
MASE	0.897680708

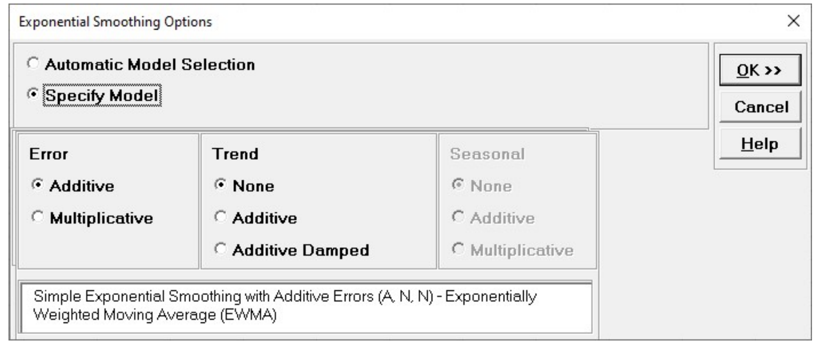


Solver used to optimize alpha and initial level parameters.

Demo of Simple Exponential Smoothing - Concentration.xlsx

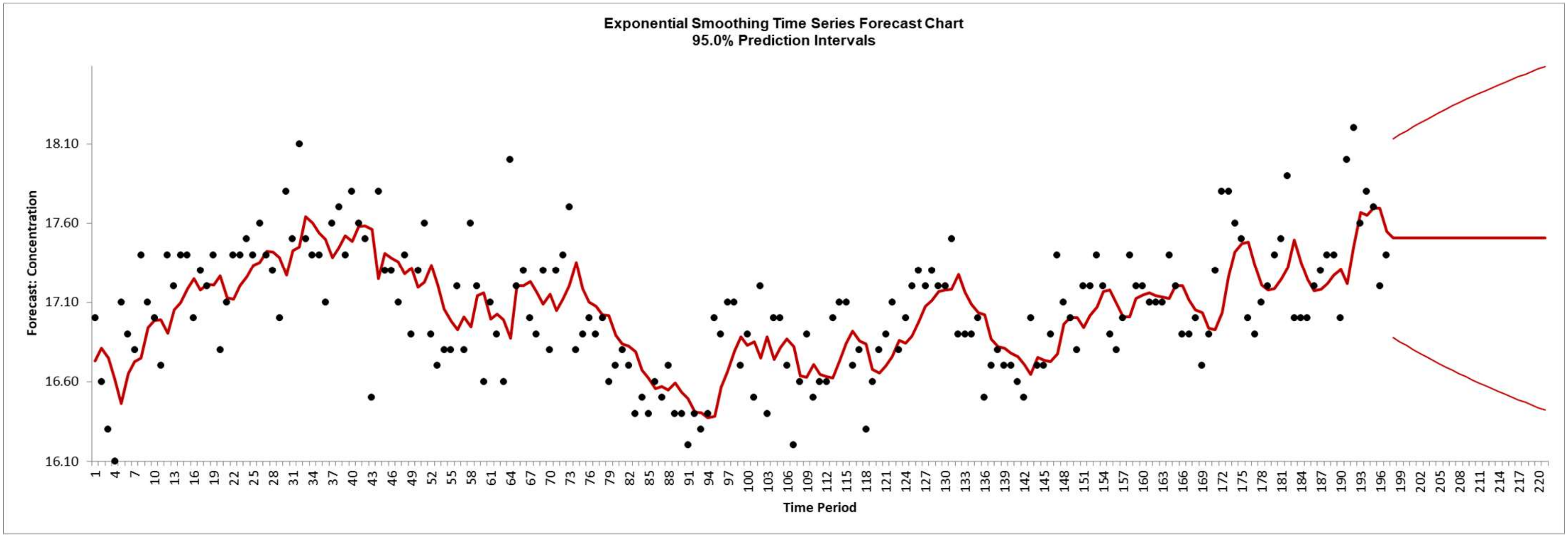
Example 1c: Box-Jenkins Series A - Chemical Process Concentration - Simple Exponential Smoothing (EWMA) Time Series Forecast





SigmaXL > Time Series Forecasting > Exponential Smoothing Forecast > Forecast

Example 1c: Box-Jenkins Series A - Chemical Process Concentration - Simple Exponential Smoothing (EWMA)



Exponential Smoothing Model: Simple Exponential Smoothing with Additive Errors (A, N, N) - Exponentially Weighted Moving Average (EWMA) - User Specified Model
Model Periods: All observations are used in the Exponential Smoothing model estimation. No withhold periods available for out-of-sample forecast accuracy evaluation.

Exponential Smoothing Model Information	
Seasonal Frequency	1
Model Selection Criterion	Specified
Box-Cox Transformation	N/A
Lambda	
Threshold	

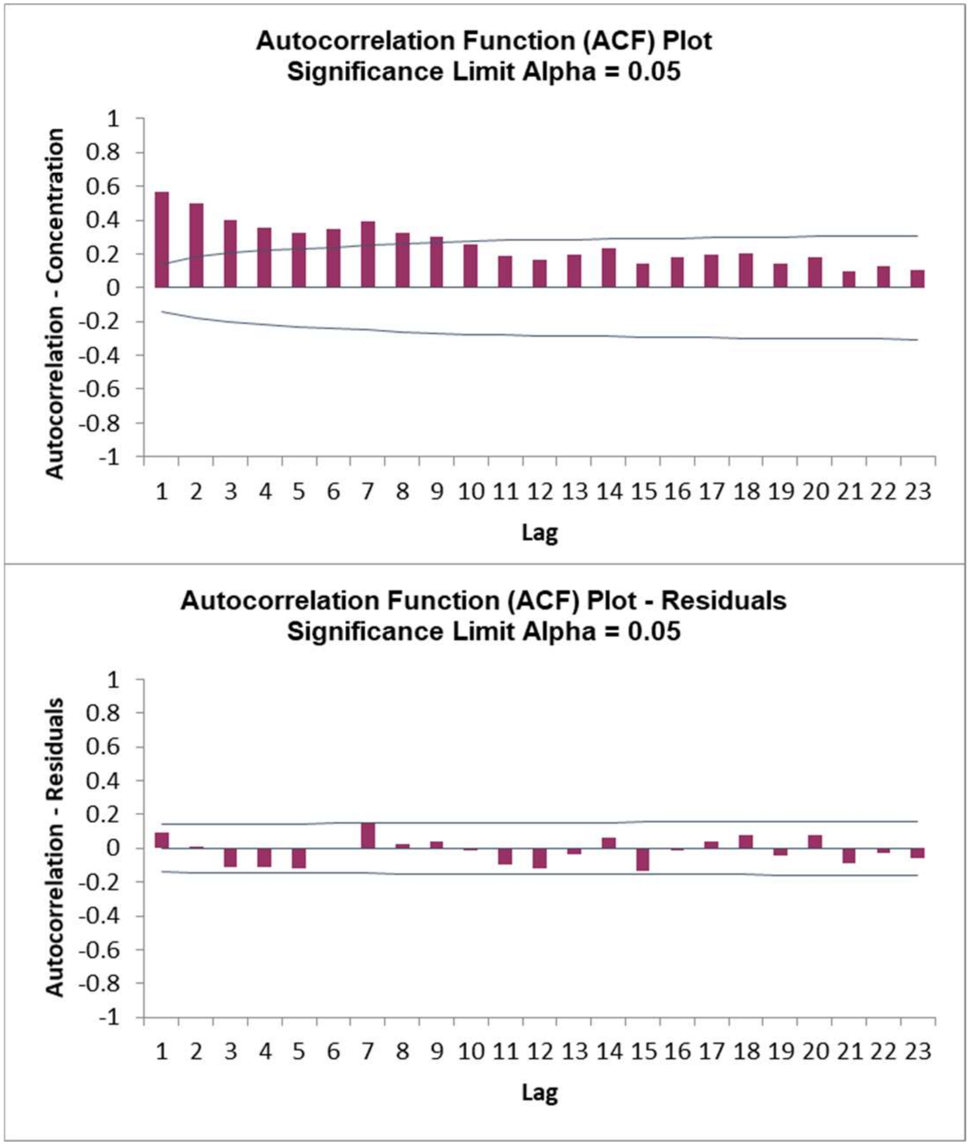
Parameter Estimates	
Term	Coefficient
alpha (level smoothing)	0.294785988
l (initial level)	16.73121246

Exponential Smoothing Model Statistics	
No. of Observations	197
DF	195
StDev	0.318189
Variance	0.101244
Log-Likelihood	-293.804
AICc	593.7316
AIC	593.6072
BIC	603.4568

Forecast Accuracy		
Metric	In-Sample (Estimation) One-Step-Ahead Forecast	Out-of-Sample (Withhold) One-Step-Ahead Forecast
N	197	
RMSE	0.316569334	
MAE	0.247329038	
MAPE	1.446520183	
MASE	0.897712804	

Simple Exponential Smoothing (EWMA) specified. 95% Prediction Intervals for forecast.

Example 1c: Box-Jenkins Series A - Chemical Process Concentration - ACF Plots (Raw Data versus Residuals)



Autocorrelation: Ljung-Box Test

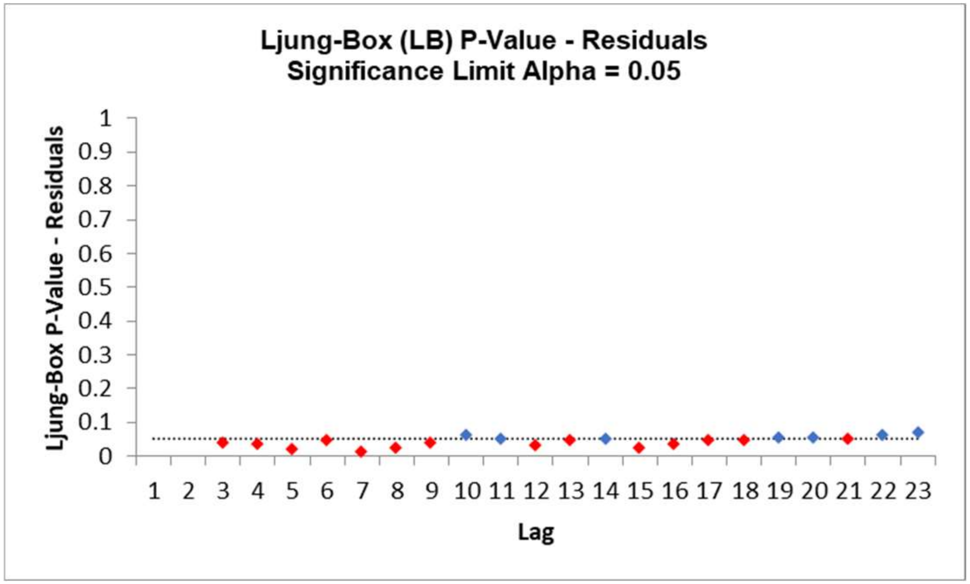
- In addition to looking at the ACF plot, we can also do a more formal test for autocorrelation by considering a whole set of r_k values as a group, rather than treating each one separately.

$$Q = T(T+2)\sum_{k=1}^{h}(T-k)^{-1}r_k^2,$$

  where h is the maximum lag being considered and T is the number of observations.

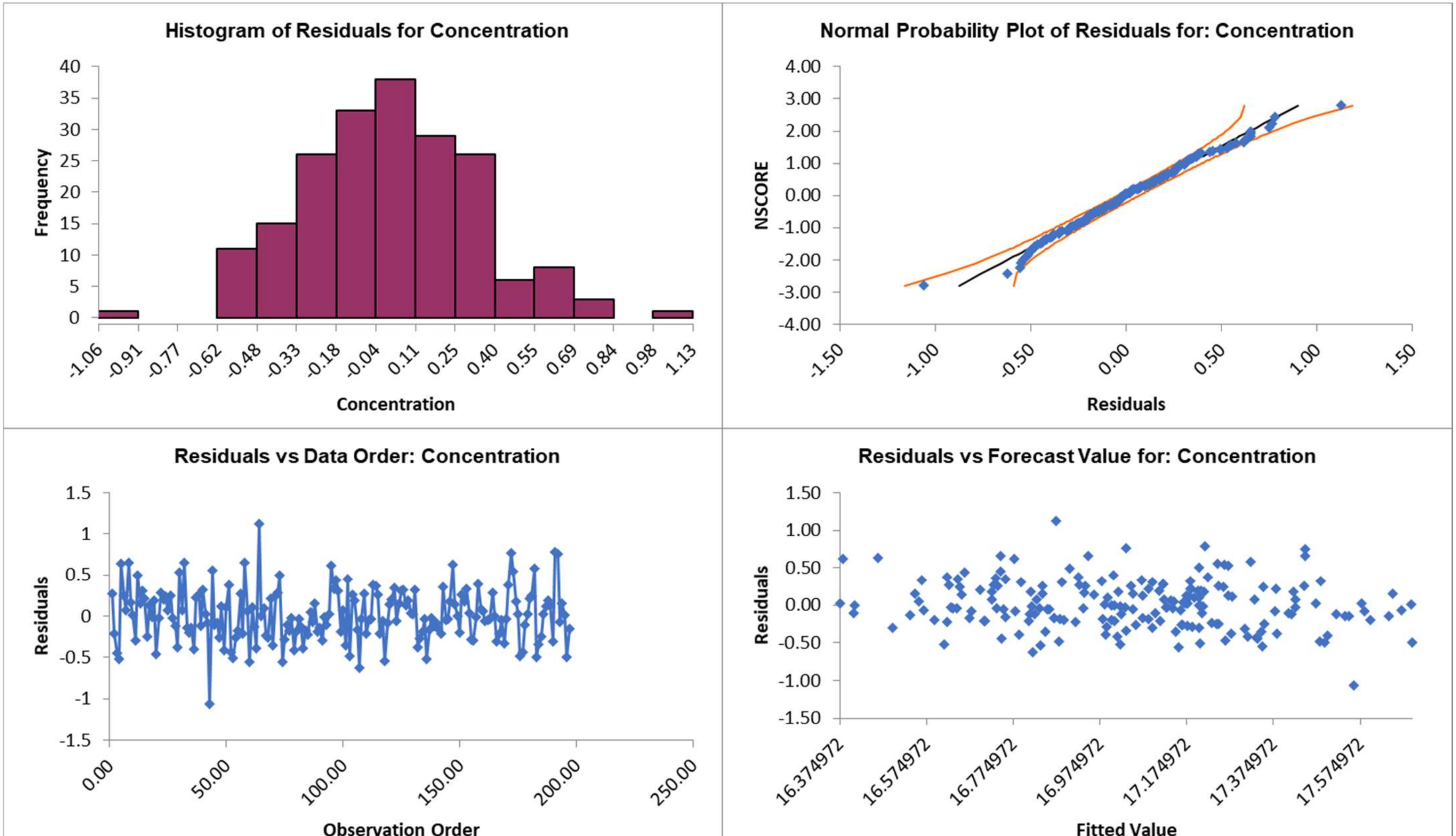
- If the autocorrelations did come from a white noise series, then Q would have a χ^2 distribution with $(h - k)$ degrees of freedom, where k is the number of parameters in the model [4].

Example 1c: Box-Jenkins Series A - Chemical Process Concentration - Ljung-Box P-Value Chart for Residuals



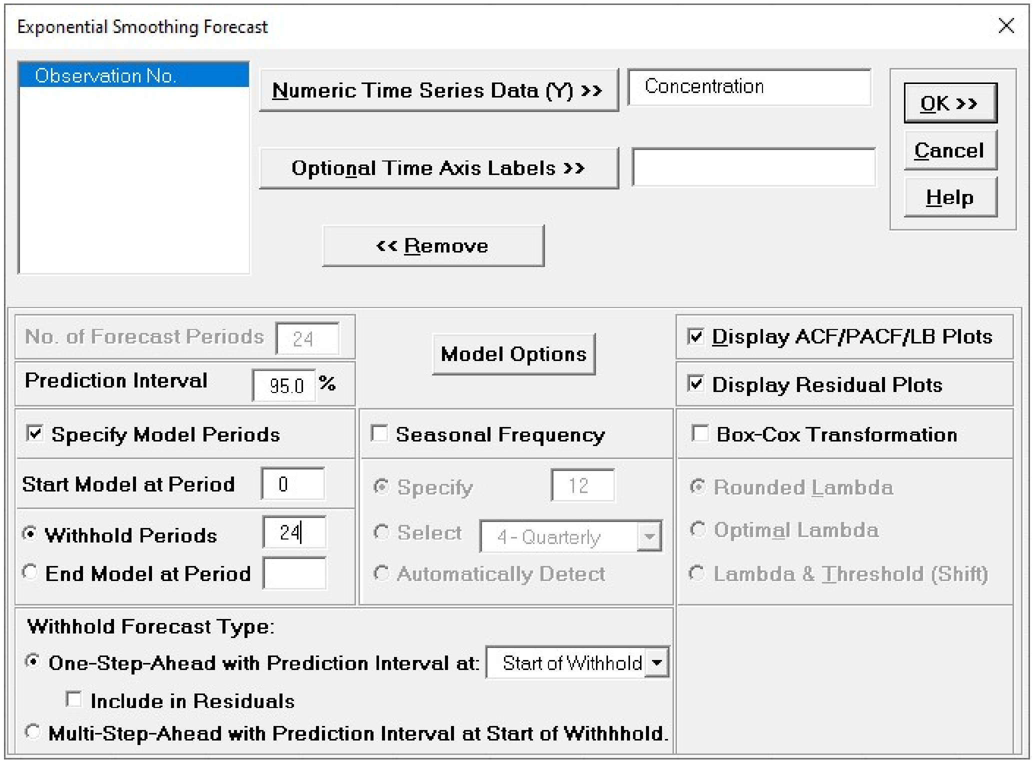
The previous ACF plots indicate that almost all of the correlation has been accounted for in the model, but the Ljung-Box plot shows that some significant autocorrelation still remains (P-Values < .05) - so the model can potentially be improved. This does not mean that the model is a bad model, it can still be very useful for prediction purposes, but the prediction intervals may not provide accurate coverage.

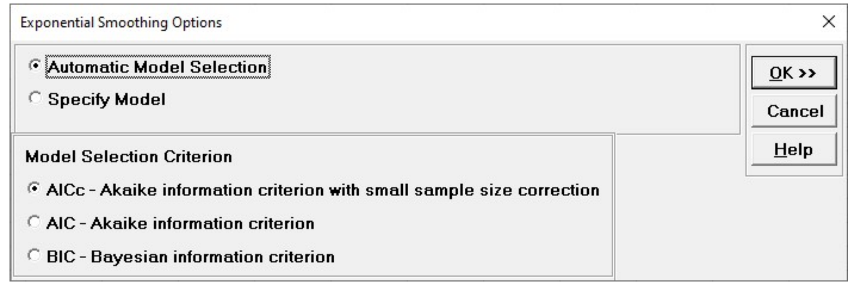
Example 1c: Box-Jenkins Series A - Chemical Process Concentration - Residuals



Residuals look good – approximately normal with equal variance.

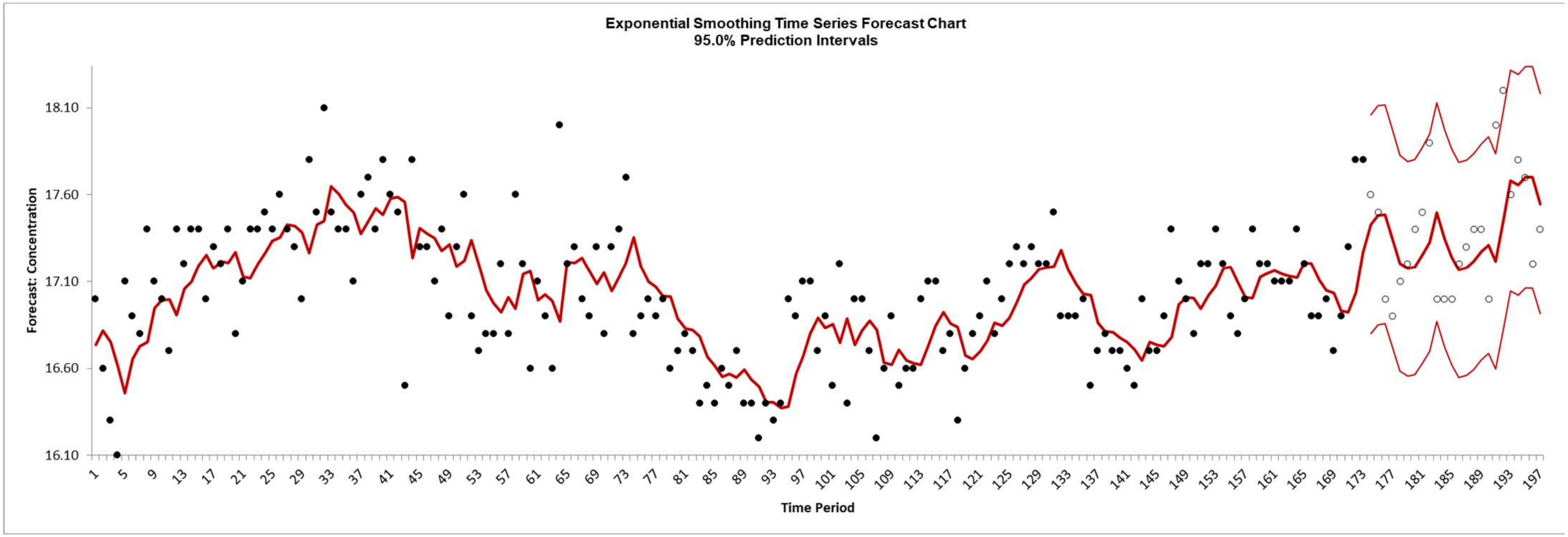
Example 1d: Box-Jenkins Series A - Chemical Process Concentration – Automatic Model Selection and Withhold Sample (One-Step-Ahead)





SigmaXL > Time Series Forecasting > Exponential Smoothing Forecast > Forecast

Example 1d: Box-Jenkins Series A - Chemical Process Concentration – Automatic Model Selection and Withhold Sample (One-Step-Ahead)



Exponential Smoothing Model: Simple Exponential Smoothing with Multiplicative Errors (M, N, N) - Model Automatically Selected
Model Periods: Model parameter estimates calculated excluding 24 withhold periods.

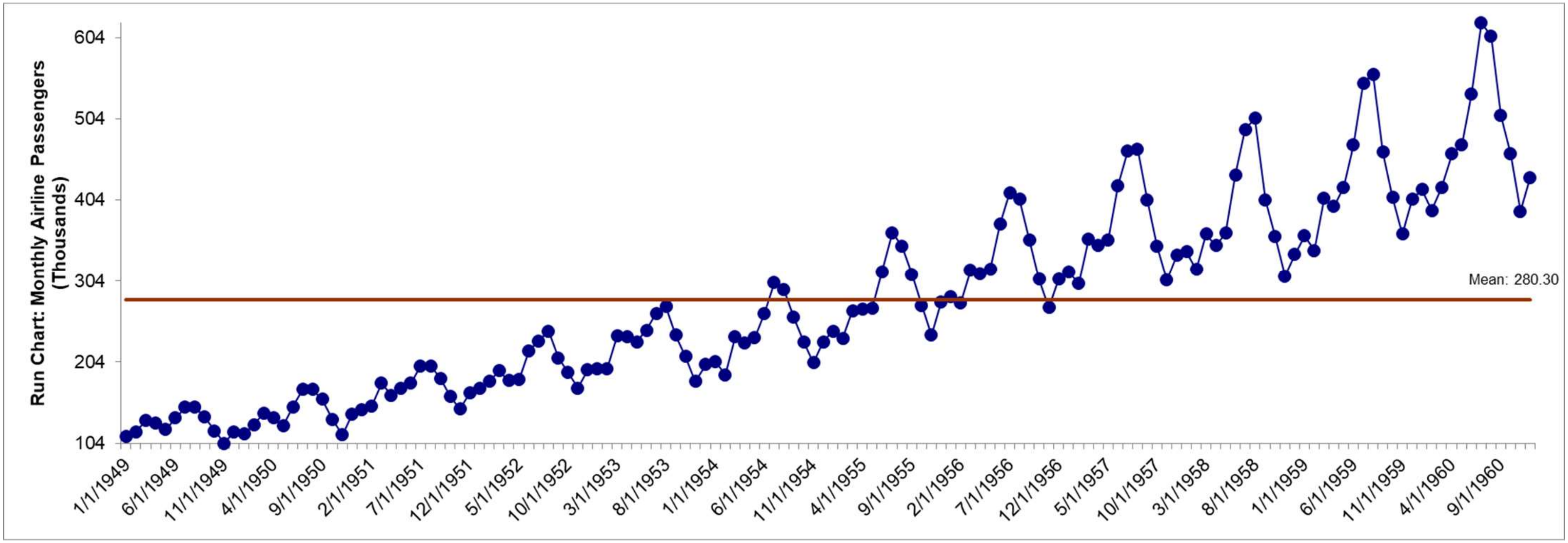
Exponential Smoothing Model Information	
Seasonal Frequency	1
Model Selection Criterion	AICc
Box-Cox Transformation	N/A
Lambda	
Threshold	

Parameter Estimates	
Term	Coefficient
alpha (level smoothing)	0.303967286
l (initial level)	16.73554259

Exponential Smoothing Model Statistics	
No. of Observations	173
DF	171
StDev	0.01841
Variance	0.000339
Log-Likelihood	-243.805
AICc	493.7523
AIC	493.6102
BIC	503.0701

Forecast Accuracy		
Metric	In-Sample (Estimation) One-Step-Ahead Forecast	Out-of-Sample (Withhold) One-Step-Ahead Forecast
N	173	24
RMSE	0.31133921	0.35209899
MAE	0.243032918	0.273389284
MAPE	1.425751706	1.567334268
MASE	0.878186174	0.987877246

Example 2a: Box-Jenkins Series G – Monthly Airline Passengers - Run Chart

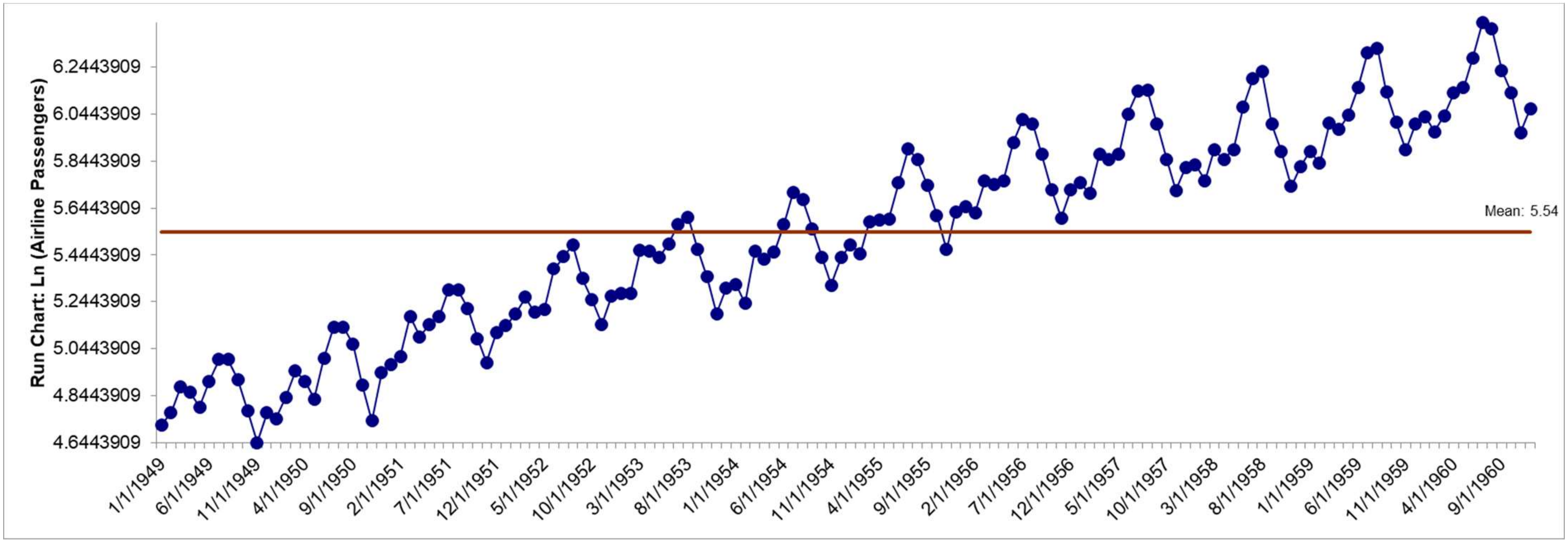


Data shows strong positive trend, strong seasonality (monthly data) and seasonal variance increases over time.

SigmaXL > Time Series Forecasting > Run Chart

Example 2a: Monthly Airline Passengers – Series G.xlsx – Monthly Airline Passengers

Example 2b: Box-Jenkins Series G – Ln(Monthly Airline Passengers) - Run Chart

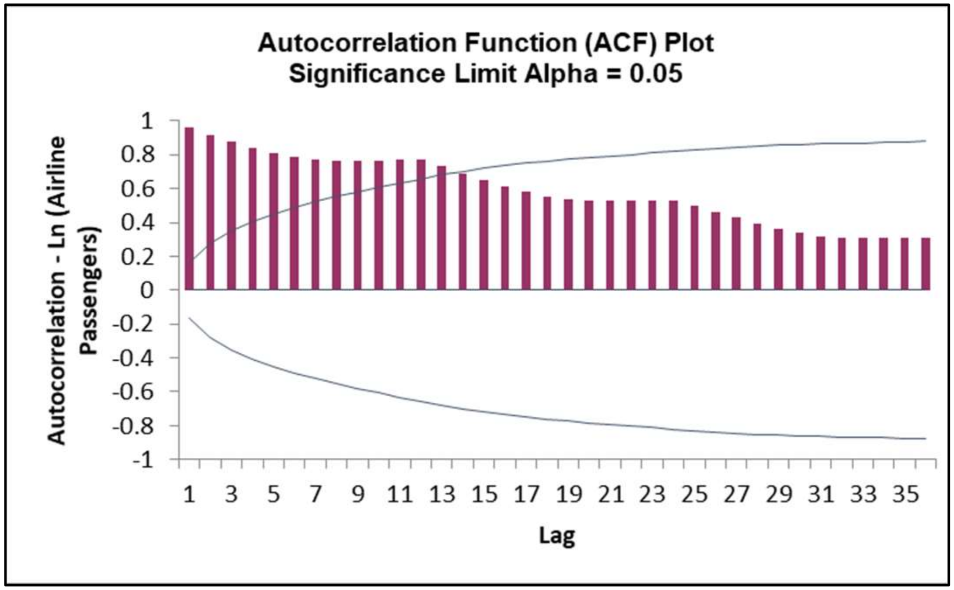


Data shows strong positive trend, strong seasonality (monthly data). Seasonal variance is now stable over time.

SigmaXL > Time Series Forecasting > Run Chart

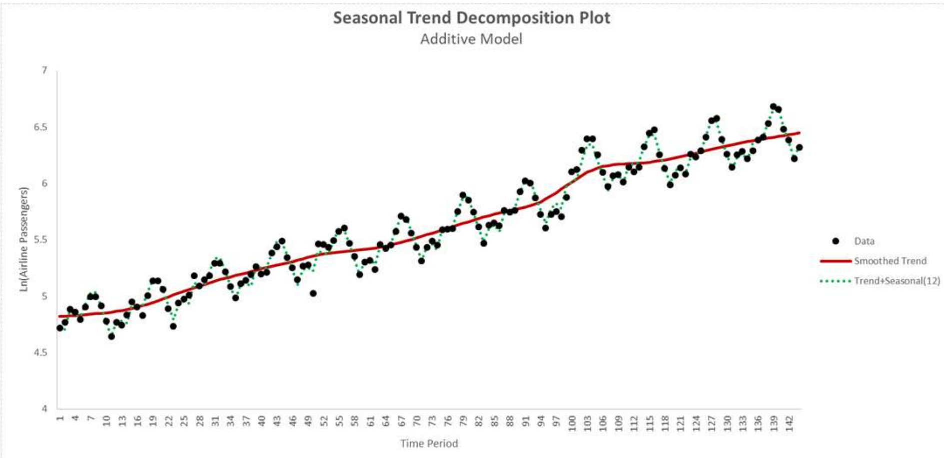
Example 2a: Monthly Airline Passengers – Series G.xlsx – Ln(Monthly Airline Passengers)

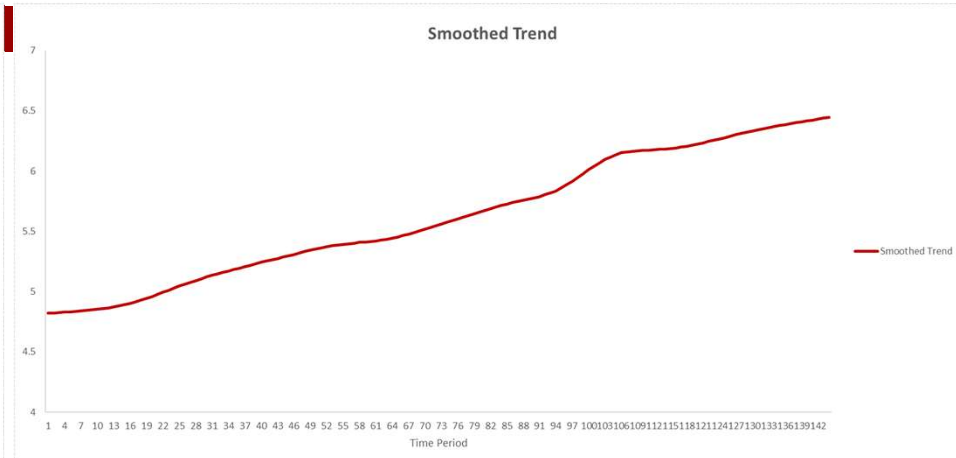
Example 2b: Box-Jenkins Series G – Ln(Monthly Airline Passengers) - Autocorrelation (ACF) Plot

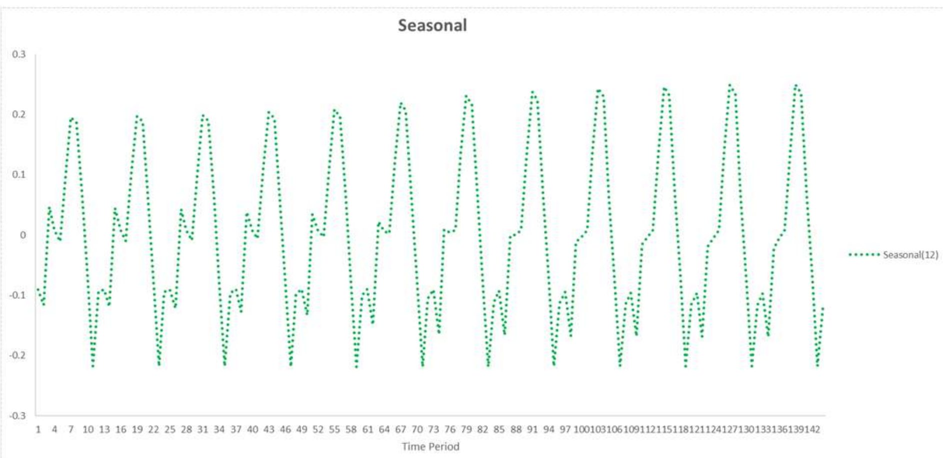


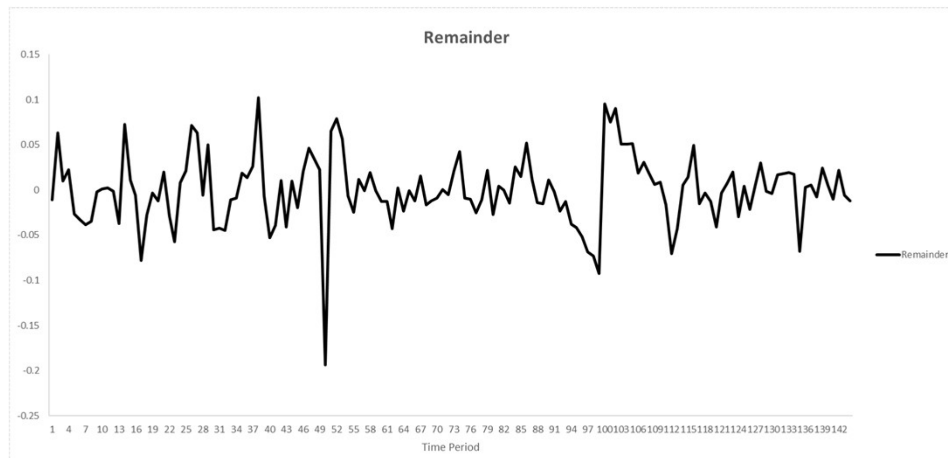
SigmaXL > Time Series Forecasting > Autocorrelation (ACF/PACF) Plots

Example 2b: Box-Jenkins Series G – Ln(Monthly Airline Passengers)



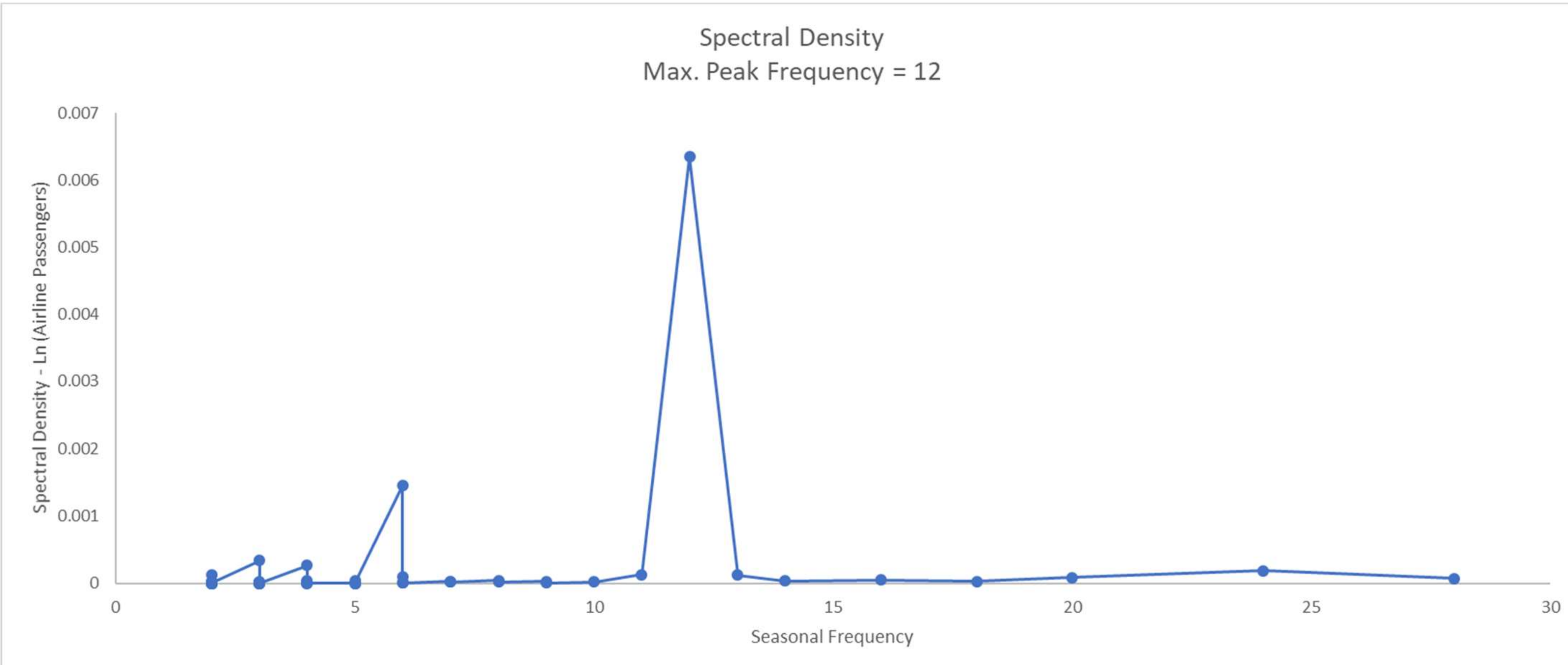






SigmaXL > Time Series Forecasting > Seasonal Trend Decomposition Plots

Example 2b: Box-Jenkins Series G – Ln(Monthly Airline Passengers) - Spectral Density Plot



SigmaXL > Time Series Forecasting > Spectral Density Plot

Error, Trend, Seasonal (ETS) Exponential Smoothing Models

- Error, Trend, Seasonal (ETS) models expand on simple exponential smoothing to accommodate trend and seasonal components as well as additive or multiplicative errors.
- Simple Exponential Smoothing is an Error Model.
- Error, Trend model is Holt's Linear, also known as double exponential smoothing.

Error, Trend, Seasonal (ETS) Exponential Smoothing Models

- Error, Trend, Seasonal model is Holt-Winters, also known as triple exponential smoothing.
  - Seasonal frequency must be specified:
    - Quarterly data = 4 (observations per year)
    - Monthly data = 12 (observations per year)
    - Daily data = 7 (observations per week)
    - Hourly data = 24 (observations per day)
  - Frequency is the number of observations per "cycle". This is the opposite of the definition of frequency in physics, or in engineering Fourier analysis, where "period" is the length of the cycle, and "frequency" is the inverse of period.

Reference: https://robjhyndman.com/hyndsight/seasonal-periods/

Error, Trend, Seasonal (ETS) models Hyndman’s Taxonomy

- Rob Hyndman has developed a complete taxonomy that describes all of the combinations of exponential smooth models in a consistent manner. [4]
- Error:
  - Additive or Multiplicative
  - The point forecasts produced by the models are identical if they use the same smoothing parameter values. Multiplicative will, however, generate different prediction intervals to accommodate change in variance.
  - An alternative to multiplicative is to use the Ln transformation (Box-Cox transformation with Lambda = 0).
  - Error models include the smoothing parameter α and initial level value.

Error, Trend, Seasonal (ETS) models Hyndman's Taxonomy

- Trend:
  - None, Additive, Additive Damped
  - Multiplicative Trend is not recommended as they tend to produce poor forecasts
  - Trend models add a smoothing parameter β and initial trend value.
  - Damped trend models add a smoothing parameter ϕ that "dampens" the trend to a flat line some time in the future.

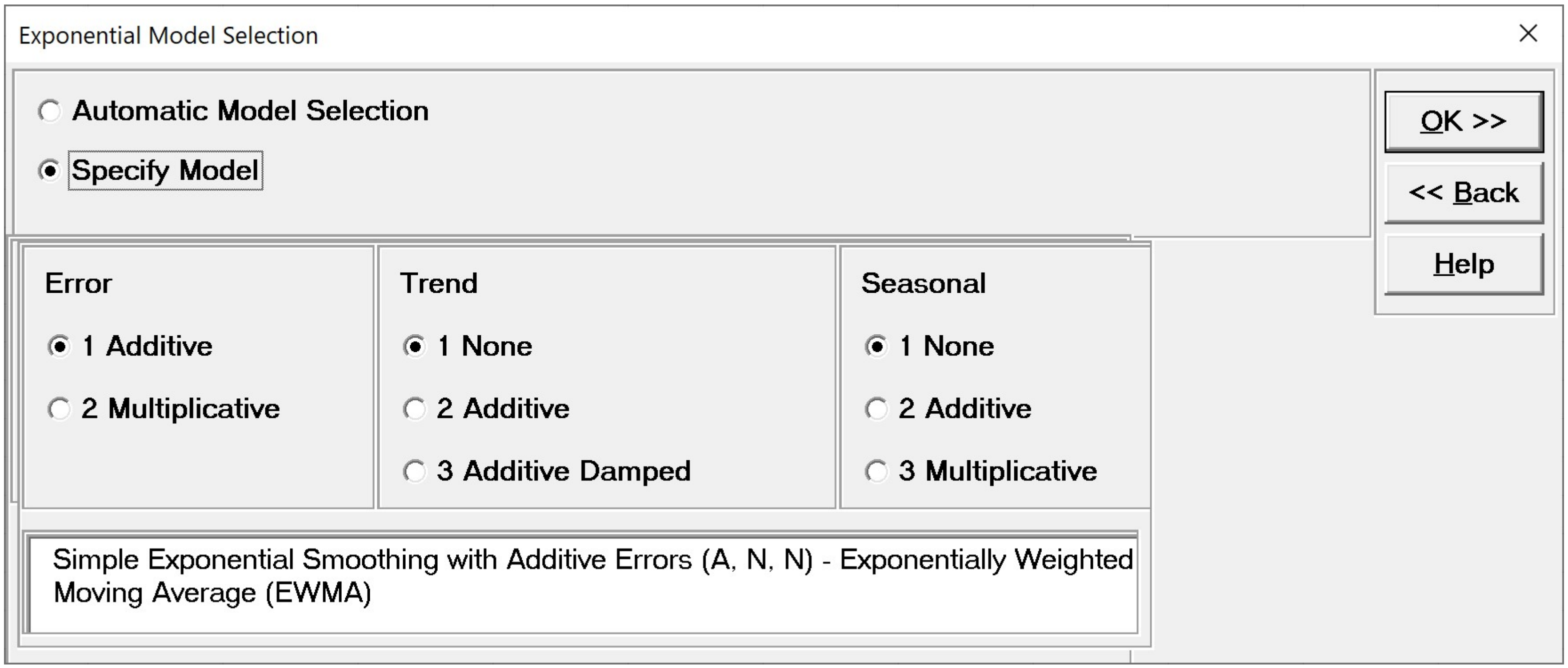
Error, Trend, Seasonal (ETS) models Hyndman's Taxonomy

- Seasonal:
  - None, Additive, Multiplicative
  - Seasonal models add a smoothing parameter γ and initial seasonal values.
    - # of initial values = seasonal frequency – 1
    - constrained to sum to 0 for additive or 12 for multiplicative

Error, Trend, Seasonal (ETS) models
Hyndman's Taxonomy

Short hand (Error, Trend, Seasonal)	Method
(A, N, N)	Simple Exponential Smoothing with Additive Errors – Exponentially Weighted Moving Average (EWMA)
(M, N, N)	Simple Exponential Smoothing with Multiplicative Errors
(A, A, N)	Additive Trend Method with Additive Errors (Holt's Linear)
(M, A, N)	Additive Trend Method with Multiplicative Errors (Holt's Linear)
(A, A, A)	Additive Trend, Additive Seasonal Method with Additive Errors (Holt-Winters)
(M, A, A)	Additive Trend, Additive Seasonal Method with Multiplicative Errors (Holt-Winters)
(A, Ad, A)	Additive Damped Trend, Additive Seasonal Method with Additive Errors
(M, Ad, A)	Additive Damped Trend, Additive Seasonal Method with Multiplicative Errors

Error, Trend, Seasonal (ETS) models Hyndman's Taxonomy

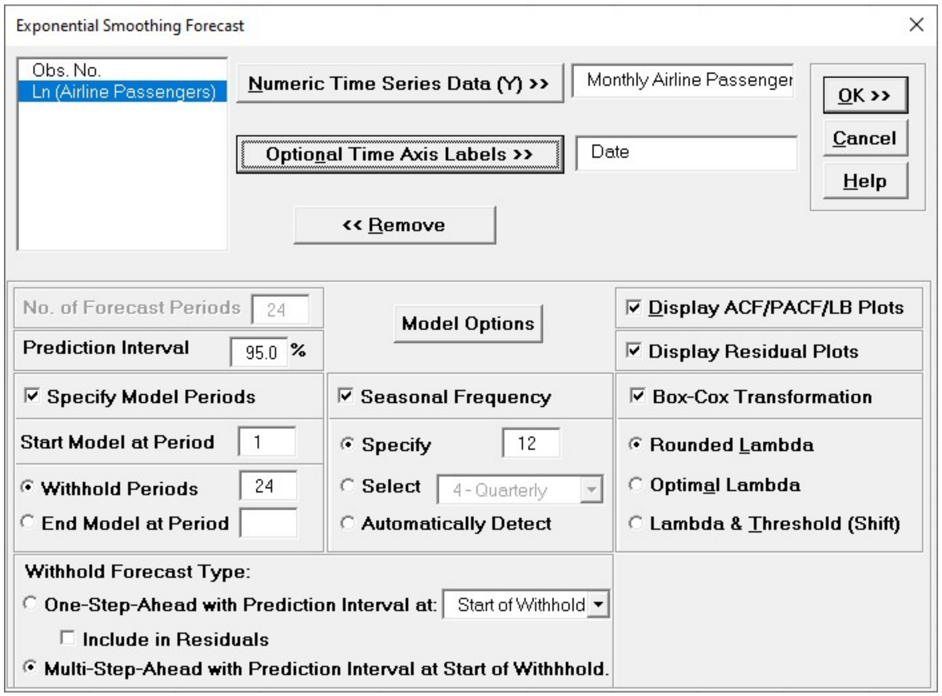


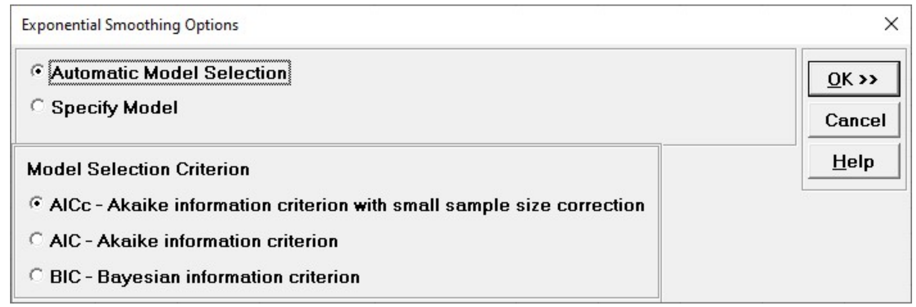
SigmaXL > Time Series Forecasting > Exponential Smoothing Forecast > Forecast

Error, Trend, Seasonal (ETS) Automatic Model Selection

- AICc is recommended as the default Information Criteria, based on forecast error performance with M3 competition data (see appendix for more information on forecast competitions).
- Some of the model combinations lead to numerical instability and are not considered in the selection process: (A,N,M) (A,A,M) (A,Ad,M)
- If a Box-Cox transformation is used, Multiplicative models are not considered.

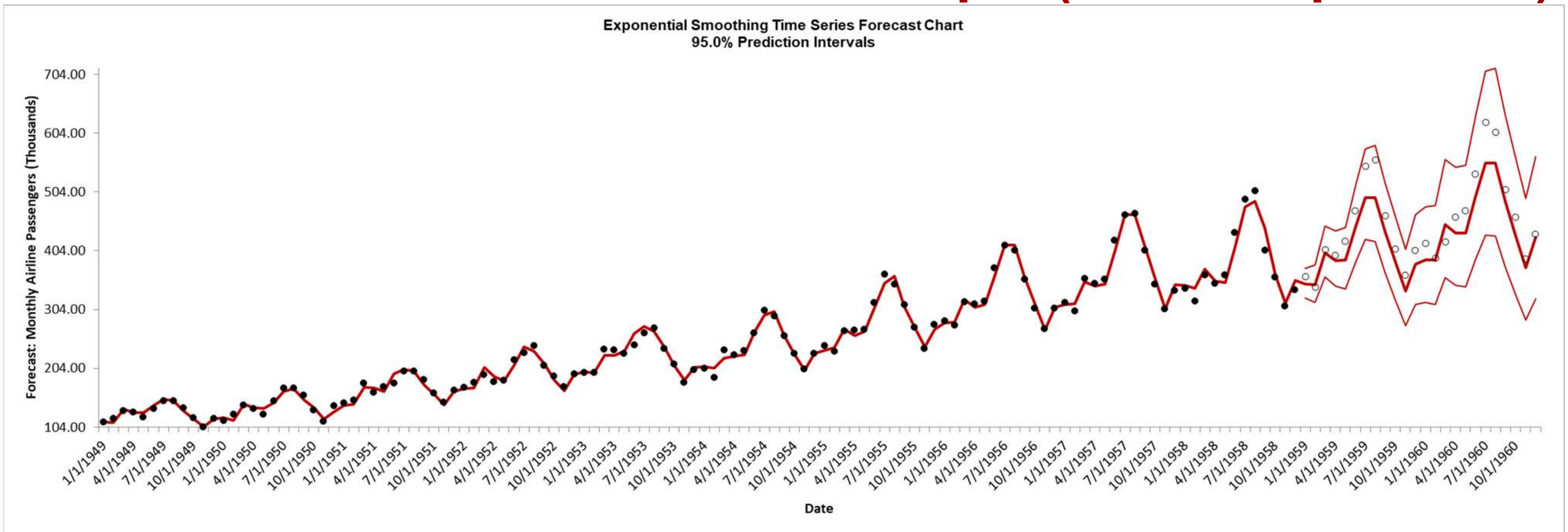
Example 2c: Box-Jenkins Series G – Monthly Airline Passengers – Automatic Model Selection, Box-Cox Transformation and Withhold Sample (Multi-Step-Ahead)





SigmaXL > Time Series Forecasting > Exponential Smoothing Forecast > Forecast

Example 2c: Box-Jenkins Series G – Monthly Airline Passengers – Automatic Model Selection, Box-Cox Transformation and Withhold Sample (Multi-Step-Ahead)



Exponential Smoothing Model: Additive Trend, Additive Seasonal Method with Additive Errors (Holt-Winters) (A, A, A) - Model Automatically Selected
Model Periods: Model parameter estimates calculated excluding 24 withhold periods.

Exponential Smoothing Model Information	
Seasonal Frequency	12
Model Selection Criterion	AICc
Box-Cox Transformation	Rounded Lambda
Lambda	0
Threshold	0

Parameter Estimates	
Term	Coefficient
alpha (level smoothing)	0.770415301
beta (trend smoothing)	0.0001
gamma (seasonal smoothing)	0.0001
l (initial level)	4.80925762
b (initial trend)	0.00935261
s1 (initial seasonal)	-0.099989196
s2 (initial seasonal)	-0.217929483
s3 (initial seasonal)	-0.076652841
s4 (initial seasonal)	0.063787774
s5 (initial seasonal)	0.198157807
s6 (initial seasonal)	0.205968013
s7 (initial seasonal)	0.108472035
s8 (initial seasonal)	-0.016634733
s9 (initial seasonal)	-0.008278553
s10 (initial seasonal)	0.033611999
s11 (initial seasonal)	-0.101347496
s12 (initial seasonal)	-0.089165324

Exponential Smoothing Model Statistics	
No. of Observations	120
DF	104
StDev	0.037144
Variance	0.00138
Log-Likelihood	116.4916
AICc	-192.983
AIC	-198.983
BIC	-151.596

Forecast Accuracy		
Metric	In-Sample (Estimation) One-Step-Ahead Forecast	Out-of-Sample (Withhold) Multi-Step-Ahead Forecast
N	120	24
RMSE	8.67099231	33.06342473
MAE	6.481327147	27.83270693
MAPE	2.730421504	5.80549634
MASE	0.226825448	0.974054552

ETS Additive Trend, Additive Seasonal Method with Additive Errors (Holt-Winters) (A, A, A) **automatically selected**. Seasonal Frequency = 12 (Monthly data).

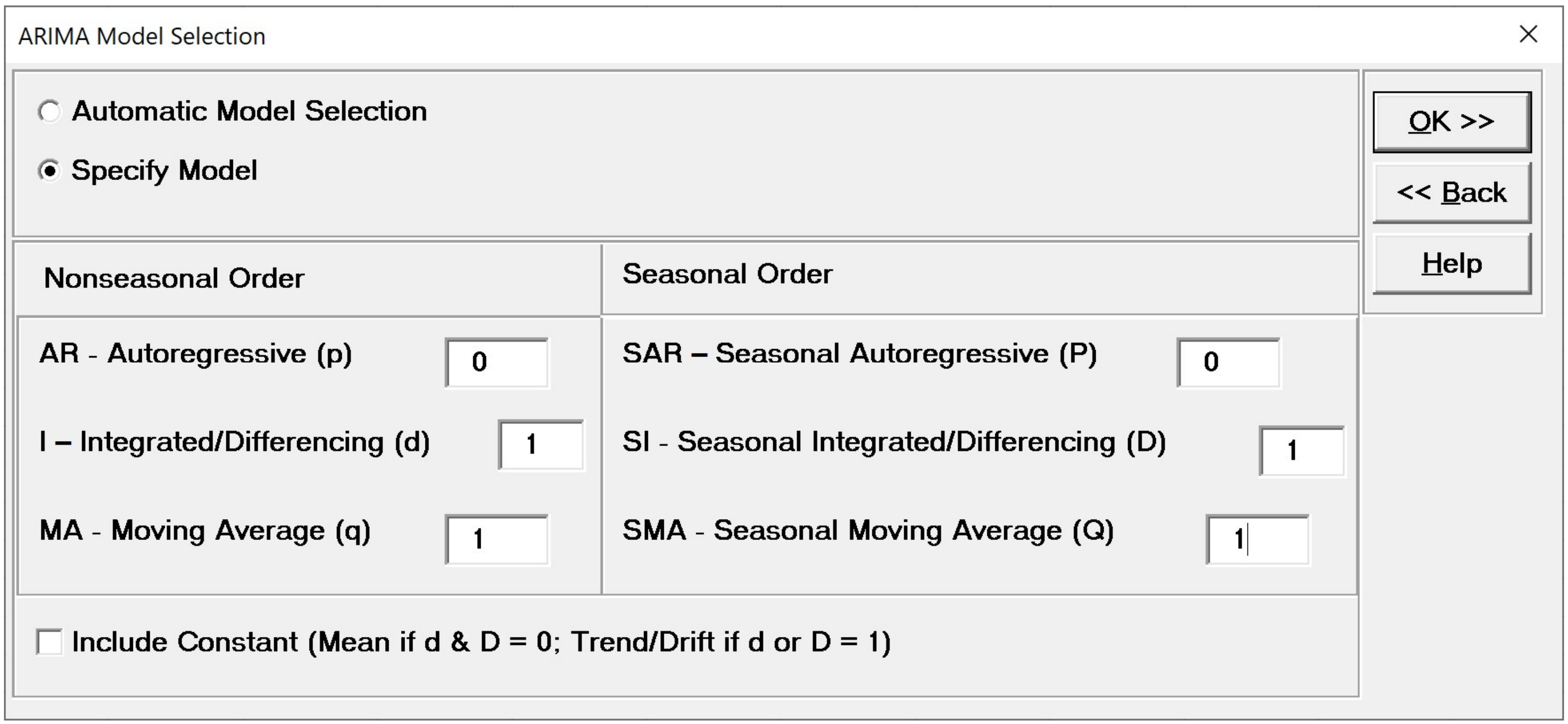
SigmaXL > Time Series Forecasting > Exponential Smoothing Forecast > Forecast

Box-Jenkins AutoRegressive Integrated Moving Average (ARIMA) Models

- An ARIMA model includes an Autoregressive (AR) component of order p, an Integrated/Differencing component of order d and a Moving Average component of order q and an optional constant.

- An ARIMA Seasonal model includes a Seasonal Autoregressive (SAR) component of order P, a Seasonal Integrated/Differencing component of order D and a Seasonal Moving Average component of order Q.

Box-Jenkins AutoRegressive Integrated Moving Average (ARIMA) Models



SigmaXL > Time Series Forecasting > ARIMA Forecast > Forecast

Box-Jenkins AutoRegressive Integrated Moving Average (ARIMA) Models - Stationarity

- ARIMA assumes that the time series is stationary, i.e., it has the property that the mean, variance and autocorrelation structure do not change over time.
- If a time series mean is not stationary (e.g. trending), this can be corrected by differencing, computing the differences between consecutive observations for non-seasonal and between consecutive periods (e.g. months) for seasonal data (Jan 2019 – Jan 2018, etc.).
- For non-seasonal, this may involve 1 or 2 orders of differencing. This order is the Integrated term *d.*
- For seasonal, this may involve 1 order of differencing. This order is the Seasonal Integrated term *D*.

Box-Jenkins AutoRegressive Integrated Moving Average (ARIMA) Models - Stationarity

- If d+D = 0, a constant term in the model is the mean.
- If d+D = 1, a constant term in the model is a trend/drift.
- If d+D > 1, a constant term would be a quadratic trend, so constant should not be included.
- It is recommended that d+D should not be > 3.
- If the variance is not stationary, use a Box-Cox transformation.
- In the Ln(Monthly Airline Passenger) data we are starting with Ln data to deal with non-stationary variance in the raw data.

Box-Jenkins AutoRegressive Integrated Moving Average (ARIMA) Models – AR

- In an autoregression model, we forecast the variable of interest using a linear combination of past values of the variable. The term autoregression indicates that it is a regression of the variable against itself.

$$y_t = c + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_p y_{t-p} + \varepsilon_t$$

where ε_t is white noise [4].

Box-Jenkins AutoRegressive Integrated Moving Average (ARIMA) Models – MA

- Rather than using past values of the forecast variable in a regression, a moving average model uses past forecast errors in a regression-like model [4].

$$y_t = c + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \cdots + \theta_q \varepsilon_{t-q}$$

- Model parameters are solved using Kalman Filters and nonlinear minimization. This permits exact calculations (backcasting is not required) and can handle missing values.

Box-Jenkins AutoRegressive Integrated Moving Average (ARIMA) Models

If we combine differencing with autoregression and a moving average model, we obtain a non-seasonal ARIMA model.

$$y'_t = c + \phi_1 y'_{t-1} + \cdots + \phi_p y'_{t-p} + \theta_1 \varepsilon_{t-1} + \cdots + \theta_q \varepsilon_{t-q} + \varepsilon_t$$

where y'_t is the differenced series.

- For seasonal, the model consists of terms that are similar to the non-seasonal components of the model. The seasonal model is ARIMA (*P,D,Q*) and combined we have ARIMA (*p,d,q*) (*P,D,Q*).

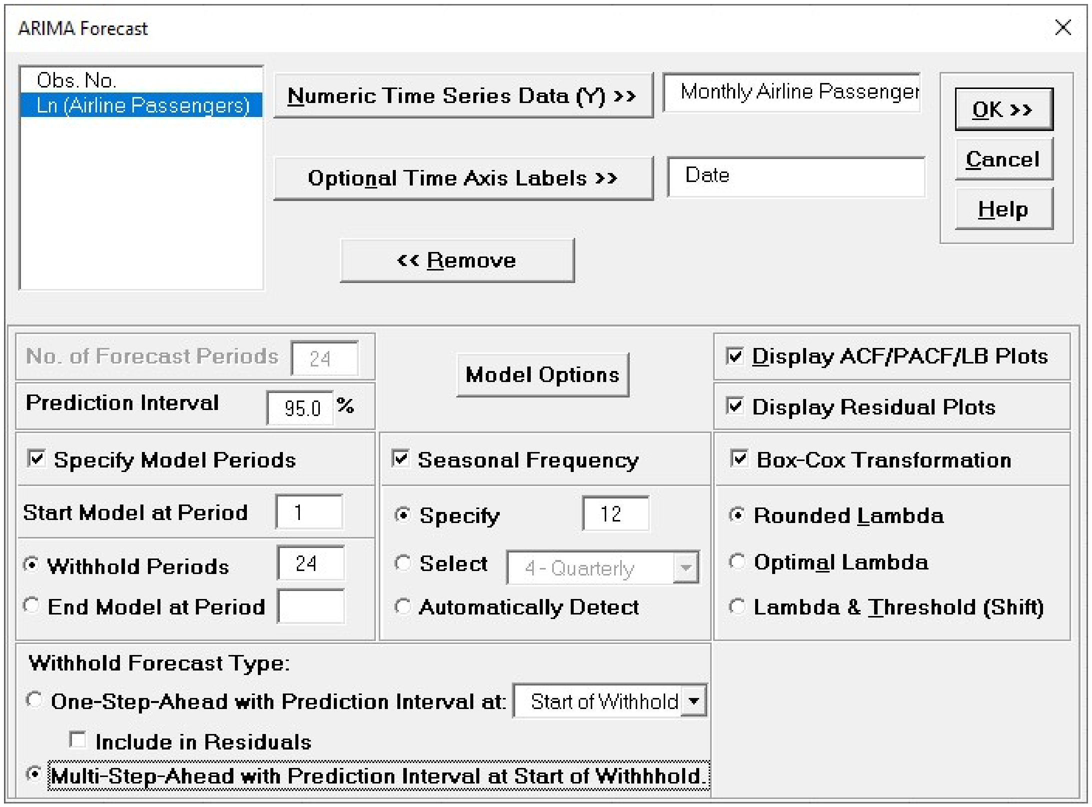
Partial Autocorrelation (PACF)

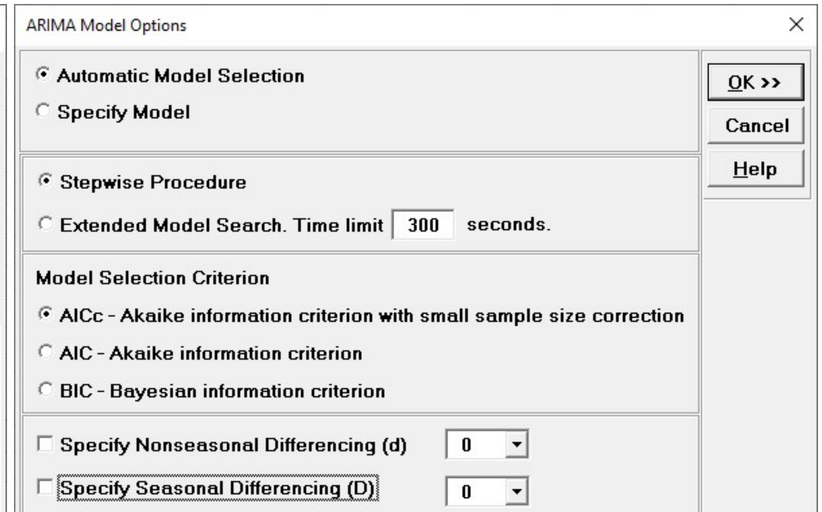
- Partial Autocorrelation plots are similar to Autocorrelation plots but adjust for correlation inherent in lags, e.g., y_t and y_{t-2} might be correlated, simply because they are both connected to y_{t-1}, rather than because of any new information contained in y_{t-2} [4].
- Each partial autocorrelation can be estimated as the last coefficient in an autoregressive model. Specifically, α_k, the kth partial autocorrelation coefficient, is equal to the estimate of ϕ_k in an AR(k) model.
- They are typically used in ARIMA to help determine the order of terms in the model, but are also useful as a general diagnostic tool.

Box-Jenkins AutoRegressive Integrated Moving Average (ARIMA) Models – Model Selection

- ACF and PACF plots may be used to assist in determining what order values to use, but this requires a high level of expertise.
- Hyndman and Khandakar [5] give a stepwise procedure to determine optimal order values:
  - Use a Seasonal strength test to determine if D=0 or 1
  - Use a test for stationarity (KPSS) to determine if d=0, 1 or 2
  - With the differenced data, apply a stepwise procedure to solve for p, q, P, Q selecting models with minimum AICc.

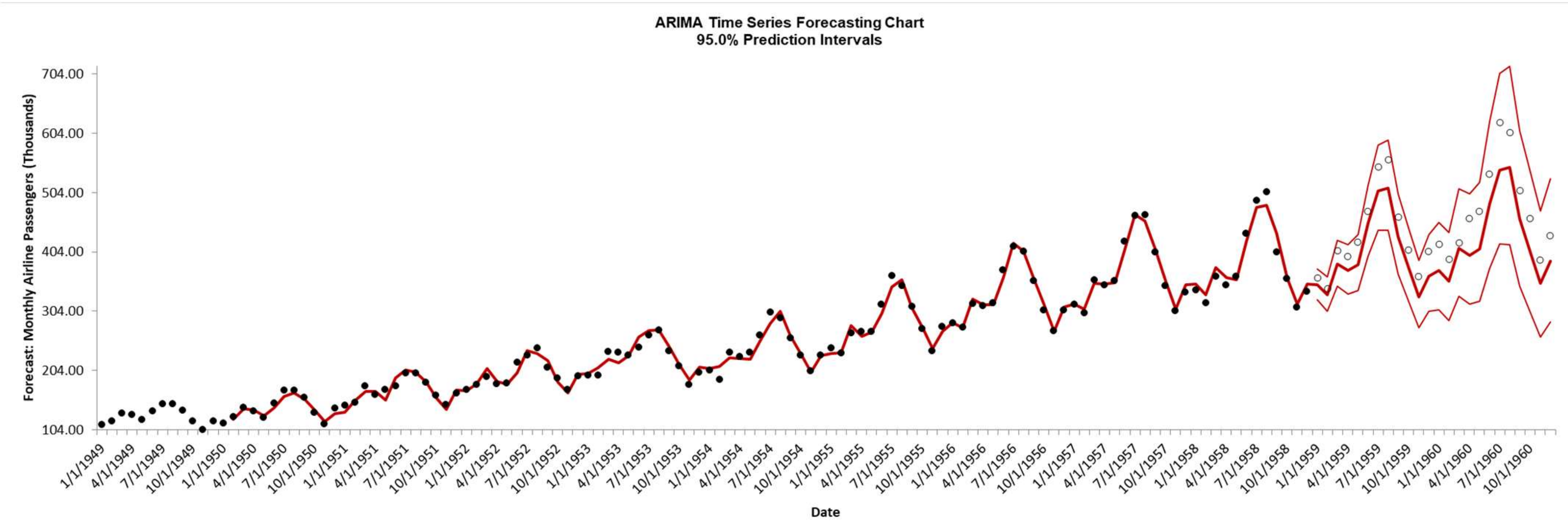
Example 2d: Box-Jenkins Series G – Monthly Airline Passengers – Automatic Model Selection, Box-Cox Transformation and Withhold Sample (Multi-Step-Ahead)





SigmaXL > Time Series Forecasting > ARIMA Forecast > Forecast

Example 2d: Box-Jenkins Series G – Monthly Airline Passengers – Automatic Model Selection, Box-Cox Transformation and Withhold Sample (Multi-Step-Ahead)



ARIMA Model: Monthly Airline Passengers (Thousands) - Model Automatically Selected
Model Periods: Model parameter estimates calculated excluding 24 withhold periods.

ARIMA Model Summary	
AR Order (p)	0
I Order (d)	1
MA Order (q)	1
SAR Order (P)	0
SI Order (D)	1
SMA Order (Q)	1
Seasonal Frequency	12
Include Constant	0
No. of Predictors	0
Model Selection Criterion	AICc
Box-Cox Transformation	Rounded Lambda
Lambda	0
Threshold	0

Parameter Estimates				
Term	Coefficient	SE Coefficient	T	P
MA_1	0.342313249	0.100902427	3.392517	0.0010
SMA_1	0.540469465	0.087677292	6.164304	0.0000

ARIMA Model Statistics	
No. of Observations	120
DF	105
StDev	0.037414431
Variance	0.00139984
Log-Likelihood	197.5047754
AICc	-388.7765411
AIC	-389.0095508
BIC	-380.9910643

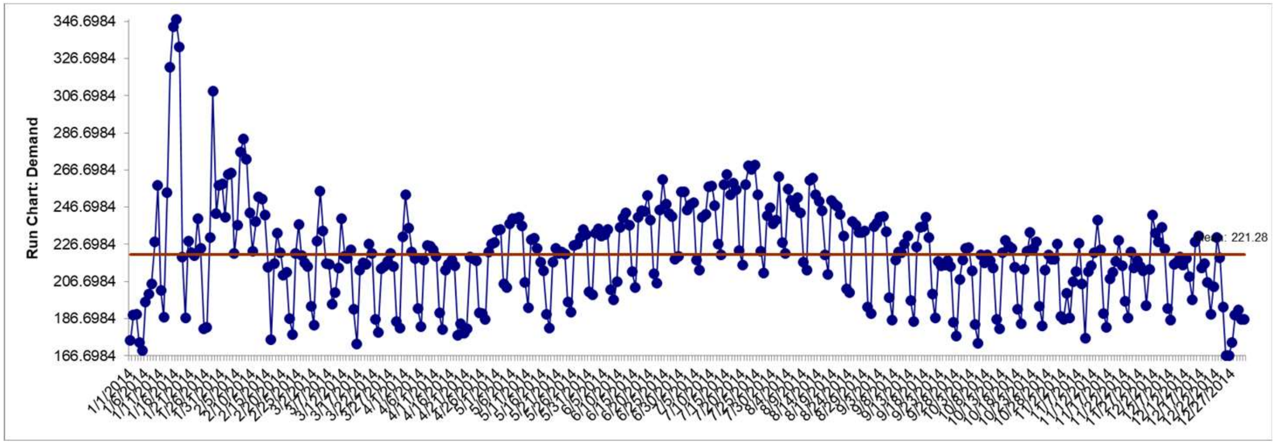
Forecast Accuracy		
Metric	In-Sample (Estimation) One-Step-Ahead Forecast	Out-of-Sample (Withhold) Multi-Step-Ahead Forecast
N	107	24
RMSE	9.44337388	43.18833723
MAE	7.384036891	39.45185993
MAPE	3.001604734	8.51734062
MASE	0.258417364	1.380687256

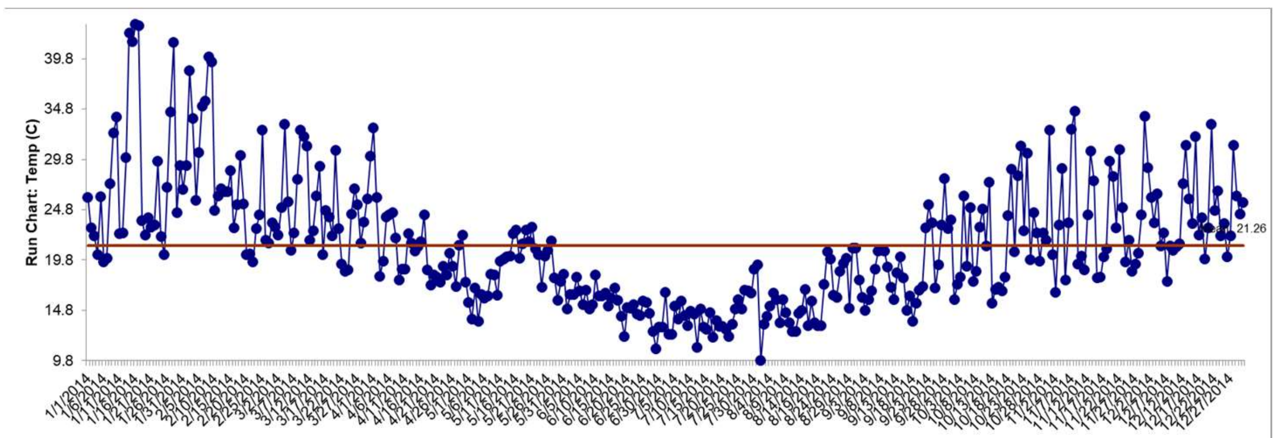
ARIMA (0,1,1) (0,1,1) **automatically selected**. Seasonal Frequency = 12 (Monthly data).

ARIMA with Predictors

- The ARIMA model supports continuous or categorical predictors, similar to multiple regression.
- In order to provide a forecast, additional predictor (X) values must be added to the dataset prior to running the analysis. The number of forecast periods will be equal to the number of additional predictor rows. Alternatively, the predictor values from a withhold sample may be used.
- As with multiple linear regression, predictors should not be strongly correlated.

Example 3: Daily Electricity Demand with Temperature and Work Day Predictors – Run Charts

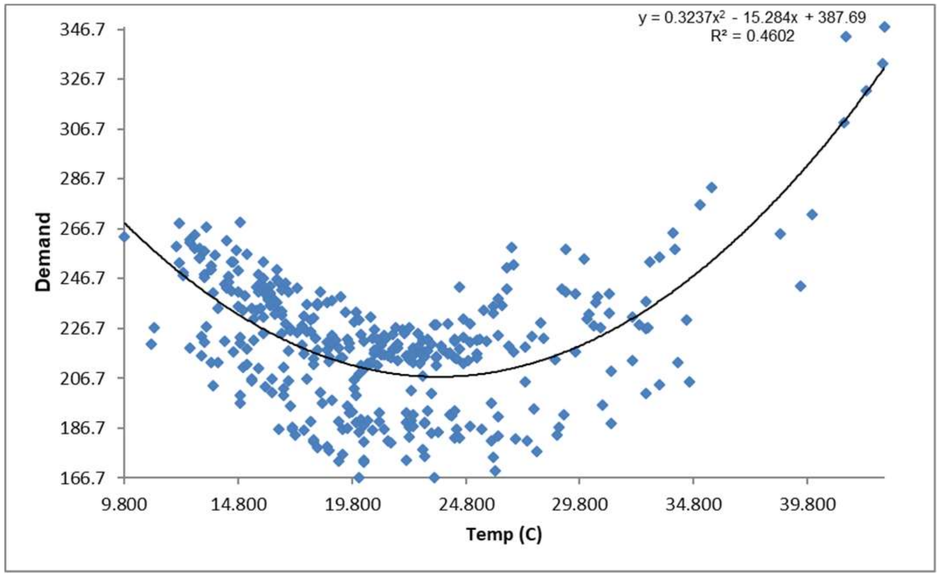


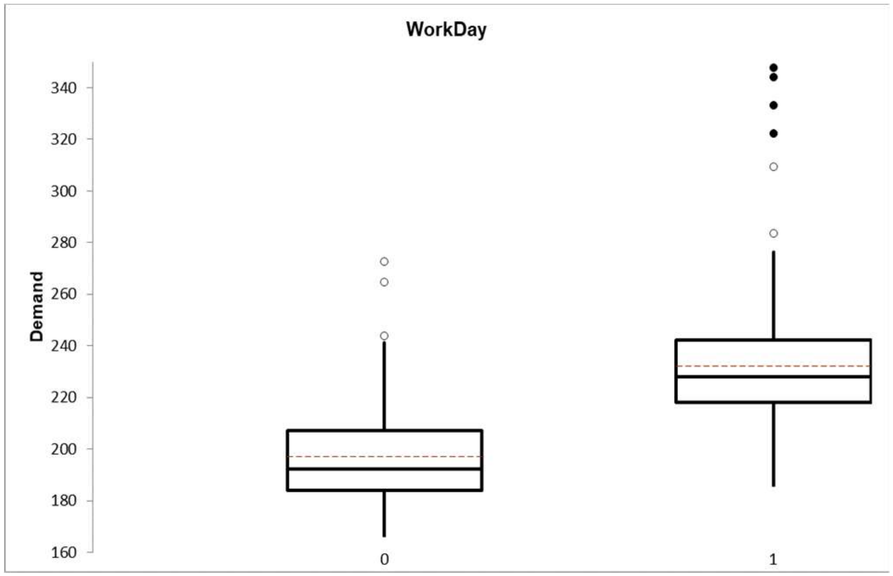


SigmaXL > Time Series Forecasting > Run Chart

Example 3: Daily Electricity Demand with Predictors – ElecDaily.xlsx
Victoria, Australia, 2014.

Example 3: Daily Electricity Demand with Temperature and Work Day Predictors – Scatterplot and Box Plot



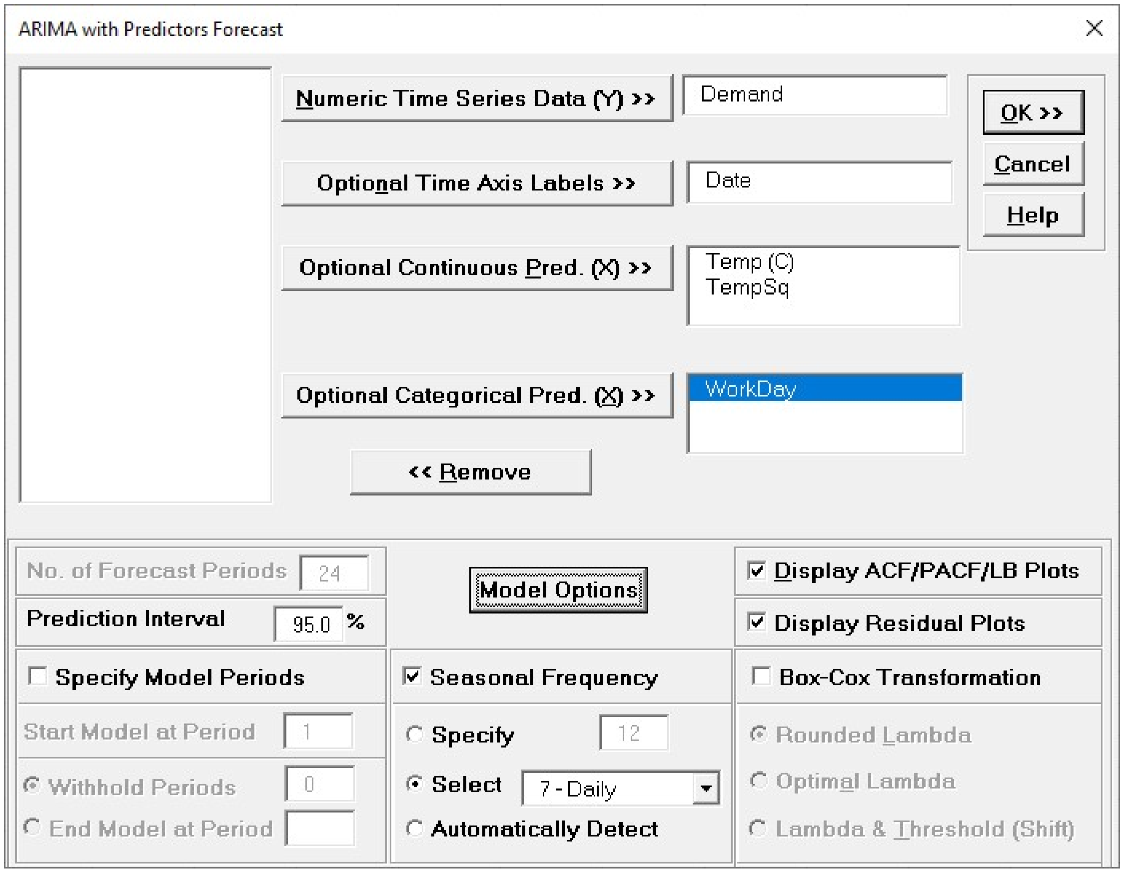


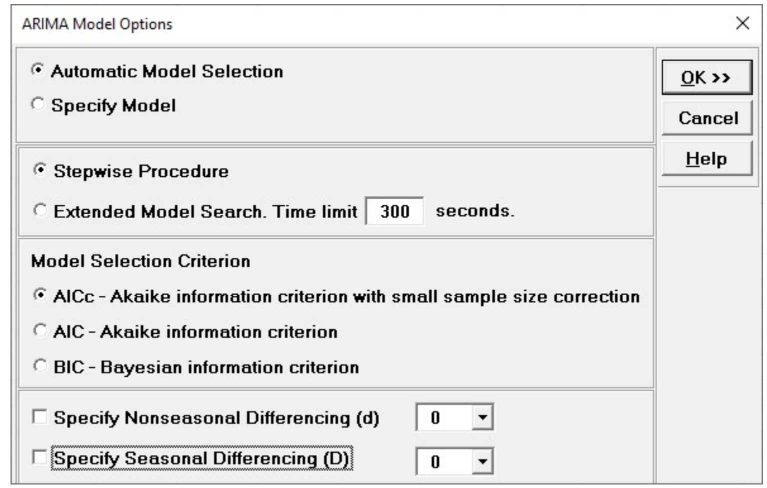
SigmaXL > Graphical Tools > Scatterplots

SigmaXL > Graphical Tools > Boxplots

Example 3: Daily Electricity Demand with Predictors – ElecDaily.xlsx
Victoria, Australia, 2014.

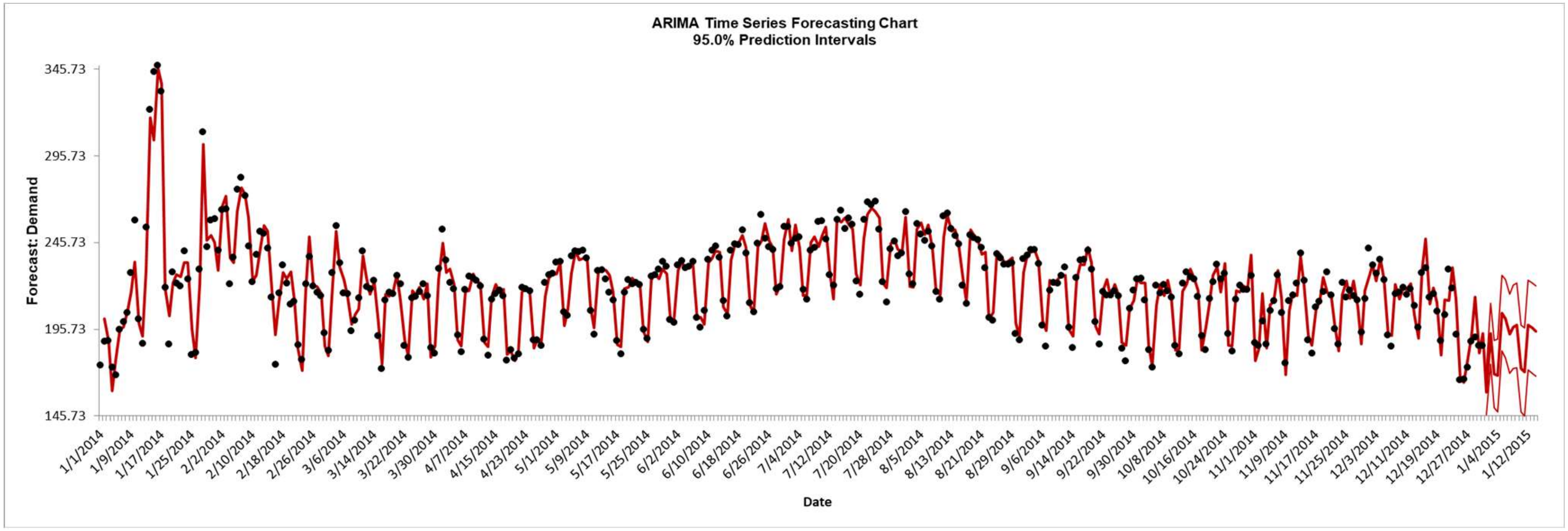
Example 3: Daily Electricity Demand with Temperature and Work Day Predictors – ARIMA Forecast with Predictors





SigmaXL > Time Series Forecasting > ARIMA Forecast > Forecast with Predictors

Example 3: Daily Electricity Demand with Temperature and Work Day Predictors – ARIMA Forecast with Predictors



ARIMA Model Summary	
AR Order (p)	2
I Order (d)	1
MA Order (q)	2
SAR Order (P)	2
SI Order (D)	0
SMA Order (Q)	0
Seasonal Frequency	7
Include Constant	0
No. of Predictors	3
Model Selection Criterion	AICc
Box-Cox Transformation	N/A
Lambda	
Threshold	

Parameter Estimates				
Term	Coefficient	SE Coefficient	T	P
AR_1	-0.063223451	0.075658448	0.83564	0.4039
AR_2	0.673128346	0.067270503	10.0063	0.0000
MA_1	0.022660844	0.043288704	0.52348	0.6010
MA_2	0.929862871	0.039474102	23.5563	0.0000
SAR_1	0.200902989	0.053912363	3.72647	0.0002
SAR_2	0.402632085	0.05676416	7.09307	0.0000
Temp (C)	-7.501559029	0.446098708	16.8159	0.0000
TempSq	0.17890261	0.008530253	20.9727	0.0000
WorkDay_1	30.56943168	1.295720007	23.5926	0.0000

What's New in SigmaXL® Version 9

Part 2 of 3: Time Series Forecasting

Questions?

References

1. Alwan, L.C., and Roberts, H.V. (1988), "Time Series Modeling for Statistical Process Control," **Journal of Business and Economic Statistics**, 6, 87-95.
2. Box, G. E. P. , Jenkins, G. M. , Reinsel, G. C. and Ljung, G.M. (2016). *Time Series Analysis, Forecasting and Control*, 5th edition, Wiley.
3. Hunter, J.S. (1986), "The Exponentially Weighted Moving Average," **Journal of Quality Technology**, 18, 203-210.
4. Hyndman, R.J., & Athanasopoulos,G. (2018). *Forecasting: principles and practice*, 2nd edition, OTexts: Melbourne, Australia. OTexts.com/fpp2.
5. Hyndman, R. J. and Y. Khandakar, (2008). "Automatic Time Series Forecasting: The forecast Package for R. " **Journal of Statistical Software**, 27(3), 1-22.

References

6. Montgomery, D. C., and Mastrengelo, C.M. (1991), "Some Statistical Process Control Methods for Autocorrelated Data," **Journal of Quality Technology**, 23, 179-204.
7. Montgomery, D. C. (2013). *Introduction to Statistical Quality Control*, 7th edition, Wiley.
8. Montgomery, D. C., Jennings, C.L., and Kulahci, M. (2015). *Introduction to Time Series Analysis and Forecasting*, 2nd edition, Wiley.
9. *NIST/SEMATECH e-Handbook of Statistical Methods*, https://www.itl.nist.gov/div898/handbook.
10. Woodall, W.H. and Faltin, F.W. "Autocorrelated Data and SPC" **ASQC Statistics Division Newsletter**, 13(4).

Lean Six Sigma Statistical Tools, Templates & Monte Carlo Simulation in Excel

What's New in SigmaXL® Version 9

Upcoming Webinars:

Part 3 of 3: Control Charts for Autocorrelated Data
Thursday, December 10, 2020 at 3 pm ET.

Or visit www.SigmaXL.com for recordings of webinars (typically available 2 weeks after scheduled webinar).