

Paper ST15  
**Basic Usage of SAS/ETS® Software to Forecast a Time Series**  
by LTC Douglas L. McAllaster  
US Army Logistics Management College, Fort Lee, Virginia

## ABSTRACT

This paper is a tutorial on using SAS/ETS to forecast a times series. We explain how PROC FORECAST generates forecasts using the three main exponential smoothing methods (single, Holt's, & Winters') and the stepwise auto-regressive method. Next, we explain usage of PROC REG (or better, PROC AUTOREG) to forecast a seasonal times series using indicator variables. Finally, we briefly overview of the capabilities and complications of using PROC ARIMA.

## INTRODUCTION

Forecasting a times series is the process of projecting a time series into the future. We can forecast two ways: either with extrapolation methods or with explanatory models. Extrapolation methods are adaptive techniques which simply project the series by responsively smoothing recently observed data. On the other hand, explanatory models attempt to fit the observed data to some mathematical model or underlying process.

## EDA

Exploratory data analysis (EDA) is the first and necessary step for proper time series analysis. We look at a times series plot of the data to visually identify the presence of three possible components: level, trend, and seasonality. First, the plot may be flat, showing only random fluctuation about some constant LEVEL. Second, the plot may show an upward (or downward) TREND over time. Finally, the plot may show some obvious SEASONALITY. The various time series plots in the appendix display these components.

## EXPONENTIAL SMOOTHING IDEAS

Exponential smoothing is the easiest and (thus?) the most widely used extrapolation method. The basic idea of this method is to project a series based mostly on recent data, i.e., without much regard to older data. We could call this a "short memory" method. This makes good common sense: we believe that sales next year depends much more on sales last year than on the sales from ten or even five years ago.

Another idea inherent in exponential smoothing is that a forecast for next period should be based on the accuracy of the most recent forecast. That is, if our most recent

forecast was too high, we adjust the next forecast downward. Likewise, if our most recent forecast was too low, we adjust the next forecast upwards. Thus, we describe smoothing methods as self-correcting, adaptive, and responsive; or, as we in the military like to say, a "fire & forget" weapon system. The appendix has plots showing the responsiveness of exponential smoothing techniques vice the fixed line created by using a regression method.

Furthermore, the short memory of exponential smoothing makes it a particularly appropriate method when the time series components (level, trend, & seasonality) are changing over time. Finally, the adaptive feature of smoothing methods is especially useful since an analyst often must forecast many dozens or hundreds of stock keeping units (SKUs) and does not want to manually determine changes in direction or tailor a separate model for each SKU. Of course, the SAS System has BY variable processing which makes for very simple coding, indeed.

## PROC FORECAST

There are three main exponential smoothing methods: single, Holt's, & Winters'. We choose the appropriate method according to the components we identified during the EDA phase. That is, we use single exponential smoothing for a series which has only the level component, Holt's exponential smoothing method for a series which also has a trend component, and Winters' exponential smoothing method for a series which also has a seasonal component.

## SES

Single exponential smoothing is (SES) the simplest of the three main exponential smoothing methods, so we begin our explanation here. A perusal of the SES equation in the appendix shows the above mentioned self-corrective nature of the method. (The second equation is more convenient for spreadsheet implementation.) By convention,  $Y_t$  is the series value and  $F_t$  is a forecast value. The alpha factor smooths the level component of the series, i.e, smooths out the past forecast errors. We call alpha a "smoothing or dampening factor" rather than a parameter since, in this context, we are not fitting an explanatory model to the data. Rather, we are simply extrapolating the series, not trying to determine a formal

statistical parameter, like a beta in the familiar regression model.

## SMOOTHING FACTOR

The analyst must specify a smoothing factor for this method. This factor controls how much relative weight the procedure gives to newer versus older data. Valid factors are between zero and unity. A factor near unity gives the vast majority of the weight to the most recent observations whereas a factor nearer zero allows the older observations to influence the forecast. Typical values are between one-tenth and one-third.

To determine a "good or best" smoothing factor, most texts recommend analysts use a trial & error approach to minimize some measure of inaccuracy, either absolute or squared error. This is fairly easy with a well designed spreadsheet which permits dynamic adjustments of the smoothing factor to see the impact the measures of accuracy. Currently, neither Microsoft® Excel nor PROC FORECAST automatically optimizes this factor to minimize squared error. Without an optimization subroutine, texts recommend that alpha should be no more than 0.4, otherwise the analyst is probably using SES on an inappropriate series, i.e., on a series which has more than the level component only. PROC FORECAST provides a default smoothing factor if the analyst omits it.

Some software packages automatically optimize the smoothing factor, e.g., Minitab and some third party MS Excel add-ins. A future release of SAS/ETS software will include a PROC TREND which will automatically optimize the smoothing factor. PROC TREND is experimental in SAS version 8.1 (& omitted from the on-line documentation). My presentation will include sample PROC TREND code if SI will provide me some preliminary documentation.

## HOLT'S METHOD

Holt extended (Brown's original) SES idea to smooth the trend component as well as the level component. That is, Holt's method is the proper one when the series has both the level & trend components. Thus, Holt's method has two dampening factors, often called alpha & beta. Alpha smooths the level component and beta smooths the trend component. Again, the equations are in the appendix.

The presentation of this paper will show MS Excel implementations of both SES and Holt's methods. These tables are helpful to see the time series component parts and to understand what the smoothing equations are doing.

Unfortunately, there are no rules of thumb for "good" smoothing factors in Holt's method. Nevertheless, some statistical software packages optimize these factors.

## WINTERS' METHOD

Winters' extended Holt's method to smooth the third and final component, seasonality. The equations (see the appendix) look a little scary and a spreadsheet version of this method is a bit cumbersome, so we omit it from the presentation. However, once an analyst understands the smoothing concept in Holt's, he should feel confident in using Winters' method, i.e., without the "black box" syndrome.

The appendix has documented examples of the code which produced the exponential smoothing forecasts (SES, Holt's, & Winters') shown. The options are a bit tricky to get Holt's method: the analyst specifies METHOD=Winters BUT provides NO third (seasonal) smoothing factor NOR includes the SEASONS=month option.

## INITIALIZATION

All of these exponential smoothing procedures require initialization. Typically, statistical software packages perform a simple linear regression against the first few observations to produce an initial forecast of the level and trend components. Thereafter, the exponential smoothing procedure adjusts future forecasts based on the accuracy of the recent forecasts. Fortunately, forecasting is not sensitive to the initialization method as long as there are ten or twenty observations, since exponential smoothing has short memory.

SAS provides various initialization options. The code in the appendix bases the initial forecast on the first one (SES), two (Holt's), or season (Winters') of data by using macro variables to pass initial values derived from a data step into the PROC FORECAST step. I use this initialization in teaching since it is easiest to implement in a spreadsheet and since exponential smoothing is robust with respect to initialization.

## WHY NOT PROC REG?

Neophyte analysts (and many experienced practitioners!) would probably apply ordinary least squares (OLS) regression to forecast a time series since every analyst learns (ok, studies) it. Actually, their results would be quite good, given a reasonable history of data and a series in which the level & trend components which are not changing drastically. The plots in the appendix show the obvious faults of using OLS regression on series whose components change significantly or on a series having seasonality. Again, a strength of exponential smoothing is

its adaptability to changing components, i.e., its short memory.

Of course, the analyst could make PROC REG responsive to recent data by generating and applying increasingly large weights to recent observations. In fact, below we will conclude that regression is indeed a good method, even despite the technical violation of the "independence of errors" assumption.

## PROC REG & SEASONALITY

Now, many analysts are also familiar with using indicator (binary) predictor variables in OLS regression. Analysts might prefer to bypass Winters' method in favor of this seasonal-indicator-variable OLS method as a very useful alternative to model seasonal data. We proceed as follows. We simply introduce eleven (seasonal) indicator variables to identify the month along with the time variable to capture the trend component. The only real trick is to first transform the original series by taking the natural log of Y in order to dampen out the typically increasing variance as time increases and then to exponentiate the forecasts to restore the original series. The appendix has a chart showing that OLS regression fits separate and parallel straight lines for each of the twelve months. (The 12th month is the one which has all eleven indicators variables equal to zero.) The appendix also has charts showing the pitfall of performing the regression without the log transformation.

## AUTO-CORRELATION

Time series data often exhibits something called auto-correlation. This means that observations are NOT independent, one from another. Now, if the obs are not independent, then neither will the residuals from OLS regression be independent. This dependency is a (technical) violation of the normal, independent, identically distributed (NII) assumption about residuals required by the OLS regression model.

There are two types of auto-correlation: positive and negative. Positive auto-correlation is more common in real data. This means that errors tend to stay positive for a while then switch negative and stay negative for a while and so on back and forth. Negative auto-correlation is not so easy to see on a plot and occurs when errors bounce back and forth between positive and negative too often. The Durbin-Watson statistic is a widely used but limited detector for auto-correlation. (PROC REG will calculate it upon request, by the DW option on the model statement.) The Durbin-Watson statistic is limited in that it detects auto-correlation at lag 1 only. Nevertheless, lag 1 correlation is the most common pattern in real data, so DW is a very useful diagnostic.

## PROC AUTOREG

The statistically correct method for handling auto-correlated data is to use the auto-regressive method. This method accounts for the dependency with an additional parameter to model the auto-correlation. SAS/ETS software includes PROC AUTOREG to fit an auto-regressive model.

Now, although times series often violate the independence assumption, the forecasts using PROC REG would be about the same as if one used PROC AUTOREG, as long as there is a reasonable history of data. The only disadvantage is that forecasts would not be as precise, i.e., confidence intervals would be wider, since OLS regression does not model the errors as well as auto-regression. The SAS/ETS User's Guide chapter on PROC AUTOREG shows this via an informative example. Using PROC AUTOREG is very similar to using PROC REG, so many SAS data analysts will find moving to PROC AUTOREG an easy step. The only really new syntax is the NLAG=# option. For example, if we specify NLAG=5, PROC AUTOREG will fit five auto-regressive parameters. Without the NLAG option PROC AUTOREG performs OLS regression.

## FORECAST VS AUTOREG

Now, PROC FORECAST can also perform auto-regression. In fact, stepwise auto-regression is its default method. However, since most neophyte forecasters may not have studied auto-regression, this default behavior is probably not where the novice should begin. Thus, perhaps the main benefit of this paper is that it presents forecasting in the typical learning sequence so the analyst can use the METHOD=stepar option without fear, as follows.

The METHOD=stepar option calculates good forecasts with minimal user input. That is, if the analyst understands that the procedure is simply a regression on time with additional parameter(s) to model correlation of residuals at various lags, then he should feel comfortable using the stepwise auto-regression results. Perhaps the only confusing thing is that the SEASONS=month options is not valid with METHOD=stepar since this method uses only time as the only predictor variable to detrend the data and models all remaining variation with auto-regressive (error correlation) parameter(s).

Syntax note. Although the procedure will fit some number of lags by default, correct BY variable processing requires specifying it via an NLAGS=# option. Thus, I recommend always specifying the number of correlations to fit via the NLAGS=# option. Again, I suggest simply specifying one more lag than the seasonal period,  $12+1=13$  for monthly data and  $4+1=5$  for quarterly data.

## ARIMA METHODS

Finally, Box and Jenkins popularized Auto-Regressive--Integrated--Moving-Average (ARIMA) methods. SAS/ETS software provides PROC ARIMA for the analyst to model data using these methods. However, proper usage of ARIMA methods on time series data requires substantially more user knowledge and input. ARIMA methods recommend a three step process: identification, estimation, and forecasting. The analyst must be familiar with prototype auto-correlation and partial auto-correlation plots. He must understand auto-regression models. He must understand moving average models. He must understand stationarity, how to achieve it with differencing, and how to confirm it with portmanteau tests (Ljung-Box statistics).

As a "quick & dirty" practitioner, I have my doubts whether this more complicated method is worth the additional effort. Sometimes an about right, robust method is better than an exactly right, rigorous model. The benefits of Keeping It Simple, Sir (KISS) often outweigh the rectitude of statistical elegance. However, for those desiring purity, the Bowerman and Makridakis texts provide good applied approaches to using these methods. The Bowerman text even provides PROC ARIMA and Minitab code to implement the methods. Furthermore, (and atypically for most software manuals) the SAS Applications Guide 1 is a very good overall reference for ARIMA methods, both its basic theory and its implementation in SAS.

## REFERENCES:

*Forecasting and Times Series: An Applied Approach*, 3rd edition, Bruce Bowerman and Richard O'Connell, Brooks/Cole 1993. (reissued in Duxbury Classic Series softcover in 2000)

*Forecasting: Methods and Applications*, 3rd edition, Makridakis, Wheelwright, & Hyndman, John Wiley & Sons, 1998.

*Applied Linear Statistical Models*, 4th edition, Neter, Kutner, Natchsheim, & Wasserman, McGraw-Hill, 1996.

SAS Institute Inc., *SAS/ETS Software: Applications Guide 1, Version 6, First Edition: Time Series Modeling and Forecasting, Financial Reporting, and Loan Analysis*, Cary, NC: SAS Institute Inc. 1991. 380pp.

SAS Institute Inc., *SAS/ETS User's Guide, Version 6, Second Edition*, Cary, NC: SAS Institute Inc. 1993. 1,022pp.

## TRADEMARK

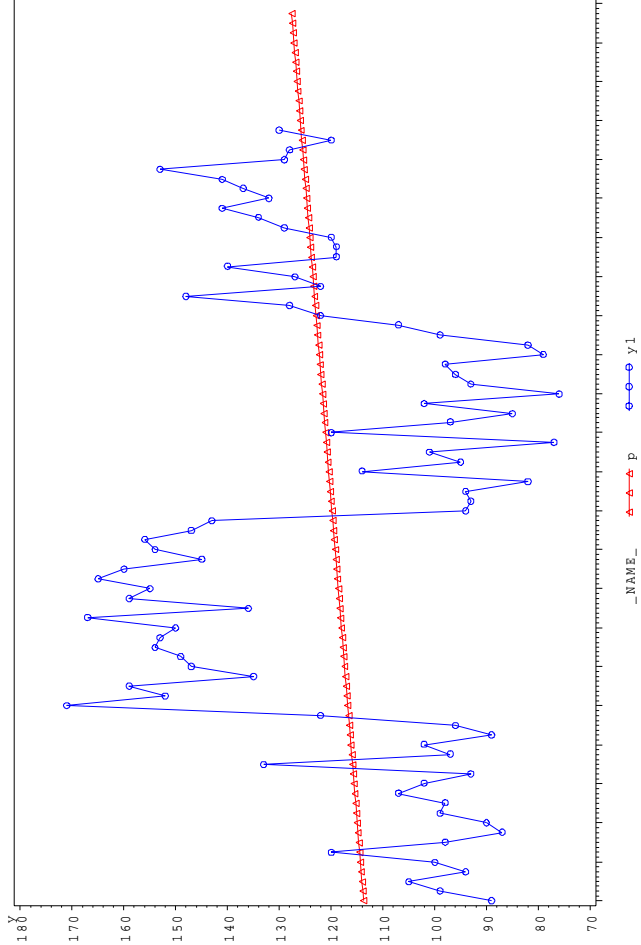
SAS is a registered trademark or trademark of SAS Institute Inc. in the USA and other countries. ® indicates USA registration.

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

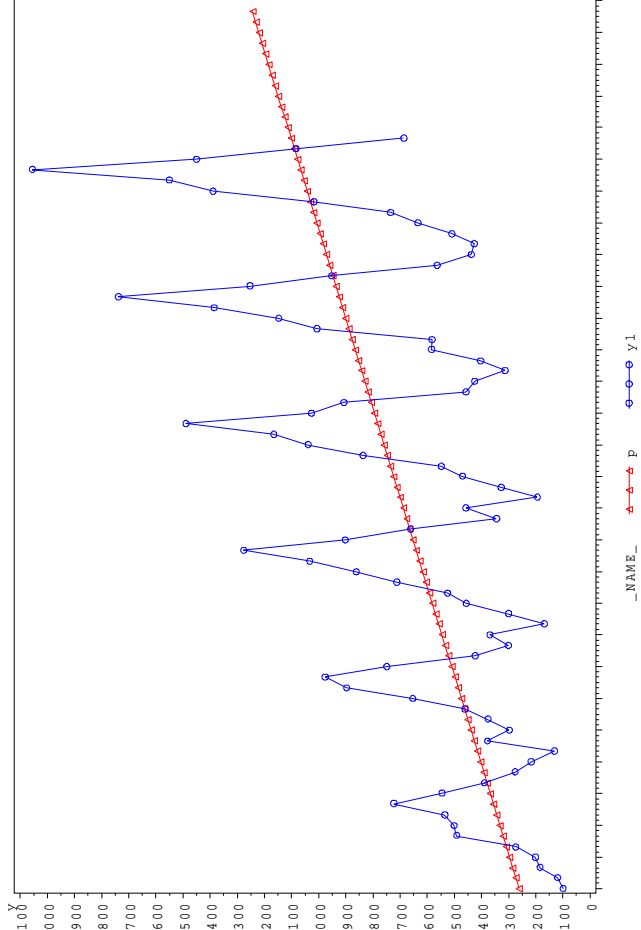
## AUTHOR CONTACT

You may contact me at: [McAllasterD@Lee.Army.Mil](mailto:McAllasterD@Lee.Army.Mil)

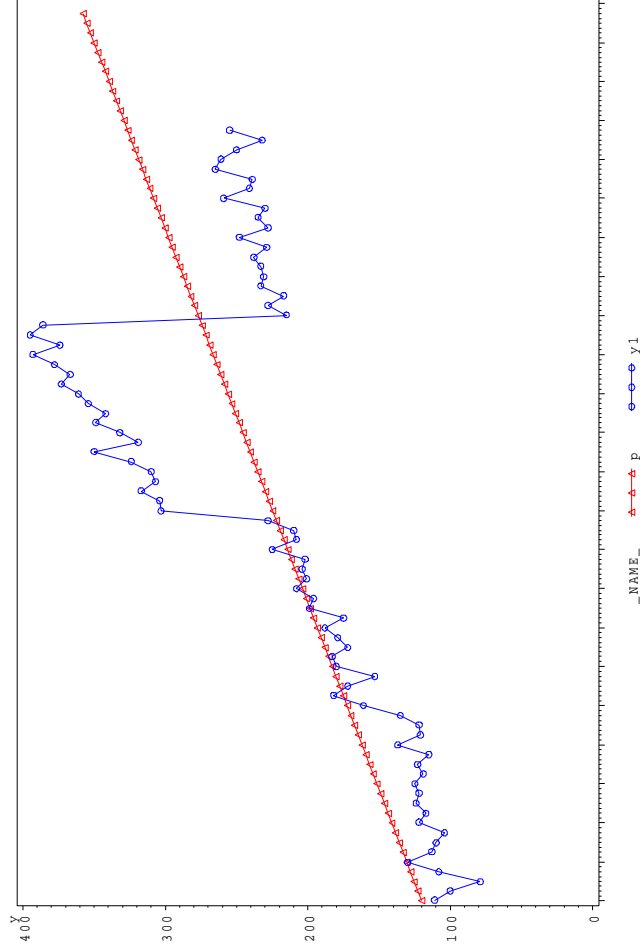
OLS Reg forecasting



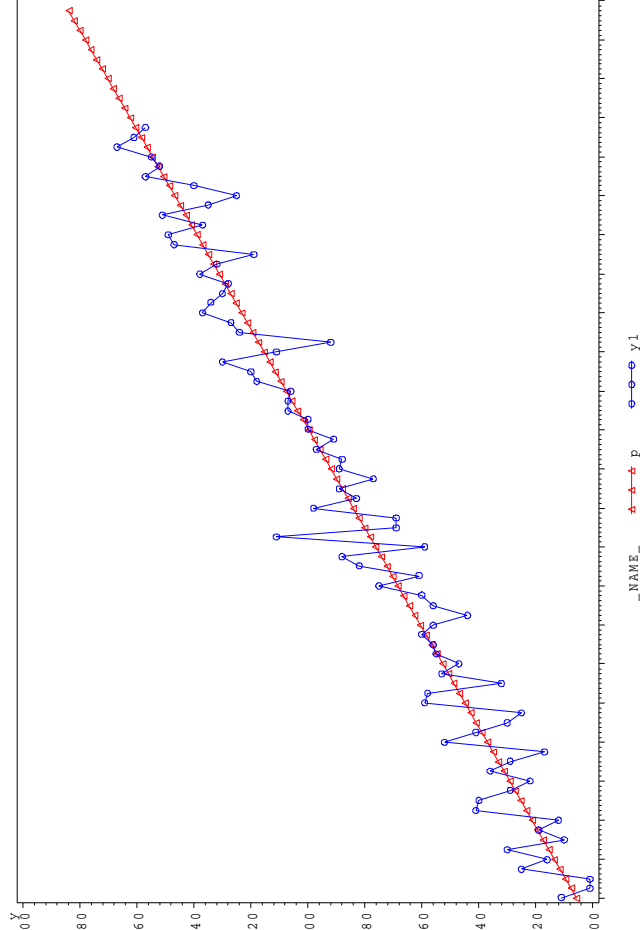
OLS Reg forecasting



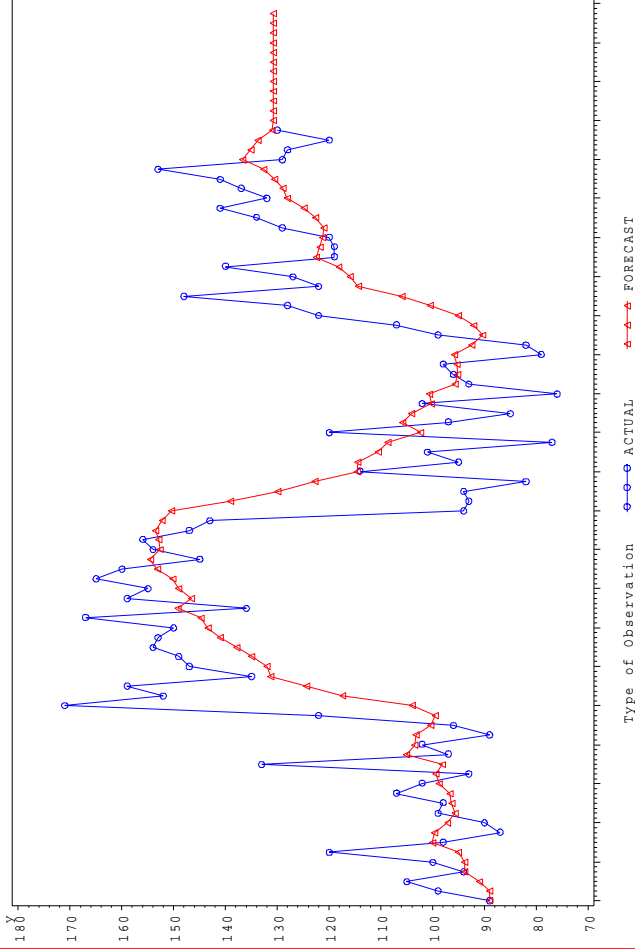
OLS Reg forecasting



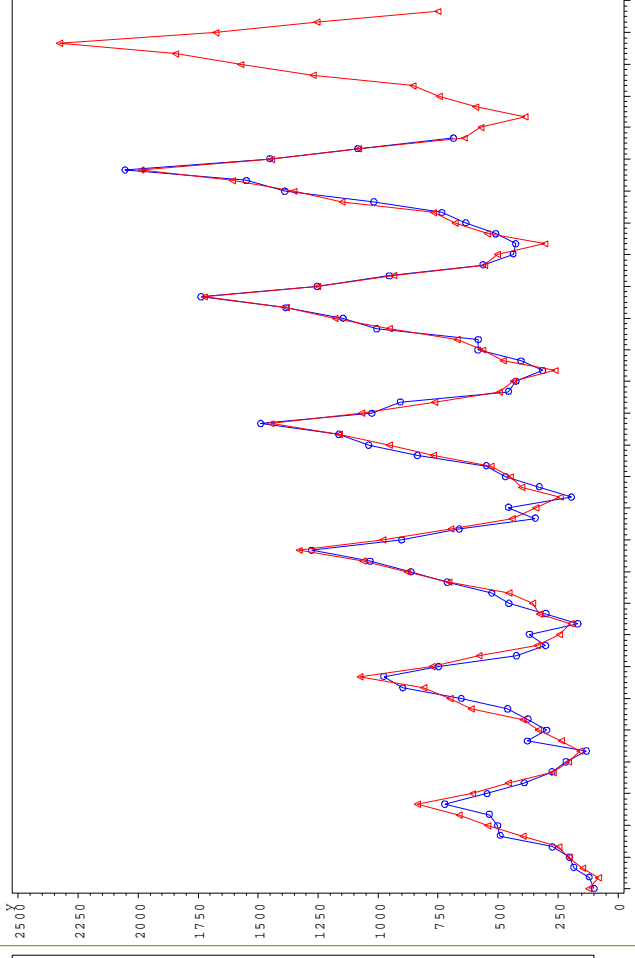
OLS Reg forecasting



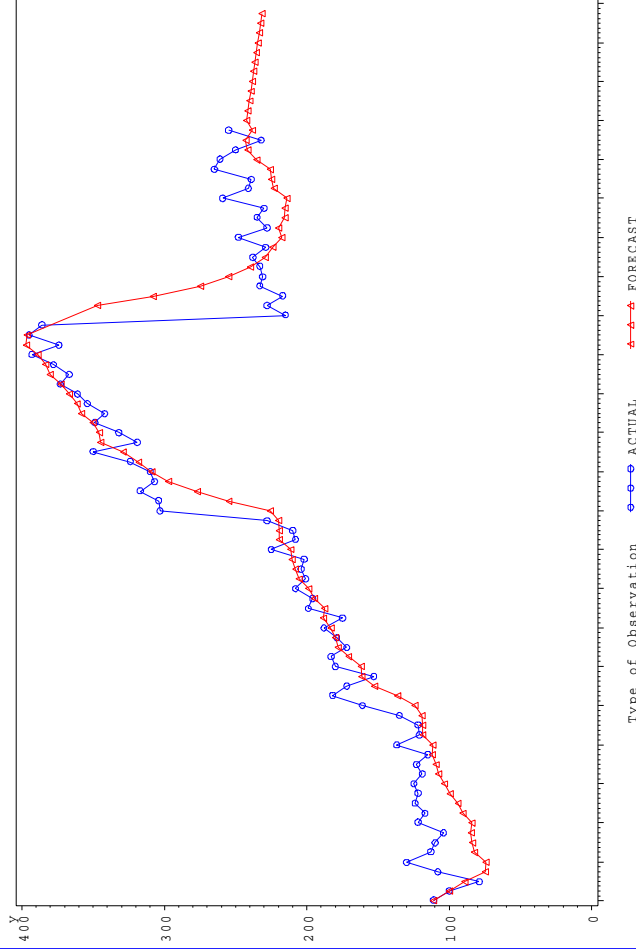
SES forecasts a LEVEL time series



Winters forecasts a SEASONAL time series



Holt forecasts a TREND time series



exponential smoothing

the basic ideas is:

give most weight to most recent obs

give less weight to older obs

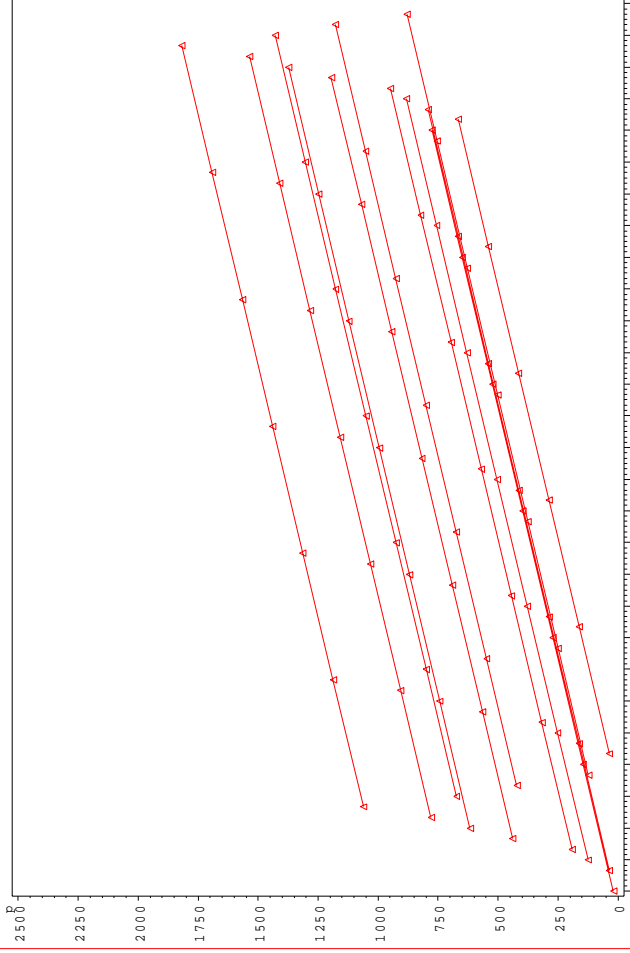
the alpha, beta, & gamma parameters

control the amount of weight given to the recent obs

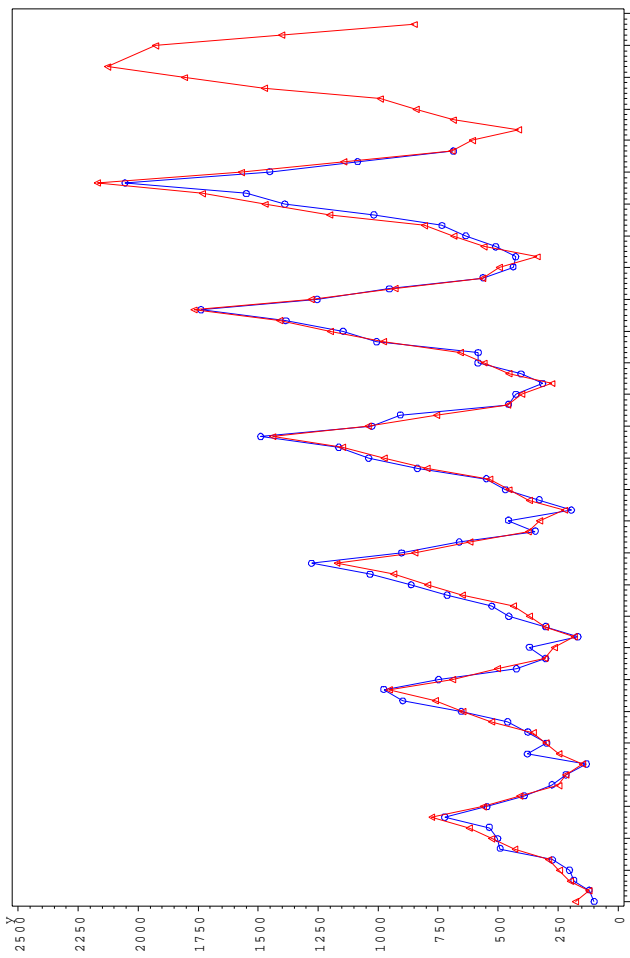
factors near unity give nearly all wt to recent obs

see other page for actual equations

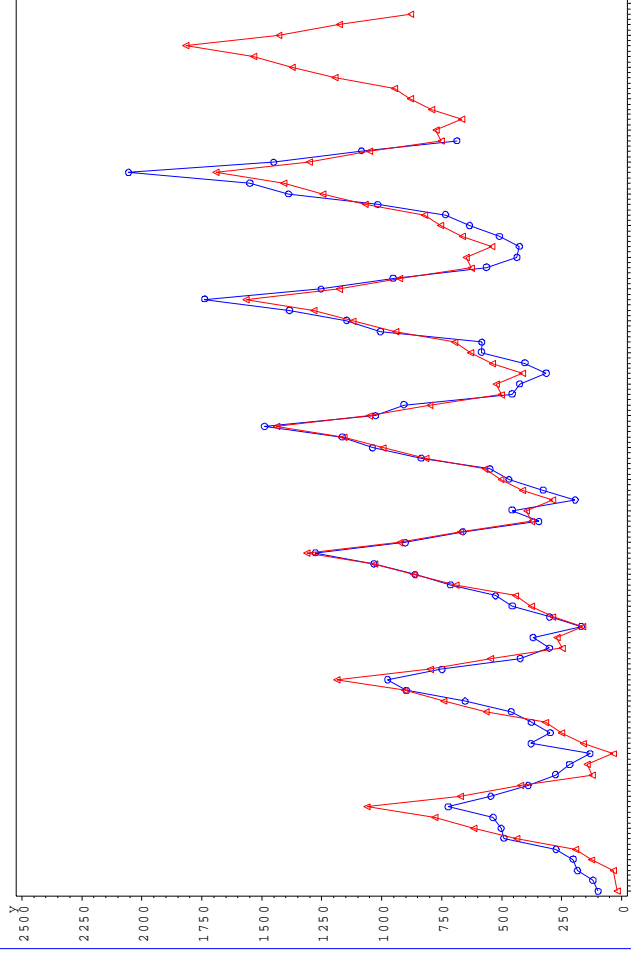
PROC REG fits parallel lines



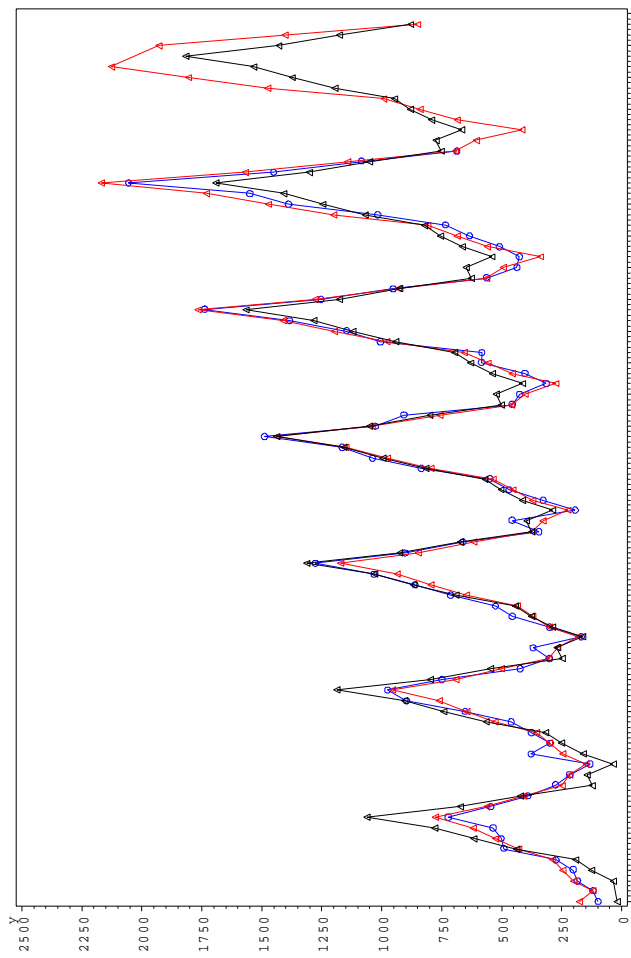
PROC REG on Log(y) w/ time & indicator vars to model seasonality



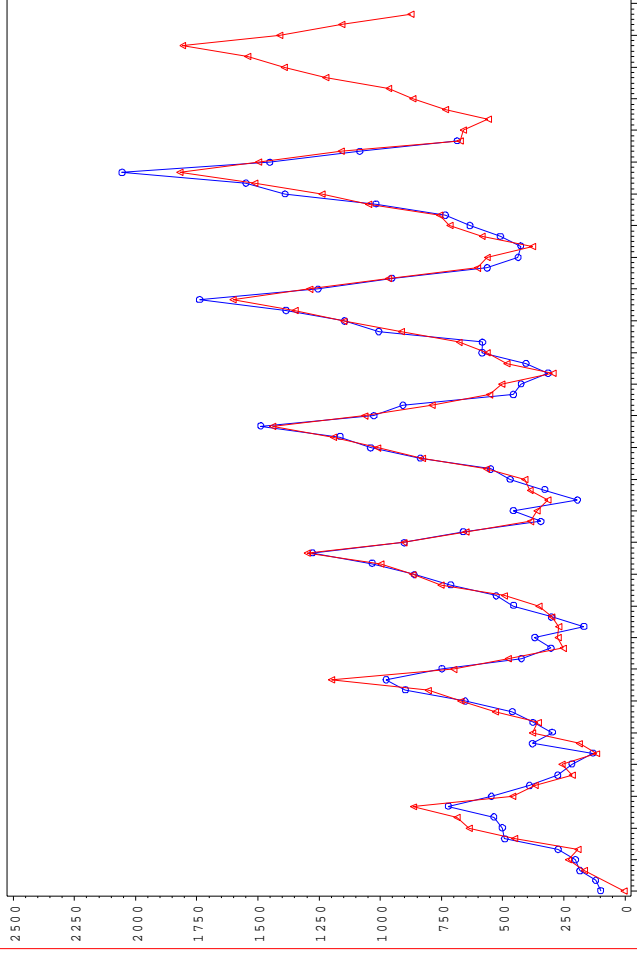
PROC REG w/ indicator vars to model seasonality



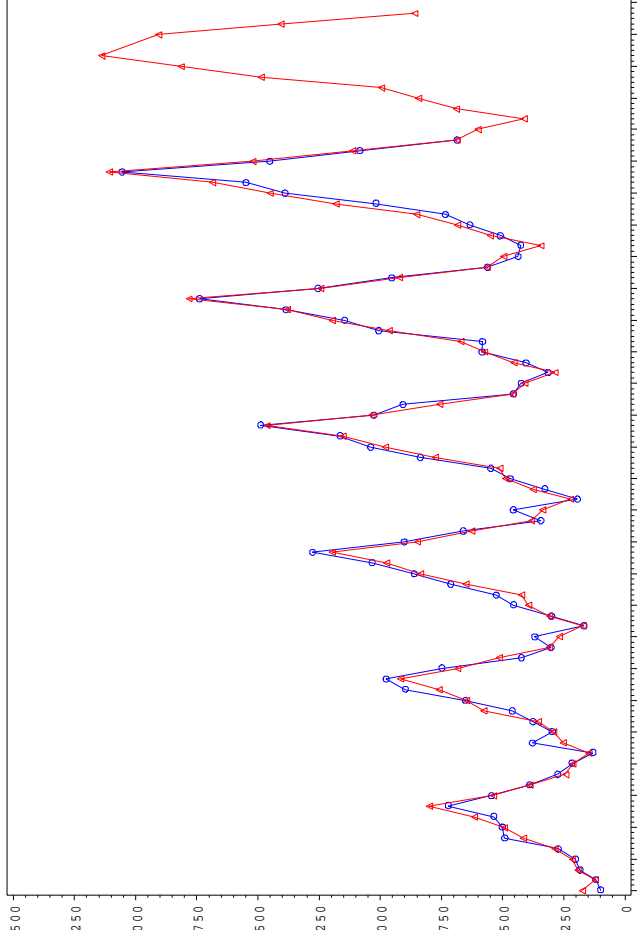
comparing regressions on ln\_y vice y on a seasonal time series



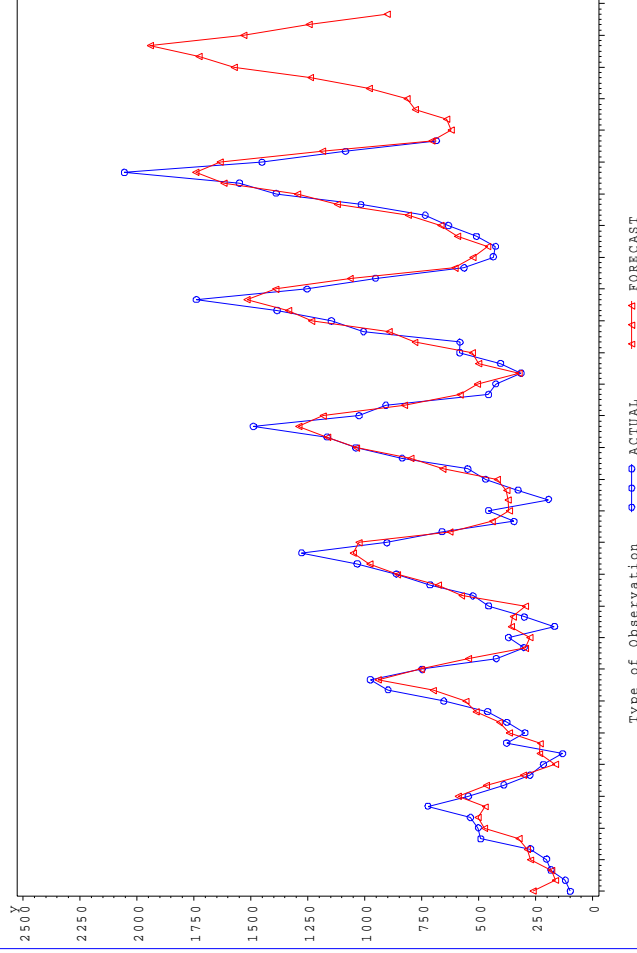
PROC AUTOREG w/ time & indicator vars & AR parms to model seasonality



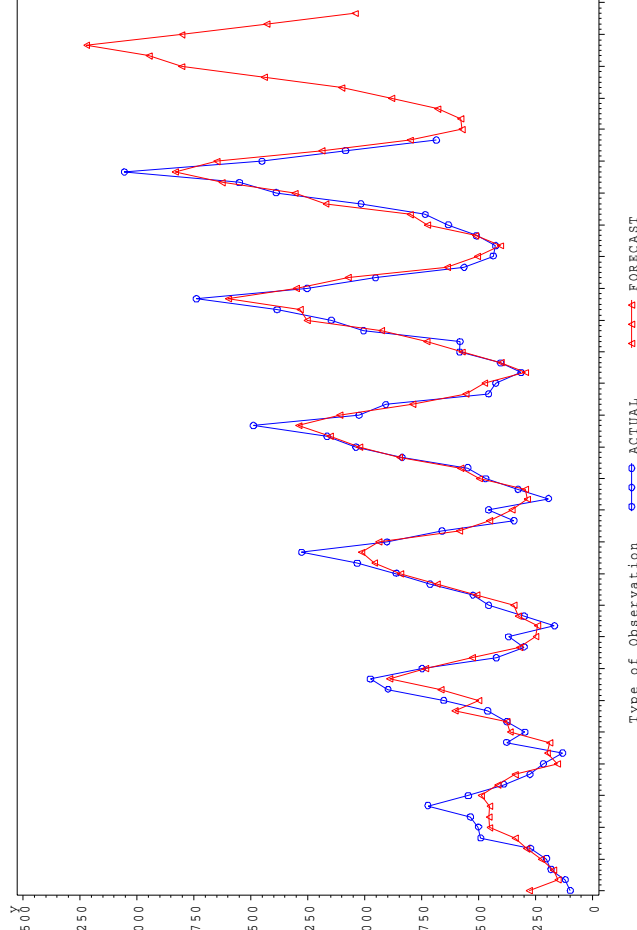
AUTOREG on LN(Y) w/ time, ind vars, & AR parms to model seasonality



PROC FORECAST method = STEPAR (default)



PROC FORECAST method = STEPAR on log(y) is much better



```

proc forecast data=table
  method=Winters
  trend=2
  interval=month
  weight=(0.20 0.10)
  lead=12
  astart=&i;level
  bstart=&i;trend
  out=out1 outfull
  outest=est1;
  /* seasons=month <--- omitting this to get Holt's method
  nstart=max <--- uses all obs to get astart & bstart */
  id date;
  var y;
  by tableid; run;

```

## Holt

```

proc forecast data=table
  method=Winters
  trend=2
  interval=month
  weight=(0.20 0.10 0.05)
  lead=12
  astart=&k;ast
  bstart=&k;bst
  out=out1 outfull
  outest=est1
  seasons=month /* including seasons=months thus gives Winters' */
  nsstart=6; /* can give # seasons for seasonal factors */
  id date;
  var y;
  by tableid; run;

```

## Winters

```

proc forecast data=table
  method=expo
  trend=1
  interval=month
  weight=(0.20)
  lead=12
  astart=&i;firstobs
  /* bstart=
  out=out1 outfull
  outest=est1;
  id date;
  var y;
  by tableid; run;

```

## SES

```

proc forecast data=table
  method=stepar
  trend=2
  interval=month
  weight=
  lead=12
  astart=&i;level
  bstart=&i;trend
  out=out1 outfull
  outest=est1
  /* seasons=month is not valid for method=stepar */
  nlags=14; /* only stepar uses nlags option */
  id date;
  var y;
  by tableid; run;

```

## STEPAR

Single exponential smoothing (SES)

$$F_{t+1} = F_t + \alpha(Y_t - F_t)$$

$$F_{t+1} = \alpha Y_t + (1 - \alpha)F_t$$

$$0 \leq \alpha \leq 1$$

Holt's Linear Method

$$L_t = \alpha Y_t + (1 - \alpha)(L_{t-1} + b_{t-1})$$

$$b_t = \beta(L_t - L_{t-1}) + (1 - \beta)b_{t-1}$$

$$F_{t+m} = L_t + b_t m$$

$$0 \leq \alpha \leq 1 \quad 0 \leq \beta \leq 1$$

Winters' Trend and (Multiplicative) Seasonality Method

$$L_t = \alpha \frac{Y_t}{S_{t-s}} + (1 - \alpha)(L_{t-1} + b_{t-1})$$

$$b_t = \beta(L_t - L_{t-1}) + (1 - \beta)b_{t-1}$$

$$S_t = \gamma \frac{Y_t}{L_t} + (1 - \gamma)S_{t-s+m}$$

$$F_{t+m} = (L_t + b_t m)S_{t-s+m}$$

$$0 \leq \alpha \leq 1 \quad 0 \leq \beta \leq 1 \quad 0 \leq \gamma \leq 1$$

1st Order Auto Regression Model, AR(1)

$$Y_t = \beta_0 + \beta_1 X_{t1} + \epsilon_t$$

$$\epsilon_t = \rho \epsilon_{t-1} + \mu_t$$

$$-1 < \rho < 1$$

$$\mu_t \sim N(0, \sigma^2) \quad \text{and independent}$$