Chapter 3 Exponential Smoothing Models

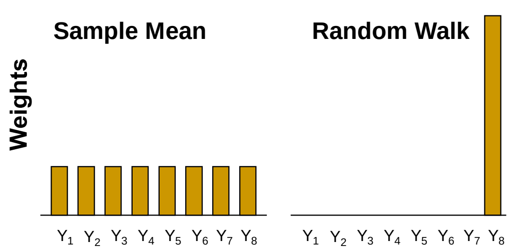
3.1 Exponential Smoothing Models

Objectives

- Explore weighted average models and exponential smoothing.
- Compare and contrast simple mean, random walk, and exponential smoothing models.

2

Weighted Average Examples



Weights applied to past values to predict Y_9

3

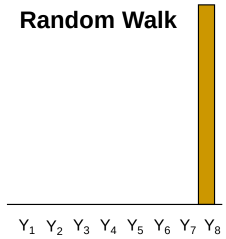
Weighted averaging is a simple and intuitive method for smoothing a time series and forecasting future values from your past observations. In weighted averaging, weights are applied to past values in such a way that they predict the value of the next time point in the series. The values of these weights are selected based on the determination of the importance of these past observations in determining the future. Look at two, very commonly used styles of weighted averages: sample mean and random walk.

Weighted Average Example: Random Walk

$$\hat{Y}_{n+1} = \sum_{t=1}^{n} w_t Y_t = Y_n$$

$$w_n = 1, w_t = 0 \text{ for } t = 1, 2, \ldots, n-1$$

Random Walk



$$\hat{Y}_9 = Y_8$$

A random walk forecast is a weighted average where all weights are 0 except the most recent, which is 1.

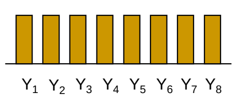
4

A common example of a weighted average is the random walk. In this setup, all weights for previous observations are 0 except for the most recent, which receives a weight of 1. In the random walk, where you will be at the next time point is related only to where you are immediately prior.

Weighted Average Example: Simple Moving Average

Sample Mean

Weights



$$\hat{Y}_9 = \frac{1}{8}\sum_{t=1}^{8} Y_t$$

$$\hat{Y}_{n+1} = \sum_{t=1}^{n} w_t Y_t = w_1 Y_1 + w_2 Y_2 + \cdots + w_n Y_n$$

$$= \sum_{t=1}^{n} \frac{1}{n} Y_t = \frac{1}{n}\sum_{t=1}^{n} Y_t = \bar{Y}$$

$$w_t = \frac{1}{n}$$

The mean is a weighted average where all weights are the same.

5

In the sample mean example, equal weight is placed on each of the previous n observations in the calculation of the prediction of the next series value. In this setup, the weight of $(1/n)$ is applied to each of the previous n observations and zero to all remaining prior observations.

Simple Moving Average

Disadvantages

- cannot be used on the first *n*-1 terms of the time series without adding other terms by some other means
- can be influenced by extreme values within the window
- requires the retaining of the most recent *n* observations to produce forecasted value

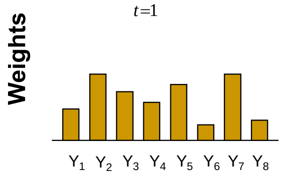
6

Despite the simplicity of the simple moving average, there are disadvantages to its usage. With the requirement of the *n* terms to produce the weighted average, this cannot be applied to the first (*n*-1) terms of the time series unless extrapolation beyond the data is applied. Another disadvantage is that a simple moving average shares the same issues with extreme values as a typical mean. Just as the mean is not robust to outliers, extreme values can make their presence known within a simple moving average. Finally, in the production of forecasts that use a simple moving average, the most recent *n* observations must be retained.

Weighted Average Example: Weighted Moving Average

$$\hat{Y}_{n+1} = \sum_{t=1}^{n} w_t Y_t = w_1 Y_1 + w_2 Y_2 + \cdots + w_n Y_n$$

$$\sum_{t=1}^{n} w_t = 1$$



The mean is a weighted average where not all weights are the same.

7

When you use a weighted moving average, it is not a requirement that the weights of the prior *n* observations be equal. In this example, the weights can vary across the previous observations in the creation of the forecasted value.

More about Weighted Moving Average

- More weight is given to the most recent terms in the time series and less to the older terms.
- Like the simple moving average, a weighted moving average cannot be used until at least *n* observations are made.
- Several methods for the handling of missing data exist.
- A weighted moving average requires the retaining of the most recent *n* observations to produce a forecasted value.

8

In many weighted moving average setups, where the weights are not equal, more weight is given to the more recent values in the series and less to the older observations. The premise is that the forecasted value is more like the observations immediately prior to it and less like those farther in the past. Even including this alteration to the weights, this setup cannot be used until the necessary *n* observations occur in the series, unless extrapolation is allowed.

If data is missing from a series, there are several methods to accommodate the data. These range from using the overall mean of the series to the one-step-ahead forecasting that is typically used in exponential smoothing (Yaffee and McGee 2000).

Like the simple moving average, the weight moving average requires the retaining of the most recent *n* observations to generate the forecasts.

Exponential Smoothing Models: Premise

- Weighted averages of past values can produce good forecasts of the future.
- The weights should emphasize the most recent data.
- Forecasting should require only a few parameters.
- Forecast equations should be simple and easy to implement.

9

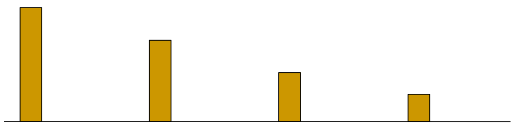
An exponential smoothing model (ESM) was first suggested by Robert Goodell Brown (1959, 1962) and added to by Charles C. Holt (1960). This model is used primarily for the creation of forecasting models for inventory control systems. ESMs formulate forecasts using a “smoothing” method of weighted averages. In the construction of the forecasts, more recent observations are given more weight than observations in the more distant past. The “exponential” is derived from the fact that weights not only diminish over time, but they do so exponentially (Fomby 2008). ESMs have the added bonus that only a few parameters are required in the forecasting model and these equations are simple to implement.

The seven common exponential smoothing models are supported by the ESM procedure. In addition, a few exponential smoothing models are not as common and are not supported. For example, triple exponential smoothing models use third differencing, that is, three differences. Such models are addressed in textbooks, but rarely provide good forecasts for real data. When they do offer acceptable results, the application is usually highly specialized.

The Exponential Smoothing Coefficient

Forecast Equation

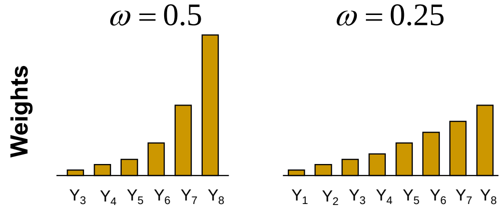
$$
\begin{aligned}
\hat{Y}_{t+1} &= \omega Y_t + (1-\omega)\hat{Y}_t \\
&= \omega Y_t + (1-\omega)[\omega Y_{t-1} + (1-\omega)\hat{Y}_{t-1}] \\
&= \omega Y_t + \omega(1-\omega)Y_{t-1} + (1-\omega)^2 \hat{Y}_{t-1} \\
&= \omega Y_t + \omega(1-\omega)Y_{t-1} + (1-\omega)^2[\omega Y_{t-2} + (1-\omega)\hat{Y}_{t-2}] \\
&= \omega Y_t + \omega(1-\omega)Y_{t-1} + \omega(1-\omega)^2 Y_{t-2} + \omega(1-\omega)^3 Y_{t-3} + \cdots
\end{aligned}
$$



10

Simple Exponential Smoothing

$\omega = 0.5$ $\omega = 0.25$

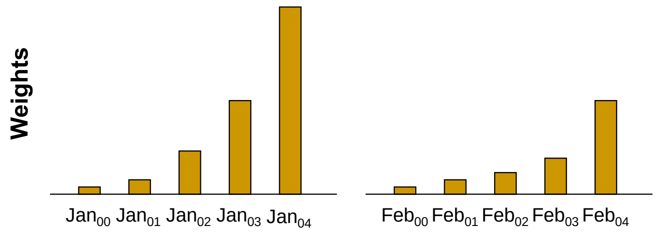


Weights applied to past values to predict Y_9

As the parameter increases, the emphasis on the most recent values increases.

11

Exponential Smoothing for Seasonal Data



Weights decay with respect to the seasonal factor.

12

To obtain the exponential decay, the absolute values of the smoothing parameter must be less than one ($|\omega| < 1$). The SAS/ETS documentation gives an explanation for weights near zero or one.

Exponential Smoothing Models (ESM)

- Models for time series with trend:
  - simple exponential smoothing
  - double (Brown) exponential smoothing
  - linear (Holt) exponential smoothing
  - damped-trend exponential smoothing
- Models for time series with seasonality:
  - seasonal exponential smoothing
- Models for time series with trend and seasonality:
  - Winters additive exponential smoothing
  - Winters multiplicative exponential smoothing

13

Because there are exactly seven models, a trial-and-error method becomes an effective strategy for model selection in the age of high-speed computers.

Simple exponential smoothing should be used when the time series data has no trend and no seasonality. The ARIMA model that is equivalent to the simple exponential smoothing model is the ARIMA(0,1,1).

Double (Brown) exponential smoothing should be used when the time series data has trend but no seasonality. The ARIMA model that is equivalent to the linear exponential smoothing model is ARIMA(0,2,2).

Seasonal exponential smoothing should be used when the time series data has no trend but has seasonality.

Winters additive or multiplicative exponential smoothing models should be used when the time series data has trend ***and*** seasonality (Fomby 2008).

ESM Parameters and Keywords

ESM	Parameters	Model=*Keyword*
Simple	ω	SIMPLE
Double	ω	DOUBLE
Linear (Holt)	ω, γ	LINEAR
Damped-Trend	ω, γ, ϕ	DAMPTREND
Seasonal	ω, δ	SEASONAL
Additive Winters	ω, γ, δ	ADDWINTERS
Multiplicative Winters	ω, γ, δ	WINTERS

14

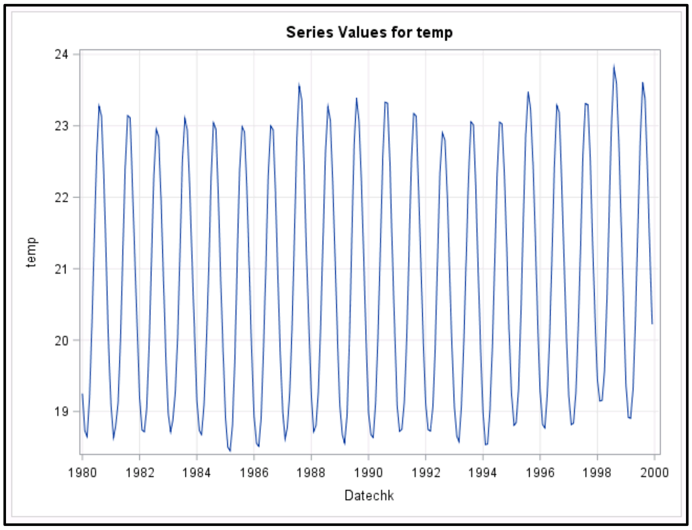
ODS Graphics

ODS Graph Name	Plot Description	PLOT= Option
ErrorACFNORMPlot	standardized autocorrelation of prediction errors	ACF
ErrorACFPlot	autocorrelation of prediction errors	ACF
ErrorHistogram	prediction error histogram	ERRORS
ErrorCorrelationPlots	prediction error plot panel	CORR
ErrorIACFNORMPlot	standardized inverse autocorrelation of prediction errors	IACF
ErrorIACFPlot	inverse autocorrelation of prediction errors	IACF
ErrorPACFNORMPlot	standardized partial autocorrelation of prediction errors	PACF
ErrorPACFPlot	partial autocorrelation of prediction errors	PACF
ErrorPeriodogramPlot	periodogram of prediction errors	PERIODOGRAM
ErrorPlot	plot of prediction errors	ERRORS
ErrorSpectralDensityPlot	combined periodogram and spectral density estimate plot	SPECTRUM
ErrorWhiteNoiseLogProbPlot	white noise log probability plot of prediction errors	WN
ErrorWhiteNoiseProbPlot	white noise probability plot of prediction errors	WN
ForecastsOnlyPlot	forecasts only plot	FORECASTSONLY
ForecastsPlot	forecasts plot	FORECASTS
LevelStatePlot	smoothed level state plot	LEVELS
ModelForecastsPlot	model and forecasts plot	MODELFORECASTS
ModelPlot	model plot	MODELS
SeasonStatePlot	smoothed season state plot	SEASONS
TrendStatePlot	smoothed trend state plot	TRENDS

15

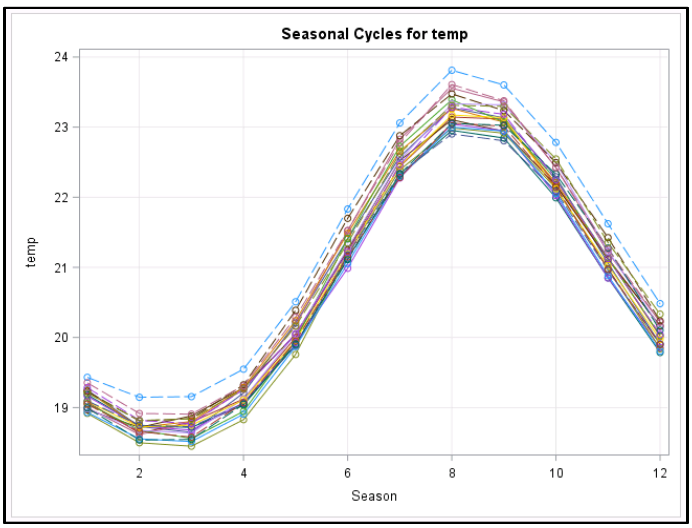
Because of ODS Graphics capabilities, the exponential smoothing procedure is capable of producing many ODS graphs during the analysis. This table displays the ODS graph name, a brief plot description, and the PLOT= option that generates the image. The ODS graph name can be used in conjunction with ODS SELECT, ODS EXCLUDE, or ODS OUTPUT statements to restrict displayed output or to produce an output data set for further use.

ESM ODS Output



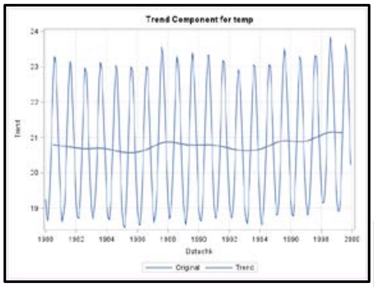
16

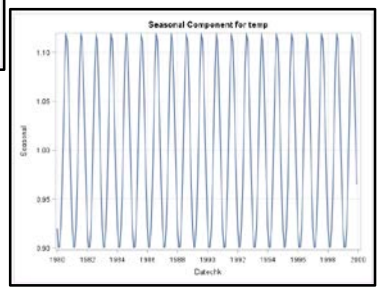
ESM ODS Output



17

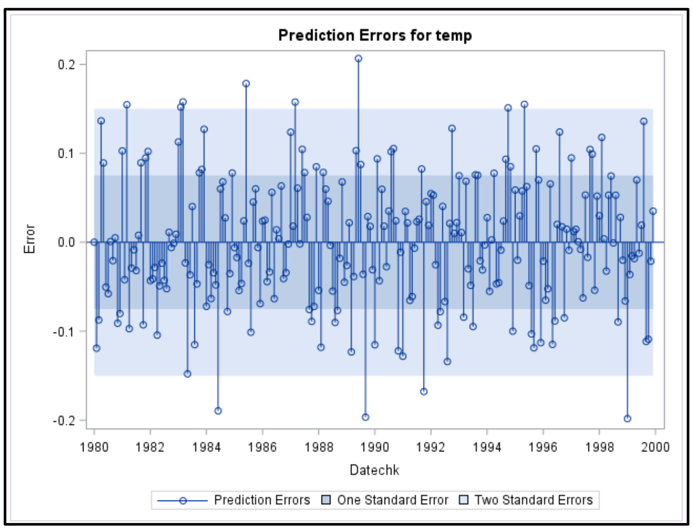
ESM ODS Output





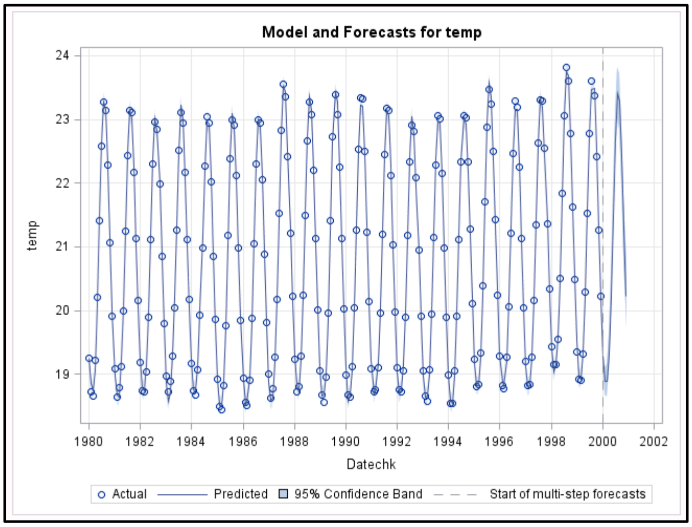
18

ESM ODS Output



19

ESM ODS Output



20

PROC ESM Syntax

```
PROC ESM DATA=SAS-data-set OUT=SAS-data-set
     OUTEST=SAS-data-set
     OUTFOR=SAS-data-set
     OUTSTAT=SAS-data-set
     OUTSUM=SAS-data-set
     SEASONALITY=n
     PLOT=option|(options)
     PRINT=option|(options)
     LEAD=n
     <options>;
     BY variables;
     ID variable INTERVAL=interval;
     FORECAST variables / MODEL=model <options>;
RUN;
```

21

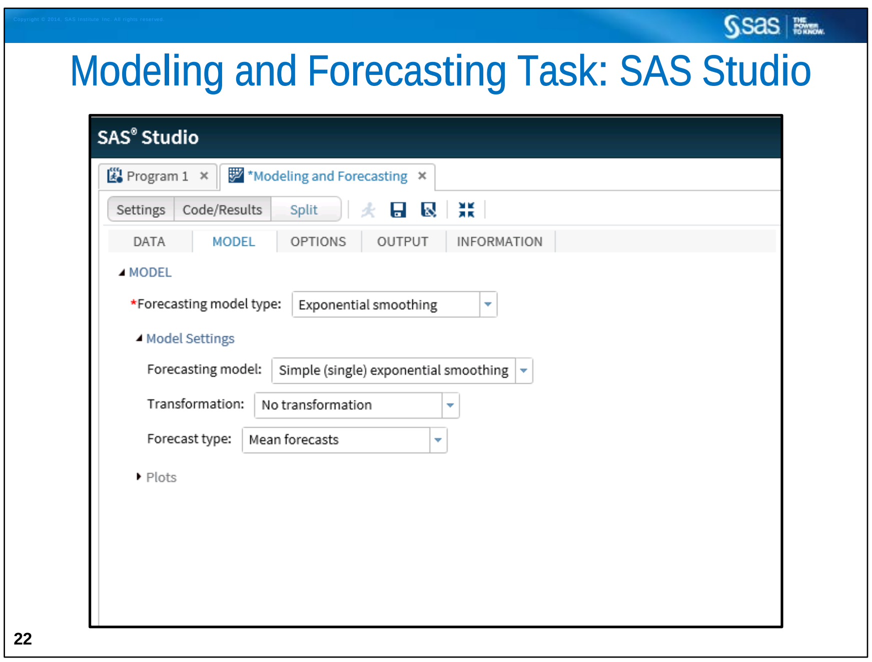
Selected ESM procedure options and statements:

OUTEST= names the output data set to contain the model parameter estimates and the associated test statistics and probability values. This data set is useful for evaluating the significance of the model parameters and understanding the model dynamics.

OUTFOR= names the output data set to contain the forecast time series components (actual, predicted, lower confidence limit, upper confidence limit, prediction error, prediction standard error). This is useful for displaying the forecasts in tabular or graphical form.

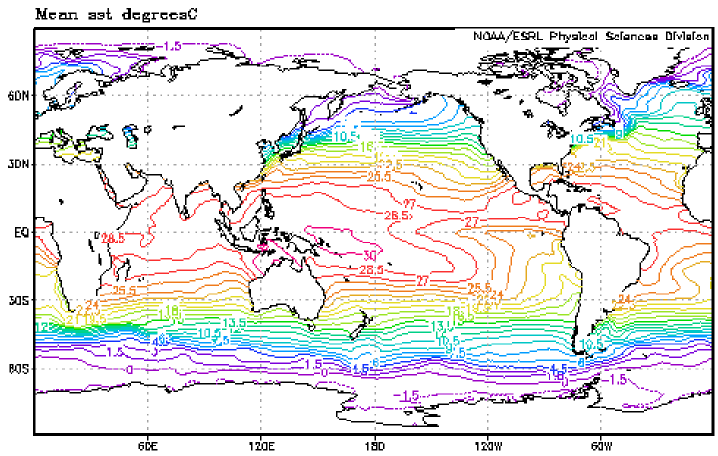
OUTSTAT= names the output data set to contain the statistics of fit (goodness of fit). This is useful for evaluating how well the model fits the series.

OUTSUM=	names the output data set to contain the summary statistics and the forecast summation. This is useful when forecasting large numbers of series and a summary of the results is needed.
SEASONALITY=	specifies the length of the seasonal cycle.
PLOT=	specifies the graphical output that is desired.
PRINT=	specifies the printed output that is desired.
LEAD=	specifies the number of periods ahead to forecast. The default is LEAD=12.
ID	names a numeric variable, assumed to be SAS data or time data valued, that identifies observations in the input and output data sets.
INTERVAL=	specifies the frequency of the input time series such as quarterly, monthly, weekly, and so on.
FORECAST	lists the numeric variables in the DATA= data set whose accumulated values are the time series to be modeled and forecast.
MODEL=	specifies the forecasting model to be used to forecast the time series. The default is MODEL=SIMPLE, which performs simple exponential smoothing.



Exponential smoothing models can be performed using the Modeling and Forecasting task in SAS Studio. After you choose the data set to analyze and set the necessary roles for variables on the DATA tab, exponential smoothing can be selected as the forecasting model type on the MODEL tab. Within the model settings, options that include the type of forecasting model to use can be selected.

Sea Surface Temperatures (SST)



24

The demonstration examines North Atlantic sea surface temperatures (SST). The data in **STSM.SST** and the image were provided by NOAA/OAR/ESRL PSD, Boulder, Colorado, USA (**http://www.esrl.noaa.gov/psd/**). The data provide monthly average sea surface temperatures for the North Atlantic.

The data file was augmented into the proper time series formatting. It contains the following variables:

Datechk SAS date variable providing the month and year of the observation

Temp numeric average sea surface temperature for that respective time point

The data spans from the mid-1800s until the present. The class focus is reduced to the 1980s and 1990s.

Analyzing Sea Surface Temperatures Using SAS Studio

In this demonstration, you use both the Time Series Exploration task and the Modeling and Forecasting task in SAS Studio to explore the data set and generate an exponential smoothing model for sea surface temperature. The **SST** data set can be found in the **STSM** library. Use the Exploration task to investigate the possible inclusion of a trend or seasonality component. Then use the Modeling and Forecasting task
to generate the appropriate exponential smoothing models. Include forecasts 12 months into the future.

1. Expand the **Tasks** area in the left navigation pane as well as the **Forecasting** subsection. Double-click the **Time Series Exploration** task. Click the **Maximized View** button.
2. On the DATA tab, select the **SST** data set within the **STSM** library.
3. Click the plus sign (+) next to **Dependent variable**. Select **temp**. Click **OK**. For this model, there is no independent variable.
4. Click and expand the **Additional Roles** section. Click the plus sign (+) next to **Time ID**. Select **Datechk**. Click **OK**. Notice that the Properties section is updated to reflect the interval of monthly data.
5. Click the **Analyses** tab. Select the **Seasonal cycles** check box. Do ***not*** clear the **Time Series** check box. Under Autocorrelation Analysis, make sure that the **Perform autocorrelation analysis** check box is selected. Change the drop-down box beside **Select plots** to display to read **Selected plots**.
6. Under Plots, select the **Autocorrelation analysis panel** and **White noise probability test (log scale)** check boxes.
7. Under Decomposition Analysis, make sure that the **Perform decomposition analysis** check box is selected. Change the drop-down box beside **Select plots** to display to read **Selected plots**.
8. Select all four check boxes (**Decomposition panel**, **Components**, **Seasonally adjusted series**, and **Seasonally adjusted series (percent change)**) under Plots.
9. When **Components** is selected, select **Trend component** and **Seasonal component** only.

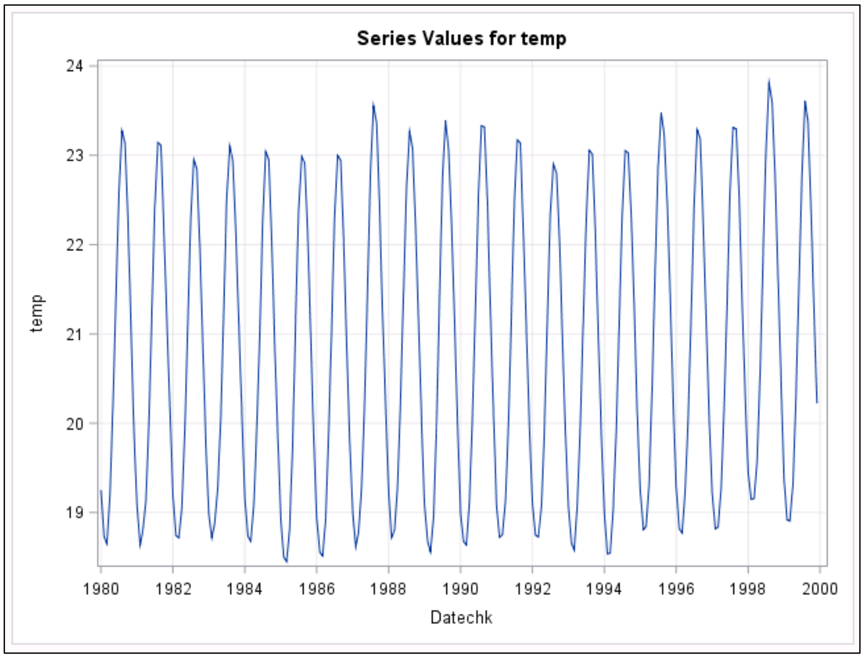
   Alternatively, you can write the SAS code directly as follows:

```
/* STSM03d01.sas */
proc timeseries data=STSM.SST
                seasonality=12
                plots=(series cycles corr wn
                       decomp sa pcsa tc sc);
   id Datechk interval=month;
   var temp;
   decomp sa pcsa tc sc / mode=multoradd;
   ods exclude WhiteNoiseProbabilityPlot;
run;
```

10. Click the running person icon to execute the task.

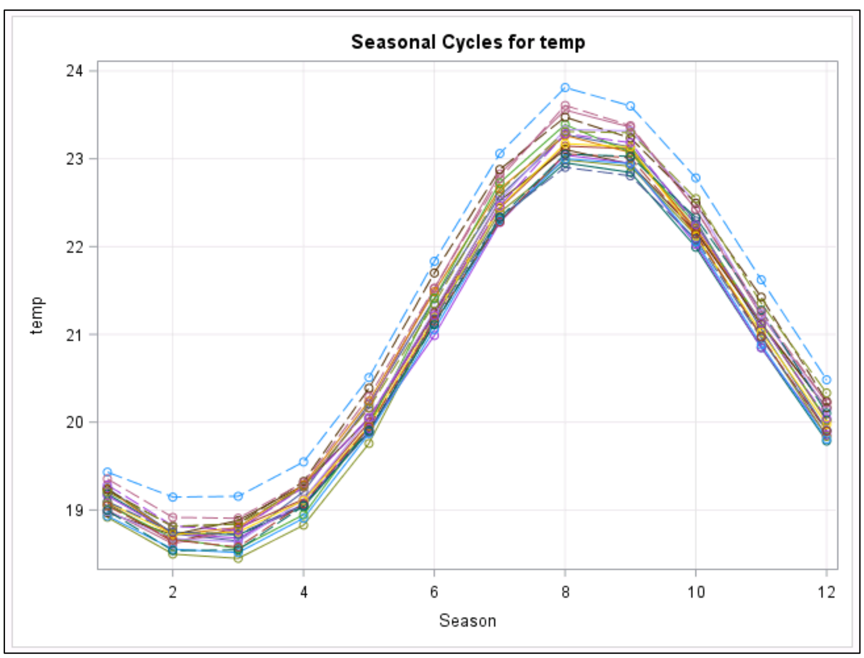
Selected Output

The output begins with a few tables that display information about the data set as well as the time series variable.

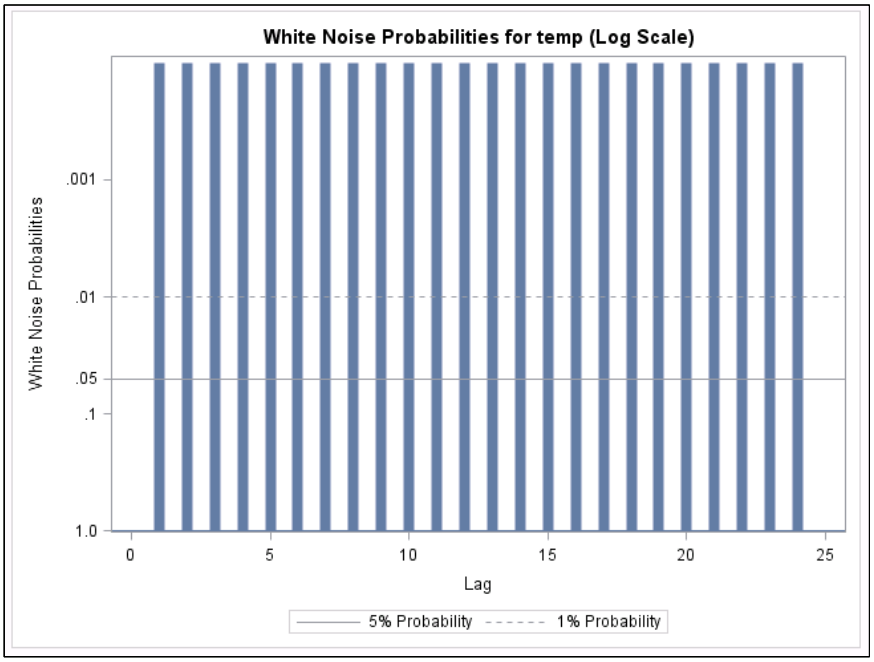


The time series plot provides a visual perspective of the dependent variable **temp** across the time ID variable **Datechk**. From this plot, you can see a clear seasonality to the data. However, it is difficult to determine whether there is a trend.

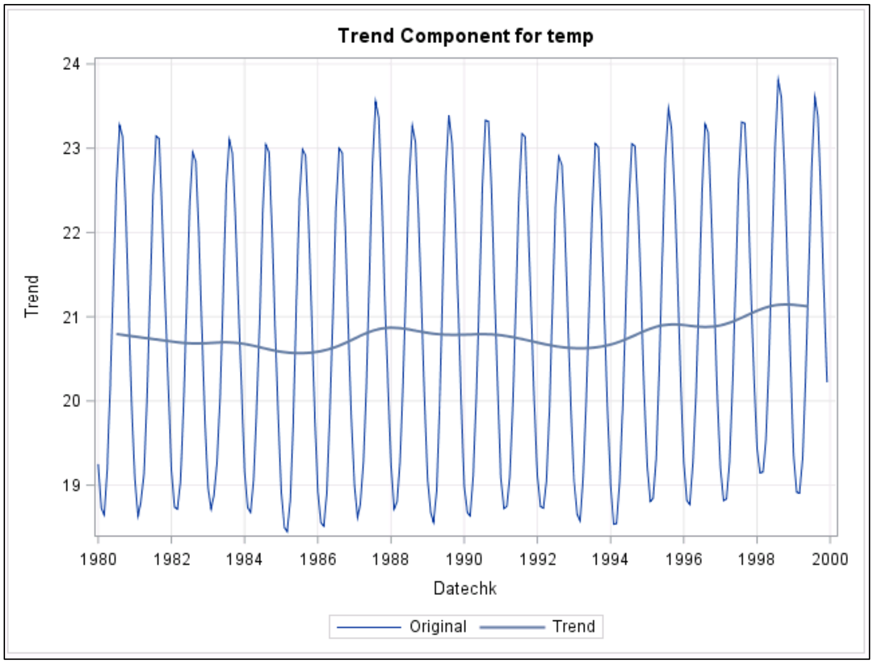
The series plot also gives an impression of whether the seasonality is additive or multiplicative. In the plot, notice that the peaks and valleys of the seasonal fluctuation appear consistent. This indicates an additive seasonality. If this plot showed peaks and valleys of the fluctuations becoming closer together or farther apart as time passed, this would indicate a multiplicative seasonality.

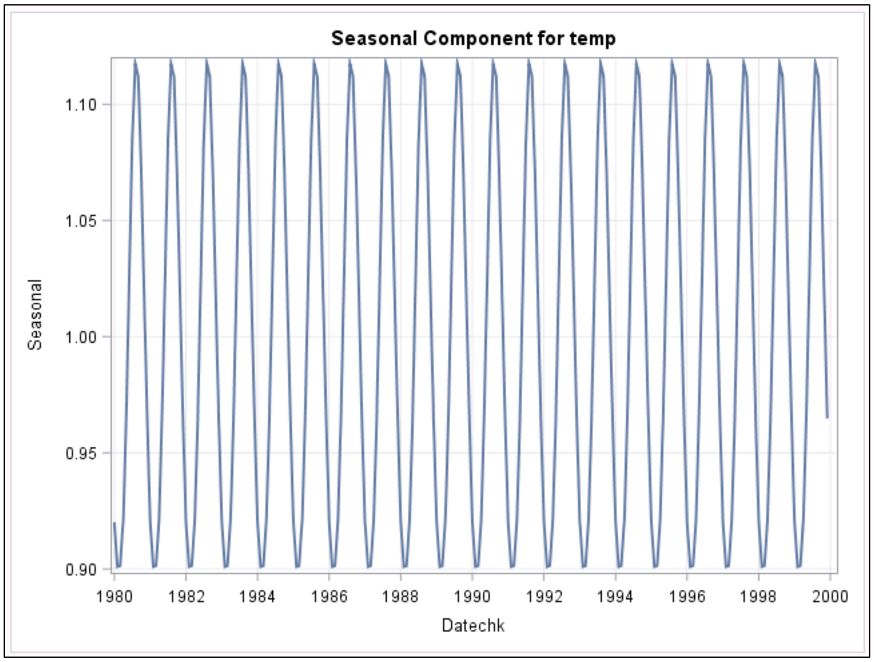


The seasonal cycles image displays a line for each year of the data. From this output, you can determine that each year followed a similar seasonal cycle. If there were lines that did not follow the pattern of the others, this could represent years when potentially some departure from the typical seasonal pattern existed. In this case, you see that the SST reached its lowest near March and its highest near August.



When the white noise probabilities are very small (*p*-values < 0.0001), it might be more advantageous to look at them using the log scale image. In either case, you can deduce that the series is not composed of only white noise. There is some attribute that can be modeled. What are these attributes? What components can you use in the modeling for this series?





Output from the decomposition analysis assists in the determination of the presence of a trend component, seasonal component, or both. From the trend component plot, you see that the trend line is rather flat. This suggests the lack of a trend in this series. From the seasonal component plot, you see a definitive seasonality that ranges from 10% below the mean at its lowest to slightly more than 10% above the mean at its highest. From these plots, you can narrow the focus to an exponential smoothing model that would incorporate seasonality but not trend.

In this case, from what you saw in the series plot earlier, this is either an additive seasonal ESM or the additive Winters model.

With the exploration complete, you can proceed to the Modeling and Forecasting task.

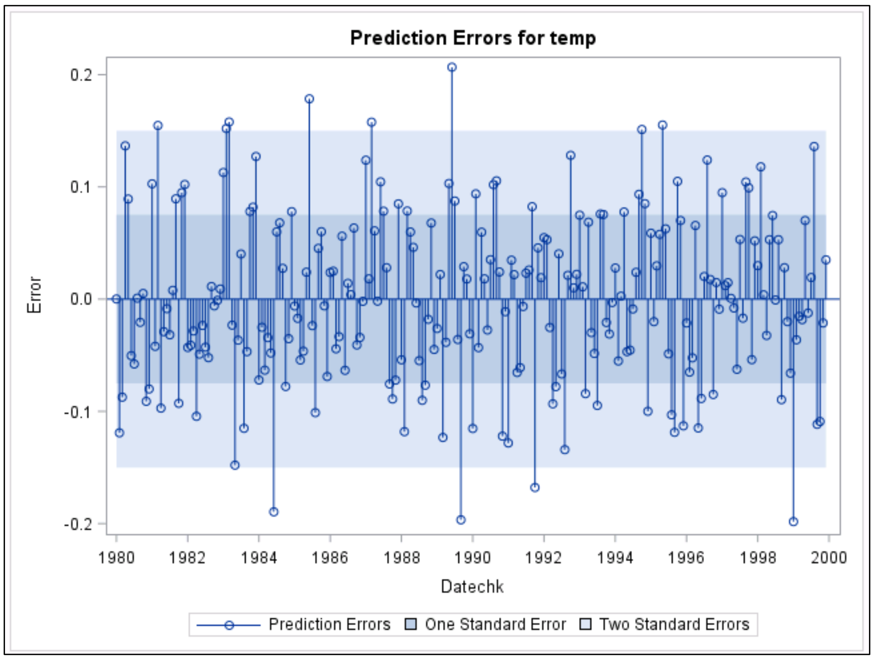
11. Exit the Maximized View. This causes the left side navigation panel to reappear.
12. Double-click the **Modeling and Forecasting** task to open it. Click the **Maximized View** button.
13. On the DATA tab, select the **SST** data set as the data set of interest. Click the plus sign (**+**) next to **Dependent variables**. Select **temp**. Click **OK**. Click and expand **Additional Roles**.
14. Click the plus sign (**+**) next to **Time ID**. Select **Datechk**. Click **OK**. The Interval area in the properties is automatically updated.
15. Click the **MODEL** tab. Next to Forecasting model type, select **Exponential smoothing** in the drop-down box. Under Model Settings, select **Additive seasonal exponential smoothing** for the Forecasting model. Do not make changes to the transformation or to the forecast type.
16. Click and expand **Plots**. Leave Select plots to display as Default plots.
17. Click the **OPTIONS** tab. Select **12** as the number of periods to forecast. (That is one year in this case.)
18. Click the **OUTPUT** tab. Select the **Create fit statistics data set** check box. Leave the name of this data set as **outstat**.

Alternatively, you can write the SAS code directly as follows:

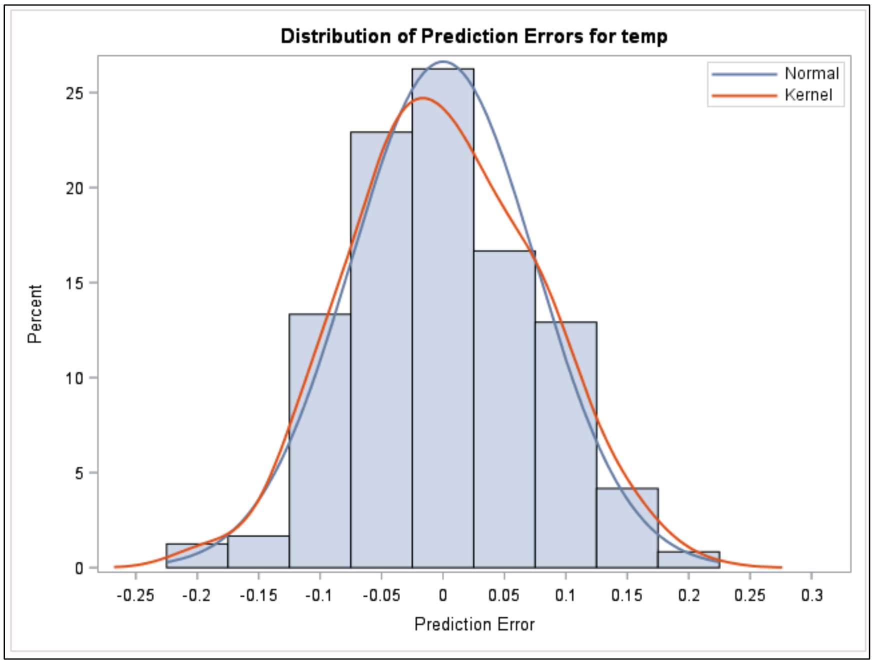
```
*Additive Seasonal Model;
proc esm data=STSM.SST
         back=0 lead=12
         seasonality=12
         plot=(corr errors modelforecasts)
         outstat=WORK.outstat;
   id Datechk interval=month;
   forecast temp / alpha=0.05 model=addseasonal;
run;
```

19. Click the running person icon to run this task.

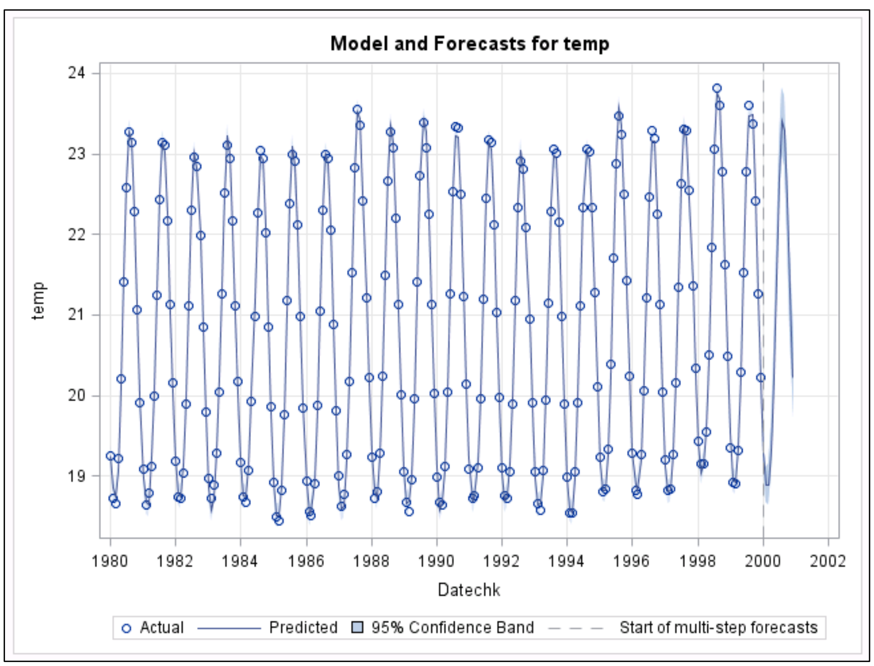
Selected Output



From the prediction errors image, you can see that the prediction error for the **temp** series generally falls within one or two standard errors from zero. A few points escape this region but not enough to cause concern.

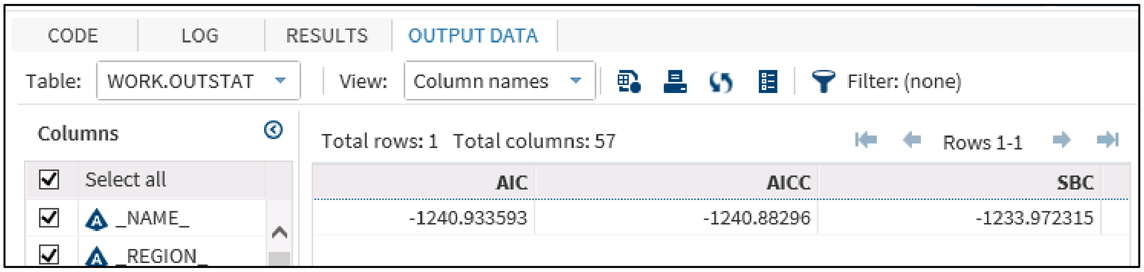


The distribution of prediction errors for the **temp** series is shown in a histogram with a superimposed normal and kernel curve. You can see that the errors do appear to be approximately normal.



You are also provided with a model and forecasts series plot for **temp**. You can see that the additive seasonal ESM does a good job of picking up the seasonality of the time series and carries this into the future forecasts.

To check the fit statistics of this additive seasonal model, check the **outstat** data set that you asked SAS to generate. Do this by clicking the **OUTPUT DATA** tab next to the RESULTS tab.



Scrolling over, you see many statistics that were generated from the model. You can choose to look at any of them but focus on the AIC (-1240.934) and SBC (-1233.972). There is another additive model that could be chosen for the time series. This is the additive Winters model. Run an additive Winters model and compare the AIC and SBC to this model.

Alternatively, you can run PROC PRINT.

```
proc print data=WORK.outstat;
    var AIC SBC;
run;
```

Obs	AIC	SBC
1	-1240.93	-1233.97

20. Click the **Model** tab and look under **Model Settings**. Change the **Forecasting model** drop-down selection to the **Winters additive** method.

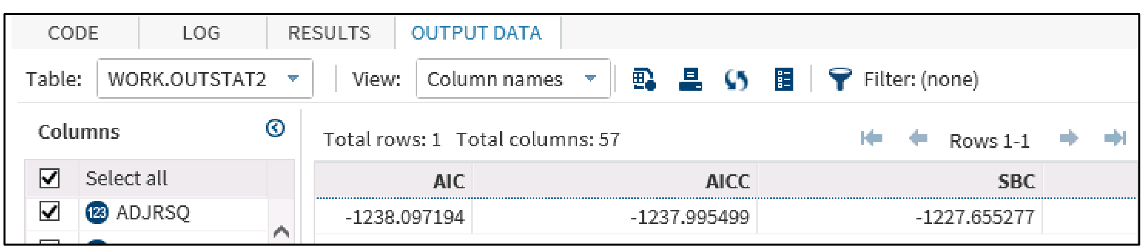
21. Under the OUTPUT tab, change the name of the fit statistics data set to **outstat2**. This enables you to retain both output data sets.

Alternatively, you can write the SAS code directly as:

```
*Additive Winters model;
proc esm data=STSM.SST
         back=0 lead=12
         seasonality=12
         plot=(corr errors modelforecasts)
         outstat=WORK.outstat2;
   id Datechk interval=month;
   forecast temp / alpha=0.05 model=addwinters;
run;
```

22. Rerun the model by clicking the running person icon.

23. Similar to the additive seasonal model, you can look at the provided images. However, click the **OUTPUT DATA** tab and look at the **outstat2** data set.



If you compare the AIC and SBC from this model, -1238.097 and -1227.655 respectively, to the additive seasonal model, the additive seasonal model has slightly smaller values and smaller is better.

Alternatively, you can run PROC PRINT.

```
proc print data=WORK.outstat;
   var AIC SBC;
run;
```

Obs	AIC	SBC
1	-1238.10	-1227.66

In both the additive seasonal model and the additive Winters model, the white noise image shows that there might still be something beyond white noise that could be incorporated into the model. This might suggest the inclusion of a predictor variable. For this, you leave the ESM methodology and incorporate another method.

End of Demonstration

Exercises

1. Determining a Trend or Seasonal Component

The data set **STSM.ConcertSales** contains weekly sales data for a new ticket provider company. The data is weekly starting in January 2005 and collected until December 2014. It is organized by the variable **Date**. The variable **sales** is already a time series with the correct accumulation for the collected time periods.

Explore this to determine the presence of a trend or seasonal component. Based on your interpretation, run the one model of the seven ESM models that you think would be useful for this 12-week forecast.

a. Are there attributes that can be modeled within this time series?

b. Is there a trend or seasonal component (or both)?

c. After you ran your selected ESM models, what were the AIC and SBC of your selections?

End of Exercises

3.01 Multiple Answer Poll

Which of the following plots can assist in determining that there are components capable of being modeled for the time series?

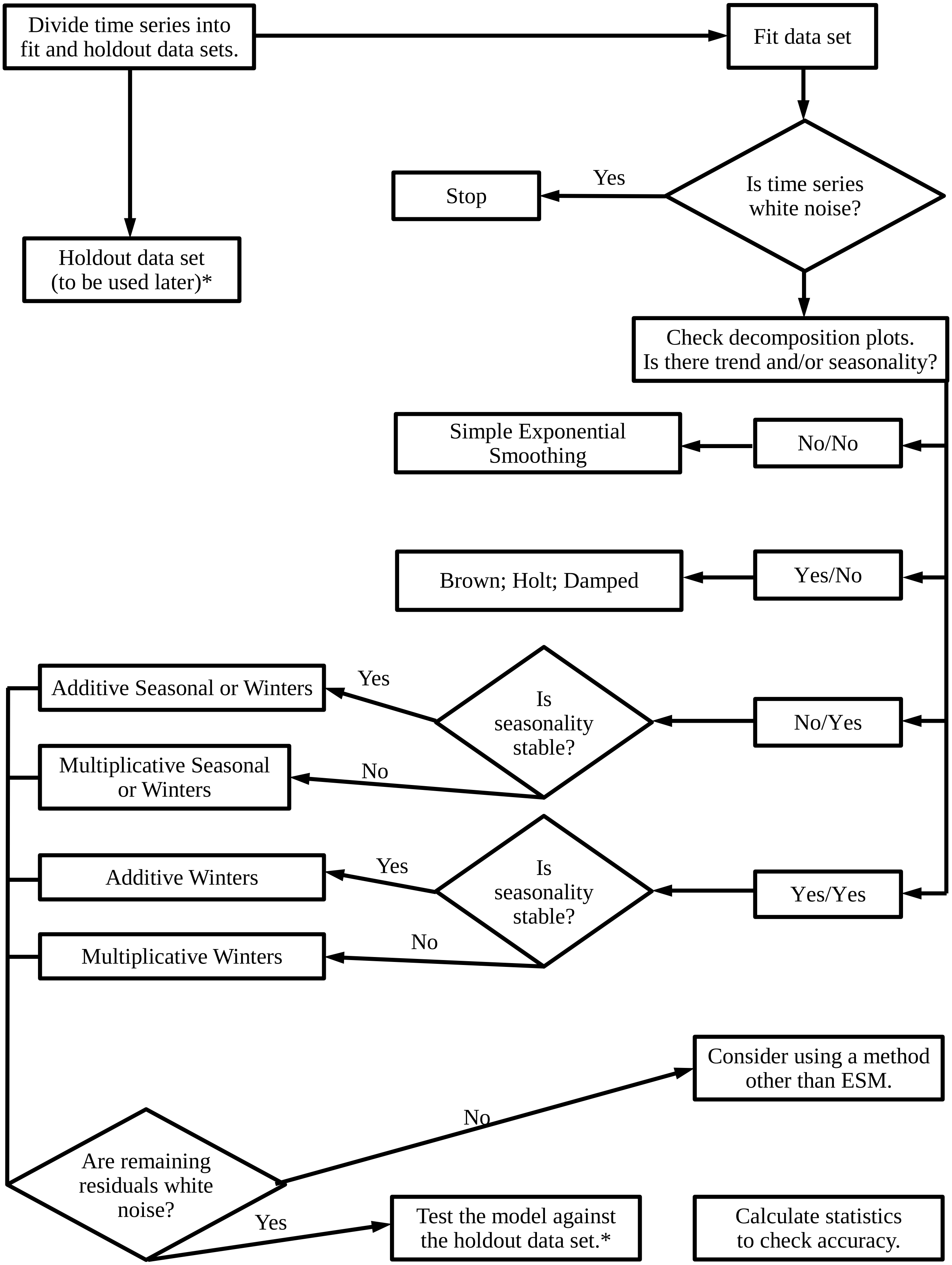
a. white noise probability
b. trend component
c. seasonal component

27

3.2 Chapter Summary

ESMs formulate forecasts using a *smoothing* method of weighted averages. In the construction of the forecasts, more recent observations are given more weight than observations in the more distant past. The *exponential* is derived from the fact that weights not only diminish over time but they do so exponentially (Fomby 2008). ESMs have the added bonus that only a few parameters are required in the forecasting model and these equations are simple to implement.

The flowchart can serve as a guide for performing exponential smoothing analysis.


Divide time series into fit and holdout data sets.
Fit data set
Holdout data set (to be used later)*
Is time series white noise?
Yes
Stop
Check decomposition plots. Is there trend and/or seasonality?
No/No
Simple Exponential Smoothing
Yes/No
Brown; Holt; Damped
No/Yes
Is seasonality stable?
Yes
Additive Seasonal or Winters
No
Multiplicative Seasonal or Winters
Yes/Yes
Is seasonality stable?
Yes
Additive Winters
No
Multiplicative Winters
Are remaining residuals white noise?
No
Consider using a method other than ESM.
Yes
Test the model against the holdout data set.*
Calculate statistics to check accuracy.