



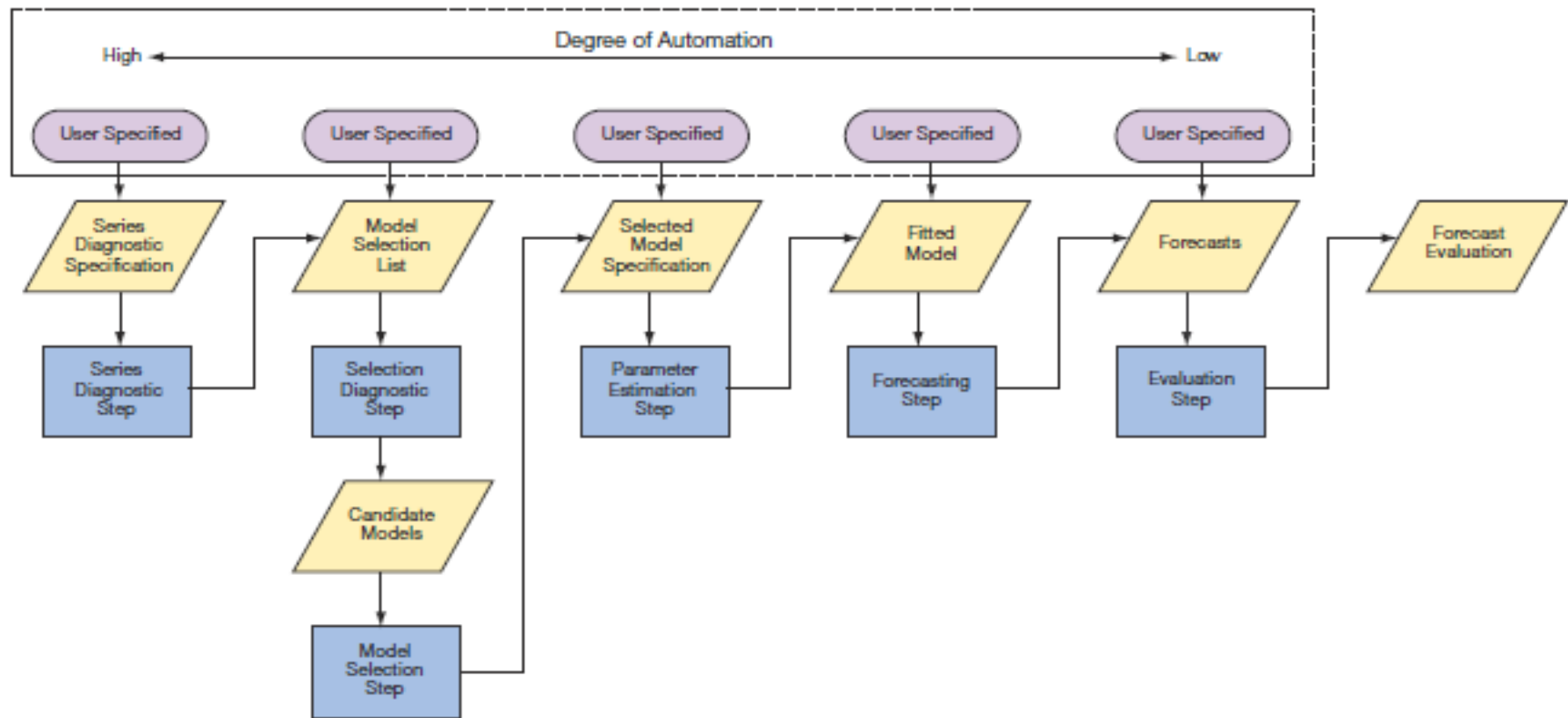
# Statistics and Samples: Judging Time Series Models Automatically

Udo Sglavo  
SAS Inc.

# What We Will Discuss Today

- Is my forecast accurate?
- Illustrative example
- Reconciliation

# Automatic Forecasting Process: The big picture

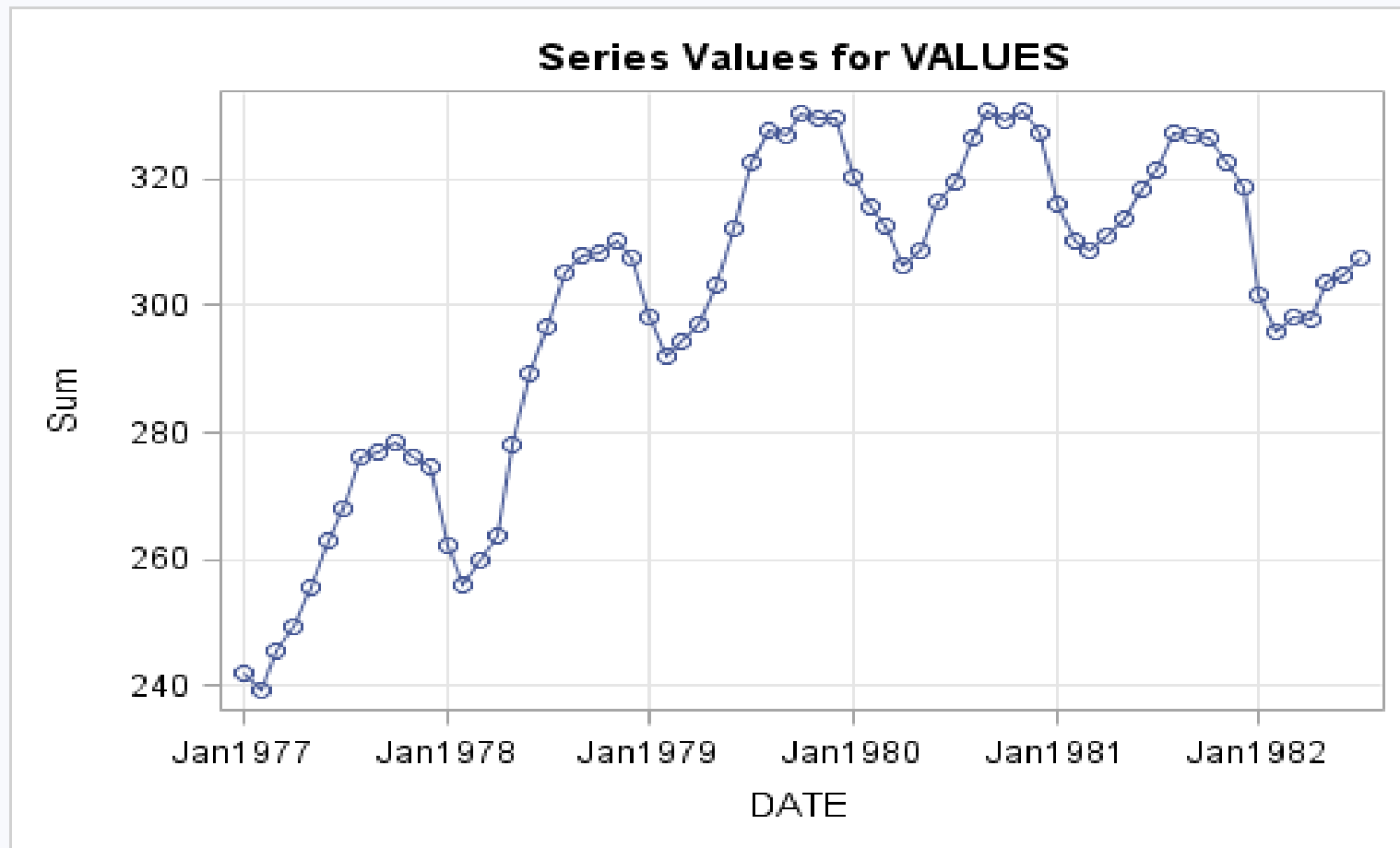


**Figure 2: Automatic forecasting information flow**

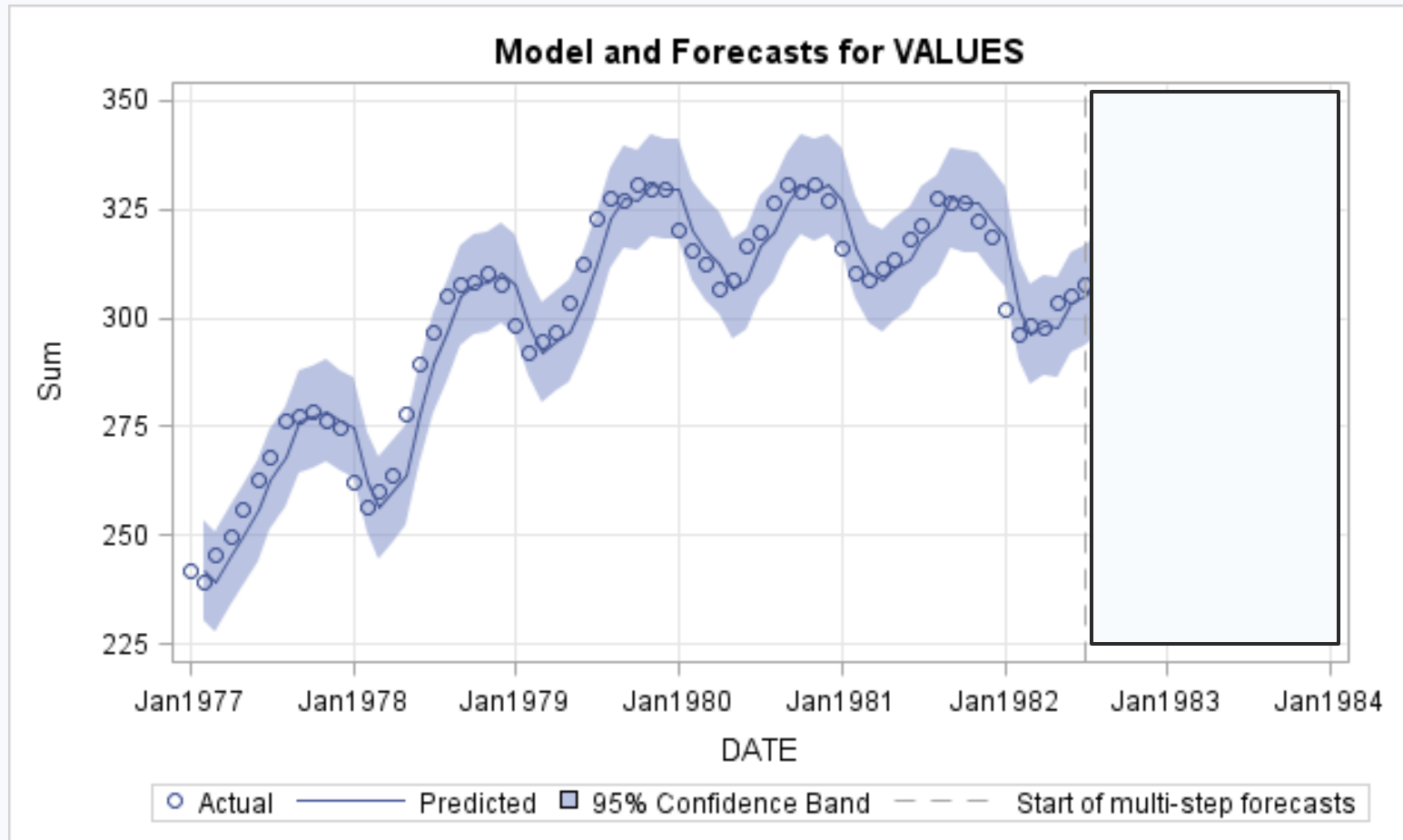
■ For details see the white paper

*'Large-Scale Automatic Forecasting Using Inputs and Calendar Events' by Michael Leonard*

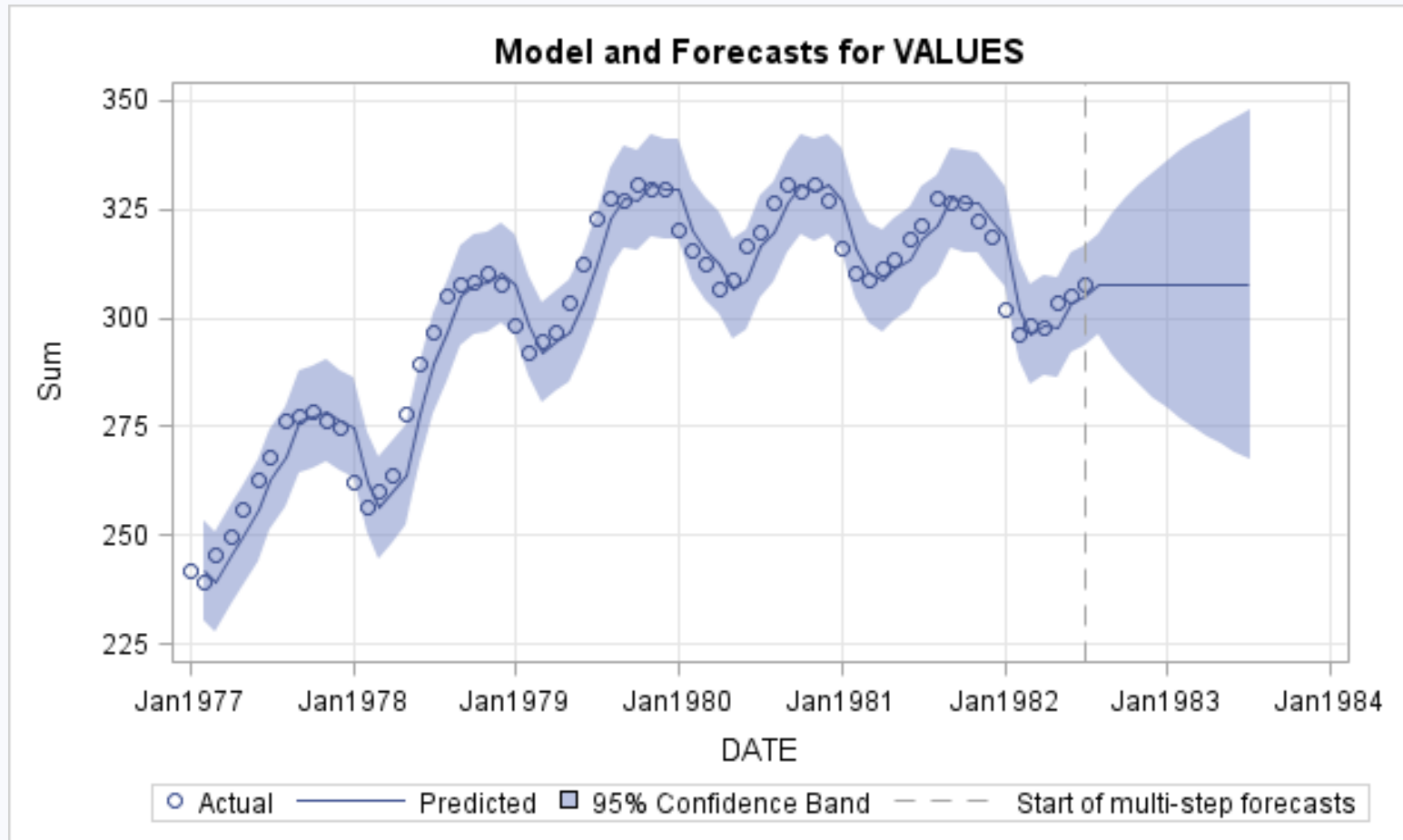
# Example Series



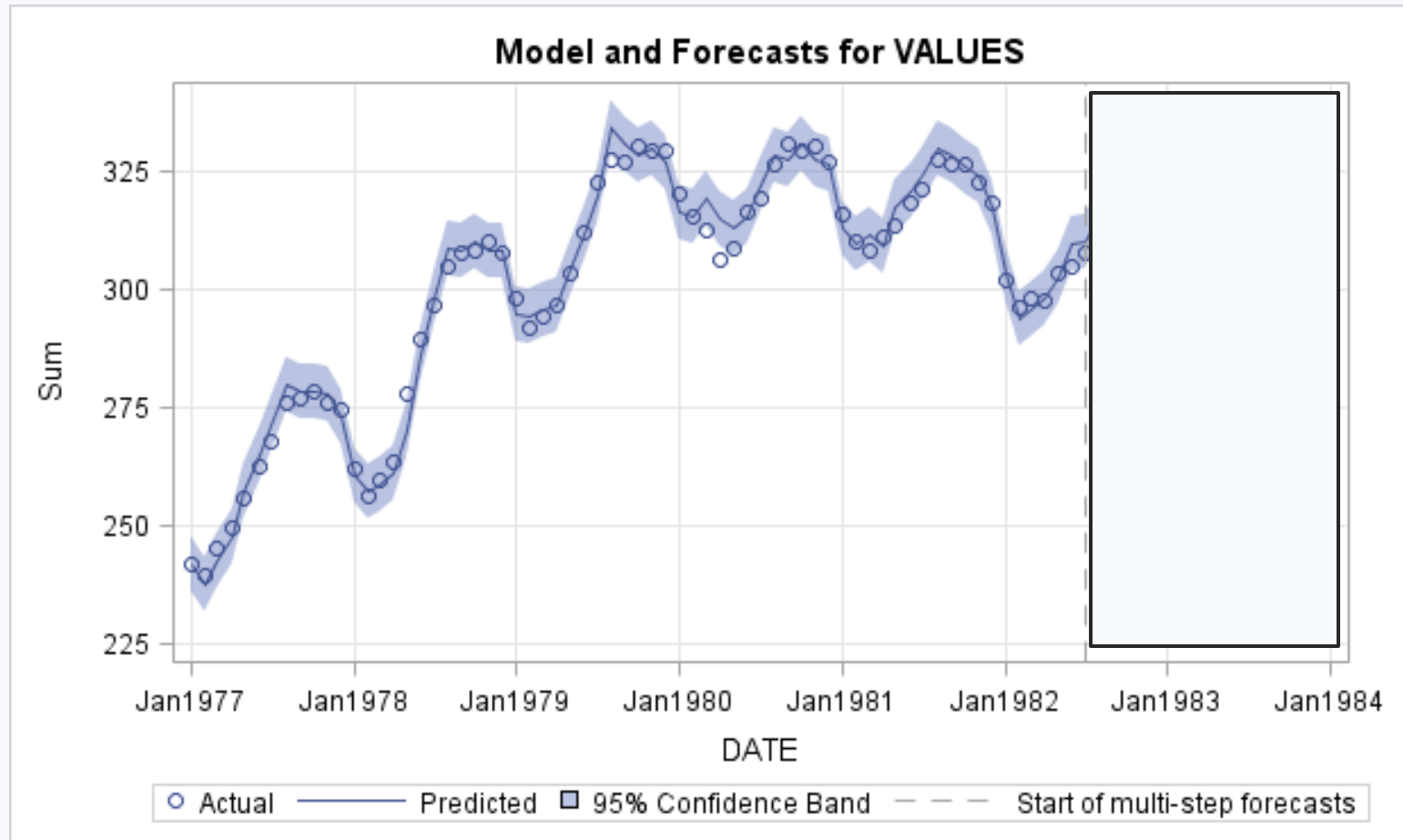
# Is This An Accurate Forecast?



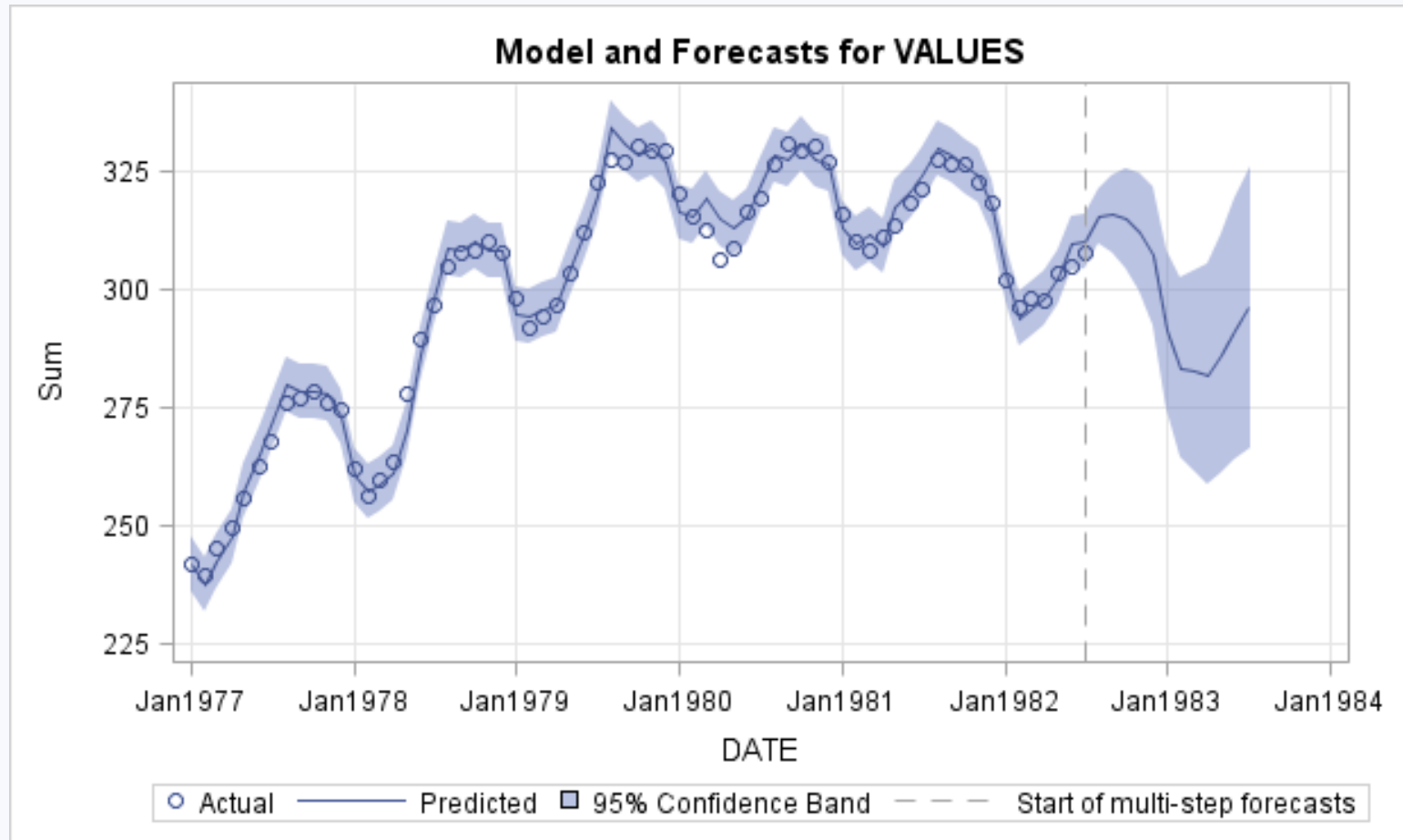
# Is This An Accurate Forecast?



# Is This An Accurate Forecast?

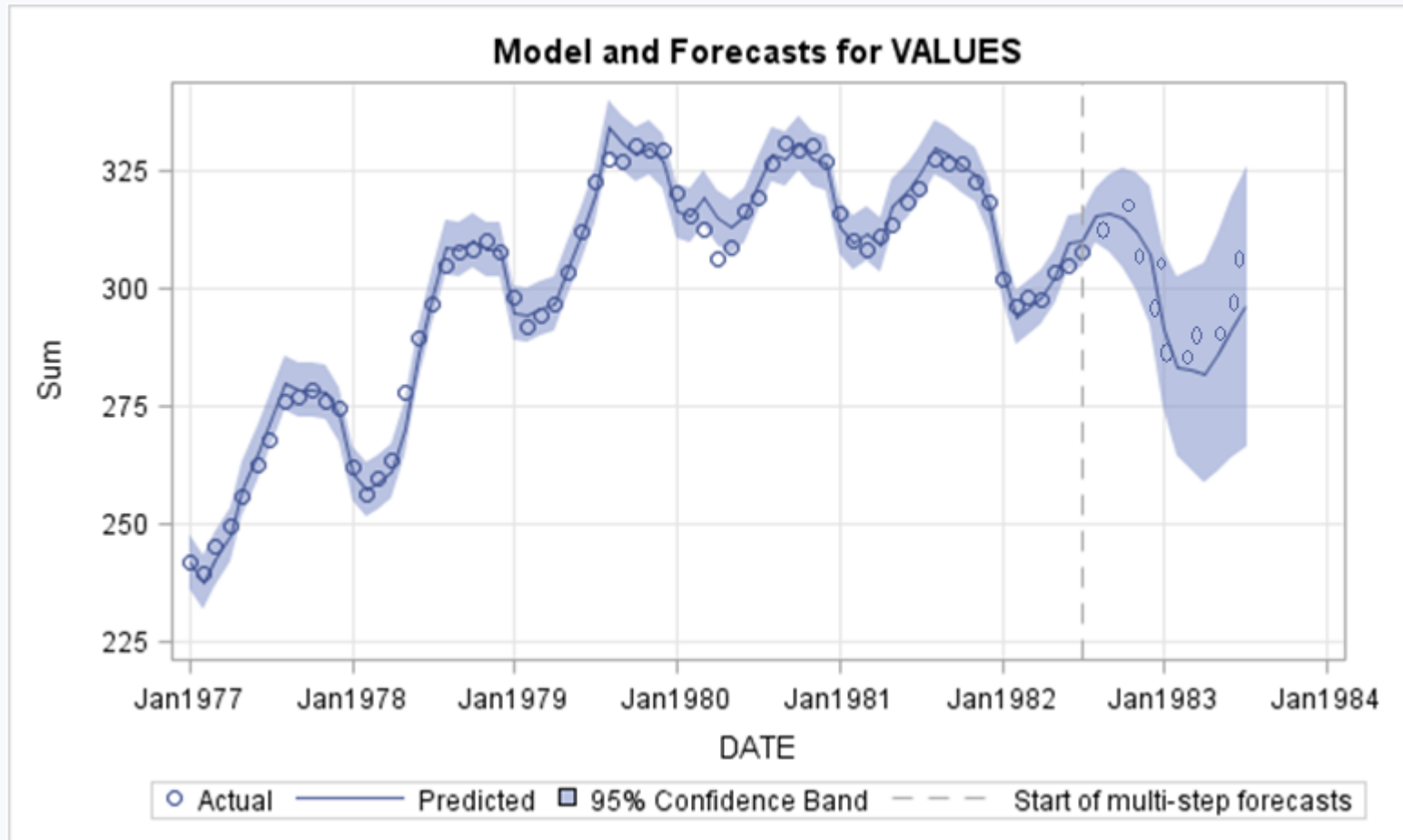


# Is This An Accurate Forecast?





# Is This An Accurate Forecast?



# Prediction Error Analysis

- Initial Fit Statistics
  - evaluate the initial model fit (fit region, in-sample)
- Selection Statistics
  - select between competing models (selection region/hold out sample, in-sample)
- Fit Statistics
  - evaluate model fit for the selected model (fit region)
- Performance Statistics
  - evaluate model selection process for the selected model (performance region, out-of-sample)

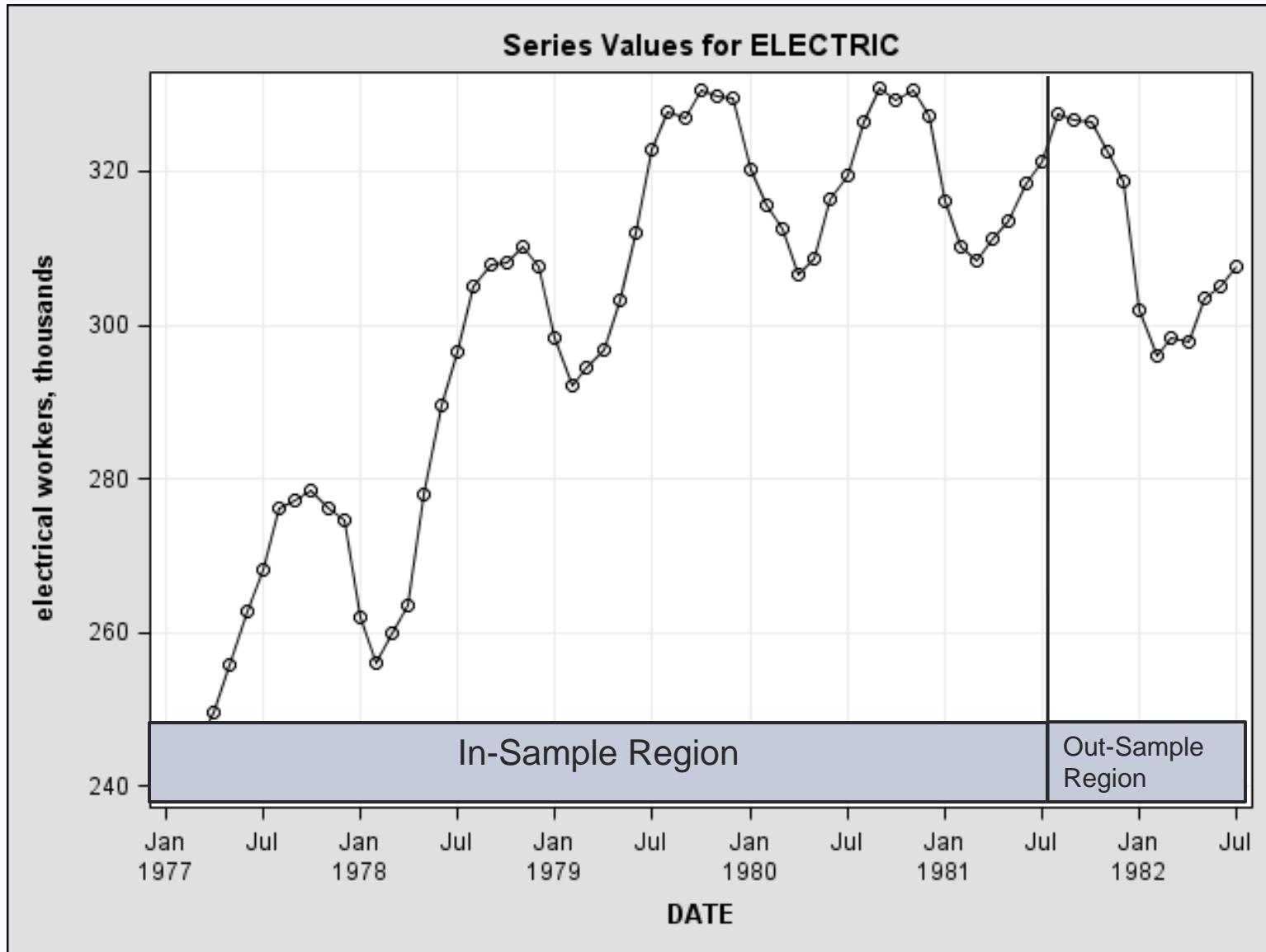
# Regions

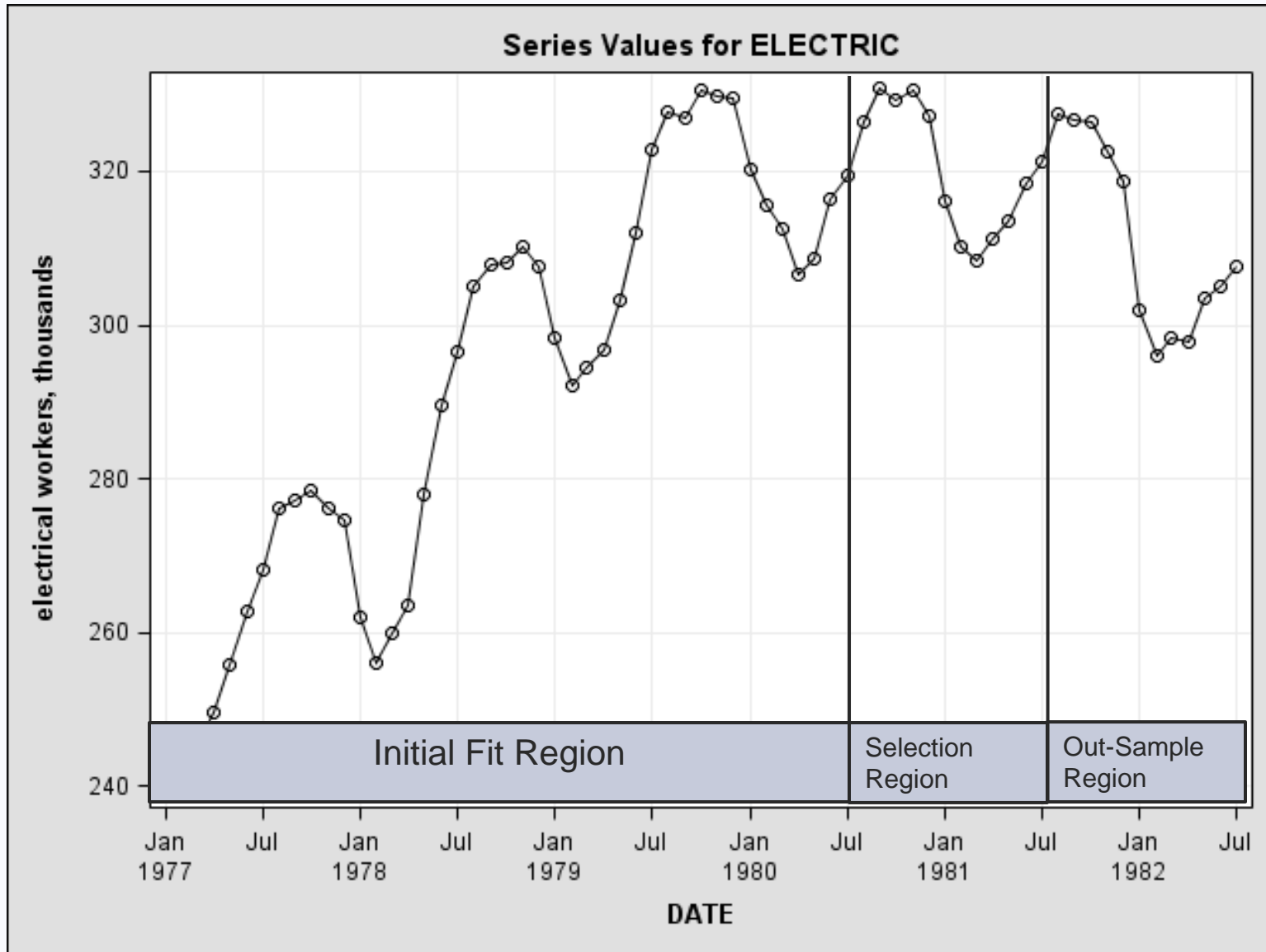
<b>Historical Time Series</b> $(1, \dots, T)$		
<b>In-Sample Region</b> $(1, \dots, T - B)$		<b>Out-Sample Region</b> $(T - B + 1, \dots, T)$
<b>Initial Fit Region</b> $(1, \dots, T - B - H)$	<b>Selection Region</b> $(T - B - H + 1, \dots, T - B)$	<b>Performance Region</b> $(T - B + 1, \dots, T)$
<b>Fit Region</b> $(1, \dots, T - B)$		<b>Forecast Horizon</b> $(T - B + 1, \dots, T - B + L)$

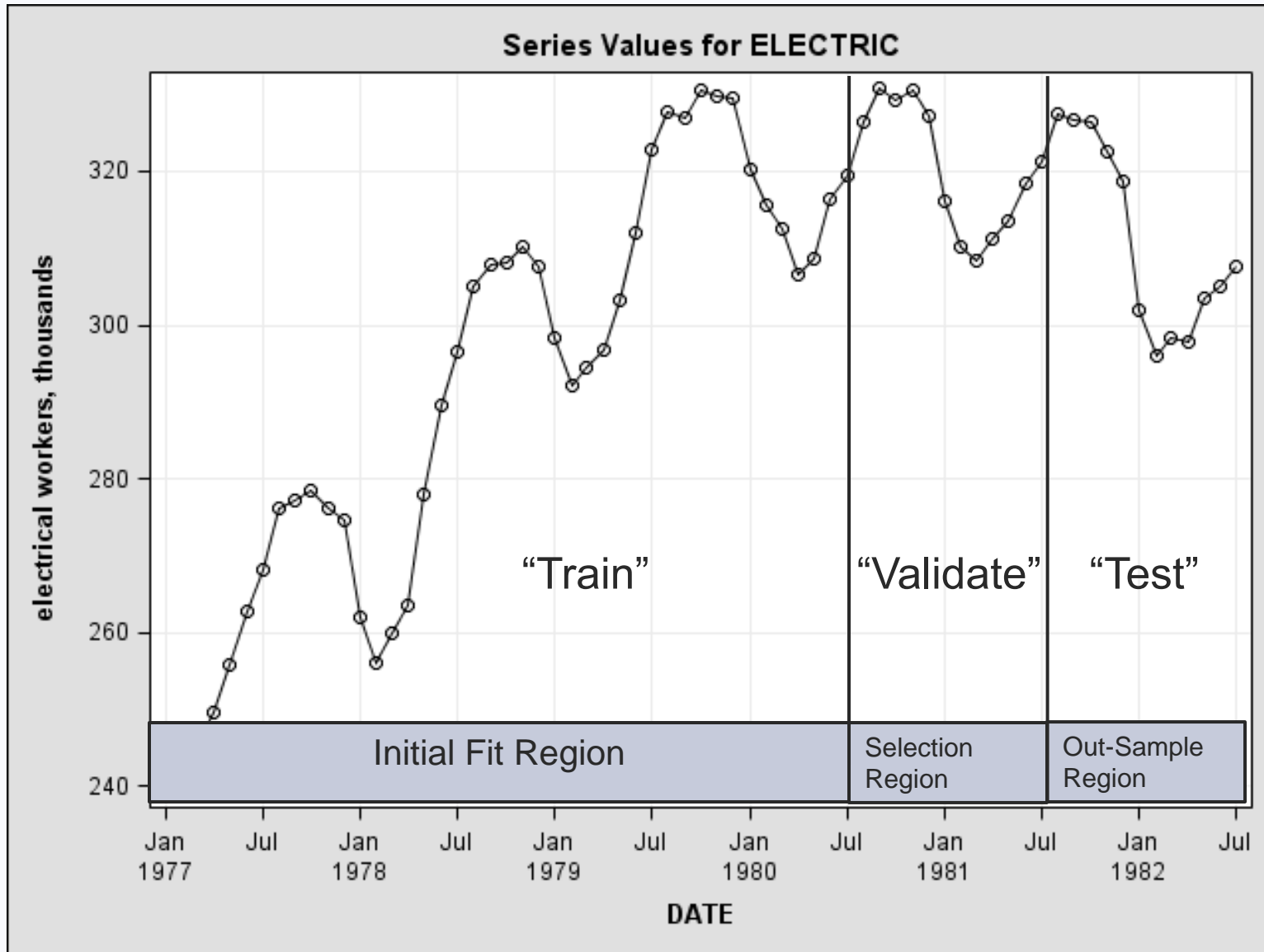
$T$  = length of time series

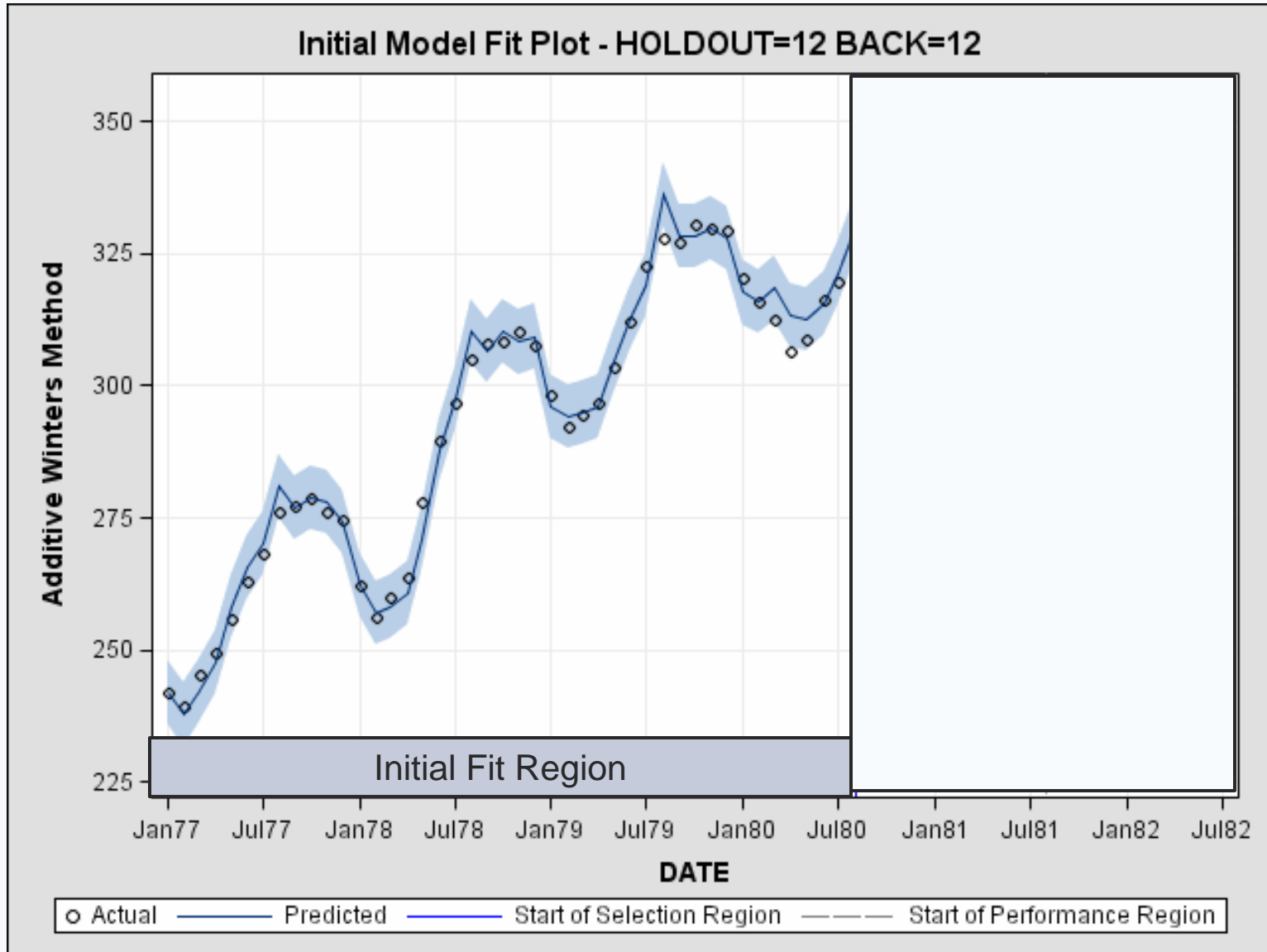
$H$  = hold-out region (in sample)

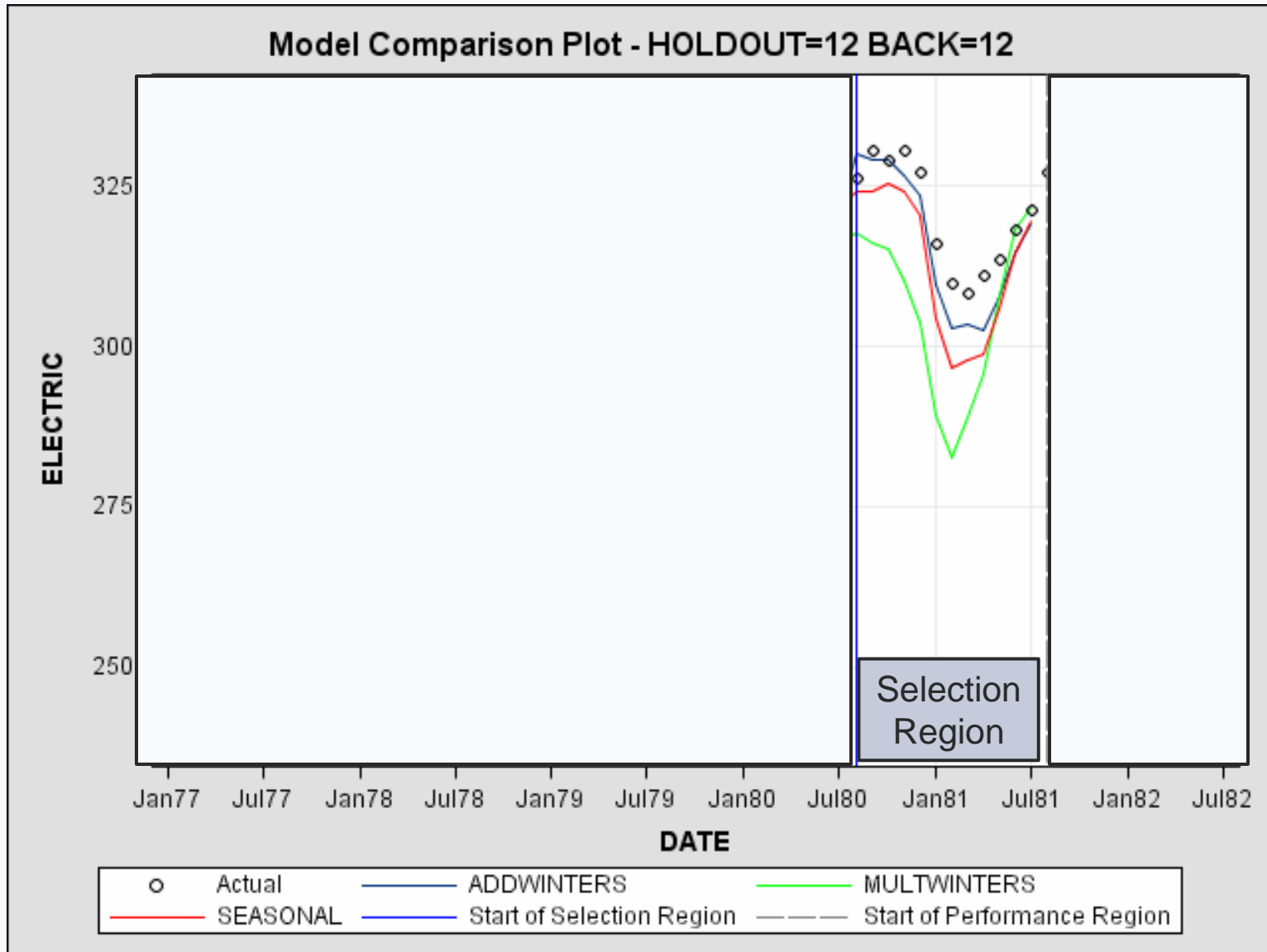
$B$  = holdback region (out-of-sample)



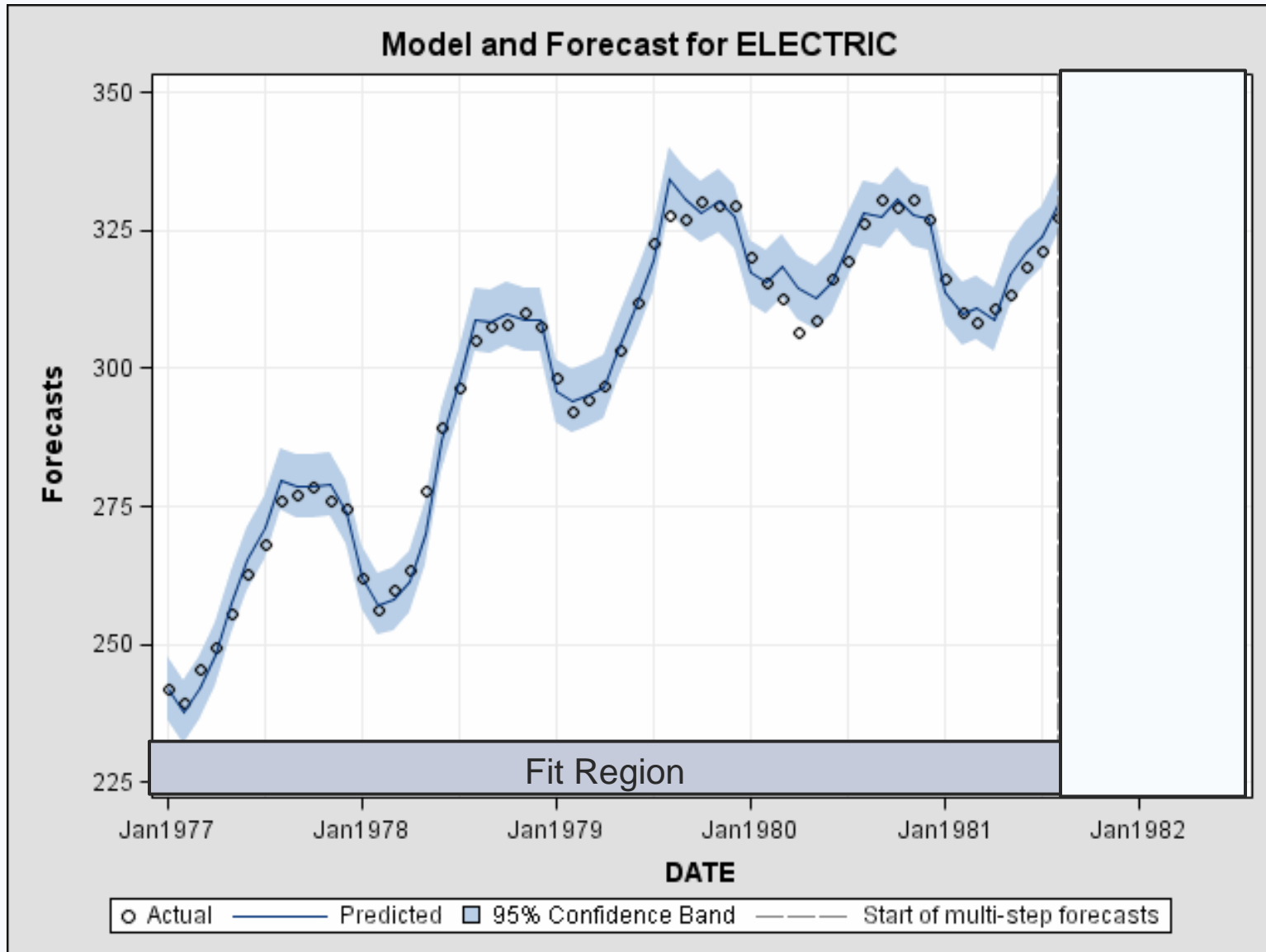


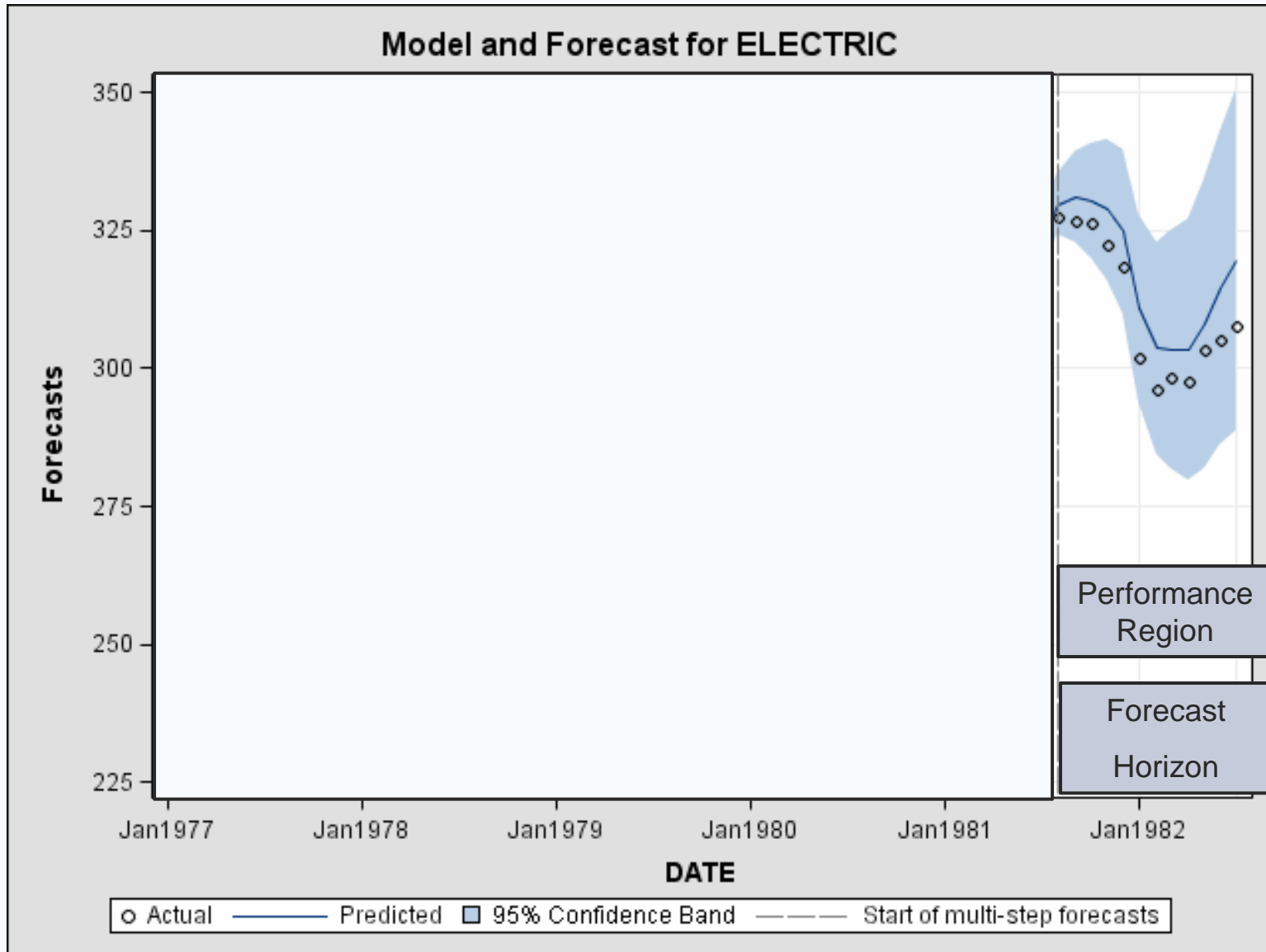












Performance Region

Forecast Horizon

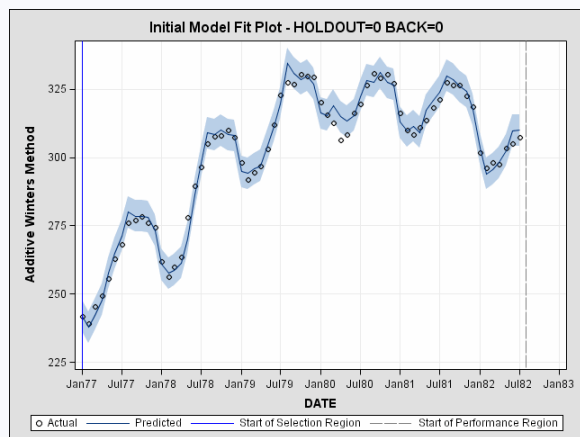
○ Actual — Predicted ■ 95% Confidence Band - - - Start of multi-step forecasts

# What We Will Discuss Today

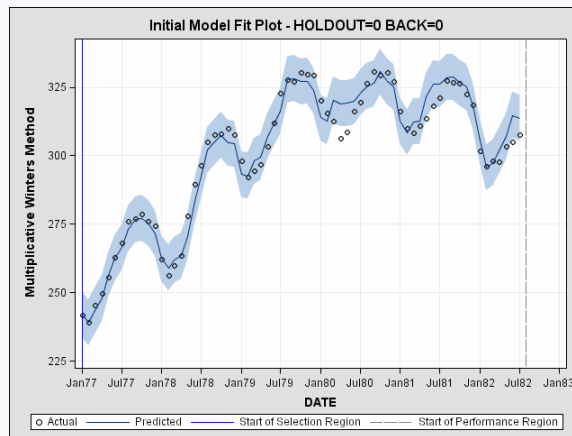
- Is my forecast accurate?
- Illustrative example
- Reconciliation

Time Regions (H = 0, B = 0)

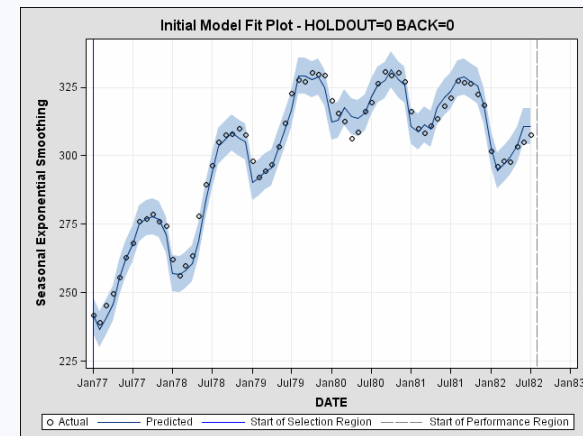
Historical Time Series (1,...,T)	Forecast Horizon (T + 1, ..., T + L)
In-Sample Region (1,...,T)	
Selection Region (1,...,T)	
Fit Region (1,...,T)	



Additive Winters Method



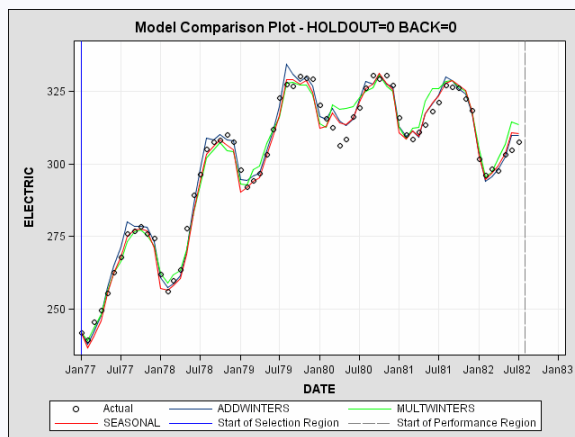
Multiplicative Winters Method



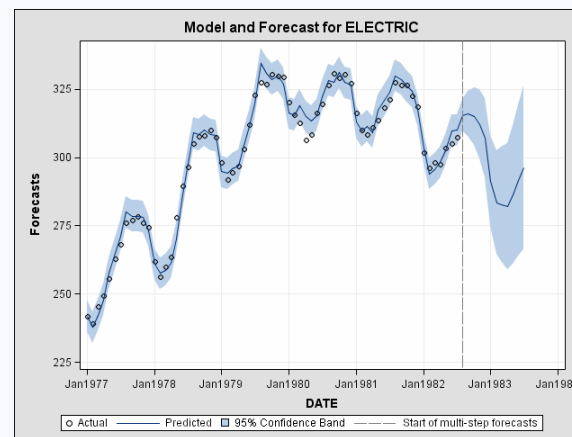
Seasonal Smoothing Method

Time Regions (H = 0, B = 0)

Historical Time Series (1,...,T)	
In-Sample Region (1,...,T)	
Selection Region (1,...,T)	
Fit Region (1,...,T)	Forecast Horizon (T + 1, ..., T + L)

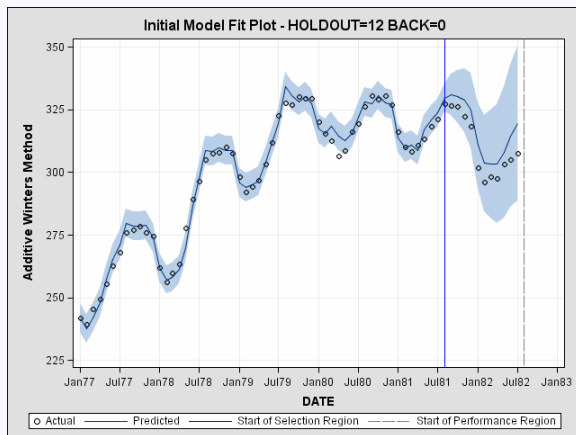
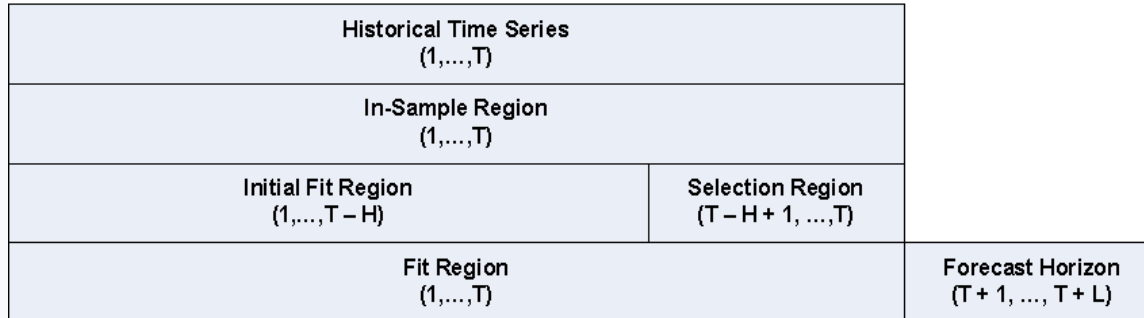


Model Comparison Plot

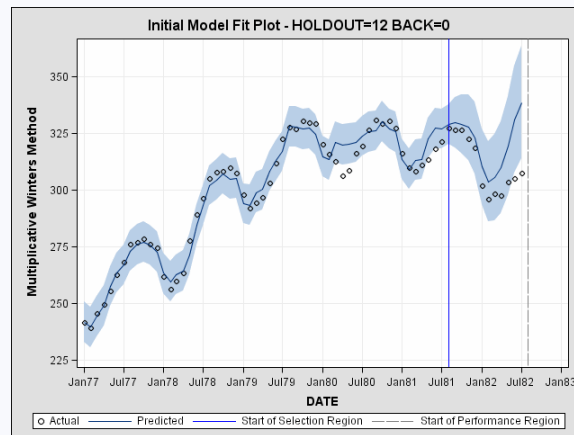


Selected Model Plot

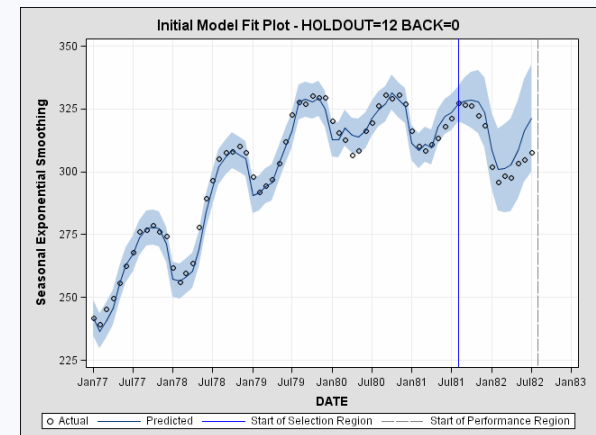
Time Regions ( $H > 0, B = 0$ )



Additive Winters Method

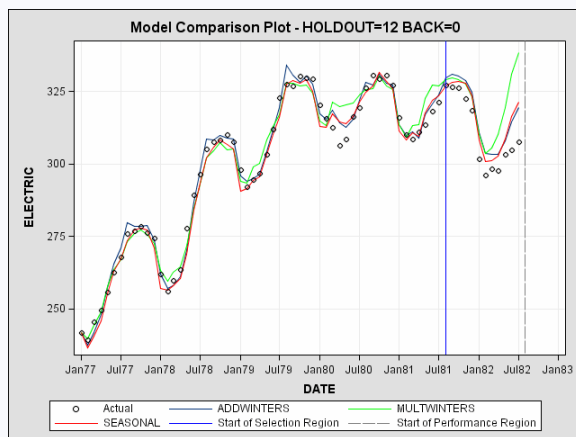
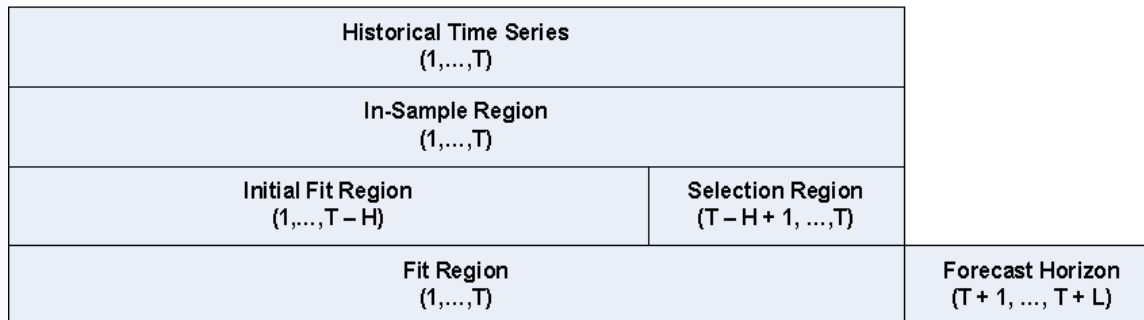


Multiplicative Winters Method

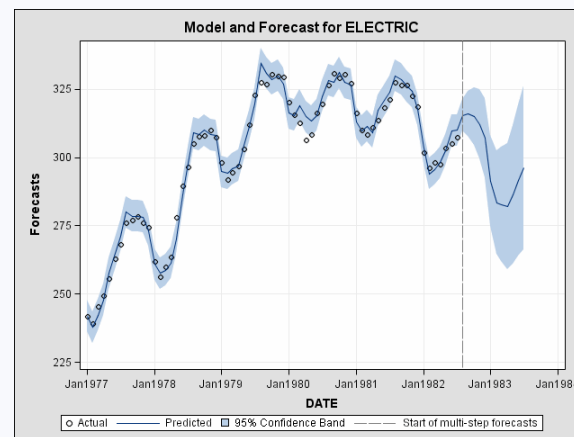


Seasonal Smoothing Method

### Time Regions (H > 0, B = 0)



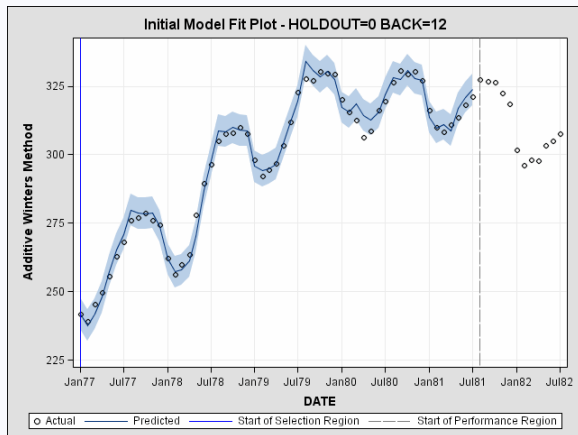
Model Comparison Plot



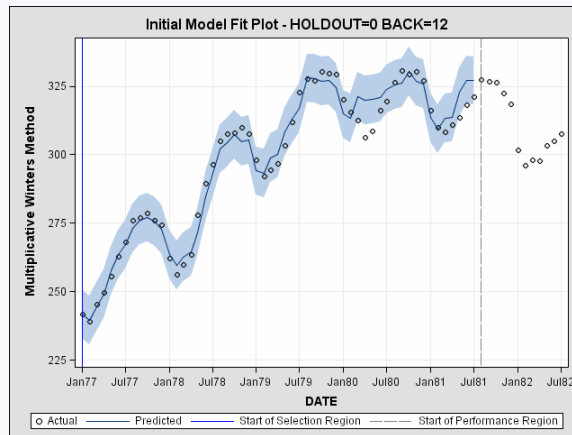
Selected Model Plot

**Time Regions (H = 0, B > 0)**

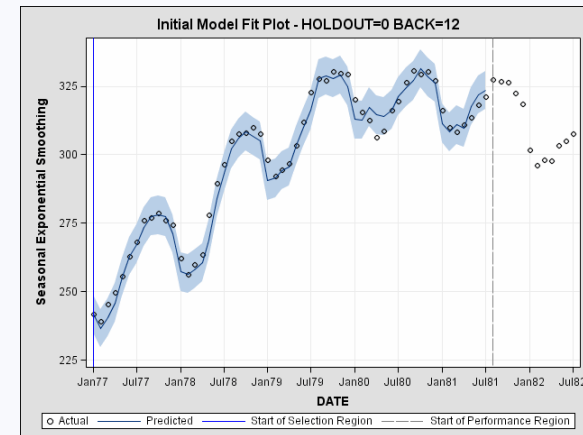
<b>Historical Time Series (1,...,T)</b>	
<b>In-Sample Region (1,...,T - B)</b>	<b>Out-Sample Region (T - B + 1,..., T)</b>
<b>Selection Region (1,...,T - B)</b>	<b>Performance Region (T - B + 1,..., T)</b>
<b>Fit Region (1,...,T - B)</b>	<b>Forecast Horizon (T - B + 1,..., T - B + L)</b>



Additive Winters Method



Multiplicative Winters Method

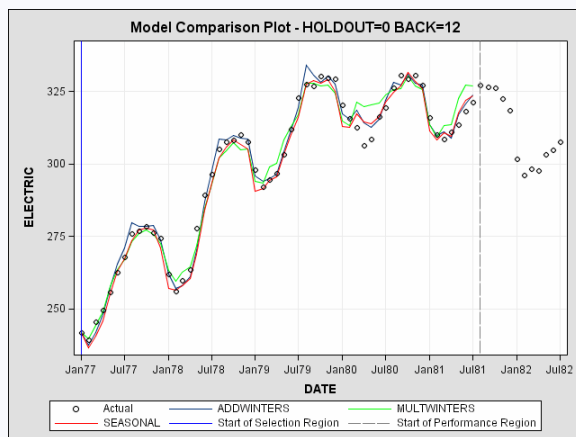


Seasonal Smoothing Method

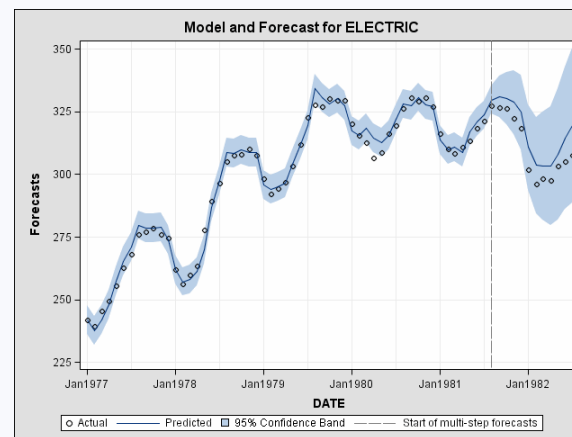


### Time Regions (H = 0, B > 0)

<b>Historical Time Series</b> (1,...,T)	
<b>In-Sample Region</b> (1,...,T - B)	<b>Out-Sample Region</b> (T - B + 1,..., T)
<b>Selection Region</b> (1,...,T - B)	<b>Performance Region</b> (T - B + 1,..., T)
<b>Fit Region</b> (1,...,T - B)	<b>Forecast Horizon</b> (T - B + 1,..., T - B + L)



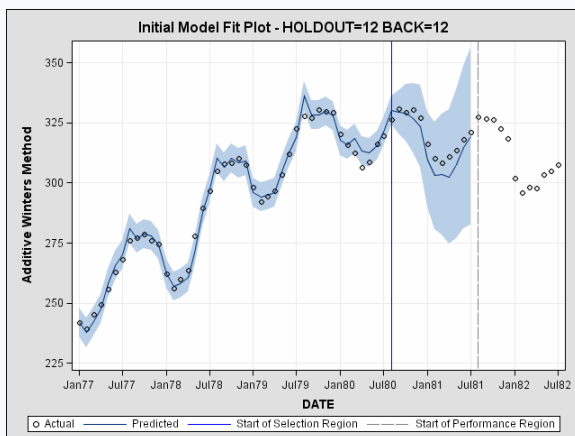
Model Comparison Plot



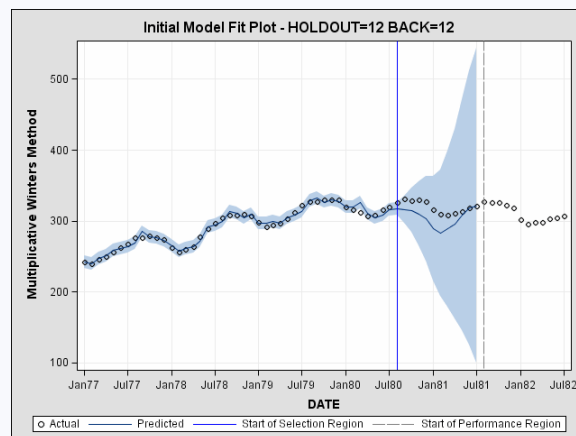
Selected Model Plot

### Time Regions (H > 0, B > 0)

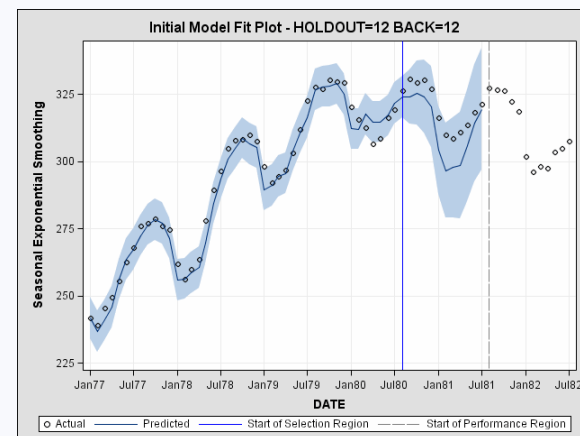
<b>Historical Time Series</b> (1,...,T)		
<b>In-Sample Region</b> (1,...,T - B)		<b>Out-Sample Region</b> (T - B + 1,..., T)
<b>Initial Fit Region</b> (1,...,T - B - H)	<b>Selection Region</b> (T - B - H + 1,...,T - B)	<b>Performance Region</b> (T - B + 1,..., T)
<b>Fit Region</b> (1,...,T - B)		<b>Forecast Horizon</b> (T - B + 1,..., T - B + L)



Additive Winters Method



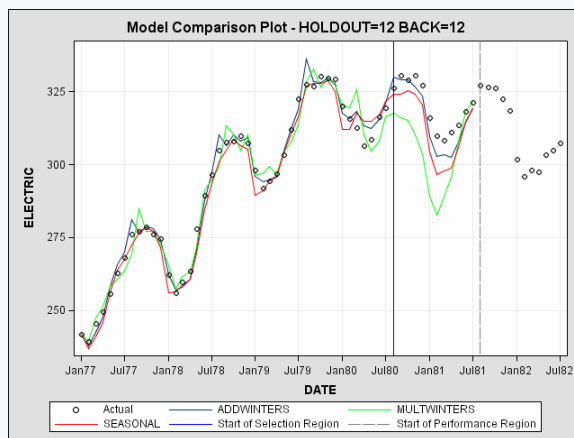
Multiplicative Winters Method



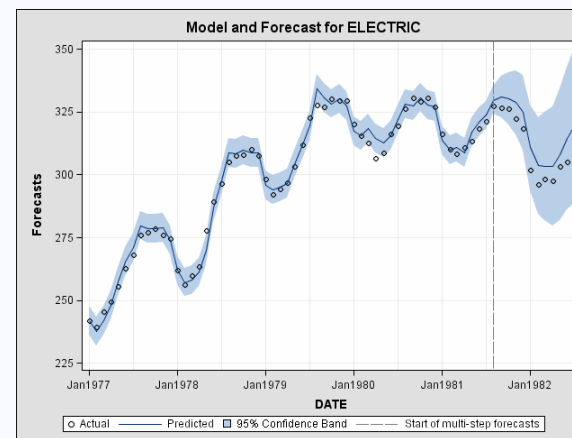
Seasonal Smoothing Method

### Time Regions ( $H > 0, B > 0$ )

<b>Historical Time Series</b> (1, ..., T)		
<b>In-Sample Region</b> (1, ..., T - B)		<b>Out-Sample Region</b> (T - B + 1, ..., T)
<b>Initial Fit Region</b> (1, ..., T - B - H)	<b>Selection Region</b> (T - B - H + 1, ..., T - B)	<b>Performance Region</b> (T - B + 1, ..., T)
<b>Fit Region</b> (1, ..., T - B)		<b>Forecast Horizon</b> (T - B + 1, ..., T - B + L)



Model Comparison Plot



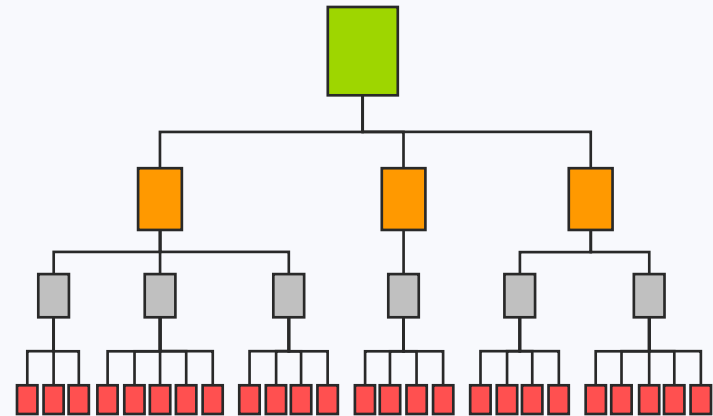
Selected Model Plot

# What We Will Discuss Today

- Is my forecast accurate?
- Illustrative example
- Reconciliation

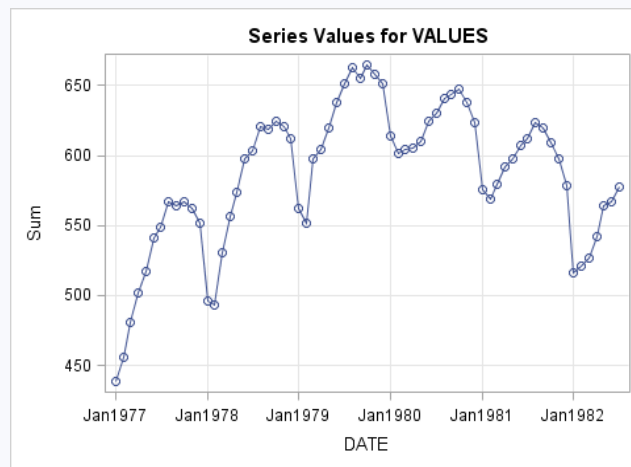
# Prediction Error Analysis

- Reconciliation Statistics
  - Computed for the predictions associated with the selected statistical time series model or the predictions associates with the reconciled predictions

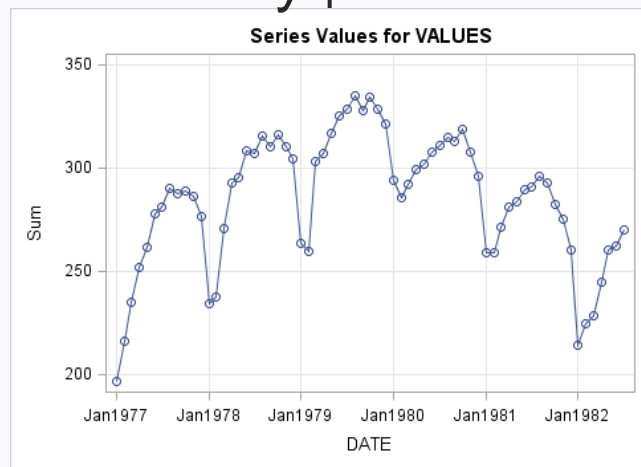


# Example Series

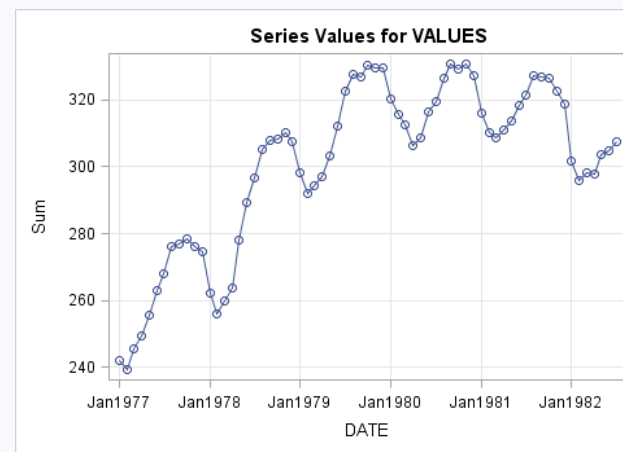
## Workers



## Masonry

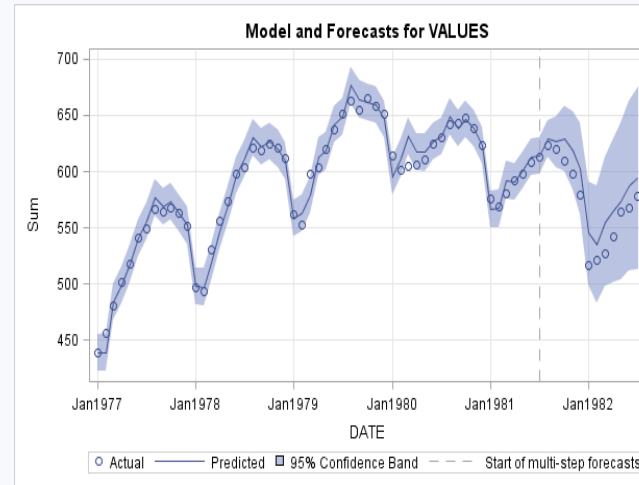


## Electric

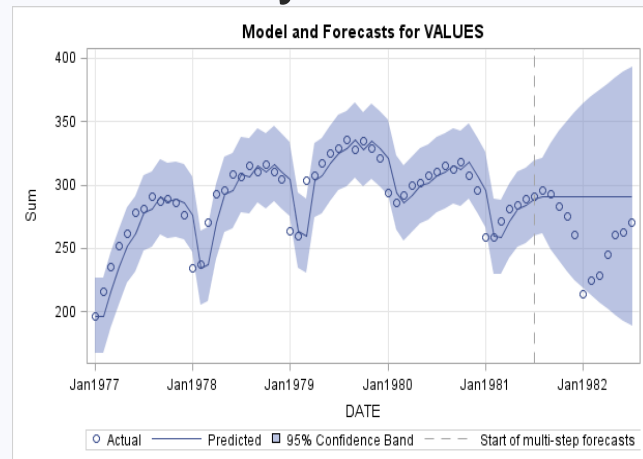


# Example Series ( $H > 0$ $B > 0$ )

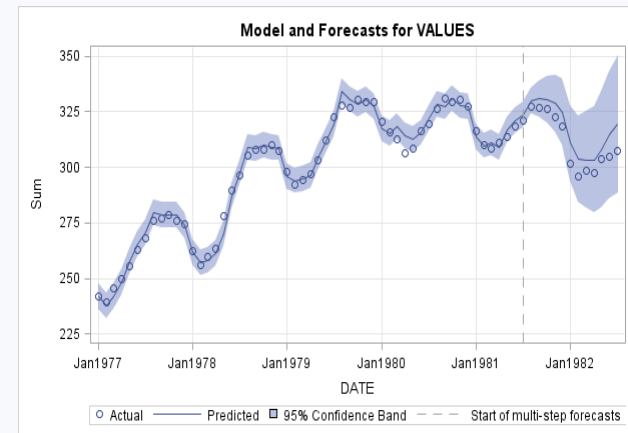
## Workers



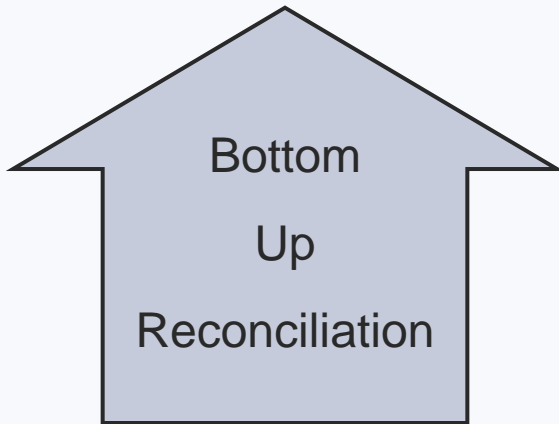
## Masonry



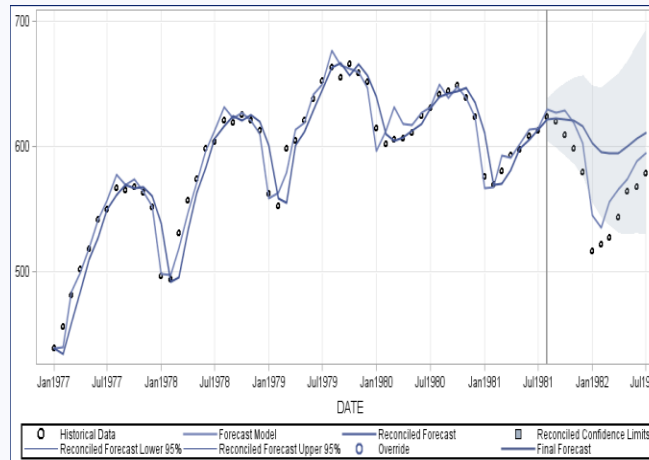
## Electric



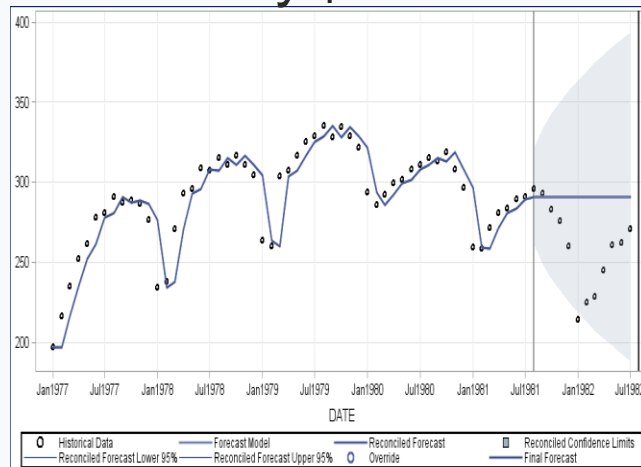
# Example Series ( $H > 0$ $B > 0$ )



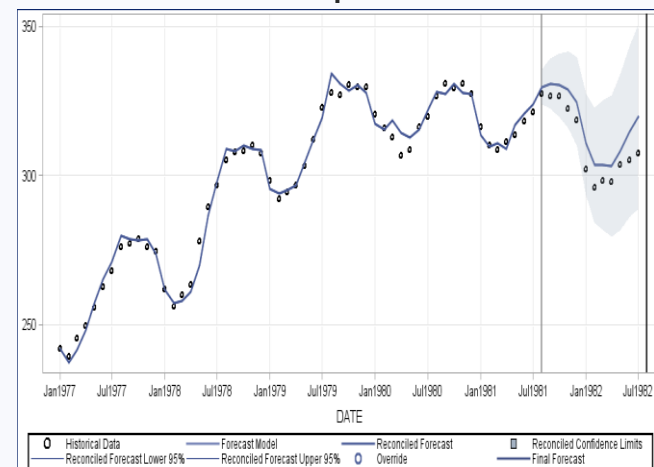
## Workers



## Masonry



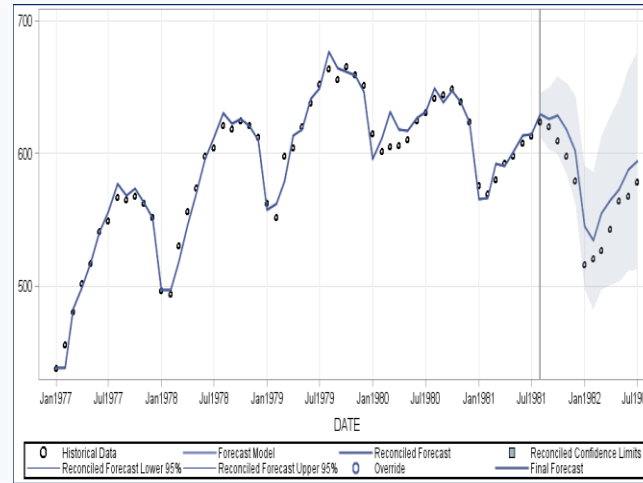
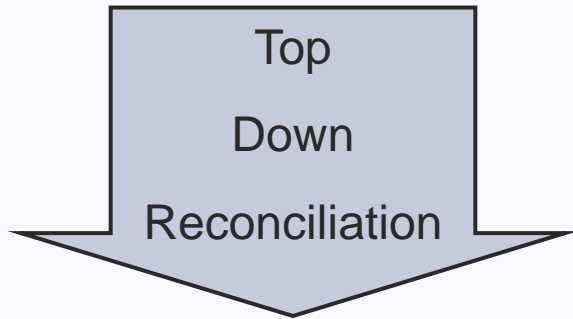
## Electric



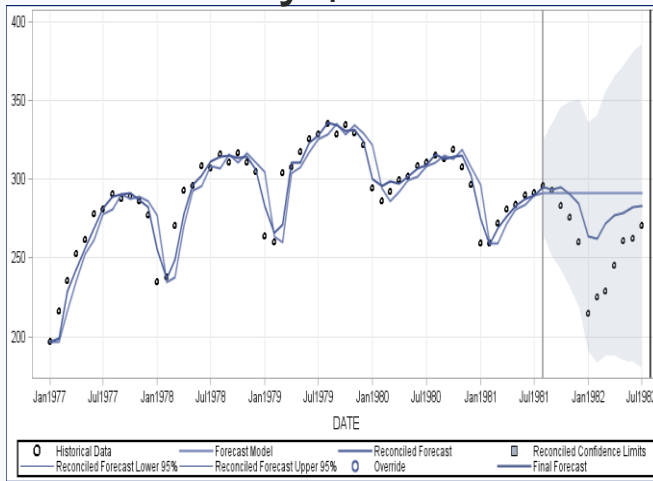


# Example Series ( $H > 0$ $B > 0$ )

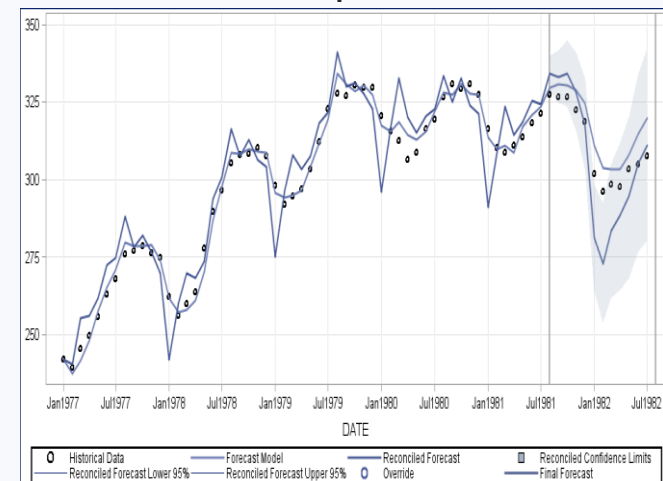
## Workers



## Masonry

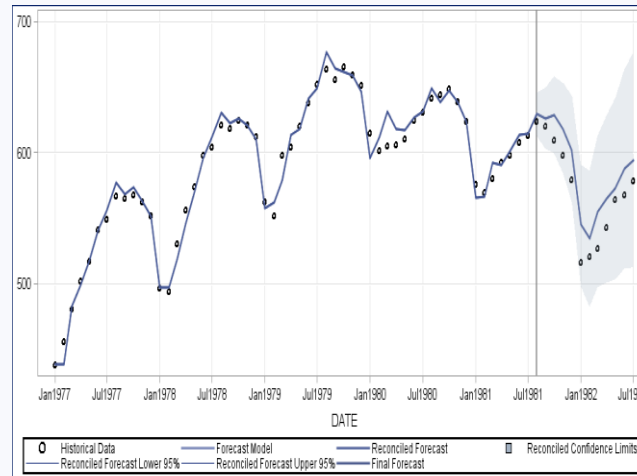
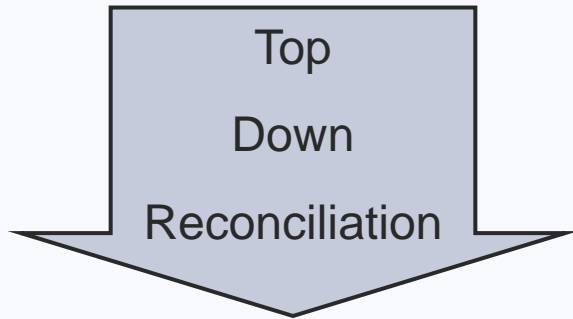


## Electric

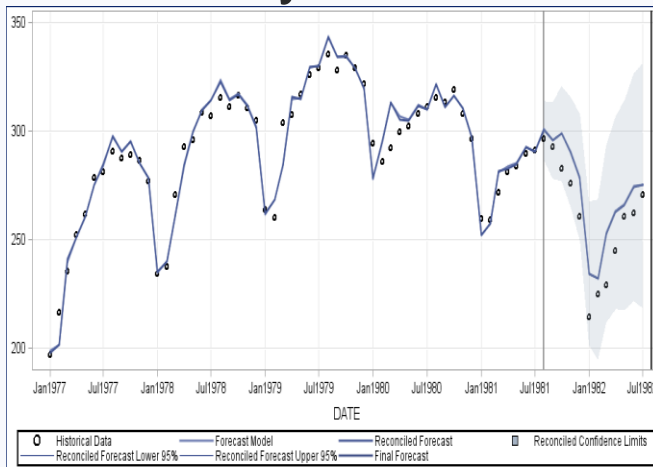


# Better Example ( $H > 0$ $B > 0$ )

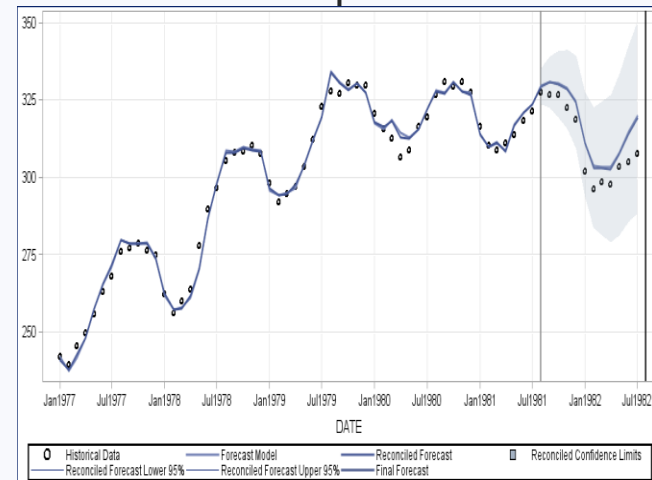
## Workers



## Masonry



## Electric



# Syntax: SAS High-Performance Forecasting

## ■ HOLDOUT=

- specifies the size of the holdout sample to be used for model selection.

## ■ BACK=

- specifies the number of observations before the end of the data that the multistep forecasts are to begin

## ■ Available in

- HPFDIAGNOSE (holdout, back)
- HPFSELECT (holdout)
- HPFENGINE(holdout, back)
- HPF (holdout, back)

# SAS Forecast Studio 3.1

**Forecasting Settings**

Time ID  
Data Preparation  
Diagnostics  
Model Generation  
**Model Selection**  
Forecast  
Reconciliation

**i** If changes are made to the settings below, models for each series will be automatically reselected. For each series, all models will be refitted, and a forecast model selected.

Use the following settings to select a forecast model for each series:

Use holdout sample for model selection:

Maximum percentage of series that holdout sample can be:

Selection criterion:

**Holdout Sample**

**Forecasting Settings**

Time ID  
Data Preparation  
Diagnostics  
Model Generation  
**Model Selection**  
**Forecast**  
Reconciliation

**i** If changes are made to the settings below, forecast models will be updated(parameters will not be changed).

Use the following settings to fit the forecast model for each series:

Number of periods to forecast (horizon):

Calculate statistics of fit over an out-of-sample range:

Confidence limit:

Allow negative forecasts

**Out-of-sample**

# Findings

- Hold-out strategies are useful for judging time series models automatically
- We need to distinguish in-sample and out-of-sample data
- Out-of-sample data allows for an unbiased performance evaluation of forecasting models
- Ultimately we want to compare against “future actual values” and re-adjust our models



2009 Washington, DC  
22-25 March 2009

