Intervention Analysis

Explanatory Variables

- In many cases, there are other factors (other than trend and seasonality) that are influencing our data series.
- We can sense these “signals” using analytics and measure their impact on our data to improve forecast accuracy.
  - HOWEVER, forecasting in this manner requires knowledge (or forecasts) of future values of these variables.

Interventions (Event Variable)

- An intervention variable is an indicator variable that contains discrete values that flag the occurrence of an event affecting the response series.
- Intervention variables can be used both to model and forecast the response series and also to analyze the impact of the intervention.
- Time series intervention analysis is used to ascertain the impact that one or more interventions have on a time series.
  - For example, the time series may be monthly revenues from the sale of a product with the intervention being the implementation of a sales promotion.

Improve Accuracy?

- Event variables enable the forecast model to accommodate **discrete shifts**, also called jumps or bangs, in time series data.
- Event variables in time series models are primarily **intercept shifters**.
- Intercept shifters are included in the model as **explanatory variables** and are based on columns of zeros and ones in the data set.

3 Types of Intervention Variables

- There are three common intervention variables:
  1. Point (or Pulse) Interventions
  2. Step Interventions
  3. Ramp Interventions

3 Types of Intervention Variables

- There are three common intervention variables:
  1. Point (or Pulse) Interventions
  2. Step Interventions
  3. Ramp Interventions

Each of these could be **deterministic** OR **stochastic**.

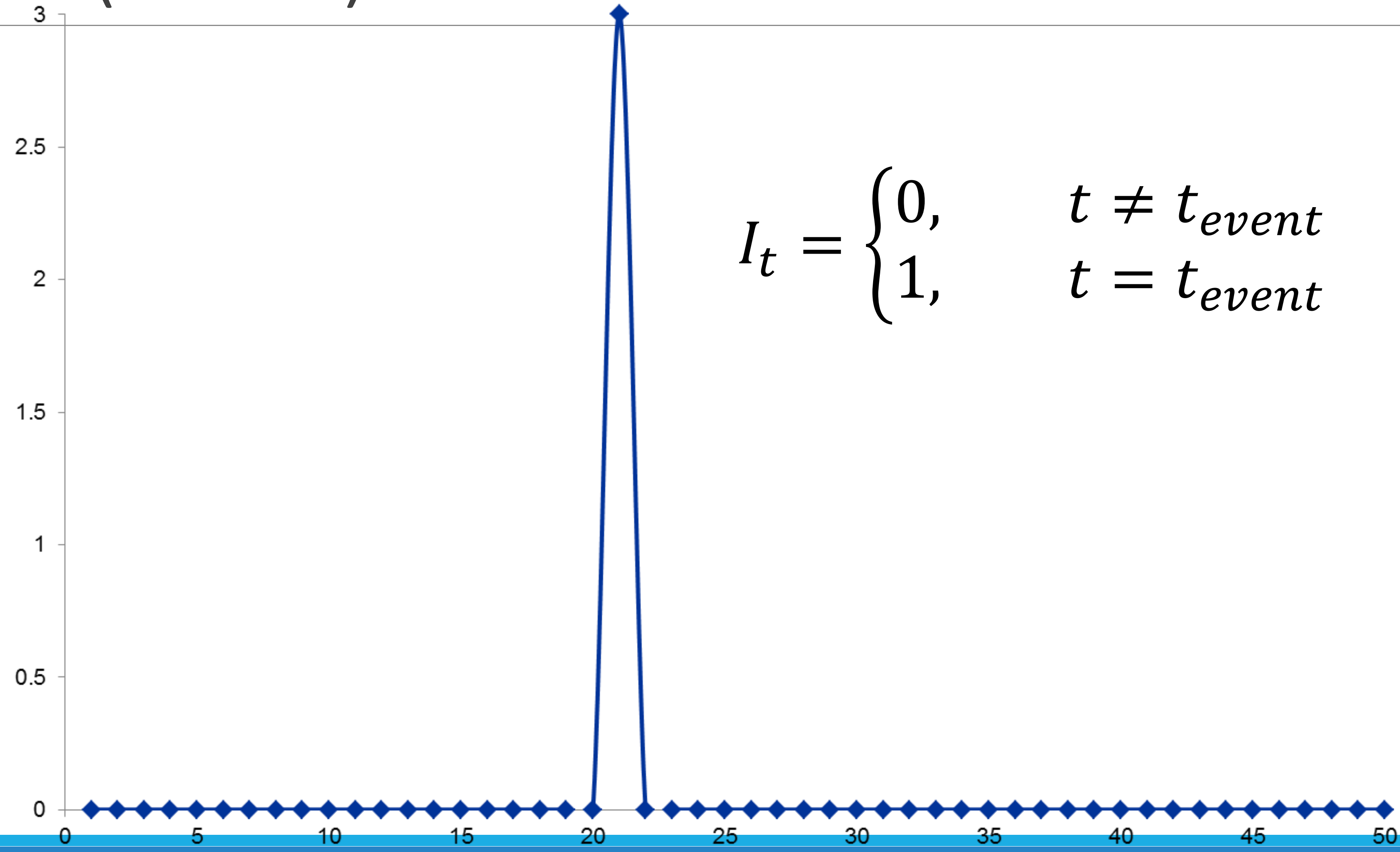
Point (Pulse) Intervention

Sales Promotion Event
T*= '01JAN2002'

Demand History For Sales	**Resulting Dummy Column = D**
01JAN2001	0
01FEB2001	0
…	…
01DEC2001	0
01JAN2002	1
01FEB2002	0
…	…
01JAN2003	0

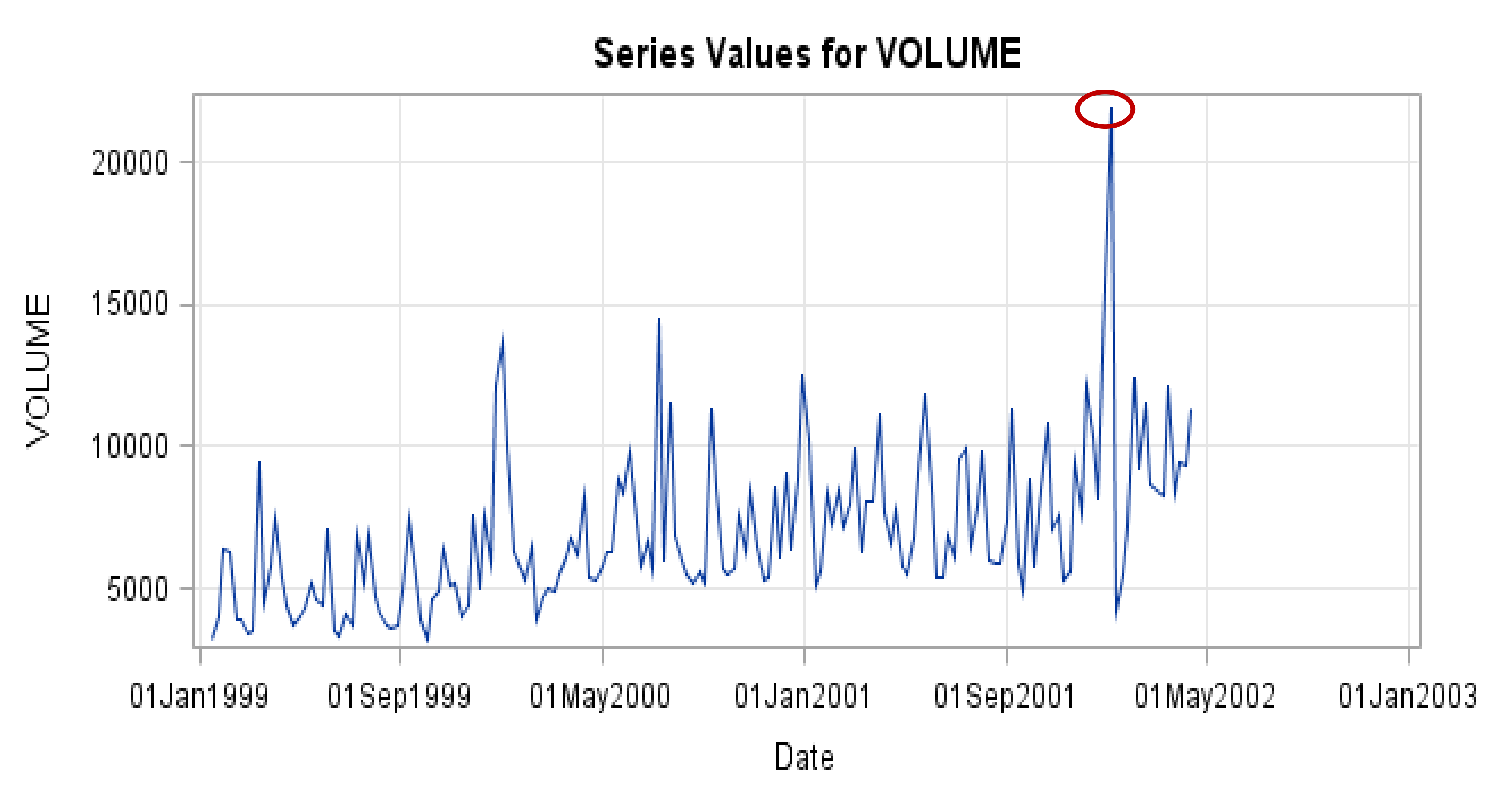
The temporary intercept shift is accomplished by adding a zero-one or dummy column to the data.

Point (Pulse) Intervention



$$I_t = \begin{cases} 0, & t \neq t_{event} \\ 1, & t = t_{event} \end{cases}$$

Point (Pulse) Intervention – Example



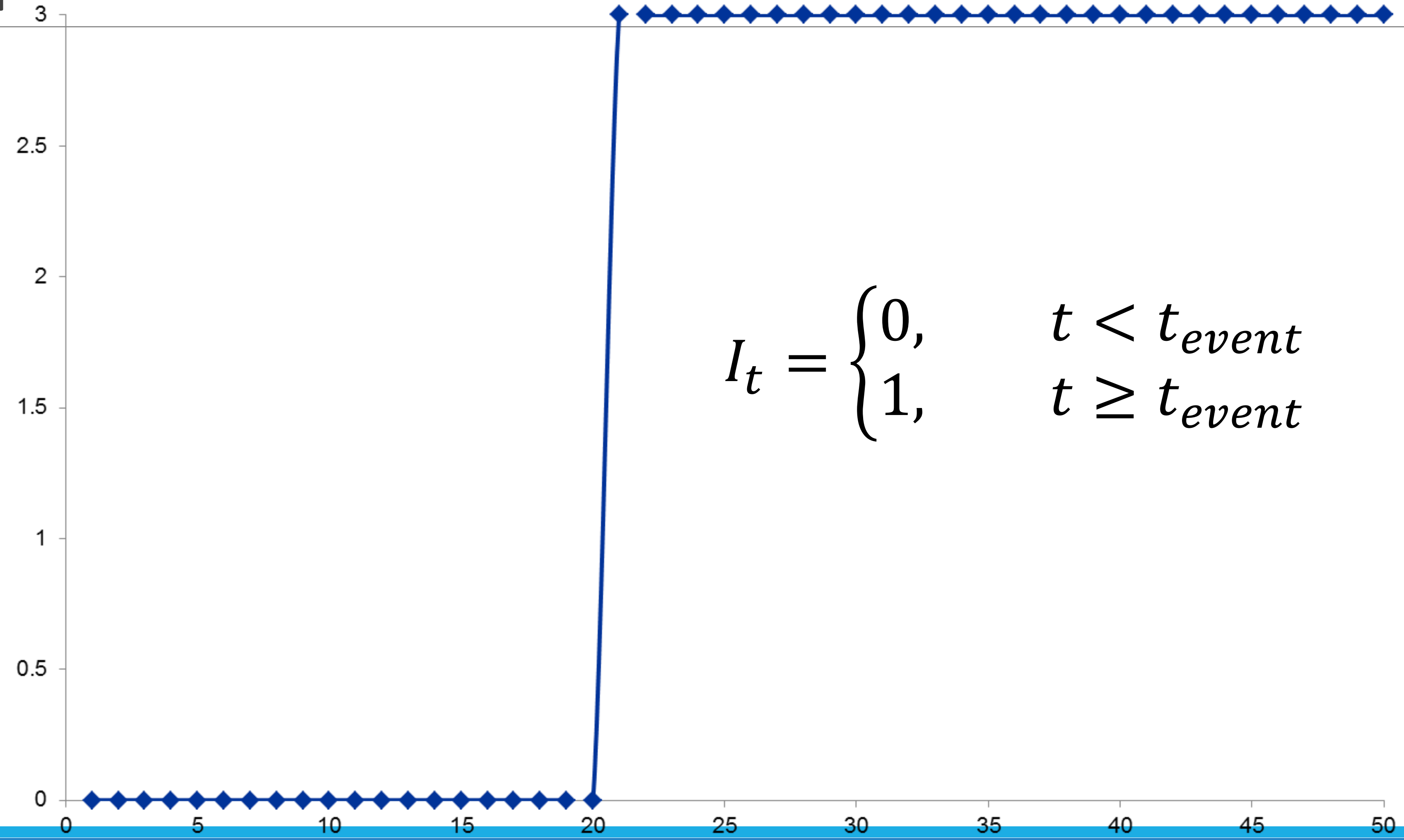
Step Intervention

New Law Enacted
T*= ’01AUG2003’d

Demand History For Sales	**Resulting Dummy Column = D**
01JAN2002	0
01FEB2002	0
…	…
01JUL2003	0
01AUG2003	1
01SEP2003	1
…	…
01JUN2004	1

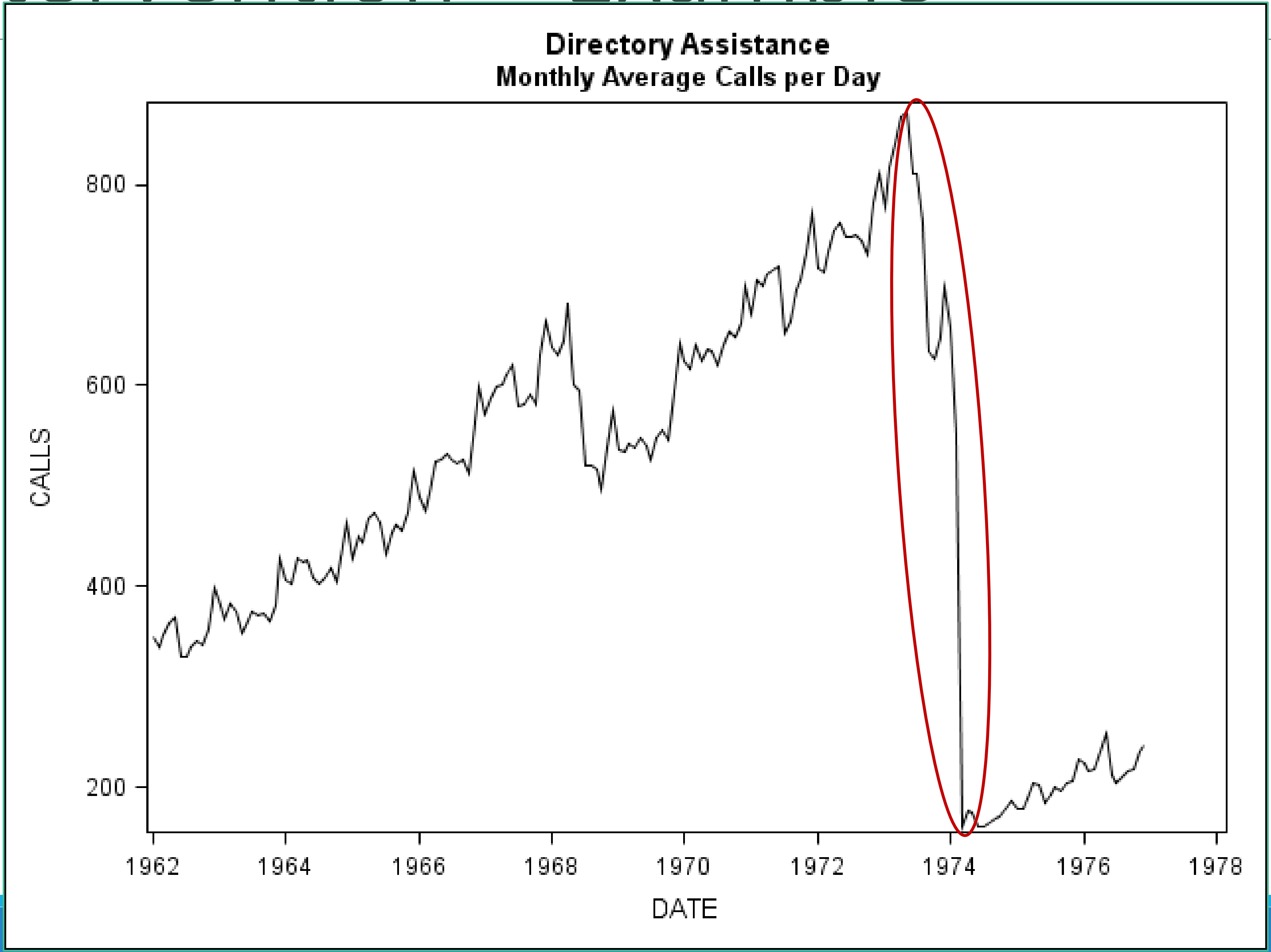
The permanent intercept shift is accomplished by adding a zero-one or dummy column to the data.

Step Intervention



$$I_t = \begin{cases} 0, & t < t_{event} \\ 1, & t \geq t_{event} \end{cases}$$

Step Intervention – Example



Ramp Intervention

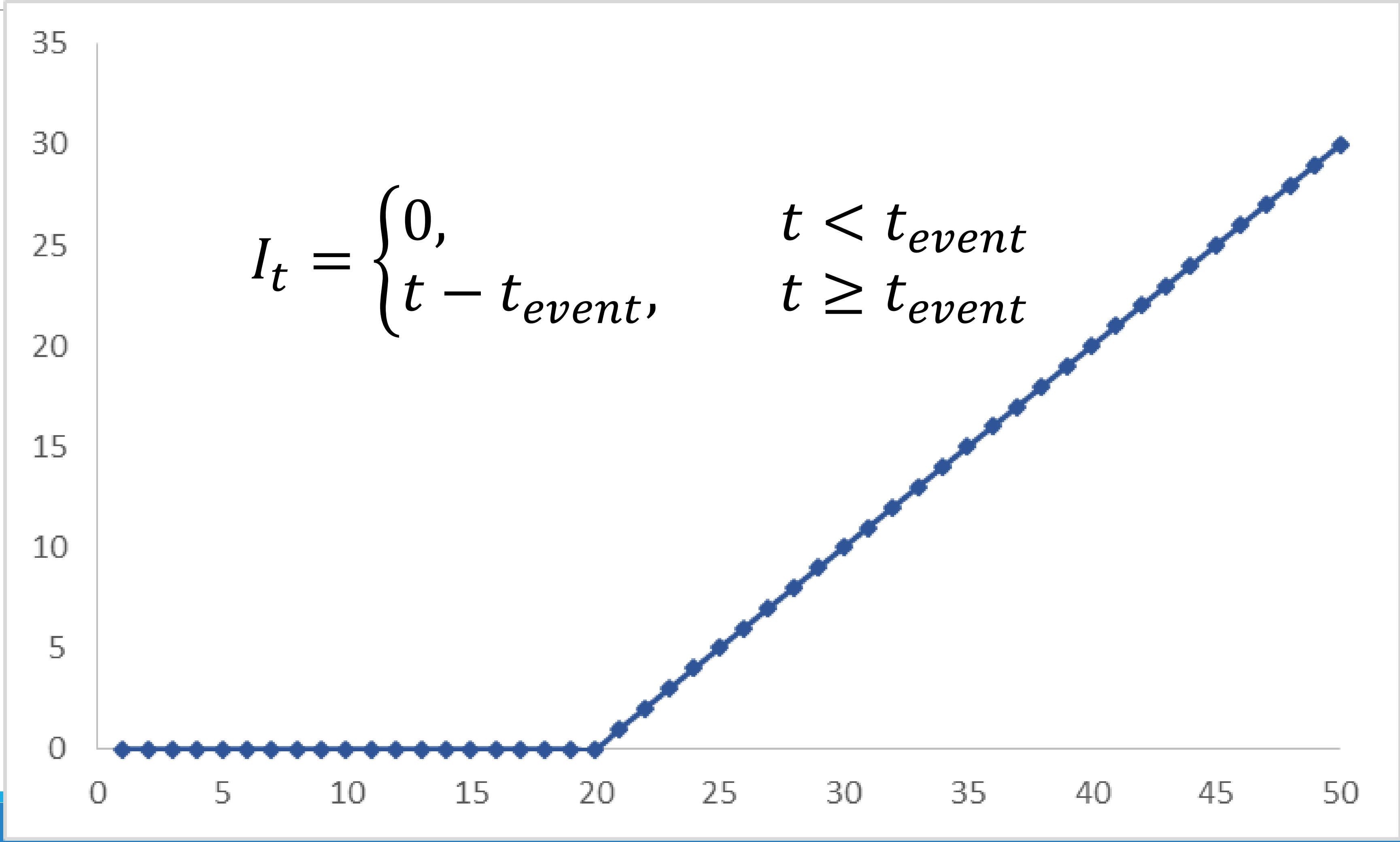
New Law Enacted
T*= '01AUG2003'd

Demand History For Sales	**Resulting Dummy Column = D**
01JAN2002	0
01FEB2002	0
…	…
01JUL2003	0
01AUG2003	1
01SEP2003	2
…	…
01JUN2004	11

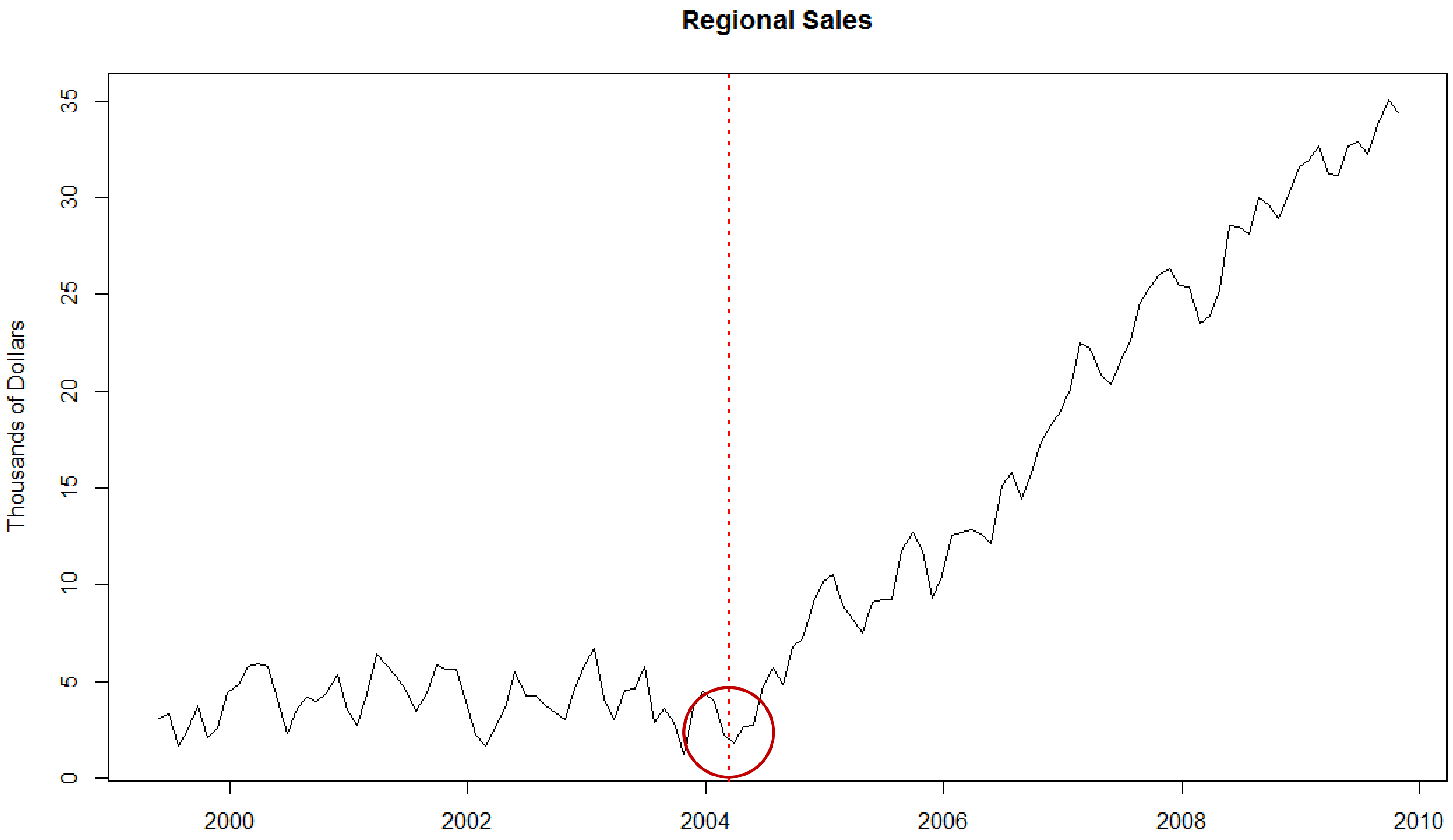
The permanent intercept shift is accomplished by adding a variable that trends after a given time.

Ramp Intervention

$$I_t = \begin{cases} 0, & t < t_{event} \\ t - t_{event}, & t \geq t_{event} \end{cases}$$



Ramp Intervention – Example

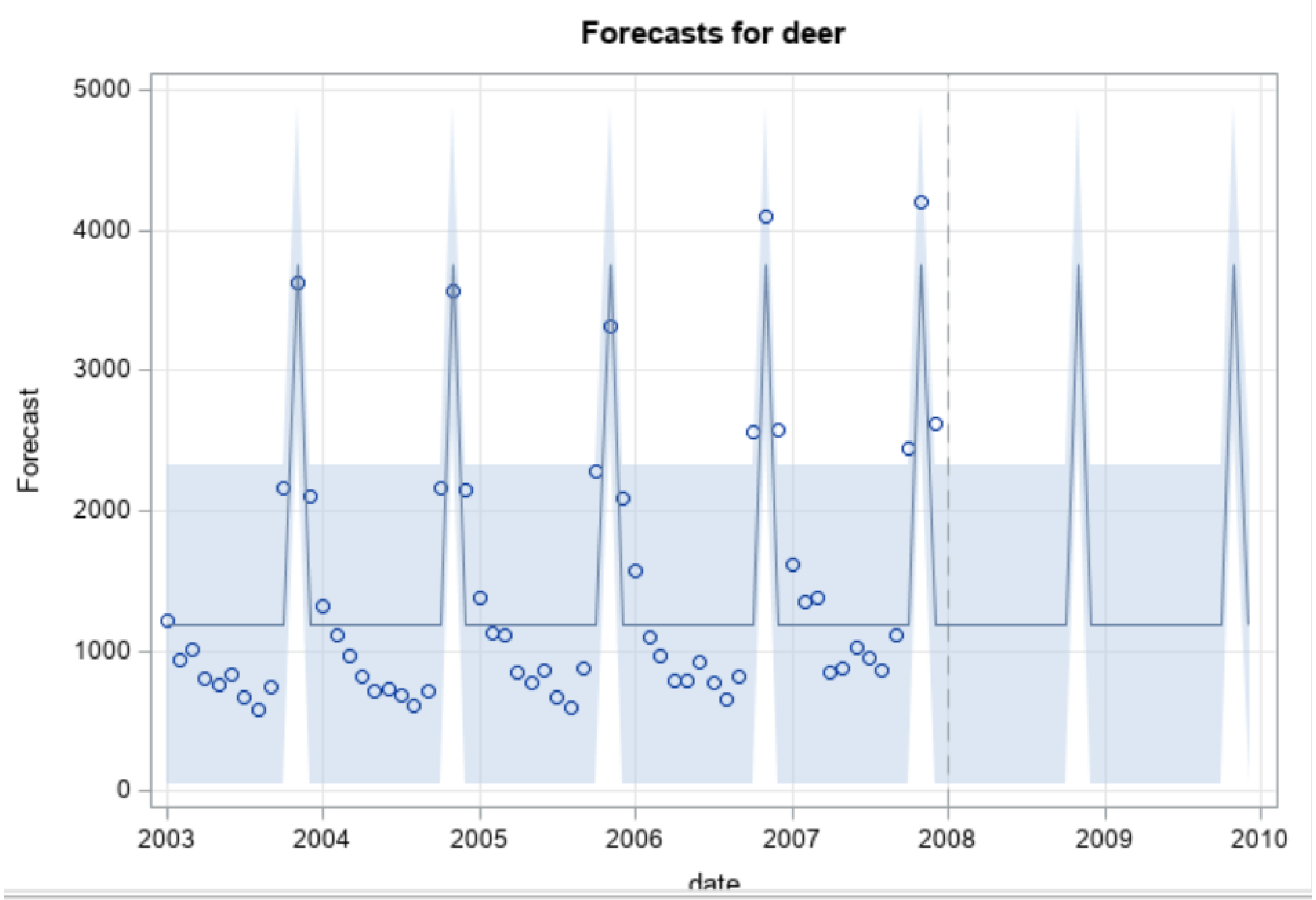


Deer Crash Data

- Deer mating season occurs each year around November.
- This results in a higher number of deer related car accidents during this time.
- Contains number of monthly deer related car wrecks from January 2003 to December 2007.

Point Intervention – Deterministic

```
proc arima data=Time.DEER2 plot(unpack)=(series(corr)
                                              forecast(all));
     identify var=deer nlag=24 crosscorr=(Nov);
     estimate input=(Nov);
     forecast lead=24 id=date interval=month;
run;
quit;
```



Forecasts for deer
Forecast
5000
4000
3000
2000
1000
0
2003
2004
2005
2006
2007
2008
2009
2010
date

Point Intervention – Deterministic

```
Deer.Model1 <- Arima(Deer.Accidents, xreg=NOV,
                     method="ML")
summary(Deer.Model1)
```

Deterministic vs. Stochastic

- There are many different structures each of these intervention points can take.
  - Previous examples were all **deterministic** – the effect had no lag structure to it.
- This intervention effect could dissipate – **stochastic** structure.
  - The intervention could have a lag structure to how it dissipates across time.

Stochastic Interventions

- These models can be evaluated with different lag structures applied to the intervention itself:

$$Y_t = \beta_0 + \frac{\omega(\mathrm{B})}{\delta(\mathrm{B})} I_t + Z_t$$

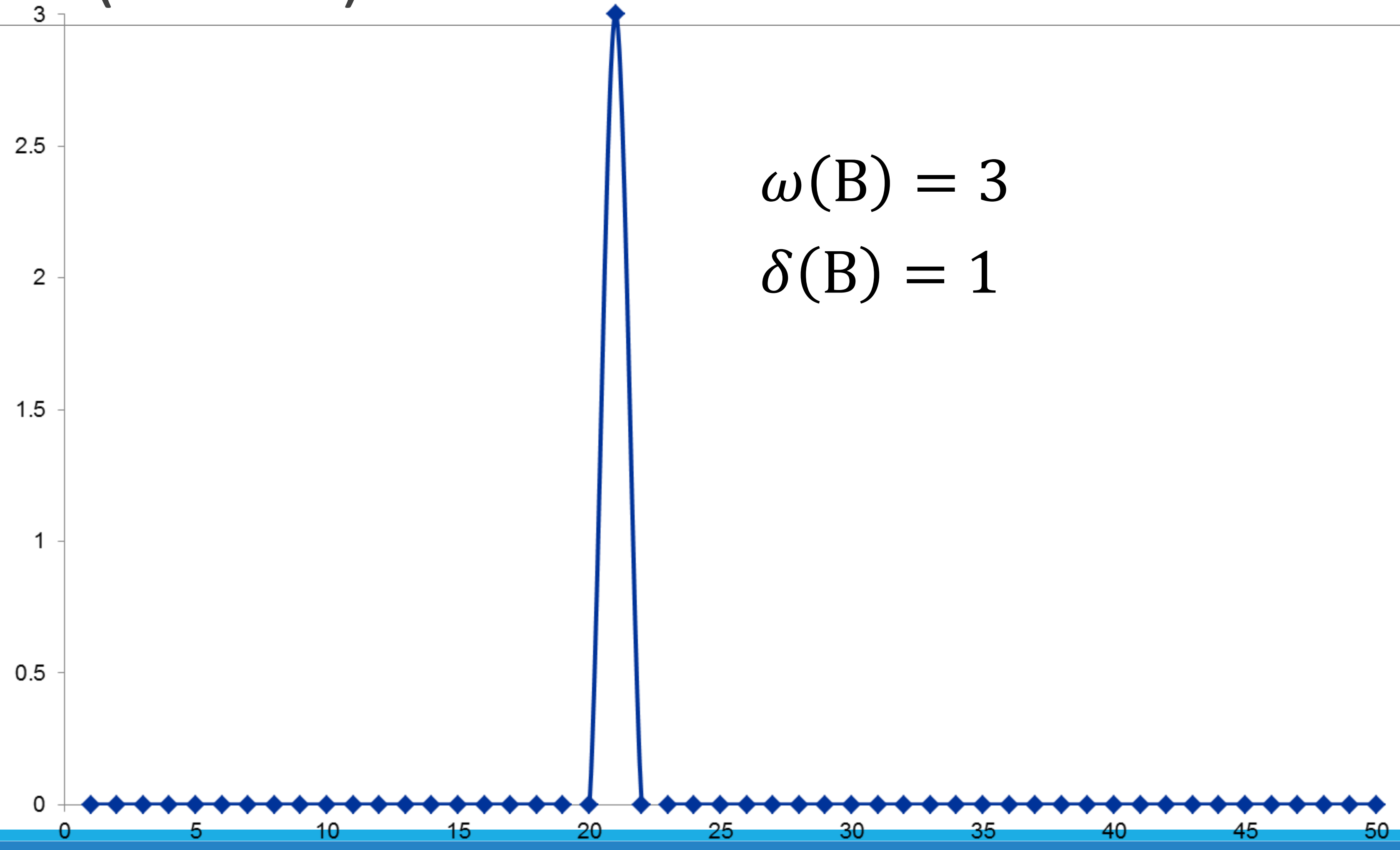
- There is the numerator term, $\omega(\mathrm{B})$, and the denominator term, $\delta(\mathrm{B})$, that help to determine how the intervention point influences the data.

Numerator vs. Denominator

$$Y_t = \beta_0 + \frac{\omega(\mathrm{B})}{\delta(\mathrm{B})} I_t + Z_t$$

- The numerator term provides information on number of lags of the Intervention variable that will be used.
- The denominator term is used when there is a “decay” effect.

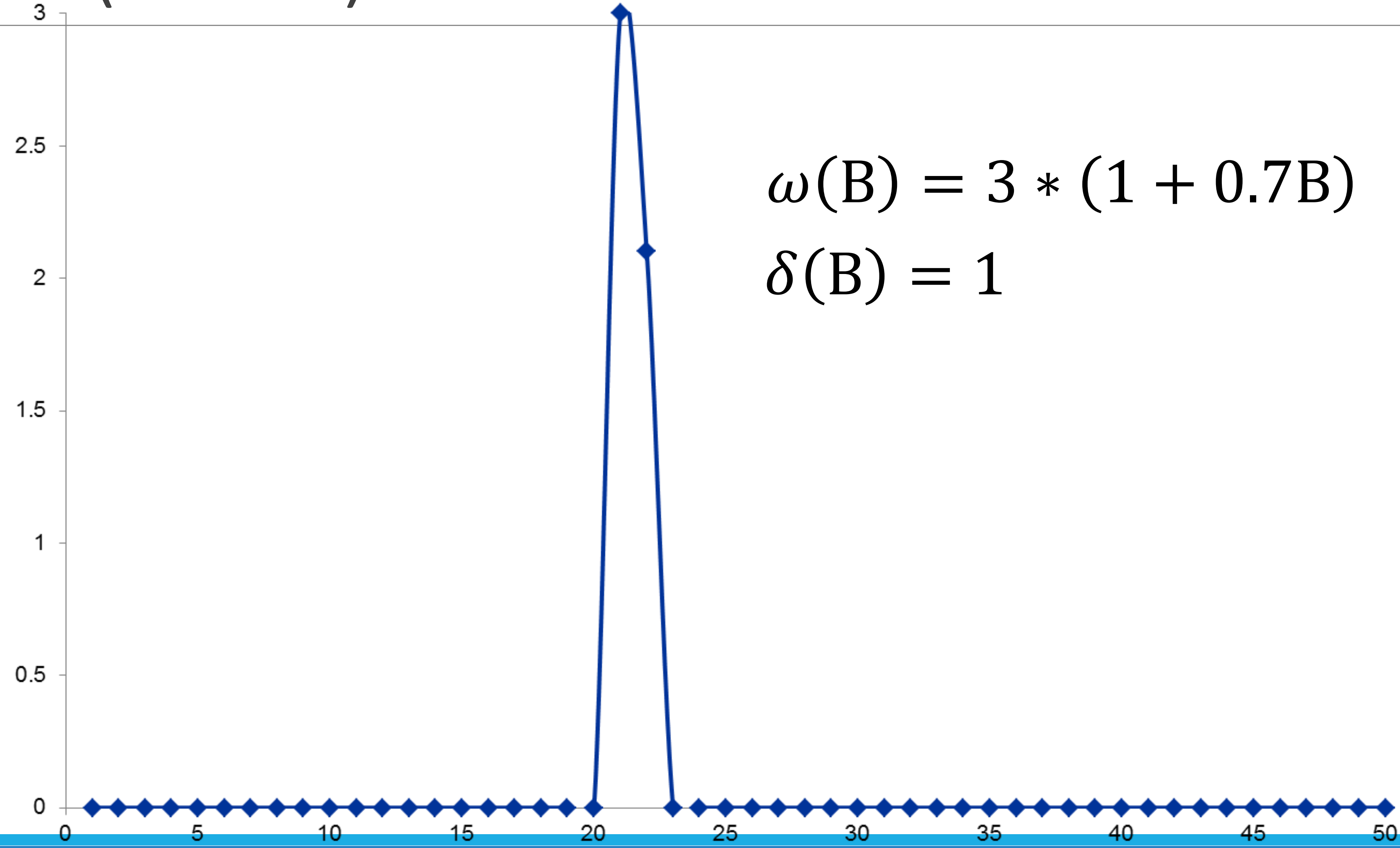
Point (Pulse) Interventions



$$\omega(\mathrm{B}) = 3$$

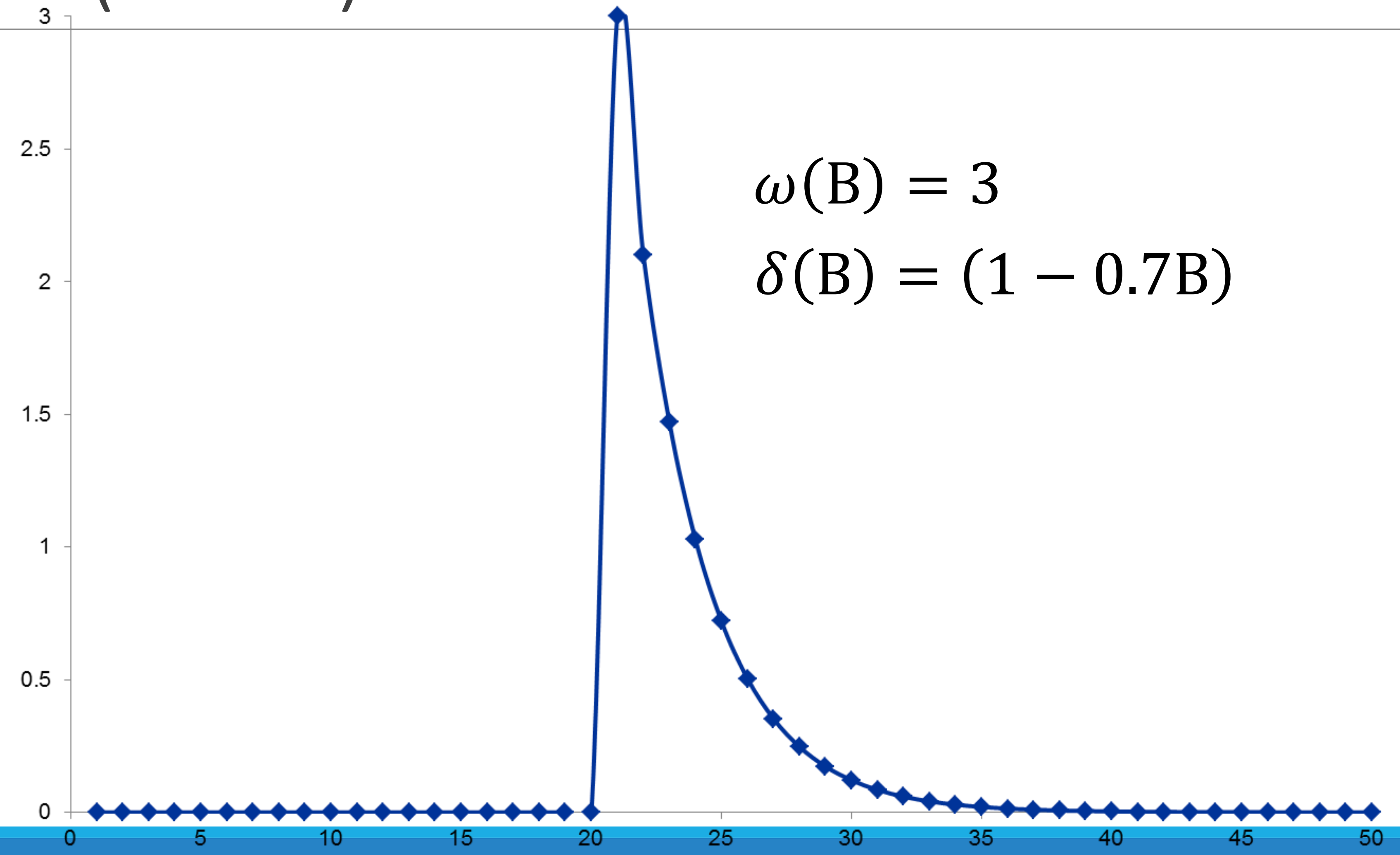
$$\delta(\mathrm{B}) = 1$$

Point (Pulse) Interventions



$$\omega(\mathrm{B}) = 3 * (1 + 0.7\mathrm{B})$$
$$\delta(\mathrm{B}) = 1$$

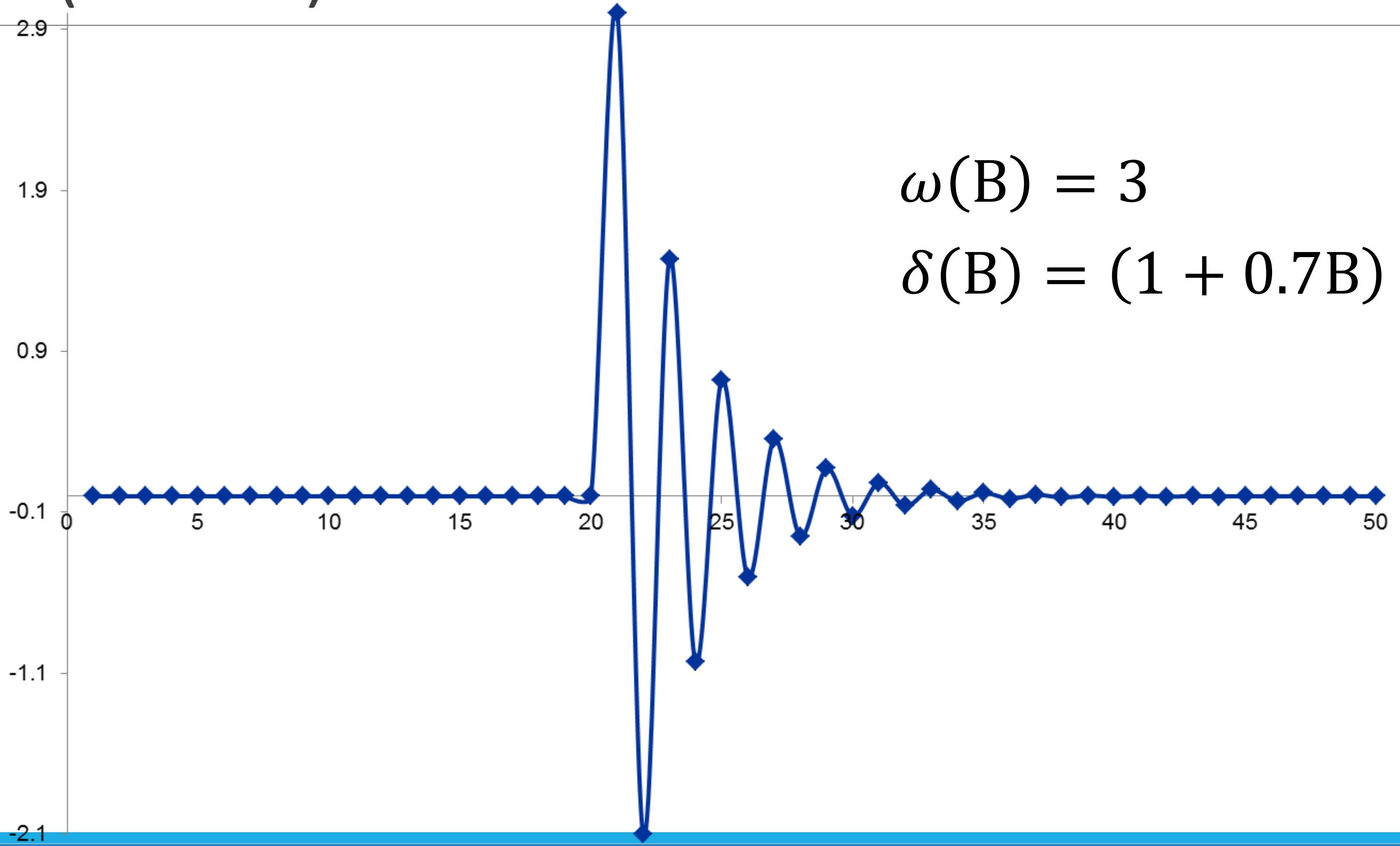
Point (Pulse) Interventions



$$\omega(\mathrm{B}) = 3$$

$$\delta(\mathrm{B}) = (1 - 0.7\mathrm{B})$$

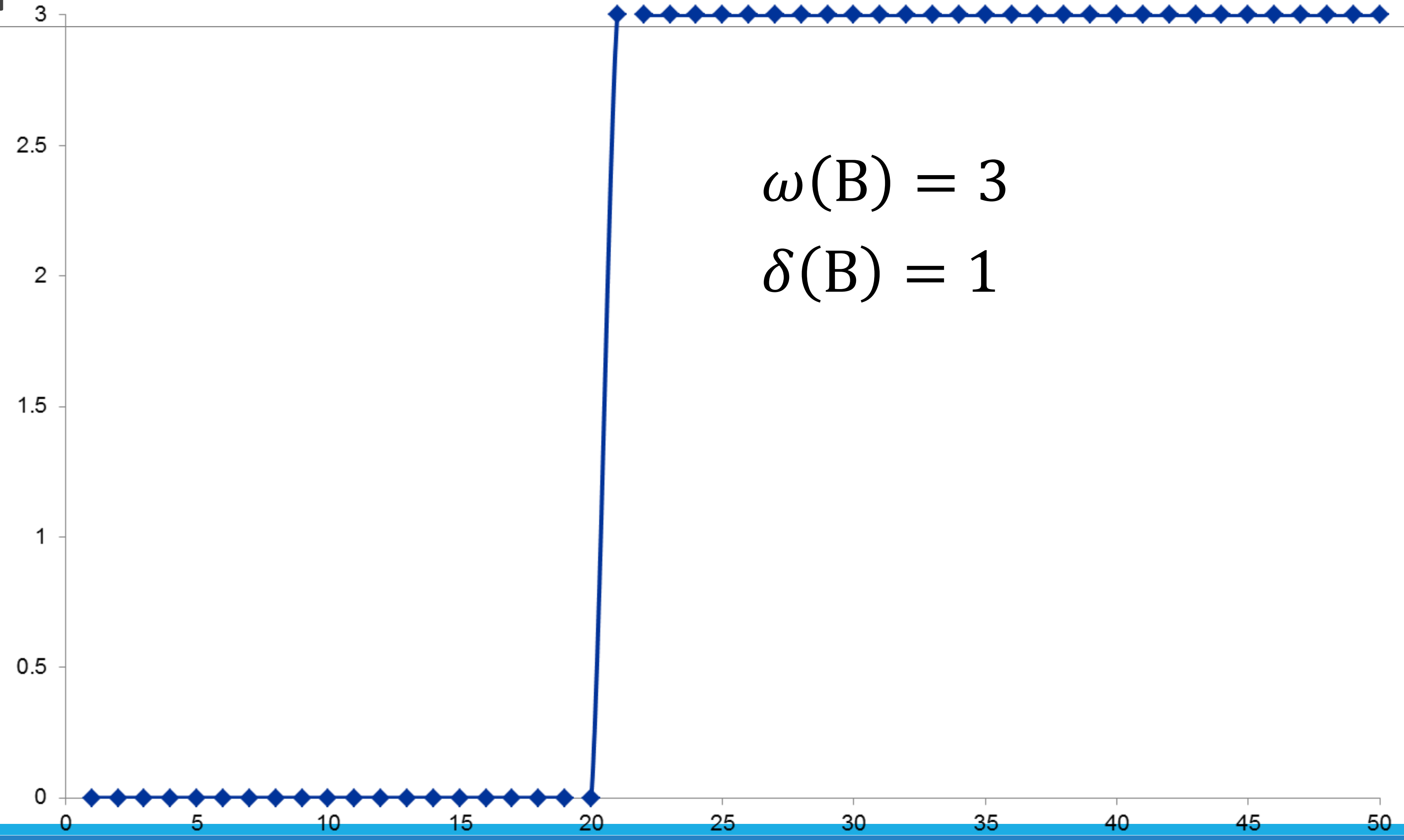
Point (Pulse) Interventions



$$\omega(\mathrm{B}) = 3$$

$$\delta(\mathrm{B}) = (1 + 0.7\mathrm{B})$$

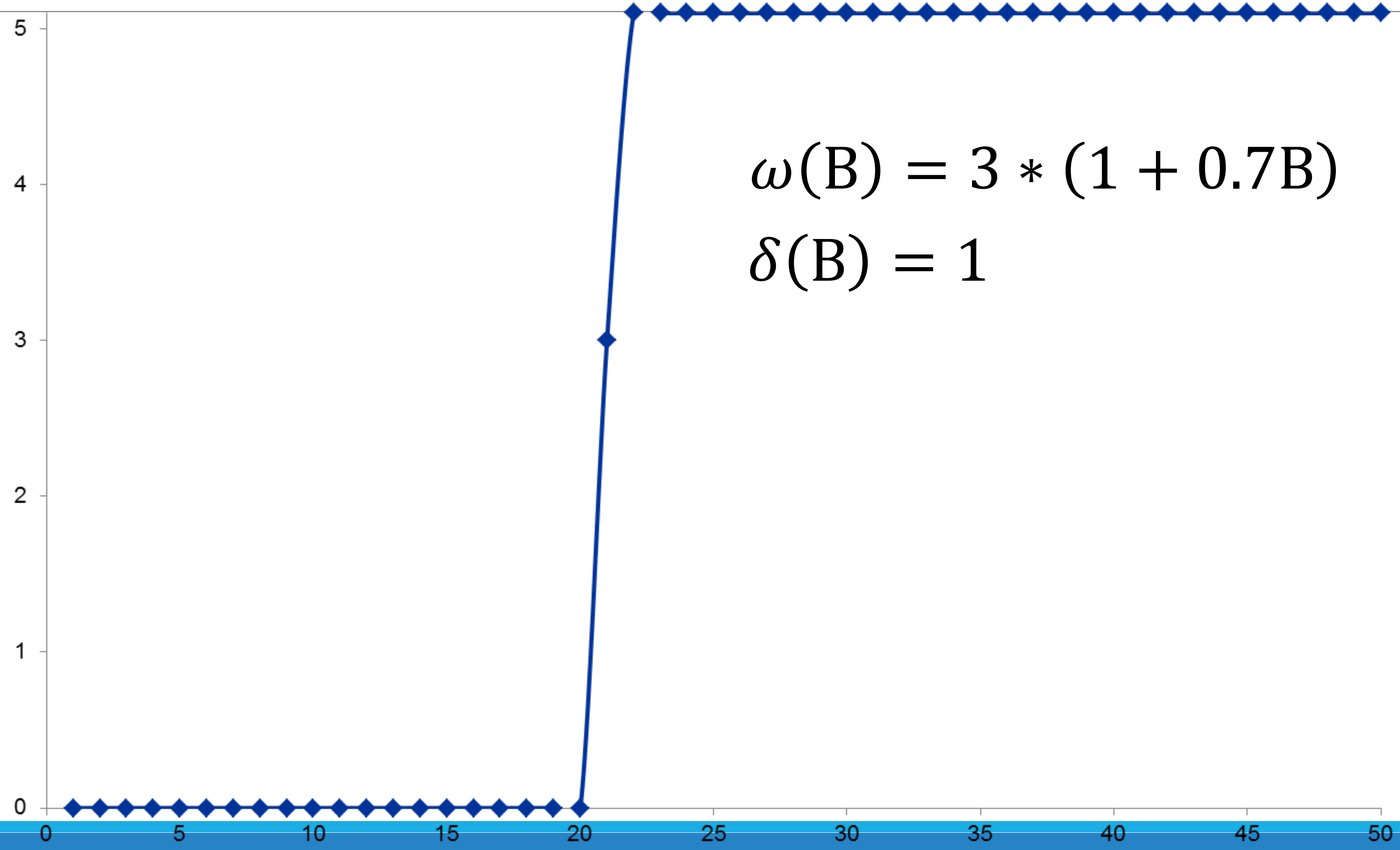
Step Interventions



$$\omega(\mathrm{B}) = 3$$

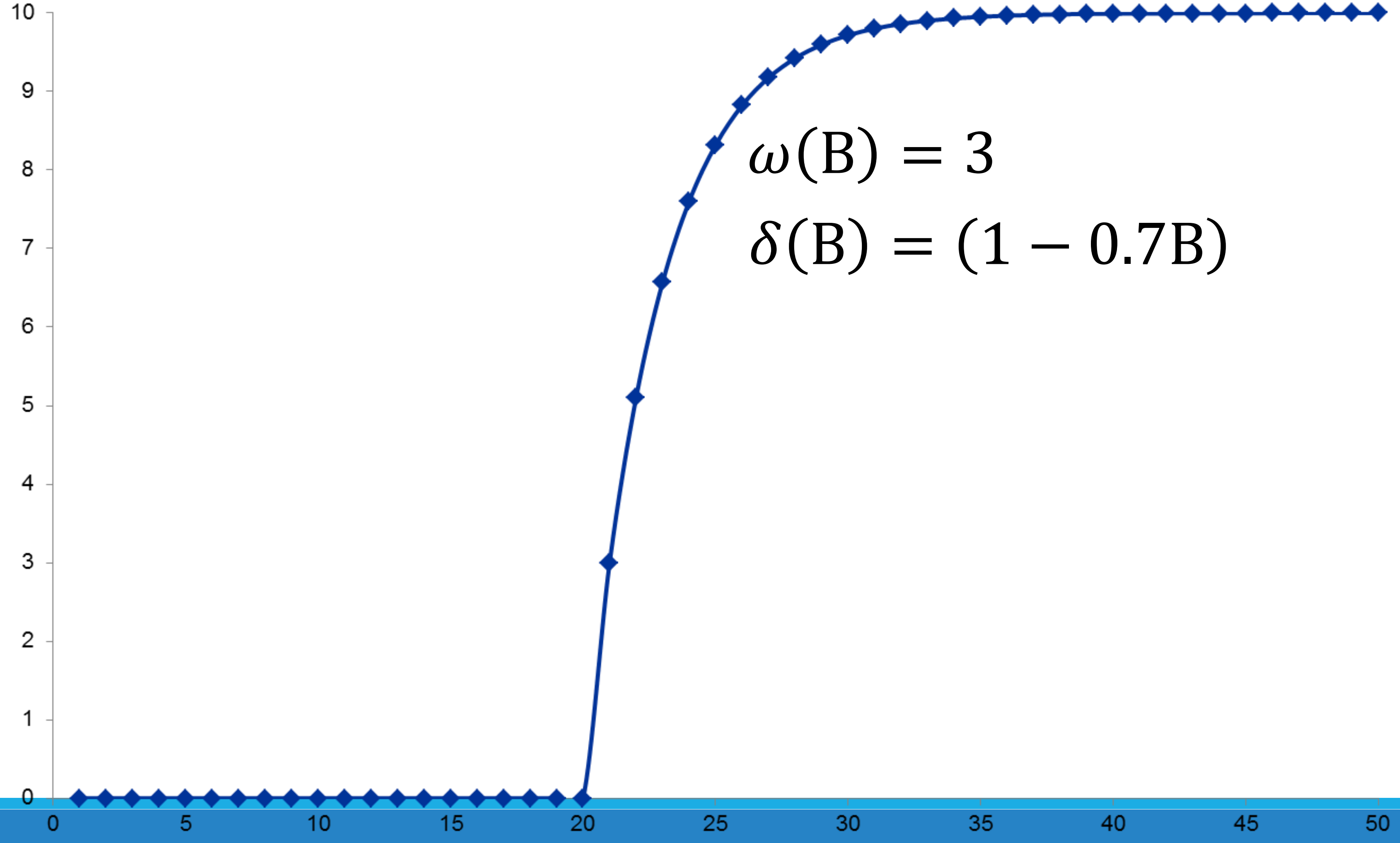
$$\delta(\mathrm{B}) = 1$$

Step Interventions

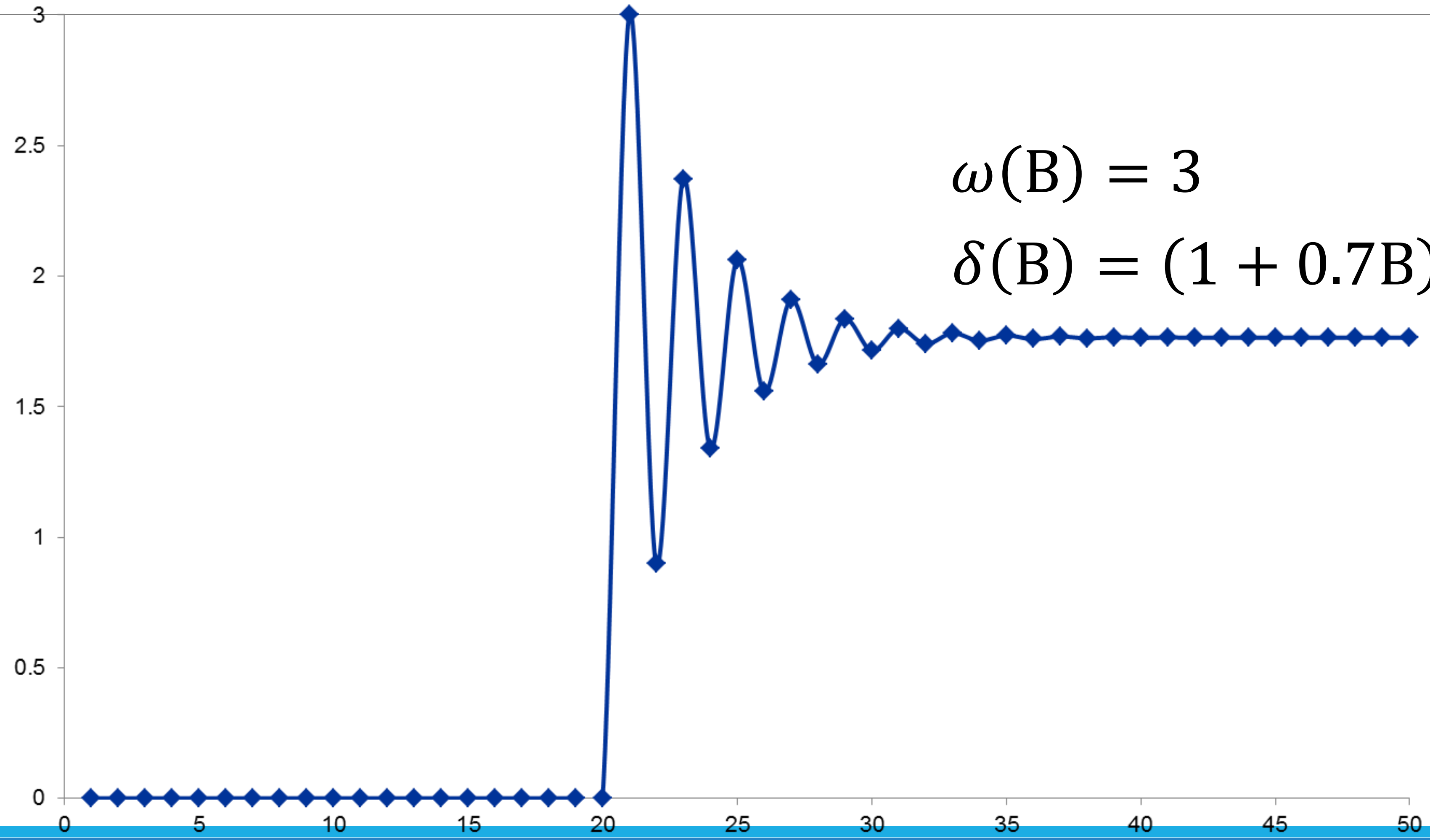


$$\omega(\mathrm{B}) = 3 * (1 + 0.7\mathrm{B})$$
$$\delta(\mathrm{B}) = 1$$

Step Interventions



Step Interventions

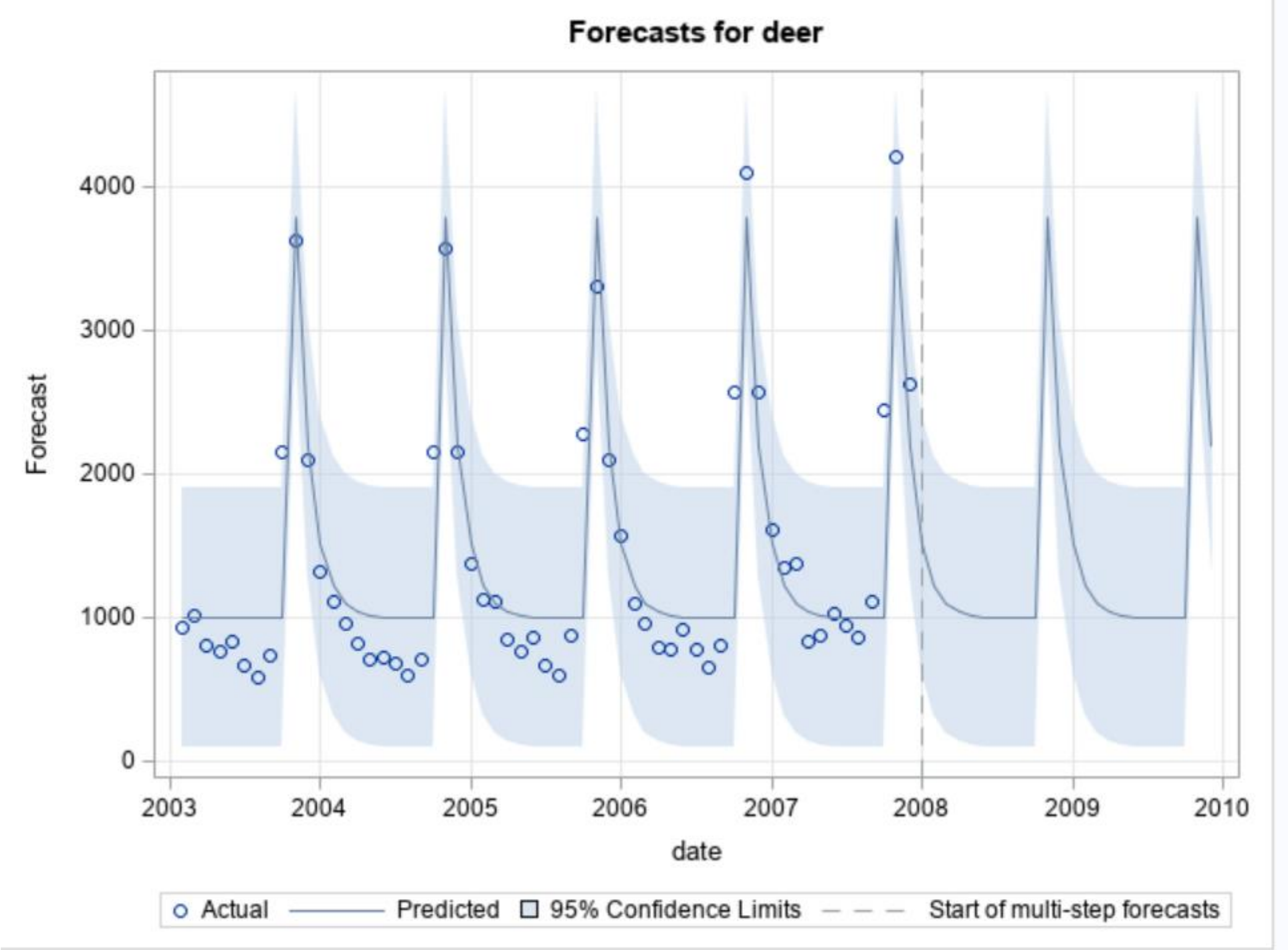


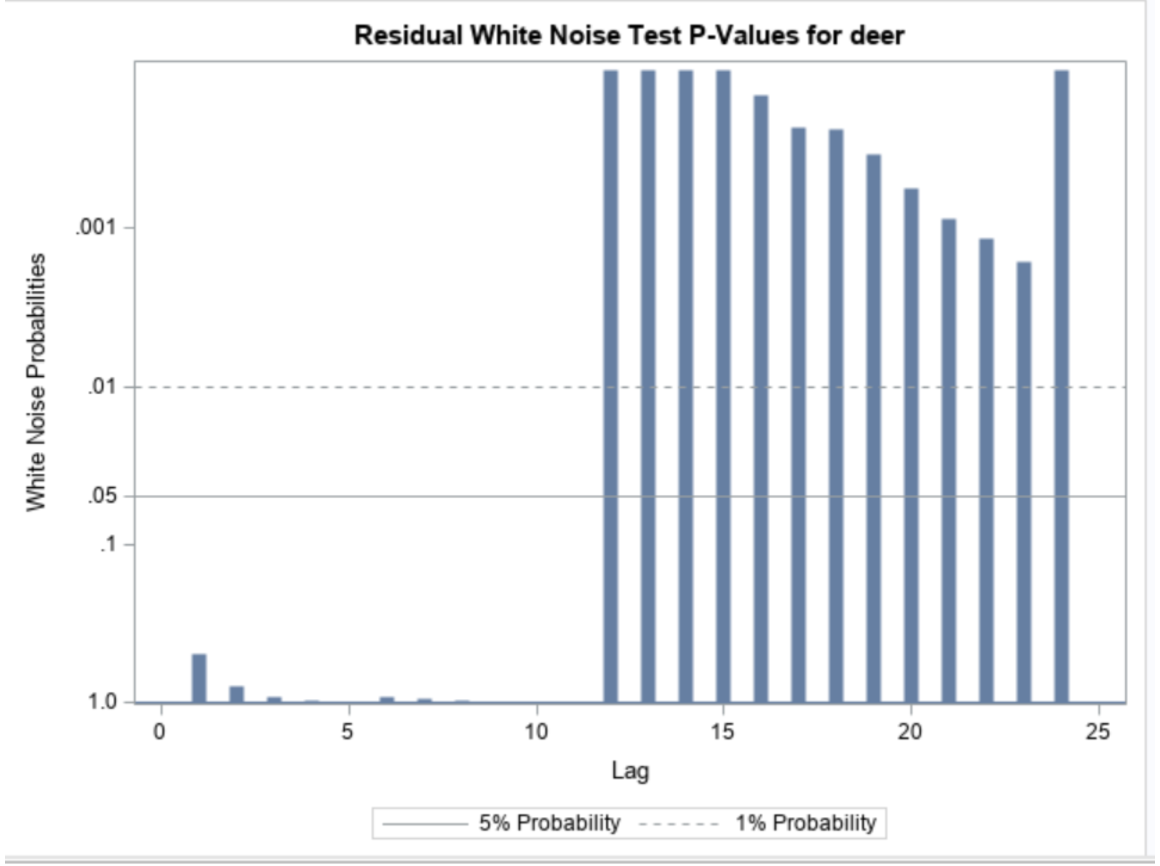
$$\omega(\mathrm{B}) = 3$$

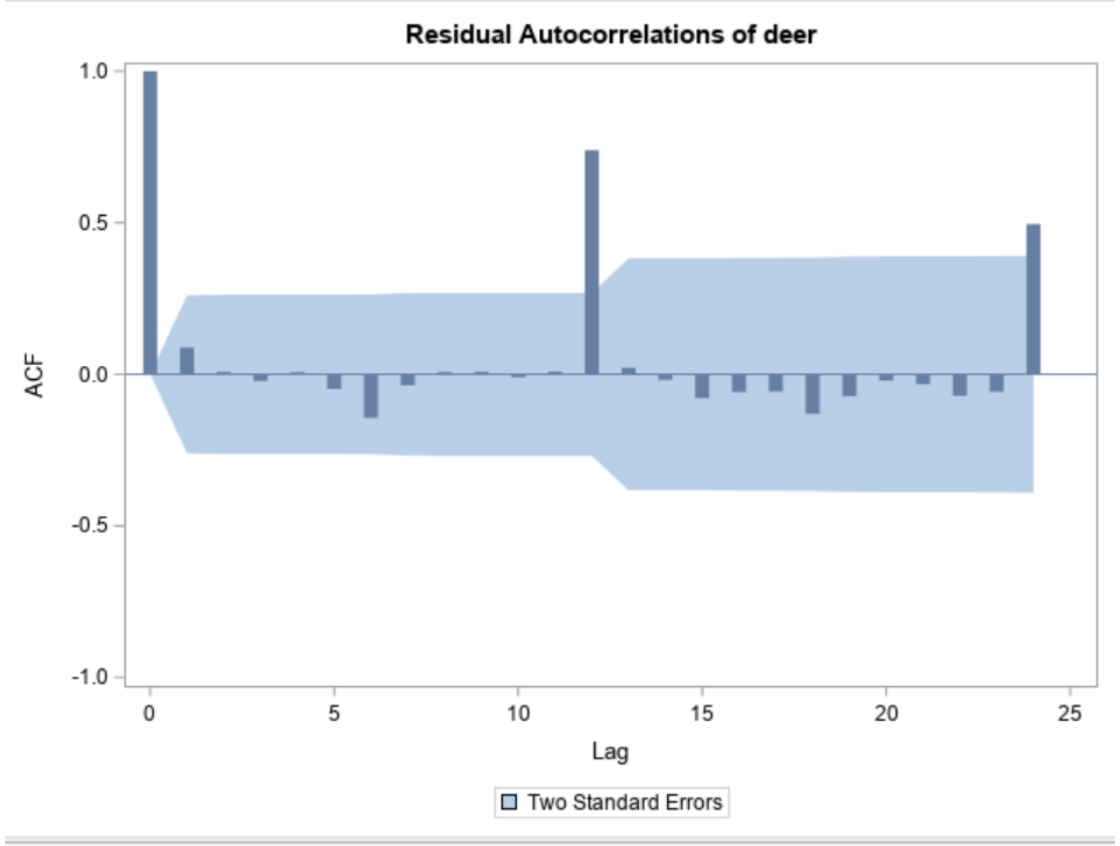
$$\delta(\mathrm{B}) = (1 + 0.7\mathrm{B})$$

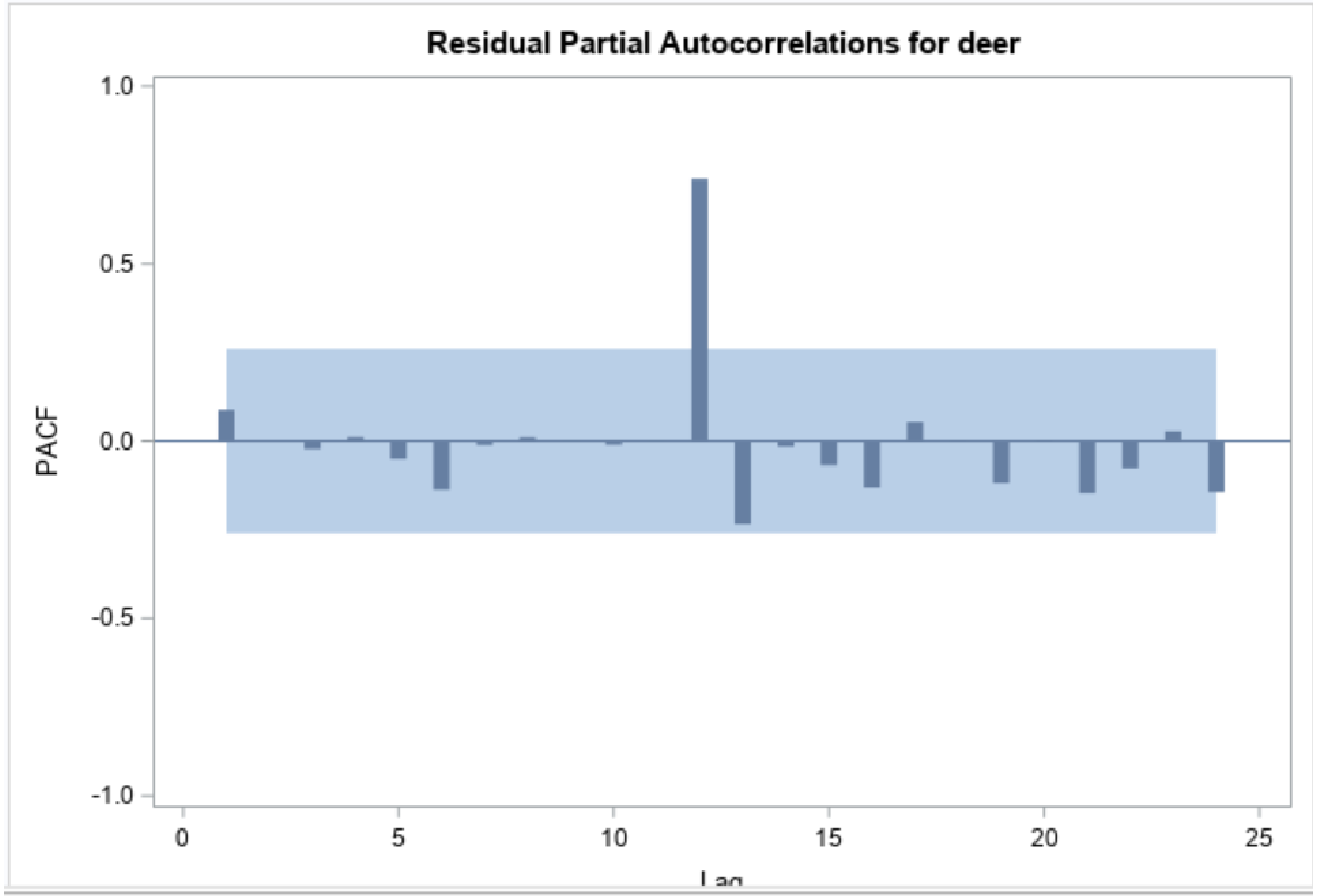
Point Intervention – Stochastic

```
proc arima data=Time.DEER2 plot(unpack)=(series(corr)
                                        forecast(all));
     identify var=deer nlag=24 crosscorr=(Nov);
     estimate input=(/(1)Nov);
     forecast lead=24 id=date interval=month;
run;
quit;
```


Forecasts for deer
Forecast
4000
3000
2000
1000
0
2003
2004
2005
2006
2007
2008
2009
2010
date
Actual
Predicted
95% Confidence Limits
Start of multi-step forecasts


Residual White Noise Test P-Values for deer
White Noise Probabilities
.001
.01
.05
.1
1.0
0
5
10
15
20
25
Lag
5% Probability
1% Probability


Residual Autocorrelations of deer
ACF
1.0
0.5
0.0
-0.5
-1.0
0
5
10
15
20
25
Lag
Two Standard Errors

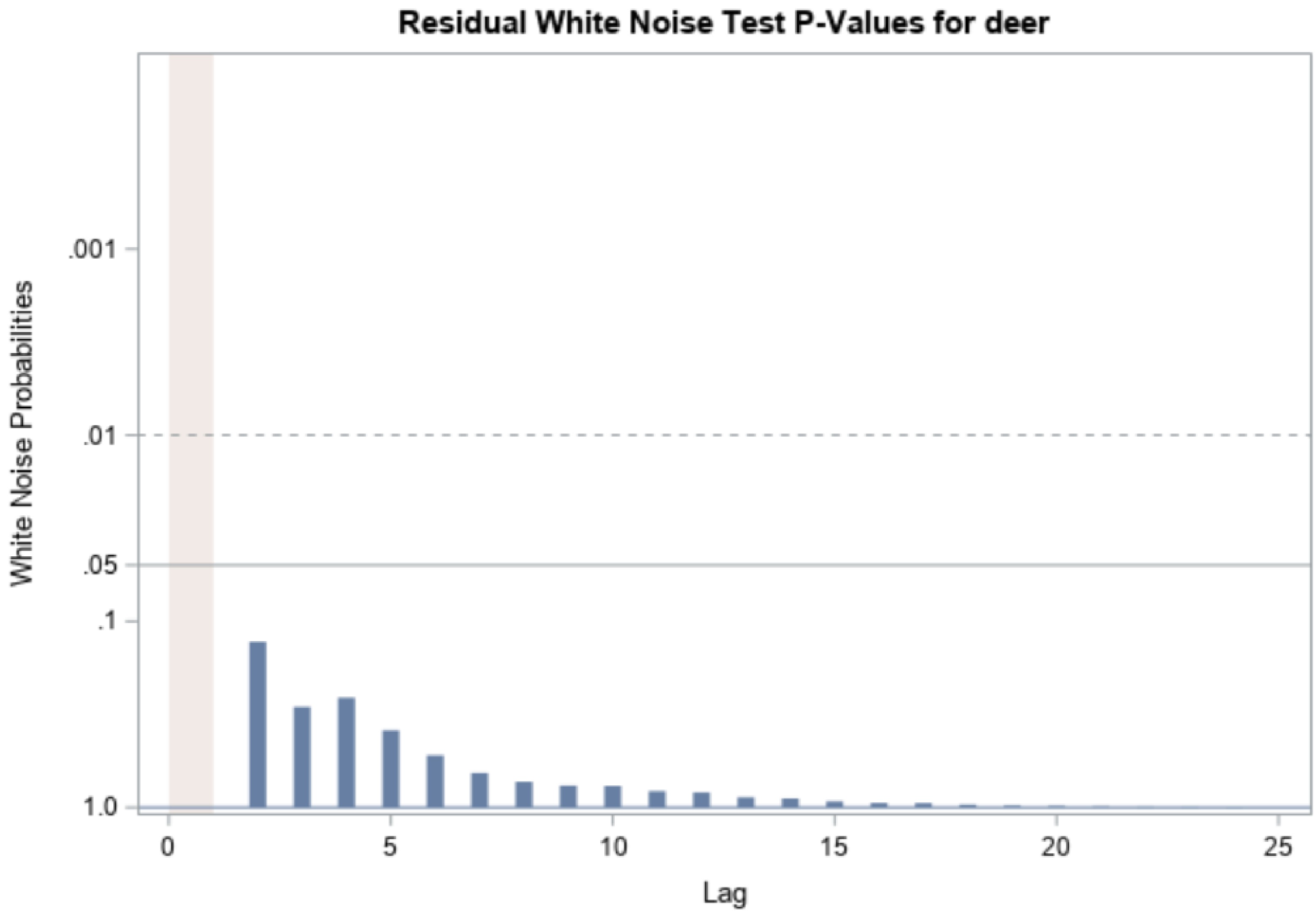

Residual Partial Autocorrelations for deer
1.0
0.5
0.0
-0.5
-1.0
PACF
0
5
10
15
20
25

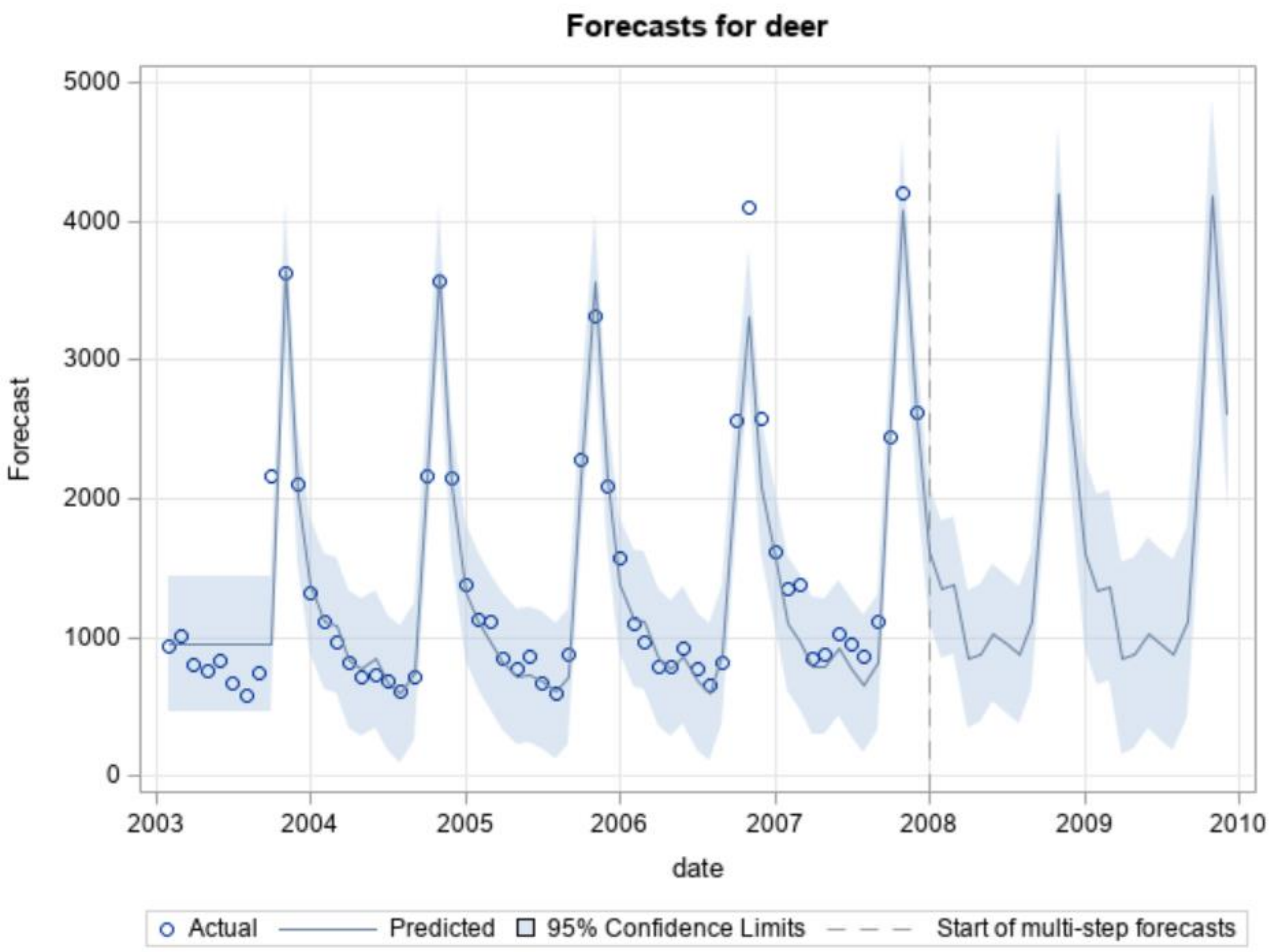
Point Intervention – Stochastic

```
Deer.Model2 <- arimax(Deer.Accidents, order=c(12,0,0),
                xtransf=NOV,
transfer=list(c(1,0)),method='ML')
summary(Deer.Model2)
```

Stochastic + Seasonal AR term

```
proc arima data=Time.DEER2 plot(unpack)=(series(corr) forecast(all));
	identify var=deer nlag=24 crosscorr=(Nov);
	estimate input=( /(1)Nov) p=(12);
	forecast lead=24 id=date interval=month;
run;
quit;
```


Residual White Noise Test P-Values for deer
White Noise Probabilities
.001
.01
.05
.1
1.0
0
5
10
15
20
25
Lag



Forecasts for deer
Forecast
5000
4000
3000
2000
1000
0
2003
2004
2005
2006
2007
2008
2009
2010
date
Actual
Predicted
95% Confidence Limits
Start of multi-step forecasts