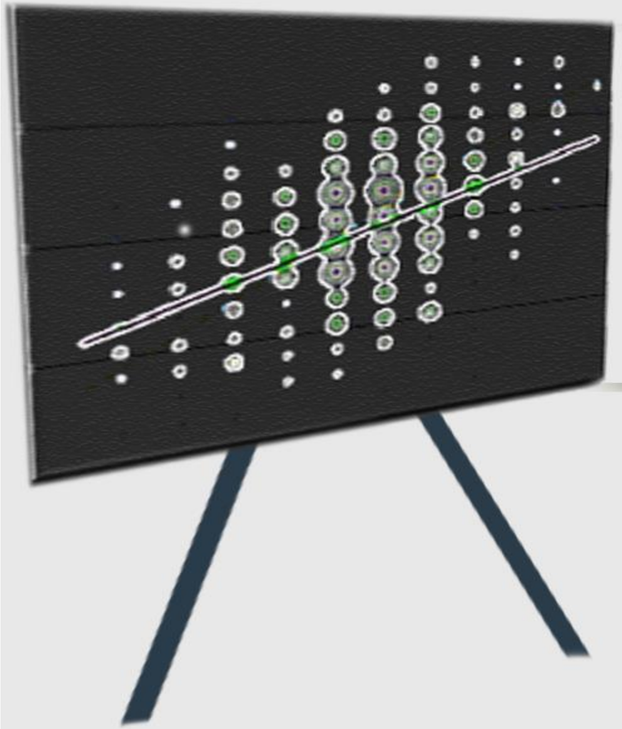


Forecasting

U. DINESH KUMAR

Those who have knowledge don’t predict.

Those who predict don’t have knowledge.

- Lao Tzu

I think there is a world market for may be 5 computers

- Thomas Watson, Chairman of IBM 1943

Computers in future weigh no more than 1.5 tons

- Popular Mechanics, 1949

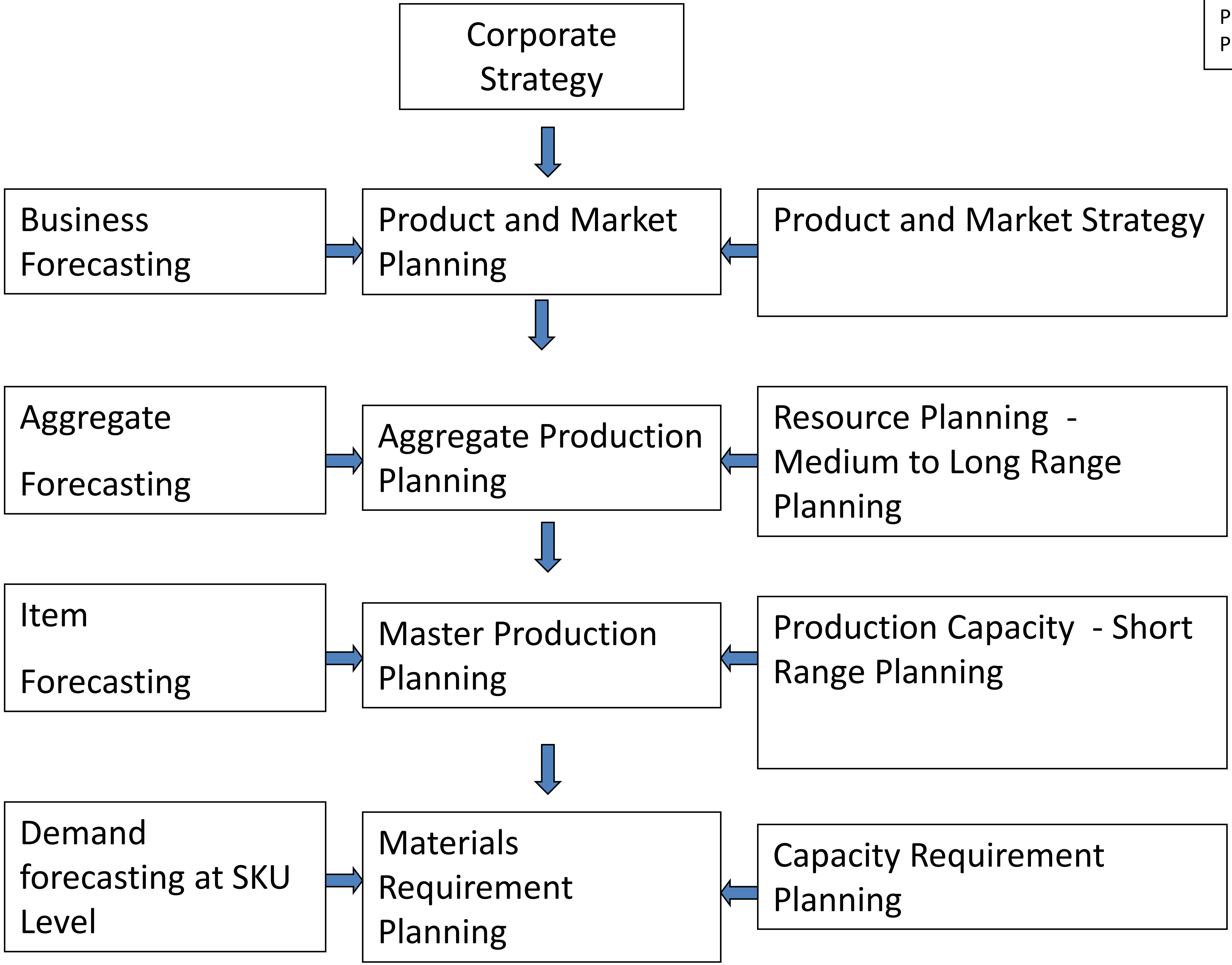
640K ought to be enough for everybody

- Bill Gates, 1981???

Forecasting

- Forecasting is a process of estimation of an unknown event/parameter such as demand for a product.
- Forecasting is commonly used to refer time series data.
- Time series is a sequence of data points measured at successive time intervals.

Predictive Analytics : QM901.1x
Prof U Dinesh Kumar, IIMB



IIMBX

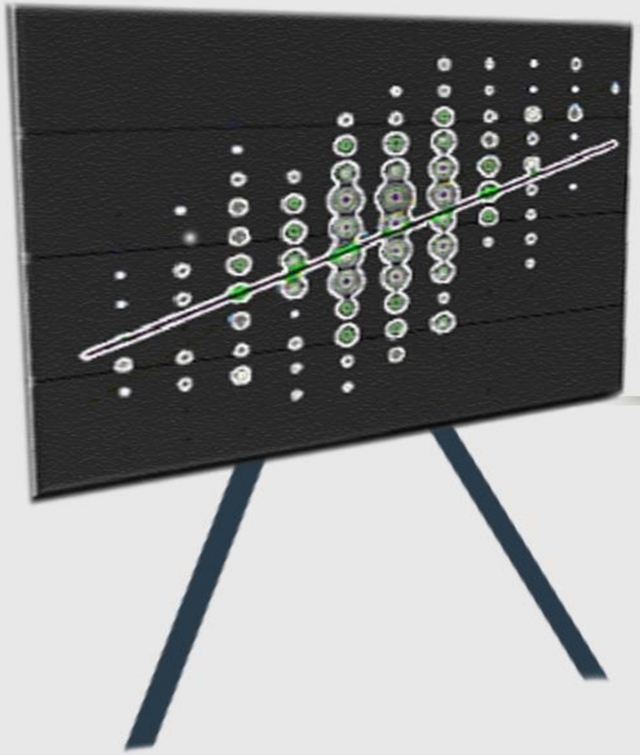
IIMB
तेजस्वि नावधीतमस्तु

Forecasting methods

- **Qualitative Techniques**.
  – Expert opinion (or Astrologers)

- **Quantitative Techniques**.
  – Time series techniques such as exponential smoothing

- **Casual Models**.
  – Uses information about relationship between system elements (e.g regression).



Time Series Analysis

U. DINESH KUMAR

IIMBX

IIMB
तेजस्वि नावधीतमस्तु

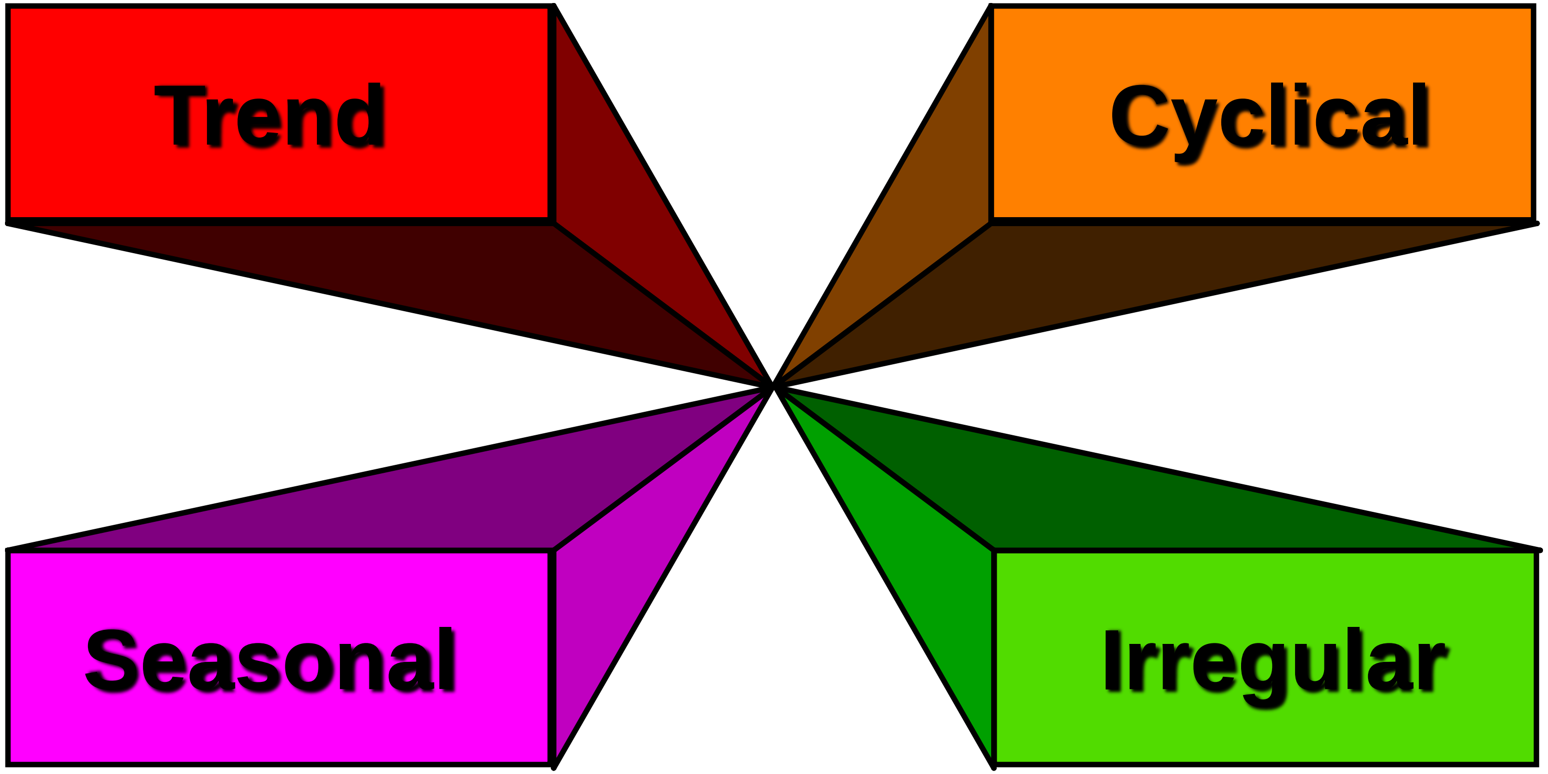
Why Time Series Analysis ?

- Time series analysis helps to identify and explain:
  - Any systemic variation in the series of data which is due to seasonality.
  - Cyclical pattern that repeat.
  - Trends in the data.
  - Growth rates in the trends.

IIMBX

IIMB
तेजस्वि नावधीतमस्तु

Time Series Components



Trend Component

- Persistent upward or downward pattern

- Due to consumer behaviour, population, economy, technology etc.

Cyclical Component

- Repeating up and down movements.
- Due to interaction of factors influencing economy such as recession.
- Usually 2-10 years duration.

Seasonal Component

- Regular pattern of up and down movements.
- Due to weather, customs, festivals etc.
- Occurs within one year.

Seasonal Vs Cyclical

- When a cyclical pattern in the data has a period of less than one year, it is referred as seasonal variation.

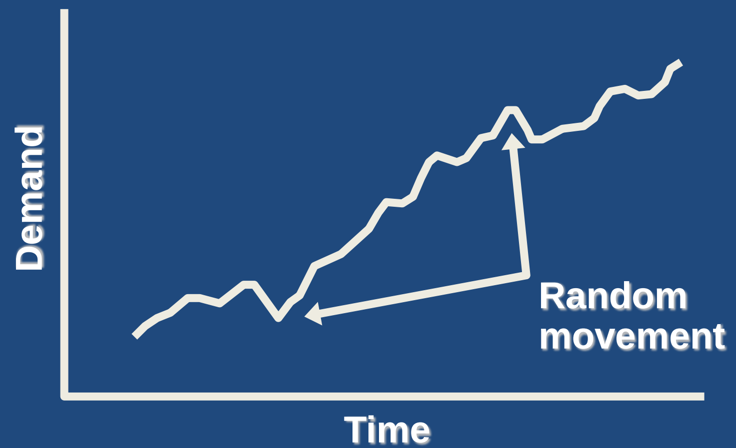
- When the cyclical pattern has a period of more than one year we refer to it as cyclical variation.

IIMBX

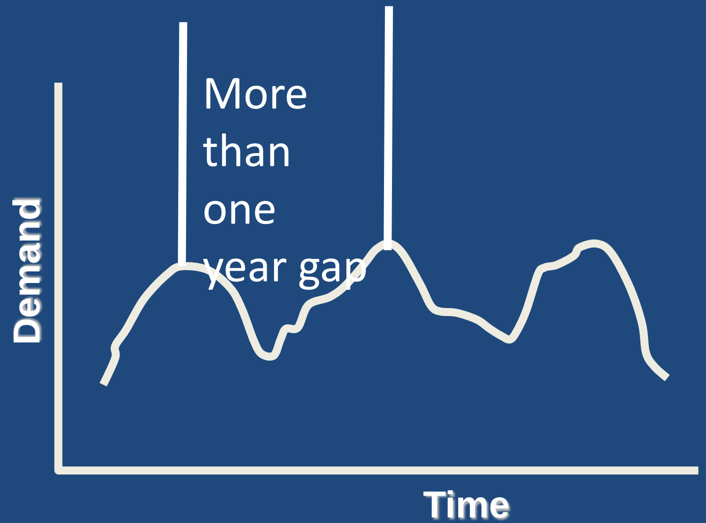
IIMB
तेजस्वि नावधीतमस्तु

Irregular Component

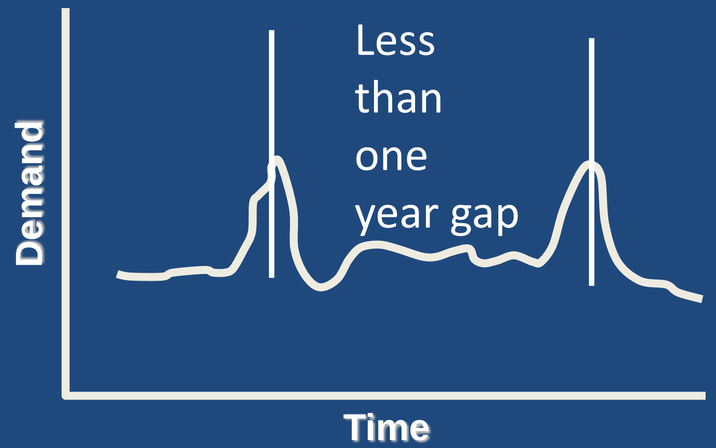
- Erratic fluctuations
- Due to random variation or unforeseen events
- White Noise



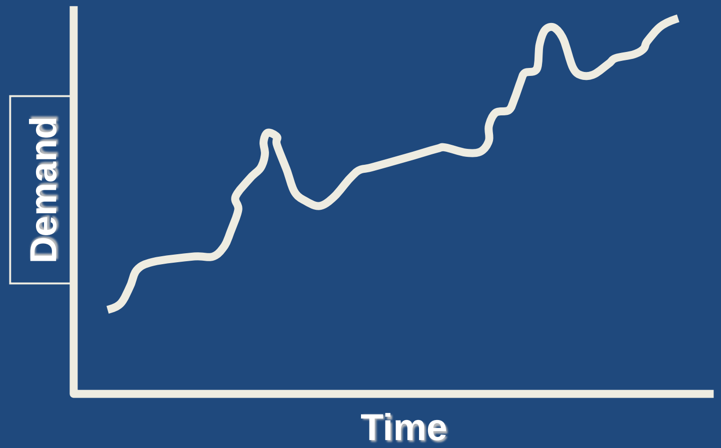
(a) Trend



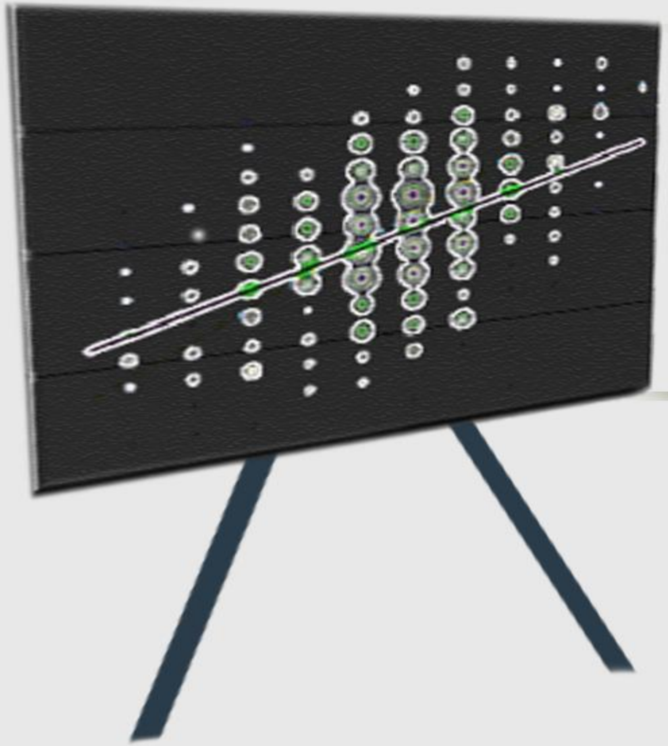
(b) Cycle



(c) Seasonal pattern



(d) Trend with seasonal pattern



Time Series Techniques for Forecasting

U. DINESH KUMAR

IIMBX

IIMB
तेजस्वि नावधीतमस्तु

Time Series Data – Additive and Multiplicative Models

1. Additive Forecasting Model

$$Y_t = \overbrace{T_t}^{Trend} + \overbrace{S_t}^{Seasonality} + \overbrace{C_t}^{Cyclical} + \overbrace{R_t}^{Random}$$

2. Multiplicative Forecasting Model

$$Y_t = \overbrace{T_t}^{Trend} \times \overbrace{S_t}^{Seasonality} \times \overbrace{C_t}^{Cyclical} \times \overbrace{R_t}^{Random}$$

IIMBX

IIMB तेजस्वि नावधीतमस्तु

Time Series Data Decomposition

Multiplicative Forecasting Model

$$Y_t = \overbrace{T_t}^{Trend} \times \overbrace{S_t}^{Seasonality}$$

$$S_t = \frac{Y_t}{T_t}$$

Y_t / T_t is called deseasonalized data

IIMBX

IIMB
तेजस्वि नावधीतमस्तु

Time Series Techniques

- Moving Average.
- Exponential Smoothing.
- Auto-regression Models (AR Models).
- ARIMA (Auto-regressive Integrated Moving Average) Models.

Moving Average Method

IIMBX

IIMB
तेजस्वि नावधीतमस्तु

Moving Average (Rolling Average)

- Simple moving average.
  - Used mainly to capture trend and smooth short term fluctuations.
  - Most recent data are given equal weights.

- Weighted moving average
  - Uses unequal weights for data

Data

- Demand for continental breakfast at the Die Another Day Hospital.
- Daily data between 1 October 2014 – 23 January 2015 (115 days)

Simple moving average

- The forecast for period t+1 (F_{t+1}) is given by the average of the 'n' most recent data.

$$F_{t+1} = \frac{1}{n} \sum_{i=t-n+1}^{t} Y_i$$

$$F_{t+1} = \text{Forecast for period } t+1$$

$$Y_i = \text{Data corresponding to time period } i$$

IIMBX

IIMB
तेजस्वि नावधीतमस्तु

Simple moving average

- The forecast for period t+1 (F_{t+1}) is given by the average of the 'n' most recent data.

$$F_{t+1} = \frac{1}{n} \sum_{i=t-n+1}^{t} Y_i$$

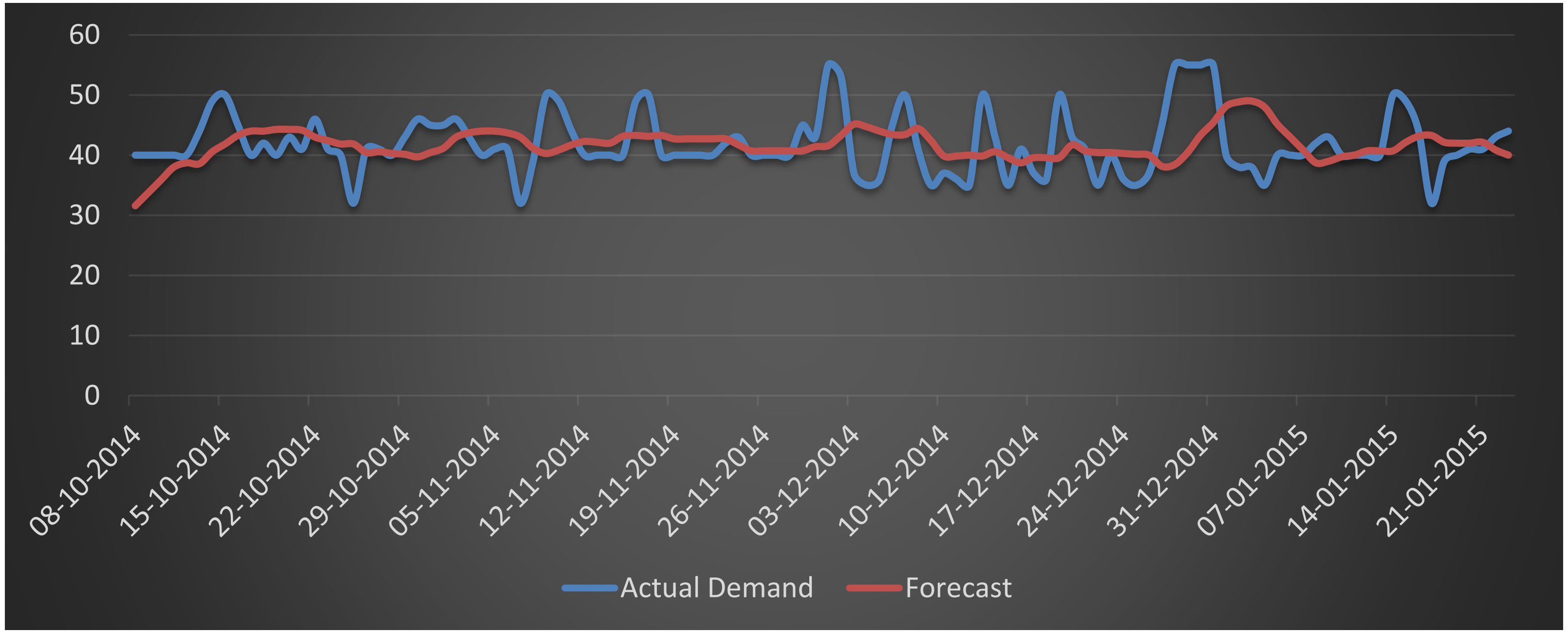
$$F_{t+1} = \text{Forecast for period } t+1$$

$$Y_i = \text{Data corresponding to time period } i$$

IIMBX

IIMB
तेजस्वि नावधीतमस्तु

Demand for Continental Breakfast at DAD Hospital – Moving Average



$$F_{t+1} = \frac{1}{7} \sum_{i=t-7+1}^{t} Y_i$$

IIMBX

IIMB
तेजस्वि नावधीतमस्तु

Measures of aggregate error

Mean absolute error **MAE**	$MAE = \frac{1}{n}\sum_{t=1}^{n}\lvert E_t \rvert$
Mean absolute percentage error **MAPE**	$MAPE = \frac{1}{n}\sum_{t=1}^{n}\left\lvert \frac{E_t}{Y_t} \right\rvert$
Mean squared error **MSE**	$MSE = \frac{1}{n}\sum_{t=1}^{n} E_t^2$
Root mean squared error **RMSE**	$RMSE = \sqrt{\frac{1}{n}\sum_{t=1}^{n} E_t^2}$

$E_t = F_t - Y_t$

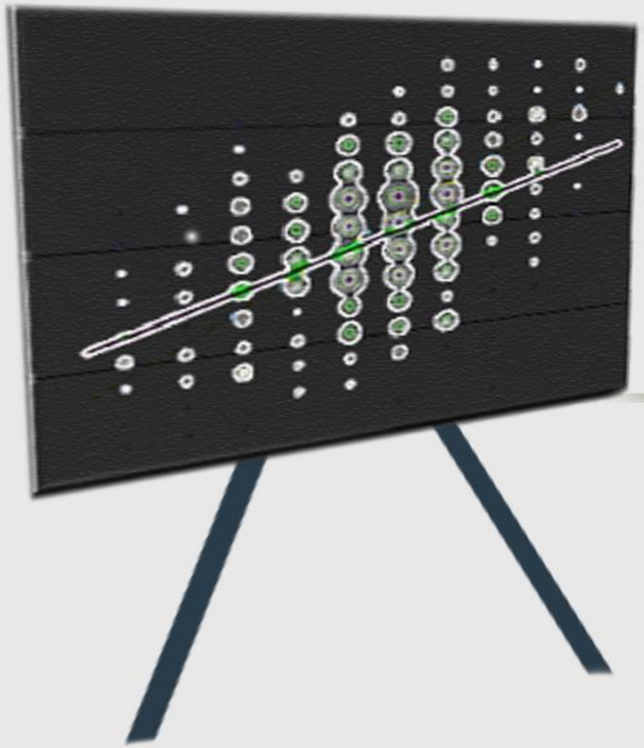
IIMBX

IIMB
तेजस्वि नावधीतमस्तु

DAD Forecasting – MAPE and RMSE

Mean absolute percentage error **MAPE**	0.1068 or 10.68%
Root mean squared error **RMSE**	5.8199

Predictive Analytics : QM901.1x
Prof U Dinesh Kumar, IIMB



Exponential Smoothing

U. DINESH KUMAR

IIMBX

IIMB
तेजस्वि नावधीतमस्तु

Exponential Smoothing

- Form of Weighted Moving Average
  - Weights decline exponentially.
  - Largest weight is given to the present observation, less weight to immediately preceding observation and so on.

- Requires smoothing constant (α)
  - Ranges from 0 to 1

Exponential Smoothing

Next forecast = $\alpha \times$ (present actual value) + $(1-\alpha) \times$ present forecast

IIMBX

IIMB
तेजस्वि नावधीतमस्तु

Simple Exponential Smoothing Equations

- Smoothing Equations

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

$$F_1 = Y_1$$

F_{t+1} is the forecasted value at time t+1

Simple Exponential Smoothing Equations

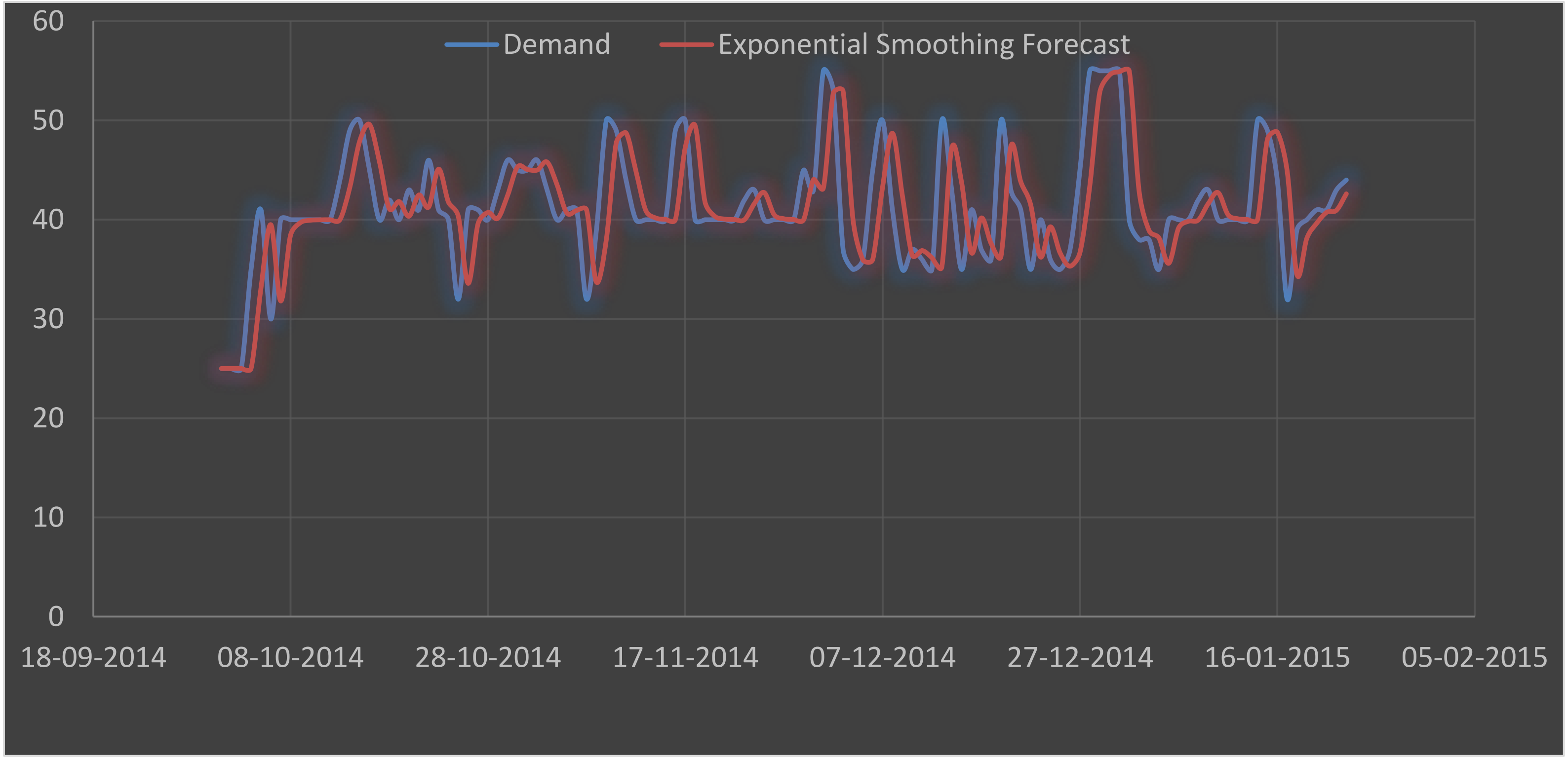
- Smoothing Equations

$$F_t = \alpha Y_{t-1} + \alpha(1-\alpha)Y_{t-2} + \alpha(1-\alpha)^2 Y_{t-3} + \ldots.$$

IIMBX

IIMB
तेजस्वि नावधीतमस्तु

Exponential Smoothing Forecast



$$F_{t+1} = 0.8098 * Y_t + (1 - 0.8098) * F_t$$

IIMBX

IIMB
तेजस्वि नावधीतमस्तु

DAD Forecasting – MAPE and RMSE Exponential Smoothing

Mean absolute percentage error **MAPE**	0.0906 or 9.06%
Root mean squared error **RMSE**	5.3806

α **= 0.8098**

IIMBX

IIMB
तेजस्वि नावधीतमस्तु

Choice of α

For “smooth” data, try a high value of a α, forecast responsive to most current data.

For “noisy” data try low a α forecast more stable—less responsive

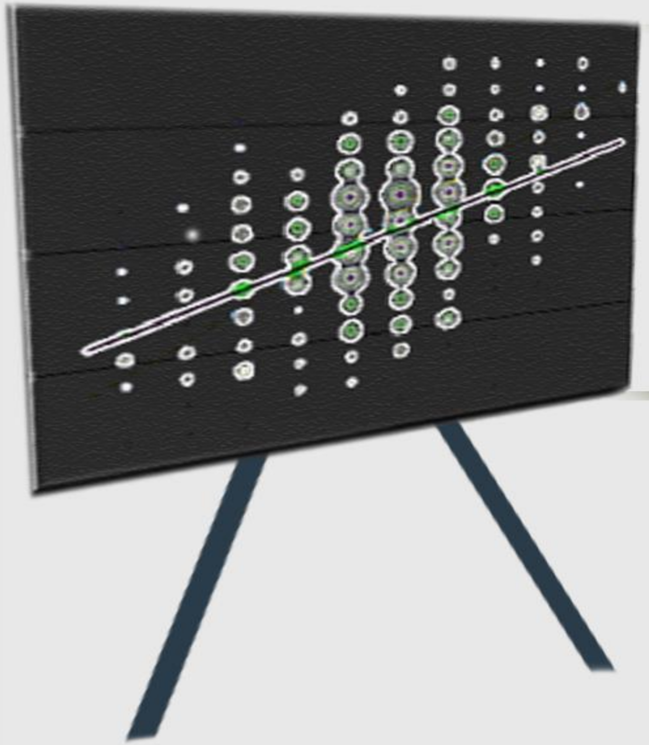
IIMBX

IIMB
तेजस्वि नावधीतमस्तु

Optimal α

$$\underset{\alpha}{Min}\left[\sqrt{\sum_{t=1}^{n}\left(Y_t - F_t\right)^2 / n}\right]$$

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



Double Exponential Smoothing

U. DINESH KUMAR

IIMBX

IIMB
तेजस्वि नावधीतमस्तु

Double exponential Smoothing – Holt's model

Simple exponential smoothing may produce consistently biased forecasts in the presence of a trend.

Holt's method (double exponential smoothing) is appropriate when demand has a trend but no seasonality.

Systematic component of demand = Level + Trend

Holt’s method

- Holt’s method can be used to forecast when there is a linear trend present in the data.

- The method requires separate smoothing constants for slope and intercept.

Predictive Analytics : QM901.1x
Prof U Dinesh Kumar, IIMB

Holt's Method

Equation for intercept or level

- Holt's Equations

$$(i) \quad L_t = \alpha \times Y_t + (1-\alpha)\times(L_{t-1} + T_{t-1})$$

$$(ii) \quad T_t = \beta \times (L_t - L_{t-1}) + (1-\beta)\times T_{t-1}$$

Equation for Slope (Trend)

- Forecast Equation

$$\hat{F}_{t+1} = L_t + T_t$$

IIMBX

IIMB
तेजस्वि नावधीतमस्तु

Initial values of L_t and T_t

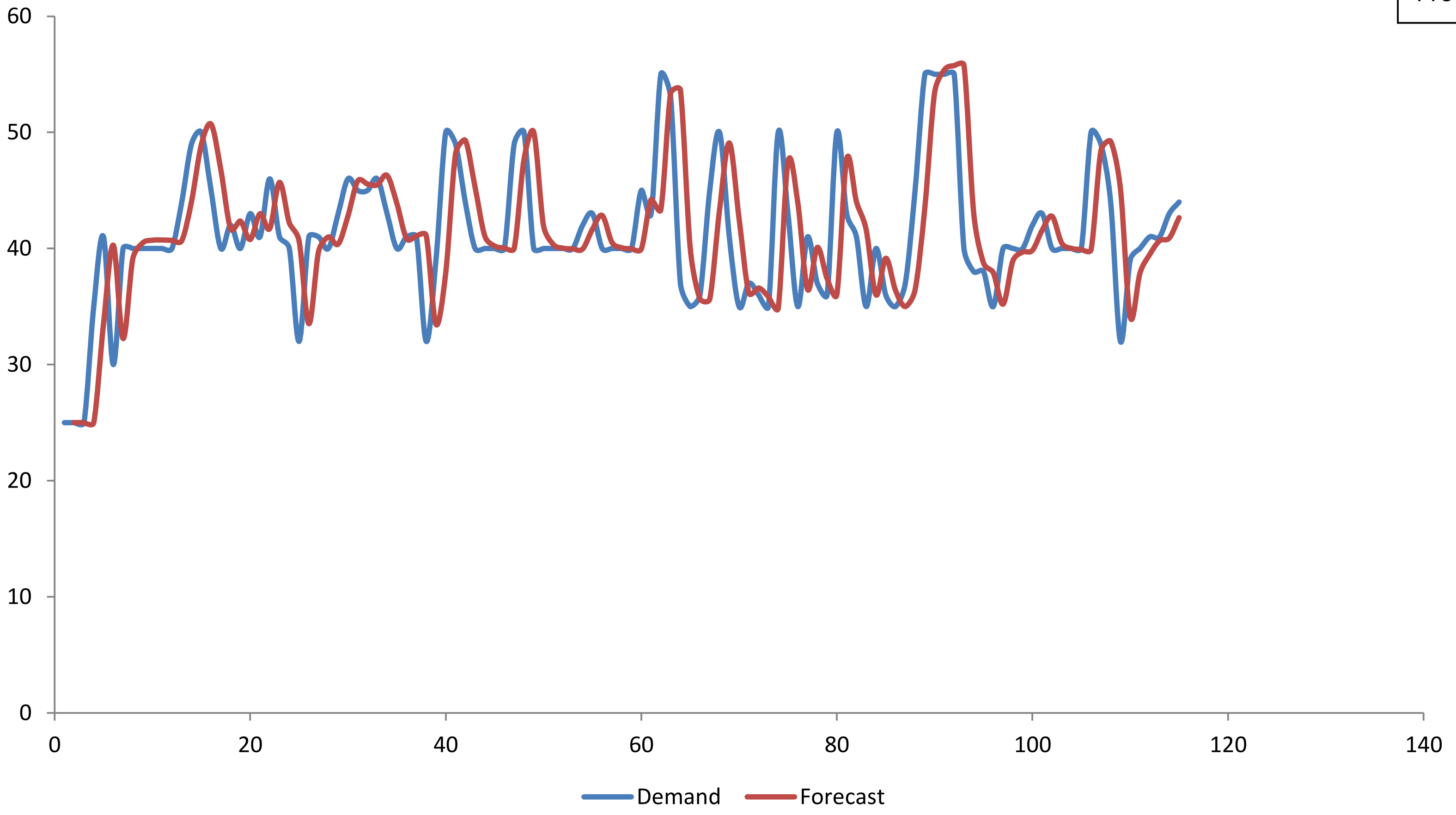
- L_1 is in general set to Y_1.

- T_1 can be set to any one of the following values (or use regression to get initial values):

$$T_1 = (Y_2 - Y_1)$$
$$T_1 = \left[(Y_2 - Y_1) + (Y_3 - Y_2) + (Y_4 - Y_3)\right] / 3$$
$$T_1 = (Y_n - Y_1)/(n-1)$$

IIMBX

IIMB तेजस्वि नावधीतमस्तु



Double exponential smoothing

α = 0.8098; β = 0.05

IIMBX

IIMB
तेजस्वि नावधीतमस्तु

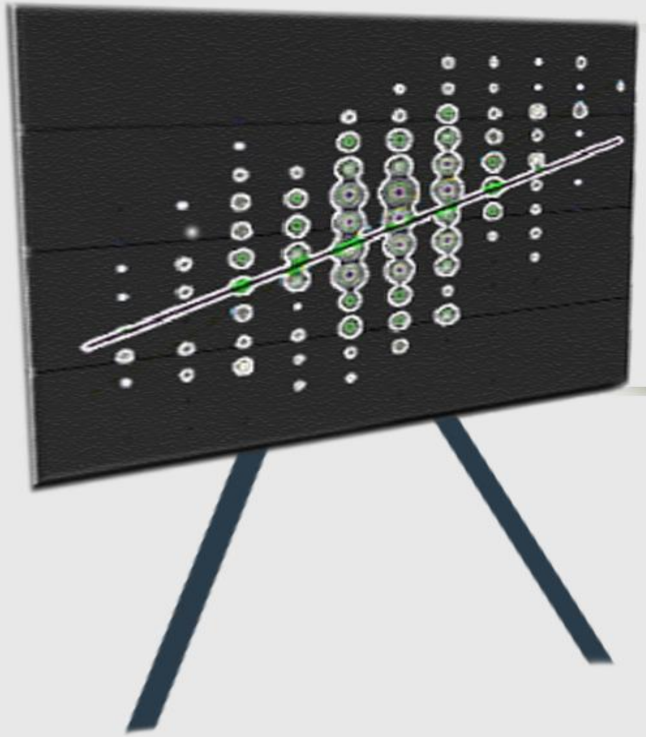
DAD Forecasting – MAPE and RMSE Double Exponential Smoothing

Mean absolute percentage error **MAPE**	0.0930 or 9.3%
Root mean squared error **RMSE**	5.5052

α = 0.8098; β = 0.05

IIMBX

IIMB
तेजस्वि नावधीतमस्तु



Forecasting Accuracy

U. DINESH KUMAR

Forecasting Accuracy

- The forecast error is the difference between the forecast value and the actual value for the corresponding period.

$$E_t = Y_t - F_t$$

$$E_t = \text{Forecast error at period t}$$

$$Y_t = \text{Actual value at time period t}$$

$$F_t = \text{Forecast for time period t}$$

Measures of aggregate error

Mean absolute error **MAE**	$MAE = \frac{1}{n}\sum_{t=1}^{n}\left\|E_t\right\|$
Mean absolute percentage error **MAPE**	$MAPE = \frac{1}{n}\sum_{t=1}^{n}\left\|\frac{E_t}{Y_t}\right\|$
Mean squared error **MSE**	$MSE = \frac{1}{n}\sum_{t=1}^{n}E_t^2$
Root mean squared error **RMSE**	$RMSE = \sqrt{\frac{1}{n}\sum_{t=1}^{n}E_t^2}$

IIMBX

IIMB
तेजस्वि नावधीतमस्तु

Forecasting Power of a Model

Theil’s coefficient (U Statistic)

$$U = \frac{\sum_{t=1}^{n} (Y_{t+1} - F_{t+1})^2}{\sum_{t=1}^{n} (Y_{t+1} - Y_t)^2}$$

The value U is the relative forecasting power of the method against naïve technique. If U < 1, the technique is better than naïve forecast If U > 1, the technique is no better than the naïve forecast.

Theil's coefficient for DAD Hospital

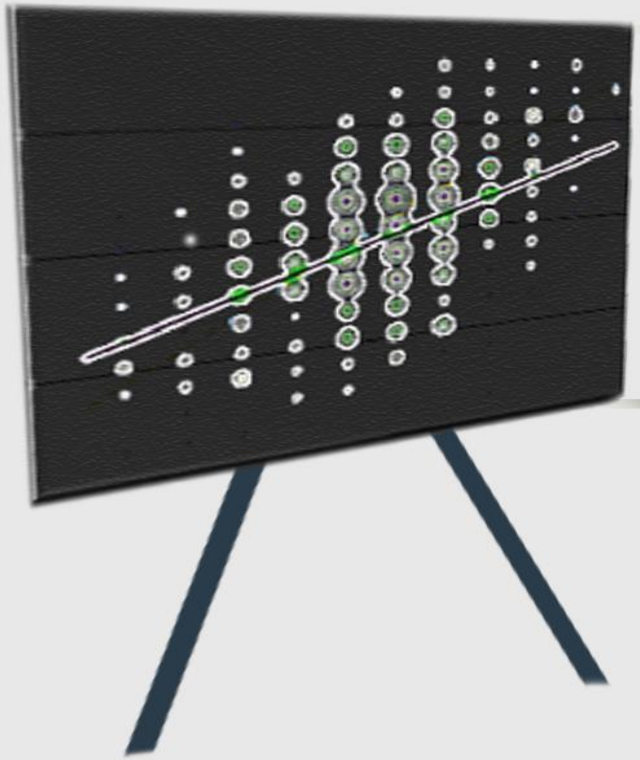
Method	U
Moving Average with 7 periods	1.221
Exponential Smoothing	1.704
Double exponential Smoothing	1.0310

Theil's Coefficient

$$U1 = \frac{\sqrt{\sum_{t=1}^{n}(Y_t - F_t)^2}}{\sqrt{\sum_{t=1}^{n} Y_t^2} + \sqrt{\sum_{t=1}^{n} F_t^2}}, \qquad U2 = \sqrt{\frac{\sum_{t=1}^{n-1}\left(\frac{F_{t+1} - Y_{t+1}}{Y_t}\right)^2}{\sum_{t=1}^{n-1}\left(\frac{Y_{t+1} - Y_t}{Y_t}\right)^2}}$$

U1 is bounded between 0 and 1, with values closure to zero indicating greater accuracy.

If **U2 = 1**, there is no difference between naïve forecast and the forecasting technique
If **U2 < 1**, the technique is better than naïve forecast
If **U2 > 1**, the technique is no better than the naïve forecast.



Time Series Data with Seasonality

U. DINESH KUMAR

IIMBX

IIMB
तेजस्वि नावधीतमस्तु

Seasonal Effect

- Seasonal effect is defined as the repetitive and predictable pattern of data behaviour in a time-series around the trend line.

- In seasonal effect the time period must be less than one year, such as, days, weeks, months or quarters.

Seasonal effect

- Identification of seasonal effect provides better understanding of the time series data.
- Seasonal effect can be eliminated from the time-series. This process is called deseasonalization or seasonal adjusting.

Seasonal Adjusting

Seasonal adjustment in multiplicative model

$$\text{Seasonal effect} = \frac{\mathrm{T_t S_t}}{\mathrm{T_t}} = \frac{Y_t}{\mathrm{T_t}} \times 100$$

Seasonal adjustment in additive model

$$\mathrm{Y_t} - S_t = \mathrm{T_t} + \mathrm{S_t} - S_t = \mathrm{T_t}$$

Seasonal Index

- Method of simple averages
- Ratio-to-moving average method

Method of simple averages

- Average the unadjusted data by period (for example daily or monthly).

- Calculate the average of daily (or monthly) averages.

- Seasonal index for day i (or month i) is the ratio of daily average of day i (or month i) to the average of daily (or monthly) averages times 100.

Predictive Analytics : QM901.1x
Prof U Dinesh Kumar, IIMB

Example: DAD Example - Demand for Continental Breakfast

5 October – 1 November 2014 Data

DAY	Week (5-11 October)	Week 2 (12-18 October)	Week 3 (19-25 OCT)	Week 4 (26 OCT - 1 NOV)
Sunday	41.00	40.00	40.00	41.00
Monday	30.00	44.00	43.00	41.00
Tuesday	40.00	49.00	41.00	40.00
Wednesday	40.00	50.00	46.00	43.00
Thursday	40.00	45.00	41.00	46.00
Friday	40.00	40.00	40.00	45.00
Saturday	40.00	42.00	32.00	45.00

IIMBX

IIMB
तेजस्वि नावधीतमस्तु

Seasonality Index

DAY	Week1 (5-11 October)	Week 2 (12-18 October)	Week 3 (19-25 OCT)	Week 4 (26 Oct-1 Nov)	Daily Average	Seasonality Index
Sunday	41.00	40.00	40.00	41.00	40.50	97.34%
Monday	30.00	44.00	43.00	41.00	39.50	94.94%
Tuesday	40.00	49.00	41.00	40.00	42.50	102.15%
Wednesday	40.00	50.00	46.00	43.00	44.75	107.55%
Thursday	40.00	45.00	41.00	46.00	43.00	103.34%
Friday	40.00	40.00	40.00	45.00	41.25	99.14%
Saturday	40.00	42.00	32.00	45.00	39.75	95.54%
					41.61	700

Average of daily averages

Total = Number of seasons x 100

IIMBX

IIMB तेजस्वि नावधीतमस्तु

Deseasonalized Data

Date	Demand	Seasonal Index	De-seasonalized Data
10/5/2014	41	97.34%	42.12
10/6/2014	30	94.94%	31.60
10/7/2014	40	102.15%	39.16
10/8/2014	40	107.55%	37.19
10/9/2014	40	103.34%	38.70
10/10/2014	40	99.14%	40.35
10/11/2014	40	95.54%	41.87

Trend

- Trend is calculated using regression on deseasonalized data.
- Deseasonalized data is obtained by dividing the actual data with its seasonality index.

SUMMARY OUTPUT						
Regression Statistics						
Multiple R	0.268288					
R Square	0.071978					
Adjusted R Square	0.036285					
Standard Error	3.715395					
Observations	28					
ANOVA						
	df	**SS**	**MS**	**F**	**Significance F**	
Regression	1	27.83729	27.83729	2.016587	0.167469	
Residual	26	358.9082	13.80416			
Total	27	386.7455				
	Coefficients	**Standard Error**	**t Stat**	**P-value**	**Lower 95%**	**Upper 95%**
Intercept	39.81731	1.442768	27.59786	8.7E-21	36.85166	42.78296
Day	0.123437	0.086923	1.420066	0.167469	-0.05524	0.30211

IIMBX

IIMB
तेजस्वि नावधीतमस्तु

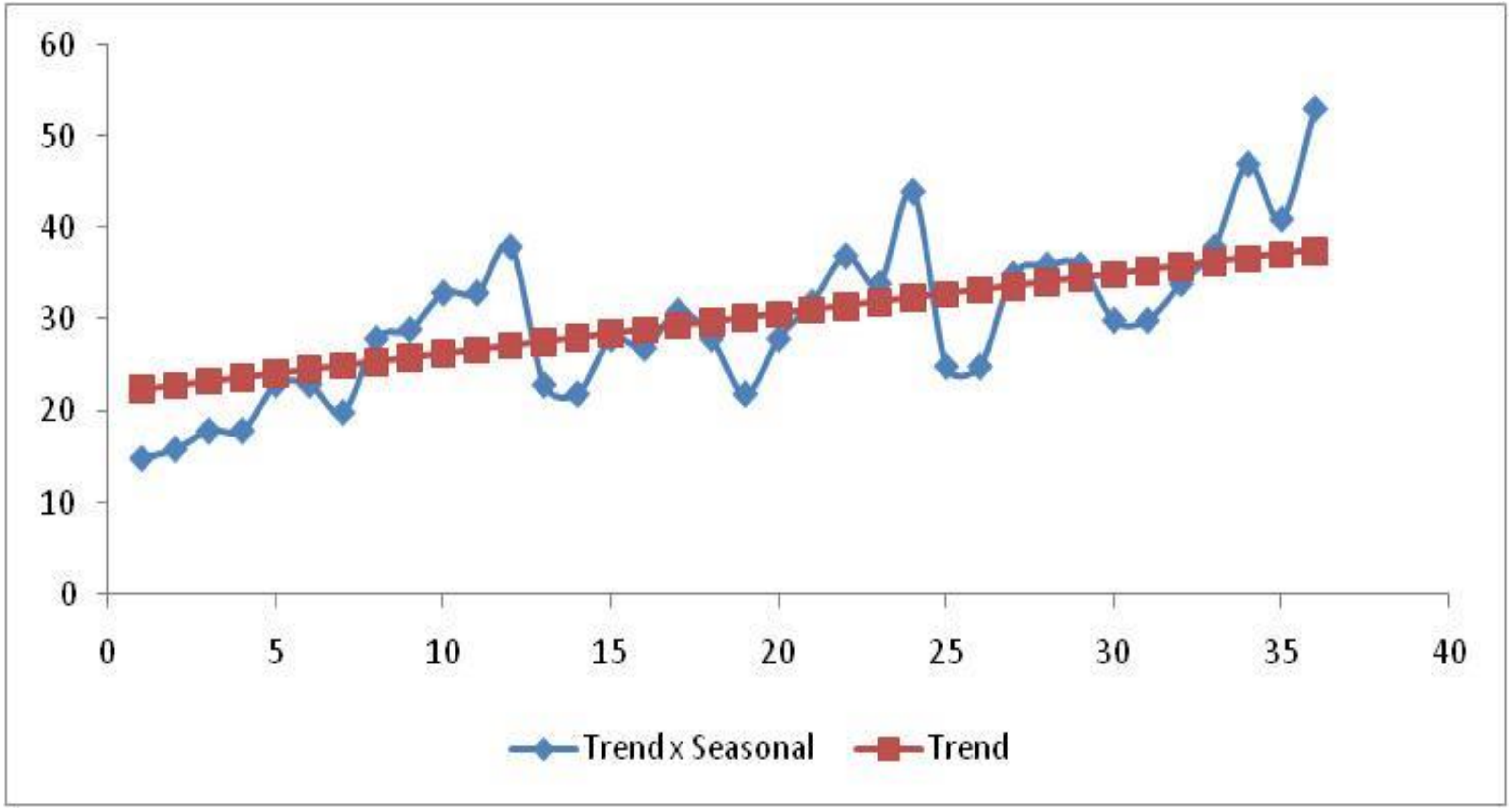
Forecast

$$F_t = T_t \text{ x } S_t$$

$$F_{29} = T_{29} \text{ x } S_{29}$$

F_{29} = (39.81 + 0.1234 x 29) x 0.9733 = **42.2372**

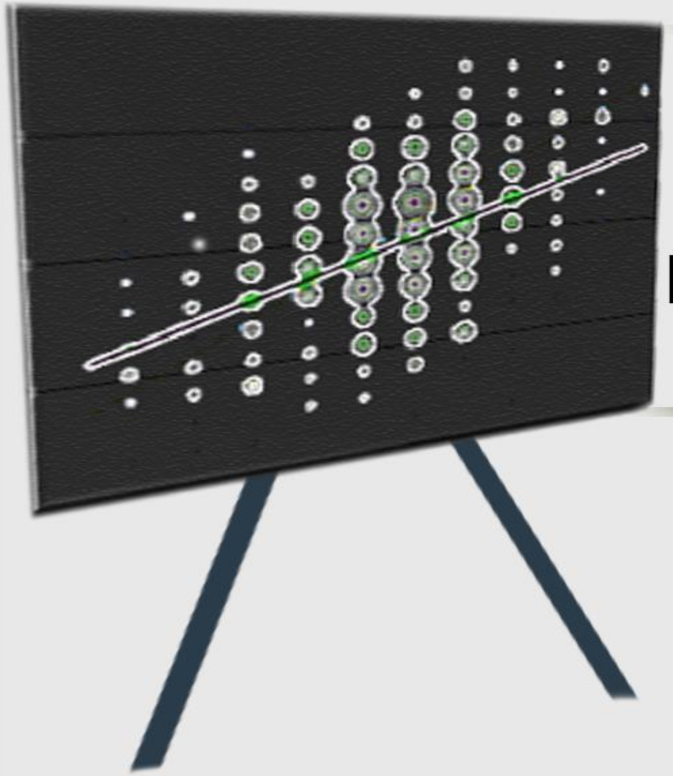
Y_{29} = 46



Forecasting using method of averages in the presence of seasonality

IIMBX

IIMB
तेजस्वि नावधीतमस्तु



Auto-regressive and Moving Average models

U. DINESH KUMAR

IIMBX

IIMB
तेजस्वि नावधीतमस्तु

Auto-correlation

- Auto correlation is a correlation of a variable observed at two time points (e.g. Y_t and Y_{t-1} or Y_t and Y_{t-3}).

- Auto-correlation of lag k, ρ_k, is given by:

$$\rho_k = \frac{\sum_{t=k+1}^{n}\left(Y_{t-k} - \bar{Y}\right)\left(Y_t - \bar{Y}\right)}{\sum_{t=1}^{n}(Y_t - \bar{Y})^2}$$

n = total number of observations

IIMBX

IIMB
तेजस्वि नावधीतमस्तु

Auto-correlation Function (ACF)

- A k-period plot of autocorrelations is called autocorrelation function (ACF) or a correlogram.

Auto-Correlation

- Auto-correlation of lag k is auto-correlation between Y_t and Y_{t+k}.
- To test whether the autocorrelation at lag k is significantly different from 0, the following hypothesis test is used:

$$H_0: \rho_k = 0$$

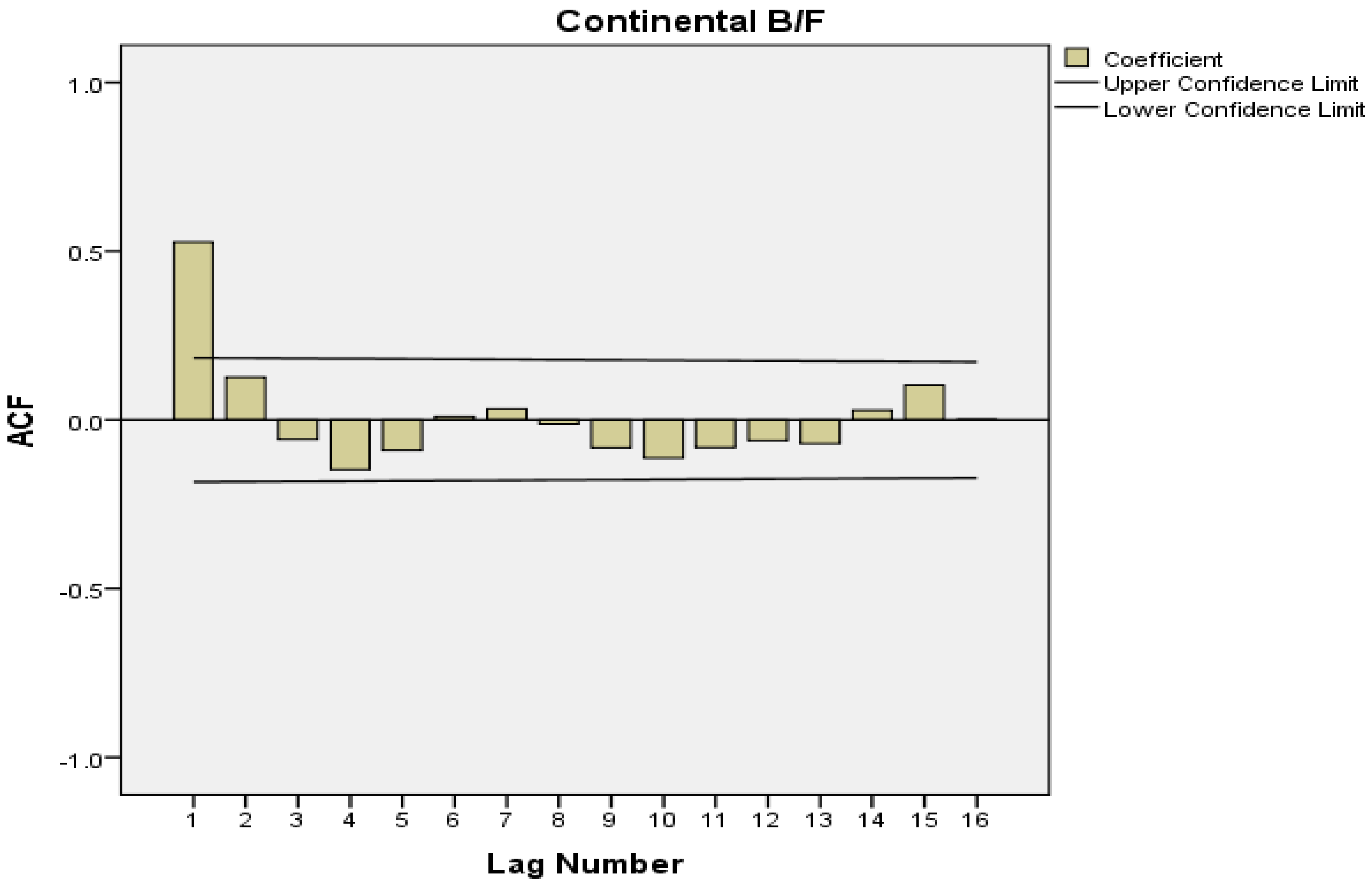
$$H_A: \rho_k \neq 0$$

- For any k, reject H_0 if $|\rho_k| > 1.96/\sqrt{n}$. Where n is the number of observations.

IIMBX

IIMB तेजस्वि नावधीतमस्तु

ACF for Demand for Continental Breakfast



Partial Auto-correlation

Partial auto correlation of lag k is an auto correlation between Y_t and Y_{t-k} with linear dependence between the intermedia values (Y_{t-k+1}, Y_{t-k+2}, …, Y_{t-1}) removed.

Partial Auto-correlation Function

A k-period plot of partial autocorrelations is called partial autocorrelation function (PACF).

Partial Auto-Correlation

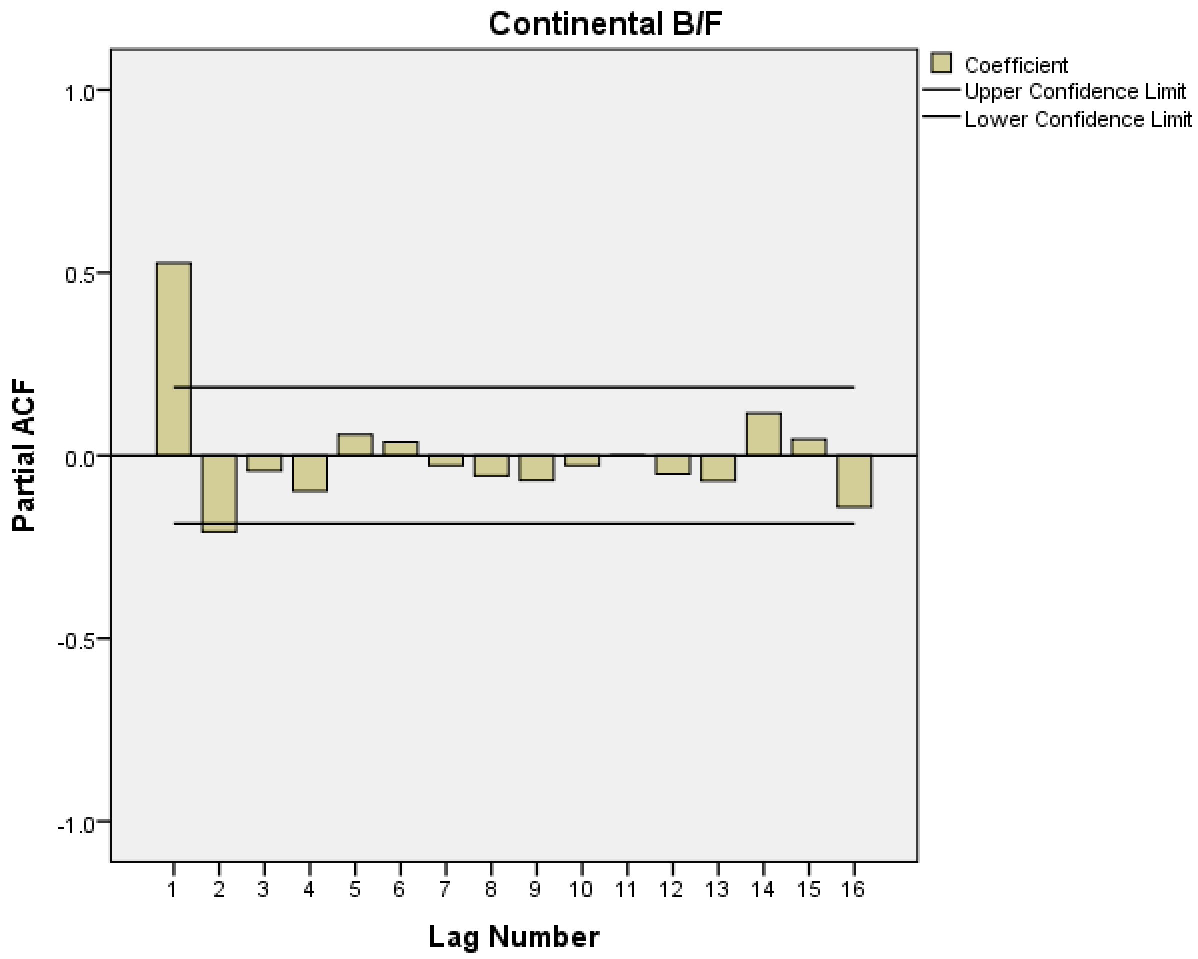
- Partial auto-correlation of lag k is auto-correlation between Y_t and Y_{t+k} after the removal of linear dependence of Y_{t+1} to Y_{t+k-1}.

- To test whether the partial autocorrelation at lag k is significantly different from 0, the following hypothesis test is used:

$$H_0: \rho_{pk} = 0$$
$$H_A: \rho_{pk} \neq 0$$

- For any k, reject H_0 if $|\rho_{pk}| > 1.96/\sqrt{n}$. Where n is the number of observations.

PACF – Demand for Continental Breakfast at DAD hospital



Stationarity

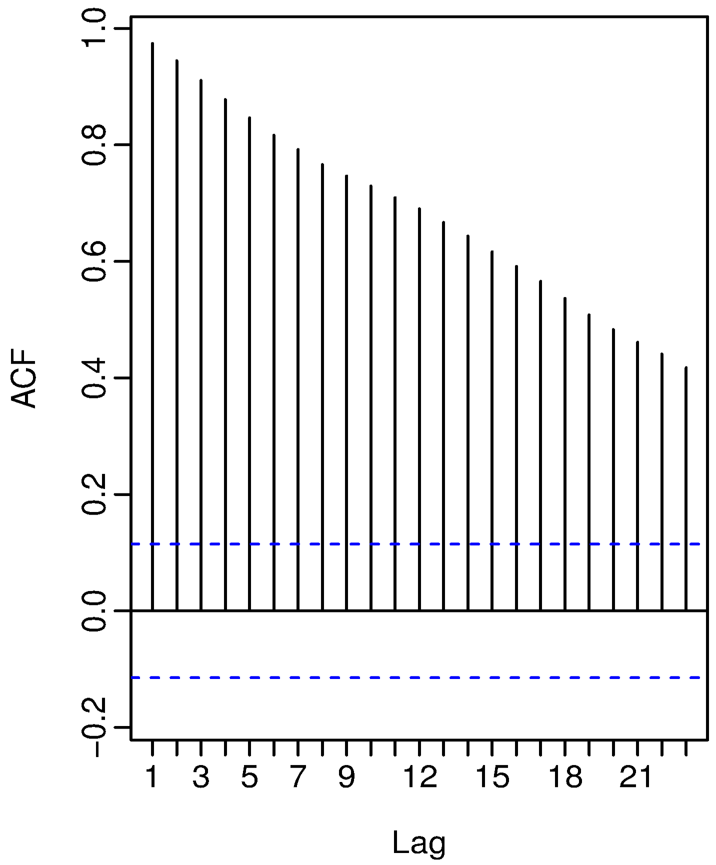
- A time series is stationary, if:
  - Mean (μ) is constant over time.
  - Variance (σ) is constant over time.
  - The covariance between two time periods (Y_t) and (Y_{t+k}) depends only on the lag k not on the time t.

We assume that the time series is stationary before applying forecasting models

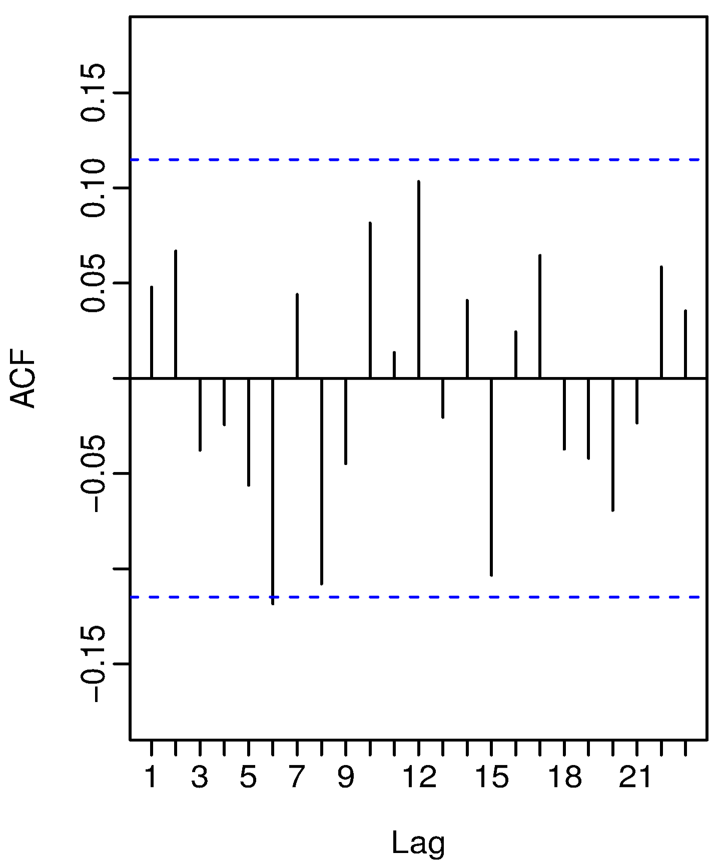
IIMBX

IIMB
तेजस्वि नावधीतमस्तु

ACF Plot of non-stationary and stationary process



Non-stationary

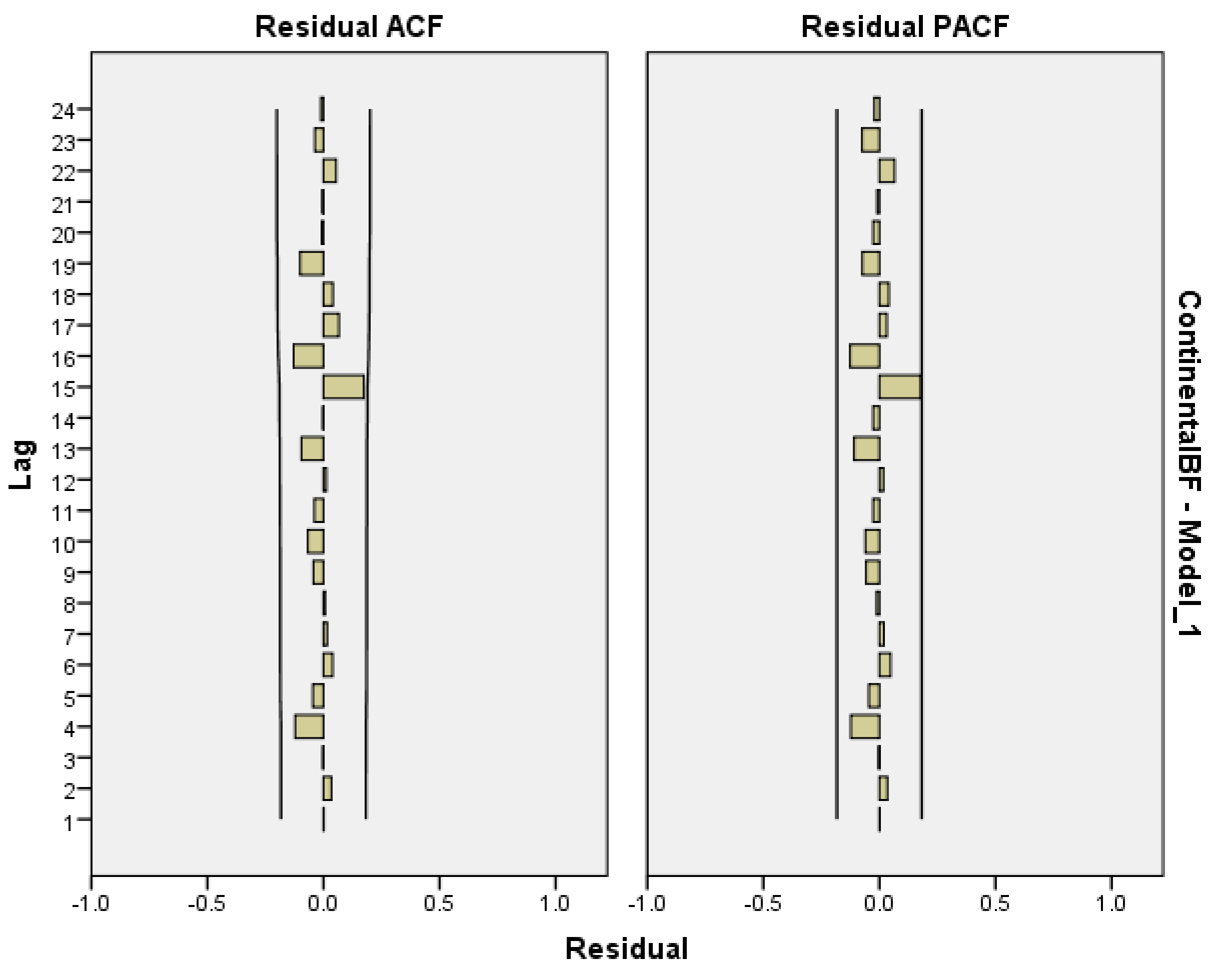


Stationary

White Noise

- White noise is a data uncorrelated across time that follow normal distribution with mean 0 and constant standard deviation σ.

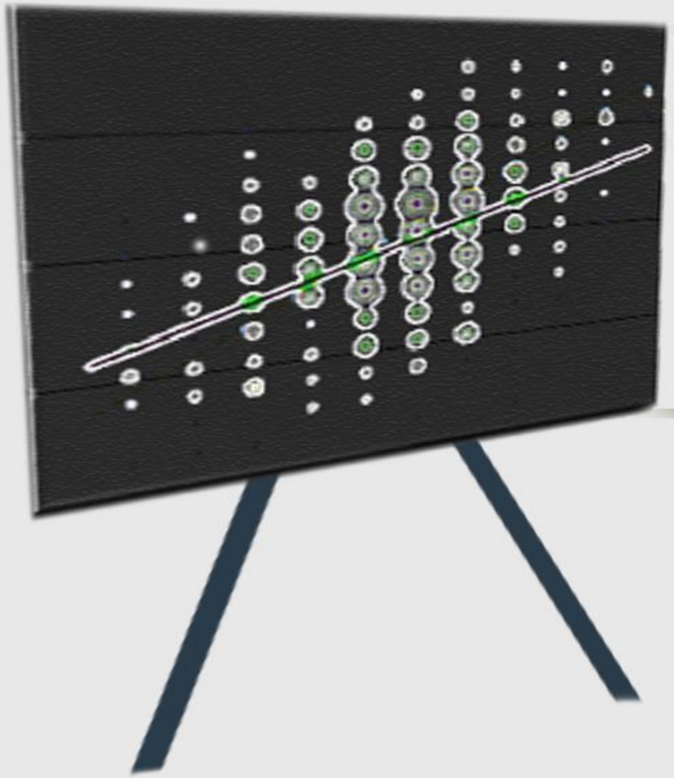
- In forecasting we assume that the residuals are white Noise.



Residual White Noise

IIMBX

IIMB
तेजस्वि नावधीतमस्तु



Auto regressive models

U. DINESH KUMAR

IIMBX

IIMB
तेजस्वि नावधीतमस्तु

Auto-Regression

- Auto-regression is a regression of Y on itself observed at different time points.

- That is, we use Y_t as the response variable and Y_{t-1}, Y_{t-2} etc. as explanatory variables.

AR(1) Parameter Estimation

$$Y_t = \phi Y_{t-1} + \varepsilon_t$$

$$\sum_{t=2}^{n} \varepsilon_t^{\,2} = \sum_{t=2}^{n} (Y_t - \phi Y_{t-1})^2$$

$$\hat{\phi} = \frac{\sum_{t=2}^{n} Y_t Y_{t-1}}{\sum_{t=2}^{n} Y_{t-1}^2}$$

OLS Estimate

IIMBX

IIMB
तेजस्वि नावधीतमस्तु

AR(1) Process

$$Y_t = \phi Y_{t-1} + \varepsilon_t$$

$$Y_t = \phi\ (\phi Y_{t-2} + \varepsilon_{t-1}) + \varepsilon_t$$

…………………………………………………………………

$$Y_t = \phi^t\ X_0 + \phi^{t-1}\ \varepsilon_1\ + \phi^{t-2}\ \varepsilon_2 + \ldots + \phi^1\ \varepsilon_{t-1} + \varepsilon_t$$

$$Y_t = \phi^t X_0 + \sum_{i=0}^{t-1} \phi^i \varepsilon_{t-i}$$

IIMBX

IIMB
तेजस्वि नावधीतमस्तु

Auto-regressive process (AR(p))

- Assume $\{Y_t\}$ is purely random with mean zero and constant standard deviation σ (White Noise).

- Then the autoregressive process of order p or AR(p) process is

$$Y_t = \phi_0 + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \ldots + \phi_p Y_{t-p} + \varepsilon_t$$

AR(p) process models each future observation as a function "p" previous observations.

IIMBX

IIMB तेजस्वि नावधीतमस्तु

Model Identification in AR(p) Process

Pure AR Model Identification

Model	ACF	PACF
AR(1)	Exponential Decay: Positive side if ø1 > 0 and alternating in sign starting on negative side if ø1 < 0.	Spike at lag 1, then cuts off to zero. Spike positive if ø1 > 0 and negative side if ø1 < 0.
AR(p)	Exponential decay: pattern depends on signs of ø1, ø2, etc	Spikes at lags 1 to p, then cuts of to zero.

$$Y_t = \phi_0 + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \ldots + \phi_p Y_{t-p} + \varepsilon_t$$

IIMBX

IIMB
तेजस्वि नावधीतमस्तु

Partial Auto-Correlation

- Partial auto-correlation of lag k is auto-correlation between Y_t and Y_{t+k} after the removal of linear dependence of Y_{t+1} to Y_{t+k-1}.

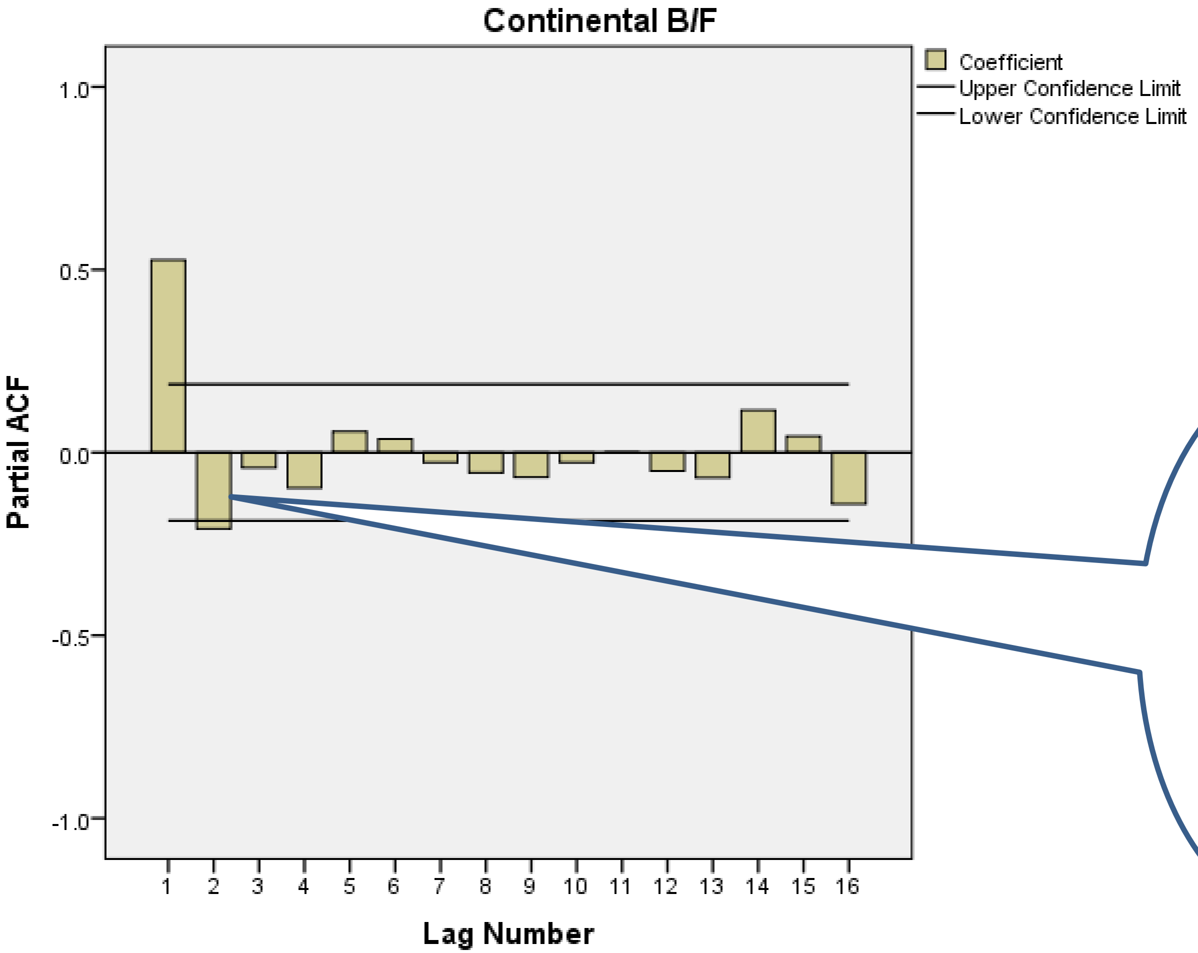
- To test whether the autocorrelation at lag k is significantly different from 0, the following hypothesis test is used:

$$H_0: \rho_k = 0$$

$$H_A: \rho_k \neq 0$$

- For any k, reject H0 if $|\rho_k| > 1.96/\sqrt{n}$. Where n is the number of observations.

PACF Function – DAD Continental Breakfast

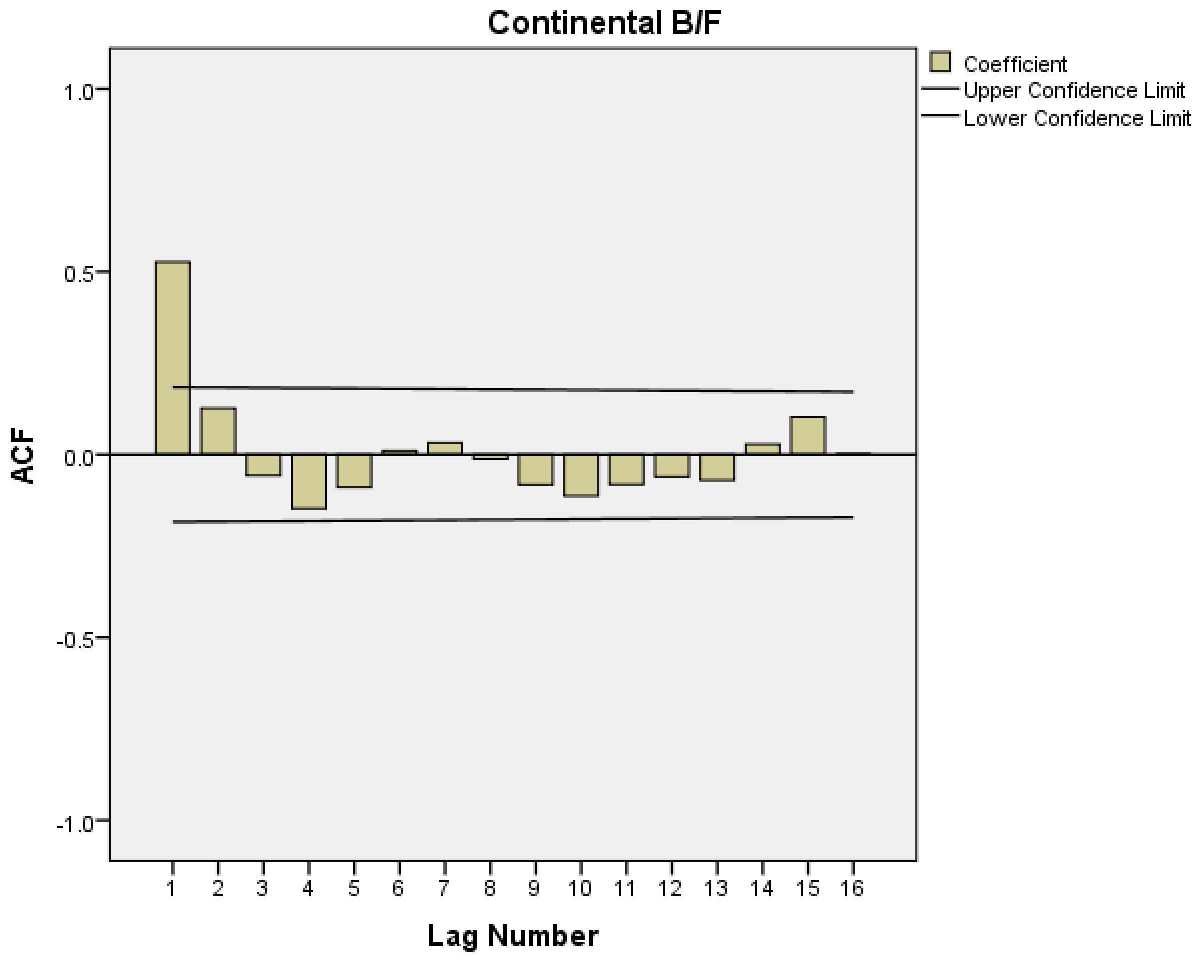


First two PAC are different from zero.

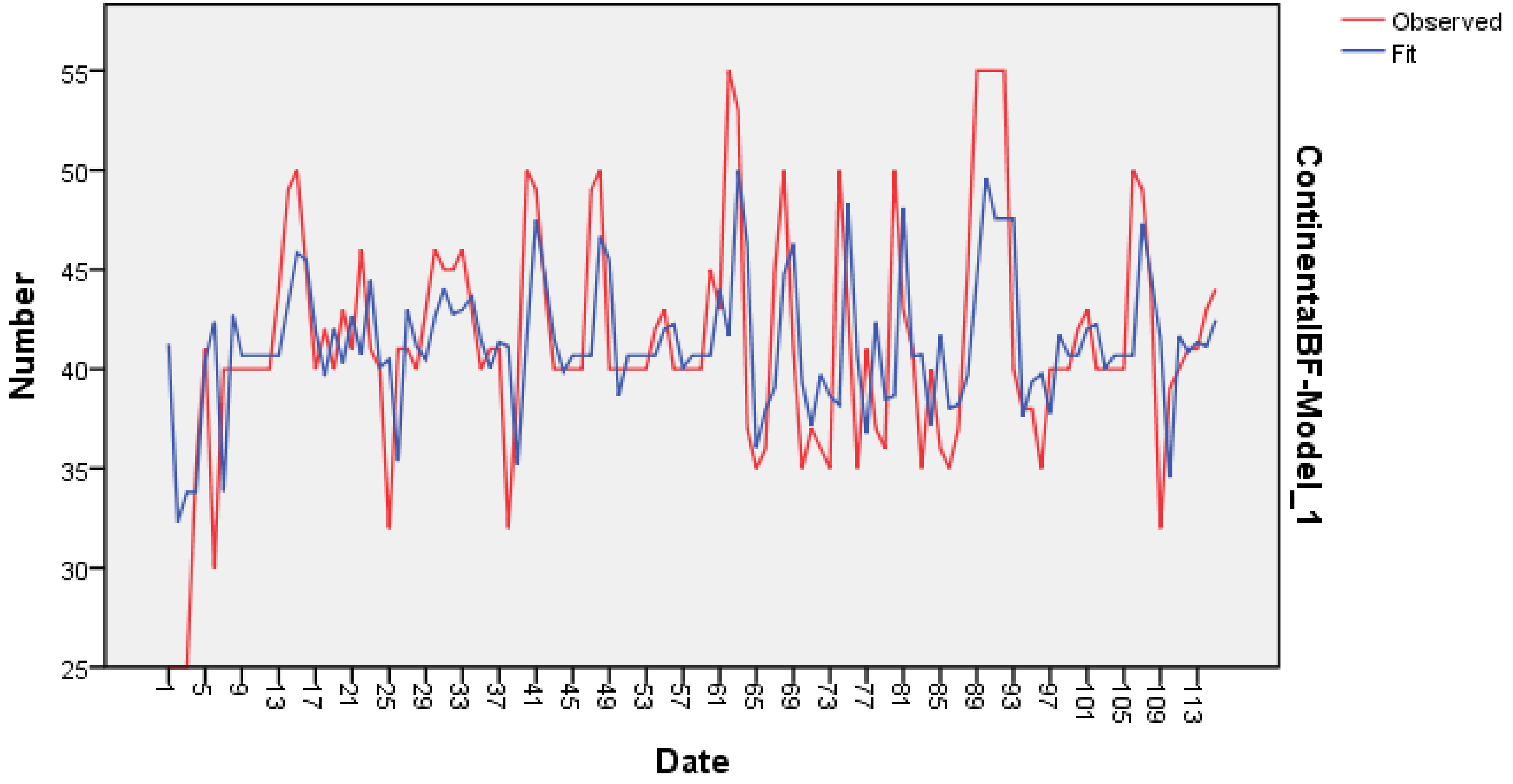
IIMBX

IIMB तेजस्वि नावधीतमस्तु

ACF Function - DAD Continental Breakfast



Actual Vs Fit- Continental Breakfast Data



IIMBX

IIMB
तेजस्वि नावधीतमस्तु

AR(2) Model

ARIMA Model Parameters

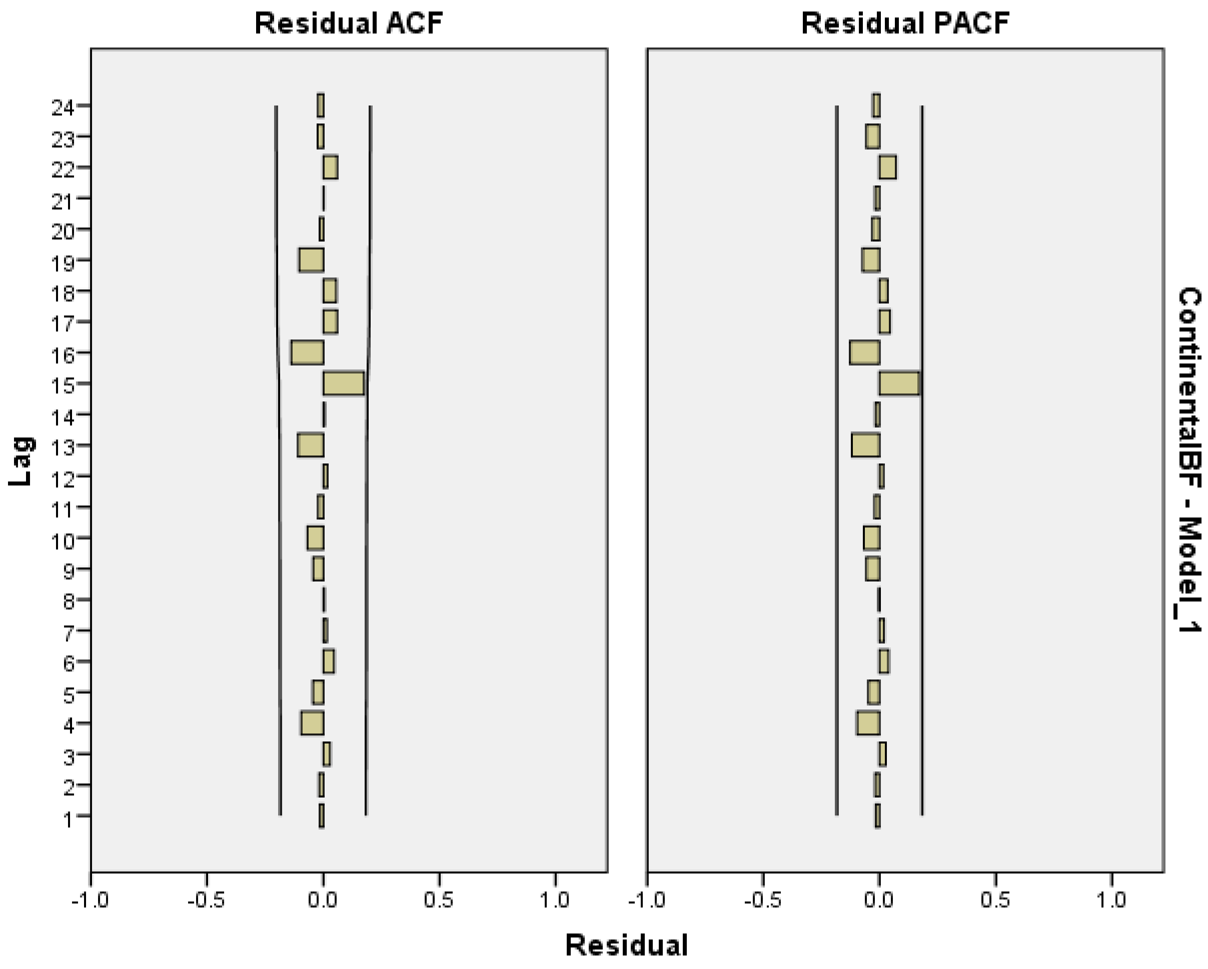
					Estimate	SE	t	Sig.
Continental B/F-Model_1	Continental B/F	No Transformation	Constant		41.252	.824	50.056	.000
			AR	Lag 1	.663	.093	7.117	.000
				Lag 2	-.204	.093	-2.200	.030

Model Fit

Fit Statistic	Mean	SE	Minimum	Maximum	Percentile						
					5	10	25	50	75	90	95
Stationary R-squared	.302	.	.302	.302	.302	.302	.302	.302	.302	.302	.302
R-squared	.302	.	.302	.302	.302	.302	.302	.302	.302	.302	.302
RMSE	4.882	.	4.882	4.882	4.882	4.882	4.882	4.882	4.882	4.882	4.882
MAPE	8.812	.	8.812	8.812	8.812	8.812	8.812	8.812	8.812	8.812	8.812
MaxAPE	65.010	.	65.010	65.010	65.010	65.010	65.010	65.010	65.010	65.010	65.010
MAE	3.486	.	3.486	3.486	3.486	3.486	3.486	3.486	3.486	3.486	3.486
MaxAE	16.252	.	16.252	16.252	16.252	16.252	16.252	16.252	16.252	16.252	16.252
Normalized BIC	3.295	.	3.295	3.295	3.295	3.295	3.295	3.295	3.295	3.295	3.295

MAPE = 8.8%

Residual White Noise



IIMBX

IIMB
तेजस्वि नावधीतमस्तु

ARIMA Model Parameters

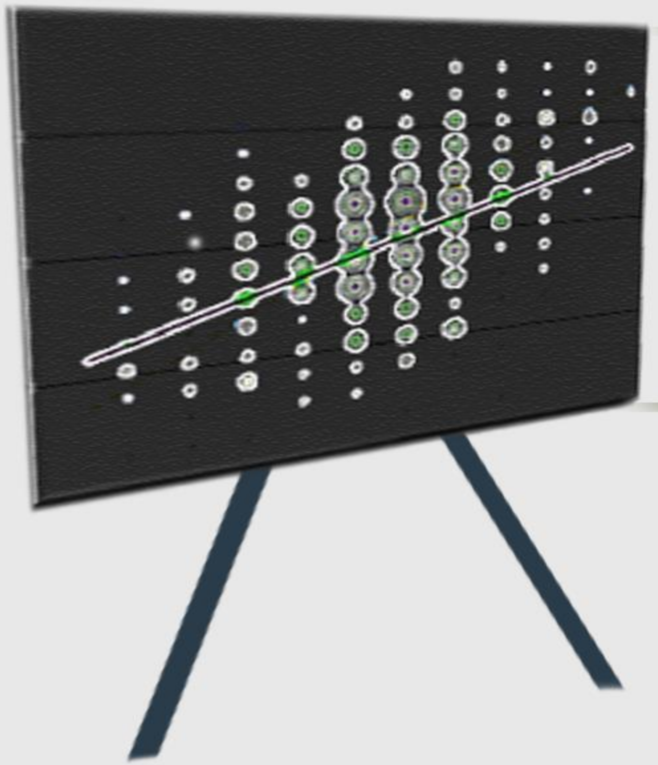
					Estimate	SE	t	Sig.
Continental B/F-Model_1	Continental B/F	No Transformation	Constant		41.252	.824	50.056	.000
			AR	Lag 1	.663	.093	7.117	.000
				Lag 2	-.204	.093	-2.200	.030

$$(Y_t - 41.252) = 0.663\,(Y_{t-1} - 41.252) - 0.204\,(Y_{t-2} - 41.252)$$

$$Y_t = (X_t - \mu)$$

IIMBX

IIMB
तेजस्वि नावधीतमस्तु



Moving average models

U. DINESH KUMAR

Moving Average Process

- A moving average process is a time series regression model in which the value at time t, Y_t, is a linear function of past errors.

- First order moving average, MA(1), is given by:

$$Y_t = \beta_0 + \beta_1 \varepsilon_{t-1} + \varepsilon_t$$

Moving Average Process MA(q)

{Yt} is a moving average process of order *q* (written MA(q)) if for some constants β0, β1, . . . βq

We have.,

$$Y_t = \beta_0 + \beta_1 \varepsilon_t + \beta_2 \varepsilon_{t-2} + \ldots + \beta_q \varepsilon_{t-q} + \varepsilon_t$$

MA(q) models each future observation as a function of "q" previous errors

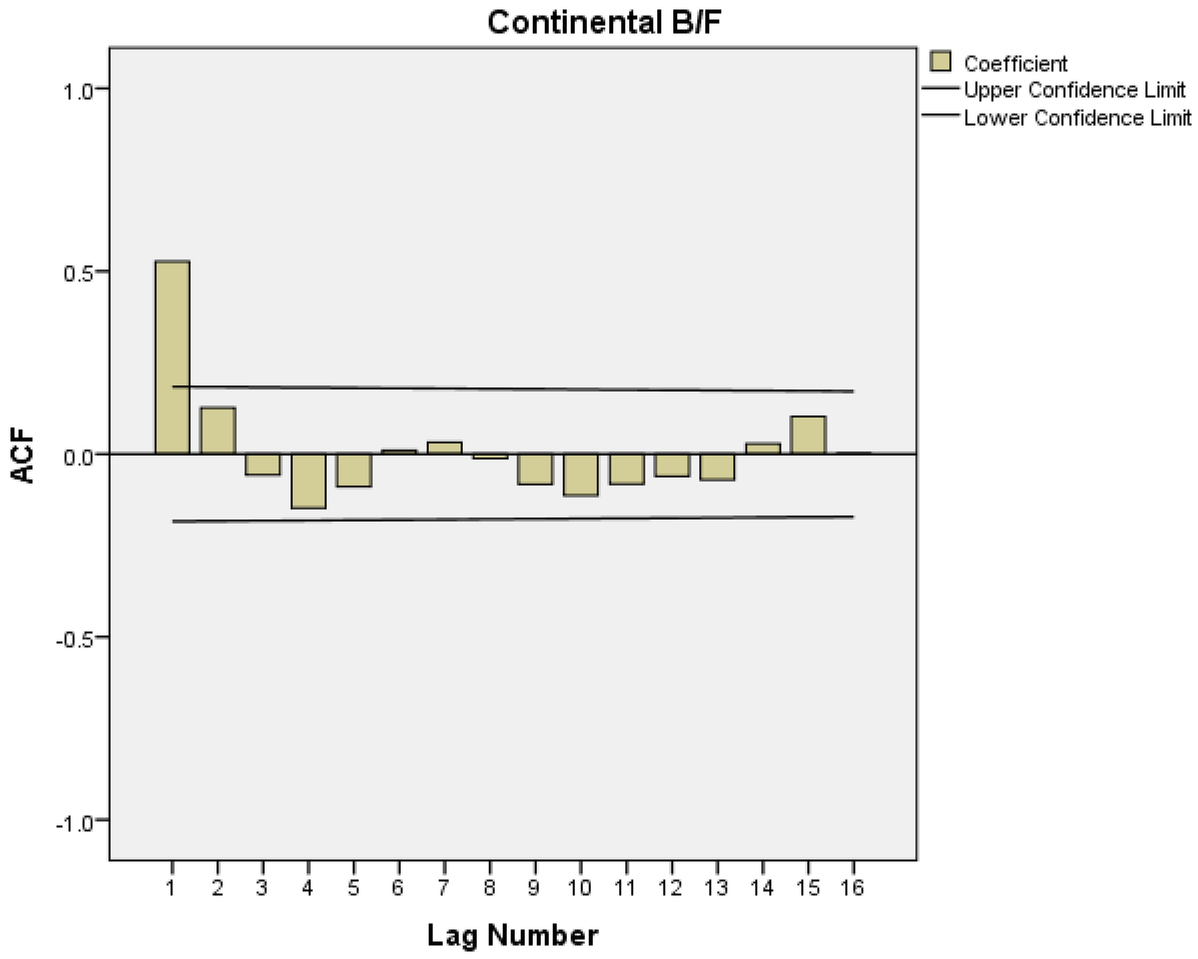
Pure MA Model Identification

Model	ACF	PACF
MA(1)	Spike at lag 1 then cuts of to zero. Spike positive if $\beta1 > 0$ and negative side if $\beta1 < 0$.	Exponential decay. On negative side if $\beta1 > 0$ on positive side if $\beta1 < 0$.
MA(q)	Spikes at lags 1 to q, then cuts off to zero.	Exponential decay or sine wave. Exact pattern depends on signs of $\beta1$, $\beta2$ etc.

IIMBX

IIMB
तेजस्वि नावधीतमस्तु

ACF Function - DAD Continental Breakfast



SPSS Output

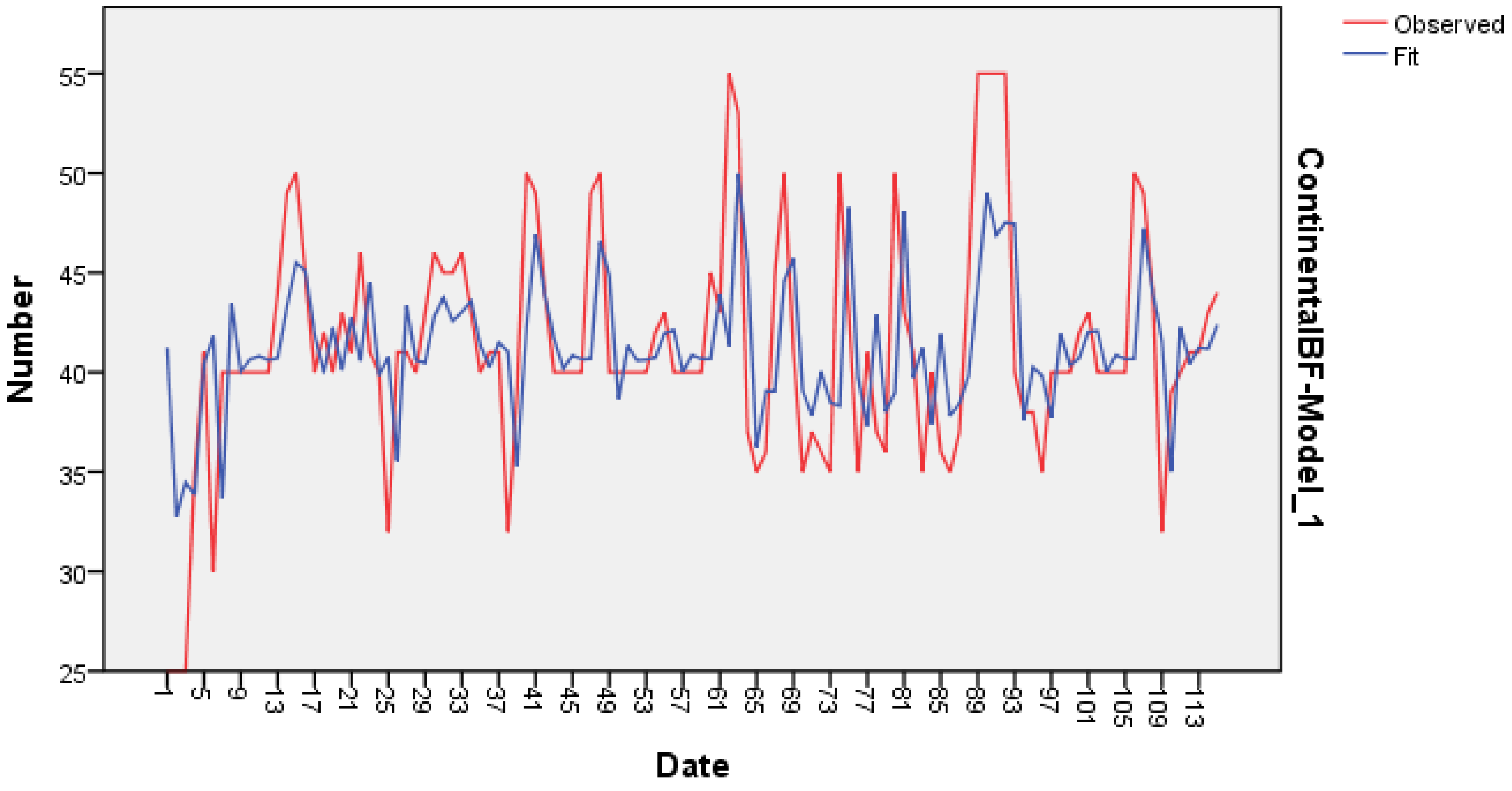
ARIMA Model Parameters

					Estimate	SE	t	Sig.
Continental B/F-Model_1	Continental B/F	No Transformation	Constant		41.240	.812	50.813	.000
			MA	Lag 1	-.651	.094	-6.940	.000
				Lag 2	-.164	.094	-1.751	.083

Model Fit

Fit Statistic	Mean	SE	Minimum	Maximum	Percentile						
					5	10	25	50	75	90	95
Stationary R-squared	.298	.	.298	.298	.298	.298	.298	.298	.298	.298	.298
R-squared	.298	.	.298	.298	.298	.298	.298	.298	.298	.298	.298
RMSE	4.896	.	4.896	4.896	4.896	4.896	4.896	4.896	4.896	4.896	4.896
MAPE	8.897	.	8.897	8.897	8.897	8.897	8.897	8.897	8.897	8.897	8.897
MaxAPE	64.960	.	64.960	64.960	64.960	64.960	64.960	64.960	64.960	64.960	64.960
MAE	3.518	.	3.518	3.518	3.518	3.518	3.518	3.518	3.518	3.518	3.518
MaxAE	16.240	.	16.240	16.240	16.240	16.240	16.240	16.240	16.240	16.240	16.240
Normalized BIC	3.301	.	3.301	3.301	3.301	3.301	3.301	3.301	3.301	3.301	3.301

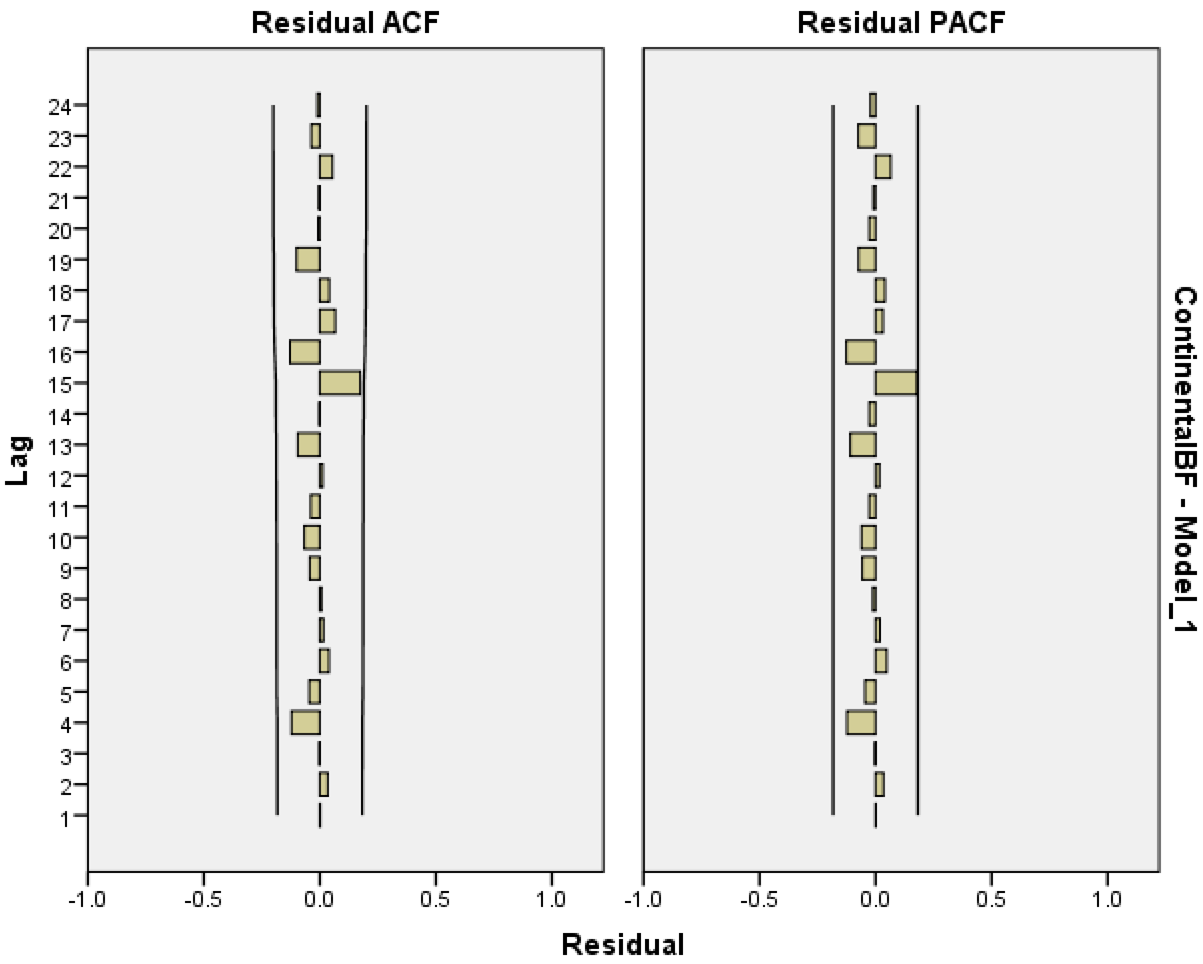
Actual Vs Fitted Value



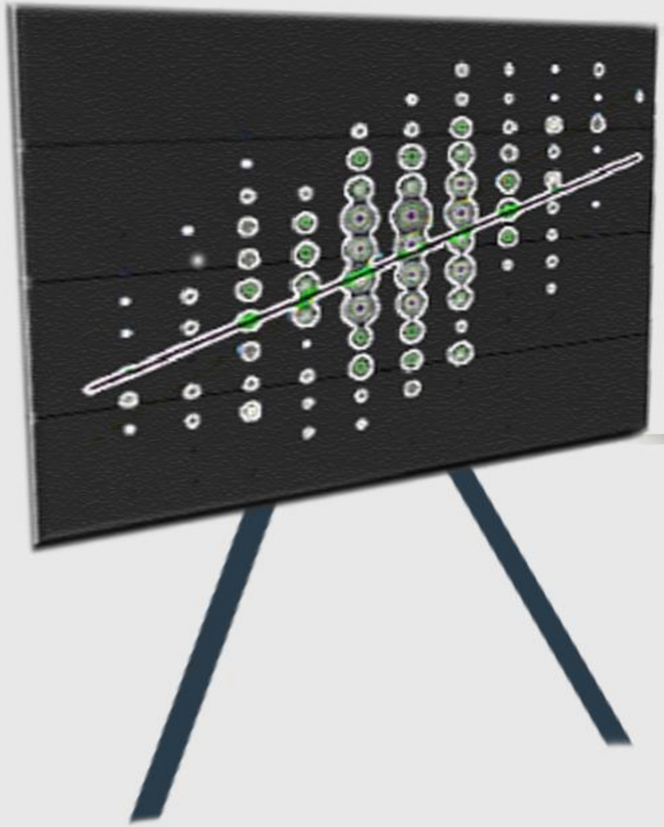
IIMBX

IIMB
तेजस्वि नावधीतमस्तु

Residual Plot



Predictive Analytics : QM901.1x
Prof U Dinesh Kumar, IIMB



ARMA Models

U. DINESH KUMAR

IIMBX

IIMB
तेजस्वि नावधीतमस्तु

ARMA(p,q) Model

AR(p) Model

$$Y_t = \alpha_0 + \alpha_1 Y_{t-1} + \alpha_2 Y_{t-2} + \ldots + \alpha_p Y_{t-p} +$$

$$+ \beta_0 + \beta_1 \varepsilon_{t-1} + \beta_2 \varepsilon_{t-2} + \ldots + \beta_q \varepsilon_{t-q} + \varepsilon_t$$

MA(q) Model

$\alpha_0 + \beta_0$ = C (Constant)

IIMBX

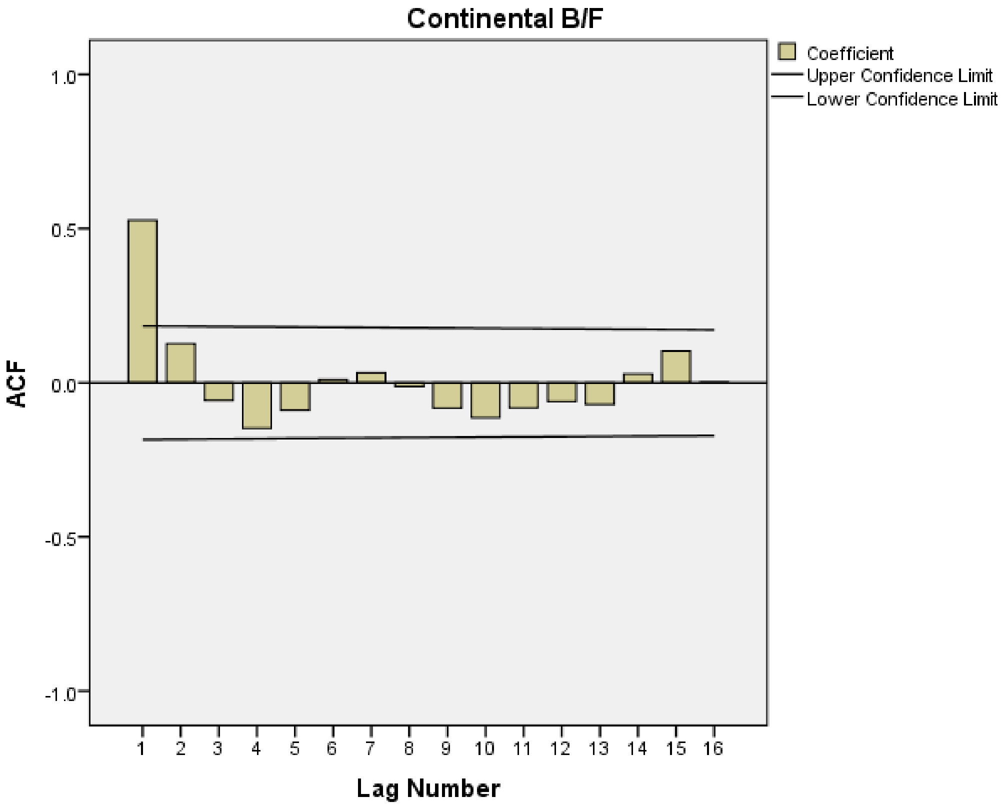
IIMB तेजस्वि नावधीतमस्तु

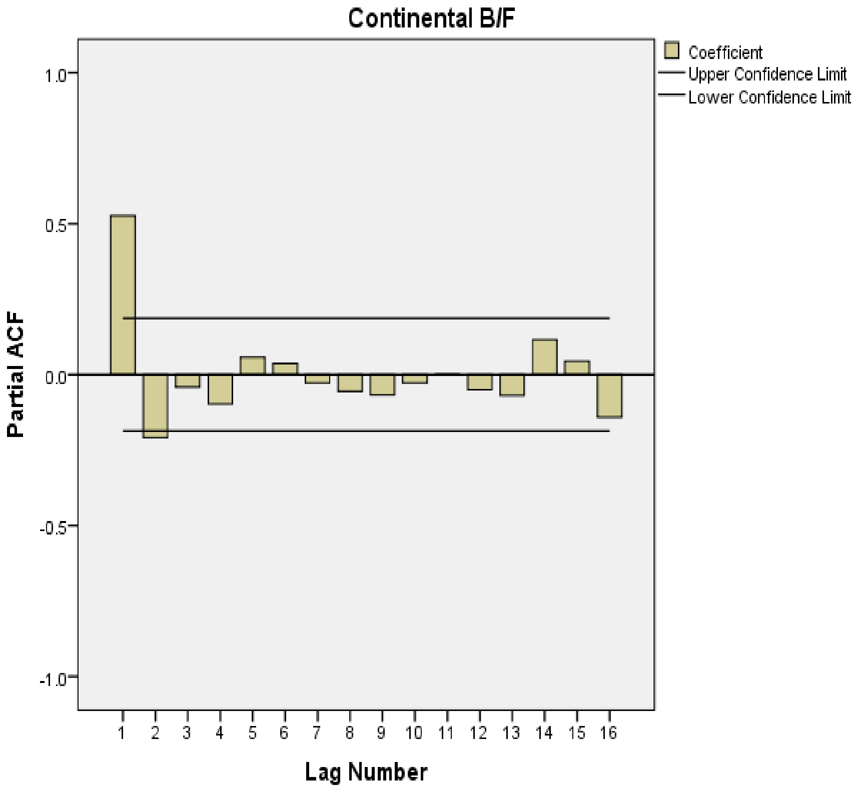
ARMA(p,q) Model Identification

- ARMA(p,q) models are not easy to identify. We usually start with pure AR and MA process. The following thump rule may be used.

Process	ACF	PACF
ARMA(p,q)	Tails off after q lags	Tails off to zero after p lags

- The final ARMA model may be selected based on parameters such as RMSE, MAPE, AIC and BIC.


Continental B/F
Coefficient
Upper Confidence Limit
Lower Confidence Limit
ACF
1.0
0.5
0.0
-0.5
-1.0
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16
Lag Number


Continental B/F
Coefficient
Upper Confidence Limit
Lower Confidence Limit
Partial ACF
1.0
0.5
0.0
-0.5
-1.0
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16
Lag Number

ACF **PACF**

ARMA(2,1) Model Output

ARIMA Model Parameters

					Estimate	SE	t	Sig.
Continental B/F-Model_1	Continental B/F	No Transformation	Constant		41.302	.742	55.698	.000
			AR	Lag 1	.983	.361	2.720	.008
				Lag 2	-.383	.193	-1.988	.049
			MA	Lag 1	.338	.385	.879	.381

Model Fit

Fit Statistic	Mean	SE	Minimum	Maximum	Percentile 5	10	25	50	75	90	95
Stationary R-squared	.304	.	.304	.304	.304	.304	.304	.304	.304	.304	.304
R-squared	.304	.	.304	.304	.304	.304	.304	.304	.304	.304	.304
RMSE	4.899	.	4.899	4.899	4.899	4.899	4.899	4.899	4.899	4.899	4.899
MAPE	8.768	.	8.768	8.768	8.768	8.768	8.768	8.768	8.768	8.768	8.768
MaxAPE	65.207	.	65.207	65.207	65.207	65.207	65.207	65.207	65.207	65.207	65.207
MAE	3.464	.	3.464	3.464	3.464	3.464	3.464	3.464	3.464	3.464	3.464
MaxAE	16.302	.	16.302	16.302	16.302	16.302	16.302	16.302	16.302	16.302	16.302
Normalized BIC	3.343	.	3.343	3.343	3.343	3.343	3.343	3.343	3.343	3.343	3.343

AR(2) Model

ARIMA Model Parameters

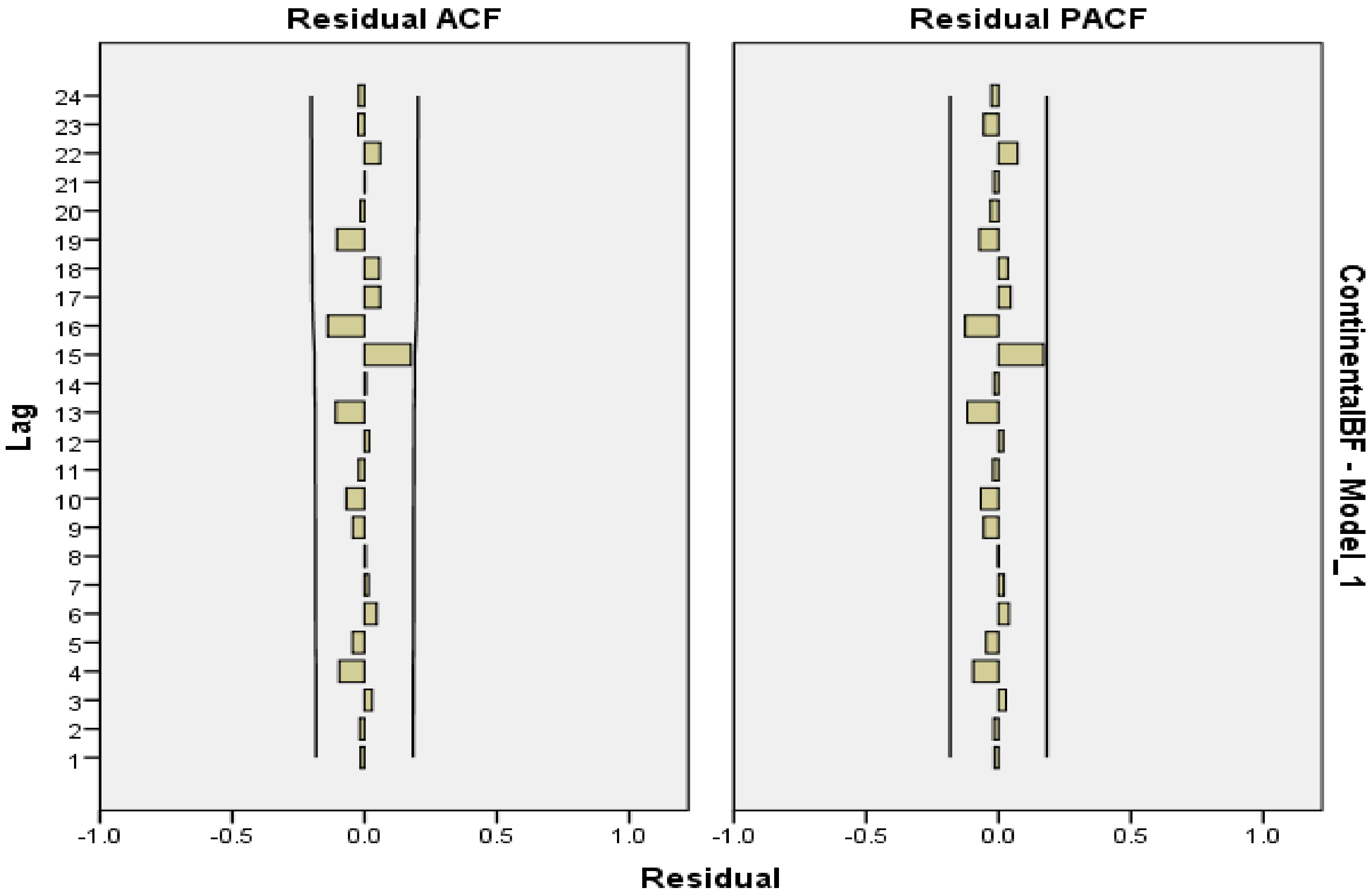
					Estimate	SE	t	Sig.
Continental B/F-Model_1	Continental B/F	No Transformation	Constant		41.252	.824	50.056	.000
			AR	Lag 1	.663	.093	7.117	.000
				Lag 2	-.204	.093	-2.200	.030

Model Fit

Fit Statistic	Mean	SE	Minimum	Maximum	Percentile						
					5	10	25	50	75	90	95
Stationary R-squared	.302	.	.302	.302	.302	.302	.302	.302	.302	.302	.302
R-squared	.302	.	.302	.302	.302	.302	.302	.302	.302	.302	.302
RMSE	4.882	.	4.882	4.882	4.882	4.882	4.882	4.882	4.882	4.882	4.882
MAPE	8.812	.	8.812	8.812	8.812	8.812	8.812	8.812	8.812	8.812	8.812
MaxAPE	65.010	.	65.010	65.010	65.010	65.010	65.010	65.010	65.010	65.010	65.010
MAE	3.486	.	3.486	3.486	3.486	3.486	3.486	3.486	3.486	3.486	3.486
MaxAE	16.252	.	16.252	16.252	16.252	16.252	16.252	16.252	16.252	16.252	16.252
Normalized BIC	3.295	.	3.295	3.295	3.295	3.295	3.295	3.295	3.295	3.295	3.295

MAPE = 8.8%

Residual White Noise



IIMBX

IIMB
तेजस्वि नावधीतमस्तु

Forecasting Model Evaluation

Akaike's information criteria

$$\boxed{AIC = -2LL + 2m}$$

Where m is the number of variables estimated in the model

Bayesian Information Criteria

$$\boxed{BIC = -2LL + m \ln(n)}$$

AIC and BIC can be interpreted as distance from true model

Where m is the number of variables estimated in the model and n is the number of observations

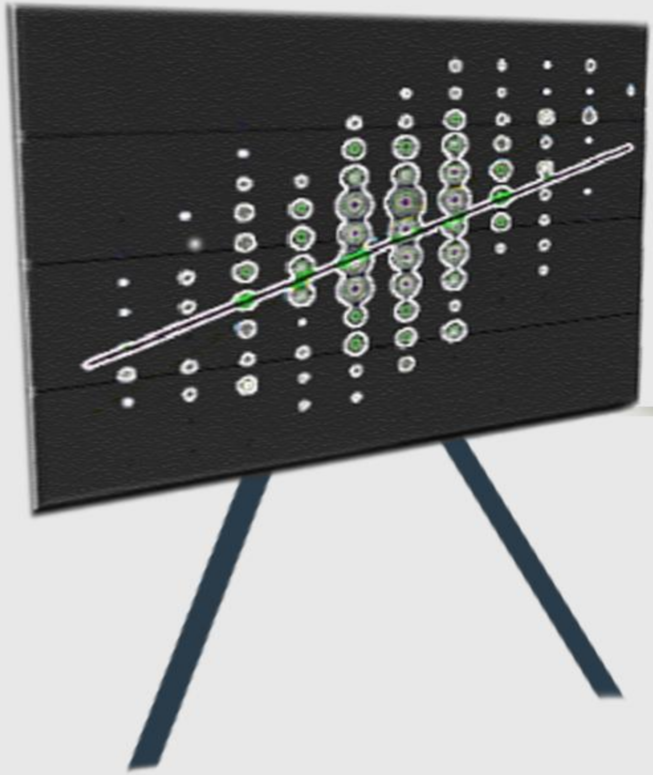
Final Model Selection

Model	AR(2)	MA(1)	ARMA(2,1)
BIC	3.295	3.301	3.343

AR(2) has the least BIC

IIMBX

IIMB
तेजस्वि नावधीतमस्तु



ARIMA Models

U. DINESH KUMAR

ARIMA

ARIMA has the following three components:

- **Auto-regressive component**: Function of past values of the time series.
- **Integration Component**: Differencing the time series to make it a stationary process.
- **Moving Average Component**: Function of past error values.

IIMBX

IIMB
तेजस्वि नावधीतमस्तु

Integration (d)

- Used when the process is non-stationary.
- Instead of observed values, differences between observed values are considered.
- When d=0, the observations are modelled directly. If d = 1, the differences between consecutive observations are modelled. If d = 2, the differences of the differences are modelled.

ARIMA (p, d, q)

- The q and p values are identified using auto-correlation function (ACF) and Partial auto-correlation function (PACF) respectively. The value d identifies the level of differencing.

- Usually p+q <= 4 and d <= 2.

Differencing

- Differencing is a process of making a non-stationary process into stationary process.

- In differencing, we create a new process X_t, where $X_t = Y_t - Y_{t-1}$.

ARIMA(p,1,q) Process

$$X_t = \overline{\alpha_0 + \alpha_1 X_{t-1} + \alpha_2 X_{t-2} + \ldots + \alpha_p X_{t-p} +}$$
$$\underline{+ \beta_0 + \beta_1 \varepsilon_{t-1} + \beta_2 \varepsilon_{t-2} + \ldots + \beta_q \varepsilon_{t-q}} + \varepsilon_t$$

Where $X_t = Y_t - Y_{t-1}$

$$\boxed{\alpha_0 + \beta_0 = \text{C (Constant)}}$$

IIMBX

IIMB
तेजस्वि नावधीतमस्तु

Ljung-Box Test

- Ljung-Box is a test on the autocorrelations of residuals. The test statistic is:

$$Q_m = n(n+2)\sum_{k=1}^{m} \frac{r_k^2}{n-k}$$

n = number of observations in the time series.
k = particular time lag checked
m = the number of time lags to be tested
r_k = sample autocorrelation function of the k^{th} residual term.

H0: The model does not exhibit lack of fit
HA: The model exhibits lack of fit

Q statistic is approximate chi-square distribution with m – p – q degrees of freedom if ARMA orders are correctly specified.

AR(2) Ljung-Box Test

Model Statistics

Model	Number of Predictors	Model Fit statistics	Ljung-Box Q(18)			Number of Outliers
		Stationary R-squared	Statistics	DF	Sig.	
Continental B/F-Model_1	0	.302	12.061	16	.740	0

P-value is 0.740, thus the model doesn't show lack of fit

IIMBX

IIMB
तेजस्वि नावधीतमस्तु

Box-Jenkins Methodology

- Identification: Identify the ARIMA model using ACF & PACF plots. This would give the values of p, q and d.
- Estimation: Estimate the model parameters (using maximum likelihood)
- Diagnostics: Check the residual for any issue such as not providing White Noise.