Introduction to Time Series Analysis

Chalermrat Nontapa

Outline

1. Objectives and Type of Forecasting
2. Components of a time series data
3. Classical Time Series and Estimation of Component
   3.1 Moving Average Method
   3.2 Least Square Method
   3.3 Seasonal Estimation
4. Box and Jenkins Method
5. Evaluation Model
6. Interpret the results for time series data with some of my research on classical time series.

1. Objectives and Type of Forecasting

Objectives

1.1 To analyze time series data in order to extract meaningful statistics and other characteristics of the data (find a mathematical model to explain time series data).

1.2 To use a model to predict future values based on previously observed values.

The benefit of the study

To help management for taking correct decisions. By providing a logical basis for planning and determining in advance the nature of future business planning, economic, and inventory management.

Type of Forecasting

- Forecasting
  - Qualitative Forecasting
  - Quantitative Forecasting
    - Regression Analysis
    - Time Series Analysis

Qualitative Forecasting

Qualitative Forecasting is a method of making predictions about a company's finances that uses judgment from experts.

Delphi is based on the principle that forecasts (or decisions) from a structured group of individuals are more accurate than those from unstructured groups.

Expert Methods is completed by an experienced UX & Usability expert, who is trained and experienced in the field and analyses the product.

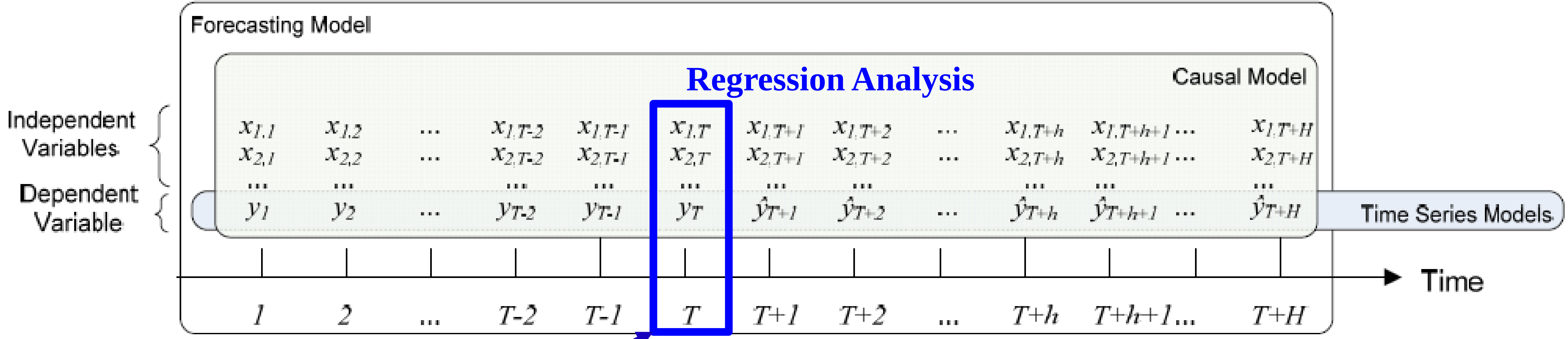
Quantitative Forecasting

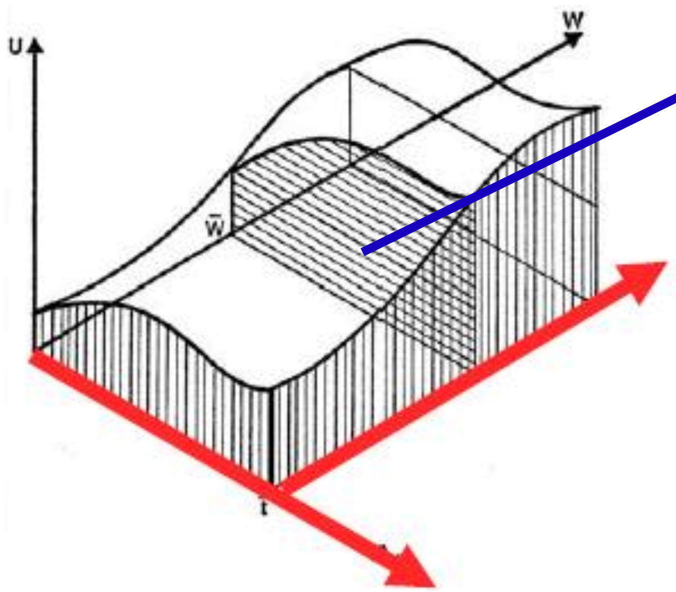
The quantitative forecasting method makes predictions about future demand based on the analysis of past historical data.

1. **Regression Analysis** is a set of statistical methods used for the estimation of relationships between a dependent variable and one or more independent variables.

2. **Time Series Forecasting** is a technique for the prediction of events through a sequence of time.

Time Series Analysis Table View





Statistical relationship is **unidirectional**

$$y_{t+1} \longleftarrow \{y_t, y_{t-1}, \dots, y_1\}$$

Dynamics of series y

values from its own series

exogenous series

$$y_{t+1} \longleftarrow \{y_t, y_{t-1}, \dots, y_1, \mathbf{X_t}, \mathbf{X_{t-1}}, \dots, \mathbf{X_1}\}$$

Three aspects of forecasting model are

1. **Sample data:** the data that we collect that describes our problem with known relationships between inputs and outputs.

2. **Learn a model:** the algorithm that we use on the sample data to create a model that we can later use over and over again.

3. **Making predictions:** the use of our learned model on new data for which we don't know the output.

Advantage & Challenge of the models

Time Series Model	Deep Learning Model
Advantages - Computational lower order of models - Converse guarantee	**Advantages** - Not dependent on linear stationary process - Detect all possible, nonlinear relationship
Challenges - Assume linear stationary process - Expensive for higher models	**Challenges** - Selection free hyperparameter - Not guaranteed to converse optimal solution

Time Series

Time series is a collection of observations over a period of time T.

$$D = \{ y_1, y_2, \ldots, y_T \} \quad \text{where} \quad t = 1, 2 \ldots, T$$

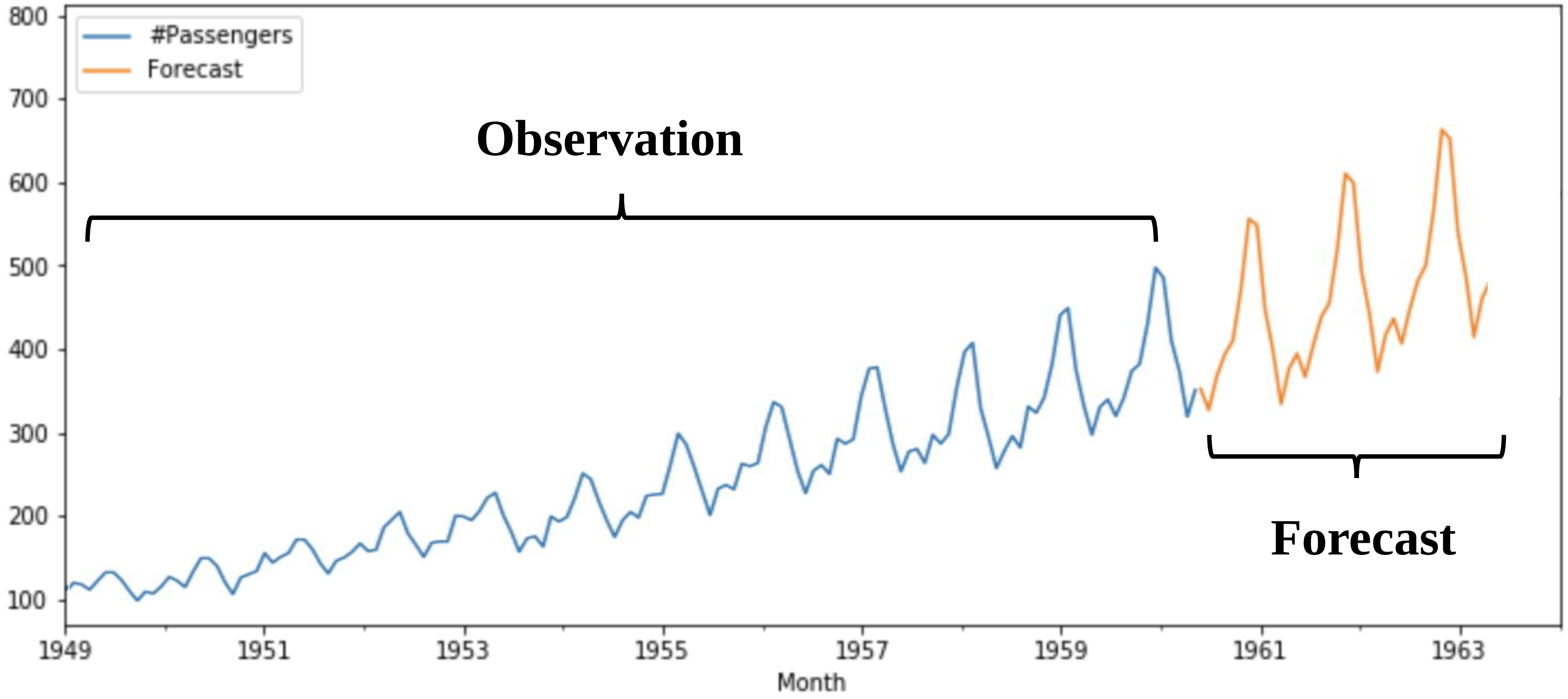


Figure 1.1: Time sequential plot of passenger dataset

The main aim of time series modeling is to study past observations and generate future values based on long-term forecasting and another variation.

Example of Time Series data set

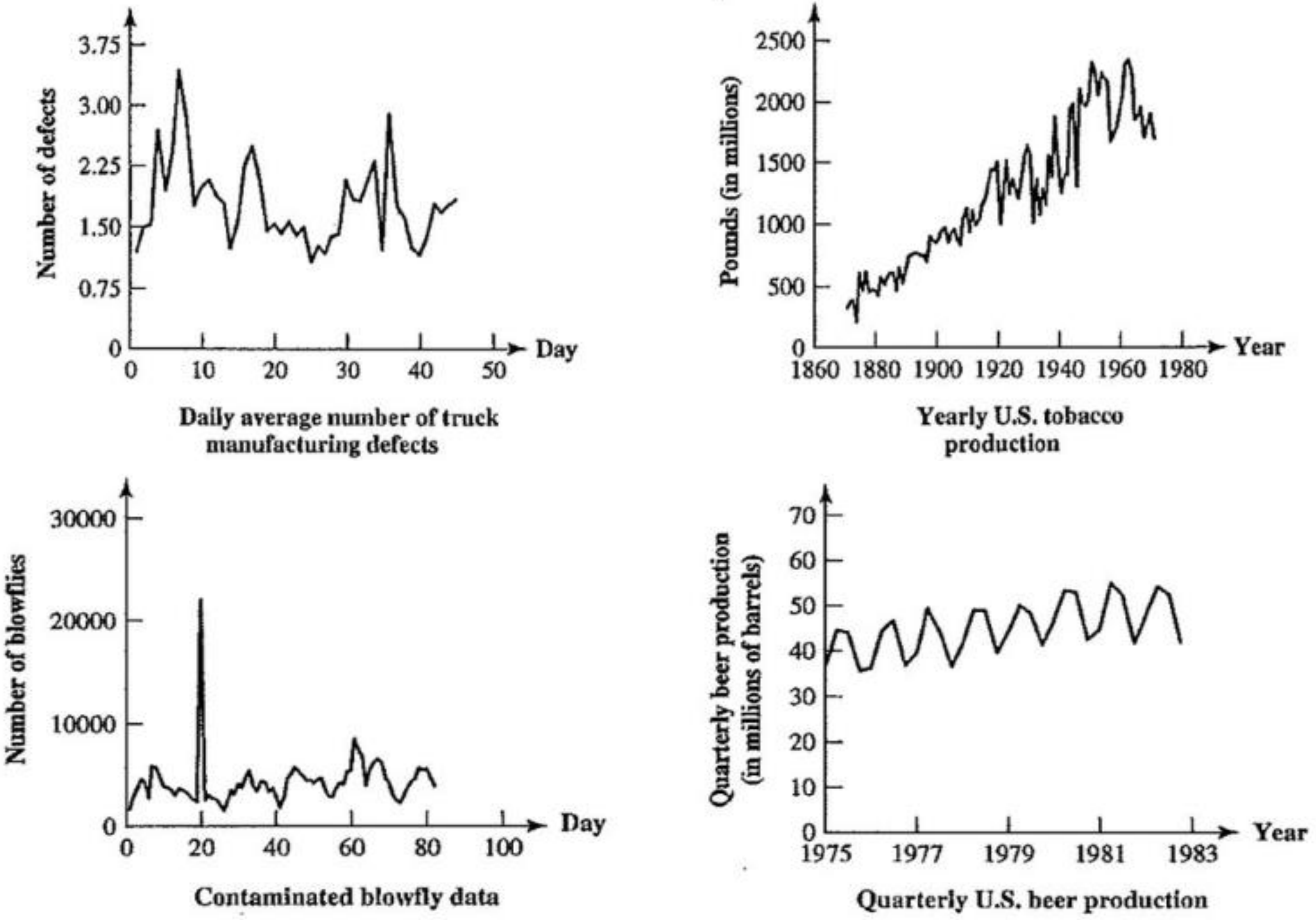


Figure 1.2: Example of time Series data

Time Series Analysis is used for many applications

- Sales Forecasting
- Economic Forecasting
- Stock Market Analysis
- Budgetary Analysis

2. Component of Time Series

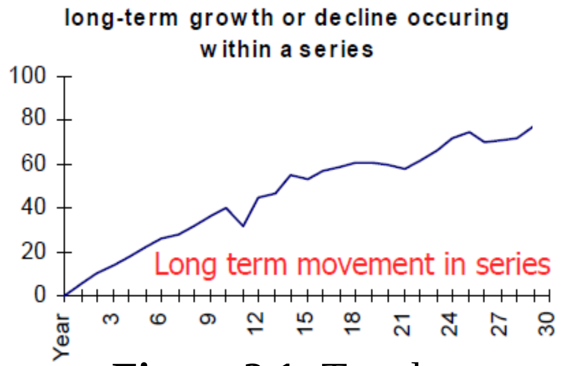


Figure 2.1: Trend

Trend is a general direction of the time series data over a long period of time.

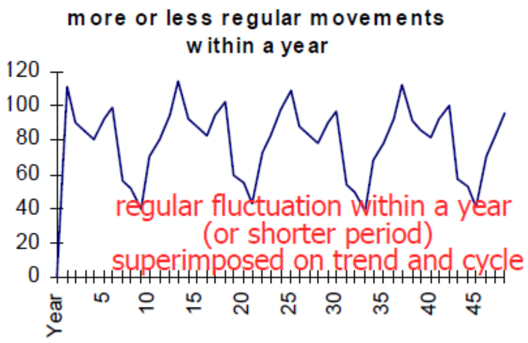


Figure 2.2: Seasonal

Seasonal component exhibits a trend that repeats with respect to timing, direction, and magnitude.

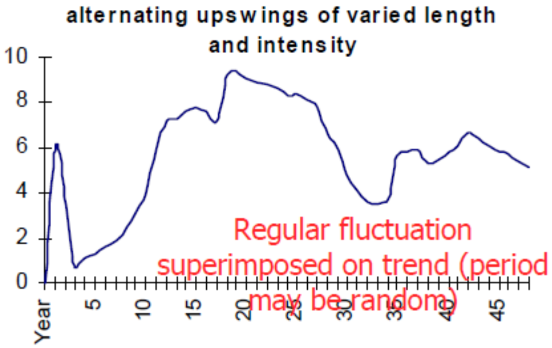


Figure 2.3: Cycle

Cycle the period of ups and downs, booms, and slums of a time series mostly observed in a time period of 3 to 12 years.

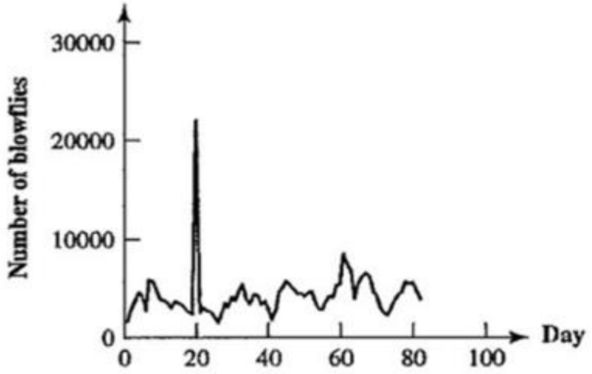


Figure 2.4: Irregular

Irregular is an unpredictable pattern and there is no regular period of time over which they occur.

3. Classical Time Serie Model

This is the classical model which separates four variation components in common forms are known as additive and multiplicative **decomposition model**, which are expressed in equations (3.2) and (3.3), respectively.

$$\boxed{y_t = f(T_t, S_t, C_t, I_t) + \varepsilon_t} \quad \text{....(3.1)}$$

$$y_t = (T_t + S_t + C_t + I_t) + \varepsilon_t \quad \text{....(3.2)}$$

$$y_t = (T_t \times S_t \times C_t \times I_t) + \varepsilon_t \quad \text{....(3.3)}$$

Techniques for fitting Trends (T_t)

The decomposition of time series is a statistical task that deconstructs a time series into several components, each representing one of the underlying categories of patterns. When we decompose a time series into components, we think of a time series as comprising three components: a trend component, a seasonal component, and residuals or noise” (containing anything else in the time series).

1. Moving Average Method

2. Least Square Method

Moving Averages Method

This method is based on the premise that if values in a time series are averaged over a sufficient period, the effect of short-term variations will be reduced. The technique for finding a moving average for a particular observation is to find the average of the *m* observations before and after the observation itself. The *k*-period simple moving averages are defined as

$$a_1 = \frac{1}{k}\sum_{i=1}^{k} y_i,\ \ a_2 = \frac{1}{k}\sum_{i=2}^{k+1} y_i,\ a_3 = \frac{1}{k}\sum_{i=3}^{k+2} y_i,\ \ldots\ , a_m = \frac{1}{k}\sum_{i=m}^{n} y_i \qquad \text{….(3.4)}$$

- When *k* is odd, the sequence a_1, a_2, a_3, … will be placed against the middle of its time period.
- When *k* is even, the sequence a_1, a_2, a_3, … of simple moving averages will be placed in the middle of two time periods. It is necessary to centralize these averages.

Moving Averages Method

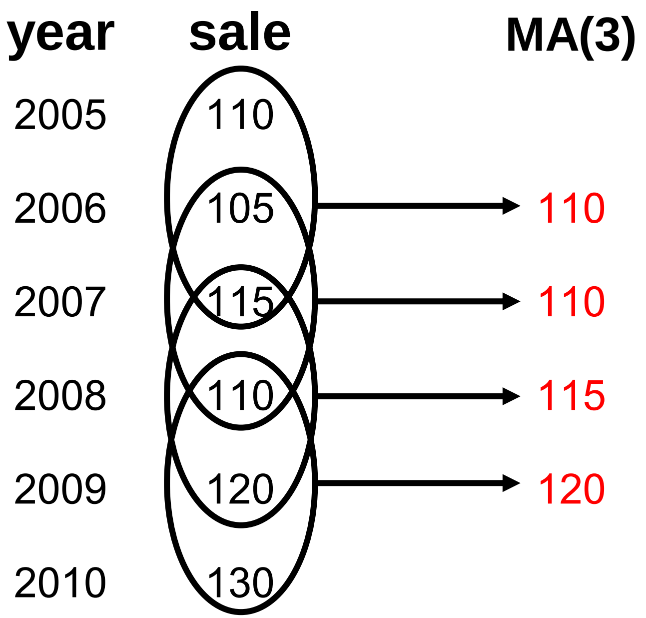
This process is continued till the last k observations have been averaged. For example, the 3-period simple moving averages are given as:

$$
\begin{aligned}
a_1 &= \frac{1}{3}(y_1 + y_2 + y_3) = \frac{1}{3}\sum_{i=1}^{3} y_i \\
a_2 &= \frac{1}{3}(y_2 + y_3 + y_4) = \frac{1}{3}\sum_{i=2}^{4} y_i \\
a_3 &= \frac{1}{3}(y_3 + y_4 + y_5) = \frac{1}{3}\sum_{i=3}^{5} y_i \\
\vdots &= \quad \vdots \\
\text{and so on}
\end{aligned}
$$

Each of these simple moving averages of the sequence $a_1, a_2, a_3, \ldots$ is placed against the middle of each successive group.

Example For the following data finds a trend by the moving averages method.

Year	2005	2006	2007	2008	2009	2010
Sale	110	105	115	110	120	130



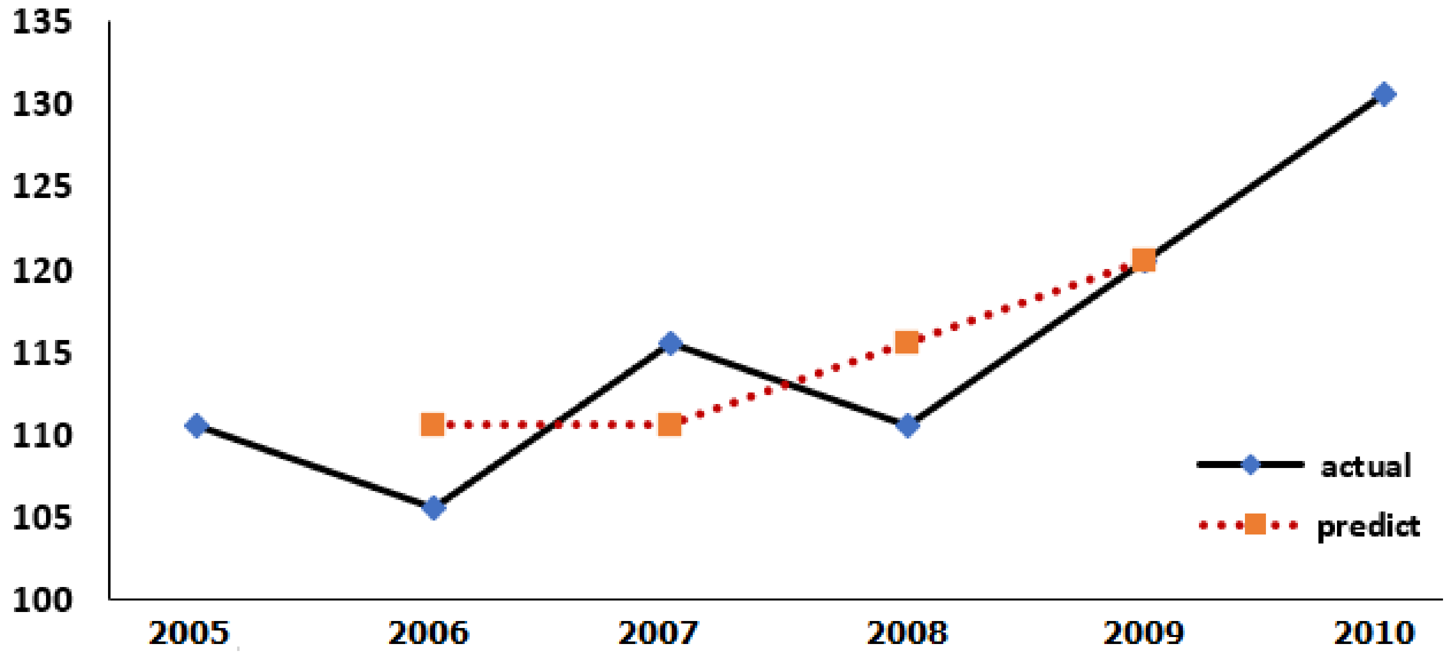


Figure 3.1: Predicted value by the moving average method

From the table, we have

(110+105+115)/3 = 110 (105+115+110)/3 = 110

(115+110+120)/3 = 115 (110+120+130)/3 = 120.

Least Square Method

Finding the best-fitting linear line by finding a and b that minimize the sum of the squares of the errors.

Minimize $$\boxed{\boldsymbol{E}(\boldsymbol{a},\boldsymbol{b}) = \sum_{i=1}^{n}(y_i - \hat{y}_i)^2 = \sum_{i=1}^{n}[(y_i - (a + bt_i)]^2} \qquad \text{….(3.5)}$$

with condition, $\dfrac{\partial E}{\partial a} = 0$ and $\dfrac{\partial E}{\partial b} = 0$

Hence, $$\boxed{a = \bar{y} - b\bar{t} \quad \text{and} \quad b = \frac{n(\sum_{i=1}^{n} t_i y_i) - (\sum_{i=1}^{n} t_i)(\sum_{i=1}^{n} y_i)}{n(\sum_{i=1}^{n} t_i^2) - (\sum_{i=1}^{n} t_i)^2}} \qquad \text{….(3.6)}$$

Steps for finding linear trends using the least square method

1. Consider the number of time series data (*n*) to set the origin of time *t*

1.1 If the number of time series data (*n*) is an even number

Set the time of two data points in the middle to be -1 and1 and the time before are -3,-5,-7,… and so on, and the time after are 3, 5, 7,… and so on.

1.2 If the number of time series data (*n*) is an odd number

Set the time of data in the middle to be 0 and the time before are -1,-2,-3,… and so on and the time after are 1, 2, 3,… and so on.

2. Compute *a* and *b*

3. Obtain (2.) to the linear trend equation

3.1 define the origin time at $t = 0$.

3.2 set up a unit of time t and y_t

Example For the following data finds a trend by the least square method.

Year	2005	2006	2007	2008	2009	2010
Sale	110	105	115	110	120	130

y	t	t^2	$y \times t$	$\hat{y}_t = 115+2(t)$
110	-5	25	-550	105
105	-3	9	-315	109
115	-1	1	-115	113
110	1	1	110	117
120	3	9	360	121
130	5	25	650	125
690	0	70	140	

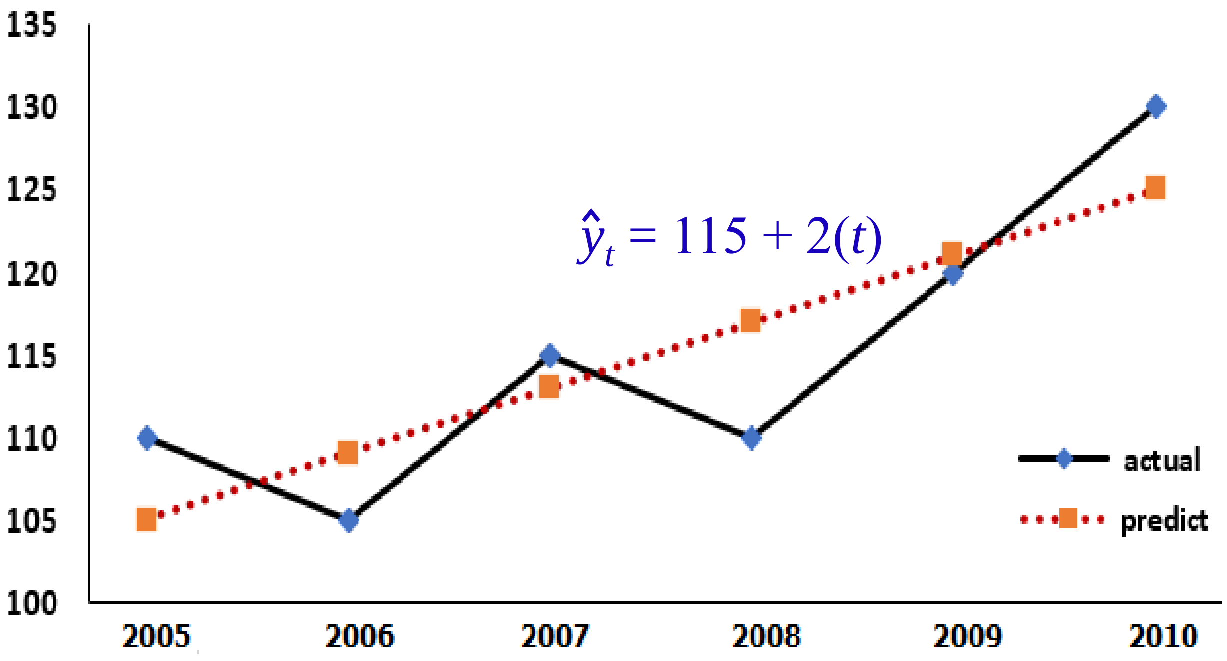


Figure 3.2: Predicted value by the least square method

From the table, we have a = 690/6 = 115 and b = 140/70 = 2.

Therefore, we have $\hat{y}_t = 115 + 2(t)$

Note that 1. origin is on 1 January 2008
2. time unit is 1/2 year
3. unit of y_t is million baht

Seasonal Approximation (S_t)

1. Obtain the trend equation and compute the trend values (y_t), by substituting them into the equation.

2. Calculate the ratio of y_t and $\bar{y}_t$ for each year.

$$S_t(\%) = \frac{y_t}{\bar{y}_t} \times 100 \qquad \text{....(3.7)}$$

3. Find the average percentage ratio for each season and the average obtained will be the seasonal index for that season (sometimes need an adjustment).

Note that, each season will have the same seasonal index.

Example The following calculator sales figures were recorded by ALL Correct Company, a calculator manufacturer, for each of the years from 1994 – 1997.

Using the simple average method determine the seasonal index for each quarter.

Year	Q_1	Q_2	Q_3	Q_4	Avg. Yearly
1994	2.8	2.1	4.0	4.5	3.35
1995	3.8	3.2	4.8	5.4	4.30
1996	4	3.6	5.5	5.8	4.73
1997	4.3	3.9	6.0	6.4	5.15

Year	Q_1	Q_2	Q_3	Q_4
1994	0.84	0.63	1.19	1.34
1995	0.88	0.74	1.12	1.26
1996	0.85	0.76	1.16	1.23
1997	0.83	0.76	1.17	1.24
Avg. Seasonal	0.85	0.72	1.16	1.27

From the table, we have

For 1994, $S_{Q1} = 2.8/3.35 = 0.84$

$S_{Q2} = 2.1/3.35 = 0.63$

$S_{Q3} = 4.0/3.35 = 1.19$

$S_{Q4} = 4.5/3.35 = 1.34$

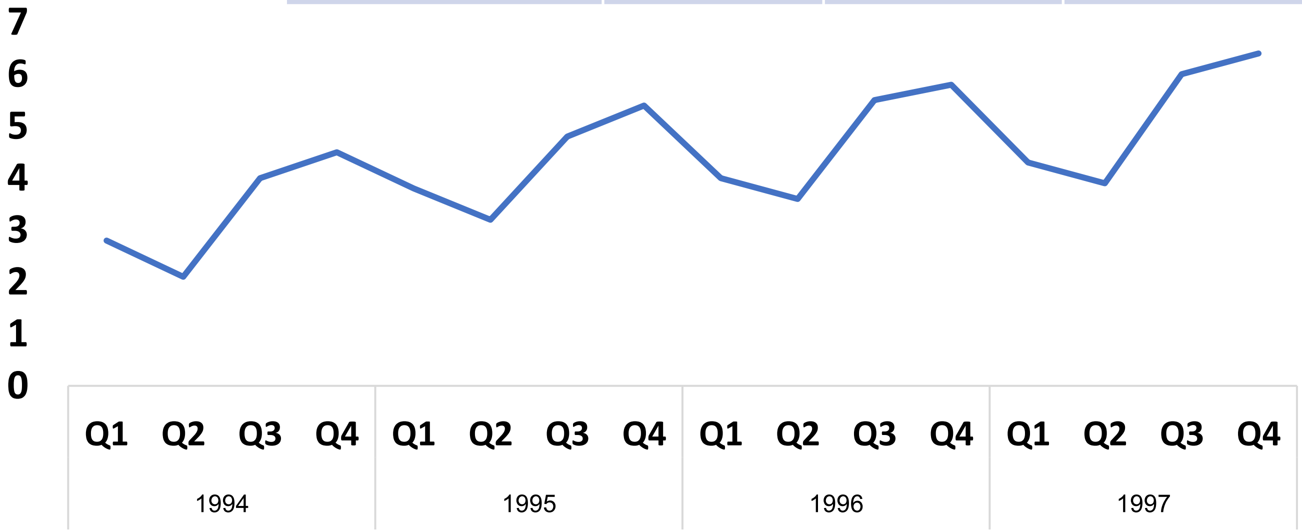


Figure 3.3: Predicted value by the least square method

Trend estimation using moving average and least square method.

Moving average method

Least square method

$$\hat{y}_t = \hat{T}_t \times \hat{S}_t$$

Year	Quarter	yt	MA(4)	CMA	yt/CMA	St	De-St	t	t^2	y * t	y(t) = 4.38 + 0.09t	Predict
1994	Q1	2.8				0.85	3.29	-15	225	-42	3.03	2.58
	Q2	2.1	3.4			0.72	2.92	-13	169	-27.3	3.21	2.31
	Q3	4	3.6	3.5	1.15	1.16	3.45	-11	121	-44	3.39	3.93
	Q4	4.5	3.9	3.7	1.20	1.27	3.54	-9	81	-40.5	3.57	4.53
1995	Q1	3.8	4.1	4.0	0.96	0.85	4.47	-7	49	-26.6	3.75	3.19
	Q2	3.2	4.3	4.2	0.76	0.72	4.44	-5	25	-16	3.93	2.83
	Q3	4.8	4.4	4.3	1.11	1.16	4.14	-3	9	-14.4	4.11	4.77
	Q4	5.4	4.5	4.4	1.23	1.27	4.25	-1	1	-5.4	4.29	5.45
1996	Q1	4	4.6	4.5	0.88	0.85	4.71	1	1	4	4.47	3.80
	Q2	3.6	4.7	4.7	0.77	0.72	5.00	3	9	10.8	4.65	3.35
	Q3	5.5	4.8	4.8	1.15	1.16	4.74	5	25	27.5	4.83	5.60
	Q4	5.8	4.9	4.8	1.20	1.27	4.57	7	49	40.6	5.01	6.36
1997	Q1	4.3	5.0	4.9	0.87	0.85	5.06	9	81	38.7	5.19	4.41
	Q2	3.9	5.2	5.1	0.77	0.72	5.42	11	121	42.9	5.37	3.87
	Q3	6				1.16	5.17	13	169	78	5.55	6.44
	Q4	6.4				1.27	5.04	15	225	96	5.73	7.28

A sequence plot between the actual value and predictive value.

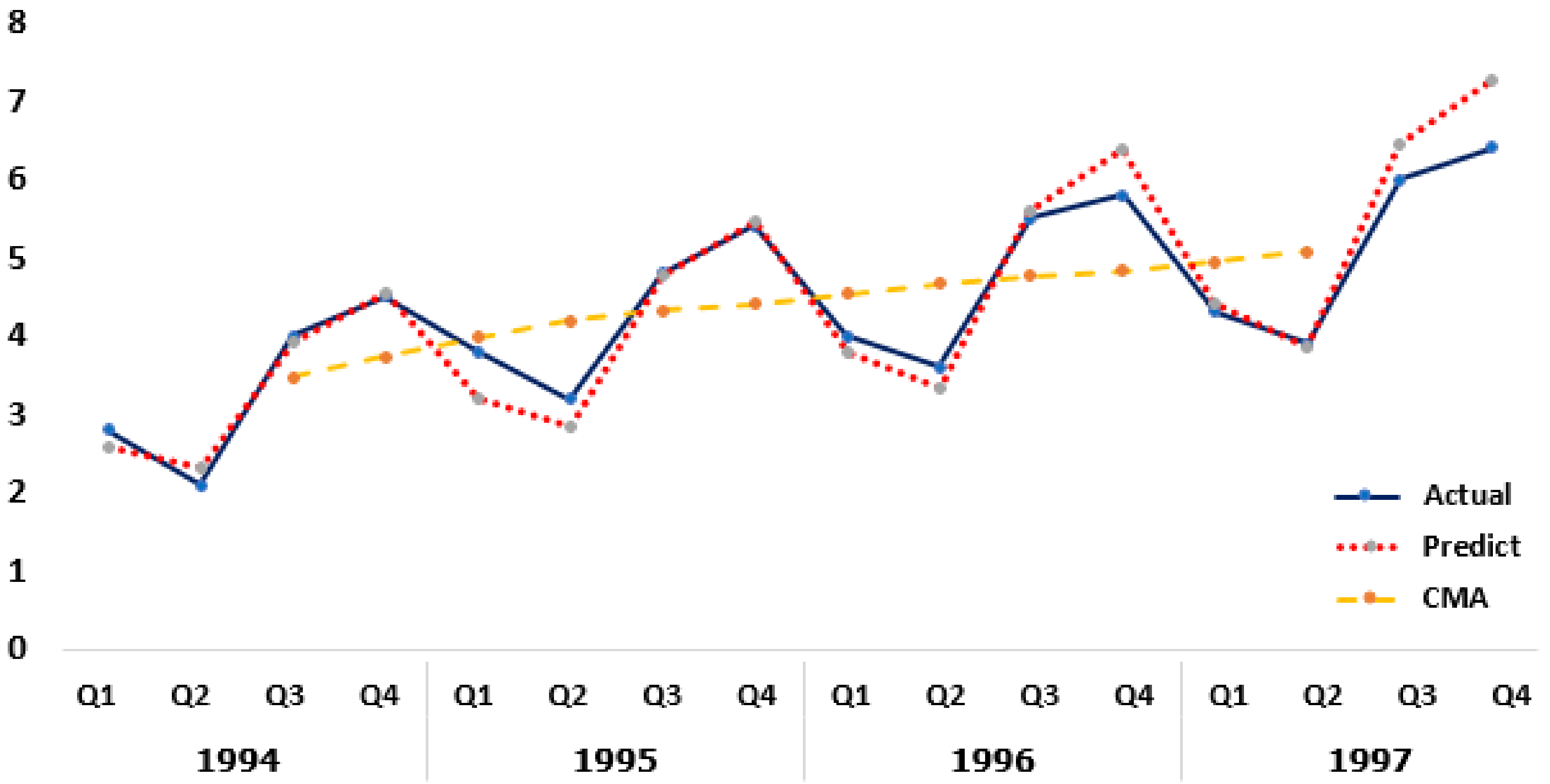
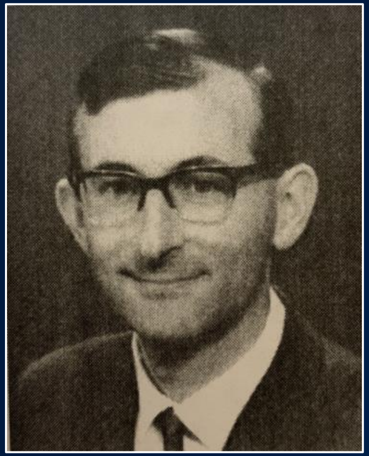


Figure 3.4: Actual value and Predicted value by the least square method

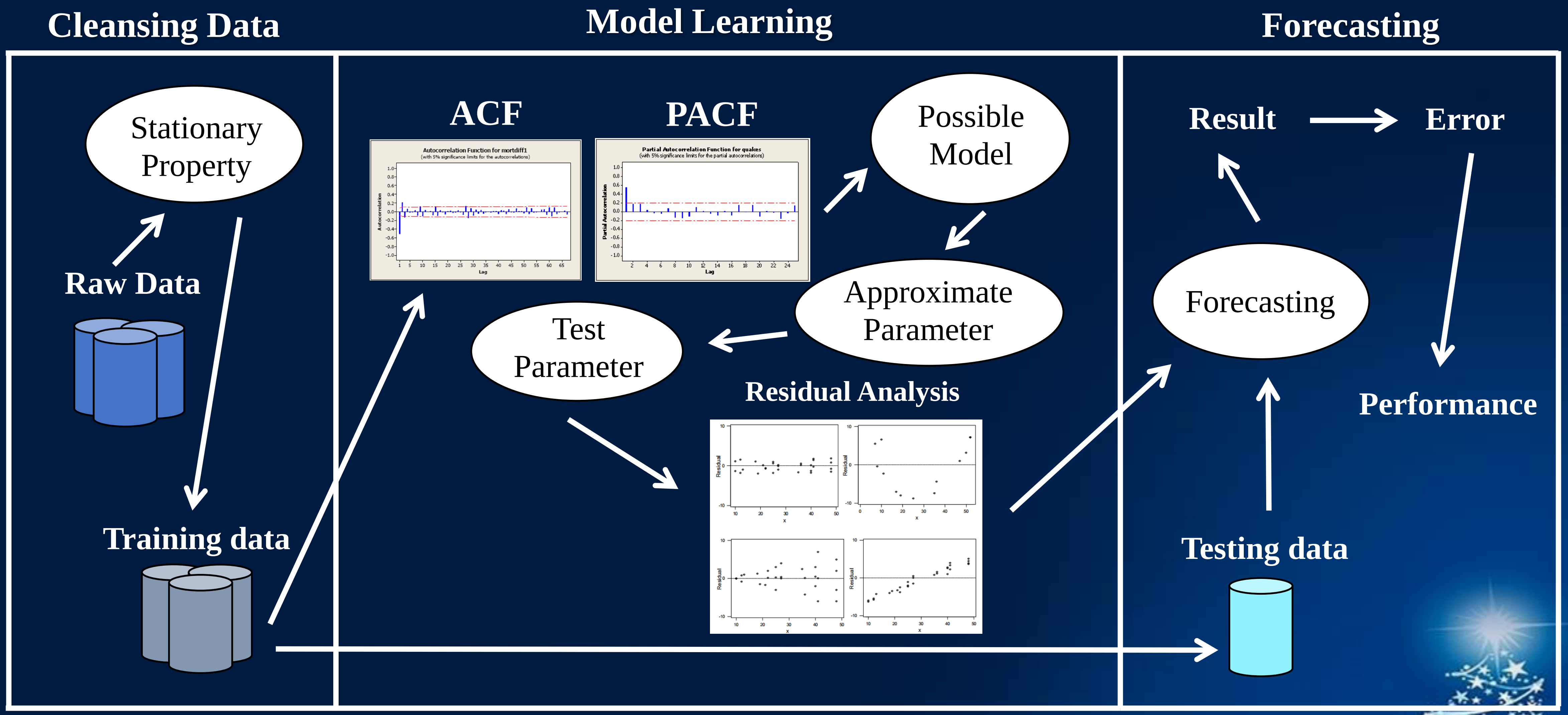
4. Box and Jenkins' Method

George Box and **Gwilym Jenkins'** apply autoregressive moving average or autoregressive integrated moving average models to find the best fit of a time series model to past values of a time series.

George Edward Pelham Box



Gwilym Jenkins


Cleansing Data
Model Learning
Forecasting
Stationary Property
Raw Data
Training data
ACF
PACF
Possible Model
Approximate Parameter
Test Parameter
Residual Analysis
Result
Error
Forecasting
Performance
Testing data

Stationary Properties

Time series y_t is stationary if it has

1. $E(y_t) = \mu$

2. $Var(y_t) = \sigma^2$ and

3. $Cov(y_t, y_{t-s}) = Cov(y_t, y_{t+s}) = \gamma_s$ are constant over time.

Backshift Operator

Backshift Operator (B) is an operator which changes an observation to the one from the previous time step.

$$By_t = y_{t-1} \qquad \text{....(4.1)}$$

In general for any integer k. This will provide a useful notation for our time series models,

$$B^k y_t = y_{t-k} \qquad \text{....(4.2)}$$

Autocorrelation Function (ACF)

Autocorrelation is a correlation coefficient between two different variables, the correlation is between two values of the same variable at times y_i and y_{i+k}.

$$r_k = \frac{\sum_{t=k+1}^{n} (y_t - \bar{y}_t)(y_{t+k} - \bar{y}_t)}{\sum_{t=1}^{n} (y_t - \bar{y}_t)^2} \qquad \text{….(4.3)}$$

where, k = 1, 2, …, n

Partial Autocorrelation Function (PACF)

Partial Autocorrelation is the correlation between observations in a time series and observations a fixed unit of time away while controlling for indirect effects.

$$r_{kk} = \begin{cases} r_1 & ; \; k = 1 \\ \dfrac{r_k - \sum_{j=1}^{k-1} (r_{k-1,j}) r_{k-j}}{1 - \sum_{j=1}^{k-1} (r_{k-1,j}) r_j} & ; \; k = 2, 3, \ldots, n \end{cases} \qquad \ldots.(4.4)$$

where, $r_{kj} = r_{k-1,j} - r_{kk}(r_{k-1.k-j}) \quad ; \; i = 1, 2, \ldots, k-1$ and $k \neq j$

Testing Hypothesis at lag *k* using T-Test

Table 2.1: Hypothesis testing of parameter for ACF and PACF

Hypothesis for Testing	t-test Statistic	Standard Error	Criteria
1. ACF H_0: $\rho_k = 0$ H_1: $\rho_k \neq 0$	$t = \frac{r_k}{S_{r_k}} \sim t_{n-m}$	$S_{r_k} \approx \frac{1}{\sqrt{n}}$	Reject H_0 $\lvert t \rvert > \left\lvert t_{\frac{\alpha}{2},n-1} \right\rvert$ p-value < 0.05
2. PACF H_0: $\rho_{kk} = 0$ H_1: $\rho_{kk} \neq 0$	$t = \frac{r_{kk}}{S_{r_{kk}}} \sim t_{n-m}$	$S_{r_{kk}} \approx \frac{1}{\sqrt{n}}$	Reject H_0 $\lvert t \rvert > \left\lvert t_{\frac{\alpha}{2},n-1} \right\rvert$ p-value < 0.05

Remark: If ACF/PACF at lag k is outsides then the data have ACF/PACF at lag k.

Estimate the Parameter

The least-squares method is used to finds the optimal parameter values by minimizing the sum of squared residuals:

Minimize $$E(\phi,\theta,\Phi,\Theta)=\sum_{t=1}^{n}(y_t-\hat{y}_t)^2 \qquad \text{....(4.5)}$$

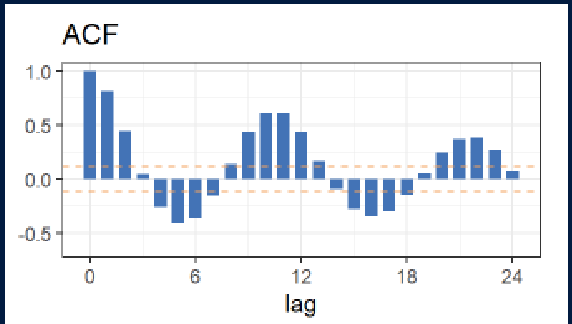
where,

$$\frac{\partial E}{\partial \phi_i}=0 \;\; ; \;\; i=1,2,\ldots,p \quad \text{and} \quad \frac{\partial E}{\partial \theta_j}=0 \;\; ; \;\; i=1,2,\ldots,q$$

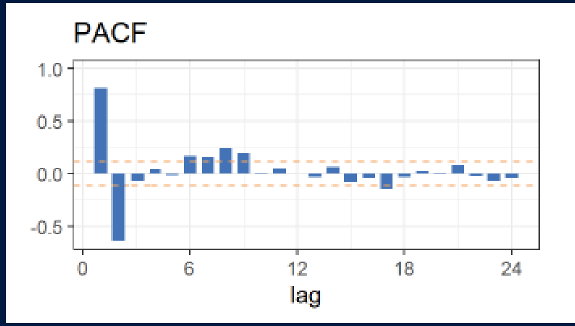
$$\frac{\partial E}{\partial \Phi_k}=0 \;\; ; \;\; k=1,2,\ldots,P \quad \text{and} \quad \frac{\partial E}{\partial \Theta_l}=0 \;\; ; \;\; l=1,2,\ldots,Q$$

Identification of Model

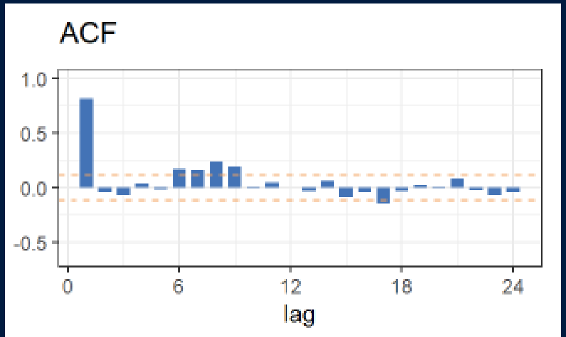
Model	Pattern of ACF	Pattern of PACF
AR(p)	Spikes decays exponentially.	Significance spikes through lags p
MA(q)	Significance spikes through lags q	Spikes decays exponentially
ARMA(p,q)	Spikes decays exponentially	Spikes decays exponentially



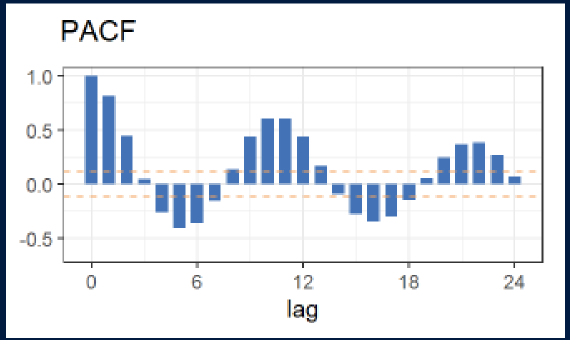
q=0



p=2



q=1



p=0

ARMA(2,0) = AR(2)

ARMA(2,1) = ARIMA(2,0,1)

ARMA(0,1) = MA(1)

Determination of model parameters

In each model, the model parameters p and q are taken from $p = 0,1,. . . ,10$ and $q = 0,1,. . . , 10$. Akaike's Information Criterion (AIC) is estimated in all cases. The parameters p and q for minimize AIC are adopted.

The AIC is given as follows:

$$\mathrm{AIC} = \ln \hat{\sigma}^2 + \frac{2}{n}(p+q) \qquad \text{….(4.6)}$$

Model Diagnostic Checking

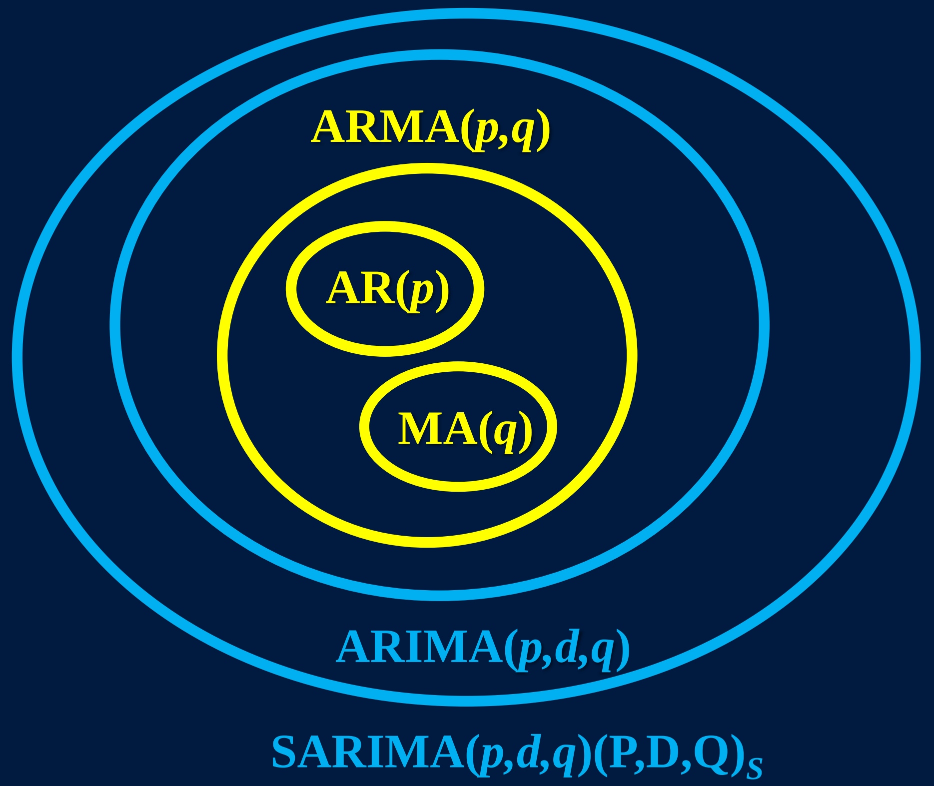
Table 2.2: Model diagnostic checking residual of error.

Hypothesis for testing	Statistic test	Criteria
1. Normal of error H_0: $\varepsilon_t \sim N(0,\sigma^2)$ H_1: $\varepsilon_t \nsim N(0,\sigma^2)$	$Z_{KS} = D_n\sqrt{n}$	Reject H_0 $\lvert Z_{KS}\rvert > \left\lvert Z_{\frac{\alpha}{2}}\right\rvert$ p-value < 0.05
2. Autocorrelation H_0: $\rho_\varepsilon(k) = 0$ H_1: $\rho_\varepsilon(k) \neq 0$	$Q = n(n+2)\sum_{k=1}^{K} \frac{r_e^2(k)}{n-k} \sim \chi^2_{K-m}$	Reject H_0 $Q > \chi^2_{K-m}$ p-value < 0.05
3. Zero mean of error H_0: $\mu_\varepsilon = 0$ H_1: $\mu_\varepsilon \neq 0$	$t = \frac{\bar{e}}{S_{\bar{e}}} \sim t_{n-1}$	Reject H_0 $\lvert t\rvert > \left\lvert t_{\frac{\alpha}{2},n-1}\right\rvert$ p-value < 0.05
4. Constant of variance H_0: $\sigma_k^2 = 0$ H_1: $\sigma_k^2 \neq 0$	$W = \dfrac{(n-k)\sum_{i=1}^{k} n_i(\bar{Z}_i - \bar{Z})^2}{(k-1)\sum_{t=1}^{n}(Z_{ij} - \bar{Z}_i)^2}$	Reject H_0 $W > F_{\alpha,(k-1,n-k)}$ p-value < 0.05

Model diagnostic checking by testing whether the estimated model conforms to the specifications of a stationary univariate process.

1. Normality
2. No Auto-correlation
3. Zero Mean
4. Constant Variance

Family of Box and Jenkins Model



Structures

1. **AR(p):** $$y_t = \phi_0 + \sum_{i=1}^{p} \phi_i y_{t-i} + \varepsilon_t$$

2. **MA(q):** $$y_t = \sum_{j=1}^{q} \theta_j \varepsilon_{t-j} + \varepsilon_t$$

3. **ARMA(p,q):** $$y_t = \phi_0 + \underbrace{\sum_{i=1}^{p} \phi_i y_{t-i}}_{\text{AR}(p)} + \underbrace{\sum_{j=1}^{q} \theta_j \varepsilon_{t-j}}_{\text{MA}(q)} + \varepsilon_t$$

4. **ARIMA(p,d,q):** $$\left(1 - \sum_{i=1}^{p} \phi_i B^i\right) \underbrace{(1-B)^d}_{\text{difference } (d)} y_t = \left(1 - \sum_{j=1}^{q} \theta_j B^j\right) \varepsilon_t$$

5. **SARIMA(p,d,q)(P,D,Q)$_S$:**

$$\phi_p(B)\underbrace{\Phi_P(B^S)}_{\text{SAR}(P)}(1-B)^d \underbrace{(1-B^S)^D}_{\text{Difference } (D)} y_t = \theta_q(B)\underbrace{\Theta_Q(B^S)}_{\text{SMA}(Q)}\varepsilon_t$$

4.1 Autoregressive Moving Average (ARMA)

Autoregressive model is a linear combination of the past values of the observation.

AR(*p*)

$$y_t = \phi_0 + \sum_{i=1}^{p} \phi_j y_{t-i} + \varepsilon_t \qquad \ldots(4.7)$$

Moving average model is a linear combination between the previous values of the errors rather than on the variable itself.

MA(*q*)

$$y_t = \theta_0 + \sum_{j=1}^{q} \theta_j \varepsilon_{t-j} + \varepsilon_t \qquad \ldots(4.8)$$

ARMA model is a hybrid between an autoregressive model AR(*p*) and Moving average model MA(*q*).

ARMA(*p*,*q*)

$$y_t = \theta_0 + \underbrace{\sum_{i=1}^{p} \phi_i y_{t-i}}_{\textbf{AR}(p)} + \underbrace{\sum_{j=1}^{q} \theta_j \varepsilon_{t-j}}_{\textbf{MA}(q)} + \varepsilon_t \qquad \ldots(4.9)$$

4.2 Autoregressive Integrated Moving Average (ARIMA)

In ARIMA models a non-stationary time series is made stationary by applying finite differencing of the observations.

ARIMA(*p*,*d*,*q*)

$$\phi_p(B)\underbrace{(1-B)^d}_{\text{Non-seasonal difference } (d)} y_t = \theta_q(B)\varepsilon_t \qquad \text{....(4.10)}$$

$$\underbrace{\left(1-\sum_{i=1}^{p}\phi_i B^i\right)}_{\text{AR}(p)} y_t' = \underbrace{\left(1-\sum_{i=j}^{q}\theta_j B^j\right)}_{\text{MA}(q)} \varepsilon_t \qquad \text{....(4.11)}$$

where, $y_t' = (1-B)^d y_t$

4.3 Seasonal Autoregressive Integrated Moving Average (SARIMA)

The ARIMA model is for non-seasonal non-stationary data. Box and Jenkins have generalized this model to deal with seasonality.

SARIMA(*p*,*d*,*q*)(*P*,*D*,*Q*)$_S$

$$\phi_p(B)\Phi_P(B^S)(1-B)^d \underbrace{(1-B^S)}_{\textbf{Seasonal difference } (D)}{}^D y_t = \theta_q(B)\Theta_Q(B^S)\varepsilon_t \qquad \ldots.(4.12)$$

$$\underbrace{\left(1-\sum_{i=1}^{p}\phi_i B^i\right)}_{\textbf{AR}(p)}\underbrace{\left(1-\sum_{k=1}^{P}\Phi_k B^k\right)}_{\textbf{SAR}(P)}y_t' = \underbrace{\left(1-\sum_{j=1}^{q}\theta_j B^j\right)}_{\textbf{MA}(q)}\underbrace{\left(1-\sum_{l=1}^{Q}\Theta_l B^l\right)}_{\textbf{SMA}(Q)}\varepsilon_t \qquad \ldots.(4.13)$$

where, $y_t' = (1-B)^d(1-B^S)^D y_t$

4.4 Seasonal Autoregressive Integrated Moving Average with Exogenous Variables (SARIMAX)

Seasonal Autoregressive Integrated Moving Average with **Exogenous Variables** (SARIMAX) model is a SARIMA model with **Exogenous Variables** $(X_{i,t})$.

$$y_t = \beta_0 + \beta_1 X_{1,t} + \beta_2 X_{2,t} + \ldots + + \beta_k X_{k,t} + \omega_t \qquad \ldots.(4.14)$$

where, $$\omega_t = \frac{\theta_q(B)\Theta_Q(B^S)}{\phi_p(B)\Phi_P(B^S)(1-B)^d(1-B^S)^D}\varepsilon_t \qquad \ldots.(4.15)$$

SARIMAX(*p,d,q*)(*P,D,Q*)$_S$

$$y_t = \beta_0 + \sum_{i=1}^{k}\beta_i X_{i,t} + \frac{\theta_q(B)\Theta_Q(B^S)}{\phi_p(B)\Phi_P(B^S)(1-B)^d(1-B^S)^D}\varepsilon_t \qquad \ldots.(4.16)$$

5. Evaluation Model

Due to the fundamental importance of time series forecasting in many practical situations, proper care should be taken while selecting a particular model. For this reason, various performance measures to estimate forecast accuracy and to compare different models.

5.1 Mean Absolute Percentage Error (MAPE)

Mean absolute percentage error is a measure of prediction accuracy of a forecasting method in statistics. It usually expresses the accuracy as a ratio. MAPE is defined by the formula:

$$MAPE\% = \frac{1}{n}\sum_{t=1}^{n}\left|\frac{y_t - \hat{y}_t}{y_t}\right| \times 100 \qquad \ldots.(5.1)$$

5.2 Mean Square Error (MSE)

Mean Squared Error is a measure of the average of the squares of the difference between the estimated values and the actual value.

$$MSE = \frac{1}{n}\sum_{t=1}^{n}(y_t - \hat{y}_t)^2 \qquad \text{….(5.2)}$$

5.3 Root Mean Square Error (RMSE)

Root Mean Squared Error is a square root of mean squared error which, measures the quality of the fit between the actual data and the predicted model.

$$RMSE = \sqrt{\frac{1}{n}\sum_{t=1}^{n}(y_t - \hat{y}_t)^2} \qquad \text{….(5.3)}$$

5.4 Reduce Error Rate (RER)

Reduce Error Rate is a measure of reducing errors for the proposed model when compares with the original model as a ratio.

$$RER\% = \left(1 - \frac{MAPE_{\Pr oposed}}{MAPE_{Original}}\right) \times 100 \qquad \ldots.(5.4)$$

5.5 Coefficient of Determination *(R^2)*

The coefficient of Determination is a statistic that will give some information about the goodness of fit of a model.

$$R^2\% = \left(1 - \frac{\sum_{t=1}^{n}\left(y_t - \hat{y}_t\right)^2}{\sum_{t=1}^{n}\left(y_t - \bar{y}_t\right)^2}\right) \times 100 \qquad \ldots.(5.5)$$