Chapter 3

Research Methodology

3.1 Research Design

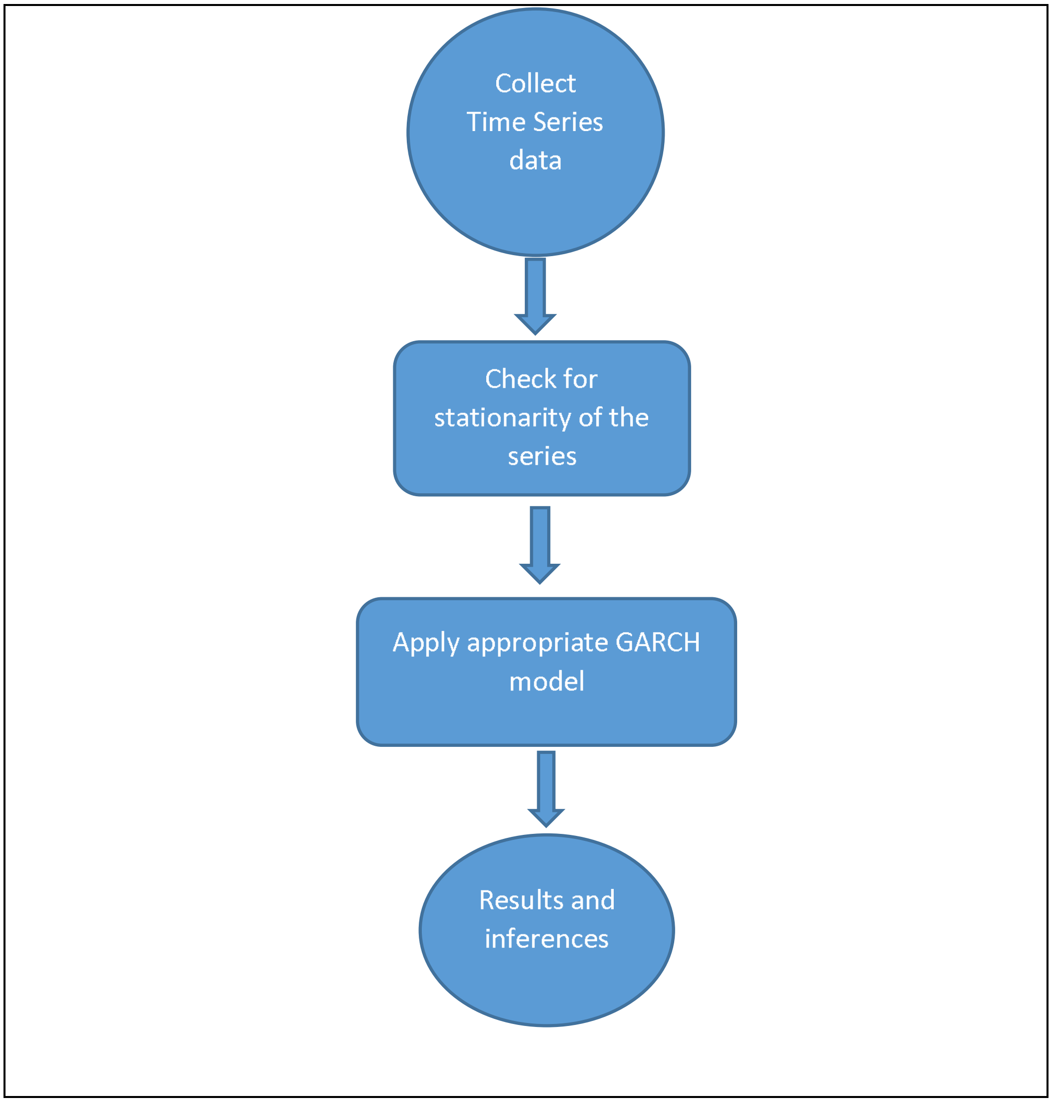

Figure 3.1 Outline of Research Design

The research design consists of time series data collection based on objective of study, analysing unit root tests for stationarity, using appropriate model for volatility check and interpretation of results. The various steps are detailed below.

3.2 Statement of the Problem

India has been exposed to various risks while catering to energy needs from imports of fossil fuels like Crude oil, Coal and LNG. India's energy policy is making us dependent on imported LNG and thus there is need to understand the volatility of prices which could affect the cost of power in India. Since most of the LNG are traded based on long-term contact contracts, there is a need to understand the volatility of Indexed LNG prices.

The Study consists of volatility check Henry Hub Index, Japan Crude Cocktail Index. There is also need to study the volatility of Time Charter rates of LNG ships and new shipbuilding prices of LNG carriers.

3.2.1 Research Questions

Understanding the literature review, research gaps and statement of the problem, the research questions could be as mentioned below

1. What is the volatility in Henry Hub and Japanese crude cocktail price indexes of Natural Gas?
2. What is the volatility in Charter prices of LNG ships?
3. What is the volatility in new Shipbuilding prices of LNG?

3.3. Objectives of the Study

1. To Study the volatility of Henry Hub and Japan Crude Cocktail prices of Natural Gas using GARCH models.

2. To analyse the Volatility of Ship Charter rates for Long Term and Short Charter rates term charter rates for LNG shipping using GARCH models.

3. To Study the Volatility of New ship building prices of LNG using GARCH models.

3.4 Scope of the study

3.4.1 Objective 1

There are many regional Gas price indexes but keeping in view India's major sourcing from Qatar and USA, have considered only Japanese Crude Cocktail and Henry Hub prices for testing volatility.

3.4.2 Objective 2

There are many classification of sizes of LNG ships based on the capacity. However, based on figure 4.5 describing present fleet size we have considered Long term ship charter rates and Short Term ship charter rates only for vessel size of 1,55,000 to 1,73,000 CBM capacity for our study.

3.4.3 Objective 3

For the study, a specific class of LNG vessels with a volume carrying capacity of between 1,55,000 to 173000 cbm have been considered. This class has been chosen because after expansion and reopening of the Panama Canal, LNG carriers with maximum capacity up to 180,000 cbm can pass through it. The identified class is applied as a filter based on the global LNG new vessel order book as shown in figure 4.7.

3.5 Data Collection

3.5.1 Objective 1

a. Ten years monthly Henry Hub natural gas prices in US dollars per million British thermal unit is taken from the website of U.S. Energy http://www.eia.gov/dnav/ng/hist/rngwhhdm.htm for the period July 2005 till June 2015.
b. Ten-year monthly Japanese Custom-cleared crude also called as Japanese Crude cocktail (JCC) prices in US dollars per barrel is taken

   http://www.paj.gr.jp/english/statis/ for the period July 2005 until June 2015.

3.5.2 Objective 2

a. The monthly long term time charter rates for LNG ships of 1,55,000 cubic meter size from February 2005 until August 2015 were obtained from Drewry Shipping Consultants Limited, Gurgaon, India in US dollars per day.

b. The monthly short-term time charter rates for LNG ships of 1,55,000 cubic meter size from February 2005 until August 2015 were obtained from Drewry Shipping Consultants Limited, Gurgaon, India in US dollars per day.

3.5.3 Objective 3

a. For modelling, new LNG ship building price volatility, 126 observations in form of monthly prices from April 2005 to August 2015 are examined in this time series. The new ship building prices were collected from Drewry Maritime Consultants, India

3.6 Tools for Analysis

To explore if the time series data is stationary the following unit root tests are applied for all the time series data used in objective 1, 2 and 3

3.6.1 Unit roots tests

3.6.1.1 Basic Unit Root theory

Let us consider an Auto regression (1) process

$$Y_t = \rho\, y_{t-1} + x_t\, \delta + \varepsilon_t \qquad \text{equation (3.1)}$$

Where x_t could be a constant trend or a constant. The parameters ρ and δ are to be estimated. ε_t is the white noise which is assumed. $|Y|$

If $|\rho| \geq 1$, then Y is a nonstationary series and the variance of Y increase with time and approaches infinity. If $|\rho| < 1$, then Y is stationary series. Therefore, the hypothesis that the series is stationary is evaluated by testing if the absolute value of ρ is less than one. The various Unit root tests are

3.6.1.2 Augmented Dicky Fuller test:

The Augmented Dicky fuller has discussed in research paper (David A. Dickey and Wayna A. Fuller, 1981) about unit root test for testing stationarity of a time series. The standard DF test is done by estimating the equation (3.1) by subtracting y_{t-1} from both the sides of the equation, then

$$\Delta y_t = \alpha y_{t-1} + x_t \delta + \varepsilon_t \qquad \text{equation (3.2)}$$

Where $\alpha = \rho - 1$. The null hypothesis and alternative hypothesis may be written as

$$H_0: \alpha = 0$$

$$H_1: \alpha < 0 \qquad \text{equation (3.3)}$$

And is evaluated by using the conventional *t*- ratio for α:

$$t_\alpha = \hat{\alpha} / (se(\hat{\alpha})) \qquad \text{equation (3.4)}$$

in the above equation (3.4) $\hat{\alpha}$ is the estimate of α and $se(\hat{\alpha})$ is the coefficient of standard error.

Dicky Fuller brought forward that under the null hypothesis of a unit root, the conventional student's *t*- distribution is not followed and these derive asymptotic results and critical values for sample sizes and various test are simulated. This is to note that the Dickey Fuller unit root test is valid when the series is Auto Regression (1) process. But if the series is correlated at higher lags then the assumption of white noise ε_t is violated. Then the Augmented Dickey Fuller test can be used to where it constructs a parametric correction for higher order correlation thereby assuming that *Y* series is an Auto Regression (*p*) process. By adding *p* Lagged difference to the independent variable *y* to the equation on the right hand of the regression, then

$$\Delta y_t = \alpha\, y_{t-1} + x_t \delta + \beta_1 \Delta y_{t-1} + \beta_2 \Delta y_{t-2} + \beta_3 \Delta y_{t-3} \ldots + \beta_p \Delta y_{t-p} + \upsilon_t \qquad \text{equation (3.5)}$$

Thereby, this new augmented condition is used to test equation (3.3) by using *t*-ratio in equation (3.4). Fuller had obtained an important result that the asymptotic distribution of the *t*- ratio for α is independent by the number of lagged first differences which are included in the Augmented Dickey Fuller regression. From the research paper (Said & Dickey, 1984) the authors demonstrate that the Augmented Dickey Fuller test is asymptotically valid in the presence of a moving average component, if in the regression sufficient lagged differences terms are included.

3.6.1.3 Kwiatkowski, Phillips, Schmidt, and Shin (KPSS) test

(Kwiatkowski, Phillips, & Peter Schmidt and Yongcheol Shin, 1992) brought forward new test where y_t is assumed to be stationary under the null. The test id based on the residuals from the original least square regression of y_t on the exogenous variables x_t:

$Y_t = x_t'\delta + u_t$ Equation (3.6)

The definition of Linear Model statistic is given by

$LM = \sum_t S(t)^2/(T^2 f_0)$ Equation (3.7)

Where f_0 is an estimator of the residual spectrum at zero frequency and where

$S(t)$ is residual function which is cumulative:

$S(t) = \sum_{r=1}^{t} \hat{\mu}_r$ Equation (3.8)

Now, based on the residuals $\hat{\mu}_r = y_t - x_t\,\delta\,(0)$. This is to note that the estimator δ used in the above calculation is different from the δ used by the GLS since this is based on a regression relating the original data and not the quasi-differenced data.

3.6.2 Modelling Objective 1

3.6.2.1 Simple Univariate GARCH Model

In this objective, for measuring the volatility, I have used simple univariate GARCH ((Bollerslev T ., 1986) and (Bollerslev T. R., 1992)) model.

The simplest GARCH (1, 1) specification used is given below.

$\ln P_t = \alpha_0 + \alpha_1 \ln P_{t-1} + e_t$ Equation (3.9)

With the error distributed as $N(0, h_t)$ where

$$h_t = \alpha + \beta e_{t-1}^2 + \gamma h_{t-1} + u_t$$ Equation (3.10)

The conditional variance equation (h_t) (Equation 2) described above is a function of three terms:

(i) The constant term, α;

(ii) News about volatility from the previous period, measured as the lag of the squared residual from the mean equation, e_{t-1}^2 (the ARCH term); and

(iii) The last period's forecast error variance, h_{t-1} (the GARCH term).

3.6.2.2 Exponential GARCH model (EGARCH)

The GARCH model successfully captures the volatility clustering, but it is an inefficient model when one wishes to analyse the *leverage effect.* Because, the conditional variance is a function only of the magnitudes of the past values but does not capture whether the effect is positive or negative. This asymmetric relationship is called *leverage effect*. It describes how a negative shock causes volatility to raise more than if a positive shock with the same magnitude had occurred. To capture this asymmetry, different models have been used. EGARCH model captures asymmetric responses of the time-varying variance to shocks and ensures that the variance is always positive. The model was developed by (Nelson, 1991).

$$\log(\sigma_t^2) = \omega + \sum_{j=1}^{q} \beta_j \log(\sigma_{t-j}^2) + \sum_{i=1}^{p} \alpha_i \left(\left| \frac{\epsilon_{t-i}}{\sigma_{t-i}} \right| - E \left| \frac{\epsilon_{t-i}}{\sigma_{t-i}} \right| \right) + \sum_{k=1}^{r} \gamma_k \frac{\epsilon_{t-k}}{\sigma_{t-k}}$$

and $\epsilon_t = \sigma_t \eta_t$ Equation (3.11)

The left-hand side is the *log* of the conditional variance. This implies that the leverage effect is exponential, rather than quadratic and that forecasts of the conditional variance are guaranteed to be nonnegative. The presence of leverage effects can be tested by the $\gamma_i < 0$. The impact is asymmetric if $\gamma_i \neq 0$.

Equation (3.11) explains a conditional variance which defines the generating process for the natural logarithm of conditional variance. EGARCH emphasizes that the asymmetric function of past innovation shocks.

γ refers to the magnitude of persistence of variance. When the closer magnitude approaches unity, the greater persistence of shocks to the volatility. The positive and negative anticipates excess the returns determine the future variance, measured by β and α. α If $\alpha > 0$, innovation in $\log(\sigma)$ is then positive when magnitude of (σ_{-j}) is larger than its expected value β β is 0, then there is a non-existence of $\beta > 0$ and statistically significant, which means the volatility of return's shock is asymmetric. The positive volatility innovations are greater $\beta<0$, and statistically significant, the volatility of return's shock is asymmetric and negative volatility innovations is larger than same magnitude of positive shocks. β represents the persistence of leverage effect.

3.6.3 Modelling for Objective 2

3.6.3.1 EGARCH Model

The mean equations of GARCH models are as follows

$$\Delta \text{logSTt} = \alpha + \beta_t \Delta \log LT_{t-1} + e_t \qquad \text{equation (3.12)}$$

$$\Delta \text{logLTt} = \alpha + \beta_t \Delta \log ST_{t-1} + e_t \qquad \text{equation (3.13)}$$

The GARCH model successfully brings forth the volatility clustering, but it is an inadequate model when we wish to analyse the leverage effect. Because, the conditional variance is a function only of the magnitudes of the past values but does not capture whether the effect is positive or negative. This asymmetric relationship is called leverage effect. This describes how a negative shock causes volatility to raise more than if a positive shock with the same magnitude had occurred. Different models have been used to capture this asymmetry. EGARCH model captures asymmetric responses of the time-varying variance to shocks and ensures that the variance is always positive. The model was developed by (Nelson, 1991) as discussed in equation (3.11).

3.6.4 Modelling for Objective 3

3.6.4.1 GARCH (1,1)

For the analysis GARCH (1, 1) model is used which has been given in Equation (3.9) and Equation (3.10) in section 3.6.2.1.

3.6.4.2 EGARCH model captures asymmetric responses of the time-varying variance to shocks and ensures that the variance is always positive. The model as discussed in Equation (3.11) has been used for analysis

3.7 Software

EVIEWS version 9 has been used for analysis