Chapter - 3

LITERATURE REVIEW

To address the challenges and issues encountered in performance evaluation of coal fired power plants, efforts have been made in this chapter to have a comprehensive review of literature covering performance evaluation techniques, productivity change, performance evaluation studies in different sectors in Indian context and performance evaluation of power utilities in generation, transmission and distribution sector of different countries. The purpose is to identify the gaps in the works done so far, so that current study can be proved to be useful for the practitioners and policy makers.

3.1 Performance Evaluation Techniques

Objective of any production process or a firm is to create value by transforming inputs to outputs. Outputs in the form of goods and services, in general are desirable implying more is better. Inputs on the other hand are valuable resources, having potential for alternative usage facilitating unspent inputs in a production process to be utilized to produce more outputs of the same kind or other indicating lesser consumption is better. The twin objectives of efficient resource utilization by a firm are (a) to produce as much output as possible from a specific quantity of input and at the same time, (b) to produce a specific quantity of output using as little input as possible (Ray, 2004). While the performance of a firm is its resource utilization ability and productivity is a descriptive measure of performance. Productivity refers to the efficiency with which resources are transformed to goods and services.

Several techniques have been used to quantitatively estimate and benchmark the productivity levels of various processes. The classic measure of productivity as the ratio of output to input, which does very well for the single input and output processes, fares badly with the increasing complexity of the modern day business processes which in reality makes use of multiple inputs to deliver multiple outputs. Various benchmarking techniques used for performance evaluation are detailed in Table 10. (Ajobhia et al., 2003; Hirschhansen and Cullman , 2005). Out of these econometrics techniques like OLS and its variants, indexing and data envelopment analysis are most commonly adopted worldwide for measuring utility performance (Shumilkina, 2010).
The techniques can broadly be categorized in to partial approach - in which one or two parameters are considered at a time and multi dimensional approach in which multiple parameters which can be broadly categorized into inputs and or outputs are taken into account simultaneously.

Partial approach uses a set of partial performance indicators (PPI) which capture the relation between two variables at a time. This approach is easy to compute and quite meaningful in single input single output contexts but looses much of its relevance for multiple input multiple output processes because of their inability to provide information about other parameters. This makes the approach also called ratio analysis, directional in nature. Ratio analysis is often criticized on the ground of subjectivity, because an analyst has the tendency to pick and choose favorable ratios in order to assess the overall performance of a firm (Malhotra and Malhotra, 2007) and have the tendency to attribute the gains to one factor, which actually was contributed by another (Cooper et al., 2007).

The limitations of individual ratios led to multi dimensional approach, in which multiple ratios are captured to provide an overall indication of performance. Common examples of this approach are Human Development Index, Comprehensive Award Scheme 2008, Altman's Z Score etc. Multi-dimensional approaches attempt to aggregate the PPIs into a composite indicator (CI). Even though the CIs has remained extremely useful for the policy planners and decision makers, devising a universally acceptable aggregation technique posses a tough challenge till date. Subjectivity associated with selection of weights lies at the centre of controversy. Nardo et al. (2005) discusses several approaches to assigning weights like Principal Component Analysis (PCA), Factor Analysis (FA), Data Envelopment Analysis (DEA), Benefits of the Doubt Approach (BoD), Unobserved Components Model (UCM), Budget
Allocation (BAL), Public Opinion, Analytic Hierarchy Process (AHP), Conjoint Analysis (CA) and aggregation techniques like Additive Aggregation Methods, Non-compensatory Multi Criteria Approach (MCA) and Geometric Aggregation.

Taking further the coefficient of resource utilization of Debreau (1951) and gauge function of Fenchel (1953), Shephard (1953) introduced the concept of distance function which was subsequently used by Farell (1957) for measuring efficiency directly from observational data. In situations where multiple inputs are used to produce multiple outputs, distance functions are used to represent the production technology and calculate the technical efficiencies or shadow prices (Kumbhakar et al., 2003). Technical efficiency (TE) of a firm reflects its ability to minimize usage of inputs to produce a given amount of output or maximize production of outputs from a given bundle on inputs. The firm which uses the least input or maximum output is called technically efficient and has a TE score of 100%.

The strength of distance function approach lies in its ability to specify multiple-input, multiple-output technology in absence of price information or in cases where cost, profit or revenue function cannot be used due to behavioral constraints (Coelli and Perelman, 1999). Distance functions can be categorized into two categories namely output distance function and input distance function. While the output distance function estimates the extent of output augmentation from the existing input set, the input distance function explores the possibility of contracting the input vector proportionally with the output vector held fixed.

Farell (1957) provided the modern efficiency measurement concepts which can accommodate multiple inputs, multiple outputs and non-constant return to scale. According to him the efficiency of a firm consists of two components: Technical Efficiency (TE) - the ability of a firm to obtain maximal output from a given set of inputs and Allocative Efficiency (AE) - the ability of a firm to use the inputs in optimal proportions, given their respective prices. These two efficiency measures when combined together provide the measure of total Economic Efficiency (EE) of a firm. Firms using optimal input and output combination are said to be fully efficient, lie on the efficiency frontier and have a TE of 1 i.e. 100%.

TE of other firms is their productivity index relative to the hypothetical firm producing maximum output possible from the same input quantity and can be
measured by measuring the distance of the production possibility set for the firm from the unit isoquant of the fully efficient firm (Ray, 2004). Some of the firms in the frontier have non-zero slacks; these firms are called weakly efficient ones and are Farrell efficient. Firms having zero slacks are Pareto-Koopmans efficient and called strongly efficient ones (Cooper et al., 2007).

Theoretically identical sets of inputs should produce identical output quantities by all firms, which are rarely observed. The set of firms producing maximum possible output occupy what is called the frontier. Estimation of the productivity gaps and identification of the factors responsible for the gap has remained a tough challenge for the researchers. The production function of a truly efficient firm is not known in practice and thus must be estimated from observations of firms in the industry concerned (Coelli, 1996).

Frontiers have been estimated using many methods, the two principal methods used are Stochastic Frontier Analysis (SFA) and Data Envelopment Analysis (DEA) (Coelli, 1996). Both the methods employ quite different methodologies for frontier estimation and efficiency measurement, each with associated strengths and weakness. The strength of SFA lies in its ability to handle statistical noise and suffers from the drawback of requiring strong assumptions as to the form of the frontier. Basic DEA does not require any assumptions for the functional form of the production function and have the disadvantage of assuming no statistical noise. The advantages of DEA emanates from its ability to identify the quantum of input and output slacks along with peer groups for benchmarking. Recent developments in the form of 2-stage and 3-Stage DEA (Yang and Politt, 2007), Stochastic DEA (Brazdik, 2005) have been reported in the literature and augmented the statistical noise handling capability of DEA.

DEA is a linear programming based, non parametric, frontier approach used to measure the relative performance level of homogenous firms consuming a bundle of identical inputs to produce similar outputs. These homogenous units are called Decision Making Units (DMU) in DEA literature and can include banks, hospitals, municipalities etc. DEA is used to evaluate the performance of each observation relative to the frontier that envelopes all of the observations. DEA was proposed by Charnes, Cooper and Rhodes (1978) and the basic DEA model assuming constant
return to scale (CRS) is also called CCR model. DEA provides information about the quantum of input reductions and output enhancements that can be achieved based on the achievements of other firms. The advantage of DEA is that it is a non-parametric approach which does not require the assumption of a functional form of the production function as required by parametric methods like Stochastic Frontier Analysis (SFA). The efficiency scores arrived are not based on theoretical considerations but based on the achievements of the efficient DMUs and the amount of possible improvement can be quantified.

The primal DEA model also called multiplier model explores to find the set of optimal weights for the inputs and outputs that maximizes the ratio of weighted outputs to weighted inputs subject to the condition that the ratio of weighted outputs to weighted inputs for other DMUs lie between 0 and 1.

The dual DEA model attempts to carve out hypothetical DMUs from the linear combination of existing DMUs, which consume not more inputs to produce at least the same level of outputs as that of the real DMU. If it is not possible to have one or more hypothetical DMUs, then it is neither possible to contract the input bundle nor augment the output levels without deteriorating other parameters. Such DMUs are called efficient and the loci of the operational parameters of such efficient units determine the efficient frontier. Otherwise the hypothetical DMU becomes a role model and the constituents of the hypothetical DMU become the peers for the real DMU, whose best practices can be emulated by benchmarking to achieve improved targets. Various models of DEA incorporating returns to scale, orientations, weight restrictions, types of variables have been formulated over the years to take care of the real world problems (Cooper et al. (2007). While the Input Oriented DEA model aim at exploring possible input contractions without deteriorating current output levels, the Output Oriented model aim at finding possible output augmentations without exceeding the currently consumed input levels.

Assuming that N DMUs produce M outputs from K inputs, for the $i^{th}$ DMU which transforms input vector $x_i$ to output vector $y_i$ and assuming variable return to scale, using input oriented DEA, the technical efficiency $\theta$ can be evaluated by solving the following linear programming problem:

\[
\text{Maximize} \sum_{j=1}^{M} \theta_j y_j \text{ subject to} \sum_{k=1}^{K} \theta_k x_{ij} \leq \sum_{j=1}^{M} \theta_j y_j \text{ for all } i = 1, 2, \ldots, N \]

\[
\theta_k \geq 0 \text{ for all } k = 1, 2, \ldots, K
\]
\[
\begin{align*}
\min_{\theta, \lambda} & \quad \theta, \\
\text{st} & \quad -y_i + Y\lambda \geq 0, \\
& \quad \theta x_i - X\lambda \geq 0, \\
& \quad N\lambda = 1 \\
& \quad \lambda \geq 0,
\end{align*}
\]

Where \( X \) is the \( KN \) input matrix, \( Y \) is the \( MXN \) output matrix, \( \lambda \) is a \( NX1 \) vector of constants. The technical efficiency \( \theta \) obtained is called the Farrell efficiency score for \( i^{th} \) DMU and \( \lambda \) is the weight vector for the DMU. Pareto Koopmans efficiency measure can be estimated by maximizing the input and output slacks (Coelli, 1996).

Ideally plants should operate at an optimal scale and at most efficient scale. However imperfect competition, constraints on technology, finance, etc. may restrict these plants operate at their most productive scale size (Coelli, 1996). Investigation of Return to Scale (RTS) properties of a plant is important to guide the managers in scaling up or scaling down their operations to make it more efficient. The initial CCR model without taking scales of operation into account was subsequently modified by incorporating Variable Return to Scale (VRS) in the BCC model by Banker, Charnes and Cooper (Banker et al., 1984) and allows investigation of RTS properties with the help of DEA. BCC model evaluates the TE under CRS and VRS assumptions with the difference attributed scale inefficiency.

DEA has been extensively used as a tool for performance evaluation, slack estimation, target setting and identification of peer units for benchmarking of many different kinds of entities engaged in many different activities in many different contexts in many different countries. Sample applications of DEA include Airports, Banking, Coal Sector, Educational Schools, Hospitals, Institutes and Universities, Gas Industry, Ports, Power Sector, Railways, Road Transportation, Telecommunication, Water Supply, maintenance activities of US Air Force bases in different geographic locations, police forces in England and Wales, (Thakur, 2007; Cooper et. al., 2007). Ramanathan (2005), Cooper et al. (2007) and Coelli (1996) are sources of excellent reference on DEA. Emrouznejad et al. (2008) surveyed the application of DEA to theoretical developments as well as real-world applications from inception to 2007. Emrouznejad and Witte (2010) proposed a unified model for non-parametric
performance modeling titled COOPER framework to make the assessments more-reliable, more repeatable and less costly.

Several studies have been reported in the literature comparing various frontier techniques. Banker et al. (1991) compared two leading frontier estimation methods: Corrected Ordinary Least Square (COLS) and DEA. The study revealed while COLS perform better for the classical distribution case with sample sizes of 50 or over; DEA performs better for all non-classical inefficiency distributions. Jamasb and Politt (2001) have surveyed the use of benchmarking methods in the OECD and few other countries and found that electricity regulators worldwide use a variety of methods for benchmarking with a notable preference for the non-parametric methods. Hirschhausen et al. (2006) applied DEA and SFA to electric distribution utilities in Germany and found high degree of correlation between both the methods.

In view of the power and appeal of the methodology, DEA based productivity management framework, comprising of performance measurement, slack estimation, target setting and benchmarking, is an ideal candidate for achieving productivity gains of Indian power plants.

3.2 Productivity Change

Productivity change refers to change in output without change in input. While increase in output is termed positive growth and indicates overall progress, decrease in output indicates regress. Solow (1956) estimated that while one-eighth of the total increase in GNP of US during 1909 to 1949 was contributed by increased capital per man-hour, remaining seven-eighths was attributable to other factors what Solow termed Technical Change. In fact Solow (1956) used the phrase Technical Change, as a short hand expression for any kind of shift in the production function. Thus slowdowns, speed-ups, improvements in the education of the labor force, and all sorts of things will appear as Technical Change. Popular methods of measuring productivity change are index number based approaches, parametric approach and non-parametric approach (Mongia and Sathaye, 1998). The drawbacks of index number based approach are requirement of prices information and assumptions concerning the behavior of producers and structure of technology. The difficulty with the parametric approach is the knowledge of the production function. In situations
where either the cost information is not available or misrepresented or the structure of the production function is not known, Malmquist Productivity Index (MPI) is used to measure TFP change. This was developed by Caves, Christensen and Divert (1982) from the Malmquist's (1953) quantitative index for consumption analysis and Farrell's (1957) distance function approach. MPI is a bilateral total factor productivity index, uses frontier approach to measure productivity change (TFPCH) between two periods and can be further decomposed to Technical Change and Efficiency Change.

Technical Change (TECHCH) measures the degree of the frontier shift i.e. changes in the efficient frontier between the two periods and captures innovation, invention and diffusion of technology, slowdowns, etc. Efficiency Change (EFFCH) measures the technical efficiency change of the firm between two periods and captures the catch-up efforts to the frontier of the firm to become industry leader. Catch-up is the ratio of (efficiency of the production point \((x_{i+1}, y_{i+1})\) with reference to \(t+1\) frontier) and (efficiency of the production point \((x_i, y_i)\) with reference to \(t+1\) frontier). Coeli (1996) and Cooper et al. (2007) provide detailed explanation to MPI. EFFCH can be further decomposed into Scale Efficiency Change (SECH) and Pure Efficiency Change (PECH). Mean of the indices is the geometric mean. While indices of more than one indicates progress, less than and equal to one indicates regress and no change respectively.

Output based MPI between two different time periods involving two production points \((x_i, y_i)\) and \((x_{i+1}, y_{i+1})\) is represented as:

\[
MPI = \sqrt{\frac{D_t(x_{i+1}, y_{i+1})}{D_t(x_i, y_i)} \cdot \frac{D_{t+1}(x_i, y_i)}{D_{t+1}(x_{i+1}, y_{i+1})}}
\]

Where \(D_t(x_i, y_i)\) is the distance function and represents the efficiency of conversion of inputs \(x_i\) to outputs \(y_i\) during the period \(t\).

Productivity change of a power plant refers to a situation when the power generation changes without changes in capacity, coal and other input parameters.

3.3 Performance Evaluation Studies in Indian Context

In the recent years researchers have applied frontier techniques like SFA and DEA for performance evaluation, slack estimation, target setting and analysing total factor
productivity change of different segments like Food Processing Industries (Ali et al., 2007), IT Industry (Mathur; 2007), Banks (Debashish, 2006; Galagdera and Edirisuriya, 2007), Sunrise Industries (Kumar, 2004), Coal Mines (Kulshreshtha and Parikh, 2004), Pharmaceutical Firms (Tripathy et al., 2009; Pannu et al., 2010), Power Utilities (Chitkara, 1999; Nag, 2004; Shanmugam and Kulshrestha, 2005; Thakur, 2005; Yadav et al., 2008). Many of the studies also attempted to unravel the determinants of inefficiency in the past to guide future policy planning and managerial intervention for improving the performance level.

Mathur (2007) analysed the performances of the Indian IT industry using DEA and computed the TFP change of the software firms during 1996 to 2006, using MPI. The results indicate that size of firms is important for exports but not for technical efficiency. The analysis also revealed that the average TFP for 1996-2006 is more than one signifying technical progress during the period.

Galagdera and Edirisuriya (2007) studied the performance of Indian commercial banks during the period 1995-2002 using DEA and MPI. The study revealed that there was no significant growth in productivity during the sample period 1995 and 2002. While the productivity grew at 5.8% in 1996, it deteriorated by approximately 0.5% per year during 1997 and 1998 indicating deterioration in bank performance. The variations of efficiency by asset size indicate, largest banks are technically more efficient and the smaller banks less efficient. They have also observed a small difference in the performance of the public and private sector banks due to modest growth in public sector banks in contrast to no growth in private sector banks.

Ali et al. (2007) estimated the relative performance of food processing industry in India. The study revealed that the mean technical efficiency of the industry at about 0.902 with average scale efficiency of 0.870. The industry was estimated to have grown at a rate of 10% per annum during 1980-81 to 2001-02. The productivity growth was mainly contributed due to shift in the production frontier attained by way of increased doses of capital input with contribution from technical efficiency change.

Tripathy et al. (2009) analysed the relative performance of the R&D intensive and non-R&D intensive Pharmaceutical firms in India using DEA, MPI and Tobit Regression. The authors have found that the introduction of the new patent regime, export of goods, inflow of FDI and the profitability of the firms as the key
determinants of efficiency in these firms. The study revealed that R&D intensity and import of capital of goods increases the efficiency of the R&D and non R&D firms respectively. Pannu et al. (2010) analysed the impact of innovation on the performance of Indian Pharmaceutical industry using DEA and MPI and found that sales growth is driven by DEA efficiency, size and age.

3.4 Performance Evaluation of Power Utilities

Numerous studies have been reported in the literature studying the performance level of power utilities of USA, China, India, Japan, Israel, Spain, Turkey, Germany and other countries in distribution as well as generation sector. Golany et al. (1994) studied the relative performance level of thermal power plants in Israel. Chitkara (1999) and Arocena and Price (2002) attempted to measure the efficiency of power generation units in India and Spain respectively. Olatubi and Dismukes (2000), Lam and Shiu (2001) and Sueyoshi (2001) studied the performance levels of electric utilities in US, China and Japan respectively.

Arocena and Price (2002) studied the effect of regulation on the public and private electricity generators in Spain. The study concluded that public coal powered generation plants were more efficient than those in the private sector under cost of service regulation. Price cap regulation have proved to be highly effective incentive mechanism for the private sector in its short run operating decisions and achieved its objective of stimulating efficiency.

Olatubi and Dismukes (2000) analysed the performance of coal fired electric power plants in US for the year 1996 using DEA with Employee Cost, Capital Cost of the plant, Cost of coal, oil and gas burnt, and Variable cost as inputs and Net generation as output, had observed that capital is over utilized by most plants in consistent with the Averch-Johnson effect, which states that firms operating in regulated industries tend to overcapitalize to maximize profits.

Using DEA and assuming individual provinces, autonomous regions and municipalities as DMUs, Lam and Shiu (2001) estimated the efficiency of China’s Thermal power generation during 1995 and 1996. The study revealed that plants in provinces and autonomous regions, which are not under the control of the State Power Corporation, achieved higher levels of efficiency.
Applying a two-stage hierarchy, Cook and Green (2005) investigated the relative operating efficiencies of a set of electric power plants. The study considered full operating hours, sudden and forced outages, equipment de-rating maintenance expenditure and occupied hours as inputs and outputs. Sarica and Or (2005) assessed the efficiency of Turkish power plants using DEA. The study revealed that private sector plants perform significantly better than the public sector plants.

Taking into account utilisation of net capacity, energy losses and operating expenses, Vaninsky (2006) applied DEA to estimate the efficiency of electric power generation in the United States during 1991 to 2004. The findings pointed to a relative stable efficiency from 1994 to 2000 at levels of 99%–100%, and a sharp plunge to 94.61% in 2004.

Yang and Politt (2007) applied six DEA models to evaluate the performance of Chinese coal-fired power plants and explore the determinants of less efficient plants. The study found significant contribution of uncontrollable variables like age of the power plants, average capacity of power generating units etc.

The study carried out by Park and Lesourd (2000) on the conventional fuel power plants in South Korea revealed the BCC efficiency of older plants is significantly lower. The study carried out by Shanmugam and Kulshrestha (2005) have also found that the effect of age on TE is negative @ 0.56% / year. Yang and Politt (2007) have also observed that age affects the technical efficiency of coal fired power plants. Because of technological developments newer units are of higher capacity and better design. However studies investigating if the performance variation is because of age or unit capacity are rare.

The operational efficiency of NTPC power plants was analysed by Chitkara (1999) using DEA. The study found that the performance of some units can be improved by renovation and repowering while the performance of some other units can be improved by extensive training of operating personnel.

Using stochastic production function methodology, the study carried out by Shanmugam and Kulshreshtha (2005) from the unbalanced panel data of 385 observations covering 56 Indian thermal power stations during the period 1991-1992 to 2000-2001 revealed that a) Efficiency varies widely across plants and regions, b)
The average technical efficiency is approximately 73%, indicating a substantial scope for increasing thermal power generation and c) The western region is technically more efficient than other regions and newer plants are more efficient compared to older plants.

Thakur and Kaushik (2004) had studied the performance of Indian electric distribution utilities i.e. the SEBs or their unbundled subsidiaries using DEA. They have found the existences of cost inefficiency i.e. majority of SEBs are not operating at the desired level of cost. Such performance analysis could help inefficient SEBs benchmark against efficient SEBs and improve their efficiency, which will reduce the cost of output, thereby benefiting the consumers. They have argued for legal provisions to maintain a free and transparent performance database for access by public on demand. In another study, Thakur (2005) had assessed the mean efficiency levels of Indian distribution utilities at 68% and the efficiency of 14 out of 26 utilities lie below the average value. She had also observed majority of the SEBs do not seem to operate at the optimum levels and there exist scale inefficiencies and suggested that restructuring and downsizing the present operations may help the utilities to reduce their scale inefficiencies.

Making use of DEA, Nag (2006) proposed a framework to estimate the carbon base line for power generation in India till the end of 11th five year plan period (2010-11) based on the Specific Coal Consumption (SCC), APC, Secondary Oil Consumption and Plant Availability.

Realising the importance of the performance improvement of coal fired power plants in India, CIIGBC undertook a study titled “Make Indian Power Plants World Class”, to catalyse and facilitate performance improvement of power generating units. The study identified best operating parameters for coal and gas based thermal power plants and collated certain best practices which can be further fine tuned and customized to move towards achieving the benchmark figures. The study also revealed that efficient utilization of the existing generation assets, could bring down the capacity addition targets by a whooping 12 GW saving investment to the tune of ₹ 48,000 Crores.

Dash et al. (2008), while exploring alternative matrices for India’s future power demand have observed that there is substantial scope for improvement of the
performance of the thermal power plants and suggested for aggressive action plans for augmenting the current output levels.

3.5 Conclusions from the Chapter

Studies have suggested existence of substantial gap in the performance level of Indian thermal power plants mainly the coal fired ones (CIIGBC, 2005; Subramanian, 2010; Sharma, 2010). Performance improvement of these plants offer a host of benefits in the form of cheaper, more and reliable power; lesser consumption of scarce fossil fuels, emission reductions, better utilization of generation assets etc. there by de-bottlenecking GDP growth and improving the quality of living of the people. Performance evaluation practices being followed in India in respect of power plants are partial in nature and based on average approach as a result identification of overall performance gap is missing.

There has been a shift in performance measurement approach a) from partial measures to overall measures in which the all round performance of firms are measured and b) average to frontier approach in which the performance of a firm is measured not in comparison to industry averages but by making comparison with the industry leaders.

Over the years DEA had emerged as a leading frontier performance evaluation tool and applied to many sectors by researchers and practitioners alike. Several studies have been reported in literature, applying DEA for performance evaluation of different segments of Indian economy also. DEA had also been applied by researchers for performance evaluation of power utilities in generation, transmission and distribution segments internationally. Regulators worldwide apply DEA for performance evaluation, benchmarking, target setting and formulating incentive mechanism for power utilities.

In the context of Indian thermal power plants, overall performance evaluation is rare and does not seem to have addressed the key issues like relative performance evaluation, slack estimation, target setting, benchmarking and productivity change.