Journal of Modelling in Management

A PCA-DEA framework for stock selection in Indian stock market

Dhanya Jothimani, Ravi Shankar, Surendra S Yadav,

A PCA-DEA Framework for Stock Selection in Indian Stock Market

1. Introduction

Portfolio is a collection of various financial assets held by an investor. In order to reduce the portfolio risk, a number of assets are held in a portfolio for the purpose of diversification. Portfolio optimization is a decision-making process where two or more conflicting objectives such as maximizing return and minimizing risks, are considered. The conflicting objectives are solved in three phases: asset selection, asset allocation, and asset management. Asset selection refers to the process of selecting a collection of assets from same or different asset classes. The asset class includes stocks, real estate, and bonds. The process of asset allocation helps the investors to decide how much money can be invested in which asset(s) to reduce the risk and maximize the return. The final step, asset management helps the investors to evaluate the portfolio and define strategies to buy, sell or hold an asset(s).

Markowitz (1952) proposed the Mean-Variance (MV) framework to solve the portfolio optimization problem using quadratic programming. The framework works on the assumption that the asset returns follow the multivariate normal distribution. It aims to find the "efficient frontier", which consists of the combination of various assets that minimizes the risk at various levels of return or minimizes return at various levels of risk. The concept of portfolio optimization has been extended by various researchers (Golmakani and Fazel, 2011; Soleimani et al., 2009; Lin and Ko, 2009; Chang et al., 2009; Cura, 2009; Chen et al., 2009; Doerner et al., 2004).

Despite its wide use and advantages, the standard MV framework has few weaknesses. Real world data does not support the assumption of multivariate normal distribution. This assumption tends to ignore the occurrence of extreme events like 1987 stock market crash (Aouni et al., 2014). An optimal portfolio constructed using MV framework includes a large number of different assets. Implementation of this portfolio with a larger number of smaller assets would increase the transaction cost. The standard MV framework was formulated as a quadratic programming problem. Existing algorithms for solving quadratic programs can find an optimal solution in a reasonable time if the number of assets is small. However, these existing algorithms fail when constraints like bounding constraints and cardinality constraints are considered in the MV framework. Hence, the standard M-V framework with consideration of few additional constraints can be classified as Non-deterministic Polynomial-time hard (NP-hard) problem (Maringer and Kellerer, 2003; Singh et al., 2010).

Today, there is a remarkable increase in the number of listings of companies on a stock exchange. For instance, as of March 2015, Bombay Stock Exchange (BSE) lists more than 5000 companies (BSE, 2015) while National Stock Exchange (NSE) lists more than 1500 companies (NSE, 2015). Various sources of information like financial statements, stock prices, and economic conditions are used by the investors to select stocks. Now, it is challenging for the investors to screen and select the most profitable stocks. Evaluation and selection of assets are considered to be important processes since they influence the asset allocation process. The scope of this paper is limited to asset selection phase of portfolio optimization. The asset type considered here is limited to stocks. The main focus of this paper is to evaluate the financial performance of the firms and, thus, aid the investors to screen and select the stocks for investment. Smith (1990) suggested the application of Data Envelopment Analysis (DEA) to the portfolio optimization problem.

Since its development by Charnes et al. (1978), DEA is a well-acknowledged technique to measure the efficiency of comparable units. But it fails to differentiate the efficient firms from the inefficient ones when the number of input and output parameters is larger than the number of comparable units. This problem is also known as *curse of dimensionality*. In order to overcome this limitation of DEA, a hybrid Principal Component Analysis (PCA) - Data Envelopment Analysis (DEA) is used. The study does not indicate that this is the only method of decision-making, but shows the application of PCA-DEA as a strategy for selection of stocks.

A brief review of previous works is provided in the next section. Section 3 discusses PCA-DEA approach followed by the methodology adopted for the study in Section 4. Section 5 presents results and discussion followed by the study's conclusion in Section 6.

2. Previous Works

The seminal work of Markowitz (1952) laid the stepping stone for portfolio optimization, which later became one of the pioneering research domains of Modern Portfolio Theory. During the past decade, many researchers have used DEA as a strategy for stock selection in various stock exchanges. For instance, DEA was used to construct portfolio(s) of stocks from the property sector in Malaysian stock market (Ismail et al., 2012). The study was carried out for the period 2004-05. The input parameters consisted of dividend yield, trading volume, liquidity, book-to-market, size, price-earnings, risk, leverage ratios, and asset utilization ratios. The output variables were return on equity and return on asset. The study concluded that the portfolio comprising of DEA-efficient firms produced positive Cumulative Annual Returns (CAR) over a long period.

Edirisinghe and Zhang (2008) proposed a relative financial strength indicator (RFSI) using DEA for firms from US technology sector. RFSI was calculated in two ways: first using the raw numbers (total assets, inventory, accounts receivable, revenue, long-term debts, net income and total liabilities) from the financial statements and second, using financial ratios (profitability, liquidity, leverage, valuation, asset utilization and growth ratios). It was observed that the RFSI calculated using latter showed a higher correlation with the stock price. Further, the portfolio constructed using RFSI of financial ratios performed better.

Chen (2008) compared two strategies, namely, size-effect and DEA as a means of stock selection, for companies listed on Taiwan Stock Exchange. In the first strategy, market equity, which is stock price times the number of outstanding, was considered as the size of the firm. The portfolio was formed by including firms of smaller size. In the second strategy, both CCR (refers to the names of the authors Charnes, Cooper, and Rhodes) and BCC (refers to the names of the authors Banker, Charnes, and Cooper) models of DEA were used to select the firms. Based on the performance of the portfolio, the study concluded that size-effect is not an appropriate strategy for portfolio optimization in Taiwan Stock Exchange. The portfolio constructed using DEA produced superior returns.

The effectiveness of DEA as a stock selection strategy was tested for value stocks (i.e., stocks with market value lower than book value) (Pätäri et al., 2010). Valuation ratios (i.e., E/P, D/P, and B/P) were taken as input and output parameters. The sample consisted of Finnish non-financial stocks over the period of 1993-2008. Three variants of DEA, namely, CCR, BCC and super-efficiency models were used. The value portfolios performed better than the market portfolio and the growth portfolios (consisting of stocks with market value higher than book value). Pätäri et al. (2012) carried out a similar study by combining both momentum and value indicators into a single efficiency score. Various combinations of input parameters (enterprise value-per-share and stock price) and output parameters (earnings before interests, depreciation and amortization (EBITDA), Book value-per-share and dividend-per-share) were used in the determination of efficiency scores. Based on the efficiency scores, the stocks were classified into 3-quantile portfolios. The top quantile portfolio performed better than the market portfolio.

Ke et al. (2008) compared the effectiveness of standard DEA and weight-restricted DEA model as a stock selection strategy in China A-share stock market. Standard DEA assumes equal importance to all the input and output parameters whereas weight-restricted DEA considers the relative importance of the input and output parameters. Weight-restricted DEA was used to increase the discriminatory power of DEA. The study concluded that weight-restricted DEA model is effective for constructing portfolios for the investors with a high risk profile. The limitation of weight-restricted DEA is that the relative importance of input and output parameters is based on the subjective judgment of the experts.

Dia (2009) proposed a four-step methodology for portfolio selection. In the first phase, the risk preferences of the investors are specified, followed by the determination of efficiency using DEA. The portfolio is formed based on the investor's risk preferences and the efficiency scores. In the final step, the investors choose the portfolio that suits their risk profile. The sample consisted of 45 stocks from Tunisia Stock Exchange. The constructed portfolios showed superior returns.

In another study by Singh et al. (2010), the portfolios obtained using DEA and ordered-weighted averaging (OWA) operator were compared for Nifty stocks during the period 2005-07. Input and output parameters were variance of the securities and expected returns, respectively. The portfolio formed using DEA showed better returns while the portfolio formed using OWA operator performed better.

Hsu (2014) proposed a four-stage integrated approach for portfolio optimization for stocks in the semiconductor section of Taiwan Stock Exchange. In the first stage, based on historical financial performances, the potential stocks were selected using DEA. In the second stage, the asset allocation was addressed using Markowitz MV framework. The investment proportion was determined using Ant Bee Colony (ABC) optimization. In the third stage, a forecasting model to predict stock price was built using Genetic Programming (GP). In the last stage, transaction rules were defined for buying, selling or holding the stocks.

The above-mentioned studies show the effectiveness of DEA as a stock selection strategy for portfolio optimization in various stock markets. Most of the studies used standard DEA and few used weight-restricted DEA. As already mentioned in the previous section, DEA suffers from the curse of dimensionality. It misclassifies the inefficient firms as efficient firms. The potential stocks for investment are decided based on the efficiency score. In order to overcome the curse of dimensionality, weight-restricted DEA has been used where the relative importance of the output and input parameters is decided based on the expert opinion, which is subjective and may be biased. Hence, misclassification of firms costs a lot to the investors. None of the above-mentioned works addressed the issue of reducing the output and input variables to increase the discriminatory power of DEA without the need for expert's judgment and its implementation for selection of stocks.

3. Principal Component Analysis – Data Envelopment Analysis (PCA-DEA)

This section briefly discusses Data Envelopment Analysis (DEA) and its limitations. It also gives an idea about Principal Component Analysis (PCA) and the formulation of PCA-DEA model.

3.1. Data Envelopment Analysis

In their seminal work, Charnes et al. (1978) and Banker et al. (1984) laid the foundation of Data Envelopment Analysis (DEA) by extending the work of Farrell (1957). DEA is a non-parametric linear programming technique that helps to determine the relative efficiencies of comparable units, which are known as decision making units (DMUs). Relative efficiency is calculated as the ratio of the weighted average of output parameters to the weighted average of input parameters. A DMU is considered to be efficient if it has an efficiency score of 1. Otherwise, it is considered to be inefficient. The fractional form of CCR model for n comparable DMUs is given in Equation 1.

$$\max \quad \eta_a = \frac{\sum_{j=1}^{J} p_{ja} y_{ja}}{\sum_{i=1}^{I} q_{ia} x_{ia}}$$

$$s.t. \quad 0 \le \frac{\sum_{j=1}^{J} p_{ja} y_{jn}}{\sum_{i=1}^{I} q_{ia} x_{in}} \le 1; \quad n = 1, 2, ..., N \qquad (1)$$

$$p_{ja}, q_{ia} \ge 0; \quad i = 1, 2, ..., I; \quad j = 1, 2, ..., J$$

where, xia represents the ith input of the ath DMU;

qia represents the weight of that input;

y_{ja} represents the j^{th} output of the a^{th} DMU;

p_{ja} represents the weight of that output;

x_{in} and y_{jn} represent the i^{th} input and j^{th} output, respectively of the n^{th} DMU.

η_a represents the relative efficiency of DMU_a with respect to other DMUs.

The above equation represents the calculation of efficiency of DMU a with respect to other DMUs, the formulation has to be repeated for remaining DMUs. The fractional DEA model is difficult to solve, hence they are converted to linear programming formats, which are simple and easy to solve. The linear programming formulations of CCR model and constant returns-to-scale additive models for n comparable units are represented in equations 2 and 3, respectively. Readers are suggested to refer Adler and Yazhemsky (2010) to read more about different DEA models and their properties.

$$\begin{aligned}&\max_{Q,P} PY^{a}\\&s.t.\quad QX^{a}=1\\&QX-PY\geq 0\\&P,Q\geq 0\end{aligned} \tag{2}$$

$$\begin{aligned}&\min_{Q,P} QX^{a}-PY^{a}\\&s.t.\quad QX-PY\geq 0\\&P\geq e\\&Q\geq e\end{aligned} \tag{3}$$

where, j outputs are denoted by Y;

i inputs are denoted by X;

e is vector of ones;

X^a represents the input column of DMU under consideration, DMU_a;

Y^a represents the output column of DMU under consideration, DMU_a;

P represents the vector of output weights;

Q represents the vector of input weights.

DEA has been applied in various fields including health organizations (Jorge et al., 2013), strategic decision making (Saen and Azadi, 2011) and performance evaluation of metallurgical firms (Tsolas, 2014). Wide applicability and use of DEA is due to its ability to handle multiple input and output parameters for evaluating the relative efficiency. But the accuracy of evaluation of efficiency is affected by the proportion of number of DMUs and number of output and input parameters. Golany and Roll (1989) recommended that the number of DMUs should be at least twice the total number of input and output parameters. Dyson et al. (2001) suggested that the number of DMUs should be more than twice the product of number of output and input parameters. The conventional way to resolve this issue is to eliminate few parameters. Elimination of even one input/output parameter affects the efficiency scores (Adler and Yazhemsky, 2010). The next approach is the use of relative importance scale for input and output parameters with the help of expert opinion as in the case of weight-restricted DEA (Ke et al., 2008). The major limitation of expert opinion is that it is biased and subjective.

Several attempts have been made by the researchers to reduce the number of input and output variables for DEA. These include integrating principal component analysis with DEA (PCA-DEA) (Ueda and Hoshiai, 1997; Adler and Golany, 2001), variable reduction using partial correlation (Jenkins and Anderson, 2003), regression-based analysis (Ruggiero, 2005) and efficiency contribution measure (ECM) (Pastor et al., 2002). Adler and Yazhemsky (2010) and Nataraja and Johnson (2011) reviewed and compared all these approaches. The performance of ECM is moderate, but the run-time is long. Regression-based analysis requires shorter run-time, but the performance is not as good as ECM. Variable reduction using partial correlation is same as that of removal of one or more input and/or output parameters, whereby the efficiency score gets affected. Of all these approaches, PCA-DEA seemed to perform better as the information loss is less. PCA-DEA retains the information in the original variables in the form of principal components. It neither eliminates an entire parameter nor requires the need for expert opinion. Due to the non-iterative nature of PCA-DEA, the run-time is smallest.

3.2. Principal Component Analysis

Principal Component Analysis (PCA) is a statistical technique that transforms a dataset into a set of linearly uncorrelated variables called principal components (PC) (Hair et al., 2009). Let the random vector of output (or input parameters) to be transformed be $Y = [Y_1, Y_2, ..., Y_p]$. The correlation matrix of vector Y is represented as C with normalized eigenvectors $l_1, l_2, ..., l_p$ and eigenvalues $\lambda_1 \geq \lambda_2 \geq ... \geq \lambda_p \geq 0$.

The linear combinations are represented as:

$$Y_{PC_i} = l_i^t Y = l_{1i}^t Y_1 + l_{2i}^t Y_2 + ... + l_{pi}^t Y_p \tag{4}$$

$$Var(Y_{PC_i}) = l_i^t C l_i, i = 1, 2, ..., p \tag{5}$$

$$Correlation \quad (Y_{PC_i}, Y_{PC_k}) = l_i^t C l_k, \quad k = 1, 2, ..., p, \quad i = 1, 2, ..., p, \quad i \neq k \tag{6}$$

where t represents the transpose operator.

The variance of the dataset is explained by PCs in a certain ratio, where ratio represents the eigenvalues of the correlation matrix *C*. Descending order of the eigenvalues indicates that the PC_1 explains the maximum variance of the dataset compared PC_2 and so on. This indicates the PC_1 is correlated with at least few of the original output (or input parameter, as the case may be). The number of PCs formed is equal to or less than the number of variables in the dataset. The orthogonality of the eigenvectors accounts for the uncorrelated principal components. In other words, PC_1 is not correlated to PC_2, which is not correlated to its lower level PCs. All the components extracted have two important properties: each component explains the maximum variance in the dataset that was not explained by its predecessor, and the components are uncorrelated with its respective preceding component.

3.3. PCA-DEA Formulation

The concept of integrating Principal Component Analysis (PCA) with Data Envelopment Analysis (DEA) was developed independently by Ueda and Hoshiai (1997) and Adler and Golany (2001, 2002). The PCA-DEA model overcomes the limitation of DEA model, which fails to discriminate the efficient DMUs from the inefficient DMUs when the number of input and output parameters is large in comparison to the number of DMUs.

The DEA models in equations 2 and 3 are transformed to accommodate the principal components, instead of the original data. As seen in equation 4, principal components are nothing but the linear combination of weighted input and output parameters. The formulation of both CCR and additive are transformed in such a way that when all the PCs are included in the PCA-DEA model, the resulting efficiency would be same as that of standard DEA. The PCA-DEA formulation for CCR and constant returns-to-scale additive models for *n* comparable units are given in equations 7 and 8, respectively (Adler and Yazhemsky, 2010).

$$max_{Q_{PC}P_{PC}}\ P_{PC}^{t}Y_{PC}^{a} \tag{7}$$

$$\begin{gathered} s.t.\ Q_{PC}^{t}X_{PC}^{a} = 1 \\ Q_{PC}^{t}X_{PC} - P_{PC}^{t}Y_{PC} \ \geq 0 \\ Q_{PC}^{t}L_{x} \ \geq 0 \\ P_{PC}^{t}L_{y} \ \geq 0 \end{gathered}$$

Q_{PC} and P_{PC} are free

$$min_{Q_{L_y}P_{L_x}} \; Q_{PC}X^a_{PC} - P_{PC}Y^a_{PC} \tag{8}$$

$$s.t.\, Q_{PC}X_{PC} - P_{PC}Y_{PC} \geq 0$$

$$Q_{PC}L_x \geq e$$

$$P_{PC}L_y \geq e$$

$$Q_{PC_i} - Q_{PC_{i+1}} \geq 0$$

$$P_{PC_i} - P_{PC_{i+1}} \geq 0$$

$$for\ i = 1, \ldots, m-1, \qquad Q_{PC}\ and\ P_{PC}\ are\ free$$

where, Q_{PC} and P_{PC} are the vectors of output and input weights, respectively. From the definition of principal components, we know that $Y_{PC} = L_yY$ and $X_{PC} = L_xX$, where L_y and L_x are the PCA coefficients of output and input data, respectively. So, $Q_{PC} \equiv Q_{PC}\,L_xX$, which implies that $Q = Q_{PC}\,L_x$ and $Q_{PC} \equiv P_{PC}\,L_yY$ implying $P = P_{PC}L_y$ (Adler and Golany, 2002). This model is equivalent to that of the standard DEA model where *PCs* explain 100% variance of the dataset.

In PCA, PCs are prioritized in descending order of importance. Constraints $Q_{PCi} - Q_{PCi+1} \geq 0$ and $P_{PCi} - P_{PCi+1} \geq 0$ ensure that the weight of PC_1 is at least equal to or greater than PC_2, the weight of PC_2 to be at least equal to or greater than PC_3 and so on. In this way, PCA increases the discriminating ability of the DEA model (Adler and Golany, 2002).

PCs are obtained through the orthogonal rotation of the coordinate system than the parallel translation. This property enables the application of PCA-DEA model to all basic DEA models such as standard radial Constant Returns to Scale and Variable Returns to Scale model, without affecting their basic properties (Adler and Yazhemsky, 2010). The PCA-DEA model is robust to sample size (Yap et al., 2013). The conventional way of variable reduction is dropping the variable(s), which would underestimate the efficiency score. There is a loss of information when a variable is not considered for the analysis. In PCA-DEA model, the PCs that do not contribute much to the variance of the data are dropped. The complete information on an input or output parameter is not lost until the principal component weight representing the variable is eliminated. In this way, the information loss is minimized. In addition, PCA-DEA is similar to weight-restricted DEA without the need for expert opinion, which is subjective.

4. Methodology

Section 4.1 deals with the process of data collection, followed by selection of input and output parameters in Section 4.2. Section 4.3 briefly explains the application of PCA-DEA model for evaluation of the efficiency of the stocks quoted on NSE.

4.1. Data Description

The data sample consists of non-financial stocks quoted on National Stock Exchange (NSE). Established in 1992, NSE is one of the leading stock exchanges in India. As of January 2015, it is the world's 12th largest stock exchange in terms of market capitalization (NSE, 2015). It was set up to bring transparency in the Indian capital market. It was the first stock exchange in the country to ease the trading facility by providing fully automated electronic trading system.

The total number of non-financial stocks quoted on NSE as of July, 2014 was 1523. On the basis of similar business functionality, the stocks were classified into 15 sectors. Investing in stocks from different sectors helps to diversify the risk. The period of study is 2006-2013.

In an exchange, there are three kinds of trading statuses: active, inactive and suspended. Active stocks are those that are being actively traded. Inactive stocks refer to those that are not listed either on NSE or BSE but on regional stock exchanges. It also refers to those stocks that are listed on NSE or BSE but are not being traded actively. The stocks that do not comply with the rules and regulations of the stock exchanges are suspended from being traded (SEBI, 2014). Hence, the stocks that were inactive and suspended during the period of study were not considered. Also, the stocks that were delisted during the study period and also those with incomplete data were removed. The stocks with alternating fiscal year ends during the study period were eliminated. For instance, the fiscal year ends of Eicher Motors were March and December during the period of 2006-08 and 2008-13, respectively. Hence, Eicher Motors was not considered for analysis. At the end of this process, the number of stocks was reduced to 523 as shown in Table 1. Both the financial statement and the stock price data were collected from ACE Equity database.

Table 1: Sectors and Number of Firms

S.No.	Sector	Number of Firms
1	Food & Beverages	33
2	Personal Products	13
3	Textile	49
4	Industrial Metals	31
5	Chemicals	76
6	Construction	27
7	Pharmaceutical	41
8	Media	21
9	ICT	47
10	Automobile	52
11	Construction Material	31
12	Consumer Services	35
13	Power	10
14	Engineering	35
15	Electric Equipment	22
	Total	**523**

4.2. Input and Output Parameters

Selection of input and output parameters plays a vital role in evaluating the relative financial performance of the firms (Golany and Roll, 1989). Hence, they must be selected appropriately. While determining the financial health of the firm, raw financial numbers such as total assets and total liabilities do not aid in differentiating the healthy stocks from the unhealthy stocks. Hence, the ratios calculated using data from the financial statement of the firms are chosen to establish the financial performance of the firms (Edirisinghe and Zhang, 2007). The financial ratios of financial firms are different from that of non-financial firms, hence, the latter group is not considered for the analysis.

The financial health of a firm can be analyzed using six aspects: liquidity, leverage, asset utilization ratio, profitability, growth ratio and valuation ratio. The generation of revenue by a firm is explained using profitability and growth ratios, hence they are classified as output parameters. Valuation ratio explains the perception of the investors towards the success of the firm. The planning and operational strategies of the firm are evaluated using asset utilization, liquidity, and leverage ratios; hence, they constitute the input parameters for PCA-DEA analysis

(Edirisinghe and Zhang, 2008). A set of 18 financial ratios is used as input and output parameters as shown in Table 2. These parameters are supported by the work of Ismail et al. (2012), Tehrani et al. (2012), Hwang et al. (2010), Edirisinghe and Zhang (2008), Feroz et al. (2003) and Powers and McMullen (2000).

4.3.PCA-DEA Framework for Stock Selection

DEA model evaluates the relative efficiency of the comparable units called Decision-Making Units (DMUs). Here, DMUs consist of firms from different sectors. As seen from Table 1, the number of DMUs in few sectors, namely, power sector and construction sector do not satisfy the rule of thumb proposed by Golany and Roll (1989). Except for Chemicals, none of the sectors satisfies the rule of thumb proposed by Dyson et al. (2001) (Refer Section 3.1). This affects the discriminatory power of the standard DEA model, thus, classifying the majority of DMUs as efficient. Integrated PCA-DEA approach overcomes this problem of DEA (Adler and Yazhemsky, 2010). Further, PCA-DEA model is robust to the size of DMUs (Yap et al., 2013). Hence, PCA-DEA is approach is used for the analysis.

Table 2: Input and Output Parameters

Parameter	Measures	Financial Ratio
Input	Liquidity Ratios	Debt to Equity Ratio
		Quick Ratio
		Current Ratio
	Leverage Ratios	Solvency Ratio I
		Solvency Ratio II
		Leverage Ratio
	Asset Utilization	Asset Turnover
		Inventory Turnover
		Receivables Turnover
Output	Profitability Ratio	Return-on-Assets
		Earnings per Share
		Net Profit Margin
		Return-on-Equity
	Growth Ratio	Revenue Growth Rate
		Earnings per Share Growth Rate
		Net Income Growth Rate
	Valuation Ratio	Price to Earnings Ratio
		Price to Book Ratio

In the data pre-processing phase, the data is normalized by dividing the elements of each input and output data by its respective mean. This ensures that the data set has similar magnitude. Data with similar magnitude helps to overcome the scaling issues and the round-off errors problems faced by few mathematical programming softwares (Sarkis, 2007).

The financial performance of the firms from various sectors during the period of study is evaluated using the integrated PCA-DEA approach. The PCA-DEA approach for stock selection is shown in Figure 1. The steps of the approach are enumerated below:

1. *Reduction using PCA:* In this phase, consider a sector for a particular year. Run Principal Component Analysis (PCA) for inputs and outputs separately.
2. *Stepwise Reduction of Information Level:* Information level refers to the amount of information that should be retained in the model. In other words, it represents the amount of variance in the input and output that is explained by their respective principal components (X_{PC} and Y_{PC}). The information level in the model is decreased by 5% at each stage starting from 100% till it reaches 80%. When the information level is reduced below 80%, the model classifies the efficient firms as inefficient. In order to avoid this overestimation bias, the efficiency is calculated till 80% information level. As mentioned in the previous section, when PCs explain 100% variance of dataset, the PCA-DEA model is nothing but the standard DEA model.
3. *Evaluation of Efficiency:* At each information level the relative efficiency of the firm is calculated using Equation 7.
4. Repeat steps 2 and 3 for remaining years.
5. Repeat steps 1, 2 and 3 for remaining sectors.
6. Examine the efficiency of the firms at various information levels (i.e., 100%, 95%, 90%, 85% and 80%). Identify the desired information level that meets the investor's desired level of discrimination.
7. *Stock Selection Criterion:* Similar to DEA model, the firms with the efficiency score of 1 are considered to be efficient in PCA-DEA model and rest are considered to be inefficient. The criterion for selecting stocks is that for a desired information level, the stocks should be efficient throughout the period of study.

Let us consider the firms from Personal Products sector for the year 2006. PCA was carried on inputs and outputs separately to obtain their respective principal components and the percentage variance explained by them. Determine the relative efficiencies using Equation 7 for 100% information level. This refers to the standard DEA model. Reduce the information level by 5%. Determine the efficiency for 95% information level. Repeat the process till the information level reaches 80% for firms from the personal sector for the year 2006. For the same sector, repeat the process for remaining years (2007-2013). The complete process[*] is repeated for other sectors for all years. Identify the desired information level and apply the stock selection criterion for all sectors.

5. Results and Discussion

This section presents the results of the PCA-DEA model. First, the PCA results are discussed, followed by the relative efficiency of the firms and recommendations to the investors. In addition, managerial implications for the firms are discussed.

5.1.Principal Component Analysis

PCA helped to increase the discriminatory power of DEA by reducing the dimensions of the linear program. The PCA results of firms from Personal Products sector for the year 2013 are shown in Table 3. From the table, it can be seen that the first principal component of output parameter PC_{y1} captures at least 77% of the variance in the data. When PC_{y2} and PC_{y3} are considered (in addition to PC_{y1}), almost 95% of the total variance is explained. In case of input, PC_{x1}, PC_{x2}, PC_{x3} and PC_{x4} explain at least 95% of the total variance in the data. In Table 3, rows 3-11 and 14-22 represent the PCA linear coefficients of inputs and outputs, respectively.

When the information level is 95% (i.e., 95% of the variance is explained by PCs). PC_{y4} to PC_{y9} of the output parameters can be dropped for efficiency calculation. It can be seen from Table 3, none of PCs dropped solely influences an entire output variable $L_{y1},...,L_{y9}$, with no contribution of other PC combinations. Though dropping of few PCs causes little loss of information but it is not same as dropping of an entire input/output parameter.

[*]Computations carried out using software available on http://pluto.huji.ac.il/ msnic/PCADEA.htm

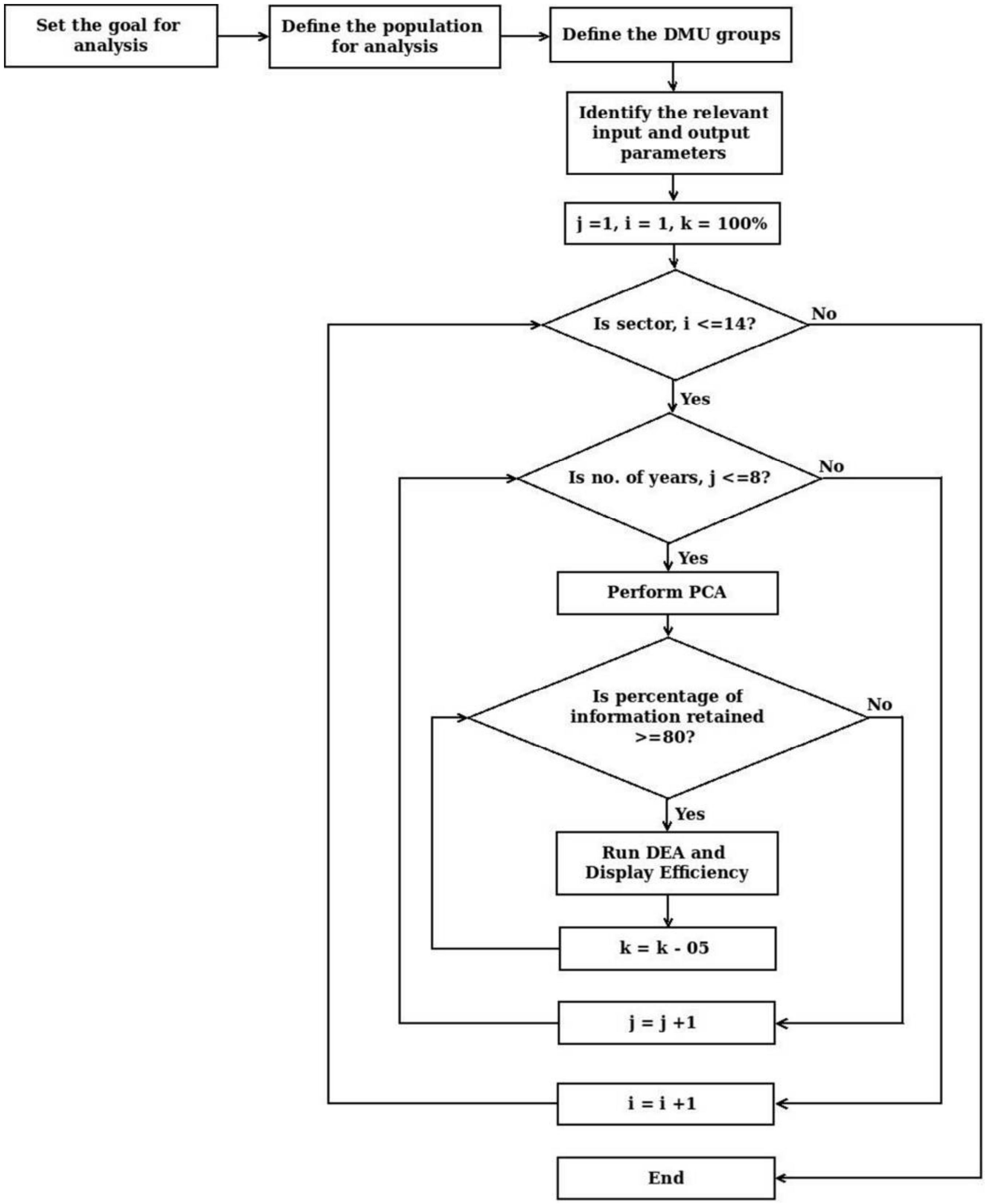

Figure 1: Proposed Approach for Stock Selection using PCA-DEA

Table 3: Principal Component Analysis for Firms under Power Sector for the Year 2006

Inputs (L_x)	**PC_{x1}**	**PC_{x2}**	**PC_{x3}**	**PC_{x4}**	**PC_{x5}**	**PC_{x6}**	**PC_{x7}**	**PC_{x8}**	**PC_{x9}**
% Correlation Explained	43.611	29.937	11.431	9.709	4.469	0.679	0.114	0.045	0.005
L_{x1}	0.4633	-0.1669	0.3106	-0.2437	0.0052	0.4377	-0.1866	-0.4032	0.4612
L_{x2}	0.2123	0.5093	-0.1276	0.1457	0.4831	-0.1558	0.5098	-0.309	0.2123
L_{x3}	-0.0005	0.2025	0.6309	0.5449	-0.3838	0.1552	0.2792	0.1208	0.0124
L_{x4}	0.1638	-0.3678	-0.2224	0.6706	-0.0484	-0.3667	-0.2747	-0.3114	0.1716
L_{x5}	-0.1501	-0.2577	0.5396	0.0879	0.7305	-0.0905	-0.1698	0.0661	-0.1926
L_{x6}	-0.2063	-0.0889	-0.3524	0.3605	0.2105	0.7802	0.0534	-0.0752	-0.184
L_{x7}	0.042	-0.6703	-0.0431	-0.0891	-0.0012	-0.0357	0.6938	0.1728	0.1629
L_{x8}	0.4129	0.0978	-0.1496	0.1645	0.1943	0.0866	-0.1806	0.7675	0.3219
L_{x9}	-0.6899	0.0575	0.0118	0.0062	0.0353	-0.0051	-0.0787	0.038	0.7152
Outputs (L_y)	**PC_{y1}**	**PC_{y2}**	**PC_{y3}**	**PC_{y4}**	**PC_{y5}**	**PC_{y6}**	**PC_{y7}**	**PC_{y8}**	**PC_{y9}**
% Correlation Explained	77.368	14.190	4.351	3.121	0.691	0.233	0.032	0.012	0.003
L_{y1}	0.3594	0.1782	0.2453	0.3335	-0.3621	-0.3695	-0.3631	0.3719	0.3604
L_{y2}	-0.2497	0.7225	0.5148	0.0629	0.1391	-0.1081	0.2366	-0.0053	-0.244
L_{y3}	-0.0282	-0.5261	0.7307	0.2167	-0.1531	0.168	0.0588	-0.2926	-0.0299
L_{y4}	-0.1865	-0.0715	0.3442	0.8339	-0.2196	0.1982	0.1483	-0.0759	-0.1764
L_{y5}	0.3652	0.0662	0.0572	0.311	0.7543	0.2316	0.0135	-0.0775	0.3653
L_{y6}	0.3371	0.309	0.0074	-0.2114	-0.4457	0.611	0.2838	-0.1142	0.2839
L_{y7}	0.1852	-0.0444	0.1179	0.0248	-0.0763	-0.5937	0.5959	-0.4203	0.2431
L_{y8}	0.0148	-0.2499	0.0721	0.0184	0.0658	0.0496	0.5924	0.7573	-0.025
L_{y9}	-0.701	0.0022	0.0083	-0.0147	-0.0108	0.0483	0.0012	0.0343	0.7104

5.2.Efficiency of the Firms

The firms that have an efficiency score of 1 are considered to be efficient, otherwise they are inefficient firms. Descriptive statistics of the relative efficiency of the firms from Food & Beverage sector during year 2006 for various information levels is shown in Table 4. First row represents the percentage of information retained in the model (i.e., the amount of variance explained by PC_xs and PC_ys. As mentioned earlier, when PC_xs and PC_ys explain 100% variance in the dataset, it is same as the standard DEA model. Hence, column 2 represents the descriptive statistics of the standard DEA model. Columns 3-6 represent the descriptive statistics of PCA-DEA model for 95% to 80% information level. Around 82% of firms are found to be DEA efficient while only 52% of the firms are found to be efficient in PCA-DEA model at 95% information level. It can be observed that the standard DEA model overestimated the inefficient firms.

Table 4: Descriptive Statistics of Efficiency of firms (Food & Beverage Sector, year 2006) for various Information Levels

	Information Level in the Model				
	100%	95%	90%	85%	80%
Minimum Efficiency	0.4314	0.2168	0.2151	0.1198	0.1198
Maximum Efficiency	1	1	1	1	1
Average Efficiency	0.9479	0.8499	0.8263	0.7353	0.7353
Efficient Firms (%)	81.82	51.52	45.45	30.30	30.30
Total Number of Firms	33	33	33	33	33

The efficiency scores of the firms from the Construction sector throughout the 8 years (i.e., 2006-2013) are shown in Table 5. The first column represents the firms, i.e., DMUs for PCA-DEA model. Columns 2 to 9 represent the efficiency of the firms for years 2006-2013 when the information level is 100% (i.e., standard DEA model). Columns 10 to 17 represent the efficiencies of firms when 95% variance of dataset is explained. The efficiency scores corresponding to 90%, 85% and 80% of information level are not shown here. From Table 5, it can be observed that 89 per cent of firms are found to be efficient using the standard DEA model (i.e., when the information level is 100%). This indicates the inability of standard DEA model to discriminate the firms when the number of DMUs (in this case, 27) is less than twice the number of input and output variables (i.e., 2*(9+9) = 36) (Dyson et al., 2001).

The firms that are efficient throughout 8 years (i.e., from 2006-2013) are considered to be the potential candidates for next stage of portfolio optimization, i.e., asset allocation. For instance, Larsen & Toubro Ltd. is found to be efficient in all 8 years, hence, can be shortlisted for investment. Whereas, Bharti Shipyard Ltd. found to be efficient in all years except 2012, is not considered for investment. The same criterion is applied to firms in remaining 13 sectors. The number of potential candidates in different sectors is shown in Table 6. The stocks from different sectors are considered in order to diversify the unsystematic risk in a portfolio, thus reducing the total risk of the portfolio. A total of 115 potential stocks are obtained using standard DEA while 41 potential stocks are obtained using PCA-DEA. This reiterates the fact that the DEA model overestimates the efficient firms.

6. Conclusion

The focus of this study was to screen and select the stocks quoted on NSE based on their historical financial performances. The financial performances were assessed in the form of financial efficiency using standard DEA and PCA-DEA with varying information levels. Though standard DEA model aided in the selection of the stocks, but it misclassified several inefficient firms as efficient. Since portfolio optimization is a decision-making process where investors select the stocks with high return and low risk, investing in an inefficient firm could be costly. In order to overcome this problem, PCA-DEA model was used, where PCA helped to reduce the number of input and output parameters with minimum loss of information. Unlike weight-restricted DEA, PCA-DEA does not employ expert's opinion to discriminate the efficiency of the firms and thus, avoiding subjectiveness of their judgments.

In addition, the results obtained from the model can be used by the firms to benchmark their performance with those that are best within their industry. The inefficient firms can try to adopt the process(es) of the efficient firms to improve their performance.

Table 5: PCA-DEA results for Construction Sector

% Information Retained	100%								95%							
DMUs	2013	2012	2011	2010	2009	2008	2007	2006	2013	2012	2011	2010	2009	2008	2007	2006
Larsen & Toubro Ltd.	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**
Mahindra Lifespace Developers Ltd.	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**
Unitech Ltd.	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**
Bharati Shipyard Ltd.	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	1	0.5004	1	1	1	1	1	1
BSEL Infrastructure Realty Ltd.	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	0.1342	1	0	0	1	1	1	1
DS Kulkarni Developers Ltd.	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	1	0.4408	0.4356	1	1	1	1	1
Era Infra Engineering Ltd.	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	0.5112	0.5665	0.6798	0.5534	0.8145	0.6748	0.5309	1
KEC International Ltd.	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	1	0.6455	0.6713	0.7372	0.3769	0.7804	0.5660	0.6327
Noida Toll Bridge Company Ltd.	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	1	1	1	0.7137	1	1	0.8347	0.5033
Pratibha Industries Ltd.	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	0.5884	0.7186	1	0.8104	0.6787	0.8704	0.4724	0.8875
Reliance Industrial Infrastructure Ltd.	**1**	**1**	**1**	**1**	**1**	**1**	**1**	**1**	1	1	1	1	1	1	0.9461	1
Ansal Properties & Infrastructure Ltd.	1	1	1	1	0.7143	1	1	1	0.2701	0.3636	0.6196	0.6331	0.3515	0.7597	1	1
BL Kashyap & Sons Ltd.	1	1	1	0.7775	1	1	1	1	0.3253	1	0.6701	0.3444	0.5862	1	1	1
Hindustan Construction Company Ltd.	1	1	1	1	1	1	1	0.6394	0.1136	0.1719	0.3119	0.4796	0.4514	0.6243	0.6306	0.5373
ITD Cementation India Ltd.	0.9084	1	1	1	1	0.6989	1	1	0.2119	0.5783	1	0.5086	0.7634	0.2937	1	0.4327
Jyoti Structures Ltd.	1	1	1	0.9943	1	0.9819	0.9166	0.7845	0.3190	0.4248	0.9467	0.6733	0.6241	0.5722	0.7020	0.6592
Kalpataru Power Transmission Ltd.	1	0.9748	1	1	0.9086	1	1	1	0.6826	0.6757	1	0.7711	0.3791	0.7854	1	1
Mukand Engineers Ltd.	1	0.8908	1	1	1	1	0.6814	0.5545	0.6184	0.4520	0.8038	1	1	0.4301	0.3473	0.1704
NCC Ltd.	0.7387	0.7731	0.6857	0.7755	0.6981	0.8443	0.7547	0.9039	0.2531	0.4240	0.3300	0.5369	0.3846	0.5794	0.6082	0.7655
Patel Engineering Ltd.	0.5714	1	0.9632	1	1	1	1	1	0.1773	0.9636	0.4532	0.5115	0.9899	0.7458	1	1
PBA Infrastructure Ltd.	0.2	1	1	0.5907	0.3945	0.4897	0.5721	0.6948	0.1047	0.2739	0.3274	0.3304	0.1979	0.3098	0.2890	0.5612
Petron Engineering Construction Ltd.	1	1	1	1	0.9872	1	0.4977	1	1	1	1	1	0.5843	0.6182	0.4428	0.8326
Punj Lloyd Ltd.	1	1	1	0.8071	1	1	0.9845	1	1	1	1	0.4540	0.6678	1	0.7714	1
Sadbhav Engineering Ltd.	1	1	1	0.8632	1	0.9433	0.8975	0.6709	0.7095	0.9875	1	0.4924	0.735	0.8704	0.7573	0.5138
Simplex Infrastructures Ltd.	0.6392	1	0.9946	0.8784	0.9664	1	1	1	0.1765	0.5520	0.6007	0.3868	0.6990	0.7810	0.4681	0.6451
SPML Infra Ltd.	1	1	0.6043	0.6351	0.7421	1	1	1	0.4674	1	0.4229	0.4252	0.4062	0.7067	1	1
Welspun Projects Ltd.	1	1	0.3715	0.5252	0.9419	0.9602	1	0.7340	0.9090	1	0	0.3669	0.7713	0.6975	0.5698	0.4102
Efficient Firms (%)	**81**	**89**	**81**	**67**	**70**	**78**	**74**	**74**	**37**	**41**	**44**	**30**	**33**	**37**	**44**	**52**

Table 6: Number of Efficient Firms in Different Sectors

		% Information Retained	
S.No.	**Sector**	**100%**	**95%**
1	Food & Beverages	10	3
2	Personal Products	8	1
3	Textile	6	3
4	Industrial Metals	7	3
5	Chemicals	7	5
6	Construction	11	3
7	Pharmaceutical	8	4
8	Media	2	1
9	ICT	7	2
10	Automobile	8	3
11	Construction Material	7	1
12	Consumer Services	11	4
13	Power	8	1
14	Engineering	9	4
15	Electric Equipment	6	3
	Total	**115**	**41**

Here, the efficiencies of two DMU groups have not been compared. It is quite possible that the efficient stocks under one DMU group might be inefficient when compared to inefficient stocks of second DMU group. For instance, Unitech Ltd is found to be efficient in Construction sector. But it may not be efficient compared to SRF Ltd., an inefficient firm in Textile sector. Two different DMU groups can be compared following the work of Banker et al. (2010) and Barrett and Donald (2003).

The study can be extended to next phase of portfolio optimization, i.e., asset allocation, where the stocks selected in this can be used as the input. The applicability of the model can be tested for stock selection in other stock exchanges.

References

Adler, N., Golany, B., 2001. Evaluation of deregulated airline networks using data envelopment analysis combined with principal component analysis with an application to Western Europe. European Journal of Operational Research 132 (2), 260 - 273.

Adler, N., Golany, B., 2002. Including principal component weights to improve discrimination in data envelopment analysis. Journal of the Operational Research Society 53 (1), 985{991.

Adler, N., Yazhemsky, E., 2010. Improving discrimination in data envelopment analysis: PCA-DEA or variable reduction. European Journal of Operational Research 202 (1), 273 - 284.

Aouni, B., Colapinto, C., Torre, D. L., 2014. Financial portfolio management through the goal programming model: Current state-of-the-art. European Journal of Operational Research 234 (2), 536 - 545.

Banker, R. D., Charnes, A., Cooper, W. W., 1984. Some models for estimating technical and scale inefficiencies in data envelopment analysis. Management Science 30 (9), 1078 - 1092.

Banker, R. D., Zheng, Z. E., Natarajan, R., 2010. DEA-based hypothesis tests for comparing two groups of decision making units. European Journal of Operational Research 206 (1), 231 - 238.

Barrett, G., Donald, S., 2003. Consistent tests for stochastic dominance. Econometrica 71 (1), 71 - 104.

BSE, 2015. Bombay Stock Exchange. www.bseindia.com, accessed on February 15, 2015.

Chang, T.-J., Yang, S.-C., Chang, K.-J., 2009. Portfolio optimization problems in different risk measures using Genetic Algorithm . Expert Systems with Applications 36 (7), 10529 - 10537.

Charnes, A., Cooper, W., Rhodes, E., 1978. Measuring the efficiency of decision making units. European Journal of Operational Research 2 (6), 429 - 444.

Chen, H., 2008. Stock selection using data envelopment analysis. Industrial Management & Data Systems 108 (9), 1255 - 1268.

Chen, J.-S., Hou, J.-L., Wu, S.-M., Chang-Chien, Y.-W., 2009. Constructing investment strategy portfolios by combination genetic algorithms. Expert Systems with Applications 36 (2, Part 2), 3824 - 3828.

Cura, T., 2009. Particle Swarm Optimization approach to portfolio optimization. Nonlinear Analysis: Real World Applications 10 (4), 2396 - 2406.

Dia, M., 2009. A portfolio selection methodology based on Data Envelopment Analysis. INFOR 47 (1), 71- 79.

Doerner, K., Gutjahr, W., Hartl, R., Strauss, C., Stummer, C., 2004. Pareto Ant Colony Optimization: A metaheuristic approach to multiobjective portfolio selection. Annals of Operations Research 131 (1-4), 79 - 99.

Dyson, R., Allen, R., Camanho, A., Podinovski, V., Sarrico, C., Shale, E., 2001. Pitfalls and protocols in DEA. European Journal of Operational Research 132 (2), 245 - 259.

Edirisinghe, N., Zhang, X., 2007. Generalized DEA model of fundamental analysis and its application to portfolio optimization. Journal of Banking & Finance 31 (11), 3311 - 3335.

Edirisinghe, N. C. P., Zhang, X., 2008. Portfolio selection under DEA-based relative financial strength indicators: Case of US industries. The Journal of the Operational Research Society 59 (6), 842 - 856.

Farrell, M. J., 1957. The measurement of productive efficiency. Journal of the Royal Statistical Society. Series A (General) 120 (3), 253 - 290.

Feroz, E. H., Kim, S., Raab, R. L., 2003. Financial statement analysis: A data envelopment analysis approach. The Journal of the Operational Research Society 54 (1), 48 - 58.

Golany, B., Roll, Y., 1989. An application procedure for DEA. Omega 17 (3), 237 - 250.

Golmakani, H. R., Fazel, M., 2011. Constrained portfolio selection using Particle Swarm Optimization . Expert Systems with Applications 38 (7), 8327 - 8335.

Hair, J. F., Black, W. C., Babin, B. J., Anderson, R. E., 2009. Multivariate Data Analysis, 7th Edition. Prentice Hall, NJ.

Hsu, C.-M., 2014. An integrated portfolio optimisation procedure based on data envelopment analysis, artificial bee colony algorithm and genetic programming. International Journal of Systems Science 45 (12), 2645 - 2664.

Hwang, S.-N., Chuang, W.-C., Chen, Y.-C., 2010. Formulate stock trading strategies using DEA: A Taiwanese Case. INFOR 48 (2), 75 - 81.

Ismail, M., Salamudin, N., Rahman, N., Kamaruddin, B., May 2012. DEA portfolio selection in Malaysian stock market. In: International Conference on Innovation Management and Technology Research (ICIMTR). pp. 739 - 743.

Jenkins, L., Anderson, M., 2003. A multivariate statistical approach to reducing the number of variables in data envelopment analysis. European Journal of Operational Research 147 (1), 51 - 61.

Jorge, M. J., de Carvalho, F. A., Jorge, M. F., de Oliveira Medeiros, R., de Souza Ferreira, D., 2013. Efficiency analysis in public health organizations in brazil. Journal of Modelling in Management 8 (2), 241 - 254.

Ke, J., Qiao, J., Wang, G., Nov 2008. Empirical analysis of portfolio optimization based on DEA model. In: International Seminar on Future Information Technology and Management Engineering, 2008 (FITME '08). pp. 490 - 493.

Lin, P.-C., Ko, P.-C., 2009. Portfolio value-at-risk forecasting with GA-based extreme value theory. Expert Systems with Applications 36 (2, Part 1), 2503 - 2512.

Maringer, D., Kellerer, H., 2003. Optimization of cardinality constrained portfolios with a hybrid local search algorithm. OR Spectrum 25 (4), 481 - 495.

Markowitz, H., 1952. Portfolio selection. The Journal of Finance 7 (1), 77 - 91.

Nataraja, N. R., Johnson, A. L., 2011. Guidelines for using variable selection techniques in data envelopment analysis. European Journal of Operational Research 215 (3), 662 - 669.

NSE, 2015. National Stock Exchange. www.nseindia.com, accessed on Feb 28, 2015.

Pastor, J. T., Ruiz, J. L., Sirvent, I., 2002. A statistical test for nested radial DEA models. Operations Research 50 (4), 728 - 735.

Pätäri, E., Leivo, T., Honkapuro, S., 2012. Enhancement of equity portfolio performance using data envelopment analysis. European Journal of Operational Research 220 (3), 786 - 797.

Pätäri, E. J., Leivo, T. H., Honkapuro, J. S., 2010. Enhancement of value portfolio performance using data envelopment analysis. Studies in Economics and Finance 27 (3), 223 - 246.

Powers, J., McMullen, P. R., 2000. Using data envelopment analysis to select efficient large market cap securities. Journal of Business and Management 7, 31 - 42.

Ruggiero, J., 2005. Impact assessment of input omission on DEA. International Journal of Information Technology & Decision Making (IJITDM) 4 (3), 359 - 368.

Saen, R. F., Azadi, M., 2011. A chance constrained data envelopment analysis approach for strategy selection. Journal of Modelling in Management 6 (2), 200 - 214.

Sarkis, J., 2007. Preparing your data for DEA. In: Zhu, J., Cook, W. (Eds.), Modeling Data Irregularities and Structural Complexities in Data Envelopment Analysis. Springer US, pp. 305 - 320.

SEBI, 2014. Common queries of investors. http://www.sebi.gov.in/cms/sebidata/attachdocs/1315458767512.pdf, accessed on: Nov 7, 2014.

Singh, A. K., Sahu, R., Bharadwaj, S., 2010. Portfolio evaluation using OWA heuristic algorithm and data envelopment analysis. The Journal of Risk Finance 11 (1), 75 - 88.

Smith, P., 1990. Data envelopment analysis applied to financial statements. Omega 18 (2), 131 - 138.

Soleimani, H., Golmakani, H. R., Salimi, M. H., 2009. Markowitz-based portfolio selection with minimum transaction lots, cardinality constraints and regarding sector capitalization using Genetic Algorithm. Expert Systems with Applications 36 (3, Part 1), 5058 - 5063.

Tehrani, R., Mehragan, M. R., Golkani, M. R., 2012. A model for evaluating financial performance of companies by Data Envelopment Analysis - A case study of 36 corporations affiliated with a private organization. International Business Research 5 (8), 8 -16.

Tsolas, I., 2014. DEA performance assessment of Greek listed metallurgical firms. Journal of Modelling in Management 9 (1), 58 - 77.

Ueda, T., Hoshiai, Y., 1997. Application of principal component analysis for parsimonious summarization of DEA inputs and/or outputs. Journal of the Operations Research Society of Japan 40 (4), 466 - 478.

Yap, G. L. C., Ismail, W. R., Isa, Z., 2013. An alternative approach to reduce dimensionality in data envelopment analysis. Journal of Modern Applied Statistical Methods 12 (1), 128 - 147.

Biography

Dhanya Jothimani is a Ph.D. scholar in the Department of Management Studies at IIT Delhi, India. She received her Bachelors of Technology (B.Tech.) from National Institute of Technology, Tiruchirappalli, India, and Masters of Technology (M.Tech.) from IIT Kharagpur, India. Her research interests include Financial Analytics, Data Mining, and Operations Research. She is a member of Institute for Operations Research and Management Science (INFORMS), American Finance Association (AFA), Decision Science Institute (DSI) and Association of Information Systems (AIS).

Ravi Shankar is a Professor of Supply Chain and Operations Management at IIT, Delhi, India. His areas of interest are supply chain management, business analytics, operations research, fuzzy modeling, six-sigma, project management, strategic technology management, etc. He has published over 300 research papers in journals and conferences, including Omega, European Journal of Operations Research, International Journal of Production Research, International Journal of Production Economics, International Journal of Supply Chain Management, International Journal of Physical Distribution & Logistics Management, IEEE Systems Man and Cybernetics Part C, Computers and Industrial Engineering, Computers and Operations Research, etc.

Surendra S. Yadav received his Bachelor of Technology (B. Tech.) from IIT, Kanpur, India, MBA from University of Delhi, India, and PhD in Management from University of Paris 1 Pantheon-Sorbonne, France. He is a Professor in the Department of Management Studies, IIT, Delhi, India. He teaches Corporate Finance, International Finance, International Business and Security Analysis & Portfolio management. His research interests are in all these areas and general management. He is the Editor-in-Chief of Journal of Advances in Management Research, published by Emerald, UK.