Prediction and performance evaluation of BDI forecasting models

–Cross efficiency, the directional distance function and the AVS utility function –

Ming-Miin Yu
National Taiwan Ocean University
2 Pei-Ning Road Keelung Taiwan 202-24, R.O.C
yumm@mail.ntou.edu.tw

Emrah Bulut
Yildiz Technical University
34349 Beşiktaş - İstanbul, Turkey
bltemrah@gmail.com

Abstract: In the study, we propose a nonparametric efficiency measurement approach for the forecasting model selection problem. Three autoregressive models and three fuzzy time series approaches are employed for the calibration of data structure to depict the trend. The directional distance function and portfolio theory are further used to evaluate the performance of BDI predictions. A directional distance function is defined that looks for possible increases in accuracy and skewness, and decreases in variance obtained by cross efficiencies of those forecasting models. We also establish a link to proper indirect accuracy- variance -skewness (AVS) utility function for various users in various utilities. An empirical section on a set of forecasting Baltic Dry Index (BDI) forecasting models serves as an illustration.

Keyword: Baltic dry index; portfolio theory; cross efficiency; the directional distance function.

1. INTRODUCTION

An international shipping freight index, Baltic Dry Index (BDI), which is expressed a freight rate for the price of the maritime transportation (such as iron ore, coal and grain by Handymax, Panamax and Capsize) and it is unique indicator of the state of the world economy. It would be highly correlated with the pricing path of all these commodities and currencies that are associated with dry cargo transportation. The BDI is an important response of maritime information for the trading and settlement of physical and derivative contracts. Since the forecasting model for the BDI influence the accuracy of the results, it plays a critical role for the shipping firms in their economic decision making and defining investment strategy. The aim of this study is to investigate the forecasting performance to select an appropriate BDI forecasting model in the shipping business by considering lower variance, probability of obtaining a large inaccuracy and higher accuracy in dimensions of goodness-of-fit, biasedness and correct sign.

Though most forecasting studies adopt various criteria to evaluate the accuracy of competing forecasting models, the assessment of these forecasting models lead to many different rankings due to a specific measure of a specific criterion (e.g., Sadorsky, 2005; 2006; Coppola, 2008; Agnolucci, 2009; Murat and Tokat, 2009; Marzo and Zagaglia, 2010). Taking all criteria into account is essential for the purposes of better performance evaluation of which model performs best overall (Ouenniche et al, 2015).

In much of the forecasting studies literature, three performance criteria have typically been used: (1) goodness-of-fit (how close the forecasts are from the actual values), (2) biasedness (whether the model tends to systematically over-estimate or under-estimate the forecasts) and (3) correct sign (the ability of a model to produce forecasts that are consistent with actuals in that forecasts reveal increase (resp. decrease) in value when actuals increase (resp. decrease) in value) (Ouenniche et al, 2015). However, one of the major problems of existing forecasting studies dealing with this three performance criteria is that their ranking lacks the consideration of the three measures simultaneously. Ouenniche et al, (2015) proposed an orientation-free super-efficiency Data Envelopment Analysis (DEA) framework to obtain a single ranking for the performance evaluation of volatility forecasting models by taking account of several criteria to fill in this gap.

To empirically implement this evaluation, Ouenniche et al, (2015) have used the slacks-based super-efficiency DEA framework for assessing the relative performance of competing volatility forecasting models in which measures of biasedness and goodness-of-fit are used as input, whereas measures of correct sign are used as output. Their approach may suffer from two problems. First, the performance of forecasting models well-diversified in terms of their performance on multiple evaluation criteria. Second, the combination of each of these models for getting better performance is invalid. We, therefore, advocate the use of cross efficiency DEA models to circumvent the first problem and portfolio theory with directional distance function to providing better prediction values with the consideration of user's preference.

Note that regardless of how one forecasts the variance and skewness of accuracy of BDI, one could not assess the relative performance of competing forecasting models and finds out which ones or the combinations have the potential of doing a good "prediction job". In order to overcome this methodological issue, our objective is therefore to address these issues by developing an accuracy-variance-skewness (AVS) framework of combination selection based on DEA cross-efficiency evaluation. The basic idea is that a forecasting model's simple efficiency score, variance and skewness of its cross-efficiencies are used to represent the prediction's accuracy, variance and low probability of obtaining a large inaccuracy characteristics. Subsequently, the AVS combination frontier approach is used to determine the forecasting model's inclusion in a portfolio-based combination in this research, which overcomes the following issues. First, the directional distance function is used to represent the accuracy, variance and skewness space and as an efficiency measure for forecasting models. The directional distance function is to account for a preference for both positive accuracy and skewness associated with an aversion to variance. Second, this directional distance function projects a combination of forecasting models for which improvements can be found, in terms of increasing accuracy and skew, decreasing variance, onto the efficient forecasting frontier and labels these inefficiency. Third, we also assess the degree of satisfaction of users' preferences in which a dual approach specifying an AVS utility function to choose among these frontier combinations. For given variance aversion and prudence parameters, we can obtain an optimal point on the boundary of the nonconvex AVS combination frontier. In sum, the proposed prediction and performance evaluation approach judges simultaneously accuracy and skewness expansions and variance contractions; namely, a portfolio-based DEA framework with cross efficiency for assessing the relative performance of competing BDI forecasting models.

The remaining structure of this article is organized as follows: Section 2 literature review. Section 3 presents cross efficiency, the directional distance function and portfolio theory. The prediction and performance evaluation of forecasting in the manner that it can incorporate accuracy, variance and skewness. Section 4 applies the proposed framework to the prediction and performance evaluation of BDI forecasting. Conclusion and future extensions are summarized in the last section.

2. LITERATURE REVIEW

2.1. Forecasting model performance evaluation

In the last decades, many studies have paid attention to reveal the appropriate forecasting methods and error measures (Armstrong & Collopy, 1992; Hibon, Meade, Makridakis, & Fildes, 1998; Hyndman & Koehler, 2006; Makridakis & Hibon, 2000; Makridakis et al., 1993). They investigated to define an appropriate method by considering features of data, proper accuracy metrics, in-sample and post-sample results, type of unit root test and length of data series. In the Makridakis and Hibbon's study (2000), more than twenty different forecasting methods and expert systems including artificial neural networks are applied to measure their accuracy for both in-sample and post-sample. They found that the results of the accuracy control in terms of post-sample for the simple forecasting methods is superior than more complex algorithm. According to the Makridakis and Hibbon, some principles can be emphasized as follows:

- Statistically sophisticated or complex methods do not necessarily provide more accurate forecasts than simpler ones.
- The relative ranking of the performance of the various methods varies according to the accuracy measure being used.
- The accuracy of the various methods depends upon the length of the forecasting horizon involved.
- The characteristics of the data series are an important factor in determining relative performance between methods.

In the existing literature, some studies also use data envelopment analysis (DEA) method to measure the performance of forecasting models by considering a specific level of a specific criterion (B Xu & Ouenniche, 2011; Bing Xu & Ouenniche, 2012a, 2012b), in which a single ranking that takes account of several criteria. However, how to find out about the multidimensional rankings with respect to different measures with user's preferences is still an issue need to be overcome.

2.2. BDI forecasting

For the shipping industry and shipping firms, forecasting plays a critical role for the investment timing, market entry-exist decisions, freight discovery and many aspects of the shipping prediction required (Bulut, Duru, & Yoshida, 2013). The shipping business research also cannot ignore the importance of the forecasting performance to reduce cost and making successful investment decisions (Bulut, Duru, & Yoshida, 2012; Bulut, 2014). Batchelor et al. (2007)

In the existing literature, there are many papers propose the forecasting model to estimate the freight rates and ship prices. Veenstra & Franses (1997) apply the Vector Autoregression (VAR) model for the forecasting of the ocean dry bulk freight rates. Tsolakis et al. (2003) use Error Correction method for the supply and demand of ship prices forecasting. Duru et al. (2012) develop a DELPHI method by using fuzzy algorithm to improve the forecasting accuracy in the dry bulk shipping index. Bulut (2014) embeds the classical VAR logics in the fuzzy time series algorithm to improve forecasting accuracy. Although many papers propose different methods to improve the forecasting accuracy, few studies investigate the performance of the forecasting methods in the shipping industry. Batchelor et al. (2007) test the performance of popular time series model in predicting spot and forward rates on major seaborne freight rates. In their study, they use just some accuracy control methods to compare results of the traditional forecasting methods such as ARIMA, VAR and VECM.

3. METHODOLOGY

We will be discussing first the DEA AVS cross-efficiency model, and then introduce the directional distance function in an attempt to analyze frontier of efficient prediction combinations using portfolio theory.

3.1. DEA AVS cross-efficiency model

Suppose there are N DMUs in a reference set, each DMU j, $j = 1,\dots,N$ has M inputs, denoted as $X_j = (x_{1j}, x_{2j}, \cdots, x_{Mj})^T$, S outputs, denoted as $Y_j = (y_{1j}, y_{2j}, \cdots, y_{Sj})^T$, where T in the super script indicates transpose. Without loss of generality, let us denote the DMU under evaluation as DMU n. The original DEA model of proposed by Charnes, Cooper, and Rhodes (CCR) (1978) can be given in (1)..

$$E_n^n = \max \sum_{s=1}^{S} u_s y_{sn} \tag{1}$$

$$\sum_{m=1}^{M} v_m x_{mj} - \sum_{s=1}^{S} u_s y_{sj} \geq 1,\quad j = 1,\dots,N \tag{1-1}$$

$$\sum_{m=1}^{M} v_m x_{mk} = 1 \tag{1-2}$$

$$u_r, v_m \geq \varepsilon,\ s = 1,\dots,S;\ m = 1,\dots,M$$

where ε is a positive non-Archimedean infinitesimal, and u_s, v_m are multipliers on outputs and inputs, respectively, to be determined by optimizing the model. By solving model (1), we can obtain the optimal solution

u_s^n, v_m^n for DMU j ($j=1,\dots,N$) . Then, the cross efficiency of DMU j using DMU n's multipliers can be calculated as in (2).

$$E_j^n = \frac{\sum_{r=1}^{R} u_r^n y_{rj}}{\sum_{m=1}^{M} v_m^n x_{ij}}, \quad n=1,\dots,N, \quad j=1,\dots,N \qquad (2)$$

For a DMU j, the return characteristic is defined as its average of cross-efficiency scores $E_j^n (n=1,\dots,N)$ which implies that the return of DMU j is determined by averaging the efficiencies of DMU j using various multipliers.

$$R_j = \frac{\sum_{n=1}^{N} E_j^n}{N}, j \in \{1,\cdots,N\} \qquad (3)$$

In this regard, each DMU is exposed to the risk of change in multipliers. The risk characteristic can then be defined as variance of its cross-efficiency scores $E_j^n (n=1,\dots,N)$ as

$$\sigma_j^2 = \frac{\sum_{n=1}^{N} (E_j^n - R_j)^2}{N}, \quad j \in \{1,\cdots,N\} \qquad (4)$$

As claimed by Lim et al. (2014) "the simple use of cross-efficiency evaluation in portfolio selection effectively considers the risk of change in weights for individual DMUs selected in a portfolio, it fails to consider the risk for the portfolio overall." In their model, Markowitz' mean-variance formulation is used to determine the DMU's inclusion in a portfolio under consideration. However, in the literature, many scholars regard that portfolio returns are generally not normally distributed. Furthermore, returns are said to be non-normally distributed and this is generally attributed to skewness. The mean-variance (MV) framework of portfolio selection effectively increases return and reduces risk but fails to contribute to a better understanding of risk preferences via its estimation of risk aversion and prudence in the long run. There is a growing literature dealing with portfolio models that account for high moment pricing effects.

To address this issue, we extend the model proposed by Lim et al. (2014) and develop a AVS framework of prediction model combination selection based on cross-efficiency evaluation. The basic idea is that a DMU's expected efficiency score, the co-variance and co-skewness of its cross-efficiencies are used to represent the DMU's return, risk and prudence characteristics. Alternatives are characterized by an expected efficiency $E\left[R_j\right]$ for $j \in \{1,\cdots,N\}$, where the co-variance matrix and co-skewness of its cross-efficiencies are as (5) and (6), respectively.

$$\Omega_{i,j} = COV\left[R_i, R_j\right], \ i,j \in \{1,\cdots,N\} \qquad (5)$$

$$CKS_{i,j,k} = E\begin{bmatrix} \left(R_i - E\left[R_i\right]\right)\left(R_j - E\left[R_j\right]\right) \\ \left(R_k - E\left[R_k\right]\right) \end{bmatrix}$$

$$i,,j,k \in \{1,\cdots,N\} \qquad (6)$$

A combination $\alpha = (\alpha_1,\dots,\alpha_N)$ _is composed by a proportion of each of these *n* alternatives ($\sum_{i=1}^{N} \alpha_i$). In general, the set of admissible combination s can be written as follows (re-write).

$$\psi = \left\{ \alpha \in R^N : \sum_{i=1}^{N} \alpha_i = 1, \alpha \ge 0 \right\} \qquad (7)$$

When a portfolio α is included, we have the return of this portfolio as:

$$\mathrm{R}(\alpha) = \sum_{i=1}^{N} \alpha_i R_i \qquad (8)$$

For a given combination, the expected accurancy, its variance, and its skewness are as follows:

$$\mathrm{E}\left[\mathrm{R}(\alpha)\right] = \sum_{i=1}^{N} \alpha_i E\left[R_i\right] = \mu(\alpha) \qquad (9)$$

$$\mathrm{Var}\left[\mathrm{R}(\alpha)\right]=E\left[(\mathrm{R}(\alpha)\text{-}\mu(\alpha))^2\right]$$
$$=\sum_{i=1}^{N}\sum_{j=1}^{N}\alpha_i\alpha_j Cov\left[R_i,R_j\right]=\sum_{i=1}^{N}\sum_{j=1}^{N}\alpha_i\alpha_j\Omega_{i,j} \tag{10}$$

$$\mathrm{Sk}\left[\mathrm{R}(\alpha)\right]=E\left[(\mathrm{R}(\alpha)\text{-}\mu(\alpha))^3\right]$$
$$=\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{k=1}^{N}\alpha_i\alpha_j\alpha_k E\left[(\mathrm{R}_i\text{-}\mu(\alpha))(\mathrm{R}_j\text{-}\mu(\alpha))(\mathrm{R}_k\text{-}\mu(\alpha))\right]$$
$$=\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{k=1}^{N}\alpha_i\alpha_j\alpha_k CKS_{i,j,k} \tag{11}$$

We denote conbination by $\alpha=(\alpha_1,\ldots,\alpha_N)\in R^N$, the function Φ is defined by $\Phi(\alpha)=\left(\mathrm{E}\left[\mathrm{R}(\alpha)\right],\mathrm{Var}\left[\mathrm{R}(\alpha)\right],Sk\left[\mathrm{R}(\alpha)\right]\right)$ to represent its expected accurancy, variance, and skewness.

To measure combination efficiency, we give a disposal representation set DR of combinations below:

$$DR=\left\{\begin{matrix}(E,V,S)\in R^3;\exists\alpha\in\psi,(E,V,S)\le\\ (E\left[R(\alpha)\right],-V\left[R(\alpha)\right],S\left[R(\alpha)\right])\end{matrix}\right\}. \tag{12}$$

to define a subset of this representation set known as the efficient frontier:

Given a directional vector $g=\left(g_E,-g_V,g_S\right)\in R^3$, the directional distance function, $S_g\left(\alpha\right)$, defined on the disposal representation set DR is

$$S_g\left(\alpha\right)=\sup\left\{\delta:\Phi\left(\alpha\right)+\delta g\in CR\right\}$$

This directional distance function represents a prediction model combination efficiency indicator.

3.2. The directional distance function and the AVS utility function

In production theory, the directional distance function measures are defined as the technical efficiencies of some point of the production possibility set and the Pareto frontier. Assume a sample of q combinations, $\alpha^1,\alpha^2,\cdots,\alpha^q$, we calculate the directional distance function for a specific combination p^k for $\alpha^1,\alpha^2,\cdots,\alpha^q$ under evaluation, $S_g(p^k)$, as the following cubic program:

$$\max\ S_g(p^k)=\delta \tag{13}$$

s.t.

$$E\left[R(\alpha)\right]\ge E\left[R(p^k)\right]+\delta g_{\mathrm{M}}, \tag{13-1}$$

$$Var\left[R(\alpha)\right]\le Var\left[R(p^k)\right]-\delta g_{\mathrm{V}} \tag{13-2}$$

$$Sk\left[R(\alpha)\right]\ge Sk\left[R(p^k)\right]+\delta g_S \tag{13-3}$$

$$\sum_{j=1}^{N}\alpha_j=1,\ \alpha_j\ge 0,\ j=1,\ldots,N \tag{13-4}$$

Given the expected accurancy, its variance, and its skewness as equations (9)–(11), model (13) can be rewritten as follows:

$$\max\ S_g(p^k)=\delta \tag{14}$$

s.t.

$$\sum_{j=1}^{N}\alpha_j E\left[R_j\right]\ge E\left[R(p^k)\right]+\delta g_{\mathrm{M}}, \tag{14-1}$$

$$\sum_{i=1}^{N}\sum_{j=1}^{N}\alpha_i\alpha_j\Omega_{i,j}\le Var\left[R(p^k)\right]-\delta g_{\mathrm{V}} \tag{14-2}$$

$$\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{k=1}^{N}\alpha_i\alpha_j\alpha_k CKS_{i,j,k}\ge Sk\left[R(p^k)\right]+\delta g_S \tag{14-4}$$

$$\sum_{j=1}^{N}\alpha_j=1,\ \alpha_j\ge 0,\ j=1,\ldots,N \tag{14-5}$$

Suppose that there is a sample of q combinations of forecasting model requires calculating model (14) for each of these q combinations in turn. These function measures the maximal feasible reduction in δ. This function seeks the simultaneous maximum reduction in its risk and expansion in its accuracy and skewness in the direction of vector g. If $\delta=0$, we say that the evaluated combination is efficient in the g direction. Feasible but inefficient firms will take on values greater than zero, reflecting the additional accuracy and skewness and reduction in risk that the particular combination could achieve if it were on the best practice frontier. In words, there exists a combination of other forecasting models that yields a higher accuracy and skewness and a lower risk; the evaluated combination is situated below the boundary, thus inefficient. The directional distance function is also a measure of combination efficiency (PE) is defined as the quantity $PE = \delta$ which not necessarily project a point on the nonconvex portfolio frontier

3.3. Efficient combination of predictions with users' preferences

Although disposal representation set representations of the technology are conceptually useful, a point obtained by Model (14) is not necessarily on the frontier maximizing the users' indirect AVS utility function. [1] If one were to have a convex representation set, imposing tangent iso-utility surfaces compatible with the set of admissible AVS combinations is required. Briec et al. (2007) defined an MVS utility function as a third-order polynomial approximation of expected utility which is relative to a convexified MVS portfolio frontier. They defined the convex representation set, CR, as follows:

$$CR = \left\{ \begin{matrix} (E, V, S) \in R^3, \forall \mu, \rho, \kappa \in R_+^3, \\ U^*(\mu, \rho, \kappa) \geq \mu E - \rho V + \kappa S \end{matrix} \right\} \quad (15)$$

Let the AVS utility function for a specific combination α be

$$\begin{aligned} &U_{(\mu,\rho,\kappa)}(\alpha) \\ &= \mu E[R(\alpha)] - \rho Var[R(\alpha)] + \kappa Sk[R(\alpha)] \end{aligned} \quad (16)$$

Given the utility function satisfies positive marginal utility of expected accuracy, skewness and negative marginal utility of risk, we have the indirect AVS utility function is as

$$U^*(\mu, \rho, \kappa) = \max \left\{ U_{(\mu,\rho,\kappa)}(\alpha), \sum_{j=1}^{N} \alpha_j = 1, \alpha_j \geq 0 \right\} \quad (17)$$

To show how the optimal combination α^* can be estimated, the nonlinear optimization program can be written as follows:

$$Max\ E[R(\alpha)] - \varphi Var[R(\alpha)] + \psi Sk[R(\alpha)] \quad (18)$$

s.t.

$$\sum_{j=1}^{N} \alpha_j = 1, \alpha_j \geq 0, \quad (18\text{-}1)$$

where $\varphi = \frac{\rho}{\mu}$ and $\psi = \frac{\kappa}{\mu}$, represent the degree of absolute risk aversion and the degree of absolute prudence, respectively. Given a set of parameters representing decision maker's absolute risk aversion and absolute prudence, the maximum value function may be estimated.

The specific combination inefficiency evaluated can be obtained by model (14). By maximizing the decision maker's direct AVS utility function (17), a unique efficient combination among those on the weakly efficient frontier may be estimated.

By alternatively choosing the convexified AVS combination frontier, the hyper-portfolio efficiency (HPE), $\hat{S}_g(\alpha)$, defined on CR can also be derived

$$\hat{S}_g(\alpha) = \sup\left\{ \hat{\delta} : \Phi(\alpha) + \hat{\delta} g \in CR \right\}$$

where the HPE guarantees reaching a point on the frontier maximizing the investor's indirect AVS utility function

The DDF introduced into portfolio analysis by Briec et al. (2007 is useful to distinguish between overall, allocative, convexity, and portfolio inefficiency when evaluating the scope for improvements in combination selection. Given the definition of overall inefficiency (OE)

$$\begin{aligned} &OE(\alpha; \mu, \rho, \kappa) = \\ &\sup\left\{ \delta : (\mu, -\rho, \kappa)(\Phi(\alpha) + \delta g) \leq U^*(\mu, \rho, \kappa) \right\} \end{aligned}$$

The relationships between overall, allocative, convexity, and combination inefficiency may be established as follows:

[1] Briec et al. (2015) show how the directional distance function is linked to the dual approach based on the specification of an MVS utility function.

The allocative inefficiency (AE) index

$$AE(\alpha;\mu,\rho,\kappa)=\mathrm{O}E(\alpha;\mu,\rho,\kappa)-\hat{S}_g(\alpha)$$

The convexity inefficiency (CE) index

$$CE(\alpha)=\hat{S}_g(\alpha)-S_g(\alpha)$$

That is

$$\mathrm{O}E=AE+CE+PE$$

3.4. Accuracy control

In this paper, the four different accuracy methods, mean absolute scaled error (MASE), median relative absolute error (MdRAE), symmetrical mean absolute percentage error (sMAPE), normalized root mean squared error (nRMSE) are applied to control the prediction results of forecasting methods, and act as inputs in the cross efficiency DEA model. The mathematical algorithm of accuracy control methods are defined as follows;

$$MASE = mean|q_t| \qquad (19)$$

$$q_t = \frac{e_t}{\frac{1}{n}\sum_{i=2}^{n}|A_i - A_{i-1}|} \qquad (19\text{-}1)$$

$$e_t = A_t - F_t \qquad (19\text{-}2)$$

where A_t is raw data series and F_t represents forecast value, subscript t represents those out-sample data, while i represents those in-sample data for estimating forecasting models.

Let $r_t = \frac{e_t}{e_t^*}$ denote the relatıve error, where e_t^* is the forecast error obtained from the benchmark method (Naive).

$$MdRAE = median|r_t| \qquad (20)$$

$$sMAPE = \frac{1}{n}\sum_{i=1}^{n}\frac{|F_i - A_i|}{\frac{(|A_i|+|F_i|)}{2}} \qquad (21)$$

$$nRMSE = \frac{\sqrt{\frac{\sum_{i=1}^{n}(F_i - A_i)^2}{n}}}{A_{\max} - A_{\min}} \qquad (22)$$

where $A_{\max} = \mathop{Max}_{i=1}^{n} A_i$ and $A_{\min} = \mathop{Min}_{i=1}^{n} A_i$

Percentage of correct direction change predictions (PCDCP) acts as output in the cross efficiency DEA model.

$$PCDCP = \sum_{t=1}^{n} z_t \Big/ n \qquad (23)$$

where z_t is a binary variable set equal to 1 if $(A_t - A_{t-1})\cdot(F_t - A_{t-1}) \succ 0$ and 0 otherwise.

4. RESULTS

4.1. BDI forecasting

In this paper, the BDI monthly data set between 01M1985 and 12M2011 is used to calculate the results of the forecasts methods and it is divided into sample period (01M1985-04M2007) and post-sample period (05M2007-12M2011). Augmented Dickey-Fuller (ADF) and Philips-Perron (PP) test are utilized for the stationary control and Table 1 clearly display that the data set is found non-stationary in 99% and %95 confidence level for PP test while it is stationary in 99% confidence level for ADF. Therefore, a differencing operation for the BDI data set is used for the transformation to obtain stationary data. Four fundamental methods of conventional time series analysis (autoregressive-AR, autoregressive integrated moving average-ARIMA, seasonal ARIMA-SARIMA and Holt-Winter`s Exponential Smoothing-HW) are applied for the accuracy measures. Among autoregressive models (AR, ARIMA-ARMA, SARIMA), significance and goodness of fit are investigated to define proper version of AR configuration. Since seasonal fluctuation exists in BDI time series, SARIMA model is found appropriate one.

Since the different fuzzy time series (FTS) approaches are widely proposed to improve the forecasting accuracy in the literature, three different FTS models (Chen`s algorithm, FILF and EFILF methods) are applied. Chen`s algorithm is utilized both raw and first different data while others are only applied stationary BDI data set.

In the empirical study, each forecasting method mentioned above is utilized to compute their accuracy measures.

Table 1: Unit root test of BDI (01M1985 – 12M2011).

	ADF Levels	PP Levels	Diffrence
t-statistics	-4.288	-3.287	-11.007
(*p**)	(*0.003*)	(*0.070*)	(*0.000*)*
1% level	-3.986*	-3.986	
5% level	-3.423	-3.423	
10% level	-3.134	-3.134	
**p*-value.			

* Stationary in %1 critical value (99% confidence level).

4.2. Forecasting model performances.

Following Hyndman and Koehle (2006), accuracy measures of scale-independent are used as input, whereas measures of correct sign are used as output. This results in a total of 5 measures for each forecasting model, as presented in Table 2.

Their superiority is based on accuracy control method. For instance, FILF method is found superior according to the results of the MASE while SARIMA has the most accurate result for MdRAE and SMAPE.

Table 2 Inputs and output of BDIM time series forecasting models

Model		Inputs				Output
		MASE	**MdRAE**	**SMAPE**	**NRMSE**	**PCDCP**
SARIMA	A1	1.040	0.850	0.090	0.210	0.583
H-W	A2	3.800	4.720	0.290	0.640	0.056
cFTS (raw)	A3	0.941	1.014	0.092	0.212	0.333
cFTS (diff)	A4	1.187	1.349	0.118	0.288	0.333
FILF	A5	0.938	1.013	0.091	0.212	0.667
EFILF	A6	0.952	1.030	0.093	0.209	0.806

According to the results of accuracy control methods, it cannot safely say that one method is found superior for all accuracy control methods. Therefore, a multidimensional framework for the performance evaluation of competing models of BDI forecasting is proposed in this study. Prior to formal modeling, we first present descriptive statistics on four measures of the scale-independent and one measure of correct sign, of six forecasting models of BDI. Table 2 exhibits the values of all measures of which serves to provide some light on the plausibility of the derivative DEA AVS cross-efficiency scoreA set of four inputs of scale-independent goodness-of-fit and one output of correct sign derived from prediction results of listed forecasting models, and compute a set of cross efficiencies based on the selected set of inputs and outputs. Table 3 shows the cross-efficiency scores obtained by using the MASE, MdRAE, SMAPE and NRMSE as inputs, PCDCP as outputs. Notice that BDI forecasting models ranked from best to worst using the CCR and CEM efficiency scores are different.

Table 3 Cross efficiency-based multidimensional rankings of BDIM time series forecasting models

CEM	A1	A2	A3	A4	A5	A6	CEM-mean	Ranking-CEM	Ranking-CCR
A1	0.876	0.720	0.876	0.662	0.747	0.747	0.771	3	2
A2	0.015	0.023	0.015	0.017	0.022	0.022	0.019	6	6
A3	0.420	0.407	0.420	0.418	0.418	0.418	0.417	4	4
A4	0.315	0.300	0.315	0.331	0.326	0.326	0.319	5	5
A5	0.841	0.816	0.841	0.840	0.846	0.846	0.838	2	3
A6	1	1	1	1	1	1	1.000	1	1

In order to take portfolio theory into consideration for evaluating the performance of BDI predictions, one would need rankings that takes account of accuracy-variance-skewness (AVS) utility function for various users in various utilities, which we provide using the proposed DEA framework. Some notes should be made from the overall, allocative, convexity, and combination inefficiency estimates in Table 4, the overall inefficiency estimate of 0.248 suggests that A1 forecasting model wastes around 24.8% of its utilities relative to the best-practice forecasting model. The decomposition indicates that 9.5% of this poor performance is due to PE, while the remaining 1.53% of the gap is due to AE* (includes CE). In other words, the PE level of 0.095 suggests that model just lost 9.5% combination inefficiency that the best-practice BDI forecasting model could make under the nonconvex combination frontier. The AE* level of 0.0153 implies the model choosing a wrong mix of accuracy, skewness, and risk given postulated risk aversion and prudence parameters. These results also show that PE levels of the tested models are well below those of HPE, and CE levels are positive. As can be seen in Table 5, our approach also provides the optimal weights for selecting a better combination of existing forecasting models to obtain a more overall-efficient forecasting result. Take model A1 as example, the users can select a combination of forecasting models A1 and A6 associated with weights of 0.681 and 0.319, respectively, to reach combination efficiency.

Table 4 Mean-Variance-Skewness Portfolio Performance-based BDIM time series forecasting models

	OE	PE	AE	HPE	CE	Ranking-HPE
A1	0.248	0.095	-0.338	0.586	0.491	5
A2	47.383	0.505	46.654	0.729	0.224	6
A3	1.204	0.131	0.724	0.480	0.349	2
A4	1.871	0.018	1.374	0.497	0.479	4
A5	0.149	0.053	-0.332	0.481	0.428	3
A6	-0.006	0.000	-0.484	0.478	0.478	1

Table 5 Combination of the BDIM time series forecasting models

	A1	A2	A3	A4	A5	A6
A1	0.681					0.319
A2		0.978		0.017	0.003	0.002
A3		0.462			0.468	0.070
A4		0.637			0.315	0.048
A5					0.726	0.274
A6						1.000

5. CONCLUSION AND FUTURE EXTENSIONS

The results above demonstrate that the proposed approach can be a promising tool for forecasting model combination selection as a means of fundamental analysis. Our results also show that the cross-efficiency with AVS approach is more effective than the one based on the simple use of CCR or cross-efficiency scores at least for this particular application. Overall, our findings consistently support the effectiveness of our approach.

REFERENCES

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

Cooper W.W., Seiford L.M. and Tone K, (2000). Data envelopment analysis: A comprehensive text with models, applications, references and DEA-solver software, Kluwer Academic PublBoston.

Xu, B., & Ouenniche, J. (2011). A Multidimensional Framework for Performance Evaluation of Forecasting Models: Context-Dependent DEA. Applied Financial Economics, 21, 1873-1890.

Xu, B., & Ouenniche, J. (2012a). Performance Evaluation of Competing Forecasting Models – A Multidimensional Framework based on Multi-Criteria Decision Analysis. Expert Systems with Applications, 39, 8312-8324.

Xu, B., & Ouenniche, J. (2012b). A Data Envelopment Analysis-based Framework for The Relative Performance Evaluation of Competing Crude Oil Prices' Volatility Forecasting Models. Energy Economics, 34, 576-583.

Briec, W., Kerstens, K. and Jokung, O. (2007).

Mean-Variance-Skewness Portfolio Performance Gauging: A General Shortage Function and Dual Approach. Management Science 53(1):135-149.

Luenberger, D. G. (1995). Microeconomic Theory. McGraw-Hill, New York.

Armstrong, J. S., & Collopy, F. (1992). Error measures for generalizing about forecasting methods: Empirical comparisons. International Journal of Forecasting *8*(1), 69–80.

Batchelor, R., Alizadeh, A., & Visvikis, I. (2007). Forecasting spot and forward prices in the international freight market. International Journal of Forecasting *23*(1), 101–114.

Bulut, E. (2014). Modeling seasonality using the fuzzy integrated logical forecasting (FILF) approach. Expert Systems with Applications, *41*(4) 1806–1812.

Bulut, E., Duru, O., & Yoshida, S. (2012). A fuzzy integrated logical forecasting (FILF) model of time charter rates in dry bulk shipping: A vector autoregressive design of fuzzy time series with fuzzy c-means clustering. Maritime Economics and Logistics *14*(3), 300–318.

Bulut, E., Duru, O., & Yoshida, S. (2013). Market entry , asset returns , and irrational exuberance : asset management anomalies in dry cargo shipping. International Journal of Shipping and Transport Logistics *5*(6), 652–667.

Duru, O., Bulut, E., & Yoshida, S. (2012). A fuzzy extended DELPHI method for adjustment of statistical time series prediction: An empirical study on dry bulk freight market case. Expert Systems with Applications *39*(1), 840–848.

Hibon, M., Meade, N., Makridakis, S., & Fildes, R. a. (1998). Generalising about univariate forecasting methods: further empirical evidence. International Journal of Forecasting *14*, 359–366.

Hyndman, R. J., & Koehler, A. B. (2006). Another look at measures of forecast accuracy. International Journal of Forecasting *22*(4), 679–688.

Makridakis, S., Chatfield, C., Hibon, M., Lawrence, M., Mills, T., Ord, K., & Simmons, L. F. (1993). The M2-competition: A real-time judgmentally based forecasting study. International Journal of Forecasting *9*(1), 5–22.

Makridakis, S., & Hibon, M. (2000). The M3-Competition : results , conclusions and implications. International Journal of Forecasting 16, 451–476.

Tsolakis, S. D., Cridland, C., & Haralambides, H. E. (2003). Econometric Modelling of Second-hand Ship Prices. Maritime Economics and Logistics *5*(4), 347–377.

Veenstra, A. W., & Franses, P. H. (1997). A co-integration approach to forecasting freight rates in the dry bulk shipping sector. Transportation Research Part A: Policy and Practice 31(6), 447–458.

Xu, B., & Ouenniche, J. (2011). A multidimensional framework for performance evaluation of forecasting models: context-dependent DEA. Applied Financial Economics 21(24), 1873–1890.

Xu, B., & Ouenniche, J. (2012a). A data envelopment analysis-based framework for the relative performance evaluation of competing crude oil prices' volatility forecasting models. Energy Economics 34(2), 576–583.

Xu, B., & Ouenniche, J. (2012b). Performance evaluation of competing forecasting models: A multidimensional framework based on MCDA. Expert Systems with Applications *39*(9), 8312–8324.