

Article

Further Study of the DEA-Based Framework for Performance Evaluation of Competing Crude Oil Prices' Volatility Forecasting Models

Zhongbao Zhou ^{1,*}, Qianying Jin ¹, Jian Peng ¹ , Helu Xiao ² and Shijian Wu ³

¹ School of Business Administration, Hunan University, Changsha 410082, China

² School of Business, Hunan Normal University, Changsha 410081, China

³ College of Economics and Management, Shandong University of Science and Technology, Qingdao 266590, China

* Correspondence: z.b.zhou@hnu.edu.cn

Received: 22 July 2019; Accepted: 31 August 2019; Published: 6 September 2019



Abstract: The super-efficiency data envelopment analysis model is innovative in evaluating the performance of crude oil prices' volatility forecasting models. This multidimensional ranking, which takes account of multiple criteria, gives rise to a unified decision as to which model performs best. However, the rankings are unreliable because some efficiency scores are infeasible solutions in nature. What's more, the desirability of indexes is worth discussing so as to avoid incorrect rankings. Hence, herein we introduce four models, which address the issue of undesirable characteristics of indexes and infeasibility of the super efficiency models. The empirical results reveal that the new rankings are more robust and quite different from the existing results.

Keywords: forecasting models; super-efficiency data envelopment analysis; undesirable factors; infeasibility

1. Introduction

Among energy sources, crude oil remains the major part of energy consumption. Its price volatility is closely related to the stability of both the macro-economy and micro-economy [1]. The importance of crude oil prices in industry hence attracts great attention among academics to forecast the volatilities of crude oil prices.

Statistical forecasting models represent one type of the main crude oil price volatility forecasting models, refer to [2,3], etc. Based on the classic Generalized Autoregressive Conditional Heteroskedasticity (GARCH) process, Kang et al. (2014) found that Component GARCH (CGARCH) and Fractionally Integrated GARCH (FIGARCH) models are useful for modelling and forecasting persistence in the volatility of crude oil prices [4]. García-Martos et al. (2013) used multivariate models to improve prediction accuracy [5]. Tang et al. (2015) proposed a novel complementary ensemble empirical mode decomposition (CEEMD) based extended extreme learning machine (EELM) ensemble model that statistically outperforms all listed benchmarks in prediction accuracy [6]. Pany and Ghoshal (2015) introduced a local linear wavelet neural network model and showed its effectiveness for price forecasting [7]. Klein and Walther (2016) compared the recently introduced Mixture Memory GARCH (MMGARCH) with classical GARCH-type models [8]. Liu et al. (2018) investigated the impact of truncated jumps on improving the forecasting ability [9].

Although there is an enormous amount of literature available on predicting the volatility of crude oil prices, less attention is devoted to evaluating their forecasting performances. In addition, among the various emphases of forecasting models, there is no consistent conclusion as to which

forecasting model outperforms all others. Regarding this, Xu and Ouenniche (2012) put forward a multidimensional performance evaluation framework for evaluating the crude oil prices' volatility forecasting models [10]. To be more specific, a super-efficiency data envelopment analysis (DEA) model proposed by Andersen and Petersen (1993) is introduced [11]. The DEA method is widely applied in various areas (refer to [12–16], etc.). This multidimensional DEA-based framework contributes significantly to obtaining a unified ranking while taking all relevant criteria into account. However, several issues discredit the applicability of this super-efficiency DEA method. First, as indicated in [17,18], the super-efficiency model appears to be infeasible when the decision-making unit (DMU) under evaluation rests outside the production possibility set spanned by the other DMUs. In this sense, the evaluation results, which include these infeasibilities, are unreliable. Second, Xu and Ouenniche (2012) failed to reflect the essence of the evaluation criteria while taking them as inputs and outputs. The details are explained in the next section [10].

The contribution of this paper is three-fold. First, the desirability characteristics of the indexes for evaluation are discussed in detail. Second, two different evaluation processes are described, and the essence of criteria in these processes are discussed correspondingly. Third, the infeasibility problem arisen in applying a super-efficiency DEA model is tackled with an advanced model proposed by [19]. The structure of this contribution can be outlined as follows. In Section 2, a super-efficiency DEA-based framework is proposed for assessing the performance of the forecasting models for crude oil prices' volatility. The essence of the evaluation criteria, namely, the desirability, is discussed in detail. The models for dealing with infeasibilities are also introduced. In Section 3, empirical results and a comparison with those in [10] are presented. Finally, concluding remarks are summarized in Section 4.

2. Materials and Methods

2.1. Super Efficiency DEA Models with Undesirable Inputs (Outputs)

While considering the forecasting process of a specific model as a production process, although all dimensions of criteria are calculated to evaluate the forecasting performance, decision makers who give preference to the ability of forecasting the correct changing direction will focus on the correct sign prediction, which means that the forecast is consistent with actual values. Consequently, it is reasonable that any measures of correct sign prediction could be regarded as desirable outputs, and any measures of goodness-of-fit and biasedness can be regarded as desirable inputs. Here, the goodness-of-fit conveys how close the forecasts are from the real values and the biasedness indicates whether over-estimation or under-estimation exists. The super-efficiency DEA model, which can discriminate among those efficient decision-making units is introduced in this subsection.

Suppose we have a set of n DMUs. Let (x_{ij}, y_{rj}) denotes the i -th input and the r -th output of the j -th DMU. For the DMU_k under evaluation. The input-oriented super-efficiency DEA model can be expressed as

$$\begin{aligned}
 \min \quad & \theta_k \\
 \text{s.t.} \quad & \sum_{\substack{j=1 \\ j \neq k}}^n \lambda_j x_{ij} \leq \theta_k x_{ik}; \forall i \\
 & \sum_{\substack{j=1 \\ j \neq k}}^n \lambda_j y_{rj} \geq y_{rk}; \forall r \\
 & \lambda_j \geq 0; \forall j \neq k \\
 & \theta_k \text{ unrestricted}
 \end{aligned} \tag{1}$$

Similarly, the output-oriented super-efficiency DEA model can be expressed as

$$\begin{aligned}
 \max \quad & \phi_k \\
 \text{s.t.} \quad & \sum_{\substack{j=1 \\ j \neq k}}^n \lambda_j x_{ij} \leq x_{ik}; \forall i \\
 & \sum_{\substack{j=1 \\ j \neq k}}^n \lambda_j y_{rj} \geq \phi_k y_{rk}; \forall r \\
 & \lambda_j \geq 0; \forall j \neq k \\
 & \phi_k \text{ unrestricted}
 \end{aligned} \tag{2}$$

The efficiency of DMU_k is obtained by comparing with a virtual benchmark. λ_j indicates the proportional weight of DMU_j that consists of this virtual benchmark. The input-oriented efficiency score is the optimal value of (1), that is $\theta_k^{Super} = \theta_k^*$. The output-oriented efficiency score is the reciprocal of the optimal value of (2), that is $\phi_k^{Super} = \frac{1}{\phi_k^*}$.

In the following, we provide two approaches to evaluate the performance of competing crude oil prices' volatility forecasting models. One is to regard the biasedness and goodness-of-fit level as inputs and the correct sign as outputs. The other is treat all of them as outputs.

2.1.1. Approach I

The inputs in this application of competing crude oil prices' volatility forecasting models' performance evaluation seem to be undesirable. The undesirability of inputs means that the increasing of the biasedness and goodness-of-fit level will lead to a corresponding drop of the correct sign. In contrast, the closer the forecasts are from the actual values, as well as less bias, the higher the correct sign should be. As for special instances, if the levels of biasedness and goodness-of-fit equal to zero, which means that the forecasting results are exactly the real prices, then the correct sign prediction reaches the highest score, that is the unity. If inefficiency exists in this process, the inputs, namely goodness-of-fit and biasedness, should be increased to improve the performance, which reveals the undesirability feature of inputs. Consequently, any measures of correct sign prediction should be treated as desirable outputs, while any measures of goodness-of-fit and biasedness should be regarded as undesirable inputs.

There are lots of discussions on undesirable inputs and undesirable outputs, please refer to [20,21] for more details. Here we treat undesirable inputs as desirable outputs, and undesirable outputs as desirable inputs by following [22]. The efficiency can be obtained by calculating the following models in Table 1, where x_{ij}^U indicates the undesirable parts of inputs. The input-oriented efficiency can be defined as $\theta_k^{Super} = \frac{1}{\theta_k^*}$ and the output-oriented efficiency is given by $\phi_k^{Super} = \frac{1}{\phi_k^*}$.

In Model 1, for an inefficient DMU, θ_k^* indicates the required proportional increase in undesirable inputs so as to become efficient. In Model 2, for an inefficient DMU, ϕ_k^* indicates the possible required proportional increase in desirable outputs in order to become efficient.

Table 1. Super efficiency DEA models with undesirable inputs.

Input-Oriented (Model 1)		Output-Oriented (Model 2)	
<i>max</i>	θ_k	<i>max</i>	ϕ_k
<i>s.t.</i>	$\sum_{\substack{j=1 \\ j \neq k}}^n \lambda_j x_{ij}^U \geq \theta_k x_{ik}^U; \forall i$ $\sum_{\substack{j=1 \\ j \neq k}}^n \lambda_j y_{rj} \geq y_{rk}; \forall r$ $\sum_{\substack{j=1 \\ j \neq k}}^n \lambda_j = 1$ $\lambda_j \geq 0; \forall j \neq k$ $\theta_k \text{ unrestricted}$	<i>s.t.</i>	$\sum_{\substack{j=1 \\ j \neq k}}^n \lambda_j x_{ij}^U \geq x_{ik}^U; \forall i$ $\sum_{\substack{j=1 \\ j \neq k}}^n \lambda_j y_{rj} \geq \phi_k y_{rk}; \forall r$ $\sum_{\substack{j=1 \\ j \neq k}}^n \lambda_j = 1$ $\lambda_j \geq 0; \forall j \neq k$ $\phi_k \text{ unrestricted}$

2.1.2. Approach II

As aforementioned, if we treat the goodness-of-fit, the biasedness and the correct sign all as outputs, evaluating the performance corresponding to the practice. That is, considering the performance evaluation process as a process without explicit inputs, we treat measures of goodness-of-fit and biasedness as undesirable outputs, while those of correct sign desirable outputs. Specifically, a better performed crude oil prices' volatility forecasting model possesses lower levels of goodness-of-fit and biasedness, while higher level of correct sign at the same time.

To formulate the situation with pure output data, we consider the DEA without explicit inputs (WEI) model developed by [23]. The super efficiency DEA WEI models with undesirable outputs are presented in Table 2, where y_p^s and y_q^b represents the p th desirable (good) output and q th undesirable (bad) output, respectively. The WEI score of Model 3 which measures the desirable outputs can be defined as $\phi_{k1}^{Super} = \frac{1}{\phi_{k1}^*}$. The WEI score of Model 4 which measures the undesirable outputs is given by $\phi_{k2}^{Super} = \phi_{k2}^* \phi_{k2}^*$.

Table 2. Super efficiency DEA WEI models with undesirable outputs.

DEA-WEI (Model 3)		DEA-WEI (Model 4)	
<i>max</i>	ϕ_{k1}	<i>min</i>	ϕ_{k2}
<i>s.t.</i>	$\sum_{\substack{j=1 \\ j \neq k}}^n \lambda_j y_{pj}^s \geq \phi_{k1} y_{pk}^s; \forall p$ $\sum_{\substack{j=1 \\ j \neq k}}^n \lambda_j y_{qj}^b \leq y_{qk}^b; \forall q$ $\sum_{\substack{j=1 \\ j \neq k}}^n \lambda_j \leq 1$ $\lambda_j \geq 0; \forall j \neq k$ $\phi_{k1} \text{ unrestricted}$	<i>s.t.</i>	$\sum_{\substack{j=1 \\ j \neq k}}^n \lambda_j y_{pj}^s \geq y_{pk}^s; \forall p$ $\sum_{\substack{j=1 \\ j \neq k}}^n \lambda_j y_{qj}^b \leq \phi_{k2} y_{qk}^b; \forall q$ $\sum_{\substack{j=1 \\ j \neq k}}^n \lambda_j \leq 1$ $\lambda_j \geq 0; \forall j \neq k$ $\phi_{k2} \text{ unrestricted}$

In Model 3, for an inefficient DMU, the super efficiency score ϕ_{k1}^* shows the required proportional increase in desirable outputs to be efficient. In Model 4, for an inefficient DMU, the efficiency score ϕ_{k2}^* indicates the required proportional decrease in undesirable outputs to be efficient.

2.2. Dealing with Infeasibility and Zero Data

As indicated in [17,18], the input-oriented super-efficiency model may be infeasible when the outputs of the DMU under evaluation rest outside the production possibility set spanned by the other DMUs. And in the same spirit, the output-oriented super-efficiency model might be infeasible when the inputs of the DMU under evaluation lies outside the production possibility set spanned by the other DMUs. The current study finds that with undesirable inputs, infeasibility in the input-oriented model (Model 1) indicates super-efficiency can be regarded as output surplus while infeasibility in the output-oriented model (Model 2) indicates super-efficiency can be regarded as undesirable input surplus. As the approach in [19,24,25], when infeasibility occurs, corresponding models of Model 1, Model 2, Model 3, and Model 4 are presented as follows—see Tables 3 and 4, where M is a large positive number defined by a decision maker 10^5 .

Table 3. Modified version of super-efficiency DEA models with undesirable inputs.

Input-Oriented (Model 5)		Output-Oriented (Model 6)	
<i>max</i>	$\tau_k - M \sum_r \beta_r$	<i>max</i>	$\gamma_k - M \sum_i t_i$
<i>s.t.</i>	$\sum_{\substack{j=1 \\ j \neq k}}^n \lambda_j x_{ij}^l \geq (1 - \tau_k) x_{ik}^l; \forall i$	<i>s.t.</i>	$\sum_{\substack{j=1 \\ j \neq k}}^n \lambda_j x_{ij}^l \geq (1 - t_i) x_{ik}^l; \forall i$
	$\sum_{\substack{j=1 \\ j \neq k}}^n \lambda_j y_{rj} \geq (1 - \beta_r) y_{rk}; \forall r$		$\sum_{\substack{j=1 \\ j \neq k}}^n \lambda_j y_{rj} \geq (1 - \gamma_k) y_{rk}; \forall r$
	$\sum_{\substack{j=1 \\ j \neq k}}^n \lambda_j = 1$		$\sum_{\substack{j=1 \\ j \neq k}}^n \lambda_j = 1$
	$\lambda_j \geq 0; \forall j \neq k$		$\lambda_j \geq 0; \forall j \neq k$
	$\beta_r \geq 0; \forall r$		$t_i \geq 0; \forall i$
	τ_k unrestricted		γ_k unrestricted

Let $I = \{i | t_i^* > 0\}$ and $R = \{r | \beta_r^* > 0\}$, as denoted in [19]. We have the output surplus index calculated with β_r^* obtained from Model 5:

$$o = \begin{cases} 0 & \text{if } R = \emptyset \\ \frac{1}{|R|} \sum_{i \in R} \left(\frac{1}{1 - \beta_r^*} \right) & \text{if } R \neq \emptyset \end{cases}$$

and undesirable input surplus index calculated with t_i^* obtained from Model 6:

$$\hat{i}_u = \begin{cases} 0 & \text{if } I = \emptyset \\ \frac{1}{|I|} \sum_{i \in I} \left(\frac{1}{1 - t_i^*} \right) & \text{if } I \neq \emptyset \end{cases}$$

Then the modified input-oriented score from Model 5 can be defined as $\check{\theta}_k^{Super} = \frac{1}{1 - \tau^*} + o$. The modified output-oriented score $\check{\phi}_k^{Super}$ from Model 6 is given by $\check{\phi}_k^{Super} = \frac{1}{\gamma^*} + \hat{i}_u$.

Table 4. Modified version of super-efficiency DEA WEI model with undesirable outputs.

Output-Oriented (Model 7)		Output-Oriented (Model 8)	
<i>min</i>	$\gamma_{k1} + M \sum_q t_q$	<i>min</i>	$\gamma_{k2} + M \left(\sum_p \beta_p + \sum_q \delta_q \right)$
<i>s.t.</i>	$\sum_{j=1}^n \lambda_j y_{pj}^g \geq (1 - \gamma_{k1}) y_{pk}^g; \forall p$ $\sum_{j=1}^n \lambda_j y_{qj}^b - t_q y_q^{b_{\max}} \leq y_{qk}^b; \forall q$ $\sum_{j=1}^n \lambda_j \leq 1$ $\lambda_j \geq 0; \forall j \neq k$ $t_i \geq 0; \forall i$ $\gamma_{k1} \text{ unrestricted}$	<i>s.t.</i>	$\sum_{j=1}^n \lambda_j y_{pj}^g \geq (1 - \beta_p) y_{pk}^g; \forall p$ $\sum_{j=1}^n \lambda_j y_{qj}^b - \delta_q y_q^{b_{\max}} \leq (1 + \gamma_{k2}) y_{qk}^b; \forall q$ $\sum_{j=1}^n \lambda_j \leq 1$ $\lambda_j \geq 0; \forall j \neq k$ $t_i \geq 0; \forall i$ $\gamma_{k2} \text{ unrestricted}$

where $y_q^{b_{\max}} = \max_{k=1}^n \{y_{qk}^b\}$. This is proposed to deal with the infeasibilities raised by zero values.

Denote $Q = \{i | t_q^* > 0\}$ and $P = \{p | \beta_p^* > 0\}$, we have the undesirable output saving index

$$\hat{i} = \begin{cases} 0 & \text{if } Q = \emptyset \\ \frac{1}{|Q|} \sum_{q \in Q} (1 + t_q^*) & \text{if } Q \neq \emptyset \end{cases}$$

and the desirable output surplus index

$$\hat{o} = \begin{cases} 0 & \text{if } P = \emptyset \\ \frac{1}{|P|} \sum_{p \in P} \left(\frac{1}{1 - \beta_p^*} \right) & \text{if } P \neq \emptyset \end{cases}$$

The output-oriented score of Model 7 can be defined as $\check{\phi}_{k1}^{Super} = \frac{1}{1 - \gamma_{k1}^*} + \hat{i}$. The modified undesirable output-oriented score $\check{\phi}_{k2}^{Super}$ of Model 8 is given by $\check{\phi}_{k2}^{Super} = \gamma_{k2}^* + \hat{o} + \hat{i}$.

3. Results

This section is dedicated to the presentation of the empirical results. Fourteen competing crude oil prices' volatility forecasting models under evaluation and corresponding performance measures used are consistent with those in [10]. The final rankings obtained from applying Model 5, Model 6, Model 3, and Model 7 are presented in the following tables. The models are solved and programmed with Matlab.

Table 5 summarizes rankings of competing forecasting models obtained by Model 5. The result show that forecasting Model 3 is always the best, no matter how we measure three criteria. Forecasting Model 1 systematically hits the second place whereas forecasting model 11 is always the worst. In addition, the rankings with MMEU and MMEO differ greatly, which suggests the performance of competing models, *HM* and *PARCH*(1, 1) for example, are very sensitive to whether one penalizes positive errors more heavily than negative ones, or vice versa.

Table 5. Rankings of competing crude oil prices’ volatility forecasting models (Model 5).

Undesirable Inputs	Desirable Output	Rank from Best to Worst
ME, MSE	PCDCP	3→1→2→5→4→7→8→12→10→9→6→14→13→11
ME, MSVolScE	PCDCP	3→1→2→5→4→7→8→12→10→9→6→14→13→11
ME, MAE	PCDCP	3→1→2→4→12→14→5→10→9→13→7→8→6→11
ME, MAVolScE	PCDCP	3→1→2→4→14→13→(12/10)→5→9→8→7→6→11
ME, MMEU	PCDCP	3→1→2→5→8→7→4→12→10→9→6→14→13→11
ME, MMEO	PCDCP	3→1→13→4→7→8→5→14→10→12→9→6→11→2
MVolScE, MSE	PCDCP	3→1→2→5→4→7→8→12→10→9→6→14→13→11
MVolScE, MSVolScE	PCDCP	3→1→2→5→4→7→8→12→10→9→6→14→13→11
MVolScE, MAE	PCDCP	3→1→2→4→12→14→5→10→9→13→7→8→6→11
MVolScE, MAVolScE	PCDCP	3→1→2→4→14→13→(12/10)→5→9→8→7→6→11
MVolScE, MMEU	PCDCP	3→1→2→5→8→7→4→12→10→9→6→14→13→11
MVolScE, MMEO	PCDCP	3→1→13→4→7→5→8→14→10→9→12→6→11→2

Table 6 summarizes the rankings of competing forecasting models obtained by Model 6. Notice that Model 1 is the best no matter how we measure goodness-of-fit, and is systematically followed by forecasting model 3. Whereas forecasting Model 6 and Model 11 are the worst. Further, the performance of Model 2 and Model 13 are relatively sensitive to whether we penalize positive errors more heavily than negative ones or vice versa.

Table 6. Rankings of competing crude oil prices’ volatility forecasting models (Model 6).

Undesirable Inputs	Desirable Output	Rank from Best to Worst
ME, MSE	PCDCP	1→3→8→5→12→10→7→9→2→14→4→13→6→11
ME, MSVolScE	PCDCP	1→3→8→5→12→10→7→9→2→14→4→13→6→11
ME, MAE	PCDCP	1→2→3→4→8→5→12→14→10→13→9→7→6→11
ME, MAVolScE	PCDCP	1→2→3→4→8→5→(14/13)→12→10→9→7→6→11
ME, MMEU	PCDCP	1→3→8→5→12→10→7→4→9→14→2→6→13→11
ME, MMEO	PCDCP	1→3→8→5→12→10→7→9→14→4→13→6→11→2
MVolScE, MSE	PCDCP	1→3→5→8→7→(12/10)→9→2→4→14→13→6→11
MVolScE, MSVolScE	PCDCP	1→3→5→8→7→(12/10)→9→2→4→14→13→6→11
MVolScE, MAE	PCDCP	1→2→3→4→5→8→12→14→13→10→7→9→6→11
MVolScE, MAVolScE	PCDCP	1→2→3→4→5→8→13→14→(12/10)→9→7→6→11
MVolScE, MMEU	PCDCP	1→3→5→8→7→(12/10)→9→4→14→2→13→6→11
MVolScE, MMEO	PCDCP	1→3→5→8→7→10→12→9→4→14→13→6→11→2

With undesirable outputs, infeasibility only exists when measuring the undesirable outputs (Model 4). Consequently, we employ Model 8 rather than Model 4 to obtain a full ranking. While using the desirable outputs-oriented, the model used remains the same, that is Model 3. Table 7 summarizes the rankings of forecasting models obtained by Model 3. Model 3 is the best and is followed by Model 5. Model 2 is systematically the worst. The rankings under MMEU and MMEO differ significantly. This indicates that the relative performance of models, *HM* and *RW* for example, are very sensitive to whether we penalize positive errors more than negative ones or vice versa.

Table 8 summarizes the rankings of competing forecasting models obtained by Model 8. Forecasting model 3 is always the best no matter what measure we use. Whereas forecasting model 2 is systematically the worst. Also, rankings under MMEU and MMEO differ significantly, which indicates the performance of models are very sensitive to whether one penalize positive errors more than negative ones or vice versa.

Table 7. Rankings of competing crude oil prices’ volatility forecasting models (Model 3).

Undesirable Outputs	Desirable Outputs	Rank from Best to Worst
ME, MSE	PCDCP	3→14→5→10→12→9→13→4→6→11→8→7→1→2
ME, MSVolScE	PCDCP	3→14→5→10→12→9→13→4→6→11→8→7→1→2
ME, MAE	PCDCP	3→5→10→12→9→14→4→6→8→13→11→1→7→2
ME, MAVolScE	PCDCP	3→5→10→12→9→14→4→13→6→8→11→7→1→2
ME, MMEU	PCDCP	3→5→10→12→9→14→4→13→6→8→11→1→7→2
ME, MMEO	PCDCP	3→2→1→11→5→8→6→9→12→10→7→14→4→13
MVolScE, MSE	PCDCP	3→14→5→10→12→9→13→4→6→11→8→7→1→2
MVolScE, MSVolScE	PCDCP	3→14→5→10→12→9→13→4→6→11→8→7→1→2
MVolScE, MAE	PCDCP	3→5→10→12→9→14→4→6→8→13→11→1→7→2
MVolScE, MAVolScE	PCDCP	3→5→10→12→9→14→4→13→6→8→11→7→1→2
MVolScE, MMEU	PCDCP	3→5→10→12→9→14→4→13→6→8→11→1→7→2
MVolScE, MMEO	PCDCP	3→2→1→11→5→8→6→12→9→10→7→14→4→13

Table 8. Rankings of competing crude oil prices’ volatility forecasting models (Model 8).

Undesirable Outputs	Desirable Outputs	Rank from Best to Worst
ME, MSE	PCDCP	3→14→5→10→12→9→13→11→6→4→8→7→1→2
ME, MSVolScE	PCDCP	3→14→5→10→12→9→13→11→6→4→8→7→1→2
ME, MAE	PCDCP	3→5→10→9→12→6→14→8→13→4→11→7→1→2
ME, MAVolScE	PCDCP	3→5→(12/10)→9→6→14→8→4→13→11→7→1→2
ME, MMEU	PCDCP	3→5→14→10→12→9→13→4→11→6→8→7→2→1
ME, MMEO	PCDCP	3→1→2→11→5→8→6→9→12→10→7→14→4→13
MVolScE, MSE	PCDCP	3→14→5→10→12→9→13→11→6→4→8→7→1→2
MVolScE, MSVolScE	PCDCP	3→14→5→10→12→9→13→11→6→4→8→7→1→2
MVolScE, MAE	PCDCP	3→5→10→9→12→6→14→8→13→4→11→7→1→2
MVolScE, MAVolScE	PCDCP	3→5→(12/10)→9→6→14→8→4→13→11→7→1→2
MVolScE, MMEU	PCDCP	3→5→14→10→12→9→13→4→11→6→8→7→2→1
MVolScE, MMEO	PCDCP	3→1→2→11→5→8→6→12→9→10→14→7→4→13

To test whether the results of our models are statistically different from that in [10], we run the Wilcoxon signed ranks tests. The results are shown in Table 9. The statistical results reveal that at 10% significance level, the rankings obtained from this contribution are significantly different from those obtained by [10]. The previous results are not considered as reliable, mainly because there are infeasibilities in their results. In addition, the desirability of inputs and outputs are not discussed in the previous studies. We have clarified this point in our new models.

Table 9. Wilcoxon signed ranks test statistics.

	Model 5	Model 6	Model 3	Model 8
ME, MSE	0.433	0.002	0.001	0.016
ME, MSVolScE	0.016	0.084	0.779	0.023
ME, MAE	0.002	0.006	0.084	0.016
ME, MAVolScE	0.001	0.001	0.002	0.016
ME, MMEU	0.221	0.002	0.004	0.016
ME, MMEO	0.002	0.013	0.144	0.500
MVolScE, MSE	0.433	0.002	0.001	0.001
MVolScE, MSVolScE	0.013	0.272	0.917	0.030
MVolScE, MAE	0.001	0.001	0.064	0.016
MVolScE, MAVolScE	0.001	0.001	0.004	0.001
MVolScE, MMEU	0.056	0.001	0.001	0.001
MVolScE, MMEO	0.001	0.013	0.465	0.001

4. Conclusions

As a multi-criteria evaluation method, the super efficiency DEA method is an innovative way for achieving a consistent evaluation of crude oil prices' volatility forecasting models. However, the infeasibility problem in the original model decreases the reliability of its ranking results. Further, the desirability of inputs and outputs are of great significance to its application. This contribution focuses on these two aspects and proposes a modified super efficiency DEA framework for evaluating forecasting models. Four models corresponding to undesirable inputs and undesirable outputs are developed, as well as the way to address the appearance of infeasibility. Several conclusions are obtained from the empirical analysis. First, the ranking of the best and the worst forecasting models appears to be robust with respect to different measures under the same model. Second, our empirical results seem to suggest that considering models with undesirable inputs, *RW* and *SAM20* outperform other forecasting models and *EGARCH*(1, 1) is systematically the worst one. With respect to models dealing with undesirable outputs, *RM* turns into one of the worst models while *SAM20* remains as the best one. Finally, the Wilcoxon signed ranks tests reveal that rankings obtained by applying our framework are significantly different from the previous studies.

Author Contributions: Z.Z. and Q.J., Methodology; J.P., Data curation; H.X. and S.W., Formal analysis.

Funding: This research is funded by the National Natural Science Foundation of China [Nos. 71771082, 71801091] and Hunan Provincial Natural Science Foundation of China [No. 2017JJ1012].

Conflicts of Interest: The authors declare no conflict of interest.

References

- Narayan, P.K.; Sharma, S.; Poon, W.C.; Westerlund, J. Do oil prices predict economic growth? New global evidence. *Energy Econ.* **2014**, *41*, 137–146. [[CrossRef](#)]
- Zhou, Z.; Lin, L.; Li, S. International stock market contagion: A CEEMDAN wavelet analysis. *Econ. Model.* **2018**, *72*, 333–352. [[CrossRef](#)]
- Zhou, Z.; Jiang, Y.; Liu, Y.; Lin, L.; Liu, Q. Does international oil volatility have directional predictability for stock returns? Evidence from brics countries based on cross-quantilogram analysis. *Econ. Model.* **2019**, *80*, 352–382. [[CrossRef](#)]
- Kang, S.H.; Kang, S.M.; Yoon, S.M. Forecasting volatility of crude oil markets. *Energy Econ.* **2009**, *31*, 119–125. [[CrossRef](#)]
- García-Martos, C.; Rodríguez, J.; Sánchez, M.J. Modelling and forecasting fossil fuels, CO₂ and electricity prices and their volatilities. *Appl. Energy* **2013**, *101*, 363–375. [[CrossRef](#)]
- Tang, L.; Dai, W.; Yu, L.; Wang, S. A Novel CEEMD-Based EELM Ensemble Learning Paradigm for Crude Oil Price Forecasting. *Int. J. Inf. Technol. Decis. Mak.* **2015**, *14*, 141–169. [[CrossRef](#)]
- Pany, P.K.; Ghoshal, S.P. Dynamic electricity price forecasting using local linear wavelet neural network. *Neural Comput. Appl.* **2015**, *26*, 2039–2047. [[CrossRef](#)]
- Klein, T.; Walther, T. Oil price volatility forecast with mixture memory GARCH. *Energy Econ.* **2016**, *58*, 46–58. [[CrossRef](#)]
- Liu, J.; Ma, F.; Yang, K.; Zhang, Y. Forecasting the oil futures price volatility: Large jum. Forecasting the oil futures price volatility: Large jumps and small jumps. *Energy Econ.* **2018**, *72*, 321–330. [[CrossRef](#)]
- Xu, B.; Ouenniche, J. A data envelopment analysis-based framework for the relative performance evaluation of competing crude oil prices' volatility forecasting models. *Energy Econ.* **2012**, *34*, 576–583. [[CrossRef](#)]
- Andersen, P.; Petersen, N.C. A procedure for ranking efficient units in data envelopment analysis. *Manag. Sci.* **1993**, *39*, 1261–1264. [[CrossRef](#)]
- Tsolas, I.E. Firm credit risk evaluation: A series two-stage DEA modeling framework. *Ann. Oper. Res.* **2015**, *233*, 483–500. [[CrossRef](#)]
- Ouenniche, J.; Xu, B.; Tone, K. DEA in performance evaluation of crude oil prediction models. *Adv. DEA Theory Appl. Ext. Forecast. Models* **2017**, 381–403. [[CrossRef](#)]
- Georgantzinou, S.K.; Giannikos, I. A modeling framework for incorporating DEA efficiency into set covering, packing, and partitioning formulations. *Int. Trans. Oper. Res.* **2019**, *26*, 2387–2409. [[CrossRef](#)]

15. Liu, W.; Zhou, Z.; Liu, D.; Xiao, H. Estimation of portfolio efficiency via DEA. *Omega* **2015**, *52*, 107–118. [[CrossRef](#)]
16. Zhou, Z.; Xiao, H.; Jin, Q.; Liu, W. DEA frontier improvement and portfolio rebalancing: An application of China mutual funds on considering sustainability information disclosure. *Eur. J. Oper. Res.* **2018**, *269*, 111–131. [[CrossRef](#)]
17. Seiford, L.M.; Zhu, J. Infeasibility of super-efficiency data envelopment analysis models. *INFOR* **1999**, *37*, 174–187. [[CrossRef](#)]
18. Chen, Y. Measuring super-efficiency in DEA in the presence of infeasibility. *Eur. J. Oper. Res.* **2005**, *161*, 545–551. [[CrossRef](#)]
19. Lee, H.S.; Zhu, J. Super-efficiency infeasibility and zero data in DEA. *Eur. J. Oper. Res.* **2012**, *216*, 429–433. [[CrossRef](#)]
20. Halkos, G.; Petrou, K.N. Treating undesirable outputs in DEA: A critical review. *Econ. Anal. Policy* **2019**, *62*, 97–104. [[CrossRef](#)]
21. Liu, W.; Zhou, Z.; Ma, C.; Liu, D.; Shen, W. Two-stage DEA models with undesirable input-intermediate-outputs. *Omega* **2015**, *56*, 74–87. [[CrossRef](#)]
22. Liu, W.; Sharp, J. DEA models via goal programming. In *Data Envelopment Analysis in the Service Sector*; Deutscher Universitäts-Verlag: Wiesbaden, Germany, 1999; pp. 79–101.
23. Liu, W.B.; Zhang, D.Q.; Meng, W.; Li, X.X.; Xu, F. A study of DEA models without explicit inputs. *Omega* **2011**, *39*, 472–480. [[CrossRef](#)]
24. Chen, Y.; Liang, L. Super-efficiency DEA in the presence of infeasibility: One model approach. *Eur. J. Oper. Res.* **2011**, *213*, 359–360. [[CrossRef](#)]
25. Cook, W.D.; Liang, L.; Zha, Y.; Zhu, J. A modified super-efficiency DEA model for infeasibility. *J. Oper. Res. Soc.* **2009**, *60*, 276–281. [[CrossRef](#)]



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).