



Assessing joint maintenance shops in the Taiwanese Army using data envelopment analysis

Shinn Sun*

Graduate School of Management Sciences, College of Management, Aletheia University, No. 32, Chen-Li Street, Tamsui, Taipei 235, Taiwan, ROC

Available online 16 March 2004

Abstract

This paper proposes an alternative data envelopment analysis (DEA) method for assessing the performance of joint maintenance shops (JMSs) in the Taiwanese Army over two 6-month periods in 2000. It further examines the possibility of using DEA over time as a basis for continuous improvement. To help productivity improvements, the DEA results for the first time period were implemented at the five selected JMSs located in Taipei, Taoyuan, Taichung, Chaiyi, and Kaoshiung. The findings show that most of previously inefficient JMSs, on average, have become relative more efficient through DEA recommendation remedial actions. Three JMSs increased their technical and pure technical efficiencies; and one JMS increased its scale efficiency. Few JMSs were operating at increasing returns to scale, indicating a potential to expand. As such, it can be a valuable benchmarking tool for the JMS managers and assist in efficiently using scarce resources to produce productivity. Finally, the Army has accepted DEA as a performance measurement tool for JMSs.

© 2004 Elsevier B.V. All rights reserved.

Keywords: Data envelopment analysis; Performance measurement; Maintenance

1. Introduction

Defense budgets have declined over past years. In respond to using scarce resources more efficiently, Taiwanese Army developed and implemented a joint maintenance logistical reform programme in 1999. The programme aims at integrating six different types of military equipment maintenance service into a joint maintenance shop (JMS) at the corps level. A JMS is responsible for overhaul maintenance of military equipment including armament, chemical equipment, communication systems, engineering equipment, vehicles and armored vehicles. To maintain military equipment, the Army operates eight JMSs nationwide.

Performance measurement has been considered important for JMS managers. Improvements in the productivity of these JMSs are highly desirable since productivity gains are translatable into increased operational capabilities. The tactical logistics command (TLC), an Army logistics agency, is responsible for controlling the performance of Army maintenance units. Over the past decade the Army has not been allocated enough defense budgets to buy the new military equipment it feels are needed. This deficit could be reduced by channeling resources saved through productivity gains into new equipment purchases.

Army JMS managers use significant resources (e.g. manpower, supplies, equipment, and facility) in the equipment maintenance process. The outputs of that process are vital to the mission of the Taiwanese

* Tel.: +886-2-26212121ext5501; fax: +886-2-26260520.

E-mail address: shinn.sun@msa.hinet.net (S. Sun).

Army. Many potential benefits are possible through performance improvements. However, to realize this potential, the Army needs an objective system to measure and monitor the performance of JMSs over time. The TLC had no creditable performance system prior to this study.

The objective of this study is to examine potential applications of data envelopment analysis (DEA) in assessing JMS productivity. Productivity/performance is defined as ‘the ability to produce the outputs or services with a minimum resource level required’. This research focuses on new methodological developments in DEA and application of an alternative DEA model to perform actual improvements in system performance. In specifically, we first present a model by integrating ‘non-discretionary (NCN)’ model (Banker and Morey, 1986) and ‘assurance region (AR)’ model (Thompson et al., 1986, 1990) and, referred to ‘NCN–AR’ model. The NCN–AR model restricts the domain of multipliers and incorporates non-discretionary input and output variables that are beyond the control of an organization’s management. Analytical benefits of the proposed include the improvement of the discriminatory power of DEA and the restriction of the multiplier domain based on a prior information incorporating decision makers’ preferences. Then, we test whether the proposed method can be used as a basis for continuous improvements in a filed experiment where the TLC took part in. The DEA NCN–AR model can, thus, proceed one more step in forging a closer association with performance measurement and become a useful system to measure and monitor JMS maintenance productivity over time.

DEA, a linear programming based model, was originally presented by Charnes et al. (1978). The DEA methodology appears to be an appropriate measurement tool for this study due, partly, to the following reasons.

- (1) Performance of the JMSs is characterized by a large number of ‘inputs’ and ‘outputs’.
- (2) Although work routines at the JMS level are generally well defined, some maintenance tasks can be performed by different combinations of such routines.
- (3) It does not require explicit functional relationships between inputs and outputs.

Due to these reasons, assessing the efficiency of the JMSs by conventional techniques (see, e.g. Roll and Sachish, 1981) proved difficult. In addition to its sound theory, transparency and reproducible computational procedure, the DEA method has several advantages over traditional approaches such as ratio analysis and regression analysis (see Sherman, 1986). A major advantage is that DEA has been empirically validated many times over. As pointed out in Golany (1988), DEA is emerging as the leading method for efficient evaluation, in terms of both the number of research papers published and the number of applications to real world problems. We shall assume throughout this paper some knowledge of DEA. Readers not familiar with DEA are referred to Charnes et al. (1994).

The rest of paper is organized as follows. Section 2 provides a review of DEA and published DEA studies of maintenance unit performance in military. Section 3 introduces a methodology section discussing the sample, selection of input and output measures, and development of an alternative DEA model. Section 4 presents the empirical results of the analysis. Section 5 concludes the paper.

2. Literature review

2.1. Basic DEA models

This subsection provides a brief review of two basic models developed in the DEA literature, namely, the Charnes, Cooper and Rhode model (referred to as the CCR model) and the Banker, Charnes and Cooper model (referred to as the BCC model). A variety of extensions to the basic DEA models can be found in Cooper et al. (2000).

Charnes et al. (1978) initially introduced the CCR model to measure the relative efficiency of decision making units (DMUs) using multiple inputs to produce multiple outputs. They addressed constant returns to scale (CRS). CRS assumes that there is a proportional change between inputs and outputs. The CRS efficiency represents technical efficiency (TE), which measures inefficiencies due to input/output configuration and as well as size of operation.

Banker et al. (1984) presented the BCC model to determine whether there are any inefficiencies attributed to disadvantageous conditions under which a DMU is operating, which are not directly related to the in-

puts and outputs, and to allow for a larger peer group to be considered. They addressed variable returns to scale (VRS). VRS assumes that there is a proportional change in inputs does not result in a proportional change in outputs. The VRS efficiency represents pure technical efficiency (PTE), that is, a measure of efficiency without scale efficiency (SE). It is thus possible to decompose TE into PTE and SE. Scale efficiency can be estimated by dividing PTE into TE.

To investigate determine the current economy of scale that a DMU is experiencing, an efficiency index must be computed when the technology exhibits non-increasing returns to scale (NIRS). This can be done by relaxing the constraint on the weigh variables in the dual form of the BCC model to be less than or equal to one. The solution to the LP problem for DMU j when $\sum_{j=1}^n \lambda = 1$ is replaced by $\sum_{j=1}^n \lambda \leq 1$ is denoted by $TE^{j,NIRS}$. Again, the LP problem is solved for each DMU. $TE^{j,NIRS}$ is then compared to $TE^{j,VRS}$. If $TE^{j,NIRS} = TE^{j,VRS}$, then the scale inefficient DMU j is experiencing DRS and would be benefited by downsizing. On the other hand, if $TE^{j,NIRS} < TE^{j,VRS}$, then the scale inefficient DMU j is operating IRS and should expand its production. For a detailed discussion on returns to scale, see Banker and Thrall (1992).

2.2. DEA studies of maintenance unit performance in military

Three main studies that have used DEA to measure the performance of military maintenance units can be found in Charnes et al. (1985), Roll et al. (1989) and Clarke (1992). Each of these papers provided important contributions to the DEA and efficiency measurement literature. In what follows we discuss these studies in terms of the production model, sample size and DEA models, and provide some key observations.

The first of these applications, the Charnes et al. (1985) research, used the input orientated CCR model (Charnes et al., 1978) to analyze fourteen tactical fighter wings (TFWs) in the US Air Force using artificial data during the period October 1981 through May 1982. The production model consisted of eight inputs and four outputs. The inputs are: number of officers, percentage of assigned officers to authorized officers in a wing, number of airmen, percentage of assigned airmen to authorized airmen in a wing, 100,000 minus total number of hours in a specific month in which

the possessed aircraft were not mission capable due to supply problems, average number of aircraft in a wing; 1000 minus number of sortie losses due to external reasons, and cannibalization rate. The outputs include total number of sorties flown by each wing in a specific month, total number of hours that the possessed aircraft were fully or partially mission capable, 100000 minus total number of hours in a specific month in which the possessed aircraft were not mission capable due to maintenance problems, and fix rate within given time intervals. Important to highlight is their attempt to capture in their production model the observed quality of maintenance operations (e.g. the total number of hours in a specific month in which the possessed aircraft in each wing were not mission capable due to maintenance problems). In addition, they used the window analysis technique to increase the number of DMUs in dealing with degrees of freedom problems caused by insufficient number of DMUs.

Roll et al. (1989) applied the input orientated CCR model to analyze five maintenance units (MUs) in the Israeli Air Force using quarterly data. The production model consisted of three inputs (labor, operational facility, spare parts consumption in dollars) and six outputs (sorties types I and II, flying hours, ratio of maximum daily sorties to average, standard deviation of daily sorties and number of cancelled flights). The emphasis of this paper is on the choice of factors to enter the analysis and on assigning numerical values to qualitative factors. This paper attempted to investigate production models with different combination of pre-selected outputs. Performance of various levels of the MUs was evaluated.

Clarke (1992) also used the input orientated CCR model to evaluate vehicle maintenance performance at seventeen bases of the United States Air Force's Tactical Air Command over the time period 1983–1986. The production model consisted of four inputs (labor hours, material costs, number of trainees, and adjusted vehicle days) and two outputs (number of in-commission days for the vehicle fleet and number of mechanics trained each year). The author discussed the results of the DEA model over time and analyzed managerial reactions to the DEA method.

There are some key observations on this previous work using DEA.

- (1) The justification of the presence of the CRS is not provided. CRS assumes there is no signifi-

cant relationship between the scale of operations and efficiency. Unfortunately, the authors of previous studies did not explain why they considered a DEA model with CRS.

- (2) The issue of RTS is not examined. Investigating RTS is important because scale inefficient units can identify their current operating economies of scale (i.e. most productive scale size, MPSS) and determine whether their operations should be downsizing or expanding. However, to date none of the published research works has addressed this issue.
- (3) None of the research cited has treated weight restrictions and the presence of non-discretionary inputs and outputs. Many researchers (e.g. Dyson and Thanassoulis, 1988; Wong and Beasley, 1990) noted that traditional DEA allows for complete weight flexibility and may result in identifying a DMU with an unrealistic weight scheme to efficient. Restricting multipliers (weights) would reduce the number of efficient DMUs. As Banker and Morey (1986) noted, there exists non-discretionary variables that have values determined by forces exogenous to the organization under evaluation, and should keep the non-controllable variables fixed at its current level. These two issues have not been addressed in previous studies.

This paper contributes five extensions to the existing research. First, we propose an alternative DEA model which considers both weight restrictions and the presence of non-discretionary inputs and outputs. Second, we present a justification for using the BCC model. Third, we empirically test the effects of DEA in evaluating maintenance performance at five JMSs of the TLC. Fourth, we investigate returns to scale for each JMS and suggest potential improvements for inefficient JMSs. Finally, we discuss what actual managerial recommendations obtained from the DEA evaluation that would lead to in improving JMSs.

3. Methodology

3.1. Sample

The sample consists of five JMSs of the Taiwanese Army's TLC; these are core maintenance units at the corps level excluding three extremely small units

which are very different from the rest by their size and activities. The five JMS were labeled A, B, . . . , E for confidentiality. A JMS in each different month was treated as a DMU. Thus, A DMU's performance in a particular month was contrasted with its performance in other months in addition to the performance of the other DMUs. To perform DEA over time we collected two sets of data covering the time periods January–June and July–December 2000. Then, we performed two DEA analyses, where each analysis examined 30 DMUs.

3.2. Output and input measures

For the purpose of building a simple and yet fair picture of maintenance operations and productivity, the TLC had decided to group maintenance services by 'importance of operation' and 'complexity of operation'. This had resulted in the following four types of maintenance operations: (1) vehicles, (2) armament, (3) armored vehicles, and (4) 'other equipment' (chemical equipment, communication systems, and engineering equipment).

It was agreed that the four categories of operations were sufficient to convey an overview of importance and complexity of operation identified and they were retained for the purposes of the assessment reported in this paper. Similarly, the other input and output indicators were aggregated into composite factors, such as labor hours, spare parts costs and personnel trained. From the review of published sources, the TLC staff managers and JMS managers selected five output measures and six input measures for this study. The output and input variables were defined as follows.

(1) Input variables

- Total number of assigned vehicle to a JMS in a specific month (x_1).
- Total number of armament assigned to a JMS in a specific month (x_2).
- Total number of armored vehicle assigned to a JMS in a specific month (x_3).
- Total number of 'other equipment' assigned to a JMS in a specific month (x_4).
- Spare parts costs (x_5).
- Total number of available labor hours (x_6).

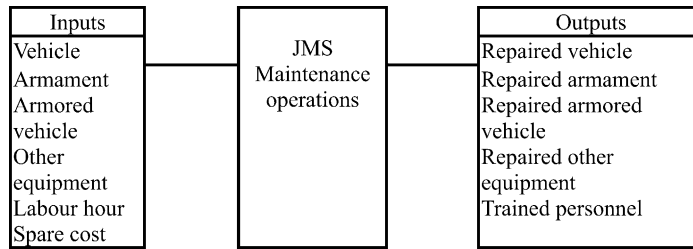


Fig. 1. General model of the joint maintenance process.

(2) Output variables

- Total number of assigned vehicles that are in serviceable condition each month (y_1).
- Total number of assigned armament that are in serviceable condition each month (y_2).
- Total number of assigned armored vehicles that are in Serviceable condition each month (y_3).
- Total number of assigned ‘other equipment’ that are in serviceable condition each month (y_4).
- Total number of personnel trained each month (y_5).

Among the input and output variables, only the amount of ‘spare parts costs’ is discretionary. Whether a JMS can obtain needed spare parts depends on defence budgets allocated to the Army and available spare parts provided by suppliers. Thus, the value of this input variable is beyond the control of the JMS managers. In addition, some outputs that can be used to measure effectiveness of JMS operations and per-

formance flexibility were not readily available. For example, average number of in-commission days for certain type of equipment used by the Army users and mean time between failures. As is typical in any empirical study, it is difficult but important that all possible inputs and outputs are completely and correctly specified. The omission of important variables produces biased results. In this paper, we base the analysis on the model of JMS maintenance operations depicted in Fig. 1.

After the cooperation of the Director of the TLC was secured, we collected monthly data for the year 2000 from five JMSs in the Taiwanese Army and the headquarters of TLC. The initial data covering the time period January–June 2000 was chosen for DEA analysis. Table 1 presents descriptive statistics for the initial data set. Table 2 shows correlations obtained. The JMS managers were then given 6 months to improve their efficiency scores. At the end of the year 2000, data covering the time period July–December

Table 1
Descriptive statistics for the five JMSs (30 observations)

	Mean	S.D.	Minimum	Maximum
Inputs				
x_1	163.6	90.54634	22	424
x_2	3196.067	1944.781	71	5849
x_3	108	97.6026	7	299
x_4	3053.8	1653.113	155	5637
x_5	8551154.7	9176527.036	238691	38674697
x_6	3976.157	4487.693	286.9	23487
Outputs				
y_1	181.5333	505.3777	11	2781
y_2	27.56667	51.86693	2	300
y_3	3.666667	9.741093	0	54
y_4	309.9	266.6728	22	1014
y_5	163.9	34.18708	110	258

Table 2
Correlation coefficients among inputs and outputs

	x_1	x_2	x_3	x_4	x_5	x_6	y_1	y_2	y_3	y_4	y_5
x_1	1										
x_2	0.526	1									
x_3	0.043	0.356	1								
x_4	0.028	-0.370	-0.473	1							
x_5	0.440	0.284	-0.222	0.040	1						
x_6	0.140	0.290	0.111	-0.094	-0.050	1					
y_1	0.531	0.266	0.410	-0.303	-0.030	0.121	1				
y_2	0.588	0.243	0.276	-0.252	0.016	0.076	0.927	1			
y_3	0.542	0.299	0.456	-0.332	-0.028	0.112	0.947	0.933	1		
y_4	0.539	0.258	0.094	0.068	0.345	-0.053	0.490	0.479	0.436	1	
y_5	0.566	-0.010	-0.100	0.522	0.212	-0.048	0.229	0.166	0.219	0.288	1

2000 were again collected and analyzed using DEA for continuous improvements.

Several points are worthy of note here.

- (1) Three operation categories (i.e. number of vehicles, armored vehicles and armament assigned to a JMS) have positively associated fix levels of the three different categories. So JMSs that perform well in one operation do so in the other two operations. 'Labor hours' has positively associated the numbers of vehicle armament, and other equipment that are in serviceable condition. The more labor hours were used, the more fix levels of these three types of assigned equipment would be increased.
- (2) 'Spare parts costs' is negatively associated with the repair levels of assigned vehicles and armored vehicles. Correlation coefficients are 0.03 and 0.028, respectively, which are not significant. It may release that that availability of spare parts for specific equipment would often affect a repair level of this specific equipment. This is as we expected.
- (3) The number of 'other equipment' assigned to a JMS is negatively associated with the repair levels of vehicles, armored vehicles and armament assigned to a JMS. It shows that as the number of 'other equipment' assigned to a JMS increases, the repair levels of other three types of equipment decrease. The JMS managers indicated that complexity of the operations, availability of spare parts and experience of workers might result in the negative correlations. The order of the relative complexity of maintaining four types of equipment

that the JMS managers considered from high to low is armored vehicles, vehicles, armament and other equipment. In practice, to obtain spare parts for armored vehicles, vehicles, armament is more difficult than to obtain spare parts for other equipment. This is because the first takes a longer time, normally 18 months, to buy materials overseas. To maintain armored vehicles, vehicles and armament need more skillful workers and longer operating hours. Knowing the fact that the more armored vehicles, vehicles, and armament assigned to a JMS the longer operating time to be, Army units would tend to assign more other equipment to a JMS. Furthermore, a JMS lacking for available spare parts and skillful workers in maintaining these three types of equipment would tend to fix more other equipment.

- (4) Total number of mechanics trained has positively associated the fix levels of four types of equipment. It is not surprising that well-trained personnel produce in higher repair levels of the four types of equipment.

Observations of negative correlations between inputs and outputs flaws DEA logic and therefore any results under these conditions are highly suspect. In order to preserve the isotone property for some outputs we must rescale the data to remove the negative correlation. A scale shift transformation was then employed by multiplying the original value of the output variable for each JMS with a scalar.

As Avkiran (2001) suggested, an alternative approach for choosing between CRS and VRS is to run the performance models under each assumption and

compare the efficiency scores. If a majority of the DMUs emerge with different scores under the two assumptions, then it is safe to assume VRS. However, if the majority of DMUs are assessed as having the same efficiency, one can employ CRS without being concerned that scale inefficiencies might confound the measure of technical efficiency. We then run the CCR and BCC models using the selected inputs and outputs. Comparing the two runs reveals different efficiency scores, thus confirming the presence of VRS. In addition, the JMS managers agreed to use the BCC model to identify the returns to scale of each JMS. Hence, the assessment was run under VRS.

3.3. DEA models

In the subsection, we present output-oriented DEA NCN–AR models for measuring the performance of JMSs. Output orientation is a natural choice because the objective of the TLC is to maximize outputs while using no more than the observed amount of any input. The development of the proposed DEA models is given in the following paragraphs.

Banker and Morey (1986) provided a modification to the basic DEA model permits the DEA solution to indicate the amount a controllable input can be reduced while keeping the non-controllable input while keeping the non-controllable input fixed at its current level. Mathematically, the primal multiplier form of the output-oriented NCN–CRS model can be represented as:

$$TE_{NCN}^{O,CRS} = \min \sum_{i \in D} v_i x_{ij} \tag{1}$$

subject to

$$\begin{aligned} \sum_{r=1}^s u_r y_{rj} - \sum_{i \in ND} v_i x_{ij} - \sum_{i \in D} v_i x_{ij} &\leq 0 \quad (j = 1, \dots, n) \\ \sum_{r=1}^s u_r y_{ro} - \sum_{i \in ND} v_i x_{io} &= 1 \\ v_i &\geq \varepsilon, \quad i \in D \quad (i = 1, \dots, m) \\ v_i &\geq 0, \quad i \in ND \\ u_r &\geq \varepsilon, \quad r = 1, \dots, s \end{aligned}$$

where $TE_{NCN}^{O,CRS}$ is the optimal value of technical efficiency of the target DMU (DMU_o) under the output-oriented NCN–CRS evaluation; x_{ij} the amount of input i to DMU j ; y_{rj} is amount of output r to DMU

j ; v_i the weight given to input i ; u_r the weight given to output r ; ε a non-Archimedean infinitesimal; s the number of outputs; m the number of inputs; and n the number of DMUs. The symbol D and ND refer to ‘discretionary’ and ‘non-discretionary’, respectively.

This NCN–CRS model can be extended in order to accommodate non-discretionary input and output variables in the following form:

$$TE_{NCN}^{O,CRS} = \min \sum_{i \in D} v_i x_{ij} \tag{2}$$

subject to

$$\begin{aligned} \sum_{r \in D} u_r y_{rj} - \sum_{i \in ND} v_i x_{ij} - \sum_{i \in D} v_i x_{ij} &\leq 0 \\ \sum_{r \in D} u_r y_{ro} - \sum_{i \in ND} v_i x_{io} &= 1 \\ v_i &\geq \varepsilon, \quad i \in D \\ v_i &\geq 0, \quad i \in ND \\ u_r &\geq \varepsilon, \quad r \in D \end{aligned}$$

Traditional CCR and BCC models place no constraints on weight attributed to each input and each output in the multiplier problem, thus allowing absolute weight flexibility. This can result in identifying a DMU with an extreme weighting scheme to be efficient. Outlier units will tend to be classified as technically efficient and zero weights will be assigned to most of their inputs/outputs. This represents a contradiction in itself because if such inputs/outputs were not important, they would not be included in the analysis. Absolute weight flexibility can result in an overestimation of technical efficiency. To overcome this problem, the assurance region (AR) model developed in Thompson et al. (1986) and further defined in Thompson et al. (1990) can be used to restrict the values that virtual weights may attain and thereby, limit the range of acceptable efficient input and output levels. The AR model provides lower and upper bounds on the admissible values of variables. These bounds take the following form.

$$\begin{aligned} \delta_r &\leq u_r / u_{ro} \leq \gamma_r \quad \text{for all } r \neq r_o, \\ \alpha_i &\leq v_i / v_{io} \leq \beta_i \quad \text{for all } i \neq i_o \end{aligned} \tag{3}$$

Here, u_{ro} and v_{io} represent primal variables which serve to establish the upper and lower bounds represented by δ_r , γ_r and by α_i , β_i for the primal variables associated with each output and input.

We now add (3) to (2). Thus, the output-oriented NCN–AR–CRS model is expressed as:

$$TE_{NCN-AR}^{O,CRS} = \min \sum_{i \in D} v_i x_{ij} \tag{4}$$

subject to

$$\begin{aligned} \sum_{r \in D} u_r y_{rj} - \sum_{i \in ND} v_i x_{ij} - \sum_{i \in D} v_i x_{ij} &\leq 0 \\ \sum_{r \in D} u_r y_{ro} - \sum_{i \in ND} v_i x_{io} &= 1 \\ \delta_r \leq u_r / u_{r0} \leq \gamma_r \text{ for all } r \neq r_o, \\ \alpha_i \leq v_i / v_{i0} \leq \beta_i \text{ for all } i \neq i_o \\ v_i &\geq \varepsilon, i \in D \\ v_i &\geq 0, i \in ND \\ u_r &\geq \varepsilon, r \in D \end{aligned}$$

where $TE_{NCN-AR}^{O,CRS}$ is the optimal value of technical efficiency of the target DMU (DMU_o) under the output-oriented NCN–AR–CRS evaluation.

Following is the primal form of the output-oriented NCN–AR–VRS model.

$$TE_{NCN-AR}^{O,VRS} = \min \sum_{i \in D} v_i x_{ij} - v_o \tag{5}$$

subject to

$$\begin{aligned} \sum_{r \in D} u_r y_{rj} - \sum_{i \in ND} v_i x_{ij} - \sum_{i \in D} v_i x_{ij} - v_o &\leq 0 \\ \sum_{r \in D} u_r y_{ro} - \sum_{i \in ND} v_i x_{io} &= 1 \\ \delta_r \leq u_r / u_{r0} \leq \gamma_r \text{ for all } r \neq r_o, \\ \alpha_i \leq v_i / v_{i0} \leq \beta_i \text{ for all } i \neq i_o \\ v_i &\geq \varepsilon, i \in D \\ v_i &\geq 0, i \in ND \\ u_r &\geq \varepsilon, r \in D \\ v_o &\text{ free in sign} \end{aligned}$$

where $TE_{NCN-AR}^{O,VRS}$ is the optimal value of pure technical efficiency of the target DMU (DMU_o) under the output-oriented NCN–AR evaluation.

Then, scale efficiency for DMU_o is obtained as:

$$SE_{NCN-AR}^O = \frac{TE_{NCN-AR}^{O,CRS}}{TE_{NCN-AR}^{O,VRS}} \tag{6}$$

It represents the proportion of inputs that can be further reduced after pure technical inefficiency is eliminated if scale adjustments are possible. It has a value of less than or equal to one. If DMU_o has a value equal to one, DMU_o is operating at the constant returns to

scale size. If SE^o is less than one, DMU_o is scale inefficient and there is potential input saving through the adjustment of its operational scale. Whether the scale inefficient DMU_o should be either downsizing or expanding depends on its current operating scale.

4. Empirical results

4.1. Restriction on output multipliers

In order to measure NCN–AR efficiency, the prior information for (3) was determined by one TLC staff manager and five JMS managers. Weight restriction is imposed on outputs because the TLC attempts to maximize the outputs while using the given input levels. The managers made a pairwise comparison of the relative importance of each output using their subjective judgments. Analytic Hierarchy Process (AHP) proposed by Saaty (1980) was used to quantifying their subject judgments. The results derived from this AHP analysis then served as a guideline for setting the upper and lower bounds of (3). The adjusted weights on outputs of the six managers are shown in Table 3.

In order to take into account the various views which the managers showed in the Table 1, all managers agreed with us to use the assurance region concept in the following way.

Let the weight for output r (y_r) be u_r ($r = 1, 2, 3, 4, 5, 6$). The ratio u_1/u_2 takes the value $0.294/0.176 = 1.670$ for Manager 1, $0.229/0.2 = 1.145$ for Manager 2, $0.243/0.162 = 1.5$ for Manager 3, $0.216/0.189 = 1.143$ for Manager 4, $0.216/0.189 = 1.404$ for Manager 5 and $0.182/0.242 = 0.752$ for manager 6. Thus, we have the range of the ratio u_1/u_2 as expressed by the lower

Table 3
Adjusted weights on outputs of six managers

Manager	y_1	y_2	y_3	y_4	y_5	Sum
Manager 1	0.294	0.176	0.265	0.147	0.118	1.000
Manager 2	0.229	0.200	0.257	0.171	0.142	1.000
Manager 3	0.243	0.162	0.270	0.189	0.135	1.000
Manager 4	0.216	0.189	0.270	0.162	0.162	1.000
Manager 5	0.219	0.156	0.281	0.156	0.188	1.000
Manager 6	0.182	0.242	0.273	0.182	0.121	1.000

and upper bounds below

$$0.752 \leq u_1/u_2 \leq 1.670$$

In the same way we can find the range of u_j/u_i for each pair (i, j) as follows:

$$0.752 \leq u_1/u_2 \leq 1.670 \quad 0.857 \leq u_2/u_4 \leq 1.330$$

$$0.666 \leq u_1/u_3 \leq 1.109 \quad 0.830 \leq u_2/u_5 \leq 2$$

$$1 \leq u_1/u_4 \leq 2 \quad \text{and } 1.429 \leq u_3/u_4 \leq 1.803$$

$$1.165 \leq u_1/u_5 \leq 2.492 \quad 1.495 \leq u_3/u_5 \leq 2.256$$

$$1.128 \leq u_3/u_2 \leq 1.506 \quad 0.830 \leq u_4/u_5 \leq 1.504$$

Use of these “paired comparisons” is in line with the principles adopted in the AHP. Using these upper and

lower bound as the assurance region constraints for output weights, we applied the NCN–AR–VRS model to two DEA analyses. The software utilized to perform the DEA analyses was the *DEA-Solver: Professional Version*, developed by Saitech Inc (2001).

4.2. Initial results

Initially, ten of the thirty DMUs identified in Table 4 were found to have pure technical efficiencies under VRS in one or more aspects of their operations and received a DEA efficiency score <100%. This means that each of the ten DMUs was inefficient relative to an efficient subset of the remaining twenty DMUs.

Table 4
Efficiency results over time

DMU code	January–June 2000				DMU code	July–December 2000			
	TE	PTE	SE	RTS		TE	PTE	SE	RTS
A-1	60.70	100	60.70	IRS	A-7	11.16	13.18	84.67	CRS
A-2	74.07	100	74.07	IRS	A-8	100	100	100	CRS
A-3	79.70	100	79.70	IRS	A-9	100	100	100	CRS
A-4	100	100	100	CRS	A-10	100	100	100	CRS
A-5	100	100	100	CRS	A-11	100	100	100	CRS
A-6	100	100	100	CRS	A-12	25.81	100	25.81	IRS
B-1	13.77	17.11	80.48	IRS	B-7	28.99	52.88	54.82	CRS
B-2	19.17	23.13	82.88	IRS	B-8	60.97	100	60.97	IRS
B-3	43.62	45.41	96.06	IRS	B-9	100	100	100	IRS
B-4	34.31	38.84	88.34	IRS	B-10	27.38	43.40	63.09	IRS
B-5	21.13	24.83	85.10	IRS	B-11	29.92	100	29.92	IRS
B-6	28.70	33.13	86.63	IRS	B-12	50.47	100	50.47	IRS
C-1	28.46	42.89	66.36	IRS	C-7	8.23	12.37	66.53	CRS
C-2	32.71	45.99	71.12	IRS	C-8	53.02	100	53.02	IRS
C-3	58.63	80.78	72.58	IRS	C-9	46.63	100	46.63	IRS
C-4	21.24	31.12	68.25	IRS	C-10	100	100	100	CRS
C-5	18.69	26.71	69.97	IRS	C-11	67.93	100	67.93	IRS
C-6	32.57	49.56	65.72	IRS	C-12	67.83	100	67.83	IRS
D-1	40.24	100	40.24	IRS	D-7	7.23	10.16	71.16	IRS
D-2	22.05	100	22.05	IRS	D-8	16.78	100	16.78	IRS
D-3	21.22	89.63	23.68	IRS	D-9	27.41	56.12	48.84	IRS
D-4	20.06	56.36	35.59	IRS	D-10	27.82	66.64	41.74	IRS
D-5	32.74	100	32.74	IRS	D-11	23.22	100	23.22	IRS
D-6	17.10	38.45	44.47	IRS	D-12	13.86	100	13.86	IRS
E-1	100	100	100	CRS	E-7	1.42	5.75	24.70	CRS
E-2	11.19	11.97	93.48	IRS	E-8	32.32	36.62	88.26	CRS
E-3	83.55	90.31	92.51	IRS	E-9	44.79	48.79	91.80	CRS
E-4	24.45	25.95	94.22	IRS	E-10	35.26	39.96	88.24	IRS
E-5	14.24	15.12	94.18	IRS	E-11	100	100	100	CRS
E-6	21.61	22.77	94.91	IRS	E-12	60.41	65.33	92.47	IRS
Mean	42.53	60.34	73.87			48.962	75.04	65.76	

Note: TE, technical efficiency; PTE, pure technical efficiency; SE, scale efficiency; RTS, returns to scale; CRS, constant returns to scale; DRS, decreasing returns to scale; IRS, increasing returns to scale.

As can be seen in Table 4, four DMUs in our sample were operating at CRS/MPSS and the remaining 26 at IRS in the first data collection period. JMSs operating at IRS had not grown larger than their most productive scale size and could consider expansion. Table 4 also shows that twenty-six of the thirty DMUs were scale inefficient with average scale efficiency score 73.87. It suggests that further potential output improvements of 12.13 percent if it is possible for a JMS to operate at the constant returns to scale technology. Among the 26 scale-inefficient JMSs, 20 did not achieve pure technical efficiencies. It should be noted, in general, it is more difficult to reduce scale inefficiency than it is to reduce technical inefficiency. Only when a DMU becomes technically efficient does it make sense to deal with scale inefficiency.

In order to examine stability and other properties of the efficiency evaluations, we performed mean–variance analysis on efficiency scores. Table 5 presents the results of mean–variance analysis. JMS A has the highest mean and low variance in its TE ratings while the other four JMSs have low means and variance. These four JMSs are due to inappropriate input/output configuration as well as size of operations. Variations in PTE ratings show that JMS A not

Table 5
Mean–variance analysis

JMS	Mean			Variance	
	1–6	7–12	Change (%)	1–6	7–12
Technical efficiency					
A	85.745	72.828	–15.064	234.952	1793.38
B	26.783	49.622	85.274	120.614	795.199
C	32.050	57.273	78.699	203.087	917.439
D	25.568	19.387	–24.175	80.006	67.038
E	42.507	45.700	7.512	1506.498	902.406
Pure technical efficiency					
A	100	85.530	–14.470	0	1256.285
B	30.408	82.713	172.010	112.898	726.176
C	46.175	85.395	84.938	365.066	1279.836
D	80.740	72.153	–10.635	714.941	1291.39
E	44.353	49.408	11.397	60.340	828.710
Scale efficiency					
A	85.745	85.080	–0.776	234.952	880.705
B	86.582	59.878	–30.842	29.212	526.572
C	69.000	66.990	–2.913	7.308	340.230
D	33.128	35.933	2.805	79.484	491.386
E	94.883	80.912	–14.724	6.944	647.328

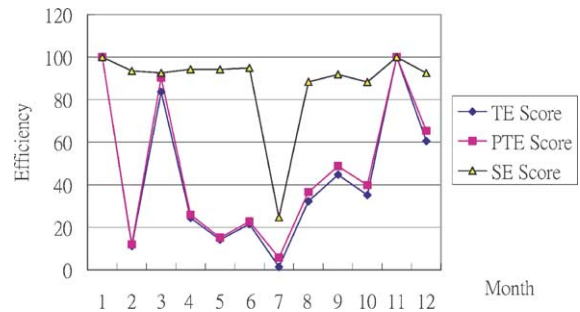


Fig. 2. Efficiency trends for the most inefficient JMS.

only has the highest mean and the lowest variance, followed by JMS D. Furthermore, JMSs B, C, and E also have low means and high variances in their PTE ratings. These three JMSs are not on pure technical efficient frontier due to inappropriate input/output configuration. JMS D has the lowest mean and high variance in its scale efficiency ratings. The inefficiency is due to not appropriate size of operation.

A closer look at each of the inefficient JMSs can be taken by depicting efficiency trends in TE, PTE, and SE scores. For brevity, Fig. 2 depicts the efficiency trends for the most inefficient JMS/JMS E with the lowest mean PTE value of 46.88. There are unstable efficiency trends for JMS E. The SE score for JMS E slightly differed during the time period January–June 2000 while its TE and PTE scores evidently differed and dramatically declined to the bottom in July. After investigating causes of this considerable change, we have found that the declining is due to lack of available spare parts for supporting its maintainability. Once a JMS obtained spare parts needed through defence acquisition undertaken in July that is a new buy for a new fiscal year, it was then in an effective operational state. Thus, each of three efficiency scores for JMS E exhibits an upward trend starting in July 2000. The Director of the TLC recognized this issue and asked his staff managers to take it into account in future acquisition and maintenance of spare part inventories. However, acquisition of spare parts for military use may not be easy due to foreigner governments’ restrictions.

The solution of the DEA models yields target input and output level which would render inefficient JMSs efficient, if not already so (see Charnes et al., 1978). Table 6 shows the potential improvements for the inefficient JMSs.

Table 6
Targets and potential improvements for the inefficient DMUs

DMU code	x_1	x_2	x_3	x_4	x_6	y_1	y_2	y_3	y_4	y_5
Target										
B-1	149.83	1036.54	58	3313	2176.60	57.63	409.94	8.60	139.07	182.10
B-2	156.79	1054.69	58	3522	2071.30	58.21	473.57	9.83	274.82	189.71
B-3	191.31	1532.52	81	3471	3988.98	99.49	785.58	16.18	122.24	160.74
B-4	169.64	1459.58	80	2877	3511.52	90.21	709.19	14.64	330.37	153.39
B-5	152.99	1365.73	77	2522	3218.31	81.94	637.15	13.22	300.94	148.26
B-6	165.68	1500.19	83	2624	3319.34	90.15	714.29	14.70	325.89	151.33
Mean	164.37	1324.88	72.83	3054	3047.68	79.61	621.62	12.86	248.89	164.26
C-1	103.63	577.21	38	2836	3341.59	41.93	243.27	5.73	193.27	139.88
C-2	104.65	705.49	36	3340	3168.50	30.57	181.34	3.88	584.79	148.39
C-3	109.58	575.64	37	3057	3514.75	43.82	256.62	6.01	202.39	142.18
C-4	102.85	575.20	38	2813	3323.23	41.61	240.74	5.68	191.99	139.61
C-5	104.42	596.21	39	2814	3327.29	42.76	251.84	5.89	195.25	139.93
C-6	111.46	629.49	39	3067	2713.50	33.34	244.60	5.56	223.83	157.61
Mean	106.09	609.87	37.83	2988	3231.48	39.01	236.40	5.46	265.25	144.60
D-3	55.92	1660	143.18	1461	391.16	14.02	46.76	1.49	107.14	122.48
D-4	64	2139.04	169.61	1575.16	416.56	19.51	98.01	2.47	146.99	119.16
D-6	81	3588.17	247.93	2670.10	343.05	17.79	90.60	2.24	250.95	118.83
Mean	66.97	2462.40	186.91	1902.09	383.59	17.11	78.46	2.07	168.36	120.16
E-2	141	1999.61	116.23	387.38	1642.34	97.25	833.09	16.69	307.81	135.16
E-3	143	2025.35	117.44	391.28	1649.43	98.69	846.85	16.95	311.92	135.58
E-4	160	2244.15	127.76	424.48	1709.63	110.87	963.87	19.19	346.89	139.18
E-5	160	2244.15	127.76	424.48	1709.63	110.87	963.87	19.19	346.89	139.18
E-6	173	2411.47	135.65	449.86	1755.67	120.18	1053.3	20.91	373.64	141.93
Mean	155	2184.95	124.97	421.55	1693.34	107.57	932.19	18.59	337.43	138.21
Potential improvement (%)										
B-1	-47.06	-76.42				423.94	819.88	859.79	133.72	1.73
B-2	-47.21	-75.95				263.80	999.90	882.91	-4.32	1.45
B-3	-36.23	-65.94			-29.39	77.66	999.90	439.20	-63.84	-1217
B-4	-36.70	-68.87			-85.05	329.59	454.06	265.99	-50.98	-16.18
B-5	-48.66	-72.03			-76.68	192.64	696.44	999.90	37.42	-18.98
B-6	-41.04	-69.56			-55.46	291.97	349.24	267.62	-15.35	-17.31
Mean	-42.82	-71.46			-61.65	263.27	719.90	619.24	6.11	-211.1
C-1	-13.64	-81.57			-29.42	222.53	312.32	90.89	-18.45	-1.49
C-2	-17.60	-83.11				239.61	571.63	-3.06	35.37	4.51
C-3	-32.36	-87.85			-21.63	-4.73	133.29	100.34	-38.67	0.13
C-4	-19.02	-89.12			-39.22	197.24	999.90	183.77	1.05	-1.69
C-5	-19.67	-89.37			-26.94	256.37	403.67	194.38	150.32	-0.76
C-6	-33.26	-89.24				101.78	389.21	85.25	-26.61	19.40
Mean	-22.59	-86.71			-29.30	168.80	468.34	108.59	17.17	3.35
D-3	-3.58		-1.93			-17.53	6.27	49.31	386.98	4.68
D-4		-17.63	-13.90	-30.36		21.93	326.13	146.53	31.24	1.84
D-6		-20.79	-17.08	-27.01		48.24	723.66	124.38	185.17	-1.79
Mean	-3.58	-19.21	-10.97	-28.69		17.55	352.02	106.74	201.13	1.58
E-2		-37.18	-51.37	-86.70	-44.88	594.67	999.90	233.78	115.25	-29.60
E-3		-52.92	-51.07	-84.01	-43.24	996.52	-0.37	239.05	-57.39	-28.64
E-4		-38.52	-48.27	-83.88	-22.71	691.90	404.64	74.49	24.78	-23.95
E-5		-41.21	-47.42	-80.42	-53.08	823.88	999.90	91.94	48.25	-21.37
E-6		-38.84	-41.78	-79.54	-24.18	999.90	431.99	596.93	0.44	-20.27
Mean		-41.73	-47.98	-82.91	-37.62	821.37	567.21	247.24	26.27	-24.77

The TARGET column shows the amount of inputs and outputs that an inefficient unit should be using or producing in order to be efficient while POTENTIAL IMPROVEMENT column shows how much, in percentage terms, an inefficient JMS's use of inputs or production of output needs to change by in order for it to be efficient. For example, the most inefficient JMS/JMS E, on average, can increase its fixed vehicle number by 821.37%, fixed armament number by 567.21%, fixed armored vehicle number by 247.24%, and number of other equipment fixed by 26.27%; while simultaneously reducing its assigned armament number by 41.73%, assigned armored vehicle number by 47.98%, number of other equipment assigned by 82.91%, labor hour by 37.62% and number of trainees by 24.77%. Similarly, such an analysis can be applied to the other inefficient JMSs.

We should note here, it is sometime difficult for an inefficient JMS to make improvement when its increasing output level or reducing input levels is over 100%. JMS managers can use the potential improvements reported in Table 6 to identify directions for improving inefficiency. In addition, we tested the stability of the efficiency scores to variable variation. We, then, have found that the more input and output variables are included in the model, the higher will be the number of DMUs with an efficiency score equal to unity.

Once this experiment was completed, the DEA results were then provided to the managers of the JMSs and to the staff managers of the TLC with instructions from us to use the results to improve their relative efficiency score over the next six months. This was done by providing them several recommendations: (1) consideration of expansion of the inefficient JMSs, (2) use of the potential improvements as a guideline for improving inefficiencies, and (3) development of an effective operational state.

4.3. Results over time

At the end of the year 2000, data were again collected and analyzed using the DEA models. Table 4 summarizes these results for the DEA models. Table 5 presents the comparison of the DEA results over time for each of JMSs in terms of mean efficiency.

In contrast to efficiency gains made from January to June 2000, the number of DMUs received TE, PTE

and SE scores of 100% in the July–December evaluation has increased from 3 to 7, 10–18, and 4–7, respectively. The mean TE score has increased from 42.53 to 48.96 (+15.12%); the mean PTE score has increased from 60.34 to 75.04 (+24.36%); and the mean SE score has decreased from 73.87 to 65.76 (–10.98%). In addition, the number of CRS has increased from 4 to 12 while the number of IRS has reduced from 26 to 18.

From Table 5, JMSs B, C, and E have increased technical and pure technical efficiencies over time while the other two JMSs have reduced their efficiency scores. In addition, JMS D has slightly increased its scale efficiency score while the other four JMSs have reduced their scale efficiency scores.

4.4. Managerial reactions to DEA

After presenting the results of DEA to LTC and JMS managers, we had their reactions to DEA. All managers agree DEA was helpful in assessing relative organizational performance, identifying inefficiencies, locating slack resources, and identifying comparable units. However, some of managers expressed the view that DEA did not address quality of operations. As for the Director of the TLC, he was happy with the DEA results and recommended that a performance measurement tool should take into account quality and efficiency of operations. Currently, DEA is implemented at JMSs and more experiments to test the effects of DEA in other TLC maintenance units are undertaken.

5. Conclusions

This research has developed an alternative DEA model to assist the TLC in evaluating the performance of JMSs and empirically tested the applicability of the proposed model. The proposed model takes into account non-controllable inputs/outputs and weight restrictions. The research findings have shown that using DEA over time can be seen as a basis for continuous improvement. On average, JMSs B, C, and E have increased technical and pure technical efficiencies through DEA recommendation remedial actions while the other two JMSs have reduced their scores. Only JMS D has slightly increased its scale efficiency

score while the other four JMSs have reduced their scores.

Since the JMSs had no creditable system prior to the DEA analysis, this study has made a stand alone contribution by presenting the DEA results. Clearly, the stricter the standards against which productivity is evaluated, the sharper the management tool. Such additional sharpness can be introduced by adding quality metrics to future DEA research efforts.

Acknowledgements

The author would like to thank Dr D.M. Stewart for editorial efforts and the two anonymous reviewers for their precious comments. Thanks also go to Dr. D.L. Bricker for helping the author with exposition of the paper. All errors and omissions remain the responsibility of the author.

References

- Avkiran, N.K., 2001. Investigating technical and scale efficiencies of Australian universities through data envelopment analysis. *Socio-Economic Planning Sciences* 35, 57–80.
- Banker, R.D., Morey, R.C., 1986. Efficiency analysis for exogenously fixed inputs and outputs. *Operations Research* 34 (4), 513–521.
- Banker, R.D., Thrall, R.M., 1992. Estimation of returns to scale using data envelopment analysis. *European Journal of Operational Research* 62 (1), 74–84.
- Banker, R.D., Charnes, A., Cooper, W.W., 1984. Some models for estimating technical and scale inefficiencies in data envelopment analysis. *European Journal of Operational Research* 30 (9), 1078–1092.
- Charnes, A., Cooper, W.W., Rhodes, E., 1978. Measuring the efficiency of decision making units. *European Journal of Operational Research* 2 (6), 429–444.
- Charnes, A., Clark, C.T., Cooper, W.W., Golany, B., 1985. A developmental study of data envelopment analysis in measuring the efficiency of maintenance units in the US Air Forces. *Annals of Operations Research* 2, 95–112.
- Charnes, A., Cooper, W.W., Lewin, A.Y., Seiford, L.M., 1994. *Data Envelopment Analysis: Theory, Methodology and Applications*. Kluwer Academic Publisher, London.
- Cooper, W.W., Seiford, L.M., Tone, K., 2000. *Data Envelopment Analysis: A Comprehensive Text with Models, Applications, References and DEA-Solver Software*. Kluwer Academic Publishers, London.
- Clarke, R., 1992. Evaluating USAF vehicle maintenance productivity over time: an application of data envelopment analysis. *Decision Sciences* 23, 376–384.
- Dyson, R.G., Thanassoulis, E., 1988. Reducing weight flexibility in data envelopment analysis. *Journal of the Operational Research Society* 39 (6), 563–576.
- Golany, B., 1988. An interactive MOLP procedure for the extension of DEA to effectiveness analysis. *Journal of Operational Research Society* 39 (8), 725–734.
- Roll, Y., Sachish, A., 1981. Productivity measurement at the plant level. *Omega* 9 (1), 37–42.
- Roll, Y., Golany, B., Seroussy, D., 1989. Measuring the efficiency of maintenance units in the Israeli Air Force. *European Journal of Operational Research* 43 (2), 136–142.
- Saaty, T.L., 1980. *The Analytic Hierarchy Process*. McGraw-Hill, New York.
- Saitech Inc., 2001. *DEA-Solver: Professional Version 2.0*, New Jersey.
- Sherman, H.D., 1986. Managing productivity of health care organizations, in: Silkman, R.H. (Ed.), *Measuring Efficiency: An Assessment of Data Envelopment Analysis, New Directions for Program Evaluation*. Jossey Bass, San Francisco, pp. 31–46.
- Thompson, R.G., Langemeier, L.N., Lee, C.T., Smith, R.M., 1990. The role of multiplier bounds in efficiency analysis with application to Kansas farming. *Journal of Econometrics* 46, 93–108.
- Thompson, R.G., Singleton, F.D., Thrall, R.M., Smith, B.A., 1986. Comparative site evaluation for locating high energy lab in Texas. *Interface* 16, 1380–1395.
- Wong, Y.-H.B., Beasley, J.E., 1990. Restricting weight flexibility in data envelopment analysis. *Journal of the Operational Research Society* 41 (9), 829–835.