



## EFFICIENCY MEASUREMENT OF BUS ROUTES IN HANOI CITY: AN APPLICATION OF DATA ENVELOPMENT ANALYSIS (DEA)

Tran Khac Duong\*, Do Quoc Cuong

University of Transport and Communications, No 3 Cau Giay Street, Hanoi, Vietnam

### ARTICLE INFO

TYPE: Research Article

Received: 9/3/2020

Revised: 15/4/2020

Accepted: 17/4/2020

Published online: 28/5/2020

<https://doi.org/10.25073/tcsj.71.4.6>

\* Corresponding author

Email: tkduong@utc.edu.vn

**Abstract.** Efficiency analysis of bus transit at the route level is critical to understand the existing performance of individual routes within a bus system and identify operational problems as well as effectively optimise their performance. This article applies the Data Envelopment Analysis (DEA) model to examine the performance of 38 bus routes in Hanoi, Vietnam. The results indicated the best and the inefficient bus routes within the given sample and identified the internal sources of inefficiency, including: *number of stops* and *vehicles*. The findings provide bus agencies in the case study with additional and useful information for decision making.

**Keywords:** Data envelopment analysis (DEA), bus performance evaluation, technical efficiency, operational effectiveness, decision making unit (DMU)

© 2020 University of Transport and Communications

### 1. INTRODUCTION

Transit agencies aim to continuously optimise their performance and improve the quality of service in order to increase transit ridership effectively [1, 2]. Measuring the performance of individual routes within a transit system plays a critical role in identifying problems in system design, operation and control, and in seeking means to increase ridership effectively. However, measuring the performance of individual transit routes is complex because multiple objectives (related to the operators, users, and community), and multiple input and output variables, exist [3]. The complexity of transit performance led to the development of a

framework by Fielding et al. [4] for transit system performance measurement. This framework consists of three dimensions; *technical efficiency*, *operational effectiveness*, and *service effectiveness* (refer to section 2). This framework allows one to compare the performance of different transit systems for a particular performance concept (such as vehicle efficiency, fuel efficiency, and operational safety) by using single ratios of service output and service input. This traditional approach cannot provide a single overall measure for transit performance evaluation [5]. The issue is addressed by using the Data Envelopment Analysis (DEA) approach, which allows one to compare the performance of different transit routes (which is considered as production units) within a transit system by building up the production frontier directly from an actual dataset and generating the efficiency scores for individual routes [1-3, 6]. In large urban areas of Vietnam (such as Hanoi and Ho Chi Minh city), there has been very little work quantitatively examining the performance of transit routes. Furthermore, there have been no studies, as far as we are aware, using the DEA for transit route performance evaluation.

This article employs the DEA model to measure the performance of individual bus routes in Hanoi, Vietnam, considering them as sub-units of a transit system. The scientific contributions of this article provide: (1) empirical understanding of bus route performance in a case study of Hanoi using the DEA model; and (2) identification of internal sources of inefficiency of given bus routes.

The article is structured as follows: Section 2 presents the review of the literature. Section 3 presents the proposed methodology, followed by the details on the dataset used for empirical analysis, discussion on the results and recommendations in section 4. Finally, the paper is concluded in section 5.

## 2. LITERATURE REVIEW

### 2.1. Transit performance concepts

Fielding *et al.* [4] have distinguished transit performance into three concepts: *technical efficiency*, *operational effectiveness*, and *service effectiveness*.

*Technical efficiency* represents the process through which service inputs are transformed into outputs. This means that a transit agency invests capital in vehicles, fuel, information systems, employees, maintenance, and other costs (service inputs). This investment produces a certain service for a community such as vehicle-km, seat-km, and seat-hours (service outputs). An agency is considered efficient if it can reduce the inputs to produce a fixed amount of outputs or increase the outputs while using similar or fewer inputs.

*Operational effectiveness* indicates the relationship between service inputs and consumed service. A transit agency spends money to offer its service, and a number of passengers (per day or week) consume its service. The transit agency will achieve higher operational effectiveness, if it increases ridership without increasing total cost of producing the services.

*Service effectiveness* examines the relationship between produced outputs and consumed service or how well a service offered by operators is consumed by a community [2]. This means that not all of the services offered (measured by vehicle-km, seat-km, and/or seat-hours) would be used by a community. If it attracts more passengers without increasing service or reduces service but still serves a similar number of passengers, it will be more effective.

## 2.2. Bus performance measurement

There are three main approaches to measure the performance of the bus system:

- Comparative Analysis (CA);
- Stochastic Frontier Analysis (SFA); and
- Data Envelopment Analysis (DEA)

The early approach applied for bus performance measurement is known as comparative analysis. This approach normally uses different key performance indicators (KPIs) to compare the performance of different bus systems with regard to different performance concepts, such as labour efficiency, vehicle efficiency, fuel efficiency, operating safety, and service consumption per expense. KPIs are defined as ratios of bus service outputs to service inputs (revenue vehicle hours per operating expense or passenger trips per revenue vehicle hour). Fielding et al. [7] defined a wide range of KPIs for comparing the performance of bus systems. Vuchic [8] provided efficiency ratios (output quantity produced per resource quantity expended) and utilisation (a ratio of demand to supply) to measure the performance of a transit system. The Transit Cooperative Research Program Report 88 [9] provided a process for developing a performance-measurement program, including both traditional and non-traditional performance indicators.

The CA approach is easy to apply for comparing the performance of bus at the route and system levels, but for a particular performance concept/indicator. The comparison, implemented for each KPI separately, leads to different levels of efficiency of one bus system for different KPIs. This approach, therefore, cannot provide a single overall measure of bus performance [5].

The latter two approaches, SFA and DEA, are frontier methods, which build up the frontier production function for evaluating the efficiency level of a set of production units with multiple inputs and outputs. SFA (a parametric approach introduced independently by Aigner et al. [10] and Meeusen and van Den Broeck [11]) uses econometric techniques, while DEA (a non-parametric approach) employs mathematical programming techniques for the frontier production function estimation. The advantage of the DEA approach is that it does not require a functional form to estimate the frontier production function. Thus, the DEA approach was widely used by researchers in transit sector in general and for bus performance measurement in particular.

## 2.3. Application of the DEA for bus performance evaluation

The application of DEA models in measuring the bus performance can be divided into two levels: (1) system; and (2) route level. At the system level, different bus systems within an area or in different nations are compared with each other, while at the route level bus routes within a system would be compared to identify the best practices (benchmarks) and inefficient routes. Comparing the performance of different bus systems plays a key role in determining the average operational efficiency of a transit system and problems related to the operation of the whole system, but cannot explore the problems related to the internal activities of each bus route. On the other hand, the performance evaluation of individual bus routes within a system substantially provides bus agencies with opportunity to understand its internal activities [6, 12], and then investigate the internal sources of inefficiency.

Chu and Fielding et al. [5] were pioneers in applying DEA models to measure the efficiency and effectiveness of public transit agencies in the United States (USA). The output data for efficiency and effectiveness assessment were annual revenue vehicle hours and annual unlinked passenger trips respectively. Based on the results of analysis, the authors reinforced the notion of Hatry [13] that in public agencies, efficiency should be evaluated separately from effectiveness.

Regarding the existing DEA literature on the field, most studies compare the performance of different bus systems (bus agencies) [5, 14-19], and a few studies focus on the performance of bus routes within a system. Sheth et al. [3] expanded the network DEA model of Färe and Grosskopf [20] to assess the performance of 60 different bus routes within a transit network in Virginia, USA. In this study, all variables related to the service provider, the users, and the community were used to compute the DEA efficiency scores. Results obtained help to rank the performance of these 60 bus routes and capture the relationship among the provider, the users, and the external and environmental variables related to the urban transit performance. Barnum et al. [6] employed the DEA model to analyse 46 bus routes of a US transit agency using weekday data. In the first stage, raw efficiency scores of individual bus routes were computed by a DEA model without considering the environmental variables. Then in the second stage, two environmental variables (population density, population), that are beyond the control of the transit agency, were used to adjust the DEA outputs (Riders and OTP). Then the adjusted DEA efficiency scores of DMUs are calculated. The results indicated that after adjusting the raw DEA scores, 20 bus routes became more efficient, 12 did not change, and 14 became less efficient. Lao et al. [1] combined the DEA model and geographic information system (GIS) to measure the performance of bus lines in a transit system. In this study, GIS was used to generate the input data for the spatial effectiveness DEA model and visualise the distribution of bus stops and routes. On the basis of operational efficiency and spatial effectiveness scores of 24 fixed bus routes, this research ranked the performance of individual bus routes and demonstrated that GIS can help to analyse the spatial variation of efficiency and effectiveness against demographic settings. More recently, 60 individual bus lines within a transit network in Thessaloniki, Greece were examined by a DEA model [2]. For model 1 and 2, input variables included trip length, span of service, and vehicles, while output variables were revenue seat-km for efficiency measure (model 1) and passengers for operational effectiveness assessment (model 2). Model 3 aimed at measuring combined effectiveness (revenue vehicle-km and vehicles are inputs and passengers is output). Along with calculating the efficiency and effectiveness scores for the three above models, this study also employed bootstrapping techniques to check robustness of DEA results for models 1 and 2. This sensitivity analysis explained that it is more reliable when correcting obtained scores for bias.

### **3. METHODOLOGY**

#### **3.1. Data Envelopment Analysis (DEA) model**

Data envelopment analysis (DEA) was developed by Charnes, Cooper, and Rhodes (CCR) in 1978 [21] and later modified by Banker, Charnes and Cooper (BCC) in 1984 [22]. It builds upon the frontier efficiency concept first elucidated in Farrell [23]. DEA is a non-parametric and empirical modelling based on linear programming and optimization. It is used widely to measure relative efficiencies of production units (termed as Decision making units, DMUs) with multi-inputs and multi-outputs.

The modelling process of DEA includes: a) identification of the production frontier (or isoquant) of a set of comparable DMUs. Within a set of comparable DMUs, those exhibiting the best use of inputs to produce outputs are identified, and would form an efficient frontier; and b) measures the efficiency level of each DMU by comparing its production function with the production frontier [24].

The CCR model measures efficiency of a DMU relative to a reference technology exhibiting constant returns to scale (CRS) whereas the BCC model exhibits variable (increasing, constant, or decreasing) returns to scale (VRS) at different points on the production frontier. Regarding bus performance, due to the constraint of capacity (for instance bus station capacity) and operating vehicle speed (because of schedule travel time), the output (passengers) might not have a constant increase when increasing the inputs (bus size, service frequency etc.). Therefore, the constant return to scale is not always existent. This article, thus, employs BCC-DEA model for empirical analysis.

### 3.2. BCC-DEA model

Suppose that each DMU<sub>j</sub> ( $j=1 \dots n$ ) uses  $m$  inputs  $x_{ij}$  ( $i=1 \dots m$ ) to generate  $s$  outputs  $y_{rj}$  ( $r=1 \dots s$ ), and the  $v_i, u_r$  are the variable weights of inputs and outputs, respectively.

This method uses the known inputs and outputs of all DMUs in the given set of data to determine the efficiency of one member DMU<sub>j</sub> ( $j=1 \dots n$ ), which is assigned as DMU<sub>0</sub>. The efficiency of DMU<sub>0</sub> is obtained by solving the following fractional programming problem  $n$  times, each DMU once.

$$\max h_0 = \frac{\sum_{r=1}^s u_r y_{r0} - u_0}{\sum_{i=1}^m v_i x_{i0}} \tag{1}$$

$$\begin{aligned} \text{Subject to: } & \frac{\sum_{r=1}^s u_r y_{rj} - u_0}{\sum_{i=1}^m v_i x_{ij}} \leq 1; & j = 1, \dots, n \\ & u_r, v_i \geq \varepsilon > 0; & r = 1, \dots, s; \quad i = 1, \dots, m. \quad u_0 \text{ free in sign} \end{aligned}$$

Where  $\varepsilon$  is a “non-Archimedean infinitesimal”, which is smaller than any positive real number such that all variables are constrained to positive values.

The objective is to obtain the input and output weights  $v_i, u_r$  as variables that maximize the ratio of DMU<sub>0</sub>, the DMU being evaluated. The value of  $h_0$  obtained from this formulation represents the efficiency score of DMU<sub>0</sub>. The constraints mean that  $h_0^*$ , being the optimal value of  $h_0$ , should not exceed 1 for all DMUs. In the case  $h_0^*=1$ , this DMU is situated on the efficiency frontier [25].

To solve this problem, the theory of Charnes *et al.* [26] is applied to convert this fractional programming problem to the linear programming (LP) model with the changes of variables  $t(\sum_{i=1}^m v_i x_{i0}) = 1; \mu_r = tu_r$  and  $\vartheta_i = tv_i$ . The above problem is replaced by the following equivalent:

$$\max h_0 = \sum_{r=1}^s \mu_r y_{r0} - \mu_0 \tag{2}$$

$$\begin{aligned} \text{Subject to: } & \sum_{i=1}^m \vartheta_i x_{i0} = 1 \\ & \sum_{r=1}^s \mu_r y_{rj} - \mu_0 - \sum_{i=1}^m \vartheta_i x_{ij} \leq 0 & j = 1, \dots, n \\ & \mu_r, \vartheta_i \geq \varepsilon > 0; & r = 1, \dots, s; \quad i = 1, \dots, m \quad \mu_0 \text{ free} \end{aligned}$$

In the case of output-oriented model, the dual problem can be expressed as follows:

$$\begin{aligned} \max \quad & \varphi - \varepsilon(\sum_{r=1}^s s_r^+ + \sum_{i=1}^m s_i^-) \tag{3} \\ \text{Subject to:} \quad & \sum_{j=1}^n \lambda_j x_{ij} + s_i^- = x_{i0} \quad i = 1, \dots, m \\ & \sum_{j=1}^n \lambda_j y_{rj} - s_r^+ = \varphi y_{r0} \quad r = 1, \dots, s; \\ & \sum_{j=1}^n \lambda_j = 1 \quad \lambda_j, s_i^+, s_i^- \geq 0, \text{ all } r, i, j \quad \varphi \text{ free} \end{aligned}$$

Where:  $(s_i^+, s_i^-)$  are the output and input slack variables. Input slack is the amount of input that one DMU could reduce to produce the same output.  $\varphi$  is the distance parameter in the output-oriented DEA model. The DMU efficiency is measured by  $1/\varphi$ .

#### 4. DATA SET AND EMPIRICAL ANALYSIS

##### 4.1. Data set

This article uses a sample of 38 bus routes in Hanoi city for empirical analysis. These bus routes include both mini bus routes (30 spaces) and medium bus routes (60 to 80 spaces). The given bus routes are shown in Table 1. Data set used in this paper is the operation of these routes during the year 2018, which is collected from Hanoi Transport Department and the website of Transerco.

Table 1. List of 38 bus routes within the data sample.

No	Bus Route	Start point - destination	No	Bus Route	Start point - destination
1	01	Gia Lam Station - Yen Nghia Station	20	47B	DHKTQD - Kieu Ky
2	02	Bac Co - Yen Nghia Station	21	48	Savico Long Bien - Nuoc Ngam Station
3	03A	Giap Bat Station - Gia Lam Station	22	07	Cau Giay - Noi Bai
4	13	Ho Tay Park - Co Nhue	23	27	Yen Nghia Station - Nam Thang Long
5	14	Bo Ho - Co Nhue	24	34	My Dinh Station - Gia Lam
6	18	DH KTQD - Long Bien - DHKTQD	25	35A	Tran Khanh Du - Nam Thang Long
7	20A	Cau Giay - Phung Station	26	55A	Times City - Bui - Cau Giay
8	22A	Gia Lam Station - Big C Thang Long	27	109	My Dinh Station - Noi Bai
9	23	Nguyen Cong Tru - Nguyen Cong Tru	28	42	Giap Bat Station - Duc Giang
10	26	Mai Dong - National Stadium	29	45	Times City - Nam Thang Long
11	31	Bach Khoa - Chem	30	49	Tran Khanh Du - My Dinh II
12	32	Giap Bat Station - Nhon	31	51	Tran Khanh Du - Cau Giay Park
13	33	Yen Nghia Station - Xuan Dinh	32	60A	Phap Van - Ho Tay Park
14	50	Long Bien - National Stadium	33	96	Nghia Do Park - Dong Anh
15	BRT01	Yen Nghia Station - Kim Ma	34	98	Yen Phu - Aeon mall Long Bien

<b>16</b>	84	My Dinh I - Linh Dam	<b>35</b>	99	Kim Ma - BVNT TU II
<b>17</b>	85	Nghia Do Park - Van Phu	<b>36</b>	104	My Dinh - Linh Dam
<b>18</b>	90	Kim Ma Station - Nhat Tan Bridge - Noi Bai Airport	<b>37</b>	105	Do Nghia - Cau Giay
<b>19</b>	08B	Long Bien - Van Phuc	<b>38</b>	106	Mo Lao - Aeon mall Long Bien

Table 2 shows the statistical description of the input and output variables of the sample for the year 2018. The variables are defined as follows:

*Route length (km)*: length of roadways from start point to destination.

*Number of stops (stop)*: the total number of bus stops along the route for one way.

*Total trips (trip)*: total number of bus trips performed on the route during the year 2018.

*Vehicles (vehicle)*: total number of bus vehicles used on the route.

*Space-km (spaces-km)*: bus vehicle capacity multiplied by total distance traversed by all vehicles on the corresponding route during a year (2018).

*Passengers*: total number of passenger trips performed on the route

Table 2. Statistical description of the inputs and outputs of the 38 bus routes.

<b>Variables</b>	<b>Input/output</b>	<b>Mean</b>	<b>Minimum</b>	<b>Maximum</b>	<b>Standard deviation</b>
<b>Route length (km)</b>	Input	19.57	13.8	31.5	4.73
<b>Number of stops (stop)</b>	Input	31.82	20	42	5.83
<b>Total trips (trip)</b>	Input	53826.24	7008	126928	28923.02
<b>Vehicles</b>	Input	11.53	6	28	6
<b>Space-km</b>	Output	66255245.6	11373984	204833205	50311570.72
<b>Passengers</b>	Output	3900952.5	300248	19164025	4054286.16

## 4.2. Model specification

In this article, the *technical efficiency* and *operational effectiveness* of given bus routes are examined on the basis of maximising the outputs. Thus, the output-oriented BCC-DEA model is adopted for empirical analysis. A DMU is defined as the performance of each bus route during the year 2018. Table 3 presents the specification of models applied and the corresponding inputs and outputs. Here, models 1 and 2 measure the *technical efficiency* and *operational effectiveness* of bus routes, respectively.

Table 3. Models and analysis framework.

<b>Model</b>	<b>Performance dimension</b>	<b>Orientation</b>	<b>Returns to scale</b>	<b>Input variables</b>	<b>Output variables</b>
<b>Model 1</b>	Technical efficiency	Output	VRS	Route length, Number of stops, Total trips, Vehicles	Space-km
<b>Model 2</b>	Operational effectiveness	Output	VRS	Route length, Number of stops, Total trips, Vehicles	Passengers

*Technical efficiency*: the output variables should present service outputs offered by the bus operator. Here, we select *space-km* because it represents the bus capacity offered by the operators. The inputs should present the resources used by bus operator to generate the service outputs. Based on the existing literature, this article uses *route length*, *number of stops*, *total trips*, and *vehicles* as inputs relevant to *space-km*. *Total trips* refer to the number of vehicles and drivers used, *vehicles*, *route length*, and *number of stops* introduce the operation and maintenance resources.

*Operational effectiveness*: the outputs should represent the service consumption, so *passengers* is selected as output. Inputs for this measure are similar to *technical efficiency*.

### 4.3. Results and discussion

The results obtained from the efficiency analysis of the aforementioned models (model 1 for technical efficiency and model 2 for operational effectiveness) are shown in **Fig. 1**. The score axis illustrates the efficiency scores of DMUs. A DMU is efficient if its score equals to 1, whereas lower score indicates that it is inefficient. In the DEA models, efficient DMUs become benchmarks for other inefficient/ineffective DMUs in the given sample. For instance, considering route 51 in model 1, its score of 0.8 indicates that it is possible to increase the outputs by 25% ( $= \frac{1-0.8}{0.8}$ ) using the similar inputs. Its benchmarks are routes 20A ( $\lambda_{20A} = 0.539$ ), 49 ( $\lambda_{49} = 0.336$ ), and BRT01 ( $\lambda_{BRT01} = 0.124$ ). The combination of 53.9%, 33.6%, and 12.4% inputs and outputs of routes 20A, 49, and BTR01, respectively can build up the virtual DMU of route 51, which locates on the production frontier.

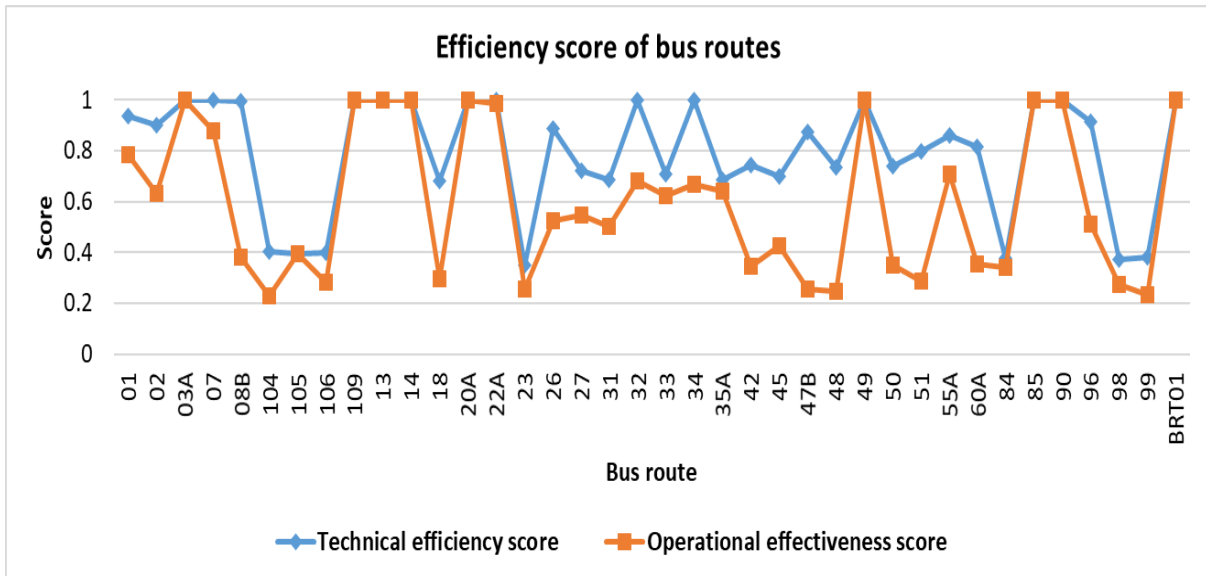


Figure 1. Efficiency scores of bus routes for model 1 and model 2.

Table 4 represents the summary statistics of the results obtained from the two models. It could be noted that the average efficiency score in model 1 is remarkably higher than those in model 2 (0.79 compared with 0.6), suggesting that bus routes considered have better performance in terms of technical efficiency. Additionally, both models witness a wide dispersion of efficiency scores because some bus routes (such as routes 104, 105, 106, 23, 98, and 99) have efficiency scores lower than 0.4.



Table 4. Efficiency scores statistics obtained for the two models.

Model	Mean	Minimum	Maximum	Standard deviation	Percentage of DMUs with score		
					< 0.5	0.5 – 0.8	0.8 - 1
Model 1	0.79	0.35	1	0.22	16.2%	27%	56.8%
Model 2	0.60	0.23	1	0.29	40.5%	38.1%	29.8%

Table 5. Slacks for several inefficient routes in models 1 and 2.

DMU	Model 1			Model 2		
	Efficiency score	Number of stops	Vehicles	Efficiency score	Number of stops	Vehicles
23	0.35	6.19	2.67	0.26	6.66	2.90
31	0.69	9.13	1.94	0.50	10.99	1.64
35A	0.68	5.75	1.71	0.64	6.99	0.83
45	0.70	4.99	1.26	0.42	5.61	1.05
84	0.38	6.87	0.79	0.34	7.52	0.33
98	0.37	0.91	0.14	0.28	1.46	0
99	0.38	6	0	0.23	6	0

Table 6. The ranking of bus routes for operational effectiveness (model 2).

DMU	Ranking	Efficiency score	DMU	Ranking	Efficiency score
03A; 13; 14; 20A; 49; 85; 90; 109; and BRT01	1	1	105	16	0.396
22A	2	0.99	08B	17	0.38
07	3	0.88	60A	18	0.36
01	4	0.79	50	19	0.35
55A	5	0.71	42	20	0.34
32	6	0.68	84	21	0.34
34	7	0.67	18	22	0.3
35A	8	0.64	51	23	0.29
02	9	0.63	106	24	0.28
33	10	0.62	98	25	0.28
27	11	0.55	23	26	0.26
26	12	0.52	47B	27	0.26
96	13	0.51	48	28	0.25
31	14	0.50	99	29	0.23
45	15	0.42	104	30	0.23

Model 1: **Fig. 1** shows that there are 13 efficient DMUs, including routes 03A, 07, 13, 14, 20A, 22A, 32, 34, 49, 85, 90, 109, and BTR01. Furthermore, there are 7 routes with poor performance (score <0.5), consisting of routes 104, 105, 106, 23, 84, 98, and 99. The remaining bus routes have fairly good performance regarding the technical efficiency.

Model 2: there are 9 efficient DMUs, including routes 03A, 13, 14, 20A, 49, 85, 90, 109, and BRT01 (the benchmarks of the sample). It is notable that there are 40.5% bus routes with poor performance (score <0.5) and 38.1% bus routes with fairly good performance (score from 0.5 to 0.8) (see Table 4). The least efficient bus routes (score < 0.3) are 18, 23, 47B, 48,

51, 98, 99, 104 and 106, which need further performance improvement. It can be observed from the results that bus routes with good performance mainly operate within the city centre (13, 14, and 85) or connect main stations (03A, 90, 109, and BRT01), while the least efficient routes mainly connect the city centre with suburban areas (47B, 98, 99, and 106). The ranking of bus routes regarding the operational effectiveness is illustrated in Table 6.

Table 5 illustrates the slacks obtained from both models 1 and 2 for several poor performance bus routes (input slack is the amount of input that one DMU could reduce to produce the same output). The results show that slacks mostly occur for *number of stops* and *vehicles*. Thus, reducing the *number of stops* and/or *vehicles* used can be one of the possible solutions to improve performance of inefficient routes. For instance, routes 23 and 31, in model 1, can reduce the number of vehicles by 2.67 and 1.94 units, respectively.

## 5. CONCLUSION

This article employs the output-oriented BCC-DEA model to provide insights into the *technical efficiency* (model 1) and *operational effectiveness* (model 2) of 38 key bus routes within the bus network in Hanoi, Vietnam. The results achieved indicate the best and the worst bus routes within the data sample. It is noted that routes 03A, 13, 14, 20A, 49, 85, 90, 109, and BRT01 become the benchmark of the sample for both technical efficiency and operational effectiveness measure. Routes 18, 23, 47B, 48, 51, 98, 99, 104 and 106, having the poorest performance in model 2, need further investigations to understand the key reasons of inefficiency, and then make appropriate decisions for performance improvement.

The empirical analysis also explains to some extent the source of inefficiency of bus route performance, including the *number of stops* and *vehicles*. This indicates the considerably low stop spacing and the excessive use of number of vehicles on some inefficient bus routes. Reduction of these resources could be a solution to optimise the performance of these bus routes. The knowledge gained helps to provide bus operators and policy makers with additional information for decision makings.

This article only uses the yearly data to evaluate the performance of 38 bus routes in Hanoi. Future studies should use a larger sample and more detailed timeframes (weekday or monthly data) for empirical analysis to obtain the more comprehensive results. Another limitation is that we do not investigate the influence of environmental factors (socio-economic factors) on the efficiency score of DMUs. This work will be performed in upcoming studies.

## ACKNOWLEDGMENT

The authors wish to sincerely thank Hanoi Transport Department and Transerco of Hanoi, which have supplied the relevant data of bus system in Hanoi, Vietnam.

## REFERENCES

- [1]. Y. Lao, L. Liu, Performance evaluation of bus lines with data envelopment analysis and geographic information systems. *Computers, Environment and Urban Systems*, 33 (2009) 247-255. <https://doi.org/10.1016/j.compenvurbsys.2009.01.005>
- [2]. G. Georgiadis, I. Politis, P. Papaioannou, Measuring and improving the efficiency and effectiveness of bus public transport systems, *Research in Transportation Economics*, 48 (2014) 84-91. <http://dx.doi.org/10.1016/j.retrec.2014.09.035>

- [3]. C. Sheth, K. Triantis, D. Teodorović, Performance evaluation of bus routes: A provider and passenger perspective, *Transportation Research Part E: Logistics and Transportation Review*, 43 (2007) 453-478. <https://doi.org/10.1016/j.tre.2005.09.010>
- [4]. G.J. Fielding, M.E. Brenner, K. Faust, Typology for bus transit, *Transportation Research Part A: General*, 19 (1985) 269-278. [https://doi.org/10.1016/0191-2607\(85\)90015-9](https://doi.org/10.1016/0191-2607(85)90015-9)
- [5]. X. Chu, G.J. Fielding, B.W. Lamar, Measuring transit performance using data envelopment analysis, *Transportation Research Part A: Policy and Practice*, 26 (1992) 223-230. [https://doi.org/10.1016/0965-8564\(92\)90033-4](https://doi.org/10.1016/0965-8564(92)90033-4)
- [6]. D.T. Barnum, S. Tandon, S. McNeil, Comparing the performance of bus routes after adjusting for the environment using data envelopment analysis, *Journal of Transportation Engineering*, 134 (2008) 77-85. [https://doi.org/10.1061/\(ASCE\)0733-947X\(2008\)134:2\(77\)](https://doi.org/10.1061/(ASCE)0733-947X(2008)134:2(77))
- [7]. G.J. Fielding, T.T. Babitsky, M.E. Brenner, Performance evaluation for bus transit, *Transportation Research Part A: General*, 19 (1985) 73-82. [https://doi.org/10.1016/0191-2607\(85\)90009-3](https://doi.org/10.1016/0191-2607(85)90009-3)
- [8]. V.R. Vuchic, *Urban transit systems and technology*, Hoboken, N.J: John Wiley & Sons, 2007.
- [9]. P. Ryus, A Summary of TCRP Report 88: A Guidebook for Developing a Transit Performance-Measurement System. TCRP Research Results Digest, 2003(56). [http://onlinepubs.trb.org/onlinepubs/tcrp/tcrp\\_report\\_88/SummaryDoc.pdf](http://onlinepubs.trb.org/onlinepubs/tcrp/tcrp_report_88/SummaryDoc.pdf)
- [10]. D. Aigner, C.K. Lovell, P. Schmidt, Formulation and estimation of stochastic frontier production function models, *Journal of Econometrics*, 6 (1977) 21-37. <http://people.stern.nyu.edu/wgreene/FrontierModeling/Reference-Papers/Aigner-Lovell-Schmidt-JE1977-ALS.pdf>
- [11]. W. Meeusen, J. van Den Broeck, Efficiency estimation from Cobb-Douglas production functions with composed error, *International economic review*, 18 (1977) 435-444. <https://www.jstor.org/stable/2525757>
- [12]. H.P. Benn, Bus route evaluation standards, 1995. <http://onlinepubs.trb.org/onlinepubs/tcrp/tsyn10.pdf>
- [13]. H.P. Hatry, Performance measurement principles and techniques: An overview for local government, *Public Productivity Review*, 4 (1980) 312-339. <https://www.jstor.org/stable/3379974>
- [14]. K. Obeng, The economic cost of subsidy-induced technical inefficiency. *International Journal of Transport Economics/Rivista internazionale di economia dei trasporti*, 21 (1994) 3-20. <https://www.jstor.org/stable/42747468>
- [15]. K. Kerstens, Technical efficiency measurement and explanation of French urban transit companies. *Transportation Research Part A: Policy and Practice*, 30 (1996) 431-452. [https://doi.org/10.1016/0965-8564\(96\)00006-7](https://doi.org/10.1016/0965-8564(96)00006-7)
- [16]. P.A. Viton, Technical efficiency in multi-mode bus transit: A production frontier analysis. *Transportation Research Part B: Methodological*, 31 (1997) 23-39. [https://doi.org/10.1016/S0191-2615\(96\)00019-7](https://doi.org/10.1016/S0191-2615(96)00019-7)
- [17]. P.A. Viton, Changes in multi-mode bus transit efficiency, 1988–1992, *Transportation*, 25 (1998) 1-21. <https://doi.org/10.1023/A:1004906024774>
- [18]. D.A. Tsamboulas, Assessing performance under regulatory evolution: A European transit system perspective. *Journal of Urban Planning and Development*, 132 (2006) 226-234. [https://doi.org/10.1061/\(ASCE\)0733-9488\(2006\)132:4\(226\)](https://doi.org/10.1061/(ASCE)0733-9488(2006)132:4(226))
- [19]. M. Arman, S. Labi, K. Sinha, Perspectives of the Operational Performance of Public Transportation Agencies with Data Envelopment Analysis Technique, *Transportation Research Record: Journal of the Transportation Research Board*, 2351(2013) 30-37. <https://doi.org/10.3141%2F2351-04>
- [20]. R. Fare, S. Grosskopf, Network DEA, *Socio-economic planning sciences*, 34 (2000) 35-49. [https://doi.org/10.1016/S0038-0121\(99\)00012-9](https://doi.org/10.1016/S0038-0121(99)00012-9)
- [21]. A. Charnes, W.W. Cooper, E. Rhodes, Measuring the efficiency of decision making units, *European Journal of Operational Research*, 2 (1978) 429-444. [http://dx.doi.org/10.1016/0377-2217\(78\)90138-8](http://dx.doi.org/10.1016/0377-2217(78)90138-8)

- [22].R.D. Banker, A. Charnes, W.W. Cooper, Some models for estimating technical and scale inefficiencies in data envelopment analysis, *Management science*, 30 (1984) 1078-1092. <https://doi.org/10.1287/mnsc.30.9.1078>
- [23].M.J. Farrell, The Measurement of Productive Efficiency. *Journal of the Royal Statistical Society. Series A (General)*, 120 (1957) 253-290. <https://doi.org/10.2307/2343100>
- [24].W.D. Cook, L.M. Seiford, Data envelopment analysis (DEA) – Thirty years on. *European Journal of Operational Research*, 192 (2009) 1-17. <http://dx.doi.org/10.1016/j.ejor.2008.01.032>
- [25].K. Tone, W.W. Cooper, L.M. Seiford, *Data Envelopment Analysis: A Comprehensive Text with Models, Applications, and DEA-Solver Software*, Kluwer Academic Publishers, 1999.
- [26].A.Charnes, W.W. Cooper, Programming with linear fractional functionals, *Naval Research logistics quarterly*, 9 (1962) 181-186. <https://doi.org/10.1002/nav.3800090303>