Efficiency as a Criterion for Optimizing Routes: A two-stage DEA routing approach

Paulo Nocera Alves Junior, University of São Paulo (USP/ESALQ), pjnocera@yahoo.com.br
Daniela Bacchi Bartholomeu, University of São Paulo (USP/ESALQ), daniela.bartholomeu@usp.br
José César Cruz Júnior, Federal University of São Carlos (UFSCAR), cesarcruz@ufscar.br
José Vicente Caixeta Filho, University of São Paulo (USP/ESALQ), jose.caixeta@usp.br


Copyright 2019 by Paulo Nocera Alves Junior, Daniela Bacchi Bartholomeu, José César Cruz Júnior, and José Vicente Caixeta Filho. All rights reserved. Readers may make verbatim copies of this document for non-commercial purposes by any means, provided that this copyright notice appears on all such copies.
ABSTRACT

The purpose of this paper is to measure the efficiency of routes, using a data envelopment analysis (DEA) model, and then use the efficiency scores as criteria for routing within a second stage. The novelty of this paper is twofold: it uses decision making units (DMU) efficiencies (including non-efficient DMUs) and it allows a vehicle routing model to select the route of minimal inverted efficiency. An empirical application was built considering routes used by producers with a final destination of São Paulo. The results showed that routing based on efficiency as a criterion leads to better results, as producers can deliver their product faster, consume less diesel, and emit less CO₂ into the environment. However, as additional travel time nears zero, the results become similar to those when using distance only.

Keywords: data envelopment analysis (DEA), vehicle routing problems (VRP), time in traffic, idle time, CO₂ emissions.

1. INTRODUCTION

The World Bank report, Best Practices in Management of International Trade Corridors (Arnold, 2006) provides a descriptive definition of transportation corridors. The most relevant physical elements of a transportation corridor may include one or more routes that connect economic centers. The most relevant functional element of a transportation corridor is that it is usually developed to support regional economic growth.

According to the Swedish Logistics Forum, the definition of Green Corridors is as follows: “Green Corridors aim at reducing environmental and climate impact while increasing safety and efficiency” (Tetraplan, 2011, apud Psaraftis and Panagakos, 2012). An important aspect of Green Corridors concepts is that they are more than just one aspect alone; they are both economically efficient and environmentally sustainable (Panagakos et al., 2016).
Measuring the economic and environmental efficiency is the challenge of the concept envolving green corridors or routes. EU Super Green Project aimed at advancing the green corridor concept through a benchmarking exercise involving Key Performance Indicators (KPIs) of sixteen European corridors (Panagakos, 2016). Other methods for benchmarking corridors and routes are related by Kengpol et al. (2014), Oliveira and Cicolin (2016) and Melo et. al (2018).

In general, the authors use Data Envelopment Analysis (DEA) combined with Analytic Hierarchy Process (AHP) or Balanced Score Card (BSC) to measure the efficiency of routes or to select efficient routes or corridors (Kengpol et al., 2014; Oliveira and Cicolin, 2016). Meanwhile, Melo et. al (2018) developed a DEA slack-based measure model for benchmarking freight transportation corridors and routes (exportation). Another alternative way is applying DEA models to measure the efficiency of routes considering Vehicle Routing Problems (VRP). The main applications of DEA in routing are related to initially generating feasible routes using heuristics for VRP, then applying DEA to select efficient routes (Kritikos and Ioannou, 2010; Juan, Huapu, Xu, Xianfeng and Huijun, 2014; and Chen, Achtari, Majkut and Sheu, 2017).

In this context, taking in consideration the concept of “green transport” (corridor or route), this paper aims to propose a different methodology to measure the efficiency of routes in a two-stage DEA routing model, using a data envelopment analysis (DEA) model to calculate the efficiencies of part of the routes (“sub-routes”) and using the efficiencies of the ranked DMUs (sub-routes) as the criterion for routing in the vehicle routing problem (VRP) model.

The novelty of this paper is twofold. First, it uses DMU efficiencies (including non-efficient DMUs) and is based on travel or idle time, so fuel consumption and emissions are not correlated to distance only. Second, it allows a vehicle routing model to select the route of minimal inverted efficiency. While a classical VRP does the same, this paper uses the efficiencies as an aggregate measure (or a composite indicator), instead of a single measure such as distance, time, or emissions. These single measures can be considered limited when, evaluating a sustainable route, for example, and instead the efficiency of a route is more suitable for solving VRP.

The case analyzed in this study was an empirical application built considering routes for urban freight transportation with a final destination of São Paulo, the biggest city in Brazil. The origin point was set in a location that represents the geographical central point of a region known as the “Fruit Trail” in the state of São Paulo (Figure 1). According to Souza-Esquerdo and
Bergamsco (2014), the region is characterized by the strong presence of family farmers (over 60% of rural properties). Considering its proximity to the state capital (approximately 75 kilometers), the city of São Paulo is an important consumer market and one of the main destinations of the region’s production. Besides its geographical location, the town of Louveira was chosen as the origin in our model because, according to Souza-Esquerdo and Bergamasco (2014), it was also the town with the largest participation in public policy programs in the region, such as the National Program for Strengthening of Family Farming (PRONAF). The most important fruits produced in the region are grape, strawberry, peach, guavas, plum, persimmon, acerola cherry, and fig. In addition to farming, the “Fruit Trail” is a popular rural tourism destination in the state, mainly because of its proximity to the capital state.

We expect that our model can helpful not only for family producers in the region, but also for any other business that have freight as one of their main costs.

Figure 1: São Paulo state “Fruit Trail” location map

Source: Souza-Esquerdo; Bergamasco (2014)
2. RELATED LITERATURE

2.1 Data Envelopment Analysis – Slack-Based Measure Model

Data envelopment analysis (DEA) is a nonparametric mathematical programming method used to measure the relative efficiency of decision-making units (DMUs) in a system with multiple criteria. In 1978, the first DEA model was developed by Charnes, Cooper and Rhodes (1978) (CCR).

Problems such as freight transportation may require simultaneous output maximization and input minimization, and the slack-based measure (SBM) model developed by Tone (2001) should be most suitable for such problems. The linear programming formulation of the SBM model is as follows (Equations (1) to (5)):

Minimize \( \tau = t - (1/m) \sum_{i=1}^{m} S_i^- / x_{i0} \) \hspace{1cm} (1)

Subject to:

\( (1/n) \sum_{j=1}^{n} S_j^+ / y_{j0} + t = 1 \) \hspace{1cm} (2)
\( \sum_{k=1}^{z} \Lambda_k x_{ik} + S_i^- - tx_{i0} = 0 \quad i = 1, 2, \ldots, m \) \hspace{1cm} (3)
\( \sum_{k=1}^{z} \Lambda_k y_{jk} - S_j^+ - ty_{j0} = 0 \quad j = 1, 2, \ldots, n \) \hspace{1cm} (4)
\( \Lambda_k \geq 0, S_i^- \geq 0, S_j^+ \geq 0 \) and \( t > 0 \) \hspace{1cm} (5)

Where \( \tau \) is efficiency, \( S_i^- \) is the slack of the \( i \)th input, \( S_j^+ \) is the slack of the \( j \)th output, \( \Lambda_k \) is the contribution of the \( k \)th DMU to the DMU under analysis, \( t \) is the model linearization factor, \( x_{i0} \) is the \( i \)th input of the DMU under analysis, \( y_{j0} \) is the \( j \)th output of the DMU under analysis. \( x_{ik} \) is the \( i \)th input of the \( k \)th DMU, \( y_{jk} \) is the \( j \)th output of the \( k \)th DMU, \( m \) is the number of inputs, \( n \) is the number of outputs, and \( z \) is the number of DMUs.

2.2 VRP

A vehicle routing problem (VRP) is an optimization problem and the model is used to obtain the minimum distance (or other desired measure such as emissions, travel time, etc.) of a route in a system with at least one origin and one destination. In 1959, the first VRP model for a truck dispatching problem was developed (Dantzig, Ramser, 1959). The model we use in this paper
is similar to the one presented in Laporte (1992), but with subtour elimination of length 1 and 2, and with extra deliveries. It is as follows, in Equations (6) to (13):

\[
\text{Minimize } \sum_{i=1}^{n} \sum_{j=1}^{n} d_{ij} x_{ij} \tag{6}
\]

Subject to:

\[
\sum_{i=1}^{n} x_{ij} - \sum_{i=1}^{n} x_{ji} = -1, \quad j = 1 \tag{7}
\]

\[
\sum_{i=1}^{n} x_{ij} - \sum_{i=1}^{n} x_{ji} = 0, \quad j = 2, \ldots, n - 1 \tag{8}
\]

\[
\sum_{i=1}^{n} x_{ij} - \sum_{i=1}^{n} x_{ji} = 1, \quad j = n \tag{9}
\]

\[
\sum_{i=1}^{n} x_{ij} + \sum_{i=1}^{n} x_{ji} \geq 2, \quad j = m_1, \ldots, m_n \tag{10}
\]

\[
x_{jj} = 0, \quad j = 1, \ldots, n \tag{11}
\]

\[
x_{ij} + x_{ji} = 1, \quad \{i=1, \ldots, n \}, \{j=1, \ldots, n \} \tag{12}
\]

\[
x_{ij} \in \{0,1\} \tag{13}
\]

Where: \(d_{ij}\) is the distance (which can be changed by another measure to be minimized) from origin \(i\) to destination \(j\), \(x_{ij}\) is the binary variable that indicates the activation of the route from origin \(i\) to destination \(j\), \(x_{ji}\) is the binary variable that indicates the activation of the route from origin \(j\) to destination \(i\), \(n\) is the number of nodes (origins and destinations), and \(m_1\) to \(m_n\) are delivery nodes (destinations of deliveries).

The objective function of the model used in the proposed method for the DEA routing approach is based on the efficiency, so it is as follows in Equation (14):

\[
\text{Minimize } \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{x_{ij}}{\tau_{ij}} \tag{14}
\]

Where \(\tau_{ij}\) is the efficiency of the route (substituting the distance \(d_{ij}\) in Equation (6)). The other parameters and variables in Equation (14) are the same as those in Equation (6).
2.3 DEA and VRP

Only few studies have combined the DEA and the VRP methods. For instance, Kritikos and Ioannou (2010) applied the IMPACT heuristic (Ioannou et al., 2001) to generate feasible routes without cargo loading and used the entire routes as decision-making units (DMUs). They then implemented a free disposal hull (FDH)—an integer DEA model for creating an efficient frontier and route selection.

Similarly, Juan et al. (2014) established a path decision set based on an overlapping section punishment search algorithm. They then developed a DEA-based multi-objective fuzzy decision-making method for the route selection of military highway transport.

Chen et al. (2017) performed a regression analysis to generate a relative utility value, using the Dijkstra algorithm to first generate an initial route design, then a DEA-based heuristic procedure to improve each route. In the model, they used a rural accessibility score, a multi-objective accessibility score, and a multi-objective utility score. They use population as outputs, and distance and service time as inputs (related to cost of fuel, labor, vehicle depreciation, and other operational expenses).

Lu and Yu (2012) ran heuristics and metaheuristics based on a genetic algorithm to solve VR problems in three instances, collecting data from the best, average, and worst cases in terms of number of vehicles, total distance, time (early and late), and computation time. Then, they applied DEA to choose the best genetic algorithm for solving VRP.

Smirlis, Zeimpekis, and Kaimakamis (2012) evaluated a software that solves VRPs, considering criteria related to routing characteristics supported by the software such as problem size, functions, GIS capabilities, performance, features, and prices. Then, they applied DEA to select the efficient software.

Cagliano, Marco, Mangano, and Zenezini (2017) make considerations about the literature for applying DEA to evaluate the efficiency of logistics service providers (LSP) and variables associated with VRP—among other factors affecting the productivity of LSP—but they do not apply DEA, and use regression analysis instead.

Xing, et al. (2013) did a review of operations research (OR) in service industries, including DEA and VRP, but did not thoroughly discuss the topics in the paper.
3. METHOD

We used the following variables in our model: distance, travel time (including time in traffic or idle time), fuel consumption, CO₂ emissions, and efficiency. For the DEA model, we used a unitary output (where output=1), and the following inputs: distance, travel time, fuel consumption, and CO₂ emissions. For the VRP model, we used efficiency as a criterion for routing. The same data was used in all comparisons. We used Matlab for the calculations.

The traditional approach found in the literature consists of generating all feasible and desirable routes by a routing algorithm and then applying DEA to select the efficient routes. The steps proposed in this paper are shown in the next section.

3.1. Proposed DEA-routing approach

The following outlines the steps of the method:

1) Making a list of nodes (origins and destinations);
2) Making a list of DMUs (routes connecting each node);
3) Collecting data;
4) Running the DEA model;
5) Inverting efficiencies (1/efficiency);
6) Running the second stage (the routing model) using the efficiency of the routes as a criterion (measure) for the optimization.

This steps of the method suggested in this paper differs from the steps exposed in the the literature, because applying DEA before generating the routes reduces the number of DEA models created to evaluate the routes (it is used every pair of nodes, i.e. each “sub-route”, instead of applying in every whole possible desirable route).

We used Google Maps to collect data (step 3) for distance and travel time (including time in traffic). Therefore, we calculated idle time based on best and worst travel times shown by Google Maps. With respect to emissions, we assumed 2.667 of kg of CO₂ emissions per litre of
diesel, based on Ecotransit (2018). Regarding fuel consumption, we assumed 3.33 liters of diesel consumed per km, based on benchmark values for fuel performance from UNEP (2009) and freight information from urban vehicle sites. We assume four liters of diesel per hour of idle consumption. The following shows variable calculations:

- **Idle Time**: \( t_{idle} = t_{worst} - t_{best} \)
- **Consumption**: \( C = d_{ij} \cdot \frac{1}{3.33} + \Delta t \cdot 4/60 \)
- **Emissions**: \( E = 2.667 \cdot d_{ij} \cdot \frac{1}{3.33} + 2.667 \cdot \Delta t \cdot 4/60 \)

### 3.2. Traditional DEA approach

We generated routes considering eight possible nodes and respecting certain rules (Figure 2). Specifically, the first node is always the origin number 1, the second and third nodes could be the numbers 2 or 3, the fourth through sixth nodes could be any of the five remaining nodes, and the last two destinations could be numbers 7 or 8. Number 0 was used to indicate shorter routes without passing by a node. Then, we verified if all delivery points were included. In total, we generated 640 routes.

**Figure 2: Map of origin and delivery locations**

Note: origin = 1; 2-8 = destinations. Source: Google Maps
4. EMPIRICAL EXAMPLE

A distribution center (DC) sends a truck whose consumption is 30 liters of diesel per 100 km and 4 liters idle to make four deliveries (n° 2, 3, 7, and 8), where number 1 is the origin of DC and number 8 is the final destination. The truck passes through all 8 locations shown in Figure 2.

An example of two routes can be seen in Figures 3 and 4. One route includes only the delivery points in Figure 3, and the other route includes part of a north-south corridor (e.g. Av. Tiradentes and Av. 9 de Julho) in Figure 4.

**Figure 3: Route including only deliveries points**

![Figure 3: Route including only deliveries points](source: Google Maps)
5. RESULTS

5.1. Proposed approach

The results of the proposed approach (routing based on efficiency) can be seen in Table 1. We can observe in Table 1 that the optimal route based on efficiency is the route that passes through the following sequence of nodes: 1 → 3 → 2 → 7 → 8 (Figure 2).

Table 1: DEA routing results

<table>
<thead>
<tr>
<th>Variable (Route)</th>
<th>X13 (1 to 3)</th>
<th>X32 (3 to 2)</th>
<th>X27 (2 to 7)</th>
<th>X78 (7 to 8)</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Distance (km)</td>
<td>76.90</td>
<td>15.50</td>
<td>16.70</td>
<td>15.00</td>
<td>124.10</td>
</tr>
<tr>
<td>Time (minutes)</td>
<td>110.00</td>
<td>50.00</td>
<td>85.00</td>
<td>75.00</td>
<td>320.00</td>
</tr>
<tr>
<td>Consumption (liters)</td>
<td>26.07</td>
<td>6.39</td>
<td>8.34</td>
<td>7.50</td>
<td>48.30</td>
</tr>
<tr>
<td>Emissions (kg of CO₂)</td>
<td>69.53</td>
<td>17.02</td>
<td>22.25</td>
<td>20.00</td>
<td>128.80</td>
</tr>
<tr>
<td>1/efficiency</td>
<td>6.03</td>
<td>1.76</td>
<td>2.46</td>
<td>2.20</td>
<td>12.45</td>
</tr>
</tbody>
</table>
We then must question what the difference is between the efficient route and the shortest route. In addition, what is the difference between making the selection based on a traditional efficiency analysis (the traditional approach in the literature) and using efficiency as a criterion for routing?

For comparison, Table 2 shows the results of the shortest route, and Table 3 the differences of the results. In topic 5.2, we comment on the traditional DEA approach.

In Table 2, we observe that the optimal route based on shortest distance is the route that passes through the following node sequence: $1 \to 2 \to 3 \to 4 \to 7 \to 8$ (Figure 4), as according to Google Maps, the distance of going from 3 to 4, then from 4 to 7 was shorter than the route selected by its algorithm going straight from 3 to 7.

### Table 2: VRP results

<table>
<thead>
<tr>
<th>Variable (Route)</th>
<th>X12 (1 to 2)</th>
<th>X23 (2 to 3)</th>
<th>X34 (3 to 4)</th>
<th>X47 (4 to 7)</th>
<th>X78 (7 to 8)</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Distance (km)</td>
<td>64.30</td>
<td>15.50</td>
<td>16.80</td>
<td>12.30</td>
<td>15.00</td>
<td>123.90</td>
</tr>
<tr>
<td>Time (minutes)</td>
<td>85.00</td>
<td>50.00</td>
<td>100.00</td>
<td>60.00</td>
<td>75.00</td>
<td>370.00</td>
</tr>
<tr>
<td>Consumption (liters)</td>
<td>21.62</td>
<td>6.39</td>
<td>9.37</td>
<td>5.96</td>
<td>7.50</td>
<td>50.84</td>
</tr>
<tr>
<td>Emissions (CO₂ kg)</td>
<td>57.67</td>
<td>17.02</td>
<td>25.00</td>
<td>15.89</td>
<td>20.00</td>
<td>135.58</td>
</tr>
<tr>
<td>1/efficiency</td>
<td>4.83</td>
<td>1.76</td>
<td>2.72</td>
<td>1.77</td>
<td>2.20</td>
<td>13.27</td>
</tr>
</tbody>
</table>

In Table 3 we find that the shortest route was 200 meters shorter than the route obtained with the proposed DEA routing method, but that the route obtained by the proposed method resulted in reduction of other inputs.

### Table 3: Differences (DEA Routing - VRP)

<table>
<thead>
<tr>
<th>Variable</th>
<th>Differences ($\sum x_{DEA_Routing} - \sum x_{VRP}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Distance (km)</td>
<td>0.20</td>
</tr>
<tr>
<td>Time (minutes)</td>
<td>-50.00</td>
</tr>
<tr>
<td>Consumption (liters)</td>
<td>-2.54</td>
</tr>
<tr>
<td>Emissions (kg of CO₂)</td>
<td>-6.78</td>
</tr>
<tr>
<td>1/efficiency</td>
<td>-0.82</td>
</tr>
</tbody>
</table>
We also tested other objective functions, optimizing each of the variables separately. Doing so, we found the results to be equal to when optimizing efficiency (with the exception of distance). However, since emissions and consumption were more correlated to distance than to travel time, depending on the idle time, other objective functions could lead to a different result based on travelled time, although optimizing consumption or emission could still lead to similar results. If the idle time is equal to zero, the other variables are even more correlated to distance and the results are the same as when optimizing by shortest distance.

5.2. Traditional DEA approach

When we used the traditional DEA approach, we found a tied efficient result between the shortest distance approach (route $1 \rightarrow 2 \rightarrow 3 \rightarrow 4 \rightarrow 7 \rightarrow 8$) and the proposed DEA method (maximum efficiency route of $1 \rightarrow 3 \rightarrow 2 \rightarrow 7 \rightarrow 8$). Therefore, it is needed to use a tiebreaking method to obtain only one route for this traditional DEA approach. The main differences between the proposed approach and the traditional DEA approach mainly relate to the order of usage of the DEA model. The traditional approach is used after generating feasible routes, similarly to VRP. Thus, in addition to the complexity of choosing feasible routes, it is needed to create one DEA model per feasible route for analysis. In the case of the final result being a tie, it is needed to use an additional tiebreaking method to choose only one route.

The proposed method is used before generating the whole set of feasible routes. First, it evaluates the efficiency of each pair of origin and destination (part of the route between two nodes), and then efficiency scores are used in a standard VRP. Therefore, it is necessary to create only one DEA model per pair of nodes before we implement the VRP. As a result, the proposed method could reduce the number of DEA models created to evaluate the routes.

6. CONCLUSIONS

When using DEA in a stage before applying VRP, each route is a DMU and characterized by input and output variables. These variables (outputs and inputs) can be related to sustainability and green variables (e.g. emissions, fuel consumption, distance, travel time, and time in traffic or idle time).
Papers regarding VRP tend to focus on specific and limited performance characteristics such as cost or individual distance. Literature related to VRP and DEA usually generates feasible routes and uses only DEA to select efficient routes. Therefore, the main contribution of this paper was to use efficiency (including that of inefficient DMUs) as a criterion in VRP (in a two-stage DEA routing model) and to let a vehicle-routing model select the route of minimal inverted efficiency, instead of using DEA to evaluate entire routes generated previously. So, it is necessary to follow fewer steps and evaluating fewer DMUs, reducing the number of DEA models created to evaluate the routes and obtaining the results faster than the approach from the literature. In addition, efficiency was understood as an aggregate measure or composite indicator, and routing was therefore not limited to a single measure.

We performed an empirical application to routes for freight transportation from a traditional fruit producing region in the state of São Paulo, setting various destinations in the capital state, which is an important consumer region.

Comparing the application using only VRP and the proposed DEA-routing, the results showed that routing based on efficiency as a criterion leads to reduced inputs (excepts for the distance in the case analyzed) when traffic jam or idle time is considered, but as the extra traveling time gets near 0, it leads to similar results as using only distance. Comparing the the DEA-VRP from the literature and the proposed DEA-routing, the results showed that the DEA-VRP from the literature selects both above-mentioned routes (minimal distance and maximal efficiency) as efficient, so it is necessary to apply an extra step of tiebreaking method to select the most efficient route, while the proposed model leads to the optimal route with fewer steps and evaluating fewer DMUs (because for large routes, the number of pair of nodes is smaller than the number of possible routes).

The final conclusion is if producers from the “Fruit Trail” region follow our recommendations, they will travel roughly the same distance indicated by traditional methods (including Google Maps), but will be able to deliver their products faster, consuming less diesel, and emitting less CO₂ into the environment. However, as additional travel time nears zero, the results become similar to those when using distance only.

For further studies, we suggest incorporating time in traffic or idle time into other scenarios including different days and hours of the day and including more variables. Future research could
also study factors such as the slope of the route, refuelling options, cargo, capacity, resources, characteristics of sustainable or green corridors, and structures that make a corridor green or sustainable—all of which would render DEA Routing more suitable than choosing only one limited aspect of the route and using separately routing. Finally, we suggest improvements to the model including developing a dynamic model or integrating both DEA and VRP in a multi-level optimization problem (instead of running the model in two stages).

ACKNOWLEDGEMENT

The authors would like to thank the São Paulo Research Foundation (FAPESP), grant #2018/20436-1, for supporting this research study.

REFERENCES


