



# Article Operational Performance Analysis of the Public Transport System over Time

José Nauri Cazuza de Sousa Júnior <sup>1,\*,†,‡</sup>, Teresa Galvão Dias <sup>2,3,‡</sup> and Mário Angelo Nunes de Azevedo Filho <sup>4,‡</sup>

- <sup>1</sup> Regulatory Agency for Delegated Public Services of the State of Ceará, Fortaleza 60822-325, Brazil
- <sup>2</sup> Department of Industrial Engineering and Management, University of Porto, 4200-465 Porto, Portugal; tgalvao@fe.up.pt
- <sup>3</sup> INESC TEC-Institute for Systems and Computer Engineering, Technology and Science, 4200-465 Porto, Portugal
- <sup>4</sup> Federal Technology Center, Department of Transport Engineering, Campus do Pici, Federal University of Ceará, Fortaleza 60440-900, Brazil; azevedo@det.ufc.br
- \* Correspondence: naurijr@det.ufc.br
- + Current address: Programa de Pós-Graduação em Engenharia de Transportes, Centro de Tecnologia, Campus do Pici, Universidade Federal do Ceará, Fortaleza 60440-900, Brasil.
- <sup>‡</sup> These authors contributed equally to this work.

Abstract: The public transport system is responsible for the displacement of a large part of the population, particularly in developing countries. This fact makes it relevant to evaluate the performance of public transport to provide an efficient and effective service. The purpose of this study is to conduct a performance evaluation of the public transport operation in the Metropolitan Region of Fortaleza (MRF), in the State of Ceará, Brazil. The analysis is based on DEA and the Malmquist index, based on three inputs (total operating time, fleet age, and the mileage traveled) and two outputs (fare revenue and number of passengers). Data were obtained through automated fare collection systems (AFCs) that were implemented in the MRF. Although there were no major fluctuations in performance during the analyzed period, the results indicate that the system's performance declined in certain years. In addition, the analysis enables a better understanding of route performance, considering the operating company or the area of operation, which helps to diagnose and comprehend the operation more effectively. By analyzing the operational performance over time, the proposed approach provides an additional contribution by offering a comprehensive overview to the involved stakeholders, fostering decision-making processes based on evidence.

**Keywords:** Malmquist index; data envelopment analysis; benchmarking; performance evaluation; public transport; performance indicators

# 1. Introduction

Mass transportation systems are crucial in large cities and metropolitan areas globally, not only because they enable the movement of large numbers of people but also because they play a significant role in addressing environmental concerns and promoting the wellbeing of citizens, according to Belwal and Belwal [1], Booysen et al. [2], Javid et al. [3], Wang et al. [4]. As supported by Belwal and Belwal [1], public transport services play a crucial role in addressing exclusion and poverty among the lower-income segments of the population in the region served. Debnath [5] states that efficient transport plays a significant role in the development of any region. Furthermore, Alonso et al. [6] highlight that a pressing challenge faced by developing countries is to promote the adoption of sustainable transportation options for daily commuting in order to mitigate pollution.

Urban centers face significant challenges in transporting people, including inadequate infrastructures, increased private vehicle ownership leading to congestion, and an



Citation: Cazuza de Sousa Júnior, J.N.; Dias, T.G.; Nunes de Azevedo Filho, M.A. Operational Performance Analysis of the Public Transport System over Time. *Infrastructures* 2023, *8*, 82. https://doi.org/10.3390/ infrastructures8050082

Academic Editor: Benedetto Barabino

Received: 30 March 2023 Revised: 23 April 2023 Accepted: 25 April 2023 Published: 28 April 2023



**Copyright:** © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https:// creativecommons.org/licenses/by/ 4.0/). inefficient public transport system [7]. Public transport, often managed and regulated by government institutions, aims to provide high-quality transportation options that address the needs of the population, as affirmed by OECD [8], which emphasizes the role of transport authorities in ensuring efficient and quality service from public or private operators. The services are crucial in facilitating travel for various activities such as work, education, leisure, and healthcare, by offering coverage in terms of space and time.

According to Sun et al. [9], evaluating the performance of bus routes can be challenging when limited indicators are available. Therefore, regular performance evaluation is essential to ensure the efficient use of available resources and the effective transportation of people. Evaluation processes can also support managers in making informed decisions by providing operational indicators that specify the details of service provision, including location, timing, and operational features.

Wei et al. [10] reinforce the importance of evaluating the public transport system by stating that it is an essential and challenging issue for researchers and transport authorities, in view of limited funding and growing public needs. Evaluating the performance can assist transportation agencies in identifying underperforming services, planning potential investments, justifying past investments, and communicating accomplishments and challenges [11,12].

De Carvalho and Marques [13] suggest that using rewards and penalties based on indicators can be an effective way to motivate and maintain a high level of performance. The authors also emphasize the importance of using simple and intuitive indicators that can be easily monitored.

Assessing the quality of services can be achieved through qualitative measures (such as user opinion surveys) or quantitative measures (such as utilizing historical service data). However, conducting qualitative assessments involves interviewing service users, which can be challenging to carry out regularly due to its subjective nature and high costs. Conversely, a quantitative evaluation relies on historical data or is obtained through field research, which is also costly and infrequently conducted, as mentioned by Zhou et al. [14].

Due to the development of new technologies, the utilization of intelligent transportation systems (ITSs) has become increasingly common, as affirmed by Qureshi and Abdullah [15]. Such systems include automated transit management systems (ATMSs), which are composed of other systems, such as automated fare collection systems (AFCs) and automated vehicle location (AVL). Gkiotsalitis et al. [16] note that public transport systems have increased the amount of equipment and sensors that allow a passive collection of large amounts of data. The utilization of massive amounts of data allows for the analysis and evaluation of the service, which is crucial in assessing the performance of public transport routes with accuracy and monitoring the quality of service provided.

Jarboui et al. [17] note that research in this field is limited to certain regions, namely Europe, North America, and Asia. The authors highlight that the countries with the most research on this topic are the United States, Norway, Germany, and India. However, it should be noted that there is already some research on this topic in other regions, albeit in small numbers. This could be due to a lack of data or a shortage of qualified professionals.

According to Litman [18], the impacts of public transport services can be diverse, including both indirect and external effects (i.e., those that affect people who do not use the service). The author also outlines four categories that can be targeted for improvements in public transport: service expansion, service enhancement, demand management, and transit-oriented development (TOD).

Kurauchi and Schmöcker [19] mention that electronic ticketing systems provide spatial and temporal information on users, vehicles, and routes. These data can be used by authorities for various purposes, such as:

- To obtain the number of passengers by fare type, which enables the measurement of the network load by the type of user and the assessment of the economic viability of the route;
- To share revenues between operators;
- To consult stop-wise distribution of passengers boarding and alighting, by fare type;
- To compute origin-destination matrices for each route, subset of routes, or the entire network.

Additionally, electronic ticketing systems offer the potential for a wide range of disaggregated or aggregated analyses per trip, depending on the intended purpose. As noted by Arbex and da Cunha [20], electronic ticketing data, as well as other data passively generated by public transport systems, allow for a much more precise and continuous technical evaluation and monitoring of the implementation of changes. Furthermore, Liu et al. [21] developed a replication methodology that includes cleaning, filtering, processing, and converting collected data into information, such as bus travel paths and subway system schedules.

According to Tsolas [22], there are only a few studies that have analyzed the efficiency of the urban transport system based on a data envelopment analysis (DEA), especially regarding production and sales processes. The authors used a two-stage DEA to evaluate the performance of a group of electric bus routes in Greece. They also note that, with the series of data available, a DEA can provide a dynamic assessment of public transport performance.

Georgiadis et al. [23], Karim and Fouad [24] highlight that there are limited studies that have utilized DEAs and considered routes as decision-making units (DMUs) in the context of public transport. The authors note that only a small percentage (1.6%) of the studies have measured the performance of routes, with the majority focusing on operators (76%) and systems (21%). Furthermore, it is worth noting that the majority of route performance analysis studies were published after 2010.

Thus, this work aims to evaluate the operational efficiency of regular public transport routes in the Metropolitan Region of Fortaleza (MRF) through a DEA and the Malmquist index. The approach fills the research gap in the use of massive data to assess the performance of public transport operations over time, providing operators and managers with information to support decision-making. Furthermore, this work proposes a performance analysis based on the routes, which is a perspective made possible by the detailed data available from the ticketing system. This approach allows for a more detailed analysis of the operational performance than those studies where DMUs are the operators or the transportation systems.

The following section provides a literature review of the DEA model and Malmquist index, along with a description of the variable selection process. Section 3 outlines the methodology and dataset employed to evaluate the operational efficiency of the service. Section 4 presents the results of the application of this methodology to the case study, followed by a detailed discussion of the results in Section 5. Lastly, the article concludes with implications and final considerations.

#### 2. Literature Review

The literature review of the DEA model, the Malmquist index, and the selection of variables.

## 2.1. DEA

The DEA is a method based on linear programming that measures the relative efficiency of entities called decision-making units (DMUs) based on input and output information. Sun et al. [9] used DEA to categorize and evaluate existing routes, noting that few studies used new technologies to assess the quality of the service. Wei et al. [10] developed a new method to evaluate the overall performance of public transport services, using a combination of mathematical programming methods, GIS-based analysis and multi-objective spatial optimization techniques. Lee et al. [25] mentioned that the DEA model has four assumptions: return to scale, convexity, free disposability, and input or output orientation. The mathematical formulation of the model is presented in Equation (1), as per Seiford and Zhu [26].

$$\max\sum_{r=1}^{5} u_r y_{ro} + u_o \tag{1}$$

subject to:

$$\sum_{i=1}^{m} v_i x_{io} = 1$$
  

$$\sum_{r=1}^{s} u_r y_{rj} - \sum_{i=1}^{m} v_i x_{ij} + u_o \le 0 \ j = 1, ..., n;$$
  

$$u_r, v_i > 0 \text{ and } u_o \text{ is free}$$

wherein,

*m*: total number of inputs; *s*: total number of outputs; *n*: total number of DMUs;  $x_{io}$ : ith input for DMU<sub>o</sub>;  $y_{ro}$ : rth output for DMU<sub>o</sub>; j = 1, 2, ..., n;  $v_i$ : weight vector of input x;  $u_r$ : weight vector of output y.

The constraints are orientation, efficiency, and non-negativity. The constraints of orientation (input-oriented DEA variable returns to scale—VRS model) lean toward minimizing the inputs, which means that all DMUs seek to minimize the number of inputs required to produce their outputs. The efficiency constraint requires that all DMUs be efficient or operate at a point of maximum efficiency, i.e., producing the maximum amount of outputs with the minimum amount of inputs. Moreover, the non-negativity constraint prevents the weights of inputs and outputs from being negative.

The DEA models proposed by Charnes (constant returns to scale—CRS) and Banker (VRS) are called standard, conventional, or basic models, according to Lovell and Rouse [27]. Additionally, there are other models adapted from these, such as the super-efficient model (Sun et al. [9], Lovell and Rouse [27]), Malmquist index (Färe et al. [28]), and network DEA (Färe et al. [29], Daraio et al. [30]). Applications and literature reviews of other models can be seen in work by Kuah et al. [31], Cavaignac et al. [32].

#### 2.2. Malmquist Index

The Malmquist index was introduced by Caves et al. [33]; it was first used in the DEA literature by Berger and Humphrey [34] and further developed by Färe et al. [28]. This is a commonly used method to measure productivity in different periods (t and t + 1). It is based on the concept of a distance function, which measures the relative efficiency of DMUs in transforming inputs into outputs. The Malmquist index calculates the productivity change over a period of time by comparing the efficiency of a DMU at two different points in time. Färe et al. [28] propose, through the geometric mean of two indices, the expression inside the square brackets in Equation (2), for periods t and t + 1. This equation represents the Malmquist index, where the first part of the equation is related to the change in technical efficiency and the second part (in the square root) is related to technological change.

$$MI^{t,t+1} = \frac{E^{t+1}(x^{t+1}, y^{t+1})}{E^{t}(x^{t}, y^{t})} \sqrt{\frac{E^{t}(x^{t}, y^{t})}{E^{t+1}(x^{t}, y^{t})}} \frac{E^{t}(x^{t+1}, y^{t+1})}{E^{t+1}(x^{t+1}, y^{t+1})}$$
(2)

wherein,

 $MI^{t,t+1}$ : Malmquist index between period t e t + 1;  $x^t, y^t$ : Inputs and outputs for period t;  $x^{t+1}, y^{t+1}$ : Inputs and outputs for period t + 1;  $E^t$ : Efficiency for period *t*;

 $E^{t+1}$ : Efficiency for period t + 1.

The decomposition of the Malmquist index into technical efficiency and technological change allows us to identify whether the increase is attributed to technological advancements, improvements in technical efficiency, or both. The assessment of technical efficiency measures how a production unit's efficiency compares to other units within the same period. It determines whether the unit is more efficient or less efficient in terms of production compared to its counterparts. On the other hand, technological change evaluates whether the technology utilized by the production unit has improved or deteriorated over time. This is determined by comparing the production efficiency in one period to that in another period.

According to Oliveira et al. [35], the index can be greater than, equal to, or less than 1, where greater than 1 indicates performance growth, while a value less than 1 reflects a reduction for periods t and t + 1.

## 2.3. Selection of Variables

The selection of variables for the application of the DEA or Malmquist index model is a topic extensively discussed in the literature, according to the number of variables and how to define the choice; Georgiadis et al. [23], De Borger et al. [36], Holmgren [37] note that there is no consensus for defining the variables. The authors state that the selection of variables is often determined by practical issues, such as data availability or the method used.

The selection of variables was based on those used in previous studies, which represent aspects related to the operation of the service, in order to analyze the following categories: availability, time, service offered, maintenance, and economics, according to European (CEN/TC320 [38]) and American (Kittelson et al. [39]) technical regulations. In addition, a correlation analysis was performed between the available variables, although Dyson et al. [40] state that omitting variables purely because they have a high correlation should be avoided.

The study chose to use only variables with absolute values, avoiding the use of ratios or rates, as recommended by Dyson et al. [40] and Emrouznejad and Amin [41]. The authors suggest an alternative to using rates or indices as inputs and outputs, which is to include the numerator as an input and the denominator as an output if the rate is used as an input. Conversely, if the rate is an output, the numerator should be used as an output and the denominator as an input. In addition to these issues, Dyson et al. [40] pointed out some pitfalls that must be avoided or be carefully handled in the application of a DEA, namely non-homogeneous units or environments, non-verification of economies of scale, a high number of inputs and outputs, use of qualitative data, undesirable inputs and outputs, and external and constrained factors. To ensure accurate results, the study aimed to follow the protocols outlined by the authors and avoid these pitfalls.

The total operating time was used in several studies, such as in work by Sheth et al. [42], Güner and Coşkun [43], Tran et al. [44]. This variable represents the actual period of service provision during the year and is related to the costs of the operation (vehicle and labor costs).

In the study by Hanumappa et al. [7], other inputs, namely fleet age and mileage traveled, were used. Fleet age is an indicator of maintenance costs, as older vehicles are expected to require more maintenance (see Taboada and Han [45]) compared to other variables, such as service quality and type of service. On the other hand, mileage traveled, such as the total operating time, relates to fleet and labor costs, but it also varies due to the traffic conditions on the route.

The number of passenger variables is the most frequently used output in the analyzed publications (see Karim and Fouad [24], Güner and Coşkun [43], Azambuja et al. [46], Sampaio et al. [47], Araújo et al. [48], Sun et al. [49], Fitzová et al. [50], Fitzová and Matulová [51]). It is a crucial variable as it represents a primary focus of the public trans-

port service provision, meeting the travel needs of passengers while providing a high quality service.

The other output variable used is fare revenue, which is utilized in work by Fitzová and Matulová [51], Zhu et al. [52], Zhang et al. [53]. Fare revenue is useful in assessing the economic viability of the service, and, similar to the number of passengers, it represents an important outcome of the public transport supply.

In addition to the studies presented, Daraio et al. [30] conducted a review of studies with the main measures used for efficiency analysis, among them, are the operating time (8.9%), fleet age (14.5%), mileage traveled (53.2%), number of passengers (16.9%), and fare revenue (11.3%). The percentages of papers reported for the variables are in parentheses.

The use of ticketing system data may underestimate passenger demand due to fare evasion (passengers with no card, passengers who use the transport several times with the same card or passengers who do not validate the card), as noted by Barabino et al. [54], National Academies of Sciences et al. [55], Lee [56].

#### 3. Materials and Methods

Monitoring and analyzing the operation of public transport networks from various perspectives is essential due to the significance of public transport in urban and metropolitan regions, and the challenges involved in meeting the population's needs with a safe, reliable, efficient, effective, and reasonably priced service. These perspectives may include, but are not limited to, availability, accessibility, time, security, information, services offered, and economic, social, and environmental factors. Figure 1 presents the steps of the research.



Figure 1. Flowchart of research steps.

The first step corresponds to the definition of the research objective, which is to evaluate the operational efficiency of regular public transport routes. The second step involves the collection of data, considering factors such as importance, availability, and cost. Step 3 entails the selection of input and output variables for the model, as described in Section 2.3. In Step 4, the approach (DEA method and Malmquist index) is chosen. Step 5 involves applying the models to the case study, followed by steps 6 and 7, which include analyzing the results, discussing the findings, and drawing conclusions, respectively.

#### 3.1. Data Envelopment Analysis and the Malmquist Index

The initial approach employed in the analysis was the conventional DEA model. As stated before, in this model, DMUs that lie on the frontier receive an efficiency index of 1, while those with an index less than 1 are deemed inefficient. The model can be CRS (constant returns to scale) or VRS (variable returns to scale) and can be either input- or output-oriented.

This approach enables the evaluation and comparison of different routes, allowing for the assessment of their performance and identification of areas that require improvement to enhance efficiency. By computing the efficiency index of each route, it is possible to scrutinize them and diagnose required adjustments to provide a better quality of service. The performance evaluation of each route is based on comparisons with other routes within the same public transport system, serving the same area and population.

After the initial analysis, the next step is to conduct a performance analysis of the routes over time to determine the variation in their efficiency. This is achieved using the Malmquist Total Productivity Index.

### 3.2. Data Set—Case Study

The case study focused on 15 out of 19 municipalities in the Metropolitan Region of Fortaleza (MRF) that have a regular metropolitan public transport service. These 15 municipalities are served by 7 different companies, as shown in Figure 2. It is important to note that the remaining four municipalities were not included in the study as they did not have a regular metropolitan public transport service during the specified period. The provision of this service plays a crucial role in integrating the municipalities within the MRF, which directly impacts economic activities in the region and its surroundings. The management and regulation of the service are the responsibility of the Government of the State of Ceará, with the task delegated to the Regulatory Agency for Delegated Public Services of the State of Ceará (ARCE).



Figure 2. Map of the Public Transport System of the Metropolitan Region of Fortaleza—Ceará.

In 2022, the metropolitan public transport system recorded over 800,000 bus trips and approximately 17 million passenger trips. The service operates using urban buses with two doors, where passengers enter through the front door and validate their transport card or pay cash to the collector. After that, all users must pass through the turnstile (ticket barrier), ensuring low fare evasion and accurate passenger counting. Additionally, multi-fare routes have a collector, and all vehicles are equipped with a surveillance camera system. The system only registers boarding data and does not record alighting information. The pricing system is based on fare zones, also known as fare rings, which are structured according to the distance covered, resulting in varying fare ranges.

The construction of the database refers to the period from 2016 to 2022. The database contains 246.1 million records (70.3 Gb). In addition to the data validation, the database also includes information about the routes, itineraries, schedules, fare values, and mileage. Each validation record contains 29 fields. For this study, the following were used:

- OperatorID : Code of the company responsible for the vehicle;
- VehicleID: Vehicle identification;
- RouteID: Route identification;
- TripDepTime: Departure date and time of the trip;
- TripFinishTime: Finish date and time of the trip;
- Fee: Fee paid by the user.

The analysis was conducted from the perspective of managers and operators, as other perspectives would have required additional data beyond what was available through the ticketing or monitoring system. From the viewpoint of managers and operators, performance metrics were derived from the desired level of service quality, which was established by defining service goals, and the actual service quality, which was measured by the level of quality achieved in daily service operations. By comparing the planned and actual service operations, managers and operators could identify inefficiencies, weaknesses, and potential improvement guidelines. While there are several other aspects of the public transport service that could be analyzed, such as cost/financial feasibility, investments, attractiveness, externalities, or flexibility to adapt to demand, this research focuses specifically on the performance of public transport operations.

Prior to conducting the data analysis, it was necessary to pre-process the raw data Zhou et al. [14], Barry et al. [57], Nunes et al. [58]. Raw data often contain incorrect, illogical, or missing records, making it necessary to prepare the data before analysis. Trépanier et al. [59] discussed the pre-processing step, which involved identifying and correcting errors before proceeding with the analysis.

In the exploratory analysis, extreme values of operating time were found, either too low or too high, which were not reasonable for the operation and were excluded from the data set. Therefore, records with travel times shorter than what could be achieved with the speed of the road (60 km/h) were removed. In addition, an analysis of outliers was performed. According to Pearson [60], outlier detection can be performed as an automatic mathematical procedure, either by the three-sigma rule, the Hampel identifier, or the boxplot rule. However, the interpretation of the results required expert knowledge and experience. This study used the boxplot rule to remove outliers for the operating time variable. working days were selected for the analysis, while weekends and holidays were excluded due to their different behaviors. Additionally, routes that did not belong to the metropolitan service or did not operate consistently throughout the year were removed from the analysis as the focus was on regular routes.

Table 1 shows the number of trips selected after each stage of data processing.

After constructing the database containing only data from regular metropolitan routes, the analysis proceeded to the next stage, which involved examining the selected variables.

Phases	2016	2017	2018	2019	2020	2021	2022
Raw data	1,307,043	1,309,493	1,220,498	1,176,309	683,971	757,634	811,309
Exclusion of outliers	1,079,537	1,079,941	1,028,157	965,857	614,075	677,839	685,941
Exclusion of holidays and weekends	810,479	814,921	814,892	730,367	475,067	515,687	526,602
Exclusion of non-regular routes	806,565	805,283	753,370	720,836	429,743	501,355	516,207

**Table 1.** Processing steps for public transport ticketing data.

The variables were selected based on the following aspects: importance in the operation of public transport, application in other studies (see Section 2), and availability of data. Based on these aspects, the following variables were calculated:

- Total operating time: Sum of the operating time of each trip per route for the entire year. The operating time was obtained by subtracting the arrival time and the departure time, provided by the ticketing system. This variable allows for analyzing temporal aspects of each route operation;
- Fleet age: Sum of the age of the vehicle used per trip and per route throughout the year. The age of the vehicle was obtained through the vehicle register provided by ARCE and the vehicle used in each trip was included in the ticketing system. This variable is related to the maintenance and comfort of the vehicles;
- Mileage traveled: Sum of the distance traveled for each trip per route for the whole year. The length of the route was obtained through the registration of the routes provided by ARCE and the number of trips was obtained by the ticketing system. This variable allows for analyzing the availability and the service offered;
- Fare revenue: Total fare revenue per trip for each route for the entire year. The fare revenue is the sum of the amount paid by the user and is provided by the ticketing system. This variable reflects the economic and financial viability of the route and service provided;
- Number of passengers: Total number of passengers per trip for each route for the whole year. The number of passengers is the sum of the number of records in the ticketing system. The variable allows for analyzing the demand of each route.

The public transport supply is represented by the input variables (total operating time, fleet age, and mileage traveled) while the demand is represented by the output variables (fare revenue and number of passengers). Table 2 provides the descriptive statistics of these variables for the year 2022.

Descriptive Statistics	Total Operating Time—In Hours	Fleet Age—In Years	Mileage Traveled—In km	Fare Revenue—In BRL	Number of Passengers
Average	10,024.35	40,598.88	214,912.55	1,332,475.09	298,168
Median	4012.46	26,428	126,000	671,289.61	99,822
Standard deviation	9785.06	37,844.34	203,021.42	1,353,761.75	339,442
Minimum	372.11	1401	4647.15	13,242.15	7155
Maximum	37,165.04	131,257	800,403.70	5,445,283.96	1,389,964
Sum	591,436.38	2,395,334	12,679,840.39	78,616,030.49	17,591,929

Table 2. Descriptive statistics of the variables used—2022.

## 4. Results

To analyze the operational efficiency of the routes, an input-oriented DEA with variable returns to scale was applied to each route per year from 2016 to 2022. Table 3 provides a yearly summary, including the number of routes analyzed, the efficient routes, and descriptive statistics of the performance index obtained.

During the study period, there were changes in the transport network, with fluctuations of up to 20 routes between 2019 and 2020, indicating adjustments made to the network. However, despite these changes, the average performance index only showed a slight fluctuation of 0.071, suggesting that the network adjustments did not have a significant impact on performance. Additionally, there has been an increase in the average operating performance index since 2018.

In a DEA approach, the difference between the input/output levels of the inefficient DMUs and the corresponding input/output levels of the efficient DMUs provide a measure of the relative inefficiency or potential for improvement of the inefficient ones. In this work, the efficient routes are used as benchmarks for the inefficient routes in terms of input–output ratios, representing the optimal utilization of inputs (operating time, fleet age, and mileage traveled) to produce outputs (number of passengers and fare revenue). For example, in 2022, route 67 operated by company 5, which connects the capital city of Fortaleza to the municipalities of Eusébio, Itaitinga, and Pacatuba, was an efficient route and served as a benchmark for 32 less efficient routes. Routes 9 and 83, operated by companies 2 and 6, respectively, serving the municipalities of Maranguape and Maracanaú, were efficient in all years and served as references for up to 40 inefficient routes in 2019. Additionally, an additional 40 routes were efficient in at least one of the seven years.

Description	2016	2017	2018	2019	2020	2021	2022
Number of routes	68	69	68	73	53	62	59
Efficient routes	15	16	12	14	14	17	13
Average	0.746	0.757	0.744	0.750	0.783	0.813	0.814
Median	0.762	0.793	0.784	0.758	0.810	0.842	0.852
Standard deviation	0.223	0.212	0.216	0.190	0.204	0.181	0.175
Minimum	0.216	0.290	0.287	0.331	0.129	0.256	0.337

Table 3. Measuring the operational performances of the bus routes.

The other analysis performed was to assess the operational performance over time through the Malmquist index. The same variables of the classic DEA model were considered. Table 4 presents the main results of this analysis. Only DMUs that operated in both periods t and t + 1 were included in the analysis. Any routes that did not operate in both periods were excluded.

Tab	le 4.	N	leasuring	g the o	perational	perf	formance of	f pu	bli	c trans	port	over	time-	-Ma	almo	quist	ind	lex.
-----	-------	---	-----------	---------	------------	------	-------------	------	-----	---------	------	------	-------	-----	------	-------	-----	------

Description	2016/2017	2017/2018	2018/2019	2019/2020	2020/2021	2021/2022	2016/2022
Number of routes	66	64	63	49	50	57	45
Routes with increasing productivity	24	34	24	7	27	32	11
Average	1.046	0.993	1.100	0.826	1.026	1.269	0.896
Median	0.912	1.013	0.956	0.805	1.027	1.036	0.751
Standard deviation	0.621	0.166	0.988	0.187	0.246	1.027	0.502
Minimum	0.229	0.349	0.317	0.443	0.547	0.146	0.147
Maximum	3.442	1.485	8.389	1.385	1.893	6.879	2.811

It is possible to see that the number of DMUs with productivity growth varied over time, with 34 routes in 2018 compared to 2017, and only 7 routes in 2020 compared to 2019. It is worth mentioning that the results for 2020 may have been affected by external factors, particularly the COVID-19 pandemic, which had its first wave between March and June 2020. The mean and median index values were consistently close to 1, except for the 2- and 3-year comparisons, where the mean and median were below 1, respectively.

When comparing only the initial (2016) and final (2022) years, only 11 routes showed productivity growth out of a total of 45 routes that remained operating during the period. The mean and median values of the index were less than 1, indicating a decrease in the system's productivity between 2016 and 2022.

#### 5. Discussion

The analysis and results obtained, both by route and over time, allow for the identification of critical areas that require attention and intervention by managers and service operators. One approach to interpreting the results is to evaluate them by company, considering that they have similar resources allocated to all routes. This can help determine measures that can be taken to minimize inefficiencies. Another aspect to consider in terms of operational performance is the spatial dimension. Routes that have similar origin– destination patterns and itineraries may serve a particular region with common demand. Therefore, it is essential to analyze what factors contribute to one route being inefficient while another similar route is efficient.

Apart from the aforementioned methods, analyzing the system's performance over time enables the establishment of correlations with external factors that may influence its operations. The outbreak of the COVID-19 pandemic in 2020, for instance, had a significant impact on public transportation. In the region under analysis, a decrease in the number of trips and routes was observed. Furthermore, the pandemic had an adverse effect on the system's productivity, which is evident in its lowest performance index (0.129) throughout the entire period studied, as well as in the mean and median values of the Malmquist index for consecutive years.

The operational performance index of three routes for company 1, obtained for each year, is presented in Figure 3. This visualization allows for a comparison of performance between the routes and reveals that routes 1 and 3 had high indices over the period, being efficient in some years. In contrast, route 2 consistently had lower values compared to the other two routes of the analyzed operator.



Figure 3. Analysis of three routes of company 1.



Figures 4 and 5 present graphs related to the performance of Route 2, which was operated by three different companies. Both the performance index (Figure 4) and the Malmquist index (Figure 5) showed a decline, indicating a decrease in operational performance over time.

Figure 4. Analysis of Route 1 operated by three companies—classic DEA.

Beginning in 2020, when route 2 was operated exclusively by company 1, an improvement in its operational performance could be observed. This suggests that there may have been an oversupply of resources for low demand, coupled with inadequate fare revenue, which negatively impacted the route's performance in the past.



Figure 5. Analysis of Route 1 operated by three companies—Malmquist index.

Despite the stability of the operating performance during the analyzed period, as indicated by the average operational performance in Table 3 and the average Malmquist index in Table 4, it is important to highlight that the public transport network experienced a reduction in supply (number of trips and routes) and demand (number of passengers), as mentioned earlier. This reduction may have resulted in a decrease in the quality of service provided and left certain regions without access to metropolitan public transport.

Therefore, when conducting the performance analysis, it is crucial to consider additional data and information to avoid drawing premature conclusions.

Therefore, by readjusting the number of trips or changing the schedule, it would be possible to enhance operational efficiency. This analysis can also be extended to other routes, identifying the critical areas that require improvement.

#### 6. Conclusions

The objective of this study was to propose a methodology for evaluating the operational performance of public transport and to apply it to the public metropolitan bus transport system of MRF. The results obtained by the DEA method can provide valuable information for managers and transport operators to make better decisions and improve the quality of service.

This research provides valuable insight for managers and operators in the specific area, as well as proposes a novel approach for analyzing operational performance that can be replicated by managers and operators in other areas or transportation systems. Additionally, it offers valuable input for researchers in the field of public transport who aim to compare methodologies and performances across different public transport systems.

By comparing the results obtained from the classic DEA and the Malmquist index, the analysis was enhanced by evaluating the performance for a specific year and its progression over time. The pre-processing analysis revealed issues in the data recording stage, emphasizing the need for the managing body and operators to improve the ticketing system process. This includes recording the start and end of each trip and the direction of the route. This improvement would help reduce the number of outliers, thereby increasing the reliability of the records.

When attempting to improve route efficiency by reducing inputs without altering output values, the model yielded significant reduction values. To minimize inefficiencies, it is suggested to reduce travel time and/or the number of trips, reorganize the network through route grouping, and review the distribution of bus stops, among other potential modifications. These modifications should consider time slots and the spatial structure of the network to avoid compromising service quality.

The study presents the limitation of not incorporating cost variables (such as operational costs, waiting time costs, etc.) due to the unavailability of disaggregated data during trip and route levels.

One potential continuation of this research involves incorporating external factors into the analysis and evaluating their impacts on public transport performance. Another possibility is to conduct studies that focus on specific companies or sets of routes that share similar characteristics. This approach could provide more nuanced insights into the factors that influence operational efficiency and help guide decision-making at a more granular level.

Author Contributions: The manuscript was written through the contributions of all authors, as follows: Conceptualization, J.N.C.d.S.J. and T.G.D.; methodology, J.N.C.d.S.J., T.G.D. and M.A.N.d.A.F.; validation, J.N.C.d.S.J., T.G.D. and M.A.N.d.A.F.; formal analysis, J.N.C.d.S.J.; investigation, J.N.C.d.S.J.; resources, J.N.C.d.S.J.; writing—original draft preparation, J.N.C.d.S.J.; writing—review and editing, J.N.C.d.S.J., T.G.D. and M.A.N.d.A.F.; visualization, J.N.C.d.S.J.; supervision, T.G.D. and M.A.N.d.A.F.; project administration, J.N.C.d.S.J. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

**Data Availability Statement:** The datasets generated and analyzed in the research are available from the corresponding author and can be furnished upon request.

Acknowledgments: We would like to thank the Agência Reguladora de Serviços Públicos Delegados do Estado do Ceará (ARCE) and the Government of the State of Ceará for their support in conducting this research.

## 14 of 16

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

### Abbreviations

The following abbreviations are used in this manuscript:

AFC	automated fare collection
ATMS	automated transit management system
CRS	constant returns to scale
DEA	data envelopment analysis
DMU	decision-making units
GIS	geographic information system
MRF	Metropolitan Region of Fortaleza
TOD	transit-oriented development
VRS	variable returns to scale

## References

- Belwal, R.; Belwal, S. Public transportation services in Oman: A study of public perceptions. J. Public Transp. 2010, 13, 1–21. [CrossRef]
- Booysen, M.; Abraham, C.; Rix, A.; Ndibatya, I. Walking on sunshine: Pairing electric vehicles with solar energy for sustainable informal public transport in Uganda. *Energy Res. Soc. Sci.* 2022, 85, 102403. [CrossRef]
- 3. Javid, M.A.; Saif Al-Khatri, H.; Said Al-Abri, S.; Ali, N.; Chaiyasarn, K.; Joyklad, P. Travelers' Perceptions on Significance of Travel Time Saving Attributes in Travel Behavior: A Case Study in Oman. *Infrastructures* **2022**, *7*, 78. [CrossRef]
- 4. Wang, Z.; Li, X.; Bobylev, N.; Li, S.; Liu, J. Coupling Coordination of "Urban Rail Transit—Social Economy" Composite System. *Infrastructures* **2022**, *7*, 56. [CrossRef]
- 5. Debnath, P. A QGIS-Based Road Network Analysis for Sustainable Road Network Infrastructure: An Application to the Cachar District in Assam, India. *Infrastructures* **2022**, *7*, 114. [CrossRef]
- 6. Alonso, F.; Faus, M.; Cendales, B.; Useche, S.A. Citizens' perceptions in relation to transport systems and infrastructures: A nationwide study in the dominican republic. *Infrastructures* **2021**, *6*, 153. [CrossRef]
- 7. Hanumappa, D.; Ramachandran, P.; Sitharam, T.; Lakshmana, S. Performance evaluation of Bangalore metropolitan transport corporation: An application of data envelopment analysis. *J. Public Transp.* **2015**, *18*, 1–19. [CrossRef]
- 8. OECD. Competition and Regulation in the Provision of Local Transportation Services, OECD Competition Policy Roundtable Background Note; OECD: Paris, France, 2022; p. 44.
- 9. Sun, D.; Chen, S.; Zhang, C.; Shen, S. A bus route evaluation model based on GIS and super-efficient data envelopment analysis. *Transp. Plan. Technol.* **2016**, *39*, 407–423. [CrossRef]
- 10. Wei, R.; Liu, X.; Mu, Y.; Wang, L.; Golub, A.; Farber, S. Evaluating public transit services for operational efficiency and access equity. *J. Transp. Geogr.* 2017, 65, 70–79. [CrossRef]
- 11. Bertini, R.L.; El-Geneidy, A. Generating transit performance measures with archived data. *Transp. Res. Rec.* 2003, 1841, 109–119. [CrossRef]
- 12. Hassan, M.N.; Hawas, Y.E.; Ahmed, K. A multi-dimensional framework for evaluating the transit service performance. *Transp. Res. Part A Policy Pract.* **2013**, *50*, 47–61. [CrossRef]
- 13. De Carvalho, G.S.D.; Marques, R.C. Strategies to foster competition for the market in the urban bus sector in developing countries. *Infrastructures* **2020**, *5*, 115. [CrossRef]
- 14. Zhou, J.; Sipe, N.; Ma, Z.; Mateo-Babiano, D.; Darchen, S. Monitoring transit-served areas with smartcard data: A Brisbane case study. *J. Transp. Geogr.* 2019, *76*, 265–275. [CrossRef]
- 15. Qureshi, K.N.; Abdullah, A.H. A survey on intelligent transportation systems. Middle-East J. Sci. Res. 2013, 15, 629–642.
- 16. Gkiotsalitis, K.; Cats, O.; Liu, T. A review of public transport transfer synchronisation at the real-time control phase. *Transp. Rev.* **2023**, *43*, 88–107. [CrossRef]
- 17. Jarboui, S.; Forget, P.; Boujelbene, Y. Public road transport efficiency: A literature review via the classification scheme. *Public Transp.* **2012**, *4*, 101–128. [CrossRef]
- 18. Litman, T. Evaluating Public Transit Benefits and Costs; Victoria Transport Policy Institute: Victoria, BC, Canada, 2015.
- 19. Kurauchi, F.; Schmöcker, J.D. Public Transport Planning with Smart Card Data; CRC Press: Boca Raton, FL, USA, 2017.
- 20. Arbex, R.O.; da Cunha, C.B. Estimaç ao da matriz origem-destino e da distribuiç ao espacial da lotaç ao em um sistema de transporte sobre trilhos a partir de dados de bilhetagem eletrônica. *Rev. Transp.* **2017**, *25*, 166.
- Liu, X.; Zhou, Y.; Rau, A. Smart card data-centric replication of the multi-modal public transport system in Singapore. J. Transp. Geogr. 2019, 76, 254–264. [CrossRef]
- 22. Tsolas, I.E. Performance Evaluation of Electric Trolley Bus Routes. A Series Two-Stage DEA Approach. *Infrastructures* **2021**, *6*, 44. [CrossRef]
- Georgiadis, G.; Politis, I.; Papaioannou, P. Measuring and improving the efficiency and effectiveness of bus public transport systems. *Res. Transp. Econ.* 2014, 48, 84–91. [CrossRef]

- Karim, Z.; Fouad, J. Measuring urban public transport performance on route level: A literature review. In Proceedings of the International Workshop on Transportation and Supply Chain Engineering (IWTSCE'18), Rabat, Morocco, 8–9 May 2018; Volume 200, p. 00021.
- 25. Lee, E.H.; Lee, H.; Kho, S.Y.; Kim, D.K. Evaluation of transfer efficiency between bus and subway based on data envelopment analysis using smart card data. *KSCE J. Civ. Eng.* 2019, 23, 788–799. [CrossRef]
- 26. Seiford, L.M.; Zhu, J. An investigation of returns to scale in data envelopment analysis. Omega 1999, 27, 1–11. [CrossRef]
- 27. Lovell, C.K.; Rouse, A. Equivalent standard DEA models to provide super-efficiency scores. J. Oper. Res. Soc. 2003, 54, 101–108. [CrossRef]
- Färe, R.; Grosskopf, S.; Lindgren, B.; Roos, P. Productivity developments in Swedish hospitals: A Malmquist output index approach. In *Data Envelopment Analysis: Theory, Methodology, and Applications*; Springer: Berlin/Heidelberg, Germany, 1994; pp. 253–272.
- 29. Färe, R.; Grabowski, R.; Grosskopf, S.; Kraft, S. Efficiency of a fixed but allocatable input: A non-parametric approach. *Econ. Lett.* **1997**, *56*, 187–193. [CrossRef]
- 30. Daraio, C.; Diana, M.; Di Costa, F.; Leporelli, C.; Matteucci, G.; Nastasi, A. Efficiency and effectiveness in the urban public transport sector: A critical review with directions for future research. *Eur. J. Oper. Res.* **2016**, 248, 1–20. [CrossRef]
- Kuah, C.T.; Wong, K.Y.; Behrouzi, F. A review on data envelopment analysis (DEA). In Proceedings of the 2010 Fourth Asia International Conference on Mathematical/Analytical Modelling and Computer Simulation, Kota Kinabalu, Malaysia, 26–28 May 2010; pp. 168–173.
- 32. Cavaignac, L.; Dumas, A.; Petiot, R. Third-party logistics efficiency: An innovative two-stage DEA analysis of the French market. *Int. J. Logist. Res. Appl.* **2021**, *24*, 581–604. [CrossRef]
- 33. Caves, D.W.; Christensen, L.R.; Diewert, W.E. The economic theory of index numbers and the measurement of input, output, and productivity. *Econom. J. Econom. Soc.* **1982**, *50*, 1393–1414. [CrossRef]
- Berger, A.N.; Humphrey, D.B. Megamergers in banking and the use of cost efficiency as an antitrust defense. *Antitrust Bull.* 1992, 37, 541–600. [CrossRef]
- 35. Oliveira, R.; Zanella, A.; Camanho, A.S. A temporal progressive analysis of the social performance of mining firms based on a Malmquist index estimated with a Benefit-of-the-Doubt directional model. *J. Clean. Prod.* **2020**, *267*, 121807. [CrossRef]
- De Borger, B.; Kerstens, K.; Costa, A. Public transit performance: What does one learn from frontier studies? *Transp. Rev.* 2002, 22, 1–38. [CrossRef]
- Holmgren, J. The effects of using different output measures in efficiency analysis of public transport operations. *Res. Transp. Bus. Manag.* 2018, 28, 12–22. [CrossRef]
- CEN/TC320; Transportation-Logistics and Services-Public Passenger Transport-Service Quality Definition, Targeting and Measurement. CEN: Brussels, Belgium, 2002.
- 39. Transportation Research Board; Kittelson and Associates, Inc. *A Guidebook for Developing a Transit Performance-Measurement System;* Transportation Research Board: Washington, DC, USA, 2003; Volume 88.
- Dyson, R.G.; Allen, R.; Camanho, A.S.; Podinovski, V.V.; Sarrico, C.S.; Shale, E.A. Pitfalls and protocols in DEA. *Eur. J. Oper. Res.* 2001, 132, 245–259. [CrossRef]
- 41. Emrouznejad, A.; Amin, G.R. DEA models for ratio data: Convexity consideration. *Appl. Math. Model.* **2009**, *33*, 486–498. [CrossRef]
- 42. Sheth, C.; Triantis, K.; Teodorović, D. Performance evaluation of bus routes: A provider and passenger perspective. *Transp. Res. Part E Logist. Transp. Rev.* **2007**, *43*, 453–478. [CrossRef]
- 43. Güner, S.; Coşkun, E. Determining the best performing benchmarks for transit routes with a multi-objective model: The implementation and a critique of the two-model approach. *Public Transp.* **2016**, *8*, 205–224. [CrossRef]
- Tran, K.D.; Bhaskar, A.; Bunker, J.; Lee, B. Data Envelopment Analysis (DEA) based transit routes performance evaluation. In Proceedings of the Transportation Research Board (TRB) 96th Annual Meeting, Washington, DC, USA, 8–12 January 2017; pp. 1–23.
- 45. Taboada, G.L.; Han, L. Exploratory data analysis and data envelopment analysis of urban rail transit. *Electronics* **2020**, *9*, 1270. [CrossRef]
- Azambuja, A.M.V.d. Análise de Eficiência na Gest ao do Transporte Urbano por Ônibus em Municípios Brasileiros. Doctorate Thesis, Universidade Federal de Santa Catarina, Florianópolis, Brazil, 2002.
- 47. Sampaio, B.R.; Sampaio, Y.; Sampaio, L. Eficiência de Sistemas de Transporte público no Nordeste com Análise Envoltória de Dados (DEA). *Rev. Econ. Nordeste* 2006, *37*, 261–275.
- 48. Araújo, C.E.F.; Martins, F.G.D.; Silva, F. Análise Exploratória da Eficiência Operacional de Empresas do Transporte Rodoviário Interestadual de Passageiros por Ônibus no Brasil. In Proceedings of the XXII Congresso de Pesquisa e Ensino em Transportes–XII ANPET, Fortaleza, Brazil, 3–9 November 2008; Panorama Nacional de Pesquisa em Transportes–CD-ROM, Volume 3.
- 49. Sun, Y.; Shi, J.; Schonfeld, P.M. Identifying passenger flow characteristics and evaluating travel time reliability by visualizing AFC data: A case study of Shanghai Metro. *Public Transp.* **2016**, *8*, 341–363. [CrossRef]
- 50. Fitzová, H.; Matulová, M.; Tomeš, Z. Determinants of urban public transport efficiency: Case study of the Czech Republic. *Eur. Transp. Res. Rev.* **2018**, *10*, 1–11. [CrossRef]

- Fitzová, H.; Matulová, M. Comparison of urban public transport systems in the Czech Republic and Slovakia: Factors underpinning efficiency. *Res. Transp. Econ.* 2020, *81*, 100824. [CrossRef]
- 52. Zhu, W.; Yang, X.; Preston, J. Efficiency measurement of bus routes and exogenous operating environment effects on efficiency. *Transp. Plan. Technol.* **2016**, *39*, 464–483. [CrossRef]
- Zhang, Y.; Martens, K.; Long, Y. Revealing group travel behavior patterns with public transit smart card data. *Travel Behav. Soc.* 2018, 10, 42–52. [CrossRef]
- 54. Barabino, B.; Lai, C.; Olivo, A. Fare evasion in public transport systems: A review of the literature. *Public Transp.* **2020**, *12*, 27–88. [CrossRef]
- 55. National Academies of Sciences, Engineering, and Medicine. *Measuring and Managing Fare Evasion*; The National Academies Press: Washington, DC, USA, 2022.
- Lee, E.H. Exploring Transit Use during COVID-19 Based on XGB and SHAP Using Smart Card Data. J. Adv. Transp. 2022, 2022, 6458371. [CrossRef]
- 57. Barry, J.J.; Freimer, R.; Slavin, H. Use of entry-only automatic fare collection data to estimate linked transit trips in New York City. *Transp. Res. Rec.* 2009, 2112, 53–61. [CrossRef]
- Nunes, A.A.; Dias, T.G.; e Cunha, J.F. Passenger journey destination estimation from automated fare collection system data using spatial validation. *IEEE Trans. Intell. Transp. Syst.* 2015, 17, 133–142. [CrossRef]
- Trépanier, M.; Tranchant, N.; Chapleau, R. Individual trip destination estimation in a transit smart card automated fare collection system. J. Intell. Transp. Syst. 2007, 11, 1–14. [CrossRef]
- 60. Pearson, R.K. Exploratory Data Analysis Using R; CRC Press: Boca Raton, FL, USA, 2018.

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.