RAIRO-Oper. Res. 57 (2023) 3061–3071
https://doi.org/10.1051/ro/2023118

RAIRO Operations Research
www.rairo-ro.org

MALMQUIST INDEX EVALUATION OF COUNTRIES: 2000–2019

Hajar Farnoudkia*

Abstract. One of the sophisticated mathematical techniques for evaluating the relative efficiency of decision-making units (DMUs) in a multi-inputs-output setting is data envelopment analysis. A non-parametric productivity index called the Malmquist index (MI) tracks changes in a DMU's overall factor output over time. This study seeks to present a broad overview of the changes in the MI for nearly all of the world's nations over 19 years beginning in 2000. This research evaluates a time series data set made up of 16 economic indexes. Each nation's MI, which compares 2000 and 2019, shows how each nation has changed over that period. One of the most important purposes of this study comparing the countries by their MI is that it allows for a fair comparison of productivity across different time periods and regions. It also provides insights into the sources of productivity growth, such as changes in technology or improvements in efficiency. As a result, the countries can allocate resources more wisely and develop more effective investment plans by understanding the elements that lead to productivity increase. By implementing some statistical techniques, nations are further divided into four categories based on their MI. Furthermore, a yearly distribution of the MI has been included to illustrate its trends between the years 2000 to 2019. Finally, the changing flow for some countries of each category is shown in three-year tracks.

Mathematics Subject Classification. 91B38, 62H30.

Received March 13, 2023. Accepted August 4, 2023.

1. Introduction

The Malmquist index (MI) and Data Envelopment Analysis (DEA) are two well-liked techniques for assessing an organization's efficiency. The MI is a non-parametric technique for tracking the evolution of a manufacturing unit's productivity over time. The DEA is a linear programming method that assesses the relative efficiency of production units based on their inputs and outputs. This section reviews how these two methods have been used and applied in different studies.

1.1. DEA

Charnes *et al.* [4] originated the DEA method with the aim of comparing the efficiencies of various production divisions. The efficiency of a manufacturing unit is determined by comparing its inputs and outputs as follows.

$$\text{Efficiency} = \frac{\text{Outputs}}{\text{Inputs}}\cdot \tag{1}$$

Keywords. Data envelopment analysis, efficiency, Malmquist index, classification.

Department of Business Administration, Baskent University, Ankara, Turkey.
*Corresponding author: hajar.farnoudkia@gmail.com, hajerfarnoudkia@baskent.edu.tr

When a unit is compared to a reference group of presumed-efficient units, the technique determines a unit's efficiency. The outputs and inputs of the units in the sample are used to determine the reference set. Each unit in the sample has its inputs and outputs optimally weighted using linear programming, and DEA uses these weights to determine each unit's efficiency score.

DEA has been extensively used in a variety of disciplines, including engineering, operations research, and finance. In the literature, DEA has been used to assess the productivity of numerous production systems, including those of banks, hospitals, factories, and educational institutions. Additionally, the public sector has employed DEA to gauge the efficiency of government agencies and divisions. Numerous studies dating back to the early 1990s have been published in the vast literature on DEA. Charnes *et al.* [4] originated the first DEA model, known as the CCR, and then Banker *et al.* [1] formulated the BCC model that extends a variable return-to-scale technology. These two models have been used in many various contexts and are the most popular DEA models. Bao *et al.* [2] proposed the slacks-based measure is another significant addition to the field in addition to these two models. Several DEA model extensions have been suggested in recent years to address some of the limitations of the existing literature. Furthermore, numerous researchers have suggested different DEA applications to tackle contemporary issues. For instance, Tone [12] used DEA to assess the efficiency of Japanese institutions and hospitals, respectively. Toloo and Ertay [10] proposed an approach to cope with the vendor selection problem under price uncertainty in DEA. Toloo [9] extended a minimax approach for finding the most efficient professional tennis players. Toloo and Tavana [11] investigated DEA methods without explicit inputs or outputs and evaluated 46 association rules from data mining without the need to solve any optimization problem. Overall, DEA has solidified as a method for assessing the efficiency of decision-making processes in settings with multiple inputs and multiple outputs. As researchers create novel models and applications to handle problems in the real world, the body of literature on DEA grows.

1.2. Malmquist index

Swedish businessman Staffan B. Malmquist created the MI in 1953. It is a non-parametric productivity indicator that tracks the evolution of a production unit's total factor productivity (TFP) over time. TFP gauges a production unit's resource utilization efficiency by comparing the output of the unit to its inputs. The catch-up (CU) and the frontier-shift (FS) components are two distance functions that are combined to create the MI. These conditions are predicated on the efficiency of a DMU concerning the frontier at two different times.

Suppose that there are n DMUs to be evaluated, where each $\text{DMU}_{(j\in\{1,\dots,n\})}$ employs m semi-positive inputs $\boldsymbol{x}_j = (x_1, \dots, x_m)$ to produce s semi-positive outputs $\boldsymbol{y}_j = (y_{1j}, \dots, y_{sj})$. The notation $\text{DMU}_o^p = (\boldsymbol{x}_o, \boldsymbol{y}_o)^p = (\boldsymbol{x}_o^p, \boldsymbol{y}_o^p)$ stands for designating DMU_o in period $p \in \{1, 2\}$. The following model measures the technical efficiency measure of DMU_o^p with respect to the technology established in period q:

$$\begin{aligned} \theta^p((\boldsymbol{x}_o, \boldsymbol{y}_o)^q) = {} & \min \theta \\ & \text{s.t.} \\ & \sum_{j=1}^{n} \lambda_j x_{ij}^p \leq \theta x_{io}^q \qquad i = 1, \dots, m \\ & \sum_{j=1}^{n} \lambda_j y_{rj}^p \geq y_{ro}^q \qquad r = 1, \dots, s \\ & \lambda_j \geq 0 \qquad j = 1, \dots, n \end{aligned} \tag{2}$$

where λ_j is the intensity variable of DMU_j and $p, q \in \{1, 2\}$. In fact, $\theta^p((\boldsymbol{x}_o, \boldsymbol{y}_o)^q)$ is the efficiency of $\text{DMU}_o^q = (\boldsymbol{x}_o, \boldsymbol{y}_o)^q$ with respect to frontier q. Hence, $\theta^1((\boldsymbol{x}_o, \boldsymbol{y}_o)^1)$ is the technical efficiency of DMU_o at the 1st time period, and analogously $\theta^2((\boldsymbol{x}_o, \boldsymbol{y}_o)^2)$ is the technical efficiency of DMU_o at the 2nd time period. In this case,

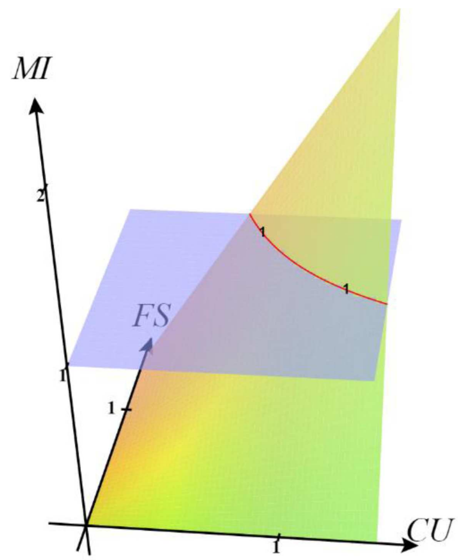

FIGURE 1. The geometric representation of MI.

the CU and FS indexes are written below, respectively.

$$\mathrm{CU}_o = \frac{\theta^2((\boldsymbol{x}_o, \boldsymbol{y}_o)^2)}{\theta^1\left((\boldsymbol{x}_o, \boldsymbol{y}_o)^1\right)} \tag{3}$$

$$\mathrm{FS}_o = \sqrt{\frac{\theta^1\left((\boldsymbol{x}_o, \boldsymbol{y}_o)^1\right)}{\theta^2\left((\boldsymbol{x}_o, \boldsymbol{y}_o)^1\right)} \times \frac{\theta^1\left((\boldsymbol{x}_o, \boldsymbol{y}_o)^2\right)}{\theta^2\left((\boldsymbol{x}_o, \boldsymbol{y}_o)^2\right)}}. \tag{4}$$

The MI of the unit under evaluation is

$$\mathrm{MI}_o = \mathrm{CU} \times \mathrm{FS}_o. \tag{5}$$

A sufficient condition to have $\mathrm{MI}_o > 1$ is $\mathrm{CU}_o > 1$ and $\mathrm{FS}_o > 1$. In other words, if both the FS and CU indexes of a DMU exceed 1, then the productivity of the DMU has improved from Period 1 to Period 2. In the case of $\mathrm{MI}_o = 1$, the DMU under evaluation's productivity level remains constant, while in the case of $\mathrm{MI} < 1$, the DMU deteriorates with respect to its productivity from time t_1 to time t_2. In this case, at least one of the MI components, *i.e.*, CU and FS is less than one. Figure 1 shows the relationship between CU (x-axis), FS (y-axis), and MI (z-axis) which is calculated by the multiplication of the values in the x and y-axis, as explained before. The red curve on the page is the isoquant, where $\mathrm{MI} = 1$. So, the values above the isoquant indicate an improvement of the related DMU and the values under the isoquant show the deterioration of the corresponding DMU.

The literature has made extensive use of the MI to evaluate output and efficiency across numerous sectors. For instance, the MI was used in research by Houshyar *et al.* [6] to evaluate the productivity growth of pomegranate production. According to the study, the productivity of farms grew over time, and the MI Index was a useful tool for assessing the efficiency of the farms. The MI was used in different research by Coelli and Rao [5] to gauge the agricultural efficiency and productivity of 93 countries. According to the research, the MI was a helpful tool for assessing shadow prices and value shares.

2. Application

The data set includes 134 nations that are represented by 16 economic indexes from 2000 to 2019, first used to compare the situation of Turkey in the world [3]. The titles and definitions of the indexes are provided in Table 1.

DEA evaluates the data, treating each nation as a DMU, and determines how each nation's efficiency changed between the years 2000 and 2019. The indices are all regarded as output because, according to the meaning of output, the more, the better. The GAMPS software was used to acquire the countries' CU and FS rates, and their multiplication results in the countries' MI. Only the first and last years are taken into account in the first part of this section to give a general idea about the countries' levels of MI growth, and the second part is devoted to MI in 3-year steps between 2001 and 2019. The MI was therefore calculated over one year, from 2000 to 2001, from 2001 to 2002, etc in the last part of this section.

2.1. Data preprocess

The first version of the data was for 140 countries with some missing values and different scales with respect to the definition of the economic index. 6 countries with a high level of missing data were removed from the analysis. The small percentage of the missing values that exist for some countries were filled by time series estimation techniques like moving average (q) that is, taking the average of q cells before the missing to estimate a value for the missing cell. Then, all indices were scaled into 0 and 100 by a function of $f(x_{ij}) = \{\frac{(x_{ij}-\min(X_j))}{(\max(X_j)-\min(X_j))} \times 99\}+1$ where small letter x_{ij} is the ith value of jth index (ith row and jth column) and $\min(X_j)$ is the minimum of the jth index *i.e.*, the lowest value of the jth column and similarly, $\max(X_j)$ is the maximum of the jth index *i.e.*, the highest value of the jth column. Note that the MI in scale-invariant means that the linear transformation of the data does not have any effect on MI.

2.2. Single 19-year period MI

The descriptive statistics of the MI and its two elements are represented in Table 2. It is seen that the mean of MI is more than 1 showing that the countries are generally improved. Also, 0.12 as the standard deviation value states that most of the MI of the countries are distributed in 0.90 and 1.14. The lower mean of the FS variable in comparison with MI and CU, states that the frontier changed a lot from 2000 to 2019, and for most countries, the value of FS is between 0.94 and 1.02 according to the small standard deviation. The values of CU and MI are close to each other, which verifies that the effect of CU is greater than FS for this data set.

The box plot of the three variables is represented in Figure 2, showing that the distribution of FS and MI is symmetric without considering the outliers, but the distribution of CU is right-skewed as can be approved by comparing the mean and the median value. So, half of the countries' CU is less than 1 where the median for the FS and MI are 0.99 and 1.01, respectively. As this study focuses on MI, the down outlier country is Gabon with a minimum MI that is 0.7414 and the upper outliers are Sierra Leone, Chad, Burundi, Brazil, Romania, Bulgaria, Ghana, and Venezuela, in order.

The countries' MI values are represented in Figure 3 by different colors and marks based on their classes. There are also horizontal and vertical lines for the value of 1, to separate the countries according to their development status from 2000 to 2019. In this part, the MI is calculated by setting t_1 to 2000 and t_2 to 2019. The countries are also classified into four groups using the confidence interval concept for normal distribution. The ISO code of the countries is used instead of their complete name for simplicity.

The countries with MI very less than 1 mean that has considerably deteriorated over the past 19 years are represented by the blue points as the first class on the map. The maximum MI for this class is 0.8454. The second class, depicted in orange, consists of MI = 1 nations with negligible regressive trends. The third class with a maximum of MI = 1.144, is depicted in grey and is made up of nations with MI > 1 showing some degree of development. The last class, which is depicted in yellow, is made up of nations with significant growth between 2001 and 2019. Figure 3 also shows that nations with both a CU and an FS that are less than 1 are

TABLE 1. Data descriptions.

No	Index name	Explanation
1	Foreign direct investment	A permanent management stake (10% or more of the voting shares) in a company working in an economy other than the investor's economy.
2	The globalization	The Commercial, Social, and Political Facets of Globalisation.
3	Freedom to trade internationally	Trade regulations, black market exchange rates, and controls on the movement of people and money are all examples of tariffs. Each element receives a number ranging from 0 to 10. The degree of independence in global trade is the average of its subcomponents.
4	Human development	A truncated indicator of typical performance in significant areas of human development is having knowledge, a high standard of living, and a long and healthy existence.
5	Current account balance	Total net exports of products and services, net primary income, and net secondary income.
6	Size of government	The average of five parts: public spending, transfers, and subsidies, state investments and enterprises, the highest marginal tax rate, and assets held by the government.
7	General government gross debt	The total amount of unpaid direct government fixed-term contractual responsibilities to others as of a given date constitutes the debt. It contains both domestic and foreign assets, such as loans, non-stock securities, and foreign currency and currency deposits.
8	General government net debt	Money and deposits, debt securities, loans, insurance, pension, and standardized guarantee plans, as well as other accounts due, are among these financial assets. Monetary gold and SDRs are also included.
9	Gross fixed capital formation	Composed of total changes in stock levels and additions to the economy's fixed assets. The purchase of plant, machinery, and equipment, the building of roads, railways, and similar structures; land improvements (fences, canals, etc.), and the construction of buildings such as schools, offices, hospitals, private homes, and commercial and industrial structures are all examples of fixed assets.
10	Economic freedom	Measures the degree of economic freedom in five main areas: public size, legal system, security of private rights, solid money, international trade independence, and regulation.
11	Sound money	The average of four aspects of access to solid money: money growth, inflation standard deviation, inflation in the most recent year, and freedom to maintain a foreign exchange bank account.
12	The legal system and property rights	The average of the legal system and property rights, including judicial independence, the impartiality of the courts, protection of property rights, military interference in politics and the rule of law, legal application of contracts, communication costs associated with real estate transactions, police credibility, business costs associated with crime, and gender inequality reduction.
13	Rule of law	It measures how much a representative's perceptions of the quality of contract execution, property rights, the police, and courts, as well as the probability of crime and violence, correspond to their perceptions of how much they trust and uphold social norms. Indicator (-2.5 weak, 2.5 strong).
14	Regulation	The average of three categories of regulations: those governing the credit market, those governing the labor market, and those governing business.
15	Political stability	The index gauges how likely it is that the government will be overthrown or destabilized through violent or illegal means, such as terrorism and politically driven violence (2.5 robust, 2.5 weak).
16	GDP per capita	Gross Domestic Product is calculated as the GDP split by the midyear population. GDP is calculated as the total of all product taxes, all product subsidies, and the overall value contributed by all industry incumbent producers. calculated without taking into account wear and tear on manufactured assets or deterioration and loss of natural resources.

TABLE 2. Descriptive statistics of various indexes.

Variable	Mean	StDev	Minimum	Median	Maximum
CU	1.03	0.11	0.74	1.00	1.40
FS	0.98	0.04	0.85	0.99	1.17
MI	1.02	0.12	0.70	1.01	1.36

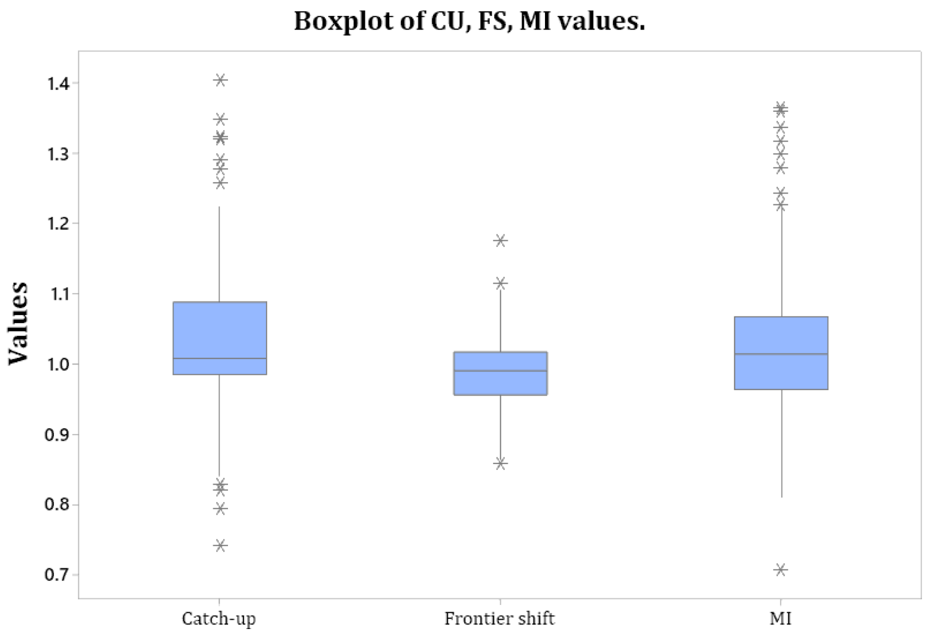

FIGURE 2. Box plot of MI, CU, and FS values of 134 countries for $t_1 = 2000$ and $t_2 = 2019$.

classified in classes 1 or 2, while those with more than one CU and FS are classified in classes 3 or 4. Figure 4 depicts the distribution of MI as a measure of development in a map plot. It can be seen that most of the deteriorated countries are in the East, and the decliners are particularly in the Middle East. Additionally, the western hemisphere is among the third and fourth classes of regions with MI > 1.

Additionally, Tukey pairwise comparison, which is one of the parts of the Analysis of Variance test (ANOVA), is employed to assess the distinguishing power of CU and FS. It is clear from Table 3 that CU could distinguish the difference between the classes by mean values, meaning that classification by only CU would not be unsuccessful. In other words, the discrimination power of CU and MI is not significantly different while FS could classify the countries into two groups: Classes 1–3 on one side and Class 4 on the other side according to Table 3.

Figure 5 approves the distinguishability of CU like MI while it shows the common intervals between the classes in the interval plot of FS of four classes indicating the weakness of FS in discriminating the classes.

2.3. Malmquist index over three years

Additionally, Figure 6 shows the 3-year MI of some countries selected from each category. Notice that MI $\gg 1$ is typically the case for countries in the fourth cluster, such as China and Cyprus, whereas MI $\ll 1$ is typically

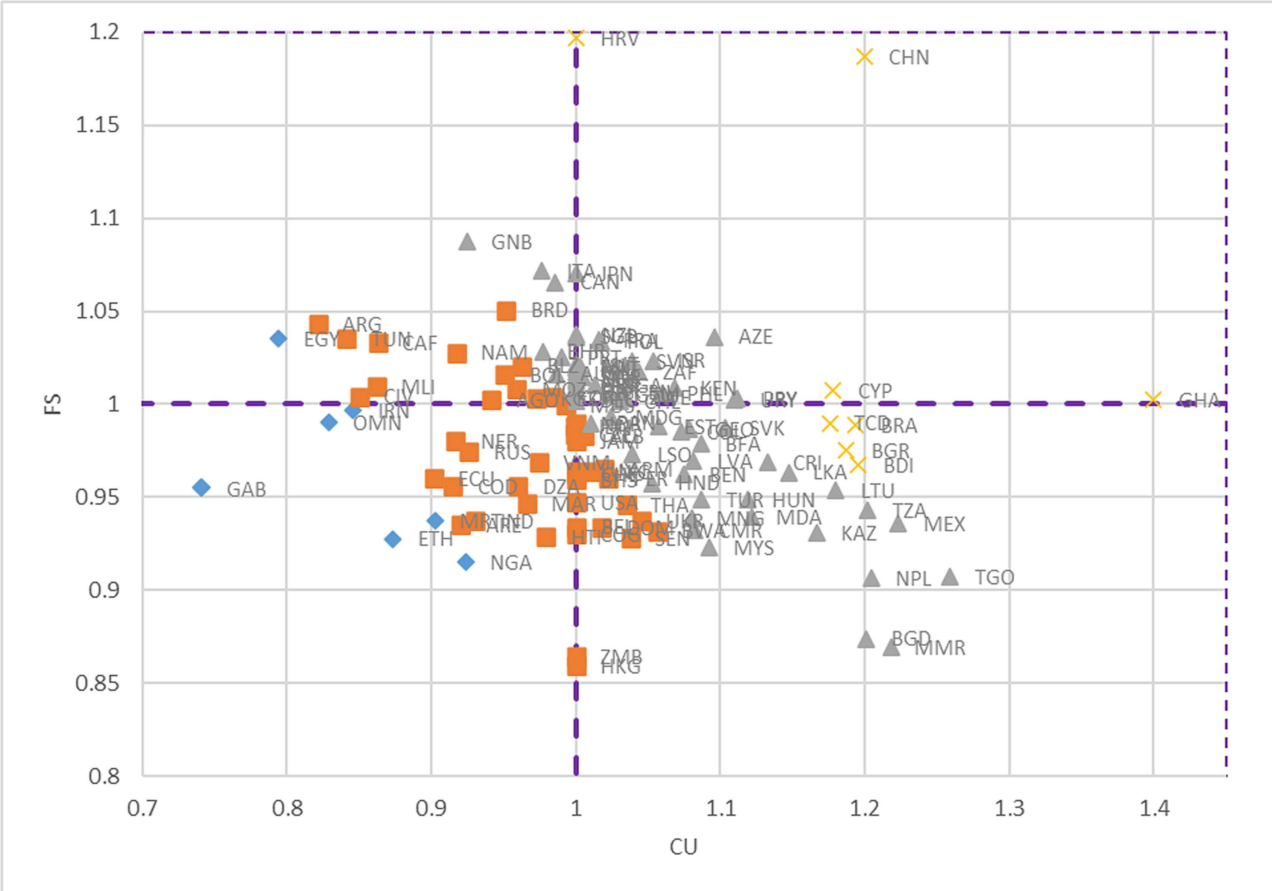

FIGURE 3. The classification plot of the countries CU on the x-axis and FS on the y-axis for $t_1 = 2000$ and $t_2 = 2019$.

the case for countries in the first cluster, such as Iran. There is no particular regulation for the second ($\text{MI} \leq 1$) and third clusters ($\text{MI} \geq 1$), where Turkey and the USA are located, respectively.

2.4. Malmquist index over a year

As stated previously, the data are annual data that began in 2001 and concluded in 2019. The percentage of nations with $\text{MI} > 1$ is therefore calculated each year and displayed in Figure 7 to help visualize how the countries' patterns of growth are changing.

Figure 7 shows the percentage of developing countries each year ($\text{MI} > 1$) (%). Conclusion: The highest level of improvement shown by the blue point in Figure 6 is 87.3%, was observed between 2004 and 2005, and was 82% in 2014–2015. The minimum increase in MI, which is only 19.4%, is seen between 2002 and 2003 shown by the green point in Figure 7. Furthermore, it is evident that when the improvement level is great for a year, it is typically low the following year. It makes perfect sense that maintaining or improving a nation with higher indexes would be more challenging than doing the same for a country with lower indexes.

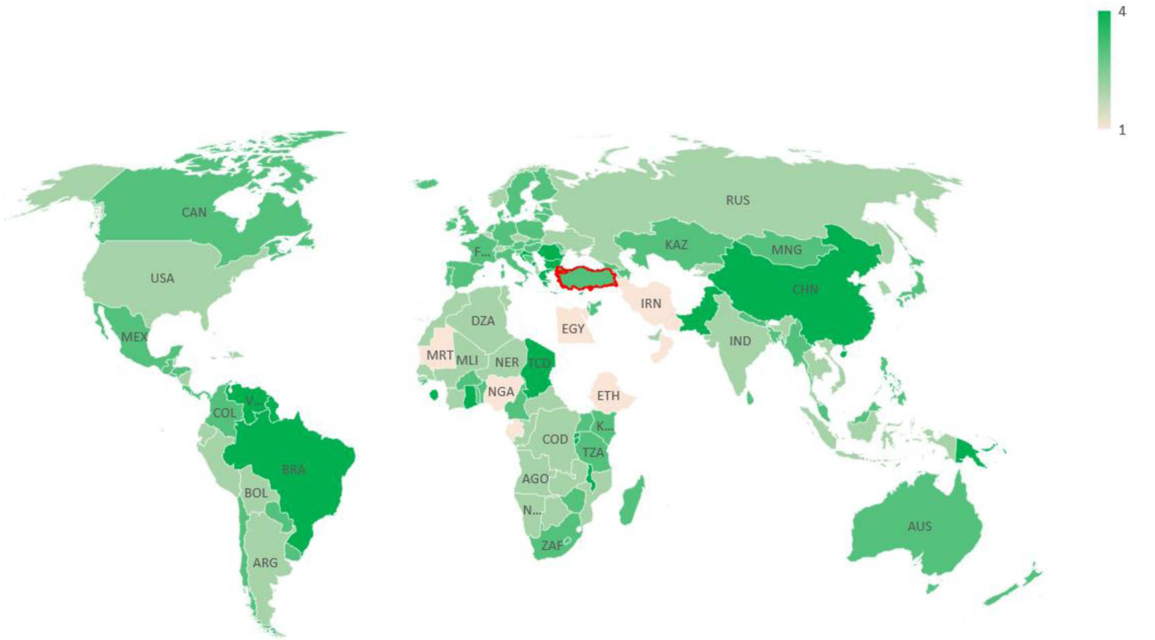

FIGURE 4. The map plot of the MI distribution of 134 countries classified into 4 categories.

TABLE 3. Grouping information of CU and MI.

Class	N	Mean (CU)	Mean (MI)	Grouping	Mean (FS)	Grouping
1	7	0.84	0.81	A	0.96	A
2	45	0.96	0.93	B	0.97	A
3	64	1.05	1.04	C	0.99	A
4	18	1.22	1.25	D	1.02	B

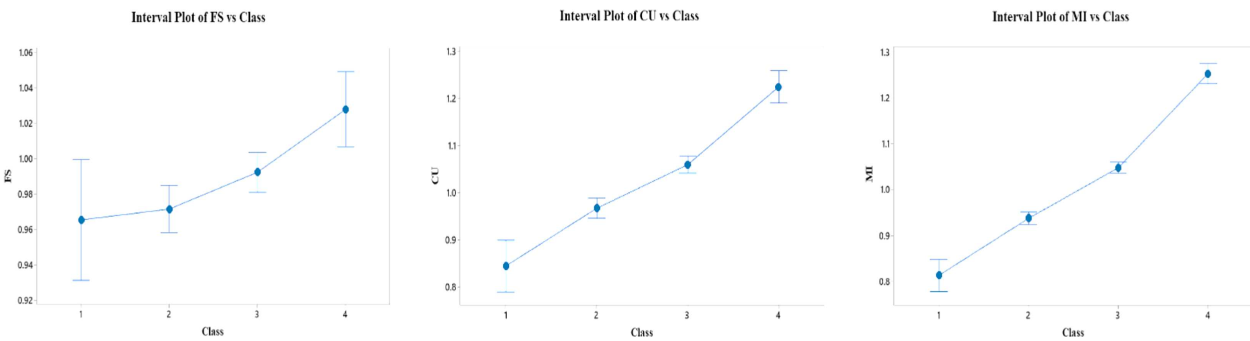

FIGURE 5. The interval plot of FS, CU, and MI, in order.

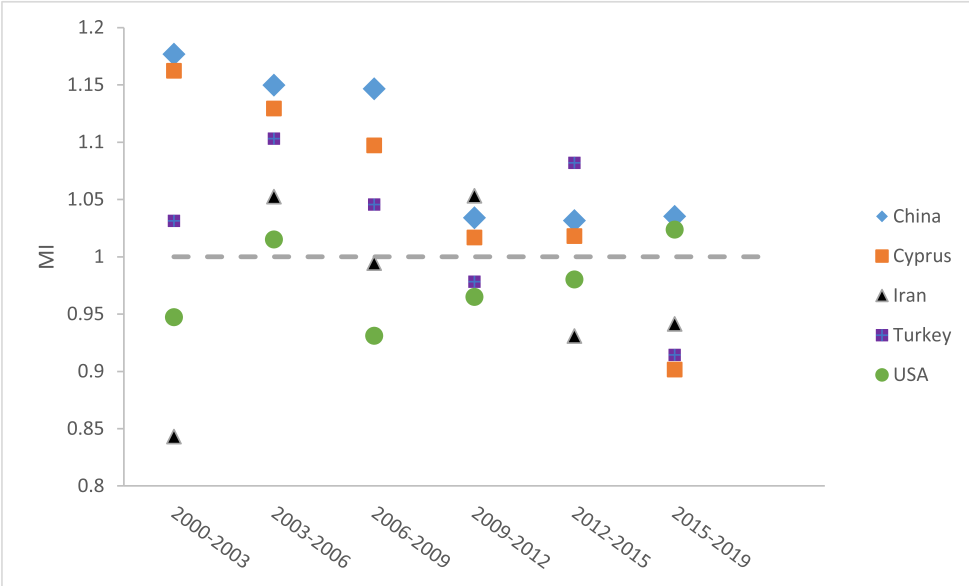

FIGURE 6. MI of five countries in three years from 2000 to 2019.

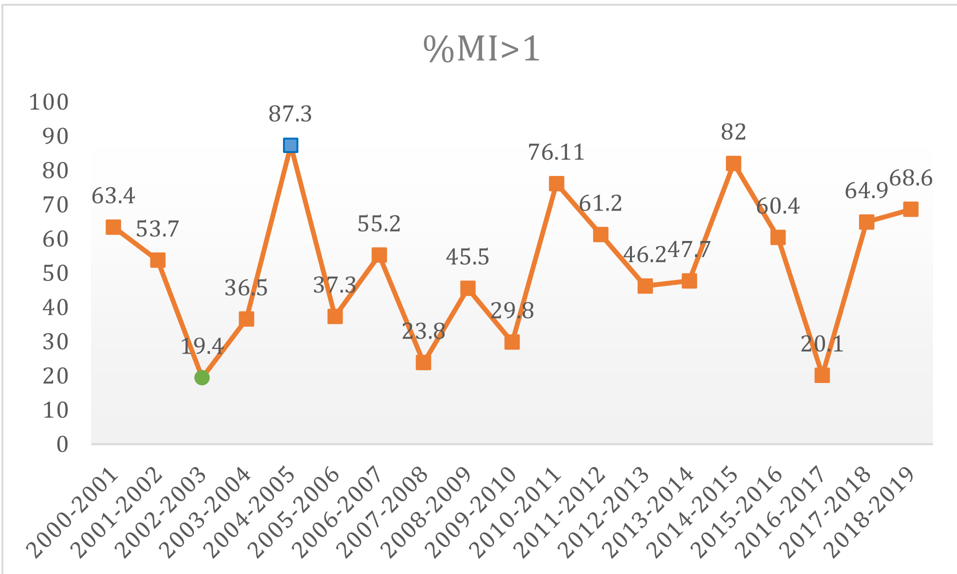

FIGURE 7. The percentage of countries with MI > 1 from 2000 to 2019.

3. DISCUSSION AND CONCLUSIONS

MI is a well-liked technique for assessing an organization's efficiency by tracking changes in a manufacturing unit's productivity over time. This study used MI to evaluate all nations in the world, with the exception of a few, in 16 economic variables from 2001 to 2019 at various time points. The first calculation is the difference between the beginning year, 2001, and the most recent time point, 2019. Based on our observations in this study, it has been noted that the majority of nations that have experienced a decline based on their MI between the years 2001 to 2019 (MI $\ll 1$) are situated in the eastern hemisphere, particularly in the Middle East region. Furthermore, the nations have been classified into four distinct groups based on their MI in comparison to a value of 1. On a map plot, each country's productivity change level is represented by a different hue. Our analysis has revealed that the distribution of the CU, which is one of the elements of the MI is consistent with the distribution of MI. However, this is not the case for the other element of MI, which is the FS. This implies that if countries were grouped based on their CU, they would belong to the same class as the grouping based on MI. Moreover, MI is also calculated in 1-year and 3-year increments to provide more information regarding the productivity of the nations between 2001 and 2019. Additionally, it is worth noting that a sharp decline in the percentage of countries that show improvement in the following year is a reasonable trend when most nations experience development in a single year. This can be attributed to the fact that if a country has achieved development in a given year, it becomes increasingly challenging to achieve further development in the subsequent year. Conversely, if a country has experienced deterioration in a given year, it becomes comparatively easier to achieve development in the following year through minor changes. In summary, this study would have been more effective by identifying the weak points that allow countries to be able to focus on improving their productivity or adopt similar strategies to those of consistently developed countries.

In future studies, the weights of each of the 16 indices will be investigated for countries based on their current status as either developed or developing nations. This will enable the identification of strong factors that contribute to productivity in developed countries, as well as weak factors that hinder productivity in developing nations. Additionally, since uncertainty is an undeniable fact in our normal life, employing a robust DEA approach to investigate the efficiency changes of countries is an interesting research direction. For a deeper discussion of robust DEA methods, we refer the reader to Salahi *et al.* [7, 8].

References

[1] R.D. Banker, A. Charnes and W.W. Cooper, Some models for estimating technical and scale inefficiencies in data envelopment analysis. *Manage. Sci.* **30** (1984) 1078–1092.

[2] C.P. Bao, T.H. Chen and S.Y. Chang, Slack-based ranking method: an interpretation to the cross-efficiency method in DEA. *J. Oper. Res. Soc.* **59** (2008) 860–862.

[3] N. Çakir and G. Küçükkocaoğlu, The Last Twenty Years of Economic Growth in Turkey (2000–2020) Turkey-China Comparative Analysis, 1st edition. Siyasal Kitabevi, Ankara (2022).

[4] A. Charnes, W.W. Cooper and E. Rhodes, Measuring the efficiency of decision-making units. *Eur. J. Oper. Res.* **2** (1978) 429–444.

[5] T.J. Coelli and D.S.P. Rao, Total factor productivity growth in agriculture: a Malmquist index analysis of 93 countries, 1980–2000. *Agric. Econ.* **32** (2005) 115–134.

[6] E. Houshyar, M. Mahmoodi-Eshkaftaki and H. Azadi, Impacts of technological change on energy use efficiency and GHG mitigation of pomegranate: application of dynamic data envelopment analysis models. *J. Cleaner Prod.* **162** (2017) 1180–1191.

[7] M. Salahi, M. Toloo and Z. Hesabirad, Robust russell and enhanced russell measures in DEA. *J. Oper. Res. Soc.* **70** (2019) 1275–1283.

[8] M. Salahi, M. Toloo and N. Torabi, A new robust optimization approach to common weights formulation in DEA. *J. Oper. Res. Soc.* **72** (2021) 1390–1402.

[9] M. Toloo, Alternative minimax model for finding the most efficient unit in data envelopment analysis. *Comput. Ind. Eng.* **81** (2015) 186–194.

[10] M. Toloo and T. Ertay, The most cost-efficient automotive vendor with price uncertainty: a new DEA approach. *Measurement* **52** (2014) 135–144.

[11] M. Toloo and M. Tavana, A novel method for selecting a single efficient unit in data envelopment analysis without explicit inputs/outputs. *Ann. Oper. Res.* **253** (2017) 657–681.

[12] K. Tone, A strange case of the cost and allocative efficiencies in DEA. *J. Oper. Res. Soc.* **53** (2002) 1225–1231.