Article

An analysis of the eco-efficiency of the agricultural industry in the Brazilian Amazon biome

Gabriela Mayumi Saiki * , André Luiz Marques Serrano * , Gabriel Arquelau Pimenta Rodrigues , Carlos Rosano-Peña , Fabiano Mezadre Pompermayer , Pedro Henrique Melo Albuquerque , Mario Del Roble Pensado-Leglise

Posted Date: 10 June 2024

doi: 10.20944/preprints202406.0411.v1

Keywords: Agricultural Industry; Amazon; Bootstrap; Eco-efficiency; DEA

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 10 June 2024 doi:10.20944/preprints202406.0411.v1

Article

An Analysis of the Eco-Efficiency of the Agricultural Industry in the Brazilian Amazon Biome

Gabriela Mayumi Saiki [1,*], André Luiz Marques Serrano [1,*], Gabriel Arquelau Pimenta Rodrigues [1], Carlos Rosano-Peña [2], Fabiano Mezadre Pompermayer [3], Pedro Henrique Melo Albuquerque [2] and Mario Del Roble Pensado-Leglise [4]

1 Professional Post-Graduate Program in Electrical Engineering (PPEE), Department of Electrical Engineering (ENE), University of Brasilia(UnB), Technology Faculty, Brasilia, 70910-900, DF, Brazil; mayumisaiki.gms@gmail.com (G.M.S.); andrelms@unb.br (A.L.M.S.); gabriel.arquelau@redes.unb.br (G.A.P.R.);
2 Faculty of Economics, Administration and Accounting (FACE), Department Administration (ADM), University of Brasilia (UnB), Brasilia, 70910-900, DF, Brazil; crosano@unb.br (C.R-P.); pedro.albuquerque@bellevuecollege.edu (P.H.M.A.)
3 Institute for Applied Economic Research (IPEA), Brasilia, 70390-025, DF, Brazil; famepomper@hotmail.com (F.M.P.)
4 Instituto Politécnico Nacional: Ciudad de México, Distrito Federal, México; mpensado@ipn.mx (M.D.R.P-L.)
* Correspondence: mayumisaiki.gms@gmail.com; andrelms@unb.br

Abstract: The exponential growth of the agricultural industry in the Amazon region has brought about notable economic advancements. However, this growth has substantially cost the region's ecosystems, manifesting in increased deforestation and biodiversity degradation within the Amazon forest. This article is dedicated to assessing the eco-efficiency of agricultural production in Amazon biome municipalities. It places particular emphasis on identifying critical determinants through the utilization of the classic Data Envelopment Analysis (DEA) model for efficiency computation, super-efficiency models for distinctive characterization, bootstrap computational techniques for robust resampling, and the Malmquist index for calculating annual eco-efficiency indices of each Decision-Making Unit (DMU). An exploration of the correlation between efficiency and meteorological attributes of the municipalities is conducted. The findings of this study reveal several significant points: 1) Eco-efficient municipalities within the Amazon Biome can serve as benchmarks for other DMUs striving to attain optimal input-output levels. 2) Most municipalities in the Amazon Biome operate closely to the productive frontier due to the prevalent technology employed in their agricultural activities. 2) The nature of the technological frontier's return suggests that small and large DMUs possess eco-efficiency potential. 4) The current dataset does not yield conclusive evidence regarding a direct correlation between the variables. Leveraging this information, strategic pathways can be formulated to drive economic development in tandem with the sustainability of Amazon Biome municipalities. These strategies promise to foster social, economic, and environmental benefits for the populace while providing valuable insights to inform future research within this thematic domain.

Keywords: agricultural industry; Amazon; bootstrap; eco-efficiency; DEA

1. Introduction

The issue of sustainable development became significantly relevant after the publication of the Club of Rome's report, "The Limits to Growth" [1]. The report warned that the planet could not support continued economic growth without exceeding the Earth's carrying capacity due to the increasing pressure on energy and natural resources, pollution, and climate change. This prompted the United Nations to establish conferences and commissions worldwide to study and discuss sustainable development and develop commitments and solutions.

Thus, in the Brundtland Report published by the United Nations World Commission on Environment and Development, entitled "Our Common Future" [2], the definition of sustainable development emerged based on the concept of the "triple bottom line" of sustainability. The three dimensions are: i) Economic dimension, which refers to the traditional search for profit and financial value creation.

 doi:10.20944/preprints202406.0411.v1

It involves assessing the organization's ability to be financially healthy. ii) Social dimension: The production unit must consider its impact on society and the communities in which it operates. This can include corporate social responsibility, the quality of relations with employees, customers, suppliers, and the contribution to social development. iii) Environment dimension: The organization assesses its environmental impact and contribution to long-term sustainability. This involves issues such as reducing greenhouse gas emissions, managing natural resources efficiently, minimizing waste, and adopting sustainable practices and products, meaning that an organization's success should not only be measured by its financial performance but also by how it contributes to people's well-being and the preservation of the environment.

Since its inception, the concept of sustainable development has been adopted in the theoretical literature. It is practiced by companies and countries seeking a balance between financial prosperity, social responsibility, and environmental preservation. From this perspective, we can highlight the study by [3], where the authors developed a survey with 170 companies in the agricultural industry of small and large multinational corporations in Germany related to social and environmental dimensions. In addition, other studies have proposed the construction of multidimensional sustainability indicators to evaluate the performance of the agricultural industry's production system. Some of them can be highlighted as: [4–7].

Moving in this direction, the 26th Conference of Nations on Climate Change, which took place in Glasgow, Scotland, referred to the conservation of forests. During this conference, according to the United Nations website[1], 110 countries that are home to 85% of the world's forests signed a declaration committing to halt and reverse deforestation by 2030. This declaration was an essential step towards the preservation of the environment. However, the most challenging task lies in achieving this regional and local commitment without compromising the sustainable development and resilience to climate change of families who live and work in forests. They are producers with limited resources to maintain and improve productivity and competitiveness. They are marginalized from the primary social services and impacted by the independence of adult children who give up and migrate to the urban environment or are fragmenting family property [8].

As a result, introducing ever stricter environmental protection regulations has caused great concern due to its social effects. This situation is no different in the Brazilian Amazon rainforest, where decision-makers face two challenges. This is due to the need to make communities more productive and efficient to offer more products with higher quality and competitive prices, and thus be able to face direct competition with agriculture on an industrial scale [9]. The second challenge arises from the evidence that agricultural intensification and expansion in forest areas can generate irreversible environmental damage that economic benefits may not offset [6,10].

Removing native primary vegetation cover is known to compromise the ecosystem service of flora and fauna, which also risks degrading. In addition, deforestation causes erosion, a reduction in soil nutrients, and an increase in greenhouse gas (GHG) emissions, which interfere with rainfall and the planet's temperature. This creates a vicious cycle, as changes in temperature and rainfall patterns can harm agricultural activity itself [8,11]. There is also the fact that increased territorial expansion and human exploitation of the environment are increasingly depleting native forests. Climate change is known to cause direct effects, such as changes in soil moisture and temperature regimes. Climate change also has immediate effects, such as changes in soil moisture and temperature regimes. However, these changes also cause what [12] calls indirect effects, such as an increase in carbon dioxide (CO_2) in the atmosphere and soil dethatching (N_2).

In the Amazon biome, this phenomenon is no different, making the issue of deforestation one of the main problems facing Brazil [13]. There are several consequences of deforestation in the Amazon. Among these consequences, we can see a decrease in rainfall (220 to 640 $mm/year$), evaporation (164

---

[1] https://news.un.org/en/story/2021/11/1104642

 doi:10.20944/preprints202406.0411.v1

to 500 $mm/year$), a slight increase in temperature (0°C to 3°C) [6,14–18]. From these inferences, it can be inferred that the impact on the biome is significant, as it causes changes in other regions and biomes across the planet, such as a decrease in precipitation, an increase in temperature, and thus even a decrease in the efficiency of local agriculture [19]. Therefore, a rapid change is needed in how these natural resources are used and preserved [20].

Another factor causing deforestation in the Amazon is mechanized farming, which promotes burning [21]. The typical fires that cause deforestation are called high-frequency fires. These fires have contributed to more than 40% of the fires detected in recent years, while on the Bolivian border and in the states of Mato Grosso, Pará, and Rondônia this figure is as high as 84% [21]. However, it is interesting to review how the agricultural industry operates and how it would be possible to improve the use of resources without failing to produce what is needed more efficiently. For this reason, it is impossible to think of reducing production to protect the environment. Instead, it is necessary to think of ways to develop production sustainably by introducing new technologies and improving planting methods.

This means that production must be performed in a way that does not harm the environment and consumes as lesser resources as possible. For this reason, one of the keywords in this research is efficiency. By efficiency, we mean production that delivers the maximum output level while spending as few resources as possible [22], thus reaching the optimum point between the ratio of input and output within the production frontier. In this way, it is expected to produce as much food as possible, consuming as few environmental resources and emitting as little CO_2 as possible. DEA will be used to assess efficiency.

This method consists of a non-parametric mathematic that, using linear programming algorithms, can compare entities that carry out similar processes of transforming *inputs* into *outputs* to define an efficiency index between them, in which the most efficient entity compared receives the index with a value equal to 1. The others receive a score below one, proportional to how less efficient it is. As a traditional technique, DEA is being popularized and has gained a lot of relevance, especially in research and academic circles. DEA is the only way to evaluate the efficiency of a transformation process by considering multiple inputs and outputs without having to transform these components into a single value, such as financial value.

In this context, this research aims to contribute to these issues by showing the possibility of producing efficiently with less environmental impact and less use of resources in agriculture in the municipalities that compose the Brazilian Amazon biome, using the latest approaches to computational models such as the bootstrap method and the stochastic frontier analysis (SFA) method. The estimates then define useful indicators for formulating and evaluating sustainability policies, such as parameters for rewarding generators of positive environmental externalities. Another purpose is to identify the correlation between weather and agricultural production efficiency. To this end, a technique for comparing efficiency over time, the *Malmquist* technique, will be used. This method also aims to examine the determinants of eco-efficiency and calculate an environmental sustainability index, meeting the Pareto optimum while simultaneously meeting economic and environmental objectives based on the best practices in the region.

2. Theoretical Framework

Efficiency is a concept based on economics, mathematics, and operational research to reach an optimum point between the ratio of inputs to be transformed (x) and the outputs generated (y), given a technological set (T) so that $T = \{(x, y)|x \text{ can produce } y\}$, in a concept of production possibilities (CPP), whose properties are described by [23]. In this way, eco-efficiency represents the factors involving economic, social, and environmental issues that will form the tripod to make up the inputs and outputs of the production process. The inputs, x, will represent the ecological costs necessary to generate the desirable products y [24]. In theory, there is a minimization of inputs (x) and a maximization of outputs (y). Nonetheless, this relationship is not generalized for cases of desirable inputs, such as

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 10 June 2024 doi:10.20944/preprints202406.0411.v1

preserved areas, and undesirable outputs, such as CO_2 emissions; this relationship is inverted given a technological set $T = \{(x, y)|x \text{ can produce } 1/y\}$.

The DMUs (Decision-Making Units) - inefficient DMUs are inserted within the CPP, formed by linear segment combinations of efficient DMUs that include a convex figure. However, they do not present an optimal performance between the ratio of inputs and outputs [25]. The upper part of the frontier segment represents the (T^{δ}) of the CPP, formed by a negative space R^{p+q} and structured by the x and y ε (R^q) of the DMUs. From this discussion, the concept of eco-efficiency emerges. It is achieved when the set of n DMUs (Decision-Making Units) that form the production frontier, composed of x and y, present the lowest possible quantity of inputs (related to the environmental cost of the process) and the highest possible amount of outputs (related to the impact that activities can have on the environment).

This distance between the inefficient point and the benchmarks made up of the frontier of efficient DMUs is called the Euclidean distance. Therefore, to achieve technical efficiency, it can be input-oriented (when inputs are minimized, without changing the level of outputs) and output-oriented (when outputs are maximized, without changing the status of inputs) [25], where the inverse of technical efficiency is the radial efficiency of [24]. Thus, the DEA model is a mathematical model that, by receiving the inputs and outputs variables of a process, can define how efficient the entity carrying it out is [26].

To do this, the model uses several entities that carry out the same transformation process to, in a comparative analysis, find the limit of efficiency and how each of these entities (DMUs) is classified about this limit. With enough DMUs, creating a scatter plot with all these DMUs distributed along a Y-axis, representing the outputs that a given DMU produces, is possible. The X-axis is the axis that quantifies the inputs that this DMU needs to consume to reach this production level. Thus, the Figure 1 is a graphic example in which DMUs A, B, C, D, and E are considered Benchmarks of excellence, i.e., those that have achieved Technical Efficiency [25].

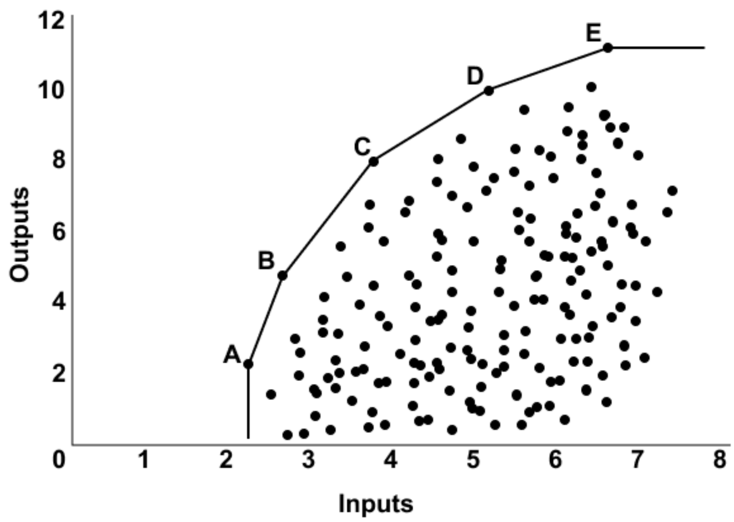

Figure 1. Scatter plot of DMU - DEA

Still analyzing the 1 graph, it can be seen how much more the DMU_E would need to produce to achieve technical efficiency and how much less it should consume, or even a combination of the two. In this context, the efficiency frontier is then generated by drawing a line between the points where it is positioned on the graph. The DEA model can calculate an efficiency index given the distance from the end to the line, and it is up to the researcher to indicate in which direction the line size should be measured. Next, the model can be oriented towards input, i.e., how much less a given DMU_i could consume to maintain the output production level.

In this way, all the DMUs are compared, and each is given an efficiency index with the symbol θ. This efficiency index ranges from 0 to 1, with 1 being the maximum, thus achieving technical

 doi:10.20944/preprints202406.0411.v1

efficiency. By multiplying θ by the inputs of thisDMU_i, we would have the ideal amount of inputs for it to achieve technical efficiency. To calculate θ, it is needed a set of n observations, considering $S_n = \{x_i,\ y_i\ to\ i = \{1, ... n\}\}$. With this set, it is possible to estimate the efficiency frontier and the estimated efficiency index $\hat{\theta}$ for each DMU after solving a series of mathematical inequalities. The model determines the technological level T by the estimator $\hat{T}_{CRS}$ and with this estimator (X_i, Y_i) for each DMU_i, we can calculate the efficiency index by solving the mathematical problem.

$$\hat{\theta}_i\,(x_i, y_i)_{OI-CRS} = max\{\hat{\theta}_i\,|\,\hat{\theta}_i(y_i\ \leq Y\lambda,\ x_i \geq X\lambda), \lambda \in \mathbb{R}^n_+\} \tag{1}$$

Where $X = [X_i, ..., X_n]$ and $Y = [Y_i, ..., Y_n]$ are the matrices representing the sets of inputs and outputs of the n DMU's observed; X_i and Y_i represent the vectors of inputs consumed and outputs produced by a given DMU_i respectively; and $\lambda = [\lambda_1, ..., \lambda_n]$ are the combinations of inputs and outputs that make it possible to achieve the most excellent efficiency.

On the other hand, the model can also be output-oriented. In this orientation, the focus is on understanding how much more a given DMU_i could produce. Therefore, the technical efficiency index, in this case, ranges from 1, for greater efficiency, to infinity and is represented by the Greek letter ϕ. However, if we calculate $1/\phi$, we also get an index from 0 to 1. The same DMU_i should receive similar efficiency indices, regardless of whether the orientation is towards inputs or outputs. As with input-oriented DEA, product-oriented DEA also allows you to calculate how much the DMU_i should produce to be technically efficient by multiplying ϕ by the DMU's outputs. Using the same principles for input-oriented DEA, we would have the following formula for output-oriented DEA:

$$\hat{\phi}_i\,(x_i, y_i)_{OO-CRS} = min\{\hat{\phi}_i\,|\,\hat{\phi}_i(y_i\ \leq Y\lambda,\ x_i \geq X\lambda), \lambda \in \mathbb{R}^n_+\} \tag{2}$$

In addition to orientation, the DEA model can vary in analyzing DMU efficiencies. [27] developed the first model, which did not consider the variation in efficiency that scale can provide in the study they carried out. They then created the DEA model called CRS, for constant return to scale.

However, there are cases where the return to scale is not constant. For example, producing the maximum batch the machinery can have in the industry is more efficient than half that batch. This is because, simultaneously, people and machines can be used for both scenarios, while the former produces twice as much as the latter. In this case, it is unfeasible for the industry in the second scenario to improve its efficiency without increasing its production scale since it cannot use half a machine to reduce its inputs.

To solve this problem, [26] created the model named after their initials, the VRS model, or VRS, variable return to scale, calculates the maximum efficiency of each DMU_i taking into account the DMU's level of production scale. Hence, in the previous example, the company could be technically efficient in both scenarios since the first and second could operate on the efficiency frontier of their scales. The concept of allocative efficiency was created for DMUs that are technically efficient and on the optimal scale. Allocatively efficient DMUs are DMUs that manage to produce more efficiently than all the others, operating on the ideal scale for the production process. They also tend to have the highest profit among all their peers [18,25,28–30].

Thus, eco-efficiency is revealed when a DMU obtains the highest possible level of desired outputs with a given level of inputs and environmental impact or requires the lowest possible amount of inputs and environmental costs to produce a given number of outputs. Its measurement results are calculated by calculating the Euclidean distance that separates each DMU from the border formed by the benchmarks. Thus, it is possible to define two measures of technical efficiency: i) Farrell's technical efficiency-oriented to maximize outputs with a given input level, (θ_0); ii) Farrell's technical efficiency aimed at minimizing the inputs with a given level of products, (θ_I). [25] efficiencies are inverse to [24] radial efficiency.

The evolution of the model is described in Figure 2.

 doi:10.20944/preprints202406.0411.v1

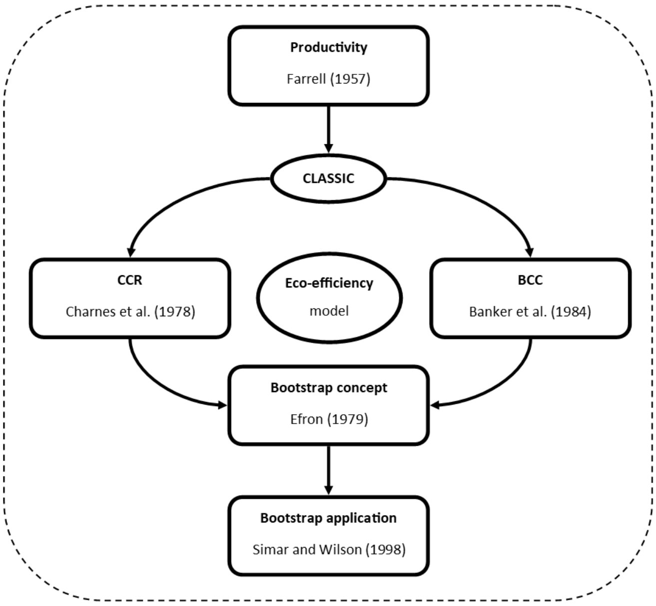

Figure 2. Framework for the evolution of the methodological theorist

2.1. Super-Efificiency and Malmquist Index

The DEA model calculates its indices based on a comparative analysis between peers. However, it has some limitations when there is a considerable range of productive units to be analyzed, making it necessary to use additional models to support the CRS, configuring the case of the super-efficiency model [18,28,29]. Thus, *outliers* can generate super-efficient DMUs. This means that these DMUs are far removed from the other DMUs and cause a distortion in the efficiency frontier. This method calculates each DMU's impact on the model to determine whether it is an atypical case, and the greater the effect, the more atypical the unit of measurement - DMU.

To make this calculation, the model removes a particular DMU_i and recalculates the efficiency of the other DMUs. Initially, the volume of the initial data set, created by the distribution of the efficiencies generated by the model, is calculated. After removing DMU_i, the book is calculated and compared with the initial volume. Therefore, the closer to 0, the more significant the impact that DMU_i will have on the total volume in the data set, which means that it was an outlier [31]. This process is repeated for all DMUs and must be done again after removing each outlier since one outlier can mask the existence of another.

The researcher must decide how many DMUs should be drawn from the database they are working with. Despite that, studies such as [32] suggest that approximately 10%.

To clarify this question, the Malmquist Index was developed, named after Sten [33]. To calculate the efficiency gain in period T compared to period $T - X$, the technical efficiency frontiers must first be calculated using DEA for the DMUs in both periods. When comparing DMU efficiencies over time, two variables can be measured: The slope of the efficiency frontier, which would mean that there is a new technology enabling DMUs to be more efficient, this phenomenon is known as the Frontier Shift Effect; The other change that can be measured is known as the Catch-up Effect, or pairing, which is when a given DMU_i has decreased its distance to the efficient frontier. The Malmquist index is the multiplication of these two variables.

The two scenarios are in Figure 3. The Frontier Shift can be seen in the efficiency frontier created by DMUs A, B, and C, representing the efficiency frontier in period $T + 1$ since it is closer to the Y-axis and further away from the X-axis. The decrease can be seen in the pairing effect in the distance from DMU_M to the efficiency frontier for the period it is in. In period T, DMU_M was at a distance of X from the border, while in period $T + 1$, it was at a distance of 0.7_X.

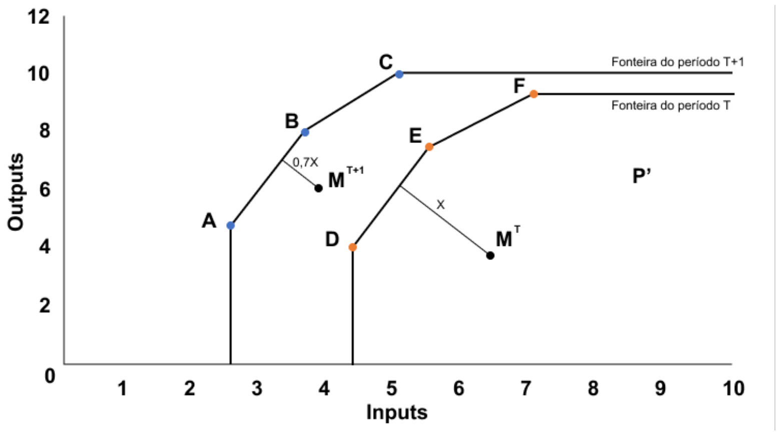

Figure 3. Frontier Shift - Malmquist Index

The matching effect, or Catch-up, is the result of continuous improvements in production processes, using the same technology. Therefore, the comparison of technical efficiency between two periods can be defined as:

$$\frac{\theta_k^t\,(x_k^t,\,y_k^t)}{\theta_k^{t-1}\,(x_k^{t-1},\,y_k^{t-1})} \tag{3}$$

Where:
$\theta_k^t\,(x_k^t,\,y_k^t)$ = DMU_K technical efficiency in a given t period;
$\theta_k^{t-1}\,(x_k^{t-1},\,y_k^{t-1})$ = DMU_K technical efficiency in a given period t+1
While the frontier shift by effect, or Frontier Shift, can be calculated through the formula:

$$\sqrt{\frac{\theta_k^{t-1}(x_k^{t-1},\,y_k^{t-1})}{\theta_k^{t}(x_k^{t-1},\,y_k^{t-1})}\,\frac{\theta_k^{t-1}(x_k^{t},\,y_k^{t})}{\theta_k^{t}(x_k^{t},\,y_k^{t})}} \tag{4}$$

Therefore, multiplying the equation 3 by the equation 4 and passing the second term of the expression 4 into the square root, the word is transformed into the product of the pairing by the frontier shift. Consequently, a result less than 1 means an improvement in the DMU's technical efficiency index, and an effect greater than 1 implies a worsening in the DMU's technical efficiency index. Therefore, the same can be said for the Malmquist Index, calculated with constant returns to scale Mo:

$$Mo = \frac{\theta_k^t\,(x_k^t,\,y_k^t)}{\theta_k^{t-1}\,(x_k^{t-1},\,y_k^{t-1})} \cdot \sqrt{\frac{\theta_k^{t-1}(x_k^{t-1},\,y_k^{t-1})}{\theta_k^{t}(x_k^{t-1},\,y_k^{t-1})}\,\frac{\theta_k^{t-1}(x_k^{t},\,y_k^{t})}{\theta_k^{t}(x_k^{t},\,y_k^{t})}} \tag{5}$$

3. Stochastic DEA Model - Confidence Intervals' Bootstrap

A significant criticism of DEA models is their deterministic nature [18] since they consider past events and do not account for stochastic effects. It would, therefore, be relevant to use statistical models in conjunction with DEA to validate hypotheses, confidence intervals, and correlation analyses [18]. With this improvement, we can have more confidence in the efficiency results and correlation with exogenous variables.

For these two reasons, the bootstrap method is used in this research. The concept of bootstrap arises from the studies of [34], and it is based on them that [35] applies the DEA stochastic model to a data generation process (DMP). Bootstrap is a statistical technique for manipulating sampling to create a probability distribution of a dependent variable from a single sample, which is necessary when the nature of the data in the original model is not exhaustively known [36]. The data generation process is described in four steps:

I. The efficiency θ_i of the original sample is calculated for each DMU_i, where $(i = 1, 2, ..., n)$ considering (x_i, y_i), considering that it is a linear programming problem (LPP) as described in the equation 1 and 2, depending on the type of orientation to be adopted.

II. A new distribution is created, generated by simulated samples $x^*{}_i = [x^*{}_1, x^*{}_2, ..., x^*{}_n]$ based on the original sample $x_i = [x_1, x_2, ..., x_n]$ of the same size, so $x^*{}_i = x_i \frac{\theta_i}{\theta^*{}_i}$. Then, the efficiency scores are calculated for each DMU_i generated, and each value is stored, causing a set of estimates $\hat{\theta}^*_{b,i}$.

$$\hat{\theta}^*_{b,i} = min\{\theta_i \mid \theta_i(y_i \ \leq Y\lambda,\ x_i \geq X\lambda), \lambda \in \mathbb{R}^n_+\} \tag{6}$$

III. The observations are replaced B times (for the estimate to be significant, usually $B \geq 1000$) with simulated comments that respect the rules of the original sample with $(B = 1, 2, ..., b)$ and the calculations are redone on top of each simulation. In this way, it is possible to understand the values that the dependent variable, or the estimator, θ^*, behaves with the variation in the sample. To estimate the standard error, the D_p error of the replications is used.

$$D_p(\theta^*) = \sqrt{\frac{\sum_{b=1}^{B}(\theta^*_b - \bar{\theta}^*)}{B-1}} \tag{7}$$

IV. With that, the confidence limits are calculated, having their intervals determined by the default mode with $\alpha = 95\%$ for each estimate, with $\{\hat{\theta}^*_{i,b}, b = 1, \ldots, B\}$, making the results even more robust and reliable [35].

3.1. Return to Scale Test

The choice of which model to use to calculate efficiency is not an arbitrary one because pre-adopting a model without investigating the nature of the behavior of the Constant Return on Scale or Variable Return on Scale (CRS or VRS) of the (T) technological frontier can cause the result to be biased. This is because, in the case of a model with constant returns to scale, it is not taken into account that certain DMUs that are close in terms of efficiency may have particularities, so total efficiency is calculated. It is analyzed that not every DMU considered efficient in the VRS model will be regarded asd efficient in the CRS model [37].

The null hypothesis H_0 considers the model to be constant (CRS), given consistent returns to scale, and the alternative hypothesis H_1 believes the model to be variable (CRS), given variable returns to scale. By design, the efficiency calculated by the constant model is always lower than the efficiency calculated by the variable model, so $\theta_{i,CRS} \leq \theta_{i,VRS}$, and therefore the estimator S is calculated by the ratio of the sum of $\theta_{i,CRS}$ and $\theta_{i,VRS}$ [18].

$$S = \frac{\sum_{i=1}^{n} \theta_{i,CRS}}{\sum_{i=1}^{n} \theta_{i,VRS}} \leq 1 \tag{8}$$

With the estimator $S = 1$, it is considered a H_0, although this value is hardly reached. Therefore, when the estimator S is close to 1, the model to be considered is CRS. In H_a, the estimator S is significantly less than 1, so VRS is the model considered. To state whether it is significantly smaller, the critical value C_α is obtained with a significance level $\alpha = 5\%$, so $S \leq C_\alpha$. Given that the natural distribution of the original sample is not exhaustively known, it is used bootstrap of the boot.sw98 package inside the R environment to calculate the efficiencies $\theta_{i,CRS}$ e $\theta_{i,VRS}$ by the resampling method to obtain the estimated value S [38]. In this way, it is decided between H_0 and H_a for the result of the equation 8.

4. Methodology of Eco-Efficiency

The concept of eco-efficiency presupposes an understanding of the production technology sector studied [39]. Therefore, as technology results from incorporating scientific innovations and individual management experiences into production processes, the status at any given time is unknown. Because

 doi:10.20944/preprints202406.0411.v1

it is unknown, it is common to describe it from the set of inputs, which produce a vector of new goods and services (outputs) in specific periods after being combined and processed.

This representation of technology often disregards the fact that some of the inputs used are returned to nature as waste and pollution, referred to here as undesirable outputs. These undesirable outputs are not subject to total recycling or total absorption by nature, i.e., they cause environmental damage. This is a consequence of the second law of thermodynamics, also called the Law of Entropy, which states that a part of the resources is always dissipated and lost with the transformation of energy or matter. Its complete reversibility to the original permitted state is not possible. Therefore, it does not consider undesirable outputs that pollute the environment, implying a partial representation of technology, which underestimates the social cost of total production and induces global results.

Thus, this section describes and analyzes the data used. The sample comprises a group of 516 municipalities in the Amazon biome - Brazil, with the application focused on eco-efficiency. The analysis is structured in stages, initially defining which variables should be considered as inputs and outputs, then processing and balancing the database using the R language. For balancing, all municipalities with unreliable numbers, missing or null data for inputs, and outputs, and even cities that did not have their data available were removed. The period used was the most recent data released by the Brazilian Institute of Geography and Statistics (IBGE) in 2017. Next, given the super-efficient DMUs, the [31] method was applied.

The model is based on the assumption that a super-efficient DMU, considered to be an outlier, masks an efficient DMU, so looping is applied to recalculate the efficiencies with an *input* orientation and remove the super-efficient DMUs until there are no more super-efficient DMUs. To do this, there is a pause criterion in the process of eliminating *outliers* given by $\frac{V_f}{V}$, with a value of less than 0.7, where a DMU could produce 70% of what it has and still be considered a technically efficient DMU. The *sdea()* function predefined in the "Benchmarking" package was used for the calculation. Given the deaR model, the efficiencies of DMUs with an output and input orientation were calculated using the processes in the RVE box.

Thus, the "make_deadata" function of the "deaR" package will allow the creation of an object adapted for calculating data envelopment analysis using the desired inputs and undesired products, where the undesired effect is minimized, and the selected input is maximized. In this study's model, the preserved hectares variable was considered a desirable input, as it is a characteristic that precedes agricultural production in the municipalities. Thus, CO_2 emissions were considered an undesirable input because they are a consequence of the agricultural industry. The "model_basic" function of the same package is applied to the object created, and the results of the efficiency index, reference DMUs, and the ideal quantity of inputs or products to achieve technical efficiency for each DMU are found. The bootstrap defines the type of return to scale of the model, and the analysis results are the RCE model. In this way, it is possible to analyze the results by calculating how much it would be possible to optimize products and inputs.

In the second stage of this research, the Malmquist index was used. The looping removal model is applied again each year. Each year would have a few DMUs considered super-efficient, thus creating a list of all the super-efficient DMUs to be removed from all the years, i.e., from the entire database. We used 0.7 for the efficiency index since 150 DMUs would be removed from the sample with this parameter, which is no more than 30% of the [32] base. To compare the time windows, processing the data is needed to give continuity to the efficiency gain, where the frontier shift index for the year T is multiplied by the index for the year $T-1$, and then the index for the year $T+1$ by the product of the two previous indices, in an accumulative manner. Using the efficiency index of the years, a cross base is created with the annual meteorological information, and the multivariate linear regression of the floor is analyzed, with the meteorological data being the independent variable and the efficiency of the year the dependent variable.

Finally, a correlation study is carried out between the efficiency of municipalities and meteorological characteristics by calculating the frontier shift and efficiency over the years. Using the Malmquist

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 10 June 2024 doi:10.20944/preprints202406.0411.v1

function, the frontier shift, pairing effect, and Malmquist index are calculated for each municipality. To calculate the frontier effect for the year, a geometric mean of all the frontier shift indices for all the DMUs in the corresponding year is used. The first year is discarded since the calculation is made by comparing efficiency in years T and $T-1$. To solve the problem, the data for 2006 is repeated, simulating 2005, where 2006 has a frontier displacement index equal to 1.

4.1. Variables

To assess the agricultural eco-efficiency of the municipalities that make up the Amazon biome, a set of variables available in the 2006 and 2016 Agricultural Census was adopted, and a time series was generated using linear regression to fill in all the missing years. The sector's classic inputs and outputs were considered, plus one positive and one negative externality.

To calculate the municipal's operational efficiencies of agricultural production, the variables used as inputs were People engaged in agricultural work, Hectares Dedicated to rural production, and Hectares Preserved, the latter being a desirable input. The outputs selected were the value of agricultural production in thousands of reais and CO_2 emissions as an undesirable output. All these variables were at the municipality level. As a general rule, the classic inputs and outputs used in the modeling were:

X_1 - People engaged in agricultural industry,
X_2 - Hectares Dedicated to agricultural production,
X_3 - Preserved Hectares,
Y_1 - Value of agricultural production in thousands of reais,
Y_2 - CO_2 emission with an undesirable *output*,
Z_1 - Average temperature,
Z_2 - Precipitation.

The area variables (hectares dedicated to production and hectares preserved), as well as the production value and people employed, come from the official Brazilian Institute of Geography and Statistics (IBGE)[2] databases, according to the Censuses carried out in 2006 and 2017. For the years in which there is no official Census information by the municipality, a simulation was made using the estimated growth in annual production in the state, according to CONAB[3] (National Supply Company) year-on-year, combined with an estimate of how much each municipality should have a share (in %) within each state, made using an arithmetic progression calculation between the municipality's claim in 2006 and the municipality's claim in 2017. All the consolidated data and results found by this research have been made available and can be consulted on figshare[4].

Figures 4 and 5 show the area dedicated to agricultural activities in each municipality as a percentage of total municipal area in 2006 and 2017, respectively. The municipalities with a more proportional area dedicated to agriculture are in the East and South borders of the Amazon Forest. From 2006 to 2017, the increase in agricultural area occurred mostly in the municipalities already with strong agricultural activities.

---

2 https://www.ibge.gov.br/estatisticas/economicas/agricultura-e-pecuaria/9827-censo-agropecuario.html
3 https://www.conab.gov.br/
4 https://doi.org/10.6084/m9.figshare.25958827

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 10 June 2024 doi:10.20944/preprints202406.0411.v1

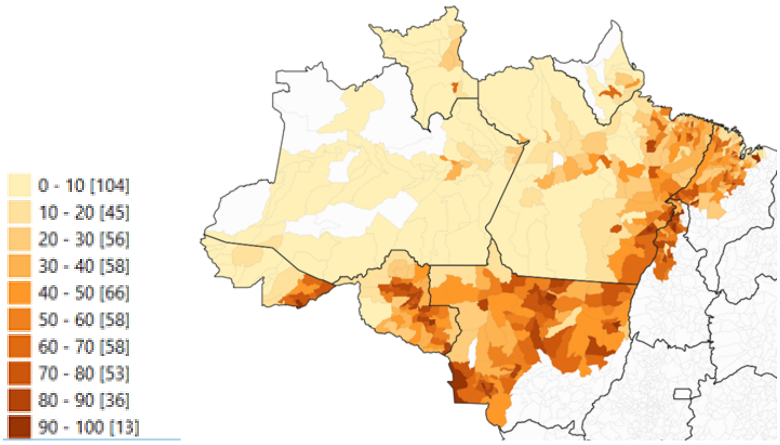

Figure 4. Dedicated area to agricultural production in 2006 - % of municipality area - [number of municipalities in each group].

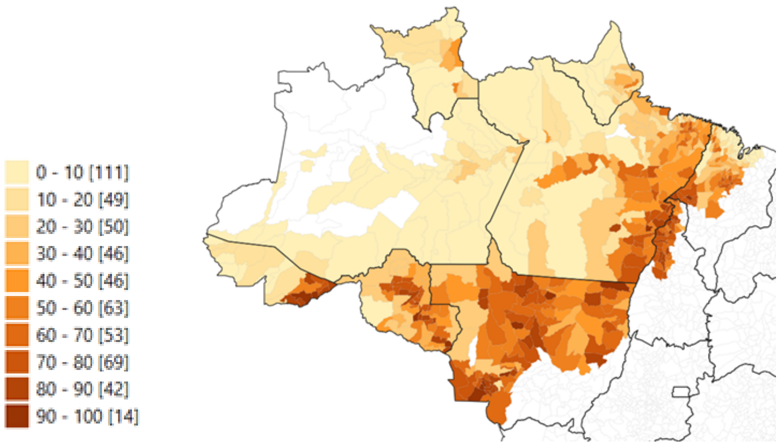

Figure 5. Dedicated area to agricultural production in 2017 - % of municipality area - [number of municipalities in each group].

Figures 6 and 7 present the protected area as a percentage of the total municipal area in 2006 and 2017, respectively. Two characteristics of these maps need attention: i) the majority of the municipalities have a very low percentage of protected area (see the number of municipalities in each group - numbers in brackets in the map caption); ii) the municipalities with the more protected area are the same with the high agricultural area. These contradictory results come from the definition of protected area variable in the Censuses, as it represents the protected area in the farms, not including Indigenous reservation areas, (environment) Conservation Units, and other public areas (with no private owners). Hence, the municipalities with known intact forests (in the center-west of the Amazon Forest area) are shown with a low percentage of protected area because they have very low areas of private farms. For this study, however, the definition of the protected area variable considering only the area in the farms is adequate since it derives from farmers' decisions on how much area is allocated for agriculture and environmental protection.

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 10 June 2024 doi:10.20944/preprints202406.0411.v1

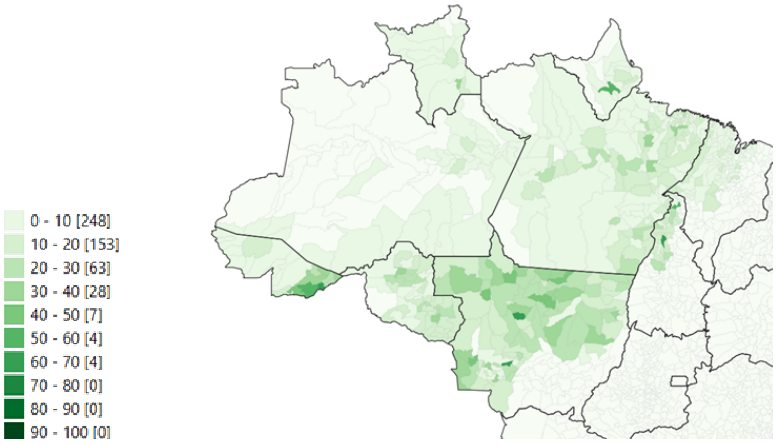

Figure 6. Protected Area in farms in 2006 - % of municipality area - [number of municipalities in each group].

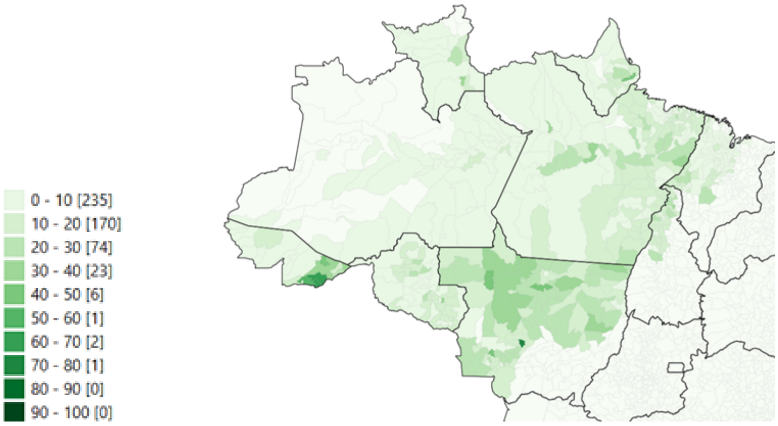

Figure 7. Protected Area in farms in 2017 - % of municipality area - [number of municipalities in each group].

The data regarding CO_2 emission was loaded from the official site of Greenhouse Gas Emission and Removal Estimation Systems (SEEG)[5]. This variable is measured in tons of carbon dioxide emitted by the agribusiness and property transformation but with an annual uninterrupted update since 2000. After that, an official database was collected from INMETRO using all the automatic measuring agencies in each state for every hour of every day to analyze the relation between the production efficiency index and the weather recorded. Then, the average temperature for each year in each state was calculated, and these values were used for each municipality that composed the Brazilian biome.

5. Results

Using the final base of 516 DMUs, after removing the outliers and super efficiencies, the bootstrap statistical testing method was used to define whether the model should be CRS or VRS. For this test, the null hypothesis is that the model should be CRS, constant returns to scale, and it should only be rejected if the value of the calculated S statistic is less than the estimated critical value, which is also defined within the bootstrap tests. In the test, 10,000 simulations are carried out, resulting in a critical value of $C_\alpha = 0.3854$. Since the estimated value for comparing the DEA models was 0.9605, it was higher than the critical value C_α. With these results, there was no statistical evidence to reject the null hypothesis, thus accepting that the best model for this problem is the CRS, constant return to scale.

5 https://plataforma.seeg.eco.br/total_emission

 doi:10.20944/preprints202406.0411.v1

The conclusion, therefore, must be that scale has no direct influence on the efficiency of agricultural production in the municipalities of the biome included in the study.

Conceptually, it can be concluded that these results make sense with what is perceived today. When analyzing the productivity of an agricultural unit, the main determining variables are soil, climate, and the technologies used. Whether the production unit is larger or smaller is no different regarding soil and climate. What determines these conditions is geography and the location of these units. As for technology, a team with greater production power could have more resources to invest in machinery. However, as the object of study is entire municipalities, this effect is diluted among the various production units within each municipality and is therefore regressed to an average.

Firstly, it is essential to analyze the current scenario of the agricultural industry in the Amazon to measure potential gains and even map out action plans to capture these gains. The results shown in Table 1 were achieved using DEA for inputs with constant returns to scale. Thus, it is possible to see that, on average, DMUs should produce 9.9% more than they did in 2017 to become efficient. In addition, it is also possible to see that the worst DMU should have approximately twice as much, or 197.5% of what it produces.

These results show how current agricultural industry practice, on average, has become more sustainable since most DMUs are close to technical efficiency, given that, in the first quartile, DMUs would only need to produce 1.5% more to achieve efficiency, and in the third quartile, this figure increases to 12.3%. Few municipalities would need to improve their production significantly to achieve technical efficiency, so only 25% of municipalities would need to increase production by more than 12.3%.

Table 1. Summary of data regarding the efficiency of municipalities

Min.	1st. Q	Median	Mean	3rd Q	Max.
1.000	1.015	1.051	1.099	1.123	1.975

Figure 8 shows the results of CRS efficiency by municipality in 2017. Some geographical concentrations of the most efficient DMUs can be seen: i) one group in the South border in the state of Mato Grosso, where highly technical soybean and corn plantations are the main agricultural activities; ii) one group in the state of Amapá (North of the map), where farming activities are only begging; and iii) one group in the West os the state of Amazonas (west of the map), where farming activities are scarce. On the other hand, less efficient DMUs are in the Southeast of Pará state, where farmers focus on raising beef cattle; and in Rondônia state (Southwest of the map), where agricultural activities are spread on several crops. It already gives some insights on how to boost eco-efficiency for agriculture in the region, which will be discussed later.

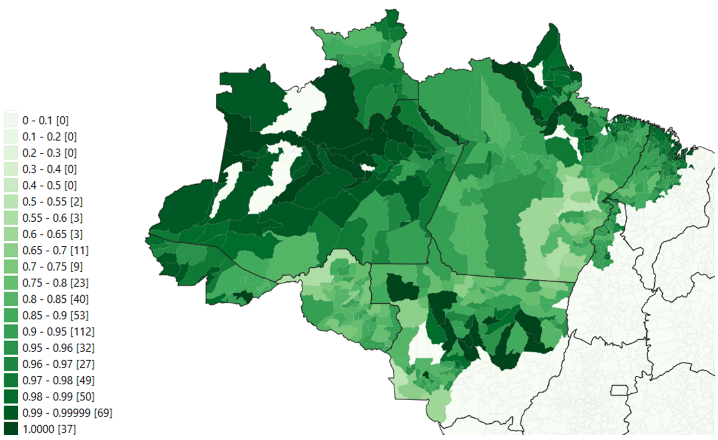

Figure 8. CRS Efficiency by the municipality in 2017 - [number of municipalities in each group].

 doi:10.20944/preprints202406.0411.v1

Table 2 shows that the variables relating to inputs and outputs between the municipalities have heterogeneous values, which coincides with the type of return to scale identified in the CRS model, where small and large producers can both be considered eco-efficient, even though they have different levels of resources and results. This is the case of Juara, which has proportionally higher values for inputs and outputs compared to the municipality of Rosário, but both are considered eco-efficient.

Table 2. Variables of efficient municipalities, inputs and outputs

ID	Municipality	Input			Output	
		Input		Desirable input	Output	Undesirable output
		Production Hectares 10^3	Employed People (10^3)	Preserved Hectares 10^2	Revenue R$ 10^3	CO_2 emission ton 10^3
2109601	Rosário	1	16.46	3	3	7
1300060	Amaturá	6	25.15	50	9	3
1600154	Pedra Branca do Amapari	33	20.43	289	16	6
1600279	Laranjal do Jari	30	11.99	259	14	13
5105101	Juara	1,522	54.69	8,461	226	2,016
2101350	Bacurituba	1	7.69	3	0	22
5107248	Santa Carmem	234	9.30	1,136	464	190
1301308	Codajás	70	38.78	610	22	9
1302108	Japurá	3	12.56	20	3	4
1302801	Maraã	4	26.26	15	17	5
1303205	Novo Airão	8	17.40	65	5	3
1303700	Santo Antônio do Içá	3	33.29	24	6	4
1304062	Tabatinga	3	87.79	12	13	2
1304203	Tefé	21	117.49	155	65	8
1500305	Afuá	134	121.28	1,201	92	37
1503002	Faro	6	3.33	43	1	30
1503101	Gurupá	57	52.83	523	27	18
1507961	Terra Alta	2	3.45	9	2	6
2106755	Miranda do Norte	20	3.04	18	1	18
5106224	Nova Mutum	766	47.29	2.630	1,957	233
5106307	Paranatinga	1,300	46.75	5.268	827	263
5107156	Reserva do Cabaçal	105	6.08	657	0	34
5107925	Sorriso	828	49.37	2,011	2,812	432
5108907	Nova Maringá	612	14.74	3,966	515	297
1200708	Xapuri	494	55.20	3,758	19	435
1300409	Barcelos	8	27.47	52	23	4
1300839	Caapiranga	4	6.20	26	5	6
1303536	Presidente Figueiredo	197	76.92	1,642	64	33
1500701	Anajás	163	29.00	1,323	20	19
1504000	Limoeiro do Ajuru	37	141.21	317	51	13
2104909	Guimarães	1	24.50	2	2	7
2111201	São José de Ribamar	1	27.06	2	1	4
5101407	Aripuanã	1,112	55.18	7,085	33	1,012
5102702	Canarana	790	26.23	4,107	756	246
5107065	Querência	834	57.70	3,813	1,546	551
1300102	Anori	22	29.31	186	15	3
5105259	Lucas do Rio Verde	316	26.05	828	1,111	55

In Table 3, you can see all the municipalities considered efficient, i.e., they have an eco-efficiency index equal to 1, resulting in 37 eco-efficient municipalities. This represents 7.17% of the initial 516 municipalities, a low figure at first. Still, considering that most of the municipalities are close to the production frontier, it can be analyzed that the DMUs considered efficient do have an optimum level of total efficiency.

 doi:10.20944/preprints202406.0411.v1

Table 3. Efficient municipalities

ID IBGE	Municipality	DEA Index CRS Output	ID IBGE	Municipality	DEA Index CRS Output	ID IBGE	Municipality	DEA Index CRS Output
2109601	Rosário	1	1304203	Tefé	1	1300839	Caapiranga	1
1300060	Amaturá	1	1500305	Afuá	1	1303536	Presidente Figueiredo	1
1600154	Pedra Branca do Amapari	1	1503002	Faro	1	1500701	Anajás	1
1600279	Laranjal do Jari	1	1503101	Gurupá	1	1504000	Limoeiro do Ajuru	1
5105101	Juara	1	1507961	Terra Alta	1	2104909	Guimarães	1
2101350	Bacurituba	1	2106755	Miranda do Norte	1	2111201	São José de Ribamar	1
5107248	Santa Carmem	1	5106224	Nova Mutum	1	5101407	Aripuanã	1
1301308	Codajás	1	5106307	Paranatinga	1	5102702	Canarana	1
1302108	Japurá	1	5107156	Reserva do Cabaçal	1	5107065	Querência	1
1302801	Maraã	1	5107925	Sorriso	1	1300102	Anori	1
1303205	Novo Airão	1	5108907	Nova Maringá	1	5105259	Lucas do Rio Verde	1
1303700	Santo Antônio do Içá	1	1200708	Xapuri	1			
1304062	Tabatinga	1	1300409	Barcelos	1			

For the DMUs that achieved a score of 1 in the efficiency index to be considered efficient, the existing outliers in the database were treated, given that the DEA's estimated efficiency frontier is sensitive to extreme values. The data cloud technique was used because it is robust for identifying outliers. In summary, the volume of the combined matrix of inputs and outputs is observed, where a significant reduction in this volume, following the removal of a DMU, would indicate that this unit is an outlier. Another technique used was to create cut-off ranges by estimating DEA with super-efficiency. Figure 9 shows the municipalities considered outliers and super-efficient in 2017, that were excluded from the Malmquist index analysis below. As can be seen, most of these DMUs are in not developed areas in Amapá and Amazonas states, and some very urbanized municipalities (but small in total area) in Pará, Maranhão, and Tocantins states (East of the map).

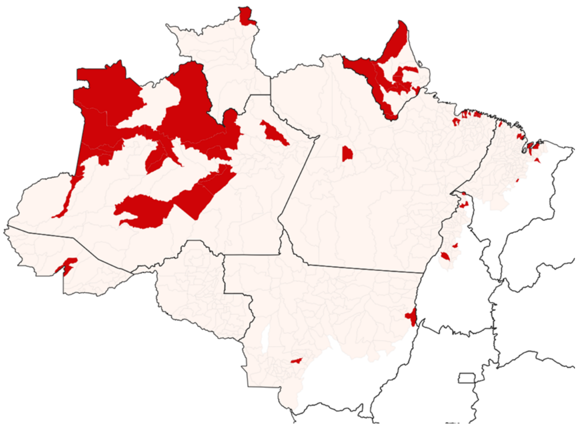

Figure 9. Municipalities considered outliers in 2017.

The Malmquist index was calculated for series from 2006 to 2017[6]. Figure 10 shows the aggregated index (2006-2017). The municipalities that have improved in the period are in the state of Mato Grosso (South of the map) where highly technical soybeans and corn plantations became the main activity. Municipalities where the eco-efficiency has worsened are in the Southeast of Pará state, a region where beef cattle raising increased considerably, and in Rondônia state (west of Mato Grosso state), where agriculture in several crops has increased. Although a great increase in the value of agricultural production was observed in these DMUs, it was accompanied by a large increase in emissions. Another important result from the Malmquist analysis is that DMUs that are very eco-efficient but have very small agricultural production had minimal progress in the period.

6 link:https://doi.org/10.6084/m9.figshare.25958827

 doi:10.20944/preprints202406.0411.v1

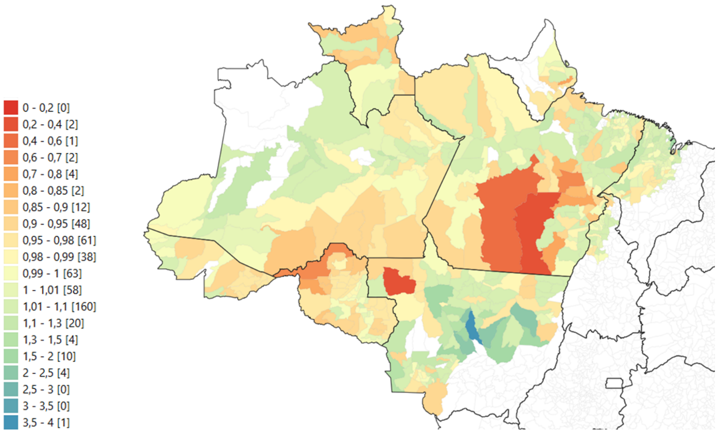

Figure 10. Malmquist Index from 2006 to 2017 - by municipality number of municipalities in each group].

In absolute terms, as the chosen orientation is towards output, the possible impact on each variable (Y_1 and Y_2) is analyzed if all the municipalities converge towards efficiency. The values are described in the Table 4, where the CO_2 emission variable represents the reduction, or saving, generating ecological gains, while the production value represents the increase needed to turn inefficient DMUs into efficient ones. Real is Brazilian currency, US$1.00 = R$4.90 in August 2023.

Table 4. Production potential gains

Variable	Value
Revenue	3,195,002,000 (R$)
CO_2 emission	25,849,560,000 (ton)

Due to the way data envelopment analysis is calculated, the efficiency index is applied equally to all variables, so if the index is 1.5, the municipality should produce 50% more of all products. Therefore, the ranking of the cities in relative values will remain static regardless of the variable being analyzed. However, this ranking can change in nominal values. With this in mind, you can see the minor efficient municipalities in the Table 5.

Table 5. Worst Municipalities Analyzed by Simulation

IBGE index	Municipality	DEA index CRS Output	IBGE index	Municipality	DEA index CRS Output
1504208	Marabá	1.975094	5102504	Cáceres	1.626327
5105507	Vila Bela da Santíssima Trindade	1.883153	1507300	São Félix do Xingu	1.601048
1505064	Novo Repartimento	1.781707	1506583	Santa Maria das Barreiras	1.573696
1100205	Porto Velho	1.703062	1500347	Água Azul do Norte	1.537795
1502764	Cumaru do Norte	1.694544			

Although it doesn't have the worst efficiency index, the municipality with the most significant potential for gains in production value is São Félix do Araguai, with a total of 182 million raises, and the municipality with the most potential for savings in CO_2 emissions is São Félix do Xingu, which could save approximately 1.7 billion tons, as can be seen in the Table 6.

Table 6. Municipality with the Greatest Potential in Production Value and Economy in CO2 Emissions

Variables	Municipalities	Potential
Revenue	São Félix do Araguai	183,353,000.00 (R$)
Co_2 emission	São Félix do Xingu	-1,683,912,000.00 (ton)

It is also possible to measure the impact of efficiency gains on input savings. However, it's important to remember that you shouldn't look at outputs and inputs simultaneously since the profits

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 10 June 2024 doi:10.20944/preprints202406.0411.v1

from these two perspectives are mutually exclusive. The results for savings in production hectares and people employed, as well as the potential for increasing the area preserved, can be found in Table 7.

Table 7. Economic potential

Variable	Value
Employed people	242,438 (un)
Production area	13,916,750 (ha)
Preserved area	6,248,745 (ha)

While the municipality with the most significant potential for gain, São Félix do Xingu, can be seen in Table 8, with the values needed to improve in each variable to achieve efficiency.

Table 8. São Félix do Xingu' economic opportunity

Variable	Quantity
Employed people	8,455 (un)
Production area	924,292 (ha)
Preserved area	578,568 (ha)

A linear regression study was carried out for the statistical modeling part. As a dependent variable, the efficiency index of the municipalities was chosen. The independent variables, or predictors, were the meteorological characteristics, volume of precipitation, average temperature, average humidity, and standard deviation of temperature over the months. The study sought to simulate the different seasons and capture the impact of this natural phenomenon and the various crops that can be planted depending on the season. It is also interesting to understand production and efficiency by state.

In Table 9, you can see the values achieved for each of the variables analyzed in 2017, as well as the state's total efficiency index, calculated by the geometric mean of the efficiencies of the municipalities in each state.

Still analyzing Table 9, it can be seen that the state with the highest efficiency is Amapá, with an index of 1.017, followed by Amazonas, which obtained a result of 1.02. The state with the worst recorded efficiency was Rondônia. Despite being the state with the third highest financial value of production, the amount of CO_2 emitted and the number of hectares used for production was much higher than ideal. It is also interesting to understand production and efficiency by state. In the Table 9, you can see the values achieved for each of the variables analyzed in 2017, as well as the state's total efficiency index, calculated by the geometric mean of the efficiencies of the municipalities in each state.

Table 9. States summary

State	Efficiency index	Input			Output	
		Input		Desirable input	Output	Undesirable output
		Production hectares 10^3	Employed People (10^3)	Preserved hectares 10^2	Revenue R\$ 10^3	CO_2 emission ton 10^3
AC	1.069	4,233	126,514	2,544	442	5,870
AM	1.023	3,733	307,201	2,150	1,294	3,102
AP	1.017	1,501	30,732	871	226	822
MA	1.041	4,652	293,574	862	865	7,794
MT	1.140	38,164	283,069	15,349	26,914	45,209
PA	1.111	27,637	951,857	10,338	6,036	42,360
RO	1.203	9,220	270,812	2,319	1,582	28,552
RR	1.067	2,636	67,070	1,181	395	1,792
TO	1.059	2,867	51,596	669	211	4,170

The ideal consumption quantities per state can be consulted in Table 10. If the states reach these values, production remains constant, with the values in the "Revenue" and "CO_2 emission" columns shown in 9 table.

Therefore, Table 10 represents the ideal level of consumption while maintaining the quantity produced. By comparing these tables, it is possible to calculate the number of inputs that could be saved or boosted in the case of the desirable input preserved area. The state of Rondônia, for example,

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 10 June 2024 doi:10.20944/preprints202406.0411.v1

could use approximately two million hectares and fifty-four thousand fewer people to maintain its production level and still increase the preserved area by about seventy thousand hectares.

On the other hand, the states could increase production. In this scenario, the states' consumption would maintain the same values in the "Production Hectares", "Employed people" and "Preserved Hectares" columns of 10 table.

Table 10. Ideal consumption by state

State	Production Hectares 10^3	Employed people (10^3)	Preserved Hectares 10^2
AC	3,903	117,526	2,738
AM	3,521	297,925	2,263
AP	1,467	29,940	888
MA	4,232	280,095	964
MT	32,929	241,189	17,856
PA	22,309	845,996	12,830
RO	7,313	216,768	3,007
RR	2,402	62,525	1,261
TO	2,651	48,023	725

The ideal production values, maintaining current consumption, can be found in Table 11. Rondônia, which was the state with the worst eco-efficiency index in the Amazon Biome, would need to increase the value of production by R$ 341 thousand reais and reduce approximately 6 thousand tons of CO_2 to reach the production frontier given by the DMUs in the North region.

Table 11. Ideal production by state

State	Revenue R$ 10^3	CO_2 emission ton 10^3
AC	470.2961	5292.594
AM	1334.37	2834.646
AP	231.9256	795.1486
MA	950.0968	7111.449
MT	28795	37305.74
PA	6792.038	32472.07
RO	1923.489	22531.48
RR	435.6717	1644.752
TO	226.3446	3834.731

In Table 12 , you can analyze the number of inputs that could be saved. In the fourth quartile, since it is the quartile with the worst DMUs, there would have to be a reduction of eleven million hectares and one hundred and seventy-nine thousand people, as well as an increase of five hundred and eighty-four thousand hectares preserved to reach the level of efficiency, provided that the level of production remained constant.

Table 12. Consumption reduction potential per Quartil

Quartil	Production Hectares 10^3	Employed People (10^3)	Preserved Hectares 10^2
1	-26	-2,374	14
2	-428	-15,480	200
3	-1,584	-45,162	651
4	-11,879	-179,423	5,384

As for the Table 13, how much more could the municipalities produce to reach the maximum possible production for each quartile? Again, quartile four would have to increase the value of production by approximately two million reais and reduce CO_2 by about twenty-three thousand tons, assuming that inputs remain constant, i.e., do not change.

Table 13. Ideal production per Quartil

Quartil	Revenue R$ 10^3	CO_2 emission ton 10^3
1	20.00	-19.06
2	251.65	-388.72
3	627.22	-2,305.30
4	2,296.13	-23,136.49

By summarizing the efficiency quartiles, it is also possible to measure the results divided into efficiency groups. This summary can be found in Table 14. Between quartile one and quartiles 2 and 3, there is no significant difference between the eco-efficiency indices determined, 1.005, 1.031, and 1.085, respectively, at which point the hypothesis that, on average there is no considerable distance between the eco-efficient and inefficient DMUs is once again confirmed, This can also be seen in the difference between the first quartile and the fourth quartile, where the eco-efficiency index values are 1.005 and 1.265 respectively, where there is no considerable discrepancy between these indices.

Table 14. Quartile summary

		Input				Output
		Input		Desirable input	Output	Undesirable output
Quartil	Efficiency index	Production Hectares 10^3	Employed People (10^3)	Preserved Hectares 10^2	Revenue R$ 10^3	CO_2 emission ton 10^3
1	1.005	12,763	463,090	6,704	13,174	8,163
2	1.031	13,374	539,044	6,133	8,420	11,606
3	1.085	19,919	578,013	7,546	7,686	27,829
4	1.265	48,586	802,278	15,900	8,684	92,074

Various combinations were made between the independent variables to find the best predictive model for the average annual efficiency index. The most assertive model with static significance, defined by the P-Value test, used Average dry bulb temperature, average dew point temperature, average relative humidity, annual rainfall volume, and standard deviation of temperature over the months.

The model calculated using these variables to predict the efficiency index achieved a P-value of 0.0007 and an R2 of 0.9721. The figures presented by the model are encouraging. However, a closer look at the statistics reveals that this relationship may not be one of cause and effect. Efficiency has increased steadily over the years without showing much variation or volatility in this growth, which tends to be explained by advancing technology for growing crops linked to the agricultural industry. At the same time, it can be said about the climate phenomena analyzed, given that the world is undergoing a gradual process of global warming. As such, the correlation may result from two phenomena showing the same trend.

6. Conclusion

The main objective of this study was to measure how advanced current eco-efficiency is in the municipalities of the Brazilian Amazon biome regarding agricultural activities, the most critical sector of the economy in the region, and on the national stage. It aims to provide a structured view of how much it is possible to maximize the region's economic development without disregarding the environmental impacts these activities can cause from 2006 to 2017, with a view to the 2030 agenda. To this end, we mapped the municipalities that should improve their efficiency. We began an analysis of how these municipalities could improve their eco-efficiency, stimulating improvements towards ecological sustainability, using the data cloud method to eliminate outliers with the potential to bias the results using bootstrap computational techniques, applying the scale return test to analyze the significance and measurement of the scores corrected for random data bias. This allows us to test the differences between the eco-efficiencies of the municipalities in the defined period from 2006 to 2017.

This provides a relevant and structured result for assessing the impacts of the agricultural industry in the municipalities that make up the Brazilian Amazon biome and evaluating how eco-efficient the industrial processes adopted in the world's leading environmental reserves are. The results point to an essential interpretation that most of the municipalities are already operating at a satisfactory level of efficiency given the technological level available; around 7.17% of the 516 municipalities would already be working on an eco-efficient scale. In addition, for the municipalities that have not yet reached the eco-efficient frontier, there is a potential for improving agricultural revenue by approximately R$ 3.195 million and decreasing 25,849,560 thousand tons of CO_2 emissions, considering the same resources, or equivalently reduce 13,915.7 Km^2 of areas used and increase preserved sites by 6,248.7 Km^2, with the same level of production, depending on the type of orientation to be used. Therefore, the research questions were answered, and the results obtained provide pertinent information for decision-making and policy definition, allowing for optimal economic and environmental development, as well as promoting actions to reduce inequality, improve the working conditions and development of the rural population, in the face of climate change undergoing in Amazon biome.

Another conclusion that can be drawn from the results is the fact that the production frontier of the municipalities results in a technological behavior of constant returns to scale (CRS), which is a relevant result for understanding factors that involve inequality between producers in the municipalities because, given the results of the model, both small, medium and large producers can be eco-efficient (total efficiency) given the CRS frontier within the Amazon Biome. Hence, these differences occur regardless of the production level, so there will be inefficient and efficient small, medium, and large producers. This statement aligns with the results presented in Table 2, where inputs and products have differences of up to three decimal places between the values for each efficient municipality.

Another research question to be answered is the influence of meteorological phenomena on efficiency gains or losses over time. This part of the research is essential because it is increasingly possible to see the changes that are happening to the weather around the planet, such as global warming, increased rainfall and droughts in different regions, phenomena such as *El Niño* can affect agricultural production and cause an under-supply of essential inputs for the world economy. Therefore, it is understood that correlations between meteorological characteristics and production efficiency would make it possible to develop alternative technologies to prepare producers better to cope with the phenomena above. However, it was impossible to prove a direct correlation between the variables. Even though the numbers were good and passed the statistical tests, it was impossible to rule out the hypothesis that made it a coincidence. Although this hypothesis cannot be ruled out, there have been studies that have been able to prove this relationship between eco-efficiency scores and climate factors, providing vital information for the environmental decision-making process, such as the work by [18], which concluded that the changes caused by the decrease in rainfall and the increase in temperature would have a positive impact on the mountainous region of the TMCF in Mexico, specifically in an area located in the Sierra Madre Oriental between 2016 and 2017.

Finally, it should be emphasized that given the study's limitations related to data availability to obtain more extended periods to determine relevant inputs and outputs and the unavailability of open access to data for scenarios other than agriculture, such as industry and commerce. The research has fulfilled its guidelines and achieved results pertinent to its data structure, providing social, environmental, and economic development indicators. According to [40], this balance between the pursuit of economic and social development, aligned with the proper use of natural resources, arouses interest in studies focused on the area to promote a set of actions that have social and environmental impacts due to the updates that have occurred in recent years in agribusiness, which leads the country towards the pursuit of sustainable development.

Therefore, for future studies, it would be interesting to evaluate this correlation, and a simulation with a larger volume of data would be necessary, considering a time series of more than a decade, where significant variations in climate change could be verified. We also recommend redoing the DEA modeling with different groups of municipalities, separating them by state or region rather than comparing all the municipalities with each other. In addition, comparing the results with other areas of the country and even other countries makes sense. Adding input or output variables to the survey would also enrich the analysis. This would allow a larger volume of data and situations to be tested.

Author Contributions: Conceptualization, Gabriela Saiki and André Serrano; Data curation, Gabriela Saiki, André Serrano and Carlos Rosano-Peña; Formal analysis, Fabiano Pompermayer; Investigation, Gabriela Saiki, André Serrano and Carlos Rosano-Peña; Methodology, Gabriela Saiki, André Serrano and Carlos Rosano-Peña; Supervision, Pedro Albuquerque and Mario Pensado-Leglise; Validation, Gabriela Saiki, André Serrano, Gabriel Rodrigues, Carlos Rosano-Peña and Fabiano Pompermayer; Writing – original draft, Gabriela Saiki, André Serrano

and Gabriel Rodrigues; Writing – review & editing, Gabriela Saiki, André Serrano, Gabriel Rodrigues and Fabiano Pompermayer.

Acknowledgments: The authors would like to thank the Brazilian National Confederation of Industry (CNI) for partially supporting this project and for their support and collaboration throughout this research project.

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

References

1. Meadows, D.H.; Meadows, D.L.; Randers, J.; Behrens, W.W. *Limites do crescimento: um relatório para o projeto do Clube de Roma sobre o dilema da humanidade*; Perspectiva São Paulo, 1972.
2. Keeble, B.R. The Brundtland Report: 'Our Common Future'. *Medicine and War* **1988**, *4*, 17–25.
3. Heyder, M.; Theuvsen, L. Determinants and effects of corporate social responsibility in German agribusiness: A PLS model. *Agribusiness* **2012**, *28*, 400 – 420. https://doi.org/10.1002/agr.21305.
4. Bockstaller, C.; Girardin, P. How to validate environmental indicators. *Agricultural Systems* **2003**, *76*, 639–653. https://doi.org/https://doi.org/10.1016/S0308-521X(02)00053-7.
5. Morse, S.; McNamara, N.; Acholo, M.; Okwoli, B. Sustainability indicators: The problem of integration. *Sustainable Development* **2001**, *9*, 1 – 15. https://doi.org/10.1002/sd.148.
6. Ângelo, H.; De Almeida, A.N.; Serrano, A.L.M. Determinants of Brazil's demand of pulpwood. *Scientia Forestalis/Forest Sciences* **2009**, *1*, 491–498.
7. Bauler, T. An analytical framework to discuss the usability of (environmental) indicators for policy. *Ecological Indicators* **2012**, *17*, 38–45. Indicators of environmental sustainability: From concept to applications, https://doi.org/https://doi.org/10.1016/j.ecolind.2011.05.013.
8. Donatti, C.I.; Harvey, C.A.; Martinez-Rodriguez, M.R.; Vignola, R.; Rodriguez, C.M. Vulnerability of smallholder farmers to climate change in Central America and Mexico: current knowledge and research gaps. *Climate and Development* **2019**, *11*, 264 – 286. https://doi.org/10.1080/17565529.2018.1442796.
9. Morton, J.F. The impact of climate change on smallholder and subsistence agriculture. *Proceedings of the National Academy of Sciences of the United States of America* **2007**, *104*, 19680 – 19685. https://doi.org/10.1073/pnas.0701855104.
10. Caldeira, A.A.; Wilbert, M.D.; Moreira, T.B.S.; Serrano, A.L.M. Brazilian State debt sustainability: An analysis of net debt and primary balance. *Revista de Administracao Publica* **2016**, *50*, 285 – 306. https://doi.org/10.1590/0034-7612151140.
11. Kummu, M.; Heino, M.; Taka, M.; Varis, O.; Viviroli, D. Climate change risks pushing one-third of global food production outside the safe climatic space. *One Earth* **2021**, *4*, 720 – 729. https://doi.org/10.1016/j.oneear.2021.04.017.
12. Butterbach-Bahl, K.; Dannenmann, M. Denitrification and associated soil N2O emissions due to agricultural activities in a changing climate. *Current Opinion in Environmental Sustainability* **2011**, *3*, 389–395. Carbon and nitrogen cycles, https://doi.org/https://doi.org/10.1016/j.cosust.2011.08.004.
13. Trigueiro, W.R.; Nabout, J.C.; Tessarolo, G. Uncovering the spatial variability of recent deforestation drivers in the Brazilian Cerrado. *Journal of Environmental Management* **2020**, *275*, 111243. https://doi.org/https://doi.org/10.1016/j.jenvman.2020.111243.
14. Dickinson, R.E.; Henderson-Sellers, A. Modelling tropical deforestation: A study of GCM land-surface parametrizations. *Quarterly Journal of the Royal Meteorological Society* **1988**, *114*, 439 – 462. https://doi.org/10.1002/qj.49711448009.
15. Henderson-Sellers, A.; Gornitz, V. Possible climatic impacts of land cover transformations, with particular emphasis on tropical deforestation. *Climatic Change* **1984**, *6*, 231 – 257. https://doi.org/10.1007/BF00142475.
16. Lean, J.; Warrilow, D. Simulation of the regional climatic impact of Amazon deforestation. *Nature* **1989**, *342*, 411 – 413. https://doi.org/10.1038/342411a0.
17. Shukla, J.; Nobre, C.; Sellers, P. Amazon deforestation and climate change. *Science* **1990**, *247*, 1322 – 1325. https://doi.org/10.1126/science.247.4948.1322.
18. Rosano-Peña, C.; Pensado-Leglise, M.D.R.; Marques Serrano, A.L.; Bernal-Campos, A.A.; Hernández-Cayetano, M. Agricultural eco-efficiency and climate determinants: application of dea with bootstrap methods in the tropical montane cloud forests of Puebla, Mexico. *Sustainable Environment* **2022**, *8*. https://doi.org/10.1080/27658511.2022.2138852.

19. Werth, D. The local and global effects of Amazon deforestation. *Journal of Geophysical Research* **2002**, *107*. https://doi.org/10.1029/2001jd000717.
20. Simberloff, D. The role of science in the preservation of forest biodiversity. *Forest Ecology and Management* **1999**, *115*, 101–111. https://doi.org/https://doi.org/10.1016/S0378-1127(98)00391-0.
21. Morton, D.; Defries, R.; Randerson, J.; Giglio, L.; Schroeder, W.; van der Werf, G. Agricultural intensification increases deforestation fire activity in Amazonia. *Global Change Biology* **2008**, *14*, 2262 – 2275. https://doi.org/10.1111/j.1365-2486.2008.01652.x.
22. Aigner, D.; Lovell, C.; Schmidt, P. Formulation and estimation of stochastic frontier production function models. *Journal of Econometrics* **1977**, *6*, 21–37. https://doi.org/https://doi.org/10.1016/0304-4076(77)90052-5.
23. Färe, R. *Fundamentals of production theory*; Springer Berlin, Heidelberg, 1988; pp. IX, 163. https://doi.org/10.1007/978-3-642-51722-8.
24. SHEPHARD, R.W. *Theory of Cost and Production Functions*; Princeton University Press, 1970.
25. Farrell, M.J. The Measurement of Productive Efficiency. *Journal of the Royal Statistical Society. Series A (General)* **1957**, *120*, 253–290. https://doi.org/10.2307/2343100.
26. Banker, R.; Charnes, A.; Cooper, W. SOME MODELS FOR ESTIMATING TECHNICAL AND SCALE INEFFICIENCIES IN DATA ENVELOPMENT ANALYSIS. *Management Science* **1984**, *30*, 1078 – 1092. https://doi.org/10.1287/mnsc.30.9.1078.
27. Charnes, A.; Cooper, W.; Rhodes, E. Measuring the efficiency of decision making units. *European Journal of Operational Research* **1978**, *2*, 429 – 444. https://doi.org/10.1016/0377-2217(78)90138-8.
28. Wilson, P.W. Detecting Outliers in Deterministic Nonparametric Frontier Models with Multiple Outputs. *Journal of Business & Economic Statistics* **1993**, *11*, 319–323. https://doi.org/10.2307/1391956.
29. Yang, Q.; Wan, X.; Ma, H. Assessing Green Development Efficiency of Municipalities and Provinces in China Integrating Models of Super-Efficiency DEA and Malmquist Index. *Sustainability* **2015**, *7*, 4492–4510. https://doi.org/10.3390/su7044492.
30. Rosano-Pena, C.; De Almeida, C.A.R.; Rodrigues, E.C.C.; Serrano, A.L.M. Spatial dependency of eco-efficiency of agriculture in São Paulo. *Brazilian Business Review* **2020**, *17*, 328 – 343. https://doi.org/10.15728/BBR.2020.17.3.5.
31. Bogetoft, P.; Otto, L. *Benchmarking with DEA, SFA, and R*; Springer New York, 2011. https://doi.org/10.1007/978-1-4419-7961-2.
32. Banker, R.D.; Chang, H. The super-efficiency procedure for outlier identification, not for ranking efficient units. *European Journal of Operational Research* **2006**, *175*, 1311 – 1320. https://doi.org/10.1016/j.ejor.2005.06.028.
33. Malmquist, S. Index numbers and indifference surfaces. *Trabajos de Estadistica* **1953**, *4*, 209 – 242. https://doi.org/10.1007/BF03006863.
34. Efron, B. Bootstrap Methods: Another Look at the Jackknife. *The Annals of Statistics* **1979**, *7*, 1–26.
35. Simar, L.; Wilson, P.W. Sensitivity analysis of efficiency scores: How to bootstrap in nonparametric frontier models. *Management Science* **1998**, *44*, 49 – 61. https://doi.org/10.1287/mnsc.44.1.49.
36. Marques Serrano, A.L.; Saiki, G.M.; Rosano-Penã, C.; Rodrigues, G.A.P.; Albuquerque, R.d.O.; García Villalba, L.J. Bootstrap Method of Eco-Efficiency in the Brazilian Agricultural Industry. *Systems* **2024**, *12*. https://doi.org/10.3390/systems12040136.
37. Simar, L.; Wilson, P.W. Non-parametric tests of returns to scale. *European Journal of Operational Research* **2002**, *139*, 115 – 132. https://doi.org/10.1016/S0377-2217(01)00167-9.
38. Simar, L. Estimating efficiencies from frontier models with panel data: A comparison of parametric, non-parametric and semi-parametric methods with bootstrapping. *Journal of Productivity Analysis* **1992**, *3*, 171 – 203. https://doi.org/10.1007/BF00158775.

 doi:10.20944/preprints202406.0411.v1

39. Basset-Mens, C.; Ledgard, S.; Boyes, M. Eco-efficiency of intensification scenarios for milk production in New Zealand. *Ecological Economics* **2009**, *68*, 1615–1625. Eco-efficiency: From technical optimisation to reflective sustainability analysis, https://doi.org/https://doi.org/10.1016/j.ecolecon.2007.11.017.
40. Rosano-Peña, C.; Guarnieri, P.; Sobreiro, V.A.; Serrano, A.L.M.; Kimura, H. A measure of sustainability of Brazilian agribusiness using directional distance functions and data envelopment analysis. *International Journal of Sustainable Development & World Ecology* **2014**, *21:3*, 210–222. https://doi.org/https://10.1080/13504509.2014.901992.