Spatial Dependency of Eco-Efficiency of Agriculture in São Paulo

Carlos Rosano-Peña[1]
gmcrosano@gmail.com | ⓘ 0000-0002-6868-9284

Carlo Aleksandr Rosano de Almeida[1]
carloaleksandr@hotmail.com | ⓘ 0000-0002-1504-8969

Evaldo César Cavalcante Rodrigues[1]
evaldocesarcr@gmail.com | ⓘ 0000-0001-7176-2940

André Luiz Marques Serrano[1]
andrelms@unb.br | ⓘ 0000-0001-5182-0496

ABSTRACT

This research presents an eco-efficiency index for the municipalities of São Paulo, indicating how much it would be possible to maximize economic and environmental objectives, taking into account the best practices for the municipalities of this region. In this vein, we used the Data Envelopment Analysis method with directional distance functions based on the classic variables of multiproduct production function and the internalization of two externalities (one positive and one negative). The study also used the tools of exploratory spatial data analysis to verify the spatial autocorrelation and spatial heterogeneity of the calculated index. The results indicate that, on average, the analyzed municipalities are able to expand the production and forested areas by 59% and also to reduce degraded areas and inputs in the same proportion. Spatial analysis demonstrated the existence of spatial heterogeneity and autocorrelation between municipalities and the formation of large clusters. Based on these results, priorities for environmental intervention in the state are defined.

KEYWORDS

Data envelopment analysis, Spatial data analysis, Agriculture, São Paulo

1. INTRODUCTION

Growth in agricultural productivity and efficiency is a relevant factor for socio-economic development. It is responsible for expanding the domestic market, increasing the production of relatively lower-priced food and raw materials, as well as demanding a greater amount of inputs and agricultural equipment and financial services (Delgado, 2001). Besides, the advancement of productivity and efficiency allowed the release of human resources for non-agricultural sectors, strengthening the concentration of urban-industrial production mode and the growth of quality and coverage of public services. This has led to changes in the political power structures in which the rural aristocracy has lost influence. Thus, a good pace of productivity growth and agricultural efficiency is a key element for countries like Brazil to achieve the standard of living of the most developed countries.

However, agricultural intensification has caused significant environmental impacts on the world's terrestrial and aquatic ecosystems (Tilman, 1999). For example, the growth of monoculture has led to forest destruction, loss of genetic biodiversity, and soil erosion, while mechanization has increased non-renewable energy consumption. Intensified use of insecticides, pesticides, and nitrogen fertilizers has increased the concentration of greenhouse gases in the atmosphere and, together with irrigation systems, led to eutrophication, pollution, and depletion of water resources (Intergovernmental Panel on Climate Change – IPCC, 2006).

In the agriculture of the richest and most populous state in Brazil, São Paulo, the environmental problem is no different. Although the problem is not geographically uniform, deforestation, soil and water contamination, and greenhouse gas (GHG) emissions in the state's agricultural sector are growing problems. Of the accumulated GHG emissions from Brazilian agriculture (12,970 Mt CO2) between 1970 and 2013, São Paulo accounts for about 9% (Institute of Forest and Agricultural Management and Certification, 2015). These emissions have predominantly been derived from beef and dairy cattle (56%), the use of synthetic fertilizers (19%), and sugarcane cultivation (10%).

In this regard, the following questions remain open: a) Is it possible to increase production while reducing environmental impacts and the use of non-renewable natural resources? b) How is eco-efficiency distributed geographically in São Paulo?

To answer these questions, several tools have been developed to measure the environmental impact of productive activities (Van Passel, Nevens, Mathijs, & Van Huylenbroeck, 2007; Van Passel, Van Huylenbroeck, Lauwers, & Mathijs, 2009). One is the life cycle accounting (LCA) method, which maps production processes and impacts (carbon footprints) on different phases, from production to consumption and recycling, identifying possible improvements. However, this method has some limitations, including the difficulties of estimating carbon emissions in each regional context and the monetary quantification of this value.

Other methods use popular efficient boundary techniques (Lampe & Hilgers, 2015). These are based on the representation of the Production Possibility Set (CPP), consisting of inputs, desired products, and unwanted products, as well as efficient and inefficient productive units. Thus, best practices will form the CPP boundary and inefficient ones (surviving due to market failures) will be below that boundary. This allows for the possibility of Pareto improvements, producing more desired products and reducing input consumption and environmental impact. Using an English expression from the business world, these improvements could be called win-win situations. These applications employ both parametric and nonparametric procedures,

which, starting from different assumptions, have advantages and disadvantages. The former uses the Stochastic Frontier Analysis (SFA) method. This estimates the CPP boundary by defining a functional econometric relationship between products and input, which allows us to decompose the boundary deviation into stochastic noise and inefficiency. Some researchers have used this method in the study of ecoefficiency: for example, Orea and Wall (2017), who made an empirical analysis with data from a sample of Spanish dairy farms, and Ho, Hoang, Wilson, and Nguyen (2018), who compared the performance of conventional and certified coffee farms in Vietnam. Nonparametric Methods, on the other hand, used Data Envelopment Analysis (DEA), which easily models both multi-product technologies and the internalization of externalities associated with the production process. It also represents technology through distance functions, but they are gauged with mathematical programming problems, without the need to define a stochastic production function and a type of distribution (behavior) of unknown errors beforehand, free from possible failures arising from these specifications. The main disadvantage of DEA is its deterministic approach, which ignores the random disturbances of the production process. Even so, the use of this method predominates in the literature on efficiency and productivity (Orea & Wall, 2017) and it is possible to find records from studies of European Community agriculture (Rybaczewska-Błażejowska & Gierulski, 2018), China (Xing, Wang, & Zhang, 2018), Chile (Angulo-Meza, Gonzalez-Araya, Iriarte, Rebolledo-Leiva, & Mello, 2019) and the United States (Dong, Mitchell, & Colquhoun, 2015), among others.

The review of the Brazilian literature showed that the use of efficient frontier techniques to study the eco-efficiency of national agriculture is still incipient. With the use of multiproduct stochastic borders, only the work of Rosano-Peña et al. (2018) was found, who used hyperbolic distance functions to estimate the eco-efficiency of Amazonian agriculture. Among the works that employ nonparametric methods, it is worth mentioning the works of Padrão, Campos, Lirio, and Silva (2012), which compare the technical and environmental efficiency of agricultural production in the Legal Amazon and estimate the opportunity cost of the Forest Code; Rosano-Peña and Daher (2015), who evaluate the eco-efficiency and sustainability of agriculture in Brazilian states; Campos, Coelho, Gomes and Mattos (2014), who study the economic and environmental performance of dairy farmers in Minas Gerais using the DEA model combined with the material balance approach and Alencar, Rosano-Peña, Guarnieri and Serrano (2019) which estimate eco-efficiency and the shadow price of greenhouse gas emissions in Brazilian pig farming.

In this context, the present work seeks to fill a gap in the literature on the eco-efficiency of agriculture in São Paulo municipalities. More precisely, its objective is to estimate an eco-efficiency indicator that, while satisfying Pareto's optimal concept, simultaneously maximizes economic and environmental objectives, based on best practices. Therefore, the DEA method, with directional distance functions used, is based on the classical variables of agricultural activity and the internalization of two externalities (one positive and one negative), since it is the most popular and suitable tool used in the literature to estimate eco-efficiency. Besides, for the examination of the results, the autocorrelation and spatial heterogeneity of the calculated index are verified through the use of exploratory spatial data analysis (ESDA) techniques.

Thus, it is believed that the results of the work can become important subsidies for the definition of integrated regional public policies consistent with the maximization of social welfare, to optimize the sustainability of São Paulo's agriculture.

2. THEORETICAL AND METHODOLOGICAL FRAMEWORK

According to Zhang, Bi, Fan, Yuan, and Ge (2008), the concept of eco-efficiency comes from the 1970s, and it has been used by Freeman, M. A., and McIntyre, J.R. as the term that expresses environmental economic efficiency. Subsequently, eco-efficiency is disseminated by World Business Council for Sustainable Development (Verfaillie & Bidwell, 2000), which defines it as the ability of a production unit (farm, industry, sector, country, etc.) to produce more and better, with less use of resources and with minimal environmental impact, thus seeking to build a more sustainable society.

From the theoretical point of view, eco-efficient units form the boundary of the production possibility set (PPS). Therefore, eco-efficiency is the ability of a company to produce a given amount of desired product with the least amount of inputs and environmental impact; or, equivalently, as the ability to maximize production with a given amount of unwanted inputs and by-products. In other words, eco-efficiency, associated with a given input combination, is achieved at the potential output frontier, an optimal point in Pareto, when there is no other production process or combination of processes that can produce the same output level, with less impacts on the environment and consuming fewer inputs. This means that eco-efficient production units are below the PPS boundary. Consequently, according to Färe, Grosskopf, and Weber (2006), a firm's level of eco-efficiency can be measured by its distance from its frontier. That is, an organization's eco-efficiency can be measured by comparing its performance with best practices.

One of the most notorious methods for estimating eco-efficiency is the directional distance function developed by Chung, Färe and Grosskopf (1997), Färe and Grosskopf (2000) and Färe et al. (2006) to include unwanted by-products. According to the authors, this method emerges as an alternative to Shephard's radial distance functions to treat desirable and undesirable outputs simultaneously and asymmetrically. Also, it allows, a priori, to define different projection directions of eco-inefficient points on the efficient frontier employing a directional vector $(g = -g_x, g_y, -g_b)$, offering a set of options for achieving eco-efficiency that can even improve a group of variables without worsening the behavior of others. The directional distance function can be expressed as follows:

$$\vec{D}\,[x, y, b; -g_x, g_y, -g_b)] = \text{Max}\,\{\beta:(x-\beta g_x, y + \beta g_y, b-\beta g_b) \in \text{PPS}\} \qquad (1)$$

where β, the optimal value to be estimated, indicates the percentage by which the evaluated productive unit could increase all the desirable products (y) and simultaneously reduce the inputs (x) and negative externalities (b) when the direction, a priori defined by the researcher, the direction vector is $(-g_x=1, g_y=1, -g_b=1)$. Therefore, β is greater than or equal to zero: β = 0 means that the evaluated unit is eco-efficient and β > 0 eco-inefficient.

Arandia and Aldanondo-Ochoa (2011), following Färe et al. (2006) state that the directional distance functions and the β for each unit evaluated can be estimated from the calculation of the following linear programming problem (PPL):

$$\vec{D} = (x, y, b; -g_x, g_y, -g_b) = \text{Max}\,\beta^i \qquad (2)$$

s.t

$$(1+\beta^i g_y) * y^i - s_y^i = Yz \tag{2.1}$$

$$(1-\beta^i g_b) * b^i + s_b^i = Bz \tag{2.2}$$

$$(1-\beta^i g_x) * x^i + s_x^i = Xz \tag{2.3}$$

$$z \geq 0 \tag{2.4}$$

where x^i, y^i e b^i denote, respectively, the input vector, the desired output and undesired output of the ith evaluated unit s^i; are the clearances of the respective variables; X is the matrix of k inputs of n evaluated units, Y denotes the matrix of desired products of order (pxn) and B is the matrix of undesired products of order (qxn); z is the intensity vector of each firm in the definition of the efficient frontier, that is, in the formation of linear combinations of best practices.

The linear programming problem (2) is characterized by treating unwanted products as inputs to avoid the congestion problems of these byproducts and works with the hypothesis of constant returns to scale, thus comparing unit performance with the highest environmental economic productivity. For more details, see Arandia and Aldanondo-Ochoa (2011).

Thus, for example, an index of $\beta^i = 0,20$ indicates that productive unit i, to be eco-efficient and achieve the highest productivity, should increase the value of the desired products (**y**) by 20%, as well as reduce the unwanted products (**b**) and the inputs (**x**) in the same proportion.

2.1. Exploratory Analysis of Spatial Data (ESDA)

Exploratory analysis of spatial data is a useful tool for the analysis of socioeconomic variables. For Anselin, Syabri, Smirnov, and Ren (2002), EDAA can be defined as the set of statistical and graphical techniques that describe and visualize spatial distributions of variables, identifying atypical local points, forms of association (spatial autocorrelation) and structures in geographical space (spatial heterogeneity).

The effects of correlation and spatial heterogeneity are relevant for the identification of patterns and anomalies of the geographical distribution of indicators that are not obvious to the first, as well as for the elaboration of local, regional and national monitoring, planning and intervention programs. Spatial correlation can be defined as the existence of a functional relationship between what happens at one point in space and what happens elsewhere; that is when the value of a variable of interest in a certain region i depends on the value of that variable in neighboring regions j. The correlation is due to the first law of geography, which, according to Tobler (1979), says: "Geographical facts are correlated, but the closest are more correlated" (p. 519). This can be explained, for example, by the impacts of communications, transport, infrastructure, agro-industry, economy of scale, as well as the effects of the diffusion process, when innovation in one municipality is imitated and popularized in others, or spillover effects, which refer to the moment when the development of a region overflows, inducing the development of the neighboring region and regional convergence (Costa, Almeida, Ferreira, & Silva, 2013). Spatial heterogeneity, in turn, seeks the effects related to spatial or regional differentiation and is defined by the existence of groupings in the space of variables of interest (Valcarce & Serrano, 2000). According to Tobler's first law of geography (1979), greater heterogeneity is expected with increasing distance. It is also interesting to note, according to Almeida (2004), that in spatial processes there is an imbrication between these two effects, since spatial heterogeneity generates spatial dependence, and spatial dependence, in turn, can lead to heterogeneity.

To analyze spatial correlation and heterogeneity, the first step is to define the weight matrix (W), which defines the spatial connectivity of a set of areas (municipalities, states, etc.). Each observation in said matrix w_{ij} represents a normalized measure of proximity between area i and area j (Baumont, Ertur, & Le Gallo, 2004). This matrix is also known as a neighborhood or contiguity matrix. From this matrix, it is possible to extract global and local spatial association measures.

The global spatial association is defined as the coincidence of spatial ubiquity of values and manifests itself when the high or low values of a variable tend to cluster in space. It can be assessed by the Moran Index (I) statistic, which estimates the degree of linear association as a whole between the observed values and the weighted average of the neighborhood values, called the lagged value (Anselin, Syabri, Smirnov, & Ren, 2002).

Moran's I is defined as:

$$I = \frac{n \sum_i \sum_j w_{ij}(x_i - \acute{x})(x_j - \acute{x})}{(\sum_i \sum_j w_{ij}) \sum_i (x_i - \acute{x})} \tag{3}$$

where: n is the total area (municipalities), w_{ij} is the neighborhood spatial weight measure, x_i and x_j denote the observed values of the variable of interest (eco-efficiency) for the municipalities i and j respectively, and $\acute{x}$ is the average of these values.

Thus, Moran's I will be computed only for neighbors in space, as established by the w_{ij} weights. Its value ranges from –1 to +1. An equal zero value indicates no spatial correlation (differences between neighbors). Positive values near the unit indicate positive spatial autocorrelation, ie the existence of areas with similar values between neighbors, and negative values near the unit indicate negative spatial autocorrelation. On the other hand, values close to zero suggest a very low spatial autocorrelation between the x value of the municipality and the value of its neighbors.

Given the Moran index, it is necessary to test the hypothesis that the result is nonzero. There are two methods for testing the hypothesis. In the first one, the distribution of variable x is observed. If evenly distributed, the probability of 0.05 is suggested as the cutoff level to reject the hypothesis of spatial autocorrelation and this probability is applied to a normal distribution. The second method applies to asymmetric data, such as eco-efficiency indexes, which use the Monte Carlo permutation test. In this case, different random permutations of the values of the variable of interest (x) are generated to the regions, resulting in a new spatial arrangement, in which the values are redistributed between the areas. Thus, an empirical distribution of Moran's I can be constructed. If the originally measured Moran I value corresponds to an "extreme" of the simulated distribution, it is a value with statistical significance.

One way to interpret Moran's statistics I is through his scatter diagram. According to Figure 1, it presents, in the Cartesian plane, the standardized value (z) of the variable x for each of the units in the abscissa and, in the ordinate, the average of the standardized value of the same variable for the neighbors of this unit w_z. The diagram is complemented by the representation of a regression line whose slope indicates the value of Moran's I, which, for the example of Figure 1, is 0.86. Thus, the greater the slope of the line relative to the horizontal axis, the greater the value of spatial autocorrelation and vice versa.

However, when the study region is large and many municipalities are analyzed, different spatial autocorrelation regimes are likely to occur in the studied subregions. This can camouflage various local patterns of spatial autocorrelation. In these cases, Moran's global indexes would not be sufficient to explain the spatial distribution of the studied region. Therefore, Anselin (1995) suggests a new indicator with the ability to observe local statistically significant linear association

patterns, indicating the existence of local spatial clusters and regions that contribute most to the existence of spatial autocorrelation. The indicator is called LISA (Local Indicator of Spatial Association) and decomposes the global autocorrelation indicator into local contributions by indicating four categories, each individually corresponding to a quadrant in Moran's scatter diagram. Thus, according to Almeida (2004), the Moran diagram can be divided into four quadrants, which correspond to four patterns of spatial local association among regions and their neighbors.
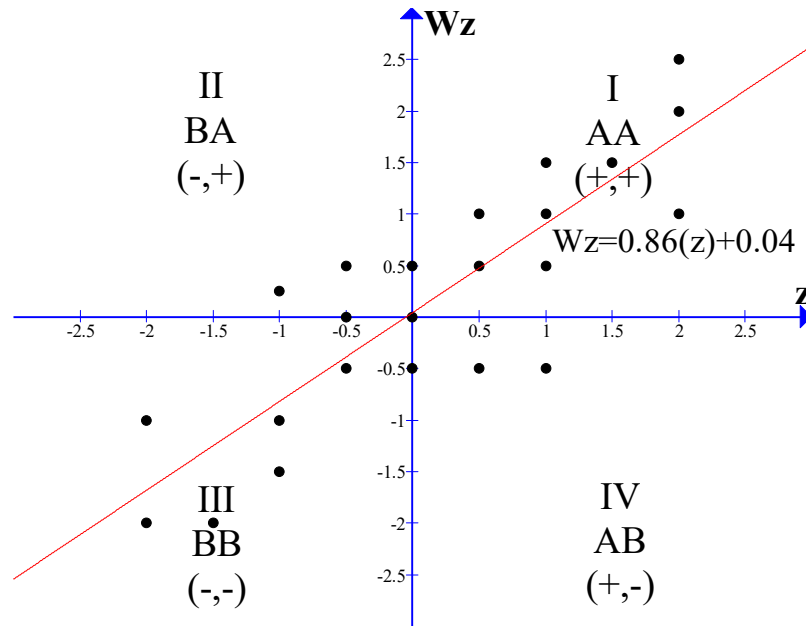

Figure 1. Moran Scatter Diagram

Quadrant I (located in the upper-right) shows the regions that present high values for the variable under analysis, surrounded by regions that also present values above the average of the variable. This quadrant is classified as high-high (HH).

Quadrant II (located in the upper-left) shows the regions with low values, surrounded by neighbors with high values. This quadrant is generally classified as low-high (LH).

Quadrant III (located in the lower-left corner) consists of regions with low values for the variable of interest, surrounded by regions with low values. This quadrant is classified as low-low (LL).

Quadrant IV (located in the lower-right corner) is formed by regions with high values for the variable under analysis, surrounded by regions with low values. This quadrant is classified as high-low (HL).

Moran's I local statistics can be obtained by the following formula:

$$I_i = \frac{(x_i - \acute{x}) \sum_j w_{ij} (x_j - \acute{x})}{\left(\sum_i (x_i - \acute{x})^2\right)/n} \tag{4}$$

Similarly to the global indicators, the significance of Moran's local index (I_i) must be evaluated using the hypothesis of normality and or simulation of random exchange distribution (Anselin, 1995). Thus, a statistically significant and positive value of the local Moran I reveals the existence of a cluster (similar municipalities, high-high or low-low). On the contrary, a negative value suggests an outlier, a municipality that is bypassed by different municipalities (high-low or low-high).

3. THE OBJECT OF STUDY AND VARIABLES

São Paulo has 645 municipalities grouped in 15 mesoregions and generates the largest GDP in Brazil. Much of its GDP comes from industry, services, finance, and, to a lesser extent, agriculture. Even so, since 2010, the state occupies second place in the ranking of the Brazilian agricultural GDP, generating more than 11% of the total. The main agricultural product from São Paulo are oranges, which reaches 80% of national production. The second is sugarcane, which accounts for 60% of the country's production. The mesoregions with the highest agricultural GDP are: Ribeirão Preto, Sao Jose do Rio Preto, Campinas, and Bauru.

As observed in Figure 2, in the state of São Paulo there are two biomes: Atlantic Forest and Cerrado. According to Maffei (2010), of the 645 municipalities of the state, 176 are part of the Cerrado and 469, of the Atlantic Forest.
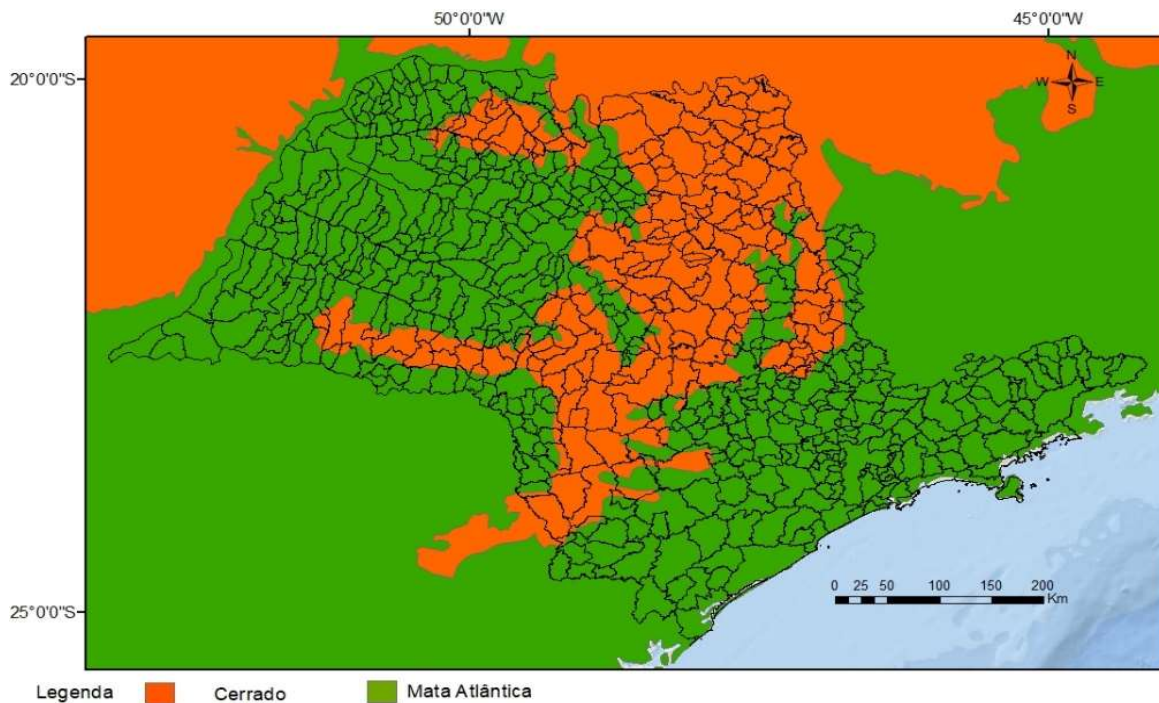
Figure 2. Map of São Paulo State Biomes (Cerrado and Atlantic Forest)
Source: Own authorship with data from the IBGE digital cartographic database. Available at: <http://www.mapas.ibge.gov.br/>.

To carry out this research, information was taken from the agricultural census conducted in 2006 (Brazilian Institute of Geography and Statistics [IBGE], 2010), since data from the latest census (2017) are not available. For the 645 municipalities of São Paulo, the classic inputs and outputs of agricultural production were considered, but one innovation was the incorporation of

a positive and a negative externality. As in most cases (Gomes, 2008), the inputs used in modeling were: x_1 – workforce on properties in number of people; x_2 – capital estimated by depreciation (10% of the value of fixed assets); x_3 – total area of establishments in hectare; x_4 – other current expenses incurred by producers, called costing. Outputs consider three types of products: y_1 – desirable product – the total value of production; y_2 – environmental desirable product – areas of natural forests and forests preserved on hectare properties; b_1 – environmental undesirable product – areas of degraded land on hectare properties (census does not include abandoned land).

4. RESULTS

Results were obtained based on the methods described and the selected variables, whose examination was conducted in two parts. In the first one, the eco-efficiency indexes are analyzed. In the second, the spatial distribution is evaluated.

4.1. ECO-EFFICIENCY INDEXES

For the 645 municipalities of São Paulo, the eco-efficiency indexes β were estimated. Table 1 shows the statistical summary of the municipalities. It is observed that 26 municipalities obtained a β = 0, indicating that these units evaluated are efficient, the state benchmarks. The average index is 0.59. This indicates that, on average, São Paulo municipalities can increase the value of total production and preserved areas by 58%, as well as reduce degraded areas and inputs by the same proportion. This can be achieved only by mimicking the region's benchmarks.

Looking at the quartiles, it is also noted that 75% of the municipalities of São Paulo have an eco-efficiency index above 0.482 and 25% have the worst rates, above 0.77. The median is 0.6633, that is, 50% of the municipalities of the state of São Paulo have an eco-efficiency level of 0.6633 or higher. The breadth of the results shows a great heterogeneity in the region, with the 26 eco-efficient municipalities characterized as outliers, as shown in Figure 3.

Table 1. *Summary of eco-efficiency indexes, β (beta)*

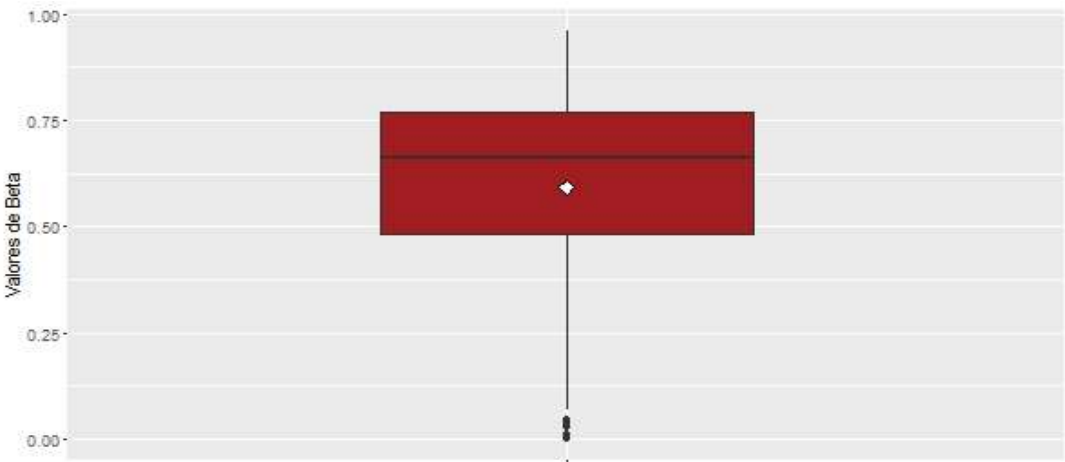|			
1º Quartile	0,4820	3º Quartile	0,7702
Average	0,5917	Standard deviation	0,2369
Standard error	0,0093	Minimum value (26 municipalities)	0
Median	0,6633	Maximum value (1 municipality)	0,9617

Figure 3. Boxplot of β from São Paulo municipalities

Table 2 also reveals the performance of the mesoregions. These results identify the Paulista South Coast as the most eco-efficient region and Presidente Prudente and Marília as the most eco-inefficient. These three mesoregions are part of the Atlantic forest biome. It is also observed that Ribeirão Preto, in the Cerrado, is the macroregion with the largest number of eco-efficient municipalities, with approximately 31% of the benchmark municipalities of São Paulo.

This is observed in more detail in Figure 4, which georeferences the stratification of the Beta eco-efficiency indexes of the municipalities. The brighter the green, the higher the efficiency (the lower the Beta); the more intense the red, the more eco-inefficient (the higher the beta). Intuitively, this map already shows the presence of spatial autocorrelation in several regions. On the one hand, it is noted the homogeneity of the eco-efficiency of the mesoregion of the Paulista South Coast. On the other hand, there is the relative homogeneity of eco-efficiency in the western regions of the state, especially in Presidente Prudente. Comparing Figure 4 with Figure 2, we note that there is no clear relationship between biomes and eco-efficiency indexes.

Table 2. *Average of* β *indexes and number of eco-efficient municipalities by mesoregions*

Mesoregion	Average	Eco-efficient	Mesoregion	Average	Eco-efficient	Mesoregion	Average	Eco-efficient
Litoral Sul Paulista	0.3097	3	Vale do Paraíba Paulista	0.5554	1	Piracicaba	0.6318	0
Ribeirão Preto	0.4835	8	Bauru	0.5854	2	São José do Rio Preto	0.6405	2
Araraquara	0.5032	2	Campinas	0.5911	2	Araçatuba	0.6558	0
Metropolitana de São Paulo	0.5126	2	Itapetininga	0.6039	1	Marília	0.6566	1
Macro Metropolitana Paulista	0.5437	2	Assis	0.6277	0	Presidente Prudente	0.7624	0

Classification — Color

0 - Eco-efficiency

[0.005 - 0.5199) Quartile with the lowest eco-inefficiency

[0.52 - 0.6685) Quartile with average eco-inefficiency

[0.6685 - 0.772) Quartile with high eco-inefficiency

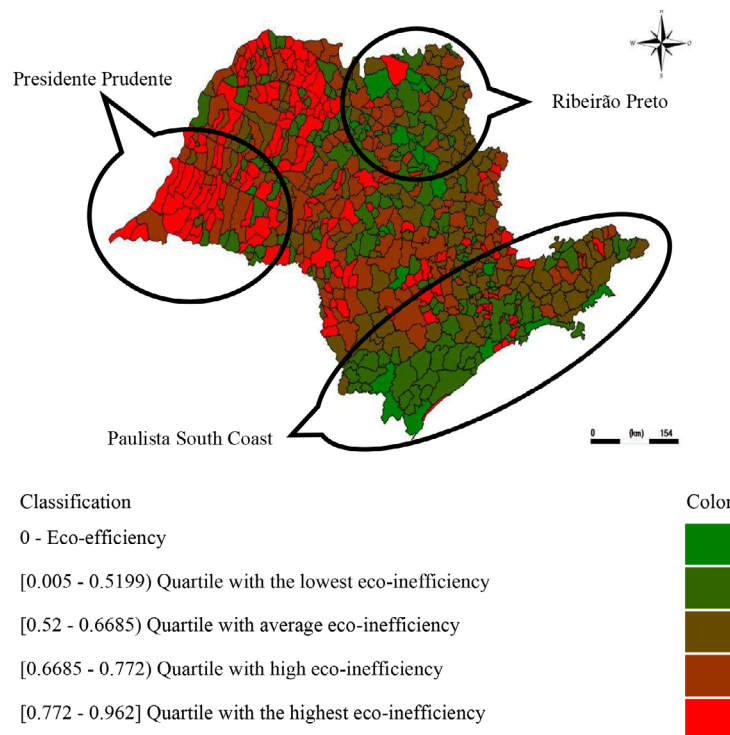[0.772 - 0.962] Quartile with the highest eco-inefficiency

Figure 4. São Paulo Beta Value Distribution

Table 3 shows the absolute values of improvements needed for the eco-efficiency of the municipalities of São Paulo. These results were obtained considering both the eco-efficiency indexes (β) and the s-slacks estimated by the PPL (2). It is evident that the economy of economic and environmental resources is substantial and the potential for growth of production is no less important (R$ 15 billion). Ribeirão Preto stands out in this table since it is the mesoregion with the highest potential level of production growth (y_1) and with 58 eco-inefficient municipalities. In the opposite direction, the metropolitan mesoregion stands out, since it is the least agricultural.

Table 3. *Improvements to the eco-efficiency of municipalities by mesoregions*

Mesoregion	x_4 – Costing (R$ 1000)	x_2 – Capital (R$ 1000)	x_1 – Labor	x_3 – Area (ha)	b_1 – Degraded areas (ha)	y_2 – Preserved Areas (ha)	y_1 – Production (R$ 1000)
São José do Rio Preto	-1093298.52	-385593.7674	-121113.31	-1501945.244	-2818.20	93301.39	2294347.732
Ribeirão Preto	-2024594.158	-291400.0873	-76799.41	-1224187.28	-2727.76	111049.39	2589903.198
Araçatuba	-499939.4213	-192328.3032	-40178.20	-842780.6204	-769.27	64995.92807	915543.693
Bauru	-2369549.61	-325219.3954	-61455.41	-1308267.347	-1590.63	128133.13	1633908.309
Araraquara	-363072.148	-106119.7609	-25515.39	-332278.1061	-242.72	30591.14	748303.235
Piracicaba	-644422.3557	-128956.0915	-20021.51	-346609.1077	-270.48	35982.66	776792.578
Campinas	-584445.3004	-347637.6772	-68574.81	-602350.3322	-883.30	61683.68	1535215.338
Presidente Prudente	-926674.7411	-242686.3095	-72928.17	-1587731.487	-772.02	80282.89	1261176.952
Marília	-173592.9446	-91427.7073	-17167.59	-444737.9349	-728.55	36332.54	423524.904
Assis	-542317.5049	-157226.0469	-38271.74	-665007.6009	-485.30	40980.62	896520.589
Itapetininga	-549819.333	-206426.2561	-45265.15	-769526.032	-1916.94	109314.37	787721.225
Macro Metropolitana	-225236.2114	-148101.3339	-46572.07	-352233.8848	-865.86	49860.39	561801.461
Vale do Paraíba Paulista	-175652.2231	-96874.22393	-26220.49	-435660.3017	-1127.04	80463.20	227654.185
Litoral Sul Paulista	-47502.6781	-20262.67816	-11120.44	-109981.4008	-360.77	38469.29	230898.187
Metropolitana de S.P.	-41927.77478	-30748.20555	-13322.90	-59184.60794	-258.71	8387.73	112949.378
Total for state	-10262044.92	-2771007.844	-684526.58	-10582481.29	-15817.50	969828.35	14996260.97

4.2. Exploratory analysis of spatial data

Municipal eco-efficiency indexes were also analyzed using spatial autocorrelation techniques. The Moran global index, based on a first-order normalized Queen neighborhood matrix for the 645 municipalities, was positive (0.2271) and statistically significant via normal distribution (p = 0.0000) Monte Carlo simulation (p = 0, 0000). This indicates the existence of positive autocorrelation among the eco-efficiency indexes of the municipalities of São Paulo. In other words, the municipalities and their neighbors as a whole have similar eco-efficiency values. Therefore, it is important to consider space as a determining factor for eco-efficiency.

For the analysis of eco-efficiency indexes in more detail, the local Moran index (LISA) was used. Results are shown in Figure 5.

In Figure 5A, each municipality is classified according to its position in relation to the quadrants of the Moran diagram. The red color represents quadrant I – Q1 (high-high), which brings together 273 municipalities, equivalent to 42.3% of the total. Green represents Q3 (low-low), which brings together 155 municipalities (24%). The sum of both 428 (66.3%) confirms the majority and the overall positive spatial dependence on environmental economic performance. The remaining municipalities are characterized by the color orange, which represents Q2 (high-low), and the color purple, which represents quadrant Q4 (low-high). These latter municipalities are outliers, as they do not follow the same spatial dependence process presented by most.

Based on this stratification, four priorities for environmental intervention in the state can be defined:

- highest priority: municipalities aggregated in Q1 – (HH) with high eco-efficiency indexes;

- first intermediate priority: municipalities aggregated in Q4 – (LH), where there is the problem of "islands of excellence", whose experiences should expand in the vicinity;

- second intermediate priority: municipalities aggregated in Q2 – (HL), where the diffusion of the best practices of the surroundings should be induced;

- lower priority: municipalities aggregated in Q3 – (LL) with low eco-inefficient indexes.

Figure 5B shows the most significant clusters (>p=0.05) of local association measures (LISA). In other words, it shows the municipalities or areas where spatial dependence is more pronounced. It is possible to verify the presence of two large low-low clusters, that is, conglomerates that have low eco-inefficiency levels in relation to the state average. They are surrounded by neighbors also with low eco-inefficiency. The largest is formed by some municipalities of the mesoregions of Paulista South Coast, Itapetininga and Metropolitana de São Paulo. The explanation of this behavior must be the low productivity, due to the fact that this region is of great slopes, with poor acid soils and low potential for extensive agriculture. This region has the largest area of forests and native vegetation still concentrated in of São Paulo, according to the Forest Inventory of Natural Vegetation of the State of São Paulo (Secretariat of the Environment/Forest Institute, 2005). The second-largest cluster is made up of the Ribeirão Preto and Araraquara mesoregions, economies strongly focused on agribusiness, with large agro-industries and logistics infrastructure, where are the municipalities with the largest sugarcane production in the state and Brazil. In addition, this second cluster is located in the Cerrado region, whereby law the native forest preservation area is greater (35%) than that required in the Atlantic Forest biome.

The map also features a large high-high eco-inefficient cluster, which should be considered as the most critical area for environmental intervention. It is formed by the mesoregion of municipalities of Presidente Prudente. According to Fundação Sistema Estadual de Dados – Seade. (2016), the economy of this region is practically based on extensive cattle raising, a sector of great environmental impact. In addition, this is one of the regions with the lowest native vegetation cover in the state (Secretaria do Meio Ambiente/Instituto Florestal, 2005).
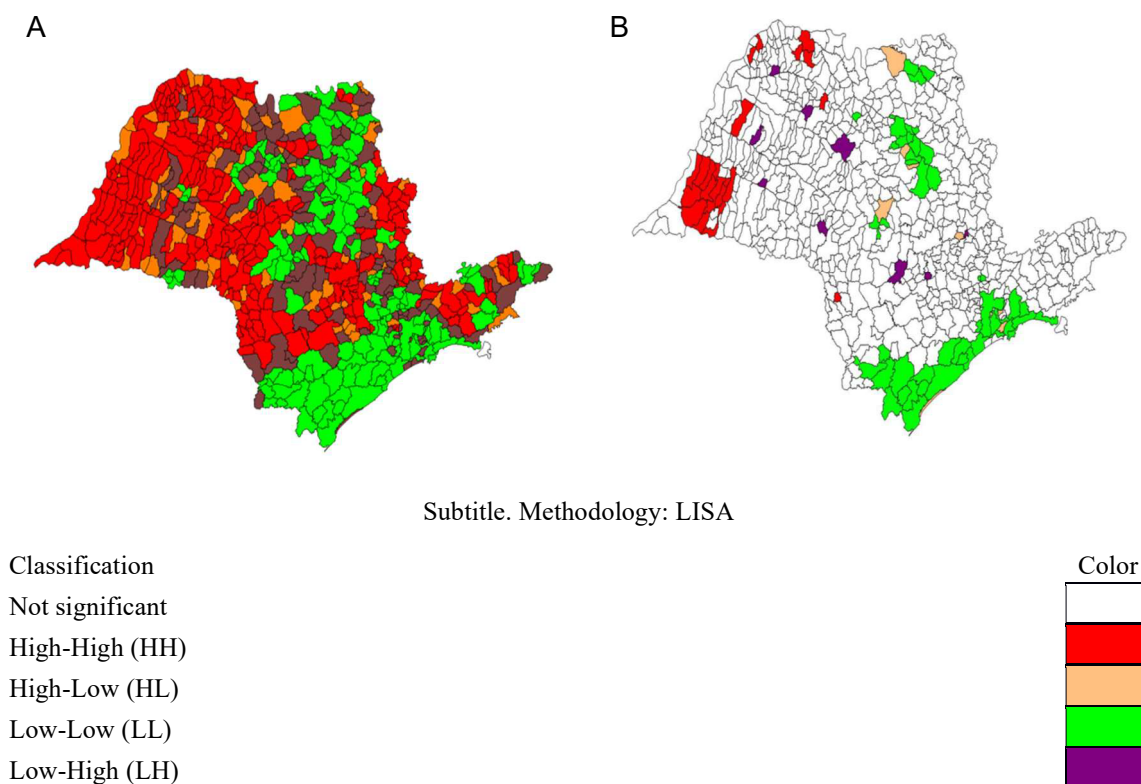
A

B

Subtitle. Methodology: LISA

Classification	Color
Not significant	
High-High (HH)	
High-Low (HL)	
Low-Low (LL)	
Low-High (LH)	

Figure 5. A, Distribution of spatial dependence classification of ecoefficiency. B, Spatial clusters for eco-efficiency index

5. CONCLUSIONS

The present research estimated an eco-efficiency index for São Paulo State's agriculture based on the non-parametric method Data Envelopment Analysis with directional distance functions and the Exploratory Spatial Data analysis techniques. Thus, it meets an important demand regarding the study of economic-environmental efficiency and its spatial distribution.

The results indicate that, on average, São Paulo municipalities can increase production and forest areas by 59%, as well as reduce degraded areas and inputs by the same proportion, only by reference to the best practices in the region. These indicators reinforce the hypothesis that it is possible to formulate policies consistent with the maximization of social welfare, while simultaneously optimizing economic and environmental objectives. Therefore, it can be concluded that the discussion of economic and environmental issues does not necessarily result in a trade-off or a zero-sum game.

The results of the spatial eco-efficiency analysis showed that there is no clear relationship among biomes and eco-efficiency indexes, but there is spatial autocorrelation among the municipalities as a whole. Best practices are found in the municipalities located in the Paulista South Coast and Ribeirão Preto mesoregions, and the worst ones are in the western region of the state.

Based on the local Moran index, two large significant LL-type clusters and one HH-type cluster were also found. The first of the LL type is located in the Paulista South Coast, Itapetininga and Metropolitana mesoregions of São Paulo and the other in the Ribeirão Preto and Araraquara mesoregions. The cluster type HH is located in the Presidente Prudente mesoregion.

These results allow the definition of priority levels for environmental intervention in the state. Also, they indicate that state level sustainable development strategies should, one one hand, consider general spatial dependence, but, on the other, need to consider local heterogeneity in defining specific policies for each region.

It is also worth mentioning that some precautions should be taken in the results found. The tools used, like any other, have limitations. On one hand, the spatial analysis was performed by municipalities, which precludes identifying infra-municipal heterogeneity at the level of productive ownership. On the other, as explained by Rosano-Peña and Daher (2015), data envelopment analysis, being a deterministic technique and estimating relative indexes concerning best practices, is very susceptible to the data used. This means that the results are conditional on the units evaluated, the variables used in the work and the principle that all other factors involved are identical. Adding or deleting units and variables may derive other results.

Finally, it is necessary to emphasize that the extension of the methodologies used opens new opportunities for future work. The expected publication of the 2017 census data will allow, for example, to analyze eco-efficiency performance over time, spillover effects and the diffusion of technological change, as well as the time course of gaps between best and worst practices with the tendency to converge or diverge.

REFERENCES

Alencar, P. A., Rosano-Peña, C., Guarnieri, P., & Serrano, A. L. M. (2019). Ecoeficiência e preço sombra das emissões de gases de efeito estufa na suinocultura brasileira. *Revista em Agronegócio e Meio Ambiente, 12*(2), 377-408.

Almeida, E. S. D. (2004). *Curso de econometria espacial aplicada*. Piracicaba: ESALQ-USP.

Angulo-Meza, L., González-Araya, M., Iriarte, A., Rebolledo-Leiva, R., & Mello, J. C. S. de (2019). A multiobjective DEA model to assess the eco-efficiency of agricultural practices within the CF+ DEA method. *Computers and Electronics in Agriculture, 161*, 151-161.

Anselin, L. (1995). Local indicators of spatial association – LISA. *Geographical analysis, 27*(2), 93-115.

Anselin, L., Syabri, I., Smirnov, O., & Ren, Y. (2002). Visualizing spatial autocorrelation with dynamically linked windows. *Computing Science and Statistics, 33*, 1-20.

Arandia, A., & Aldanondo-Ochoa, A. (2011). Pollution shadow prices in conventional and organic farming: an application in a Mediterranean context. *Spanish Journal of Agricultural Research, 9*(2), 363-376.

Baumont, C., Ertur, C., & Le Gallo, J. (2004). Spatial analysis of employment and population density: the case of the agglomeration of Dijon 1999. *Geographical analysis, 36*(2), 146-176.

Campos, S. A. C., Coelho, A. B., Gomes, A. P., & Mattos, L. B. de (2014). Eficiência e custos associados à adequação ambiental para a produção láctea em Minas Gerais. O*rganizações Rurais e Agroindustriais, 16*(3), 324-342.

Chung, Y. H., Färe, R., & Grosskopf, S. (1997). Productivity and undesirable outputs: a directional distance function approach. *Journal of Environmental Management, 51*(3), 229-240.

Costa, C. C. de M., Almeida, A. L. T. de, Ferreira, M. A. M., & Silva, E. A. (2013). Determinantes do desenvolvimento do setor agropecuário nos municípios. *Revista de Administração, 48*(2), 295-309.

Delgado, G. C. (2001). Expansão e modernização do setor agropecuário no pós-guerra: um estudo da reflexão agrária. *Estudos avançados, 15*(43), 157-172.

Dong, F., Mitchell, P. D., & Colquhoun, J. (2015). Measuring farm sustainability using data envelope analysis with principal components: The case of Wisconsin cranberry. *Journal of environmental management, 147*, 175-183.

Färe, R., & Grosskopf, S. (2000). Theory and application of directional distance functions. *Journal of productivity analysis, 13*(2), 93-103.

Färe, R., Grosskopf, S., & Weber, W. L. (2006). Shadow prices and pollution costs in US agriculture. *Ecological economics, 56*(1), 89-103.

Fundação Sistema Estadual de Dados – Seade. *Economia*. http://produtos.seade.gov.br/produtos/anuario

Gomes, E. G. (2008). Uso de modelos DEA em agricultura: revisão da literatura. *Engevista, 10*, 27-51.

Ho, T. Q., Hoang, V. N., Wilson, C., & Nguyen, T. T. (2018). Eco-efficiency analysis of sustainability-certified coffee production in Vietnam. *Journal of cleaner production, 183*, 251-260.

Instituto Brasileiro de Geografia e Estatística - IBGE. (2010). *Censo agropecuário 2006*. Rio de Janeiro: IBGE.

Instituto de Manejo e Certificação Florestal e Agrícola (2015). *Documento de análise: evolução das emissões de gases de efeito estufa no Brasil (1970-2013): setor de agropecuária*. São Paulo: Observatório do Clima.

Lampe, H. W., & Hilgers, D. (2015). Trajectories of efficiency measurement: A bibliometric analysis of DEA and SFA. *European Journal of Operational Research, 240*(1), 1-21.

Maffei, F. (2010). *Diversidade e uso do habitat de comunidades de anfíbios anuros em Lençóis Paulista, Estado de São Paulo*. (Dissertação de mestrado, Universidade Estadual Paulista, Botucatu, SP, Brasil).

Orea, L., & Wall, A. (2017). A Parametric Approach to Estimating Eco-Efficiency. *Journal of agricultural economics, 68*(3), 901-907.

Padrão, G. A., Campos, S. A. C., Lirio, V. S., & Silva, M. L. (2012). Environmental efficiency and opportunity cost of the Forest Code for the Amazon. In *RIO+ 20-International Society for Ecological Economics Conference*. International Society for Ecological Economics Conference, Rio de Janeiro.

Intergovernmental Panel on Climate Change – IPCC. (2006). N_2O emission from managed soils and CO_2 emission from lime and urea application. In: *Guidelines for National Greenhouse Gas Inventories. Agriculture, Forestry and Other Land Use*. Geneva, Switzerland: IPCC. p. 11.1-11.54, vol. 4.

Rosano-Peña, C., & Daher, C. E. (2015). The Impact of Environmental Regulation and Some Strategies for Improving the Eco-Efficiency of Brazilian Agriculture. In *Decision Models in Engineering and Management* (pp. 295-322). Springer, Cham.

Rosano-Peña, C., Serrano, A. L. M., Britto, P. A. P. de, Franco, V. R., Guarnieri, P., & Thomé, K. M. (2018). Environmental preservation costs and eco-efficiency in Amazonian agriculture: Application of hyperbolic distance functions. *Journal of cleaner production, 197*, 699-707.

Rybaczewska-Błażejowska, M., & Gierulski, W. (2018). Eco-efficiency evaluation of agricultural production in the EU-28. *Sustainability, 10*(12), 4544.

Secretaria do Meio Ambiente/Instituto Florestal. (2005). *Inventário Florestal da Vegetação Natural do Estado de São Paulo*. São Paulo: Imprensa oficial.

Tilman, D. (1999). Global environmental impacts of agricultural expansion: the need for sustainable and efficient practices. *Proceedings of the National Academy of Sciences, 96*(11), 5995-6000.

Tobler, W. R. (1979). Smooth pycnophylactic interpolation for geographical regions. *Journal of the American Statistical Association, 74*(367), 519-530.

Valcarce, E. V., & Serrano, R. M. (2000). *La utilidad de la Econometría Espacial en el ámbito de la Ciencia Regional. Documento de Trabajo 2000-13, Mayo* Barcelona: Universidad de Barcelona.

Van Passel, S., Nevens, F., Mathijs, E., & Van Huylenbroeck, G. (2007). Measuring farm sustai¬nability and explaining differences in sustainable efficiency. *Ecological Economics, 62*(1), 149-161.

Van Passel, S., Van Huylenbroeck, G., Lauwers, L., & Mathijs, E. (2009). Sustainable value assess¬ment of farms using frontier efficiency benchmarks. *Journal of Environmental Management, 90*, 3057-3069.

Verfaillie, H. A. & Bidwell, R. (2000). *Measuring eco-efficiency: a guide to reporting company performance*. Geneva, Switzerland: World Business Council for Sustainable Development.

Xing, Z., Wang, J., & Zhang, J. (2018). Expansion of environmental impact assessment for eco-efficiency evaluation of China's economic sectors: An economic input-output based frontier approach. *Science of the Total Environment, 635*, 284-293.

Zhang, B., Bi, J., Fan, Z., Yuan, Z., & Ge, J. (2008). Eco-efficiency analysis of industrial system in China: A data envelopment analysis approach. *Ecological Economics, 68*(1-2), 306-316.