Available online at www.centmapress.org

Int. J. Food System Dynamics 14 (4), 2023, 431-442

DOI: https://dx.doi.org/10.18461/ijfsd.v14i4.H6

INTERNATIONAL JOURNAL ON FOOD SYSTEM DYNAMICS

Sustainability of Soybean Farms Participating in the Agro Plus Program in Minas Gerais State, Brazil: An Application of Cluster and Principal Component Analyzes

Antônio Consentino Teixeira Oliveira[a], Aziz Galvão da Silva Júnior[a], Zhang Min[b]

[a]*Department of Rural Economics, Federal University of Viçosa, Brazil.*
[b]*College of Agriculture, Nanjing Agricultural University, China.*
antonio.consentino@ufv.br; aziz@ufv.br; springzm@njau.edu.cn

Received April 2023, accepted September 2023, available online November 2023

ABSTRACT

Brazil is the world's largest producer and exporter of soy. The 2022/23 harvest reached 154 million tons and the soy complex (soybean and processed products) was responsible for 19 % of Brazilian agribusiness total exports, contributing with US$ 60 billion to the trade balance. The sustainability of agricultural production is a key issue for the European and Chinese markets. The Agro Plus program, former Soja Plus, was set up in the early 2010s by the Brazilian Vegetable Oil Industries and Farmers Associations (ABIOVE) with the objective of improving the sustainability of soy production. The Agro Plus has been implemented in 5.300 farms nationwide using a checklist which comprises around 230 indicators divided into Social and Environmental and Rural Construction major themes. In Minas Gerais State, the program is coordinate by the Federal University of Viçosa (UFV) and the Farmers´ Association. The objective of this study was to identify critical indicators and groups of farms, allowing the discussion and proposition of individual and collective actions. The Cluster and the Principal Component Analysis (PCA) methods were used. Based on the updated version of the checklist applied in 123 farms during the 2021/22 season, three groups of farms (A, B and C) were identified comprising of 18, 77 and 22 farms respectively. PCA analysis was carried out for each major theme. The first three PCs explain 67% of the variance of Socio Environmental themes and 70% of Rural Building themes. Identified critical indicators and the analysis of farms´ groups allowed the proposition of focused capacity building and distribution of information material to specific group of farms. The UFV team shared the results with the Coordination of the Program and actions for the next Agro Plus assessment campaign will be discussed in a National Seminar to be held in early 2023. It would be highly opportune to include data from other states and to discuss the results considering the requirements of specific market, such as the Chinese one.

Keywords*: Multivariate analysis; soybean chain sustainability; brazilian soybean.*

1 Introduction

In the last four harvests, Brazil has surpassed the United States as the world's largest producer of soy. The country accounted for about 30% of world production and reached a harvest of 154 million tons in 2022/23. The soybean production chain plays a key role in the Brazilian economy. In 2022, the soy complex exported the value of US$ 60 billion, equivalent to 19% of total Brazilian agribusiness exports (MDIC, 2023). Around 70% of soybean production are exported to China, which is the most important commercial partner of the Brazilian economy.

The competitiveness of Brazilian soy production is the result of favorable conditions as: availability of climate and land, adequate technologies and supply of modern inputs, existing of an advanced processing agro-industrial sector, innovative financing services and the competence and efforts of rural producers. However, the issue of sustainability is a growing concern of buyers, not only from the European Union, the second largest market for the soybean production chain and main market for Brazilian meal, but also from the Brazilian largest trading partner, China (ABIOVE, 2023; FEFAC, 2021; Silva Júnior, 2016; WWF, 2015; Visser, 2014; Brown-Lima, 2010; MAPA, 2021).

China is the world's largest soybean importer. In 2020, the world's soybean imports reached 167 million tons. China's imports were 103 million tons, equivalent to 62% of world export market. In 2019, China imported 58 million tons of soybeans from Brazil, and Brazil became China's largest source of soybean imports. Imported soybeans will help alleviate the shortage of agricultural resources in China and promote China's food security development. As a leader in world trade, China should take a more active role and participate in the initiative to promote the sustainable production of Brazilian soybeans (Min, 2022).

For the soy chain, the International Trade Center (ITC) Standard Map lists 93 certification systems related to the sustainability of this commodity (ITC, 2023). Certification systems are important instruments, but by definition they require compliance with all mandatory requirements, including the labor and environmental requirements of producing countries. But on the other hand, certification is a process that inevitably excludes producers who do not meet the minimum requirements. Furthermore, most certification does not consider the specific edaphoclimatic conditions and the availability of natural, human and financial resources of the rural properties.

In this context, the Soja Plus Program was created in 2011 by the Brazilian Association of Vegetable Oil Industries (ABIOVE) in partnership with the Soy and Corn Producers Association of Mato Grosso State (APROSOJA-MT). In 2022, the program was extended to others production chains and named as Agro Plus. The Agro Plus objective is to improve the management of rural properties, considering the conditions of each producer. The program is a nationwide, transparent and participatory initiative, aiming at meeting market demands for sustainable products. The program is based on gradual and continuous improvement of the environmental, social and economic aspects of production through evaluation of more than 230 indicators constructed by brazilian law and international requierments. Currently, the program is present in the main production states, i.e Mato Grosso, Mato Grosso do Sul, Bahia, Minas Gerais, Goiás, Maranhão and Rio Grande do Sul, and was implemented in 6.2 thousand farms, totalizing 5.2 million ha and around 18.5 million ton of soybean (ABIOVE, 2023). Besides Farmers Association, several organizations are partners of the initiative including the larger Bank in Brazil, the Banco do Brasil, and international organizations such as Solidaridad, The Nature Conservancy, EU Vegetable Oil and Proteinmeal Industry Association (FEDIOL), The European Feed´s Manufacturers Association (FEFAC) and the China Soybean Industry Association.

The Agro Plus program works through a diagnostic visit to the property, when applicators evaluate the indicators, quantify each of the checklist's themes and guide the necessary adaptations to the farmer. Generally, after months a second visit is made to check the evolution of the farm and help with the continuous improvement indicators. Also, training to farmers and labors, materials such a set of around 50 sign plates, first aid boxes and control sheets are provided to farmers free of charge. In Minas Gerais State, the program has attended more than 400 farms.

The thematic lines of action of the Agro Plus program include:

- Quality of life at work;
- Best production practices;
- Economic viability;
- Product quality;
- Social responsibility and;
- Rural constructions.

The 230 indicators evaluated in diagnostic visits are binary data, the answer to which is "comply" or "does not comply". These indicators are grouped into specific economic, socio-environmental and rural constructions themes. By meeting the indicators that comprise it, the percentage of each theme is measured, which generates a rich quantitative database. The data used in this work were collected in the 21/22 harvest, when the Agro Plus checklist was composed

of 19 themes; 9 environmental issues and 10 rural constructions and carried-out anonymously and confidentially to preserve the farmers situation.

The Agro Plus has been consolidating itself as the largest private assistance program in the Brazil's agribusiness. It operates with the main agricultural product Brazilian for more than a decade and consists of one of the main initiatives for improvement of social and environmental conditions for soybean production in Brazil. As already elucidated, these issues are essential in global markets and are closely linked to product quality. Given the above, it is timely and important to seek to make it more effective with scientific information.

2 Objective

The general objective of the study is to contribute to the improvement of the Agro Plus Program in the state of Minas Gerais, through the identification of groups of farms and critical indicators for the implementation of individual and collective actions. Specifically, the objectives are:

a) Identify the most critical and most appropriate topics on partner farms.

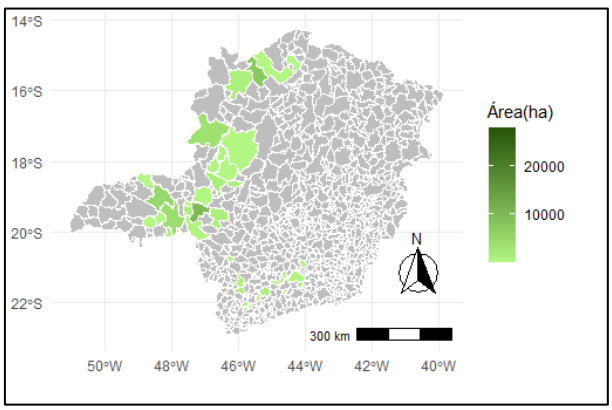
b) Group farms considering their similarities through analysis multivariate.

c) Summarize the information and highlight the most important ones in the database by through principal component analysis (PCA).

d) Propose individual and collective actions to improve partner farms and greater customization of the program in Minas Gerais.

3 Methodology

Data were obtained through the application of the Agro Plus program between 2021 and 2022 in 70 municipalities in the state of Minas Gerais, Southeast of Brazil and the sixth larger soybean producer (CONAB, 2023). The program visited 123 farms in that season and the soybean area (hectars) municipality where are farms visited can be consulted below at the Minas Gerais state maps.



Source: authors

Figure 1. Sum of soybean production area of Agro Plus farms by municipalities evaluated at the season 2021/22.

Mean and standard deviation of themes. Initially, univariate analyses were performed to evaluate each checklist theme individually. Mean and standard deviation were measured for each theme, allowing an initial perspective of the farm's situation in the Minas Gerais State. In this way, it was intended to identify the most critical themes, most adequate and more dispersed between farms to then proceed with multivariate analyses. Furthermore, illustrative maps of Minas Gerais were designed to know the situation of the indicators geographically. The proposal was to identify possible regionalities in the themes complying and to find out which actions, for example, are most necessary in each municipality. The maps were built using the R programming language version 4.1.0.

Cluster Analysis is a multivariate statistical technique that groups objects with based on their similarities (Rossoni et al., 2021). Thus, groups are created with objects similar to each other and different from objects in other groups (Rodrigues, Fachel, and Passuelo, 2012). That analysis by group facilitates understanding the database and can direct decision-making problem under study. Given Agro Plus' extensive database, this analysis is useful to facilitate the evaluation of

farms through groups, allowing them to be categorized and identify the most critical themes in certain groups to plan future actions of the program. It becomes possible to direct the necessary actions without need to evaluate farm by farm, maintaining managerial assistance customized service provided by the program and optimizing resources.

Hierarchical grouping methodology was used, which, according to Rohlf (1970), forms several homogeneous groups that come together in groups each increasingly larger and more heterogeneous. In this method, it is not possible to exchange objects of groups as in non-hierarchical methods and the number of groups is chosen arbitrarily according to the application for the job, although there are ways to measure the most appropriate number of groups. To group objects based on their similarity, it is needed to measure the similarity or dissimilarity between them. The dissimilarity measures make it possible to calculate how different two objects are, which allows, subsequently, group and separate the least and most dissimilar respectively (Crispim, Fernandes, and Aalbuquerque, 2019). Measure adopted was the Euclidean distance, which consists of an extrapolation of the Pythagoras. Each variable used to calculate the distance is considered as a "cateto" of the triangle, so that it is given by:

$$D_{x,y} = \sqrt{\Sigma_k^{\ n} (T_{kx} - T_{ky})^2}$$

where T_{kx} is the percentage reached in theme K by farm X and T_{ky} is the percentage reached in the same theme K by farm Y.

Then, the farms were grouped by the linkage criteria (single, complete, average, centroid and minimum variance) whose dendrogram resulted in the highest cophenetic correlation with the dissimilarity matrix. Finally, ANOVA was carried out to verify the difference of the averages of each cluster. As the Agro Plus database is extensive, this analysis is useful to facilitate the assessment of farms by groups, allowing them to be categorized and identify the most critical issues in certain groups. This helps to optimize the assistance provided by the program and its customization.

Principal Component Analysis (PCA) is an exploratory statistical data tool, which uses multiple observations of *p* variables of *n* objects (Hotelling, 1933; Jollife, 1973). It is a technique widely used to analyze large databases, reducing a given number of correlated variables into new uncorrelated (orthogonal) variables, the principal components (PC). In this way, it is possible to condense the information while preserving the maximum variation of the original data set, which facilitates its description. Briefly, linear combinations of the *n x p* dimension matrix are made, when it is possible to gather the total variance present in just some of the *p* columns, forming the PCs. Thus, the description of the data can be done with few variables, which gather most of the information contained in the *p* original variables.

$$\sum_{j=1}^{p} a_j \mathbf{x}_j = \mathbf{Xa}$$

The simple linear combination is described above. **X** is the *n x p* matrix, $\mathbf{x}_j$ is the observation of j variable, $\mathbf{a}_j$ is a constant of the vector and **Xa** is the new matrix which variance is summed in less vectors. Jolliffe and Cadima (2016) provides a very detailed description of the technique.

Examples of studies that use PCA to address the issue of agricultural sustainability are those by Dong et al. (2015); Dong, Mitchell, and Colquhoun (2015). The PCA is an interesting tool, since measuring sustainability involves several indicators and its isolated interpretation may not be representative. In this sense, the analysis can help Agro Plus in evaluating and understanding the situation of the partner farms. In addition, identifying the most relevant variables in each PC and the most correlated with each other provides important information, which indicates ways to improve the technical assistance provided to rural properties.

Both of multivariate analyzes were carried out in the R programming language. It is worth noting that missing observations impair the quality of clustering and principal component analysis. As some of specific themes are not applicable in some farms, the databases have missing values. Thus, themes that were missing at more than 40 % of farms present in the database were removed from the multivariate analyses as a criterion, which as specified in the results section. Additionally, it was excluded the outliers objects, that have a discrepant variance, because they also harm the multivariate analysis quality.

4 Results and Discussion

4.1 Univariate analysis

Among the 123 farms visited in 2021, the worst results of the Socio-environmental themes (Table 1) were indicators related to: i) Risk Management in the Work Environment (T5); ii) Training and Capacity Building (T6) and; iii) Waste Management And Good Practices (T7). The risk management (T5) theme in the update version of the checklist has more indicators than previous years, which currently better detail the actions of the owners to minimize the risks associated with the functions of the employees. It can be noted that it is one of the topics that most deserves Agro Plus focus to guide and training participants, especially considering the low standard deviation and consequent uniformity between farms. Training and Capacity Building (T6) is a topic of the updated version of the checklist, which was not evaluated at the farms before and includes some indicators that were in other topics before, such as risk management and accident prevention. Maybe for this, the most low mean mean in this. An important information also is the difference between the standard-deviation (sd). Training and Capacity Building has the worst mean and higher sd, that indicates the dispersion around the mean is the biggest It does it means there are farms comply it satisfactorily and farms very less in that. This does not occurs with Waste Management and Good Practices for example, which also has a low mean but sd demonstrates the farms are generally near of that value.

Furthermore, Agrochemicals (T1), Work Contract (T2), Health and Safety (T3), Environmental Regularization (T8) and Financial Control (T9) are those socio-environmental themes that have more suitable indicators. The most Agro Plus partners farmers knows to manage pesticides appropriately, respect the laborers' dignity and the Brazilian environmental legislation.

Table 1.
Mean and standard deviation of the themes evaluated at Agro Plus program (%).

Social and Environmental themes	**Mean**	**Standard Deviation**
Agrochemicals T1	86.99	8.37
Work Contract T2	78.47	17.40
Health and Safety Management T3	78.32	19.18
Work Environment T4	69.76	18.45
Risk Management in the Workplace T5	54.03	21.11
Training and Capacity Building T6	45.99	30.13
Waste Management and Good Practices T7	61.87	14.73
Environmental Regularization T8	80.41	16.23
Financial Control T9	87.61	19.68
Rural Constructions themes		
Accommodation for Employees T10	69.24	30.55
Housing for Employees T11	94.75	12.15
Living Area T12	71.94	38.06
Agrochemical Deposit T13	62.88	27.33
Agrochemical Packaging Deposit T14	43.02	34.88
Silos and Dryers T15	57.32	37.38
Fuel Filling Point T16	65.97	24.22
Maintenance, Washing and Oil Change Area T17	62.53	32.02
Machinery Shed T18	61.39	24.02
PPE Laundry T19	17.66	32.95

Source: authors.

The averages related to the indicators of the Rural Constructions major theme were lower than those of the Socio-Environmental theme. The Agrochemical Packaging Storage (T14) and PPE Laundry (T19) indicators obtained the lowest averages and demonstrate that they have great potential for improvement in the soybean farm in Minas Gerais state. A few farms has an appropriated construction to storage, for until one year, the pesticides empty packaging. Similarly, PPE laundry is a critical point because many farms do not has an exclusive place to wash the PPE as required the Brazilian norm, then they are with 0 %. The them Silos and Dryers (T15), which was recently added to the checklist, also stood out negatively. Some properties do not yet have these constructions, for this reason the dissemination of information and training are important for new constructions to reach the appropriate standard.

There was not identified any predominance of good or bad situation in the Minas Gerais regions. The complying of each theme is specific in each municipality and made possible knows where and what is necessary to improve the soybean sustainability. The maps of the averages of some themes (best and worst) are shown in figures 2 and 3.

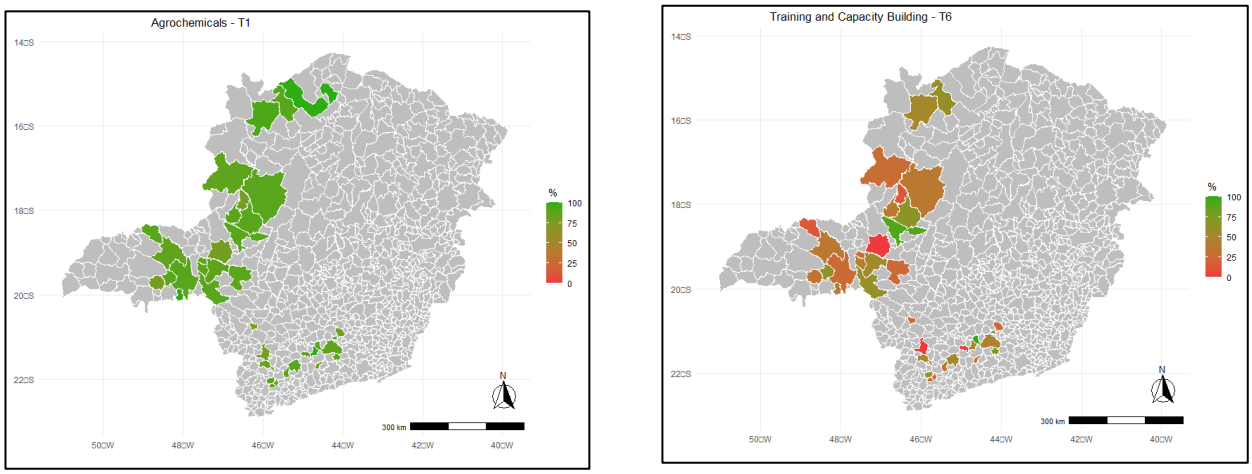


Figure 2. Mean of Socio-environmental themes T1 and T6 by municipalities (Source: authors).

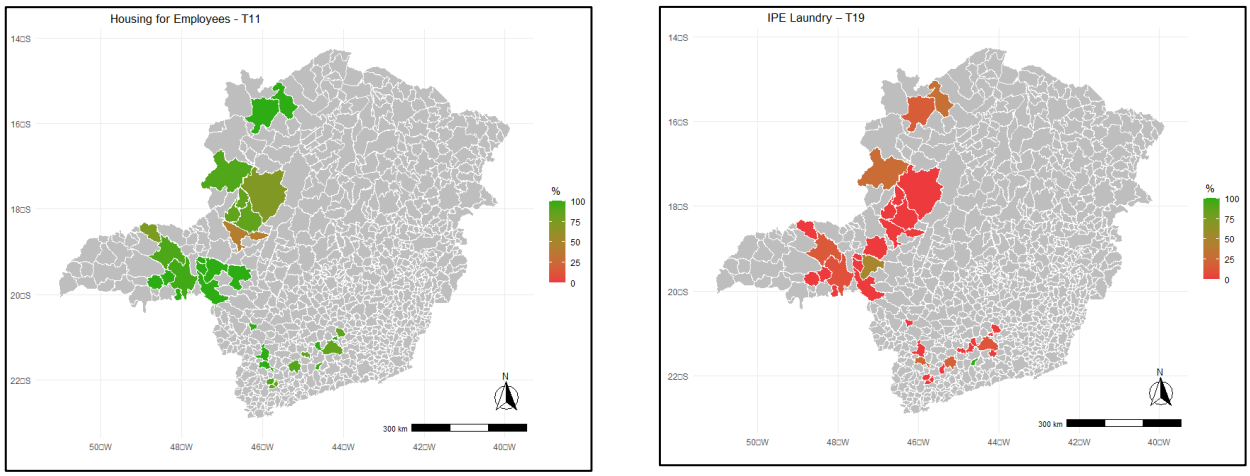


Figure 3. Mean of Rural Construction themes T11 and T19 by municipalities (Source: authors).

4.2 Cluster Analysis

Six (6) farms with discrepant observations were eliminated, leaving 117 for cluster analysis. Also, it excluded the specific themes T10, T15 and T17 due the same reason as the database from the old checklist. The average linkage criteria had the highest cophenetic correlation, however it was 0.618. This reveals a not very good fit of the method in relation to the dissimilarity calculated by the Euclidean distance. With a greater number of farms. the clustering would have better preserved the dissimilarity and the correlation would have been greater. In any case, it is relevant to analyze and interpret the groups, since the cophenetic correlation is not so low and the averages of the themes are significantly different between the groups (p-value<0.05).

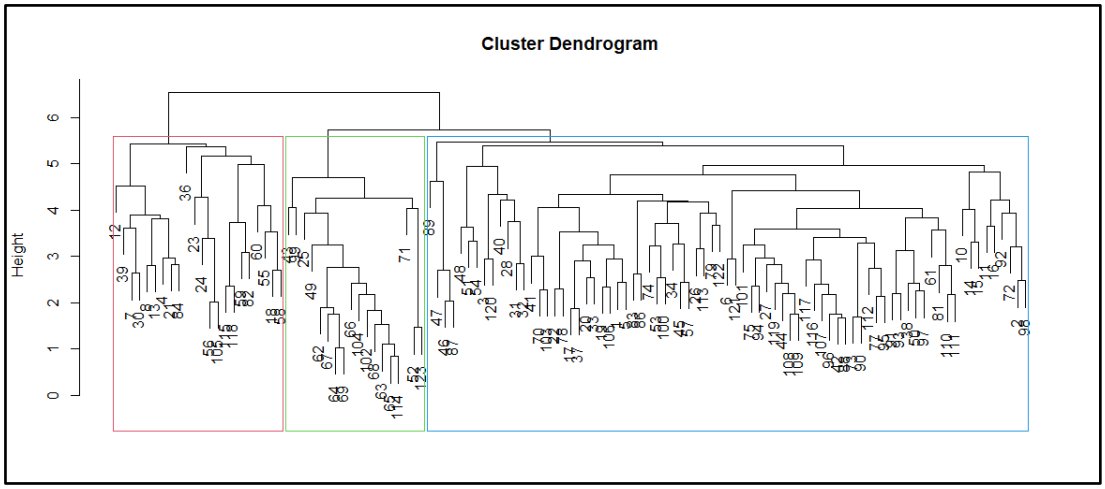


Figure 4. Cluster dendrogram (Source: authors).

The best number of groups to be interpreted. based on the dendrogram. was three and named group A, B and C, encompassing 18, 77 and 22 farms respectively.

Table 2.
Mean (%) by themes.

Groups	T1	T2	T3	T4	T5	T6	T7	T8	T9	T11	T12	T13	T14	T16	T18	T19
A	97	90	97	83	83	87	66	91	93	98	97	93	88	82	84	69
B	85	80	77	71	53	42	64	83	88	96	74	66	43	67	62	10
C	84	58	62	53	30	21	50	65	77	91	34	30	12	50	32	1
p-value*	<0.05															

*ANOVA showed a significant difference between groups for all themes at 5% probability.

Source: authors

Group A had the best averages for all themes. It contains the farms with the highest compliance with the indicators. Only the themes T7 (Waste Management and Good Practices) T19 (PPE laundry) were under 80 %, meaning that improvements are necessary in these indicators in some farms of group A. As was discussed at the 4.1 section, T7 has a low mean and standard-deviation, then even the better farms are below in it.

Most of the farms assisted are at an intermediate management level, constituting group B. They have high averages in the best-scored themes throughout the state, but very low in T5, T6, T14 and T19. Here, it was possible to identify the farms are worst in Training and Capacity Building for example. It was commented that theme has the higher standard-deviation, what reveals larges differences among the farms in then. The analysis was able to separate which farms needs to improve in T6 (groups B and C) and which ones do not needs it (group A). The future actions to group B should focus on risk management, employee training, waste management and adapting the Agrochemical Packaging Storage (T14) and PPE laundry (T19) facilities.

Finally, group C is made up of properties that are less covered by the legislation and that need to improve in many aspects. Even so, it is worth highlighting positively T1, T9 and T11, whose averages are very high in these 22 farms, which shows the potential they have in adapting the other indicators. With managerial assistance and guidance, the producer is able to bring his property fully into compliance.

With the results of this analysis, it can be said that most farms have ample scope to improve their sustainability, especially focusing on the aforementioned points. To optimize and customize the assistance provided by the Agro Plus program, it is important to train: 1) all properties in terms of Waste Management and Good Practices (T7) and IPE Laundry; (T19) 2) those belonging to groups B and C regarding risk management, employee training and packaging deposit and; 3) those in group C, specifically regarding the work environment, living area, agrochemical deposit and machine shed.

4.3 Principal Component Analysis (PCA)

A PCA analysis was performed for each of the major themes of the checklist. Similarly as in cluster analysis, the six farm outliers and T10, T15 and T17 were excluded.

Socio Environmental themes

Nine (9) social and environmental themes were included in the PCA. It was possible to gather 67 % of the variance in the first three PCs, whose eigenvalues are greater than 0.7 (Jolliffe, 1973).

Table 3.
Information of each PC considering Socio-environmental themes.

PC	Eigenvalue	Explained variance (%)	Accumulated explained variance(%)
1	3.9655	44.0612	44.0612
2	1.0884	12.0935	56.1547
3	0.9809	10.8984	67.0531

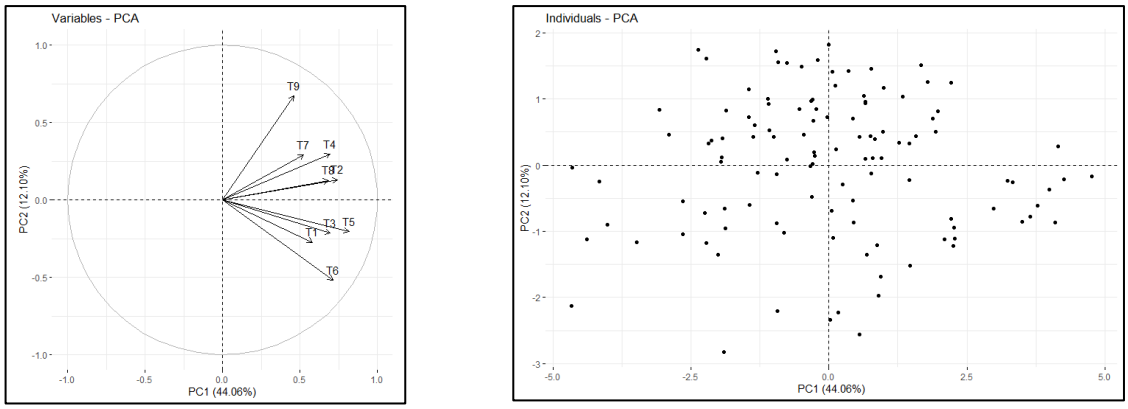
Source: authors.

Table 4.
Significant correlations between themes and PCs at 1% probability (p-value<0.01).

PC	T1	T2	T3	T4	T5	T6	T7	T8	T9
1	0.5790	0.7419	0.6939	0.6926	0.8154	0.7119	0.5208	0.6821	0.4589
2	-0.2742	-	-	0.2955	-	-0.5195	0.2900	-	0.6712
3	0.3246	-	0.2928	-	-	-	-0.6737	-	0.4757

Source: authors

By the socio-environmental major theme, 44% of total variance was explained by PC1. T5, T2, T6, T3 and T4 are the most significant themes. They are related to social aspects, pointing out the strong impact they have on the sustainability of properties. In addition, T5 (*Risk Management in the Workplace*) and T6 (*Training and Capacity Building*) also have the lowest means, 54 and 46% respectively, as pointed out in table 1. This reinforces the importance of risk management indicators and employee training in the evolution of socio-environmental management. The theme with the highest average is the most correlated with PC2. T9 (*Financial Control*) contributes with more than 40% of the variance explained by it and has an average of 87% across farms, demonstrating how it is uncorrelated with others. By PC3, T7 (*Waste Management and Good Practices)* was the only one most relevant to it, and is also little associated with the other themes.



(a) (b)

Figure 5. PC1 x PC2 distribution of vectors (a) and objects (b).
(Socio-environmental major themes; source authors)

It is important to emphasize the distribution of the farms (b), which is rather homogeneous above the axes and, thus, above of the average for the PC1 and PC2 and for the themes that compose them consequently. It shows those farms above the mean are more similar each other than those with lows percentage. It is essential provide a customized assistance for the worst farms in management level primarily. There is no predominance, in number of farms, regarding the management level.

Rural Construction themes

More than 70 % of the total variance was summarized in the 3 PCs, all with eigenvalues greater than 0.7.

Table 5.
Information of each PC considering Rural Construction themes (updated checklist).).

PC	Eigenvalue	Explained variance (%)	Accumulated explained variance (%)
1	2.9739	42.4838	42.4838
2	1.0870	15.5288	58.0125
3	0.8637	12.3389	70.3515

Source: authors.

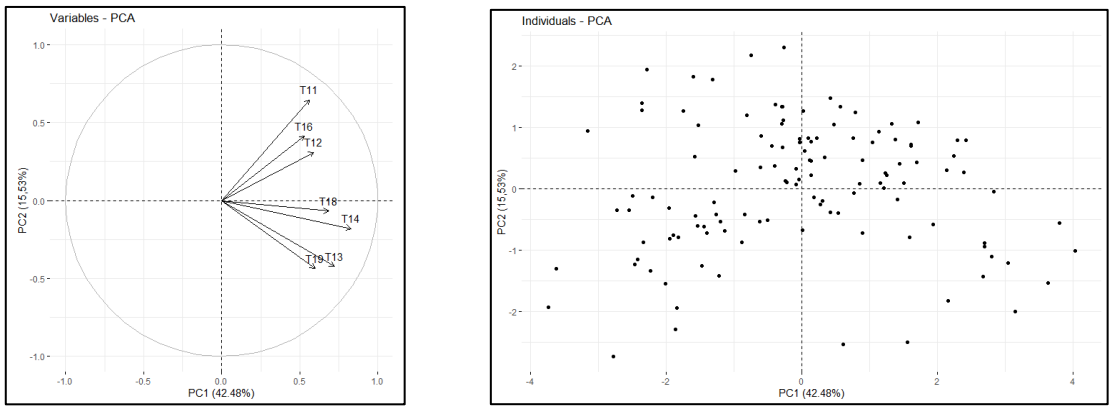
Table 6.
Significant correlations between themes and PCs at 1% probability (p-value<0.01).

PC	T11	T12	T13	T14	T16	T18	T19
1	0.5622	0.5871	0.7216	0.8287	0.5302	0.6838	0.5974
2	0.6447	0.3090	-0.4232	-	0.4113	-	-0.4364
3	-	0.6238	-	-	-0.6083	-	-

Source: authors.

PC1 has variance mainly derived from themes T14, T13, T18 and T19. Together, T13 (*Agrochemical Deposit*) and T14 (*Packaging Deposit*) add up to more than 40 % of its total variance and have a great influence on the total variation of

the database. presenting themselves as crucial constructions in improving sustainability. Especially, if one considers the low averages of the two themes in the 123 farms, of 63 and 43% respectively, as pointed out in table 1. T11 (*Housing for Employees*) is the theme with the highest average among the farms (94%) and weakly correlated with the others. He composed, almost alone, PC2 with more than 38% of contribution to it. Although T19 (*IPE Laundry*) is more associated with PC1, it is worth mentioning its moderate negative correlation with PC2, as it is the most urgent topic to be improved (average of only 18%). As T11 (*Housing for Employees*) and T19 (P*PE Laundry*) are oppositely correlated to PC2, there is a tendency in some farms: despite the high compliance to T11, they also have the smaller value at T19. In practice, this discrepancy signaled by PC2 indicates that it does not matter how adequate housing for employees is; PPE laundry is critical on most farms. It is further evidence of how widespread and targeted T19 compliances have to be even on properties with other well-suited constructions.



(a) (b)

Figure 6. PC1 x PC2 distribution of vectors (a) and objects (b).
(Rural Construction major themes, source: authors)

The distribution of farms in the first quadrant (those with high scores in the PC1 and PC2 themes) is less dispersed than in the others, in the same way of PCA done for the socio-environmental themes. Those that are below average, are in very particular situations and have different themes to be worked on. It is essential to know and analyze the producer's reality, mainly in those with worse indicators, in order to guide him in the most effective way.

5 Conclusions

The results of the analyzes using the selected methodologies allowed discussing and proposing measures to increase the efficiency of the Agro Plus program. The identification of indicators with lower averages in each municipality will make it possible to direct the elaboration of didactic material and the intensification of training for different regions. The themes more critical are Risk Management, Training and Capacity Building, Waste Management and Good Practices, Agrochemical Packaging Storage and PPE Laundry.

There are groups of farms with relevant differences. Even the better farms need improvements in Waste Management and PPE Laundry. Most of farms (66 %) are at an intermediary management level.

The results of the analysis of principal components made it possible to know the correlation between the themes and the situation of most farms easily. There is no correlation between the more suitable theme and those less suitable. It is indispensable to assistant all of soybean farms to advance in the sustainability.

The expansion of the analysis, including data from other Brazilian states, will allow a more robust evaluation of the program and comparisons between the different regions.

Acknowledgments

To ABIOVE for coordinating the Agro Plus program at the national level and for the opportunity to contribute to it through scientific statistical results. Also, to CNPq, the main Brazilian research promotion agency, for granting a scholarship for this study.

References

ABIOVE (2023). Associação Brasileira das Indústrias de Óleos Vegetais. Estatísticas. Available at: <http://abiove.org.br/sobre/> (accessed 04.01.1023).

BRASIL Portaria SEPRT nº 22.677 (2020). NR-31 - Segurança e saúde no trabalho na agricultura, pecuária, silvicultura, exploração florestal e aquicultura. Diário Oficial da República Federativa do Brasil, Brasília, DF, 22 out. 2020. Available at: <www.gov.br/trabalho-e-previdencia/pt-br/normas-regulamentadoras> (access 03.11.2023).

Brown-Lima, C., Cooney, M., Cleary (2010). An overview of the Brazil-China soybean trade and its strategic implications for conservation. The Nature Conservancy: Arlington. Available at: <https://www.readkong.com/page/an-overview-of-the-brazil-china-soy> (accessed 01.10. 1023).

CONAB – Companhia Nacional de Abastecimento (2023). Historical Harvest Series - Soya.. Available at: <//www.conab.gov.br/info-agro/safras/serie-historica-das-safras/itemlist/category/911-soy>. (accessed 14 Aug. 2023).

Dong, F., Mitchell, P.D., Knuteson, D., Wyman, J., Bussan, A.J., Conley, J. (2015). Assessing sustainability and improvements in US Midwestern soybean production systems using a PCA–DEA approach. *Renewable Agriculture and Food Systems*, ***31***(6): 524–539.

Dong. F., Mitchell. P., Colquhoun. J. (2014). Measuring farm sustainability using data envelopment analysis with principal component: The case of Wisconsin Cranberry. *Journal of Environmental Management*, **147**: 175–183.

FEFAC (2021). European Compound Feed Manufacturers' Federation. Soy Sourcing Guidelines 2021. Available at: <fefac.eu/wp-content/uploads/2021/02/FEFAC-Soy> (accessed 01.10.2022).

Hotelling, H. (1933). Analysis of a complex of statistical variables into principal components. *Journal of Educational Psychology*, **24**(6): 417-441.

IBGE (2022). Available at:<www.ibge.gov.br> (accessed 01.10.2023).

Jolliffe, I.T. (1973). Discarding variables in a principal component analysis. II: Real data. *Applied Statistics*, **22**(1): 21-31.

Jolliffe, I.T., Cadima, J. (2016). Principal component analysis: a review and recent developments. *Philos. Trans. R. Soc.*, **374:** 2065.

Kayser, H.F. (1958). The varimax criterion for analytic rotation in factor analysis. *Psychometrika*, **23**(3).

MAPA (2021). Com foco na sustentabilidade. Brasil e China discutem ações para a agricultura em diálogo bilateral. Available at: <https://www.gov.br/agricultura/pt-br/assuntos/noticias/coml> (accessed 04.10.2022).

MIDIC (2023). Available at: <http://comexstat.mdic.gov.br/pt/home> (accessed 01.10.2023).

Min, Z., Silva Junior, A.G., Suya, Y., Aijun. L. (2022). Analysis on the sustainable production of soybean in Brazil. *Research in Food Issues,* **29**: 28-30.

Rohlf, F.J. (1970). Adaptive hierarchical clustering schemes. *Systematic Zoology,* **18**: 58-82.

Rodrigues, A., Fachel, J.M.G., Passuelo, A.C. (2009). Estatística espacial e análise de cluster em dados de desastres naturais: mapeamento das inundações no Rio Grande do Sul entre 2003 e 2009. *Revista Iniciação Científica*, **10**(1): 48-67.

Rossoni, R.A., Moraes, M.L., Catellan, R. (2021). O Perfil da Modernização da Agricultura do Paraná: Uma análise de Cluster. IGepec. Toledo. Edição Especial: *58º Congresso da SOBER*, **25**: 29-45.

Silva Júnior, A.G., Zanasi, C., Souza Jr., W., Ajona, J.V.G. (2016). Matching Brazilian soybean production to the EU sustainability standards' requirements. Compliance of the SojaPlus management program with the FEFAC guidelines. 10th EAAE-IGLS Forum. Innsbruck-Austria.

Visser, C.L.M., Schreuder, R., Stoddard, F. (2014). The EU's dependency on soya bean import for the animal feed industry and potential for EU produced alternatives. *Oilseed and Fat Crops and Seeds*, **21**(4): D407.

WWF (2015). Solving the soy problem. Available at: <http://wwf.panda.org/what_we_do/footprint/agriculture/soy/> (accessed 01.10.2023).