

**UNIVERSIDADE ESTADUAL DE PONTA GROSSA
SETOR DE CIÊNCIAS AGRÁRIAS E TECNOLOGIA
PROGRAMA DE PÓS-GRADUAÇÃO EM AGRONOMIA**

SALATHIEL ANTUNES TEIXEIRA

**CARACTERIZAÇÃO E DETERMINAÇÃO DA EFICIÊNCIA EM SISTEMA AGRÍCOLA
DE CULTIVO DE SOJA POR MEIO DE ANÁLISE ENVOLTÓRIA DE DADOS E
MÉTODOS MULTIVARIADOS.**

PONTA GROSSA - PR

2020

SALATHIEL ANTUNES TEIXEIRA

CARACTERIZAÇÃO E DETERMINAÇÃO DA EFICIÊNCIA EM SISTEMA AGRÍCOLA
DE CULTIVO DE SOJA POR MEIO DE ANÁLISE ENVOLTÓRIA DE DADOS E
MÉTODOS MULTIVARIADOS.

Dissertação de Mestrado apresentada como requisito para obtenção do título de mestre em Agronomia do Programa de Pós Graduação em Agronomia da Universidade Estadual de Ponta Grossa.

Orientador: Prof. Dr. Carlos Tadeu dos Santos Dias.

PONTA GROSSA - PR

2020

T266 Teixeira, Salathiel Antunes
Caracterização e determinação da eficiência em sistema agrícola de cultivo da soja por meio de análise envoltória de dados e métodos multivariados / Salathiel Antunes Teixeira. Ponta Grossa, 2020.
84 f.

Dissertação (Mestrado em Agronomia - Área de Concentração: Agricultura), Universidade Estadual de Ponta Grossa.

Orientador: Prof. Dr. Carlos Tadeu dos Santos Dias.

1. Energia na agricultura. 2. Balanço de energia. 3. Eficiência energética. 4. Análises multivariadas. 5. Análise envoltória. I. Dias, Carlos Tadeu dos Santos. II. Universidade Estadual de Ponta Grossa. Agricultura. III.T.

CDD: 630

**PRÓ-REITORIA DE PESQUISA E PÓS-GRADUAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM AGRONOMIA**

CERTIFICADO DE APROVAÇÃO

Título da Dissertação: **Caracterização e determinação da eficiência em sistema agrícola de cultivo de soja por meio de análise envoltória de dados e métodos multivariados.**

Nome: **Salathiel Antunes Teixeira**

Orientador: **Carlos Tadeu dos Santos Dias**

Aprovado pela Comissão Examinadora:

Prof. Dr. Carlos Tadeu dos Santos Dias

Prof. Dr. Fabricio Tondello Barbosa

Prof. Dr. Luis Miguel Schiebelbein

Data da Realização: 28 de fevereiro de 2020.

*Ainda que ele não entenda os símbolos,
as equações, os gráficos e as ideias, mas
para que compreenda que a educação
dignifica e engrandece o ser humano, ao
meu filho deixo como exemplo e dedico.*

AGRADECIMENTOS

Ao meu filho Henrique que, com a ingenuidade característica das crianças, não entendia as aflições que me atormentavam mas ainda assim me oferecia abraços. A minha esposa Luciane que me suportou e apoiou durante o período em que estive dedicado a realizar meu Mestrado.

Aos meus pais, Adão e Elizabete, que proporcionaram condições para que eu permanecesse em Ponta Grossa durante a realização da pós-graduação e sempre estiveram ao meu lado dando apoio moral.

Ao Professor Dr. Carlos Tadeu dos Santos Dias que me orientou, ajudou e ensinou pacientemente durante essa jornada, e de quebra me proporcionou a possibilidade de levar o exemplo de um profissional digno do seu currículo e com a humildade e modéstia compatível com envergadura intelectual de um verdadeiro *gentleman*.

A Professora Dra. Neyde Fabíola Balarezo Giarola e a Camila Moreira que sempre estiveram à disposição para dirimir eventuais dúvidas com relação ao regimento bem como dando suporte nos trâmites do curso de pós-graduação.

Aos Professores Dr. Fabrício Tondello Barbosa, Dr. Sérgio da Costa Saab e Dra. Carolina Weigert Galvão que contribuíram com o andamento desta dissertação ao avaliar o pré-projeto apresentado como requisito parcial para aprovação na disciplina de Seminários II.

Ao Professor Dr. Luis Miguel Schiebelbein que gentilmente cedeu espaço nas suas aulas e me acompanhou no estágio de docência realizando sábias observações.

Ao Me. Santos Henrique Brant Dias cujo o apoio foi determinante para que eu tivesse condições e realizasse este curso de pós-graduação.

A colega de pós-graduação Daiane Hyeda que na minha ausência fez a gentileza de encaminhar as cópias desta dissertação para a apreciação da coordenação e componentes da banca de avaliação.

Universidade Estadual de Ponta Grossa e aos profissionais que nela exercem suas funções proporcionando ensino, infraestrutura e ambiente favoráveis à realização de trabalhos como este.

Aos produtores rurais da cidade de Boa Esperança-PR, especialmente aos do distrito de Paranaguçu, que confiaram informações sobre suas propriedades permitindo a realização deste trabalho.

À CAPES - Coordenação de Aperfeiçoamento de Pessoal de Nível Superior pela bolsa de estudos concedida.

RESUMO

O desenvolvimento de técnicas de análise e gestão é importante. Mesmo que o alvo deste trabalho esteja centrado em uma análise energética de um sistema de produção as técnicas aqui exploradas podem ser aplicadas em outras circunstâncias. O objetivo geral do trabalho foi caracterizar e comparar a eficiência de um grupo de produtores da região de Boa Esperança, estado do Paraná, no sistema de produção da cultura da soja utilizando a análise envoltória de dados e métodos multivariados sob a óptica da energia empregada no sistema. Os dados referentes a quantidades de energia utilizados e produzidos no sistemas foram convertidos em equivalentes de energia por unidade de área para realização do balanço de energia. Em seguida foi realizada revisão bibliográfica sobre a análise envoltória de dados (DEA – *data envelopment analysis*). Posteriormente, para obter valores de eficiência, foi utilizada a DEA no conjunto de dados obtidos junto aos 27 produtores nas safras 2014/2015, 2015/2016 e 2016/2017. A análise multivariada de variância, regressão logística discriminante e análise de agrupamento no plano das componentes principais foram utilizadas para verificar diferenças entre classes de eficiência e padrão de agrupamentos considerando as diferentes variáveis. Nas análises multivariadas as variáveis foram: 1.produtividade, 2.pesticidas (fungicidas, herbicidas, inseticidas, adjuvantes e espalhantes), 3.nutrientes (N, K₂O, P₂O₅ e micronutrientes), 4.sementes e 5.operacionais (mão-de-obra, máquinas e combustível). O balanço de energia na produção foi positivo considerando a média dos 27 produtores em três safras. O total de energia empregada na produção de soja foi de 40.888,98 MJ ha⁻¹. Foi constatado um saldo de 53.634,69 MJ ha⁻¹ debitando da produtividade equivalente em energia os gastos equivalentes em energia. A razão entre os totais de saída e entradas foi de 2,31. Verificou-se em uma abordagem matemática que é possível obter valores de eficiência construindo valores de eficiência relativa entre elementos de um grupo de unidades formadoras de decisão (DMU's – *decision making units*). Utilizando DEA com orientação para entradas nos 27 produtores de soja nas safras 2014/2015, 2015/2016 e 2016/2017 o máximo valor de eficiência foi de 100 % e o menor valor foi de 71,37 %. Foram definidas duas classes: “eficientes” e “ineficientes”. Os vetores de médias associados as classes “eficientes” e “ineficientes” apresentam diferença significativa pelo teste Lambda de Wilks ao nível de 5% segundo a análise multivariada de variância. Intervalos simultâneos com pelo menos 95% de confiança estimados pelo método de Bonferroni para a diferença entre os vetores de médias indicam que as médias das variáveis “operacionais” e “produtividade” são diferentes. A regressão logística estimada foi capaz de realizar a distinção entre as classes “eficientes” e “ineficientes” em mais de 75% do total dos casos avaliados. Na análise de agrupamentos por componentes principais foram identificados 3 agrupamentos considerando a média dos agrupamentos. O agrupamento 1 é caracterizado pelo equilíbrio entre custo e produtividade de energia por unidade de área. O agrupamento 2 se caracteriza em média por valores de produtividades em equivalentes em energia baixos e altos dispêndios de equivalentes em energia com pesticidas e operacionais. O agrupamento 3 possui DMU's eficientes e ineficientes. As DMU's deste grupo possuem altas (baixas) produtividades associadas a altos (baixos) dispêndios de energia.

Palavras-chave: Energia na agricultura; Balanço de energia; Eficiência energética; Análises multivariadas; Análise envoltória; Soja.

ABSTRACT

The development of analysis and management techniques is important. Even if the aim of this work is centered on an energy analysis of a production system, the techniques explored here can be applied in other circumstances. The general objective of the work was to characterize and compare the efficiency of a group of farmers in the region of Boa Esperança, state of Paraná, in the soybean production system using data envelopment analysis and multivariate methods from the perspective of the energy used in the system. The data referring to the amounts of energy used and produced in the systems were converted into energy equivalents per unit of area to perform the energy balance. Then, a bibliographic review on data envelopment analysis (DEA) was carried out. Subsequently, to obtain efficiency values, DEA was used in the set of data obtained from the 27 producers in the 2014/2015, 2015/2016 and 2016/2017 seasons. Multivariate analysis of variance, discriminant logistic regression and clustering analysis in the plane of the principal components were used to verify differences between efficiency classes and clustering pattern considering the different variables. In multivariate analyzes the variables were: 1.productivity, 2.pesticides (fungicides, herbicides, insecticides, adjuvants and spreaders oils), 3.nutrients (N, K₂O, P₂O₅ and micronutrients), 4.seeds and 5.operational (labor, machinery and fuel). The energy balance in production was positive considering the average of 27 producers in three seasons. The total energy used in the production of soybeans was 40,888.98 MJ ha⁻¹. A balance of 53,634.69 MJ ha⁻¹ was found, debiting equivalent energy expenditures from energy equivalent productivity. The ratio between the totals of output and inputs was 2.31. It was found in a mathematical approach that it is possible to obtain efficiency values by building relative efficiency values between elements of a group of decision making units (DMU's). Using DEA with input orientation in the 27 soybean producers in the 2014/2015, 2015/2016 and 2016/2017 harvests, the maximum efficiency value was 100% and the lowest value was 71.37%. Two classes were defined: "efficient" and "inefficient". The mean vectors associated with the "efficient" and "inefficient" classes show a significant difference by the Wilks Lambda test at the level of 5% according to the multivariate analysis of variance. Simultaneous intervals with at least 95% confidence estimated by the Bonferroni method for the difference between the mean vectors indicate that the means of the "operational" and "productivity" variables are different. The estimated logistic regression was able to distinguish between the "efficient" and "inefficient" classes in more than 75% of the total cases evaluated. In the clustering analysis by principal components, three clusters were identified considering the average of clusters. The balance between energy cost and productivity per unit area characterizes cluster 1. Cluster 2 is characterized on average by productivity values in low energy equivalents and high expenditure of energy equivalents with pesticides and operational. Cluster 3 has efficient and inefficient DMU's. The DMU's in this group have high (low) productivity associated with high (low) energy expenditures.

Keywords: Energy in agriculture; Energy balance; Energy efficiency; Multivariate analysis; Envelopment analysis; Soybean.

LISTA DE FIGURAS

- Figura 1.2.1 – Análise sobre a utilização de pesticidas para um grupo de 27 produtores de soja durante três safras de cultivo em Boa Esperança-PR. Valores em percentual da quantidade total de ingredientes ativos. Esquerda classes de pesticidas e à direita classe herbicida por mecanismo de ação..... 25
- Figura 1.3.1 – Porcentagem da quantidade de energia empregada em relação ao total de entradas dos grupos indicados na horizontal (a) e seus sub itens e em relação ao total de energia dos quatro grupos (b) exceto a água..... 27
- Figura 4.3.1 – Resíduos multivariados (a) segundo a distribuição de qui-quadrado (χ^2) versus distância de Mahalanobis e distribuição de resíduos univariado (b) segundo os quantis amostrais e teóricos..... 53
- Figura 4.3.2 – Intervalos simultâneos com pelo menos 95% de confiança estimados pelo método de Bonferroni para a diferença entre os vetores de médias das classes eficiente e ineficiente nas variáveis estudadas..... 54
- Figura 4.3.3 – Valores de probabilidades estimados pela função logística para as amostras de 28 DMU's eficientes (representadas pela cor azul e letras E) e das 32 DMU's ineficientes (representadas em vermelho e letras I). Barras azul e vermelha sólidas representam a probabilidade média das classes..... 56
- Figura 4.3.4 – Biplot com as duas componentes principais que representam 69,7% da variabilidade dos dados..... 57
- Figura 4.3.5 – Dendrograma de agrupamento das DMU's avaliadas pelo método da média..... 58
- Figura 4.3.6 – Screeplot da soma de quadrados interna em função do número de grupos..... 59
- Figura 4.3.7 – Projeção dos agrupamentos 1, 2 e 3 no plano das componentes principais. À esquerda os planos com as médias (pontos maiores) no centro das elipses com 0,95 de confiança da média e à direita projeção das variáveis..... 60

LISTA DE TABELAS

Tabela 1.2.1 - Fatores de conversão de equivalentes em energia por unidade de massa para produção de máquinas e equipamentos por diferentes autores.....	17
Tabela 1.2.2 – Fatores de tempo utilizados para estimativa de equivalentes em energia por unidade de área (EEUA) para máquinas e equipamentos utilizados em sistemas de produção agrícola na região de Boa Esperança-PR.....	18
Tabela 1.2.3 – Equivalentes de energia por unidade de massa ($MJ\ kg^{-1}$) de nitrogênio (N), fósforo (P_2O_5), potássio (K_2O) e micronutrientes diluídos segundo diferentes autores.....	21
Tabela 1.2.4 – Equivalentes de energia por unidade de massa ($MJ\ kg^{-1}$) de ingrediente ativo (i. a.) de diferentes classes pesticidas empregadas na produção de soja.....	24
Tabela 1.2.5 – Determinação do fator de conversão em equivalente em energia por unidade de massa ($MJ\ kg^{-1}$) dos diferentes mecanismos de ação da classe herbicida de acordo com a participação dos mecanismos no grupo amostral.....	25
Tabela 1.3.1 – Balanço de equivalentes em energia para a produção de soja.....	26
Tabela 3.3.1 – Valores de eficiência técnica (CRS) orientado para entradas calculados nas três safras (S1, S2 e S3), estimados para as janelas (J-1 e J-2) e a Eficiência.....	46
Tabela 4.3.1 – Resumo da análise de variância multivariada.....	53
Tabela 4.3.2 – Resultados da função logística discriminante.....	55

SUMÁRIO

INTRODUÇÃO.....	11
OBJETIVOS.....	14
OBJETIVO GERAL.....	14
OBJETIVOS ESPECÍFICOS.....	14
CAPÍTULO 1 – AVALIAÇÃO DO SISTEMA DE PRODUÇÃO DA SOJA EM EQUIVALENTES DE ENERGIA E O BALANÇO ENERGÉTICO.....	15
1.1 INSUMOS E PRODUTOS AGRÍCOLAS EM EQUIVALENTES DE ENERGIA	15
1.2 VALORES DE CONVERSÃO PARA EQUIVALENTES EM ENERGIA.....	16
1.2.1 Máquinas e implementos agrícolas e seus equivalentes em energia.....	16
1.2.2 Mão-de-obra na agricultura e equivalentes em energia.....	18
1.2.3 Equivalentes em energia para combustível.....	19
1.2.4 Equivalentes em energia para macronutrientes e micronutrientes.....	20
1.2.5 Sementes e grãos produzidos e seus equivalentes em energia.....	22
1.2.6 Energia química da água.....	22
1.2.7 Equivalentes em energia para agrotóxicos utilizados na agricultura.....	23
1.2.8 Balanço de energia.....	25
1.3 RESULTADOS E DISCUSSÃO.....	26
1.4 CONCLUSÕES.....	29
CAPÍTULO 2 – ESTADO DA ARTE: ASPECTOS ELEMENTARES SOBRE ANÁLISE ENVOLTÓRIA DE DADOS.....	30
2.1 REVISÃO DE LITERATURA.....	30
2.1.1 Análise Envoltória de Dados – <i>Data Envelopment Analysis</i> (DEA).....	30
2.1.2 Aplicações de análise envoltória de dados.....	31
2.2 ANÁLISE ENVOLTÓRIA DE DADOS (DEA) – ASPECTOS MATEMÁTICOS	33
2.3 A MODELAGEM DO PROBLEMA PELA PROGRAMAÇÃO LINEAR.....	34
2.4 O CONJUNTO <i>P</i> DE POSSIBILIDADES DE PRODUÇÃO.....	35
2.5 O CARÁTER ENVOLTÓRIO (<i>ENVELOPMENT</i>) DA ANÁLISE DE DADOS...	36

2.6	O AJUSTE DOS FATORES DE PRODUÇÃO.....	40
2.7	A ANÁLISE COM ORIENTAÇÃO PARA SAÍDAS (<i>OUTPUT ORIENTED</i>)....	41
2.8	ANÁLISE ENVOLTÓRIA COM RETORNOS VARIÁVEIS POR ESCALA.....	42
CAPÍTULO 3 – UTILIZAÇÃO DE DEA NA DETERMINAÇÃO DE EFICIÊNCIA DE PRODUÇÃO DE SOJA CONSIDERANDO EQUIVALENTES EM ENERGIA.....		44
3.1	MOTIVAÇÃO.....	44
3.2	METODOLOGIA.....	44
3.3	RESULTADOS E DISCUSSÃO.....	45
3.4	CONCLUSÕES.....	47
CAPÍTULO 4 – ESTUDOS MULTIVARIADOS NA CARACTERIZAÇÃO E AVALIAÇÃO DE CLASSES DE EFICIÊNCIA OBTIDAS POR MEIO DE DEA.....		48
4.1	MOTIVAÇÃO.....	48
4.2	METODOLOGIA.....	48
4.2.1	Análise de variância multivariada.....	49
4.2.2	Regressão logística discriminante.....	50
4.2.3	Análise de agrupamento por componentes principais.....	51
4.3	RESULTADOS E DISCUSSÃO.....	52
4.3.1	Análise de variância multivariada e regressão logística.....	52
4.3.2	Análise de agrupamento por componentes principais.....	56
4.4	CONCLUSÕES.....	62
CONSIDERAÇÕES FINAIS.....		63
REFERÊNCIAS.....		64
APÊNDICE A – Demonstrativo de programação e cálculos no R Software.		72

INTRODUÇÃO

É comum verificar na literatura registros a respeito da importância da agricultura na atualidade. Campanhas publicitárias veiculadas no horário nobre em grandes veículos de mídia como a televisão, jornais e revistas não são raras. Miranda (2013) afirma que “nunca a população mundial foi tão grande, nem viveu tanto. Nunca se produziram nem se consumiram tantos alimentos”.

Contudo, a tecnificação da agricultura de produção em escala pode ser verificada em contextos históricos em outros países. Exemplo é a descrição das fazendas “*bonanza*” no então extremo noroeste dos Estados Unidos da América (MORRIS, 2006, p. 117). Essas recebiam a denominação de *bonanza* em função dos lucros elevados que proporcionavam. Eram verdadeiras “fazendas fábricas” que tinham um sistema de produção organizado e gerenciado como empresas visando constante evolução a fim de melhorar eficiência de processos reduzindo custos e maximizando o lucro.

Os preços de mercadorias agrícolas são altamente suscetíveis a alterações por vários fatores inerentes ao processo produtivo. No caso de *commodities* agrícolas os preços apresentam sazonalidade bem como volatilidade. A alteração de preços está associada a fatores como clima, armazenamento, produção da safra, tamanho do período entre safras, questões alfandegárias entre países e evolução tecnológica (RIBEIRO; SOSNOSKI; OLIVEIRA, 2010).

Com um comércio global de *commodities* agrícolas, o produtor rural vê reduzida a possibilidade de interferir nos preços de seu produto ficando condicionado aos ditames de preço regidos pela oferta e demanda do produto no mercado e a sazonalidade dos preços. Para Lunkes (2003) o preço exerce grande influência nas decisões de compra dos consumidores.

Empresas de segmentos diferenciados têm grandes cuidados na determinação de preços de seus produtos e, em geral, baseiam-se na apuração de seus custos de produção. Dessa forma é importante que o produtor rural tenha um controle de custos das suas atividades. Tal matéria é abordada e tema de estudo há algum tempo (SCHUH, 1976).

A evolução digital proporciona ferramentas de gestão que visam gerar informação e fomentar a tomada de decisão do gestor rural (MARTIN et al., 1998) racionalizando dessa maneira o processo produtivo e aumentando a competitividade nos negócios agropecuários. Outro ponto fundamental da informatização e melhoria de processos de controle é a capacidade de obter informações precisas a respeito do custeio de produção. O custeio pode ser definido

como a apuração de custos dos bens, mercadorias e serviços de entidades públicas e privadas, determinada por diferentes metodologias, para atribuir o custo do produto (CARARETO; JAYME; TAVARES, 2006).

Valor é o que os compradores se dispõem a pagar por um produto. Valor superior provém de: (1) benefícios únicos que compensam preços superiores ou de (2) oferta de preços inferiores aos dos concorrentes para benefícios equivalentes. Tais constatações levam as duas estratégias de vantagem competitiva básicas: diferenciação e liderança de custos (BRITO; BRITO, 2012; PETERAF; BARNEY, 2003).

Como afirmado anteriormente o produtor rural dificilmente terá empuxo para alterar significativamente preços no mercado de *commodities*. Dessa forma, a estratégia de vantagem competitiva que lhe resta é diferenciar-se com relação a custos de produção. Portanto, toda técnica que permita avaliar os índices de eficiência e/ou análise de eficiência da produção deve ser considerada (NACHILUK, 2018) para que agregue na tomada de decisão e avaliação do perfil tecnológico utilizado.

Assim, é fundamental o desenvolvimento de técnicas robustas de análise ainda que a adoção esteja condicionada a perfis de gestão mais elevados ou que a aplicabilidade imediata não seja viável. Esse é o propósito de desenvolver a aplicação da análise envoltória de dados em sistemas agrícolas. Mesmo que o alvo principal deste trabalho esteja centrado em uma análise energética de um sistema de produção a técnica pode ser aplicada em outras análises de cunho econômico como será relatado na revisão.

Atualmente outros pontos, além da óbvia maximização de lucros, norteiam os sistemas de gerenciamento de negócios. Preocupações com indicadores ambientais (CAMPOS; MELO, 2008) que estejam harmonizados com demandas sociais e econômicas (RODRIGUES, 2010) são aspectos fundamentais a serem considerados no processo de gestão. Empresas do agronegócio estão sujeitas a essas demandas socioambientais.

Outro aspecto que ganha força no contexto de racionalização de utilização de recursos é a eficiência energética de sistemas. A energia é considerada uma importante base material do crescimento econômico e social de uma região (MOUSAVI-AVVAL et al., 2011). Estudos de âmbito nacional (COSTA et al., 2017) e internacional (CHEN et al., 2006) abordam aspectos importantes sobre a utilização de energia a fim de estabelecer as bases de uma economia racional e ecologicamente correta.

A produção agropecuária tem importante papel na produção e utilização de energia. A produção de energia por meio de resíduos de cultura (COSTA et al., 2017) ou biomassa e

resíduo florestal (DIAS JÚNIOR et al., 2017) vêm se tornando alternativa para muitas empresas. Do ponto de vista de consumo, trabalhos para avaliar a eficiência energética dos sistemas agrícolas consideram os insumos utilizados do ponto de vista energético (MOUSAVI-AVVAL et al., 2011, 2013; OMMANI, 2011). Assim, pode-se ter um balanço de energia considerando entradas e saídas de energia de um sistema agrícola.

Neste contexto, buscar-se-á avaliar a eficiência de produtores agrícolas. As entradas e saídas no sistema serão comparadas a fim de avaliar a eficiência. Para isso, sugere-se a utilização da análise envoltória de dados e técnicas de análise multivariada a fim de estabelecer uma ordem e/ou grupos de eficiência em uma amostra de produtores de soja do município de Boa Esperança-PR. A hipótese de nulidade considera que não existe diferença significativa entre os produtores nas safras avaliadas.

OBJETIVOS

OBJETIVO GERAL

Caracterizar e comparar a eficiência de um grupo de produtores da região de Boa Esperança-PR no sistema de produção da cultura da soja utilizando a análise envoltória de dados e métodos multivariados sob a óptica da energia empregada no sistema.

OBJETIVOS ESPECÍFICOS

- i. Determinar o balanço energético na cultura da soja comparando os índices obtidos com os valores referenciados em bibliografia.
- ii. Descrever em detalhes os dois principais métodos de análise envoltória de dados (CRS – *constant return to scale* e VRS – *variable return to scale*).
- iii. Determinar numericamente por meio da análise envoltória de dados (DEA) valores de eficiência para cada elemento do grupo de produtores avaliados.
- iv. Determinar se existe diferença significativa entre os índices de eficiência atribuídos aos elementos do grupo (ou dos subgrupos e classes de eficiência) os caracterizando por meio de métodos de análise multivariada.

CAPÍTULO 1 – AVALIAÇÃO DO SISTEMA DE PRODUÇÃO DA SOJA EM EQUIVALENTES DE ENERGIA E O BALANÇO ENERGÉTICO

1.1 INSUMOS E PRODUTOS AGRÍCOLAS EM EQUIVALENTES DE ENERGIA

Atualmente a preocupação com a escassez de recursos naturais e suas consequências são tema de diversas reflexões e estudos. Estes têm a intenção de repensar o modelo de desenvolvimento alicerçado na exploração de combustíveis fósseis que se consolidou após a revolução industrial.

A indústria, comércio e transporte global têm evidente dependência de insumos e combustíveis derivados de petróleo. As implicações do consumo crescente são abordadas de um ponto de vista da sustentabilidade do sistema bem como da necessidade de racionalização para não causar prejuízos irreversíveis ao ambiente (ROCKSTRÖM et al., 2009).

A atividade agrícola está altamente associada ao desenvolvimento das civilizações. Porém o desenvolvimento das atividades agrícolas em civilizações como egípcia e na região mesopotâmica está associada à utilização da água como insumo no suporte a produção. Egípcios, romanos (HANSEN, 1983) e sírios (JOMEHPOUR, 2009; LIGHTFOOT, 1996, 2000) tiveram destacada e reconhecida habilidade em utilizar a água para atividades agrícolas e outras demandas urbanas. Nota-se dessa forma a importante relação entre a atividade agrícola e o consumo de recursos com o objetivo de produção.

A agricultura moderna têm como referencial a revolução verde (MATSON, 1998). Nesse modelo de exploração agrícola há uma intensa utilização de mecanização na operacionalização dos cultivos. A condução das áreas de produção é feita utilizando insumos industrializados tais como fertilizantes químicos, pesticidas e sementes híbridas e/ou melhoradas, etc. Uma evolução significativa neste modelo foi a introdução de sementes geneticamente modificadas popularizadas no início do novo milênio.

A agricultura cada vez mais é alvo de observações dos mais diversos setores da sociedade com relação as suas práticas e possíveis impactos. Ações de compensação estão sendo utilizadas para retribuir quem utiliza técnicas de manejo que impactam positivamente sobre os recursos. Em inglês o termo *payments of environmental services (PES)* é utilizado. No Brasil o termo comum é pagamento por serviços ambientais (PSA). Exemplo de programa de PSA é o *Vittel PES* (SATTTLER; MATZDORF, 2013), na França, em que a companhia *Nestlé*

Waters fez um acordo com produtores rurais para prevenir a contaminação de aquíferos por nitrato devido à intensificação da agricultura.

Neste contexto de conscientização e racionalização o consumo de recursos nos sistemas de produção agrícola têm sido estudado do ponto de vista energético (GUZMÁN; GONZÁLEZ DE MOLINA, 2015; MONDANI et al., 2017). Um exemplo é a “*Emergy analysis*” (EURICH, 2011, 2016; ORTEGA et al., 2005) onde as entradas e saídas do sistemas são convertidas em equivalentes de energia. Ainda, outros exemplos da literatura (MOUSAVI-AVVAL et al., 2011, 2013; OMMANI, 2011) demonstram que para se analisar o balanço de energia entre insumos e produtos é necessário possuir bons referenciais de valores para as conversões.

O objetivo é apresentar valores de conversões em equivalentes de energia das várias classes de produtos e atividades utilizados como fatores de produção e o resultado do balanço de energia considerando as entradas e saídas no cultivo da cultura da soja.

1.2 VALORES DE CONVERSÃO PARA EQUIVALENTES EM ENERGIA

1.2.1 Máquinas e implementos agrícolas e seus equivalentes em energia

As máquinas e equipamentos agrícolas têm associados a si um valor agregado que é tema de diversas considerações como depreciação, vida útil, etc. Esses pontos são conhecidos e compreendidos pelos produtores rurais em maior ou menor escala de precisão a depender do nível de instrução do usuário destes equipamentos.

O que não está consolidado perante a comunidade como um todo, e não apenas ao grupo de produtores e usuários de equipamentos agrícolas, é que há uma quantidade de energia associada a um determinado equipamento. Essa preocupação com a quantificação da energia associada aos equipamentos agrícolas pode ser verificada em clássicos da literatura a respeito do tema (PIMENTEL, 1992; STOUT, 1989).

É possível verificar em vários relatos da literatura atual (GHORBANI et al., 2011; KAZEMI et al., 2015; MONDANI et al., 2017; NASSIRI; SINGH, 2009; SHARIFI, 2018; TABATABAEEFAR et al., 2009) considerações a respeito de equivalentes de energia para máquinas e equipamentos utilizados na operacionalização de sistemas agrícolas. Contudo, nestes casos valores genéricos e que não fazem distinção entre equipamentos utilizados em

distintas operações criam um caráter minimalista que pode ser depreciativo de acordo com a profundidade pretendida na análise.

Exemplos de trabalhos com uma distinção entre os diferentes equipamentos podem ser encontrados (MOUSAVI-AVVAL et al., 2011) e tornam-se mais consistentes para o caso em que diferentes máquinas e equipamentos são utilizados nas operações. Dessa forma, a partir dos valores da Tabela 1.2.1 serão determinados os equivalentes em energia para a amostra deste trabalho.

Tabela 1.2.1 - Fatores de conversão de equivalentes em energia por unidade de massa para produção de máquinas e equipamentos por diferentes autores

Máquina	$MJ\ kg^{-1}$
Trator	93,61*
Colhedora automotriz	87,63*
Outros equipamentos	62,70*

Fonte: *Média de dois autores (CANAKCI et al., 2005; MOUSAVI-AVVAL et al., 2011).

A quantidade de energia associada a um equipamento será utilizada durante o ciclo de vida útil deste. Em outras palavras, deverá haver um fator adicional que considere a vida útil do equipamento (em horas por exemplo) relacionado a taxa de utilização desse equipamento por unidade de área (horas por hectare) possibilitando assim realizar a estimativa de energia por unidade de área empregada em um cultivo específico. Esta é uma consideração já definida pela literatura (MOUSAVI-AVVAL et al., 2011) necessitando apenas de adequações a realidade dos produtores que compõem a amostra deste trabalho.

Na Tabela 1.2.2 são apresentados valores de vida útil e tempo por unidade de área para as máquinas e operações considerados como fatores adicionais para os 27 produtores de soja nas três safras de soja de 2014/2015 até 2016/2017. É possível verificar qual o cálculo (através da análise dimensional) para obter cada equivalente em energia por unidade de área (EEUA) por produtor. O demonstrativo de cálculo é feito realizando a multiplicação

$$EEUA (MJ\ ha^{-1}) = (MJ\ kg^{-1}) \cdot kg \cdot h^{-1} \cdot (h\ ha^{-1}) \quad (1.2.1)$$

em que todos os valores referentes as unidades entre parênteses à direita da igualdade poderão ser encontradas nas Tabela 1.2.1 e Tabela 1.2.2. A unidade h^{-1} está relacionada com a vida útil dos equipamentos apresentados na Tabela 1.2.2 e finalmente a unidade kg refere-se a massa dos

equipamentos utilizados pelos produtores (obtido de site e catálogos de fabricantes dos equipamentos informados e/ou equivalentes no caso de modelos fora de linha). Alerta-se ainda que para cada operação haverá diferentes valores de EEUA. Por exemplo, semeadura corresponde a um valor de equivalente em energia assim como a pulverização (operação de cultivo) terá outro valor. Isso decorre da imediata percepção de que máquinas diferentes realizam operações diferentes em tempos diferentes.

Tabela 1.2.2 – Fatores de tempo utilizados para estimativa de equivalentes em energia por unidade de área (EEUA) para máquinas e equipamentos utilizados em sistemas de produção agrícola na região de Boa Esperança-PR

Operações	Equipamento	Vida útil - horas* (h)	Tempo por unidade de área** (h ha ⁻¹)
Colheita	Colhedora automotriz	7000	0,855
	Trator 1	7000	0,601
Semeadura	Semeadora	3500	0,601
	Trator 2	9000	0,206
Cultivo	Pulverizador	3000	0,125
	Aplicador de sólidos	482	0,081

Notas: * Considerando: 20,5 anos de vida útil de uma colhedora automotriz trabalhando 400 ha ano⁻¹; 14,6 anos de vida útil para uma semeadora que realiza semeadura de 400 ha ano⁻¹; 29,1 anos para uma Trator 1 que realiza a mesma área da semeadora; 10 anos de vida útil para um pulverizador que realiza a pulverização de 2400 ha ano⁻¹; 15 anos de vida útil para um aplicador de sólidos que realiza a aplicação a lanço de 400 ha ano⁻¹ e 15,6 anos de vida útil para um Trator 2 que realiza as operações de pulverização e aplicação de sólidos a lanço. ** Considerando: velocidades de 2,17 km h⁻¹ e 5,4 m de largura da plataforma de corte na colheita; velocidade de 3,70 km h⁻¹ e largura de 4,5 m na semeadura; velocidade de 4,44 km h⁻¹ e largura de 18 m na operação de pulverização; velocidade de 6,90 km h⁻¹ e largura de 18 m na aplicação de sólidos (fertilizantes) à lanço.

Fonte: valores médios observados em 27 produtores de soja em Boa Esperança-PR nas safras 2014/2015 até 2016/2017.

1.2.2 Mão-de-obra na agricultura e equivalentes em energia

Na operacionalização dos sistemas de manejo e execução dos itinerários técnicos das mais diversas culturas (HIRAKURI et al., 2012) a atividade agrícola tem inerentemente a sua realização a participação de mão-de-obra. Essa utilização de mão-de-obra deve ser considerada no balanço de energia. A depender da cultura e nível de mecanização adotada será utilizada maior ou menor quantidade de mão-de-obra.

Na literatura são diversos os exemplos de trabalhos que consideram essa entrada de energia para as mais variadas culturas e sistemas de produção (BEHESHTI TABAR; KEYHANI; RAFIEE, 2010; MONDANI et al., 2017; SINGH et al., 2008). Para este trabalho o valor utilizado na conversão em equivalentes de energia da mão-de-obra utilizada na produção de soja é de 1,77 MJ h⁻¹ obtido a partir da média de diversos autores (CHENG et al., 2017;

GHORBANI et al., 2011; KAZEMI et al., 2015; MONDANI et al., 2017; MOUSAVI-AVVAL et al., 2011; NASSIRI; SINGH, 2009; SAMANI et al., 2018; SHARIFI, 2018; STOUT, 1989; TABATABAEFFAR et al., 2009; WANG et al., 2014). Este número reflete quanto de energia um indivíduo dispense em uma hora de trabalho.

Contudo, verifica-se que esse valor necessita ser adequado dimensionalmente a fim de proporcionar um equivalente em energia por unidade de área. Essa adequação depende basicamente do tempo dispensado por unidade de área e do número de indivíduos que atuaram no processo (semeadura, pulverização, colheita, etc.). Fundamentalmente o tempo dispensado em operações segundo informações coletadas na amostra de produtores é destinado para atividades de semeadura, colheita, pulverização e aplicação de sólidos a lanço. Assim, os tempos dispensados pelos produtores por unidade de área será considerado o mesmo apresentado na Tabela 1.2.2 anteriormente. Já o número de participantes na(s) operação(ões) foram informados por cada um dos produtores. Logo para realizar as estimativas de equivalentes em energia por unidade de área tem-se o seguinte demonstrativo de cálculo

$$EEUA (MJ ha^{-1}) = MJ h^{-1} \sum_{i=1}^4 n_i \cdot (h ha^{-1})_i \quad (1.2.2)$$

em que $MJ h^{-1}$ é um valor constante ($1,77 MJ h^{-1}$), n_i denota o número de pessoas envolvidas na atividade i , a unidade $h ha^{-1}$ denota tempo por unidade de área (Tabela 1.2.2) e o índice i se refere ao número de operações (semeadura, colheita, pulverizações e aplicação de sólidos a lanço) por unidade de área para cada um dos produtores.

1.2.3 Equivalentes em energia para combustível

Perante um sistema de produção agrícola altamente dependente de mecanização e da crescente preocupação com a utilização de combustíveis procedentes de fontes não renováveis, a consideração da entrada de energia por meio de combustível é fundamental para a análise do sistema agrícola de produção de soja.

Nesta amostra de produtores de soja em estudo as atividades de semeadura, cultivo e colheita são feitos predominantemente com métodos mecanizados. E na totalidade dos produtores avaliados a fonte de energia utilizada na movimentação de máquinas é proveniente da combustão de diesel nos motores de máquinas agrícolas.

A utilização de motores diesel na agricultura atualmente é indispensável. Uma economia agrícola funcional de mercados globalizados é intrínseca ao uso de mecanização agrícola baseada em motores ciclo diesel (SMIL, 2010, p. 147). Não será difícil verificar a presença desse combustível fóssil em trabalhos de balanço de energia de culturas agrícolas diversas. Assim, o equivalente em energia por litro de diesel foi de $51,44 \text{ MJ L}^{-1}$ obtido por meio da média dos valores divulgados por diversos autores (GHORBANI et al., 2011; KAZEMI et al., 2015; MONDANI et al., 2017; MOUSAVI-AVVAL et al., 2011; NASSIRI; SINGH, 2009; PIMENTEL, 1992; SAMANI et al., 2018; SHARIFI, 2018; TABATABAEEFAR et al., 2009).

A adequação deste valor médio ($51,44 \text{ MJ L}^{-1}$) para obter um equivalente em energia por unidade de área dependerá do tipo de operação realizada e do consumo médio (L ha^{-1}) informado por cada produtor para as referidas operações de tal forma que o demonstrativo de cálculo é dado a seguir

$$EEUA (\text{MJ ha}^{-1}) = \text{MJ L}^{-1} \sum_{i=1}^4 (\text{L ha}^{-1})_i \quad (1.2.3)$$

em que MJ L^{-1} é um valor constante ($51,44 \text{ MJ L}^{-1}$) o índice i refere-se as quatro operações (semeadura, colheita, pulverizações e aplicação de sólidos a lanço).

1.2.4 Equivalentes em energia para macronutrientes e micronutrientes

A utilização de fertilizantes químicos industrializados é uma prática extensamente difundida no itinerário técnico de grande parte das culturas e na produção de *commodities* como a soja na escala comercial. Portanto, torna-se fundamental a sua consideração em equivalentes de energia para análise do sistema de cultivo de uma *commodity*. Um exemplo (PIMENTEL, 1992, p. 20) é análise feita sobre a quantidade média de diversas fontes de entrada de energia na produção de um hectare de milho nos Estados Unidos no ano de 1983. Dentre diversas entradas encontram-se os valores para nitrogênio (N), fósforo (P_2O_5) e potássio (K_2O).

Stout (1989, pp. 60, 67) também detalha as quantidades em equivalentes de energia para produção, transporte e distribuição de fertilizantes fosfatados e potássicos. Trata os valores em “*oil equivalents*” que podem ser convertidos em unidades de Joules e seus múltiplos seguindo os padrões do sistema internacional de medidas. Os autores deste e do parágrafo anterior são nitidamente clássicos no estudo de energia e uso na agricultura. Contudo, optou-se pela utilização de valores médios de valores disponibilizados por diversos autores e que podem ser verificados na Tabela 1.2.3.

Tabela 1.2.3 – Equivalentes de energia por unidade de massa ($MJ\ kg^{-1}$) de nitrogênio (N), fósforo (P_2O_5), potássio (K_2O) e micronutrientes diluídos segundo diferentes autores

Nutrientes	$MJ\ kg^{-1}$
N	72,8*
P_2O_5	18,7*
K_2O	11,7*
Micronutrientes	120,0**

Fontes: * Média de diversos autores (FISCHER, 1999; MOUSAVI-AVVAL et al., 2011; ORTIZ-CAÑAVATE; HERNANZ, 1999; PIMENTEL, 1992; SAMANI et al., 2018; SHARIFI, 2018; STOUT, 1989). ** Em acordo com Mondani et al. (2017)

Para encontrar os valores de equivalentes em energia por unidade de área para cada um dos nutrientes foi necessário decompor a massa utilizada de cada um destes nutrientes descritos na Tabela 1.2.3. Apesar de ser um processo dispendioso de tempo, a partir das informações de quantidades e tipos de fertilizantes utilizados (formulação e elementos simples) por cada um dos produtores isso é possível. Dessa forma a EEUA para N, P_2O_5 e K_2O pode ser obtida da seguinte maneira

$$EEUA (MJ\ ha^{-1}) = MJ\ kg^{-1} \sum_{i=1}^2 (kg\ ha^{-1})_i \quad (1.2.4)$$

em que o índice i de 1 até 2 refere-se as operações de adubações de base e aplicação de fertilizantes em cobertura, $MJ\ kg^{-1}$ se refere aos fatores constantes da Tabela 1.2.3 e $kg\ ha^{-1}$ é a massa de N, P_2O_5 ou K_2O utilizada por produtor amostrado.

No caso específico dos micronutrientes diluídos aplicados por via líquida nos procedimentos de pulverização existe uma grande diversidade de produtos e formulações comerciais usadas pelos produtores. Dessa forma seguindo o mesmo raciocínio apontado na relação (1.2.4) foram analisados os valores de $kg\ ha^{-1}$ de produto(s) comercial(ais) considerando a densidade de $1\ kg\ L^{-1}$ independente de marcas e micronutrientes fornecidos na formulação. Tal providência decorre da compreensão de que obter um valor relacionado à cada micronutriente em determinada formulação seria demasiadamente moroso de ser obtido, sendo desnecessária tal profundidade de distinção para este trabalho.

1.2.5 Sementes e grãos produzidos e seus equivalentes em energia

Os grãos produzidos durante um cultivo de um cereal ou oleaginosa, como no caso da soja neste estudo, carrega consigo uma quantidade de energia característica. O mesmo acontece para as sementes que são utilizadas como fonte de propagação elementar dando origem as plantas no processo de instalação dos cultivos. Valores para realizar essas conversões serão utilizados de acordo com a bibliografia (MOUSAVI-AVVAL et al., 2011; ORTEGA et al., 2005). No caso deste trabalho será considerado o valor de 25 MJ kg^{-1} tanto para sementes como para a grãos produzidos. Dessa forma o EEUA será obtido por meio da seguinte relação

$$EEUA (\text{MJ ha}^{-1}) = \text{MJ kg}^{-1} \cdot (\text{kg ha}^{-1}) \quad (1.2.5)$$

em que o valor entre parênteses se refere à quantidade de sementes utilizada na semeadura ou produção colhida por unidade de área por cada um dos produtores componentes da amostra.

1.2.6 Energia química da água

Como discutido anteriormente a água é objeto de iniciativas (SATTLER; MATZDORF, 2013) que visam preservar ou garantir sua qualidade e/ou disponibilidade. Para os estudos de balanço de energia é comum encontrar ponderações a respeito da quantidade de energia associada desse importante recurso para produção agrícola (MOUSAVI-AVVAL et al., 2011; SHARIFI, 2018; WANG et al., 2014).

Para estimar o equivalente em energia por unidade de área neste caso seria necessário registro em cada produtor sobre dados que permitissem estimar o consumo de água ou a evapotranspiração da cultura. Tal constatação remete a imediata impossibilidade de obter tal informação com tamanha precisão. Contudo, o INMET (Instituto Nacional de Meteorologia) disponibilizou registros de velocidade do vento, temperatura e umidade do ar do município de Goioêre-PR, código da estação INMET A825 e OMN 86917 (INMET, 2019).

Os registros desta estação (a mais próximo de Boa Esperança-PR com dados disponibilizados das safras desta pesquisa) foram utilizados na estimativa da evapotranspiração de referência (ET_o) diária (CONCEIÇÃO, 2006) e total do período entre setembro e fevereiro de cada safra. Tal valor foi utilizado como referencial médio de demanda de água na região para a cultura da soja. A média das safras do somatório das ET_o diárias estimadas pelo método de

Penman-Monteith foi de 0,637 m (637,15 mm) por safra. Ainda que não seja a estimativa de evapotranspiração da cultura este valor de evapotranspiração de referência situa-se dentro das faixa de necessidade hídrica da cultura pelo que se verifica na literatura (CARVALHO et al., 2013; FARIAS; NEPOMUCENO; NEUMAIER, 2007).

Dessa maneira é possível determinar a energia química da água nos mesmos termos utilizado na literatura (CHENG et al., 2017) e descrito a seguir

$$EEUA (MJ ha^{-1}) = 4.94 \times 10^{-6} MJ kg^{-1} \cdot \left\{ 10.000 m^2 ha^{-1} \cdot 1000 kg m^3 \right\} \cdot \bar{m} \quad (1.2.6)$$

em que o valor \bar{m} se refere à média nas três safras do somatório das ETo diárias estimadas pelo método de Penman-Monteith (CONCEIÇÃO, 2006) no período entre setembro e fevereiro, o valor de $4,94 \times 10^{-6} MJ kg^{-1}$ é a energia livre de Gibbs (constante) e entre colchetes são adequações de unidades.

1.2.7 Equivalentes em energia para agrotóxicos utilizados na agricultura

Um dos mais polêmicos pontos da agricultura de larga escala repousa sobre a necessidade e utilização de pesticidas para o cultivo. O objetivo nesta seção não é debater a viabilidade ou periculosidade do uso destes produtos. Dessa forma, compreenda-se que os termos pesticidas, agrotóxicos adotados por alguns autores (SCORZA JÚNIOR, 2006) serão considerados sinônimas para o termo defensivos agrícolas sem qualquer prejuízo ao conteúdo pretendido nesta seção.

A utilização desses defensivos agrícolas nas suas variadas classes como inseticidas, fungicidas, herbicidas, etc. têm relação direta com o desenvolvimento da agricultura moderna alicerçada no paradigma da revolução verde (MATSON, 1998; SCORZA JÚNIOR, 2006). Por isso, é razoável que a análise considere as diferentes classes como realizado em alguns estudos (MOUSAVI-AVVAL et al., 2011, 2013).

Foram solicitados para produtores de soja a área cultivada, quantidade e produtos utilizados na cultura que continham algum ingrediente ativo das classes fungicida, inseticida e herbicida e os respectivos adjuvantes. O inventário de utilização de pesticidas para cada produtor foi determinado para as seguintes safras: 2014/2015, 2015/2016 e 2016/2017. Herbicidas, fungicidas e inseticidas foram considerados em Kg de ingredientes

ativos segundo as descrições em bula dos fabricantes. Adjuvantes em Kg de produto comercial arbitrando a densidade de $1\text{ Kg }L^{-1}$ de produto comercial.

Os dados obtidos com os produtores foram convertidos em equivalentes em energia por unidade de área utilizando os fatores de conversão descritos na Tabela 1.2.4. No caso específico dos herbicidas foi realizada uma consideração de acordo com a participação dos principais mecanismos de ação herbicida utilizados no manejo de plantas daninhas pelo grupo de produtores (Figura 1.2.1). Conforme se verificou, da quantidade total de quilogramas de ingredientes ativos (i.a.) de defensivos utilizados pelo grupo de produtores, pouco mais de 75% foram da classe de herbicidas.

Ainda com relação a classe herbicida, ao analisar os mecanismos de ação utilizados pelo grupo nas três safras (Figura 1.2.1), verifica-se respectivamente os valores de 59,39% e 25,14% da quantidade total de quilogramas de i.a. para os mecanismos de ação Inibidores da EPSPs (Enol Piruvato Shiquimato Fosfato sintase) e auxinas sintéticas (mimetizadores das auxinas). Esses números representam a dependência do manejo de controle de plantas daninhas das moléculas de glifosato e 2,4-D. Assim, torna-se pertinente a realização das considerações apontadas anteriormente e registradas na literatura (TEIXEIRA; DIAS, 2019) em acordo com a Tabela 1.2.5.

Tabela 1.2.4 – Equivalentes de energia por unidade de massa ($MJ\text{ kg}^{-1}$) de ingrediente ativo (i. a.) de diferentes classes pesticidas empregadas na produção de soja

Classe pesticida	$MJ\text{ Kg}^{-1}$
Adjuvantes	38,7 ^a
Fungicidas	111,0 ^b
Herbicidas	369,5 ^c
Inseticidas	203,8 ^d

Fontes: ^a Reineke, Stockfisch e Märlander (2013). ^{b,d} Média de vários autores (FISCHER, 1999; MOUSAVI-AVVAL et al., 2011; ORTIZ-CANAVATE; HERNANZ, 1999; PIMENTEL, 1992; STOUT, 1989). ^c Determinado de acordo com a participação dos mecanismos de ação herbicidas empregados pelos produtores (TEIXEIRA; DIAS, 2019).

Essa mesma análise não será viável para as classes inseticida e fungicida em função dos dados clássicos (PIMENTEL, 1992; STOUT, 1989) disponíveis. Para inseticidas, os valores referência de equivalência em energia por unidade de massa de i.a. remetem a gerações ultrapassadas como grupos químicos dos piretróides e organofosforados. Na classe de fungicidas as moléculas são em geral fungicidas multissítios. Estes inseticidas e fungicidas são utilizados pelos produtores analisados mas não são os principais protagonistas no itinerário técnico utilizado no manejo de pragas e doenças. Dessa forma, valores médios (Tabela 1.2.4) da literatura atendem as necessidades para os objetivos pretendidos neste trabalho.

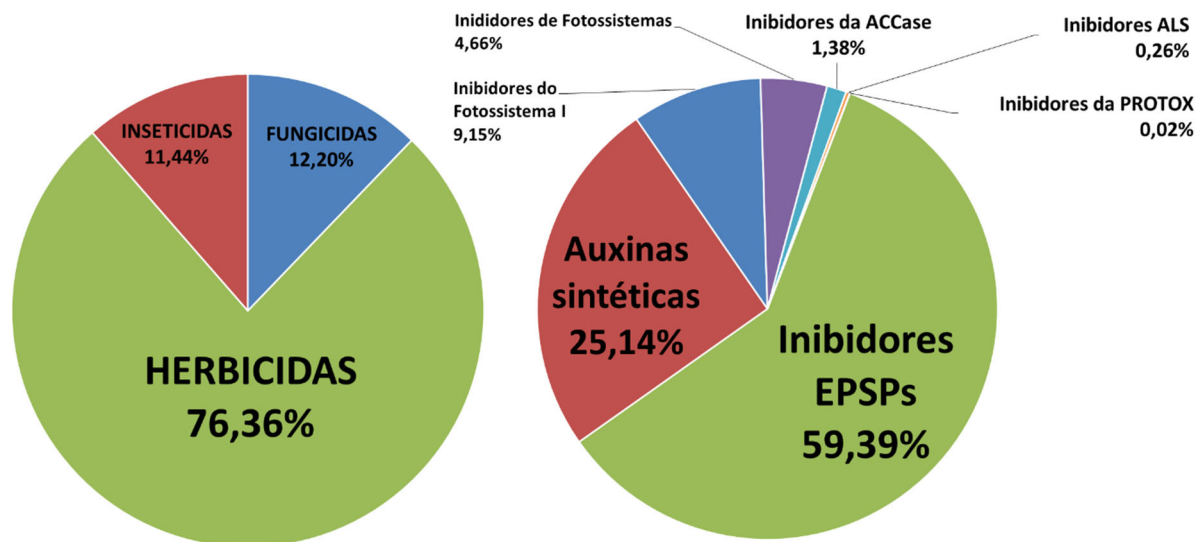
Tabela 1.2.5 – Determinação do fator de conversão em equivalente em energia por unidade de massa ($MJ\ kg^{-1}$) dos diferentes mecanismos de ação da classe herbicida de acordo com a participação dos mecanismos no grupo amostral.

Mecanismo de ação	Médias ^a	Participação ^b	Determinado
Inibidores da EPSPs	518,0	59,34%	261,0
Mimetizadores de auxina	196,4	25,14%	49,4
Inibidor do FS I	426,8	9,15%	39,1
Inibidores de FS I + FS II	261,9	4,66%	15,3
Outros	275,8	1,71%	4,7
Total		100,00%	369,5

Notas: EPSPs – Enol piruvato shiquimato fosfato sintase; FS I – Fotossistema I; FS II – Fotossistema II.

Fontes: ^aMédias ($MJ\ kg^{-1}$) baseadas em diferentes ingredientes ativos herbicidas (PIMENTEL, 1992; HELSEL, 2018) agrupados por mecanismo de ação. ^b Participação por mecanismo de ação herbicida nos 27 produtores avaliados durante as safras de 2014/2015, 2015/2016 e 2016/2017.

Figura 1.2.1 – Análise sobre a utilização de pesticidas para um grupo de 27 produtores de soja durante três safras de cultivo em Boa Esperança-PR. Valores em percentual da quantidade total de ingredientes ativos. Esquerda classes de pesticidas e à direita classe herbicida por mecanismo de ação.



Fonte: Determinado de acordo com a participação dos mecanismos de ação herbicidas empregados pelos produtores (TEIXEIRA; DIAS, 2019).

1.2.8 Balanço de energia

Para obtenção do balanço de energia foram realizadas as médias dos 27 produtores nas três safras. Foram considerados como entradas os equivalentes em energia por unidade de área com: sementes, nutrientes (N, P_2O_5 , K_2O e micronutrientes), pesticidas (herbicidas, inseticidas, fungicidas

e adjuvantes – óleos e espalhantes adesivos), mão-de-obra, mecanização (Trator 1, Trator 2, semeadora adubadora de plantio direto, pulverizador de arrasto, aplicador de sólidos à lanço e colhedora autopropelida) e energia química da água (energia livre de Gibbs). A saída de energia: produtividade de soja em equivalente de energia por unidade de área.

1.3 RESULTADOS E DISCUSSÃO

Os resultados de balanço de energia considerando as diferentes entradas e saídas de energia do sistema de cultivo de soja em 27 produtores durante as safras de 2014/2015, 2015/2016 e 2016/2017 é apresentado a seguir (Tabela 1.3.1).

Tabela 1.3.1 – Balanço de equivalentes em energia para a produção de soja

Entradas	Quantidade ¹ (Unidade ha ⁻¹)	Equivalentes em Energia ¹ (MJ ha ⁻¹)
Sementes		1.311,09
Sementes (kg)	52,44	1.311,09
Nutrientes		2.633,71
N (kg)	6,75	490,94
P ₂ O ₅ (kg)	67,46	1.263,41
K ₂ O (kg)	60,71	731,03
Micronutrientes (kg de p. c.) ²	1,24	148,32
Pesticidas		1.421,35
Herbicidas (kg de i.a.) ³	3,26	1.206,02
Inseticidas (kg de i.a.) ³	0,53	108,90
Fungicidas (kg de i.a.) ³	0,57	63,66
Adjuvantes (kg de p. c.) ²	1,11	42,77
Operacionais		4.055,03
Mecanização	----	319,19
Combustível (L)	72,50	3.729,96
Mão-de-obra (h)	3,32	5,88
Água		31.467,80
Energia química (Livre de Gibbs)	----	31.467,80
Total - Entradas		40.888,98
Saída		
Produtividade (kg)	3780,95	94.523,68
Total - Saída		94.523,68
Balanço		+ 53.634,69
Razão (Saída:Entradas)		2,31 (~ 2,3:1)

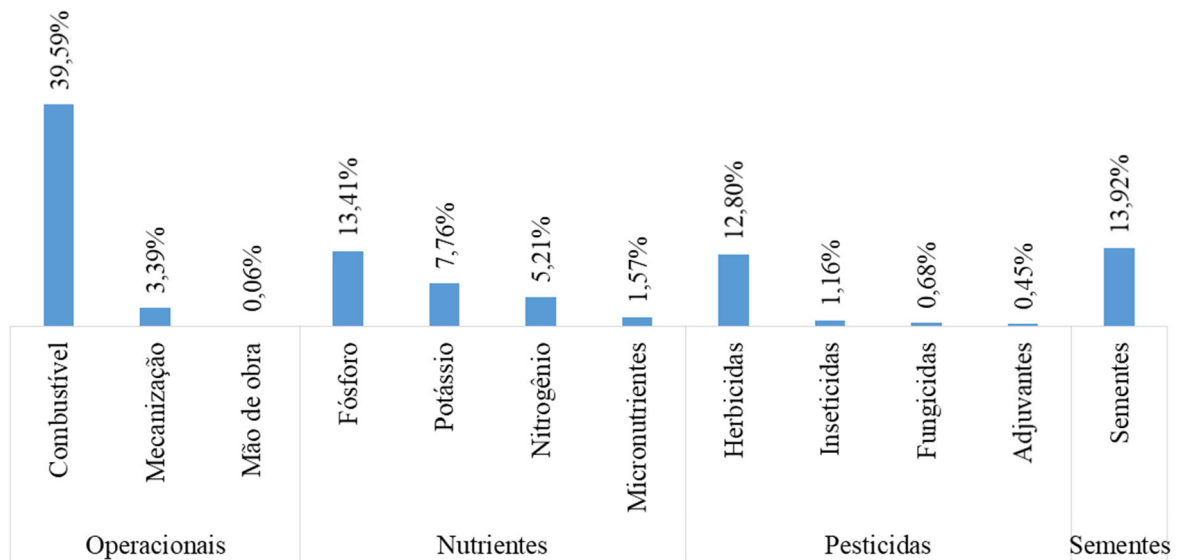
Notas: ¹ Média de 27 produtores de soja durante as safras de 2014/2015, 2015/2016 e 2016/2017; ² p. c. – produto comercial considerando densidade de 1 kg L⁻¹ para produtos em formulações líquidas; ³ i. a. – ingrediente(s) ativo(s).

Fonte: O autor.

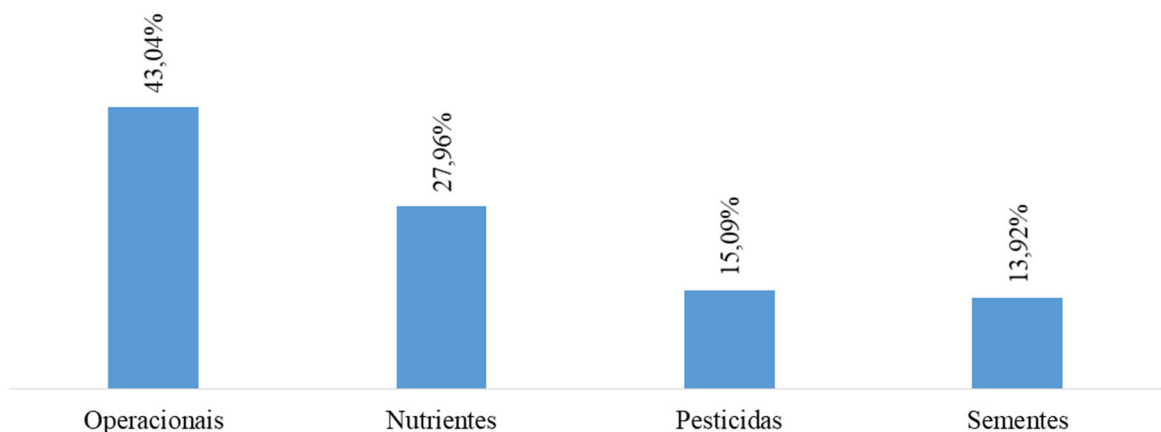
O balanço de energia na produção de soja foi positivo considerando a média dos 27 produtores nas três safras analisadas (Tabela 1.3.1). O total de energia empregado na produção de soja foi de $40.888,98 \text{ MJ ha}^{-1}$. Porém, a energia química da água responde por $31.467,80 \text{ MJ ha}^{-1}$ que foi estimada como sendo um valor constante para todos os produtores já que não se trata de cultura irrigada e não há interferência de manejo nesta quantidade. Logo as demais fontes de entrada de energia no sistema se referem à $9.421,18 \text{ MJ ha}^{-1}$ (Figura 1.3.1).

Figura 1.3.1 – Porcentagem da quantidade de energia empregada em relação ao total de entradas dos grupos indicados na horizontal (a) e seus sub itens e em relação ao total de energia dos quatro grupos (b) exceto a água.

(a) Em relação ao total de energia dos grupos indicados na horizontal.



(b) Em relação aos total de energia dos quatro grupos indicados no eixo horizontal.



Nota: Fósforo – P_2O_5 , Potássio – K_2O , Nitrogênio – N em acordo com o que se apresentou na Tabela 1.3.1.
Fonte: O autor.

Pela Figura 1.3.1 é possível verificar que do total de energia a que se referem os quatro grupos excetuando-se a energia química da água, as despesas operacionais totalizaram 43,04% do equivalente em energia. Dentro deste grupo o combustível representou 39,59%, mecanização 3,39% e mão-de-obra 0,06%. Considerando os dados da Tabela 1.3.1 se observa uma dependência da mecanização no manejo da cultura da soja e a redução da mão-de-obra que se restringe apenas as operações mecanizadas envolvendo poucas pessoas nas diversas etapas (plantio, cultivo, colheita). Tal constatação é observada ao comparar os resultados expostos na Tabela 1.3.1 aos resultados publicados na literatura (FELICI et al., 2006; ROMANELLI; NARDI; SAAD, 2012) sobre a produção de soja no Brasil.

A Figura 1.3.1 revela que os gastos em energia com nutrientes representam 27,96% ao desconsiderar a energia química da água. Os gastos com fósforo (P_2O_5) são de 13,41%, potássio (K_2O) 7,76%, nitrogênio (N) 5,21% e com micronutrientes 1,57%. Os resultados apresentados na Tabela 1.3.1 são maiores do que os apresentados para cultura da soja em resultados no estado do Paraná (ROMANELLI; NARDI; SAAD, 2012) e de localidades do Irã (MOUSAVI-AVVAL et al., 2011). Contudo, verifica-se neste trabalho que as produtividades foram superiores o que pode ser reflexo dos maiores investimentos em nutrição.

Com relação aos pesticidas (Figura 1.3.1) empregados no cultivo da soja verifica-se que esse grupo representa 15,09% da energia ao excetar-se a energia química da água. Os herbicidas correspondem a 12,80%, inseticidas 1,16%, fungicidas 0,68% e adjuvantes 0,45%. A ordem de grandeza de consumo de pesticidas está de acordo com o que é apresentado em outros trabalhos (MOUSAVI-AVVAL et al., 2011, 2013) mas mostra a evolução no consumo de pesticidas em relação a registros anteriores (FELICI et al., 2006; ROMANELLI; NARDI; SAAD, 2012) principalmente da classe herbicida baseado em poucos produtos no manejo químico de plantas daninhas (TEIXEIRA; DIAS, 2019).

O balanço de energia na produção de soja foi positivo (Tabela 1.3.1). A razão aproximada entre saída e entradas foi de 2,3:1. Este resultado está em acordo com o que é exposto em outros trabalhos envolvendo soja (FELICI et al., 2006; MOUSAVI-AVVAL et al., 2011, 2013). Este resultado é adequado ao que se espera de um sistema de produção do ponto de vista da sua racionalidade, pois, um custo energético maior do que se produz não será razoável.

Finalmente é importante salientar que este trabalho não contemplou custos de energia provenientes de atividades de correção solo, transporte e ou armazenamento de produtos. Tais

custos podem ser importantes na composição de um balanço de energia impactando no saldo de energia podendo alterar a razão entre as entradas e saídas.

1.4 CONCLUSÕES

- O balanço de energia da cultura considerando os 27 produtores de soja nas safras 2014/2015, 2015/2016 e 2016/2017 foi positivo.
- Foi constatado um saldo de $53.634,69 \text{ MJ ha}^{-1}$ debitando do equivalente em energia da produtividade os gastos equivalentes em energia operacionais, com sementes, com pesticidas, com nutrientes e a energia química da água utilizados na produção.
- A razão entre os totais de saída e entradas foi de 2,31. Essa razão aproximada de 2,3:1 indica que para cada MJ investido se produziu $2,3 \text{ MJ}$ por unidade de área indicando racionalidade de um ponto de vista do balanço de energia.
- Os custos energéticos na produção de soja em ordem decrescente foram com água, operacionais, nutrientes, pesticidas e sementes.
- Os dispêndios de energia com Combustível (39,59%), Fósforo (13,41%), Sementes (13,92%) e Herbicidas (12,80%) correspondem a mais de 75% do total de energia ao desconsiderar a água.

CAPÍTULO 2 – ESTADO DA ARTE: ASPECTOS ELEMENTARES SOBRE ANÁLISE ENVOLTÓRIA DE DADOS

2.1 REVISÃO DE LITERATURA

Neste capítulo é feita uma abordagem sobre o estado da arte a respeito da análise envoltória de dados e suas aplicações em sistemas agrícolas. É possível encontrar vários trabalhos publicados que utilizam os modelos de análise envoltória a fim de realizar a avaliação sobre a eficiência de produção ou sistemas de produção.

Essa técnica tem a vantagem de construir valores de eficiência relativa, o que é mais razoável do ponto de vista de gestão do que a utilização de metas intangíveis. Uma abordagem matemática a respeito dos modelos mais simples é realizada com o intuito de criar um ambiente propício a compreensão de como os valores de eficiência são obtidos bem como o seu significado.

2.1.1 Análise Envoltória de Dados – *Data Envelopment Analysis* (DEA)

A aplicação de DEA permite encontrar índices de eficiência relativa entre elementos de um grupo. Esses elementos ou organizações em estudo são chamados de unidades formadoras de decisão (DMU – *decision making unit*) (COOPER; SEIFORD, 2007, p. 22). O que determina a eficiência destas DMU é a relação entradas (insumos para produção) e saídas (produtos).

Na essência a DEA consiste em transformar um problema fracional em um problema de programação linear (COOPER; SEIFORD, 2007, p. 23) utilizando os conceitos de álgebras linear, vetorial e de matrizes. A resolução deste problema para determinar os índices de eficiência é feita facilmente utilizando computadores e softwares (COELLI, 1996; HŘEBÍČEK et al., 2015; SILVA et al., 2017).

O modelo básico de DEA é o que assume retorno constante por escala (*constant return to scale* - CRS) proposto por Charnes, Cooper e Rhodes (1978) e em alguns trabalhos identificado pelas iniciais destes autores (CCR). Do modelo básico surgiu a análise com base em retorno variável por escala (*variable return to scale* – VRS) e também associada às iniciais BCC dos autores do trabalho (BANKER; CHARNES; COOPER, 1984).

A partir dos valores de eficiência técnica (CRS) e eficiência técnica pura (VRS) é possível determinar a escala de eficiência (SE – *scale efficiency*) pela razão entre a eficiência técnica e eficiência técnica pura (COOPER; SEIFORD, 2007, p. 153; MOUSAVI-AVVAL et al., 2011). Nos tópicos a seguir será feita uma abordagem a respeito de aplicações e de aspectos matemáticos da análise envoltória de dados.

2.1.2 Aplicações de análise envoltória de dados

A análise envoltória de dados (*data envelopment analysis* – DEA) é uma técnica que permite a avaliação do desempenho e eficiência dos mais variados sistemas. O desempenho de CEO's (*chief executive officers*) e suas práticas de gestão (CHEN; GREGORIOU; ROUAH, 2016), eficiência de organizações de transporte (YU; CHEN, 2016) e o impacto da incorporação de índices de satisfação de consumidores sobre a eficiência de custos de redes de distribuição (GROWITSCH et al., 2016) são alguns dos exemplos de empregos de DEA.

No Brasil estudos de eficiência de produção de citrus utilizaram DEA como ferramenta de análise concluindo que boa parte dos produtores de citrus da região de São Paulo são ineficientes e as variáveis que mais contribuem para a eficiência são a escolaridade e a experiência como produtor (CLEMENTE; LÍRIO; GOMES, 2015).

Um trabalho associado a empresas familiares e não familiares utilizou DEA para determinar a eficiência destes dois diferentes grupos (GUSE et al., 2015). As conclusões deste trabalho indicam que, apesar de não significativos, empresas não familiares apresentam resultados mais eficientes que empresas familiares no curto prazo devido à gestão profissionalizada das empresas não familiares priorizar o desempenho de curto prazo.

Determinar as eficiências relativas de terminais intermodais de transporte de grãos por meio de DEA foi objeto de estudo visando avaliar a eficiência em dois períodos do ano (SANTOS et al., 2016). Os resultados apontaram para um baixo nível de eficiência dos terminais intermodais devido à ociosidade na operação. Entrevistas com gerentes de algumas unidades estudadas apontaram, dentre outras causas externas, a baixa disponibilidade de vagões para transporte ferroviário. Os autores ainda apontaram outras três hipóteses a serem consideradas em estudos futuros: produção de grãos na região é insuficiente para a capacidade dos terminais, infraestrutura do terminal superestimada na sua construção e poder de mercado de companhias multinacionais com maior participação no transporte ferroviário.

Batista et al. (2016) fez uso de DEA para determinar o índice de eficiência produtiva em cultivos intercalados envolvendo a cenoura. Pontuou que o modelo baseado em DEA calculou a eficiência relativa das unidades produtivas usando programação linear proporcionando estimativa de quais unidades têm as melhores práticas de consórcio.

Da mesma forma a avaliação de eficiência de horticultores agroecológicos (OLIVEIRA et al., 2014) foi feita utilizando a análise envoltória de dados. Neste estudo, sete lotes de terras apresentaram índices de eficiência média superiores à média geral. Por meio desse indicativo constatou-se que esse desempenho superior era devido a estratégias diferenciadas de produção nos lotes com variedade de produtos e intensificação de produção atingindo o público consumidor com maior frequência nas feiras.

Estudos da cadeia produtiva animal também utilizaram DEA na determinação de índices de desempenho (GOMES et al., 2015). Os resultados permitiram apontar fontes de ineficiência em mão-de-obra de baixa qualificação e uso de touros de qualidade questionável, fatores comuns em sistemas extensivos. Estes são alguns dos principais gargalos nos sistemas de produção animal como um todo.

A cana-de-açúcar é a cultura que com maior frequência aparece em estudos envolvendo DEA. Contudo, o emprego da técnica fica restrito a avaliação dos sistemas operacionais de industrialização ou custos de produção da cultura. Um exemplo é a avaliação entre usinas no território brasileiro (SALGADO JUNIOR et al., 2014). Tamanho e localização das usinas foram considerados como fatores que alteram a eficiência operacional. Concluiu-se que o tamanho influencia a eficiência em função do maior emprego de tecnologia. Já a localização, proporciona condições edafoclimáticas para maiores teores de sacarose na cana, influenciando a eficiência operacional da usina.

Um estudo de eficiência operacional de usinas de cana-de-açúcar foi conduzido usando DEA identificando as unidades mais eficientes dentre um grupo de 355 unidades (JUNIOR; CARLUCCI; GRESPAN, 2014). Além de apontar a tendência de eficiência de usinas localizadas no estado de São Paulo, foi possível apontar, em termos de eficiência operacional, os estados de Alagoas, Pernambuco e algumas regiões de Minas Gerais, Paraná e Mato Grosso do Sul com potencial para investimento em novas usinas.

Pereira & Silveira (2016) também realizam um estudo aprofundado da eficiência produtiva de usinas de cana-de-açúcar fazendo uso de DEA. Suas conclusões foram basicamente que a componente tecnológica desempenha papel fundamental na eficiência das unidades que se destacaram. Afirma-se ainda a importância do investimento em pesquisa e

desenvolvimento no setor para não tornar-se apenas um tomador de pacotes tecnológicos padrões de mercado.

Os custos de produção de cana-de-açúcar em diferentes regiões de produção foram objetos de estudo e foram avaliados com a utilização de DEA (PEREIRA; TAVARES, 2017). Neste caso foram consideradas as produções de cada região (saídas) e custos (entradas) com mecanização, mão-de-obra, suprimentos, *leasing*, administrativos, remuneração da terra e capital. Uma das conclusões é que a região centro sul teve maior índice de tecnificação.

Produtores de soja no Irã foram avaliados de acordo com a eficiência energética na produção de soja utilizando DEA (MOUSAVI-AVVAL et al., 2011, 2013). Foi avaliado um grupo de produtores sendo estes classificados de acordo com os índices de eficiência. Trabalho semelhante na cultura do milho (OMMANI, 2011) foi realizado com produtores do Irã. No Brasil as iniciativas que estudam ou fazem uma análise energética na soja (FELICI et al., 2006) e milho (LORENZONI MARTINS et al., 2015) podem ser encontradas, porém sem realizar uma comparação entre diferentes unidades (produtores).

2.2 ANÁLISE ENVOLTÓRIA DE DADOS (DEA) – ASPECTOS MATEMÁTICOS

O modelo inicial desenvolvido para DEA (CHARNES; COOPER; RHODES, 1978) é o que assume o retorno constante por escala (CRS – *constant return to scale*). Surgiu com base no do trabalho pioneiro de Farrell do final da década de 50. Inicialmente discute-se aqui o modelo básico orientado para entradas (minimizar as entradas satisfazendo no mínimo os níveis observados das saídas) apresentando conceitos elementares como conjunto possibilidades de produção que fornece a idealização de CRS.

Seja $\bar{y} = (y_1, y_2, \dots, y_s)$ um vetor em um espaço R^s que representa as saídas (*outputs*) de uma atividade de produção (de soja por exemplo) e $\bar{x} = (x_1, x_2, \dots, x_r)$ um vetor em um espaço R^r que representa as entradas (*inputs*) de uma atividade de produção (insumos utilizados como fatores de produção da soja por exemplo). Os índices s e r sobrescritos representam o números de variáveis de saída e entradas respectivamente. Pode-se descrever a máxima eficiência θ como uma razão entre o produto escalar de \bar{y} e \bar{x} pelos vetores $\bar{u} = (u_1, u_2, \dots, u_s)$ e $\bar{v} = (v_1, v_2, \dots, v_r)$ que representam respectivamente pesos ótimos para insumos e produtos, dessa forma tem-se

$$\max_{\bar{u}, \bar{v}} \theta = \frac{\bar{u} \cdot \bar{y}}{\bar{v} \cdot \bar{x}} \quad (2.2.1)$$

sujeito as seguintes restrições:

$$\frac{\vec{u} \cdot \vec{y}}{\vec{v} \cdot \vec{x}} \leq 1 \quad (2.2.2)$$

$$\begin{cases} \vec{u}, \vec{v} \neq \vec{0} \\ \vec{u}, \vec{v} \geq \vec{0} \end{cases} \quad (2.2.3)$$

em que lê-se: maximize θ em relação a \vec{u}, \vec{v} (2.2.1). Note que a condição de restrição imposta por (2.2.2) é o raio de eficiência, ou seja, a eficiência máxima será 1 (100%). A condição de (2.2.3) é dita “semipositividade” da solução vetorial (COOPER; SEIFORD, 2007, p. 42) pois impõem que algumas componentes dos vetores podem ser nulas, porém não todas, e as componentes não nulas serão positivas. Em resumo as componentes destes vetores \vec{u} e \vec{v} são pesos ótimos que maximizam o valor da eficiência θ .

A presença de vetores e as operações características de álgebra vetorial podem causar receio ao leitor menos dedicado à matemática. Porém, a forma apresentada anteriormente é a forma fracional do problema que pode ser visualizado como um problema de programação linear com algum manuseio.

2.3 A MODELAGEM DO PROBLEMA PELA PROGRAMAÇÃO LINEAR

É possível descrever a máxima eficiência θ de maneira diferenciada. O produto escalar de \vec{y} e \vec{x} pelos vetores $\vec{\tau} = (\tau_1, \tau_2, \dots, \tau_s)$ e $\vec{\varphi} = (\varphi_1, \varphi_2, \dots, \varphi_r)$. Estes dois últimos vetores representam respectivamente os pesos ótimos para produtos e insumos neste caso. Dessa forma tem-se

$$\max_{\vec{\tau}, \vec{\varphi}} \theta = \vec{\tau} \cdot \vec{y} \quad (2.3.1)$$

sujeito as seguintes restrições:

$$\vec{\varphi} \cdot \vec{x} = 1 \quad (2.3.2)$$

$$\vec{\tau} \cdot \vec{y} \leq \vec{\varphi} \cdot \vec{x} \quad (2.3.3)$$

$$\vec{\tau} \geq \vec{0} \quad (2.3.4)$$

$$\vec{\varphi} \geq \vec{0} \quad (2.3.5)$$

As equações de (2.3.1) até (2.3.5) compõe a versão do problema em uma forma de programação linear. É possível perceber que, com algum manuseio algébrico, as equações e condições de restrição de (2.3.1) até (2.3.5) refletem as mesmas condições do problema fracional descrito anteriormente entre (2.2.1) até (2.2.3) sendo a transformação entre a modelagem pela programação linear reversível à visão fracional do problema.

Suponha uma solução ótima dada por $\bar{\tau} = \bar{\tau}'$ e $\bar{\varphi} = \bar{\varphi}'$ para o problema de programação linear – equações (2.3.1) até (2.3.5). Dessa forma tem-se um valor $\theta = \theta'$ ótimo (maximizado) para solução do problema. Assim, graças a reversibilidade da transformação citada no parágrafo anterior, será possível uma solução para as equações (2.2.1) até (2.2.3) com vetores $\bar{u} = \bar{\tau}'$ e $\bar{v} = \bar{\varphi}'$. Portanto, a solução $\theta = \theta'$ é satisfatória para ambos: o problema fracional e o de programação linear (COOPER; SEIFORD, 2007). Detalhes matemáticos do modelo CRS (CHARNES; COOPER; RHODES, 1978) e peculiaridades de variações deste podem ser encontrados na literatura (CAMIOTO; MARIANO; REBELATTO, 2014; CAMIOTO; REBELATTO; ROCHA, 2015; MARIANO; REBELATTO, 2014).

2.4 O CONJUNTO P DE POSSIBILIDADES DE PRODUÇÃO

Até aqui toda a discussão a respeito da descoberta de um valor de θ que represente um índice de eficiência esteve centrado na análise e discussão de dados referentes a entradas e saídas de um sistema apenas. Em simples palavras, as análises anteriores representam uma empresa, corporação, mecanismo, etc. Contudo, é razoável que em um sistema de avaliação de eficiência operacional mais de uma entidade seja avaliada. Nesse contexto é que se estabelece o conjunto de possibilidades de produção.

Inicialmente relembre-se que os vetores \bar{y} e \bar{x} têm caráter semipositivo de acordo com as restrições estabelecidas em (2.2.3). Assim, considera-se que existe ao menos uma componente positiva e não nula para os vetores. Esta é uma consideração inteiramente razoável pois não há sentido em conceber uma atividade produtiva sem ao menos uma entrada e saída ou vice-versa.

Assim, suponha que n entidades estejam sendo avaliadas. Cada uma das n entidades é chamada de DMU (*decision making unit*). Portanto, a cada DMU está associado um par de vetores \bar{y}_j e \bar{x}_j com $j = 1, \dots, n$. Dessa forma um par de vetores (\bar{x}, \bar{y}) representa um ponto em um espaço multidimensional R^{s+r} suficiente para expressar as quantidades de entradas e saídas. O conjunto desses pontos (\bar{x}, \bar{y}) formam o conjunto de possibilidades de produção e cada um

dos pontos representa uma atividade. O conjunto P de possibilidades de produção é expresso por

$$P = \{(\bar{x}, \bar{y}) \mid \bar{x} \geq X\bar{\lambda}^t, \bar{y} \leq Y\bar{\lambda}^t, \bar{\lambda}^t > \bar{0}\} \quad (2.4.1)$$

em que X é uma matriz de dimensão $s \times n$ com as s entradas das n DMU's avaliadas, Y é uma matriz de dimensão $r \times n$ com as r saídas das n DMU's avaliadas e $\bar{\lambda}^t$ é um vetor não negativo com n componentes transposto.

O conjunto (2.4.1) satisfaz quatro propriedades propostas para o conjunto P (COOPER; SEIFORD, 2007, p. 42). Não serão todas mencionadas aqui, porém, uma especial atenção é destinada a segunda que afirma: “Se uma atividade (\bar{x}', \bar{y}') pertence à P , então a atividade $(b\bar{x}', b\bar{y}')$ pertence a P para qualquer escalar positivo b . Esta propriedade caracteriza o retorno constante por escala”.

2.5 O CARÁTER ENVOLTÓRIO (*ENVELOPMENT*) DA ANÁLISE DE DADOS

A definição do conjunto P de possibilidades de produção permite avaliar o problema de programação linear definido anteriormente nas equações (2.3.1) até (2.3.5) considerando as n DMU's. Para isso considere os vetores \bar{u}, \bar{v} definidos anteriormente e as matrizes X e Y definidas em (2.4.1) para entradas e saídas das n DMU's. Assim, para a DMU_o têm-se

$$\max_{\bar{u}, \bar{v}} \bar{u} \cdot \bar{y}_o \quad (2.5.1)$$

sujeito as seguintes restrições:

$$\bar{v} \cdot \bar{x}_o = 1 \quad (2.5.2)$$

$$\bar{u}Y \leq \bar{v}X \quad (2.5.3)$$

$$\bar{u} \geq \bar{0} \quad (2.5.4)$$

$$\bar{v} \geq \bar{0} \quad (2.5.5)$$

em que \bar{y}_o e \bar{x}_o são respectivamente os vetores referentes as saídas (produção) e entradas (insumos) da DMU_o.

Note que as equações (2.5.1) até (2.5.5) correspondem às equações (2.3.1) até (2.3.5). Apenas foram escritas incluindo as matrizes das n DMU's analisadas. Dessa forma são variáveis

desse problema de maximização os vetores \vec{u}, \vec{v} . Esse é o “formato de multiplicadores” do problema de programação linear.

Como mencionado procura-se desenvolver a idéia de eficiência orientado para entradas, ou seja, reduzir \vec{x}_o ao menor nível possível sem reduzir \vec{y}_o na DMU_o. Da definição determinada pelas desigualdades do conjunto P em (2.4.1) percebe-se que uma atividade qualquer atinge valor máximo quando

$$(\vec{x}, \vec{y}) = (X\vec{\lambda}^t, Y\vec{\lambda}^t) \quad (2.5.6)$$

Para a DMU_o procura-se reduzir \vec{x}_o ao menor nível possível sem reduzir \vec{y}_o . Dessa forma é preciso encontrar um índice que equalize a atividade da DMU_o da seguinte maneira

$$(\theta\vec{x}_o, \vec{y}_o) = (X\vec{\lambda}^t, Y\vec{\lambda}^t) \quad (2.5.7)$$

em que θ é um escalar. Considerando agora a orientação para entradas, as propriedades do conjunto P expressas em (2.4.1) e os critérios de semipositividade dos vetores \vec{x}_o e \vec{y}_o é possível expressar θ no seguinte problema de programação linear

$$\min_{\theta, \vec{\lambda}^t} \theta \quad (2.5.8)$$

sujeito as seguintes restrições:

$$\theta\vec{x}_o - X\vec{\lambda}^t \geq \vec{0} \quad (2.5.9)$$

$$Y\vec{\lambda}^t \geq \vec{y}_o \quad (2.5.10)$$

$$\vec{\lambda}^t > \vec{0} \quad (2.5.11)$$

Observa-se em (2.5.9) que o vetor $X\vec{\lambda}^t$ limita (inferiormente) o vetor proveniente do produto $\theta\vec{x}_o$. Considere três possibilidades a seguir:

- Supor $\vec{x}_o < X\vec{\lambda}^t$ implica que $\theta > 1$ para não violar a definição $\vec{x}_o \geq X\vec{\lambda}^t$ de (2.4.1). Porém θ não deve ultrapassar o valor máximo do raio de eficiência igual a 1, logo essa alternativa não é viável.
- Supondo $\vec{x}_o = X\vec{\lambda}^t$ implica que o valor θ deve ser neutro para multiplicação atendendo a definição $\theta\vec{x}_o \geq X\vec{\lambda}^t$ de (2.4.1), logo, $\theta = 1$.

- c. Supondo $\bar{x}_o > X\bar{\lambda}^t$ implica que existe um valor suficientemente pequeno de $\theta < 1$ e que não viola a definição $\theta\bar{x}_o \geq X\bar{\lambda}^t$ de (2.4.1). Ou seja, quando minimiza-se θ até atingir um valor que $\theta\bar{x}_o = X\bar{\lambda}^t$ se reduz as entradas ao menor nível possível sem alterar o nível das saídas.

Analisando (2.5.10) observa-se que o vetor $Y\bar{\lambda}^t$ está limitado (inferiormente) pelo vetor \bar{y}_o . Os limites discutidos ($\theta\bar{x}_o \geq X\bar{\lambda}^t$ e $Y\bar{\lambda}^t \geq \bar{y}_o$) impõem os valores ótimos para as variáveis θ e $\bar{\lambda}^t$ no problema de programação linear expresso por (2.5.8) até (2.5.11). Essa representação é a “forma envoltória” da análise (envelopada). As equivalências entre os formatos “multiplicadores” e “envoltório” é discutida com maior profundidade na literatura (COOPER; SEIFORD, 2007).

Analisando as restrições (2.5.9) e (2.5.10) se nota que os vetores $\theta\bar{x}_o$, \bar{y}_o da DMU_o na condição de menor valor de θ podem ser expressos como combinação linear respectivamente de \bar{x}_j e \bar{y}_j com $j = 1, \dots, n$ tendo como coeficientes as entradas de $\bar{\lambda}^t$ como segue

$$\begin{cases} \theta\bar{x}_o = \sum_{j=1}^n \lambda_j \bar{x}_j \\ \bar{y}_o = \sum_{j=1}^n \lambda_j \bar{y}_j \end{cases} \quad \forall \lambda_j \neq 0, \quad j=1, \dots, n \quad (2.5.12)$$

Porém na condição em que $\theta = 1$ as equações (2.5.12) tornam-se,

$$\begin{cases} \bar{x}_o = \sum_{j=1}^n \lambda_j \bar{x}_j \\ \bar{y}_o = \sum_{j=1}^n \lambda_j \bar{y}_j \end{cases} \quad \forall \lambda_j \neq 0, \quad j=1, \dots, n \quad (2.5.13)$$

Em (2.5.13) se identifica uma possibilidade importante. Como se assume o valor de $\theta = 1$ para a DMU_o é possível então que θ não possa ser minimizado abaixo de 1 (índice máximo de raio de eficiência), contudo, os fatores de produção tenham de ser “adequados” por qualquer combinação linear com algum(s) λ_j com índice $j \neq o$. Essa condição é conhecida como eficiência fraca (*weak efficiency*).

Essa constatação é importante para se compreender o conceito de “mix” de eficiência. É fácil visualizar que na condição de $\theta < 1$ o lado esquerdo da equação é alterado. Dessa forma, haverá necessariamente uma combinação linear (“mix”) que altera a parte direita da igualdade em (2.5.12).

Assim, se os fatores de produção de uma DMU_k qualquer necessitam ser adequados por qualquer combinação linear que envolva um λ_j com índice $j \neq k$, essa DMU_k não atingiu a eficiência plena (total) mesmo que assuma o valor máximo de raio de eficiência $\theta = 1$ (SUEYOSHI; GOTO, 2018, p. 27).

O parágrafo anterior é uma interpretação da definição de eficiência “forte” (“*strong efficiency*”) ou eficiência Pareto-Koopmans (COOPER; SEIFORD, 2007, p. 45) que diz: “Uma DMU é totalmente eficiente se, e somente se, não for possível melhorar qualquer entrada ou saída sem interferir em alguma outra entrada ou saída”. Uma análise para esta definição é dada a seguir.

Análise – De imediato, se para a DMU_o $\theta < 1$ as componentes do vetor \bar{x}_o são reduzidas pelo produto $\theta\bar{x}_o$. Assim, é feito um ajuste nas entradas e nesse caso a DMU_o não está no nível máximo de eficiência radial. Também $\lambda_o \neq 0$, pois, caso contrário os vetores \bar{x}_o e \bar{y}_o são “adequados” por combinação com outro(s) λ_j com índice $j \neq o$. Portanto assuma o menor índice de eficiência radial possível $\theta = 1$ e $\lambda_o \neq 0$, então de (2.5.13)

$$\begin{cases} \bar{x}_o = \sum_{j=1}^{o-1} \lambda_j \bar{x}_j + \lambda_o \bar{x}_o + \sum_{j=o+1}^n \lambda_j \bar{x}_j \\ \bar{y}_o = \sum_{j=1}^{o-1} \lambda_j \bar{y}_j + \lambda_o \bar{y}_o + \sum_{j=o+1}^n \lambda_j \bar{y}_j \end{cases} \quad \lambda_j = 0 \quad \forall j \neq o, \quad j = 1, \dots, o, \dots, n \quad (2.5.14)$$

Logo

$$\begin{cases} \bar{x}_o = \lambda_o \bar{x}_o \\ \bar{y}_o = \lambda_o \bar{y}_o \end{cases} \quad \lambda_j = 0 \quad \forall j \neq o, \quad j = 1, \dots, o, \dots, n \quad (2.5.15)$$

$$\Rightarrow \lambda_o = 1$$

Assim, a DMU_o não terá alterada suas componentes de entradas e saídas. Por outro lado, assuma o menor índice $\theta = 1$, $\lambda_o \neq 0$ e algum outro $\lambda_k \neq 0$. Dessa maneira tem-se

$$\begin{cases} \bar{x}_o = \lambda_o \bar{x}_o + \lambda_k \bar{x}_k \\ \bar{y}_o = \lambda_o \bar{y}_o + \lambda_k \bar{y}_k \end{cases} \quad \lambda_j = 0 \quad \forall j \neq o, k \quad j = 1, \dots, o, k, \dots, n \quad (2.5.16)$$

$$\Rightarrow$$

$$(1 - \lambda_o) = \lambda_k$$

Por suposição $\lambda_k \neq 0$, logo $\lambda_o \neq 1$, e dessa forma os vetores \bar{x}_o e \bar{y}_o são alterados em uma combinação linear que envolve os vetores \bar{x}_k e \bar{y}_k pertencentes a DMU_k . Assim, a DMU_o não pode estar na condição de eficiência total. ■

2.6 O AJUSTE DOS FATORES DE PRODUÇÃO

As interpretações de (2.5.12) até (2.5.16) possibilitaram compreender que se uma DMU_k qualquer não apresenta eficiência total os fatores de produção dessa DMU_k são ajustados por uma combinação linear de vetores (fatores de produção) de outra(s) DMU_j com $j \neq k$.

Logo, existem dois vetores de ajuste (entradas e saídas) cujas componentes são interpretadas como “folgas” (“*slacks*”) nos fatores de produção. Quando referentes as entradas o vetor ajuste indica os excessos nos fatores de produção. Para as saídas o vetor ajuste indica as deficiências na produção. Pode-se expressar esses vetores de ajuste da seguinte forma

$$\bar{a}^+ = \theta \bar{x} - X\bar{\lambda}^t \quad (2.6.1)$$

$$\bar{a}^- = Y\bar{\lambda}^t - \bar{y} \quad (2.6.2)$$

em que \bar{a} é o vetor ajuste de produção, cujo o sinal sobrescrito indica negativo para deficiências (saídas) e positivo para excessos (entradas).

Ainda, se o sinal sobrescrito é positivo \bar{a} pertence ao mesmo espaço dimensional de \bar{x} e se negativo pertence ao mesmo espaço dimensional de \bar{y} , portanto, o sinal sobrescrito denota que apesar de ambos terem o propósito de ajuste dos fatores de produção não se tratam do mesmo vetor. Se uma DMU qualquer não necessita de ajustes nos fatores de produção, ou seja, está na condição de eficiência total, então $\bar{a}^+ = \bar{0}$ e $\bar{a}^- = \bar{0}$. Caso isto não esteja claro é recomendável avaliar o raciocínio desenvolvido a partir de (2.5.13).

Dessa maneira as componentes dos vetores provenientes de (2.6.1) e (2.6.2), quando não nulas, fornecem as quantidades necessárias para adequação de fatores de produção para os

menores níveis possíveis para o retorno constante por escala considerando orientação para entradas como é proposto desde o início da discussão.

Não serão levadas adiante as demonstrações e análises, contudo, existe farto referencial a respeito dos ajustes e folgas (*slacks*) disponibilizados na literatura (BOGETOFT PEDERSEN; OTTO, 2011, p. 127; COOPER; SEIFORD, 2007, p. 44; SUEYOSHI; GOTO, 2018).

2.7 A ANÁLISE COM ORIENTAÇÃO PARA SAÍDAS (*OUTPUT ORIENTED*)

Até o subtítulo anterior foi realizada a interpretação e análise considerando a orientação para entradas. É possível que algum estudo específico necessite uma interpretação inversa: com uma orientação para saídas. Nesse caso, é objetivo determinar os índices de eficiência das DMU's em estudo buscando maximizar as saídas sem alterar o nível observado para as entradas.

De fato a resolução deste problema é possível empregando a análise envoltória de dados nos termos já estabelecidos. Por isso a compreensão do modelo de retorno constante por escala orientado para entradas é essencial. Considere então a DMU_o e que é necessário resolver o seguinte problema de programação linear

$$\max_{\zeta, \bar{\beta}^t} \zeta \quad (2.7.1)$$

sujeito as seguintes restrições:

$$\bar{x}_o - X\bar{\beta}^t \geq \bar{0} \quad (2.7.2)$$

$$Y\bar{\beta}^t - \zeta \bar{y}_o \geq \bar{0} \quad (2.7.3)$$

$$\bar{\beta}^t > \bar{0} \quad (2.7.4)$$

É possível encontrar uma solução ótima para as equações (2.7.1) até (2.7.4) diretamente das equações (2.5.8) a (2.5.11) relacionando as soluções na seguinte forma

$$\bar{\beta}^t = \zeta \bar{\lambda}^t, \quad \zeta = \frac{1}{\theta} \quad (2.7.5)$$

Perceba que com (2.7.5) e alguma manipulação algébrica é possível retornar as equações (2.5.8) a (2.5.11) definidas anteriormente no problema de orientação para entradas (COOPER; SEIFORD, 2007, p. 58). É possível a utilização da solução dada por (2.7.5)

evidenciando que solucionar o problema de programação com orientação para entradas permite encontrar a solução do problema de programação linear com orientação para saídas (*output oriented*).

2.8 ANÁLISE ENVOLTÓRIA COM RETORNOS VARIÁVEIS POR ESCALA

Iniciou-se essa revisão destacando o estudo de Charnes, Cooper e Rhodes que possui suas bases no trabalho realizado por Farrell. Esses estudos consolidaram DEA e atualmente este é o principal método não paramétrico (THANASSOULIS, 2001, p. 9) que possibilita realizar a avaliação de DMU's.

Contudo, modelos alternativos de determinação de índices de eficiência podem ser encontrados e que evoluíram do trabalho de Charnes, Cooper e Rhodes. Com ampla difusão, cita-se aquela que possibilita avaliar DMU's considerando retorno variável por escala (VRS – *variable return to scale*) proposto por Banker, Charnes e Cooper (1984).

Como dito, no modelo VRS a fronteira de eficiência é ajustada por alteração em algumas das restrições de resolução do problema de programação linear adicionando um novo componente ao conjunto solução. Assim, o conjunto solução P_v em que o índice subscrito faz referência ao caráter variável do retorno por escala é dado por

$$P_v = \left\{ (\bar{x}, \bar{y}) \mid \bar{x} \geq X\bar{\lambda}^t, \bar{y} \leq Y\bar{\lambda}^t, \bar{\lambda} \cdot \bar{e} = 1, \bar{\lambda}^t > \bar{0} \right\} \quad (2.8.1)$$

em que X é uma matriz de dimensão $s \times n$ com as s entradas das n DMU's avaliadas, Y é uma matriz de dimensão $r \times n$ com as r entradas das n DMU's avaliadas, $\bar{\lambda}^t$ é um vetor não negativo com n componentes transposto e \bar{e} é um vetor com valor 1 para as n componentes. Percebe-se em relação a (2.4.1) que apenas o produto escalar entre \bar{e} e o vetor $\bar{\lambda}$ é acrescido na descrição do conjunto de possibilidades de produção.

Isso acrescenta às condições de restrição do problema de programação linear descrito em (2.5.8) até (2.5.11) um novo fator. Numa interpretação literal, apenas uma nova condição é imposta na solução do problema para uma DMU_o. Este novo problema pode ser então escrito da seguinte forma orientado para entradas

$$\min_{\theta, \bar{\lambda}^t} \theta \quad (2.8.2)$$

sujeito as seguintes restrições:

$$\theta \bar{x}_o - X\bar{\lambda}^t \geq \bar{0} \quad (2.8.3)$$

$$Y\bar{\lambda}^t \geq \bar{y}_o \quad (2.8.4)$$

$$\bar{\lambda}^t > \bar{0} \quad (2.8.5)$$

$$\bar{\lambda} \cdot \bar{e} = 1 \quad (2.8.6)$$

Essa nova condição de restrição (2.8.6) impõe que a soma das componentes do vetor $\bar{\lambda}$ deve ser igual a 1. Apesar de simples, essa é uma alteração importante pois adequa a fronteira de eficiência. Uma consulta à bibliografia (COOPER; SEIFORD, 2007, p. 88; MOUSAVI-AVVAL et al., 2011) pode elucidar a ideia de ajuste da fronteira de eficiência com interpretações geométricas para exemplos de poucas variáveis.

No caso em que o objetivo é avaliar a DMU_o orientado as saídas a alteração nas condições de restrição são análogas as apresentadas em (2.8.3) até (2.8.6) e dessa forma as equações considerando retorno variável por escala são

$$\max_{\zeta, \bar{\beta}^t} \zeta \quad (2.8.7)$$

sujeito as seguintes restrições:

$$\bar{x}_o - X\bar{\beta}^t \geq \bar{0} \quad (2.8.8)$$

$$Y\bar{\beta}^t - \zeta \bar{y}_o \geq \bar{0} \quad (2.8.9)$$

$$\bar{\beta}^t > \bar{0} \quad (2.8.10)$$

$$\bar{\beta} \cdot \bar{e} = 1 \quad (2.8.11)$$

com a solução ótima na forma em que é descrita em (2.7.5). Com isso, tem-se elementos para compreender a DEA por um prisma da matemática bem como os elementos básicos permitindo a operacionalização e interpretações dos principais resultados obtidos em cálculos numéricos.

CAPÍTULO 3 – UTILIZAÇÃO DE DEA NA DETERMINAÇÃO DE EFICIÊNCIA DE PRODUÇÃO DE SOJA CONSIDERANDO EQUIVALENTES EM ENERGIA

3.1 MOTIVAÇÃO

A cultura da soja se consolida como a grande cultura tipo exportação e que gera receitas importantes para o agronegócio. Durante a safra 2016/2017 Brasil e Estados Unidos foram responsáveis por 83% das exportações de soja em todo o mundo. O Brasil foi líder exportando aproximadamente 63 milhões de toneladas. Mais de 85% desse total foi destinado à países asiáticos com destaque para a China que foi destino de 77% das exportações brasileiras (GALE; VALDES; ASH, 2019).

Assim, a utilização de métodos ou técnicas para determinação da eficiência em unidades de produção de grãos pode ser valiosa. Conhecer quais unidades de um determinado grupo tem maior ou menor eficiência permite ao gestor determinar quais os motivos destas diferenças. Ainda, pode-se customizar soluções em unidades com baixos índices de eficiência visando reduzir custos e melhorar a competitividade.

Diante do exposto, o objetivo do Capítulo 3 é determinar índices de eficiência de retorno constante por escala (CRS) utilizando a análise envoltória de dados (DEA) com orientação para entradas nos 27 produtores de soja considerando as safras 2014/2015, 2015/2016 e 2016/2017.

3.2 METODOLOGIA

Foram solicitados para produtores de soja a área cultivada e produção, produtos e quantidades utilizadas na cultura, máquinas empregadas no processo de produção e o consumo de combustível e finalmente mão-de-obra durante a semeadura, cultivo e colheita. O inventário foi determinado para cada produtor para as seguintes safras: 2014/2015, 2015/2016 e 2016/2017. Uma amostra de 27 produtores foi utilizada na elaboração das análises apresentadas neste Capítulo.

Para aplicar a DEA se considerou como entradas (*inputs*): adjuvantes, espalhantes, fungicidas, herbicidas, inseticidas, micronutrientes, nitrogênio (N), fósforo, (P_2O_5), potássio (K_2O), sementes, combustível, mão-de-obra e mecanização (agrupados os valores de trator 1, trator 2, colhedora, semeadora, aplicador de sólidos e pulverizador) em um total de 14 variáveis

de entrada. Como saída (*output*) foi analisado a produtividade da cultura da soja (uma variável de saída apenas). Todas convertidas em equivalentes em energia por unidade de área como descrito no Capítulo 1.

Os valores de eficiência foram determinados por DEA orientado para entradas no modelo CRS (eficiência técnica) com o pacote “rDEA” (SIMM; BESSTREMYANNAYA; SIMM, 2016) utilizando o R Software (R CORE TEAM, 2017). Para os cálculos de índices DEA de eficiência foi determinada a janela de análise nos termos que são propostos em outros trabalhos (CAMIOTO; MARIANO; REBELATTO, 2014; FERRAZ et al., 2018). Para este caso, o número total de safras é três, o tamanho da janela é duas safras e as janelas são duas: janela 1 composta pelas safras 2014/2015 e 2015/2016; janela 2 composta pelas safras 2015/2016 e 2016/2017. Tudo isso em acordo com as equações

$$p = \frac{k + 1}{2} \quad (3.2.1)$$

$$w = k - p + 1 \quad (3.2.2)$$

em que p é o tamanho da janela, k é o número total de períodos (safras) e w o número de janelas. A média das duas safras compõe a média da janela. A média das duas janelas compõe o índice de eficiência (DEA-CRS orientado para entradas) de cada um dos 27 produtores avaliados.

Com os índices de eficiências calculados as DMU’s representadas pelos 27 produtores foram classificadas em eficientes e ineficientes. Para realizar tal classificação foi estabelecido o valor de corte de 98% cujos valores superiores foram considerados eficientes e os valores inferiores ineficientes.

3.3 RESULTADOS E DISCUSSÃO

Segundo os resultados das médias dos cálculos dos índices DEA-CRS apresentados na Tabela 3.3.1 foram 9 as DMU’s que atingiram o valor de 100% de eficiência considerando o retorno constante por escala: as DMU’s 1, 4, 5, 12, 16, 20, 22, 23 e 27. Sem atingir o valor máximo mas acima dos 98% estão as DMU’s 3, 9, 17, 18. Portanto, as DMU’s classificadas como eficientes formam um grupo constituído de 13 unidades (produtores). Por outro lado, foram classificadas como ineficientes 14 unidades sendo elementos deste grupo as DMU’s 2, 6, 7, 8, 10, 11, 13, 14, 15, 19, 21, 24, 25 e 26.

Tabela 3.3.1 – Valores de eficiência técnica (CRS) orientado para entradas calculados nas três safras (S1, S2 e S3), estimados para as janelas (J-1 e J-2) e a Eficiência

Identificação	CRS-S1	CRS-S2	CRS-S3	J-1	J-2	Eficiência ¹	Classificação
DMU 1	100,00	100,00	100,00	100,00	100,00	100,00	Eficiente
DMU 2	74,56	72,08	100,00	73,32	86,04	79,68	Ineficiente
DMU 3	100,00	100,00	98,39	100,00	99,19	99,59	Eficiente
DMU 4	100,00	100,00	100,00	100,00	100,00	100,00	Eficiente
DMU 5	100,00	100,00	100,00	100,00	100,00	100,00	Eficiente
DMU 6	80,22	100,00	90,89	90,11	95,44	92,78	Ineficiente
DMU 7	100,00	95,28	97,34	97,64	96,31	96,97	Ineficiente
DMU 8	84,76	76,46	87,91	80,61	82,19	81,40	Ineficiente
DMU 9	95,27	100,00	100,00	97,63	100,00	98,81	Eficiente
DMU 10	85,27	77,00	74,82	81,13	75,91	78,52	Ineficiente
DMU 11	81,12	100,00	100,00	90,56	100,00	95,28	Ineficiente
DMU 12	100,00	100,00	100,00	100,00	100,00	100,00	Eficiente
DMU 13	97,15	91,30	92,62	94,22	91,96	93,09	Ineficiente
DMU 14	77,96	65,58	76,35	71,77	70,97	71,37	Ineficiente
DMU 15	100,00	90,81	99,26	95,41	95,03	95,22	Ineficiente
DMU 16	100,00	100,00	100,00	100,00	100,00	100,00	Eficiente
DMU 17	96,09	100,00	100,00	98,05	100,00	99,03	Eficiente
DMU 18	100,00	99,07	95,33	99,53	97,20	98,37	Eficiente
DMU 19	90,50	100,00	100,00	95,25	100,00	97,62	Ineficiente
DMU 20	100,00	100,00	100,00	100,00	100,00	100,00	Eficiente
DMU 21	83,41	80,74	92,43	82,07	86,59	84,33	Ineficiente
DMU 22	100,00	100,00	100,00	100,00	100,00	100,00	Eficiente
DMU 23	100,00	100,00	100,00	100,00	100,00	100,00	Eficiente
DMU 24	90,62	92,65	100,00	91,64	96,33	93,98	Ineficiente
DMU 25	82,81	100,00	96,17	91,41	98,09	94,75	Ineficiente
DMU 26	92,40	88,81	91,39	90,61	90,10	90,35	Ineficiente
DMU 27	100,00	100,00	100,00	100,00	100,00	100,00	Eficiente

Notas: ¹Eficiência estimada a partir da média dos valores encontrados para as janelas de observação (CAMIOTO; MARIANO; REBELATTO, 2014; FERRAZ et al., 2018).

Fonte: O autor.

Registre-se que os índices de eficiência 100% atingidos por algumas DMU's não significam que todas estas atingiram a eficiência plena. A única informação certa a respeito deste valor é que as que atingiram 100% estão sobre a fronteira de eficiência (COOPER; SEIFORD, 2007; MOUSAVI-AVVAL et al., 2011). Somente aquelas que não são expressas como combinações lineares de outras, ou seja, dependem apenas de seus fatores de produção para atingir os resultados de eficiência é que atingiram a eficiência plena como é apresentado e discutido com maior profundidade no Capítulo 2.

Sobre o critério de classificação das DMU's em eficientes e ineficientes foi arbitrado um valor adequado a realidade da demanda deste estudo. A literatura não determina especificamente que apenas as que atingem 100% serão eficientes. Exatamente ao contrário disso, como exposto no parágrafo anterior, é possível que uma DMU que atinge o valor de 100% necessite de adequação nos seus fatores de produção (quantidades das variáveis) para atingir a eficiência plena (COOPER; SEIFORD, 2007).

Portanto, considera-se que as DMU's entre 98% e 100% estão próximas da eficiência o suficiente para serem consideradas dentro do grupo de DMU's "eficientes". Ainda, estes valores de eficiência são índices relativos (COOPER; SEIFORD, 2007) comparando as DMU's umas com as outras a fim de determinar a eficiência intragrupo.

3.4 CONCLUSÕES

- As DMU's 1, 4, 5, 12, 16, 20, 22, 23 e 27 atingiram o valor de 100% de eficiência segundo o índice de eficiência técnica (CRS) orientado para entradas.
- Pelo mesmos índices as DMU's 2, 6, 7, 8, 10, 11, 13, 14, 15, 19, 21, 24, 25 e 26 obtiveram índices abaixo de 98% de eficiência técnica.
- Adotando o critério de corte de 98% são 13 DMU's eficientes (eficientes e muito próximas da eficiência) e 14 DMU's ineficientes dividindo as unidades em grupos próximos da equidade de elementos.
- O máximo valor de eficiência foi de 100% (DMU's 1, 4, 5, 12, 16, 20, 22, 23 e 27) e o menor valor foi de 71,37% (DMU 14).

CAPÍTULO 4 – ESTUDOS MULTIVARIADOS NA CARACTERIZAÇÃO E AVALIAÇÃO DE CLASSES DE EFICIÊNCIA OBTIDAS POR MEIO DE DEA

4.1 MOTIVAÇÃO

No Capítulo 2 foi introduzida a análise envoltória de dados e os principais modelos de cálculo de eficiências utilizando a técnica. Trata-se de uma metodologia onde a avaliação de diversas variáveis permite estabelecer uma ordem de eficiência entre vários indivíduos como se observa nos resultados apresentados no Capítulo 3.

Contudo, apesar de se revelar uma técnica poderosa e que pode ser explorada em profundidade no estudo de eficiência nos mais diversos setores agrícolas, é necessário ponderar a respeito de resultados anteriores realizando alguns questionamentos: existe diferença entre os elementos classificados como eficientes e não eficientes? Como estes se agrupam segundo as variáveis estudadas?

Estas são questões que serão abordadas no decorrer deste Capítulo. Antes, vale lembrar que a DEA não foi explorada ao máximo. Existem meios de estimar “o consumo ideal” de cada uma das variáveis analisadas para cada DMU. Esse seria o ajuste nos fatores de produção (BOGETOFT PEDERSEN; OTTO, 2011, p. 127; COOPER; SEIFORD, 2007, p. 44; SUEYOSHI; GOTO, 2018) abordado no Capítulo 2. Essa análise dos ajustes dos fatores de produção não será feita neste trabalho de dissertação.

Assim, o objetivo no presente capítulo é determinar se existe diferença significativa entre os grupos (eficientes e ineficientes) por meio de técnicas de análise multivariada de variância comparando os vetores de médias pelo método de Bonferroni, buscando uma função que discrimine as classes por meio da análise da função logística, caracterizando e avaliando agrupamentos por meio da análise de agrupamento por componentes principais.

4.2 METODOLOGIA

Para atingir o objetivo proposto nesta parte do trabalho foram realizados testes estatísticos multivariados (MANLY; ALBERTO, 2017). Contudo, para realização destes testes é necessário atender pressupostos visando resultados fidedignos ao final de cada um dos testes. Desta forma, os métodos e testes serão detalhados em separado nos subitens a seguir. Todos os

gráficos, testes e análises foram realizados no R Software (R CORE TEAM, 2017; WICKHAM et al., 2019b).

4.2.1 Análise de variância multivariada

A análise de variância multivariada (manova) foi realizada a fim de identificar se existe diferença entre os vetores de médias associados as classes “eficiente” e “ineficiente” determinadas através dos índices estimados por DEA-CRS orientado para entradas para cada uma das DMU’s.

A hipótese de nulidade (H_0) para a análise de variância multivariada é de que não existe diferença entre os vetores de médias das variáveis associados as diferentes classes de eficiência determinadas por meio dos índices estimados pela análise envoltória de dados. Ou seja,

$$H_0 : \vec{\mu}_1 = \vec{\mu}_2 \quad (4.2.1)$$

em que $\vec{\mu}_i$ refere-se aos vetores de médias das duas classes (ineficiente e eficiente).

Devido ao grande número de variáveis foi necessário agrupar algumas no intuito de reduzir esse número e assim viabilizar a análise. Na manova as variáveis usadas foram: 1.produtividade, 2.pesticidas (fungicidas, herbicidas, inseticidas, adjuvantes e espalhantes), 3.nutrientes (N, K₂O, P₂O₅ e micronutrientes), 4.sementes e 5.operacionais (mão-de-obra, máquinas e combustível).

Todas as variáveis têm unidades de equivalentes em energia por unidade de área ($MJ ha^{-1}$) possibilitando seu agrupamento. Este agrupamento segue a padronização apresentada na Tabela 1.3.1 com exceção da energia química da água que, como apontado na metodologia do Capítulo 1, foi considerado como um valor médio para todas as DMU’s (produtores) não tendo sentido incluí-la na manova.

Os pacotes “dplyr” e “readxl” (WICKHAM et al., 2019a, 2019c) foram utilizados na leitura e manuseio de dados. O pacote “car” (FOX et al., 2019) para execução da manova utilizando o teste de Wilks (MANLY; ALBERTO, 2017). A análise de homocedasticidade foi realizada utilizando o pacote estatístico “biotools” (SILVA, 2017) pelo teste M de Box. A análise da distribuição dos resíduos foi realizada pelos testes de Shapiro-Wilk no caso univariado e pelo teste de Mardia no caso multivariado por meio da função “mvn” no pacote estatístico “MVN” (KORKMAZ; GOKSULUK; ZARARSIZ, 2014).

Para atender os critérios de normalidade dos resíduos uni e multivariados bem como a homocedasticidade das matrizes de covariâncias foram necessários dois procedimentos. O primeiro diz respeito a padronização dos dados pela amplitude de acordo com a seguinte equação

$$y_{i(j)} = \frac{x_{i(j)} - \min x_{(j)}}{\max x_{(j)} - \min x_{(j)}} \quad (4.2.2)$$

em que y denota o valor padronizado, x faz referência aos valores originais da variável j , i é o índice do número de elementos da variável variando de 1 até 81 (27 DMU's em 3 safras) e j faz referência as 5 variáveis descritas anteriormente. O segundo procedimento foi a exclusão de pontos discrepantes (*outliers*). Com isso, foram utilizados dados de 28 DMU's classificadas como eficientes e 32 DMU's como ineficientes bem como seus respectivos vetores já descrito no terceiro parágrafo deste tópico.

O contraste entre os vetores de médias das classes eficiente e ineficiente foi avaliado utilizando o método de Bonferroni para determinar os intervalos simultâneos de confiança ao nível de 95% para cada uma das cinco variáveis segundo a seguinte relação

$$\left(\bar{X}_{k(j)} - \bar{X}_{l(j)}\right) \pm t_{n-g, 1-\alpha/(pg(g-1))} \sqrt{\left(\frac{1}{n_k} + \frac{1}{n_l}\right) \cdot \frac{W_{jj}}{(n-g)}} \quad (4.2.3)$$

em que n_k e n_l são os número de elementos das classes eficientes e ineficientes, n é o número total de elementos das duas classes, g é o número de classes (duas apenas), W_{jj} são as componentes da diagonal principal da matriz da soma de quadrados interna estimados na manova, p é o número de variáveis, t é o valor da estatística t -Student nas condições especificadas, α é a significância desejada, \bar{X} é a média dos k e l elementos na variável j e o sinal \pm designando os limites superior e inferior.

4.2.2 Regressão logística discriminante

Uma abordagem utilizando a regressão logística foi aplicada no intuito de obter uma função para discriminar os grupos “eficientes” e “ineficientes”. Chama-se de função logística aquela que estima uma probabilidade de sucesso em função de quaisquer variáveis importantes

à determinado estudo (MANLY; ALBERTO, 2017). Para o caso em estudo neste trabalho o sucesso é dito como uma DMU que é classificada como eficiente.

Tomando o mesmo conjunto de dados utilizados na análise de manova, as classes de eficiência foram convertidas em um binário (eficientes =1 e ineficientes = 0). Uma regressão logística foi realizada para 60 DMU's (28 eficientes e 32 ineficientes). As variáveis independentes foram produtividade, pesticidas, nutrientes, sementes e operacionais. A variável dependente é dada pelo binário envolvendo as classificações.

Um teste de qui-quadrado foi realizado para avaliar quanto um caso eficiente está relacionado às variáveis anteriormente descritas. Em seguida foram estimados a constante e os coeficientes que multiplicam as variáveis dependentes no modelo linear geral. Estes coeficientes também foram analisados pelo teste de qui-quadrado a fim de verificar quais variáveis são significativamente importantes na estimativa da probabilidade (e consequentemente da discriminação) entre as DMU's eficientes e ineficientes.

A efetividade da regressão logística em discriminar elementos dos dois grupos (eficientes e ineficientes) em função das variáveis foi avaliada através de um gráfico onde são apresentadas as probabilidades encontradas. Todos os cálculos foram realizados no R Software (R CORE TEAM, 2017) utilizando as funções base do sistema como sugere a literatura (MANLY; ALBERTO, 2017) sem necessidade de pacotes estatísticos além dos já citados anteriormente neste Capítulo.

4.2.3 Análise de agrupamento por componentes principais

Uma análise de agrupamento por componentes principais (KASSAMBARA, 2017) foi efetuada com intenção de verificar relações de “proximidade” entre as DMU's. A análise de agrupamento é uma técnica importante para uma análise sobre proximidades de elementos em banco de dados multivariados. A abordagem por análise de agrupamento por componentes principais permite combinar os métodos de análise de componentes principais e de agrupamento em análises multivariadas (HUSSON; JOSSE; PAGES, 2010).

As análises foram realizadas utilizando as variáveis independentes produtividade, pesticidas, nutrientes, sementes e operacionais. Todas as variáveis tiveram os dados padronizados pela amplitude. Como as análises de agrupamento e de componentes principais têm caráter exploratório e não demandam repetições (MANLY; ALBERTO, 2017) o

agrupamento por componentes principais foi realizado utilizando os valores médios das três safras para cada DMU.

Para executar o agrupamento por componentes principais no R Software foi seguido o algoritmo (KASSAMBARA, 2017): 1.análise de componentes principais; 2.escolha do número de grupos baseado na soma de quadrados interna; 3.análise de agrupamento e dendrograma; 4.análise de agrupamento no plano das componentes principais.

O método de agrupamento utilizado foi o da distância média entre grupos (MANLY; ALBERTO, 2017). O número de agrupamentos foi definido por meio da soma de quadrados interna usando o algoritmo “*k-means*” (JAIN, 2010). No manuseio de dados foram usados pacotes disponíveis no R Software (WICKHAM et al., 2019a, 2019c). A análise de componentes principais, análise de agrupamento e o agrupamento por componentes principais foram executados utilizando as funções nos pacotes “factoextra” e “FactoMineR” (HUSSON et al., 2019; KASSAMBARA; MUNDT, 2017).

4.3 RESULTADOS E DISCUSSÃO

4.3.1 Análise de variância multivariada e regressão logística

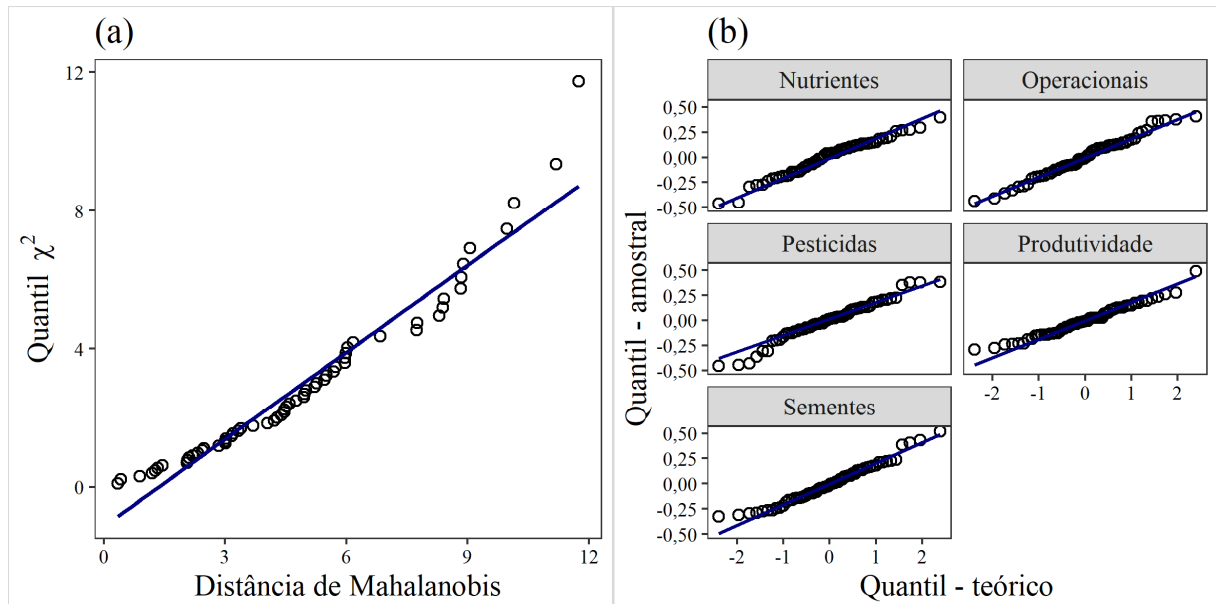
A análise de variância foi realizada com a intenção de verificar se existe diferença entre os vetores de médias das 5 variáveis associados às classes eficiente e ineficiente. Pelos resultados do teste M de Box as matrizes de covariância são consideradas homogêneas ($p = 0,6177$). A distribuição dos resíduos univariados é considerada normal segundo os resultados dos testes de Shapiro-Wilk ($p > 0,05$). O mesmo se observa para a distribuição dos resíduos multivariados segundo o teste de Mardia ($p > 0,05$ para curtose e assimetria). A distribuição dos resíduos pode ser verificada na Figura 4.3.1.

O quadro resumo da análise de variância multivariada (Tabela 4.3.1) aponta resultado significativo com valor de $p < 0,05$ utilizando o teste de Wilks. Portanto, aceita-se a hipótese alternativa em detrimento da hipótese de nulidade. Existe diferença significativa ao nível de 5% entre o contraste envolvendo os vetores de médias das classes (eficiente e ineficiente).

Um ponto importante a ressaltar é que o resultado de diferença entre os vetores de médias associados com as classes fundamenta a decisão tomada de realizar o corte entre eficientes e ineficientes ao valor de 98%. Com esse resultado, é possível afirmar que, mesmo

que algumas DMU's não tenham atingido o valor de 100%, classificá-las como eficientes não alterou significativamente a média das variáveis.

Figura 4.3.1 – Resíduos multivariados (a) segundo a distribuição de qui-quadrado (χ^2) versus distância de Mahalanobis e distribuição de resíduos univariados (b) segundo os quantis amostrais e teóricos.



Fonte: O autor.

Assim, as DMU's que não tenham atingido a eficiência máxima não se distanciam das que atingem 100% de tal maneira a serem consideradas pontos discrepantes e/ou alterem os resultados da análise de variância multivariada. Ainda que 21 vetores tenham sido excluídos para adequação dos pressupostos de normalidade e homocedasticidade apenas um vetor de médias era de DMU's que tiveram eficiência entre 98% e 100%.

Tabela 4.3.1 – Resumo da análise de variância multivariada

Fonte de variação	g. l.	Λ de Wilks	F Estimado	g. l. Numerador	g. l. Denominador	p (>F)
Classes	1	0,8093	2,5449	5	54	0,03861*

Notas: *Significativo ao nível de 5%; Λ – Lambda; g. l. – grau de liberdade; F – estatística de Fischer; p – valor de probabilidade.

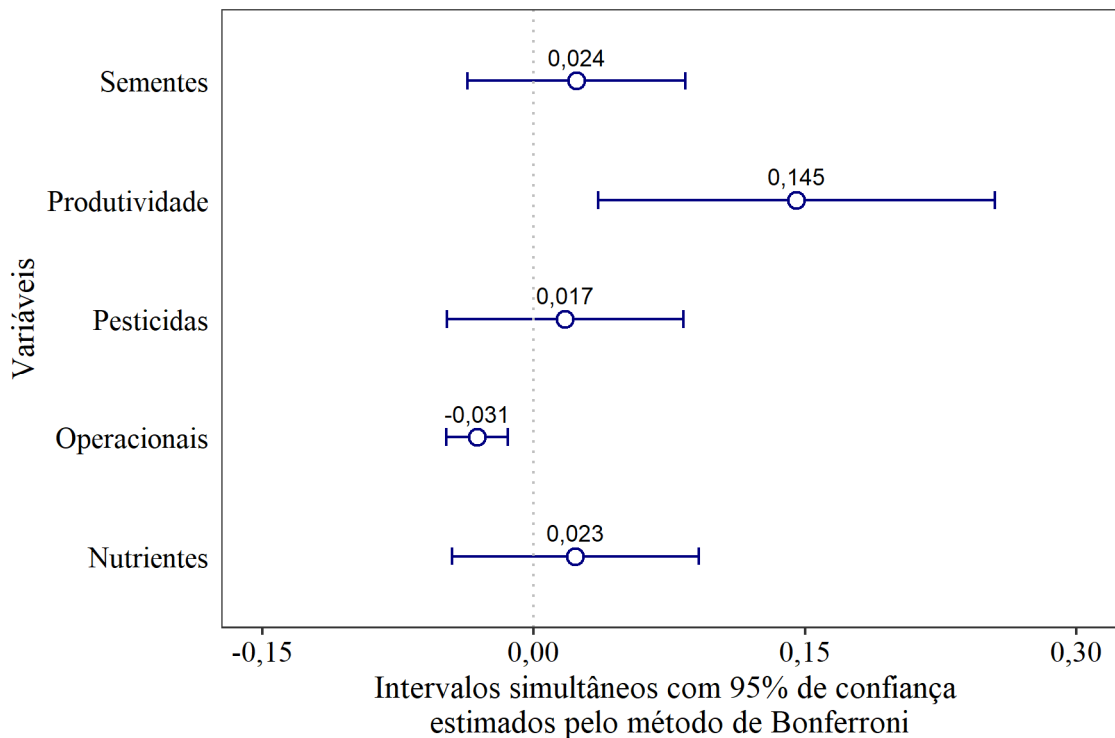
Fonte: O autor.

Buscando verificar maiores detalhes a respeito do contraste entre os vetores de médias das DMU's eficientes contra as ineficientes foi estimado pelo método de Bonferroni o intervalo simultâneo de confiança das variáveis utilizadas na manova. Os detalhes são apresentados na Figura 4.3.2. Como se observa, as variáveis operacionais e produtividade excluem o zero pelos

intervalos de confiança. Logo, a diferença entre as médias das duas classes nestas duas variáveis é estatisticamente diferente. Vale lembrar que os valores são os padronizados, por isso uma ordem de grandeza tão pequena.

Portanto, considerando os resultados da Figura 4.3.2 percebe-se por estes resultados que a diferença entre as classes eficiente e ineficiente parece estar associada as quantidades de energia utilizadas por unidade de área nas operações e a quantidade de energia produzida por unidade de área (produtividade). Tais constatações acabam não sendo expostas pelos resultados dos índices de eficiência estimados pela análise envoltória de dados apresentados no capítulo anterior na Tabela 3.3.1.

Figura 4.3.2 – Intervalos simultâneos com pelo menos 95% de confiança estimados pelo método de Bonferroni para a diferença entre os vetores de médias das classes eficiente e ineficiente nas variáveis estudadas.



Fonte: O autor.

Diante da indicação de diferenças significativa entre os vetores de médias na manova foi realizada uma regressão logística discriminante a fim de avaliar as variáveis pelo teste de qui-quadrado e verificar se, por meio destas variáveis, seria possível definir função probabilística que separa os dois grupos como sugere a literatura (MANLY; ALBERTO, 2017). A Tabela 4.3.2 apresenta os resultados das estimativas dos coeficientes que multiplicam as variáveis, o erro padrão, estatística de qui-quadrado (χ^2) e o valor-p de probabilidade.

Tabela 4.3.2 – Resultados da função logística discriminante

Variável	Coefficiente	Erro padrão	χ^2	p ($> \chi^2$)
Constante	-2,481	1,608	----	----
Produtividade	6,688	2,265	8,720	0,0030**
Pesticidas	-0,797	1,803	0,195	0,6590
Nutrientes	-1,001	2,178	0,211	0,6460
Sementes	0,155	1,747	0,008	0,9290
Operacionais	-0,940	1,476	0,406	0,5240

Notas: **Significativo ao nível de 0,01 de probabilidade.

Fonte: O autor.

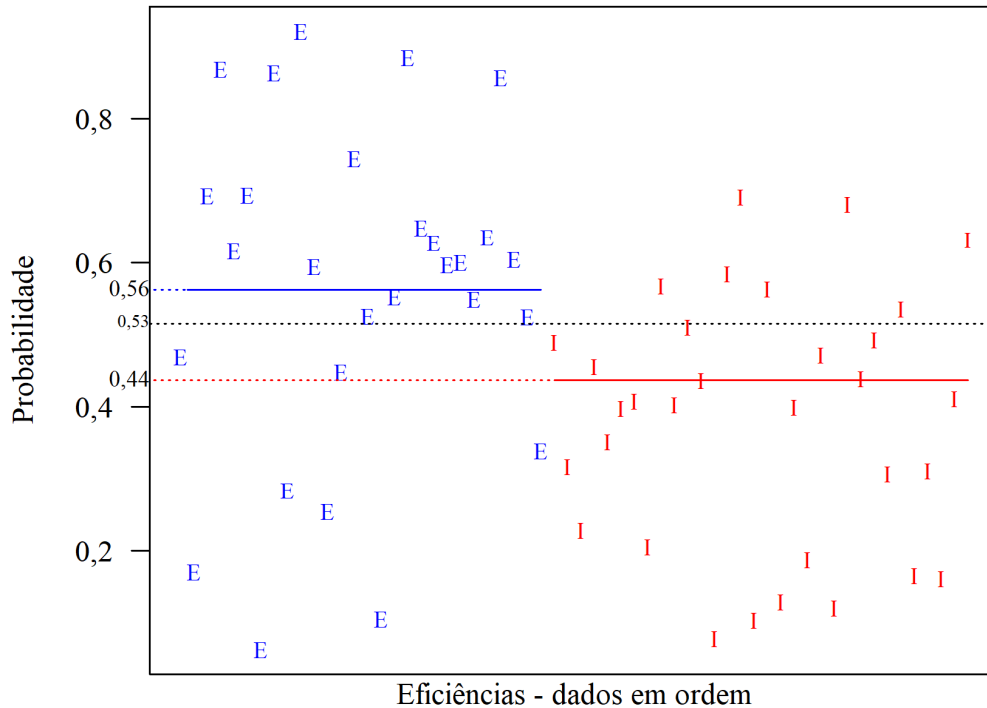
Segundo os resultados da função logística discriminante apenas a variável produtividade foi significativa pelo teste de qui-quadrado. Tal resultado confirma a importância da variável produtividade apresentado anteriormente (Figura 4.3.2). Contudo, por esse teste a variável que retrata o consumo de energia por unidade de área com as operações não apresentou resultado significativo.

Isto não significa que os gastos operacionais deixaram de ser importante para distinção entre os grupos e sim que seu peso na determinação de um valor de probabilidade por meio da função logística para as DMU's serem enquadradas como eficiente ou ineficiente não é significativo. Os valores de probabilidades e a disposição dos grupos das duas classes podem ser visualizados na Figura 4.3.3 a seguir.

Na Figura 4.3.3 se verifica os resultados de probabilidades encontrados pela função logística discriminante. A linha pontilhada preta marcando 0,53 de probabilidade representa o valor arbitrário (MANLY; ALBERTO, 2017) definido em função do número de elementos nas classes eficiente e ineficiente. Caso o número de elementos fosse o mesmo essa probabilidade seria de 0,5. No entanto, no caso em estudo tratam-se de 28 valores de probabilidade da classe eficiente em um total de 60 (0,47). Logo a probabilidade de uma ineficiente é de 0,53.

As linhas azul (probabilidade de 0,56) e vermelha (probabilidade de 0,44) representam as médias de probabilidades de todos as estimativas de E e I respectivamente (Figura 4.3.3). Em média, a função logística foi capaz de distinguir entre as classes eficiente e ineficiente. Contudo, esta distinção acaba sendo prejudicada em função da má classificação de alguns elementos. Para a classe eficiente foram 8 probabilidades estimadas abaixo de 0,53 o que representa cerca de 71% de acertos. Por outro lado cerca de 78% das probabilidades representam a classe ineficiente adequadamente. No total dos 60 valores de probabilidade estimados foram 75% de acertos.

Figura 4.3.3 – Valores de probabilidades estimados pela função logística para as amostras de 28 DMU's eficientes (representadas pela cor azul e letras E) e das 32 DMU's ineficientes (representadas em vermelho e letras I). Barras azul e vermelha sólidas representam a probabilidade média das classes.



Fonte: O autor.

4.3.2 Análise de agrupamento por componentes principais

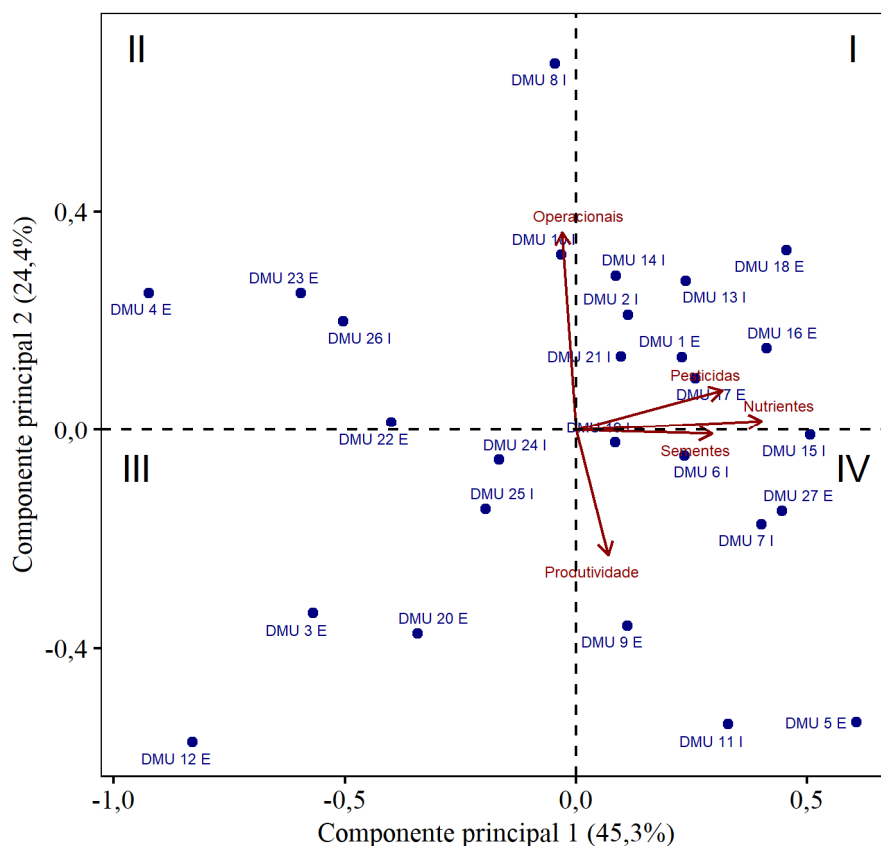
É de se supor que exista uma relação de proximidade entre algumas DMU's de forma que as DMU's eficientes tenham características em comum. O mesmo raciocínio vale para DMU's ineficientes. Para verificar a existência desse padrão entre as classes de eficiência das DMU's foi realizada a análise de agrupamento por componentes principais.

As duas componentes principais mais representativas (Figura 4.3.4) acumulam 69,7% da variação dos dados. O quadrante I se caracteriza pelos maiores consumos de energia com pesticidas e nutrientes (DMU's 1E, 2I, 13I, 14I, 16E, 17E, 18E, 21I em que E e I as denotam como eficientes e ineficientes em acordo com o critério apresentado no capítulo anterior). No quadrante II predominam DMU's com maiores dispêndios de energia operacional e baixa produtividade (DMU's 4E, 8I, 10I, 22E, 23E e 26I). O quadrante III revela as DMU's que possuem baixos consumos de equivalentes em energia com pesticidas, nutrientes e sementes mantendo bons índices de equivalentes em energia de produtividade (DMU's 3E, 12E, 20E, 24I e 25I). No quadrante IV se encontram DMU's com altos valores de produtividade em

equivalente de energia, alto consumo de equivalentes em energia com sementes e baixo consumo de energia com operações (DMU's 5E, 6I, 7I, 9E, 11I, 15I, 19I e 27E).

Em resumo o quadrante I pode ser dito como de alto custo energético com pesticidas e nutrientes. O quadrante II alto custo de operações e baixa produtividade em equivalentes de energia. O quadrante III baixo custo de pesticidas, nutrientes e sementes com bons índices de produtividade. Quadrante IV altos índices de produtividade com alto custo de sementes considerando os respectivos equivalentes em energia.

Figura 4.3.4 – Biplot com as duas componentes principais que representam 69,7% da variabilidade dos dados.



Notas: Vetores vermelhos representam a projeção das variáveis segundo componentes principais. As letras E (eficiente) e I (Ineficiente) associadas aos nomes das DMU's indicam como estão classificadas quanto à eficiência.

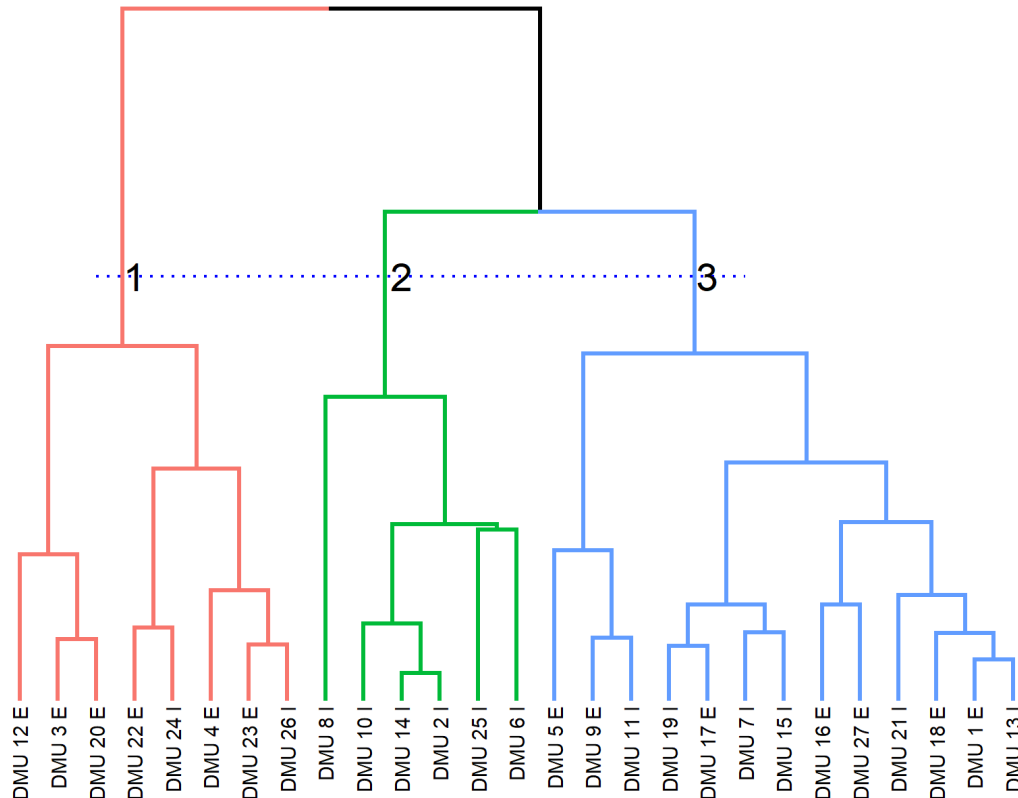
Fonte: O autor.

Nota-se pela Figura 4.3.4 que as DMU's 24 e 25 classificadas como ineficientes estão no quadrante III. Tal constatação deixa claro que o total de 69,7% da variabilidade total explicada pelas componentes principais não é suficiente para a completa caracterização de resultados. Outro ponto a ser questionado é: existem grupos que podem ser visualizados como comuns no contexto das componentes principais? Pois, se nos quadrantes existem DMU's eficientes e ineficientes é de se supor que elas estejam "próximas" de alguma forma. Tentando

identificar esta relação de proximidade é apresentado na Figura 4.3.5 os agrupamentos realizados pela distâncias entre as médias dos grupos. Os três agrupamentos definidos em função da soma de quadrados interna (inércia) são apresentados na

Figura 4.3.6.

Figura 4.3.5 – Dendrograma de agrupamento das DMU's avaliadas pelo método da média.



Nota: As letras E (eficiente) e I (Ineficiente) associadas aos nomes das DMU's indicam como estão classificadas quanto à eficiência. Os números de 1 até 3 indicam o número dos agrupamentos.

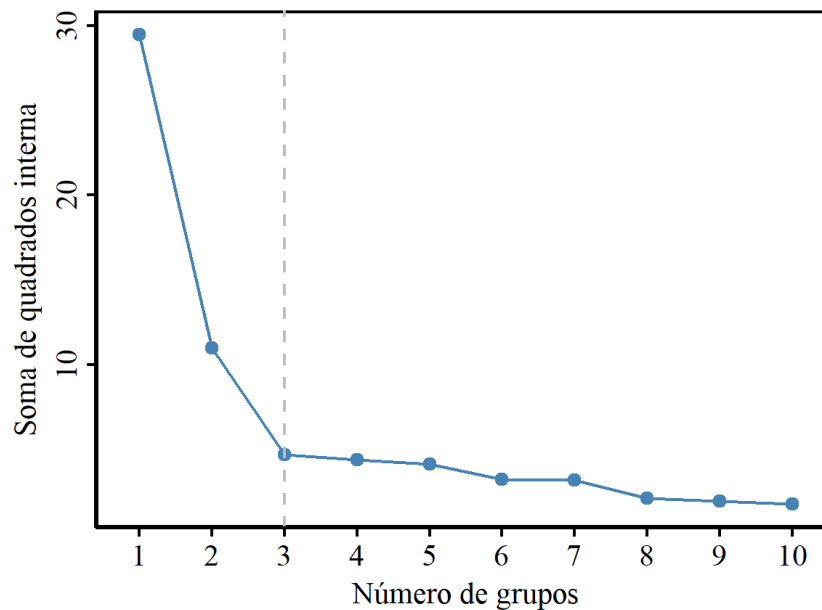
Fonte: O autor.

Acerca dos agrupamentos apresentados na Figura 4.3.5 é possível verificar padrões em relação as classes de eficiência. Note que no agrupamento 1 (vermelho) predominam as DMU's eficientes sendo exceções as DMU's 24 e 26. No agrupamento 2 (verde) as DMU's são todas classificadas como ineficientes. No agrupamento 3 (azul) existe uma diversidade de DMU's classificadas como eficientes e ineficientes. As médias destes agrupamentos serão comparadas utilizando as três componentes principais (Figura 4.3.7) que juntas totalizam cerca de 89% da variabilidade e criam um ambiente mais favorável as interpretações de cada grupo.

Observando os planos formados pelas três componentes principais na coluna à direita da Figura 4.3.7 é possível constatar que existe uma forte correlação das variáveis pesticidas, nutrientes e sementes com a componente principal 1 (gráficos da coluna à direita) sendo

interpretada com uma medida de consumo destas variáveis. A componente principal 2 pode ser interpretada como um contraste entre as variáveis produtividade e operacionais. A componente principal 3 é interpretada como medidas das variáveis produtividade e operacionais em contraste com o consumo de pesticidas. As três componentes principais juntas explicam 89,4% da variabilidade total dos dados.

Figura 4.3.6 – Screeplot da soma de quadrados interna em função do número de grupos.



Fonte: O autor.

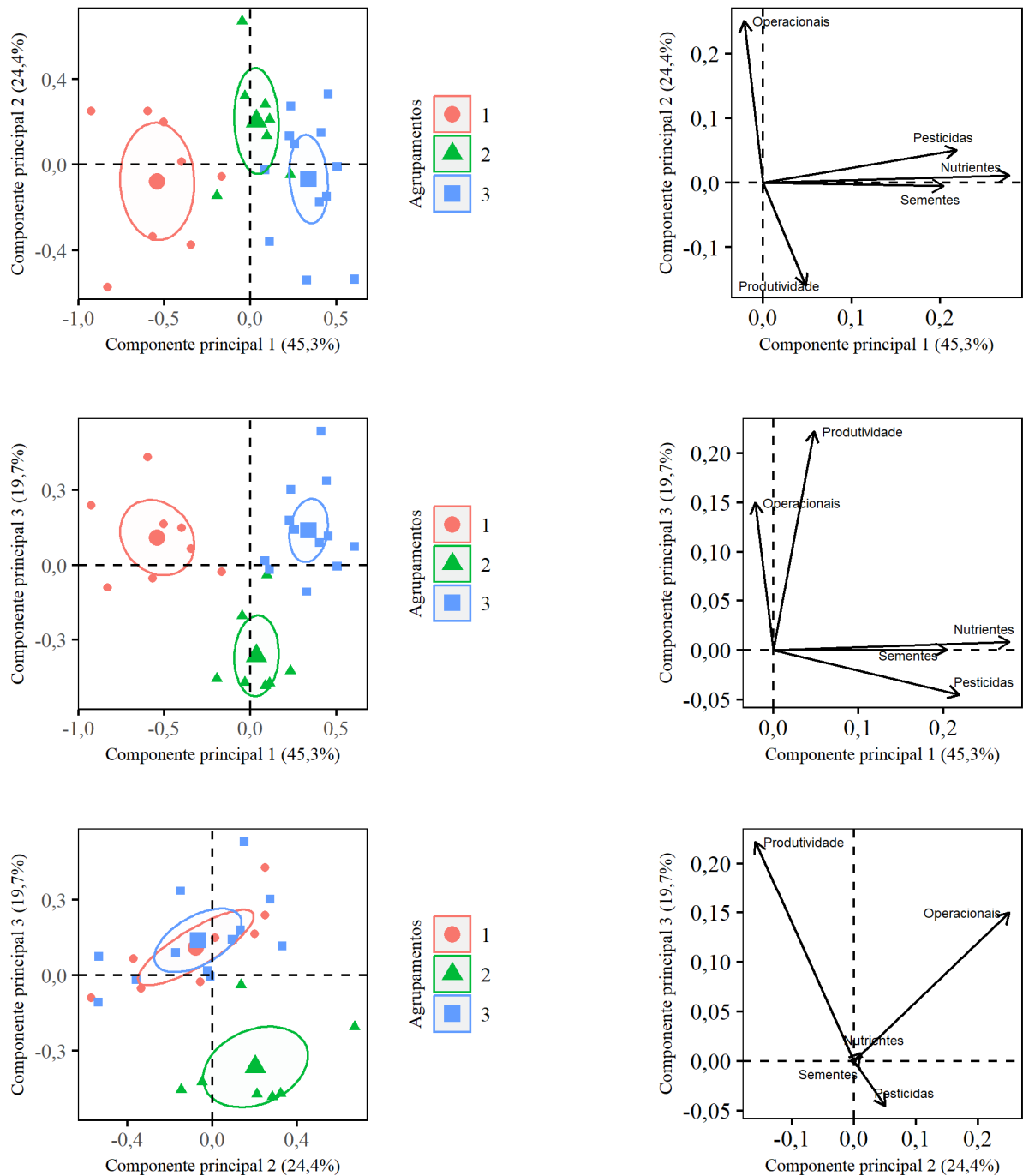
Nos gráficos na coluna à esquerda da Figura 4.3.7 verificam-se as projeções de pontos e da média dos agrupamentos (nos centros das elipses de confiança). O agrupamento 1 está localizado nos quadrantes onde permite-se dizer que este grupo em média possui bons índices de produtividade com baixos custos com pesticidas, nutrientes e sementes. Perceba que este não possui os melhores índices de produtividade em média.

As unidades ineficientes do agrupamento 2 (Figura 4.3.7). Os elementos deste grupo são os que apresentam altos dispêndios de energia nas variáveis pesticidas e operacionais com os menores valores de produtividade de equivalentes em energia por unidade de área. Esse é exatamente o comportamento que se espera em média de um grupo composto por unidades ineficientes: altos custos e baixas produções.

O agrupamento 3 é o que possui as DMU's com as melhores produtividades. Confirma-se pelos pontos azuis extremos no quarto quadrante do gráfico da primeira linha a esquerda (Figura

4.3.7). Note que nesta mesma coluna na segunda linha os pontos azuis são os pontos extremos na direção do vetor produtividade. O mesmo se observa no gráfico da terceira linha no seu segundo quadrante.

Figura 4.3.7 – Projeção dos agrupamentos 1, 2 e 3 no plano das componentes principais. À esquerda os planos com as médias (pontos maiores) no centro das elipses com 0,95 de confiança da média e à direita projeção das variáveis.



Fonte: O autor.

Assim, fica evidente que apesar do bom desempenho médio em produtividade das DMU's do agrupamento 3, muitas destas DMU's possuem gastos consideráveis com pesticidas, nutrientes, sementes e operacionais. Quando estes gastos de algumas DMU's não são tão grandes suas produtividades ficam abaixo da média do grupo. Isso explica a razão da heterogeneidade de classes de eficiência neste agrupamento. Em resumo, alguns elementos do grupo produzem bem gastando muito e outros elementos gastam pouco mas não obtêm bons índices de produtividade.

Diante de tais constatações fica claro que o paradigma de manejo nas DMU's do agrupamento 2 precisa de revisão visando melhorar seu desempenho em relação aos agrupamentos 1 e 3. Nesta mesma conjuntura verifica-se que existem adequações a serem realizadas pelos elementos do agrupamento 3 em função da heterogeneidade na classificação de eficiências. É necessário identificar quais DMU's produzem bem mais gastam muito e, naturalmente, quais são os fatores que levam a tal comportamento.

Com relação ao agrupamento 1 é necessário identificar quais as (boas) práticas que estão fazendo com que os elementos deste grupo fazem que em média ele se destaque dos demais. Uma análise dos dados permite identificar que existe uma racionalidade na utilização de equivalentes em energia ligados aos pesticidas, nutrientes e sementes. Para se ter dimensão disto as DMU's do agrupamento 3 precisaram em média de 76,8%, 74,5% e 57,6% mais do equivalente em energia respectivamente de pesticidas, nutrientes e sementes para produzir em média 3,44% mais que as DMU's do agrupamento 1. Já o agrupamento 2 necessitou em média de 75,7%, 55,2% e 37,6% mais do equivalente em energia respectivamente de pesticidas, nutrientes e sementes para produzir em média -19,9% que as DMU's do agrupamento 1.

Desta forma, é coerente afirmar que as boas práticas que fazem com que o agrupamento 1 se destaque na média está relacionado a utilização de defensivos agrícolas (pesticidas) utilizados com parcimônia, utilização racional de adubos e fertilizantes no manejo nutricional da cultura e utilização de populações de plantas adequadas resultando em economia de sementes no instante da semeadura.

As análises multivariadas apresentadas permitem crer que o aproveitamento da análise envoltória de dados em sistema de produção agrícola pode ser uma boa ferramenta de gestão. Contudo, parece que o pleno alcance da visualização de resultados acontece quando existe uma combinação entre a análise envoltória e outras técnicas multivariadas. Talvez essa multidisciplinaridade seja a alternativa mais adequada para a tomada de decisões assertivas a

respeito de qualquer modelo de gestão em qualquer setor de produção indo ao encontro de uma visão sistêmica dos processos.

4.4 CONCLUSÕES

- Os vetores de médias associados as classes “eficientes” e “ineficientes” apresentam diferença significativa pelo teste Λ de Wilks ao nível de 5% segundo a análise multivariada de variância. Intervalos simultâneos com pelo menos 95% de confiança estimados pelo método de Bonferroni para a diferença entre os vetores de médias indicam que as médias das variáveis “operacionais” e “produtividade” são diferentes.
- A regressão logística estimada foi capaz de realizar a distinção entre as classes “eficientes” e “ineficientes” em mais de 75% do total dos casos avaliados. Dos coeficientes estimados para regressão logística apenas aquele relacionado a produtividade foi significativo ao nível de 1% pelo teste de qui-quadrado.
- Na análise de agrupamentos por componentes principais foram identificados 3 agrupamentos considerando a distância média.
- As DMU's do agrupamento 1 são classificadas na maioria como eficientes. Este grupo é caracterizado pelo equilíbrio entre custo e produtividade de energia por unidade de área.
- O agrupamento 2 é composto de elementos que apresentam altos dispêndios de energia nas variáveis pesticidas e operacionais com os menores valores de produtividade de equivalentes em energia por unidade de área.
- O agrupamento 3 possui uma heterogeneidade contendo DMU's eficientes e ineficientes. As DMU's deste grupo possuem altas (baixas) produtividades associadas a altos (baixos) dispêndios de energia.

CONSIDERAÇÕES FINAIS

Este trabalho revela que o estudo de sistemas agrícolas considerando os equivalentes em energia pode ser realizado com intuito de verificar a racionalidade do sistema de produção. Novos estudos considerando as metodologias apresentadas e em acordo com outros autores constroem material bibliográfico que serve de referencial para avaliar como evoluem estes sistemas em função das novas técnicas e impactos de novas tecnologias.

Trabalhos que considerem outras regiões produtoras de soja podem contribuir para verificação de padrões regionais de consumo de energia no manejo da cultura da soja. Ainda, as metodologias empregadas bem como alguns índices expostos neste trabalho podem ser imediatamente utilizados na avaliações de outras culturas como milho, trigo, algodão, etc. Um trabalho mais audacioso seria a possibilidade de avaliar sistemas de cultivos como a sucessão soja/milho segunda safra (tão importante em contexto nacional), soja/trigo, integrações (lavoura-pecuária, lavoura-pecuária-floresta).

A DEA pode ser uma ferramenta auxiliar de avaliação e gestão não só do ponto de vista de racionalização do sistema de produção. Esta pode ser empregada nos mais diversos setores agrícolas que demandam estimativas de eficiências relativas entre elementos de um grupo. Sistemas de armazenagem de grãos, gestão de mão-de-obra, eficiência de máquinas agrícolas de usinas de açúcar/álcool, etc.

Contudo, apesar da DEA se mostrar uma ferramenta poderosa, é fundamental perceber que a tomada de decisão utilizando dados multivariados deve obrigatoriamente passar por análises estatísticas multivariadas. Os grandes bancos de dados possuem necessariamente características e padrões que podem ser identificados e melhor analisados se fazendo o uso de análise de agrupamento, análise de componentes principais, análises discriminantes, etc. que na essência possuem elementos matemáticos e/ou estatísticos.

Bancos de dados públicos disponíveis na internet ou que podem ser solicitados aos órgãos públicos de setores ligados a agricultura podem subsidiar novos trabalhos que envolvam as técnicas de análise multivariada utilizadas nesta dissertação bem como outras não abordadas. Por exemplo, no estado do Paraná, é fato que um sistema digitalizado de registro de receitas agronômicas funciona já algum tempo. Este deve ser um rico banco de dados para utilização de técnicas de análise multivariada.

REFERÊNCIAS

- BANKER, R. D.; CHARNES, A.; COOPER, W. W. Some models for estimating technical and scale inefficiencies in data envelopment analysis. **Management science**, [s. l.], v. 30, n. 9, p. 1078–1092, 1984.
- BATISTA, T. M. de V.; BEZERRA NETO, F.; PORTO, V. C. N.; BARROS JÚNIOR, A. P.; SILVA, Í. N.; SILVA, M. L.; LIMA, J. S. S.; OLIVEIRA, E. Q. Bio-agro-economic returns from carrot and salad rocket as intercrops using hairy woodrose as green manure in a semi-arid region of Brazil. **Ecological Indicators**, [s. l.], v. 67, p. 458–465, 2016.
- BEHESHTI TABAR, I.; KEYHANI, A.; RAFIEE, S. Energy balance in Iran's agronomy (1990–2006). **Renewable and Sustainable Energy Reviews**, [s. l.], v. 14, n. 2, p. 849–855, 2010.
- BOGETOFT PEDERSEN, P.; OTTO, L. S. **Benchmarking with DEA, SFA, and R**. 1. ed. New York, NY: Springer, 2011.
- BRITO, R. P.; BRITO, L. A. L. Vantagem competitiva e sua relação com o desempenho: uma abordagem baseada em valor. **Revista de Administração Contemporânea**, [s. l.], v. 16, n. 3, p. 360–380, 2012.
- CAMIOTO, F. de C.; MARIANO, E. B.; REBELATTO, D. A. do N. Efficiency in Brazil's industrial sectors in terms of energy and sustainable development. **Environmental Science & Policy**, [s. l.], v. 37, p. 50–60, 2014.
- CAMIOTO, F. de C.; REBELATTO, D. A. do N.; ROCHA, R. T. Análise da eficiência energética nos países do BRICS: um estudo envolvendo a Análise por Envoltória de Dados. **Gestão & Produção**, [s. l.], v. 23, n. 1, p. 192–203, 2015.
- CAMPOS, L. M. de S.; MELO, D. A. Indicadores de desempenho dos Sistemas de Gestão Ambiental (SGA): uma pesquisa teórica. **Production**, [s. l.], v. 18, n. 3, p. 540–555, 2008.
- CANAKCI, M.; TOPAKCI, M.; AKINCI, I.; OZMERZI, A. Energy use pattern of some field crops and vegetable production: Case study for Antalya Region, Turkey. **Energy Conversion and Management**, [s. l.], v. 46, n. 4, p. 655–666, 2005.
- CARARETO, E. S.; JAYME, G.; TAVARES, M. P. Z. Gestão Estratégica de Custos: custos na tomada de decisão. **Revista de Economia da UEG**, [s. l.], v. 2, n. 2, p. 24, 2006.
- CARVALHO, I. R.; KORCELSKI, C.; PELISSARI, G.; HANUS, A.; ROSA, G. Demanda hídrica das culturas de interesse agrônômico. **Enciclopédia biosfera, Centro Científico Conhecer-Goiânia**, [s. l.], v. 9, n. 17, p. 969, 2013.
- CHARNES, A.; COOPER, W. W.; RHODES, E. Measuring the efficiency of decision making units. **European journal of operational research**, [s. l.], v. 2, n. 6, p. 429–444, 1978.
- CHEN, G. Q.; JIANG, M. M.; CHEN, B.; YANG, Z. F.; LIN, C. Emergy analysis of Chinese agriculture. **Agriculture, Ecosystems & Environment**, [s. l.], v. 115, n. 1–4, p. 161–173, 2006.

CHEN, Y.; GREGORIOU, G. N.; ROUAH, F. D. Efficiency Persistence of Bank and Thrift CEOs Using Data Envelopment Analysis. In: **Data Envelopment Analysis**. [s.l.] : Springer, 2016. p. 1–16.

CHENG, H.; CHEN, C.; WU, S.; MIRZA, Z. A.; LIU, Z. Emergy evaluation of cropping, poultry rearing, and fish raising systems in the drawdown zone of Three Gorges Reservoir of China. **Journal of Cleaner Production**, [s. l.], v. 144, p. 559–571, 2017.

CLEMENTE, F.; LÍRIO, V. S.; GOMES, M. F. M. Technical efficiency in Brazilian citrus production. **Bio-based and Applied Economics**, [s. l.], v. 4, n. 2, p. 165–178, 2015.

COELLI, T. **A guide to DEAP version 2.1: a data envelopment analysis (computer) program**. 1996. Disponível em: <<http://www.owl.net.rice.edu/~econ380/DEAP.PDF>>. Acesso em: 18 maio. 2018.

CONCEIÇÃO, M. A. F. **Roteiro de cálculo da evapotranspiração de referência pelo método de Penman-Monteith-FAO**. - Portal Embrapa. 2006. Disponível em: <<https://www.embrapa.br/busca-de-publicacoes/-/publicacao/541867/roteiro-de-calculo-da-evapotranspiracao-de-referencia-pelo-metodo-de-penman--monteith-fao>>. Acesso em: 8 jul. 2019.

COOPER, W. W.; SEIFORD, L. M. **Data Envelopment Analysis: A Comprehensive Text with Models, Applications, References and DEA-Solver Software**. 2. ed. [s.l.] : Springer-Verlag US, 2007.

COSTA, M. A. B.; BATAGHIN, F. A.; FRANCI, A. A. B.; BRASSOLATTI, T. F. Z. Ecoeficiência energética no setor sucroalcooleiro. **A Cartilha da bioeletricidade**, [s. l.], v. 2, n. 4, p. 21, 2017.

DIAS JÚNIOR, A. F.; LANA, A. Q.; BERNARDES, F. F.; BRITO, J. O. Madeira, energia e geração de termoeletricidade. **ENERGIA NA AGRICULTURA**, [s. l.], v. 32, n. 2, p. 157, 2017.

EURICH, J. **Balço energético em agrossistemas de produção familiar**. 2011. Dissertação (Mestrado em Agricultura) - Universidade Estadual de Ponta Grossa, Ponta Grossa, 2011. Disponível em: <<http://tede2.uepg.br/jspui/handle/prefix/2210>>. Acesso em: 3 jul. 2019.

EURICH, J. **Externalidades de um sistema de produção agrícola e pecuário**. 2016. Tese (Doutorado em Agricultura) - Universidade Estadual de Ponta Grossa, Ponta Grossa, 2016. Disponível em: <<http://tede2.uepg.br/jspui/handle/prefix/2294>>. Acesso em: 9 jul. 2019.

FARIAS, J. R. B.; NEPOMUCENO, A. L.; NEUMAIER, N. Ecofisiologia da soja. **Embrapa Soja-Circular Técnica (INFOTECA-E)**, [s. l.], 2007.

FELICI, P. H. N.; CORONATO, R.; GAZZONI, D.; RALISCH, R. Balço energético das culturas de girassol e soja para produção de biocombustível. In: 2006, **Anais...** . In: EMBRAPA SOJA-ARTIGO EM ANAIS DE CONGRESSO (ALICE). : In: JORNADA ACADÊMICA DA EMBRAPA SOJA, 2., 2006, Londrina. Resumos expandidos. Londrina: Embrapa Soja, 2006., 2006.

FERRAZ, D.; MORALLES, H. F.; CAMPOLI, J. S.; OLIVEIRA, F. C. R.; REBELATTO, D. A. do N. Economic Complexity and Human Development: DEA performance measurement in

Asia and Latin America. **Gestão & Produção**, [s. l.], 2018. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0104-530X2018005008101&lng=en&tlng=en>. Acesso em: 16 ago. 2018.

FISCHER, J. **Energy inputs in Swiss agriculture**. 1999. Disponível em: <<http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.707.6655&rep=rep1&type=pdf>>. Acesso em: 14 set. 2018.

FOX, J.; WEISBERG, S.; PRICE, B.; ADLER, D.; BATES, D.; BAUD-BOVY, G.; BOLKER, B.; ELLISON, S.; FIRTH, D.; FRIENDLY, M.; GORJANC, G.; GRAVES, S.; HEIBERGER, R.; LABOISSIERE, R.; MAECHLER, M.; MONETTE, G.; MURDOCH, D.; NILSSON, H.; OGLE, D.; RIPLEY, B.; VENABLES, W.; WALKER, S.; WINSEMIUS, D.; ZEILEIS, A.; R-CORE. **car: Companion to Applied Regression**. 2019. Disponível em: <<https://CRAN.R-project.org/package=car>>. Acesso em: 11 jul. 2019.

GALE, F.; VALDES, C.; ASH, M. **Interdependence of China, United States, and Brazil in Soybean Trade**. 2019. Disponível em: <<http://www.ers.usda.gov/publications/pub-details/?pubid=93389>>. Acesso em: 11 dez. 2019.

GHORBANI, R.; MONDANI, F.; AMIRMORADI, S.; FEIZI, H.; KHORRAMDEL, S.; TEIMOURI, M.; SANJANI, S.; ANVARKHAH, S.; AGHEL, H. A case study of energy use and economical analysis of irrigated and dryland wheat production systems. **Applied Energy**, [s. l.], v. 88, n. 1, p. 283–288, 2011.

GOMES, E. G.; ABREU, U. G. P.; MELLO, J. C. C. B. S.; CARVALHO, T. B.; ZEN, S. Economic and socio-environmental performance assessment of beef cattle production systems: a data envelopment analysis (DEA) approach with weight restrictions. **Revista Brasileira de Zootecnia**, [s. l.], v. 44, n. 6, p. 219–225, 2015.

GROWITSCH, C.; JAMASB, T.; MÜLLER, C.; WISSNER, M. Social Cost Efficient Service Quality: Integrating Customer Valuation in Incentive Regulation—Evidence from the Case of Norway. In: **Data Envelopment Analysis**. [s.l.] : Springer, 2016. p. 71–91.

GUSE, J. C.; POLITELO, L.; LEITE, M.; SILVA, T. P. Propriedade Familiar e Desempenho de Empresas do Agronegócio Brasileiro. **Revista em Agronegócio e Meio Ambiente**, [s. l.], v. 8, n. 2, p. 271, 2015.

GUZMÁN, G. I.; GONZÁLEZ DE MOLINA, M. Energy Efficiency in Agrarian Systems From an Agroecological Perspective. **Agroecology and Sustainable Food Systems**, [s. l.], v. 39, n. 8, p. 924–952, 2015.

HANSEN, R. D. Water and waste water in imperial Rome. **Journal of the American Water Resources Association**, [s. l.], v. 19, n. 2, p. 263–269, 1983.

HELSEL, Z. R. **Energy Use and Efficiency in Pest Control, Including Pesticide Production, Use, and Management Options - eXtension**. 2016. Disponível em: <<https://articles.extension.org/pages/62513/energy-use-and-efficiency-in-pest-control-including-pesticide-production-use-and-management-options>>. Acesso em: 5 set. 2018.

HIRAKURI, M. H.; DEBIASI, H.; PROCÓPIO, S. de O.; FRANCHINI, J. C.; CASTRO, C. De. Sistemas de produção: conceitos e definições no contexto agrícola. **Documentos n. 335**, Londrina: Embrapa Soja. [s. l.], v. 14, n. 1, Londrina: Embrapa Soja, p. 24, 2012.

HŘEBÍČEK, J.; TRENZ, O.; CHVÁTALOVÁ, Z.; SOUKOPOVÁ, J. Optimization of corporate performance using data envelopment analysis with Maple. In: **Engineering Optimization 2014**. London: CRC Press, 2015. p. 763–767.

HUSSON, F.; JOSSE, J.; LE, S.; MAZET, J. **FactoMineR: Multivariate Exploratory Data Analysis and Data Mining**. 2019. Disponível em: <<https://CRAN.R-project.org/package=FactoMineR>>. Acesso em: 12 jul. 2019.

HUSSON, F.; JOSSE, J.; PAGES, J. Principal component methods-hierarchical clustering-partitional clustering: why would we need to choose for visualizing data. **Applied Mathematics Department**, [s. l.], p. 1–17, 2010.

INMET. :: **INMET - Instituto Nacional de Meteorologia** :: 2019. Disponível em: <<http://www.inmet.gov.br/portal/index.php?r=estacoes/estacoesAutomaticas>>. Acesso em: 8 jul. 2019.

JAIN, A. K. Data clustering: 50 years beyond K-means. **Pattern recognition letters**, [s. l.], v. 31, n. 8, p. 651–666, 2010.

JOMEHPOUR, M. Qanat irrigation systems as important and ingenious agricultural heritage: case study of the qanats of Kashan, Iran. **International Journal of Environmental Studies**, [s. l.], v. 66, n. 3, p. 297–315, 2009.

JUNIOR, A. P. S.; CARLUCCI, F. V.; GRESPAN, C. A. Investment Potential for New Sugarcane Plants in Brazil Based on Assessment of Operational Efficiency. **International Food and Agribusiness Management Review**, [s. l.], v. 17, n. 2, p. 41–64, 2014.

KASSAMBARA, A. **Practical Guide To Principal Component Methods in R: PCA, M(CA), FAMD, MFA, HCPC, factoextra**. [s.l.] : CreateSpace Independent Publishing Platform, 2017. Disponível em: <<https://books.google.com.br/books?id=eFEyDwAAQBAJ>>

KASSAMBARA, A.; MUNDT, F. **factoextra: Extract and Visualize the Results of Multivariate Data Analyses**. 2017. Disponível em: <<https://CRAN.R-project.org/package=factoextra>>. Acesso em: 12 jul. 2019.

KAZEMI, H.; KAMKAR, B.; LAKZAEI, S.; BADSAR, M.; SHAHBYKI, M. Energy flow analysis for rice production in different geographical regions of Iran. **Energy**, [s. l.], v. 84, p. 390–396, 2015.

KORKMAZ, S.; GOKSULUK, D.; ZARARSIZ, G. MVN: An R package for assessing multivariate normality. **The R Journal**, [s. l.], v. 6, n. 2, p. 151–162, 2014.

LIGHTFOOT, D. R. Syrian qanat Romani: history, ecology, abandonment. **Journal of Arid Environments**, [s. l.], v. 33, n. 3, p. 321–336, 1996.

LIGHTFOOT, D. R. The Origin and Diffusion of Qanats in Arabia: New Evidence from the northern and southern Peninsula. **The Geographical Journal**, [s. l.], v. 166, n. 3, p. 215–226, 2000.

- LORENZONI MARTINS, F. G.; BARBOSA, J. A.; DA SILVA CARVALHO, R. C.; VELOSO, A. V.; MARIN, D. B. Análise energética da produção de milho para silagem cultivado em diferentes espaçamentos. **ENERGIA NA AGRICULTURA**, [s. l.], v. 30, n. 4, p. 418, 2015.
- LUNKES, R. J. Uma contribuição à formação de preços de venda. **Revista Brasileira de contabilidade**, [s. l.], n. 141, p. 50–57, 2003.
- MANLY, B. F. J.; ALBERTO, J. A. N. **Multivariate statistical methods: a primer**. Fourth edition ed. Boca Raton: CRC Press, Taylor & Francis Group, 2017.
- MARIANO, E. B.; REBELATTO, D. A. do N. Transformation of wealth produced into quality of life: analysis of the social efficiency of nation-states with the DEA's triple index approach. **Journal of the Operational Research Society**, [s. l.], v. 65, n. 11, p. 1664–1681, 2014.
- MARTIN, N. B.; SERRA, R.; OLIVEIRA, M. D. M.; ÂNGELO, J. A.; OKAWA, H. Sistema Integrado de Custos Agropecuários - CUSTAGRI. **Informações Econômicas do estado de São Paulo - Instituto de Economia Agrícola**, [s. l.], v. 28, n. 1, p. 7–28, 1998.
- MATSON, P. A. Integration of Environmental, Agronomic, and Economic Aspects of Fertilizer Management. **Science**, [s. l.], v. 280, n. 5360, p. 112–115, 1998.
- MIRANDA, E. E. **Agricultura no Brasil do século XXI**. [s.l.] Metalivros, 2013.
- MONDANI, F.; ALEAGHA, S.; KHORAMIVAFA, M.; GHOBADI, R. Evaluation of greenhouse gases emission based on energy consumption in wheat Agroecosystems. **Energy Reports**, [s. l.], v. 3, p. 37–45, 2017.
- MORRIS, C. R. **Os magnatas: como Andrew Carnegie, John D. Rockefeller, Jay Gould e J. P. Morgan inventaram a supereconomia americana**. 2. ed. Porto Alegre, RS: L&PM, 2006.
- MOUSAVI-AVVAL, S. H.; RAFIEE, S.; JAFARI, A.; MOHAMMADI, A. Optimization of energy consumption for soybean production using Data Envelopment Analysis (DEA) approach. **Applied Energy**, [s. l.], v. 88, n. 11, p. 3765–3772, 2011.
- MOUSAVI-AVVAL, S. H.; RAFIEE, S.; JAFARI, A.; MOHAMMADI, A. Energy saving of field operations in soybean production using data envelopment analysis approach. **African Journal of Agricultural Research**, [s. l.], v. 8, n. 46, p. 5776–5784, 2013.
- NACHILUK, K. Tecnologia: mais do mesmo? **Análises e indicadores do agronegócio**, [s. l.], v. 13, n. 8, 2018. Disponível em: <<http://www.iea.sp.gov.br/ftp/iea/aia/AIA-45-2018.pdf>>. Acesso em: 4 ago. 2018.
- NASSIRI, S. M.; SINGH, S. Study on energy use efficiency for paddy crop using data envelopment analysis (DEA) technique. **Applied Energy**, [s. l.], v. 86, n. 7–8, p. 1320–1325, 2009.
- OLIVEIRA, E.; ANDRADE, F. V.; MELLO, J. C.; MACHADO, T. B.; PEREIRA, C. R. Avaliação da eficiência de horticultores agroecológicos utilizando análise envoltória de dados. **Horticultura Brasileira**, [s. l.], v. 32, n. 3, p. 336–342, 2014.

OMMANI, A. R. Productivity of energy consumption in agricultural productions: A case study of corn farmers of Ahwaz Township, Iran. **African Journal of Agricultural Research**, [s. l.], v. 6, n. 13, p. 2945–2949, 2011.

ORTEGA, E.; CAVALETT, O.; BONIFÁCIO, R.; WATANABE, M. Brazilian Soybean Production: Emery Analysis With an Expanded Scope. **Bulletin of Science, Technology & Society**, [s. l.], v. 25, n. 4, p. 323–334, 2005.

ORTIZ-CAÑAVATE, J.; HERNANZ, J. Energy analysis and saving. In: KITANI, O. (Ed.). **CIGR Handbook of agricultural engineering**. [s.l.] : ASAE St. Joseph, MI, 1999. v. 5p. 13–42.

PEREIRA, C. N.; SILVEIRA, J. M. F. J. Análise Exploratória da Eficiência Produtiva das Usinas de Cana-de-açúcar na Região Centro-Sul do Brasil. **Revista de Economia e Sociologia Rural**, [s. l.], v. 54, n. 1, p. 147–166, 2016.

PEREIRA, N. A.; TAVARES, M. Efficiency of major producing regions of sugar cane through Data Envelopment Analysis (DEA). **Custos e @gronegocio on line**, [s. l.], v. 13, p. 37–70, 2017.

PETERAF, M. A.; BARNEY, J. B. Unraveling the resource-based tangle. **Managerial and Decision Economics**, [s. l.], v. 24, n. 4, p. 309–323, 2003.

PIMENTEL, D. Chapter 2 - Energy Inputs in Production Agriculture. In: FLUCK, R. C. (Ed.). **Energy in Farm Production**. Energy in World Agriculture Amsterdam: Elsevier, 1992. p. 13–29.

R CORE TEAM. **R: A language and environment for statistical computing**. **R Foundation for Statistical Computing**. 2017. Disponível em: <<https://www.R-project.org/>>. Acesso em: 1 mar. 2017.

REINEKE, H.; STOCKFISCH, N.; MÄRLÄNDER, B. Analysing the energy balances of sugar beet cultivation in commercial farms in Germany. **European Journal of Agronomy**, [s. l.], v. 45, p. 27–38, 2013.

RIBEIRO, C. D. O.; SOSNOSKI, A. A. K. B.; OLIVEIRA, S. M. Um modelo hierárquico para previsão de preços de commodities agrícolas. **Revista Produção Online**, [s. l.], v. 10, n. 4, p. 719–733, 2010.

ROCKSTRÖM, J.; STEFFEN, W.; NOONE, K.; PERSSON, Å.; CHAPIN, F. S.; LAMBIN, E. F.; LENTON, T. M.; SCHEFFER, M.; FOLKE, C.; SCHELLNHUBER, H. J.; NYKVIST, B.; DE WIT, C. A.; HUGHES, T.; VAN DER LEEUW, S.; RODHE, H.; SÖRLIN, S.; SNYDER, P. K.; COSTANZA, R.; SVEDIN, U.; FALKENMARK, M.; KARLBERG, L.; CORELL, R. W.; FABRY, V. J.; HANSEN, J.; WALKER, B.; LIVERMAN, D.; RICHARDSON, K.; CRUTZEN, P.; FOLEY, J. A. A safe operating space for humanity. **Nature**, [s. l.], v. 461, n. 7263, p. 472–475, 2009.

RODRIGUES, W. A. Sistemas de gestão ambiental: compatibilização dos objetivos econômicos, estratégicos, sociais e ambientais nas organizações empresariais. **Interface tecnológica**, [s. l.], v. 7, n. 1, p. 11, 2010.

ROMANELLI, T. L.; NARDI, H. de S.; SAAD, F. A. Material embodiment and energy flows as efficiency indicators of soybean (*Glycine max*) production in Brazil. **Engenharia Agrícola**, [s. l.], v. 32, n. 2, p. 261–270, 2012.

SALGADO JUNIOR, A. P.; CARLUCCI, F. V.; NOVI, J. C.; PACAGNELLA JUNIOR, A. C. Aplicação da análise envoltória de dados (AED) na avaliação da eficiência operacional relativa entre usinas de cana-de-açúcar no território brasileiro. **Engenharia Agrícola**, [s. l.], v. 34, n. 5, p. 826–843, 2014.

SAMANI, B. H.; CHOUBIN, S.; GHASEMI-VARNAMKHAZI, M.; ABEDI, A. Analysis of energy consumption and end-use application of rapeseed in an agricultural production system in Izeh-Khuzestan. **Engineering in Agriculture, Environment and Food**, [s. l.], v. 11, n. 3, p. 101–108, 2018.

SANTOS, A. B.; SPROESSER, R. L.; BATALHA, M. O.; CAMPEÃO, P.; PEREIRA, M. W. G. Are the grain intermodal terminals in Brazil's Northeastern region efficient? **Custos e @gronegócio on line**, [s. l.], v. 12, n. 2, p. 64–83, 2016.

SATTLER, C.; MATZDORF, B. PES in a nutshell: From definitions and origins to PES in practice—Approaches, design process and innovative aspects. **Ecosystem Services**, [s. l.], v. 6, p. 2–11, 2013.

SCHUH, G. E. Considerações teóricas sobre custos de produção na agricultura. **Agricultura em São Paulo**, [s. l.], v. 23, p. 97–119, 1976.

SCORZA JÚNIOR, R. **Pesticidas, agricultura e recursos hídricos**. 2006. Disponível em: <<https://ainfo.cnptia.embrapa.br/digital/bitstream/item/38751/1/CT200612.pdf>>. Acesso em: 19 set. 2018.

SHARIFI, M. Energy inputs - Yield relationship and cost analysis of melon production in Khorasan Razavi province of Iran. **Engineering in Agriculture, Environment and Food**, [s. l.], v. 11, n. 3, p. 109–113, 2018.

SILVA, A. R. **biotools: Tools for Biometry and Applied Statistics in Agricultural Science**. 2017. Disponível em: <<https://CRAN.R-project.org/package=biotools>>. Acesso em: 11 jul. 2019.

SILVA, V. S.; KAKIHARA, A. A. S. B. de S.; POKER, J. H.; MILANI FILHO, M. A.; BELLI, M. M. Análise de eficiência dos serviços de hospitais públicos nas capitais brasileiras. **Revista Produção Online**, [s. l.], v. 17, n. 3, p. 1090, 2017.

SIMM, J.; BESSTREMYANNAYA, G.; SIMM, M. J. **Package 'rDEA'**. 2016. Disponível em: <<ftp://200.236.31.7/CRAN/web/packages/rDEA/rDEA.pdf>>. Acesso em: 3 ago. 2018.

SINGH, K.; PRAKASH, V.; SRINIVAS, K.; SRIVASTVA, A. Effect of tillage management on energy-use efficiency and economics of soybean (*Glycine max*) based cropping systems under the rainfed conditions in North-West Himalayan Region. **Soil and Tillage Research**, [s. l.], v. 100, n. 1–2, p. 78–82, 2008.

SMIL, V. **Two prime movers of globalization: the history and impact of diesel engines and gas turbines**. Cambridge, Mass: MIT Press, 2010.

STOUT, B. A. CHAPTER II - Energy Use. In: STOUT, B. A. (Ed.). **Handbook of Energy for World Agriculture**. Amsterdam: Pergamon, 1989. p. 50–94.

SUEYOSHI, T.; GOTO, M. **Environmental Assessment on Energy and Sustainability by Data Envelopment Analysis**. 1. ed. [s.l.] : John Wiley & Sons Ltd, 2018. Disponível em: <<https://books.google.com.br/books?id=abdJDwAAQBAJ>>. Acesso em: 28 ago. 2018.

TABATABAEFAR, A.; EMAMZADEH, H.; VARNAMKHAJASTI, M.; RAHIMIZADEH, R.; KARIMI, M. Comparison of energy of tillage systems in wheat production. **Energy**, [s. l.], v. 34, n. 1, p. 41–45, 2009.

TEIXEIRA, S.; DIAS, C. T. dos S. Manejo fitossanitário de plantas daninhas, insetos e doenças em soja e equivalentes energéticos. In: ANAIS DO V CONGRESSO BRASILEIRO DE FITOSSANIDADE 2019, Curitiba. **Anais...** . In: V CONGRESSO BRASILEIRO DE FITOSSANIDADE. Curitiba Disponível em: <<http://fitossanidade.fcav.unesp.br/seer/index.php/anaisconbraf/article/view/279>>. Acesso em: 7 set. 2019.

THANASSOULIS, E. **Introduction to the theory and application of data envelopment analysis: a foundation text with integrated software**. 1. ed. Boston: Springer, 2001.

WANG, X.; CHEN, Y.; SUI, P.; GAO, W.; QIN, F.; ZHANG, J.; WU, X. Emergy analysis of grain production systems on large-scale farms in the North China Plain based on LCA. **Agricultural Systems**, [s. l.], v. 128, p. 66–78, 2014.

WICKHAM, H.; BRYAN, J.; ATTRIBUTION), Rs. (Copyright holder of all R. code and all C. Code without explicit copyright; CODE), M. K. (Author of included R.; CODE), K. V. (Author Of included libxls; CODE), C. L. (Author Of included libxls; CODE), B. C. (Author Of included libxls; CODE), D. H. (Author Of included libxls; CODE), E. M. (Author Of included libxls. **readxl: Read Excel Files**. 2019a. Disponível em: <<https://CRAN.R-project.org/package=readxl>>. Acesso em: 11 jul. 2019.

WICKHAM, H.; CHANG, W.; HENRY, L.; PEDERSEN, T. L.; TAKAHASHI, K.; WILKE, C.; WOO, K.; YUTANI, H.; RSTUDIO. **ggplot2: Create Elegant Data Visualisations Using the Grammar of Graphics**. 2019b. Disponível em: <<https://CRAN.R-project.org/package=ggplot2>>. Acesso em: 11 jul. 2019.

WICKHAM, H.; FRANÇOIS, R.; HENRY, L.; MÜLLER, K.; RSTUDIO. **dplyr: A Grammar of Data Manipulation**. 2019c. Disponível em: <<https://CRAN.R-project.org/package=dplyr>>. Acesso em: 11 jul. 2019.

YU, M.-M.; CHEN, L.-H. Assessment of Transportation Performance: A Network Structure. In: **Data Envelopment Analysis**. [s.l.] : Springer, 2016. p. 17–43.

APÊNDICE A – Demonstrativo de programação e cálculos no R Software.

```

# # Importando planilha
# library(readxl)
# data <- read_excel("Tratamento_dados_soja.xlsx",
#                   sheet = "Equivalentes",
#                   range = "b1:x82")
#
# # Agrupando as variáveis de mecanização em "data2"
# library(dplyr)
# data2 <- data %>%
#   mutate(Mecanização= `Trator 1`+`Trator 2`+ Colhedora
#          + Semeadora + Aplicador + Pulverizador)
#
# data2 <- data2[,-(16:21)]
#
# #####
# ##### Obtendo os valores CRS-DEA para as três safras#####
# #####
# # Dataset
# deadf <- data2[,-c(2,5)]
#
# # Safra 1
# s1 <- deadf[deadf$Safra == "2014/2015",]
#
# X <- s1[,4:16]
# Y <- s1[,3]
#
# library(rDEA)
# crs <- dea(XREF = X, YREF = Y, X=X, Y=Y, model = "input", RTS = "contant")
# crs_s1 <- crs$thetaOpt
#
# # Safra 2
# s2 <- deadf[deadf$Safra == "2015/2016",]
#
# X <- s2[,4:16]
# Y <- s2[,3]
#
# library(rDEA)
# crs <- dea(XREF = X, YREF = Y, X=X, Y=Y, model = "input", RTS = "constant")
# crs_s2 <- crs$thetaOpt
#
# # Safra 3
# s3 <- deadf[deadf$Safra == "2016/2017",]
#
# X <- s3[,4:16]
# Y <- s3[,3]
#
# library(rDEA)
# crs <- dea(XREF = X, YREF = Y, X=X, Y=Y, model = "input", RTS = "contant")
# crs_s3 <- crs$thetaOpt
#
# # Agrupando os dados de CRS em um quadro resumo
# quadro_resumo <- as.data.frame(cbind(crs_s1=100*round(crs_s1, 4),
#                                     crs_s2=100*round(crs_s2, 4),
#                                     crs_s3=100*round(crs_s3, 4)))
# quadro_resumo <- as.data.frame(cbind(Identificação=c(DMU[1:27]),
#                                     quadro_resumo))
#
# # Calculando os valores de CRS de acordo com as Janelas de eficiência
# library(dplyr)
# quadro_resumo <- quadro_resumo %>%
#   mutate(Jan_1 = round((crs_s1+crs_s2)/2,2),
#          Jan_2 = round((crs_s2+crs_s3)/2,2),
#          media_jan=round((Jan_1+Jan_2)/2, 2),
#          sd=round(sqrt((Jan_1-media_jan)^2+(Jan_2-media_jan)^2/2), 2))
#
# # Dados exportados em arquivo *.csv
# write.csv(quadro_resumo, file = "valores_DEA_CRS.csv", row.names = FALSE)
#
# #####
# ##### MANOVA DAS CLASSES DE EFICIÊNCIA #####
# #####
# # Importando dados

```

```

# library(readxl)
# df <- as.data.frame(read_excel("valores_crs.xlsx",
#                               sheet = "Equivalentes_class_efic",
#                               range = "a1:r82"))
# df$class_efic <- as.factor(df$class_efic)
#
# # Agrupando variáveis:
# # Produção - prod
# # Água (energia química) - agua
# # Químicos (defensivos) - pestic
# # Nutrientes (macros e Micros) - nutri
# # Sementes - semente
# # Operacionais (Mecanização, Mão-de-obra, combustível) - operac
# library(dplyr)
# df <- df %>%
#   mutate(prod = Produção,
#          pestic = Adjuvantes+Herbicidas+Inseticidas+Fungicidas+Espalhantes,
#          nutri = Micronutrientes+N+K2O+P2O5,
#          semente = Sementes,
#          operac = Combustível+Mecanização+`Mão-de-obra`)
#
# dfmanova <- select(df, DMU, class_efic, prod, pestic, nutri, semente, operac)
#
# normalit<-function(m){
#   (m - min(m))/(max(m)-min(m))
# }
#
# dfmanova <- dfmanova %>%
#   mutate_each(funs(normalit), prod, pestic, nutri, semente, operac)
#
# # Foi reduzido tamanho da amostra no intuito de atender
# # Os pressupostos de Homocedasticidade e normalidade
#
# dfmanova <- dfmanova %>%
#   filter( pestic > 0.1 & pestic < 2100 &
#          nutri > 0.1 & nutri < 3980 &
#          semente > 0.1 &
#          prod<1000 & prod>0.23 &
#          operac>0.1)
#
# summary(dfmanova)
#
# # Definindo as variaveis dependentes (vdepend) e indepedentes (vindepend)
# vdepend <- dfmanova$class_efic
# vindepend <- as.matrix((dfmanova[,2:6]))
#
# #Hmocedasticidade
# library(biotoools)
# boxM(vindepend, vdepend)
#
# # Realizando a manova com os quatro testes
# library(car)
# fit <- lm(vindepend~vdepend, data = dfmanova)
# tabela <- Manova(fit)
# summary(tabela, multivariate = TRUE)
# Manova(fit, test.statistic="Wilks")
#
# # Avaliação dos resíduos
# residuos <- data.frame(residuals(fit))
#
# # Teste de normalidade multivariada - Mardia Curtose e Skewness (p-0,05)
# # Teste de normalidade univariada Shapiro-wilk (p-0,05)
# library(MVN)
# mardia.tes <- mvn(residuos[,1:5],
#                  mvnTest = c("mardia"),
#                  multivariatePlot = "qq",
#                  covariance = FALSE,
#                  multivariateOutlierMethod = "adj")
# mardia.tes
#
# # Plot Mahalanobis vs Q-sqr
# residuos <- residuos[,1:5]
# as.matrix(residuos)

```

```

# centro <- colMeans(as.matrix(residuos))
# cov <- cov(as.matrix(residuos))
# mahalnbis <- mahalnobis(as.matrix(residuos), centro, cov)
#
#
# residuos <- residuos[,1:5]
# residuos <- cbind(residuos, mahalnbis)
# residuos <- residuos[order(residuos$mahalnbis,
#                             residuos$pestic,
#                             residuos$nutri,
#                             residuos$semente,
#                             residuos$operac,
#                             decreasing = FALSE),]
# Qsqr <- qchisq(ppoints(nrow(residuos)), df=3)
# residuos <- cbind(residuos, Qsqr)
# residuos
#
#
# # PLOT
# options(OutDec = ",")
# library(ggplot2)
# ggplot(residuos,
#        aes(residuos$mahalnbis,
#            residuos$Qsqr))+
#   labs(#title = "Plot dos resíduos multivariados",
#        #subtitle = expression("Quantil  $\sim$ italic(chi)2 $\sim$ vs Distância de Mahalanobis"),
#        x = "Distância de Mahalanobis",
#        y = expression("Quantil  $\sim$ italic(chi)2))+
#   geom_point(shape=21, size=3, stroke=1.25) +
#   geom_smooth(col="blue", se = FALSE, method = "lm")+
#   theme_bw()+
#   theme(text = element_text(family = "serif", size = 14),
#         panel.grid = element_blank(),
#         axis.text = element_text(size = 14, colour = "black"))
#
# # Preparando dados para plotar resíduos univariados
# Classe <- as.character(dfmanova$class_efic)
#
# prod <- cbind(Classe,rep("Produtividade",nrow(dfmanova)))
# pestic <- cbind(Classe,rep("Pesticidas",nrow(dfmanova)))
# nutri <- cbind(Classe,rep("Nutrientes",nrow(dfmanova)))
# semente <- cbind(Classe,rep("Sementes",nrow(dfmanova)))
# operac <- cbind(Classe,rep("Operacionais",nrow(dfmanova)))
#
# plot.resi <- as.data.frame(rbind(prod, pestic, nutri, semente, operac))
#
# resi.col <- c(residuos$prod,
#              residuos$pestic,
#              residuos$nutri,
#              residuos$semente,
#              residuos$operac)
#
# plot.resi <- cbind(plot.resi, resi.col)
#
# # Plot Univariados
# library(ggplot2)
# ggplot(plot.resi,
#        aes(sample=plot.resi$resi.col))+
#   labs(title = "(b)",
#        subtitle = "Univariate residuals"
#        ,x = "Quantile - theoretic"
#        ,y = "Quantile - sample")+
#   geom_qq(shape=21, size=1.5, stroke=1.0, color="gray25")+
#   geom_qq_line(col="black", size=0.75)+
#   facet_wrap(~plot.resi$V2, ncol = 2, scales = "free_y")+
#   theme_bw()+
#   theme(text = element_text(size = 20,
#                             family = "serif"),
#         axis.text = element_text(colour = "black"),
#         panel.grid = element_blank())
#
#
# # Dados organizados em colunas
# coluna_valores <- as.vector(cbind(dfmanova$prod,
#                                   dfmanova$pestic,

```

```

#                                     dfmanova$nutri,
#                                     dfmanova$semente,
#                                     dfmanova$operac))
#
# dados.colun <- cbind(plot.resi, coluna_valores)
# names(dados.colun) <- c("Classe", "Variáveis", "Resíduos", "Valor")
#
# # Multivariados e univariados na mesma figura
# library(ggplot2)
# library(gridExtra)
#
# png(filename = "resíduos.png", width = 20, height = 10, units = "cm", res = 300)
# grid.arrange(arrangeGrob(ggplot(resíduos,
#                               aes(resíduos$mahlnbis,
#                                   resíduos$qsqr))+
#                               labs(title = "(a)"
#                                   ,x = "Distância de Mahalanobis"
#                                   ,y = expression("Quantil ~italic(chi)^2))+
#                               geom_point(shape=21, size=2, stroke=0.75, color="black")+
#                               geom_smooth(col="navy", se = FALSE, method = "lm")+
#                               theme_bw()+
#                               theme(text = element_text(family = "serif", size = 14),
#                                       axis.text = element_text(size = 9, color = "black"),
#                                       panel.grid = element_blank(),
#                                       plot.background = element_rect(colour = "gray85"))),
#               ggplot(plot.resi,
#                       aes(sample=plot.resi$resi.col))+
#               labs(title = "(b)"
#                   ,x = "Quantil - teórico"
#                   ,y = "Quantil - amostral")+
#               geom_qq(shape=21, size=2, stroke=0.75, color="black")+
#               geom_qq_line(col="navy", size=0.75)+
#               facet_wrap(~plot.resi$V2, ncol = 2)+
#               theme_bw()+
#               theme(text = element_text(size = 14,
#                                       family = "serif"),
#                       axis.text = element_text(colour = "black", size = 9),
#                       panel.grid = element_blank(),
#                       plot.background = element_rect(colour = "gray85")),
#               ncol=2)
# dev.off()
#
# #####
# # Intervalos de confiança simplificados de Bonferroni#
# #####
#
# n <- dim(dfmanova)[1] / length(unique(dfmanova$class_efic))
# total.means <- colMeans(dfmanova[,3:7])
# class.group <- data.matrix(split(dfmanova[,3:7], dfmanova$class_efic))
# class.means <- sapply(class.group, function(x) {
#   apply(x, 2, mean)
# }, simplify = 'data.frame')
# class.means
# rownames(class.means) <- c("Produtividade",
#                            "Pesticidas",
#                            "Nutrientes",
#                            "Sementes",
#                            "Operacionais")
#
# # Soma dos quadrados - Matriz B: entre samples (Between samples)
# # B = matrix(data = 0, nrow = 3, ncol = 3)
# # for (i in 1:dim(B)[1]) {
# #   for (j in 1:i) {
# #     B[i,j] <- n * sum((Pest.means[i,] - total.means[i]) *
# #                     (Pest.means[j,] - total.means[j]))
# #     B[j,i] <- n * sum((Pest.means[j,] - total.means[j]) *
# #                     (Pest.means[i,] - total.means[i]))
# #   }
# # }
# # B
#
# # Soma dos quadrados - Matriz W: dentro dos samples (within samples)
# # W = matrix(data = 0, nrow = 5, ncol = 5)
# # for (i in 1:dim(w)[1]) {

```

```

# for (j in 1:i) {
#   b <- c()
#   for (k in class.group) {
#     a <- sum((k[,i] - mean(k[,i])) * (k[,j] - mean(k[,j])))
#     b <- append(b, a)
#   }
#   w[i,j] <- sum(b)
#   w[j,i] <- sum(b)
# }
# }
#
# w
#
# diagW <- as.vector(diag(w))
#
# n <- nrow(dfmanova)
# p <- ncol(dfmanova[,-(1:2)])
# g <- nlevels(dfmanova$class_efic)
# alfa <- 0.05
#
# b <- c()
# for (i in 1:5) {
#   a <- qt((1-(alfa/(p*g*(g-1)))), (n-g))*
#     sqrt((w[i]/(n-g))*((1/nrow(dfmanova[dfmanova$class_efi=="Eficiente",]))+
#       (1/nrow(dfmanova[dfmanova$class_efi=="Ineficiente",])))))
#   b <- append(b, a)
# }
#
# LS <- (class.means[,1]-class.means[,2])+b
# DI <- (class.means[,1]-class.means[,2])
# LI <- (class.means[,1]-class.means[,2])-b
#
# ISBon <- as.data.frame(cbind(DI, LI, LS))
# ISBon
#
# options(OutDec = ",")
# library(ggplot2)
# library(ggrepel)
# png(filename = "diferença_médias_bonferroni.png", width = 15,
#   height = 10, units = "cm", res = 300)
# ggplot(ISBon, aes(ISBon$DI, rownames(ISBon)))+
#   geom_errorbarh(aes(xmin=ISBon$DI-b,
#     xmax=ISBon$DI+b),
#     height=0.15,
#     col="navy")+
#   geom_point(size=2.5, shape=21, fill="white", stroke=0.75,col="navy")+
#   geom_text_repel(label=round(ISBon$DI,3),
#     nudge_x = 0.0, nudge_y = 0.2,
#     show.legend = FALSE,
#     segment.alpha = 0,
#     label.size = 0.75,
#     size=3)+
#   geom_vline(xintercept = 0, lty="dotted",
#     col="gray", size=0.5)+
#   scale_x_continuous(
#     name = "Intervalos simultâneos com 95% de confiança \nestimados pelo método de Bonferroni",
#     limits = c(-0.15, 0.3), breaks = seq(-0.15,0.3,0.15))+
#   labs(y="Variáveis")+
#   theme_bw()+
#   theme(text = element_text(size = 12,
#     family = "serif"),
#     axis.text = element_text(colour = "black",
#       size = 11),
#     axis.ticks.y = element_blank(),
#     axis.title = element_text(size = 12),
#     panel.grid = element_blank())
# dev.off()
#
#####
##### Função logística discriminante #####
#####
#
# # Eficiência
# dfmanova$class_efic <- as.factor(dfmanova$class_efic)
# # Converting the variable "Eficiência" into a binary variable

```

```

## 0 = Ineficiente, 1 = Eficiente
# dfmanova$class_bin <- ifelse(dfmanova$class_efic == "Ineficiente",0,1)
## Intercept model (no covariates)
# logitn.dfmanova <- glm(dfmanova$class_bin ~ 1,
#                        family=binomial(link="logit"))
## Logistic regression model with the covariates included
## Invoking glm with binomial distr. for the error and logit link
# logit.dfmanova <- glm(dfmanova$class_bin ~ dfmanova$prod+
#                      dfmanova$pestic +dfmanova$nutri +
#                      dfmanova$semente + dfmanova$operac,
#                      family=binomial(link="logit"))
## Testing by Chi-square the global significance
## of the model with 4 covariates, X5-X7 and X9
# anova(logitn.dfmanova,logit.dfmanova, test="Chisq")
## Summary of glm (logit model), z values displayed
# (summ.logitsk <- summary.glm(logit.dfmanova))
## Computing chi-squared statistic = z^2
# chisq.statsk <- summ.logitsk$coefficients[,3]^2
## Replacing the z-values column with chi-squared values
# summ.logitsk$coefficients[,3] <- chisq.statsk
# colnames(summ.logitsk$coefficients)[3]="Chi-squared"
# colnames(summ.logitsk$coefficients)[4]="P(> Chi^2)"
## Table showing estimates and chi-squared values
# summ.logitsk
#
## Estimation of probabilities for every class
# estim.probs <- predict(logit.dfmanova,type="response")
## Data frame: class and estimated probability
# mat.probs <- cbind(data.frame(dfmanova$class_efic),estim.probs)
# rownames(mat.probs)=c(1:nrow(dfmanova))
# library(dplyr)
# mat.probs <- mat.probs %>% arrange(dfmanova.class_efic)
## Mean probability of Eficientes
# meanp.Efic <- mean(estim.probs[c(1:28)])
# meanp.Efic
## Mean probability of Ineficientes
# meanp.Inefic <- 1-meanp.Efic
# meanp.Inefic
## Labels of points in the plot of class in data order vs. probability
# point.dfmanova <- ifelse(mat.probs$dfmanova.class_efic == "Eficiente","E","I")
## Colors of points
# colvec <- c(rep("blue",28),rep("red",32))
## Plot: class in data order vs prob.
# options(OutDec = ",")
# png(filename = "reg_log.png", height = 14, width = 15, units = "cm", res = 300)
# plot(c(1:60),estim.probs,pch=point.dfmanova,col=colvec,las=1, cex=0.75,
#      xlab=" ",ylab="Probabilidade",xaxt="n", family="serif")
# title(xlab = "Eficiências - dados em ordem",line = 0,
#       cex.main=1, adj=0.5, family="serif")
## Showing mean probability of sex from the male
# segments(1.5,meanp.Efic,28,meanp.Efic,col="blue")
# segments(-1.0,meanp.Efic,1.5,meanp.Efic,col="blue", lty = "dotted")
## Showing mean probability of skulls from the female
# segments(29,meanp.Inefic,60,meanp.Inefic,col="red")
# segments(-1.0,meanp.Inefic,29,meanp.Inefic,col="red", lty = "dotted")
# abline(h=0.515,lty="dotted")
# mtext("0,53",side=2,at=c(-3,0.52),las=1, cex = 0.6, family="serif")
# mtext("0,56",side=2,at=c(-3,0.565),las=1, cex = 0.75, family="serif")
# mtext("0,44",side=2,at=c(-3,0.44),las=1, cex = 0.75, family="serif")
# dev.off()
#
#
#####
##### PCA #####
#####
#
## Importando dados
# library(readxl)
# df <- read_excel("valores_crs.xlsx",
#                 sheet = "Equivalentes_class_efic",
#                 range = "a1:r82")
# df <- as.data.frame(df)
# df$class_efic <- as.factor(df$class_efic)
# df$class_efic <- as.factor(df$class_efic)
# rownames(df) <- paste(df$class_efic, df$DMU, 1:nrow(df), sep = " ")

```



```

#
# # df para agrupamento por todas as variáveis em PC
# dfgrpd <- df
# library(dplyr)
# dfgrpd <- as.data.frame(dfgrpd[,-(2:3)] %>%
#                       group_by(DMU, class_efic) %>%
#                       summarise_all(funs(mean)))
# dfgrpd <- dfgrpd %>%
#   mutate(DMU = DMU,
#          prod = Produção,
#          quim = Adjuvantes+Herbicidas+Inseticidas+Fungicidas+Espalhantes,
#          nutri = Micronutrientes+N+K2O+P2O5,
#          semente = Sementes,
#          operac = Combustível+Mecanização+`Mão-de-obra`)
#
# #levels(dfgrpd$class_efic) <- c("A", "B", "M")
#
# rownames(dfgrpd) <- paste(dfgrpd$DMU, dfgrpd$class_efic, sep = " ")
#
# dfgrpd <- select(dfgrpd, class_efic, prod, quim, nutri, semente, operac)
#
# # Análise de principais componentes - Realizada com os 81 elementos amostrais.
# library(factoextra)
# library(FactoMiner)
# pca.dfgrpd <- PCA(dfgrpd[,-1],
#                  scale.unit = TRUE,
#                  graph = TRUE)
# get_eig(pca.dfgrpd)
#
# # Scree plot
# fviz_screplot(pca.dfgrpd,
#               geom="line", linetype= "dotted",
#               choice="variance",
#               addlabels=TRUE)+
#   theme(panel.grid = element_blank(),
#         panel.background = element_rect(fill=NULL, colour = "black"))
#
# # Variáveis
# fviz_pca_var(pca.dfgrpd,
#              col.var = "contrib",
#              gradient.cols = c("navy","gray45", "darkred"),
#              repel = TRUE)+
#   theme(panel.grid = element_blank(),
#         panel.background = element_rect(fill=NULL, colour = "black"))
#
# # Biplot
# fviz_pca_biplot(pca.dfgrpd,
#                 geom.ind = c("point"), #show points only
#                 palette = c("#00AFBB", "#FC4E07", "#665E5E"),
#                 col.ind = dfgrpd$class_efic, # color by groups
#                 addEllipses = TRUE, # Ellipses
#                 ellipse.type= "confidence",
#                 legend.title = "Grupos de eficiências")+
#   theme(panel.grid = element_blank(),
#         panel.background = element_rect(fill=NULL, colour = "black"))
#
# #####
# ##### AGRUPAMENTO #####
# #####
# # Importando dados
# library(readxl)
# df <- read_excel("valores_crs.xlsx",
#                 sheet = "Equivalentes_class_efic",
#                 range = "a1:r82")
# df <- as.data.frame(df)
# df$class_efic <- as.factor(df$class_efic)
# df$class_efic <- as.factor(df$class_efic)
# rownames(df) <- paste(df$class_efic, df$DMU, 1:nrow(df), sep = " ")
#
# normalit<-function(m){
#   (m - min(m))/(max(m)-min(m))}
#
#
# # # df para agrupamento por todas as variáveis em PC
# df.agrup <- df

```

```

#
# library(dplyr)
# df.agrup <- as.data.frame(df.agrup[,-(2:3)] %>%
#                           group_by(DMU, class_efic) %>%
#                           summarise_all(funs(mean)))
#
# df.agrup <- df.agrup %>%
#   mutate(DMU = DMU,
#          class_efic = class_efic,
#          Produtividade = Produção,
#          Pesticidas = Adjuvantes+Herbicidas+Inseticidas+Fungicidas+Espalhantes,
#          Nutrientes = Micronutrientes+N+K2O+P2O5,
#          Sementes = Sementes,
#          Operacionais = Combustível+Mecanização+`Mão-de-obra`)%>%
#   mutate_each(funs(normalit), Produtividade, Pesticidas, Nutrientes,
#              Sementes, Operacionais)
#
# levels(df.agrup$class_efic) <- c("E", "I")
#
# rownames(df.agrup) <- paste(df.agrup$DMU, df.agrup$class_efic, sep = " ")
#
# df.agrup <- select(df.agrup, class_efic, Produtividade,
#                   Pesticidas, Nutrientes, Sementes, Operacionais)
#
#
# # Prin comp analysis
# library(factoextra)
# library(FactoMiner)
# pca.df.agrup <- PCA(df.agrup[, -c(1)],
#                    scale.unit = FALSE,
#                    graph = FALSE,
#                    axes = c(1,2),
#                    ncp = 4)
# get_eig(pca.df.agrup)
#
#
# # Variaveis
# fviz_pca_var(pca.df.agrup,
#             col.var = "contrib",
#             gradient.col = c("#1B118A", "#A68E60", "#FA1212"),
#             repel = TRUE)+
#   theme(panel.grid = element_blank(),
#         panel.background = element_rect(fill = NA, colour = "black"))
#
# # Biplot
# options(OutDec = ",")
# library(ggrepel)
# png(filename = "biplot_dmus.png", width = 12, height = 12,
#     units = "cm", res = 300)
# fviz_pca_biplot(pca.df.agrup,
#               repel = TRUE,
#               labelsize = 2,
#               geom.ind = c("point", "text"),          #show points only
#               #palette = c("#00AFBB", "#FC4E07", "#665E5E"),
#               col.var = "darkred",
#               col.ind = "navy",                      # color by groups
#               #addEllipses = TRUE,                   # Ellipses
#               ellipse.type = "confidence",
#               legend.title = "Grupos de eficiências",
#               title = "" )+
#   labs(x="Componente principal 1 (45,3%)",
#        y="Componente principal 2 (24,4%)")+
#   annotate("text", x=0.6, y=0.7, label='I', size=5)+
#   annotate("text", x=-0.95, y=0.7, label='II', size=5)+
#   annotate("text", x=-0.95, y=-0.07, label='III', size=5)+
#   annotate("text", x=0.6, y=-0.07, label='IV', size=5)+
#   theme(text = element_text(family = "serif"),
#         panel.grid = element_blank(),
#         panel.background = element_rect(fill = NA, colour = "black"),
#         axis.text = element_text(size = 10, colour = "black"),
#         axis.title = element_text(size = 10))
# dev.off()
#
# # Agrupamento por PC

```

```

# hcpc.df.agrup <- HCPC(pca.df.agrup,
#                       method = "average",
#                       graph = FALSE,
#                       min = 3,
#                       max = 3)
#
# # Corte
# library(ggplot2)
# ggsave("wss.png",
#        fviz_nbclust(hcpc.df.agrup$data.clust,
#                     FUNcluster = kmeans,
#                     method = "wss")+
#        labs(title = "", y="Soma de quadrados interna", x="Número de grupos")+
#        geom_vline(xintercept = 3, linetype="dashed", color="gray")+
#        theme(panel.background = element_rect(fill = NA, color = "black", size = 1),
#              panel.grid = element_blank(),
#              axis.text.y = element_text(hjust = 0.5, angle = 90),
#              text = element_text(family = "serif", size = 10)),
#        width = 10, height = 8, units = "cm", dpi = 300)
#
# # Dendrograma
# library(ggsci)
# library(ggrepel)
# library(gridExtra)
#
# png(filename = "dendro.png", width = 15, height = 12,
#      units = "cm", res = 300)
# fviz_dend(hcpc.df.agrup,
#           color_labels_by_k = FALSE,
#           rect = TRUE,
#           lwd = 0.75,
#           cex = 0.5,
#           labels_track_height = 0.5,
#           main = " ") +
#   scale_y_continuous(limits = c(-0.0025,0.02), name = " ") +
#   annotate("segment",
#           x = 3, xend = 20,
#           y = 0.012, yend = 0.012,
#           colour = "blue", lty="dotted", size=0.5)+
#   annotate("text", y=0.012, x=4, label="1", size=5)+
#   annotate("text", y=0.012, x=11, label="2", size=5)+
#   annotate("text", y=0.012, x=19, label="3", size=5)+
#   theme(axis.text = element_blank(),
#         axis.ticks = element_blank(),
#         text = element_text(family = "serif"),
#         plot.background = element_rect(colour = NULL, size = 0.75))
# dev.off()
#
# #
# fviz_cluster(hcpc.df.agrup,
#              axes = c(1,2),
#              geom = c("point", "text"),
#              labels_size = 5,
#              point_size = 1,
#              repel = TRUE,
#              ellipse = TRUE,
#              ellipse.type = "confidence",
#              ellipse.level = 0.95,
#              ellipse.alpha = 0.01,
#              #star.plot=TRUE,
#              show.clust.cent = TRUE,
#              palette="RdBuG",
#              main = "Análise de agrupamento")+
#   scale_x_continuous(limits = c(-1,1), breaks = seq(-1,1,0.5))+
#   scale_y_continuous(limits = c(-0.8,0.8), breaks = seq(-0.8,0.8,0.4))+
#   geom_vline(xintercept = 0, linetype="dashed", color="gray")+
#   geom_hline(yintercept = 0, linetype="dashed", color="gray")+
#   theme(panel.background = element_rect(fill = NA, color = "black", size = 1),
#         panel.grid = element_blank(),
#         text = element_text(family = "serif"),
#         plot.background = element_rect(colour = "black", size = 0.75))
#
# fviz_pca_var(pca.df.agrup,
#              labels_size = 3,
#              axes = c(1,2),

```

```

#           col.var = "contrib",
#           gradient.col = c("#1B118A", "#A68E60", "#FA1212"),
#           repel = TRUE)+
# scale_x_continuous(limits = c(-0.5,0.5), breaks = seq(-0.5,0.5,0.25))+
# scale_y_continuous(limits = c(-0.4,0.4), breaks = seq(-0.4,0.4,0.2))+
# theme(panel.grid = element_blank(),
#       text = element_text(family = "serif"),
#       panel.background = element_rect(fill = NA, colour = "black"),
#       plot.background = element_rect(colour = "black", size = 0.75))
#
# # Agrupamento no plano de PCs
# library(gridExtra)
# library(ggsci)
#
# p1 <- fviz_cluster(hcpc.df.agrup,
#                   axes = c(1,2),
#                   labelsize = 3,
#                   geom = "point",
#                   repel = TRUE,
#                   ellipse = TRUE,
#                   ellipse.type = "confidence",
#                   ellipse.alpha = 0.01,
#                   #star.plot=TRUE,
#                   show.clust.cent = TRUE,
#                   palette="RdBuG",
#                   main = "Análise de agrupamento")+
# geom_vline(xintercept = 0, linetype="dashed", color="black")+
# geom_hline(yintercept = 0, linetype="dashed", color="black")+
# scale_fill_discrete(name = "Agrupamentos", labels = c("1", "2", "3"))+
# scale_color_discrete(name = "Agrupamentos", labels = c("1", "2", "3"))+
# scale_shape_discrete(name = "Agrupamentos", labels = c("1", "2", "3"))+
# labs(x="Componente principal 1 (45,3%)",
#      y="Componente principal 2 (24,4%)",
#      title = " ")+
# guides(colour = guide_legend(title.position = "left",title.hjust = 0.5),
#        fill = guide_legend(title.position = "left",title.hjust = 0.5),
#        shape = guide_legend(title.position = "left",title.hjust = 0.5)) +
# theme(panel.background = element_rect(fill = NA, color = "black"),
#       panel.grid = element_blank(),
#       legend.title = element_text(size = 8, angle = 90),
#       legend.position = "right",
#       legend.direction = "vertical",
#       axis.title= element_text(size = 8, colour = "black"),
#       text = element_text(family = "serif"))
#
# p2 <- fviz_pca_var(pca.df.agrup,
#                   axes = c(1,2),
#                   labelsize=2,
#                   #col.var = "contrib",
#                   gradient.col = c("#1B118A", "#A68E60", "#FA1212"),
#                   repel = TRUE)+
# labs(x="Componente principal 1 (45,3%)",
#      y="Componente principal 2 (24,4%)",
#      title = " ")+
# theme(text = element_text(family = "serif"),
#       panel.grid = element_blank(),
#       panel.background = element_rect(fill = NA, colour = "black"),
#       axis.text = element_text(size = 10, colour = "black"),
#       axis.title = element_text(size = 8))
#
# p3 <- fviz_cluster(hcpc.df.agrup,
#                   axes = c(1,3),
#                   labelsize = 3,
#                   geom = "point",
#                   repel = TRUE,
#                   ellipse = TRUE,
#                   ellipse.type = "confidence",
#                   ellipse.alpha = 0.01,
#                   #star.plot=TRUE,
#                   show.clust.cent = TRUE,
#                   palette="RdBuG",
#                   main = "Análise de agrupamento")+
# geom_vline(xintercept = 0, linetype="dashed", color="black")+
# geom_hline(yintercept = 0, linetype="dashed", color="black")+

```

```

# scale_fill_discrete(name = "Agrupamentos", labels = c("1", "2", "3"))+
# scale_color_discrete(name = "Agrupamentos", labels = c("1", "2", "3"))+
# scale_shape_discrete(name = "Agrupamentos", labels = c("1", "2", "3"))+
# labs(x="Componente principal 1 (45,3%)",
#      y="Componente principal 3 (19,7%)",
#      title = "")+
# guides(colour = guide_legend(title.position = "left",title.hjust = 0.5),
#        fill = guide_legend(title.position = "left",title.hjust = 0.5),
#        shape = guide_legend(title.position = "left",title.hjust = 0.5)) +
# theme(panel.background = element_rect(fill = NA, color = "black"),
#        panel.grid = element_blank(),
#        legend.title = element_text(size = 8, angle = 90),
#        legend.position = "right",
#        legend.direction = "vertical",
#        axis.title= element_text(size = 8, color = "black"),
#        text = element_text(family = "serif"))
#
# p4 <- fviz_pca_var(pca.df.agrup,
#                   axes = c(1,3),
#                   labels=2,
#                   #col.var = "contrib",
#                   gradient.col = c("#1B118A", "#A68E60", "#FA1212"),
#                   repel = TRUE)+
# labs(x="Componente principal 1 (45,3%)",
#      y="Componente principal 3 (19,7%)",
#      title = "")+
# theme(text = element_text(family = "serif"),
#        panel.grid = element_blank(),
#        panel.background = element_rect(fill = NA, colour = "black"),
#        axis.text = element_text(size = 10, colour = "black"),
#        axis.title = element_text(size = 8))
#
# p5 <- fviz_cluster(hcpc.df.agrup,
#                   axes = c(2,3),
#                   labels= 3,
#                   geom = "point",
#                   repel = TRUE,
#                   ellipse = TRUE,
#                   ellipse.type = "confidence",
#                   ellipse.alpha = 0.01,
#                   #star.plot=TRUE,
#                   show.clust.cent = TRUE,
#                   palette="RdBuG",
#                   main = "Análise de agrupamento")+
# geom_vline(xintercept = 0, linetype="dashed", color="black")+
# geom_hline(yintercept = 0, linetype="dashed", color="black")+
# scale_fill_discrete(name = "Agrupamentos", labels = c("1", "2", "3"))+
# scale_color_discrete(name = "Agrupamentos", labels = c("1", "2", "3"))+
# scale_shape_discrete(name = "Agrupamentos", labels = c("1", "2", "3"))+
# labs(x="Componente principal 2 (24,4%)",
#      y="Componente principal 3 (19,7%)",
#      title = "")+
# guides(colour = guide_legend(title.position = "left",title.hjust = 0.5),
#        fill = guide_legend(title.position = "left",title.hjust = 0.5),
#        shape = guide_legend(title.position = "left",title.hjust = 0.5)) +
# theme(panel.background = element_rect(fill = NA, color = "black"),
#        panel.grid = element_blank(),
#        legend.title = element_text(size = 8, angle = 90),
#        legend.position = "right",
#        legend.direction = "vertical",
#        axis.title= element_text(size = 8, colour = "black"),
#        text = element_text(family = "serif"))
#
# p6 <- fviz_pca_var(pca.df.agrup,
#                   axes = c(2,3),
#                   labels=2,
#                   #col.var = "contrib",
#                   gradient.col = c("#1B118A", "#A68E60", "#FA1212"),
#                   repel = TRUE)+
# labs(x="Componente principal 2 (24,4%)",
#      y="Componente principal 3 (19,7%)",
#      title = "")+
# theme(text = element_text(family = "serif"),
#        panel.grid = element_blank(),
#        panel.background = element_rect(fill = NA, colour = "black"),

```

```

#       axis.text = element_text(size = 10, colour = "black"),
#       axis.title = element_text(size = 8))
#
# p6
#
# lay <- rbind(c( 1, 1, 1, 1,NA,2,2,2),
#             c( 1, 1, 1, 1,NA,2,2,2),
#             c( 1, 1, 1, 1,NA,2,2,2),
#             c( 1, 1, 1, 1,NA,2,2,2),
#             c( 1, 1, 1, 1,NA,2,2,2),
#             c( 1, 1, 1, 1,NA,2,2,2),
#             c( 3, 3, 3, 3,NA,4,4,4),
#             c( 3, 3, 3, 3,NA,4,4,4),
#             c( 3, 3, 3, 3,NA,4,4,4),
#             c( 3, 3, 3, 3,NA,4,4,4),
#             c( 3, 3, 3, 3,NA,4,4,4),
#             c( 3, 3, 3, 3,NA,4,4,4),
#             c( 5, 5, 5, 5,NA,6,6,6),
#             c( 5, 5, 5, 5,NA,6,6,6),
#             c( 5, 5, 5, 5,NA,6,6,6),
#             c( 5, 5, 5, 5,NA,6,6,6),
#             c( 5, 5, 5, 5,NA,6,6,6),
#             c( 5, 5, 5, 5,NA,6,6,6)
# )
#
# library(gridExtra)
# png(filename = "agrup_planos.png", width = 16, height = 19,
#     units = "cm", res = 300)
# grid.arrange(grobs=list(p1,p2,p3,p4,p5,p6), layout_matrix=lay)
# dev.off()
#

```