



UNIVERSIDADE
ESTADUAL de LONDRINA

LARA MARIE GUANAIS SANTOS

**GERAÇÃO DE MAPAS DE ZONAS DE POTENCIAL PRODUTIVO UTILIZANDO
DADOS DE PRODUTIVIDADE E SUA RELAÇÃO COM VARIÁVEIS DE SOLO**

Londrina
2024

LARA MARIE GUANAIS SANTOS

GERAÇÃO DE MAPAS DE ZONAS DE POTENCIAL PRODUTIVO UTILIZANDO
DADOS DE PRODUTIVIDADE E SUA RELAÇÃO COM VARIÁVEIS DE SOLO

Tese apresentada ao Programa de Pós-Graduação em Agronomia da Universidade Estadual de Londrina, como requisito parcial à obtenção do título de Doutor em Agronomia.

Orientadora: Prof^a. Dr^a. Maria de Fátima Guimarães.

Coorientador: Dr. Hevandro Colonhese Delalibera – IDR - Paraná

Londrina
2024

Ficha de identificação da obra elaborada pelo autor, através do Programa de Geração Automática do Sistema de Bibliotecas da UEL

L318g Guanais Santos, Lara Marie.
GERAÇÃO DE MAPAS DE ZONAS DE POTENCIAL PRODUTIVO UTILIZANDO DADOS DE PRODUTIVIDADE E SUA RELAÇÃO COM VARIÁVEIS DE SOLO / Lara Marie Guanais Santos. - Londrina, 2024.
61 f. : il.

Orientador: Profa. Dra. Maria de Fátima Guimarães.
Coorientador: Dr. Hevandro Colanhese Delalibera.
Tese (Doutorado em Agronomia) - Universidade Estadual de Londrina, Centro de Ciências Agrárias, Programa de Pós-Graduação em Agronomia, 2024.
Inclui bibliografia.

1. Zonas de manejo - Tese. 2. Colheita de grãos - Tese. 3. Atributos do solo - Tese. 4. Método Jackknife - Tese. I. Guimarães, Profa. Dra. Maria de Fátima. II. Delalibera, Dr. Hevandro Colanhese. III. Universidade Estadual de Londrina. Centro de Ciências Agrárias. Programa de Pós-Graduação em Agronomia. IV. Título.

CDU 63

LARA MARIE GUANAIS SANTOS

GERAÇÃO DE MAPAS DE ZONAS DE POTENCIAL PRODUTIVO UTILIZANDO
DADOS DE PRODUTIVIDADE E SUA RELAÇÃO COM VARIÁVEIS DE SOLO

Tese apresentada ao Departamento de
Agronomia da Universidade Estadual de
Londrina, como requisito parcial à obtenção
do título de Doutora em Agronomia.

BANCA EXAMINADORA

Orientadora: Profa. Dra. Maria de Fátima
Guimarães.
Universidade Estadual de Londrina - UEL

Prof. Dr. Gustavo Di Chiacchio Faulin
Faculdade de Tecnologia "Shunji Nishimura"
de Pompeia – FATEC Pompeia

Prof. Dr. Ricardo Ralisch
Universidade Estadual de Londrina – UEL

Profa. Dra. Marisa Silveira Almeida Renaud
Faulin
Faculdade de Tecnologia "Shunji Nishimura"
de Pompeia – FATEC Pompeia

Prof. Dr. Alexandre de Moura Guimarães
Faculdade de Tecnologia "Shunji Nishimura"
de Pompeia – FATEC Pompeia

Londrina, 29 de abril de 2024.

***Dedico este trabalho aos meus pais
Ivone Guanais e José Sebastião e ao
meu esposo Caio Fernando.***

AGRADECIMENTOS

À Deus, pilar em minha vida. Sem Ele nada seria possível.

Aos meus pais Ivone Guanais Mineiro Santos e José Sebastião dos Santos, que com muito esforço e dedicação venceram os momentos difíceis e sempre me proporcionaram o melhor.

Ao meu esposo Caio Fernando dos Santos Oliveira, por todo apoio e companheirismo.

Aos meus orientadores e mentores Maria de Fátima Guimaraes e Hevandro Colhese Delalibera pela generosidade em contribuir para meu desenvolvimento intelectual, pessoal e de minha pesquisa.

A todos os docentes da Universidade Estadual de Londrina pelo conhecimento cedido a mim, em especial aos professores Maria de Fátima Guimaraes, João Tavares Filho, Maurício Ursi Ventura e Claudemir Zucareli pelo carinho e apoio.

Ao programa de pós graduação em Agronomia pela oportunidade de desenvolver minha pesquisa nesta universidade maravilhosa.

À CAPES, pela concessão da bolsa.

Ao IDR – Paraná de Londrina, pela concessão do método analítico de análise de dados e álgebra de mapas.

Ao senhor Agnaldo Favoni e seus filhos Vinicius Andrade Favoni e Douglas Andrade Favoni pela concessão dos dados.

À Secretária Helenice, pela paciência e todo auxílio.

Aos técnicos do Laboratório de Solos Marcio Praxedes e João Machado, pela amizade e auxílio na realizações das análises.

Ao Laboratório de solos do IDR – Paraná de Londrina, pelo auxílio na execução das análises físicas, em especial à Profa. Dra. Graziela Moraes de Cesare Barbosa e os analistas Vilson José Teixeira Rozendo e Dr. José Francirlei de Oliveira.

À companheira de pós Tamires Firmino e ao meu esposo pelo auxílio na coleta de minhas amostras de solo.

Aos amigos da pós-graduação Fernanda Paduan, Gabriel Shimizu, Smaylla El Kadri, Gianne Gidodi, Tamires Firmino e Stephanie Locatelli pela amizade e apoio em momentos difíceis e de alegrias.

Às estagiárias do laboratório de solo pelo auxílio.

Aos professores Adriana da Silva e Thadeu Rodrigues de Melo por me auxiliarem nas minhas análises, com a escrita e conselhos.

Aos meus amigos de trabalho da Faculdade de tecnologia de Pompéia – Fatec Shunji Nishimura de Pompéia pelo carinho, apoio e acolhida sempre.

Serei eternamente grata!

***A vida me ensinou a nunca desistir,
nem ganhar, nem perder mas procurar
evoluir.***

- Charlie Brown Jr.

GUANAIS SANTOS, Lara Marie. **Geração de mapas de zonas de potencial produtivo utilizando dados de produtividade e sua relação com variáveis de solo.** 2024. 61 folhas. Doutorado (Doutorado em Agronomia) – Universidade Estadual de Londrina, Londrina, 2024.

RESUMO

A definição de unidade de gestão diferenciada pode ser feita com base em dados de diversas fontes como, imagens aéreas, análise de solo, uso de sensores terrestres e mapas de produtividade. Os mapas de produtividade, cujos dados são obtidos por meio de sensores acoplados às colhedoras, possibilitam a obtenção da produtividade de uma área georreferenciada com grande densidade de dados. Contudo, em meio a estes dados, existem aqueles que não representam o comportamento do fenômeno colheita e, portanto, devem ser eliminados para não alterarem a qualidade e a confiabilidade da informação fornecida pelo mapa. Este trabalho tem por intuito validar uma metodologia de tratamento, álgebra de mapas e segregação em zonas de manejo de mapas a partir de dados de produtividade advindos de monitores de colheita, executado com base em procedimentos estatísticos consagrados. Para este trabalho foi confeccionado um mapa de potencial produtivo de uma área agrícola localizada em Cândido Mota-SP; estes mapas, feitos a partir dos históricos de safras de soja e milho obtidos por monitor de colheita, foram utilizados como base para a distribuição dos pontos de amostragem para a análise físico-química do solo. Foi possível concluir que o estudo avaliando da relação entre o mapa de zonas de potencial produtivo, gerado por meio da metodologia proposta por Guanais Santos, et al. (2023) utilizando mapas de produtividade como dados de entrada, com as variáveis de solo, apresentou relação de 84%. Conclui-se também que as variáveis com maior poder de modelamento estão ligadas a estrutura do solo e a movimentação de água e ar, sendo a macroporosidade a mais importante. Por fim, também pode-se afirmar que a metodologia de separação em zonas de potencial produtivo é útil para a alocação e redução do número de pontos amostrais em uma área agrícola, direcionando a amostragem por meio de informação quantitativa, evitando amostragens em grade, de forma aleatória e empírica e em grande quantidade.

Palavras-chave: Unidades de gestão diferenciadas; Colheita de grãos; Atributos do solo; Zonas de manejo; Método Jackknife

GUANAIS SANTOS, Lara Marie. **Generation of maps of productive potential zones using productivity data and its relationship with soil variables.** 2024. 61 sheets. PhD (PhD in Agronomy) - State University of Londrina, Londrina, 2024.

ABSTRACT

The definition of a differentiated management unit can be made based on data from different sources such as aerial images, soil analysis, use of terrestrial sensors and productivity maps. Productivity maps, whose data are obtained through sensors attached to harvesters, make it possible to obtain the productivity of a georeferenced area with a high density of data. However, among this data, there are those that do not represent the behavior of the harvest phenomenon and, therefore, must be eliminated so as not to alter the quality and reliability of the information provided by the map. This work aims to validate a treatment methodology, map algebra and segregation into map management zones based on productivity data from harvest monitors, carried out based on established statistical procedures. For this work, a map of the productive potential of an agricultural area located in Cândido Mota-SP was created; These maps, made from historical soybean and corn harvests obtained by a harvest monitor, were used as a basis for the distribution of sampling points for the physical-chemical analysis of the soil. It was possible to conclude that the study evaluating the relationship between the map of productive potential zones, generated through the methodology proposed by Guanais Santos, et al. (2023) using productivity maps as input data, with soil variables, presented a relationship of 84%. It is also concluded that the variables with the greatest modeling power are linked to soil structure and the movement of water and air, with macroporosity being the most important. Finally, it can also be stated that the methodology of separation into zones of productive potential is useful for allocating and reducing the number of sampling points in an agricultural area, directing sampling through quantitative information, avoiding grid sampling, of randomly and empirically and in large quantities.

Keywords: Differentiated management units; Grain harvest; Soil attributes; Management areas; Jackknife method.

FIGURAS

Figura 1 – Sensores necessários para obtenção de um mapa de produtividade.	17
Figura 2 – Distribuição normal padrão (Laplace-Gauss ou Gaussiana) e probabilidades do contido.	20
Figura 3 – Mapas de demonstração da área agrícola utilizada para o estudo.	23
Figura 4 – Mapas de escore-z com a separação nas tres zonas de potencial produtivo, para cada cultura e safra.	25
Figura 5 – Mapa de zonas de potencial produtivo e sua subdivisão em tres classes com a distribuição dos 21 pontos de coleta de solo.	26
Figura 6 – Diagrama da representação das grandezas das variáveis de acordo com a amostra	28
Figura 7 – <i>Representação gráfica da escala ou grandeza do conjunto de variáveis para cada amostra.</i>	30
Figura 8 – Exemplo de algumas variáveis consideradas covariantes, como matéria orgânica com carbono orgânico e densidade do solo com porosidade total do solo. .	31
Figura 9 – Representação das variáveis consideradas multicolineares.	32
Figura 10 – Poder de Modelamento das variáveis para o conjunto avaliado na PCA.	33
Figura 11 – Disposição gráfica bidimensional das componentes 1 e 2, 1 e 3 e 2 e 3, para os scores a esquerda e para os loadings a direita.	35
Figura 12 – Representação gráfica do grau de poder de modelagem dos loadings dentro do modelo PCA.	37
Figura 13 – Representação gráfica das projeções de classes apresentada pelo modelo SIMCA.	38
Figura 14 – Representação gráfica da distância entre classes para as relações entre as três componentes dimensionais.	39
Figura 15 – Poder de discriminação ou diferenciação entre as variáveis, conforme apresentado pela SIMCA.	40
Figura 16 – Disposição gráfica do poder de modelagem das variáveis para a estimativa geral das classes nominais.	41
Figura 17 - Histogramas dos índices de produtividade da cultura da soja e diagrama de frequência ordenada dos índices de desvio padronizado e seus respectivos graus de ajuste da reta, referente ao fenômeno colheita mecanizada de grãos antes e após da aplicação do método jackknife.	46
Figura 18 - Histogramas dos índices de produtividade da cultura do milho e diagrama	

de frequência ordenada dos índices de desvio padronizado e seus respectivos graus de ajuste da reta, referente ao fenômeno colheita mecanizada de grãos antes e após da aplicação do método jackknife.	47
Figura 19 - Histogramas dos índices de produtividade do mapa de zonas de potencial produtivo e diagrama de frequência ordenada dos índices de desvio padronizado e seus respectivos graus de ajuste da reta, referente ao fenômeno colheita mecanizada de grãos antes e após da aplicação do método jackknife.....	48
Figura 20 - Representação gráfica da disposição das amostras a direita e das variáveis direita para a classe de alto potencial produtivo.	49
Figura 21 - Disposição gráfica do poder de modelagem das variáveis para a classe de alto potencial produtivo.	50
Figura 22 - Representação gráfica da disposição das amostras a direita e das variáveis direita para zona de médio potencial produtivo.....	51
Figura 23 - Disposição gráfica do poder de modelagem das variáveis para a zona de médio potencial produtivo.	52
Figura 24 - Representação gráfica da disposição das amostras a esquerda e das variáveis direita para zona de baixo potencial produtivo.....	53
Figura 25 - Disposição gráfica do poder de modelagem das variáveis para zona de baixo potencial produtivo.	54

TABELAS

Tabela 1 – Predição analítica de enquadramento das amostras com base nas três classes (Zonas de Potencial Produtivo) cada amostra se enquadra (Best) e qual outra classe a referida amostra poderia ser realocada (NextBest)	43
Tabela 2 – Modelo de validação quantitativa do enquadramento das amostras nas classes analíticas nominado como Misclassification.....	44
Tabela 3 – Variáveis utilizadas nos modelos multivariados com suas respectivas zonas de potencial produtivo.	61

SUMÁRIO

RESUMO.....	9
ABSTRACT.....	10
1. INTRODUÇÃO.....	14
2. REVISÃO BIBLIOGRÁFICA	15
2.1 AGRICULTURA DE PRECISÃO	15
2.2 MAPAS DE PRODUTIVIDADE.....	16
2.3 INCONSISTÊNCIA E ERROS NOS DADOS DE PRODUTIVIDADE	18
2.4 MÉTODO ESTATÍSTICO PARA TRATAMENTO E INTERPRETAÇÃO DE DADOS.....	19
2.5 INTERPOLAÇÃO DOS DADOS E ÁLGEBRA DE MAPAS	21
2.6 INTERPRETAÇÃO DE MAPAS DE PRODUTIVIDADE.....	22
3. MATERIAL E MÉTODOS.....	23
4. RESULTADOS E DISCUSSÃO.....	27
CONCLUSÃO	54
REFERÊNCIAS.....	55
ANEXO.....	61

1. INTRODUÇÃO

A agricultura de precisão (AP) tornou-a uma ferramenta amplamente utilizada para auxiliar na averiguação das características e atributos das áreas agrícolas, bem como para tratá-las de maneira individualizadas por meio de critérios de variabilidade espacial e temporal.

Esta variabilidade pode ser oriunda de fatores físicos, químicos, biológicos e de outros fatores. Com base em informações ou apontamentos georreferenciados torna-se mais fácil visualizar o problema e efetuar o controle localizado, fornecendo a solução mais adequada para cada local e situação.

A subdivisão das áreas em regiões de características diferentes é chamada de unidades de gestão diferenciada (UGD). Para esta subdivisão da área pode-se utilizar mapas de produtividade. Assim, as informações oriundas de colhedoras instrumentadas se mostra como importante ferramenta, pois, trata-se de informação bastante detalhada que aponta a resposta da cultura para todos os fatores, sendo eles estáveis, com pouca variabilidade dentro das safras, ou variáveis, com ocorrência diferenciada dentro das safras.

Porém, nas informações concebidas pelo monitor de colheita, podem existir valores que não representam a realidade local e, portanto, não representam a população estatística dos dados e podem comprometer a análise posterior dos mapas. Existem diversos métodos que visam identificar estes dados, chamados de *outliers* ou pontos discrepantes e que devem ser excluídos para obter uma análise mais representativa do fenômeno avaliado. Porém, esta exclusão deve ser feita com critério, utilizando métodos adequados à natureza e ao conjunto de dados.

Assim, na hipótese de que o mapa de zonas de potencial produtivo tem correlação com as variáveis de solo e é adequado para a realização de amostragem e de inferência e recomendação de manejo da área, este trabalho tem por objetivo verificar a existência de relação entre a metodologia de tratamento e interpretação de dados de produtividade proposta por Guanais Santos, et al. (2023), com variáveis de solo, visando sua validação quanto a utilidade na inferência de manejo na área.

2. REVISÃO BIBLIOGRÁFICA

2.1 AGRICULTURA DE PRECISÃO

Melhorar a produção e a qualidade do rendimento das culturas, reduzindo os custos operacionais e a poluição ambiental, é o objetivo fundamental na agricultura de precisão (AP) (CHLINGARYAN, 2018; AWAN, 2016). O potencial de crescimento e rendimento depende de muitos atributos de produção diferentes, como clima, propriedades do solo, topografia, irrigação e manejo de fertilizantes (CHLINGARYAN, 2018). A necessidade de gestão oportuna e precisa desses insumos para grandes campos agrícolas levou ao aumento da adoção de tecnologias de sensoriamento remoto e proximal (CAMPBELL e WYNNE, 2011).

A AP foi introduzida no Brasil nos anos 80 na aplicação de técnicas para racionalizar o uso de fertilizantes, variando as taxas dentro do campo, a fim de atender às necessidades diferenciadas de cada parte da área. Isto foi possível devido a melhores métodos de investigação, como levantamentos de solo e mapeamento, fotografia aérea (ROBERT, 2002).

AP pode ser definida como uma estratégia de gerenciamento (LI e CHUNG, 2015), que utiliza técnicas e sensores para identificar a variabilidade espacial e temporal das lavouras, auxiliando a gestão e manejo das áreas agrícolas (ZHANG e KOVACS, 2012).

Para aplicação da AP existem tecnologias como, o *Global Navigation Satellite System* (GNSS), ou sistema global de navegação por satélite, sistemas de informação geográficas (SIG) e diversos sensores para avaliar a variabilidade do local, do solo e da planta, que fornecem informações para ajudar os produtores no gerenciamento mais preciso e assertivo do sistema agrícola (PATHAK et al., 2019). Um dos benefícios gerados é a redução de custos da produção agrícola, proporcionando um aumento da competitividade nos cenários nacional e internacional e benefícios ao meio ambiente (RIBEIRO, et al. 2020).

Em um modelo tradicional de gerenciamento das propriedades, as áreas são tratadas como homogêneas (SRINIVASAN, 2006), ou seja, a variabilidade da área não é considerada. Com isso, as operações são realizadas de maneira uniforme, independentemente da possível heterogeneidade do campo. Embora esse tipo de manejo seja mais rápido e de fácil implementação, implica em

custos desnecessários devido à aplicação inadequada de insumos e não permite o total aproveitamento do potencial da área (KOUTSOS e MENEXES, 2017).

Ao longo dos anos foram realizados inúmeros estudos demonstrando os benefícios da substituição das práticas tradicionais pelas tecnologias de AP, apontando os ganhos com a utilização desta tecnologia (BOYER et al. 2011; PANAGOPOULOS et al. 2014; WEST e KOVACS, 2017; WACHOWIAK et al., 2017).

2.2 MAPAS DE PRODUTIVIDADE

Alcançar o máximo rendimento das culturas a um custo mínimo com um ecossistema saudável é um dos principais objetivos da produção agrícola. A detecção precoce e o gerenciamento de problemas associados a restrições de rendimento de culturas podem ajudar a aumentar a produtividade e o lucro subsequente, e estimar o rendimento é importante para as decisões de negócios e gerenciamento de culturas (CHLINGARYAN, 2018).

Desde o início da AP, o monitoramento do rendimento está entre as ferramentas ou técnicas mais praticadas pelos produtores para examinar a variabilidade espacial em seus campos agrícolas (KHOSLA et al., 2010). Em uma pesquisa realizada pela Corporação de Pesquisa e Desenvolvimento de Grãos (GRDC), em 2014, na Austrália, constatou-se que a adoção média nacional do mapeamento de rendimento era de 29% da área cultivada (UMBERS et al. 2015).

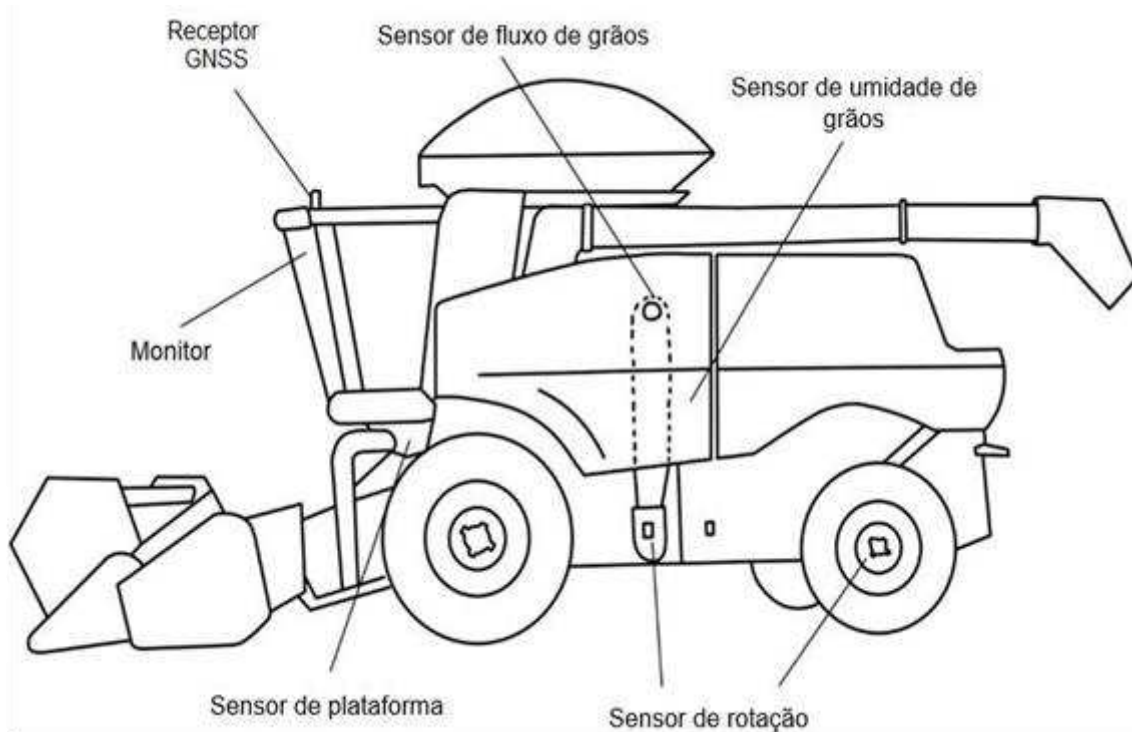
No Brasil, cerca de 45% dos produtores de grãos utilizam pelo menos uma técnica da AP. Destes, cerca de 31% avaliam a produtividade dos talhões através de mapas de produtividade obtidos por monitores de colheita (MOLIN, 2017).

Com os avanços tecnológicos da AP, os mapas de produtividade tornaram-se economicamente acessíveis aos agricultores, porém, é necessário que a operação de colheita seja realizada por uma colhedora equipada com monitor de rendimento (SOUZA et al., 2016).

O termo monitor de colheita representa um conjunto de componentes acoplados à colhedora, como: sensor de produtividade, sensor de umidade de grãos, sensor de velocidade de colheita, sensor indicador da posição da plataforma da colhedora, receptor GNSS e o monitor que está conectado a todos

os sensores citados, (SILVA et al., 2011). Os componentes constam da Figura 1.

Figura 1 – Sensores necessários para obtenção de um mapa de produtividade.



Fonte: Próprio autor

Os mapas de produtividade integram os efeitos de várias variáveis espaciais, como propriedades do solo, taxas de fertilizantes, atributos topográficos, condições climáticas e ocorrência de infestações por insetos e doenças (SOUZA et al., 2016). Estes mapas de rendimento em si não são ferramentas de conhecimento ou decisão, para tanto, é necessário incorporá-los aos processos de tomada de decisão, análise e planejamento geral das operações da propriedade (SOUZA et al., 2016).

Para obtenção de um bom mapa é necessário decidir como os dados serão interpolados. Alguns métodos não preservam necessariamente a variabilidade do conjunto de dados original, em função de uma estimativa dos intervalos entre os pontos amostrais e podem produzir resultados diferentes para o mesmo conjunto de dados (SOUZA et al., 2016).

Na literatura existem técnicas que avaliam a relação entre rendimento e atributos do solo, sendo eles fatores químicos, físicos ou em relação a declividade, para identificar as principais limitações da produção agrícola em uma

determinada área ou região (HORNUNG et al., 2006), como Kravenchenko e Bullock (2000), Sudduth et al. (1995) e Coleman e Tadesse (1995). Com essa relação, é possível estabelecer as áreas com potencial produtivo diferenciadas.

2.3 INCONSISTÊNCIA E ERROS NOS DADOS DE PRODUTIVIDADE

Durante a coleta de dados pela colhedora podem ocorrer erros (MENEGATTI e MOLIN, 2004). Estes erros podem ser oriundos da dinâmica de colheita da colhedora, medição constante da umidade e rendimento, precisão da informação de posicionamento e erros do operador da máquina.

Segundo Moore (1998), Molin, 2002 e Lyle et al. (2014), o tempo de enchimento da colhedora é aquele gasto entre o corte da planta pela barra de corte da plataforma e a passagem do grão pelo sensor de produtividade. O fluxo de grãos medido no sensor de produtividade deve ser correlacionado com a posição da colhedora quando a cultura foi colhida. Este tempo é pré-estabelecido pelo usuário ao monitor de colheita. Caso se estabeleça um tempo incorreto para esse processo, isto resultará em erros no cálculo da produtividade com efeito de deslocamento da sua coordenada.

Da mesma forma, a velocidade da colhedora é um grande causador de erros de leitura induzidos pelos operadores das máquinas. Alterações repentinas de velocidade afetam significativamente a precisão das medições de produtividade, o que pode ocasionar leituras de produtividades discrepantes em distâncias próximas (LYLE et al., 2014).

Ao finalizar a passada da máquina, há o tempo de esvaziamento da colhedora, que é o mesmo processo do tempo de enchimento, mas em sentido contrário, ou seja, o operador começa a retardar o processo de colheita no final da passada da colhedora para realizar a manobra de cabeceira, e o material que já havia sido colhido termina de ser trilhado (MICHELAN, 2007).

Outra possível fonte de erro é a diminuição da quantidade de satélites enviando sinal para o receptor GNSS ou do sinal de correção, isto pode acontecer em áreas próximas a galpões, ou em dias com muitas nuvens, ou também de acordo com o tipo de sinal utilizado (RTX ou RTK). Este erro tem relação com a interferências da ionosfera e troposfera, erros do relógio do satélite e do receptor, erros na órbita dos satélites e erros de reflexão do sinal ou

multicaminhamento (MICHELAN, 2007).

A delimitação errada da largura de plataforma também é considerada uma fonte de erro, pois implica em uma estimativa normalmente inferior de produtividade, uma vez que o cálculo é feito multiplicando-se a largura da barra de corte da colhedora pela distância percorrida num determinado tempo (MALDANER ET AL. 2018).

Menegatti e Molin (2002) cita a umidade dos grãos como sendo outro fator de erro em mapas de produtividade, que deve ser considerado na interpretação, porque ela é variável no talhão, seja por uma heterogeneidade no estágio fenológico da cultura ou por variações durante os períodos do dia, devido a umidade relativa do ar e temperatura do ambiente. Este problema é agravado quando a colheita é iniciada em condições de alta umidade e, no decorrer do processo, os grãos secam a uma umidade constante.

Outros trabalhos também apontam a importância da calibração dos sensores presentes na colhedora, evitando a aferição incorreta do material colhido (BAZZI ET AL., 2008; SPEKKEN ET AL., 2013; MALDANER ET AL. 2018).

Tendo em vista todas essas fontes de erro torna-se necessária realização de procedimentos analíticos que mitiguem o efeito negativo dos erros na qualidade dos dados sem que estes gerem informações que não representem a realidade do evento avaliado. Com essa prática é possível reduzir a quantidade de erros nos conjuntos de dados, e conseqüentemente, tornar os mapas mais confiáveis (MENEGATTI e MOLIN, 2004; MICHELAN et al., 2007).

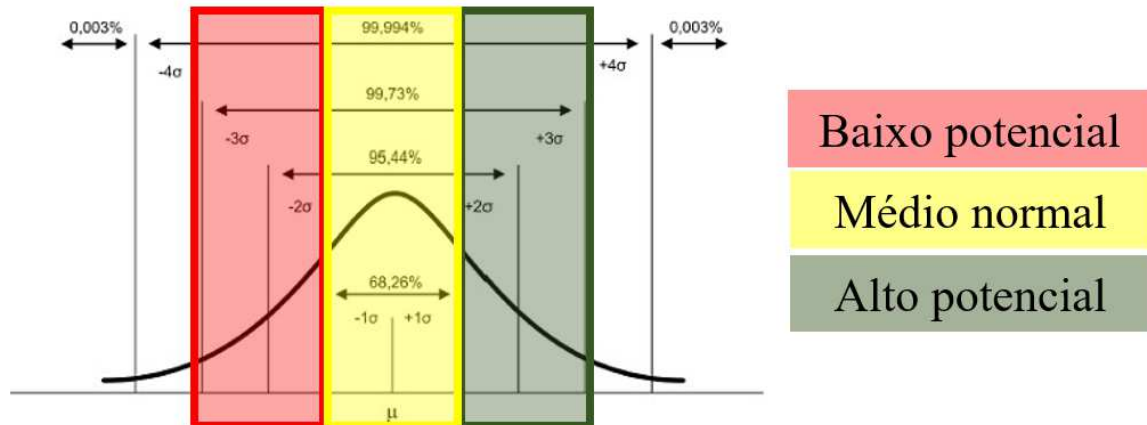
2.4 MÉTODO ESTATÍSTICO PARA TRATAMENTO E INTERPRETAÇÃO DE DADOS

O método *Jackknife*, proposto por Wu *et al.* (1986), é composto por análises de desvios, como por exemplo os Desvios Padronizados, baseado no teorema de Laplace, ou os Desvios Studentizados (ou teorema do limite central), onde ambos efetuam a remoção de amostras consideradas não pertencentes a população.

A análise dos Desvios Padronizados desenvolvida pela estatística *z*, tem como média 0 e desvio padrão 1 para dados absolutamente contínuos se obtendo os escores-*z* (Equação 1) e, conforme Wu *et al.* (1986) pode ser utilizada como método de base estatística para exclusão de amostras consideradas não

pertencentes à população analisada, utilizando como parâmetros os valores críticos da distribuição teórica da tabela “z” (HARTER, 1960), sendo o método de exclusão dos dados fora do valor crítico pré determinado pelo analista, chamado de *Jackknife*

Figura 2 – Distribuição normal padrão (Laplace-Gauss ou Gaussiana) e probabilidades do contido.



Fonte: adaptado de Portal Action; Onde, a cor vermelha remete a classificação da área em baixo potencial produtivo, em amarelo a classificação da área em médio potencial produtivo e a cor verde na classificação da área em alto potencial produtivo. Classificação apresentada posteriormente no trabalho.

Eq. 1

$$Z = \frac{x_i - \bar{x}_i}{\sigma}$$

Sendo:

Z – escore-z

x_i – Amostra

\bar{x}_i – Média amostral ou do tratamento

σ – Desvio padrão

Para esta análise pela estatística z, aplicou-se o valor crítico de -3,09 a 3,09 (bilateral), ou seja, tem 99,8% de probabilidade de estar contido dentro da população normal padrão distribuída, nos quais, dados que apresentam índice escore-z fora deste intervalo são considerados como não pertencentes à população e, portanto, devem ser excluídos (SNEDECOR e COCHRAN, 1980).

É importante que estes métodos de análise e remoção de dados sejam realizados em planilha eletrônica aplicando cálculos referenciados para

que, após cada remoção de valor extremo, os índices sejam recalculados, pois, com a remoção dos dados, as medidas de tendência central podem se mover de posição dentro da curva de distribuição, bem como os índices finais calculados a partir destas medidas (DELALIBERA *et al.*, 2017).

Também é de suma importância iniciar sempre a remoção pelos índices mais extremos, pois, a cada remoção de valores discrepantes, os índices podem ser movidos para o centro ou para a extremidade da distribuição, podendo, para os casos onde os valores se apresentem no limiar de corte, serem movidos para dentro ou para fora do intervalo crítico de corte escolhido (DELALIBERA *et al.*, 2017).

Ao final da aplicação do método *Jackknife* (WU *et al.*, 1986), é possível observar que os índices resultantes são um escalamento de medidas dos dados entre $\pm 3,09$, com média 0 e desvio padrão 1.

Além de aplicar *Jackknife* reamostragem, é importante avaliar o grau de linearidade da frequência ordenada de índices, z -score, para cada mapa de colheita, que é facilmente obtido ordenando z -score do mais baixo para o mais alto e aplicando uma dispersão para ajustar a equação linear, onde o grau de linearidade é representado pelo R^2 e os dados ajustam-se a uma linha reta (GUANAIS SANTOS *et al.*, 2023).

2.5 INTERPOLAÇÃO DOS DADOS E ÁLGEBRA DE MAPAS

A interpolação de dados é um método geoestatístico para estimar valores, ou seja, realiza estimativas entre os intervalos dos pontos amostrais. Um passo importante na geração de um bom mapa temático é o método de interpolação de dados.

Dois dos métodos comuns para interpolação de dados incluem o inverso da distância (IDW) e krigagem (ISAAKS e SRIVASTAVA, 1989). Outros métodos incluem mínima curvatura (BAKKALI e AMRANI, 2008), polígonos de Thiessen, vizinhos mais próximos, regressão polinomial (GOODALE *et al.*, 1998) e média móvel (GALLICHAND *et al.*, 1992).

A krigagem, originalmente desenvolvida por Krige (1951), onde para ser usada é necessário que exista dependência espacial, aferida pelo semivariograma (SILVA, *et al.*, 2008), partindo do princípio que as médias não são

constantes. Uma variável regionalizada não estacionária pode ser considerada como constituída por dois componentes: 1) a tendência, que consiste no valor médio esperado dessa variável, dentro de uma certa vizinhança, que varia sistematicamente; e 2) o resíduo, que é a diferença entre os valores reais e o valor do deslocamento causado pela tendência. Nestes casos, a krigagem é executada sobre os resíduos, o que impõe o ajuste do variograma sobre os resíduos para que se possa descrever a estrutura da variação sem tendência (OLIVEIRA, 2015).

O IDW por sua vez, realiza a estimativa da variável ao longo do espaço, ponderando pesos a cada um dos “n” postos mais próximos, que é função do inverso de uma potência da distância, ou seja, quanto mais próximo do ponto a ser estimado, maior o peso atribuído ao ponto amostrado (JIMENEZ e DOMECCO, 2008). Esse método simples e clássico é recomendado para regiões com alta densidade de amostras (DEUS et al., 2010).

Já o método da interpolação por redes de triangulação irregular (TIN) é composto por pontos ou dados conhecidos que se ligam três a três, resultando em triângulos irregulares. Estes pontos simbolizam um conjunto de locais referenciados geograficamente, podendo ser pontos descontínuos, picos, divisores de água e outros (MAUNE e BLACK, 2001). Uma forma de minimizar erros na interpolação é evitar triângulos longos e finos, ou seja, com ângulos agudos. O ideal é que os triângulos sejam o mais equilátero possível (CARMO, 2015).

Segundo Landim (2000), o método TIN é de estimação direta, pois, os contornos derivam do padrão original dos dados, fiel aos dados originais, não permitindo a extrapolação, e as estimativas limitam-se estritamente a área amostrada.

2.6 INTERPRETAÇÃO DE MAPAS DE PRODUTIVIDADE

A AP é um processo cíclico e contínuo. As observações, interpretações, avaliações e implementações estarão sempre presentes a cada safra (COOK e BRAMLEY, 1998).

Na interpretação de mapas de produtividade, além do interesse na similaridade dos mapas, é interessante o estudo das similaridades de classes ou áreas dentro dos mapas. Essas diferenças entre classes ou áreas são observáveis através da investigação das associações existentes entre as categorias dentro de

cada classe e/ou entre as categorias de diferentes classes (CARVALHO, 2002).

Os mapas de produtividade contêm informações essenciais de diagnose da variabilidade da lavoura e, conseqüentemente, da eficiência do uso das técnicas da AP (MOLIN et al., 2015). Estes mapas de rendimento devem ser incorporados aos processos de tomada de decisão, análise e planejamento geral das operações na propriedade agrícola (SOUZA et al., 2016).

A correta interpretação dos mapas é mais efetiva após a observação das informações relacionadas à variabilidade, sendo então possível uma intervenção com vistas à utilização de operações localizadas (MOLIN, 2000).

3. MATERIAL E MÉTODOS

Os dados utilizados são provenientes de seis safras da cultura do milho (safras 2017, 2019 e 2020) e da soja (safras 2015, 2016 e 2019), de uma área agrícola de 23 ha, conforme a Figura 3, sob Plantio Direto, em Latossolo Vermelho (Santos et al., 2018), com 52% de argila, localizada no município de Cândido Mota – SP, ($22^{\circ}53'57.10''S$, $50^{\circ}23'58.11''O$), 368 m de altitude, clima Cfa (KÖPPEN, 1931).

Figura 3 – Mapas de demonstração da área agrícola utilizada para o estudo.



Os dados foram obtidos por meio de uma colhedora de grãos da marca John Deere, modelo STS 9750, equipada com sensor de produtividade tipo

placa de impacto; sensor de umidade por capacitância; sensor de velocidade tipo efeito Hall, instalado nas rodas movidas; sensor de inclinação pelo receptor GNSS; e, monitor modelo GS4 ® John Deere que apresenta e grava dados, disponibilizadas em formato de planilha eletrônica.

A produtividade medida pelo sensor de massa foi corrigida para 13% de umidade pelo método proposto por Aguiar (1977), conforme equação 2.

Eq. 2

$$\text{Peso Convertido} = \frac{100 - \text{Umidade mensurada}}{100 - 13} * \text{Peso Inicial}$$

Após a correção da umidade, os dados foram tratados pelo método *Jackknife*, metodologia descrita por Guanais Santos, et al. (2023), onde, no Excel, foi aplicado a análise de desvios pelo método dos Desvios Padronizados, visando estimar o índice de desvio pontual de cada leitura de rendimento, para posteriormente compará-la com a distribuição padrão “z”. Com o uso deste método os dados de produtividade (kg ha⁻¹) são convertidos para índice de Desvios (adimensional). Todos os valores abaixo de -3,09 ou acima de 3,09 foram considerados não pertencentes a população conforme a distribuição “z” e, portanto, foram removidos conforme sugere o método *Jackknife* (WU et al., 1986). Com a remoção dos dados, foi necessário recalcular os índices (DELALIBERA et al., 2017).

Após a remoção de pontos discrepantes foram gerados mapas com base nos índices de desvio por meio do *software SIG QGIS 3.22.8*, Figura 4. Os dados foram interpolados pelo método Redes de Triangulação Irregular (TIN). Utilizou-se o modo Desvio Padrão com 1 desvio para mais e para menos para realizar a classificação do mapa em três zonas.

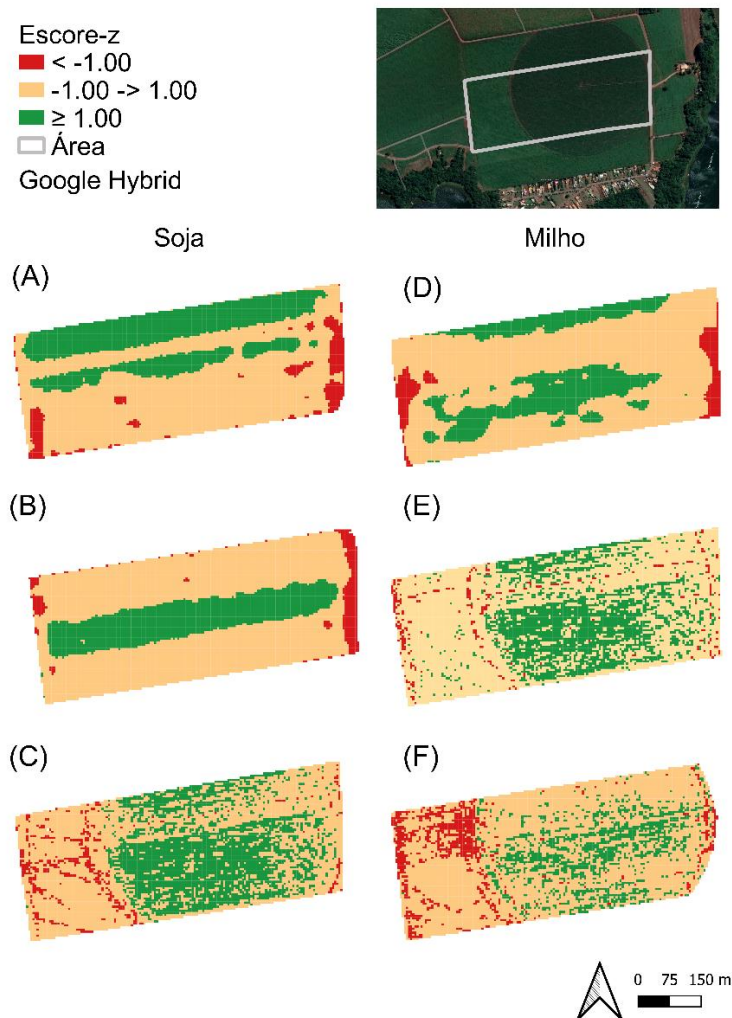
O método TIN foi utilizado pois não houve o interesse em estimar dados nos intervalos entre os pontos. Em dados de produtividade há uma grande densidade e uma proximidade dos pontos e o desejado foi obter uma informação fiel aos dados. Este método utiliza um menor número de pontos vizinhos para o cálculo da interpolação mantendo o nível de acurácia de uma grade regular mais densa (MIKHAIL et al., 2001).

As áreas com índices menores que -1 desvio foram consideradas de baixo índice de potencial produtivo, entre -0,999 a 0,999, consideradas de potencial

médio normal local e, acima de 1 desvio, consideradas de alto índice de potencial produtivo.

Em uma distribuição normal perfeita, 68,26% das ocorrências se concentraram na área do gráfico demarcada por um desvio padrão à direita e um desvio padrão à esquerda da linha média. Não necessariamente essa porcentagem de representação ocorre na distribuição de um fenômeno medido, porém, para este caso, foi considerado que os dados compreendidos em cada intervalo definido (<-1; -1 a 1 e >1 desvio) correspondem a níveis de potencial produtivo distintos.

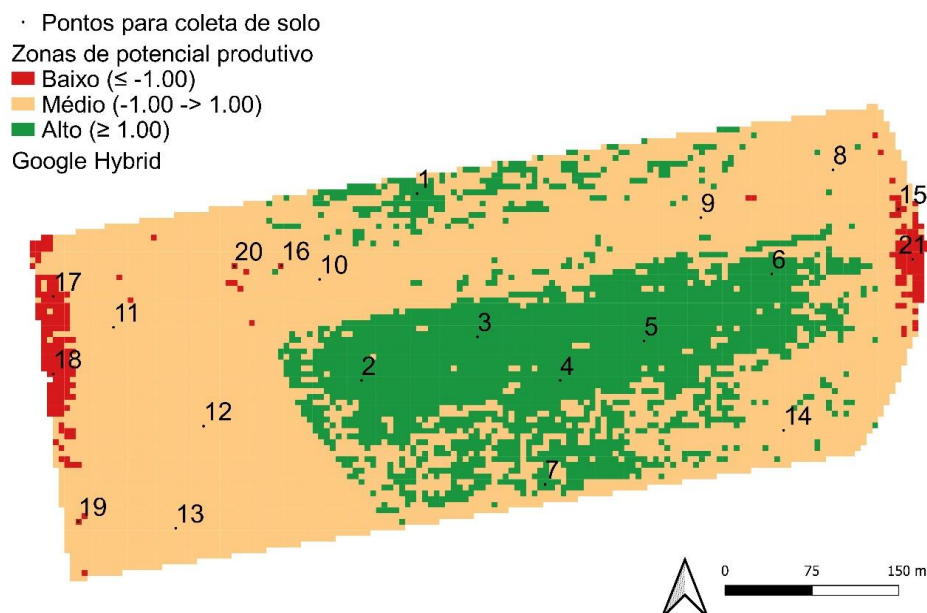
Figura 4 – Mapas de escore-z com a separação nas tres zonas de potencial produtivo, para cada cultura e safra.



(A) Refere-se à safra de 2015 da cultura da soja; (B) Refere-se à safra de 2016 da cultura da soja; (C) Refere-se à safra de 2019 da cultura da soja; (D) Refere-se à safra de 2017 da cultura do milho; (E) Refere-se à safra de 2019 da cultura do milho; (F) Refere-se à safra de 2020 da cultura do milho.

A partir da obtenção do mapa de índices de colheita corrigido (filtrado), foi realizada a álgebra entre os mapas de soja e milho, Figura 5, visando a obtenção de um único mapa que consiste na média dos índices das culturas. O intuito da realização dessa álgebra de mapas de índices de colheita foi proporcionar a diluição de efeitos de fatores não controlados e obter uma informação resumida para facilitar a interpretação dos resultados de forma conjunta e a tomada de decisão ser mais assertiva.

Figura 5 – Mapa de zonas de potencial produtivo e sua subdivisão em três classes com a distribuição dos 21 pontos de coleta de solo.



Para quantificar as áreas correspondentes aos potenciais produtivos, foi efetuado a conversão de *raster* para vetor com a ferramenta *raster pixels to polygons*, em seguida selecionada cada classe de interesse para quantificá-las em hectares por meio da ferramenta resumo estatístico.

Com base neste mapa médio, foram definidos os locais e a quantidade de amostras a serem coletadas para o decorrer do trabalho. Foi realizada a coleta de 21 amostras de solos a campo, conforme apresenta a Figura 5, consistindo de 7 amostras para cada zona de potencial produtivo. As coletas de amostra de solo foram retiradas em três profundidades de coleta 0 - 5cm, 5 - 10cm e 10 - 20cm. Os pontos enumerados de 1 a 7 foram posicionados na área de alto

potencial produtivo, de 8 a 14 posicionados na área de médio potencial produtivo e de 15 a 21 na área de baixo potencial produtivo.

Nos 21 pontos amostrais, em ambas as profundidades, foram avaliados os atributos químicos carbono orgânico (CO), matéria orgânica (MO), pH CaCl, pH água, pH SMP, H+Al, Al, cálcio (Ca), magnésio (Mg), fósforo (P) e potássio (K), pela metodologia proposta por Pavan et al. (1992); os atributos físicos granulometria (DONAGEMMA et al., 2017) com agitação de 200 rpm, argila dispersa em água (ADA), pelo método da pipeta, de acordo com a lei de Stokes após 20g de solo e 100ml de água destilada terem sido agitadas por 1 hora em agitador orbital horizontal (BARRAL, et al., 1988) à 200 rpm (MELO et al., 2015).

O diagnóstico rápido da estrutura do solo (DRES) foi realizado nos 21 pontos amostrais seguindo a metodologia proposta por Ralisch et al. (2017). Com as subdivisões das camadas, segundo a metodologia do DRES, foi realizada a análise da estabilidade de agregados (Embrapa, 1997) em cada uma das camadas, onde foram feitas 2 repetições de laboratório para cada camada. Utilizaram-se as peneiras de malhas 8,0 – 4,0 – 2,0 – 1,0 – 0,5 e 0,25 mm de diâmetro. O solo retido em cada peneira foi quantificado (EMBRAPA, 2017).

Também foram coletadas 21 amostras de solo indeformadas na camada de 0 - 10 cm de profundidade sendo uma para cada ponto de coleta, utilizando um amostrador cilíndrico de 99,47 cm³, para determinação da densidade do solo, tensão de retenção de água, macroporosidade e microporosidade pelo método da mesa de tensão (EMBRAPA, 1997).

Foram levantadas 41 variáveis no total, submetidas à análise estatística multivariada realizada no programa computacional Pirouette 4.5 ® Infometrix, Inc. Foram realizadas duas análises baseadas em escalonamento multidimensional, sendo uma análise exploratória, seguida pela aplicação de um método classificatório. Como método exploratório utilizou-se a análise *Principal Component Analysis* (PCA) e como método classificatório *Soft Independent Modeling of Class Analogy* (SIMCA).

O intuito destas análises foi verificar a existência de relação entre as zonas de potencial produtivo do mapa, com as variáveis de solo.

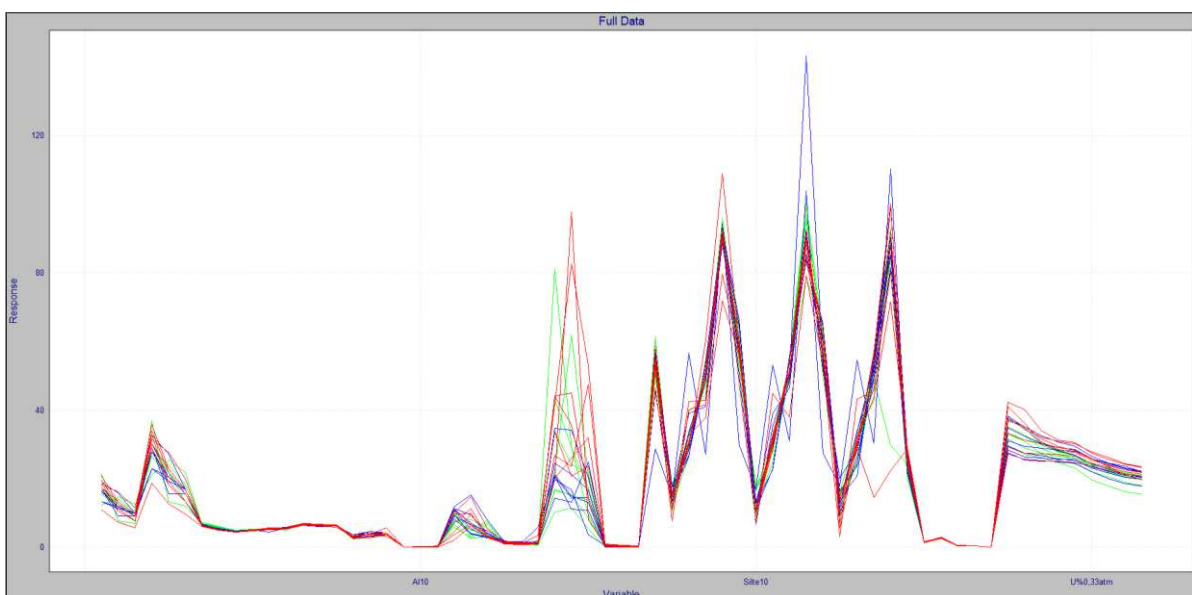
4. RESULTADOS E DISCUSSÃO

Com a aplicação da análise PCA é possível verificar a possível

relação de agrupamento entre as amostras em três grupos de potencial produtivo, ou seja, a distribuição multidimensional das amostras devem se dispor em forma de agrupamentos no espaço analítico, nesta análise também é possível verificar o comportamento dos dados e a existência de variáveis colineares e covariantes, permitindo escolher o pré-processamento, transformação e exclusão de variáveis que tendenciam e atrapalham o modelo. Com a comprovação deste agrupamento foi realizada a SIMCA como uma derivação dentro do modelo PCA.

Previamente à aplicação da PCA, foi necessário avaliar a distribuição das grandezas das variáveis (*Loadings*) conforme a amostra de entrada (*Scores*), representado na Figura 6. Ao observar variáveis com medidas de grandezas distintas, foi necessário aplicar um método de padronização dos dados conhecido como *Autoscale* ou autoescalamamento, que nada mais é que a conversão das variáveis para seus valores de *Score-z*, antes da aplicação da PCA.

Figura 6 – Diagrama da representação das grandezas das variáveis de acordo com a amostra .



Também nesta avaliação preliminar, conforme relatado acima, é possível avaliar a existência de variáveis covariantes e a existência de multicolinearidade no conjunto de dados. Esta avaliação prévia é importante, visto que a existência das relações citadas pode proporcionar erros de interpretação ou de julgamento do modelo gerado, sendo esta interferência muito forte na SIMCA. Portanto, neste momento, deve-se decidir por excluir parte das variáveis que

apresentem esses comportamentos de covariância e colinearidade.

As 41 variáveis foram utilizadas como parâmetros de entrada do tipo *Loadings* nas análises PCA para estimativa das inter-relações entre as zonas de potencial produtivo e as variáveis de solo. Esse modelo analítico visou verificar, por meio de métodos estatísticos, a qualidade e o grau de representatividade do método proposto para a geração do mapa de zonas de potencial produtivo, na determinação da variabilidade espacial de solo e planta, por meio de entradas de origem, natureza e métodos analíticos distintos, como, produtividade das culturas e variáveis de solo, para no caso desta validação, se propor a aplicação do Mapa de Zonas de Potencial Produtivo como método confiável, de fácil execução e de baixo custo, para direcionar a coleta de amostras de solo e realizar manejos distintos a partir dos resultados e da informação gerada.

A análise exploratória utilizada foi a PCA, capaz de representar graficamente as relações entre conjuntos de amostras e suas respectivas variáveis, possibilitando a redução da dimensionalidade efetiva do conjunto, encontrando as combinações lineares das variáveis independentes que representam a quantidade de variação máxima do conjunto de amostras (HAIR et al., 2009).

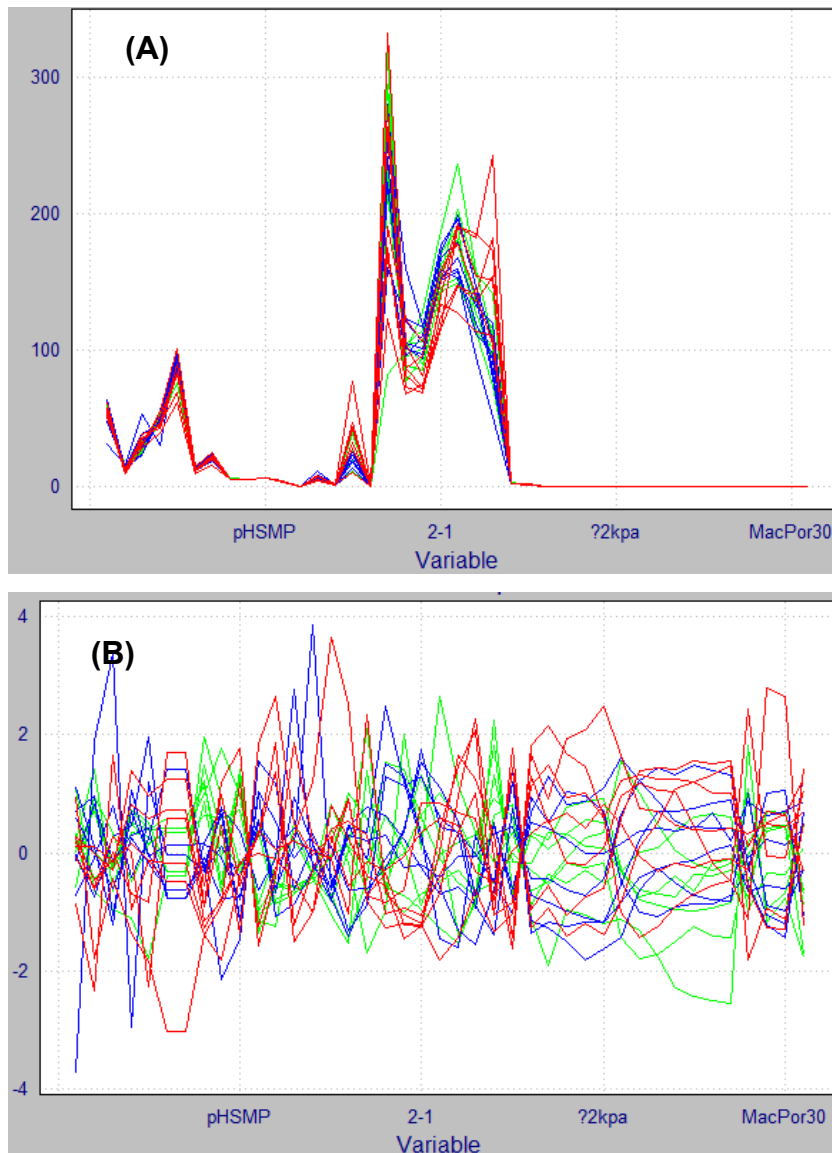
Como método classificatório, aplicou-se a SIMCA, a qual desenvolve modelos de componentes principais para cada classe atribuída pelo analista ao conjunto, testando a pertinência de cada amostra a sua classe proposta e atribuindo novas classes para cada amostra, conforme sua distância do grupo ou classe mais próxima, ou mesmo pode ditar se aquela amostra não pertence a nenhum dos grupos existentes no conjunto.

Ainda, esta técnica fornece um rico conjunto de diagnósticos específicos para cada classe, como o poder de modelamento dos *loadings* para cada classe que, diferente da PCA, no qual esta estimativa é genérica para o conjunto, na SIMCA, determina-se as variáveis mais importantes para cada classe ou conjunto de agrupamento separadamente, que, neste trabalho as classes são representadas pelas zonas de potencial produtivo. A estrutura de variância obtida para cada classe fornece a complexidade da categoria e até pode revelar os eventos que proporcionam a diferenciação ou separação entre as categorias ou grupos.

Previamente ao início da aplicação da PCA, foram avaliados os parâmetros referentes a natureza, a distribuição, a escala e a pressuposição da existência de covariância e multicolinearidade das variáveis de entrada que compõem

as amostras. Os dados apresentados na Figura 7 correspondem a representação gráfica da escala ou grandeza do conjunto de variáveis para cada amostra.

Figura 7 – Representação gráfica da escala ou grandeza do conjunto de variáveis para cada amostra.



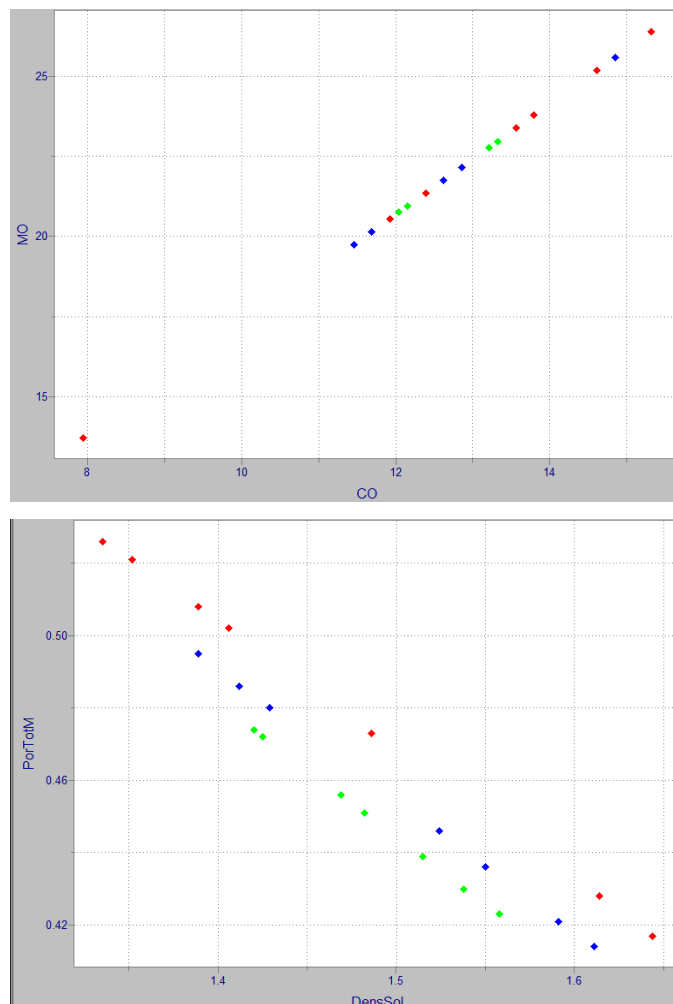
(A) Representação gráfica das diferenças de escala real de medida entre as variáveis para cada amostra; (B) representação gráfica das variáveis para cada amostra após a aplicação do pré-processamento por Autoescalamento.

Visto que a PCA utiliza da distância euclidiana para estimar a variabilidade intrínseca do conjunto de dados, conjuntos de variáveis que possuem escala de medida ou grandezas distintas proporcionam erros no desenvolvimento da análise, portanto, foi necessária a aplicação de um método de pré-processamento

das amostras, visando padronizar suas grandezas. Neste, decidiu-se aplicar o método conhecido como Autoescalamento, que se trata de uma derivação da padronização de dados conhecida como *escore-z*, advinda da estatística *z*, permitindo que apenas a variabilidade intrínseca de cada variável permaneça no conjunto analítico.

Ainda foram consideradas covariantes as variáveis pH + Al e pH SMP, Areia e Argila, MO e CO, Porosidade total do solo e Densidade do solo (DensSol), Saturação por base (?Sat) conforme exemplos na Figura 8. Destas foram excluídas as variáveis pH SMP, Areia, CO e Densidade do solo.

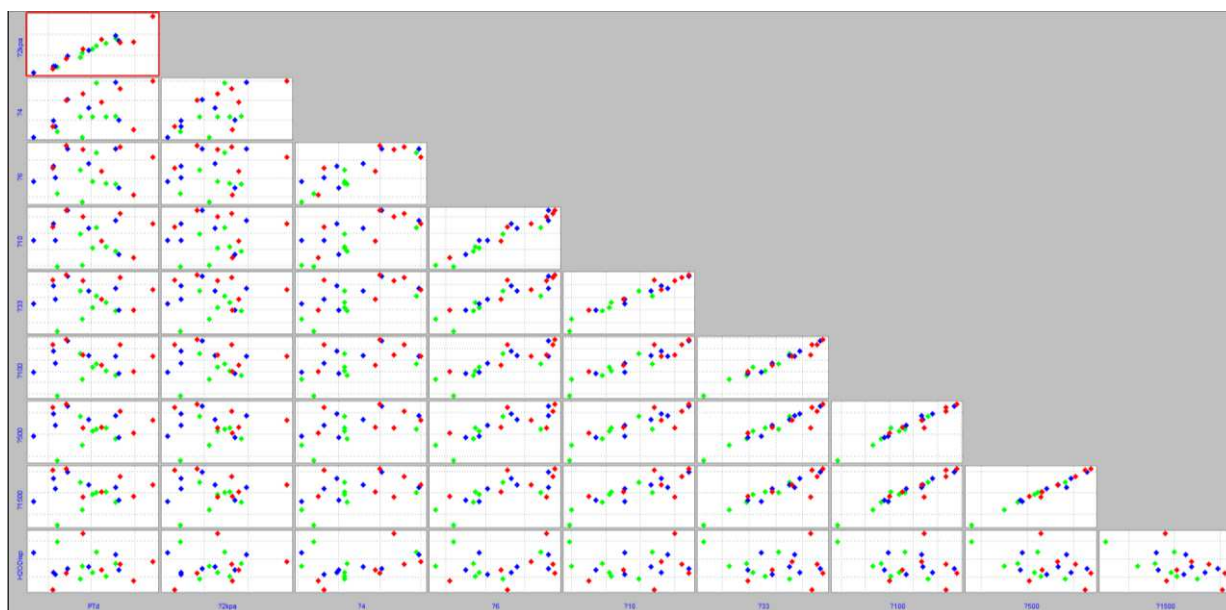
Figura 8 – Exemplo de algumas variáveis consideradas covariantes, como matéria orgânica com carbono orgânico e densidade do solo com porosidade total do solo.



(MO) Matéria orgânica; (CO) carbono orgânico; (DensSol) densidade do solo; (Porosidade total do solo) porosidade total do solo. Onde, a cor vermelha remete a classificação da área em baixo potencial produtivo, em amarelo a classificação da área em médio potencial produtivo e a cor verde na classificação da área em alto potencial produtivo.

Conforme apresentado na Figura 9, foram consideradas como multicolineares diversas variáveis resultantes da estimativa da curva de retenção e destas, foram excluídas Θ_{2kPA} (θ_{2kpa}), 4, 6, 100 e Θ_{500} (θ_{500}), referente a retenção de água na naquela determinada tensão.

Figura 9 – Representação das variáveis consideradas multicolineares.



(Θ_{2kPA} (θ_{2kpa}), 4, 6, 100 e Θ_{500} (θ_{500})): se tratam das leituras obtidas para cada tensão exercidas nas amostras indeformadas pela mesa de tensão ou pelo extrator de Richards.

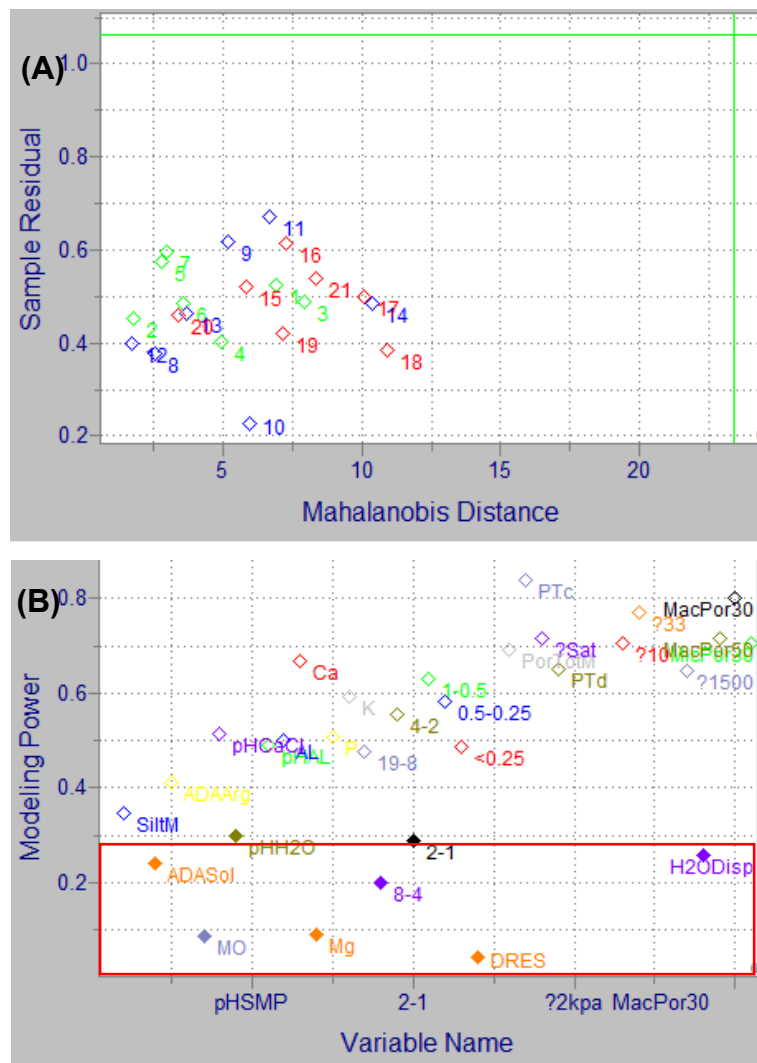
O processo de verificação e exclusão de variáveis covariantes e/ou multicolineares previamente à realização da PCA e da SIMCA é relevante, pois, as variáveis com estas propriedades proporcionam erros de estimativa das análises por deslocar o centro gravitacional do modelo em tendência das mesmas e, erroneamente se apresentarem como importantes dentro das componentes geradas. Entretanto, no poder de modelagem da PCA, estas se apresentaram como poder médio ou baixo.

Após as exclusões mencionadas anteriormente, foi realizada uma PCA, aplicando o Autoescalamento como pré-processamento. Nesta primeira análise, foi verificado primeiramente dois aspectos da PCA, ainda como uma fase de filtragem ou exclusão de dados, antes de realizar qualquer interpretação. O primeiro aspecto avaliado foi com relação a Distância de Mahalanobis, visando verificar a existência de amostras (*Scores*) discrepantes, conforme a Figura 10, que, neste caso, não foram detectadas amostras com essa característica.

O segundo aspecto consistiu em avaliar o Poder de Modelamento das variáveis para o conjunto avaliado na PCA, apresentado na Figura 10. Para este foram excluídas variáveis com poder de modelamento menor que 0,3 (30%), visto que estas variáveis apresentam baixa relevância para o modelo e acabam poluindo o mesmo, dificultando a interpretação dos resultados. Quanto mais próximo de um está o poder de modelagem da variável, mais ela contribui para a explicação do modelo.

Com base neste preceito foram excluídas as variáveis Argila, Argila dispersa em água (ADASol), MO, pH H₂O, Mg, 8-4, 2-1, onde as duas últimas variáveis correspondem a malha das peneiras utilizadas na análise de estabilidade de agregado e DRES.

Figura 10 – Poder de Modelamento das variáveis para o conjunto avaliado na PCA.



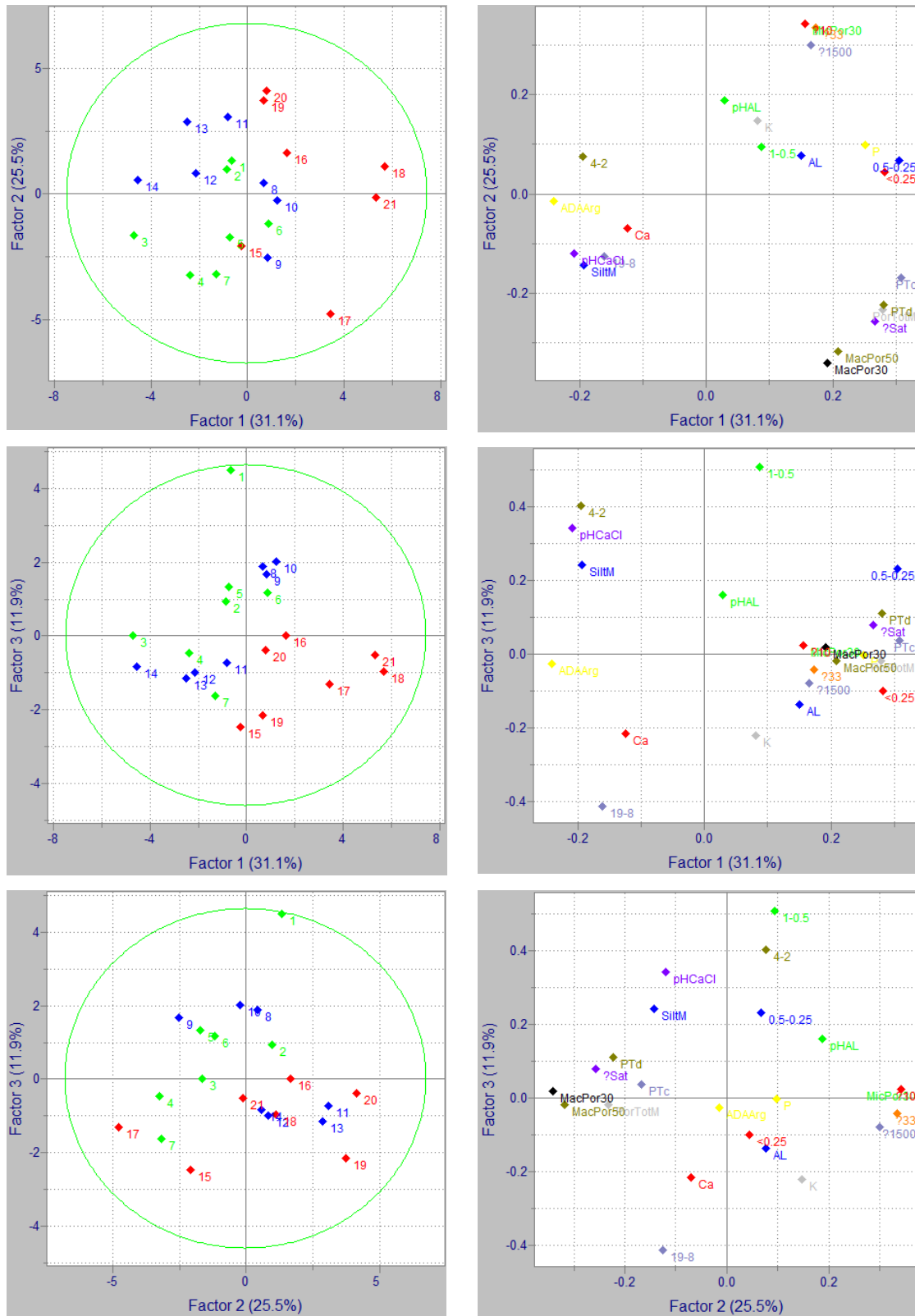
(A) Disposição gráfica da distância de *Mahalanobis* onde as linhas verdes são referentes aos valores críticos; (B) Disposição gráfica do grau de poder de modelamento por variável, onde, as destacadas em vermelho representam as variáveis que foram excluídas.

Após esta segunda fase de exclusão de variáveis não representativas, a PCA foi executada novamente, também utilizando o método de Autoescalamento como pré-processamento e, aplicado na forma de classes para a PCA, as referências de cada amostra com sua respectiva Zona de Potencial Produtivo, visando que, durante esta análise, as disposições gráficas das amostras apareceram identificadas por cores, onde, as amostras de 1 a 7, em verde claro, representam as amostras da zona de alto potencial produtivo, de 8 a 14, em azul, as amostras da zona de potencial médio normal e, de 15 a 21, em vermelho, as amostras da zona de baixo potencial produtivo. Todas as figuras que apresentam apenas essas três cores para as amostras (*loadings*), aplica-se esta regra citada acima.

Após a aplicação da PCA para este conjunto de 21 amostras (*scores*) pré-classificadas em três Zonas de Potencial Produtivo e 23 variáveis (*loadings*) restantes, verificou-se que as três primeiras componentes explicam cerca de 68,6% da variabilidade intrínseca do conjunto, e são suficientes para explicar os eventos avaliados.

A Figura 11 apresenta a disposição gráfica da explicação bidimensional entre as componentes 1 e 2, 1 e 3 e 2 e 3, fatores estes que correspondem ao ângulo de visão bidimensional do conjunto, ou seja, a projeção específica do ambiente multidimensional das interações para os *scores* e para os *loadings*. Nestas disposições gráficas é possível observar a formação ou separação dos grupos de amostras na mesma ordem das classes pré-identificadas por cores das zonas de potencial produtivo.

Figura 11 – Disposição gráfica bidimensional das componentes 1 e 2, 1 e 3 e 2 e 3, para os scores a esquerda e para os loadings a direita.



Pontos enumerados de 1 a 7 foram posicionados na área de alto potencial produtivo, de 8 a 14 na área de médio potencial produtivo e de 15 a 21 na área de baixo potencial produtivo.

Com a apresentação dos três ângulos de visão, observa-se a aglomeração dos conjuntos de dados e seus respectivos fatores mais representativos a essa interação.

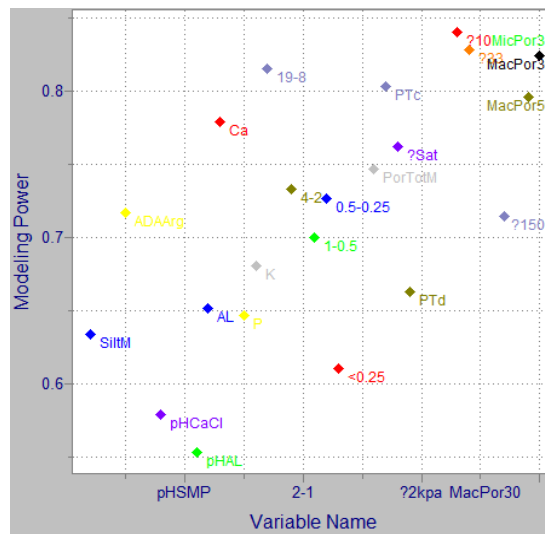
Entretanto, a baixa dissimilaridade entre o conjunto das amostras, independente da classe a que pertencem, dificulta o entendimento das inter-relações dos *scores* com os *loadings*, possibilitando apenas um entendimento superficial de que para a classe 1, destacada em verde (Zona de Alto Potencial Produtivo), possui sua variabilidade regida por pH CaCl, Silte e 4-2 (correspondente a malha das peneiras utilizadas na análise de estabilidade de agregados).

Da mesma forma, a classe 3, destacada em vermelho (Zona de Baixo Potencial Produtivo), possui sua variabilidade regida por diversas variáveis ligadas as medidas de potencial matricial (ψ_{1500} , ψ_{33}), que corresponde a porcentagem de umidade no extrator de Richards dados em kPa, correlacionas a estrutura do solo, as medidas da malha das peneiras ($<0,25$) e a presença do elemento Al.

Contudo, o comentado no parágrafo anterior podem ser afirmações consideradas frágeis, visto a baixa dissimilaridade dos grupos de amostras formados e das próprias amostras.

Esse fator é explicado pela composição do poder de modelagem de cada *loading* no conjunto, conforme apresentado na Figura 12, no qual apenas a variável ψ_{33} apresenta maior poder de modelamento em relação a composição do modelo genérico da PCA, com poder de modelamento maior que 0,8. Não considerando a relevância direta das variáveis para cada classe ou conjunto separadamente, e sim para o conjunto amostral.

Figura 12 – Representação gráfica do grau de poder de modelagem dos *loadings* dentro do modelo PCA.



Com a PCA, foi possível verificar o poder de modelagem das variáveis analisadas para verificar quais delas melhor responderia a separação em três grupos, Figura 12.

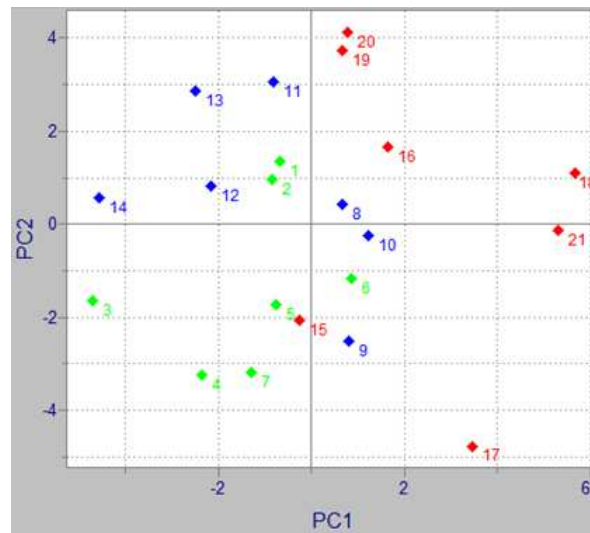
Com esse processamento de dados observou-se já uma separação de atributos correlacionados a características físicas do solo, como Micro Porosidade (MicPor30), Macro Porosidade (MacPor30) e Capacidade de campo. Por se tratar de uma área com tratos culturais realizados de maneira homogênea por todo o talhão, as variáveis químicas do solo não apresentaram maiores correlações com os agrupamentos.

Segundo Kiehl (1979), os macroporos são designados para aeração, movimentação de água e penetração das raízes, enquanto os microporos são responsáveis por armazenar a água no solo. Em relação à textura, quando nos solos predominam partículas maiores, como a areia, predominam também macroporos, e em solos predominantemente argilosos, predominam os microporos (VIEIRA et al., 1988).

Para um entendimento mais aprofundado e para se obter mais resultados quantitativos sobre a representatividade do modelo PCA e suas interpretações, se desdobrou a PCA pelo modelo SIMCA, visando testar a validade e congruência de cada classe atribuída (Zonas de Potencial Produtivo) aos *scores* (amostras coletadas dentro de cada zona à campo), bem como determinar as variáveis (*loadings*) que são importantes dentro de cada classe.

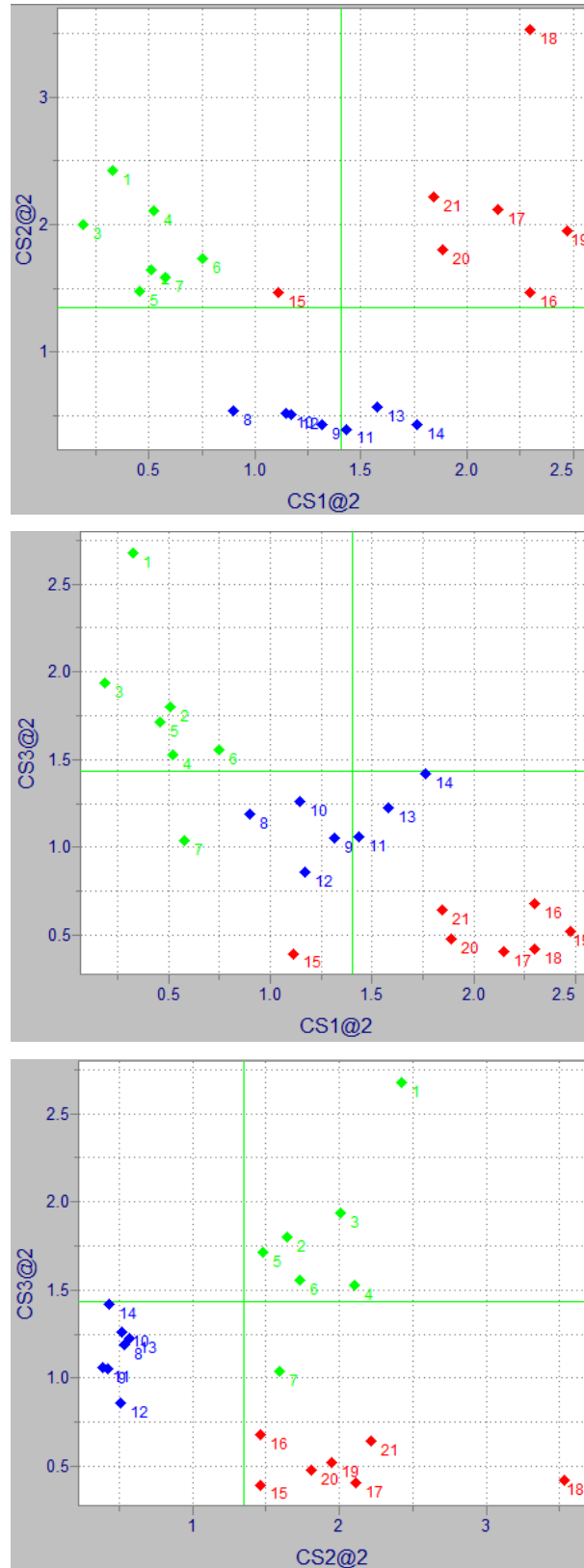
A Figura 13 apresenta a representação gráfica da componente 1 com a componente 2 para a projeção de classes da modelagem geral do conjunto de amostras da SIMCA e, assim como a PCA, devem necessariamente ser iguais (Figura 11 componentes 1 e 2 dos scores).

Figura 13 – Representação gráfica das projeções de classes apresentada pelo modelo SIMCA.



Na representação dos modelos derivativos de distâncias entre classes nominais da SIMCA, é possível observar que, mesmo com a baixa dissimilaridade entre as amostras, o modelo foi capaz de fazer uma boa separação entre os grupos de amostras com base em suas respectivas classes nominais, como se pode observar na Figura 14.

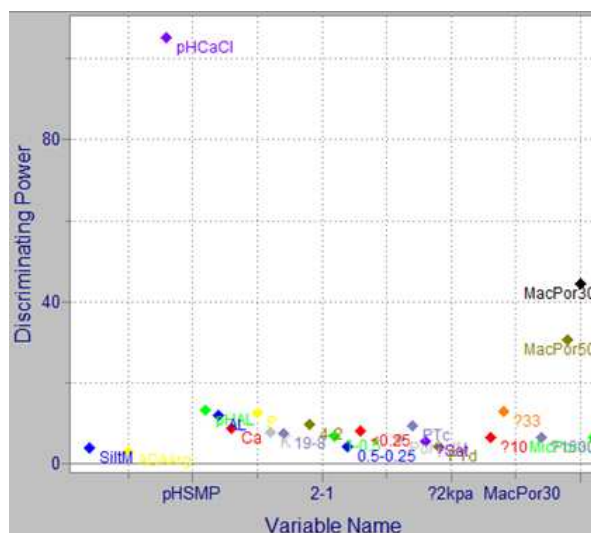
Figura 14 – Representação gráfica da distância entre classes para as relações entre as três componentes dimensionais.



Este fato comprova que o modelo gerado foi eficiente em explicar as relações e as diferenças para o conjunto de dados testado, o qual apresenta baixa dissimilaridade, podendo observar que em ambas as dimensões, as amostras se agrupam com sua respectiva classe nominal (cores). Observando a Figura 14, também se tem que pelas classes resultantes das zonas de manejo, que há uma forte relação entre o histórico de produtividade da área, representado de forma agrupada e resumida pelo mapa de zonas de potencial produtivo, realmente é regido pelas variáveis de solo constantes no modelo, sendo que a produtividade e o potencial produtivo da área, são resultados de uma complexa interação múltipla destas variáveis, tornando o evento produtividade das culturas, de difícil correlação direta com variáveis isoladas e de interpretação complexa.

Também é possível observar que, na representação gráfica do poder de discriminação das variáveis (Figura 15), que a variável pH CaCl se destaca dentre as outras variáveis, com grau de 105,2, ou seja, é uma variável que apresenta comportamento distinto das demais e, pode ser explicado por seu baixo coeficiente de variação, onde o valor apresentado na Tabela 3 com 2,04, 1,03 e 1,63 para as três zonas.

Figura 15 – Poder de discriminação ou diferenciação entre as variáveis, conforme apresentado pela SIMCA.



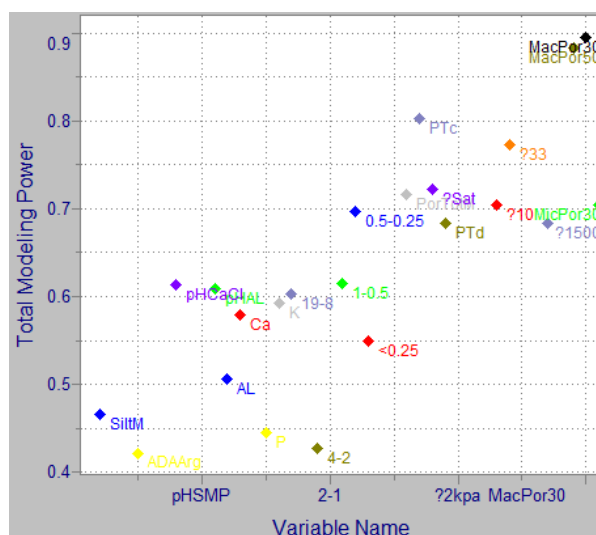
Este resultado também pode explicar o fato de esta mesma variável ter aparecido como importante na explicação da variabilidade intrínseca do conjunto para a classe 1 na PCA. Entretanto, na sequência da interpretação da SIMCA é

apresentado que esta interpretação não é adequada para uma visão mais estreita das classes, conforme já comentado.

A SIMCA também apresenta o poder de modelagem das variáveis para a estimativa geral das classes nominais, conforme a Figura 16, que diferentemente do apresentado na PCA (Figura 12), que apresenta o poder de modelagem com referência no conjunto amostral e, mesmo com essa diferença matemática e de visão analítica, a SIMCA corrobora parcialmente com o apresentado nos resultados da PCA, sendo referência importante para a validação entre as diferentes variáveis analisadas.

Nas Figuras 12 e 16, se observa que as variáveis relacionadas à translocação e armazenamento de água no solo se apresentam como as mais relevantes para o contexto geral da explicação da variabilidade intrínseca da área, ou seja, mais de 70% das relações interdimensionais que proporcionam o resultado da formação das classes de zonas de manejo estudadas estão relacionadas as variáveis MacPor50, MacPor30, MicPor30, porosidade total calculada (PTc), θ_{33} kPa e θ_{10} kPa.

Figura 16 – Disposição gráfica do poder de modelagem das variáveis para a estimativa geral das classes nominais.



Para o quesito de classificação preditiva da SIMCA, que avalia com base nas distâncias entre as amostras e as classes nominais, qual melhor alocação de classe para cada amostra, conforme a Tabela 1 como *Best* (indicação analítica

principal do enquadramento da amostra nas classes nominais) e, *Next Best*, que apresenta em quais outras classes nominais cada amostra poderia se enquadrar ou estaria mais próxima, lembrando que quando esta referência indica o valor zero, a amostra não se enquadra em nenhuma das classes indicadas pelo analista.

A Tabela 1, apresenta a validação do método de estimativa para a geração do mapa de Zonas de Potencial Produtivo, pois, essa parte da SIMCA testa o enquadramento de cada amostra em uma classe pré-cadastrada pelo analista, que neste caso específico são as três Zonas de Potencial Produtivo, sempre mencionada no corpo do texto como classe nominal, onde a coluna *Best*, apresenta a classe no qual o modelo analítico enquadrou cada amostra em teste pontual uma a uma e, a coluna *Next Best*, se refere à classe no qual o modelo sugere que cada amostra pode ser reenquadrada, caso o enquadramento de classe analítica (*Best*) não corrobore com a classe nominal. Quando o valor zero é apresentado na coluna *Best* e/ou na coluna *Next Best*, a amostra não se encaixa em nenhuma das classes nominais sugeridas.

Tabela 1 – Predição analítica de enquadramento das amostras com base nas três classes (Zonas de Potencial Produtivo) cada amostra se enquadra (*Best*) e qual outra classe a referida amostra poderia ser realocada (*NextBest*).

Potencial produtivo	Número da Amostra	Best (classe)	NextBest (classe)
Alto (classe 1)	1	1	0
	2	1	0
	3	1	0
	4	1	0
	5	1	0
	6	1	0
	7	1	3
Médio (classe 2)	8	2	1
	9	2	3
	10	2	1
	11	2	3
	12	2	3
	13	2	3
	14	2	3
Baixo (classe 3)	15	3	1
	16	3	0
	17	3	0
	18	3	0
	19	3	0
	20	3	0
	21	3	0

De acordo com a Tabela 1, todas as amostras se enquadram nas suas respectivas classes nominais (amostras de 1 a 7), pertencem a classe 1 ou zona de alto potencial produtivo, conforme estimada pela metodologia proposta; amostras de 8 a 14 pertencem à classe 2 ou zona de potencial médio normal; amostras de 15 a 21 pertencem à classe 3, ou zona de baixo potencial produtivo, com base na SIMCA a nível de 95% de probabilidade, aceito 5% de erro probabilístico. Ainda na mesma tabela, é possível observar que o modelo analítico também indica que as amostras 7 (pertencente a classe 1), 9, 11, 12, 13 e 14 (pertencentes à classe 2) são próximas e poderiam ser enquadradas na classe 3, caso alguma apresentasse o valor 0 na coluna *Best*. O mesmo acontece com as amostras 8, 10 (pertencentes à classe 2) e 15 (pertencentes a classe 3), as quais apresentam proximidade da classe 1.

Continuando a avaliação das saídas da SIMCA, Tabela 2, há uma estimativa de auto validação dos seus modelos, a qual verifica possíveis amostras

classificadas erroneamente, chamada de *Misclassification*.

Nesta tabela, a identificação de valores numéricos diferentes de zero fora da diagonal da tabela, indicam que o modelo pode ter errado o enquadramento de classe de alguma das amostras. Já os números 7, representam o número total de amostras enquadradas corretamente em sua respectiva classe analítica. Sobre a coluna não correspondência (*No Match*), a existência de qualquer número diferente de zero indica a presença de amostras que não se enquadram em nenhuma das classes nominais apresentadas pelo analista.

Tabela 2 – Modelo de validação quantitativa do enquadramento das amostras nas classes analíticas nominado como *Misclassification*.

Zonas de potencial produtivo	Pred1	Pred2	Pred3	No match
Alto	7	0	0	0
Médio	0	7	0	0
Baixo	0	0	7	0

A partir do resultado obtido pela autoavaliação do modelo gerado pela SIMCA, demonstramos que eles representam a realidade dos eventos analisados a 5% de probabilidade de erro.

Após a interpretação do modelo geral da SIMCA, bem como o teste de enquadramento das amostras nas classes de zonas de potencial produtivo sugeridas e sua auto validação, podemos dar sequência à interpretação dos desdobramentos por analogia de classe ou zonas de potencial produtivo.

Ambas as tabelas citadas acima, comprovam a eficácia da metodologia proposta para identificação de zonas de potencial produtivo e suas correlações com os fatores de solo. Todas as amostras distribuídas nas zonas de baixo e alto potencial produtivo se agrupam em seus respectivos grupos, sem a possibilidade de similaridade com outro grupo, ou seja, há uma total distinção das zonas bem como suas relações diretas com as variáveis de solo.

Apenas as amostras da zona de médio potencial produtivo apresenta a possibilidade de pertencer a mais de um grupo, o que se mostra um comportamento normal, visto que se trata de uma zona de transição.

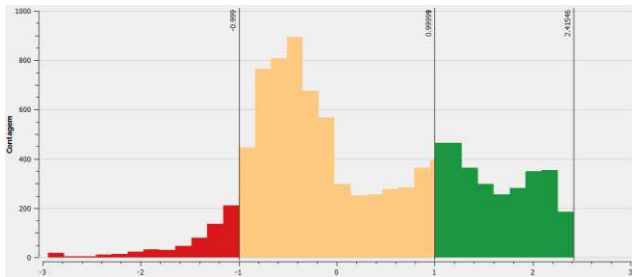
Este resultado aponta que a proposto por Guanais Santos, et al. (2023) é eficaz para relacionar o potencial produtivo da área com as variáveis de solo,

classificando a área em apenas três subdivisões e permitindo além da inferência de amostragem no campo, a inferência sobre o manejo da área. Com isso pode-se reduzir a quantidade de amostras de solo por zona de potencial produtivo e em quantidade reduzida, sem a perda de informações importantes e permitindo que o responsável técnico não se limite apenas a análise físico-química da área durante seu diagnóstico visando a recomendação técnica.

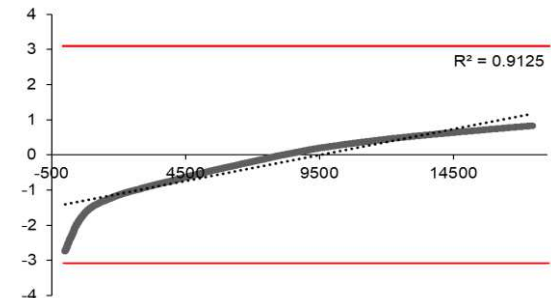
Em sequência, nas Figuras 17, 18 e 19, são apresentados os histogramas de frequência dos escores-z, que permite a avaliação da qualidade do manejo e os gráficos do grau de linearidade da frequência ordenada dos escores-z, demonstrado pelo índice R^2 , que ajuda no entendimento da exclusão de dados discrepantes das entradas do mapa.

Figura 17 - Histogramas dos índices de produtividade da cultura da soja e diagrama de frequência ordenada dos índices de desvio padronizado e seus respectivos graus de ajuste da reta, referente ao fenômeno colheita mecanizada de grãos antes e após da aplicação do método jackknife.

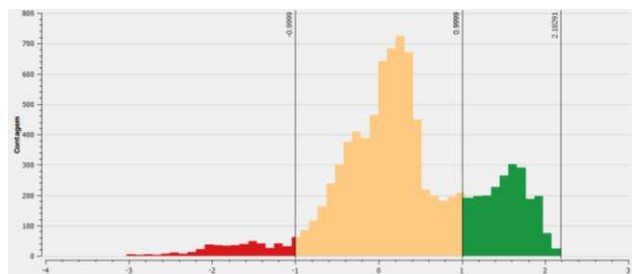
(A)



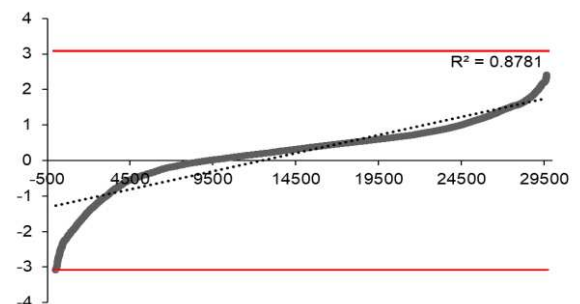
(D)



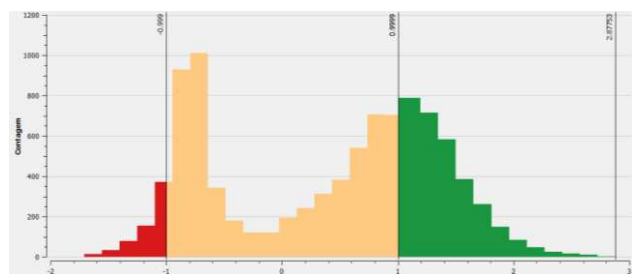
(B)



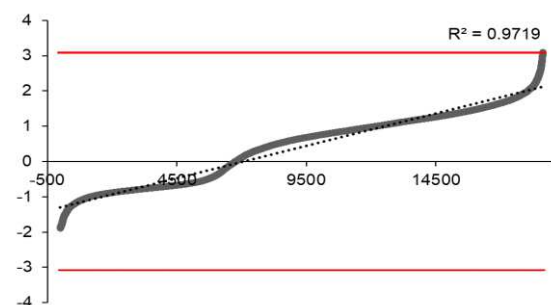
(E)



(C)

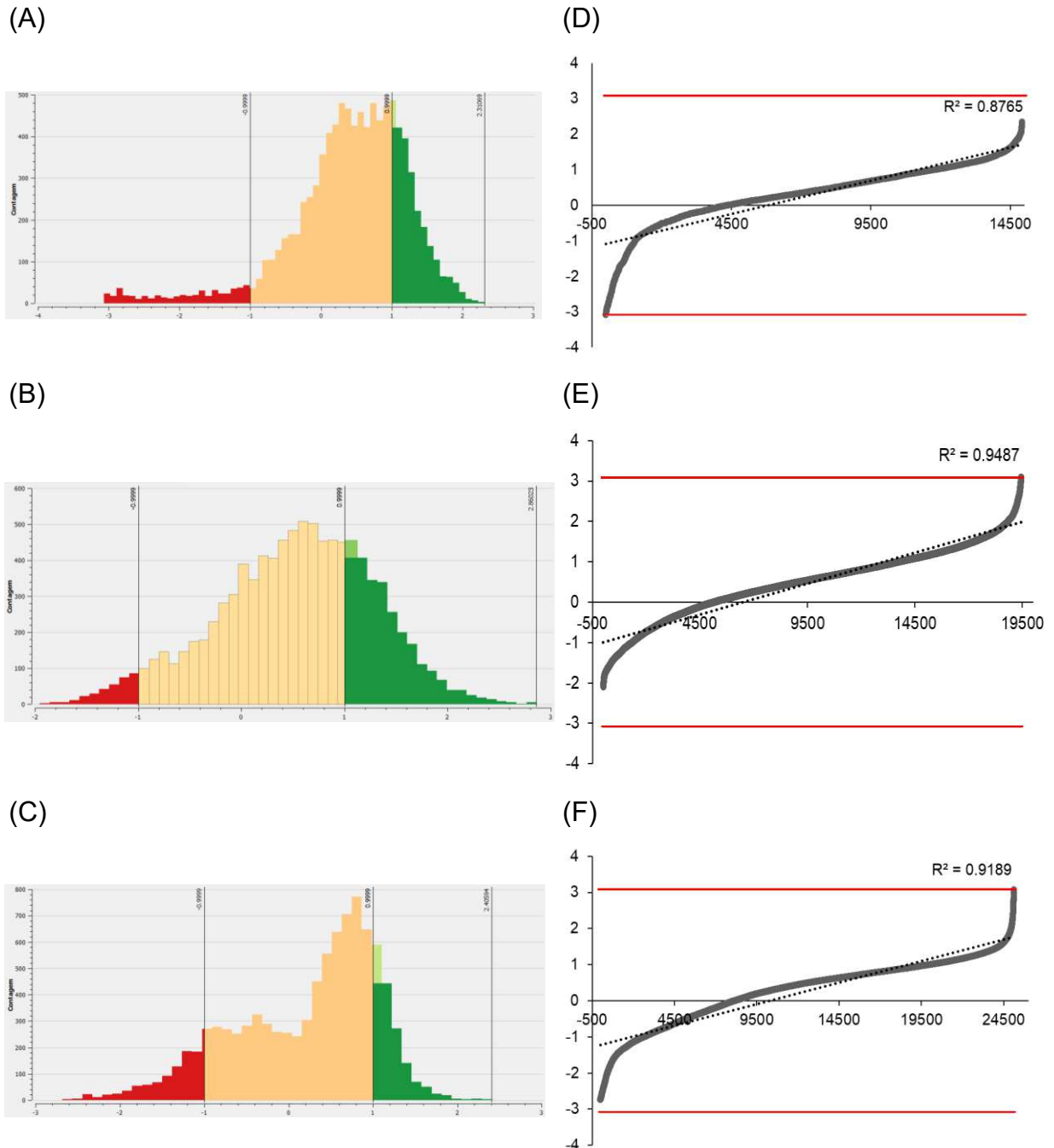


(F)



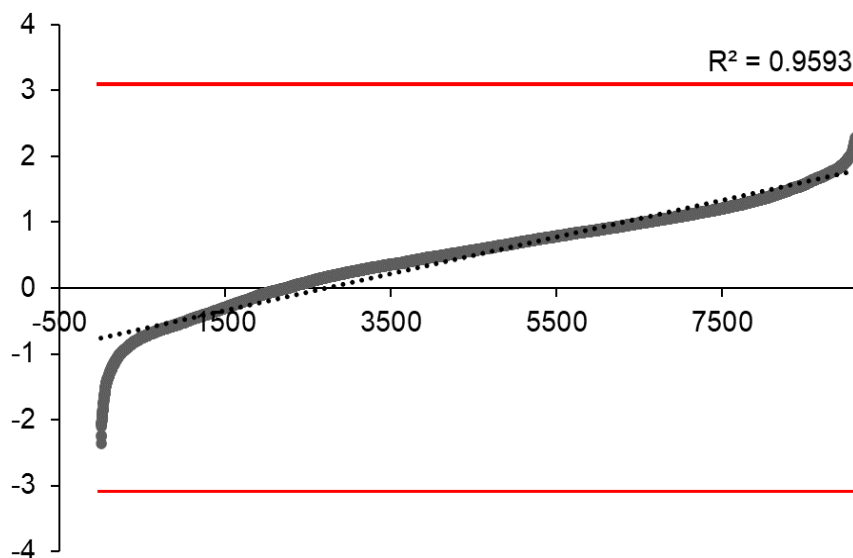
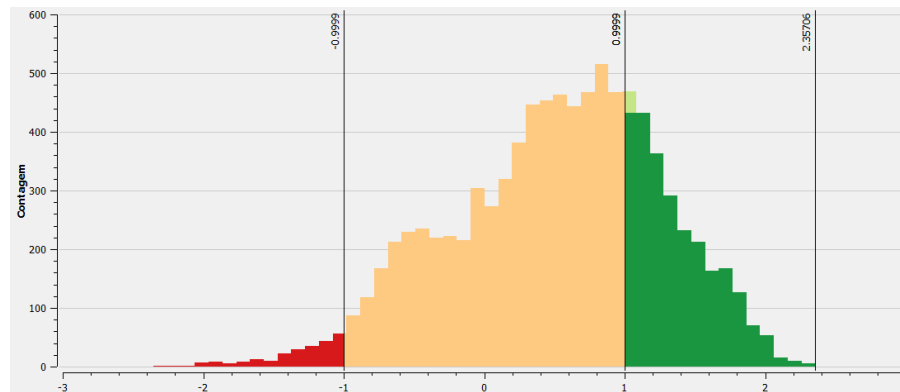
De A à C - Corresponde aos histogramas dos mapas após a interpolação dos dados das safras 2015, 2016 e 2019, respectivamente; De D à F - Corresponde aos diagramas de frequência ordenada de desvios padronizados das culturas das safras 2015, 2016 e 2019, respectivamente.

Figura 18 - Histogramas dos índices de produtividade da cultura do milho e diagrama de frequência ordenada dos índices de desvio padronizado e seus respectivos graus de ajuste da reta, referente ao fenômeno colheita mecanizada de grãos antes e após da aplicação do método jackknife.



De A à C - Corresponde aos histogramas dos mapas após a interpolação dos dados das safras 2017, 2019 e 2020, respectivamente; De D à F - Corresponde aos diagramas de frequência ordenada de desvios padronizados das culturas das safras 2015, 2019 e 2020, respectivamente.

Figura 19 - Histogramas dos índices de produtividade do mapa de zonas de potencial produtivo e diagrama de frequência ordenada dos índices de desvio padronizado e seus respectivos graus de ajuste da reta, referente ao fenômeno colheita mecanizada de grãos antes e após a aplicação do método jackknife.

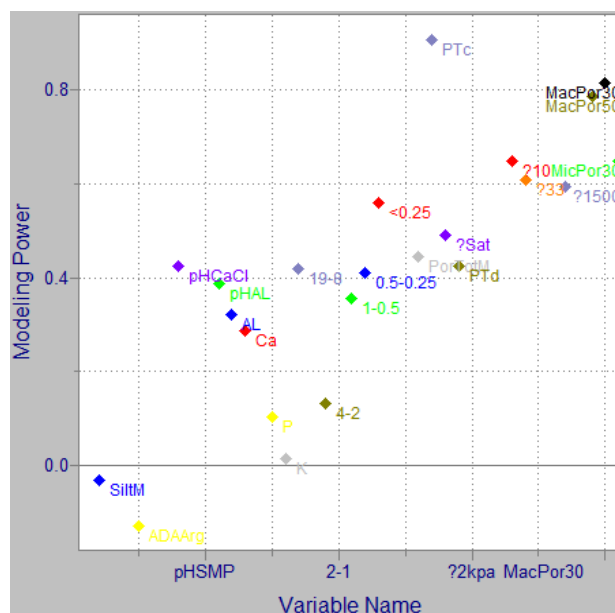


Ainda aprofundando este estudo de caso, foram feitos os desdobramentos da SIMCA para cada uma das zonas de potencial produtivo (classes) para a avaliação da variável que melhor responderia a cada uma das zonas.

Na figura 20, que corresponde ao desdobramento da primeira classe, de alto potencial produtivo, a distância de *Mahalanobis* não apresentou amostras discrepantes e o modelo apresentou explicação acumulada para os três fatores existentes igual a 84,5% das relações interespecíficas entre as variáveis desta classe. A princípio, a interpretação do modelo se apresenta confusa, pois as variáveis acabam se relacionando de forma mais estreita, cada qual com uma amostra e o entendimento do conceito de classe de forma mais holística fica prejudicado.

No entanto, a não ser que o analista tenha interesse em interpretar amostra por amostra, o que não se aplica a este estudo, o relevante é analisar a estimativa do poder de modelagem, conforme apresentado na Figura 21.

Figura 21 - Disposição gráfica do poder de modelagem das variáveis para a classe de alto potencial produtivo.



Com esta disposição gráfica do poder de modelagem verifica que a variável porosidade total calculada ($PTc = 0,90$) é a variável de maior importância para essa classe, seguido pelas variáveis MacPor30 e MacPor50 (0,81 e 0,78, respectivamente), seguindo a tendência das respostas de poder de modelamento da PCA e a global da SIMCA.

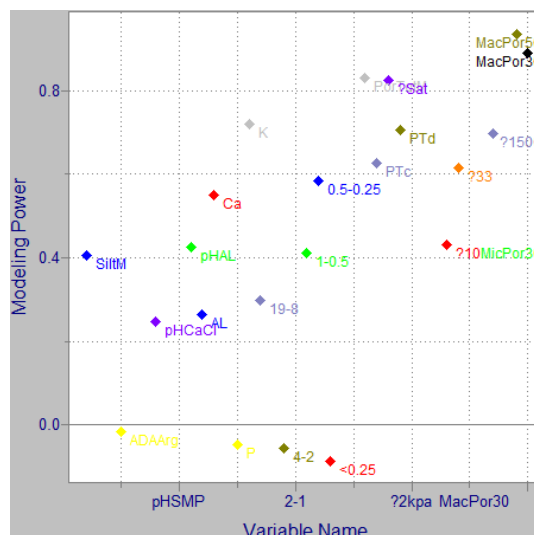
Para a classe 2, a distância de *Mahalanobis* também não demonstrou a presença de amostras discrepantes. As três componentes existentes no modelo apresentaram explicação acumulada da variância de 84,7% conforme a Figura 22.

Ainda, é possível verificar que as variáveis Silte, da área de médio potencial produtivo, classe 2, e ADA, da área de alto potencial produtivo, classe 1, não respondem no modelo apresentando até valores negativos. Fato este que pode estar relacionado à condição de que a área avaliada apresenta a mesma classe textural de solo.

É válido ressaltar que a interpretação das inter-relações das variáveis, amostra por amostra, não contribui com a elucidação da hipótese testada neste trabalho, visto que esta já foi resolvida conforme relatado anteriormente, nas Tabelas 1 e 2.

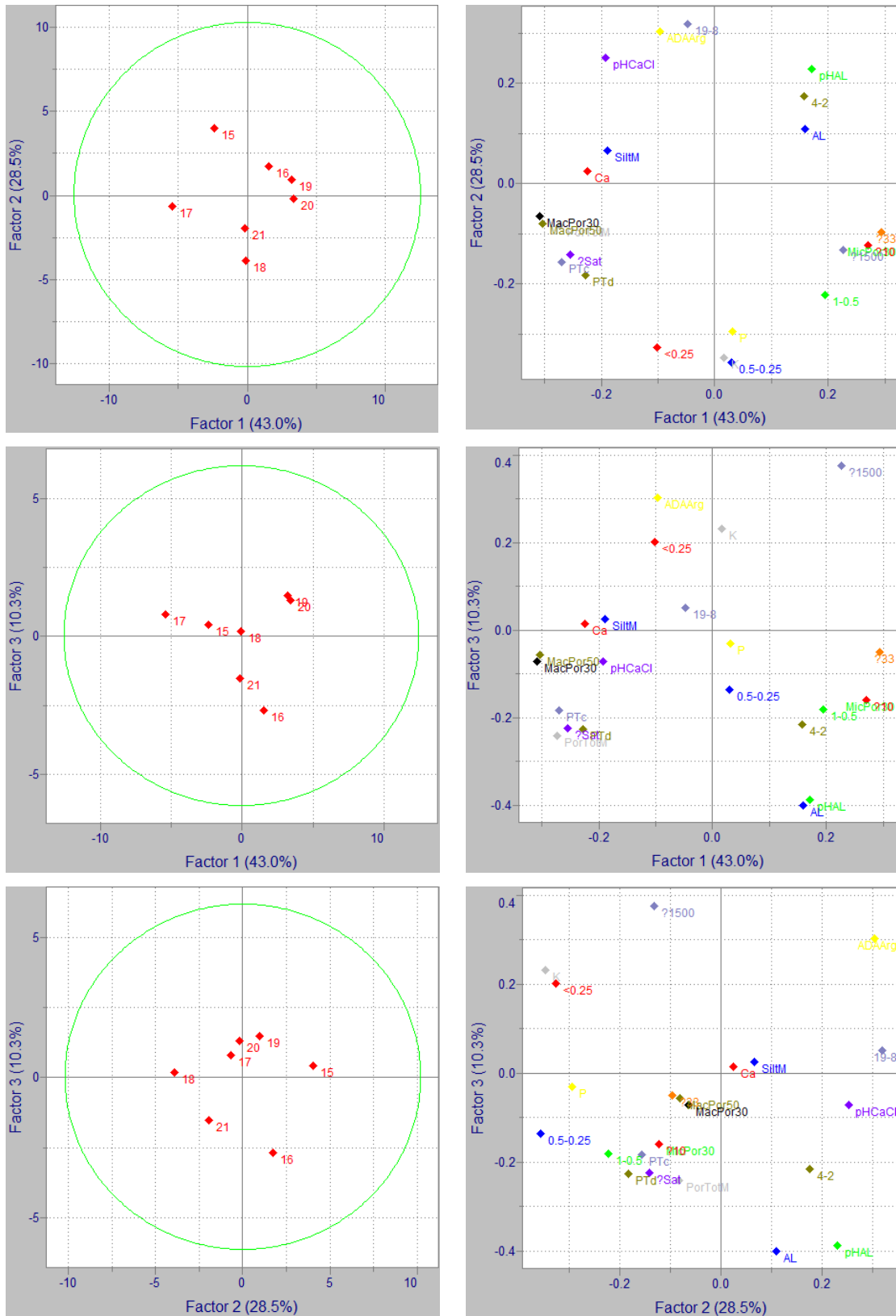
No entanto, aproveitando o contexto analítico, a Figura 23 apresenta o poder de modelagem das variáveis para a classe 2 e, nesta é possível verificar que, como para a classe 1, a variabilidade desta zona de potencial produtivo é explicada pelas variáveis ligadas à porosidade do solo e condutividade hidráulica, na ordem MacPor50 (0,94), MacPor30 (0,89), Porosidade total do soloM (0,83) e Θ Sat (?Sat = 0,83). E, como na classe 1, as variáveis ADA, Arg, P, 4 - 2 e <0,25, não contribuem com a explicação do modelo e, até poderiam ser excluídas.

Figura 23 - Disposição gráfica do poder de modelagem das variáveis para a zona de médio potencial produtivo.



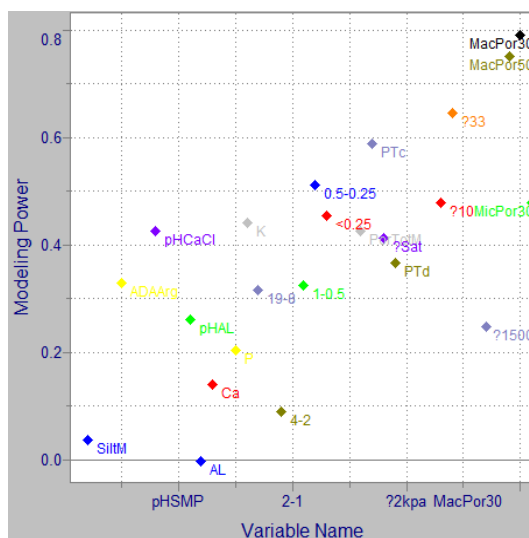
Para a classe 3, de baixo potencial produtivo, a distância de *Mahalanobis* não apresentou amostras discrepantes com explicação acumulada da variância de 81,9%, conforme a Figura 24.

Figura 24 - Representação gráfica da disposição das amostras à esquerda e das variáveis direita para zona de baixo potencial produtivo.



Assim como nas outras duas classes, o modelo gerado para a classe 3 também apresenta as mesmas variáveis como sendo as de maior poder de modelamento sendo MacPor30 (0,79), MacPor50 (0,75) e Θ_{33} ($\Theta_{33} = 0,65$), correlatas a estrutura do solo.

Figura 25 - Disposição gráfica do poder de modelagem das variáveis para zona de baixo potencial produtivo.



Em muitos dos resultados apresentados há a repetição da variável Macro Poro, variável esta que está correlaciona a estrutura de solo, nos aponta como possível problema o manejo da área. Em estudo realizado por Cereta e Aita (2010) indica que as principais formas de degradação física do solo estão ligadas ao mal manejo da área, como selamento superficial, compactação e adensamento, danos à estrutura, alterações na porosidade e permeabilidade, erosão, inundação e drenagem de áreas úmidas. Ainda a implementação de manejos e cultivos que contribuam para a estruturação e aumento da macroporosidade do solo são relevantes para o aumento de produtividade da área em estudo.

CONCLUSÃO

O estudo de avaliação da relação entre o mapa de zonas de potencial produtivo, gerado por meio da metodologia proposta por Guanais Santos, et al. (2023) utilizando mapas de produtividade como dados de entrada, com as variáveis de solo, apresentou relação de 84%. Comprovando que esta metodologia é válida e útil para a

realização de intervenções na área, conforme estudo de caso.

O estudo de caso realizado, apresentou que as variáveis com maior poder de modelamento estão ligadas à estrutura do solo e a movimentação de água e ar, sendo a macroporosidade a mais importante.

Também se pode afirmar que a metodologia de separação em zonas de potencial produtivo é útil para a alocação e redução do número de pontos amostrais em uma área agrícola, direcionando a amostragem por meio de informação quantitativa, evitando amostragens em grade, de forma aleatória e/ou empírica e em grande quantidade.

REFERÊNCIAS

- AGUIAR, P. A. A. **Padronização da umidade dos grãos na experimentação agrícola**. In: Embrapa Semiárido-Resumo em anais de congresso (ALICE). In: EMPRESA BRASILEIRA DE PESQUISA AGROPECUÁRIA. Centro de Pesquisa Agropecuária do Tópico Semi-Árido. Petrolina. Resumo de atividades de pesquisa. Petrolina, 1977.
- AWAN, S. 2016. **A review of the past and future of PA in the UK**. Research review No. 87. Agriculture and Horticulture Development Board, Stoneleigh Park, Kenilworth.
- BAZZI, C. L., SOUZA, E. G., OPAZO, M. A. U., NÓBREGA, L. H., PINHEIRO NETO, R. Influência da distância entre passadas de colhedora equipada com monitor de colheita na precisão dos mapas de produtividade na cultura do milho. **Engenharia Agrícola**, v. 28, n. 2, p. 355-363, 2008.
- BAKKALI, S., AMRANI, M. About the use of spatial interpolation methods to denoising Moroccan resistivity data phosphate "Disturbances" map. **Acta Montanistica Slovaca**, v.13, n. .7, p. 216–222, 2008.
- BOYER, C. N., WADE, B., SOLIE, J. B., RAUN, W. R. **Profitability of variable rate nitrogen application in wheat production**. PA v.12, p.473–487. 2011. doi:10.1007/s11119-010-9190-5.
- CAMPBELL, J. B.; WYNNE, R. H. **Introduction to remote sensing**. Guilford Press, 2011.
- CARMO, E. J.; RODRIGUES, D. D.; SANTOS, G. R. Avaliação dos interpoladores krigagem e topo to raster para geração de modelos digitais de elevação a partir de um "as built". **Boletim de Ciências Geodésicas**, v. 21, p. 674-690, 2015.
- CERETTA, C. A.; AITA, C. **Manejo e conservação do solo**. 2010.
- CARVALHO, J. R. P.; VIEIRA, S. R.; MORAN, R. C. C. P. Análise de

correspondência-uma ferramenta útil na interpretação de mapas de produtividade. **Revista brasileira de ciência do solo**, v. 26, p. 435-443, 2002.

CHLINGARYAN, A.; SUKKARIEH, S., WHELAN, B. Machine learning approaches for crop yield prediction and nitrogen status estimation in precision agriculture: A review, **Computers and Electronics in Agriculture**, v. 151, p. 61-69, 2018. <https://doi.org/10.1016/j.compag.2018.05.012>.

COLEMAN, T. L.; TADESSE, W. Differentiating soil physical properties from multiple band DOQ data. **Soil science**, v. 160, n. 2, p. 81-91, 1995.

COOK, S. E., BRAMLEY, R. G. V. Precision agriculture-opportunities, benefits and pitfalls of site specific crop management in Australia. **Australian Journal of Experimental Agriculture**, p.753-763, 1998.

DELALIBERA, H. C.; HIGASHIBARA, L. R.; RALISCH, R. **Estimativa de perdas na colheita mecanizada de grãos: Adequação da metodologia para plataformas universais modernas**. In: Anais do XLVI Congresso Brasileiro de Engenharia Agrícola – CONBEA, 2017.

DEUS, B. V.; ZEILHOFER, P.; ARAUJO, G. C.; SANTOS, A. S. L. **Interpolação pluviométrica na Bacia do Alto e Médio Rio Teles Pires: uma análise de séries históricas e interpoladores**. In: SIMPÓSIO BRASILEIRO DE CIÊNCIAS GEODÉSICAS E TECNOLOGIAS DA GEOINFORMAÇÃO, 3., 2010, Recife. Anais... Recife: UFPE, 2010. p.1-7.

DONAGEMMA, G. K.; VIANA, J. H. M.; DE ALMEIDA, B. G.; RUIZ, H. Á.; KLEIN, V. A.; DECHEN, S. C. F.; FERNANDES, R. B. A. **Análise granulométrica**. EMBRAPA Solos, 2017.

EMPRESA BRASILEIRA DE PESQUISA AGROPECUÁRIA - EMBRAPA. Serviço Nacional de Levantamento e Conservação de Solo. **Manual de métodos de análise de solo**. Rio de Janeiro, Ministério da Agricultura, 1997.

EMPRESA BRASILEIRA DE PESQUISA AGROPECUÁRIA - EMBRAPA. Centro Nacional de Pesquisa de Solos. **Manual de métodos de Análise de Solo**. 3.ed. Brasília: EMBRAPA, p. 573. 2017.

GALLICHAND, J., BUCKLAND, G. D., MARCOTTE, D., HENDRY, M. J. Spatial interpolation of soil salinity and sodicity for a saline soil in Southern Alberta. **Canadian Journal of Soil Science**, v. 72, p. 503 – 516, 1992.

GOODALE, C. L., ABER, J. D., OLLINGER, S. V. Mapping monthly precipitation, temperature, and solar radiation for Ireland with polynomial regression and a digital elevation model. **Climate Research**, v. 10, p. 35 – 49, 1998.

GUANAIS SANTOS, L. M., SAAB, O. J. G. A., GUIMARÃES, M. D. F., RALISCH, R., & DELALIBERA, H. C. Methodology for estimating productive potential zones from productivity data. **Semina: Ciênc. Agrár**, v. 44, n. 3, p. 1001 - 1016. 2023. Doi: 10.5433/1679-0359.2022v44n3p1001.

HARTER, H. L. Tables of range and studentized range. **The Annals of**

Mathematical Statistics, p. 1122 - 1147, 1960.

HAIR, J. F.; BLACK, W. C.; BABIN, B. J.; ANDERSON, R. E.; TATHAM, R. L. **Análise multivariada de dados**. Bookman editora, 2009.

HORNUNG, A., KHOSLA, R., REICH, R., INMAN, D., WESTFALL, D. Comparison of site-specific management zones: Soil color based and yield based. **Agronomy Journal**, v.98, p. 405–417. 2006.

ISAAKS, E. H., SRIVASTAVA, R. M. 1989. **An Introduction to Applied Geostatistics**. New York: Oxford University Press. Jones, N. L., R. J. Davis, and W. Sabbah. 2003. A comparison of three-dimensional interpolation techniques for plume characterization. *Ground Water* v.41, p. 411–419. 2003.

JIMENEZ, K. Q.; DOMEQ, F. M. **Estimação de chuva usando métodos de interpolação**. Porto Alegre: Instituto de Pesquisas Hidráulicas. Universidade Federal do Rio Grande do Sul, p.16, 2008.

KHOSLA, R., WESTFALL, D., REICH, R., MAHAL, J. S., GANGLOFF, W. J. 2010. **Spatial variation and site-specific management zones**. In: *Geo-Statistical Applications in Precision Agriculture*, ed. M. Oliver, pp. 195–219. Amsterdam: Springer.

KIEHL, E. J. **Manual de edafologia**. São Paulo: Agronômica Ceres, 1979. 262p.

KÖPPEN, W. P. **Grundriss der klimakunde**. W. de Gruyter, 1931.

KOUTSOS, T., MENEXES, G. **Beneficia da adoção de tecnologias agrícolas de precisão. Uma revisão sistemática**. In: 18º Congresso Florestal Panhellenic e Workshop Internacional “Tecnologia da Informação, Desenvolvimento Sustentável, Rede Científica e Proteção da Natureza”. Edessa, Grécia. 2017.

Krige D (1951) **Uma abordagem estatística para alguns problemas básicos de avaliação de minas em Witwatersrand**. *J Southern Afr Inst Min Metall* 52(6):119–139.

KRAVCHENKO, A. N.; BULLOCK, D. G.; BOAST, C. W. **Joint multifractal analysis of crop yield and terrain slope**. *Agronomy Journal*, v. 92, n. 6, p. 1279-1290, 2000.

LANDIM, P. M. B. **Introdução aos métodos de estimação espacial para confecção de mapas**. Lab. Geomatématica, DGA, IGCE, UNESP/Rio Claro, Texto Didático 03, 2000.

LI, M., CHUNG, S. 2015. Special issue on precision agriculture. **Comput Electron Agric.**112, pp. 1.

LYLE, G.; BRYAN, B. A., OSTENFORD, B. Post-processing methods to eliminate erroneous grain yield measurements: review and directions for future development. **Precision agriculture**, v. 15, n. 2, p. 377-402, 2014.

MALDANER, L. F., CORRÊDO, L. P., TAVARES, T. R., MENDEZ, L. G., DUARTE,

- C., MOLIN, J. P. Identificando e filtrando outliers em conjuntos de dados espaciais. **In: Anais da 14ª Conferência Internacional sobre Agricultura de Precisão**, Montreal, QC, Canadá, 2018. p. 24-27.
- MAUNE, T. A., BLACK, E. W. DEM user requirements. In: MAUNE, T.A. (Ed). **Digital elevation model technologies and applications**. Editora Bethesda, MD, p. 449 - 473. 2001.
- MELO, T. R., MACHADO, W., TAVARES FILHO, J. Correlation os cationic indeces with clay dispersion degree of two soil from Brazil fertilized with chicken manure. **International Journal of Plant and Soil Science**, v. 4, n.4, p. 338 - 351, 2015.
- MOLIN, J. P.; MENEGATTI, L. A. A. **Mapeamento da resistência elétrica do solo por sensor de contato**. 2º SIMPÓSIO INTERNACIONAL DE AGRICULTURA DE PRECISÃO, Viçosa, MG, Brasil, Anais, v. 1, 2002.
- MENEGATTI, L. A. A., MOLIN, J. P. Remoção de erros em mapas de produtividade via filtragem de dados brutos. **Revista Brasileira de Engenharia Agrícola e Ambiental**, v. 8, n. 1, p. 126 - 134. 2004.
- MICHELAN, R., SOUZA, E. G., OPAZO, M. A. U. Determinação e remoção do tempode atraso em mapas de colheita de milho. **Acta Scientiarum**. Agronomy, v. 29, n. 2, p. 147 - 155, 2007.
- MOLIN, J. P. **Geração e interpretação de mapas de produtividade para agricultura de precisão**. BORÉM, A.; GIÚDICE, MP; QUEIROZ, DM; MANTOVANI, EC, p. 237 - 58, 2000.
- MOLIN, J. P. Definição de unidades de manejo a partir de mapas de produtividade. **Engenharia agrícola**, Jaboticabal, v. 22, n. 1. p. 83 - 92, 2002.
- MOLIN, J. P., DO AMARAL, L. R., COLAÇO, A. **Agricultura de precisão**. Oficina deTextos, 2015.
- MOLIN, J. P. **Agricultura de precisão: números do mercado brasileiro**. Boletim Técnico, v. 3, p. 7, 2017.
- MOORE, M. **An investigation into the accuracy of yield maps and their subsequent use in crop management**. Silsoe: Silsoe College, 1998.
- OLIVEIRA, R. P.; GREGO, C. R.; BRANDÃO, Z. N. **Geostatística aplicada na agricultura de precisão utilizando o Vesper**. 2015.
- PANAGOPOULOS, Y., et al. Assessing the cost-effectiveness of irrigation water management practices in water stressed agricultural catchments: The case of Pinios. **Agricultural Water Management**, v.139, p.31 – 42. 2014.
- PATHAK, H. S., BROWN, P., BEST, T. Uma revisão sistemática da literatura sobre os fatores que afetam o processo de adoção da agricultura de precisão. **Precision Agriculture**, v.20, p.1292 – 1316. 2019.
- PAVAN, M.A.; BLOCH, M.F.; ZEMPULSKI, H.C.; MIYAZAWA, M. & ZOCOLER, D.C.

Manual de análise química do solo e controle de qualidade. Instituto Agronômico do Paraná, (IAPAR Londrina), 1992.

RALISCH, R., et al. **Diagnóstico Rápido da Estrutura do Solo: DRES.** EMBRAPA Soja, 2017.

ROBERT, P. C. Precision agriculture: a challenge for crop nutrition management. **Plant and Soil**, v. 247, p.143 - 149, 2002.

SILVA, F. M. et al. Variabilidade espacial de atributos químicos e produtividade da cultura do café em duas safras agrícolas. **Ciência e Agrotecnologia**, v. 32, p. 231 - 241, 2008.

SILVA, D. V. R, ANTUNES, M. U. F, GIROTTO, J., FELIPE, P. O. M., SANGOI, L. F., GIOTTO, E. **Mapas de produtividade da agricultura de precisão.** 2011. Disponível em < <https://mundogeo.com/2000/01/01/mapas-de-produtividade-na-agricultura-de-precisao/> >. Acesso em: 03 de agosto de 2020.

SNEDECOR G. W., COCHRAN, W. G. Statistical Methods. Iowa State University Press. **Biometry**, ed.7, p.593, 1980.

SOUZA, E. G., BAZZI, C. L., KHOSLA, R., URIBE-OPAZO, M. A., REICH, R. M. Interpolation type and data computation of crop yield maps is important for precision crop production, **Journal of Plant Nutrition**, v. 39, n. 4, p. 531 - 538, 2016.

SPEKKEN, M., ANSELMINI, A. A., MOLIN, J. P. **A simple method for filtering spatial data.** In: EUROPEAN CONFERENCE ON PRECISION AGRICULTURE, 9, Lleida. Anais ... Lleida: Fundação Universitária de Lleida. Sem página. 2013.

SRINIVASAN, A., et al. **Manual de agricultura de precisão: princípios e aplicações** . Imprensa CRC, 2006.

SUDDUTH, K. A. et al. **Sensoriamento por indução eletromagnética como indicador de produtividade em solos argilosos.** In: Gestão específica do local para sistemas agrícolas. Madison, WI, EUA: Sociedade Americana de Agronomia, Crop Science Society of America, Soil Science Society of America, p. 671 - 681. 1995.

UMBERS, A., WATSON, P. E WATSON, D. **Relatório de pesquisa de práticas agrícolas.** 2015.

VIEIRA, L. S.; SANTOS, P. C. T. C.; VIEIRA, N. F. **Solos: propriedades, classificação e manejo.** Brasília: MEC/ABEAS, 1988.

WACHOWIAK, M. P., WALTERS, D. F., KOVACS, J. M., WACHOWIAK-SMOLÍKOVÁ, R., JAMES, A. L. Análise visual e imagens de sensoriamento remoto para apoiar a pesquisa baseada na comunidade para agricultura de precisão em áreas emergentes. **Computers and Electronics in Agriculture**, v. 143, p. 149 - 164, 2017.

WEST, G. H., KOVACS, K. Addressing Groundwater Declines with Precision Agriculture: An Economic Comparison of Monitoring Methods for Variable-Rate

Irrigation. **Water**, v.9, p.28. 2017.

WU, C. J. et al. Jackknife, bootstrap and other resampling methods in regression analysis. **The Annals of Statistics**, v. 14, n. 4, p. 1261-1295, 1986.

ZHANG, C., KOVACS, J. M., The application of small unmanned aerial systems for precision agriculture: a review. **Precision Agriculture**. v.13, n.6, p.693–712. 2012.

ANEXO

Tabela 3 – Variáveis utilizadas nos modelos multivariados com suas respectivas zonas de potencial produtivo.

Zona de potencial produtivo	Classe	Identificação da amostra	SiltM	ADA base teor de argila (Arg)	pH CaCl	pH + AL	AL	Ca	K	P	19 - 8	4 - 2	1 - 0.5	0.5 - 0.25	<0.25	Porosidade total (PorTot)	Porosidade total calculada (PTc)	Porosidade total determinada (PTd)	ΘSat	Θ10	Θ33	Θ1500	MacPor50	MacPor30	MicPor30									
			%	(cmolc dm ⁻³)						(mg dm ⁻³)					Malha das peneiras para avaliação da estabilidade de agregados (mm)					cm ³ .cm ³					Pressão no extrator de richards (kPa)					cm ³ .cm ³				
Alto	1	1	15.94	94.20	6.10	3.73	0.03	5.34	0.28	35.03	92.57	133.68	246.35	165.62	135.32	0.44	0.45	0.51	0.33	0.41	0.36	0.30	0.02	0.04	0.04	0.41								
		2	14.41	91.35	5.97	3.81	0.04	5.37	0.32	27.27	207.03	98.22	174.25	141.09	133.15	0.43	0.45	0.49	0.32	0.40	0.37	0.32	0.04	0.05	0.05	0.40								
		3	14.53	92.85	5.87	3.15	0.02	7.55	0.45	26.49	266.88	104.61	182.38	101.64	76.57	0.42	0.39	0.46	0.30	0.36	0.31	0.24	0.01	0.04	0.04	0.36								
		4	12.91	92.72	5.90	3.07	0.04	5.61	0.08	9.09	328.74	77.97	161.70	129.08	86.29	0.45	0.44	0.49	0.33	0.36	0.33	0.27	0.07	0.08	0.08	0.36								
		5	11.96	89.40	5.83	3.21	0.04	5.57	0.09	15.25	188.53	122.89	172.14	123.60	83.63	0.46	0.47	0.52	0.35	0.38	0.35	0.30	0.08	0.09	0.09	0.38								
		6	11.51	65.81	5.80	3.19	0.03	5.49	0.31	42.51	174.03	96.35	202.00	150.00	115.41	0.47	0.46	0.50	0.35	0.38	0.34	0.30	0.07	0.08	0.08	0.38								
		7	14.94	89.73	5.73	3.26	0.06	7.08	0.31	13.50	383.82	83.31	140.48	97.99	81.55	0.47	0.47	0.53	0.37	0.38	0.34	0.28	0.08	0.09	0.09	0.38								
		Média	13.74	88.01	5.89	3.35	0.04	6.00	0.26	24.16	234.51	102.43	182.76	129.86	101.70	0.45	0.45	0.50	0.34	0.38	0.34	0.29	0.05	0.07	0.07	0.38								
		DP	1.64	9.94	0.12	0.30	0.01	0.91	0.13	12.18	99.07	20.08	33.74	24.71	25.52	0.02	0.03	0.02	0.03	0.03	0.02	0.03	0.03	0.03	0.03	0.03	0.02							
		CV	11.96	11.29	2.04	8.84	35.31	15.23	51.33	50.41	42.25	19.60	18.46	19.03	25.10	4.39	6.02	4.47	7.54	4.71	6.45	9.19	56.53	37.73	4.71									
Médio	2	8	13.13	87.27	5.60	3.98	0.02	5.04	0.21	29.72	175.57	104.54	204.76	149.82	83.41	0.48	0.48	0.50	0.35	0.41	0.37	0.32	0.06	0.07	0.07	0.41								
		9	14.94	87.60	5.50	4.19	0.10	5.47	0.09	14.29	167.95	103.02	186.92	145.12	118.22	0.50	0.48	0.53	0.38	0.37	0.34	0.29	0.09	0.11	0.11	0.37								
		10	14.67	92.97	5.50	3.99	0.08	6.45	0.11	20.86	154.39	117.45	191.02	144.55	73.53	0.49	0.51	0.53	0.37	0.42	0.37	0.31	0.08	0.09	0.09	0.42								
		11	11.29	112.01	5.60	3.79	0.04	6.82	0.35	16.49	237.62	92.32	163.57	134.90	115.71	0.45	0.45	0.47	0.31	0.43	0.39	0.34	0.01	0.02	0.02	0.43								
		12	12.89	85.58	5.53	3.79	0.06	7.38	0.37	13.41	281.93	101.15	156.10	121.74	75.35	0.44	0.43	0.46	0.30	0.39	0.36	0.31	0.03	0.04	0.04	0.39								
		13	12.71	91.03	5.47	3.81	0.04	8.34	0.37	16.55	213.31	120.69	135.83	96.21	132.27	0.41	0.42	0.46	0.28	0.41	0.38	0.33	0.01	0.01	0.01	0.41								
		14	16.92	102.72	5.47	3.35	0.06	11.64	0.37	17.47	248.93	118.74	159.55	103.10	43.43	0.42	0.42	0.43	0.27	0.39	0.35	0.28	0.02	0.03	0.03	0.39								
		Média	13.79	94.17	5.52	3.84	0.06	7.31	0.27	18.40	211.38	108.27	171.11	127.92	91.70	0.45	0.45	0.48	0.32	0.40	0.37	0.31	0.04	0.05	0.05	0.40								
		DP	1.85	9.73	0.06	0.26	0.03	2.21	0.13	5.54	47.43	10.76	23.94	21.45	31.41	0.03	0.03	0.04	0.05	0.02	0.02	0.02	0.02	0.03	0.04	0.02								
		CV	13.42	10.33	1.03	6.82	47.79	30.27	47.61	30.10	22.44	9.94	13.99	16.77	34.26	7.23	7.59	7.83	13.77	4.49	5.11	6.41	76.59	73.23	4.49									
Baixo	3	15	12.88	99.69	5.40	3.73	0.06	7.00	0.15	26.01	403.04	77.23	126.00	100.29	88.46	0.51	0.48	0.51	0.37	0.39	0.36	0.30	0.07	0.09	0.09	0.39								
		16	11.56	77.53	5.30	4.30	0.16	7.48	0.22	32.65	210.60	95.67	190.50	148.22	102.54	0.47	0.48	0.49	0.33	0.42	0.39	0.29	0.05	0.06	0.06	0.42								
		17	13.13	94.94	5.30	2.99	0.05	10.03	0.43	26.04	167.56	74.32	156.40	160.71	228.24	0.53	0.54	0.55	0.41	0.37	0.34	0.29	0.16	0.17	0.17	0.37								
		18	11.68	60.04	5.23	3.11	0.05	7.15	0.67	92.43	160.20	74.45	185.35	173.13	219.21	0.50	0.51	0.53	0.38	0.42	0.39	0.33	0.08	0.09	0.09	0.42								
		19	7.92	86.38	5.20	3.71	0.13	4.39	0.38	34.60	245.11	79.65	164.53	145.06	167.52	0.42	0.43	0.46	0.28	0.41	0.39	0.34	0.02	0.02	0.02	0.41								
		20	11.89	86.21	5.17	3.59	0.06	5.77	0.44	31.47	167.33	85.38	197.67	153.31	146.55	0.43	0.44	0.47	0.29	0.43	0.40	0.34	0.01	0.01	0.01	0.43								
		21	9.05	65.41	5.17	3.71	0.11	3.89	0.34	42.00	191.86	69.81	175.92	177.68	178.42	0.52	0.49	0.57	0.42	0.41	0.37	0.32	0.07	0.08	0.08	0.41								
		Média	11.16	81.46	5.25	3.59	0.09	6.53	0.38	40.74	220.81	79.50	170.91	151.20	161.56	0.48	0.48	0.51	0.35	0.41	0.38	0.32	0.07	0.08	0.08	0.41								
		DP	1.95	14.68	0.09	0.44	0.04	2.08	0.17	23.44	85.74	8.64	24.50	25.55	53.45	0.04	0.04	0.04	0.06	0.02	0.02	0.02	0.05	0.05	0.05	0.02								
		CV	17.43	18.02	1.63	12.18	49.32	31.83	45.03	57.52	38.83	10.86	14.34	16.90	33.08	9.19	7.90	8.28	16.00	5.13	5.54	6.97	75.80	68.62	5.13									