

UNIVERSIDADE TECNOLÓGICA FEDERAL DO PARANÁ
PROGRAMA DE PÓS-GRADUAÇÃO EM TECNOLOGIAS COMPUTACIONAIS
PARA O AGRONEGÓCIO - PPGTCA

RODRIGO WERNER DALL'AGNOL

**APLICAÇÃO WEB INTEGRADA AO SOFTWARE R PARA ANÁLISE
ESPACIAL E GERAÇÃO DE MAPAS TEMÁTICOS EM
AGRICULTURA DE PRECISÃO**

DISSERTAÇÃO

MEDIANEIRA - PR

2017

RODRIGO WERNER DALL'AGNOL

**APLICAÇÃO WEB INTEGRADA AO SOFTWARE R PARA ANÁLISE
ESPACIAL E GERAÇÃO DE MAPAS TEMÁTICOS EM
AGRICULTURA DE PRECISÃO**

DISSERTAÇÃO

Dissertação apresentada ao Programa de Pós-Graduação em Tecnologias Computacionais para o Agronegócio - PPGTCA – da Universidade Tecnológica Federal do Paraná – UTFPR – Campus Medianeira, como requisito parcial para obtenção do título de Mestre em Tecnologias Computacionais para o Agronegócio / Área de Concentração: Tecnologias Computacionais Aplicadas à Produção Agrícola e Agroindústria.

Orientador: Prof. Dr. Claudio L. Bazzi.

MEDIANEIRA - PR

2017

Dados Internacionais de Catalogação na Publicação

D144a

Dall'Agnol, Rodrigo Werner

Aplicação web integrada ao software R para análise espacial e geração de mapas temáticos em agricultura de precisão / Rodrigo Werner Dall'Agnol -2018
85 f. : il. ; 30 cm.

Orientador: Claudio Leones Bazzi.

Dissertação (Mestrado) – Universidade Tecnológica Federal do Paraná. Programa de Pós-Graduação em Tecnologias Computacionais para o Agronegócio. Medianeira, 2017.

Inclui bibliografias.

1. Agricultura de precisão. 2. Serviços da Web. 3. Interpolação. 4. Tecnologias Computacionais - Dissertações. I. Bazzi, Claudio Leones. orient. II. Universidade Tecnológica Federal do Paraná. Programa de Pós-Graduação em Tecnologias Computacionais para o Agronegócio. IV. Título.

CDD: 004

Biblioteca Câmpus Medianeira
Marci Lucia Nicodem Fischborn 9/1219



TERMO DE APROVAÇÃO

APLICAÇÃO WEB INTEGRADA AO SOFTWARE R PARA ANÁLISE ESPACIAL E GERAÇÃO DE MAPAS TEMÁTICOS EM AGRICULTURA DE PRECISÃO

Por

RODRIGO WERNER DALL'AGNOL

Essa dissertação foi apresentada as treze horas e trinta minutos, do dia treze de dezembro de dois mil e dezessete, como requisito parcial para a obtenção do título de Mestre em Tecnologias Computacionais para o Agronegócio, Linha de Pesquisa Tecnologias Computacionais Aplicadas à Produção Agrícola, no Programa de Pós-Graduação em Tecnologias Computacionais para o Agronegócio - PPGTCA, da Universidade Tecnológica Federal do Paraná. O candidato foi arguido pela Banca Examinadora composta pelos professores abaixo assinados. Após deliberação, a Banca Examinadora considerou o trabalho aprovado.

Prof. Dr. Claudio Leones Bazzi (Orientador – PPGTCA)

Prof. Dr. Paulo Sérgio Graziano Magalhães (Membro Interno – PPGTCA)

Dr. Luis Henrique Bassoi (Membro Externo – EMBRAPA INSTRUMENTAÇÃO)

A via original com as assinaturas encontra-se na secretaria do programa.

AGRADECIMENTOS

Gostaria de enfatizar meu agradecimento ao meu orientador Prof. Dr. Claudio Leones Bazzi, pela confiança, pela oportunidade, dedicação e ensinamentos em todas as etapas deste trabalho.

Ao Prof. Nelson Miguel Betzek, pela disponibilidade e por todos os ensinamentos que contribuíram para realização desse projeto.

A todos meus colegas de trabalho, pela força, vibração e colaboração.

A todos os professores do PPGTCA que forneceram ensinamentos valiosos.

E meu reconhecimento a minha família, pelo suporte, compreensão e apoio durante essa jornada.

"Faça o seu melhor, na condição que você tem, até que você tenha condições melhores, para fazer melhor ainda."

Mario Sergio Cortella

RESUMO

DALL'AGNOL, Rodrigo Werner. **Aplicação web integrada ao software R para análise espacial e geração de mapas temáticos em agricultura de precisão.** 2017. Dissertação (Mestrado em Tecnologias Computacionais para o Agronegócio) - Programa de Pós-graduação em Tecnologias Computacionais para o Agronegócio, Universidade Tecnológica Federal do Paraná.

O desenvolvimento de tecnologias agrícolas que possibilitem o aumento da produtividade e a redução de custo de produção, obtendo-se maior lucratividade é o objetivo de produtores e pesquisadores de todo o mundo. Utilizando a agricultura de precisão como fundamento de aplicação de tecnologias no campo, este estudo apresenta o desenvolvimento e a implementação de uma ferramenta computacional em ambiente web para análise espacial de dados agrícolas e geração de mapas temáticos. O software tem como característica possibilitar na etapa de análise geoestatística a seleção automática de parâmetros do semivariograma e realizar a indicação do melhor modelo ajustado, não sendo necessário o ajuste manual de modelos ao semivariograma experimental. A partir da integração com o software estatístico R, utilizando os dados de atributos físicos e químicos do solo e a produtividade de três propriedades agrícolas, foi realizada a análise geoestatística e estatística convencional e a geração de mapas temáticos utilizando as técnicas de interpolação conhecidas como krigagem ordinária e o inverso da distância elevada a uma potência. A comparação realizada com software comercial mostrou que a ferramenta desenvolvida na pesquisa é uma interessante opção para geração de mapas temáticos destinados a fornecer informações sobre a variabilidade espacial e temporal dos atributos de solo e planta, possibilitando assim uma tomada de decisão mais assertiva por parte dos produtores para correção dos atributos analisados.

Palavras-chave: Agricultura de precisão, análise geoestatística, interpolação de dados.

ABSTRACT

DALL'AGNOL, Rodrigo Werner. **Integrated web application to R software for spatial analysis and generation of thematic maps in precision agriculture**. 2017. Dissertação (Mestrado em Tecnologias Computacionais para o Agronegócio) - Programa de Pós-graduação em Tecnologias Computacionais para o Agronegócio, Universidade Tecnológica Federal do Paraná.

The development of agricultural technologies that make it possible to increase productivity and reduce the cost of production, achieving greater profitability is the goal of producers and researchers from all over the world. Using precision agriculture as a base for application of technologies in the field, this study presents the development and implementation of a computational tool in the web environment for spatial analysis of agricultural data and generation of thematic maps. The software allows the automatic selection of parameters of the semivariogram in the geostatistical analysis step and the indication of the best adjusted model, and it is not necessary to manually adjust the models to the experimental semivariogram. From the integration with the statistical software R, using the data of physical and chemical attributes of the soil and the productivity of three farms, the geostatistical analysis and conventional statistics and the generation of thematic maps were performed using interpolation techniques known as kriging and the inverse of the elevated distance to a power. The comparison with commercial software showed that the tool developed in the research is an interesting option for generating thematic maps to provide information on the spatial and temporal variability of soil and plant attributes, thus enabling a more assertive decision making by the producers to correct the analyzed attributes.

Key-words: Precision agriculture, geostatistical analysis, data interpolation.

LISTA DE ILUSTRAÇÕES

Figura 1 - Estrutura Semivariograma.	27
Figura 2 - Modelos teóricos.	28
Figura 3 - Representação gráfica do modelo esférico.	29
Figura 4 - Representação gráfica dos modelos Família Matérn.	30
Figura 5 - Representação gráfica modelo exponencial.	31
Figura 6 - Representação gráfica modelo gaussiano.	32
Figura 7 - Mapas temáticos representativo da distribuição do elemento Fósforo em uma propriedade agrícola.	38
Figura 8 - Representação geral do projeto com destaque as funcionalidades foco deste trabalho.	44
Figura 9 - Diagrama das etapas do projeto.	45
Figura 10 - Áreas utilizadas para pesquisa.	47
Figura 11 - Representação da estrutura e funcionamento do módulo proposto.	49
Figura 12 - Gráficos descritivos que apresentam o comportamento dos dados para interpretação do usuário.	51
Figura 13 - Opções de análises disponíveis no módulo desenvolvido.	54
Figura 14 - Representação da tela com os dados retornados para escolha da área a ser analisada.	55
Figura 15 - Representação dos dados retornados ao usuário para escolha da amostra a ser analisada.	56
Figura 16 - Tela para realização da análise geoestatística.	58
Figura 17 - Código de ligação entre o módulo desenvolvido e o software R.	59
Figura 18 - Tela de visualização das análises realizadas.	60
Figura 19 - Representação gráfica da estatística descritiva dos dados apresentados na tela de visualização.	61
Figura 20 - Representação gráfica dos modelos ajustados e do melhor modelo ajustado.	62
Figura 21 - Tabela com os valores dos doze melhores modelos.	63
Figura 22 - Apresentação do mapa interpolado por Krigagem Ordinária.	64
Figura 23 - Tela para realização da interpolação pelo inverso da distância elevado a uma potência.	65
Figura 24 - Apresentação do mapa interpolado por inverso da distância elevado a uma potência.	66
Figura 25 - Comparativo por interpolador dos mapas temáticos da Área A.	68
Figura 26 - Análise geoestatística Área A.	69
Figura 27 - Comparativo por interpolador dos mapas temáticos da Área B.	70
Figura 28 - Análise geoestatística Área B.	71
Figura 29 - Comparativo por interpolador dos mapas temáticos da Área C.	72
Figura 30 - Análise geoestatística Área C.	73
Figura 31 - Comparativo por software dos mapas temáticos da Área A - Interpolador IDP.	75
Figura 32 - Comparativo por software dos mapas temáticos da Área A - Interpolador KRIG.	76
Figura 33 - Comparativo por software dos mapas temáticos da Área C - Interpolador IDP.	77
Figura 34 - Comparativo por software dos mapas temáticos da Área A - Interpolador KRIG.	78

LISTA DE TABELAS

Tabela 1 - Identificação das variáveis (atributos) e períodos de coleta, para cada área experimental.....	47
---	----

SUMÁRIO

1 INTRODUÇÃO	14
1.1 OBJETIVO GERAL	17
1.2 OBJETIVOS ESPECÍFICOS	17
2 REVISÃO DE LITERATURA	18
2.1 AGRICULTURA DE PRECISÃO	18
2.2 SISTEMA DE INFORMAÇÃO GEOGRÁFICA.....	20
2.3 SOFTWARE ESTATÍSTICO R	21
2.4 ANÁLISE EXPLORATÓRIA DE DADOS.....	22
2.5 GEOSTATÍSTICA.....	24
2.6 INTERPOLAÇÃO ESPACIAL DE DADOS	35
2.7 MAPAS TEMÁTICOS	38
3 MATERIAIS E MÉTODOS	40
3.1 FERRAMENTAS E TECNOLOGIAS EMPREGADAS	40
3.2 CONTEXTO EM QUE O PROJETO SE INSERE.....	42
3.3 ÁREAS EXPERIMENTAIS E DADOS UTILIZADOS NO ESTUDO.....	46
3.4 FUNCIONAMENTO DO MÓDULO PROPOSTO	48
3.5 ANÁLISE EXPLORATÓRIA E GEOESTATÍSTICA	50
3.6 INTERPOLAÇÃO DE DADOS.....	52
3.7 MAPAS TEMÁTICOS	53
3.8 COMPARAÇÃO DE MAPAS TEMÁTICOS ENTRE FERRAMENTAS	53
4 RESULTADOS E DISCUSSÃO	54
4.1 INTERFACE E PROCEDIMENTOS DISPONÍVEIS NO MÓDULO	54
4.2 CONSUMO DE DADOS PARA REALIZAÇÃO DAS ANÁLISES PELO MÓDULO DESENVOLVIDO	55
4.3 PROCEDIMENTO PARA REALIZAÇÃO DA ANÁLISE GEOESTATÍSTICA E INTERPOLAÇÃO DE DADOS POR KRIGAGEM ORDINÁRIA.....	57
4.4 PROCEDIMENTO PARA REALIZAÇÃO DA INTERPOLAÇÃO DE DADOS PELO INVERSO DA DISTÂNCIA ELEVADO A UMA POTÊNCIA.....	65
4.5 APLICAÇÃO DO MÓDULO SOBRE AS ÁREAS EXPERIMENTAIS E SEUS ATRIBUTOS.....	67
4.6 COMPARATIVO ENTRE O MÓDULO DESENVOLVIDO E O SOFTWARE ARCMAP.....	74
5 CONCLUSÃO	79
REFERÊNCIAS	80

1 INTRODUÇÃO

O conhecimento da relação entre a produtividade das culturas agrícolas e de atributos do solo e do relevo, assim como das principais limitações da produção, torna-se fundamental quando se tem como objetivo o manejo racional da produção, reduzindo a exaustão química e a degradação dos atributos físicos do solo. Por meio do manejo racional, verifica-se o aumento do rendimento na produção das principais culturas, quando comparado à realidade da década de 1990, não estando este atrelado a somente ao aumento da produção biológica das culturas, mas sim da melhor adaptação dos cultivares e do uso intensivo de tecnologia, em especial as relacionadas à fertilização (HORTON, 2000).

Esse aumento da produção agrícola promoveu tanto o desenvolvimento da mecanização no campo como o cultivo de áreas cada vez maiores. Devido às limitações de tempo e da produção em áreas maiores, a condição de se ter conhecimento localizado do talhão passou a ser dificultado, tornando a identificação e o conhecimento de características de solo em locais específicos insuficientes, comparado a quando se trabalha com áreas menores de fácil identificação. Segundo Molin (2012) apesar das diferenças locais, normalmente realiza-se o manejo e aplicação de insumos considerando a sua disponibilidade média, o que pode causar problemas em certos locais por receberem quantidades insuficientes, e em outros, por receberem superdosagens, podendo propiciar problemas ambientais (COELHO *et al.*, 2009).

Para resolver este tipo de problema utiliza-se um conceito chamado agricultura de precisão (AP) que tem por objetivo avaliar as características do solo e planta de forma local, visando proporcionar uma aplicação mais exata de insumos, no momento adequado e em quantidades necessárias com o mínimo de desperdício. A agricultura de precisão corresponde a um conjunto de tecnologias e procedimentos utilizados para que as lavouras e os sistemas de produção sejam aperfeiçoados, tendo como principal elemento o gerenciamento e monitoramento intensivo do campo (SWINTON, 1998). Os recursos de agricultura de precisão têm sido desenvolvidos a partir da evolução de tecnologias de hardware e software, geoprocessamento, sistemas de posicionamento global e sensores, que proporcionam a determinação de uma série de elemento envolvidos na produção,

em uma forma nova, rápida e com custos acessíveis de visualizar uma propriedade agrícola detalhando características específicas dessa área (LI *et al.*, 2002).

Para praticar adequadamente a AP, é necessário estar ciente de que a variabilidade espacial das lavouras é algo mais complexo do que apenas mapear. Existe a necessidade de obter o conhecimento das características de manejo de uma propriedade e de seus dados de produtividade, fatores químicos e físicos do solo. Segundo Silva (2003) deve-se considerar que a variabilidade espacial das características de solo e planta precisa ser incluída aos procedimentos e técnicas aplicados na agricultura para aperfeiçoar a aplicação de insumos por meio de interpretações obtidas de várias fontes de dados, tais como análises laboratoriais, dados empíricos, recomendações agrônômicas, dados de relevo e clima, entre outros. Esse conhecimento é muito importante para um manejo eficiente, pois o solo, por mais regular que seja, apresenta variações nos seus atributos mesmo em áreas homogêneas e essas variações podem ter valores suficientes para interferir na produtividade das culturas.

Na área de solos, quando realizados estudos da variabilidade espacial e temporal, a adoção de ferramentas tecnológicas associada a esse conhecimento torna-se positiva. Para gerar as informações sobre a variabilidade espacial e temporal de uma propriedade agrícola, é necessário conhecimento sobre os mecanismos disponíveis para efetuar esse processo, tais como a geoestatística que leva em consideração o comportamento espacial das variáveis, considerando que os valores de uma variável estão de alguma forma relacionada à sua distribuição espacial. Entre as metodologias contidas na estatística espacial, a geoestatística destaca-se por utilizar toda a informação disponível, inclusive à posição geográfica dos dados amostrais, caracterizando assim a variável em estudo. O principal objetivo da geoestatística é modelar as variações espaciais contínuas, assumindo uma estrutura de correlação espacial da variável analisada (DIGGLE e RIBEIRO, 2000). Técnicas de geoestatísticas para análise da distribuição espacial de dados têm sido utilizadas para estudar a variabilidade espacial dos atributos do solo, levando a um melhor entendimento dos processos de interação entre solo, planta e atmosfera (WENDROTH *et al.*, 2001).

Softwares de análise espacial de dados buscam armazenar e analisar dados de atributos coletados em diferentes períodos de tempo ou safras, gerando assim informações úteis para tomada de decisão, apesar de que, a maior parte

destes softwares não foi desenvolvida para se trabalhar com o segmento agrícola e são adaptados para serem utilizados na agricultura. Além disso, a maior parte das ferramentas utilizadas não é gratuita, tendo-se a necessidade de realizar a aquisição de licença de uso, que acaba em muitos casos sendo um limitador, devido à necessidade de aporte financeiro de produtores menores. Com a evolução das ferramentas livres tais como o software estatístico R que possuem potencial comprovado para realizar esses procedimentos, sua utilização torna-se vantajosa e isso é comprovado por empresas de desenvolvimento de software comerciais que realizaram a integração com o software estatístico R para incrementar seus produtos.

O software estatístico R é um dos programas utilizados para análises estatísticas. Sua principal vantagem é ser um programa livre, ou seja, possui livre distribuição e código aberto, além de possuir como fator importante que seus pacotes e bibliotecas são escritos e desenvolvidos por pesquisadores das mais diversas áreas do conhecimento. Dentre os pacotes destacam-se os desenvolvidos para análises geoestatísticas, o geoR e o gstat. No entanto o aprendizado para se utilizar o software estatístico R não é fácil e alguns esforços serão necessários para que se consiga tirar algum benefício (SILVA, 2012).

A criação de uma interface para facilitar a utilização dos recursos de uma ferramenta como o software estatístico R buscando alto nível de usabilidade e principalmente ser funcional pode tornar os processos realizados mais simples. A necessidade da criação de uma interface amigável ao usuário é fundamental para utilização e comunicação entre o homem e computador, no qual são realizadas as operações visando atingir um objetivo comum. A interface é parte do sistema computacional que determina como os usuários operam e controlam um sistema e quando uma interface é bem projetada, ela se torna acessível, agradável e controlável, tornando as ações dos usuários mais seguras.

Neste contexto, o intuito desse projeto foi desenvolver uma ferramenta de análise espacial de dados, em plataforma web, que faz parte de um projeto maior que contempla outros mecanismos de análise e coleta de dados de onde serão utilizados os dados armazenados nesta estrutura.

1.1 OBJETIVO GERAL

Desenvolver um software em plataforma web, que contenha uma interface amigável e funcional para geração de mapas temáticos e que permita executar funcionalidades de análise espacial de dados e métodos de interpolação, disponíveis no software estatístico R, de forma transparente para o usuário.

1.2 OBJETIVOS ESPECÍFICOS

Os objetivos específicos definidos para este trabalho foram:

- a) Integrar a ferramenta web desenvolvida com recursos disponibilizados pelo software estatístico R destinados à análise espacial de dados e à geração de mapas temáticos.
- b) Construir uma interface amigável e funcional, para que o usuário possa realizar a geração de mapas temáticos fazendo uso dos interpoladores Krigagem ordinária ou Inverso da distancia elevado a uma potência, de forma fácil e objetiva.
- c) Construir procedimentos automatizados para criação de semivariogramas, mecanismos de ajustes e seleção de melhores modelos de forma automática, fazendo uso de estatísticas da validação cruzada.
- d) Efetuar teste de geração e apresentação de mapas temáticos, utilizando a ferramenta desenvolvida, fazendo uso de dados reais já obtidos em propriedades agrícolas localizadas na região oeste do Paraná, e realizar a comparação com software comercial a fim de garantir a confiabilidade da ferramenta.

2 REVISÃO DE LITERATURA

2.1 AGRICULTURA DE PRECISÃO

Segundo Stafford (2000) a variabilidade temporal e espacial dos fatores do solo e da cultura dentro de uma propriedade rural tem sido acompanhada por séculos. Antes mesmo da etapa da mecanização agrícola, o tamanho pequeno das propriedades permitiu que os agricultores diversificassem seus manejos anualmente. Entretanto, com o crescimento das áreas e a mecanização intensiva, tornou-se cada vez mais difícil ter um conhecimento da variabilidade pertencente à propriedade sem que houvesse um desenvolvimento revolucionário nas tecnologias utilizadas.

A AP é um conceito relativamente novo de gestão agrícola, introduzido em meados da década de 1980, que além de atrair um interesse considerável, iniciou uma revolução na gestão de recursos agrícolas (ROBERT, 1999). Desde então, tem sido atrelado a este conjunto de técnicas e procedimentos que visam gerenciar a variabilidade espacial da lavoura, vários termos para nomear esta forma de gerenciar as áreas produtoras, tais como: "gerenciamento de precisão", "cultivo de precisão" ou "agricultura de precisão". Apesar disso, é consenso que se trata de uma forma de gerenciar orientada a dados e com uso de tecnologia intensiva (MURAKAMI *et al.*, 2007).

Apesar dos benefícios do uso da AP, esta vem sendo adotada com cautela devido a vários fatores relacionados a questões educacionais da natureza ou mesmo pela baixa instrução de produtores que acabam por ter dificuldades em lidar com a nova tecnologia (GRIFFIN *et al.*, 2004). Já para Pedersen *et al.* (2003) um dos grandes problemas tem sido a dificuldade em verificar os ganhos econômicos e ambientais significativos, em especial para produtores de pequeno porte, pela necessidade inicial de alto investimento.

Segundo (SARAIVA *et al.*, 1998; SORENSEN *et al.*, 2011; MURAKAMI, 2006) os requisitos mais importantes que incluem a AP são:

- a) Sistemas de gestão e de apoio à decisão tem o dever de ser projetados para atender às necessidades específicas dos agricultores.

- b) Os sistemas devem ter uma interface de usuário simples que permita a customização para diferentes perfis de usuário. Uma interface de usuário amigável é especialmente importante para aqueles usuários que são inexperientes com software.
- c) Automatizações de métodos simples de usar para processamento de dados são necessários. Os sistemas devem permitir a inclusão e programação de novos métodos automatizados de acordo com regras definidas pelo usuário.
- d) O usuário deve ter permissão e controle completo, sempre que desejar, ter acesso aos parâmetros de processamento e funções de análise. Os usuários mais experientes podem desejar controlar e tentar novas soluções.
- e) A introdução de conhecimento especializado deve ser possível. Isto pode oferecer a oportunidade de afinar os sistemas às condições locais, e incluem conhecimentos, práticas e preferências do usuário.
- f) Serão necessários sistemas de computador mais integrados e mais padronizados. Isso pode reduzir o investimento técnico, a curva de aprendizagem, e a necessidade de apoio técnico.
- g) Apoio à integração fácil, sem custo, incluindo interoperabilidade com outros pacotes de software, outras fontes de dados (tais como dados meteorológicos, dados de mercado), local ou remotamente através da Internet, usando padrões abertos de dados, interfaces e protocolos. Isto é especialmente importante para integrar sistemas legados e sistemas distribuídos.
- h) Escalabilidade, para atender diferentes necessidades.
- i) Suporte para meta-dados para permitir a troca de dados entre aplicações.
- j) Baixo custo.

De acordo com Koch e Khosla (2003) as principais práticas ligadas à AP são o mapeamento da produtividade e a amostragem georreferenciada de parâmetros do solo e planta como ferramentas de diagnóstico, tratamento localizado e ações de manejo, além de outras ferramentas como sensoriamento direto ou remoto (terrestre, aéreo ou orbital) que também é utilizado no levantamento da variabilidade. As ações de manejo podem ser aplicadas por meio de mapas de recomendação que são

gerados previamente a aplicação ou até mesmo em tempo real, a partir das medições realizadas pelos sensores no momento da aplicação (GODWIN; MILLER, 2003).

A própria redução no uso de insumos pode ser vantajosa tanto ambientalmente como economicamente (WANG *et al.*, 2003). Bongiovanni e Lowerberg-Deboer (2001) apresentam a ideia de que aplicar insumos somente onde é necessário seria vantajoso ambientalmente já foi comprovada em diversos estudos que avaliaram a taxa variável de aplicação de nitrogênio, fósforo, herbicidas e inseticidas. Por manter a produção segura ambientalmente e intensiva, a AP é considerada um dos meios para se alcançar a segurança alimentar (GEEBERS e ADAMCHUCK, 2010).

A AP pressupõe a elaboração e a utilização de mapas que descreve as características e as propriedades de uma dada área. Esses mapas representam espacialmente a variabilidade de propriedades do solo, produtividade e características observadas em campo. Com base nas informações contidas nesses mapas, serão tomadas as decisões que envolvem o manejo localizado dessas áreas (SILVA *et al.*, 2003).

Para Guedes Filho (2010) as propriedades agrícolas brasileiras capazes de absorver as tecnologias de AP são de grande porte. Porém, pode também ser praticada nas pequenas propriedades, uma vez que nessas, os campos também apresentam variabilidade espacial e temporal.

É improvável que um único sistema proprietário vai atender a todas estas necessidades, devido a sua complexidade e abrangência. É por isso que uma plataforma de software aberto é uma solução mais adequada para o problema (MURAKAMI, 2006). Agricultura entrou em uma nova era em que a chave do sucesso é o acesso a informações adequadas para a tomada de decisão (MAGNE, 2010).

2.2 SISTEMA DE INFORMAÇÃO GEOGRÁFICA

Um sistema de informação geográfica (SIG) é constituído por um conjunto de ferramentas que possuem como objetivo adquirir, armazenar, recuperar,

transformar e apresentar informações espaciais. Esses dados geográficos descrevem objetos do mundo real em termos de posicionamento, com relação a um sistema de coordenadas, seus atributos não aparentes (como a cor, pH, custo, incidência de pragas, etc.) e das relações topológicas existentes. Portanto, um SIG pode ser utilizado em estudos relativos ao meio ambiente e recursos naturais, na pesquisa da previsão de determinados fenômenos ou no apoio a decisões de planejamento, considerando a concepção de que os dados armazenados representam um modelo do mundo real (CÂMARA e ORTIZ, 1998).

Os SIGs consistem em sistemas de suporte à decisão, integrando dados referenciados espacialmente em um ambiente de respostas a problemas e que realizam funções de análises espaciais baseados nos atributos das entidades armazenadas na base de dados e a partir de módulos específicos, geram um conjunto de dados estimados utilizando-se dos pontos amostrados de coleta da informação, em que as coordenadas geográficas ou locais representam a posição dos pontos (CÂMARA e MEDEIROS, 1996).

Um SIG pode, ainda, ser definido como um sistema provido de quatro grupos de aptidões para manusear dados georreferenciados: entrada, gerenciamento, manipulação e análise, e saída. Os dados são georreferenciados quando estes possuem basicamente duas características: dimensão física e localização espacial (CÂMARA e ORTIZ, 1998).

2.3 SOFTWARE ESTATÍSTICO R

Desenvolvido originalmente por Ross Ihaka e Robert Gentleman no departamento de estatística da Universidade de Auckland, Nova Zelândia, o software R é uma linguagem e ambiente para computação estatística e gráfica, que fornece ampla variedade de técnicas estatísticas e gráficas de modo extensível. De acordo com R Development Core Team (2016) o software R está disponível como software livre sob os termos da GNU - General Public License da Free Software Foundation em forma de código fonte. Ele pode ser compilado e executado em plataformas UNIX e sistemas similares (incluindo FreeBSD e Linux), sistema Windows e MacOS.

Diferentemente da maioria dos softwares estatísticos o R é utilizado fazendo uso de códigos que são executados por meio de linhas de comando.

A integração do R com outras linguagens de programação como C, C++ e FORTRAN é utilizada para melhorar o desempenho de alguns procedimentos. Além de interagir com estas linguagens, o R permite a comunicação com outros softwares como Microsoft Excel, OpenBugs e Statistical Analysis System (SAS). Outras integrações também são realizadas com o objetivo de fornecer novas ferramentas de visualização de dados gerados (SILVA, 2012).

A linguagem de programação R é uma ferramenta importante para o desenvolvimento em ambientes de análises numéricas e aprendizado de máquina. O R ficou mais rápido com o tempo, servindo hoje como uma linguagem para anexar a diferentes conjuntos de dados, ferramentas e pacotes de software. É considerada uma das melhores formas de criar análises de alta qualidade, possuindo flexibilidade e potência para se trabalhar com dados (SABINO, 2014).

Essa plataforma abrange todos os tipos de análises estatísticas para as mais variadas áreas do conhecimento, como economia, biologia, genética, ciências sócias, ciências agrárias. Possuindo como uma das maiores qualidades grande quantidade de pacotes estatísticos disponíveis e que possibilitam manipular dados de forma eficaz e armazená-los. O software estatístico R é extensível e oferece funcionalidades para desenvolvedores criarem suas próprias ferramentas e métodos. O software estatístico R é aderido por profissionais da área acadêmica e por quem busca soluções mais personalizadas (SILVA, 2012).

2.4 ANÁLISE EXPLORATÓRIA DE DADOS

A finalidade da análise exploratória de dados é examinar os dados previamente à aplicação de qualquer técnica estatística. Desta forma o analista consegue um entendimento básico de seus dados e das relações existentes entre as variáveis analisadas. A análise exploratória extrai informações de um conjunto de dados sem levar em consideração o peso das possíveis suposições de um modelo probabilístico. As técnicas de representação gráfica desempenham um importante papel neste tipo de abordagem (LAURETTO, 2011).

Grego (2014) define que antes dos dados serem submetidos à análise geoestatística, é de grande importância e tem como objetivo a identificação prévia dos dados discrepantes, além da normalidade da distribuição e a variação dos dados. Pois muitas vezes os valores discrepantes ocorrem por mau contato do sensor utilizado na medição, erro de amostragem, análise laboratorial ou simplesmente erro de digitação.

De acordo com Ribeiro Junior (1995), a análise exploratória possui as etapas de análise descritiva geral e espacial. A primeira corresponde à descrição geral dos dados sem considerar a posição das amostras, enquanto a segunda refere-se à posição das amostras georeferenciadas. Algumas etapas elencadas como importantes referentes à análise exploratória e que devem ser cumpridas são a construção de diagramas de ramas e folhas e os gráficos de box-plot, sendo instrumentos importantes para se detectar no conjunto de dados os pontos discrepantes.

A análise exploratória dos dados é importante para verificar se os dados seguem determinada normalidade, se os dados analisados não possuem pontos discrepantes que podem interferir no resultado de análises. Neste sentido, são utilizados os parâmetros de tendência central e medidas de dispersão, com o objetivo de verificar o padrão de distribuição dos dados estudados. Segundo Andriotti (2003) se utiliza os parâmetros de coeficiente de variação, assimetria e curtose para se definir a necessidade de transformação logarítmica para um melhor ajuste nas informações.

A maneira comum de representação gráfica de dados amostrais é o histograma, que consiste em um grupo de retângulos que tem por base o eixo das abscissas e por ponto médio o valor central do intervalo de classe. A construção de histogramas é uma das representações preliminares de dados estatísticos em que se pode observar a distribuição dos dados e verificar se o conjunto de dados apresenta normalidade. Os histogramas assim como os mapas de distribuição são ferramentas para descrição espacial dos dados através da sua visualização.

2.5 GEOSTATÍSTICA

A geoestatística surgiu na África do Sul, quando Daniel Krige em 1951 trabalhando com dados de concentração de ouro, verificou que para encontrar sentido nas variâncias em estudo era preciso levar em conta as distâncias entre as amostras, e assim surgiu o conceito da teoria das variáveis regionalizadas, que leva em consideração a localização geográfica e a dependência espacial (GREGO, 2014). Já Matheron em 1965 colocou essa ideia em termos matemáticos e desenvolveu a geoestatística, que corresponde a um ramo da estatística que utiliza conceito de variáveis regionalizadas na avaliação de variabilidade espacial. Além disso, esta análise não se limita apenas em obter um modelo de dependência espacial, mas permite que a partir deste modelo possam-se estimar valores em locais onde não foram realizadas coletas amostrais (SRIVASTAVA, 1996).

Matheron (1971) diz que uma variável regionalizada é uma função $f(x)$ do ponto x , mas também é uma função irregular na qual se têm dois aspectos contraditórios ou complementares: um aspecto aleatório, cuja irregularidade não permite prever as variações de um ponto a outro e um aspecto estruturado, que reflete as características estruturais do fenômeno regionalizado. Para Matheron, essa teoria tem dois objetivos teoricamente, descrever a correção espacial e na prática resolver problemas de estimativa de uma variável regionalizada com base em uma amostra.

A geoestatística permite modelar o valor de um atributo, dentro de uma região, como uma função aleatória (CRESSIE, 1993) e é utilizada para realizar o estudo dos atributos espacialmente georreferenciados e está fundamentada em conceitos básicos da estatística clássica, levando em consideração as coordenadas geográficas dos pontos amostrais estudados e a dependência espacial que esses elementos contêm entre si (ASSUMPÇÃO *et al.*, 2007).

A geoestatística tem como objetivo entender o comportamento dos dados espaciais relacionados à variabilidade pertencente ao fenômeno em estudo (CHILÈS e DELFNER, 1999). De acordo com Kavanagh (2001) a geoestatística se configura como sendo uma estatística que leva em consideração não apenas o valor coletado para uma variável, mas também sua posição expressa por um sistema de coordenadas. O valor tomado por uma variável localizada espacialmente por suas

coordenadas pode ser predito por valores em posições vizinhas que pareçam não ser estocasticamente independentes entre si.

Segundo Matheron (1963), o ponto de partida para o desenvolvimento da geoestatística foi devido á inabilidade da estatística clássica em considerar o aspecto espacial de um fenômeno, que constitui uma característica importante num estudo geológico. O valor obtido por uma variável localizada espacialmente por suas coordenadas pode ser predito por valores em posições vizinhas que pareçam não ser aleatoriamente independentes entre si (KAVANAGH, 2001).

Para Grossi Sad (1986) a geoestatística dedica-se ao estudo da distribuição e da variabilidade dos valores, em função de sua adequação espacial ou temporal, ou seja, de valores regionalizados que se modificam no espaço e que são providos de certa estrutura.

No estudo da variabilidade espacial de determinado atributo é fundamental conhecer o valor dele em toda a área sob estudo. Diante da inviabilidade de coletar valores em toda a área, a partir de dados pontuais obtidos por amostragem, é realizada a predição do valor do atributo em pontos não amostrados utilizando métodos de interpolação. Na interpolação por krigagem os pesos são determinados a partir de uma análise espacial, baseada no semivariograma experimental (CRESSIE, 1993).

Após a construção do semivariograma experimental, faz-se necessário ajustar valores por uma curva atenuada média, função de alguns parâmetros que quantificam a continuidade espacial e devem ser estimados (SOARES, 2000). Entre os modelos teóricos mais utilizados para definir a estrutura de variabilidade espacial, pode-se citar o esférico, exponencial, gaussiano e família Matérn. Além de ajustar um modelo teórico ao semivariograma experimental é importante avaliar a qualidade do ajuste.

Segundo Isaaks e Srivastava (1989), o critério de validação cruzada é uma técnica de avaliação de erros de estimativas que permite comparar os valores estimados com os amostrados que além de indicar a qualidade do ajuste, indicam também à avaliação de estacionariedade e do peso da presença de dados atípicos (BASTIANE, 2012).

2.5.1 Semivariograma

O semivariograma é a ferramenta da geoestatística que permite modelar e verificar o grau de dependência espacial entre amostras. As amostras georreferenciadas são coletadas em diversos pontos que podem estar regularmente distribuídos ou não. No caso de haver dependência espacial, o valor de cada ponto deve apresentar uma relação com valores obtidos por seus vizinhos e a influência entre eles será tanto quanto maior ou menor forem às distâncias que os separam (SANTOS *et al.*, 2011).

Para Vieira (1998), o semivariograma é um “medidor” do grau de semelhança entre vizinhos, mostra a medida do grau de dependência espacial entre amostras ao longo de um suporte específico, contudo para sua construção são utilizados valores das variáveis obtidas no campo em diferentes distâncias. Para que a variável avaliada tenha dependência espacial é necessário que o semivariograma seja crescente com a distância. Além disso, é exigido que a hipótese intrínseca seja válida, ou seja, que o semivariograma após ter um aumento com a distância, se estabilize no valor da variância dos dados. A maneira como o semivariograma cresce com a distância até o alcance é que define o comportamento espacial da variável estudada.

O estimador clássico do semivariograma proposto por Matheron (1963) é representado na equação (1).

$$\hat{\gamma}(h) = \frac{1}{2N_{(h)}} \sum_{i=1}^{N(h)} [Z(x_i) - Z(x_i + h)]^2 \quad (1)$$

Em que

$\hat{\gamma}(h)$ é o valor da estimativa da semivariância;

$Z(x_i)$ o valor da variável no ponto x_i ;

$Z(x_i) + h$ é o valor da variável no ponto $x_i + h$;

$N(h)$ é o número de pares separados por uma determinada distância h .

De acordo com Druck *et al.* (2004), o semivariograma representado na Figura 1 é um gráfico no qual a semivariância dos dados cresce em função da

distância(h) até determinado ponto onde não se observa mais a continuidade espacial.

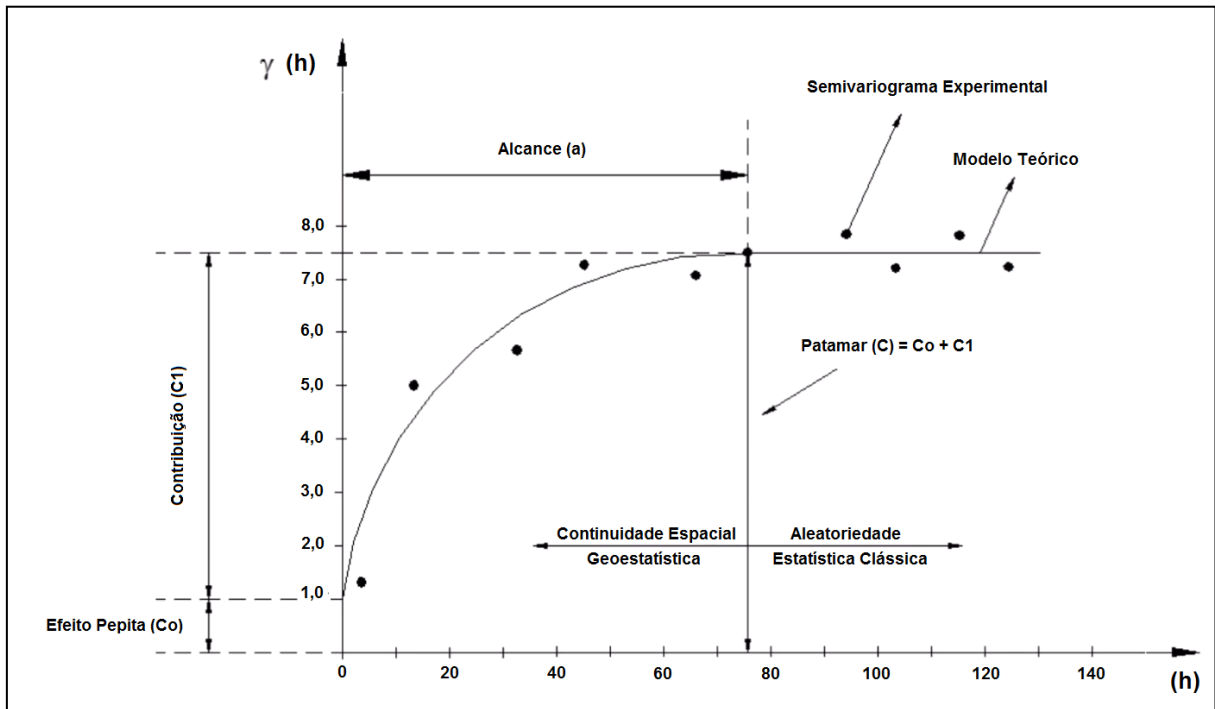


Figura 1- Estrutura Semivariograma.

Fonte: Adaptado de Druck *et al.* (2004).

O ponto no eixo $x(h)$ é denominado alcance (a) do semivariograma, é a distância h , dentro do qual as amostras apresentam-se correlacionadas espacialmente. A partir deste valor o gráfico se estabiliza, ou seja, torna-se aproximadamente constante. No eixo $y(h)$ é denominado patamar ($C_0 + C_1$) e, deste ponto em diante, considera-se que não existe mais dependência espacial entre as amostras, porque a variância da diferença entre pares de amostras torna-se aproximadamente constante. O efeito pepita (C_0) é uma descontinuidade na origem, representa as variações locais de pequena escala, tais como erros de amostragens e de medidas (DRUCK *et al.* 2004) e a contribuição (C_1), também conhecida como a variância da dispersão, representa as diferenças espaciais entre os valores de C_0 e o patamar (CRESSIE, 1993).

Gráfico do semivariograma é uma série de pontos discretos para a qual, uma função contínua deve ser ajustada. Ajuste de modelo ao semivariograma é um dos

aspectos mais importantes das aplicações da geoestatística porque os cálculos de geoestatística dependem do valor do modelo do semivariograma para cada distância especificada (GREGO, 2014).

2.5.2 Modelos Teóricos e métodos de ajuste

Existem diversos modelos na literatura aplicáveis a diferentes fenômenos com continuidade espacial. Estes modelos são chamados de modelos teóricos, tais como esférico, exponencial, gaussiano, Família Matérn, entre outros, que são utilizados para o ajuste de um modelo teórico ao semivariograma experimental (DIGGLE e RIBEIRO, 2000). McBratney e Webster (1986) indicam que os modelos mais adequados para os mais variados tipos de situações, na maioria dos casos, são os modelos esférico, exponencial e gaussiano, que estão representados na Figura 2.

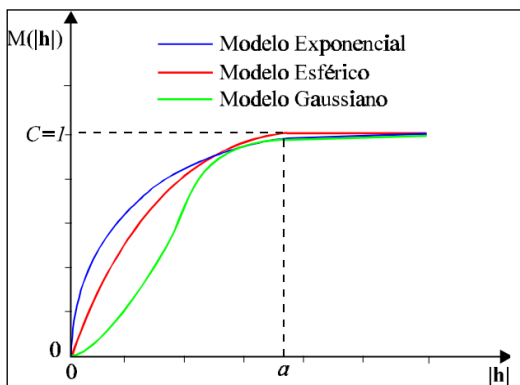


Figura 2 - Modelos teóricos.

Fonte: Adaptador de Isaaks e Srivastava (1989)

O modelo esférico segundo Andriotti (2003) é o modelo mais comumente utilizado nas variáveis em geociências. Apresenta um crescimento rápido na origem, onde tem um comportamento linear (Equação 2). Está representado graficamente na Figura 3.

$$\gamma(h) = \begin{cases} 0, & \text{se } h = 0 \\ C_0 + C_1 \left[\frac{3}{2} \left(\frac{h}{a} \right) - \frac{1}{2} \left(\frac{h}{a} \right)^3 \right], & \text{se } 0 < h \leq a \\ C_0 + C_1, & \text{se } h > a \end{cases} \quad (2)$$

Em que

$\gamma(h)$ é a semivariância para um determinado h ;

C_0 é o efeito pepita; C_1 é a contribuição; a é o alcance;

h são os valores das distâncias.

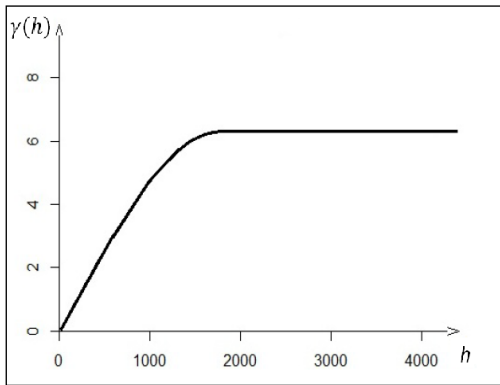


Figura 3 - Representação gráfica do modelo esférico.
Fonte: Adaptador de Andriotti (2003)

A família dos modelos de Matérn é uma generalização de outros modelos teóricos e tem como característica fundamental a inclusão de um parâmetro (k) que determina a suavização (MINASNY; MCBRATNEY, 2005) é definida na Equação 3.

$$\gamma(h) = \begin{cases} 0, & \text{se } h = 0 \\ \varphi_1 + \varphi_2 \left[1 - \left(2^{k-1} \tau(k) \right)^{-1} \left(\frac{h}{\varphi_3} \right)^k K_k \left(\frac{h}{\varphi_3} \right) \right], & \text{se } h > a \end{cases} \quad (3)$$

A Família Matérn é válida para φ_3 , $k > 0$ e corresponde à função de correlação exponencial quando $k = 0,5$; e a função de correção gaussiana quando $k \rightarrow \infty$. A Figura 4 apresenta o modelo Família Matérn para $k = 0,1$; $k = 0,2$; $k = 0,5$; $k = 1$; $k = 2$.

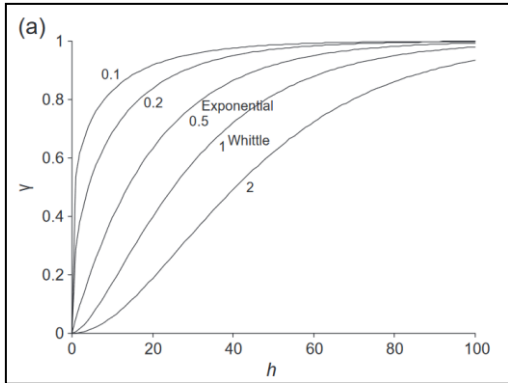


Figura 4 - Representação gráfica dos modelos Família Matérn.
Fonte: Adaptado de Minasny e McBraatney (2005)

Modelo exponencial este modelo apresenta comportamento linear na origem. O alcance tem significado puramente analítico, sendo o patamar alcançado pela curva de forma assintótica, ou seja, teoricamente, quando h tende ao infinito (ANDRIOTTI, 2003), representado na Equação 4.

$$\gamma(h) = \begin{cases} 0, & \text{se } h = 0 \\ \gamma(h) = C_0 + C_1[1 - e^{-3(h/a)}], & \text{se } h \neq 0 \end{cases} \quad (4)$$

Em que

$\gamma(h)$ é a semivariância para um determinado h ;

C_0 é o efeito pepita;

C_1 é a contribuição;

a é o alcance;

h são os valores das distâncias.

Apesar de o modelo exponencial crescer junto da origem mais rápido que o modelo esférico, as estruturas apresentam uma continuidade espacial maior. Sendo assim, maiores prolongamentos, resultantes das maiores correlações espaciais a grandes distâncias de h (SOARES, 2000), a Figura 5 representa sua estrutura.

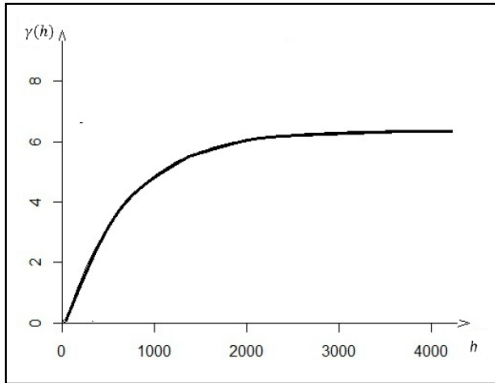


Figura 5 - Representação gráfica modelo exponencial.
Fonte: Adaptador de Andriotti (2003)

Já o modelo gaussiano tem comportamento parabólico nas vizinhanças da origem e reflete grande continuidade da variável estudada, devido ao seu leve crescimento. Trata-se de um modelo transitivo, muitas vezes utilizado para modelar fenômenos extremamente contínuos (ISAAKS e SRIVASTAVA, 1989), sendo representado pela Equação 5 e Figura 6.

$$\gamma(h) = \begin{cases} 0, & \text{se } h = 0 \\ C_0 + C_1[1 - e^{-3(h/a)^2}], & \text{se } h \neq 0 \end{cases} \quad (5)$$

Em que

$\gamma(h)$ é a semivariância para um determinado h ;

C_0 é o efeito pepita; C_1 é a contribuição;

a é o alcance;

h são os valores das distâncias.

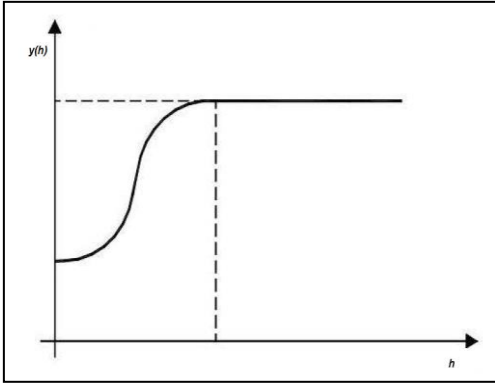


Figura 6 - Representação gráfica modelo gaussiano.
Fonte: Adaptado de Minasny e McBratney (2005)

Por meio do semivariograma experimental, é possível definir o modelo que melhor descreve o comportamento dos dados no espaço (JOURNEL e HUIJBREGTS, 1978). Em seguida, cria-se a preocupação para o ajuste da função matemática ao semivariograma experimental. De acordo com Mc Bratney e Webster (1986), em vários trabalhos na área de ciência do solo, a função matemática e o procedimento de ajuste dos parâmetros, se mencionados, foram simplesmente declarados de forma subjetiva, em que, em praticamente 100% dos casos, o semivariograma foi estimado de forma visual.

Até a década de 80, o ajuste dos modelos espaciais era comumente feito de forma visual também chamado de "sentimento", sem nenhum procedimento matemático. Posteriormente, com a capacidade de recursos computacionais, outros métodos de ajuste, foram e estão sendo estudados (GUEDES, 2012). Dentre os métodos de ajuste do modelo espacial ao semivariograma experimental, destacam-se os Métodos: Quadrados Mínimos Ordinários (*Ordinary Least Squares* – OLS) e Quadrados Mínimos Ponderados (*Weight Least Squares* – WLS). Esses métodos foram desenvolvidos com o intuito de retirar ou atenuar o caráter de subjetividade na estimativa dos parâmetros do semivariograma de forma visual (CRESSIE, 1993).

O método de quadrados mínimos ordinários (OLS) é um dos estimadores mais utilizados na estatística e baseia-se no princípio dos quadrados mínimos, introduzido por Gauss em 1809 (BUSSAB e MORETIN, 2002). O método consiste em obter os valores dos parâmetros de um modelo que minimizam a soma do quadrado da diferença entre os valores observados e os estimados (MELLO *et al.*, 2005). De acordo com David (1977) o método dos quadrados mínimos ponderados

(WLS) é uma tentativa de ponderar a semivariância, uma vez que a ponderação pela variância das estimativas das semivariâncias não é conhecida.

Com os resultados obtidos torna-se necessário a comparação por meio dos modelos teóricos para se avaliar qual deles se ajusta melhor com os dados do estudo. Após a definição do modelo que melhor se ajustou, os parâmetros encontrados são utilizados para estimar os pontos não amostrados.

2.5.3 Validação Cruzada e o Índice de seleção de Interpoladores (ISI)

Ao realizar o ajuste de um modelo teórico aos pontos de um semivariograma experimental pode haver dúvida se esse ajuste é o melhor. A validação cruzada é uma técnica de avaliação de erros de estimativa que permite comparar valores, estimados e amostrados, usando-se somente a informação disponível na amostra de dados coletada. Essas comparações são úteis para ajudar a escolher entre os diferentes modelos de estimação. Em várias situações práticas, é necessário verificar os resultados de diferentes modelos e escolher o que mais se aproxima do semivariograma experimental (ISAAKS e SRIVASTAVA, 1989).

Na aplicação da validação cruzada, o modelo é testado nos locais das amostras existentes. O valor da amostra, em determinada localização, é temporariamente descartado do conjunto de dados da amostra, o valor na mesma localização é então, estimado por krigagem, usando-se as amostras restantes. Esse procedimento pode ser visto como um experimento no qual se imita o processo de estimação ao suporte que nunca se obtém uma amostra naquela localização. Com a estimação calculada, pode-se compará-la ao valor da amostra que foi inicialmente removida do conjunto de dados amostrais. Esse procedimento é repetido para todas as amostras disponíveis, esse método é conhecido como método de "deixar um fora" (ISAAKS e SRIVASTAVA, 1989, p.561).

Após conhecidos os valores obtidos da amostragem bem como estes estiverem estimados para cada ponto, a variância total da estimativa é conhecida, permitindo realizar a avaliação de qualidade e o erro de estimação pode ser calculado e conhecido como Erro Médio (GONÇALVES, 2001). Assume como melhor modelo o que tiver o menor valor de erro médio (EM – Equação 6) e do

desvio-padrão do erro médio (DP_{EM} – Equação 7), obtidos por meio do resultado do cálculo da validação cruzada proposta por Isaaks e Srivastava (1989).

$$EM = \frac{1}{n} \sum_{i=1}^n Z(s_i) - \hat{Z}(s_{(i)}) \quad (6)$$

$$DP_{EM} = \sqrt{\frac{1}{n} \sum_{i=1}^n \left(Z(s_i) - \hat{Z}(s_{(i)}) \right)^2} \quad (7)$$

Em que

n é o número de dados;

$Z(s_i)$ é o valor observado no ponto s_i ;

$\hat{Z}(s_{(i)})$ o valor predito por krigagem ordinária no ponto s_i .

Para Montes *et al.* (2005) a ideia da validação cruzada consiste em validar a associação da capacidade do modelo ajustado do semivariograma com a imprecisão do atributo não amostrado.

Os valores de EM e DP_{EM} calculados para cada conjunto de parâmetros são armazenados e utilizados para determinar o índice de seleção de interpoladores (ISI; Bier *et al.*, 2017). É considerado o melhor modelo de ajuste para o semivariograma aquele que apresentar menor valor para o ISI (Equação 8).

$$ISI = \left\{ \frac{abs(EM)}{\max_{i=1}^j [abs(EM)]} + \frac{[DP_{EM} - \min_{i=1}^j (DP_{EM})]}{\max_{i=1}^j [abs(DP_{EM})]} \right\} \quad (8)$$

Em que

EM – erro médio;

DP_{EM} – desvio-padrão dos erros médios;

$abs(EM)$ – valor em módulo do erro médio da validação cruzada;

$\min_{i=1}^j$ – menor valor encontrado entre os j modelos comparados;

$\max_{i=1}^j$ – maior valor encontrado entre os j modelos comparados.

2.6 INTERPOLAÇÃO ESPACIAL DE DADOS

A interpolação é uma técnica utilizada para a estimativa do valor de um atributo em locais não amostrados, a partir de pontos amostrados em uma mesma área ou região. A interpolação espacial permite converter dados de observações pontuais em campos contínuos, produzindo padrões espaciais que podem ser comparados com outras entidades espaciais contínuas. O raciocínio da interpolação que representa seu conceito base é que, na média, os valores de um atributo tendem a ser similares em locais mais próximos do que em locais mais distantes. Esse conceito também fundamenta a base das relações espaciais entre fenômenos geográficos, utilizando a correlação espacial como fonte para diferenciar os atributos estimados (CÂMARA e MEDEIROS, 1998).

A interpolação é um procedimento de estimação de valores de algum atributo em locais não amostrados baseados nas informações obtidas em pontos amostrados. A interpolação espacial converte dados de observações pontuais em campos contínuos e produz padrões espaciais que podem ser comparados com outras entidades espaciais contínuas (BAZZI, 2009).

Os métodos de interpolação mais comuns e de uso em conjuntos de dados relacionados à agricultura de precisão são: vizinho mais próximo, inverso das distâncias elevado a uma potência (IDP) e krigagem ordinária. A diferença entre os métodos de interpolação é a maneira como são definidos os pesos as diferentes amostras (MOLIN, 2008).

2.6.1 Inverso das distâncias elevado a uma potência (IDP)

O inverso das distâncias elevado a uma potência implementa a suposição de que as coisas mais próximas entre si são mais parecidas do que as mais distantes. Para prever um valor para um determinado local não medido, o IDP como é conhecido usa os valores amostrados a sua volta, que contém um maior peso do que os valores mais distantes, ou seja, cada ponto possuiu uma influência no novo

ponto, que diminui na medida em que a distância aumenta (JAKOB e YOUNG, 2006).

O método inverso das distâncias elevado a uma potência possui dois casos mais conhecidos: o inverso da distância (ID) e o inverso do quadrado da distância (IQD). A diferença entre eles é o expoente a que o cálculo é submetido. A potência é utilizada para atenuar a influência dos pontos distantes. É um processo puramente matemático, no qual os dados são ponderados de tal forma que o efeito entre eles diminui conforme a distância aumenta (MAZZINI e SCHETTINI, 2009).

Neste método de interpolação os pesos são definidos como o inverso da distância elevado a uma potência (Equação 9), em que a distância é a que separa o valor interpolado dos valores observados. Este método faz com que os pesos dos dados sejam avaliados durante o processo de interpolação e a influência de cada ponto amostrado é inversamente proporcional à distância do ponto a ser estimado (MAZZINI e SCHETTINI, 2009).

$$\hat{Z}_i = \frac{\sum_{i=1}^n \left(\frac{1}{d_i^p} * Z_i \right)}{\sum_{i=1}^n \left(\frac{1}{d_i^p} \right)} \quad (9)$$

Em que

\hat{Z}_i - valor interpolado;

Z_i - valor do atributo amostrado;

d_i^p – distância euclidiana entre o i-ésimo ponto de vizinhança e o ponto amostrado elevado a uma potência.

Esse método tem algumas variações, a diferença é apenas o expoente ao qual o cálculo é submetido. Ambos os métodos, inverso da distância ou inverso do quadrado da distância são métodos considerados de fácil aplicação, entretanto são menos precisos que a krigagem ordinária, pelo motivo de que não consideram o padrão da estrutura da dependência espacial (SILVA *et al.* 2008).

2.6.2 Krigagem ordinária

Krigagem ordinária é considerada um método de interpolação mais complexo, pois faz uso da análise geoestatística para efetuar a interpolação, o que em muitos casos é uma grande vantagem sobre outros métodos. A krigagem define o grau de dependência ou correlação espacial entre as amostras por meio dos parâmetros ajustados no semivariograma obtidos na etapa de análise geoestatística (CRESSIE, 1993). Uma vez modelado o semivariograma, é possível verificar o nível de características dos dados, e então definir os melhores pesos para as amostras. De acordo com Druck *et al.* (2004), o método da krigagem tem seus pesos determinados a partir da análise espacial, baseada no semivariograma experimental, sendo necessário constatar a dependência espacial através da análise geoestatística, e então realizar as inferências para os locais não amostrados.

De acordo com Grego e Oliveira (2014) a krigagem realiza a interpolação dos valores com as condições de estimativa sem tendência e com a variância mínima em relação aos valores conhecidos, levando em consideração a estrutura da variabilidade espacial que é encontrada para o atributo. A eficiência da krigagem deve-se ao fato de prover um algoritmo dos erros associados aos resultados obtidos, através de um modelo contínuo de variação espacial. Este método de interpolação é considerado robusto e flexível e procura demonstrar as tendências direcionais que os dados apresentam (CRESSIE, 1993). A krigagem é definida por Krige (1951) Equação 10

$$\hat{Z}(x_0) = \sum_{i=1}^n \lambda_i * Z(x_i) \quad (10)$$

Em que

$\hat{Z}(x_0)$ - é o valor estimado em determinado local;

λ_i - é o peso atribuído aos valores amostrais;

$Z(x_i)$ - é o valor do atributo amostrado;

n - é o número de localidades vizinhas empregadas para a interpolação do ponto, em que o somatório dos pesos λ_i deve ser igual a um.

A krigagem considera além da posição do ponto estimado, os valores e a posição dos pontos amostrados, bem como a variabilidade espacial da variável. Para isso se faz necessário utilizar os parâmetros obtidos no processo referente à análise geoestatística.

2.7 MAPAS TEMÁTICOS

Para se realizar a confecção de um mapa temático, é utilizado pontos amostrais de determinada área, de acordo com a variável regionalizada de interesse, que pode ser: produtividade da cultura, teor de água, potássio, nitrogênio, resistência do solo à penetração, entre outros atributos físicos e químicos do solo e produtividade. Mesmo com esses pontos amostrais não se tem uma definição de como está se comportando toda abrangência da área em estudo, apenas da parte amostrada. Para que se torne possível visualizar o comportamento da variável por toda a área, é necessário que se utilize algum método de interpolação. Abaixo a representação do mapa temático da variabilidade de fósforo em uma propriedade agrícola (Figura 7).

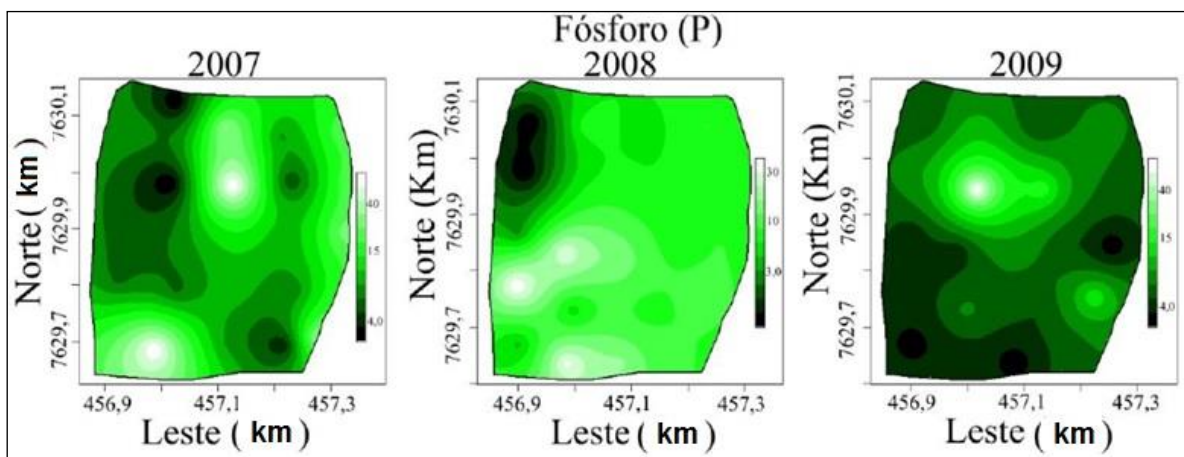


Figura 7 - Mapas temáticos representativo da distribuição do elemento Fósforo em uma propriedade agrícola.

Fonte: Adaptado de Ferraz (2012).

Os mapas temáticos apresentam padrões espaciais de um conjunto de atributos e enfatizam a variação espacial da distribuição geográfica (SPERANZA, 2010), abrangendo em sua elaboração as etapas de coleta de dados, análise, interpretação e representação de informações. Os mapas temáticos são elaborados com a utilização de técnicas que tem como objetivo melhorar a visualização e comunicação, para representarem fenômenos de qualquer natureza, geograficamente distribuídos sobre a superfície terrestre (ARCHELA e THÉRY, 2008).

3 MATERIAIS E MÉTODOS

3.1 FERRAMENTAS E TECNOLOGIAS EMPREGADAS

Para o desenvolvimento desse estudo foi necessário à utilização de ferramentas livres e com recursos que contemplasse os requisitos propostos, sendo assim a linguagem de programação Java, o sistema gerenciador de banco de dados PostgreSQL e o framework de desenvolvimento web Vrapator foram escolhidos, considerando a estrutura inicial que estava sendo desenvolvida e a perspectiva de utilização de ferramentas livres.

3.1.1 Linguagem de Programação Java

A linguagem Java foi projetada para ser utilizada no desenvolvimento de aplicações que utilizem o mínimo de recursos do sistema e que possam ser executadas em diferentes plataformas de hardware e software, tornando-se uma das linguagens preferidas para realizar a implementação de aplicativos baseados na internet e software para dispositivos que se comunicam por rede, além de ser uma linguagem orientada a objetos que torna mais rápidas as atividades de programação e manutenção, maior facilidade para reutilização de código o que cria possibilidade de se construir sistemas muito mais complexos, pela incorporação de funções prontas (DEITEL, 2010).

De acordo com Gonçalves (2007) Java é a base para praticamente todos os tipos de aplicações em rede e é o padrão global para o desenvolvimento e distribuição de ferramentas móveis e incorporadas, jogos, conteúdo baseado na web e softwares corporativos. É uma linguagem gratuita e de código aberto que contém diversos editores e ambientes de desenvolvimento gratuitos (NetBeans, Eclipse, JCreator) e servidores de aplicação (Tom Cat, Jboss).

3.1.2 Banco de Dados PostgreSQL

O PostgreSQL é um sistema gerenciador de banco de dados objeto relacional gratuito e cujo o código fonte é aberto, podendo ser executado em todos os principais sistemas operacionais, incluindo Linux, UNIX (AIX, BSD, HP-UX, SGI IRIX, Mac OS X, Solaris, Tru64) e Windows. Tem interfaces de programação nativas para C / C ++, Java, .Net, Perl, Python, Ruby, Tcl, ODBC, entre outros (POSTGRESQL, 2016).

A escolha do PostgreSQL para realizar o desenvolvimento proposto neste trabalho é por possuir extensão para se trabalhar com dados espaciais (Postgis), contendo muitos recursos para se aplicar na área de banco de dados geográficos, além de possuir uma documentação com informações suficientes para se trabalhar com dados espaciais. É uma ferramenta flexível para manipular altos volumes de dados o que é extremamente necessário quando se está utilizando dados geográficos.

3.1.3 Extensão Espacial Postgis

O mecanismo de extensibilidade do PostgreSQL permite incorporar capacidades adicionais ao sistema de forma a torná-lo mais flexível, como a utilização das funcionalidades espaciais, incorporadas pela extensão PostGIS, que foi desenvolvida pela empresa Canadense Refrations Research Inc (OLIVEIRA, 2007). A fim de servir como uma alternativa a produtos comerciais de custos elevados a comunidade de software livre vêm desenvolvendo a extensão espacial PostGIS. Atualmente, a empresa Refrations Research Inc mantém o time de desenvolvimento dessa extensão que explora vários dos recursos oferecidos pelo PostgreSQL (POSTGIS, 2016).

3.1.4 Framework de desenvolvimento Vraptor4

Para realizar a implementação do módulo com padrões de desenvolvimento, foram utilizados *frameworks* que fornecem conjuntos de procedimentos já implementados que aceleram o processo e garantem a confiabilidade da aplicação.

Para realizar a implementação dos processos do módulo desenvolvido nessa pesquisa de forma ágil e eficiente foi utilizado o *framework* de desenvolvimento Vraptor 4, ele tem como característica e foco principal a produtividade, segue a arquitetura MVC e contém uma curva de aprendizagem curta, além disso, é aberto e baseado nas melhores práticas e tem uma estrutura que possibilita a fácil integração com outras ferramentas e é brasileiro (CAVALCANTI, 2014).

3.1.5 Framework de front-end Bootstrap

Já para a estrutura visual do módulo foi utilizado o *framework de front-end* Bootstrap desenvolvido pela equipe do Twitter, o Bootstrap é um *framework front-end* de código aberto (*open source*) é um conjunto de ferramentas criadas para facilitar o desenvolvimento de sites e sistemas web. Compatível com as linguagens de programação HTML5 e CSS3, que além de ser gratuito fornece agilidade para construção de templates, inclui modelos de design baseados em HTML e CSS para tipografia, formulários, botões, tabelas, navegação, bem como plugins Javascript opcionais, a união dessas tecnologias atendem todas as necessidades para o desenvolvimento do módulo dessa pesquisa.

3.2 CONTEXTO EM QUE O PROJETO SE INSERE

Essa pesquisa faz parte de um projeto maior idealizado pelo Prof. Dr. Claudio Leones Bazzi e desenvolvido por alunos da Universidade Tecnológica

Federal do Paraná, campos Medianeira que envolve vários módulos integrados a API AgDataBox é um conjunto de rotinas e padrões de programação fornecidos para ser utilizados por varios serviços por onde os dados são coletados e processados. AgDataBox API contém módulos que contemplam mecanismos de coleta de amostras em campo por meio de sensores, definição de áreas, dados climáticos, controle operacional.

A API utiliza recursos como o GPS (Sistema de posicionamento global) para obter informações como latitude e longitude, recurso de câmera para análise de imagens, interpolação de dados e apresentação de mapas temáticos entre outros módulos em desenvolvimento.

A estrutura da API possibilita as interações de coleta, manipulação e inserção de dados, para que os módulos que fazem parte da aplicação consigam ter acesso a essas informações, para assim possibilitar a utilização desses dados em cálculos, conversões o que for necessário para a geração dos resultados, depois as informações geradas por esses procedimentos serão retornadas para a API e armazenados na base de dados para que assim os outros módulos tenham acesso.

A Figura 8 permite ter uma visão geral de todos os módulos que compõem esse projeto e em destaque a parte do projeto em que está localizada a aplicação desenvolvida nesta pesquisa.

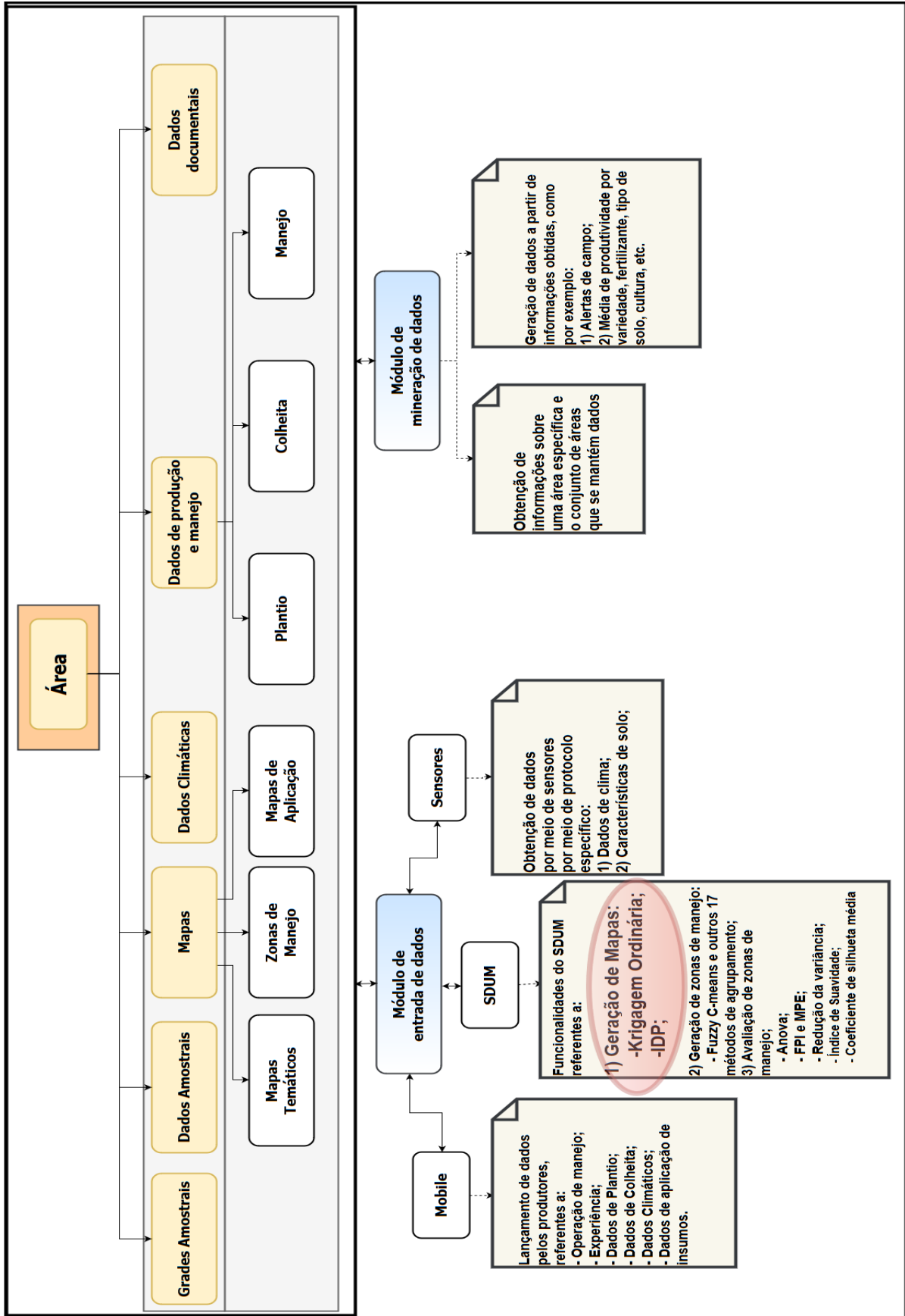


Figura 8 - Representação geral do projeto com destaque as funcionalidades foco deste trabalho.

Fonte: Autoria Própria.

Na Figura 8 pode ser observada a área de atuação da presente pesquisa (em destaque) que envolvem os procedimentos geoestatísticos e interpolação de dados. Dentro do processo de interpolação de dados e geração de mapas temáticos, foco do presente trabalho, a estrutura da pesquisa é definida nas etapas apresentadas de maneira estruturada na Figura 9.

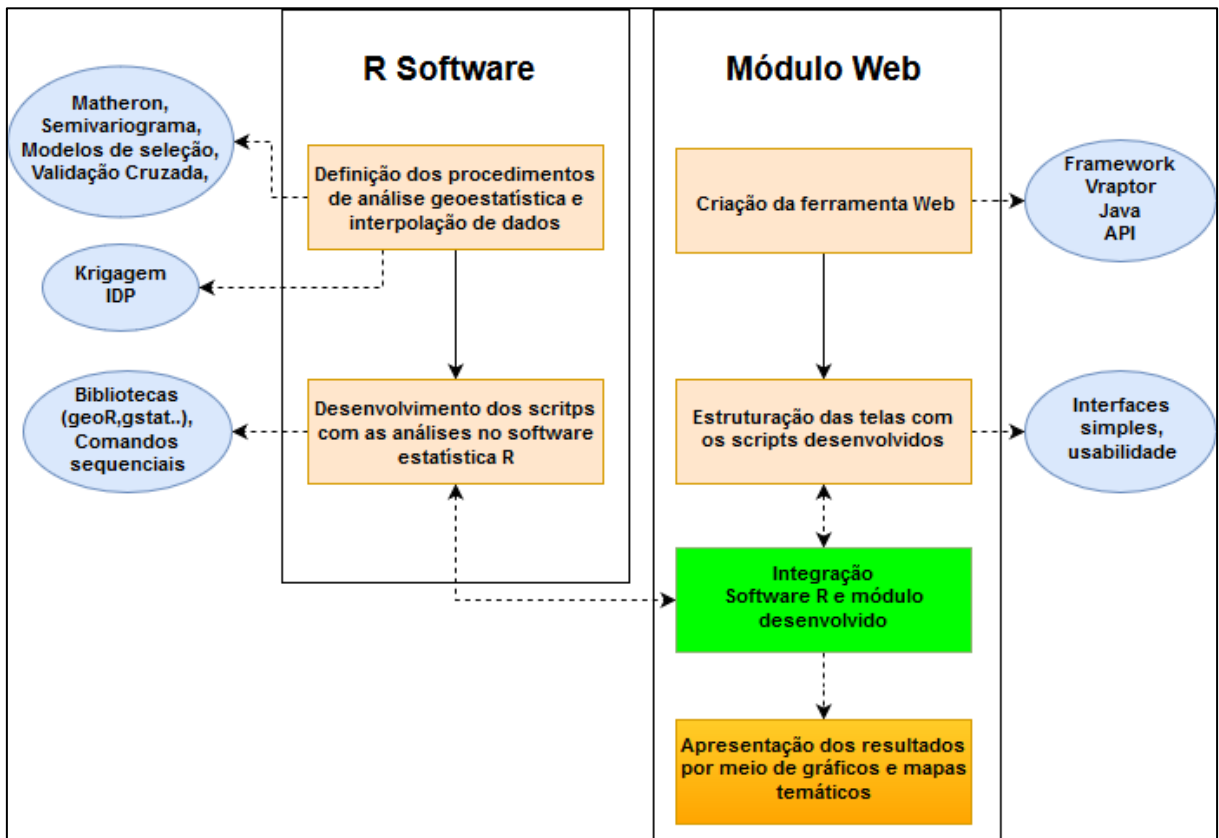


Figura 9 - Diagrama das etapas do projeto.
Fonte: Autoria Própria.

Como representado na Figura 9 o desenvolvimento desta pesquisa foi realizado em duas etapas, a primeira dentro da ferramenta estatística R com colaboração do professor Nelson Miguel Betzek da Universidade Tecnológica Federal do Paraná, que tem seu projeto inserido na área de geoestatística, essas informações foram utilizadas para realizar pesquisa, organização, análise e definição dos procedimentos que envolvem análise estatística, geoestatística e procedimentos de interpolação, além da busca por bibliotecas e pacotes que fornecem funcionalidades para realizar esses processos como, por exemplo, geoR e gStat ,

essas duas bibliotecas possuem funções e procedimentos de análise estatística e geoestatística. Foi necessário também realizar a integração do software estatístico R com a base de dados, no caso a base de dados escolhida foi o PostgreSQL por onde é realizada a coleta das informações para realização das análises propostas, essa funcionalidade de conexão e interação é fornecida pela biblioteca RPostgreSQL.

O procedimento mais complexo foi realizar a adaptação dos scripts e criar a integração com a segunda etapa do projeto, que foi o a criação da ferramenta web por meio da linguagem de programação Java utilizando os *frameworks* de desenvolvimento Vraptor e o Bootstrap que fornece conjunto de componentes para implementação de uma interface completa com alta usabilidade e a apresentação dos resultados de forma simples e objetiva.

3.3 ÁREAS EXPERIMENTAIS E DADOS UTILIZADOS NO ESTUDO

Foram utilizados dados coletados de três áreas agrícolas localizadas no Estado do Paraná, Brasil. Esses dados já foram utilizados em outras pesquisas como Schenatto *et al.* (2016) e Gavioli *et al.* (2016). Essas áreas têm seus solos classificados como LATOSSOLO VERMELHO Distroférico típico e cultivados em sistema de plantio direto.

A área A com 15 ha, localizada no município de Céu Azul, localização geográfica central de 25°06'32" S e 53°49'55" W e altitude média de 460 m; área B: com 9,9 ha, localizada no município de Serranópolis do Iguaçu, localização geográfica central de 25°24'28" S e 54°00'17" W e altitude média de 363 m; Área C também localizada no município de Serranópolis do Iguaçu com centro geográfico sob as coordenadas 25°24'17"S e 54°00'18"O e elevação média de 363 m, possui aproximadamente 20,9 ha onde foram definidos 73 pontos amostrais, essas três áreas estão representadas na Figura 10 como (A, B e C).



Figura 10 - Áreas utilizadas para pesquisa.
Fonte: Autoria Própria.

Foram utilizadas para testes, variáveis comumente estudadas, variando-as em relação às áreas, conforme apresentado na Tabela 1. Realizou-se coleta em 40 (2,67 pontos ha⁻¹), 42 (4,24 pontos ha⁻¹) e 73 (3,7 pontos ha⁻¹) pontos amostrais para as áreas A, B e C, respectivamente.

Tabela 1 - Identificação das variáveis (atributos) e períodos de coleta, para cada área experimental.

Variável (Atributo)	Área A (40 amostras) 2012	Área B (42 pontos) 2012	Área C (73 amostras) 2013
Areia	X		
Argila	X		
Cálcio			X
Cobre		X	X
Fósforo		X	
Manganês			X
Matéria Orgânica	X	X	
Potássio		X	X
Produtividade Soja	X	X	X
Zinco			X

A localização dos pontos amostrais foi obtida a partir de um aparelho receptor GPS da marca Trimble, modelo Juno SB. As amostras de solo foram coletadas à profundidade de 0 - 0,2 m como forma de evitar a camada superficial, geralmente com presença de matéria orgânica e alteração da quantidade de partículas finas por lixiviação. A resistência mecânica do solo à penetração (RSP) foi determinada para as profundidades de 0 - 0,1 m, 0,1 - 0,2 m e 0,2 - 0,3 m, utilizando-se um penetrômetro eletrônico de compactação do solo Falker PenetroLOG PLG1020. Os dados referentes à altitude das três áreas foram obtidos

utilizando-se uma estação total eletrônica de alta precisão Topcon GPT-7505, sendo posteriormente calculados os valores de declividade em função da altitude dos pontos amostrais. Posteriormente, as amostras foram encaminhadas ao laboratório para análise e obtenção de dados de textura (argila, silte e areia), e características químicas do solo.

Os dados de produtividade para a área A foram determinados por meio de um monitor de colheita acoplado a uma colhedora da marca CASE IV. Já para as áreas B e C, a produtividade foi determinada por meio da colheita manual de uma área de amostragem de 1 m² em cada um dos pontos amostrais.

As informações processadas não se limitam apenas as utilizadas nesta pesquisa, podem envolver todos os elementos necessários que proporcionam o conhecimento da área em estudo, características de solo como, por exemplo, dados relacionados às características químicas do solo (C, pH, H+Al, Ca, Mg, K, Cu, Zn, Fe, Mn), macro e micro nutrientes, todas as informações geradas a partir das análises das amostras coletadas em campo, serão armazenados na base de dados, com seus pontos georreferenciados.

3.4 FUNCIONAMENTO DO MÓDULO PROPOSTO

Para descrever os procedimentos realizados para análise dos dados, os tópicos descritos a seguir foram colocados na ordem de processamento, a Figura 12 tem por objetivo representar e esclarecer em forma sequencial e visual os processos realizados pelo módulo.

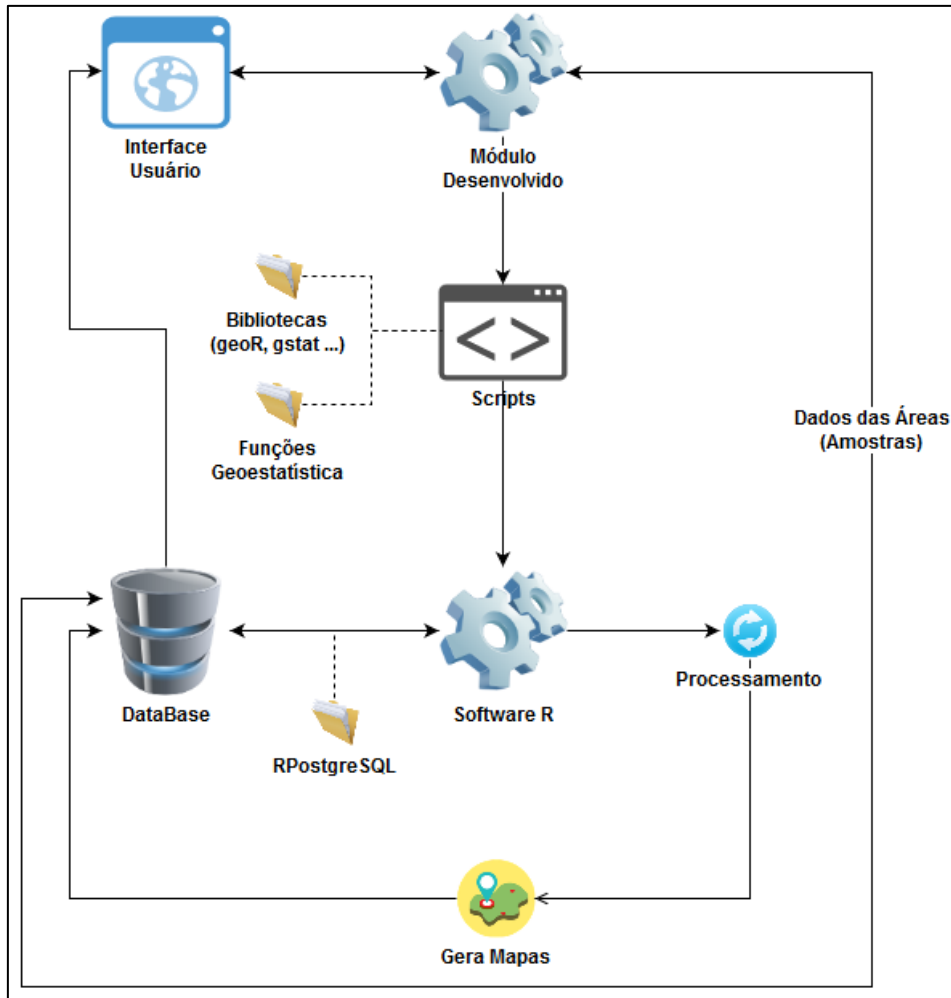


Figura 11 - Representação da estrutura e funcionamento do módulo proposto.
Fonte: Autoria Própria.

O fluxo realizado começa com uma interface disponibilizada pelo módulo web, onde por meio de campos em formato de lista o usuário (agricultor ou pesquisador) faz a seleção da propriedade agrícola e do atributo a ser analisado, esses dados de propriedade e do atributo estão armazenados e são consultados da base de dados pertencente à API AgDataBox.

Por meio da interação e a partir de um botão disponível na interface é acionado o script da ferramenta estatística R, onde os dados fornecidos na interface do usuário são encaminhados e preenchidos dentro dos parâmetros existentes no script desenvolvido no software estatístico R. Feito isso é iniciada a execução, realizando o carregamento das bibliotecas e pacotes dentro do software estatístico R e os processamentos que executam os procedimentos de análise estatística,

geoestatística, interpolação dos dados e geração dos mapas temáticos.

Depois os dados são armazenados no banco de dados e disponibilizados para consulta e visualização. As etapas de análises realizadas pelos scripts desenvolvidos dentro do software estatístico R são detalhadas nas próximas seções.

3.5 ANÁLISE EXPLORATÓRIA E GEOESTATÍSTICA

Foram implementados procedimentos no software estatístico R, capazes de realizar a análise exploratória e geoestatística dos dados e identificar automaticamente o melhor ajuste para o semivariograma experimental. A etapa de geoestatística tem uma interface própria, disponibilizada com parâmetros específicos (área, atributo, tamanho do pixel, número de lags, número de pares entre outros), para usuários com conhecimento mais avançado sobre os procedimentos geoestatísticos está disponível a alteração dos valores desses campos, caso o usuário não tenha conhecimento necessário os campos já possuem valores padrões definidos pela pesquisa, após o preenchimento desses campos é acionado por meio de um botão o envio dos parâmetros para o script R onde é iniciado o processo de análise.

Para iniciar o processo da análise geoestatística primeiramente é feito o carregamento das bibliotecas (geoR, gstat) e outras pacotes como STAT para realizar cálculos estatísticos, SP e SPDEP utilizado para se trabalhar com dados espaciais, o RPostgreSQL que possibilita a criação da conexão da ferramenta R com o gerenciador de banco de dados PostgreSQL. Esse pacote fornece o driver compatível com Database interface (DBI) para que o R consiga acessar as informações da base de dados PostgreSQL (R DEVELOPMENT CORE TEAM, 2016).

Depois de criada a conexão é realizada uma consulta na base de dados utilizando os parâmetros selecionados pelo usuário do módulo web, os parâmetros são a área e a amostra que será utilizada para o processamento. Essa consulta retornará os dados da área referentes à amostra com o posicionamento georreferenciado (latitude, longitude) e o valor dos pontos amostrais.

A próxima etapa tem por finalidade realizar a análise exploratória dos dados a fim de verificar se há normalidade, gerando gráficos para representatividade da variabilidade ou evidenciar qualquer discrepância no conjunto de dados utilizado, evitando assim seguir no processo sem que sejam necessárias análises mais aprofundadas ou de transformação dos dados. Então são calculadas as estatísticas de média, mediana, quartis, coeficientes assimetria e curtose, desvio padrão e coeficiente de variação, sendo apresentados gráficos descritivos (histograma, box-plot, post-plot) representado na Figura 12, que buscam facilitar o entendimento do comportamento de dados que se busca estudar.

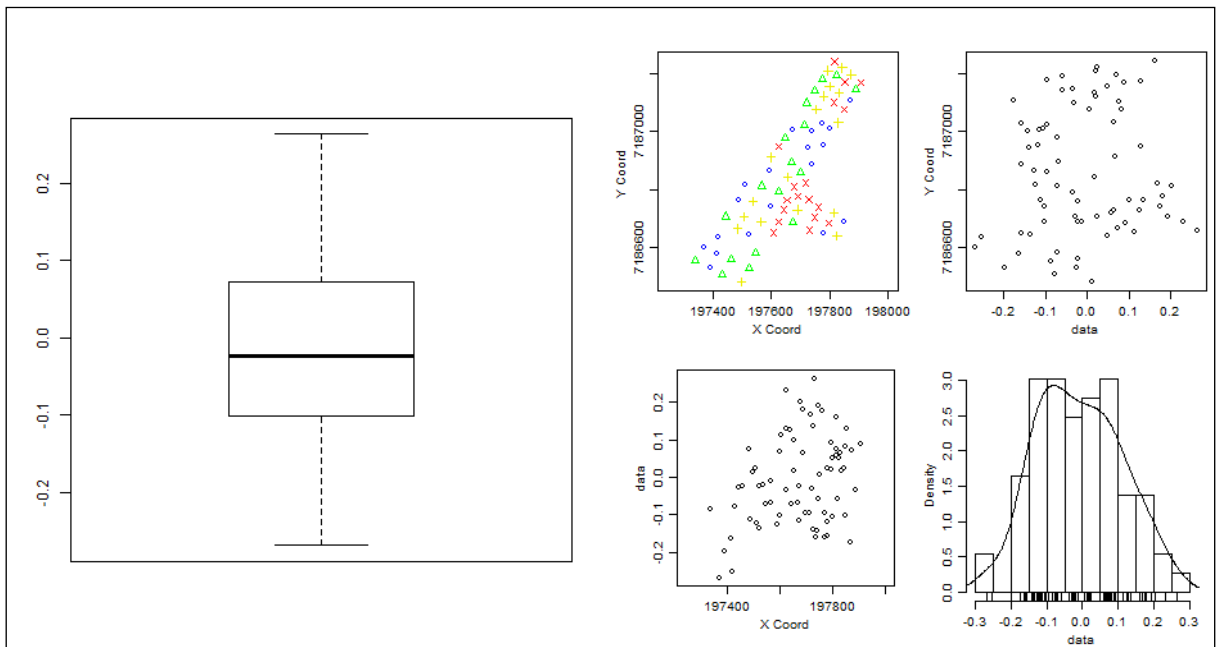


Figura 12 - Gráficos descritivos que apresentam o comportamento dos dados para interpretação do usuário.

Fonte: Autoria Própria.

Essas informações são importantes para que se tenha uma visão do comportamento do atributo em estudo, possibilitando caso haja necessidade remover valores discrepantes que possam influenciar nos resultados da análise geoestatística e interpolação dos dados.

Na etapa de análise espacial dos dados, utilizou-se os recursos disponíveis no pacote geoR (R DEVELOPMENT CORE TEAM, 2016), que fornece funções para

análises geoestatísticas de dados permitindo a criação do semivariograma experimental, em que são ajustados os modelos teóricos e definidos os parâmetros de efeito pepita, contribuição, patamar e alcance utilizados pelo interpolador krigagem ordinária. A qualidade da estimativa de valores para pontos não amostrados que são gerados pela krigagem ordinária depende do ajuste realizado no modelo aplicado ao semivariograma.

Foi disponibilizada uso de seis modelos (esférico, gaussiano, exponencial, Matérn 1.0, Matérn 1.5 e Matérn 2.0) e dois métodos de ajuste do semivariograma (Quadrados Mínimos Ordinários - Ordinary Least Squares – OLS; Quadrados Mínimos Ponderados - Weight Least Squares – WLS) totalizando 12 diferentes formas de análise. Para cada um dos modelos, 25 conjuntos diferentes de parâmetros (cinco valores iniciais para o parâmetro contribuição e cinco para alcance) foram utilizados, totalizando 300 diferentes ajustes analisados para a identificação do melhor modelo para o semivariograma.

A análise para identificar o melhor modelo foi efetuada por meio do desenvolvimento de procedimentos que realizam para cada modelo de ajuste, variação de valores dos parâmetros do semivariograma, calcula-se o erro médio (EM – Equação 6) e o desvio padrão do erro médio (DP_{EM} – Equação 7). E para definir qual o melhor modelo de ajuste para o semivariograma é aplicado o índice de seleção de interpoladores (ISI - Equação 8).

3.6 INTERPOLAÇÃO DE DADOS

A interpolação de dados possibilita a construção de mapas temáticos que representam a variabilidade do atributo em estudo, nesta pesquisa como já abordado anteriormente é utilizado dois interpoladores o inverso da distância elevado a uma potência e a krigagem ordinária. Para isso foi utilizado o estimador clássico de Matheron (1963) para o cálculo das semivariâncias e o alcance foi limitado à metade da distância entre os pontos conhecido como *cutoff* = 50%. Já o ajuste do semivariograma experimental necessário para interpolação por krigagem foi realizado pelo melhor modelo identificado na análise geoestatística e seus

parâmetros. Já para o método de interpolação inverso da distância elevado a uma potência foi realizado a utilização dos procedimentos implementados no software geoR, aplicando sobre as informações armazenadas no banco de dados PostgreSQL.

3.7 MAPAS TEMÁTICOS

Após ser realizada a interpolação dos dados, tanto pelo interpolador inverso da distância elevado a uma potência (IDP) ou a Krigagem ordinária, é criado na base de dados os pontos gerados que compõem a estrutura que representa as características do atributo estudado, com essas informações armazenadas o módulo desenvolvido faz a geração e apresentação do mapa temático referente à área e os atributos processados.

3.8 COMPARAÇÃO DE MAPAS TEMÁTICOS ENTRE FERRAMENTAS

Para determinar a acuracidade dos resultados obtidos pela ferramenta desenvolvida nesta pesquisa, foi utilizado o módulo ArcMap 9.3 (ESRI, 2016), que faz parte do software ArcGis, é uma das ferramentas mais utilizadas no mundo para se trabalhar com dados georreferenciados, o que torna sua utilização para comparações dos resultados muito válida. A ferramenta é utilizada para gerar mapas tanto para interpolação com o inverso da distância elevado a uma potência como para realização da análise geoestatística e interpolação por krigagem ordinária, o que possibilitou a comparação dos mapas gerados pelas duas ferramentas.

4 RESULTADOS E DISCUSSÃO

4.1 INTERFACE E PROCEDIMENTOS DISPONÍVES NO MÓDULO

Seguindo a estrutura organizacional definida para o módulo desenvolvido para facilitar o processo de análise, foi desenvolvida uma interface simplificada que é um dos objetivos do presente projeto, para isso foi utilizado o *framework* de *front-end* Bootstrap, realizada a estruturação de um ambiente simples para seleção dos métodos de análise contendo duas opções, a primeira referente aos procedimentos de análise geoestatística e krigagem ordinária e a segunda opção para a interpolação pelo inverso da distancia elevado a uma potência representado na Figura 13.



Figura 13 - Opções de análises disponíveis no módulo desenvolvido.
Fonte: Autoria Própria.

Além das opções para realizar as análises, também foi disponibilizada uma interface simplificada para visualização das análises já realizadas anteriormente, cada uma das opções de visualização com suas determinadas análises, seguindo o mesmo padrão o que torna sua utilização mais simples.

4.2 CONSUMO DE DADOS PARA REALIZAÇÃO DAS ANÁLISES PELO MÓDULO DESENVOLVIDO

A parte mais importante quando se deseja realizar a análise de dados é possuir uma estrutura com informações bem definidas. O repositório utilizado foi o AgDataBox API que possui as informações das áreas, amostras, atributos, os pontos amostrais de cada atributo e seus valores. Essas informações são carregadas por meio de uma consulta que é realizada pelo módulo desenvolvido diretamente na API, onde o módulo encapsula essa informação retornada pela consulta em uma lista e a envia para a tela de visualização do usuário contendo o conjunto de áreas pertencentes ao usuário, representado na Figura 14.

Interface de usuário para análise de dados. O menu lateral contém opções como "GEOESTATÍSTICA / KRIGAGEM ORDINÁRIA", "GERAR ANÁLISE GEOESTATÍSTICA", "VISUALIZAR ANÁLISES GEOESTATÍSTICAS", "ESTATÍSTICA / INVERSO DA DISTÂNCIA", "GERAR ANÁLISES IDW" e "VISUALIZAR ANÁLISES IDW". O painel principal mostra "Informações" com o usuário "Rodrigo". Há campos para "Área" (menu suspenso com opções: Área 1 Schenatto, Selecione uma Área, Área 1 Schenatto, Tasca_1, Área 2 Schenatto) e "Amostra" (menu suspenso com opção: Selecione uma amostra). Abaixo há uma aba "Configurações".

Figura 14 - Representação da tela com os dados retornados para escolha da área a ser analisada.

Fonte: Autoria Própria.

Depois de realizada a seleção da área que o usuário deseja analisar um procedimento criado no sistema automaticamente vai retornar no campo ao lado, as amostras armazenadas pertencentes à área que foi selecionada, conforme representado na Figura 15.

The screenshot shows a web application interface. On the left is a sidebar with navigation options: 'Início', 'GEOESTATÍSTICA / KRIGAGEM ORDINÁRIA', 'GERAR ANÁLISE GEOESTATÍSTICA', 'VISUALIZAR ANÁLISES GEOESTATÍSTICAS', 'ESTATÍSTICA / INVERSO DA DISTÂNCIA', 'GERAR ANÁLISES IDW', and 'VISUALIZAR ANÁLISES IDW'. The main content area is titled 'Informações' and displays 'Usuário : Rodrigo'. Below this, there are two dropdown menus: 'Área:' with 'Área 1 Schenatto' selected, and 'Amostra:' with a dropdown menu open. The dropdown menu lists several options, with 'Fósforo_area1_2012_certo' highlighted in blue. Below the dropdowns is a 'Descrição:' field. At the bottom, there is a 'Configurações' tab.

Figura 15 - Representação dos dados retornados ao usuário para escolha da amostra a ser analisada.

Fonte: Autoria Própria.

Esse procedimento é comum para as duas formas de análise fornecidas, tanto para análise geoestatística/krigagem ordinária como para a interpolação pelo inverso da distância elevado uma potência, sendo que ambos os métodos necessitam dos dados de área, pontos amostrais e seus respectivos valores. O campo descrição também comum nas duas análises tem como objetivo fornecer uma identificação própria para o usuário referente à análise que está sendo realizada.

4.3 PROCEDIMENTO PARA REALIZAÇÃO DA ANÁLISE GEOESTATÍSTICA E INTERPOLAÇÃO DE DADOS POR KRIGAGEM ORDINÁRIA

A tela para realizar a análise geoestatística contém as informações da área, amostra e descrição conforme abordado anteriormente, juntamente com parâmetros tamanho de pixel que tem influência sobre o tamanho do pixel do mapa a ser gerado, lags automáticos caso esse campo estiver selecionado a opção com valor sim, a rotina computacional implementada será responsável por determinar o número de lags, se o usuário optar em estipular um número determinado de lags deverá selecionar a opção não e realizar o preenchimento do parâmetro número de lags também disponível, outros parâmetros como cutoff, número de pares, além dos dois parâmetros, intervalo de alcance e intervalo de contribuição que tem influência sobre a quantidade de verificações a serem realizados, esses parâmetros estão diretamente ligados ao procedimento de comparação e recomendação do melhor modelo ajustado ao semivariograma.

Os dados referentes a esses parâmetros já estão com valores padrões preenchidos para facilitar as análises de usuários que não possuem um conhecimento aprofundado das características específicas para se realizar análise geoestatística. Depois de verificado por parte do usuário os valores de cada parâmetro é necessário simplesmente clicar sobre o botão "gerar", essa ação dará início ao processo de análise, a interface está representada na Figura 16.

Informações

Usuário : **Rodrigo**

Área: Tasca_1 Amostra: Areia2013

Descrição: AreiaArea3

Configurações

Pixel

Tam. X: 5 Tam. Y: 5

Avançado

ISI	Nº De Lags	Cutoff
Sim	11	50
Transformar dados?	Nº De Pares	Nº Intervalo Alcance
Sim	30	5
Lags Automáticos	Nº Mínimo Sequência Contribuição	Nº Intervalo Contribuição
Sim	0	5
Estimador	Nº Mínimo Sequência Alcance	
Matheron - Classical	0	

Gerar

Figura 16 - Tela para realização da análise geoestatística.
Fonte: Autoria Própria.

Após acionado o botão "gerar", os parâmetros preenchidos na tela são enviados para a integração do módulo desenvolvido com software R (Figura 17), onde os parâmetros enviados pelo usuário preenchem as variáveis dentro do script R, para assim realizar a análise em questão, nesse caso a de geoestatística.

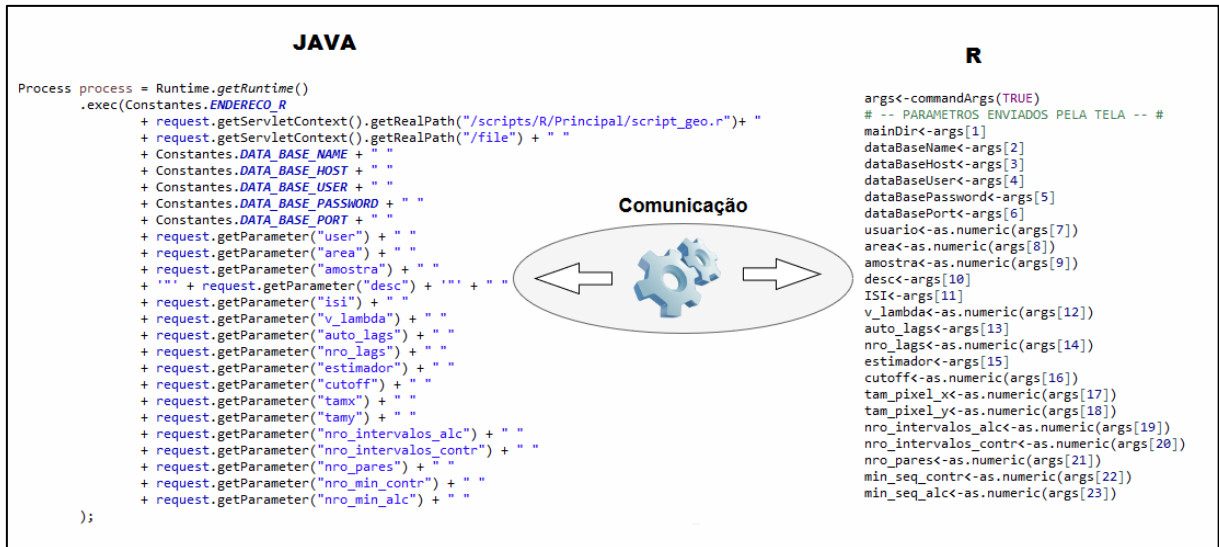
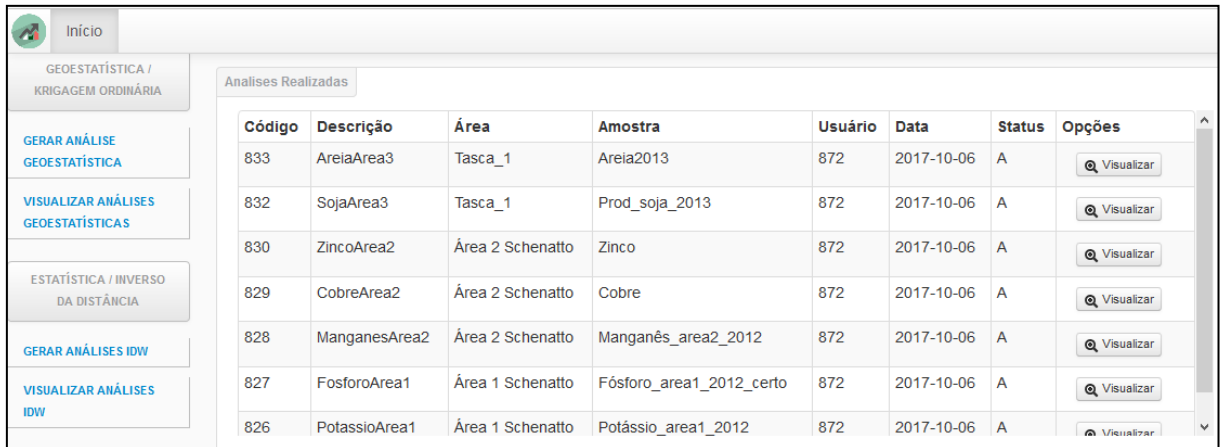


Figura 17 - Código de ligação entre o módulo desenvolvido e o software R.
Fonte: Autoria Própria.

Com o processo iniciado o software R faz as interações necessárias para armazenamento dos dados em banco de dados e gera as apresentações das estatísticas dentro do servidor, quando esse processo termina é gerado um retorno por parte do software R de confirmação para o módulo web dar sequência no processo. O módulo web verifica esse código de confirmação, caso o processo tenha terminado em erro é apresentado uma mensagem para alertar o usuário sobre um possível problema nos parâmetros passados ou com os dados armazenados, caso tenha terminado com sucesso o usuário é redirecionado a tela de visualização das análises, onde estão disponíveis todas as análises que determinado usuário realizou.

Na tela de visualização das análises tanto para análises geoestatísticas como para análise e interpolação pelo inverso da distância elevado a uma potência, as análises aparecem com as seguintes informações: a descrição definida pelo usuário no momento da análise; área e a amostra selecionadas; data de realização. Essas informações facilitam a identificação de determinada análise para futuras consultas, tornando possível um acompanhamento da variabilidade temporal aplicada a uma determinada área e atributo. Ao lado direito da tela possui um botão visualizar é por ele onde são carregadas as informações de cada análise, sendo assim para cada análise há um botão para realizar seu carregamento (Figura 18).



Código	Descrição	Área	Amostra	Usuário	Data	Status	Opções
833	AreiaArea3	Tasca_1	Areia2013	872	2017-10-06	A	Visualizar
832	SojaArea3	Tasca_1	Prod_soja_2013	872	2017-10-06	A	Visualizar
830	ZincoArea2	Área 2 Schenatto	Zinco	872	2017-10-06	A	Visualizar
829	CobreArea2	Área 2 Schenatto	Cobre	872	2017-10-06	A	Visualizar
828	ManganesArea2	Área 2 Schenatto	Manganês_area2_2012	872	2017-10-06	A	Visualizar
827	FosforoArea1	Área 1 Schenatto	Fósforo_area1_2012_certo	872	2017-10-06	A	Visualizar
826	PotassioArea1	Área 1 Schenatto	Potássio_area1_2012	872	2017-10-06	A	Visualizar

Figura 18 - Tela de visualização das análises realizadas.
Fonte: Autoria Própria.

Depois de carregado os dados da análise (nesse caso da análise geoestatística), é exibido à estatística descritiva dos dados da amostra (Figura 19). As representações gráficas geradas são: o histograma que já possibilita uma visualização da distribuição dos dados; o box-plot importante ferramenta para visualização de pontos discrepantes (*outliers*); o gráfico de pontos que representa a distribuição dos dados sobre um eixo.

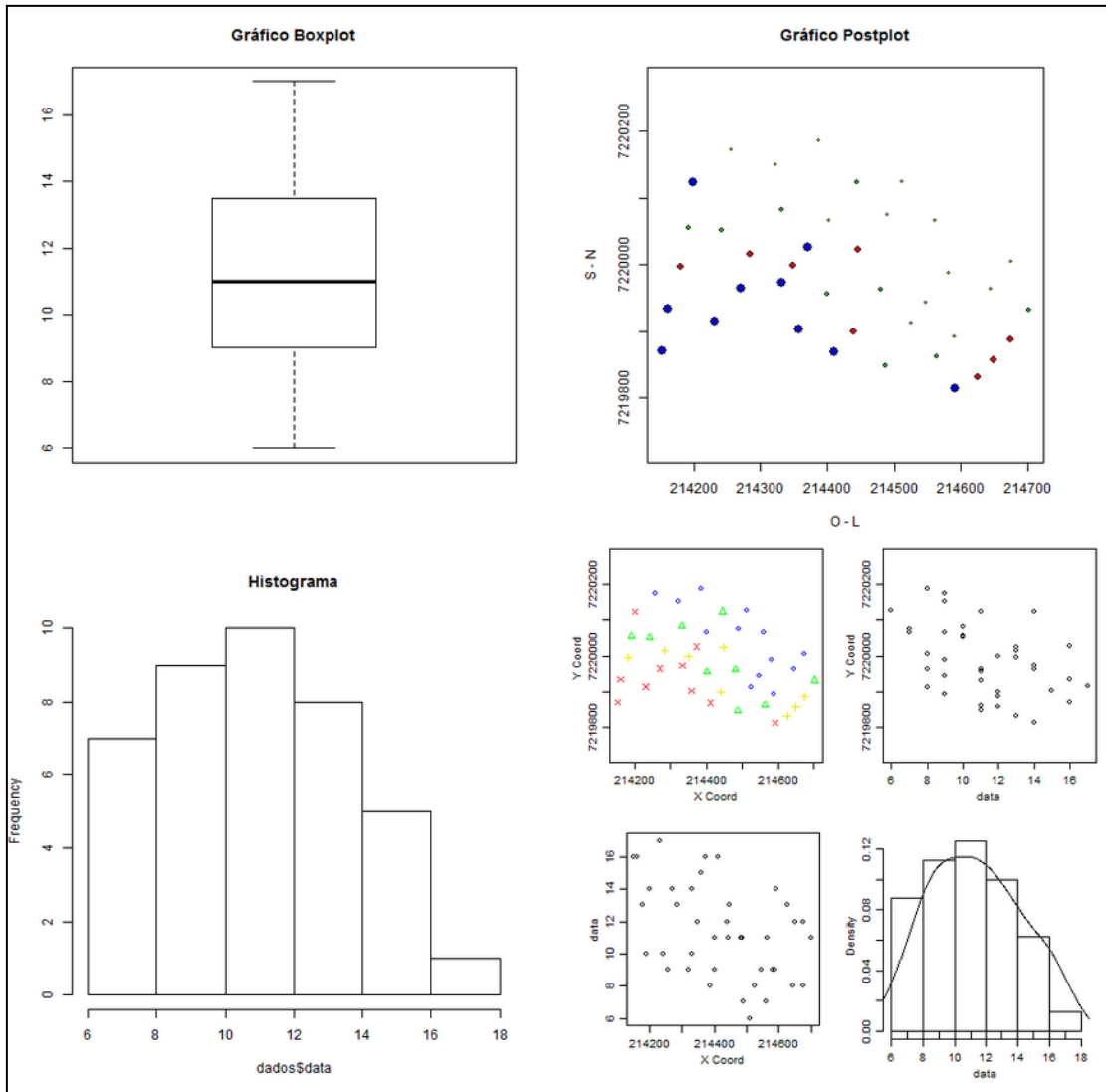


Figura 19 - Representação gráfica da estatística descritiva dos dados apresentados na tela de visualização.

Fonte: Autoria Própria.

A análise geoestatística possui como característica específica o semivariograma, sendo assim foram utilizadas duas representações gráficas; a primeira envolve os 12 melhores modelos que foram classificados em meio a 300 comparações realizadas utilizando o ISI equação (8), sendo cada modelo definido com uma cor para facilitar a visualização; a segunda é a representação gráfica do melhor modelo aplicado ao semivariograma Figura 20.

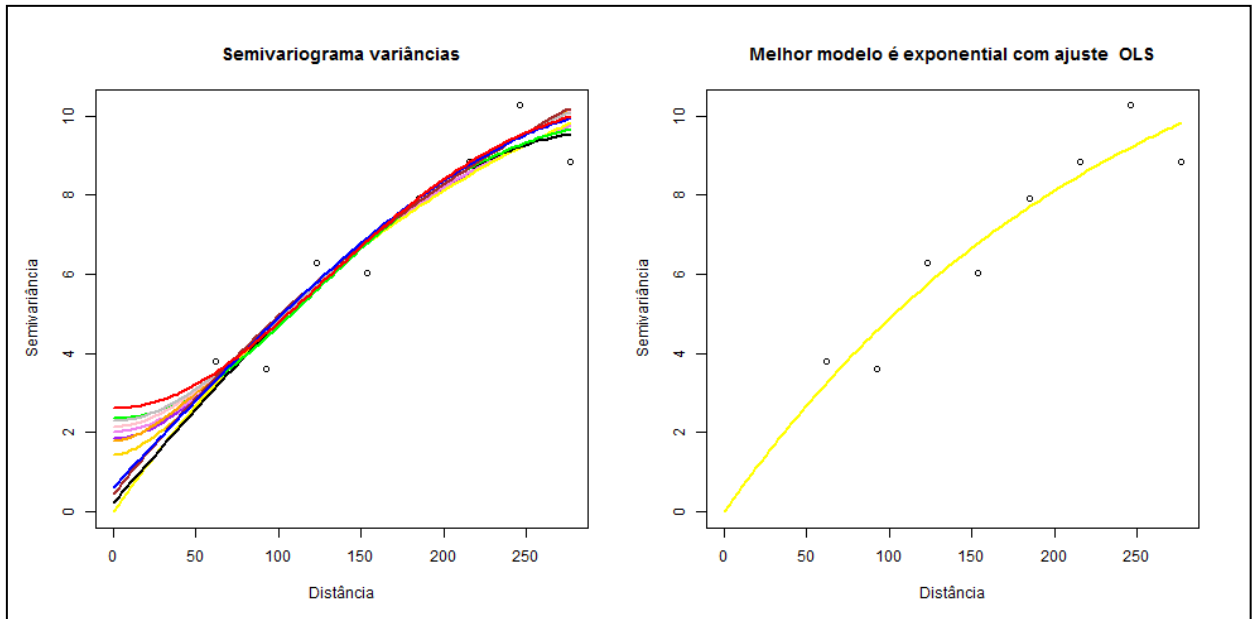



Figura 20 - Representação gráfica dos modelos ajustados e do melhor modelo ajustado.
Fonte: Autoria Própria.

Além da representação gráfica também é apresentado uma tabela com as informações dos 12 melhores modelos de semivariograma da análise selecionada. O módulo conforme definido em seu objetivo faz a recomendação do melhor modelo ajustado, destacando o com a cor verde, representado na Figura 21.

Análises Realizadas							
Código	Descrição	Área	Amostra	Usuário	Data	Status	Opções
833	AreiaArea3	Tasca_1	Areia2013	872	2017-10-06	A	





Linhas da Análise								
Código	Modelo	Método	Min. ICE	Contribuição	Alcance	Val. Kappa	ISI	Opções
5941	exp	ols	0.999997150483478	3.60714285714286	276.384099263849	0.5	0.0654084215363553	
5947	exp	wl	0.999821836710699	6.93571428571429	172.740062039906	0.5	0.22268292659928	
5938	matern	ols	0.999990153071718	10.2642857142857	276.384099263849	1.0	0.329788573566742	
5948	sph	wl	0.999983963484957	6.93571428571429	276.384099263849	0.5	0.463405159588249	

Figura 21 - Tabela com os valores dos doze melhores modelos.
Fonte: Autoria Própria.

As linhas de cada análise contêm informações do modelo, método e seus valores de contribuição, alcance, valor kappa, além do valor de ISI o qual é utilizado para verificar qual é o melhor modelo. Em cada linha contém um botão "gerar mapa" para realizar a geração do mapa utilizando o método de interpolação krigagem ordinária.

Quando acionado esse botão os valores exibidos na linha selecionada são enviados para o script do software R, onde esses valores são atribuídos as variáveis do script, e realizam os procedimentos de interpolação por krigagem ordinária. Seguindo o mesmo processo de integração realizado na análise geoestatística o módulo aguarda o retorno do processamento por parte do software R. Caso o retorno gerado estiver em erro o módulo exibe ao usuário uma mensagem alertando a existência de algum problema no processo. Caso tenha concluído com sucesso o módulo redireciona o usuário para a tela de visualização e apresentação do mapa gerado conforme representado na Figura 22.

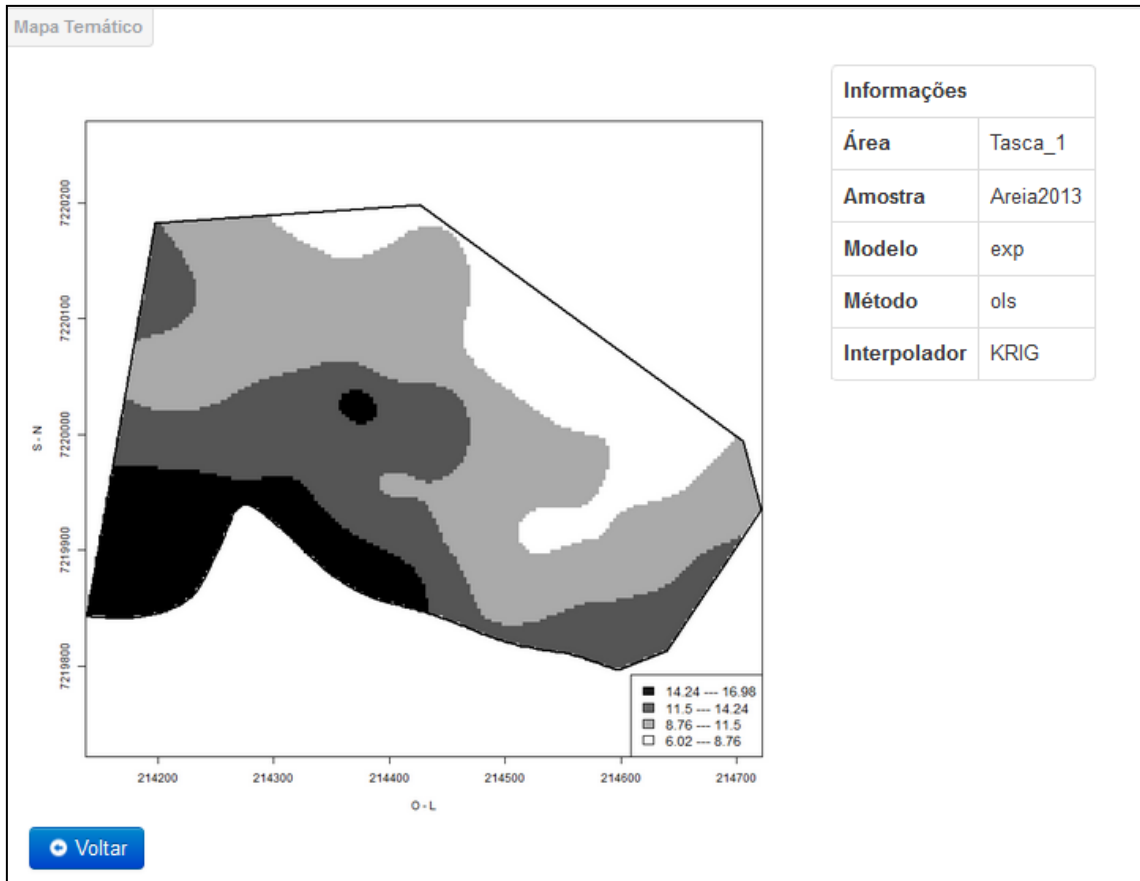


Figura 22 - Apresentação do mapa interpolado por Krigagem Ordinária.
Fonte: Autoria Própria.

Na tela de visualização do mapa as informações de área, amostra, modelo e método utilizado estão visíveis em forma de tabela ao lado do mapa. O mapa gerado é classificado em quatro classes diferenciadas por uma escala de cor que facilita sua visualização, o mapa contém uma legenda que representa cada cor e o intervalo de valores a que pertence.

4.4 PROCEDIMENTO PARA REALIZAÇÃO DA INTERPOLAÇÃO DE DADOS PELO INVERSO DA DISTÂNCIA ELEVADO A UMA POTÊNCIA

Com o objetivo de atender tanto a análise geoestatística como a análise estatística foi implementado a interpolação pelo inverso da distância elevado a uma potência. A tela para realização da análise é semelhante à de análise geoestatística, porém com seus parâmetros específicos, representado na Figura 23.

The screenshot shows a web-based software interface for performing Inverse Distance Weighted (IDW) interpolation. The interface is organized into a sidebar on the left and a main workspace on the right. The sidebar contains a navigation menu with the following items: 'Início', 'GEOESTATÍSTICA / KRIGAGEM ORDINÁRIA', 'GERAR ANÁLISE GEOESTATÍSTICA', 'VISUALIZAR ANÁLISES GEOESTATÍSTICAS', 'ESTATÍSTICA / INVERSO DA DISTÂNCIA', 'GERAR ANÁLISES IDW', and 'VISUALIZAR ANÁLISES IDW'. The main workspace is divided into two sections: 'Informações' and 'Configurações'. In the 'Informações' section, the user 'Rodrigo' is logged in. The 'Área' is set to 'Tasca_1' and the 'Amostra' is 'Areia2013'. The 'Descrição' field contains 'ArealDWTasca'. The 'Configurações' section is further divided into 'Pixel' and 'Avançado' sub-sections. The 'Pixel' section has input fields for 'Tam. X' (5), 'Tam. Y' (5), and 'Raio' (0). The 'Avançado' section has input fields for 'Expoente' (1) and 'Nº Vizinhos' (50). A blue 'Gerar' button is positioned at the bottom right of the interface.

Figura 23 - Tela para realização da interpolação pelo inverso da distância elevado a uma potência.

Fonte: Autoria Própria.

O procedimento para análise também necessita que sejam fornecidos os dados da área e da amostra juntamente com o valor da descrição, tamanho do pixel, números de vizinhos, e como característica específica o valor do expoente que deseja utilizar. Depois de realizada a análise e interpolação pelo inverso da distância elevando a uma potência o mapa gerado é exibido ao usuário (Figura 24).

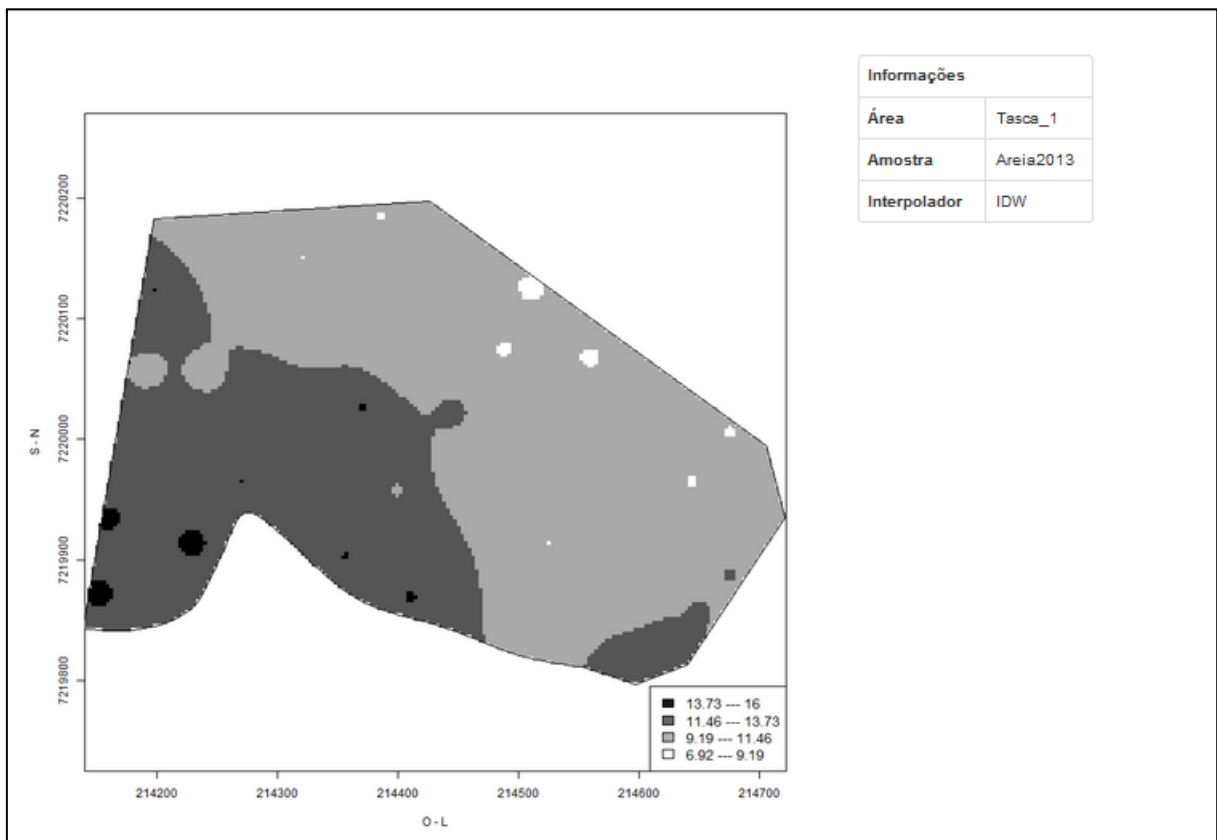


Figura 24 - Apresentação do mapa interpolado por inverso da distância elevado a uma potência.

Fonte: Autoria Própria.

O mapa possui as mesmas características do mapa gerado pelo interpolador geoestatístico, sua classificação em uma escala de cores, a legenda para identificação e a tabela com as informações da análise gerada. O processo de geração e visualização das análises foi projetado para tornar a complexidade do que seria uma análise desse nível no software R em algo que um usuário com conhecimento básico consiga gerar para obter os resultados que se deseja.

4.5 APLICAÇÃO DO MÓDULO SOBRE AS ÁREAS EXPERIMENTAIS E SEUS ATRIBUTOS

Para verificar o potencial e possibilitar uma perspectiva do desempenho da integração entre o módulo desenvolvido em um ambiente web e o software estatístico R, foram realizadas várias análises, tanto por geoestatísticas/krigagem ordinária como por interpolação pelo inverso da distância elevado a uma potência destacando a complexidade do processo geoestatístico que, por possuir uma alta sequência de comparações e procedimentos, necessita de utilização de grande quantidade de recurso de processamento em comparação a análise pelo inverso da distância a uma potência.

Para cada uma das três áreas utilizadas nessa pesquisa, foram avaliados quatro atributos, da área A foram utilizados os seguintes elementos: areia, argila, matéria orgânica (MO) e produtividade de soja; da área B os elementos de cálcio (Ca), fósforo (P), potássio (K) e a produtividade de soja; já da área C foram utilizados os elementos de cobre (Cu), ferro (Fe), potássio (K) e produtividade de soja. Com isso totalizando três áreas e doze atributos, os mapas destes atributos foram gerados pela krigagem ordinária e pelo inverso da distância elevado a uma potência, a potência definida foi a de valor um, gerando assim vinte e quatro mapas temáticos que representam a variabilidade espacial dos atributos utilizados. As representações de cada área estão agrupadas em formato comparativo para cada atributo e interpolador, acompanhada das imagens complementares da análise geoestatística realizada, representados nas Figuras 25 a Figura 30.

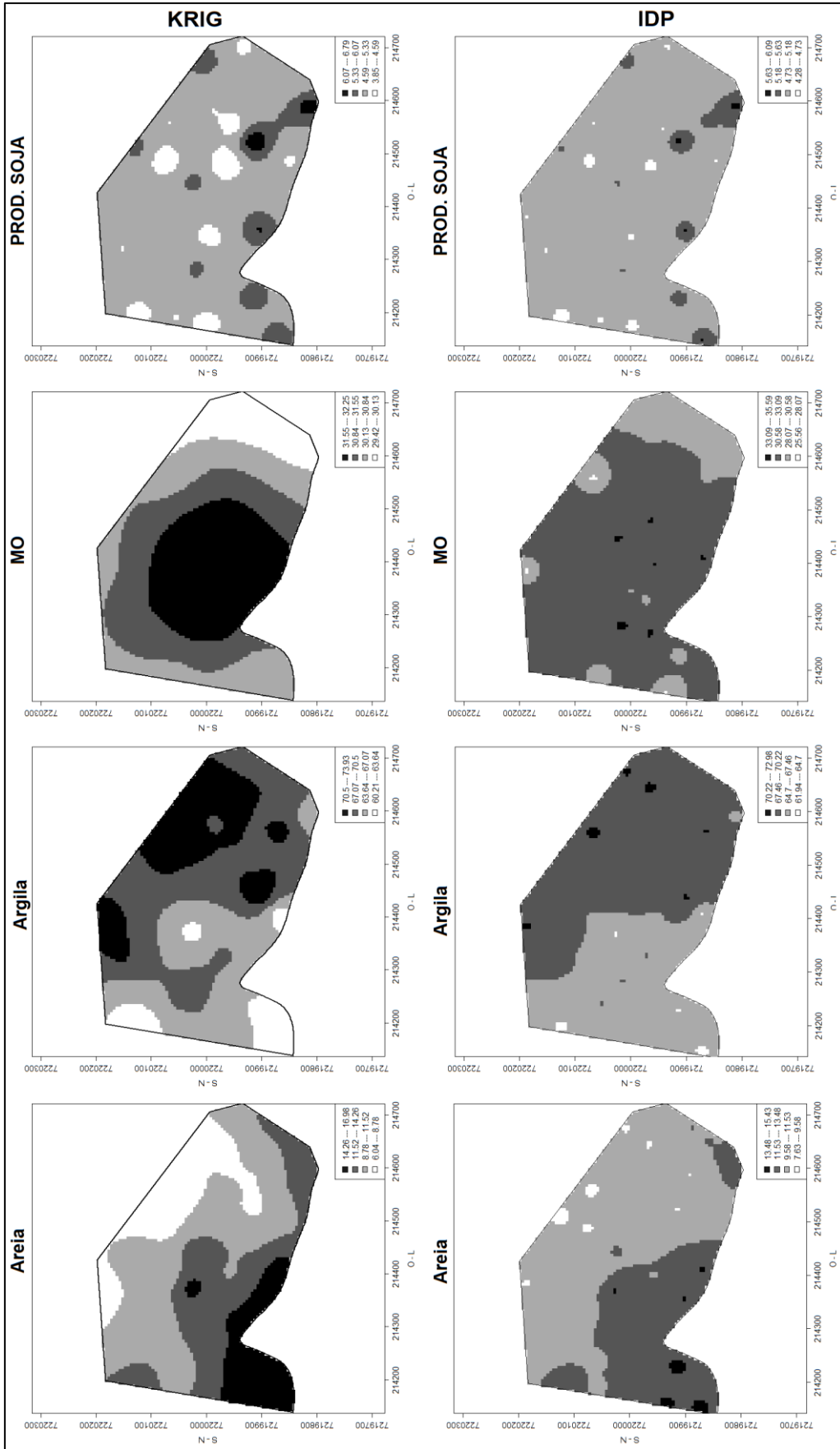


Figura 25 - Comparativo por interpolador dos mapas temáticos da Área A.
 Fonte: Autoria Própria.

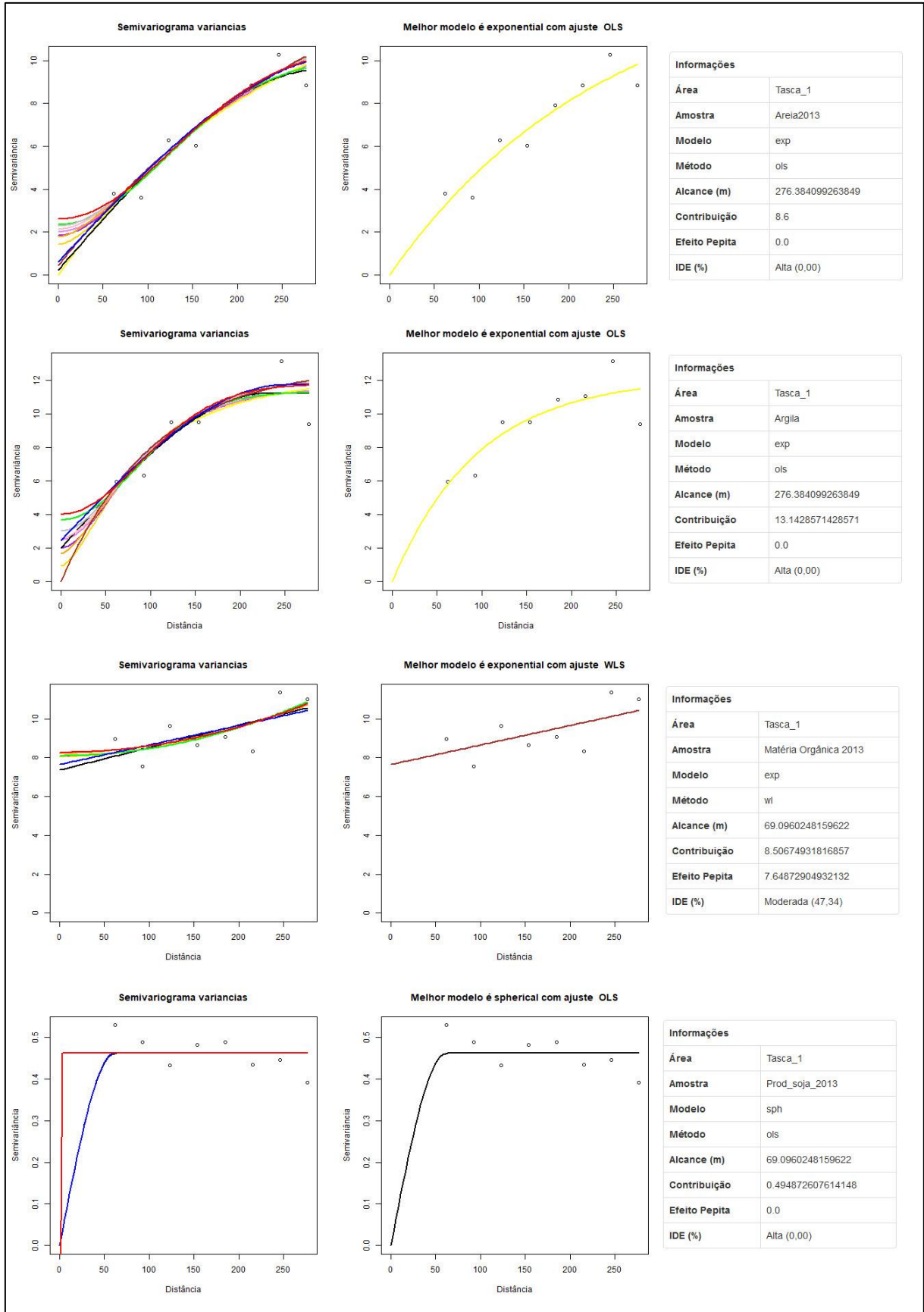


Figura 26 - Análise geostatística Área A.
Fonte: Autoria Própria.

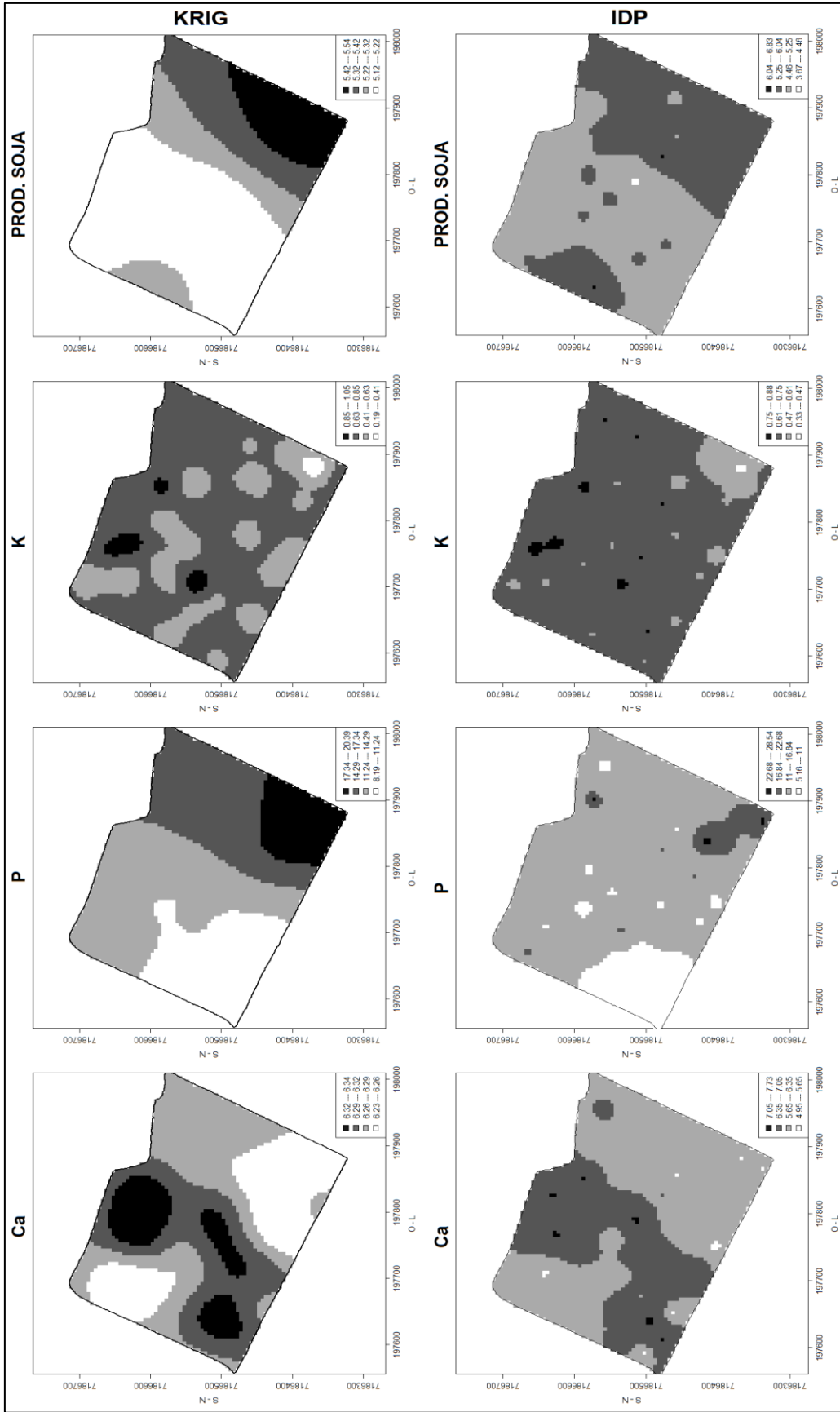


Figura 27 - Comparativo por interpolador dos mapas temáticos da Área B.
 Fonte: Autoria Própria.

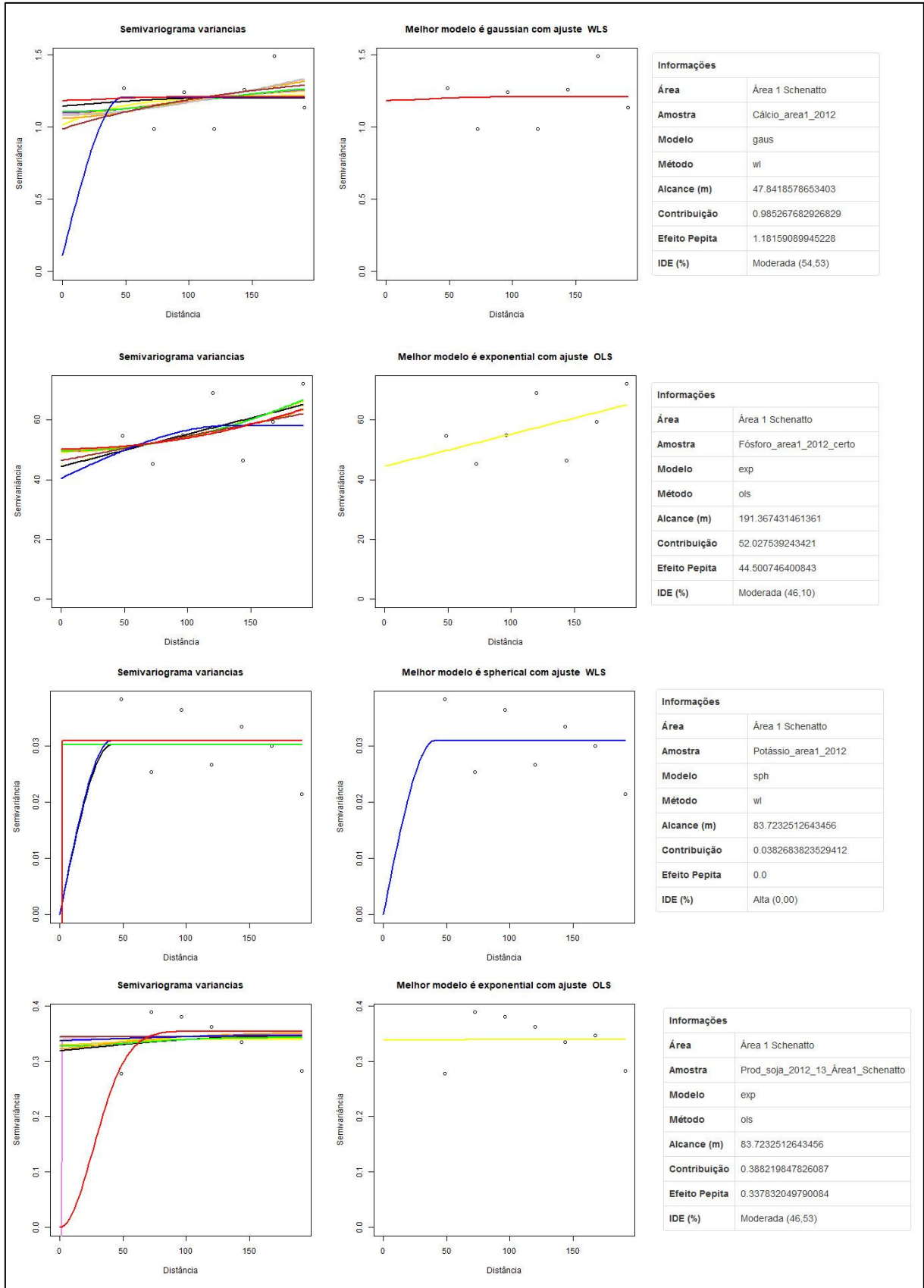


Figura 28 - Análise geoestatística Área B.
Fonte: Autoria Própria.

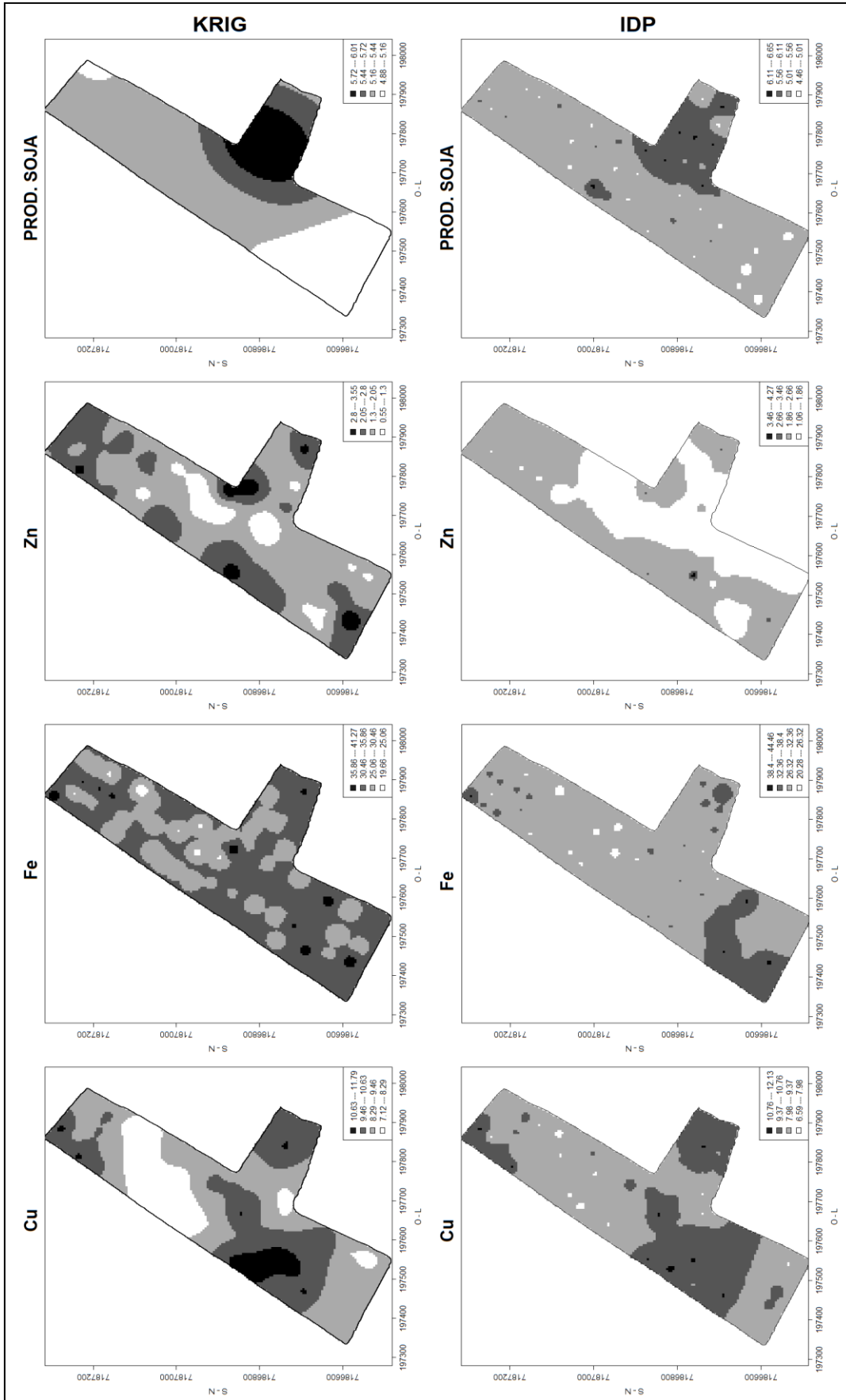


Figura 29 - Comparativo por interpolador dos mapas temáticos da Área C.
 Fonte: Autoria Própria.

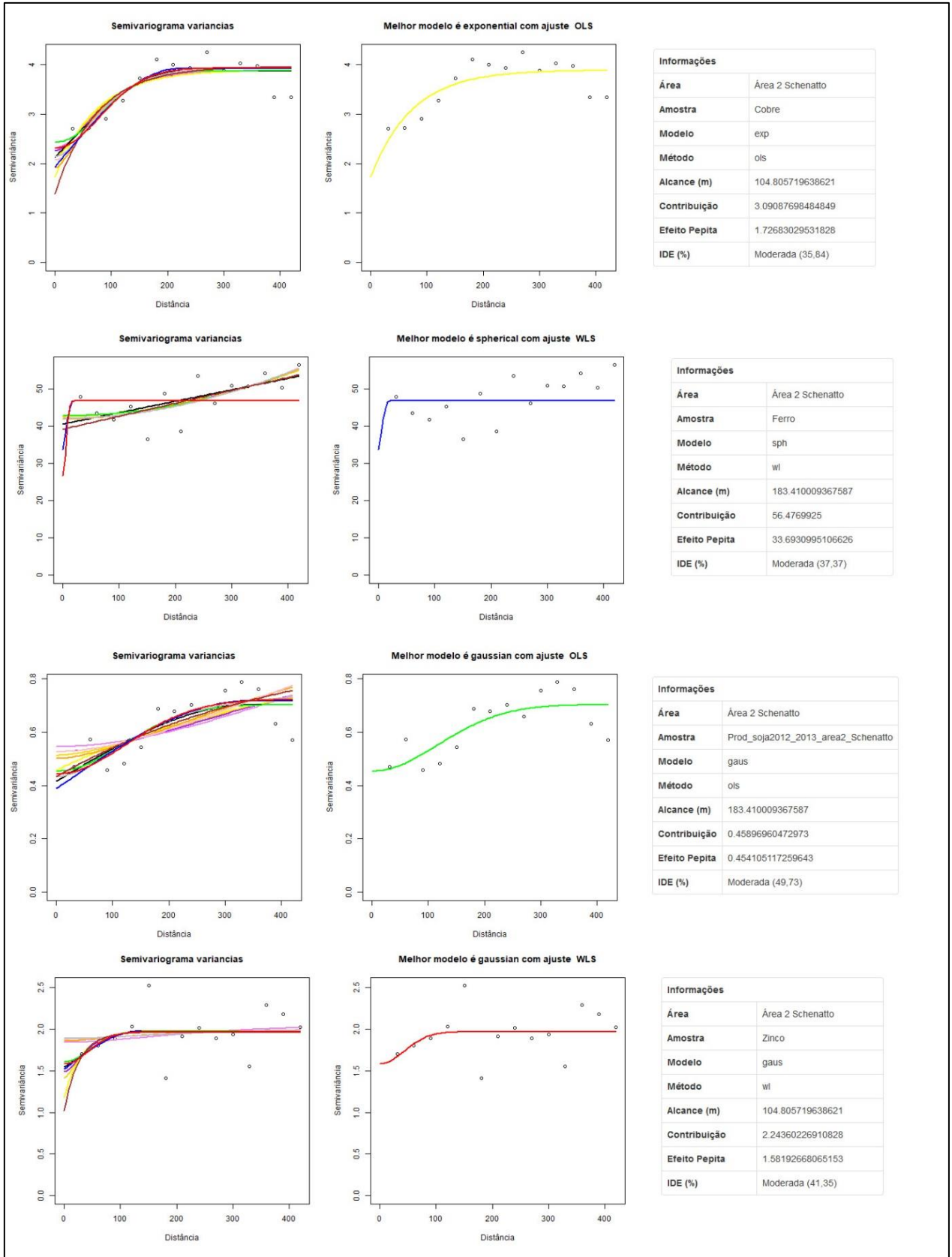


Figura 30 - Análise geostatística Área C.
 Fonte: Autoria Própria.

4.6 COMPARATIVO ENTRE O MÓDULO DESENVOLVIDO E O SOFTWARE ARCMAP

Além da geração dos mapas gerados acima pelo módulo, foi utilizado o software ArcMap 9.3 (ESRI, 2016) para realizar a comparação e validação dos mapas gerados, a proposta foi comparar as feições geradas pelos mapas do módulo com os gerados pelo ArcMap, levando em consideração os mesmos dados e parâmetros disponíveis para os atributos das áreas A e C.

Como pode ser visto nas Figuras 31 a 34 foi estruturada uma comparação entre os atributos de cada propriedade em ambos os softwares e interpoladores, como pode-se notar os mapas possuem certa semelhança, baseado na mesma classificação, algumas diferenças são visíveis porém em grande parte se assemelham, o que permite concluir que o módulo desenvolvido gera uma representatividade confiável em relação a outros softwares normalmente utilizados para esse tipo de análise e representação de dados geográficos.

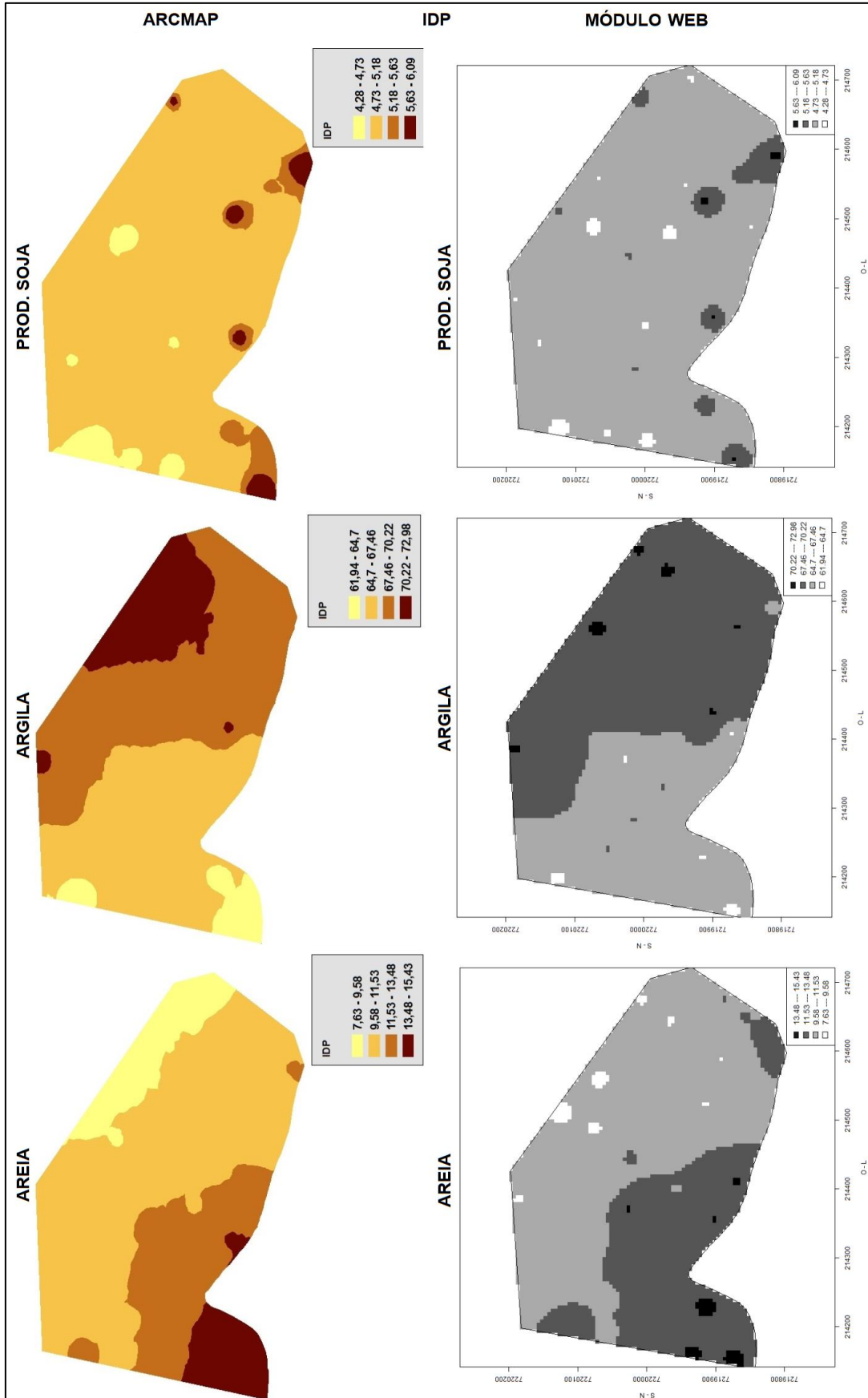


Figura 31 - Comparativo por software dos mapas temáticos da Área A - Interpolador IDP.
 Fonte: Autoria Própria.

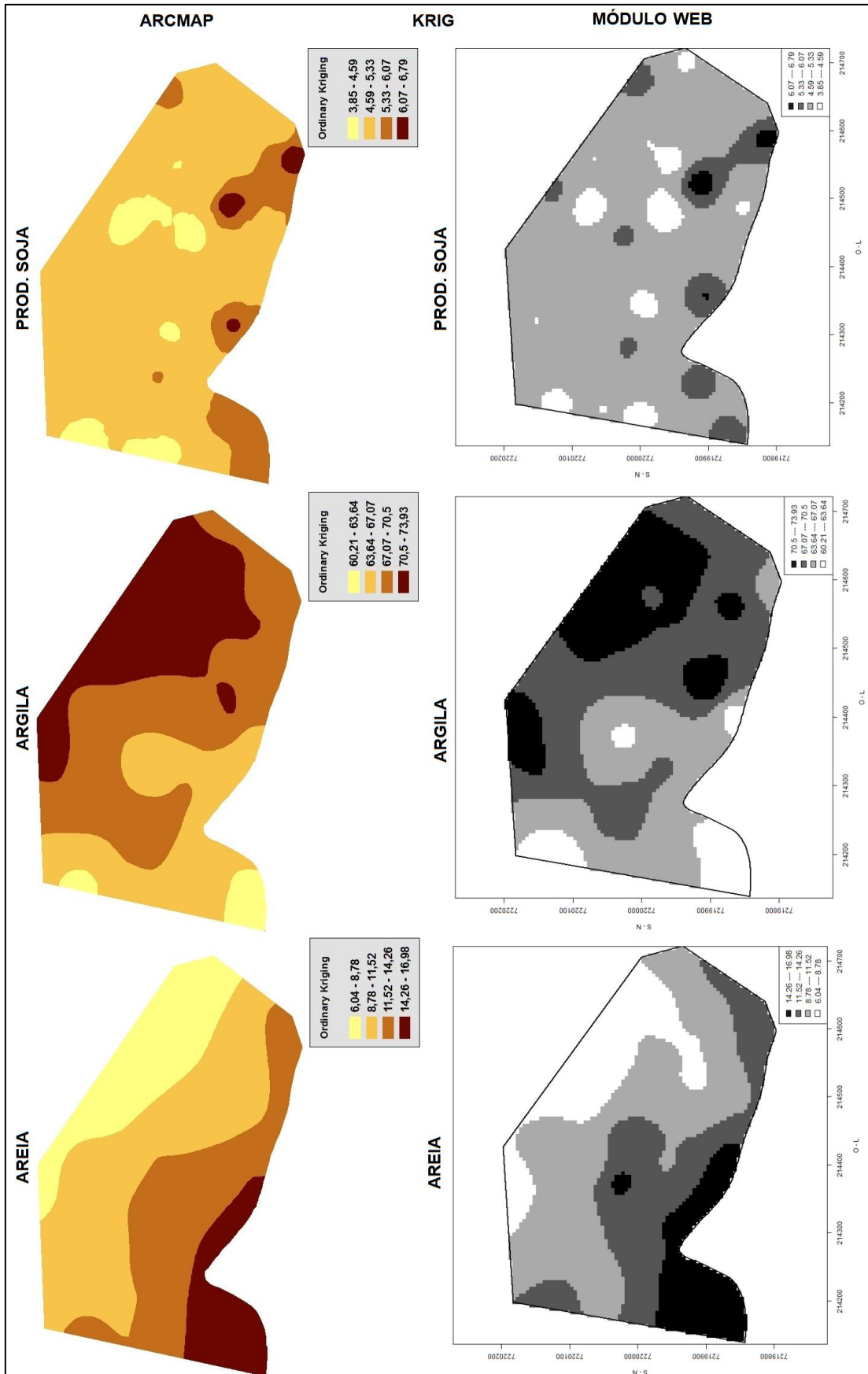


Figura 32 - Comparativo por software dos mapas temáticos da Área A - Interpolador KRIG.
 Fonte: Autoria Própria.

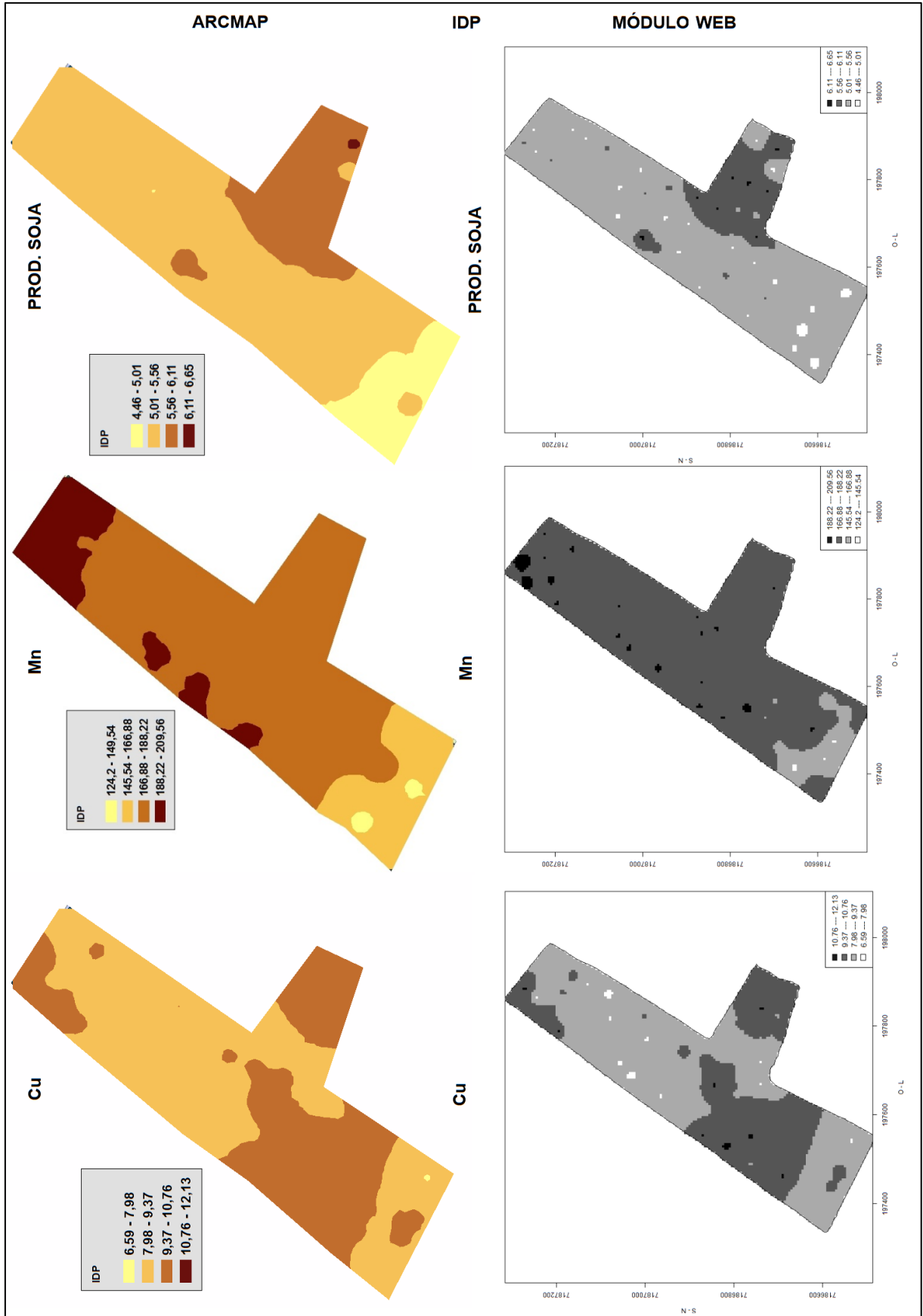


Figura 33 - Comparativo por software dos mapas temáticos da Área C - Interpolador IDP.
 Fonte: Autoria Própria.

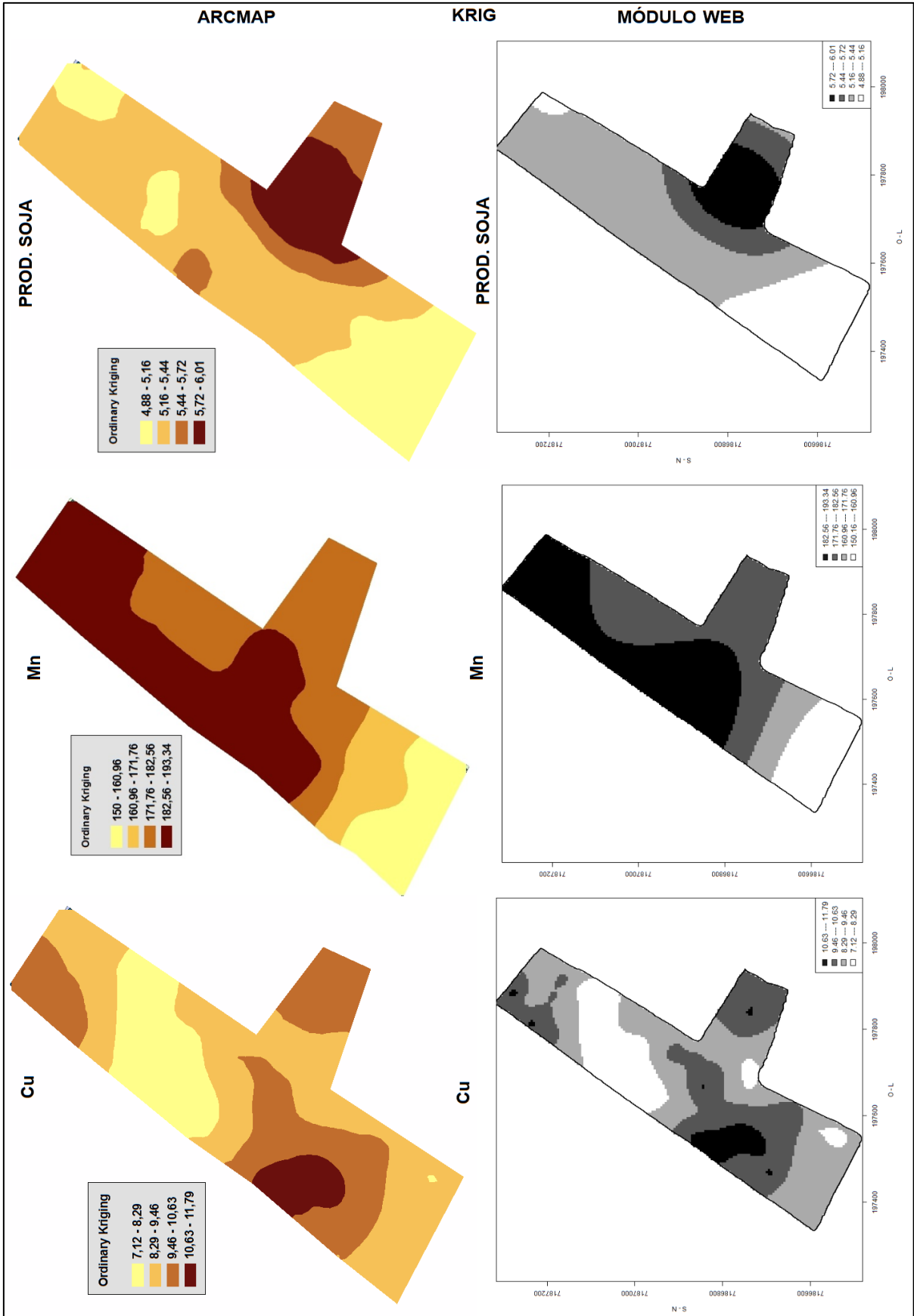


Figura 34 - Comparativo por software dos mapas temáticos da Área A - Interpolador KRIG.
 Fonte: Autoria Própria.

5 CONCLUSÃO

Procedimentos de análise estatística e geoestatística são mecanismos complexos, ainda mais quando não se tem ferramentas que efetuem algumas etapas de forma automática, além disso, em alguns casos possuem custo para se realizar esse processo, a ferramenta desenvolvida na pesquisa buscou facilitar o processo de geração de mapas temáticos de determinada área e atributo, sua interface mostrou-se ser muito simples de usar, além de automatizar processos que normalmente seriam feitos manualmente por parte do usuário, por exemplo, o ajuste do melhor modelo do semivariograma.

As funções computacionais implementadas foram eficientes e capazes de identificar o melhor ajuste para o semivariograma a ser utilizado na krigagem ordinária, bem como a utilização da interpolação pelo inverso da distância elevado a uma potência. A estrutura organizacional do sistema permitiu o armazenamento e o gerenciamento de informações de forma organizada, tal que suas interfaces intuitivas facilitaram a realização das análises, além de possibilitar consultas de mapas de forma facilitada.

Além disso, pode-se concluir que o software R e os componentes utilizados podem ser usados seguramente para realizar outros tipos de análises em ambientes de ensino e pesquisa. A integração com a ferramenta estatística R mostrou-se extremamente poderosa, essa comunicação torna possível a implementação de diversos processos não só no quesito estatístico e geoestatístico conforme abordado na pesquisa, mais para outras áreas do conhecimento, essa reutilização de implementações já existentes no software R é o que engrandece o valor da proposta apresentada.

A comparação realizada entre os mapas gerados no software ArcMap e do módulo web desenvolvido comprovaram a acuracidade dos procedimentos implementados dentro da ferramenta estatística R.

REFERÊNCIAS

- ANDRIOTTI, J. L. S. **Fundamentos de estatística e geoestatística**. São Leopoldo: Ed. Unisinos, 2003.
- ARCHELA, R. S.; THÉRY, H. Orientação metodológica para construção e leitura de mapas temáticos. **Revista Confins**.
- ASSUMPÇÃO, R, A. B.; URIBE-OPAZO, M. A.; SOUZA, E. G.; JOHANN, J. A. Uso da krigagem indicatriz na avaliação da probabilidade da produtividade de soja segundo os padrões regional, estadual e nacional. **Acta Scientiarum Agronomy**, Maringá, v. 29, n. 2, p.165-171, 2007.
- BAZZI, C. L.; SOUZA, E. G.; OPAZO, M. A. U.; SANTOS, D.; KONOPATZKI, M. R. S. **Uso da validação cruzada na seleção de modelos de semivariogramas**. XXXVIII CONGRESSO BRASILEIRO DE ENGENHARIA AGRÍCOLA, Petrolina-PE. Agosto/2009.
- BIER, V. A.; SOUZA, E. G. Interpolation selection index for delineation of thematic maps. **Computers and Electronics in Agriculture**, v. 136, p. 202-209, 2017.
- BUSSAB, W.O.; MORETTIN, P.A. **Estatística básica**. 5.ed. São Paulo: Saraiva, 2002. 526p.
- BONGIOVANNI, R.; LOWENBERG-DEBOER, J. Precision agriculture: Economics of nitrogen management in corn using site-specific crop response estimates from a spatial regression model. **American Agricultural Economists Association Annual Meeting**, Chicago, Illinois, 2001.
- CÂMARA, G.; MEDEIROS J. S. **Geoprocessamento para projetos ambientais**. São Jose dos Campos: INPE, 1996.
- CÂMARA, G.; MEDEIROS, J. S. **Princípios básicos em geoprocessamento**. In: Sistemas de informações geográficas: aplicações na agricultura. 2. ed. ver. ampl. Brasília, DF: Embrapa-SPI: Embrapa-CPAC, pp.3-11, 1998.
- CÂMARA, G.; ORTIZ, M.J. "Sistemas de Informação Geográfica para Aplicações Ambientais e Cadastrais: Uma Visão Geral". In: Souza E SILVA, M., "**Cartografia, Sensoriamento e Geoprocessamento**", cap. 2, pp.59-88. Lavras, UFLA/SBEA, 1998.
- CAVALCANTI, L.; **VRaptor: Desenvolvimento ágil para web com Java**, Editora Casa do Código, 2014.
- CHILES, J. P., DELFINER, P., 1999. Geostatistics: Modelling Spatial Uncertainty. Wiley, New York.

COELHO, Edy Célia et al . Influência da densidade amostral e do tipo de interpolador na elaboração de mapas temáticos. **Acta Sci., Agron.**, Maringá , v. 31, n. 1, p. 165-174, Mar. 2009.

CRESSIE, N.A.C. **Statistics for Spatial Data**. Edição Revisada. New York: John Wiley & Sons, 1993. 900p.

DAVID, M. **Geostatistical Ore Reserve Estimation**. Amsterdam: Elsevier, 1977. 205p.

DEBASTIANI, F. **Influência local em modelos espaciais lineares com distribuição da família de contornos elípticos**. Dissertação (Mestrado em estatística) – Universidade Federal de Pernambuco, 2012.

DEITEL H., DEITEL P. **Java - Como Programar - 8ª Ed.** 2010.

DIGGLE, P. J.; RIBEIRO JÚNIOR, P. J. Model Based Geostatistics. In: SINAPE, 14., 2000, Caxambu. **Anais...** São Paulo: Associação Brasileira de Estatística, 2000.

DOERGE, T. A. **Site-specific management guidelines**. Norcross: Potash & Phosphate Institute, 2000. 135p.

DRUCK, S.; CARVALHO, M. S. **Análise Espacial de Dados Geográficos**. Brasília. EMBRAPA, 2004. 209 p.

FERRAZ, Gabriel A. E S. et al . Variabilidade espacial e temporal do fósforo, potássio e da produtividade de uma lavoura cafeeira. **Eng. Agríc.**, Jaboticabal , v. 32, n. 1, p. 140-150, Feb. 2012 .

GAVIOLI A., SOUZA E. G., BAZZI C.L., GUEDES L.P., SCHENATTO K., **Optimization of management zone delineation by using spatial principal components**, Computers and Electronics in Agriculture, Volume 127, Pages 302-310, ISSN 0168-1699, 2016;

GONÇALVES, A. C. A.; FOLEGATTI, Marcos Vinicius ; MATA, José de Deus Viana da . Análise exploratória e geoestatística da variabilidade de propriedades físicas de um argissolo vermelho. **Acta Scientiarum**, Maringá, v. 23, n.5, p. 1149-1157, 2001.

GONÇALVES, E. "**Desenvolvendo Aplicações Web com JSP, Servlets, Javasever Faces, Hibernate, EJB 3 Persistence e Ajax**."Rio de Janeiro – Editora Ciência Moderna, 2007.

GREGO, C. R.; OLIVEIRA, R. P. de; VIEIRA, S. R. **Geoestatística aplicada a agricultura de precisão**. Embrapa Solos, 2014. 80p.

GRIFFIN, T. W., et al. **Adoption, profitability, and making better use of precision farming**, Department of Agricultural Economics, Purdue University, 2004

GROSSI SAD, J.H. **Fundamentos sobre variabilidade dos depósitos minerais**. Rio de Janeiro. DNPM/CPRM - GEOSOL, 1986. 141p.

GODWIN, R.J; MILLER, P.C.H. A review of the technologies for mapping within-field variability. **Biosystems Engineering**, London, v.84, n.4,p. 393-407, 2003.

GEBBERS, R.; ADAMCHUK, V. I. Precision agriculture and food security. **Science**, Washington, v. 327, n. 5967, p. 828-831, 2010.

GUEDES FILHO, Osvaldo; VIEIRA, S. R. ; CHIBA, M. K. ; MELLIS, E. V. ; DECHEN, S. C. F. ; De MARIA, I. C. . Variabilidade espacial dos teores foliares de nutrientes e da produtividade da soja em dois anos de cultivo em um latossolo vermelho. **Revista Brasileira de Ciência do Solo** (Impresso), v. 34, p. 1503-1514, 2010.

GUEDES, Isabel Carolina de Lima et al . TÉCNICAS GEOESTATÍSTICAS E INTERPOLADORES ESPACIAIS NA ESTRATIFICAÇÃO DE POVOAMENTOS DE Eucalyptus sp. **Ciênc. Florest.**, Santa Maria , v. 22, n. 3, p. 541-550, Sept. 2012 .

HORTON, P. Prospects for crops improvement through the genetic manipulation of photosynthesis: morphological and biochemical aspects of light capture. **Journal of Experimental Botany**, v.51, p. 475-485. 2000.

ISAAKS, E. H.; SRIVASTAVA, R. M. **An introduction to applied geostatistics**. New York: Oxford Universit, 1989. 561p.

JAKOB, A. A. E. YOUNG, A. F. **O uso de métodos de interpolação espacial de dados nas análises sociodemográficas**. XV Encontro Nacional de Estudos Populacionais. Brasil, 2006.

JOURNAL, A.G; HUIJBREGTS, C.J. **Mining geostatistics**. London: Academic Press, 1978. 600p.

KAVANAGH, C. R. Estimacão de parâmetros no ajuste de modelos teóricos a semivariogramas experimentais. Cascavel, PR, 2001. 85f. Dissertação (Mestrado em Engenharia Agrícola) – **Centro de Ciências Exatas e Tecnológicas**, Universidade Estadual do Oeste do Paraná.

KESTRING, B. F. **Análise geoestatística de mapas temáticos da produtividade da soja com diferentes grades amostrais**. Dissertação (Mestrado em Engenharia Agrícola), Universidade Estadual do Oeste do Paraná, Cascavel, 2011. 15 p.

KRIGE, D.G. **A statistical approach to some basic mine evaluation problems on the Witwatersrand**. J Chem Metall Min Soc Afi 1951; 52: 119-39.

KOCH, B.; KHOSLA, R. The role of precision agriculture in cropping systems. **Journal of Crop Production**. Philadelphia, v. 9, p. 361-381, 2003.

LAURETTO, M., 2011, **Análise Exploratória de Dados**, USP - Universidade de São Paulo.

LI, H.; LASCANO, R. J.; BOOKER, J.; WILSON, L. T.; BRONSON, K. F.; SEGARRA, E. State-space description of heterogeneity: water and nitrogen use in cotton. **Soil Science Society of America Journal**, Madison, 2002.

MAGNE, M.A., Cerf, M., Ingrand, S., 2010. A conceptual model of farmers' informational activity: a tool for improved support of livestock farming management. **Animal** 4, 842–852.

MATHERON, G. 1963. Principles of Geostatistics. **Economic Geology**, v. 58, p. 1246-1266.

MATHERON, G. **The Theory of Regionalized Variables and Its Applications**, Universidade da Virgínia, 1971.

MAZZINI, P. L. F.; SCHETTINI, C. A. F. Avaliação de metodologias de interpolação espacial aplicadas a dados hidrográficos costeiros quase sinóticos. **Brazilian Journal of Aquatic Science and Technology**, v.13, n.1, p.53-64, 2009.

MC BRATNEY, A.G; WEBSTER, A.G. Choosing functions for semi-variograms and fitting them to sampling estimates. **Journal of Soil Science**, v.37, p.617-639, 1986.

MELLO, J.M.; BATISTA, J.L.F.; RIBEIRO JÚNIOR., P.J.; OLIVEIRA, M.S. Ajuste e seleção de modelos espaciais de semivariograma visando à estimativa volumétrica de *Eucalyptus grandis*. **SCIENTIA FORESTALIS**. n. 69, p.25-37, dez. 2005.

MINASNY, B; MCBRATNEY, A. **The Matérn function as a general model for soil variogramas**. *Geoderma*. v. 128, n. 3-4, p. 192-207, 2005.

MOLIN, J. P. Agricultura de precisão no Brasil: estado atual e perspectivas. Cuiabá, **Anais...**, CONGRESSO BRASILEIRO DE SOJA, 6.,2012, Cuiabá.

MOLIN, J.P. **Agricultura de precisão: o gerenciamento da variabilidade**. Piracicaba: ESALQ/USP, 2008. 83 p.

MONTES, F.; HERNÁNDEZ, J. M.; CAÑELLAS, ISABEL. **A geostatistical approach to cork production sampling estimation in *Quercus suber* forests**. *Canadian Journal of Forest Research*, vol.35, p.2787-2796, 2005.

MURAKAMI, E., et al. An infrastructure for the development of distributed service oriented information systems for precision agriculture, 2007. **Comput. Electron. Agric.**

OLIVEIRA, T. C. A ; et al . Indexação dos dados espaciais do Banco de Dados do Inventário de Minas Gerais. In: Simpósio Brasileiro de Sensoriamento Remoto, 2007, Florianópolis-SC. **Anais**. v. 9. p. 5973-5981.

PEDERSEN, S.M. et al., 2003. Adoption of precision farming in Denmark. In: Stafford, J., Werner, A. (Eds.), **Proceedings of the 4th European Conference on Precision Agriculture**. Academic Publishers, Wageningen, pp. 533–538.

POSTGRESQL (2016). **Sobre o PostgreSQL**. Disponível em <<https://www.postgresql.org.br/sobre>>

POSTGIS (2016). **Sobre o PostGIS**. Disponível em <<http://postgis.net/>>

R Development Core Team (2016). R: A language and environment for statistical computing. **R Foundation for Statistical Computing**, Vienna, Austria. ISBN 3-900051-07-0, URL <http://www.R-project.org>.

RIBEIRO JUNIOR, P. J. **Métodos geoestatísticos no estudo da variabilidade espacial de parâmetros de solo**. Dissertação (Mestrado em estatística e Experimentação Agrônômica) - Escola Superior de Agricultura Luiz de Queiroz, Universidade de São Paulo, Piracicaba, 1995.

ROBERT, P.C., 1999. Precision agriculture: research needs and status in the USA. In: Stafford, J.V. (Ed.), **Proceedings of the 2nd European Conference on Precision Agriculture**, Part 1. Academic Press, SCI/Sheffield, pp. 19–33.

SABINO, Claudia Vilhena Schayer; LAGE, Ludmila Vieira; ALMEIDA, Katiane Cristina de Brito. **Uso de métodos estatísticos robustos na análise ambiental**. Eng. Sanit. Ambient., Rio de Janeiro, v. 19, n. spe, p. 87-94, 2014.

SANTOS, P. C. et al. O emprego da geostatística na determinação do tamanho "ótimo" de amostras aleatórias com vistas a obtenção de estimativas dos volumes dos fustes de espécies florestais em Paragominas, estado do Pará. **Acta Amaz.**, v. 41, n. 2, p.213-222, 2011.

SARAIVA, A.M., et al., 1998. An object model for field information systems. In: **Proceedings of the 4th International Conference on Precision Agriculture**, St. Paul, MN, USA. ASA/CSSA/SSSA, pp. 1355–1366.

SCHENATTO K. et al., Data interpolation in the definition of management zones. **Acta Scientiarum**, ISSN 1806-2563, 2016.

SILVA, V. R; REICHERT, J. M.; STORCK, L.. Variabilidade espacial das características químicas do solo e produtividade de milho em um Argissolo Vermelho-Amarelo distrófico arênico. **Rev. Bras. Ciênc. Solo**, Viçosa, v. 27, n. 6, p. 1013-1020, Dec. 2003.

SILVA, S. A.; SOUZA LIMA, J. S.; SOUZA, G. S.; OLIVEIRA, R. B. **Avaliação de interpoladores estatísticos e determinísticos na estimativa de atributos do solo em agricultura de precisão**. Idesia, Arica, v. 26, n. 2, p. 75-81, 2008.

SILVA e SILVA, L. G.. **Novas Ferramentas para Visualização Georreferenciada de Dados: Uma integração entre R e Google Maps (Earth)**. 2012.

SILVA, S. G. L. **Novas ferramentas para visualização georreferenciada de dados: uma integração entre R e Google Maps**. (Mestrado em Ciências) - Universidade de Federal de Minas Gerais, Belo Horizonte, 2012.

SOARES, A. **Geoestatística para as ciências da terra e do ambiente**. Portugal: IST Press, 2000.186 p.

SPERANZA, E. A. et al. **Métodos e ferramentas de interpolação para geração de mapas georreferenciados para irrigação de precisão**. Anais CONGRESSO BRASILEIRO DE AGRICULTURA DE PRECISÃO, 2010, Ribeirão Preto.

SWINTON, S. M., and J. Lowenberg-DeBoer. 1998. Evaluating the Profitability of Site-Specific Farming. *J. Prod. Agric.* 11:439-446. doi:10.2134/jpa1998.0439

SORENSEN C. G. Functional requirements for a future farm management information system. **Computers and Electronics in Agriculture**. Volume 76, Issue 2, May 2011, Pages 266–276.

STAFFORD, J.V., EVANS, K., 2000. Spatial distribution of potato cyst nematode and the potential for varying nematicide application. **Proceedings of Fifth International Conference on Precision Agriculture**, July 16/19, 2000. Bloomington, MN, USA.

SRIVASTAVA, R. M. **Describing spatial variability using geostatistics analysis**. In: SRIVASTAVA, R. M.; ROUHANI, S.; CROMER, M. V.; JOHNSON, A. I.; DESBARATS, A. J. (Ed.). *Geostatistics for environmental and geotechnical applications*. West Conshohocken: ASTM, 1996. p. 13-19. <http://dx.doi.org/10.1520/STP16110S0>

VIERIA, S.R. **Geoestatística aplicada a agricultura de precisão**. In: GIS Brasil 98, Curitiba, PR: Sagres, 1998, 53p.

WANG, D.; PRATO, T.; QIU, Z; SUDDUTH, K. A. Economic and environmental evaluation of variable rate nitrogen and lime application for claypan soil fields. *Precision Agriculture*, **Secaucus**, v. 4, p. 35-52, 2003.

WENDROTH, O.; JÜRSCHIK, P.; KERSEBAUM, K. C.; REUTER, H.; VAN KESSEL, C.; NIELSEN, D. R. Identifying, understanding and describing spatial processes in agricultural landscapes – four case studies. **Soil & Tillage Research**, Oxford, v. 58, n. 3-4, p. 113- 127, 2001.