

Marystella Duarte Correia

Modelagem geoestatística da distribuição de carbono do solo e biomassa de herbáceas em sistema silvopastoril

Recife, 25 de fevereiro de 2013



UNIVERSIDADE FEDERAL RURAL DE PERNAMBUCO
PRÓ-REITORIA DE PESQUISA E PÓS-GRADUAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM BIOMETRIA E ESTATÍSTICA APLICADA

Modelagem geoestatística da distribuição de carbono do solo e biomassa de herbáceas em sistema silvopastoril

Dissertação apresentada ao Programa de Pós-Graduação em Biometria e Estatística Aplicada como exigência parcial à obtenção do título de Mestre.

Área de Concentração: Biometria e Estatística Aplicada

Orientador: Prof. Dr. Rômulo Simões Cezar Menezes

Coorientador: Prof. Dr. Ricardo Alves de Olinda

Recife-PE, 25 de fevereiro de 2013

UNIVERSIDADE FEDERAL RURAL DE PERNAMBUCO
PRÓ-REITORIA DE PESQUISA E PÓS-GRADUAÇÃO
PROGRAMA DE PÓS-GRADUAÇÃO EM BIOMETRIA E ESTATÍSTICA APLICADA

**Modelagem geoestatística da distribuição de carbono do solo e biomassa de
herbáceas em sistema silvopastoril**

Marystella Duarte Correia

Dissertação julgada adequada para obtenção do título de mestre em Biometria e Estatística Aplicada, defendida e aprovada por unanimidade em 25/02/2013 pela Comissão Examinadora.

Orientador:

Dr. Rômulo Simões Cezar Menezes
Universidade Federal de Pernambuco

Banca Examinadora:

Dr. Ricardo Alves de Olinda
Universidade Estadual da Paraíba
DE-UEPB

Dr. Dário Costa Primo
Universidade Federal de Pernambuco
DEN-UFPE

Dr. Gustavo Henrique Esteves
Universidade Estadual da Paraíba
DE-UEPB

Dedico

Clarice e Francisco
Thávyla e Thácyla
Alex Ramalho
Prof. Ricardo

Agradecimentos

Não consigo entender o porquê deste capítulo ser tão difícil de escrever. Vem um filme na cabeça de tudo o que se passou, os olhos já começam a encher com lágrimas, a tela do computador começa a embaçar, os óculos ficam sujos, retira o óculos, cega, não enxerga direito!!! Coloca de novo, e assim vai. Memorizar para sempre em palavras todas as pessoas que acompanharam o desenrolar do negócio, é legal, mas é difícil. Sou teimosa, vamos ver no que vai dar!

Quando se termina a graduação, fica uma pergunta no ar "e agora?" Vamos trabalhar, passam-se alguns meses e a perturbação na cabeça continua "estudei tanto, e agora?" Vamos estudar para o mestrado, "Eita!" O bicho pegou, é aí onde entram duas pessoas importantes, olham para mim e dizem: - "Vá para a UFCG, entre na sala de aula e estude, é o que você quer, então vá!" Profa. Ruth Nascimento e Prof. Gil Luna. Vem outras pessoas que protagonizaram esse acesso Profa. Grayce-Mary e Profa. Ana Cristina Brandão. Nossa!!! Foram palavras, ensinamentos, conversas fundamentais no processo de graduação-mestrado. Pronto, chegou o dia da prova, e agora, ir para Recife e ficar onde? Na casa de minha prima que é uma "riqueza"!

Diego não vai trabalhar me leva na UFRPE, depois, vão me buscar, Lucinha logo pergunta pela prova, respondo que "fiz tudo e que tinha gostado, mas, era muito difícil passar, pois tinha muita gente". Pelo menos tentei! Chega o grande dia! Resultado final, que ansiedade triste, nossa que alegria! Acabei meus créditos de celular, agradecendo as pessoas que me ajudaram a conquistar essa vitória, vem logo na cabeça, vai ser complicado, mas vou conseguir, e foi!

Difícil encontrar uma pessoa para ficar na loja, treinar e deixá-la capacitada para trabalhar na empresa RC Tiras com dois "amigos patrões" Robson Lins e Ana Carla Barreto. Conseguimos, já posso ir tranquila! Aliery.

E agora, vou ficar onde em Recife? Gerlúcia liga dizendo que eu poderia ficar no apartamento deles, tá bom eu fico por alguns meses, até encontrar um apartamento para dividir. Quem disse que Diego e Lucinha me deixaram sair de lá? Iria ficar complicado, além do aluguel, das despesas de Recife, ainda tinha as despesas das meninas, as viagens para Campina, o dinheiro da bolsa não dava! Fiquei morando com eles. Só Deus mesmo para

retribuir tudo o que vocês três fizeram por mim, a terceira pessoa que falei e que ainda não surgiu, o nome dela Maria Júlia, "pirrainha", você não tem noção de como sua companhia me ajudou, do quanto seu abraço me deu forças, eu precisava de uma terceira filha perto de mim e apareceu! Meus primos, eu os amo demais, não sei o que faço para agradecer tudo o que vocês três fizeram por mim, a única coisa que posso pedir é que Deus sempre esteja com vocês!

Primeiro dia de aula! Amizades a primeira vista (Danila e Ana Clara), concorrentes? Pode até ser, mas o carinho (Priscilla, Gabi), a cumplicidade (Érika, Nyedja e Renata) e a lealdade (Neto), não há concorrência para esses, quantos momentos bons passamos juntos, estudando, às vezes falando besteiras, rindo para o vento, lágrimas foram enxugadas e tentei enxugar de alguém, surge nesse momento, uma mascote, a sobrinha do departamento "Sophia", "Eita" pimentinha, nos seminários, você era a atriz principal rs rs! Mas como resolver todas as questões, tem momentos que não sai um $2 + 2$ help! Samuel e Rodrigo! Ainda bem que vocês estavam sempre no departamento!

A tarefa mais difícil! Deixar nem que sejam por alguns dias da semana, Thávyla e Beatriz, minhas filhas, meu orgulho, minha vida, meu oxigênio, minha razão de viver, lutar e vencer. Por vocês eu mudo de lugar, de cidade, de estado, de país, tudo em busca de proporcionar uma vida melhor, de estudo e de lazer. A mamãe ama demais as duas, vocês pra mim são apenas, tudo! Mas como iria viajar? Vou deixá-las com quem? Mainha, Painho (pra variar rs rs rs), exemplo de vida para nós três, só tenho que agradecer por tudo o que fizeram, desde o divórcio até hoje, se não fosse a ajuda de vocês, a caminhada teria sido mais espinhosa, graças a Deus que tenho pais maravilhosos! Minha eterna gratidão. Sei que em alguns momentos fui ignorante, perdão e muito obrigada! Se por elas eu mudo de lugar, por você Alex Ramalho, eu fico em Campina, meu escape, meu companheiro, minha força, não precisa falar nada, só sua presença me dá ânimo para continuar, você, não sei como, recarrega minhas forças! Desculpa a todos, pela minha ausência, pelas minhas conversas chatas da dissertação, enfim. Amo demais!

Obrigada ao meu cunhado Renato e minha amada irmã Claristela, por terem filhos lindos! Pedro e Clarice, como é bom tê-los por perto. Nos momentos juntos, vocês só transmitem alegria! Telinha, nossa família real é simplesmente "real" e única, graças a Deus que fazemos parte, sempre ajudando uma a outra em todos os momentos e em todas as circunstâncias. Minha irmã muito obrigada!

Aos meus amigos Elaine, Vivian, Pollyanna, Kely, Frade, Corrinha, Marcinha, Lourdinha, Mara, Francisca e Zuleide, raramente nos encontramos, mas o que vivemos no

passado, conta como pontos até os dias de hoje.

Chega o dia de "quem vai ser meu orientador" Pesquisa no Lattes, gostei dos trabalhos do Prof. Rômulo Menezes (adorava quando desenhava as árvores no quadro e no papel), envia-se e-mail, marca-se reunião, tudo certo! Rômulo meu orientador, chega o momento de definir o trabalho "Estatística Espacial" e agora? Extra! Extra! Procura-se desesperadamente um professor que saiba do assunto! Extra! Extra! Vamos à procura: UFRPE...UFPE...UFPB...UFCG e pra salvar a pátria conversando com o Prof. Gil Luna UEPB, indica-se Prof. Ricardo Olinda ESALQ-UEPB, mas não tem o e-mail, e agora? Prof. Gustavo Esteves, entra na história, me dar os parabéns do mestrado e me envia o e-mail do Prof. Ricardo, lá vai a aluna (Eu) enviar um e-mail na maior cara de pau para o Prof. Ricardo, não o conhecia, (ele não sabia a batalha que iria enfrentar, se soubesse não tinha respondido ao e-mail rs, rs, rs), mas enfim, estava Prof. Ricardo em São Paulo e respondeu ao e-mail, já marcando uma reunião, começou o trabalho, começaram os números a ter cara de dissertação, os resultados foram surgindo, foi aumentando a quantidade de páginas e estou agora, escrevendo a última página. Graças a Deus que existem o Prof. Gil e Prof. Gustavo para socorrer e Prof. Ricardo para iluminar. A vocês meus orientadores, que fizeram parte da minha formação, o meu "Muito obrigada", minha gratidão com vocês será eterna!

Mas a pesquisa precisa de algo que não sei onde encontrar. Manda-se e-mail para Prof. Rômulo, que o encaminha para um grupo de pesquisadores do IPA e orientandos dele, surgem Érik, Dário, Júlio, Keneddy e Tiago, vão-se à procura de valores de referência, como indicador de fertilidade do solo, pede-se livros emprestados a Keneddy UFCG, procuro, leio e nada, não encontro nada. Bate desespero! O grupo consegue! Ufa! Com dificuldade, mas conseguiu! Sem esses valores, seria impossível continuar a pesquisa. Muito obrigada a vocês!

Aos professores Eufrázio, Tatijana e Borko (UFRPE), Everardo (UFPE) e Rosângela (UFCG), o conhecimento transmitidos ficarão para sempre guardados na memória. Aos funcionários do DEINFO em especial ao secretário Marco, pela documentação exigida durante o período de curso, pelas dicas de qual ônibus pegar para resolver as coisas em Recife. A Zuleide, pelas conversas. Aos funcionários do DEN-UFPE, em especial ao porteiro, ao vigilante e a "nossa alegria", pelas inúmeras conversas descontraídas, enquanto aguardava o prof. Rômulo. À Capes, pela bolsa de estudos.

À Deus, por toda essa trajetória! Só tenho a agradecer!

Marystella Duarte Correia.

"Não fiz o melhor, mas fiz tudo para que o melhor fosse feito."

Martin Luther King

"Nenhum obstáculo é tão grande se a sua vontade de vencer for maior."

Autor desconhecido

Resumo

Estudos de algumas regiões semiáridas têm apontado o efeito benéfico de certas espécies de árvores em sistemas silvopastoris, por promoverem a formação de fertilidade do solo e aumentarem a sustentabilidade da produtividade agrícola. Neste trabalho, foram selecionadas três espécies arbóreas encontradas em pastagens de Capim buffel (*Cenchrus ciliaris*), juazeiro (*Zyziphus joazeiro*), umbuzeiro (*Spondias tuberosa*) e algaroba (*Prosopis juliflora*) da caatinga, bioma exclusivamente brasileiro, concentrado na região Nordeste do Brasil. O semivariograma é a parte central dos estudos geoestatísticos, capaz de descrever a variação espacial, além de ser o ponto chave na interpolação dos dados por Krigagem. Daí a importância do ajuste e seleção dos modelos. O presente estudo foi conduzido em uma fazenda experimental silvopastoril em Custódia PE, onde foram avaliadas as diferenças nas características químicas do solo e do sub-bosque herbáceo, entre áreas embaixo e fora da copa destas árvores. Amostras de solo (0-15 cm) e do estrato herbáceo foram coletadas, para avaliação da conservação de carbono e biomassa de herbáceas, bem como a dependência espacial destas variáveis. As espécies arbóreas Algarobas foram incluídas nos campos cultivados, o juazeiro e umbuzeiro já estavam incluídas no local, a fim de explicar se as espécies arbóreas ou o Capim buffel consegue manter e/ou preservar o carbono no solo, verificando também o crescimento do capim na região de estudos, classificando-se a dependência espacial e mapeando-se os dados observados com suas probabilidades.

Palavras-chave: Semivariograma; Krigagem; Fertilidade do solo.

Abstract

Studies of some semiarid regions have pointed the beneficial effect of certain tree species in silvopastoral systems, by promoting the formation of soil fertility and increase the sustainability of agricultural productivity. In this study, we selected three tree species found in pastures of Grass buffel *Cenchrus ciliaris*, juazeiro *Zyziphus joazeiro*, umbuzeiro *Spondias tuberosa* and algaroba *Prosopis juliflora* caatinga, exclusively Brazilian biome, concentrated in the Northeast region of Brazil. The semivariogram is the central part of geostatistical studies, able to describe the spatial variation, besides being the key point in the data interpolation by kriging. Hence the importance of setting and selection of models. This study was conducted on an experimental farm in Custody silvopastoral PE, which we evaluated the differences in the chemical characteristics of the soil and herbaceous understory, between areas under and outside the canopy of these trees. Soil samples (0-15 cm) and the herbaceous layer were collected for evaluation of carbon storage and biomass of herbaceous and spatial dependence of these variables. Tree species were included in Algarobas cultivated fields, and the juazeiro umbuzeiro were already included on the site in order to explain whether the tree species or buffel grass can maintain and / or preserve soil carbon, also checking the growth of grass in region studies, classifying and mapping the spatial dependence is observed data with their probabilities.

Keywords: Semivariogram, kriging, Soil Fertility.

Lista de Figuras

| | | |
|----|---|-------|
| 1 | Amostragem em duas dimensões | p. 19 |
| 2 | Parâmetros do semivariograma | p. 21 |
| 3 | Gráfico descritivo do padrão espacial do juazeiro, variável carbono, com os dados originais. | p. 43 |
| 4 | Gráfico de possíveis transformações de variáveis por meio da transformação ótima de <i>Box-Cox</i> (Esquerdo) e a verificação de dependência espacial por meio do envelope simulado (Direito) da variável carbono da espécie juazeiro. | p. 43 |
| 5 | Gráfico descritivo do padrão espacial do juazeiro, variável carbono, com os dados transformados. | p. 44 |
| 6 | Modelo ajustado ao semivariograma conforme o método da máxima verossimilhança para a variável carbono da espécie juazeiro. | p. 45 |
| 7 | Comparação dos valores observados com o valor de referência para solos com baixo teor de carbono, na espécie juazeiro. | p. 46 |
| 8 | Comparação dos valores observados com o valor de referência para solos com médio teor de carbono, na espécie juazeiro. | p. 46 |
| 9 | Comparação dos valores observados com o valor de referência para solos com alto teor de carbono, na espécie juazeiro. | p. 47 |
| 10 | Gráfico descritivo do padrão espacial do juazeiro, variável biomassa de herbáceas, com os dados originais. | p. 48 |
| 11 | Gráfico de possíveis transformações de variáveis por meio da transformação ótima de <i>Box-Cox</i> (Esquerdo) e a verificação de dependência espacial através do envelope simulado (Direito) da variável biomassa de herbáceas da espécie juazeiro. | p. 49 |
| 12 | Gráfico descritivo do padrão espacial do juazeiro, variável biomassa de herbáceas, com os dados transformados. | p. 50 |

| | | |
|----|--|-------|
| 13 | Modelo ajustado ao semivariograma conforme o Método de Máxima Verossimilhança para a variável biomassa de herbáceas da espécie juazeiro. | p. 51 |
| 14 | Comparação dos valores observados com a média do valor de referência com efeito de crescimento de pasto em área com juazeiro. | p. 52 |
| 15 | Gráfico descritivo do padrão espacial do umbuzeiro, variável carbono, com os dados originais. | p. 53 |
| 16 | Gráfico de possíveis transformações de variáveis por meio da transformação ótima de <i>Box-Cox</i> (Esquerdo) e a verificação de dependência espacial através do envelope simulado (Direito) da variável carbono da espécie Umbuzeiro. | p. 54 |
| 17 | Gráfico descritivo do padrão espacial do umbuzeiro, variável carbono, com os dados transformados. | p. 55 |
| 18 | Modelo ajustado ao semivariograma conforme o método de máxima verossimilhança para a variável carbono da espécie umbuzeiro. | p. 56 |
| 19 | Comparação dos valores observados com o valor de referência para solos com baixo teor de carbono, na espécie umbuzeiro. | p. 57 |
| 20 | Comparação dos valores observados com o valor de referência para solos com médio teor de carbono, na espécie umbuzeiro. | p. 58 |
| 21 | Comparação dos valores observados com o valor de referência para solos com alto teor de carbono, na espécie umbuzeiro. | p. 59 |
| 22 | Gráfico descritivo do padrão espacial do umbuzeiro, variável herbáceas, com os dados originais. | p. 60 |
| 23 | Gráfico de possíveis transformações de variáveis por meio da transformação ótima de <i>Box-Cox</i> (Esquerdo) e a verificação de dependência espacial através do envelope simulado (Direito) da variável biomassa de herbáceas da espécie umbuzeiro. | p. 61 |
| 24 | Gráfico descritivo do padrão espacial do umbuzeiro, variável biomassa de herbáceas, com os dados transformados. | p. 62 |
| 25 | Modelo ajustado ao semivariograma conforme o método da máxima verossimilhança para a variável biomassa de herbáceas da espécie umbuzeiro. | p. 63 |

| | | |
|----|--|-------|
| 26 | Comparação dos valores observados com a média do valor de referência com efeito de crescimento de pasto em área com umbuzeiro. | p. 64 |
| 27 | Gráfico descritivo do padrão espacial do algaroba, variável carbono, com os dados originais. | p. 65 |
| 28 | Gráfico de possíveis transformações de variáveis por meio da transformação ótima de <i>Box-Cox</i> (Esquerdo) e a verificação de dependência espacial através do envelope simulado (Direito) da variável carbono da espécie algaroba. | p. 66 |
| 29 | Gráfico descritivo do padrão espacial do algaroba, variável carbono, com os dados transformados. | p. 67 |
| 30 | Modelo ajustado ao semivariograma conforme o método da máxima verossimilhança para a variável carbono da espécie algaroba. | p. 68 |
| 31 | Comparação dos valores observados com o valor de referência para solos com médio teor de carbono, na espécie Algaroba. | p. 69 |
| 32 | Comparação dos valores observados com o valor de referência para solos com médio teor de carbono, na espécie Algaroba. | p. 70 |
| 33 | Comparação dos valores observados com o valor de referência para solos com médio teor de carbono, na espécie algaroba. | p. 71 |
| 34 | Gráfico descritivo do padrão espacial do algaroba, variável biomassa de herbáceas, com os dados originais. | p. 72 |
| 35 | Gráfico de possíveis transformações de variáveis por meio da transformação ótima de <i>Box-Cox</i> (Esquerdo) e a verificação de dependência espacial através do envelope simulado (Direito) da variável biomassa de herbáceas da espécie algaroba | p. 72 |
| 36 | Gráfico descritivo do padrão espacial do algaroba, variável biomassa de herbáceas, com os dados transformados. | p. 73 |
| 37 | Modelo ajustado ao semivariograma conforme o método da máxima verossimilhança para a variável biomassa de herbáceas da espécie algaroba. | p. 73 |
| 38 | Comparação dos valores observados com a média do valor de referência com efeito de crescimento de pasto em área com algaroba | p. 74 |

Lista de Tabelas

- 1 Estatística descritiva: mínimo, máximo, média, mediana, desvio padrão (DP), coeficiente de variação (CV), curtose, assimetria de amostras de solo da variável carbono da espécie juazeiro. p. 42
- 2 Estimativa dos parâmetros associados aos modelos por meio da máxima verossimilhança, assumindo a média da variável carbono constante sobre a região de estudo na espécie juazeiro. p. 45
- 3 Estatística descritiva: mínimo, máximo, média, mediana, desvio padrão (DP), coeficiente de variação (CV), curtose, assimetria de amostras de solo da variável biomassa de herbáceas da espécie juazeiro. p. 47
- 4 Estimativa dos parâmetros associados aos modelos por meio da máxima verossimilhança, assumindo-se a média da variável biomassa de herbáceas constante sobre a região de estudo na espécie juazeiro. p. 49
- 5 Estatística descritiva: mínimo, máximo, média, mediana, desvio padrão (DP), coeficiente de variação (CV), curtose, assimetria de amostras de solo da variável carbono da espécie umbuzeiro. p. 53
- 6 Estimativa dos parâmetros associado aos modelos por meio da máxima verossimilhança, assumindo-se a média da variável carbono constante sobre a região de estudo na espécie umbuzeiro. p. 55
- 7 Estatística descritiva: mínimo, máximo, média, mediana, desvio padrão (DP), coeficiente de variação (CV), curtose, assimetria de amostras de solo da variável biomassa de herbáceas da espécie umbuzeiro. p. 57
- 8 Estimativa dos parâmetros associado aos modelos por meio da máxima verossimilhança, assumindo-se a média como um polinômio de primeira ordem sobre as coordenadas (x, y). p. 59
- 9 Estatística descritiva: mínimo, máximo, média, mediana, desvio padrão (DP), coeficiente de variação (CV), curtose, assimetria de amostras de solo da variável carbono da espécie algaroba. p. 62

- 10 Estimativa dos parâmetros associado aos modelos por meio da máxima verossimilhança, assumindo a média da variável carbono constante sobre a região de estudo na espécie algaroba. p. 64
- 11 Estatística descritiva: mínimo, máximo, média, mediana, desvio padrão (DP), coeficiente de variação (CV), curtose, assimetria de amostras de solo da variável biomassa de herbáceas da espécie algaroba. p. 67
- 12 Estimativa dos parâmetros associado aos modelos por meio da máxima verossimilhança, assumindo a média da variável biomassa de herbáceas constante sobre a região de estudo na espécie algaroba. p. 70

Sumário

| | | |
|----------|---|-------|
| 1 | Introdução | p. 1 |
| 2 | Fundamentação Teórica | p. 4 |
| 2.1 | Bioma caatinga | p. 4 |
| 2.2 | Marco histórico | p. 9 |
| 2.3 | Estatística Espacial | p. 10 |
| 2.3.1 | Análise de Pontos | p. 11 |
| 2.3.2 | Análise Espaço Temporal | p. 11 |
| 2.3.3 | Análise de Superfícies | p. 11 |
| 2.4 | Geoestatística | p. 12 |
| 2.5 | Campos aleatórios | p. 13 |
| 2.5.1 | Variograma e Semivariograma | p. 18 |
| 2.5.2 | Parâmetros do semivariograma | p. 21 |
| 2.5.3 | Modelos geoestatísticos gaussianos | p. 22 |
| 2.5.4 | Índice de dependência espacial | p. 23 |
| 2.5.5 | Função de máxima verossimilhança | p. 24 |
| 2.5.6 | Funções de correlações | p. 26 |
| 2.5.7 | Critério de Informação de Akaike - (<i>AIC</i>) | p. 27 |
| 2.5.8 | Critério de Informação Bayesiana - (<i>BIC</i>) | p. 28 |
| 2.5.9 | Krigagem | p. 28 |
| 2.6 | Espécies vegetais estudadas | p. 33 |
| 3 | Material e métodos | p. 39 |

| | | |
|----------|--|-------|
| 3.1 | Área de estudo | p. 39 |
| 3.2 | Software | p. 40 |
| 3.3 | Valores de referência - Teor do solo | p. 41 |
| 4 | Resultados e discussão | p. 42 |
| 4.1 | Juazeiro | p. 42 |
| 4.2 | Umbuzeiro | p. 52 |
| 4.3 | Algaroba | p. 61 |
| 5 | Considerações finais | p. 75 |
| | Referências | p. 76 |
| | Apêndice | p. 80 |

1 Introdução

A caatinga é um bioma exclusivamente brasileiro, que ocorre na região semiárida, em grande parte localizada na região Nordeste do País. Apresenta grande diversidade de ambientes, o que propicia uma rica biodiversidade apresentando muitas espécies endêmicas de alto valor biológico Brasil (2005), muitas ainda desconhecidas e/ou, não catalogadas (ALVES *et al.*, 2009). Segundo Oliveira *et al.* (2009), em seu aspecto fisionômico a caatinga apresenta uma cobertura vegetal arbustiva a arbórea, pouco densa e geralmente espinhosa. Sua variabilidade espacial na composição e no arranjo de seus componentes botânicos é resposta aos processos de sucessão e de diversos fatores ambientais, onde a densidade de plantas, a composição florística e o potencial do estrato herbáceo variam em função das características de solo (ARAÚJO, 1986).

A modelagem estatística é utilizada em diversos campos do conhecimento para tentar descrever o comportamento de um ou mais atributos que não podem ser descritos exclusivamente por modelos determinísticos. De forma geral, os modelos estatísticos tentam explicar, o máximo possível, a variabilidade dos processos estocásticos através de uma ou mais variáveis explanatórias que possuam alguma associação com a resposta de interesse Fonseca (2008), traz resultados diferentes daqueles obtidos pela estatística clássica, sendo os primeiros geralmente mais robustos por incorporarem a dimensão espacial. Para sua análise são necessárias pelo menos as informações sobre a localização e os atributos, que são valores associados aos dados independentemente da forma como sejam medidos, e parte-se do pressuposto que os dados são espacialmente dependentes (KREMPI, 2004).

A variabilidade espacial de algumas características do solo vem sendo uma das preocupações de pesquisadores desde o início do século XX, a análise da variabilidade de atributos físicos do solo pode ser realizada por meio da estatística descritiva, no início do século se utilizavam grandes quantidades de dados amostrais. Esta ferramenta, entretanto, não considera a distribuição dos dados no espaço, ao contrário da estatística espacial, que considera a dependência espacial entre as amostras e a sua localização geográfica (VIEIRA, 2000).

Segundo Vieira (2000), numa análise geoestatística, os valores de alcance dos semi-variogramas experimentais informam as distâncias mínimas entre as amostras, para assim serem consideradas dependentes. A independência ou não dos pontos, é fundamental para a escolha do teste estatístico.

No estudo geoestatístico dois aspectos são apresentados quando se avalia a continuidade espacial, o primeiro, fundamenta-se na modelagem da dependência espacial considerando as pressuposições da distribuição gaussiana e determinando estimativas paramétricas através dos estimadores de máxima verossimilhança. O segundo é o não paramétrico, que independe de métodos rígidos, e que baseiam-se em tratamentos de dados ordenados, de tal forma que cada valor é representado por um número que caracteriza a sua posição na sequência. Estas duas linhas de estudos, utilizam-se do semivariograma para identificar a dependência espacial e do interpolador *krigagem*, que possibilita a partir de observações pontuais informações para grande extensões de terra, baseados nas observações da variável a ser estimada em pontos não amostrados (OLIVEIRA, 2003).

Os atributos físicos e químicos do solo influenciam diretamente e indiretamente o crescimento e o desenvolvimento das plantas. Por esse motivo a avaliação da variabilidade espacial destes atributos tem-se tornado importante ferramenta na determinação de estratégias de manejo do solo.

O sistema silvopastoril é a combinação intencional de árvores, pastagem e gado numa mesma área ao mesmo tempo e manejados de forma integrada, com o objetivo de incrementar a produtividade por unidade de área. Nesses sistemas, ocorrem interações em todos os sentidos e em diferentes magnitudes, apresentam grande potencial de benefícios econômicos e ambientais para os produtores e para a sociedade. São sistemas multifuncionais, onde existe a possibilidade de intensificar a produção pelo manejo integrado dos recursos naturais evitando sua degradação, além de recuperar sua capacidade produtiva (SILVA, 2004).

Portanto, uma proposta deste trabalho foi utilizar os estimadores de máxima verossimilhança, tomando-se como valores iniciais os parâmetros do semivariograma, para avaliar a estrutura de dependência espacial em variáveis de fertilidade do solo. O presente trabalho foi desenvolvido com o objetivo específico de explorar e avaliar a união de ferramentas da estatística espacial, com o propósito de analisar as relações entre os aspectos da distribuição espacial de carbono orgânico e da biomassa de herbáceas em sistema silvopastoril, a fim de explicar se a espécie arborea ou o Capim buffel consegue manter e/ou preservar o carbono no solo, verificando também o crescimento do capim na região de estudos,

classificando-se a dependência espacial e mapeando-se os dados observados com suas probabilidades.

2 Fundamentação Teórica

2.1 Bioma caatinga

A caatinga ocupa oficialmente mais de 800.000 km^2 do território brasileiro. Estende-se pela totalidade do estado do Ceará (100%) e mais de metade da Bahia (54%), da Paraíba (92%), de Pernambuco (83%), do Piauí (63%) e do Rio Grande do Norte (95%), quase metade de Alagoas (48%) e Sergipe (49%), além de pequenas porções em Minas Gerais (2%) e do Maranhão (1%). A caatinga é muito rico em biodiversidade, tanto vegetal quanto animal. Nos períodos sem chuva, cerca de 8 meses por ano, ela "adormece" e suas folhas caem. Depois, com as primeiras chuvas, ela rebrota rapidamente. Cerca de 28 milhões de brasileiros habitam esse bioma, sendo que aproximadamente 38% vivem no meio rural (CRUZ, 2010).

A região do bioma caatinga tem sido utilizada pelo homem há vários séculos, mas ainda são pouco conhecidas as riquezas ambientais da região. Milhares de hectares da vegetação nativa estão agora ameaçados de extinção. O crescente desmatamento surge para produção de lenha, carvão e para o uso do solo, para a pecuária e agricultura. Geralmente esses desmatamentos ocorrem de forma indiscriminada e desordenada, que aliada às adversidades climáticas, provocam desequilíbrios ambientais. Apesar disso, o bioma caatinga ainda possui aproximadamente 40 a 50% do seu território com cobertura vegetal (LINHARES; GEWANDSZBAJDER, 1998).

Conforme Linhares e Gewandszajder (1998), nas regiões de caatinga, o clima é quente com prolongadas estações secas e o regime de chuvas influencia na vida de animais e vegetais. A diversidade de espécies é menor, comparado a outros biomas como a mata atlântica e a amazônia. Estudos recentes revelam um alto número de espécies endêmicas, isto é, espécies que só ocorrem naquela região. A vegetação se caracteriza por arbustos tortuosos, com aspecto seco e esbranquiçado por quase todo ano. O manejo da caatinga consiste numa atividade potencial para a região, sobretudo, torna-se uma necessidade ecológica, econômica e social.

Fauna

Segundo Linhares e Gewandszajder (1998), a maioria dos animais do bioma tem hábitos noturno, o que evita que se exponham em horas mais quentes. Os lagartos são muito comuns: 47 espécies deles já foram catalogadas. Entre elas estão o calango verde (*Cnemidophorus ocellifer*) e o calanguinho (*Tropidurus torquatus*). Entre os répteis destacam-se as serpentes. Até agora foram encontradas 45 espécies de serpentes, a cascavel (*Crotalus durissus*) é uma das mais vistas na caatinga.

Algumas aves são típicas da caatinga, tais como o carcará (*Polyborus plancus*), a asa-branca (*Dendrocygna autumnalis*) e a gralha-canção (*Cyanocorax cyanopogon*). Neste bioma, vivia a ararinha azul (*Cyanopsitta spixii*), vista pela última vez na natureza em 2000 e considerada extinta pelo Ibama. Outra ave em estado de conservação crítico é a arara-azul-de-Lear (*Anodorhynchus leari*), encontrada apenas em uma pequena área, nos municípios de Canudos, Euclides da Cunha e Jeremoabo, no interior da Bahia. Ameaçada pela perda do hábitat e captura para exportação, ela vive nas palmeiras licuri (*Syagrus coronata*), cujos frutos são seu principal alimento, e faz seus ninhos em cavidades nos paredões de arenito.

Os anfíbios são animais numerosos na caatinga, dos mais conhecidos, pode-se citar o sapo cururu (*Rhinella jimi*) e a gia (*Leptodactylus ocellatus*). Entre as árvores secas e em terrenos pedregosos, vivem onças (*Panthera onca*), gatos selvagens (*Felis silvestris catus*), capivaras (*Hydrochoerus hydrochoeris*), gambás (*Didelphis marsupialis*), preás (*Cavia aperea*), macacos-prego (*Cebus apella*), e o veado catingueiro (*Mazama gouazoubira*), também ameaçado de extinção.

Vegetação

Quando fala-se em caatinga sempre vem a imagem de um ambiente árido, seco, com árvores quase sem folhas e esbranquiçadas. Realmente é assim que a vegetação da caatinga se apresenta em grande parte do ano. Existem várias espécies de plantas no bioma, algumas só existem na caatinga. Se a devastação continuar, talvez em um futuro breve, nem chegaremos a conhecer e estudar o bioma, por isso o grande interesse em preservar, pois o mesmo é importante para o país, principalmente para o Nordeste (LINHARES; GEWANDSZAJDER, 1998).

Conforme Linhares e Gewandszajder (1998), em época de chuvas, a caatinga muda seu aspecto, a paisagem fica verde e aparecem flores. A caatinga apresenta três estratos: arbóreo, arbustivo e o herbáceo. A flora dos sertões é constituída por espécies com longa história de adaptação ao calor e à seca. A vegetação é composta por plantas xerófitas.

Espécies que acabaram desenvolvendo mecanismos para sobreviverem em um ambiente com poucas chuvas e baixa umidade. Espinhos estão presentes em muitas espécies vegetais. Nos cactos, por exemplo, eles são folhas que se modificaram ao longo da evolução, fazendo com que a perda de água pela transpiração seja menor.

Segundo Linhares e Gewandszajder (1998), algumas plantas simplesmente perdem suas folhas na estação seca para evitar a perda de água. Por isso, parece que toda a vegetação está morta, sem folhas, sem verde, só caules e troncos secos e retorcidos. Na verdade, as plantas permanecem vivas, utilizando-se, por exemplo, suas raízes bem desenvolvidas para obter água armazenada no solo. Outras espécies como o mandacaru (*Cereus jamacaru*), a coroa-de-frade (*Melocactus zehntneri*), o xique-xique (*Pilosocereus gounellei*), o juazeiro (*Zizyphus Joazeiro*), o umbuzeiro (*Spondias tuberosa Arruda*) e a aroeira (*Schinus molle L.*), desenvolvem raízes na superfície, o que lhes permitem, no período das chuvas, absorver o máximo possível da água que cai sobre os terrenos, elas mesmas armazenam água, é o caso dos cactos, em que são muito representativos da vegetação da caatinga. Mas não são os únicos. A criação e a ampliação de unidades de conservação é outro ponto indicado como fator importante para a proteção do bioma. A caatinga tem apenas 7% de áreas protegidas, sendo que 2% são de proteção integral e os outros 5% são de unidades de conservação de uso sustentável.

Solo

Conforme Linhares e Gewandszajder (1998), de forma geral, o solo é rico em minerais, mas pobre em matéria orgânica, já que a decomposição desta matéria é acelerada pelo calor e pela luminosidade intensa durante todo ano. Fragmentos de rochas são frequentes na superfície, o que dá ao solo um aspecto pedregoso. Com muitas variedades e tamanhos de pedras, dificilmente o solo consegue armazenar a água que cai no período das chuvas.

A presença de minerais no solo da caatinga é garantia de fertilidade em um ambiente que sofre com a falta de chuvas. Por isso, nos poucos meses em que a chuva cai, algumas regiões secas, rapidamente se transformam, dando espaço a árvores verdes e algumas gramíneas.

Relevo

Segundo Linhares e Gewandszajder (1998), o relevo da caatinga apresenta duas formações dominantes: os planaltos e as grandes depressões. É comum existirem fragmentos de rochas na superfície do solo. Esses fragmentos são encontrados com mais frequência nos estados da Paraíba, Pernambuco, Rio Grande do Norte e Alagoas. O pla-

nalto da Borborema é uma formação que se destaca com altitudes variando em média, entre 650 e 1000 metros. Em alguns pontos, esta marca é ultrapassada: o pico de Jabre, na Paraíba, chega a 1.197 metros e o pico do Papagaio, em Pernambuco, a 1.260 metros. O planalto é uma grande barreira para as nuvens carregadas de umidade que vêm do oceano Atlântico em direção ao interior. Quando essas nuvens encontram este "paredão", elas se condensam, provocando chuvas nas regiões mais baixas do lado oriental do planalto, ou seja, o lado voltado para o oceano. Isto dificulta a ocorrência de chuvas do lado ocidental, que é marcado pela seca. Este lado seco é o que compõe a maior parte da área do bioma caatinga.

Água

Os rios que fazem parte da caatinga brasileira são, em grande maioria temporários. Neste bioma, onde há escassez de chuva durante maior parte do ano, os rios que nascem na região ficam secos por longos períodos. Os que nascem em outros lugares, como o São Francisco e o Parnaíba, são fundamentais para a vida na caatinga, pois atravessam os terrenos quentes e secos. Para enfrentar a falta de água nas estações secas, os moradores constroem poços, cacimbas e açudes. Na maior parte das vezes, só conseguem obter água salobra, imprópria para consumo (LINHARES; GEWANDSZBAJDER, 1998).

A situação do semiárido brasileiro, tende a se agravar com o aquecimento global, que terá reflexos imediatos sobre a disponibilidade hídrica. Os impactos poderão ser muito graves. Uma das estimativas é a diminuição em aproximadamente 20% do volume de chuvas que cai sobre a região todos os anos, hoje a média anual está em torno de 750 mm, a redução ainda não transforma o semiárido em deserto, uma redução nesse índice significa que pode-se conhecer, é uma diminuição muito significativa para uma região que apresenta problemas sérios para acumular água em lençóis subterrâneos, por causa da presença de rochas cristalinas em 70% da área. Só em 30% dela, como na região do aquífero Gurguéia (PI), há capacidade de armazenamento de águas subterrâneas (LYRA *et al.*, 2009).

O armazenamento dessa água em mananciais de superfície, como foi à opção até agora, principalmente as grandes barragens, sempre apresentou o inconveniente da enorme evaporação, o problema do semiárido é mais a perda por evaporação, também por transpiração de plantas e animais do que a falta de precipitação. Calculam-se que haja 3 mm de evaporação para cada 1 mm de precipitação. A eventual elevação da temperatura aumentará essa diferença (LYRA *et al.*, 2009).

Conforme Lyra *et al.* (2009), o futuro do semiárido passará pelo aumento da capta-

ção da água de chuva em reservatórios fechados. As cisternas, que desempenham um importante papel na vida dos sertanejos, têm a capacidade de armazenar águas pluviais, fazendo com que essa água armazenada tenha uma função bem determinada dentro do cotidiano das pessoas.

Clima

O clima da caatinga é chamado de semiárido. São características desse bioma: o clima quente, a baixa umidade e o pouco volume pluviométrico. Os períodos de seca, podem chegar a oito ou nove meses por ano. Este clima irregular influencia o curso dos rios, que secam em determinadas épocas, diminuindo a disponibilidade de água para plantas, animais e para os moradores do bioma, aumentando assim a aridez. O clima é então um fator determinante na caatinga, ele acaba definindo a paisagem e os hábitos dos seres que sobrevivem neste bioma (LINHARES; GEWANDSZBAJDER, 1998).

A caatinga localizada em área de clima semiárido, apresenta temperaturas médias anuais que oscilam entre 25°C e 29°C e com médias pluviométricas inferiores aos 800 mm, já que são características desse tipo de clima a baixa umidade e a quantidade reduzida de chuvas. As chuvas ocorrem no início do ano e o poder de recuperação do bioma é muito rápido, surgem pequenas plantas e as árvores ficam cobertas de folhas. A rigidez climática das caatingas é conferida principalmente pela irregularidade na distribuição destas chuvas no tempo e no espaço (MECABO *et al.*, 2012).

A paisagem mais comum da caatinga é a que ela apresenta durante a seca. O solo, raso e pedregoso, é composto por vários tipos diferentes de rochas. Na estação seca a temperatura do solo pode chegar até 60°C. Segundo Mecabo *et al.* (2012), as variações em temperatura são muito menos extremas durante a estação chuvosa, e também durante certos períodos quando a neblina se forma, especialmente à noite nas áreas de maior altitude, durante a estação seca. Este clima irregular influencia o curso dos rios, que secam em determinadas épocas; diminuindo a disponibilidade de água para plantas, animais e para os homens; aumenta a aridez do ambiente. As secas são cíclicas e prolongadas, interferindo de maneira direta na vida de uma população de, aproximadamente, 25 milhões de habitantes, que convivem com os longos períodos de estiagem e irregularidade climática.

Em razão da semiáridéz e do predomínio de rios temporários, era de se esperar que a biota aquática da caatinga fosse pouco diversificada, mas a caatinga apresenta uma fauna e uma flora bastante rica.

2.2 Marco histórico

Conforme Vieira (2000), a variabilidade em fenômenos espaciais vem sendo uma das preocupações de pesquisadores, praticamente desde o início do século XX. Em 1910 Smith estudou a disposição de parcelas no campo em experimentos de rendimento de variedades de milho, numa tentativa de eliminar o efeito de variações do solo. No ano de 1913 Montgomery preocupado com o efeito do nitrogênio no rendimento do trigo, fez um experimento em 224 parcelas, medindo o rendimento de grãos. Vários outros autores em 1915 como Robinson e Lloyd e em 1918 Waynick estudou a variabilidade espacial de nitrificação no solo, em 1919 Waynick e Sharp estudaram variações de nitrogênio total e carbono orgânico no solo.

Segundo Landim (2003), as raízes da Geoestatística estão na indústria de minérios, na década de 50, quando o engenheiro de minas Daniel Krige (1951) concluiu que a variância dos dados possuía uma estruturação que dependia da amostragem, ele não conseguia encontrar sentido nas variâncias, se não levasse em conta a distância entre as amostras. Conforme Deutsch (2002), o estatístico Sichel (1992), juntamente com Daniel Krige desenvolveram na década de 1950, novos métodos de estimação para reservas minerais espalhadas. Entre 1957 e 1962, o estatístico Matheron (1963), baseado nestas observações, desenvolveu uma teoria, a qual ele chamou de Teoria das Variáveis Regionalizadas, que posteriormente receberia do Centre de Morphologie Mathematique em Fontaineblau, França (1962 e 1963) o nome de Geoestatística.

Entre 1968 e 1970, foi desenvolvida a Teoria da Krigagem Universal, este método foi desenvolvido por Matheron na década de 60, e recebeu este nome, em homenagem ao engenheiro de minas sul-africano Daniel G. Krige, que primeiro desenvolveu e implementou esta fórmula de inferência em 1951 (MELLO, 2004), maiores detalhes na subseção 2.5.9. Em 1972, Matheron criou a teoria Intrínseca de Ordem K , aplicada à meteorologia. Entre 1972 e 1973 surgiram os princípios da Análise Convexa, visando-se maximizar as reservas recuperáveis das jazidas subterrâneas. Em 1974 nasceu a teoria das funções de recuperação e, baseada nela, a Geoestatística não-linear aplicada na seleção de reservas recuperáveis (ROCHA, 2005).

Segundo Druck *et al.* (2004), um conceito chave na compreensão e análise dos fenômenos espaciais que é a dependência espacial foi criada por TOBLER (1970), a qual foi considerada primeira lei da geografia, onde diz que "*todas as coisas são parecidas, no entanto, coisas mais próximas se parecem mais que coisas mais distantes*". Independente do tipo de ocorrência, sejam naturais ou não, apresentam entre si uma relação que depende

da distância.

Conforme Landim (2003), a escola norte-americana se apoiava essencialmente na Estatística clássica e variáveis independentes, a escola sul-africana admitia a existência de correlações espaciais e também a influência dos tamanhos das amostras, já a escola francesa, uniu as duas anteriores e tentou corrigir seus problemas.

Segundo Rocha (2005), uma, dentre as muitas vantagens da aplicação da Geoestatística, é o fato de ela necessitar e incentivar a interdisciplinaridade, assegurando uma maior troca de informações entre geólogos, engenheiros, matemáticos, agrônomos e estatísticos e uma melhor interpretação da realidade geológica em estudo. De início, a aplicação era apenas para as situações em geologia mineira, estendeu-se a várias áreas nesses últimos anos com aplicação em climatologia, geologia ambiental, geotecnia, hidrogeologia, pedologia, entre outros (LANDIM, 2003).

2.3 Estatística Espacial

A análise espacial objetiva mensurar propriedades e relacionamentos considerando a localização espacial do fenômeno em estudo, ou seja, permite estudar, explorar e modelar fenômenos geográficos, é composta por um conjunto de procedimentos encadeados cuja finalidade é a escolha de um modelo inferencial que considera explicitamente os relacionamentos espaciais presentes no fenômeno (DRUCK *et al.*, 2004).

A aplicação da estatística espacial era apenas para situações em geologia mineira na lavra e prospecção, mas posteriormente se estendeu para outros campos, especialmente nesses últimos anos, com aplicações em climatologia, geologia ambiental, geotecnia, hidrogeologia, pedologia, ganhando impulso em áreas distintas da mineração e da geologia a partir de 1980, com grande aplicabilidade na ciência do solo. No Brasil destaca-se alguns trabalhos pioneiros desenvolvidos pelos pesquisadores Sidney Rosa Vieira, Paulo Libardi e Klaus Reichardt, todos realizados na década de 80 (GUIMARÃES, 2004). Atualmente a aplicação e a utilização dessa metodologia de análise de dados está difundida em vários ramos das ciências, entre elas estão as humanas, as biológicas e as ciências exatas.

Conforme Druck *et al.* (2004), entende-se por dependência espacial, o fato de que a maior parte das ocorrências naturais ou sociais apresentam entre si uma relação que depende da distância. A ideia é verificar como a dependência espacial varia, a partir da comparação entre os valores de uma amostra e de seus vizinhos, é formada por três grandes áreas de estudo: geoestatística, dados de área e processos pontuais, que são

utilizadas conforme o tipo de dados em questão (FONSECA, 2008).

2.3.1 Análise de Pontos

Processos pontuais são fenômenos expressos por meio de ocorrência identificadas como pontos localizados no espaço. São exemplos desse tipo de dados a localização da ocorrência de casos de doenças e a localização de indivíduos de uma determinada espécie (OLINDA, 2008).

2.3.2 Análise Espaço Temporal

O modelo espaço-temporal pode ser aplicado de maneira tradicional, estimando-se os parâmetros por mínimos quadrados, ou seja, pelos procedimentos usuais de regressão múltipla (SZWARCOWALD *et al.*, 2001). É importante notar que este modelo é apropriado para aplicação em dados de áreas onde as ponderações espaciais podem ser definidas com base em diferentes critérios de vizinhança, como o da contiguidade ou distância entre áreas geográficas. Conforme Olinda (2008), são fenômenos associados aos dados de levantamentos populacionais, como censos, e que, originariamente, referem-se a indivíduos localizados em pontos específicos no espaço. Normalmente, esses pontos são agregados em unidades de análises, usualmente delimitadas por polígonos fechados, tais como: setores censitários, municípios e microrregiões.

2.3.3 Análise de Superfícies

Análise de superfície é gerada a partir de um processo de interpolação dos dados pontuais. O objetivo desta análise é reconstruir a superfície na qual as amostras foram retiradas. Neste caso, tem-se interesse em padrões nos valores dos atributos, e não mais nos padrões das localizações das observações (BAILEY; GATRELL, 1995).

Conforme Camargo (2002), estes dados estão disponíveis na forma de amostras pontuais, e para utilizá-los de forma efetiva em um ambiente de Geoprocessamento, necessita-se de um procedimento de interpolação, para gerar uma representação na forma de grade regular. As amostras são valores representativos do fenômeno estudado, usualmente obtidas a partir de levantamento de campo, e que apresentam consistência de metodologia e unidade, essas amostras podem representar tanto variáveis naturais (como teor de argila no solo), como também socioeconômicas (taxa de homicídios).

As superfícies contínuas podem ser estimadas a partir de um conjunto de amostras de campo, que podem estar regularmente ou irregularmente distribuídas, permitindo-se que dados disponíveis sob a forma pontual sejam interpolados gerando-se uma superfície. Este tipo de dado espacialmente contínuo é geralmente referenciado como dado geoestatístico.

A modelagem de tendências ou variação em larga escala se faz necessária quando a etiologia de um fenômeno deve ser estudada e aonde a estimação da tendência é importante na compreensão do fenômeno. Segundo Olinda (2008), um exemplo desse tipo de dados são medidas da concentração de um elemento químico no solo.

2.4 Geoestatística

Segundo Landim (2003), Krige em 1951 observou as variâncias de dados de mineração de ouro obtidas por meio da abordagem clássica de amostragem, não faziam sentido se não considerasse as distâncias entre as amostras. Foi então, que em 1963 Matheron, baseado nestas observações, desenvolveu a Teoria das Variáveis Regionalizadas. Conforme Mello (2004), ela foi definida como uma função espacial numérica, variando-se de um local para outro, apresentando-se continuidade aparente e cuja variação não pode ser representada por uma simples função matemática. A aplicação dessa teoria a problemas voltados para a geologia e mineração recebeu o nome de Geoestatística.

Atualmente o termo geoestatística é consagrado como um tópico especial da estatística aplicada que trata de problemas referentes às variáveis regionalizadas, as quais têm um comportamento espacial mostrando características intermediárias entre as variáveis verdadeiramente aleatórias e aquelas totalmente determinísticas. Esta metodologia apresenta uma aparente continuidade no espaço (LANDIM, 2003). Preocupa-se, portanto, com a estimativa da variação regionalizada em uma dimensão, duas dimensões ou três dimensões. Como todas as técnicas estatísticas, a Geoestatística baseia-se em um conceito probabilístico (GUIMARÃES, 2004).

As vantagens reconhecidas da Geoestatística sobre outras técnicas convencionais de predição são o estudo da variabilidade espacial, a suavização, o desagrupamento, a determinação da anisotropia, a precisão e a incerteza. São justamente nos problemas onde a estatística clássica tem limitações, que o uso da Geoestatística tem suas maiores aplicações (VIEIRA, 2000).

2.5 Campos aleatórios

Um campo aleatório ou uma função aleatória é um processo estocástico definido em algum espaço $S \subset \mathbb{R}^d$, ou seja, uma função cujos valores são realizações de variáveis aleatórias em qualquer ponto do domínio (SCHIMIDT; SANZO, 2006), ou em outras palavras, uma família ou coleção de variáveis aleatórias, em que cada um dos seus membros podem ser identificados ou localizados de acordo com a mesma métrica (SCHABENBERGER; GOTWAY, 2005). Um campo aleatório pode ser definido por

$$Z(s_i) : s_i \in S \subset \mathbb{R}^d, \quad (2.1)$$

em que $Z(s_i)$ é o valor da variável regionalizada do atributo Z obtido na localização s_i do espaço sob estudo S e $d \geq 1$ é a dimensão do campo aleatório, para esta análise utiliza-se o \mathbb{R}^2 . Segundo Schimidt e Sanso (2006) e Le e Zidek (2006), a descrição de um campo aleatório é obtida por meio das distribuições acumuladas finito-dimensionais F , para qualquer conjunto de amostras (s_1, s_2, \dots, s_n) da variável aleatória pertencentes à região S e qualquer inteiro n :

$$F_{S_1, S_2, \dots, S_n(z_1, z_2, \dots, z_n)} \equiv P[(Z(s_1) \leq z_1, Z(s_2) \leq z_2), \dots, Z(s_n) \leq z_n)]. \quad (2.2)$$

Devido a sua simplicidade inferencial, a distribuição de probabilidade gaussiana é uma das mais utilizadas na literatura, tomando-se por base apenas uma amostra, visto que a mesma é o resultado único de uma função casual. Sendo assim, um campo aleatório é dito ser gaussiano. Para qualquer conjunto finito de localizações $s = (s_1, s_2, \dots, s_n)$ pertencente a S , $Z(s)$ segue uma distribuição normal, para índices $i = 1, 2, \dots, n$ é uma distribuição gaussiana n -variada e é completamente especificada pelo vetor de médias $n \times 1$, denotado por μ , e pela matriz de covariâncias $n \times n$, denotada por Σ , que no contexto de geoestatística possui comportamento de que quanto maior a distância euclidiana entre duas localizações s_l e s_k , menor a correlação entre $Z(s_l)$ e $Z(s_k)$. Para ser considerada válida, precisa ser positiva definida, para tanto cada um de seus elementos devem ser dados por uma função de covariância que não é fácil encontrar uma forma para gerar esse comportamento e, ao mesmo tempo, assegurar que a matriz de covariâncias fique positiva. Diggle e Ribeiro (2007) mostram maiores detalhes sobre campos aleatórios gaussianos.

Em análises geoestatística geralmente, não é possível ter mais de uma realização do processo, assim sendo, outras suposições devem ser impostas sobre o campo aleatório gaussiano para a realização de inferências. A restrição mais utilizada é que o processo é estacionário, ou seja, a distribuição de probabilidade associada ao campo aleatório

não depende da grandeza de escala das coordenadas, logo, a distribuição conjunta de $(Z(s_1), Z(s_2), \dots, Z(s_n))$ é igual a distribuição conjunta de $(Z(s_1 + \mathbf{h}), Z(s_2 + \mathbf{h}), \dots, Z(s_n + \mathbf{h}))$, para qualquer incremento $\mathbf{h} = \|s_i - s_j\|, \forall i \neq j$ com direção e orientação específica em um espaço de uma, duas ou três dimensões. Outra definição menos restritiva é que a média do campo aleatório é igual em toda a região sob estudo e a correlação entre $Z(s_l)$ e $Z(s_k)$, para quaisquer s_l e s_k , só depende da distância entre as localizações, ou seja, a grandeza de escala de Z não influencia na estrutura de correlação espacial. Esse tipo de estacionariedade é conhecido na literatura como estacionariedade de segunda ordem. Uma observação importante é que a primeira restrição implica na segunda, no entanto, o contrário não é válido, a não ser que o processo espacial seja gaussiano, que produz equivalência entre as duas restrições (DIGGLE; RIBEIRO, 2007). No entanto, nem sempre é fácil verificar as restrições de estacionariedade forte ou fraca, logo, outra possibilidade menos restritiva é assumir que os incrementos $[Z(s_i) - Z(s_i + \mathbf{h})]$ possuem estacionariedade. Esta característica é denominada de estacionariedade intrínseca (SCHABENBERGER; GOTWAY, 2005).

Conforme Fonseca (2008), um campo aleatório é dito ser intrinsecamente estacionário se para todo s_i pertencente a S se

$$\text{i) } E[Z(s_i)] = \mu$$

$$\text{ii) } Var[Z(s_i) - Z(s_i + \mathbf{h})] = 2\gamma(\mathbf{h}),$$

sendo

$$\text{i) } \gamma(\mathbf{h}) = Var(Z(s_i)) - Cov(Z(s_i); Z(s_i + \mathbf{h})),$$

denominado de semivariograma.

Em alguns casos a suposição de estacionariedade não é válida, sendo assim, diversas maneiras são propostas para contornar este problema. Quando a média do processo estocástico não é constante na região de estudo, geralmente, utiliza-se a inclusão de covariáveis na modelagem, onde a média é tratada como efeito fixo, que influenciam a variável de interesse, e sua interpretação é igual a de modelos lineares. Já para problemas com variâncias e covariâncias não constantes, uma técnica mais simples é a utilização de transformação nos valores observados do campo aleatório (CHRISTENSEN *et al.*, 2001). Quando o fenômeno em estudo revela diferentes padrões de dependência espacial, ou seja, apresenta uma variabilidade que não é a mesma em todas as direções, o fenômeno em estudo

é chamado de anisotrópico, em problemas práticos de geoestatística não é fácil identificar tal característica nos dados observacionais.

Um campo aleatório gaussiano é dito ser homogêneo se ele for estacionário e seu padrão de correlações não depende das direções. Utilizando-se essa suposição o processo estocástico fica bastante restritivo. Com a suposição de homogeneidade de um campo aleatório gaussiano, é necessário estabelecer uma função matemática que dependa apenas das distâncias entre as localizações amostradas do espaço sob estudo e que estruture a matriz de covariâncias, de forma que esta seja positiva e com o comportamento empírico utilizado em geoestatística.

Estacionariedade

Conforme GUIMARÃES (2004), uma variável aleatória é estacionária se o desenvolvimento desse processo no tempo ou no espaço ocorrer de maneira mais ou menos homogênea, com oscilações aleatórias contínuas em torno de um valor médio. As características de um processo estacionário independe da origem adotada. Pode-se definir uma função aleatória $Z(s)$ como estacionária, se todos os momentos estatísticos são invariantes para toda mudança de origem.

Estatisticamente pode-se dizer que, se o processo é estacionário de ordem k , então:

$$E[Z^k(s)] = \mu_k(s), \quad (2.3)$$

em que s é uma constante, confirmando-se a estacionariedade de ordem k , ele será estacionário também nas demais ordens inferiores a k .

Para estudos de geoestatística necessita-se, como restrição máxima, que o primeiro e o segundo momento em relação à origem sejam constantes, ou seja, exige-se no máximo a estacionariedade de segunda ordem. Se a esperança matemática de uma variável aleatória é constante, independente da origem que se toma no espaço ou no tempo, pode-se dizer que a variável é estacionária de primeira ordem e, portanto, a média será a mesma para todo o processo, isto é,

$$E[Z(s)] = \mu(s) = \mu \quad (2.4)$$

Estacionariedade 1ª ordem

Sejam $Z(s_i)$ e $Z(s_i + h)$ dois valores de uma variável regionalizada obtidos dos pontos s_i e $s_i + h$, separados por uma distância h . Sendo assim, um campo aleatório é estacionário

de primeira ordem se para qualquer deslocamento do vetor \mathbf{h} , a esperança matemática é constante em relação à origem (GUIMARÃES, 2004), ou seja,

$$E[Z(s_i)] = E[Z(s_i + \mathbf{h})] = \mu. \quad (2.5)$$

Estacionariedade 2ª ordem

Segundo Rocha (2005) e GUIMARÃES (2004), uma variável regionalizada é estacionária de segunda ordem, se além de cumprir a estacionariedade de primeira ordem, a covariância (Cov) entre as variáveis $Z(s_i)$ e $Z(s_i + \mathbf{h})$, separados por um vetor distância \mathbf{h} , depende apenas de \mathbf{h} . Em que:

A média é constante

$$E[Z(s_i)] = \mu \quad (2.6)$$

e o segundo momento existe

$$E[Z^2(s_i)] < \infty \quad (2.7)$$

Para cada par $Z(s_i), Z(s_i + \mathbf{h})$ a função de covariância existe e depende apenas de \mathbf{h}

$$\begin{aligned} Cov[Z(s_i), Z(s_i + \mathbf{h})] &= E[(Z(s_i) - \mu) \times (Z(s_i + \mathbf{h}) - \mu)] \\ &= E[Z(s_i) \times Z(s_i + \mathbf{h}) - \mu \times Z(s_i + \mathbf{h}) - \mu \times Z(s_i) + \mu^2] \\ &= E[Z(s_i) \times Z(s_i + \mathbf{h})] - \mu \times E[Z(s_i + \mathbf{h})] - \mu \times E[Z(s_i)] + \mu^2 \end{aligned}$$

Da Equação (2.5) tem-se que:

$$\begin{aligned} Cov(\mathbf{h}) &= E[Z(s_i) \times Z(s_i + \mathbf{h})] - \mu^2 - \mu^2 + \mu^2 \\ &= E[Z(s_i) \times Z(s_i + \mathbf{h})] - \mu^2, \end{aligned} \quad (2.8)$$

ou seja,

$$E[Z(s_i) \times Z(s_i + \mathbf{h})] = Cov(\mathbf{h}) + \mu^2 \quad (2.9)$$

Na Equação (2.8), a estacionariedade da covariância implica na estacionariedade da variância:

$$\begin{aligned}
Var[Z(s_i)] &= E[Z(s_i) - \boldsymbol{\mu}]^2 \\
&= E[Z^2(s_i) - 2\boldsymbol{\mu} \times Z(s_i) + \boldsymbol{\mu}^2] \\
&= E[Z(s_i) \times Z(s_i + \mathbf{0})] - 2\boldsymbol{\mu}^2 + \boldsymbol{\mu}^2 \\
&= E[Z(s_i) \times Z(s_i + \mathbf{0})] - \boldsymbol{\mu}^2 \\
&= Cov(\mathbf{0})
\end{aligned} \tag{2.10}$$

A estacionariedade da covariância implica na estacionariedade da variância; isto é, $VarZ(s_i) = Cov(\mathbf{0})$ e do variograma que é definido como:

$$\begin{aligned}
2\gamma(\mathbf{h}) &= E[Z(s_i + \mathbf{h}) - Z(s_i)]^2 \\
&= E[Z^2(s_i) - 2Z(s_i) \times Z(s_i + \mathbf{h}) + Z^2(s_i + \mathbf{h})] \\
&= E[Z^2(s_i)] - 2E[Z(s_i) \times Z(s_i + \mathbf{h})] + E[Z^2(s_i + \mathbf{h})]
\end{aligned} \tag{2.11}$$

em que $2\gamma(\mathbf{h})$ representa a função conhecida como variograma, que será detalhada na Seção 2.5.1.

Da Equação (2.9) obtém-se:

$$E[Z(s_i) \times Z(s_i + \mathbf{h})] = Cov(\mathbf{h}) + \boldsymbol{\mu}^2 \tag{2.12}$$

e de (2.10) tem-se que:

$$E[Z^2(s_i)] = E[Z(s_i) \times Z(s_i + \mathbf{0})] = Cov(\mathbf{0}) + \boldsymbol{\mu}^2 \tag{2.13}$$

Substituindo as Equações (2.12) e (2.13) na Equação (2.11), obtém-se:

$$2\gamma(\mathbf{h}) = Cov(\mathbf{0}) + \boldsymbol{\mu}^2 - 2[Cov(\mathbf{h}) + \boldsymbol{\mu}^2] + E[Z^2(s_i + \mathbf{h})] \tag{2.14}$$

e como $E[Z^2(s_i)] = E[Z^2(s_i + \mathbf{h})]$, tem-se que

$$\begin{aligned}
2\gamma(\mathbf{h}) &= Cov(\mathbf{0}) + \boldsymbol{\mu}^2 - 2[Cov(\mathbf{h}) + \boldsymbol{\mu}^2] + Cov(\mathbf{0}) + \boldsymbol{\mu}^2 \\
&= 2Cov(\mathbf{0}) - 2Cov(\mathbf{h})
\end{aligned} \tag{2.15}$$

de onde segue que

$$\gamma(\mathbf{h}) = Cov(\mathbf{0}) - Cov(\mathbf{h}) \tag{2.16}$$

em que $\gamma(\mathbf{h})$ representa o semivariograma na teoria das variáveis regionalizadas, que será detalhada na Seção 2.5.1.

O coeficiente de correlação entre $Z(s_i + \mathbf{h})$ e $Z(s_i)$, chamado de correlograma ou autocorrelograma, é definido por

$$r(\mathbf{h}) = \frac{C(\mathbf{h})}{C(\mathbf{0})} = 1 - \frac{\gamma(\mathbf{h})}{C(\mathbf{0})} \quad (2.17)$$

Se ocorrer a estacionariedade de segunda ordem, o correlograma (autocorrelograma) e o variograma (semivariograma) serão ferramentas correspondentes na determinação da dependência espacial. Mas se a estacionariedade de segunda ordem não é atendida o autocorrelograma não pode ser usado, pois, o denominador da função autocorrelograma é uma variância e, neste caso, $C(\mathbf{0}) \neq$ constante (GUIMARÃES, 2004).

Estacionariedade Intrínseca

Como na hipótese anterior, aqui se admite que $E[Z(s_i)] = \boldsymbol{\mu}(s_i) = \boldsymbol{\mu}, \forall(s_i)$. Além disso, admite-se que a variância das diferenças depende somente do vetor distância

$$Var[Z(s_i) - Z(s_i + \mathbf{h})] = E[Z(s_i) - Z(s_i + \mathbf{h})]^2 = 2\gamma(\mathbf{h}) \quad (2.18)$$

Por ser a menos restritiva e requerer apenas a existência e estacionariedade do variograma, sem nenhuma restrição quanto à existência de variância finita, esta hipótese é a mais frequentemente admitida em Geoestatística. Em se tratando das hipóteses da Krigagem Universal, admite-se que $\boldsymbol{\mu}(s_i)$ é a tendência principal e que $C(\mathbf{h})$ e $\gamma(\mathbf{h})$ possuem estacionariedade dentro de uma vizinhança de tamanho restrito. Além disso, supõe-se que $E[Z(s_i)] = \boldsymbol{\mu}(s_i)$, que $\boldsymbol{\mu}(s_i)$ deixa de ser estacionária, variando-se de modo regular dentro de tal vizinhança. Não somente a covariância e o variograma são definidos a partir de valores experimentais, mas também o tamanho da vizinhança onde as hipóteses mantêm-se válidas (DAVID, 1977).

2.5.1 Variograma e Semivariograma

Conforme Frade (2011), a hipótese de aleatoriedade dos dados não pode ser aceita antes que se prove a inexistência de correlação entre os pontos amostrais. O variograma é uma ferramenta básica de suporte às técnicas de krigagem, que permite representar quantitativamente a variação de um fenômeno regionalizado no espaço (CAMARGO, 2002). Segundo Guerra (1988), existem três tipos de variogramas:

- i) experimental ou observado é o primeiro gráfico que deve ser feito e é obtido a partir do conjunto de dados do experimento vindos de um processo de amostragem sobre as coordenadas geográficas;
- ii) o verdadeiro é desconhecido e representa a situação real;
- iii) teórico é utilizada para verificar qual modelo que melhor se ajusta ao variograma experimental, e fazer inferências sobre o verdadeiro variograma.

Conforme Camargo *et al.* (2002), supondo que $Z(s_i)$ represente o valor da variável para o local s_i , em que s_i é o vetor $(s_1; s_2)$ e $Z(s_i + \mathbf{h})$ representa o valor da mesma variável para alguma distância \mathbf{h} , em qualquer direção conforme ilustra a Figura 1.

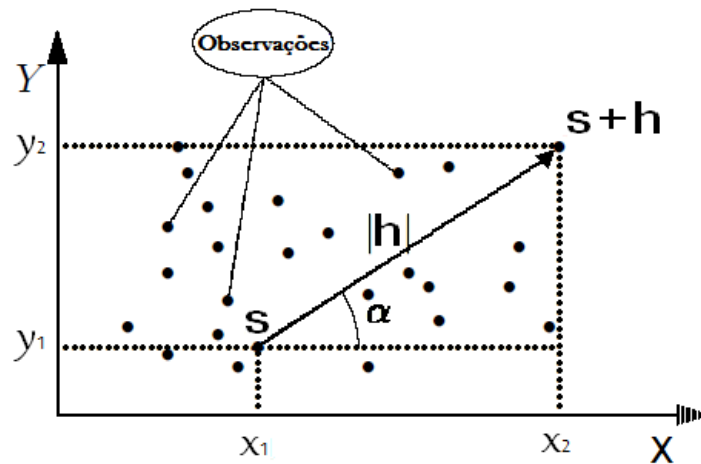


Figura 1: Amostragem em duas dimensões

A função variograma, denominada de $2\gamma(\mathbf{h})$, é a esperança matemática do quadrado da diferença entre pares de pontos separados por uma distância \mathbf{h}

$$2\gamma(\mathbf{h}) = E[Z(s_i) - Z(s_i + \mathbf{h})]^2 \quad (2.19)$$

Se as condições da hipótese de estacionariedade são contempladas, o variograma ($2\gamma(\mathbf{h})$) pode ser estimado a partir dos dados amostrais:

$$2\hat{\gamma}(\mathbf{h}) = \frac{1}{N(\mathbf{h})} \sum_{i=1}^{N(\mathbf{h})} [Z(s_i) - Z(s_i + \mathbf{h})]^2, \quad (2.20)$$

em que: $2\hat{\gamma}(\mathbf{h})$ é o variograma; $N(\mathbf{h})$ é o número de pares de valores medidos, $Z(s_i)$ e $Z(s_i + \mathbf{h})$, separados por um vetor distância \mathbf{h} ; $Z(s_i)$ e $Z(s_i + \mathbf{h})$ são valores da i -ésima

observação da variável regionalizada, coletados nos pontos s_i e $s_i + \mathbf{h}$, $i = 1, 2, \dots, n$ separados pelo vetor \mathbf{h} .

A metade da função variograma é denominada de função semivariograma. O estimador da semivariância é $[\hat{\gamma}(\mathbf{h})]$ é igual a média aritmética das diferenças ao quadrado entre pares de valores experimentais, em todos os pontos separados pela distância \mathbf{h}

$$\hat{\gamma}(\mathbf{h}) = \frac{1}{2N(\mathbf{h})} \sum_{i=1}^{N(\mathbf{h})} [Z(s_i) - Z(s_i + \mathbf{h})]^2, \quad (2.21)$$

em que $[\hat{\gamma}(\mathbf{h})]$ é a semivariância estimada para cada distância (\mathbf{h}), $N(\mathbf{h})$ é o número de pares de pontos separados por uma distância \mathbf{h} , $Z(s_i)$ é o valor da variável regionalizada no ponto s_i e $Z(s_i + \mathbf{h})$ é o valor no ponto $s_i + (\mathbf{h})$. A função semivariograma permite gerar o gráfico da semivariância em função da distância (\mathbf{h}), denominado de semivariograma experimental, o qual permite interpretar a continuidade espacial da variável regionalizada.

O semivariograma é uma ferramenta da geoestatística que permite verificar e modelar a dependência espacial de uma variável. Uma aplicação imediata do semivariograma é a utilização das informações geradas por ele na interpolação, ou seja, na estimativa de dados e posterior mapeamento da variável (GUIMARÃES, 2004).

Os semivariogramas são preferidos para caracterizar a estrutura de continuidade espacial da característica avaliada, por exigirem hipóteses de estacionaridade menos restritivas Hipótese Intrínseca (MELLO, 2004). O semivariograma representa uma função de semivariâncias em relação às respectivas distâncias. A semivariância é definida como a metade da variância de diferenças entre observações de uma variável aleatória Z , separadas por uma distância \mathbf{h} . Assim, valores baixos indicam menor variabilidade.

Por meio do semivariograma experimental, o pesquisador é capaz de definir o modelo que melhor descreve o comportamento dos dados no espaço (JOURNEL; HUIJBREGTS, 1978). Em seguida, sua preocupação se volta para o ajuste da função matemática ao semivariograma experimental ou aos dados.

Até a década de 80, o ajuste do modelo espacial ao semivariograma experimental, era usualmente feito de forma visual a "sentimento", sem nenhum procedimento matemático. Posteriormente, com o aumento da capacidade dos recursos computacionais, outros métodos de ajuste, sem subjetividade, foram e estão sendo estudados. Dentre estes métodos, destacam-se os Métodos dos Quadrados Mínimos Ordinários, Ponderados e o Método de Máxima Verossimilhança. Pode-se dizer que estes métodos tiveram o intuito de retirar e/ou atenuar o caráter de subjetividade na estimação dos parâmetros do semivariograma.

Após a construção do semivariograma experimental, a etapa seguinte é estudar o melhor modelo do semivariograma. Conforme McBratney e Webster (1986), a seleção se dá por meio de técnicas quantitativas como, por exemplo, o Critério de Informação de Akaike, que em inglês é designado pela sigla AIC *Akaike's Information Criterion* e a comparação de modelos sob a teoria *Bayesiana*, como o *Bayesian Information Criterion* BIC de Schwarz AKAIKE (1983), maiores detalhes nas subseções 2.5.7 e 2.5.8.

2.5.2 Parâmetros do semivariograma

Os parâmetros do semivariograma são utilizados para interpretar o fenômeno regionalizado. Estes parâmetros são representados pela Figura 2 que ilustra um semivariograma. De posse do semivariograma é possível ajustar uma função matemática que expressa a estrutura da dependência espacial, estas funções serão descritas na subseção 2.5.6.

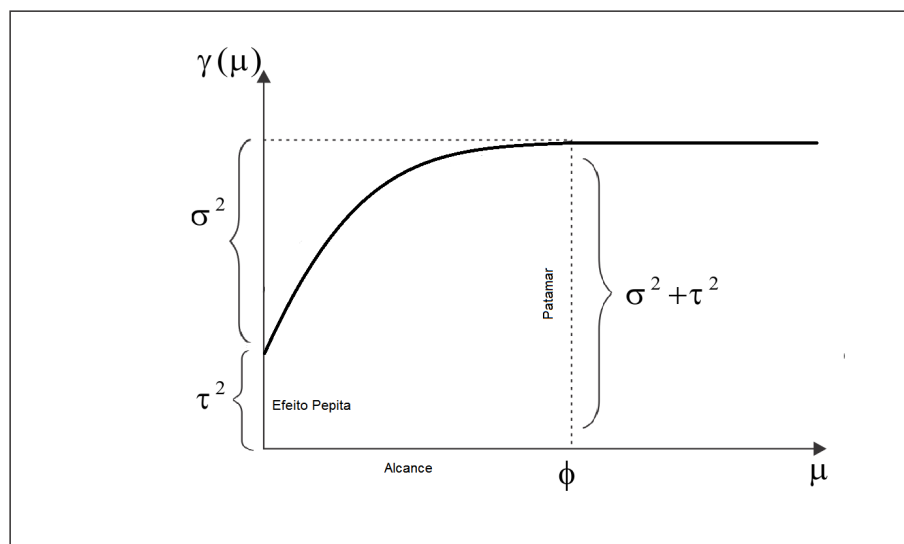


Figura 2: Parâmetros do semivariograma

- i) Alcance (ϕ): distância dentro da qual as amostras apresentam-se correlacionadas espacialmente.
- ii) Patamar ($\tau^2 + \sigma^2$): é o valor do semivariograma correspondente a seu alcance (ϕ). Deste ponto em diante, considera-se que não existe mais dependência espacial entre as amostras, porque a variância da diferença entre pares de amostras ($Var[Z(s_i) - Z(s_i + \mathbf{h})]$) torna-se invariante com a distância.
- iii) Efeito Pepita (τ^2): o ideal seria que, $\gamma(0) = 0$, entretanto, na prática, à medida que \mathbf{h} tende para 0 (zero), $\gamma(\mathbf{h})$ se aproxima de um valor positivo chamado Efeito Pepita τ^2 ,

que revela a descontinuidade do semivariograma para distâncias menores do que a menor distância entre as amostras. Parte desta descontinuidade pode ser também devida a erros de medição (ISAAKS; SRIVASTAVA, 1989), mas é impossível quantificar se a maior contribuição provém dos erros de medição ou da variabilidade de pequena escala não captada pela amostragem.

iv) Contribuição (σ^2): é a diferença entre o patamar ($\tau^2 + \sigma^2$) e o Efeito Pepita (τ^2).

Quando o semivariograma é constante e igual ao patamar para qualquer valor de h , tem-se a ausência total de dependência espacial, ou seja, tem-se o efeito pepita puro.

O semivariograma onde as semivariâncias crescem, sem limite, para todos os valores de h , é chamado de semivariograma sem patamar definido. Este semivariograma indica que a hipótese de estacionariedade de segunda ordem não foi atendida e, provavelmente, trabalha-se com a hipótese intrínseca (fenômeno com capacidade infinita de dispersão). Ele indica também que a máxima distância h entre as amostras não foi capaz de exibir toda a variância dos dados e provavelmente existe tendência dos dados para determinada direção. Uma alternativa para solucionar esse problema de tendência é remover esta tendência e verificar se a variável resíduo apresenta semivariograma com patamar (estacionariedade de segunda ordem).

2.5.3 Modelos geoestatísticos gaussianos

Considerando-se que em alguma área S exista um campo aleatório gaussiano Z , o processo existe mas não é observável, é necessário fazer uma amostragem de n localizações espaciais dentro da área S e observar valores de atributo de interesse nas localizações amostradas. Sendo assim, existe um vetor $Y(\mathbf{s})$ de dimensão $n \times 1$ de valores observados em $\mathbf{s} = (s_1, s_2, \dots, s_n)$, que segundo Diggle e Ribeiro (2007), pode ser definido por

$$Y(\mathbf{s}) = \boldsymbol{\mu}(x) + Z(\mathbf{s}) + \boldsymbol{\varepsilon}, \quad (2.22)$$

em que $\boldsymbol{\mu}(x) = \mathbf{X}\boldsymbol{\beta}$, sendo \mathbf{X} uma matriz $n \times q$ contendo $q - 1$ possíveis covariáveis, $\boldsymbol{\beta}$ um vetor $q \times 1$ de parâmetros associados a \mathbf{X} , $Z(\mathbf{s})$ um campo aleatório gaussiano que possui vetor de médias $n \times 1$ nulo e matriz de covariâncias $\boldsymbol{\Sigma}$ de dimensão $n \times n$, em que cada elemento $\Sigma_{i,j}$ é igual a $Cov[Z(s_i); Z(s_j)]$, para todo s_i e s_j pertencentes a \mathbf{s} , e $\boldsymbol{\varepsilon}$ um vetor $n \times 1$ de ruídos brancos, que por suposição são independentes e identicamente distribuídos (*iid*) com distribuição de probabilidade normal com média zero e desvio padrão σ .

A variação estocástica de uma quantidade física nem sempre é bem descrita por uma distribuição Gaussiana (DIGGLE; RIBEIRO, 2007). Uma das maneiras mais simples de entender o modelo Gaussiano é assumir que depois de aplicar uma transformação para os dados originais, um valor positivo na variável resposta. Uma classe útil de transformações é da família Box-Cox (BOX; COX, 1964).

$$Y^* = \begin{cases} \frac{(Y^\lambda - 1)}{\lambda}, & \text{se } \lambda \neq 0 \\ \log Y, & \text{se } \lambda = 0 \end{cases} \quad (2.23)$$

em que: λ é o parâmetro da transformação e Y a variável resposta.

Outra extensão simples para o modelo básico é permitir que haja uma variação média espacial, por exemplo, substituindo-se o μ constante por um modelo de regressão linear para a esperança condicional de Y_i dado $Z(s_i)$.

Anisotropia

Segundo Camargo (2002), a anisotropia pode ser facilmente constatada por meio da observação dos semivariogramas obtidos para diferentes direções. Quando os semivariogramas seguem uma mesma direção 0° , 45° , 90° e 135° , a distribuição dos dados é denominada de isotrópica. Neste caso, um único modelo é suficiente para descrever a variabilidade espacial do fenômeno em estudo.

Por outro lado, se os semivariogramas não são iguais em todas as direções, é denominada anisotrópica. Se a anisotropia é observada e é refletida pelo mesmo patamar com diferentes alcances do mesmo modelo, então ela é denominada geométrica.

Existe ainda um outro tipo de anisotropia em que os semivariogramas apresentam os mesmos alcances e diferentes patamares. Neste caso, a anisotropia é denominada zonal. Como a isotropia, a anisotropia zonal também é um caso menos frequente nos fenômenos naturais. O mais comum é encontrar combinações da anisotropia zonal e geométrica, denominada anisotropia combinada.

2.5.4 Índice de dependência espacial

Para analisar o grau da dependência espacial do atributo em estudo, pode-se utilizar o Índice de Dependência Espacial $IDE(\%)$, definido pela seguinte equação:

$$IDE(\%) = \frac{\tau^2}{\tau^2 + \sigma^2} \times 100, \quad (2.24)$$

em que:

- i) τ^2 Efeito pepita ou *nugget*;
- ii) σ^2 Componente estrutural;
- iii) $\tau^2 + \sigma^2$ Patamar.

Conforme a classificação quanto ao grau de dependência espacial da variável em estudo, segundo os intervalos propostos por Cambardella *et al.* (1994), para avaliar a % do efeito pepita:

- i) $IDE \leq 25\%$ se a razão entre o efeito pepita e o patamar for menor ou igual a 25%, diz-se que a variável tem forte dependência espacial;
- ii) $25\% < IDE < 75\%$ se a razão entre o efeito pepita e o patamar estiver entre 25% e 75%, tem-se uma dependência espacial moderada;
- iii) $IDE > 75\%$ se a razão entre o efeito pepita e o patamar estiver entre 75% e 100%, a variável apresenta fraca dependência;
- iv) $IDE = 100\%$ se a razão entre efeito pepita e patamar for igual a 100%, em que o semi-variograma apresenta efeito pepita puro, diz-se que a variável possui independência espacial.

2.5.5 Função de máxima verossimilhança

Assumindo-se que o campo aleatório possui estacionariedade, pode-se optar por estimadores de máxima verossimilhança, que consiste em utilizar os valores observados da variável resposta para encontrar um vetor $\hat{\theta}$ que seja o ponto de máximo da função de verossimilhança associada a θ . Por simplicidade matemática, normalmente utiliza-se o logaritmo da função de verossimilhança para fazer a estimação, que associada a (2.14) tem a seguinte forma:

$$l(\theta; Y(s)) = -0,5(n \ln(2\pi) + (\ln(|\Sigma_Y|) + (Y(s) - \mathbf{X}\beta)^t \sum_Y^{-1} (Y(s) - \mathbf{X}\beta)) \quad (2.25)$$

No contexto de geoestatística, Diggle e Ribeiro (2007), propuseram a utilização da reparametrização $\nu = \frac{\tau}{\sigma}$, a qual facilita a estimação de θ . O vetor de parâmetros a ser

estimado passa a ser $\theta^* = (\beta, \sigma^2, \phi^*, \nu^2)$ e \sum_Y pode ser escrita como $\sigma^2 \mathbf{V}$, em que \mathbf{V} é uma matriz $n \times n$ que depende apenas de ν e ϕ^* . O logaritmo da função de verossimilhança fica da seguinte forma:

$$l(\beta^*; Y(s)) \propto -0,5(n \ln(\sigma^2) - \ln(|\mathbf{V}|) - \sigma^{-2}(Y Y^t(\mathbf{s} - \mathbf{X}\beta)^t \mathbf{V}^{-1}(Y(\mathbf{s}) - \mathbf{X}\beta)) \quad (2.26)$$

sendo que existem formas analíticas apenas para os estimadores de β e σ^2 .

$$\begin{aligned} \hat{\beta} &= (\mathbf{X}^t \mathbf{V}^{-1} \mathbf{X})^{-1} (\mathbf{X}^t \mathbf{V}^{-1} Y(\mathbf{s})) \\ \hat{\sigma}^2 &= n^{-1} (Y(\mathbf{s}) - \mathbf{X} \hat{\beta})^t \mathbf{V}^{-1} (Y(\mathbf{s}) - \mathbf{X} \hat{\beta}). \end{aligned} \quad (2.27)$$

Observe que $\hat{\beta}$ e $\hat{\sigma}^2$ são funções dos demais parâmetros e além das formas fechadas para os estimadores, é possível encontrar a matriz de informação de Fischer observada para os mesmos, logo, pode-se encontrar a matriz de covariâncias associada a $(\hat{\beta}, \hat{\sigma}^2)$ (DIGGLE; RIBEIRO, 2007). Para ϕ^* e ν^2 não existem formas analíticas para os estimadores. Utilizando $\hat{\beta}$ e $\hat{\sigma}^2$ em (2.19), tem-se o logaritmo da função de verossimilhança concentrada, que depende apenas de $\theta_c = (\nu, \phi^*)$.

Para calcular $\hat{\theta}_c$ é possível utilizar métodos numéricos de maximização de funções, como por exemplo, o método de NELDER e R. (1986), o qual calcula $\hat{\phi}^*$ e $\hat{\nu}^2$ e a matriz Hessiana estimada, denotada por \mathbf{H} . Com os parâmetros da função de máxima verossimilhança concentrada estimados, por invariância é possível encontrar as estimativas de β , σ^2 e τ^2 .

Utilizando-se as propriedades assintóticas dos estimadores de máxima verossimilhança pode-se encontrar a distribuição de probabilidade de $\hat{\theta}$, que é $N(\theta, \Sigma_\theta)$, sendo $\Sigma_\theta = \Delta^t \Sigma_{\theta^*} \Delta$, em que a i -ésima coluna de Δ é o vetor $\frac{\partial l(\theta_i)}{\partial \theta^*}$ e:

$$\Sigma_{\theta^*} = \begin{bmatrix} \Sigma_{\beta, \sigma^2} & \mathbf{0} \\ \mathbf{0}^t & \Sigma_{\theta_c} \end{bmatrix}$$

sendo Σ_{β, σ^2} a matriz de covariâncias de $(\hat{\beta}, \hat{\sigma}^2)$, que possui forma analítica, $\Sigma_{\theta_c} = -\mathbf{H}^{-1}$ é a matriz de covariâncias de $\hat{\theta}_c$, e $\mathbf{0}$ uma matriz de zeros, pois não é possível recuperar as covariâncias entre (β, σ^2) e θ_c .

Geralmente o objetivo final dos estudos com dados georeferenciados é calcular as predições espaciais em localizações não amostradas. Esse processo de predição é denominado de krigagem e é baseado nas propriedades inferenciais da distribuição normal,

maiores detalhes na subseção 2.5.9.

2.5.6 Funções de correlações

Conforme Mello (2004), a partir do semivariograma é possível ajustar uma função matemática que expressa a estrutura de dependência espacial da característica avaliada, é necessário ajustá-lo a um modelo teórico que deve fornecer soluções estáveis para o estimador (krigagem). Isto quer dizer que as covariâncias têm de ser definidas positivamente. A condição de não negatividade do modelo, limita o conjunto de funções usadas na modelagem do semivariograma experimental. Dentre várias funções de correlações existentes na literatura, destaca-se neste trabalho: Gaussiano, Esférico, Circular e a família de funções de Matérn com $kappa$ ¹ (0,5) e (1,0), em que a função Matérn com $kappa$ igual a 0,5 é a função exponencial.

Matérn

Essa família de funções de correlação foi proposta por MATÉRN (1986) e possui a seguinte forma:

$$\rho(\mathbf{h}) = \{2^{v-1}\Gamma(v)\}^{-1} \left(\frac{\mathbf{h}}{\phi}\right)^v K_v\left(\frac{\mathbf{h}}{\phi}\right), \quad (2.28)$$

sendo a função gama definida por $\Gamma(w) = \int_0^\infty e^{-t}t^{w-1}dt$, $K_v(\cdot)$ é a função Bessel de ordem v , definida por $x^2 \frac{\partial^2 y}{\partial x^2} + x \frac{\partial y}{\partial x} + (x^2 - v^2)y = 0$, \mathbf{h} a distância euclidiana entre duas localizações quaisquer do campo aleatório, os parâmetros dessa função são $\phi > 0$ e $v > 0$. Sendo que o ϕ está vinculado ao alcance das correlações, e maiores valores indicam dependência espacial de maior alcance. Já o segundo parâmetro k está relacionado à suavidade do processo, de forma que quanto maior o valor do parâmetro, maior a suavidade.

Gaussiano

Segundo Landim (2003), a função de correlação gaussiana é um modelo transitivo, muitas vezes usado para modelar fenômenos extremamente contínuos (CAMARGO, 2002). Sua formulação é definida por

$$\rho(\mathbf{h}) = \exp\left(-\frac{\mathbf{h}}{\phi}\right)^2. \quad (2.29)$$

Semelhante ao modelo exponencial, o modelo gaussiano atinge o patamar assintoticamente e o parâmetro ϕ é definido como o alcance prático ou distância na qual o valor do modelo é 95% do patamar. A curva é parabólica junto à origem e a tangente nesse ponto

¹Suavização analítica.

é horizontal, o que indica pequena variabilidade para curtas distâncias (LANDIM, 2003).

Circular

Segundo McBratney e Webster (1986), a função de correlação circular é válida em \mathbb{R} e em \mathbb{R}^2 mas não em \mathbb{R}^3 e define-se da seguinte forma:

$$\rho(\mathbf{h}) = \begin{cases} \frac{2}{\pi}(\text{sen}^{-1}(\mathbf{h} + \sqrt{1 - \mathbf{h}^2})), & \mathbf{h} > 0 \\ 1, & \mathbf{h} = 0 \end{cases} \quad (2.30)$$

Esférico

A função esférica é definida pela seguinte função:

$$\rho(\mathbf{h}) = 1 - 1,5 \left(\frac{\mathbf{h}}{\phi}\right) + 0,5 \left(\frac{\mathbf{h}}{\phi}\right)^3 \quad (2.31)$$

em que: \mathbf{h} é a distância, e o ϕ é o parâmetro de alcance.

Pode-se afirmar que equivale à função de distribuição normal da estatística clássica (LANDIM, 2003).

2.5.7 Critério de Informação de Akaike - (AIC)

Na validação do desempenho dos modelos, o Critério de informação de *Akaike* - *Akaike's Information Criterion (AIC)* tem sido bastante utilizado para diferentes estruturas de covariâncias. AKAIKE (1983), relacionou a discrepância, medida que existe entre o modelo verdadeiro e o modelo aproximado, com a máxima verossimilhança, que é o que possui melhor ajuste para os dados observados.

O princípio do Critério de informação de *Akaike* é selecionar uma combinação de variáveis exploratórias a modelos para a função de correlação que minimize o valor do *AIC*. É importante observar que, em muitas situações clássicas, tais como regressão linear ou em modelos de séries temporais, o *AIC* não é uma condição consistente para a seleção de modelos, ou seja, como o crescente aumento do tamanho das amostras, há uma alta probabilidade de que um modelo selecionado pelo *AIC* não corresponda ao verdadeiro modelo. O critério de informação de *Akaike* é expresso por:

$$AIC = -2 \times [L(\boldsymbol{\theta}; y)] + 2p, \quad (2.32)$$

em que, $L(\boldsymbol{\theta}; y)$ é a função de verossimilhança maximizado e p é o número de parâmetros. Segundo este critério, o melhor modelo é o que possui menor valor de *AIC*.

2.5.8 Critério de Informação Bayesiana - (*BIC*)

Segundo Carlin e Louis (2000), a aplicação deste critério leva à escolha do modelo que minimiza menos duas vezes o valor esperado a posteriori da log-verossimilhança mais uma função de penalização, que depende da dimensão da amostra e do número de parâmetros do modelo.

Conforme Olinda (2008), a comparação de modelos sob a teoria Bayesiana pode ser feita a partir de medidas de adequabilidade, como o *Bayesian Information Criterion (BIC)* de Schwarz, os quais são aproximações do fator de Bayes. Ao estimar os parâmetros do modelo usando estimativa da máxima verossimilhança, é possível aumentar a probabilidade de adicionar parâmetros, que podem resultar em *overfitting*².

O *BIC* resolve este problema por meio da introdução de um termo de penalidade para o número de parâmetros do modelo definido por

$$BIC = -2 \times (\log L(\boldsymbol{\theta}; y)) + p \log(n), \quad (2.33)$$

em que, n é o número de observações, ou equivalente ao tamanho da amostra, p é o número de parâmetros livres a serem estimados e $\log L(\boldsymbol{\theta}; y)$ é o valor maximizado da função de verossimilhança para o modelo estimado. Menor valor do *BIC* indica o melhor ajuste do modelo.

2.5.9 Krigagem

O método de krigagem foi desenvolvido para solucionar problemas de mapeamentos geológicos, mas com o tempo, expandiu-se para outros campos, como por exemplo, mapeamento de solos, mapeamento hidrológico, mapeamento atmosférico e outros campos correlatos. Segundo Lima (2006), utiliza-se a variância da krigagem para definir o intervalo de confiança do tipo gaussiano.

O estimador espacial denominado de krigagem, tem como base os dados amostrados da variável regionalizada e as propriedades estruturais do semivariograma obtido a partir destes dados. Krigagem é um método de inferência espacial, o qual estima dados em pontos não amostrados a partir de pontos amostrados, levando-se em consideração a estrutura de dependência espacial do fenômeno em estudo. A estimativa pontual é obtida pela seguinte expressão:

²Ocorre quando um modelo estatístico descreve um erro aleatório ou ruído.

$$Z_{s_0}^* = \sum_{i=1}^n \lambda_i Z(s_i),$$

em que λ_i são os pesos de krigagem definidos conforme o semivariograma e $z(s_i)$ são os pontos amostrados. Conhecido o semivariograma da variável, e havendo dependência espacial, podem-se interpolar valores em qualquer posição no campo de estudo, sem tendência e com variância mínima (VIEIRA, 2000).

O estimador espacial denominado krigagem, tem como base os dados amostrados da variável regionalizada e as propriedades estruturais do semivariograma obtido a partir destes dados. O método fornece, além dos valores estimados, o erro associado a tal estimativa, o que o distingue dos demais algoritmos à disposição (LANDIM, 2003).

Krigagem Simples

É utilizada quando as médias locais são relativamente constantes em toda área em estudo e de valor muito semelhante à média da população que é conhecida. A média da população é utilizada para cada estimativa local, em conjunto com os pontos vizinhos estabelecidos como necessários para a estimativa. Considere amostras Z , em n pontos distintos, com coordenadas representadas pelo vetor $s = (s_1, s_2, \dots, s_n)$, em que s_i identifica uma posição em duas dimensões representadas pelos pares de coordenadas (s_{i1}, s_{i2}) , para $i = 1, 2, \dots, n$. Assim, tem-se um conjunto de valores $Z(s_i)$ $i = 1, 2, \dots, n$. Supondo que o objetivo é estimar o valor de Z no ponto s_0 , o valor desconhecido de $Z(s_0)$ pode ser estimado a partir de uma combinação linear dos n valores observados, adicionado a um parâmetro, λ_0 (JOURNEL; HUIJBREGTS, 1978).

$$Z_{s_0}^* = \lambda_0 + \sum_{i=1}^n \lambda_i Z(s_i).$$

Então, tem-se:

$$\begin{aligned} E[Z_{s_0}^*] &= E[\lambda_0 + \sum_{i=1}^n \lambda_i Z(s_i)] \\ &= \lambda_0 + \sum_{i=1}^n \lambda_i E[Z(s_i)]. \end{aligned} \quad (2.34)$$

Deseja-se um estimador não tendencioso, isto é,

$$E[Z_{s_0} - Z_{s_0}^*] = 0 \quad (2.35)$$

ou seja,

$$E[Z_{s_0}] = E[Z_{s_0}^*] \quad (2.36)$$

Substituindo a Equação (2.28) em (2.30), obtém-se o parâmetro λ_0

$$\lambda_0 = E[Z_{s_0}] - \sum_{i=1}^n \lambda_i \times E[Z(s_i)] + \sum_{i=1}^n \lambda_i \times Z(s_i). \quad (2.37)$$

a média μ é conhecida e constante a priori no método de krigagem simples, então

$$E[Z_{s_0}] = E[Z(s_i)] = \mu. \quad (2.38)$$

Substituindo a Equação (2.32) em (2.31), o estimador de krigagem simples fica

$$\begin{aligned} Z_{s_0}^* &= \mu - \sum_{i=1}^n \lambda_i \times \mu + \sum_{i=1}^n \lambda_i \times Z(s_i) \\ &= \mu + \sum_{i=1}^n \lambda_i \times [Z(s_i) - \mu]. \end{aligned} \quad (2.39)$$

Conforme Journel (1988), minimizando a variância do erro ($Var[Z_{s_0} - Z_{s_0}^*]$), os pesos λ_i , são obtidos a partir do seguinte sistema de equações, denominados sistemas de krigagem simples:

$$\sum_{i=1}^n \lambda_i \times Cov(s_i, s_j) = Cov(s_i, s_0),$$

para $i = 1, 2, \dots, n$ em que, $Cov(s_i, s_j)$ refere-se à função de covariância correspondente a um vetor \mathbf{h} , com origem em s_i e extremidade no ponto s_0 a ser estimado.

Por exemplo, para $n = 2$, o sistema de krigagem simples constitui-se de duas incógnitas (λ_1, λ_2), a saber:

$$\begin{cases} \lambda_1 Cov_{11} + \lambda_2 Cov_{12} = Cov_{10} \\ \lambda_1 Cov_{21} + \lambda_2 Cov_{22} = Cov_{20}. \end{cases}$$

A correspondente variância minimizada do erro, denominada variância de krigagem

simples σ_{KS}^2 , é dada por (JOURNEL; HUIJBREGTS, 1978)

$$\begin{aligned}\sigma_{KS}^2 &= \text{Var}[Z_{s_0} - Z_{s_0}^*] \\ &= \text{Cov}(0) - \sum_{i=1}^n \lambda_i \times \text{Cov}(s_i, s_0).\end{aligned}\quad (2.40)$$

Em notação matricial, o sistema de krigagem é escrito como:

$$\mathbf{K} \times \boldsymbol{\lambda} = \mathbf{v} \quad \Rightarrow \quad \boldsymbol{\lambda} = \mathbf{K}^{-1} \times \mathbf{v}, \text{ com}$$

$$\mathbf{K} = \begin{bmatrix} \text{Cov}_{11} & \text{Cov}_{12} & \text{Cov}_{13} & \cdots & \text{Cov}_{1n} \\ \text{Cov}_{21} & \text{Cov}_{22} & \text{Cov}_{23} & \cdots & \text{Cov}_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \text{Cov}_{n1} & \text{Cov}_{n2} & \text{Cov}_{n3} & \cdots & \text{Cov}_{nn} \end{bmatrix}, \quad \boldsymbol{\lambda} = \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \end{bmatrix} \quad e \quad \mathbf{v} = \begin{bmatrix} \text{Cov}_{10} \\ \text{Cov}_{20} \\ \vdots \\ \text{Cov}_{n0} \end{bmatrix}$$

em que \mathbf{K} e \mathbf{v} são matrizes das covariâncias e $\boldsymbol{\lambda}$ o vetor dos pesos. A variância de krigagem simples é definida por

$$\sigma_{KS}^2 = \text{Cov}(0) - \boldsymbol{\lambda}^T \times \mathbf{v}.\quad (2.41)$$

Krigagem Ordinária

A krigagem ordinária é a forma mais simples de krigagem e pode responder satisfatoriamente à maioria dos problemas de estimativas. Exige que o modelo obedeça algumas condições, como: estacionariedade intrínseca e que as médias verdadeiras dos dados sejam constantes, porém desconhecidas (OLIVEIRA; SERIGATTO, 2004).

Considerando-se uma superfície Z , em n pontos distintos, com coordenadas representadas pelo vetor s , define-se um conjunto de valores $Z(s_i)$, $i = 1, \dots, n$, em que s_i identifica uma posição. Supondo-se que se quer estimar o valor de Z em um ponto s_0 , ou seja, $\hat{Z}_{(s_0)}$, este valor pode ser obtido, de acordo com Deutsch e Journel (1992), a partir de uma combinação linear dos n valores observados, adicionado a um parâmetro λ_0 , como se segue:

$$Z_{s_0}^* = \lambda_0 + \sum_{i=1}^n \lambda_i Z(s_i).\quad (2.42)$$

Deseja-se um estimador não tendencioso, isto é:

$$E[Z_{s_0} - Z_{s_0}^*] = 0. \quad (2.43)$$

Esta igualdade implica que as esperanças de Z_0 e $Z_{s_0}^*$ sejam iguais. Aplicando-se a Equação (2.41) em (2.42), obtêm-se:

$$E[Z_{s_0}] = E[\lambda_0 + \sum_{i=1}^n \lambda_i \times Z(s_i)] \Rightarrow \mu = \lambda_0 + \sum_{i=1}^n \lambda_i \times \mu \quad (2.44)$$

Diferente da krigagem simples, a krigagem ordinária não requer o prévio conhecimento da média μ . Assim, para que a igualdade da Equação (2.43) seja satisfeita é necessário que:

$$\lambda_0 = 0 \quad e \quad \sum_{i=1}^n \lambda_i = 1,$$

Portanto, o estimador de krigagem ordinária é:

$$Z_{s_0}^* = \sum_{i=1}^n \lambda_i \times Z(s_i), \text{ com } \sum_{i=1}^n \lambda_i = 1.$$

Minimizando a variância do erro ($Var[Z_{s_0} - Z_{s_0}^*]$) na condição de $\sum_{i=1}^n \lambda_i = 1$, os pesos λ_i são obtidos a partir do seguinte sistema de equações, denominado sistema de krigagem ordinária:

$$\begin{cases} \sum_{j=1}^n \lambda_j \times Cov(s_i, s_j) - \alpha = Cov(s_i, s_0) \text{ para } i = 1, \dots, n \\ \sum_{j=1}^n \lambda_j = 1 \end{cases} \quad (2.45)$$

em que $Cov(s_i, s_j)$ e $Cov(s_i, s_0)$ são, respectivamente, a semivariância entre os pontos s_i, s_j e entre os pontos s_i, s_0 ; α é o multiplicador de Lagrange necessário para a minimização da variância do erro.

A correspondente variância minimizada do erro, denominada variância de krigagem ordinária σ_{K0}^2 , é definida pela seguinte expressão (DEUTSCH; JOURNAL, 1992):

$$\sigma_{K0}^2 = Var[Z_{s_0} - Z_{s_0}^*] = Cov(0) - \sum_{i=1}^n \lambda_i \times Cov(s_i, s_0) - \alpha \quad (2.46)$$

O sistema de krigagem ordinária Equação (2.48) pode ser descrito em notação matricial como:

$$\mathbf{K} \times \boldsymbol{\lambda} = \mathbf{v} \quad \Rightarrow \quad \boldsymbol{\lambda} = \mathbf{K}^{-1} \times \mathbf{v}, \text{ com}$$

$$\mathbf{K} = \begin{bmatrix} Cov_{11} & Cov_{12} & Cov_{13} & \cdots & Cov_{1n} & 1 \\ Cov_{21} & Cov_{22} & Cov_{23} & \cdots & Cov_{2n} & 1 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ Cov_{n1} & Cov_{n2} & Cov_{n3} & \cdots & Cov_{nn} & 1 \\ 1 & 1 & 1 & \cdots & 1 & 0, \end{bmatrix}, \boldsymbol{\lambda} = \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \\ \alpha \end{bmatrix} \text{ e } \mathbf{v} = \begin{bmatrix} Cov_{10} \\ Cov_{20} \\ \vdots \\ Cov_{n0} \\ 1, \end{bmatrix}$$

em que \mathbf{K} e \mathbf{v} são matrizes das covariâncias e $\boldsymbol{\lambda}$ o vetor dos pesos.

A variância de krigagem ordinária, na forma matricial, é dada por:

$$\sigma_{K0}^2 = Cov(0) - \boldsymbol{\lambda}^T \times \mathbf{v}. \quad (2.47)$$

2.6 Espécies vegetais estudadas

***Cenchrus ciliaris* L. (Capim buffel)**

Várias espécies de gramíneas têm sido avaliadas para a formação de pastagens, buscando-se elevada produtividade e persistência, entre elas, o capim-buffel (*Cenchrus ciliaris* L.) tem se mostrado altamente adaptado à seca, associando rápida germinação e estabelecimento com precocidade na produção de sementes e capacidade de entrar em dormência no período seco ARAÚJO (1986), é uma gramínea forrageira, que apresenta excelente palatabilidade para os animais.

A implantação deste capim cultivado em zonas semiáridas do Nordeste, tem demonstrado, de um modo geral, que é mais fácil estabelecer o capim buffel em áreas de caatinga recém-desmatadas. O desmatamento manual com queima uniforme, tem revelado ser mais eficiente e de menor custo quando comparado ao desmatamento mecânico, principalmente pela conservação da camada mais superficial do solo (INFOPIEDIA, 2003).

Segundo Infopedia (2003), o plantio normalmente é realizado por sementes, sendo de modo geral manual, coloca-se em torno de 70 sementes por metro linear. No plantio manual gastam-se 10 a 15 quilos de sementes para plantar um hectare. Para se obter uma boa germinação, as sementes de buffel devem ser plantadas pelo menos seis meses após colhidas, em razão da dormência fisiológica que apresentam. Normalmente, os produtores utilizam as sementes colhidas no ano anterior para o plantio de novas áreas.

A adubação química deve ser realizada de acordo com o resultado da análise de fertilidade do solo. No entanto, estudos demonstraram que o uso da adubação à base de fósforo, acelera o crescimento do sistema radicular das plantas novas, o que é bastante desejável nas regiões semiáridas. No caso da adubação orgânica, a quantidade recomendada depende da disponibilidade de esterco na propriedade (JUNIOR; SILVA, 2012).

Conforme Junior e Silva (2012), o controle de ervas daninhas pode ser feito manual, mecânico ou químico. No controle manual é comum o uso da enxada, roçadeira, foice, estrovenga e chibanca, entre outros. No controle mecânico é utilizada a roçadeira acoplada ao trator, e no químico, utilizam-se herbicidas específicos. A utilização de fogo controlado pode ser realizada a cada três ou quatro anos, em áreas de topografia plana à suavemente ondulada.

Como toda cultura, a produtividade de Capim buffel varia de acordo com as condições do clima e do solo da região. No entanto, produtividades de 4 a 12 toneladas de matéria seca/ha/ano têm sido obtidas no semiárido nordestino. Junior e Silva (2012) mostra que além do pasto direto, o buffel pode ser também utilizado na forma de feno na alimentação de animais. O melhor feno é obtido quando a parte aérea das plantas encontra-se com cerca de 35 dias. Neste período o teor de proteína bruta na planta apresenta-se mais elevado. Nas regiões tropicais, as pastagens formadas com o capim buffel podem proporcionar ganho de peso diário variando de 200 a 400 g/animal.

A preservação total ou parcial do capim buffel durante o período chuvoso, associado ao uso da caatinga, é uma prática que tem revelado bons resultados. Essa associação, além de preservar a caatinga, possibilita o aproveitamento do seu elevado potencial forrageiro durante a época chuvosa (JUNIOR; SILVA, 2012).

***Spondias tuberosa* Arruda (Umbuzeiro)**

O umbuzeiro é uma fruteira nativa da caatinga de grande importância para região. Pertence a Família das *Anacardiáceas*, xerófila e endêmica do semiárido nordestino. Tem uma capacidade de sobrevivência em períodos de seca, suas raízes superficiais exploram aproximadamente 1m de profundidade, isso ocorre devido ao fato de que o umbuzeiro retira água e nutrientes de suas batatas ou xilopódios que é constituído de tecido lacunoso que armazena água, mucilagem, glicose, tanino, amido, ácidos, entre outras substâncias. Seu porte pode alcançar mais de 7m de altura com copa medindo até 22m de diâmetro projetando sombra densa sobre o solo, vive em média 100 anos e pode até armazenar dois mil litros de água em suas raízes. Seu tronco é atrofiado e retorcido com diâmetro de 0,30 a 1,39m (FRANCISCO, 2012b).

Originário dos chapadões semiáridos do Nordeste brasileiro, nas regiões do Agreste (Piauí), Cariris (Paraíba) e Caatinga (Pernambuco e Bahia), a planta encontrou boas condições para seu desenvolvimento encontrando-se, em maior número, nos Cariris Velhos, seguindo desde o Piauí à Bahia e até norte de Minas Gerais.

Conforme Francisco (2012b), o caule, com casca cor cinza, tem ramos novos lisos e ramos velhos com ritidomas (casca externa morta que se destaca) as folhas são verdes, as flores são brancas e perfumadas. A frutificação inicia-se em período chuvoso. Os frutos do umbuzeiro são drupas lisas ou levemente pilosas e arredondados, com peso variando de 5,5 a 130 gramas, possui um diâmetro médio 3,0cm, forma arredondada a ovalada, é constituída por casca, polpa e caroço. Sua polpa é quase aquosa quando madura. Sua semente segue a forma do fruto, peso de 1,0 a 2,0 gramas e seu diâmetro entre 1,2 a 2,4cm, quando despulpada. Ele é muito perecível, é utilizado na fabricação de polpa para sucos, doces, geléias, de vinho, de vinagre, de acetona, de concentrado para sorvete e uma grande variedade de produtos.

O umbuzeiro perde totalmente as folhas durante a época seca e reveste-se de folhas após as primeiras chuvas. A floração pode iniciar-se após as primeiras chuvas independentemente da planta estar ou não enfolhada, a abertura das flores dá-se entre 0:00 horas e 4:00 horas (com pico às 2:00 horas). Cerca de 60 dias após a abertura da flor o fruto estará maduro (FRANCISCO, 2012b).

Segundo Francisco (2012b), o umbuzeiro requer clima quente, temperatura entre 12°C e 38°C e 400mm a 800mm de chuva (entre novembro e fevereiro), podendo viver em locais com chuvas de 1.600mm/ano. Vegeta bem em solos não úmidos, profundos, bem drenados.

***Prosopis juliflora* Swartz DC. (Algaroba)**

É uma árvore que cresce razoavelmente bem nos desertos americanos e em alguns africanos, sendo uma espécie xerófila. Introduzida no Brasil por volta de 1942, no município de Serra Talhada, sertão de Pernambuco. Dessa primeira tentativa não se tem informações sobre o sucesso, o que demonstra que há fortes indícios de que tenha fracassado. Adaptou-se muito bem no Nordeste brasileiro. A algarobeira é uma planta seletiva, pouco exigente em água, cuja ocorrência, em sua forma natural, se dá em zonas tropicais áridas, que não chegam a alcançar índices de 100mm (SILVA, 2012).

Essa característica é de extrema importância para o nordeste brasileiro, uma vez que a precipitação pluviométrica média anual dessa região gira em torno de 750 mm e, embora seja baixa para outras espécies vegetais, já é 7,5 vezes maior do que essa espécie

necessita para ocorrer (SILVA, 2012).

Conforme Silva (2012), devido a essa pequena exigência em água, comprovada capacidade de medrar em solos de baixa fertilidade e de condições físicas imprestáveis a outras culturas, evidencia-se as grandes potencialidades desta leguminosa como fonte geradora de alimentos para o homem e para os animais, constituindo-se em importante fonte de desenvolvimento para as regiões áridas e semiáridas do planeta. Alcança melhor desenvolvimento em solos de aluvião, iniciando boa produção a partir do 5^a ano de plantio. O fruto é uma vagem do tipo achatada e mais ou menos curva, com média de 20 cm de comprimento, porém foram encontrados frutos com até 36 cm de comprimento no Cariri paraibano.

Segundo Campo (2012), da vagem da algaroba se obtém uma farinha que pode ser usada na alimentação humana. Vários outros produtos alimentícios podem ser obtidos da algaroba como o mel, licor e "café". O gênero *Prosopis* tem múltiplos usos, como na produção de madeira, tanino, gomas, tinturaria e produz vagem com ampla utilização na alimentação animal.

Sendo as espécies adaptadas a climas áridos e semiáridos, a germinação, geralmente acontece durante a estação chuvosa. A espécie possui dois sistemas de raiz, uma raiz funda e um tapete de raízes laterais onde absorve águas de chuvas.

Sua raiz principal chega a atingir grandes profundidades, pois, sendo uma planta de regiões áridas, busca encontrar o lençol freático, retirando água do subsolo para a superfície, pode sobreviver em áreas com baixa precipitação anual ou períodos secos muito prolongados.

Suas raízes podem captar água do solo ou outras fontes de água permanentes dentro dos primeiros anos. E tem sido pelo as suas múltiplas aplicações e usos, além de outras características importantes, que a algarobeira é reconhecidamente no meio rural nordestino, como "Planta Mágica", de valor precioso para o nordestino e tem sido recomendada por conceituados pesquisadores e técnicos da área para a região do polígono das secas (SILVA, 2012).

Conforme Campo (2012), no reflorestamento a algarobeira é uma excelente essência florestal para o semiárido do nordeste. Importante para evitar a erosão e a desertificação, se cultivada de forma planejada e orientada, resiste à seca, é de fácil fixação, cresce rápido, produz madeira de qualidade, lenha, carvão vegetal, mourões, ripas, caibros, dormentes, entre outras coisas, no entanto o mais importante é a função de fertilizar o solo através do nitrogênio do ar.

***Zizyphus joazeiro* Mart. (Juazeiro)**

É uma espécie de árvore abundante no Nordeste brasileiro. Possui copa larga e alta, adapta-se bem em clima quente, vive em terras semiúmidas, semiáridas e cresce melhor em terrenos mais úmidos onde pode chegar aos quinze metros de altura. É uma planta que não perde totalmente as folhas durante o ano, suas folhas são verdes, brilhosas, as bordas são serrilhadas e podem chegar a dez centímetros. As flores são pequenas, tem uma coloração amarela para verde (FRANCISCO, 2012a).

Segundo Francisco (2012a), os frutos são muito apreciados pelos sertanejos em época de fome, são pequenos, arredondados, amarelos quando maduros, sua polpa é esbranquiçada, são adocicados e ricos em vitamina C, sendo consumidos por aves, animais e pelo homem. Dos frutos secos, pode ser feito o vinho moscatel, os ramos servem de alimentos para ovinos, bovinos e caprinos em qualquer época, essa planta deve ser utilizada apenas como recurso alimentar alternativo durante a época seca, no período de maior escassez, as ramas de juazeiro são ricas em proteína digestível, em hidratos de carbono e até em celulose digestível.

Conforme Francisco (2012a), esta planta tem sido empregada na medicina popular. Cascas e folhas são usadas internamente, por via oral para alívio de problemas como expectorante, no tratamento de bronquites, tosses e de úlceras gástricas, externamente, para limpeza dos cabelos, como tônico capilar anticaspa e para clarear a pele do rosto.

Nos dentes, usando um pouco do pó que se prende à escova de dente molhada, serve como creme dental apresenta efeito mais eficaz na diminuição da placa dental do que os dentífrícos convencionais, desestabilizando a placa dental e exercendo uma ação antimicrobiana sobre *Streptococcus mutans*, principal germe causador da cárie dentária. As folhas e as cascas, quando agitadas com água produzem abundante espuma devido a sua propriedade espumígena. Os resultados de ensaios farmacológicos revelam uma atividade antifebril em coelhos usados como animais de experiência (FRANCISCO, 2012a).

De todas as árvores do Nordeste brasileiro, o juazeiro é a planta símbolo da caatinga. É uma planta perfeitamente adaptada ao clima seco, com nítida preferência por solos férteis de várzeas e beira de rios. Aparece de maneira espontânea no Piauí, Ceará, Rio Grande do Norte, Paraíba, Pernambuco, Alagoas, Sergipe, Bahia e norte de Minas Gerais (SILVA *et al.*, 2011).

Conforme Francisco (2012a), é uma árvore de crescimento vagaroso e de vida longa, podendo passar de 100 anos. Uma das poucas árvores da caatinga que não perde as folhas durante a estação seca. Como a sua floração ocorre nos meses mais secos do ano

(novembro e dezembro), quando a maioria das espécies da caatinga encontra-se desfolhada e sem flores ela é quase a única espécie a fornecer néctar às abelhas, também é bastante utilizado como madeira para marcenaria e construções rurais devido a sua durabilidade e resistência.

3 Material e métodos

3.1 Área de estudo

O estudo foi realizado em 1996, em uma fazenda de gado de corte pertencente à Agropecuária Jaçanã, situada na Cidade de Custódia, PE (8°13.70' ao Sul e 37°44.70' a Oeste). A precipitação média de chuva no local é 740 milímetros e a temperatura média anual é de 26° C (dados a partir dos registros da fazenda). A vegetação dentro da área da Agropecuária, inicialmente consistia em uma vegetação, de caatinga semidecídua seca, a qual foi desmatada por tratores em 1984, e em seguida foram retirados os entulhos como tocos e pedras que porventura ficaram no local dentre os 5.000 *ha* (MENEZES; SALCEDO, 1999).

Conforme Tiessen *et al.* (2003), em uma área da fazenda com aproximadamente 3.000 *ha*, foram plantadas no ano seguinte as árvores *Prosopis juliflora* Swartz DC. (Algarobas), a uma distância de 10m por 10m, consorciadas com *Cenchrus ciliares* L. (Capim buffel). O pasto nunca recebeu adubação, aplicação de agrotóxicos e nem foram registrados incêndio no local.

Em outra área da fazenda, com cerca de 2.000 *ha*, foram preservadas algumas espécies nativas da caatinga como: *Ziziphus joazeiro* Mart., (Juazeiro) e *Spondias tuberosa* Arruda (Umbuzeiro), com uma distância média entre as árvores variando cerca de 30m a 40m. A idade das árvores preservadas neste local foram estimadas em pelo menos 50 anos de idade (Dados levantados junto aos trabalhadores da fazenda).

Segundo Tiessen *et al.* (2003), as espécies arbóreas nativas representam situações diferentes em termos de diferenças de potencial no solo e características herbáceas em cada sistema, uma vez que as árvores *P. juliflora* foram plantadas ao mesmo tempo das pastagens. Enquanto que as árvores nativas já estavam presentes quando a pastagem foi estabelecida. A lotação de animais, nos pastos da fazenda situava-se, na época do estudo, em torno de 0,17% animais por *ha*, bem abaixo da densidade normal em pastagens de capim buffel na região.

As espécies analisadas são:

- i) *Spodias tuberosa* (Umbuzeiro);
- ii) *Ziziphus joazeiro* (Juazeiro ou Juá);
- iii) *Prosopis juliflora* (Algaroba).

Segundo Menezes e Salcedo (1999), dentro das áreas de estudo, foram selecionados três indivíduos de *Z. joazeiro*, *P. juliflora* e *S. tuberosa*. Deu-se preferência por árvores maduras isoladas uma das outras. Para cada árvore, estabeleceram-se 81 pontos de amostragem regularmente distribuídos, onde a espécie arborea sempre está no centro da área.

A coleta de dados foi realizada por pesquisadores e estudantes do Departamento de Energia Nuclear da UFPE. As amostras de solo foram coletadas na profundidade de 0-15 cm, com as devidas coordenadas X e Y , as amostras foram secas ao ar e passadas numa peneira de 2mm, analisou-se quanto ao total de carbono orgânico (C). Do mesmo modo, as amostras de biomassa de herbáceas, foram tomadas em cada posição da grade (0,7cm x 0,7cm), toda a biomassa viva em pé foi cortada ao nível do solo e colocadas em sacos de papel, as amostras foram secas em estufa durante 48 horas a $60^{\circ}C$ (MENEZES; SALCEDO, 1999).

3.2 Software

Após a coleta dos dados e das análises realizadas em laboratório, utilizou-se o software Microsoft Excel 2007 para a organização dos dados. A variável carbono está referenciada em g/kg , no momento da análise no software R, teve-se a necessidade de obter o % dos dados, dividiu-se então o resultado por 10. Para a variável biomassa de herbáceas, referenciou-se em m^2 . Por não haver nenhum valor de referência que pôde-se comparar aos valores observados, optou-se em utilizar a média dos dados.

Baixou-se o programa estatístico gratuito R por ser muito rico em ferramentas estatísticas e que pode ser encontrado pelo site www.r-project.org, *version 2.15.0*, para assim obter os resultados das análises estatísticas descritivas e espacial, alguns pacotes são necessários para a obtenção dos parâmetros, das funções de correlações e dos respectivos envelopes simulados como o geoR, consta no apêndice a programação utilizada.

Retorna-se ao Microsoft Excel, já com os valores dos parâmetros adquiridos no software R, obteve-se os resultados do índice de dependência espacial. Ao terminar as análises estatísticas e as imagens projetadas pelo mesmo, as palavras que são em inglês foram retiradas e inseridas nas imagens em português, este procedimento foi realizado no Microsoft Paint.

3.3 Valores de referência - Teor do solo

O solo no bioma caatinga para ser considerado como fértil, há valores de referências para fazerem-se comparações com os valores observados. O material do solo será considerado como orgânico, quando o teor de carbono for igual ou superior a $80g/kg$, avaliado na fração terra fina seca ao ar Embrapa (2009), como dito anteriormente os valores de referências foram divididos por 10, são assim descritos: baixo teor de fertilidade 0,8%, médio teor de fertilidade¹ 0,9% a 1,4% e alto teor de fertilidade $\geq 1,4\%$ (NORDESTE, 2007).

¹Analisou-se a média do considerado médio teor de fertilidade.

4 Resultados e discussão

Inicia-se a análise exploratória da estatística descritiva e espacial do teor de fertilidade do solo, através da variável carbono e o crescimento de Capim buffel, através da variável biomassa de herbáceas, em áreas pertencentes ao bioma caatinga.

4.1 Juazeiro

Variável Carbono

Apresenta-se na Tabela 1 a estatística descritiva, os valores de média e mediana são aproximadamente iguais. Com o afastamento da variável em relação a um valor central, tem-se então uma distribuição assimétrica à direita, apresentou-se uma curva com uma distribuição platicúrtica, diferindo da curva normal (mesocúrtica). De acordo com o critério proposto por Wilding e Dress (1983), o coeficiente de variação classifica a variabilidade como muito alta $CV > 30\%$. O valor mínimo e máximo indicam a não existência de problemas amostrais com os dados.

Tabela 1: Estatística descritiva: mínimo, máximo, média, mediana, desvio padrão (DP), coeficiente de variação (CV), curtose, assimetria de amostras de solo da variável carbono da espécie juazeiro.

| Mínimo | Máximo | Média | Mediana | DP | CV | Curtose | Assimetria |
|--------|--------|-------|---------|-------|--------|---------|------------|
| 0,240 | 1,217 | 0,687 | 0,657 | 0,217 | 31,551 | 2,683 | 0,475 |

O gráfico do canto superior esquerdo da Figura 3 representa os dados nos quartis da variável carbono, ou seja, " + " 1º quartil, " Δ " 2º quartil, " o " 3º quartil e " × " 4º quartil, essa imagem condiz com a ideia de que existe padrão espacial na variável. O gráfico do canto superior direito e do canto inferior esquerdo, apresentam a dispersão dos dados em torno da média, observa-se que os dados estão dispersos. Com base no gráfico do canto inferior direito, a análise exploratória do pressuposto de gaussianidade da resposta foi analisado, o qual mostra a densidade amostral dos dados desconsiderando-se o possível padrão

espacial. Esta inspeção não garante a tal hipótese mas serve para mostrar desvios óbvios, caso existam. Como existe uma fuga da distribuição gaussiana nos resíduos, necessita-se de possíveis transformações na variável, é através do Box-Cox que encontra-se o λ , que fará com que esta transformação seja possível, como também os intervalos de confiança.

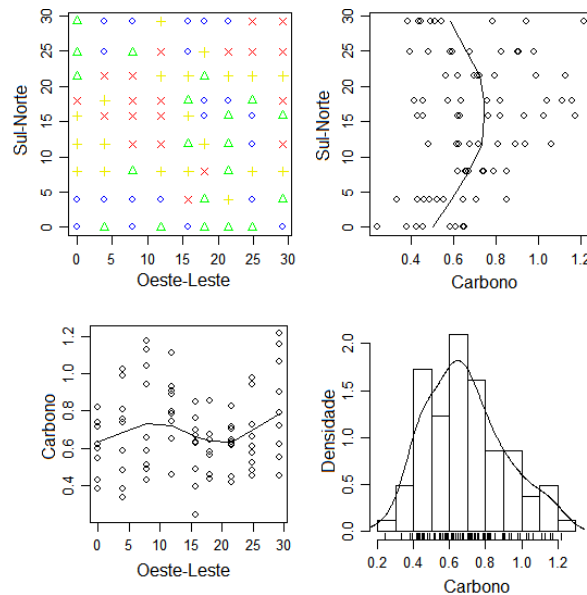


Figura 3: Gráfico descritivo do padrão espacial do juazeiro, variável carbono, com os dados originais.

A transformação ótima de *Box-Cox* (lado esquerdo) da Figura 4, por meio do logaritmo da função da verossimilhança, mostra qual o valor do parâmetro que mais aproxima os resíduos à distribuição normal. Como os intervalos de confiança $[-0,24; 0,94]$ para λ não contém o valor 1, é preciso utilizar uma transformação na variável em estudo, neste caso, para a variável, foi encontrado o valor de $\lambda = 0,35$.

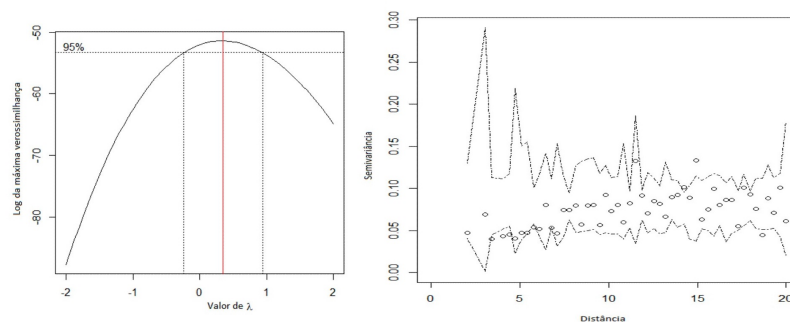


Figura 4: Gráfico de possíveis transformações de variáveis por meio da transformação ótima de *Box-Cox* (Esquerdo) e a verificação de dependência espacial por meio do envelope simulado (Direito) da variável carbono da espécie juazeiro.

Após a transformação dos dados de carbono, percebe-se na Figura 5 que os dados

estão menos dispersos em torno da média que o gráfico da figura 3, gráficos estes localizados no canto superior direito e no canto inferior esquerdo. A análise exploratória do pressuposto de gaussianidade da resposta foi analisado, para o gráfico do canto inferior direito, percebe-se uma mudança em relação ao anterior.

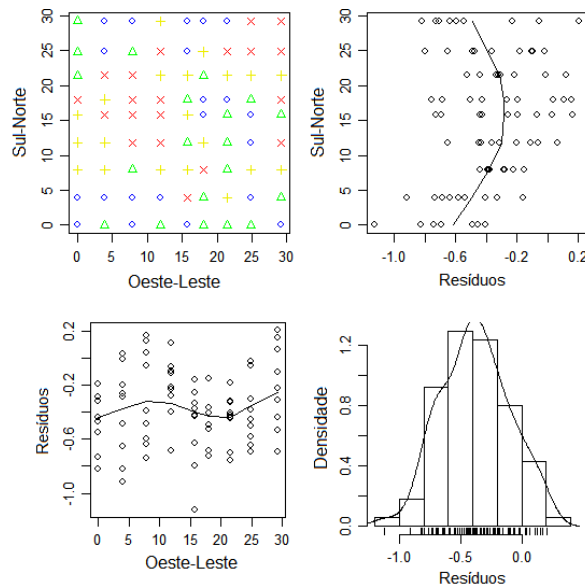


Figura 5: Gráfico descritivo do padrão espacial do juazeiro, variável carbono, com os dados transformados.

A existência de dependência espacial pode ser verificada por meio do envelope simulado. A imagem do lado direito da Figura 4 mostra alguns pontos fora do intervalo, indicando assim uma dependência espacial do carbono no solo, vale salientar que foram utilizadas 1.000 simulações. Para que haja dependência espacial, deve haver ao menos um ponto do variograma fora do envelope simulado, deste modo, se isso ocorrer rejeita-se a hipótese nula (H_0), de que não há dependência espacial (DIGGLE; RIBEIRO, 2007).

As dependências espaciais observadas para os valores de carbono mostram que as análises da estatística clássica, em que as observações são consideradas aleatórias e independentes espacialmente, devem ser substituídas por análises espaciais, que levam em consideração as relações entre observações vizinhas.

Na Tabela 2, observa-se as estimativas da máxima verossimilhança para os parâmetros dos modelos ajustados, considerando-se a média constante sob a região de estudo. O estimador do parâmetro $\hat{\beta}$ apresenta valores próximos para as funções de correlações: Esférico, Gaussiano e Circular. As funções *Matérn* com *kappas* (0,5) e (1,0), demonstram valores discrepantes, comparados com os demais.

Para todas as características avaliadas, o modelo que ajustou melhor ao semivario-

Tabela 2: Estimativa dos parâmetros associados aos modelos por meio da máxima verossimilhança, assumindo a média da variável carbono constante sobre a região de estudo na espécie juazeiro.

| Modelos | $\hat{\beta}$ | $\hat{\tau}^2$ | $\hat{\sigma}^2$ | $\hat{\phi}$ | <i>AIC</i> | <i>BIC</i> | IDE |
|------------|---------------|----------------|------------------|--------------|------------|------------|-----|
| Matérn 0,5 | -0,415 | 0,018 | 0,068 | 7,423 | -39,82 | -30,24 | 21% |
| Matérn 1,0 | -0,412 | 0,027 | 0,060 | 5,194 | -41,03 | -31,45 | 31% |
| Esférico | -0,398 | 0,023 | 0,055 | 14,878 | -43,06 | -33,49 | 30% |
| Gaussiano | -0,398 | 0,033 | 0,051 | 7,840 | -43,98 | -34,40 | 39% |
| Circular | -0,392 | 0,024 | 0,054 | 12,348 | -43,37 | -33,80 | 31% |

grama experimental para a variável carbono onde está inserida o juazeiro foi a função de correlação gaussiano indicando o menor valor de *AIC* e *BIC* para todos os modelos espaciais. Na Figura 6, encontra-se o semivariograma experimental ajustado conforme o método da máxima verossimilhança.

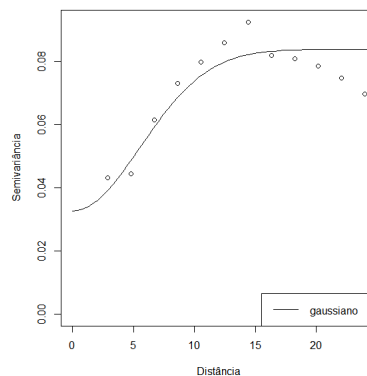


Figura 6: Modelo ajustado ao semivariograma conforme o método da máxima verossimilhança para a variável carbono da espécie juazeiro.

Para verificar o grau de dependência espacial entre as características de fertilidade de solo, utilizou-se o efeito pepita τ^2 sobre o patamar $\sigma^2 + \tau^2$. Obteve-se então, um grau de dependência espacial moderada 39%, conforme a classificação de Cambardella *et al.* (1994). Nota-se que a variável carbono apresenta dependência espacial, que pode ser descrita pelo modelo Gaussiano com alcance de aproximadamente 7,840m, ou seja, amostras de carbono de solo selecionadas a distâncias inferiores são espacialmente dependentes.

Conforme visto anteriormente, a krigagem é uma metodologia utilizada para estimar o valor de uma variável de interesse em um local onde não foi possível fazer a observação. É obtida por meio de uma interpolação, que utiliza a correlação existente entre valores dos dados obtidos em pontos máximos. Com os modelos finais estabelecidos, foram calculadas as previsões espaciais para o índice de fertilidade de solo em uma malha de 81

localizações espaciais, sendo que as estimativas paramétricas foram substituídas nas fórmulas de krigagem. Uma análise importante que se dispõe com a utilização dos mapas preditivos de krigagem são os de probabilidades condicionais marginais, pois eles podem prever a partir de um determinado valor numérico as chances de ocorrência de um evento na região de estudos.

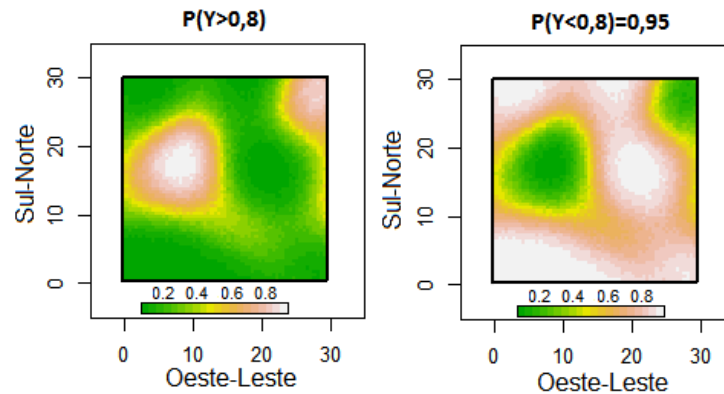


Figura 7: Comparação dos valores observados com o valor de referência para solos com baixo teor de carbono, na espécie juazeiro.

Analisando-se a Figura 7 percebe-se que as amostras ficaram condizentes com os dados observados, ou seja, localizações com $P > 0,8\%$ possuem entre 80 – 100% de carbono no solo (parte clara da imagem esquerda).

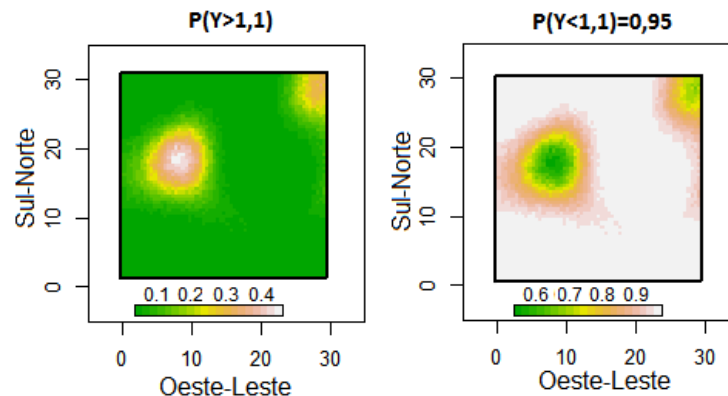


Figura 8: Comparação dos valores observados com o valor de referência para solos com médio teor de carbono, na espécie juazeiro.

Seguindo a ideia do gráfico anterior, a Figura 8 (lado esquerdo) com $P > 1,1\%$ o "verde" é a cor predominante com aproximadamente 10% de concentração de carbono dentro da área.

A Figura 9 representa o mapa preditivo, a imagem do lado direito, confirma através da cor mais clara, que há concentração de carbono em quase toda região com $P < 1,4\%$, ou

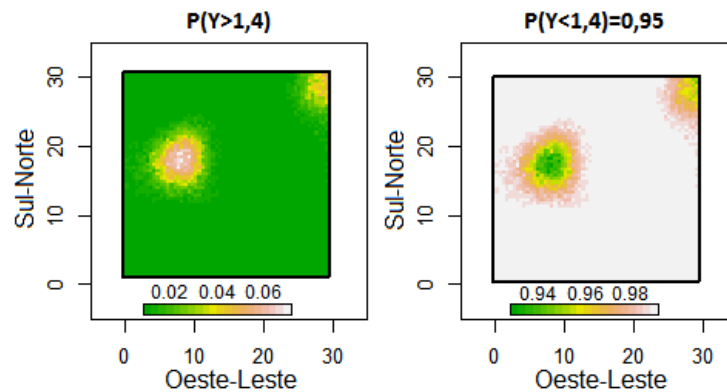


Figura 9: Comparação dos valores observados com o valor de referência para solos com alto teor de carbono, na espécie juazeiro.

seja, entre 98 – 100% da área.

Variável Biomassa de Herbáceas

Pode-se observar por meio da Tabela 3 a estatística descritiva. Não houve um bom comportamento das estimativas dos parâmetros de média, há valores discrepantes, uma vez que os valores de média e mediana não se aproximam. Com relação ao afastamento da variável a um valor central (curtose), tem-se então uma distribuição assimétrica à direita, a curva para esta variável tem-se uma distribuição leptocúrtica, de acordo com o critério proposto por Wilding e Dress (1983), o coeficiente de variação classifica a variabilidade como muito alta $CV > 30\%$. O valor mínimo e máximo indicam a não existência de problemas amostrais com os dados.

Tabela 3: Estatística descritiva: mínimo, máximo, média, mediana, desvio padrão (DP), coeficiente de variação (CV), curtose, assimetria de amostras de solo da variável biomassa de herbáceas da espécie juazeiro.

| Mínimo | Máximo | Média | Mediana | DP | CV | Curtose | Assimetria |
|--------|--------|---------|---------|--------|--------|---------|------------|
| 7,68 | 301,04 | 101,317 | 83,92 | 60,382 | 59,597 | 4,388 | 1,177 |

Da mesma forma que analisou-se o carbono, o gráfico do canto superior esquerdo da Figura 10 representa os dados nos quartis da variável biomassa de herbáceas, ou seja, ”+” 1º quartil, ”Δ” 2º quartil, ”o” 3º quartil e ”×” 4º quartil, a imagem condiz com a ideia de que existe padrão espacial. O gráfico do canto superior direito e do canto inferior esquerdo, apresentam a dispersão dos dados em torno da média, observa-se que os dados estão dispersos. Com base no gráfico do canto inferior direito, a análise exploratória do pressuposto de gaussianidade da resposta foi analisado, o qual mostra a densidade amostral dos dados desconsiderando-se o possível padrão espacial. Esta inspeção não garante

a tal hipótese mas serve para mostrar desvios óbvios, caso existam. Como existe uma fuga da distribuição gaussiana nos resíduos, necessita-se de possíveis transformações na variável, é através do *Box-Cox* que encontra-se o λ , que fará com que esta transformação seja possível, como também os valores do intervalo de confiança.

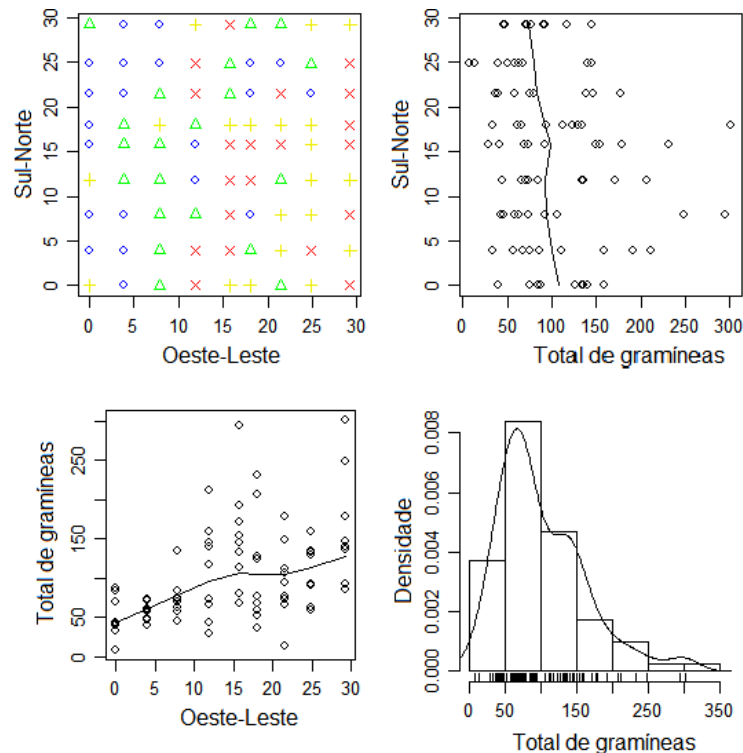


Figura 10: Gráfico descritivo do padrão espacial do juazeiro, variável biomassa de herbáceas, com os dados originais.

A transformação ótima de *Box-Cox* (lado esquerdo) da Figura 11, por meio do logaritmo da função de verossimilhança, mostra qual o valor do parâmetro que mais aproxima os resíduos à distribuição normal. Como os intervalos de confiança $[0,04; 0,58]$ para λ não contém o número 1, é preciso utilizar uma transformação na variável em estudo, neste caso, para esta variável, foi encontrado o valor de $\lambda = 0,31$.

Após os dados serem transformados, percebe-se que o gráfico do canto superior esquerdo da Figura 12 representa os quartis da variável já transformados, ou seja, ”+” 1º quartil, ” Δ ” 2º quartil, ”o” 3º quartil e ” \times ” 4º quartil, essa imagem condiz com a ideia de que existe padrão espacial na variável. O gráfico do canto superior direito e do canto inferior esquerdo, apresentam a dispersão dos dados em torno da média, observa-se que os dados estão menos dispersos que o gráfico da Figura 10. A análise exploratória do pressuposto de gaussianidade da resposta foi analisado para o gráfico do canto inferior direito, o qual mostra as densidades amostrais dos dados, desconsiderando-se o possível padrão espacial. Esta inspeção não garante a tal hipótese mas serve para mostrar desvios

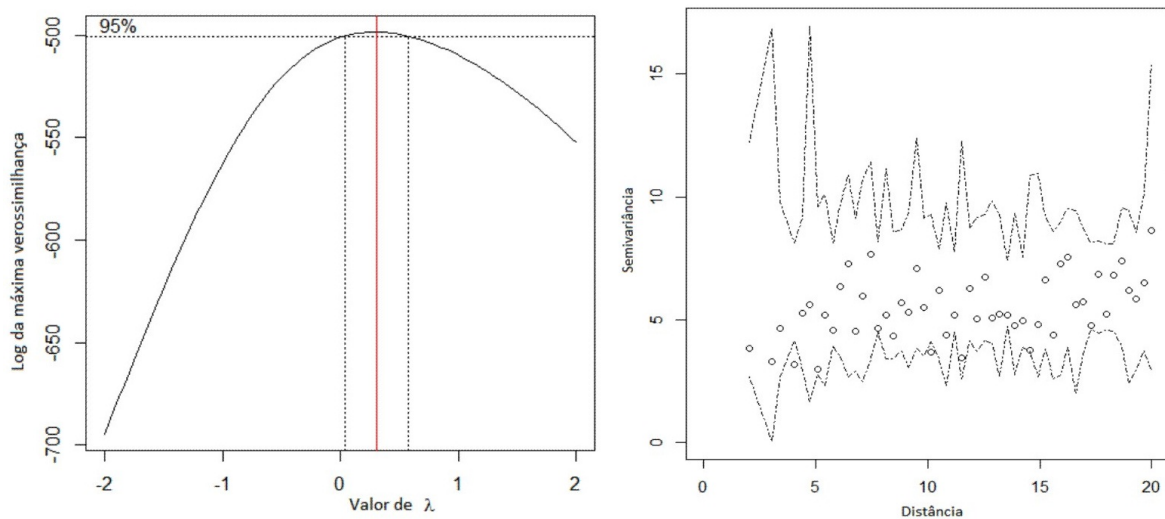


Figura 11: Gráfico de possíveis transformações de variáveis por meio da transformação ótima de *Box-Cox* (Esquerdo) e a verificação de dependência espacial através do envelope simulado (Direito) da variável biomassa de herbáceas da espécie juazeiro.

óbvios, caso existam.

A dependência espacial pode ser verificada por meio do gráfico utilizando envelopes simulados e variogramas empíricos, Figura 11 (lado direito) nessa pesquisa foram utilizadas 1.000 simulações. Alguns pontos estão fora do envelope simulado indicando assim uma dependência espacial de biomassa de herbáceas. Para que haja dependência, deve haver ao menos um ponto do variograma fora do envelope simulado, deste modo, se isso ocorrer rejeita-se H_0 , de que não há dependência espacial (DIGGLE; RIBEIRO, 2007), as análises clássicas devem ser substituídas por análises espaciais, que levam em consideração as relações entre observações vizinhas, ou seja, dependência espacial.

Tabela 4: Estimativa dos parâmetros associados aos modelos por meio da máxima verossimilhança, assumindo-se a média da variável biomassa de herbáceas constante sobre a região de estudo na espécie juazeiro.

| Modelos | $\hat{\beta}$ | $\hat{\tau}^2$ | $\hat{\sigma}^2$ | $\hat{\phi}$ | AIC | BIC | IDE |
|------------|---------------|----------------|------------------|--------------|-------|-------|-----|
| Matérn 0,5 | 9,716 | 2,078 | 4,201 | 6,938 | 859,0 | 868,6 | 33% |
| Matérn 1,0 | 9,749 | 2,541 | 3,591 | 4,095 | 859,2 | 868,6 | 41% |
| Esférico | 9,732 | 2,993 | 3,952 | 23,582 | 860,6 | 870,2 | 43% |
| Gaussiano | 5,099 | 0,324 | 0,359 | 6,119 | 862,6 | 872,2 | 47% |
| Circular | 9,782 | 2,968 | 4,221 | 20,738 | 860,7 | 870,3 | 41% |

Dando sequências as análises, pode-se observar por meio da Tabela 4 as estimativas da máxima verossimilhança para os parâmetros dos modelos ajustados considerando-se

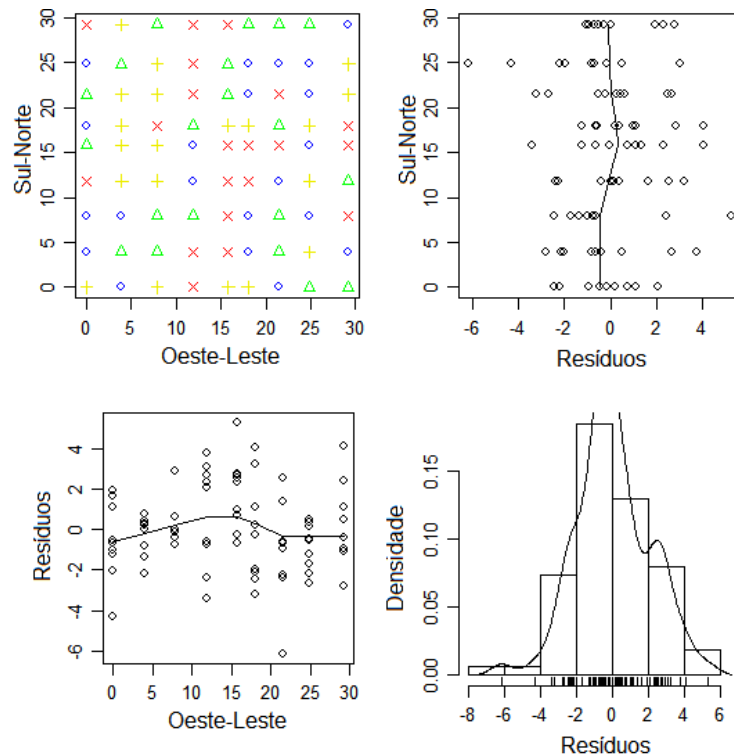


Figura 12: Gráfico descritivo do padrão espacial do juazeiro, variável biomassa de herbáceas, com os dados transformados.

a forma para a matriz de delineamento X : média constante sob a região de estudo. A estimativa do estimador do parâmetro $\hat{\beta}$ apresenta valores próximos para as funções de correlações: *Matérn* com *kappas* 0,5 e 1,0, *Esférico* e *Circular*, a função *Gaussiano*, demonstra valores discrepantes, comparados com os demais. O próximo passo é selecionar o melhor modelo geoestatístico, para todas as características avaliadas, o modelo que ajustou melhor ao semivariograma experimental, para a variável biomassa de herbáceas, onde está inserida o juazeiro, foi a função de correlação *Matérn* com *kappa* 0,5, esta função chega a ser considerada pela literatura como uma função exponencial, segundo (MELLO, 2004).

Na Figura 13 encontra-se o semivariograma experimental ajustado conforme o método da máxima verossimilhança, a análise variográfica apresentada mostrou que todas as características estudadas nos procedimentos até aqui descritos, apresentam estruturas espaciais bem definidas, ou seja, existe uma função estrutural com semivariância de comportamento modelável.

Para verificar o grau de dependência espacial, utilizou-se da relação entre o efeito pepita τ^2 dividido pelo patamar $\sigma^2 + \tau^2$. Pela classificação de Cambardella *et al.* (1994), quando a relação está entre $25\% < IDE < 75\%$, há uma moderada dependência espacial

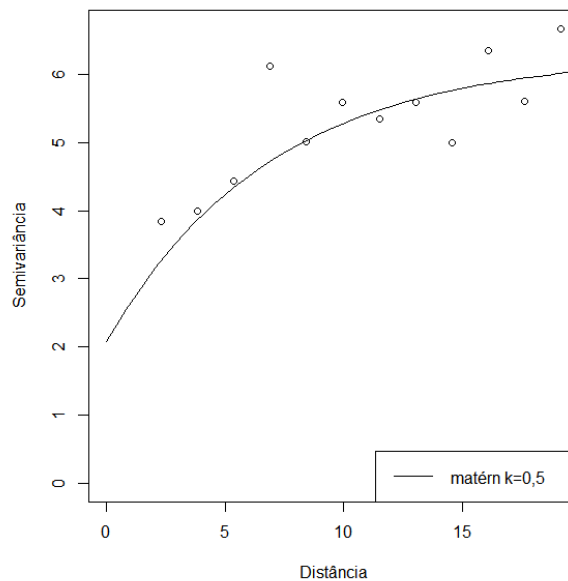


Figura 13: Modelo ajustado ao semivariograma conforme o Método de Máxima Verossimilhança para a variável biomassa de herbáceas da espécie juazeiro.

33%. Nota-se que a variável biomassa de herbáceas apresenta dependência espacial, que pode ser descrita pelo modelo *Matérn* com κ 0,5 com alcance de aproximadamente de $6,938m$, ou seja, amostras de biomassa de herbáceas selecionadas a distâncias inferiores são espacialmente dependentes.

Os resultados obtidos demonstraram que há dependência espacial de biomassa de herbáceas em função da presença dos indivíduos de juazeiro.

Conforme visto anteriormente, a krigagem é uma interpolação. Com os modelos finais estabelecidos, foram calculadas as previsões espaciais para o efeito de crescimento de capim no solo em um malha de 81 localizações espaciais, sendo que as estimativas paramétricas foram substituídas nas fórmulas de krigagem.

Pode-se observar por meio da Figura 14 o mapa preditivo do campo aleatório para o efeito de crescimento de capim. Outra análise importante que se dispõe com a utilização do mapa de krigagem é o mapa de probabilidade condicional, pois ele prediz a partir de um determinado valor numérico as chances de ocorrência de um evento na região de estudos. Observa-se pela figura do lado esquerdo, a área mais clara da imagem confirma a existência de Capim buffel com $P(Y > 101, 30)$, ou seja há uma concentração entre $80 - 100g/m^2$ na região, ou seja, uma concentração de biomassa de herbáceas. A explicação que se tem para este resultado, é que o sol nasce do lado direito e do mesmo lado encontra-se ao sul um riacho intermitente, ou seja, o mesmo consegue manter parte do solo sempre

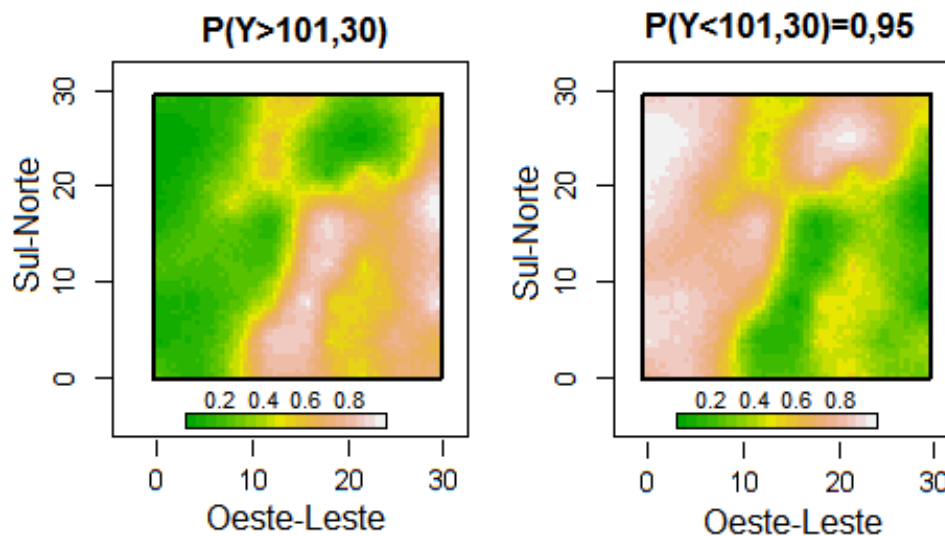


Figura 14: Comparação dos valores observados com a média do valor de referência com efeito de crescimento de pasto em área com juazeiro.

molhado, o sol se põe do lado oeste o calor excessivo da tarde não prejudica o crescimento do capim.

Visando quantificar as relações solo-planta em sistemas silvopastoril, vários estudos têm demonstrado que a diminuição da fertilidade do solo no semiárido pode ser revertida e ou prevenida através da introdução e/ou preservação de espécies arbóreas. Esse mesmo estudo realizado em épocas diferentes com Menezes e Salcedo (1999) e Menezes *et al.* (2002) avaliaram a fertilidade do solo, e encontraram "ilhas de fertilidade" nas áreas sob a influência das copas de árvores presentes em pastagens de capim buffel. Ou seja, observou-se que com a preservação do juazeiro (*Ziziphus joazeiro*) nas pastagens aumentou os níveis de matéria orgânica e nutrientes na camada superficial do solo, como também, maior produção de biomassa pelas herbáceas sob a copa das árvores.

4.2 Umbuzeiro

Variável Carbono

Os resultados obtidos demonstraram que há dependência espacial do carbono em função da presença dos indivíduos do umbuzeiro. Apresenta-se na Tabela 5 a estatística descritiva. Houve um bom comportamento das estimativas dos parâmetros de média, uma vez que os valores de média e mediana são próximos. Com relação ao afastamento da variável em relação a um valor central, tem-se então uma distribuição assimétrica à direita, apresentou-se uma curva com uma distribuição leptocúrtica. De acordo com o critério pro-

posto por Wilding e Dress (1983), o coeficiente de variação classifica a variabilidade entre $20\% < CV < 30\%$ como alta. O valor mínimo e máximo indicam a não existência de problemas amostrais com os dados.

Tabela 5: Estatística descritiva: mínimo, máximo, média, mediana, desvio padrão (DP), coeficiente de variação (CV), curtose, assimetria de amostras de solo da variável carbono da espécie umbuzeiro.

| Mínimo | Máximo | Média | Mediana | DP | CV | Curtose | Assimetria |
|--------|--------|-------|---------|-------|--------|---------|------------|
| 0,480 | 1,412 | 0,860 | 0,843 | 0,242 | 28,162 | 2,288 | 0,427 |

Da mesma forma com o juazeiro, a Figura 15 representa os dados nos quartis da variável carbono, ou seja, " + " 1º quartil, " Δ " 2º quartil, " o " 3º quartil e " \times " 4º quartil. O gráfico do canto superior direito e do canto inferior esquerdo, apresentam a dispersão dos dados em torno da média, observa-se que os dados estão dispersos. Existe uma fuga da distribuição gaussiana nos resíduos, necessita-se de possíveis transformações na variável, é através do *Box-Cox* que encontra-se o λ , que fará com que esta transformação seja possível, como também os valores do intervalo de confiança.

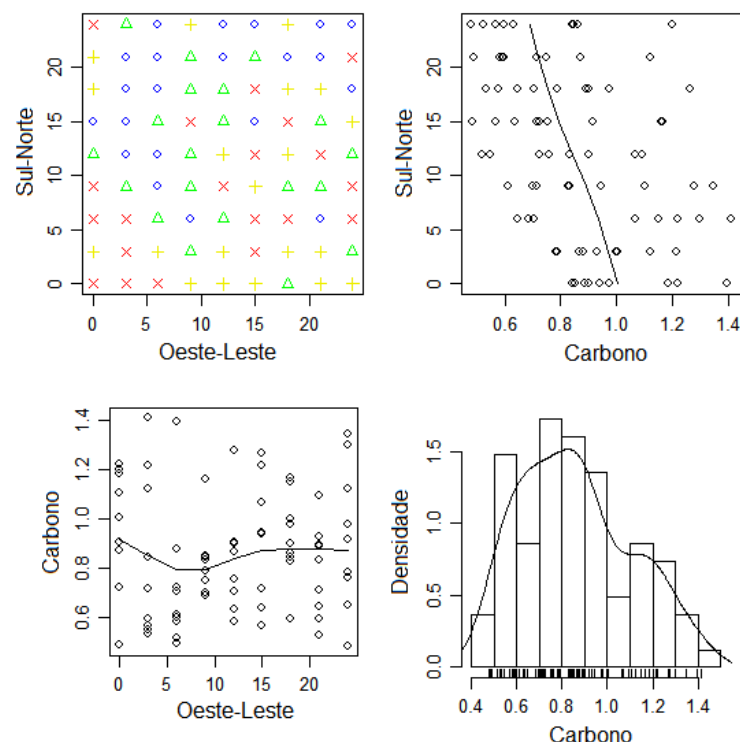


Figura 15: Gráfico descritivo do padrão espacial do umbuzeiro, variável carbono, com os dados originais.

A transformação ótima de *Box-Cox* lado esquerdo da Figura 16, por meio do logaritmo da função de verossimilhança, mostra qual o valor do parâmetro que mais aproxima-se os

resíduos à distribuição normal. Os intervalos de confiança também apresenta-se com essa transformação $[-0,73; 0,84]$ para λ não contém o valor 1, é preciso utilizar uma transformação nas variáveis, neste caso, para a variável foi encontrado o valor de $\lambda = 0,06$. A existência de dependência espacial pode ser verificada através do gráfico utilizando envelopes simulados e variogramas empíricos (Direito), nessa pesquisa foram utilizadas 1.000 simulações. Observa-se que, alguns pontos estão fora do envelope indicando assim uma dependência espacial de carbono no solo. Para que haja dependência, deve haver ao menos um ponto do variograma fora do envelope simulado, deste modo, se isso ocorrer rejeita-se hipótese nula (H_0), de que não há dependência espacial (DIGGLE; RIBEIRO, 2007).

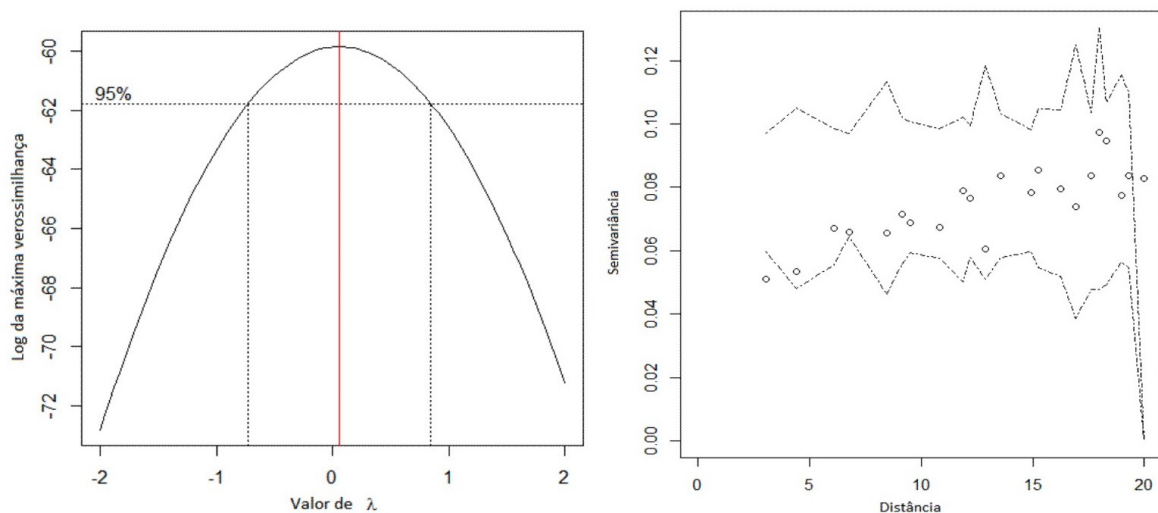


Figura 16: Gráfico de possíveis transformações de variáveis por meio da transformação ótima de *Box-Cox* (Esquerdo) e a verificação de dependência espacial através do envelope simulado (Direito) da variável carbono da espécie Umbuzeiro.

Após a transformação aos dados de carbono, percebe-se que o gráfico do canto superior esquerdo da Figura 16 condiz com a ideia de que existe padrão espacial na variável. O gráfico do canto superior direito e do canto inferior esquerdo, apresentam a dispersão dos dados em torno da média, observa-se que os dados estão menos dispersos que o gráfico da Figura 17.

A dependência espacial observada para os valores de carbono mostra que as análises da estatística clássica, em que as observações são consideradas aleatórias e independentes espacialmente, devem ser substituídas por análises espaciais, que levam em consideração as relações entre observações vizinhas.

Na Tabela 6, observa-se as estimativas da máxima verossimilhança para os parâmetros dos modelos ajustados, considerando-se a média constante sob a região de estudo.

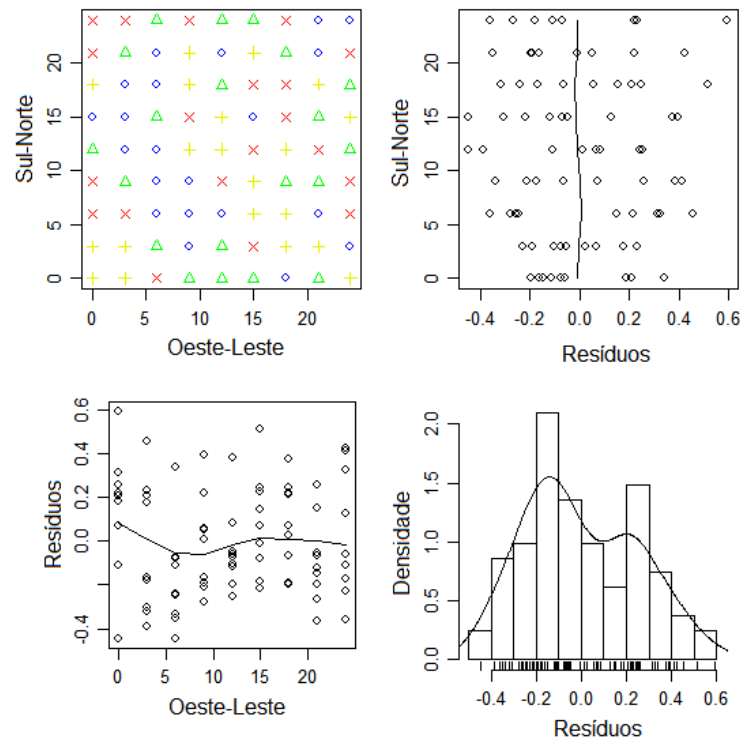


Figura 17: Gráfico descritivo do padrão espacial do umbuzeiro, variável carbono, com os dados transformados.

O estimador do parâmetro $\hat{\beta}$ apresenta valores próximos para todas as funções de correlações. O próximo passo é selecionar o melhor modelo geoestatístico. Para todas as características avaliadas, o modelo que ajustou melhor ao semivariograma experimental, para a variável carbono onde está inserida o umbuzeiro, foi a função de correlação *Matérn* com $kappa$ (1,0), onde apresenta-se o menor valor de *AIC* e o *BIC*.

Tabela 6: Estimativa dos parâmetros associado aos modelos por meio da máxima verossimilhança, assumindo-se a média da variável carbono constante sobre a região de estudo na espécie umbuzeiro.

| Modelos | $\hat{\beta}$ | $\hat{\tau}^2$ | $\hat{\sigma}^2$ | $\hat{\phi}$ | <i>AIC</i> | <i>BIC</i> | IDE |
|------------|---------------|----------------|------------------|--------------|------------|------------|-----|
| Matérn 0,5 | -0,160 | 0,025 | 0,054 | 4,659 | -14,99 | -5,415 | 32% |
| Matérn 1,0 | -0,161 | 0,036 | 0,043 | 3,438 | -15,05 | -5,471 | 46% |
| Esférico | -0,170 | 0,029 | 0,046 | 8,749 | -14,38 | -4,806 | 39% |
| Gaussiano | -0,166 | 0,043 | 0,034 | 5,659 | -14,79 | -5,209 | 57% |
| Circular | -0,170 | 0,031 | 0,045 | 7,891 | -14,51 | -4,936 | 41% |

Da mesma forma que analisou-se o juazeiro, faz-se para o umbuzeiro, o grau de dependência espacial, entre as características de fertilidade de solo, utilizou-se da relação entre o efeito pepita τ^2 e o patamar $\sigma^2 + \tau^2$, nota-se que o carbono apresenta dependência espacial moderada 46% de acordo com Cambardella *et al.* (1994), que pode ser descrita

pelo modelo *Matérn* com $kappa$ 1,0 com alcance de 3,438m, ou seja, amostras de carbono de solo selecionadas à distâncias inferiores estão correlacionadas entre si.

A análise variográfica apresentada na (Figura 18), mostrou que todas as características estudadas nos procedimentos até aqui descritos, apresentam estruturas espaciais bem definidas, ou seja, existe uma função estrutural com semivariância de comportamento modelável.

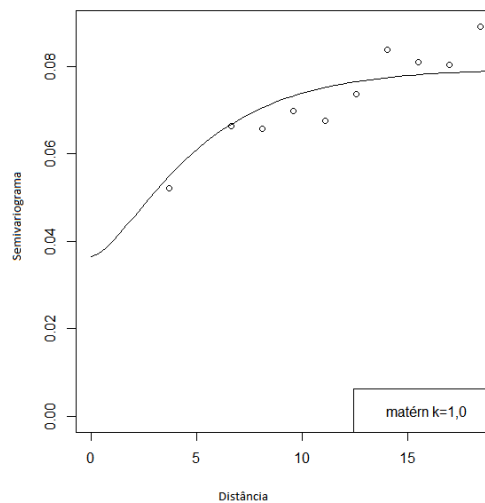


Figura 18: Modelo ajustado ao semivariograma conforme o método de máxima verossimilhança para a variável carbono da espécie umbuzeiro.

A krigagem é uma metodologia utilizada para estimar o valor de uma variável de interesse em local onde não foi possível fazer a observação. Com os modelos finais estabelecidos, foram calculadas as predições espaciais para o índice de fertilidade de solo em uma malha de 81 localizações espaciais, sendo que as estimativas paramétricas foram substituídas nas fórmulas de krigagem.

As Figuras 19, 20 e 21, apresentam as krigagens para diferentes valores de referência 0,8% como baixo, 1,1% médio e 1,4% alto, respectivamente, valores esses considerados o limite de especificação para a comparação de fertilidade do solo com os valores observados. Outra análise importante que se dispõe com a utilização, são os mapas de probabilidades condicionais, pois eles podem prever um valor numérico as chances de ocorrência de um evento na região de estudos.

Através da Figura 19 observa-se que a área mais clara do mapa do lado esquerdo apresenta uma $P(Y > 0,8\%)$ 80 – 100% de carbono no solo.

Analisa-se por meio da Figura 20 uma concentração de carbono com $P(Y > 1,1\%)$ de

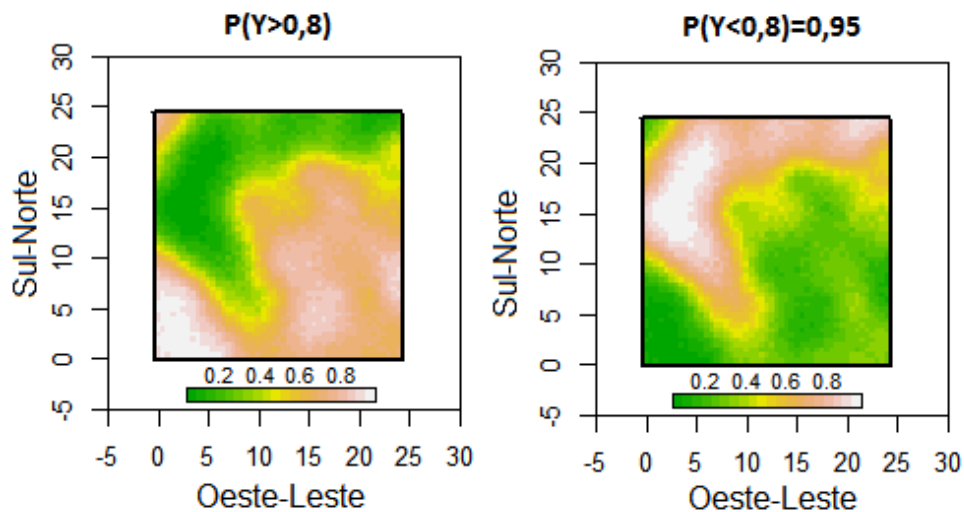


Figura 19: Comparação dos valores observados com o valor de referência para solos com baixo teor de carbono, na espécie umbuzeiro.

carbono entre 10 – 20%.

A Figura 21 confirma os que as imagens anteriores afirmam, a imagem do lado direito com $P < 1,4\%$ a cor clara predomina o local, ou seja, há concentração de carbono entre 95 – 100% em quase toda região.

Variável Biomassa de Herbáceas

Da mesma forma com as demais variáveis, segue-se as análises para esta variável. Os resultados obtidos demonstraram que há dependência espacial de biomassa de herbáceas em função da presença dos indivíduos do umbuzeiro. Pode-se observar por meio da Tabela 7 as estatísticas descritivas. Não houve um bom comportamento das estimativas dos parâmetros de média, uma vez que os valores de média e mediana não se aproximam. Com relação ao afastamento da variável em relação a um valor central, tem-se então uma distribuição assimétrica à direita $0,381 > 0$, sua distribuição é platicúrtica, o valor mínimo e máximo indicam a não existência de problemas amostrais com os dados. De acordo com o critério proposto por Wilding e Dress (1983), o coeficiente de variação classifica a variabilidade como muito alta $CV > 30\%$.

Tabela 7: Estatística descritiva: mínimo, máximo, média, mediana, desvio padrão (DP), coeficiente de variação (CV), curtose, assimetria de amostras de solo da variável biomassa de herbáceas da espécie umbuzeiro.

| Mínimo | Máximo | Média | Mediana | DP | CV | Curtose | Assimetria |
|--------|--------|--------|---------|---------|--------|---------|------------|
| 17,68 | 509,50 | 241,70 | 227,60 | 154,465 | 63,915 | 2,009 | 0,381 |

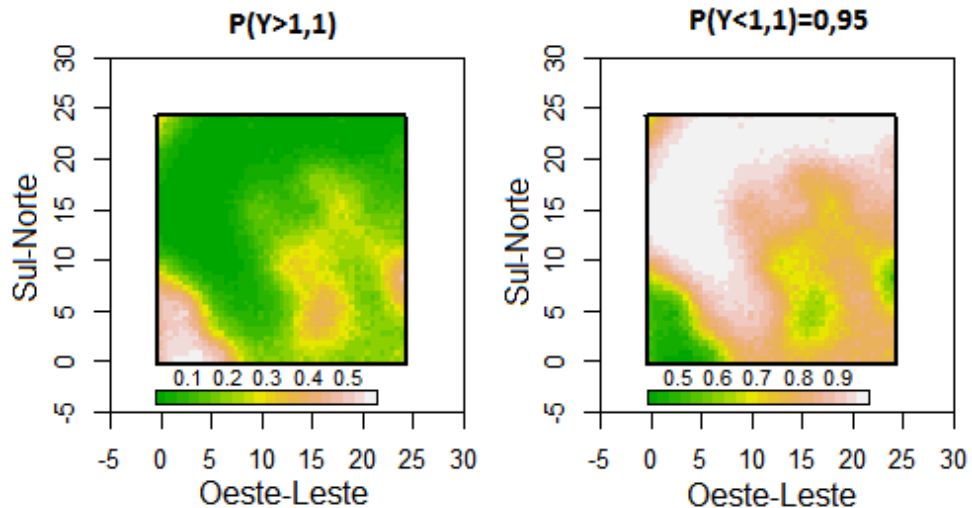


Figura 20: Comparação dos valores observados com o valor de referência para solos com médio teor de carbono, na espécie umbuzeiro.

O gráfico do canto superior esquerdo da Figura 23 representa os dados os quartis da variável biomassa de herbáceas, ” + ” 1º quartil, ” Δ ” 2º quartil, ” o ” 3º quartil e ” \times ” 4º quartil. No canto superior direito e no canto inferior esquerdo, apresentam a dispersão dos dados em torno da média. O gráfico do canto inferior direito, a análise exploratória do pressuposto de gaussianidade da resposta foi analisado, o qual mostra a densidade amostral dos dados desconsiderando-se o possível padrão espacial. Esta inspeção não garante a tal hipótese mas serve para mostrar desvios óbvios, caso existam. Como existe uma fuga da distribuição gaussiana nos resíduos, necessita-se de possíveis transformações na variável, é através do *Box-Cox* que encontra-se os valores do intervalo de confiança e de λ que fará com que esta transformação se torne possível.

A Figura 24 do lado esquerdo indica a transformação ótima de *Box-Cox* por meio do logaritmo da função da verossimilhança, mostra qual o valor do parâmetro que mais aproxima-se os dados à uma distribuição normal. Como os intervalos de confiança $[0, 15; 0, 73]$ para λ não contém o valor 1, foi encontrado o valor de $\lambda = 0,43$. A existência de dependência espacial pode ser verificada por meio do gráfico envelope simulado e variograma empírico (lado direito), nessa pesquisa foram utilizadas 1.000 simulações. Alguns pontos estão fora do envelope indicando assim uma dependência espacial de biomassa de herbáceas. Como dito anteriormente, para que haja dependência, deve haver ao menos um ponto do variograma fora, se isso ocorrer rejeita-se (H_0), de que não há dependência espacial (DIGGLE; RIBEIRO, 2007).

Após a transformação com o *Box-Cox* Figura 24 aos dados de biomassa de herbáceas, a imagem condiz com a ideia de que existe padrão espacial na variável. O gráfico do

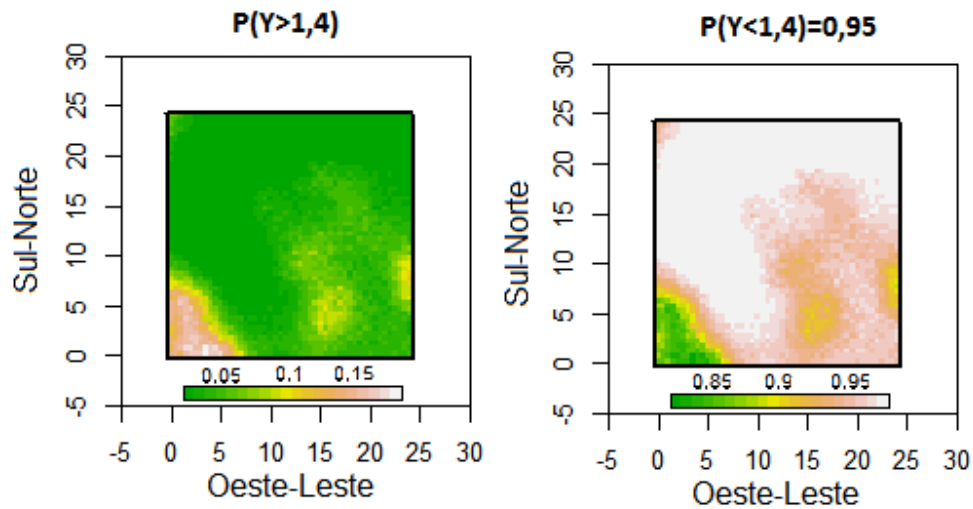


Figura 21: Comparação dos valores observados com o valor de referência para solos com alto teor de carbono, na espécie umbuzeiro.

canto superior direito e do canto inferior esquerdo, apresentam a dispersão dos dados em torno da média, observa-se que os dados estão menos dispersos que o gráfico da Figura 22. A análise exploratória do pressuposto de gaussianidade da resposta foi analisado para o gráfico do canto inferior direito, o qual mostra as densidades amostrais dos dados, desconsiderando-se o possível padrão espacial. Esta inspeção não garante a tal hipótese mas serve para mostrar desvios óbvios, caso existam.

As dependências espaciais observadas para os valores de biomassa de herbáceas mostram que as análises estatísticas clássicas, em que as observações são consideradas aleatórias e independentes espacialmente, devem ser substituídas por análises espaciais, que levam em consideração as relações entre observações vizinhas.

Tabela 8: Estimativa dos parâmetros associado aos modelos por meio da máxima verossimilhança, assumindo-se a média como um polinômio de primeira ordem sobre as coordenadas (x, y) .

| Modelos | $\hat{\beta}_0$ | $\hat{\beta}_1$ | $\hat{\beta}_2$ | $\hat{\tau}^2$ | $\hat{\sigma}^2$ | $\hat{\phi}$ | AIC | BIC | IDE |
|------------|-----------------|-----------------|-----------------|----------------|------------------|--------------|------|------|-----|
| Matérn 0,5 | 19,716 | 0,059 | 0,065 | 0,000 | 50,824 | 1,932 | 1029 | 1044 | 0% |
| Matérn 1,0 | 19,747 | 0,058 | 0,060 | 14,985 | 36,096 | 1,610 | 1029 | 1044 | 29% |
| Esférico | 19,534 | 0,061 | 0,066 | 18,766 | 32,444 | 6,813 | 1027 | 1042 | 37% |
| Gaussiano | 19,733 | 0,057 | 0,051 | 26,722 | 24,971 | 3,700 | 1028 | 1042 | 52% |
| Circular | 19,534 | 0,060 | 0,066 | 19,875 | 31,567 | 6,167 | 1027 | 1042 | 39% |

Dando seqüências as análises, pode-se observar por meio da Tabela 8, as estimativas da máxima verossimilhança para os parâmetros dos modelos ajustados. A estimativa do estimador dos parâmetros $\hat{\beta}_0$, $\hat{\beta}_1$ e $\hat{\beta}_2$ apresentam valores próximos para todas as funções

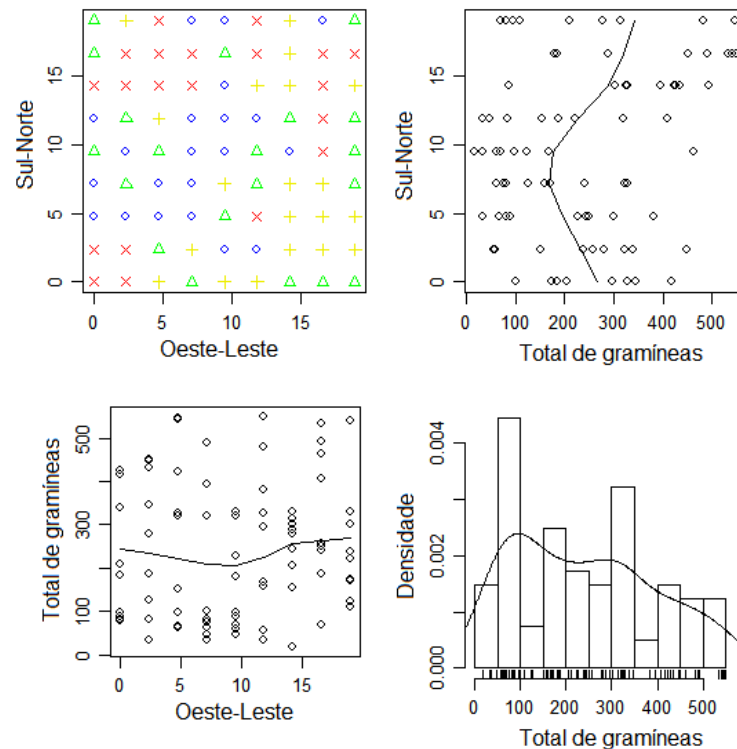


Figura 22: Gráfico descritivo do padrão espacial do umbuzeiro, variável herbáceas, com os dados originais.

de correlações. Para todas as características avaliadas, os modelos que ajustou-se melhor ao semivariograma experimental para a variável biomassa de herbáceas onde está inserida o umbuzeiro foram as funções de correlações Circular e Esférico, onde os valores de AIC e BIC são iguais, partindo do princípio a escolha do melhor ajuste é por meio do valor do logaritmo da função da verossimilhança, ou seja, o menor valor do log, indica melhor ajuste da função aos dados. O modelo Esférico apresentou um $\log = -508,00$, o Circular $\log = -507,6$, o melhor modelo para a variável em estudo é o Esférico. Na (Figura 25), encontra-se o semivariograma experimental ajustado conforme o método da máxima verossimilhança, a análise variográfica apresentada mostrou que todas as características estudadas nos procedimentos até aqui descritos, apresentam estruturas espaciais bem definidas, ou seja, existe uma função estrutural com semivariância de comportamento modelável.

Utilizou-se da relação entre o efeito pepita τ^2 sobre o patamar $\sigma^2 + \tau^2$, para verificar o grau de dependência espacial, entre as características de fertilidade de solo. Pela classificação de Cambardella *et al.* (1994), 39% está entre $25\% < IDE < 75\%$, indica portanto uma moderada dependência espacial, amostras de biomassa de herbáceas selecionadas a distâncias inferiores a $6,813m$ são espacialmente dependentes.

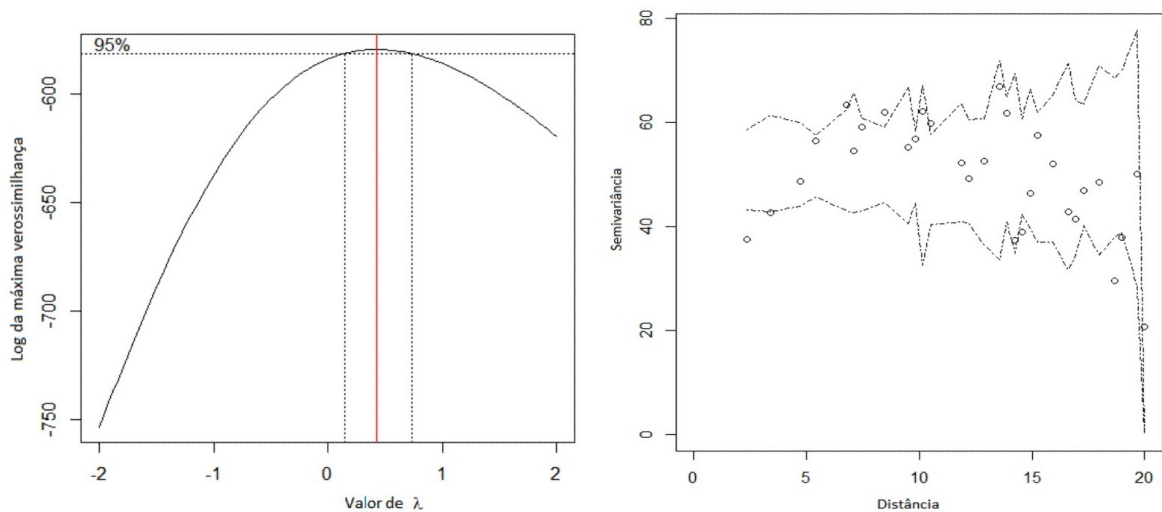


Figura 23: Gráfico de possíveis transformações de variáveis por meio da transformação ótima de *Box-Cox* (Esquerdo) e a verificação de dependência espacial através do envelope simulado (Direito) da variável biomassa de herbáceas da espécie umbuzeiro.

Conforme visto anteriormente com a espécie juazeiro, a krigagem é uma metodologia utilizada para estimar o valor de uma variável de interesse em um local onde não foi possível fazer a observação, que por sua vez é obtida por meio de uma interpolação. Com os modelos finais estabelecidos, foram calculadas as previsões espaciais para o efeito de crescimento de capim no solo em um malha de 81 localizações, sendo que as estimativas paramétricas foi substituída na fórmula da krigagem.

Pode-se observar por meio da Figura 26 o mapa preditivo do campo aleatório para o efeito de crescimento de capim. Outra análise importante que se dispõe com a utilização de krigagem é a probabilidade condicional, pois ela prediz a partir de um determinado valor numérico as chances de ocorrência de um evento na região de estudos. O lado direito da figura, indica $P < 241,70g/m^2$, observa-se que há maior concentração de Capim buffel próximo e abaixo da copa do umbuzeiro.

4.3 Algaroba

Variável Carbono

Analisou-se o juazeiro e o umbuzeiro, da mesma forma faz-se-á com a algaroba. Os resultados obtidos demonstraram que há dependência espacial do carbono em função da presença dos indivíduos da algaroba.

Apresenta-se na Tabela 9 a estatística descritiva. Não houve um bom comportamento

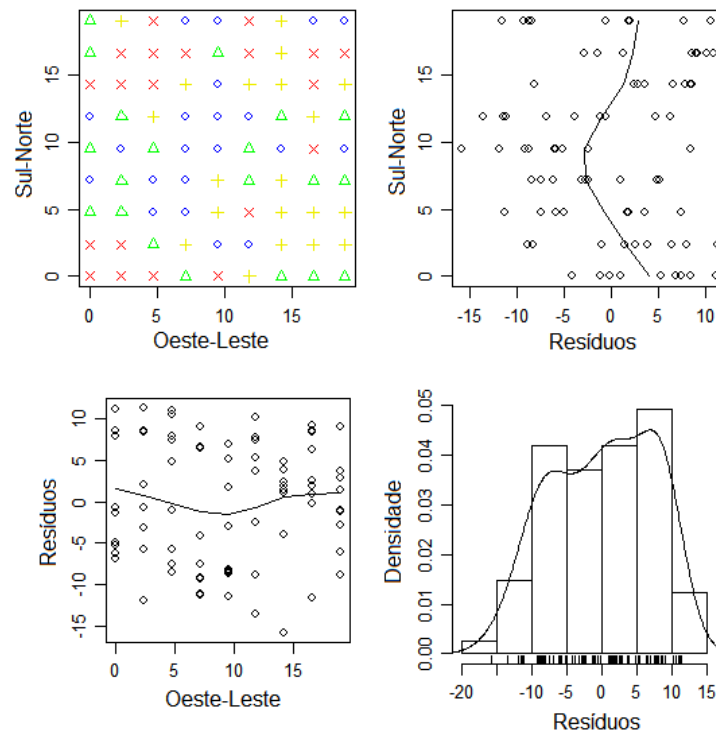


Figura 24: Gráfico descritivo do padrão espacial do umbuzeiro, variável biomassa de herbáceas, com os dados transformados.

das estimativas dos parâmetros de média, uma vez que os valores de média e mediana não se aproximam. Com relação ao afastamento da variável a um valor central, tem-se então uma distribuição assimétrica à direita, sua distribuição é platicúrtica, o valor mínimo e máximo indicam a não existência de problemas amostrais com os dados. De acordo com o critério proposto por Wilding e Dress (1983), o coeficiente de variação classifica a variabilidade dos dados como muito alta $CV > 30\%$.

Tabela 9: Estatística descritiva: mínimo, máximo, média, mediana, desvio padrão (DP), coeficiente de variação (CV), curtose, assimetria de amostras de solo da variável carbono da espécie algaroba.

| Mínimo | Máximo | Média | Mediana | DP | CV | Curtose | Assimetria |
|--------|--------|-------|---------|-------|--------|---------|------------|
| 0,516 | 3,804 | 1,520 | 1,476 | 0,540 | 35,555 | 3,059 | 0,553 |

O gráfico do canto superior esquerdo da Figura 27 representa os dados nos quartis da variável carbono, ” + ” 1º quartil, ” Δ ” 2º quartil, ” o ” 3º quartil e ” x ” 4º quartil, a imagem condiz com a ideia de que existe padrão espacial. Observou-se por meio do gráfico do canto superior direito e do canto inferior esquerdo a dispersão dos dados em torno da média, encontra-se um pouco dispersa. A análise exploratória do pressuposto de gaussianidade da resposta foi analisado por meio do gráfico do lado direito inferior, o qual mostra a densidade amostral dos dados desconsiderando-se o possível padrão espacial.

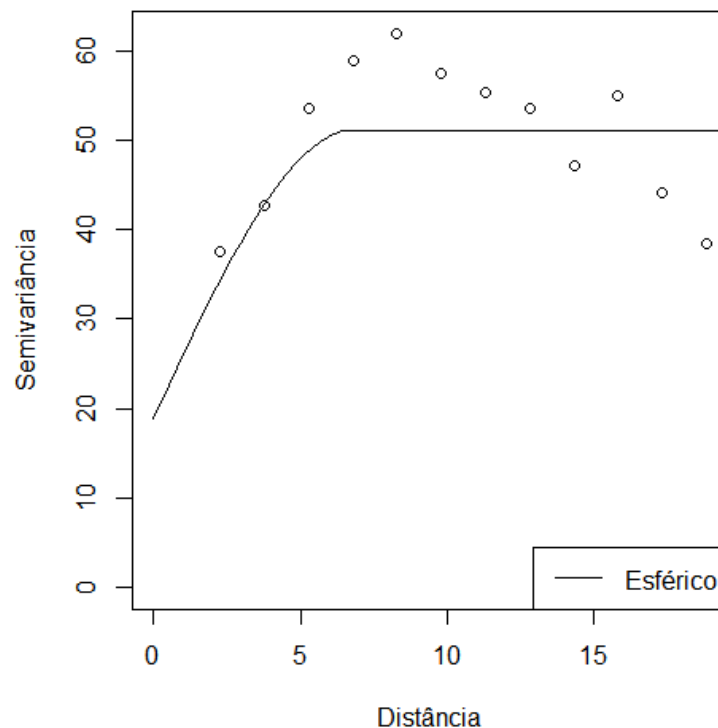


Figura 25: Modelo ajustado ao semivariograma conforme o método da máxima verossimilhança para a variável biomassa de herbáceas da espécie umbuzeiro.

Esta inspeção não garante a tal hipótese mas serve para mostrar desvios óbvios, caso existam. Como existe uma fuga da distribuição gaussiana nos resíduos, necessita-se de possíveis transformações na variável, é através do *Box-Cox* que encontra-se o valor do λ , como também os valores do intervalo de confiança.

Através da transformação ótima de *Box-Cox* (lado esquerdo) da Figura 28, por meio do logaritmo da função da verossimilhança, mostra qual o valor do parâmetro que mais aproxima-se os resíduos à distribuição normal. O valor 1 não está dentro do intervalo confiança $[-0,09; 0,91]$, neste caso, para a variável foi encontrado o valor de $\lambda = 0,39$. A existência de dependência espacial pode ser verificada através do gráfico utilizando envelope simulado e variograma empírico (lado direito), nessa pesquisa foram realizadas 1.000 simulações. Há dependência espacial de carbono no solo, quando se tem ao menos um ponto fora do envelope simulado, deste modo, se isso ocorrer rejeita-se hipótese nula, de que não há dependência espacial (DIGGLE; RIBEIRO, 2007).

Logo após a transformação, Figura 29, observa-se que os dados estão menos dispersos que o gráfico da Figura 27. A análise exploratória do pressuposto de gaussianidade da resposta foi analisado para o gráfico do canto inferior direito, o qual mostra as densidades amostrais dos dados, desconsiderando-se o possível padrão espacial. Esta inspeção não

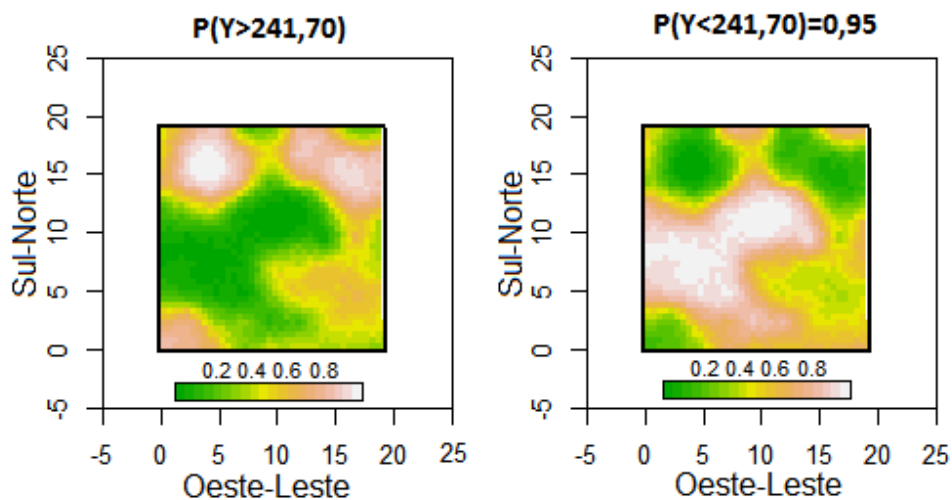


Figura 26: Comparação dos valores observados com a média do valor de referência com efeito de crescimento de pasto em área com umbuzeiro.

garante a tal hipótese mas serve para mostrar desvios óbvios, caso existam.

As dependências espaciais observadas para os valores de carbono mostram que as análises da estatística clássica, em que as observações são consideradas aleatórias e independentes espacialmente, devem ser substituídas por análises espaciais, que levam em consideração as relações entre observações vizinhas.

Na Tabela 10 observa-se as estimativas da máxima verossimilhança para os parâmetros dos modelos ajustados, considerando-se a média constante sob a região de estudo. O estimador do parâmetro $\hat{\beta}$ apresenta valores próximos para todas as funções de correlações. O próximo passo é selecionar o melhor modelo geoestatístico. Para todas as características avaliadas, o modelo que ajustou melhor ao semivariograma experimental, foi a função de correlação *Matérn* com *kappa* (0,5), este modelo, refere-se ao exponencial segundo Mello (2004).

Tabela 10: Estimativa dos parâmetros associado aos modelos por meio da máxima verossimilhança, assumindo a média da variável carbono constante sobre a região de estudo na espécie *algaroba*.

| Modelos | $\hat{\beta}$ | $\hat{\tau}^2$ | $\hat{\sigma}^2$ | $\hat{\phi}$ | <i>AIC</i> | <i>BIC</i> | IDE |
|------------|---------------|----------------|------------------|--------------|------------|------------|-----|
| Matérn 0,5 | 0,386 | 0,038 | 0,139 | 3,185 | 112,8 | 122,4 | 1% |
| Matérn 1,0 | 0,387 | 0,074 | 0,103 | 2,566 | 113,0 | 122,5 | 2% |
| Esférico | 0,405 | 0,000 | 0,163 | 4,400 | 114,4 | 124,0 | 0% |
| Gaussiano | 0,394 | 0,101 | 0,075 | 5,293 | 113,4 | 122,9 | 3% |
| Circular | 0,405 | 0,000 | 0,164 | 3,908 | 113,2 | 122,8 | 0% |

Para verificar o grau de dependência espacial entre as características de fertilidade de

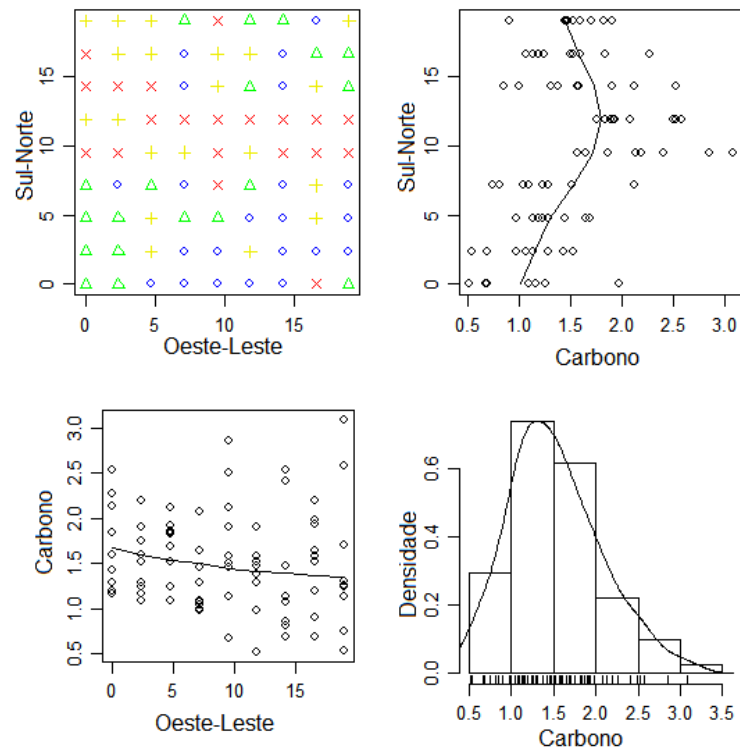


Figura 27: Gráfico descritivo do padrão espacial do algaroba, variável carbono, com os dados originais.

solo, utilizou-se o efeito pepita τ^2 sobre o patamar $\sigma^2 + \tau^2$. Obteve-se então, um grau de dependência forte 1%, conforme a classificação de Cambardella *et al.* (1994), que pode ser descrita pelo modelo *Matérn* com *kappa* 0,5 com alcance de aproximadamente 3,185m, ou seja, amostras de carbono de solo selecionadas a distâncias inferiores são espacialmente dependentes.

A análise variográfica apresentada na Figura 30 mostrou-se que todas as características estudadas nos procedimentos até aqui descritos, apresentam estruturas espaciais bem definidas, ou seja, existe uma função estrutural com semivariância de comportamento modelável.

A krigagem é uma metodologia utilizada para estimar o valor de uma variável de interesse em um local onde não foi possível fazer a observação. É obtida por meio de uma interpolação, que utiliza a correlação existente entre valores dos dados obtidos em pontos máximos. Com os modelos finais estabelecidos, foram calculadas as previsões espaciais para o índice de fertilidade de solo em uma malha de 81 localizações espaciais, sendo que as estimativas paramétricas foram substituída para a fórmula de krigagem.

As Figuras 31, 32 e 33 apresentam os mapas preditivos dos campos aleatórios para diferentes valores de referência para o teor de fertilidade do solo. Outra análise importante

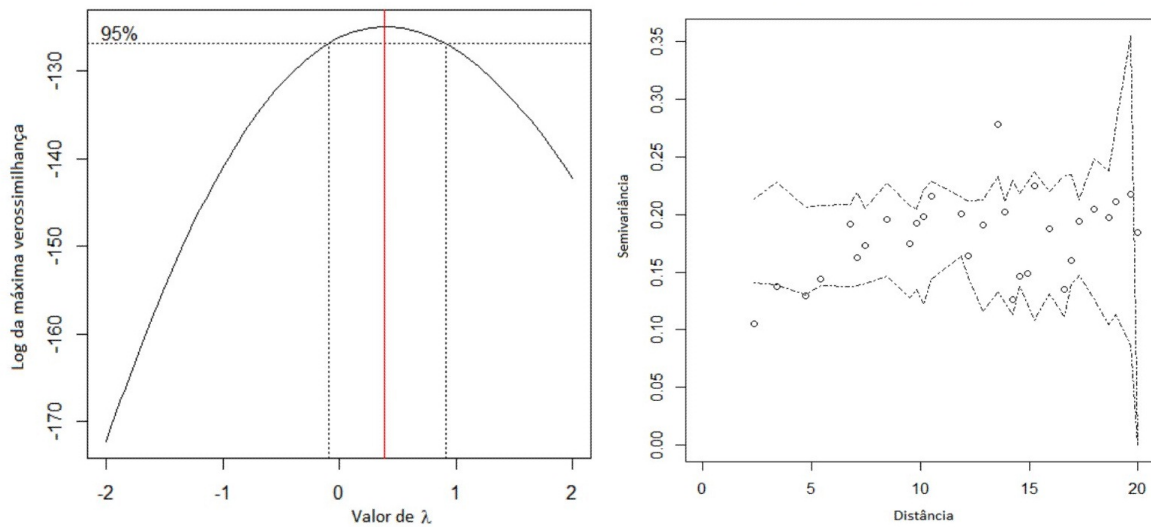


Figura 28: Gráfico de possíveis transformações de variáveis por meio da transformação ótima de *Box-Cox* (Esquerdo) e a verificação de dependência espacial através do envelope simulado (Direito) da variável carbono da espécie algaroba.

que se dispõe com a utilização de krigagem são as probabilidades condicionais, pois eles podem prever a partir de um determinado valor numérico as chances de ocorrência de um evento na região de estudos. Vale salientar que todas as observações aqui descritas, são em função do limite de especificação para baixo 0,8%, médio 1,1% e alto teor de fertilidade do solo 1,4%, respectivamente.

Na Figura 31, observa-se uma variedade de probabilidades dentro da área onde está inserida a algaroba. Percebe-se uma concentração muito forte de carbono na cor mais clara em quase toda a região de estudos, com probabilidade entre 90 – 100% de baixo teor de carbono no solo, havendo tão somente uma concentração bem pequena na região sul-leste com probabilidade entre 40 – 70%.

Seguindo a mesma análise, a Figura 32 apresenta uma concentração de carbono, abaixo da copa onde está inserida a algaroba, como também do lado leste e oeste, com probabilidade entre 80 – 100% de carbono no local para médio teor de fertilidade.

A Figura 33 confirma a concentração de carbono abaixo da copa onde está inserida a árvore com um índice de 80 – 100% de carbono.

Variável Biomassa de Herbáceas

Os resultados obtidos demonstraram que há dependência espacial de biomassa de herbáceas em função da presença dos indivíduos da algaroba. Pode-se observar por meio

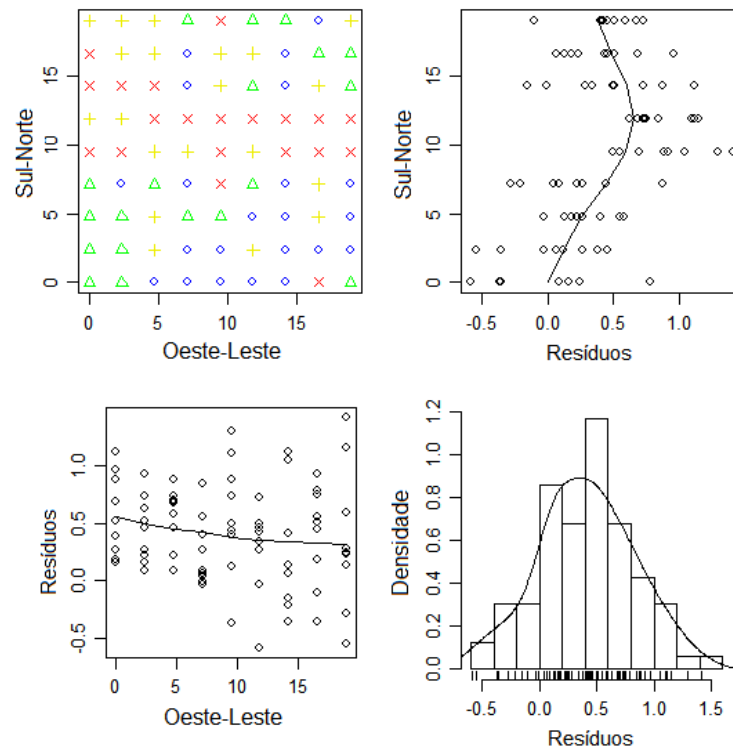


Figura 29: Gráfico descritivo do padrão espacial do algaroba, variável carbono, com os dados transformados.

da Tabela 11 a estatística descritiva. Não houve um bom comportamento, uma vez que os valores de média e mediana não se aproximam. Com relação ao afastamento da variável a um valor central, tem-se então uma distribuição assimétrica à direita, sua distribuição é platicúrtica, diferindo da curva normal. Os valores mínimos e máximos indicam a não existência de problemas amostrais com os dados. De acordo com o critério proposto por Wilding e Dress (1983), o coeficiente de variação classifica a variabilidade como muito alta $CV > 30\%$.

Tabela 11: Estatística descritiva: mínimo, máximo, média, mediana, desvio padrão (DP), coeficiente de variação (CV), curtose, assimetria de amostras de solo da variável biomassa de herbáceas da espécie algaroba.

| Mínimo | Máximo | Média | Mediana | DP | CV | Curtose | Assimetria |
|--------|--------|--------|---------|--------|-------|---------|------------|
| 9,26 | 485,30 | 180,00 | 155,90 | 129,24 | 71,80 | 2,67 | 0,74 |

Faz-se necessário a análise de dados da variável biomassa de herbáceas, da mesma forma que analisou-se as demais variáveis ditas anteriormente, o gráfico do canto superior esquerdo da Figura 34 representa o " + " 1º quartil, " Δ " 2º quartil, " o " 3º quartil e " x " 4º quartil, a imagem condiz com a ideia de que existe padrão espacial. O gráfico do canto superior direito e do canto inferior esquerdo, apresentam a dispersão dos dados

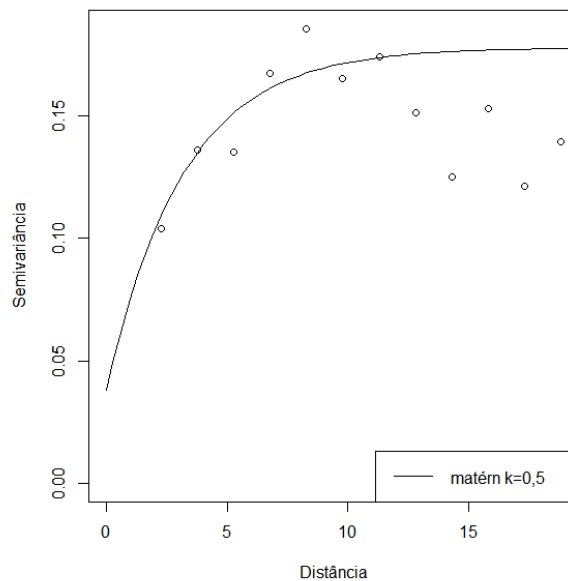


Figura 30: Modelo ajustado ao semivariograma conforme o método da máxima verossimilhança para a variável carbono da espécie algaroba.

em torno da média, onde os mesmos estão dispersos. Com base no gráfico do canto inferior direito, a análise exploratória do pressuposto de gaussianidade da resposta foi analisado, o qual mostra a densidade amostral dos dados desconsiderando-se o possível padrão espacial. Existe uma fuga da distribuição gaussiana nos resíduos, necessita-se de possíveis transformações, é através do *Box-Cox* que encontra-se o valor de λ que fará com que esta transformação seja possível e dos valores do intervalo de confiança.

A transformação ótima de *Box-Cox* lado esquerdo da Figura 35, por meio do logaritmo da função de verossimilhança apresenta qual o valor do parâmetro que mais aproxima-se os dados à distribuição normal. Os intervalos de confiança $[0, 19; 0, 65]$ não contém o valor 1, neste caso, foi encontrado o valor de $\lambda = 0,43$. A existência de dependência espacial pode ser verificada por meio do gráfico utilizando envelope simulado e variograma empírico (lado direito), nessa pesquisa foram utilizadas 1.000 simulações. Observa-se que, alguns pontos estão fora do envelope indicando assim uma dependência espacial de biomassa de herbáceas. Para que haja dependência, deve haver ao menos um ponto do variograma fora do intervalo, deste modo, se isso ocorrer rejeita-se hipótese nula (H_0), de que não há dependência espacial (DIGGLE; RIBEIRO, 2007).

Após a transformação com o *Box-Cox* aos dados de biomassa de herbáceas. O gráfico da Figura 36 do canto superior direito e do canto inferior esquerdo, apresentam a dispersão dos dados em torno da média, observa-se que os dados estão menos dispersos que o

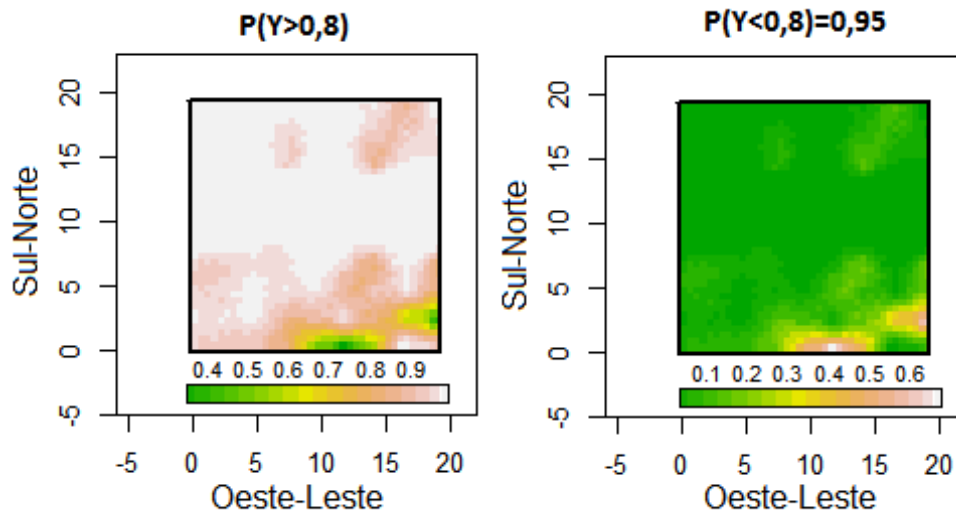


Figura 31: Comparação dos valores observados com o valor de referência para solos com médio teor de carbono, na espécie Algaroba.

gráfico da Figura 34. A análise exploratória do pressuposto de gaussianidade da resposta foi analisado para o gráfico do canto inferior direito, o qual mostra as densidades amostrais dos dados, desconsiderando-se o possível padrão espacial. Esta inspeção não garante a tal hipótese mas serve para mostrar desvios óbvios, caso existam.

As dependências espaciais observadas para os valores de biomassa de herbáceas mostram que as análises estatísticas clássicas, em que as observações são consideradas aleatórias e independentes espacialmente, devem ser substituídas por análises espaciais, que levam em consideração as relações entre observações vizinhas.

Dando sequências as análises, pode-se observar por meio da Tabela 12 as estimativas da máxima verossimilhança para os parâmetros dos modelos ajustados. A estimativa do estimador do parâmetro $\hat{\beta}$ apresenta valores próximos para todas as funções de correlações. O próximo passo é selecionar o melhor modelo geoestatístico. Para todas as características avaliadas, o modelo que ajustou melhor ao semivariograma experimental para a variável biomassa de herbáceas, onde está inserida a algaroba, foi a função de correlação Esférico, apresentou-se o menor valor de *AIC* e *BIC*.

Para verificar o grau de dependência espacial, entre as características de biomassa de herbáceas, utilizou-se da relação entre o efeito pepita τ^2 e o patamar $\sigma^2 + \tau^2$, para o grau de dependência espacial obteve-se 63%, ou seja, é considerada uma dependência moderada, pela classificação de Cambardella *et al.* (1994), com alcance de aproximadamente 13,260m, ou seja, amostras de biomassa de herbáceas selecionadas a distâncias inferiores são espacialmente dependentes.

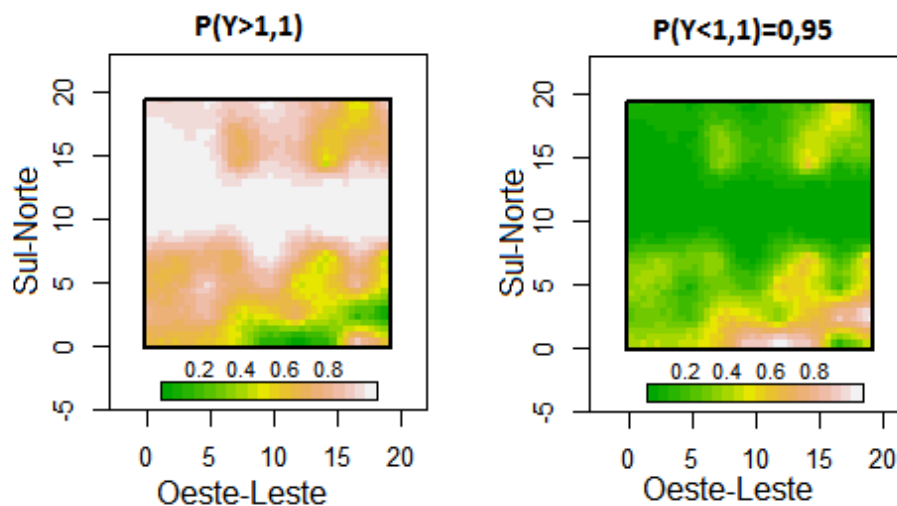


Figura 32: Comparação dos valores observados com o valor de referência para solos com médio teor de carbono, na espécie Algaroba.

Tabela 12: Estimativa dos parâmetros associado aos modelos por meio da máxima verossimilhança, assumindo a média da variável biomassa de herbáceas constante sobre a região de estudo na espécie algaroba.

| Modelos | $\hat{\beta}$ | $\hat{\tau}^2$ | $\hat{\sigma}^2$ | $\hat{\phi}$ | AIC | BIC | IDE |
|------------|---------------|----------------|------------------|--------------|-------|--------|-----|
| Matérn 0,5 | 18,524 | 29,815 | 22,802 | 5,306 | 993,0 | 1003,0 | 57% |
| Matérn 1,0 | 18,555 | 34,198 | 18,801 | 4,084 | 992,9 | 1002,0 | 65% |
| Esférico | 18,280 | 32,830 | 19,370 | 13,260 | 992,3 | 1002,0 | 63% |
| Gaussiano | 18,480 | 37,046 | 16,216 | 7,649 | 992,6 | 1002,0 | 70% |
| Circular | 19,930 | 34,400 | 28,470 | 22,140 | 993,3 | 1003,0 | 55% |

A análise variográfica apresentada na Figura 37 mostrou que todas as características estudadas nos procedimentos até aqui descritos, apresentam estruturas espaciais bem definidas, ou seja, existe uma função estrutural com semivariância de comportamento modelável.

A metodologia utilizada para estimar o valor de uma variável de interesse em um local onde não foi possível fazer a observação, é chamada de krigagem, ela é obtida através de uma interpolação, que utiliza a correlação existente entre valores dos resíduos obtidos em pontos máximos. Com os modelos finais estabelecidos, foram calculadas as previsões espaciais para a biomassa de herbáceas em um malha de 81 localizações espaciais, sendo que as estimativas paramétricas foram substituídas na fórmula da krigagem.

Pode-se observar por meio da Figura 38 o mapa preditivo do campo aleatório para o valor de referência para a biomassa de herbáceas. Outra análise importante que se dispõe com a utilização dos mapas de krigagem são os mapas de probabilidades condicionais,

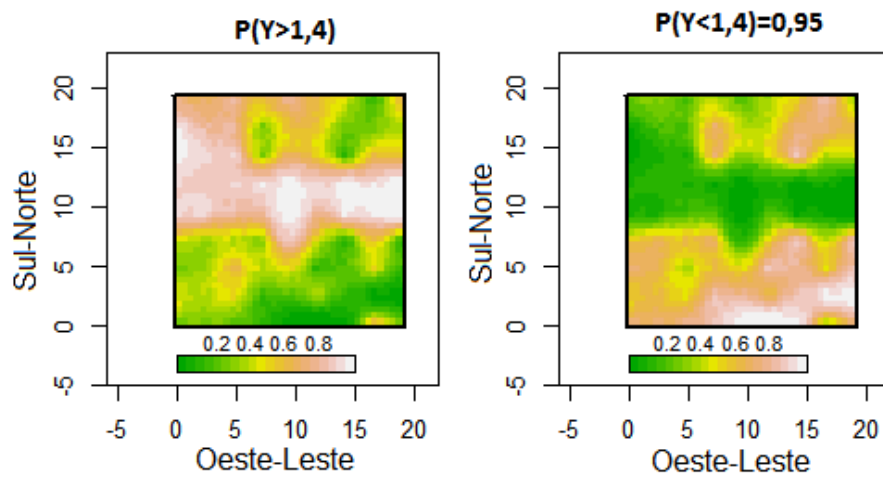


Figura 33: Comparação dos valores observados com o valor de referência para solos com médio teor de carbono, na espécie algaroba.

pois eles podem prever a partir de um determinado valor numérico as chances de ocorrência de um evento na região de estudos, a imagem do lado direito confirma que o capim está concentrado no centro-sul da região com um índice de 70 – 100% para $180,00 \text{ g/m}^2$.

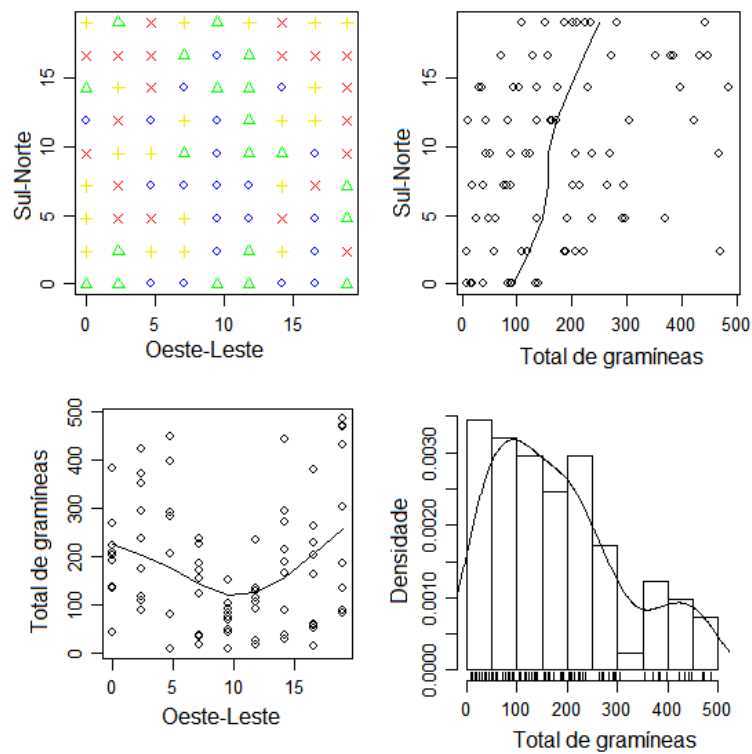


Figura 34: Gráfico descritivo do padrão espacial do algaroba, variável biomassa de herbáceas, com os dados originais.

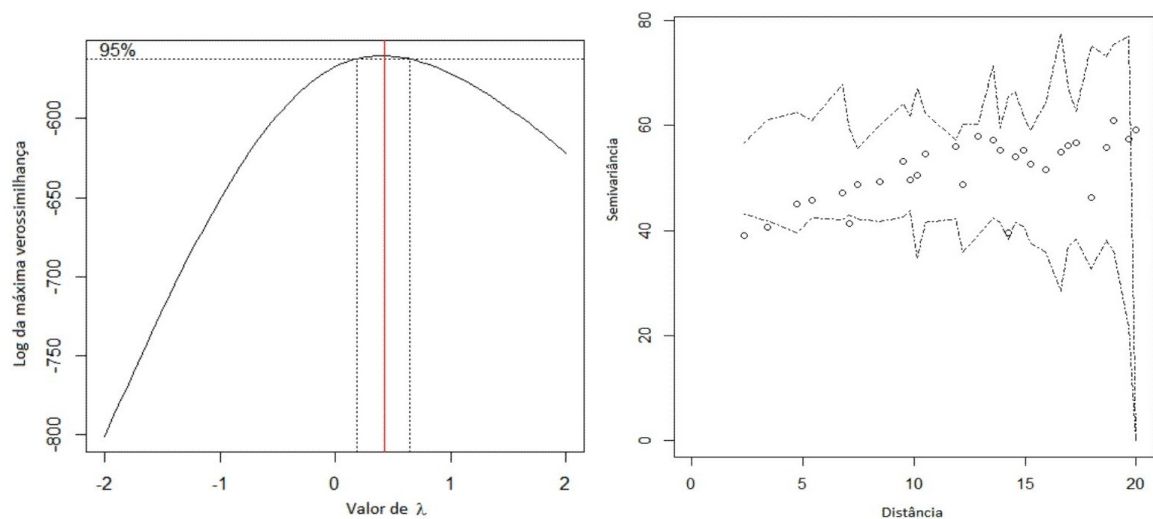


Figura 35: Gráfico de possíveis transformações de variáveis por meio da transformação ótima de *Box-Cox* (Esquerdo) e a verificação de dependência espacial através do envelope simulado (Direito) da variável biomassa de herbáceas da espécie algaroba

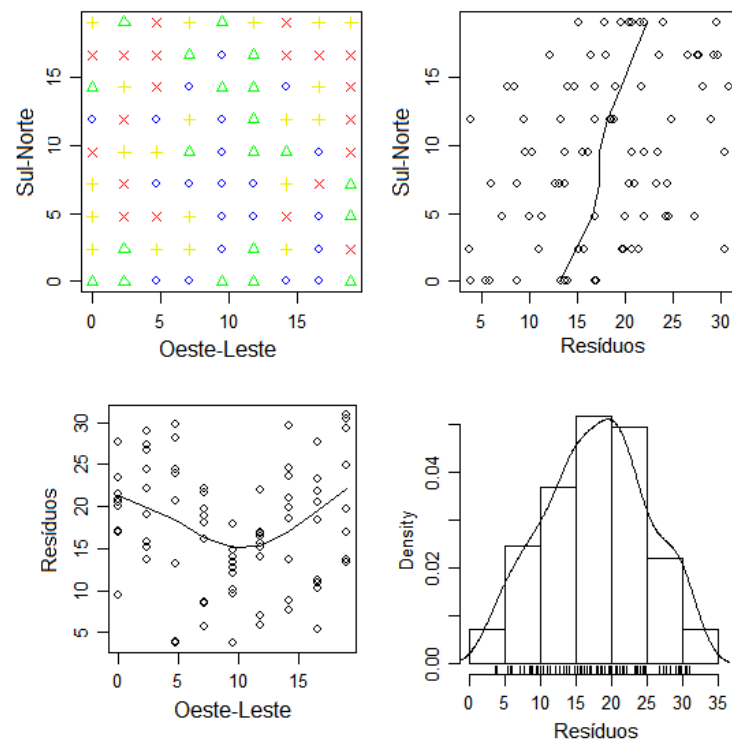


Figura 36: Gráfico descritivo do padrão espacial do algaroba, variável biomassa de herbáceas, com os dados transformados.

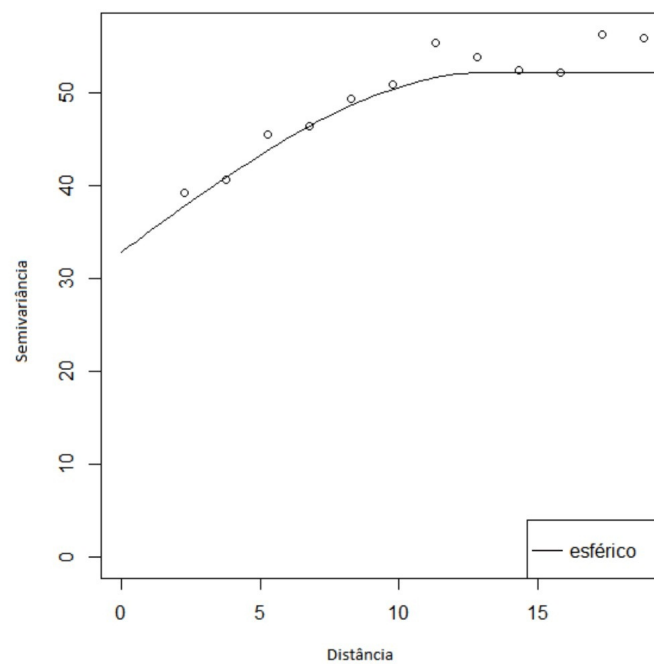


Figura 37: Modelo ajustado ao semivariograma conforme o método da máxima verossimilhança para a variável biomassa de herbáceas da espécie algaroba.

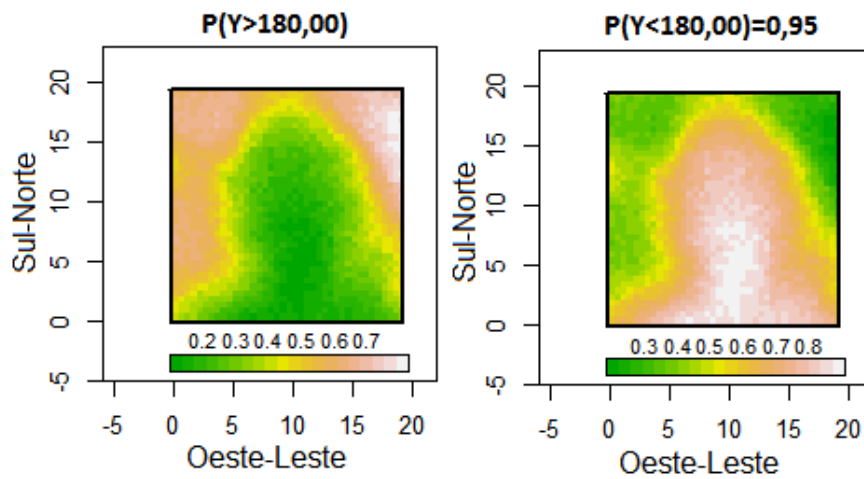


Figura 38: Comparação dos valores observados com a média do valor de referência com efeito de crescimento de pasto em área com algaroba

5 Considerações finais

A partir dos métodos apresentados neste trabalho chega-se as seguintes conclusões: com o semivariograma foi possível verificar e modelar a dependência espacial através de funções de correlação que expressam a estrutura de dependência espacial das características avaliadas. A técnica da krigagem é considerada uma boa metodologia de interpolação de dados para as variáveis estudadas, tornou-se um fator importante, pois possibilitou realizar predições a partir dos dados amostrais.

Visando quantificar as relações solo-planta em sistemas silvopastoril, vários estudos têm demonstrado que a diminuição da fertilidade do solo no semiárido pode ser revertida e/ou prevenida através da introdução e ou preservação de espécies arbóreas.

Avaliou-se a fertilidade do solo, e encontrou-se "ilhas de fertilidade" nas áreas sob a influência das copas de árvores presentes em pastagens de capim buffel, ou seja, observou-se que o plantio de algaroba nas pastagens aumentou os níveis de matéria orgânica e nutrientes na camada superficial do solo. Concluí-se que parte dos nutrientes acumulados na camada superficial do solo sob a copa de indivíduos de juazeiro são provenientes de camadas mais profundas. O plantio de Capim buffel consorciado ao umbuzeiro, consegue-se preservar o carbono no solo.

Referências

- AKAIKE, H. Information measures and model selection. *International Statistical Institute*, v. 44, p. 277–291, 1983.
- ALVES, J. J. A.; ARAÚJO, M. A. d.; NASCIMENTO, S. S. d. Degradação da caatinga: uma investigação ecogeográfica. *Revista Caatinga*, v. 22, n. 3, p. 126–135, 2009.
- ARAÚJO, F. J. A. Manipulação da vegetação lenhosa da caatinga com fins pastoris. simpósio sobre caatinga e sua exploração racional, 1984. *Feira de Santana*, p. 327–343, 1986.
- BAILEY, T.; GATRELL, A. Interactive spatial data analysis. *Longman Scientific and Technica*, p. 413, 1995.
- BOX, P.; COX, D. R. An analysis of transformations. *Journal of the Royal Statistical Society*, p. 211–252, 1964.
- BRASIL, P. Ministério do meio ambiente, secretaria de recursos hídricos. programa de ação nacional de combate à desertificação e mitigação dos efeitos da seca. *Brasília.DF*, p. 213, 2005.
- CAMARGO, E. C. G. Geoestatística: Fundamentos e aplicações. *Análise Espacial de Dados Geográficos*, p. 12–25, 2002.
- CAMARGO, E. C. G.; FUCKS, S. D.; CÂMARA, G. Análise espacial de superfícies. p. 37, 2002.
- CAMBARDELLA, C. A.; MOORMAN, T. B.; NOVAK, J. M.; PARKIN, T. B.; KARLEN, D. L.; TURCO, R. F.; KONOPKA, A. E. Field-scale variability of soil properties in central iowa soils. p. 11, 1994.
- CAMPO, N. R. N. do. 2012. Disponível em: <<http://www.nordeste rural.com.br/nordeste rural/matler.asp?newsId=6578>>.
- CARLIN, B.; LOUIS, T. Bayes and empirical bayes methods for data analysis. *Chapman and Hall*, v. 2, 2000.
- CHRISTENSEN, O. F.; DIGGLE, P. J.; RIBEIRO, P. J. Analysing positive.valued spatial data the transformed gaussian model. *Geostatistics for environmental applications, Boston*, p. 287–298, 2001.
- CRUZ, L. Biosfera. *Instituto de Matemática Pura e Aplicada. Mestrado em Matemática (opção Computação Gráfica), Rio de Janeiro*, p. 44, 2010. Disponível em: <w3.impa.br/lcruz/courses/pi/terrenos-biosfera.pdf>.
- DAVID, M. Geostatistical ore reserve estimation. *Elsevier Scientific, New York*, p. 6, 1977.

- DEUTSCH, C. Geostatistical reservoir modeling. *Applied Geoestatics Series*, 2002.
- DEUTSCH, C. V.; JOURNEL, A. G. Geostatistical software library and users guide. *Oxford University Press*, p. 184, 1992.
- DIGGLE, P. J.; RIBEIRO, P. J. *Model-Based geostatistics*. [S.l.]: New York: Springer, 2007. 230 p.
- DRUCK, S.; CARVALHO, M. S.; CAMARA, G.; MONTEIRO, A. M. V. Análise espacial de dados geográficos. *EMBRAPA*, p. 230, 2004.
- EMBRAPA. *Sistema Brasileiro de Classificação de Solos*. [S.l.]: Centro Nacional de Pesquisa de Solos - Rio de Janeiro, RJ, 2009. 412 p.
- FONSECA, B. H. F. Um estudo sobre estimação e predição em modelos geoestatísticos bivariados. p. 74, 2008. Piracicaba. Dissertação (Mestrado) - Escola Superior de Agricultura "Luiz de Queiroz", Universidade de São Paulo.
- FRADE, D. D. R. Modelagem geoestatística na determinação da variabilidade espacial das propriedades físico-químicas do solo. p. 45, 2011. Monografia. Universidade Estadual da Paraíba.
- FRANCISCO, P. S. Juazeiro. 2012. Disponível em: <<http://www.portalsaofrancisco.com.br/alfa/juazeiro/juazeiro.php>>.
- FRANCISCO, P. S. Umbu. 2012. Disponível em: <<http://www.portalsaofrancisco.com.br/alfa/umbu/umbu-1.php>>.
- GUERRA, P. A. G. Geoestatística operacional. p. 154, 1988. Brasília: Departamento Nacional da Produção Mineral.
- GUIMARÃES, E. C. Geoestatística básica e aplicada. *UFU-FAMAT Núcleo de Estudos Estatísticos e Biométricos*, p. 77, 2004. Disponível em: <<http://pt.scribd.com/doc/61263725/Geoestatistica-Basica-e-Aplicada>>.
- INFOPEDIA. Gramineas. *Porto Editora*, 2003. Disponível em: <[http://www.infopedia.pt/\\$gramineas](http://www.infopedia.pt/$gramineas)>.
- ISAAKS, E.; SRIVASTAVA, R. M. An introduction to applied geostatistics. *New York: Oxford University Press*, p. 561, 1989.
- JOURNEL, A. G. New distance measures: The route toward truly non-gaussian geoestatics. *Mathematical Geology, Vol.2, n.04*, p. 459–475, 1988.
- JOURNEL, A. G.; HUIJBREGTS, C. J. Mining geostatistics. *London: Academic Press*, p. 600, 1978.
- JUNIOR, I. S. d. O.; SILVA, V. M. da. Capim buffel. *IPA - Instituto Agrônomo de Pernambuco*, 2012. Disponível em: <<http://www.ipa.br/resp34.php>>.
- KREMPI, A. P. Explorando recursos de estatística espacial para análise de acessibilidade da cidade de bauru. *Dissertação (Mestrado) Escola de Engenharia de São Carlos, USP*, p. 94, 2004.

LANDIM, P. M. B. *Análise estatística de dados geológicos*. [S.l.]: 2 edição revista e ampliada. São Paulo, 2003.

LE, D. N.; ZIDEK, J. V. Statistical analysis of environmental space-time processes. *New York: Springer*, p. 327, 2006.

LIMA, D. L. Sensoriamento remoto e geoestatístico na caracterização espaço-temporal de plantas aquáticas. presidente prudente. dissertação (mestrado) - universidade estadual paulista. p. 100, 2006.

LINHARES, S.; GEWANDSZBAJDER, F. *Biologia hoje* - vol 3. são paulo. *Editora Atica*, 1998. Disponível em: <<http://www.invivo.fiocruz.br/cgi/cgilua.exe/sys/start.htm?infoid=962&sid=2>>.

LYRA, L. H. d. B.; LIMA, D. L.; SILVA, S. d. S.; XAVIER, T. S.; LIMA, V. d. C. A questão do semi-Árido e o bioma caatinga. p. 5, 2009.

MATHERON, G. Principles of geostatistics. *Economic mGeology*, v. 58, p. 1246–1266, 1963.

MATÉRN, B. Spatial variation. *Verlag, Berlin: Springer*, p. 365, 1986.

MCBRATNEY, A. G.; WEBSTER, A. G. Choosing functions for semi-variograms and fitting them to sampling estimates. *Journal of Soil Science*, 617-39, p. 37, 1986.

MECABO, A.; JUNIOR, G.; TOMASI, I.; LOHANE, I.; CENCI, N.; KEIKO, N.; DESIREE, T. Bioma caatinga. *Instituto Federal Catarinense - Campus Camboriú*, p. 14, 2012.

MELLO, J. M. de. *Geoestatística aplicada ao inventário florestal*. 2004. Disponível em: <<http://www.teses.usp.br/teses/disponiveis/11/11150/tde-06122004-100612/pt-br.php>>.

MENEZES, R.; SALCEDO, I.; ELLIOTT, E. Microclimate and nutrient dynamics in a silvopastoral system of semiarid northeastern brazil. *Agroforestry Systems*, p. 27–38, 2002.

MENEZES, R. S. C.; SALCEDO, I. H. Influence of tree species on the herbaceous understory and soil chemical characteristics in a silvopastoral system in semi-arid northeastern brazil. *R. Bras. Ci. Solo*, p. 817–826, 1999.

NELDER, J.; R., M. A simplex method for function minimization. *Comput. J. 7: p. 308-313*, 1965. *Spinger*, p. 365, 1986.

NORDESTE, B. do. Agenda do produtor rural. p. 271, 2007.

OLINDA, R. A. d. Métodos para análise de independência entre marcas e pontos em processos pontuais marcados. *Dissertação (Mestrado) Universidade Federal de Lavras UFLA MG*, p. 93, 2008.

OLIVEIRA, A. M. S.; SERIGATTO, E. M. Tutorial usando arcgis: Geostatistical analyst. *Viçosa, MG: UFV*, 2004.

OLIVEIRA, M. C. N. de. Métodos de estimação de parâmetros em modelos geoestatísticos com diferentes estruturas de covariâncias: uma aplicação ao teor de cálcio no solo. *Tese apresentada a Escola Superior de Agricultura Luiz de Queiroz*, p. 153, 2003.

- OLIVEIRA, W. M. d.; CHAVES, I. d. B.; LIMA, E. R. V. d. Índices espectrais de vegetação de caatinga em um neossolo litólico do semiárido paraibano. *Simpósio Brasileiro de Sensoriamento Remoto, 14, 2009, Natal. Anais... Natal: INPE*, p. 2103–2110, 2009.
- ROCHA, A. C. B. A geoestatística aplicada à avaliação e caracterização de reservatórios petrolíferos. p. 84, 2005. Campina Grande. Dissertação(Mestrado), Universidade Federal de Campina Grande.
- SCHABENBERGER, O.; GOTWAY, C. Statistical methods for spatial data analysis. *Boca Raton: Chapman and Hall / CRC*, p. 488, 2005.
- SCHIMIDT, A. M.; SANSO, B. Modelagem bayesiana da estrutura de covariância de processos espaciais e espaço-temporais. *Simpósio Nacional de Probabilidade e Estatística. Associação Brasileira de Estatística . Caxambú, Minicurso... São Paulo*, p. 151, 2006.
- SICHEL, H. Anatomy of the generalized inverse gaussian-poisson distribution with special applications to bibliometric studies. *Information processing e management, Oxford*, n. 1 v. 28, p. 5–17, 1992.
- SILVA, C. G. d. Algaroba. 2012. Disponível em: <www.ct.ufpb.br/laboratorios/lpfd/index.php?option=com_content&view=article&id=72&Itemid=70>.
- SILVA, E. D. da; MOREIRA, J. M.; SANTOS SILVA, C. dos; OLIVEIRA, C. F. T. de; SILVA, S. P. da; SOUZA, S. B. de; SANTOS, J. F. dos; PARANHOS, M. G. Plantar árvores para colher o futuro. *Agroecologia na Borborema*, p. 37, 2011. Disponível em: <<http://aspta.org.br/wp-content/uploads/2011/11/Cartilha-Plantar-árvores-para-colher-o-futuro.pdf>>.
- SILVA, V. P. d. Sistemas silvipastoris embrapa florestas. *Centro Nacional de Pesquisa de Florestas. Colombo PR Brasil*, 2004. Disponível em: <www.cnpf.embrapa.br/pesquisa/safs/index.htm>.
- SZWARCWALD, C. L.; BASTOS, F. I.; BARCELLOS, C.; ESTEVES, M. A. P.; CASTILHO, E. A. de. Dinâmica da epidemia de aids no município do rio de janeiro, no período de 1988-1996: uma aplicação de análise estatística espaço-temporal. *Cad. Saúde Pública, Rio de Janeiro, set-out, 2001*, p. 18, 2001.
- TIESSEN, H.; MENEZES, R.; SALCEDO, I.; B.WICK. Organic matter transformations and soil fertility in a treed pasture in semiarid ne brazil. *Plant and Soil*, p. 195–205, 2003.
- TOBLER, W. R. A computer movie simulating urban growth in the detroit region. *Economic Geography, Worcester*, v. 46, n. 1, p. 234–240, 1970.
- VIEIRA, S. R. Geoestatística em estudos de variabilidade espacial. 2000.
- WILDING, L. P.; DRESS, L. Spatial variability and pedology. In L.P. Wilding, N. Smeck and G.F. Hall (eds.). *Pedogenesis and Soil Taxonomy. Wageningen. Netherlands*, p. 83–116, 1983.

Apêndice

Exemplo do programa utilizado para as análises das variáveis sobre qualidade do solo.

```
## Carregando o pacote geoR
require(akima)
require(geoR)
require(MASS)
require(moments)
require(akima)
require(tcltk)

#rm(list=ls(all=TRUE))
## importacao de dados para o R
solo= read.table('tree2.txt', head=T)
solo[1:2,]
head(solo)

#####
# Análise descritiva dos dados para variável Carbono
geosolo <- as.geodata(solo, coords.col=c(1, 2), data.col=3)
class(geosolo)
attach(geosolo)
summary(geosolo)
plot(geosolo, low=T)
title("Dados originais")
#savePlot('fig1.png', type="png")

##### verificando os pressupostos
geosolo
boxcox(geosolo)
abline(v=0.35, col="red")
```

```

title("Transformação Box-Cox")
#savePlot('fig2.png',type="png")
plot(geosolo, low=T,lam=0.35)#trend é o efeito de tendência linear
title("Dados transformados")
#savePlot('fig3.png',type="png")

points(geosolo, pt.div="quintile", xlab="leste", ylab="norte")
title("Quantis dos dados transformados")
#savePlot('fig4.png',type="png")

media<-mean(solo$Carbono)
minimo<-min(solo$Carbono)
maximo<-max(solo$Carbono)
desvio<-sd(solo$Carbono)
CV<-100*sd(solo$Carbono)/mean(solo$Carbono)
curtose<-kurtosis(solo$Carbono)
assimetria<-skewness(solo$Carbono)

Analise_descritiva<-data.frame(media,minimo,maximo,desvio,CV,curtose,assimetria)
## Definindo uma borda (arbitrária) na area clicando em pontos
bor <- locator(type="p", pch=21)
polygon(bor)
#geosolo$borders <- with(bor, cbind(x,y));geosolo$borders
#points(geosolo, pt.div="quintile", xlab="leste", ylab="norte")
#summary(geosolo)

bord = read.table('borda.txt')
geosolo$borders <- with(bord, cbind(V1,V2))
points(geosolo, pt.div="quintile", xlab="leste", ylab="norte")

#####
# Verificando graficamente a dependência espacial com 1000 simulações
var1 = variog(geosolo, max.dist=20,lam=0.35)
plot (var1)

var2 = variog(geosolo, option="cloud",lam=0.35)

```

```
plot (var2)

var3 = variog(geosolo,uvec=seq(0,20,l=60),lam=0.35)
plot(var3)

env.var = variog.mc.env(geosolo, obj.v=var3, nsim=1000)
plot(var3, env=env.var)
title("verificando a dependência espacial através de simulações")
savePlot('fig6.png',type="png")

## Variograma visual apenas para dar "chutes" iniciais aos parâmetros
## do semivariograma
v1 = variog(geosolo, max.dist=25,lam=0.35)
plot(v1)
## Ajuste (visual) de modelo para o variograma
ef1 = eyefit(v1)
summary(ef1)

## Estimação de parâmetros (por máxima verossimilhança)
## valores iniciais ef sugerido visualmente pelo variograma
## vendo o função do modelo ajustado por máxima verossimilhança( MV)

m11 <- likfit(geosolo, ini=c(0.07,3.89), nug=0.0092,cov.model= "matern",
kappa=0.5,lam=0.35)
m11
summary(m11)

m12 <- likfit(geosolo,ini=c(0.07,3.89), nug=0.0092,cov.model= "matern",
kappa=1,lam=0.35)
m12
summary(m12)

m13 <- likfit(geosolo, ini=c(0.07,3.89), nug=0.0092,cov.model="spherical",
lam=0.35)
m13
summary(m13)
```

```

m14 <- likfit(geosolo, ini=c(0.07,3.89), nug=0.0092,cov.model="gaussian",
lam=0.35)
m14
summary(m14)

m15 <- likfit(geosolo, ini=c(0.07,3.89), nug=0.0092,cov.model="circular",
lam=0.35)
m15
summary(m15)

plot(variog(geosolo, max.dist=25,lam=0.35),ylab="Semivariância",xlab="Distância")
lines.variomodel(m11,col="red")
lines.variomodel(m12,col="yellow")
lines.variomodel(m13,col="green")
lines.variomodel(m14,col="black")
lines.variomodel(m15,col="blue")
legend("bottomright", c("gaussiano"), lty=1,col=c("black"))
title("Ajuste do variograma com diferentes função de correlação")
#savePlot('fig7.png',type="png")

#####
## malha de pontos de predição
#args(pred_grid)

gr <- pred_grid(geosolo$borders, by=0.5)

points(geosolo)
points(gr, col=2, pch=19, cex=0.3)

#args(krige.conv)
#args(krige.control)

gr0 <- gr[.geoR_inout(gr, geosolo\borders),]
points(gr0, col=3, pch=19, cex=0.3)
title("Malha de pontos para predição")

```

```
#savePlot('fig8.png',type="png")

dim(gr)
dim(gr0)

KC <- krige.control(type="OK", obj.model=m14)
OC <- output.control(n.post=1000,thres=0.8)
geosolo.kc <- krige.conv(geosolo,locations=gr0, krige=KC,output=OC)

### SIMULANDO!!

gr <- pred_grid(geosolo\$borders, by=0.5)
s.out = output.control(n.predictive = 1000, n.post = 1000, quant=0.95, thres=0.8)
geosolo.kc = krige.conv(geosolo,loc=gr, krige= krige.control(obj=m14),
output = s.out)

#argumenos existentes na função "krige.conv"
names(geosolo.kc)

par(mfrow=c(2,2))

image(geosolo.kc, col=terrain.colors(20),
val=(1-geosolo.kc\$probabilities.simulations),
main="Map of P(Y > 0.8)",ylim=c(-5,35),xlim=c(-5,35),
x.leg=c(-0,22), y.leg=c(-4,-2.5))
#savePlot('fig11.png',type="png")

image(geosolo.kc, col=terrain.colors(20),
val=geosolo.kc\$probabilities.simulations,
main="Map of P(Y < y)=0,95",ylim=c(-5,35),xlim=c(-5,35),
x.leg=c(-0,22), y.leg=c(-4,-2.5))
#savePlot('fig12.png',type="png")
```