



MARCÍLIA BRUNA DOS REIS TEIXEIRA

**COMPARAÇÃO ENTRE ESTIMADORES DE
SEMIVARIÂNCIA**

**LAVRAS - MG
2013**

MARCÍLIA BRUNA DOS REIS TEIXEIRA

**COMPARAÇÃO ENTRE ESTIMADORES DE
SEMIVARIÂNCIA**

Dissertação apresentada à Universidade Federal de Lavras, como parte das exigências do Programa de Pós-Graduação em Estatística e Experimentação Agropecuária, área de concentração em Estatística Aplicada, para a obtenção do título de Mestre.

Orientador
Dr. João Domingos Scalon

**LAVRAS - MG
2013**

**Ficha Catalográfica Elaborada pela Divisão de Processos Técnicos da
Biblioteca da UFLA**

Teixeira, Marcília Bruna dos Reis.

Comparação entre estimadores de semivariância / Marcília Bruna dos Reis Teixeira. – Lavras : UFLA, 2013.

122p. : il.

Dissertação (mestrado) – Universidade Federal de Lavras, 2013.

Orientador: João Domingos Scalon.

Bibliografia.

1. Geoestatística. 2. Semivariograma. 3. Dependência espacial. I. Universidade Federal de Lavras. II. Título.

CDD – 519.53

MARCÍLIA BRUNA DOS REIS TEIXEIRA

**COMPARAÇÃO ENTRE ESTIMADORES DE
SEMIVARIÂNCIA**

Dissertação apresentada à Universidade Federal de Lavras, como parte das exigências do Programa de Pós-Graduação em Estatística e Experimentação Agropecuária, área de concentração em Estatística Aplicada, para a obtenção do título de Mestre.

APROVADA em 26 de fevereiro de 2013.

Dr. Marcelo Silva de Oliveira UFLA

Dr. José Márcio de Mello UFLA



Dr. João Domingos Scalen
Orientador

LAVRAS - MG

2013

Aos meus amados pais, Lia e Wagner.
à minha querida irmã, Marília.
e à minha linda sobrinha, Melina.
DEDICO.

AGRADECIMENTOS

A Deus.

Aos meus pais, pelo carinho, pela dedicação e pelo apoio incondicional.

À minha irmã, Marília, pelo incentivo e conversas, pelo exemplo de determinação.

A Melina, por proporcionar os melhores sorrisos e pelos momentos de descontração mais divertidos.

Ao Vitor, pelo apoio constante, positividade e amor. Por sempre estar ao meu lado nos momentos difíceis.

Ao meu orientador, João Domingos Scalon, pelo suporte fornecido para a edificação deste trabalho, pelas sugestões e conversas sempre tão amigas. E pelos e-mails tão bem humorados.

Ao professor Marcelo Silva de Oliveira, pelo período de orientação, no qual me apresentou a Geoestatística. Por suas sugestões, ensinamentos e disponibilidade.

Aos professores José Márcio de Mello e Daniel Furtado Ferreira, pelos conhecimentos compartilhados, pela disponibilidade, pelas sugestões e críticas.

À Universidade Federal de Lavras (UFLA) e ao Programa de Pós-Graduação em Estatística e Experimentação Agropecuária, pela oportunidade concedida para a realização do mestrado.

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) e ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), pela concessão da bolsa de estudos, essencial para essa conquista.

A todos os professores do Departamento de Ciências Exatas (DEX), que participaram, direta ou indiretamente de minha formação. Assim como aos demais funcionários, em especial à Josi, que sempre tem informações e respostas para todos os problemas. E ao Marcelo do LCC, pelos preciosos conhecimentos computacionais compartilhados.

Ao professor João Marcos Louzada, pela ajuda e pelos ensinamentos compartilhados.

Aos colegas de pós-graduação, por caminharmos juntos nesta estrada.

A Laís Daniele Silva, pela positividade, por compartilhar meus anseios e sonhos e por estar ao meu lado diante das dificuldades.

A Ruth Vilela Reis, por sempre me incentivar e acreditar que seria possível concluir esta jornada.

A Cristina Henriques Nogueira, pela amizade, por estar comigo nesta caminhada, pelos risos e por compartilhar a tensão dos dias de estudo intenso.

A Solange Riveli, por suas conversas tranquilas e por todo o apoio.

“Um ‘bom modelo’ é um modelo
que começa por ‘não ser mau’ e
que depois dá bons resultados”

António Brotas

RESUMO

Esta dissertação foi elaborada com o objetivo de comparar os seguintes estimadores de semivariância: estimador clássico de Matheron, estimador robusto de Cressie e Hawkins, estimador das medianas de Cressie, estimador das diferenças de Haslett, estimador altamente robusto de Genton, estimador Pairwise, estimador New-1 e estimador New-2. As comparações foram realizadas a fim de verificar o desempenho dos estimadores diante de banco de dados com características como tendência, quantidade de *outliers* e tamanho das amostras. Foram abordados bancos de dados com malhas quadradas, sendo adotados *grides* de tamanho 5×5 , 10×10 e 20×20 . Os dados discrepantes foram em contaminações de 1%, 5%, 10% e a opção sem *outliers*. Por fim, os cenários sem tendência, com tendência linear e com tendência quadrática. A combinação dessas características resultou em 36 situações experimentais, sobre as quais foram realizadas 500 simulações. Todo o trabalho computacional foi realizado utilizando-se o *software* R. Nas estimações foram consideradas duas situações, uma que englobava todas as distâncias h existentes e uma segunda abordagem que considerava um *cutoff* de 50%, ou seja, apenas metade das distâncias foi utilizada, a fim de eliminar as estimativas das maiores distâncias que costumam ter grandes oscilações. A comparação foi realizada por meio do Erro Médio Absoluto e do Erro Médio Quadrático das estimativas de semivariância, sendo estes resultados aparados em 10% (5% em cada extremo), no intuito de eliminar possíveis oscilações do *software* nas simulações. Ao final, os resultados foram sumarizados por meio da média e da mediana para realizar as comparações. De forma geral, o melhor estimador foi o New-1, em casos de pouca contaminação. Para casos com maiores quantidades de *outliers*, os estimadores robusto de Cressie e Hawkins, das medianas e o de Genton apresentaram bons desempenhos. É importante destacar que, o ideal, é considerar um tipo de banco de dados específico para verificar qual foi o melhor estimador, visto que os resultados variam de acordo com as características encontradas nos dados.

Palavras-chave: Geoestatística. Semivariograma. Dependência Espacial.

ABSTRACT

This dissertation was designed with the purpose of comparing the following variogram estimators: Matheron's classic estimator, Cressie and Hawkin's robust estimator, Haslett's difference estimator, Genton's highly robust estimator, Pairwise estimator, New-1 estimator and New-2 estimator. The comparisons were carried out in order to verify the performance of the estimators up against the data bank with characteristics as a trend, amount of outliers and size of the samples. Data bank with square grids were approached, grids of the sizes of 5×5 , 10×10 and 20×20 being adopted. The discrepant data were in contaminations of 1%, 5%, 10% and the options with no outliers and at last, the scenarios without a bias, with a linear bias and with a quadratic bias. The combination of those characteristics resulted into 36 experimental situations over which 500 simulations were conducted. All the computer work was undertaken by utilizing the R software. In the estimations, two situations were taken into account, one which encompassed all the existing h distances and one second approach which considered a cutoff of 50%, that is, only a half of the distances was utilized in order to eliminate the estimates of the largest distances which use to have great oscillations. The comparison was undertaken by means of the Absolute Mean Error and of the Root Mean Squared Error of the estimates of semi variance, these results being trimmed by 10% (5% in each end) with the purpose of eliminating possible oscillations of the software in the simulations. At the final, the results were summarized by means of the mean and of the median to do the comparisons. In general, the best estimator was the New-1, in cases of little contamination. For cases with the greatest amounts of outliers, Cressie and Hawkins' robust estimators, estimators of the medians and Genton's estimator presented good performances. It is important to stand out that the ideal is to consider a sort of particular databank to verify what was the best estimator, since that the results ranged according to the characteristics found in the data.

Keywords: Geostatistics. Semivariogram. Spatial Dependence.

LISTA DE ILUSTRAÇÕES

Figura 1	Mapas de distribuição das observações	26
Figura 2	Janela móvel bidimensional para um campo isotrópico .	51
Figura 3	Mapa da localização espacial dos dados	76
Gráfico 1	Gráficos de dispersão <i>versus</i> coordenadas, com tendência linear	27
Gráfico 2	Gráfico de dispersão <i>versus</i> coordenadas, sem tendência	27
Gráfico 3	Gráficos de dispersão <i>versus</i> coordenadas, com tendência quadrática	28
Gráfico 4	Semivariograma	33
Gráfico 5	Semivariograma sem patamar	34
Gráfico 6	Semivariograma com efeito pepita puro	34
Gráfico 7	Modelo esférico	38
Gráfico 8	Modelo exponencial	39
Gráfico 9	Modelo gaussiano	40
Gráfico 10	EMQ <i>versus</i> quantidade de <i>outliers</i> dos dados com n=25 e sem tendência	61
Gráfico 11	EMQ <i>versus</i> quantidade de <i>outliers</i> dos dados com n=25 e com tendência linear	62
Gráfico 12	EMQ <i>versus</i> quantidade de <i>outliers</i> dos dados com n=25 e com tendência quadrática	63
Gráfico 13	Recorte do gráfico EMQ <i>versus</i> quantidade de <i>outliers</i> dos dados com n=25 e com tendência quadrática	64
Gráfico 14	EMQ <i>versus</i> quantidade de <i>outliers</i> dos dados com n=100 e sem tendência	65
Gráfico 15	EMQ <i>versus</i> quantidade de <i>outliers</i> dos dados com n=100 e com tendência linear	66
Gráfico 16	EMQ <i>versus</i> quantidade de <i>outliers</i> dos dados com n=100 e com tendência quadrática	67
Gráfico 17	Recorte do gráfico EMQ <i>versus</i> quantidade de <i>outliers</i> dos dados com n=100 e com tendência quadrática	68

Gráfico 18	EMQ <i>versus</i> quantidade de <i>outliers</i> dos dados com n=400 e sem tendência	69
Gráfico 19	EMQ <i>versus</i> quantidade de <i>outliers</i> dos dados com n=400 e com tendência linear	70
Gráfico 20	EMQ <i>versus</i> quantidade de <i>outliers</i> dos dados com n=400 e com tendência quadrática	71
Gráfico 21	Boxplot	77
Gráfico 22	Gráficos de dispersão dos dados <i>versus</i> as coordenadas .	77
Gráfico 23	Semivariograma da variável índice de sítio	79
Gráfico 24	Semivariograma da variável índice de sítio	80

SUMÁRIO

1	INTRODUÇÃO.....	14
2	REFERENCIAL TEÓRICO.....	17
2.1	A Geoestatística.....	18
2.1.1	Variáveis regionalizadas.....	21
2.1.2	Processo estocástico.....	22
2.1.3	Análise geoestatística básica.....	23
2.1.3.1	Análise exploratória.....	24
2.1.3.2	Análise variográfica.....	28
2.2	O Semivariograma.....	31
2.2.1	Semivariância.....	31
2.2.2	Semivariograma.....	32
2.2.3	Parâmetros do Semivariograma.....	32
2.2.4	Semivariograma Experimental.....	34
2.2.4.1	Isotropia e anisotropia.....	36
2.2.5	Semivariograma teórico.....	36
2.2.5.1	Modelos teóricos de semivariograma.....	37
2.2.5.2	Ajuste do modelo teórico.....	40
2.3	Estimadores de semivariância.....	41
2.3.1	Estimador clássico de Matheron.....	42
2.3.2	Estimador robusto de Cressie e Hawkins.....	44
2.3.3	Estimador das medianas de Cressie.....	46
2.3.4	Estimador das diferenças de Haslett.....	47
2.3.5	Estimador altamente robusto de Genton.....	48
2.3.6	Estimador Pairwise.....	49
2.3.7	Estimadores New-1 e New-2.....	50
3	METODOLOGIA.....	53
3.1	Simulação.....	54
3.1.1	Tamanho da amostra.....	55
3.1.2	Quantidade de <i>outliers</i>	55
3.1.3	Tendência.....	56
3.1.4	Comparação.....	57
3.2	Softwares.....	58
4	RESULTADOS E DISCUSSÕES.....	60
4.1	Resultados das comparações em <i>grides</i> 5x5.....	60
4.2	Resultados das comparações em <i>grides</i> 10x10.....	64
4.3	Resultados das comparações em <i>grides</i> 20x20.....	69

4.4	Discussão.....	72
4.5	Aplicação.....	75
5	CONCLUSÃO	81
	REFERÊNCIAS	83
	ANEXOS	87

1 INTRODUÇÃO

A busca do aperfeiçoamento dos métodos de análise de dados visa otimizar os resultados e, conseqüentemente, as implicações dos mesmos, como custos, utilização de recursos, aumento de produção, entre outros.

A metodologia Geoestatística aborda a detecção da estrutura da variabilidade espacial em superfícies contínuas. Esta detecção tem grande influência nas análises, visto que, na estatística clássica, assume-se a independência dos dados, contudo, essa independência, comumente, não condiz com a realidade. Assim, esta metodologia tende a trazer melhoras significativas nas análises realizadas.

O anseio do pesquisador sempre é o de aperfeiçoar a metodologia utilizada e a busca de uma melhor detecção da estrutura da variabilidade retornará resultados mais precisos e análises mais acuradas.

A Geoestatística busca estudar dois objetivos principais: conseguir detectar a estrutura da variabilidade espacial e uma medida da correlação existente entre os dados, por meio do semivariograma e ser capaz, por meio de um processo discreto de amostragem, de medir a precisão de toda predição ou estimação, obtendo uma estimativa contínua da área amostrada, pela técnica de krigagem.

O semivariograma é um gráfico das semivariâncias (γ) em função das distâncias (h) entre os pares de pontos. Este gráfico é extremamente importante nesta metodologia, pois é por meio dele que é possível detectar a estrutura da dependência espacial. Muitos autores enfatizam esta importância e ressaltam a necessidade de zelo na construção do semivariograma. Nesse sentido, é visível a importância de uma maior compreensão desse método, em busca de otimizar sua utilização.

Existem vários estimadores de semivariância, um dos mais utilizados é o modelo clássico de Matheron. Contudo, é comum senso que este estimador não é robusto diante de características adversas nos dados, como, por exemplo, dados discrepantes. Assim, outros estimadores surgiram, no intuito de tentar aprimorar as estimativas.

É natural que cada estimador possua características específicas, retornando resultados diferentes uns dos outros. Também é razoável acreditar que, diante de diferentes tipos de bancos de dados, estimadores diferentes apresentem os melhores resultados. Neste contexto, é importante saber qual se destaca diante de determinada característica nos dados trabalhados.

Sob a égide desse discurso, aparecem questões inerentes às diferenças dos estimadores: Qual é melhor? Em quê cada um é melhor? Surge uma necessidade de comparar os estimadores no intuito de detectar qual é melhor em cada situação e, assim, otimizar a construção do semivariograma por meio da utilização dos estimadores mais adequados. Afinal, a utilização de estimadores que reflitam a variabilidade mais próxima do valor real do fenômeno em estudo é fundamental para a confiabilidade das análises realizadas.

Assim, torna-se interessante e útil a comparação de estimadores diante de diferentes situações, por meio de bancos de dados com perfis pré-selecionados. É importante a utilização de estimadores que levem em conta toda a informação disponível nos dados, modelos que se ajustem melhor às características a serem trabalhadas. É muito útil saber quando utilizar cada estimador, no intuito de potencializar a confiabilidade dos resultados.

No rastro dessas discussões, surgiu a seguinte questão que direcionou este estudo: Qual estimador de semivariância otimiza a detecção da variabilidade espacial de um banco de dados com determinadas características específicas? Para responder a tal questionamento, esta dissertação foi elaborada com o objetivo específico de realizar um levantamento bibliográfico dos estimadores de semivariância disponíveis na literatura. Foram encontrados oito estimadores, sendo eles: estimador clássico de Matheron, estimador robusto de Cressie e Hawkins, estimador das medianas de Cressie, estimador e das diferenças de Haslett, estimador altamente robusto de Genton, estimador Pairwise, estimador New-1 e estimador New-2. Foram realizados estudos das características e propriedades de cada estimador.

Objetivou-se também encontrar características de diferentes tipos de dados de superfícies contínuas, a fim de escolher aquelas que seriam utilizadas para comparar os estimadores. Foram selecionadas algumas características referentes ao tamanho das amostras, quantidade de *outliers* e tendência.

O objetivo geral foi o de, por meio de simulação, comparar os oito estimadores utilizando o Erro Médio Absoluto e o Erro Médio Quadrático das estimativas de semivariância. Vale destacar que, nas simulações, foram abordados diferentes tipos de bancos de dados, com a combinação das características selecionadas. Também foi realizado o estudo de um banco de dados real, com o objetivo de abordar a análise geoestatística básica e de aplicar os resultados obtidos.

2 REFERENCIAL TEÓRICO

A Estatística Espacial é uma área que tem demonstrado amplo desenvolvimento. Um dos grandes diferenciais desta área são as técnicas e modelos que utilizam a referência espacial associada com cada valor dos dados. A variabilidade espacial, por exemplo, tem se tornado, cada vez mais, uma das preocupações dos pesquisadores. O desenvolvimento tecnológico permite uma facilidade maior na obtenção de dados espaciais e impulsiona fortemente os estudos e a utilização da estatística espacial. A maior capacidade computacional e o desenvolvimento de *softwares* específicos facilitam os intensos cálculos que a área exige.

A análise espacial é classificada, por autores como Bailey e Gatrell (1995), Cressie (1993) e Waller e Gotway (2004), em quatro categorias. Esta categorização diz respeito à natureza estocástica da observação. As categorias são: dados de processos pontuais, dados de área, dados de superfícies contínuas e dados de interação espacial. O que distingue estas categorias é o tipo de dado aleatório e, portanto, é natural que em cada um dos tipos existam métodos estatísticos diferentes para analisar os dados. Esta dissertação está concentrada apenas na análise dos dados de superfícies contínuas, também conhecida como Geoestatística.

A Geoestatística é uma metodologia que permite identificar a existência ou não de dependência espacial entre as observações. Ela pode ser aplicada em mapeamentos, orientação de futuras amostragens e modelagens, permitindo, assim, estimar o valor do atributo em locais não amostrados, facilitando a gestão de recursos naturais, entre outros exemplos de aplicação possíveis.

Como exemplo específico, sabe-se que as propriedades do solo diferenciam-se por seus padrões de variabilidade espacial e temporal. Assim, a adoção de técnicas e métodos de gerenciamento mais eficazes tende a aumentar a produção. Este melhor gerenciamento pode resultar em redução de custos para o agricultor, por exemplo, fazendo com que os insumos agrícolas sejam aplicados somente onde, de fato, se faz necessário e não

indiscriminadamente em toda a extensão da área cultivada. A principal contribuição da Geoestatística, neste processo, é a gerência adequada do solo, por meio de um estudo detalhado das áreas de plantio, construindo-se mapas temáticos que auxiliam na aplicação correta dos insumos, diminuindo os impactos e otimizando os custos.

Outro exemplo da utilização da Geoestatística vem da estimativa de reservas, que é uma parte fundamental do planejamento na área de mineração. O conhecimento geológico que se tem a respeito do bem mineral a ser explorado é essencial, pois, para investir na exploração e decidir se o empreendimento será lucrativo ou não, deve-se estimar o tamanho da reserva, processo este que necessita de precisão.

Soares (2000) enfatiza que os avanços e saltos qualitativos desta metodologia são diretamente proporcionais aos investimentos que se fazem na investigação e na divulgação da disciplina. Este impulso deve-se também ao desenvolvimento da tecnologia computacional, já que a metodologia requer muitos cálculos.

Enfim, a Geoestatística é uma área de grande utilidade e importância, sendo relevante o seu estudo e o aprimoramento de seus métodos de análise. Nesta dissertação, objetiva-se o estudo do semivariograma, que é um gráfico utilizado nesta metodologia e que será definido e detalhado mais a frente. A princípio, serão apresentados alguns pontos relevantes da metodologia, para que, a partir deste conhecimento, possamos aprofundar no assunto específico, tema deste trabalho.

2.1 A Geoestatística

Soares (2000) delinea o objeto da Geoestatística como a caracterização dos fenômenos espaciais naturais e a quantificação da incerteza desse conhecimento. O autor explica que, partindo de um conjunto discreto e limitado de dados experimentais, a Geoestatística permite a concepção de modelos que visam descrever as diferentes realidades de cada estudo, definindo-a assim, como um conjunto de métodos,

técnicas e instrumentos estatísticos que caracterizam os fenômenos espaciais naturais. É importante enfatizar que, nesta abordagem, os dados devem ser referenciados espacialmente.

De forma geral, a ideia é que quanto mais próximos estejam os pontos amostrados, mais parecidos serão seus valores. Assim, uma medida amostrada em um determinado ponto mantém relações de dependência com medidas tomadas em pontos adjacentes, obtendo-se uma estrutura de correlação. Em outras palavras, observações apresentam características de acordo com a posição espacial que ocupam e não somente com relação ao valor que assumem. A partir desta concepção é que a Geoestatística descreve e modela estas relações entre distâncias e dependência.

Apesar de sua complexidade formal, basicamente, a Geoestatística se fundamenta em dois conceitos: semivariograma e krigagem, sendo o primeiro um gráfico das semivariâncias *versus* as distâncias, que permite descrever a estrutura da variabilidade espacial de um fenômeno regionalizado no espaço, enquanto o segundo é um interpolador que prediz, não-tendenciosamente e com variância mínima, os valores não observados (CRESSIE, 1993).

A Geoestatística tem sido cada vez mais utilizada. E, como foi exemplificado anteriormente, esta área possui várias e importantes aplicações. Um destaque nesta metodologia, ressaltado por Pontes (2002, p. 35), é que, nela, "a correlação espacial entre as observações não é considerada um incômodo a ser evitado, mas sim uma fonte de informações que melhora a análise dos dados". A dependência espacial sempre esteve no centro das atenções da experimentação. A dificuldade em lidar com este fator aumenta consideravelmente o erro experimental, sendo visível a necessidade de uma análise mais acurada. Assim, vale enfatizar que a Geoestatística apresenta muitas vantagens como metodologia que considera a dependência espacial.

Andriotti (2003, p. 92) discursa sobre os benefícios desta teoria, afirmando que

As vantagens reconhecidas da Geoestatística sobre outras técnicas convencionais de predição são os seguintes: o estudo da variabilidade espacial (a análise de um variograma é a única técnica disponível para medir a variabilidade espacial de uma variável regionalizada), a suavização (a estimação Geoestatística suaviza ou faz a regressão de valores preditos, baseada na proporção da variabilidade total da amostra levada em conta pelas variações aleatórias; quanto maior for a variabilidade da amostra, menos as observações individuais representam a sua vizinhança imediata e mais elas são suavizadas, com maior incerteza associada), o desagrupamento (ou efeito de anular as concentrações localizadas de observações), a determinação da anisotropia (os comportamentos da variabilidade nas diferentes direções são considerados), a precisão (a krigagem fornece valores precisos sobre as áreas ou pontos a serem avaliados) e a incerteza (estimativa obtida por meio da krigagem associa a margem de erro que acompanha a estimativa).

Em resumo, a Geoestatística vem se destacando, adquirindo cada vez mais adeptos e tem trazido vários resultados satisfatórios. Grondona e Cressie (1991) afirmam que a Geoestatística vem propiciando resultados eficientes e confiáveis. Hoje em dia, ela está sendo aplicada em diversas áreas das ciências humanas, biológicas e exatas (GUIMARÃES, 2004), tornando-se popular em vários campos da ciência e da indústria, em que existe necessidade de se avaliar dados espacialmente ou temporalmente correlacionados.

Para melhor compreender esta metodologia, é importante definirmos alguns conceitos. Nas próximas subseções serão tratados alguns tópicos importantes da Geoestatística.

2.1.1 Variáveis regionalizadas

Segundo Guerra (1988), as variáveis regionalizadas são como uma função que varia de um lugar a outro no espaço, mantendo a suavidade, ou seja, exibindo um aspecto de continuidade. Camargo (1998) especifica que uma variável regionalizada é uma função numérica com distribuição espacial, que varia de um ponto a outro com continuidade aparente.

Na teoria das variáveis regionalizadas, $Z(x)$ pode ser definida como uma variável aleatória que assume diferentes valores Z em função do vetor de posição x dentro de uma certa região. O conjunto formado pelas variáveis $Z(x)$ medidas em uma determinada área pode ser considerado uma função aleatória $Z(x)$, uma vez que, segundo Isaaks e Srivastava (1989), são variáveis aleatórias, regionalizadas e assume-se que a dependência entre elas é especificada por algum mecanismo probabilístico.

Assim, a teoria das variáveis regionalizadas pressupõe que a variação de uma variável pode ser expressa pela soma de três componentes: **a)** uma componente estrutural, associada a um valor médio constante ou a uma tendência constante; **b)** uma componente aleatória, espacialmente correlacionada, e **c)** um ruído aleatório ou erro residual. Seja x um vetor de coordenadas espaciais, então, o valor da variável Z , em x , é dada por Burrough (1987 apud CAMARGO, 1998),

$$Z(x) = m(x) + \varepsilon'(x) + \varepsilon'', \quad (2.1)$$

em que:

- a)** $m(x)$ é uma função determinística que descreve a componente estrutural de Z em x ;
- b)** $\varepsilon'(x)$ é um termo estocástico, que varia localmente e depende da localização espacial;
- c)** ε'' é um ruído aleatório não correlacionado, que possui média zero e variância σ^2 e que, às vezes, apresenta normalidade.

Andriotti (2003) elucida que a Geoestatística trata as variáveis como regionalizadas, podendo, mesmo, ser considerada uma aplicação prática da Teoria das Variáveis Regionalizadas.

2.1.2 Processo estocástico

Nos modelos geoestatísticos, os dados amostrais são interpretados como provenientes de um processo aleatório. Um valor localizado espacialmente em x_1 (denominação genérica de um conjunto de coordenadas geográficas) é interpretado como uma realização $z(x_1)$ da variável aleatória $Z(x_1)$

Segundo Oliveira (1991), um processo estocástico (real) é uma coleção $\{Z(x) : x \in R \subset \mathbb{R}^p\}$ de variáveis aleatórias reais, definidas sobre um mesmo espaço de probabilidade, indexadas em um subconjunto R do espaço vetorial p -dimensional \mathbb{R}^p .

O autor salienta que a Geoestatística baseia-se no pressuposto de que a distribuição espacial e/ou temporal de uma variável z em uma região R é uma realização $\{z(x) : x \in R \subset \mathbb{R}^p\}$ do processo estocástico $\{Z(x) : x \in R \subset \mathbb{R}^p\}$.

Neste contexto, na Geoestatística trabalha-se com algumas hipóteses, sendo elas a hipótese intrínseca e a estacionariedade de segunda ordem.

Segundo Cressie (1993), a hipótese intrínseca é definida por

- (i) a esperança matemática existe e não depende da referência x :

$$E[Z(x)] = \mu, \quad \forall x;$$

- (ii) para todo h , a variância da diferença $[Z(x+h) - Z(x)]$ existe e não depende de x :

$$Var[Z(x+h) - Z(x)] = E([Z(x+h) - Z(x)]^2) = 2\gamma(h),$$

sendo μ uma constante real e $2\gamma(h)$ o valor da variância, que será melhor definida mais a frente.

Observe que de **(i)** temos que $E[Z(x+h) - Z(x)] = 0$. E, devido a isso, que $Var[Z(x+h) - Z(x)] = E([Z(x+h) - Z(x)]^2) - (E[Z(x+h) - Z(x)])^2 = E([Z(x+h) - Z(x)]^2)$

Já a estacionariedade de segunda ordem pode ser definida por

- (i)** a esperança matemática $E[Z(x)]$ existe e não depende da referência x , assim,

$$E[Z(x)] = \mu = \text{constante};$$

- (ii)** para cada par $\{Z(x), Z(x+h)\}$, a covariância (Cov) existe e depende apenas da distância de separação h ,

$$Cov(Z(x+h), Z(x)) = \kappa(h).$$

A hipótese de estacionariedade de segunda ordem é muito restritiva, pois assume a existência de uma covariância e, portanto, de uma variância finita. Nesta hipótese, a estrutura de dependência espacial estabiliza a partir de uma determinada distância h entre os pontos amostrados. Mas, existem muitos fenômenos que têm uma capacidade infinita de dispersão, que não têm nem uma variância *a priori*, nem uma covariância. Assim, muitas vezes, é utilizada a hipótese intrínseca (JOURNEL; HUIJBREGTS, 1978), que não exige esta estabilização da estrutura de variabilidade espacial.

2.1.3 Análise geoestatística básica

Para realizar a análise básica Geoestatística de um banco de dados, assim como na Estatística Clássica, primeiramente, é necessário realizar um estudo exploratório do mesmo. Após conhecer melhor os dados, prossegue-se com a detecção da estrutura da variabilidade espacial, por meio da construção do semivariograma. Ao final, podem-se estimar pontos não

amostrados e criar mapas pelo processo de krigagem. Estes tópicos serão mais bem explanados nas subsecções que seguem.

2.1.3.1 Análise exploratória

Em qualquer estudo de dados é importante, inicialmente, realizar uma análise exploratória. É indispensável que o pesquisador conheça a variável em estudo, aliando o conhecimento prévio sobre a variável que está sendo trabalhada e as análises estatísticas realizadas. Assim, além de haver um aprimoramento do estudo e da interpretação dos dados, há uma melhor detecção de erros e incoerências nos dados, análises e resultados.

Para uma análise inicial, é interessante observar algumas estatísticas descritivas, como **medidas de localização**: média, valor mínimo, quartil inferior, mediana, quartil superior e valor máximo, entre outros; **medidas de dispersão**: variância e desvio padrão; **medidas de forma**: coeficiente de assimetria, coeficiente de curtose e coeficiente de variação, entre outros. Estes valores ajudam a descrever e a conhecer melhor a variável em estudo. Também é interessante utilizar análise gráfica como, por exemplo, o gráfico de dispersão.

Um estudo que deve ser realizado com atenção é em relação à existência de *outliers*. Esta é uma das características utilizadas, nesta dissertação, na comparação dos estimadores de semivariância. Os valores discrepantes afetam a análise dos dados, mascarando as estatísticas calculadas. É importante detectar estes valores e estudá-los, a fim de decidir qual a melhor forma de analisá-los. Existem teorias que indicam a retirada dos *outliers*, a sua substituição por um valor médio e a utilização de estimadores robustos, entre outras. O ideal é o pesquisador tentar conhecer o banco de dados e a variável estudada, pois, assim, pode-se detectar se houve erro na amostragem ou se o fenômeno possui alguma variação ou característica que deve ser investigada.

Existem vários métodos de detecção de dados discepanetes. Inclusive, na literatura encontram-se métodos específicos no caso de dados

georeferenciados. Contudo, como não é a especificidade da dissertação, por simplicidade será utilizado o gráfico *box-plot* para realizar esse tipo de análise.

Outra característica que influencia as estimativas da semivariância é a presença de tendência nos dados avaliados, ou seja, a não estacionariedade da variável. Ela influencia fortemente a construção deste gráfico, pois dificulta o ajuste de modelos adequados e mascara o verdadeiro comportamento espacial da variável. A tendência pode até omitir a existência do patamar e, por isso, deve-se fazer uma análise prévia dos dados, garantindo a ausência de tendência.

A partir do gráfico de distribuição das observações é possível observar a localização espacial dos pontos amostrados, nele, os pontos são identificados de acordo com o quantil ao qual pertencem. Por meio deste gráfico pode-se proceder a uma análise preliminar baseada em inspeção visual, a qual permite a identificação de alguns problemas, tais como escassez ou excesso de observações em determinado local, ocorrência de valores discrepantes, etc. Na Figura 1, observa-se uma malha quadrada de 10 x 10 pontos, totalizando 100 pontos. No gráfico **(a)** é possível verificar indícios claros de tendência, sendo fácil perceber que os dados estão dispostos de forma organizada, de acordo com o valor da variável observada, o que caracteriza a existência de uma tendência. Neste caso, deve-se retirar a tendência e trabalhar com os resíduos. Já em **(b)** tem-se um gráfico já com os resíduos, no qual isso não ocorre. Neste gráfico, os dados estão dispostos de forma aleatória, o que indica a não existência de tendência.

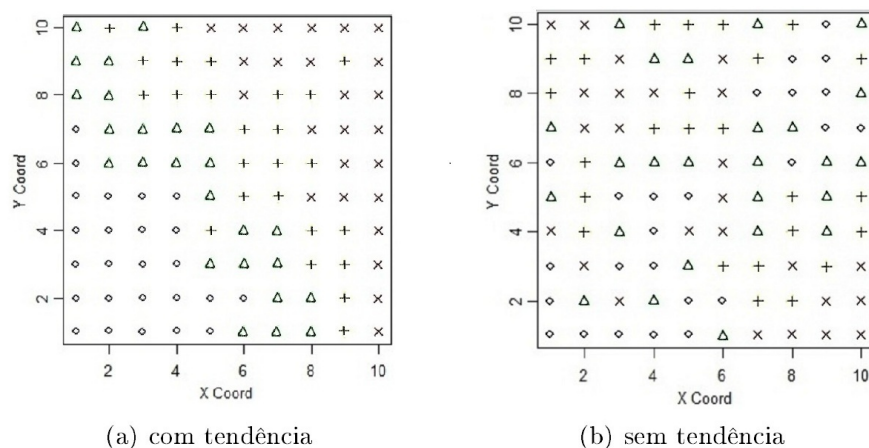


Figura 1 Mapas de distribuição das observações

em que

- a) ○ representa as observações que estão no primeiro quartil;
- b) △ representa as observações que estão entre o primeiro e o terceiro quartis;
- c) + representa as observações que estão entre o terceiro e o quarto quartis;
- d) × representa as observações que estão acima do quarto quartil.

É muito comum utilizar o mapa da distribuição das observações e gráficos de dispersão dos dados *versus* suas coordenadas para detectar a existência (ou não existência) de tendência. No Gráfico 1 e no Gráfico 2 encontram-se exemplos de gráficos de dispersão *versus* as coordenadas cartesianas, sendo que, no primeiro, verifica-se a existência de tendência linear nos valores, visto que a disposição inclinada dos dados demonstra um comportamento linear nas duas direções. Neste caso, deve-se retirar a tendência e trabalhar com os resíduos. Já no Gráfico 2 pode-se dizer que não há tendência, pois a forma que os dados estão dispostos pode ser representada por uma linha paralela ao eixo de referência.

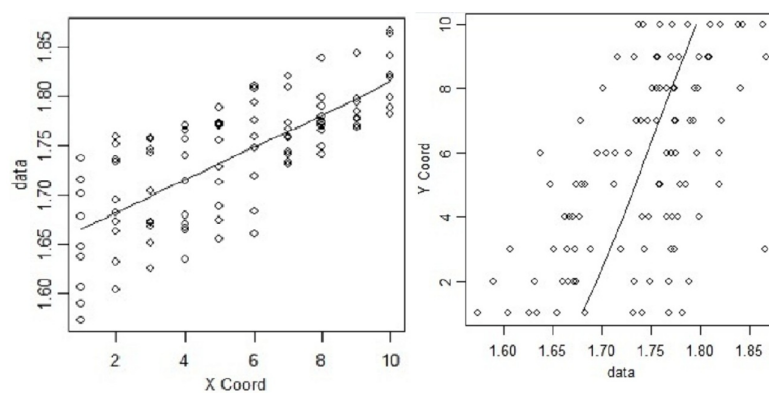


Gráfico 1 Gráficos de dispersão *versus* coordenadas, com tendência linear

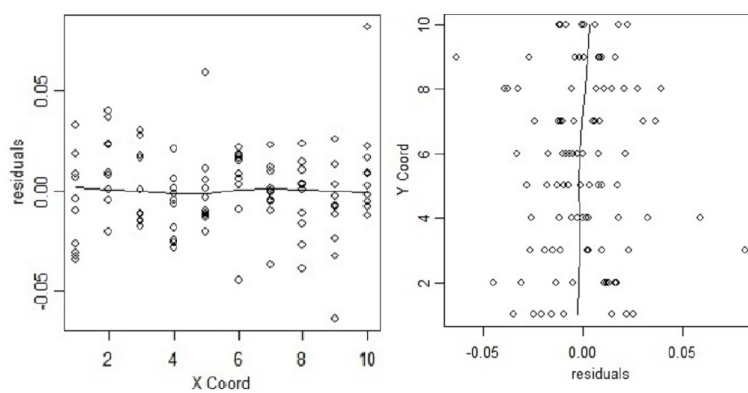


Gráfico 2 Gráfico de dispersão *versus* coordenadas, sem tendência

Também vale destacar que existem outros tipos de tendência. Por exemplo, no Gráfico 3, os pontos demonstram um formato parabólico, o que indica a existência de uma dependência quadrática.

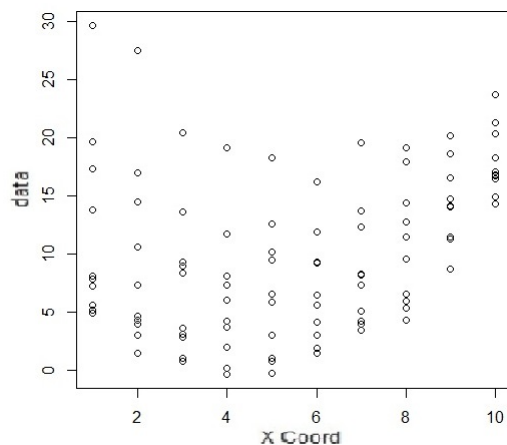


Gráfico 3 Gráficos de dispersão *versus* coordenadas, com tendência quadrática

2.1.3.2 Análise variográfica

Existem várias funções que podem ser utilizadas para estudar a relação existente entre as variáveis, tais como a covariância (que gera o covariograma), a correlação (que gera o correlograma) e a semivariância (que gera o semivariograma). Guimarães (2004) destaca que, na Geoestatística, a medida normalmente utilizada é a semivariância.

A covariância é uma medida que verifica como duas variáveis estão relacionadas. Segundo Journel e Huijbregts (1978), sob a hipótese de estacionaridade de segunda ordem, a covariância e a variância são dois métodos equivalentes para caracterizar as correlações entre duas variáveis correlacionadas $Z(x+h)$ e $Z(x)$ separadas por uma distância h . A variância é um caso particular da covariância quando $h = 0$, ou seja,

$$Cov(Z(x), Z(x)) = E[Z^2(x)] - \mu^2 = Var[Z(x)]. \quad (2.2)$$

A correlação é a covariância normada pela variância e $\rho(0) = 1$:

$$\rho(h) = \frac{\kappa(h)}{Var[Z(x)]} = 1 - \frac{\gamma(h)}{Cov(0)}. \quad (2.3)$$

Outra forma de definir a correlação, é por meio da relação existente entre κ e γ

$$\rho(h) = \frac{\kappa(h)}{\kappa(0)} = 1 - \frac{\gamma(h)}{\kappa(0)}. \quad (2.4)$$

Ainda é importante notar que, se ocorre a estacionaridade de segunda ordem, o correlograma e o variograma serão funções correspondentes para a detecção da estrutura de variabilidade espacial. Mas, se a estacionaridade de segunda ordem não é atendida, não é possível utilizar o covariograma (GUIMARÃES, 2004). Soares (2000) explica que a preferência da utilização do semivariograma na prática Geoestatística é derivada do fato de que em muitos processos aleatórios não existe a covariância.

Ao contrário da covariância e da correlação, a semivariância é uma medida de dissimilaridade, ou seja, o valor da semivariância é maior à medida que as variáveis estão menos associadas, fato este que facilita a interpretação dos resultados.

Segundo Andriotti (2003), o semivariograma pode ser interpretado como uma técnica estatística utilizada na verificação da presença de dependência espacial entre os pontos amostrais georreferenciados espacialmente, representado por um gráfico da função semivariância *versus* a sua distância. É comum delimitar uma distância máxima, para qual a semivariância é estimada, sendo esta chamada de *Cutoff*. As estimativas de semivariância de distâncias maiores que a distância máxima costumam ter uma estimativa menos precisa, devido à menor quantidade de pares de pontos utilizados e acabam perdendo o sentido prático (CLARK, 1979).

Após a análise exploratória, é feita a análise variográfica, na qual realiza-se a construção do semivariograma. Duas formas de obter esta construção são comumente encontradas na literatura. A primeira, é a realização de um ajuste direto, por meio do Método de Máxima Verossimilhança. Contudo, este método gera de forma direta o desenho do semivariograma, o que impossibilita fazer um estudo mais detalhado da estrutura de variabilidade espacial, já que não é possível detectar oscilações e/ou discrepâncias nos valores da semivariância.

Assim, neste trabalho, recomenda-se a utilização da segunda forma de construção do semivariograma, que é feita utilizando-se estimadores de semivariância. Por meio deles, é possível construir uma nuvem de pontos que caracteriza a estrutura de dependência especial. Como necessitamos de uma representação contínua desta dependência, então, posteriormente, ajustamos um modelo teórico de semivariograma. Por meio da nuvem de pontos, pode-se decorrer um estudo mais detalhado da variabilidade, possibilitando que o processo de construção do semivariograma seja mais minucioso e interativo.

Para verificar a qualidade do ajuste, comumente utiliza-se a validação cruzada, que é uma forma de checar as suposições sobre o modelo, podendo-se verificar se a modelação foi correta e/ou bem ajustada. No procedimento de validação cruzada, uma amostra é removida e posteriormente estimada por meio das demais. Após esta estimação, o valor real desta amostra é reintroduzido nos dados e o processo se repete para cada uma das outras amostras. No final, comparam-se as amostras reais com os valores estimados para cada uma, respectivamente. Quanto mais próximo de zero for o valor médio das diferenças, melhor terá sido o ajuste.

Vale destacar que alguns cuidados devem ser tomados para se realizar a análise da continuidade espacial. É importante observar a qualidade e o tipo de amostragem e estudar os *outliers*, observando se são erros da coleta de dados, ou se, de fato, representam alguma anormalidade no fenômeno estudado. Deve-se tentar reduzir os erros de amostragem, pois estes contribuem para o aumento da descontinuidade na origem (efeito pepita). Deve-se ficar atento ao fato de que diferentes tipos de dados referentes ao mesmo atributo talvez não tenham a mesma estrutura de variabilidade, devendo ser estudados separadamente.

Por meio da estrutura de variabilidade fornecida pelo semivariograma, é possível estimar pontos não amostrados e construir mapas de contorno por meio da krigagem, que é um interpolador. Ela é uma técnica usada na Geoestatística para estimar valores para locais não

amostrados, considerando os parâmetros do semivariograma, que resulta em valores sem tendência e com variância mínima. Contudo, não é de interesse, neste trabalho, aprofundar neste tópico.

2.2 O Semivariograma

Um pilar fundamental da abordagem Geoestatística referente-se à detecção, por meio do semivariograma, do modo e do grau da variabilidade espacial. Pontes (2002) explica que, por meio deste gráfico, são conduzidos ajustes, em busca de determinar um modelo de variabilidade dentro de uma região, fornecendo, assim, parâmetros adequados ao entendimento da estrutura de dependência espacial e ao procedimento de krigagem.

2.2.1 Semivariância

A semivariância é uma medida do nível de dependência entre duas variáveis regionalizadas, $Z(x_i)$ e $Z(x_i + h)$, que é representada por $\gamma(h)$ e é definida como a metade da esperança matemática do quadrado da diferença entre os valores de pontos no espaço, separados pelo vetor distância h , isto é,

$$\gamma(h) = \frac{1}{2}E[Z(x+h) - Z(x)]^2 = \frac{1}{2}\{Var[Z(x+h) - Z(x)]\}. \quad (2.5)$$

Como exemplifica Guimarães (2004), a ideia é de que quanto mais próximos estiverem os pontos amostrados, maior será a semelhança entre eles e, portanto, menor a semivariância; e, quanto mais distantes estiverem os pontos amostrados, menor será a semelhança e, conseqüentemente, maior a dispersão (variância).

Na literatura, é possível encontrar abordagens que adotam a utilização de $2\gamma(h)$, que é a variância, a qual gera o variograma. No entanto, por conveniência matemática, é mais comum a utilização da semivariância

$(\gamma(h))$, que gera o semivariograma. Contudo, a utilização de ambos retorna resultados equivalentes.

2.2.2 Semivariograma

O semivariograma é um gráfico da função semivariância *versus* as distâncias h . Ele pode ser interpretado como uma técnica estatística utilizada na verificação da presença de dependência espacial entre pontos amostrais georreferenciados.

Segundo Campoza (1990), existem três tipos de semivariograma para cada variável regionalizada, que são:

1. observado ou experimental, que é obtido a partir do conjunto de dados disponíveis das variáveis locais;
2. variograma verdadeiro, que é o real, e representa o verdadeiro comportamento da variável regionalizada, o qual é sempre desconhecido;
3. variograma teórico, o qual é descrito por funções matemáticas, e é utilizado para ajustar o variograma experimental.

Contudo, quando a construção do semivariograma é feita diretamente por meio do Método de Máxima Verossimilhança, apenas os tópicos **(2)** e **(3)** são considerados.

2.2.3 Parâmetros do Semivariograma

Um semivariograma empírico com características muito próximas do ideal pode ser observado no Gráfico 4. Os padrões apresentados representam o que, intuitivamente, se espera de dados de campo, isto é, que observações mais próximas apresentem mais semelhanças no comportamento do que aquelas mais distantes entre si. Dessa maneira, é esperado que γ aumente com a distância h .

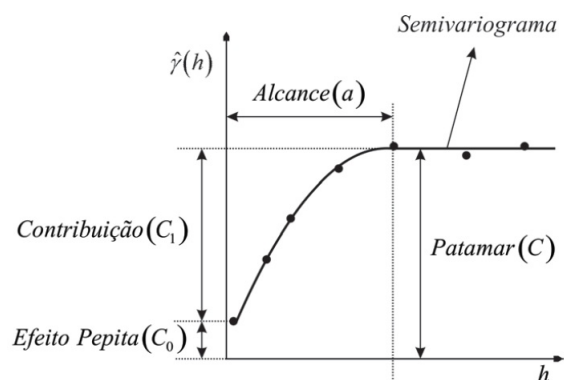


Gráfico 4 Semivariograma

O semivariograma apresenta os seguintes parâmetros (GUIMARÃES, 2004):

- a)** alcance (a): é a distância h , dentro da qual as amostras apresentam-se correlacionadas espacialmente. A partir deste valor gráfico estabiliza, ou seja, torna-se aproximadamente constante. As medições realizadas em distâncias maiores que a , são consideradas independentes entre si. Alguns autores denominam o “alcance” de “amplitude”;
- b)** efeito pepita (C_0): é uma descontinuidade na origem. O valor da semivariância no ponto $h = 0$ é nula, visto que a variância de um valor amostrado com ele mesmo é zero. Contudo, a curva do semivariograma próximo da origem costuma sofrer uma descontinuidade. Isso acontece devido a variações que podem ocorrer a distâncias menores do que a menor distância amostrada, assim como também pode ser proveniente de erros na amostragem (ISAACS; SRIVASTAVA, 1989), erros na análise laboratorial, entre outros fatores;
- c)** contribuição (C_1): é a diferença entre o patamar (C) e o efeito pepita (C_0);
- d)** patamar (C): $C_0 + C_1$ é o valor da semivariância correspondente à distância a . Deste ponto em diante, considera-se que não existe mais dependência espacial entre as amostras porque a variância da diferença

entre pares de amostras ($Var[Z(x+h) - Z(x)]$) torna-se invariante com a distância.

Também existem os casos em que o semivariograma não tem patamar (Gráfico 5), ou seja, existe dispersão infinita e os casos em que não há ou não foi detectada dependência espacial, havendo assim um efeito pepita puro (Gráfico 6).

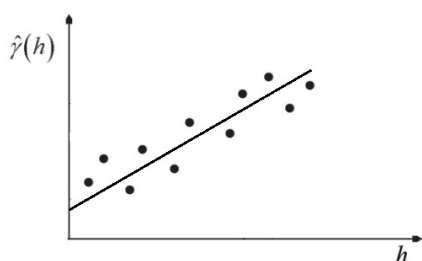


Gráfico 5 Semivariograma sem patamar

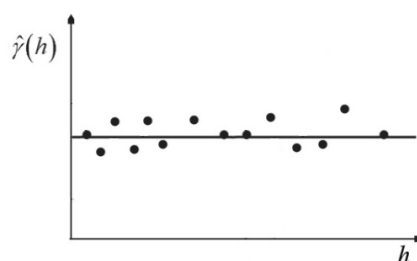


Gráfico 6 Semivariograma com efeito pepita puro

2.2.4 Semivariograma Experimental

O semivariograma experimental é representado por uma nuvem de pontos obtidos por meio de um estimador de semivariância. Trabalhos como o de Mingoti e Rosa (2008) apontam a relevância da escolha do estimador de semivariância para a qualidade final da análise.

Um dos estimadores de semivariância mais comumente utilizados na literatura é o modelo clássico proposto por Matheron (1962). Segundo Genton (1998) e Mingoti e Rosa (2008), o estimador de variograma clássico proposto por Matheron é muito afetado pela presença de *outliers* nos dados, não apresentando propriedades de robustez, nem sendo suficiente para fazer modificações simples, como os propostos por Cressie e Hawkins (1980) a fim de alcançar robustez.

Existem outros estimadores na literatura, tendo muitos sido elaborados com o intuito de serem mais robustos, como os de Cressie e Hawkins (1980), que foi construído para ser robusto contra *outliers*; das Medianas (CRESSIE, 1993); o estimador altamente robusto de Genton (1998), que busca ser mais robusto diante de *outliers*; o estimador das diferenças proposto por Haslett (1997), utilizado num contexto de séries temporais, especialmente para dados não estacionários; o estimador de Pairwise (ISAAKS; SRIVASTAVA, 1989) que tenta controlar os dados discrepantes por meio de uma ponderação e os estimadores New-1 e New-2 (LI; LAKE, 1994) que, segundo os autores, são consideravelmente robustos.

Os semivariogramas podem apresentar diversas características. Por exemplo, os semivariogramas que apresentam estacionaridade de segunda ordem para a variável exibem o patamar claro e bem definido. Se o semivariograma for constante e igual ao patamar para qualquer valor de h , ele é denominado por efeito pepita puro. Nele há ausência total de dependência espacial, ou seja, a dependência espacial, se existir, será manifestada à distância h (ou tempo) menor do que o menor espaçamento entre amostras.

Os semivariogramas podem apresentar flutuações, sendo, então, chamados de semivariogramas periódicos ou cíclicos. E eles indicam uma periodicidade nos dados. Existem também os que não possuem patamar definido, o que mostra que a hipótese de estacionaridade de segunda ordem não foi atendida e, provavelmente, estamos trabalhando com a hipótese intrínseca (fenômeno com capacidade infinita de dispersão). Ele pode indicar ainda a necessidade de se amostrar a maiores distâncias, ou a existência de tendência nos dados.

Pode haver também um semivariograma com mais de uma estrutura de variância, que é chamado de semivariogramas com estruturas entrelaçadas ou semivariogramas imbricados.

A distância máxima para qual a semivariância é estimada é chamada *Cut-off*. Os pontos que estão posicionados além desta distância máxima são considerados não influentes (ISAAKS; SRIVASTAVA, 1989). O

cálculo das semivariâncias não deve exceder a valores de h superiores a, aproximadamente, metade da maior distância (CLARK, 1979; SOARES, 2000).

2.2.4.1 Isotropia e anisotropia

Outro fator de importância relevante na análise estrutural é a anisotropia, que exerce grande influência na continuidade espacial dos fenômenos naturais, pois, num meio anisotrópico, as propriedades físicas são diferentes ao longo de direções distintas. Se os variogramas para diferentes direções tiverem o mesmo comportamento, então, há isotropia. Ou seja, o variograma não depende da direção em que se estuda o fenômeno (ANDRIOTTI, 2003). Quando o variograma experimental revelar um comportamento muito diferente, quando construído considerando-se diferentes direções, então, neste caso, há anisotropia. Para analisar em qual categoria os dados se enquadram, basta construir variogramas direcionais e observar se estes diferem entre si. Os variogramas direcionais mais frequentemente utilizados são os referentes a 0° , 45° , 90° e 135° do eixo adotado pelo pesquisador.

A anisotropia pode ser zonal, geométrica ou uma combinação de ambas. Na anisotropia geométrica, o patamar é constante e os alcances são distintos. Na anisotropia zonal, o alcance é único, mas com patamares distintos em diversas direções. É mista quando a análise resulta em variogramas com patamares e alcances distintos em direções distintas

2.2.5 Semivariograma teórico

Soares (2000) afirma que a etapa da escolha do semivariograma teórico é extremamente importante, pois é neste ponto que se faz a síntese das características estruturais do fenômeno espacial num modelo único. O modelo teórico, diferentemente do semivariograma experimental, representa

a estrutura de dependência espacial de forma contínua, fornecendo as estimativas de semivariância para todas as possíveis distâncias h . Segundo o autor, não se trata de obter um modelo contínuo $\gamma(h)$ de uma função aleatória estacionária $Z(x)$, mas sim de modelizar e sintetizar os principais padrões de continuidade espacial representativos do fenômeno em estudo.

2.2.5.1 Modelos teóricos de semivariograma

Os modelos variográficos teóricos podem ser divididos, segundo o patamar, em dois grupos, os que atingem um patamar, sendo coerentes com a hipótese de estacionariedade de 2^a ordem, e os que não têm patamar. Para os primeiros, pode-se fazer uma subdivisão em três categorias:

- a) modelos com comportamento linear próximo à origem;
- b) modelos que apresentam comportamento parabólico junto à origem, e
- c) modelos de comportamento constante, o efeito pepita.

Segundo Soares (2000), os variogramas têm de ser definidos positivos, para que se tenham soluções estáveis no processo de krigagem. Esta condição limita a escolha dos modelos teóricos a um conjunto de funções autorizadas. Aqui serão apresentados apenas três modelos, sendo estes com patamar.

- 1 **Modelo esférico:** segundo Andriotti (2003), este é o modelo mais comumente utilizado nas variáveis em geociências. Apresenta um crescimento rápido na origem, onde tem um comportamento linear.

$$\gamma(h) = \begin{cases} 0, & \text{se } h = 0 \\ C_0 + C_1 \left[\frac{3}{2} \left(\frac{h}{a} \right) - \frac{1}{2} \left(\frac{h}{a} \right)^3 \right], & \text{se } 0 < h \leq a \\ C_0 + C_1, & \text{se } h > a \end{cases} \quad (2.6)$$

em que

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

C_0 é o efeito pepita;

C_1 é a contribuição;

a é o alcance;

h são os valores das distâncias.

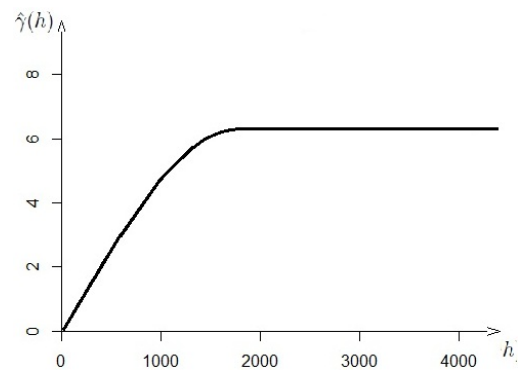


Gráfico 7 Modelo esférico

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

$$\gamma(h) = \begin{cases} 0, & \text{se } h = 0 \\ \gamma(h) = C_0 + C_1 [1 - e^{-3(h/a)}], & \text{se } h \neq 0 \end{cases} \quad (2.7)$$

em que

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

C_0 é o efeito pepita;
 C_1 é a contribuição;
 a é o alcance;
 h são os valores das distâncias.

Apesar de o modelo crescer junto da origem mais rápido que o modelo esférico, as estruturas apresentam uma maior continuidade espacial, isto é, maiores prolongamentos, resultantes das maiores correlações espaciais a grandes distâncias de h (SOARES, 2000).

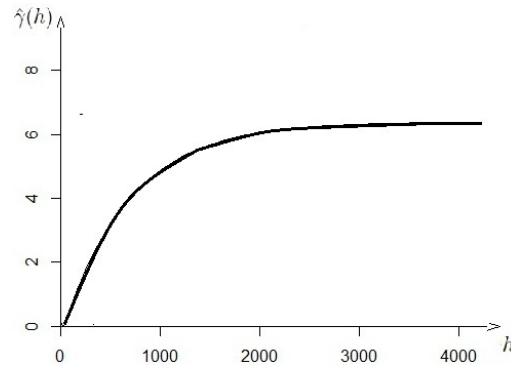


Gráfico 8 Modelo exponencial

3 Modelo gaussiano: tem comportamento parabólico nas vizinhanças da origem e reflete uma grande continuidade da variável estudada, devido ao seu crescimento suave. O modelo gaussiano é um modelo transitivo, muitas vezes utilizado para modelar fenômenos extremamente contínuos (ISAACS; SRIVASTAVA, 1989).

$$\gamma(h) = \begin{cases} 0, & \text{se } h = 0 \\ \gamma(h) = C_0 + C_1 \left[1 - e^{-3(h/a)^2} \right], & \text{se } h \neq 0 \end{cases} \quad (2.8)$$

em que

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

C_0 é o efeito pepita;

C_1 é a contribuição;

a é o alcance;

h são os valores das distâncias.

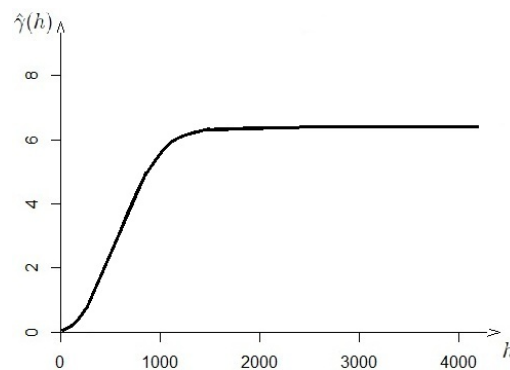


Gráfico 9 Modelo gaussiano

2.2.5.2 Ajuste do modelo teórico

Esta etapa é muito importante, pois é quando se define a estrutura de variabilidade final, que vai ser utilizada na krigagem. Para ajustar o modelo, devem-se utilizar conhecimentos prévios do processo físico, e não apenas realizar um ajustamento automático.

Schabenberger e Gotway (2004) explicam que o semivariograma não é apenas um dispositivo para derivar a estrutura de dependência espacial e para construir a matriz de variância-covariância de $Z(x)$, que é necessária para o modelo com base em inferências estatísticas. É um instrumento estrutural que, por si só, transmite muita informação sobre o comportamento de um fenômeno. Por exemplo, semivariogramas

que aumentam lentamente a partir da origem e ou apresentam um comportamento quadrático perto da origem implicam processos mais suaves do que aqueles nas quais as semivariâncias apresentam comportamento linear perto da origem.

Chilès e Delfiner (1999) ressaltam que, além de examinar o comportamento do variograma a grandes distâncias, é igualmente interessante analisar o seu comportamento próximo da origem, porque ele está ligado à continuidade e à regularidade espacial da variável regionalizada. Segundo os autores,

- a) um comportamento parabólico caracteriza um alto grau de continuidade espacial. Se esse comportamento persistir ao longo de grandes distâncias, também pode estar associado à presença de um desvio forte;
- b) no comportamento linear, a variável regionalizada é contínua, pelo menos em pequenas distâncias, e é menos regular do que no caso anterior;
- c) o efeito pepita significa que a variável regionalizada, geralmente, não é contínua e é, portanto, muito irregular. Pode ser também devido à não amostragem em pequenas distâncias;
- d) no caso de efeito pepita puro ou ruído branco indica que não existe uma correlação entre os pontos.

Assim, deve-se conhecer bem a variável em estudo para que, a partir das características específicas dela, seja possível ajustar um modelo teórico.

2.3 Estimadores de semivariância

Nesta seção, são apresentados os oito estimadores de semivariâncias comparados nesta dissertação.

2.3.1 Estimador clássico de Matheron

O estimador de semivariograma clássico, desenvolvido por Matheron (1962), foi construído a partir das médias dos quadrados dos incrementos do processo, e fundamenta-se no método dos momentos. Segundo Cressie (1993) e Li e Lake (1994), este estimador é não viciado, mas é influenciado pela presença de *outliers*. O fato de as médias amostrais não serem robustas faz com que o estimador de Matheron também não seja robusto. O estimador é definido pela seguinte expressão:

$$2\hat{\gamma}_M(h) = \frac{1}{N(h)} \sum_{i=1}^{N(h)} (Z(x_i + h) - Z(x_i))^2 \quad (2.9)$$

em que

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

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

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

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

Segundo Schabenberger e Gotway (2004), uma das propriedades atraentes do estimador de Matheron é a simplicidade dos cálculos utilizados na estimação.

O estimador de Matheron é obtido por meio do método dos momentos. Mas, antes de mostrar o processo de obtenção deste estimador, são necessárias as seguintes definições (MOOD, 1950):

- 1 se U é uma variável aleatória, o r -ésimo momento populacional de U , usualmente denotado por μ'_r , é definido como

$$\mu'_r = E[U^r]$$

se a esperança existir;

2 sejam U_1, U_2, \dots, U_n uma amostra aleatória da densidade $f(\cdot)$. Então, o r -ésimo momento amostral centrado em 0, denotado por M'_r , é definido como

$$M'_r = \frac{1}{n} \sum_{i=1}^n U_i^r.$$

Pelo método dos momentos, os estimadores são obtidos igualando-se os momentos populacionais aos respectivos momentos amostrais.

$$\mu'_r = M'_r.$$

Para realizar a demonstração do estimador de Matheron, vamos considerar o conjunto $\{U_1, U_2, \dots, U_{N(h)}\}$, tal que $U_i = (Z(x_i + h) - Z(x_i))^2$. Vamos começar a demonstração encontrando os momentos populacional e amostral.

O primeiro momento populacional é dado por

$$\mu'_1 = E[U^1] = E[(Z(x_i + h) - Z(x_i))^2] = 2\gamma(h),$$

pois, por definição,

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

e

$$\text{Var}[Z(x_i + h) - Z(x_i)] = E[(Z(x_i + h) - Z(x_i))^2] - (E[Z(x_i + h) - Z(x_i)]).$$

Usando a propriedade de que a esperança de subtrações é igual à subtração das esperanças, tem-se que

$$\text{Var}[Z(x_i + h) - Z(x_i)] = E[(Z(x_i + h) - Z(x_i))^2] - (E[Z(x_i + h)] - E[Z(x_i)]).$$

Como assumimos a estacionariedade da média, então, $E[Z(x_i + h)] = E[Z(x_i)] = m$. Assim,

$$\begin{aligned} 2\gamma(h) &= \text{Var}[Z(x_i + h) - Z(x_i)] = E[(Z(x_i + h) - Z(x_i))^2] - (m - m) \\ 2\gamma(h) &= \text{Var}[Z(x_i + h) - Z(x_i)] = E[(Z(x_i + h) - Z(x_i))^2], \end{aligned}$$

portanto,

$$2\gamma(h) = E[(Z(x_i + h) - Z(x_i))^2].$$

O primeiro momento amostral é dado por

$$M'_1 = \frac{1}{n} \sum_{i=1}^n U_i^1 = \frac{1}{N(h)} \sum_{i=1}^{N(h)} (Z(x_i + h) - Z(x_i))^2.$$

Igualando-se o momento populacional ao momento amostral tem-se:

$$\mu'_1 = M'_1$$

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

que é o estimador proposto por Matheron.

2.3.2 Estimador robusto de Cressie e Hawkins

Devido à sensibilidade do estimador de Matheron frente aos *outliers*, Cressie e Hawkins (1980) quiseram propor um estimador mais robusto. Para tanto, eles removeram o quadrado presente nos incrementos do estimador de Matheron. Basicamente, a ideia era a de que a raiz quadrada diminuísse a contaminação por valores discrepantes (CRESSIE, 1993). Este estimador é conhecido como estimador robusto e é definido pela seguinte expressão, na qual o denominador corrige viés sob gaussianidade.

$$2\hat{\gamma}_{CH}(h) = \left[\frac{1}{N(h)} \sum_{i=1}^{N(h)} |(Z(x_i + h) - Z(x_i))|^{\frac{1}{2}} \right]^4 / \left(0.457 + \frac{0,494}{N(h)} \right) \quad (2.10)$$

em que

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

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

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

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

No entanto, apesar de o estimador de Cressie e Hawkins atenuar o efeito de dados contaminados, Genton (1998) afirma que ele ainda é sensível à presença de *outliers* nos dados.

A construção do estimador robusto de Cressie e Hawkins parte da ideia de que se $Z(x)$ é um processo gaussiano, então,

$$\frac{Z(x_i+h)-Z(x_i)}{\sqrt{2\gamma(h)}} \sim N(0, 1).$$

Elevando ao quadrado, tem-se que

$$\frac{[Z(x_i+h)-Z(x_i)]^2}{2\gamma(h)} \sim \chi_1^2.$$

Os autores verificaram que a raiz quarta de $(Z(x_i+h) - Z(x_i))^2$ tem distribuição aproximadamente gaussiana, com média

$$E[|Z(x_i+h) - Z(x_i)|^{\frac{1}{2}}] \approx \frac{1}{2}\pi^{-\frac{1}{2}}\Gamma(0,75)\gamma(h)^{\frac{1}{4}}.$$

Além disso, o valor esperado de $\frac{1}{N(h)} \sum_{i=1}^{N(h)} |Z(x_i+h) - Z(x_i)|^{\frac{1}{2}}$ elevado a quarta potência é

$$E\left\{\left(\frac{1}{N(h)} \sum_{i=1}^{N(h)} |Z(x_i+h) - Z(x_i)|^{\frac{1}{2}}\right)^4\right\} \approx 2\hat{\gamma}(h)\left(0,457 + \frac{0,494}{N(h)} + \frac{0,045}{(N(h))^2}\right).$$

Segundo os autores, o termo $\frac{0,045}{(N(h))^2}$ não afeta significativamente a estimativa e, assim, ele foi omitindo da expressão. Então, isolando-se $2\gamma(h)$, obteve-se o estimador de Cressie Hawkins:

$$2\hat{\gamma}(h) = \left[\frac{1}{N(h)} \sum_{i=1}^{N(h)} |(Z(x_i+h) - Z(x_i))^{\frac{1}{2}}| \right]^4 / \left(0,457 + \frac{0,494}{N(h)}\right).$$

Segundo Schabenberger e Gotway (2004), uma vez que se calcula primeiramente a raiz quadrada das diferenças, o estimador é menos afetado por valores discrepantes do que a média das diferenças de quadrados do estimador de Matheron. O autor explica que o estimador robusto não é imparcial, mas o termo no denominador garante uma maior estabilidade.

2.3.3 Estimador das medianas de Cressie

Cressie propôs um segundo modelo. Substituindo a média pela mediana no estimador robusto de Cressie e Hawkins, obtem-se o estimador das medianas,

$$2\hat{\gamma}_{Md}(h) = \frac{\text{med} \left[|(Z(x_i + h) - Z(x_i))|^{\frac{1}{2}} \right]^4}{0.457} \quad (2.11)$$

em que

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

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

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

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

$\text{med}\{\bullet\}$ denota a mediana da sequência $\{\bullet\}$.

Emerson e Hoaglin (1983) comentam que a média amostral não oferece proteção contra erros grosseiros. Devido ao fato de dar o peso de $\frac{1}{n}$ a cada uma das observações na amostra, a presença de um único valor discrepante distorce seriamente o valor da média. Segundo os autores, a mediana, por outro lado, é muito mais tolerante a erros grosseiros. Quase a metade de um conjunto de números pode ser de valores discrepantes, sem alterar muito o valor da mediana.

Este estimador é obtido seguindo a mesma ideia do estimador robusto de Cressie e Hawkins: se $Z(x)$ for gaussiano, então,

$$\frac{[Z(x_i+h) - Z(x_i)]^2}{2\gamma(h)} \sim \chi_1^2,$$

de onde temos que

$$[Z(x_i + h) - Z(x_i)]^2 \sim \chi_1^2 2\gamma(h).$$

Seja Q_θ o quantil empírico de ordem θ e $F_{\chi_1^2}$ o quantil de ordem θ da distribuição χ_1^2 , então,

$$2\hat{\gamma}(h) = \frac{\{Q_\theta(Z(x_i+h)-Z(x_i))^2:(x_i, x_i+h) \in N(h)\}}{F_{\chi_1^2}},$$

quando $\theta = \frac{1}{2}$, tem-se que

$$2\hat{\gamma}_{Md}(h) = \frac{\text{med}[(Z(x_i+h)-Z(x_i))^{\frac{1}{2}}]^4}{0,457}.$$

2.3.4 Estimador das diferenças de Haslett

Este estimador foi proposto por Haslett (1997) para utilização em séries temporais, com o objetivo principal de reconhecer processos ARMA.

$$2\hat{\gamma}_H(h) = \frac{1}{N(h) - 1} \sum_{i=1}^{N(h)} (d_{hi} - \bar{d}_h)^2 \quad (2.12)$$

em que

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

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

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

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

$d_{hi} = (Z(x_i + h) - Z(x_i))$.

O estimador das diferenças de Haslett surgiu baseado na função de variância

$$s^2 = \frac{\sum(t_i - \bar{t})^2}{n-1}.$$

Considerando-se $d_{hi} = (Z(x_i + h) - Z(x_i))$ e sendo \bar{d}_h a média de todos os pares com distância h , então, a variância dos valores d_{hi} é

$$2\hat{\gamma}(h) = \frac{1}{N(h) - 1} \sum_{i=1}^{N(h)} (d_{hi} - \bar{d}_h)^2$$

2.3.5 Estimador altamente robusto de Genton

Genton (1998) propôs um estimador denominado altamente robusto que, segundo o autor, é resistente a *outliers*, utilizando a teoria dos estimadores de escala M para obter robustez. A ideia fundamental para a construção deste estimador foi a de estimar diretamente a variância dos incrementos, utilizando um estimador de escala robusto e com boa eficiência em modelos normais. Assim, foi utilizado o estimador de escala Q_n , proposto por Rousseeuw e Croux (1993 apud GENTON, 1998). O estimador de Genton é definido como

$$2\hat{\gamma}_G(h) = (Q_{N(h)})^2 \quad (2.13)$$

em que

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

$Q_{N(h)} = 2,2191 \{(|V_i(h) - V_j(h)|; i < j\}_{(k)}$;

$V(h) = Z(x + h) - Z(x)$;

2,2191 é a consistência da distribuição gaussiana;

$k = \left(\frac{[\frac{N(h)}{2}] + 1}{2} \right)$;

$[\frac{N(h)}{2}]$ denota a parte inteira de $\frac{N(h)}{2}$.

O estimador de escala Q_n é definido por

$$Q_n(W_1, W_2, \dots, W_n) = c\{|W_i - W_j| : 1 \leq i < j \leq n\}_{(k)}$$

sendo $k = \left(\frac{[\frac{n}{2}] + 1}{2}\right)$.

Se considerarmos $W_i = Z(x_i + h) - Z(x_i)$ e $W_j = Z(x_j + h) - Z(x_j)$, temos o estimador de Genton

$$Q_n = c\{|Z(x_i + h) - Z(x_i) - (Z(x_j + h) - Z(x_j))| : i < j\}_{(k)}$$

em que $k = \left(\frac{[\frac{N(h)}{2}] + 1}{2}\right)$ e $c=2,2191$.

2.3.6 Estimador Pairwise

O estimador Pairwise é um estimador de variância relativo. Segundo Isaaks e Srivastava (1989), ele contribui para produzir uma melhor visualização da continuidade espacial. A diferença entre ele e o da equação 2.9 é o denominador, que serve para reduzir a influência de valores muito grandes. Este ajustamento, contudo, é feito separadamente para cada par de valores de amostra, utilizando-se a média dos dois valores. O estimador é definido por

$$\hat{\gamma}_{PR}(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} \left(\frac{Z(x_i + h) - Z(x_i)}{\frac{Z(x_i + h) + Z(x_i)}{2}} \right)^2 \quad (2.14)$$

em que

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

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

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

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

Li e Lake (1994) afirmam que esse estimador se caracteriza por diminuir o efeito de dados discrepantes, porém, apresenta a desvantagem de a soma dos valores da variável Z nos pontos de coordenadas x e $(x + h)$ não poder ser nula. Além disso, quando a média se aproxima de zero,

podem ocorrer semivariogramas com “saltos” nas semivariâncias. Segundo os autores, o estimador de Pairwise obteve sucesso prático, apesar de ser incerta sua teoria.

2.3.7 Estimadores New-1 e New-2

A falta de precisão das estimativas das semivariâncias está diretamente ligada ao número cada vez menor de pares de observações, com o aumento gradativo das distâncias h . Sustentados por este fato, Li e Lake (1994) propuseram dois estimadores de semivariância, o estimador New-1, que é definido pela expressão

$$2\hat{\gamma}_{N1}(h) = \frac{2}{n} \sum_{i=1}^n \left\{ \frac{1}{2m} \sum_{j \in D_{i,h}} (Z(x_i) - Z(x_j))^2 \right\} \quad (2.15)$$

em que

$\hat{\gamma}_{N1}(h)$ é o valor da semivariância estimada pelo New-1;

$Z(x_i)$ e $Z(x_j)$ são os valores da variável nos respectivos pontos i e j ;

n é o número total de dados;

$D_{i,h}$ é o índice de um conjunto de valores de dados em uma janela móvel $\Delta i, h$ (de tamanho h centrada no ponto bloco i), excluindo o ponto x_i ;

m é o número de dados em $D_{i,h}$,

e o estimador New-2, que é definido por

$$2\hat{\gamma}_{N2}(h) = 2(\hat{\gamma}_{N1}(h) + \frac{h}{d} \hat{\gamma}'_{N1}(h)); \quad (2.16)$$

em que

$\hat{\gamma}_{N2}(h)$ é o valor da semivariância estimada pelo New-2;

$\hat{\gamma}'_{N1}(h)$ é a derivada de $\hat{\gamma}_{N1}(h)$ em relação a h , calculada pelo Método da Diferença Central;

h vetor de distâncias;

d é a dimensão no espaço euclidiano.

O grande diferencial dos estimadores New-1 e New-2 se deve ao fato de que eles usam todos os dados na obtenção de cada estimativa, independentemente do h . Segundo os autores, os dois estimadores propostos são imparciais, robustos e resistentes à contaminação.

Na Figura 2 observa-se um exemplo, de duas dimensões, da janela móvel. A ideia é movimentar a janela. Os estimadores New utilizam todos os pontos na janela, realizando as diferenças entre o ponto x_i com todos os pontos x_j da janela, exceto com o próprio ponto x_i .

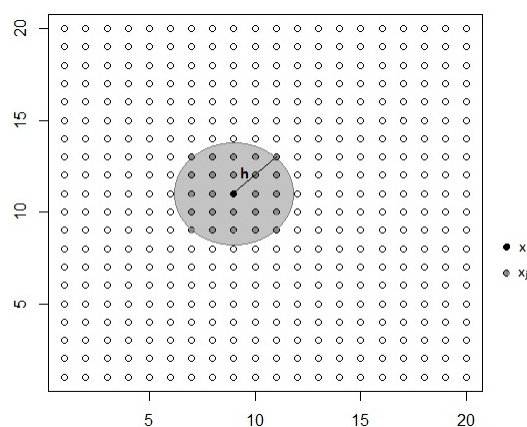


Figura 2 Janela móvel bidimensional para um campo isotrópico

Para construir os novos estimadores, Li e Lake (1994) definiram uma nova semivariância, tomando um $(d - 1)$ -ésimo momento da definição semivariância:

$$\gamma_N(h) = \frac{1}{\int_0^h \xi^{d-1} d\xi} \int_0^h \xi^{d-1} \gamma(\xi) d\xi$$

Segundo os autores, esta definição pode ser interpretada como uma média ponderada de $\gamma(\xi)$ sobre $(0, h)$, com uma função de ponderação $\xi^{d-1} / \int_0^h \xi^{d-1} d\xi$, em que d é a dimensão no espaço euclidiano.

Utilizando a definição da função de distribuição de probabilidade condicional e a definição de esperança, os autores chegaram à seguinte relação:

$$E[u(S, T)] = E\{E[u(S, T)|_S]\}$$

onde $E[u(S, T)]$ é a esperança conjunta, $E[u(S, T)|_S]$ é a esperança condicional e $E\{\}$ é a esperança marginal. Calculando:

$$E\left\{\frac{1}{2 \int_0^h \xi^{d-1} d\xi} \int_0^h \xi^{d-1} E[(Z(x) - Z(x + \xi))^2 |_{Z(x)}] d\xi\right\}$$

chegou-se ao estimador New-1.

O segundo estimador é proveniente da seguinte relação entre a $\gamma(h)$ e $\gamma_N(h)$:

$$\gamma(h) = \gamma_N(h) + \frac{h}{d} \hat{\gamma}'_N(h)$$

em que $\hat{\gamma}'_N(h)$ é a derivada de $\gamma_N(h)$ em relação a h

Por meio da combinação do New-1 e da relação anterior, realizando as devidas substituições, tem-se o estimador New-2.

3 METODOLOGIA

A realização deste trabalho constituiu-se de quatro etapas. Na primeira, foi realizado um levantamento bibliográfico, no intuito de identificar diferentes estimadores de semivariância disponíveis na literatura. Como a comparação dos estimadores requer um conhecimento detalhado sobre cada modelo, então, decorreu um período de estudo das propriedades e características de cada um. Neste estudo foram realizadas as demonstrações referentes à forma de obtenção de cada estimador, as quais estão disponíveis em subseções da subseção 2.3. Nesta etapa foram escolhidos os seguintes estimadores de semivariância:

- 1 estimador clássico de Matheron;
- 2 estimador robusto de Cressie e Hawkins;
- 3 estimador das medianas de Cressie;
- 4 estimador das diferenças de Haslett;
- 5 estimador altamente robusto de Genton;
- 6 estimador Pairwise;
- 7 estimador New-1 e
- 8 estimador New-2.

Para fazer as comparações entre os estimadores, foram realizadas simulações de diferentes tipos de bancos de dados, com o objetivo de verificar o comportamento de cada estimador diante de determinada característica. Assim, a segunda etapa constituiu-se de trabalhos de simulação. Neste período, foram definidos quais tipos de banco de dados seriam gerados e quais características seriam consideradas ao realizar as simulações. Todas as simulações foram feitas no *software* R (R DEVELOPMENT CORE

TEAM, 2012), utilizando-se o pacote *Randomfields* (SCHLATHER, 2006). Os detalhes das simulações encontram-se na subseção 3.1.

Na terceira etapa, já com os resultados das simulações, foram realizadas as comparações dos estimadores de semivariância. Após as comparações, foram elaboradas tabelas, indicando quais estimadores tiveram melhor desempenho diante de cada tipo de banco de dados abordado no estudo. Também foram construídos gráficos, no intuito de facilitar a análise dos resultados.

Ao final, como última etapa, foi realizado um estudo, a título de exemplo, com o estudo de um banco de dados reais. Foi realizada uma análise geoestatística básica dos dados. Para a construção do semivariograma, foi utilizado o estimador de semivariância sugerido nos resultados deste trabalho, considerando-se os resultados obtidos nas comparações.

3.1 Simulação

Todas as simulações foram realizadas por meio do pacote *RandomFields* (SCHLATHER, 2006) e foram baseadas no modelo teórico esférico. Por meio da função *GaussRF* foram simulados processos gaussianos estacionários.

```
simul.Gauss <-c(GaussRF(x=x,y=y,model=model,grid=T,
                      gridtriple=FALSE,method=NULL,
                      param=param))
```

Nas simulações, foram abordados casos de dados com tendência e, também, situações sem a existência da mesma. Assim, é importante destacar a alteração que ocorre em um dos comandos da simulação. No caso de não se utilizar a tendência, o parâmetro usado é $param = c(mean, variance, nugget, scale, \dots)$. Contudo, alternativamente, quando se adota o uso de tendência, usa-se $param = c(variance, nugget, scale, \dots)$.

3.1.1 Tamanho da amostra

Os bancos de dados abordados, foram simulados em formato de malhas regulares quadradas, com três diferentes tipos de tamanho, sendo:

- a) gride 5x5, 25 dados;
- b) gride 10x10, 100 dados;
- c) gride 20x20, 400 dados.

Estes tamanhos foram escolhidos, pois é comum, na literatura, encontrar indicações de que se deve trabalhar com mais de 30 observações. Assim, aqui estariam sendo abordados um banco de dados considerado pequeno (25 dados), um de porte médio (100 dados) e um relativamente grande (400 dados).

3.1.2 Quantidade de *outliers*

As contaminações dos dados, com valores discrepantes, foram realizadas em quatro situações distintas, referentes à porcentagem de *outliers* presente no banco de dados, sendo:

- a) sem *outliers*;
- b) 1% de *outliers*;
- c) 5% de *outliers*;
- d) 10% de *outliers*.

É importante observar que no banco de dados com tamanho 25, as porcentagens de dados discrepantes não correspondem a valores inteiros. Assim, foi considerado o menor inteiro superior ao valor calculado.

A princípio, os dados foram simulados, sendo posteriormente contaminados com valores discrepantes, por meio da função do pacote

simframe. Nesta função, a porcentagem pré-definida de dados foi contaminada de forma aleatória, sendo estes valores substituídos por valores discrepantes, maiores que o valor máximo e/ou menores que o valor mínimo, gerados por uma distribuição uniforme.

O comando utilizado na contaminação foi o seguinte:

```
cc <- DARContControl(target = "simul.Gauss", epsilon=epsilon,
                     un = function(x) x - x runif(100*epsilon,4,8))
simul.Gauss<-contaminate(simul.Gauss, cc)
```

em que “simul.Gauss” são os dados nos quais “epsilon” por cento estão sendo contaminados com a função $function(x)$, que, no caso, é a “runif”, que gera números provenientes de uma distribuição uniforme.

No *script* há também um segundo comando deste, pois a contaminação é realizada tanto com números menores que o valor mínimo, quanto com números maiores que o valor máximo. Para assegurar a quantidade de *outliers* foi utilizado o box-plot como forma de detecção dos pontos discrepantes. Foi construído um comando de repetição para assegurar que a simulação só seria terminada quando fosse simulado um banco de dados com a quantidade de *outliers* correta.

Como o gráfico boxplot só necessita de medidas de posição, como a mediana e os quartis, associadas ao conjunto de dados e como a mediana e os quartis são resistentes à existência de alguns valores discrepantes, o boxplot também é, em geral, resistente à influência desses valores discrepantes. Mais especificamente, aproximadamente 25% dos dados podem ser substituídos por *outliers*, sem gerar grande perturbação na mediana, nos quartis e, conseqüentemente, no boxplot (EMERSON; STRENIO, 1983).

3.1.3 Tendência

A tendência foi outra característica abordada nas simulações. Foram consideradas as seguintes situações:

- a) dados sem tendência;

- b) dados com tendência linear;
- c) dados com tendência quadrática.

Um dos problemas na determinação do semivariograma é a presença de tendência nos dados avaliados, ou seja, a não estacionariedade da variável, dificultando o ajuste de modelos inadequados ao comportamento espacial da variável. Assim, esta característica deve ser estudada com atenção, ao se realizar a análise de um banco de dados.

Andriotti (2003) argumenta que, quando a variável apresentar tendência, esta deve ser removida antes do ajuste do semivariograma passando-se a utilizar os resíduos para a detecção da estrutura de variabilidade. Contudo, o objetivo, neste trabalho, é averiguar o comportamento dos estimadores diante da tendência, no intuito de verificar uma possível robustez, de algum deles, diante desta característica. Por isso, a tendência não foi retirada ao se realizar as análises.

3.1.4 Comparação

Como foram utilizadas três situações de tendência, quatro situações de *outliers* e três situações de tamanhos da amostra, foram obtidos 36 tipos de banco de dados provenientes das combinações das três características supracitadas. Para cada tipo de banco de dados foram realizadas 500 simulações.

As comparações foram realizadas utilizando-se as seguintes expressões, referentes a erros médios:

1 erro médio absoluto:

$$EMA = \frac{1}{N} \sum_{i=1}^N 2|\gamma(h) - \hat{\gamma}(h)|$$

2 erro médio quadrático:

$$EMQ = \frac{1}{N} \sum_{i=1}^N 2(\gamma(h) - \hat{\gamma}(h))^2$$

Após a simulação de cada banco de dados, foram obtidas as estimativas de cada um dos oito estimadores de semivariância. Como o estimador New-2 não calcula a semivariância referente à última distância, não foi calculada a estimativa de semivariância para a maior distância nos demais estimadores. Foram consideradas duas situações. Na primeira, foram calculados os erro médio absoluto e o erro quadrático médio, considerando-se as $N(h) - 1$ estimativas e, em um segundo momento, ambos foram calculados, contudo utilizando-se um *cutoff* de 50%.

O interessante, ao se observar estas duas situações, é que, nas maiores distâncias, o número de pares é pequeno, o que, comumente, aumenta o erro das estimativas. Assim, um estimador pode ser bom, contudo, seu erro em distâncias maiores pode mascarar um bom comportamento em distâncias menores.

3.2 Softwares

O trabalho de simulação e análise de dados foi realizado utilizando-se funções disponíveis em várias bibliotecas do *software* R (R DEVELOPMENT CORE TEAM, 2012). Foram utilizados os seguintes pacotes:

package RandomFields: é um pacote desenvolvido por Schlather (2006), que realiza simulações e análises de campos aleatórios;

package geoR: desenvolvido por Ribeiro Júnior e Diggle (2001), é denominado *Analysis of geostatistical data* e por meio dele é possível fazer um estudo geoestatístico básico dos dados;

package simFrame: desenvolvido por Alfons, Templ e Filzmoser (2010), foi utilizado para gerar as contaminações com *outliers* nos bancos de dados.

O *software* R é um programa de código aberto, ou seja, que pode ser modificado conforme o interesse do usuário, permitindo que se criem rotinas de cálculos e análises estatísticas. Assim, além das funções disponíveis nas bibliotecas, também foram construídas algumas funções, a fim de automatizar processos repetitivos e realizar as análises necessárias.

É importante destacar que, para realizar as comparações, foram construídas as funções referentes aos oito estimadores de semivariância abordados. Em anexo, encontra-se um *script* contendo estas funções.

4 RESULTADOS E DISCUSSÕES

Foram gerados 36 experimentos, provenientes das combinações entre os três tamanhos de amostras abordados, sem tendência e com tendência linear e quadrática e também das quatro porcentagens de *outliers*. Os resultados foram sumarizados em 72 tabelas, as quais encontram-se em anexo.

Nas subseções seguintes, estes resultados são descritos e discutidos. Também são apresentados gráficos, referentes ao Erro Médio Quadrático de cada estimador de semivariância, de acordo com os tamanhos das amostras, para melhor visualização dos erros.

4.1 Resultados das comparações em *grides* 5x5

As análises referentes aos bancos de dados, de tamanho 25, foram sumarizadas em três gráficos, um para cada tipo de tendência. No Gráfico 10 encontra-se a análise dos dados sem tendência. Nele, é possível perceber a influência da utilização do *cutoff*, o qual diminui, consideravelmente, o erro da maioria dos estimadores. Esta diferença é muito realçada, principalmente, nos estimadores altamente robusto de Genton e das medianas de Cressie. O estimador New-1 é o que apresenta sofrer menor influência do *cutoff*.

Também, pode-se perceber que o estimador New-1, na maior parte do tempo, obteve os menores erros. O estimador de Cressie e Hawkins demonstrou bom desempenho, de acordo com o aumento da quantidade de *outliers*, em casos com *cutoff*.

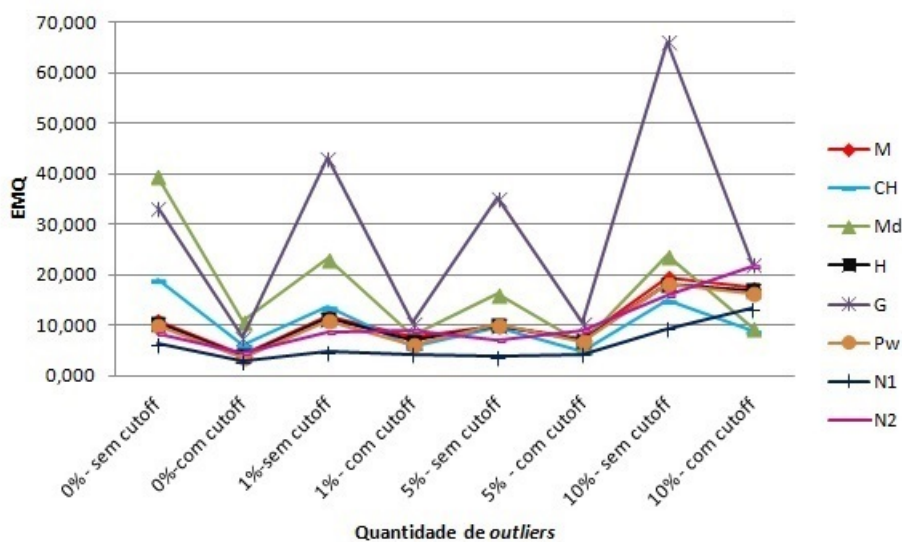


Gráfico 10 EMQ *versus* quantidade de *outliers* dos dados com $n=25$ e sem tendência

Nos dados com tendência linear, apresentados no Gráfico 11, pode-se perceber que o estimador New-1 destacou-se nos casos de 0% e 1% de *outliers*. Nas demais contaminações, os estimadores das medianas e de Cressie e Hawkins, com destaque para o segundo, obtiveram melhor desempenho.

Também foi possível verificar que a utilização do *cutoff* reduz os erros na maioria dos estimadores, sendo o estimador New-1 o menos afetado pela utilização desta técnica. A utilização do *cutoff* é muito influente nos resultados de estimador de Genton.

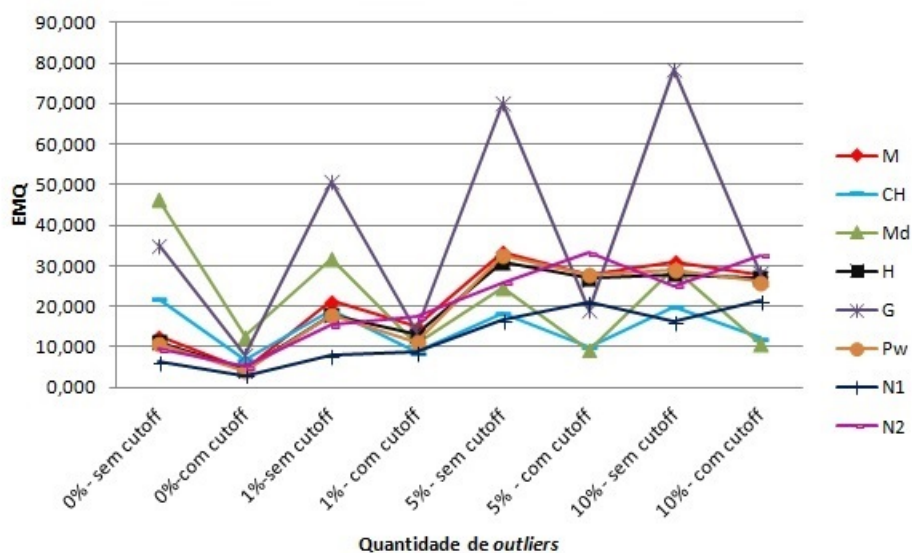


Gráfico 11 EMQ versus quantidade de outliers dos dados com $n=25$ e com tendência linear

No Gráfico 12, tem-se os dados com tendência quadrática. É possível perceber que o estimador Pairwise é extremamente afetado pela tendência quadrática, principalmente, quando esta vem acompanhada de altas contaminações por outliers.

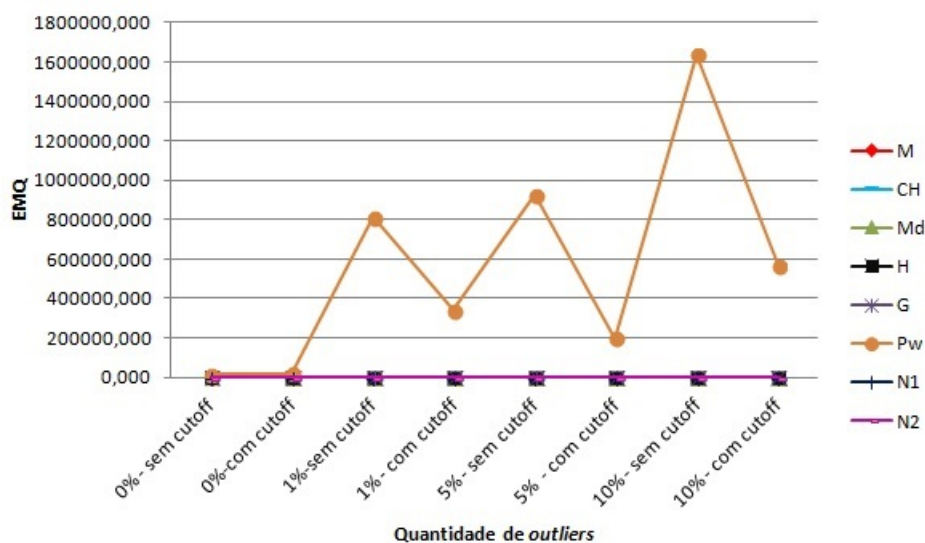


Gráfico 12 EMQ *versus* quantidade de *outliers* dos dados com $n=25$ e com tendência quadrática

Para melhor analisarmos os resultados referentes à tendência quadrática, foi realizado um recorte no Gráfico 12, que pode ser observado no n25qa. Neste recorte, podemos perceber que, o estimador New-1 apresentou o melhor desempenho. Nos casos com baixa contaminação de *outliers*, os estimadores de Haslett e New-2 demonstraram bons resultados. Contudo, nas demais contaminações, o estimador de Cressie e Hawkins se destacou.

Também é possível perceber que, diante da tendência quadrática, foram obtidos erros maiores do que nas situações sem tendência ou com tendência linear.

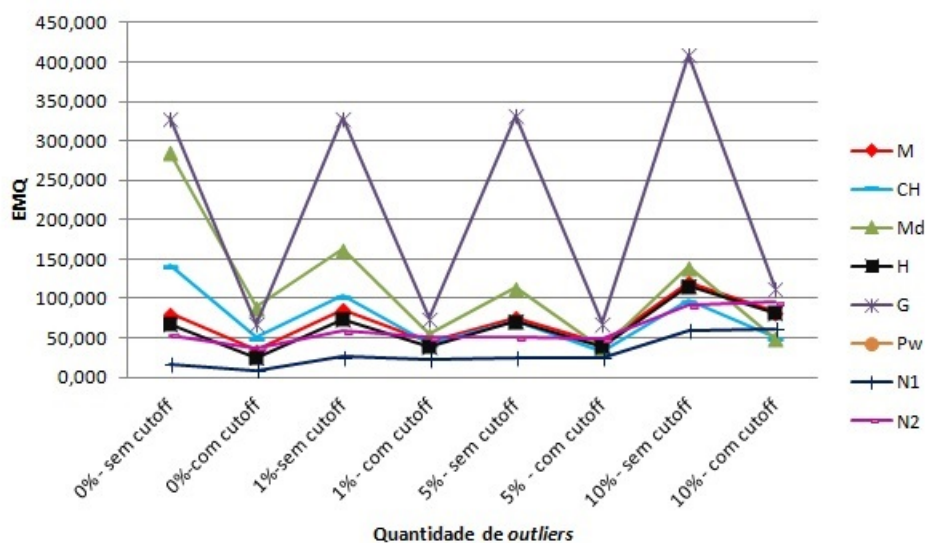


Gráfico 13 Recorte do gráfico EMQ *versus* quantidade de *outliers* dos dados com $n=25$ e com tendência quadrática

De forma geral, em situações sem *outliers*, o estimador New-1 apresentou os melhores resultados, na maioria dos casos. O estimador das Medianas e o estimador de Cressie e Hawkins apresentaram bons desempenhos em casos de contaminação com 5% e 10% de *outliers*, chegando a obter os melhores resultados em algumas situações.

Também, pode-se observar que, o estimador de Genton não obteve bons resultados em amostras pequenas. O estimador Pairwise foi influenciado, fortemente, pela tendência quadrática, apresentando grandes oscilações em seus resultados, principalmente quando havia altas contaminações com *outliers*.

4.2 Resultados das comparações em *grides* 10x10

Os resultados, referentes às análises dos bancos de dados de tamanho 100, também foram sumarizados em três gráficos, sendo um para cada

tipo de tendência. No Gráfico 14 tem-se os resultados referentes as bancos de dados sem tendência. Pode-se perceber que as contaminações de 10% de *outliers* causaram forte influência nos resultados, ocasionando um aumento nos erros. A utilização de *cutoff* gerou uma redução nos erros das estimativas.

O estimador New-1 apresentou o melhor desempenho em nas quantidades de 0%, 1% e 5% de *outliers*, sendo seguido de perto pelo estimador New-2. Os estimador de Cressie e Hawkins demonstrou bons resultados na contaminação de 10% de *outliers*. O estimador das medianas também apresentou bons resultados nesta configuração dos dados.

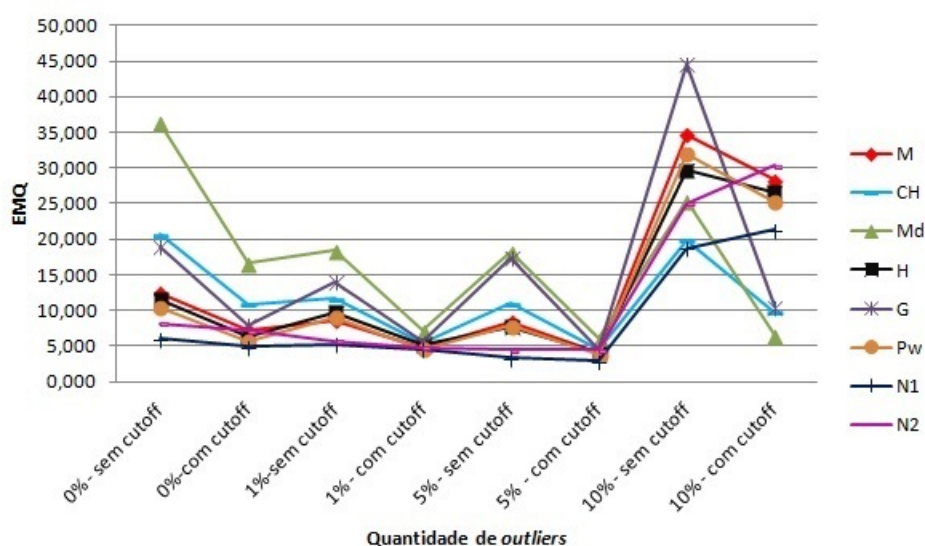


Gráfico 14 EMQ *versus* quantidade de *outliers* dos dados com $n=100$ e sem tendência

Nos casos de tendência linear, apresentados no Gráfico 15, o estimador New-1 apresentou os melhores resultados em quase todas as situações. Em baixas contaminações por *outliers*, o estimador New-2 também obteve erros baixos. Contudo, em grandes quantidades de *outliers*, os estimadores das medianas e de Cressie e Hawkins foram os que

apresentaram melhor desempenho. A utilização de *cutoff* reduziu os erros na grande maioria dos casos.

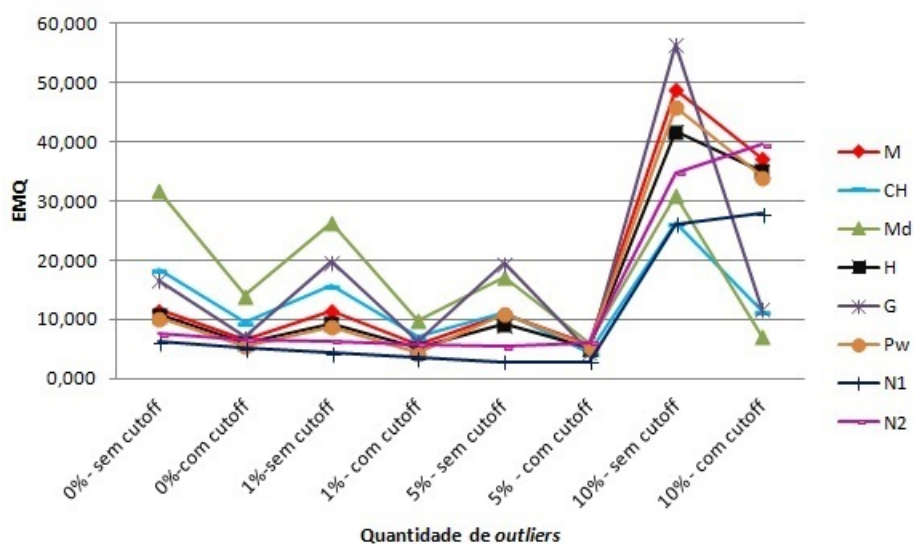


Gráfico 15 EMQ *versus* quantidade de *outliers* dos dados com $n=100$ e com tendência linear

No Gráfico 16, pode-se perceber que a tendência quadrática e sua combinação com os *outliers* influenciaram fortemente o estimador Pairwise, que sofreu grandes oscilações, com destaque nos casos de 10% de *outliers*.

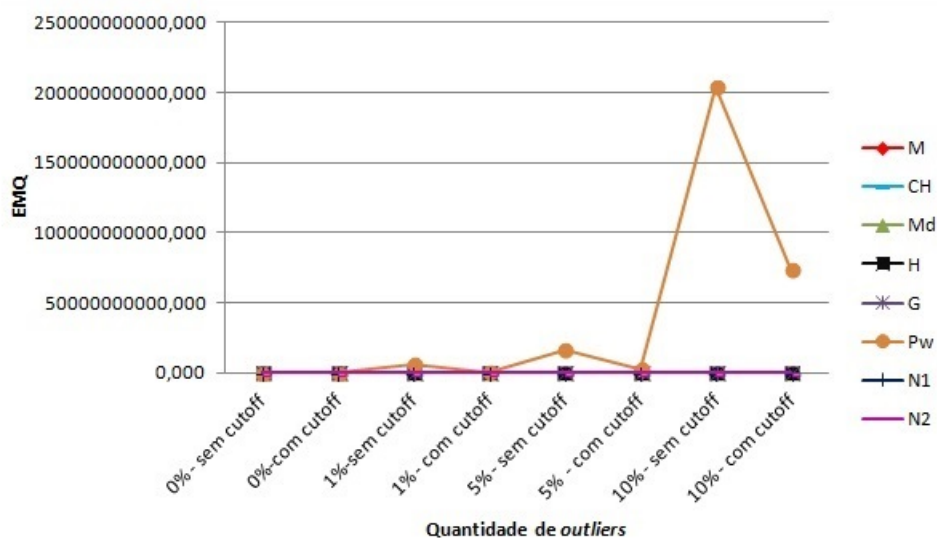


Gráfico 16 EMQ *versus* quantidade de *outliers* dos dados com $n=100$ e com tendência quadrática

Assim, para melhor analisar os resultados, foi realizado um recorte no gráfico anterior, que pode ser observado no Gráfico 17. Pode-se perceber que os estimadores New-1 e de Haslett revezaram os melhores resultados, em baixas contaminações. O estimador de Genton também apresentou bom desempenho em todas as situações, obtendo os menores erros nos casos de alta contaminação por *outliers*. A utilização de *cutoff* apresentou bons resultados.

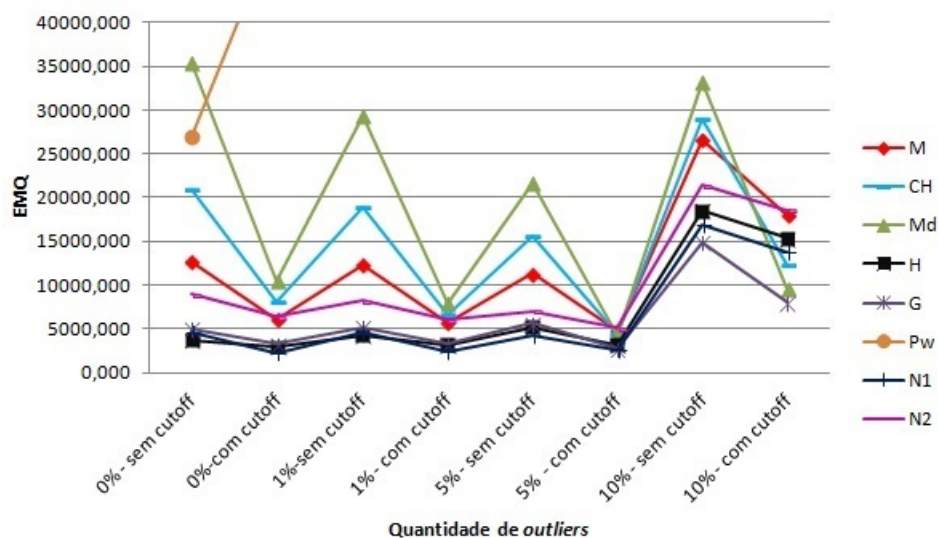


Gráfico 17 Recorte do gráfico EMQ versus quantidade de outliers dos dados com n=100 e com tendência quadrática

De forma geral, o estimador New-1 obteve os melhores resultados nas situações sem outliers, com 1% e 5% de contaminação, nos casos sem tendência e com tendência linear. Sendo que, os estimadores das medianas e de Cressie e Hawkins alternaram os melhores resultados nos casos de 10% de contaminação. Contudo, diante da tendência quadrática os estimadores New-1 e de Halett demonstraram bons resultados para situações de 0%, 1% e 5% de outliers, sendo seguidos de perto pelos bons resultados do estimador de Genton. Nas contaminações de 10%, o estimador de Genton obteve o melhor desempenho.

Também foi possível observar que a existência de outliers e, principalmente, de tendência, ocasionaram aumentos nos erros. O estimador Pairwise sofre grandes oscilações nos casos de tendência quadrática. O estimador de Genton apresentou melhores resultados nos bancos de dados de tamanho 100 do que nos de tamanho 25.

4.3 Resultados das comparações em *grides* 20x20

Nesta subseção, são apresentados os gráficos referentes às análises dos resultados dos bancos de dados de tamanho 400. Neste tamanho de amostra, a contaminação com *outliers* causou grandes oscilações no estimador Pairwise. Assim, por simplicidade, serão apresentados apenas os recortes dos gráficos, a fim de melhor analisar os resultados.

No Gráfico 18, pode-se observar que o estimador New-1, seguido do estimador New-2, apresentou os melhores resultados nos casos de 0% e 1% de *outliers*. O estimador das medianas, seguido do estimador de Genton, demonstrou melhor desempenho nas demais contaminações. A utilização de *cutoff* contribuiu para a redução dos erros.

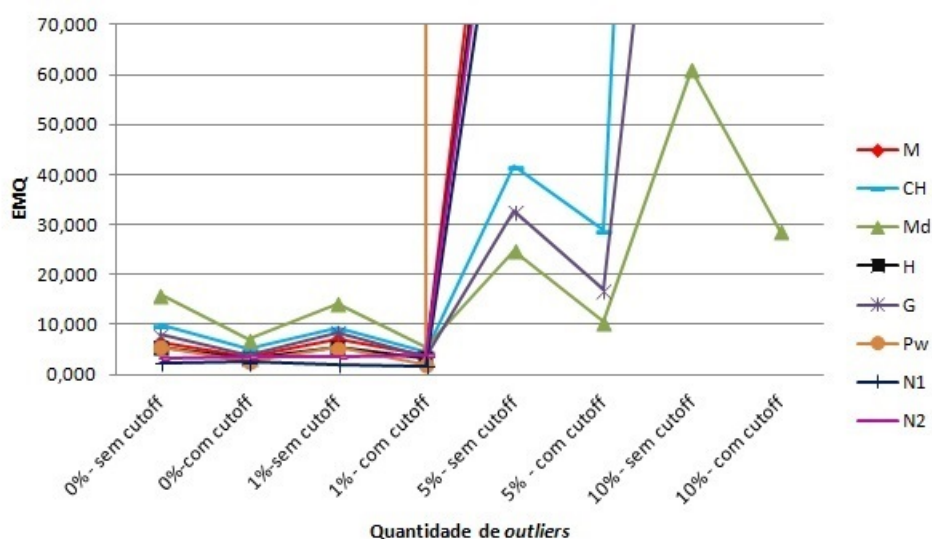


Gráfico 18 EMQ versus quantidade de *outliers* dos dados com $n=400$ e sem tendência

Nos casos de tendência linear, apresentados no Gráfico 19, as oscilações foram ainda mais fortes. Nas situações sem *outliers* e com 1% de contaminação, o estimador New-1 apresentou o melhor desempenho.

Sendo o segundo melhor, o estimador New-2. Já nos casos de 5% e 10% de *outliers*, o estimador que se destacou foi o das medianas, sendo seguido pelo estimador de Genton. A utilização do *cutoff* reduziu os erros, enquanto que a existência da tendência linear aumentou seus valores.

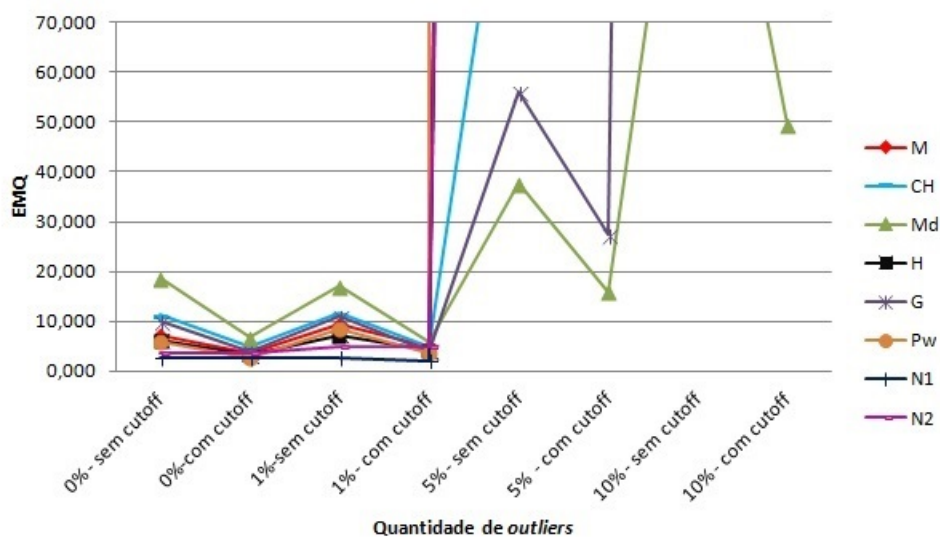


Gráfico 19 EMQ *versus* quantidade de *outliers* dos dados com $n=400$ e com tendência linear

No gráfico Gráfico 20, pode-se observar os resultados referentes aos dados com tendência quadrática. Nesta configuração os erros foram bem maiores. Nos casos sem *outliers* o estimador Pairwise apresentou melhor desempenho. Com 1%, 5% e 10% de contaminação, o estimador de Genton foi o melhor, sendo seguido pelo estimador New-1. A utilização do *cutoff* diminuiu os erros.

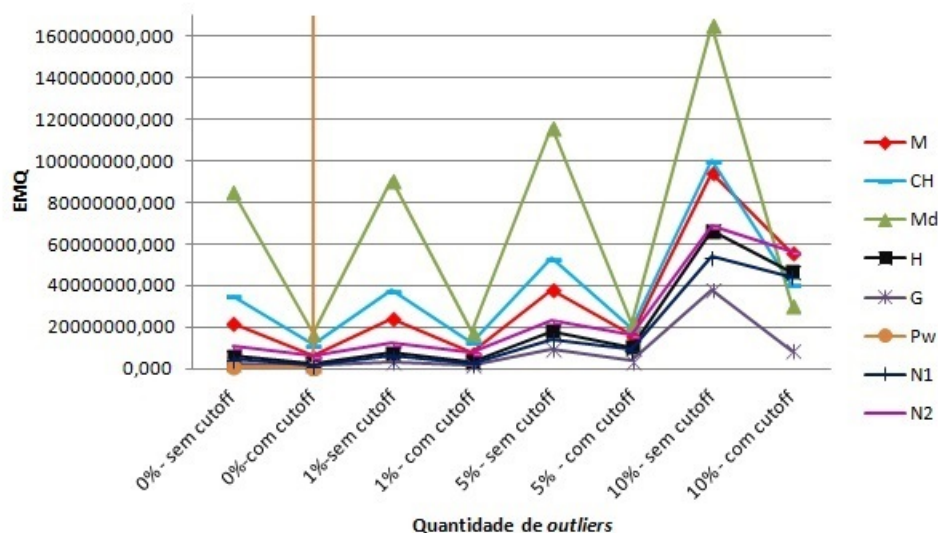


Gráfico 20 EMQ versus quantidade de outliers dos dados com $n=400$ e com tendência quadrática

De forma geral, o estimador New-1 apresentou melhor desempenho nos casos sem tendência e com tendência linear, ambos com 0% e 1% de outliers. Já nos casos de 5% e 10% de contaminação, o estimador das medianas foi o melhor, seguido pelo de Genton. O estimador altamente robusto de Genton teve o melhor desempenho na tendência quadrática, exceto no caso sem outliers, em que o Pairwise foi o melhor.

O *cutoff* apresentou bons resultados, reduzindo os erros. Os outliers e a tendência contribuíram no aumento dos erros e oscilações das estimativas. O estimador Pairwise sofreu grandes alterações diante de dados com contaminações elevadas. Enquanto que, o estimador de Genton demonstrou melhores resultados em bancos de dados maiores.

4.4 Discussão

Na literatura estudada, observou-se que a característica mais mencionada como influente na escolha do estimador de semivariância é o *outlier*. Genton (1998) comenta que em estudos de ciências aplicadas mostra-se que os dados medidos contêm, em média, entre 10% a 15% de valores periféricos, devido a erros grosseiros, erros de medição, gravação defeituosa, etc.

O autor explica que, apesar de análises exploratórias identificarem *outliers* nos dados, esta abordagem é subjetiva e informal. Além de haver perda de informação, o uso de técnicas mais robustas, como solução deste problema, tem maior credibilidade.

Schabenberger e Gotway (2004) defendem que, a menos que uma medição tenha sido mal amostrada, a remoção de observações extremas não é o melhor caminho. Além disso, retirá-la reduz o número de pares disponíveis nas classes. A fim de manter as observações, mas reduzir sua influência negativa, pode-se escolher uma estatística que é menos afetada pelo valor discrepante.

No estudo aqui realizado foi possível perceber a grande influência exercida pelos *outliers*. E, como defendem Genton (1998) e Schabenberger e Gotway (2004), a escolha de um estimador robusto é uma opção viável, visto que, em casos de contaminação dos dados, encontramos estimadores que obtiveram resultados significativamente melhores que os de outros estimadores.

Silva, Guimarães e Tavares (2008) observaram que a presença de tendência nos dados avaliados dificulta o ajuste de modelos adequados ao comportamento espacial da variável. Os autores demonstram como há alteração no comportamento de semivariogramas esféricos quando os dados amostrais apresentam diferentes tipos de tendência. Ainda segundo Silva, Guimarães e Tavares (2008), nos semivariogramas com tendência linear e quadrática não foi possível determinar o alcance da dependência. Nestes casos, o semivariograma deve ser ajustado com os resíduos e, posteriormente,

a tendência deve ser novamente adicionada.

Nas comparações aqui realizadas também foi possível identificar o quanto a tendência mascara o comportamento da dependência espacial. Os erros tornaram-se muito maiores quando havia tendência linear e, principalmente, nos casos de tendência quadrática.

Também foi possível perceber que o tamanho da amostra influenciou o desempenho dos estimadores. Um exemplo disso é o estimador de Genton, que teve um desempenho ruim nos bancos de dados pequenos, contudo, nos maiores, o desempenho foi muito melhor, em relação aos demais estimadores.

Trabalhos como o de Mingoti e Rosa (2008) apontam a relevância da escolha do estimador de semivariância para a qualidade final da análise. Em seu trabalho, os autores compararam alguns estimadores e assinalaram diferenças entre eles. Vale destacar que, apesar da similaridade entre o trabalho aqui desenvolvido com o de Mingoti e Rosa (2008), há diferenças significativas nas características dos dados trabalhados (tendência e tamanho de amostra) e nos estimadores utilizados (Pairwise, New-1 e New-2).

No trabalho de Mingoti e Rosa (2008) foram adotados bancos de dados de tamanho 100, sendo realizadas 100 simulações com casos sem *outliers* e com *outliers*. Foram abordados apenas os cinco primeiros estimadores apresentados nesta dissertação. Os autores indicaram a utilização dos estimadores de Genton e o das medianas para os casos com *outliers* e o de Matheron e o de Haslett, para os casos sem *outliers*.

Considerando a mesma situação do trabalho de Mingoti e Rosa (2008), e desconsiderando os estimadores diferentes, foi possível constatar algumas diferenças nos resultados. Nos casos com *outliers*, o estimador de Cressie e Hawkins foi o que obteve os melhores resultados; já no caso sem *outliers*, também se indica a utilização do estimador de Haslett.

Outro fato interessante é que Li e Lake (1994), quando propuseram os estimadores New-1 e New-2, afirmaram que o segundo era melhor. Todavia, assim como no trabalho de Marins (2006), o estimador New-1

apresentou desempenho melhor que o estimador New-2, o qual também se apresentou como um bom estimador.

Genton (1998) também discute a importância da escolha do estimador a ser utilizado. O autor afirma que o estimador de variograma robusto melhora significativamente a estimativa do variograma que, segundo o autor, é uma etapa crucial da predição espacial, pois determina os pesos de krigagem. Ele enfatiza a importância de o estimador se manter próximo ao variograma verdadeiro, mesmo quando *outliers* estão presentes nos dados. Caso contrário, a krigagem pode produzir mapas com informações erradas.

De fato, foi possível perceber que a escolha do estimador pode exercer grande influência nos resultados. Assim, é importante fazer um estudo detalhado para se escolher o estimador a ser utilizado. Soares (2000, p. 9) faz uma colocação muito interessante acerca da existência de diferenças entre modelos, assim como denominações de “melhor modelo”:

Os modelos geoestatísticos, tal como acontece na maioria das ciências aplicadas, estão “limitados” ao debate do mundo real dos fenômenos naturais. A robustez e a eficácia dos modelos são assim válidas pelo fenômeno que está a ser estudado. Assim, um modelo é considerado bom, ou mau, na medida do veredito que a natureza decida conferir aos seus resultados. Este quadro tem como corolário alguns princípios orientadores da prática geoestatística: não existe um modelo universalmente bom para todas as situações, isto é, o mesmo modelo pode ser bom para um caso de estudo e mau noutro; não é condição suficiente um modelo assentar num corpo teórico rico e coerente para que produza bons resultados.

Assim como mostra Soares (2000, p. xi), nenhum estimador é totalmente bom, nem totalmente mau. Sua eficácia depende do fenômeno estudado. Em cada situação um estimador é o melhor. Assim, apenas devemos saber em qual situação deve-se utilizar cada um.

4.5 Aplicação

Para exemplificação, foi tomado um banco de dados referente a eucaliptos, fornecido gentilmente pelo professor José Márcio de Mello. Foi estudada uma amostra de 161 parcelas, que se encontra em anexo. Na tabela constam:

- parcelas[**a**] - número da parcela;
- longitude[**b**] - em UTM (metros);
- latitude[**c**] - em UTM (metros);
- índice de sítio[**d**] - em metros.

Assim a variável em estudo é o índice de sítio. Ela refere-se à classe de produtividade da área, sendo a medida da qualidade do sítio baseada na altura média das árvores dominantes no povoamento, em uma determinada idade de referência.

Mello (2004) realizou um trabalho que abrangia a mesma variável aqui abordada. O autor também ressaltou a importância da utilização de técnicas geoestatísticas no estudo da dependência espacial de características dendrométricas para eucaliptos.

As análises foram realizadas no *software* R. Detalhes da análise de dados georreferenciados, como comandos e funções, utilizando este *software* podem ser encontradas em literaturas como Bivand, Pebesma e Rubio (2008) e Diggle e Ribeiro Júnior (2007).

A disposição dos dados pode ser verificada na Figura 3, por meio da qual pode-se observar a configuração dos dados no espaço. É possível perceber que os dados não encontram-se regularmente espaçados e que alguns encontram-se um pouco dispersos dos demais.

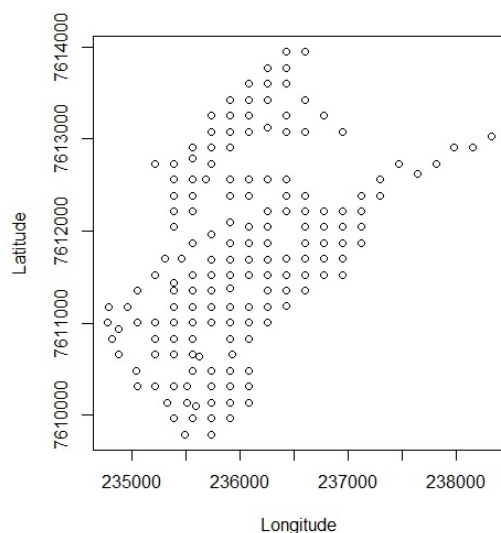


Figura 3 Mapa da localização espacial dos dados

A princípio, foi realizada uma análise descritiva dos dados. Em relação às coordenadas, tem-se que a longitude variou de 234.778 a 238.330, enquanto a latitude foi de 7.609.791 à 7.613.946.

Uma breve análise descritiva, em relação à variável índice de sítio, foi realizada. Observou-se que o valor mínimo foi de 17 m e o máximo, de 33 m. Assim, tem-se uma amplitude de 16 m. A média foi de 26,78 m e a mediana, 27 m, sendo estes valores muito parecidos.

Para detectar dados discrepantes foi realizado um estudo do gráfico boxplot, por meio do qual foi possível verificar a existência de um *outlier*, como mostrado no Gráfico 21. Através deste gráfico, também foi possível verificar a distribuição dos dados e os quartis, assim como a proximidade dos valores da média e da mediana.

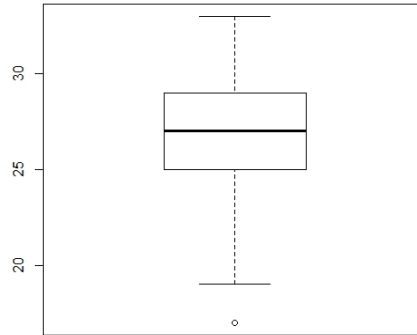
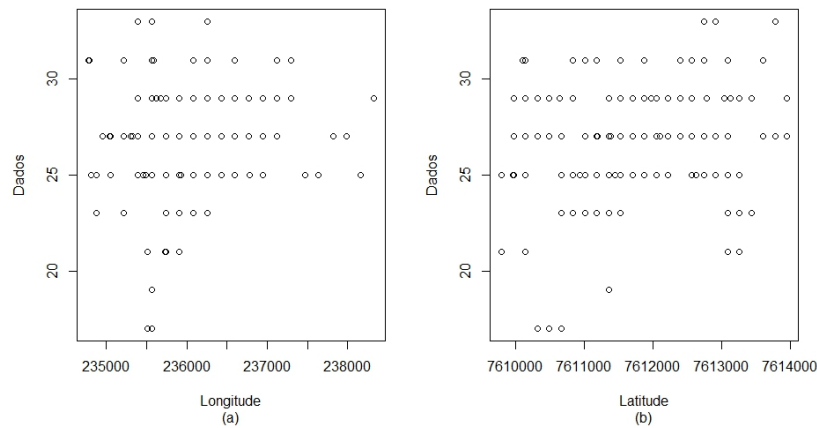


Gráfico 21 Boxplot

Também foi averiguado se há existência de tendência nos dados. Para tanto, foram plotados os gráficos de dispersão dos dados *versus* a longitude e a latitude. Observando o Gráfico 22, é possível perceber que os dados encontram-se dispostos de forma aleatória. Assim, verificou-se que não há existência de tendência em nenhuma das duas direções.

Gráfico 22 Gráficos de dispersão dos dados *versus* as coordenadas

A construção do semivariograma requer a escolha de um estimador de semivariância. Para tanto, verificou-se nos dados características de tendência, quantidade de *outliers* e tamanho da amostra. Assim, foram considerados os seguintes fatores:

- a) o tamanho da amostra, que foi considerada como próximo de 100 dados;
- a) a quantidade de *outliers*, aproximadamente 1%;
- a) a não existência de tendência.

Diante destas características e levando em consideração os resultados das simulações, apresentados anteriormente nesta dissertação, indica-se a utilização do estimador New-1, proposto por Li e Lake (1994), e que é definido pela expressão 2.15

Quando os dados são escassos ou de forma irregular, como é o caso, o número $N(h)$ de pares distintos de uma distância h pode não ser suficiente para obter uma estimativa estável em h . Neste caso a literatura (ANDRIOTTI, 2003; SOARES 2000) indica que se trabalhe com classes de h , assim, foram adotadas treze classes de mesma amplitude. As estimativas realizadas por meio do New-1 podem ser verificadas no gráfico Gráfico 23, em que o eixo das abcissas apresenta os pontos médios das classes adotadas.

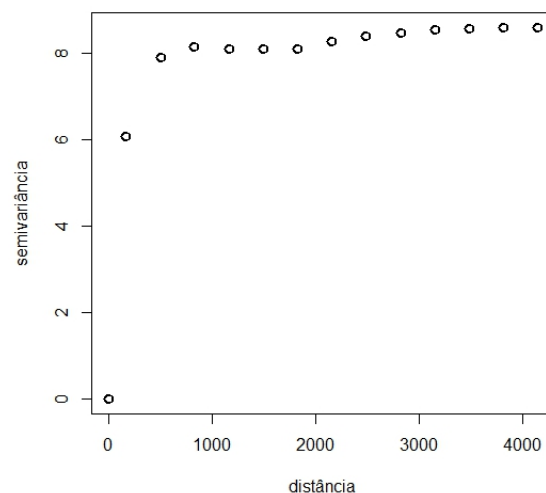


Gráfico 23 Semivariograma da variável índice de sítio

Assim como no trabalho de Mello (2004), foi possível perceber certa estrutura de dependência espacial na variável índice de sítio. Um fator interessante, que pode ser observado no gráfico acima, é o comportamento regular das estimativas obtidas por meio do New-1.

As análises geoestatísticas comumente utilizam o estimador de Matheron e o estimador Robusto de Cressie e Hawkins. A adoção destes estimadores, muitas vezes, é realizada devido à facilidade de aplicação e ao fato destes estimadores estarem fortemente difundidos na literatura. Contudo, através deste trabalho, percebemos que existem diferenças relevantes nos resultados, de acordo com o estimador utilizado. Diante das características do banco de dados abordado neste exemplo, os resultados das simulações sugeriram a utilização do estimador New-1, pois este apresentou os menores erros nesta situação. No Gráfico 24 pode-se comparar a diferença das estimativas deste três estimadores de semivariância.

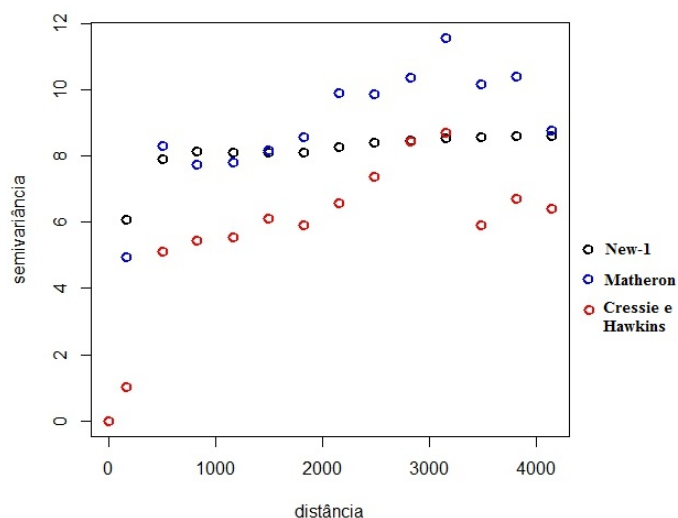


Gráfico 24 Semivariograma da variável índice de sítio

Neste gráfico, pode-se observar que a estrutura de variabilidade espacial apresentou patamar diferente em cada estimador. O estimador de Matheron apresentou alguns valores próximos aos do estimador New-1, contudo estabilizou as estimativas em torno do valor 10, diferentemente do estimador New-1, que estabilizou em torno de 8. O estimador robusto de Cressie e Hawkins apresentou valores mais baixos e uma estrutura com mais oscilações, pode-se dizer que seu patamar encontra-se em torno do valor 6.

Assim, considerando que, nestas configurações, o estimador New-1 apresenta os menores erros, a utilização do estimador de Matheron ou do estimador de Cressie e Hawkins acarreta um aumento significativo do erro das estimativas.

5 CONCLUSÃO

Por meio das simulações realizadas nesta dissertação, foi possível observar a influência da tendência na construção do semivariograma. A combinação de tendência e *outliers* foi extremamente negativa para alguns estimadores, podendo-se ressaltar o Pairwise, que foi muito afetado, principalmente, na presença de tendência quadrática.

O estimador New-1, de forma geral, apresentou os melhores resultados em situações sem tendência e com tendência linear, na ausência de *outliers*, com 1% e, às vezes, com 5% de contaminação. Já os estimadores das medianas e robusto de Cressie e Hawkins demonstraram um melhor comportamento, em casos de contaminações mais severas, de 5% e 10% de *outliers* nos bancos de dados de tamanho 25 e 100.

Nas amostras de tamanho 400, o estimador New-1 apresentou bons resultados nos casos sem tendência e com tendência linear, com 0% e 1% de contaminação. Os estimadores de Genton, Haslett e das medianas revezaram os menores erros nos casos de 5% e 10% de contaminação, para este tamanho de amostra.

Na tendência quadrática, o estimador que se destacou foi o New-1, nos bancos de dados de tamanho 25 e 100, sendo que no segundo tamanho, os estimadores de Haslett e Genton também obtiveram bons resultados, alternando os menores erros. Já nas amostras de tamanho 400, o estimador com melhor desempenho foi o de Genton, com exceção dos casos sem *outliers*, em que o estimador Pairwise apresentou o melhor desempenho.

Outro fator observado foi que, a utilização do *cutoff* não foi relevante para a comparação entre os estimadores. Contudo, sua utilização reduz consideravelmente o erro das estimativas em um mesmo estimador.

A média e a mediana apresentaram resultados que divergiram em poucas situações. Em um contexto geral, não foi percebida diferença significativa entre seus resultados.

A literatura, muitas vezes, apresenta o estimador de Matheron como o único estimador existente. Contudo, com o presente trabalho, mostrou-

se que, dependendo das condições dos dados, a escolha do estimador de semivariância tem grande influência no resultado final.

No exemplo aqui apresentado, foi possível perceber que, a utilização do estimador de Matheron, ou mesmo o de Cressie e Hawkins, ao invés do estimador New-1, que foi recomendado pelos resultados das simulações, modificaria significamente o semivariograma final, chegando a alterar o patamar. Assim, para redução dos erros e melhor construção do semivariograma, recomenda-se cautela e atenção na escolha do estimador de semivariância, sempre considerando as características dos dados estudados.

REFERÊNCIAS

ALFONS, A.; TEMPL, M.; FILZMOSER, P. An object-oriented framework for statistical simulation: the R package simFrame. **Journal of Statistical Software**, Los Angeles, v. 37, n. 3, p. 1-36, 2010.

ANDRIOTTI, J. L. S. **Fundamentos de estatística e geoestatística**. São Leopoldo: UNISINOS, 2003. 165 p. (Coleção Manual Universitário).

BAILEY, T. C.; GATRELL, A. C. **Interactive spatial data analysis**. Harlow: Longman Scientific & Technical, 1995. 413 p.

BIVAND, R. S.; PEBESMA, E. J.; GOMEZ-RUBIO, V. **Applied spatial data analysis with R**. New York: Springer, 2008. 374 p.

CAMARGO, E. Geoestatística: fundamentos e aplicações. In: _____. **Geoprocessamento em projetos ambientais**. São José dos Campos: INPE, 1998. p. 1-36.

CAMPOZANA, F. P. **Modelagem probabilística e simulação de reservatórios**. 1990. 164 p. Tese (Doutorado em Engenharia de Petróleo) - Universidade Estadual de Campinas, Campinas, 1990.

CHILÈS, J.; DELFINER, P. **Geostatistics: modeling spatial uncertainty**. New York: Wiley, 1999. 695 p. (Wiley Series in Probability and Statistics).

CLARK, I. **Practical geostatistics**. London: Applied Science, 1979. 129 p.

CRESSIE, N. A. C. **Statistics for spatial data**. New York: J. Wiley, 1993. 900 p.

CRESSIE, N. A. C.; HAWKINS, D. M. Robust estimation of the variogram. **Mathematical Geology**, New York, v. 12, n. 2, p. 115-125, 1980.

DIGGLE, P. J.; RIBEIRO JUNIOR, P. J. R. **Model based geostatistics**. New York: Springer, 2007. 230 p.

EMERSON, J. D.; HOAGLIN, D. C. Análise de tabelas bidimensionais através de medianas. In: HOAGLIN, D. C.; MOSTELLER, F.; TUKEY, J. W. (Ed.). **Análise exploratória de dados: técnicas robustas**. Lisboa: Salamandra, 1983. p. 167-212.

EMERSON, J. D.; STRENIO, J. Caixas de bigodes e comparação de colecções. In: HOAGLIN, D. C.; MOSTELLER, F.; TUKEY, J. W. (Ed.). **Análise exploratória de dados: técnicas robustas**. Lisboa: Salamandra, 1983. p. 59-96.

GENTON, M. G. Highly robust variogram estimation. **Mathematical Geology**, New York, v. 30, n. 2, p. 213-221, 1998.

GRONDONA, M. O.; CRESSIE, N. Using spatial considerations in the analysis of experimentos. **Technometrics**, Alexandria, v. 33, n. 4, p. 381-392, 1991.

GUERRA, P. A. G. **Geoestatística operacional**. Brasília: Ministério das Minas e Energia, 1988. 145 p.

GUIMARÃES, E. C. **Geoestatística básica e aplicada**. Uberlândia: UFU/FAMAT, 2004. 77 p. Apostila.

HASLETT, J. On the sample variogram and the sample autocovariance for non-stationary time series. **The Statistician**, Washington, v. 46, n. 4, p. 475-485, 1997.

ISAAKS, E. H.; SRIVASTAVA, R. M. **Applied geostatistics**. Oxford: Oxford University, 1989. 560 p.

JOURNEL, A. G.; HUIJBREGTS, C. J. **Mining geostatistics**. San Diego: Academic, 1978. 600 p.

LI, D.; LAKE, L. W. A moving window semivariance estimator. **Water Resources Research**, Washington, v. 30, n. 5, p. 1479-1490, May 1994.

MARINS, A. **Métodos de estimação da função semivariância aplicados a dados simulados e reais de produtividade da soja e de atributos físicos de um latossolo**. 2006. 140 p. Dissertação (Mestrado em Engenharia Agrícola) - Universidade Estadual do Oeste do Paraná, Cascavel, 2006.

MATHERON, G. **Traité de géostatistique appliquée**. Paris: Technip, 1962. v. 14.

MELLO, J. M. **Geoestatística aplicada ao inventário florestal**. 2004. 110 p. Tese (Doutorado em Recursos Florestais, Silvicultura e Manejo Florestal) - Escola Superior de Agricultura "Luiz de Queiroz", Piracicaba, 2004.

MINGOTI, S. A.; ROSA, G. A note on robust and non-robust variogram estimators. **Revista Escola de Minas**, Ouro Preto, v. 61, n. 1, p. 87-95, 2008.

MOOD, A. M. **Introduction to the theory of statistics**. New York: McGraw-Hill, 1950. 564 p.

OLIVEIRA, M. S. **Planos amostrais para variáveis espaciais utilizando geoestatística**. 1991. 100 p. Dissertação (Mestrado em Estatística) - Universidade Estadual de Campinas, Campinas, 1991.

PONTES, J. M. **A geoestatística: aplicações em experimentos de campo.** 2002. 82 p. Dissertação (Mestrado em Estatística e Experimentação Agropecuária) - Universidade Federal de Lavras, Lavras, 2002.

R DEVELOPMENT CORE TEAM. **R: a language and environment for statistical computing.** Vienna: R Foundation for Statistical Computing, 2012. Disponível em: <<http://www.R-project.org>>. Acesso em: 20 dez. 2012.

RIBEIRO JUNIOR, P. J. R.; DIGGLE, P. J. GeoR: a package for geostatistical analysis. **R-NEWS**, New York, v. 1, n. 2, p. 14-18, June 2001.

SCHABENBERGER, O.; GOTWAY, C. **Statistical methods for spatial data analysis: texts in statistical science.** New York: Taylor & Francis, 2004. 488 p.

SCHLATHER, M. Randomfields: simulation and analysis of random fields: r package version 2.0.59. **Journal of Statistical Software**, Los Angeles, v. 1, n. 2, p. 18-20, June 2006.

SILVA, A. R. D.; GUIMARÃES, E. C.; TAVARES, M. Efeito de tendência no ajuste de semivariogramas esféricos. **FAMAT em Revista**, Uberlândia, v. 10, p. 107-116, abr. 2008.

SOARES, A. **Geoestatística para as ciências da terra e do ambiente.** 2. ed. Lisboa: IST, 2000. 214 p. (Ensino da Ciência e da Tecnologia, 9).

WALLER, L. A.; GOTWAY, C. A. **Applied spatial statistics for public health data.** New Jersey: Wiley-Interscience, 2004. 494 p.

ANEXOS

ANEXO A - Script da simulação

```
#####
####          SIMULAÇÃO: COM OUTLIER E COM TENDÊNCIA LINEAR      ####
####          #####
#####

## Packages utilizados:
  library(RandomFields)
  library(geoR)
  require(simFrame)
#####
m<-500
result1<-data.frame(EQM1=rep(0,m),EQM2=rep(0,m),
  EQM3=rep(0,m),EQM4=rep(0,m),EQM5=rep(0,m),
  EQM6=rep(0,m),EQM7=rep(0,m),EQM8=rep(0,m))
result2<-data.frame(EAM1=rep(0,m),EAM2=rep(0,m),
  EAM3=rep(0,m),EAM4=rep(0,m),EAM5=rep(0,m),
  EAM6=rep(0,m),EAM7=rep(0,m),EAM8=rep(0,m))
result3<-data.frame(EQM1.5=rep(0,m),EQM2.5=rep(0,m),
  EQM3.5=rep(0,m),EQM4.5=rep(0,m),EQM5.5=rep(0,m),
  EQM6.5=rep(0,m),EQM7.5=rep(0,m),EQM8.5=rep(0,m))
result4<-data.frame(EAM1.5=rep(0,m),EAM2.5=rep(0,m),
  EAM3.5=rep(0,m),EAM4.5=rep(0,m),EAM5.5=rep(0,m),
  EAM6.5=rep(0,m),EAM7.5=rep(0,m),EAM8.5=rep(0,m))

for(k in 1:m)
{
repeat
{
#####
## Parâmetros para simulação:
  model      <- "sph" # modelo esférico
  media      <- 10
  variancia  <- 4
  pepita     <- 0.5
  contribuicao <- 4
  patamar    <- pepita+contribuicao # sill
  alcance    <- 6
  realizacao <- 1 # número de realizações
  param      <- c(variance=variancia,nugget=pepita,
```

```

                                scale=alcance)
trend      <- function(x,y) media + 0.05*x
## definindo as coordenadas:
esp1 <- 1.0 # espaçamento entre os pontos
x <- seq(1,10,esp1)
y <- seq(1,10,esp1)
## Construindo o GRIDE onde os dados serão gerados:
coordxy <- cbind(x,y)
coordXY <- expand.grid(X=x,Y=y)
plot(coordXY) # visualizando o gride

#####
## simulando:
simul.Gauss <- c(GaussRF(x=x,y=x,model=model,grid=T,
                        param=param,trend=trend))

#####
## Contaminando os dados com outliers:
simul.Gauss<-data.frame(simul.Gauss)

cc1<- DARContControl(target = "simul.Gauss",
epsilon=(0.03) , fun = function(x) x - x + runif(3,2,6))
simul.Gauss<-contaminate(simul.Gauss, cc1)
simul.Gauss<-simul.Gauss[,1]
simul.Gauss<-data.frame(simul.Gauss)

cc2 <- DARContControl(target = "simul.Gauss",
epsilon=(0.02) , fun = function(x) x - x + runif(2,15,20))
simul.Gauss<-contaminate(simul.Gauss, cc2)
simul.Gauss<-simul.Gauss[,1]
dados<-data.frame(cbind(coordXY,simul.Gauss))

bp<-boxplot(simul.Gauss) # boxplot
if (length(bp$out)==5) break
}
bp$out # quantidade de outliers

#####
## cálculo das distâncias e organização dos dados
## para aplicação dos estimadores de semivariância

## Achando as distâncias:
distancia=as.factor(as.matrix(dist(dados[,1:2],
method = "euclidean", diag = T, upper = T)))

```

```

n=length(distancia)
hh<-as.numeric(levels(distancia)) # níveis da distância

## trocando por uma sequência de nº inteiros
levels(distancia)<-1:length(levels(distancia))

## diferença e soma dos pares de pontos:
va.rep1<-rep(simul.Gauss,each=sqrt(n))
va.rep2<-c(rep(simul.Gauss,sqrt(n)))
seq1<-rep(1:sqrt(length(va.rep1)),each=sqrt(length(va.rep1)))
seq2<-rep(1:sqrt(length(va.rep2)),sqrt(length(va.rep2)))
v1<-data.frame(v1=ifelse(seq1<=seq2,(va.rep1-va.rep2),"a"))
v2<-data.frame(v2=ifelse(seq1<=seq2,(va.rep1+va.rep2),"a"))
v3<-data.frame(ifelse(seq1<=seq2,distancia,"a"))
dif<-as.numeric(v1[v1!="a"]) # diferença entre os pares de pts
soma<-as.numeric(v2[v2!="a"]) # soma entre os pares de pts
h<-as.numeric(v3[v3!="a"]) # distância entre os pares de pts
cutoff=ceiling(0.5*length(hh))

## Cálculo do modelo teórico:
esferico<-ifelse(hh<=alcance,2*(pepita+(contribuicao*(1.5*
(hh/alcance)-0.5*((hh/alcance)^3))), 2*patamar)
esferico[1]=0
esferico<-esferico[1:(length(hh)-1)] # sem o último ponto

#####
###          Estimador Clássico de Matheron          ###
#####
qdif=((dif)^2) # quadrado da diferença
dados1<-data.frame(cbind(h,qdif))
colnames(dados1) <- c("h","qdif")

Est.Mat <- matrix(0,length(hh))
for(i in 1:length(hh))
{
Est.Mat[i] <- (sum(dados1$qdif[dados1$h== i]))/
length(dados1$qdif[dados1$h== i])
}
Est.Mat
Est.Mat<-Est.Mat[1:(length(hh)-1)]
EQM1<-(sum((esferico-Est.Mat)^2))/(length(hh)-1)
EQM1.5<-(sum((esferico[1:cutoff]-Est.Mat[1:cutoff])^2))/cutoff
EAM1<-(sum(abs(esferico-Est.Mat)))/(length(hh)-1)
EAM1.5<-sum(abs(esferico[1:cutoff]-Est.Mat[1:cutoff]))/cutoff

```

```
#####
###          Estimador Robusto de Cressie e Hawkins          ###
#####
    radif=(abs(dif))^0.5 # raiz quadrada do valor absoluto das dif.
    dados2<-data.frame(cbind(h,radif))
    colnames(dados2) <- c("h","radif")

    Est.CH <- matrix(0,length(hh))
    for(i in 1:length(hh))
    {
    Est.CH[i] <- (((sum(dados2$radif[dados2$h== i]))/
        length(dados2$radif[dados2$h== i]))^4/
        (0.457+(0.494/(length(dados2$radif[dados2$h== i])))))
    }
    Est.CH
    Est.CH<-Est.CH[1:(length(hh)-1)]
EQM2<-(sum((esferico-Est.CH)^2))/(length(hh)-1)
EQM2.5<-(sum((esferico[1:cutoff]-Est.CH[1:cutoff])^2))/cutoff
EAM2<-(sum(abs(esferico-Est.CH)))/(length(hh)-1)
EAM2.5<-sum(abs(esferico[1:cutoff]-Est.CH[1:cutoff]))/cutoff

#####
###          Estimador das Medianas          ###
#####
    radif=(abs(dif))^0.5 # raiz quadrada do valor absoluto das dif.
    dados3<-data.frame(cbind(h,radif))
    colnames(dados3) <- c("h","radif")

    Est.Med <- matrix(0,length(hh))
    for(i in 1:length(hh))
    {
    Est.Med[i] <- (median(dados3$radif[dados3$h== i]))^4/(0.457)
    }
    Est.Med
    Est.Med<-Est.Med[1:(length(hh)-1)]
EQM3<-(sum((esferico-Est.Med)^2))/(length(hh)-1)
EQM3.5<-(sum((esferico[1:cutoff]-Est.Med[1:cutoff])^2))/cutoff
EAM3<-(sum(abs(esferico-Est.Med)))/(length(hh)-1)
EAM3.5<-sum(abs(esferico[1:cutoff]-Est.Med[1:cutoff]))/cutoff

#####
###          Estimador das Diferenças de Haslett          ###
#####
```

```

dados4<-data.frame(cbind(h,dif))
colnames(dados4) <- c("h","dif")

Est.Has <- matrix(0,length(hh))
for(i in 1:length(hh))
{
Est.Has[i] <- sum((dados4$dif[dados4$h== i]-
mean(dados4$dif[dados4$h== i]))^2)/((
length(dados4$dif[dados4$h== i])-1))
}
Est.Has
Est.Has<-Est.Has[1:(length(hh)-1)]
EQM4<-(sum((esferico-Est.Has)^2))/(length(hh)-1)
EQM4.5<-(sum((esferico[1:cutoff]-Est.Has[1:cutoff])^2))/cutoff
EAM4<-(sum(abs(esferico-Est.Has)))/(length(hh)-1)
EAM4.5<-sum(abs(esferico[1:cutoff]-Est.Has[1:cutoff]))/cutoff

#####
###          Estimador Altamente Robusto de Genton          ###
#####

dados5<-data.frame(cbind(h,dif))
colnames(dados5) <- c("h","Vh")

Est.Gen <- matrix(0,length(hh))
for(i in 1:length(hh))
{
vh.rep1<-rep(dados5$Vh[dados5$h== i],
each=length(dados5$Vh[dados5$h== i]))
vh.rep2<-rep(dados5$Vh[dados5$h== i],
length(dados5$Vh[dados5$h== i]))
vh.seq1<-rep(1:length(dados5$Vh[dados5$h== i]),
each=length(dados5$Vh[dados5$h== i]))
vh.seq2<-rep(1:length(dados5$Vh[dados5$h== i]),
length(dados5$Vh[dados5$h== i]))
qn.2<-data.frame(qn.2=ifelse(vh.seq1<vh.seq2,
(2.2191*abs(vh.rep1-vh.rep2))^2,"a"))
Qn.2<-as.numeric(qn.2[qn.2!="a"])
Est.Gen[i] <- quantile(Qn.2,((floor(length
(dados5$Vh[dados5$h== i])/2)+1)/2)/
length(dados5$Vh[dados5$h== i]))
}
Est.Gen
Est.Gen<-Est.Gen[1:(length(hh)-1)]
EQM5<-(sum((esferico-Est.Gen)^2))/(length(hh)-1)

```

```

EQM5.5<-(sum((esferico[1:cutoff]-Est.Gen[1:cutoff])^2))/cutoff
EAM5<-(sum(abs(esferico-Est.Gen)))/(length(hh)-1)
EAM5.5<-sum(abs(esferico[1:cutoff]-Est.Gen[1:cutoff]))/cutoff

#####
###          Estimador de Pairwise          ###
#####
  div<-(dif/soma)^2
  dados6<-data.frame(cbind(h,div))
  colnames(dados6) <- c("h","div")

  Est.Pw <- matrix(0,length(hh))
  for(i in 1:length(hh))
  {
  Est.Pw[i] <-4*(sum(dados6$div[dados6$h== i]))/
    (length(dados6$div[dados6$h== i]))
  }
  Est.Pw
  Est.Pw<-(Est.Pw[1:(length(hh)-1)]*(Est.Mat[2]/Est.Pw[2]))
EQM6<-(sum((esferico-Est.Pw)^2))/(length(hh)-1)
EQM6.5<-(sum((esferico[1:cutoff]-Est.Pw[1:cutoff])^2))/cutoff
EAM6<-(sum(abs(esferico-Est.Pw)))/(length(hh)-1)
EAM6.5<-sum(abs(esferico[1:cutoff]-Est.Pw[1:cutoff]))/cutoff

#####
###          Estimador New-1          ###
#####
  va.rep1<-rep(simul.Gauss,each=sqrt(n))
  va.rep2<-c(rep(simul.Gauss,sqrt(n)))
  seq1<-rep(1:sqrt(length(va.rep1)),each=sqrt(length(va.rep1)))
  seq2<-rep(1:sqrt(length(va.rep2)),sqrt(length(va.rep2)))

  new1.dif<-data.frame(new1.dif=ifelse(seq1!=seq2,
    ((va.rep1-va.rep2)^2),"a"))
  New1.dif<-as.numeric(new1.dif[new1.dif!="a"])
  new1.h<-data.frame(new1.h=ifelse(seq1!=seq2,distancia,"a"))
  New1.h<-as.numeric(new1.h[new1.h!="a"])
  new1.seq<-data.frame(new1.seq=ifelse(seq1!=seq2,seq1,"a"))
  New1.seq<-as.numeric(new1.seq[new1.seq!="a"])

  dados7<-data.frame(cbind(New1.h,New1.dif,New1.seq))
  colnames(dados7)<-c("h","Nwd","seq")

  tab<-data.frame(tapply(dados7$Nwd, list(dados7$h,

```

```

                                dados7$seq), mean))
Est.New1.<-matrix(0,(length(hh)-1))
for (i in 1:(length(hh)-1))
{
  Est.New1.[i]<-2*(mean(((colMeans(tab[1:i,],na.rm = TRUE))/2),
                                na.rm = TRUE ))
}
Est.New1.<-matrix(c(0,Est.New1.))
Est.New1<-matrix(c(Est.New1.[1:(length(hh)-1)]))
EQM7<-(sum((esferico-Est.New1)^2))/(length(hh)-1)
EQM7.5<-(sum((esferico[1:cutoff]-Est.New1[1:cutoff])^2))/cutoff
EAM7<-(sum(abs(esferico-Est.New1)))/(length(hh)-1)
EAM7.5<-sum(abs(esferico[1:cutoff]-Est.New1[1:cutoff])/cutoff

#####
###                               Estimador New-2                               ###
#####
  dados8<-data.frame(cbind(hh,Est.New1.))
  colnames(dados8)<-c("h","New1.")

  Est.New2<-matrix(0,(length(hh)-2))
  for (i in 1:(length(hh)-2))
  {
    Est.New2[i]<- Est.New1.[i+1] +((hh[i+1])/2)*((Est.New1.[i+2]-
                                                    Est.New1.[i])/(hh[i+2]-hh[i]))
  }
  Est.New2<-matrix(c(0,Est.New2))
  Est.New2
EQM8<-(sum((esferico[1:(length(hh)-1)]-Est.New2)^2))/(length(hh)-1)
EQM8.5<-(sum((esferico[1:cutoff]-Est.New2[1:cutoff])^2))/cutoff
EAM8<-(sum(abs(esferico[1:(length(hh)-1)]-Est.New2)))/(length(hh)-1)
EAM8.5<-sum(abs(esferico[1:cutoff]-Est.New2[1:cutoff])/cutoff
#####
result1[k,]<-c(EQM1,EQM2,EQM3,EQM4,EQM5,EQM6,EQM7,EQM8)
result2[k,]<-c(EAM1,EAM2,EAM3,EAM4,EAM5,EAM6,EAM7,EAM8)
result3[k,]<-c(EQM1.5,EQM2.5,EQM3.5,EQM4.5,EQM5.5,EQM6.5,
               EQM7.5,EQM8.5)
result4[k,]<-c(EAM1.5,EAM2.5,EAM3.5,EAM4.5,EAM5.5,EAM6.5,
               EAM7.5,EAM8.5)
}
result1
result2
result3
result4

```

ANEXO B - Tabelas com os resultados

Resultados das comparações em *grides* 5x5Tabela 1B E.M.Q. e E.M.A de dados com com n=25, sem *outliers*, sem tendência e sem *Cutoff*.

Est.	Todas as distâncias			
	EMQ		EMA	
	Média	Mediana	Média	Mediana
$\hat{\gamma}_M(h)$	10,993	7,781	2,417	2,167
$\hat{\gamma}_{CH}(h)$	19,128	9,707	2,948	2,417
$\hat{\gamma}_{Md}(h)$	39,548	15,142	3,924	2,955
$\hat{\gamma}_H(h)$	10,418	9,137	2,477	2,352
$\hat{\gamma}_G(h)$	33,315	14,480	3,481	2,767
$\hat{\gamma}_{Pw}(h)$	10,006	6,996	2,333	2,084
$\hat{\gamma}_{N1}(h)$	6,272	4,968	1,995	1,946
$\hat{\gamma}_{N2}(h)$	8,318	6,215	2,222	2,073

Tabela 2B E.M.Q. e E.M.A de dados com simulação com n=25, sem *outliers*, sem tendência e *Cutoff* de 50%.

Est.	<i>Cutoff</i> de 50%			
	EMQ		EMA	
	Média	Mediana	Média	Mediana
$\hat{\gamma}_M(h)$	4,042	2,712	1,455	1,269
$\hat{\gamma}_{CH}(h)$	5,957	3,397	1,679	1,405
$\hat{\gamma}_{Md}(h)$	10,892	5,420	2,153	1,744
$\hat{\gamma}_H(h)$	3,959	3,232	1,479	1,350
$\hat{\gamma}_G(h)$	7,672	4,673	1,879	1,623
$\hat{\gamma}_{Pw}(h)$	3,665	2,564	1,396	1,225
$\hat{\gamma}_{N1}(h)$	2,875	2,258	1,294	1,252
$\hat{\gamma}_{N2}(h)$	4,482	2,899	1,548	1,371

Tabela 3B E.M.Q. e E.M.A de dados com $n=25$, com 1% *outliers*, sem tendência e sem *Cutoff*.

Est.	Todas as distâncias			
	EMQ		EMA	
	Média	Mediana	Média	Mediana
$\hat{\gamma}_M(h)$	11,957	8,833	2,610	2,410
$\hat{\gamma}_{CH}(h)$	13,579	9,562	2,659	2,372
$\hat{\gamma}_{Md}(h)$	23,259	13,787	3,263	2,847
$\hat{\gamma}_H(h)$	11,345	9,336	2,555	2,439
$\hat{\gamma}_G(h)$	43,124	20,709	3,972	3,258
$\hat{\gamma}_{Pw}(h)$	10,915	9,225	2,503	2,367
$\hat{\gamma}_{N1}(h)$	4,818	3,564	1,743	1,540
$\hat{\gamma}_{N2}(h)$	8,676	6,620	2,277	2,086

Tabela 4B E.M.Q. e E.M.A de dados com $n=25$, com 1% *outliers*, sem tendência e *Cutoff* de 50%.

Est.	<i>Cutoff</i> de 50%			
	EMQ		EMA	
	Média	Mediana	Média	Mediana
$\hat{\gamma}_M(h)$	7,588	4,481	2,054	1,725
$\hat{\gamma}_{CH}(h)$	5,882	3,598	1,747	1,511
$\hat{\gamma}_{Md}(h)$	8,081	5,339	1,984	1,783
$\hat{\gamma}_H(h)$	6,969	4,142	1,951	1,630
$\hat{\gamma}_G(h)$	10,460	5,897	2,212	1,902
$\hat{\gamma}_{Pw}(h)$	6,105	3,610	1,841	1,546
$\hat{\gamma}_{N1}(h)$	4,258	2,646	1,570	1,291
$\hat{\gamma}_{N2}(h)$	9,010	5,846	2,243	1,883

Tabela 5B E.M.Q. e E.M.A de dados com $n=25$, com 5% *outliers*, sem tendência e sem *Cutoff*.

Est.	Todas as distâncias			
	EMQ		EMA	
	Média	Mediana	Média	Mediana
$\hat{\gamma}_M(h)$	9,972	8,483	2,416	2,266
$\hat{\gamma}_{CH}(h)$	9,696	7,846	2,317	2,158
$\hat{\gamma}_{Md}(h)$	16,126	11,557	2,816	2,598
$\hat{\gamma}_H(h)$	9,998	8,677	2,436	2,311
$\hat{\gamma}_G(h)$	35,346	20,776	3,720	3,276
$\hat{\gamma}_{Pw}(h)$	10,088	8,485	2,436	2,283
$\hat{\gamma}_{N1}(h)$	3,792	3,237	1,531	1,458
$\hat{\gamma}_{N2}(h)$	7,027	5,687	2,001	1,849

Tabela 6B E.M.Q. e E.M.A de dados com $n=25$, com 5% *outliers*, sem tendência e *Cutoff* de 50%.

Est.	<i>Cutoff</i> de 50%			
	EMQ		EMA	
	Média	Mediana	Média	Mediana
$\hat{\gamma}_M(h)$	7,240	5,270	2,042	1,824
$\hat{\gamma}_{CH}(h)$	4,680	3,255	1,609	1,424
$\hat{\gamma}_{Md}(h)$	6,298	4,837	1,790	1,649
$\hat{\gamma}_H(h)$	7,025	4,909	1,998	1,764
$\hat{\gamma}_G(h)$	10,468	6,851	2,289	2,017
$\hat{\gamma}_{Pw}(h)$	6,696	4,656	1,956	1,720
$\hat{\gamma}_{N1}(h)$	4,220	2,758	1,558	1,274
$\hat{\gamma}_{N2}(h)$	8,826	6,832	2,207	1,989

Tabela 7B E.M.Q. e E.M.A de dados com $n=25$, com 10% *outliers*, sem tendência e sem *Cutoff*.

Est.	Todas as distâncias			
	EMQ		EMA	
	Média	Mediana	Média	Mediana
$\hat{\gamma}_M(h)$	19,481	14,942	3,445	3,176
$\hat{\gamma}_{CH}(h)$	14,959	10,919	2,868	2,672
$\hat{\gamma}_{Md}(h)$	23,950	15,335	3,293	2,984
$\hat{\gamma}_H(h)$	18,174	14,237	3,319	3,090
$\hat{\gamma}_G(h)$	66,267	45,039	5,343	4,822
$\hat{\gamma}_{Pw}(h)$	18,123	14,578	3,353	3,107
$\hat{\gamma}_{N1}(h)$	9,347	6,875	2,377	2,059
$\hat{\gamma}_{N2}(h)$	15,872	12,426	3,006	2,650

Tabela 8B E.M.Q. e E.M.A de dados com $n=25$, com 10% *outliers*, sem tendência e *Cutoff* de 50%.

Est.	<i>Cutoff</i> de 50%			
	EMQ		EMA	
	Média	Mediana	Média	Mediana
$\hat{\gamma}_M(h)$	17,521	13,804	3,356	3,179
$\hat{\gamma}_{CH}(h)$	9,069	6,346	2,259	2,014
$\hat{\gamma}_{Md}(h)$	9,630	6,168	2,162	1,930
$\hat{\gamma}_H(h)$	16,954	13,788	3,262	3,095
$\hat{\gamma}_G(h)$	22,216	16,070	3,514	3,192
$\hat{\gamma}_{Pw}(h)$	16,244	12,937	3,235	3,069
$\hat{\gamma}_{N1}(h)$	13,244	10,759	2,976	2,853
$\hat{\gamma}_{N2}(h)$	21,699	18,486	3,684	3,538

Tabela 9B E.M.Q. e E.M.A de dados com $n=25$, com sem *outliers*, com tendência linear e sem *Cutoff*.

Est.	Todas as distâncias			
	EMQ		EMA	
	Média	Mediana	Média	Mediana
$\hat{\gamma}_M(h)$	12,711	8,384	2,580	2,291
$\hat{\gamma}_{CH}(h)$	21,972	11,432	3,177	2,669
$\hat{\gamma}_{Md}(h)$	46,206	18,279	4,288	3,249
$\hat{\gamma}_H(h)$	11,194	9,464	2,574	2,467
$\hat{\gamma}_G(h)$	35,160	14,033	3,563	2,749
$\hat{\gamma}_{Pw}(h)$	10,801	7,531	2,435	2,188
$\hat{\gamma}_{N1}(h)$	6,539	5,231	2,045	1,989
$\hat{\gamma}_{N2}(h)$	9,390	6,736	2,349	2,125

Tabela 10B E.M.Q. e E.M.A de dados com $n=25$, com sem *outliers*, com tendência linear e *Cutoff* de 50%.

Est.	<i>Cutoff</i> de 50%			
	EMQ		EMA	
	Média	Mediana	Média	Mediana
$\hat{\gamma}_M(h)$	4,608	2,887	1,536	1,293
$\hat{\gamma}_{CH}(h)$	6,877	3,750	1,795	1,504
$\hat{\gamma}_{Md}(h)$	12,556	5,714	2,307	1,856
$\hat{\gamma}_H(h)$	4,291	3,560	1,539	1,435
$\hat{\gamma}_G(h)$	8,111	4,437	1,891	1,574
$\hat{\gamma}_{Pw}(h)$	4,045	2,787	1,470	1,288
$\hat{\gamma}_{N1}(h)$	3,040	2,313	1,355	1,276
$\hat{\gamma}_{N2}(h)$	5,208	3,112	1,637	1,432

Tabela 11B E.M.Q. e E.M.A de dados com $n=25$, com 1% *outliers*, com tendência linear e sem *Cutoff*.

Est.	Todas as distâncias			
	EMQ		EMA	
	Média	Mediana	Média	Mediana
$\hat{\gamma}_M(h)$	21,374	13,343	3,448	3,017
$\hat{\gamma}_{CH}(h)$	18,969	10,793	3,052	2,600
$\hat{\gamma}_{Md}(h)$	31,836	14,551	3,654	2,923
$\hat{\gamma}_H(h)$	17,976	13,597	3,200	2,944
$\hat{\gamma}_G(h)$	50,701	25,441	4,368	3,471
$\hat{\gamma}_{Pw}(h)$	17,901	13,596	3,203	3,054
$\hat{\gamma}_{N1}(h)$	7,920	4,750	2,194	1,793
$\hat{\gamma}_{N2}(h)$	15,415	8,864	2,962	2,395

Tabela 12B E.M.Q. e E.M.A de dados com $n=25$, com 1% *outliers*, com tendência linear e *Cutoff* de 50%.

Est.	<i>Cutoff</i> de 50%			
	EMQ		EMA	
	Média	Mediana	Média	Mediana
$\hat{\gamma}_M(h)$	15,007	9,546	2,886	2,564
$\hat{\gamma}_{CH}(h)$	8,626	4,323	2,052	1,643
$\hat{\gamma}_{Md}(h)$	10,638	5,655	2,204	1,813
$\hat{\gamma}_H(h)$	12,955	8,068	2,652	2,275
$\hat{\gamma}_G(h)$	13,786	7,679	2,544	2,106
$\hat{\gamma}_{Pw}(h)$	11,285	6,762	2,482	2,139
$\hat{\gamma}_{N1}(h)$	8,660	4,524	2,246	1,817
$\hat{\gamma}_{N2}(h)$	17,343	11,925	3,154	2,809

Tabela 13B E.M.Q. e E.M.A de dados com $n=25$, com 5% *outliers*, com tendência linear e sem *Cutoff*.

Est.	Todas as distâncias			
	EMQ		EMA	
	Média	Mediana	Média	Mediana
$\hat{\gamma}_M(h)$	33,383	24,736	4,427	4,057
$\hat{\gamma}_{CH}(h)$	18,428	11,906	3,079	2,752
$\hat{\gamma}_{Md}(h)$	24,735	13,898	3,334	2,901
$\hat{\gamma}_H(h)$	30,899	23,779	4,230	3,933
$\hat{\gamma}_G(h)$	70,253	37,644	5,067	4,231
$\hat{\gamma}_{Pw}(h)$	32,459	23,158	4,403	4,005
$\hat{\gamma}_{N1}(h)$	16,640	9,998	3,221	2,685
$\hat{\gamma}_{N2}(h)$	25,998	18,841	3,933	3,536

Tabela 14B E.M.Q. e E.M.A de dados com $n=25$, com 5% *outliers*, com tendência linear e *Cutoff* de 50%.

Est.	<i>Cutoff</i> de 50%			
	EMQ		EMA	
	Média	Mediana	Média	Mediana
$\hat{\gamma}_M(h)$	27,955	20,218	4,120	3,801
$\hat{\gamma}_{CH}(h)$	9,906	5,812	2,289	1,897
$\hat{\gamma}_{Md}(h)$	9,236	5,843	2,109	1,822
$\hat{\gamma}_H(h)$	26,868	19,087	3,982	3,582
$\hat{\gamma}_G(h)$	19,189	12,072	3,050	2,698
$\hat{\gamma}_{Pw}(h)$	27,876	20,241	4,133	3,755
$\hat{\gamma}_{N1}(h)$	21,019	14,246	3,659	3,347
$\hat{\gamma}_{N2}(h)$	33,221	24,806	4,523	4,168

Tabela 15B E.M.Q. e E.M.A de dados com $n=25$, com 10% *outliers*, com tendência linear e sem *Cutoff*.

Est.	Todas as distâncias			
	EMQ		EMA	
	Média	Mediana	Média	Mediana
$\hat{\gamma}_M(h)$	31,023	24,906	4,339	4,121
$\hat{\gamma}_{CH}(h)$	19,850	13,000	3,233	2,878
$\hat{\gamma}_{Md}(h)$	29,488	15,191	3,515	3,014
$\hat{\gamma}_H(h)$	27,897	23,791	4,084	3,903
$\hat{\gamma}_G(h)$	78,365	51,645	5,734	5,111
$\hat{\gamma}_{Pw}(h)$	29,020	22,394	4,233	3,892
$\hat{\gamma}_{N1}(h)$	16,281	11,821	3,254	2,988
$\hat{\gamma}_{N2}(h)$	25,171	21,119	3,906	3,707

Tabela 16B E.M.Q. e E.M.A de dados com $n=25$, com 10% *outliers*, com tendência linear e *Cutoff* de 50%.

Est.	<i>Cutoff</i> de 50%			
	EMQ		EMA	
	Média	Mediana	Média	Mediana
$\hat{\gamma}_M(h)$	27,789	21,943	4,226	4,058
$\hat{\gamma}_{CH}(h)$	12,213	8,001	2,603	2,250
$\hat{\gamma}_{Md}(h)$	10,732	6,509	2,252	1,945
$\hat{\gamma}_H(h)$	27,068	21,474	4,117	3,888
$\hat{\gamma}_G(h)$	28,090	19,360	3,867	3,470
$\hat{\gamma}_{Pw}(h)$	26,053	20,823	4,092	3,944
$\hat{\gamma}_{N1}(h)$	21,420	16,862	3,829	3,717
$\hat{\gamma}_{N2}(h)$	32,677	27,688	4,608	4,533

Tabela 17B E.M.Q. e E.M.A de dados com $n=25$, sem *outliers*, com tendência quadrática e sem *Cutoff*.

Est.	Todas as distâncias			
	EMQ		EMA	
	Média	Mediana	Média	Mediana
$\hat{\gamma}_M(h)$	81,377	41,583	6,288	5,327
$\hat{\gamma}_{CH}(h)$	142,825	72,367	8,006	6,737
$\hat{\gamma}_{Md}(h)$	285,360	148,343	10,997	9,487
$\hat{\gamma}_H(h)$	67,685	30,851	5,401	4,364
$\hat{\gamma}_G(h)$	329,194	184,561	11,231	9,403
$\hat{\gamma}_{Pw}(h)$	15167,783	21,616	14,674	3,953
$\hat{\gamma}_{N1}(h)$	17,232	6,394	3,005	2,236
$\hat{\gamma}_{N2}(h)$	54,193	24,962	5,390	4,423

Tabela 18B E.M.Q. e E.M.A de dados com n=25, sem *outliers*, com tendência quadrática e *Cutoff* de 50%.

Est.	<i>Cutoff</i> de 50%			
	EMQ		EMA	
	Média	Mediana	Média	Mediana
$\hat{\gamma}_M(h)$	33,968	17,439	4,074	3,435
$\hat{\gamma}_{CH}(h)$	51,124	30,618	4,937	4,334
$\hat{\gamma}_{Md}(h)$	88,817	56,066	6,343	5,736
$\hat{\gamma}_H(h)$	25,084	12,591	3,334	2,739
$\hat{\gamma}_G(h)$	68,416	41,064	5,572	4,859
$\hat{\gamma}_{Pw}(h)$	17576,433	10,781	17,522	2,667
$\hat{\gamma}_{N1}(h)$	9,173	3,831	2,168	1,739
$\hat{\gamma}_{N2}(h)$	37,950	19,493	4,507	3,872

Tabela 19B E.M.Q. e E.M.A de dados com n=25, com 1% *outliers*, com tendência quadrática e sem *Cutoff*.

Est.	Todas as distâncias			
	EMQ		EMA	
	Média	Mediana	Média	Mediana
$\hat{\gamma}_M(h)$	86,519	56,831	7,130	6,450
$\hat{\gamma}_{CH}(h)$	103,559	53,702	7,148	5,877
$\hat{\gamma}_{Md}(h)$	162,283	87,864	8,505	7,150
$\hat{\gamma}_H(h)$	74,141	50,482	6,478	5,911
$\hat{\gamma}_G(h)$	329,542	199,433	11,705	10,462
$\hat{\gamma}_{Pw}(h)$	813230,981	540,156	108,897	11,955
$\hat{\gamma}_{N1}(h)$	25,942	18,442	4,236	4,035
$\hat{\gamma}_{N2}(h)$	58,334	40,829	6,242	5,899

Tabela 20B E.M.Q. e E.M.A de dados com n=25, com 1% *outliers*, com tendência quadrática e *Cutoff* de 50%.

Est.	<i>Cutoff</i> de 50%			
	EMQ		EMA	
	Média	Mediana	Média	Mediana
$\hat{\gamma}_M(h)$	45,147	35,524	5,377	5,207
$\hat{\gamma}_{CH}(h)$	42,474	31,156	4,859	4,664
$\hat{\gamma}_{Md}(h)$	55,160	40,697	5,238	4,905
$\hat{\gamma}_H(h)$	39,162	31,137	4,931	4,745
$\hat{\gamma}_G(h)$	74,905	53,823	6,292	5,799
$\hat{\gamma}_{Pw}(h)$	340606,709	233,819	89,108	8,876
$\hat{\gamma}_{N1}(h)$	22,566	17,259	3,964	3,796
$\hat{\gamma}_{N2}(h)$	51,634	40,394	5,957	5,747

Tabela 21B E.M.Q. e E.M.A de dados com n=25, com 5% *outliers*, com tendência quadrática e sem *Cutoff*.

Est.	Todas as distâncias			
	EMQ		EMA	
	Média	Mediana	Média	Mediana
$\hat{\gamma}_M(h)$	75,775	52,413	6,774	6,165
$\hat{\gamma}_{CH}(h)$	71,464	43,076	6,039	5,372
$\hat{\gamma}_{Md}(h)$	113,605	67,086	6,894	6,036
$\hat{\gamma}_H(h)$	70,785	49,677	6,382	5,865
$\hat{\gamma}_G(h)$	332,494	221,142	11,842	10,661
$\hat{\gamma}_{P_w}(h)$	927219,284	697,514	114,917	12,966
$\hat{\gamma}_{N_1}(h)$	25,614	17,158	4,250	3,898
$\hat{\gamma}_{N_2}(h)$	50,946	37,681	5,913	5,635

Tabela 22B E.M.Q. e E.M.A de dados com n=25, com 5% *outliers*, com tendência quadrática e *Cutoff* de 50%.

Est.	<i>Cutoff</i> de 50%			
	EMQ		EMA	
	Média	Mediana	Média	Mediana
$\hat{\gamma}_M(h)$	42,479	32,718	5,272	4,944
$\hat{\gamma}_{CH}(h)$	32,230	22,449	4,290	4,022
$\hat{\gamma}_{Md}(h)$	37,683	24,610	4,310	3,801
$\hat{\gamma}_H(h)$	39,598	31,617	5,018	4,781
$\hat{\gamma}_G(h)$	69,358	52,384	6,192	5,955
$\hat{\gamma}_{P_w}(h)$	198763,584	252,002	70,441	9,560
$\hat{\gamma}_{N_1}(h)$	24,719	19,186	4,175	4,020
$\hat{\gamma}_{N_2}(h)$	49,392	38,208	5,848	5,522

Tabela 23B E.M.Q. e E.M.A de dados com n=25, com 10% *outliers*, com tendência quadrática e sem *Cutoff*.

Est.	Todas as distâncias			
	EMQ		EMA	
	Média	Mediana	Média	Mediana
$\hat{\gamma}_M(h)$	120,381	95,058	9,027	8,642
$\hat{\gamma}_{CH}(h)$	96,973	71,815	7,420	6,912
$\hat{\gamma}_{Md}(h)$	140,605	98,214	7,868	7,291
$\hat{\gamma}_H(h)$	115,673	93,165	8,726	8,192
$\hat{\gamma}_G(h)$	409,843	320,918	14,041	13,655
$\hat{\gamma}_{P_w}(h)$	1635305,160	1859,067	157,975	20,019
$\hat{\gamma}_{N_1}(h)$	59,465	49,936	6,874	6,690
$\hat{\gamma}_{N_2}(h)$	92,420	77,691	8,373	7,988

Tabela 24B E.M.Q. e E.M.A de dados com $n=25$, com 10% *outliers*, com tendência quadrática e *Cutoff* de 50%.

Est.	<i>Cutoff</i> de 50%			
	EMQ		EMA	
	Média	Mediana	Média	Mediana
$\hat{\gamma}_M(h)$	83,278	71,388	7,726	7,548
$\hat{\gamma}_{CH}(h)$	51,181	40,939	5,721	5,474
$\hat{\gamma}_{Md}(h)$	49,580	32,961	4,957	4,540
$\hat{\gamma}_H(h)$	81,849	70,386	7,600	7,455
$\hat{\gamma}_G(h)$	113,432	87,080	8,270	7,925
$\hat{\gamma}_{Pw}(h)$	566978,601	920,898	117,144	16,800
$\hat{\gamma}_{N1}(h)$	61,682	53,065	6,902	6,681
$\hat{\gamma}_{N2}(h)$	95,722	84,249	8,442	8,247

Resultados das comparações em *grides* 10x10

Tabela 25B E.M.Q. e E.M.A de dados com $n=100$, sem *outliers*, sem tendência e sem *Cutoff*.

Est.	Todas as distâncias			
	EMQ		EMA	
	Média	Mediana	Média	Mediana
$\hat{\gamma}_M(h)$	12,376	8,497	2,659	2,393
$\hat{\gamma}_{CH}(h)$	20,627	11,732	3,183	2,697
$\hat{\gamma}_{Md}(h)$	36,285	18,207	3,972	3,248
$\hat{\gamma}_H(h)$	11,580	9,556	2,699	2,501
$\hat{\gamma}_G(h)$	18,916	12,128	2,990	2,713
$\hat{\gamma}_{Pw}(h)$	10,426	8,064	2,500	2,327
$\hat{\gamma}_{N1}(h)$	6,073	4,164	2,058	1,864
$\hat{\gamma}_{N2}(h)$	8,235	5,263	2,252	1,951

Tabela 26B E.M.Q. e E.M.A de dados com n=100, sem *outliers*, sem tendência e *Cutoff* de 50%.

Est.	<i>Cutoff</i> de 50%			
	EMQ		EMA	
	Média	Mediana	Média	Mediana
$\hat{\gamma}_M(h)$	7,237	4,341	2,023	1,760
$\hat{\gamma}_{CH}(h)$	10,920	5,565	2,355	1,947
$\hat{\gamma}_{Md}(h)$	16,695	7,790	2,797	2,225
$\hat{\gamma}_H(h)$	6,348	4,378	1,952	1,719
$\hat{\gamma}_G(h)$	7,873	5,212	2,115	1,870
$\hat{\gamma}_{Pw}(h)$	5,745	3,990	1,849	1,681
$\hat{\gamma}_{N1}(h)$	5,025	3,678	1,834	1,728
$\hat{\gamma}_{N2}(h)$	7,154	4,523	2,020	1,768

Tabela 27B E.M.Q. e E.M.A de dados com n=100, 1% *outliers*, sem tendência e sem *Cutoff*.

Est.	Todas as distâncias			
	EMQ		EMA	
	Média	Mediana	Média	Mediana
$\hat{\gamma}_M(h)$	8,682	7,153	2,298	2,184
$\hat{\gamma}_{CH}(h)$	11,684	8,781	2,556	2,376
$\hat{\gamma}_{Md}(h)$	18,510	11,846	2,993	2,673
$\hat{\gamma}_H(h)$	9,743	9,004	2,530	2,475
$\hat{\gamma}_G(h)$	14,071	10,546	2,660	2,517
$\hat{\gamma}_{Pw}(h)$	9,041	8,175	2,404	2,330
$\hat{\gamma}_{N1}(h)$	5,301	4,120	1,989	1,927
$\hat{\gamma}_{N2}(h)$	5,686	4,462	1,972	1,868

Tabela 28B E.M.Q. e E.M.A de dados com com n=100, com 1% *outliers*, sem tendência e *Cutoff* de 50%.

Est.	<i>Cutoff</i> de 50%			
	EMQ		EMA	
	Média	Mediana	Média	Mediana
$\hat{\gamma}_M(h)$	4,677	3,460	1,697	1,538
$\hat{\gamma}_{CH}(h)$	5,540	3,871	1,809	1,636
$\hat{\gamma}_{Md}(h)$	7,129	4,909	2,015	1,772
$\hat{\gamma}_H(h)$	5,210	4,225	1,817	1,714
$\hat{\gamma}_G(h)$	5,598	4,263	1,845	1,689
$\hat{\gamma}_{Pw}(h)$	4,613	3,839	1,724	1,656
$\hat{\gamma}_{N1}(h)$	4,646	3,847	1,808	1,791
$\hat{\gamma}_{N2}(h)$	4,768	3,451	1,742	1,586

Tabela 29B E.M.Q. e E.M.A de dados com $n=100$, com 5% *outliers*, sem tendência e sem *Cutoff*.

Est.	Todas as distâncias			
	EMQ		EMA	
	Média	Mediana	Média	Mediana
$\hat{\gamma}_M(h)$	8,327	6,421	2,191	2,049
$\hat{\gamma}_{CH}(h)$	10,965	8,466	2,440	2,292
$\hat{\gamma}_{Md}(h)$	18,269	12,313	2,944	2,779
$\hat{\gamma}_H(h)$	7,712	7,040	2,193	2,112
$\hat{\gamma}_G(h)$	17,422	11,123	2,690	2,563
$\hat{\gamma}_{Pw}(h)$	7,881	6,469	2,172	2,079
$\hat{\gamma}_{N1}(h)$	3,410	2,718	1,591	1,549
$\hat{\gamma}_{N2}(h)$	4,508	3,436	1,754	1,626

Tabela 30B E.M.Q. e E.M.A de dados com $n=100$, com 5% *outliers*, sem tendência e *Cutoff* de 50%.

Est.	<i>Cutoff</i> de 50%			
	EMQ		EMA	
	Média	Mediana	Média	Mediana
$\hat{\gamma}_M(h)$	4,090	3,229	1,655	1,544
$\hat{\gamma}_{CH}(h)$	4,510	3,832	1,706	1,631
$\hat{\gamma}_{Md}(h)$	6,047	5,204	1,937	1,836
$\hat{\gamma}_H(h)$	4,056	3,524	1,662	1,576
$\hat{\gamma}_G(h)$	4,577	3,789	1,720	1,605
$\hat{\gamma}_{Pw}(h)$	3,764	3,236	1,600	1,522
$\hat{\gamma}_{N1}(h)$	3,006	2,596	1,454	1,443
$\hat{\gamma}_{N2}(h)$	4,429	3,632	1,700	1,578

Tabela 31B E.M.Q. e E.M.A de dados com $n=100$, com 10% *outliers*, sem tendência e sem *Cutoff*.

Est.	Todas as distâncias			
	EMQ		EMA	
	Média	Mediana	Média	Mediana
$\hat{\gamma}_M(h)$	34,571	28,202	4,863	4,579
$\hat{\gamma}_{CH}(h)$	19,871	13,021	3,202	2,808
$\hat{\gamma}_{Md}(h)$	25,388	14,154	3,205	2,841
$\hat{\gamma}_H(h)$	29,736	25,426	4,493	4,332
$\hat{\gamma}_G(h)$	44,537	30,015	4,037	3,658
$\hat{\gamma}_{Pw}(h)$	31,969	26,396	4,610	4,408
$\hat{\gamma}_{N1}(h)$	18,809	16,336	3,973	3,884
$\hat{\gamma}_{N2}(h)$	25,112	22,245	4,441	4,314

Tabela 32B E.M.Q. e E.M.A de dados com $n=100$, com 10% *outliers*, sem tendência e *Cutoff* de 50%.

Est.	<i>Cutoff</i> de 50%			
	EMQ		EMA	
	Média	Mediana	Média	Mediana
$\hat{\gamma}_M(h)$	28,267	26,055	4,735	4,713
$\hat{\gamma}_{CH}(h)$	9,764	7,487	2,535	2,328
$\hat{\gamma}_{Md}(h)$	6,525	5,138	1,963	1,838
$\hat{\gamma}_H(h)$	26,667	24,967	4,560	4,509
$\hat{\gamma}_G(h)$	10,526	7,743	2,601	2,372
$\hat{\gamma}_{Pw}(h)$	25,263	22,701	4,432	4,290
$\hat{\gamma}_{N1}(h)$	21,346	19,313	4,193	4,116
$\hat{\gamma}_{N2}(h)$	30,288	27,994	4,866	4,856

Tabela 33B E.M.Q. e E.M.A de dados com $n=100$, sem *outliers*, com tendência linear e sem *Cutoff*.

Est.	Todas as distâncias			
	EMQ		EMA	
	Média	Mediana	Média	Mediana
$\hat{\gamma}_M(h)$	11,573	8,355	2,613	2,391
$\hat{\gamma}_{CH}(h)$	18,406	11,255	3,080	2,703
$\hat{\gamma}_{Md}(h)$	31,695	15,344	3,769	3,093
$\hat{\gamma}_H(h)$	10,873	9,547	2,653	2,551
$\hat{\gamma}_G(h)$	16,700	11,074	2,824	2,609
$\hat{\gamma}_{Pw}(h)$	10,352	8,729	2,514	2,408
$\hat{\gamma}_{N1}(h)$	6,269	4,619	2,110	2,006
$\hat{\gamma}_{N2}(h)$	7,643	5,656	2,214	2,027

Tabela 34B E.M.Q. e E.M.A de dados com $n=100$, sem *outliers*, com tendência linear e *Cutoff* de 50%.

Est.	<i>Cutoff</i> de 50%			
	EMQ		EMA	
	Média	Mediana	Média	Mediana
$\hat{\gamma}_M(h)$	6,646	4,356	1,948	1,711
$\hat{\gamma}_{CH}(h)$	9,422	5,372	2,210	1,868
$\hat{\gamma}_{Md}(h)$	14,076	6,590	2,589	2,082
$\hat{\gamma}_H(h)$	6,149	4,425	1,933	1,763
$\hat{\gamma}_G(h)$	7,010	4,858	2,006	1,787
$\hat{\gamma}_{Pw}(h)$	5,585	4,031	1,827	1,695
$\hat{\gamma}_{N1}(h)$	5,203	4,351	1,897	1,883
$\hat{\gamma}_{N2}(h)$	6,619	4,343	1,966	1,734

Tabela 35B E.M.Q. e E.M.A de dados com $n=100$, com 1% *outliers*, com tendência linear e sem *Cutoff*.

Est.	Todas as distâncias			
	EMQ		EMA	
	Média	Mediana	Média	Mediana
$\hat{\gamma}_M(h)$	11,446	8,545	2,558	2,326
$\hat{\gamma}_{CH}(h)$	15,708	10,782	2,859	2,587
$\hat{\gamma}_{Md}(h)$	26,391	15,784	3,442	3,052
$\hat{\gamma}_H(h)$	9,361	7,939	2,409	2,292
$\hat{\gamma}_G(h)$	19,741	12,581	2,933	2,705
$\hat{\gamma}_{Pw}(h)$	8,806	7,464	2,297	2,202
$\hat{\gamma}_{N1}(h)$	4,365	2,835	1,736	1,515
$\hat{\gamma}_{N2}(h)$	6,388	4,600	2,040	1,816

Tabela 36B E.M.Q. e E.M.A de dados com $n=100$, com 1% *outliers*, com tendência linear e *Cutoff* de 50%.

Est.	<i>Cutoff</i> de 50%			
	EMQ		EMA	
	Média	Mediana	Média	Mediana
$\hat{\gamma}_M(h)$	5,758	4,250	1,874	1,735
$\hat{\gamma}_{CH}(h)$	7,054	4,887	1,998	1,823
$\hat{\gamma}_{Md}(h)$	9,811	6,008	2,269	1,946
$\hat{\gamma}_H(h)$	5,250	4,056	1,825	1,673
$\hat{\gamma}_G(h)$	6,275	4,582	1,931	1,768
$\hat{\gamma}_{Pw}(h)$	4,295	2,945	1,643	1,449
$\hat{\gamma}_{N1}(h)$	3,584	2,348	1,541	1,408
$\hat{\gamma}_{N2}(h)$	5,829	4,178	1,908	1,709

Tabela 37B E.M.Q. e E.M.A de dados com $n=100$, com 5% *outliers*, com tendência linear e sem *Cutoff*.

Est.	Todas as distâncias			
	EMQ		EMA	
	Média	Mediana	Média	Mediana
$\hat{\gamma}_M(h)$	10,891	7,562	2,459	2,179
$\hat{\gamma}_{CH}(h)$	11,042	8,193	2,441	2,260
$\hat{\gamma}_{Md}(h)$	17,177	12,071	2,892	2,717
$\hat{\gamma}_H(h)$	9,266	7,610	2,314	2,186
$\hat{\gamma}_G(h)$	19,489	11,601	2,742	2,515
$\hat{\gamma}_{Pw}(h)$	10,955	7,949	2,458	2,198
$\hat{\gamma}_{N1}(h)$	2,883	1,784	1,356	1,134
$\hat{\gamma}_{N2}(h)$	5,341	3,402	1,781	1,476

Tabela 38B E.M.Q. e E.M.A de dados com $n=100$, com 5% *outliers*, com tendência linear e *Cutoff* de 50%.

Est.	<i>Cutoff</i> de 50%			
	EMQ		EMA	
	Média	Mediana	Média	Mediana
$\hat{\gamma}_M(h)$	5,650	3,762	1,894	1,618
$\hat{\gamma}_{CH}(h)$	4,446	3,304	1,658	1,495
$\hat{\gamma}_{Md}(h)$	5,930	4,834	1,889	1,755
$\hat{\gamma}_H(h)$	5,108	3,642	1,805	1,575
$\hat{\gamma}_G(h)$	4,567	3,602	1,683	1,554
$\hat{\gamma}_{Pw}(h)$	5,462	3,732	1,864	1,614
$\hat{\gamma}_{N1}(h)$	2,898	2,140	1,343	1,179
$\hat{\gamma}_{N2}(h)$	6,112	4,360	1,915	1,618

Tabela 39B E.M.Q. e E.M.A de dados com $n=100$, com 10% *outliers*, com tendência linear e sem *Cutoff*.

Est.	Todas as distâncias			
	EMQ		EMA	
	Média	Mediana	Média	Mediana
$\hat{\gamma}_M(h)$	48,840	42,377	5,808	5,592
$\hat{\gamma}_{CH}(h)$	26,367	16,203	3,616	3,033
$\hat{\gamma}_{Md}(h)$	31,080	17,736	3,448	3,002
$\hat{\gamma}_H(h)$	41,870	35,583	5,368	5,171
$\hat{\gamma}_G(h)$	56,389	35,053	4,449	3,778
$\hat{\gamma}_{Pw}(h)$	45,866	34,880	5,525	5,128
$\hat{\gamma}_{N1}(h)$	26,140	22,533	4,685	4,587
$\hat{\gamma}_{N2}(h)$	34,796	30,597	5,274	5,224

Tabela 40B E.M.Q. e E.M.A de dados com $n=100$, com 10% *outliers*, com tendência linear e *Cutoff* de 50%.

Est.	<i>Cutoff</i> de 50%			
	EMQ		EMA	
	Média	Mediana	Média	Mediana
$\hat{\gamma}_M(h)$	37,296	32,912	5,471	5,361
$\hat{\gamma}_{CH}(h)$	11,436	7,762	2,710	2,390
$\hat{\gamma}_{Md}(h)$	7,039	4,841	2,013	1,763
$\hat{\gamma}_H(h)$	35,105	30,984	5,269	5,202
$\hat{\gamma}_G(h)$	11,660	7,780	2,689	2,314
$\hat{\gamma}_{Pw}(h)$	34,102	30,658	5,184	5,154
$\hat{\gamma}_{N1}(h)$	27,880	25,093	4,800	4,758
$\hat{\gamma}_{N2}(h)$	39,517	35,253	5,613	5,500

Tabela 41B E.M.Q. e E.M.A de dados com n=100, sem *outliers*, com tendência quadrática e sem *Cutoff*.

Est.	Todas as distâncias			
	EMQ		EMA	
	Média	Mediana	Média	Mediana
$\hat{\gamma}_M(h)$	12752,621	12354,197	101,954	101,053
$\hat{\gamma}_{CH}(h)$	20877,824	20123,347	127,007	126,802
$\hat{\gamma}_{Md}(h)$	35294,022	33910,680	158,302	157,610
$\hat{\gamma}_H(h)$	3751,976	3582,363	56,381	55,724
$\hat{\gamma}_G(h)$	5018,018	4766,611	63,525	62,851
$\hat{\gamma}_{Pw}(h)$	27020,473	869,320	27,320	23,182
$\hat{\gamma}_{N1}(h)$	4647,962	4518,155	61,966	61,725
$\hat{\gamma}_{N2}(h)$	8938,227	8743,356	87,265	87,026

Tabela 42B E.M.Q. e E.M.A de dados com n=100, sem *outliers*, com tendência quadrática e *Cutoff* de 50%.

Est.	<i>Cutoff</i> de 50%			
	EMQ		EMA	
	Média	Mediana	Média	Mediana
$\hat{\gamma}_M(h)$	6155,941	6004,099	68,692	68,727
$\hat{\gamma}_{CH}(h)$	8044,969	7787,765	77,601	77,463
$\hat{\gamma}_{Md}(h)$	10422,838	9911,711	86,987	86,333
$\hat{\gamma}_H(h)$	2960,246	2769,067	48,275	47,376
$\hat{\gamma}_G(h)$	3285,800	3028,324	50,701	49,585
$\hat{\gamma}_{Pw}(h)$	51149,725	843,834	31,373	24,729
$\hat{\gamma}_{N1}(h)$	2205,319	2140,376	41,265	41,153
$\hat{\gamma}_{N2}(h)$	6441,946	6278,662	70,509	70,430

Tabela 43B E.M.Q. e E.M.A de dados com n=100, com 1% *outliers*, com tendência quadrática e sem *Cutoff*.

Est.	Todas as distâncias			
	EMQ		EMA	
	Média	Mediana	Média	Mediana
$\hat{\gamma}_M(h)$	12253,721	11832,103	99,942	98,899
$\hat{\gamma}_{CH}(h)$	18909,163	17844,726	119,677	117,862
$\hat{\gamma}_{Md}(h)$	29278,492	27706,958	142,379	141,103
$\hat{\gamma}_H(h)$	4312,845	4155,070	60,258	59,983
$\hat{\gamma}_G(h)$	5185,718	5058,876	64,139	64,835
$\hat{\gamma}_{Pw}(h)$	5979985612,290	1518142,000	3550,548	273,807
$\hat{\gamma}_{N1}(h)$	4538,691	4419,813	62,434	62,295
$\hat{\gamma}_{N2}(h)$	8271,181	8007,052	84,913	84,126

Tabela 44B E.M.Q. e E.M.A de dados com n=100, com 1% *outliers*, com tendência quadrática e *Cutoff* de 50%.

Est.	<i>Cutoff</i> de 50%			
	EMQ		EMA	
	Média	Mediana	Média	Mediana
$\hat{\gamma}_M(h)$	5729,691	5576,571	67,614	67,688
$\hat{\gamma}_{CH}(h)$	6630,193	6337,523	71,221	70,691
$\hat{\gamma}_{Md}(h)$	7866,312	7417,404	76,151	74,783
$\hat{\gamma}_H(h)$	3220,471	3125,754	51,418	51,528
$\hat{\gamma}_G(h)$	3296,644	3188,482	50,907	50,705
$\hat{\gamma}_{Pw}(h)$	699378547,565	741494,400	2135,662	281,840
$\hat{\gamma}_{N1}(h)$	2366,655	2334,876	44,219	44,153
$\hat{\gamma}_{N2}(h)$	6001,523	5819,620	69,716	69,569

Tabela 45B E.M.Q. e E.M.A de dados com n=100, com 5% *outliers*, com tendência quadrática e sem *Cutoff*.

Est.	Todas as distâncias			
	EMQ		EMA	
	Média	Mediana	Média	Mediana
$\hat{\gamma}_M(h)$	11328,813	11056,734	95,779	95,664
$\hat{\gamma}_{CH}(h)$	15563,142	14874,339	105,356	104,897
$\hat{\gamma}_{Md}(h)$	21728,176	20357,339	117,289	116,707
$\hat{\gamma}_H(h)$	5163,559	4877,448	65,588	65,020
$\hat{\gamma}_G(h)$	5643,242	5286,952	65,092	63,915
$\hat{\gamma}_{Pw}(h)$	16708301254,522	9649753,000	6601,077	708,293
$\hat{\gamma}_{N1}(h)$	4208,219	4092,111	61,732	61,207
$\hat{\gamma}_{N2}(h)$	7070,779	6941,425	79,741	79,455

Tabela 46B E.M.Q. e E.M.A de dados com n=100, com 5% *outliers*, com tendência quadrática e *Cutoff* de 50%.

Est.	<i>Cutoff</i> de 50%			
	EMQ		EMA	
	Média	Mediana	Média	Mediana
$\hat{\gamma}_M(h)$	4847,001	4713,300	64,032	63,764
$\hat{\gamma}_{CH}(h)$	4368,924	4197,242	58,573	57,806
$\hat{\gamma}_{Md}(h)$	4292,386	4138,309	56,150	55,458
$\hat{\gamma}_H(h)$	3239,329	3169,410	52,999	52,761
$\hat{\gamma}_G(h)$	2728,261	2594,766	46,330	46,181
$\hat{\gamma}_{Pw}(h)$	2919581716,234	3397678,000	4328,116	602,389
$\hat{\gamma}_{N1}(h)$	2545,646	2471,558	47,746	47,395
$\hat{\gamma}_{N2}(h)$	5099,805	4994,229	66,462	66,204

Tabela 47B E.M.Q. e E.M.A de dados com n=100, com 10% *outliers* e com tendência quadrática e sem *Cutoff*.

Est.	Todas as distâncias			
	EMQ		EMA	
	Média	Mediana	Média	Mediana
$\hat{\gamma}_M(h)$	26532,200	26235,535	154,215	153,529
$\hat{\gamma}_{CH}(h)$	29009,381	27762,788	151,699	151,701
$\hat{\gamma}_{Md}(h)$	33115,605	31173,429	152,110	150,834
$\hat{\gamma}_H(h)$	18528,815	18159,621	128,638	128,330
$\hat{\gamma}_G(h)$	14874,375	13881,256	105,640	104,768
$\hat{\gamma}_{Pw}(h)$	203795388479,3	173732800	27142,8	2663,5
$\hat{\gamma}_{N1}(h)$	16921,133	16761,361	127,586	127,155
$\hat{\gamma}_{N2}(h)$	21415,827	21202,400	143,005	142,733

Tabela 48B E.M.Q. e E.M.A de dados com n=100, com 10% *outliers*, com tendência quadrática e *Cutoff* de 50%.

Est.	<i>Cutoff</i> de 50%			
	EMQ		EMA	
	Média	Mediana	Média	Mediana
$\hat{\gamma}_M(h)$	17927,362	17834,258	128,891	128,707
$\hat{\gamma}_{CH}(h)$	12268,569	12158,987	101,753	101,638
$\hat{\gamma}_{Md}(h)$	9647,193	9221,670	85,325	84,186
$\hat{\gamma}_H(h)$	15329,933	15230,849	119,559	118,995
$\hat{\gamma}_G(h)$	8038,352	7676,635	80,925	80,015
$\hat{\gamma}_{Pw}(h)$	73390023054,4	52112630	21704,9	2123,0
$\hat{\gamma}_{N1}(h)$	13825,910	13557,960	114,341	113,574
$\hat{\gamma}_{N2}(h)$	18574,324	18453,766	131,805	131,508

Resultados das comparações em *grides* 20x20

Tabela 49B E.M.Q. e E.M.A de dados com n=400, sem *outliers* e sem tendência e sem *Cutoff*.

Est.	Todas as distâncias			
	EMQ		EMA	
	Média	Mediana	Média	Mediana
$\hat{\gamma}_M(h)$	6,419	4,853	1,868	1,660
$\hat{\gamma}_{CH}(h)$	9,909	6,442	2,199	1,899
$\hat{\gamma}_{Md}(h)$	16,070	10,272	2,651	2,305
$\hat{\gamma}_H(h)$	5,548	4,544	1,811	1,675
$\hat{\gamma}_G(h)$	7,983	6,018	1,966	1,770
$\hat{\gamma}_{Pw}(h)$	5,520	4,468	1,760	1,607
$\hat{\gamma}_{N1}(h)$	2,455	1,609	1,303	1,152
$\hat{\gamma}_{N2}(h)$	3,390	1,879	1,420	1,154

Tabela 50B E.M.Q. e E.M.A de dados com n=400, sem *outliers*, sem tendência e *Cutoff* de 50%.

Est.	<i>Cutoff</i> de 50%			
	EMQ		EMA	
	Média	Mediana	Média	Mediana
$\hat{\gamma}_M(h)$	3,691	1,930	1,459	1,171
$\hat{\gamma}_{CH}(h)$	5,217	2,609	1,682	1,309
$\hat{\gamma}_{Md}(h)$	7,162	3,611	1,937	1,530
$\hat{\gamma}_H(h)$	3,267	2,067	1,402	1,226
$\hat{\gamma}_G(h)$	4,023	2,285	1,519	1,266
$\hat{\gamma}_{Pw}(h)$	2,737	1,820	1,303	1,119
$\hat{\gamma}_{N1}(h)$	2,491	1,877	1,334	1,237
$\hat{\gamma}_{N2}(h)$	3,616	1,950	1,444	1,145

Tabela 51B E.M.Q. e E.M.A de dados com n=400, com 1% *outliers* e sem tendência e sem *Cutoff*.

Est.	Todas as distâncias			
	EMQ		EMA	
	Média	Mediana	Média	Mediana
$\hat{\gamma}_M(h)$	6,973	5,185	1,952	1,744
$\hat{\gamma}_{CH}(h)$	9,247	6,241	2,134	1,909
$\hat{\gamma}_{Md}(h)$	14,305	9,025	2,502	2,244
$\hat{\gamma}_H(h)$	5,552	4,540	1,778	1,637
$\hat{\gamma}_G(h)$	8,546	6,362	1,964	1,815
$\hat{\gamma}_{Pw}(h)$	5,473	4,576	1,719	1,621
$\hat{\gamma}_{N1}(h)$	2,079	1,220	1,177	1,033
$\hat{\gamma}_{N2}(h)$	3,580	2,154	1,483	1,285

Tabela 52B E.M.Q. e E.M.A de dados com $n=400$, com 1% *outliers*, sem tendência e *Cutoff* de 50%.

Est.	<i>Cutoff</i> de 50%			
	EMQ		EMA	
	Média	Mediana	Média	Mediana
$\hat{\gamma}_M(h)$	4,026	2,532	1,589	1,358
$\hat{\gamma}_{CH}(h)$	4,471	2,956	1,622	1,410
$\hat{\gamma}_{Md}(h)$	5,471	3,620	1,760	1,559
$\hat{\gamma}_H(h)$	3,370	2,111	1,456	1,243
$\hat{\gamma}_G(h)$	3,532	2,361	1,462	1,287
$\hat{\gamma}_{Pw}(h)$	2,113	1,516	1,177	1,039
$\hat{\gamma}_{N1}(h)$	1,707	1,081	1,067	0,957
$\hat{\gamma}_{N2}(h)$	3,958	2,413	1,566	1,326

Tabela 53B E.M.Q. e E.M.A de dados com $n=400$, com 5% *outliers* e sem tendência e sem *Cutoff*.

Est.	Todas as distâncias			
	EMQ		EMA	
	Média	Mediana	Média	Mediana
$\hat{\gamma}_M(h)$	151,109	138,505	11,374	11,095
$\hat{\gamma}_{CH}(h)$	41,749	32,740	5,240	4,923
$\hat{\gamma}_{Md}(h)$	24,823	16,316	3,307	2,943
$\hat{\gamma}_H(h)$	138,437	127,778	10,888	10,606
$\hat{\gamma}_G(h)$	32,607	23,382	3,949	3,768
$\hat{\gamma}_{Pw}(h)$	218167023,301	283895,200	455,753	72,424
$\hat{\gamma}_{N1}(h)$	120,328	115,990	10,812	10,722
$\hat{\gamma}_{N2}(h)$	131,520	125,353	11,242	11,039

Tabela 54B E.M.Q. e E.M.A de dados com $n=400$, com 5% *outliers*, sem tendência e *Cutoff* de 50%.

Est.	<i>Cutoff</i> de 50%			
	EMQ		EMA	
	Média	Mediana	Média	Mediana
$\hat{\gamma}_M(h)$	133,253	129,622	11,263	11,224
$\hat{\gamma}_{CH}(h)$	28,917	24,927	4,861	4,766
$\hat{\gamma}_{Md}(h)$	10,680	7,142	2,520	2,214
$\hat{\gamma}_H(h)$	128,884	124,422	11,076	11,012
$\hat{\gamma}_G(h)$	17,078	12,799	3,488	3,288
$\hat{\gamma}_{Pw}(h)$	39865132,652	129763,050	318,223	65,019
$\hat{\gamma}_{N1}(h)$	115,302	110,111	10,552	10,411
$\hat{\gamma}_{N2}(h)$	133,804	130,094	11,306	11,270

Tabela 55B E.M.Q. e E.M.A de dados com n=400, com 10% *outliers* e sem tendência e sem *Cutoff*.

Est.	Todas as distâncias			
	EMQ		EMA	
	Média	Mediana	Média	Mediana
$\hat{\gamma}_M(h)$	3195,113	3108,967	54,540	54,321
$\hat{\gamma}_{CH}(h)$	382,867	355,846	18,024	17,746
$\hat{\gamma}_{Md}(h)$	61,017	43,449	5,616	5,319
$\hat{\gamma}_H(h)$	3129,772	3037,837	53,967	53,717
$\hat{\gamma}_G(h)$	172,469	124,740	10,106	9,621
$\hat{\gamma}_{Pw}(h)$	2336074792050,700	2773875000	44312,455	5973,769
$\hat{\gamma}_{N1}(h)$	2974,145	2914,761	54,161	53,805
$\hat{\gamma}_{N2}(h)$	3037,922	2977,390	54,635	54,306

Tabela 56B E.M.Q. e E.M.A de dados com n=400, com 10% *outliers*, sem tendência e *Cutoff* de 50%.

Est.	<i>Cutoff</i> de 50%			
	EMQ		EMA	
	Média	Mediana	Média	Mediana
$\hat{\gamma}_M(h)$	3025,717	2969,035	54,308	54,015
$\hat{\gamma}_{CH}(h)$	315,366	306,658	17,274	17,242
$\hat{\gamma}_{Md}(h)$	28,832	23,533	4,640	4,527
$\hat{\gamma}_H(h)$	3005,000	2946,183	54,117	53,770
$\hat{\gamma}_G(h)$	94,820	84,535	9,101	8,873
$\hat{\gamma}_{Pw}(h)$	484634206927,085	714920300	31332,121	4867,044
$\hat{\gamma}_{N1}(h)$	2941,304	2870,310	53,703	53,258
$\hat{\gamma}_{N2}(h)$	3045,335	2976,865	54,536	54,125

Tabela 57B E.M.Q. e E.M.A de dados com n=400, sem *outliers* e com tendência linear e sem *Cutoff*.

Est.	Todas as distâncias			
	EMQ		EMA	
	Média	Mediana	Média	Mediana
$\hat{\gamma}_M(h)$	7,289	4,747	1,953	1,658
$\hat{\gamma}_{CH}(h)$	11,370	6,750	2,324	1,988
$\hat{\gamma}_{Md}(h)$	18,749	10,426	2,821	2,325
$\hat{\gamma}_H(h)$	6,231	4,826	1,878	1,687
$\hat{\gamma}_G(h)$	9,917	6,325	2,096	1,829
$\hat{\gamma}_{Pw}(h)$	5,802	4,453	1,795	1,642
$\hat{\gamma}_{N1}(h)$	2,577	1,641	1,325	1,150
$\hat{\gamma}_{N2}(h)$	3,662	2,130	1,475	1,218

Tabela 58B E.M.Q. e E.M.A de dados com n=400, sem *outliers*, com tendência linear e *Cutoff* de 50%.

Est.	<i>Cutoff</i> de 50%			
	EMQ		EMA	
	Média	Mediana	Média	Mediana
$\hat{\gamma}_M(h)$	3,580	2,171	1,451	1,221
$\hat{\gamma}_{CH}(h)$	4,970	2,680	1,668	1,346
$\hat{\gamma}_{Md}(h)$	6,820	3,616	1,922	1,554
$\hat{\gamma}_H(h)$	3,241	2,013	1,404	1,183
$\hat{\gamma}_G(h)$	3,932	2,318	1,513	1,233
$\hat{\gamma}_{Pw}(h)$	2,698	1,899	1,301	1,138
$\hat{\gamma}_{N1}(h)$	2,568	1,863	1,349	1,215
$\hat{\gamma}_{N2}(h)$	3,555	2,223	1,444	1,236

Tabela 59B E.M.Q. e E.M.A de dados com n=400, com 1% *outliers* e com tendência linear e sem *Cutoff*.

Est.	Todas as distâncias			
	EMQ		EMA	
	Média	Mediana	Média	Mediana
$\hat{\gamma}_M(h)$	9,432	6,620	2,252	1,954
$\hat{\gamma}_{CH}(h)$	11,618	7,570	2,347	2,040
$\hat{\gamma}_{Md}(h)$	17,059	10,775	2,692	2,325
$\hat{\gamma}_H(h)$	7,229	5,159	1,996	1,756
$\hat{\gamma}_G(h)$	10,991	7,153	2,143	1,879
$\hat{\gamma}_{Pw}(h)$	8,377	5,836	2,080	1,828
$\hat{\gamma}_{N1}(h)$	2,716	1,460	1,316	1,135
$\hat{\gamma}_{N2}(h)$	4,968	2,819	1,746	1,478

Tabela 60B E.M.Q. e E.M.A de dados com n=400, com 1% *outliers*, com tendência linear e *Cutoff* de 50%.

Est.	<i>Cutoff</i> de 50%			
	EMQ		EMA	
	Média	Mediana	Média	Mediana
$\hat{\gamma}_M(h)$	5,043	2,959	1,775	1,513
$\hat{\gamma}_{CH}(h)$	5,038	2,782	1,704	1,415
$\hat{\gamma}_{Md}(h)$	5,974	3,396	1,815	1,500
$\hat{\gamma}_H(h)$	4,224	2,417	1,624	1,349
$\hat{\gamma}_G(h)$	3,965	2,350	1,532	1,300
$\hat{\gamma}_{Pw}(h)$	3,699	2,129	1,511	1,243
$\hat{\gamma}_{N1}(h)$	2,012	1,131	1,127	0,980
$\hat{\gamma}_{N2}(h)$	4,970	2,953	1,746	1,466

Tabela 61B E.M.Q. e E.M.A de dados com $n=400$, com 5% *outliers* e com tendência linear e sem *Cutoff*.

Est.	Todas as distâncias			
	EMQ		EMA	
	Média	Mediana	Média	Mediana
$\hat{\gamma}_M(h)$	1293,623	1195,405	33,721	32,951
$\hat{\gamma}_{CH}(h)$	124,741	104,228	9,550	9,202
$\hat{\gamma}_{Md}(h)$	37,618	23,774	4,171	3,720
$\hat{\gamma}_H(h)$	1249,898	1147,638	33,148	32,567
$\hat{\gamma}_G(h)$	55,942	38,160	5,267	4,816
$\hat{\gamma}_{Pw}(h)$	1569596547765,0	652968450,0	30922,249	3187,467
$\hat{\gamma}_{N_1}(h)$	1090,806	1091,717	32,673	32,938
$\hat{\gamma}_{N_2}(h)$	1143,756	1135,604	33,353	33,504

Tabela 62B E.M.Q. e E.M.A de dados com $n=400$, com 5% *outliers*, com tendência linear e *Cutoff* de 50%.

Est.	<i>Cutoff</i> de 50%			
	EMQ		EMA	
	Média	Mediana	Média	Mediana
$\hat{\gamma}_M(h)$	1128,634	1127,322	32,988	33,214
$\hat{\gamma}_{CH}(h)$	86,604	78,783	8,744	8,637
$\hat{\gamma}_{Md}(h)$	15,969	11,256	3,137	2,917
$\hat{\gamma}_H(h)$	1113,157	1116,906	32,755	33,062
$\hat{\gamma}_G(h)$	27,341	21,989	4,480	4,390
$\hat{\gamma}_{Pw}(h)$	264661786618,720	217383350,0	22535,006	2543,941
$\hat{\gamma}_{N_1}(h)$	1062,052	1064,662	32,149	32,442
$\hat{\gamma}_{N_2}(h)$	1135,116	1136,102	33,147	33,417

Tabela 63B E.M.Q. e E.M.A de dados com $n=400$, com 10% *outliers* e com tendência linear e sem *Cutoff*.

Est.	Todas as distâncias			
	EMQ		EMA	
	Média	Mediana	Média	Mediana
$\hat{\gamma}_M(h)$	192190,655	187758,885	421,593	420,450
$\hat{\gamma}_{CH}(h)$	4327,589	3838,045	59,620	58,473
$\hat{\gamma}_{Md}(h)$	142,707	71,334	7,603	6,853
$\hat{\gamma}_H(h)$	190917,245	186238,470	420,172	419,634
$\hat{\gamma}_G(h)$	1190,275	224,550	14,282	12,741
$\hat{\gamma}_{Pw}(h)$	47399901459839,2	61660940000,0	198283,6	30986,1
$\hat{\gamma}_{N_1}(h)$	177538,826	175148,020	417,886	417,209
$\hat{\gamma}_{N_2}(h)$	180142,635	177925,900	420,183	419,923

Tabela 64B E.M.Q. e E.M.A de dados com $n=400$, com 10% *outliers*, com tendência linear e *Cutoff* de 50%.

Est.	<i>Cutoff</i> de 50%			
	EMQ		EMA	
	Média	Mediana	Média	Mediana
$\hat{\gamma}_M(h)$	177555,892	174080,020	415,561	413,600
$\hat{\gamma}_{CH}(h)$	3257,172	3158,657	55,749	55,322
$\hat{\gamma}_{Md}(h)$	49,347	39,741	6,120	5,850
$\hat{\gamma}_H(h)$	177251,238	173894,755	415,206	413,561
$\hat{\gamma}_G(h)$	156,569	134,545	11,644	11,277
$\hat{\gamma}_{Pw}(h)$	8514805528516,1	27192500000,0	155203,4	27554,7
$\hat{\gamma}_{N1}(h)$	175945,732	171927,095	414,809	412,276
$\hat{\gamma}_{N2}(h)$	178762,526	175651,175	417,421	416,054

Tabela 65B E.M.Q. e E.M.A de dados com $n=400$, sem *outliers* e com tendência quadrática e sem *Cutoff*.

Est.	Todas as distâncias			
	EMQ		EMA	
	Média	Mediana	Média	Mediana
$\hat{\gamma}_M(h)$	21971230,204	21965027,000	4084,330	4085,617
$\hat{\gamma}_{CH}(h)$	35002575,344	34967737,500	5247,717	5248,128
$\hat{\gamma}_{Md}(h)$	85327168,411	85303558,000	7766,376	7775,019
$\hat{\gamma}_H(h)$	6068553,244	6061678,500	2212,808	2213,951
$\hat{\gamma}_G(h)$	2791871,913	2783452,000	1534,798	1533,756
$\hat{\gamma}_{Pw}(h)$	964698,032	703083,750	665,178	710,884
$\hat{\gamma}_{N1}(h)$	4768693,027	4763896,500	1973,949	1973,929
$\hat{\gamma}_{N2}(h)$	11017538,131	11014344,000	3033,248	3032,218

Tabela 66B E.M.Q. e E.M.A de dados com $n=400$, sem *outliers*, com tendência quadrática e *Cutoff* de 50%.

Est.	<i>Cutoff</i> de 50%			
	EMQ		EMA	
	Média	Mediana	Média	Mediana
$\hat{\gamma}_M(h)$	6163167,673	6158002,500	2175,026	2174,971
$\hat{\gamma}_{CH}(h)$	11564834,544	11572226,000	2920,304	2919,962
$\hat{\gamma}_{Md}(h)$	17040339,427	16941271,000	3503,883	3499,314
$\hat{\gamma}_H(h)$	2440133,698	2441052,000	1372,087	1372,200
$\hat{\gamma}_G(h)$	1745934,551	1752795,500	1187,790	1190,945
$\hat{\gamma}_{Pw}(h)$	654945,679	485025,600	527,777	596,550
$\hat{\gamma}_{N1}(h)$	1826785,942	1826813,000	1185,963	1186,492
$\hat{\gamma}_{N2}(h)$	6230585,184	6223135,500	2178,194	2178,681

Tabela 67B E.M.Q. e E.M.A de dados com n=400, com 1% *outliers* e com tendência quadrática e sem *Cutoff*.

Est.	Todas as distâncias			
	EMQ		EMA	
	Média	Mediana	Média	Mediana
$\hat{\gamma}_M(h)$	24206174,042	24094225,500	4361,051	4356,166
$\hat{\gamma}_{CH}(h)$	37775348,571	37673627,500	5468,885	5465,454
$\hat{\gamma}_{Md}(h)$	91153855,698	88893647,500	7984,789	7942,200
$\hat{\gamma}_H(h)$	7519634,684	7379807,500	2518,210	2505,699
$\hat{\gamma}_G(h)$	3347223,676	3276449,500	1665,124	1659,887
$\hat{\gamma}_{Pw}(h)$	462990227768639	245911600000	592961,9	59072,9
$\hat{\gamma}_{N1}(h)$	6055542,244	6049842,000	2283,755	2282,951
$\hat{\gamma}_{N2}(h)$	12785715,100	12752241,500	3321,205	3317,547

Tabela 68B E.M.Q. e E.M.A de dados com n=400, com 1% *outliers*, com tendência quadrática e *Cutoff* de 50%.

Est.	<i>Cutoff</i> de 50%			
	EMQ		EMA	
	Média	Mediana	Média	Mediana
$\hat{\gamma}_M(h)$	7513515,216	7505185,500	2475,175	2478,298
$\hat{\gamma}_{CH}(h)$	12776716,396	12757246,000	3089,635	3085,512
$\hat{\gamma}_{Md}(h)$	17915687,502	17921497,000	3598,669	3599,285
$\hat{\gamma}_H(h)$	3411837,722	3413561,000	1691,802	1692,667
$\hat{\gamma}_G(h)$	2016656,800	2012558,000	1277,607	1277,353
$\hat{\gamma}_{Pw}(h)$	147492174069211	76904615000	487000,5	43771,5
$\hat{\gamma}_{N1}(h)$	2697222,027	2703167,500	1512,212	1513,616
$\hat{\gamma}_{N2}(h)$	7575400,164	7566304,000	2479,657	2480,874

Tabela 69B E.M.Q. e E.M.A de dados com n=400, com 5% *outliers* e com tendência quadrática e sem *Cutoff*.

Est.	Todas as distâncias			
	EMQ		EMA	
	Média	Mediana	Média	Mediana
$\hat{\gamma}_M(h)$	37934505,533	37739813,500	5695,216	5694,262
$\hat{\gamma}_{CH}(h)$	53096203,307	52730508,000	6538,732	6524,959
$\hat{\gamma}_{Md}(h)$	116415070,464	114012373,500	8889,400	8882,765
$\hat{\gamma}_H(h)$	18038749,876	17391001,500	4031,217	4003,787
$\hat{\gamma}_G(h)$	9375833,162	7682993,500	2458,343	2414,467
$\hat{\gamma}_{Pw}(h)$	5587815423140100	2656203000000	2264946,7	184612,1
$\hat{\gamma}_{N1}(h)$	14358710,878	14318667,000	3687,823	3683,898
$\hat{\gamma}_{N2}(h)$	23235330,144	23200077,000	4652,106	4655,108

Tabela 70B E.M.Q. e E.M.A de dados com n=400, com 5% *outliers*, com tendência quadrática e *Cutoff* de 50%.

Est.	<i>Cutoff</i> de 50%			
	EMQ		EMA	
	Média	Mediana	Média	Mediana
$\hat{\gamma}_M(h)$	16026715,880	15994020,500	3833,325	3829,304
$\hat{\gamma}_{CH}(h)$	19303425,920	19285328,500	3888,298	3885,422
$\hat{\gamma}_{Md}(h)$	21954597,373	21887656,000	4011,254	4010,996
$\hat{\gamma}_H(h)$	10409763,220	10423612,000	3132,044	3138,114
$\hat{\gamma}_G(h)$	3655359,767	3666905,000	1720,435	1722,928
$\hat{\gamma}_{Pw}(h)$	1580824059237030	313811350000	1561076,3	110186,9
$\hat{\gamma}_{N1}(h)$	9253890,920	9243770,500	2972,711	2974,552
$\hat{\gamma}_{N2}(h)$	16127942,716	16103721,000	3852,178	3849,177

Tabela 71B E.M.Q. e E.M.A de dados com n=400, com 10% *outliers* e com tendência quadrática e sem *Cutoff*.

Est.	Todas as distâncias			
	EMQ		EMA	
	Média	Mediana	Média	Mediana
$\hat{\gamma}_M(h)$	55822008,911	50251709,500	7293,769	6973,007
$\hat{\gamma}_{CH}(h)$	40443293,740	39067647,500	5777,535	5691,890
$\hat{\gamma}_{Md}(h)$	30671600,067	30573233,000	4776,458	4776,104
$\hat{\gamma}_H(h)$	46673885,778	41400900,500	6687,799	6351,812
$\hat{\gamma}_G(h)$	8572424,264	8542422,500	2625,626	2623,894
$\hat{\gamma}_{Pw}(h)$	1018931818562660	2886566500000	1611476,3	289513,3
$\hat{\gamma}_{N1}(h)$	44463157,073	39470722,000	6548,917	6232,538
$\hat{\gamma}_{N2}(h)$	56385098,551	50687621,000	7343,921	7006,306

Tabela 72B E.M.Q. e E.M.A de dados com n=400, com 10% *outliers*, com tendência quadrática e *Cutoff* de 50%.

Est.	<i>Cutoff</i> de 50%			
	EMQ		EMA	
	Média	Mediana	Média	Mediana
$\hat{\gamma}_M(h)$	98266550,038	98245052,500	9809,462	9816,180
$\hat{\gamma}_{CH}(h)$	81646671,676	81645648,500	8498,041	8498,452
$\hat{\gamma}_{Md}(h)$	60472332,627	60407841,000	6848,998	6844,307
$\hat{\gamma}_H(h)$	88439679,049	88226238,000	9320,371	9311,964
$\hat{\gamma}_G(h)$	30761060,660	30607439,500	5067,358	5061,237
$\hat{\gamma}_{Pw}(h)$	89705145,378	89350514,500	9384,863	9376,037
$\hat{\gamma}_{N1}(h)$	85829359,978	86002832,000	9199,928	9215,972
$\hat{\gamma}_{N2}(h)$	99242314,751	99316862,500	9866,122	9874,293

ANEXO C - Dados utilizados no exemplo

Tabela 1C Dados de eucalipto

Parcela	Longitude	Latitude	Índice de sítio
1	236256	7612563	29
2	236429	7612560	27
3	236429	7612387	29
4	236601	7612387	29
5	236601	7612215	27
6	236774	7612215	29
7	236601	7612042	27
8	235910	7611866	25
9	236083	7611866	31
10	236259	7611866	29
11	236259	7612042	29
12	236083	7612042	27
13	236259	7612213	29
14	236432	7612213	25
15	236083	7612386	29
16	235910	7612386	27
17	235910	7612560	29
18	236083	7612560	27
19	236429	7611696	25
20	236601	7611696	29
21	236774	7611696	29
22	236947	7611696	29
23	237120	7611869	31
24	236947	7611523	29
25	236774	7611523	29
26	236601	7611523	31
27	236429	7611523	27
28	236253	7611520	27
29	235455	7611696	25
30	235564	7611866	31
31	235910	7612093	27
32	235737	7611966	29
33	235564	7611520	25
34	235737	7611520	25
35	235737	7611693	25
36	235910	7611693	27
37	236083	7611693	27
38	236259	7611693	25

Tabela 1C, continua

39	236083	7611520	25
40	235910	7611520	25
41	235910	7611377	27
42	235391	7611441	25
43	236256	7611350	23
44	236429	7611350	27
45	236601	7611350	29
46	236083	7611177	25
47	236256	7611177	25
48	236428	7611183	27
49	235910	7611004	25
50	236083	7611004	25
51	236256	7611004	27
52	235910	7610828	23
53	236083	7610828	25
54	235931	7610655	25
55	235048	7611002	25
56	234787	7611175	31
57	234778	7611002	31
58	234872	7610929	25
59	235048	7611348	27
60	235218	7611523	23
61	235306	7611696	27
62	234954	7611177	27
63	235218	7610656	27
64	235218	7610829	31
65	235391	7610829	29
66	235391	7610656	25
67	235564	7610832	25
68	235218	7611002	27
69	235391	7611175	25
70	235391	7611002	27
71	235564	7611002	25
72	235564	7611175	27
73	235737	7611175	23
74	235737	7611002	23
75	235564	7611348	19
76	235391	7611348	25
77	235737	7611348	23
78	235910	7611175	23
79	235391	7612045	29
80	235391	7612215	29

Tabela 1C, continua

81	235564	7612215	27
82	235564	7612387	31
83	235391	7612387	29
84	235618	7610638	29
85	235589	7610100	31
86	235391	7612560	25
87	235564	7612560	31
88	235564	7610656	17
89	235564	7610483	17
90	235512	7610310	17
91	235512	7610137	21
92	235218	7612733	31
93	235391	7612733	33
94	235564	7612906	33
95	235564	7612782	29
96	235737	7610483	27
97	235910	7610483	27
98	235910	7610310	27
99	235910	7610137	29
100	235910	7609967	27
101	236083	7610137	31
102	236083	7610310	27
103	236083	7610483	29
104	235737	7610137	25
105	235737	7610310	27
106	235218	7610310	27
107	235045	7610310	27
108	235042	7610483	27
109	234875	7610656	23
110	234814	7610829	25
111	235910	7612906	25
112	235391	7609965	25
113	235734	7609791	21
114	235737	7609961	25
115	235564	7609964	29
116	235325	7610137	27
117	235488	7609791	25
118	235391	7610310	29
119	235737	7612906	25
120	235737	7612733	25
121	235679	7612560	29
122	235737	7613252	25

Tabela 1C, conclusão.

123	235910	7613252	21
124	235737	7613079	21
125	235910	7613079	23
126	236083	7613252	23
127	236083	7613079	25
128	236083	7613425	23
129	235910	7613425	23
130	236083	7613598	31
131	236259	7613598	31
132	236259	7613425	29
133	236432	7613252	29
134	236259	7613121	29
135	236432	7613079	27
136	236601	7613079	31
137	236429	7613946	29
138	236601	7613946	27
139	236429	7613771	27
140	236256	7613771	33
141	236429	7613598	27
142	236601	7613425	29
143	236774	7613252	29
144	236947	7613079	25
145	237469	7612733	25
146	237818	7612733	27
147	237642	7612618	25
148	237985	7612906	27
149	238158	7612906	25
150	238330	7613027	29
151	237296	7612560	31
152	237120	7612387	29
153	237120	7612215	27
154	237296	7612387	29
155	236947	7612215	27
156	236777	7612042	25
157	236601	7611869	25
158	236774	7611869	27
159	236947	7611869	29
160	237120	7612042	29
161	236947	7612042	25