



THIAGO JUNIOR FURTADO GARCIA

**GEOESTATÍSTICA APLICADA ÀS NORMAIS
CLIMATOLÓGICAS DE TEMPERATURAS
MÉDIAS COMPENSADAS NO BRASIL**

LAVRAS - MG

2015

THIAGO JUNIOR FURTADO GARCIA

**GEOESTATÍSTICA APLICADA ÀS NORMAIS CLIMATOLÓGICAS
DE TEMPERATURAS MÉDIAS COMPENSADAS NO BRASIL**

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 e Experimentação Agropecuária, para a obtenção do título de Mestre.

Orientador

Dr. Marcelo Silva de Oliveira

LAVRAS – MG

2015

**Ficha catalográfica elaborada pelo Sistema de Geração de Ficha Catalográfica da Biblioteca
Universitária da UFLA, com dados informados pelo(a) próprio(a) autor(a).**

Garcia, Thiago Junior Furtado.

Geoestatística aplicada às normais climatológicas de
temperaturas médias compensadas no Brasil / Thiago Junior
Furtado Garcia. – Lavras : UFLA, 2015.

67 p. : il.

Dissertação (mestrado acadêmico)–Universidade Federal de
Lavras, 2015.

Orientador: Marcelo Silva de Oliveira.

Bibliografia.

1. Geoestatística. 2. Normais Climatológicas. 3. Krigagem. 4.
Mudanças climáticas. I. Universidade Federal de Lavras. II. Título.

THIAGO JUNIOR FURTADO GARCIA

**GEOESTATÍSTICA APLICADA ÀS NORMAIS CLIMATOLÓGICAS
DE TEMPERATURAS MÉDIAS COMPENSADAS NO BRASIL**

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 e Experimentação Agropecuária, para a obtenção do título de Mestre.

APROVADA em 27 de fevereiro de 2015

Dr. Marcelo Ângelo Cirillo UFLA

Dr. José Márcio de Mello UFLA

Dr. Marcelo Silva de Oliveira
Orientador

LAVRAS – MG

2015

AGRADECIMENTOS

A Deus. A Minha noiva. A minha família. À coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) pela concessão da bolsa de estudos, a Universidade Federal de Lavras, todos os professores e ao Programa de Pós-Graduação em Estatística e Experimentação Agropecuária.

RESUMO

O estudo e a preocupação com o clima vêm se tornando cada vez mais frequentes. Isso porque a vida de forma geral é dependente de condições favoráveis do clima no planeta. Este trabalho estudou as Normais de Temperaturas Médias, no período de 1931-1960 a 1961-1990, utilizando-se da metodologia geoestatística (ramo da estatística espacial). Objetivou-se estudar e entender a distribuição espacial das Normais no território brasileiro, para cada um desses períodos, visando avaliar uma possível mudança das médias de temperatura no país. Com o uso da geoestatística foi possível conhecer características determinísticas (tendência) e características da distribuição espacial como: semivariâncias e semivariograma, modelo de ajuste de semivariogramas, anisotropia, validação de modelos e estudo do melhor modelo de interpolação espacial, a *Krigagem*. Por meio da *Krigagem*, foram gerados dois mapas de interpolação para as temperaturas médias respectivas aos dois períodos. Num terceiro mapa, gerado com as diferenças entre os dois primeiros, foram verificadas as regiões afetadas por mudanças de médias de temperatura. Concluímos o aumento das temperaturas médias em torno de 0,2 a 1,33°C. No período estudado (1931-1990), as médias mundiais aumentaram cerca de 0,55°C. Assim pode-se dizer que o Brasil sofreu mudanças de acréscimos de médias de temperatura do período de 1931 a 1990 compatíveis com as mudanças observadas num contexto mundial. Algumas áreas como a Região Sul, Minas, centro e o norte do Nordeste tiveram mudanças de temperatura fora do padrão da análise do mundial.

Palavras-chave: Geoestatística. Temperatura. Semivariograma. Normais Climatológicas. *Krigagem*. Mudanças climáticas.

ABSTRACT

The study and the concern over the climate have been becoming more and more frequent. That because life in general is dependent upon favorable conditions of climate on the planet. That work studied the Average Temperature Normals in the period of 1931-1960 to 1961-1990 by utilizing of the geostatistical methodology (a branch of spatial statistics). It was aimed to study and understand then spatial distribution of the Normals in the Brazilian territory for each of those periods aiming to evaluate possible change of the means of temperature in the country. From the use of Geoestatistics, it was possible to know deterministic characteristics (trends) and spatial distribution characteristics such as: semivariances and semivariogram, fit model of semivariograms, anisotropy, model validation and study of the best spatial interpolation model, *Kriging*. Through Kriging, two interpolation maps for the average temperatures respective to the two periods were generated. On a third map, generated with the differences between the two first ones, it was found that the regions affected by changes of means of temperature. We conclude the increase of the average temperatures around 0.2 to 1.33°C. In the studied period (1931-1990), the world means increased about 0.55°C. Thus, one can say that Brazil underwent changes of additions of means of temperature of the period of 1931 to 1990 compatible with the changes observed in a world context. Some areas like the South Region, Minas Gerais, Central and Northern Northeast had changes of temperature outside the pattern of the analysis of the world one.

Palavras-chave: Geoestatistics. Temperature. Semivariogram. Climate Normals. *Krigagem*. Climate changes.

SUMÁRIO

1	INTRODUÇÃO.....	8
2	REFERENCIAL TEÓRICO.....	10
2.1	Algumas informações sobre temperatura.....	10
2.2	Normais climatológicas.....	10
2.3	Caracterizações da região em estudo.....	12
2.4	Estudo e preocupação sobre as temperaturas e o clima.....	14
2.5	Geoestatística.....	19
2.6	Variável regionalizada.....	20
2.7	Estacionariedade.....	23
2.8	Variograma/ Semivariograma.....	25
2.9	Modelagem de Semivariograma.....	32
2.10	Auto validação.....	36
2.10.1	Gráfico: Medido vs Estimado.....	37
2.10.2	Erro Médio, Erro Médio Reduzido e Desvio Padrão dos Erros.....	37
2.11	Distância útil variográfica.....	39
2.12	Grau de dependência espacial.....	39
2.13	Tendência.....	41
2.14	Anisotropia e isotropia.....	42
2.15	<i>Krigagem</i>	43
2.15.1	<i>Krigagem</i> simples e <i>Krigagem</i> ordinária.....	44
2.15.2	<i>Krigagem</i> universal.....	47
3	MATERIAL E MÉTODOS.....	50
3.1	Material.....	50
3.2	Métodos.....	50
4	RESULTADOS E DISCUSSÃO.....	52
4.1	Análise preliminar dos dados.....	52
4.2	Discussões iniciais.....	54
4.3	Análise espacial.....	55
4.4	<i>Krigagens</i> das normais.....	58
4.5	Discussões finais.....	59
5	CONCLUSÕES.....	62
	REFERÊNCIAS.....	64

1 INTRODUÇÃO

A preocupação com o clima, principalmente no que diz respeito ao aumento da temperatura mundial, é um dos temas mais discutidos e atuais no mundo inteiro. A sobrevivência da população humana, flora e fauna dependem de características específicas da natureza. Entre essas características encontra-se, evidentemente, o clima do nosso planeta.

Existem e, estão sendo feitos, muitos artigos e trabalhos sobre o clima de países, continentes, regiões, etc. e é claro sobre o clima no mundo: (BARCELLOS et al., 2009; MARENGO, 2008; NOBRE, 2012; SATHLER; REIS, 2010).

Recentemente, o professor Richard Muller e sua equipe de cientistas, com base na Universidade da Califórnia em Berkeley realizaram um estudo com mais de 14,4 milhões de observações coletadas, em mais de 44,4 mil localidades do mundo inteiro e concluíram um aumento da temperatura mundial. Eles, ainda, fizeram o cruzamento de *plots* de variáveis para informações sobre as causas dessa elevação, chegando a uma forte evidência da relação entre as emissões de gases poluentes e o aumento da temperatura no planeta (HICKMAN, 2012).

Nesse trabalho, por meio da teoria geoestatística, foram analisadas as temperaturas médias no Brasil, utilizando os dados de Normais climatológicas dos períodos 1931-1960, 1961-1990 e avaliou-se uma possível elevação da temperatura pelos mapas gerados ao fim de todas as análises geoestatísticas. Entre tais análises se encontram: construção de semivariogramas ou correlogramas, estudo de isotropia e anisotropia, tendência e dependência espacial da variável em estudo e finalmente a *Krigagem*.

A construção de mapas por meio da *Krigagem*, referentes às duas Normais e um terceiro mapa gerado pela diferença entre os dois primeiros, junto

a todas as informações obtidas das análises forneceram informações sobre a temperatura no Brasil durante o período de 1931 a 1990.

Com os resultados da análise geoestatística fez-se uma reflexão sobre uma possível mudança de temperatura no país. Buscou-se responder se as médias de temperatura no Brasil, a exemplo do que se concluiu sobre as médias mundiais no artigo da equipe do Projeto Berkeley: A New estimate of the average Earth Surface Land Temperature Spanning 1753 to 2011, também sofreram um aumento e se as regiões brasileiras sofreram diferentes mudanças em suas médias de temperatura.

2 REFERENCIAL TEÓRICO

2.1 Algumas informações sobre temperatura

É comum a confusão que se faz entre tempo e clima. Portanto, antes de qualquer coisa é importante que se conheça ambos os conceitos. Define-se o tempo como “o estado físico das condições atmosféricas em um determinado momento e local”. O clima é definido como o estudo médio do tempo para determinado período em um dado local (INSTITUTO NACIONAL DE METEOROLOGIA - INMET, 2014). O clima tende a ser estável, duradouro, não sofrendo mudanças repentinas como, por exemplo, pode sofrer o tempo. Vários elementos compõem o clima: temperatura, umidade, ventos, chuvas, massas de ar, ação do homem, entre outros. Portanto, para que se conheça satisfatoriamente o clima de determinada região, é necessário que antes se tenha conhecimento de seus elementos de composição.

A temperatura é um elemento de composição do clima. Entende-se a temperatura como um estado térmico do ar atmosférico, o calor existente na temperatura é predominantemente causado pelo Sol. A irradiação também libera calor na atmosfera afetando a temperatura. Cinco fatores são os principais causadores de variação na temperatura: altitude, latitude, correntes marinhas, Continentalidade e Maritimidade e as estações do ano.

2.2 Normais climatológicas

Por volta de 1860, iniciaram-se as primeiras observações regulares meteorológicas. No Brasil, essas observações iniciaram-se sistematicamente a partir de 1910 com a unificação das atividades do observatório Nacional da Marinha e do Telégrafo Nacional (INMET, 2014).

Em 1872, o Comitê Meteorológico Internacional tomou a decisão de agrupar valores médios climatológicos em um período uniforme de modo a compatibilizar os dados coletados em várias estações. De acordo com as Regulamentações Técnicas, as Normais são definidas como: “valores médios calculados para um período relativamente longo e uniforme, compreendendo no mínimo três décadas consecutivas”. Tem-se ainda a definição de padrões climatológicos normais: “médias de dados climatológicos por períodos consecutivos de 30 anos”. As primeiras Normais climatológicas brasileiras, correspondentes aos anos de 1931-1960, foram publicadas pelo escritório de meteorologia do Ministério da Agricultura em 1970. Em 1992, o INMET¹ reuniu informações sobre pressão atmosférica, temperatura máxima, mínima e média, umidade relativa, nebulosidade, precipitação máxima em 24 horas, evaporação total e insolação total de 209 estações meteorológicas e publicou as Normais Climatológicas 1961–1990 (INMET, 2014).

Utilizou-se neste trabalho as Normais Climatológicas de temperaturas médias compensadas. O cálculo diário dos dados para a constituição dessas Normais até o ano 1937 foi feito de acordo com a equação (1). Os índices representam as horas de coletas: 7, 14 e 21 horas. A partir de 1937, esse cálculo foi feito pela equação (2), que além das temperaturas recolhidas às 12 e 24 horas, também leva em consideração as temperaturas máxima e mínima diárias (BRASIL, 1969).

Equação de Medição de temperatura média compensada até o ano de 1937:

$$T_m = \frac{(T_7 + T_{14} + 2T_{21})}{4} \quad (1)$$

¹ Instituto Nacional de Meteorologia. O INMET é um órgão governamental que contém informações completas sobre o tempo e o clima em todo Brasil.

Equação de Medição de temperatura média compensada a partir do ano de 1937:

$$T_m = \frac{(T_{12} + 2T_{24} + T_{\text{máx}} + T_{\text{mín}})}{5} \quad (2)$$

2.3 Caracterizações da região em estudo

Segundo o Instituto Brasileiro de Geografia e Estatística - IBGE (2014), o Brasil é uma Federação constituída por 26 Estados e o Distrito Federal. Localiza-se no subcontinente da América do Sul. Seu território possui dimensão continental. Possui área de 8 514 876 km², sendo o quinto maior país em extensão do mundo. Cerca de 90% do território está localizado entre os trópicos de Capricórnio e Câncer. É atravessado ao norte pela Linha do Equador e ao sul pelo Trópico de Capricórnio, assim sendo, situa-se em zonas de latitudes baixas onde prevalecem climas quentes e úmidos com temperaturas médias de 20°C. Possui, também, a quinta maior população do mundo, cerca de 200,4 milhões de habitantes. Como consequência da grande extensão de seu território, diversidade de formas de relevo e altitudes e, também, a dinâmica das correntes e massas de ar, o país possui grande variedade de climas.

Alvares et al. (2014), desenvolveram um sistema de informação para identificar tipos de clima no Brasil de acordo com o critério de Köppen (1936).

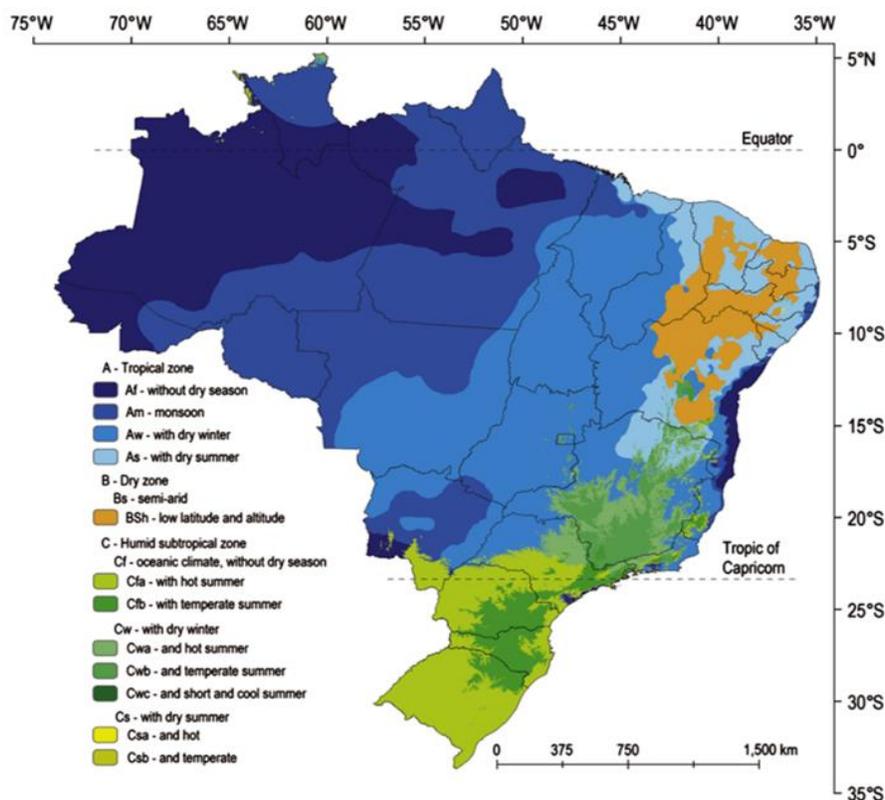


Figura 1 Climas no Brasil de acordo com o critério de Köppen (1936)

Fonte: Alvares et al. (2014).

Segundo o (IBGE), as massas de ar que interferem mais diretamente no país são a equatorial (continental e atlântica), a tropical (continental e atlântica) e a polar atlântica. Por esse motivo, são verificados no país desde climas superúmidos quentes, provenientes das massas equatoriais, (caso de grande parte da região Amazônica), até climas semiáridos muito fortes, próprios do sertão nordestino. Temos, então, como principais tipos climáticos brasileiros: subtropical, semiárido, equatorial úmido, equatorial semiúmido, tropical e tropical de altitude.



Figura 2 Mapa com climas e correntes do Brasil

Fonte: Estados... (2014).

2.4 Estudo e preocupação sobre as Temperaturas e o Clima

O matemático Jean Baptiste Joseph Fourier foi o pioneiro (início do século XIX) na observação de que a atmosfera terrestre atua como uma estufa no planeta Terra, o que possibilita um ambiente de sobrevivência para plantas e animais. Porém, essa “estufa” contém gases, como vapor d’água e gás carbônico, que funcionam como uma espécie de vidro, pois permitem a entrada da radiação solar (o que aquece a Terra), mas dificultam a perda de calor para o espaço. O próprio Fourier já havia constatado a elevação da temperatura terrestre em decorrência da absorção de calor pelos gases atmosféricos.

Berkeley Earth Surface Temperature Project (BEST), ou projeto Berkeley de Temperatura da Superfície Terrestre como ficou conhecido no

Brasil, foi lançado e presidido pelo professor e cientista físico Richard Muller. Sua equipe conta com Robert Rohde (físico), Judith Curry (climatologista) e outros físicos e estatísticos. O projeto lançado em outubro de 2011 destinou-se a responder críticas de céticos do clima sobre a qualidade de registros coletados das temperaturas mundiais, mas acabou trazendo à tona uma pertinente discussão sobre o aquecimento global. O projeto confirma um aumento da superfície da Terra, ao longo dos últimos 250 anos, tendo encontrado como melhor explicação para esse aumento a combinação das mudanças nos níveis de dióxido de carbono e atividades vulcânicas (CARBON BRIEF, 2013).

No artigo proveniente do Projeto Berkeley, A New estimate of the average Earth Surface Land Temperature Spanning 1753 to 2011, Rohde et al. (2013) estimaram a temperatura média da Terra para o período de 1753 a 2011, e, ainda, procurando diminuir o viés das estimativas, utilizaram um número maior de estações de coleta quando comparados a outros três grandes projetos: The National Aeronautics and Space Administration Goddard Institute for Space Studies (NASA GISS), The National Oceanic and Atmospheric Administration (NOAA) e uma colaboração do Hadley Centre of the UK Meteorological Office com The Climate Research Unit of East Anglia (HadCRU). A equipe do BEST coletou cerca de 14,4 milhões de dados de médias de temperaturas mensais em cerca de 44.455 estações.

Segundo Rohde et al. (2013, p. 1), “a temperatura média da superfície da Terra é um indicador chave de mudanças climáticas”. Os autores trazem no artigo New Estimate um gráfico com as temperaturas médias em décadas e afirmam não haver discordâncias estatísticas significantes entre os quatro grupos. Porém, como se pode notar no gráfico, o Projeto Berkeley estende as estimativas para os anos de 1753.

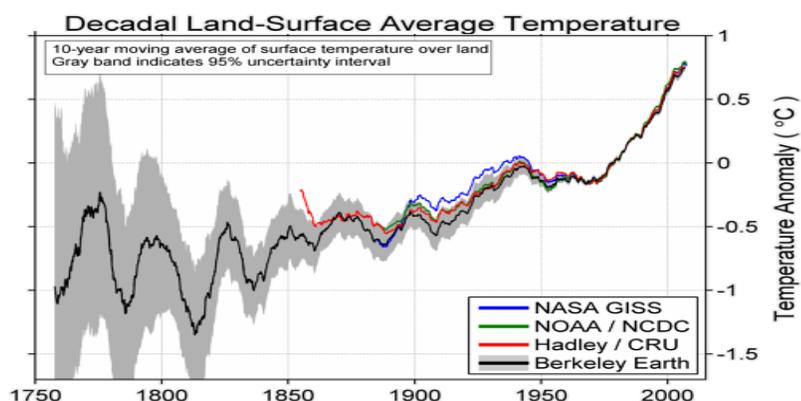


Figura 3 Temperaturas médias globais em décadas nos períodos de 1763-2011.
Fonte: Rohde et al. (2013).

Para o período compreendido entre 1950 e 2000, NASA GISS, NOAA e HadCRU reportaram um aumento da temperatura de superfície do planeta entre 0,81 e 0,93°C. No projeto Berkeley, as estimativas para esse mesmo período são de $0,90 \pm 0,05^\circ\text{C}$ com uma probabilidade de 95% de confiança. A alteração do aquecimento da superfície, relatada pelo Projeto Berkeley, é consistente com resultados anteriores, porém possui menos incertezas. No mapa global seguinte, gerado pela equipe de Muller, pode-se avaliar o aumento das médias entre as décadas de 1951-1960 e 2001-2010.

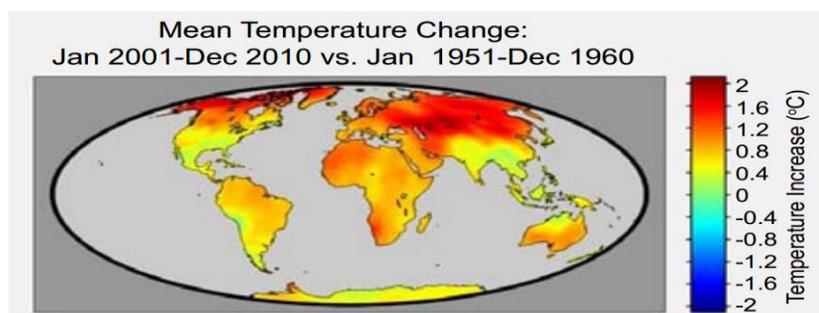


Figura 4 Alteração das Temperaturas médias globais 1951-2011
Fonte: Rohde et al. (2013).

Ainda, segundo Rohde et al. (2013), *The diurnal temperature*, que é a diferença diária entre a temperatura máxima e a temperatura mínima ($D = T_{\text{máx}} - T_{\text{mín}}$) sofreu uma queda lenta no período de 1900 para 1980, e segue aumentando até o período mais recente da análise (ano de 2011). Os cientistas do projeto não esperavam por esse comportamento, considerando-o inesperado.

A maior contribuição do Projeto Berkeley foi utilizar-se de modelos estatísticos “simples”, para descobrir se as mudanças de temperatura global se modelavam mais similarmente: com as mudanças nos níveis de dióxido de carbono (por fatores antropogênicos), atividade solar e ou atividades vulcânicas (CARBON BRIEF, 2013).

Para Rohde et al. (2013), a melhor modelagem para a contribuição da atividade vulcânica no período é de $-1,5 \pm 0,5^{\circ}\text{C}$ por 100 mg de sulfato atmosférico. A contribuição da atividade antropogênica, que é relativa à emissão de CO_2 pelo homem (proveniente de amostras atmosféricas e ar preso em gelo polar), é de $3 \pm 0,3^{\circ}\text{C}$. Quando os autores incluíram informações da atividade solar (dados assumidos pelo IPCC²), eles concluíram sua não significância ao modelo de ajuste das temperaturas. Desse modo, a melhor explicação ao aumento de temperatura, em níveis mundiais no período de 1753 a 2011, deve-se a uma combinação linear simples, composta pela atividade vulcânica (baseada em estimativas da injeção de sulfato na atmosfera), e um termo antropogênico representado pelo logaritmo de CO_2 .

O professor Muller enfatiza que o resultado do estudo do projeto Berkeley não prova que o aquecimento global seja causado pelos gases poluentes lançados na atmosfera pelos seres humanos, porém, é atualmente sua melhor explicação (HICKMAN, 2012).

² IPCC. Intergovernmental Panel on Climate Change. Painel Intergovernamental sobre Mudanças Climáticas da ONU.

Ainda, segundo Hickman (2012), Muller afirma que, pelas análises de sua equipe no BEST, pode ocorrer um aumento na temperatura de 1,5°C nos próximos 50 anos e que esse aumento pode acontecer em menos de 20 anos caso a China continue utilizando o carvão mineral de forma desenfreada.

Em conclusão, os cientistas do Projeto Berkeley afirmam que o modelo matemático (linha vermelha), proveniente da combinação linear entre emissões de sulfato vulcânico (responsável por quedas curtas) e o logaritmo natural de CO₂ (responsável por crescimento gradual), tem a melhor aproximação aos dados de temperaturas globais (Berkeley Earth Average, linha preta) ao longo dos anos de 1763 a 2011. Afirmam, ainda, que a inclusão da atividade solar no modelo foi considerada não significativa (ROHDE et al., 2013).

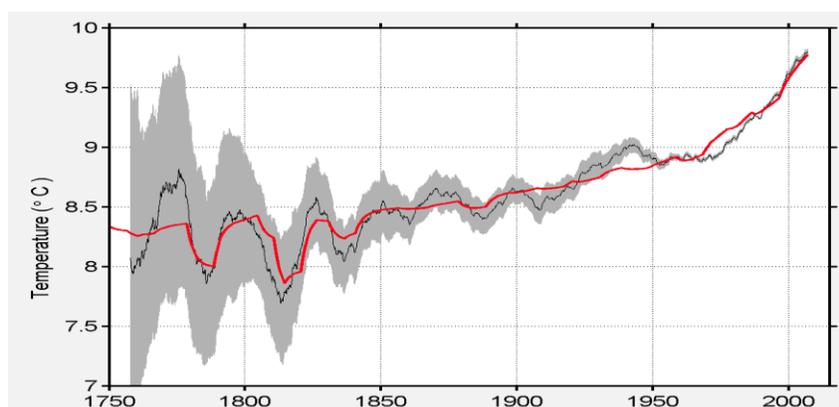


Figura 5 Temperaturas médias em décadas provenientes do (Berkeley Earth average) comparado com a combinação linear de atividades vulcânicas (emissões de sulfato) e o logaritmo natural de CO₂

Fonte: Rohde et al. (2013).

A elevação das temperaturas médias mundiais relatadas pelo Projeto Berkeley foram a inspiração para a realização desse trabalho, onde buscou-se responder sobre uma possível mudança das médias de temperatura no Brasil. O gráfico seguinte foi modificado do original do BEST. Percebe-se a relação das

temperaturas mundiais (cor preta) desde 1750 até 2010 e a curva da combinação da atividade vulcânica e das emissões de CO₂. Pela análise da equipe do BEST, no período de 1930 a 1990, período estudado nesse trabalho para as temperaturas no Brasil, a temperatura média mundial elevou-se em cerca de 0,55°C.

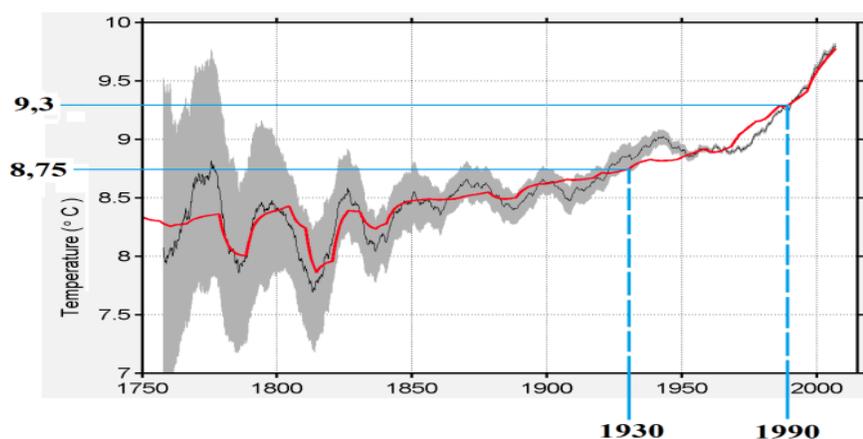


Figura 6 Temperaturas médias em décadas provenientes do (BEST) comparado com a combinação linear de atividades vulcânicas (emissões de sulfato) e o logaritmo natural de CO₂.

Fonte: Rohde et al. (2013).

2.5 Geoestatística

No ano de 1951, Daniel G. Krige, trabalhando com dados de concentração de ouro, chegou à conclusão de que a variância de seus dados dependia da distância de sua amostragem. Krige percebeu que, para formalizar as análises do estudo, precisaria levar em consideração a distância entre as observações. Surge desse ponto o conceito de geoestatística, a qual utiliza a localização geográfica, ou georreferenciamento dos dados e considera a

dependência espacial (CAMARGO, 2001). A geoestatística é uma das ramificações da estatística espacial.

Rivoirard (2005) nos traz que de 1954 a 1963 o matemático e engenheiro Professor Georges Matheron, trabalhando em pesquisa geológica em depósitos de ouro na França e Argélia, desenvolveu e formalizou a maioria dos conceitos da teoria que nomeou de geoestatística.

Clark (1979) nos conta que a geoestatística surgiu para fazer a estimação de ouro em reservas de minas e essa foi sua utilidade por um período considerável. Contudo, suas técnicas podem ser usadas em qualquer situação em que medidas amostrais sejam feitas em um espaço e de onde se espera que essa posição no espaço afete as medidas e, ainda, que exista relação entre os pontos amostrados e seus vizinhos.

Para Isaaks e Srivastava (1989), a fundamentação geoestatística baseia-se na esperança de que, em média, amostras retiradas próximas umas das outras sejam mais similares do que amostras retiradas em pontos distantes.

Matheron desenvolveu amplamente em suas teses o que conhecemos por geoestatística “clássica” (geoestatística linear baseada na estacionariedade da covariância e na função de semivariograma). Ele compartilhava do ideal de Krige de que não fazia sentido trabalhar com as variâncias amostradas sem considerar as distâncias que separavam essas amostras. Ele desenvolveu, então, em 1963, a teoria das variáveis regionalizadas, outro nome dado à geoestatística.

2.6 Variável Regionalizada

Iniciamos este capítulo com uma explicação sobre as simbologias que serão utilizadas neste trabalho. A variável $Z(x_i)$ representa a ocorrência da variável Z em uma referência de espaço (posição) bidimensional (x_i) .

Usualmente, na geoestatística, utiliza-se a letra h para representar distâncias entre pontos. Assim sendo, h é comumente chamada de *lag*, ou seja, o *lag* é um vetor de distância entre pontos amostrados.

A teoria sobre as variáveis regionalizadas pressupõe que cada dado $z(x_i)$ pode ser modelado como uma variável aleatória. Assim sendo, $Z(x_i)$ possui média, variância e uma função de densidade acumulada. O conjunto de variáveis aleatórias $\{Z(x_i), x_i \in A\}$ é uma função aleatória, também, chamada de processo aleatório ou processo estocástico.

Para Druck et al. (2004), uma variável regionalizada é uma variável distribuída no espaço, cujos valores são considerados realizações de uma função aleatória (ou processo aleatório, ou ainda, processo estocástico).

A variável regionalizada pode, também, ser entendida como a representação do conjunto de variáveis aleatórias existentes em uma região de estudo que trata de um mesmo fenômeno, assumindo diferentes valores em função da localização onde é amostrada no campo. Por exemplo, neste trabalho existem inúmeras variáveis aleatórias de temperatura média compensada, obtidas no território brasileiro, que é a região de estudo.

Segundo Vicente (2004), as variáveis regionalizadas apresentam continuidade espacial. A continuidade espacial existe pela razão de que variáveis próximas possuem valores similares e, com o aumento da distância entre essas variáveis, essa similaridade nos valores tende a desaparecer. Se existe continuidade espacial no fenômeno estudado, os dados de pontos vizinhos se assemelharão, enquanto dados de pontos distantes serão distintos. Funções ordinárias são usadas para representar essa continuidade espacial.

Para Mello (2004), cada variável aleatória $Z(x_i)$ pode ser expressa pelo somatório de três componentes: uma componente estrutural, uma componente aleatória e um erro aleatório.

Assim,

$$Z(x_i) = \mu(x_i) + \varepsilon'(x_i) + \varepsilon'' \quad (3)$$

Onde:

$\mu(x_i)$ – função determinística. Descreve a componente estrutural de Z em x_i ; está associado a um valor médio constante ou uma tendência.

$\varepsilon'(x_i)$ – Componente aleatória, espacialmente correlacionada. É nessa componente que a geoestatística atua.

ε'' - erro aleatório ou residual (ruído), não correlacionado. Possui distribuição normal $(0, \sigma^2)$.

A seguir uma figura de Burrough (1987), modificada por Camargo (2001), que ilustra as três componentes fundamentais da variação espacial. Em a) temos uma variação abrupta da componente determinística, na figura b) essa mesma componente apresenta tendência constante.

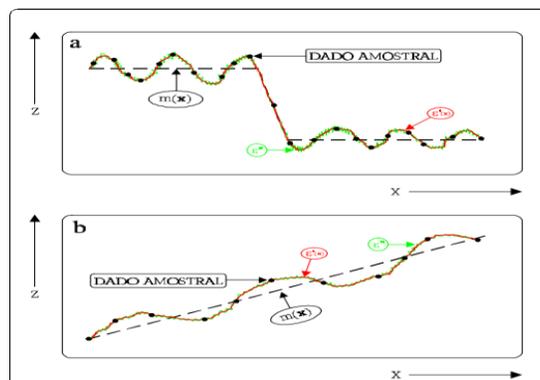


Figura 7 Componentes de uma variável regionalizada

Fonte: Burrough (1985 apud CAMARGO, 2001).

2.7 Estacionariedade

Guimarães (2004), declara que um processo (ou uma variável) é estacionário se seu desenvolvimento no espaço (ou tempo) ocorre de maneira relativamente homogênea. Para Andriotti (1988), isso implica que os fatores influenciadores do fenômeno estudado ajam de forma similar em toda área de estudo ou, ainda, que a variável regionalizada do estudo seja homogênea do ponto de vista estatístico. Observa-se que as características do processo estacionário independem da origem adotada.

Processos não estacionários não atendem as demandas acima citadas e as características do processo dependem da origem adotada. É importante perceber que, dentro de processos não estacionários, podem existir regiões de processos estacionários. Na análise geoestatística exige-se em restrição máxima à existência de estacionariedade de segunda ordem.

Caracteriza-se, como uma variável estacionária de primeira ordem, uma variável aleatória que possua esperança matemática constante, independentemente de sua origem no espaço e, conseqüentemente, possua a mesma média para toda região do estudo.

$$E [Z(x_i)] = E [Z(x_i+h)] = m_1(x_i) = \mu = \text{constante} \quad (4)$$

Segundo Guimarães (2004), uma variável é considerada estacionária de segunda ordem se:

A média é constante (estacionariedade de primeira ordem):

$$E [Z(x_i)] = E [Z(x_i+h)] = m_1(x_i) = \mu = \text{constante} \quad (5)$$

O segundo momento existe:

$$E [Z^2(x_i)] < \infty \quad (6)$$

Em cada par $\{Z(x_i), Z(x_i+h)\}$, a função covariância existe e só depende de h .

$$\text{Cov} [Z(x_i), Z(x_i + h)] = \text{Cov}(h) = C(h) \quad (7)$$

Segundo Landin (2013), a covariância é uma medida de relação mútua entre duas variáveis. Tratando-se da geoestatística a covariância mede a relação entre valores de uma mesma variável aleatória, amostradas em diferentes pontos e separados por um vetor distância h .

Segundo Druck et al. (2004), em uma variável estacionária de segunda ordem, a covariância existe, só depende de h e é dada por:

$$\begin{aligned} \text{Cov} (h) &= \text{Cov} [Z(x_i), Z(x_i + h)] = \\ &= E [Z(x_i) \cdot Z(x_i + h)] - E[Z(x_i)] E[Z(x_i + h)] = \\ &= E [Z(x_i) \cdot Z(x_i + h)] - m^2 \end{aligned} \quad (8)$$

A estacionariedade da covariância implica estacionariedade da variância:

$$\begin{aligned} \text{Cov} [Z(x_i), Z(x_i)] &= E [Z(x_i) \cdot Z(x_i)] - E [Z(x_i)] \cdot E [Z(x_i)] = \\ &= E [Z(x_i)^2] - (E [Z(x_i)])^2 = \text{Var} Z(x_i) \end{aligned}$$

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

$$= E [(Z(x_i))^2] - m^2 = \text{Cov}[Z(x_i), Z(x_i + 0)] = \text{Cov}(0) \quad (9)$$

Quando a estacionariedade de ordem 2 é assumida, infere-se a existência de uma variância e, por conseqüente, a existência de uma variância *a priori* finita. Porém, algumas variáveis regionalizadas provenientes de fenômenos físicos possuem uma capacidade de dispersão infinita.

Para Journel e Huijbregts (1978), existem fenômenos físicos, com capacidade infinita de dispersão, que não atendem totalmente as estacionariedades de ordem 1 e 2. Nessas situações, a hipótese intrínseca exige apenas existência de estacionariedade do semivariograma (semivariograma com patamar definido, explicando toda a variância amostral), sem qualquer restrição quanto à existência de variância finita, sendo a mais adotada em trabalhos geoestatísticos.

Em resumo, a verificação da hipótese intrínseca permite a criação de um modelo de semivariograma empírico satisfatório dentro de um limite espacial onde as relações de dependência espacial são homogêneas.

Como se percebe, devem ser tomados alguns cuidados em relação às hipóteses adotadas e aos fundamentos geoestatísticos iniciais. Isso possibilitará a criação de um bom modelo de semivariograma que influenciará diretamente na qualidade das interpolações espaciais futuras.

2.8 Variograma/ Semivariograma

Druck et al. (2004), trazem que se as estacionariedades de ordem 1 e 2 existirem, logo, a determinação da função $C(h)$ é suficiente para se caracterizar a variável regionalizada. Segundo esses autores, para que seja determinada a função $C(h)$ utiliza-se a função auxiliar conhecida por variograma: $2\gamma(h)$.

$$\begin{aligned}
2\gamma(h) &= E [(Z(x_i) - Z(x_i+h))^2] = \\
&= E [(Z(x_i))^2 - 2Z(x_i) \cdot Z(x_i+h) + (Z(x_i+h))^2] \\
&= E [(Z(x_i))^2] - 2E [Z(x_i) \cdot Z(x_i+h)] + E [(Z(x_i+h))^2]
\end{aligned} \tag{10}$$

Da equação (8) obtém-se:

$$E [Z(x_i) \cdot Z(x_i+h)] = C(h) + m^2 \tag{11}$$

Da equação (9), retira-se:

$$E [(Z(x_i))^2] = E [(Z(x_i+h))^2] = C(0) + m^2 \tag{12}$$

Utilizando (11) e (12) em (10) obtém-se:

$$\begin{aligned}
2\gamma(h) &= C(0) + m^2 - 2(C(h) + m^2) + C(0) + m^2 = 2C(0) + 2m^2 - 2C(h) - 2m^2 \\
2\gamma(h) &= 2C(0) - 2C(h) \quad \text{ou} \quad \gamma(h) = C(0) - C(h)
\end{aligned} \tag{13}$$

$\gamma(h)$ é chamado de **semivariograma**. A semivariância em algum *lag* h , é a covariância em 0 menos a covariância em h .

Bohling (2005), afirma que, sob a condição de estacionariedade de segunda ordem (médias e variâncias constantes espacialmente), existem três relações entre as funções de covariância, correlação e semivariograma.

$$C(0) = \text{Cov} [Z(x_i), Z(x_i)] = \text{Var} [Z(x_i)] \tag{14}$$

$$\rho(h) = C(h)/C(0) = \text{Correlograma} \tag{15}$$

$$\gamma(h) = C(0) - C(h) \tag{16}$$

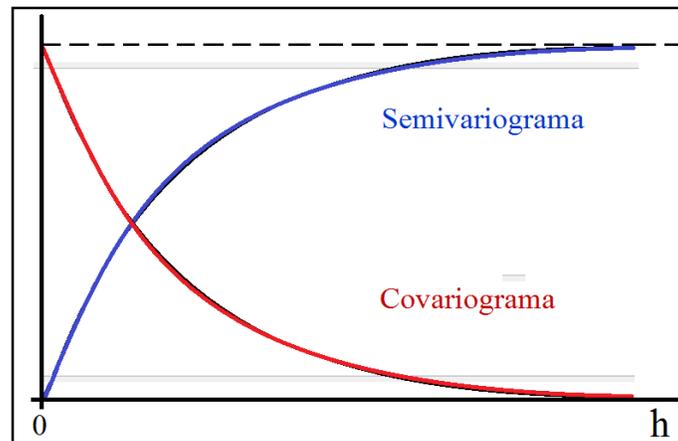


Figura 8 Relação Covariograma vs Semivariograma

Fonte: Modificado de Bohling (2005).

Pode-se notar que a covariância em $h = 0$, $C(0)$ deve ser igual à variância global da variável em estudo. O correlograma deverá apresentar-se como a função de covariância escalada pela variância. Para utilizar o covariograma é necessário que a variável estudada seja estacionária de segunda ordem. Como na prática isso dificilmente é alcançado, a maioria dos trabalhos geoestatísticos utiliza os variogramas.

O semivariograma deve apresentar-se como a função de covariância de “ponta cabeça”. Ao contrário de analistas de séries temporais, que preferem trabalhar tanto com a função de covariância ou o correlograma, geoestatísticos normalmente trabalham com os semivariogramas. Isto é primeiramente porque o semivariograma, que calcula a média do quadrado das diferenças da variável, tende a filtrar a influência de uma variação média no espaço. Além disso, o semivariograma pode ser aplicado sempre que, as diferenças da variável regionalizada, $Z(x_i + h) - Z(x_i)$, são de segunda ordem. Esta forma de estacionariedade é referenciada como hipótese intrínseca, é uma exigência mais fraca da própria variável, ou seja, o semivariograma pode ser definido em alguns casos em que a função de covariância não pode (semivariâncias que continuam crescendo infinitamente) (BOHLING, 2005, p. 13).

Para Andriotti (1988), o variograma é uma função intrínseca que estrutura o fenômeno estudado através de relações estatísticas de covariância existentes entre as amostras espaçadas por sucessivos valores de h (lag). Segundo o autor, pode-se descrever matematicamente o variograma como a esperança matemática do quadrado das diferenças entre todos os pares de pontos presentes na área de estudo, tomados a uma distância h um dos outros.

$$\text{Var} [Z(x_i + h) - Z(x_i)] = E [(Z(x_i + h) - Z(x_i))^2] = 2\gamma(h) \quad (17)$$

Os professores e autores Yamamoto e Landim (2013) chamam atenção para a confusão que se faz entre variograma e semivariograma. Na prática, o semivariograma divide a média das diferenças ao quadrado por dois:

$$\gamma(h) = 1/2 E [(Z(x_i + h) - Z(x_i))^2] \quad (18)$$

Assim, $2 \gamma(h)$ fica conhecido por variograma, enquanto $\gamma(h)$ é chamado de semivariograma. Muitos autores tratam o semivariograma pelo nome de variograma, mas sempre consideram a divisão por dois em seus cálculos. Neste trabalho utiliza-se o termo semivariograma.

O semivariograma é a ferramenta geoestatística mais utilizada, uma vez que exige hipóteses de estacionariedades menos restritivas (Hipótese Intrínseca). O estudo dos modelos de semivariogramas experimentais (empíricos) permite ao pesquisador definir um modelo que possa descrever da melhor forma o comportamento de seus dados no espaço. A utilização do método dos momentos é o mais usual para se estimar semivariâncias.

Vários modelos existem na literatura. Modelos muito utilizados são Cressie e Hawkins (1993) e Matheron (1962).

O semivariograma experimental Clássico de Matheron (1962) - amplamente utilizado por se tratar de um modelo eficiente e didático para situações homogêneas - é um estimador não enviesado para as semivariâncias, porém exige que a variável regionalizada seja um processo estocástico intrinsecamente estacionário. Além disso, o estimador de Matheron é severamente afetado por *outliers*. Outro tipo muito utilizado de semivariograma empírico é o Estimador de Cressie e Hawkins (1993). Na presença de valores amostrais discrepantes, que não podem ser ignorados pelo pesquisador, pode-se fazer uso do Estimador de Cressie e Hawkins. Esse estimador é considerado “robusto”, pois não é afetado pelos *outliers*, sendo frequentemente utilizado em simulações de distribuições assimétricas.

Estimador não tão conhecido, mas também utilizado, é o Estimador de Pairwise. Ele tende a diminuir os efeitos de *outliers*, porém não pode ser aplicado em situações onde a soma dos valores da variável Z nos pontos (x_i) e $(x_i + h)$ seja zero. Outro ponto negativo reside no fato de que, quando as médias estão se aproximando de zero, as semivariâncias podem apresentar saltos.

Respectivamente, os modelos empíricos de Matheron, Cressie e Hawkins e Pairwise.

$$\hat{\gamma}_M(h) = \frac{1}{2N(h)} \sum^{N(h)} E[Z(x_i) - Z(x_i+h)]^2 \quad (19)$$

$$\hat{\gamma}_C(h) = \frac{\left\{ \frac{1}{N(h)} \sum^{N(h)} E[Z(x_i) - Z(x_i+h)]^{\frac{1}{4}} \right\}^4}{0,914 + \frac{0,988}{N(h)}} \quad (20)$$

$$\hat{\gamma}_P(h) = \frac{2}{N(h)} \sum^{N(h)} \left[\frac{Z(x_i) - Z(x_i+h)}{Z(x_i) + Z(x_i+h)} \right]^2 \quad (21)$$

Onde: $N(h)$ = Número de pares de dados utilizados no cálculo da semivariância correspondente a uma distância h , e $x = (x_i, y_i) \in S \subset \mathbb{R}^2$ são posições amostrais.

Os estimadores New 1 e New 2 dos autores Li e Lake (1994) trabalham no sentido de reduzir erros quando as amostras apresentam-se muito distantes.

Na figura seguinte, observa-se o comportamento ideal esperado de um semivariograma experimental (ajustado por um modelo). Esse semivariograma indica estacionariedade de 2ª ordem, uma vez que atinge um patamar bem definido.

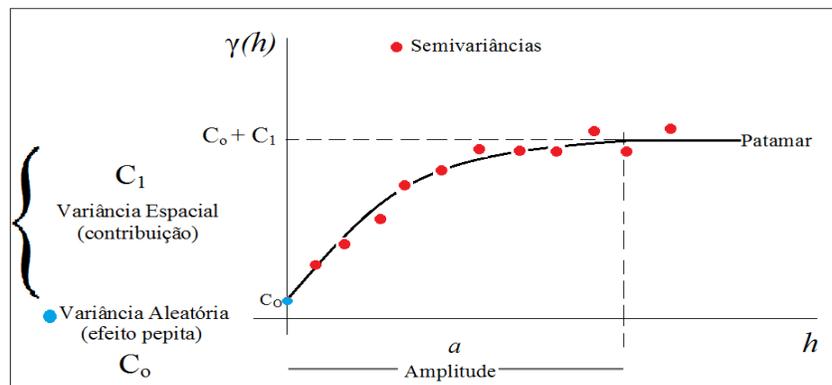


Figura 9 Semivariograma e seus componentes

Parâmetros do semivariograma de acordo com Camargo (2001):

- a) C_0 - Efeito pepita (*nugget*) – Semivariância em $h = 0$. Idealmente $\gamma(0) = 0$, porém, na prática, isso não ocorre. Quando h tende para

zero, $\gamma(h)$ se aproxima de um valor positivo conhecido por efeito pepita³. O efeito pepita representa a semivariância de distâncias próximas de zero, menores do que a menor distância amostrada.

- b) C_1 – Variância Espacial (Contribuição) – é a variação causada pelo espaço. Diferença entre o patamar e o efeito pepita.
- c) C – Patamar (*Sill*) – Valor onde o semivariograma se estabiliza correspondente à semivariância do alcance, $\gamma(a)$. O patamar é a soma de efeito pepita e contribuição espacial. $C = C_0 + C_1$.
- d) a - Amplitude ou alcance (*Range*) – distância perante a qual existe correlação espacial dentre as amostras.

Caso o semivariograma seja constante e igual ao patamar para qualquer valor de h , tem-se o chamado **efeito pepita puro**. Nesse caso, não existe dependência espacial nos dados amostrados. Guimarães (2004) diz que uma vez que se observe efeito pepita puro, em um semivariograma dos resíduos, pode-se afirmar que a superfície de tendência é a melhor representação espacial para a variável. Um caso especial e muito importante de semivariograma é o **semivariograma sem patamar**. Nesse tipo de semivariograma, as semivariâncias crescem sem limites. Isso indica que a hipótese de estacionariedade de 2ª ordem não foi atendida. Assim, muito provavelmente o fenômeno estudado tem capacidade infinita de dispersão (hipótese intrínseca). No semivariograma sem patamar, a máxima distância h entre as amostras não é capaz de exibir toda a variância contida nos dados. Pode-se ainda dizer com grande probabilidade que existe tendência nos dados para alguma direção. Essa tendência deve ser retirada para que possa ser fornecido um semivariograma

³ Efeito Pepita – variância em h muito próximo de zero. Krige percebeu que dentro de uma mesma pepita de ouro (distância ou lag $h = 0$) havia uma variação espacial, essa variação para distâncias próximas de 0 (zero) ficou conhecida por efeito pepita. Uma homenagem a Daniel G. Krige.

com patamar. Outro tipo de semivariograma é o **cíclico**. Ele apresenta flutuações nas semivariâncias. Por fim, temos um tipo de semivariograma onde as estruturas de semivariância são entrelaçadas. Esse caso de semivariograma geralmente ocorre em populações mistas, ou seja, uma população inicial que pode ser vista como duas ou mais populações, esse é o semivariograma **imbricado** ou **entrelaçado** (GUIMARÃES, 2004).

Na figura seguinte, retirada de Guimarães (2004), temos as representações dos semivariogramas citados:

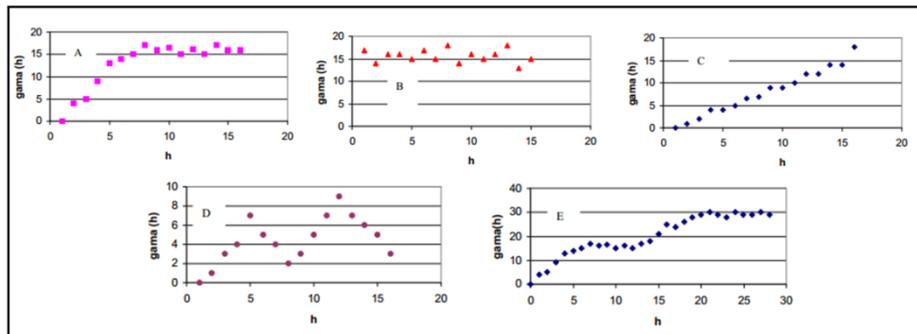


Figura 10 1- Com patamar 2- Pepita puro 3- Sem patamar 4- Cíclico 5- Imbricado

Fonte: Guimarães (2004).

2.9 Modelagem de Semivariograma

Após a construção de um semivariograma experimental, deve-se proceder ao ajuste de um modelo matemático que estimará os parâmetros desse semivariograma. Landim (2006) afirma que a modelagem do semivariograma, ou seja, o ajuste do semivariograma experimental por uma função matemática é uma passo fundamental na análise variográfica. O autor diz ainda que esse processo envolve várias tentativas e que a experiência do pesquisador tem um peso muito grande nesse ajuste.

Bohling (2005) afirma que, para o bem da geoestatística, é preciso que o semivariograma empírico adotado para os dados seja ajustado por um acertado modelo matemático. Segundo Mello (2004), esses modelos matemáticos de ajustes de semivariogramas, conhecidos por modelos teóricos, devem fornecer soluções estáveis para o estimador linear (*Krigagem*). As covariâncias devem ser definidas positivamente, para que a matriz de covariâncias possa ser invertida no processo da *Krigagem*. Essa condição de não negatividade do modelo reduz o número de funções utilizadas na modelagem dos semivariogramas empíricos.

Os modelos que possuem um patamar definido (*finite sill*) são conhecidos como modelos transitivos, ou modelos de transição e possuem funções de covariâncias dadas por:

$$\text{Cov}(h) = C(0) - \gamma(h) \quad (22)$$

Modelo Esférico (Matheron): Andriotti (1988) traz que a inclinação da tangente na origem ($h=0$) é $3C_1/2a$. Apresenta um rápido crescimento na origem, onde apresenta um comportamento linear. Clark (1979) afirma que o modelo esférico é conhecido como modelo ideal de semivariograma, é o modelo mais comum de ajuste, representando para a geoestatística o que a distribuição Normal é para a estatística. É também conhecido por modelo de Matheron.

$$\gamma(h) = \begin{cases} C_0 + C_1 \left[\frac{3}{2} \left(\frac{h}{a} \right) - \frac{1}{2} \left(\frac{h}{a} \right)^3 \right], & \text{se } 0 \leq h \leq a; \\ C_0 + C_1, & \text{se } h > a. \end{cases} \quad (23)$$

Modelo Exponencial: Para Andriotti (1988), a inclinação da tangente na origem é C_1/a . Nesse modelo a corresponde a 95% de patamar alcançado. Clark

(1979) afirma que o modelo exponencial, quando comparado ao modelo esférico, cresce mais lentamente na origem, apresenta continuidade nos primeiros valores de h e não chega a alcançar seu patamar. O patamar é atingido apenas assintoticamente.

$$\gamma(\mathbf{h}) = \begin{cases} C_0 + C_1 \left[1 - e^{\left(-3\frac{h}{a}\right)} \right], & \text{se } 0 \leq \mathbf{h} \leq a; \\ C_0 + C_1, & \text{se } \mathbf{h} > a. \end{cases} \quad (24)$$

Modelo Gaussiano (Parabólico): Curva parabólica próxima à origem, tangente horizontal nesse ponto. Possui um ponto de inflexão na sua curva. Isso implica pequena variabilidade para pequenas distâncias. Nesse modelo a corresponde a 95% de patamar alcançado.

$$\gamma(\mathbf{h}) = \begin{cases} C_0 + C_1 \left[1 - e^{-3\left(\frac{h}{a}\right)^2} \right], & \text{se } 0 \leq \mathbf{h} \leq a; \\ C_0 + C_1, & \text{se } \mathbf{h} > a. \end{cases} \quad (25)$$

A relação que se dá entre os alcances dos três modelos, até agora apresentados, segundo Andriotti (1988):

$$a \text{ (esférico)} < a \text{ (Gaussiano)} < a \text{ (exponencial)}$$

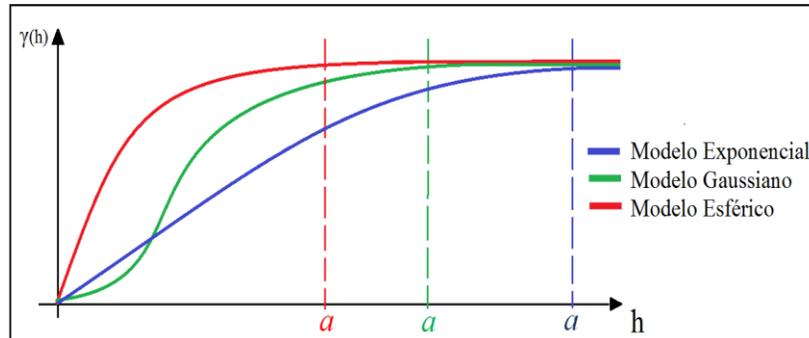


Figura 11 Semivariogramas Transitivos

Segundo Clark (1979), uma propriedade interessante sobre os modelos transitivos é que o valor do patamar, C , é igual à variância amostral. Tomando uma amostra aleatória de observações independentes pode-se calcular sua variância amostral por:

$$S^2 = \frac{1}{n-1} \sum_i^n (x_i - \bar{x}_i)^2, \text{ onde } \bar{x}_i = \frac{1}{n} \sum_i^n x_i \quad (26)$$

Então, S^2 e C serão ambos estimadores da mesma variância populacional.

Para Clark (1979), outro modelo existente é o Efeito Pepita Puro. Usado para descrever o semivariograma de efeito puro (fenômeno com uma baixa dependência espacial). Efetivamente, o efeito pepita é um modelo esférico com um alcance de influência muito pequeno. Seu modelo matemático é dado por:

$$\begin{aligned} \gamma(0) &= 0 \\ \gamma(h) &= C_0 \quad \text{quando } h > 0 \end{aligned} \quad (27)$$

Ainda, segundo Clark (1979), existem modelos que não atingem um patamar e continuam aumentando com a distância. O modelo linear é o mais simples deles:

$$\gamma(h) = C_0 + ph^\alpha \quad (28)$$

Onde p representa a inclinação da reta dada no semivariograma e $\alpha = 1$. Apresenta variância, *a priori*, infinita.

O nome de modelo linear generalizado é utilizado para $0 < \alpha < 2$. Alguns autores se referem a esses modelos como “funções *power*” (Andriotti, 1988).

Clark (1979) nos apresenta outro modelo sem patamar, o modelo de Wjjsian, em que o semivariograma é linear se representado graficamente contra o logaritmo da distância:

$$\gamma(h) = 3\alpha \log_e(h) \quad (29)$$

2.10 Auto Validação

Usualmente o ajuste do semivariograma é feito visualmente, o chamado “ajuste a sentimento”. Entretanto recomenda-se que sejam realizadas algumas técnicas existentes que ajudam na adequação da escolha de um modelo. As técnicas de autovalidação consistem em fazer comparações entre os valores amostrados e os valores estimados provenientes das interpolações espaciais.

Vieira (2000) afirma que, na análise geoestatística, existe sempre certo grau de incerteza sobre as hipóteses assumidas ou sobre os parâmetros ajustados nos modelos utilizados. Tal incerteza é o erro de estimativa. O erro de estimativa pode ser avaliado pelo procedimento de autovalidação ou “Jack-Knifing”. Essa

técnica consiste em ignorar, um por um, cada valor já amostrado e prever um valor interpolado pela técnica de *Krigagem*, utilizando-se de um modelo proposto. Segundo esse mesmo autor, a execução do “jack-knifing” permite verificar se a *Krigagem* é mesmo representativa da variabilidade e se as hipóteses assumidas são verdadeiras. Então, pode-se escolher o melhor modelo, dentre os modelos propostos, comparando os valores preditos e amostrados por meio de diferentes procedimentos. A seguir apresentam-se três deles.

2.10.1 Gráfico: Medido vs Estimado

Para Vieira (2000), se em todos os n locais amostrados tem-se um valor medido $Z(x_i)$ e se, também, são conhecidos os valores estimados pela *Krigagem* $Z^*(x_i)$, então é possível construir um gráfico dos valores pareados de $Z(x_i)$ e $Z^*(x_i)$. Pode-se calcular a regressão linear entre eles pela equação:

$$Z^*(x_i) = a + b Z(x_i) \quad (30)$$

onde a é a interseção, b é o coeficiente angular da reta. Assim sendo, se $Z(x_i)$ e $Z^*(x_i)$ são idênticos, então a seria nulo e b seria igual a 1 (um). O gráfico $Z^*(x_i) = f\{Z(x_i)\}$ seria uma série de pontos alinhados em 1:1 (VIEIRA, 2000).

2.10.2 Erro Médio, Erro Médio Reduzido e Desvio Padrão dos Erros

Com os valores estimados provenientes da autovalidação e os dados amostrados, pode-se conhecer a variância total da estimativa, isso permite o cálculo do erro de estimação, o Erro Médio. O erro médio indicará a eficácia do ajuste dos modelos de semivariograma, baseando-se no fato de que os erros,

provenientes do processo de *Krigagem* otimizado, apresentam média nula e variância mínima. Equação do Erro Médio:

$$EM = \frac{1}{n} \sum_{i=1}^n (Z(x_i) - Z^*(x_i)) \quad (31)$$

Outra maneira utilizada para se calcular os resíduos da autovalidação, é o erro médio reduzido ou padronizado (CRESSIE, 1993).

$$ER = \frac{1}{n} \sum_{i=1}^n \frac{ (Z(x_i) - Z^*(x_i)) }{ \sigma_i (x_i) } \quad (32)$$

Onde $\sigma_i (x_i)$ é o desvio padrão da *Krigagem* na posição x_i , não sendo considerada a observação $Z(x_i)$. Outro parâmetro importante, para as análises de qualidade de modelos de semivariograma por meio da autovalidação, é o desvio padrão dos erros, obtido pela equação:

$$S_{ER} = \sqrt{ \frac{1}{n} \sum_{i=1}^n \left\{ \frac{ (Z(x_i) - Z^*(x_i)) }{ \sigma_i (x_i) } \right\}^2 } \quad (33)$$

Para Cressie (1993), na busca pela escolha do melhor modelo de ajuste, deve-se procurar pelo modelo que possua EM e ER o mais próximo de zero e desvio padrão dos erros (S_{ER}) mais próximo de 1.

2.11 Distância útil variográfica

É comum deparar-se com semivariogramas que têm características de parâmetros bem comportados, mas que têm em seus extremos finais (últimos *lags*), semivariâncias discrepantes. Por isso, pensa-se logo na possibilidade de outliers ou valores amostrais que possam ser ignorados por razões diversas. Até que ponto (h) o semivariograma deve ser representado? Existe uma distância mínima de cálculo de semivariância para os dados amostrais?

A distância máxima para cálculo da semivariância deve ser, logicamente, à máxima distância de coleta da amostra. Porém, adotar menores distâncias de semivariâncias, como, por exemplo, 40, 50, 60 e 80% da distância máxima já podem ser utilizadas. A utilização de distâncias semivariográficas menores que a distância máxima são úteis para evitar trabalhar com semivariâncias finais discrepantes.

Isto se justifica pelo fato de que a grandes distâncias o número de pares para o cálculo da semivariância reduz-se drasticamente, fazendo com que a estimativa da semivariância tenha pouca precisão, esse valor de distância máxima pode ser alterado pelo pesquisador (GUIMARÃES, 2004, p. 37).

Contudo, é necessário levar em conta o bom senso com o georreferenciamento e os objetivos da análise espacial.

2.12 Grau de dependência espacial

A dependência espacial manifesta-se como uma correlação entre as amostras que diminui, à medida que a distância entre os pontos amostrais aumenta e desaparece completamente, quando as distâncias entre os pontos amostrais é suficientemente grande, os quais, neste caso, são

considerados estatisticamente independentes (AMARAL, 2010, p. 9).

O pesquisador geoestatístico, na observação do semivariograma modelado, pode determinar o grau de dependência espacial da variável em estudo. O grau de dependência espacial (GDE), segundo Cambardella et al. (1994), é simples e diretamente a razão entre a variância espacial C_1 e o patamar $C_0 + C_1$, dado em porcentagem.

$$GDE = \frac{C_1}{C_0 + C_1} \cdot 100\% \quad (34)$$

Cambardella et al. (1994) classificam o GDE da seguinte maneira:

Forte dependência espacial: Contribuição representa entre 75% e 100% do patamar.

$$0.75 < GDE < 1.00.$$

Moderada dependência espacial – Contribuição representa entre 25% e 75% do patamar.

$$0.25 \leq GDE \leq 0.75$$

Fraca dependência espacial – Contribuição representa menos que 25% do patamar.

$$GDE < 0.25$$

2.13 Tendência

Em uma variável regionalizada $Z(x_i) = \mu(x_i) + \varepsilon'(x_i) + \varepsilon''$, verifica-se que $\mu(x_i)$ representa a componente determinística estrutural de Z , ou seja, um valor médio ou uma tendência. Quando a média ou tendência desse fenômeno é constante, na região do estudo, e a variância é constante, dentro dos limites de continuidade espacial (estacionariedade), podemos utilizar as técnicas geoestatísticas de *Krigagem Ordinária e Simples* sem qualquer problema. Entretanto, existem situações especiais, onde essa tendência não é considerada constante. Usualmente fala-se que os dados são tendenciosos. Para Andriotti (1988), nesse tipo de situação, as semivariâncias, fatalmente, irão crescer mais rapidamente que h – indicativo de presença de deriva - e, nesse caso, o semivariograma experimental irá sobrestimar o semivariograma real. Bohling (2005) traz que, se os semivariogramas continuam subindo sem parar, mesmo além da variância global, isso é um forte indicativo de tendência espacial na variável, o que resulta em correlações negativas. Segundo o autor, pelo menos duas opções são recomendadas para os casos de dados tendenciosos:

- a) Ignorar o problema e seguir as análises utilizando uma modelagem linear ou potencial;
- b) Modelar a superfície de tendência e trabalhar com os resíduos da tendência. Para isso a geoestatística utiliza-se da *Krigagem Universal*, a qual será apresentada posteriormente.

Algumas maneiras de visualizar a tendência nos dados: Semivariograma apresenta uma curva acentuada em alguma direção que não seja a direção do patamar; Plotagem de gráficos das coordenadas individualmente vs variáveis respostas, nesse caso deve ser observado se há tendência na dispersão.

2.14 Anisotropia e Isotropia

Para Druck et al. (2004), a variabilidade ou distribuição espacial de certos fenômenos pode ocorrer mais intensamente numa direção e menos intensamente em outra direção. Essa variação da continuidade espacial em função da direção é conhecida como anisotropia. Segundo os autores: “A anisotropia é uma característica muito frequente nos elementos da natureza”.

Druck et al. (2004) trazem que existem várias formas para a detecção de anisotropia. Uma maneira simples e direta pode ser feita pelo cálculo dos semivariogramas experimentais em várias direções. As direções comumente adotadas são as direções 0° (na direção das ordenadas), 90° (direção das abscissas), 45° e 135° (diagonais principais). Então, plotam-se todos os semivariogramas em um único gráfico e visualmente faz-se a avaliação de suas similaridades ou dissimilaridades.

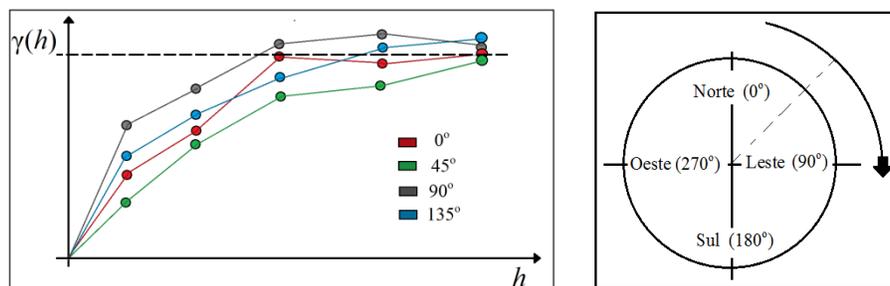


Figura 12 Semivariogramas direcionais

Como se observa, h é um vetor, logo possui magnitude e direção. Assim, o semivariograma dependerá dessa magnitude e direção. Caso seja observada a identidade, similaridade dos semivariogramas, em diferentes direções, pode-se concluir a isotropia do fenômeno (VIEIRA, 2000).

Para Isaaks e Srivastava (1989), nos fenômenos isotrópicos, os semivariogramas experimentais só dependem da distância h e não da direção. A modelagem de um semivariograma experimental omnidirecional é indicada nesses casos. Para Vieira (2000), os fenômenos isotrópicos representam situações de trabalho bem mais simples quando comparados aos fenômenos anisotrópicos.

2.15 Krigagem

Segundo Gaetan e Guyon (2009), o ponto central da geoestatística é a predição de $Z(x_i)$ em toda região de estudo R utilizando apenas um número finito de observações. Para Franklin (2014), a Krigagem é um método de predição espacial para pontos não amostrados que se utilizam das observações já amostradas e de seus respectivos pesos. Pesos que são, segundo Cressie (1993), atribuídos, baseados nas distâncias entre o ponto estimado e seus vizinhos e na própria distância entre esses vizinhos, de acordo com a variabilidade espacial expressa no semivariograma. Ou seja, o interpolador que utiliza o semivariograma em sua modelagem é conhecido como Krigagem. Segundo Isaaks e Srivastava (1989), a interpolação por Krigagem é o mais importante recurso geoestatístico.

Melo (2004) nos traz que este método foi desenvolvido por Georges Matheron na década de 1960. Seu nome é uma homenagem ao engenheiro de minas Daniel G. Krige que primeiramente introduziu esse método de inferência em 1951 em seus trabalhos. O método foi inicialmente utilizado na solução de mapeamentos geológicos, mas expandiu-se sendo muito utilizado em mapeamentos de solos, hidrológicos e atmosféricos.

Para Vieira (2000), a Krigagem se diferencia de outros métodos de interpolação pela maneira como atribui pesos às amostras. Por exemplo, na

interpolação linear simples, os pesos são todos iguais a $1/n$ (sendo n o número total de amostras). Já na interpolação pelo Inverso do quadrado da distância, os pesos são definidos como o inverso do quadrado da distância que separa o ponto a ser interpolado dos valores amostrados. A *Krigagem* é um estimador de média móvel ponderada, seus pesos são determinados pela análise espacial baseada no semivariograma experimental. Tal estimador é considerado ótimo pela maneira como os pesos são distribuídos. As formas mais conhecidas da família de *Krigagem* são: *Krigagem* Simples, *Krigagem* Ordinária e *Krigagem* Universal.

2.15.1 *Krigagem* Simples e *Krigagem* Ordinária

A *Krigagem* simples assume que a média é conhecida e constante em toda área de estudo. Assim, $\mu(x_i) = m$. Dificilmente trabalha-se com uma região de estudo totalmente homogênea, assim sendo, a *Krigagem* Simples, basicamente, é utilizada na ciência dos solos, onde as áreas estudadas são preparadas para sua homogeneidade.

A *Krigagem* ordinária (KO) não requer conhecimento prévio da média. Em vez de assumir que a média é constante em todo o domínio, pode-se assumir que ela é constante na vizinhança local de cada ponto de estimação, desde que se defina adequadamente a vizinhança local.

Franklin (2014) afirma que, no caso de se trabalhar com processos de estacionariedade intrínseca, com uma média constante desconhecida e sem presença de tendência, usa-se a interpolação da *Krigagem* ordinária, cujo estimador é:

$$Z^*(x_0) = \sum_{i=1}^n w_i Z(x_i) \quad (35)$$

w_i são os pesos ótimos da KO. Como já foi dito sobre as *Krigagens*, deseja-se trabalhar com os melhores estimadores não viesados. Assim encontra-se o best linear unbiased predictor (BLUP) pela minimização da variância do erro e pela condição de não viés. Para que a KO seja otimizada, seu estimador não pode ser tendencioso e deve ter variância mínima.

$$E [Z^*(x_0) - Z(x_0)] = E [Z^*(x_0)] - E [Z(x_0)] = \mu - \mu = 0 \quad (36)$$

$$\text{Var} [Z^*(x_0) - Z(x_0)] = E [(Z^*(x_0) - Z(x_0))^2] = \text{mínima} \quad (37)$$

Desenvolvendo a expressão da esperança do erro (36), chega-se à condição de não viés, também chamada de restrição:

$$\begin{aligned} E [Z^*(x_0)] - E [Z(x_0)] &= E \left[\sum_{i=1}^n w_i \cdot Z(x_i) \right] - E [Z(x_0)] \\ &= \sum_{i=1}^n w_i E [Z(x_i)] - E [Z(x_0)] = \sum_{i=1}^n w_i \mu - \mu \\ &= \mu \left(\sum_{i=1}^n w_i - 1 \right) = 0 \quad \rightarrow \quad \sum_{i=1}^n w_i = 1, \forall \mu \end{aligned} \quad (38)$$

As equações (36) e (37) representam, respectivamente, as condições de não tendência e variância mínima. Essas condições devem ser satisfeitas e, por isso, são utilizadas como princípios básicos na dedução das equações de *Krigagem* (VIEIRA, 2000). Assim, efetuando-se a minimização da variância do

erro de estimativa, sob a condição de não viés ou de restrição (38) com a utilização da técnica dos multiplicadores de Lagrange, gera-se o sistema de equações da KO escrito em termos da função de semivariograma (LANDIM, 2006):

$$\left\{ \begin{array}{l} \sum_{j=1}^n w_j \gamma(\mathbf{x}_i, \mathbf{x}_j) + \mathbf{L} = \gamma(\mathbf{x}_i, \mathbf{x}_0) \\ \sum_{j=1}^n w_j = 1 \end{array} \right. \quad i = 1, 2, \dots, n$$

(39)

A variância de *Krigagem* Ordinária:

$$\sigma_{\text{KO}}^2 = \sum_{i=1}^n w_i \gamma(\mathbf{x}_i, \mathbf{x}_0) + \mathbf{L}$$

(40)

Vieira (2000) traz que a resolução do sistema de *Krigagem* da equação (39) pode ser escrito em notação matricial como:

$$[A][w] = [B] \quad (41)$$

$$[w] = [A]^{-1}[B] \quad (42)$$

onde [A] é a matriz com as semivariâncias das diferenças entre todos os pontos utilizados na estimação de $Z^*(x_0)$; [w] a matriz coluna com os pesos e o multiplicador de Lagrange (L) e [B] é a matriz com as semivariâncias das diferenças entre os valores amostrados e o ponto a ser estimado.

O sistema (39) contém $n + 1$ equações e $n + 1$ incógnitas. Uma única solução desse sistema produz n pesos (w), e um multiplicador de Lagrange (L). As matrizes $[A]$, $[B]$ e $[w]$:

$$\begin{array}{ccc}
 [A] & [B] & [w] \\
 \left[\begin{array}{cccccc}
 \gamma(x_1, x_1) & \gamma(x_1, x_2) & \dots & \gamma(x_1, x_n) & 1 \\
 \gamma(x_2, x_1) & \gamma(x_2, x_2) & \dots & \gamma(x_2, x_n) & 1 \\
 \vdots & \vdots & & \vdots & \vdots \\
 \gamma(x_n, x_1) & \gamma(x_n, x_2) & \dots & \gamma(x_n, x_n) & 1 \\
 1 & 1 & \dots & 1 & 0
 \end{array} \right] & \left[\begin{array}{c}
 \gamma(x_1, x_0) \\
 \gamma(x_2, x_0) \\
 \vdots \\
 \gamma(x_n, x_0) \\
 1
 \end{array} \right] & \left[\begin{array}{c}
 w_1 \\
 w_2 \\
 \vdots \\
 w_n \\
 L
 \end{array} \right]
 \end{array}$$

Figura 13 Matrizes do sistema de *Krigagem* Ordinária

A equação (40) correspondente à variância da estimativa na KO, expressa em notação matricial:

$$\sigma^2 = [w]^t [B] \quad (43)$$

2.15.2 *Krigagem* Universal

Quando nos deparamos com uma média (μ) que não seja constante em nosso processo espacial, precisamos expandir nossos recursos para dar conta de uma média variável (FRANKLIN, 2014). Andriotti (1988) utiliza-se dos termos **deriva ou tendência** para designar essa média variável.

A *Krigagem* com tendência, formalmente conhecida como *Krigagem* Universal (KU), é essencialmente a *Krigagem* ordinária. Porém a *Krigagem* universal, não só estima uma média local na vizinhança do ponto de estimação,

como também modela uma tendência, de ordem linear ou maior, nas coordenadas (x,y) dos dados (BOHLING, 2005).

Polinômios lineares, quadráticos, cúbicos e outras funções podem ser utilizados na explicação (modelagem) da tendência. A modelagem da tendência de primeira ordem pode ser feita, segundo Bohling (2005):

$$\mu(x_i) = m(x,y) = a_0 + a_1x + a_2y \quad (44)$$

A modelagem da tendência de segunda ordem, por Landin e Corsi (2001):

$$\mu(x_i) = m(x,y) = a_0 + a_1x + a_2y + a_3x^2 + a_4xy + a_5y^2 \quad (45)$$

Na prática o uso direto da KU exige a construção dos modelos de tendência. A introdução desses modelos no sistema de equações lineares da KU dificulta a sua implementação prática. Por causa disso, Andriotti (1988) recomenda uma variação da técnica de KU, aproximadamente igual a essa, na qual, após a modelagem da tendência, a variável estudada é decomposta em duas partições. A primeira $\mu(x_i)$ é referente a uma tendência principal (drift) no ponto e a segunda (ε) é chamada de resíduo. Assim podemos escrever a variável $Z(x_i)$ como:

$$Z(x_i) = \mu(x_i) + \varepsilon(x_i) \quad (46)$$

Guimarães (2004) traz que se $\mu(x_i) = \text{constante}$, então, o semivariograma de $Z(x_i)$, usando as observações reais, será idêntico ao semivariograma dos resíduos $\varepsilon(x_i)$. Contudo, se existir tendência nos dados (linear, quadrática, cúbica, etc.), o semivariograma dos resíduos apresentar-se-á como uma melhor

estruturção para a definição dos parâmetros e pode produzir estimativas com menor variância. Aí se faz útil manobra na KU que subtrai a função de ajuste de tendência dos dados originais e obtém os erros (antes aleatórios) e, a seguir, o semivariograma é remodelado a partir dos erros. Diggle e Ribeiro Junior (2007) corroboram essa afirmação, recomendando que seja modelada e, posteriormente, retirada a tendência dos dados por regressão e, então, faz-se a análise dos resíduos pela *Krigagem* ordinária.

3 MATERIAL E MÉTODOS

3.1 Material

Os dados de temperaturas médias compensadas, utilizados nas análises geoestatísticas, foram obtidos por meio do INMET e na biblioteca da Universidade Federal de Lavras. Nas Normais 1931-1960, trabalhamos com 157 estações meteorológicas, obtidas em todos os estados brasileiros, já nas Normais 1961-1990 trabalhamos com 221 amostras. As análises estatísticas e geoestatísticas foram realizadas pelos Softwares ARCGIS 10.1 e R (versão 3.1.2, 2014).

3.2 Métodos

As análises iniciais foram exploratórias dos dados de temperaturas. Desse modo buscou-se: número de estações, média, mediana e moda, valores máximos e mínimos, histograma, variância e desvio padrão dos dados.

Numa segunda etapa, fez-se a construção da dispersão dos dados em mapas para que pudessem observados possíveis aglomerados, padrões e tendências.

Em seguida, iniciou-se a construção dos semivariogramas empíricos por meio dos modelos de Matheron e Cressie e Hawkins até as distâncias plausíveis. Os modelos empíricos foram, então, modelados pelos modelos exponencial, gaussiano e esférico. A seguir, por auto-validação, os modelos de ajustes foram comparados e aqueles escolhidos (apresentaram menores erros e variâncias) foram utilizados nos processos de *Krigagem* respectivos.

No momento final da análise realizou -se a *Krigagem* para cada uma das Normais, o que originaram dois mapas interpolados do território brasileiro. O

último passo consistiu na construção de um mapa com as diferenças entre as temperaturas estimadas nos dois períodos. Esse mapa permitiu ao pesquisador avaliar se existiam mudanças de temperaturas médias entre os períodos das Normais estudadas

4 RESULTADOS E DISCUSSÃO

Iniciaram-se as análises estatísticas e geoestatísticas com o intuito de gerar conclusões sobre a distribuição espacial da variável em estudo.

4.1 Análise preliminar dos dados

Inicialmente foi feita uma análise sobre as características dos dados por meio da ferramenta *summary* do programa R. Na tabela a seguir pôde ser feita uma comparação inicial entre os dados das Normais 1931-1960 e 1961-1990. Os dados das Normais 1961-1990 apresentaram outliers, esses foram retirados.

Tabela 1 Análise exploratória de Dados para as Normais 1931-1960 e 1960-1990

Normais	Número de amostras	Mínimo	Máximo	Média	Mediana	1º Quartil	3º Quartil	Coefficiente de Variação
1931-1960	157	14,50	28,00	21,99	22,20	19,30	25,00	2,10%
1961-1990	219	14,70	29,40	22,99	23,40	20,80	25,55	1,34%

A seguir os histogramas gerados pelo R correspondentes às normais em estudo:

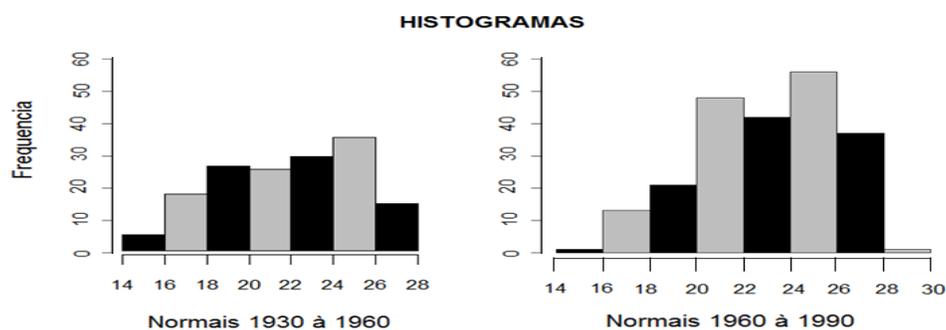


Figura 14 Histogramas 1930-1960 e 1960-1990

A seguir apresentam-se as dispersões das duas Normais. As dispersões buscam auxiliar na percepção, indicação de tendências.

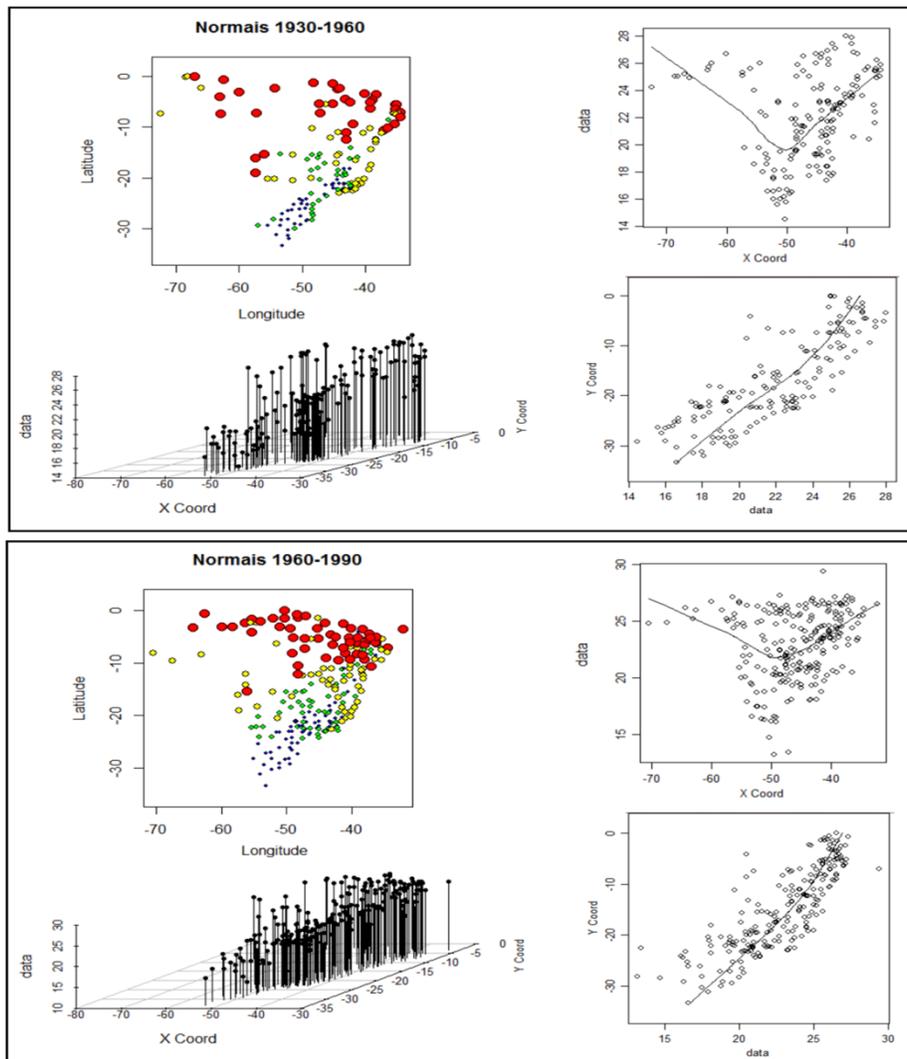


Figura 15 Dispersões para as Normais 1931-1960 e 1961-1990

4.2 Discussões Iniciais

Uma análise da tabela 1 nos permite verificar mudanças do período de 1931-1960 para 1961-1990:

- a) Aumento do número de amostras (estações de coleta);
- b) As temperaturas mínima e máxima sofreram aumento;
- c) Aumento de média e mediana de 1931-1960 para 1961-1990;
- d) Quartis foram elevados;
- e) Coeficiente de variação (CV) diminuiu.

O aumento no número de amostras é evidentemente esperado por vários motivos como: aumento da população e ocupação de áreas maiores do país, avanço tecnológico, preocupação em se coletar amostras em mais pontos para melhorar a qualidade das análises.

A elevação das temperaturas, mínima, máxima, mediana e média sugerem uma elevação da temperatura no país. Como essas informações são provenientes de uma simples análise exploratória, faz-se o uso de técnicas da geoestatística para corroborar esses resultados iniciais.

A diminuição do coeficiente de variação indica uma menor variabilidade da média nos dados de 1961-1990 em relação à média nos dados de 1931-1990.

Pela análise dos histogramas das Normais de 1931-1960 e 1961-1990:

- a) Percebe-se claramente que as classes de temperaturas mais baixas (entre 14 e 20°C) tiveram uma significativa diminuição;
- b) As classes de temperaturas médias e altas (20 a 28°C) tiveram uma significativa elevação;

- c) Destacam-se, pela alta frequência, três classes nas normais 1961-1990. Essas classes têm temperaturas variando de 20 a 28 °C ;
- d) Surgimento de uma nova classe para temperaturas entre 28 e 30°C.

Analisando os Mapas de Dispersão:

- a) As regiões centro-oeste e norte têm o número de amostras bem menor que as demais regiões;
- b) Os mapas de dispersão ajudam na percepção da disposição das amostras no mapa do território brasileiro. Pode-se ter uma ideia sobre possíveis regiões de temperaturas mais semelhantes, pois as estações recebem cores diferentes, de acordo com o quartil a que pertencem (azul para 1ºquartil, verde para 2ºquartil, amarelo para 3ºquartil e vermelho para 4ºquartil);
- c) É bem notável, nas duas Normais, a presença de uma tendência quadrática na direção longitudinal e uma tendência linear na direção latitudinal. Porém, percebe-se que ambas as tendências são suavizadas das Normais de 1931-1960 para as de 1961-1990, contudo continuam existindo e, por isso, elas foram modeladas para a realização das *Krigagens*.

4.3 Análise Espacial

- a) A tendência de segunda ordem foi retirada nos dados das duas Normais para que os semivariogramas pudessem ser utilizados futuramente nas *Krigagens*;

- b) Semivariogramas omnidirecionais foram adotados, pois o software adotado nas análises (R) não realizou estudo aprofundado anisotrópico;
- c) Os semivariogramas foram construídos no R, utilizando-se dos modelos empíricos de Matheron e Cressie e Hawkins nas distâncias de 40, 50, 60, 70, 80, 90 e 100% da distância máxima. Os modelos foram comparados visualmente e aqueles com melhores relações de crescimento semivariográfico até atingir um patamar foram adotados:
- Cressie e Hawkins com 60% da distância máxima para as normais 1931-1960;
 - Cressie e Hawkins com 40% da distância máxima para as normais 1961-1990;

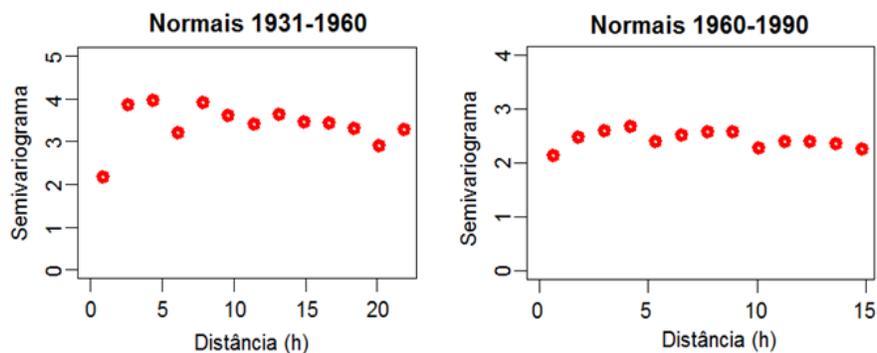


Figura 16 Semivariogramas empíricos para as Normais 1931-1960 e 1961-1990

Para as modelagens dos semivariogramas empíricos foram testados os modelos exponencial, gaussiano e esférico. A figura seguinte ilustra os ajustes dos modelos para cada Normal.

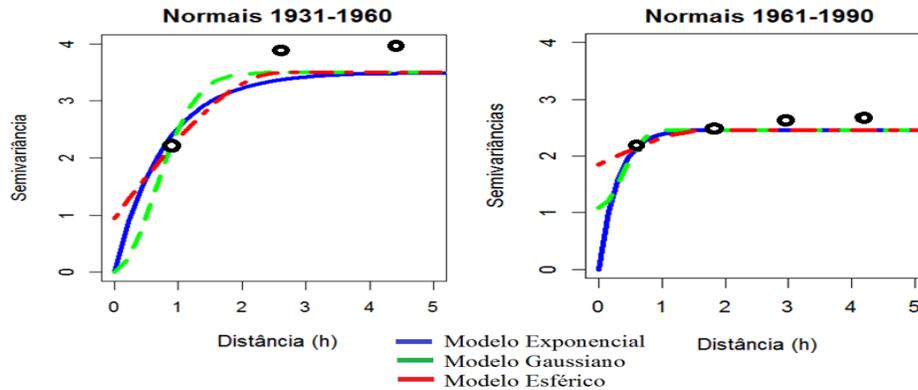


Figura 17 Semivariogramas modelados para as Normais 1931-1960 e 1961-1990

Para concluirmos a utilização de um desses modelos utilizamos as informações provenientes das tabelas 2 e 3.

Tabela 2 Informações sobre os parâmetros, erros e inclinação da reta de autovalidação dos modelos testados para as Normais 1931-1960

Modelos	a	C_0	C_1	GDE	EM	ER	Inclinação
“Exp.”	0,781	0	3,496	100%	0,0176	0,0018	1,017
“Gaus.”	0,886	0	3,496	100%	0,5484	1,8046	0,1046
“Esf.”	2,589	0,920	2,577	73,7%	0,0228	0,0066	1,0127

Tabela 3 Informações sobre os parâmetros, erros e inclinação da reta de autovalidação dos modelos testados para as Normais 1961-1990

Modelos	a	C_0	C_1	GDE	EM	ER	Inclinação
“Exp.”	0,288	0	2,463	100%	0,0036	0,0013	0,9934
“Gaus.”	0,484	1,046	1,417	57,5%	0,0035	0,0012	0,992
“Esf.”	1,776	1,852	0,611	24,8%	0,0044	0,0015	1,0015

A qualidade de nossos modelos ajustados foi verificada pela autovalidação, conforme nos indica Vieira (2000). Realizou-se a técnica de “Jack Knifing”. Cada ponto amostrado foi também estimado por meio da *Krigagem*. Isso permitiu o cálculo de regressão linear entre os pares de dados. Os melhores ajustes foram obtidos, quando os valores do coeficiente angular (inclinação da reta) foram os mais próximos de 1 (um) e os erros médio e padronizado foram os mais próximos de 0 (zero). A qualidade de estimativa dos modelos pode ser medida pelo julgamento desses parâmetros (VIEIRA, 2000). Assim, os modelos de ajuste escolhidos para representarem a distribuição espacial das variáveis de Normais de temperaturas médias foram:

- a) Exponencial com parâmetros: 0,781 de alcance, 0 (zero) de efeito pepita e 3,486 de contribuição espacial para as Normais 1931-1960;
- b) Exponencial com parâmetros: 0,288 de alcance, 0 (zero) de efeito pepita e 2,463 de contribuição espacial para as Normais 1961-1990.

4.4 *Krigagens das Normais*

A *Krigagem* das Normais foi o passo final das análises deste trabalho. É importante que os modelos de semivariograma ajustados tenham sido bem escolhidos, pois a qualidade da *Krigagem* depende diretamente desse fato.

Como podem ser notados na figura 15, os dados, tanto para as Normais 1931-1960, tanto para as Normais 1961-1990, possuem tendência linear na direção latitudinal e uma tendência quadrática na direção longitudinal (é bom notar que essas tendências sofreram uma amenização, porém continuam existindo). Por isso, a *Krigagem* universal foi adotada, pois ela considera a tendência presente nos dados para o cálculo dos pesos de *Krigagem*.

Realizou-se, então, para cada uma das Normais climatológicas, uma *Krigagem* Universal que estimou valores de temperaturas médias para toda a região do território brasileiro em cada um desses períodos. Como o intuito do trabalho foi verificar uma possível mudança nas temperaturas médias no Brasil, durante os períodos das Normais adotadas, foi criado um mapa do território estudado com as diferenças existentes entre os mapas de *Krigagens* respectivos aos períodos de 1931-1960 e 1961-1990.

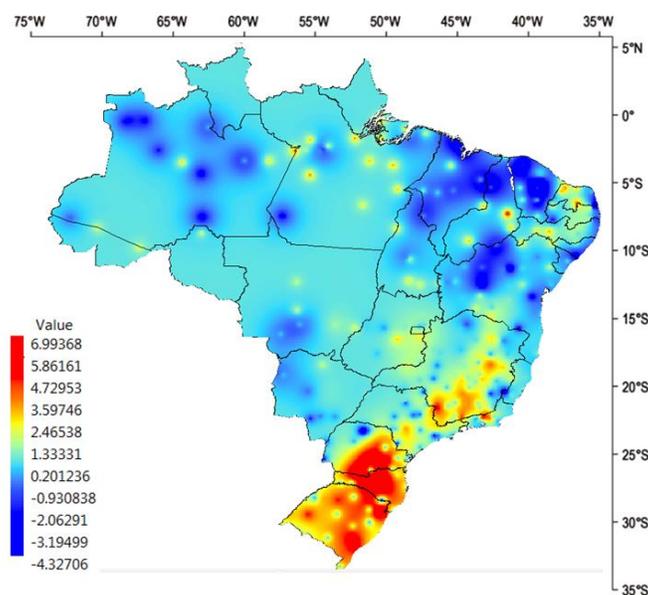


Figura 18 Mapa das diferenças de temperaturas médias para as Normais estudadas

4.5 Discussões finais

Pela análise do mapa de diferenças de temperaturas (Figura 18) observaram-se algumas características pertinentes às conclusões deste trabalho:

- a) As regiões Norte e Centro-Oeste formam um aglomerado de coloração mais homogênea. Tal aglomerado apresenta-se num tom de azul claro que denota por mudanças de temperaturas entre 0,20 e 1,33°C de acréscimo;
- b) Na região Nordeste, os estados do Maranhão e Ceará apresentam coloração azul e azul escuro, em grande parte do território, assim como a região norte do Piauí e centro-norte da Bahia. Essas áreas apontam por uma mudança das médias de temperatura entre -0,93 e -2,06°C;
- c) Rio Grande do Norte e o centro de Pernambuco têm áreas em amarelo, isso aponta por acréscimos de 2,46 a 3,6°C. O restante da região Nordeste tem coloração azul claro apontando acréscimos de 0,20 a 1,33°C;
- d) Na região Sudeste, os estados de São Paulo, Rio de Janeiro e Espírito Santo apresentam, na maior parte do território, a coloração de azul claro, o que indica mudanças entre 0,20 e 1,33°C de acréscimo. No estado de São Paulo, há uma mancha amarela localizada no centro, o que aponta acréscimos de 2,46 a 3,6°C. Essa mancha atinge as estações de Manduri e Jacarezinho. No Estado do Rio de Janeiro existe uma mancha alaranjada, o que aponta acréscimos de 3,6 a 4,73°C. A mancha localiza-se em volta da capital do estado;
- e) O estado de Minas Gerais possui a região mais heterogênea. As microrregiões do Sul de Minas, Zona da Mata, Centro (Campos das Vertentes), Rio Doce e Jequitinhonha apresentam manchas em amarelo e laranja, o que indica aumentos de temperatura em torno de 2,46 até 4,72°C. O restante das microrregiões (Centro Oeste, Triângulo, Alto Paranaíba, Noroeste e Norte) apresenta coloração azul claro (acrécimos de 0,20 e 1,33°C);

- f) A região Sul do país é a que mais chama atenção. Apresenta acréscimos acentuados de temperatura em praticamente toda sua extensão. Os acréscimos são indicados pelas colorações em laranja e vermelho (acrécimos de 3,6 a 6,99°C). Somente parte localizada no noroeste do Paraná não apresenta esse padrão de aumento tão elevado;
- g) Faz-se uso da análise dos mapas já apresentados nas figuras (1) e (2) para que sejam discutidos possíveis fatores de mudanças do clima no período estudado.

5 CONCLUSÕES

A análise do mapa de diferenças nos permite verificar uma mudança das temperaturas médias no Brasil entre os períodos de 1931-1960 e 1961-1990. A grande maioria do território apresenta mudanças por acréscimos de temperaturas, o que pode ser notado visualmente por todas as áreas cobertas pelas cores: azul claro, verde claro, amarelo, laranja e vermelho. As áreas em azul e azul escuro ocupam pequenas porções do território brasileiro. Assim, poucas áreas do Brasil tiveram diminuição das médias de temperaturas.

As duas grandes manchas indicativas de diminuição de médias de temperatura por volta de $-0,93$ e $-3,19^{\circ}\text{C}$ localizam-se no Nordeste brasileiro, uma nos estados do Maranhão, norte do Piauí e Ceará e a segunda na região centro-norte da Bahia.

Na região Sudeste, o estado de Minas Gerais se destaca por apresentar uma faixa de aumento entre $2,46$ e $4,73^{\circ}\text{C}$. Essa faixa inicia-se largamente no extremo sul do estado e vai se afunilando para o norte, na divisa do estado com a Bahia. Os climas da zona subtropical úmida de verão temperado e de verão quente (Figura 1) posicionam-se, coincidentemente, nessa área, o que nos leva a crer que tais climas foram mais afetados pelas mudanças de temperatura do período.

A região Sul destaca-se por ter grandes elevações de temperatura. Pode-se notar que a mancha na região Sul em amarelo e vermelho, que denota acréscimos de $2,46$ a $6,99^{\circ}\text{C}$, posiciona-se sobre o clima subtropical úmido (Figura 2). Observa-se que somente essa região possui o clima controlado por massas de ar tropicais e polares, sendo afetado pelas correntes marinhas formadas em regiões polares. Assim, o aumento das temperaturas das correntes polares poderia ser um dos fatores contribuintes da elevação das temperaturas na região Sul do país.

Pensando na região do território brasileiro como um todo, há uma forte indicação de aumento de temperaturas médias em torno de 0,2 a 1,33°C. Isso é perceptível pela coloração azul claro que ocupa grandes áreas em todo o mapa. Quando se analisa o gráfico do BEST da Figura 6, pode-se notar um acréscimo médio de 0,55°C nas temperaturas mundiais. Assim, pode-se dizer que o Brasil sofreu mudanças de acréscimos de médias de temperatura do período de 1930 a 1990 compatíveis com as mudanças observadas num contexto mundial. Já as áreas da Região Sul, Minas e as duas manchas de decréscimo do Nordeste, que tiveram mudanças de temperatura fora do padrão da análise do BEST, possivelmente sofreram com fatores naturais (correntes marinhas, climas mais propensos a mudanças climáticas) e ou sociais (queimadas, desmatamentos, criação de pastagens, etc.) que provocaram, e ou, intensificaram essas distorções.

REFERÊNCIAS

- ALVARES, C. A. et al. Koppen's climate classification map for Brazil. **Meteorologische Zeitschrift**, Berlin, v. 22, n. 6, p. 711-728, Jan. 2014.
- AMARAL, L. P. **Geoestatística na caracterização do solo e da vegetação em floresta ombrófila mista**. 2010. 133 p. Dissertação (Mestrado em Agronomia) - Universidade Estadual do Centro-Oeste, Guarapuava, 2010.
- ANDRIOTTI, J. L. S. **Introdução à geoestatística**. Porto Alegre: Acta Geológica Leopoldensia, 1988. 81 p.
- BARCELLOS, C. et al. Mudanças climáticas e ambientais e as doenças infecciosas: cenários e incertezas para o Brasil. **Epidemiologia e Serviços de Saúde**, Brasília, v. 18, n. 3, set. 2009. Disponível em: <http://scielo.iec.pa.gov.br/scielo.php?script=sci_arttext&pid=S1679-49742009000300011&lng=pt&nrm=iso>. Acesso em: 4 mar. 2015.
- BOHLING, G. **Introduction to geostatistics variograms analysis**. Survey: Kansas Geological, 2005. 20 p.
- BRASIL. Ministério da Agricultura. **Normais climatológicas**. Rio de Janeiro, 1969. 5 v.
- BURROUGH, P. A. **Principles of geographical information systems for land resources assessment**. Oxford: Clarendon, 1987. 193 p.
- CAMARGO, E. C. G. Geoestatística: fundamentos e aplicações. In: _____. **Geoprocessamento em projetos ambientais**. São José dos Campos: INPE, 2001. cap. 5, p. 1-36.
- CAMBARDELLA, C. A. et al. Field scale variability of soil properties in Central Iowa soils. **Soil Science Society American Journal**, Madison, v. 47, n. 5, p. 1501-1511, Sept. 1994.
- CARBON BRIEF. **Berkeley earth surface temperature (BEST) study**. Disponível em: <[http://www.carbonbrief.org/profiles/berkeley-earth-surface-temperature-\(best\)-study/](http://www.carbonbrief.org/profiles/berkeley-earth-surface-temperature-(best)-study/)>. Acesso em: 10 dez. 2013.
- CLARK, I. **Practical geostatistics**. London: Applied Science, 1979. 430 p.

CRESSIE, N. A. C. **Statistics for spatial data**. New York: J. Wiley, 1993. 928 p.

CRESSIE, N. A. C.; HAWKINS, D. M. **Robust estimation of the variogram: mathematical geology**. New York: J. Wiley, 1993. 900 p.

DIGGLE, P. J.; RIBEIRO JUNIOR, P. J. **Model-based geostatistics**. New York: Springer, 2007. 228 p.

DRUCK, S. et al. (Ed.). **Análise espacial de dados geográficos**. Brasília: EMBRAPA, 2004. Disponível em: <<http://www.dpi.inpe.br/gilberto/livro/analise/>>. Acesso em: 10 nov. 2014.

ESTADOS e capitais do Brasil: listas das siglas, bandeiras, regiões, estados e capitais do Brasil. Disponível em: <<http://www.estadosecapitaisdobrasil.com/>>. Acesso em: 10 nov. 2014.

FRANKLIN, M. **Solution to ordinary and Universal Kriging Equations**. Disponível em: <<http://www-hsc.usc.edu/~mereditf/KrigingEquations.pdf>>. Acesso em: 10 nov. 2014.

GAETAN, C.; GUYON, X. **Spatial statistics and modeling**. New York: Springer, 2010. 302 p.

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

HICKMAN, L. **Novo estudo sobre clima muda opinião de cientistas céticos**. Disponível em: <<http://www.oeco.org.br/guardian-environment-network/26295-novo-estudo-sobre-clima-muda-opiniao-de-cientistas-ceticos>>. Acesso em: 10 dez. 2012.

INSTITUTO BRASILEIRO DE GEOGRAFIA E ESTATÍSTICA. **Contagem populacional**. Disponível em: <<http://www.ibge.gov.br>>. Acesso em: 10 nov. 2014.

INSTITUTO NACIONAL DE METEOROLOGIA. **Normais climatológicas do Brasil: período 1961-1990**. Disponível em: <<http://www.inmet.gov.br>>. Acesso em: 10 nov. 2014.

ISAAKS, E. H.; SRIVASTAVA, R. M. **Applied geostatistics**. New York: Oxford University, 1989. 592 p.

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

KOPPEN, W. Das geographische system der klimate. In: KOPPEN, W. R.; GEIGER, R. (Ed.). **Handbuch der klimatologie**. Berlin: Gebruder Borntrager, 1936. p. 1-44.

LANDIM, P. M. B. Sobre geoestatística e mapas. **Terrae Didática**, Campinas, v. 2, n. 1, p. 19-33, 2006.

LANDIM, P. M. B.; CORSI, A. C. **Cálculo de superfícies de tendência, por regressão polinomial, pelo SURFER 6**. Rio Claro: UNESP, 2011. 11 p. (Texto Didático, 5).

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

MARENGO, J. A. Água e mudanças climáticas. **Estudos Avançados**, São Paulo, v. 22, n. 63, p. 83-96, 2008. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0103-40142008000200006&lng=pt&nrm=iso>. Acesso em: 25 abr. 2015.

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

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

NOBRE, C. A. **Fundamentos científicos das mudanças climáticas**. São José dos Campos: INPE, 2012. 44 p.

RIVOIRARD, J. **Concepts and methods of geostatistics**. Paris: Centre de Géostatistique, Ecole des Mines de Paris, 2005. 37 p.

ROHDE, R. et al. A new estimate of the average earth surface land temperature spanning 1753 to 2011. **Geoinformatics & Geostatistics: An Overview**, Foster City, v. 1, n. 1, p. 1-7, 2013.

SATHLER, D.; REIS, R. Mudanças climáticas globais: as dimensões humanas em perspectiva. **Revista Brasileira de Estudos de População**, São Paulo, v. 27, n. 1, p. 237-238, jun. 2010. Disponível em: <http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0102-30982010000100015&lng=en&nrm=iso>. Acesso em: 25 abr. 2015.

VICENTE, J. **Estudo comparativo de métodos geoestatísticos aplicados em agricultura de precisão**. 2004. 163 p. Dissertação (Mestrado em Ciências Cartográficas) - Universidade Estadual Paulista, Presidente Prudente, 2004.

VIEIRA, S. R. Geoestatística em estudos de variabilidade espacial do solo. In: NOVAIS, R. F.; ALVAREZ, V. H.; SCHAEFER, G. R. (Ed.). **Tópicos em ciência do solo**. Viçosa, MG: Sociedade Brasileira de Ciência do Solo, 2000. v. 1, p. 1-54.

YAMAMOTO, J. K.; LANDIM, P. M. B. **Geoestatística: conceitos e aplicações**. São Paulo: Oficina de Textos, 2013. 215 p.