



MINISTÉRIO DA CIÊNCIA, TECNOLOGIA E INOVAÇÃO
INSTITUTO NACIONAL DE PESQUISAS ESPACIAIS

sid.inpe.br/mtc-m21b/2014/05.07.17.24-TDI

APLICAÇÃO DO MÉTODO DE KRIGING PARA ESTIMAR CAMPOS DE RADIAÇÃO SOLAR: UM ESTUDO PARA O NORDESTE DO BRASIL

Roque Magalhães Brito dos Santos

Dissertação de Mestrado do Curso de Pós-Graduação em Meteorologia, orientada pelo Dr. Enio Bueno Pereira, aprovada em 26 de maio de 2014.

URL do documento original:

<<http://urlib.net/8JMKD3MGP5W34M/3G9DAR2>>

INPE
São José dos Campos
2014

PUBLICADO POR:

Instituto Nacional de Pesquisas Espaciais - INPE

Gabinete do Diretor (GB)

Serviço de Informação e Documentação (SID)

Caixa Postal 515 - CEP 12.245-970

São José dos Campos - SP - Brasil

Tel.:(012) 3208-6923/6921

Fax: (012) 3208-6919

E-mail: pubtc@sid.inpe.br

CONSELHO DE EDITORAÇÃO E PRESERVAÇÃO DA PRODUÇÃO INTELLECTUAL DO INPE (RE/DIR-204):**Presidente:**

Marciana Leite Ribeiro - Serviço de Informação e Documentação (SID)

Membros:

Dr. Gerald Jean Francis Banon - Coordenação Observação da Terra (OBT)

Dr. Amauri Silva Montes - Coordenação Engenharia e Tecnologia Espaciais (ETE)

Dr. André de Castro Milone - Coordenação Ciências Espaciais e Atmosféricas (CEA)

Dr. Joaquim José Barroso de Castro - Centro de Tecnologias Espaciais (CTE)

Dr. Manoel Alonso Gan - Centro de Previsão de Tempo e Estudos Climáticos (CPT)

Dr^a Maria do Carmo de Andrade Nono - Conselho de Pós-Graduação

Dr. Plínio Carlos Alvalá - Centro de Ciência do Sistema Terrestre (CST)

BIBLIOTECA DIGITAL:

Dr. Gerald Jean Francis Banon - Coordenação de Observação da Terra (OBT)

REVISÃO E NORMALIZAÇÃO DOCUMENTÁRIA:

Maria Tereza Smith de Brito - Serviço de Informação e Documentação (SID)

Yolanda Ribeiro da Silva Souza - Serviço de Informação e Documentação (SID)

EDITORAÇÃO ELETRÔNICA:

Maria Tereza Smith de Brito - Serviço de Informação e Documentação (SID)

André Luis Dias Fernandes - Serviço de Informação e Documentação (SID)



MINISTÉRIO DA CIÊNCIA, TECNOLOGIA E INOVAÇÃO
INSTITUTO NACIONAL DE PESQUISAS ESPACIAIS

sid.inpe.br/mtc-m21b/2014/05.07.17.24-TDI

APLICAÇÃO DO MÉTODO DE KRIGING PARA ESTIMAR CAMPOS DE RADIAÇÃO SOLAR: UM ESTUDO PARA O NORDESTE DO BRASIL

Roque Magalhães Brito dos Santos

Dissertação de Mestrado do Curso de Pós-Graduação em Meteorologia, orientada pelo Dr. Enio Bueno Pereira, aprovada em 26 de maio de 2014.

URL do documento original:

<<http://urlib.net/8JMKD3MGP5W34M/3G9DAR2>>

INPE
São José dos Campos
2014

Dados Internacionais de Catalogação na Publicação (CIP)

Santos, Roque Magalhães dos.

Sa59a Aplicação do método de kriging para estimar campos de radiação solar: um estudo para o nordeste do Brasil / Roque Magalhães Brito dos Santos. – São José dos Campos : INPE, 2014.
xxvi + 88 p. ; (sid.inpe.br/mtc-m21b/2014/05.07.17.24-TDI)

Dissertação (Mestrado em Meteorologia) – Instituto Nacional de Pesquisas Espaciais, São José dos Campos, 2014.

Orientador : Dr. Enio Bueno Pereira.

1. Radiação solar. 2. Interpolação kriging. 3. Energias renováveis. I. Título.

CDU 521.521.1(812/813)



Esta obra foi licenciada sob uma Licença [Creative Commons Atribuição-NãoComercial 3.0 Não Adaptada](https://creativecommons.org/licenses/by-nc/3.0/).

This work is licensed under a [Creative Commons Attribution-NonCommercial 3.0 Unported License](https://creativecommons.org/licenses/by-nc/3.0/).

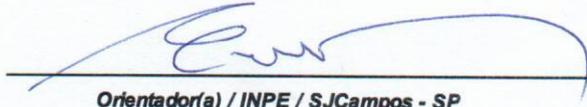
Aprovado (a) pela Banca Examinadora
em cumprimento ao requisito exigido para
obtenção do Título de **Mestre** em
Meteorologia

Dr. Juan Carlos Ceballos



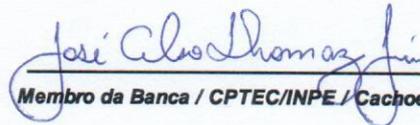
Presidente / INPE / Cachoeira Paulista - SP

Dr. Enio Bueno Pereira



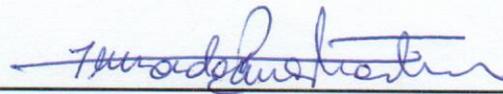
Orientador(a) / INPE / SJCampos - SP

Dr. José Celso Thomaz Júnior



Membro da Banca / CPTEC/INPE / Cachoeira Paulista - SP

Dr. Fernando Ramos Martins



Convidado(a) / UNIFESP / Santos - SP

Este trabalho foi aprovado por:

maioria simples

unanimidade

Aluno (a): **Roque Magalhães Brito dos Santos**

São José dos Campos, 26 de Maio de 2014

“A persistência é o menor caminho para o êxito”

Charles Chaplin

Esse trabalho é dedicado a todos os meus amigos, pessoas que foram importantes e contribuíram de alguma forma para esta realização.

Principalmente, dedico este trabalho à Adriana, ... desde adolescentes fomos amigos, companheiros, namorados e, hoje, somos uma família.

... à Bianca, hoje, razão da minha existência.

AGRADECIMENTOS

Ao meu orientador, Dr. Enio Pereira e ao Dr. Fernando Martins, por todo o apoio e confiança depositada. Especialmente, pela compreensão e paciência de uma orientação em grande parte à distância.

À Pós-Graduação em Meteorologia, em especial a Dra. Maria Paulete pela coordenação, e aos professores pelo apoio e suporte; às estimadas secretárias Luana e Simone pela paciência e colaboração. À CAPES, pelo apoio financeiro e ao Instituto Nacional de Meteorologia (INMET) pela disponibilização dos dados.

A todos os meus amigos, parceiros de trabalho e à grande família que constitui nessa etapa. Especialmente, ao Francisco (ô Chico), muitas vezes meu mestre nessa caminhada, assim como todos os amigos do laboratório: Chica, Lucía, Silvia, Alice e Rafael. Agradecimentos leais ao Jefferson Souza pelas conversas, amparos e dicas em programação.

Aos amigos Denis e Fabio pelas infinitas horas de bate papo e trocas de experiências em pesquisa e docência.

À Adriana, pelo apoio absoluto em todas as etapas da minha vida, pela paciência dedicada, pelo amor incondicional!

RESUMO

Foram feitas estimativas de campos de irradiação solar em superfície para o território do Nordeste brasileiro para os anos de 2008 a 2011. Para tanto, foi empregado um método de estimativa baseado em interpolação por krigagem. Para avaliação de seu desempenho, foi feita uma intercomparação dos valores estimados por interpolação com valores medidos em pontos amostrais representados por Plataformas Coletoras de Dados (PCDs) e estimativas realizadas pelo modelo físico de transferência radiativa BRASIL-SR. A avaliação foi realizada através de índices estatísticos aliados ao método de validação cruzada e a análise visual de mapas da variabilidade espacial da irradiação solar. Os dados brutos das PCDs foram selecionados e organizados com base em uma análise qualitativa dos mesmos. Os dados medidos passaram pelo mesmo procedimento de controle de qualidade adotado pela *Baseline Surface Radiation Network* (BSRN). Na interpolação por krigagem, o ajuste cuidadoso do semivariograma é fundamental para o bom desempenho dessa técnica. As séries de dados amostrais não apresentaram anisotropia e tendência significativas e não influenciaram nas estimativas por interpolação. Os resultados mostraram que as interpolações por krigagem foram capazes de explicar mais de 70 % da variabilidade da irradiância medida pelas PCDs na série de 2010 e 60 % na série de 2011. Considerando o período todo, o Erro Quadrático Médio Absoluto – EQMA ponderado pela média dos valores medidos – apresentou valores menores em torno de 5 % de erro, o que caracteriza um desempenho admissível do método de krigagem mostrando que os desvios foram pouco acentuados. Ainda, esses resultados estatísticos qualificaram a metodologia empregada na adequação dos dados utilizados nesse estudo. A análise visual dos mapas de irradiação para os 4 anos nos pontos amostrais, permitiu verificar que houve pouca variabilidade espacial da irradiância em superfície com níveis de irradiância média anual em torno de 6000 Wh/m². No entanto, em todos os mapas pôde-se observar um gradiente acentuado na região norte do Estado da Bahia fato esse, que pode ser atribuído a anomalias nas séries de dados medidos pelas PCDs e que passaram pela qualificação empregada. O modelo BRASIL-SR apresentou um desempenho melhor comparado ao método de krigagem na intercomparação com os dados medidos. O coeficiente de determinação r^2 apresentou valor próximo de 0,74, mostrando que 74 % da variabilidade dos dados medidos foi explicada pelas estimativas do modelo. Os resíduos indicaram que o modelo superestimou os valores em 11, dos 14 pontos de referência. Quando comparado ao método de krigagem, as estimativas apresentaram um grau elevado de correlação e coeficiente de determinação. Em relação ao EQMA, o erro calculado ficou abaixo 5 %. Os mapas de variabilidade espacial apresentaram maior amplitude no método de krigagem.

APPLYING THE METHOD TO KRIGING TO ESTIMATE FIELDS OF SOLAR RADIATION: A STUDY FOR NORTHEASTERN BRAZIL

ABSTRACT

Field estimates of solar irradiation at surface for the Northeastern territory of Brazil for the years 2008 and 2011 were made. It was used the Kriging interpolation method. An intercomparison was made using site specific values of incident solar irradiation estimated by interpolation of data collected by automatic meteorological stations (WMS's) and estimates made by the physical radiative transfer model BRAZIL-SR. The evaluation was performed using statistical indices combined with the cross validation method and the visual analysis of maps of the spatial variability of solar irradiation. The selected raw data were reformatted and organized based on a qualitative procedure and then submitted to the same data quality control adopted by the Baseline Surface Radiation Network (BSRN). In Kriging procedure, the careful adjustment of the semivariogram is critical. The measured data showed no anisotropy and trends that could otherwise have prejudiced the estimates by interpolation. The results showed that the Kriging interpolations were able to explain more than 70 % of the variability of the irradiance measured by the number of WMS's in 2010 and 60 % in 2011. Considering the whole period, the Absolute Mean Square Error (ARMSE) normalized by average of measured values had lower rates of around 5 % error, which characterizes an acceptable performance of the method showing that the deviations were minor. Furthermore, these statistical results qualified the methodology employed in the adequacy of the data used in this study. The visual analysis of maps of irradiation for the site specific points in four years showed that there was little spatial variability of the surface irradiance levels with average annual irradiance around 6000 Wh/m². However, all maps could presented sharp gradient in the Northern region of State of Bahia. This can be attributed to anomalies in the data measured by the WMS's that were not detected by data quality control. The BRAZIL-SR model performed better when compared to the Kriging method in estimating the site specific data. The coefficient of determination r^2 showed a value close to 0.74, showing that 74% of the variability of the measured data was explained by the model estimates. The residues indicated that the model overestimated the values in 11 of the 14 reference points. When compared to the Kriging method, estimates showed a high degree of correlation and coefficient of determination. Regarding the ARMSE, the calculated error was below 5%. The maps obtained by using the Kriging method showed greater spatial variability in amplitude.

LISTA DE FIGURAS

	<u>Pág.</u>
Figura 2.1 – Espectro solar fora da atmosfera, comparado com o espectro de um corpo negro a 6000 K e com o espectro solar no nível do mar. A área sombreada corresponde a absorção da radiação pelas moléculas listadas. Fonte: Modificado de Robinson (1966) citado por Guarnieri (2006).....	11
Figura 2.2 – Exemplo: semivariograma experimental. Fonte: adaptado de Camargo (1997) p. 13.	18
Figura 2.3 – Representação gráfica de semivariogramas experimentais sobrepostos juntamente com os modelos teóricos discutidos. Fonte: adaptado de Camargo (1997) p. 19.	21
Figura 2.4 – Representação gráfica de anisotropia geométrica. As linhas sólidas em ambas as direções são os modelos teóricos de ajuste dos semivariogramas experimentais Fonte: adaptada de Camargo (1997), p. 25.....	22
Figura 2.5 – Região semiárida do Nordeste Brasileiro. Essa região é caracterizada principalmente pela forte insolação, temperaturas relativamente altas e regime de chuvas marcado pela escassez, irregularidade e concentração das precipitações num curto período, em média, de três meses, apresentando reservas de água insuficientes em seus mananciais. Fonte: Adaptado de MOURA et al., (2007).....	29
Figura 2.6 – Período Chuvoso: normais climatológicas de nebulosidade (em décimos) para os meses de Dezembro, Janeiro, Fevereiro e Março. Fonte: Adaptada de INMET (2012).	31

Figura 2.7 – Período seco: normais climatológicas de nebulosidade (em décimos) para os meses de Junho, Julho, Agosto e Setembro. Fonte: Adaptada de INMET (2012).	32
Figura 3.1 – Diagrama esquemático demonstrando os passos para a qualificação dos dados utilizados nesse estudo.....	36
Figura 3.2 - Diagrama esquemático demonstrando a primeira etapa para a escolha das PCDs a serem utilizadas neste estudo – aquelas que contêm todos os dados medidos entre as 9h e 16h.	37
Figura 3.3 - Diagrama esquemático demonstrando os passos para a escolha das séries de dados com no mínimo 2/3 de dados medidos pelas PCDs para (a) séries mensais e (b) séries anuais.	38
Figura 3.4 – Arranjo espacial das PCDs que passaram pelos testes de qualificação.	39
Figura 3.5 – Numero de PCDs qualificadas distribuídas por meses num ano. A linha tracejada em vermelho representa o limite mínimo aceitável para a utilização da série anual no processo de interpolação espacial por krigagem. As séries de 2005, 2006 e 2007 foram descartadas.	40
Figura 3.6 - Distribuição espacial das PCDs de comparação na etapa (I) – os pontos em destaque representam os pontos amostrais usados nas comparações com o resultado da KO.	50
Figura 3.7 – Gráfico de dispersão: dados medidos contra a sobreposição das estimativas produzidas através do ajuste de 3 modelos variográficos. A comparação foi realizada para o período de 2010 utilizando os pontos amostrais listados na Tabela 3.4.	52

- Figura 3.8 – Valores dos desvios das estimativas por KO – semivariograma experimental ajustado pelo modelo esférico – considerando as direções de 0° , 45° , 90° , 135° e 180° para cada ponto amostral.....53
- Figura 3.9 – Representação esquemática do processo de escolha dos pontos amostrais. Na figura, o triângulo em verde – ponto da PCD A417 – é escolhido como ponto amostral pois atende ao quesito de conter num raio de até 100 km, no mínimo 4 PCDs circundantes (triângulos em amarelo).....56
- Figura 3.10 – Distribuição espacial dos pontos amostrais disponíveis. Em destaque as 14 PCDs (pontos amostrais de referência) resultantes.57
- Figura 3.11 - Distribuição espacial dos pontos amostrais disponíveis. Em destaque, estão as PCDs utilizadas como pontos amostrais para comparação. A figura geométrica em destaque liga os pontos das 5 PCDs usadas para a demonstração visual dos semivariogramas ajustados.....60
- Figura 3.12 – Ajuste do semivariograma experimental para o ponto amostral: PCD A303 ano de 2010. A curva de cor preta representa o semivariograma experimental dos dados de entrada. A curva azul representa o ajuste do modelo teórico (esférico) aos dados. Uma unidade do eixo x corresponde a 111,25 km.61
- Figura 3.13 – Ajuste do semivariograma experimental para o ponto amostral: PCD A310 ano de 2010. A curva de cor preta representa o semivariograma experimental dos dados de entrada. A curva azul representa o ajuste do modelo teórico (esférico) aos dados. Uma unidade do eixo x corresponde a 111,25 km.62
- Figura 3.14 – Ajuste do semivariograma experimental para o ponto amostral: PCD A340 ano 2010. A curva de cor preta representa o semivariograma experimental dos dados de entrada. A curva azul representa o ajuste do

modelo teórico (esférico) aos dados. Uma unidade do eixo x corresponde a 111,25 km.	63
Figura 3.15 – Ajuste do semivariograma experimental para o ponto amostral: PCD A417 ano 2010. A curva de cor preta representa o semivariograma experimental dos dados de entrada. A curva azul representa o ajuste do modelo teórico (esférico) aos dados. Uma unidade do eixo x corresponde a 111,25 km.	64
Figura 3.16 – Ajuste do semivariograma experimental para o ponto amostral: PCD A436 ano 2010. A curva de cor preta representa o semivariograma experimental dos dados de entrada. A curva azul representa o ajuste do modelo teórico (esférico) aos dados. Uma unidade do eixo x corresponde a 111,25 km.	65
Figura 4.1 – Gráfico de dispersão e ajuste linear dos valores estimados pelos valores medidos para os anos de 2008, 2009, 2010 e 2011.....	68
Figura 4.2 - Semivariogramas produzidos com os dados das séries anuais de 2008, 2009, 2010 e 2011. Uma unidade do eixo x corresponde a 111,25 km.....	69
Figura 4.3 – Variabilidade espacial da irradiação global média anual para a série de 2008 - valores em Wh/m ² obtidos por krigagem ordinária a partir dos 116 pontos de dados amostrais. Os pontos destacados em preto representam a disposição espacial dos pontos amostrais de referência.	71
Figura 4.4 – Variabilidade espacial da irradiação global média anual para a série de 2009 - valores em Wh/m ² obtidos por krigagem ordinária a partir dos 116 pontos de dados amostrais. Os pontos destacados em preto representam a disposição espacial dos pontos amostrais de referência.	72

Figura 4.5 – Variabilidade espacial da irradiação global média anual para a série de 2010 - valores em Wh/m ² obtidos por krigagem ordinária a partir dos 116 pontos de dados amostrais. Os pontos destacados em preto representam a disposição espacial dos pontos amostrais de referência.	73
Figura 4.6 – Variabilidade espacial da irradiação global média anual para a série de 2011 - valores em Wh/m ² obtidos por krigagem ordinária a partir dos 116 pontos de dados amostrais. Os pontos destacados em preto representam a disposição espacial dos pontos amostrais de referência.	74
Figura 4.7 – Resíduos: os valores observados nos pontos amostrais foram subtraídos dos valores estimados pelo modelo BRASIL-SR.	76
Figura 4.8 – Representação de dispersão dos valores estimados (eixo y) contra os valores observados (eixo x).	76
Figura 4.9 – Mapa da irradiação global média anual para a o ano de 2010. Estimativas realizado pelo modelo BRASIL-SR.	77
Figura 4.10 – Gráfico de dispersão e ajuste por regressão linear dos valores estimados por krigagem e pelo modelo BRASIL-SR.	79
Figura 4.11 – Resíduos obtidos pela subtração das estimativas do modelo BRASIL-SR das estimativas por krigagem ordinária para cada ponto amostral de referência.	79
Figura 4.12 – Comparação visual dos mapas de irradiação média anual produzidos através das estimativas para o ano de 2010 por: krigagem ordinária (esquerda) e o modelo numérico BRASIL-SR (direita).....	80

LISTA DE TABELAS

	<u>Pag.</u>
Tabela 1.1 – Comparação de desempenho entre algoritmos de interpolação espacial. A escala de valores de 1 (melhor método) até 5 (pior método) foi adotada para classificação.....	6
Tabela 3.1 - Discriminação dos arquivos de dados fornecidos pelo INMET em números de arquivos com dados amostrais por unidade federal da região Nordeste do Brasil.....	34
Tabela 3.2 - Limites utilizados para derivar os produtos de radiação GEWEX. Valores fisicamente possíveis. Os valores fora dos intervalos indicados foram tratados como ausentes. S_0 é a constante solar ajustada para distância Terra-Sol. μ é o cosseno do ângulo zenital.....	35
Tabela 3.3 - Semelhante à Tabela 3.2, apresentando os intervalos de "Extremamente raros" para qualificação de dados coletados de irradiação solar.....	35
Tabela 3.4 - PCDs escolhidas como pontos de referência onde foram aplicados os pontos amostrais das médias anuais do ano 2010.....	50
Tabela 3.5 – Índices estatísticos para cada um dos modelos variográficos. O melhor desempenho é observado no modelo Esférico.	51
Tabela 3.6 – Apresentação dos resíduos (Wh/m^2) nos três casos de entradas de dados para todas os pontos amostrais: A221, A317, A327 e A446. A última linha representa a média das estimativas em cada ponto.	54
Tabela 3.7 – Pontos amostrais resultantes (PCDs de referência) e suas posições geográficas.	58

Tabela 4.1 – Índices estatísticos que avaliam o desempenho das estimativas em relação aos 14 pontos de dados amostrais de referência para as séries de dados de 2008, 2009, 2010 e 2011.....	68
Tabela 4.2 – Estimativas (Wh/m^2) do modelo numérico BRASIL-SR em relação aos dados medidos para o ano de 2010 nos pontos de referência. Abaixo, são apresentados os valores do coeficiente de determinação (r^2) e do Erro Quadrático Médio Absoluto (EQMA) em porcentagem.	75
Tabela 4.3 – Valores de irradiação média anual (Wh/m^2) estimados por krigagem ordinária e pelo modelo numérico BRASIL-SR para os pontos amostrais de referência o ano de 2010. Abaixo, são mostrados os valores dos índices estatísticos.....	78

SUMÁRIO

	<u>Pag.</u>
1. - INTRODUÇÃO.....	1
1.1 - Objetivos.....	6
1.2 - Organização do trabalho	7
2. - FUNDAMENTAÇÃO TEÓRICA.....	9
2.1 - Radiação solar	9
2.1.1 - Grandezas radiativas básicas	10
2.1.2 - Interações da radiação com a atmosfera.....	11
2.2 - Métodos de interpolação	12
2.2.1 - Métodos de interpolação e estimativas para modelagem de dados e superfícies	13
2.3 - Método de interpolação por krigagem	13
2.3.1 - Variograma e semivariograma	14
2.3.2 - Parâmetros do semivariograma.....	17
2.3.3 - Modelos teóricos de semivariogramas.....	19
2.3.4 - Modelo efeito pepita	20
2.3.5 - Modelo esférico	20
2.3.6 - Modelo exponencial.....	20
2.3.7 - Modelo gaussiano	21
2.3.8 - Anisotropia	22
2.4 - Métodos de krigagem.....	24
2.4.1 - Krigagem Simples (KS)	24
2.4.2 - Krigagem Ordinária (KO)	27

2.5 - Clima na região Nordeste do Brasil	29
3. - ANÁLISE EXPLORATÓRIA E QUALIFICAÇÃO DOS DADOS	33
3.1 - Estrutura dos dados.....	33
3.2 - Qualificação dos dados	34
3.3 - Análise qualitativa dos dados caracterizados.....	37
3.4 - Preenchimento de dados faltantes.....	41
3.4.1 - Índice de transmissividade kt.....	42
3.4.2 - Preenchimento de falhas nas séries de diárias utilizando o índice kt.....	42
3.4.3 - Cálculo da radiação <i>TOA</i> e do índice kt	43
3.4.4 - Aplicação do método de preenchimento de falhas	46
3.5 - Análise de desempenho: índices estatísticos e validação cruzada	47
3.6 - Interpolação espacial.....	48
3.6.1 - Etapa I.....	49
3.6.2 - Etapa II	55
3.7 - Definição das PCDs de referência.....	55
3.8 - Krigagem	58
4. - RESULTADOS E DISCUSSÕES	67
4.1 - Intercomparação: krigagem x dados observados.....	67
4.2 - Intercomparação: estimativas modelo BRASIL-SR x dados observados.....	75
4.3 - Intercomparação: krigagem x modelo BRASIL-SR	78
5. - CONCLUSÕES E CONSIDERAÇÕES FINAIS	81
REFERÊNCIAS BIBLIOGRÁFICAS	85

1. INTRODUÇÃO

A radiação solar é o principal modulador da complexa dinâmica atmosférica. Além de fornecer a energia necessária para os movimentos que impulsionam a circulação atmosférica equilibrando o clima do planeta, fornece energia abundante para todos os seres vivos que aqui habitam, desde a fotossíntese das plantas, até as células fotovoltaicas usadas na conversão de radiação solar em energia elétrica. O estudo de sua natureza e interação com os constituintes atmosféricos até atingir e aquecer a superfície implica em fundamentação de informações a serem aplicadas em áreas diversas do conhecimento humano. Dentre elas, se destacam o estudo da variabilidade climática, o ciclo hidrológico e a pesquisa por fontes alternativas de energia. Assim sendo, o conhecimento aprofundado acerca do potencial de radiação solar torna-se relevante e sua importância evidente.

O saldo de radiação é o fator principal na modelagem do ciclo completo de energia disponível na atmosfera (TRENBERTH; SOLOMON, 1994). Essa energia é responsável por alimentar os sistemas atmosféricos que ocorrem na troposfera. A radiação solar, o calor sensível e o calor latente são as fontes adiabáticas de calor que originam a energia potencial disponível e que impulsionam a dinâmica atmosférica (CHANG et al.; 2002; TRENBERTH et al., 2001). O aquecimento da superfície impulsiona os gradientes de temperatura implicando em movimentos ascendentes e descendentes do ar que, por sua vez, cria diferenças de pressão obtendo como resposta o escoamento do ar originando os ventos – que transportam além de distintos materiais particulados, calor e umidade através da atmosfera.

Além da área de Ciências Atmosféricas a radiação solar desempenha papel importante em outras áreas da atividade humana como, por exemplo, a agropecuária e a arquitetura, que necessitam de informação sobre a radiação solar incidente na superfície para o planejamento e obtenção da maior eficiência energética (PEREIRA et al., 2006). Pesquisas acerca dos níveis de radiação solar contribuem para elaboração de projetos para sistemas de irrigação de culturas; de refrigeração de alimentos, vacinas e remédios; aquecimento e iluminação artificial;

conforto térmico e iluminação natural em projetos de construção civil (MARTINS et al., 2005). Ainda se pode considerar a aplicação dos níveis de radiação solar como fator determinante para as estimativas do saldo de evapotranspiração numa dada região. A evapotranspiração por sua vez, é fundamental para compreensão e modelagem do ciclo hidrológico aplicado ao estudo do balanço hídrico de micro bacias. O fenômeno de evapotranspiração compreende os efeitos da evaporação superficial das superfícies aquáticas e o efeito de transpiração das plantas (CAMARGO; CAMARGO, 2000) e, desse modo, o déficit entre ao valor de evapotranspiração estimado e o valor climatologicamente esperado para um dado período e região está ligado diretamente à seca agrícola (BLAIR; BRUNINI, 2007), fator que ratifica a importância de estudos acerca dos níveis de radiação solar em superfície.

Estudos sobre os níveis de radiação solar em superfície são especialmente relevantes em pesquisas por fontes de energia renováveis (TIBA et al., 2000). O aumento da demanda energética aliado a possibilidade de redução da oferta de combustíveis convencionais e a crescente atenção dada a degradação ambiental estão motivando a comunidade científica pesquisar e desenvolver tecnologias para aproveitamento de fontes alternativas de energia que sejam menos poluentes, renováveis e que minimizem o impacto ambiental (PEREIRA et al., 2006). Ainda hoje, o Brasil possui na energia hidráulica sua maior fonte de geração de eletricidade: as usinas hidroelétricas respondem por 76,9 % da potência total instalada no país (BRASIL, 2013). No entanto, como observados nas últimas décadas, as reservas brasileiras de geração de energia elétrica por hidroelétricas não foram capazes de suprir a demanda em épocas de seca severa. Além disso, a obtenção de energia a partir de fontes como a queima de biomassa polui e degrada a atmosfera e a energia nuclear, produz resíduos radioativos altamente tóxicos. Essas alternativas energéticas claramente manifestar-se inadequadas ou ineficientes para suprir a demanda futura por energia (MARTINS et al., 2005) sob a perspectiva hoje difundida como sustentável. Nesse sentido, a transformação de energia solar em energia elétrica tem se tornado cada vez mais atraente e vem sendo amplamente discutida no Brasil nos

últimos anos (BRASIL, 2013; PEREIRA et al., 2006; MARTINS et al., 2005). Em alguns países europeus como por exemplo, a Alemanha (VIANA et al., 2011), a aplicação desse recurso natural já é uma realidade fazendo com que o país seja hoje, referência em pesquisa e desenvolvimento de tecnologias, investimentos em pesquisas de base e aplicação do recurso solar para a geração de energia elétrica e aquecimento.

Em virtude da posição geográfica do Brasil, grande parte de seu território está contido na região tropical sul do planeta – entre o Equador e o Trópico de Capricórnio. Esse fato serve como estímulo para a investigação dos níveis de radiação solar incidente em superfície no território brasileiro. Desse modo, particularmente, a região Nordeste do Brasil (NEB) está localizada em uma posição privilegiada geograficamente. Primeiro porque está afastada da região de floresta amazônica, situação que lhe priva das constantes formações de nuvens devida à alta concentração de umidade característica de regiões com florestas densas e, segundo, porque a parte norte da região NEB situa-se próxima à linha do Equador, o que lhe proporciona regularidade na intensidade de radiação incidente durante boa parte do ano. A região NEB apresenta registros de temperaturas elevadas durante todo o ano e as amplitudes térmicas máximas observadas são em torno de 6°C (FERREIRA; MELLO, 2005). Ainda, do ponto de vista climático a região NEB é considerada semiárida por apresentar substanciais variações temporal e espacial da precipitação pluviométrica, assim como elevadas temperaturas ao longo do ano (AZEVEDO et al., 1998).

Um obstáculo considerável para um correto levantamento dos níveis de radiação solar incidente em superfície é a disponibilidade de informações confiáveis necessárias para entender a sua distribuição sazonal e espacial assim como a influência do clima em sua variabilidade (COSTA, 2012; MARTINS; PEREIRA, 2011). Estimativas dos níveis de radiação solar incidente em superfície, medidos pela irradiância, são mais apuradas quando os dados são obtidos através de sensores específicos denominados piranômetros. Outros equipamentos destinados a medidas dos níveis de irradiação solar incidente, tais como os heliógrafos e

actinógrafos também são empregados, porém estão caindo em desuso pois fornecem informação com um grau de incerteza muito superior aos piranômetros. A instalação e operação destes equipamentos possui custo elevado e, adicionando a necessidade de manutenção periódica e uma localização adequada, a utilização desses sensores pode dificultar ou até mesmo inviabilizar projetos que dependam dessas medidas (MARTINS; PEREIRA, 2011; GUARNIERI et al., 2006).

Alguns projetos vinculados a pesquisas científicas atuam no sentido de melhorar a qualidade das medidas de dados solarimétricos com o objetivo de torna-los mais confiáveis quanto a sua utilização em trabalhos e pesquisas científicas nas áreas da meteorologia, agrometeorologia, climatologia e até em áreas aplicadas como das energias renováveis. Isso é feito ampliando e fornecendo suporte aos sensores que compõe redes de Estações Meteorológicas de Observação de Superfície Automáticas (para simplificação, doravante denominadas por PCDs – sigla para Plataforma Coletoras de Dados) junto a metodologias para o tratamento e qualificação dos dados coletados pelas PCDs e, ainda, através de investimento em pesquisa e desenvolvimento de modelos computacionais que utilizam de dados meteorológicos e de satélites para estimar a irradiância solar em superfície.

No cenário brasileiro, podemos destacar dois projetos que objetivaram a coleta de dados a nível nacional, o controle de qualidade desses dados, sua organização e disseminação a nível nacional e internacional, bem como na geração de mapas. São eles o projeto SWERA e o projeto SONDA – descritos a seguir.

O projeto SWERA (*Solar and Wind Energy Resource Assessment*), financiado pelo Programa das Nações Unidas para o Meio Ambiente (PNUMA) teve como objetivo fundamental facilitar a inclusão de fontes de energia renováveis na matriz energética de um grupo de países-piloto selecionados para esse projeto. O projeto contou com a colaboração de treze países divididos em três grupos regionais: África, América Latina e Ásia. O Instituto Nacional de Pesquisas Espaciais (INPE) foi responsável pela coordenação das atividades do projeto para o Brasil e para a América Latina (MARTINS et al., 2005). Mais

informações em: <http://en.openei.org/apps/SWERA/>. O projeto SONDA (Sistema de Organização Nacional de Dados), financiado pela FINEP e mantido pelo INPE através do suporte financeiro da Petrobras, é um projeto ainda em andamento ligado à área de pesquisas em clima e meteorologia, todavia verifica-se uma vertente desse projeto direcionada ao suporte de atividades na área de energias renováveis. O projeto tem como objetivo principal o desenvolvimento de uma base de dados completa, integrada e de alta confiabilidade que contemple as necessidades dos setores da sociedade envolvidos com a pesquisa, o desenvolvimento, o planejamento e o investimento em uso e aplicações no setor energético (MARTINS et al., 2005). Mais informações em: <http://sonda.ccst.inpe.br/>.

Os dados coletados pelos sensores radiométricos são pontuais e representam apenas uma pequena área ao redor dos mesmos. O Instituto Nacional de Meteorologia (INMET) é responsável pela operação de uma malha de estações medidoras e disponibiliza para consulta mapas com a distribuição espacial das PCDs distribuídas por toda a extensão territorial brasileira que pode ser consultado aqui: <http://www.inmet.gov.br/portal/index.php?r=estacoes/mapaEstacoes>. Pode-se considerar o emprego dessa malha de dados para alimentar métodos de interpolação espacial que permitam estimar os valores diários de irradiância em localidades onde não há coleta de dados. Outra forma de mapear a disponibilidade e variabilidade espacial da irradiação solar incidente é fazer uso de modelos numéricos para a solução da equação de transferência radiativa na atmosfera utilizando dados obtidos com o uso de satélites ambientais em órbita da Terra. Uma terceira maneira consiste no emprego de modelos atmosféricos numéricos de escala reduzida para esse fim. Contudo, Martins e Pereira (2011) mostraram que a confiabilidade das avaliações de irradiação solar obtidas por interpolação espacial é menor que a confiabilidade dos dados obtidos através dos modelos numéricos físicos alimentados por dados de satélites quando a densidade da rede de coleta de dados é tal que a distância entre estações próximas é superior a 50 km.

Na literatura, são encontrados diferentes tipos de métodos matemáticos para realizar procedimentos de ajuste de uma função a pontos não amostrados a fim de se obter valores que, de alguma maneira não se pode medir. O método de krigagem – termo cunhado do francês *krigeage* e do inglês *kriging* em homenagem a Daniel G. Krige, engenheiro pioneiro na aplicação de técnicas estatísticas em estudos geológicos aplicados a mineração – é largamente utilizado como estimador de dados não amostrados. Landim (2000) realizou testes comparativos entre 5 métodos de interpolação mais usuais (Triangulação, Inverso da Distância, Superfície/Tendência, Mínima Curvatura e Krigagem) demonstrando, através de uma comparação por desempenho nas estimativas (desvio em relação aos dados medidos) que, para dados irregularmente espaçados e uma malha de pontos maior que 30, a krigagem obteve o melhor desempenho em precisão geral como mostra a Tabela 1.1.

Tabela 1.1 – Comparação de desempenho entre algoritmos de interpolação espacial. A escala de valores de 1 (melhor método) até 5 (pior método) foi adotada para classificação.

Algoritmo	Fidelidade aos dados	Suavidade das curvas	Velocidade de computação	Precisão geral
Triangulação	1	5	1	5
Inverso da Distância	3	4	2	4
Superfície/Tendência	5	1	3	3
Mínima Curvatura	4	2	4	2
Krigagem	2	3	5	1

Fonte: Adaptado de Landim (2000)

1.1 Objetivos

Com base no que foi tratado acima, este trabalho tem como objetivo testar e avaliar o desempenho do método de interpolação por krigagem através das estimativas de irradiância. Posteriormente, realizar uma intercomparação com estimativas

fornecidas por um modelo numérico, o modelo de transferência radiativa BRASIL-SR, utilizado para elaboração do Atlas Brasileiro de Energia Solar (Pereira et al., 2006) frente aos dados observados – medidos pelas PCDs. Deste modo, será possível avaliar o desempenho do método de krigagem e do modelo BRASIL-SR diante das observações da rede de PCDs. Os resultados, além de avaliarem os métodos e a metodologia empregada, podem contribuir disponibilizando elementos que possam atender a demanda por informações sobre os níveis de irradiação solar nos setores de atividade humana discutidos anteriormente.

1.2 Organização do trabalho

A seguir, são apresentadas as etapas a cumpridas de modo a atingir os objetivos propostos.

- Manipulação e pré-processamento dos dados observacionais de irradiação solar disponibilizados pelo INMET para o NEB: esse tratamento consiste na adequação dos dados em forma matricial – as linhas representam os dados horários de irradiação para cada dia juliano correspondente no período de 2005 a 2011; as colunas representam as horas do período diurno abrangendo o tempo de 6h as 19h (disponibilizadas no formato UTC e convertidas para a hora local - civil).
- Controle de qualidade dos dados segundo especificações de órgãos internacionais específicos: após a etapa de pré-processamento dos dados, realização do cálculo do total acumulado diário, e dos valores médios mensais e anuais da irradiação solar em cada PCD.
- Desenvolvimento de uma metodologia adequada para escolha das PCDs de referência.

- Desenvolvimento, adaptação e implementação dos códigos computacionais necessários para a produção dos semivariogramas e a realização da interpolação. Em seguida, o ajuste dos semivariogramas dos dados de irradiação solar observados e sua aplicação no método de interpolação por krigagem.
- Realização de intercomparações: dados observados x dados estimados por krigagem; dados observados x dados estimados pelo modelo BRASIL-SR (para o ano de 2010); dados estimados por krigagem x dados estimados pelo modelo BRASIL-SR (para o ano de 2010).
- Quantificação dos desvios observados nas intercomparações realizadas na etapa anterior e análise qualitativa dos resultados obtidos.

2. FUNDAMENTAÇÃO TEÓRICA

2.1 Radiação solar

Hoje sabemos com clareza a importância do Sol no surgimento, desenvolvimento e manutenção da vida na Terra. A radiação emitida da superfície solar atinge nosso planeta constantemente sob a forma de radiação eletromagnética nutrindo os processos químicos, térmicos e dinâmicos que ocorrem tanto na atmosfera, quanto na superfície dos continentes e nos oceanos.

Tradicionalmente a energia é definida como a capacidade de realizar trabalho ou exercer uma ação. Ou ainda, podemos defini-la como um estado dinâmico num sistema fechado, podendo variar de acordo com a alteração de certas propriedades da matéria como massa e temperatura. A energia que recebemos do Sol, se propaga pelo espaço através de ondas eletromagnéticas. A energia solar é convertida por meio da interação da radiação eletromagnética com a matéria e seu aspecto ondulatório é particularmente interessante a Meteorologia (VAREJÃO, 2006).

A radiação eletromagnética é caracterizada por grandezas físicas específicas como o comprimento de onda (λ) ou a frequência (ν). Essas grandezas constituem o espectro eletromagnético e se relacionam através da proporção $c = \lambda\nu$, onde c é a velocidade da luz no meio. O comprimento de onda é a distância que separa dois máximos sucessivos numa onda e a frequência é o número de repetições de máximos (ou mínimos) que passam por um ponto de referência (TIPLER, 1978). O espectro eletromagnético da radiação solar envolve comprimentos de onda que vão desde a ordem de 10^{-12} m (região dos raios gama) até próximo a 10^5 m (região das ondas longas de rádio) sendo, o espectro visível, uma faixa espectral que contém os comprimentos de onda (ou as frequências) que sensibilizam o olho humano. A faixa espectral do visível compreende comprimentos de onda por volta de $0,36 \mu m$ até $0,74 \mu m$.

2.1.1 Grandezas radiativas básicas

A taxa de transferência de energia pela radiação eletromagnética é chamada de fluxo de radiação ou fluxo radiante, expresso em unidades de energia por unidade de tempo: Joules por segundo $J.s^{-1}$ ou Watts (W). Assim, a densidade de fluxo radiante é o fluxo radiante que atravessa uma unidade de área, e é expressa em Watts por metro quadrado ($W.m^{-2}$) sendo então, designada irradiância (E) quando representa o fluxo emergindo ou incidindo em uma área (LIOU, 2002). A irradiância incidente medida em superfície possui contribuição de muitas fontes com direções diferentes e, fundamentalmente, muitas vezes é necessário identificar a parte da irradiância segundo uma direção específica ou um comprimento de arco específico. Esse feixe de radiação é definido como radiância (L) sendo a irradiância por unidade de ângulo sólido (WALLACE; HOBBS, 1977). Essa grandeza permite detalhar as porções de radiação provenientes de cada direção em torno de um ponto de observação (GUARNIERI, 2006).

Como a radiação solar atinge a superfície com inclinações distintas, o ângulo zenital (θ_z) é inserido de modo a corrigir essa discrepância. O ângulo θ_z é o ângulo entre a normal à superfície – ou o zênite do observador que faz a medida – e a direção do feixe incidente. Desse modo, obtemos a componente da radiância na direção normal a superfície de incidência multiplicando a radiância pelo cosseno do ângulo zenital ($L \cdot \cos(\theta_z)$) (LIOU, 2002).

Em termos de comprimento de onda (ou monocromático), pode-se escolher obter a irradiância em unidades de energia por área por comprimento de onda e por esterradiano. Assim, a irradiância monocromática de energia radiante é definida pelo componente normal de radiância integrada ao longo de todo o hemisfério por unidade de ângulo sólido $d\Omega$ (LIOU, 2002).

$$E_\lambda = \int_0^\pi L_\lambda \cos(\theta_z) d\Omega \quad ; \quad \text{onde} \quad d\Omega = \sin(\theta_z) d\theta d\varphi \quad (2.3)$$

Portanto, a irradiância total, ou seja, levando em conta todos os comprimentos de onda, é obtida por meio da integração da radiância em todo o espectro eletromagnético.

2.1.2 Interações da radiação com a atmosfera

Ao atravessar a atmosfera, a radiação solar interage com as partículas que a constitui sofrendo o processo de atenuação. Ela é absorvida, espalhada, refratada ou refletida pelas nuvens, vapor d'água e partículas presentes na atmosfera – como os aerossóis. Segundo Liou (2002), os processos mais importantes que influenciam na intensidade da radiação solar incidente em superfície são a absorção e o espalhamento atmosféricos.

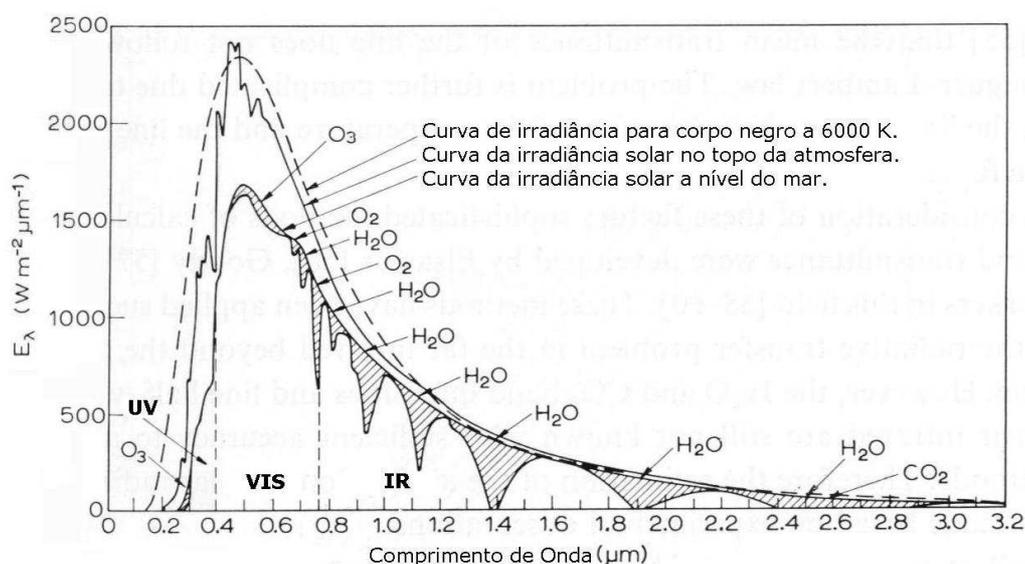


Figura 2.1 – Espectro solar fora da atmosfera, comparado com o espectro de um corpo negro a 6000 K e com o espectro solar no nível do mar. A área sombreada corresponde a absorção da radiação pelas moléculas listadas.

Fonte: Modificado de Robinson (1966) citado por Guarnieri (2006).

A área entre as curvas (linhas contínuas) da Figura 2.1 representa a redução sofrida pela radiação solar incidente, durante sua passagem através da atmosfera. Essa redução é dividida em duas partes: a área não sombreada representa o espalhamento sofrido pela radiação e a área sombreada a absorção pelas moléculas do ar, primariamente por vapor d'água (H₂O), dióxido de carbono (CO₂), ozônio (O₃) e oxigênio (O₂).

2.2 Métodos de interpolação

Neste capítulo são apresentados alguns dos métodos de interpolação mais usuais. No entanto, o foco é mantido no método de interpolação empregado nesse trabalho, a krigagem. Todavia, não são discutidos os algoritmos matemáticos dos métodos, mas sim sua ideia, função e aplicação prática.

A rigor, interpolar uma função $f(x)$ consiste em aproximar essa função por outra função $g(x)$, escolhida entre uma classe de funções definidas *a priori* e que satisfaça algumas propriedades. A função $g(x)$ é então usada em substituição à função $f(x)$.

Deste modo, a necessidade de se efetuar esta substituição surge em várias situações, como por exemplo, quando são conhecidos somente valores numéricos da função para um conjunto de pontos e é necessário calcular o valor da função em um ponto não tabelado (RUGGIERO; ROCHA, 1996). Desse modo, empregar métodos de interpolação, é fazer uma estimativa do valor de uma variável em locais não amostrados, a partir de locais amostrados numa mesma área (BURROUGH; McDONNELL, 1998).

2.2.1 Métodos de interpolação e estimativas para modelagem de dados e superfícies

Não é difícil encontrar descrições e aplicações sobre diferentes metodologias de interpolação. Como são muitas as classificações, 3 práticas interpoladoras são comumente citadas (HARTKAMP et al., 1999; MAZZINI; SCHETTINI, 2009): globais ou locais, exatos ou suavizantes e determinísticos ou estocásticos.

Interpoladores globais ponderam todos os pontos do mesmo local amostrado, o que permite interpolar o valor da função em qualquer ponto dentro do domínio dos dados originais de modo que a adição ou remoção de um valor tem consequências no domínio de definição da função, ou seja, afeta todo o mapa. Os interpoladores locais consistem em funções definidas para porções determinadas do mapa, portanto a alteração de um valor afeta localmente os pontos próximos ao mesmo (MAZZINI; SCHETTINI, 2009).

Interpoladores exatos, a rigor de sua classificação, respeitam os dados originais de modo que a predição no ponto amostrado deverá ser igual ao valor medido. Já os interpoladores suavizantes (ou *smoothing*), por outro lado, produzem suavização das curvas da superfície originada, de modo que os prováveis erros são minimizados.

Interpoladores estocásticos fazem uso da teoria da probabilidade e incorporam critérios estatísticos na determinação dos pesos atribuídos aos pontos amostrais para o cálculo das interpolações. Interpoladores determinísticos, no entanto, não fazem uso da probabilidade e, ao invés disso, geram uma combinação linear dos valores amostrados baseando-se apenas na geometria da distribuição espacial dos dados amostrados (MAZZINI; SCHETTINI, 2009).

2.3 Método de interpolação por krigagem

O método de interpolação por krigagem é caracterizado como um modelo interpolador estocástico. Segundo Hartkamp et al. (1999), interpoladores

estocásticos fazem uso da teoria das probabilidades – são incorporados critérios estatísticos que atribuem pesos aos pontos amostrais durante a interpolação. A diferença no método de krigagem está no processo de interpolação, onde são utilizadas técnicas de geoestatística. Na geoestatística são empregados métodos estatísticos na investigação da variação espacial de determinadas grandezas físicas, as chamadas *variáveis regionalizadas* (MATHERON, 1971). Segundo Landim (2003), pode-se dizer, em último caso, que a krigagem é um método de estimativa por médias móveis.

Conforme Landim e Sturaro (2002), a metodologia geoestatística procura extrair, de uma aleatoriedade aparente dos dados medidos, as propriedades estruturais probabilísticas do componente regionalizado, ou seja, é obtida uma função de correlação entre os valores situados em determinada distância e direção no espaço amostrado.

Os valores estimados através do método de krigagem são obtidos num processo de estimativas por médias móveis de medidas distribuídas no espaço a partir de valores de seus arredores e a função que relaciona a dependência espacial é o semivariograma (LANDIM; STURARO, 2002). A krigagem – estimador que se baseia numa série de técnicas de análise de regressão, sejam elas lineares ou não – procura minimizar a variância estimada a partir de um modelo prévio que leva em conta a dependência estocástica (não determinista) entre os dados distribuídos no espaço (LANDIM, 2003). Finalmente, o método de krigagem fornece um mapa que possibilita examinar a variabilidade espacial da variável que se quer estudar.

2.3.1 Variograma e semivariograma

O variograma é um instrumento fundamental que fornece suporte às técnicas de krigagem admitindo uma representação quantitativa da alteração de uma variável regionalizada no espaço (HUIJBREGTS, 1975).

Seja uma variável regionalizada $Z(i)$ coletada em diversos pontos i regularmente distribuídos. Segundo Landim (2003), o valor de cada ponto está relacionado de algum modo com valores obtidos a partir de pontos situados a certa distância, sendo aceitável concluir que a influência é tanto maior quanto menor for a distância entre os pontos.

Assim, define-se um vetor distância $\Delta \vec{h}$ que possui orientação específica e seu grau de relação entre pontos numa certa direção pode ser atribuído pela covariância. O uso desse artifício, embora exista covariância entre todas as distâncias possíveis ao longo de h , permite estipular que somente sejam considerados valores entre pontos regularmente espaçados por múltiplos inteiros de $\Delta \vec{h}$. Desse modo, a covariância entre valores encontrados nessas distâncias separadas por $\Delta \vec{h}$, ao longo de h , pode ser representada por

$$C(h) = C(\Delta \vec{h}) = \frac{1}{n} \sum_{i=1}^n Z_i Z_{i+h} \quad (2.4)$$

Como n é o número total de pares de valores comparados, a covariância é igual à média dos produtos-cruzados dos valores $Z(i)$ encontrados nos pontos i pelos valores $Z(i+h)$ encontrados nos pontos $i+h$ distantes a um intervalo $\Delta \vec{h}$. A covariância dependerá do tamanho do vetor h e, se $h = 0$, $C(h)$ passa a representar a variância de modo que $C(0) = \frac{1}{n} \sum_{i=1}^n (Z_i Z_{i+0})^2 = \text{var}(Z)$.

Desse modo, pode-se calcular uma função, denominada *semivariância* (Γ), definida como a metade da variância das diferenças.

$$\Gamma(\vec{h}) = \Gamma(h) = \frac{1}{2n} \sum_{i=1}^n (Z_{i+h} - Z_i)^2 \quad (2.5)$$

Onde $\Gamma(h)$ fornece, num gráfico contra as distâncias h , o assim chamado *semivariograma experimental*.

Sabemos que a variância é dada por: $\text{var}(Z) = \sum_{i=1}^n \frac{Z_i^2}{n} - (\sum_{i=1}^n Z_i/n)^2$, assim, pode-se representar $\Gamma(h)$ por

$$\Gamma(h) = \left[\frac{1}{2n \sum_{i=1}^n (Z_{i+h} - Z_i)^2} \right] - \left[\frac{1}{2n \sum_{i=1}^n (Z_{i+h} - Z_i)} \right]^2 \quad (2.6)$$

Como a média da variável regionalizada $\mathbf{z}(\mathbf{i})$ é também a média da variável regionalizada $\mathbf{z}(\mathbf{i} + \mathbf{h})$, pois se trata das mesmas observações, as variáveis são tomadas apenas em \mathbf{i} e $\mathbf{i} + \mathbf{h}$. Assumindo estacionaridade, pode-se reescrever

$$\left[\frac{1}{2n \sum_{i=1}^n (Z_{i+h} - Z_i)} \right]^2 = 0$$

e

$$\Gamma(h) = \left[\frac{1}{2n \sum_{i=1}^n (Z_{i+h} - Z_i)^2} \right] = \frac{1}{2n} \left[\left(\sum_{i=1}^n Z_{i+h}^2 \right) + \left(\sum_{i=1}^n Z_i^2 \right) \right] - \frac{1}{n} \left(\sum_{i=1}^n Z_{i+h} - Z_i \right)$$

Portanto,

$$\Gamma(h) = \mathcal{C}(0) - \mathcal{C}(h) \quad (2.7)$$

Na Equação 2.7, a semivariância é representada em função da variância $\mathcal{C}(0)$ e da covariância $\mathcal{C}(h)$. Isso nos permite avaliar o crescimento da covariância em função

do vetor h . De acordo com o que já foi exposto, o vetor h apresentando-se infinitamente pequeno faz que a variância seja mínima e a covariância máxima. Desse modo, haverá um valor h para as quais ambas podem apresentar valores aproximadamente iguais, entretanto, à medida que h aumenta, a covariância diminui ao passo que a variância aumenta. Isso acontece porque ocorre progressivamente maior independência entre os pontos com o aumento da distância (LANDIM, 2003).

Se os dados forem estacionários, ou seja, se não forem observadas mudanças bruscas, a semivariância varia de 0 (zero), quando $h = 0$, até um valor igual à variância das observações para um alto valor de h . A distância segundo a qual $\Gamma(h)$ atinge um patamar (denominado soleira ou “*sill*” em inglês), *a priori* igual à variância dos dados, que se dá o nome de alcance ou amplitude (*range*). Geralmente, representa-se a soleira (*sill*) por C e o alcance (*range*) por α .

O semivariograma é, portanto, uma medida da variabilidade da variável em relação à distância. Essa variabilidade é consideravelmente desigual quando consideradas diferentes direções (LANDIM, 2003). Na construção do semivariograma é indispensável dispor de um conjunto de valores obtidos a intervalos regulares dentro de um mesmo suporte geométrico.

2.3.2 Parâmetros do semivariograma

Como já foi dito, o semivariograma experimental é arranjado plotando a semivariância $\Gamma(h)$ contra h (Figura 2.2). Dessa forma, é possível uma visualização do comportamento espacial da variável em função de alguns parâmetros característicos do semivariograma. São eles: alcance (α), patamar (C) e efeito pepita (C_0).

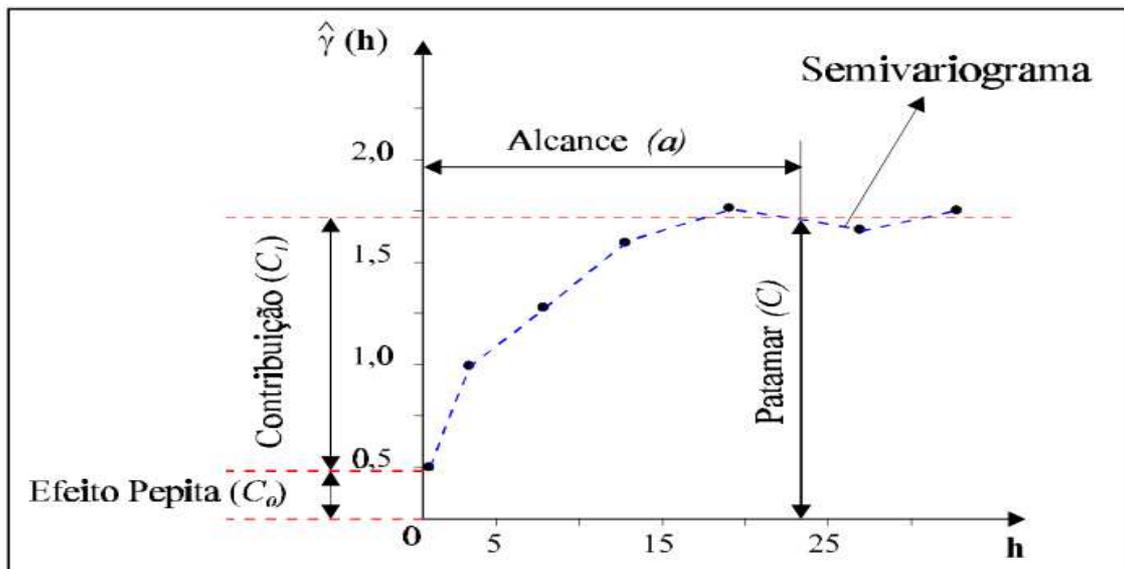


Figura 2.2 – Exemplo: semivariograma experimental.

Fonte: adaptado de Camargo (1997) p. 13.

- *Alcance*: distância dentro da qual as amostras apresentam-se correlacionadas espacialmente (na Figura 2.2, o alcance ocorre próximo a 25 em unidades de h).
- *Patamar*: trata-se do valor máximo para a semivariância correspondente ao seu *alcance* (a). Essa informação é importante, pois, desse ponto em diante a dependência espacial entre as amostras é desconsiderada.
- *Efeito pepita*: este parâmetro aponta a descontinuidade do semivariograma para distâncias menores do que a menor distância entre as amostras. Alguns autores atribuem essa descontinuidade a erros de medição e é consenso que é pouco provável quantificar se a maior contribuição de C_0 provém dos erros de medição, ou da variabilidade de pequena escala não captada pela amostragem (CAMARGO, 1997).
- *Contribuição* (C_1): diferença entre o patamar e o efeito pepita.

2.3.3 Modelos teóricos de semivariogramas

Depois de construído o variograma experimental, é necessário ajustá-lo junto a um modelo teórico. Na verdade, é o modelo que se ajusta ao semivariograma experimental. Os modelos teóricos frequentemente usados são: modelo esférico, exponencial e gaussiano (LANDIM, 2003; CAMARGO, 1997). O gráfico do semivariograma experimental como na Figura 3.1 é constituído por uma série de valores sobre os quais se objetiva ajustar uma função. É importante que o modelo ajustado represente a tendência de $\Gamma(h)$ em relação à h . Deste modo, as estimativas a partir da krigagem serão mais precisas e, portanto mais confiáveis (CAMARGO, 1997) do ponto de vista estatístico.

O procedimento de ajuste é interativo e requer experiência e resignação na adequação dos parâmetros ao modelo teórico. Trata-se de um ajuste visual e, por muitas vezes de tentativa e erro alterando-se o ajuste entre os valores dos parâmetros do semivariograma. Assim, dependendo do ajuste obtido, pode ou não redefinir o modelo, até obter um que seja considerado satisfatório.

Os modelos apresentados nesta seção são considerados modelos básicos, designados de modelos isotrópicos por Isaaks e Srivastava (1989). Estão divididos em dois tipos: modelos com patamar (ou com soleira) e modelos sem patamar (ou sem soleira). Modelos do primeiro tipo são referenciados na geoestatística como modelos transitivos. Alguns dos modelos transitivos atingem o patamar (C) assintoticamente. Para tais modelos, o alcance (a) é arbitrariamente definido como a distância correspondente a 95% do patamar. Modelos do segundo tipo não atingem o patamar, e continuam aumentando enquanto a distância aumenta. Tais modelos são utilizados para modelar fenômenos que possuem capacidade infinita de dispersão.

2.3.4 Modelo efeito pepita

Conforme tratado anteriormente, muitos semivariogramas experimentais apresentam certa descontinuidade em sua origem. Teoricamente, quando $h = 0$, o valor do semivariograma é estritamente zero. Porém, quando h tende a zero, o valor do semivariograma pode ser significativamente maior que zero, isto é, ocorre uma descontinuidade na origem. Essa descontinuidade pode ser normalizada e modelada através do efeito pepita (C_0) como segue:

$$\Gamma_0(h) = \begin{cases} 0 & \text{se } h = 0 \\ 1 & \text{se } h \neq 0 \end{cases} \quad (2.8)$$

2.3.5 Modelo esférico

Neste modelo, descrito pelas Equações 2.9 e 2.10, a inclinação da reta tangente junto à origem ($h \cong 0$) é $3C/2a$; segundo Landim (2003), trata-se do modelo mais comum e equivale a função de distribuição normal da estatística básica.

$$\Gamma(h) = C \left[\frac{3}{2} \left(\frac{h}{a} \right) - \frac{1}{2} \left(\frac{h}{a} \right)^3 \right] \quad (\text{quando } h < a) \quad (2.9)$$

$$\Gamma(h) = C \quad (\text{quando } h \geq a) \quad (2.10)$$

2.3.6 Modelo exponencial

Este modelo, descrito pela Equação 2.11, atinge o patamar assintoticamente, com o alcance prático definido como a distância na qual o valor do modelo é 95% do patamar (LANDIM, 2003).

$$\Gamma(h) = C \left[1 - e^{-3h/a} \right] \quad (2.11)$$

2.3.7 Modelo gaussiano

O modelo gaussiano é um modelo transitivo, muitas vezes usado para modelar fenômenos extremamente contínuos (ISAAKS; SRIVASTAVA, 1989). É descrito por:

$$\Gamma(h) = C \left[1 - e^{(-3h/a)^2} \right] \quad (2.12)$$

Semelhante no modelo exponencial, o modelo gaussiano atinge o patamar assintoticamente e o parâmetro a é definido como o alcance prático ou distância na qual o valor do modelo é 95% do patamar (LANDIM, 2003).

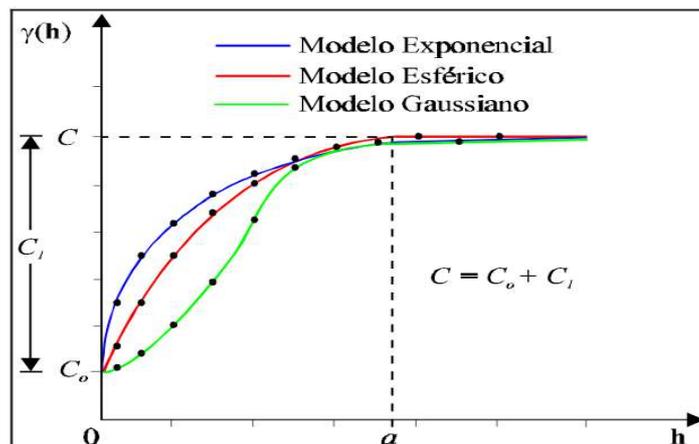


Figura 2.3 – Representação gráfica de semivariogramas experimentais sobrepostos juntamente com os modelos teóricos discutidos.

Fonte: adaptado de Camargo (1997) p. 19.

Até este ponto foram expostos os principais modelos básicos “normalizados”, os quais são utilizados para modelar ou ajustar o semivariograma experimental. Na prática, os semivariogramas experimentais possuem valores de efeito pepita (C_0)

maiores que zero e valores de patamar (C) maiores que a unidade, conforme ilustrado na Figura 2.3.

2.3.8 Anisotropia

Em muitos casos nos deparamos com variáveis que possuem a característica de alterar sua magnitude com a direção. A *anisotropia* pode ser facilmente identificada através da observação dos semivariogramas experimentais obtidos para direções distintas. Segundo Camargo (1997) e Landim (2003), nas aplicações e análises geoestatísticas são identificados 3 tipos de anisotropias: geométrica, zonal e combinada.

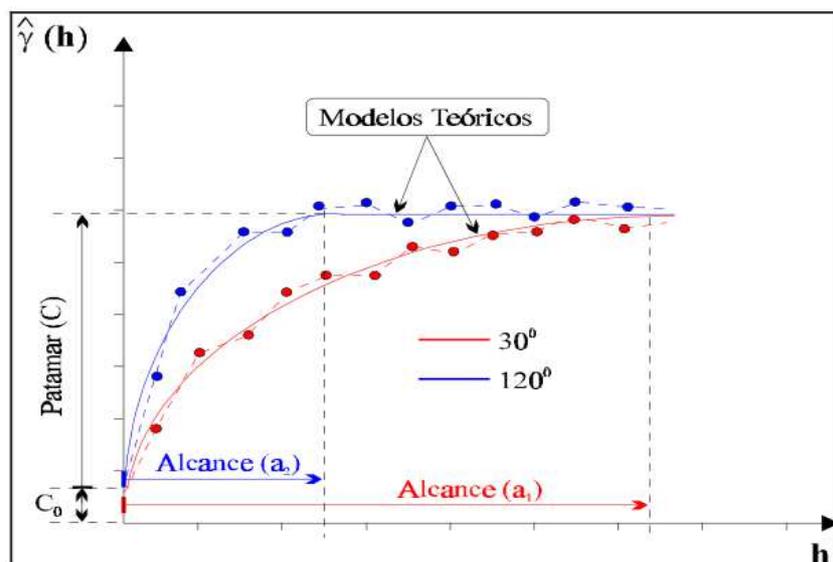


Figura 2.4 – Representação gráfica de anisotropia geométrica. As linhas sólidas em ambas as direções são os modelos teóricos de ajuste dos semivariogramas experimentais

Fonte: adaptada de Camargo (1997), p. 25.

A ocorrência de anisotropia geométrica proporciona a característica de se obter semivariogramas com diferentes valores de α e valores muito próximos de C . Para elucidação, Camargo (1997) ilustrou o problema avaliando semivariogramas obtidos para direções diferentes (30° e 120°). Considerando os semivariogramas apresentados na Figura 2.4, observa-se que o semivariograma que atinge primeiro o patamar (azul) se refere à direção de 120° e o semivariograma com maior alcance (vermelho) se refere à direção de 30°.

O fenômeno de anisotropia zonal apresenta situação oposta à anisotropia geométrica, ou seja, os valores para C são diferentes enquanto que os valores para α são muito próximos. Esse tipo de anisotropia é pouco observado. No entanto, é mais frequente se observar a anisotropia combinada que, como seu nome sugere, é uma combinação dos dois tipos de anisotropia observados – zonal e geométrico.

Então, segundo Landim (2003), para o emprego do semivariograma são demandadas as seguintes hipóteses a respeito da malha de dados:

- As diferenças entre pares de valores de amostras são determinadas apenas pela orientação espacial relativa dessas amostras;
- O interesse é focalizado na média e na variância das diferenças, constituindo esses dois parâmetros como sendo condicionados unicamente da orientação (*hipótese intrínseca*);
- Por conveniência, assume-se que os valores da área de estudo não apresentam tendência que possa afetar os resultados e, assim, a atenção do método fica apenas com a variância da diferença entre as amostras.

2.4 Métodos de krigagem

São várias as formas (ou tipos) de krigagem. As mais usuais devidas sua simplicidade perante a complexidade dos outros tipos são a *krigagem simples* e a *krigagem ordinária* (CAMARGO, 1997; LANDIM, 2003). Descrever todos os tipos de krigagem encontrados na literatura assim como suas particularidades não é objetivo desse trabalho. Sendo assim, é descrito a seguir, sucintamente, apenas os métodos mais simples e usuais – a terminologia assim como os símbolos e equações aqui demonstrados são como em Landim (2003).

2.4.1 Krigagem Simples (KS)

Considere uma superfície sobre a qual é observada determinada propriedade do solo, Z , em n pontos distintos, com coordenadas representadas pelo vetor \mathbf{x} . Assim, tem-se um conjunto de valores $\{Z(\mathbf{x}_i), i = 1, \dots, n\}$, onde \mathbf{x}_i , identifica uma posição em duas dimensões representada pelos pares de coordenadas (x_i, y_i) . Suponha que se objetive estimar o valor de Z no ponto \mathbf{x}_0 . O valor desconhecido de $Z(\mathbf{x}_0)$ pode ser estimado a partir de uma combinação linear dos n valores observados, adicionado a um parâmetro, λ_0 (JOURNEL, 1988), assim, podemos escrever

$$Z_{\mathbf{x}_0}^* = \lambda_0 + \sum_{i=1}^n \lambda_i Z(\mathbf{x}_i) \quad (2.13)$$

Deseja-se obter um estimador não tendencioso, ou seja,

$$E[Z_{\mathbf{x}_0} - Z_{\mathbf{x}_0}^*] = 0 \quad (2.14)$$

No entanto, essa relação impõe que as médias do valor medido (Z_{x_0}) e do valor estimado ($Z_{x_0}^*$) sejam iguais,

$$E[Z_{x_0}] = E[Z_{x_0}^*] \quad (2.15)$$

Assim, substituindo a Equação 2.13 no argumento do lado direito da equação 2.15, podemos obter λ_0 como segue,

$$E[Z_{x_0}^*] = E\left[\lambda_0 + \sum_{i=1}^n \lambda_i Z(x_i)\right] = \lambda_0 + \sum_{i=1}^n \lambda_i E[Z(x_i)] \quad (2.16)$$

Substituindo a Equação 2.16 em 2.15,

$$\lambda_0 = E[Z(x_0)] - \sum_{i=1}^n \lambda_i E[Z(x_i)] \quad (2.17)$$

Assim, inserindo λ_0 na Equação 2.13, obtém-se o estimador:

$$Z_{x_0}^* = E[Z(x_0)] - \sum_{i=1}^n \lambda_i E[Z(x_i)] + \sum_{i=1}^n \lambda_i Z(x_i) \quad (2.18)$$

A característica do método de *krigagem simples* é supor que a média (m) é conhecida e constante *a priori*, ou seja,

$$E[Z(x_0)] = E[Z(x_i)] = m \quad (2.19)$$

Substituindo a Equação 3.19 em 3.18, o estimador de krigagem simples fica:

$$Z_{x_0}^* = m + \sum_{i=1}^n \lambda_i [Z(x_i) - m] \quad (2.20)$$

Journel (1988) mostra que ao minimizar a variância do erro ($var[Z_{x_0} - Z_{x_0}^*]$), os pesos λ_i são obtidos a partir do seguinte sistema de equações denominado sistema de krigagem simples.

$$\sum_{j=1}^n \lambda_j C(x_i, x_j) = C(x_i, x_0) \quad (2.21)$$

onde,

- $C(x_i, x_j)$ refere-se a função de covariância correspondente ao vetor \mathbf{h} , com origem em x_i e extremidade em x_j ;
- $C(x_i, x_0)$ refere-se a função de covariância correspondente ao vetor \mathbf{h} , com origem em x_i e extremidade em x_0 .

A correspondente variância minimizada do erro, apresentada por Journel (1998) e denominada variância de krigagem simples (σ_{ks}^2) é dada por:

$$\sigma_{ks}^2 = var[Z_{x_0} - Z_{x_0}^*] = C(\mathbf{0}) - \sum_{i=1}^n \lambda_i C(x_i, x_0) \quad (2.22)$$

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

$$\mathbf{K} \cdot \boldsymbol{\lambda} = \mathbf{k} \rightarrow \boldsymbol{\lambda} = \mathbf{K}^{-1} \cdot \mathbf{k} \quad (2.23)$$

onde

$$\mathbf{K} = \begin{pmatrix} C_{11} & \dots & C_{1n} \\ \vdots & \ddots & \vdots \\ C_{n1} & \dots & C_{nn} \end{pmatrix}, \quad \boldsymbol{\lambda} = \begin{bmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{bmatrix}, \quad \mathbf{k} = \begin{bmatrix} C_{10} \\ \vdots \\ C_{n0} \end{bmatrix}$$

\mathbf{K} e \mathbf{k} são matrizes das covariâncias e $\boldsymbol{\lambda}$ o vetor dos pesos. Assim, a variância de krigagem simples é

$$\sigma_{k_s}^2 = C(\mathbf{0}) - \boldsymbol{\lambda}^T \cdot \mathbf{k} \quad (2.24)$$

2.4.2 Krigagem Ordinária (KO)

A *krigagem ordinária* é uma técnica de estimativa linear para uma variável regionalizada que satisfaz a hipótese intrínseca (LNADIN, 2003). Em contraste com a *krigagem simples* que, sob a hipótese de estacionaridade de segunda ordem (LNADIN, 2003), exige que a média seja conhecida nesse caso a média é desconhecida.

Assim como no método de krigagem simples, o valor desconhecido de $\mathbf{Z}(\mathbf{X}_0)$ pode ser estimado por uma combinação linear dos n valores observados acrescentados de um parâmetro λ_0 (ver equação 2.13), de modo que o estimador seja não tendencioso o que estabelece a condição de médias iguais. Pode-se escrever,

$$E[Z_{x_0}] = E\left[\lambda_0 + \sum_{i=1}^n \lambda_i Z(x_i)\right] \Rightarrow m = \lambda_0 + \sum_{i=1}^n \lambda_i m \quad (2.25)$$

Como o método de KO não requer o conhecimento da média m , para que a Equação 2.25 tenha solução aceitável,

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

Journel (1988) mostra que minimizando a variância do erro levando em conta a condição $\sum_{i=1}^n \lambda_i = 1$, os pesos λ_i são obtidos a partir do seguinte sistema de equações (CAMARGO, 1997).

$$\begin{cases} \sum_{j=1}^N \lambda_j C(\mathbf{x}_i, \mathbf{x}_j) - \alpha = C(\mathbf{x}_i, \mathbf{x}_0) & \text{para } i = 1, \dots, n \\ \sum_{j=1}^n \lambda_j = 1 \end{cases} \quad (2.26)$$

onde,

- $C(\mathbf{x}_i, \mathbf{x}_j)$ e $C(\mathbf{x}_i, \mathbf{x}_0)$ são definidos como no sistema de equações para a KS;
- α é o multiplicador de Lagrange para a minimização da variância do erro.

A variância minimizada do erro (variância de krigagem ordinária) σ_{ko}^2 é obtida como em Journel (1988) citado por Camargo (1997).

$$\sigma_{ko}^2 = var[Z_{x_0} - Z_{x_0}^*] = C(\mathbf{0}) - \sum_{j=1}^N \lambda_j C(\mathbf{x}_j, \mathbf{x}_0) - \alpha \quad (2.27)$$

Assim, finalmente é obtido o sistema matricial para o cálculo usado no algoritmo de calculo da krigagem ordinária.

$$\mathbf{K} \cdot \boldsymbol{\lambda} = \mathbf{k} \quad \Rightarrow \quad \boldsymbol{\lambda} = \mathbf{K}^{-1} \cdot \mathbf{k} \quad (2.28)$$

de modo que \mathbf{K} e \mathbf{k} são as matrizes das covariâncias e $\boldsymbol{\lambda}$ o vetor dos pesos.

$$\mathbf{K} = \begin{pmatrix} C_{11} & C_{12} & \dots & C_{1n} & 1 \\ C_{21} & C_{22} & \dots & C_{2n} & 1 \\ \vdots & \vdots & \dots & \vdots & \vdots \\ C_{n1} & C_{n2} & \dots & 1 & 1 \\ 1 & 1 & \dots & 1 & 0 \end{pmatrix}, \quad \boldsymbol{\lambda} = \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_n \\ \alpha \end{pmatrix} \text{ e } \mathbf{k} = \begin{pmatrix} C_{10} \\ C_{20} \\ \vdots \\ C_{n0} \\ 1 \end{pmatrix}$$

De modo que a variância na krigagem ordinária é obtida por

$$\sigma_{k_0}^2 = C(0) - k^T \cdot k \quad (2.29)$$

2.5 Clima na região Nordeste do Brasil

A maior parte do território brasileiro está compreendida em regiões intertropicais. Essa característica atribui a essa região um grande potencial de incidência solar praticamente o ano todo (PEREIRA et al., 2006). Nesse sentido, as delimitações espaciais que compreendem a área da região Nordeste do Brasil (NEB) foram escolhidas para a observação da variabilidade espacial e as estimativas de irradiância.

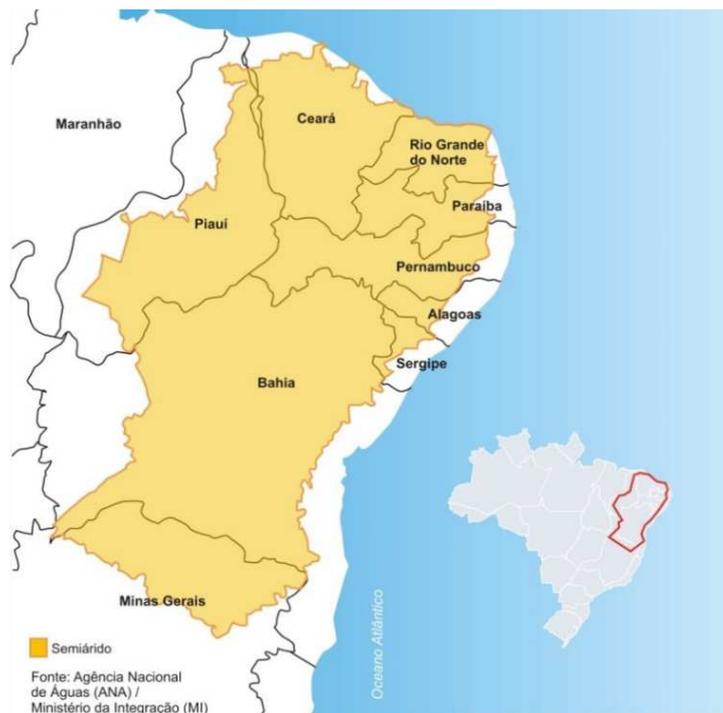


Figura 2.5 – Região semiárida do Nordeste Brasileiro. Essa região é caracterizada principalmente pela forte insolação e temperaturas relativamente elevadas o ano todo.

Fonte: Adaptado de (MOURA et al., 2007).

O NEB localiza-se entre as latitudes 1°02' S a 18°20' S e longitudes 34°47' W a 48°45' W, e possui uma área com cerca de 1,5 milhão de km² (18.2 % do território nacional). Aproximadamente 760.000 km² desta área apresentam um clima semiárido (MOURA et al., 2007). Apesar de sua localização, o NEB não apresenta uma distribuição de chuvas típica de áreas equatoriais sendo influenciado principalmente por três tipos de clima: clima litorâneo úmido (se estendendo do litoral da Bahia ao litoral do Rio Grande do Norte); clima tropical (em áreas dos estados da Bahia, Ceará, Maranhão e Piauí); e clima tropical semiárido (em todo o sertão nordestino) (CAVALCANTI et al., 2009).

Como principais características, o clima da região semiárida é quente e seco, com temperatura média anual de 27° C e 2500 horas/ano de insolação média. A precipitação anual de chuvas varia de 400 a 600 mm, contrastando com uma intensa taxa de evapotranspiração de 2.500 mm/ano. (GALVÃO, 1967).

Há, entretanto, uma variabilidade interanual na precipitação e, portanto, na cobertura de nuvens, fator que influencia diretamente no estudo da radiação solar incidente na superfície. Cavalcanti, et al. (2009), Molion e Bernardo (2002), Vianello e Alves (1991) entre outros, mostraram também que a Zona de Convergência Intertropical (ZCIT) além de vários outros fenômenos que atuam nessa região possui influência considerável no regime de chuvas.

Pesquisadores como Cavalcante, et al. (2009), Moura, et al.(2007) e Azevedo, et al. (1998) descreveram pelo menos cinco mecanismos como principais influencias no regime de precipitação no NEB: 1) Eventos El Niño-Oscilação Sul (ENOS); 2) Temperatura da Superfície do Mar (TSM) na bacia do oceano Atlântico, Ventos Alísios e Pressão ao Nível do Mar (PNM); 3) Zona de Convergência Intertropical (ZCIT) sobre o oceano Atlântico; 4) Frentes Frias e 5) Vórtices Ciclônicos de Altos Níveis (VCAN).

Como foi apontado, o território do NEB é consideravelmente extenso e seu regime de chuvas é influenciado por sistemas atmosféricos distintos. Desse ponto de vista,

é interessante avaliar a climatologia da nebulosidade nessa região. As Figuras 2.6 e 2.7 exibem normais climatológicas de nebulosidade sobre todo o território do NEB para os meses de: Dezembro, Janeiro, Fevereiro, Março (Figura 2.6) e Junho, Julho, Agosto, Setembro (Figura 2.7). Moura, et al. (2007) apresenta uma climatologia da precipitação mensal da região NEB onde o período chuvoso é predominante no verão assim como o período seco é acentuado no inverno.

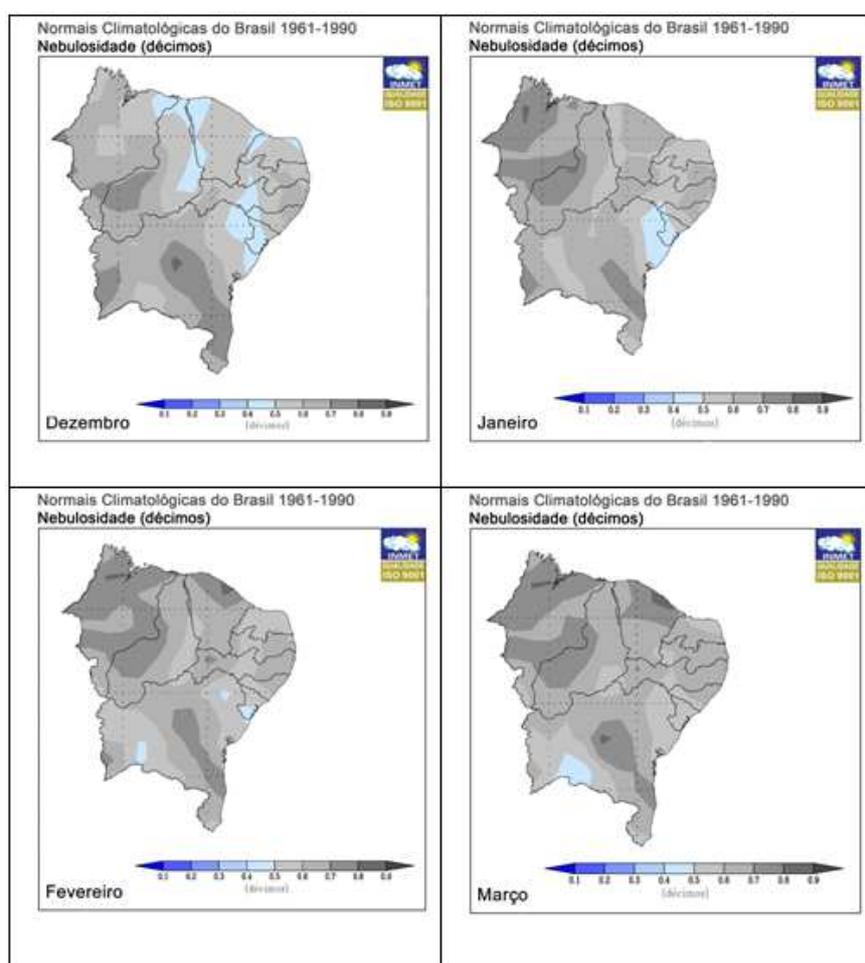


Figura 2.6 – Período Chuvoso: normais climatológicas de nebulosidade (em décimos) para os meses de Dezembro, Janeiro, Fevereiro e Março.

Fonte: Adaptada de INMET (2012).

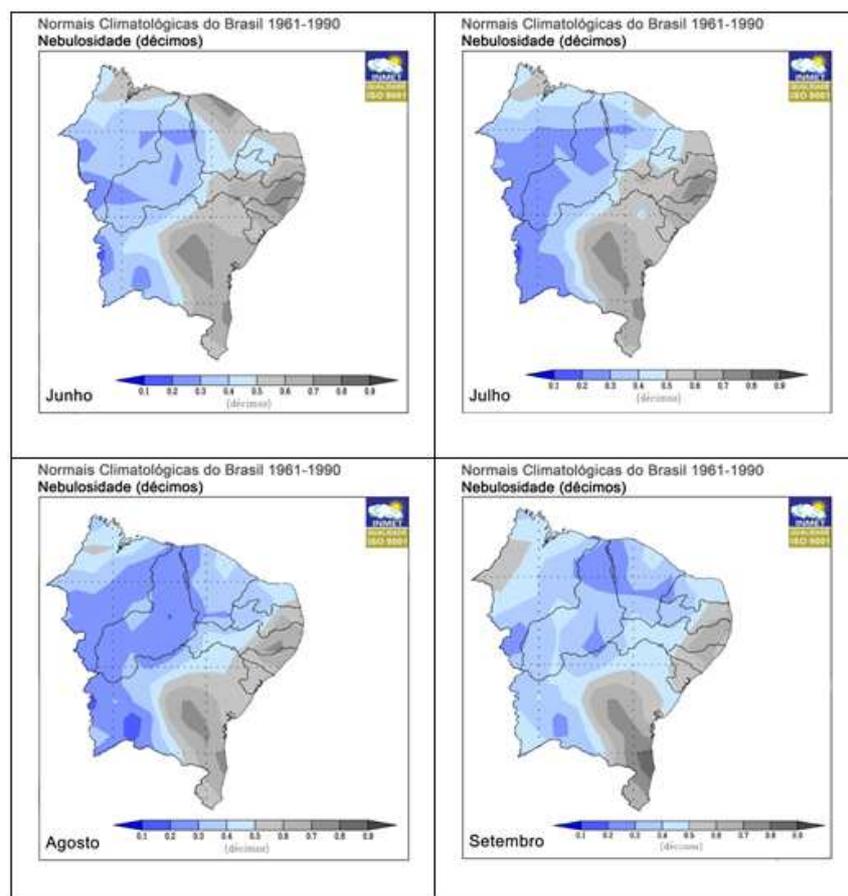


Figura 2.7 – Período seco: normais climatológicas de nebulosidade (em décimos) para os meses de Junho, Julho, Agosto e Setembro.

Fonte: Adaptada de INMET (2012).

Pode-se observar que no período chuvoso da climatologia exibida na Figura 2.6 à ocorrência de nebulosidades moderadas em todo o território nordestino. Entretanto, a Figura 2.7 revela para os meses secos (junho, julho, agosto e setembro) a ocorrência de nebulosidade fraca principalmente no sertão nordestino. Espera-se que o aproveitamento das informações sobre os recursos solares no período seco, assim como no período chuvoso sejam largamente difundidos e utilizados nos diversos setores de atividades humana.

3. ANÁLISE EXPLORATÓRIA E QUALIFICAÇÃO DOS DADOS

Este capítulo é dedicado a descrever a metodologia aplicada no tratamento dos dados medidos pelas PCDs e usados nas interpolações. O tratamento dos dados consistiu em analisar os arquivos estruturalmente realizando uma análise exploratória que consistiu em adequação, validação/qualificação e preenchimento de eventuais falhas nas séries de dados. Todo o tratamento e análise dos dados foram realizados em ambiente Matlab© (MathWorks Inc. TM)

3.1 Estrutura dos dados

Os dados utilizados nessa pesquisa foram disponibilizados pelo Instituto Nacional de Meteorologia (INMET) após solicitação prévia. A base original de dados é constituída por valores médios horários de irradiação solar global e precipitação medidos em estações do tipo PCD operadas pelo INMET e distribuídas espacialmente pela região Nordeste do Brasil. Ao todo, foram disponibilizados para a região de estudo 118 arquivos de dados separados por estado como apresentado na Tabela 3.1. Cada arquivo é dotado da coordenada geográfica (latitude e longitude) e a altitude do local onde a PCD está posicionada – representando um ponto: dado amostral.

As séries temporais de dados correspondem a um período de sete anos. Têm início em 1 de janeiro de 2005 e fim em 31 de dezembro de 2011. A base de dados apresentou séries incompletas em todas os pontos amostrais – tanto por falhas na aquisição de dados quanto por carência de manutenção. Falhas por carência de manutenção representaram lacunas nas séries de dados por períodos maiores: desde dias até meses sem dados medidos.

A ocorrência de falhas longas nas séries de dados implicou na adoção de uma metodologia para garantir que ao menos 2/3 dos dados utilizados nas interpolações sejam dados medidos pelas PCDs – e não preenchidos por regressão linear. Desse

modo, foi possível escolher somente os pontos amostrais que atenderam critérios mínimos para utilização de seus dados. Esse critério e o método para preenchimento de dados faltantes serão discutidos com detalhes adiante. Esse processo reduziu a amplitude da série de 7 para 4 anos.

Tabela 3.1 - Discriminação dos arquivos de dados fornecidos pelo INMET em números de arquivos com dados amostrais por unidade federal da região Nordeste do Brasil.

Estado (UF)	Nº de arquivos
Alagoas (AL)	6
Bahia (BA)	38
Ceará (CE)	14
Maranhão (MA)	11
Paraíba (PB)	8
Pernambuco (PE)	11
Piauí (PI)	18
Rio Grande do Norte (RN)	7
Sergipe (SE)	5

3.2 Qualificação dos dados

Alguns fatores podem afetar a confiabilidade dos dados medidos por equipamentos automáticos como os piranômetros que compõem as PCDs. Eventuais defeitos no equipamento de medida, interrupções de energia, descargas elétricas, mudanças no ambiente onde está o equipamento medidor – como projeções de sombras por árvores ou prédios – podem alterar as medições.

A qualificação dos dados foi realizada segundo as especificações da BSRN (*Baseline Surface Radiation Network* - <http://www.bsrn.awi.de/en/home/bsrn/>) e da WMO (*World Meteorological Organization*). A BSRN é um projeto coordenado pelo do Programa Mundial de Pesquisas Climáticas (WCRP sigla em inglês) em parceria como o GEWEX (sigla em inglês para *Energy and Water Cycle Experiment*) com o objetivo de detectar alterações relevantes no campo de radiação medido na

superfície da Terra e que podem estar associados a mudanças climáticas. A organização WMO (em português: Organização Meteorológica Mundial) é um órgão internacional vinculado as Nações Unidas que coordena pesquisas e fornece suporte para as áreas de Meteorologia, clima e ciências correlatas.

Roesch et al. (2011) avaliando os registros do banco de dados de irradiação disponibilizados pela BSRN para o cálculo de médias mensais, determinou os limites aceitáveis para eventos extremamente raros e fisicamente possíveis. Os limites são exibidos nas Tabelas 3.2 e 3.3 para a qualificação dos dados como apresentados por MacArthur (2004).

Tabela 3.2 - Limites utilizados para derivar os produtos de radiação GEWEX. Valores fisicamente possíveis. Os valores além dos intervalos indicados foram tratados como ausentes. S_0 é a constante solar ajustada para distância Terra-Sol. μ é o cosseno do ângulo zenital.

Parâmetro	Limite inferior	Limite superior
Global	-4 Wm^{-2}	$1,50 \cdot S_0 \cdot \mu^{1,2} + 100 \text{ Wm}^{-2}$
Difusa	-4 Wm^{-2}	$0,95 \cdot S_0 \cdot \mu^{1,2} + 100 \text{ Wm}^{-2}$
Direta	-4 Wm^{-2}	$S_0 \cdot \mu^{1,2}$

Fonte: Adaptada de MacArthur (2004).

Tabela 3.3 - Semelhante à Tabela 3.2, apresentando os intervalos de "Extremamente raros" para qualificação de dados coletados de irradiação solar.

Parâmetro	Limite	Limite superior
Global	-2 Wm^{-2}	$1.2 \cdot S_0 \cdot \mu^{1,2} + 50 \text{ Wm}^{-2}$
Difusa	-2 Wm^{-2}	$0.75 \cdot S_0 \cdot \mu^{1,2} + 0 \text{ Wm}^{-2}$
Direta	-2 Wm^{-2}	$0.95 \cdot S_0 \cdot \mu^{1,2} + 10 \text{ Wm}^{-2}$

Fonte: Adaptada de MacArthur (2004).

Os dados disponibilizados pelo INMET foram tratados inicialmente como brutos, ou seja, dados que não foram submetidos a um processo de qualificação no que diz respeito à verificação por testes de confiabilidade. Após a análise estrutural dos arquivos, foi empregado em toda a série de dados um processo de qualificação similar ao adotado para estações automáticas de superfície participantes da BSRN como definido por Roesch et al. (2011) e descrito em MacARTHUR (2004). A Figura 3.1 apresenta um diagrama de blocos que exemplifica o algoritmo de qualificação pelo qual foram submetidos os dados amostrais.

Desse algoritmo resultaram as PCDs que contêm apenas dados medidos dentro dos limites impostos no método de qualificação. No procedimento a seguir, esses dados são analisados quantitativamente de modo a definir as “melhores” PCDs a serem utilizadas nas comparações representando os dados observados.

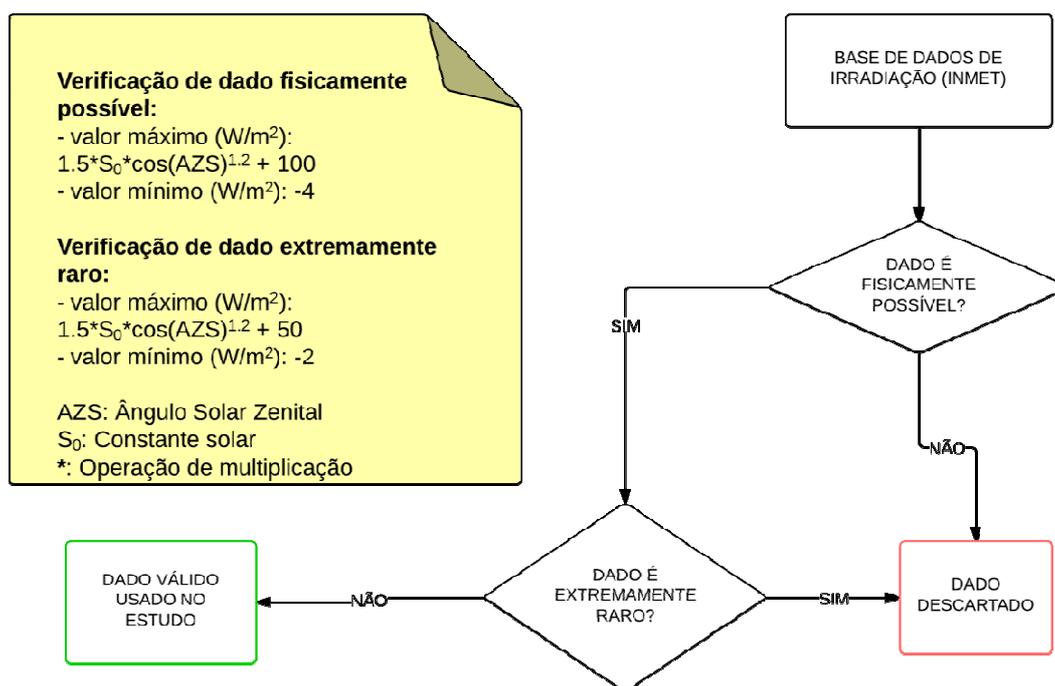


Figura 3.1 – Diagrama esquemático demonstrando os passos para a qualificação dos dados utilizados nesse estudo.

3.3 Análise qualitativa dos dados caracterizados

Nessa etapa foi aplicado um diagnóstico para seleção das PCDs de acordo com a disponibilidade de dados após o procedimento de qualificação. Desse modo, foram selecionadas as PCDs que apresentaram uma quantidade de dados observados que atendam os seguintes critérios:

- a) **Representatividade dos horários de maior intensidade na incidência de radiação solar nas estações de coleta de dados:** a maior intensidade do fluxo de irradiação ocorre nas horas próximas ao meio dia solar (MS) e, a fim de se obter uma banda segura de medições em torno da hora que corresponde ao MS local, foi considerado como período de maior intensidade de irradiação solar as horas que compreendem o intervalo de 9h as 16h. Esse artifício garantiu que no período de maior intensidade de irradiação não há preenchimento de falhas – por regressão linear – na coleta de dados para determinação do total diário de irradiação solar, o que, teoricamente, aumentou a confiabilidade do valor de total diário a ser utilizado na interpolação espacial. A Figura 3.2 demonstra o algoritmo usado na execução dessa etapa.

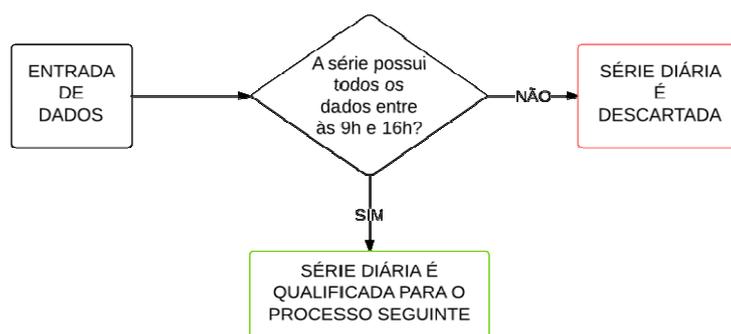


Figura 3.2 - Diagrama esquemático demonstrando a primeira etapa para a escolha das PCDs a serem utilizadas neste estudo – aquelas que contêm todos os dados medidos entre as 9h e 16h.

b) Representatividade de 2/3 dos dados medidos: esse critério assume como fator limitante (aceitável) o número de falhas inferior a 1/3 da extensão total das séries, ou seja, foram consideradas como aceitáveis as PCDs que continham séries de dados medidos com no mínimo 20 dias num mês e 8 meses num ano. As Figuras 3.3 (a) e 3.3 (b) demonstram o digrama de blocos do algoritmo usado nessa etapa.

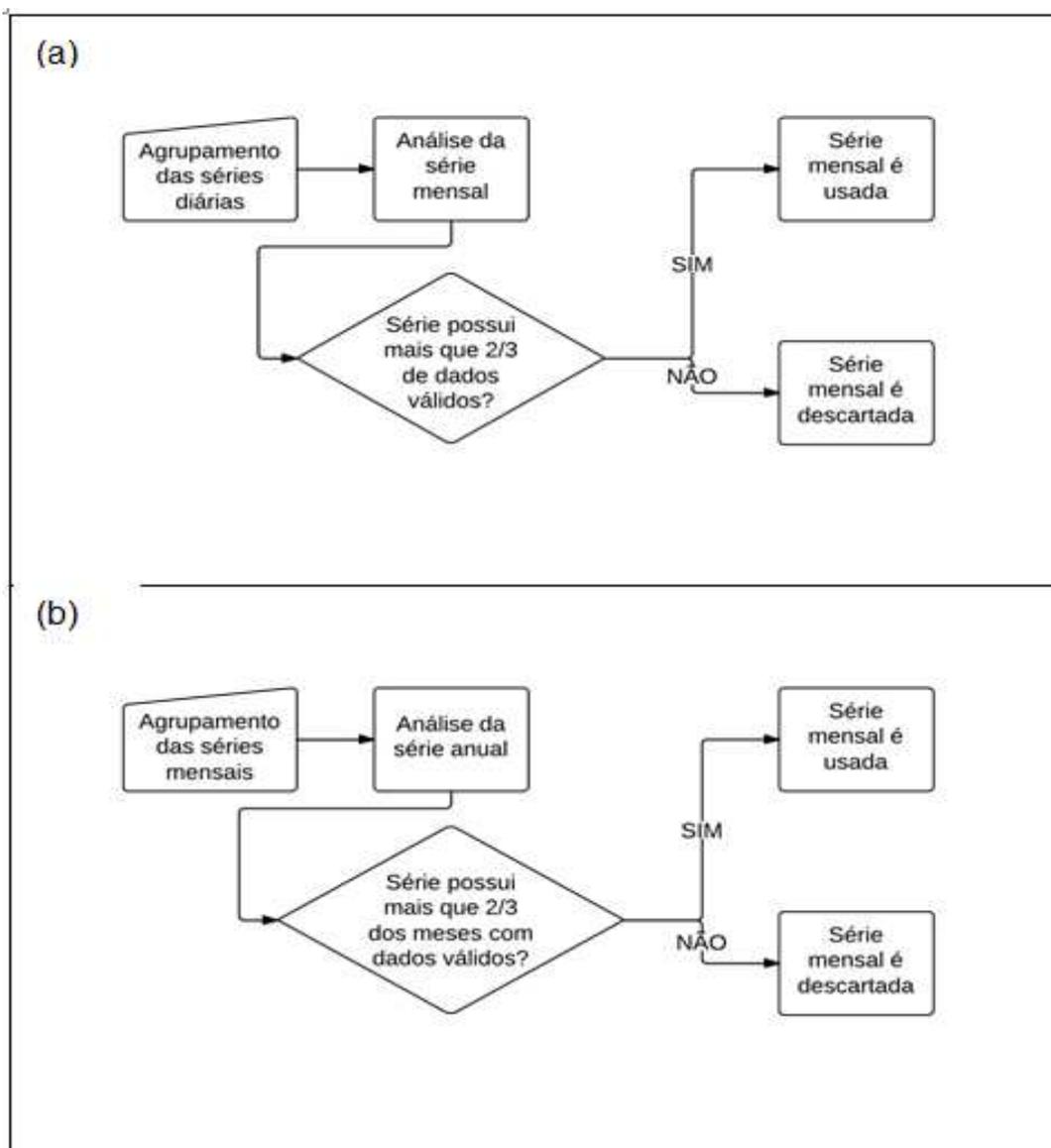


Figura 3.3 - Diagrama esquemático demonstrando os passos para a escolha das séries de dados com no mínimo 2/3 de dados medidos pelas PCDs para (a) séries mensais e (b) séries anuais.

Assim, como resultado dessa metodologia aplicada as séries de dados observados em todas as PCDs (espaço amostral), foram excluídos dois dos pontos interpoladores – estações medidoras que não atenderam os critérios adotados. Assim, um total de 116 pontos interpoladores (PCDs) foram selecionados para alimentar o procedimento de interpolação espacial por krigagem (Figura 3.4). Vale lembrar que a disponibilidade de dados em cada PCD não era contínua e coincidente com as demais PCDs. Dessa etapa resultou o seguinte panorama: para a séries anuais de dados (2005 a 2011), ficaram disponíveis 116 pontos interpoladores onde, cada ponto interpolador representa os dados amostrais medidos por uma PCD.

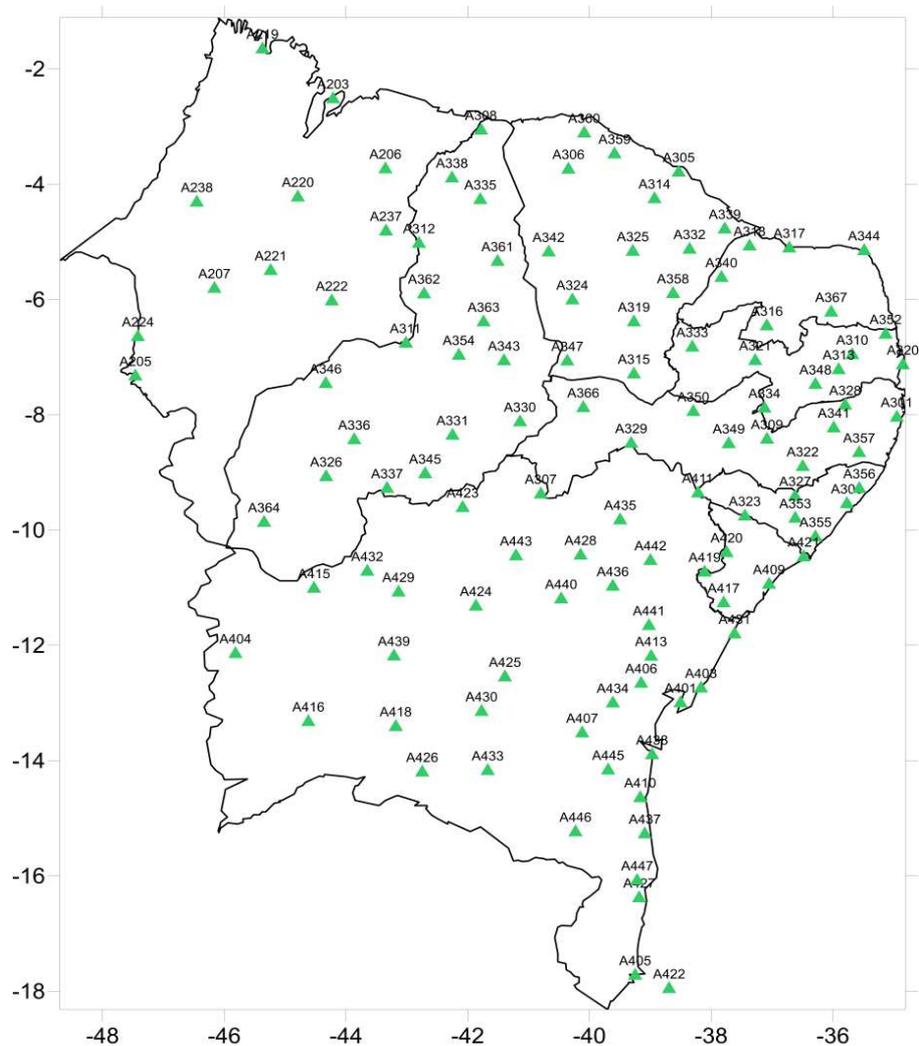


Figura 3.4 – Arranjo espacial das PCDs que passaram pelos testes de qualificação.

No entanto, Landim (2000) recomenda que o número de pontos amostrados não deve ser menor do que 30 na aplicação do método de interpolação por krigagem, devido o número insuficiente de pares de amostras para a modelagem do semivariograma. Sendo assim, visando garantir a qualidade mínima dos resultados das interpolações, foi adotado um limite inferior de, no mínimo, 40 pontos interpoladores (40 PCDs). A Figura 3.5 apresenta a disponibilidade de PCDs por meses ao longo de dos anos de 2005 a 2011.

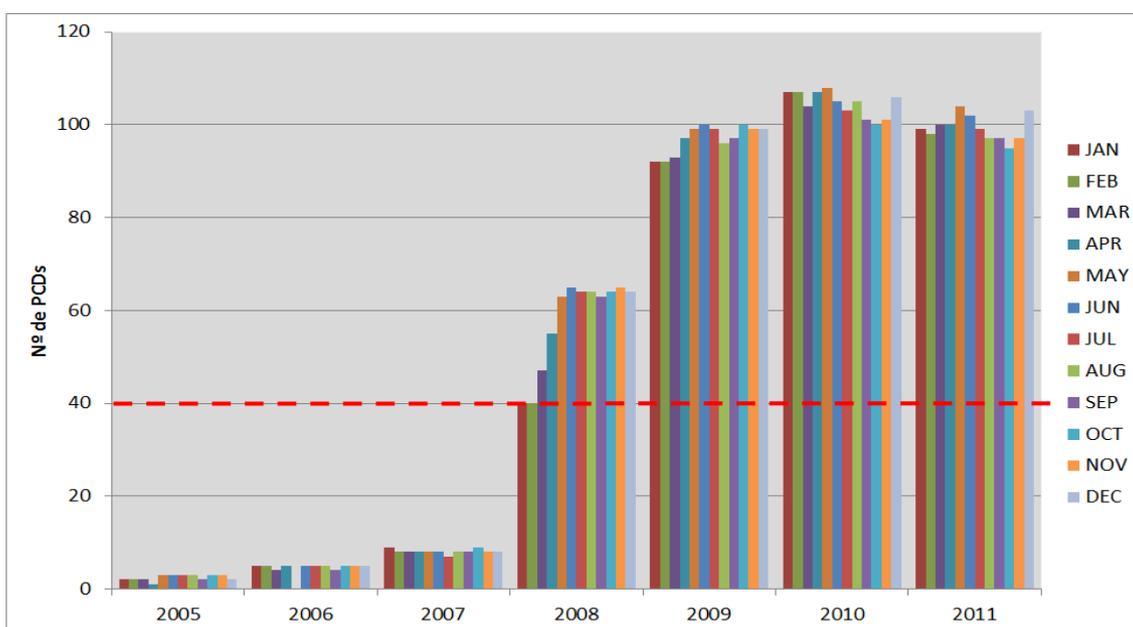


Figura 3.5 – Numero de PCDs qualificadas distribuídas por meses num ano. A linha tracejada em vermelho representa o limite mínimo aceitável para a utilização da série anual no processo de interpolação espacial por krigagem. As séries de 2005, 2006 e 2007 foram descartadas.

Assim, tomando como referência um grupo não inferior a 40 PCDs qualificadas, foram eliminadas do estudo as séries de dados correspondentes aos anos de 2005, 2006 e 2007. Isso ocorreu, pois nesses anos o número de pontos amostrais resultantes dos processos de qualificação e ajuste foi inferior a 40 (Figura 3.5).

3.4 Preenchimento de dados faltantes

Após as etapas de qualificação e ajuste dos dados, restou a verificação das séries diárias a fim de preencher eventuais falhas nessas séries. Como no processo de ajuste foram selecionadas as PCDs que contêm todos os dados no período de maior intensidade na incidência de radiação solar, nessa etapa foi executada uma rotina para preencher as séries diárias que contenham falhas em horários fora desse intervalo.

Falhas em séries temporais de dados meteorológicos e climatológicos podem ocorrer por diversos motivos, mau funcionamento dos instrumentos de medição, interrupções de energia elétrica, substituição dos instrumentos da estação, etc (CÁRDENAS; KRAINSK, 2011). Nos dados aqui avaliados, além das possíveis fontes de falhas apontadas, medidas que não passaram no processo de qualificação foram excluídas das séries.

Diversos são os métodos para atribuir dados faltantes em séries temporais. Fernandez (2007) apresentou diferentes metodologias estatísticas para o preenchimento de falhas em séries de dados meteorológicos diários. Entre elas, técnicas matemáticas como regressão linear simples e múltipla são habitualmente usadas e produzem resultados satisfatórios perante outros métodos (FERNANDEZ, 2007; CÁRDENAS; KRAINSKI, 2011).

Desse modo, nesse estudo foi aplicado um método de regressão linear simples baseado no índice de transmissividade k_t . O índice k_t é calculado com base na radiação no topo da atmosfera numa dada posição aparente do Sol na esfera celeste. Esse método simples garante adequação e confiabilidade nos resultados das atribuições de valores para o preenchimento das falhas nas séries de dados em medições de irradiação solar.

3.4.1 Índice de transmissividade k_t

O índice de transmissividade k_t assume que há uma relação linear entre as condições de céu claro e céu totalmente nublado. O índice k_t é definido pela razão entre a irradiação global (R_g) e a irradiação solar incidente no topo da atmosfera (TOA) como mostra a Equação 3.1 – por convenção, a radiação TOA é representada aqui como I_0 .

$$k_t = \frac{R_g}{I_0} \quad (3.1)$$

Esse índice pode variar de acordo com a nebulosidade e quantidade de aerossóis na atmosfera (QUERINO et al., 2010). Em condições de céu totalmente claro e sem a presença de aerossóis, k_t aproxima-se de 1 e em situação de céu total ou parcialmente nublado onde R_g tende a diminuir quando medida em superfície, k_t aproxime-se de 0. É importante ressaltar que em condições reais, os valores do índice k_t nunca chegarão a 1 e 0. Na verdade, mesmo com o céu totalmente claro e sem nuvens ainda há presença de gases e material particulado que interagem com a radiação. Em relação a dias completamente nublados, há a presença da radiação difusa que não deixa o índice chegar a 0.

3.4.2 Preenchimento de falhas nas séries de diárias utilizando o índice k_t

Nessa etapa, foram apurados todos os dias das séries resultantes de todas as etapas anteriores a fim de identificar e preencher as falhas nas medidas utilizando o método de regressão linear simples baseado no índice de k_t .

A aplicação da metodologia para preenchimento dos dados faltantes teve início no cálculo de I_0 e, em seguida, foram calculados os índices kt para todas as medidas horárias. Por fim, foi feita uma averiguação das séries diárias a fim de atribuir dados em possíveis falhas nas séries. As equações e constantes empregadas em todos os cálculos estão descritas como em Iqbal (1983).

3.4.3 Cálculo da radiação TOA e do índice kt

É possível determinar precisamente I_0 em cada posição aparente do Sol no céu ao longo do dia. Para tanto, são necessárias informações da órbita da Terra, medidas de tempo e da posição aparente do Sol na esfera celeste. Diante dessas informações, aplicando a Equação 3.2 pôde-se obter o valor de I_0 em cada hora do dia.

$$I_0 = S_0 \cos(\theta_z) \quad (3.2)$$

S_0 é a constante solar e θ_z o ângulo zenital.

O cálculo aplicado para conseguir as informações da órbita terrestre possui variação diária e possibilita obter S_0 para cada dia juliano. Certamente, devida à trajetória elíptica de nosso planeta, é necessária uma correção diária no cálculo da distância Terra-Sol ($r_{2sun,i}$). A equação utilizada para essa correção é dada pela Equação 3.3 apresentada a seguir.

$$r_{2sun,i} = r1 + r2 \cos(\gamma) + r3 \text{sen}(\gamma) + r4 \cos(2\gamma) + r5 \text{sen}(2\gamma) \quad (3.3)$$

Onde o índice i representa o valor do i -ésimo dia do ano; $r1, r2, r3, r4$ e $r5$ são constantes dadas por: $r1 = 1,0001$; $r2 = 0,034221$; $r3 = 0,00128$; $r4 = 0,000719$;

$r_5 = 0,000077$; γ representa o deslocamento angular diário do Sol medido em *rad* e dado pela Equação 3.4.

$$\gamma_i = \frac{[2\pi(d_{\text{juliano}} - 1)]}{365} \quad (3.4)$$

Desse modo, o fator de correção $r_{2\text{sun}}$ representa a posição da Terra em sua órbita ao longo do ano.

Portanto, pode-se calcular o valor de S_0 em cada dia do ano a partir de seu valor médio e do fator $r_{2\text{sun}}$. O valor médio de S_0 utilizado nesse estudo foi de 1367 W/m^2 (IQBAL, 1983).

$$S_{0,i} = 1367 r_{2\text{sun},i} \quad ; \quad i = 1, 2, \dots, n \quad (3.5)$$

O cosseno do ângulo zenital θ_z é obtido através das equações do tempo (d_{tempo}), ângulo horário (ω) e declinação solar (δ). A Equação 3.6 apresenta a equação do tempo. Essa equação corrige o tempo solar médio de modo ajustá-lo ao tempo solar verdadeiro, ou seja, o tempo decorrido após ser completa uma culminação do Sol no céu no referencial do observador – meio dia solar.

$$d_{\text{tempo},i} = c1 + c2 \cos(\gamma) - c3 \text{sen}(\gamma) - c4 \cos(2\gamma) - c5 \text{sen}(2\gamma) \quad (3.6)$$

Como γ é medido em radianos, é necessária uma conversão na unidade de medida de d_{tempo} para a unidade de tempo medida em horas. Para tanto, aplica-se a Equação 3.7.

$$d_{\text{tempo}_h} = d_{\text{tempo}} \frac{180}{15\pi} \quad (3.7)$$

As coordenadas de posição do Sol na esfera celeste necessários para a obtenção do cosseno do ângulo zenital foram calculadas em seguida. Tratam-se do ângulo horário ω e do ângulo de declinação solar δ . Para obter o ângulo horário primeiramente calcula-se o valor do tempo associado à hora legal¹ (Equação 3.8).

$$t_{i,j} = d_{\text{tempo}_h} + \left(\frac{\text{lon} + 45^\circ}{15} + h_{i,j} \right) \quad (3.8)$$

onde lon é a longitude do local onde está a PCD, h representa a hora – local, medida na PCD e j é um índice que indexa a hora percorrendo o período de insolação (como mencionado anteriormente, fica definido como período de insolação, o período de 6h às 19h, fazendo que j varie de 1 a 14).

Assim, o ângulo horário é dado por

$$\omega_{i,j} = 12 - 15t_{i,j} \quad (3.9)$$

Para o cálculo da declinação utiliza-se a Equação 3.10 onde são definidas as constantes:

$$d1 = 0,006918, d2 = 0,399912, d3 = 0,70257, d4 = 0,006758, d5 = 0,000907, d6 = 0,002697 \text{ e } d7 = 0,001480$$

$$\delta_i = d1 - d2 \cos(\gamma_i) + d3 \text{sen}(\gamma_i) - d4 \cos(2\gamma_i) + d5 \text{sen}(2\gamma_i) - d6 \cos(3\gamma_i) + d7 \text{sen}(3\gamma_i) \quad (3.10)$$

Dessa maneira é possível calcular o cosseno do ângulo zenital como segue:

¹ Tempo do fuso horário mais próximo, ou seja, hora civil do meridiano central do fuso. Fonte: <http://astro.if.ufrgs.br/tempo/tempo.htm>. Acesso em 11/09/2013.

$$\cos(\theta_z)_{i,j} = \text{sen}(\delta_i) \text{sen}\left(\text{lat} \frac{\pi}{180^\circ}\right) + \cos(\delta_i) \cos\left(\text{lat} \frac{\pi}{180^\circ}\right) \cos\left(\omega_{i,j} \frac{\pi}{180^\circ}\right) \quad (3.11)$$

onde lat representa a latitude do local onde esta a PCD.

Finalmente, pôde-se calcular a radiação no topo da atmosfera aplicando a Equação 3.2 de modo a atentar-se aos índices que representam as variações diárias (i) e horárias (j).

$$I_{0_{i,j}} = S_{0_i} \cos(\theta_z)_{i,j} \quad (3.12)$$

Assim, sabendo o valor de I_0 para cada hora do dia e utilizando a Equação 3.1, foi possível obter a matriz do índice kt para todas as medidas de R_g ao longo da série de dados.

3.4.4 Aplicação do método de preenchimento de falhas

Nesta ocasião, foi analisada cada série diária a fim de localizar falhas na série. Primeiramente, foram determinados os índices kt das extremidades do intervalo com dados disponíveis. Como já foi definido pela aplicação dos procedimentos descritos anteriormente, o período de maior intensidade de irradiação solar não apresenta falhas e, portanto, haverá índices kt obtidos por dados medidos para os valores correspondentes às 9h e 16h. Assim, foram aplicados os valores do índice kt correspondentes aos dois extremos de cada falha encontrada, para o preenchimento da mesma.

3.5 Análise de desempenho: índices estatísticos e validação cruzada

A seguir são apresentados os métodos utilizados para medir, quantitativamente, os desempenhos das estimativas por KO. Foi aplicado o método de validação cruzada aliado a análise dos índices estatísticos coeficiente de correlação (r^2), **VIES** e o **EQMA**.

O algoritmo empregado na validação cruzada foi aplicado da seguinte maneira: a cada rodada (N =1 até o número total de pontos amostrais de referência) foi retirado o ponto amostral conhecido da série total dos dados e foram realizadas as estimativas empregando todos os outros pontos. Em seguida, os valores estimados para os pontos amostrais são comparados com os valores reais – retirados anteriormente. Na comparação são empregados os índices estatísticos de modo a medir o desempenho das estimativas.

Os índices estatísticos aplicados são definidos a seguir:

O coeficiente de determinação (r^2): é uma medida descritiva da qualidade do ajuste obtido. Esse índice possibilita uma medida da proporção da variabilidade (ou da variância) dos dados medidos que pôde ser explicada pelos dados estimados por cada modelo. O coeficiente de determinação pode ser obtido através do ajuste de uma reta num gráfico de dispersão dos valores medidos contra os valores estimados e assume um intervalo de -1 a 1 ($-1 \leq r^2 \leq +1$), onde:

- Seus valores negativos indicam que as estimativas possuem tendência a subestimar os dados medidos;
- Seus valores positivos indicam que as estimativas possuem tendência a superestimar os dados medidos;
- Os valores próximos a 0 indicam os piores resultados onde as estimativas não podem explicar a variabilidade dos dados observados;

- Valores próximos as extremidades -1 e +1 são considerados ideias, pois significam que as estimativas puderam explicar a variabilidade dos dados medidos.

VIÉS: o *VIÉS* é obtido pela diferença entre o valor estimado e o valor medido. Como pode ser entendido como resíduo, o *VIÉS* aponta erros sistemáticos ou tendências em regiões onde os valores foram superestimados ou subestimados;

EQMA (Erro Quadrático Médio Absoluto): o *EQMA* é uma forma de avaliar a diferença entre o valor estimado e o valor medido. É absoluto, pois avalia o erro em relação ao valor medido ponderado pela média da série (Equação 3.13).

$$EQMA = \left[\frac{\sqrt{1/N \sum_{i=1}^N (R_{e_i} - R_{m_i})^2}}{\bar{R}_m} \right] \quad (3.13)$$

3.6 Interpolação espacial

Nesta seção é descrita a metodologia empregada no processo de interpolação. Na execução da interpolação foi usado o software Surfer© (Golden SoftwaereTM Inc.) e uma adaptação do código disponibilizado por Wolfgang Schwanghart^{2,3} em ambiente Matlab© tanto para a interpolação por krigagem ordinária bem como na geração dos semivariogramas e mapas. Devido ao arranjo dos dados sob a forma de média anual dos totais diários (soma das médias horárias acumuladas num dia), a unidade de medida dos dados, assim como das estimativas passa a ser em Wh/m² (lê-se Watts hora por metro quadrado). O procedimento de interpolação foi dividido em duas etapas:

² Copyright© 2010, Wolfgang Schwanghart All rights reserved.

³ Disponível em: <http://www.mathworks.com/matlabcentral/fileexchange/authors/25218>

- I. Análise quantitativa dos modelos variográficos mais usuais e averiguação de tendência e anisotropia na distribuição espacial dos dados;
- II. Aplicação do método de interpolação segundo definições e resultados da etapa anterior.

A etapa (I) teve objetivo de escolher o modelo variográfico mais adequado entre os modelos mais usuais: esférico, gaussiano e exponencial e, ainda, verificar a presença de anisotropia e tendência significativas na distribuição espacial dos dados amostrais e seu impacto nas estimativas. A etapa (II) consistiu na realização da KO para pontos escolhidos como referência (aqui, estações ou PCDs de referência são os pontos amostrais escolhidos para representar os dados medidos diante das intercomparações com as estimativas) – o processo de escolha dos pontos de referência é discutido adiante.

3.6.1 Etapa I

Em busca de uma representação espacial homogênea, nesta etapa foram escolhidos 6 pontos amostrais distribuídos espacialmente de modo a cobrir uma área considerável do NEB como mostram a Figura 3.6 e a Tabela 3.4. Desses pontos, 5 foram usados para a determinação do modelo variográfico e 4 foram usados para a averiguação da presença de anisotropia. Assim, esses pontos amostrais foram utilizados na comparação com as estimativas produzidas por KO na escolha do modelo variográfico mais adequado para os dados respectivos as médias anuais do ano de 2010. O ano de 2010 foi escolhido para essa comparação pois, de todos os anos disponíveis, esse foi o período que apresentou o menor número de falhas na série.

O desempenho dos modelos variográficos teóricos foi verificado pelo método de validação cruzada onde foi realizada uma análise dos índices estatísticos: r^2 , *VIES* e o *EMQA*.

O ajuste do modelo variográfico ao semivariograma experimental é feito de forma visual. Foram realizados múltiplos ajustes dos semivariogramas para encontrar o modelo mais adequado. Para todos os pontos amostrais de referência (Tabela 3.4), foi avaliado cada um dos três modelos variográficos (esférico, gaussiano e exponencial). O procedimento consistiu no ajuste visual, variando os parâmetros (C_0 , a_0 , C : *Efeito pepita*, *Alcance* e *Contribuição* respectivamente) do semivariograma e, posteriormente, foi aplicada a validação cruzada a cada interação, para cada modelo e para cada ponto amostral (PCD). Após a análise da validação cruzada foi possível avaliar o desempenho dos modelos obtendo o índice r^2 para a média anual do período – ano de 2010.

Tabela 3.5 – Índices estatísticos para cada um dos modelos variográficos. O melhor desempenho é observado no modelo Esférico.

Índices estatísticos	Modelos variográficos		
	Esférico	Gaussiano	Exponencial
r^2	0,70	0,31	0,41
<i>EQMA (%)</i>	6,64	12,31	0,43

A Tabela 3.5 apresenta os resultados da escolha do modelo variográfico e a Figura 3.7 mostra a dispersão das estimativas em relação aos dados amostrais. Entre os três modelos variográficos, o modelo esférico foi o que apresentou o melhor desempenho – em relação ao valor de r^2 – explicando cerca de 70 % da variância dos dados amostrais. Já o modelo exponencial, apresentou o menor valor do *EQMA (%)* mas, no entanto, explicou apenas 40 % da variabilidade dos dados amostrais. Assim, no geral, o modelo variográfico esférico obteve o melhor desempenho. Desse modo, mostrou-se que o modelo esférico foi superior aos outros modelos justificando sua escolha no ajuste dos semivariogramas experimentais para as estimativas de irradiação de todo o NEB.

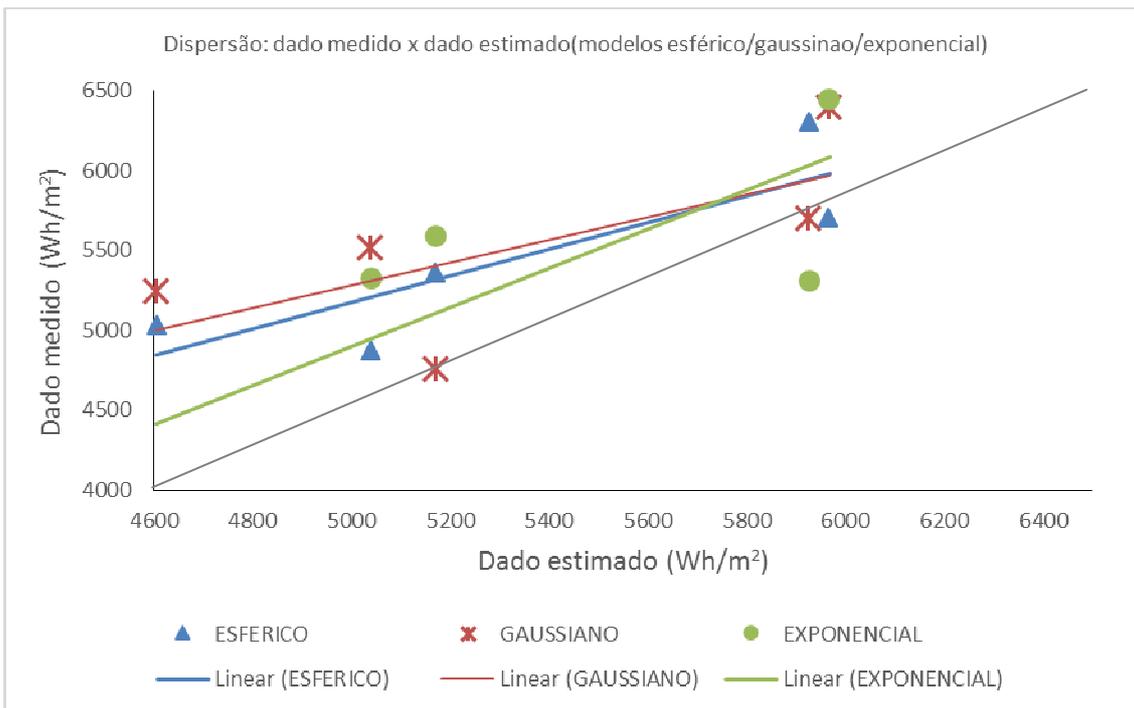


Figura 3.7 – Gráfico de dispersão: dados medidos contra a sobreposição das estimativas produzidas através do ajuste de 3 modelos variográficos. A comparação foi realizada para o período de 2010 utilizando os pontos amostrais listados na Tabela 3.4.

A seguir, foi feita uma verificação a fim de identificar a presença de anisotropia nas séries de dados e sua influência nas estimativas. Para tanto, no ajuste dos semivariogramas, foram fixadas direções intermediárias compreendendo o arco de 0° a 180° dividido em frações 45°. Para a representatividade espacial dos dados medidos, foram escolhidos 4 pontos amostrais: as PCDs A221, A317, A327 e A446. Essas PCDs foram escolhidas pois suas localizações apresentam condições distintas que possam vir a influenciar as estimativas tais como: proximidade a costa (A317); proximidade ao limite geográfico da região (A221); ausência de PCDs circundantes em uma dada direção (446) e maior densidade de PCDs próximas.

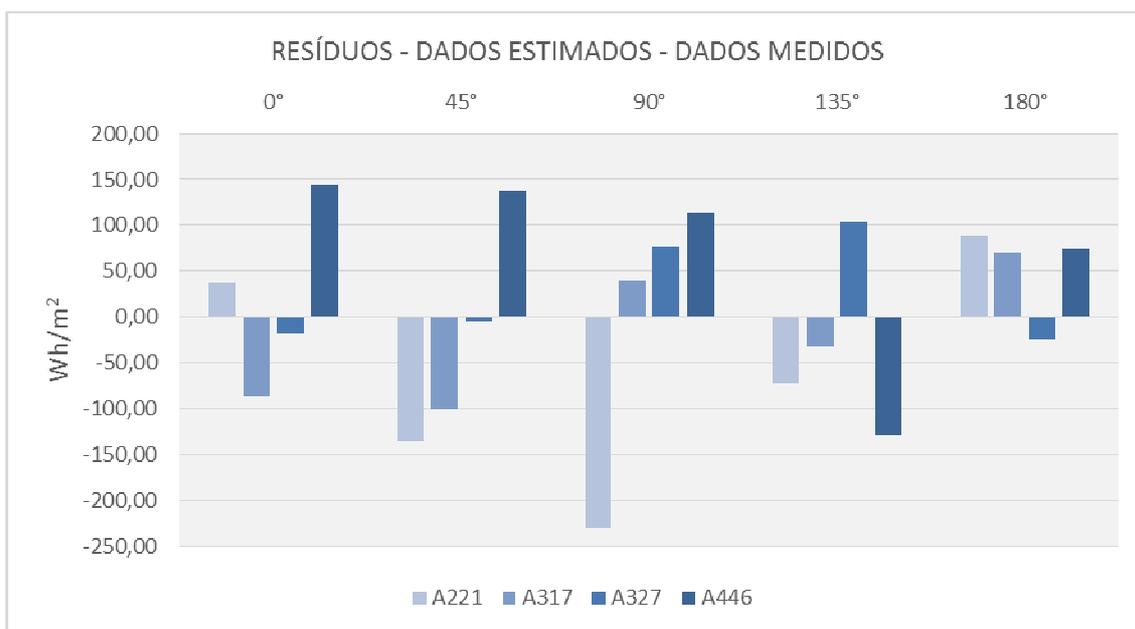


Figura 3.8 – Valores dos desvios das estimativas por KO – semivariograma experimental ajustado pelo modelo esférico – considerando as direções de 0°, 45°, 90°, 135° e 180° para cada ponto amostral.

A Figura 3.8 mostra os desvios das estimativas em relação aos dados medidos nos pontos amostrais. De acordo com os resultados, é possível observar que o maior desvio – relativo a PCD A221 – apresentou valor em torno de 230 Wh/m². Esse valor corresponde a apenas 4,5 % do valor médio das estimativas para esse ponto. Em todos os pontos amostrais pôde-se observar a presença de anisotropia – mais acentuada nos pontos das PCDs A221 e A446 provavelmente, pela ausência de PCDs próximas, principalmente no sentido sul. Contudo, apesar da presença de anisotropia ser identificada nos resultados, foi descartada a sua influência significativa no resultado das estimativas. Desse modo, pôde-se assumir um tipo de variograma onidimensional – sem direção privilegiada.

Por fim, foi feita uma investigação a respeito da existência de tendência nos dados de entrada. O procedimento de ajuste dos semivariogramas foi similar ao da comparação entre os modelos variográficos teóricos. Foram comparados três casos na entrada dos dados, todos eles incorporados no código do Software Surfer©. A

escolha por um caso ou outro é feita simplesmente escolhendo o tipo de entrada de dados antes da geração do semivariograma.

- *Caso 1*: entrada de dados sem tendência;
- *Caso 2*: entrada de dados com tendência linear (modelada linearmente);
- *Caso 3*: entrada de dados com tendência quadrática (modelada de 2ª ordem).

Os pontos amostrais correspondentes as PCDs A327 e A446 não apresentaram variações significativas entre os casos mas, no entanto, a PCD A446 subestimou os dados, o que pode ser explicado por sua posição – extremidade sul da região NEB – e pela ausência de PCDs circundantes em regiões próximas. As maiores variações entre os casos de tendência foram observadas nos pontos correspondentes as PCDs A221 e A317.

Tabela 3.6 – Apresentação dos resíduos (Wh/m²) nos três casos de entradas de dados para todas os pontos amostrais: A221, A317, A327 e A446. A última linha representa a média das estimativas em cada ponto.

TENDÊNCIA (TIPO)	A221	A317	A327	A446
Caso1	6,81	40,20	6,70	-181,26
Caso 2	41,70	5,27	4,53	-171,91
Caso 3	-34,94	49,71	2,02	-131,20
MÉDIA	5106,86	5433,79	5064,93	4719,65

Considerando os valores mínimos e máximos em cada caso, as diferenças entre os resíduos das estimativas foram de 16 % e 10 % para os pontos A221 e A317 respectivamente. Em relação à média das estimativas, o desvio máximo observado (A317: $49,71 \text{ Wh/m}^2 - 5,27 \text{ Wh/m}^2 = 44,44 \text{ Wh/m}^2$) representa um valor menor do que 1 %.

Assim, a validação cruzada mostrou que não houve diferenças significativas nas estimativas em relação à entrada de dados com ou sem tendência nos pontos

amostrais considerados. Por isso, foi assumido que os dados de entrada para a krigagem não possuem tendência significativa.

Portanto, a etapa I definiu o modelo teórico e as características dos dados a serem usados no ajuste dos semivariogramas experimentais. Os resultados mostraram que o modelo teórico esférico obteve o melhor desempenho diante dos outros modelos testados e, ainda, que se pôde desconsiderar a presença de tendência e anisotropia significativas na série de dados.

3.6.2 Etapa II

Essa etapa consistiu na delimitação dos pontos amostrais mais adequados considerando os dados de todos os 116 pontos amostrais disponíveis, no ajuste dos semivariogramas, realização da KO para distintos pontos amostrais escolhidos como referência e a produção dos mapas de variabilidade espacial de irradiação solar.

3.7 Definição das PCDs de referência

Antes de iniciar o processo de krigagem, foram escolhidas entre os 116 pontos amostrais disponíveis as PCDs que serviram como pontos amostrais de referência, representando os dados observados.

Foi criado um algoritmo que possibilitou escolher entre a rede de estações, os pontos amostrais que continham no mínimo 4 estações medidoras próximas a uma distância de até 100 km do ponto (Figura 3.9). A delimitação de distâncias menores do 100 km não foi possível devido a disponibilidade das PCDs resultantes do processo de qualificação e, pelo mesmo motivo, não foi levada em consideração a direção das PCDs circundantes ao ponto amostral. A distribuição espacial dos

pontos amostrais resultantes desse processo é apresentada na Figura 3.10 e a relação desses pontos é listada na Tabela 3.7.

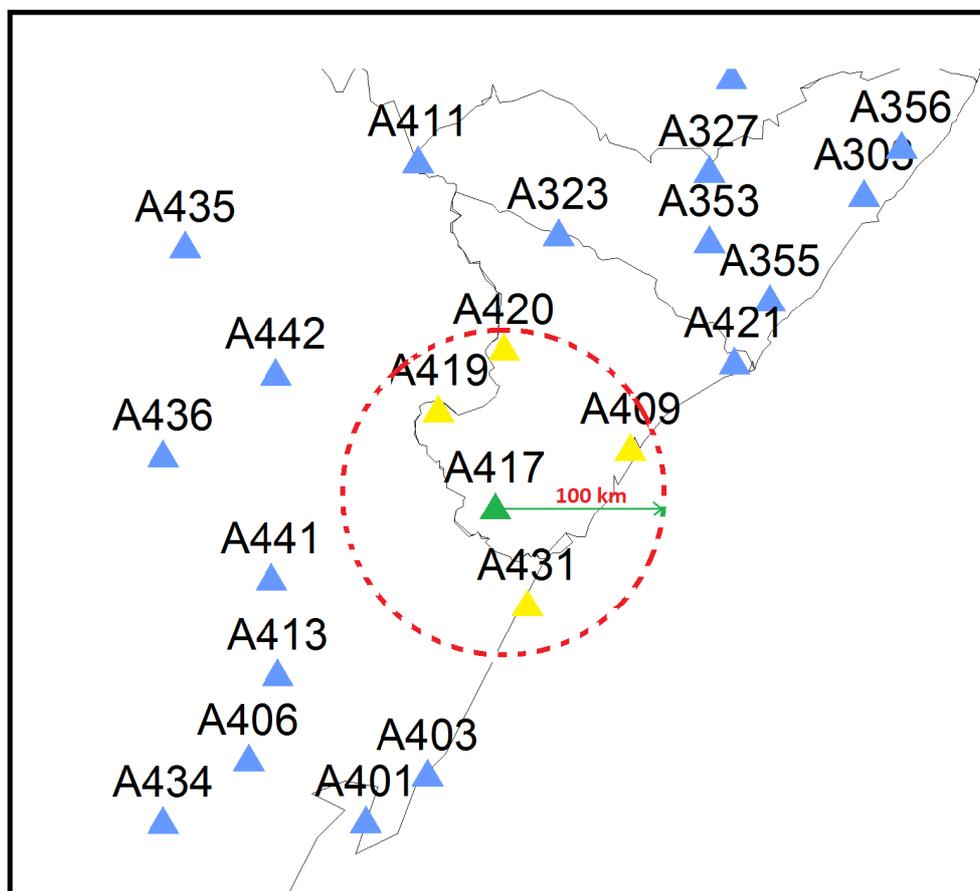


Figura 3.9 – Representação esquemática do processo de escolha dos pontos amostrais. Na figura, o triângulo em verde – ponto da PCD A417 – é escolhido como ponto amostral pois atende ao quesito de conter num raio de até 100 km, no mínimo 4 PCDs circundantes (triângulos em amarelo).

Tabela 3.7 – Pontos amostrais resultantes (PCDs de referência) e suas posições geográficas.

N	ID	LON	LAT
1	A303	-35,78	-9,55
2	A310	-35,69	-6,95
3	A323	-37,43	-9,75
4	A327	-36,62	-9,4
5	A328	-35,78	-7,85
6	A340	-37,82	-5,6
7	A341	-35,99	-8,25
8	A348	-36,29	-7,5
9	A353	-36,62	-9,8
10	A355	-36,29	-10,1
11	A357	-35,57	-8,65
12	A417	-37,79	-11,25
13	A420	-37,73	-10,4
14	A436	-39,62	-11

3.8 Krigagem

A metodologia utilizada na interpolação foi estruturada com base nas médias anuais calculadas através das médias dos totais diários. As interpolações foram realizadas em 4 etapas, percorrendo os anos de 2008, 2009, 2010 e 2011.

O artifício de comparação do desempenho das interpolações por KO foi baseado no método de validação cruzada descrito no Capítulo 3.5. Foi elaborado um algoritmo para encontrar e extrair de cada série anual de dados os valores de irradiação medidos referentes aos pontos amostrais. Assim, para cada um dos 14 pontos ($n = 1, \dots, 14$) amostrais foi realizada a interpolação – se $n = 1$, os dados referentes ao ponto amostral 1 são retirados da série e é realizada a interpolação sem esses dados. Esse procedimento foi realizado para todos os 14 pontos interpoladores para cada uma das séries anuais, resultando em: $14 \times 4 = 56$ ajustes dos semivariogramas. Em seguida, foi feita a comparação quantitativa dos valores estimados para os pontos amostrais em relação aos valores medidos. A estimativa é tanto melhor quanto menor for o desvio em relação ao dado observado.

O ajuste dos semivariogramas experimentais ao modelo esférico foi o primeiro passo do processo das interpolações. Esse procedimento foi realizado variando a

cada iteração ($n = 1, \dots, 14$) os parâmetros de ajuste do modelo teórico: a_0 (alcance) e C_0 (patamar/contribuição).

Devido ao número extenso de semivariogramas produzidos a cada interpolação para cada ano para cada ponto amostral, fica inviável a apresentação de todos os 56 semivariogramas nesse trabalho.

Como já foi mencionado no Capítulo 3.6 (Etapa I), a série de dados que representa o ano de 2010 apresentou o menor número de falhas entre as séries. Buscando adequar a apresentação dos resultados optou-se por apresentar os semivariogramas para o ano de 2010 em 5 pontos – dos 14 pontos amostrais.

A escolha dos pontos amostrais foi feita como segue:

- Foi realizada uma análise visual da distribuição dos 14 pontos amostrais (Figura 3.10) diante da representatividade espacial e foram selecionadas as PCDs: A303, A310, A340, A417 e A436.

Na Figura 3.11 são apresentados os pontos amostrais de referência distribuídos espacialmente pela região NEB. Através da análise visual da distribuição espacial dos pontos, foi possível escolher 5 pontos amostrais de forma que esses pontos possam representar os demais de maneira “homogênea”, ou seja, o objetivo foi escolher os pontos mais afastados espacialmente de modo a compreender uma região que contenha os demais pontos amostrais de referência.

As Figuras apresentadas a seguir mostram os semivariogramas ajustados para os pontos de referência para o ano de 2010. A distribuição dos pontos nos semivariogramas experimentais representada pela curva na cor preta representa a distribuição espacial em unidades de posição geográfica – em graus ($^\circ$) de latitude e longitude – unidade do eixo ‘ x ’ – dos pares de dados utilizados a cada atribuição dos pesos dos estimadores em função da variância (eixo ‘ y ’). Esse aspecto é o mesmo em todos os semivariogramas pois depende dos dados de entrada e não do

ajuste. O que muda, e pode (e deve) ser ajustado é o modelo variográfico esférico – curvas na cor azul – Figuras 3.12, 3.13, 3.14 e 3.15.

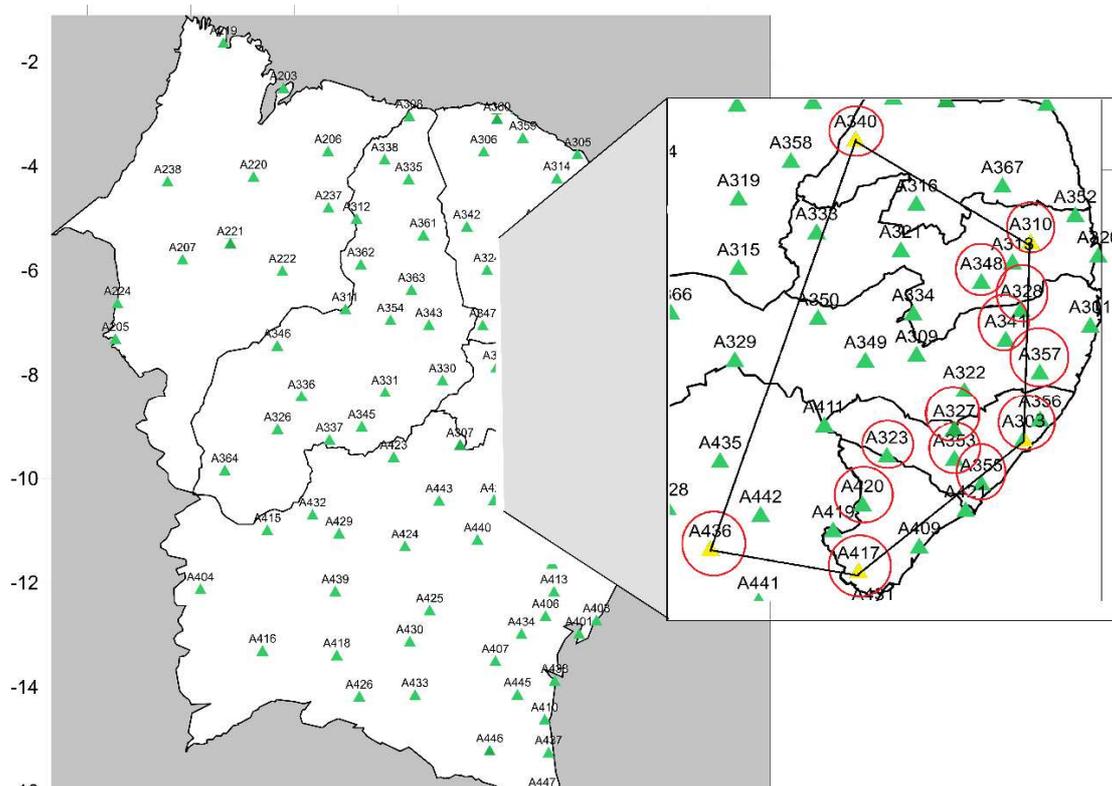


Figura 3.11 - Distribuição espacial dos pontos amostrais disponíveis. Em destaque, estão as PCDs utilizadas como pontos amostrais para comparação. A figura geométrica em destaque liga os pontos das 5 PCDs usadas para a demonstração visual dos semivariogramas ajustados.

Considerando o diâmetro da Terra igual a 12.745 km aproximadamente, sua circunferência possui um perímetro de 40.030,17 km. Dividindo esse valor por 360, obtemos a equivalência de 1° de latitude ou longitude em km: 1° = 111,20 km.

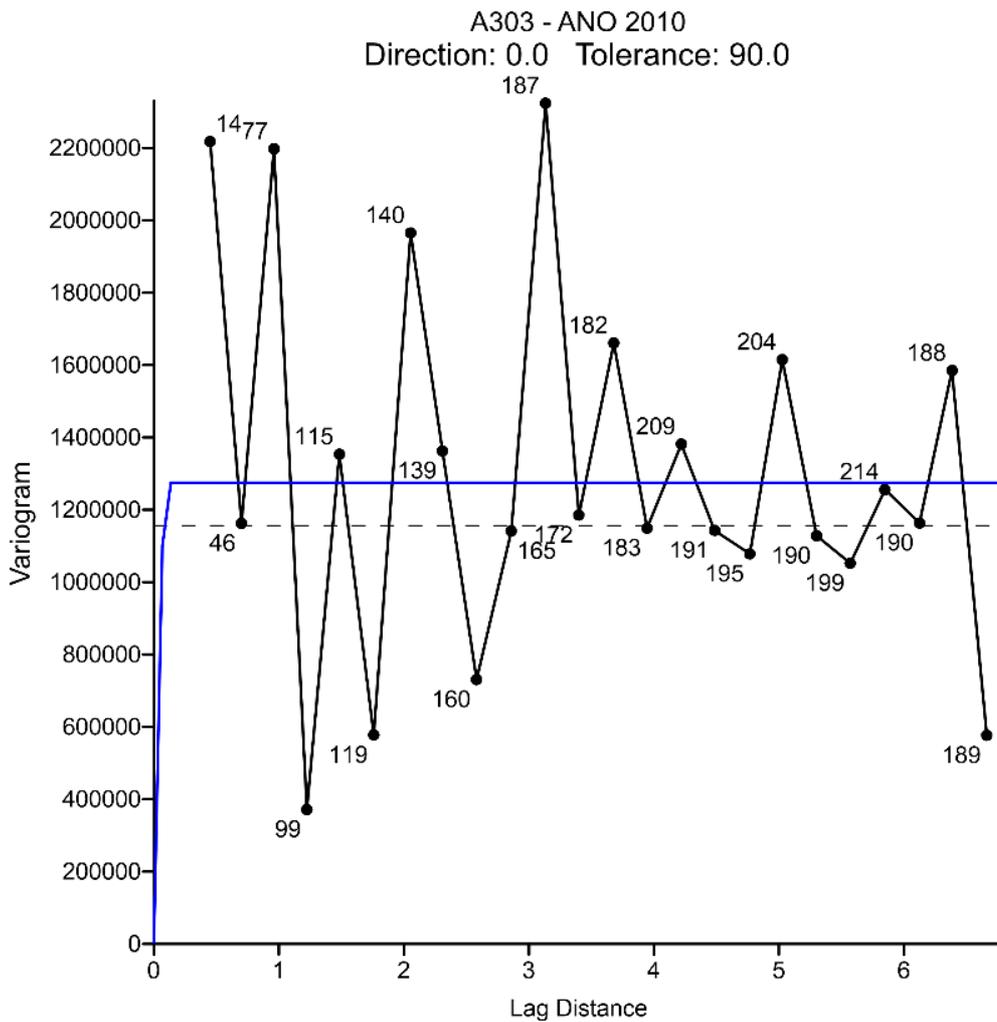


Figura 3.12 – Ajuste do semivariograma experimental para o ponto amostral: PCD A303 ano de 2010. A curva de cor preta representa o semivariograma experimental dos dados de entrada. A curva azul representa o ajuste do modelo teórico (esférico) aos dados. Uma unidade do eixo x corresponde a 111,25 km.

Na Figura 3.12 é apresentado o ajuste para o ponto A303. É possível observar que o alcance (a_0) ajustado para esse ponto é pequeno, seu valor foi: $a_0 = 0.1^\circ$ o que corresponde a aproximadamente 11,12 km. Esse ponto é localizado próximo à costa (ver Figura 3.10) e não há estações medidoras próximas em todo o seu lado leste. A presença da PCD A356 (a 36 km de distância) pode explicar o alto grau de dependência desse ponto em relação a pontos vizinhos.

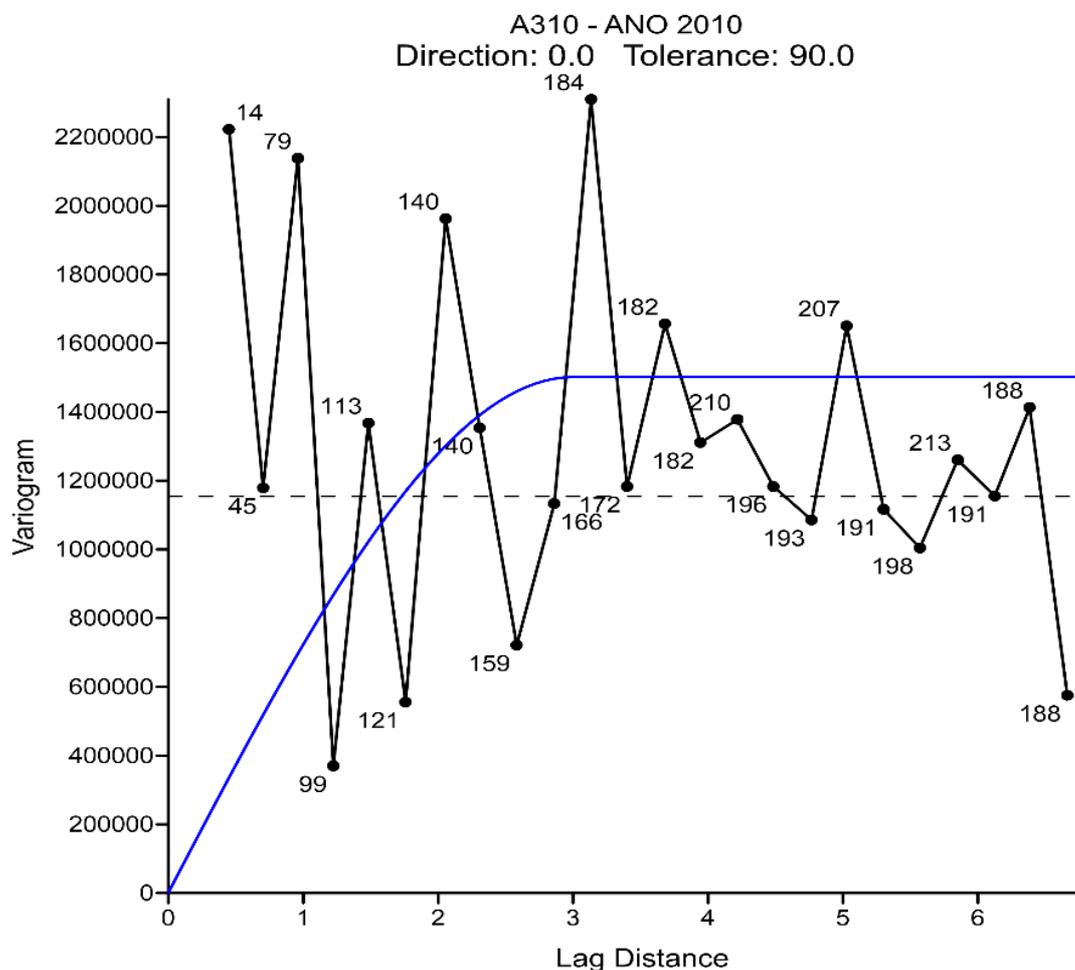


Figura 3.13 – Ajuste do semivariograma experimental para o ponto amostral: PCD A310 ano de 2010. A curva de cor preta representa o semivariograma experimental dos dados de entrada. A curva azul representa o ajuste do modelo teórico (esférico) aos dados. Uma unidade do eixo x corresponde a 111,25 km.

A Figura 3.13 apresenta o ajuste do semivariograma para o ponto A310. O valor de a_0 foi de 3° , o que corresponde a 333 km aproximadamente. O número de pares de pontos amostrais com estações medidoras próximas até essa distância contribuiu para explicar a dependência espacial nesse ponto amostral.

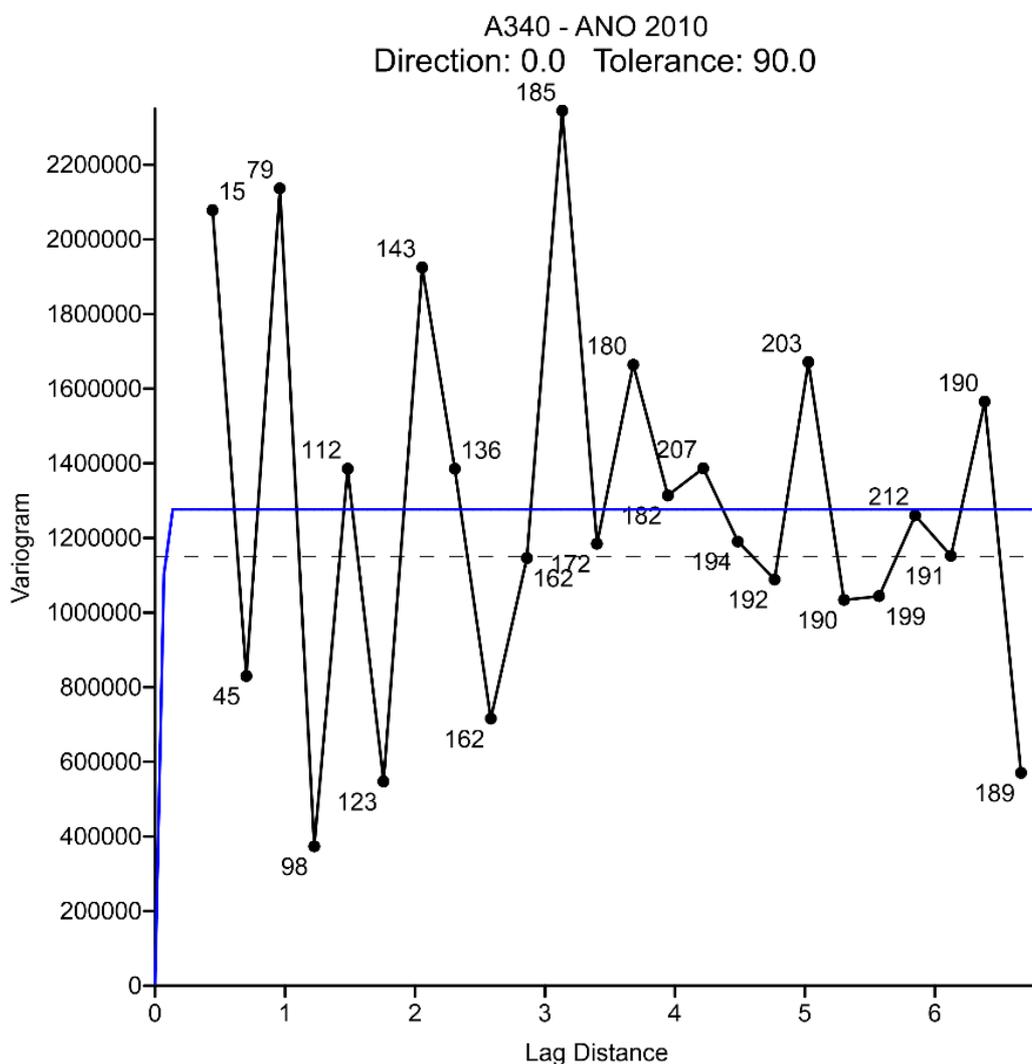


Figura 3.14 – Ajuste do semivariograma experimental para o ponto amostral: PCD A340 ano 2010. A curva de cor preta representa o semivariograma experimental dos dados de entrada. A curva azul representa o ajuste do modelo teórico (esférico) aos dados. Uma unidade do eixo x corresponde a 111,25 km.

O ajuste apresentado na Figura 3.14 representa o ponto amostral A340. Nesse ponto foi observado um alto grau de dependência com a distância a pontos vizinhos para explicar sua variância ($a0 = 0.1^\circ = 11,11 \text{ km}$). Nesse caso, também é observada a ausência de pontos amostrais vizinhos do seu lado leste (ver figura 3.10).

O ponto amostral A340 apresentou 4 pontos com estações próximas a menos de 100 km em sua vizinhança, são elas as PCDs A318, A332, A358 e A339 – essa configuração pode ser observada na Figura 3.10. Esse resultado indica que apesar de esse ponto atender o critério para ser utilizado como ponto amostral de comparação, seus resultados podem atribuir viés as estimativas pois não há estações próximas numa distância inferior a 79 km – esse fato pode ser comprovado no Capítulo 4, onde são apresentados os resultados das interpolações e a produção dos mapas.

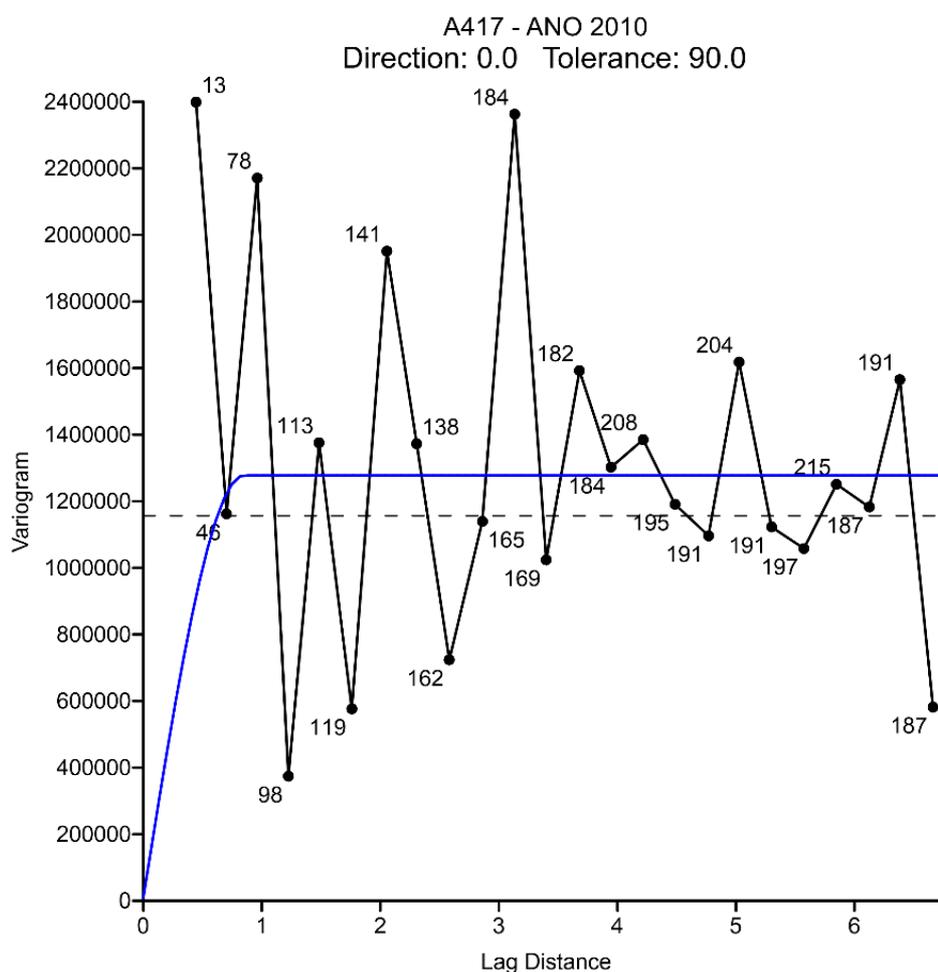


Figura 3.15 – Ajuste do semivariograma experimental para o ponto amostral: PCD A417 ano 2010. A curva de cor preta representa o semivariograma experimental dos dados de entrada. A curva azul representa o ajuste do modelo teórico (esférico) aos dados. Uma unidade do eixo x corresponde a 111,25 km.

A Figura 3.15 apresenta o ajuste para o ponto amostral A417. Esse ponto apresentou alcance $a_0 = 0.85^\circ = 94 \text{ km}$. Como esse ponto amostral apresenta PCDs em sua vizinhança com distâncias inferiores ao valor de a_0 , pôde-se ajustar o modelo esférico ao semivariograma experimental de modo a utilizar ao menos 46 pares de pontos estimadores nas estimativas que explicaram a variância e a dependência espacial nesse ponto.

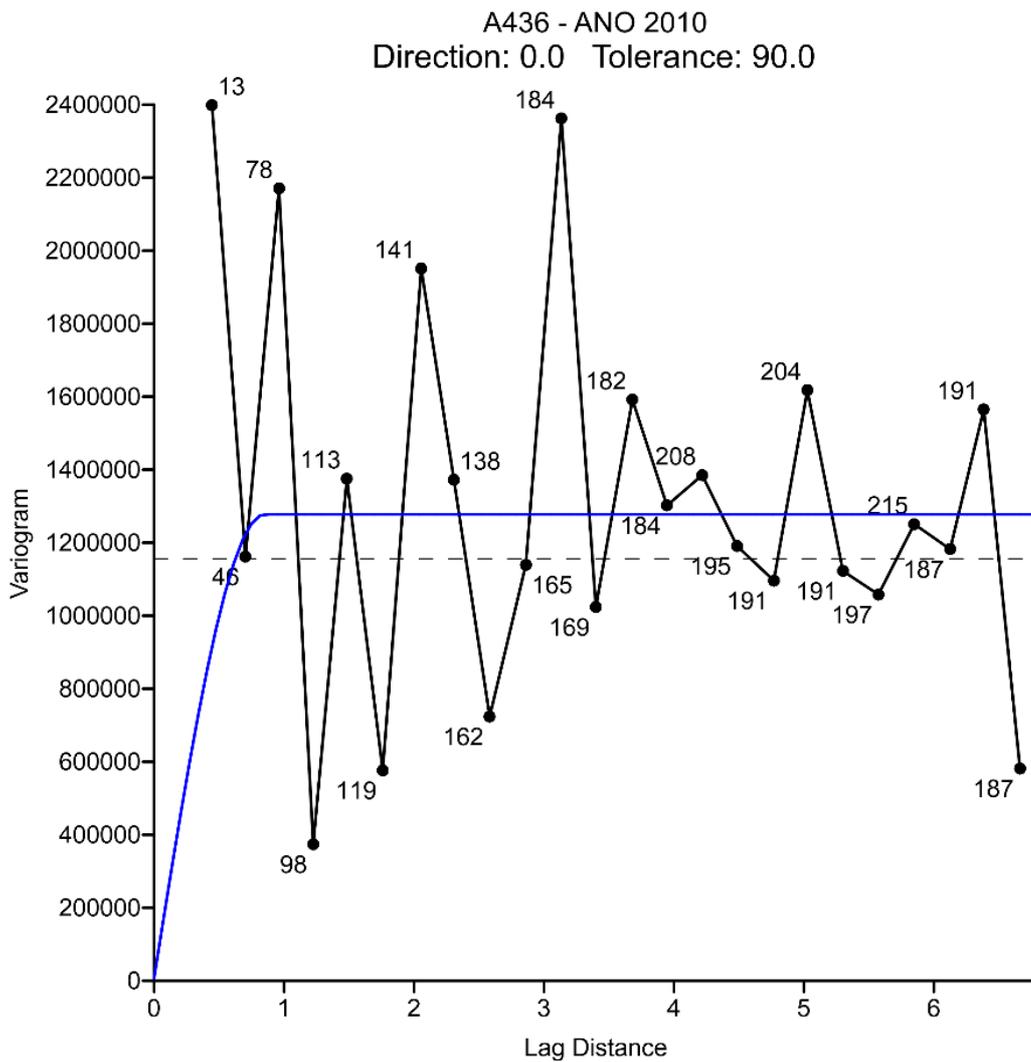


Figura 3.16 – Ajuste do semivariograma experimental para o ponto amostral: PCD A436 ano 2010. A curva de cor preta representa o semivariograma experimental dos dados de entrada. A curva azul representa o ajuste do modelo teórico (esférico) aos dados. Uma unidade do eixo x corresponde a 111,25 km.

A Figura 3.16 apresenta o ajuste para o ponto amostral A436. Assim como o ponto amostral A417 discutido anteriormente, esse ponto apresentou valor de $a_0 = 0.85^\circ = 94 \text{ km}$. Fato que lhe permitiu ajuste muito próximo ao discutido acima.

4. RESULTADOS E DISCUSSÕES

Neste capítulo, são exibidos os resultados obtidos através das intercomparações dos métodos de estimativa utilizados (krigagem e os resultados do modelo BRASIL-SR para o ano de 2010) em relação aos dados coletados pela rede de PCDs operadas pelo INMET – os pontos amostrais. Primeiramente é discutido o desempenho da krigagem em relação aos dados amostrais no período estabelecido pela metodologia de tratamento dos dados - 2008 a 2011. Em seguida, é discutido o desempenho das estimativas do modelo BRASIL-SR em relação aos dados amostrais para o ano de 2010. Posteriormente, apresenta-se a comparação quantitativa das estimativas dois métodos (krigagem e modelo numérico) para o ano de 2010.

O desempenho do método de estimativa foi avaliado utilizando um conjunto de índices estatísticos aliados ao método de validação cruzada. Assim, para minimizar os erros na interpretação dos resultados, como mostrado por Weber et al. (1982), foram calculados índices estatísticos que avaliam os valores estimados em relação aos dados medidos. Portanto, os resultados apresentados são avaliados pelo coeficiente de determinação (r^2) – obtido no ajuste dos pontos ao plotar os dados observados contra os dados estimados –, o Erro Quadrático Médio Absoluto (EQMA) e a variabilidade espacial da irradiância através da visualização dos mapas de irradiação solar.

4.1 Intercomparação: krigagem x dados observados

A intercomparação foi realizada avaliando as médias anuais para os anos de 2008, 2009, 2010 e 2011; A avaliação estatística referente às médias anuais é apresentada na Tabela 4.1 e o ajuste dos pontos para obtenção do coeficiente de determinação é apresentado na Figura 4.1

Tabela 4.1 – Índices estatísticos que avaliam o desempenho das estimativas em relação aos 14 pontos de dados amostrais de referência para as séries de dados de 2008, 2009, 2010 e 2011.

Índices	ANO-2008	ANO-2009	ANO-2010	ANO-2011
r^2	0,40	0,46	0,72	0,59
<i>EQMA (%)</i>	5,89	5,83	5,02	5,66

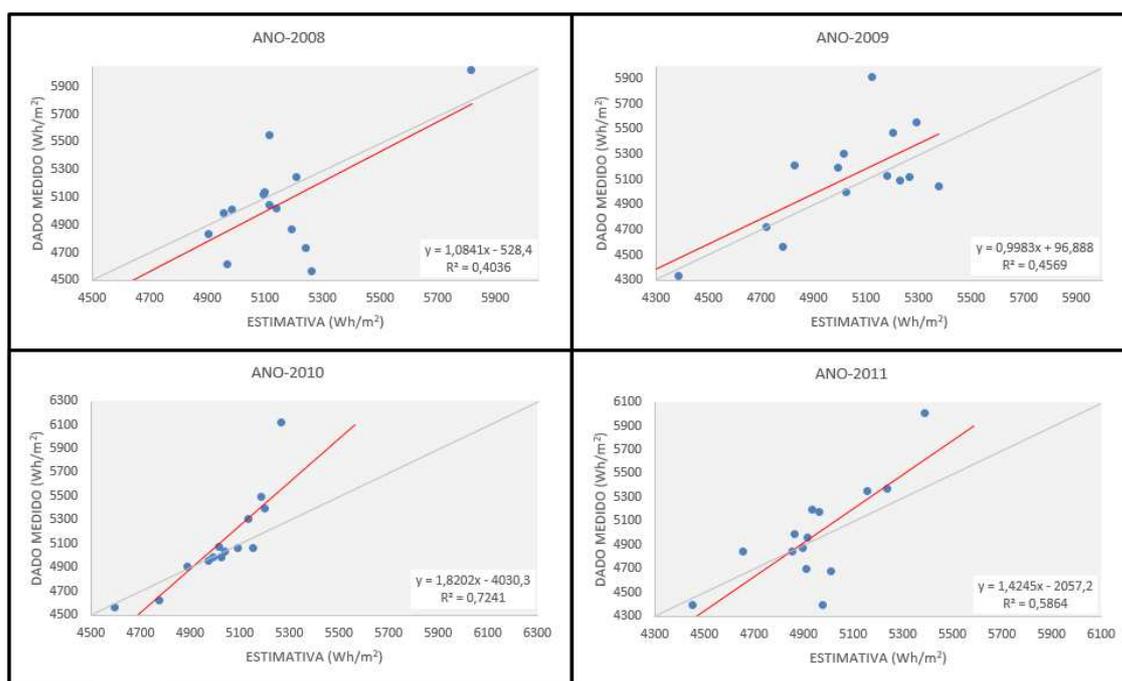


Figura 4.1 – Gráfico de dispersão e ajuste linear dos valores estimados pelos valores medidos para os anos de 2008, 2009, 2010 e 2011. A linha diagonal delimita a região onde os valores foram superestimados (inclinação da reta de ajuste acima da linha diagonal) ou subestimados (inclinação abaixo da linha diagonal).

A Tabela 4.1, mostra os valores do coeficiente de determinação r^2 . Pôde-se ver que os valores estimados puderam explicar a variabilidade dos dados medidos em 40 % na série de 2008, 46 % na série de 2009, 72 % na série de 2010 e 59 % na série de 2011. Em relação ao *EQMA (%)*, considerando todas as séries anuais seu valor calculado oscilou entre 5 e 6 % mostrando que, em relação ao valor médio dos

pontos amostrais para cada ano, as estimativas produzidas sofreram desvios pouco significativos em relação aos dados amostrais.

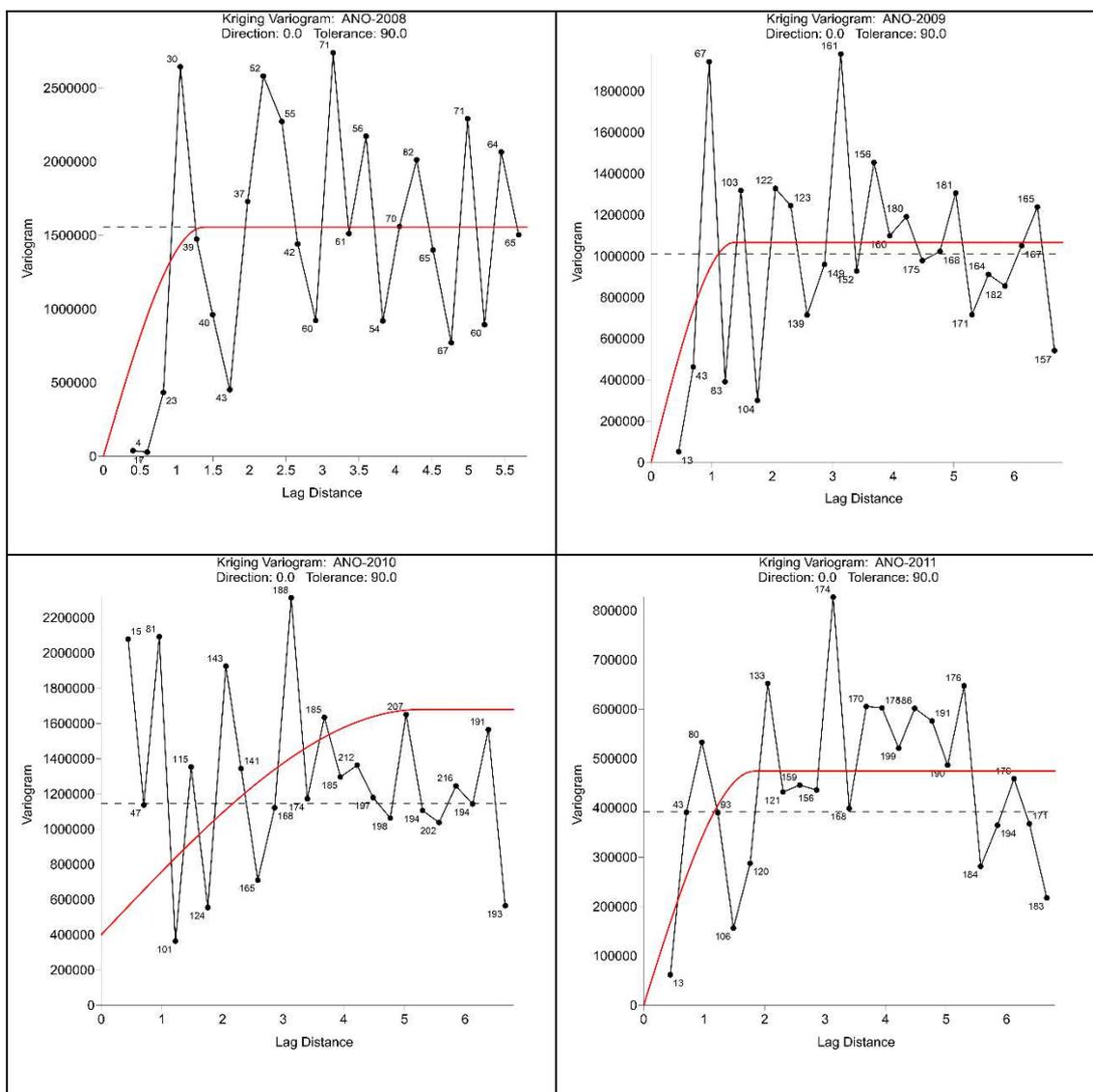


Figura 4.2 - Semivariogramas produzidos com os dados das séries anuais de 2008, 2009, 2010 e 2011. Uma unidade do eixo x corresponde a 111,25 km.

Esses resultados indicam que as estimativas produzidas por krigagem superestimaram os valores medidos nos anos de 2008 e 2009 e, para anos com maior densidade de pontos amostrais (estações medidoras com dados medidos

validados) como no caso dos anos de 2010 e 2011, os saldos de radiação foram superestimados em escalas de até 5000 Wh/m² aproximadamente – acima desses valores os dados foram subestimados.

A variabilidade espacial da irradiação solar em cada ano pode ser observada através dos semivariogramas e mapas das médias anuais dos totais diários apresentados nas figuras adiante. Os semivariogramas apresentados na Figura 4.2 foram ajustados durante o processo de krigagem e contribuíram para a obtenção das estimativas e produção dos mapas.

A análise da variabilidade espacial pode ser vista nos mapas de irradiação solar nas Figuras 4.3, 4.4, 4.5 e 4.6. Pode-se observar altos níveis de irradiância em todo o território do NEB como era de se esperar. A região central do NEB apresentou a maior homogeneidade em todas as séries proporcionando valores entre 5000 e 7000 Wh/m² por ano. Os níveis de irradiância mais baixos foram observados nas regiões de fronteira entre a região NEB e à costa oceânica e as divisas com a região norte e sudeste – vale ressaltar que nessas regiões de fronteira não há pontos amostrais além do limite territorial da região Nordeste.

É possível observar nas Figuras 4.3, 4.4 e 4.6, próximo a região central, uma diminuição relevante dos níveis de irradiância representada por um gradiente acentuado – variando de 1000 a 5000 Wh/m². O aspecto desses mapas indica que algum ponto amostral no estado da Bahia pode ter influenciado negativamente as estimativas recebendo um peso muito maior diante dos pontos vizinhos. Como todos os dados fornecidos pelas PCDs usadas como pontos amostrais passaram pelo processo de qualificação, algumas das PCDs nessa região pode ter apresentado erro sistemático significativo que não foi identificado no processo de análise e qualificação dos dados.

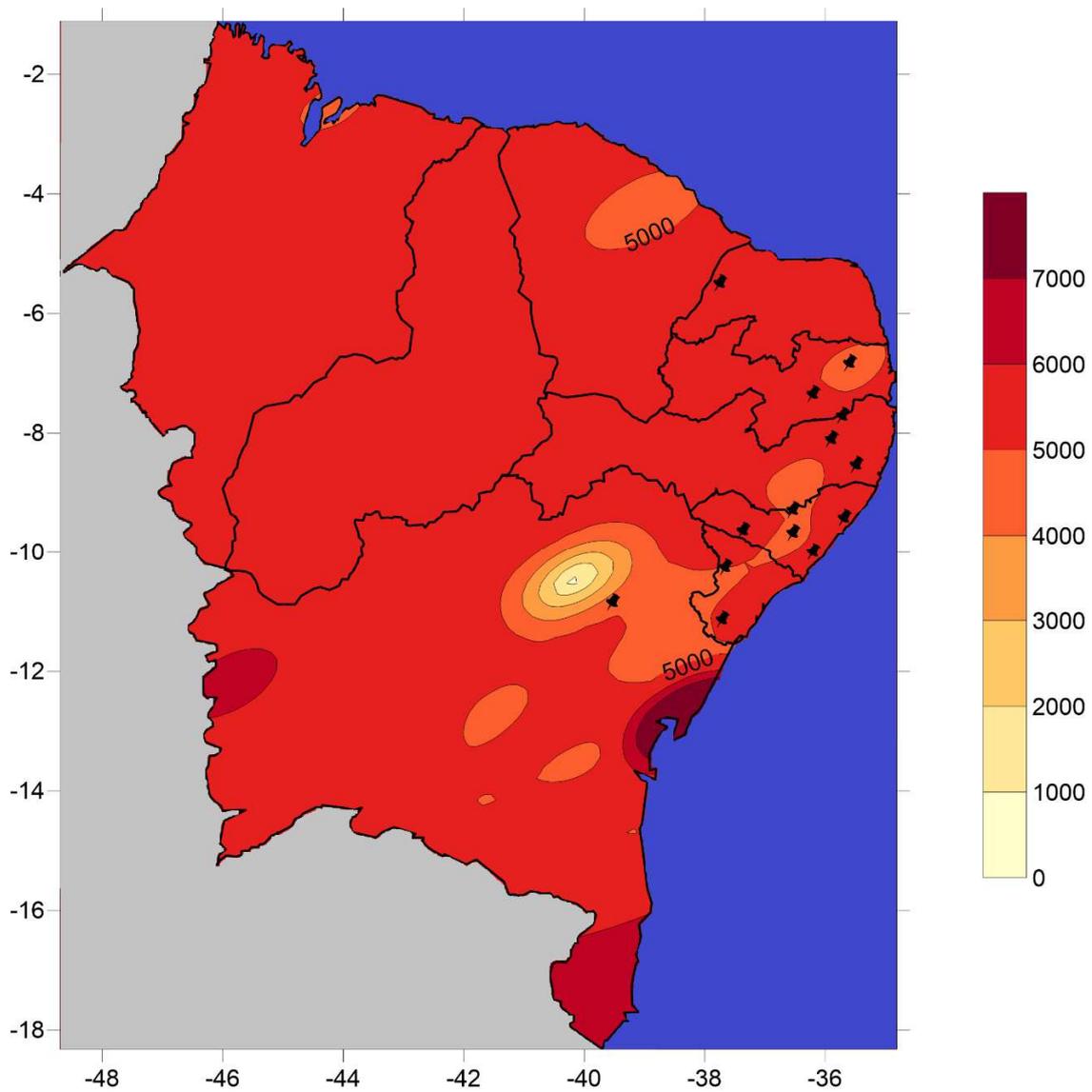


Figura 4.3 – Variabilidade espacial da irradiação global média anual para a série de 2008 - valores em Wh/m² obtidos por krigagem ordinária a partir dos 116 pontos de dados amostrais. Os pontos destacados em preto representam a disposição espacial dos pontos amostrais de referência.

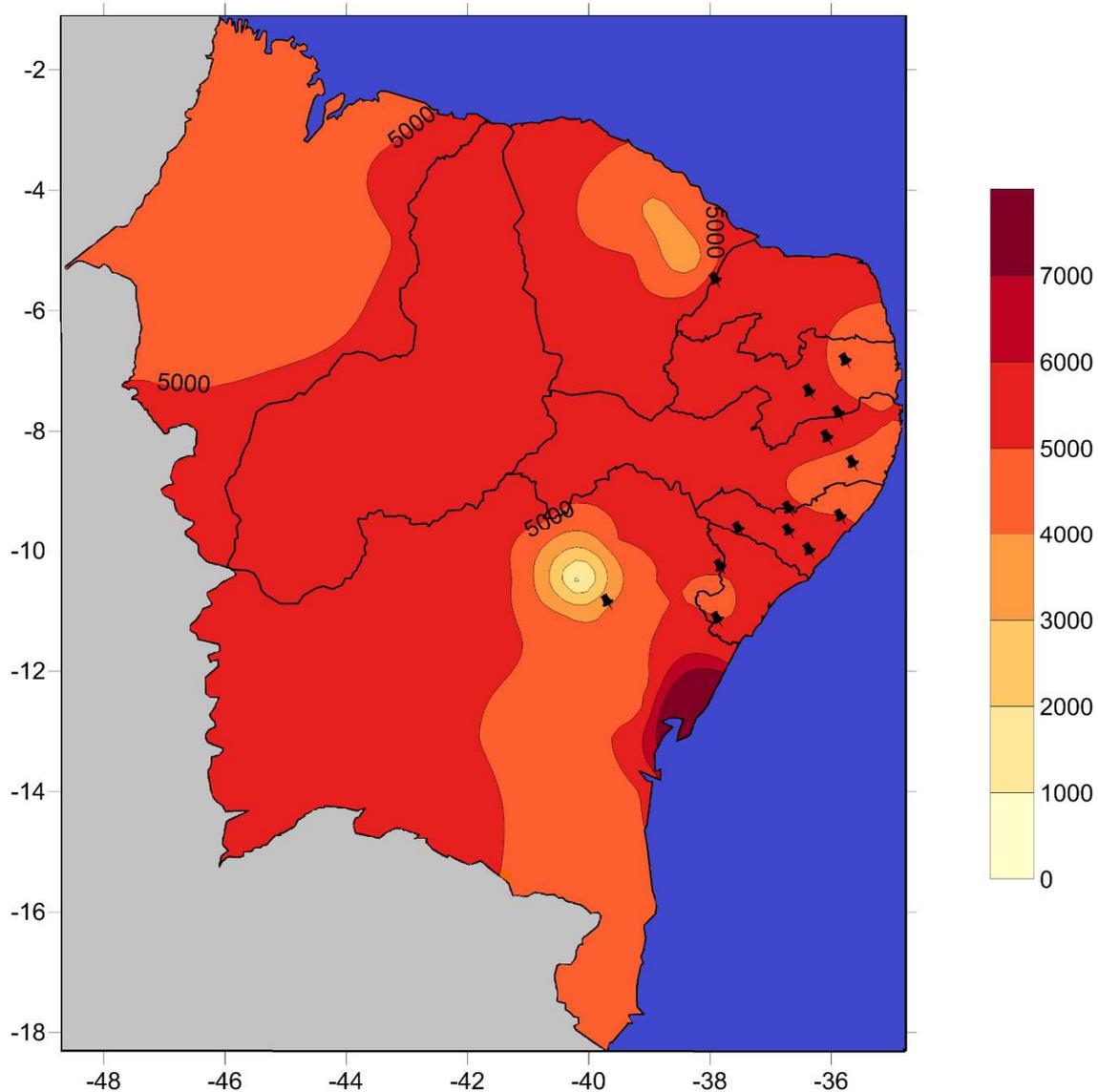


Figura 4.4 – Variabilidade espacial da irradiação global média anual para a série de 2009 - valores em Wh/m² obtidos por krigagem ordinária a partir dos 116 pontos de dados amostrais. Os pontos destacados em preto representam a disposição espacial dos pontos amostrais de referência.

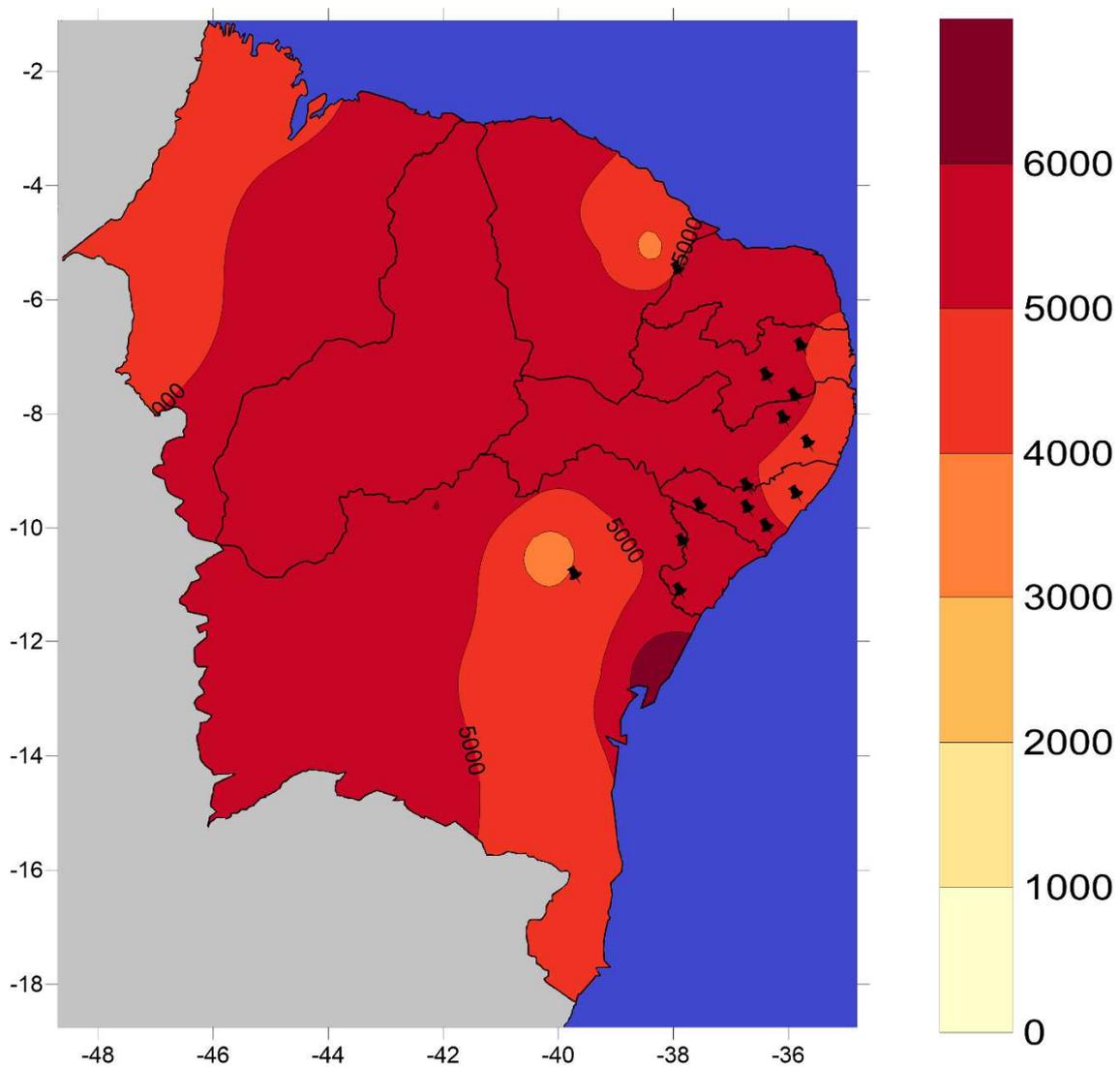


Figura 4.5 – Variabilidade espacial da irradiação global média anual para a série de 2010 - valores em Wh/m² obtidos por krigagem ordinária a partir dos 116 pontos de dados amostrais. Os pontos destacados em preto representam a disposição espacial dos pontos amostrais de referência.

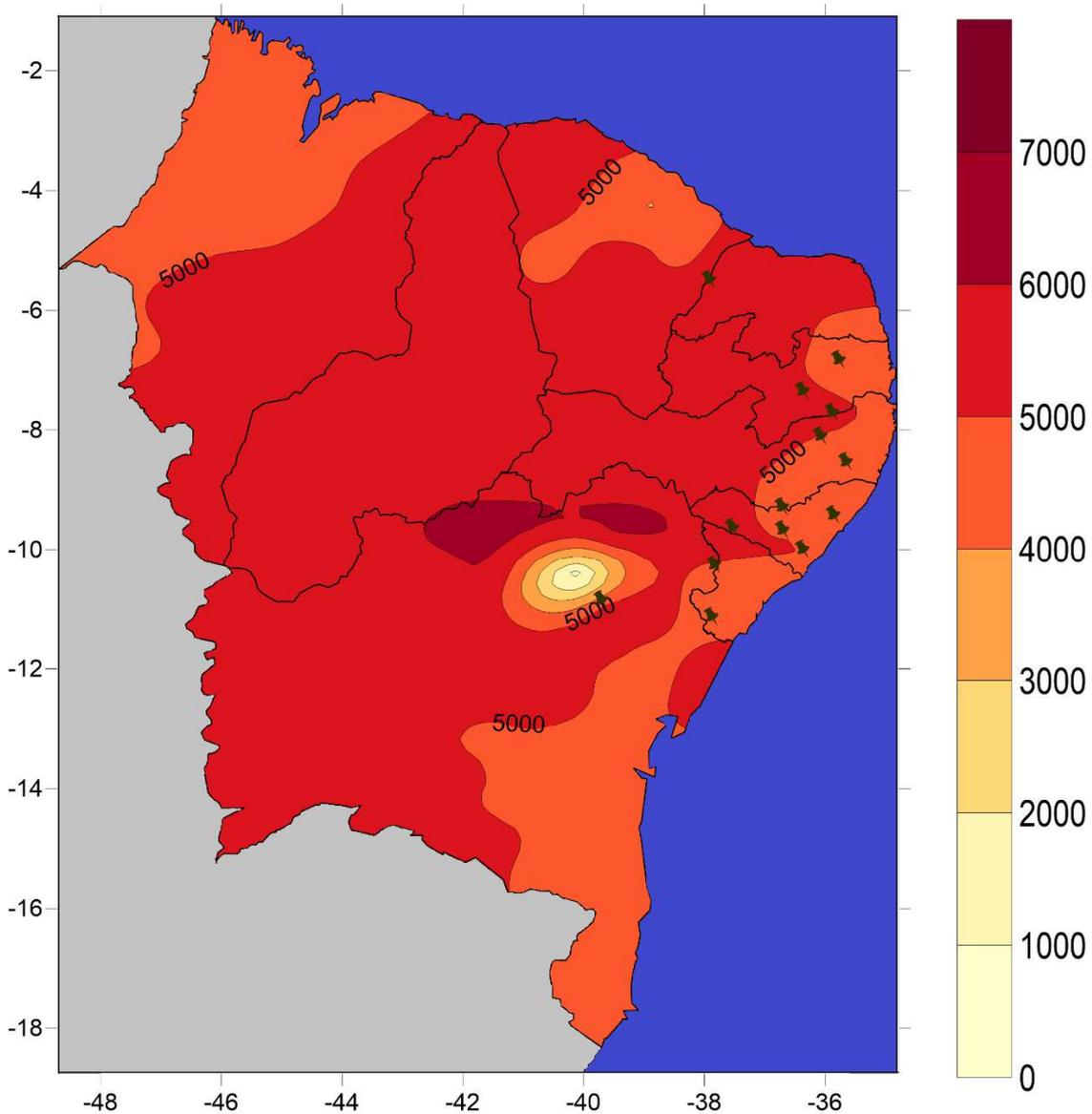


Figura 4.6 – Variabilidade espacial da irradiação global média anual para a série de 2011 - valores em Wh/m² obtidos por krigagem ordinária a partir dos 116 pontos de dados amostrais. Os pontos destacados em preto representam a disposição espacial dos pontos amostrais de referência.

4.2 Intercomparação: estimativas modelo BRASIL-SR x dados observados

Nessa etapa, foi feita a comparação das estimativas produzidas pelo modelo numérico de transferência radiativa BRASIL-SR em relação aos dados observados. A série de dados que contém as estimativas do modelo BRASIL-SR é o ano de 2010 – definido aqui como o período de análise desse subcapítulo.

A análise dos índices estatísticos mostra que as estimativas do modelo BRASIL-SR pôde explicar cerca de 74 % da variabilidade dos dados observados. Esses resultados são apresentados na Tabela 4.2

Tabela 4.2 – Estimativas (Wh/m^2) do modelo numérico BRASIL-SR em relação aos dados medidos para o ano de 2010 nos pontos de referência. Abaixo, são apresentados os valores do coeficiente de determinação (r^2) e do Erro Quadrático Médio Absoluto (EQMA) em porcentagem.

ID-PCD	ANO-2010	BRASIL-SR
A303	5069,64	5109,58
A310	4628,05	5154,92
A323	5494,44	5318,25
A327	5038,53	5196,42
A328	5304,90	5117,83
A340	6115,57	5689,33
A341	5067,97	5076,67
A348	5394,46	5376,58
A353	5063,02	5228,17
A355	4988,30	5230,67
A357	4562,25	5011,42
A417	4954,30	5033,00
A420	4986,24	5161,67
A436	4909,59	4999,83
r^2		0,74
EQMA (%)		4,92

Os resíduos, assim como o ajuste para obter o coeficiente de determinação são apresentados nas Figuras 4.7 e 4.8. Na Figura 4.7 podem-se ver as PCDs que superestimaram e subestimaram os valores. O módulo dos maiores valores dos resíduos é observado entre 400 Wh/m² (PCD de A340) e 500 Wh/m² (PCD de ID A310), o que representa menos de 10 % dos valores máximos medidos e estimados.

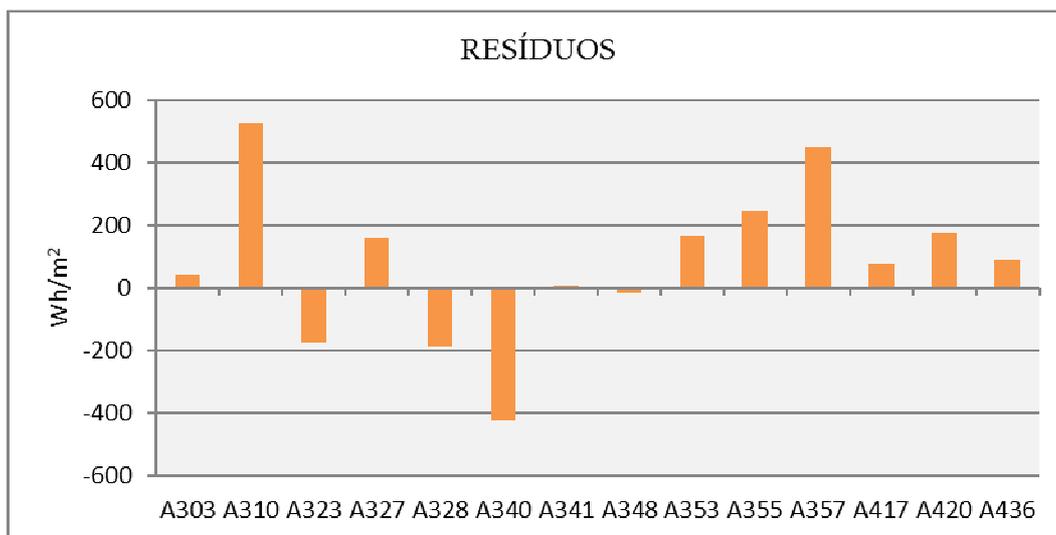


Figura 4.7 – Resíduos: os valores observados nos pontos amostrais foram subtraídos dos valores estimados pelo modelo BRASIL-SR.

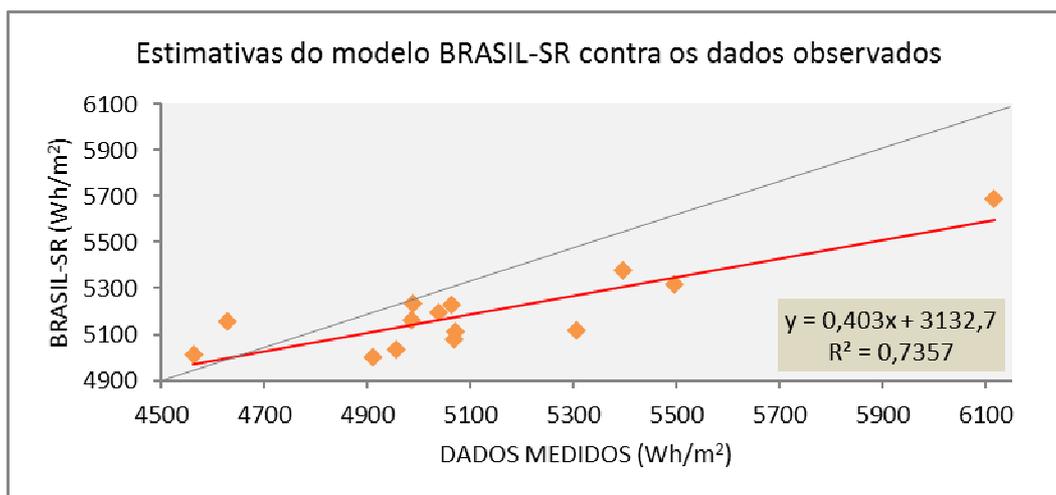


Figura 4.8 – Representação de dispersão dos valores estimados (eixo y) contra os valores observados (eixo x).

O mapa apresentado na Figura 4.9 mostra a variabilidade espacial da irradiância em superfície no NEB para o ano de 2010, estimativas feitas pelo modelo BRASIL-SR. É possível observar certa homogeneidade dos valores entre 5000 e 6000 Wh/m².

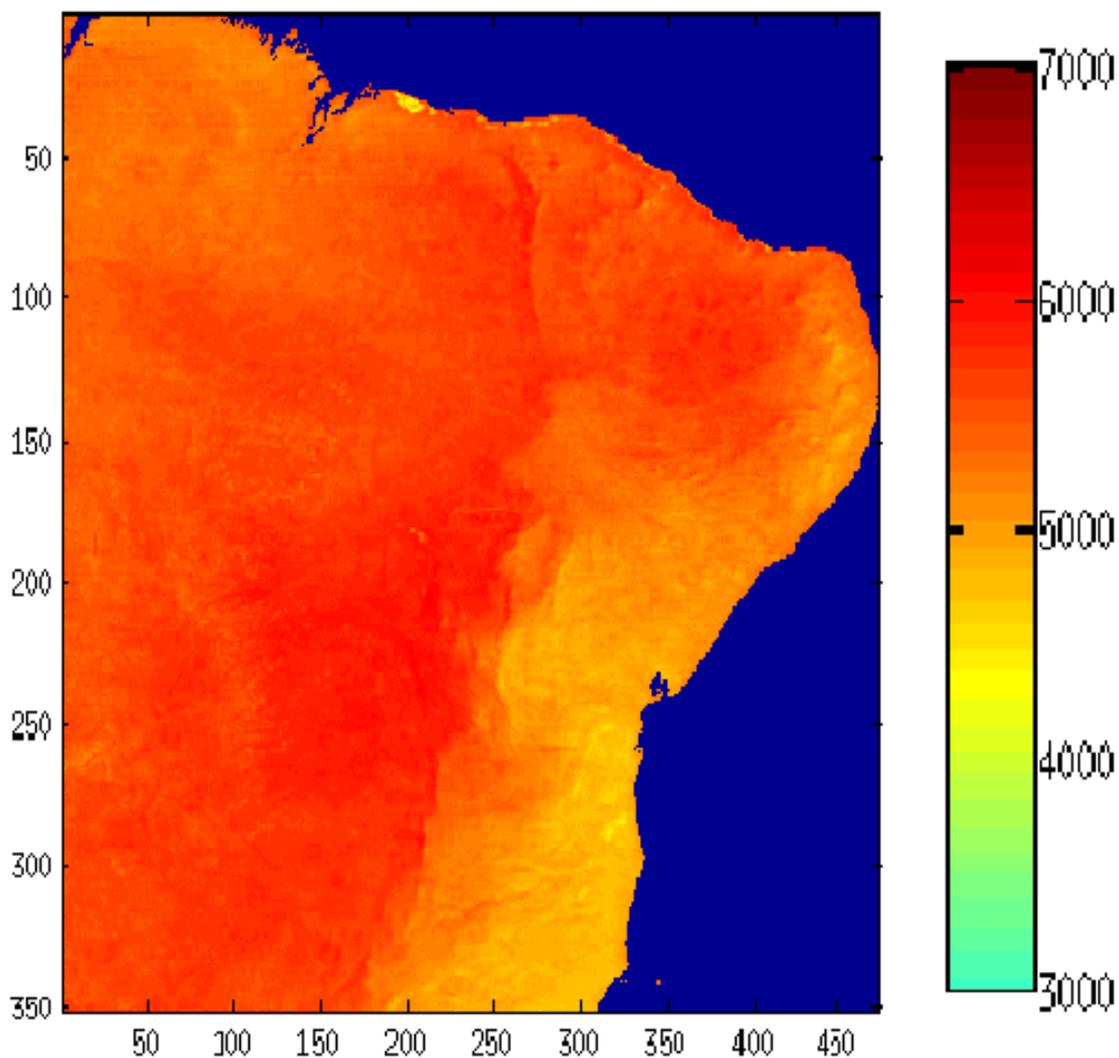


Figura 4.9 – Mapa da irradiação global média anual para o ano de 2010. Estimativas realizadas pelo modelo BRASIL-SR.

4.3 Intercomparação: krigagem x modelo BRASIL-SR

Essa análise confronta as estimativas realizadas através da interpolação por krigagem e as estimativas realizadas pelo modelo BRASIL-SR para o ano de 2010.

A Tabela 4.3 apresenta os resultados quantitativos das estimativas de irradiância média anual por krigagem ordinária e pelo modelo BRASIL-SR para o ano de 2010, assim como os desvios. A Figura 4.10 apresenta o gráfico de dispersão ajustado por regressão linear e o coeficiente de determinação. A Figura 4.11 mostra o comportamento dos desvios.

O valor de r^2 mostra que o desempenho dos métodos foi considerado apropriado pois ambos foram capazes de explicar cerca de 74 % da variabilidade observada nas estimativas e apresentaram uma forte correlação.

Tabela 4.3 –Valores de irradiação média anual (Wh/m²) estimados por krigagem ordinária e pelo modelo numérico BRASIL-SR para os pontos amostrais de referência o ano de 2010. Abaixo, são mostrados os valores dos índices estatísticos.

ID	Krigagem (K)	BRASIL-SR (B)	DESVIOS (K-B)
A303	5069,6394	5109,58	-39,9406
A310	4628,0509	5154,92	-526,8691
A323	5494,4376	5318,25	176,1876
A327	5038,5299	5196,42	-157,8901
A328	5304,9001	5117,83	187,0701
A340	6115,5731	5689,33	426,2431
A341	5067,9738	5076,67	-8,6962
A348	5394,4588	5376,58	17,8788
A353	5063,0204	5228,17	-165,1496
A355	4988,2954	5230,67	-242,3746
A357	4562,2451	5011,42	-449,1749
A417	4954,2981	5033,00	-78,7019
A420	4986,2374	5161,67	-175,4326
A436	4909,5907	4999,83	-90,2393
$r^2 = 0,74$ (correlação = 0,86)		EQMA (%) = 4,92	

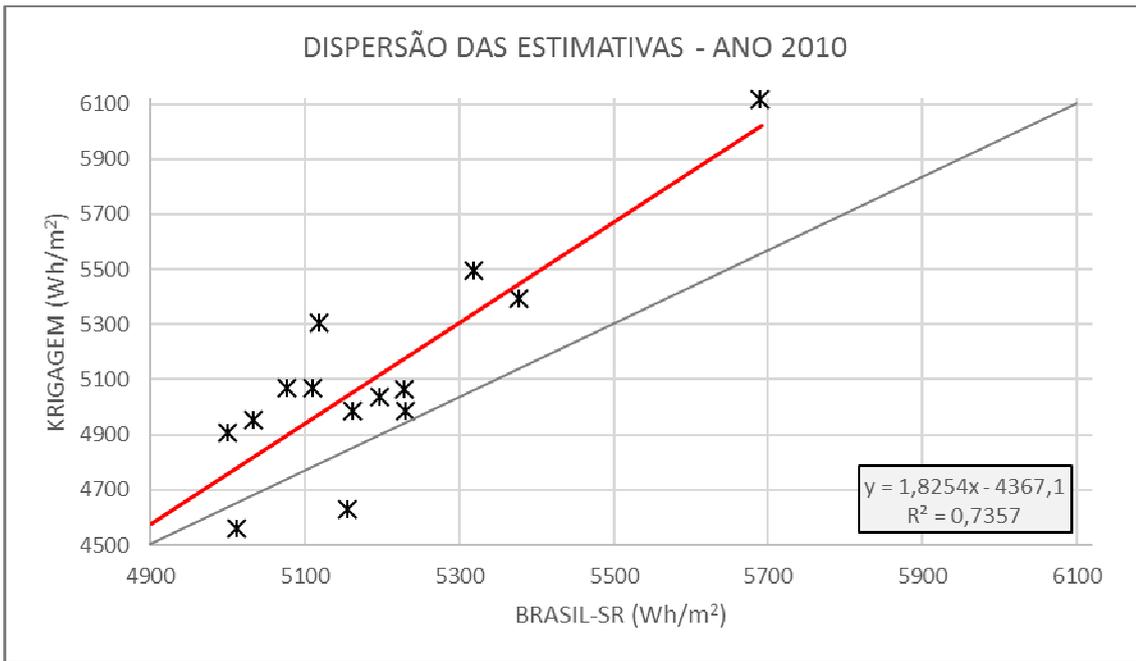


Figura 4.10 – Gráfico de dispersão e ajuste por regressão linear dos valores estimados por krigagem e pelo modelo BRASIL-SR.

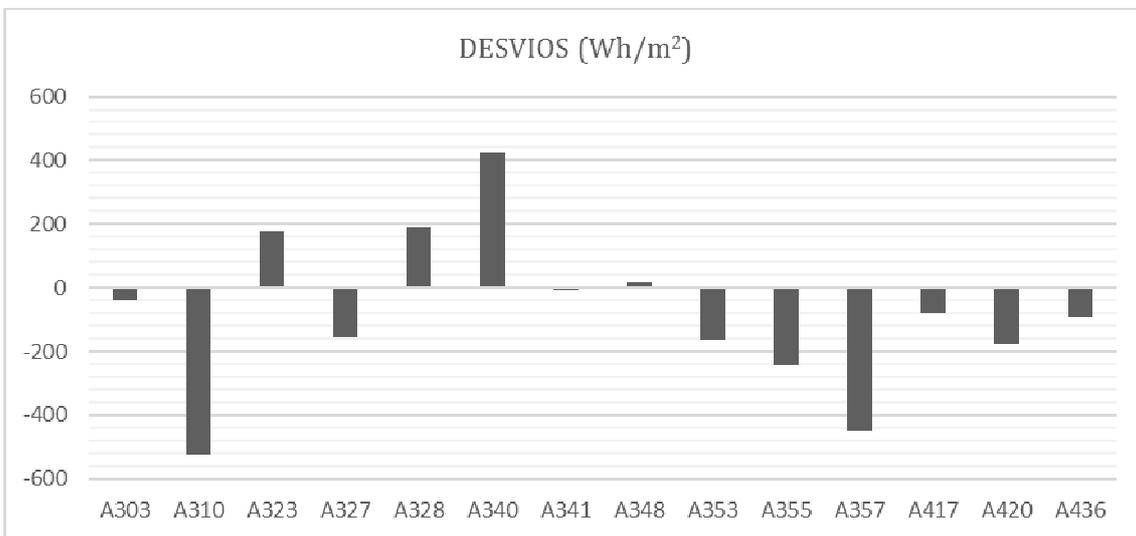


Figura 4.11 – Resíduos obtidos pela subtração das estimativas do modelo BRASIL-SR das estimativas por krigagem ordinária para cada ponto amostral de referência.

A Figura 4.12 exibe a comparação visual da variabilidade espacial da irradiância em superfície estimada por krigagem ordinária e pelo modelo BRASIL-SR.

De acordo com os mapas, pode-se observar um potencial elevado indicando altos níveis de irradiância disponíveis para região NEB no período de 2010. Os valores das estimativas do modelo BRASIL-SR apresentaram pouca variabilidade espacial, com valores entre 4500 Wh/m² e 6000 Wh/m². Os valores estimados por krigagem apresentaram maior amplitude na variabilidade espacial com valores de irradiância entre aproximadamente 3000 Wh/m² e aproximadamente 6000 Wh/m². De um modo geral, na região central da região NEB os dois métodos estimaram valores próximos as 6000 Wh/m².

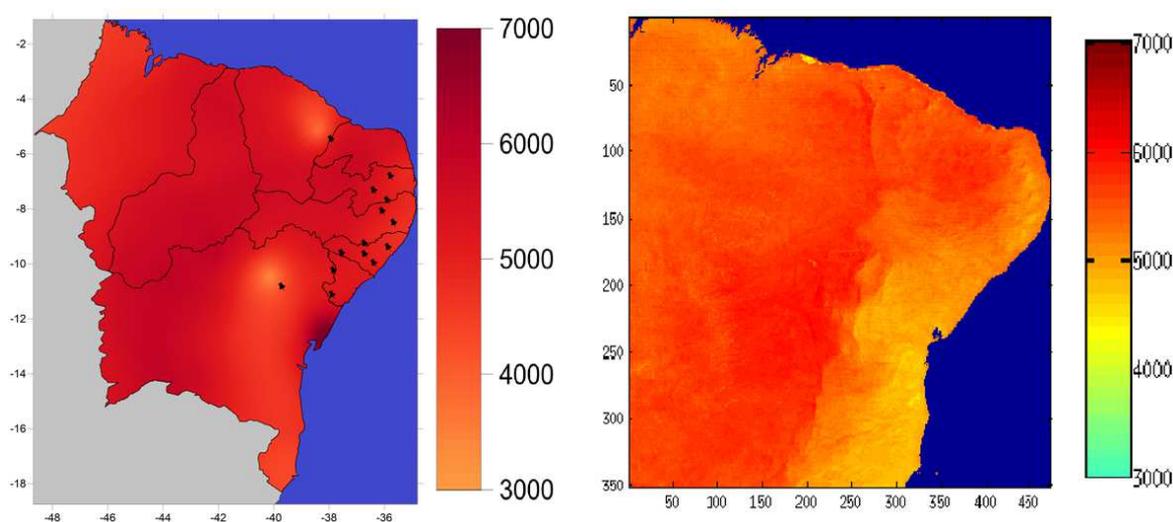


Figura 4.12 – Comparação visual dos mapas de irradiação média anual produzidos através das estimativas para o ano de 2010 por: krigagem ordinária (esquerda) e o modelo numérico BRASIL-SR (direita).

5. CONCLUSÕES E CONSIDERAÇÕES FINAIS

Pôde-se estimar campos de radiação para o nordeste brasileiro aplicando o método de interpolação Kriging. A validação das estimativas foi feita através do método de validação cruzada aliado a análise de índices estatísticos aplicados as séries dos dados estimados em relação a pontos amostrais.

De acordo com os resultados apresentados, pôde-se notar a sensibilidade do método de krigagem ao ajuste local do modelo teórico do semivariograma e seu efeito nas estimativas da irradiância. Nas estimativas produzidas para os totais diários no período analisado (2008, 2009, 2010 e 2011), em dois anos (2010 e 2011) o coeficiente de determinação calculado foi próximo a 0,7 e 0,6 respectivamente. É certo que as séries de dados mais completas e que possuíram o menor número de falhas foram exatamente as séries de 2010 e 2011. Para esses anos, de acordo com o coeficiente de determinação, as estimativas produzidas pelo método de krigagem puderam explicar a variabilidade dos dados nos pontos amostrais em 70 % e 60 %. As séries de dados 2008 e 2009 apresentaram maior número de falhas e, portanto, tiveram reduzida a confiabilidade das estimativas nesse período. Para esses anos o coeficiente de correlação indicou que as estimativas puderam explicar, no máximo, 46 % da variabilidade dos dados medidos.

Assim, o método de krigagem apresentou um bom desempenho para as estimativas que se utilizaram de séries de dados medidos mais completas e, portanto, com um menor número de falhas como nos casos das séries de 2010 e 2011. No entanto, para as séries de 2008 e 2009 as estimativas não apresentaram o mesmo nível de confiabilidade. Esse fato pode ser atribuído, num primeiro momento, ao maior número de falhas nessas séries. Uma investigação a respeito do comportamento do método krigagem em função do número de falhas na série pode ser uma sugestão a ser anexada em uma possível extensão desse trabalho.

Os níveis de irradiância observados nos mapas de irradiância solar média anual dos totais diários nos quatro anos estiveram ente 6000 e 7000 Wh/m². O *EQMA*

mostrou desvios nas estimativas em torno de 5 % em relação aos dados nos pontos amostrais. A análise visual dos mapas de irradiância evidencia certa uniformidade nos campos de radiação exceto por uma região ao norte do estado da Bahia onde se pode ver um gradiente acentuado no campo de irradiância diminuindo em direção ao centro. De certo, alguma (ou algumas) estação medidora nessa região produziu erros sistemáticos nas estimativas. De algum modo, estações medidoras com dados possivelmente suspeitos passaram pelos testes de qualificação e análise dos dados. Com valores que destoaram dos demais (ao redor), provavelmente esses pontos atribuíram o maior peso nos cálculos dos valores estimados ao redor, fato esse, que necessita ser investigado em trabalhos posteriores a esse.

Pôde-se ainda, avaliar as estimativas do modelo de transferência radiativa BRASIL-SR em relação a amostra dos dados de irradiância e em relação as estimativas por interpolação Kriging. O coeficiente de determinação calculado sobre as estimativas para a série de 2010 indicou que o modelo BRASIL-SR conseguiu explicar cerca de 74% da variabilidade dos dados medidos pela rede de PCDs.

Os resultados aqui apresentados contribuem na busca por informações a respeito desse importante recurso natural que, como foi mostrado, é abundante em algumas regiões de nosso país como é o caso da região nordeste. Assim, além de possibilitar contribuições para exploração da radiação solar para a região NEB, os resultados positivos puderam corroborar a metodologia empregada no tratamento e qualificação dos dados medidos pela rede de PCDs do INMET.

Esse trabalho pôde disponibilizar subsídios que podem colaborar para a exploração do recurso solar disponível para a região Nordeste brasileira apresentando mapas de variabilidade espacial de irradiância em superfície, obtidos através das estimativas por interpolação Kriging e amparadas pelas intercomparações realizadas com as estimativas do modelo numérico BRASIL-SR para o ano de 2010 e os dados amostrais. Desse modo, os resultados aqui apresentados forneceram uma validação positiva das estimativas do modelo BRASIL-SR para o ano de 2010,

reforçando a viabilidade do emprego desse modelo em estimativas dos níveis de radiação solar em locais onde não há rede de estações automáticas

Acerca do tratamento imposto aos dados disponibilizados pelo INMET, os resultados desse trabalho ainda podem evidenciar a qualificação dos dados e a escolha dos pontos interpoladores como positivas apoiados pelos índices estatísticos r^2 e $EQMA$ – calculados a partir das estimativas em relação aos dados amostrais. A variabilidade espacial observada nos mapas possibilitou a comparação visual da distribuição e intensidade dos níveis de irradiância no NEB. Ainda, pôde-se verificar a qualidade das medições realizadas por estações automáticas operadas pelo INMET. Constatou-se, pela metodologia adotada, que dos 7 anos de dados medidos disponibilizados para esse estudo, apenas 4 puderam ser empregados. Das séries restantes, em todas haviam falhas e necessitaram se preenchidas. O uso do índice kt foi empregado para a imputação de dados faltantes e, diante dos resultados, produziu efeito positivo sobre as estimativas em pelo menos 2 das séries anuais.

Por fim, algumas sugestões para trabalhos futuros e que possam fazer uso dos resultados aqui alcançados podem ser discutidas:

A princípio, a mesma metodologia de interpolações por krigagem pode ser empregada para um estudo de campos de radiação numa área maior. Essa área pode ser expandida para todo o território nacional.

Uma verificação da sensibilidade e desempenho do método de krigagem aplicado a séries de dados de irradiância com maior número de dados imputados – utilizando o índice kt . Sob a perspectiva da disponibilidade de dados de radiação solar confiáveis ainda ser insuficiente, resultados positivos poderiam corroborar para a melhora de estimativas a partir de dados já disponíveis.

Adequação e ajuste desse método/metodologia de interpolação por krigagem para tratar e imputar valores em séries extensas de dados climatológicos que alimentam parâmetros de modelo numéricos.

REFERÊNCIAS BIBLIOGRÁFICAS

- AZEVEDO, P. V. De; SILVA, B. B., Da; RODRIGUES, M. F. G. Previsão das chuvas de outono no estado do Ceará. **Revista Brasileira de Meteorologia**, v. 13, n. 1, p. 19-30, 1998.
- BLAIR, G. C.; BRUNINI, O. Caracterização do regime de evapotranspiração real, em escala decenal, no estado de São Paulo. **Revista Brasileira de Meteorologia**, v. 22, n. 1, p. 75-82, 2007.
- BRASIL. Empresa de Pesquisa Energética. **Balanco Energético Nacional 2013 – Ano base 2012: Relatório Síntese**. Rio de Janeiro: 2013. 55 p.
- BURROUGH, P.; McDONNELL, R. **Principles of Geographical Information Systems: spatial information systems and geostatistics**. Oxford: University Press, 1998. 333 p. ISBN 0-19-823366-3.
- CAMARGO, E. C. G. **Desenvolvimento, implementação e teste de procedimentos geostatísticos (Krigagem) no sistema de processamento de informações georreferenciadas (SPRING)**. 1997. 146 p. (INPE-6410-TDI/620). Dissertação (Mestrado em Sensoriamento Remoto) - Instituto Nacional de Pesquisas Espaciais, São Jose dos Campos, 1997. Disponível em: <<http://urlib.net/6qtX3pFwXQZ3r59YD6/GPfam> >. Acesso em: 26 jun. 2012.
- CAMARGO, A. P.; CAMARGO, M. B. P. **Uma revisão analítica da evapotranspiração potencial**. Bragantina, Campinas, v. 59, n. 2, p. 125-137, 2000.
- CARDENAS, R. R.; KRAINSKI, E. T. Preenchimento de falhas em bancos de dados meteorológicos diários: uma comparação de abordagens. In: CONGRESSO BRASILEIRO DE AGROMETEOROLOGIA, 17., 2011, Guarapari – ES. **Anais...** Pelotas, RS: SBA, 2011.
- CAVALCANTE, F. A. et al. (orgs). **Tempo e clima no Brasil**. São Paulo: Oficina de Textos, 2009. 432 p.
- CHANG, E. K. M; LEE, S.; SWANSON, K. L. Storm track dynamics. **Journal of Climate**, v. 15, n. 16, p. 2163-2183, 2002.
- COSTA, R. S. **Influência dos aerossóis atmosféricos na quantificação do recurso energético solar: experimentos em modelo de transferência radiativa**. 2012. 247 p. (sid.inpe.br/mtc-m19/2012/03.28.14.21-TDI). Tese (Doutorado em Meteorologia) – Instituto Nacional de Pesquisas Espaciais (INPE), São José dos Campos, 2004. Disponível em: <http://urlib.net/8JMKD3MGP7W/3BK7PGB>. Acesso em: 20 fev. 2014.

FERNANDEZ, M. N. **Preenchimento de falhas em séries temporais**. 2007. p. Dissertação (Mestrado em Engenharia Oceânica) – Fundação Universidade Federal do Rio Grande, 2007. Disponível em: <http://repositorio.furg.br:8080/handle/1/3441>. Acesso em 24 Fev. 2014.

FERREIRA, A. G.; MELLO, N. G. Principais sistemas atmosféricos atuantes sobre a região nordeste do Brasil e a influência dos oceanos pacífico e atlântico no clima da região. **Revista Brasileira de Climatologia**, v. 1, n. 1, p. 15-28, 2005.

GALVÃO, M. V. Regiões climáticas do Brasil. **Revista Brasileira de Geografia**. v. 29. n. 1. p. 3-36, 1967. Disponível em: http://biblioteca.ibge.gov.br/visualizacao/periodicos/115/rbg_1967_v29_n1.pdf

GUARNIERI, R. A. **Emprego de redes neurais artificiais e regressão linear múltipla no refinamento das previsões de radiação solar do modelo ETA**. Dissertação (Mestrado em Meteorologia) - Instituto Nacional de Pesquisas Espaciais (INPE), São José dos Campos, 2006. Disponível em: <<http://urlib.net/6qtX3pFwXQZGivnJSY/LiF8s>>. Acesso em: 28 jul. 2014.

HARTKAMP, A.D.; DE BEURS, K.; STEIN, A. e WHITE, J.W. **Interpolation techniques for climate variables**. Mexico, D. F.: CIMMYT, 1999. 34p. (NRG-GIS Series 99-01).

HUIJBREGTS, C. J. Regionalized variables and quantitative analysis of spatial data. In: DAVIS, J. C.; McCULLAGH, M. J.(eds). **Display and analysis of spatial data**. New York: John Wiley, 1975. p. 38-53.

IQBAL, M. **An introduction to solar radiation**. Elsevier, 1983.

ISAACS, E. H.; SRIVASTAVA, R. M. **An introduction to applied geostatistics**. New York: Oxford University Press, 1989. 560 p.

JOURNEL, A. G. **Fundamentals of geostatistics in five lessons**. California: Stanford Center for Reservoir Forecasting Applied Sciences department, 1988.

LANDIM, P. M. B. **Introdução aos métodos de estimação espacial para confecção de mapas**. Rio Claro: DGA, IGCE, UNESP. Lab. Geomatemática, 2000. 20 p. Texto Didático.

_____. **Análise estatística de dados geológicos**. 2. ed. São Paulo: Editora UNESP, 2003. 253 p. ISBN(85-7139-504-7).

LANDIM, P. M. B.; STURARO, J. R. **Krigagem indicativa aplicada à elaboração de mapas probabilísticos de riscos**. Rio Claro: Unesp. Lab. Geomatemática, 2002. 20 p. (Texto Didático 06). Disponível em:

http://www.rc.unesp.br/igce/geologia/GAA01001/aulas_files/kindicativa.pdf.

Acesso em: 12 fev. 2014.

LIU, K. N. **An introduction to atmospheric radiation**. 2. ed. San Diego: Academic Press, 2002.

MARTINS, F. R.; PEREIRA, E. B.; ABREU, S. L. D.; COLLE, S. Mapas de irradiação solar para o Brasil – resultados do projetos SWERA. In: SIMPÓSIO BRASILEIRO DE SENSORIAMENTO REMOTO, 12. (SBSR), 2005, Goiânia. **Anais...** São José dos Campos: INPE, 2005. p. 3137-3146. CD-ROM, On-line. ISBN 85-17-00018-8. (INPE-12716-PRE/8006). Disponível em: <<http://urlib.net/ltid.inpe.br/sbsr/2004/11.19.15.49>>. Acesso em: 15 maio 2014.

MARTINS, F. R., PEREIRA, E. B. Estudo comparativo da confiabilidade de estimativas de irradiação solar para o sudeste brasileiro obtidas a partir de dados de satélite e por interpolação/extrapolação de dados de superfície. **Revista Brasileira de Geofísica**. v. 29, n. 2, p. 265-276, 2011.

MATHERON, G. **The theory of regionalized variables and its applications**. Paris: Les Cahiers du Centre de Morphologie Mathématique de Fontaniebleu, 1971. 211 p.

MAZZINI, P. L. F.; SCHETTINI, C. A. F. Avaliação de metodologias de interpolação espacial aplicadas a dados hidrográficos costeiros quase sinóticos. **Brazilian Journal of Aquatic Science and Technology**, Itajaí, v. 13, n. 1, p. 53-64, jul. 2009. ISSN 1983-9057. Disponível em: <<http://siaiweb06.univali.br/seer/index.php/bjast/article/view/1336/1058>>. Acesso em: 12 Fev. 2014. doi:<http://dx.doi.org/10.14210/bjast.v13n1.p53-64>.

McARTHUR LJB. **Baseline Surface Radiation Network (BSRN)**. Operations Manual. 2004. WMO/TD-No. 879, WCRP/WMO. 188 p.

MOURA, M. S. B. et al. Clima e água de chuva no semi-árido. In: BRITO, L. T. de L.; MOURA, M. S. B.; GAMA, G. F. B. (Eds.). **Potencialidades da água de chuva no semi-árido brasileiro**. Petrolina: Embrapa Semi-Árido, 2007. Cap. 2, 37-59 p. Disponível em: <http://www.alice.cnptia.embrapa.br/handle/doc/159649>. Acesso em 08 fev. 2014.

MOLION, L. C. B.; BERNARDO, S. de O. Uma revisão da dinâmica das chuvas no Nordeste Brasileiro. **Revista Brasileira de Meteorologia**, v. 17, n. 1, p. 1-10, 2002.

OLIVEIRA, F.K.S.; SARAIVA, M. F. O. **Astronomia e astrofísica**. 2. ed. Porto Alegre: Editora Livraria da Física, 2004. 557 p.

PEREIRA, E. B.; MARTINS, F. R.; ABREU, S. L. A.; RUTHER, R. **Atlas brasileiro de energia solar**. São José dos Campos: INPE, 2006. 60p.

ROESCH, A.; WILD, M.; OHMURA, A.; DUTTON, E.G.; LONG, C.N.; ZHANG, T. Assessment of BSRN radiation records for the computation of monthly means. **Atmospheric Measurement Techniques**, v.4, p. 339–354, 2011.

RUGGIERO, M. A. G.; ROCHA, V. L. **Cálculo numérico: aspectos teóricos e computacionais**. 2. ed. São Paulo: Editora Pearson, 1996. 406 p.

TIBA, C. et al. **Atlas solarimétrico do Brasil: banco de dados terrestres**. Recife: Editora Universitária da UFPe, 2000. 111 p. Mapas.

TIPLER, P. A. **Física**. Rio de Janeiro: Editora Guanabara Dois S. A., 1978. p. 516-999. v. 2

TRENBERT, K. E.; CARON, J. M.; D. P. STEPANIAK. The atmospheric energy budget and implications for surface fluxes and ocean heat transports. **Climate Dynamics**, v. 17, p. 259-276, 2001.

TRENBERT, K. E.; SOLOMON, A. The global heat balance: heat transports in the atmosphere and ocean. **Climate Dynamics**, v. 10, p. 107-134, 1994.

VAREJÃO, S. M. A. **Meteorologia e climatologia**. Recife, 2006. Versão digital 2, cap. X. 449 p. Disponível em:
http://www.agritempo.gov.br/modules.php?name=Downloads&d_op=viewdownload&cid=19. Acesso em: 23 dez. 2013.

VIANA, T. S.; RUTHER, R.; MARTINS, F. R.; PEREIRA, E. B. Assessing the potential of concentrating solar photovoltaic generation in Brazil with satellite-derived direct normal irradiation. **Energy Solar**, v. 85, p. 486-495, 2011.

VIANELLO, R. L.; ALVES, A. R. **Meteorologia básica e aplicações**. Viçosa: Universidade Federal de Viçosa, 1991. 448 p.

WALLACE, J. M.; HOBBS, P. V. **Atmospheric science: an introductory survey**. Academic Press, 2006.