

# Análise comparativa entre diferentes métodos de interpolação para estimativa da temperatura média do ar no Brasil

Igor Balteiro Pereira de Campos <sup>(1)</sup>

<sup>(1)</sup> Universidade Federal do Rio de Janeiro – Brasil, igor@lma.ufrj.br

## 1. Introdução

Em estudos climatológicos, uma das maiores dificuldades relaciona-se com a obtenção de valores de elementos climáticos onde não existem estações de medição. Assim, para contornar tal problema, utilizam-se técnicas de interpolação que transformam valores pontuais em superfícies contidas no limite da malha de pontos considerada, sendo, estas técnicas, conhecidas por técnicas de interpolação espacial.

A interpolação espacial é o procedimento para se estimar valores de propriedades de locais não-amostrados, baseando-se em valores de dados observados em dados conhecidos (Burrough, 1986). Os métodos de interpolação não possuem uma classificação muito clara, mas podem ser classificados, de forma didática, em: (1) globais ou locais; (2) exatos ou suavizantes; e (3) determinísticos ou estocásticos (Hartkamp *et al.*, 1999).

As classificações quanto aos métodos serem globais ou locais, determinísticos ou estocásticos, exatos ou suavizantes é, em termos práticos, subjetiva, pois a grande maioria dos softwares de interpolação permite ao usuário alterar, para o mesmo método, os parâmetros de suavização, número de dados observados a serem utilizados no raio de influência, dentre outros fatores, fazendo, assim, que o método de interpolação mude de classificação.

Hartkamp *et al.* (1999), apresentam que quando os dados são abundantes e regularmente distribuídos, as metodologias de interpolação geram resultados similares. Por outro lado, quando os dados são escassos, há uma grande variação de resultados gerados. Portanto, a questão de qual é o melhor método a ser aplicado sobre um determinado conjunto de dados amostrados torna-se pertinente.

## 2. Objetivo

O objetivo deste trabalho é avaliar o desempenho de algoritmos de interpolação para a estimativa de campos de temperatura média do ar no território brasileiro.

## 3. Metodologia

A área de estudo é o território brasileiro, que se situa na faixa que vai de 5.27°N até 33.74°S,

aproximadamente, e 35°W até 74°W, aproximadamente. Foram utilizados dados de 206 estações meteorológicas pertencentes ao Instituto Nacional de Meteorologia (INMET) e referem-se às Normais Climatológicas de 1961 a 1990, sendo escolhidos todos os meses do ano, além da média anual para este estudo. Também foram utilizados dados de aeroportos para fazer a validação dos métodos. A Figura 1 mostra a distribuição das estações do INMET (pontos vermelhos) e os aeroportos selecionados (triângulos roxos).

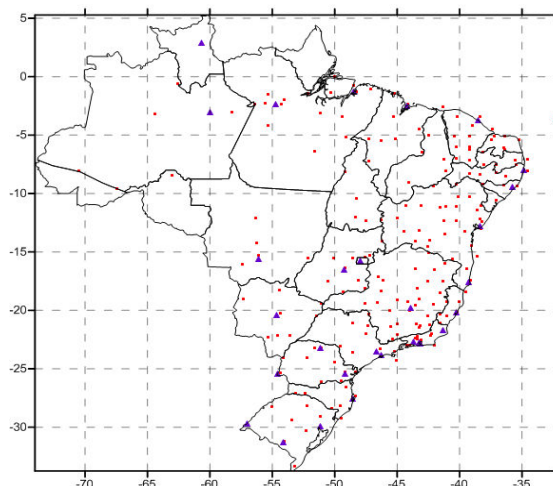


Fig. 1 – Área de estudo.

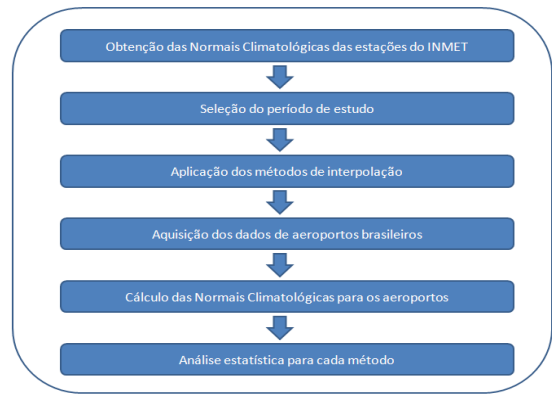
Os métodos de interpolação utilizados foram escolhidos seguindo o critério de serem os mais encontrados na literatura e em programas de computador de análise de dados espaciais. Tais métodos são encontrados no software Surfer<sup>®</sup> 9 (Golden Software, 2009). Este trabalho não tem o objetivo de descrever os algoritmos matemáticos empregados em cada método, mas sim sua idéia, função e aplicação prática. Dos 12 algoritmos de interpolação presentes no software Surfer<sup>®</sup> foram selecionados três, dos quais a seguir se faz uma breve descrição (Mazzini e Schettini, 2009):

- 1) Kriging: método de interpolação estocástico que utiliza geoestatística para efetuar a interpolação, o que, em muitos casos, é uma grande vantagem sobre outros métodos. O termo ‘geoestatística’ foi dado por Matheron em seu trabalho para a solução de problemas espaciais

voltados para a mineração (Oliveira, 1991). Matheron foi influenciado através dos trabalhos realizados pelo engenheiro de minas Daniel G. Krige, na África do Sul, surgindo daí o nome do método. A Krigagem define o grau de dependência ou correlação espacial entre as amostras através do semivariograma. Uma vez modelado o semivariograma, é possível verificar o nível de anisotropia dos dados, e então definir os melhores pesos para as amostras. Kriging pode ser um interpolador tanto exato como suavizador. O método tenta expressar tendência sugerida pelos dados;

- 2) Minimum Curvature (Spline): método muito aceito e utilizado atualmente com o nome derivado de uma ferramenta flexível do desenho técnico. Diferencia-se dos outros métodos polinomiais, pois não utiliza apenas um polinômio de grande ordem para a interpolação de todo o conjunto de dados, mas divide a série de dados em subconjuntos e utiliza polinômios de pequenas ordens para cada subconjunto. A soma ou junção destes polinômios de pequenas ordens é que forma a interpolação sobre todo o domínio. O grau mais utilizado é o grau 3, o spline cúbico. Este método, que é muito utilizado em geociências, gera curvas mais suaves ao mesmo tempo em que tenta honrar ao máximo os dados. Entretanto, não é um interpolador exato. Outras vantagens do Minimum Curvature são a boa convergência, aproximações precisas das derivações, e boa estabilidade na presença de erros de aproximação;
- 3) Radial Basis Function: é um conjunto de métodos interpoladores. A maioria dos métodos são derivações de spline, com características similares uns dos outros. O método de derivação multiquadrático é o padrão automático no Surfer<sup>®</sup>, pois é considerado o melhor na maioria dos casos.

A metodologia empregada neste trabalho é apresentada na Figura 2.



**Fig. 2** – Metodologia.

Para validação dos campos interpolados, foram utilizados dados de 29 aeroportos brasileiros. A partir destes dados calcularam-se as Normais Climatológicas, também para o período de 1961 a 1990, para cada aeroporto. Este procedimento foi feito através de rotinas na linguagem Shell Script. Após este procedimento foi feita a análise estatística para cada método, utilizando os seguintes parâmetros estatísticos:

- 1) Cálculo da Tri-Média (T-M) dos dados observados e previstos para cada período;
- 2) Cálculo do Intervalo Inter-Quartilício (IQR) para os mesmos;
- 3) Cálculo da diferença entre a T-M do previsto e a T-M do observado;
- 4) Cálculo do Erro Absoluto Médio (MAE) - (Equação 1);
- 5) Cálculo da Raiz do Erro Quadrático Médio (RMSE) - (Equação 2); e
- 6) Cálculo do Índice de Concordância (*d*) - (Equação 3) → varia entre 0.0 a 1.0. Quanto mais próximo de 1.0 melhor é a concordância entre o campo observado e o campo previsto.

Esses métodos de validação seguem o procedimento descrito por Willmott (1984). A validação calculou os erros estatísticos em 29 pontos-controle, assumindo como dados observados os dados dos aeroportos e, como dados previstos (interpolados), os dados gerados pelos métodos em estudo.

O uso do MAE como método validador é, às vezes, preferível em relação ao uso do RMSE, pois é menos sensível à valores extremos; entretanto, o RMSE é uma medida de erro comumente computada em aplicações geográficas.

$$MAE = N^{-1} \sum_{i=1}^N |P_i - O_i| \quad \text{(Equação 1)}$$

$$RMSE = N^{-1} \sum_{i=1}^N [(P_i - O_i)^2]^{-1/2} \text{ (Equação 2)}$$

$$d = 1 - \left[ \frac{\sum_{i=1}^N (P_i - O_i)}{PE} \right],$$

$$\text{onde } PE = \sum_{i=1}^N (|P_i - \bar{O}| + |O_i - \bar{O}|)^2 \text{ (Equação 3),}$$

onde P = previsto (interpolado) e O = observado.

#### 4. Resultados

Nas Figuras 3, 4 e 5 são apresentados os campos interpolados de temperatura média do ar para os três métodos em estudo, tendo sido selecionado os meses de janeiro, julho e a média anual para ilustração.

A partir da análise visual das Figuras 3, 4 e 5 nota-se um comportamento similar entre os métodos Kriging e Minimum Curvature e, percebe-se também, que o algoritmo Radial Basis Function foi o que mais suavizou o campo de temperatura média. Observa-se que todos mantiveram a mesma feição climatológica, onde a faixa equatorial apresentou os maiores valores e, na faixa das latitudes médias, apresentam-se as temperaturas mais baixas.

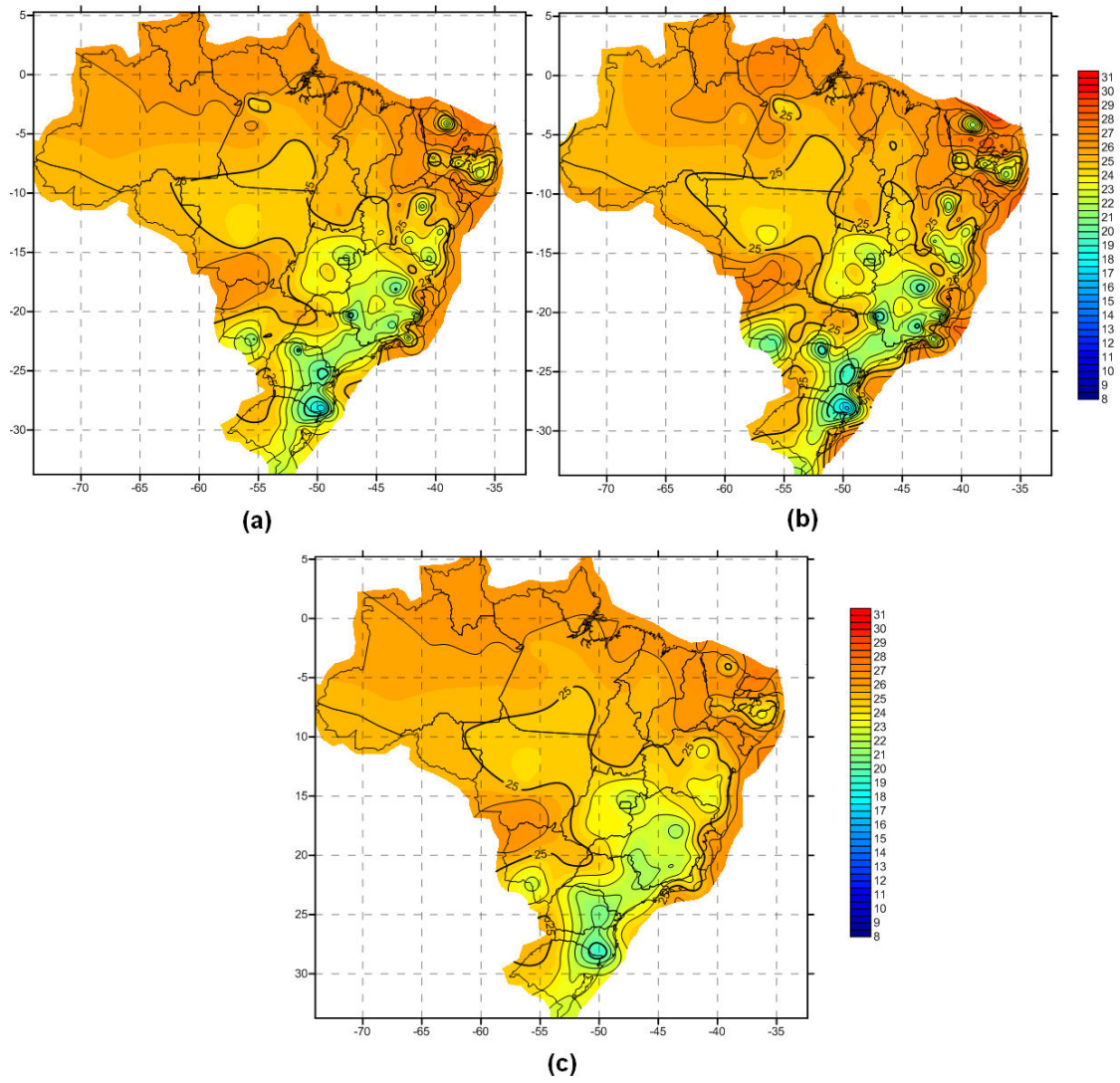
Foram gerados os resíduos, em forma de histogramas (Figuras 6, 7, 8, 9 e 10) e tabela (Tabela 1), de cada campo interpolado, para verificar se cada método tende a honrar ou não os dados originais.

A partir das figuras 6, 7, 8, 9 e 10 percebe-se que os métodos de Kriging e Minimum Curvature tendem a honrar os dados originais, pois seus valores ficaram muito abaixo da faixa de imprecisão adotada ( $\pm 1.0^\circ\text{C}$ ). Já o algoritmo Radial Basis Function teve um desempenho inferior aos métodos já citados, tendo seus valores de resíduo situados na faixa de  $-6^\circ\text{C}$  a  $4^\circ\text{C}$ . A tabela 1 mostra, em forma de percentual, a quantidade de pontos agrupados em intervalos de classe.

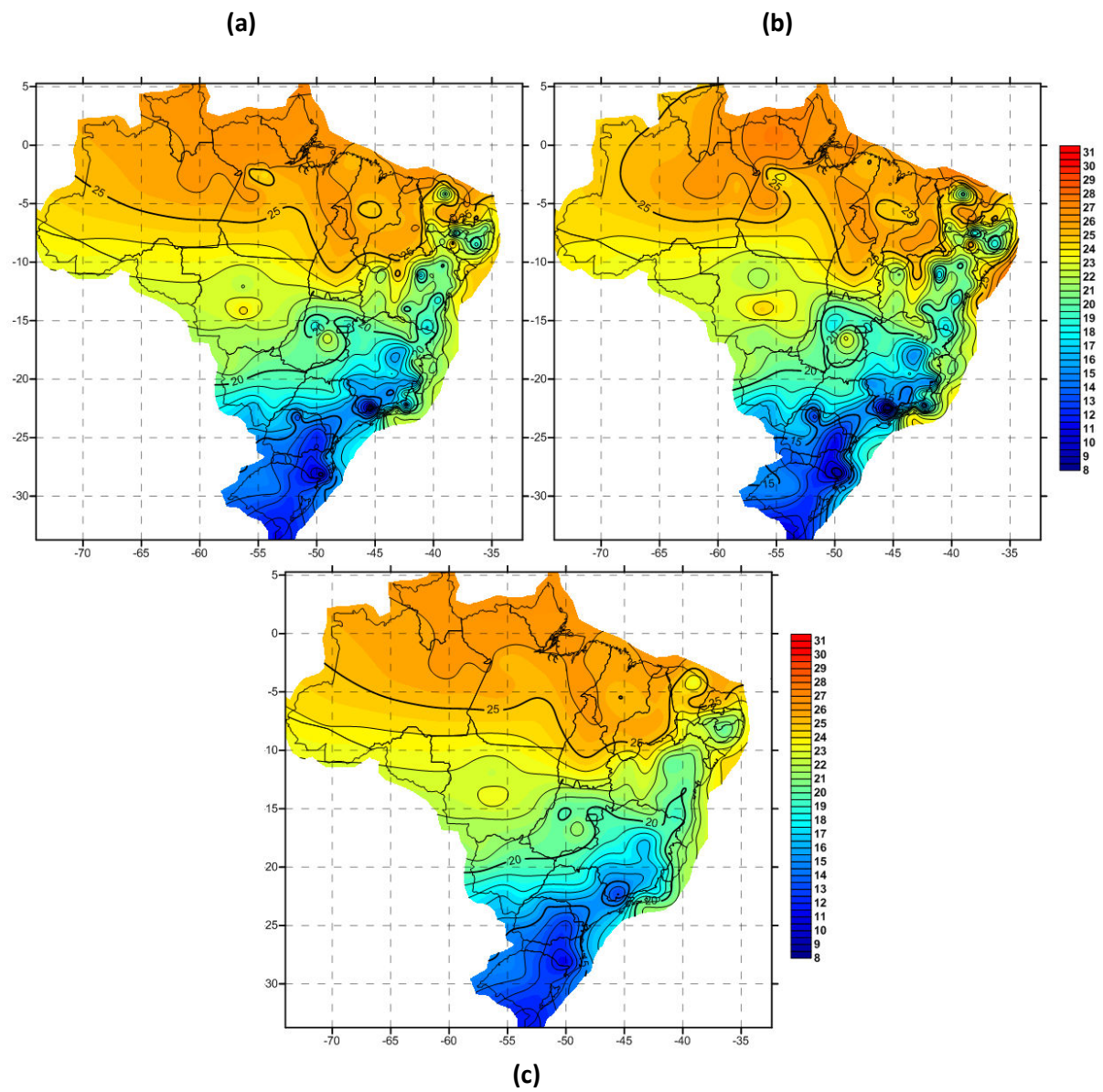
Frequencia relativa dos resíduos por classes para cada período					
Período	Método	Classes			
		(-6.0)   - (-1.0)	(-1.0)   - (0.0)	0.0   - 1.0	1.0   - 6.0
Janeiro	Kriging	0.00	0.57	0.43	0.00
	Minimum Curvature	0.00	0.47	0.53	0.00
	Radial Basis Function	0.12	0.44	0.34	0.10
Fevereiro	Kriging	0.00	0.57	0.43	0.00
	Minimum Curvature	0.00	0.50	0.50	0.00
	Radial Basis Function	0.11	0.45	0.34	0.10
Marco	Kriging	0.00	0.55	0.45	0.00
	Minimum Curvature	0.00	0.46	0.54	0.00
	Radial Basis Function	0.13	0.42	0.33	0.12
Abril	Kriging	0.00	0.54	0.46	0.00
	Minimum Curvature	0.00	0.57	0.43	0.00
	Radial Basis Function	0.00	0.47	0.53	0.00
Maio	Kriging	0.12	0.42	0.34	0.12
	Minimum Curvature	0.00	0.56	0.44	0.00
	Radial Basis Function	0.00	0.48	0.52	0.00
Junho	Kriging	0.13	0.42	0.33	0.12
	Minimum Curvature	0.00	0.58	0.42	0.00
	Radial Basis Function	0.00	0.49	0.51	0.00
Julho	Kriging	0.11	0.48	0.29	0.12
	Minimum Curvature	0.00	0.57	0.43	0.00
	Radial Basis Function	0.00	0.48	0.52	0.00
Agosto	Kriging	0.00	0.47	0.53	0.00
	Minimum Curvature	0.14	0.46	0.30	0.10
	Radial Basis Function	0.00	0.56	0.44	0.00
Setembro	Kriging	0.00	0.49	0.51	0.00
	Minimum Curvature	0.11	0.46	0.31	0.12
	Radial Basis Function	0.00	0.55	0.45	0.00
Outubro	Kriging	0.00	0.49	0.51	0.00
	Minimum Curvature	0.12	0.44	0.34	0.10
	Radial Basis Function	0.00	0.57	0.43	0.00
Novembro	Kriging	0.00	0.48	0.52	0.00
	Minimum Curvature	0.10	0.49	0.32	0.09
	Radial Basis Function	0.10	0.48	0.32	0.09
Dezembro	Kriging	0.00	0.56	0.44	0.00
	Minimum Curvature	0.00	0.48	0.52	0.00
	Radial Basis Function	0.13	0.44	0.33	0.10
Anual	Kriging	0.00	0.56	0.44	0.00
	Minimum Curvature	0.00	0.50	0.50	0.00
	Radial Basis Function	0.12	0.44	0.33	0.11

**Tabela 1** – Frequência relativa de pontos por faixa de resíduos para cada período.

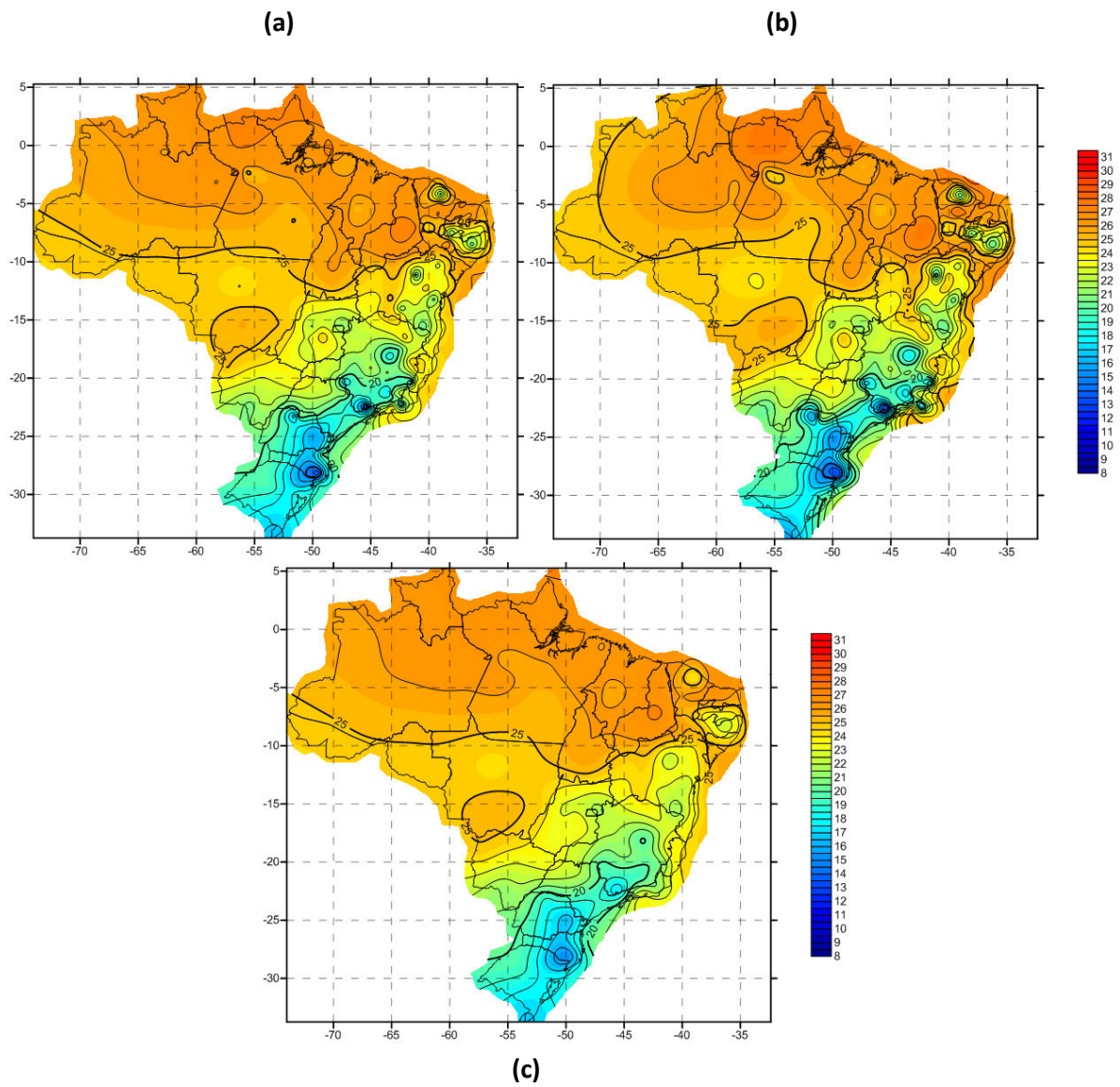
Através da Tabela 1 nota-se também, que além do que já foi mencionado acima, os métodos de Kriging e Minimum Curvature tendem, respectivamente, a superestimar e subestimar os valores originais. O método Radial Basis Function tende a superestimar os dados originais. Após a análise dos resíduos, foram aplicados os testes estatísticos descritos por Willmott (1984) para validar os métodos em estudo. A Tabela 2 mostra a síntese de todos os resultados obtidos.



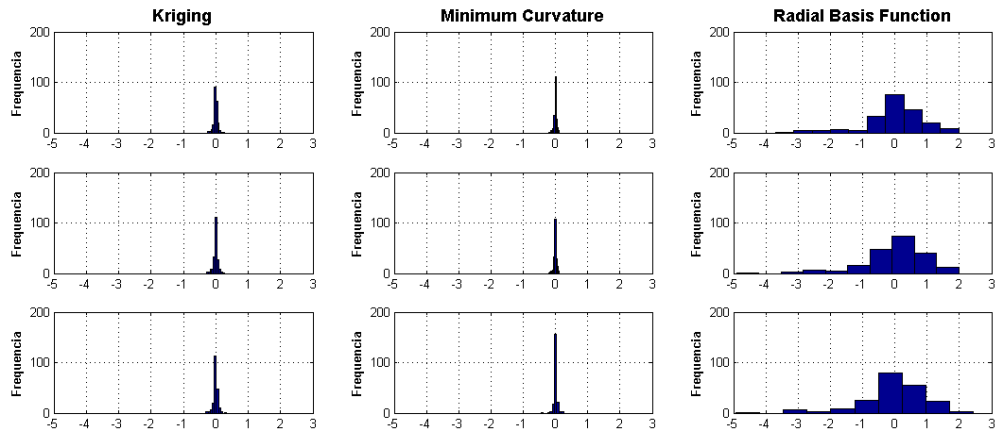
**Fig. 3** – Campo de temperatura média do ar em janeiro utilizando os métodos: (a) Kriging, (b) Minimum Curvature e (c) Radial Basis Function.



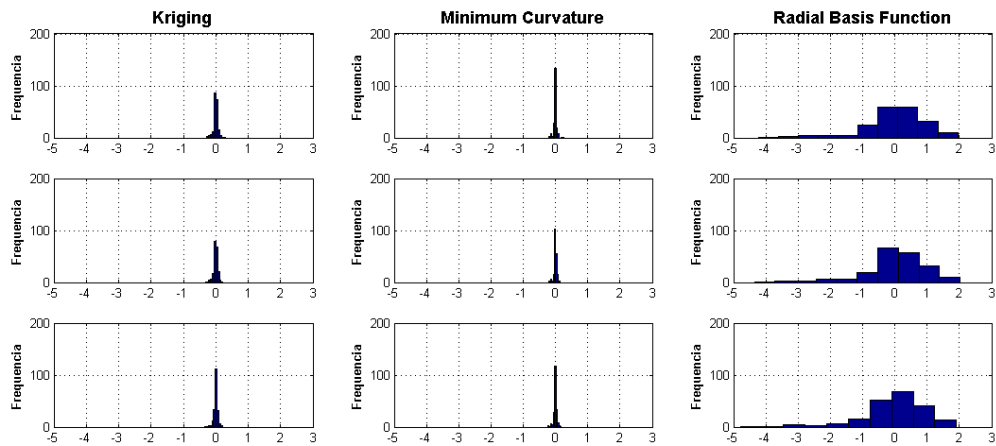
**Fig. 4** – Campo de temperatura média do ar em julho utilizando os métodos: (a) Kriging, (b) Minimum Curvature e (c) Radial Basis Function.



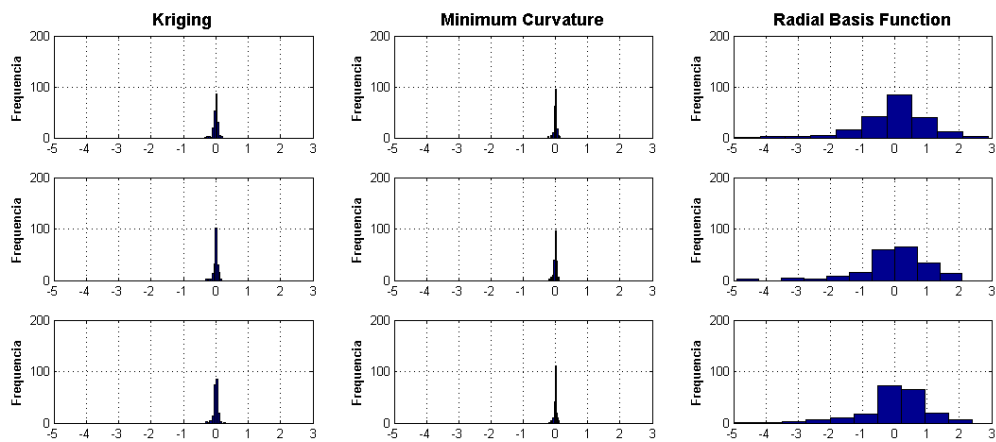
**Fig. 5** – Campo de temperatura média do ar na média anual utilizando os métodos: (a) Kriging, (b) Minimum Curvature e (c) Radial Basis Function.



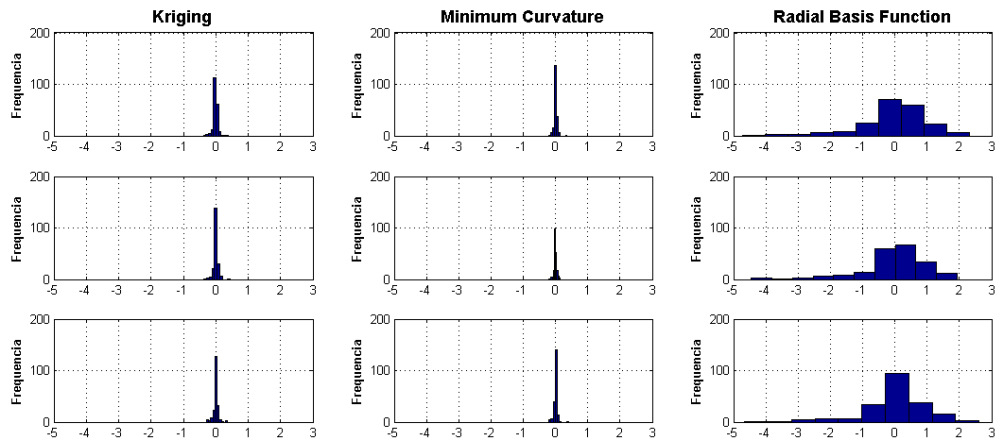
**Fig. 6** – Histogramas para os métodos de: Kriging, Minimum Curvature e Radial Basis Function. Cada linha representa um período, seguindo de cima para baixo: Janeiro, Fevereiro e Março.



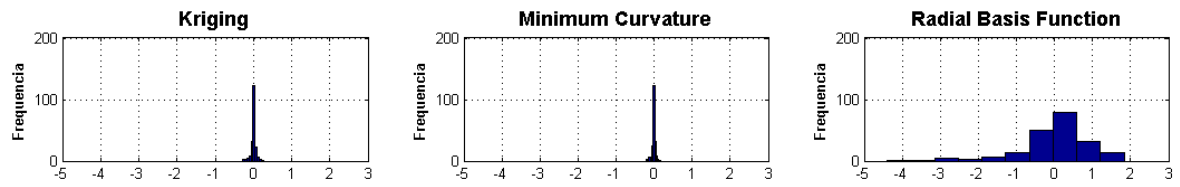
**Fig. 7** – Histogramas para os métodos de: Kriging, Minimum Curvature e Radial Basis Function. Cada linha representa um período, seguindo de cima para baixo: Abril, Maio e Junho.



**Fig. 8** – Histogramas para os métodos de: Kriging, Minimum Curvature e Radial Basis Function. Cada linha representa um período, seguindo de cima para baixo: Julho, Agosto e Setembro.



**Fig. 9** – Histogramas para os métodos de: Kriging, Minimum Curvature e Radial Basis Function. Cada linha representa um período, seguindo de cima para baixo: Outubro, Novembro e Dezembro.



**Fig. 10** – Histogramas para os métodos de: Kriging, Minimum Curvature e Radial Basis Function para a média anual de temperatura.



Período	Método	N	T-M Observado	T-M Previsto	Diferença (T-Mprev - T-Mobser)	IQR (Observado)	IQR (Previsto)	Diferença (IQRprev - IQRobser)	MAE	RMSE	d
Janeiro	Kriging	29	26.2	25.2	-1.1	2.2	2.2	0.0	1.61	2.03	0.68
	Minimum Curvature	29	26.2	25.2	-1.0	2.2	2.3	0.1	1.59	2.07	0.69
	Radial Basis Function	29	26.2	24.8	-1.5	2.2	2.6	0.4	1.83	2.13	0.65
Fevereiro	Kriging	29	26.3	24.7	-1.6	2.0	2.9	0.8	1.70	2.10	0.65
	Minimum Curvature	29	26.3	24.9	-1.4	2.0	2.6	0.5	1.66	2.13	0.66
	Radial Basis Function	29	26.3	24.7	-1.6	2.0	2.8	0.7	1.90	2.24	0.61
Março	Kriging	29	25.9	24.5	-1.3	2.4	3.4	1.1	1.64	2.07	0.73
	Minimum Curvature	29	25.9	24.5	-1.3	2.4	3.4	1.0	1.66	2.12	0.73
	Radial Basis Function	29	25.9	24.5	-1.4	2.4	3.1	0.7	1.85	2.20	0.69
Abril	Kriging	29	24.5	23.6	-0.9	3.5	4.5	1.0	1.70	2.12	0.91
	Minimum Curvature	29	24.5	23.7	-0.7	3.5	4.6	1.1	1.77	2.21	0.90
	Radial Basis Function	29	24.5	23.0	-1.5	3.5	4.9	1.4	1.87	2.24	0.90
Maio	Kriging	29	22.7	21.9	-0.8	5.4	6.6	1.3	1.70	2.29	0.92
	Minimum Curvature	29	22.7	22.2	-0.5	5.4	6.3	0.9	1.78	2.36	0.92
	Radial Basis Function	29	22.7	21.3	-1.4	5.4	6.6	1.2	1.91	2.37	0.92
Junho	Kriging	29	21.3	20.5	-0.8	6.2	7.3	1.1	1.64	2.25	0.91
	Minimum Curvature	29	21.3	20.9	-0.4	6.2	7.2	1.0	1.76	2.31	0.90
	Radial Basis Function	29	21.3	19.9	-1.4	6.2	7.4	1.2	1.84	2.32	0.90
Julho	Kriging	29	21.1	20.2	-0.8	5.4	6.6	1.3	1.77	2.37	0.92
	Minimum Curvature	29	21.1	20.7	-0.4	5.4	6.8	1.4	1.87	2.46	0.91
	Radial Basis Function	29	21.1	19.3	-1.8	5.4	7.0	1.6	1.96	2.45	0.92
Agosto	Kriging	29	21.9	21.0	-0.9	3.9	5.9	2.1	1.57	2.06	0.93
	Minimum Curvature	29	21.9	21.3	-0.6	3.9	5.6	1.8	1.65	2.15	0.93
	Radial Basis Function	29	21.9	20.2	-1.7	3.9	6.1	2.3	1.85	2.18	0.93
Setembro	Kriging	29	22.5	21.4	-1.1	4.0	6.1	2.1	1.43	1.86	0.96
	Minimum Curvature	29	22.5	21.8	-0.7	4.0	5.8	1.8	1.50	1.99	0.95
	Radial Basis Function	29	22.5	20.9	-1.7	4.0	5.6	1.6	1.64	1.94	0.96
Outubro	Kriging	29	23.6	22.5	-1.1	2.8	5.2	2.4	1.59	2.02	0.88
	Minimum Curvature	29	23.6	22.7	-0.9	2.8	5.2	2.4	1.61	2.14	0.86
	Radial Basis Function	29	23.6	22.0	-1.6	2.8	5.1	2.3	1.77	2.08	0.87
Novembro	Kriging	29	24.8	23.7	-1.0	2.2	3.7	1.4	1.55	1.96	0.82
	Minimum Curvature	29	24.8	24.0	-0.7	2.2	3.4	1.1	1.56	2.08	0.80
	Radial Basis Function	29	24.8	23.2	-1.6	2.2	3.8	1.5	1.71	2.05	0.81
Dezembro	Kriging	29	25.6	24.5	-1.1	2.3	2.8	0.6	1.65	2.07	0.90
	Minimum Curvature	29	25.6	24.6	-1.0	2.3	3.0	0.8	1.58	2.10	0.89
	Radial Basis Function	29	25.6	24.1	-1.5	2.3	3.4	1.1	1.84	2.18	0.89
Média Anual	Kriging	29	25.6	23.0	-2.7	3.5	4.6	1.1	1.58	2.03	0.87
	Minimum Curvature	29	25.6	23.3	-2.3	3.5	4.7	1.2	1.61	2.10	0.87
	Radial Basis Function	29	25.6	22.4	-3.3	3.5	5.0	1.4	1.78	2.14	0.86

N = número de aeroportos

T-M = Tri-Média

IQR = Intervalo Inter-Quartilico

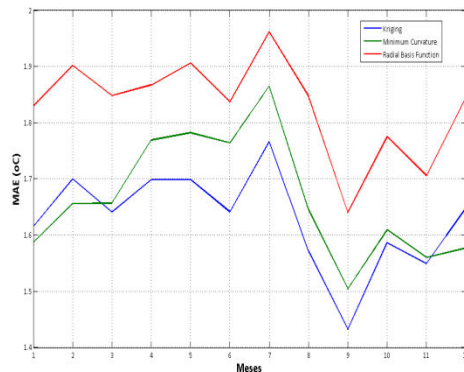
**Tabela 2 - Análise estatística de cada método para cada período**

Para o cálculo das medidas de posição e de dispersão foram escolhidos Tri-Média (T-M) e o Intervalo Inter-Quartilico (IQR), respectivamente. Essa escolha é devida ao fato destes métodos serem robustos e resistentes em relação aos métodos de posição (média, mediana) e dispersão (desvio-padrão, variância) clássicos.

A partir da Tabela 2, percebe-se que:

- 1) Todos os métodos subestimaram os dados originais, o que pode ser notado pela diferença entre a T-M dos dados previstos e a T-M dos dados observados. Destacam-se os métodos Minimum Curvature e Radial Basis Function por apresentarem à menor e a maior diferença, respectivamente, entre o previsto e o observado;
- 2) O IQR dos métodos foi muito semelhante, destacando apenas o Radial Basis Function que obteve os maiores valores de IQR em todo o período de estudo;
- 3) O Erro Médio Absoluto (MAE) foi muito similar para os três métodos, destacando-se os métodos de Kriging e Radial Basis Function por apresentarem, na maioria do período de estudo, o menor valor e o maior valor, respectivamente;
- 4) A Raiz do Erro Quadrático Médio (RMSE) obteve comportamento similar para os métodos em comparação ao MAE, tendo tido apenas valores maiores;
- 5) O Índice de Concordância obteve valores altos para os três métodos, excetuando os meses de Janeiro, Fevereiro e Março, onde se observa valores mais baixos para os três algoritmos. Nota-se também, que os métodos Kriging e Minimum Curvature obtiveram valores, praticamente, idênticos para os períodos estudados;

Ainda em relação ao Erro Médio Absoluto (MAE), construiu-se um gráfico (Figura 11) para analisar a variação sazonal deste índice para os três algoritmos em avaliação.



**Fig. 11** – Variação sazonal do Erro Médio Absoluto para os algoritmos Kriging, Minimum Curvature e Radial Basis Function.

A partir da Figura 11 nota-se que, além do que já foi dito anteriormente, os três métodos obtiveram um melhor e um pior desempenho em Setembro e Julho, respectivamente.

A partir das constatações feitas acima, pode-se notar que os algoritmos obtiveram um bom desempenho. Além disso, percebe-se que existe uma diferença no desempenho de cada método na escala temporal, o que pode ser constatado em todos os meses analisados, tanto para os valores de erros gerados quanto para valores de IQR.

Deve-se ter em mente que o número de estações utilizadas é muito baixo em relação à área do território brasileiro, sendo este fator um limitador para os algoritmos, já que Campos e Ruivo (não publicado) demonstraram em um estudo com esta mesma finalidade orientado para o preenchimento do campo de temperatura da superfície do mar, um desempenho superior do algoritmo Radial Basis Function em relação ao algoritmo Minimum Curvature, utilizando mais de 625 mil pontos de dados observados (dados oriundos de satélite).

## 5. Conclusões

O presente trabalho teve como objetivo comparar a eficácia de diferentes métodos de interpolação para a estimativa do campo de temperatura média do ar para o território brasileiro. Os métodos escolhidos estão presentes no software Surfer<sup>®</sup>, a saber: Kriging, Minimum Curvature e Radial Basis Function. Foram utilizados dados oriundos das Normais Climatológicas de 1961 a 1990 para todos os meses do ano, além da média anual. Para a validação dos métodos, utilizaram-se dados de aeroportos presentes no território brasileiro, para o mesmo período de estudo.

A análise dos dados revelou que todos os métodos de interpolação geraram valores fora da faixa de imprecisão adotada, destacando-se a subestimação

de todos os métodos. Os métodos Kriging e Minimum Curvature obtiveram desempenho semelhante em quase todo o período de estudo.

Destaca-se que, para o mês de Julho, todos os algoritmos obtiveram um desempenho inferior comparado aos outros meses do ano.

Percebe-se, também, que o algoritmo Radial Basis Function obteve um comportamento semelhante aos outros métodos, excetuando-se os meses de Janeiro, Maio e Julho, sendo assim necessário um estudo mais aprofundado deste algoritmo para saber seu comportamento ao longo do ano e com um volume maior de dados de entrada.

Portanto, conclui-se que, dentre os métodos estudados, os mais indicados para geração do campo de temperatura média do ar no território brasileiro são os algoritmos Kriging e Minimum Curvature.

## 6. Bibliografia

Burrough, P. A. 1986. **Principals of geographical information system for land resources assessment**. Oxford, Calderon Press, 194p.

Golden Software Inc. 2009. **Surfer User's Guide**. Colorado, USA, 619p.

Hartkamp, A. D.; De Beurs, K.; Stein, A.; White. J. W. 1999. **Interpolation techniques for climate variables**. NRG-GIS Series 99-01. Mexico, D. F.: CIMMTY, 34p.

Mazzini, P. L. F.; Schettini, C. A. F. **Avaliação de metodologias de interpolação espacial aplicadas a dados hidrográficos costeiros quase-sinóticos**. Braz. J. Aquat. Sci. Technol., 2009, p53-64.

Oliveira, M. S. de. 1991. **Planos amostrais para variáveis espaciais utilizando geoestatística**. Tese de mestrado – Instituto de Matemática, Estatística e Ciência da Computação – UNICAMP, Campinas, 100p.

Willmott, C. J. 1984. **On the evaluation of model performance in physical geography**. In Spatial Statistics Models, Ed. L. Gaile, and C. J. Willmott, p443-460.