Revista Brasileira de Climatologia
Brazilian Journal of Climatology
OPEN ACCESS
ISSN 2237-8642

ESTIMATIVA DE TEMPERATURA MÉDIA PARA O ESTADO DA BAHIA

Average temperature estimate for the state of Bahia

Estimación de la temperatura media del estado de Bahía

Jessica Almeida Monteiro Arruda
Universidade Estadual de Feira de Santana
jamarruda@uefs.br

Rosângela Leal Santos
Universidade Estadual de Feira de Santana
rosaleal@uefs.br

Resumo: Este trabalho objetivou desenvolver um modelo matemático considerando variáveis temporais e espaciais capaz de estimar da temperatura média do ar no Estado da Bahia. O modelo desenvolvido leva em consideração a série de Fourier incompleta com uso dos três primeiros harmônicos e combinação linear dos parâmetros altitude, latitude e longitude. A regressão linear múltipla foi utilizada para ajustar os dados das 26 estações meteorológicas estudadas. Para testar a confiabilidade do modelo foi realizada uma correlação entre os valores estimados e a temperatura observada em cada estação. A análise dos dados foi feita mediante os valores encontrados em comparação com os de referência e a confecção de gráficos com auxílio de software estatístico e de planilha eletrônica. Ao final foram desenvolvidos 6 modelos utilizando a divisão geográfica de mesorregiões, pois, devido as enormes diferenças de altitude, que é o parâmetro de maior significância para o modelo estatístico, não foi possível definir um único modelo para o Estado.

Palavras-chave: Estatística Aplicada. Análise de dados. Séries Infinitas.

Abstract: This study aimed to develop a mathematical model incorporating temporal and spatial variables to estimate mean air temperature in the State of Bahia. The developed model considers an incomplete Fourier series using the first three harmonics combined with linear parameters of altitude, latitude, and longitude. Multiple linear regression was applied to adjust data from the 26 meteorological stations analyzed. To assess model reliability, a correlation analysis was performed between estimated values and observed temperatures at each station. Data analysis was conducted by comparing modeled values with reference data and generating graphs using statistical software and spreadsheet tools. Ultimately, six regional models were developed based on mesoregion geographical divisions, as the state's significant altitudinal variations—the most statistically significant parameter—precluded the use of a single unified model for the entire territory.

Keywords: Applied statistics. Data analysis. Infinite series.

Resumen: Este estudio tuvo como objetivo desarrollar un modelo matemático que considere variables temporales y espaciales para estimar la temperatura media del aire en el Estado de Bahía. El modelo desarrollado incorpora una serie de Fourier incompleta utilizando los tres primeros armónicos, combinados linealmente con los parámetros de altitud, latitud y longitud. Se empleó regresión lineal múltiple para ajustar los datos de las 26 estaciones meteorológicas analizadas. Para evaluar la confiabilidad del modelo, se realizó un análisis de correlación entre los valores estimados y las temperaturas observadas en cada estación. El procesamiento de datos incluyó la comparación de los valores modelados con datos de referencia y la generación de gráficos mediante software estadístico y hojas de cálculo. Finalmente, se desarrollaron seis modelos regionales basados en divisiones geográficas por mesorregiones, ya que las marcadas variaciones altitudinales—el parámetro estadísticamente más significativo—impidieron definir un único modelo para todo el territorio.

Palabras clave: Estadística aplicada. Análisis de datos. Series infinitas.

Submetido em: 17/01/2024
Aceito para publicação em: 14/04/2025
Publicado em: 22/04/2025

1. INTRODUÇÃO

De acordo com Capuchinho et al. (2019) a temperatura do ar é o fator climático responsável por exercer influência em várias atividades que movimentam a economia, sobretudo em setores como engenharia, produção de energia, economia e agropecuária (MEDEIROS; HOLANDA, 2019). Pode-se ainda afirmar, segundo Medeiros et al. (2005) que a temperatura do ar é o fator climático que possui maior influência sobre os processos fisiológicos dos seres vivos. A variação da temperatura do ar, nas escalas espacial e temporal, leva em consideração vários fatores como: latitude, altitude, continentalidade e maritimidade que são parâmetros de análise do clima de uma determinada região (GOMES et al., 2014).

Atualmente, a temperatura do ar é monitorada por estações meteorológicas convencionais ou automáticas, as quais realizam medições diárias. No entanto, a dispersão espacial das estações e lacunas temporais frequentemente resultam em falhas nos registros de dados (FILGUEIRAS *et al*., 2016). No Brasil, a rede do Instituto Nacional de Meteorologia (INMET) enfrenta desafios significativos, incluindo a escassez de séries históricas contínuas e uma cobertura territorial insuficiente. Diante dessas limitações, torna-se essencial aplicar técnicas estatísticas para o ajuste de dados, como a análise de regressão linear múltipla, que permite estimar valores ausentes com base em parâmetros geográficos (latitude, longitude e altitude), garantindo assim maior confiabilidade às séries climáticas (PEDRO JÚNIOR *et al.,* 1994). Este estudo busca [objetivo da pesquisa], contribuindo para o avanço metodológico na reconstrução de dados meteorológicos em regiões com escassez de monitoramento.

A escassez de dados observacionais de temperatura do ar em escala nacional tem impulsionado o desenvolvimento de metodologias para reconstrução e preenchimento de lacunas em séries meteorológicas. Dentre essas técnicas, destaca-se a análise de regressão múltipla, amplamente empregada por sua capacidade de correlacionar variáveis termométricas com parâmetros fisiográficos espaciais ou temporais de forma isolada (ANTONINI *et al.,* 2019). Essa abordagem estatística tem se mostrado fundamental para gerar produtos climáticos contínuos, especialmente em regiões com monitoramento insuficiente, como observado no território brasileiro.

Diante da escassez de dados meteorológicos confiáveis, este estudo propõe sete modelos matemáticos distintos para estimar a temperatura média mensal do ar no estado da Bahia, estratificados por mesorregiões. Essa abordagem multirregional justifica-se pela

significativa variabilidade altimétrica do território baiano, a qual impede a aplicação de um modelo único sem comprometer a precisão das estimativas. A segmentação por mesorregiões visa, portanto, reduzir os erros sistemáticos associados às particularidades topográficas locais, garantindo maior robustez aos resultados.

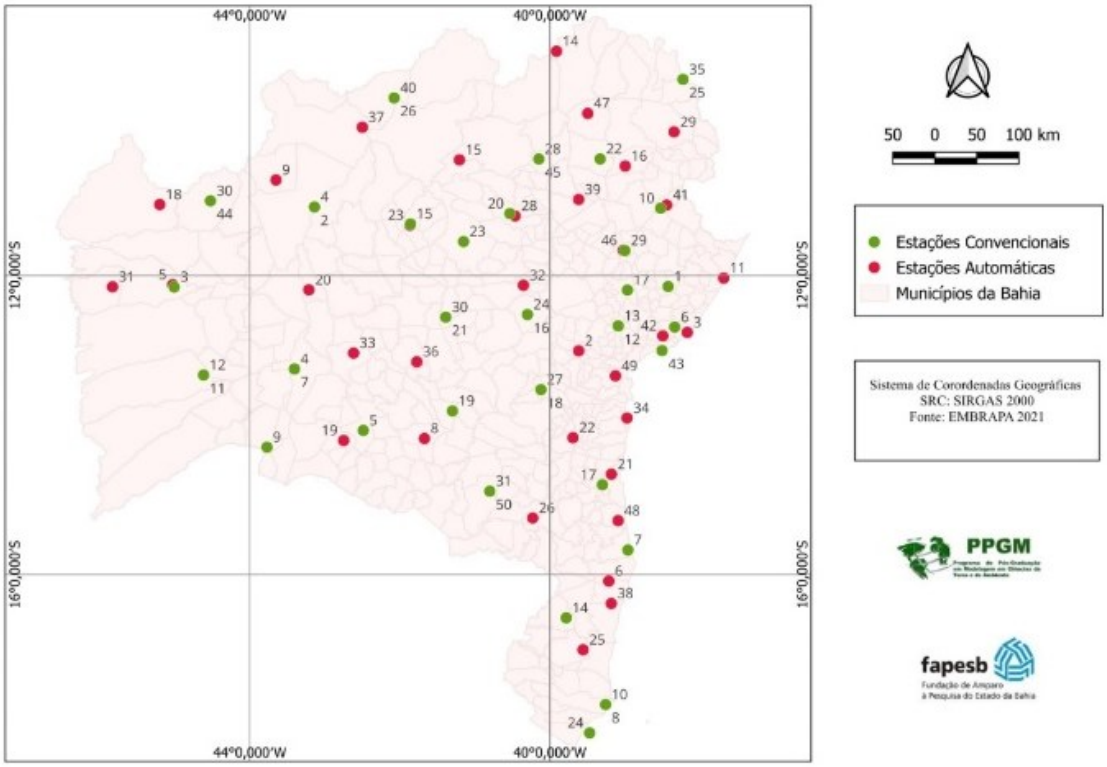
2. METODOLOGIA

2.1. Área de estudo

A área de abrangência desse estudo é todo o Estado da Bahia. Foram utilizados dados das 26 estações meteorológicas ativas no estado. Os dados das séries históricas foram obtidos no site do INSTITUTO NACIONAL DE METEOROLOGIA - INMET.

A Figura 1 ilustra a distribuição espacial das estações meteorológicas utilizadas neste estudo, evidenciando a disparidade na cobertura observacional ao longo do território baiano. Nota-se uma concentração irregular das estações, com amplas áreas desprovidas de monitoramento, o que limita a representatividade espacial dos dados. Para contornar essa limitação, incorporou-se um Modelo Digital de Elevação (MDE) do estado, derivado de cartas topográficas disponibilizadas pelo IBGE, visando caracterizar a influência do relevo na variabilidade térmica regional.

Figura 1 - Localização da área de estudo e localização espacial das estações meteorológicas ativas utilizadas.



Fonte: Elaborado pelos autores (2023).

2.2. Dados Meteorológicos

Para o desenvolvimento desta pesquisa, foram obtidos através do site do Instituto Nacional de Meteorologia - INMET os dados de temperatura média mensal de 50 estações automáticas e 31 estações convencionais disponíveis para o estado da Bahia, no período de 1999 a 2019, totalizando 20 anos de dados.

Após a coleta e análise dos dados das estações meteorológicas, foi possível identificar e descartar as estações cujos dados não eram suficientes para esta pesquisa, e os motivos de exclusão foram: séries com mais de 8 anos sem dados contínuos ou vazios de mais de dois anos entre os anos. Então, na Tabela 1 estão dispostas as estações que foram efetivamente utilizadas para este trabalho, cuja dados de latitude, longitude e altitude de cada estão dispostas sendo que C representa as estações convencionais e A as estações automáticas.

Tabela 1 - Estações Meteorológicas automáticas e convencionais utilizadas.

N°	Estação	Latitude	Longitude	Altitude	Tipo
1	Alagoinhas	-12,15	-38,43	129,78	C
2	Barra	-11,08	-43,14	407,50	A
3	Barreiras	-12,12	-45,03	474,17	A
4	Bom Jesus da Lapa	-13,25	-43,41	447,75	C
5	Caitité	-14,07	-42,49	876,29	C
6	Caravelas	-17,74	-39,26	605,00	C
7	Caranhanha	-14,30	-43,77	455,25	C
8	Cipó	-11,10	-38,52	143,53	C
9	Correntina	-13,33	-44,62	551,71	C
10	Cruz das Almas	-12,68	-39,09	219,69	C
11	Feira de Santana	-12,20	-38,97	229,52	C
12	Guaratinga	-16,58	-39,78	197,88	C
13	Irecê	-11,31	-41,86	747,22	C
14	Itaberaba	-12,52	-40,30	250,11	C
15	Itiruçu	-13,53	-40,12	757,42	A
16	Jacobina	-11,17	-40,53	486,37	C
17	Lençóis	-12,56	-41,39	438,09	C
18	Monte Santo	-10,44	-39,33	465,84	C
19	Morro do Chapéu	-11,55	-41,15	1001,66	C
20	Paulo Afonso	-9,38	-38,23	255,04	C
21	Remanso	-9,63	-42,08	397,40	C
22	Salvador (Ondina)	-13,01	-38,51	47,35	C
23	Senhor do Bonfim	-10,44	-40,15	532,09	C
24	Serrinha	-11,67	-39,00	362,39	C
25	Sta. R de Cássia (Ibipetuba)	-11,00	-44,52	450,01	C
26	Vitória da Conquista	-14,89	-40,80	879,47	C

Fonte: Elaborado pelos autores (2023).

Portanto, foram utilizadas 26 estações meteorológicas, sendo 23 convencionais (C) e 3 automáticas (A).

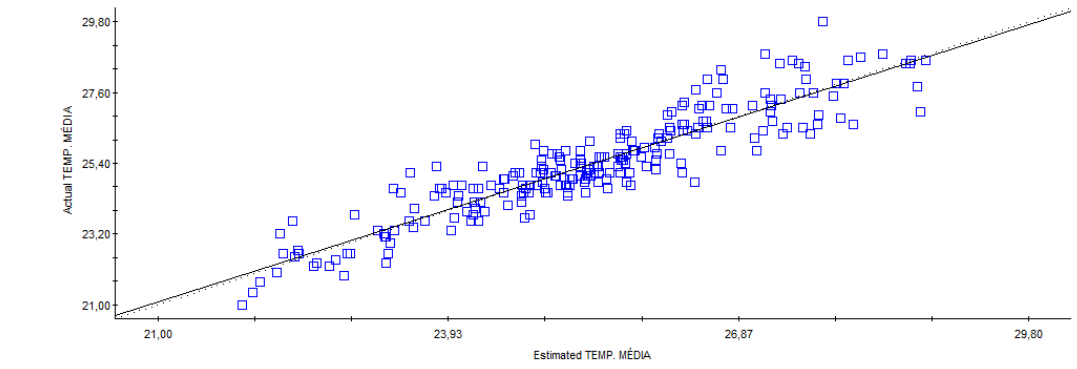
2.3. Preenchimento de falhas

Para preenchimento de dados das séries históricas foi utilizado o software GS+ utilizando krigagem, e como resultado obteve-se uma tabela de temperaturas mantendo as conhecidas, e extraindo o valor interpolado para os locais com falhas.

Assim, os dados de entradas foram divididos em X, Y e Z que corresponde ao mês, ano e valor da temperatura média mensal referente a este mês, deste ano, respectivamente. Após a primeira iteração foi calculado o valor do IDE, Índice de Dependência Espacial, e logo em seguida avaliado e observado, a depender do valor obtido, elimina-se os valores discrepantes e gera-se uma nova iteração até obter um valor de IDE máximo para aquele conjunto de dados.

Na Figura 2 pode-se observar o gráfico de dispersão dos valores da temperatura média mensal de Barreiras, que representou melhor ajuste e melhor valor de IDE de 0,951.

Figura 2- Gráfico de dispersão dos valores observados e estimados de Barreiras.



Fonte: Elaborado pelos autores (2023).

Após os preenchimentos de falhas dos dados das 26 estações meteorológicas selecionadas, estes dados foram unificados em uma planilha com um total 6553 linhas e 10 colunas, para determinação do modelo de regressão.

2.4. Modelo de Regressão

Para compor o modelo de estimativa da temperatura média diária do ar, em função do espaço e do tempo, utilizou-se como dados da variável dependente, os valores médios da temperatura média diária, para cada mês do ano (To_{ip}) do período de observação das séries históricas consideradas. Foram utilizadas como variáveis independentes a altitude (Alt), latitude (Lat), longitude (Long) e a sequência dos meses do ano (t_i). O modelo será desenvolvido por meio de uma combinação linear da Alt, Lat e Long com a série trigonométrica de Fourier incompleta com os três primeiros harmônicos:

$$(1) \qquad Tm_i = \beta_0 + \beta_1 \mathrm{Alt} + \beta_2 \mathrm{Lat} + \beta_3 \mathrm{Long} + \sum_{n=1}^{3}\left[a_n \cos\left(n \cdot w \cdot t_i\right) + b_n \mathrm{sen}\left(n \cdot w \cdot t_i\right)\right]$$

em que: o Tm_i é a temperatura média do ar do mês i (°C); i é o i-ésimo mês do ano; β_0 é o coeficiente linear, e β_1, β_2 e β_3 são os coeficientes angulares das variáveis espaciais; Alt é a altitude do local (m); Lat é a latitude do local em valores positivos (graus decimais); Long é a longitude do local em valores positivos (graus decimais); a_n e b_n são os coeficientes dos termos da série trigonométrica de Fourier incompleta; t_i é o mês do ano da observação i, que varia de 1 a 12 meses; *n* é o número de harmônicos da série trigonométrica de Fourier; *w* é a frequência angular fundamental igual a 2π/12 em rad/s. Com os coeficientes angulares a_n e b_n da série trigonométrica de Fourier, pode-se obter a amplitude de cada harmônica, igual a $\left(a_n^2 + b_n^2\right)^{1/2}$, que representam as amplitudes anual, semestral e trimestral da temperatura média mensal do ar.

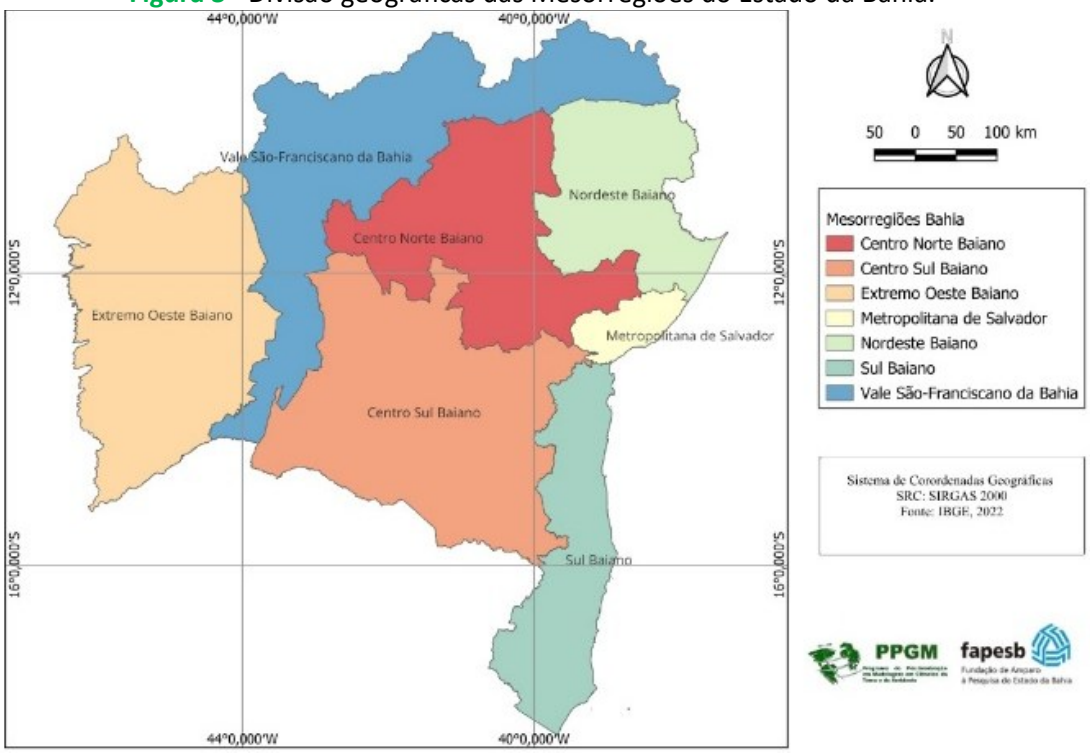
Utilizou-se séries históricas de dados de temperatura média mensal do ar com o mínimo de oito anos de observações, de períodos não uniformizados, para demonstrar a aplicabilidade do modelo proposto no trabalho, as 26 estações meteorológicas na Tabela 1.

Os valores das temperaturas obtidas totalizaram 6552 dados, formando uma matriz de 6553 linhas e dez colunas que são correspondentes a: temperatura média mensal estimada, altitude, latitude, longitude e os valores dos componentes da série de Fourier nos três primeiros harmônicos: cos(*wt*), cos(*2wt*), cos(*3wt*), sen(*wt*), sen(*2wt*) e sen(*3wt*).

Em seguida, subdividiu-se o modelo geral em modelos menores de modo que atendesse a classificação das Mesorregiões do Estado da Bahia, cuja Figura 3 representa as

distribuições geográficas delas, obtendo assim, um modelo de regressão para cada Mesorregião, uma divisão político-administrativa existente, mas que por apresentar locais com pequenas diferenças de altitude, tornou-se ideal para a proposta desse trabalho. A obtenção dos coeficientes do modelo pela análise de Regressão Linear Múltipla foi obtida através do software RStudio em acompanhamento do software de Planilha Eletrônica, Excel.

Figura 3 - Divisão geográficas das Mesorregiões do Estado da Bahia.



Fonte: Elaborado pelos autores (2023).

2.5. Testes de Validação

O software de Planilha Eletrônica Excel foi utilizado para efetuar a determinação dos coeficientes do modelo pela análise de Regressão Linear Múltipla. O grau de ajuste do modelo foi avaliado pelo coeficiente de determinação, e a verificação da significância dos coeficientes pelo teste t de Student, a 5% de probabilidade de erro. Para avaliar o grau de desvio do modelo, em relação à linha 1:1, foi utilizado o recurso da Regressão Linear Simples, aplicado aos pares de valores estimados equação (1) e observados. A quantificação das diferenças entre os valores de temperatura média diária do ar, estimados pelo modelo equação (1), e os valores observados nas séries históricas das estações meteorológicas estudadas foi feita com os seguintes índices estatísticos: média dos erros (MBE), variância dos erros (s^2), raiz do

quadrado médio dos erros (RMSE) e média dos erros absolutos (MAE) definidos, respectivamente, pelas seguintes equações: $s^2 = \left[\sum_{i=1}^{N}\left(Tm_i - To_{ip} - MBE\right)^2\right] / \left(N-1\right)$

$$MAE = \sum_{i=1}^{N}\left|Tm_i - To_{ip}\right| / N \text{ - } MBE = \sum_{i=1}^{N}\left(Tm_i - To_{ip}\right) / N \text{ - } RMSE = \sqrt{\sum_{i=1}^{N}\left(Tm_i - To_{ip}\right)^2 / N}$$

A qualidade preditiva do modelo será avaliada quanto à precisão, exatidão e desempenho. A precisão será quantificada pelo coeficiente de correlação de Pearson (r), calculado por meio da seguinte equação: $r = \sqrt{\sum_{i=1}^{N}\left(Tm_i - \bar{O}\right)^2 / \left[\sum_{i=1}^{N}\left(Tm_i - \bar{O}\right)^2 + \sum_{i=1}^{N}\left(To_{ip} - Tm_i\right)^2\right]}$.

em que: N é o número total de observações; e $\bar{O}$ é a média dos valores de temperatura média diária observados (°C).

A exatidão será quantificada pelo índice de concordância (d) de Willmott (Willmott, 1982), calculado por meio da equação: $d = 1 - \left[\sum_{i=1}^{N}\left(Tm_i - \bar{O}\right)^2 / \sum_{i=1}^{N}\left(\left|Tm_i - \bar{O}\right| + \left|To_{ip} + \bar{O}\right|\right)^2\right]$.

O desempenho do modelo foi quantificado pelo índice de desempenho c = r×d (Camargo & Camargo, 2000). Os indicadores de concordância e de desempenho foram interpretados de acordo com Camargo & Sentelha (1997).

3. DESENVOLVIMENTO

Foi realizado um modelo classificado como geral, sendo este um único modelo para todo o Estado da Bahia, mesmo com as grandes diferenças de altitude de uma estação para outra. Após a confecção do modelo verificou-se que o índice de consistência do modelo foi baixo, desta forma, foi feita uma subdivisão da Bahia.

O estado da Bahia foi divido geograficamente pelo Instituto Brasileiro de Geografia e Estatística (IBGE) em sete mesorregiões. Após a confecção do modelo foi constatado que o índice de desempenho melhora quando analisado por cada mesorregião.

3.1. Modelo Geral

3.1.1. Dados Observados e Montagem do Modelo

Após realização do ajuste do modelo foi realizado através de análise de regressão linear múltipla, os dez coeficientes significativos foram obtidos e estão disposto na Tabela 2.

Tabela 2 - Valores dos coeficientes da RLM.

Coeficientes do modelo	Valor Obtido	Coeficientes do modelo	Valor Obtido
β0	12,64	**a2**	-0,11
β1	-0,01	**a3**	-0,01
β2	-0,38	**b1**	0,69
β3	0,47	**b2**	0,2
a1	0,43	**b3**	-0,06

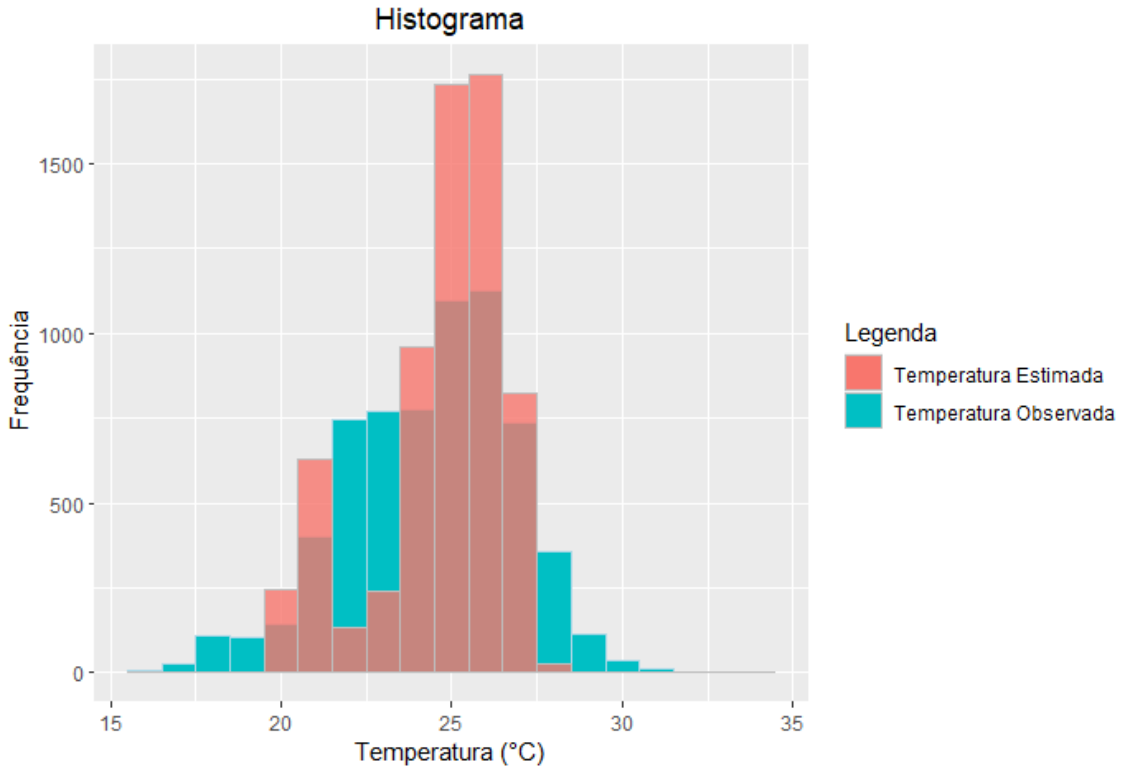
Fonte: Elaborado pelos autores (2023).

Ambos com um erro padrão associado próximos de zero, indicando que os componentes se fazem necessários na representação da variação espaço-temporal da temperatura média mensal do ar no Estado de Bahia.

Analisando a Tabela 2 no que diz respeito da Altitude, o coeficiente β1 indica que a Tmi varia 0.63°C a cada 100 m de variação positiva em altitude, assim, com o aumento da altitude a Tmi diminui. No caso da latitude a cada aumento de 1° há uma diminuição de aproximadamente 0.38 °C na temperatura do ar. Já em relação a longitude observou-se um acréscimo de 0.47 °C na temperatura aproximadamente para cada aumento de 1°.

Realizou-se a elaboração da Figura 4, na qual é possível observar através do histograma, a distribuição dos dados de temperatura observada e de temperatura estimada.

Figura 4 - Histograma da distribuição de temperatura.



Fonte: Elaborado pelos autores (2023).

Observa-se que há uma grande discrepância entre os valores que foram estimados e os observados, indicando uma alta variabilidade nos dados estimados e observados.

3.1.2. Avaliação do Modelo

O grau de correlação do modelo é a diferença entre o par: valor observado-valor estimado e mostra a dispersão das estimativas no que tange a regressão linear, porém não é através deste coeficiente que é medido o grau de confiabilidade do modelo, pois a magnitude do desvio entre o valor observado e o estimado não está necessariamente relacionado pelo modelo de regressão. Isso posto, quanto mais próximo de 0 for o coeficiente linear (termo independente) e quanto mais próximo de 1 for o coeficiente angular menor é o desvio entre o par de valores observado: estimado.

Ao gerar a regressão deste modelo geral foram obtidos valores para coeficiente angular e linear estão dispostos na Tabela 3, representados por Intercept e temp_est,

respectivamente. A significância é representada pelo p-value, que representa a probabilidade que a variável não seja relevante para o modelo, está na Tabela 3 como p-value (Pr(>F)) e tem um valor menor que 0,05, ou seja, a relação entre as duas variáveis é significativa.

Tabela 3 - Anova entre os dados observados e estimados.

ANOVA TEST (Coeficientes)				
	Estimate	Std. Error	t value	Pr(>|t|)
Intercept	0,89734	0,25645	3,499	0,00047
temp_est	0,9529	0,01306	92,001	<2e-16
Multiple R-Squared		0,5637		
Adjusted R-Squared		0,5637		
p-value		<2.2e-16		

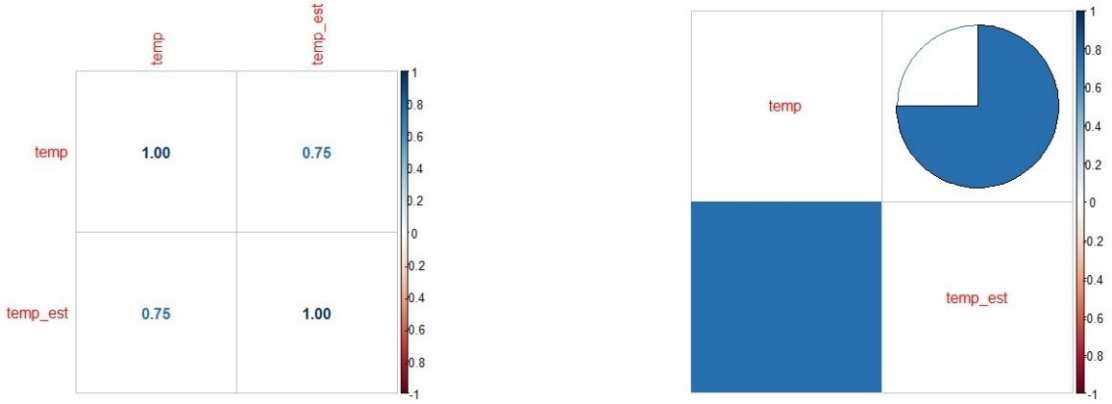
Fonte: Elaborado pelos autores (2023).

O valor do coeficiente Adjusted R-Squared indica o quanto o modelo explica a variância da temperatura estimada a partir do regressor incluso no modelo, neste caso, o quanto a Temperatura Estimada pode ser justificada pela Temperatura Obtida. O modelo linear explica apenas 56,37% da temperatura estimada, indicando que o modelo tem uma variação de 43,63% decorrente de outras variáveis, ou seja, que há uma grande variabilidade dos valores.

Outro coeficiente a ser observado na Tabela 3 é o F-Statistic, que indica o quanto o modelo criado é diferente do modelo nulo, ou seja, um modelo de regressão onde a variável preditora é a medida da variável de resultado, Temperatura Obtida e Temperatura Estimada, respectivamente. Como o valor de F é maior que 0,05 significa que o modelo criado não é bom, e a regressão para estimativa da temperatura não é confiável.

Quando se avalia a Linearidade das variáveis, esta indica o grau de associação entre ambas e é representada pelo coeficiente de correlação de Pearson. Na Figura 5 tem-se os dados obtidos entre os valores observados e estimados.

Figura 5 - Linearidade entre os coeficientes estimados e observados



Fonte: Elaborado pelos autores (2023).

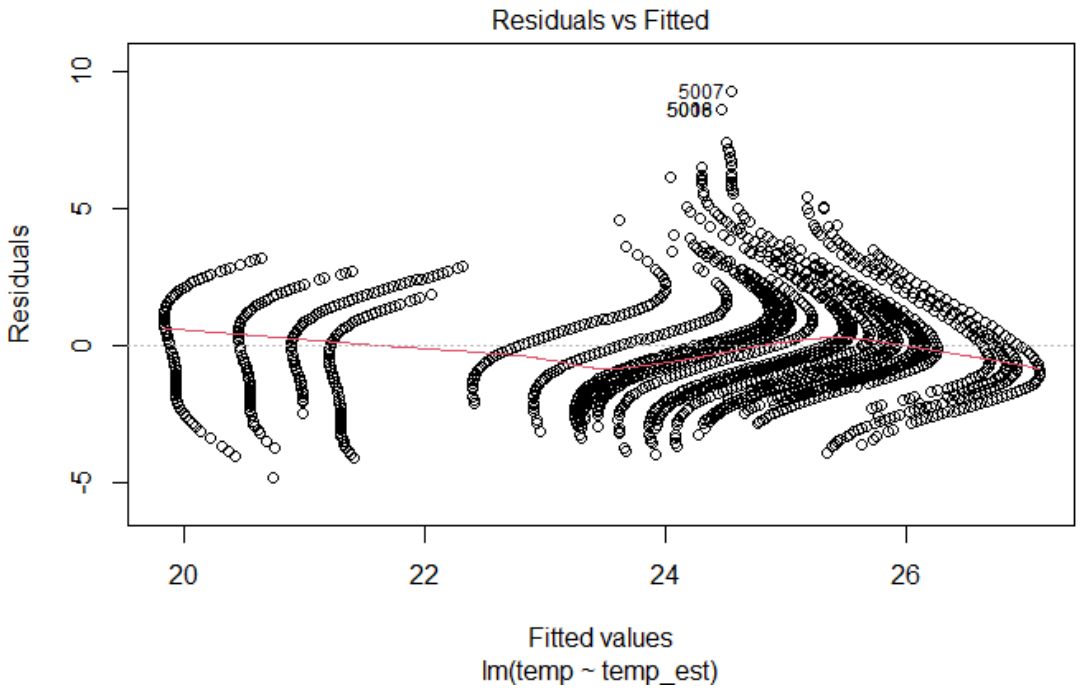
Convencionou-se que para verificar a correlação das variáveis, a regra abaixo é utilizada para avaliar os coeficientes de correlação:

- 0.9 para mais ou para menos indica uma correlação muito forte.
- 0.7 a 0.9 positivo ou negativo indica uma correlação forte.
- 0.5 a 0.7 positivo ou negativo indica uma correlação moderada.
- 0.3 a 0.5 positivo ou negativo indica uma correlação fraca.
- 0 a 0.3 positivo ou negativo indica uma correlação desprezível.

Neste caso, como o coeficiente é de 0.7508297, aproximadamente, 0.75 representado na Figura 3, existe uma correlação forte entre a variável estimada e a observada.

Na Figura 6 é possível observar a homocedacidade dos resíduos, ou seja, a hipótese da variância constante com os valores previsto e dos resíduos em cada previsão, não apresentando de fato nenhum padrão ou tendência como preconiza a literatura.

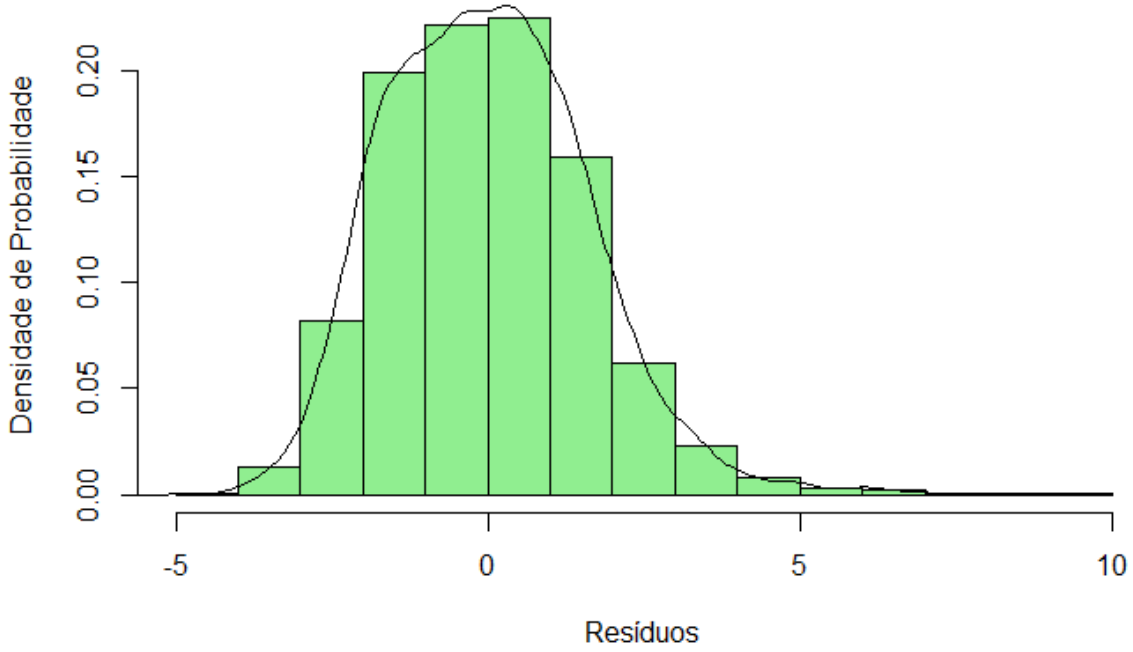
Figura 6 - Homocedacidade dos resíduos.



Fonte: Elaborado pelos autores (2023).

Analisou-se também a normalidade dos resíduos que em um modelo de regressão linear deve apresentar normalidade, ou seja, a distribuição dos resíduos deve ser aproximar de uma distribuição normal. Pela Figura 7 é possível visualizar que os resíduos não indicam nenhuma assimetria e possuem a aparência de uma curva normal.

Figura 7- Normalidade dos resíduos.



Fonte: Elaborado pelos autores (2023).

3.1.3. Validação dos dados do Modelo

Os índices calculados para a validação dos dados do modelo de regressão linear criado foram: média dos erros (MBE), variância dos erros (s^2), raiz do quadrado médio dos erros (RMSE) e média dos erros absolutos (MAE) e estão dispostos na Tabela 4.

Tabela 4 - Validação do modelo.

Descrição da medida estatística	Símbolo	Ajuste do Modelo
Média dos erros (°C)	MBE	0,26535
Variância dos erros	s^2	2,60786
Raiz do quadrado médio dos erros (°C)	RMSE	1,63643
Média dos erros absolutos (°C)	MAE	1,32015
Média dos valores observados (°C)	O	24,41960
Média dos valores estimados (°C)	P	24,7
Coeficiente de correlação	r	0,76765
Coeficiente de Wilmott	d	1
Índice de desempenho	c	0,76765
Número de observações	N	6553

Fonte: Elaborado pelos autores (2023).

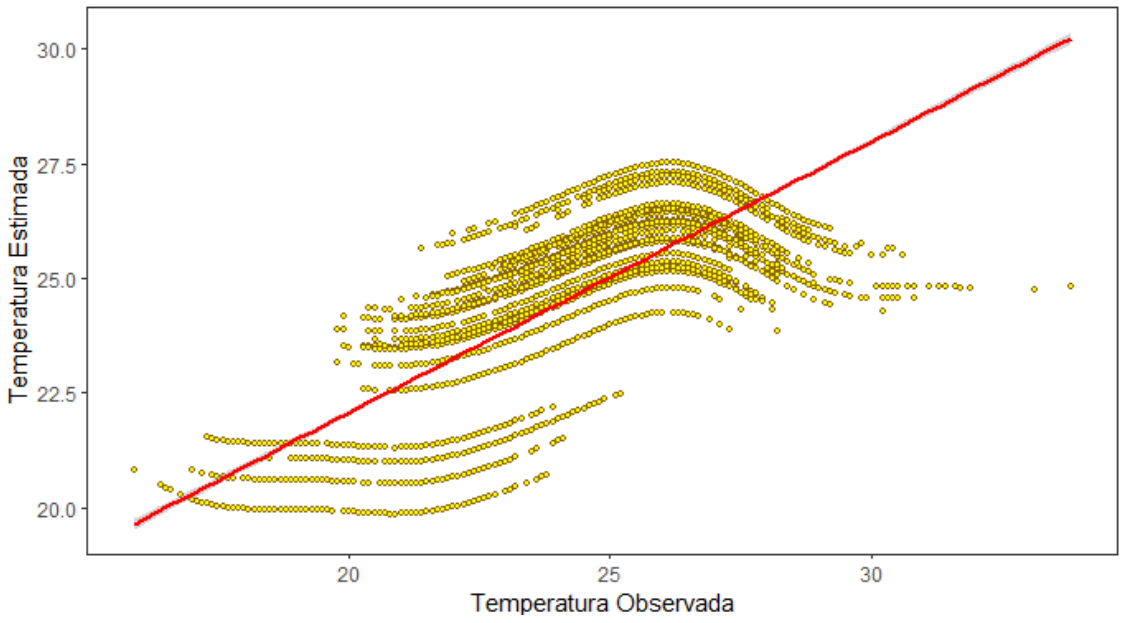
Quando analisamos as diferenças entre os valores de temperatura média mensal estimados e os observados nas séries históricas, a magnitude do erro absoluto médio é de 1.32 °C e a variância dos erros de 2.61, aproximadamente, disposta na Tabela 4, indicando que a dispersão média dos dados foi de quase 3 °C, o que representa uma grande variabilidade.

No que se refere a aplicação desse modelo de ajuste de dados espaciais e climáticos, para todo o Estado da Bahia, resultou em d = 1 e r = 0.77 aproximadamente, mostrando que tem uma forte correlação, indicando uma boa exatidão e precisão do modelo, onde a precisão é representada pelo coeficiente de correlação de Pearson, r, e a exatidão, pelo coeficiente de Wilmont, d. Esses valores demonstram que o modelo foi capaz de estimar a temperatura média mensal do ar com relativa exatidão e precisão.

O desempenho do modelo, medido pelo índice de desempenho, c, foi de 0.77 aproximadamente classificado como muito bom, mesmo com a grande diferença de altitude empregada.

Na Figura 8 é possível observar o gráfico de dispersão das variáveis estimadas e observadas, quanto mais forte for a correlação linear mais o gráfico irá se parecer com uma reta, logo observa-se a dispersão da variável temperatura, e o modelo de regressão linear criado para avaliação desses dados, cujos resultados estão dispostos na Tabela 4.

Figura 8- Gráfico de Dispersão entre as variáveis estimadas e observadas.



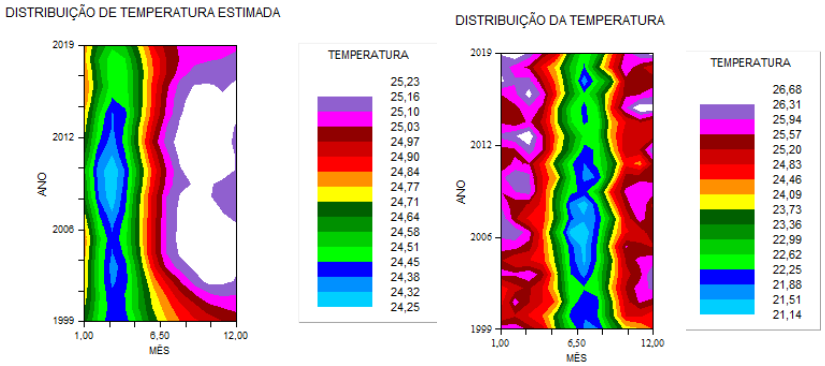
Fonte: Elaborado pelos autores (2023).

O modelo se torna inadequado para fornecer estimativas de temperatura confiáveis e que possam de fato, ser utilizadas, indicando a necessidade de ter um outro modelo capaz de gerar um alto índice de consistência.

3.1.4. Comparativo entre os Modelos

Para uma melhor visualização entre os dados de temperatura estimada e observada elaborou-se o gráfico dispostos na Figura 9 em duas dimensões (2D) e no qual é possível verificar a discrepância que o modelo geral proporcionou na estimativa dos dados inviabilizando, graficamente o modelo.

Figura 9 - Gráfico em 2D das temperaturas estimada e observada da região CNB.



Fonte: Elaborado pelos autores (2023).

Analisando a Figura 9 é possível notar que há uma diferença de 3.11°C entre o valor médio da temperatura mínima em cada caso e de 1.45° C na temperatura máxima, sendo que a temperatura observada apresenta os valores máximos e mínimos, respectivamente de 21.14°C e 26,68 °C em detrimento de 24,25°C e 25,23°C para o modelo estimador.

Fica evidenciando tanto numérica como graficamente que o Modelo Geral não consegue efetuar a estimativa da temperatura com consistência.

3.2. Modelo das Mesorregiões

3.2.1. Dados Observados e Montagem do Modelo

Foi realizado o ajuste do modelo e obtido os coeficientes para montagem da equação de estimativa da Temperatura Média Mensal, ambos com um erro padrão associado próximos de zero, indicando que os componentes se fazem necessários na representação da variação espaço-temporal da temperatura média mensal do ar no Estado de Bahia estudando as mesorregiões.

Na Tabela 5 estão dispostos os coeficientes da Equação (1) para cada mesorregião, sendo Centro Norte Baiano (CNB), Centro Sul Baiano (CSB), Extremo Oeste Baiano (EOB), Nordeste Baiano (NB), Região Metropolitana de Salvador (RMS), Sul Baiano (SB) e Vale São Franciscano (VSF).

Tabela 5 - Valores dos coeficientes do modelo geral e das mesorregiões.

	Parâmetros Estimados do Modelo							
Coeficientes do modelo	**Geral**	**CNB**	**CSB**	**EOB**	**NB**	**RMS**	**SB**	**VSF**
β0	12,6367	21,1765	20,5721	25,0063	51,4898	24,2580	24,0199	27,1876
β1 (alt)	**-0,0064**	**-0,0003**	**-0,0004**	**0,0000**	**0,0017**	**-0,0001**	**0,0000**	**-0,0004**
β2 (lat)	**-0,3790**	**-0,0063**	**-0,0014**	**0,0018**	**-0,0395**	**0,0000**	**0,0000**	**-0,0047**
β3 (long)	**0,4685**	**0,0456**	**0,0378**	**0,0000**	**-0,6959**	**0,0000**	**0,0000**	**0,0131**
a1	0,4294	2,1712	3,1733	-1,7325	-0,8355	-0,4399	-0,0383	-3,5149
a2	-0,1102	2,8424	1,5386	2,9340	3,2231	3,3188	3,3339	-0,9886
a3	-0,0133	-1,4442	-1,1619	1,0387	0,5186	0,2930	0,0290	-0,7159
b1	0,6874	-0,4635	0,9323	-0,5652	-0,9233	-1,0500	-1,0628	1,3378
b2	0,2028	0,6300	-0,1589	-0,3465	-0,1960	-0,1281	-0,0177	0,4530
b3	-0,0591	-0,1080	-0,6095	-0,0215	0,1941	0,2638	0,2731	0,3152

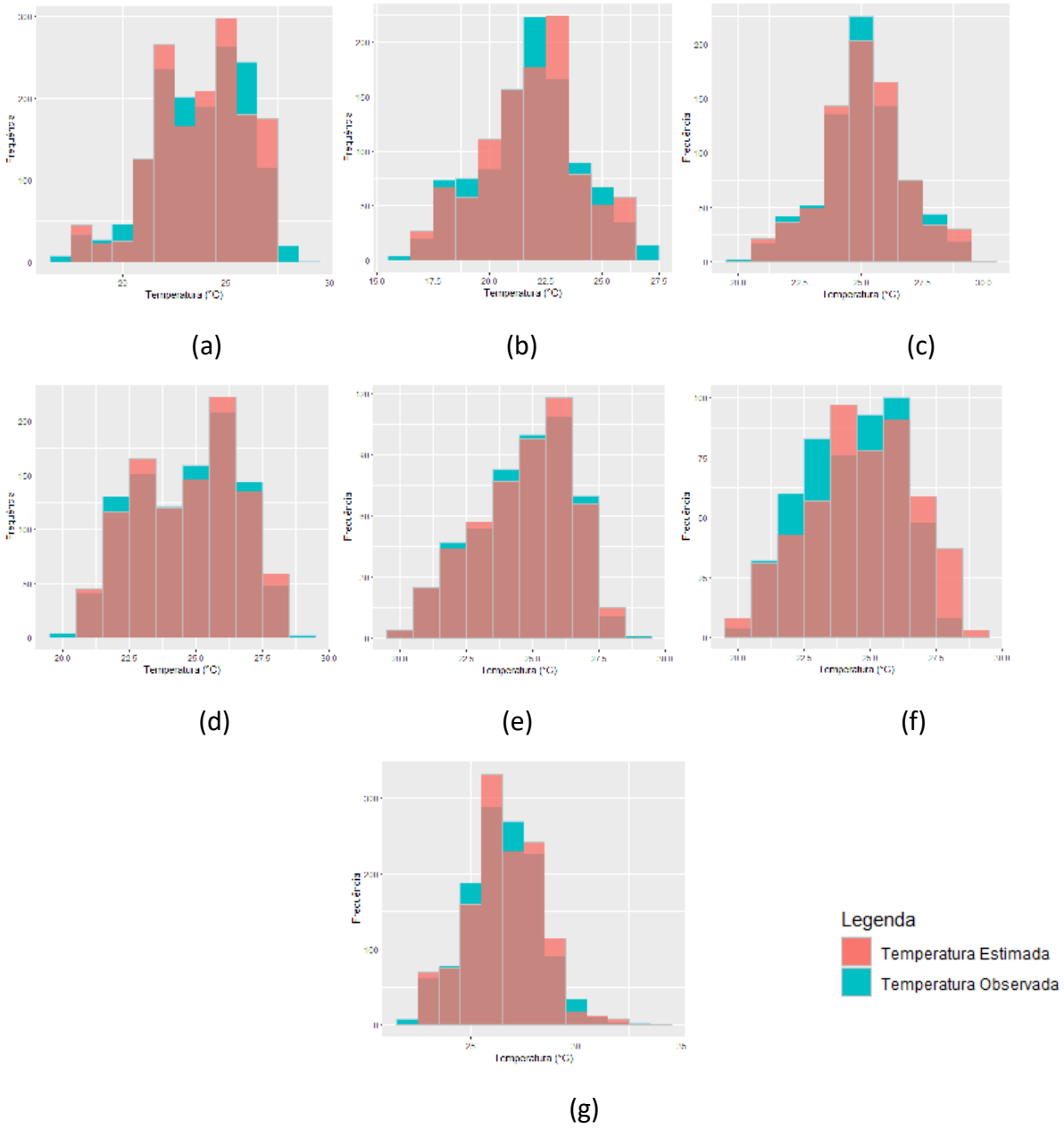
Fonte: Elaborado pelos autores (2023).

Com os coeficientes obtidos na regressão linear se observou que a Altitude, o coeficiente que a acompanha indica que a Tmi varia 0.002 °C até aproximadamente 0.04 °C no EOB e no VSF a cada 100 m de variação positiva em altitude, assim, pode-se inferir que com o aumento da altitude a Tmi diminui. No caso da latitude a cada aumento de 1° há uma diminuição variando de aproximadamente 0.0063 °C até 0°C de variação na temperatura do

ar quando analisados CNB e RMS ou SB, respectivamente. Já em relação a longitude observou-se um acréscimo de 0.7 °C até 0 °C na temperatura para cada aumento de 1°.

Elaborou-se a Figura 10 (a), (b), (c), (d), (e), (f) e (g) em que correspondem as mesorregiões CNB, CSB, EOB, NB, RMS, SB e VSF, respectivamente na qual é possível observar através do histograma, a distribuição dos dados de temperatura observada e estimada.

Figura 10 - Histogramas de distribuição de temperatura.

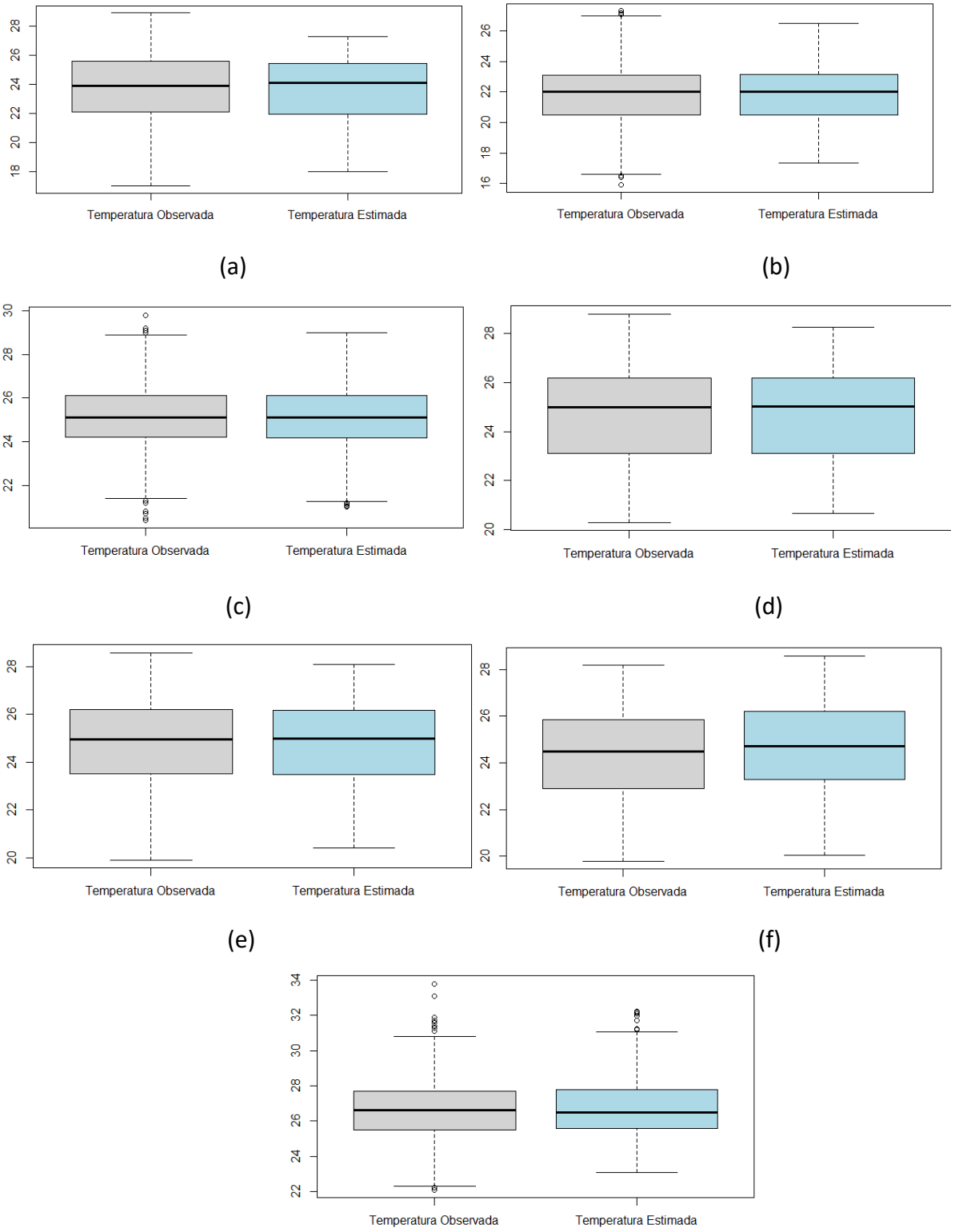


Fonte: Elaborado pelos autores (2023).

Na Figura 11 observamos o comportamento da temperatura obtida, a posição da média desses valores e, os valores máximos e mínimos da temperatura após a estimativa dos

dados de temperatura média mensal para o Estado da Bahia, onde (a), (b), (c), (d), (e), (f) e (g) correspondem as mesorregiões CNB, CSB, EOB, NB, RMS, SB e VSF, respectivamente.

Figura 11 - BoxPlot comparativo da temperatura observada e estimada.



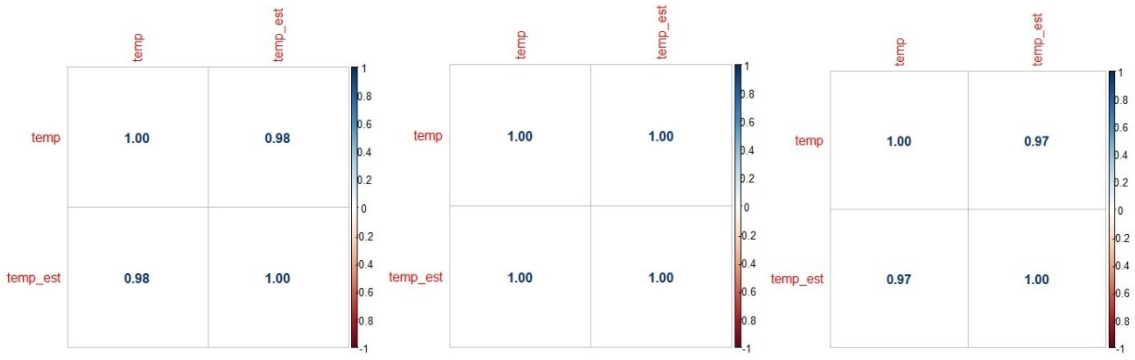
Fonte: Elaborado pelos autores (2023).

Observa-se que há uma pequena discrepância entre os valores que foram estimados, para os valores observados, e que essa diferença não foi tão significativa como no modelo geral, o que pode ser justificado pelo coeficiente R-Squared disposto na Tabela 6.

3.2.2. Avaliação do Modelo

Quando se avalia a Linearidade das variáveis, esta indica o grau de associação entre ambas e é representada pelo coeficiente de correlação de Pearson. Na Figura 12 (a), (b) e (c) tem-se os dados obtidos entre os valores observados e estimados, sendo as mesorregiões CNB, CSB e VSF referente a (a), EOB, NB, RMS a (b) e SB a (c).

Figura 12 - Linearidade entre os coeficientes estimados e observados para as mesorregiões.

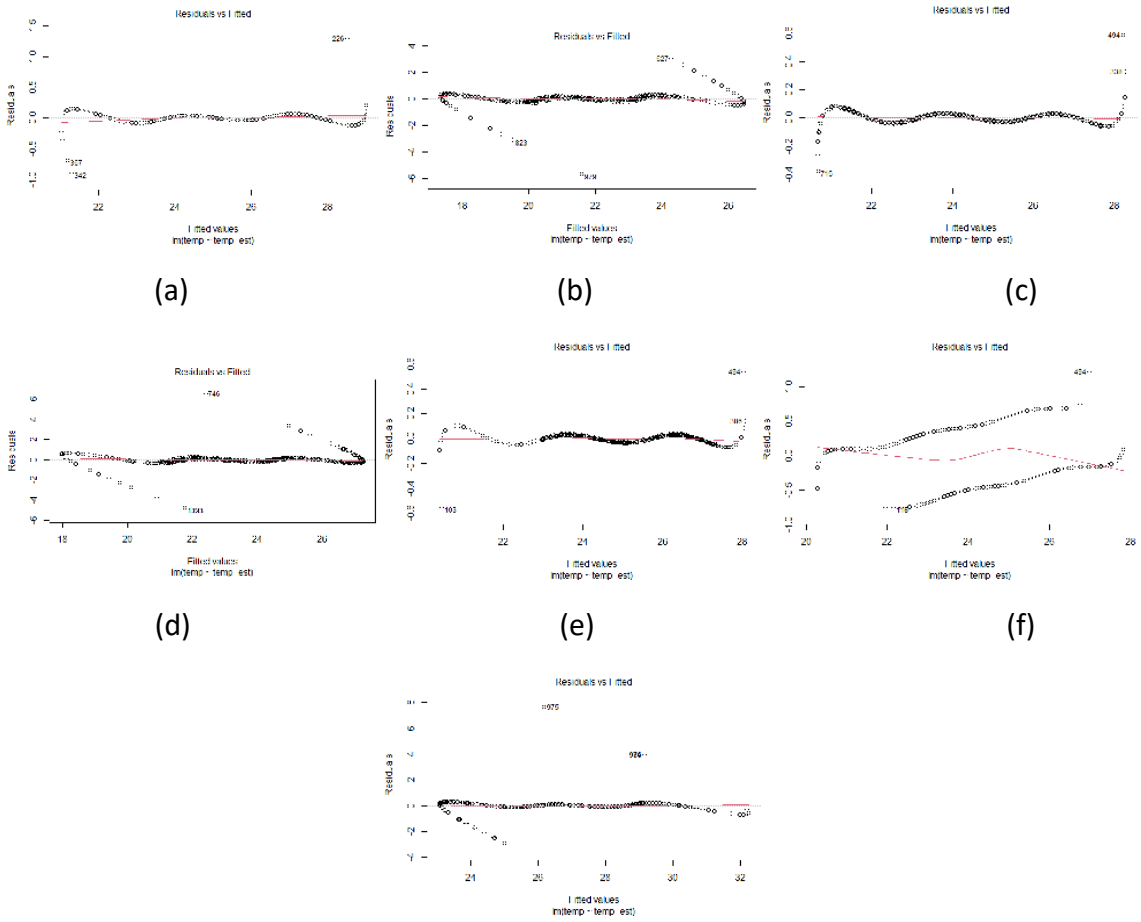


(a) (b) (c)

(b) **Fonte:** Elaborado pelos autores (2023).

Neste caso, conforme Figura 12 todos os coeficientes são maiores que 0.9, existe uma correlação muito forte entre a variável estimada e a observada e a Figura 13 observamos que a homocedacidade dos resíduos, ou seja, a hipótese da variância constante com os valores previsto e dos resíduos em cada previsão, não apresentando de fato nenhum padrão ou tendência como preconiza a literatura, onde (a), (b), (c), (d), (e), (f) e (g) correspondem as mesorregiões CNB, CSB, EOB, NB, RMS, SB e VSF, respectivamente tanto para Figura 13 como para as Figura 14 e 15.

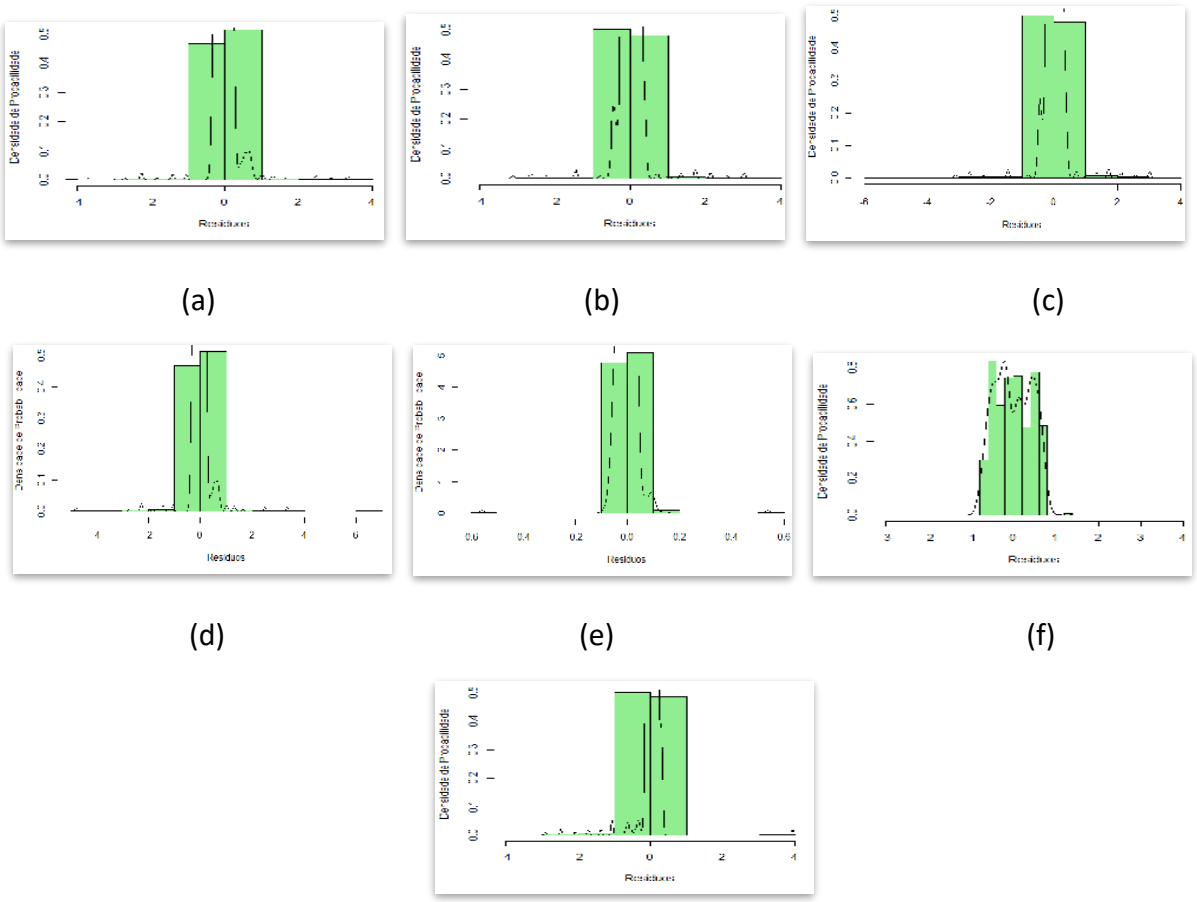
Figura 13- Homocedacidade dos resíduos para as mesorregiões



Fonte: Elaborado pelos autores (2023).

Pela Figura 14 é possível visualizar que os resíduos indicam assimetria e possuem a aparência próxima de uma curva normal, apresentando pequenos desvios em relação a distribuição, apesar dos resíduos apresentarem pequenos desvios na normalidade, não tem influência significativa no modelo de regressão.

Figura 14 - Normalidade dos resíduos das mesorregiões.



(a) (b) (c)

(d) (e) (f)

(g)

Fonte: Elaborado pelos autores (2023).

3.2.3. Validação dos dados do Modelo

Os índices calculados para a validação dos dados dos modelos de regressão linear criado foram: média dos erros (MBE), variância dos erros (s^2), raiz do quadrado médio dos erros (RMSE) e média dos erros absolutos (MAE) e estão dispostos na Tabela 6.

Tabela 6 - Valores dos coeficientes do modelo geral e das mesorregiões.

Descrição da medida estatística	Ajuste do Modelo							
	Geral	**CNB**	**CSB**	**EOB**	**NB**	**RMS**	**SB**	**VSF**
Média dos erros - MBE (°C)	0,265	0,000	0,000	0,000	0,000	0,000	0,309	0,000
Variância dos erros - s^2	**2,608**	**0,160**	**0,150**	**0,010**	**0,000**	**0,000**	**0,230**	**0,120**
Raiz do quadrado médio dos erros - RMSE (°C)	1,636	0,400	0,150	0,080	0,040	0,050	0,570	0,340
Média dos erros absolutos – MAE (°C)	**1,320**	**0,180**	**0,030**	**0,040**	**0,030**	**0,030**	**0,480**	**0,110**
Média dos valores observados – O (°C)	24,420	23,740	24,420	25,130	24,700	24,730	24,330	26,570
P-Média dos valores estimados - P (°C)	24,700	23,740	21,830	25,130	24,700	24,730	24,640	26,570
Coeficiente de correlação - r	0,751	0,980	0,980	1,000	1,000	1,000	0,970	0,980
Coeficiente de Wilmontt - d	0,768	1,000	1,000	1,000	1,000	1,000	1,000	1,000
Índice de desempenho - c	**0,768**	**0,990**	**0,990**	**1,000**	**1,000**	**1,000**	**0,970**	**0,980**
Número de observações - N	6553	1512	1008	756	1008	504	504	1260

Fonte: Elaborado pelos autores (2023).

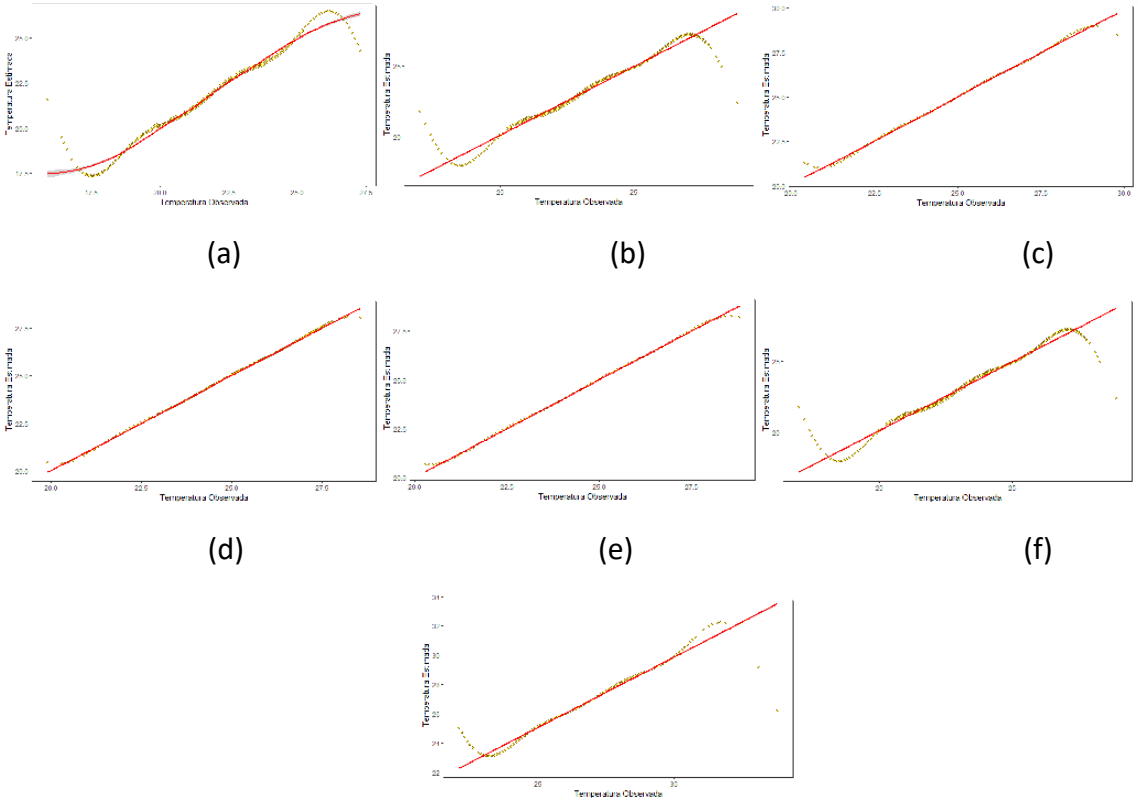
Quando se analisa as diferenças dos valores de temperatura mensal estimados e observados, tem-se que a maior magnitude do erro absoluto médio foi 0.48°C em detrimento de 1.32°C do modelo geral. Uma variância dos erros máxima de 0.23 já no modelo geral esse valor foi de 2.608, indicando que a dispersão média dos dados foi de cerca de 0.2 °C.

No que se refere a aplicação desse modelo de ajuste de dados espaciais e climáticos, de uma forma geral, as mesorregiões, obtiveram d = 1 e r variando entre 0.97 e 1.00, mostrando que existe uma correlação muito forte, indicando uma boa exatidão e precisão do modelo. Esses valores demonstram que os modelos subdivididos foram capazes de estimar a temperatura média mensal do ar com boa exatidão e precisão.

O desempenho do modelo, medido pelo índice de desempenho, c, variou de 0.97 a 1.00 aproximadamente classificado como muito bom, já o do modelo geral foi de 0.77 aproximadamente, indicando que uma correlação forte entre as variáveis.

Na Figura 15 observa-se o gráfico de dispersão das variáveis estimadas e observadas, quanto mais forte for a correlação linear mais o gráfico irá se parecer com uma reta.

Figura 15 - Gráfico de Dispersão entre as variáveis estimadas e observadas das mesorregiões



(a) (b) (c)

(d) (e) (f)

(g)

Fonte: Elaborado pelos autores (2023).

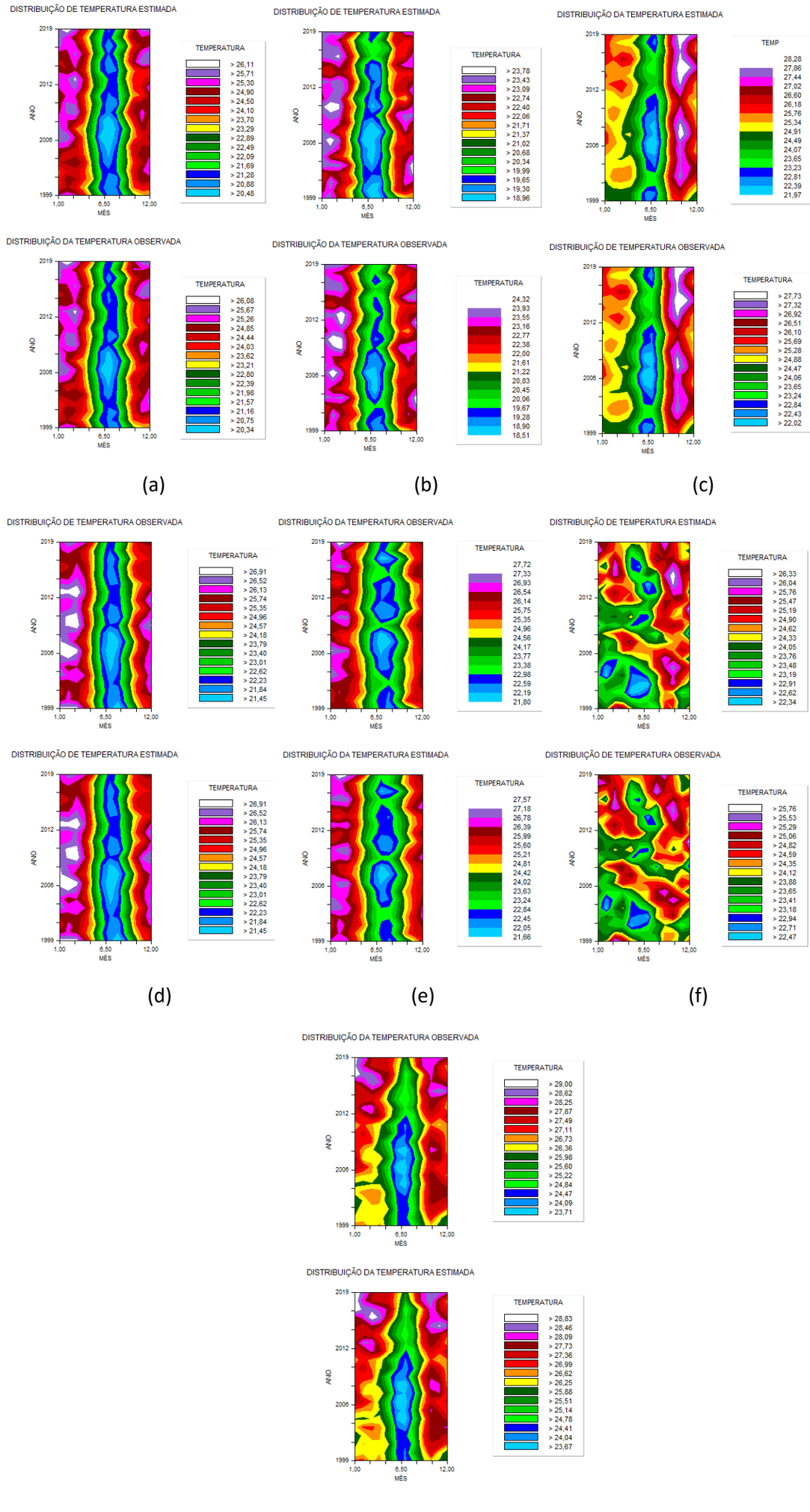
Na Figura 15 observamos que a dispersão da variável temperatura, tanto a observada quanto a estimada, e que as mesorregiões com melhor aproximação foram (b), (d) e (e) correspondentes a CSB, NB e RMS que, de acordo com os resultados dispostos na Tabela 6 apresentam índice de desempenho igual a 1, valor estimado e observado iguais.

A avaliação do modelo demonstrou que uma grande os critérios estabelecidos pelo método dos mínimos quadrados foram atendidos, e que o modelo será capaz de realizar previsões para novos dados, com uma grande representatividade delas.

3.2.4. Comparativo entre os Modelos

Para uma melhor visualização entre os dados de temperatura estimada e observada elaborou-se os gráficos dispostos na Figura 16, em duas dimensões (2D), nas quais é possível verificar a discrepância entre os valores estimados e observados verticalmente.

Figura 16 - Gráfico em 2D das temperaturas estimada e observada das mesorregiões.



Fonte: Elaborado pelos autores (2023).

Analisando a Figura 16 é possível notar que há uma diferença máxima 0.57°C entre o valor médio da temperatura máxima na mesorregião SB e de 0.45° C na temperatura mínima da mesorregião CSB, indicando que o modelo consegue descrever muito bem a estimativa.

Fica evidenciando tanto numérica como graficamente que os modelos das mesorregiões conseguem efetuar a estimativa da temperatura com consistência.

4. CONSIDERAÇÕES FINAIS

Concluímos que o modelo geral não conseguiu realizar previsões com uma boa exatidão e precisão, obteve-se um valor de índice de desempenho de 0.76. Este valor pode ser justificado por querer descrever em um modelo único uma região tão heterogênea como a Bahia, com fator limitante predominante sendo a altitude, pois há uma variação de altitude de cerca de 1000m entre estações meteorológicas.

Ao realizar as subdivisões em mesorregiões que possui características de altitude com uma discrepância de no máximo 350 m de altitude entre uma estação meteorológica para outra o índice desempenho obtido foi acima de 0.99, já as mesorregiões que apresentaram uma discrepância maior que 350m de altitude, o índice de desempenho obtido, foi um valor acima de 0.97, ambos caracterizando uma excelente precisão e exatidão dos modelos obtidos.

REFERÊNCIAS

ANTONINI, J.C.A. *et al*. **Modelo matemático para estimativa da temperatura média diária do ar no Estado de Goiás**. Pesquisa Agropecuária Brasileira, v.44, n.4, p.331-338, 2009. DOI https://doi.org/10.1590/S0100-204X2009000400001. Disponível em: https://www.scielo.br/j/pab/a/zynGJvC9ndLzbN7QXyccZ4p/?lang=pt#. Acesso em: 15 nov. 2021.

CAMARGO, A.P. de; CAMARGO, M.B.P. de. **Uma revisão analítica da evapotranspiração potencial**. Bragantia, v.59, p.125-137, 2000. Disponível em: https://www.scielo.br/j/brag/a/94MFPHnhYKGKRds4cMKFBBt/abstract/?lang=pt. Acesso em: 01 out. 2021.

CAMARGO, A.P. de; SENTELHAS, P.C. Avaliação de desempenho de diferentes métodos de estimativa da evapotranspiração potencial no Estado de São Paulo, Brasil. **Revista Brasileira de Agrometeorologia**, v.5, p.89-97, 1997.

CAPUCHINHO, F. F. *et al*. Coordenadas geográficas e altitude na estimativa da temperatura do ar no estado de Goiás. **Revista Brasileira de Agricultura Irrigada-RBAI**, v. 13, n. 2, p. 3346-3357, 2019.

FILGUEIRAS, R.; *et al*. Predição da temperatura do ar por meio de sensoriamento remoto orbital. **Revista Brasileira de Cartografia**, v. 68, n. 8, 2016.

GOMES, D.P. *et al*. Estimativa da temperatura do ar e da evapotranspiração de referência do estado do Rio de Janeiro. **Irriga**, 2014, v.19 (2), 302-314.

MEDEIROS, R. M.; HOLANDA, R. M. Classificação climática e balanço hídrico pelo método de Köppen e Thornthwaite do município de Barbalha, Ceará, Brasil. **Revista Equador**, v. 8, n.3, p. 19-43, 2019.

MEDEIROS, S. DE. S. *et al*. Estimativa e espacialização das temperaturas do ar mínimas, médias e máximas na Região Nordeste do Brasil. **Revista Brasileira de Engenharia Agrícola e Ambiental**, Campina Grande, v.9, n.2, p.247-255, 2005.

PEDRO JÚNIOR, M. J. *et al*. Determinação da temperatura-base, graus-dia e índice biometeorológico para a videira 'Niagara Rosada'. **Revista Brasileira de Agrometeorologia**, Santa Maria, v. 2, p. 51-56, 1994.