



XIX Congresso Brasileiro de Agrometeorologia

23 a 28 de agosto de 2015

Lavras – MG – Brasil

Agrometeorologia no século 21:

O desafio do uso sustentável dos biomas brasileiros



Correção de falhas de dados de temperatura média mínima utilizando métodos estatísticos

¹Yohana de Oliveira Medeiros; ²Pietros André Balbino dos Santos; ³Paulo Henrique Fernandes Figueiredo; ⁴Luiz Gonsaga de Carvalho; ⁵Gabriel Carvalho Matias; ⁶Fernanda Pereira; Jodean Alves da Silva⁷.

¹Graduando em Agronomia, Depto. de Agricultura, UFLA, Lavras – MG Fone: (32) 8427 – 5237yohanamedeiros@hotmail.com

²Eng. Agrícola, Mestrando, Depto.de Engenharia, UFLA, Lavras – MG pietros.balbino@gmail.com

³Graduando em Eng. Agrícola, Depto.de Engenharia, UFLA, Lavras-MG paulo.engagricola@gmail.com

⁴Eng. Agrícola, Prof. Associado, Depto. de Engenharia, UFLA, Lavras - MGlgonsaga@deg.ufla.br

⁵Graduando em Eng. Agrícola, Depto.de Engenharia, UFLA, Lavras-MG gabrielcarvim@hotmail.com

⁶Graduando em Zootecnia, Depto.deZootecnia, UFLA, Lavras - MG, fernandazoo@hotmail.com.br

⁷Eng. Agrônomo, Doutorando, Depto.de Engenharia, UFLA, Lavras – MG jodean2013@gmail.com

RESUMO: A ausência de registros de dados meteorológicos é um problema recorrente, seja ele por falhas dos equipamentos, falta de calibração e falha humana. No preenchimento dessas falhas de dados se destaca procedimentos estatísticos e geoestatístico. O objetivo deste trabalho é avaliar técnicas estatísticas para o preenchimento de falhas em uma série de dados de temperatura média mínima. Os dados utilizados neste trabalho foram registrados pela estação meteorológica convencional situada na cidade de Lavras – Minas Gerais pertencente à rede do INMET (Instituto Nacional de Meteorologia). Usando-se dados de temperatura média mínima dos “vizinhos próximos” estimou-se, usando técnicas estatísticas como média ponderada e inverso da distancia, e comparou com valores reais observados ajustando assim modelos de estimativa. Os valores estimados de temperatura mínima apresentaram uma boa correlação com os valores observados, sendo que a correlação se enfraquecia conforme se considerava uma maior quantidade de dias de influencia sobre o valor a ser estimado, no entanto todos os modelos mostram-se satisfatório para estimativa da temperatura média mínima, com r^2 variando entre 0,76 a 0,84.

PALAVRAS-CHAVE: Técnicas estatísticas, série de dados, estação climatológica.

Data error correction average minimum temperature using statistical methods

ABSTRACT The absence of meteorological data records is a recurring problem, whether by equipment failure, lack of calibration and human error. In filling these data gaps stand out statistical and geostatistical procedures. The objective of this study is to evaluate statistical techniques for gap filling in a series of minimum temperature data. The data used in this work were recorded by conventional weather station in the city of Lavras - Minas Gerais belonging to the network of INMET (National Institute of Meteorology). Using minimum temperature data of "nearest neighbors" was estimated using statistical techniques such as inverse and weighted average distance, and compared with actual values observed adjusting well estimation models. The estimated values of minimum temperature showed a good correlation with the observed values, and the correlation is weakened as is considered a higher number of days of influence on the value to be estimated, however all models show satisfactory for estimation of the mean minimum temperature, with r^2 ranging from 0.76 to 0.84.

KEYWORDS: statistical techniques, data series, climatological station

Acompanhar a variação dos dados climáticos é necessário no meio agrícola para simular o crescimento da cultura e assim se preparar para cenários futuros. Para isso é necessárias longas séries de dados para geração de um melhor modelo. No entanto, uma das maiores deficiências que limitam o uso de modelos em larga escala é a falta de dados de solo e clima Ritchie et al., (1990)

Dados climatológicos podem ser obtidos por meios automáticos (estação automática), onde as informações são obtidas a partir de vários sensores 24 horas por dia e em intervalos de tempo variados, ou manualmente (estação convencional), onde há a necessidade de um observador para anotar as medidas captadas pelos equipamentos. Entretanto, diversos problemas podem ocorrer, impedindo que o observador realize alguma leitura em determinado horário do dia, provocando assim falhas nos registros dos dados Ventura, (2012). Pickering et al. (1994) destaca os seguintes fatores que levam à falha de dados: erros no formato, medidas não realizadas e dados eliminados devido ao erro de coleta.

Existem diversos estudos na linha de estimativa de temperatura, com técnicas distintas e que apresentam resultados satisfatório, Oliveira Neto et al. (2002), utilizando altitude longitude e latitude do local, ajustaram equações de regressão para estimar os valores médios mensais de temperatura mínima, média e máxima.

Todavia, estes métodos apresentam modelos que levam em consideração latitude, longitude e altitude, mas não apresentam formas de se estimar a temperatura sem estas informações. Sendo assim, objetivou-se no presente trabalho desenvolver modelos que estimem a temperatura mínima do dia em função das temperaturas medias diárias coletadas em dias anteriores e posteriores, e assim estimar a temperatura mínima para dias faltantes ou que apresente discrepâncias no dado coletado.

MATERIAIS E MÉTODOS

A metodologia proposta consiste em obter dados da estação meteorológica convencional e comparar com dados obtidos por meio de técnicas estatísticas. Os dados meteorológicos foram coletados na estação meteorológica convencional situada na Universidade Federal de Lavras, Lavras, Minas Gerais, Brasil, no período de 1/1/2007 a 31/12/2007. As técnicas estatísticas usadas para estimar o dado faltante foram: média simples e o inverso do quadrado da distancia. A média simples se deu da seguinte forma, obteve a média dos dados anteriores e posteriores ao dia a ser estimado, totalizando no máximo oito dados, quatro anteriores e quatro posteriores, podendo observar nas equações 1, 2, 3 e 4.

$$X_{p_1} = \frac{(X_{n-1} + X_{n+1})}{2} \quad (1) \quad X_{p_3} = \frac{(X_{n-3} + X_{n-2} + X_{n-1} + X_{n+1} + X_{n+2} + X_{n-3})}{6} \quad (3)$$

$$X_{p_2} = \frac{(X_{n-2} + X_{n-1} + X_{n+1} + X_{n+2})}{4} \quad (2) \quad X_{p_4} = \frac{(X_{n-4} + X_{n-3} + X_{n-2} + X_{n-1} + X_{n+1} + X_{n+2} + X_{n-3} + X_{n-4})}{8} \quad (4)$$

X_p : variável estimada; X_{n-1} : valor da variável n - ésim anterior ao dado faltante; X_{n+1} : valor da variável n - ésim posterior ao dado faltante.

A estimativa pelo inverso do quadrado da distancia, conforme as equações 5, 6 e 7, foram feitas ponderando pesos de menor valor conforme o dado se distancia, em dias, do dado a ser estimado. Usou-se valores e ponderação d (Figura 1) que melhor se ajustaram.

$$X_n = \frac{\frac{(X_{n-1} + X_{n+1})}{d_1^2} + \frac{(X_{n+2} + X_{n-2})}{d_2^2}}{d_1^2 + d_2^2} \quad (5)$$

$$X_n = \frac{\frac{(X_{n-1} + X_{n+1})}{d_1^2} + \frac{(X_{n+2} + X_{n-2})}{d_2^2} + \frac{(X_{n+3} + X_{n-3})}{d_3^2}}{d_1^2 + d_2^2 + d_3^2} \quad (6)$$

$$X_n = \frac{\frac{(X_{n-1} + X_{n+1})}{d_1^2} + \frac{(X_{n+2} + X_{n-2})}{d_2^2} + \frac{(X_{n+3} + X_{n-3})}{d_3^2} + \frac{(X_{n+4} + X_{n-4})}{d_4^2}}{d_1^2 + d_2^2 + d_3^2 + d_4^2} \quad (7)$$

X_p : variável estimada; X_{n-1} : valor da variável n - ésima anterior ao dado faltante; X_{n+1} : valor da variável n - ésima posterior ao dado faltante.

O d é o valor a ser ponderado á media do dia, levando em consideração que a influencia da temperatura o dia, na n-ésima posição, é inversamente proporcional à distancia (em dias) do ponto estimado.

Tabela1- Valores do d ajustados para as equações 5, 6 e 7.

	Equações		
	5	6	7
d_1	1	1	1
d_2	1,4	1,2	1,1
d_3	-	1,4	1,3
d_4	-	-	1,4

As análises estatísticas foram realizadas por meio do Ms Excel®, e os dados foram submetidos a teste de correlação e exatidão para obtenção da equação que melhor se ajuste.

RESULTADOS E DISCUSSÃO

Na Figura 2, estão dispostas a dispersão dos valores de temperatura mínima estimada (T_{min} Estimada) pelas equações 1, 2, 3 e 4 em relação a médias diárias de temperatura mínima observada (T_{min}). Foram realizada comparação entre T_{min} estimado e T_{min} observado e assim gerado modelos para a estimativa da temperatura.

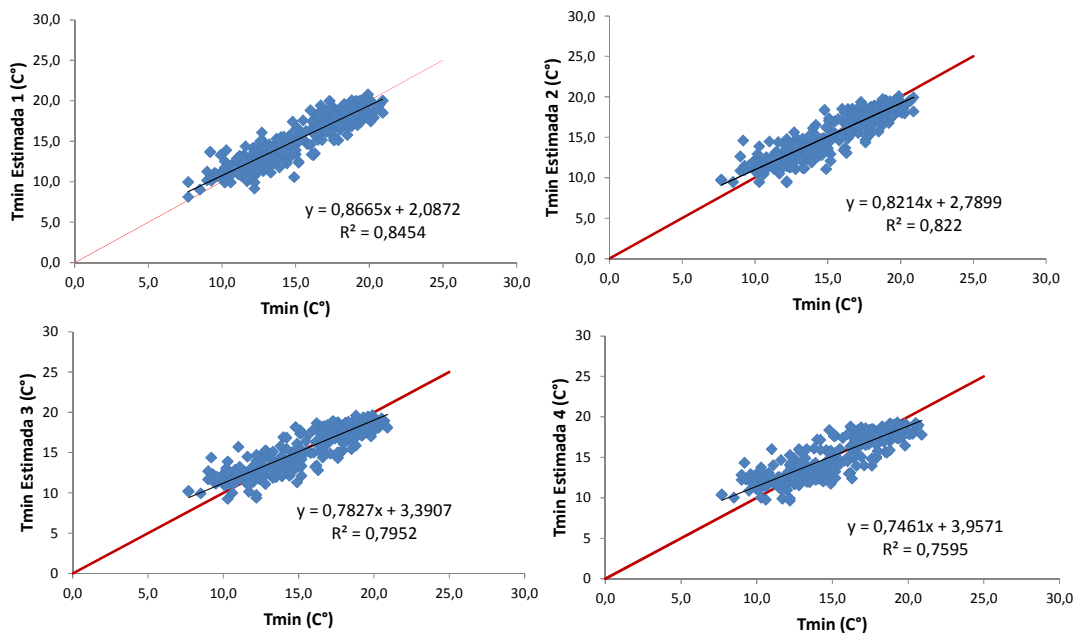


Figura 2 –Diagrama de dispersão dos valores de temperatura mínima observada e estimadas por médias entre valores vizinhos em $^{\circ}C$, as equações lineares e correspondente índice estatísticos de avaliação de ajuste coeficiente de determinação R^2 .

Figura 3 encontra-se valores dispersos de temperatura mínima estimada (Tmin Estimada) pelas equações 5, 6 e 7 em relação à média diária de temperatura mínima observada (Tmin) e os respectivos modelos gerados para a estimativa da temperatura do dia.

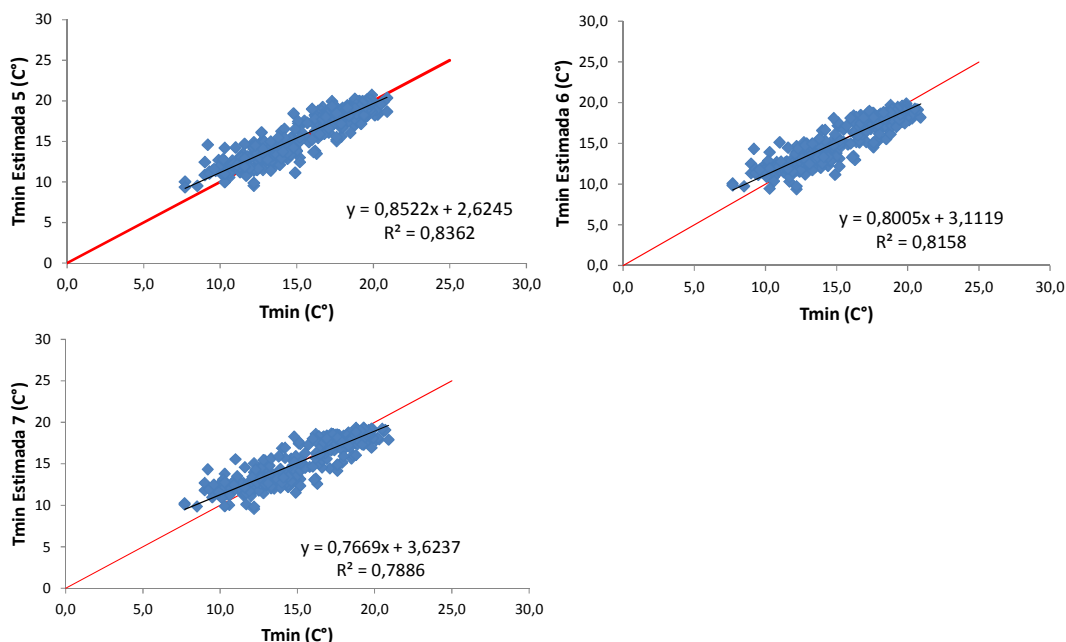


Figura 3–Diagrama de dispersão dos valores de temperatura mínima observada e estimadas por médias entre valores vizinhos ponderados pela distancia em C°, as equações lineares e correspondente índice estatísticos de avaliação de ajuste coeficiente de determinação R².

O índice de concordância de Willmont, Willmont (1981) fornece o grau de exatidão entre as variáveis envolvidas, uma vez que está relacionada à diferença entre os valores estimados em relação aos valores observados, e quanto mais próximo de 1 indica uma confiança perfeito. Tabela 1 e Tabela 2 estão apresentados os valores de Willmont, todas as equações obteve valores acima de 0,93, desempenho este considerado ótimo.

Tabela 1 –Estatística do coeficiente de Willmot e Nash para os modelos propostos

	Tmin estimada (C°)			
	1	2	3	4
Willmot	0,96	0,95	0,94	0,93
Nash	0,83	0,78	0,73	0,67

Tabela 2 - Estatística do coeficiente de Willmot e Nash para os modelos propostos

	Tmin estimada (C°)		
	5	6	7
Willmot	0,95	0,94	0,94
Nash	0,80	0,77	0,72

De acordo com Zhang et al. (2007), o valor do coeficiente Nash-Sutcliffe (Nash) indica o ajustamento dos dados simulados aos observados na reta 1:1, podendo variar de $-\infty$ a 1. Os valores observado na



XIX Congresso Brasileiro de Agrometeorologia

23 a 28 de agosto de 2015

Lavras – MG – Brasil

Agrometeorologia no século 21:



O desafio do uso sustentável dos biomas brasileiros

Tabela 1, Nash ajustado entre os valores estimados pelas equações 1, 2, 3 e 4 e os valores observados, variam entre 0,67 e 0,83, sendo o maior valor observado para o ajuste da equação 1 e o pior desempenho para a equação 4. Na Tabela 2 encontram-se os valores do Nash para as equações 5, 6 e 7, com valores variando entre 0,8 e 0,72.

Segundo Silva et al. (2008), quando o valor de Nash resultar valores maior que 0,75, o desempenho do modelo é considerado bom. Para Nash entre 0,36 e 0,75, o desempenho é considerado aceitável, enquanto valores de Nash inferiores a 0,36 fazem com que o modelo seja julgado como inaceitável. Os valores encontrados para as equações 1 e 2 (Tabela 1) e equações 5 e 6 (Tabela 2) apresentam um desempenho bom, e os outros modelos apresentam desempenho aceitável, considerando a classificação proposta por Silva et al. (2008).

CONCLUSÕES

Os valores estimados de temperatura mínima apresentaram uma boa correlação com os valores observados, sendo que a correlação se enfraquecia conforme se considerava uma maior quantidade de dias de influência sobre o valor a ser estimado.

Todos os modelos mostram-se satisfatório para estimativa da temperatura média mínima, com r^2 variando entre 0,76 a 0,84. A equação 1 para a estimativa da temperatura mínima, onde considera a média entre um dia posterior e anterior ao dia a ser estimado, apresenta o melhor desempenho. Considerando valores de ponderação, a equação 5 obteve um melhor desempenho em relação a equação 6 e 7.

REFERÊNCIAS BIBLIOGRÁFICAS

OLIVEIRA NETO, S N. de; REIS, G. G. dos; REIS, M. das G. F.; LEITE, H. G.; COSTA, J. M. N. da. Estimativa da temperatura mínima, média e máxima do território brasileiro situado entre 16 e 24° latitude sul e 48 e 60° longitude oeste. **Engenharia na Agricultura**, v. 10, n. 1, p. 8-17, 20102.

RITCHIE, J.T., GODWIN, D.C. and SINGH, U. Soil and weather inputs for the IBSNAT crop models. In: Proceedings of IBSNAT Symposium of Decision Support System for Agrotechnology Transfer. Dept. of Agronomy and Soil Science, College of Tropical Agriculture and Human Resources, University of Hawaii, Honolulu, HI., p.31-45, 1990.

SILVA, P.M.O.; MELLO, C.R.; SILVA, A.M. & COELHO, G. Modelagem da hidrografia de cheia em uma bacia hidrográfica da região Alto Rio Grande. *R. Bras. Eng. Agric. Amb.*, 12:258-265, 2008.

VENTURA, T. M. **Preenchimento de falhas de dados micrometeorológicos utilizando técnicas de inteligência artificial**. Dissertação (Dissertação em Física Ambiental) - UFMT, 2012.

WILLMOTT, C. J. et al. Statistics for the evaluation and comparison of models. *Journal of Geophysical Research*, Ottawa, v. 90, n. C5, p. 8995-9005, 1985.

ZHANG, H. G.; Fu, S. H.; Fang, W. H.; Imura, H.; Zhang, X. C. (2007), Potential effects of climate change on runoff in the Yellow River Basin of China. *Transactions of the ASABE*, 50, 911-918.