



Previsão da temperatura média mensal de Uberlândia, MG, com modelos de séries temporais

Maria I. S. Silva¹, Ednaldo C. Guimarães² & Marcelo Tavares²

RESUMO

Modelos de séries temporais têm sido amplamente usados no estudo de variáveis climatológicas, como temperatura e precipitação. Diversos são os objetivos traçados neste trabalho a fim de se analisar a série de temperatura média mensal da cidade de Uberlândia, MG, descrevendo seus componentes, e fazer previsões para períodos subseqüentes através de modelos ajustados para a série. A análise permitiu identificar, na série, a presença dos componentes, tendência e sazonalidade. Modelos do tipo SARIMA foram ajustados e, por meio dos critérios AIC (Akaike Information Criterion), BIC (Bayesian Information Criterion) e MSE (Mean Square Error) foi selecionado o modelo SARIMA (3,1,0)(0,1,1) para fins de previsão.

Palavras-chave: modelo SARIMA, regressão, sazonalidade, tendência

Forecast of monthly mean temperatures in Uberlândia, Minas Gerais, Brazil using time series models

ABSTRACT

Time series models have been used for climatological variables such as temperature and rainfall. In this paper, time series of monthly mean temperatures for the municipality of Uberlândia, Minas Gerais, were analyzed describing their components and making forecasts using the SARIMA models. The analysis showed trend and seasonality components. The SARIMA model was adjusted by AIC (Akaike Information Criterion), BIC (Bayesian Information Criterion) and MSE (Mean Square Error) criterion. The results showed that the SARIMA (3,1,0)(0,1,1) was a good model for forecasting.

Key words: regression, SARIMA model, seasonality, trend

¹ UFLA/DEX, CEP 37200-000 Lavras, MG. Fone: (35) 3829 1653. E-mail: maissilva@gmail.com. Bolsista do CNPq

² UFU/FAMAT. Campus Santa Mônica, CEP 38400-902 Uberlândia, MG. Fone: (34) 3239 4156. E-mail: ecg@ufu.br; mtavares@ufu.br

INTRODUÇÃO

Uma série temporal é um conjunto de observações ordenadas em intervalos de tempo, comumente iguais. Essas observações apresentam dependência serial e constituem, um dos objetivos do estudo de séries temporais, analisar e modelar essa dependência.

Aplicada a fatores climatológicos, a análise de séries temporais atrai especial interesse, pois é inegável que tais fatores exercem papel significativo no sucesso ou fracasso de muitos empreendimentos. Alguns modelos estatísticos mais utilizados em análise de séries temporais foram apresentados por Latorre & Cardoso (2001).

A habilidade de antecipar como o clima irá mudar de um ano para o outro, possibilita melhor gerenciamento da agricultura, recursos hídricos e atividade pesqueira, além da possibilidade de contribuição relevante nos campos dos transportes, abastecimento, turismo e lazer. Pela incorporação das previsões climáticas nas decisões gerenciais, a humanidade se tornará melhor adaptada aos ritmos irregulares do clima. Alterações climáticas em geral, em especial o aquecimento global, têm sido motivo de grandes preocupações da comunidade científica, embora vários estudos comprovem que muitas tendências de aquecimento observadas nos registros meteorológicos expressam mais os efeitos urbanos locais que o aquecimento global da atmosfera (Back, 2001); Desta forma, torna-se importante detectar se o aumento nos registros reflete um aquecimento significativo ou uma variação aleatória natural.

Cavalcanti et al. (2006) desenvolveram um modelo de estimativa da temperatura do ar em função das coordenadas geográficas e das anomalias de temperaturas da superfície do mar, para séries temporais das médias mensais de temperatura no Nordeste brasileiro. Considerando a mesma região, Medeiros et al. (2005) obtiveram estimativas da temperatura do ar mínima, média e máxima, a partir de equações de regressão múltipla, com ajuste quadrático, gerando mapas temáticos a partir das equações ajustadas.

Ferraz et al. (1999) utilizaram modelos de séries temporais na previsão da série de precipitações pluviais mensais da cidade de Lavras, MG. Dentre os modelos ajustados, modelo de regressão, modelo de alisamento exponencial de Holt-Winters e modelo SARIMA determinístico, este último apresentou melhores valores de previsão, segundo os resultados obtidos por esses autores.

Com o objetivo de identificar tendências anuais da temperatura e precipitação pluvial no estado de Santa Catarina, Back (2001) utilizou o teste de Mann-Kendall e a análise de regressão, indicando esses métodos como eficientes em estudos de séries temporais. Rigozo & Nordemann (1998) usaram o método da regressão iterativa na procura de periodicidades em séries temporais de temperatura média global do ar na superfície da Terra. Também foram estabelecidas relações entre os fenômenos geofísicos envolvidos.

Cargnelutti Filho et al. (2006) destacam a importância de se determinar o tamanho ótimo de amostra necessário para estimar a temperatura média mensal de uma localidade, além da variabilidade temporal e espacial, que também deve ser considerada.

Dada a grande importância de previsões confiáveis da temperatura do ar para diversos empreendimentos, vários estudos são conduzidos no intuito de se obter modelos que conduzam a estimativas cada vez mais precisas em diversas regiões do Brasil (Cavalcanti et al., 2006; Ferraz et al., 1999; Marin et al., 2003; Medeiros et al., 2005; Oliveira Neto et al., 2002; Pedro Júnior et al., 1991; Sedyama & Melo Júnior, 1998).

No Triângulo Mineiro, particularmente na cidade de Uberlândia, o comportamento das variáveis climatológicas é ainda pouco estudado e a base de dados é pequena em relação a outras cidades, apesar da grande importância agrícola da região no Estado.

Analisou-se, no presente estudo, uma série de temperatura média mensal da cidade de Uberlândia, MG, com o objetivo de descrever o comportamento da série por meio do seu modelo de decomposição, verificando a existência dos componentes sazonalidade e/ou tendência; identificar e estimar modelos de previsão aos quais os dados se ajustem e fazer previsões de temperatura para períodos subsequentes utilizando-se esses modelos.

MATERIAL E MÉTODOS

Os dados avaliados fazem parte do Banco de Dados Climatológicos do Instituto de Geografia - Universidade Federal de Uberlândia (IGUFU/UFU). A Estação climatológica está situada no Campus Santa Mônica da UFU e faz parte da rede de estações climatológicas do Instituto Nacional de Meteorologia (INMET), 5º Distrito de Meteorologia (5º DISME), órgão ligado ao Ministério da Agricultura.

A cidade de Uberlândia, MG, se encontra na latitude de 18° 55' 25" S, longitude de 48° 17' 19" W. Segundo a classificação de Köppen, o clima da região é do tipo Aw, isto é, tropical quente úmido com inverno frio e seco. O total médio de chuva no mês mais seco fica em torno de 60 mm e no mês mais chuvoso em torno de 250 mm e o total anual médio fica entre 1500 a 1600 mm. Os meses de verão (dezembro a fevereiro) são responsáveis por aproximadamente 50% da precipitação anual da cidade. A temperatura média mensal nos meses de inverno atinge 18 °C enquanto nos meses mais quentes a média fica em torno de 23 °C, com média das máximas por volta de 28 a 29 °C (Silva et al., 2003).

A cidade de Uberlândia localiza-se no Triângulo Mineiro e conta com uma população estimada de aproximadamente 600.000 habitantes, constituindo-se em uma cidade pólo para a região. A cidade apresenta desenvolvimento relevante no setor agrícola além de um alto grau de desenvolvimento industrial e comercial.

A variável analisada foi a temperatura média mensal registrada em Uberlândia, no período de janeiro de 1981 a dezembro de 2003 totalizando, assim, uma série de 276 observações. Os valores referentes ao ano de 2003 foram reservados para comparação com as previsões, analisando-se então a série com 264 dados.

Inicialmente, a série foi analisada graficamente e em seguida se utilizaram os testes não paramétricos de Cox – Stu-

art e de Mann – Kendall, além da análise de regressão múltipla e teste de Fisher (Morettin & Toloi, 2004), no intuito de investigar a presença de tendência e periodicidades nos dados. Após identificar os componentes da série, utilizou-se o operador diferença para torná-la estacionária.

Modelos da classe SARIMA foram ajustados e, a qualidade do ajuste verificada por meio do teste de Box Pierce, com base na estatística Q da função de autocorrelação.

Os modelos ajustados foram usados para previsões e, por meio dos critérios AIC (Akaike Information Criterion), BIC (Bayesian Information Criterion) e MSE (Mean Square Error) indicou-se aquele que melhor se ajustou e o que forneceu previsões com menor erro quadrático médio de previsão.

RESULTADOS E DISCUSSÃO

Uma inspeção visual na Figura 1 permite uma análise preliminar do comportamento da série sob estudo.

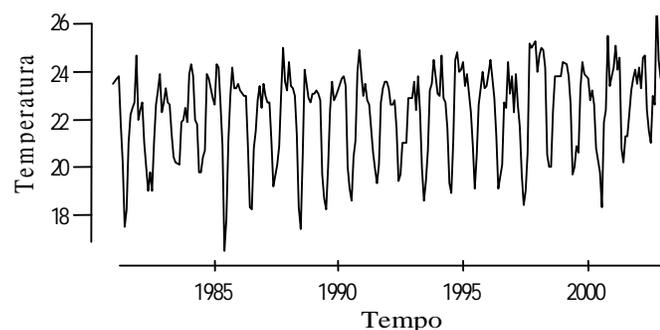


Figura 1. Representação gráfica da série de temperatura média mensal de Uberlândia

Nesta primeira análise se observa pequeno indício de tendência crescente nos dados, além de ficar bastante evidente a existência de sazonalidade, conforme esperado, em se tratando de séries climatológicas. A significância estatística da tendência foi verificada pelo teste do sinal, pelo teste de Mann – Kendall e pela regressão múltipla.

Para o teste do sinal a série foi dividida em $c = 132$ pares (Z_i, Z_{i+c}) , obtendo-se, assim, $T_2 = 93$ sinais positivos $(Z_i < Z_{i+c})$, 35 sinais negativos e 4 igualdades, sendo $n = 128$ o número de pares em que $Z_i \neq Z_{i+c}$. O teste é baseado em rejeitar H_0 (H_0 : não existe tendência) se $T_2 \geq n - t$, sendo t um valor tabelado obtido por meio da aproximação da binomial pela distribuição normal; em outras palavras, não se rejeita a existência de tendência se o número de sinais positivos for muito maior que o de números negativos. Usando-se a aproximação normal, obteve-se o valor de t de 73,3 e então se rejeita a hipótese H_0 , concluindo-se haver tendência positiva na série.

O teste de Mann-Kendall, descrito em Back (2001), apresentou um valor da estatística de teste $u(t)$ dado por 3,57. A hipótese nula de ausência de tendência pode ser rejeitada com probabilidade α_1 calculada por meio de uma tabela da normal reduzida, tal que $\alpha_1 = \text{prob}(|u| > |u(t)|)$ considerando-se, em ge-

ral, um nível de significância de 0,05. Consultando-se a tabela, obteve-se $\alpha_1 = 0,00018$ permitindo, portanto, a conclusão de que existe tendência na série e, pelo sinal de $u(t)$, que esta tendência é positiva.

O método da regressão múltipla verifica a existência de tendência e sazonalidade e os resultados estão apresentados na Tabela 1. O teste F para o modelo de regressão foi altamente significativo ($\text{Pr} > F = 0,0001$), indicando bom ajuste desse modelo aos dados. Os dados da Tabela 1 mostram que, na sua maioria, os coeficientes da regressão referentes aos meses (α_j), foram significativos, confirmando a existência de sazonalidade determinística nos dados. O coeficiente b_1 , referente aos anos, também foi significativo, confirmando a existência de tendência na série.

Tabela 1. Parâmetros da análise de regressão múltipla da variável temperatura como função dos meses e dos anos

Coef	Estimativa	Erro Padrão	T	Prob > T
β_0	22,66	0,19	117,60	<0,0001
β_1	0,062	0,008	7,97	<0,0001
α_1	0,13	0,24	0,53	0,5986
α_2	0,26	0,24	1,07	0,2844
α_3	0,08	0,24	0,32	0,7493
α_4	-0,61	0,24	-2,52	0,0123
α_5	-2,64	0,24	-10,94	<0,0001
α_6	-4,15	0,24	-17,20	<0,0001
α_7	-4,02	0,24	-16,64	<0,0001
α_8	-2,28	0,24	-9,45	<0,0001
α_9	-0,64	0,24	-2,64	0,0089
α_{10}	0,55	0,24	2,28	0,0236
α_{11}	0,25	0,24	1,05	0,2929

Como alternativa para verificar a presença de periodicidades na série, utilizou-se o teste de Fisher descrito em Morettin & Toloi (2004), de acordo com o qual o maior período, de 12 meses, referente ao componente sazonal, foi significativo a um nível de significância menor que 0,0001; conclui-se portanto, que existe sazonalidade de 12 meses na série, conforme em geral ocorre em séries de temperatura.

Identificados os componentes tendência e componente sazonal, utilizou-se o operador diferença (Morettin & Toloi, 2004) para tornar a série estacionária e a partir daí ajustar um modelo a essa nova série.

Após tomar uma diferença simples na série verificou-se, graficamente, que a tendência foi eliminada porém, tomando-se uma diferença de doze para eliminar a sazonalidade, notou-se que ainda existiam “lags” múltiplos de doze significativos na função de autocorrelação; concluiu-se, desta forma, que seria necessário incluir no modelo um componente sazonal que considere a existência de sazonalidade estocástica, além da determinística, retirada com uma diferença de doze.

O modelo sugerido para esta série é do tipo auto-regressivo integrado de médias móveis sazonal, ou seja, SARIMA(p,d,q)(P,D,Q) que pode ser representado pela expressão:

$$\phi(B)\Phi(B^{12})\Delta^d\Delta_{12}^D Z_t = \theta(B)\Theta(B^{12})a_t$$

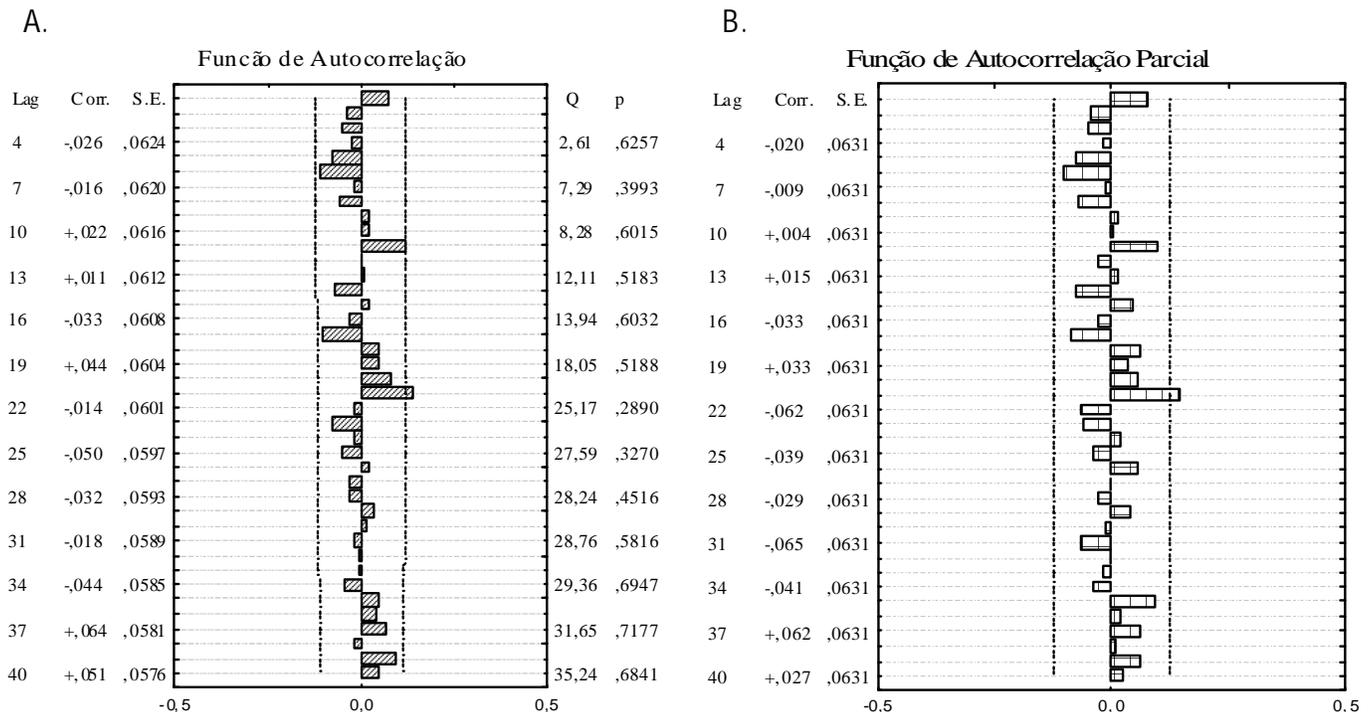


Figura 2. Função de autocorrelação (2A) e função de autocorrelação parcial (2B) dos resíduos do modelo ajustado SARIMA(0,1,1)(0,1,1)

em que:

$\phi(B) = 1 - \phi_1(B) - \phi_2(B^2) - \dots - \phi_p(B^p)$ é o operador auto-regressivo de ordem p

$\Phi(B^{12}) = 1 - \Phi_1(B^{12}) - \Phi_2(B^{12})^2 - \dots - \Phi_p(B^{12})^p$ é o operador de médias móveis de ordem q

$\Phi(B^{12}) = 1 - \Phi_1(B^{12}) - \Phi_2(B^{12})^2 - \dots - \Phi_p(B^{12})^p$ é o operador auto-regressivo sazonal de ordem P

$\Theta(B^{12}) = 1 - \Theta_1(B^{12}) - \Theta_2(B^{12})^2 - \dots - \Theta_q(B^{12})^q$ é o operador de médias móveis sazonal de ordem Q

Δ^d é o operador diferença simples, d indicando o número de diferenças

Δ_{12}^D é o operador diferença sazonal, D indicando o número de diferenças sazonais

a_t é o resíduo que, eventualmente, pode ser ruído branco, $BZ_t = Z_{t-1}$, e $\Delta = 1 - B$.

Através da análise das funções de autocorrelação e de autocorrelação parcial da série com uma diferença simples e uma diferença sazonal de 12, verificaram-se as ordens dos modelos que poderiam ser ajustados.

O primeiro modelo ajustado foi um SARIMA(0,1,1)(0,1,1). As estimativas dos dois parâmetros do modelo foram significativas e seus valores são: $q_1 = 0,74794$ e $Q_1 = 0,73212$. O modelo é dado pela expressão:

$$(1-B^{12})(1-B)Z_t = (1 - 0,7479B)(1 - 0,73212B^{12})a_t$$

Para verificar a qualidade do ajuste, os gráficos das funções de autocorrelação e de autocorrelação parcial dos resíduos foram analisados e estão apresentados na Figura 2.

Analisando-se visualmente os gráficos da Figura 2A e 2B, nota-se que, em sua maioria, os “lags” foram não significativos (dentro do limite de confiança), indicando série de resíduos estacionária e possivelmente ruído branco. A estatística Q da função de autocorrelação para o “lag” 40 foi de 35,24; portanto, menor que o valor de qui-

quadrado com 38 graus de liberdade; assim, o teste de Box Pierce ($Q < \chi_{38}^2 = 53,38$) é não significativo e, portanto, aceita-se a hipótese de que o resíduo é um ruído branco. Conclui-se, desta forma, que os dados se ajustam bem a este modelo.

O modelo SARIMA(2,1,0)(0,1,1), apresentou parâmetros estimados, todos significativos com valores $\phi_1 = -0,5307$, $\phi_2 = -0,2753$ e $Q_1 = 0,76275$; e expressão dada por:

$$(1 + 0,5307B + 0,2753B^2)(1-B^{12})(1-B)Z_t = (1 - 0,7627B^{12})a_t$$

O ajuste foi analisado adotando-se os mesmos procedimentos que no modelo anterior. O teste de Box Pierce, dado por $Q = 44,34 < \chi_{36}^2 = 52,19$, foi também não significativo; conclui-se, então, que o modelo apresentou bom ajuste mas o “lag” 3 da função de autocorrelação parcial significativo indica que um novo modelo pode ser ajustado, aumentando-se a ordem auto-regressiva do modelo.

Com base nesta última informação, o novo modelo ajustado foi o SARIMA(3,1,0)(0,1,1) cujas estimativas dos parâmetros, todas significativas, foram $\phi_1 = -0,5863$, $\phi_2 = -0,3846$, $\phi_3 = -0,2168$ e $Q_1 = 0,74807$. O teste de Box Pierce com 36 graus de liberdade apresentou o valor de $Q = 42,84 < \chi_{36}^2 = 50,998$, indicando que o resíduo do modelo é um ruído branco, o que confirma a qualidade do ajuste.

A Tabela 2 apresenta os valores de temperatura previstos para o ano de 2003, por meio dos três modelos ajustados, e os valores observados de temperatura nesse ano.

De maneira geral, as previsões a partir dos três modelos se aproximaram dos valores reais da série apresentando, porém, tendência a superestimar este valor observado.

A escolha do melhor modelo foi feita através dos critérios AIC, BIC e MSE, cujos valores estão apresentados na Tabela 3.

De acordo com os critérios AIC e BIC, o SARIMA(0,1,1)(0,1,1) foi o modelo mais bem ajustado, seguido dos modelos SARIMA(3,1,0)(0,1,1) e SARIMA(2,1,0)(0,1,1); nessa ordem, porém;

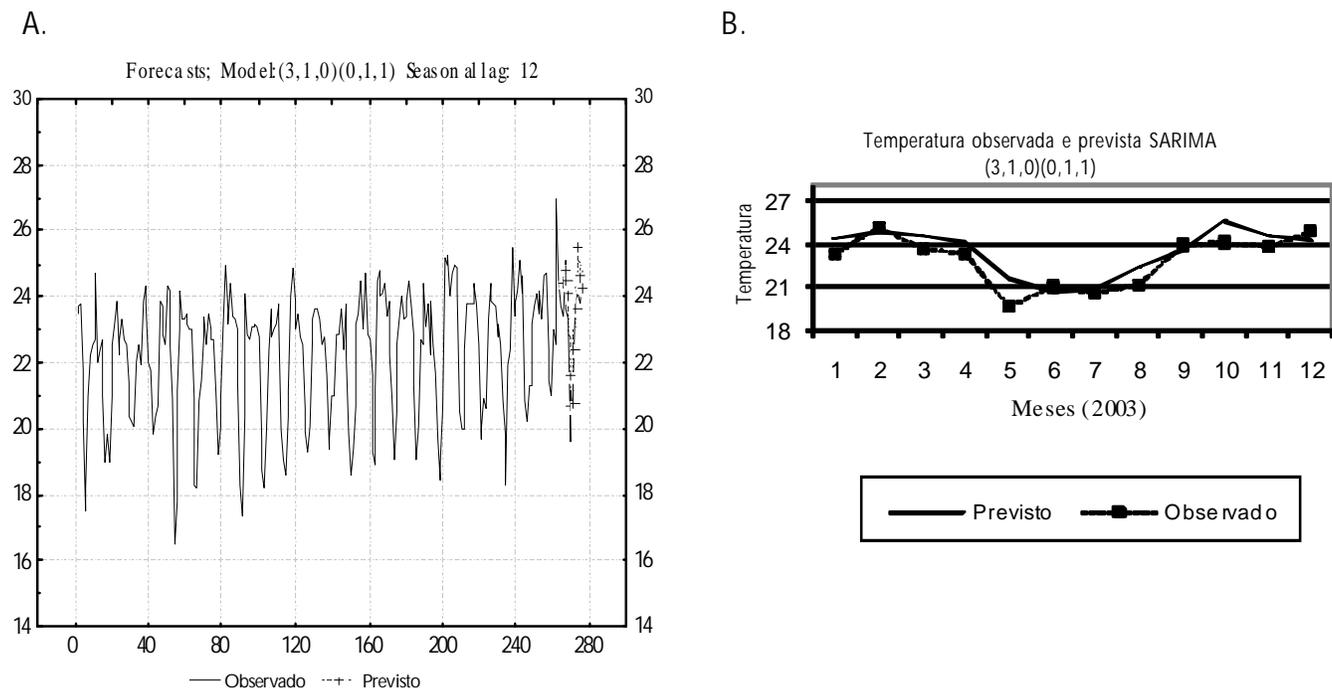


Figura 3. Temperatura estimada e observada em todo o período (3A.), apenas em 2003 (3B)

Tabela 2. Valores de temperatura média, observados e previstos para o ano de 2003, com os modelos M_1 : SARIMA(0,1,1)(0,1,1); M_2 : SARIMA(2,1,0)(0,1,1); M_3 : SARIMA(3,1,0)(0,1,1)

Ano 2003		Valores previstos		
Mês	Valor real	M_1	M_2	M_3
Jan	23,4	24,80	24,88	24,44
Fev	25,1	24,72	24,52	24,8
Mar	23,6	24,75	24,43	24,52
Abr	23,3	24,44	24,2	24,09
Mai	19,6	21,86	21,61	21,6
Jun	21,1	20,99	20,68	20,76
Jul	20,5	21,06	20,8	20,81
Ago	21,2	22,67	22,41	22,41
Set	23,9	23,85	23,67	23,64
Out	24,1	25,85	25,55	25,59
Nov	23,8	24,84	24,62	24,61
Dez	24,8	24,49	24,29	24,27

Tabela 3. Critérios de escolha do melhor ajuste e melhor modelo de previsão

Modelo	AIC	BIC	MSE
SARIMA(0,1,1)(0,1,1)	-0,179	-0,159	1,379
SARIMA(2,1,0)(0,1,1)	0,008	0,029	1,076
SARIMA(3,1,0)(0,1,1)	-0,067	-0,037	0,963

o modelo SARIMA(3,1,0)(0,1,1) teve o menor erro quadrático médio de previsão; portanto, o mais indicado para fazer previsões; além disso, os valores de AIC e BIC para este último modelo também foram pequenos comparados com os valores do modelo de melhor ajuste. As previsões com este modelo e os valores observados de temperatura são apresentados na Figura 3.

Na Figura 3A é possível observar que a tendência e a sazonalidade identificadas foram reproduzidas pelos valo-

res estimados e plotados juntamente com a série completa. A Figura 3B possibilita melhor comparação entre valores observados e estimados.

A série em estudo, para fins de previsão pode, então, ser descrita pelo modelo SARIMA(3,1,0)(0,1,1), cuja expressão é dada por:

$$(1 + 0,5863B + 0,3846B^2 + 0,2168B^3)(1-B^{12})(1-B)Z_t = (1 - 0,74807B^{12})a_t$$

Considerando a importância da previsão da temperatura para o planejamento de várias atividades, espera-se que a eficiência dos modelos ajustados propicie incentivo para estudos futuros ligados ao planejamento dessas atividades na região ou em outras localidades, considerando-se também estimativas diárias que, provavelmente, serão mais eficazes que as mensais abordadas aqui.

CONCLUSÕES

1. Os modelos SARIMA foram bem ajustados e as previsões, embora ligeiramente superiores, forneceram valores próximos dos observados, reproduzindo a periodicidade e a tendência crescente detectadas na série.

2. Entre os modelos ajustados, o modelo SARIMA(3,1,0)(0,1,1) foi considerado mais adequado para prever a temperatura média mensal no município de Uberlândia, MG.

LITERATURA CITADA

Back, A. J. Application of statistical analysis for identification of climatic tendency. Pesquisa Agropecuária Brasileira, v.36, n.5, p.717-726, 2001.

- Cargnelutti Filho, A.; Matzenauer, R.; Maluf, J. R. T. Variabilidade temporal e espacial do tamanho de amostra da temperatura mínima do ar no Rio Grande do Sul, Brasil. *Ciência Rural*, v.36, n.4, p.1156-1163, 2006.
- Cavalcanti, E. P.; Silva, V. P. R.; Sousa, F. A. S. Programa computacional para a estimativa da temperatura do ar para a Região Nordeste do Brasil. *Revista Brasileira de Engenharia Agrícola e Ambiental*, v.10, n.1, p.140-147, 2006.
- Ferraz, M. I. F.; Sáfyadi, T.; Lage, G. Uso de modelos de séries temporais na previsão de séries de precipitação pluviométricas mensais no município de Lavras, MG. *Revista Brasileira de Agrometeorologia*, v.7, n.2, p.259-267, 1999.
- Latorre, M. do R. D. de O.; Cardoso, M. R. A. Time series analysis in epidemiology: an introduction to methodological aspects. *Revista Brasileira de Epidemiologia*, v.4, n.3, p.145-152, 2001.
- Marin, F. R.; Pandorf, H.; Ferreira, A. S. Estimativa das temperaturas máximas, mínimas e médias mensais para o Brasil. In: Congresso Brasileiro de Agrometeorologia, 13, 2003, Santa Maria. Anais... Santa Maria: SBA, 2003. p.761.
- Medeiros, S. S.; Cecílio, R. A.; Melo Júnior, J. C. F.; Silva, J. L. C. Estimativas e espacialização das temperaturas do ar mínimas, médias e máximas na Região Nordeste do Brasil. *Revista Brasileira de Engenharia Agrícola e Ambiental*, v.9, n.2, p.247-255, 2005.
- Morettin P. A., Toloí C. M. C. *Análise de séries temporais*. São Paulo: Edgard Blücher, 2004. 535p.
- Oliveira Neto, S. N.; Reis, G. G.; Reis, M. G. F.; Leite, H. G.; Costa, J. M. N. Estimativa de temperaturas 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.1-4, p.8-17, 2002.
- Pedro Jr., M. J.; Mello, M. H. A.; Ortolani, A. A.; Alfonsi, R. R.; Sentelhas, P. C. Estimativa das temperaturas médias mensais, das máximas e das mínimas para o Estado de São Paulo. Campinas: IAC, 1991. 11p. Boletim Técnico, n.142.
- Rigozo, N. R.; Nordemann, D. J. R. Análise por regressão iterativa de periodicidades em séries temporais de registros geofísicos. *Revista Brasileira de Geofísica*, v.16, n.2-3, p.149-158, 1998.
- Sediyama, G. C.; Melo Jr., J. C. F. Modelos para estimativas das temperaturas normais mensais médias, máximas, mínimas, e anual no Estado de Minas Gerais. *Revista Engenharia na Agricultura*, v.6, n.1, p.57-61, 1998.
- Silva, J. W.; Guimarães, E. C.; Tavares, M. Variabilidade temporal da precipitação mensal e anual na estação climatológica de Uberaba, MG. *Ciência e Agrotecnologia*, v.27, n.3, p.665-674, 2003.