

Evolução dos óbitos por neoplasias malignas de pele nas regiões sul e sudeste do brasil: aplicação da metodologia de box-jenkins

Evolution of deaths from malignant skin neoplasms in the south and southeast regions of brazil: application of the box-jenkins methodology

Vinícius Pinheiro Nunes^I, Ângela Pellegrin Ansuji^{II}, Angela Isabel dos Santos Dullius^{III}, Fernando de Jesus Moreira Júnior^{IV}, Maria Emilia Camargo^V

RESUMO

Nos conjuntos de dados cujas observações encontram-se ordenadas em instantes de tempo, indica-se a análise de séries temporais. Técnica que permite identificar padrões de sazonalidade, tendência, ciclos e mudanças de nível, sendo possível gerar modelos explicativos capazes de descrever o comportamento da sequência de observações, obtendo previsões acerca da série analisada. Esta pesquisa aplicou a metodologia de Box- Jenkins nas séries de óbitos por neoplasias malignas de pele entre o período de 1996 e 2014, ocorridos nas regiões Sul e Sudeste do Brasil, afim de identificar padrões, gerar modelos explicativos e obter previsões. Evidenciaram-se algumas características como presença de tendência crescente, ausência de sazonalidade e comportamento não estacionário nas séries de óbitos. A metodologia de Box-Jenkins demonstrou boa qualidade de predição para as séries da região Sul e Sudeste as quais apresentaram, os modelos ARIMA (4,1,0) com constante e ARIMA (0,1,1) com constante como modelos melhores ajustados aos dados. Concluiu-se que os óbitos por neoplasias malignas de pele nas regiões Sul foram explicados apenas por componentes AR e a região Sudeste por componentes MA, ambas com componente de integração.

Palavras-chave: Neoplasias malignas de pele. Séries temporais. Metodologia de Box-Jenkins

ABSTRACT

In data sets whose observations are sorted in moments of time, the analysis of temporal series is indicated. Technique that allows to identify patterns of seasonality, tendency, cycles and changes of level, being possible to generate explanatory models capable of describing the behavior of the sequence of observations, obtaining predictions about the series analyzed. This research applied the Box-Jenkins methodology in the series of deaths due to malignant skin neoplasms between 1996 and 2014, occurring in the south and southeast regions of Brazil, in order to identify patterns, generate explanatory models and obtain predictions. Some characteristics were evidenced as the presence of increasing tendency, absence of seasonality and non-stationary behavior in the death series. The Box-Jenkins methodology showed good prediction quality for the South and Southeast series, which presented the ARIMA models (4, 1, 0) with constant and ARIMA (0, 1, 1) with constant as better models adjusted to the data. It was concluded that deaths from malignant skin neoplasms in the southern regions were explained only by AR components and the Southeast region by MA components, both with integration component.

Keywords: Neoplasms of the skin. Time series. Box-Jenkins methodology

^I Universidade Federal de Santa Maria, Brasil; e-mail: vini.estadistica@gmail.com;

^{II} Universidade Federal de Santa Maria, Brasil. e-mail: angelaansuj@yahoo.com;

^{III} Universidade Federal de Santa Maria, Brasil; e-mail: angeladullius@gmail.com;

^{IV} Universidade Federal de Santa Maria, Brasil. e-mail: fmjunior777@yahoo.com.br;

^V Universidade de Caxias do Sul, Brasil. e-mail: mariaemiliappga@gmail.com



1 INTRODUÇÃO

O uso de técnicas estatísticas em estudos e pesquisas tem se tornado cada vez mais importante em diversas áreas do conhecimento como: economia, biologia, agronomia, administração, epidemiologia, física, entre outras áreas. O processo de escolha da(s) técnica(s) que será(ão) utilizada(s), entre as possíveis existentes, dependerá dos objetivos e interesses de cada pesquisador e das especificidades da variável ou das variáveis que se pretende analisar. Nos conjuntos de dados cujas observações encontram-se ordenadas em instantes de tempo, eventualmente em intervalos equidistantes, indica-se a análise de séries temporais. Esse tipo de análise permite identificar padrões como: sazonalidade, tendência, ciclos e mudanças de nível. Além disso, é possível gerar modelos explicativos capazes de descrever o comportamento da sequência de observações, e ainda obter previsões acerca da série analisada (MORETTIN; TOLOI, 2004; SOUZA; CAMARGO, 1996; BOX; JENKINS; REINSEL, 1994).

Do ponto de vista analítico, os modelos utilizados na análise de séries de tempo se diferenciam dos modelos de regressão linear e dos modelos lineares generalizados (MLG) devido à sua capacidade de analisar a dependência temporal presente no conjunto de dados. Em outras palavras, essa dependência pode ser compreendida como a influência que as observações passadas exercem sobre as observações presentes. Na análise de séries temporais univariadas têm-se os modelos autorregressivo (AR), de médias móveis (MA), e autorregressivo e de médias móveis (ARMA), descrito na sua forma geral por Whittle (1951) e, posteriormente, popularizado por Box e Jenkins (1970). O sucesso da metodologia de Box e Jenkins se deve a generalidade que os autores apresentaram a análise de modelos ARMA. Essa, possibilitou analisar não apenas processos estacionários utilizando os modelos ARMA, mas também os não estacionários. A análise dos processos não estacionários tornou-se possível a partir da inclusão de uma ordem integrada (I), que consiste em realizar sucessivas diferenças até tornar o processo estável, originando, assim, o modelo autorregressivo integrado de médias móveis (ARIMA), amplamente utilizado na análise de séries de tempo (MADDALA, 2001).

Diante disso, esta pesquisa propõe-se a realizar um estudo utilizando a metodologia de Box e Jenkins nas séries de óbitos por neoplasias malignas de pele, ocorridos nas regiões Sul e Sudeste do Brasil. Assim, esta pesquisa justifica-se pela necessidade de realizar estudos quanto a evolução dos óbitos por neoplasias malignas de pele nas regiões Sul e Sudeste, as quais apresentam as maiores incidências comparado as demais regiões, com o propósito de evidenciar e estimular a implementação de novas políticas públicas no combate e prevenção dessa doença.

2 METODOLOGIA

De acordo com a Classificação estatística internacional de doenças e problemas relacionados à saúde – Décima revisão (CID-10), estabelecida pela Organização Mundial da Saúde (OMS), as neoplasias malignas de pele correspondem aos melanomas e outras (os) neoplasias (tumores) malignas (os) da pele (C43-C44). Essas neoplasias malignas de pele estão descritas no Capítulo II da CID-10, relacionado as neoplasias (tumores), e dividem-se em dois grandes grupos: os melanomas malignos da pele (C43) e outros tipos de neoplasias da pele (C44), as quais compreendem a todas as doenças especificadas no quadro 1 (ORGANIZAÇÃO MUNDIAL DA SAÚDE, 1994). Esta pesquisa se caracteriza-se como um estudo quantitativo, descritivo e inferencial, aplicado a um contexto epidemiológico cujo objetivo foi analisar as frequências absolutas de óbitos por neoplasias malignas de pele, nas regiões Sudeste e Sul do Brasil, entre os anos de 1996 e 2015. Tratou-se de um estudo de séries temporais univariado, pois as regiões foram analisadas em um espaço de tempo, cujo enfoque foi descrever o problema no período analisado, comparar as regiões analisadas, realizar a modelagem da série de dados e obter previsões de novos casos, explicadas unicamente pela dependência temporal existente entre seus períodos passados. Além disso, pode ser definido como um estudo ecológico, pois as frequências absolutas referentes aos óbitos foram analisadas considerando o nível agregado, ou seja, a ocorrência de óbitos em grupos específicos, que nesse contexto, refere-se as regiões Sul e Sudeste do Brasil. As observações das séries foram agregadas em períodos trimestrais, equidistantes, ao longo do período analisado.

Realizou-se também uma pesquisa bibliográfica relacionada à análise de séries temporais, utilizando materiais impressos e eletrônicos, tais como: livros, artigos científicos e páginas da internet, os quais permitiram um maior conhecimento acerca do assunto estudado. Assim, a partir desse tipo de pesquisa foi possível obter subsídios substanciais para realização de uma análise correta, estatisticamente bem executada e que expresse de forma mais fidedigna possível o contexto estudado. Quanto ao contexto relacionado as informações referentes as neoplasias malignas de pele, foi efetuada uma pesquisa documental baseada na consulta de documentos oficiais, registros e relatórios. Esse tipo de pesquisa possibilitou um maior aporte da mensuração de informações associadas a esse contexto pesquisado, contribuindo no processo teórico do desenvolvimento de justificativas que reforçam a necessidade da realização deste estudo.

Para o estudo considerou-se todos os óbitos que apresentaram como causa básica às neoplasias malignas de pele, ou seja, os dados populacionais relacionados a essa doença. Esses óbitos correspondem aos melanomas malignos de pele (C43) e a outros tipos de neoplasias malignas de pele (C44). Utilizou-se 80 observações de periodicidade trimestral, entre os anos de 1996 a 2015. Utilizou-se apenas os óbitos ocorridos a partir de 1996, ano que marca o início da utilização da CID-10 como critério de classificação de doenças e agravos relacionados a saúde no Brasil. Analisou-se as informações referentes às regiões brasileiras do Sudeste, formada pelos estados do Espírito Santo (ES), Minas Gerais (MG), Rio de Janeiro (RJ) e São Paulo (SP) e Sul, formada pelos estados do Paraná (PR), Rio Grande do Sul (RS) e Santa Catarina (SC). Os registros dos óbitos foram obtidos através da base de dados do SIM, disponibilizada pelo DATASUS no TABNET. Estes dados tratam-se de informações oficiais brasileiras e estão disponíveis de forma pública. Deste modo, não houve a necessidade de avaliação e aprovação do Comitê de Ética em Pesquisa (CEP) da Universidade Federal de Santa Maria (UFSM).

Os procedimentos metodológicos utilizados foram a metodologia de séries de óbitos por neoplasias malignas de pele nas regiões Sul e Sudeste foi realizado utilizando os seguintes procedimentos metodológicos:

(1) Foram analisados os gráficos das séries originais a fim de observar e evidenciar características a partir da evolução das observações ao longo do tempo;

(2) Foram analisadas a FAC e FACP das séries originais a fim de evidenciar características quanto a estacionariedade e possíveis ordens do modelo de melhor ajuste;

(3) Foi realizado o teste de raiz unitária KPSS para avaliar a estacionariedade das séries analisadas. Quando verificado a não estacionariedade, ou seja, presença de raiz unitária, foi necessário a realização de sucessivas diferenças, $I(d)$, até tornar o processo estacionário;

(4) Foram novamente analisadas a FAC e a FACP, considerando as séries de órbitos na ordem $I(d)$ em que foi evidenciada a ausência de raiz unitária (estacionariedade), com a intenção de identificar as ordens p e q dos termos AR e MA do modelo ARIMA (p,d,q) ;

(5) Foram ajustados modelos ARIMA (p,d,q) e comparados a partir dos critérios de informação, AIC e BIC, a fim de identificar o modelo de maior parcimônia para cada região;

(6) Foi realizado o diagnóstico dos resíduos dos modelos mais parcimoniosos, analisando os pressupostos de ausência de autocorrelação, normalidade através do teste Jarque Bera (Jarque Bera, 1987) e utilização do teste $Q'(k)$ de Ljung-Box (Ljung-Box, 1978) para verificar se os resíduos são i.i.d;

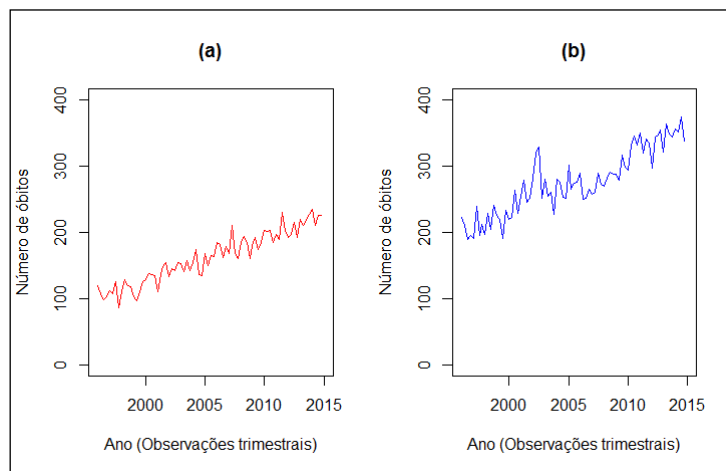
(7) Por fim, foram obtidas as previsões acerca do modelo de melhor ajuste, para as regiões Sul e Sudeste, e avaliada a qualidade das previsões por meio do MAPE e U de Theil (Theil, 1966);

Os gráficos, os testes realizados e os modelos ajustados foram obtidos utilizando o RStudio versão 1.0.136, software livre de ambiente de desenvolvimento integrado para R, utilizando os pacotes stats, graphics, astsa, urca, forecast e tseries. Quanto ao nível de significância, foi considerado 5% em todas as análises realizadas.

3 RESULTADOS E DISCUSSÃO

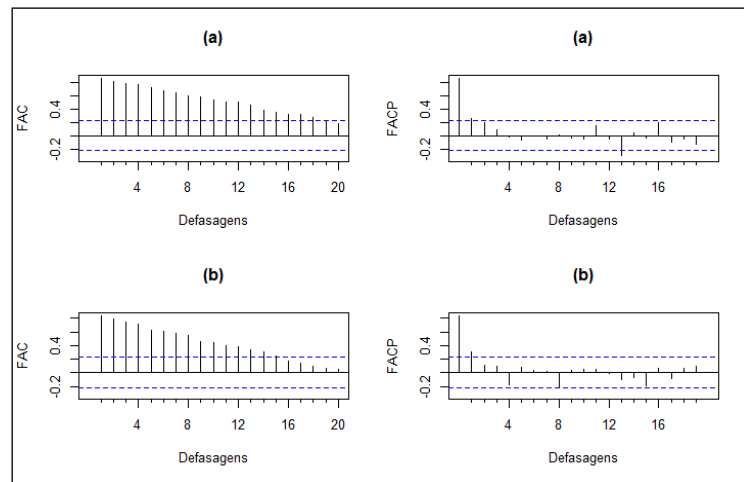
Analisou-se, inicialmente, os gráficos das séries trimestrais dos óbitos por neoplasias malignas de pele, entre 1996 e 2014, nas regiões Sul (a) e Sudeste (b), apresentados na Figura 1. Conforme observado, pode-se verificar uma possível tendência de crescimento em ambas as regiões, assim como a inexistência de sazonalidade nas séries. Essas mesmas características na evolução da série temporal foram observadas em Akhtar e Rozi (2009) e Liu *et al.* (2011).

Figura 1 – Série trimestral de óbitos por neoplasias malignas de pele nas regiões Sul (1a) e Sudeste (1b) do Brasil entre 1996 e 2014.



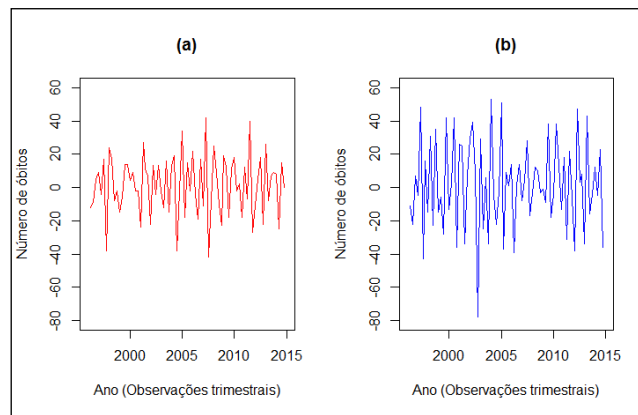
Quanto ao comportamento observado na FAC e na FACP (Figura 2) das séries das regiões analisadas, observou-se características de um processo não estacionário. Verificou-se um decaimento bastante lento na FAC, tanto na série da região Sul quanto na série da região Sudeste, cujas autocorrelações tornaram-se não significativas, respectivamente, apenas a partir das defasagens 20 e 16. Sendo assim, essas características reforçam a teoria de não estacionariedade na evolução das séries.

Figura 2 – Função de autocorrelação (FAC) e Função de autocorrelação parcial (FACP) da Série trimestral de óbitos por neoplasias malignas de pele nas regiões Sul (a) e Sudeste (b) do Brasil entre 1996 e 2014.



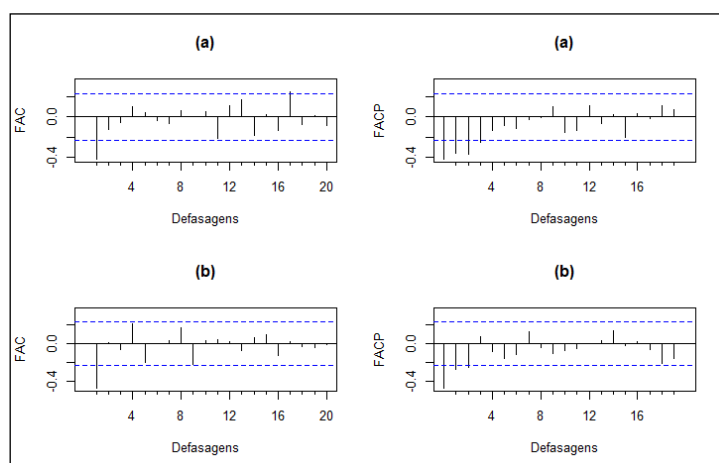
O Teste de raiz unitária KPSS truncado na segunda defasagem apresentou p -valor = 0,01 demonstrando que há evidências para rejeitar a estacionariedade da série da região Sul, ao nível de significância de 5%. Esse mesmo resultado foi obtido para a série da região Sudeste. Sendo assim, houve a necessidade de realização da primeira diferença em ambas as séries temporais afim de torna-las estacionárias e livres de tendência (Figura 3). Analisando-se os gráficos da Figura 3 observou-se que a primeira diferença das séries de óbitos por neoplasias malignas de pele ocorridos nas regiões Sul (a) e Sudeste (b) apresentam um comportamento em torno de uma média e variância constante, ou seja, a análise gráfica indica que as séries se tornaram estacionárias após a realização da primeira diferença. Essa afirmação foi comprovada pelo teste KPSS cujo p -valor $> 0,1$ para ambas as séries em primeira diferença. Assim como observado nesse estudo, nas séries analisadas por Akhtar e Rozi (2009) e Zhang et al (2017), a realização da primeira diferença também foi suficiente para tornar o processo estacionário. Esse resultado satisfaz a afirmação feita por Makridakis, Whellwright e Hyndman (1998) e Morettin e Toloí (2004), de que uma ou duas diferenças, em geral, seriam suficientes para tornar uma série temporal estacionária. Porém, as vezes pode fugir à regra, como pode ser observado em Liu *et al.* (2011), na qual foram necessárias três diferenças para estacionarizar a série de incidência de febre hemorrágica com síndrome renal na China.

Figura 3 – Primeira diferença da Série trimestral de óbitos por neoplasias malignas de pele nas regiões Sul (a) e Sudeste (b) do Brasil entre 1996 e 2014.



Sabe-se, partir do que foi realizado até então, que os modelos ajustados para as regiões Sul e Sudeste deve representam um ARIMA com p componentes AR, uma componente I e q componentes MA, ou seja, ARIMA ($p,1,q$). Utilizou-se a FAC e FACP (Figura 4) para indicar um possível modelo explicativo para as séries de óbitos por neoplasias malignas de pele ocorridos nas regiões Sul e Sudeste, em primeira diferença. A FACP e a FAC (Figura 4) da região Sul apresentam, respectivamente, 1 e 3 defasagens fora dos limites de confiança. Desse modo, o modelo ARIMA (3,1,1) é um possível candidato a melhor ajuste para a série dessa região. Para a região Sudeste, a FAC e a FACP (Figura 4), apresentam, respectivamente, 2 e 1 defasagens fora dos limites de confiança, indicando assim um ARIMA (2,1,1) como possível modelo de melhor ajuste.

Figura 4 – Função de autocorrelação (FAC) e Função de autocorrelação parcial (FACP) da primeira diferença da série trimestral de óbitos por neoplasias malignas de pele em primeira diferença nas regiões Sul (a) e Sudeste (b) do Brasil, entre 1996 e 2014.



Após o ajuste de diversos modelos com número de parâmetros em torno dos modelos indicados na análise da Figura 4, foi constatado que os modelos com melhores ajustes para as séries das regiões Sul e Sudeste foram os apresentados no Quadro 1. Entre os três modelos apresentados para cada região (Quadro 3) o modelo ARIMA (4,1,0) com constante mostrou-se o mais parcimonioso considerando o BIC para a série de dados da região Sul, e o modelo ARIMA (0,1,1) com constante apresentou maior parcimônia para a série de dados da região Sudeste.

Com isso, pode-se observar que os modelos ajustados (Quadro 2) considerados os mais parcimoniosos para as séries de órbitos das regiões Sul e Sudeste, são explicados por diferentes componentes. A região Sul não apresentou componentes de MA, enquanto a região Sudeste não apresentou componentes AR.

O modelo justado para a região Sudeste apresentou semelhanças com o modelo ajustado por Liu et al. (2011), apenas divergem no número de diferenças necessárias para estacionarizar a série, mas em ambos os modelos uma componente MA foi suficiente para explicar a evolução das séries. O modelo ajustado para a região Sul não possui semelhanças com nenhum dos ajustados por Earnest et al. (2005), Akhtar e Rozi (2009) e ZHANG et. al (2017), em todos esses estudos estão presentes tanto componentes AR, quanto componentes MA.

Quadro 1 – Critério de informação Bayesiano (BIC) dos três modelos de melhor ajuste as séries de órbitos por neoplasias malignas de pele nas regiões sul e sudeste, entre 1996 e 2014.

Modelos	Região Sul	Modelos	Região sudeste
ARIMA(4,1,0) com constante	619,46	ARIMA(0,1,1) com constante	677,60
ARIMA(3,1,0) com constante	620,66	ARIMA(1,1,1) com constante	681,92
ARIMA(5,1,0) com constante	621,66	ARIMA(0,1,2) com constante	681,92

Quadro 2 – Parâmetros estimados dos modelos ARIMA(4,1,0) com constante (Região Sul) e ARIMA(0,1,1) com constante (Região Sudeste).

Modelo Ajustado (Região Sul)	Parâmetros estimados
ARIMA(4,1,0) com constante	Constante = 1,6014
	$\hat{\phi}_1 = -0,8150$
	$\hat{\phi}_2 = -0,7382$
	$\hat{\phi}_3 = -0,5957$
	$\hat{\phi}_4 = -0,2686$
Modelo Ajustado (Região Sudeste)	Parâmetros estimados
ARIMA(0,1,1) com constante	Constante = 2,0297
	$\hat{\theta}_1 = -0,7267$

Após a escolha do modelo mais parcimonioso para cada região, foi realizada a análise de diagnóstico dos resíduos desses modelos. Na Figura 5 os resíduos dos modelos podem ser observados graficamente e na Figura 6 as FAC e FACP, onde verificou-se que os resíduos estão dentro dos limites de confiança.

Figura 5 – Resíduos dos modelos ajustados para as séries de óbitos por neoplasias malignas de pele nas regiões Sul (a) e Sudeste (b) do Brasil.

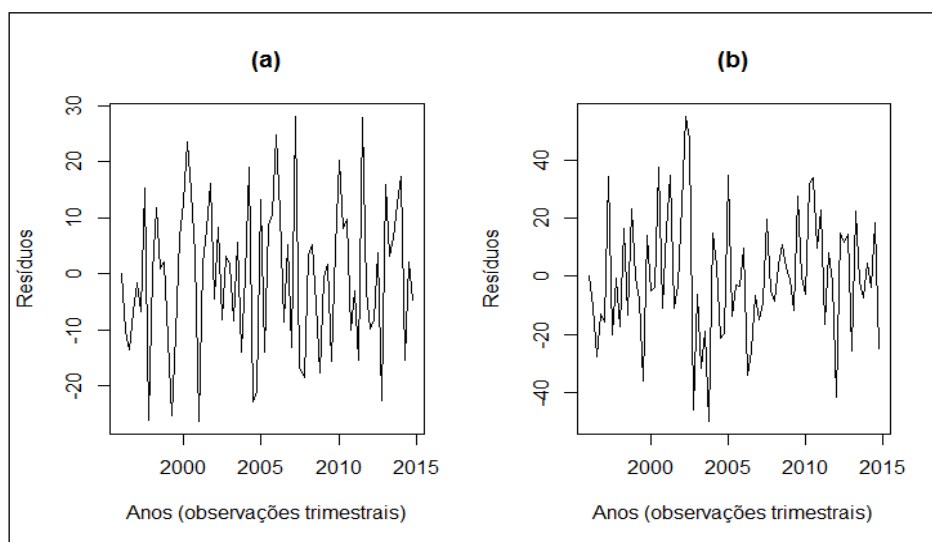
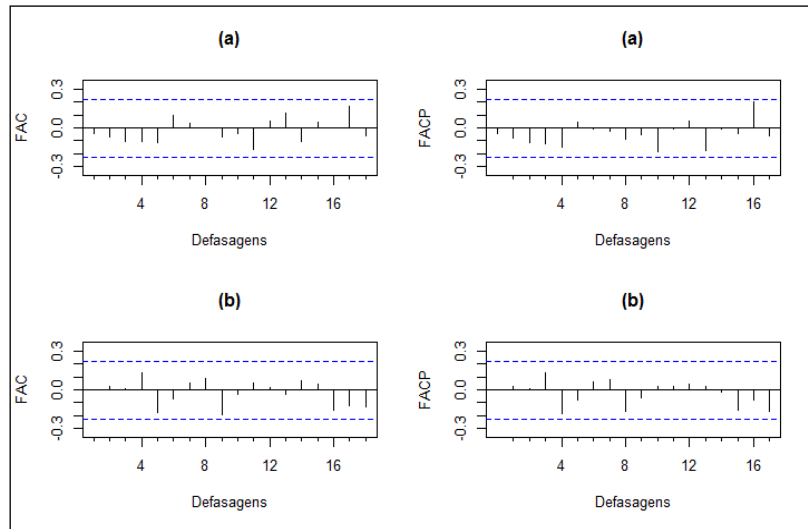


Figura 6 – Função de autocorrelação (FAC) e Função de autocorrelação parcial (FACP) dos resíduos dos modelos ARIMA ajustados para os óbitos por neoplasias malignas de pele nas regiões Sul (a) e Sudeste (b)

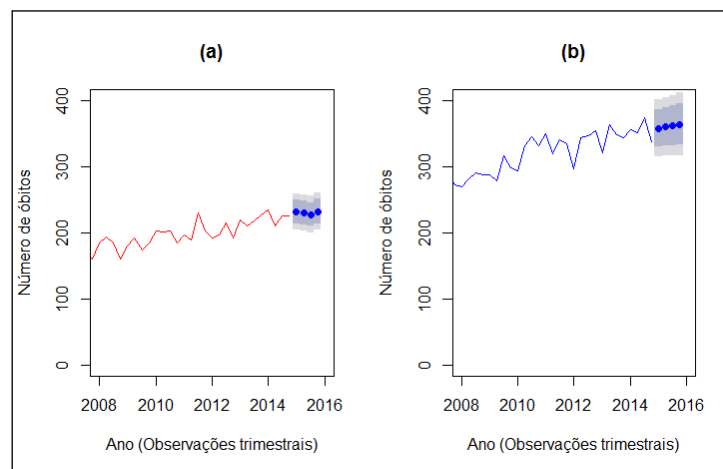


O resultado do teste Jarque-Bera, para os resíduos, apresentou p-valor = 0,5617 e p-valor = 0,8152, respectivamente, dos modelos ajustados para as regiões Sul e Sudeste. Logo, pode-se inferir que não há evidências para rejeitar a hipótese de normalidade dos resíduos dos modelos ARIMA (4,1,0) com constante e ARIMA (0,1,1) com constante, considerando um nível de significância de 5%. O resultado do teste Q' de Ljung-Box (Ljung-Box ,1978) resultou em um p-valor igual a 0,7056 e 0,5511, para os respectivos resíduos dos modelos das regiões Sul e Sudeste. Com isso, também pode-se inferir que não há evidências para rejeitar a hipótese de que os resíduos são, i.i.d , ou seja, estão ausentes de autocorrelação. Sendo assim, conclui-se que os modelos ajustados são adequados para explicar as séries de óbitos por neoplasias malignas de pele nas regiões Sul e Sudeste. O último procedimento da metodologia de Box-Jenkins diz respeito a obtenção de previsões acerca do modelo escolhido. As previsões de óbitos por neoplasias malignas de pele obtidas para os quatro trimestres de 2015 nas regiões Sul (a) e Sudeste (b) estão descritas no Quadro 3 e graficamente na Figura 7.

Quadro 3 – Previsões dos óbitos por neoplasias malignas de pele para os trimestres do ano de 2015 nas regiões Sul e Sudeste considerando, respectivamente, os modelos ARIMA(4,1,0) com constante e ARIMA(0,1,1).

Região Sul	Óbitos reais	Previsões	Limite inferior	Limite Superior
1º Trimestre	213	231,6174	204,7569	258,4779
2º Trimestre	229	229,7422	202,4261	257,0582
3º Trimestre	252	227,7973	200,3187	255,2758
4º Trimestre	235	232,3967	204,4132	260,3801
Região Sudeste	Óbitos reais	Previsões	Limite inferior	Limite Superior
1º Trimestre	377	357,9866	315,3216	400,6517
2º Trimestre	383	360,0163	315,7870	404,2456
3º Trimestre	413	372,0460	326,3059	417,7861
4º Trimestre	399	364,0757	316,8731	411,2783

Figura 7- Gráfico de previsões para as séries trimestral de óbitos por neoplasias malignas ocorridos nas regiões Sul (a) e Sudeste (b).



4 CONCLUSÕES

Conforme observado, este estudo possibilitou evidenciar algumas características como presença de tendência crescente, ausência de sazonalidade e comportamento não estacionário nas séries de óbitos por neoplasias malignas de pele tanto na região Sul, quanto na região Sudeste. O modelo de melhor ajuste e mais parcimonioso para os óbitos por neoplasias malignas de pele na região Sul foi o ARIMA (4,1,0) com constante e o modelo de melhor ajuste para a região Sudeste foi o ARIMA (0,1,1). Em ambos os modelos, as previsões obtidas apresentaram baixos

MAPE's (respectivamente, 6,29% e 7,13%) e mostraram um contínuo crescimento nos óbitos. Quanto a comparação dos modelos obtidos para as regiões Sul e Sudeste, observou-se a presença de diferentes componentes. Os óbitos por neoplasias malignas de pele na região Sul foram explicados apenas por componentes AR e os óbitos na região Sudeste apenas por componentes MA, sendo ambas com a necessidade de uma diferenciação para tornar a série estacionária.

Em geral, os modelos ARIMA apresentaram boa qualidade de ajuste e de previsões para as séries de óbitos por neoplasias malignas de pele nas regiões Sul e Sudeste do Brasil. A metodologia de Box-Jenkins mostrou-se bastante satisfatória, no monitoramento e detecção de padrões da evolução das séries analisadas. Sendo assim, os modelos ajustados podem ser utilizados como justificativa para a formulação de novas políticas públicas proporcionando combater, prevenir e/ou estabilizar a mortalidade por neoplasias malignas de pele, não somente nas regiões Sul e Sudeste, mas também nas demais regiões brasileiras.

É importante salientar que este foi o primeiro estudo relacionado as neoplasias malignas de pele que se utiliza de modelos de previsão, mais especificamente os modelos ARIMA. Porém, é sabido que existem modelos de previsão mais recentes e que possivelmente consigam ajustar-se melhor as variações das séries analisadas. Por isso, sugere-se a modelagem dessas séries utilizando os modelos GARMA (p,q) e GLARMA (p,q) pois, como as séries temporais são de dados de contagem, esses podem reduzir os resíduos dos modelos e melhorar ainda mais a qualidade das previsões.

REFERÊNCIAS

- AKHTAR, S.; ROZI, S.. An autoregressive integrated moving average model for short-term prediction of hepatitis C virus seropositivity among male volunteer blood donors in Karachi, Pakistan. **World Journal of Gastroenterology: WJG**, v. 15, n. 13, p. 1607, 2009.
- BOX, G. E. P.; JENKINS, G. M. **Time series analysis: forecasting and control**. San Francisco: Ed: Holden-Day, 575 p., 1970.
- BOX, G. E. P.; JENKINS, G. M.; REINSEL, G. C. **Time series analysis: forecasting and control, 3º ed**. Prentice Hall: New Jersey, 589 p., 1994.
- JARQUE, C. M.; BERA, A. K. A test for normality of observations and regression residuals. **International Statistical Review**, Woodbury, v.55, p. 163-172, 1987.
- LJUNG, G.; BOX, G. E. P. On a measure of lack off it in time series models. **Biometrika**, v. 65, p. 297-303, 1978.
- LIU, Q. et al. Forecasting incidence of hemorrhagic fever with renal syndrome in China using ARIMA model. **BMC infectious diseases**, v. 11, n. 1, p. 218, 2011.
- MADDALA, G. S. **Introduction to Econometrics**. 3.ed. New York: John Wiley & Sons, 664 p., 2001.
- MAKRIDAKIS, S.; WHEELWRIGHT, S. C.; HYNDMAN, R. J. **Forecasting: Methods and Applications**. 3rd ed. Ed. Wiley & Sons, 642 p., 1998.
- MATHERS, C. D. et al. Counting the dead and what they died from: an assessment of the global status of cause of death data. **Bulletin of the world health organization**, v. 83, n. 3, p. 171-177c, 2005.
- MATOS, K. S. L de; VIEIRA, S. L.; **Pesquisa educacional: o prazer de conhecer**. Fortaleza: Demócrito Rocha, 143p., 2001.
- MORETTIN, P. A.; TOLOI, C. M. C. **Análise de Séries Temporais**. São Paulo: Blücher, 535 p., 2004.
- SOUZA, R. C.; CAMARGO, M. E. **Análise de Séries Temporais: Os Modelos ARIMA, Ijuí: SEDIGRAF**. 242 p., 1996.
- THEIL, H. **Applied economic forecasting**. Amsterdam: North-Holland, 1966.
- WHITTLE, P. **Hypotesis testing in the time series analysis**. Thesis, Uppsala University, Almqvist and Wiksell, Uppsala, 1951.

ZHANG, X. et al. Tendência preditiva das caries da primeira infância na China continental: uma abordagem combinada de modelagem meta-analítica e matemática baseada em pesquisas epidemiológicas. **Scientific Reports**, v. 7, n. 1, p. 6507, 2017.