

MODELOS DE SÉRIES TEMPORAIS APLICADOS A PREVISÃO DE RADIAÇÃO SOLAR

Time series models applied to solar radiation forecasting

Eliezer José Balbinot¹; Jaque Willian Scotton²; Simone Maffini Cerezer³;
Claodomir Antonio Martinazzo⁴.

¹ Mestrando do Programa de Pós-Graduação em Modelagem Matemática, Universidade Regional do Noroeste do Estado do Rio Grande do Sul, E-mail: eliezer.balbinot@hotmail.com

² Doutorando do Programa de Pós-Graduação em Modelagem Computacional, Universidade Federal do Rio Grande.

³ Professora do Instituto Federal de Educação, Ciência e Tecnologia do Rio Grande do Sul - Campus Canoas.

⁴ Professor do Departamento de Ciências Exatas e da Terra, Universidade Regional Integrada do Alto Uruguai e das Missões.

Data do recebimento: 12/04/2017 - Data do aceite: 30/06/2017

RESUMO: O presente trabalho tem por objetivo investigar a possibilidade de utilizar os modelos Box e Jenkins para descrever o comportamento da série média dos valores de radiação global horizontal média diária mensal (MJ/m^2) na superfície terrestre por ano, registrados para o Leste dos Estados Unidos da América, no período de 1961 a 1990, e disponíveis no banco de dados SAMSON. Através da análise estatística realizada, utilizando-se os critérios de Akaike e Schwarz, os modelos escolhidos que melhor descrevem o comportamento da série analisada, livre de sazonalidade, foram o AR(1) e MA(1). A análise estatística foi realizada com o auxílio do software Minitab 15.0.

Palavras-chave: Séries Temporais. Radiação Solar. Modelos Box e Jenkins.

ABSTRACT: This work aims to investigate the possibility of using the Box and Jenkins models to describe the behavior of the average series values of daily and monthly average of the horizontal global radiation (MJ / m^2) on land surface per year, recorded for the Eastern of the United States of America, from 1961 to 1990, and available in the SAMSON database. Through the statistical analysis, using the Akaike and Schwarz criteria, the chosen models

that best describe the behavior of the analyzed series, seasonally free, were the AR (1) and the MA (1). Statistical analysis was performed using the software Minitab 15.0.

Keywords: Time series. Solar radiation. Models Box and Jenkins.

Introdução

A radiação solar desempenha um importante papel para a vida humana. É por meio dela que diversos fenômenos meteorológicos, tanto de pequena quanto de larga escala, são desencadeados. A energia do sol, por exemplo, faz com que ocorra a evaporação, permitindo que o ciclo das águas se desenvolva. Além de ter fundamental importância na circulação atmosférica, a radiação é ainda uma alternativa energética inesgotável, sendo uma excelente fonte de luz e de calor.

Com base na importância da radiação solar e considerando as transformações climáticas que vem ocorrendo no planeta, é fundamental realizar pesquisas nessa área, tendo em vista que o aumento da radiação influencia diretamente a vida de todos os seres vivos. Para prever o índice de radiação solar, existem ramos da estatística que fornecem a possibilidade de analisar dados e prever o comportamento do objeto em estudo.

Pesquisadores como Silva et al. (2010); Anjos (2013); Ceballos e Macedo (2014), Zhitorchuk e Shanina (1994), têm se dedicado ao estudo da radiação solar através de modelos de séries temporais. Tais modelos possibilitam, através do estudo do comportamento das variações ocorridas no passado, prever futuras variações, além de prever os futuros valores dos dados em estudo. (MORETTIN; TOLOI, 2004).

Neste sentido, o presente trabalho visa investigar a possibilidade de utilizar os modelos Box e Jenkins (1976) para descrever valores de radiação solar. Esta investigação foi realizada com o auxílio do software

Minitab 15.0. Verificou-se, inicialmente, o comportamento da série em estudo, através da construção do gráfico estabelecido pela média dos valores de radiação solar global horizontal, média diária mensal (MJ/m^2) na superfície terrestre por ano. A próxima etapa do trabalho consistiu-se em desconsiderar a componente sazonal e ajustar a série dessazonalizada aos diferentes modelos de séries temporais. Por meio de testes estatísticos, identificaram-se dois modelos que melhor descreveram os valores da série e, após identificados esses modelos, realizaram-se as previsões de valores de radiação solar correspondentes às médias mensais de 1990.

Material e Métodos

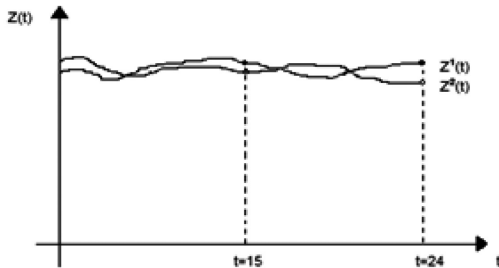
Séries Temporais

Séries temporais são conjuntos de observações ordenadas no tempo, que podem ser contínuas ou discretas. Uma série é dita contínua quando as observações são feitas continuamente no tempo, porém, quando as medidas são tomadas em intervalos de tempo regularmente espaçados, a série é denominada discreta. (MORETTIN; TOLOI, 2004).

Podem-se encontrar exemplos de séries temporais nos mais diversos campos de domínio. Em economia pode-se falar de taxas de juro, produto interno bruto; em meteorologia, registro de temperatura diária, registro de precipitação mensal; em medicina, leitura de eletrocardiograma, leitura de eletroencefalograma, entre outros campos. Em relação à representação gráfica de uma série temporal,

Spiegel (1994) afirma que o gráfico é descrito por um ponto que se move com o decorrer do tempo. A Figura 1 apresenta um exemplo da representação gráfica de duas séries temporais, $Z^1(t)$ e $Z^2(t)$.

Figura 1 - Representação gráfica de séries temporais.



Fonte: Morretin e Toloí (2004).

Elementos das Séries Temporais

Conforme descrito em Morettin e Toloí (2004), as séries temporais são compostas por quatro elementos básicos, também denominados componentes de uma série temporal, que podem ser classificados em:

- a) tendência: refere-se ao movimento dos dados a longo prazo, para cima ou para baixo;
- b) variações cíclicas: variações com certo grau de regularidade, porém com período diferente de um ano;
- c) variações sazonais: os fenômenos sazonais estão associados às estações do ano, difere das variações cíclicas por considerar o período de um ano;
- d) variações irregulares: são variações aleatórias, que não apresentam regularidade.

Segundo Reboita (2005), nos estudos climatológicos é interessante remover a componente sazonal das séries temporais, pois

ela é muito intensa o que acaba mascarando as outras componentes das séries. A autora afirma, ainda, que há duas finalidades para o fato de se isolar a componente sazonal de uma série temporal. Uma é remover o padrão a fim de estudar as variações cíclicas e irregulares e a outra é identificar os fatores sazonais de forma que eles possam ser levados em conta na tomada de decisões.

Tipos de Modelos

Os modelos para séries temporais podem ser classificados em modelos paramétricos, nos quais o número de parâmetros é finito e, modelos não-paramétricos, os quais envolvem um número infinito de parâmetros. Na classe de modelos paramétricos, a análise é feita no domínio do tempo. Dentre estes modelos, os mais frequentemente utilizados são os modelos de erro, o auto-regressivo e de média móvel (ARMA), o modelo auto-regressivo integrado e de média móvel (ARIMA), modelos de memória longa (ARFIMA), modelos estruturais e modelos não-lineares. Os modelos não-paramétricos mais usados são a função de auto-covariância e sua transformada de Fourier. (MORETTIN; TOLOI, 2004).

Neste artigo, apresentam-se os modelos ARIMA de Box e Jenkins (1976), envolvendo seus casos particulares, ou seja, o modelo auto-regressivo, média móvel e auto-regressivo e de média móvel.

Modelo Auto-Regressivo

Segundo Silva (2005), o modelo auto-regressivo (*Auto-Regressive*) foi criado com a ideia de que a presente observação da série Z_t pode ser explicada como uma função das p observações passadas, $Z_{t-1}, Z_{t-2}, \dots, Z_{t-p}$, em que p determina o número de passos entre as observações passadas e a previsão da próxima observação.

Com base em Morettin e Tolo (2004), denota-se o modelo auto-regressivo de ordem p por AR (p) e, expressa-se a estrutura auto-regressiva por:

$$Z_t = \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + \dots + \phi_p Z_{t-p} + a_t$$

sendo, ϕ_i são parâmetros da estrutura, $i = 1, \dots, p$ (ordem da estrutura); a_t é o ruído branco com média zero e variância σ_a^2 .

Modelo de Média Móvel

O modelo de média móvel é formado pela combinação linear do ruído branco, a_t , ocorrido no período corrente e nos períodos passados. (SILVA, 2005). Conforme apresentado em Morettin e Tolo (2004), denota-se o modelo média móvel de ordem q , por MA (q) (*Moving Average*) e expressa-se a estrutura média móvel por:

$$Z_t = \mu + a_t - \theta_1 a_{t-1} - \dots - \theta_q a_{t-q}$$

em que, θ_i são parâmetros da estrutura, $i = 1, \dots, p$ (ordem da estrutura); a_t é o ruído branco com média zero e variância σ_a^2 .

Modelo Auto-Regressivo e de Média Móvel

O modelo auto-regressivo e de média móvel (*Auto-Regressive Moving Average*), designado por ARMA, representa a combinação dos modelos AR e MA, em que Z_t é descrito por seus valores passados e pelos ruídos branco corrente e passados. (SILVA, 2005). A estrutura geral ARMA (p, q) é expressa por:

$$Z_t = \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + \dots + \phi_p Z_{t-p} + a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \dots - \theta_q a_{t-q}$$

Modelo Auto-Regressivo Integrado e de Média Móvel

Segundo Carvalho (2005), no modelo ARIMA (*Auto-Regressive Integratd Moving Average*), a série temporal é gerada por um processo estocástico cuja natureza pode ser representada através de um modelo. A notação utilizada para designar esse modelo é ARIMA (p, d, q), em que p é o número de termos auto-regressivo, d o número de diferenciações para que a série torne-se estacionária e q o número de termos de média móvel. A estrutura geral ARIMA (p, d, q) é expressa por:

$$\phi(B)\Delta^d Z_t = \theta(B)a_t$$

sendo, $\phi(B)$ representa o operador auto-regressivo de ordem p ; $\theta(B)$ representa o operador média móvel de ordem q ; a_t é o ruído branco; $\Delta = 1 - B$ representa o operador diferença.

Uma metodologia bastante utilizada na análise de modelos paramétricos é conhecida como abordagem de Box e Jenkins, no qual consiste em ajustar os modelos ARIMA (p, d, q) a conjunto de dados. Segundo Silva (2005), o objetivo da metodologia de Box e Jenkins é determinar os três componentes (p, d, q) que configuram qualquer estrutura do modelo. O procedimento para utilização da metodologia de previsão consiste nos seguintes passos:

- a) Realização da diferenciação da série original, para torná-la estacionária;
- b) Identificação dos valores p e q através da análise das funções de autocorrelação e de autocorrelação parcial estimadas e da estimação dos parâmetros $\phi_i, i = 1, \dots, p$ e $\theta_i, i = 1, \dots, q$. A análise destas funções permite a tomada de decisão acerca do número de parâmetros auto-

regressivo e do número de parâmetros de média móvel que devem ser escolhidos de maneira a se obter uma estrutura parcimoniosa, ou seja, conter um número pequeno de parâmetros dentre todas as estruturas que se ajustem aos dados da série;

- c) Realização das previsões (obtendo os novos valores da série) e os intervalos de confiança para as mesmas.

Sazonalidade

A sazonalidade representa os fenômenos que ocorrem no período máximo de um ano. Segundo Morettin e Toloi (2004), observa-se que em séries sazonais ocorrem relações entre observações para meses sucessivos em um ano particular e entre as observações para o mesmo mês em anos sucessivos.

Denota-se por *período*, mesmo que o padrão não seja exatamente periódico. Assim, Z_t é relacionada com Z_{t-1}, Z_{t-2}, \dots , mas também com Z_{t-s}, Z_{t-2s}, \dots , implicando, deste modo, que as séries sazonais são caracterizadas por apresentarem correlação alta em *lags* sazonais, isto é, *lags* que são múltiplos do período s . No entanto, um ajuste sazonal, removerá esta correlação ou pelo menos removerá grande parte dela.

Função de Autocorrelação e Autocorrelação Parcial

Segundo Reboita (2005), a função de autocorrelação (fac) é a ferramenta primária para a caracterização de relações entre valores de dados na aproximação do domínio temporal. A fac, denotada por $\rho(\tau)$, é definida como:

$$\rho(\tau) = \frac{\gamma(\tau)}{\gamma(0)}, \text{ onde } \gamma(0) = \text{Var}(Z_t).$$

Logo, $\rho(\tau)$ será uma medida padrão de dependência com $|\rho(\tau)| \leq 1$, para $\tau \in Z$. Como

na prática se tem somente uma amostra de observações, o coeficiente de autocorrelação é estimado por:

$$\hat{\rho}(\tau) = \frac{\sum_{t=0}^{n-\tau} (Z_t - \bar{Z}) \times (Z_{t-\tau} - \bar{Z})}{\sum_{t=0}^{n-\tau} (Z_{t-\tau} - \bar{Z})^2},$$

sendo n é o número de observações da amostra.

A função que associa cada valor “ τ ” com o seu respectivo coeficiente de autocorrelação é chamada de função de autocorrelação. Seu gráfico é comumente chamado de correlograma.

Segundo Morettin e Toloi (2004), os processos AR (p), MA (q) e ARMA (p, q) apresentam fac com características especiais, as quais serão úteis no processo de identificação do modelo. Deste modo:

- (i) um processo AR (p) tem fac que decai de acordo com exponenciais e/ou senóides amortecidas, infinita em extensão;
- (ii) um processo MA (q) tem fac finita, no sentido que ela apresenta um corte após o *lag* q ;
- (iii) um processo ARMA (p, q) tem fac infinita em extensão, a qual decai de acordo com exponenciais e/ou senóides amortecidas após o *lag* $q - p$.

Outro instrumento utilizado para facilitar o procedimento de identificação do modelo é a função de autocorrelação parcial (facp). Uma descrição detalhada da facp pode ser encontrada em Morettin e Toloi (2004).

Crítérios de Informação de Akaike (AIC)

O Critério de Akaike (1974) é recomendado quando $n / N \geq 40$, ou seja, quando o resultado da divisão entre o número de termos da série pelo número de parâmetros

do modelo ajustado for maior ou igual a 40. (BELLO, 2010). Deste modo, o cálculo do AIC é determinado por:

$$AIC = n \times \ln(\hat{\sigma}_N^2) + 2 \times (N + 1)$$

sendo, $\hat{\sigma}_N^2$ é o estimador de máxima verossimilhança da variância do erro, expresso por:

$$\hat{\sigma}_N^2 = \frac{\sum_{i=1}^n (y_i - \hat{\mu}_i)^2}{n}$$

em que, y_i é o i -ésimo valor da resposta; $\hat{\mu}_i$ é a estimativa de y_i ; n é o número de termos da série; N é o número de parâmetros do modelo ajustado.

Critério de Informação de Schwarz (SIC)

Outro critério de informação para a seleção de modelos é apresentado em Schwarz (1978), definido por:

$$SIC = \ln(\hat{\sigma}_N^2) + \left[\left(\frac{\ln n}{n} \right) \times (N) \right]$$

sendo, n o número de termos da série; $\hat{\sigma}_N^2$ é o estimador de máxima verossimilhança da variância do erro; N é o número de parâmetros do modelo ajustado.

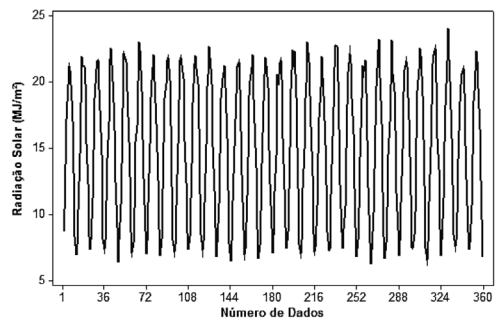
Resultados e Discussão

Com o objetivo de utilizar modelos de séries temporais para descrever valores de radiação solar, foram reunidos, inicialmente, dados do período de 1961 a 1990 do banco de dados SAMSON (*Solar and Meteorological Surface Observational Network*), para o Leste dos Estados Unidos da América (EUA), fornecidos pelo Centro Nacional de Dados Climáticos (Asheville, EUA). Os dados utilizados foram medidos em 79 estações de coleta, registrados como valores horários

e disponibilizados como médias mensais, totalizando 28440 médias. Algumas das principais grandezas registradas pelas estações foram a radiação extraterrestre global (Wh/m²); radiação horizontal global (Wh/m²); nebulosidade (parcela do céu coberta por nuvens – em décimo); temperatura de bulbo úmido (°C); umidade relativa (em percentuais); velocidade do vento (m/s); altitude (m) e latitude (°, ', ").

Para o ajuste dos diferentes modelos, considerou-se, neste trabalho, a série temporal formada pela média dos valores de radiação global horizontal média diária mensal (MJ/m²) na superfície terrestre por ano (Figura 2). Esse ajuste foi realizado com o auxílio do software Minitab 15.0.

Figura 2 - Média dos valores de radiação global horizontal média diária mensal (MJ/m²) na superfície terrestre por ano.



Com base na Figura 2, observa-se que a série não apresenta tendência, o que permite considerar a série estacionária. Observa-se também um aparente comportamento sazonal, que será identificado com a análise do correlograma.

As Figuras 3 (a) e 3 (b) apresentam a função de autocorrelação e a função de autocorrelação parcial dos valores de radiação solar para o Leste dos EUA, respectivamente. Analisando esses gráficos, observa-se que a série possui um comportamento aproximadamente periódico, havendo semelhanças a cada $s = 12$ meses. Para este comportamento senoidal, típico de um modelo sazonal, é

aconselhável remover a componente sazonal, conforme descrito em Reboita (2005). O software Minitab 15.0 utiliza o método da média móvel centrada de 12 meses, que consiste em retirar além das influências sazonais, as irregularidades e alguns ciclos. A Figura 4 apresenta a série livre de sazonalidade, resultado comprovado pela fac (Figura 5 (a)) e facp (Figura 5 (b)) da série dessazonalizada.

Figura 3 - Função de autocorrelação (a) e função de autocorrelação parcial (b), para a série de radiação global horizontal média diária mensal (MJ/m²) na superfície terrestre por ano.

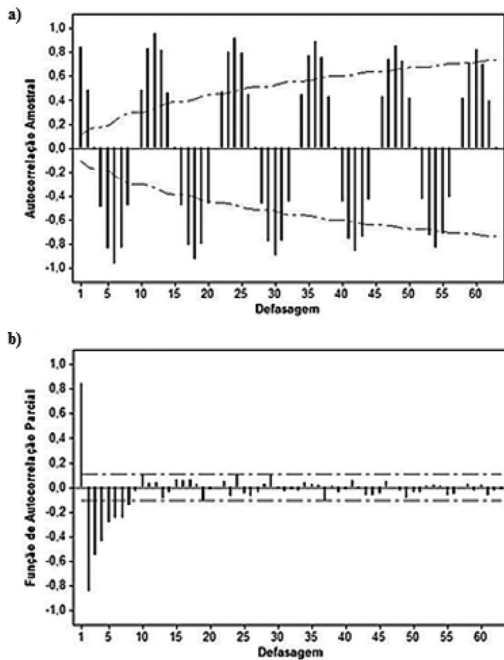


Figura 4 - Média dos valores dessazonalizados de radiação global horizontal média diária mensal (MJ/m²) na superfície terrestre por ano.

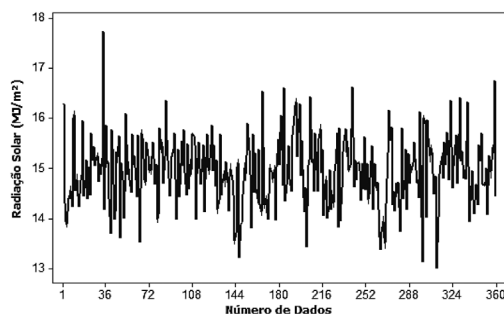
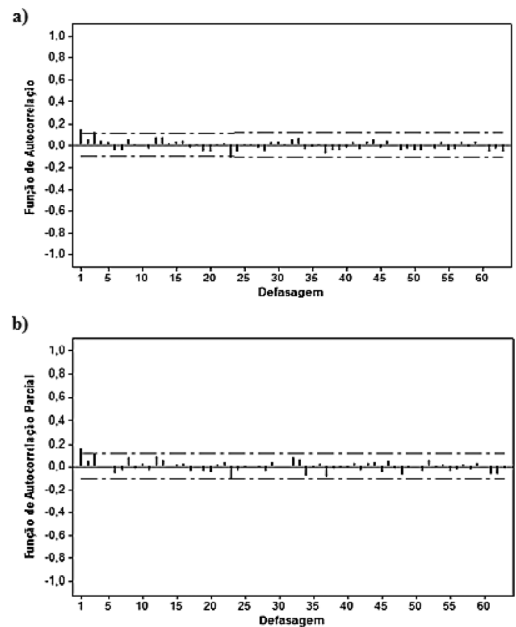


Figura 5 - Função de autocorrelação (a) e função de autocorrelação parcial (b) para a série dessazonalizada de radiação global horizontal média diária mensal (MJ/m²) na superfície terrestre por ano.



Verificação da Qualidade do Ajuste para Diferentes Modelos de Séries Temporais

Para a escolha do modelo apropriado, ajustou-se a série dessazonalizada aos seguintes modelos: AR (1), AR (2), MA (1), MA (2), ARMA (1,1), ARMA (1,2), ARMA (2,1) e ARMA (2,2). Posteriormente, com base nos critérios de comparação de Akaike (1974) e de Schwarz (1978), determinou-se os melhores modelos. A Tabela I apresenta os resultados obtidos.

Analisando os dados apresentados na Tabela I, percebe-se que os modelos AR (1) e MA (1) apresentaram os menores valores para os critérios de AIC e SIC, sendo escolhidos para descrever a série dos valores dessazonalizados de radiação global horizontal média diária mensal (MJ/m²) na superfície terrestre por ano (Figura 4). A Tabela II apresenta os

parâmetros estimados e seus respectivos p-valores, erro médio absoluto (MAE) e erro quadrático médio (EQM) dos dois modelos escolhidos.

Tabela I - Critérios de comparação para a escolha do melhor modelo.

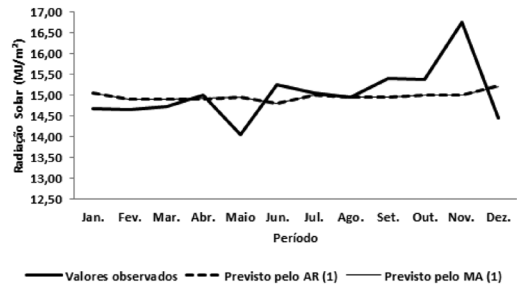
Modelo	AIC	SIC
AR (1)	-280,44	-0,78
AR (2)	-278,83	-0,76
MA (1)	-279,88	-0,77
MA (2)	-278,12	-0,76
ARMA (1,1)	-280,26	-0,76
ARMA (1,2)	-276,12	-0,74
ARMA (2,1)	-278,34	-0,75
ARMA (2,2)	-277,46	-0,74

Tabela II - Estimativas dos parâmetros, p-valor, MAE e EQM dos dois modelos escolhidos.

Modelo	Estimativas dos Parâmetros	p-valor	MAE	EQM
AR (1)	$\phi_1 = 0,1524$		0,48	0,48
	Constante = 12,6657	0,004 0,000		
MA (1)	$\theta_1 = -0,1459$		0,47	0,49
	Constante = 14,943	0,006		

Considerando os modelos ajustados, realizaram-se previsões dos valores de radiação solar correspondente às médias mensais de 1990. A Figura 6 apresenta a comparação entre os dados observados e previstos pelos modelos escolhidos.

Figura 6 - Valores observados e previstos pelos modelos AR (1) e MA (1) de radiação global horizontal média diária mensal (MJ/m^2) na superfície terrestre para os meses de Jan. a Dez. do ano de 1990.



Considerações Finais

Neste trabalho foram abordados modelos de séries temporais para descrever valores de radiação solar registrados para o Leste dos EUA no período de 1961 a 1990 para 79 estações de coleta. Os resultados obtidos, através dos critérios de AIC e SIC, foram determinantes para a escolha dos modelos AR (1) e MA (1).

As previsões realizadas foram avaliadas com duas medidas de erro utilizadas na literatura para previsões de séries temporais, sendo elas o EQM e o MAE. Pode-se perceber através destas medidas que a eficiência nas previsões está diretamente relacionada aos parâmetros de cada modelo. Ressalta-se, também, que para estudos climatológicos é interessante considerar a série dessazonalizada, ideia desenvolvida por Reboita (2005), pois a componente sazonal pode mascarar outros elementos da série.

REFERÊNCIAS

- ANJOS, P. S. **Correlações de longo alcance em séries temporais da velocidade do vento e radiação solar em Fernando de Noronha, Brasil**. 2013. 63 f. Dissertação de mestrado (Mestrado em Biometria e Estatística Aplicada) – Universidade Federal Rural de Pernambuco, Recife, 2013.
- AKAIKE, H. **A new look at the statistical model identification**. IEEE Transactions on Automatic Control., Boston, v.19, n.6, p.716-723, 1974.
- BELLO, L. H. A. Critérios de Informação e Seleção de Modelos. In: **Modelagem em Experimentos Mistura-Processo para Otimização de Processos Industriais**, 2010, pg. 93 – 112.
- CARVALHO, J. V. Análise de Séries Temporais. In: **Modelagem Temporal das medidas de vazão de drenos na Barragem de Funil (RJ) utilizando Redes Neurais e Métodos Estatísticos**, 2005, p. 29 - 70.
- CEBALLOS, J. C.; MACEDO, L. O. Uma base de dados de radiação solar na América do Sul estimada por satélite (Modelo GL1.2/CPTEC). In: CONGRESSO BRASILEIRO DE ENERGIA SOLAR, 2014, Recife. **Anais...** Recife:[s.n.], 2014.
- MORETTIN, P. A.; TOLOI, C. M. C. **Análise de Séries Temporais**. São Paulo: Edgard Blücher, 2004.
- SILVA, R. A., et al. Estudo da variabilidade da radiação solar no nordeste do Brasil. **Revista Brasileira de Engenharia Agrícola e Ambiental**, v. 14, n. 5, p. 501-509, 2010.
- REBOITA, M. S. **Introdução à Estatística Aplicada à Climatologia: Análise de Séries Temporais**, São Paulo, 2005.
- SCHWARZ, G. Estimating the dimensional of a model. **Annals of Statistics**, v.6, n.2, p.461- 464, 1978.
- SILVA, M. O. Metodologia de Previsão de Séries Temporais- Box & Jenkins. In: **Uma aplicação de árvores de decisão, rede neurais e knn para a identificação de modelos ARMA não-sazonais e sazonais**, 2005, p.15- 40.
- SPIEGEL, M. R. **Estatística**. 3. ed. São Paulo: Makron Books, 1994.
- ZHITORCHUK, Y. V.; SHANINA, I. N. Study of linear trends of temporal series of solar radiation. **Izvestiâ Akademii nauk SSSR**, v.30, p.389-396, 1994.

