



PREVISÃO DE CUSTOS DE UM CONDOMÍNIO DE APARTAMENTOS UTILIZANDO A METODOLOGIA DE BOX E JENKINS PARA SÉRIES TEMPORAIS

COST FORECASTING OF A CONDOMINIUM OF APARTMENTS USING THE METHODOLOGY OF BOX AND JENKINS FOR TEMPORARY SERIES

Carlos dos Santos*

Resumo

Conhecer bem para onde vai o dinheiro é essencial nos tempos atuais. A maior preocupação dos condôminos é o valor de condomínio, que deixou de apenas assustar para passar a "amedrontar" quem vive nesse tipo de moradia. A previsão desses custos pode auxiliar a reduzir os mesmos. Do exposto, o presente trabalho tem por finalidade realizar a previsão do custeio com energia elétrica, água, folha de pagamento, entre outros, de um condomínio de apartamentos do município de Foz do Iguaçu, para os 12 meses subseqüentes, utilizando o método de Box e Jenkins para a previsão de séries temporais.

Palavras-chave: Box e Jenkins, previsão, séries temporais.

Abstract

Knowing where money is going is essential in today's times. The biggest concern of the condominium owners is the condominium value, which stopped from scaring to start to "frighten" those who live in this type of housing. The forecast of these costs may help to reduce them. From the above, the purpose of this study is to forecast the cost of electric power, water, payroll, among others, of a condominium of apartments in the municipality of Foz do Iguaçu, for the following 12 months, using the Box method and Jenkins for the time series forecast.

Keywords: Box and Jenkins, forecast, time series.

* Professor da Universidade Estadual do Oeste do Paraná - Campus de Foz do Iguaçu. E-mail: sancars@bol.com.br



1 Introdução

O controle de custos tem por finalidade o fornecimento de informações que contribuem para auxiliar no setor administrativo e auxilia na lucratividade da empresa de acordo com o seu planejamento adequado. Para que uma empresa tenha um bom desempenho com o controle de custos, precisa-se de uma análise muito rigorosa em todos os setores da empresa, já que o controle de custos não está somente ligado à produção (Martins,2011).

A taxa de condomínio, de apenas mais um item nas despesas mensais, passou a disputar a linha de frente com alimentação, escola e aluguel. E quem são os vilões do condomínio: salários e encargos, serviços públicos, com especial destaque para as contas d'água e conservação de elevadores, etc., (<http://www.controlpred.com.br/nossomanual.pdf>, acessado em 24/02/2017). A previsão desses custos pode auxiliar no controle ou até mesmo na redução dos seus valores. Do exposto, o presente trabalho tem os seguintes objetivos:

- Geral: Estudar e aplicar a metodologia de Box e Jenkins para séries temporais;
- Específico: Realizar a previsão da taxa de condomínio individual, de um conjunto de apartamentos.

1.1 Série temporal

Pode-se dizer que uma série temporal $Z_1, Z_2, Z_3, \dots, Z_t$, é um conjunto de observações ordenadas no tempo (t), e que apresentam dependência serial, ou seja, dependência entre instantes de tempo. A notação $Z_1, Z_2, Z_3, \dots, Z_t$, indica uma série de tamanho T . O instante T geralmente indica o último instante disponível (MORETTIN & TOLOI, 2006). São exemplos de séries temporais:

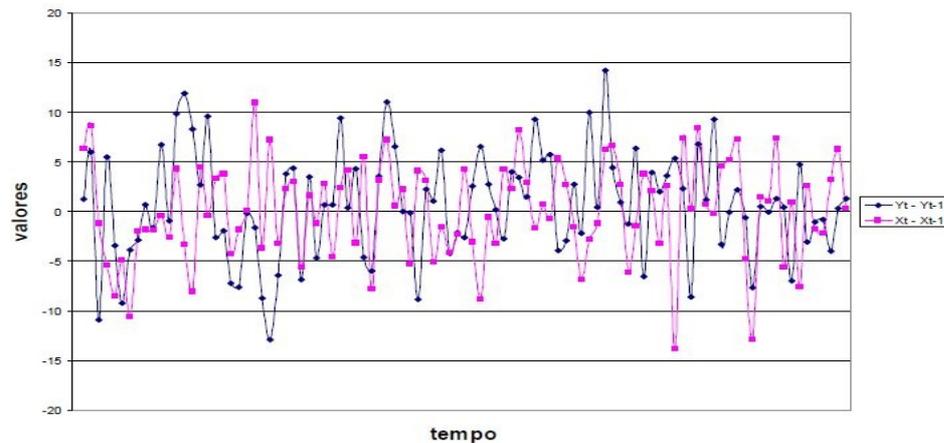
- Registro de acidentes, num cruzamento;
- Índice Pluviométrico, por mês, na cidade de Foz do Iguaçu;
- Produção de energia, por dia, na Usina Hidrelétrica de Itaipu;
- Registro de marés no porto de Paranaguá.

A Série temporal é discreta se forem obtidos valores em intervalos de tempo iguais, como mostram os exemplos ii e iii , enquanto i e iv, são exemplos de séries temporais contínuas, pois as ocorrências podem ser observadas a qualquer instante do intervalo de tempo.

1.2 Estacionariedade

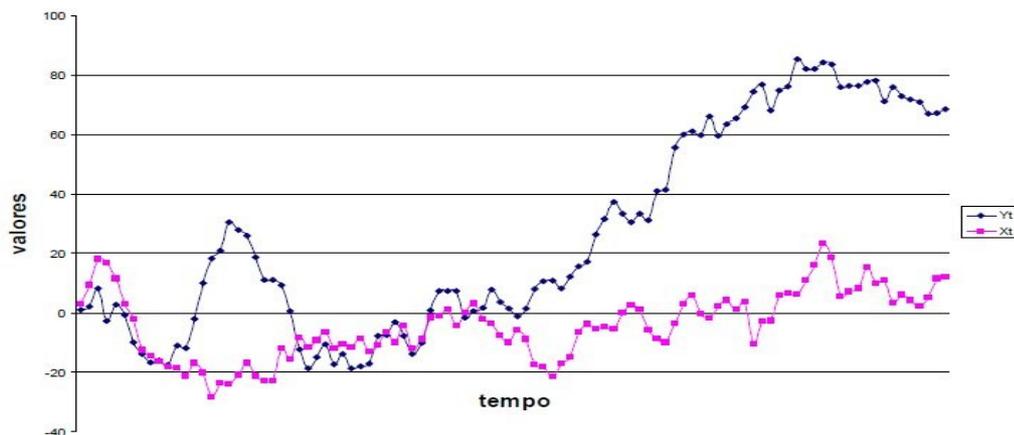
Uma das suposições que se faz a respeito de uma série temporal é a de que a mesma deve ser estacionária, ou seja, desenvolve-se no tempo, aleatoriamente, ao redor de uma média constante, refletindo alguma forma de equilíbrio estável (MORETTIN & TOILOI, 2006), conforme mostra a figura 3.4.

FIGURA 1 – Comportamento de uma série estacionária

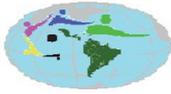


Fonte: Elaboração própria

Figura 2 – Comportamento de uma série não estacionária



Fonte: Elaboração própria



1.3 Autocorrelação

Segundo Bueno (2008), coeficiente de autocovariância e de uma única variável em defasagem j (número de intervalos de tempo defasados) é dada por:

$$\text{cov}(Z_t; Z_{t+j}) = E(Z_t - \mu_{t+j}) \cdot (Z_{t+j} - \mu_{t+j})$$

No caso de uma série estacionária, a média e a variância não variam no tempo. Dessa forma, a autocovariância entre duas observações quaisquer depende somente do número de intervalos defasados e é dada por

$$\text{cov}(Z_t; Z_{t+j}) = E(Z_t - \mu) \cdot (Z_{t+j} - \mu)$$

O coeficiente de autocorrelação é utilizado para verificar se há correlação de uma única variável em defasagem j , no tempo. Sua expressão é dada por:

$$\rho_j = \frac{\text{COV}(Z_t, Z_{t+j})}{\text{DP}(Z_t) \cdot \text{DP}(Z_{t+j})}$$

em que:

$\text{DP}(Z_t)$ é o desvio padrão da série $Z_1, Z_2, Z_3, \dots, Z_t$; e

$\text{DP}(Z_{t+j})$ é o desvio padrão da série com defasagem j .

O estimador do coeficiente de autocorrelação ρ_j é representado por r_j , cuja expressão é dada por:

$$\hat{\rho}_j = \frac{\sum_{t=1}^n (Z_t - \bar{Z})(Z_{t+j} - \bar{Z})}{\sqrt{\sum_{t=1}^n (Z_t - \bar{Z})^2} \sqrt{\sum_{t=1}^n (Z_{t+j} - \bar{Z})^2}}$$

Em que:

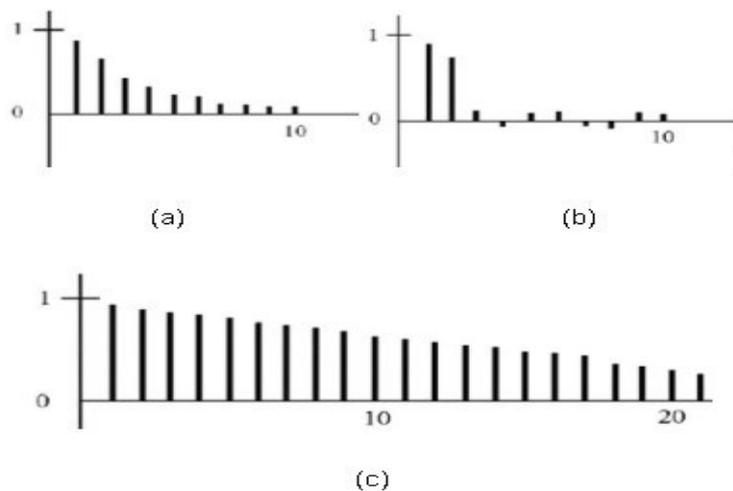
\bar{Z} é a média amostral da série;



Z_t é o valor amostral observado no instante t ;
 Z_{t+j} é o valor amostral observado, no instante $t + j$.

Quando a equação do coeficiente de correlação e suas estimativas são colocadas para todas as defasagens, então a equação é chamada de função de autocorrelação (FAC). O correlograma corresponde ao gráfico da FAC, no qual é possível realizar um diagnóstico da condição de estacionariedade da série em estudo. Uma série é estacionária se os valores de da FAC caem lentamente (figura 3a), ou se há um truncamento abrupto, ou um padrão de decaimento rápido (figura 3b). Uma série é não estacionária se há um padrão de decaimento lento para zero (figura 3c).

FIGURA 3 – Correlogramas de verificação de estacionariedade (a e b) ou a não estacionariedade (c) da série da série



Fonte: Elaboração própria

1.4 Autocorrelação parcial

A autocorrelação parcial mede a dependência linear entre Z_t e Z_{t+k} , eliminando o efeito das variáveis intermediárias $Z_{t+1}, Z_{t+2}, \dots, Z_{t+k-1}$.

1.5 Ruído Branco

O ruído branco é uma sequência onde cada valor tem média zero, variância constante e não tem relação com qualquer valor observado da própria série, ou seja, a autocorrelação é igual a zero. A média zero é conveniência, pois seria possível



especificar um ruído branco cuja média fosse diferente de zero. O processo é um ruído branco porque sua função espectral é horizontal como a luz branca e causa alterações na série como as ondas eletromagnéticas produzem ruídos na sintonização de um rádio. Portanto, um ruído branco é, ao mesmo tempo, temporalmente homogêneo, estacionário e sem memória (BUENO, 2008).

1.6 Modelos auto-regressivos AR(p)

Os modelos AR(p) (ou modelos auto regressivos de ordem p), têm a estrutura

$$\phi(B) \cdot Z_t = a_t \Rightarrow (1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p) \cdot Z_t = a_t$$
$$Z_t = \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + \dots + \phi_p Z_{t-p} + a_t$$

em que:

B é o operador retardo, ou operador translação para o passado, denotado por B e definido por: $B \cdot Z_t = Z_{t-1}$, ..., $B^p \cdot Z_t = Z_{t-p}$

$Z_t, Z_{t-1}, \dots, Z_{t-p}$, são variáveis aleatórias com distribuição $\phi_1, \phi_2, \dots, \phi_p$ são os coeficientes do modelo, os quais são estimados pelo método de mínimos quadrados, minimizando a soma dos quadrados dos erros (diferenças entre os valores da observados na série e os estimados pelo modelo); a_t é o ruído branco.

1.7 Modelos Médias Móveis MA(q)

Segundo Barros (2004), o modelo MA(q) (médias móveis de ordem q) é sempre estacionário e sua expressão é dada do por:

$$Z_t = \theta(B) \cdot a_t \Rightarrow Z_t = (1 + \theta_1 B + \theta_2 B^2 + \dots + \theta_q B^q) \cdot a_t$$
$$Z_t = a_t + \theta_1 a_{t-1} + \theta_2 a_{t-2} + \dots + \theta_q a_{t-q}$$

Onde:

θ_i são parâmetros da estrutura, $i=1, \dots, q$ (a ordem da estrutura)
 a_t é ruído branco com média zero e variância σ_a .

1.8 Modelo Auto-Regressivo Médias Móveis (ARMA)

Esse modelo é uma combinação dos dois anteriores onde Z_t é descrito por seus valores passados e pelos ruídos branco corrente e passados.

A estrutura geral ARMA(p,q) é expressa por:



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

Em que:

ϕ_i são os parâmetros da estrutura auto-regressiva, $i=1, \dots, p$
 θ_i são os parâmetros da estrutura médias móveis, $i=1, \dots, q$
 a_t é o ruído branco

1.9 Modelo Auto-Regressivo Integrado de Médias Móveis (ARIMA)

O modelo ARIMA (p, d, q) (*Auto-Regressive Integrated Moving Average*) é adequado para a previsão de séries temporais cujo processo estocástico não é estacionário. Logo, a série original passará por algumas diferenciações a fim de torná-la estacionária (Box & Jenkins, 1994).

O número necessário de diferença para tornar uma série estacionária é denominado ordem de integração (d).

A estrutura geral ARIMA(p, d, q) com a constante θ_0 é expressa por:

$$(1 - \phi_1 B - \dots - \phi_p B^p)(1 - B)^d Z_t = \theta_0 + (1 - \theta_1 B - \dots - \theta_q B^q) a_t$$

$$\phi(B) \nabla^d Z_t = \theta_0 + \theta(B) a_t$$

Em que:

B é operador de translação para o passado definido como $B^m Z_t = Z_{t-m}$;

i) $\phi(B) = 1 - \phi_1(B) - \dots - \phi_p(B)^p$ representa o operador auto-regressivo de ordem p ;

ii) $\theta(B) = 1 - \theta_1(B) - \dots - \theta_q(B)^q$ representa o operador médias móveis de ordem q ;

iii) a_t é o ruído branco;

d representa o número de diferenças

$\nabla = 1 - B$ representa o operador diferença

Este operador diferença é definido como:

$$Z_t - Z_{t-1} = Z_t - B Z_t = (1 - B) Z_t = \nabla Z_t$$

$$\nabla^d = (1 - B)^d$$

1.10 Modelo Sazonal Auto-Regressivo Integrado de Médias Móveis (SARIMA)

Os modelos ARIMA exploram a autocorrelação entre os valores da série em instantes sucessivos, mas, quando os dados são observados em períodos inferiores a um ano, a série também pode apresentar autocorrelação para uma estação de sazonalidades. Os modelos que contemplam as séries que apresentam autocorrelação sazonal são conhecidos como SARIMA. O modelo SARIMA contém



uma parte não sazonal com parâmetros (p,d,q) e uma sazonal com parâmetros (P,D,Q). Logo, o modelo geral para a nossa série, com estação sazonal 5, é escrito da forma:

$$\phi(B)\Phi(B^5)(1-B^5)^D(1-B)^d Z_t = \theta(B)\Theta(B^5)\alpha_t$$

em que:

- i) B é operador de translação para o passado definido como $B^m Z_t = Z_{t-m}$;
- ii) $\phi(B) = 1 - \phi_1(B) - \dots - \phi_p(B)^p$ é a parte auto-regressiva não sazonal ou estacionária de ordem p;
- iii) $\Phi(B^5) = 1 - \Phi_1 B^5 - \dots - \Phi_p(B)^p$ é a parte auto-regressiva sazonal de ordem p e ordem sazonal 5;
- iv) $(1 - B^5)^D$ é a parte de integração sazonal de ordem D e estação sazonal 5;
- v) $(1 - B)^d$ é a parte de integração não-sazonal de ordem d;
- vi) $\theta(B) = 1 - \theta_1(B) - \dots - \theta_q(B)^q$ é a parte não sazonal de médias móveis de ordem q;
- vii) $\Theta(B^5) = 1 - \Theta_1 B^5 - \dots - \Theta_Q(B)^Q$ é a parte sazonal de médias móveis de ordem Q e estação sazonal de ordem 5.

$$\phi(B)\Phi(B^5)(1-B^5)^D(1-B)^d Z_t = \theta(B)\Theta(B^5)\alpha_t$$

O modelo SARIMA(2,2,0)x(2,2,0) fica:

$$(1 - \phi_1 B - \phi_2 B) (1 - \Phi_1 B - \Phi_2 B) (1-B^5)^2 (1-B^5)^2 = \alpha_t$$

1.11 Método de Box e Jenkins

Neste trabalho foi utilizada a metodologia de Box e Jenkins (1970). Tal metodologia é bastante utilizada na análise de modelos paramétricos e consiste em ajustar modelos auto regressivos integrados de médias móveis ARIMA (p, q, d) a um conjunto de dados (MORETTIN & TOLOI, 2006).

Segundo Bueno (2008), os termos auto regressivos correspondem a defasagens da série transformada, e os termos médias móveis (“moving averages”) referem-se a defasagens dos erros aleatórios. Já, o termo “integrado” refere-se ao processo de diferenciar a série original para torná-la estacionária.

O objetivo da metodologia Box & Jenkins é determinar os três componentes que configuram qualquer estrutura que são: p parâmetros auto-regressivos, d processos de diferenciação (integração) e q parâmetros de médias móveis. De uma forma geral, a notação apresentada por Box & Jenkins é do tipo ARIMA (p, d, q). Por exemplo, a estrutura ARIMA (1, 2, 0) significa que esta possui um parâmetro auto-regressivo, duas diferenciações a partir da série original e nenhum parâmetro de



médias móveis. O procedimento padrão para utilização da metodologia de previsão ARIMA consiste nos seguintes passos, segundo Box & Jenkins:

1. Realiza-se a diferenciação da série original tantas vezes quanto necessário para torná-la estacionária;

2. Identifica-se os valores p e q através da análise das funções de autocorrelação (FAC) e de autocorrelação parcial (FACP) estimadas e a estimação dos parâmetros (ϕ_i , $i = 1, \dots, p$ e θ_i , $i = 1, \dots, q$). A análise destas funções permite a tomada de decisão acerca do número p de parâmetros auto-regressivos e do número q de parâmetros de médias móveis que devem ser escolhidos de maneira a se obter uma estrutura parcimoniosa (ou seja, uma estrutura que tenha o menor número de parâmetros dentre todas as estruturas que se ajustem aos dados da série);

3. Realizam-se as previsões (obtenção dos novos valores da série) e os intervalos de confiança para as mesmas. Para verificar se é necessária a diferenciação da série original, é comum utilizar o gráfico da mesma. Na maioria dos casos, a série pode ser não-estacionária quanto ao nível ou quanto à inclinação. Quando a série for não-estacionária quanto ao nível, para torná-la estacionária basta tomar a primeira diferença e, quando for não-estacionária quanto à inclinação, basta tomar a segunda diferença (Souza e Camargo, 2004). Por vezes é necessário realizar uma transformação logarítmica na série original para torná-la estacionária. Deve-se considerar que, na metodologia desenvolvida por Box & Jenkins, os estágios da identificação e estimação se sobrepõem. Muitas vezes, ao se sobrefixar uma estrutura na etapa de identificação, aquela será corrigida na estimação, quando a significância dos parâmetros estimados será avaliada. Na maioria dos casos, o resultado da identificação não será preciso e mais de uma. Metodologia de Previsão de Séries Temporais - Box & Jenkins 20 estrutura será identificada. Esta imprecisão é consequência de que, na prática, o ajuste a situações reais induz a alguns erros. A identificação visa encontrar uma classe de estruturas a serem ajustadas e validadas (BOX, JENKINS & REINSEL, 1994) e (SOUZA E CAMARGO, 2004). Procura-se então, encontrar um subconjunto adequado de estruturas ARIMA (p, d, q).

A estratégia para a construção do modelo é baseada em um ciclo iterativo, no qual a escolha da estrutura do modelo é baseada nos dados observados. Os estágios do ciclo iterativo são:

- a) uma classe geral de modelos é considerada para análise (especificação);
- b) há identificação de um modelo com base na análise de autocorrelações e autocorrelações parciais;
- c) finalmente, há a verificação ou diagnóstico do modelo ajustado, através de uma análise de resíduos, para saber se este é adequado para previsão.

Caso o modelo não seja adequado, o ciclo é repetido, retornando à fase de identificação. Um procedimento que muitas vezes é utilizado é o de identificar não apenas um modelo, mas um conjunto deles que serão estimados e verificados. Se o propósito é a previsão, será escolhido o modelo com menor erro quadrático médio (MORETTIN & TOLOI, 2006).



Ha certa dificuldade na aplicação do método de Box e Jenkins, por ser laborioso. Assim, requer a utilização de programas computacionais adequados. Na análise dos dados foram utilizados os programas Statgraphics.

1.12 Medidas da acurácia do erro

As estatísticas MAPE, MAE e RMSE são informações importantes para compararmos os modelos de suavização e decidirmos se estes representam um bom ajuste aos nossos dados (disponível em <<http://www.portalaction.com.br/series-temporais>>, acessado em 25/10/2017).

A média percentual absoluta do erro (MAPE - Mean Absolute Percentage Error) expressa a acurácia do erro em porcentagem. Por exemplo, se temos um MAPE de 7%, quer dizer que o nosso ajuste esta errado em 7% dos dados. Para calcular o MAPE utilizamos a seguinte expressão:

$$\text{MAPE} = \frac{\sum_{t=1}^n |(y_t - \hat{y}_t) / y_t|}{n} \times 100 \text{ se } y_t \neq 0$$

onde y_t são os nossos dados, \hat{y}_t são os ajustes e n é o número de observações. Note que se $y_t = \hat{y}_t$ obtemos MAPE = 0, ou seja, quanto menor o MAPE melhor é o nosso ajuste.

O erro médio absoluto (MAE - Mean Absolute error) representa o erro médio dos valores ajustados pelo modelo em relação aos valores observados. Por exemplo, se estamos ajustando uma série temporal de visitas durante o tempo e encontramos um MAE de 72, significa que a diferença dos valores ajustados em relação aos observados foi, em média, igual a 72. Para encontrarmos o MAE, realizamos o seguinte cálculo:

$$\text{MAE} = \frac{\sum_{t=1}^n |y_t - \hat{y}_t|}{n}$$

onde y_t são os nossos dados, \hat{y}_t são os ajustes e n é o número de observações. Note que se $y_t = \hat{y}_t$ obtemos MAE = 0, ou seja, quanto menor o MAE melhor é o ajuste do modelo aos dados.



Raiz do erro quadrático da média (RMSE – Root Mean Squared error). Esta medida de acurácia é bastante comum em ajustes de séries temporais. Para calcularmos o RMSE utilizamos a seguinte expressão:

$$RMSE = \sqrt{\frac{\sum_{t=1}^n (y_t - \hat{y}_t)^2}{n}}$$

onde y_t são os nossos dados, \hat{y}_t são os ajustes e n é o número de observações. Note que se $y_t = \hat{y}_t$ obtemos $RMSE = 0$, ou seja, quanto menor o RMSE melhor é o nosso ajuste.

1.12 Resultados e discussões

Utilizando o programa Statgraphics foram digitadas 52 observações mensais passadas e o modelo de previsão selecionado foi o ARIMA (0,1,1), com diferença de ordem não sazonal 1, para prever os valores da taxa de condomínio, por apartamento, de 12 meses futuros. Foram feitas simples diferenças de ordem 1.

A Tabela 1 resume a significância estatística dos termos no modelo de previsão. Os termos com valores P inferiores a 0,05 são, estatisticamente, significativamente diferentes de zero no nível de confiança de 95,0%. O valor P para o termo MA(1) é inferior a 0,05, portanto, é significativamente diferente de 0. O valor P para o termo constante é inferior a 0,05, portanto, é significativamente diferente de 0. O desvio padrão estimado do ruído branco de entrada é igual a 33.502. Nota-se que ocorrem as seguintes estimativas: $\theta_0 = 2,75059$ e $\theta_1 = 0,98586$. Logo, a equação do modelo ARIMA(0,1,1) fica

$$(1 - B) Z_t = \theta_0 + (1 - \theta B) a_t$$

$$(1 - B) Z_t = 2,75059 + a_t - 0,98586 a_{t-1}$$

Ou

$$Z_t - B Z_t = \theta_0 + a_t - \theta B a_t$$

$$Z_t - Z_{t-1} = \theta_0 + a_t - \theta a_{t-1}$$

$$Z_t = \theta_0 + Z_{t-1} + a_t - \theta a_{t-1}$$

$$Z_{t+h} = \theta_0 + Z_{t+h-1} + a_{t+h} - \theta a_{t+h-1}, \forall h$$

$$Z_{t+h} = 2,75059 + Z_{t+h-1} + a_{t+h} - 0,98586 a_{t+h-1}$$



Tabela 1 – Análise dos coeficientes do modelo ARIMA(0,1,1)

Parâmetro	Estimati va	Erro padrão	T	Valor P
MA(1)	0,98586	0,0181657	54,270	0,000000
Mean	2,75059	0,27928	4	0,000000
Constant	2,75059		5	

Fonte: Elaboração própria

A tabela 2 mostra algumas estatísticas que envolvem o erro ou resíduo, ou seja, as diferenças entre os valores observados e os ajustados pelo modelo.

- (1) raiz do erro quadrático médio (RMSE)
- (2) erro absoluto médio (MAE)
- (3) média percentual absoluta do erro (MAPE)
- (4) o erro médio (ME)
- (5) o erro médio percentual (MPE)

Esta tabela compara os resultados de cinco modelos de previsão diferentes:

- (A) ARIMA(0,1,1) com constante
- (B) Média constante = 344,275
- (C) Curva de tendência $S = \exp(5,86539 + -0,403023 / t)$
- (D) Média móvel simples de três termos
- (E) Suavização exponencial simples com alfa = 0,2323

Observando as medidas de acurácia do erro, o modelo com o menor erro quadrático médio (RMSE) é o modelo A. O modelo com o menor erro médio absoluto médio (MAE) é o modelo A. O modelo com menor média percentual absoluta do erro (MAPE) é o modelo A. O MAPE, expressa a acurácia do erro em percentagem. Neste caso, temos um MAPE de aproximadamente 7%, quer dizer que o ajuste do modelo ARIMA(0,1,1) está errado em apenas 7% dos dados, um resultado bom.

Tabela 2 – Comparação dos resíduos de 5 modelos .

Model	RMSE	MAE	MAPE
(A)	33,3317	24,334	7,10209
(B)	52,309	43,6272	12,8791
(C)	47,9566	39,282	11,4661
(D)	40,8722	31,1205	8,79196
(E)	37,2958	27,3132	7,68942

Fonte: Elaboração própria



A tabela 3 resume os resultados de cinco testes executados nos resíduos para determinar se cada modelo é adequado para os dados. Um “OK” significa que o modelo passa o teste. Um * significa que ele falha no nível de confiança de 95%. Dois * significa que ele falha no nível de confiança de 99%. Três * significa que ele falha no nível de confiança de 99,9%. Observe que o modelo atualmente selecionado, modelo A, passa 5 testes. Uma vez que nenhum teste é estatisticamente significativo ao nível de confiança de 95% ou mais, o modelo ARIMA(0,1,1) provavelmente é adequado para os dados.

Tabela 3 - Testes dos resíduos

Modelo	RMSE	RUNS	RUNM	AUTO	MEAN	VAR
(A)	33,3317	OK	OK	OK	OK	OK
(B)	52,309	OK	***	***	***	OK
(C)	47,9566	OK	**	***	***	OK
(D)	40,8722	*	OK	OK	OK	OK
(E)	37,2958	OK	OK	OK	OK	OK

Onde:

RMSE = raiz do erro do quadrado médio;

RUNS = Teste para corridas excessivas para cima e para baixo;

RUNM = Teste para corridas excessivas acima e abaixo da mediana ;

AUTO = Teste de Box-Pierce para autocorrelação excessiva

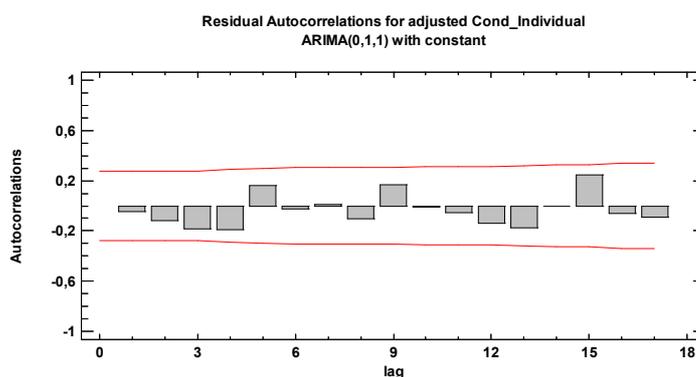
MEAN = Teste para a diferença na média da metade;

VAR = Teste para diferença de variação 1ª metade para 2ª metade.

Fonte: Elaboração própria

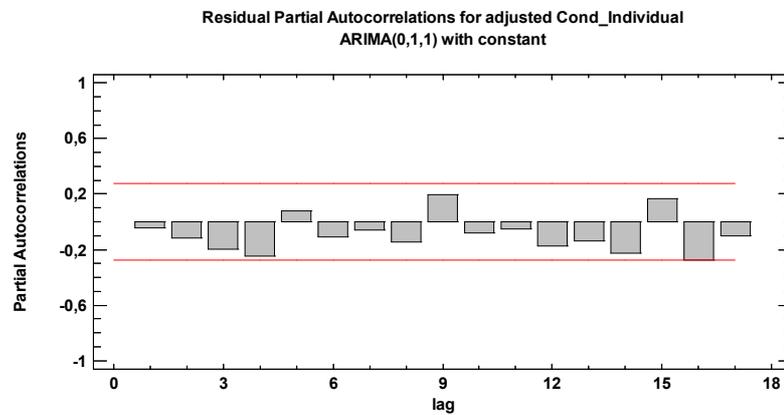
Também foram plotados os gráficos da função de autocorrelação (FAC) e da função auto correlação parcial (FACP), dos erros ou resíduos, nas figuras 4 e 5, as quais mostram que não existe autocorrelação alta entre os mesmos, ou seja, os resíduos são independentes. A figura 6 mostra os resíduos se distribuindo de forma equilibrada em relação à linha central, o que mostra um comportamento aleatório dos mesmos.

Figura 4 - Gráfico da autocorrelação dos resíduos



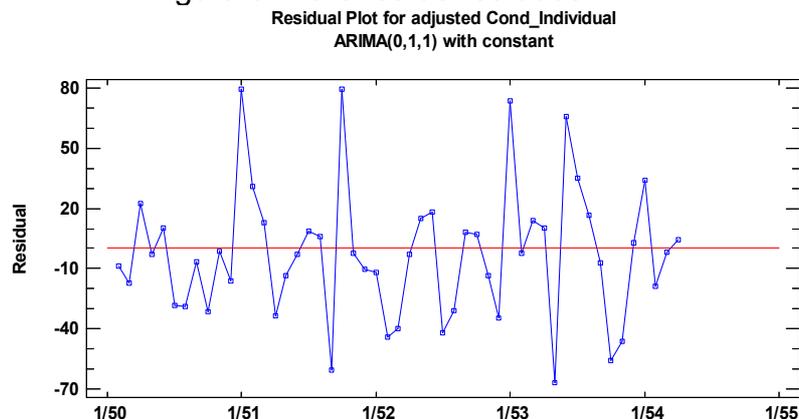
Fonte: Elaboração própria

Figura 5 - Gráfico da autocorrelação parcial dos resíduos



Fonte: Elaboração própria

Figura 6 - Gráfico de resíduos



Fonte: Elaboração própria

Três testes foram realizados para determinar se os resíduos formam ou não uma sequência aleatória, como mostra a tabela 3. Uma sequência de números aleatórios geralmente é chamada de ruído branco, pois contém contribuições iguais em muitas frequências. O primeiro teste fez a contagem do número de vezes que a sequência estava acima ou abaixo da mediana. O número dessas corridas é igual a 24, em comparação com um valor esperado de 26,0 se a sequência fosse aleatória.



Uma vez que o valor $P = 0,668167$ é maior do que $0,05$, não podemos rejeitar a hipótese de que os resíduos são aleatórios no nível de confiança de 95% ou superior.

O segundo teste conta o número de vezes que a sequência cresceu ou decresceu. O número dessas corridas é igual a 28, em comparação com um valor esperado de 33,6667 se a sequência fosse aleatória. Uma vez que o valor $P = 0,0806011$ é maior do que $0,05$, não podemos rejeitar a hipótese de que a série é aleatória no nível de confiança de 95,0% ou superior.

O Teste de Box-Pierce é baseado na soma dos quadrados dos primeiros 24 coeficientes de autocorrelação. Uma vez que o valor $P = 0,578877$ para este teste é maior do que $0,05$, não podemos rejeitar a hipótese de que a série de erros é aleatória no nível de confiança de 95% ou superior.

Uma vez que ocorre um comportamento aleatório dos resíduos para os três testes, o modelo selecionado ARMA(0,1,1), captura toda a estrutura nos dados e pode ser utilizado para a previsão.

Tabela 3. Testes de aleatoriedade dos resíduos

Teste	Estatística	Valor P
A	0,428661	0,668167
B	1,74721	0,080601
C	14,2664	0,578877

*A = teste do número de corridas acima e abaixo da mediana

*B = teste do número de corridas para cima e para baixo

*C = teste de Box-Pierce

Fonte: Elaboração própria

As previsões e seus respectivos intervalos de 95% confiança para os doze meses subsequentes, ou seja, de maio de 2017 a abril de 2018 estão apresentados na tabela 4.

Tabela 4 - Previsões e respectivos intervalos de 95% de confiança para a taxa de condomínio, por apartamento

Mês/ano	Previsão da taxa de condomínio(R\$)	Limite inferior de 95% de confiança	Limite superior de 95% de confiança
Maio/2017	418,398	351,073	485,722
Junho/2017	421,148	353,816	488,48
Julho/2017	423,899	356,56	491,237
Agosto/2017	426,649	359,304	493,994
Setembro/2017	429,400	362,048	496,752
Outubro/2017	432,150	364,792	499,509
Novembro/2017	434,901	367,536	502,266
Dezembro/2017	437,652	370,28	505,024
Janeiro/2018	440,402	373,023	507,781
Fevereiro/2018	443,153	375,767	510,538
Março/2018	445,903	378,511	513,296
Abril/2018	448,654	381,255	516,053

Fonte: Elaboração própria



2 Conclusão

Através de várias tentativas, aplicando o método de Box e Jenkins, foi achado o modelo ARIMA(0,1,1), com $\theta_0 = 2,75059$ e $\theta_1 = 0,98586$, cujos valores são significativos, com valores P menores que 0,05, e a equação de previsão foi $Z_{t+h} = 2,75059 + Z_{t+h-1} + a_{t+h} - 0,98586a_{t+h-1}$. Esse modelo foi escolhido devido ao fato de, através dele, serem achadas previsões da taxa de condomínio com erros não correlacionados e com distribuição normal. O gráfico de resíduos e o teste de Box-Pierce mostraram que os erros ou resíduos se comportam de forma aleatória. Este modelo ainda foi comparado com outros e foi o que apresentou o menor MAPE (erro médio percentual absoluto), menor raiz do erro quadrático médio (RMSE) e menor erro médio absoluto médio (MAE). Alguns dos valores previstos da taxa de condomínio pelo modelo, de maio a outubro de 2017 foram comparados com os valores reais e apresentaram diferenças bem baixas, bem como todos valores reais caíram dentro do intervalo de confiança proposto, indicando que o modelo proposto serve para a previsão dessas taxas, nesta situação específica.

Referências

- BARROS, M. **Processos Estocásticos**. Rio de Janeiro. Papel Virtual, 2004.
- BOX, G. E. P., Jenkins, G. M., Reinsel, G. (1994). **Time Series Analysis: Forecasting and Control**. Third Edition. Englewood Cliffs: Prentice Hall.
- BUENO, R. L..S. **Econometria de Séries Temporais**. São Paulo: Cengage Learning, 2008.
- MARTINS, Elizeu. **Contabilidade de Custos**. 8ªEd. São Paulo, 2011.
- MORETTIN, P. A., TOLOI, C. M. C. **Análise de Séries Temporais**. São Paulo, Blucher, 2006.
- KORMANN, A. C. M.; PORTELLA, K. F.; PEREIRA, P. N. and SANTOS, R. P.. **Estudo do desempenho de quatro sistemas de materiais de reparo para estruturas hidráulicas de concreto de barragens**. 2003, vol.49, n.309, pp. 48-54. Disponível em < <http://www.periodicos.capes.gov.br>>, acessado em 07 de setembro de 2012.
- PORTAL ACTION. **Séries Temporais**. Disponível em <www.portalaction.com.br/series-temporais>, acessado em 25 de outubro de 2017.



REVISTA ORBIS LATINA
ISSN: 2237 6976



90

SOUZA, R. C; CAMARGO, M. E. **Análise e previsão de séries temporais: os modelos ARIMA**. 2ª Ed. Rio de Janeiro: Gráfica e Editora Regional, 2004.

Recebido em 06/11/2017

Aprovado em 19/12/2017