



## Avaliação da Precisão dos Modelos ARIMA com e sem Transformação Estabilizadora da Variância na Previsão de Séries Temporais Anuais

Discente: Rafael Guilherme Fernandes de Lima Silva

Orientador: Prof. Dr. Marcelo Ruy

### Resumo:

Dentre as aplicações de Análise de Séries Temporais, a previsão de valores futuros é uma das mais utilizadas, principalmente em se tratando de séries econômicas e financeiras. Nas previsões, a precisão do método utilizado é um dos fatores críticos em sua adoção. Um dos métodos clássicos e mais utilizados é o modelo Autorregressivo Integrado de Médias Móveis (ARIMA – *Autoregressive Integrated Moving Average*). Este método, diferentemente de alguns outros existentes, é aplicável a séries que possuem variância constante. Caso a variabilidade dos dados mude com o tempo ou dependa de seu nível, é usual aplicar à série uma transformação não linear (p. ex., logaritmo), visando a estabilização de sua variância. Do ponto de vista exclusivo das previsões alguns autores afirmam que a transformação estabilizadora da variância tem pouca influência na precisão das mesmas. Este trabalho teve como objetivo testar esta hipótese em séries temporais anuais. Para tanto, foram tomadas aleatoriamente 40 séries temporais anuais da M3 *Competition* e às mesmas foram ajustados modelos ARIMA com e sem transformação e a precisão de suas previsões pontuais foram calculadas. A conclusão foi que ambos os métodos têm em média o mesmo grau de precisão.

**Palavras chave:** Séries Temporais Univariadas; Métodos de Previsão; Modelos ARIMA; Transformações Estabilizadoras da Variância.

## Evaluating Forecast Accuracy of ARIMA Models with and without Variance Stabilizing Transformations in Annual Time Series

### Abstract:

Among the applications of Time Series Analysis, forecasts of future values is one of the most used, especially in economic and financial series. In forecasting, the precision of the method used is one of the critical factors in its adoption. One of the classic and most used methods is ARIMA model (autoregressive integrated moving average). This method, unlike some existing ones, is applicable to series that have constant variance. If the variability of the data changes over time or depends on its level, it is usual to apply a non-linear transformation (eg, logarithm) to the series, in order to stabilize its variance. From the standpoint of forecasting, Hyndman and Athanasopoulos (2018) argue that the variance stabilizing transformations have little influence on their accuracy. The objective of this work was to test this hypothesis in annual time series. For this, 40 annual time series of the M3 Competition (MAKRIDAKIS AND HIBON, 2000) were randomly selected, adjusted and transformed with ARIMA models, and the accuracy of their predictions were calculated. The conclusion was that both methods have on average the same degree of precision.

**Key words:** Univariate Time Series; Forecast Methods; ARIMA Models; Variance Stabilizing Transformations.

### 1. Introdução

De acordo com Morettin e Toloí (2006), série temporal é qualquer conjunto de observações ordenadas no tempo. Uma série temporal é contínua quando o registro das observações é feito continuamente ao longo do tempo. Quando as observações são tomadas em pontos específicos



do tempo, a série é discreta. Assim, contínua e discreta não se relacionam ao tipo de variável sendo medida, mas ao tempo das medições.

Segundo Morettin e Toloí (2006), os principais objetivos da análise de séries temporais são: descrever o comportamento da série; investigar seu mecanismo gerador; procurar periodicidades nos dados; e fazer previsões de seus valores futuros, sendo este último um dos mais utilizados, principalmente em se tratando de séries econômicas e financeiras.

Previsões são requeridas em muitas situações tais como projetar a demanda futura de determinado produto ou o nível de estoque de certo item. A precisão das previsões é um fator crítico para, dentre outras coisas, reduzir custos e fornecer um melhor serviço aos clientes (MAKRIDAKIS; HIBON, 2000).

Para Hyndman e Athanasopoulos (2018), os principais métodos de previsão de séries temporais são a decomposição em componentes, a suavização exponencial e o modelo autorregressivo integrado de médias móveis (ARIMA – *Autoregressive Integrated Moving Average*).

Os métodos de decomposição de séries temporais consideram que as mesmas são compostas de quatro padrões ou elementos básicos: tendência, variações cíclicas, sazonalidade e variações irregulares. A ideia é decompor a série em suas componentes básicas de variação, analisar cada uma separadamente e, então, recombina-las, a fim de descrever ou prever a variável de interesse. Hyndman e Athanasopoulos (2018) descrevem diversos métodos de decomposição, sendo o mais comum o método da decomposição clássica baseado em médias móveis centradas.

Morettin e Toloí (2006) explicam que a suavização exponencial é uma classe de métodos de previsão que se baseiam na ideia de que observações passadas contêm informações sobre o padrão da série temporal. O propósito é distinguir um padrão de comportamento de qualquer outro ruído que possa estar contido nas observações e, então, usar esse padrão para prever valores futuros da série. Hyndman e Athanasopoulos (2018) explicam que as previsões feitas pelos métodos de suavização exponencial são médias ponderadas de valores passados, com os pesos decaindo exponencialmente, à medida que as observações se tornam mais antigas. Os principais métodos de suavização exponencial são a suavização exponencial simples (para séries localmente constantes), a suavização de Holt (para séries com tendência apenas) e a suavização de Holt-Winters (para séries com sazonalidade).

De acordo com Hyndman e Athanasopoulos (2018), ao contrário dos métodos anteriores que modelam os componentes da série, o modelo ARIMA explora outra característica distintiva das mesmas: o fato de dados próximos terem maior relação entre si do que dados separados por grandes intervalos de tempo. Esta característica é resumida matematicamente por uma grandeza que varia entre  $\pm 1$  denominada autocorrelação serial, que seria o grau de associação linear entre valores defasados da série. A autocorrelação na defasagem 1 (dados separados por 1 unidade de tempo) seria a associação entre  $y_t$  e  $y_{t-1}$ , na defasagem 2 entre  $y_t$  e  $y_{t-2}$  e assim sucessivamente. O modelo ARIMA utiliza esta autocorrelação em diversas defasagens para prever os valores futuros da série baseado em seus valores passados e presente.

Diferentemente dos outros métodos, o modelo ARIMA é aplicável a séries que possuem variância constante. Caso a variabilidade dos dados mude com o tempo ou dependa de seu nível, é usual aplicar à série uma transformação não linear (p. ex., logaritmo), visando a estabilização de sua variância. Entretanto Hyndman e Athanasopoulos (2018) afirmam que a transformação estabilizadora da variância tem pouca influência na precisão das previsões pontuais do modelo ARIMA, o que remete aos objetivos deste trabalho.

## 2. Objetivo

Este trabalho tem como objetivo geral testar se há ou não diferença significativa entre as precisões das previsões pontuais geradas por modelos ARIMA com e sem transformação estabilizadora da variância, quando aplicados às mesmas séries temporais anuais (portanto, sem sazonalidade). Como objetivo específico, estaremos testando as seguintes hipóteses estatísticas:

$H_0$ : Os dois métodos possuem em média a mesma precisão

$H_1$ : Os dois métodos não possuem em média a mesma precisão

## 3. Referencial Teórico

### 3.1. Modelo ARIMA Não Sazonal

Na análise de regressão múltipla, a variável de interesse é escrita como uma combinação linear de variáveis independentes. Já em um modelo autorregressivo, a variável de interesse é escrita como uma combinação linear de seus valores passados. Assim, um modelo autorregressivo de ordem  $p$ , representado compactamente por  $AR(p)$ , é escrito como:

$$y_t = c + \phi_1 \cdot y_{t-1} + \phi_2 \cdot y_{t-2} + \dots + \phi_p \cdot y_{t-p} + \varepsilon_t \quad (1)$$

Onde  $c$  é o intercepto e  $\varepsilon_t$  uma sequência de variáveis aleatórias independentes igualmente distribuídas denominada erro ou resíduo (HYNDMAN; ATHANASOPOULOS, 2018). A equação (1) é semelhante à da regressão, exceto que as variáveis predictoras são valores defasados de  $y_t$ .

Ao invés de utilizar os valores passados da variável de interesse, um outro tipo de modelo utiliza uma média ponderada de distúrbios aleatórios passados e presente na equação:

$$y_t = \mu + \varepsilon_t + \theta_1 \cdot \varepsilon_{t-1} + \theta_2 \cdot \varepsilon_{t-2} + \dots + \theta_q \cdot \varepsilon_{t-q} \quad (2)$$

Onde  $\mu$  é a média do processo gerador da série temporal. O modelo descrito pela equação (2) é denominado de modelo de médias móveis de ordem  $q$  e é representado por  $MA(q)$ . Os valores de  $y_t$  podem ser interpretados como uma média móvel ponderada do erro presente e dos seus últimos  $q$  valores.

A combinação do modelo autorregressivo  $AR(p)$  com o de médias móveis  $MA(q)$  resulta no modelo denominado  $ARMA(p,q)$ . Para muitas séries encontradas na prática, para a obtenção de um modelo com um número não muito grande de parâmetros, a inclusão de termos autorregressivos e de médias móveis é uma solução adequada. Os modelos  $AR(p)$  e  $MA(q)$  são casos especiais do modelo  $ARMA(p,q)$  tomando-se  $q = 0$  e  $p = 0$ , respectivamente.

Os 3 modelos apresentados são aplicáveis a séries que possuem média constante (estacionárias na média). Séries com tendência estocástica podem ser transformadas em séries estacionárias na média tomando-se diferenças entre seus valores sucessivos, como indica a equação (3):

$$\Delta y_t = y_t - y_{t-1} \quad (3)$$

Se após aplicar-se a 1ª diferença da equação (3), a série estabilizar sua média, ela é dita integrada de ordem 1. Caso isso não ocorra, aplica-se a 2ª diferença, isto é, a equação (3) duas vezes:  $\Delta(\Delta y_t)$ . E assim sucessivamente até que se obtenha uma série com média constante. Uma série integrada de ordem  $d$  é aquela que necessitou de  $d$  diferenças para tornar-se estacionária na média. Se uma série integrada de ordem  $d$  puder ser modelada por um processo  $ARMA(p,q)$ ,

ter-se-á um modelo ARIMA( $p,d,q$ ) não sazonal, cuja equação é dada por:

$$\Delta^d y_t = c + \phi_1 \cdot \Delta^d y_{t-1} + \dots + \phi_p \cdot \Delta^d y_{t-p} + \varepsilon_t + \theta_1 \cdot \varepsilon_{t-1} + \dots + \theta_q \cdot \varepsilon_{t-q} \quad (4)$$

Onde o intercepto  $c$  é dado por:

$$c = (1 - \phi_1 - \dots - \phi_p) \cdot \mu \quad (5)$$

De acordo com Hyndman e Athanasopoulos (2018), o procedimento geral para ajustar modelos ARIMA aos dados é o seguinte: (a) transforme os dados para estabilizar a variância, caso necessário; (b) se a série for não estacionária na média, tome diferenças até que isso ocorra; (c) determine  $p$  e  $q$ ; (d) ajuste o melhor modelo do passo anterior e teste se os resíduos são ruído branco (variáveis aleatórias independentes igualmente distribuídas); (e) implemente o modelo para executar as previsões.

Segundo Hyndman e Athanasopoulos (2018), o passo (b) é feito por meio de análise gráfica ou mais objetivamente utilizando-se os chamados testes de raízes unitárias, amplamente disponíveis em *software* estatístico. Os passos (d) e (e) também estão disponíveis em programas voltados à análise de séries temporais.

Tradicionalmente o passo (c) é feito por meio da análise do padrão de autocorrelação dos dados. Isto se deve ao fato de que cada modelo ARMA( $p,q$ ) exibe um comportamento particular teórico para cada combinação de  $p$  e  $q$ . Na prática, entretanto, tal método não é simples. Primeiramente por ser subjetivo e dependente da experiência do analista. Além disso, o padrão das autocorrelações amostrais pode divergir bastante do padrão teórico devido ao erro amostral.

Para se contornar tal limitação, normalmente é utilizado um método baseado em uma função penalizadora (MORETTIN; TOLOI, 2006). A ideia básica é determinar a combinação de  $p$  e  $q$  que minimiza uma função que indica a falta de ajuste do modelo. Com o atual poder computacional, tal tarefa leva alguns poucos segundos. Basta ajustar todos os modelos possíveis com  $p$  e  $q$  variando entre, por exemplo, 0 e 5 e verificar qual combinação minimiza a função penalizadora. As funções mais utilizadas são o critério de informação de Akaike corrigido (AICc) e o critério de informação bayesiano (BIC).

A operação de diferenciação pode ser utilizada para transformar uma série não estacionária na média em uma série estacionária. Séries não estacionárias na variância possuem sua função variância dependente do tempo ou da média da série. Para transformar tais séries em estacionárias, faz-se necessário o uso de transformações não lineares (Figura 1). O tipo de transformação depende da relação entre o desvio-padrão da série e sua média.

A família de transformações estabilizadoras da variância é denominada transformação de Box-Cox e é definida como:

$$z = (y_t^\lambda - 1)/\lambda \quad (6)$$

Onde  $\lambda$  é um parâmetro a ser determinado a partir dos dados. Assim, se  $\lambda$  for, por exemplo,  $\frac{1}{2}$ , tem-se a transformação raiz quadrada (Figura 1). Pode-se provar utilizando Cálculo que quando  $\lambda$  for zero a equação (6) reduz-se ao logaritmo natural de  $y_t$ . A figura 2 mostra um exemplo de uma série heteroscedástica tornada homoscedástica por meio de uma transformação logarítmica.

Todos os passos de (a) a (e) do procedimento geral de ajuste dos modelos ARIMA aos dados podem ser executadas no *software* R (R CORE TEAM, 2019). O pacote *forecast* desenvolvido por Hyndman e Khandakar (2008) automatiza todo o processo e facilita ainda mais a análise,

principalmente quando o objetivo é analisar diversas séries simultaneamente. Pacotes são coleções de funções, dados e códigos compilados e que ampliam as capacidades originais do *software*. Normalmente são programados por estatísticos computacionais e não raro são o estado da arte em determinada área da Estatística. Como os pacotes e o próprio *software* são de código aberto, eles estão disponíveis à verificação da comunidade científica com relação a sua correção, eficácia, etc.

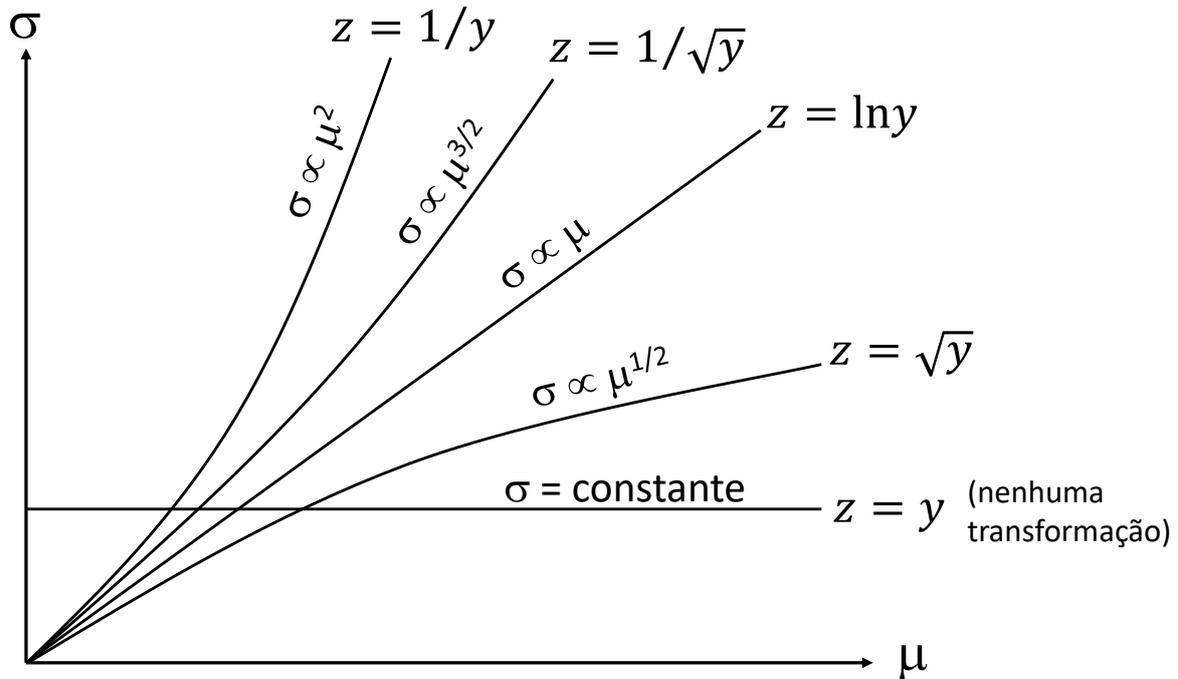


Figura 1: Transformações Estabilizadoras da Variância. Fonte: Ruy (2019).

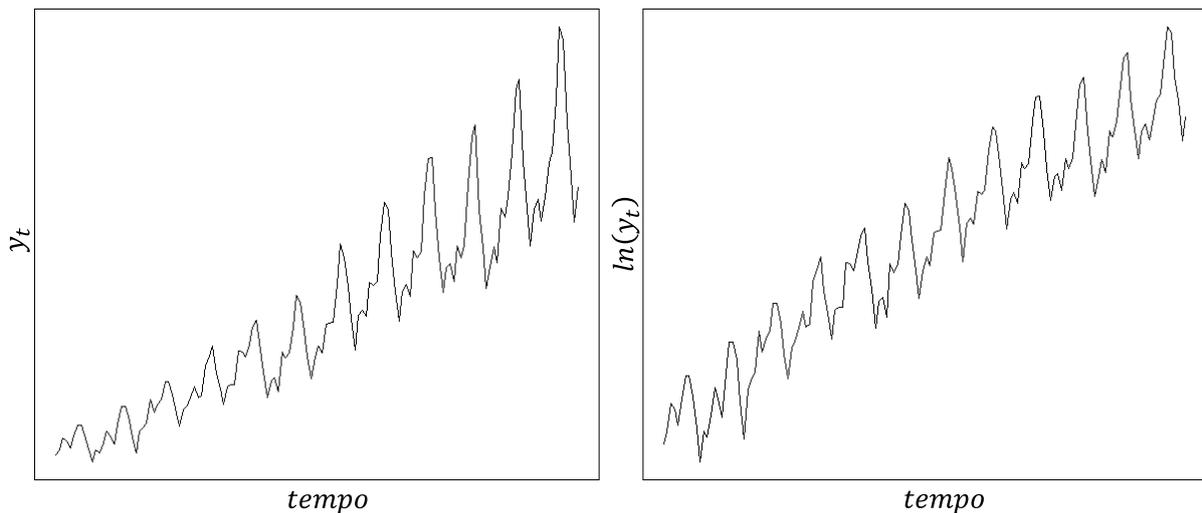


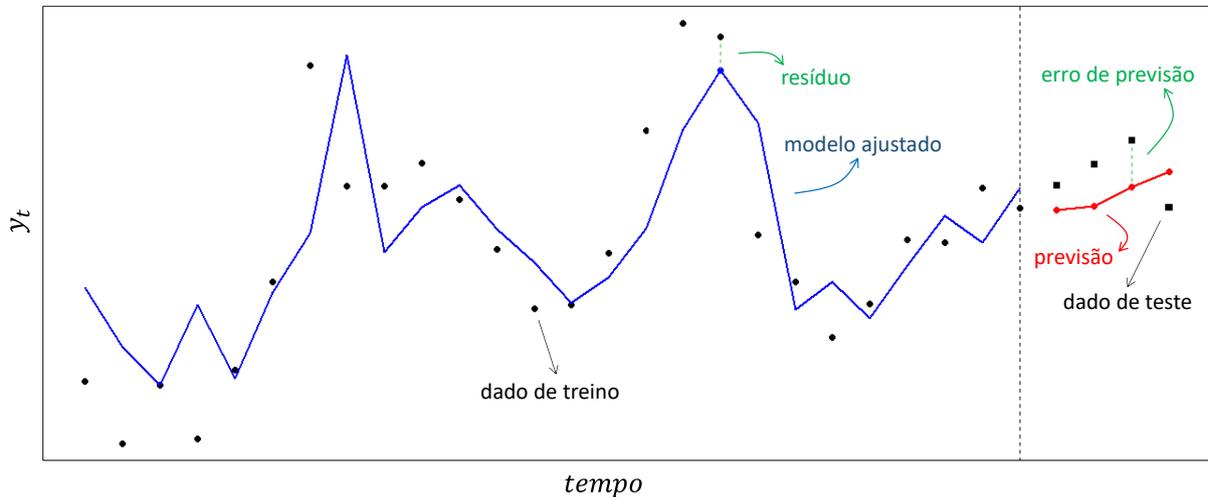
Figura 2: Exemplo de transformação estabilizadora da variância. Fonte: Elaboração própria.

### 3.2. Avaliando o Desempenho das Previsões

Para Hyndman e Athanasopoulos (2018), é importante avaliar a precisão das previsões utilizando previsões genuínas. Dessa forma, os resíduos do modelo não são uma indicação confiável do quão grande os erros de previsão provavelmente serão. A precisão somente pode

ser determinada considerando-se como o modelo desempenha em dados novos e que não foram utilizados no seu ajuste.

Assim, uma prática comum é separar os dados históricos disponíveis em duas partes: os dados de treino, usado para estimar o modelo, e os dados de teste, utilizados para se avaliar a precisão das previsões. Uma vez que os dados de teste não foram usados no ajuste do modelo, eles são um indicador da qualidade do modelo de previsão em dados novos (Figura 3).



**Figura 3:** Desempenho das previsões – conceitos básicos. *Fonte:* Elaboração própria.

O erro de previsão é a diferença entre o valor observado e o previsto. O erro de previsão é diferente do resíduo do modelo, pois eles são aplicados em dados diferentes (teste e treino, respectivamente) (Figura 3). Segundo Hyndman e Athanasopoulos (2018), a precisão das previsões pode ser avaliada de três maneiras: por meio dos erros dependentes de escala, dos erros percentuais ou dos erros escalonados.

Os erros dependentes de escala estão na mesma unidade de medidas da série original, sendo, portanto, os mais simples e intuitivos de serem utilizados. A desvantagem é que não podem ser usados para fazer comparações entre séries que envolvam unidades diferentes. Os principais erros deste tipo são o erro absoluto médio e a raiz quadrada do erro quadrático médio.

Os erros percentuais não possuem a desvantagem anterior, pois são livres de unidade. Por outro lado, eles são indefinidos quando  $y_t = 0$  e assumem que a série foi medida em uma escala que possui um zero absoluto (escala de razão). Os principais erros deste tipo são o erro percentual absoluto médio e o erro percentual absoluto médio simétrico.

Os erros escalonados são uma alternativa aos erros percentuais. Hyndman e Athanasopoulos (2018) propõem o uso do erro escalonado absoluto médio (MASE – *mean absolute scaled error*). O MASE é definido como a média aritmética dos valores absolutos dos  $q_j$  dados pela equação 7. O pacote *forecast* calcula automaticamente o MASE.

$$q_j = \frac{e_j}{\frac{1}{T} \cdot \sum_{t=2}^T |y_t - y_{t-1}|} \quad (7)$$



#### 4. Metodologia

Inicialmente, foi montado um banco de dados com as séries temporais anuais a serem analisadas. As séries foram retiradas da M3 *Competition* (MAKRIDAKIS; HIBON, 2000). Esta foi uma competição onde *experts* foram convidados a fornecer previsões para 3003 séries temporais reais utilizando métodos variados. As séries da M3 *Competition* foram escolhidas por representarem uma variedade de aplicações de séries temporais (havia séries financeiras, micro e macro econômicas, demográficas e industriais). Das 3003 séries, 645 eram anuais. Das 645 anuais, 40 foram sorteadas por meio de amostragem aleatória simples para serem analisadas. Uma amostra de tamanho 40 foi escolhida por ser considerada uma amostra grande o suficiente para justificar as aproximações de grandes amostras para os testes paramétricos e não paramétricos (MONTGOMERY; RUNGER, 2012).

De posse das 40 séries, cada uma foi dividida em duas partes: os dados de treino e os de teste. Seguindo o mesmo procedimento da M3 *Competition*, os dados de teste foram sempre os últimos 6 valores das séries, ou seja, o horizonte de tempo para as previsões foi de 6 anos. Nos dados anuais originais, nenhuma série de treino tinha menos de 14 termos, o que é suficiente para estimar os modelos ARIMA (MAKRIDAKIS; HIBON, 2000).

A cada uma das 40 séries de treino foram ajustados dois tipos de modelos: ARIMA aos dados de treino brutos (sem transformação) e ARIMA aos dados de treino previamente transformados pela transformação de Box-Cox. Os parâmetros  $p$ ,  $d$ ,  $q$  e  $\lambda$  (quando se aplicar) foram todos determinados automaticamente pelo pacote *forecast* versão 8.7 com o *software* R versão 3.6.0.

Uma vez ajustados ambos os modelos aos dados de teste, o software fez previsões 6 períodos a frente. Dadas as previsões e os dados de teste, o pacote *forecast* calcula automaticamente diversos erros de previsão. Neste artigo, o MASE será utilizado para se comparar a precisão dos dois métodos de previsão. No Apêndice 1 tem-se os comandos utilizados exemplificados em uma das 40 séries analisadas.

#### 5. Resultados

A tabela 1 apresenta o MASE calculado para cada uma das 40 séries utilizando ambos os métodos, com e sem transformação estabilizadora da variância.

De acordo com Montgomery e Runger (2012), para se testar a hipótese de que os dois métodos têm em média a mesma precisão *versus* que as precisões médias são diferentes, deve-se analisar as diferenças entre os MASE da tabela 1 (última coluna). Isto se deve ao fato de se ter duas amostras dependentes (emparelhadas). Como cada série foi analisada por dois métodos diferentes, os dados foram coletados aos pares. O teste paramétrico adequado para este teste de hipótese é o teste  $t$  emparelhado. Entretanto, este teste tem como suposição que as diferenças seguem uma distribuição normal.

Aplicando-se o teste de normalidade de Shapiro-Wilks disponível no *software* R, rejeita-se a hipótese de normalidade das diferenças da (valor- $p = 2,492 \times 10^{-9}$ , inferior ao menor nível de significância tradicionalmente utilizado de 1%) (Tabela 1). Logo o teste  $t$  emparelhado não pode ser utilizado, devendo ser substituído por um teste não paramétrico equivalente. Os testes não paramétricos são livres de distribuição, no sentido que seu desenvolvimento não é baseado na hipótese de normalidade (MONTGOMERY; RUNGER, 2012).

Segundo Montgomery e Runger (2012), o teste não paramétrico equivalente é o teste de Wilcoxon do posto sinalizado para observações emparelhadas. Em linhas gerais, o teste leva em consideração a quantidade de diferenças positivas e negativas, bem como a magnitude dessas diferenças. Intuitivamente, se ambos os métodos com e sem transformação (Tabela 1)



forem equivalentes, a quantidade de diferenças positivas e negativas serão próximas. Por outro lado, caso o método sem transformação for mais preciso que o com transformação, o número de diferenças negativas será muito maior e vice-versa.

**Tabela 1:** MASE das Previsões sem e com Transformação Estabilizadora da Variância

Série	Sem Transformação	Com Transformação	Diferença entre os MASE
1	0,8679	1,3261	- 0,4582
2	1,9954	0,7198	1,2756
3	10,2518	6,3461	3,9057
4	2,0015	2,1640	- 0,1626
5	0,7601	0,4630	0,2971
6	4,7118	4,9063	- 0,1945
7	1,6236	2,5908	- 0,9672
8	7,2703	4,7106	2,5597
9	24,0093	29,5570	- 5,5477
10	5,5127	8,9390	- 3,4262
11	5,5725	13,0062	- 7,4337
12	2,5494	2,6149	- 0,0655
13	1,4316	1,4259	0,0057
14	1,2710	0,9931	0,2780
15	1,8543	2,4959	- 0,6417
16	1,6208	2,0151	- 0,3943
17	0,4343	0,7030	- 0,2687
18	10,0324	10,0754	- 0,0430
19	2,2512	2,1037	0,1475
20	3,6623	2,4807	1,1816
21	3,1040	2,9718	0,1322
22	0,6819	1,0617	- 0,3798
23	2,1262	0,9176	1,2086
24	0,8191	1,9233	- 1,1042
25	2,3196	2,2278	0,0918
26	0,8257	0,7873	0,0383
27	2,5251	2,4607	0,0644
28	3,7269	6,1044	- 2,3775
29	3,5714	3,3628	0,2086
30	1,6102	1,4150	0,1952
31	2,5297	2,5184	0,0113
32	0,0835	0,2645	- 0,1810
33	1,4040	1,0672	0,3367
34	4,9370	4,4500	0,4869
35	1,0192	16,0421	- 15,0229
36	1,2143	1,2143	0,0000
37	0,9457	0,8625	0,0833
38	1,9039	1,9047	- 0,0007
39	1,4150	1,1807	0,2344
40	1,8537	2,0541	- 0,2004

Fonte: Elaboração Própria

O *software* R possui uma função que calcula o teste de Wilcoxon. Aplicando este teste às diferenças da Tabela 1 encontra-se um valor-p de 0,6302, superior ao maior nível de significância tradicionalmente utilizado de 10%.

Logo, falha-se em rejeitar a hipótese alternativa  $H_1$ , ou seja, não há evidências que os dois métodos não possuam em média a mesma precisão. Dessa forma, o presente estudo corrobora a hipótese de Hyndman e Athanasopoulos (2018) de que a transformação estabilizadora da variância tem pouca influência na precisão das previsões pontuais do modelo ARIMA.



## 6. Conclusão

O presente artigo teve como objetivo testar a hipótese de Hyndman e Athanasopoulos (2018) de que a transformação estabilizadora da variância tem pouca influência na precisão das previsões pontuais do modelo ARIMA. Para tanto, 40 séries anuais foram selecionadas por meio de amostragem aleatória simples, modelos ARIMA com e sem transformação foram ajustados aos dados de treino, previsões com os modelos foram executadas e suas precisões foram calculadas comparando-as com os dados de teste. Utilizando o teste não paramétrico de Wilcoxon, determinou-se que a hipótese não foi rejeitada. Este resultado requer ser analisado do ponto de vista de seu grau de generalização e de suas limitações.

Do ponto de vista estatístico, o resultado é generalizável para a população de séries anuais das quais a amostra foi retirada, ou seja, as 645 séries anuais da *M3 Competition*.

Porém, mais interessante é a generalização do estudo para a teoria. Baseado em trabalhos anteriores e em suas próprias pesquisas, Hyndman e Athanasopoulos (2018) formularam uma hipótese que foi testada e replicada neste artigo em um contexto diferente dos anteriores. Quando ocorrem replicações independentes, os resultados passam a ser válidos para um número maior de contextos e a hipótese inicial sai fortalecida. A consequência lógica é que ela deva ser válida para um número maior de situações do que aquelas testadas até então. Assim, cada vez que a hipótese não é rejeitada em testes especificamente projetados para refutá-la, ela se torna mais robusta ou com maior grau de corroboração (ALVES-MAZZOTTI; GEWANDSZNAJDER 2002). Desta forma, é provável que este resultado seja válido para muitas séries encontradas na prática e que tenham características semelhantes às séries anuais da *M3 Competition*.

Com relação às limitações do trabalho a primeira delas é com relação ao ajuste automático dos modelos ARIMA utilizando a minimização do AICc e a ausência da análise residual (ver seção 3.1). É possível que uma analista experiente utilizando conjuntamente o método tradicional de identificação (determinação de  $p$  e  $q$  usando o padrão de autocorrelação dos dados) e a análise residual chegasse a modelos ajustados melhores do que os identificados automaticamente pelo *software* e que nesses modelos a precisão das previsões diferisse significativamente. Assim, os resultados obtidos devem ser interpretados neste contexto. Neste trabalho, o uso do método automático de seleção se deveu à grande quantidade de séries analisadas simultaneamente. Dessa forma, uma possibilidade para trabalhos futuros seria avaliar o impacto do método de identificação (tradicional *versus* minimização de função penalizadora) na precisão das previsões.

Outra questão é que a diferença entre a precisão dos métodos possa depender da quantidade de heteroscedasticidade e/ou de outros fatores presente nos dados. Na amostra utilizada, a quantidade de heteroscedasticidade e demais fatores de ruído foram controlados/isolados usando o teste emparelhado. Assim, o resultado é prontamente generalizável para a população de 645 séries da *M3 Competition* e provavelmente válido para séries semelhantes. Um resultado muito mais abrangente poderia ser alcançado modelando-se explicitamente a heteroscedasticidade e/ou outros fatores utilizando a teoria estatística do planejamento de experimentos (MONTGOMERY; RUNGER, 2012). Esta última opção poderia ser alcançada em trabalhos futuros manipulando-se séries simuladas, ao invés de usar dados observacionais como no presente artigo.

## Referências

ALVES-MAZZOTTI, A. J.; GEWANDSZNAJDER, F. *O Método nas Ciências Naturais e Sociais: pesquisa quantitativa e qualitativa*. 2. Ed. São Paulo: Pioneira Thomson Learning, 2002.



HYNDAMN, R. J.; ATHANASOPOULOS, G. *Forecasting: principles and practice*. 2. Ed. Melbourne: OTexts, 2018. Disponível em: <<https://otexts.org/fpp2/>>. Acesso em: 14 set. 2018.

HYNDMAN R. J.; KHANDAKAR, Y. Automatic Time Series Forecasting: the forecast package for R. *Journal of Statistical Software*, 27(3), 2008.

MAKRIDAKIS, S.; HIBON, M. The M3 Competition: results, conclusions and implications. *International Journal of Forecasting*. 16(4), p.451-476, 2000.

MONTGOMERY, D. C.; RUNGER, G.C. *Estatística Aplicada e Probabilidade para Engenheiros*. 5. Ed. Rio de Janeiro: LTC, 2012.

MORETTIN, P. A.; TOLOI, C. M. C. *Análise de Séries Temporais*. 2. Ed. São Paulo: Blucher, 2006.

R CORE TEAM (2019). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing: Vienna, Austria.. Disponível em <<http://www.R-project.org>>, 2019.

RUY, M. *Análise de Séries Temporais: modelos probabilísticos lineares*. Uberlândia: FAGEN/UFU, 2019. 149 p. Notas de Aula da Disciplina Análise de Dados III.

## Apêndice 1 – Comandos utilizados no *software* R para o cálculo dos erros de previsões

```
### Instala o pacote forecast (única vez)
install.packages("forecast", dependencies = TRUE)

### Carrega o pacote na memória no início de cada sessão
library(forecast)

### Pré-processamento dos dados
## A série x abaixo é a 40ª série analisada na Tabela 1
x = ts(c(797, 816, 1034, 1013, 1162, 2088, 2456, 2583, 2763, 2743,
2981, 2804, 3084, 3157, 3313, 3595, 4317, 4687, 5179, 5504, 5836,
6098, 5306), start = 1, frequency = 1)

n = length(x)
treino = window(x, start = 1, end = (n-6))
teste = window(x, start = (n-5), end = n)

## Ajuste, previsão e erro de previsão

## Sem transformação
m1 = auto.arima(treino, stepwise = FALSE, approximation = FALSE)
p1 = forecast(m1, h = 6)
accuracy(p1, teste)

## Com transformação de Box-Cox
l1ik = BoxCox.lambda(treino, method = c("loglik"))
m2 = auto.arima(treino, stepwise = FALSE, approximation = FALSE,
lambda = l1ik, biasadj = TRUE)
p2 = forecast(m2, h = 6, lambda = l1ik, biasadj = TRUE)
accuracy(p2, teste)
```