

Combinação de previsões aplicada à modelagem de operações: um estudo de caso em um terminal portuário

Combination of forecasts applied to modeling operations: a case study in a port terminal

Icaro Romolo Sousa Agostino¹

Renata Silva Araujo²

Maiara de Oliveira Noronha³

Jose Inácio Dos Santos Fonseca⁴

Adriano Mendonça Souza⁵

1 Engenheiro de Produção pela Universidade Ceuma (2017). Mestrando em Engenharia de Produção pela Universidade Federal de Santa Maria (UFSM). Atualmente é Pesquisador Bolsista (CAPES) no Laboratório de Análise e Modelagem Estatística da UFSM. Universidade Federal de Santa Maria icaroagostino@gmail.com

2 Graduanda em Arquitetura e Urbanismo na Universidade Estadual do Maranhão. Universidade Estadual do Maranhão renata.2araujo@gmail.com

3 Mestre em Engenharia de Produção - Universidade Federal de Santa Maria. Pós-graduanda no curso de Especialização em Estatística e Modelagem Quantitativa - Universidade Federal de Santa. Bacharel em Gestão Ambiental pela Universidade Federal do Pampa- RS. Universidade Federal de Santa Maria mai.ol.noronha@gmail.com

4 Graduando do curso de Administração pela Faculdade Pitagoras São Luís. Faculdade Pitagoras São Luís inacioc21@gmail.com

5 Possui graduação em Matemática (1988 - Faculdade Imaculada Conceição), Especialização em Estatística e Modelagem Quantitativa (1991 - UFSM) e Mestrado em Engenharia de Produção (1993 - UFSM). Doutorando em Engenharia de Produção (2000 - UFSC) e post-doc no Instituto Superior da Ciência do Trabalho e da Empresa - ISCTE - Business School - Lisbon Institute - Portugal. Atualmente é Professor Titular do Departamento de Estatística (UFSM) atuando nos cursos de Especialização em Estatística e Modelagem Quantitativa (UFSM) e no Mestrado em Engenharia de Produção (UFSM) assim como nos demais cursos de graduação da instituição. Universidade Federal de Santa Maria amsouza.sm@gmail.com

Resumo

O objetivo deste estudo foi ajustar um modelo de previsão combinado para prever o comportamento de um sistema de descarga de minério em um terminal portuário. Foram utilizadas observações de um mês de operação, coletadas a cada hora, totalizando 720 observações. Os procedimentos metodológicos foram realizados em três etapas: (i) ajuste dos modelos concorrentes por meio da metodologia de Box-Jenkins; (ii) combinação dos modelos selecionados pelos métodos de média aritmética, *Least Squares Weights* e *Inverse Mean Squared*; (iii) comparação das medidas de acurácia de cada previsão para seleção do melhor modelo. Foram obtidos 20 modelos concorrentes, dos quais foram selecionados três para a etapa de combinação. O melhor modelo foi obtido pelo método de combinação *Inverse Mean Squared* apresentando um erro absoluto médio de 2,017%. Os métodos de combinação proporcionaram um aumento da acurácia das previsões, gerando subsídios à tomada de decisão para melhorias no planejamento do processo produtivo.

Palavras-chave: Box-Jenkins; Combinação de previsões; *Inverse Mean Squared*.

Abstract

The aim of this study was to adjust a combined forecasting model to verify the behavior of an ore discharge system at a port terminal. It was used observations of one month of operation, collected every hour, totaling 720 observations. The methodological procedures were performed in three stages: (i) adjustment of the competitors models by the Box-Jenkins methodology; (ii) combination of the models selected by arithmetic average methods, *Least Squares Weights* and *Inverse Mean Squared*; (iii) compare of the accuracy measures of each forecast to select the best model. Twenty concurrent models were obtained, of which three were selected for a combination stage. The best model was obtained by the *Inverse Mean Squared* method, with an average absolute error of 2.017%. The combination methods provide an increase in the guarantee of the forecasts, generating subsidies to the decision making for improvements in the planning of the productive process.

Keywords: Box-Jenkins; Combination of forecasts; *Inverse Mean Squared*.

1 Introdução

A atividade portuária no Brasil possui um importante papel na cadeia logística nacional, sobretudo na exportação de commodities, sendo responsável por aproximadamente 90% das exportações totais do país (CODEBA, 2017). Nesse contexto, a alta produtividade dos portos conduz menores custos em relação aos demais transportes, beneficiando a produção nacional em comércios exteriores. Contudo, a gestão da atividade portuária se torna complexa necessitando de subsídios que fomentem o controle de operações objetivando a eficiência em relação aos custos, tempo, nível de serviço, qualidade e identificação dos elementos críticos para concepção de padrões e metas (Gonzalez & Trujillo, 2008).

Nesse cenário, as técnicas de previsão são apresentadas com o intuito de solucionar problemas diversos envolvendo riscos, pois possibilitam gerar subsídios que auxiliem os decisores na formulação de estratégias nos processos produtivos, com o foco na definição de melhores formas de utilização dos recursos limitados buscando eficiência na utilização dos mesmos (Slack, Brandon-Jones, & Johnston, 2015). Nesse contexto, a previsão consiste em um processo metodológico que objetiva definir os dados futuros com base em modelos estatísticos, matemáticos, econométricos e subjetivos (Martins & Laugeni, 2009).

Uma alternativa para obtenção de previsões mais eficazes consiste em combinar mais de um modelo de previsão resultando em previsões estatisticamente superiores em termos de acurácia, comparados a técnicas individuais (Hibon & Evgeniou, 2005). Diversos estudos apresentam resultados demonstrando que diferentes métodos de combinações obtiveram resultados mais acurados em relação a previsões individuais (Clemem, 1989; Makridakis & Hibon, 2000; Stock & Watson, 2004; Amendola & Storti, 2008; Jeong

& Kim, 2009; Costantini & Pappalardo, 2010; Wallis, 2011).

Dessa forma, o objetivo deste estudo foi ajustar um modelo de previsão capaz de prever o comportamento de um sistema de descarga de minério em um terminal portuário, utilizando como metodologia a combinação de modelos preditores ARIMA, a partir dos métodos média aritmética, *Least Squares Weights* e *Inverse Mean Squared*. Como resultado buscou-se um modelo validado pelas estatísticas de ajuste que possa servir de suporte ao processo de tomada de decisão.

Além dessa introdução, o presente artigo está dividido em mais quatro seções. Na seção seguinte é apresentada uma revisão de literatura, fundamentada em dois tópicos relevantes para a discussão desta pesquisa, o primeiro trata de uma breve revisão acerca dos modelos para previsão de séries temporais, enquanto o segundo aborda a modelagem ARIMA, assim como a abordagem para combinação de previsões. Na terceira seção são descritos os passos metodológicos adotados, na seção posterior é apresentado o estudo de caso, em que foi descrito o processo de produção e caracterizada a variável estudada. Na quinta seção são apresentados os resultados obtidos, e por fim na última seção são apresentadas as considerações finais.

2 Revisão de literatura

A seguir apresenta-se a contextualização teórica a respeito dos temas: modelos para previsão de séries temporais, modelos autorregressivos integrados de médias móveis e combinação de previsão.

2.1 Modelos para previsão de séries temporais

Realizar previsões é uma atividade indispensável no planejamento de atividades e nos proces-

tos de tomada decisão (Werner, 2005; Wang & Chang, 2010). A previsão pode ser entendida como a busca de informações sobre o comportamento futuro de uma variável através de um processo racional envolvendo incertezas. Nesse contexto, a previsão consiste em um processo metodológico que objetiva definir os dados futuros com base em modelos estatísticos, matemáticos, econométricos e/ou subjetivos (Martins & Laugeni, 2009; Moreira, 2014).

Dessa forma, comumente os modelos de previsão são divididos em duas classes genéricas: qualitativas e quantitativas (Klimberg, Lawrence & Lal, 2009). Técnicas qualitativas são largamente utilizadas em diversos contextos, devido a facilidade de aplicação e de apresentarem baixos custos relativos. Tais métodos normalmente são utilizados em situação de escassez de dados históricos, sendo, de uma forma geral, menos robustos quando comparados a métodos quantitativos. Dentre os métodos qualitativos se destacam a abordagem de painel, o método Delphi, o planejamento de cenários, dentre outros (Rosienkiewicz, Chlebus, & Detyna, 2017).

As técnicas quantitativas utilizam dados históricos para a realização de previsões, podendo utilizar modelos baseados em séries temporais ou modelos causais. Enquanto modelos causais buscam incluir variáveis explicativas para prever o comportamento da variável estudada, os modelos de séries temporais utilizam o comportamento dos dados ao longo do tempo para extrair um padrão que possa explicar o comportamento futuro da variável estudada (Rosienkiewicz, Chlebus, & Detyna, 2017).

Dentre os métodos de previsão existentes, a metodologia de Box-Jenkins se destaca como uma abordagem poderosa na solução de muitos problemas de previsão, pois comumente oferece previsões acuradas para séries temporais (Montgomery, Jennings, & Kulahci, 2015).

2.2 Modelos autorregressivos integrados de médias móveis

Os modelos Autoregressivos Integrados de Médias Móveis, também denominados de ARIMA (*Auto Regressive Integrated Moving Average*), foram desenvolvidos por Box & Jenkins (1970), sendo caracterizados por captar o comportamento da correlação seriada entre os valores da série temporal, e assim realizar previsões futuras (Gujarati & Porter, 2011).

Um pressuposto básico para utilização da metodologia de Box-Jenkins é da estacionariedade da série temporal utilizada, garantindo a significância dos parâmetros estimados ao longo dos períodos. Dessa forma, caso a série possua tendência, não sendo estacionária, se faz necessário a aplicação de diferenças entre os valores a fim de torná-la estacionária (Bueno, 2008). A verificação da estacionariedade se dá a priori por análise gráfica, sendo confirmada pela aplicação de testes de raízes unitárias. Os testes Augmented Dick-Fuller – ADF (Dickey, 1984) e Kwiatkowski, Phillips, Schmidt e Shin – KPSS (Kwiatkowski, Phillips, Schmidt, & Shin, 1992) comumente são utilizados a fim de verificar a estacionariedade em séries temporais, sendo sugerido a utilização conjunta de ambos para obter resultados mais acurados sobre o grau de estacionariedade dos dados (Souza, 2016).

A metodologia resulta da combinação dos seguintes componentes, também denominados de filtros: autorregressivo (AR), filtro de integração (I) e o componente de médias móveis (MA). Tais modelos constituem um ciclo iterativo, permitindo a escolha do melhor modelo baseado nos dados da série estudada, a partir dos seguintes procedimentos: (i) identificação, consistindo na determinação do modelo que melhor descreve o comportamento da série, por meio da análise das funções de autocorrelação e autocorrelação parcial; (ii) estimação, realizando a estimação dos parâmetros autorre-

gressivos, assim como de médias móveis; (iii) validação, consistindo na análise da adequação do modelo ajustado ao comportamento real da série, em que são analisados os resíduos do modelo; (iv) previsão, que somente é realizada, quando as etapas anteriores forem satisfatórias (Souza, 2016).

Usualmente os modelos ARIMA (p, d, q) são representados pela equação 1, em que Z_t representa a série temporal modelada, B representa o operador retroativo, d a ordem de integração, ϕ é termo que representa o parâmetro autoregressivo de ordem p, θ representa o parâmetro de médias móveis de ordem q.

$$\phi(B)\Delta^d Z_t = \theta(B)\epsilon_t \tag{1}$$

Os resíduos ϵ_t gerados pelo modelo estimado são definidos como a diferença dos valores reais e dos valores previstos da série. Para um modelo adequadamente especificado, o ruído gerado deve possuir média zero, variância constante e não apresentar autocorrelação, denominado de ruído branco RB $\sim (0, \sigma^2)$ (Bueno, 2008).

Para mensurar o ajuste do modelo escolhido, o critério *Akaike's Information Criteria* - AIC (Akaike, 1973) comumente é utilizado para a escolha do melhor modelo entre os modelos ARIMA concorrentes. Tal critério leva em consideração o número de parâmetros dos modelos e a variância dos erros gerados, dessa forma, o modelo que apresentar o menor valor de AIC terá o melhor ajuste (Morettin, 2011). O critério de AIC é dado pela equação:

$$AIC(p, q) = \ln \sigma_{p,q}^2 + \frac{2(p + q)}{n} \tag{2}$$

Em que: p e q são os parâmetros conhecidos; n é o tamanho da amostra; ln o logaritmo neperiano e σ^2 a variância estimada dos erros.

A fim de verificar a acurácia do modelo ajustado, algumas medidas são utilizadas no processo de previsão. Tais medidas são capazes de medir o desempenho de um modelo a partir dos erros gerados em relação aos valores reais da série (Montgomery, Jennings, & Kulahci, 2015). As medidas de acurácia utilizadas nessa pesquisa para a avaliação da precisão dos modelos foram: *Mean Absolute Error* (MAE); *Mean Absolute Percentage Error* (MAPE); e o coeficiente de *U-Theil*, que avalia o desempenho da previsão em relação à previsão ingênua ou trivial, em que $U \geq 1$, o erro do modelo ajustado é maior ou igual que de uma previsão ingênua; $U < 1$, o erro do modelo ajustado é menor que de uma previsão ingênua. As equações das medidas de acurácia são ilustradas na Tabela 1.

Tabela 1: Medidas de acurácia

Siglas	Equações
MAE	$\frac{\sum_{t=1}^n \epsilon_t }{n} \tag{3}$
MAPE	$\frac{\sum_{t=1}^n \epsilon_t / Z_t * 100}{n} \tag{4}$
U-Theil	$\frac{\sqrt{\sum_{t=1}^n (\epsilon_t)^2}}{\sqrt{\sum_{t=1}^n (Z_t - Z_{t-1})^2}} \tag{5}$

Fonte: Montgomery, Jennings & Kulahci (2015).

Em que: ϵ_t representa o erro (diferença entre o valor estimado pelo modelo e o valor real); Z_t representa a variável modelada.

2.3 Combinação de previsão

Na busca por modelos cada vez mais acurados, pesquisadores têm colocado esforços em desenvolver métodos capazes de aumentar a precisão de previsões em diversas áreas (Firmino, Mattos Neto, & Ferreira, 2014). Nesse contexto, Martins

& Werner (2014) argumentam que a combinação de previsões é uma abordagem atraente, uma vez que, ao contrário de se preocupar em selecionar o melhor modelo para a série estudada, o pesquisador deve buscar quais modelos podem contribuir com o aumento de acurácia uma vez combinados.

Nesse mesmo sentido, Rausser & Oliveira (1976) e Clemem (1989) comentam que comumente no processo de realização de previsões, em que são ajustados 2 ou 3 modelos, utiliza-se alguma medida de acurácia para selecionar o melhor e os demais são descartados, muitas vezes perdendo alguma informação capturada pelos modelos não selecionados. Dessa forma, a previsão combinada é uma técnica a ser considerada na busca de um modelo mais preciso (Souza, Souza, Ferreira, & Menezes, 2011).

A média aritmética simples (MS) é um dos mais populares métodos empregados na combinação de previsões (Martins & Werner, 2014). Dessa forma, a combinação por média consiste em uma combinação linear das previsões individuais com pesos fixos, conforme a equação (6):

$$F_c = (F_1 + F_2 + \dots + F_n)/n \quad (6)$$

Onde: F_c é a previsão combinada; F_1, F_2, \dots, F_n são as previsões individuais; n o número de previsões individuais utilizadas na combinação.

Outra abordagem sugerida por Granger & Ramanathan (1984) para agregar previsões trata de combinar previsões em casos da existência de mais de dois preditores individuais através de uma combinação linear utilizando um modelo de regressão sem restrição para os pesos, denominado *Least Squares Wights* (LS) sendo estes estimados por Mínimos Quadrados em relação a série original. A equação (7) apresenta o modelo de combinação:

$$F_c = w_0 + w_1F_1 + w_2F_2 + \dots + w_nF_n \quad (7)$$

Onde: F_c é a previsão combinada; w_0 é o intercepto da regressão; w_1, w_2, \dots, w_n são os pesos estimados para cada previsão individual; F_1, F_2, \dots, F_n são as previsões individuais; n o número de previsões individuais utilizadas na combinação.

Uma terceira abordagem para combinação de previsões considerada neste estudo é o método *Inverse Mean Squared* (IMS), proposto por Stock & Watson (2001). O método utiliza pesos diferentes para cada horizonte de previsão, baseando-se no inverso do erro quadrático médio computado por cada previsão. Os pesos para a realização da combinação das previsões são dados pela equação (8):

$$w_i = \frac{1/MSE_i^k}{\sum_{j=1}^n 1/MSE_j^k} \quad (8)$$

Onde: w_i é o peso dado para cada previsão “ i ”; MSE é o erro quadrático médio cometido por cada previsão individual; k é o horizonte da previsão; n o número de previsões individuais utilizadas na combinação.

3 Metodologia

Este estudo se caracteriza como aplicado, pois visa a aplicação de conceitos teóricos em casos práticos com o objetivo de resolver problemas específicos. A abordagem adotada é quantitativa visto que apresenta resultados quantificáveis com a aplicação de métodos matemáticos. Quanto aos objetivos a abordagem adotada foi exploratória, pois visa explicar e gerar hipóteses sobre o problema em estudo (Prodanov & Freitas, 2013; Gil,

2010). A pesquisa adota como procedimento técnico o estudo de caso, em que o estudo proposto apresenta poucos objetivos e é realizado a partir de um caso particular de modo a permitir o seu conhecimento amplo e detalhado (Gil, 2010).

Os dados utilizados para modelagem foram coletados a partir do banco de dados de apontamentos do sistema de automação industrial, sendo o principal sistema de controle de produção da organização. Os dados são referentes a um mês de operação, com observações a cada hora, gerando uma série temporal com 720 observações.

Foi aplicada uma transformação logarítmica na série temporal com o objetivo de estabilizar a variância dos dados, conforme sugerido por Montgomery, Jennings & Kulaheci (2015) em casos de dados com variâncias não constantes ao longo do tempo.

A série temporal foi dividida em duas partes, a primeira correspondente ao período de treino com cerca de 80% das observações e segunda a fase de teste com cerca de 20% das observações. A fase de treino foi utilizada para o ajuste dos modelos, enquanto a de teste para a avaliação das médias de acurácia.

O desenvolvimento da pesquisa foi dividido em três etapas: (i) inicialmente foram ajustados modelos concorrentes utilizando a metodologia de Box-Jenkins; (ii) posteriormente a combinação dos modelos selecionados e, por fim, (iii) comparação das medidas de acurácia de cada previsão para seleção do melhor modelo. A Figura 1 ilustra os passos adotados.

A modelagem ARIMA seguiu os passos metodológicos propostos por Box-Jenkins (1970). Inicialmente foi examinado o comportamento da série estudada, investigando sua estacionariedade e os parâmetros de autocorrelação.

Foram identificados os filtros a serem utilizados por meio da análise das funções de autocorrelação (FAC) e autocorrelação parcial (FACP) e suas defasagens. Realizou-se a estimação dos parâmetros dos modelos concorrentes por meio do Método da Máxima Verossimilhança.

Foram analisados os resíduos dos modelos ajustados, utilizando como critério para seleção aqueles que apresentaram menores valores de AIC, parâmetros significativos ($p < 0,05$), e características de ruído branco. Após a seleção dos modelos realizou-se as previsões individuais.

Para a combinação das previsões foram adotadas três abordagens: combinação por média aritmética (MS), pelo método *Least Squares Weights* (LS), assim como pelo método *Inverse Mean Squared* (IMS). A avaliação final dos modelos se deu pela análise das medidas de acurácia MAE, MAPE e *U-Theil*, sendo possível comparar o desempenho tanto das previsões individuais quanto combinadas e selecionar o melhor predictor para o processo modelado.

O software utilizado para tratamento e modelagem dos dados foi o *Eviews 9 S.V.*

4 Estudo de caso

O estudo foi desenvolvido em uma gerência de operação portuária, modelando parte de um sistema logístico integrado de movimentação de minério de ferro (contendo mina, ferrovia e porto),

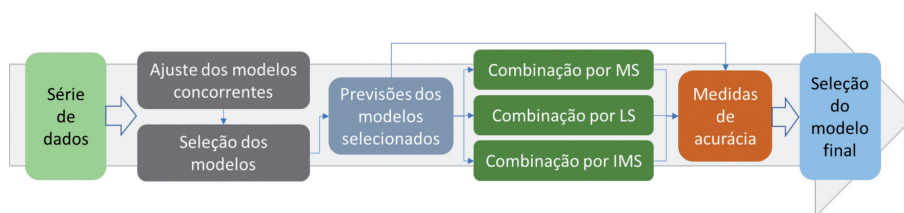


Figura 1: Etapas para modelagem e combinação das previsões

Fonte: Os autores.

escoando toda a produção da mina de Carajás (Brasil), que atualmente é considerada a maior mina a céu aberto do mundo. O transporte do material, após extração, é realizado via sistema ferroviário por composição de três lotes de 110 vagões formando uma composição de 330 vagões. Tal composição é desmembrada via Terminal Ferroviário e cada lote é manobrado e posicionado em um Virador de Vagões para realizar a descarga do minério.

O processo inicial portuário se caracteriza pela descarga do minério utilizando Viradores de Vagões, em que o equipamento industrial gira um par de vagão por vez, direcionando o material para um sistema de correias transportadoras. O modelo proposto contempla a operação de descarga de minério. A variável modelada representa o Tempo Médio de Descarga (TMD) do sistema de viradores de vagões a cada hora, o indicador é representado pela equação (9):

$$TMD = TTD / N^{\circ} \text{ Lotes} \tag{9}$$

Em que: TTD representa o tempo total de descarga; N° Lotes representa o número total de lotes operados no período, sendo cada lote constituído de 110 vagões.

A variável TMD é considerada um indicador chave de produção, sendo o principal parâmetro de controle do sistema de descarga portuário. Devido ao processo de descarga ser fornecedor dos processos portuários posteriores de empilhamento e embarque, o controle sobre a variável TMD torna-se um importante fator a ser abordado nos processos decisórios, tais como o dimensionamento de produção e o planejamento de rotinas de manutenção, desta forma o modelo ajustado será utilizado como previsor do comportamento futuro da variável, servindo como uma ferramenta de apoio à tomada de decisão.

5 Resultados e discussão

Inicialmente foi realizada a inspeção gráfica da série original, conforme a Figura 2, representando o comportamento da variável TMD ao longo de um mês de operação. É possível observar que a mesma é estacionária, possuindo média constante e sem tendência. Para confirmar tal hipótese foram realizados os testes de raízes unitária ADF e KPSS.

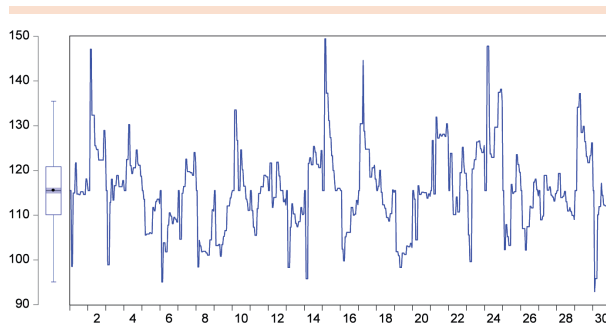


Figura 2: Tempo Médio de Descarga (TMD)
Fonte: Os autores.

Na Tabela 2 apresentam-se os resultados dos testes de raízes unitárias ADF e KPSS realizados, sendo que a hipótese de estacionariedade da série original foi confirmada, pois os dois testes convergem seus resultados indicando que a série é estacionária em nível.

Tabela 2: Resultado dos Testes de Raízes Unitárias

	ADF ^a	KPSS ^b
Série em nível	-7,211 ($p < 0,000$) ^c	0,1455 ^d

^aH0: a série possui uma raiz unitária; $\alpha = 0,05$; ^bH0: a série é estacionária; $\alpha = 0,05$;
^cValor crítico para o teste ADF: -3,4783; ^dValor crítico para o teste KPSS: 0,4630
Fonte: Elaborado pelos autores.

Posteriormente analisou-se as funções de autocorrelação (FAC) e autocorrelação parcial (FACP) da série original, conforme Figura 3. A série possui uma estrutura de autocorrelação significativa, com dependência entre os períodos Z_t e Z_{t-1} , assim como autocorrelação entre os resíduos ϵ_t e $\epsilon_{t-1}, \epsilon_{t-2}, \dots, \epsilon_{t-18}$.

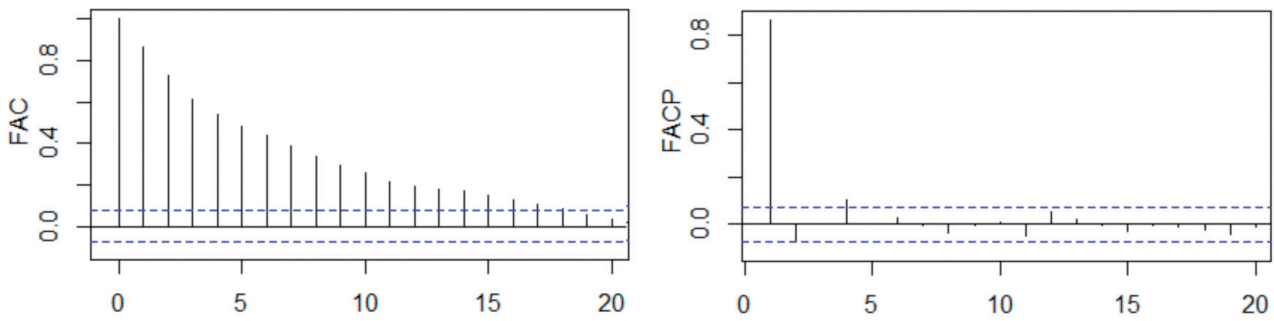


Figura 3 - FAC e FACP
Fonte: Os autores.

Em seguida foram ajustados 20 modelos concorrentes com objetivo de selecionar modelos candidatos para a etapa de combinação das previsões. Na Figura 4 estão dispostos os valores de AIC para os modelos concorrentes ajustados.

Os modelos foram ordenados de acordo com os valores de AIC. Na Tabela 3 são apresentados os 5 modelos com menores valores de AIC, assim como seus parâmetros, o nível de significância (*p-value*), e a verificação das características dos ruídos gerados por cada modelo.

Para a etapa de combinação foram selecionados os modelos ARIMA(1,0,2), ARIMA(3,0,1) e ARIMA(2,0,3), pois entre os modelos ajustados, além de possuírem menores valores de AIC, apresentaram todos os parâmetros significativos ($p < 0,05$), e características de ruído branco, sendo então considerados modelos adequadamente ajustados e com

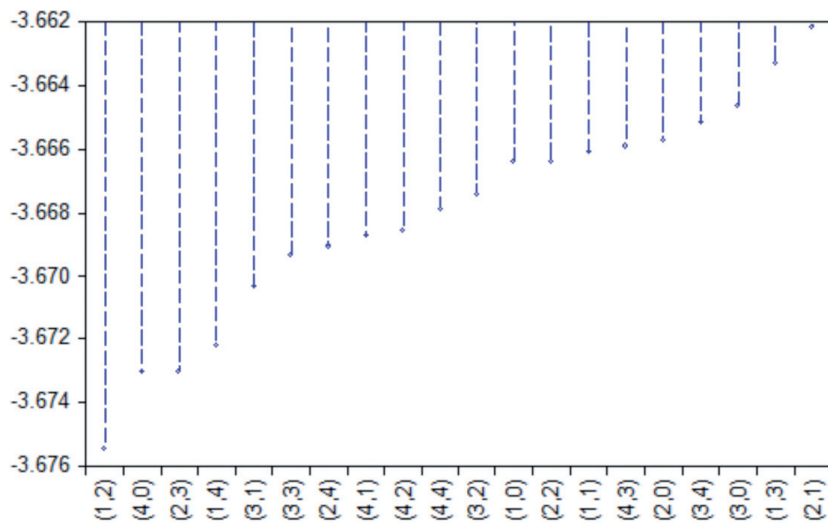


Figura 4: Valores de AIC dos modelos concorrentes
Fonte: Os autores.

Tabela 3: Modelos Ajustados

Modelo	AR(p)	p-valor	MA(q)	p-valor	RB	AIC
ARIMA (1,0,2)	$\phi_1 = 0,913$	< 0,001	$\theta_2 = -0,114$ $\theta_3 = -0,147$	0,016 < 0,001	Sim	-3,675
ARIMA (4,0,0)	$\phi_1 = 0,907$ $\phi_2 = -0,098$ $\phi_3 = -0,047$ $\phi_4 = 0,113$	< 0,001 0,284 0,315 0,003	-	-	Não	-3,673
ARIMA (2,0,3)	$\phi_1 = 1,134$ $\phi_2 = -0,134$	< 0,001 < 0,001	$\theta_1 = -0,188$ $\theta_2 = -0,160$ $\theta_3 = -0,167$	0,033 < 0,001 < 0,001	Sim	-3,673
ARIMA (1,0,4)	$\phi_1 = 0,906$	< 0,001	$\theta_1 = 0,006$ $\theta_2 = -0,113$ $\theta_3 = -0,145$ $\theta_4 = 0,033$	0,974 0,029 0,001 0,668	Não	-3,672
ARIMA (3,0,1)	$\phi_1 = 1,756$ $\phi_2 = -0,884$ $\phi_3 = 0,128$	< 0,001 < 0,001 < 0,001	$\theta_1 = -0,815$	< 0,001	Sim	-3,670

Fonte: Os autores.

capacidade de contribuir para o aumento da acurácia dos modelos combinados.

Foi então realizada a combinação dos três modelos selecionados pelos métodos da média aritmética simples (MS), *Least Squares Weights* (LS), assim como pelo método *Inverse Mean Squared* (IMS), obtendo seis previsões diferentes. Os resultados das medidas de acurácia MAE, MAPE e *U-Theil* para a fase de teste de todos os modelos são apresentados na Tabela 4.

Tabela 4: Medidas de acurácia dos modelos de previsão analisados

Modelo	MAE	MAPE	U-Theil
ARIMA (1,0,2)	2,442	2,103	0,0197
ARIMA (3,0,1)	2,404	2,074	0,0201
ARIMA (2,0,3)	2,359	2,031	0,0199
Média Simples	2,385	2,057	0,0200
Inverse Mean Squared	2,344	2,017	0,0196
<i>Least Squares Weights</i>	2,466	2,118	0,0197

Fonte: Os autores.

Entre as previsões avaliadas, o modelo ARIMA (2,0,3) se mostrou superior nas medidas MAE e MAPE em relação aos demais modelos ARIMA avaliados. Para a estatística de *U-Theil* o modelo ARIMA (1,0,2) apresentou menores valores que as demais previsões individuais. Entre os métodos de combinação avaliados, o menos eficaz foi a combinação por Média Simples, apresentando menor acurácia em relação aos demais métodos de combinação.

Dessa forma, a melhor previsão foi obtida pela combinação dos três modelos ARIMA pelo método *Inverse Mean Squared* em destaque na Tabela 4, apresentando menores valores para as estatísticas de erro, com o MAE de 2,344, MAPE de 2,017. Para a medida de *U-Theil* todos os modelos apresentaram valores menores que 1, indicando como bons previsores em relação a previsões ingênuas, sendo o menor valor obtido pela combinação *Inverse Mean Squared*.

Tais resultados vão de encontro com a literatura, pois a combinação das previsões possibilitou um aumento da precisão dos resultados, gerando previsões estatisticamente superiores comparadas às previsões individuais, pois comumente no processo de realização de previsões, em que são ajustados 2 ou 3 modelos, utiliza-se alguma medida de acurácia para selecionar o melhor e os demais são descartados, muitas vezes perdendo alguma informação capturada pelos modelos não selecionados.

Dessa forma, foi possível evidenciar que os modelos não selecionados como ótimos no processo de ajustes de modelos correntes podem contribuir positivamente na previsão final quando utilizado algum procedimento de combinação capaz de reduzir a perda de informação dos modelos individuais.

Entre os métodos adotados, o método *Inverse Mean Squared* utiliza pesos diferentes para cada horizonte de previsão, o que permitiu uma abordagem mais acurada devido a variabilidade da série temporal em estudo, em que os modelo de combinação com pesos fixos utilizados não foram eficazes em reduzir os erros gerados pelos modelos individuais.

No contexto industrial, estudos com combinação de previsões tem trazido resultados substanciais, como no estudo realizado por Martins & Werner (2014) que comparou previsões individuais e combinadas para 50 séries reais do setor industrial, obtendo resultados superiores com previsões combinadas.

Outros estudos recentes, como o realizado por Aye, Balcilar, Gupta, & Majumdar (2015) chegaram a resultados semelhantes combinando previsões no setor do varejo, assim como Pang & Gekba (2017) que utilizaram a combinação de previsões para obter resultados mais acurados aplicados à transferência de containers em um porto localizado na Indonésia encontrando

resultados superiores em relação às previsões individuais.

Por fim, os resultados desse estudo vão de encontro com os resultados encontrados por Jacobs, Souza & Zanini (2016) que em seu estudo utilizaram modelos baseados em redes neurais, assim como a modelagem ARIMA em conjuntos com diversos métodos de combinação para prever a demanda agregada de uma indústria de lácteos. Os resultados encontrados pelos autores demonstraram, que para o caso estudado, o método *Inverse Mean Squared* obteve a melhor acurácia, apresentado menores erros para as estatísticas de erro absoluto e relativo, assim como no presente estudo.

Quanto a aplicabilidade, as abordagens para combinação de previsão em geral são de simples utilização, não exigindo grandes custos computacionais para sua implementação e ainda assim apresentando relativa facilidade de compreensão com ganhos significativos em termos de acurácia. Dessa forma, a combinação de previsões é uma abordagem atrativa em detrimento a modelos de previsão menos sofisticados.

6 Considerações finais

Os métodos de previsão são importantes ferramentas no processo de tomada de decisão nas organizações, gerando subsídios que auxiliam gestores no gerenciamento de processos, possibilitando decisões mais assertivas baseadas em modelos estatísticos com embasamento científico.

A modelagem ARIMA foi capaz de capturar o comportamento temporal da variável estudada, e os métodos de combinação proporcionaram um aumento da precisão final das previsões, possibilitando combinar três modelos selecionados e adequadamente ajustados a série de dados. Dessa forma, os resultados capacitam o modelo combi-

nado para ser utilizado pelos gestores da empresa no dimensionamento de produção e no planejamento de rotinas de manutenção, servindo como uma ferramenta de apoio à tomada de decisão.

Como sugestão para futuras pesquisas recomenda-se a utilização de outras técnicas para previsão, como os modelos de redes neurais artificiais, assim como modelos de suavização exponencial. Além disso a utilização de outros métodos para a combinação de previsões pode trazer resultados acurados, contribuindo como aumento de precisão dos modelos.

Referências

- Akaike, H. (1973). Information theory and an extension of the maximum likelihood principle. *2nd International Symposium on Information Theory*.
- Amendola, A., & Storti, G. (2008). A GMM procedure for combining volatility forecasts. *Computational Statistics & Data Analysis*, 52, 3047-3060.
- Aye, G. C., Balcilar, M., Gupta, R., & Majumdar, A. (2015). Forecasting Aggregate Retail Sales: The Case of South Africa. *International Journal of Production Economics*, 160, 66-79.
- Box, G., & Jenkins, G. (1970). *Time Series Analysis, forecast and control*. San Francisco: Holden-Day.
- Bueno, R. L. (2008). *Econometrica de séries temporais*. São Paulo: Cengage Learning.
- Clemem, R. T. (1989). Combining forecasts: A review and annotated bibliography. *International Journal of Forecasting*, 5, 559-583.
- CODEBA. (2017). *Sistema portuária Brasileiro*. Acesso em 17 de dezembro de 2017, disponível em Companhia das Docas do Estado da Bahia - Autoridade portuária: <http://www.codeba.com.br/eficiente/sites/portalcodoba/pt-br/site.php?secao=sistemaportuariabrasileiro>
- Costantini, M., & Pappalardo, C. (2010). A hierarchical procedure for the combination of forecasts. *International Journal of Forecasting*, 26, 725-743.
- Dickey, D. A. (1984). Power of Unit Root Tests. *Proceedings of business and economic statistics Sections*, 74, 427-431.
- Firmino, P. R., Mattos Neto, P. S., & Ferreira, T. A. (2014). Correcting and combining time series forecasters. *Neural Networks*, 50, 1-11.

- Gil, A. C. (2010). *Como elaborar projetos de pesquisa* (Vol. 5). São Paulo: Atlas.
- Gonzalez, M. M., & Trujillo, L. (2008). Reforms and infrastructure efficiency in Spain's container ports. *Transportation Research Part A: Policy and Practice*, 42(1), 243-257.
- Granger, C., & Ramanathan, R. (1984). Improved Methods of Combining Forecasts. *Journal of Forecasting*, 3, 197-204.
- Gujarati, D. N., & Porter, D. C. (2011). *Econometria Básica* (5 ed.). Porto Alegre: AMGH.
- Hibon, M., & Evgeniou, T. (2005). To combine or not to combine: selecting among forecasts and their combinations. *International Journal of Forecasting*, 21, 15-24.
- Jacobs, W., Souza, A. M., & Zanini, R. R. (2016). Combination of Box-Jenkins and MLP/RNA models for forecasting. *IEEE Latin America Transactions*, 14, 1870-1878.
- Jeong, D. I., & Kim, Y.-O. (2009). Combining single-value streamflow forecasts – A review and guidelines for selecting techniques. *Journal of Hydrology*, 377, 284-299.
- Klimberg, R. K., Lawrence, K. D., & Lal, T. (2009). Using data envelopment analysis (DEA) to forecast bank performance. In *Advances in business and management forecasting* (pp. 53-61). Emerald Group Publishing Limited.
- Kwiatkowski, D., Phillips, P., Schmidt, P., & Shin, Y. (1992). Testing the null hypothesis of stationarity against the alternative of a unit root. *Journal of Econometrics*, 54, 159-178.
- Makridakis, S., & Hibon, M. (2000). The M3-Competition: results, conclusions and implications. *International Journal of Forecasting*, 16, 451-476.
- Martins, P. G., & Laugeni, F. P. (2009). *Administração da Produção* (2 ed.). São Paulo: Saraiva.
- Martins, V. M., & Werner, L. (2014). Comparação de previsões individuais e suas combinações: um estudo com séries industriais. *Production*, v. 24, n. 3, 618-627.
- Montgomery, D. C., Jennings, C. L., & Kulahci, M. (2015). *Introduction to Time Series Analysis and Forecasting*. New Jersey: John Wiley & Sons.
- Moreira, D. A. (2014). *Administração da Produção e Operações* (2 ed.). São Paulo: Cengage Learning.
- Morettin, P. A. (2011). *Econometria financeira - Um curso em séries temporais financeiras* (2 ed.). São Paulo: Blucher.
- Pang, G., & Gekba, B. (2017). Forecasting container throughput using aggregate or terminal-specific data? The case of Tanjung Priok Port, Indonesia. *International Journal of Production Research*, 55, 2454-2469.
- Prodanov, C. C., & Freitas, E. c. (2013). *Metodologia do trabalho científico: métodos e técnicas da pesquisa e do trabalho acadêmico* (2 ed.). Novo Hamburgo: Universidade Feevale.
- Rausser, G. C., & Oliveira, R. A. (1976). An econometric analysis of wilderness area use. *Journal of the American Statistical Association*, 71, 276-285.
- Rosienkiewicz, M., Chlebus, E., & Detyna, J. (2017). A Hybrid Spares Demand Forecasting Method Dedicated to Mining Industry. *Applied Mathematical Modelling*, 49, 87-107.
- Slack, N., Brandon-Jones, A., & Johnston, R. (2015). *Administração da Produção*. São Paulo: Atlas.
- Souza, F. (2016). *Modelos de Previsão: aplicações à energia elétrica - ARIMA - ARCH - AI e ACP*. Curitiba: Appris.
- Souza, A. M., Souza, F. M., Ferreira, N., & Menezes, R. (2011). Electrical energy supply for Rio Grande do Sul, Brazil, using forecast combination of weighted eigenvalues. *GEPROS: Gestão da Produção, Operações e Sistemas*, 6, 23.
- Stock, J. H., & Watson, M. (2001). A comparison of linear and nonlinear univariate models for forecasting. In R. F. Engle, & H. White, *Festschrift in Honour of Clive Granger* (pp. 1-44). Cambridge: Cambridge University Press.
- Stock, J. W., & Watson, M. W. (2004). Combination Forecasts of Output Growth in a Seven-Country Data Set. *Journal of Forecasting*, 23, 405-430.
- Wallis, K. F. (2011). Combining forecasts – forty years later. *Applied Financial Economics*, 21, 33-41.
- Wang, F.-K., & Chang, K.-K. (2010). Adaptive neuro-fuzzy inference system for combined forecasts in a panel manufacturer. *Expert Systems with Applications*, 37, 8119-8126.
- Werner, L. (2005). Um modelo composto para realizar previsão de demanda através da integração da combinação de previsões e do ajuste baseado na opinião. *Tese (Doutorado) - Universidade Federal do Rio Grande do Sul*.



Período	Série Original	ARIMA(1,0,2)	ARIMA(2,0,3)	ARIMA(3,0,1)	Least Squares Weights	Inverse Mean Squared	Média Simples
697	115.508	124.447	125.719	125.762	124.212	125.320	125.309
698	115.508	114.643	115.487	115.798	114.456	115.318	115.309
699	92.927	116.145	117.040	117.065	115.980	116.757	116.750
700	95.852	95.547	95.599	95.173	95.584	95.438	95.440
701	95.852	99.711	99.830	100.319	99.636	99.956	99.953
702	95.852	100.515	100.358	99.084	100.683	99.979	99.986
703	110.095	97.696	97.270	98.479	97.640	97.817	97.815
704	110.095	111.672	111.142	111.621	111.715	111.476	111.478
705	111.933	109.683	108.951	109.374	109.769	109.332	109.336
706	111.933	110.326	109.477	111.223	110.288	110.343	110.342
707	111.933	112.077	111.156	111.116	112.248	111.442	111.450
708	117.146	111.583	110.880	111.264	111.668	111.238	111.242
709	114.444	116.630	116.083	116.257	116.710	116.320	116.323
710	114.444	113.782	113.248	113.228	113.882	113.415	113.420
711	112.287	113.827	113.422	113.790	113.860	113.678	113.680
712	112.287	112.666	112.286	111.888	112.778	112.275	112.280
713	112.089	112.506	112.265	112.234	112.553	112.333	112.335
714	112.089	112.515	112.281	112.056	112.583	112.281	112.284
715	112.089	112.350	112.149	112.086	112.393	112.193	112.195
716	112.089	112.357	112.156	112.085	112.401	112.197	112.199
717	113.817	112.339	112.139	112.084	112.381	112.185	112.187
718	112.771	113.895	113.749	113.707	113.926	113.783	113.784
719	112.771	112.740	112.552	112.522	112.777	112.603	112.605
720	112.771	112.780	112.600	112.700	112.802	112.692	112.693

Apêndice A: Previsões calculadas para os últimos 24 períodos (697 – 720)

Recebido em 2 fev. 2018 / aprovado em 28 mar. 2018

Para referenciar este texto

Agostino, I. R. S., Araujo, R. S., Noronha, M. O., Fonseca, J. I. S., & Souza, A. M. Combinação de previsões aplicada à modelagem de operações: um estudo de caso em um terminal portuário. *Exacta*, São Paulo, v. 17, n. 1, p. 99-110. jan./mar. 2019. Disponível em: <<https://doi.org/10.5585/ExactaEP.v17n1.8327>>