

## Técnicas de Redes Neurais Artificiais e Regressão Linear Múltipla aplicadas na previsão de valores de importação e exportação no setor de bebidas no Brasil

Suellen Teixeira Zavadzki de Pauli, Mariana Kleina, Wagner Hugo Bonat

**Resumo:** As técnicas de previsão têm aplicação em diversas áreas e possibilitam ter-se uma estimativa a respeito de uma variável de interesse. A previsão de séries temporais, mais especificamente, consiste na realização de estimativas a respeito de períodos de tempos futuros por meio de dados históricos. Um dos ganhos de tais previsões está na possibilidade de realizar um bom planejamento, além de evitar gastos desnecessários e possuir um acompanhamento da informação estudada. Neste artigo serão aplicadas as técnicas de redes neurais artificiais e regressão linear múltipla a fim de prever valores de importação e exportação no setor de bebidas no Brasil. Os dados avaliados são mensais e contemplam o período de janeiro de 1996 à maio de 2019. As informações foram separadas em um período de treinamento, o que contempla 75% dos dados e validação, com 25%. Os resultados foram avaliados por meio da Raiz do Erro Quadrático Médio, além de gráficos de acompanhamento.

**Palavras chave:** Previsão, Redes Neurais Artificiais, Regressão Linear múltipla, Bebidas, Exportação.

## Techniques of Artificial Neural Networks and Multiple Linear Regression applied in the prediction of import and export values in the Brazilian beverage sector

**Abstract:** The forecasting techniques have application in several areas and make it possible to have an estimate about a variable of interest. Time series forecasting, more specifically, consists of having historical data and making estimates of future time periods. One of the gains of such predictions is the possibility of having planning, avoiding unnecessary expenses and having a follow up of the studied information. In this paper we will apply artificial neural networks techniques and multiple linear regression in order to predict import and export values in the beverage sector in Brazil. The data evaluated are monthly and cover the period from January 1996 to May 2019. The information was separated into a training period, which represented 75% and validation, with 25%. The results were evaluated using the root mean square error, as well as follow-up graphs.

**Key-words:** Prediction, Artificial Neural Networks, Multiple Linear Regression, Beverage, Export.

### 1. Introdução

A necessidade de previsão de informações ocorre em diversos setores, incluindo negócios, indústria, medicina, política, ciências sociais entre outros mais. Alguns destes problemas envolvem dados dispostos em séries temporais, o que constitui de uma sequência cronológica de observações da variável de estudo. Um ponto importante que faz a previsão de futuros eventos ser tão relevante é pela necessidade de um embasamento para planejamento e processos para tomadas de decisões. Com os modelos é possível resumir os padrões dos dados e estimar a relação entre comportamentos atuais e anteriores e neste processo há um erro da previsão (MONTGOMERY, 2016).

Um exemplo de aplicação com dados históricos em que é possível e interessante utilizar técnicas de estimação é com relação a previsão de demanda. Ter uma estimativa do que será

demandado evita gastos desnecessários, afinal, espera-se que o volume de produção seja mais assertivo. Outro exemplo que pode seguir na mesma linha é o de estimativa de gastos com determinado produto em uma empresa, como o setor de bebidas, por exemplo. Pode-se portanto estimar valores que serão gastos em exportação e importação do setor de bebidas no Brasil.

Um dos principais insumos da produção de bebidas é a água e, tendo no Brasil acesso fácil a este insumo, é possível encontrar indústrias deste setor em todo o território nacional. Além disso, a produção acaba sendo consumida em grande parte pelo mercado interno. No entanto, com o crescimento de renda há mudanças de consumo, o que muitas vezes faz o consumidor buscar produtos diferenciados e, devido a este fator, ocorre o aumento da importação no setor. As últimas pesquisas, por exemplo, indicaram o vinho como a principal bebida estrangeira procurada (CERVIERI et al., 2014). Em meio a este cenário, tem-se a importância de previsões de valores de exportação e importação do setor de bebidas no Brasil.

O presente artigo tem o objetivo de comparar técnicas de Redes Neurais Artificiais e a Regressão Linear Múltipla a fim de avaliar qual tem melhor resultado de previsão para a aplicação em dados de valores em importação e exportação de bebidas no Brasil, o valor previsto foi estimado sempre para o próximo mês considerando os dados históricos de meses anteriores. Para comparação dos modelos foi utilizada a técnica de Raiz do Erro Quadrático Médio.

## 2. Materiais e métodos

Nesta seção serão abordados os conceitos referentes as técnicas aplicadas para previsão, o método utilizado para a comparação dos resultados assim como informações referentes ao objeto de estudo.

### 2.1 Regressão Linear Múltipla

O modelo de Regressão Linear Múltipla (RLM) é uma técnica estatística muito comum em aplicações que relacionam um conjunto de duas ou mais variáveis. Para este modelo, tem-se que  $p$  variáveis explanatórias  $X_1, X_2, \dots, X_p$  são relacionadas com uma variável independente  $Y$  e tem-se que a matrix dos dados é obtida de uma amostra aleatória simples com  $n$  observações  $(x_{i1}, x_{i2}, \dots, x_{ip}, y_i)$ , em que  $i = 1, 2, \dots, n$ . O modelo pode ser expresso conforme a equação (1) (JOBSON, 1991).

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_p x_{ip} + \varepsilon_i, \quad (1)$$

Em que  $\varepsilon_i$  são valores de erro do modelo, os quais são mutuamente independentes e identicamente distribuídos, cuja  $E[\varepsilon_i] = 0$  e  $V[\varepsilon_i] = \sigma_\varepsilon^2$ .

### 2.2 Rede Neural Artificial – Multilayer Perceptron

As Redes Neurais Artificiais (RNA), consideradas estruturas de computação flexíveis e aproximadores universais, na previsão de problemas de séries temporais possuem grau elevado de precisão. A técnica constitui de estruturas auto adaptáveis, não havendo a necessidade de grandes suposições iniciais ao modelo, além de poderem ser utilizadas em séries não-lineares (KHASHEI & BIJARI, 2010).

Os modelos de RNAs têm a capacidade de aprendizagem, que ocorre devido aos processos iterativos dos ajustes. Não há um único algoritmo de RNA, sendo assim, diversos modelos já foram desenvolvidos com variação no processo de aprendizagem e configuração da

arquitetura (BRAGA et al., 2000).

As redes são compostas por neurônios e estes são organizados hierarquicamente em camadas. A primeira camada é denominada camada de entrada, a qual recebe informações das variáveis explicativas. A camada final é a camada de saída e as intermediárias são denominadas camadas ocultas. As conexões no percurso da rede são ponderadas por pesos, então, os neurônios da próxima camada recebem a soma dos sinais da camada anterior ponderada pelos pesos, esta informação passa então por uma função de ativação (COOPER, 2010).

A aprendizagem da rede *Multilayer Perceptron* (MLP) é supervisionada por correção de erro. Nesta rede, a regra de propagação ocorre pelo produto interno das entradas ponderadas pelos pesos adicionada do termo *bias*. A saída da camada anterior é a entrada da próxima (HAYKIN, 2001).

O treinamento da MLP ocorre nas etapas: *feedforward*, *backpropagation* e ajuste dos pesos (FAUSSET, 1994). Na fase de *feedforward*, tem-se  $n$  variáveis de entrada  $x_j$  com  $j = 1, 2 \dots n$  para cada  $k$  neurônios da camada oculta, e os respectivos pesos sinápticos  $w_{kj}$ , o que resulta na equação (2) (HAYKIN, 2001).

$$u_k = \sum_{j=1}^n x_j w_{kj} + b_k \quad (2)$$

Também tem-se o acréscimo do termo *bias*  $b_k$  ao final, o que aumenta ou diminui a entrada líquida da função de ativação, também possibilita o deslocamento no hiperplano da origem no eixo das abscissas (HAYKIN, 2001). O sinal de saída do neurônio  $k$  passa por uma função de ativação  $\varphi(u_k)$ . Uma função comumente utilizada é a sigmóide logística, cuja equação está representada em (3).

$$\varphi(u_k) = \frac{1}{1 + e^{-u_k}} \quad (3)$$

O mesmo processo ocorre para a camada de saída. Após este processo, tem-se uma saída computada pela rede, sendo que a diferença entre esta e o valor real corresponde ao erro. A fim de minimizar o erro os pesos são ajustados e ocorre a fase *backpropagation*, em que atualiza-se os pesos e a informação passa pela rede com os novos pesos, isto ocorre até o número de iterações determinado. A intenção final é definir o conjunto de pesos que assegure que, para cada conjunto de entradas, a saída produzida pela rede será o mais próximo do valor real (COOPER, 2010).

### 2.3 Rede Neural Artificial – Elman

A Elman é uma rede dinâmica recorrente cuja estrutura contém, além das camadas da rede MLP, mais uma camada de contexto, a qual abrange etapas de tempo adjacentes, desta forma introduzindo uma noção de tempo ao modelo (SHEN et al., 2015; LIPTON et al., 2015).

Para esta rede, a camada de entrada possui duas partes, sendo elas os neurônios de entrada, assim como nas MLP e os de contexto, estes últimos possuem uma cópia das ativações dos neurônios anteriores da camada oculta (TOHA & TOKHI, 2009). Tanto os neurônios de entrada quanto os de contexto ativam os neurônios ocultos, então os neurônios ocultos ativam os neurônios de saída, os quais também se alimentam para ativar os neurônios de contexto, desta forma, a camada de contexto permite que a Elman se adapte às características dinâmicas. Assim, os nós de contexto podem ser usados para memorizar as ativações anteriores dos nós

da camada oculta (CHUANJIN et al., 2017; SHEN et al., 2015).

## 2.4 Raiz do Erro Quadrático Médio

A Raiz do Erro Quadrático Médio (REQM) é utilizada para mensuração da precisão da previsão. Na equação (4), tem-se  $y_j$ , que corresponde ao valor real da informação  $j$ , em que,  $j = 1, 2 \dots n$ , sendo  $n$  o tamanho do conjunto de dados e  $\hat{y}_j$  é o valor predito (CHAÂBANE, 2013).

$$RMSE = \sqrt{\frac{\sum_{j=1}^n (y_j - \hat{y}_j)^2}{n}} \quad (4)$$

## 2.5 Metodologia

O estudo foi feito no *software* R (R CORE TEAM, 2019) e, para a importação dos dados, o pacote utilizado foi o *ipeadatar*, o qual possui acesso ao banco de dados do Instituto Brasileiro de Pesquisa Econômica Aplicada (Ipea) (GOMES, 2019). A informação referente a valores de importação e exportação de bebidas é captada da Fundação Centro de Estudos do Comércio Exterior – FUNCEX. Para a aplicação das redes neurais, o pacote utilizado foi RSNNS.

Quanto a periodicidade dos dados, tanto para importação quanto exportação de bebidas, tem-se informação mensal de valores de janeiro de 1996 até maio de 2019, o que constitui um volume de 281 informações mensais para cada uma das duas bases. Na Figura 1 é possível visualizar o comportamento da série histórica, primeiramente de exportação e posteriormente importação. Para a aplicação, os dados foram divididos em um período de treinamento, com 75% da informação e outro de validação, com 25%, a divisão da série ocorre em agosto de 2013.

A previsão sempre foi realizada para um mês à frente, considerando como variáveis de entrada os meses anteriores a este.

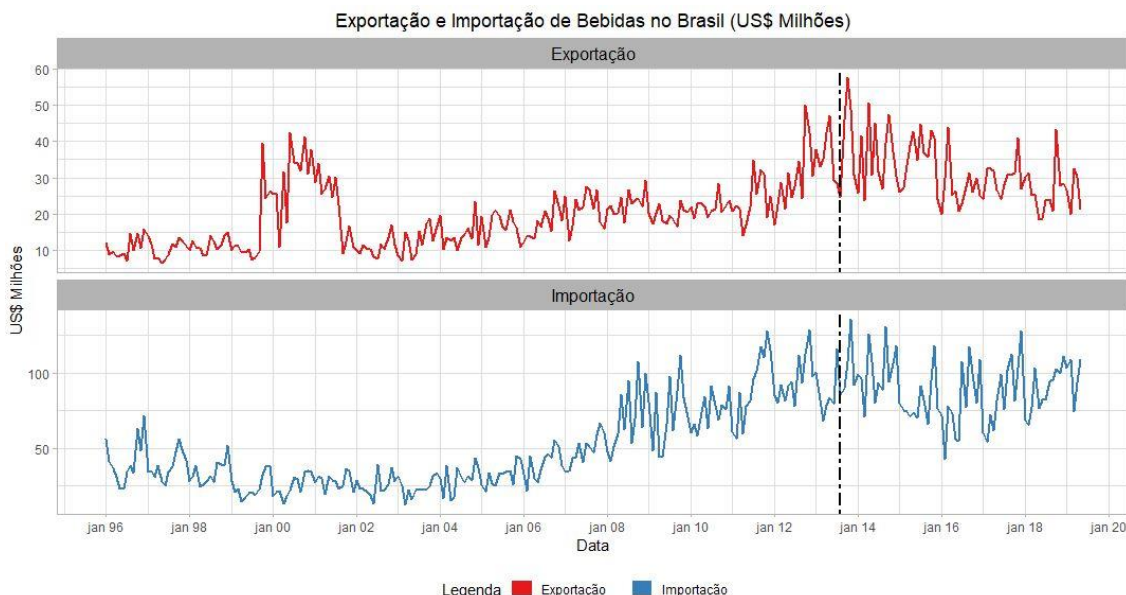


Figura 1 – Série histórica de exportação e importação

Com a Figura 2 é possível verificar a correlação entre as variáveis históricas. Tem-se que a correlação com dados anteriores é maior quanto mais próximos os dias, a maior correlação observada para exportação foi de 0,74 e, para importação, 0,84. Tem-se também que a

correlação entre importação e exportação observada foi de no máximo 0,63.

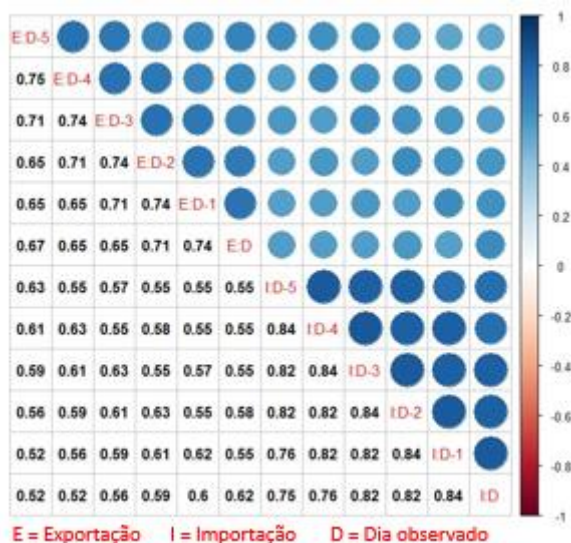


Figura 2 – Gráfico de correlação

### 3. Resultados

Foram testadas duas técnicas de RNAs: MLP e Elman, além da Regressão Linear Múltipla. Cada modelo foi treinado recebendo de 1 a 5 variáveis históricas como entrada do modelo, a fim de avaliar qual possibilidade teria melhor desempenho. No caso das RNAs, o treinamento foi realizado com 50.000 iterações, a estrutura foi testada com 1 a 10 neurônios na camada oculta e também 20 e 30. A função de ativação utilizada para ambas as redes foi sigmóide logística. Os resultados da raiz do erro quadrático médio obtidos para a previsão de exportação constam na Figura 3.

1 VARIÁVEL												
	1 Nó	2 Nós	3 Nós	4 Nós	5 Nós	6 Nós	7 Nós	8 Nós	9 Nós	10 Nós	20 Nós	30 Nós
MLP	8.12	7.97	8.32	8.57	8.76	8.82	8.85	8.86	8.87	8.86	8.89	8.89
Elman	7.89	7.90	8.35	8.41	8.47	8.46	8.45	8.53	8.41	8.50	8.45	8.46
MLR	9.21											
2 VARIÁVEIS												
MLP	8.24	8.21	8.84	8.73	8.16	8.52	8.47	8.47	8.19	8.62	8.81	8.80
Elman	8.09	8.17	8.73	9.69	9.55	10.20	8.59	8.60	14.33	22.04	23.68	8.56
MLR	8.65											
3 VARIÁVEIS												
MLP	8.23	8.32	8.33	8.92	9.09	9.18	9.20	9.08	8.76	9.22	8.80	8.79
Elman	8.09	8.22	8.47	11.06	9.97	8.52	10.18	8.42	17.87	43.49	43.35	54.43
MLR	8.56											
4 VARIÁVEIS												
MLP	8.20	8.28	8.86	8.85	9.35	8.98	8.89	8.91	11.88	11.06	10.41	9.34
Elman	8.23	8.20	9.66	11.44	16.71	30.71	10.87	16.16	25.18	13.59	118.43	43.11
MLR	9.59											
5 VARIÁVEIS												
MLP	8.12	8.25	8.35	9.21	9.91	9.18	9.54	9.87	15.00	9.65	17.10	17.38
Elman	8.26	9.12	8.94	12.74	13.70	13.89	13.66	48.11	24.44	31.79	40.43	46.11
MLR	14.54											

Figura 3 – RMSE para os dados de exportação

Os valores de RMSE obtidos para os modelos de previsão de importação podem ser observados na Figura 4. Ao observar-se as figuras, os menores valores estão em verde. É

possível perceber que modelos com menos variáveis e, no caso das RNAs, com menos nós foram os que os resultados foram mais próximos do real para ambas as bases de dados.

1 VARIÁVEL												
	1 Nó	2 Nós	3 Nós	4 Nós	5 Nós	6 Nós	7 Nós	8 Nós	9 Nós	10 Nós	20 Nós	30 Nós
MLP	19.09	19.38	19.93	19.92	19.88	19.88	20.84	20.88	20.76	20.86	21.07	21.11
Elman	18.95	18.94	19.21	18.56	22.47	19.66	19.58	19.23	19.49	199.12	19.95	21.56
MLR	23.14											
2 VARIÁVEIS												
MLP	19.50	21.10	19.79	19.79	20.53	19.78	19.73	20.63	20.70	19.17	22.07	23.08
Elman	19.01	19.39	55.01	19.47	19.40	27.69	21.90	22.43	26.92	25.38	84.06	55.59
MLR	21.10											
3 VARIÁVEIS												
MLP	19.12	19.09	19.05	20.20	20.11	20.55	20.17	19.88	19.98	23.19	22.71	21.71
Elman	19.00	19.86	20.22	20.61	21.28	26.45	32.55	25.98	305.51	24.08	28.61	58.81
MLR	20.10											
4 VARIÁVEIS												
MLP	19.39	20.09	20.76	21.60	21.02	20.87	21.37	20.87	20.89	20.94	24.77	21.99
Elman	19.44	20.38	21.21	20.79	570.99	27.04	28.52	28.64	179.44	33.80	282.86	30.97
MLR	20.96											
5 VARIÁVEIS												
MLP	19.49	19.26	20.65	20.90	20.96	20.59	24.65	23.44	23.20	23.19	22.84	22.94
Elman	19.46	20.43	21.72	27.83	28.25	21.57	30.82	25.73	48.95	39.87	46.88	53.63
MLR	23.89											

Figura 4 – RMSE para os dados de importação

Após todos os testes, foram selecionados os melhores resultados, ou seja, menores RMSE. Para o modelo MLP, com os dados de exportação a melhor resposta foi obtida com uma variável e dois nós, para os dados de importação 3 variáveis e 3 nós. O modelo Elman treinado com uma variável e um nó foi o melhor resultado para exportação, para importação foi uma variável e quatro nós. O modelo regressão linear múltipla foi melhor com três variáveis tanto com dados de exportação quanto de importação. Entre os três modelos, a rede Elman foi melhor para as duas bases de dados.

Na Figura 5, é possível observar os erros, em US\$ Milhões, dos 3 modelos nas 2 bases de dados. Não há grandes indícios de melhor modelo com esta informação. Também pode-se notar uma certa dispersão nos dados.

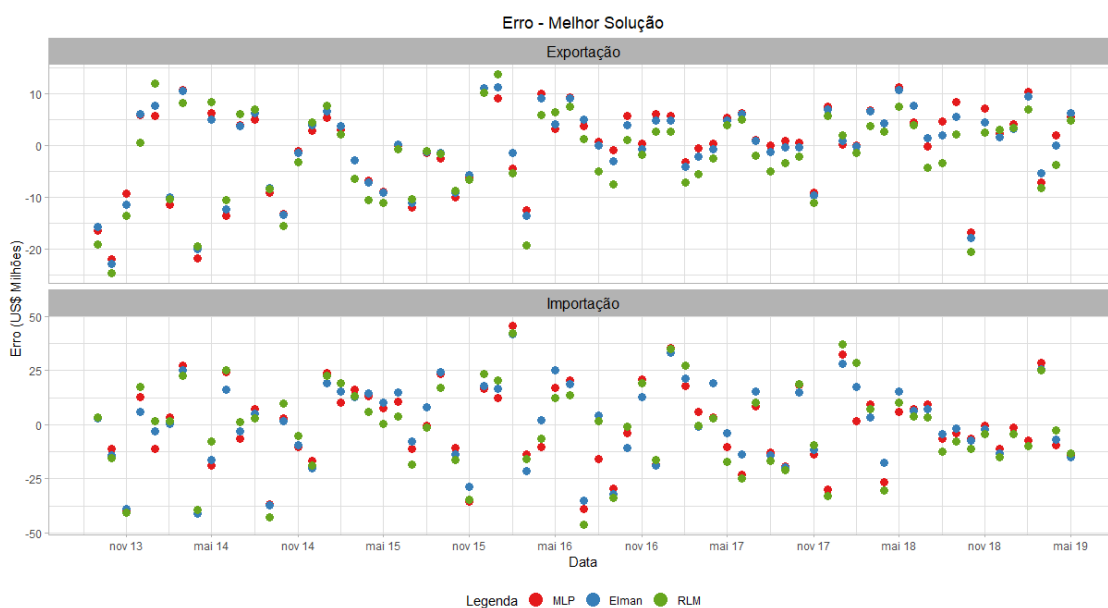


Figura 5 – Erro dos modelos na validação

Ao observar-se a distribuição de frequência dos erros na Figura 6, pode-se notar que não há simetria. Também é possível perceber que para exportação há uma tendência para a direita, ou seja, tem-se uma maior frequência de superestimação dos valores. O erro para esta base é inferior ao de importação devido a amplitude dos erros.

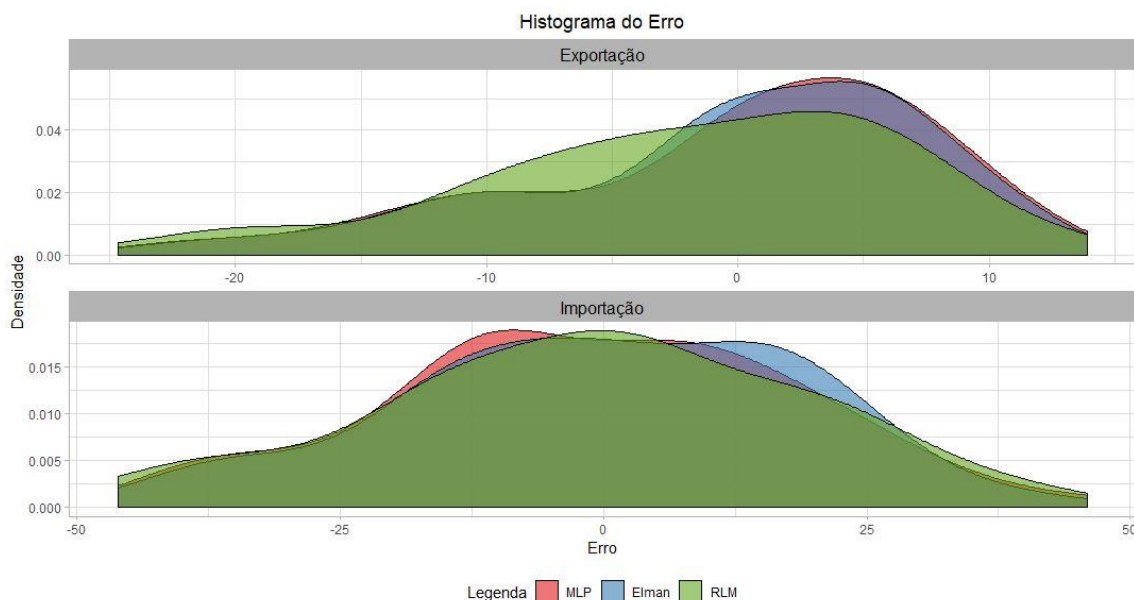


Figura 6 – Distribuição do erro dos modelos no período de validação

Na Figura 7 é possível visualizar a previsão dos modelos cujos RMSE foram os menores, também tem-se o valor real no período de validação. Para ambas as previsões, ~~mesmo~~ as melhores soluções não estão próximas do realizado, isto ocorre devido às grandes flutuações de valores, sendo assim os modelos não conseguiram acompanhar comportamentos bruscos.

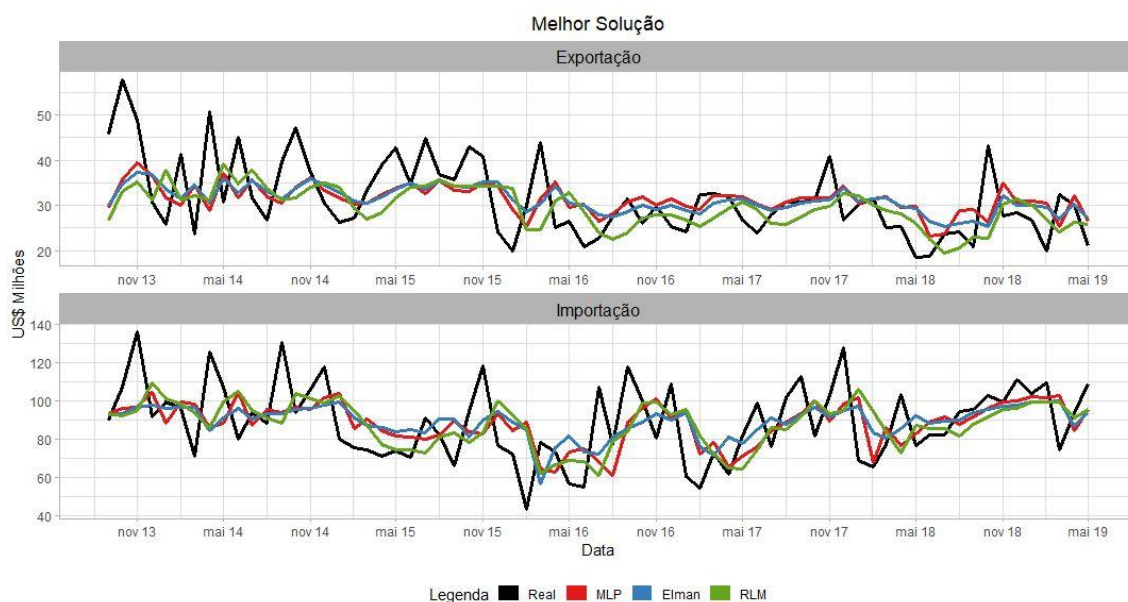


Figura 7 – Previsão versus real no período de validação

#### 4. Discussão

A previsão de valores de exportação e importação no setor de bebidas no Brasil é um exemplo de aplicação para estimação monetária em meses futuros e que tem grande relevância para o planejamento de gastos, por exemplo. Há diversas técnicas que podem ser aplicadas para a previsão de séries temporais e, para este artigo foram escolhidas as técnicas de redes neurais artificiais e regressão linear múltipla.

Os modelos foram aplicados com diferentes configurações e ao final, foi selecionado aquele com melhor resultado de RMSE. Entre todas as tentativas, a rede Elman foi a técnica com resultados melhores, porém, ainda assim os valores previstos foram distantes do real.

O comportamento dos dados não foi completamente absorvido pelos modelos, sugere-se então para estudos futuros a aplicação com outros modelos de inteligência computacional e também com diferentes variáveis que tenham relação à mudança de comportamento do consumidor de bebidas para produtos estrangeiros.

#### AGRADECIMENTOS

Os autores agradecem aos Professores do Programa de Pós-Graduação em Engenharia de Produção da Universidade Federal do Paraná (PPGEP) e aos membros do Grupo de Pesquisa Operacional (GPO), pelo apoio, empenho e dedicação fornecidos em todo o desenvolvimento da pesquisa. Também à CAPES pela bolsa de estudo que vêm sendo concedida ao primeiro autor.

#### Referências

- BRAGA, A. P.; CARVALHO, A. P. L.; LUDEMIR, T. B. **Redes Neurais Artificiais: Teoria e Aplicações**. 2. ed. Rio de Janeiro. Ed. LTC, 2000.
- CERVIERI J. O.; TEIXEIRA J. J. R.; GALINARI, R. RAWEL E. L.; SILVEIRA C. T. J. O setor de bebidas no Brasil. Banco Nacional de Desenvolvimento Econômico e Social. Rio de Janeiro, 2014.
- CHAÂBANE, A. A novel auto-regressive fractionally integrated moving average–least-squares support vector machine model for electricity spot prices prediction. **Journal of Applied Statistics**, v.41, p. 635-651, 2013.
- CHUANJIN, Y.; YONGLE, L.; MINGJIN, Z. An improved Wavelet Transform using Singular Spectrum Analysis for wind speed forecasting based on Elman Neural Network. **Energy Conversion and Management**, v.148, p. 895-904, 2017.
- COOPER, J. C. B. Artificial neural networks versus multivariate statistics: An application from economics. **Journal of Applied Statistics** v.26, pp. 909-921, 2010.
- FAUSSET, L. **Fundamentals of Neural Network: Architecture, Algorithm, and Application**. New Jersey: Prentice Hall, 1994.
- GOMES, L. E. S. **Package ‘ipeadatar’**. Disponível em: <<https://cran.r-project.org/web/packages/ipeadatar/ipeadatar.pdf>> Acesso em: 23 jul. 2019.
- HAYKIN, S. **Redes Neurais: princípios e práticas**. Tradução de Paulo Martins Engel. Porto Alegre. Ed. Bookman, 2001.
- JOBSON, J.D. **Applied Multivariate Data Analysis**. Volume II: Categorical and Multivariate Methods. Springer Texts in Statistics. New York. p. 219-398, 1991.



- KHASHEI, M.; BIJARI, M. An artificial neural network (p, d, q) model for timeseries forecasting. **Expert Systems with Applications**, v.37, p. 479-489, 2010.
- LIPTON, Z. C.; BERKOWITZ, J.; ELKAN, C. A Critical Review of Recurrent Neural Networks for Sequence Learning, Cornell university, arXiv:1506.00019, 2015.
- MONTGOMERY, D. C.; JENNINGS, C. L.; KULAHCI, M. **Introduction to Time Series Analysis and Forecasting**. Ed. Wiley, 2016.
- R CORE TEAM (2019). R: **A language and environment for statistical computing**. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.
- SHEN, C.; SONG, R.; LI, J.; ZHANG, X.; TANG, J.; SHI, Y.; LIU, J.; CAO, H. Temperature drift modeling of MEMS gyroscope based on genetic - Elman neural network. **Mechanical Systems and Signal Processing** v.72–73, pp. 897-905, 2015.
- TOHA, S. F.; TOKHI, M. O. MLP and Elman Recurrent Neural Network Modelling for the TRMS. **7th IEEE International Conference on Cybernetic Intelligent Systems**, 2008.