

Universidade Federal do Paraná

Amábile Galdino Leandro

Cynthia dos Santos Vidal

**Comparação de métodos estatísticos para a  
modelagem do mix de marketing**

**Curitiba**

**2022**

Amábile Galdino Leandro  
Cynthia dos Santos Vidal

## **Comparação de métodos estatísticos para a modelagem do mix de marketing**

Trabalho de Conclusão de Curso apresentado à disciplina Laboratório B do Curso de Graduação em Estatística da Universidade Federal do Paraná, como exigência parcial para obtenção do grau de Bacharel em Estatística.

Orientador: Prof. Dr. Cesar Augusto Taconeli

Curitiba  
2022

# Agradecimentos

Agradecemos aos nossos amigos Eduardo Cock, Flávio Freire, Juliana Botega, Rafael Teodoro que nos acompanharam ao longo desses anos de graduação, nos apoiaram e dividiram momentos bons (e ruins) durante este período.

Aos nossos pais, pelo apoio e por nunca terem medido esforços para nos proporcionar um ensino de qualidade.

Ao nosso orientador, Professor Doutor Cesar Augusto Taconeli, por ter desempenhado tal função com dedicação, paciência e compreensão.

Ao Professor Doutor Fernando Lucambio Pérez, pela disponibilidade em aceitar o convite para participar da banca deste trabalho.

Por fim, um agradecimento à parceria que nós, Amábile e Cynthia, desenvolvemos mutuamente. Não somente ao longo deste trabalho, mas sim, durante toda a graduação.

*"O jeito é deletar – ou apagar – e começar de novo.  
Sempre melhora. Sempre".  
(Moacir Scliar)*

# Resumo

Diante do problema da alocação de investimento em publicidade, os gestores de marketing utilizam de métodos estatísticos para compreender os impactos das ações de mídia na receita das empresas. A modelagem do mix de marketing engloba uma variedade de métodos estatísticos que são aplicados aos dados de publicidade para mensurar o impacto de táticas de marketing nas vendas de uma empresa e realizar previsões. O método estatístico mais popular é a regressão linear múltipla. Entretanto esta abordagem possui algumas desvantagens, sendo as principais: linearidade é questionável no dia a dia das ações de marketing - por exemplo, o aumento do investimento em determinado canal de mídia e a receita - e que os conjuntos de dados de marketing possuem estrutura temporal que não é captada por esse tipo de modelo. Estudos recentes apontam métodos alternativos para proporcionar uma compreensão mais ampla do impacto de covariáveis na receita da empresa. O presente trabalho teve como objetivo comparar os ajustes dos modelos ARIMA e Redes Neurais Recorrentes em um conjunto de dados de marketing. Ao realizar a comparação entre os valores observados e as previsões de cada um dos métodos, foi identificado que o modelo ARIMA obteve um desempenho mais satisfatório, quando comparado às Redes Neurais Recorrentes.

**Palavras-chave:** Inferência Estatística. Redes Neurais Recorrentes. Séries Temporais. ARIMA.

# Lista de ilustrações

Figura 1 – Exemplo de gráficos de ACF e PACF . . . . .	12
Figura 2 – Imagem do modelo de Mcculloch e Pitts. Fonte:(BARRETO, 2002) . .	13
Figura 3 – Imagem de neurônio artificial, Fonte:(BARRETO, 2002) . . . . .	14
Figura 4 – Representação de rede neural artificial de camada única. Fonte: (GUTI- ERREZ, 2003) . . . . .	15
Figura 5 – Representação de neural artificial de múltiplas camadas. Fonte:(GUTIERREZ, 2003) . . . . .	16
Figura 6 – Representação de rede neural recorrente. Fonte:(GUTIERREZ, 2003) .	16
Figura 7 – Gráfico de correlação entre as variáveis . . . . .	20
Figura 8 – Gráfico de comparação dos dados originais e agregados . . . . .	21
Figura 9 – Gráfico de decomposição da variável demanda . . . . .	21
Figura 10 – Gráfico de comparação da série de demanda original e estacionária . .	22
Figura 11 – Análise de resíduos para o modelo ARIMA(3,0,3) . . . . .	25
Figura 12 – Gráfico com os valores reais e previstos do modelo ARIMA . . . . .	26
Figura 13 – Gráfico com os valores reais e previstos do modelo RNR . . . . .	27
Figura 14 – Gráfico com os valores reais e previstos do modelo RNR e ARIMA . .	29

# Lista de tabelas

Tabela 1 – Lista de configurações testadas para o modelo ARIMA . . . . .	23
Tabela 2 – Tabela com os resultados do AIC dos modelos ajustados . . . . .	25
Tabela 3 – Valores observados e previstos para demanda . . . . .	26
Tabela 4 – Tabela de valores previstos para RNR . . . . .	27
Tabela 5 – Comparação dos resultados do RNR e ARIMA . . . . .	28

# Sumário

	<b>Lista de ilustrações</b>	<b>5</b>
	<b>Lista de tabelas</b>	<b>6</b>
<b>1</b>	<b>INTRODUÇÃO</b>	<b>8</b>
<b>2</b>	<b>MATERIAL E MÉTODOS</b>	<b>10</b>
<b>2.1</b>	<b>Metodologia</b>	<b>10</b>
2.1.1	Séries temporais	10
2.1.1.1	Estacionariedade	10
2.1.1.2	ARIMA	11
2.1.1.3	Modelos de regressão dinâmica	12
2.1.2	Redes Neurais	13
2.1.3	Métricas para comparação entre os modelos	17
<b>2.2</b>	<b>Base de dados</b>	<b>18</b>
<b>2.3</b>	<b>Análise e Preparação dos dados</b>	<b>19</b>
2.3.1	Preparação do conjunto de dados	19
2.3.2	Séries temporais	22
2.3.3	Redes Neurais Recorrentes	23
<b>3</b>	<b>RESULTADOS E DISCUSSÃO</b>	<b>25</b>
<b>3.1</b>	<b>Modelos</b>	<b>25</b>
<b>3.2</b>	<b>Redes Neurais Recorrentes</b>	<b>27</b>
<b>3.3</b>	<b>Comparação dos resultados</b>	<b>28</b>
<b>4</b>	<b>CONCLUSÃO</b>	<b>30</b>
	<b>REFERÊNCIAS</b>	<b>32</b>

# 1 Introdução

Um problema comum que os gerentes de *marketing* enfrentam é a alocação ótima dos investimentos em publicidade. Apesar da grande atenção que a publicidade online tem recebido nos últimos anos - principalmente por conta das mídias sociais e novas formas de campanhas na internet - estudos sobre o impacto do investimento das ações de mídia na receita das empresas não são tão recentes assim.

O conceito de *Marketing Mix* data da década de 1950, quando Neil Borden, descreveu o gerente de *marketing* como “misturador de ingredientes”, que cria uma combinação de estratégias de *marketing* para viabilizar uma empresa. Entretanto, esse conceito foi popularizado por Jerome McCarthy, que definiu o *mix* de *marketing* como as ‘variáveis que um gerente pode controlar para influenciar as vendas ou participação do volume de vendas que essa empresa possui no setor (MARSCHNER; MCCARTHY, 1972).

A modelagem do *mix* de *marketing*, do termo em inglês, *marketing mix modeling* (MMM) - engloba uma variedade de métodos estatísticos que são empregados na modelagem de dados históricos de vendas em função de variáveis de *marketing* e variáveis de controle, por exemplo condição climática, período do ano ou mês e competição de mercado (CHAN; PERRY, 2017). A MMM permite às empresas avaliar os resultados de seus esforços nas diferentes mídias, como as redes sociais, anúncios na internet, televisão, rádio, jornais, entre outras. Sendo assim, o objetivo do MMM é mensurar o impacto das variáveis de *mix* de *marketing* sobre as métricas de negócio, para proporcionar uma compreensão mais ampla das variáveis que afetam os resultados, para direcionar estratégias para a distribuição e economia de recursos e, assim, maximizar o lucro.

A definição do modelo capaz de refletir adequadamente relação entre a variável resposta e as covariáveis é uma etapa fundamental da MMM. A técnica de modelagem comumente utilizada é regressão linear múltipla, uma vez que o ajuste do modelo, a estimação e interpretação dos parâmetros ocorrem de forma acessível (WOLFE; CROTTS, 2011; CHAN; PERRY, 2017; MENDEZ-SUAREZ, 2021), além de estar disponível em um grande número de *softwares* estatísticos.

Apesar da popularidade, a regressão linear múltipla possui algumas desvantagens: as previsões ficam limitadas a um intervalo de tempo delimitado pelos dados usados no ajuste do modelo, a suposição de linearidade é questionável no dia a dia das ações de marketing (WIGREN, 2019; MHITAREAN, 2017). Uma outra limitação deve-se à característica comum dos conjuntos de dados da MMM que são os erros correlacionados, devido à coleta de dados ao longo do tempo.

Estudos mais recentes apontam métodos alternativos que podem ser utilizados na MMM, são exemplos: regressão por séries temporais, modelos bayesianos hierárquicos, modelagem de equações estruturais com estimação por mínimos quadrados parciais (PLS-

SEM), regressão não-linear e XGBoost. (CHORNOUS G., 2021; WIGREN, 2019; MENDEZ-SUAREZ, 2021). Mesmo com o desenvolvimento e divulgação de técnicas estatísticas alternativas, a regressão linear múltipla ainda é utilizada amplamente, mesmo quando há evidências da violação das suposições do modelo. O uso incorreto da modelagem estatística pode levar a estratégias de *marketing* equivocadas. Uma abordagem estatisticamente adequada no ajuste de modelos de *mix* de *marketing* pode auxiliar na tomada de decisões mais assertivas na alocação de investimento em publicidade.

Com base neste contexto, o objetivo geral deste trabalho foi analisar um conjunto de dados de marketing em duas abordagens distintas: séries temporais e redes neurais recorrentes. Foram elencados os seguintes objetivos específicos:

- a) Revisar e aprofundar os conhecimentos sobre os séries temporais e redes neurais recorrentes;
- b) Comparar o desempenho preditivo das abordagens por séries temporais e redes neurais recorrentes;
- c) Elaborar recomendações para a aplicação de tais técnicas aos dados do mix de marketing.

## 2 Material e Métodos

Neste capítulo é apresentado um contexto geral sobre o conjunto de dados e os métodos utilizados para a análise e comparação dos resultados. São abordados tópicos sobre as séries temporais e redes neurais recorrentes.

### 2.1 Metodologia

#### 2.1.1 Séries temporais

Segundo (BOSHNAKOV, 2016) uma série temporal é uma sequência temporal ou cronológica de observações em uma variável de interesse.

A análise de séries temporais é uma ferramenta importante para entender e prever tendências futuras em dados ao longo do tempo. Existem diversas técnicas de modelagem para este tipo de dado. A análise das séries temporais pode ser feita através da sua decomposição e envolve o entendimento de 3 componentes: tendência, sazonalidade e ruído aleatório (MAKRIDAKIS; WHEELWRIGHT; HYNDMAN, 1998)

A **tendência** indica à direção na qual uma série temporal se comporta ao longo de um intervalo de tempo, ou seja, se ela apresenta um comportamento crescente, decrescente ou permanece estável em um período.

A **sazonalidade** em uma série temporal pode ser observada quando os dados apresentam um padrão de repetição conforme um fator sazonal como semana, mês, trimestre ou ano. Esse fenômeno pode ocorrer por diversos motivos como clima, datas comemorativas, características do mercado, etc.

Já os **ruídos aleatórios** são deslocamento esporádico da série temporal e correspondem à parcela de incerteza da demanda.

##### 2.1.1.1 Estacionariedade

Uma suposição comum em séries temporais é a de estacionariedade. Uma série é dita estacionária quando ela se desenvolve no tempo de modo que a escolha de uma origem nos tempos não seja considerada importante (MORETTIN; TOLOI, 2004).

Assim, uma série temporal com tendência, ou com sazonalidade, não será estacionárias pois afetará o valor da série temporal em momentos diferentes. Por outro lado, uma série de ruído branco é estacionária, visto que não importa o intervalo de tempo em que é observada, ela apresenta padrão estável em qualquer ponto no tempo (HYNDMAN; ATHANASOPOULOS, 2018). Essa suposição também pode ser levada em consideração nos casos em que a média, variância e autocorrelação estimadas possam ser usadas como base para ocorrências futuras.

Uma técnica comum para transformar uma série temporal não-estacionária em estacionária é a diferenciação. Essa técnica envolve calcular a diferença entre uma observação e sua observação anterior. É uma abordagem amplamente utilizada para análise de dados de séries temporais quando as observações não são estacionárias.

#### 2.1.1.2 ARIMA

O modelo ARIMA é uma abordagem comumente utilizada para análise de séries temporais. A sigla ARIMA é a abreviação de "*Auto-Regressive Integrated Moving Average model*", o que significa que é um modelo que combina características auto-regressivas, integradas e de médias móveis.

Neste modelo o termo AR( $p$ ) corresponde ao modelo autorregressivo, ou seja, usa a relação dependente entre uma observação e um número  $p$  de observações defasadas. O termo I( $d$ ) refere-se ao uso de diferenciação de observações originais para tornar a série temporal estacionária e o termo MA( $q$ ) correspondente ao modelo de médias móveis, onde  $q$  representa o tamanho da janela de média móvel, configurando a dependência entre uma observação e um erro residual do modelo de média móvel aplicado a observações defasadas. (GUTIERREZ, 2003; HYNDMAN; ATHANASOPOULOS, 2018)

Uma das principais ferramentas para identificar modelos ARIMA é o PACF (*Partial Autocorrelation Function*) e o ACF (*Autocorrelation Function*), que são a função de autocorrelação parcial e a função de autocorrelação respectivamente. Essas funções são medidas de correlação entre valores de séries atuais e anteriores, que indicam quais valores de série anteriores são mais informativos para prever valores futuros.

O ACF mede a correlação entre um valor da série e seus valores anteriores para diferentes intervalos entre observações. Já a PACF mede a correlação entre um valor da série e seus valores anteriores, desconsiderando a correlação entre intervalos mais curtos. E com essas funções é possível identificar características de uma série temporal, a fim de identificar os parâmetros de um modelo ARIMA( $p,d,q$ ).

Por exemplo, quando os valores da ACF começam próximos da unidade e declinam-se de maneira gradual e lentamente à medida que aumenta o valor de  $k$ , há indicações da existência de raiz unitária, sendo portanto, não estacionária. Porém, caso os coeficientes declinem de maneira acelerada à medida que aumenta o valor de  $k$ , então pode-se dizer que a característica de estacionariedade (MATOS, 2000). Um exemplo pode ser visto na Figura 1.

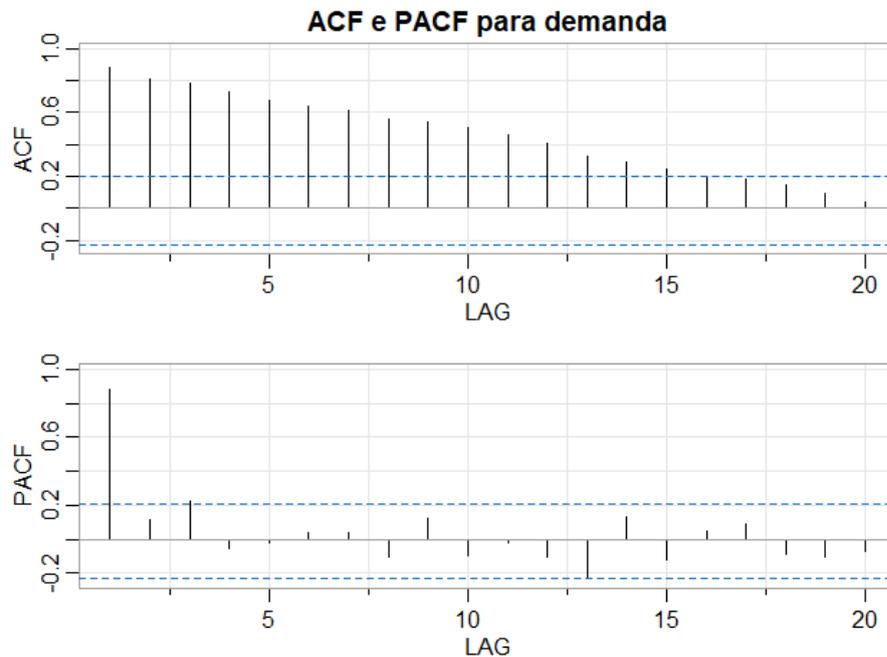


Figura 1 – Exemplo de gráficos de ACF e PACF

### 2.1.1.3 Modelos de regressão dinâmica

O modelo de regressão dinâmica é uma técnica de modelagem estatística que se concentra na análise de séries temporais. A diferença em relação à autoregressão é que o modelo de regressão dinâmica permite que outras variáveis explicativas sejam incluídas, além dos valores passados da variável dependente.

No contexto do trabalho realizado, foi escolhido esta classe de modelo por ser considerado o mais adequado para o conjunto de dados disponível devido à possibilidade de incluir outras variáveis.

Os modelos de regressão dinâmica são apropriados para situações em que há uma correlação entre a variável de interesse e possíveis variáveis exógenas, e quando a estrutura de correlação dos dados indica que os erros não podem ser considerados independentes.(SOUZA; KLIDZIO, 2013)

No processo de modelagem, são estimadas equações separadas. Esses modelos levam em conta a natureza dinâmica dos dados e permitem que as relações entre as variáveis mudem ao longo do tempo, incluindo tanto variáveis exógenas quanto defasagens da variável de interesse. Os erros são modelados por meio de processos auto-regressivos, possibilitando a captura de dependências temporais entre as variáveis e correlações entre os erros.(SOUZA; KLIDZIO, 2013). Como na equação 2.1.

Um modelo de regressão dinâmica pode ter a seguinte estrutura (HYNDMAN; ATHANASOPOULOS, 2018):

$$y'_t = \beta x'_t + \eta' \quad (2.1)$$

onde

$$\eta'_t = \phi_1 \eta'_{t-1} + \varepsilon_t \quad (2.2)$$

é um erro AR(p).

## 2.1.2 Redes Neurais

Embora não exista uma definição formal, as redes neurais artificiais (RNAs) costumam ser ilustradas como uma estrutura similar ao do cérebro, capaz de simular o aprendizado humano, porém em um ambiente computacional. As RNAs consistem em um modo de abordar a solução de problemas usando inteligência artificial. O primeiro modelo de neurônio data de 1943 e foi proposto por McCulloch e Pitts, que interpretaram o funcionamento do neurônio como sendo um circuito binário e é apresentado na Figura 2 (BARRETO, 2002):

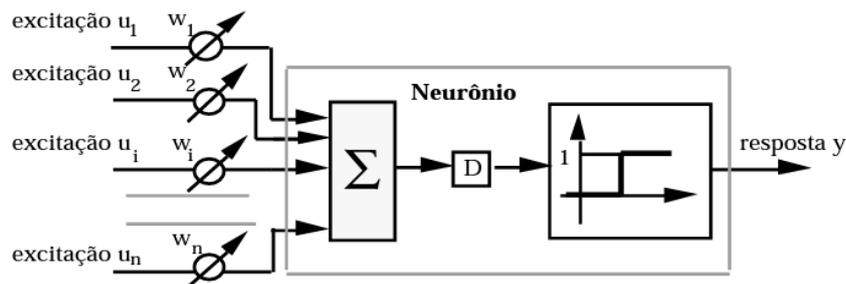


Figura 2 – Imagem do modelo de McCulloch e Pitts. Fonte:(BARRETO, 2002)

Nesse modelo, a entrada do neurônio é binária e as variáveis de entrada são combinadas por uma soma ponderada, produzindo a entrada efetiva do neurônio:

$$entrada \quad efetiva = \sum_1^n w_i u_i \quad (2.3)$$

O resultado na entrada efetiva sofre um retardo  $D$  (algumas vezes este retardo é desprezado tendo-se um neurônio estático) e serve de argumento a uma função chamada de função de transferência neste caso de saída binária 0 1 para dar a resposta do neurônio. Uma generalização do modelo de McCulloch e Pitts é exibido na Figura 3:

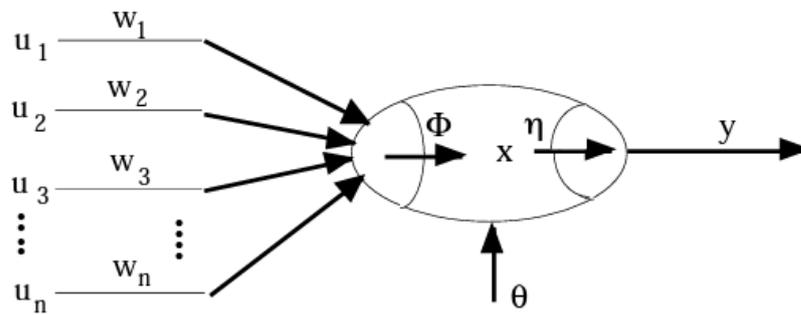


Figura 3 – Imagem de neurônio artificial, Fonte:(BARRETO, 2002)

As entradas  $w_i u_i$  são combinadas usando uma função  $\Phi$ , para produzir um estado de ativação do neurônio que através da função  $\eta$  vai produzir a saída do neurônio. Na literatura, as redes neurais, em geral, podem ser classificadas em três estruturas de rede fundamentalmente diferentes: redes alimentadas adiante com camada única, redes alimentadas diretamente com múltiplas camadas e redes recorrentes (L TAVARES M, 2016).

Redes neurais (*Feedforward*) com camada única: é a forma mais simples de rede neural, em que os neurônios são organizados em uma única camada. A saída de cada um dos neurônios constituem uma saída da rede, isto é, as entradas são enviadas diretamente à saída. A rede é sempre alimentada adiante, pois a camada de nós fonte fornece os sinais de entrada para a camada de saída, sem existir uma realimentação.

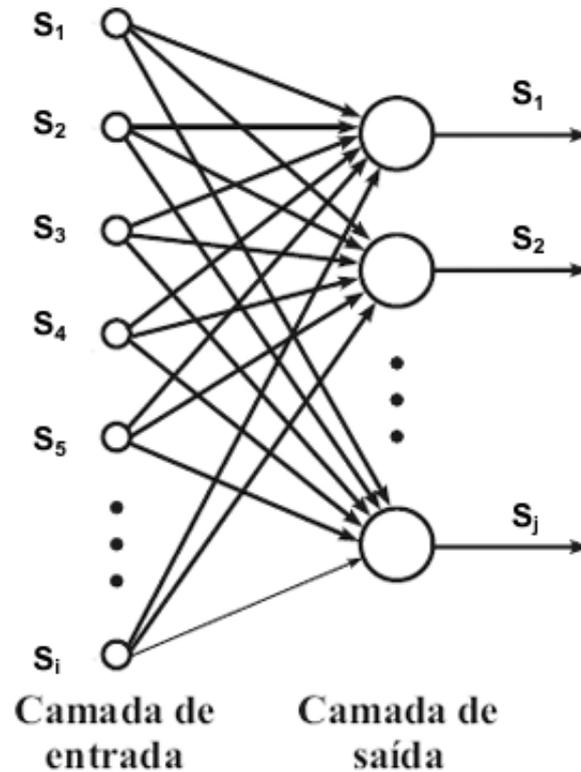


Figura 4 – Representação de rede neural artificial de camada única. Fonte: (GUTIERREZ, 2003)

Já as Redes Neurais *Feedforward* de Múltiplas camadas, são compostas estruturalmente por camadas de neurônios artificiais, divididas em 3 grupos: camada de entrada (*Input layer*), camada escondida (*Hidden layer*) e camada de saída (*Output layer*). Essa é a arquitetura mais freqüentemente encontrada na literatura referente a redes neurais artificiais. (LEMLEY J.; BAZRAFKAN, 2017; LECUN YANN;BENGIO, 2017; DIAS, 2022), como representado Figura 5, logo abaixo. Cada camada é composta por um ou mais neurônios que estão conectados aos neurônios da camada seguinte, assim os sinais se propagam até que cheguem à saída da rede.

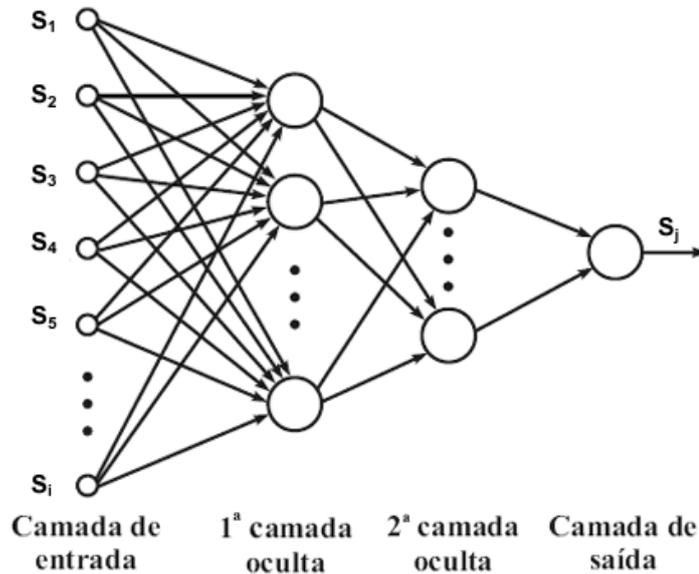


Figura 5 – Representação de neural artificial de múltiplas camadas. Fonte:(GUTIERREZ, 2003)

As Redes Neurais Recorrentes (RNR) diferem das duas anteriores, pois possuem pelo menos um laço de realimentação, o que a torna capaz de armazenar e processar informações no tempo. É essa característica que torna essa arquitetura útil para o processamento de dados sequenciais, como som, dados de séries temporais ou linguagem natural, pois assim o neurônio possui não apenas informações das entradas atuais, mas também informações das entradas anteriores.

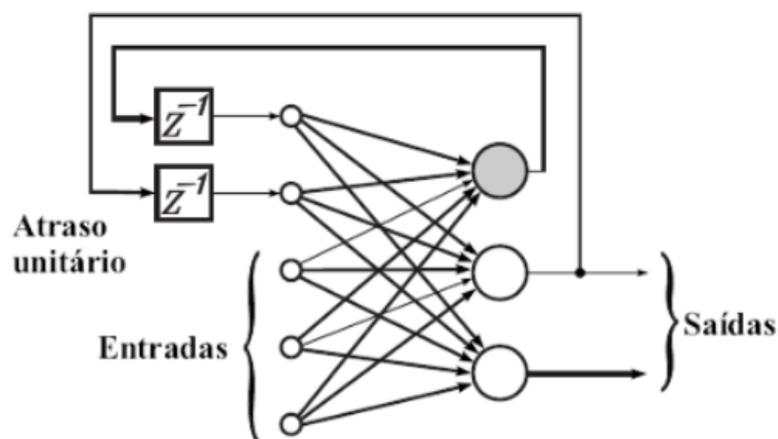


Figura 6 – Representação de rede neural recorrente. Fonte:(GUTIERREZ, 2003)

A Figura 6 acima, representa uma rede deste tipo, em que um neurônio oculto (destacado em cinza) e dois neurônios de saída. Uma arquitetura de redes neurais recorrentes que surgiu na tentativa resolver o problema de dissipação do gradiente é a *Gated Unit*

*Recurrent* (GRU). Esse tipo de arquitetura utiliza *gates* (portões) para determinar, a partir do valor atual de referência e sua memória, qual será o valor de ativação e o valor previsto em sequência (VASCO, 2020), possuindo menos parâmetros, e por isso, geralmente é treinada mais rapidamente ou com um quantidade menor de dados.

Para as redes neurais, há alguns parâmetros que determinam a estrutura do modelo, são utilizados para controlar o processo de aprendizado, os chamados **hiperparâmetros**. Eles são definidos manualmente, antes da etapa de treinamento.

Além do número de camadas, há outros hiperparâmetros associados à estrutura das redes neurais. São eles:

**Número de neurônios:** quantidade unidades ocultas dentro de uma camada. Um número menor de unidades pode causar falta de adequação.

**Tamanho do lote** (ou *Batch Size*, do inglês): refere-se a quantidade de amostras que serão inseridas para treinamento no modelo a cada iteração. Por exemplo, um lote de tamanho 32, corresponde dizer que os 32 primeiros dados da base de treinamento são inseridas na rede e treinadas, e assim sucessivamente, até o final de base de treinamento.

**Otimizador:** algoritmos capazes de realizar ajustes de maximização e minimização nos erros obtidos através de uma função de custo. O otimizador é um dos parâmetros, que mais impactam diretamente no desempenho geral da rede neural. Para RNR, os otimizadores **Adam** (*Adaptive Moment Estimation*), o **SGD** (*Stochastic Gradient Descent*) e o **RMSProp** são mais frequentemente utilizados.

**Tamanho de sequência:** define o comprimento das sequências de dados de entrada no modelo. Como as RNRs aprendem o conjunto sequencial de dados, a rede lembra das sequências que foram inseridas.

**Número de épocas** é o número de vezes que todos os dados de treinamento são mostrados à rede durante esta etapa.

**Dropout** é o parâmetro que reduz o sobreajuste (*overfitting*), e evita que o modelo se ajuste de forma excessiva ao conjunto de dados utilizado para o treinamento.

### 2.1.3 Métricas para comparação entre os modelos

De acordo com a literatura (ARMSTRONG; COLLOPY, 1992) há algumas medidas de acurácia para identificar a precisão de um método de previsão ou como critério de seleção de modelos de previsão. Uma das métricas que mais se destacam é, a raiz do erro quadrático médio (RMSE – *Root Mean Square Error*), que é definida pela seguinte equação:

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (2.4)$$

Em que o  $n$  é o número de observações,  $y(i)$  é a  $i$ -ésima medição e  $y(i)$  é sua previsão correspondente.

O modelo considerado adequado é aquele que apresentar menor erro, independentemente da medida utilizada para os cálculos.

Uma outra métrica utilizada é o Erro percentual médio absoluto (MAPE - *Mean Absolute Percentage Error*), ela que mostra a porcentagem de erro em relação aos valores reais. Nesta equação há o cálculo do valor da média da divisão entre a diferença entre o valor real e o predito sobre o valor real.

$$MAPE = \frac{100\%}{n} \sum_{t=1}^n \left| \frac{A_t - F_t}{A_t} \right| \quad (2.5)$$

O modelo considerado adequado é aquele que apresentar menor erro, independentemente da medida utilizada para os cálculos.

## 2.2 Base de dados

Para comparar as abordagens do modelo ARIMA e RNR no contexto de ajustes de modelos e previsão da demanda, foi utilizado uma base de dados de uma empresa do setor de telecomunicações da China, no período compreendido entre janeiro de 2010 e fevereiro de 2017, com granularidade diária, totalizando 2613 observações e 17 variáveis. Nele, há informações das variáveis de mídia e marketing, como demanda, volume de produtos em estoque, vendas, valor unitário do produto, investimento em canais de mídia como SMS, internet, TV, rádio e jornal e as respectivas audiências. Abaixo, detalhamento das variáveis de mídia e marketing:

**Demand:** Quantidade de consumidores que estão dispostos e aptos a comprar produtos;

**Unit Price:** Preço unitário do produto;

**Supply Data:** Quantidade total de produtos disponível para os consumidores;

**Sales:** Valor das vendas, na unidade monetária da China (Yuan);

**Cost SMS:** Investimento em campanhas via SMS;

**Cost Newspaper:** Investimento em campanhas em jornais;

**Cost Radio:** Investimento em campanhas em rádio;

**Cost TV:** Investimento em campanhas de TV;

**Cost Internet:** Investimento em campanhas na internet;

**GRP Newspaper:** Pontos de audiência bruta em jornal;

**GRP SMS:** Pontos de audiência bruta em SMS;

**GRP Radio:** Pontos de pontos de audiência bruta em rádio;

**GRP Internet:** Pontos de audiência bruta na internet;

**GRP TV:** Pontos de audiência bruta em TV.

O *Gross Rating Points* (GRP) é definido como a quantidade de pontos de audiência bruta, que indica a intensidade com a qual o anunciante está se comunicando com o público utilizando determinada propaganda durante programação da TV aberta. Porém para esse trabalho, as métricas de GRP não foram utilizadas na etapa de modelagem e análise devido incertezas sobre a origem da variáveis.

Além disso, o base de dados traz informações macroeconômicas, que são detalhadas a seguir:

**Consumer Price Index (CPI)** : este índice refere-se à variação média ao longo do tempo nos preços pagos pelos consumidores urbanos por uma cesta de mercado de bens e serviços de consumo.

**Commodity Channel Index (CCI)** : índice relacionado ao nível atual dos preços em relação ao seu valor médio. O CCI é relativamente alto quando os preços estão muito acima da média. O CCI é relativamente baixo quando os preços estão muito abaixo da média.

**Producer Price Index (PPI)** : índice utilizado para medir a variação média ao longo do tempo nos preços de venda recebidos pelos produtores domésticos por sua produção.

Para a análise, tomou-se a variável *Demand* para ser a variável resposta. As demais variáveis foram consideradas variáveis explicativas, com exceção das métricas de GRP que foram descartadas devido a incertezas sobre a suas origens.

Para a realização desse trabalho foi utilizado o *software* estatístico de programação livre R 4.2.2 com a interface gráfica **RStudio** com o objetivo de auxiliar nas etapas de análise e modelagem dos dados .

## 2.3 Análise e Preparação dos dados

### 2.3.1 Preparação do conjunto de dados

Inicialmente foi feita uma análise descritiva dos dados tendo *Demand* como a variável resposta, buscando identificar o comportamento dos dados.

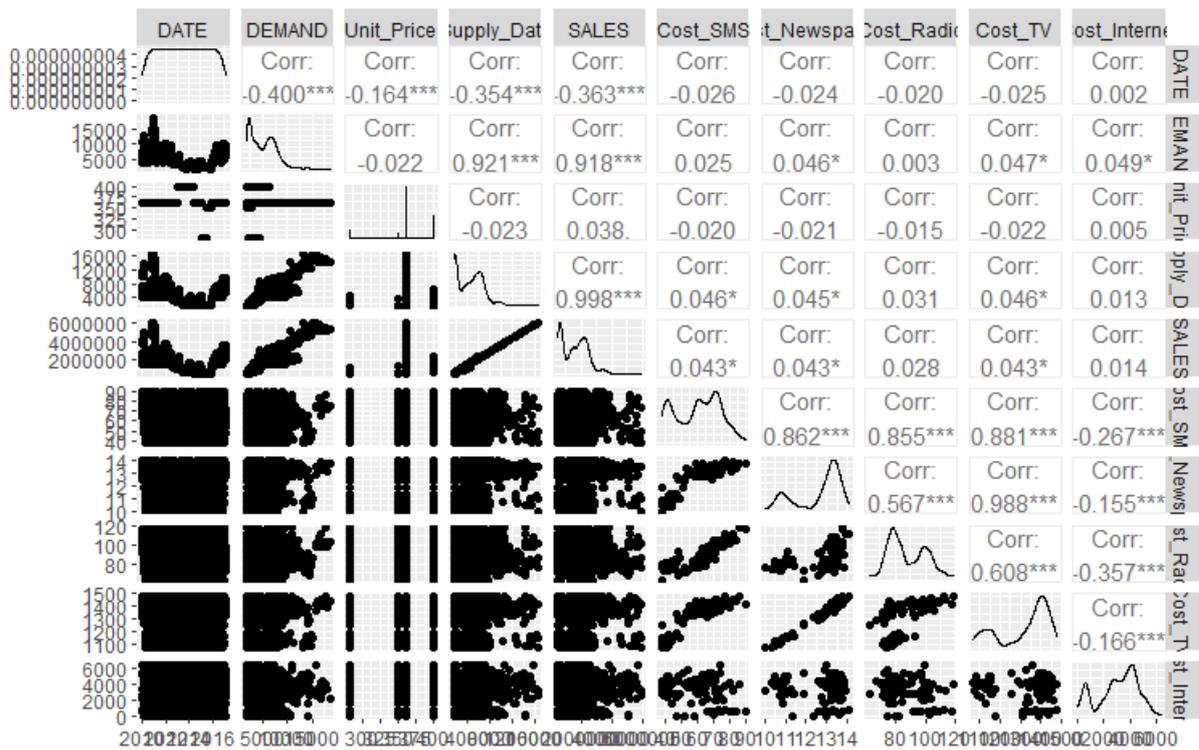


Figura 7 – Gráfico de correlação entre as variáveis

Foi observado que as variáveis *Supply data*, *Sales* e *Demand* estão fortemente correlacionadas entre si, apresentando um comportamento semelhante. Essas variáveis apresentaram valores elevados nos anos de 2010 a 2012, seguidos de uma queda nos resultados de 2013 a 2016. Além disso, as variáveis de custo também apresentaram forte correlação entre si e um comportamento cíclico ao longo do período da base de dados, incluindo o período de queda nos resultados.

Notou-se que as variáveis macroeconômicas apresentam mudanças de valor apenas em intervalos mensais, diferentemente das outras variáveis do conjunto de dados, que exibem mudanças diárias.

Durante a etapa de análise descritiva, com o objetivo de suavizar o comportamento da série temporal e levando em conta que as variáveis macroeconômicas apresentam valores mensais, foi decidido fazer a agregação mensal dos dados diários disponíveis na base de dados. Isso reduziu o número de observações de 2.613 para 86. A Figura 8, logo abaixo, apresenta a suavização da série temporal com a agregação das observações.

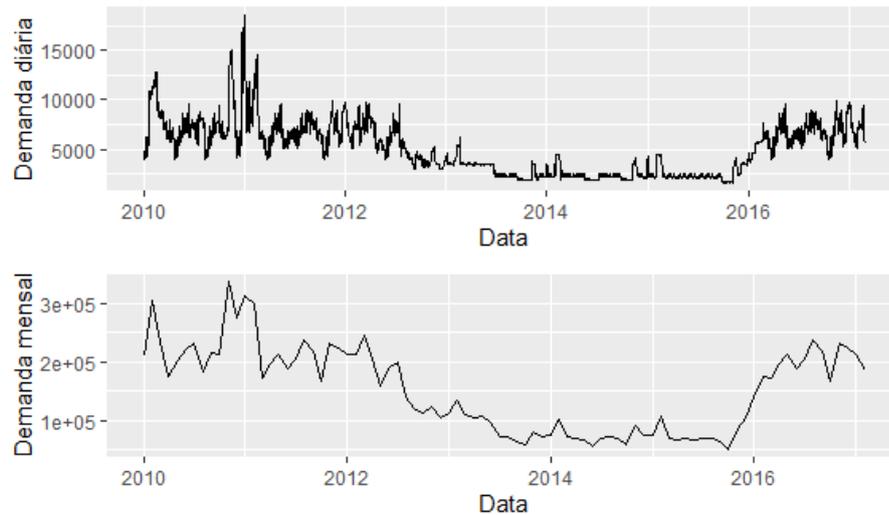


Figura 8 – Gráfico de comparação dos dados originais e agregados

Com a agregação mensal dos dados originais, observa-se que o comportamento da variável *Demand* original foi reproduzido. Entretanto, a série ainda permaneceu um comportamento bastante irregular, que pode ser identificado na Figura 8.

Após a etapa de agregação, foi realizada uma análise exploratória dos dados, em que foram analisados o ACF e PACF com o objetivo de identificar o seu comportamento das variáveis e suas características ao longo do tempo, como estacionariedade ou não-estacionariedade, sazonalidade e a presença de componentes autoregressivos e de média móvel. Ainda, foi realizada uma análise de decomposição das variáveis com o objetivo de compreender melhor os comportamentos de tendência e sazonalidade das variáveis, como pode ser observado na Figura 9.

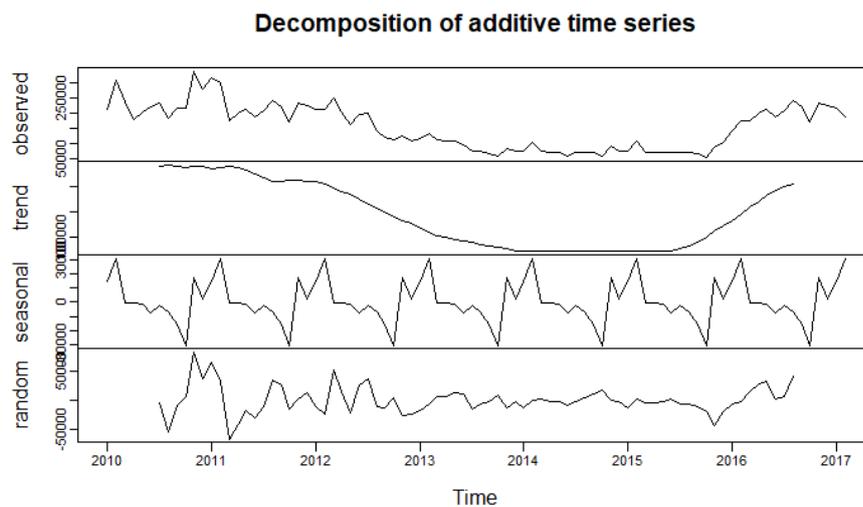


Figura 9 – Gráfico de decomposição da variável demanda

A Figura 9 apresenta a decomposição da variável resposta, no qual se pode observar uma tendência de redução nos valores de demanda de 2013 a 2016. Além disso, há uma

sazonalidade anual nos resultados, onde há valores mais altos no início do ano, seguidos de uma queda ao longo dos meses e um aumento no final do ano, caracterizando um pico.

### 2.3.2 Séries temporais

Concluída a etapa de análise exploratória dos dados, foi aplicado o modelo de regressão dinâmica, mais especificamente ao da equação 2.1. Como descrito em (HYNDMAN; ATHANASOPOULOS, 2018), uma consideração importante ao ajustar um modelo de regressão dinâmica é que todas as variáveis no modelo devem primeiro ser estacionárias. Ao ajustar um modelo, sem que a condição de estacionariedade não for satisfeita, os coeficientes estimados não serão estimativas consistentes. Portanto, primeiramente foram diferenciadas todas as variáveis do nosso modelo, criando assim um “modelo diferenciado”.

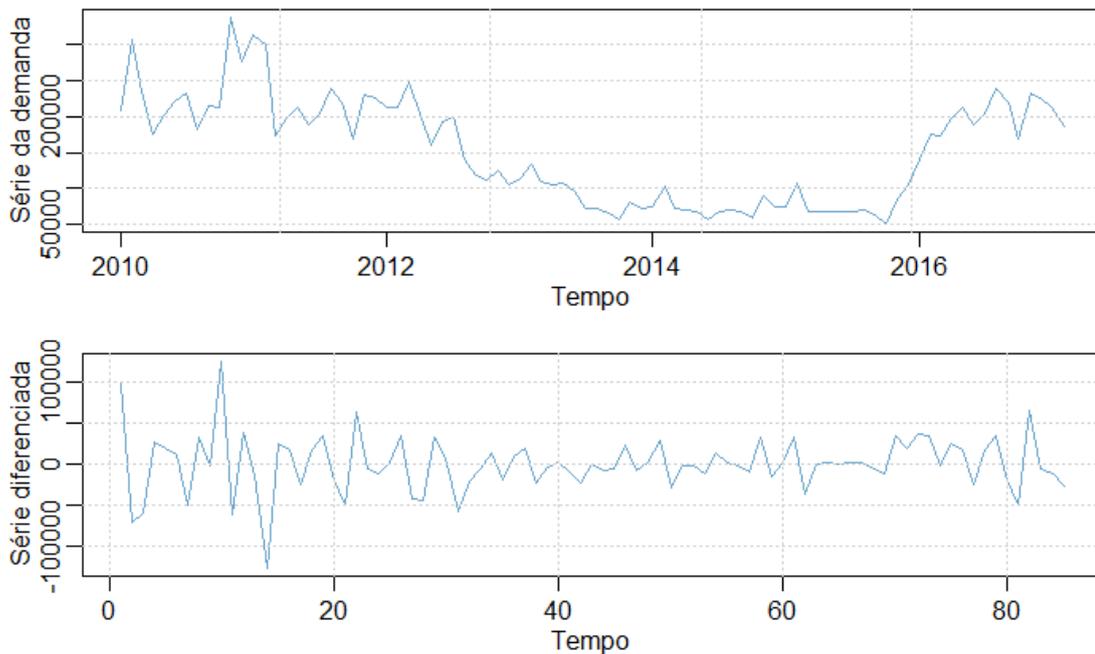


Figura 10 – Gráfico de comparação da série de demanda original e estacionária

Na Figura 10 observa-se o efeito da aplicação da diferença para transformar os dados em uma série temporal estacionária quando aplicado à variável demanda.

Realizada a diferenciação para todas as variáveis, uma nova base de dados foi construída apenas com a série diferenciada e foi feita a separação em treino e teste, na qual as primeiras 77 observações (ou seja, 77 meses) foram alocadas para treino e 8 observações finais ficaram para teste.

Ao fim das etapas de transformação e de separação do conjunto de dados em treino e teste, foram testadas diversas configurações do modelo ARIMA com as covariáveis para verificar qual possuía o melhor ajuste na base de treino. Os critérios utilizados para o ajustes foram baseados nos resultados obtidos na análise exploratória dos dados. As

configurações que foram consideradas estão indicadas na tabela 1.

Modelo	Configuração
1	(1,0,1)
2	(2,0,1)
3	(3,0,1)
4	(3,0,2)
5	(3,0,3)
6	(3,0,0)
7	(3,0,4)

Tabela 1 – Lista de configurações testadas para o modelo ARIMA

Para a seleção do melhor ajuste foi levado em consideração o AIC e a análise de resíduos para verificação de existência de autocorrelação entre os resíduos. Além disso foi utilizado o RMSE para auxiliar na escolha do melhor modelo através da análise da comparação entre os valores reais e os valores previstos para a base de teste dos modelos ajustados.

Por fim, para concluir a análise de séries dos dados, foi realizada a predição para os 8 últimos meses utilizando o modelo selecionado e realizada uma nova transformação das observações para seus valores originais para fins de visualização do ajuste do modelo, visto que a análise foi conduzida com os dados estacionários e para a comparação dos valores preditos com os valores reais a transformação para a escala original se torna mais interessante.

### 2.3.3 Redes Neurais Recorrentes

Para a modelagem utilizando as Redes Neurais Recorrentes, o conjunto de dados com as 86 observações, foi dividido da seguinte maneira: conjunto de treinamento com 64 primeiras observações; conjunto de validação com as observações de índice 65 a 78 e conjunto de teste com últimas 8 observações.

Antes de iniciar o treinamento, foi necessário fazer a padronização (z-score), com o objetivo de manter os dados na mesma escala. Cada observação foi subtraída pela média da série e em seguida dividida pelo desvio padrão da mesma, através da equação:

$$z = \frac{x - \mu}{\sigma} \quad (2.6)$$

Com os dados já tratados, foi ajustado um modelo de rede neural GRU em linguagem R, utilizando a o pacote *torch*. O modelo utilizado possui 4 camadas (1 de entrada, 2 intermediárias e 1 de saída). A camada de entrada possui 12 neurônios (todas as variáveis das bases de dados) e 8 neurônios na camada de saída (a previsão da demanda para 8 meses). Nas duas camadas intermediárias há 32 neurônios. Neste estudo, os seguintes

hiperparâmetros foram definidos: otimizador foi o Adam, o dropout como 0, para o número de épocas tomou-se 15 e para o tamanho do lote tomou-se 32.

## 3 Resultados e Discussão

### 3.1 Modelos

Dentre os diversos modelos analisados o que apresentou o melhor ajuste, com base no menor RMSE, foi o ARIMA(3,0,3), sendo ele composto por 3 parâmetros autoregressivos e 3 parâmetros de média móvel. Na Tabela 2 pode-se observar os modelos testados e seus respectivos valores de AIC, critério o qual foi utilizado para a seleção do melhor modelo.

Modelo	Configuração	AIC
1	(1,0,1)	667,1052
2	(2,0,1)	665,9898
3	(3,0,1)	661,1335
4	(3,0,2)	654,4687
5	(3,0,3)	648,6669
6	(3,0,0)	663,6413
7	(3,0,4)	680,7990

Tabela 2 – Tabela com os resultados do AIC dos modelos ajustados

Este modelo foi validado pela análise de resíduos, em que foi verificado o ACF e PACF, normalidade e estacionariedade do erro. Como pode ser observado na Figura 11 abaixo não houve autocorrelação entre os resíduos.

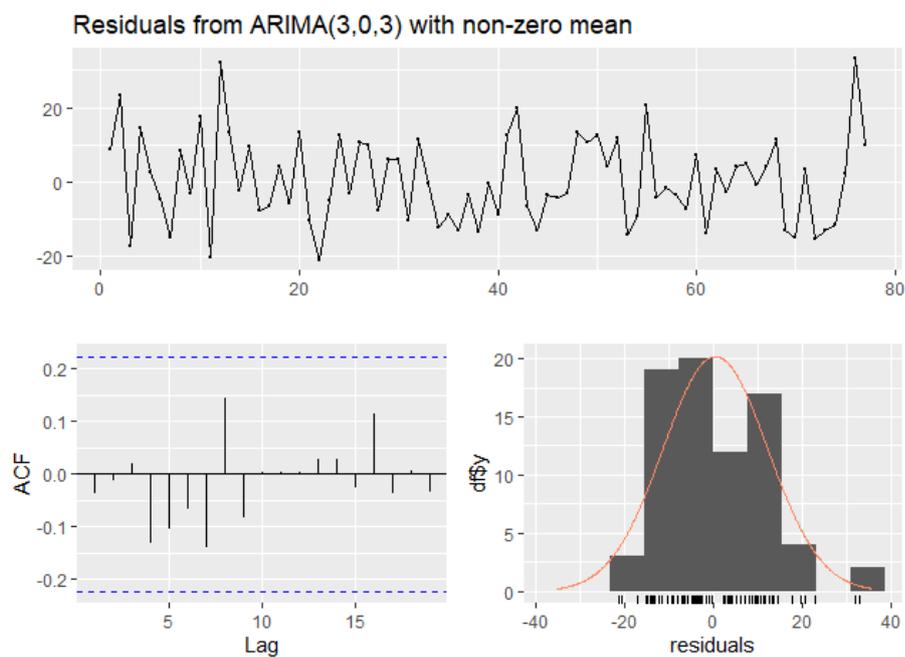


Figura 11 – Análise de resíduos para o modelo ARIMA(3,0,3)

Toda essa etapa foi realizada ainda com os dados estacionários, após a análise de resíduos e seleção do modelo calculou-se a predição para a base de teste e obteve-se então uma predição com os dados estacionários. Como mencionado na sessão de materiais e métodos, foi realizado um cálculo para retornar a escala original dos dados.

A Tabela 3 abaixo apresenta os resultados já com os valores em sua escala original, em que é possível verificar os valores previstos e os valores reais dos últimos 8 meses do conjunto de dados analisado.

Tabela 3 – Valores observados e previstos para demanda

Mês	Valor Real	Valor previsto
jul/2016	204545	190726,1
ago/2016	237869	211738,1
set/2016	218513	205560,2
out/2016	168110	197891,5
nov/2016	231870	234989,6
dez/2016	224974	230247,2
jan/2017	213122	211023,7
fev/2017	185612	183522,6

A Figura 12, por sua vez, traz os dados reais (em azul) e os dados preditos (em vermelho), em que é possível observar o quanto a predição se assemelha aos valores observados.

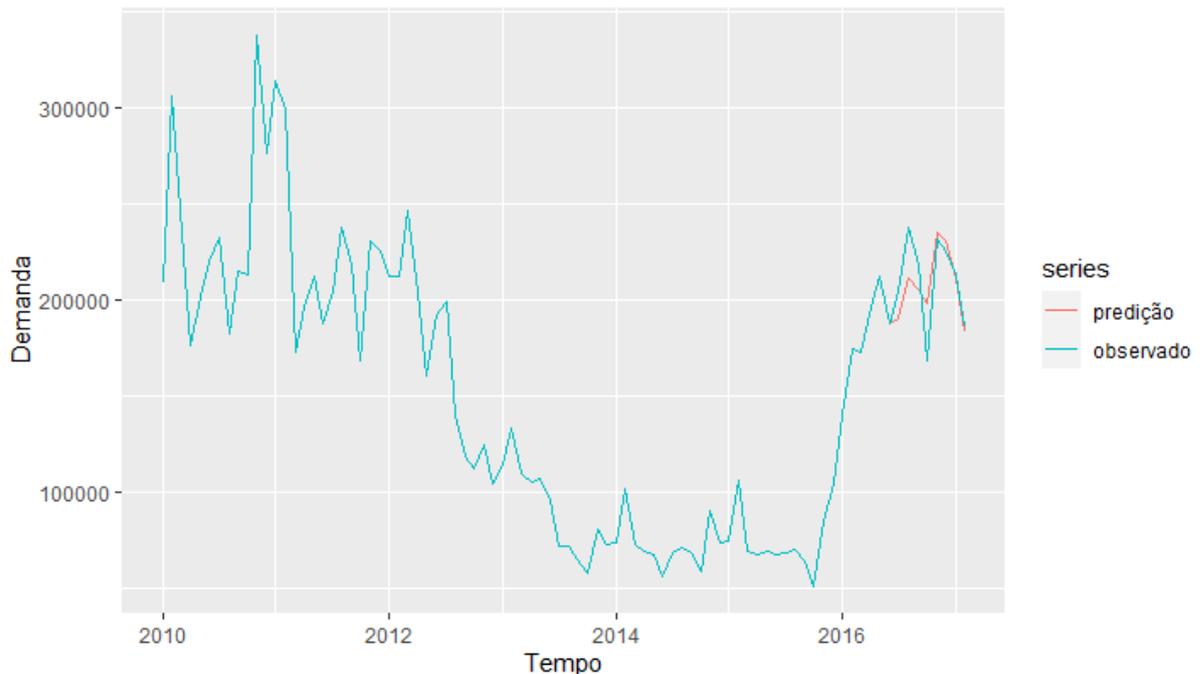


Figura 12 – Gráfico com os valores reais e previstos do modelo ARIMA

Com as informações da Tabela 3 e da Figura 12, verifica-se que o modelo de regressão dinâmica ajustado conseguiu prever bem os dados para os últimos 8 meses. O MAPE e RMSE foram baixos comparado aos outros modelos ajustados, indicando um bom resultado.

## 3.2 Redes Neurais Recorrentes

Na Tabela 4 abaixo é possível verificar os valores previstos através do método de redes neurais recorrentes:

Mês	Valor Real	Valor Previsto
jul/2016	204545	171353,1
ago/2016	237869	168123,1
set/2016	218513	167969
out/2016	168110	165288,5
nov/2016	231870	161884,9
dez/2016	224974	159054,5
jan/2017	213122	152365,3
fev/2017	185612	145833,6

Tabela 4 – Tabela de valores previstos para RNR

Na Figura 13 tem-se a representação visual do comportamento da previsão para os últimos 8 meses.

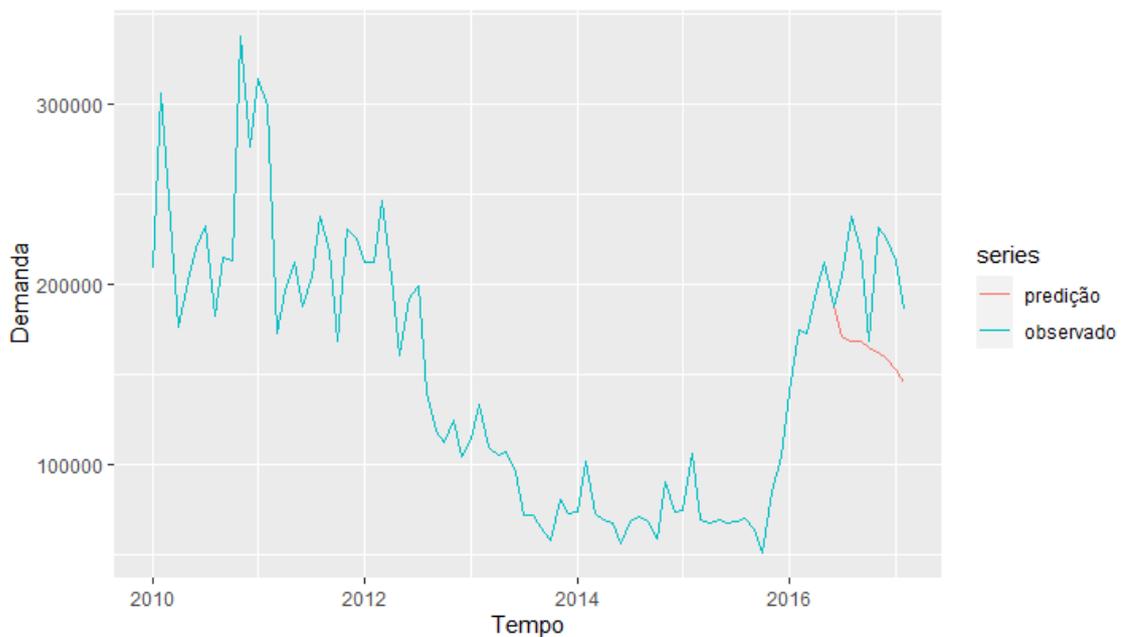


Figura 13 – Gráfico com os valores reais e previstos do modelo RNR

O cenário obtido com o modelo da RNR apresentou resultados abaixo do esperado para o horizonte de 8 meses. Pode-se observar que, de forma geral, as previsões ficaram subestimadas. Tal fato pode ter ocorrido devido a base de dados possuir um volume pequeno para a utilização de redes neurais recorrentes ou até mesmo devido ao comportamento da variável *Demand*, que a partir da metade de 2012 até o fim de 2015, sofreu uma queda brusca.

### 3.3 Comparação dos resultados

Para avaliação, comparação e escolha do modelo com melhor desempenho preditivo, foi utilizada a métrica Raiz do Erro Quadrático Médio (RMSE). O RMSE refere-se ao cenário que avalia o horizonte de previsão dos últimos 8 meses. Na Tabela 5, traz os valores da previsão e o RMSE para cada um dos métodos.

Mês	Valor Real	ARIMA(3,0,3)	RNR
jul/2016	204545	190726,1	171353,1
ago/2016	237869	211738,1	168123,1
set/2016	218513	205560,2	167969
out/2016	168110	197891,5	165288,5
nov/2016	231870	234989,6	161884,9
dez/2016	224974	230247,2	159054,5
jan/2017	213122	211023,7	152365,3
fev/2017	185612	183522,6	145833,6
	RMSE	15711,48	53672,43
	MAPE	5,90%	22,45%

Tabela 5 – Comparação dos resultados do RNR e ARIMA

Na Figura 14 abaixo compara a predição realizada por ambos os métodos, em azul tem-se os dados observados, em vermelho os valores preditos para o modelo ARIMA e em verde os valores preditos para as RNR.

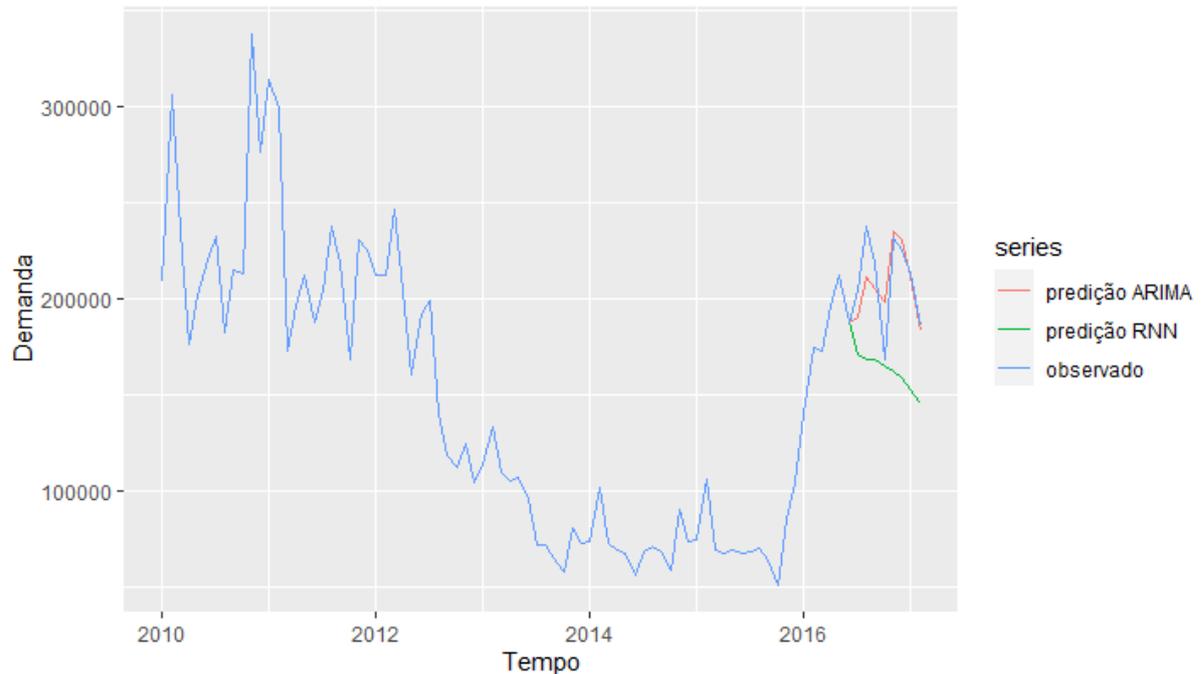


Figura 14 – Gráfico com os valores reais e previstos do modelo RNR e ARIMA

Ao observar a Tabela 5 nota-se um desempenho satisfatório para o modelo ARIMA(3,0,3), o qual produziu um menor RMSE para o horizonte de 8 meses de previsão. Este modelo apresentou valores mais próximos aos valores reais do conjunto de dados. Por outro lado, os valores obtidos pelo método RNR não demonstraram um bom desempenho.

Tais resultados podem estar associados a forma como cada análise é conduzida e a natureza dos dados observados. Para o RNR, por exemplo, a base foi dividida em validação, treino e teste. Além disso, a validação do modelo ocorreu em um cenário atípico da série temporal o que pode ter influenciado em valores mais baixos em seus resultados preditivos. Além disso, para a aplicação das redes neurais recorrentes não é necessário que seja satisfeita algumas condições para a aplicação do modelo.

Enquanto no modelo ARIMA foi necessário que pressupostos fossem satisfeitos para que o modelo fosse ajustado e validado, como a estacionariedade das observações e a análise de resíduos. Ademais, para o modelo ARIMA houve apenas a separação em treino e teste, não sendo necessário uma etapa de validação como a qual ocorreu no método de RNRs.

Outro ponto que pode ter influenciado nos resultados da comparação entre os dois métodos foi o volume de dados utilizados, visto que uma grande quantidade de observações é recomendável para a aplicação de métodos de aprendizagem de máquina e aprendizado profundo. A agregação da série para observações mensais, pode ter levado a resultados não tão satisfatórios para o método de redes neurais, enquanto o modelo ARIMA não precisa, necessariamente, de um volume massivo para realizar ajuste e previsão.

## 4 Conclusão

Em conclusão, este trabalho procurou abordar duas técnicas que podem ser aplicadas para dados de MMM e comparar seus resultados para a previsão. Para isso, foram feitas análises e pesquisas sobre séries temporais como modelos ARIMA e sobre redes neurais, em especial as redes neurais recorrentes, também conhecidas como RNR. Buscou-se comparar o desempenho das previsões dos métodos utilizados.

Os resultados obtidos utilizando os métodos de modelagem ARIMA e Rede Neural indicam que o modelo ARIMA(3,0,3) é mais adequado para modelar os dados da análise realizada. Isso pode ser atribuído à quantidade de dados necessários para a construção de um modelo ARIMA, que é menor do que o necessário para uma rede neural, além das diferenças nas características das técnicas de modelagem utilizadas em cada método, que podem ter influenciado nos resultados.

Com isso, propõe-se para trabalhos futuros, utilizar outras agregações do conjunto de dados, como uma abordagem semanal ou até mesmo com a base em seu formato original (diário) com a finalidade de comparar novamente os métodos e o desempenho das redes neurais recorrentes com a presença de mais observações no processo de aprendizagem.

Além disso, uma outra sugestão é a inclusão do *Gross Rating Point* para cada mídia, como variáveis explicativas no modelo, com o objetivo de verificar o impacto dessas variáveis no resultado final dos modelos. Assim como a possibilidade de comparar outros métodos de aprendizagem de máquina e aprendizagem profunda com relação às séries temporais, a fim de verificar um resultado mais satisfatório com o uso de outro método.

Este estudo é relevante porque apresenta novas formas de lidar com aplicação de métodos estatísticos em dados de mix de marketing. Recorrendo a métodos um pouco além dos popularmente usados para este tipo de contexto, como os modelos de regressão múltipla.

De maneira geral pode-se dizer que este trabalho atendeu à proposta inicial. Primeiramente, foi aprofundado os conhecimentos de séries temporais e RNR, bem como a consolidação de alguns conceitos. Em seguida, foram apresentadas etapas da pesquisa desde a coleta dos dados e suas variáveis, até a modelagem e resultados obtidos com os métodos aplicados, comparando o desempenho de cada um dos modelos ajustados.

Além disso foi realizada uma discussão sobre o desempenho preditivo de cada uma das abordagens e a recomendações sobre futuras aplicações para dados de mix de marketing.

O estudo realizado apresenta algumas limitações. A primeira, deve-se ao fato de que os dados vieram de um repositório aberto, sem a identificação da empresa e produtos comercializados. Com isso, não foi possível elaborar suposições sobre o comportamento atípico da série em alguns períodos. A segunda é atribuída à capacidade computacional

e a falta de diversificação de materiais de referência para a codificação dos modelos das redes neurais recorrentes. Foram realizadas as tentativas da utilização de outros pacotes/bibliotecas, entretanto não ocorreu o progresso na codificação. Com isso, o estudo não realizou outros comparativos de desempenho, tornando-se refém do pacote *torch*.

# Referências

- ARMSTRONG, J.; COLLOPY, F. Error measures for generalizing about forecasting methods: Empirical comparisons. *International Journal of Forecasting*, v. 8, n. 1, p. 69–80, 1992. ISSN 0169-2070. Disponível em: <<https://www.sciencedirect.com/science/article/pii/016920709290008W>>.
- BARRETO, J. Introdução às redes neurais artificiais. Laboratório de Conexionismo e Ciência Cognitivas UFSC - Departamento de Informática e de Estatística, 2002.
- BOSHNAKOV, G. Introduction to time series analysis and forecasting, 2nd edition, wiley series in probability and statistics, by douglas c.montgomery, cheryl l.jennings and muratkulahci (eds). published by john wiley and sons, hoboken, nj, usa, 2015. total number of pag: Introduction to time series analysis and forecasting, 2nd edition, wiley series in probability and statistics, by douglas c. montgomery, cheryl l. jen. *Journal of Time Series Analysis*, v. 37, 08 2016.
- CHAN, D.; PERRY, M. Challenges and opportunities in media mix modeling. *Google Inc.*, p. 1–16, 2017. Disponível em: <<https://ai.google/research/pubs/pub45998>>.
- CHORNOUS G., F. Y. Marketing mix modeling for pharmaceutical companies on the basis of data science technologies. *KTH Royal Institute of Technology*, ACCESS Press, p. 274–289, 8 2021. ISSN 2683-1007.
- DIAS, R. Estimativa da vazão de produção de poços utilizando técnicas de inteligência artificial. Pontifícia Universidade Católica - RJ, 2022.
- GUTIERREZ, J. L. C. *Monitoramento da instrumentação da barragem de Corumbá I por redes neurais e modelos de Box Jenkins*. Pontificada Universidade Católica do Rio de Janeiro - PUC-RIO, 2003. Disponível em: <[https://www.maxwell.vrac.puc-rio.br/4244/4244\\_5.PDF](https://www.maxwell.vrac.puc-rio.br/4244/4244_5.PDF)>.
- HYNDMAN, R.; ATHANASOPOULOS, G. *Forecasting: Principles and Practice*. 2nd. ed. Australia: OTexts, 2018.
- L TAVARES M, E. e. a. F. Redes neurais artificiais: Princípios básicos artificial neural networks: Basic principles. Revista Eletrônica Científica Inovação e Tecnologia, 2016.
- LECUN YANN;BENGIO, Y. G. Deep learning. *IEEE Consumer Electronics Magazine*, v. 6, n. 2, 2017.
- LEMLEY J.; BAZRAFKAN, S. C. P. Deep learning for consumer devices and services: Pushing the limits for machine learning, artificial intelligence, and computer visions. *IEEE Consumer Electronics Magazine*, v. 6, n. 2, 2017.
- MAKRIDAKIS, S.; WHEELWRIGHT, S.; HYNDMAN, R. J. *Forecasting: methods and applications*. [S.l.]: John Wiley Sons, 1998.
- MARSCHNER, D. C.; MCCARTHY, E. J. Basic marketing: A managerial approach. *Journal of Marketing*, v. 36, 1972. ISSN 00222429.
- MATOS, O. C. *Practical Time Series Analysis*. [S.l.]: Atlas, 2000.

MENDEZ-SUAREZ, M. Marketing mix modeling using pls-sem, bootstrapping the model coefficients. *Mathematics*, v. 9, 2021. ISSN 22277390.

MHITAREAN, E. *Marketing Mix Modelling from the multiple regression perspective*. Dissertação (Mestrado) — KTH, Mathematical Statistics, 2017.

MORETTIN, P. A.; TOLOI, C. M. d. C. *Análise de séries temporais*. [S.l.]: Edgard Blucher, 2004.

SOUZA, A. M.; KLIDZIO, R. Modelos de previsão aplicados ao controle de qualidade com dados autocorrelacionados. *Revista Produção Online*, v. 13, n. 4, 2013.

VASCO, L. P. Um estudo de redes neurais recorrentes no contexto de previsões no mercado financeiro. Trabalho de Conclusão de Curso– Universidade Federal de São Carlos–UFSCar), 2020.

WIGREN, F. C. R. Marketing mix modelling : A comparative study of statistical models. *Diva-Portal.Org*, 2019.

WOLFE, M. J.; CROTTS, J. C. Marketing mix modeling for the tourism industry: A best practices approach. *International Journal of Tourism Sciences*, v. 11, 2011. ISSN 1598-0634.