

FUNDAÇÃO GETULIO VARGAS  
ESCOLA de ECONOMIA de SÃO PAULO

Antonio Sergio Ferreira Bonato

Utilização de Modelos Não Lineares  
para Previsão do ICMS Paulista

São Paulo  
2023

**Antonio Sergio Ferreira Bonato**

# **Utilização de Modelos Não Lineares para Previsão do ICMS Paulista**

Dissertação para obtenção do grau de mestre apresentada à Escola de Economia de São Paulo

Área de concentração: Economia

Orientador: Emerson Fernandes Marçal

Coorientador: Pedro Luís Valls Pereira

**São Paulo  
2023**

Aluno, Nome C.

Utilização de Modelos Não Lineares para Previsão  
do ICMS Paulista

[152](#) páginas

Dissertação (Mestrado) - Escola de Economia de  
São Paulo.

1. Modelos de Mudança Markoviana
2. Redes Neurais Recorrentes LSTM
3. Model Confidence Set

**Antonio Sergio Ferreira Bonato**

# **Utilização de Modelos Não Lineares para Previsão do ICMS Paulista**

Dissertação para obtenção do grau de mestre  
apresentada à Escola de Economia de São Paulo.  
Área de concentração: Economia

E aprovada em 24/07/2023  
Pela comissão organizadora

---

Prof. Dr.

Emerson Fernandes Marçal

---

Prof. Dr.

Pedro Luís Valls Pereira

---

Prof. Dr.

Bruno Tebaldi de Queiroz Barbosa

*Dedico este trabalho à minha esposa Cristina e aos meus filhos Lucca, Caio, Bruno e  
Filipe.*

# Agradecimentos

Agradeço ao meu orientador, ao meu coorientador e aos meus colegas da Secretaria da Fazenda e Planejamento do Estado de São Paulo, especialmente ao economista André Grotti, por ter compartilhado sua profunda experiência na previsão da arrecadação do ICMS paulista.

No Brasil, até o passado é incerto.

Pedro Malan, Ministro da Fazenda (1995-2002)

# Resumo

Atualmente, na Secretaria da Fazenda e Planejamento do Estado de São Paulo, a previsão de arrecadação do ICMS paulista é feita usando-se um modelo SARIMA com variáveis exógenas, o SARIMAX. Mas o ICMS é não linear e há várias quebras estruturais controladas com diversas *dummies*. A consequência é um modelo sobre ajustado, poco adequado para previsões de mais longo prazo (mais do que alguns meses, muito menos para um ano). A proposta deste trabalho é utilizar modelos não lineares na previsão do ICMS paulista, utilizando redes neurais recorrentes LSTM (*Long Short-Term Memory*) e modelos Markov-Switching, fazendo um contraponto entre estes dois modelos. E utilizar, como base de comparação, o modelo SARIMAX atual.

**Palavras-chave:** Modelos de Mudança Markoviana. Redes Neurais Recorrentes. LSTM. Model Confidence Set. ICMS.

# Abstract

Currently, at the Department of Treasure and Planning of the State of São Paulo, the forecast of ICMS collection in São Paulo is made using a SARIMA model with exogenous variables, the SARIMAX. But ICMS is non-linear and there are several structural breaks controlled with different *dummies*. The consequence is an overfitted model, unsuitable for long-term forecasts (more than a few months, let alone a year). The purpose of this work is to use non-linear models to forecast the ICMS in São Paulo, using LSTM recurrent neural networks (*Long Short-Term Memory*) and Markov-Switching models, making a counterpoint between these two models. And use, as baseline for comparison, the current SARIMAX model.

**Keywords:** Markov-Switching Models. Recurrent Neural Networks. LSTM. Model Confidence Set. VAT.

# Listas de Figuras

2.1	IVA real comparado com o IVA previsto na Espanha, de 2013 a 2018.	6
2.2	Probabilidade suavizada de retrações (áreas acinzentadas) e expansões dos ciclos de negócio brasileiros. Dados anuais de 1900 a 1999.	17
3.1	PIB real trimestral dos EUA - em log, no painel de cima, e em taxa de crescimento, no painel de baixo - com a datação dos ciclos de retração (sombreados) e crescimento feitas pela NBER.	27
3.2	Modelo não linear de um neurônio. Adaptada de (HAYKIN, 1998)	34
3.3	Estrutura de um bloco LSTM. Adaptado de (WANG; ZHU; LI, 2019).	41
3.4	Série de arrecadação mensal nominal do ICMS em R\$, de janeiro de 2003 a dezembro de 2022	43
3.5	Séries de arrecadação mês a mês do ICMS, de 2003 a 2022; as linhas horizontais azuis representam a média de cada mês.	44
3.6	Comparação entre as previsões de mercado do Bacen e os valores reais da série do IGP-M.	54
3.7	Comparação entre as previsões de mercado do Bacen e os valores reais da série do Dólar.	54
3.8	Comparação entre as previsões de mercado do PIB do Bacen, convertidas em IBC-Br, e os valores reais da série do IBC-Br.	56
4.1	Séries temporais utilizadas nos experimentos de 2022.	62
4.2	Séries em primeira diferença.	63
4.3	Séries em segunda diferença.	63
4.4	Série do ICMS em nível, com e sem sazonalidade	64
4.5	Séries em nível, com e sem sazonalidade: IBC-Br, IGP-M e Dólar	65
4.6	Sumário do Modelo SARIMA(0,1,1)(2,0,1)	67

4.7	Análise dos resíduos do Modelo SARIMA(0,1,1)(2,0,1) . . . . .	68
4.8	Sumário do Modelo SARIMA(0,1,1)(1,0,1) . . . . .	68
4.9	Análise dos resíduos do Modelo SARIMA(0,1,1)(1,0,1) . . . . .	68
4.10	Sumário do Modelo SARIMAX(0,1,1)(2,0,0) . . . . .	70
4.11	Correlogramas dos resíduos do Modelo SARIMAX(0,1,1)(2,0,0) . . . . .	70
4.12	Sumário do Modelo SARIMAX(0,1,1)(1,0,1) . . . . .	71
4.13	Correlogramas dos resíduos do Modelo SARIMAX(0,1,1)(1,0,1) . . . . .	71
4.14	Resultados do modelo de Markov com Variância Fixa . . . . .	72
4.15	Resultados do modelo de Markov com Mudança de Variância . . . . .	73
4.16	Resultados do modelo de Markov com Mudança de Média e Variância . . . . .	74
4.17	Gráfico da perda em treino e validação do modelo RNN-LSTM 1 mês . . . . .	75
4.18	Gráfico da perda em treino e validação do modelo RNN-LSTM 12 meses . . . . .	75
4.19	Métricas RMSE dos modelos univariados . . . . .	77
4.20	Métricas MAPE dos modelos univariados . . . . .	77
4.21	Previsões do Modelo Naïve comparada com o Realizado . . . . .	78
4.22	Previsões dos Modelos SARIMA e SARIMAX comparadas com o Realizado . . . . .	78
4.23	Previsões dos Modelos de Mudança Markoviana comparadas com o Realizado . . . . .	79
4.24	Previsões dos Modelos RNN-LSTM comparadas com o Realizado . . . . .	79
4.25	Sumário do Modelo SARIMAX(0,1,1)(2,0,1) . . . . .	82
4.26	Análise dos resíduos do Modelo SARIMAX(0,1,1)(2,0,1) . . . . .	82
4.27	Sumário do Modelo SARIMAX(0,1,1)(1,0,1) . . . . .	83
4.28	Análise dos resíduos do Modelo SARIMAX(0,1,1)(1,0,1) . . . . .	83
4.29	Resultados do modelo de Markov com Variância Fixa . . . . .	84
4.30	Resultados do modelo de Markov com Mudança de Variância . . . . .	85
4.31	Resultados do modelo de Markov com Mudança de Variância . . . . .	86
4.32	Gráfico da perda em treino e validação do modelo RNN-LSTM . . . . .	87
4.33	Métricas RMSE dos modelos de previsão . . . . .	88
4.34	Métricas MAPE dos modelos de previsão . . . . .	89
4.35	Previsões dos Modelos univariados com regressão múltipla com Variáveis Previstas comparadas com o Realizado . . . . .	90
4.36	Métricas RMSE e MAPE dos modelos SARIMAX com o uso de variáveis exógenas com valores previstos pelo mercado e valores reais. . . . .	90

4.37	Métricas RMSE e MAPE do modelo RNN-LSTM com o uso de variáveis exógenas com valores previstos pelo mercado e valores reais. . . . .	91
4.38	Métricas RMSE e MAPE dos modelos de mudança markoviana com o uso de variáveis exógenas com valores previstos pelo mercado e valores reais. . .	91
4.39	Comparação dos menores MAPE por experimento por ano independente de estratégia. . . . .	93
4.40	Apenas o experimento com menor MAPE por ano independente de estratégia. . . . .	93
4.41	Menor MAPE por estratégia por ano independente de experimento. . . . .	94
4.42	MAPE por estratégia por ano no experimento univariado. . . . .	95
4.43	Menor MAPE por estratégia por ano no experimento univariado. . . . .	95
4.44	MAPE por estratégia por ano no experimento univariado com regressão múltipla. . . . .	96
4.45	Menor MAPE por estratégia por ano no experimento univariado com regressão múltipla. . . . .	97
4.46	MAPE por ano por estratégia no experimento univariado. . . . .	97
4.47	MAPE por ano por estratégia no experimento univariado com regressão múltipla. . . . .	98
4.48	Valores de arrecadação, em R\$, previstos pelo modelo da estratégia Naïve comparados com os valores reais arrecadados. . . . .	99
4.49	Valores de arrecadação, em R\$, previstos pelo modelos das estratégias SARIMA/SARIMAX comparados com os valores reais arrecadados. . . . .	99
4.50	Valores de arrecadação, em R\$, previstos pelos modelos da estratégia Markov-Switching comparados com os valores reais arrecadados. . . . .	100
4.51	Valores de arrecadação, em R\$, previstos pelos modelos da estratégia RNN-LSTM comparados com as previsões do modelo Naïve. . . . .	101
4.52	Valores de arrecadação, em R\$, previstos pelos modelos da estratégia RNN-LSTM comparados com os valores reais arrecadados. . . . .	101
4.53	MAPE dos modelos considerando-se as previsões de arrecadação de dez anos produzida por cada um deles. . . . .	102
4.54	Resultados do ranqueamento dos modelos pelo método MCS. . . . .	103
4.55	Resultados do ranqueamento dos modelos pelo menor MAPE a posteriori. .	104

4.56 Valores de arrecadação, em R\$, previstos pelos modelos com os melhores MAPEs por ano com escolha a posteriori comparados com os valores reais arrecadados . . . . .	105
4.57 MAPE da previsão de 10 anos realizada pelos modelos com menores MAPE por ano com escolha a posteriori comparada com a do melhor modelo escolhido pelo MCS . . . . .	105

# **Lista de Tabelas**

2.1	MAPE obtidos para a previsão um mês à frente com os modelos de Alisamento Exponencial de Holt-Winters (HW), SARIMA e Rede Neural Artificial MLP . . . . .	7
2.2	MAPE obtidos para a previsão quatro meses à frente com os modelos de Alisamento Exponencial de Holt-Winters (HW), SARIMA e Rede Neural Artificial MLP . . . . .	7
2.3	Erro Relativo da para a previsão da arrecadação anual do ICMS-RJ em 2019. Os modelos SARIMA e VAR são os utilizados pela SEFAZ-RJ para a previsão oficial de arrecadação. . . . .	10
2.4	Erro Relativo da para a previsão da arrecadação anual do ICMS-RJ em 2018. O modelo SEFAZ-RJ se refere ao modelo, não informado pelo artigo, utilizado para a previsão oficial de arrecadação . . . . .	11
2.5	RMSE dos Modelos Arima e RNN-LSTM obtidos por (SIAMI-NAMINI; TAVAKOLI; NAMIN, 2018) . . . . .	13
2.6	MSE dos Modelos de (YAN; OUYANG, 2018) . . . . .	14
2.7	RSS dos Modelos de (HANSSON, 2017) . . . . .	14
2.8	Resultados da previsão do PIB-Brasil, fora da amostra, do 2º Tri de 1989 ao 1º Tri de 2000 . . . . .	18
2.9	Resultados da previsão obtidos por (ROCHA; PEREIRA, 2015) . . . . .	19
2.10	Resultados da previsão obtidos por (NGUYEN, 2020) . . . . .	19
3.1	Cálculos realizados sobre os valores da série do ICMS, conforme (SILVA; FIGUEIREDO, 2020) . . . . .	50
4.1	Resultados dos testes ADF e KPSS para as séries em nível. . . . .	62
4.2	Resultados dos testes ADF e KPSS para as séries em primeira diferença. .	62

4.3	Resultados dos testes ADF e KPSS para as séries em segunda diferença. . . . .	63
4.4	Outliers detectados automaticamente . . . . .	64
4.5	Métricas dos modelos de previsão univariados . . . . .	76
4.6	Escolha dos modelos univariados de 2022 e seus critérios. . . . .	80
4.7	Métricas dos modelos de previsão univariados com regressão múltipla . .	88
4.8	Escolha dos modelos univariados com regressão múltipla de 2022 e seus critérios. . . . .	92
4.9	MAPE dos modelos considerando-se as previsões de arrecadação de dez anos produzida por cada um deles. . . . .	102
4.10	Modelos com Menores MAPE por ano com escolha a posteriori, independente de experimento. . . . .	104
4.11	MAPE da previsão de 10 anos realizada pelos modelos com menores MAPE por ano com escolha a posteriori comparada com a do melhor modelo escolhido pelo MCS . . . . .	104
A.1	Arrecadação nominal do ICMS de janeiro de 2003 a dezembro de 2006 . . .	109
A.2	Arrecadação nominal do ICMS de janeiro de 2007 a dezembro de 2012 . . .	110
A.3	Arrecadação nominal do ICMS de janeiro de 2013 a dezembro de 2018 . . .	111
A.4	Arrecadação nominal do ICMS de janeiro de 2019 a dezembro de 2022 . . .	112

# Sumário

<b>1</b>	<b>Introdução</b>	<b>1</b>
1.1	Objetivo . . . . .	2
1.2	Justificativa . . . . .	2
1.3	Abordagem . . . . .	2
1.4	Organização do Trabalho . . . . .	3
<b>2</b>	<b>Bibliografia fundamental</b>	<b>4</b>
2.1	Previsões de Arrecadação . . . . .	4
2.1.1	Estimativas e Previsões do IVA . . . . .	4
2.2	Redes Neurais Artificiais . . . . .	6
2.2.1	Experimentos Iniciais no Uso de Redes Neurais na Previsões do ICMS	6
2.2.2	Experimentos Recentes no Uso de Redes Neurais na Previsões do ICMS . . . . .	8
2.2.3	Outros Experimentos de Previsão com Redes Neurais Recorrentes LSTM . . . . .	12
2.3	Modelos de Mudança Markoviana . . . . .	15
2.4	Avaliação de Modelos . . . . .	20
<b>3</b>	<b>Metodologia</b>	<b>21</b>
3.1	Fundamentação Teórica . . . . .	21
3.1.1	Modelos ARIMA . . . . .	21
3.1.2	Modelos de Mudança Markoviana . . . . .	27
3.1.3	Redes Neurais Recorrentes LSTM . . . . .	33
3.1.4	Model Confidence Set . . . . .	41
3.2	Metodologia . . . . .	42

3.2.1	Dados . . . . .	42
3.2.2	Experimentos . . . . .	45
3.2.3	Escolha dos Melhores Modelos . . . . .	59
3.2.4	Model Confidence Set . . . . .	60
<b>4</b>	<b>Resultados</b>	<b>61</b>
4.1	Análise dos Dados . . . . .	61
4.1.1	Estacionariedade . . . . .	61
4.1.2	Sazonalidade . . . . .	63
4.1.3	Outliers . . . . .	64
4.1.4	Teste de Causalidade de Granger . . . . .	65
4.2	Experimento Univariado . . . . .	66
4.2.1	Modelo Naïve . . . . .	66
4.2.2	Modelos SARIMA e SARIMAX . . . . .	66
4.2.3	Modelos de Mudança Markoviana . . . . .	71
4.2.4	Modelo RNN LSTM . . . . .	73
4.2.5	Previsões . . . . .	76
4.3	Experimento Univariado com Regressão Múltipla . . . . .	79
4.3.1	Modelos SARIMAX . . . . .	80
4.3.2	Modelos de Mudança Markoviana . . . . .	84
4.3.3	Modelo RNN LSTM . . . . .	86
4.3.4	Previsões . . . . .	87
4.3.5	Comparação dos Experimentos de 2022 . . . . .	92
4.4	Comparação Geral dos Experimentos . . . . .	92
4.4.1	Todas as Estratégias e Experimentos . . . . .	94
4.4.2	Desempenho Anual dos Experimentos . . . . .	96
4.5	Ranking dos Modelos pelo MCS . . . . .	101
4.5.1	Rankeamento dos Modelos a Posteriori . . . . .	103
<b>5</b>	<b>Conclusão</b>	<b>106</b>
<b>A</b>	<b>Dados de Arrecadação do ICMS</b>	<b>109</b>

<b>B Sumários dos Modelos de Mudança Markoviana</b>	<b>113</b>
B.1 Experimento Univariado 2022 . . . . .	113
B.1.1 Modelo de Mudança Markoviana com Variância Fixa . . . . .	113
B.1.2 Modelo de Mudança Markoviana com Mudança de Variância . . . . .	117
B.1.3 Modelo de Mudança Markoviana com Componente Média-Variância	122
B.2 Experimento Univariado com Regressão Múltipla 2022 . . . . .	127
B.2.1 Modelo de Mudança Markoviana com Variância Fixa . . . . .	127
B.2.2 Modelo de Mudança Markoviana com Mudança de Variância . . . . .	134
B.2.3 Modelo de Mudança Markoviana com Componente Média-Variância	138
<b>Referências Bibliográficas</b>	<b>147</b>

# Capítulo 1

## Introdução

O ICMS é o Imposto sobre a Circulação de Mercadorias e Serviços. Classificado no grupo de Impostos sobre a Produção e a Circulação, apresenta-se hoje como um dos mais significativos tributos brasileiros, tanto no aspecto financeiro (quase 25% de toda a receita tributária), quanto no da técnica de Administração Tributária ([DINIZ, 2012](#)).

A receita tributária nominal total do Estado de São Paulo, em 2022, foi de 238 bilhões de reais. Destes, aproximadamente 85% são oriundos do arrecadação do ICMS, que foi de 201.2 bilhões de reais ([TRIBUTÁRIA, 2023](#)). Por este motivo é tão importante para a administração pública ser capaz de prever corretamente a arrecadação do ICMS.

Existe ainda a obrigação legal, imposta aos gestores públicos pela Lei de Responsabilidade Fiscal<sup>1</sup> (LRF), que afirma em seu artigo 11, *caput*:

Constituem requisitos essenciais da responsabilidade na gestão fiscal a instituição, previsão e efetiva arrecadação de todos os tributos da competência constitucional do ente da Federação.

E, em seu artigo 12, *caput*:

As previsões de receita observarão as normas técnicas e legais, considerarão os efeitos das alterações na legislação, da variação do índice de preços, do crescimento econômico ou de qualquer outro fator relevante e serão acompanhadas de demonstrativo de sua evolução nos últimos três anos, da projeção para os dois seguintes àquele a que se referirem, e da metodologia de cálculo e premissas utilizadas.

---

<sup>1</sup>Lei de Responsabilidade Fiscal: <[http://www.planalto.gov.br/ccivil\\_03/leis//lcp/lcp101.htm](http://www.planalto.gov.br/ccivil_03/leis//lcp/lcp101.htm)>

## 1.1 Objetivo

O objetivo deste trabalho é o de experimentar o uso de métodos de previsão não lineares para a previsão da receita tributária do ICMS do Estado de São Paulo. Os modelos a serem testados são os modelos de mudança markoviana, ou *Markov-Switching Models* e os modelos de Redes Neurais Recorrentes com neurônios LSTM. Não serão utilizadas variáveis *dummy* de intervenção nestes modelos.

Um objetivo secundário do trabalho é o de obter uma datação dos ciclos de expansão e retração da arrecadação do estado com o uso os modelos de mudança markoviana.

Este trabalho não tem como objetivo realizar um estudo teórico dos modelos não lineares, mas sim o de ser um trabalho empírico para determinar se o uso dos modelos não lineares é mais adequado à previsão da série de arrecadação do ICMS do estado que o uso de modelos lineares com variáveis *dummy* de intervenção.

## 1.2 Justificativa

O processo gerador da série temporal do ICMS, intuitivamente, é não linear. Ele muda em resposta aos ciclos econômicos de expansão e retração da economia, bem como aos processos inflacionários. Entretanto, o método utilizado atualmente pela SEFAZ-SP é linear, um modelo SARIMA com variáveis exógenas denominado SARIMAX. As quebras estruturais da série, que introduzem as não linearidades, são modeladas com o uso de variáveis *dummies*. Isto faz com que o modelo fique sobre ajustado (*overfitted*), o que provavelmente o torna menos adequado para previsões para além do mês seguinte, quanto mais para previsões de 12 meses. E previsões de 12 meses são fundamentais para o planejamento e a execução orçamentária.

## 1.3 Abordagem

Este trabalho irá adotar como abordagem a realização de 20 experimentos de previsão, sendo 10 univariados e 10 univariados com regressão múltipla, para os anos de 2013 a 2022. Em cada experimento serão utilizados os modelos lineares SARIMA e SARIMAX, e os modelos não lineares Markov-Switching e Redes Neurais Recorrentes LSTM.

Os modelos resultantes terão seus resultados ranqueados pelo seu poder preditivo com

o uso do Model Confidence Set.

## 1.4 Organização do Trabalho

Este trabalho está organizado da seguinte forma. O capítulo 1 contém a introdução, o objetivo do trabalho e sua justificativa. No capítulo 2 é feita a revisão da literatura, com artigos selecionados sobre a previsão de arrecadação, previsões utilizando-se Redes Neurais Recorrentes e previsões utilizando-se Modelos de Mudança Markoviana. O uso de Model Confidence Set para a seleção de modelos também é resenhado.

O capítulo 3 apresenta uma breve revisão teórica dos modelos e conceitos utilizados neste trabalho, e depois apresenta os dados utilizados e a metodologia que foi seguida na execução dos experimentos. O capítulo 4 apresenta os resultados obtidos na análise dos dados e na execução dos experimentos do ano de 2022, bem como o resultado geral de todos os experimentos e o ranqueamento dos modelos pelo MCS.

A conclusão, finalmente, comenta os resultados obtidos e aponta possíveis caminhos para trabalhos futuros.

Nos apêndices são apresentados os dados mensais da receita de ICMS do estado utilizados no trabalho e os sumários dos modelos de mudança markoviana de 2022.

Os resultados detalhados dos experimentos de 2013 a 2021 não estão descritos neste texto como foram descritos os de 2022 apenas por uma questão de brevidade. Entretanto, todas as análises estão disponíveis para consulta em um repositório do Github, assim como todo o código fonte do trabalho, bem como as bases de dados utilizadas.

O link do repositório do Github é : <[https://github.com/asbonato/mpe\\_fgv\\_2023.git](https://github.com/asbonato/mpe_fgv_2023.git)>.

# Capítulo 2

## Bibliografia fundamental

Este capítulo irá abordar, com maior ou menor minúcia, os trabalhos acadêmicos utilizados na elaboração desta dissertação, a depender de sua relevância imediata para o tema em questão.

### 2.1 Previsões de Arrecadação

O ICMS é o imposto brasileiro sobre valor agregado, ou IVA, em inglês, VAT<sup>1</sup>. É o principal imposto do Brasil, conforme ([SILBERSCHNEIDER, 2000](#)).

#### 2.1.1 Estimativas e Previsões do IVA

Uma técnica de estimativa do IVA baseada em indicadores macroeconômicos foi apresentada por ([JENKINS et al., 2000](#)), no capítulo 8 de sua obra. Segundo os autores, é importante para os países poderem estimar a base do seu IVA pois este tipo de imposto vem substituindo tributos sobre receita e sobre o consumo de alguns produtos em diversos países. Há diferentes modos de se calcular o IVA devido, mas o mais comum tem sido o método do valor da nota fiscal emitida na transação. São apresentadas três tipos de estimativas da base do IVA: a abordagem agregada, a abordagem setorial e a abordagem de entrada e saída.

- abordagem agregada: inicia-se pelo PIB. O IVA é estimado subtraindo-se as exportações e somando-se as importações de modo a refletir o consumo.

---

<sup>1</sup> *Value Added Tax, ou Imposto de Valor Agregado*

- abordagem setorial: se baseia no valor adicionado por cada setor da economia e assim reflete a contribuição setorial.
- abordagem de entrada e saída: a base do IVA é estimada empregando-se a equivalência da base do imposto sobre valor agregado a um tributo sobre vendas no varejo cobrado sobre o preço final de venda dos produtos e serviços.

No final do capítulo ([JENKINS et al., 2000](#)) apresentam um estudo de caso de aplicação da abordagem de entradas e saídas no México em 1997.

Outro trabalho que apresenta um refinamento da abordagem de estimativa da base de cálculo do IVA por meio de entradas e saídas é o de ([MINH, 2007](#)), que mostra a aplicação do método no sistema tributário da Romênia.

Um experimento de previsão do IVA espanhol utilizando-se de indicadores macroeconômicos de Contas Nacionais coletados para estimativa do PIB foi realizado por ([VASCO; LÓPEZ, 2018](#)). Os autores utilizaram indicadores de diferentes áreas, como Demanda Doméstica, Mercado de Trabalho e Indicadores de Atividade e, devido ao alto grau de dependência linear entre estes indicadores, o que causava problemas de multicolinearidade no modelo, fizeram uso da técnica de PCA<sup>2</sup> - proposto por ([PEARSON, 1901](#)) e que vem passando por atualizações - conforme ([JOLLIFFE; CADIMA, 2016](#)), para reduzirem o conjunto de dados a três variáveis.

A escolha dos indicadores macroeconômicos se deu por meio do estudo da correlação entre cada um deles e a arrecadação do IVA. Os indicadores analisados foram o Índice de Produção Industrial em Bens de Consumo, as Importações de Bens e Serviços, os Registros de Veículos de Passageiros, as Vendas das Grandes Empresas de Consumo, o Investimento Imobiliário, as Chegadas de Turistas Estrangeiros, a Folha de Pagamento, o Emprego Formal, o Consumo de Energia Elétrica, o Percentual de Utilização da Capacidade Industrial e Pesquisas de Tendências de Negócios.

Após o tratamento das séries citadas, removendo ciclos e tendências, utilizaram-se de modelos ARIMA Sazonais (SARIMA) para realizar as previsões. Os autores concluíram que as séries de Índice de Produção Industrial em Bens de Consumo, Percentual de Utilização da Capacidade Industrial, Registros de Veículos de Passageiros e Pesquisas de Tendências de Negócios adicionaram informação antecessora significativa para o modelo. Já as demais variáveis tiveram sua relação contemporânea com a série do IVA explorada.

---

<sup>2</sup>do inglês *Principal Component Analysis*, ou Análise de Componente Principal

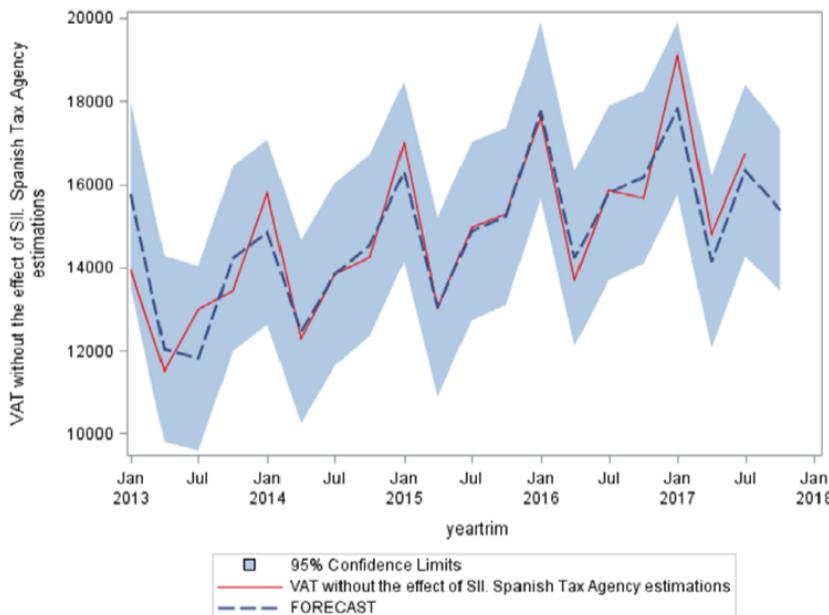


Figura 2.1: IVA real comparado com o IVA previsto na Espanha, de 2013 a 2018.

Os autores fizeram previsões 1 mês a frente, de maneira recursiva, e obtiveram os resultados reportados na figura 2.1.

## 2.2 Redes Neurais Artificiais

Há muitos trabalhos publicados sobre a aplicação de métodos não lineares de aprendizado de máquina na previsão econômica, a sua maioria focada em métodos baseados em Redes Neurais, principalmente os mais antigos ([MASINI; MEDEIROS; MENDES, 2023](#)).

Nesta seção serão abordados alguns trabalhos sobre o uso de Redes Neurais na previsão do ICMS.

### 2.2.1 Experimentos Iniciais no Uso de Redes Neurais na Previsão do ICMS

Em 2005 ([CONTRERAS, 2005](#)) publicou um estudo de previsão usando redes neurais artificiais MLP<sup>3</sup> para os estados de Pernambuco, Rio de Janeiro e São Paulo. Utilizando-se dos dados mensais de arrecadação destes três estados no período compreendido entre julho de 1994 e novembro de 2004, foram realizadas previsões um mês e quatro meses à frente. Para realizar as previsões foram usados modelos SARIMA, ver ([ENDERS, 2014](#)); Ali-

<sup>3</sup>*Multi-Layer Perceptron*, ou Perceptrons Multi Camada

samento Exponencial de Holt-Winters, ver ([HYNDMAN; ATHANASOPOULOS, 2018](#)) e Redes Neurais Articiais MLP, ver ([MITCHELL, 1997](#)) e ([DUDA; HART; STORK, 2012](#)). A métrica utilizada foi o MAPE<sup>4</sup>, Erro Percentual Médio Absoluto, expresso pela equação 2.1

$$MAPE = \frac{1}{n} \sum_{i=1}^n \frac{|y_i - \hat{y}_i|}{y_i} \quad (2.1)$$

onde  $n$  é o número total de observações,  $y_i$  é o valor real da observação e  $\hat{y}_i$  é o valor predito.

O autor obteve maior precisão na previsão um passo à frente com o alisamento exponencial para o ICMS de Pernambuco e Rio de Janeiro, mas as redes neurais MLP se saíram melhor para o ICMS de São Paulo<sup>5</sup>, conforme tabela 2.1. Porém, para múltiplos passos, o alisamento exponencial se saiu melhor para PE, o SARIMA para o RH e, novamente, a rede neural MLP para o ICMS do estado de São Paulo<sup>6</sup>, conforme a tabela 2.2.

Série	HW	SARIMA	Rede MLP
ICMS-PE	3.3489	4.8448	4.6070
ICMS-RJ	6.5078	10.0483	7.2851
ICMS-SP	3.9701	1.1906	1.7806

Tabela 2.1: MAPE obtidos para a previsão um mês à frente com os modelos de Alisamento Exponencial de Holt-Winters (HW), SARIMA e Rede Neural Artificial MLP

Série	HW	SARIMA	Rede MLP
ICMS-PE	3.3569	4.0248	3.6205
ICMS-RJ	5.4974	6.8847	6.9249
ICMS-SP	5.2460	2.2296	1.6440

Tabela 2.2: MAPE obtidos para a previsão quatro meses à frente com os modelos de Alisamento Exponencial de Holt-Winters (HW), SARIMA e Rede Neural Artificial MLP

A previsão da arrecadação federal do Brasil foi realizada por ([FREITAS; CIARELLI; SOUZA, 2009](#)) utilizando redes neurais. Os autores utilizaram os dados obtidos a partir da

<sup>4</sup> Mean Absolute Percentage Error

<sup>5</sup> Valores de MAPE compilados a partir das tabelas 4.8, 4.15 e 4.20 do trabalho de ([CONTRERAS, 2005](#)), sempre da coluna prev-o.

<sup>6</sup> Valores de MAPE compilados a partir das tabelas 4.8, 4.15 e 4.20 do trabalho de ([CONTRERAS, 2005](#)), sempre da coluna prev-m.

série histórica "Arrecadação das receitas federais - receita bruta R\$ (milhões)" divulgada pelo IPEA, contendo 176 observações mensais no período de julho de 1994 a fevereiro de 2009. Os modelos utilizados basearam-se em Redes Neurais Autorregressivas (RNAR) e Redes Neurais Probabilísticas (RNP). Foram realizadas previsões para 12 meses, recursivas, mês a mês. Os resultados foram comparados com os obtidos por um modelo ARIMA. O MAPE obtido para o modelo RNAR, para 12 meses - de janeiro a dezembro de 2001, fora da amostra, foi 6,79%. Para o modelo ARIMA foi 7,74%. E, para o modelo RNP, foi 7,87%. Portanto, o melhor modelo foi o RNAR.

Em um estudo publicado em 2006, ([SISNANDO; FREITAS, 2006](#)) utilizaram Redes Neurais MLP para a previsão e avaliação do desempenho de contribuintes do ICMS do estado do Ceará, reportando bons resultados, embora o experimento não tenha sido na previsão da arrecadação propriamente dita.

## 2.2.2 Experimentos Recentes no Uso de Redes Neurais na Previsões do ICMS

Em 2020 ([SILVA; FIGUEIREDO, 2020](#)) reportaram resultados expressivos com o uso de modelos LSTM na previsão de arrecadação do ICMS fluminense. As RNN<sup>7</sup> LSTM<sup>8</sup>, de acordo com ([HOCHREITER; SCHMIDHUBER, 1997](#)), são uma variação das redes neurais recorrentes tradicionais que, de acordo com ([SCHMIDHUBER, 2015](#)), tem tido ótimos resultados na previsão de séries temporais.

Em seu trabalho, as autoras realizaram 10 experimentos com Redes Neurais, sendo 8 com RNN-LSTM e 2 com redes neurais MLP convencionais. Em cada experimento foram feitas variações do número de camadas, número de neurônios da camada escondida e funções de ativação, conforme as configurações a seguir:

- 1 e 2 camadas;
- funções de ativação ReLU e Sigmoide nas redes RNN-LSTM e função Tangente Hiperbólica na rede MLP.
- algoritmo de otimização Adam nas redes RNN-LSTM e SGD<sup>9</sup>

---

<sup>7</sup>*Recurrent Neural Networks*, ou Redes Neurais Recorrentes

<sup>8</sup>*Long Short-Term Memory*, ou Memória de Curto Prazo Longa

<sup>9</sup>*Stochastic Gradient Descent*, ou Descida Estocástica do Gradiente

- de 1 a 100 neurônios na primeira camada escondida e apenas um neurônio na segunda camada escondida, caso houvesse.
- *mini-batch* igual a 32.
- taxa de aprendizado igual a 0.001.
- foi utilizado *early stopping* para interromper o treinamento.
- não foi reportado o uso de *dropout*.

Os dados da arrecadação nominal mensal do ICMS foram obtidos no site da SEFAZ-RJ. O período se estende de janeiro de 2002 a dezembro de 2019. Os conjuntos de treinamento, validação e testes foram divididos nos períodos de janeiro de 2002 a dezembro de 2016 para treinamento, janeiro de 2017 a dezembro de 2018 para validação e janeiro de 2019 a dezembro de 2019 para testes. Os autores utilizaram apenas a série do ICMS na previsão, sem variáveis exógenas. Para a criação do conjunto de características a ser trabalhado na rede neural eles utilizaram a série do ICMS com defasagens (*lags*) de 1, 2, 3, 4, 6 e 12 meses, além de médias móveis de 2, 3, 6 e 12 meses.

A previsão realizada foi de 12 meses, pelo método *multi-step* (WANG; ZHU; LI, 2019) com janela de 12 meses: faz-se a previsão do mês de janeiro de 2019 utilizando-se os dados dos últimos 12 meses - janeiro a dezembro de 2018 - e incorpora-se esta previsão na série. Para a previsão de fevereiro de 2019 utiliza-se a janela de fevereiro de 2018 a janeiro de 2019. E assim por diante. Para a previsão de dezembro de 2019, todos os dados, exceto o de dezembro de 2018, que é um valor realizado, são dados previstos.

O melhor modelo encontrado pelos autores for a RNN-LSTM com 2 camadas e 50 neurônios e função de ativação Sísmoide. Na tabela 2.3 se veem os resultados obtidos, expressos em termos de Erro Relativo, calculado conforme a equação 2.2.

$$\text{Erro Relativo} = \frac{| \text{Previsão} - \text{Real} |}{\text{Real}} \quad (2.2)$$

O experimento de (SILVA; FIGUEIREDO, 2020) com o ICMS-RJ foi renovado por (DORNELAS; CAMPOS; FIGUEIREDO, 2022), porém desta vez utilizando também variáveis exógenas. Os autores afirmam terem testado diversos dados de produtos e serviços que possuíam cobrança de ICMS e testaram a correlação de Pearson (PEARSON, 1896) de cada uma das séries com a do ICMS e selecionaram as que apresentaram resultados

Modelo	Erro Relativo (%)
SARIMA e VAR	1.49
Rede LSTM	0.13
Rede MLP	0.85

Tabela 2.3: Erro Relativo da para a previsão da arrecadação anual do ICMS-RJ em 2019. Os modelos SARIMA e VAR são os utilizados pela SEFAZ-RJ para a previsão oficial de arrecadação.

mais próximos de 1 - quanto mais próximo de 1 maior é a relação entre as variáveis, quanto mais próximo de -1, menor a relação. Procedendo desta forma escolheram como exógenas a série do PIB, obtida do site do IPEA<sup>10</sup>, a série de consumo de energia elétrica comercial na região sudeste, obtida no site da Empresa de Pesquisa Energética<sup>11</sup> e a série do ICMS-RJ, obtida no site da SEFAZ-RJ<sup>12</sup>. Todas as séries contêm dados mensais de janeiro de 2002 a dezembro de 2019. Os dados foram divididos em três conjuntos: treinamento, de janeiro de 2002 a dezembro de 2016, validação de janeiro de 2017 a dezembro de 2018 e testes, de janeiro de 2019 a dezembro de 2019.

De maneira análoga ao trabalho anterior, foi utilizado o método de janelamento deslizante (*multi-step*) para a realização das previsões. Também foi utilizado o método *multi-step* direto, conforme (CHEVILLON, 2007), de forma que a previsão 2 meses foi feita utilizando o mesmo conjunto de entradas de 24 meses anteriores, sem uso de nenhum processo autorregressivo ou atualização das entradas.

Foram conduzidos três estudos de caso com Redes Neurais Recorrentes LSTM: um cenário univariado utilizando apenas a série do ICMS, um cenário univariado com uma variável exógena utilizando a série do ICMS e a do PIB e um terceiro cenário utilizando a série do ICMS e duas exógenas, as séries do PIB e do consumo de energia elétrica.

Em todos os cenários foram feitas 10 execuções, utilizando *mini-batch* igual a 32, algoritmos otimizadores ADAM com taxa de aprendizado 0.001 e SGD com taxa de aprendizado 0.01. Os modelos foram construídos com duas camadas, a primeira com variações de 400, 500, 550, 600 ou 700 neurônios. Já a segunda camada sempre foi usada com 12 neurônios para os 12 meses de previsão. As funções de ativação utilizadas nas camadas foram ou Sigmoide ou ReLU e foi utilizado *early stopping* para interromper o treinamento

---

<sup>10</sup><http://www.ipeadata.gov.br>

<sup>11</sup><http://www.epe.gov.br>

<sup>12</sup><http://www.fazenda.rj.gov.br/sefaz>

e evitar sobre ajuste.

O menor MAPE em validação, para o cenário univariado, foi de 6.95% para o modelo com 550 neurônios, otimizador SGD e função de ativação ReLU. Para o cenário com duas variáveis exógenas, o menor MAPE foi de 6.17%, obtido com o modelo com 400 neurônios, otimizador Adam e função de ativação Sigmoide. Finalmente, o menor MAPE para o cenário com três variáveis exógenas, de 5.21%, foi resultado do modelo com 500 neurônios, otimizador Adam e função de ativação Sigmoide.

Os resultados, para cada cenário, obtidos no conjunto de testes estão expressos na tabela 2.4. O melhor modelo foi a rede RNN-LSTM com Exógenas com 3 variáveis.

Modelo	Erro Relativo (%)
SEFAZ-RJ	5.83
Rede LSTM Univariada	3.29
Rede LSTM com Exógenas ICMS e PIB	2.43
Rede LSTM com Exógenas ICMS, PIB e Energia Elétrica	1.65

Tabela 2.4: Erro Relativo da para a previsão da arrecadação anual do ICMS-RJ em 2018. O modelo SEFAZ-RJ se refere ao modelo, não informado pelo artigo, utilizado para a previsão oficial de arrecadação

Em 2022 a SEFAZ-RS publicou em seu site um experimento realizado por ([DORNELLES; SCHWARTZER; BRAATZ, 2022](#)) acerca do uso de redes neurais LSTM para a previsão da arrecadação do ICMS no Rio Grande do Sul. Os autores utilizaram a série univariada com as manipulações de *lags* e médias móveis reportadas por ([SILVA; FIGUEIREDO, 2020](#)) e a série do ICMS de janeiro de 2009 a dezembro de 2020 para treino e validação. As previsões foram feitas para os meses de janeiro a agosto de 2021. Porém foram feitas mês a mês, com o modelo sendo treinado novamente com os dados reais até o mês anterior para cada previsão.

A arquitetura de rede utilizada foi de duas camadas, com 80 neurônios LSTM em cada uma, com *dropout* de 5%. O tamanho do *batch* utilizado foi 5. O otimizador foi o Adam, com taxa de aprendizado padrão - 0.001 - e função de ativação padrão - Tangente Hiperbólica. Como resultados, os autores reportaram um erro acumulado, nos 8 meses de previsão, de -2.33% em comparação com um erro acumulado de -8.24% obtido com os métodos atuais de previsão da SEFAZ-RS, não mencionados no artigo.

### 2.2.3 Outros Experimentos de Previsão com Redes Neurais Recorrentes LSTM

Uma comparação entre modelos ARIMA e modelos RNN-LSTM na previsão de séries temporais foi publicada por (SIAMI-NAMINI; TAVAKOLI; NAMIN, 2018). Os autores utilizaram séries temporais obtidas no website do Yahoo Finance<sup>13</sup> de Janeiro de 1985 a agosto de 2018, incluindo o índice Nikkei 225 (N225), o índice composto da NASDAQ (IXIC), o índice Hang Seng (HIS), o índice de preços de commodities da S&P 500 (GSPC) e o índice médio industrial da Dow Jones (DF). Além disso, os autores obtiveram séries do Federal Reserve Bank of St. Louis<sup>14</sup> e do Fundo Monetário Internacional<sup>15</sup>, com dados de janeiro de 1967 a julho de 2017, acerca de serviços de saúde para consumidores urbanos (MC), habitação para consumidores urbanos (HO), peso da venda de dólares americanos frente a outras moedas estrangeiras (EX), consumo urbano de alimentos e bebidas (FB), estoque M1 de moeda nos EUA (MS) e transportes para consumidores urbanos (TR). Diferentemente as anteriores, a série MS vai de janeiro de 1959 a Julho de 2017, e a série TR vai de janeiro de 1947 a julho de 2017.

Todas as séries foram divididas em 70% para treinamento e 30% para testes, sendo que a maior delas tinha um total de 1698 observações, e a menor, 258 observações. A métrica utilizada para a comparação das séries foi o RMSE<sup>16</sup>, Raiz do Erro Quadrático Médio, que pode ser visto na equação 2.3.

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (2.3)$$

onde  $n$  é o número total de observações,  $y_i$  é o valor real da observação e  $\hat{y}_i$  é o valor predito.

Para previsão foi utilizada a técnica do *rolling forecast*, que tem como objetivo realizar a previsão do próximo mês fora da amostra, adicioná-lo ao conjunto de treinamento, treinar novamente o modelo, prever o próximo mês e assim por diante. Para todas as séries foi utilizado um modelo ARIMA e um modelo RNN-LSTM com uma camada com 4 neurônios LSTM. Os resultado obtidos, para cada série, podem ser vistos na tabela 2.5.

---

<sup>13</sup><https://finance.yahoo.com>

<sup>14</sup><https://fred.stlouisfed.org>

<sup>15</sup><http://www.imf.org/external/index.htm>

<sup>16</sup>*Root Mean Squared Error*

Os autores reportaram redução média de 87.44% no RMSE para as séries do mercado de ações; e redução média de 84,39% nas séries de dados econômicos.

Série	RMSE ARIMA	RMSE RNN-LSTM	% de Redução no RMSE
N225	766.45	105.31	-86.26
IXIIC	135.61	22.21	-83.62
HSI	1306.95	141.69	-89.16
GSPC	55.30	7.81	-85.87
DJ-Mensal	516.98	77.64	-84.98
DJ-Semanal	287.60	30.61	-89.36
<b>Média</b>	<b>511.48</b>	<b>64.21</b>	<b>-87.44</b>
MC	0.81	0.80	-1.11
HO	0.52	0.43	-17.62
EX	1.29	0.25	-80.48
FB	0.48	0.40	-16.94
MS	30.23	3.17	-89.51
TR	2.67	0.57	-78.70
<b>Média</b>	<b>6.00</b>	<b>0.94</b>	<b>-84.39</b>

Tabela 2.5: RMSE dos Modelos Arima e RNN-LSTM obtidos por ([SIAMI-NAMINI; TAVAKOLI; NAMIN, 2018](#))

Uma comparação do poder preditivo entre modelos de aprendizado de máquina em séries financeiras foi realizada por ([YAN; OUYANG, 2018](#)). Os autores utilizaram a série do preço diário de fechamento do Índice Composto da Bolsa de Xangai no período de 04 de janeiro de 2012 a 30 de junho de 2017. As 1347 observações foram divididas em 1077 para treinamento, 135 para validação e 135 para testes.

Os modelos utilizados foram o KNN<sup>17</sup>, ANN-MLP<sup>18</sup>, SVM<sup>19</sup> e RNN-LSTM. A métrica utilizada foi o MSE<sup>20</sup>, Erro Quadrático Médio, que pode ser visto na equação 2.4.

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

onde  $n$  é o número total de observações,  $y_i$  é o valor real da observação e  $\hat{y}_i$  é o valor predito.

---

<sup>17</sup> *K-Nearest Neighbors*, ou K-Vizinhos mais Próximos

<sup>18</sup> *Artificial Neural Network-Multi Layer Perceptron*, ou Rede Neural Artificial com Perceptron Multi Camada

<sup>19</sup> *Support Vector Machines*, ou Máquinas de Vetores de Suporte

<sup>20</sup> *Mean Squared Error*

Os resultados obtidos, na previsão do conjunto de treinamento, foram superiores para a RNN-LSTM, ANN-MLP, SVM e KNN, como pode ser visto na tabela 2.6.

Modelo	MSE
RNN-LSTM	0.0030
ANN-MLP	0.0040
SVM	0.0113
KNN	0.0465

Tabela 2.6: MSE dos Modelos de (YAN; OUYANG, 2018)

A previsão de séries do mercado de ações brasileiro, americano e sueco foi feita por (HANSSON, 2017) utilizando RNN-LSTM. O autor utilizou as séries S&P 500 Index, Bovespa 50 Index e OMX Index, de 02 de janeiro de 2009 a 28 de abril de 2017, com aproximadamente 2100 observações em cada uma. Os modelos de rede neural utilizados foram dois. Um modelo RNN-LSTM de camada única, denominado LSTM pelo autor, com 4 neurônios LSTM. E um modelo de 3 camadas escondidas, denominado *Deep LSTM*, com 4 neurônios LSTM na primeira camada, 40 na segunda e 100 na terceira. Além disso o autor utilizou um modelo linear ARMA(1,1)-GJRGARCH(1,1)<sup>21</sup>. A métrica utilizada foi o RSS<sup>22</sup>, Soma Residual de Quadrados, expressa pela equação 2.5.

$$RSS = \sum_{i=1}^n (y_i - \hat{y}_i)^2 \quad (2.5)$$

onde  $n$  é o número total de observações,  $y_i$  é o valor real da observação e  $\hat{y}_i$  é o valor predito.

A tabela 2.7 apresenta os resultados obtidos. A previsão realizada foi de um dia à frente. Os resultados dos três modelos são muito similares, tendo o autor concluído que redes LSTM não são um bom preditor para os retornos do mercado de ações.

Modelo	RSS SP500	RSS Bovespa	RSS OMX
ARMA-GJRGARCH	0.0008	0.0015	0.0019
LSTM	0.0009	0.0015	0.0019
Deep LSTM	0.0090	0.0015	0.0019

Tabela 2.7: RSS dos Modelos de (HANSSON, 2017)

<sup>21</sup>Glosten-Jagannathan-Runkle GARCH

<sup>22</sup>Residual Sum of Squares

Um trabalho comparando os resultados de previsão da inflação americana foi publicado por ([ALMOSOVA; ANDRESEN](#), ). Os autores utilizaram a série mensal do Índice de Preços ao Consumidor do Estados Unidos, US CPI, obtida no banco de dados FRED<sup>23</sup> de janeiro de 1960 a junho de 2020, e realizaram a previsão de 12 meses, mês a mês, com retreino do modelo.

Os modelos utilizados foram um Passeio Aleatório a título de modelo naïve, um modelo autoregressivo, uma rede neural MLP com uma camada, denominada NN, uma rede neural recorrente LSTM, um modelo SARIMA e um modelo de mudança markoviana autoregressivo denominado MS-AR.

Após a análise dos resultados, os autores concluíram que a rede neural recorrente LSTM foi ligeiramente mais eficiente do que os modelos AR, NN e MS-AR. Porém, o modelo SARIMA e o modelo LSTM tiveram desempenho similar nos experimentos realizados.

Sumarizando, os resultados apresentados por ([SILVA; FIGUEIREDO, 2020](#)) e por ([DORNELAS; CAMPOS; FIGUEIREDO, 2022](#)) são corroborados pelos trabalhos de ([SIAMI-NAMINI; TAVAKOLI; NAMIN, 2018](#)) e ([YAN; OUYANG, 2018](#)), mas são contrapostos pelos trabalhos de ([HANSSON, 2017](#)) e ([ALMOSOVA; ANDRESEN,](#) ) na previsão de séries financeiras e econômicas.

O uso de *ensembles* pode melhorar a previsão de redes LSTM. Em seu trabalho, ([BOROVKOVA; TSIAMAS, 2019](#)) utilizaram um *ensemble* de redes LSTM para reduzir o sobre ajuste na previsão de séries de tempo multivariadas. Entretanto, esta abordagem não foi avaliada neste trabalho.

## 2.3 Modelos de Mudança Markoviana

Um importante uso dos modelos de mudança markoviana é datação de ciclos econômicos. ([CHAUVET, 2002](#)) utilizou um modelo Markov-Switching para fazer a previsão dos ciclos econômicos brasileiros, a partir dos dados do PIB com frequência anual de 1900 a 1999, obtidos nos bancos de dados da Fundação Getulio Vargas (FGV) e dos dados do PIB real quadrimestral do primeiro trimestre de 1980 até o primeiro trimestre de 2000, obtida no IBGE.

---

<sup>23</sup>*Federal Reserve Bank of St. Louis*

De modo a capturar a variação cíclica da economia brasileira, a autora utilizou o log do PIB real,  $\tilde{y}_t$ , modelado como a soma de dois componentes, conforme descrito por ([HAMILTON, 1989](#)): um termo de tendência markoviana  $\tilde{n}_t$  e um componente gaussiano,  $\tilde{z}_t$ :

$$\tilde{y}_t = \tilde{n}_t + \tilde{z}_t \quad (2.6)$$

A tendência markoviana é determinada por:

$$\tilde{n}_t = \tilde{n}_{t-1} + \mu_{st} \quad (2.7)$$

onde  $st$  é uma cadeia de Markov não observável de primeira ordem com dois estados e  $\mu_{st}$  é o termo de *drift* dependente de estado. O termo *drift*  $\mu_{st}$  assume o valor  $\mu_0$  quando a economia está em uma fase de baixo crescimento - ou recessão ( $s_t = 0$ ) - e  $\mu_1$  quando a economia está crescendo, ou seja, em expansão ( $s_t = 1$ ). A mudança de um estado para outro depende unicamente do estado anterior.

Estes mudanças são governadas por uma matriz de probabilidades  $P_2$  com elementos  $p_{ij} = pr[s_t = j | s_{t-1} = i]$ , onde  $i$  denota a  $i$ -ésima coluna e  $j$  a  $j$ -ésima linha. Cada coluna de  $P_2$  soma um, de modo que  $1'_2 P_2 = 1'_2$ , onde  $1'_2$  é um vetor coluna de uns.

O componente gaussiano segue um processo ARIMA( $r, 1, 0$ ) com média zero:

$$\tilde{z}_t = \tilde{z}_{t-1} + \phi_1(\tilde{z}_{t-1} - \tilde{z}_{t-2}) + \cdots + \phi_r(\tilde{z}_{t-r} - \tilde{z}_{t-r-1}) + \epsilon_t \quad (2.8)$$

onde  $\epsilon_t \sim iidN(0, \sigma^2)$  e  $\epsilon_t$  é independente em  $n_{t+j}$ ,  $\forall j$ .

Para lidar com as quebras estruturais da economia brasileira, como por exemplo os planos de estabilização das décadas de 80 e 90, a autora propôs lidar com a mudança do log do PIB ( $y_t$ ) como uma cadeia de Markov oculta e configurar o processo autorregressivo da equação [2.8](#) igual a zero ( $r = 0$ ). Deste modo o log do PIB brasileiro,  $\tilde{y}_t$ , é função de um processo integrado que segue uma cadeia de Markov e um processo de ruído branco.

O modelo é:

$$\begin{aligned} \tilde{y}_t &= \tilde{n}_t + \tilde{z}_t \\ \tilde{n}_t &= \tilde{n}_{t-1} + \mu_{st} \\ \tilde{z}_t &= \tilde{z}_{t-1} + \epsilon_t \end{aligned} \quad (2.9)$$

Tirando a primeira diferença de 2.9 tem-se:

$$y_t = \mu_{st} + \epsilon_t \quad (2.10)$$

A datação de ciclos de crescimento e de retração encontrados pela autora podem ser vistos na figura 2.2.

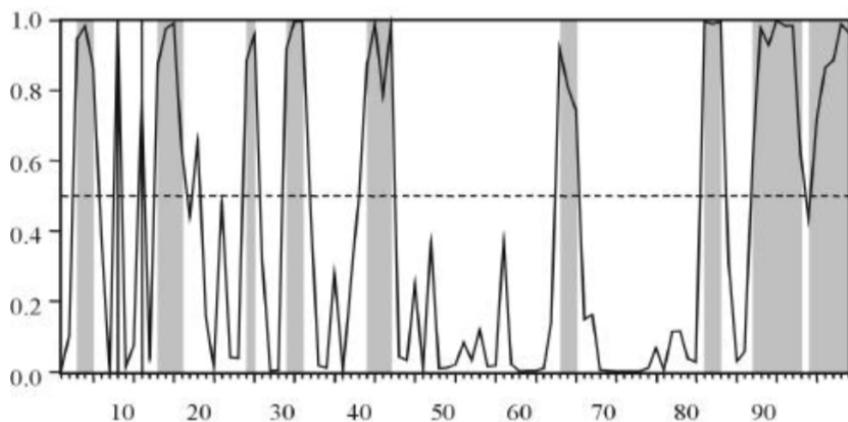


Figura 2.2: Probabilidade suavizada de retracções (áreas acinzentadas) e expansões dos ciclos de negócios brasileiros. Dados anuais de 1900 a 1999.

A autora compara o poder preditivo fora da amostra do modelo Markov-Switching com modelos lineares autorregressivos, mais especificamente um ARMA(1,1) e um AR(5). Ela conclui que o modelo Markov-Switching tem RMS<sup>24</sup> e MAE<sup>25</sup> menores do que dos modelos lineares, além de possuir um coeficiente de determinação ajustado  $\bar{R}^2$  bem maior, conforme resultados apresentados na figura 2.8. Analogamente, (DOORNIK, 2013) utiliza o modelo Markov-Switch para a datação de ciclos do PIB norte-americano, mas não chega a fazer previsões em seu artigo. Entretanto, ambos os artigos confirmam os resultados obtidos por (HAMILTON, 1989) no uso do modelo Markov-Switching para datação de ciclos econômicos. Mais ainda, Doornik estende o modelo Markov-Switching criando um modelo com variação independente da média e da variância, de modo que o modelo proposto por Hamilton passe a ter validade para o PIB americano no também no período que vai de 1984 a 2007, conhecido por *Grande Moderação*.

Alguns modelos não lineares foram avaliados quanto ao seu poder de previsão da série mensal da produção industrial brasileira, PIM, série publicada pelo IBGE (ROCHA; PEREIRA, 2015). Os autores utilizaram dados mensais da PIM de janeiro de 1975 a

<sup>24</sup>Root mean squared error, ou Raiz do erro quadrático médio

<sup>25</sup>Mean absolute error, ou Erro absoluto médio

Desempenho da Previsão	Markov-Switching AR(0)	ARMA(1,1)	AR(5)
$\bar{R}^2$	0.714	0.118	0.224
RMS	0.996	2.674	2.107
MAE	0.759	1.939	1.563

Tabela 2.8: Resultados da previsão do PIB-Brasil, fora da amostra, do 2º Tri de 1989 ao 1º Tri de 2000

fevereiro de 2014. O horizonte de previsão foi de setembro de 2008 a fevereiro de 2014, num total de 66 períodos, com previsões um mês à frente e reestimativa do modelo a cada previsão.

Os modelos utilizados foram o de mudança markoviana autorregressivo, Markov-Switching Autoregressive (MSAR), o LSTAR, *Logistic Smooth Transition Autoregressive Model* e o modelo *Autometrics* com Saturação de Dummies (DSAT), que foi usado também com uma variação com apenas dummies degrau (*step dummies*), o DSAT-step. E dois modelos naïve, o DD12, onde a série foi diferenciada com 12 meses de diferença, isto é, cada mês é diferenciado com o mesmo mês do ano anterior; e um AR(p), onde p é o número de defasagens encontrado pelo critério de informação de Schwarz a cada treinamento do modelo.

O objetivo do trabalho é o de testar o poder preditivo destes modelos não lineares frente à quebras estruturais das séries. A quebra escolhida, determinante do período de previsão selecionado, foi a Crise Financeira do Subprime de 2008. Para comparar os modelos foi usado o teste de Diebold & Mariano, ver (DIEBOLD; MARIANO, 2002).

Como resultado, os autores reportam que os modelos com melhor desempenho foram os modelos naïve lineares. O modelo que obteve o melhor MAPE foi o *Autometrics* com Saturação de Dummies com apenas dummies degrau, o DSAT-step, como pode ser visto na tabela 2.9. Mas o teste de Diebold & Mariano mostrou que o modelo com melhor resultado geral foi o AR(p).

Um trabalho acerca da previsão do preço do ouro no Canadá foi publicado por (NGUYEN, 2020). O autor obteve a série mensal da cotação do ouro no Canadá, em dólares americanos (USD) e dólares canadenses (CAD), de janeiro de 1979 a maio de 2020, no website do *World Gold Council*<sup>26</sup>. O autor realizou previsões de um mês utilizando três modelos:

---

<sup>26</sup><https://www.gold.org/goldhub/data/gold-prices>

Modelo	MAPE %
LSTAR	100.033
MSAR-BIC	109.010
DSAT	104.760
DSAT-step	89.327
AR(p)	93.937
DD12	118.961

Tabela 2.9: Resultados da previsão obtidos por (ROCHA; PEREIRA, 2015)

- um modelo explanatório, uma regressão linear simples utilizando o preço do ouro em USD como variável explicada e a inflação do Canadá, expressa pelo CPI<sup>27</sup> como variável explicativa;
- um modelo autorregressivo AR(1);
- um modelo de mudança markoviana autorregressivo, MS-AR, com dois regimes, um sem o uso de variáveis exógenas, e outro regime utilizando o preço do ouro em dólares americanos, o preço em dólares canadenses e o índice de inflação CPI.

O autor obteve o menor MAPE com o modelo MS-AR, como pode ser visto na tabela 2.10.

Modelo	MAPE %
Explanatório	7.31
AR(1)	2.99
MS-AR	2.61

Tabela 2.10: Resultados da previsão obtidos por (NGUYEN, 2020)

Outros trabalhos, cujo objetivo não era o de realizar previsões, aplicaram o modelo Markov-Switching na alocação de portfólios com mudança de regime (OLIVEIRA; PEREIRA, 2018) ou utilizaram este modelo para estudar o contágio financeiro entre retornos do mercado de ações (ROTTA; PEREIRA, 2016) sob o modelo econômico RSDC<sup>28</sup> de (PELLETIER, 2006), que decompõe as covariâncias entre múltiplas séries temporais em correlações e desvios-padrão, e faz com que a matriz de correlação siga um modelo de alternância de regimes (SILVA, 2013), mostrando mais uma vez o poder do modelo Markov-Switching na datação de ciclos.

<sup>27</sup>Consumer Price Index, disponível na base do FRED

<sup>28</sup>Regime switching dynamic correlation, ou Correlação Dinâmica com Mudança de Regime

## 2.4 Avaliação de Modelos

A questão da avaliação de modelos de previsão é antiga na Econometria. ([ELLIOTT; TIMMERMANN, 2008](#)) propõem um arcabouço para avaliação de modelos a partir de suas funções de perda. Neste mesmo sentido, ([DIEBOLD; MARIANO, 2002](#)) propõem uma técnica para avaliação de modelos dois a dois.

Já ([HANSEN; LUNDE; NASON, 2011](#)) definiram um processo para determinar, dentro de um conjunto de modelos, qual é o melhor modelo, em termos de critérios especificados pelo usuário por meio de uma função de perda escolhida. A este processo denominaram MCS, *Model Confidence Set*.

Este procedimento de escolha de melhores modelos foi utilizado por ([VIDAL; KRIST-JANPOLLER, 2020](#)), que utilizou Redes Neurais LSTM para prever a volatilidade do ouro. Diversos modelos de previsão de volatilidade foram ranqueados por ([GONZÁLEZ-RIVERA; LEE; MISHRA, 2004](#)) utilizando o MCS. Modelos para a previsão da produção industrial agregada e desagregada foram avaliados por ([PRINCE; MARÇAL; PEREIRA, 2022](#)) com o uso do MCS.

# Capítulo 3

## Metodologia

### 3.1 Fundamentação Teórica

Esta seção irá apresentar um resumo dos modelos lineares ARIMA, SARIMA e suas versões com variáveis exógenas, o ARIMAX e o SARIMAX. Será apresentada também uma visão geral dos modelos não lineares de mudança markoviana, o Markov-Switching AR e o Markov-Switching VAR. Além disso, as Redes Neurais Artificiais, as Redes Neurais Recorrentes e os neurônios LSTM serão aprofundados.

#### 3.1.1 Modelos ARIMA

Um resumo dos modelos ARIMA é apresentado por ([SIAMI-NAMINI; TAVAKOLI; NAMINI, 2018](#)). Para uma visão mais detalhada, veja ([ENDERS, 2014](#)) e ([HYNDMAN; ATHANASOPOULOS, 2018](#)).

Os modelos ARIMA, *Autorregressive Integrated Moving Average*<sup>1</sup>, são uma generalização dos modelos ARMA, *Autorregressive Moving Average*<sup>2</sup>, que combinaram processos Autorregressivos (AR) com processos de Média Móvel (MA) e criam um modelo composto para séries temporais. Assim, um modelo  $ARIMA(p, d, q)$  captura os elementos chave do modelo:

- *AR*: autorregressão. Um modelo de regressão que usa a dependência entre uma observação e um número de observações defasadas (*lags*) ( $p$ ).

---

<sup>1</sup>Média Móvel Integrada Autorregressiva

<sup>2</sup>Média Móvel Autorregressiva

- I: integração. Para tornar as séries estacionárias tirando as diferenças entre cada observação e sua observação anterior ( $d$ ) vezes.
- MA: média móvel. Uma abordagem que leva em conta a dependência entre observações e o termo de erro residual quando o modelo de média móvel é usado nas observações defasadas (*lags*) ( $q$ ).

Uma representação simples de um modelo AR de ordem  $p$ , i. e.,  $AR(p)$ , pode ser escrita como uma equação linear dada por:

$$y_t = c + \sum_{i=1}^p \phi_i y_{t-i} + \epsilon_t \quad (3.1)$$

onde  $y_t$  é a variável estacionária,  $c$  é constante, os termos  $\phi_i$  são coeficientes de autocorrelação para os lags  $1, 2, \dots, p$  e  $\epsilon_t$  são os resíduos, que devem ser um ruído branco gaussiano com média zero e variância constante  $\sigma^2$ , i.e.,  $\sim N(0, \sigma^2)$ .

Um modelo MA de ordem  $q$ , i.e.,  $MA(q)$ , pode ser escrito na forma:

$$y_t = \mu + \sum_{i=1}^q \theta_i \epsilon_{t-i} \quad (3.2)$$

onde  $\mu$  é a esperança de  $y_t$  (geralmente assumida como zero), os termos  $\theta_i$  são os pesos aplicados aos valores atuais e anteriores dos termos estocásticos da série e  $\theta_0 = 1$ . Novamente, assume-se que  $\epsilon_t$  são um ruído branco gaussiano com média zero e variância constante  $\sigma^2$ , i.e.,  $\sim N(0, \sigma^2)$ .

Combinando-se os dois modelos tem-se o ARMA de ordem  $(p,q)$ :

$$y_t = c + \sum_{i=1}^p \phi_i y_{t-i} + \epsilon_t + \sum_{i=1}^q \theta_i \epsilon_{t-i} \quad (3.3)$$

onde  $\phi_i \neq 0$ ,  $\theta_i \neq 0$  e  $\sigma_\epsilon^2 > 0$ . Os parâmetros  $p$  e  $q$  são as ordens do *AR* e do *MA*, respectivamente.

Um ponto extremamente importante para a estabilidade dos modelos ARIMA é a estacionariedade. Uma condição necessária, mas não suficiente, é ter o parâmetro  $\phi_i$  do componente autorregressivo, ou seja, as raízes do polinômio dentro do círculo de raiz unitária, i.e.,  $|\phi_i| < 1$ .

O modelo ARIMA é capaz de lidar com séries não estacionárias devido ao seu componente integrado. De fato, o componente integrado envolve a diferenciação da série para

converter a série não estacionária em estacionária. Assim, a forma geral do modelo é  $ARIMA(p,d,q)$ , com a equação:

$$y_t' = c + \sum_{i=1}^p \phi_i y_{t-i}' + \epsilon_t + \sum_{i=1}^q \theta_i \epsilon_{t-i} \quad (3.4)$$

onde  $y_t'$  é a série diferenciada (e que pode ser diferenciada mais de uma vez).

## Modelos SARIMA

Os dados das séries de tempo podem ser carregados de componentes sazonais, que afetam a estacionariedade e interferem nas previsões. Por este motivo é necessário o uso de um modelo ARIMA sazonal, ou SARIMA (*Seasonal ARIMA*), que incorpora tanto fatores sazonais quanto não sazonais em um modelo multiplicativo. A forma geral do modelo SARIMA é denotada por  $ARIMA(p,d,q) \times (P,D,Q)S$ , onde  $p$  é a ordem do *AR* não sazonal,  $d$  é a diferenciação não sazonal,  $q$  é a ordem *MA* não sazonal,  $P$  é a ordem do *AR* sazonal,  $D$  é a diferenciação sazonal,  $Q$  é a ordem *MA* não sazonal e  $S$  é o tempo de repetição do padrão sazonal.

Diferenciação sazonal ([HYNDMAN; ATHANASOPOULOS, 2018](#)) é a diferença entre uma observação e a observação prévia do mesmo período sazonal, ou seja:

$$y_t' = y_t - y_{t-s} \quad (3.5)$$

onde  $s$  é o número de períodos sazonais. A equação do modelo SARIMA, então, é:

$$y_t' = c + \sum_{i=1}^p \phi_i y_{t-i}' + \sum_{i=1}^q \theta_i \epsilon_{t-i} + \sum_{i=1}^P \alpha_i y_{t-si}' + \sum_{i=1}^Q \eta_i \epsilon_{t-si} + \epsilon_t \quad (3.6)$$

onde  $\alpha_i$  é o *AR* sazonal e  $\eta_i$  são os pesos do *MA* sazonal e  $y_t'$  é a série diferenciada. Note que a série pode ser diferenciada mais de uma vez e ter diferenciações diferentes para a série e para o componente sazonal, isto é,  $d \neq D$  pode acontecer.

O passo mais importante na estimativa de um modelo SARIMA é identificar os valores de  $(p,d,q)$  e  $(P,D,Q)$ . Com base no gráfico da série pelo tempo, por exemplo, nota-se que a variância cresce com o tempo, e deveriam ser usadas transformações como o uso de defasagens e diferenciação para estabilizar a variância. Então, pelo método de Box and

Jenkins, ver ([BOX et al., 2015](#)), deve-se utilizar a função de autocorrelação (ACF<sup>3</sup>) para medir o quanto linearmente dependentes são duas observações da série separadas por um lag  $p$  e a função de autocorrelação parcial (PACF<sup>4</sup>) para determinar quantos termos autoregressivos  $q$  são necessários.

## SARIMAX

O modelo SARIMAX leva em conta os componentes sazonais, como o SARIMA, mas também considera variáveis exógenas no modelo como regressoras. O modelo SARIMAX pode ser representado por:

$$y'_t = c + \sum_{i=1}^p \phi_i y'_{t-i} + \sum_{i=1}^q \theta_i \epsilon_{t-i} + \sum_{i=1}^r \beta_i x_{i_t} + \sum_{i=1}^P \alpha_i y'_{t-si} + \sum_{i=1}^Q \eta_i \epsilon_{t-si} + \epsilon_t \quad (3.7)$$

onde os  $\beta_i$  são os coeficientes da matriz de variáveis exógenas  $X$ .

## Auto-Arima

Uma técnica eficiente para encontrar os melhores parâmetros  $(p,d,q)$  e  $(P,D,Q)$  de um modelo SARIMA é a função Auto-Arima, disponível no software estatístico EViews, no pacote *forecast*([HYNDMAN; KHANDAKAR, 2008](#)) do R e também na biblioteca *pmdarima*([SMITH et al., 2017](#)) do Python. Alguns trabalhos reportam o uso bem sucedido desta função, como ([YERMAL; BALASUBRAMANIAN, 2017](#)), que usaram esta função no EViews, ([KHAN; ALGHULAIAKH, 2020](#)) que a utilizaram no R e ([ARUNKUMAR et al., 2021](#)) que utilizaram a da *pmdarima*.

Em suma, esta função executa, de modo iterativo, o cálculo dos modelos com diversas variações de defasagens e integrações  $(p,d,q)$  e  $(P,D,Q)$  e, mediante algum critério de informação, retorna o melhor modelo. Os critérios de informação geralmente utilizados são o Critério de Informação de Akaike (AIC) e o Critério de Informação de Schwarz (SIC), também bastante conhecido como Critério de Informação Bayesiano (BIC).

---

<sup>3</sup>Autocorrelation Function

<sup>4</sup>Partial Autocorrelation Function

## Critérios de Informação AIC e BIC

Os critérios de informação mostram quão bom é o modelo em explicar a relação entre as variáveis. Tanto o AIC quanto o BIC são critérios de informação que avaliam a qualidade do modelo por darem crédito aos modelos com menor erro enquanto aplicam penalidades nos modelos com muitos parâmetros.

O AIC é representado pela equação 3.8:

$$AIC = -2 \log L(\hat{\theta}) + 2K \quad (3.8)$$

onde  $\log L(\hat{\theta})$  representa a função de log verossimilhança e K é o número total de parâmetros do modelo.

O BIC é outro critério similar ao AIC, mas impõe uma penalidade menor no número de parâmetros do que o AIC. Em ambos o menor valor de log verossimilhança representam o melhor modelo que tem o maior valor de log verossimilhança. O BIC é representado matematicamente por 3.9.

$$BIC = -2 \log L(\hat{\theta}) + K \log N \quad (3.9)$$

onde N é o número de observações. Como  $\log N$  sempre será maior que 2 (afinal, o número de observações teria que ser menor ou igual a 7 para que isso não acontecesse), o critério de Schwarz (BIC) sempre irá selecionar um modelo mais parcimonioso. Para maiores detalhes, ver ([ENDERS, 2014](#)).

## Testes de Estacionariedade

Há duas maneiras de eliminar a tendência. Uma série estacionária de tendência pode ser transformada em uma série estacionária removendo-se a tendência determinística. As séries com raiz unitária, algumas vezes chamadas de séries estacionárias de diferença, podem ser transformadas em uma série estacionária por diferenciação. Utilizar a técnica incorreta para tornar a série estacionária pode causar problemas séries, como afirma ([ENDERS, 2014](#)).

Para se determinar como proceder, podem ser feitos dois testes de raiz unitária, o ADF e o KPSS.

**Teste de raiz unitária Dickey-Fuller Aumentado (ADF)** O teste de raiz unitária geralmente utilizado é o proposto por ([DICKEY; FULLER, 1979](#)), cujas hipóteses nula e alternativa estão descritas a seguir:

$$H_0 : \text{a série tem uma raiz unitária.}$$

$$H_1 : \text{a série não tem raiz unitária.}$$

**Teste de raiz unitária Kwiatkowski-Phillips-Schmidt-Shin (KPSS)** Um procedimento alternativo para o teste de estacionariedade foi proposto por ([KWIATKOWSKI et al., 1992](#)). Eles criaram um teste cuja hipótese nula é a inexistência de raiz unitária, i. e., a série é estacionária. Ver ([KIRCHGÄSSNER; WOLTERS; HASSLER, 2012](#)) para detalhes. As hipóteses do teste são:

$$H_0 : \text{a série é estacionária de tendência.}$$

$$H_1 : \text{a série não é estacionária.}$$

**Interpretação dos testes** A combinação dos dois testes pode ter os seguintes resultados com as seguintes consequências:

- Caso 1: Ambos os testes concluem que a série não é estacionária, então a série não é estacionária
- Caso 2: Ambos os testes concluem que a série é estacionária, então a série é estacionária
- Caso 3: o KPSS indica estacionariedade e o ADF indica não estacionariedade; a série é estacionária de tendência. A tendência precisa ser removida para tornar a série estritamente estacionária. A série sem tendência deve ser verificada quanto à estacionariedade.
- Caso 4: KPSS indica não estacionariedade e ADF indica estacionariedade; a série é estacionária de diferença. A diferenciação deve ser usada para torná-la estacionária. A série diferenciada deve ser verificada quanto à estacionariedade.

Nos casos 3 e 4 a série precisa ter sua estacionariedade verificada novamente pelos dois testes e, caso ainda não seja estacionária, deve passar pelo método de diferenciação ou remoção de tendência conforme o processo descrito acima.

### 3.1.2 Modelos de Mudança Markoviana

O conteúdo desta seção está grandemente baseado no trabalho de (DOORNIK, 2018). Mais detalhes sobre os Markov-Switching Models podem ser obtidos em (ENDERS, 2014).

Os modelos de mudança de regime permitem que alguma parte do modelo dependa do estado da economia (o regime). Por exemplo, a média da regressão do modelo pode ser diferente quando há recessões e expansões.

Nos modelos de mudança markoviana, a especificação dentro de cada regime é linear. E probabilidades de transição que regem como o modelo se move de um regime para o outro são estimadas. Assim, a série de tempo resultante é não linear.

A título de motivação do problema de mudanças de regime, considere os gráficos que estão na figura 3.1, que mostram o PIB trimestral dos EUA em log ao longo do tempo e também sua taxa de crescimento, a primeira diferença dos logs:

$$\Delta y_t = \log(Y_t) - \log(Y_{t-1})$$

As áreas sombreadas são as recessões determinadas pelo comitê de datação de ciclos do NBER<sup>5</sup>. O pico mais recente apresentado é foi em dezembro de 2007, seguido pela recessão associada à crise financeira indo até junho de 2009.

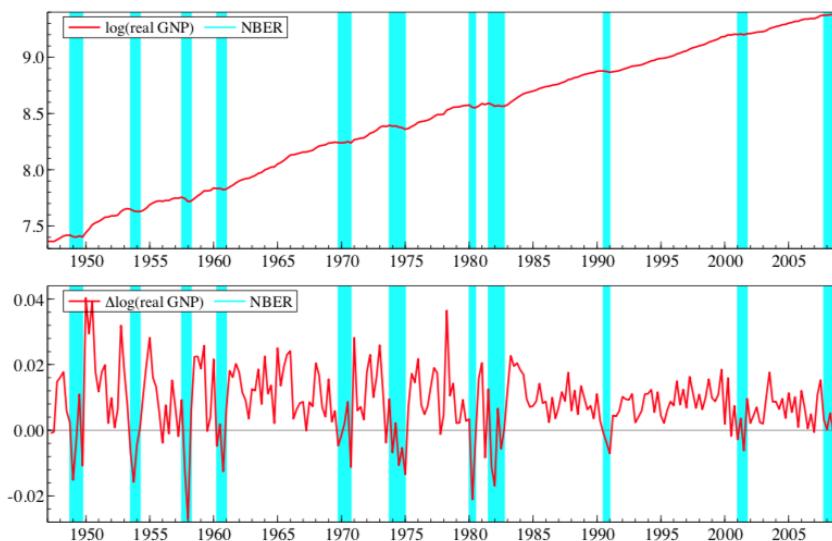


Figura 3.1: PIB real trimestral dos EUA - em log, no painel de cima, e em taxa de crescimento, no painel de baixo - com a datação dos ciclos de retração (sombreados) e crescimento feitas pela NBER.

O objetivo dos modelos de mudança de regime é permitir diferentes comportamentos

<sup>5</sup> National Bureau of Economic Research - [www.nber.org](http://www.nber.org)

em diferentes estados da natureza, e ao mesmo tempo estimar quando há a transição de um estado para o outro. Deste modo, ao permitir que a taxa de crescimento da média em uma recessão seja diferente do normal, é possível classificar a economia com estando em recessão ou em um crescimento normal: um modelo de dois regimes compatível com o ciclo de datação da NBER. Mas médias diferentes podem não ser o bastante, é preciso também ter uma dinâmica diferente entre os dois regimes. Ou talvez três regimes seriam necessários para uma descrição satisfatória da série.

Um modelo simples de mudança de regime seria:

$$\begin{aligned} \text{Regime 0 : } & y_t = \mu(0) + \rho y_{t-1} + \epsilon_t \\ \text{Regime 1 : } & y_t = \mu(1) + \rho y_{t-1} + \epsilon_t \\ & \epsilon_t \sim N(0, \sigma^2) \end{aligned} \tag{3.10}$$

É claro que o número de regimes é arbitrário. Escrevendo  $S_t$  para a variável que denota o regime, então a média pode ser escrita como uma função de  $S_t$ :

$$\mu(S_t) = \begin{cases} 1 & \text{if } S_t = 0 \text{ recessão} \\ 2 & \text{if } S_t = 1 \text{ expansão} \end{cases} \tag{3.11}$$

A interpretação mostra a arbitrariedade da numeração: só se pode chamar o regime 0 de recessão se  $\mu_0 < \mu_1$ . Caso contrário é necessário mudar a classificação.

**Probabilidades de Transição** No modelo de mudança markoviana, a variável aleatória não observável  $S_t$  segue uma cadeia de Markov, definida pela probabilidade de transição entre os estados  $S$  (ou regimes):

$$p_{i|j} = P[S_{t+1} = i | S_t = j], \quad i, j = 0, \dots, S - 1 \tag{3.12}$$

Então a probabilidade de se mover do estado  $j$  em um período para o estado  $i$  no próximo apenas depende do estado anterior: não há benefício em se saber a história completa do regime:

$$P[S_{t+1} = i | S_t = j, S_{t-1}, S_{t-2}, \dots] = P[S_{t+1} = i | S_t = j] \tag{3.13}$$

Como o sistema tem que estar em um dos estados  $S$  tem-se:

$$\sum_{i=0}^{S-1} p_{i|j} = 1 \quad (3.14)$$

A matriz  $S \times S$  de probabilidades de transição é denotada por  $\mathbf{P}$ :

$$\mathbf{P} = (p_{i|j}) \quad (3.15)$$

com probabilidades condicionais somando 1 nas colunas. Por exemplo, para  $S = 2$ :

$$\mathbf{P} = \left( \begin{array}{c|cc} & S_t = 0 & S_t = 1 \\ \hline S_{t+1} = 0 & p_{0|0} & p_{0|1} \\ S_{t+1} = 1 & p_{1|0} & p_{1|1} \\ \hline \Sigma & 1 & 1 \end{array} \right)$$

Probabilidades redundantes podem ser substituídas. Por exemplo:

$$\mathbf{P} = \left( \begin{array}{c|cc} & S_t = 0 & S_t = 1 \\ \hline S_{t+1} = 0 & p_{0|0} & 1 - p_{1|1} \\ S_{t+1} = 1 & 1 - p_{0|0} & p_{1|1} \\ \hline \Sigma & 1 & 1 \end{array} \right)$$

onde  $0 \leq p_{0|0}, p_{1|1} \leq 1$ .

**MS-AR e MS-DR: dois tipos de mudança markoviana na média** Distingue-se entre dois tipos de modelos de mudança markoviana: os modelos de mudança markoviana de regressão dinâmica (MS-DR) e os modelos de mudança markoviana autorregressivos.

A especificação do MS-DR segue o modelo de regressão dinâmica na especificação das dinâmicas. Um exemplo simples sem regressores é:

$$y_t = \nu(S_t) + \alpha y_{t-1} + \epsilon_t, \quad \epsilon_t \sim IIN[0, \sigma^2] \quad (3.16)$$

O modelo MS-AR relacionado é escrito como:

$$y_t - \mu(S_t) = \rho[y_{t-1} - \mu(S_{t-1})] + \epsilon_t, \quad \epsilon_t \sim IIN[0, \sigma^2] \quad (3.17)$$

Sem a mudança de regime ambas as especificações são idênticas:  $\nu = (1 - \rho)\mu$ ,  $\alpha = \rho$  e podem ser escritas uma como a outra. Mas este não é o caso para modelos de mudança markoviana: o intercepto no modelo MS-AR depende dos últimos dois regimes:

$$\mu(S_t) - \rho\mu(S_{t-1}) \quad (3.18)$$

Isto dá quatro possíveis valores para o intercepto:

	$S_{t-1} = 0$	$S_{t-1} = 1$
$S_t = 0$	$\mu(0) - \rho\mu(0)$	$\mu(0) - \rho\mu(1)$
$S_t = 1$	$\mu(1) - \rho\mu(0)$	$\mu(1) - \rho\mu(1)$

Um modo de lidar com isso é recodificar o regime de dois períodos de classificação em estado de 4 períodos  $N_t$ :

	$S_{t-1} = 0$	$S_{t-1} = 1$
$S_t = 0$	$N + t = 0$	$N + t = 1$
$S_t = 1$	$N + t = 2$	$N + t = 3$

Em geral, o modelo MS-AR( $p$ ) requer um vetor de estados de dimensão  $N = S^{(1+p)}$  para obter uma representação de Markov com avaliação de verossimilhança para  $S$  regimes e autorregressão de ordem  $p$ . Portanto, o modelo MS-AR é efetivamente infactível para dados mensais ou autoregressões de alta ordem. O modelo MS-DR não tem este problema, pois o número de estados é igual ao número de regimes:  $N = S$ .

A diferença conceitual entre os modelos MS-DR e MS-AR é que o MS-DR se ajusta imediatamente ao novo regime, enquanto o MS-AR se ajusta gradualmente.

Regressores e defasagens adicionais podem ser adicionados a ambas especificações. Um exemplo de um modelo MS-DR( $p$ ) é:

$$y_t = \nu(S_t) + \sum_{i=1}^p \alpha_i y_{t-p} + \mathbf{x}'_t \beta + \epsilon_t, \quad \epsilon_t \sim IIN[0, \sigma^2] \quad (3.19)$$

O modelo MS-AR relacionado é escrito como:

$$y_t - \mu(S_t) + \mathbf{x}'_t \gamma = \sum_{i=1}^p \rho_i (y_{t-i} - \mu(S_{t-i}) - \mathbf{x}'_{t-i} \gamma) + \epsilon_t, \quad \epsilon_t \sim IIN[0, \sigma^2] \quad (3.20)$$

**Heterocedasticidade dos modelos de mudança markoviana** Os modelos de mudança markoviana não apresentam problemas diante do erro de variância (heterocedasticidade). Existirão  $S$  termos de variância em vez de apenas um. Por exemplo, para um MS-DR simples com dois regimes:

$$\begin{aligned} \text{Regime 0 : } y_t &= \mu(0) + \rho y_{t-1} + \sigma(0)\epsilon_t, & \epsilon_t &\sim N[0,1] \\ \text{Regime 1 : } y_t &= \mu(1) + \rho y_{t-1} + \sigma(1)\epsilon_t, & \epsilon_t &\sim N[0,1] \end{aligned} \quad (3.21)$$

Nesta especificação, a média e a variância mudam simultaneamente.

**Estimação dos modelos de mudança markoviana** A verossimilhança dos modelos de mudança markoviana pode ser avaliada eficientemente usando o procedimento de filtro de ([HAMILTON, 1990](#)) e o procedimento de suavização de ([KIM, 1994](#)).

A log verossimilhança, que é função dos parâmetros em [3.19](#) ou [3.20](#) bem como das probabilidades de transição  $p_{i|j}$ , pode então ser maximizada sujeita às restrições de que as probabilidades estão variando entre 0 e 1 e sempre somam um. Para esta estimativa ([DOORNIK, 2018](#)) sugere o uso de programação não linear.

**Tipos de variância** Os modelos podem lidar com alguns tipos de variância. Neste trabalho serão utilizados três tipos, que se diferenciam um do outro de acordo com a variação ou não da média ou da variância:

- **Variância fixa:** a mudança de regime acontece apenas na média; a variância é a mesma para todos os regimes.
- **Variância variável:** cada regime tem sua média e sua variância, mas a esta variação não é independente; por exemplo, um modelo com dois regimes tem uma média e uma variância para o regime 0 e uma média e uma variância para o regime 1.
- **Variação na média e na variância:** os regimes tem variação independente de média e de variância; estes modelos são conhecidos como modelos de mudança markoviana com estrutura de componentes e são explicados em mais detalhes a seguir.

**Modelos de mudança markoviana com estrutura de componentes** Dados macroeconômicos que incluem o período da Grande Moderação requerem dois ou mais regimes de variância e dois ou mais regimes de média. Entretanto, estes não coincidem: a média pode flutuar mais rapidamente que a variância. O modelo de mudança markoviana com componente média-variância foi proposto por ([DOORNIK, 2018](#)).

Um modelo restrito de quatro estados pode ser construído permitindo-se que a média e a variância flutuem com uma estrutura de componentes:

	$S_t^m = 0 : \mu_0$	$S_t^m = 1 : \mu_1$
$S_{t+1}^m = 0 : \mu_0$	$p_{0 0}^m$	$p_{0 1}^m$
$S_{t+1}^m = 1 : \mu_1$	$p_{1 0}^m$	$p_{1 1}^m$
	$S_t^v = 0 : \sigma_0$	$S_t^v = 1 : \sigma_1$
$S_{t+1}^v = 0 : \sigma_0$	$p_{0 0}^v$	$p_{0 1}^v$
$S_{t+1}^v = 1 : \sigma_1$	$p_{1 0}^v$	$p_{1 1}^v$

Este é um modelo restrito de quatro regimes:

	$S_t^v = 0$	$S_t^v = 1$
$S_t^m = 0$	$S_t^m = 0$	$S_t^m = 1$
$S_t = 0$	$S_t = 1$	$S_t = 2$
$S_{t+1} = 0$	$p_{0 0}^v \mathbf{P}^m$	$p_{0 1}^v \mathbf{P}^m$
$S_{t+1} = 2$	$p_{1 0}^v \mathbf{P}^m$	$p_{1 1}^v \mathbf{P}^m$

Em vez de quatro médias e quatro variâncias há apenas dois de cada. Mais ainda, a matriz de transição é restrita a:

$$\mathbf{P} = \mathbf{P}^v \otimes \mathbf{P}^m \quad (3.22)$$

que tem apenas quatro probabilidades livres, em vez das 12 do modelo MS(4) irrestrito.

### 3.1.3 Redes Neurais Recorrentes LSTM

As *Long Short-Term Memory* (LSTM) são um tipo de Rede Neural Recorrente com a capacidade de lembrar valores dos estágios anteriores para uso futuro (HOCHREITER; SCHMIDHUBER, 1997). Antes de detalhar o que são as redes LSTM, é preciso falar dos conceitos de Redes Neurais Artificiais, Redes Neurais Recorrentes e o embasamento matemático destes modelos. Para isto, esta seção irá se basear nos trabalhos de (SCHMIDHUBER, 2015), (BENGIO; GOODFELLOW; COURVILLE, 2017), (HAYKIN, 1998) e (MITCHELL, 1997).

#### Redes Neurais Artificiais

As Redes Neurais Artificiais (ANN<sup>6</sup>) consistem de pelo menos três camadas:

- camada de entrada (*input layer*)
- camada escondida (*hidden layer*)
- camada de saída (*output layer*)

O número de características<sup>7</sup> do conjunto de dados determina a dimensionalidade ou o número de nós da camada de entrada. Estes nós são conectados por ligações conhecidas por *sinapses* com os nós criados na camada escondida - ou camadas escondidas, pois pode haver mais de uma. Estas ligações *sinápticas* carregam alguns pesos para cada nó na camada de entrada. Os pesos, basicamente, têm o papel de decidir qual sinal, ou entrada, irá passar e qual não irá passar. Os pesos também mostram a força da camada escondida. Uma rede neural aprende, basicamente, ajustando os pesos para cada *sinapse*.

Nas camadas escondidas, os nós aplicam uma função de ativação (por exemplo, sigmoide ou tangente hiperbólica) sobre a soma ponderada das entradas para transformar as entradas em saídas, ou valores preditos. A camada de saída gera um vetor de probabilidades para as várias saídas e seleciona aquela com menor taxa de erro, i. e., minimizando as diferenças entre os valores esperados e os preditos, também conhecido como custo, usando uma função chamada softmax.

As atribuições do vetor de pesos e os consequentes erros obtidos pelo treinamento da rede pela primeira vez podem não ser os melhores. Para encontrar os valores otimizados

---

<sup>6</sup>Artificial Neural Networks

<sup>7</sup>features

para os erros, estes são retro propagados<sup>8</sup> na rede da camada de saída em direção das camadas escondidas e como resultado os pesos são ajustados. O procedimento é repetido várias vezes - ou várias *épocas* - com as mesmas observações e os pesos são reajustados até que haja uma melhora nos valores preditos e, como consequência, uma redução do custo. Quando a função de custo é minimizada, o modelo está treinado.

**Redes *Feedforward* de Camada Única** A equação que descreve este tipo de rede é:

$$y_k = \varphi\left(\sum_{i=0}^m w_i x_i\right) \quad (3.23)$$

onde  $m$  é o número de entradas,  $y_k$  é a saída,  $\varphi(\cdot)$  é uma função de ativação,  $x_i$  é uma entrada e  $w_i$  é o peso correspondente. Fixando-se  $x_0 = 1$ ,  $w_{k0}$  torna-se um parâmetro de viés, que tem por objetivo aplicar uma transformação afim em  $\mathbf{X} \cdot \mathbf{W}$ . Este viés, assim como o restante dos pesos, são parâmetros ajustáveis dentro da rede. Uma representação gráfica de um neurônio pode ser vista na figura 3.2.

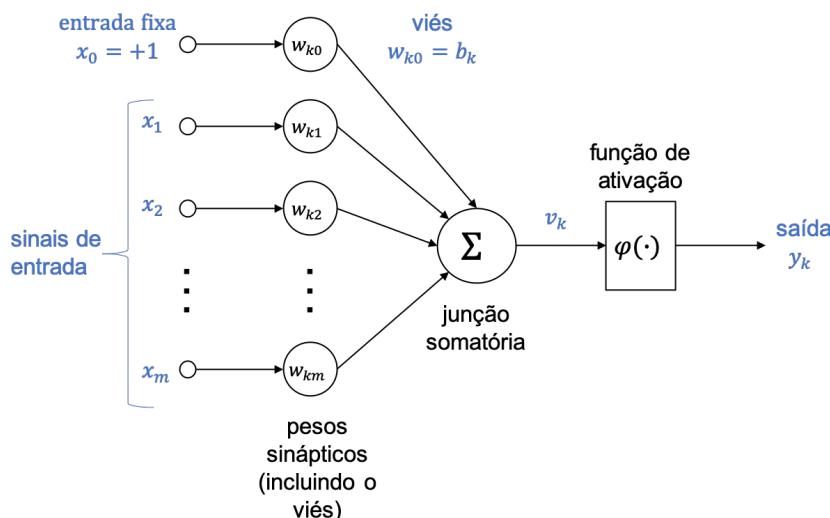


Figura 3.2: Modelo não linear de um neurônio. Adaptada de (HAYKIN, 1998)

**Funções de Ativação** A função de ativação  $\varphi(\cdot)$  da equação 3.23 pode ter diferentes características.

**Função linear** Se  $\varphi(\cdot)$  tem a forma de uma função linear:

<sup>8</sup>back propagated

$$\varphi(a) = a, \quad (3.24)$$

a equação 3.23 se torna uma regressão linear. Geralmente é usada na camada de saída com um único neurônio em tarefas de regressão.

**Função degrau** Se  $\varphi(\cdot)$  é uma função binária dada por:

$$\varphi(a) = \begin{cases} 1 & \text{se } a > 0 \\ 0 & \text{caso contrário} \end{cases} \quad (3.25)$$

a equação 3.23 se torna um classificador do tipo Perceptron; ver (HAYKIN, 1998) para mais detalhes.

**Função sigmoide** Se  $\varphi(\cdot)$  tem a forma de uma função sigmoide, que é a função logística, não linear:

$$\varphi(a) = \frac{1}{1 + e^{-a}} \quad (3.26)$$

a equação 3.23 se torna uma regressão logística quando usada na camada de saída com um único neurônio em tarefas de classificação binária. Entretanto, pode também ser utilizada nas camadas escondidas.

**Função tangente hiperbólica** Se  $\varphi(\cdot)$  tem a forma de uma função tangente hiperbólica:

$$\varphi(a) = \frac{e^a - e^{-a}}{e^a + e^{-a}} \quad (3.27)$$

uma função utilizada na camada de saída.

**Função ReLU** Se  $\varphi(\cdot)$  tem a forma de uma função Unidade Linear Retificada:

$$\varphi(a) = \max(0, \frac{1}{1 + e^{-a}}) \quad (3.28)$$

é a função de ativação mais utilizada nas camadas escondidas (BENGIO; GOODFELLOW; COURVILLE, 2017)..

**Função softmax** Se  $\varphi(\cdot)$  tem a forma de uma função softmax, ela geralmente é utilizada na última camada para tarefas de classificação multi classes. Esta função normaliza a saída de cada neurônio da última camada pela soma das saídas de todos os neurônios da última camada, fazendo com que a soma de todas as saídas seja 1. Deste modo, o neurônio cuja resultado possuir o valor mais alto fará com que a classe escolhida seja a classe que ele representa.

$$\varphi(a)_i = \frac{\exp(a_i)}{\sum_{j=1}^n \exp(a_j)} \quad (3.29)$$

onde  $i$  é o neurônio à que a função se refere e  $n$  é o total de neurônios da última camada.

**Redes *Feedforward* Multi Camadas** A equação que descreve este tipo de rede é:

$$y_k = \rho\left(\sum_{j=0}^M w_{kj}^{(2)} \varphi(a_j)\right) \quad (3.30)$$

onde,

$$a_j = \sum_{i=0}^m w_{ij}^{(1)} x_i \quad (3.31)$$

onde  $M$  é o número de neurônios da camada. Esta estrutura é bem similar à apresentada na equação 3.23. Mas a função  $\rho(\cdot)$  é linear, e a grande diferença é que a saída da camada interna é multiplicada novamente por um vetor de pesos e empacotada em função de ativação  $\varphi(\cdot)$ . Note que a saída da primeira camada é usualmente não linear, o que implica em que mais camadas tornam a saída mais complexa.

**Treinamento da Rede** O treinamento da rede pode ser descrito pela definição de uma função de perda (função de erro) e a utilização da retropropagação com descida do gradiente. A retropropagação é o método de encontrar o gradiente da função de erro, pela rede, de trás para frente. A descida do gradiente é um método simples de otimização de pesos.

**Derivação da retropropagação** Considere-se a seguinte soma:

$$a_j = \sum_i w_{ji} z_i \quad (3.32)$$

Esta soma é uma representação de uma camada da rede, onde  $w_{ij}$  são os pesos e  $z_i$  são as entradas. Quando  $z_i$  é a entrada inicial da rede ela é chamada de  $x_i$  e a saída da rede é chamada de  $y_i$ .

Considere-se agora

$$z_j = \varphi(a_j) \quad (3.33)$$

onde a camada está empacotada por uma função de ativação. De modo a otimizar os pesos deve-se definir uma função de erro adequada e encontrar suas derivadas. Funções de erro geralmente utilizadas nas regressões são o MSE e o MAPE. Ver equações 2.4 e 2.1

$$E = \sum_i E^n \quad (3.34)$$

Além disso,

$$E^n = E^n(y_1, \dots, y_c) \quad (3.35)$$

A derivada de cada erro com respeito ao peso, usando a regra da cadeia, é:

$$\frac{\partial E^n}{\partial w_{ji}} = \frac{\partial E^n}{\partial a_j} \frac{\partial a_j}{\partial w_{ji}} \quad (3.36)$$

A primeira parte do lado direito da equação 3.36 é o *delta*:

$$\delta_j \equiv \frac{\partial E^n}{\partial a_j} \quad (3.37)$$

A segunda parte do lado direito da equação 3.36 se torna, claramente:

$$\frac{\partial a_j}{\partial w_{ji}} = z_i \quad (3.38)$$

Substituindo-se 3.37 e 3.38 em 3.36, tem-se:

$$\frac{\partial E^n}{\partial w_{ji}} = \delta_j z_i \quad (3.39)$$

Isto quer dizer que para se calcular a derivada da função de erro é necessário multiplicar o *delta* pelas entradas. O *delta* para a camada de entrada se torna:

$$\delta_k \equiv \frac{\partial E^n}{\partial a_k} = \varphi'(a_k) \frac{\partial E^n}{\partial y_k} \quad (3.40)$$

O *delta* para as camadas escondidas é:

$$\delta_j \equiv \frac{\partial E^n}{\partial a_j} = \sum_k \frac{\partial E^n}{\partial a_k} \frac{\partial a_k}{\partial a_j} \quad (3.41)$$

Assim, o algoritmo geral da retropropagação é:

$$\delta_j = \varphi'(a_j) \sum_k w_{kj} \delta_k \quad (3.42)$$

A fórmula da retropropagação se propaga para trás através da rede e modo a calcular o gradiente. A variação em  $a_j$  vai através de  $a_k$ . Em uma primeira vista pode parecer complicado mas o *delta* para a camada de saída é conhecido, então pode-se recursivamente utilizar a equação da retropropagação 3.42 para calcular o gradiente para qualquer estrutura de rede *feedforward*.

Uma vez que um lote (*batch*) é inserido na rede e o algoritmo de retropropagação da rede é executado juntamente com o algoritmo de otimização, uma época (*epoch*) está completa.

**Descida do gradiente** O algoritmo de retropropagação encontra o gradiente da função de erro, mas é necessário ter um algoritmo de otimização. Analiticamente, é possível encontrar um conjunto de pesos,  $\mathbf{w}$ ,  $\nabla E = 0$ . Pode não ser possível encontrar  $\mathbf{w}$  dada a dimensionalidade do problema. Quando isso acontece, o algoritmo de descida do gradiente é usado para se chegar o mais próximo possível. O método começa com um vetor de pesos aleatórios,  $\mathbf{w}^{(0)}$ , e calcula o gradiente para a função de erro naquele ponto. Então o algoritmo dá um pequeno passo,  $\tau$ , na direção oposta à do gradiente e novamente calcula o novo ponto  $\mathbf{w}^\tau$ . Isto pode ser expresso como:

$$\Delta \mathbf{w}^\tau = -\eta \nabla E|_{w_\tau} \quad (3.43)$$

onde  $\eta$  é a taxa de aprendizagem (*learning rate*). Este algoritmo simples tem pro-

blemas com otimizações não convexas, o que é comum nas redes neurais. Entretanto, já alguns ajustes tais como a descida estocástica do gradiente (SGD<sup>9</sup>) e modo a superar este problema.

**Regularização** Regularização, na terminologia de redes neurais, se refere ao processo de regular as camadas de rede de modo a prevenir o sobre ajuste (*overfitting*). *Dropout* é o método de regularização mais comum, baseado na ideia simples de descartar um percentual. i.e., desistir de parte do aprendizado, de modo a reduzir os graus de liberdade (os parâmetros ajustáveis) da rede. A motivação do *dropout* está no fato de que se o conjunto de dados de treinamento tem 1000 observações e a rede consiste de 1000 nós, ela pode armazenar todas a informação de entrada e, portanto, não precisa construir um modo generalizado de descrever os dados. Se 500 nós são removidos, a rede já não consegue armazenar toda a entrada e então precisa encontrar modos inteligentes de descrever os dados. O método de *dropout* é usado geralmente nas redes profundas, que tem uma grande quantidade de parâmetros ajustáveis.

## Redes Neurais Recorrentes

As Redes Neurais Recorrentes são uma classe de redes neurais que podem ser pensadas sob o conceito de *feedback loops* (laços de retorno). Quando os dados são processados por uma das camadas escondidas em uma rede eles são enviados de volta e de novo processadas pela camada junto com a nova entrada. Há vários tipos de redes recorrentes dependendo de como elas são estruturadas. Um tipo de rede recorrente são as redes LSTM (*long short-term memory*) que serão discutidas na próxima seção.

Redes neurais recorrentes são usadas quando os dados são de natureza sequencial, por exemplo, séries de tempo. No caso de dados com estrutura autoregressiva as redes neurais recorrentes podem ser usadas. Um exemplo onde redes neurais recorrentes são usadas no processamento de linguagem natural (NLP<sup>10</sup>, pois a construção de frases é sequencial, e a ordem dos elementos na frase faz diferença.

---

<sup>9</sup> *Stochastic Gradient Descent*

<sup>10</sup> *Natural Language Processing*

## Redes LSTM

As Redes Neurais Recorrentes LSTM foram introduzidas por ([HOCHREITER; SCHMIDHUBER, 1997](#)). Uma das vantagens do modelo LSTM comparado às redes neurais recorrentes tradicionais é a habilidade de capturar estruturas autorregressivas de comprimentos arbitrários. A rede recorrente regular precisa saber quantos *loops* de *feedback* que estarão presentes, o que não é necessário quando se lida com redes LSTM.

Em vez de nós regulares de redes neurais, a rede LSTM consiste de blocos LSTM. Considere o seguinte sistema de equações:

$$\left\{ \begin{array}{l} f_t = \sigma(W_f \cdot [x_t, h_{t-1}] + b_f) \\ i_t = \sigma(W_i \cdot [x_t, h_{t-1}] + b_i) \\ o_t = \sigma(W_o \cdot [x_t, h_{t-1}] + b_o) \\ c_t = f_t \circ c_{t-1} + i_t \circ \tanh(W_c \cdot [x_t, h_{t-1}] + b_c) \\ h_t = o_t \circ \tanh(c_t) \end{array} \right. \quad (3.44)$$

onde  $\sigma(\cdot)$  é a função sigmoide, ver [3.26](#),  $\tanh(\cdot)$  é a função tangente hiperbólica, ver [3.27](#),  $x_t$  é o vetor de entrada,  $h_t$  é o vetor de saída,  $c_t$  é um vetor de estados de células.  $W$  são os pesos e  $b$  são os vieses.  $f_t$ ,  $i_t$  e  $o_t$  são chamados portões (*gates*) do bloco. Note que o não é uma multiplicação matricial, mas um produto de Schur (ou produto de Hadamard), i.e., um produto elemento a elemento.

O sistema de equações [3.44](#) representa um bloco LSTM. Esquematicamente, um bloco LSTM pode ser representado com na figura [3.3](#). É possível ver as propriedades recorrentes do bloco ao notar que  $h + t - 1$ , i.e., o vetor de saída para o período  $t - 1$  é incluso no cálculo do vetor de saída  $h_t$ .

O primeiro portão do bloco LSTM,  $f_t$ , é o portão de esquecimento (*forget gate*). Este portão decide qual informação será esquecida no estado da célula,  $c_t$ . Nota-se que a expressão linear está empacotada em uma função sigmoide, que é limitada entre 0 e 1. Se a otimização configura  $f_t$  para 0 o valor é *esquecido* no cálculo. O portão de entrada,  $i$ , regula qual nova informação será armazenada no estado da célula. O portão de saída é  $o_t$ . O vetor de saída  $h_t$  é uma combinação do portão de saída e o estado da célula que contém a informação armazenada.

A chave das redes LSTM são as funções de ativação que ajudam o bloco a armazenar e

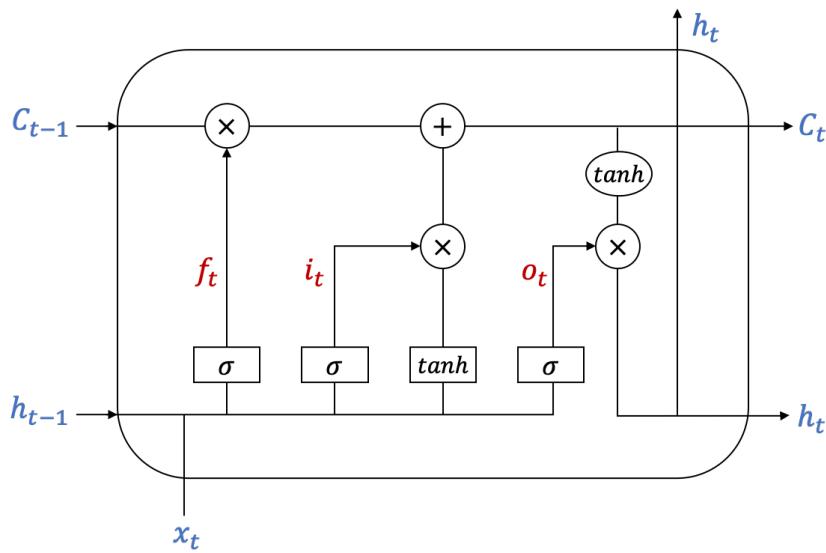


Figura 3.3: Estrutura de um bloco LSTM. Adaptado de ([WANG; ZHU; LI, 2019](#)).

esquecer informação. O algoritmo de otimização então ajusta os parâmetros de acordo com o necessário. Segundo ([HOCHREITER; SCHMIDHUBER, 1997](#)), o algoritmo baseado em gradiente das redes LSTM forçam que haja um fluxo constante - sem explodir nem desaparecer - dos erros através do estado interno dos blocos.

### 3.1.4 Model Confidence Set

O Model Confidence Set (MCS) é um método estatístico definido por ([HANSEN; LUNDE; NASON, 2011](#)) para determinar, dentro de um conjunto de modelos  $\mathcal{M}^*$ , qual é o melhor modelo, denominado  $\mathcal{M}^0$ , em termos de critérios especificados pelo usuário por meio de uma função de perda  $L$  escolhida. O MCS define

$$d_{ij,t} \equiv L_i - L_j \quad (3.45)$$

em um determinado período  $t$  e

$$\mu_{i,j} \equiv E(d_{ij,t}). \quad (3.46)$$

O modelo  $i$  é preferido ao modelo  $j$  se  $\mu_{ij} < 0$ . O conjunto de modelos superiores é dado por

$$\mathcal{M}^* \equiv \{ \in \mathcal{M}^0 : \mu_{i,j} \leq 0 \text{ para todo } j \in M^0 \} \quad (3.47)$$

As hipóteses nula do teste é:

$$H_{0,\mathcal{M}} : \mu_{ij} = 0 \text{ para todo } i, j \in \mathcal{M}. \quad (3.48)$$

E a hipótese alternativa é:

$$H_{a,\mathcal{M}} : \mu_{ij} \neq 0 \text{ para todo } i, j \in \mathcal{M}. \quad (3.49)$$

O modelo MCS trabalha com uma *teste de equivalência*  $\delta_{\mathcal{M}}$  que é usado para testar a hipótese nula para cada modelo e uma *regra de eliminação*  $\epsilon_{\mathcal{M}}$  que remove o modelo do conjunto caso a nula seja rejeitada. Este procedimento é repetido até que nenhum modelo seja rejeitado pela regra de equivalência. Os modelos restantes são então ranqueados pelos p-valores.

## 3.2 Metodologia

Nesta seção serão descritos os dois experimentos de previsão do ICMS-SP realizados neste trabalho, um experimento univariado, apenas com a série do ICMS-SP, e um experimento univariado com regressão múltipla, incluindo como variáveis exógenas as séries do índice de inflação IGP-M, do índice de atividade econômica IBC-Br, da cotação do Dólar e dos dias úteis de cada mês.

### 3.2.1 Dados

#### ICMS-SP

A série de arrecadação utilizada é a série atualmente em uso para a previsão oficial da receita de ICMS do estado de São Paulo, vista na figura 3.4, proveniente da Secretaria da Fazenda deste mesmo estado, com os valores nominais arrecadados, em R\$, dos meses de janeiro de 2003 a dezembro de 2022. Esta série difere um pouco da série disponível no portal da Secretaria da Fazenda do Estado<sup>11</sup>, pois há alguns ajustes realizados com relação a pagamentos diferidos, que nos dados publicados estão computados no mês de recebimento, mas na série utilizada para previsão são computados no mês de competência.

---

<sup>11</sup><https://portal.fazenda.sp.gov.br/acessoinformacao/Paginas/Relatórios-da-Receita-Tributária.aspx>

A título de reproduzibilidade dos experimentos, a série modificada está disponível no Apêndice A deste trabalho.

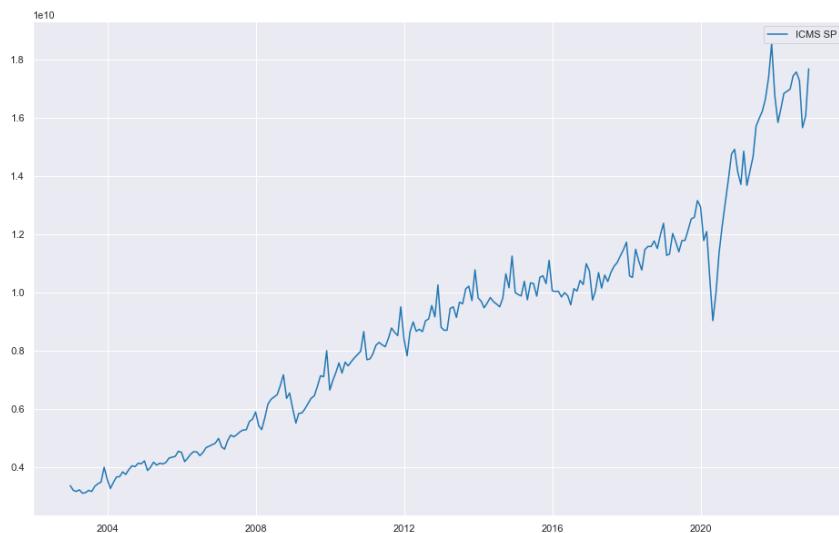


Figura 3.4: Série de arrecadação mensal nominal do ICMS em R\$, de janeiro de 2003 a dezembro de 2022

## Variáveis Exógenas

A receita do ICMS é impactada pelo crescimento (ou contração) do PIB, pela variação de preços da economia e pela carga tributária do ICMS ((ESTADO), 2022). Esta é a razão da escolha do IGP e do Dólar, pois medem a variação de preços, e do IBC-Br, pois mede o PIB. A série de Dias Úteis foi incluída pois a quantidade de dias úteis do mês tem implicações na arrecadação. Pode-se notar na figura 3.5 que, no mês de fevereiro, a arrecadação é sempre menor.

Todas as séries das variáveis exógenas foram obtidas no website do IPEA<sup>12</sup>. A seguir, uma breve descrição de cada uma.

**IGP-M** A série do IGP-M usada foi a denominada *Índice geral de preços do mercado (IGP-M) - geral: índice (ago. 1994 = 100)*. Esta série, gerada pela Fundação Getulio Vargas, é uma média ponderada de três outros índices de preços, cujas pesos correspondem a parcelas da despesa interna bruta, calculadas com base nas Contas Nacionais (GAMA, 2020):

---

<sup>12</sup><https://www.ipeadata.gov.br>

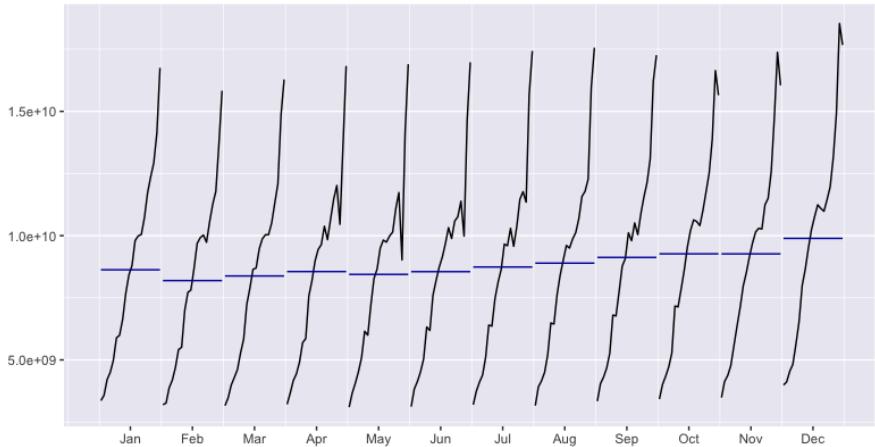


Figura 3.5: Séries de arrecadação mês a mês do ICMS, de 2003 a 2022; as linhas horizontais azuis representam a média de cada mês.

- IPA-M: Índice de Preços ao Produtor Amplo - Mercado, com peso de 60%, que equivale ao valor adicionado pela produção de bens agropecuários e industriais nas transações comerciais do produtor.
- IPC-M: Índice de Preços ao Consumidor - Mercado, com peso de 30%, equivalem ao valor adicionado pelo setor varejista e pelos serviços destinados ao consumo das famílias.
- INCC-M: Índice Nacional de Custo da Construção, com peso de 10%, equivale ao valor adicionado pela indústria de construção civil.

O IGP apresenta-se em três versões: IGP-10, IGP-DI e IGP-M. A diferença entre eles é o período de apuração. No caso do IGP-M, este período está compreendido entre os dias 21 do mês anterior e o dia 20 do mês de referência.

É importante destacar que a versão IGP-DI do índice seria mais adequada pois comprehende o mês de referência cheio. Entretanto, a escolha pelo IGP-M se deu por esta versão ser a que tem previsões disponíveis no Sistema de Expectativas de Mercado do Bacen<sup>13</sup>. E estas previsões foram utilizadas nas previsões do ICMS do experimento univariado com regressão múltipla, como será visto em breve.

**IBC-Br** A série do IBC-Br utilizada foi a denominada *Índice real da atividade econômica do Banco Central (IBC-Br): índice real (2002=100)*. Trata-se de um indicador mensal

<sup>13</sup><https://www3.bcb.gov.br/expectativas2/#/consultaSeriesEstatisticas>

gerado pelo Banco Central com o objetivo de mensurar a evolução contemporânea da atividade econômica do país.

**Dólar** A série do Dólar utilizada foi a denominada *Taxa de câmbio - R\$ / US\$ - comercial - venda - média* com frequência mensal. A série, elaborada pelo Bacen, é composta pela média no período a partir das séries de taxas diárias.

A variável Dólar, como será visto na seção que descreve os experimentos, foi omitida nos modelos de mudança markoviana. Isto porque, nos experimentos realizados, os modelos com regressão múltipla sem dólar tiveram melhor desempenho do que os modelos que utilizaram a variável Dólar.

**Dias Úteis** A série utilizada foi a denominada *Número de dias úteis*, gerada pelo IPEA, levando em conta os feriados nacionais para a contabilização dos dias úteis. Os feriados considerados são: Confraternização Universal (Ano Novo), Terça-feira de Carnaval, Paixão de Cristo, Tiradentes, Dia do Trabalho, Corpus Christi, Independência do Brasil, Nossa Senhora Aparecida, Finados, Proclamação da República e Natal.

### 3.2.2 Experimentos

Foram realizados dois experimentos, um univariado, trabalhando apenas com a série do ICMS, e outro univariado com regressão múltipla, trabalhando com todas as séries mencionadas. Estes dois experimentos foram repetidos para os anos de 2013 a 2022, gerando 10 conjuntos de previsões para cada experimento, duas para cada ano.

Os experimentos foram, em sua maioria, executados em um *Jupyter Notebook*<sup>14</sup> rodando Phyton 3. Os gráficos foram elaborados, em sua grande maioria, utilizando a biblioteca *Seaborn*([WASKOM, 2021](#)).

Roteiro dos 20 experimentos:

1. Separar os dados em amostra de treino (ou *in-sample*) e amostra de teste (*out-sample*) para o ano de referência.
2. Obter as séries exógenas a partir das previsões do IGP-M mensal, PIB trimestral e Dólar Mensal no Sistema de Expectativas de Mercado do Bacen para os experimentos univariados com regressão múltipla.

---

<sup>14</sup><https://jupyter.org>

3. Projetar o IBC-Br a partir do PIB trimestral para os experimentos univariados com regressão múltipla.
4. Descrever as séries em termos de estacionariedade e sazonalidade em ambos os experimentos.
5. Encontrar as *dummies* de intervenção usando o X13-ARIMA-SEATS em ambos os experimentos.
6. Montar a matriz de exógenas somente com as *dummies* no experimento univariado e com as *dummies* mais as séries exógenas no univariado com regressão múltipla.
7. Rodar o modelo Naïve no experimento univariado.
8. Rodar o Auto-arima e encontrar os melhores parâmetros baseado nos critérios de informação AIC e BIC com e sem exógenas em ambos os experimentos.
9. Rodar o modelo SARIMA do experimento univariado.
10. Rodar os modelos SARIMAX de ambos os experimentos.
11. Rodar os modelos Markov-Switching de ambos os experimentos
12. Rodar os modelos RNN-LSTN de ambos os experimentos em validação para escolher a arquitetura.
13. Rodar os modelos RNN-LSTN de ambos os experimentos em testes com a arquitetura escolhida.
14. Para cada um dos modelos e ano de referência, em ambos experimentos, realizar a previsão fora da amostra para 12 meses e coletar as métricas RMSE e MAPE.
15. Montar uma matriz com o MAPE dos modelos escolhidos de cada estratégia, experimento e ano de referência.
16. Rodar o MCS para ranquear os modelos de acordo com a capacidade de previsão de cada um.

## Experimento Univariado

No experimento univariado foram adotadas 5 estratégias de modelos:

- modelo Naïve, como *baseline*, que é a média mensal da arrecadação do ano anterior;
- modelo SARIMA apenas com a série do ICMS;
- modelo SARIMAX com a série do ICMS e variáveis *dummy* de intervenção;
- modelo Markov-Switching apenas com a série do ICMS;
- modelo RNN-LSTM apenas com a série do ICMS.

**Log Natural** Os valores de arrecadação da série são muito elevados, na ordem de  $10^9$  e  $10^{10}$ . Por este motivo a série foi colocada em log.

**Métricas** Todos os modelos tiveram seu desempenho medido em termos do MAPE, descrito na equação 2.1 e do RMSE, descrito na equação 2.3. Para o cálculo de ambos foram utilizadas funções da biblioteca Scikit Learn<sup>15</sup> (PEDREGOSA et al., 2011).

**Conjuntos de Treinamento e Teste** A série do ICMS foi dividida em dois conjuntos, treinamento e teste, para as estratégias SARIMA, SARIMAX e Markov-Switching. O conjunto de treinamento se iniciou em janeiro de 2003 e terminou um ano antes do ano de referência do experimento. Por exemplo, o conjunto de treinamento de 2013 iniciou-se em janeiro de 2003 e terminou em dezembro de 2012. O de 2022 iniciou-se também em janeiro de 2003, mas terminou em dezembro de 2021.

Já o conjunto de testes, em todos os anos, compreendeu o período ao qual o experimento se referia. Assim, o conjunto de testes de 2013 iniciou-se em janeiro de 2013 e terminou em dezembro de 2013. O de 2022, iniciou-se em janeiro de 2022 e terminou em dezembro de 2022.

Os modelos SARIMA, SARIMAX e Markov-Switching foram treinados com estes conjuntos de treinamento e tiveram sua acurácia avaliada contra o conjunto de testes. Deste modo, todas as previsões foram feitas fora da amostra (*out-of sample*).

Para as Redes Neurais, além dos conjuntos de treinamento e testes, foi criado também um conjunto de validação. Neste caso, o conjunto de treinamento parou dois anos antes

---

<sup>15</sup><https://scikit-learn.org/stable/index.html>

do ano de referência. O de testes continuou igual. E o de validação compreendeu os 12 meses entre o conjunto de treinamento e o de testes. No caso de 2013, por exemplo, o treinamento passou a terminar em dezembro de 2011. Validação iniciou-se em janeiro de 2012 e terminou em dezembro de 2012. E testes continuou igual, começando em janeiro de 2013 e terminando em dezembro de 2013. De modo análogo, em 2022 o conjunto de treinamento parou em dezembro de 2020. O de validação, foi de janeiro a dezembro de 2021. E o de testes, de janeiro a dezembro de 2022.

**Estacionariedade** Os testes de raiz unitária ADF e KPSS foram realizados na série do ICMS em todos os 10 experimentos univariados. Mais sobre os testes na seção [3.1.1](#).

**Sazonalidade** A análise da sazonalidade da série do ICMS foi feita em todos os 10 experimentos univariados utilizando-se do modelo X13-ARIMA-SEATS ([MONSELL; LY-TRAS; FINDLEY, 2013](#)). Foi encontrada sazonalidade na série em todos os experimentos. Entretanto, a série dessazonalizada não foi utilizada no modelos, pois o SARIMA e o SARIMAX lidam com a sazonalidade. O modelo Markov-Switching utilizou a técnica de dummies sazonais centradas. E a RNN-LSTM aprende por si só a lidar com a sazonalidade.

**Dummies de Intervenção** O X13-ARIMA-SEATS aponta outliers na série. Os outliers apontados foram utilizados no modelo SARIMAX como *dummies* de intervenção. Mas nos modelos Markov-Switching e RNN-LSTM não foram usados pois a proposta deste trabalho é de estudar se modelos não lineares tratam melhor as mudanças de regime na série de arrecadação do que variáveis *dummy* nos modelos lineares.

**Auto-arima** A função auto-arima da biblioteca do Python *pmdarima*([SMITH et al., 2017](#)) foi utilizada para determinar os melhores parâmetros  $(p,d,q)$  e  $(P,D,Q)$  dos modelos SARIMA e SARIMAX. A função foi rodada quatro vezes em cada experimento univariado, duas vezes usando somente a série do ICMS para o modelo SARIMA, duas vezes usando a série do ICMS e as *dummies* de intervenção para o modelo SARIMAX. E, para cada modelo, usou alternadamente o critério de Akaike e o critério de Schwarz. Mais sobre Auto-arima e critérios de informação em [3.1.1](#).

**SARIMA** O modelo SARIMA foi executado para cada um dos experimentos univariados uma ou duas vezes, a depender dos resultados dos parâmetros  $(p,d,q)$  e  $(P,D,Q)$  definidos na execução do Auto-arima. Para cada execução foi feita a previsão para 12 meses e calculados o MAPE (eq. 2.1) e o RMSE (eq. 2.3), fora da amostra, comparando com o realizado no ano de referência do experimento.

Foi utilizada a função ARIMA da biblioteca *statsmodels*<sup>16</sup> do Python para rodar o modelo.

**SARIMAX** O modelo SARIMAX também foi executado para cada um dos experimentos univariados uma ou duas vezes, a depender dos resultados dos parâmetros  $(p,d,q)$  e  $(P,D,Q)$  definidos na execução do Auto-arima. De modo análogo, para cada execução foi feita a previsão para 12 meses e calculados o MAPE e o RMSE, fora da amostra, comparando com o realizado no ano de referência do experimento.

**Markov-Switching** No experimento univariado o modelo de mudança markoviana executado foi o Markov-Switching Dinâmico MS-DR. O motivo da escolha do modelo dinâmico é que, segundo (DOORNIK, 2018), o modelo Markov-Switching Autorregressivo MS-AR é infactível para dados mensais, o que é caso destes experimentos.

O treinamento foi feito com a série de treinamento. Cada modelo utilizou duas mudanças de regime e as seguintes variáveis:

- $\log(ICMS)$  em primeira diferença - Y
- Constante - mudança de regime (regime switching)
- $\log(ICMS)$  em primeira diferença com uma defasagem - (regime switching)
- 11 *dummies* sazonais centradas - X (variável exógena)

De modo análogo a (CHAUDET, 2002), foram usadas séries estacionárias nos modelos de mudança markoviana. A série do ICMS, de acordo com os testes ADF e KPSS realizados, apresenta raiz unitária em nível mas é estacionária em primeira diferença. Por este motivo foi usada em primeira diferença no experimento univariado.

Foram treinados três modelos. A diferença entre eles está na variação da variância. O modelo *fixed* usa a variância fixa, com mudança de regime somente na média. O modelo

---

<sup>16</sup><https://www.statsmodels.org/stable/index.html>

com mudança de variância, *switching variance*, varia tanto a média quanto a variância, mas de maneira dependente, onde cada um dos dois regimes possui sua média e variância fixas. O modelo *mean variance component* varia tanto a média como a variância, mas de maneira independente, tendo um regime para média e outro para variância. Mais sobre este assunto na seção 3.1.2

A biblioteca *statsmodels*, embora possua funções para modelos de mudança markoviана, estes modelos não tem o método *forecast* implementado. Por isso toda a modelagem e previsão para estes modelos foi realizada no software OxMetrics 8<sup>17</sup>, usando o pacote PcGive.

As previsões de 12 meses foram feitas fora da amostra e comparadas com os valores realizados constantes da série de testes do ano de referência do experimento. Foram computadas as métricas MAPE e o RMSE.

**RNN-LSTM** Para os modelos RNN-LSTM univariados foi usada somente a série do ICMS como variável alvo (*target*) e transformações matemáticas da própria série na matriz de características (*features*), a saber: defasagens 1, 2, 3, 4, 6 e 12 e médias móveis 2, 3, 6 e 12 meses, conforme tabela 3.1.

#### Cálculos Realizados sobre os Valores da Série $x$

$$MM12 = \frac{x_{n-1} + x_{n-2} + \dots + x_{n-11} + x_{n-12}}{12}$$

$$MM6 = \frac{x_{n-1} + x_{n-2} + \dots + x_{n-5} + x_{n-6}}{6}$$

$$MM3 = \frac{x_{n-1} + x_{n-2} + x_{n-3}}{3}$$

$$MM2 = \frac{x_{n-1} + x_{n-2}}{2}$$

$$L12 = x_{n-12}$$

$$L6 = x_{n-6}$$

$$L4 = x_{n-4}$$

$$L3 = x_{n-3}$$

$$L2 = x_{n-2}$$

$$L1 = x_{n-1}$$

Tabela 3.1: Cálculos realizados sobre os valores da série do ICMS, conforme (SILVA; FIGUEIREDO, 2020)

A matriz de características será criada de duas formas diferentes para dois experimentos univariados diferentes. O primeiro, denominado RNN-LSTM-1-mês, onde cada mês

---

<sup>17</sup><https://www.doornik.com/products.html>

é previsto com os dados do mês anterior. E outro, denominado RNN-LSTM-12-meses, onde cada mês é previsto com os dados dos 12 meses anteriores.

Entretanto, a série do conjunto de testes não poderá ser usada para a criação da matriz de características de testes, pois neste caso se estaria usando os dados reais para a realização da previsão. Então, neste caso, será usada a técnica de previsão *multi-step* ([CHEVILLON, 2007](#)), que consiste em, tomando como exemplo o experimento de 2022 para o modelo RNN-LSTM-1-mês, realizar a previsão de janeiro de 2022 com os dados de dezembro de 2022, adicionar o resultado à série e criar a matriz de características de janeiro de 2022 a partir da nova série. Fazer então a previsão para fevereiro de 2022, adicionar à série e criar a matriz de características para fevereiro de 2022. E assim por diante até dezembro de 2022.

Já para o experimento de 2022, modelo RNN-LSTM-12-meses, realizar a previsão de janeiro de 2022 com os dados de janeiro a dezembro de 2022, adicionar o resultado à série e criar a matriz de características de janeiro de 2022 a partir dele. Fazer então a previsão para fevereiro de 2022, utilizando os dados de fevereiro de 2021 a janeiro de 2022 (previsto), adicionar à série e criar a matriz de características para fevereiro de 2022. E assim por diante até dezembro de 2022.

É importante frisar que não há retreinamento da rede. Nos dois casos, RNN-LSTM-1-mês e 12-meses, a rede é treinada uma única vez.

Em cada rede foram testadas diferentes arquiteturas com diferentes números de camadas e de neurônios. Além disso, diversos hiperparâmetros foram testados: funções de ativação, tamanho de lote (*batch*), algoritmo de ajuste de pesos, taxa de aprendizado, etc.

Chegou-se a duas arquiteturas principais, uma com uma camada LSTM e outra com três camadas LSTM. Nas camadas LSTM, a função de ativação para todas foi a tangente hiperbólica. Em todas as camadas LSTM foi usada a regularização por meio de *dropout*. As duas arquiteturas têm uma camada Densa na última camada, com um neurônio e função de ativação linear. A função de perda das duas arquiteturas é o MAPE e o otimizador Adam com taxa de aprendizado, em geral, de  $10^{-5}$ . Ver a seção [3.1.3](#) para maiores detalhes.

Os tamanho do lote variou entre 4 e 16 em cada experimento. O número de épocas foi fixado em um valor alto, 1000 épocas. Mas, para evitar sobre ajuste, foi usada a técnica de *parar antes* (*early stop*). Se a rede ficasse mais de 100 épocas sem melhorar a perda em

validação o treinamento era parado. E, para tentar fazer a rede aprender mais, foi usada a técnica de reduzir a taxa de aprendizado caso a rede encontrasse um platô (*reduce LR on plateau*). No caso, em 10% a cada cem épocas sem melhora da perda em validação.

A decisão de qual arquitetura usar para cada experimento univariado e cada modelo (1-mês e 12-meses) foi tomada a partir dos resultados do MAPE sobre o conjunto de validação. Foram treinadas redes nas duas arquiteturas para cada modelo com o MAPE calculado sobre o conjunto de validação, utilizando previsão *multi-step*. A arquitetura com melhor MAPE em validação de cada modelo foi a selecionada para ser a utilizada no experimento do ano de referência e ter o MAPE e o RMSE calculados e comparados com o conjunto de testes.

Para execução das redes foram utilizadas as bibliotecas *Keras*<sup>18</sup> e *TensorFlow*<sup>19</sup>.

## Experimento Univariado com Regressão Múltipla

No experimento univariado foram adotadas 3 estratégias de modelos:

- modelo SARIMAX com a série do ICMS, variáveis exógenas IGP-M, IBC-Br, Dólar e Dias Úteis e variáveis *dummy* de intervenção;
- modelo Markov-Switching com a série do ICMS, variáveis exógenas IGP-M, IBC-Br e Dias Úteis, mas sem as dummies de intervenção; Dólar não foi usado pois piorava as previsões;
- modelo RNN-LSTM com a série do ICMS, variáveis exógenas IGP-M, IBC-Br, Dólar e Dias Úteis, e também sem as dummies de intervenção.

**Conjuntos de Treinamento e Teste** A série do ICMS, no experimento univariado com regressão múltipla, também foi posta em log e foi dividida do mesmo modo que no experimento univariado, em treinamento e teste.

Os dados do IGP-M, IBC-Br e Dólar foram os dados reais para o conjunto de treinamento e os dados do Sistema de Expectativas de Mercado do BC para o conjunto de testes. Deste modo o experimento se aproximou mais do caso real, no qual a previsão é feita a priori, sem que nenhuma das séries - com exceção de Dias Úteis - tenha sido realizada no período de testes.

---

<sup>18</sup><https://keras.io>

<sup>19</sup><https://www.tensorflow.org>

**Previsão das Séries Exógenas** Cada uma das previsões das séries exógenas será detalhada a seguir.

**IGP-M** As estimativas do IGP-M são as obtidas nas Séries Estatísticas do Sistema de Expectativas de Mercado do Bacen usando-se os parâmetros de pesquisa:

- Grupo de Indicadores = Índices de Preços
- Periodicidade = Mensal
- Indicador(es) = IGP-M
- Selecione uma Base de Cálculo = Expectativas informadas nos últimos 30 dias
- Estatística = Mediana
- Período em que as projeções foram feitas:
  - Data inicial = primeiro dia útil de dezembro do ano anterior ao ano de referência do experimento
  - Data final = último dia útil de dezembro do ano anterior ao ano de referência do experimento

Foram então utilizados os valores da variação do IGP-M, de janeiro a dezembro do ano de referência do experimento, constantes última linha da tabela obtida na pesquisa, isto é, do último dia útil do mês.

Como estes valores estão em variação percentual em relação ao mês anterior, e a série realizada que está sendo usada neste trabalho contém o valor índice, cada valor de variação foi dividido por 100 e ao resultado foi somado um. Depois, o valor de janeiro foi multiplicado pelo valor realizado de dezembro do ano anterior da série do IGP-M, obtendo-se o valor do índice de janeiro. O valor de variação de fevereiro foi multiplicado pelo valor de janeiro obtido resultando o valor de fevereiro. E assim por diante. A comparação de previsto com realizado dos 10 anos de experimentos está na figura 3.6.

**Dólar** As estimativas do Dólar são as obtidas nas Séries Estatísticas do Sistema de Expectativas de Mercado do Bacen usando-se os parâmetros de pesquisa:

- Grupo de Indicadores = Taxas

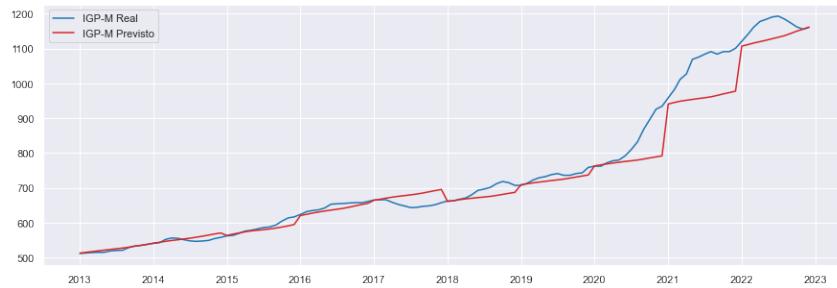


Figura 3.6: Comparação entre as previsões de mercado do Bacen e os valores reais da série do IGP-M.

- Periodicidade = Mensal
- Indicador(es) = Câmbio
- Selecione uma Base de Cálculo = Expectativas informadas nos últimos 30 dias
- Estatística = Mediana
- Período em que as projeções foram feitas:
  - Data inicial = primeiro dia útil de dezembro do ano anterior ao ano de referência do experimento
  - Data final = último dia útil de dezembro do ano anterior ao ano de referência do experimento

Foram então utilizados os valores do Dólar, de janeiro a dezembro do ano de referência do experimento, constantes última linha da tabela obtida na pesquisa, isto é, do último dia útil do mês.

A comparação de previsto com realizado dos 10 anos de experimentos está na figura 3.7.



Figura 3.7: Comparação entre as previsões de mercado do Bacen e os valores reais da série do Dólar.

**IBC-Br** O Sistema de Expectativas de Mercado do Bacen não tem estimativas do IBC-Br, mas do PIB. Entretanto, a variação trimestral do PIB foi utilizada para estimar a variação mensal do IBC-Br conforme se explica a seguir. Inicialmente, foram obtidos os valores de variação do PIB no Bacen usando-se os parâmetros de pesquisa:

- Grupo de Indicadores = Atividade
- Periodicidade = Trimestral
- Indicador(es) = PIB Total
- Selecione uma Base de Cálculo = Expectativas informadas nos últimos 30 dias
- Estatística = Mediana
- Período em que as projeções foram feitas:
  - Data inicial = primeiro dia útil de dezembro do ano anterior ao ano de referência do experimento
  - Data final = último dia útil de dezembro do ano anterior ao ano de referência do experimento

Para estimar o IBC-Br, foi feito o seguinte procedimento: utilizando-se a série do conjunto de treinamento do IBC-Br foi rodado um modelo SARIMA e feita a previsão do IBC-Br para o ano de referência do experimento. Este índice previsto do IBC-Br foi somado em grupos de três meses para se ter o acumulado do índice por trimestre. Então, cada mês foi dividido pelo valor acumulado do trimestre ao qual pertence para se ter a participação percentual daquele mês no IBC-Br acumulado do trimestre.

Procedimento análogo foi realizado com a série realizada do IBC-Br. Os valores de índice do ano anterior ao ano de referência foram somados em grupos de três meses de modo a se obter o índice acumulado do trimestre.

Então, da tabela obtida na pesquisa no Sistema de Expectativas, na última linha, que representa o último dia útil do mês anterior ao mês de referência, foram pegos os dados da variação do PIB de cada um dos trimestres do ano de referência. Como são variações percentuais, cada um foi multiplicado por 100 e foi somado 1 no resultado.

Cada valor de variação trimestral então foi multiplicado pelo acumulado do trimestre do ano anterior ao ano de referência. Assim, obteve-se o IBC-Br trimestral previsto do

ano de referência. Para distribuir o valor trimestral nos meses que compõem o trimestre, foi utilizada a participação percentual calculada a partir do SARIMA mencionado.

A comparação de previsto com realizado dos 10 anos de experimentos está na figura 3.8.



Figura 3.8: Comparação entre as previsões de mercado do PIB do Bacen, convertidas em IBC-Br, e os valores reais da série do IBC-Br.

**Estacionariedade** Os testes de raiz unitária ADF e KPSS que já haviam sido feitos na série do ICMS no experimento univariado foram aplicados nas séries do IGP-M, Dólar, IBC-Br e Dias Úteis.

**Sazonalidade** A análise da sazonalidade das séries do ICMS que foi feita em todos os 10 experimentos univariados utilizando-se do modelo X13-ARIMA-SEATS ([MONSELL; LYTRAS; FINDLEY, 2013](#)) foram repetidas nas séries das variáveis exógenas do modelo univariado com regressão múltipla. Foram encontradas sazonalidades nas séries, porém as séries dessazonalizadas, novamente, não foram utilizadas no modelos, pois o SARIMA e o SARIMAX lidam com a sazonalidade. O modelo Markov-Switching utilizou a técnica de dummies sazonais centradas. E a RNN-LSTM aprende por si só a lidar com a sazonalidade.

**Dummies de Intervenção** O X13-ARIMA-SEATS aponta os mesmos outliers na série do ICMS no experimento univariado com regressão múltipla, como era de se esperar. Os outliers apontados foram utilizados somente no modelo SARIMAX como *dummies* de intervenção. Nos modelos Markov-Switching e RNN-LSTM não foram usados pelos mesmos motivos que não foram usados no experimento univariado: a proposta deste trabalho é de estudar se modelos não lineares tratam melhor as mudanças de regime na série de arrecadação do que variáveis *dummy* nos modelos lineares.

**Auto-arima** A função auto-arima da biblioteca do Python *pmdarima*(SMITH et al., 2017–) foi utilizada para determinar os melhores parâmetros  $(p,d,q)$  e  $(P,D,Q)$  do modelo SARIMAX usando novamente os critérios de informação de Akaike e o de Schwarz.

**SARIMAX** O modelo SARIMAX foi executado para cada um dos experimentos univariados com regressão múltipla uma ou duas vezes, a depender dos resultados dos parâmetros  $(p,d,q)$  e  $(P,D,Q)$  definidos na execução do Auto-arima. Foram usados como exógenas as séries do IGP-M, Dólar, IBC-Br e Dias Úteis. Todas em nível e sem dessazonalizar. De modo análogo, para cada execução foi feita a previsão para 12 meses e calculados o MAPE e o RMSE, fora da amostra, comparando com o realizado no ano de referência do experimento.

**Markov-Switching** No experimento univariado com regressão múltipla o modelo de mudança markoviana executado foi o Markov-Switching Dinâmico MS-DR, assim como no experimento univariado. E, de modo análogo aos modelos SARIMAX, as previsões de mercado das séries do IGP-M e IBC-Br foram utilizadas na previsão do ICMS; as previsões com as séries reais do IGP-M e IBC-Br foram feitas somente a título de comparação de resultados. A variável Dólar não foi utilizada pois resultou em previsões piores nos experimentos realizados.

O treinamento foi feito com a série de treinamento. Cada modelo utilizou duas mudanças de regime e as seguintes variáveis:

- $\log(ICMS)$  em primeira diferença - Y
- Constante - mudança de regime (regime switching)
- IBC-Br realizado em primeira diferença - (regime switching)
- IGP-M realizado em segunda diferença - (regime switching)
- Dias Úteis em nível - X (variável exógena)
- 11 *dummies* sazonais centradas - X (variável exógena)

Assim como no experimento univariado, a diferença entre os três modelos está na variação da variância. O modelo *fixed* usa a variância fixa, com mudança de regime somente na média. O modelo *switching variance* varia tanto a média quanto a variância,

mas de maneira dependente, onde cada um dos dois regimes possui sua média e variância fixas. O modelo *mean variance component* varia tanto a média como a variância, mas de maneira independente, tendo um regime para médias e outro para variâncias.

As séries utilizadas foram tornadas estacionárias para uso no modelo. Segundo os testes ADF e KPSS realizados, Dias Úteis é estacionária em nível, ICMS e IBC-Br são estacionárias em primeira diferença mas IGP-M e Dólar necessitam de dias diferenciações para se tornarem estacionárias.

Conforme (FRANCQ; ZAKOÏAN, 2001), as condições de estacionariedade local, i.e., estacionariedade dentro de cada regime de um modelo de mudança markoviana, não são necessárias para obtenção da estacionariedade global. Os autores mostram que as condições de estacionariedade e a função de autocovariância deste processo não linear coincidem com os de um ARMA padrão. Outros autores, como (STELZER, 2009) e (AKNOUCHE; FRANCQ, 2022) discutem a estacionariedade e a ergodicidade dos modelos Markov-Switching ARMA chegando a conclusões semelhantes. E (NELSON; PIGER; ZIVOT, 2001) afirmam que o teste ADF, embora padeça de algumas distorções quando utilizado sob mudança de regime markoviano, pode ser usado para testar a presença de raiz unitária nestes modelos.

Estes experimentos também foram executados no software OxMetrics 8<sup>20</sup>, usando o pacote PcGive, pelas mesmas razões expostas no experimento univariado.

As previsões de 12 meses foram feitas fora da amostra e comparadas com os valores realizados constantes da série de testes do ano de referência do experimento. Foram computadas as métricas MAPE 2.1 e o RMSE 2.3.

**RNN-LSTM** Para os modelos RNN-LSTM univariados com regressão múltipla foi usada a série do ICMS como variável alvo (*target*) e as séries das variáveis exógenas IGP-M, IBC-Br, Dólar e Dias Úteis na matriz de características (*features*) - contemporâneas, com uma defasagem e com duas defasagens - totalizando 12 colunas.

Como treinamento e validação, foram utilizadas como features as séries reais das variáveis exógenas. Como testes, foram utilizadas as versões previstas das séries.

De modo análogo ao experimento univariado, foram testadas diferentes arquiteturas com diferentes números de camadas e de neurônios. Além disso, diversos hiperparâmetros foram testados: funções de ativação, tamanho de lote (*batch*), algoritmo de ajuste de

---

<sup>20</sup><https://www.doornik.com/products.html>

pesos, taxa de aprendizado, etc.

Chegou-se a duas arquiteturas principais, semelhantes às do experimento univariado, uma com uma camada LSTM e outra com três camadas LSTM. Nas camadas LSTM, a função de ativação para todas foi a tangente hiperbólica. Em todas as camadas LSTM foi usada a regularização por meio de *dropout*. As duas arquiteturas têm uma camada Densa na última camada, com um neurônio e função de ativação linear. A função de perda das duas arquiteturas é o MAPE e o otimizador Adam com taxa de aprendizado, em geral, de  $10^{-5}$ . Ver a seção 3.1.3 para maiores detalhes.

Os tamanho do lote variou entre 4 e 16 em cada experimento. O número de épocas foi fixado em um valor alto, 1000 épocas. Mas, para evitar sobre ajuste, foi usada a técnica de *parar antes* (*early stop*). Se a rede ficasse mais de 100 épocas sem melhorar a perda em validação, o treinamento era parado. E, para tentar fazer a rede aprender mais, foi usada a técnica de reduzir a taxa de aprendizado caso a rede encontrasse um platô (*reduce LR on plateau*). No caso, em 10% a cada cem épocas sem melhora da perda em validação.

A decisão de qual arquitetura usar para cada experimento univariado foi tomada a partir dos resultados do MAPE sobre o conjunto de validação. Foram treinadas redes nas duas arquiteturas para cada modelo com o MAPE calculado sobre o conjunto de validação. A arquitetura com melhor MAPE em validação de cada modelo foi a selecionada para ser a utilizada no experimento do ano de referência e ter o MAPE e o RMSE calculados e comparados com o conjunto de testes.

Para execução das redes novamente foram utilizadas as bibliotecas *Keras*<sup>21</sup> e *TensorFlow*<sup>22</sup>.

### 3.2.3 Escolha dos Melhores Modelos

Ao final de cada experimento alguns modelos foram executados por estratégia. No experimento univariado, podem haver dois modelos SARIMA e dois modelos SARIMAX resultantes das escolhas realizadas utilizando-se o Auto-arima. Ou não, pois o Auto-arima pode ter gerado os mesmos parâmetros  $(p,d,q)$  e  $(P,D,Q)$ . Podem haver três modelos Markov-Switching com diferentes variações de média e variância. Ou não, pois às vezes o modelo não converge. E há sempre dois modelos de RNN-LSTM, um que usou 1 mês

---

<sup>21</sup><https://keras.io>

<sup>22</sup><https://www.tensorflow.org>

para a previsão e outro que usou 12 meses.

De modo análogo, a estratégia univariada com regressão múltipla pode resultar em dois modelos SARIMAX e três modelos Markov-Switching. Mas apenas 1 modelo RNN-LSTM.

Foi necessário, então, definir-se um critério para a escolha de apenas um modelo por estratégia. Escolher aqueles com menor MAPE estava fora de questão, pois se estaria fazendo um exercício de futurologia pois para se ter o MAPE é necessário ter os dados da realização do ICMS, algo que não estaria disponível no caso real, onde as previsões estariam sendo feitas no ano anterior ao ano de referência.

Esta questão foi resolvida da seguinte maneira. Para os modelos estatísticos, onde há disponibilidade dos critérios de informação, foram escolhidos os modelos com o menor critério de informação de Schwarz (BIC) de cada estratégia. Assim, quando havia dois SARIMAX, aquele com menor BIC foi o escolhido. Quando havia três Markov-Switching, idem, o com menor SIC - *Schwarz Information Criteria* - foi o escolhido.

No caso das redes univariadas, onde não há critério de informação, o critério de escolha foi o menor MAPE em validação. Assim, a rede que apresentou o menor MAPE no conjunto de dados de validação foi a escolhida.

Deste modo restou um modelo representando cada estratégia em cada experimento.

### 3.2.4 Model Confidence Set

O Model Confidence Set ainda não se encontra disponível em nenhuma biblioteca Python. Por este motivo, sua execução foi feita no R, utilizando-se o pacote MCS ([BERNARDI; CATANIA, 2018](#)).

Como função de perda foi utilizado o MAPE de cada modelo. O RMSE foi descartado porque ele aumenta naturalmente com o aumento da arrecadação que ocorre no passar dos anos. Uma matriz com 8 linhas, cada uma com uma das 8 estratégias, 5 do experimento univariado - Naïve, SARIMA, SARIMAX, Markov-Switching e RNN-LSTM - e 3 do experimento univariado com regressão múltipla - SARIMAX, Markov-Switching e RNN-LSTM. E 10 colunas, cada uma com os valores do MAPE de 2013 a 2022.

# Capítulo 4

## Resultados

Este capítulo apresenta os resultados dos experimentos. A seção 4.1 apresenta os principais resultados da análise descritiva das séries de tempo. A seção 4.2 apresenta os resultados do experimento univariado. A seção 4.3 apresenta os resultados do experimento univariado com regressão múltipla. E, finalizando o capítulo, a seção 4.4 compara os resultados obtidos.

As séries temporais utilizadas podem ser vistas na figura 4.1, que mostra todas as séries em nível. A série do ICMS está em log. Os gráficos do IGP-M, IBC-Br e Dólar apresentam a série realizada de janeiro de 2003 a dezembro de 2022 e a série prevista de janeiro de 2022 a dezembro de 2022.

### 4.1 Análise dos Dados

Esta seção irá apresentar os resultados obtidos na análise descritiva dos dados para os experimentos univariado e univariado com regressão múltipla de 2022. Os resultados dos demais experimentos se encontram no Github deste trabalho<sup>1</sup>.

#### 4.1.1 Estacionariedade

Os testes de raiz unitária realizados, ADF e KPSS, apresentaram os resultados listados na tabela 4.1 para o ano de 2022, que foram avaliados segundo os critérios apresentados na seção 3.1.1.

---

<sup>1</sup><[https://github.com/asbonato/mpe\\_fgv\\_2023](https://github.com/asbonato/mpe_fgv_2023)>

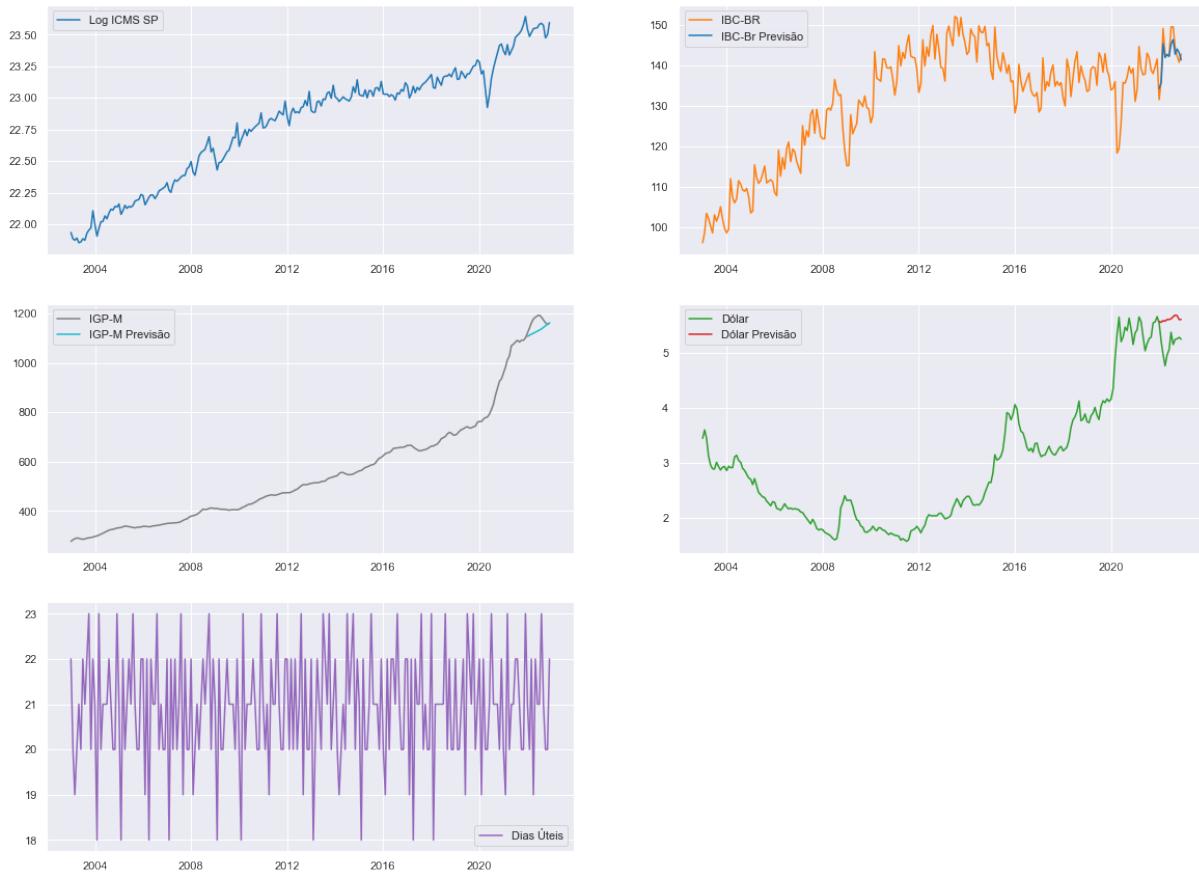


Figura 4.1: Séries temporais utilizadas nos experimentos de 2022.

Série	ADF	KPSS	Estacionariedade
ICMS	Não rejeita $H_0$	Rejeita $H_0$	Não estacionária.
IBC-BR	Não rejeita $H_0$	Rejeita $H_0$	Não estacionária.
IGP-M	Não rejeita $H_0$	Rejeita $H_0$	Não estacionária.
Dólar	Não rejeita $H_0$	Rejeita $H_0$	Não estacionária.
Dias Úteis	Rejeita $H_0$	Não rejeita $H_0$	Estacionária.

Tabela 4.1: Resultados dos testes ADF e KPSS para as séries em nível.

Foram feitos os testes em primeira diferença das séries que não foram estacionárias em nível. Os resultados estão na tabela 4.2.

Série	ADF	KPSS	Estacionariedade
ICMS	Rejeita $H_0$	Não rejeita $H_0$	Estacionária.
IBC-BR	Rejeita $H_0$	Não rejeita $H_0$	Estacionária.
IGP-M	Rejeita $H_0$	Rejeita $H_0$	Não estacionária.
Dólar	Rejeita $H_0$	Rejeita $H_0$	Não estacionária.

Tabela 4.2: Resultados dos testes ADF e KPSS para as séries em primeira diferença.

As séries colocadas em primeira diferença estão representadas na figura 4.2.

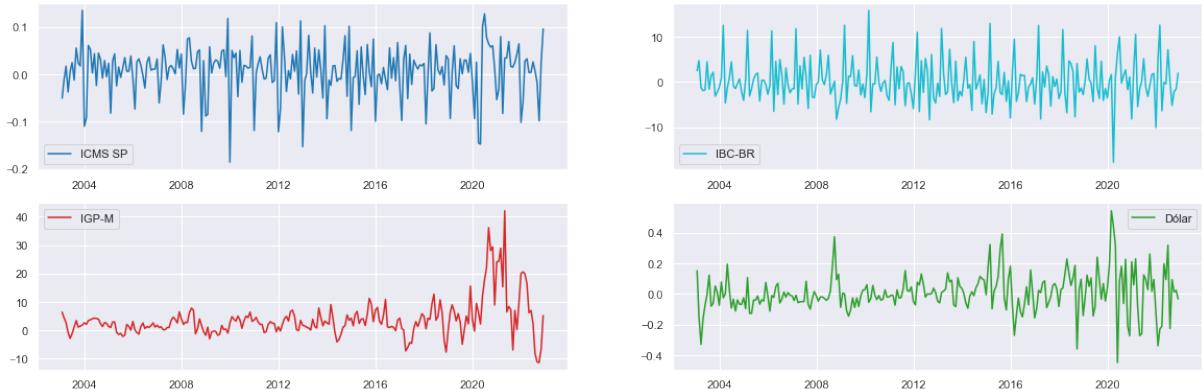


Figura 4.2: Séries em primeira diferença.

Foram feitos os testes em segunda diferença das séries que não foram estacionárias em primeira diferença. Os resultados estão na tabela 4.3.

Série	ADF	KPSS	Estacionariedade
IGP-M	Rejeita $H_0$	Não rejeita $H_0$	Estacionária.
Dólar	Rejeita $H_0$	Não rejeita $H_0$	Estacionária.

Tabela 4.3: Resultados dos testes ADF e KPSS para as séries em segunda diferença.

Apenas a série dos Dias Úteis é estacionária em nível. Todas as demais somente são estacionárias em diferença.

As séries em segunda diferença estão representadas na figura 4.3.



Figura 4.3: Séries em segunda diferença.

#### 4.1.2 Sazonalidade

O estudo de sazonalidade das séries, feito utilizando-se o método X13 Arima Seats mostrou sazonalidade na série do ICMS, como pode ser verificado na figura 4.4.

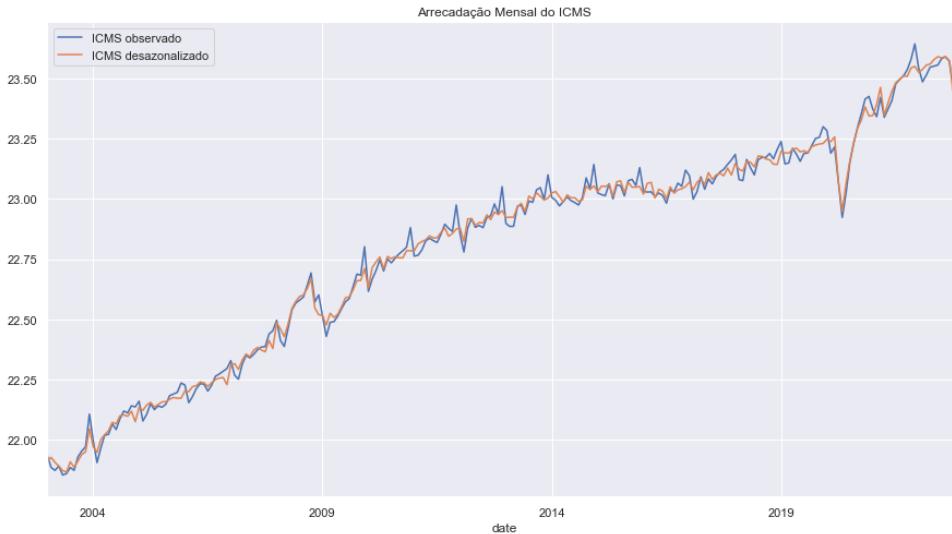


Figura 4.4: Série do ICMS em nível, com e sem sazonalidade

As séries do IBC-Br e de Dias Úteis também apresentaram sazonalidade. Já as séries do IGP-M e do Dólar apresentaram sazonalidade muito baixa, como se observa na figura 4.5.

### 4.1.3 Outliers

O algoritmo do ARIMA Seats X13 encontrou 5 outliers automaticamente na série do ICMS, todos estatisticamente significativos a 1%. Estes outliers possuem explicação econômica, conforme a tabela 4.4 e foram utilizados nos modelos SARIMAX dos experimentos univariado e univariado com regressão múltipla como *dummies* de intervenção.

As *dummies* LS são do tipo Mudança de Nível, ou *Level Shift*. Elas valem zero até a um mês antes da data de início e valem um da data de início até o final da série. A dummy AO é um outlier aditivo, ou *Additive Outlier*. Ela vale um apenas no mês indicado.

Variable	Estimate	Error	t-value	Explicação
LS2008.Nov	-0.0065	0.00113	-5.75	Crise do Subprime
LS2020.Apr	-0.0073	0.00120	-6.12	V da Covid - descida
AO2020.Mai	-0.0063	0.00124	-5.03	Vale da Covid
LS2020.Ago	0.0059	0.00114	5.17	V da Covid - subida
LS2022.Oct	-0.0059	0.00122	-4.88	Mudanças legais na alíquota do ICMS

Tabela 4.4: Outliers detectados automaticamente

Obs.: os períodos das *dummies* estão identificados no seu nome.



Figura 4.5: Sériess em nível, com e sem sazonalidade: IBC-Br, IGP-M e Dólar

As mudanças legais a que se refere a variável LS2022.Oct estão relacionadas à Lei Complementar 194/2022<sup>2</sup>, que impôs um limite de cobrança do ICMS nos serviços essenciais e causou um grande impacto nas alíquotas do estado de São Paulo, que eram de 25%, mas que depois da lei não podem passar de 18%.

#### 4.1.4 Teste de Causalidade de Granger

Os testes de causalidade de Granger foram realizados entre a série do ICMS e as séries das variáveis exógenas. Para a realização dos testes foram usadas as séries dessazonalizadas e estacionarizadas, colocando-as em diferenças conforme 4.1.1.

Os resultados obtidos foram:

- IBC-Br granger causa ICMS a partir do lag 1 e ICMS não granger causa IBC-Br
- IGP-M granger causa ICMS a partir do lag 2 e ICMS não granger causa IGP-M.

<sup>2</sup>disponível em <[https://www.planalto.gov.br/ccivil\\_03/leis/lcp/lcp194.htm](https://www.planalto.gov.br/ccivil_03/leis/lcp/lcp194.htm)>

- Dólar granger causa ICMS a partir do lag 3 e ICMS granger causa Dólar a partir do lag 1.
- Dias Úteis granger causa ICMS a partir do lag 1 e ICMS não granger causa Dias Úteis.

## 4.2 Experimento Univariado

Este experimento teve como objetivo realizar uma análise univariada da série de arrecadação do ICMS paulista. Foi utilizada a série do ICMS com sazonalidade e em nível, deixando para os modelos dessazonalizarem a série e torná-la estacionária. Para os modelos SARIMA, SARIMAX e modelos de Mudança Markoviana, a série foi separada em amostras de treinamento (*in-sample*) e de teste. Para o modelo de Rede Neural Recorrente LSTM a amostra foi separada em treinamento, validação e teste. Mais detalhes sobre a separação dos conjuntos de dados na subseção 3.2.2.

### 4.2.1 Modelo Naïve

Um modelo simples, denominado modelo naïve, foi criado para fins de comparação com os modelos de previsão. Este modelo utilizou a arrecadação média do ano anterior e considerou este valor como o previsto para os demais meses. Em 2021 a arrecadação média foi de R\$ 15.47 bilhões.

### 4.2.2 Modelos SARIMA e SARIMAX

Foram utilizados os modelos SARIMA e SARIMAX. Ambos conseguem trabalhar com séries sazonais e não estacionárias, pois estes modelos dessazonalizam e diferenciam a série por si próprios. A vantagem desta abordagem é que as previsões apresentadas já estão em nível e contém sazonalidade. E a diferença entre ambos é que o SARIMA não emprega variáveis exógenas, ao contrário do SARIMAX. Para a previsão de 12 meses, então, foi utilizada a série de treinamento, uma vez sem as *dummies* de intervenção - modelo SARIMA - e depois com as *dummies* de intervenção - modelo SARIMAX.

## Auto-arima dos Modelos sem *Dummies*

Utilizando-se a ferramenta Auto-arima para determinar o melhor modelo ARIMA baseando-se em critérios de informação, obtiveram-se os resultados reportados a seguir.

- Critério de Akaike AIC: SARIMA(0,1,1),(2,0,1)
- Critério de Schwarz BIC: SARIMA(0,1,1)(1,0,1)

Ambos apresentaram as mesmas defasagens de AR e MA, e a mesma diferenciação. A mudança foi apenas no componente AR da sazonalidade.

### Modelo SARIMA(0,1,1),(2,0,1) sugerido pelo critério de Akaike

O teste de Ljung-Box tem estatística 0.00, mostrando uma baixa correlação dos resíduos. São significativos a 5% os coeficientes do lag 1 da média móvel e os coeficientes do lag 12 e 24 do AR e do lag 12 do MA, todos do ARIMA sazonal. O sumário dos resultados está na figura 4.6.

SARIMAX Results						
Dep. Variable:	y	No. Observations:	228			
Model:	ARIMA(0, 1, 1)x(2, 0, 1, 12)	Log Likelihood	394.129			
Date:	Wed, 19 Apr 2023	AIC	-778.258			
Time:	18:01:51	BIC	-761.133			
Sample:	01-01-2003 - 12-01-2021	HQIC	-771.348			
Covariance Type:	opg					
	coef	std err	z	P> z	[0.025	0.975]
ma.L1	-0.2707	0.044	-6.111	0.000	-0.357	-0.184
ar.S.L12	1.2834	0.128	10.050	0.000	1.033	1.534
ar.S.L24	-0.2847	0.111	-2.564	0.010	-0.502	-0.067
ma.S.L12	-0.9668	0.262	-3.688	0.000	-1.481	-0.453
sigma2	0.0017	0.000	5.862	0.000	0.001	0.002
Ljung-Box (L1) (Q):	0.00	Jarque-Bera (JB):			146.38	
Prob(Q):	0.99	Prob(JB):			0.00	
Heteroskedasticity (H):	1.64	Skew:			-0.75	
Prob(H) (two-sided):	0.03	Kurtosis:			6.64	

Figura 4.6: Sumário do Modelo SARIMA(0,1,1)(2,0,1)

A análise dos resíduos mostra que não há autocorrelação significativa, que os resíduos são estacionários e tem uma distribuição próxima à normal, conforme a figura 4.7.

### Modelo SARIMA(0,1,1),(1,0,1) sugerido pelo critério de Schwarz

O teste de Ljung-Box tem estatística 0.00, mostrando uma baixa correlação dos resíduos. São significativos a 5% os coeficientes do lag 1 da média móvel e os coeficientes sazonais do lag 12 tanto do AR quanto do MA. O sumário dos resultados está na figura 4.8.

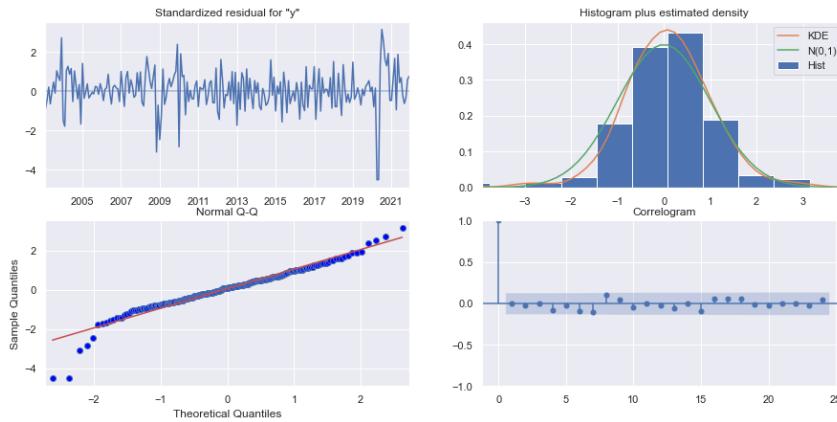


Figura 4.7: Análise dos resíduos do Modelo SARIMA(0,1,1)(2,0,1)

```
=====
Dep. Variable:                      y      No. Observations:                  228
Model:                 ARIMA(0, 1, 1)x(1, 0, 1, 12)   Log Likelihood:          391.426
Date:                   Wed, 19 Apr 2023   AIC:                         -774.851
Time:                       18:01:53      BIC:                         -761.151
Sample:             01-01-2003 - 12-01-2021   HQIC:                        -769.323
Covariance Type:                opg
=====
              coef    std err      z   P>|z|      [ 0.025   0.975 ]
ma.L1     -0.2783    0.039   -7.228   0.000    -0.354   -0.203
ar.S.L12    0.9075    0.045   20.136   0.000     0.819    0.996
ma.S.L12   -0.5914    0.086   -6.845   0.000    -0.761   -0.422
sigma2     0.0018    0.000   16.059   0.000     0.002    0.002
=====
Ljung-Box (L1) (Q):                  0.00   Jarque-Bera (JB):        122.22
Prob(Q):                           0.99   Prob(JB):                  0.00
Heteroskedasticity (H):               1.67   Skew:                     -0.74
Prob(H) (two-sided):                0.03   Kurtosis:                  6.28
=====
```

Figura 4.8: Sumário do Modelo SARIMA(0,1,1)(1,0,1)

A análise dos resíduos mostra que não há autocorrelação significativa, que os resíduos são estacionários e tem uma distribuição próxima à normal, conforme a figura 4.9.

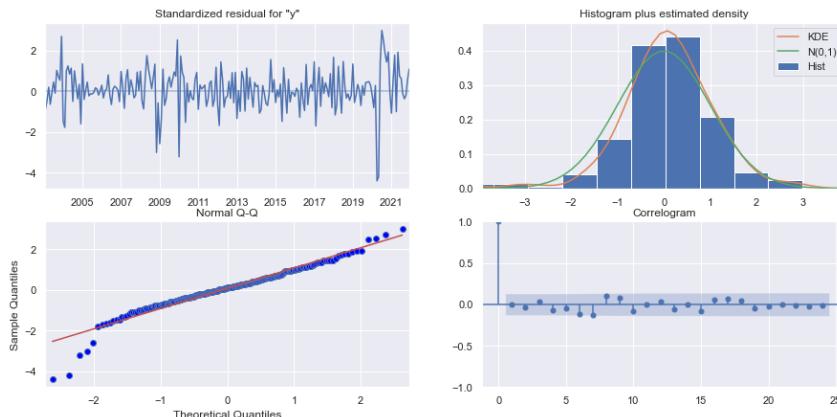


Figura 4.9: Análise dos resíduos do Modelo SARIMA(0,1,1)(1,0,1)

### Auto-arima dos Modelos com *Dummies*

Utilizando-se a ferramenta Auto-arima para determinar o melhor modelo ARIMA baseando-se em critérios de informação, considerando agora o uso das *dummies* de intervenção, obtiveram-se os resultados reportados a seguir.

- Critério de Akaike AIC: SARIMAX(0,1,1),(2,0,0)
- Critério de Schwarz BIC: SARIMAX(0,1,1)(1,0,1)

Ambos apresentaram as mesmas defasagens de AR e MA dos modelos sem *dummies*, e a mesma diferenciação. A mudança aconteceu nos componentes AR e MA da sazonalidade.

### Modelo SARIMAX(0,1,1),(2,0,0) sugerido pelo critério de Akaike

O teste de Ljung-Box tem estatística 0.67, mostrando há correlação dos resíduos. E são significativos a 5% os coeficientes do lag 1 da média móvel e os coeficientes sazonais do lag 12 e 24 do AR. O sumário dos resultados está na figura 4.10.

No sumário as variáveis exógenas são referenciadas por  $x_1, \dots, x_4$ . Esta referência é dada pela ordem em que as variáveis aparecem na matrix de exógenas. Assim:

- $x_1$  é a *dummy* LS2008.Nov
- $x_2$  é a *dummy* LS2020.Abr
- $x_3$  é a *dummy* AO2020.Mai
- $x_4$  é a *dummy* LS2020.Ago

As quatro *dummies* LS2008.Nov, LS2020.Apr, AO2020.Mai e LS2020.Ago são estatisticamente significativas. Já a *dummy* LS2022.Out não foi considerada por valer zero em todas as observações do conjunto de dados de treinamento.

A análise dos resíduos mostra que há autocorrelação significativa no lag 7, que os resíduos são estacionários e tem uma distribuição próxima à normal, conforme a figura 4.11.

SARIMAX Results						
Dep. Variable:	y	No. Observations:	228			
Model:	ARIMA(0, 1, 1)x(2, 0, [ ], 12)	Log Likelihood	432.284			
Date:	Mon, 24 Apr 2023	AIC	-846.568			
Time:	01:58:56	BIC	-815.743			
Sample:	01-01-2003 - 12-01-2021	HQIC	-834.130			
Covariance Type:	opg					
coef	std err	z	P> z	[0.025	0.975]	
x1	-0.1445	0.027	-5.390	0.000	-0.197	-0.092
x2	-0.1628	0.026	-6.176	0.000	-0.214	-0.111
x3	-0.1386	0.042	-3.264	0.001	-0.222	-0.055
x4	0.1318	0.029	4.479	0.000	0.074	0.189
const	0	2.83e-08	0	1.000	-5.55e-08	5.55e-08
ma.L1	-0.4858	0.051	-9.539	0.000	-0.586	-0.386
ar.S.L12	0.5432	0.062	8.814	0.000	0.422	0.664
ar.S.L24	0.2198	0.078	2.836	0.005	0.068	0.372
sigma2	0.0012	0.000	11.140	0.000	0.001	0.001
Ljung-Box (L1):	0.67	Jarque-Bera (JB):	4.21			
Prob(Q):	0.41	Prob(JB):	0.12			
Heteroskedasticity (H):	0.90	Skew:	-0.12			
Prob(H) (two-sided):	0.64	Kurtosis:	3.62			

Figura 4.10: Sumário do Modelo SARIMAX(0,1,1)(2,0,0)

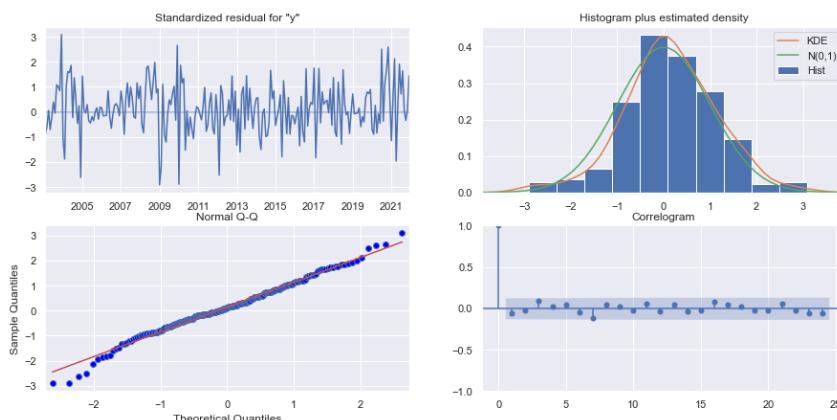


Figura 4.11: Correlogramas dos resíduos do Modelo SARIMAX(0,1,1)(2,0,0)

### Modelo SARIMAX(0,1,1),(1,0,1) sugerido pelo critério de Schwartz

O teste de Ljung-Box tem estatística 0.82, mostrando que há correlação dos resíduos. E são significativos a 5% os coeficientes do lag 1 da média móvel e os coeficientes sazonais do lag 12 e 24 do AR. O sumário dos resultados está na figura 4.12.

As quatro *dummies* LS2008.Nov, LS2020.Apr, AO2020.Mai e LS2020.Ago são estatisticamente significativas. Já a *dummy* LS2022.Out não foi considerada por valer zero em todas as observações do conjunto de dados de treinamento.

A análise dos resíduos mostra que ainda há autocorrelação significativa no lag 7, que os resíduos são estacionários e tem uma distribuição próxima à normal, conforme a figura 4.13.

Dep. Variable:	y	No. Observations:	228			
Model:	ARIMA(0, 1, 1)x(1, 0, 1, 12)	Log Likelihood	433.782			
Date:	Wed, 19 Apr 2023	AIC	-849.564			
Time:	18:03:56	BIC	-818.740			
Sample:	01-01-2003 - 12-01-2021	HQIC	-837.126			
Covariance Type:	opg					
	coef	std err	z	P> z	[0.025	0.975]
x1	-0.1380	0.030	-4.640	0.000	-0.196	-0.080
x2	-0.1638	0.028	-5.778	0.000	-0.219	-0.108
x3	-0.1385	0.043	-3.239	0.001	-0.222	-0.055
x4	0.1366	0.030	4.483	0.000	0.077	0.196
const	0	3134.010	0	1.000	-6142.547	6142.547
ma.L1	-0.4972	0.051	-9.672	0.000	-0.598	-0.396
ar.S.L12	0.9052	0.039	23.297	0.000	0.829	0.981
ma.S.L12	-0.4574	0.091	-5.038	0.000	-0.635	-0.279
sigma2	0.0012	0.000	11.186	0.000	0.001	0.001
Ljung-Box (L1) (Q):	0.82	Jarque-Bera (JB):	4.03			
Prob(Q):	0.36	Prob(JB):	0.13			
Heteroskedasticity (H):	0.96	Skew:	-0.06			
Prob(H) (two-sided):	0.86	Kurtosis:	3.64			

Figura 4.12: Sumário do Modelo SARIMAX(0,1,1)(1,0,1)

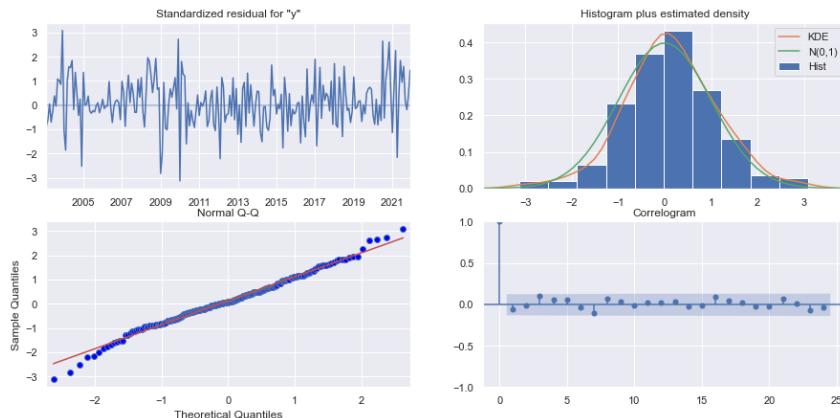


Figura 4.13: Correlogramas dos resíduos do Modelo SARIMAX(0,1,1)(1,0,1)

### 4.2.3 Modelos de Mudança Markoviana

Os modelos de mudança markoviana, conforme detalhado na metodologia, usaram apenas a série do ICMS em primeira diferença no experimento univariado, sendo ICMS em primeira diferença com uma defasagem a variável de mudança de regime. Foram usadas 2 mudanças de regime e não foram utilizadas as *dummies* de intervenção, apenas *dummies* sazonais centradas. Cada um dos três modelos usou um tipo de variação na variância: variância fixa (*fixed variance*), mudança de variância (*switching variance*) e componente média-variância (*mean-variance component*).

## Modelo de Mudança Markoviana com Variância Fixa

Este modelo mostrou praticamente apenas um regime, o regime 1. A série ficou 212 meses neste regime e apenas 14 meses no regime 0. O coeficiente da variável de mudança em ambos os regimes é parecido em intensidade, mas tem o sinal trocado:  $diff(\log(icms))$  com uma defasagem tem o valor 0.56 no regime 0 e  $-0.53$  no regime 1. Todas as variáveis, inclusive as *dummies* sazonais centradas, foram significativas a 5%, exceto a constante no regime 0 e a dummy número 3. A variância do modelo foi 0.034. Portanto, o regime mais duradouro é o de menor média.

A análise dos resíduos mostra que não há normalidade e há autocorrelação significativa, conforme pode ser visto na figura 4.14 e no resultado dos testes no apêndice B.

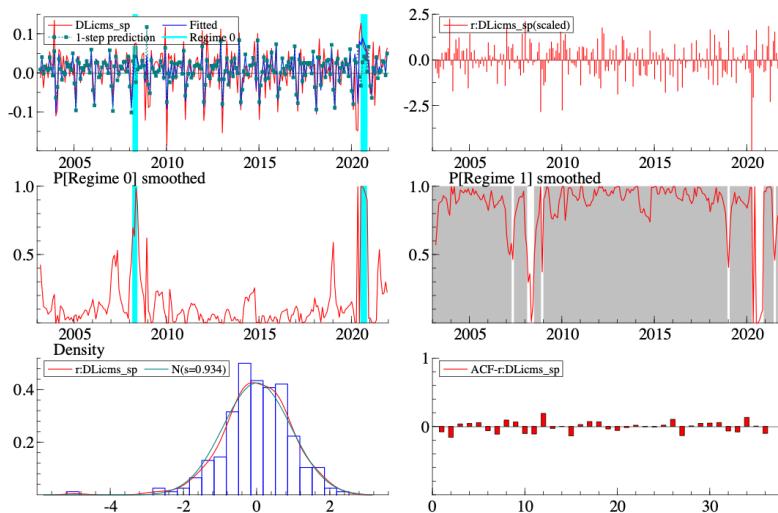


Figura 4.14: Resultados do modelo de Markov com Variância Fixa

Os sumários de cada modelo estão no Apêndice B.

## Modelo de Mudança Markoviana com Mudança de Variância

Este modelo mostrou alternância de regimes, com predominância do regime 0. A série ficou 196 meses neste regime e apenas 30 meses no regime 1. O coeficiente da variável de mudança,  $diff(\log(icms))$  com uma defasagem, é  $-0.51$  no regime 0 e 0.089 no regime 1. Todas as variáveis, inclusive as *dummies* sazonais centradas (com exceção à de número 4), foram significativas a 5%, exceto a constante no regime 1. A variância do modelo foi 0.03 no regime 0 e 0.07 no regime 1. Portanto, o regime mais duradouro é o que tem menor média e menor variância.

A análise dos resíduos mostra que não há normalidade e que há autocorrelação significativa, conforme a figura 4.15 e o resultado dos testes no apêndice B.

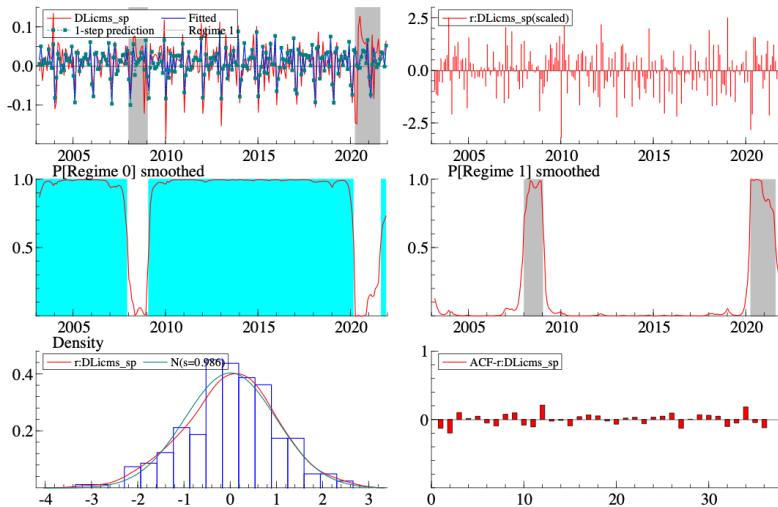


Figura 4.15: Resultados do modelo de Markov com Mudança de Variância

### Modelo de Mudança Markoviana com Componente Média-Variância

Este modelo mostrou alternância nos regimes de média, mas com predominância do regime 0, e total dominância do regime 0 de variância. A série ficou os 134 meses no regime 0 de média e 210 meses no regime 0 de variância, e 92 meses no regime 1 de média e apenas 16 meses no regime 1 de variância. Em ambos os regimes o coeficiente da variável de mudança,  $\text{diff}(\log(icms))$  com uma defasagem, foi negativo:  $-0.13$  no regime 0 (embora não estatisticamente significativa) e, no regime 1,  $-0.67$ . Todas as variáveis, inclusive as *dummies* sazonais centradas, foram significativas a 5%, exceto a variável de mudança no regime 0 e a dummy de número 4. A variância do modelo foi 0.029 no regime 0 e 0.086 no regime 1. Portanto, o regime mais duradouro é o que tem maior média e maior variância.

A análise dos resíduos mostra normalidade, porém há autocorrelação significativa, conforme figura 4.16 e testes realizados.

#### 4.2.4 Modelo RNN LSTM

Foram testadas diversas arquiteturas de rede com diversos hiperparâmetros. A arquitetura que apresentou melhores resultados em validação, em ambos os experimentos, foi a arquitetura de uma camada. O modelo RNN-LSTM-1-mês teve MAPE de 10.97% e o

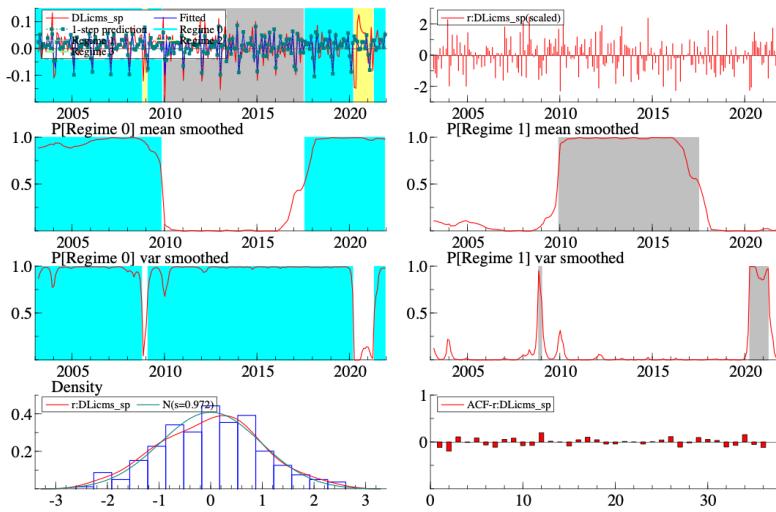


Figura 4.16: Resultados do modelo de Markov com Mudança de Média e Variância

modelo RNN-LSTM-12-Meses obteve MAPE de 9.56%, contra um MAPE de 20.75% para o modelo RNN-LSTM-1-mês com arquitetura com 3 camadas escondidas, e um MAPE de 24.72% para o modelo RNN-LSTM-12-meses, também com arquitetura com 3 camadas escondidas.

A arquitetura e os hiperparâmetros escolhidos foram:

- uma camada escondida com 600 neurônios LSTM;
- dropout de 25
- função de perda MAPE;
- otimizador Adam com taxa de aprendizagem começando em  $10^{-5}$ ;
- camada densa de saída com 1 neurônio com função de ativação linear.
- 700 épocas com *early-stop* com 100 épocas de paciência.
- redução da taxa de aprendizagem em 10% em caso de atingimento de platô, com paciência de 10 épocas.

### Modelo RNN-LSTM 1 mês

O modelo rodou 108 épocas até parar. O resultado da perda em treino e validação está na figura 4.17. A curva não aparenta nem sub ajuste (*underfit*), e nem sobre ajuste (*overfit*).

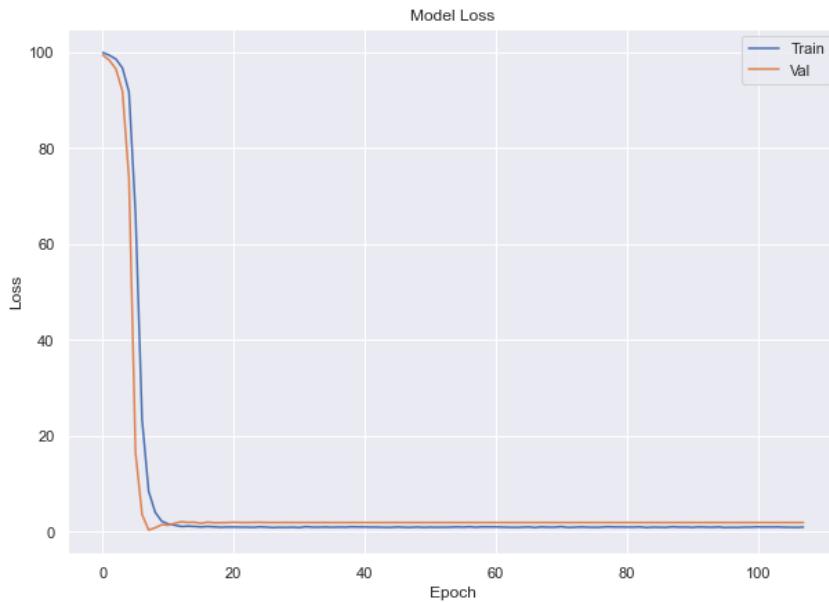


Figura 4.17: Gráfico da perda em treino e validação do modelo RNN-LSTM 1 mês

### Modelo RNN-LSTM 12 meses

O modelo rodou 106 épocas até parar. O resultado da perda em treino e validação está na figura 4.18. A curva não aparenta nem sub ajuste (*underfit*), e nem sobre ajuste (*overfit*).

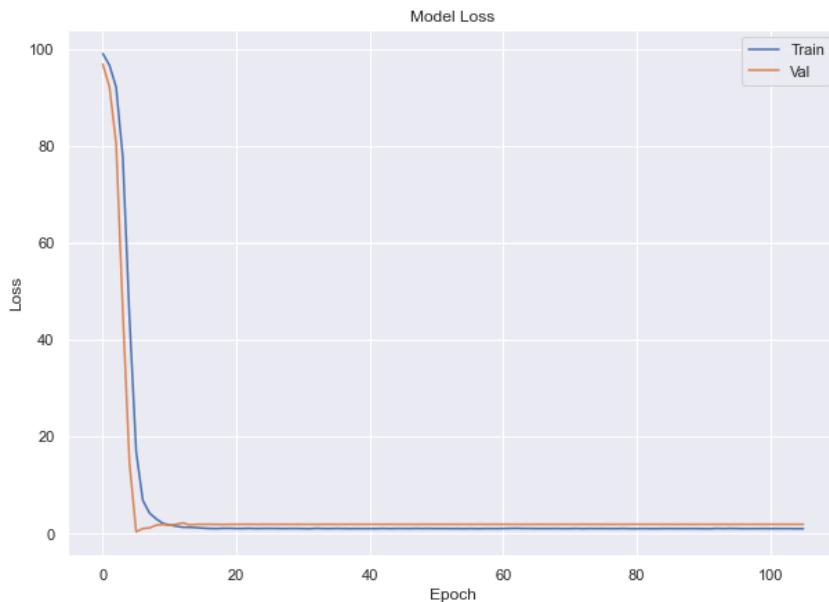


Figura 4.18: Gráfico da perda em treino e validação do modelo RNN-LSTM 12 meses

### 4.2.5 Previsões

As previsões foram realizadas sobre o conjunto de testes (*out-of-sample*). As métricas RMSE e MAPE estão na tabela 4.5.

	RMSE	MAPE
Mean	1.452835e+09	0.076073
SARIMA(0,1,1),(2,0,1)	1.866081e+09	0.088529
SARIMA(0,1,1),(1,0,1)	1.850802e+09	0.085239
SARIMAX(0,1,1),(2,0,0)	2.139219e+09	0.100347
SARIMAX(0,1,1),(1,0,1)	1.993289e+09	0.093292
Markov-Fixed	1.460961e+09	0.066127
Markov-Switching-Var	1.581809e+09	0.073473
Markov-Mean-Var	1.929173e+09	0.093325
RNN-LSTM-1-mês	7.753787e+08	0.037627
RNN-LSTM-12-meses	8.466023e+08	0.045839

Tabela 4.5: Métricas dos modelos de previsão univariados

### RMSE

Analizando-se o RMSE na tabela e na figura 4.19. Pode-se notar que:

- O menor RMSE foi do modelo RNN-LSTM de 1 mês, seguido do modelo RNN-LSTM de 12 meses.
- O terceiro menor foi o do modelo Markov-Switching com variância fixa.

### MAPE

Analizando-se o MAPE na tabela e na figura 4.20. Pode-se notar que:

- O menor MAPE foi do modelo RNN-LSTM de 12 meses, seguido do modelo RNN-LSTM de 1 mês em segundo,
- O modelo de Markov-Switching com variância fixa em terceiro.

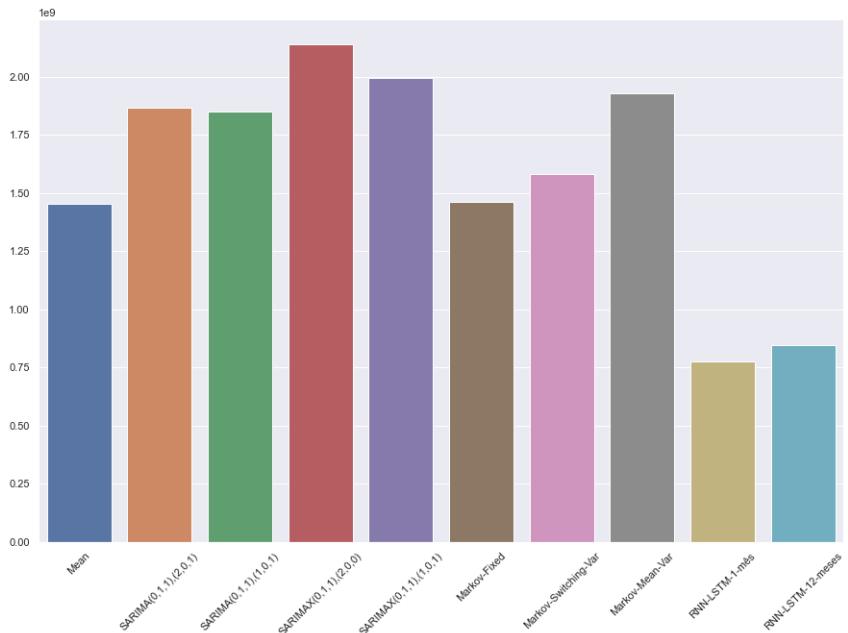


Figura 4.19: Métricas RMSE dos modelos univariados

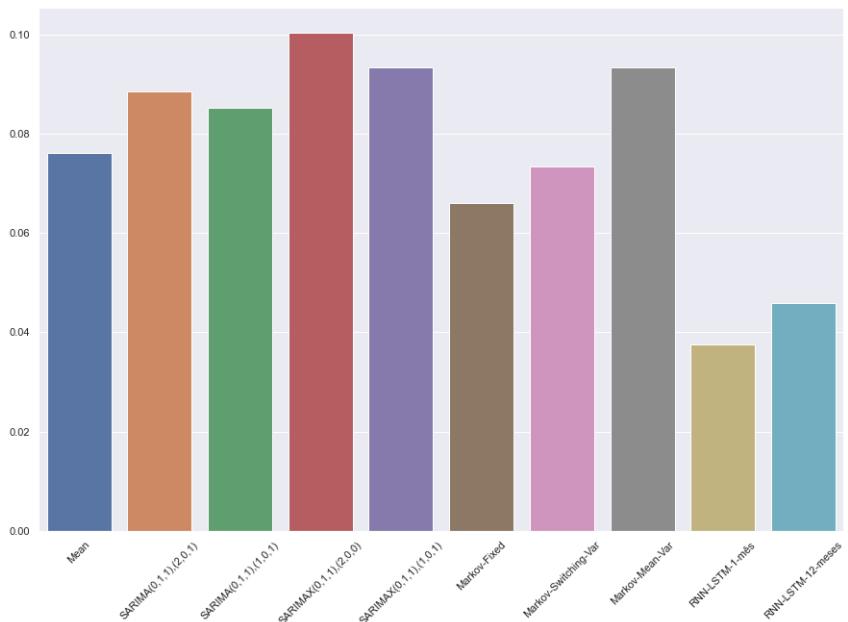


Figura 4.20: Métricas MAPE dos modelos univariados

### Comparação das Previsões

Embora os modelos RNN-LSTM tenham apresentados as melhores métricas MAPE e RMSE, são os modelos que tem menor variabilidade mês a mês no período de 12 meses, como se pode ver no gráfico da figura 4.24. Tanto os modelos SARIMA e SARIMAX, apresentados na figura 4.22 quanto os modelos de Mudança Markoviana, mostrados na figura 4.23 apresentam maior variabilidade, porém nenhum deles é capaz de prever a

queda de arrecadação ocorrida em Outubro de 2022 devido às mudanças nas alíquotas do ICMS, como já foi mencionado. Vale notar que os modelos SARIMA e SARIMAX, neste ano, obtiveram performance pior do que o modelo Naïve mostrado em 4.21.

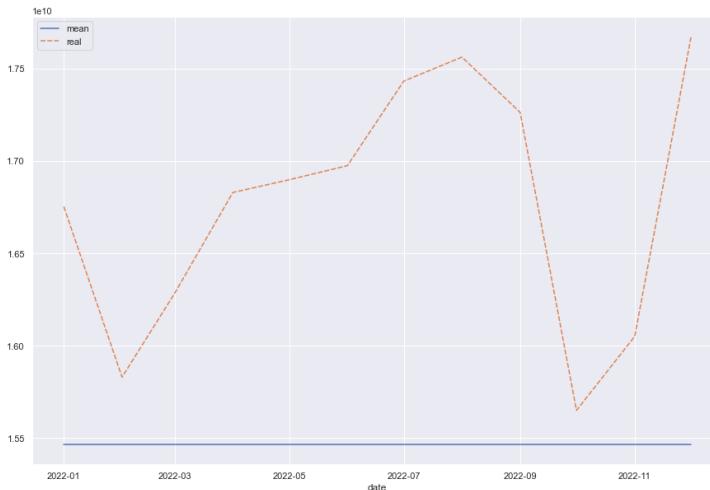


Figura 4.21: Previsões do Modelo Naïve comparada com o Realizado

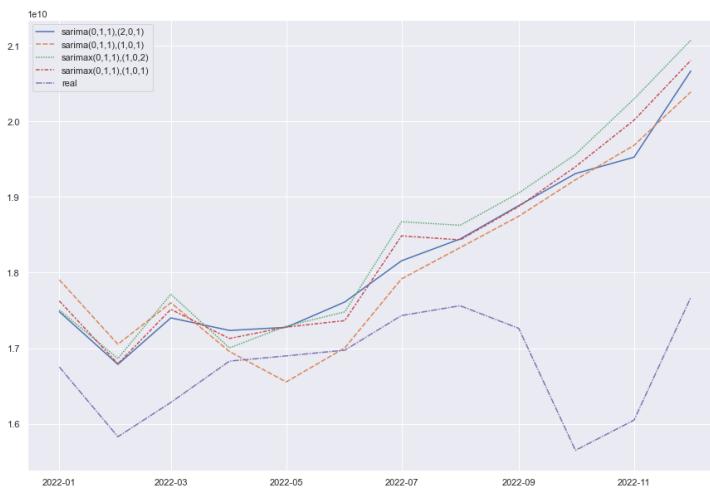


Figura 4.22: Previsões dos Modelos SARIMA e SARIMAX comparadas com o Realizado

## Seleção dos Modelos Univariados de 2022

De cada uma das estratégias foi escolhido um modelo. O critério de escolha para SARIMA, SARIMAX e Markov-Switching foi o menor Critério de Informação de Schwarz (BIC). As RNN-LSTM foi o menor MAPE em validação. Os resultados estão na tabela 4.6.

É importante notar que o menor BIC fez com que o modelo SARIMA com menor MAPE fosse escolhido, assim como o SARIMAX com menor MAPE. No caso do modelo de mudança markoviana aconteceu o oposto. O modelo de maior MAPE foi o escolhido,

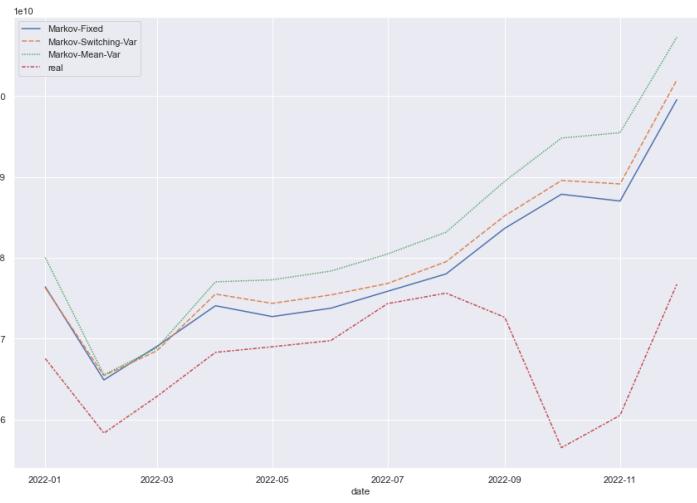


Figura 4.23: Previsões dos Modelos de Mudança Markoviana comparadas com o Realizado

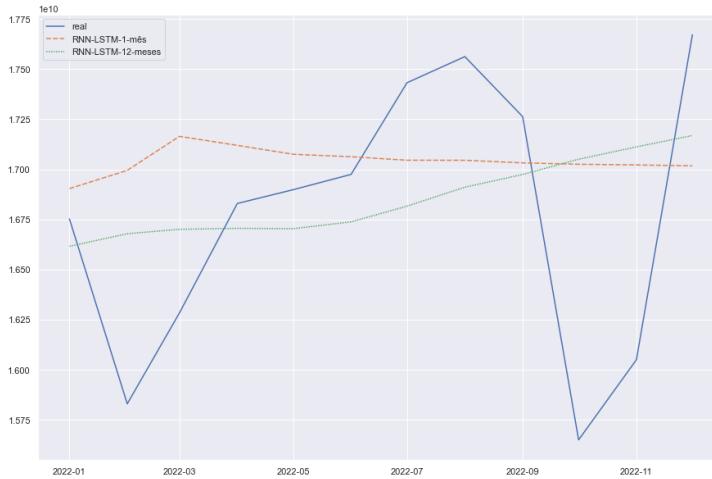


Figura 4.24: Previsões dos Modelos RNN-LSTM comparadas com o Realizado

e não o de menor MAPE. Porém, não seria possível saber isto de antemão caso a previsão tivesse sido feita no ano de 2021. De modo análogo, o menor MAPE em validação fez com que a rede neural LSTM com maior MAPE em teste fosse a escolhida. Novamente, não haveria como saber isto a priori.

## 4.3 Experimento Univariado com Regressão Múltipla

Este experimento teve como objetivo realizar uma análise univariada com regressão múltipla da série de arrecadação do ICMS paulista, para tentar melhorar as previsões, utilizando as séries do IGP-M, IBC-Br, Dólar e Dias Úteis. Nos modelos SARIMAX, todas as séries foram utilizadas com sazonalidade e em nível, deixando para os modelos dessazonalizarem a série e torná-la estacionária, de forma análoga ao experimento univariado.

Modelo	Estratégia	BIC	MAPE Val	Escolha
Mean	naive			<b>Sim</b>
SARIMA(0,1,1),(2,0,1)	sarima	-761.133000		Não
SARIMA(0,1,1),(1,0,1)	sarima	-761.151000		<b>Sim</b>
SARIMAX(0,1,1),(2,0,0)	sarimax	-815.743000		Não
SARIMAX(0,1,1),(1,0,1)	sarimax	-818.740000		<b>Sim</b>
Markov-Fixed	markov-switching	-3.29793748		Não
Markov-Switching-Var	markov-switching	-3.34189434		Não
Markov-Mean-Var	markov-switching	-3.353027446		<b>Sim</b>
RNN-LSTM-1-mês	rnn-lstm		0.1097476	Não
RNN-LSTM-12-meses	rnn-lstm		0.0955967	<b>Sim</b>

Tabela 4.6: Escolha dos modelos univariados de 2022 e seus critérios.

Já nos modelos de mudança markoviana as séries foram postas em diferença, mas não foram dessazonalizadas a priori, pois optou-se pelo uso de dummies sazonais centradas. Nos modelos RNN-LSTM não houve diferenciação nem dessazonalização, mas o modelo não demanda transformações prévias.

Nos modelos SARIMAX e modelos de Mudança Markoviana, a série foi separada em amostras de treinamento (*in-sample*) e de teste (*out-of-sample*). Para o modelo de Rede Neural Recorrente LSTM a amostra foi separada em treinamento, validação e teste. Mais detalhes sobre os conjuntos de dados na subseção [3.2.2](#).

Foram criadas duas amostras de teste, em todos os experimentos. Uma delas contém os dados reais de 2022 das variáveis exógenas - Dias Úteis, IBC-Br, IGP-M e Dólar. A outra contém as previsões do obtidas no Sistema de Expectativas de Mercado do Bacen em 27/12/2021 (com exceção de Dias Úteis). A maneira como os dados foram obtidos está descrita na seção [3.2.1](#).

### 4.3.1 Modelos SARIMAX

Para a previsão de 12 meses foi utilizada a série de treinamento com as *dummies* de intervenção e com os conjuntos de testes com as séries exógenas previstas e as reais.

## Auto-arima

Utilizando-se a ferramenta Auto-arima para determinar o melhor modelo SARIMAX baseando-se em critérios de informação, obteve-se os resultados reportados a seguir.

- Critério de Akaike AIC: SARIMAX(0,1,1),(2,0,1)
- Critério de Schwarz BIC: SARIMAX(0,1,1)(1,0,1)

Foram os mesmos modelos com as mesmas defasagens sugeridas no experimento univariado.

### Modelo SARIMAX(0,1,1),(2,0,1) sugerido pelo critério de Akaike

O teste de Ljung-Box tem estatística 0.85, mostrando há correlação dos resíduos. E são significativos a 5% os coeficientes do lag 1 da média móvel e o coeficiente sazonal do lag 24 do AR. O sumário dos resultados está na figura 4.25.

No sumário as variáveis exógenas são referenciadas por  $x1, \dots, x8$ . Esta referência é dada pela ordem em que as variáveis aparecem na matriz de exógenas. Assim:

- x1 é IBC-Br;
- x2 é IGP-M;
- x3 é Dólar;
- x4 é Dias Úteis;
- x5 é a *dummy* LS2008.Nov
- x6 é a *dummy* LS2020.Abr
- x7 é a *dummy* AO2020.Mai
- x8 é a *dummy* LS2020.Ago

As 4 dummies LS2008.Nov, LS2020.Apr, AO2020.Mai e LS2020.Ago são estatisticamente significativas. Já a dummy LS2022.Out não foi considerada pois ela vale zero em todas as observações do conjunto de dados de teste.

As variáveis exógenas do Dólar e do IBC-BR não foram significativas. As demais, IGP-M e Dias Úteis foram.

SARIMAX Results						
Dep. Variable:	y	No. Observations:	228			
Model:	ARIMA(0, 1, 1)x(2, 0, 1, 12)	Log Likelihood	452.855			
Date:	Mon, 21 Aug 2023	AIC	-877.710			
Time:	11:12:44	BIC	-829.761			
Sample:	01-01-2003 - 12-01-2021	HQIC	-858.362			
Covariance Type:	opg					
coef	std err	z	P> z	[0.025	0.975]	
x1	0.0007	0.001	0.471	0.638	-0.002	0.003
x2	0.0006	0.000	2.960	0.003	0.000	0.001
x3	0.0307	0.016	1.951	0.051	-0.000	0.062
x4	0.0086	0.002	3.900	0.000	0.004	0.013
x5	-0.1277	0.040	-3.200	0.001	-0.206	-0.049
x6	-0.1733	0.041	-4.271	0.000	-0.253	-0.094
x7	-0.1352	0.052	-2.597	0.009	-0.237	-0.033
x8	0.1186	0.046	2.592	0.010	0.029	0.208
const	0	8.25e-08	0	1.000	-1.62e-07	1.62e-07
ma.L1	-0.4539	0.055	-8.254	0.000	-0.562	-0.346
ar.S.L12	0.3433	0.216	1.586	0.113	-0.081	0.767
ar.S.L24	0.3665	0.145	2.529	0.011	0.082	0.651
ma.S.L12	0.0777	0.245	0.317	0.751	-0.402	0.557
sigma2	0.0010	8.9e-05	11.735	0.000	0.001	0.001
<hr/>						
Ljung-Box (L1) (Q):	0.85	Jarque-Bera (JB):	12.32			
Prob(Q):	0.36	Prob(JB):	0.00			
Heteroskedasticity (H):	0.80	Skew:	-0.21			
Prob(H) (two-sided):	0.34	Kurtosis:	4.06			
<hr/>						

Figura 4.25: Sumário do Modelo SARIMAX(0,1,1)(2,0,1)

A análise dos resíduos mostra autocorrelação significativa no lag 7, que os resíduos são estacionários e tem uma distribuição próxima à normal, conforme a figura 4.26.

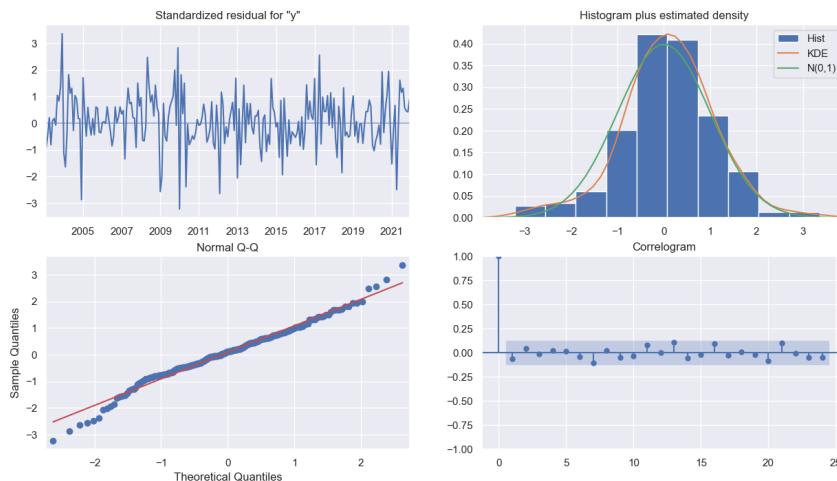


Figura 4.26: Análise dos resíduos do Modelo SARIMAX(0,1,1)(2,0,1)

### Modelo SARIMAX(0,1,1),(1,0,1) sugerido pelo critério de Schwartz

O teste de Ljung-Box tem estatística 1.26, mostrando que há correlação dos resíduos. E são significativos a 5% os coeficientes do lag 1 da média móvel e o coeficiente sazonal do lag 12 do AR. O sumário dos resultados está na figura 4.27.

As 4 dummies LS2008.Nov, LS2020.Apr, AO2020.Mai e LS2020.Ago são estatisticamente significativas. Já a dummy LS2022.Out não foi considerada pois ela vale 0 em todas as observações do conjunto de dados de teste.

A variável exógena do IBC-Br não foi significativa. As demais, Dólar, IGP-M e Dias Úteis foram.

SARIMAX Results						
Dep. Variable:	y	No. Observations:	228			
Model:	ARIMA(0, 1, 1)x(1, 0, 1, 12)	Log Likelihood	450.326			
Date:	Mon, 21 Aug 2023	AIC	-874.653			
Time:	11:12:49	BIC	-830.128			
Sample:	01-01-2003 - 12-01-2021	HQIC	-856.686			
Covariance Type:	opg					
	coef	std err	z	P> z	[0.025	0.975]
x1	6.355e-05	0.001	0.051	0.959	-0.002	0.003
x2	0.0006	0.000	2.846	0.004	0.000	0.001
x3	0.0365	0.015	2.508	0.012	0.008	0.065
x4	0.0089	0.002	4.021	0.000	0.005	0.013
x5	-0.1445	0.032	-4.489	0.000	-0.208	-0.081
x6	-0.1860	0.036	-5.230	0.000	-0.256	-0.116
x7	-0.1375	0.044	-3.145	0.002	-0.223	-0.052
x8	0.1170	0.042	2.770	0.006	0.034	0.200
const	0	8.1e-08	0	1.000	-1.59e-07	1.59e-07
ma.L1	-0.4560	0.054	-8.427	0.000	-0.562	-0.350
ar.S.L12	0.7573	0.078	9.651	0.000	0.603	0.911
ma.S.L12	-0.2263	0.129	-1.749	0.080	-0.480	0.027
sigma2	0.0011	9.28e-05	11.552	0.000	0.001	0.001
Ljung-Box (L1) (Q):	1.26	Jarque-Bera (JB):	6.28			
Prob(Q):	0.26	Prob(JB):	0.04			
Heteroskedasticity (H):	0.86	Skew:	-0.09			
Prob(H) (two-sided):	0.52	Kurtosis:	3.79			

Figura 4.27: Sumário do Modelo SARIMAX(0,1,1)(1,0,1)

A análise dos resíduos mostra que não há autocorrelações significativas, que os resíduos são estacionários e tem uma distribuição próxima à normal, conforme a figura 4.28.

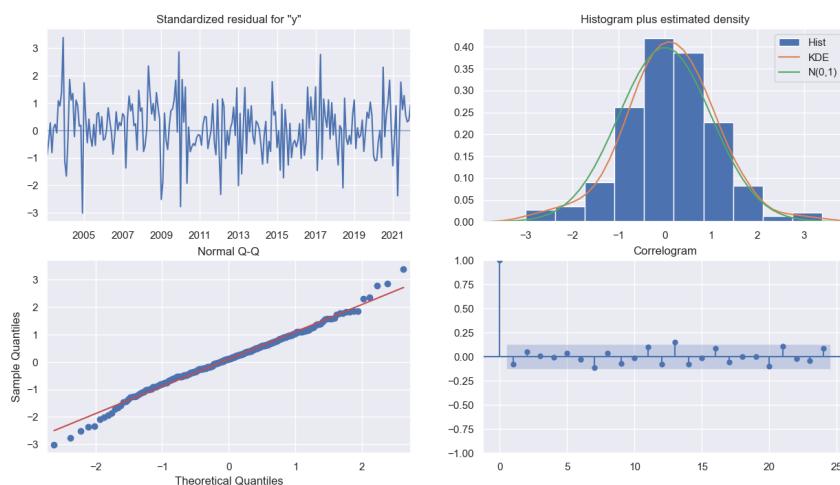


Figura 4.28: Análise dos resíduos do Modelo SARIMAX(0,1,1)(1,0,1)

### 4.3.2 Modelos de Mudança Markoviana

No experimento univariado com regressão múltipla os três modelos de mudança markoviana utilizaram as mesmas variáveis, a série de treinamento do ICMS e IBC-Br e IGP-M como mudança de regime. Dólar foi deixado de lado pois piorava as previsões. Dias Úteis foi usada como variável exógena. Foram adotadas duas mudanças de regime. Não foram utilizadas *dummies* de intervenção, apenas *dummies* sazonais centradas. Foram usados três tipos de variância, variância fixa (*fixed variance*), mudança de variância (*switching variance*) e componente média-variância (*mean-variance component*).

Os resultados são apresentados a seguir. O sumário de cada um dos modelos está no Apêndice B.

#### Modelo de Mudança Markoviana com Variância Fixa

Este modelo mostrou elevada alternância de regimes, ficando 158 meses no regime 0 e 68 meses no regime 1, com duração média de 2.39 meses no regime 0 e 1.03 meses no regime 1. A variável dummy sazonal centrada 3 não foi significativa a 5%. Todas as demais foram estatisticamente significativas. IGP-M teve coeficiente 0.0006 no regime 0 e -0.0015 no regime 1. IBC-Br teve coeficiente 0.0028 no regime 0 e 0.0070 no regime 1. A variância do modelo foi 0.03.

A análise dos resíduos mostra normalidade mas há autocorrelação significativa, conforme pode ser visto na figura 4.29 e no resultado dos testes no apêndice B.

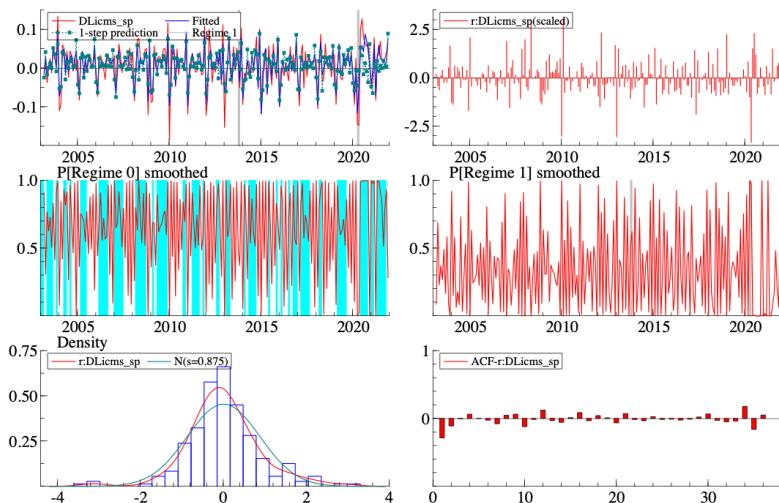


Figura 4.29: Resultados do modelo de Markov com Variância Fixa

## Modelo de Mudança Markoviana com Mudança de Variância

Este modelo mostrou alternância de regimes, ficando 171 meses no regime 0 e 55 meses no regime 1, com duração média de 9.50 meses no regime 0 e 3.24 meses no regime 1. As variáveis dummies sazonais centradas 3 e 8 não foram significativas a 5%, assim como IGP-M no regime 1 e IBC-BR no regime 0. Todas as demais foram estatisticamente significativas. IGP-M teve coeficiente  $-0.0008$  no regime 0 e  $0.0005$  no regime 1. IBC-Br teve coeficiente  $0.0006$  no regime 0 e  $0.0077$  no regime 1. A variância do modelo foi 0.02 no regime 0 e 0.06 no regime 1.

A análise dos resíduos não mostra normalidade e há autocorrelação significativa, conforme pode ser visto na figura 4.30 e nos testes no apêndice B.

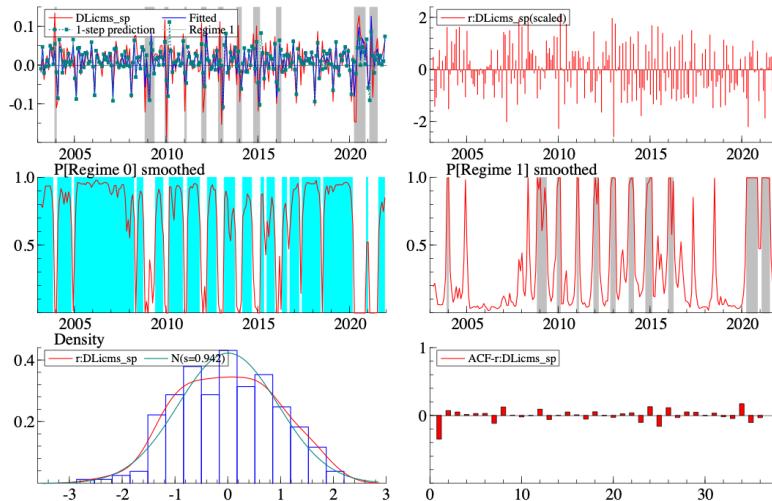


Figura 4.30: Resultados do modelo de Markov com Mudança de Variância

## Modelo de Mudança Markoviana com Componente Média-Variância

Este modelo mostrou considerável alternância de regimes de variância, mas houve predominância do regime 0 de média. No regime de média o modelo ficou 189 meses no regime 0 e 37 meses no regime 1, com duração média de 5.40 meses no regime 0 e 1.09 meses no regime 1. Já no regime de variância o modelo ficou 105 meses no regime 0 e 121 meses no regime 1, com duração média de 5.00 meses no regime 0 e 5.76 meses no regime 1.

A variável dummy sazonal centrada 3 não foi significativa a 5%, assim como o IBC-BR no regime 0. Todas as demais foram estatisticamente significativas. IGP-M teve coeficiente 0.0007 no regime 0 e  $-0.0015$  no regime 1. IBC-Br teve coeficiente 0.0009 no

regime 0 e 0.009 no regime 1. A variância do modelo foi 0.009 no regime 0 e 0.045 no regime 1.

A análise dos resíduos não mostra normalidade e há autocorrelação significativa, conforme pode ser visto na figura 4.31 e nos testes no apêndice B.

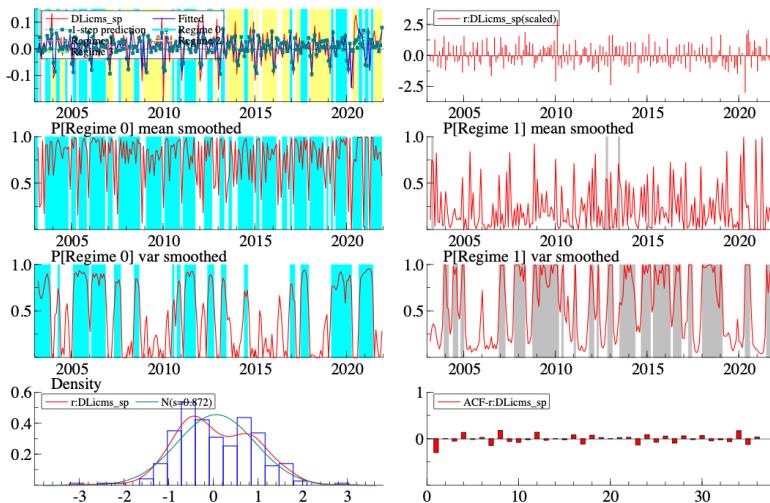


Figura 4.31: Resultados do modelo de Markov com Mudança de Variância

### 4.3.3 Modelo RNN LSTM

No experimento univariado com regressão múltipla o conjunto de dados das características foram apenas as variáveis IBC-Br, IGP-M, Dólar e Dias Úteis. Elas foram adicionadas ao conjunto de dados com lags 0 (contemporâneas), com lag1 e lag 2, perfazendo 12 colunas de características. Com isso, as 2 primeiras amostras do conjunto de treinamento foram perdidas e este se iniciou em Março de 2003.

Foram feitos dois experimentos, um com um conjunto de testes usando os valores previstos para 2022 do IBC-Br, IGP-M e Dólar e outro com os dados reais do ano de 2022.

Foram testadas, em validação, diversas arquiteturas de rede com diversos hiperparâmetros. A arquitetura que apresentou melhores resultados foi a com três camadas escondidas, com MAPE de 6.55%, contra um MAPE de 6.99% da arquitetura com uma só camada escondida.

A arquitetura e os hiperparâmetros utilizados foram:

- 3 camadas escondidas com 600 neurônios LSTM;

- dropout de 15%;
- função de perda MAPE;
- otimizador Adam com taxa de aprendizagem começando em  $10^{-5}$ ;
- camada de saída com 1 neurônio com função de ativação linear.
- 1000 épocas com *early-stop* com 100 épocas de paciência.
- redução da taxa de aprendizagem em 10% em caso de atingimento de platô, com paciência de 25 épocas.

### Modelo RNN-LSTM

O modelo rodou 108 épocas até parar. O resultado da perda em treino e validação está na figura 4.32. A curva não aparenta nem sub ajuste (*underfit*), e nem sobre ajuste (*overfit*).

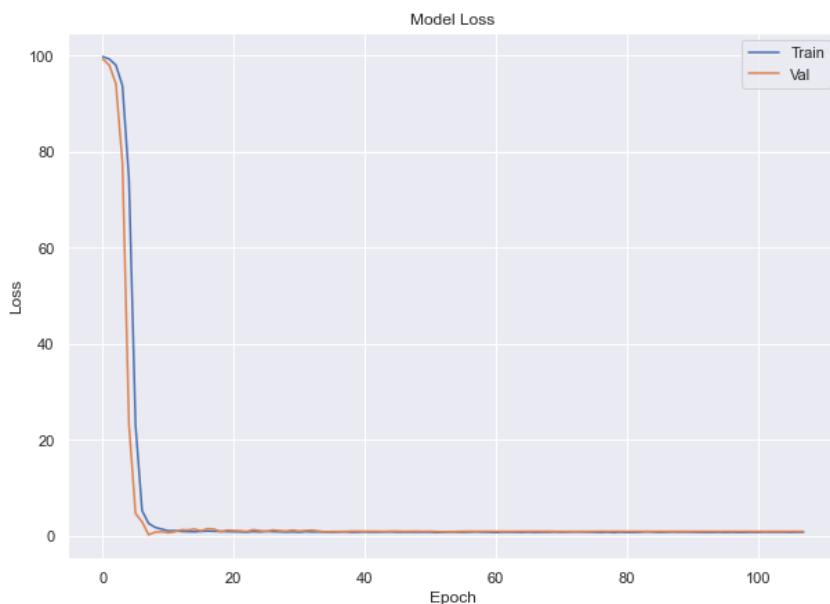


Figura 4.32: Gráfico da perda em treino e validação do modelo RNN-LSTM

#### 4.3.4 Previsões

As previsões foram realizadas sobre o conjunto de testes (*out-of-sample*). As métricas RMSE e MAPE estão na tabela 4.7.

	RMSE	MAPE
SARIMAX(0,1,1)(2,0,1)	1.406283e+09	0.062002
SARIMAX(0,1,1),(1,0,1)	1.451333e+09	0.065229
Markov-Fixed	1.555024e+09	0.076957
Markov-Switching-Var	1.727172e+09	0.092085
Markov-Mean-Var	1.499231e+09	0.074070
RNN-LSTM	6.599282e+08	0.035033

Tabela 4.7: Métricas dos modelos de previsão univariados com regressão múltipla

## RMSE

Analizando-se o RMSE na tabela e na figura 4.33. Pode-se notar que:

- O menor RMSE foi do modelo RNN-LSTM, seguido do modelo SARIMAX(0,1,1)(2,0,1).
- O terceiro menor foi o do modelo SARIMAX(0,1,1),(1,0,1).

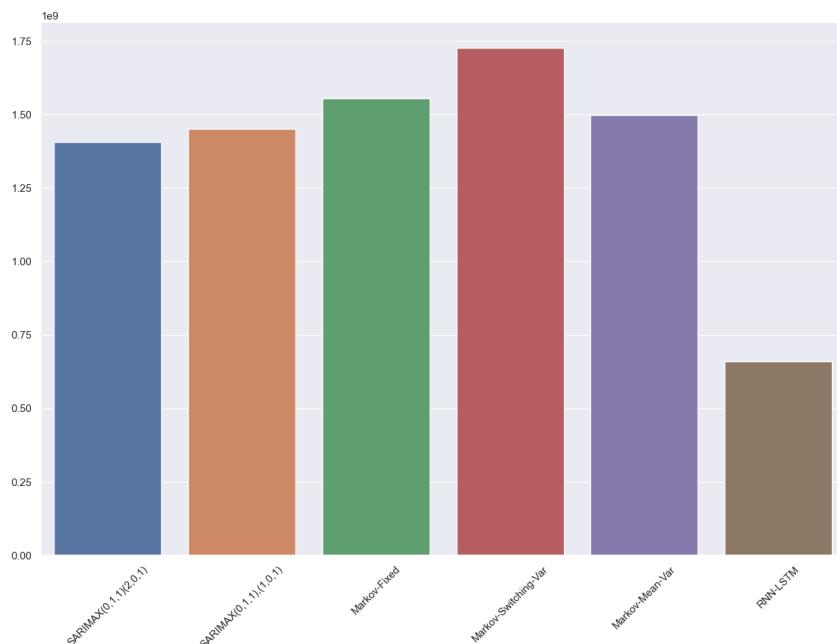


Figura 4.33: Métricas RMSE dos modelos de previsão

## MAPE

Analizando-se o MAPE na tabela e na figura 4.34. Pode-se notar que:

- O menor MAPE foi do modelo RNN-LSTM, seguido do modelo SARIMAX(0,1,1)(2,0,1).
- O terceiro menor foi o modelo SARIMAX(0,1,1),(1,0,1).

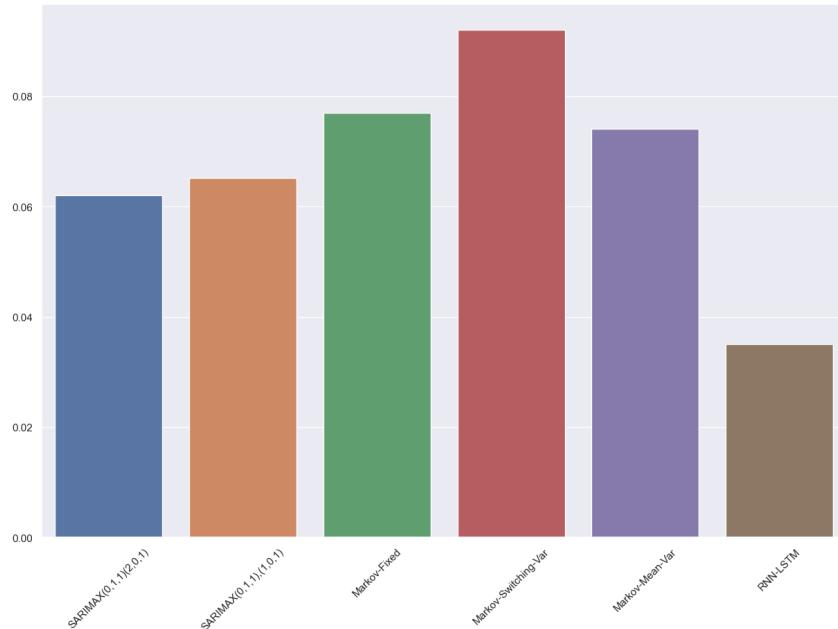


Figura 4.34: Métricas MAPE dos modelos de previsão

### Comparação das Previsões

Neste experimento todos os modelos apresentaram maior variabilidade mês-a-mês. O de melhor desempenho foi o modelo RNN-LSTM. E o de pior desempenho foi o de Mudança Markoviana com Mudança de Variância.

### Análise do Desempenho das Variáveis Exógenas Previstas e Reais

Conforme descrito na seção 3.2.1 as previsões univariadas com regressão múltipla foram realizadas com o uso de previsões das variáveis exógenas obtidas no Sistema de Expectativas de Mercado do Bacen. Com o objetivo de comparar o desempenho das previsões de arrecadação vis a vis as previsões do Bacen, estas foram feitas utilizando-se também os valores realizados das exógenas.

Como é possível observar na figura 4.36, os valores de RMSE e MAPE dos modelos SARIMAX são bastante semelhantes, mostrando que o uso das previsões do IBC-Br, IGP-M e Dólar do Bacen não comprometeu a qualidade das previsões.

Comportamento semelhante ocorre com o modelo RNN-LSTM. Os valores de RMSE e MAPE dos modelos são bastante semelhantes, como é possível observar na figura 4.37.

Por outro lado, nos modelos de mudança markoviana houve diferença notável nos resultados obtidos com as variáveis previstas e reais para os modelos com variação na variância e variação na média e na variância. Já o modelo com variância fixa ficou bas-



Figura 4.35: Previsões dos Modelos univariados com regressão múltipla com Variáveis Previstas comparadas com o Realizado

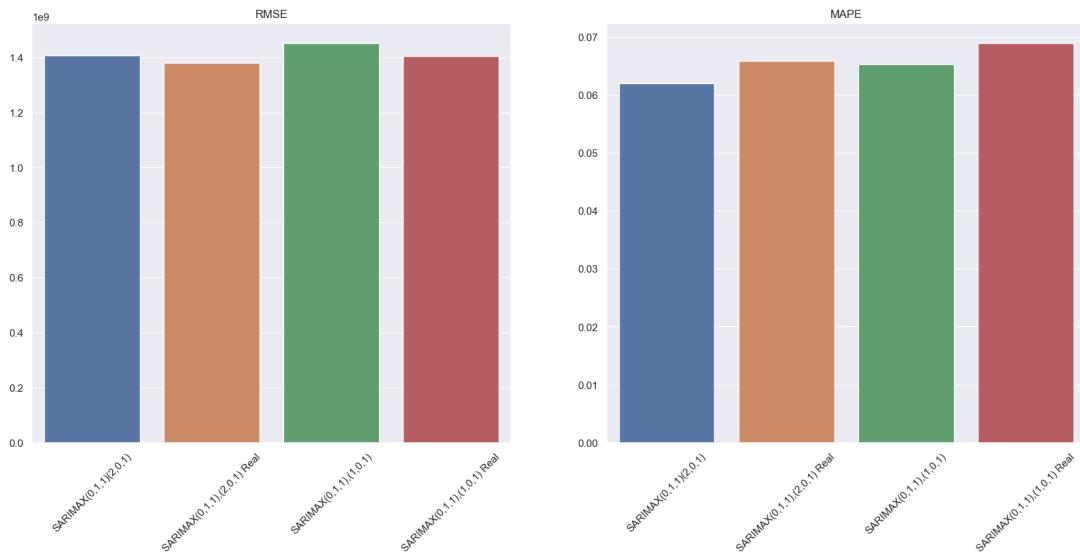


Figura 4.36: Métricas RMSE e MAPE dos modelos SARIMAX com o uso de variáveis exógenas com valores previstos pelo mercado e valores reais.

tante semelhante tanto para as variáveis exógenas previstas quanto para as realizadas, conforme a figura 4.38.

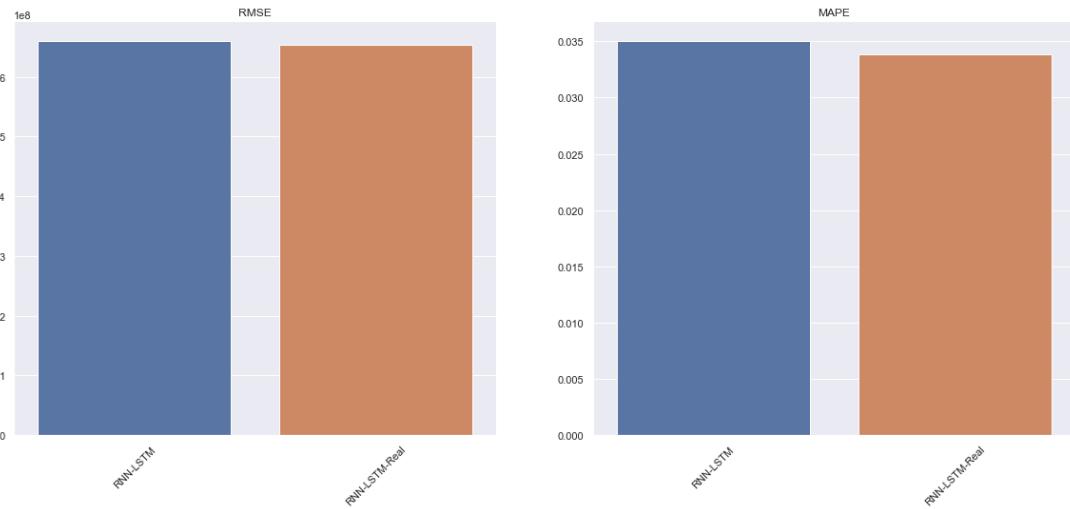


Figura 4.37: Métricas RMSE e MAPE do modelo RNN-LSTM com o uso de variáveis exógenas com valores previstos pelo mercado e valores reais.

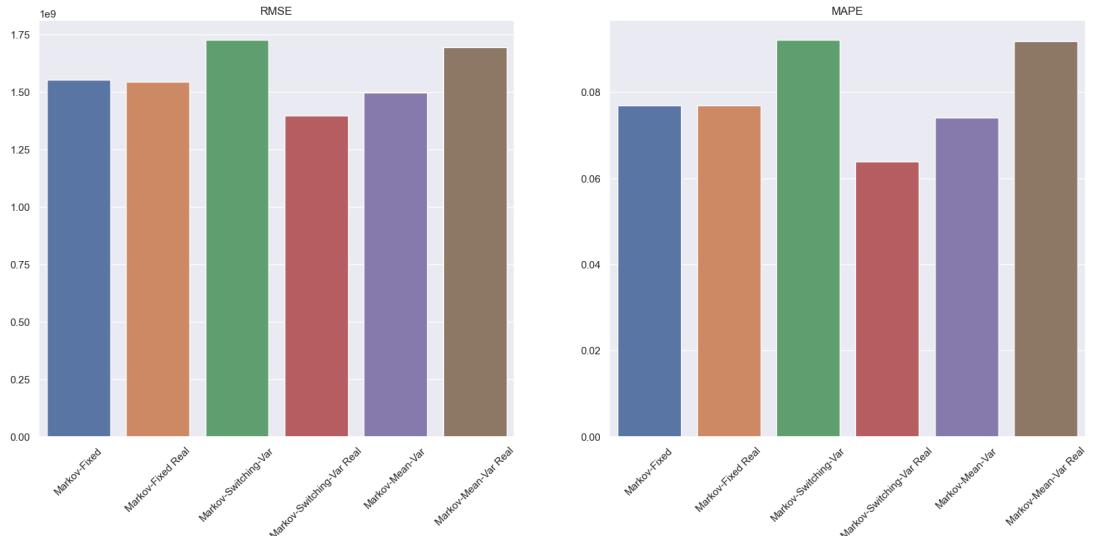


Figura 4.38: Métricas RMSE e MAPE dos modelos de mudança markoviana com o uso de variáveis exógenas com valores previstos pelo mercado e valores reais.

### Seleção dos Modelos Univariados com Regressão Múltipla de 2022

De cada uma das estratégias foi escolhido um modelo. O critério de escolha para SARIMA, SARIMAX e Markov-Switching foi o menor Critério de Informação de Schwarz (BIC). As RNN-LSTM foi o menor MAPE em validação. Os resultados estão na tabela 4.8.

De modo análogo ao experimento univariado, o menor BIC não proporcionou sempre as melhores escolhas. O modelo SARIMAX escolhidos não foi os com menor MAPE, mas no caso do modelo de Markov-Switching, foi. Novamente, é importante deixar claro que não seria possível saber isto de antemão caso a previsão tivesse sido feita no ano de 2021. Como foi feito apenas um modelo de rede neural LSTM, não houve escolha a ser feita.

Modelo	Estratégia	BIC	MAPE Val	Escolha
SARIMAX(0,1,1)(2,0,1)	sarimax	-829.761000		Não
SARIMAX(0,1,1),(1,0,1)	sarimax	-830.128000		<b>Sim</b>
Markov-Fixed	markov-switching	-3.30064232		Não
Markov-Switching-Var	markov-switching	-3.33453345		Não
Markov-Mean-Var	markov-switching	-3.35649623		<b>Sim</b>
RNN-LSTM	rnn-lstm	0.06551987		<b>Sim</b>

Tabela 4.8: Escolha dos modelos univariados com regressão múltipla de 2022 e seus critérios.

#### 4.3.5 Comparação dos Experimentos de 2022

Comparando-se os modelos SARIMA e SARIMAX em ambos experimentos, estes obtiveram performance melhor no experimento univariado com regressão múltipla, com o melhor modelo SARIMAX univariado com regressão múltipla apresentando um MAPE de 6.20% contra um MAPE de 8.52% do melhor SARIMA/SARIMAX univariado.

Os modelos de Mudança Markoviana obtiveram desempenho melhor que os modelos SARIMA/SARIMAX no experimento univariado com regressão múltipla, mas tiveram desempenho pior no experimento univariado, com o modelos de variância fixa se saindo melhor no experimento univariado e o modelo com mudança de média e variância se saindo melhor no experimento univariado com regressão múltipla. O univariado teve MAPE de 6.61%, e o univariado com regressão múltipla teve MAPE de 6.38%.

As Redes Neurais Recorrentes LSTM apresentaram a melhor performance em ambos experimentos, com MAPE de 3.76% no experimento univariado e MAPE de 3.50% no univariado com regressão múltipla.

O melhor modelo, em todos os experimentos, foi a RNN-LSTM 12 Meses com Variáveis Exógenas.

### 4.4 Comparação Geral dos Experimentos

Os experimentos univariados e univariados com regressão múltipla foram repetidos para os anos de 2013 a 2021, além do ano de 2022. Os resultados de cada experimento podem

ser vistos nos dados do projeto disponíveis no Github<sup>3</sup>.

Analizando-se os resultados por experimento por ano, os menores MAPE se encontram, em sua maioria, nos experimentos univariados, como se pode notar no gráfico da figura 4.39.

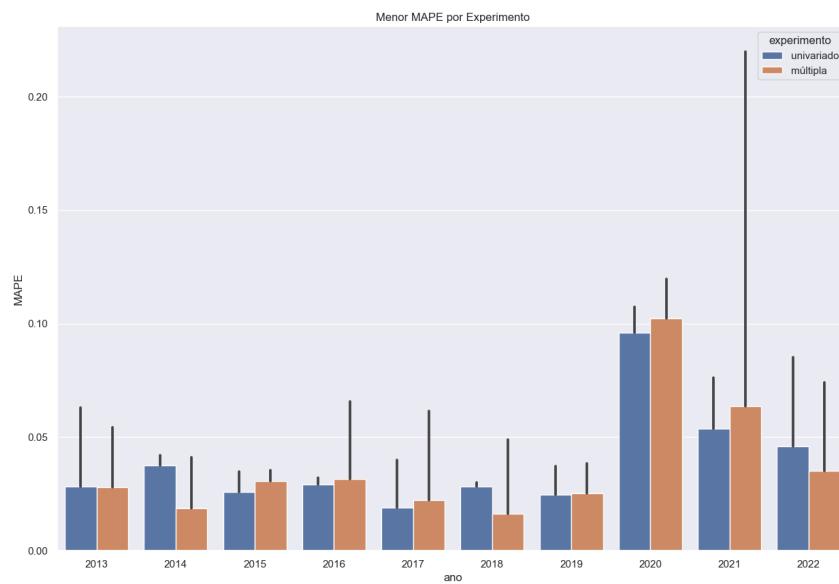


Figura 4.39: Comparação dos menores MAPE por experimento por ano independente de estratégia.

Mostrando apenas o melhor do ano, pode-se ter certeza de que a maioria dos menores MAPE foram obtidos nos experimentos univariados. Isso pode ser visto na figura 4.40.

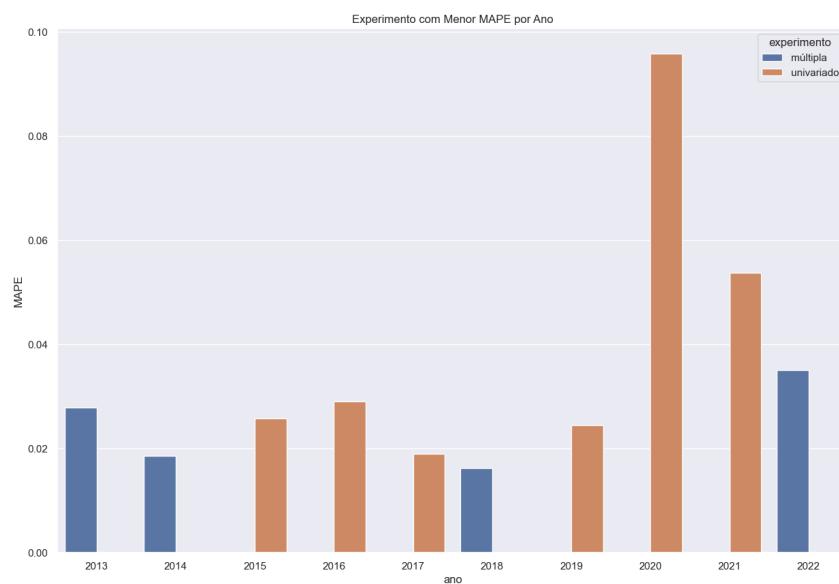


Figura 4.40: Apenas o experimento com menor MAPE por ano independente de estratégia.

<sup>3</sup><[https://github.com/asbonato/mpe\\_fgv\\_2023.git](https://github.com/asbonato/mpe_fgv_2023.git)>

Com relação às estratégias, analisando-se as de menor MAPE por ano, independente de experimento - gráfico da figura 4.41, percebe-se que os modelos de mudança markoviana são os melhores, se saindo melhor 50% das vezes. Os modelos SARIMAX tem o melhor desempenho 30% das vezes. E a rede neural só tem o melhor desempenho no último ano. Vale destacar que em 2016, ano da crise econômica do segundo mandato da presidente Dilma Rousseff, a melhor estratégia foi a Naïve. A tendência da arrecadação é subir todo ano, como se vê na figura 3.4. Mas, devido à queda da atividade econômica, isto não ocorreu em 2016, ficando a média mensal de arrecadação do ano, R\$ 10.237 bilhões, menor do que a média de 2015, R\$ 10.099 bilhões e esta queda não tinha como ser identificada pelos modelos treinados com dados até o ano anterior.

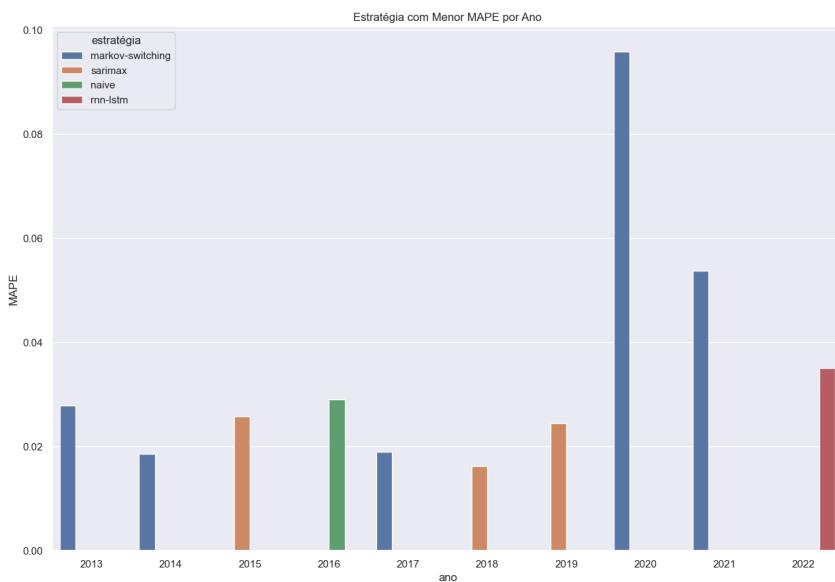


Figura 4.41: Menor MAPE por estratégia por ano independente de experimento.

#### 4.4.1 Todas as Estratégias e Experimentos

Independente da estratégia ou do experimento, é possível notar um baixo desempenho dos modelos nos anos de 2020 e 2021, anos em que a crise da Covid-19 e a decorrente crise inflacionária mudaram radicalmente os cenários de previsão.

#### Experimento Univariado

Agrupados por ano, o desempenho de cada estratégia no experimento univariado pode ser observada na figura 4.42. Os modelos SARIMAX tem melhor desempenho no início, de 2013 a 2015 e depois voltam a ter bom desempenho em 2019. Os modelos SARIMA

não tiveram o melhor desempenho em nenhum dos experimentos univariados, mostrando a importância do uso das *dummies* de intervenção. Os modelos de mudança markoviana só tem bom desempenho em 2017 e 2018, e depois em 2020 e 2021. Nos demais anos, os modelos Markov-Switching tem os piores desempenhos. O modelo RNN-LSTM só tem performance superior em 2022.

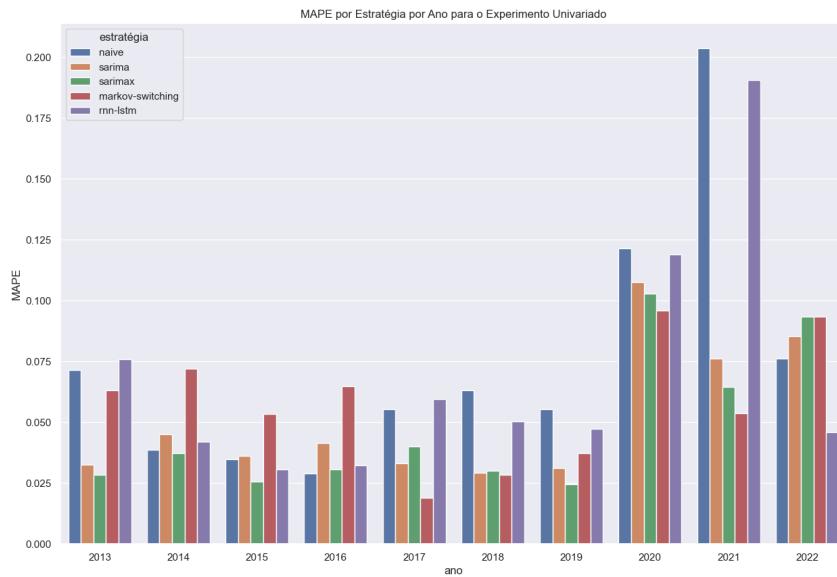


Figura 4.42: MAPE por estratégia por ano no experimento univariado.

O melhor modelo do experimento univariado de cada ano está na figura 4.43.

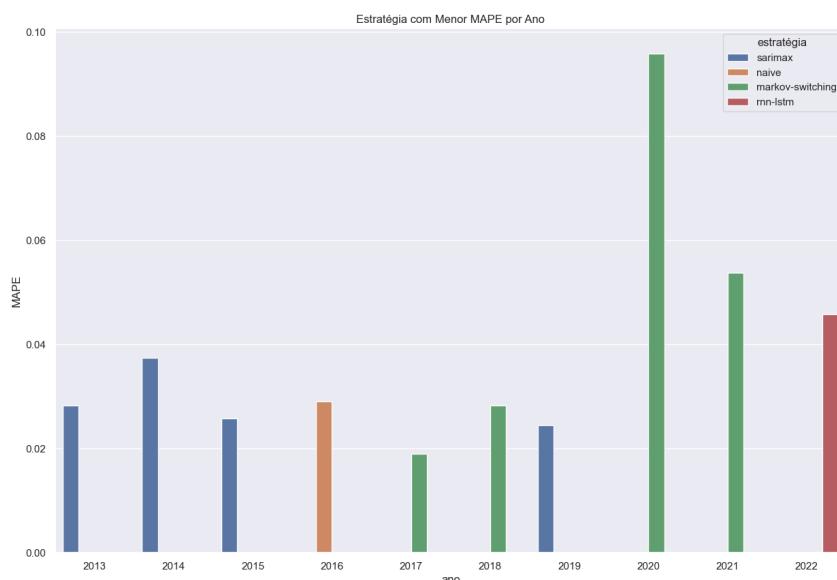


Figura 4.43: Menor MAPE por estratégia por ano no experimento univariado.

## Experimento Univariado com Regressão Múltipla

Focando agora no experimento univariado com regressão múltipla, o desempenho de cada estratégia, agrupado por ano, pode ser observada na figura 4.44. Os modelos de mudança markoviana tem bom desempenho no geral, sendo os melhores em 2014, 2015, 2017, 2020 e 2021. O modelo SARIMAX tem o melhor desempenho em 2016. em 2018 e 2019. O modelo RNN-LSTM, tem performance superior em 2015 e 2022.

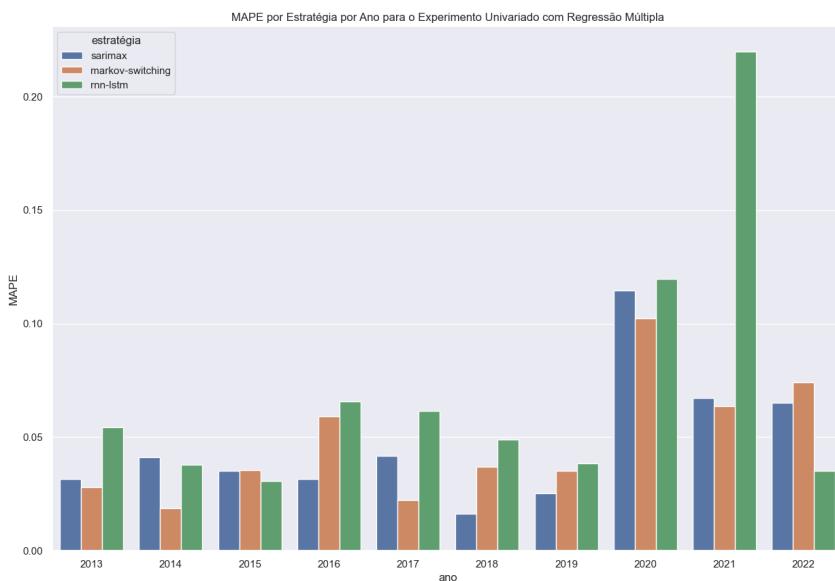


Figura 4.44: MAPE por estratégia por ano no experimento univariado com regressão múltipla.

O melhor modelo do experimento univariado com regressão múltipla de cada ano está na figura 4.45.

### 4.4.2 Desempenho Anual dos Experimentos

Como já apontado, todos os modelos, em ambos experimentos, comprehensivelmente perderam muito do seu poder preditivo nos anos de 2020 a 2022.

#### Experimento Univariado

Os modelos SARIMA e SARIMAX obtiveram bom desempenho de 2013 a 2019. Os modelos não-lineares já têm performance irregular, desempenhando mal nos anos iniciais e finais do experimento, mas comportando-se bem de 2017 a 2019. As redes neurais foram bem em 2015 e 2016, e tiveram o melhor desempenho de 2022, ano em que todas as outras estratégias foram ruins.

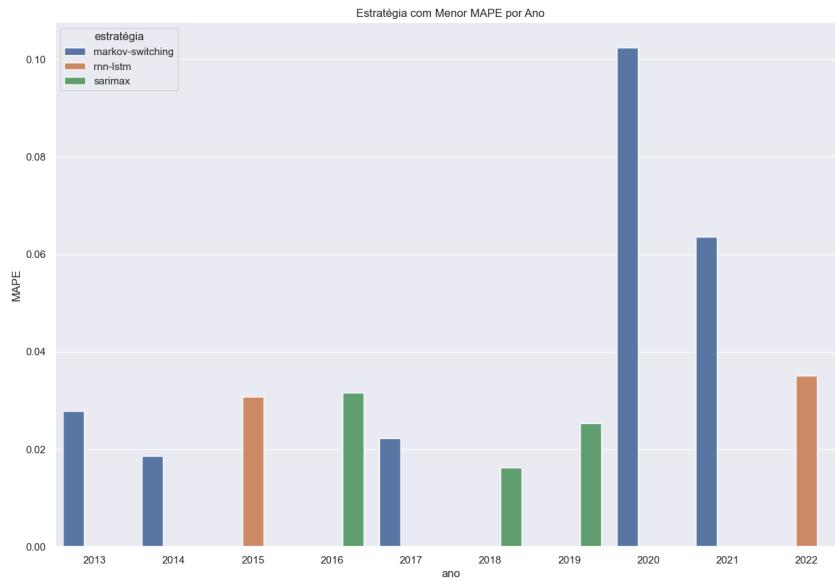


Figura 4.45: Menor MAPE por estratégia por ano no experimento univariado com regressão múltipla.

O desempenho dos modelos univariados pode ser visto na figura 4.46.

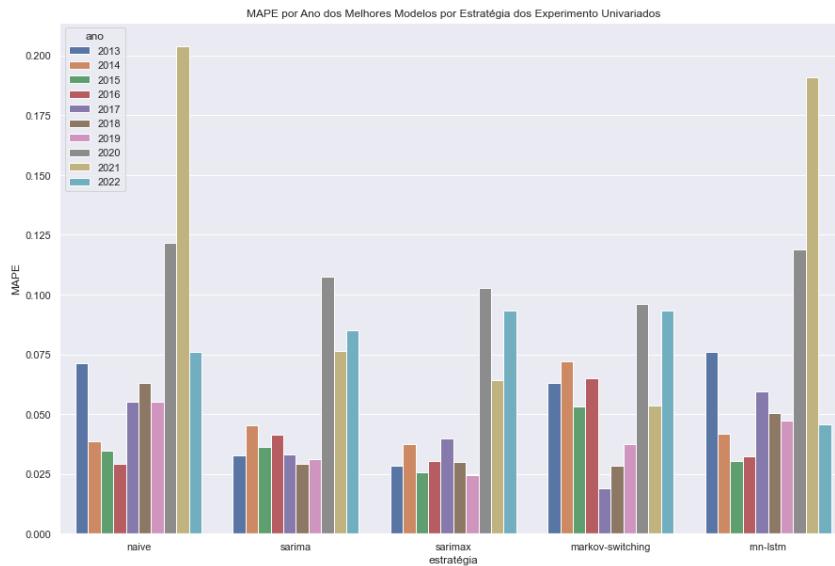


Figura 4.46: MAPE por ano por estratégia no experimento univariado.

## Experimento Univariado com Regressão Múltipla

Ao contrário do experimento univariado, nem todos os modelos obtiveram bom desempenho de 2014 a 2016. O modelo SARIMAX, novamente, se desempenhou melhor de 2013 a 2019. O modelo de mudança markoviana também se comportou bem neste período, exceto em 2016, quando teve um dos piores desempenhos. A RNN-LSTM foi o pior modelo em todos os anos, exceto o último, no qual foi o melhor.

Em 2020 e 2021 todos os modelos foram ruins, mas o Markov-Switching foi o melhor entre eles. Mas há uma regularidade de desempenho do modelo SARIMAX que precisa ser destacada. O desempenho dos modelos univariados com regressão múltipla pode ser visto na figura 4.47.

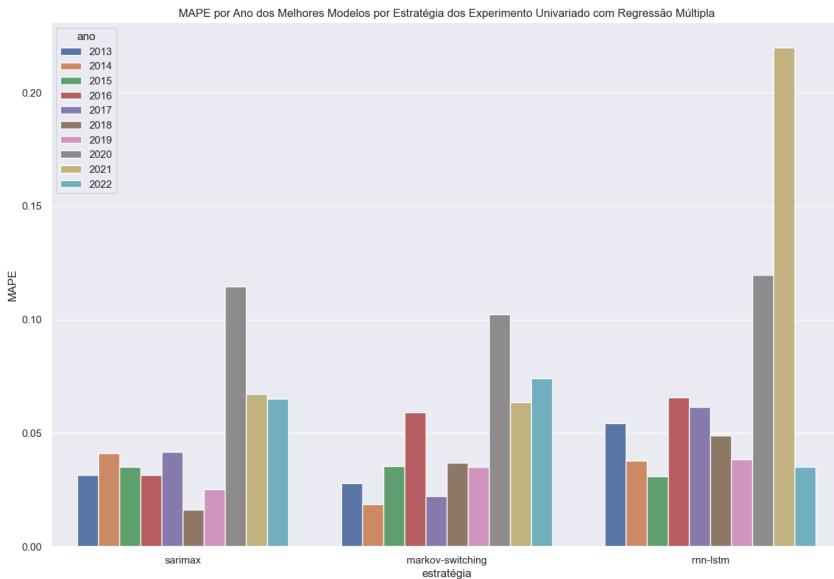


Figura 4.47: MAPE por ano por estratégia no experimento univariado com regressão múltipla.

## Comparação das Previsões

É possível comparar os valores previstos por cada um dos modelos escolhidos no período de 2013 a 2022 com os valores reais da arrecadação.

**Estratégia Naïve** A estratégia Naïve, que é a média mensal do ano anterior, sempre ficou abaixo dos valores reais arrecadados, mostrando que a arrecadação de um ano sempre é maior que a do ano anterior. Exceto no ano de 2016, como resultado da crise econômica do segundo governo da presidente Dilma Rousseff, onde a arrecadação foi menor do que a do ano anterior. Aliás, este ponto já havia sido apontado na comparação dos MAPEs, pois em 2016 o melhor MAPE foi o da estratégia Naïve. A comparação é mostrada na figura 4.48.

**Estratégias SARIMA/SARIMAX** Os modelos da estratégia SARIMA do experimento univariado e da estratégia SARIMAX dos modelos univariado e univariado com

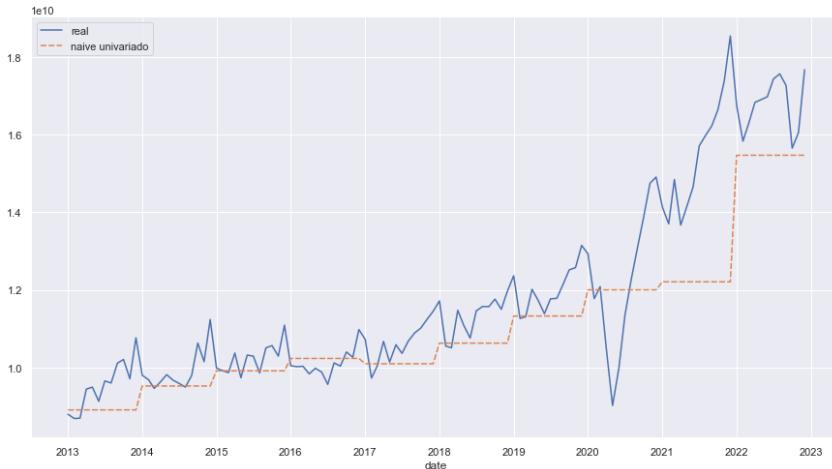


Figura 4.48: Valores de arrecadação, em R\$, previstos pelo modelo da estratégia Naïve comparados com os valores reais arrecadados.

regressão múltipla apresentaram curvas bastante semelhantes, que acompanharam de maneira similar a curva da arrecadação real, com exceção de dois pontos importantes: a queda decorrente da pandemia de COVID-19 em 2020 e a queda de arrecadação devido às mudanças de alíquota de combustíveis e outros preços administrados em 2022. Nota-se que a subida inflacionária de 2021 foi bem acompanhada pelos modelos, embora o realizado tenha sido maior do que previsto, e esta mesma escalada foi reproduzida nas previsões de 2022. As curvas podem ser vistas na figura 4.49.



Figura 4.49: Valores de arrecadação, em R\$, previstos pelo modelos das estratégias SA-RIMA/SARIMAX comparados com os valores reais arrecadados.

**Estratégias Markov-Switching** De modo análogo aos modelos das estratégias SA-RIMA/SARIMAX, os modelos da estratégia de mudança markoviana univariado e univa-

riado com regressão múltipla apresentaram curvas que acompanharam de maneira similar a curva da arrecadação real, com exceção dos mesmos dois pontos, a queda da COVID-19 em 2020 e a queda da mudanças de alíquota em 2022. O modelo univariado seguiu melhor a subida inflacionária de 2021, mas ficou muito mais alto que a arrecadação real de 2013 a 2016. Já o univariado com regressão múltipla ficou mais próximo da curva real neste período, mas ficou muito abaixo em 2021. Assim como os modelos SARIMA/SARIMAX, os modelos de mudança markoviana também reproduziram a subida de 2021 em 2022, o que não aconteceu na realidade. As curvas podem ser vistas na figura 4.50.

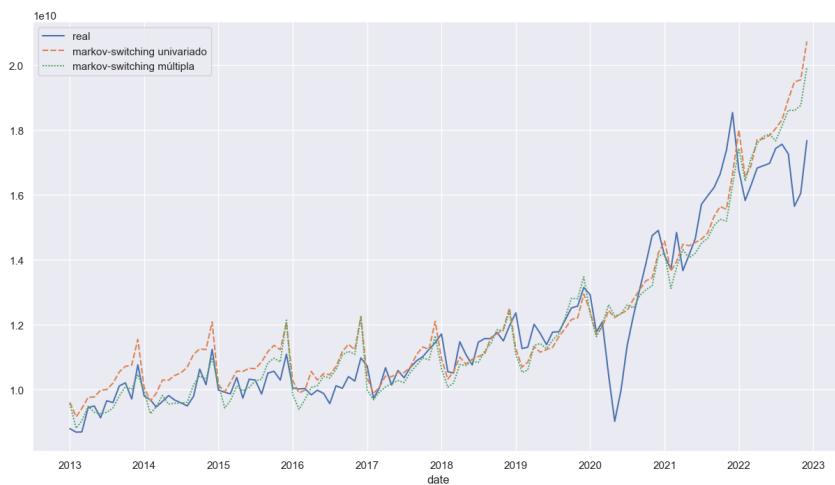


Figura 4.50: Valores de arrecadação, em R\$, previstos pelos modelos da estratégia Markov-Switching comparados com os valores reais arrecadados.

**Estratégias RNN-LSTM** As curvas de previsões produzidas pelos modelos da estratégia de Redes Neurais Recorrentes LSTM ficou muito semelhante à curva produzida pela estratégia Naïve, como pode ser visto na figura 4.51.

A curva da rede univariada com regressão múltipla ainda mostra uma certa variabilidade, mas a univariada ficou tão achatada como a da média.

Em comparação com a curva da arrecadação real, pode-se notar que o ano de 2021 ficou muito abaixo do real, pois a rede não foi capaz de acompanhar a subida da arrecadação causada pela inflação. As curvas estão na figura 4.52.

## MAPE dos Dez Anos

Foi calculado o MAPE de cada uma das séries de previsão produzidas pelos modelos de cada uma das estratégias. Como pode ser visto na tabela 4.9, os melhores MAPEs ficaram



Figura 4.51: Valores de arrecadação, em R\$, previstos pelos modelos da estratégia RNN-LSTM comparados com as previsões do modelo Naïve.



Figura 4.52: Valores de arrecadação, em R\$, previstos pelos modelos da estratégia RNN-LSTM comparados com os valores reais arrecadados.

com os modelos SARIMA/SARIMAX, seguidos dos modelos de mudança markoviana, ficando por último as redes neurais LSTM e os modelos da estratégia Naïve. O gráfico de barras da figura 4.53 mostra este ranking de modelos por MAPE.

## 4.5 Ranking dos Modelos pelo MCS

Os resultados do Model Confidence Set podem ser vistos abaixo. A estratégia de Redes Neurais, de ambos os experimentos, foi descartada, assim como o modelo Naïve, como esperado. Dos modelos ranqueados, os de melhor poder preditivo ao longo dos 10 anos de experimentos foram os da estratégia SARIMAX univariada com regressão múltipla, ficando logo atrás o SARIMAX univariado. Os modelos Markov-Switching ficaram em

Estratégia	MAPE
SARIMAX Reg. Múltipla	0.046986
Markov-Switching Reg. Múltipla	0.047526
SARIMAX Univariado	0.047816
SARIMA Univariado	0.051806
Markov-Switching Univariado	0.058087
RNN-LSTM Univariado	0.070504
RNN-LSTM Reg. Múltipla	0.071202
Naïve Univariado	0.074857

Tabela 4.9: MAPE dos modelos considerando-se as previsões de arrecadação de dez anos produzida por cada um deles.

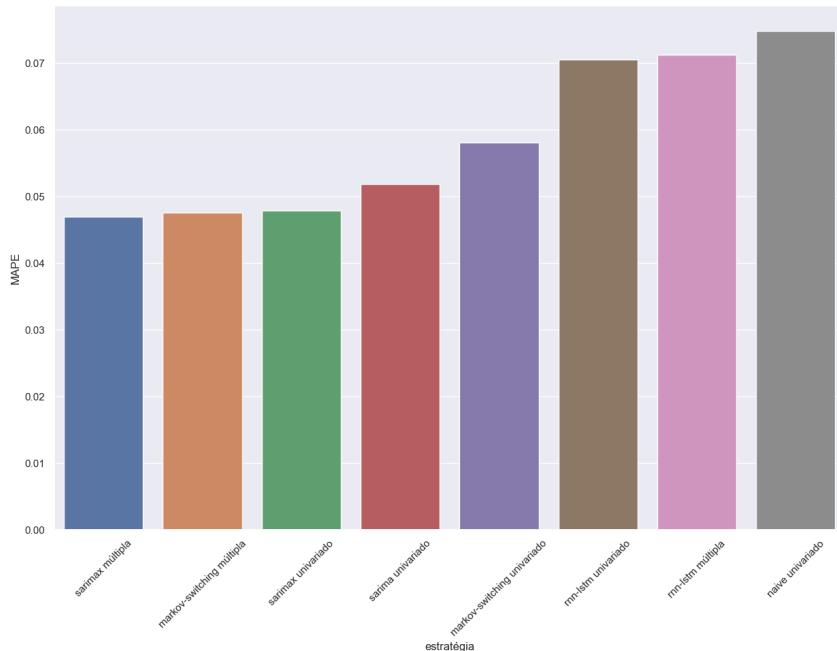


Figura 4.53: MAPE dos modelos considerando-se as previsões de arrecadação de dez anos produzida por cada um deles.

terceiro e quinto lugares, com as Redes Neurais com Variáveis Exógenas ficando em quinto lugar.

O ranking foi:

1. SARIMAX Reg. Múltipla
2. SARIMAX Univariado
3. Markov-Switching Reg. Múltipla
4. SARIMA Univariado

## 5. Markov-Switching Univariado

Os resultados do Model Confidence Set podem ser vistos na figura 4.54.

```
Model naive univariado eliminated 2023-08-21 16:55:09.108596
Model rnn-lstm múltipla eliminated 2023-08-21 16:55:13.583389
Model rnn-lstm univariado eliminated 2023-08-21 16:55:17.589689
#####
Superior Set Model created      :
      Rank_M      v_M    MCS_M Rank_R      v_R    MCS_R     Loss
sarima univariado        4  1.3095335 0.4948      5  3.1235342 0.0000 0.05180554
sarimax univariado       2 -1.4896768 1.0000      3  0.3269794 1.0000 0.04771008
markov-switching univariado 5  1.4147725 0.4182      4  1.6145359 0.4724 0.05808682
sarimax múltipla          1 -1.9975809 1.0000      1 -0.1841693 1.0000 0.04698628
markov-switching múltipla   3 -0.8509486 1.0000      2  0.1841693 1.0000 0.04752559
p-value :
[1] 0.4182
#####

```

Figura 4.54: Resultados do ranqueamento dos modelos pelo método MCS.

### 4.5.1 Ranqueamento dos Modelos a Posteriori

Realizando-se uma análise a posteriori dos melhores modelos, levando-se em conta os que obtiveram o menor MAPE em cada ano, independentemente de experimento, obtém-se um resultado diferente.

Os modelos de mudança markoviana tem o melhor desempenho preditivo em 60% dos casos, e o SARIMAX apenas em 20%. Os modelos univariados com regressão múltipla também predominam sobre os univariados em 60% das vezes. A tabela 4.10 e a figura 4.55 apresentam os resultados.

A curva de arrecadação dos dez anos de experimentos considerando as previsões realizadas pelos modelos com melhor MAPE está representada na figura 4.56. Pode-se notar que a previsão ainda não conseguiu acompanhar a queda da COVID em 2020, mas a queda da mudança de alíquota em 2022 foi acompanhada.

O MAPE resultante desta previsão de 10 anos dos melhores modelos pode ser visto na tabela 4.11 e na figura 4.57 em comparação com o MAPE do melhor modelo escolhido pelo Model Confidence Set.

Ano	Modelo	Experimento	Estratégia	MAPE
2013	Markov-Mean-Var	univ. reg. múltipla	markov-switching	0.027844
2014	Markov-Mean-Var	univ. reg. múltipla	markov-switching	0.018645
2015	SARIMAX(0,1,1),(1,0,1)	univariado	sarimax	0.025731
2016	Mean	univariado	naive	0.029089
2017	Markov-Fixed	univ. reg. múltipla	markov-switching	0.016581
2018	SARIMAX(0,1,1),(1,0,1)	univ. reg. múltipla	sarimax	0.016283
2019	Markov-Fixed	univ. reg. múltipla	markov-switching	0.023556
2020	Markov-Switching-Var	univariado	markov-switching	0.095877
2021	Markov-Mean-Var	univariado	markov-switching	0.053760
2022	RNN-LSTM	univ. reg. múltipla	rnn-lstm	0.035033

Tabela 4.10: Modelos com Menores MAPE por ano com escolha a posteriori, independente de experimento.

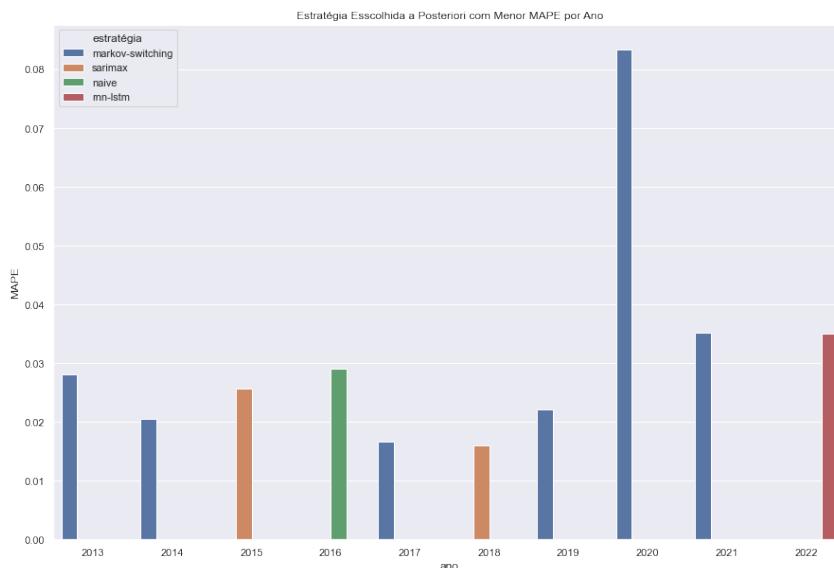


Figura 4.55: Resultados do ranqueamento dos modelos pelo menor MAPE a posteriori.

Estratégia	MAPE
Melhor MAPE	0.034240
SARIMAX Univariado	0.047816

Tabela 4.11: MAPE da previsão de 10 anos realizada pelos modelos com menores MAPE por ano com escolha a posteriori comparada com a do melhor modelo escolhido pelo MCS



Figura 4.56: Valores de arrecadação, em R\$, previstos pelos modelos com os melhores MAPEs por ano com escolha a posteriori comparados com os valores reais arrecadados.

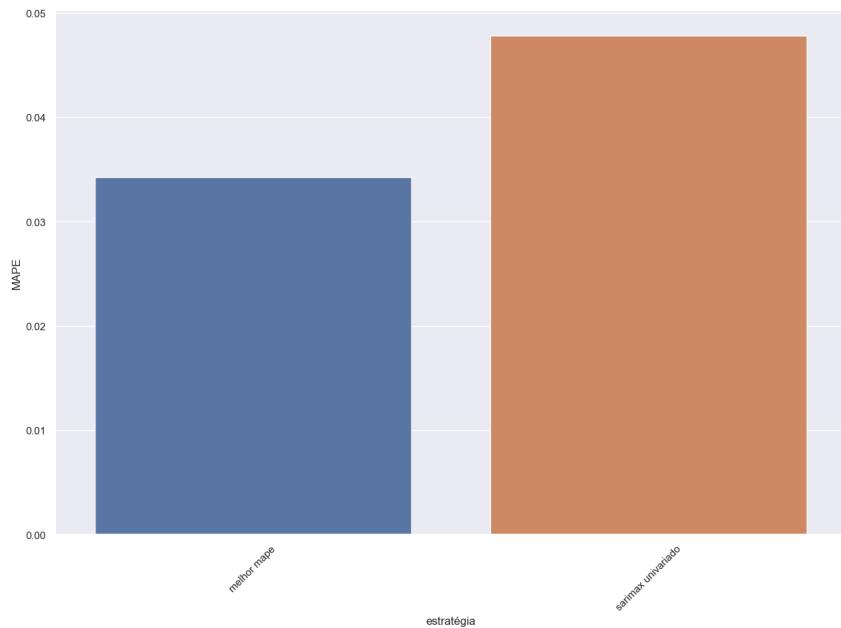


Figura 4.57: MAPE da previsão de 10 anos realizada pelos modelos com menores MAPE por ano com escolha a posteriori comparada com a do melhor modelo escolhido pelo MCS

# Capítulo 5

## Conclusão

O objetivo deste trabalho é o de comparar modelos lineares e não lineares na previsão 12 meses do ICMS paulista durante um período de 10 anos, de 2013 a 2022. A hipótese é de que os modelos não lineares teriam melhor desempenho uma vez que se intui que o processo gerador da série de arrecadação do ICMS não seja linear.

Foram realizados experimentos univariados, somente com a série do ICMS e *dummies de intervenção*, e univariados com regressão múltipla, incluindo-se as séries exógenas do IBC-Br, IGP-M, Dólar e quantidade de dias úteis do mês. E foram testados modelos lineares SARIMA e SARIMAX, e modelos não lineares de mudança markoviana e redes neurais recorrentes LSTM. As *dummies* de intervenção foram deixadas de fora dos modelos não lineares.

A escolha dos melhores modelos de cada ano foi feita baseada no critério de informação de Schwarz para os modelos estatísticos e no melhor MAPE em validação para as redes neurais. Comparando-se a métrica MAPE para todos os modelos, foi notado um melhor desempenho de cada estratégia dependendo do ano analisado, com os modelos de mudança markoviana obtendo desempenho superior em metade dos períodos. A rede neural, porém, foi superior aos modelos lineares somente no último ano de experimentos, 2022.

O uso do Model Confidence Set, entretanto, mostrou um resultado diferente, criando um ranking com os cinco melhores modelos no qual os modelos não lineares aparecem na terceira e na quinta posição. Os dois primeiros são lineares, a saber:

1. SARIMAX Reg. Múltipla
2. SARIMAX Univariado

### 3. Markov-Switching Reg. Múltipla

### 4. SARIMA Univariado

### 5. Markov-Switching Univariado

As redes neurais foram eliminadas. Embora tenham sido as melhores em 2022, tiveram um desempenho muito irregular nos outros anos, comparando-se ao modelo Naïve, que era simplesmente a média mensal de arrecadação do ano anterior. As redes neurais são instáveis e difíceis de parametrizar. Uma arquitetura que obteve bom desempenho em um ano simplesmente não funciona no ano seguinte, tornando difícil sua utilização. Além disso, conforme ([CERQUEIRA; TORGÓ; SOARES, 2019](#)), o desempenho preditivo dos modelos de machine learning aumenta quando a amostra é grande, o que não era o caso deste trabalho, que possuía 240 amostras apenas.

Na análise geral de desempenho de todos os modelos, e não apenas dos selecionados por meio do BIC, os modelos de mudança markoviana univariados com regressão múltipla foram os que obtiveram o melhor resultado preditivo de acordo com o MAPE. Infelizmente, situação que só pode ser observada a posteriori, isto é, quando já se tem o realizado a mão, o que torna a previsão inútil.

Conclui-se, portanto, a partir dos resultados obtidos, que tanto os modelos não lineares de mudança markoviana sem *dummies* de intervenção quanto os modelos lineares SARIMAX com *dummies* de intervenção tem resultados similares na previsão de arrecadação do ICMS paulista. As redes neurais, porém, não se mostraram eficazes. Além disso, o uso das variáveis exógenas escolhidas - Dias Úteis, IBC-Br e IGP-M - melhorou as previsões, pois os modelos do experimento univariado com regressão foram superiores aos seus respectivos modelos no experimento univariado. Já o Dólar melhorou as previsões dos modelos SARIMAX, mas não as dos modelos de mudança markoviana.

Um resultado adicional esperado deste trabalho era uma datação de ciclos de negócio do estado de São Paulo com base na arrecadação a partir dos modelos de mudança markoviana. O objetivo era o de datar os períodos de expansão e retração da arrecadação, a luz dos trabalhos de ([CHAUDET, 2002](#)) com o PIB brasileiro e ([HAMILTON, 1989](#)) e ([DOORNIK, 2013](#)) com o PIB americano. Houve a datação de ciclos, é verdade, mas não foram coincidentes com as datas de recessão e expansão do PIB e não houve uma grande alternância de ciclos, mas sim dois ciclos longos na maioria das vezes, ou um ciclo bem

curto seguido de outro longo.

O que pode explicar a dificuldade na datação dos ciclos é que a arrecadação é influenciada não apenas pela atividade econômica, mas o componente inflacionário também tem grande impacto. E há um terceiro componente, que não é diretamente dependente das variáveis econômicas, que são as alíquotas de imposto. A combinação destes fatores provavelmente prejudicou o desempenho dos modelos de mudança markoviana. Uma forma de se contornar o problema seria utilizar a arrecadação real, deflacionada, e analisar as mudanças de ciclo causadas apenas pelas componentes do ciclo de negócios e eventuais variações legais na alíquota. Porém, isto foi deixado para trabalhos posteriores.

Embora os resultados das redes neurais não tenham sido animadores, o uso de modelos de aprendizado de máquina não está descartado. Como trabalhos futuros tem-se o caminho apontado por ([MASINI; MEDEIROS; MENDES, 2023](#)), que mostra o crescente uso de modelos de Florestas Aleatórias (*Random Forests*) na previsão de séries de tempo. E novos experimentos podem ser feitos com redes neurais recorrentes LSTM utilizando-se séries de faturamento por setor da economia com dados provenientes da base da Nota Fiscal Eletrônica, a NF-e.

A questão das mudanças de regime também pode continuar sendo explorada, por exemplo com o uso de modelos autorregressivos com limiares TAR<sup>1</sup>. Mais sobre o model TAR em ([ENDERS, 2014](#)), capítulo 7. É interessante investigar se existe ganho de previsão utilizando-se este tipo de modelo com os limiares se iniciando nos pontos onde o X13 Arima SEATS indica a presença de *outliers* do tipo mudança de nível, como os apontados em 2008 e 2020.

Entretanto, a reforma tributária aprovada no momento da finalização deste trabalho provavelmente irá transformar o cenário arrecadatório de tal modo que ainda haverá muitas mudanças a serem estudadas nos modelos de previsão de arrecadação do Estado de São Paulo.

---

<sup>1</sup> *Threshold Auto Regressive*

# Apêndice A

## Dados de Arrecadação do ICMS

Data	Valor Nominal R\$	Data	Valor Nominal R\$
Fev/2003	3.196.557.724,20	Fev/2005	3.877.723.655,10
Mar/2003	3.160.010.802,40	Mar/2005	3.987.763.190,46
Abr/2003	3.215.307.355,38	Abr/2005	4.165.236.057,77
Mai/2003	3.098.930.221,02	Mai/2005	4.065.209.078,51
Jun/2003	3.117.682.343,84	Jun/2005	4.129.911.714,52
Jul/2003	3.196.875.781,96	Jul/2005	4.103.806.190,53
Ago/2003	3.159.728.212,41	Ago/2005	4.157.096.900,59
Set/2003	3.340.886.550,63	Set/2005	4.308.014.895,07
Out/2003	3.423.217.817,18	Out/2005	4.337.226.225,49
Nov/2003	3.484.680.639,34	Nov/2005	4.367.002.421,36
Dez/2003	3.989.413.564,99	Dez/2005	4.540.354.524,22
Jan/2004	3.575.094.652,55	Jan/2006	4.500.855.445,15
Fev/2004	3.262.938.246,12	Fev/2006	4.183.616.162,92
Mar/2004	3.469.352.922,05	Mar/2006	4.294.516.987,67
Abr/2004	3.656.956.768,36	Abr/2006	4.438.785.048,97
Mai/2004	3.667.439.628,98	Mai/2006	4.527.037.198,06
Jun/2004	3.831.564.821,35	Jun/2006	4.516.993.117,50
Jul/2004	3.742.950.400,95	Jul/2006	4.389.310.788,52
Ago/2004	3.916.334.996,46	Ago/2006	4.500.399.917,52
Set/2004	4.041.519.027,37	Set/2006	4.669.949.268,70
Out/2004	4.010.443.601,21	Out/2006	4.712.939.035,12
Nov/2004	4.129.992.676,10	Nov/2006	4.766.434.794,35
Dez/2004	4.110.988.326,04	Dez/2006	4.820.778.027,05

Tabela A.1: Arrecadação nominal do ICMS de janeiro de 2003 a dezembro de 2006

Data	Valor Nominal R\$	Data	Valor Nominal R\$
Jan/2007	4.980.597.529,73	Jan/2010	6.638.984.545,44
Fev/2007	4.689.489.271,70	Fev/2010	6.985.707.804,74
Mar/2007	4.610.978.827,36	Mar/2010	7.236.913.493,86
Abr/2007	4.909.166.624,86	Abr/2010	7.571.526.621,90
Mai/2007	5.092.643.328,00	Mai/2010	7.224.327.645,09
Jun/2007	5.037.811.780,39	Jun/2010	7.598.451.199,73
Jul/2007	5.112.757.722,85	Jul/2010	7.472.720.997,40
Ago/2007	5.210.713.898,12	Ago/2010	7.617.332.169,92
Set/2007	5.271.335.316,22	Set/2010	7.753.536.508,94
Out/2007	5.276.543.462,31	Out/2010	7.857.816.124,39
Nov/2007	5.562.282.430,73	Nov/2010	7.975.403.666,83
Dez/2007	5.638.954.129,38	Dez/2010	8.650.174.168,75
Jan/2008	5.887.631.046,01	Jan/2011	7.683.174.282,14
Fev/2008	5.412.915.906,78	Fev/2011	7.712.329.698,15
Mar/2008	5.282.664.678,27	Mar/2011	7.885.237.227,41
Abr/2008	5.692.820.799,06	Abr/2011	8.185.187.568,04
Mai/2008	6.151.430.280,00	Mai/2011	8.275.920.566,50
Jun/2008	6.324.942.790,00	Jun/2011	8.194.212.355,95
Jul/2008	6.401.730.742,70	Jul/2011	8.131.284.177,03
Ago/2008	6.484.860.548,25	Ago/2011	8.415.752.465,10
Set/2008	6.803.165.624,60	Set/2011	8.771.883.265,52
Out/2008	7.166.597.572,30	Out/2011	8.622.925.726,59
Nov/2008	6.353.422.822,77	Nov/2011	8.513.016.137,28
Dez/2008	6.542.012.157,16	Dez/2011	9.499.312.159,16
Jan/2009	5.991.830.200,46	Jan/2012	8.415.596.140,01
Fev/2009	5.507.604.548,07	Fev/2012	7.817.540.140,34
Mar/2009	5.839.373.950,24	Mar/2012	8.641.820.818,32
Abr/2009	5.856.156.442,25	Abr/2012	8.978.636.524,33
Mai/2009	5.995.709.494,65	Mai/2012	8.658.720.393,98
Jun/2009	6.182.688.710,69	Jun/2012	8.728.689.587,72
Jul/2009	6.358.778.500,59	Jul/2012	8.646.154.936,04
Ago/2009	6.442.940.045,10	Ago/2012	9.020.436.696,97
Set/2009	6.770.176.236,50	Set/2012	9.082.476.429,42
Out/2009	7.132.104.717,67	Out/2012	9.544.432.745,91
Nov/2009	7.100.588.837,73	Nov/2012	9.154.402.343,69
Dez/2009	7.991.949.426,17	Dez/2012	10.250.974.504,50

Tabela A.2: Arrecadação nominal do ICMS de janeiro de 2007 a dezembro de 2012

Data	Valor Nominal R\$	Data	Valor Nominal R\$
Jan/2013	8.799.359.392,64	Jan/2016	10.049.097.279,23
Fev/2013	8.690.433.378,92	Fev/2016	10.023.125.262,37
Mar/2013	8.698.444.956,87	Mar/2016	10.034.865.566,64
Abr/2013	9.446.983.707,28	Abr/2016	9.838.064.791,47
Mai/2013	9.497.988.773,88	Mai/2016	9.984.410.117,04
Jun/2013	9.133.476.764,70	Jun/2016	9.883.704.182,40
Jul/2013	9.657.128.835,02	Jul/2016	9.569.317.048,65
Ago/2013	9.604.406.882,29	Ago/2016	10.121.992.736,36
Set/2013	10.115.359.517,45	Set/2016	10.039.723.981,38
Out/2013	10.209.787.506,00	Out/2016	10.403.468.908,44
Nov/2013	9.712.794.926,82	Nov/2016	10.265.349.291,54
Dez/2013	10.767.619.675,75	Dez/2016	10.980.038.207,93
Jan/2014	9.803.278.286,71	Jan/2017	10.728.863.702,74
Fev/2014	9.687.760.169,38	Fev/2017	9.732.299.924,95
Mar/2014	9.465.054.142,84	Mar/2017	10.040.514.202,37
Abr/2014	9.635.429.590,53	Abr/2017	10.675.676.975,84
Mai/2014	9.819.079.302,47	Mai/2017	10.143.032.856,54
Jun/2014	9.677.264.495,05	Jun/2017	10.588.513.888,33
Jul/2014	9.594.589.921,30	Jul/2017	10.366.894.922,35
Ago/2014	9.500.059.699,22	Ago/2017	10.684.258.307,78
Set/2014	9.796.970.052,46	Set/2017	10.891.863.732,32
Out/2014	10.633.628.201,90	Out/2017	11.017.213.456,19
Nov/2014	10.152.442.434,94	Nov/2017	11.242.731.882,83
Dez/2014	11.240.824.439,12	Dez/2017	11.449.887.807,12
Jan/2015	9.987.825.321,50	Jan/2018	11.718.355.597,73
Fev/2015	9.919.476.771,77	Fev/2018	10.555.550.874,05
Mar/2015	9.871.726.570,29	Mar/2018	10.511.071.630,25
Abr/2015	10.376.399.630,25	Abr/2018	11.475.965.434,78
Mai/2015	9.740.005.805,87	Mai/2018	11.081.345.886,61
Jun/2015	10.323.070.506,13	Jun/2018	10.761.981.308,79
Jul/2015	10.294.871.977,88	Jul/2018	11.462.216.151,09
Ago/2015	9.864.942.434,35	Ago/2018	11.574.145.377,49
Set/2015	10.510.189.118,94	Set/2018	11.570.676.831,84
Out/2015	10.566.387.056,66	Out/2018	11.762.024.640,15
Nov/2015	10.296.081.459,38	Nov/2018	11.501.356.472,25
Dez/2015	11.094.379.335,88	Dez/2018	11.967.867.023,03

Tabela A.3: Arrecadação nominal do ICMS de janeiro de 2013 a dezembro de 2018

Data	Valor Nominal R\$	Data	Valor Nominal R\$
Jan/2019	12.366.848.957,36	Jan/2021	14.133.473.972,75
Fev/2019	11.265.979.239,52	Fev/2021	13.701.530.194,44
Mar/2019	11.309.998.289,14	Mar/2021	14.843.084.321,24
Abr/2019	12.016.958.191,07	Abr/2021	13.667.803.234,53
Mai/2019	11.733.622.424,54	Mai/2021	14.146.225.904,01
Jun/2019	11.387.443.318,14	Jun/2021	14.652.653.103,66
Jul/2019	11.770.991.485,82	Jul/2021	15.708.379.019,07
Ago/2019	11.784.031.575,09	Ago/2021	15.975.322.472,46
Set/2019	12.147.385.372,31	Set/2021	16.224.603.498,88
Out/2019	12.517.176.554,82	Out/2021	16.640.679.330,81
Nov/2019	12.574.528.195,31	Nov/2021	17.376.032.210,00
Dez/2019	13.146.256.505,87	Dez/2021	18.537.720.178,94
Jan/2020	12.929.806.414,52	Jan/2022	16.752.501.424,61
Fev/2020	11.776.855.482,29	Fev/2022	15.829.795.775,37
Mar/2020	12.085.931.795,23	Mar/2022	16.283.523.068,62
Abr/2020	10.455.595.055,24	Abr/2022	16.829.366.573,90
Mai/2020	9.023.528.772,57	Mai/2022	16.898.411.987,88
Jun/2020	9.983.543.650,99	Jun/2022	16.974.555.107,17
Jul/2020	11.348.726.905,97	Jul/2022	17.432.098.806,50
Ago/2020	12.277.307.781,12	Ago/2022	17.562.240.974,27
Set/2020	13.098.349.990,47	Set/2022	17.262.865.345,76
Out/2020	13.877.865.858,91	Out/2022	15.650.000.000,00
Nov/2020	14.745.896.345,78	Nov/2022	16.050.000.000,00
Dez/2020	14.906.401.854,84	Dez/2022	17.671.351.427,79

Tabela A.4: Arrecadação nominal do ICMS de janeiro de 2019 a dezembro de 2022

# Apêndice B

## Sumários dos Modelos de Mudança Markoviana

### B.1 Experimento Univariado 2022

#### B.1.1 Modelo de Mudança Markoviana com Variância Fixa

Ox Professional version 8.20 (OS\_X\_64/U) (C) J.A. Doornik, 1994-2019

----- Switching - PcGive 2.08 session started at 17:39:01 on 23-05-2023 -----

Switching( 1) Modelling DLicms\_sp by MS(2)

The dataset is: /Users/asbonato/mpe-fgv-eesp/\_dissertacao/

codigo/ final\_diff/ox/variaveis\_previsao.xlsx

The estimation sample is: 2003-03-01 - 2021-12-01

	Coefficient	Std.Error	t-value	t-prob
CSeasonal	-0.0912685	0.01317	-6.93	0.000
CSeasonal_1	-0.148519	0.01563	-9.50	0.000
CSeasonal_2	-0.0632533	0.01481	-4.27	0.000
CSeasonal_3	-0.0245879	0.01311	-1.87	0.062
CSeasonal_4	-0.0599106	0.01400	-4.28	0.000
CSeasonal_5	-0.0601961	0.01365	-4.41	0.000

CSeasonal_6	-0.0490031	0.01274	-3.85	0.000
CSeasonal_7	-0.0465071	0.01279	-3.64	0.000
CSeasonal_8	-0.0274183	0.01257	-2.18	0.030
CSeasonal_9	-0.0287015	0.01292	-2.22	0.027
CSeasonal_10	-0.0591054	0.01267	-4.67	0.000
Constant(0)	0.00913191	0.009848	0.927	0.355
Constant(1)	0.00822816	0.002744	3.00	0.003
DLicms_sp_1(0)	0.563556	0.2016	2.79	0.006
DLicms_sp_1(1)	-0.537388	0.08973	-5.99	0.000

	Coefficient	Std.Error		
sigma	0.0342371	0.002040		
p_{0 0}	0.568296	0.1982		
p_{1 1}	0.927749	0.06309		
log-likelihood	421.45175			
no. of observations	226	no. of parameters	18	
AIC	-3.57036947	SC	-3.29793748	
mean(DLicms_sp)	0.00777758	se(DLicms_sp)	0.0532509	

Linearity LR-test Chi^2(4) = 34.624 [0.0000]\*\* approximate  
upperbound: [0.0000]\*\*

Transition probabilities  $p_{i|j} = P(\text{Regime } i \text{ at } t+1 | \text{Regime } j \text{ at } t)$

	Regime 0,t	Regime 1,t
Regime 0,t+1	0.56830	0.072251
Regime 1,t+1	0.43170	0.92775

Used uniform probabilities to start recursion

Std.Error based on numerical Hessian matrix

SQPF using analytical derivatives (eps1=0.0001; eps2=0.005):

Strong convergence

Used starting values:

-0.10047	-0.12889	-0.054481	-0.027353
-0.064455	-0.049311	-0.042220	-0.038367
-0.023862	-0.027764	-0.057827	0.010708
0.0062286	-0.26958	-0.29464	0.040407
0.86903	0.73389		

Regime classification based on smoothed probabilities

Regime 0	months	avg.prob.
2007-05-01 - 2007-05-01	1	0.533
2008-03-01 - 2008-06-01	4	0.799
2008-12-01 - 2008-12-01	1	0.624
2019-01-01 - 2019-01-01	1	0.592
2020-05-01 - 2020-05-01	1	1.000
2020-07-01 - 2020-11-01	5	0.964
2021-07-01 - 2021-07-01	1	0.539

Total: 14 months (6.19%) with average duration of 2.00 months.

Regime 1	months	avg.prob.
2003-03-01 - 2007-04-01	50	0.876
2007-06-01 - 2008-02-01	9	0.821
2008-07-01 - 2008-11-01	5	0.748
2009-01-01 - 2018-12-01	120	0.925
2019-02-01 - 2020-04-01	15	0.894
2020-06-01 - 2020-06-01	1	1.000
2020-12-01 - 2021-06-01	7	0.929
2021-08-01 - 2021-12-01	5	0.744

Total: 212 months (93.81%) with average duration of 26.50 months.

Descriptive statistics for scaled residuals:

Normality test: Chi^2(2) = 32.333 [0.0000]\*\*  
 ARCH 1-1 test: F(1,206) = 2.1995 [0.1396]  
 Portmanteau(36): Chi^2(36) = 58.671 [0.0099]\*\*

Forecasting DLicms\_sp from 2022(1) to 2023(1)

Horizon	Forecast	(SE)	Actual
1	-0.049667	0.033932	-0.10126
2	-0.067522	0.036879	-0.056654
3	0.024870	0.036896	0.028260
4	0.029343	0.036956	0.032972
5	-0.0077358	0.037332	0.0040943
6	0.0059610	0.037559	0.0044958
7	0.011969	0.036586	0.026598
8	0.012185	0.035860	0.0074379
9	0.031189	0.036420	-0.017194
10	0.022693	0.036326	-0.098087
11	-0.0044875	0.036868	0.025238
12	0.064933	0.037041	0.096236

Standard errors based on M=10000 replications

mean(Error) = -0.010133 RMSE = 0.042794  
 SD(Error) = 0.041577 MAPE = 90.694

Forecasts of probabilities to be in each regime

Horizon P[Regime 0] P[Regime 1]

1	0.18403	0.81597
2	0.16354	0.83646
3	0.15337	0.84663
4	0.14833	0.85167
5	0.14583	0.85417
6	0.14459	0.85541
7	0.14397	0.85603
8	0.14367	0.85633

9	0.14352	0.85648
10	0.14344	0.85656
11	0.14340	0.85660
12	0.14339	0.85661

Forecasts for each regime

Horizon	y[Regime 0]	y[Regime 1]
1	0.0092068	-0.062945
2	-0.079326	-0.051592
3	-0.043953	0.027572
4	0.014646	0.023696
5	0.012348	-0.0095437
6	0.010767	0.0080333
7	0.021069	0.0097806
8	0.029371	0.011338
9	0.053138	0.029590
10	0.065249	0.018498
11	0.041671	-0.0059453
12	0.087488	0.066296

### B.1.2 Modelo de Mudança Markoviana com Mudança de Variância

Ox Professional version 8.20 (OS\_X\_64/U) (C) J.A. Doornik, 1994-2019

----- Switching - PcGive 2.08 session started at 17:52:31 on 23-05-2023 -----

Switching( 1) Modelling DLicms\_sp by MS(2)

The dataset is: /Users/asbonato/mpe-fgv-eesp/\_dissertacao/codigo/final\_diff/ox/variaveis\_previsao.xlsx

The estimation sample is: 2003-03-01 - 2021-12-01

Coefficient	Std.Error	t-value	t-prob
-------------	-----------	---------	--------

CSeasonal	-0.0920864	0.01230	-7.48	0.000
CSeasonal_1	-0.146678	0.01168	-12.6	0.000
CSeasonal_2	-0.0698385	0.01138	-6.14	0.000
CSeasonal_3	-0.0171556	0.01110	-1.55	0.124
CSeasonal_4	-0.0560157	0.01119	-5.01	0.000
CSeasonal_5	-0.0613883	0.01091	-5.63	0.000
CSeasonal_6	-0.0544991	0.01095	-4.98	0.000
CSeasonal_7	-0.0465791	0.01077	-4.32	0.000
CSeasonal_8	-0.0278081	0.01079	-2.58	0.011
CSeasonal_9	-0.0292780	0.01091	-2.68	0.008
CSeasonal_10	-0.0578339	0.01095	-5.28	0.000
Constant(0)	0.00999123	0.002338	4.27	0.000
Constant(1)	0.00962218	0.01260	0.764	0.446
DLicms_sp_1(0)	-0.510043	0.06281	-8.12	0.000
DLicms_sp_1(1)	0.0892876	0.1960	0.456	0.649

	Coefficient	Std.Error
sigma(0)	0.0307195	0.001778
sigma(1)	0.0694208	0.01040
p_{0 0}	0.986492	0.01003
p_{1 1}	0.923053	0.06923

log-likelihood	429.129143
no. of observations	226
no. of parameters	19
AIC	-3.62946145
SC	-3.34189434
mean(DLicms_sp)	0.00777758
se(DLicms_sp)	0.0532509

Linearity LR-test Chi^2(5) = 49.979 [0.0000]\*\* approximate  
 upperbound: [0.0000]\*\*

Transition probabilities  $p_{i|j} = P(\text{Regime } i \text{ at } t+1 \mid \text{Regime } j \text{ at } t)$   
 Regime 0,t    Regime 1,t

Regime 0,t+1	0.98649	0.076947
Regime 1,t+1	0.013508	0.92305

Used uniform probabilities to start recursion

Std.Error based on numerical Hessian matrix

SQPF using analytical derivatives (eps1=0.0001; eps2=0.005):

Strong convergence

Used starting values:

-0.097330	-0.12631	-0.054402	-0.028773	-0.067244
-0.051927	-0.045079	-0.041595	-0.027560	-0.031289
-0.059264	0.012410	0.0025034	-0.22941	-0.30668
0.029841	0.054859	0.84399	0.69807	

Regime classification based on smoothed probabilities

Regime 0	months	avg.prob.
2003-03-01 - 2007-12-01	58	0.965
2009-02-01 - 2020-03-01	134	0.979
2021-09-01 - 2021-12-01	4	0.684

Total: 196 months (86.73%) with average duration of 65.33 months.

Regime 1	months	avg.prob.
2008-01-01 - 2009-01-01	13	0.905
2020-04-01 - 2021-08-01	17	0.891

Total: 30 months (13.27%) with average duration of 15.00 months.

Descriptive statistics for scaled residuals:

Normality test: Chi^2(2) = 2.1254 [0.3455]

ARCH 1-1 test: F(1,205) = 1.9379 [0.1654]

Portmanteau(36): Chi^2(36) = 71.039 [0.0004]\*\*

Forecasting DLicms\_sp from 2022(1) to 2023(1)

Horizon	Forecast	(SE)	Actual
1	-0.050389	0.040214	-0.10126
2	-0.063550	0.043432	-0.056654
3	0.018402	0.043033	0.028260
4	0.040816	0.042745	0.032972
5	-0.0065778	0.042896	0.0040943
6	0.0059615	0.042921	0.0044958
7	0.0080511	0.041622	0.026598
8	0.015143	0.040619	0.0074379
9	0.031119	0.041081	-0.017194
10	0.023321	0.040835	-0.098087
11	-0.0022200	0.041328	0.025238
12	0.065737	0.041421	0.096236

Standard errors based on M=10000 replications

mean(Error) =	-0.011140	RMSE =	0.042887
SD(Error) =	0.041415	MAPE =	94.410

Forecasts of probabilities to be in each regime

Horizon P[Regime 0] P[Regime 1]

1	0.74534	0.25466
2	0.75487	0.24513
3	0.76353	0.23647
4	0.77141	0.22859
5	0.77858	0.22142
6	0.78510	0.21490
7	0.79103	0.20897
8	0.79643	0.20357
9	0.80134	0.19866
10	0.80580	0.19420
11	0.80986	0.19014
12	0.81355	0.18645

Forecasts for each regime

Horizon	y[Regime 0]	y[Regime 1]
1	-0.060173	-0.021756
2	-0.051066	-0.084069
3	0.021129	-0.012793
4	0.036989	0.046254
5	-0.0099604	0.012667
6	0.0086132	0.0042949
7	0.0060291	0.010437
8	0.015267	0.018905
9	0.029326	0.038432
10	0.020686	0.038706
11	-0.0034631	0.010174
12	0.066688	0.065461

Regime classification based on smoothed probabilities

Regime 0	months	avg.prob.
2003-03-01 - 2007-12-01	58	0.965
2009-02-01 - 2020-03-01	134	0.979
2021-09-01 - 2021-12-01	4	0.684

Total: 196 months (86.73%) with average duration of 65.33 months.

Regime 1	months	avg.prob.
2008-01-01 - 2009-01-01	13	0.905
2020-04-01 - 2021-08-01	17	0.891

Total: 30 months (13.27%) with average duration of 15.00 months.

Descriptive statistics for scaled residuals:

Normality test: Chi^2(2) = 2.1254 [0.3455]

ARCH 1-1 test: F(1,205) = 1.9379 [0.1654]

Portmanteau(36): Chi^2(36) = 71.039 [0.0004]\*\*

### B.1.3 Modelo de Mudança Markoviana com Componente Média-Variância

Ox Professional version 8.20 (OS\_X\_64/U) (C) J.A. Doornik, 1994-2019

---- PcGive 15.20 session started at 1:13:51 on 6-07-2023 ----

---- Switching - PcGive 2.08 session started at 1:13:57 on 6-07-2023 ----

Switching( 1) Modelling DLicms\_sp by MS\_Component(2, 2)

The dataset is: /Users/asbonato/mpe-fgv-eesp/\_dissertacao/codigo/final/ox/variaveis\_previsao.xlsx

The estimation sample is: 2003-03-01 - 2021-12-01

	Coefficient	Std.Error	t-value	t-prob
CSeasonal	-0.0797995	0.01383	-5.77	0.000
CSeasonal_1	-0.147890	0.01177	-12.6	0.000
CSeasonal_2	-0.0524688	0.01192	-4.40	0.000
CSeasonal_3	-0.00907992	0.01094	-0.830	0.408
CSeasonal_4	-0.0504436	0.01104	-4.57	0.000
CSeasonal_5	-0.0531475	0.01082	-4.91	0.000
CSeasonal_6	-0.0466426	0.01092	-4.27	0.000
CSeasonal_7	-0.0424776	0.01072	-3.96	0.000
CSeasonal_8	-0.0229310	0.01086	-2.11	0.036
CSeasonal_9	-0.0253598	0.01085	-2.34	0.020
CSeasonal_10	-0.0507058	0.01129	-4.49	0.000
Constant(0)	0.0112985	0.003041	3.72	0.000
Constant(1)	0.00801728	0.003314	2.42	0.016
DLicms_sp_1(0)	-0.127460	0.08775	-1.45	0.148
DLicms_sp_1(1)	-0.674951	0.07653	-8.82	0.000

	Coefficient	Std.Error
sigma(0)	0.0292255	0.001919

sigma(1)	0.0861463	0.01728	
p_{0 0}	0.990014	0.01064	
p_{1 1}	0.984998	0.01487	
pv_{0 0}	0.982809	0.01618	
pv_{1 1}	0.824913	0.1328	
log-likelihood	435.807718		
no. of observations	226	no. of parameters	21
AIC	-3.67086476	SC	-3.35302744
mean(DLicms_sp)	0.00777758	se(DLicms_sp)	0.0532509

Linearity LR-test Chi^2(7) = 63.336 [0.0000]\*\* approximate  
 upperbound: [0.0000]\*\*

Mean-variance component model

Mean Transition probabilities  $p_{i|j} = P(\text{Regime } i \text{ at } t+1 | \text{Regime } j \text{ at } t)$

	Regime 0,t	Regime 1,t
Regime 0,t+1	0.99001	0.015002
Regime 1,t+1	0.0099862	0.98500

Variance Transition probabilities  $pv_{i|j} = P(\text{Regime } i \text{ at } t+1 | \text{Regime } j \text{ at } t)$

	Regime 0,t	Regime 1,t
Regime 0,t+1	0.98281	0.17509
Regime 1,t+1	0.017191	0.82491

Transition probabilities  $p_{i|j} = P(\text{State } i \text{ at } t+1 | \text{State } j \text{ at } t)$

	State 0,t	State 1,t	State 2,t	State 3,t
State 0,t+1	0.97299	0.014744	0.17334	0.0026266
State 1,t+1	0.0098145	0.96807	0.0017484	0.17246
State 2,t+1	0.017019	0.00025789	0.81668	0.012375

State 3,t+1    0.00017167    0.016933    0.0082377    0.81254

Independent regimes for means and variances:

State	Mean	sigma
0	0	0
1	1	0
2	0	1
3	1	1

Used uniform probabilities to start recursion

Std.Error based on numerical Hessian matrix

SQPF using numerical derivatives (eps1=0.0001; eps2=0.005):

Strong convergence

Used starting values:

-0.097330	-0.12631	-0.054402	-0.028773	-0.067244
-0.051927	-0.045079	-0.041595	-0.027560	-0.031289
-0.059264	0.012410	0.0025034	-0.22941	-0.30668
0.029841	0.054859	0.84399	0.69807	0.90000
0.90000				

Mean regime classification based on smoothed probabilities

Mean regime 0                      months avg.prob.

2003-03-01 - 2009-11-01        81        0.934

2017-08-01 - 2021-12-01        53        0.957

Total: 134 months (59.29%) with average duration of 67.00 months.

Mean regime 1                      months avg.prob.

2009-12-01 - 2017-07-01        92        0.944

Total: 92 months (40.71%) with average duration of 92.00 months.

Variance regime classification based on smoothed probabilities

Variance regime 0                      months avg.prob.

2003-03-01 - 2008-10-01	68	0.966
2009-02-01 - 2020-03-01	134	0.977
2021-05-01 - 2021-12-01	8	0.850

Total: 210 months (92.92%) with average duration of 70.00 months.

Variance regime 1	months	avg.prob.
2008-11-01 - 2009-01-01	3	0.783
2020-04-01 - 2021-04-01	13	0.940

Total: 16 months (7.08%) with average duration of 8.00 months.

Descriptive statistics for scaled residuals:

Normality test: Chi^2(2) = 0.33212 [0.8470]

ARCH 1-1 test: F(1,203) = 1.3688 [0.2434]

Portmanteau(36): Chi^2(36) = 67.586 [0.0011]\*\*

Forecasting DLicms\_sp from 2022(1) to 2023(1)

Horizon	Forecast	(SE)	Actual
1	-0.029373	0.030707	-0.10126
2	-0.083974	0.032446	-0.056654
3	0.019872	0.032499	0.028260
4	0.047333	0.033084	0.032972
5	0.0014053	0.033692	0.0040943
6	0.0060962	0.034150	0.0044958
7	0.011766	0.033379	0.026598
8	0.014885	0.033086	0.0074379
9	0.033796	0.033950	-0.017194
10	0.027806	0.033948	-0.098087
11	0.0034231	0.034447	0.025238
12	0.058669	0.034629	0.096236

Standard errors based on M=10000 replications

mean(Error) = -0.013297 RMSE = 0.047269

SD(Error) = 0.045360 MAPE = 83.332

Forecasts of probabilities to be in each regime

Horizon	P[Regime 0]	P[Regime 1]	P[Regime 2]	P[Regime 3]
1	0.94395	0.025188	0.029315	0.0015495
2	0.92391	0.033966	0.040032	0.0020891
3	0.90641	0.042380	0.048452	0.0027610
4	0.89096	0.050483	0.055041	0.0035157
5	0.87719	0.058318	0.060170	0.0043179
6	0.86481	0.065915	0.064137	0.0051422
7	0.85355	0.073296	0.067178	0.0059711
8	0.84324	0.080480	0.069482	0.0067928
9	0.83372	0.087479	0.071201	0.0075993
10	0.82486	0.094303	0.072454	0.0083857
11	0.81655	0.10096	0.073338	0.0091490
12	0.80873	0.10746	0.073929	0.0098877

Forecasts for each regime

Horizon	y[Regime 0]	y[Regime 1]	y[Regime 2]	y[Regime 3]
1	-0.028337	-0.067050	-0.028337	-0.067050
2	-0.084568	-0.046205	-0.084568	-0.046205
3	0.018021	0.035147	0.018021	0.035147
4	0.048334	0.023627	0.048334	0.023627
5	0.0031065	-0.0099612	0.0031065	-0.0099612
6	0.0061672	0.010005	0.0061672	0.010005
7	0.012282	0.0030338	0.012282	0.0030338
8	0.015668	0.011904	0.015668	0.011904
9	0.034783	0.025464	0.034783	0.025464
10	0.029917	0.013883	0.029917	0.013883
11	0.0051916	-0.0036466	0.0051916	-0.0036466
12	0.059049	0.058891	0.059049	0.058891

## B.2 Experimento Univariado com Regressão Múltipla 2022

### B.2.1 Modelo de Mudança Markoviana com Variância Fixa

Switching( 4) Modelling DLicms\_sp by MS(2)

The dataset is: /Users/asbonato/mpe-fgv-eesp/\_dissertacao/codigo/revisao/ox/variaveis\_previsao\_markov.xlsx

The estimation sample is: 2003-03-01 - 2021-12-01

	Coefficient	Std.Error	t-value	t-prob
dias_uteis	0.00828187	0.003024	2.74	0.007
CSeasonal	-0.111468	0.01180	-9.45	0.000
CSeasonal_1	-0.107391	0.01487	-7.22	0.000
CSeasonal_2	-0.0998874	0.01721	-5.80	0.000
CSeasonal_3	-0.00977535	0.01253	-0.780	0.436
CSeasonal_4	-0.0863195	0.01161	-7.44	0.000
CSeasonal_5	-0.0526886	0.01201	-4.39	0.000
CSeasonal_6	-0.0859970	0.01360	-6.33	0.000
CSeasonal_7	-0.0632033	0.01180	-5.36	0.000
CSeasonal_8	-0.0341061	0.01206	-2.83	0.005
CSeasonal_9	-0.0580994	0.01183	-4.91	0.000
CSeasonal_10	-0.0602471	0.01348	-4.47	0.000
Constant(0)	-0.156055	0.06409	-2.43	0.016
Constant(1)	-0.187720	0.06326	-2.97	0.003
D222_igp_m(0)	0.000628313	0.0002104	2.99	0.003
D222_igp_m(1)	-0.00151287	0.0005907	-2.56	0.011
D22_ibc_br(0)	0.00283669	0.001281	2.21	0.028
D22_ibc_br(1)	0.00699882	0.001205	5.81	0.000

	Coefficient	Std.Error
sigma	0.0298550	0.001738
p_{0 0}	0.500412	0.1225

p_{1 1}	0.0495863	0.04882	
log-likelihood	429.8882		
no. of observations	226	no. of parameters	21
AIC	-3.61847965	SC	-3.30064232
mean(DLicms_sp)	0.00777758	se(DLicms_sp)	0.0532509

Linearity LR-test Chi^2(5) = 30.328 [0.0000]\*\* approximate  
 upperbound: [0.0000]\*\*

Transition probabilities  $p_{i|j} = P(\text{Regime } i \text{ at } t+1 | \text{Regime } j \text{ at } t)$

	Regime 0,t	Regime 1,t
Regime 0,t+1	0.50041	0.95041
Regime 1,t+1	0.49959	0.049586

Used uniform probabilities to start recursion

Std.Error based on numerical Hessian matrix

SQPF using analytical derivatives (eps1=0.0001; eps2=0.005):

Strong convergence

Used starting values:

0.0076729	-0.10772	-0.10089	-0.089863	-0.0089143
-0.076083	-0.045593	-0.077485	-0.053848	-0.016092
-0.052088	-0.050984	-0.15572	-0.15822	0.00038169
0.00027233	0.0042246	0.0046074	0.038577	0.86903
0.73389				

Regime classification based on smoothed probabilities

Regime 0	months	avg.prob.
2003-03-01 - 2003-03-01	1	0.913
2003-05-01 - 2003-09-01	5	0.693

2003-11-01 - 2003-12-01	2	0.954
2004-02-01 - 2004-02-01	1	0.848
2004-04-01 - 2004-06-01	3	0.891
2004-08-01 - 2004-09-01	2	0.848
2004-11-01 - 2004-11-01	1	0.985
2005-01-01 - 2005-01-01	1	1.000
2005-03-01 - 2005-07-01	5	0.703
2005-09-01 - 2005-09-01	1	0.881
2005-11-01 - 2005-11-01	1	0.901
2006-01-01 - 2006-01-01	1	0.982
2006-03-01 - 2006-09-01	7	0.714
2006-11-01 - 2006-11-01	1	0.970
2007-01-01 - 2007-07-01	7	0.756
2007-09-01 - 2007-09-01	1	0.984
2007-11-01 - 2007-11-01	1	1.000
2008-01-01 - 2008-01-01	1	1.000
2008-03-01 - 2008-10-01	8	0.764
2008-12-01 - 2008-12-01	1	0.907
2009-02-01 - 2009-03-01	2	0.690
2009-05-01 - 2009-12-01	8	0.746
2010-02-01 - 2010-02-01	1	1.000
2010-04-01 - 2010-04-01	1	0.989
2010-06-01 - 2010-06-01	1	0.989
2010-08-01 - 2010-08-01	1	0.862
2010-10-01 - 2010-12-01	3	0.838
2011-02-01 - 2011-02-01	1	0.986
2011-04-01 - 2011-09-01	6	0.722
2011-11-01 - 2011-12-01	2	0.788
2012-02-01 - 2012-02-01	1	0.928
2012-04-01 - 2012-04-01	1	0.988
2012-06-01 - 2012-06-01	1	0.964
2012-08-01 - 2012-10-01	3	0.851

2012-12-01 - 2012-12-01	1	1.000
2013-02-01 - 2013-05-01	4	0.799
2013-07-01 - 2013-07-01	1	0.952
2013-09-01 - 2013-09-01	1	0.992
2013-12-01 - 2013-12-01	1	0.995
2014-02-01 - 2014-03-01	2	0.876
2014-05-01 - 2014-10-01	6	0.745
2014-12-01 - 2014-12-01	1	0.994
2015-02-01 - 2015-02-01	1	0.999
2015-04-01 - 2015-04-01	1	0.997
2015-06-01 - 2015-07-01	2	0.845
2015-09-01 - 2015-09-01	1	0.991
2015-11-01 - 2015-12-01	2	0.772
2016-02-01 - 2016-03-01	2	0.903
2016-05-01 - 2016-05-01	1	0.997
2016-07-01 - 2016-08-01	2	0.811
2016-10-01 - 2017-01-01	4	0.762
2017-03-01 - 2017-09-01	7	0.753
2017-11-01 - 2017-11-01	1	0.982
2018-01-01 - 2018-01-01	1	1.000
2018-03-01 - 2018-05-01	3	0.915
2018-07-01 - 2018-07-01	1	0.999
2018-09-01 - 2018-09-01	1	0.983
2018-11-01 - 2018-11-01	1	0.804
2019-01-01 - 2019-01-01	1	0.991
2019-03-01 - 2019-09-01	7	0.710
2019-11-01 - 2019-11-01	1	0.978
2020-01-01 - 2020-01-01	1	0.999
2020-03-01 - 2020-03-01	1	0.999
2020-06-01 - 2020-11-01	6	0.997
2021-01-01 - 2021-03-01	3	0.993
2021-05-01 - 2021-11-01	7	0.858

Total: 158 months (69.91%) with average duration of 2.39 months.

Regime 1	months	avg.prob.
2003-04-01 - 2003-04-01	1	0.692
2003-10-01 - 2003-10-01	1	0.585
2004-01-01 - 2004-01-01	1	0.921
2004-03-01 - 2004-03-01	1	0.579
2004-07-01 - 2004-07-01	1	0.733
2004-10-01 - 2004-10-01	1	0.533
2004-12-01 - 2004-12-01	1	0.991
2005-02-01 - 2005-02-01	1	0.634
2005-08-01 - 2005-08-01	1	0.540
2005-10-01 - 2005-10-01	1	0.506
2005-12-01 - 2005-12-01	1	0.809
2006-02-01 - 2006-02-01	1	0.556
2006-10-01 - 2006-10-01	1	0.694
2006-12-01 - 2006-12-01	1	0.905
2007-08-01 - 2007-08-01	1	0.569
2007-10-01 - 2007-10-01	1	0.846
2007-12-01 - 2007-12-01	1	0.912
2008-02-01 - 2008-02-01	1	0.739
2008-11-01 - 2008-11-01	1	0.997
2009-01-01 - 2009-01-01	1	0.638
2009-04-01 - 2009-04-01	1	0.625
2010-01-01 - 2010-01-01	1	0.999
2010-03-01 - 2010-03-01	1	0.560
2010-05-01 - 2010-05-01	1	0.852
2010-07-01 - 2010-07-01	1	0.657
2010-09-01 - 2010-09-01	1	0.511
2011-01-01 - 2011-01-01	1	0.976
2011-03-01 - 2011-03-01	1	0.555
2011-10-01 - 2011-10-01	1	0.514
2012-01-01 - 2012-01-01	1	0.967

2012-03-01 - 2012-03-01	1	0.764
2012-05-01 - 2012-05-01	1	0.870
2012-07-01 - 2012-07-01	1	0.607
2012-11-01 - 2012-11-01	1	0.707
2013-01-01 - 2013-01-01	1	0.998
2013-06-01 - 2013-06-01	1	0.877
2013-08-01 - 2013-08-01	1	0.740
2013-10-01 - 2013-11-01	2	0.568
2014-01-01 - 2014-01-01	1	0.944
2014-04-01 - 2014-04-01	1	0.727
2014-11-01 - 2014-11-01	1	0.852
2015-01-01 - 2015-01-01	1	0.975
2015-03-01 - 2015-03-01	1	0.578
2015-05-01 - 2015-05-01	1	0.860
2015-08-01 - 2015-08-01	1	0.765
2015-10-01 - 2015-10-01	1	0.628
2016-01-01 - 2016-01-01	1	0.509
2016-04-01 - 2016-04-01	1	0.928
2016-06-01 - 2016-06-01	1	0.807
2016-09-01 - 2016-09-01	1	0.776
2017-02-01 - 2017-02-01	1	0.817
2017-10-01 - 2017-10-01	1	0.609
2017-12-01 - 2017-12-01	1	0.931
2018-02-01 - 2018-02-01	1	0.786
2018-06-01 - 2018-06-01	1	0.985
2018-08-01 - 2018-08-01	1	0.796
2018-10-01 - 2018-10-01	1	0.665
2018-12-01 - 2018-12-01	1	0.587
2019-02-01 - 2019-02-01	1	0.916
2019-10-01 - 2019-10-01	1	0.642
2019-12-01 - 2019-12-01	1	0.756
2020-02-01 - 2020-02-01	1	0.894

2020-04-01 - 2020-05-01	2	0.987
2020-12-01 - 2020-12-01	1	0.999
2021-04-01 - 2021-04-01	1	1.000
2021-12-01 - 2021-12-01	1	0.718

Total: 68 months (30.09%) with average duration of 1.03 months.

Descriptive statistics for scaled residuals:

Normality test: Chi^2(2) = 39.113 [0.0000]\*\*  
 ARCH 1-1 test: F(1,203) = 15.440 [0.0001]\*\*  
 Portmanteau(36): Chi^2(36) = 58.332 [0.0107]\*

Forecasting DLicms\_sp from 2022(1) to 2023(1)

Horizon	Forecast	(SE)	Actual
1	-0.057221	0.029855	-0.10126
2	-0.043221	0.029855	-0.056654
3	0.019853	0.029855	0.028260
4	0.028303	0.029855	0.032972
5	-0.0045142	0.029855	0.0040943
6	0.015434	0.029855	0.0044958
7	-0.0027499	0.029855	0.026598
8	0.028113	0.029855	0.0074379
9	0.020393	0.029855	-0.017194
10	0.0086083	0.029855	-0.098087
11	-0.0026502	0.029855	0.025238
12	0.070554	0.029855	0.096236
mean(Error) = -0.010730		RMSE = 0.038653	
SD(Error) = 0.037134		MAPE = 118.13	

Forecasts of probabilities to be in each regime

Horizon P[Regime 0] P[Regime 1]

1	0.82336	0.17664
---	---------	---------

2	0.57990	0.42010
3	0.68946	0.31054
4	0.64016	0.35984
5	0.66234	0.33766
6	0.65236	0.34764
7	0.65685	0.34315
8	0.65483	0.34517
9	0.65574	0.34426
10	0.65533	0.34467
11	0.65551	0.34449
12	0.65543	0.34457

Forecasts for each regime

Horizon	y[Regime 0]	y[Regime 1]
1	-0.040293	-0.13613
2	-0.022842	-0.071351
3	0.023176	0.012476
4	0.051227	-0.012477
5	0.010845	-0.034642
6	0.033333	-0.018155
7	0.010109	-0.027365
8	0.043903	-0.0018411
9	0.044531	-0.025586
10	0.027278	-0.026888
11	0.018596	-0.043078
12	0.092496	0.028816

## B.2.2 Modelo de Mudança Markoviana com Mudança de Variância

Switching( 5 ) Modelling DLicms\_sp by MS(2)

The dataset is: /Users/asbonato/mpe-fgv-eesp/\_dissertacao/codigo/revisao/ox/variaveis\_previsao\_markov.xlsx

The estimation sample is: 2003-03-01 - 2021-12-01

	Coefficient	Std.Error	t-value	t-prob
dias_uteis	0.0122092	0.002601	4.69	0.000
CSeasonal	-0.0571874	0.01234	-4.63	0.000
CSeasonal_1	-0.0990276	0.01198	-8.26	0.000
CSeasonal_2	-0.0461487	0.01344	-3.43	0.001
CSeasonal_3	0.0174788	0.009686	1.80	0.073
CSeasonal_4	-0.0616174	0.01004	-6.14	0.000
CSeasonal_5	-0.0365472	0.009919	-3.68	0.000
CSeasonal_6	-0.0593045	0.01023	-5.79	0.000
CSeasonal_7	-0.0471353	0.009562	-4.93	0.000
CSeasonal_8	-0.0117433	0.009425	-1.25	0.214
CSeasonal_9	-0.0326074	0.009338	-3.49	0.001
CSeasonal_10	-0.0265912	0.01068	-2.49	0.014
Constant(0)	-0.243869	0.05391	-4.52	0.000
Constant(1)	-0.251952	0.05570	-4.52	0.000
D222_igp_m(0)	-0.000781277	0.0003390	-2.30	0.022
D222_igp_m(1)	0.000471544	0.0004258	1.11	0.269
D22_ibc_br(0)	0.000557369	0.0009635	0.578	0.564
D22_ibc_br(1)	0.00767356	0.001871	4.10	0.000

	Coefficient	Std.Error
sigma(0)	0.0194272	0.001740
sigma(1)	0.0590377	0.005834
p_{0 0}	0.843963	0.05101
p_{1 1}	0.699170	0.08291

log-likelihood	436.428164		
no. of observations	226	no. of parameters	22
AIC	-3.66750588	SC	-3.33453345
mean(DLicms_sp)	0.00777758	se(DLicms_sp)	0.0532509

Linearity LR-test Chi^2(6) = 43.408 [0.0000]\*\* approximate  
 upperbound: [0.0000]\*\*

Transition probabilities  $p_{\{i|j\}} = P(\text{Regime } i \text{ at } t+1 \mid \text{Regime } j \text{ at } t)$

	Regime 0,t	Regime 1,t
Regime 0,t+1	0.84396	0.30083
Regime 1,t+1	0.15604	0.69917

Used uniform probabilities to start recursion

Std.Error based on numerical Hessian matrix

SQPF using analytical derivatives (eps1=0.0001; eps2=0.005):

Strong convergence

Used starting values:

0.0093769	-0.10436	-0.094334	-0.083049	-0.012052
-0.078148	-0.050295	-0.073674	-0.059081	-0.026572
-0.052534	-0.051362	-0.18733	-0.19972	0.00034632
0.00043968	0.0018668	0.0058835	0.028860	0.049659
0.84399	0.69807			

Regime classification based on smoothed probabilities

Regime 0	months	avg.prob.
2003-03-01 - 2003-11-01	9	0.869
2004-02-01 - 2004-11-01	10	0.820
2005-01-01 - 2008-04-01	40	0.894
2008-06-01 - 2008-10-01	5	0.771
2009-06-01 - 2009-11-01	6	0.768
2010-03-01 - 2010-12-01	10	0.863
2011-03-01 - 2011-11-01	9	0.871
2012-04-01 - 2012-10-01	7	0.841
2013-03-01 - 2013-10-01	8	0.731

2014-03-01 - 2014-09-01	7	0.784
2015-03-01 - 2015-05-01	3	0.795
2015-07-01 - 2015-12-01	6	0.693
2016-05-01 - 2016-08-01	4	0.713
2016-10-01 - 2017-04-01	7	0.859
2017-06-01 - 2018-06-01	13	0.901
2018-08-01 - 2020-03-01	20	0.907
2020-12-01 - 2021-01-01	2	0.524
2021-08-01 - 2021-12-01	5	0.761

Total: 171 months (75.66%) with average duration of 9.50 months.

Regime 1	months	avg.prob.
2003-12-01 - 2004-01-01	2	0.998
2004-12-01 - 2004-12-01	1	0.984
2008-05-01 - 2008-05-01	1	1.000
2008-11-01 - 2009-05-01	7	0.847
2009-12-01 - 2010-02-01	3	0.977
2011-01-01 - 2011-02-01	2	0.997
2011-12-01 - 2012-03-01	4	0.909
2012-11-01 - 2013-02-01	4	0.981
2013-11-01 - 2014-02-01	4	0.961
2014-10-01 - 2015-02-01	5	0.933
2015-06-01 - 2015-06-01	1	0.571
2016-01-01 - 2016-04-01	4	0.887
2016-09-01 - 2016-09-01	1	0.526
2017-05-01 - 2017-05-01	1	0.855
2018-07-01 - 2018-07-01	1	0.996
2020-04-01 - 2020-11-01	8	1.000
2021-02-01 - 2021-07-01	6	0.999

Total: 55 months (24.34%) with average duration of 3.24 months.

Descriptive statistics for scaled residuals:

Normality test: Chi^2(2) = 3.2603 [0.1959]  
 ARCH 1-1 test: F(1,202) = 7.1889 [0.0079]\*\*  
 Portmanteau(36): Chi^2(36) = 73.239 [0.0002]\*\*

### B.2.3 Modelo de Mudança Markoviana com Componente Média-Variância

Switching( 6) Modelling DLicms\_sp by MS\_Component(2, 2)  
 The dataset is: /Users/asbonato/mpe-fgv-eesp/\_dissertacao/  
 codigo/revisao/ox/variaveis\_previsao\_markov.xlsx  
 The estimation sample is: 2003-03-01 - 2021-12-01

	Coefficient	Std.Error	t-value	t-prob
dias_uteis	0.0118064	0.001607	7.34	0.000
CSeasonal	-0.0991486	0.01627	-6.09	0.000
CSeasonal_1	-0.104224	0.009184	-11.3	0.000
CSeasonal_2	-0.0495728	0.008044	-6.16	0.000
CSeasonal_3	0.00771271	0.007884	0.978	0.329
CSeasonal_4	-0.0543568	0.007319	-7.43	0.000
CSeasonal_5	-0.0476456	0.007836	-6.08	0.000
CSeasonal_6	-0.0724977	0.008135	-8.91	0.000
CSeasonal_7	-0.0534669	0.007390	-7.24	0.000
CSeasonal_8	-0.0168900	0.008171	-2.07	0.040
CSeasonal_9	-0.0459084	0.008472	-5.42	0.000
CSeasonal_10	-0.0308459	0.007294	-4.23	0.000
Constant(0)	-0.241060	0.03300	-7.30	0.000
Constant(1)	-0.254761	0.03436	-7.41	0.000
D222_igp_m(0)	0.000680142	9.837e-05	6.91	0.000
D222_igp_m(1)	-0.00150856	0.0002081	-7.25	0.000
D22_ibc_br(0)	0.000887334	0.0006112	1.45	0.148
D22_ibc_br(1)	0.00889256	0.0005683	15.6	0.000

Coefficient Std.Error

sigma(0)	0.00877889	0.001361	
sigma(1)	0.0453226	0.004027	
p_{0 0}	0.742525	0.07544	
p_{1 1}	0.107148	0.1387	
pv_{0 0}	0.741497	0.09235	
pv_{1 1}	0.800660	0.07548	
log-likelihood	444.330494		
no. of observations	226	no. of parameters	24
AIC	-3.71973888	SC	-3.35649623
mean(DLicms_sp)	0.00777758	se(DLicms_sp)	0.0532509

Linearity LR-test Chi^2(8) = 59.213 [0.0000]\*\* approximate  
 upperbound: [0.0000]\*\*

#### Mean-variance component model

Mean Transition probabilities  $p_{i|j} = P(\text{Regime } i \text{ at } t+1 | \text{Regime } j \text{ at } t)$

	Regime 0,t	Regime 1,t
Regime 0,t+1	0.74252	0.89285
Regime 1,t+1	0.25748	0.10715

Variance Transition probabilities  $pv_{i|j} = P(\text{Regime } i \text{ at } t+1 | \text{Regime } j \text{ at } t)$

	Regime 0,t	Regime 1,t
Regime 0,t+1	0.74150	0.19934
Regime 1,t+1	0.25850	0.80066

Transition probabilities  $p_{i|j} = P(\text{State } i \text{ at } t+1 | \text{State } j \text{ at } t)$

	State 0,t	State 1,t	State 2,t	State 3,t
State 0,t+1	0.55058	0.66205	0.14801	0.17798
State 1,t+1	0.19092	0.079450	0.051325	0.021359

State 2,t+1	0.19194	0.23081	0.59451	0.71487
State 3,t+1	0.066558	0.027698	0.20615	0.085789

Independent regimes for means and variances:

State	Mean	sigma
0	0	0
1	1	0
2	0	1
3	1	1

Used uniform probabilities to start recursion

Std.Error based on numerical Hessian matrix

SQPF using numerical derivatives (eps1=0.0001; eps2=0.005):

Strong convergence

Used starting values:

0.0093769	-0.10436	-0.094334	-0.083049	-0.012052
-0.078148	-0.050295	-0.073674	-0.059081	-0.026572
-0.052534	-0.051362	-0.18733	-0.19972	0.00034632
0.00043968	0.0018668	0.0058835	0.028860	0.049659
0.84399	0.69807	0.80000	0.80000	

Mean regime classification based on smoothed probabilities

Mean regime 0	months	avg.prob.
2003-03-01 - 2003-03-01	1	0.888
2003-06-01 - 2003-06-01	1	0.961
2003-08-01 - 2004-11-01	16	0.852
2005-01-01 - 2005-04-01	4	0.960
2005-06-01 - 2005-11-01	6	0.869
2006-01-01 - 2006-11-01	11	0.926
2007-01-01 - 2007-04-01	4	0.925
2007-06-01 - 2007-06-01	1	0.971
2007-08-01 - 2008-10-01	15	0.867

2008-12-01 - 2008-12-01	1	0.904
2009-02-01 - 2009-12-01	11	0.789
2010-02-01 - 2010-12-01	11	0.891
2011-02-01 - 2011-10-01	9	0.825
2011-12-01 - 2011-12-01	1	0.957
2012-02-01 - 2012-02-01	1	0.908
2012-04-01 - 2012-09-01	6	0.906
2012-12-01 - 2012-12-01	1	0.991
2013-02-01 - 2013-05-01	4	0.925
2013-08-01 - 2013-10-01	3	0.778
2013-12-01 - 2013-12-01	1	0.955
2014-02-01 - 2014-12-01	11	0.794
2015-02-01 - 2015-04-01	3	0.977
2015-06-01 - 2016-03-01	10	0.822
2016-05-01 - 2016-05-01	1	0.970
2016-07-01 - 2016-10-01	4	0.755
2016-12-01 - 2017-01-01	2	0.934
2017-03-01 - 2017-11-01	9	0.892
2018-01-01 - 2018-10-01	10	0.822
2018-12-01 - 2019-01-01	2	0.855
2019-03-01 - 2019-04-01	2	0.877
2019-06-01 - 2020-01-01	8	0.917
2020-03-01 - 2020-03-01	1	0.989
2020-05-01 - 2020-11-01	7	0.879
2021-01-01 - 2021-03-01	3	0.998
2021-05-01 - 2021-12-01	8	0.832

Total: 189 months (83.63%) with average duration of 5.40 months.

Mean regime 1	months	avg.prob.
2003-04-01 - 2003-05-01	2	0.737
2003-07-01 - 2003-07-01	1	0.827
2004-12-01 - 2004-12-01	1	0.565
2005-05-01 - 2005-05-01	1	0.841

2005-12-01 - 2005-12-01	1	0.656
2006-12-01 - 2006-12-01	1	0.589
2007-05-01 - 2007-05-01	1	0.622
2007-07-01 - 2007-07-01	1	0.832
2008-11-01 - 2008-11-01	1	0.927
2009-01-01 - 2009-01-01	1	0.546
2010-01-01 - 2010-01-01	1	0.762
2011-01-01 - 2011-01-01	1	0.715
2011-11-01 - 2011-11-01	1	0.521
2012-01-01 - 2012-01-01	1	0.661
2012-03-01 - 2012-03-01	1	0.819
2012-10-01 - 2012-11-01	2	0.539
2013-01-01 - 2013-01-01	1	0.624
2013-06-01 - 2013-07-01	2	0.617
2013-11-01 - 2013-11-01	1	0.534
2014-01-01 - 2014-01-01	1	0.507
2015-01-01 - 2015-01-01	1	0.599
2015-05-01 - 2015-05-01	1	0.679
2016-04-01 - 2016-04-01	1	0.731
2016-06-01 - 2016-06-01	1	0.651
2016-11-01 - 2016-11-01	1	0.563
2017-02-01 - 2017-02-01	1	0.725
2017-12-01 - 2017-12-01	1	0.683
2018-11-01 - 2018-11-01	1	0.553
2019-02-01 - 2019-02-01	1	0.505
2019-05-01 - 2019-05-01	1	0.845
2020-02-01 - 2020-02-01	1	0.804
2020-04-01 - 2020-04-01	1	1.000
2020-12-01 - 2020-12-01	1	0.972
2021-04-01 - 2021-04-01	1	1.000

Total: 37 months (16.37%) with average duration of 1.09 months.

Variance regime classification based on smoothed probabilities

Variance regime 0	months	avg.prob.
2003-03-01 - 2003-11-01	9	0.758
2004-04-01 - 2004-05-01	2	0.703
2004-10-01 - 2004-10-01	1	0.584
2005-02-01 - 2005-12-01	11	0.851
2006-02-01 - 2006-11-01	10	0.860
2007-05-01 - 2007-09-01	5	0.754
2008-06-01 - 2008-09-01	4	0.761
2010-04-01 - 2010-04-01	1	0.510
2010-07-01 - 2010-08-01	2	0.583
2010-10-01 - 2010-12-01	3	0.768
2011-03-01 - 2011-10-01	8	0.846
2012-03-01 - 2012-03-01	1	0.590
2012-06-01 - 2012-10-01	5	0.771
2013-03-01 - 2013-06-01	4	0.751
2014-06-01 - 2014-08-01	3	0.701
2015-04-01 - 2015-04-01	1	0.509
2016-05-01 - 2016-05-01	1	0.671
2016-12-01 - 2017-03-01	4	0.705
2017-07-01 - 2017-12-01	6	0.820
2019-03-01 - 2020-04-01	14	0.802
2020-09-01 - 2021-06-01	10	0.866

Total: 105 months (46.46%) with average duration of 5.00 months.

Variance regime 1	months	avg.prob.
2003-12-01 - 2004-03-01	4	0.765
2004-06-01 - 2004-09-01	4	0.841
2004-11-01 - 2005-01-01	3	0.889
2006-01-01 - 2006-01-01	1	0.724
2006-12-01 - 2007-04-01	5	0.856
2007-10-01 - 2008-05-01	8	0.958

2008-10-01 - 2010-03-01	18	0.866
2010-05-01 - 2010-06-01	2	0.938
2010-09-01 - 2010-09-01	1	0.664
2011-01-01 - 2011-02-01	2	0.817
2011-11-01 - 2012-02-01	4	0.785
2012-04-01 - 2012-05-01	2	0.757
2012-11-01 - 2013-02-01	4	0.923
2013-07-01 - 2014-05-01	11	0.803
2014-09-01 - 2015-03-01	7	0.904
2015-05-01 - 2016-04-01	12	0.867
2016-06-01 - 2016-11-01	6	0.791
2017-04-01 - 2017-06-01	3	1.000
2018-01-01 - 2019-02-01	14	0.884
2020-05-01 - 2020-08-01	4	1.000
2021-07-01 - 2021-12-01	6	0.877

Total: 121 months (53.54%) with average duration of 5.76 months.

Descriptive statistics for scaled residuals:

Normality test: Chi^2(2) = 0.60539 [0.7388]

ARCH 1-1 test: F(1,200) = 9.5478 [0.0023]\*\*

Portmanteau(36): Chi^2(36) = 81.372 [0.0000]\*\*

Forecasting DLicms\_sp from 2022(1) to 2023(1)

Horizon	Forecast	(SE)	Actual
1	-0.061505	0.032365	-0.10126
2	-0.059151	0.031013	-0.056654
3	0.040666	0.030280	0.028260
4	0.027618	0.029882	0.032972
5	0.011874	0.029667	0.0040943
6	0.0038390	0.029550	0.0044958
7	-0.011592	0.029487	0.026598

8	0.025783	0.029453	0.0074379
9	0.026594	0.029434	-0.017194
10	-0.00084467	0.029424	-0.098087
11	0.0083201	0.029418	0.025238
12	0.060328	0.029415	0.096236
mean(Error) = -0.0099826		RMSE = 0.037164	
SD(Error) = 0.035798		MAPE = 96.401	

#### Forecasts of probabilities to be in each regime

Horizon	P[Regime 0]	P[Regime 1]	P[Regime 2]	P[Regime 3]
1	0.29567	0.058912	0.51005	0.13537
2	0.30138	0.090198	0.47035	0.13807
3	0.31984	0.091794	0.45700	0.13137
4	0.32789	0.094617	0.44818	0.12931
5	0.33252	0.095882	0.44366	0.12793
6	0.33500	0.096606	0.44117	0.12722
7	0.33634	0.096993	0.43983	0.12683
8	0.33707	0.097203	0.43910	0.12662
9	0.33747	0.097317	0.43870	0.12651
10	0.33768	0.097379	0.43849	0.12645
11	0.33780	0.097412	0.43837	0.12642
12	0.33786	0.097431	0.43831	0.12640

#### Forecasts for each regime

Horizon	y[Regime 0]	y[Regime 1]	y[Regime 2]	y[Regime 3]
1	-0.040781	-0.14745	-0.040781	-0.14745
2	-0.053344	-0.078785	-0.053344	-0.078785
3	0.030902	0.074653	0.030902	0.074653
4	0.040875	-0.018328	0.040875	-0.018328
5	0.017510	-0.0076717	0.017510	-0.0076717
6	0.011805	-0.023784	0.011805	-0.023784
7	-0.0097696	-0.017911	-0.0097696	-0.017911

8	0.031150	0.0071704	0.031150	0.0071704
9	0.041481	-0.025027	0.041481	-0.025027
10	0.0063161	-0.025676	0.0063161	-0.025676
11	0.019006	-0.028736	0.019006	-0.028736
12	0.072197	0.019170	0.072197	0.019170

# Referências Bibliográficas

- AKNOUCHE, A.; FRANCQ, C. Stationarity and ergodicity of markov switching positive conditional mean models. *Journal of Time Series Analysis*, Wiley Online Library, v. 43, n. 3, p. 436–459, 2022.
- ALMOSOVA, A.; ANDRESEN, N. Nonlinear inflation forecasting with recurrent neural networks. *Journal of Forecasting*, n/a, n. n/a. Disponível em: <<https://onlinelibrary.wiley.com/doi/abs/10.1002/for.2901>>.
- ANAS, J. et al. Business cycle analysis with multivariate markov switching models. *Growth and Cycle in the Eurozone*, p. 249–260, 2004.
- ARUNKUMAR, K. et al. Forecasting the dynamics of cumulative covid-19 cases (confirmed, recovered and deaths) for top-16 countries using statistical machine learning models: Auto-regressive integrated moving average (arima) and seasonal auto-regressive integrated moving average (sarima). *Applied Soft Computing*, v. 103, p. 107161, 2021. ISSN 1568-4946. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S1568494621000843>>.
- BELLONE, B. et al. Classical estimation of multivariate markov-switching models using msvarlib. *User Guide*, 2005.
- BENGIO, Y.; GOODFELLOW, I.; COURVILLE, A. *Deep learning*. [S.l.]: MIT press Cambridge, MA, USA, 2017. v. 1.
- BERNARDI, M.; CATANIA, L. The model confidence set package for r. *International Journal of Computational Economics and Econometrics*, v. 8, n. 2, p. 144–158, 2018. Disponível em: <<https://www.inderscienceonline.com/doi/abs/10.1504/IJCEE.2018.091037>>.
- BOROVKOVA, S.; TSIAMAS, I. An ensemble of lstm neural networks for high-frequency stock market classification. *Journal of Forecasting*, Wiley Online Library, v. 38, n. 6, p. 600–619, 2019.
- BOX, G. E. et al. *Time series analysis: forecasting and control*. [S.l.]: John Wiley & Sons, 2015.
- CERQUEIRA, V.; TORGÓ, L.; SOARES, C. *Machine Learning vs Statistical Methods for Time Series Forecasting: Size Matters*. 2019.
- CHAUDET, M. *The Brazilian business and growth cycles*. Revista Brasileira de Economia [online], 2002. Disponível em: <<https://www.scielo.br/j/rbe/a/DBb7SnLr58KMbX764nLqWDN/?lang=en#>>. Acesso em: 08-Set-2022.

CHEVALLIER, J. Global imbalances, cross-market linkages, and the financial crisis: A multivariate markov-switching analysis. *Economic Modelling*, Elsevier, v. 29, n. 3, p. 943–973, 2012.

CHEVILLON, G. Direct multi-step estimation and forecasting. *Journal of Economic Surveys*, Wiley Online Library, v. 21, n. 4, p. 746–785, 2007.

CONTRERAS, J. C. S. *Previsão de arrecadação do ICMS através de redes neurais no Brasil*. Dissertação (Mestrado) — Universidade Federal de Pernambuco, 2005.

DICKEY, D. A.; FULLER, W. A. Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association*, Taylor & Francis, v. 74, n. 366a, p. 427–431, 1979. Disponível em: <<https://doi.org/10.1080/01621459.1979.10482531>>.

DIEBOLD, F. X.; MARIANO, R. S. Comparing predictive accuracy. *Journal of Business & economic statistics*, Taylor & Francis, v. 20, n. 1, p. 134–144, 2002.

DINIZ, P. *ICMS de São Paulo Comentado*. [S.l.]: São Paulo: Ferreira, 2012.

DOORNIK, J. A. A markov-switching model with component structure for us gnp. *Economics Letters*, v. 118, n. 2, p. 265–268, 2013. ISSN 0165-1765. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0165176512005885>>.

DOORNIK, J. A. Econometric analysis with markov-switching models. *PCGiveTM15*, v. 4, 2018.

DORNELAS, A. F.; CAMPOS, L. D.; FIGUEIREDO, K. Modelos para previsão tributária utilizando redes neurais lstm. In: SBC. *Anais do XIX Encontro Nacional de Inteligência Artificial e Computacional*. [S.l.], 2022. p. 95–105.

DORNELLES, G. Z.; SCHWARTZER, F. R.; BRAATZ, J. Redes neurais aplicadas na previsão de receita de icms no rio grande do sul. 2022.

DUDA, R.; HART, P.; STORK, D. *Pattern Classification*. Wiley, 2012. ISBN 9781118586006. Disponível em: <<https://books.google.com.br/books?id=Br33IRC3PkQC>>.

ELLIOTT, G.; TIMMERMANN, A. Economic forecasting. *Journal of Economic Literature*, v. 46, n. 1, p. 3–56, 2008.

ENDERS, W. *Applied Econometric Time Series*. Wiley, 2014. (Wiley Series in Probability and Statistics). ISBN 9781118808566. Disponível em: <<https://books.google.com.br/books?id=udDxoAEACAAJ>>.

(ESTADO), S. P. *Projeto de Lei nº 247/2022 - Diretrizes Orçamentárias - Anexo II - Riscos Fiscais*. Assembleia Legislativa do Estado de São Paulo, 2022. Disponível em: <<https://www.al.sp.gov.br/norma/204520>>. Acesso em: 16-Mai-2023.

FRANCQ, C.; ZAKOÏAN, J.-M. Stationarity of multivariate markov-switching arma models. *Journal of Econometrics*, v. 102, n. 2, p. 339–364, 2001. ISSN 0304-4076. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0304407601000574>>.

FREITAS, F. D.; CIARELLI, P. M.; SOUZA, A. F. Previsão da arrecadação federal com redes neurais. In: *Anais do XIX Congresso Brasileiro de Redes Neurais*. Ouro Preto, MG: SBRN, 2009. p. 1–5.

GAMA, M. A. *Macroeconomia brasileira: que país é esse?* 1. ed. [S.l.]: Ed. do Autor, 2020. ISBN 978-65-00-11265-8.

GONZÁLEZ-RIVERA, G.; LEE, T.-H.; MISHRA, S. Forecasting volatility: A reality check based on option pricing, utility function, value-at-risk, and predictive likelihood. *International Journal of Forecasting*, v. 20, n. 4, p. 629–645, 2004. ISSN 0169-2070. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0169207003001420>>.

HAMILTON, J. D. A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica*, [Wiley, Econometric Society], v. 57, n. 2, p. 357–384, 1989. ISSN 00129682, 14680262. Disponível em: <<http://www.jstor.org/stable/1912559>>.

HAMILTON, J. D. Analysis of time series subject to changes in regime. *Journal of econometrics*, Elsevier, v. 45, n. 1-2, p. 39–70, 1990.

HANSEN, P. R.; LUNDE, A.; NASON, J. M. The model confidence set. *Econometrica*, [Wiley, Econometric Society], v. 79, n. 2, p. 453–497, 2011. ISSN 00129682, 14680262. Disponível em: <<http://www.jstor.org/stable/41057463>>.

HANSSON, M. On stock return prediction with lstm networks. 2017.

HAYKIN, S. *Neural networks: a comprehensive foundation*. [S.l.]: Prentice Hall PTR, 1998.

HOCHREITER, S.; SCHMIDHUBER, J. Long short-term memory. *Neural computation*, MIT Press, v. 9, n. 8, p. 1735–1780, 1997.

HYNDMAN, R.; ATHANASOPOULOS, G. *Forecasting: principles and practice*. OTexts, 2018. ISBN 9780987507112. Disponível em: <[https://books.google.com.br/books?id=\\_bBhDwAAQBAJ](https://books.google.com.br/books?id=_bBhDwAAQBAJ)>.

HYNDMAN, R. J.; KHANDAKAR, Y. Automatic time series forecasting: the forecast package for r. *Journal of statistical software*, v. 27, p. 1–22, 2008.

JENKINS, G. P. et al. Tax analysis and revenue forecasting. *Cambridge, Massachusetts: Harvard Institute for International Development, Harvard University*, 2000.

JOLLIFFE, I. T.; CADIMA, J. Principal component analysis: a review and recent developments. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, v. 374, n. 2065, p. 20150202, 2016. Disponível em: <<https://royalsocietypublishing.org/doi/abs/10.1098/rsta.2015.0202>>.

KHAN, S.; ALGHULAIKH, H. Arima model for accurate time series stocks forecasting. *International Journal of Advanced Computer Science and Applications*, Science and Information (SAI) Organization Limited, v. 11, n. 7, 2020.

KIM, C.-J. Dynamic linear models with markov-switching. *Journal of econometrics*, Elsevier, v. 60, n. 1-2, p. 1–22, 1994.

KIRCHGÄSSNER, G.; WOLTERS, J.; HASSELER, U. *Introduction to modern time series analysis*. [S.l.]: Springer Science & Business Media, 2012.

KWIATKOWSKI, D. et al. Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root? *Journal of Econometrics*, v. 54, n. 1, p. 159–178, 1992. ISSN 0304-4076. Disponível em: <<https://www.sciencedirect.com/science/article/pii/030440769290104Y>>.

MASINI, R. P.; MEDEIROS, M. C.; MENDES, E. F. Machine learning advances for time series forecasting. *Journal of economic surveys*, Wiley Online Library, v. 37, n. 1, p. 76–111, 2023.

MINH, T. L. Estimating the vat base: method and application. *Tax Notes International*, v. 46, p. 42, 2007.

MITCHELL, T. *Machine Learning*. McGraw-Hill Education, 1997. (McGraw-Hill international editions - computer science series). ISBN 9780070428072. Disponível em: <<https://books.google.com.br/books?id=xOGAngEACAAJ>>.

MONSELL, B.; LYTRAS, D.; FINDLEY, D. Getting started with x-13arima-seats input files (accessible version). *United States Census Bureau*, v. 8, 2013.

NELSON, C. R.; PIGER, J.; ZIVOT, E. Markov regime switching and unit-root tests. *Journal of Business & Economic Statistics*, [American Statistical Association, Taylor & Francis, Ltd.], v. 19, n. 4, p. 404–415, 2001. ISSN 07350015. Disponível em: <<http://www.jstor.org/stable/1392275>>.

NGUYEN, B. A. *Markov Regime-Switching in Forecasting Models*. Tese (Doutorado) — Carleton University, 2020.

OLIVEIRA, A. B.; PEREIRA, P. L. V. Asset allocation with markovian regime switching: Efficient frontier and tangent portfolio with regime switching. v. 38, n. 1, 2018.

PEARSON, K. Vii. mathematical contributions to the theory of evolution.—iii. regression, heredity, and panmixia. *Philosophical Transactions of the Royal Society of London. Series A, containing papers of a mathematical or physical character*, The Royal Society London, n. 187, p. 253–318, 1896.

PEARSON, K. Liii. on lines and planes of closest fit to systems of points in space. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, Taylor & Francis, v. 2, n. 11, p. 559–572, 1901. Disponível em: <<https://doi.org/10.1080/14786440109462720>>.

PEDREGOSA, F. et al. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, v. 12, p. 2825–2830, 2011.

PELLETIER, D. Regime switching for dynamic correlations. *Journal of Econometrics*, v. 131, n. 1-2, p. 445–473, 2006. Disponível em: <<https://EconPapers.repec.org/RePEc:eee:econom:v:131:y:2006:i:1-2:p:445-473>>.

PRINCE, D. de; MARÇAL, E. F.; PEREIRA, P. L. V. Forecasting industrial production using its aggregated and disaggregated series or a combination of both: Evidence from one emerging market economy. *Econometrics*, v. 10, n. 2, 2022. ISSN 2225-1146. Disponível em: <<https://www.mdpi.com/2225-1146/10/2/27>>.

- ROCHA, J. V.; PEREIRA, P. L. V. Forecast comparison with nonlinear methods for brazilian industrial production. 2015. Disponível em: <<http://hdl.handle.net/10438/13862>>.
- ROTTA, P. N.; PEREIRA, P. L. V. Analysis of contagion from the dynamic conditional correlation model with markov regime switching. *Applied Economics*, Routledge, v. 48, n. 25, p. 2367–2382, 2016. Disponível em: <<https://doi.org/10.1080/00036846.2015.1119794>>.
- SCHMIDHUBER, J. Deep learning in neural networks: An overview. *Neural networks*, Elsevier, v. 61, p. 85–117, 2015.
- SIAMI-NAMINI, S.; TAVAKOLI, N.; NAMIN, A. S. A comparison of arima and lstm in forecasting time series. In: IEEE. *2018 17th IEEE international conference on machine learning and applications (ICMLA)*. [S.l.], 2018. p. 1394–1401.
- SILBERSCHNEIDER, W. Some notes about the problem of the icms evasion in brazil. *George Washington University. Minerva Program*, 2000.
- SILVA, P.; FIGUEIREDO, K. Aprendizado profundo aplicado na previsão de receita tributária utilizando variáveis endógenas. In: *Anais do XVII Encontro Nacional de Inteligência Artificial e Computacional*. Porto Alegre, RS, Brasil: SBC, 2020. p. 414–425. ISSN 2763-9061. Disponível em: <<https://sol.sbc.org.br/index.php/eniac/article/view/12147>>.
- SILVA, R. M. d. *Relações no mercado internacional de soja em grão: Preços, volatilidades e fluxo de informações*. Tese (Doutorado) — Universidade de São Paulo, 2013.
- SISNANDO, S. R. A.; FREITAS, M. A. de S. Previsão e avaliação do desempenho dos contribuintes do icms do estado do ceará utilizando as redes neurais artificiais. *Revista econômica do Nordeste*, v. 37, n. 1, p. 131–149, 2006.
- SMITH, T. G. et al. *pmdarima: ARIMA estimators for Python*. 2017–. [Online; accessed 17-Mai-2023]. Disponível em: <<http://www.alkaline-ml.com/pmdarima>>.
- STELZER, R. On markov-switching arma processes—stationarity, existence of moments, and geometric ergodicity. *Econometric Theory*, Cambridge University Press, v. 25, n. 1, p. 43–62, 2009.
- TRIBUTÁRIA, C. d. A. *Relatório da Receita Tributária do Estado de São Paulo - janeiro-2023*. Secretaria da Fazenda e Planejamento do Estado de São Paulo, 2023. Disponível em: <[https://portal.fazenda.sp.gov.br/acessoinformacao/\\_layouts/download.aspx?SourceUrl=/acessoinformacao/Downloads/Relatorios-da-Receita-Tributaria/2023/janeiro/INTERNET\\_janeiro23.pdf](https://portal.fazenda.sp.gov.br/acessoinformacao/_layouts/download.aspx?SourceUrl=/acessoinformacao/Downloads/Relatorios-da-Receita-Tributaria/2023/janeiro/INTERNET_janeiro23.pdf)>. Acesso em: 22-Mai-2023.
- VASCO, C. G.; LÓPEZ, C. P. Determining macroeconomic indicators to implement a short-term forecasting model for vat revenue. *Papeles de trabajo del Instituto de Estudios Fiscales. Serie economía*, Instituto de Estudios Fiscales, n. 1, p. 4–22, 2018.
- VIDAL, A.; KRISTJANPOLLER, W. Gold volatility prediction using a cnn-lstm approach. *Expert Systems with Applications*, Elsevier, v. 157, p. 113481, 2020.
- WANG, Y.; ZHU, S.; LI, C. Research on multistep time series prediction based on lstm. In: IEEE. *2019 3rd International Conference on Electronic Information Technology and Computer Engineering (EITCE)*. [S.l.], 2019. p. 1155–1159.

WASKOM, M. L. seaborn: statistical data visualization. *Journal of Open Source Software*, The Open Journal, v. 6, n. 60, p. 3021, 2021. Disponível em: <<https://doi.org/10.21105/joss.03021>>.

YAN, H.; OUYANG, H. Financial time series prediction based on deep learning. *Wireless Personal Communications*, Springer, v. 102, n. 2, p. 683–700, 2018.

YERMAL, L.; BALASUBRAMANIAN, P. Application of auto arima model for forecasting returns on minute wise amalgamated data in nse. In: *2017 IEEE International Conference on Computational Intelligence and Computing Research (ICCIC)*. [S.l.: s.n.], 2017. p. 1–5.