

UNIVERSIDADE FEDERAL DO ABC
PROGRAMA DE PÓS-GRADUAÇÃO EM ECONOMIA

THIAGO DO CARMO NUNES

**COMPARAÇÃO DA PERFORMANCE PREDITIVA DE
MODELOS ARIMA, VAR E VEC EM SÉRIES TEMPORAIS
COINTEGRADAS**

São Bernardo do Campo, SP

2023

THIAGO DO CARMO NUNES

**COMPARAÇÃO DA PERFORMANCE PREDITIVA DE MODELOS
ARIMA, VAR E VEC EM SÉRIES TEMPORAIS COINTEGRADAS**

Dissertação apresentada ao Programa de
Pós-Graduação em Economia da
Universidade Federal do ABC, como
requisito à obtenção do título de Mestre em
Economia.

Área de Concentração: Macroeconomia e
Finanças

Orientador: Prof. Dr. Guilherme de Oliveira
Lima Cagliari Marques

São Bernardo do Campo, SP

2023

Sistema de Bibliotecas da Universidade Federal do ABC
Elaborada pelo Sistema de Geração de Ficha Catalográfica da UFABC com os dados
fornecidos pelo(a) autor(a).

Nunes, Thiago do Carmo

COMPARAÇÃO DA PERFORMANCE PREDITIVA DE MODELOS ARIMA,
VAR E VEC EM SÉRIES TEMPORAIS COINTEGRADAS / Thiago do Carmo
Nunes. — 2023.

46 fls.

Orientador: Guilherme de Oliveira Lima Cagliari Marques

Dissertação (Mestrado) — Universidade Federal do ABC, Programa de
Pós-Graduação em Economia, São Bernardo do Campo, 2023.

1. Econometria. 2. Séries Temporais. 3. Finanças. 4. Modelos Preditivos.
5. Performance. I. Marques, Guilherme de Oliveira Lima Cagliari. II.
Programa de Pós-Graduação em Economia, 2023. III. Título.

Esse exemplar foi revisado e alterado em relação a versão original de acordo com as observações levantadas pela banca examinadora no dia da defesa, sob responsabilidade única do(a) autor(a) e com a anuência do(a) (co)orientador(a).

São Bernardo do Campo, 17 de Outubro de 2023

Thiago do Carmo Nunes

Nome Completo e Assinatura do(a) autor(a)

Guilherme de Oliveira Lima Cagliari Marques

Nome Completo e Assinatura do(a) orientador(a)



MINISTÉRIO DA EDUCAÇÃO

Fundação Universidade Federal do ABC

Avenida dos Estados, 5001 – Bairro Santa Terezinha – Santo André – SP
CEP 09210-580 · Fone: (11) 4996-0017

FOLHA DE ASSINATURAS

Assinaturas dos membros da Banca Examinadora que avaliou e aprovou a Defesa de Dissertação de Mestrado do candidato, THIAGO DO CARMO NUNES realizada em 22 de Julho de 2022:



Documento assinado digitalmente

RICARDO BUSCARIOLLI PEREIRA

Data: 18/10/2023 23:09:51-0300

Verifique em <https://validar.iti.gov.br>

Prof.(a) RICARDO BUSCARIOLLI PEREIRA

UNIVERSIDADE FEDERAL DO ABC

Documento assinado digitalmente



SERGIO NARUHIKO SAKURAI

Data: 22/07/2022 19:57:09-0300

Verifique em <https://verificador.iti.br>

Prof.(a) SERGIO NARUHIKO SAKURAI

Prof.(a) BRUNO DE PAULA ROCHA

UNIVERSIDADE FEDERAL DO ABC

Prof.(a) CHANG CHIANN

UNIVERSIDADE DE SÃO PAULO



Documento assinado digitalmente

GUILHERME DE OLIVEIRA LIMA CAGLIARI MARQUES

Data: 22/07/2022 12:06:48-0300

Verifique em <https://verificador.iti.br>

Prof.(a) GUILHERME DE OLIVEIRA LIMA CAGLIARI MARQUES

UNIVERSIDADE FEDERAL DO ABC - Presidente

* Por ausência do membro titular, foi substituído pelo membro suplente descrito acima: nome completo, instituição e assinatura



Universidade Federal do ABC

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Código de Financiamento 001

AGRADECIMENTOS

Dedico esse trabalho a minha mãe Aparecida, que em todos os momentos me apoiou e me deu forças para iniciar e concluir esse trabalho.

Ao meu pai Geraldo, que sempre foi meu exemplo de vida.

Aos meus amigos de trabalho, que desde o início do projeto me deram forças para concluir.

Ao meu orientador Guilherme, por toda sabedoria transmitida e por estar sempre presente.

Ao professor Ramon Garcia, que desde o início dessa jornada foi um grande facilitador e coordenador do mestrado em economia da UFABC.

RESUMO

O objetivo desta dissertação é comparar a capacidade preditiva dos modelos univariados ARIMA com os modelos multivariados VAR e VEC sob a hipótese de cointegração.

A análise foi baseada em simulações de Monte Carlo, utilizando métricas estatísticas para avaliar a performance preditiva dos modelos. Para fornecer suporte empírico às conclusões obtidas via simulações, foi realizado um experimento de modelagem com séries temporais reais de cambio (USD/BRL) e inflação acumulada em 12 meses (IPCA).

As simulações mostraram que a capacidade preditiva dos modelos em dados fora da amostra é, em média, equivalente aos valores dos intervalos de confiança a 95%. No experimento com dados reais, a capacidade preditiva dos modelos foi similar, considerando os intervalos de confiança. Portanto houve um alinhamento entre as conclusões simuladas e com dados reais. Os resultados também se mostraram alinhados com as práticas de mercado no qual, mesmo sabendo a especificação teórica mais adequada ao conjunto de dados no processo de modelagem, deve-se atentar para experimentos de performance fora da amostra para maior assertividade nas previsões.

Na comparação entre modelos univariados contra dois modelos multivariados em dados cointegrados, era esperado que os modelos VEC performassem melhor em relação ao VAR e o ARIMA, fato esse não observado. A amplitude dos intervalos de confiança aponta para a necessidade de se realizar validações mais complexas, previamente à escolha da estrutura de modelagem para predição.

Concluiu-se que, mesmo em um conjunto de séries cointegradas, onde pela literatura econométrica existe uma orientação para modelagem via VEC, para maior assertividade das previsões é necessário realizar experimentos de validação fora da amostra afim de escolher o modelo com maior capacidade preditiva.

Palavras-Chave: Modelos Preditivos; Cointegração; ARIMA; VAR; VEC

ABSTRACT

This work aims to compare the predictive power of the univariate ARIMA models with the multivariate VAR and VEC models under the cointegration hypothesis.

The analysis was conducted on Monte Carlo simulations using statistical methods to evaluate the predictive power. A model was estimated using exchange rate (USD/BRL) and 12-month cumulative inflation (IPCA) to provide empirical support to check the Monte Carlo simulation insights.

The model's predictive power on out-of-sample data is similar to Monte Carlo simulations' 95% confidence intervals. The conclusion of the exchange rate and inflation data experiment aligns with simulations. The predictive power of the models was similar within the confidence intervals of the mean. However, the size of the confidence intervals demands more complex validations before choosing the modeling specification for prediction.

When comparing univariate models against two multivariate models in cointegrated data, it was expected that the VEC models would perform better than VAR and ARIMA, which should have been observed. The results are also in line with market best practices, where even knowing the appropriate theoretical model specification, we should pay attention to out-of-sample performance experiments for greater assertiveness in predictions.

The work concluded that even in a set of cointegrated series, where there is a guide for modeling VEC in the econometric literature, to get more assertiveness of predictions, it is necessary to do out-of-sample experiments outside the sample to choose the model with greater predictive power.

Keywords: Forecast Models; Cointegration; ARIMA; VAR; VEC

LISTA DE FIGURAS

<i>Figura 1 - Estratégia de Validação Fora da Amostra Validation</i>	12
<i>Figura 2 - Fluxograma do Processo de Simulação</i>	15
<i>Figura 3 - Performance média variando o número de observações (MAE)</i>	16
<i>Figura 4 - Performance média variando o número de observações (MSE)</i>	17
<i>Figura 5 - Performance média variando o tamanho da amostra de validação (MAE)</i>	17
<i>Figura 6 - Performance média variando o número de variáveis do sistema cointegrado (MAE)</i>	18
<i>Figura 7 - Performance média aplicando uma média móvel 12 na série temporal (MAE)</i>	19
<i>Figura 8 - Acurácia global dos Modelos, Diebold Mariano</i>	20
<i>Figura 9 - Acurácia dos Modelos VAR vs ARIMA , Diebold Mariano</i>	20
<i>Figura 10 - Acurácia dos Modelos VEC vs ARIMA, Diebold Mariano</i>	21
<i>Figura 11 - Acurácia dos Modelos VEC vs VAR , Diebold Mariano</i>	21
<i>Figura 12 – Curva de Distribuição acumulada de resíduos MAE para conjuntos de dados com 500 observações e 3 séries cointegradas</i>	22
<i>Figura 13 - Curva de Distribuição acumulada de resíduos MAE para conjuntos de dados com 250 observações e 3 séries cointegradas</i>	23
<i>Figura 14 - USDBRL vs IPCA Séries Temporais em 2 eixos verticais</i>	24
<i>Figura 15 - USDBRL vs IPCA Séries Temporais Segmentados por Período de Treino e Teste</i>	25
<i>Figura 16 - ΔUSDBRL vs ΔIPCA Séries Temporais Segmentados por Período de Treino e Teste</i>	26
<i>Figura 17 - Performance out-of-sample por Modelo (MAE) USDBRL ~ IPCA</i> 27	
<i>Figura 18 - Performance out-of-sample por Modelo (MSE) USDBRL ~ IPCA</i> 28	

LISTA DE TABELAS

<i>Tabela 1 - GRID de parâmetros da simulação</i>	<i>10</i>
<i>Tabela 2 - Parâmetros fixos da simulação.....</i>	<i>11</i>
<i>Tabela 3 - Parâmetros para modelagem</i>	<i>24</i>
<i>Tabela 4 – ADF test.....</i>	<i>25</i>
<i>Tabela 5 – Johansen-Procedure.....</i>	<i>26</i>
<i>Tabela 6 – Modelos Escolhidos</i>	<i>27</i>

SUMÁRIO

1. Introdução	1
2. Revisão Bibliográfica	2
3. Metodologia	5
3.1. Linha do Tempo dos Modelos	5
3.2. Descrição dos Modelos	7
3.3. Metodologia de Simulação	10
3.4. Validação dos Modelos	12
4. Resultados	15
4.1. Análise dos Resíduos (MAE e MSE)	15
4.2. Teste de Diebold Mariano	19
4.3. Teste de KS e ECDF's	22
5. Estudo de Caso Real	23
6. Conclusões	29
7. Referências	31
8. Repositório	34

1. Introdução

Os modelos estatísticos de previsão são ferramentas usadas por cientistas de diversas áreas para antecipar tendências sobre o comportamento futuro de variáveis como: (i) demanda de atendimentos de emergência (Jones et al., 2002), (ii) consumo de combustível de carros em ambientes urbanos (Afana et al., 2018) e (iii) previsões climáticas (Hargreaves et al., 2004). No domínio da economia, os modelos de previsão são utilizados tanto por instituições do setor público quanto do setor privado. Entidades como os bancos centrais, por exemplo, utilizam modelos de previsão para antecipar tomadas de decisão acerca da política monetária (Nyoni, 2018). No setor privado, a capacidade de prever um padrão de comportamento futuro traz vantagens competitivas para as empresas, como por exemplo o aumento da eficiência dos investimentos (Chen et al., 2017).

Apesar dos protocolos econométricos para estimação de modelos envolvendo dados cointegrados estarem claros na literatura econométrica, em áreas de Finanças (Abdullah & Dwivedy, 2012) e de Riscos do mercado financeiro (Araz et al., 2020) é observado o amplo uso de modelos preditivos para resolver problemas, onde a escolha da sua estrutura é feita somente comparando a capacidade preditiva de uma pré-seleção de modelos em períodos fora da amostra. Um dos riscos desta estratégia é que mesmo com modelos mais modernos pode-se observar falhas endógenas e exógenas que demandam a atenção dos cientistas. (Ioannidis et al., 2020) explora aprendizados de competições de modelagem que são úteis para todas as áreas onde são utilizados modelos preditivos sem uma alta maturidade e amplo conhecimento “de negócio”, e a seleção dos modelos se baseiam somente em performance fora da amostra.

Este trabalho visa comparar a capacidade preditiva dos modelos clássicos Modelo Auto-Regressivo Integrado e de Médias Móveis (ARIMA), Modelos Vetoriais Autorregressivos (VAR) e Modelo Vetorial de Correção de Erros (VEC) em conjuntos de dados cointegrados e validar as performances destes modelos em dados fora da amostra, uma vez que os protocolos econométricos

envolvendo dados cointegrados possuem indicações prévias de qual modelo deve ser utilizado.

Com este intuito, foram analisados os ganhos de performance em relação à capacidade preditiva com séries de dados simulados. Foram utilizados procedimentos de simulação de Monte Carlo para gerar os conjuntos de séries temporais cointegradas baseados em coeficientes lineares e termos de perturbação aleatórios. As simulações mostraram que a capacidade preditiva dos modelos em dados fora da amostra é, em média, equivalente aos valores dos intervalos de confiança a 95%.

Para dar maior suporte às conclusões obtidas via simulações, foi também realizado um experimento de modelagem com séries temporais reais de cambio (USD/BRL) e inflação acumulada em 12 meses (IPCA). No experimento com dados reais, a capacidade preditiva dos modelos foi similar, considerando os intervalos de confiança. Portanto houve um alinhamento entre as conclusões simuladas e com dados reais.

Os resultados se mostraram alinhados com as práticas de mercado no qual, mesmo sabendo a especificação teórica mais adequada ao conjunto de dados no processo de modelagem, deve-se atentar para experimentos de performance fora da amostra para maior assertividade nas predições.

A seguir, é apresentado na seção 2 a realização da revisão bibliográfica; na seção 3, a metodologia abordando uma linha do tempo dos modelos mais utilizados, o design destes modelos e estratégia de simulação e validação; na seção 4, os resultados das simulações e na seção 5 são descritas as conclusões.

2. Revisão Bibliográfica

Em (Fanchon & Wendel, 1992) foram comparados os modelos VAR e VEC em séries de dados cointegrados para modelar preço de gado. A conclusão é que o modelo VAR gerou erros quadráticos médios (MSE) ligeiramente menores do que o VEC. A comparação entre VAR e VEC depende das defasagens

incluídas no modelo de previsão, portanto dependendo dos parâmetros escolhidos, um modelo performa melhor em relação ao outro.

Em (Bachmeier & Swanson , 2005) foram avaliados modelos ARIMA, VAR e VEC utilizando a técnica de validação fora da amostra para testar a capacidade preditiva dos modelos em dados da inflação norte-americana de 1959 até 1999. Foram também utilizadas simulações de Monte Carlo para gerar estatísticas sobre os modelos. Os autores concluíram que quando os parâmetros do vetor de cointegração utilizado são estimados, o modelo VEC sempre apresenta uma capacidade preditiva inferior aos do modelo VAR em primeiras diferenças. Isso sugere que trabalhos anteriores que indicaram que os modelos VEC não têm capacidade de previsão melhor que os modelos VAR pode ser, em parte, devido à presença de erro de estimativa dos parâmetros do vetor de cointegração.

Esse resultado sustenta a hipótese de que, ao encontrar modelos VEC com performance inferior aos modelos VAR, pode estar associado ao erro de estimação do vetor de cointegração e não necessariamente à estrutura de modelagem como um todo.

No artigo (Ng, Fan & Wong ,2011) foram utilizados modelos VEC para confirmar a influência do capital público no capital privado de construção civil através dos efeitos de gastos públicos sobre o investimento privado (efeito *crowding-in*). O Produto Interno Bruto (PIB) real, investimento do setor público em construção civil e o desemprego são variáveis que influenciam o investimento privado em construção em Hong Kong. O estudo compara regressões lineares, em uma estrutura de defasagens muito próxima ao modelo VAR, com o modelo VEC e apresenta através da comparação de performance fora da amostra, que o modelo com vetor de correção de erros é a estrutura mais adequada para projetar séries reais não estacionárias e variáveis cointegradas.

Em (Anggraeni, Boga e Mahananto , 2018) foram comparadas as capacidades preditivas de modelos ARIMA e estruturas VAR para prever preços de arroz em conjuntos de dados fora da amostra, utilizando indicadores de performance MAPE (Erro Percentual Médio Absoluto). No estudo em questão os modelos ARIMA estimados obtiveram uma performance 15% melhor do que os modelos VAR.

Em (Tehreem et al., 2020) foram estudados o impactos da temperatura média, demanda de energia, valor agregado setorial e crescimento populacional na qualidade dos recursos hídricos e na taxa de mortalidade no Paquistão. Os autores utilizaram a estatística do traço proposta por (Johansen, 1988) para identificação de relações de cointegração entre demanda de energia, valor agregado da indústria e a qualidade dos recursos hídricos. Para detectar cointegração em séries temporais nas simulações deste trabalho, serão utilizadas as estatísticas dos autovalores e autovetores também propostas por (Johansen, 1988).

A versatilidade de modelos autorregressivos é comprovada pela ampla variedade de temas em que eles são aplicados, e oferecem soluções satisfatórias. Foram citados trabalhos no campo da Economia, porém em outras áreas, como na Saúde, tem-se observado um amplo uso dessa classe de modelos. O ano de 2020 foi de alta demanda para a área de modelos de previsão. A pandemia do vírus COVID-19 exigiu um uso massivo de modelos de previsão nas áreas de saúde, finanças e setores industriais. Em (Hu et al. , 2020) apresenta-se métodos de inteligência artificial para prever a duração da pandemia de Covid-19 na China utilizando a validação fora da amostra como método de seleção de modelos. Em (Salisu et al. , 2020) propõe-se em modelos clássicos, como os modelos VAR, para prever o comportamento dos estoques de petróleo durante a pandemia de COVID-19.

Além da pesquisa de vacinas e medicamentos para mitigar doenças como a Covid-19, governos e autoridades centrais utilizaram de modelos autorregressivos para planejarem seus sistemas de saúde para suportar surtos de outras doenças, como por exemplo o estudo de (Wangdi et al., 2010), onde é apresentado uma seleção de modelos ARIMA para previsão, buscando avaliar a incidência da malária nos distritos endêmicos.

Além da academia, outros setores da economia utilizam sua infraestrutura de pesquisa e, até mesmo, patrocinam competições de modelos de previsão para entender padrões nos dados relacionados a COVID-19, utilizando como estratégia e critério de performance a validação fora da amostra. Em (Ioannidis et al., 2020) destaca que, mesmo com modelos mais modernos para prever

dados relacionados às pandemias, a tarefa continua sendo árdua, visto que existem falhas endógenas e exógenas que precisam de atenção dos cientistas. O trabalho elencou que dentre as razões mais comuns de falha dos modelos de previsão, estão: (i) insuficiência das “covariáveis” (*features*) e a falta de informações claras sobre o problema; (ii) premissas erradas no modelo, levantando a necessidade de uso de modelos probabilísticos em apoio às séries de tempo; (iii) modelos utilizados com base em dados de epidemias anteriores, sendo que a disponibilidade e qualidade da informação estavam em patamares muito piores das atuais; (iv) falta de expertise em Epidemiologia nas empresas e por cientistas que desenvolveram os modelos. Os aprendizados do artigo são úteis para qualquer área onde são utilizados modelos preditivos sem uma alta maturidade e amplo conhecimento “de negócio”, onde a seleção dos modelos se baseiam somente em performance fora da amostra.

A partir da revisão bibliográfica, percebeu-se que a validação fora da amostra é a principal técnica para se validar modelos de previsão. Isto é, para testar sua acurácia preditiva. Dada a divergência de conclusão entre os trabalhos, não se tem evidências suficientes para pressupor que nenhum dos modelos, ARIMA, VAR ou VEC, possuam genericamente uma capacidade preditiva superior aos demais sem que o contexto seja levado em consideração. O contexto de negócio e as condições de seleção dos parâmetros de cada estrutura parecem ser determinantes para escolha do melhor modelo preditivo.

3. Metodologia

A sessão de metodologia está organizada para: (i) apresentar uma linha do tempo evidenciando a evolução dos modelos escolhidos como alvo desse trabalho, (ii) o design e as especificações mais técnicas dos modelos, (iii) a metodologia de simulação e (iv) as estratégias de validação dos modelos.

3.1. Linha do Tempo dos Modelos

Um dos problemas básicos da econometria aplicada às séries temporais é formular modelos de previsão. Em (Box & Jenkins, 1970) popularizaram a versão mais conhecida dos modelos de previsão baseada em dados univariados, o Modelo Auto-Regressivo Integrado e de Médias Móveis (ARIMA). Sua proposta é modelar parametricamente a média condicional de uma série temporal baseando-se em seu comportamento passado. Os autores propuseram um algoritmo de fácil implementação contendo basicamente quatro etapas: identificação, estimação, verificação e previsão.

Apesar da análise univariada de séries temporais ser amplamente utilizada na econometria aplicada, as variáveis macroeconômicas tipicamente apresentam comportamentos dinâmicos e interdependentes com outras variáveis. Este comportamento dinâmico é observado em indicadores de atividade econômica, inflação e instrumentos de política monetária aplicados em curvas de retorno de investimentos (Diebold et al., 2006). Para resolver um sistema de variáveis interdependentes, (Sims, 1980) desenvolveu os modelos Vetoriais Autorregressivos (VAR). Em (Lütkepohl, 1991) explica-se que a estabilidade dos modelos VAR pressupõe a condição de estacionariedade entre as variáveis integrantes do sistema. A estrutura VAR contribuiu no desenvolvimento e representação matemática de conceitos relacionados às séries temporais macroeconômicas estacionárias.

No entanto, para séries econômicas não estacionárias devido a presença de tendências estocásticas, (Engle & Granger, 1987) mostraram a necessidade do sistema satisfazer uma condição estatística denominada por eles como “cointegração”. Informalmente, séries temporais que apresentam uma relação estrutural entre si compartilhando uma mesma tendência estocástica por meio de uma combinação linear, são ditas cointegradas. Segundo (Enders, 2014) a cointegração possui interpretação econômica de “uma relação de longo prazo entre as variáveis”. (Engle & Granger, 1987) também demonstraram que, se um conjunto de variáveis for integrado de primeira ordem e cointegrado, é possível representá-lo sob a forma de um Modelo Vetorial de Correção de Erros (VEC). A principal diferença do modelo VEC em relação ao seu antecessor VAR é a presença de um termo de ajuste de longo prazo, também conhecido como termo

de correção de erros. (Johansen, 1988) demonstrou uma forma elegante de testar se séries temporais são cointegradas, a partir do modelo VEC.

Um fato estilizado em Finanças é que existem séries temporais onde a causalidade não ocorre de forma bidirecional. Quando os dados compartilharem uma mesma tendência estocástica, existe exogeneidade em determinado sentido entre as variáveis (Rogalski & Vinso, 1977). Por exemplo, como nos pressupostos oriundos da Teoria Econômica, é possível imaginar que uma série temporal referente aos dados de inadimplência de uma pequena carteira de crédito de um banco e a taxa de desemprego do país possam ser correlacionadas. Nesse sistema, pressupõe-se que, apesar de ser possível um compartilhamento de tendência entre as séries, a relação deve ser unidirecional, ou seja, somente a taxa de desemprego seria capaz de influenciar a trajetória da inadimplência e não o contrário.

3.2. Descrição dos Modelos

Esta dedica-se a descrever resumidamente os modelos ARIMA, VAR e VEC, uma vez que esses modelos são bastante abordados pela literatura econométrica. Para um maior aprofundamento nestes modelos, serão indicadas algumas referências bibliográficas renomadas.

Os modelos ARIMA são processos estocásticos constituídos para representar a média condicional de uma série temporal. Esta classe de modelos é utilizada para descrever a trajetória de séries temporais a partir de suas próprias observações e de estimativas de choques em períodos passados. Segundo (Brockwell & Davis, 1991), a estrutura do modelo é composta por dois componentes: um autorregressivo e um de médias móveis. Uma estrutura ARIMA é representada pela Equação 1:

$$\Delta^d Y_t = C_0 + \sum_{i=1}^p \phi_i * \Delta^d Y_{t-i} + \sum_{j=1}^q \theta_j \epsilon_{t-j} + \epsilon_t, \quad (1)$$

sendo: p a ordem autorregressiva; d a ordem de diferenciação; q a ordem das médias móveis; Δ operador diferença $\Delta^d = (1 - B)^d Y_t$, com $B^j Y_t = Y_{t-j}$, é

responsável por estacionarizar a série Y_t ; C_0 o intercepto; ϕ o coeficiente linear dos termos autorregressivos; Θ coeficiente linear dos termos de médias móveis; ϵ_t é um processo ruído branco. Em econometria o método de estimação mais comum para esta classe de modelos é a Máxima Verossimilhança.

De acordo com (Morettin & Toloi, 2004), modelos ARIMA possuem duas condições importantes de serem estudados: estacionariedade e invertibilidade. A condição básica de estacionariedade em uma estrutura autorregressiva é que as raízes do processo gerador da série devem ser, em módulo, maiores do que 1, ou seja, fora do círculo unitário. A invertibilidade diz sobre a característica de se transformar um modelo AR em MA sob algumas condições.

Os modelos VAR, introduzidos por (Sims, 1980), são estruturas multivariadas representadas por um sistema de equações simultâneas que capturam a existência de relações de interdependência entre variáveis. A estrutura do VAR correlaciona observações atuais de uma variável com suas observações passadas e as observações passadas de outras variáveis do sistema. A forma mais simples do VAR é expressa em sua estrutura bivariada de grau um conforme representado pela Equação 3:

$$VAR(1) : \begin{pmatrix} Y_t \\ X_t \end{pmatrix} = \begin{pmatrix} C_0 \\ C_1 \end{pmatrix} + \begin{pmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{pmatrix} * \begin{pmatrix} Y_{t-1} \\ X_{t-1} \end{pmatrix} + \begin{pmatrix} \epsilon_{y,t} \\ \epsilon_{x,t} \end{pmatrix}, \quad (2)$$

$$VAR(1) : \begin{cases} Y_t = C_0 + \phi_{11} * Y_{t-1} + \phi_{12} * X_{t-1} + \epsilon_{y,t} \\ X_t = C_0 + \phi_{21} * Y_{t-1} + \phi_{22} * X_{t-1} + \epsilon_{x,t} \end{cases}, \quad (3)$$

sendo: C o intercepto; ϕ o coeficiente linear dos autorregressivos; ϵ resíduo ou um processo ruído branco.

(Lütkepohl, 1991) explica que a estabilidade dos modelos VAR pressupõe a condição de estacionariedade entre as variáveis integrantes do sistema. Portanto, para um sistema de variáveis não cointegradas e integradas de grau um, deve-se modelar o VAR em primeiras diferenças, conforme representado pela Equação 5:

$$\Delta VAR(1) : \begin{pmatrix} \Delta^1 Y_t \\ \Delta^1 X_t \end{pmatrix} = \begin{pmatrix} C_0 \\ C_1 \end{pmatrix} + \begin{pmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{pmatrix} * \begin{pmatrix} \Delta^1 Y_{t-1} \\ \Delta^1 X_{t-1} \end{pmatrix} + \begin{pmatrix} \epsilon_{y,t} \\ \epsilon_{x,t} \end{pmatrix}, \quad (4)$$

$$\Delta VAR(1) : \begin{cases} \Delta^1 Y_t = C_0 + \phi_{11} * \Delta^1 Y_{t-1} + \phi_{12} * \Delta^1 X_{t-1} + \epsilon_{y,t} \\ \Delta^1 X_t = C_0 + \phi_{21} * \Delta^1 Y_{t-1} + \phi_{22} * \Delta^1 X_{t-1} + \epsilon_{x,t} \end{cases}, \quad (5)$$

Modelos VAR também são comumente estimados utilizando o método da Máxima Verossimilhança. Para os conjuntos de variáveis cointegrados, a abordagem econométrica mais adequada é representá-las em um modelo VEC. Segundo (Enders, 2014), incluir o termo de correção de erros no modelo VAR permite equilibrar o ajuste dinâmico de curto prazo e longo prazo entre as variáveis do sistema. Conforme (Engle & Granger, 1987), o termo de correção de erros entre duas variáveis Y_t e X_t se apresenta na forma da Equação 6 e sua premissa básica é que o termo seja estacionário.

$$\hat{\tau}_y * Y_{t-1} + \hat{\tau}_x * X_{t-1} = I(0), \quad (6)$$

Sendo: $\hat{\tau}_n$ os coeficientes lineares do termo de correção de erros; Y_{t-1} e X_{t-1} variáveis aleatórias cointegradas; $I(0)$ uma série temporal não integrada, ou seja, estacionária.

Portanto o modelo VEC bivariado de ordem um pode ser descrito a partir da representação da Equação 8:

$$VEC(1) : \begin{pmatrix} \Delta Y_t \\ \Delta X_t \end{pmatrix} = \begin{pmatrix} C_0 \\ C_1 \end{pmatrix} + \begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix} * \begin{pmatrix} Y_{t-1} \\ X_{t-1} \end{pmatrix} + \begin{pmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{pmatrix} * \begin{pmatrix} \Delta Y_{t-1} \\ \Delta X_{t-1} \end{pmatrix} + \begin{pmatrix} \epsilon_{y,t} \\ \epsilon_{x,t} \end{pmatrix}, \quad (7)$$

$$VEC(1) : \begin{cases} \Delta Y_t = C_0 + (\lambda_1 * Y_{t-1} + \lambda_2 * X_{t-1}) + \phi_{11} * \Delta Y_{t-1} + \phi_{12} * \Delta X_{t-1} + \epsilon_{y,t} \\ \Delta X_t = C_0 + (\lambda_1 * Y_{t-1} + \lambda_2 * X_{t-1}) + \phi_{21} * \Delta Y_{t-1} + \phi_{22} * \Delta X_{t-1} + \epsilon_{x,t} \end{cases}, \quad (8)$$

sendo: C o intercepto; ϕ o coeficiente linear dos autorregressivos; λ o coeficiente linear do termo de longo prazo; ϵ resíduo ou um processo Ruído Branco. Modelos VEC também são estimados pelo método de Máxima Verossimilhança.

Para seleção da melhor estrutura de modelagem, foi utilizado o critério de informação bayesiano (BIC) proposto por (Schwarz, 1978), que é dado pela equação:

$$BIC = -2 \log f(x_n | \theta) + p \log n, \quad (9)$$

sendo: $f(x_n|\theta)$ é o modelo escolhido, p é o número de parâmetros a serem estimados e n é o número de observações da amostra.

A identificação da cointegração entre as séries temporais foi feita com o teste dos autovalores e autovetores de (Johansen, 1988). O teste proposto pelo autor é mais genérico e cobre situações em que testes mais simplificados não conseguem validar a presença de cointegração. Em situações nas quais tem-se mais de duas variáveis, o teste pode apresentar também mais de um vetor de cointegração, ou seja, dois processos estocásticos agindo como uma relação de longo prazo entre os dados.

3.3. Metodologia de Simulação

Todas as etapas computacionais foram feitas utilizando o software R v3.6.3 (R Core Team, 2016). Na primeira etapa do processo de simulação iniciou-se uma semente para que cada etapa do processo fosse reproduzível. Foram definidos conjuntos de parâmetros fixos e uma grade de parâmetros variáveis para a simulação. A grade de parâmetros contém cinco variáveis: (i) **n_obs**: o número de observações de cada série temporal variando de 100, 250, 500 e 1000 unidades; (ii) **n_ma**: suavização ou não das séries temporais via média móvel de 12 períodos; (iii) **n_fold_ts**: proporção da amostra de dados reservado para “treino” do modelo variando entre 85%, 90% e 95%; (iv) **n_ts**: o número de covariáveis no sistema variando de dois a até três séries temporais.

A grade de parâmetros para as simulações de Monte Carlo está descrita na Tabela 1.

Tabela 1 - GRID de parâmetros da simulação. Fonte: Próprio Autor

Parâmetros	Rótulo	Valores
Número de Observações	n_obs	[1000,500,250,100]
Suavização por média móvel	n_ma	[0,12]
Proporção da amostra reservada para treino do modelo	n_fold_ts	[85%,90%,95%]
Número de Variáveis	n_ts	[2,3]

Para cada conjunto de parâmetros da grade da Tabela 1 foram gerados 1000 conjuntos de dados para estimação e validação dos modelos. Para os testes de hipótese foi utilizado um nível de confiança de 95%. A grade de parâmetros fixos para as simulações é descrita na Tabela 2.

Tabela 2 - Parâmetros fixos da simulação. Fonte: Próprio Autor

Parâmetros	Valor
Número de iterações por conjunto de parâmetros	1000
Semente Inicial do Processo	2022
Nível de Confiança usada nos Testes de Hipótese	95%

Dado os conjuntos de parâmetros acima, para cada iteração do processo é selecionada uma combinação de variáveis e iniciada a simulação. Dada uma nova semente, o algoritmo gera o conjunto de dados cointegrado.

A simulação do sistema de dados cointegrados foi realizada seguindo o procedimento descrito na sessão 12.2.4 de modelagem de séries financeiras de (Zivot & Wang, 2006). Os sistemas cointegrados podem ser simulados utilizando a representação triangular proposta por (Phillips, 1991). Por exemplo, um sistema cointegrado bivariado $Y_t = (y_{1t}, y_{2t})$ com um vetor de cointegração $\beta = (1, -\beta_2)$ tem a sua representação seguindo a Equação 10:

Sistema Cointegrado Bivariado:

$$\begin{cases} y_{1t} = \beta_2 * y_{2t} + u_t & , \text{onde } u_t \sim I(0) \\ y_{2t} = y_{2t-1} + v_t & , \text{onde } v_t \sim I(0) \end{cases} \quad (10)$$

A primeira equação descreve a relação de equilíbrio de longo prazo com um termo de perturbação $I(0)$. A segunda equação especifica uma tendência estocástica comum e também um termo de perturbação $I(0)$. Um sistema cointegrado trivariado $Y_t = (y_{1t}, y_{2t}, y_{3t})$ com um vetor de cointegração $\beta = (1, -\beta_2, -\beta_3)$ tem a sua representação seguindo a Equação 11:

Sistema Cointegrado Trivariado:

$$\begin{cases} y_{1t} = \beta_2 * y_{2t} + \beta_3 * y_{3t} + u_t & , \text{onde } u_t \sim I(0) \\ y_{2t} = y_{2t-1} + v_t & , \text{onde } v_t \sim I(0) \\ y_{3t} = y_{3t-1} + w_t & , \text{onde } w_t \sim I(0) \end{cases} \quad (11)$$

Após gerar as séries cointegradas, foram realizados dois testes sobre as hipóteses de estacionariedade e de cointegração.

O teste de estacionariedade escolhido foi o Teste de Dickey–Fuller Aumentado (Hatanaka, 1996), mais conhecido como teste ADF. Para que o fluxo de modelagem continue, é necessário que todas as séries simuladas Y_1 , Y_2 e Y_3 sejam $I(1)$, ou seja, integradas de ordem um. Caso essa condição seja satisfeita a um nível de confiança de 95% o algoritmo continua para a próxima etapa. Caso contrário, é gerado um novo conjunto de dados cointegrados até que essa condição seja satisfeita.

A próxima etapa do processo é validar se as séries são cointegradas, utilizando as estatísticas dos autovalores e autovetores de Johansen (Søren, 1991). Se a relação de cointegração não for satisfeita a nível de confiança de 95%, é gerado um novo conjunto de dados cointegrados até que essa condição seja satisfeita. O intuito desta validação é selecionar apenas as séries que cumpririam as condições necessárias em uma análise empírica de cointegração.

3.4. Validação dos Modelos

Para validar a capacidade preditiva dos modelos, utilizou-se a metodologia de validação fora da amostra: foram separados os conjuntos de dados de treino e teste para validar os resultados preditivos dos modelos variando o tamanho do período de teste conforme Figura 1.

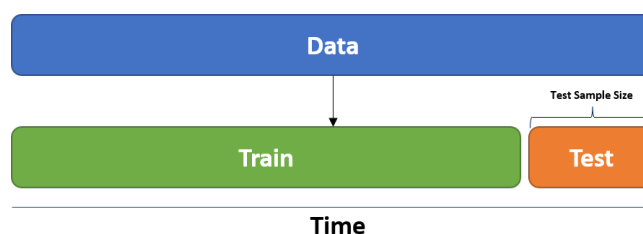


Figura 1 - Estratégia de Validação Fora da Amostra. Fonte: Próprio Autor

A validação fora da amostra é essencial para saber se o modelo proposto está realmente generalizando de forma adequada aos dados ou está somente fazendo um ajuste nos dados de treinamento sem que **este** reflita em performance nos períodos de teste. O tamanho da amostra de teste será definido como um parâmetro da simulação. Para isso, foram reservados entre 5%, 10% e 15% de todo o conjunto de dados mais recente para validar o modelo.

Após a construção das bases de treino e teste, foram estimados os três modelos: ARIMA (p,q,n), VAR (p) em primeiras diferenças e um VEC (p) , sendo (p,q,n) os parâmetros escolhidos que minimizaram o critério de informação bayesiano BIC. Gerados os objetos de modelos, foram realizadas as previsões e comparadas com os dados reservados para teste. Para mensurar a qualidade do ajuste foram utilizadas as seguintes métricas para avaliação da capacidade preditiva dos modelos:

- MAE (Mean Absolute Error);
- MSE (Mean Squared Error);
- Estatística de Diebold Mariano;
- Curvas de Distribuição Acumulada e distância de Kolmogorov-Smirnov (KS);

O MAE (Mean Absolute Error) é definido como erro absoluto médio, e pode ser calculado a partir da Equação 12.

$$MAE = \frac{\sum_{t=1}^T |\hat{Y}_t - Y_t|}{T} = \frac{\sum_{t=1}^T |\hat{\epsilon}_t|}{T} \quad (12)$$

Sendo: T o número de observações preditas; \hat{Y}_t a série de dados predita; Y_t a série de dados observada; $\hat{\epsilon}_t$ os resíduos do modelo.

O MSE (Mean Square Error) é definido como o erro quadrático médio, e pode ser calculado a partir da Equação 13

$$MSE = \frac{\sum_{t=1}^T (\hat{Y}_t - Y_t)^2}{T} = \frac{\sum_{t=1}^T (\hat{\epsilon}_t)^2}{T} \quad (13)$$

As métricas acima (MAE, MSE) não levam em consideração a direção dos erros, somente a amplitude do termo de perturbação em relação à média. O MAE é menos sensível a valores atípicos. Por outro lado, o MSE, apesar de ser mais sensível a valores extremos, pune com mais intensidade erros de maior magnitude.

Para gerar os intervalos de confiança das estatísticas estimadas, foi utilizado o método de *Bootstrap* para gerar os intervalos de confiança empíricos descritos em (Mooney & Duval, 1993). Os parâmetros de *Bootstrap* utilizados foram: 1000 iterações; amostra de 20% da população por iteração e intervalos de confiança de 95%.

A Estatística de *Diebold-Mariano* (Diebold, 2013) compara previsões de dois modelos em períodos fora da amostra para checar se a acurácia entre ambos é estatisticamente igual. O teste proposto inicialmente por Diebold & Mariano (1995), aplica uma função de custo $f(e)$ sobre os erros de previsão de dois modelos e testa se o valor esperado $E[g(t)] = 0$, sendo $g(t)$ definido como $g(t) = f(e_{modelo\ 1}) - f(e_{modelo\ 2})$. Como o teste será utilizado em uma grande massa de modelos simulados, a ideia é utilizar a estatística de *Diebold-Mariano* para identificar, dentro de uma amostra grande de previsões, qual o percentual destas possuem acurácia estatisticamente iguais.

A curva de distribuição acumulada, ou ECDF, descreve como as probabilidades são associadas a valores de uma variável aleatória. A curva descreve a distribuição da probabilidade de uma variável aleatória e é uma alternativa visual para observar as distribuições de dados entre as amostras de previsão. A partir da ECDF foi estimada a distância de KS (Daniel, Wayne W, 1990). A distância de Kolmogorov-Smirnov é usada como estatística no teste KS, que testa a probabilidade de a hipótese nula de duas amostras serem retiradas da mesma distribuição. O objetivo é entender se a distribuição de erros entre os dois modelos testados é igual.

Após a validação das projeções os resultados são armazenados em bases de dados dedicadas. Esse processo se repete até que todas as combinações de parâmetros sejam testadas. O algoritmo completo de simulação desenvolvido para esse trabalho está descrito no fluxograma da Figura 2.

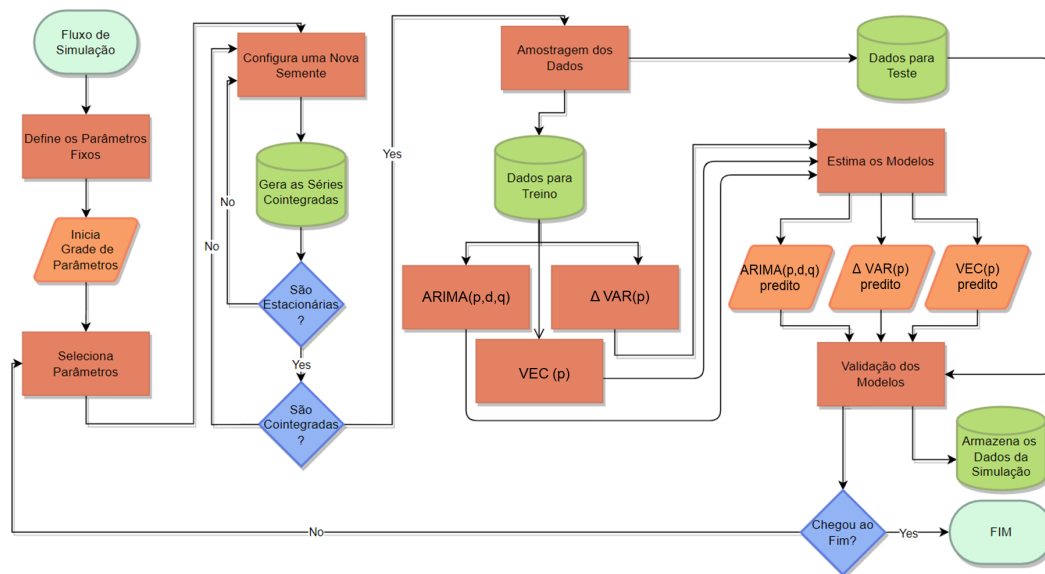


Figura 2 - Fluxograma do Processo de Simulação. Fonte: Próprio Autor

4. Resultados

4.1. Análise dos Resíduos (MAE e MSE)

Em conjuntos de dados envolvendo séries temporais com 100 observações, o modelo ARIMA possui performance similar ao modelo VAR e melhor do que o VEC. Entretanto, quando a amostra fica igual ou superior a 250 observações, a performance dos modelos é similar, conforme Figura 3. A hipótese sob a menor performance do modelo VEC pode estar ligada à presença de erro de estimativa de parâmetro de vetor de cointegração, pois conforme discutido por (Bachmeier & Swanson, 2005), quanto menor o volume de dados de treino, intuitivamente, maior seria o erro desta estimativa. Quando o algoritmo encontra um volume grande o suficiente de observações, como por exemplo 500, os modelos possuem performance similar.

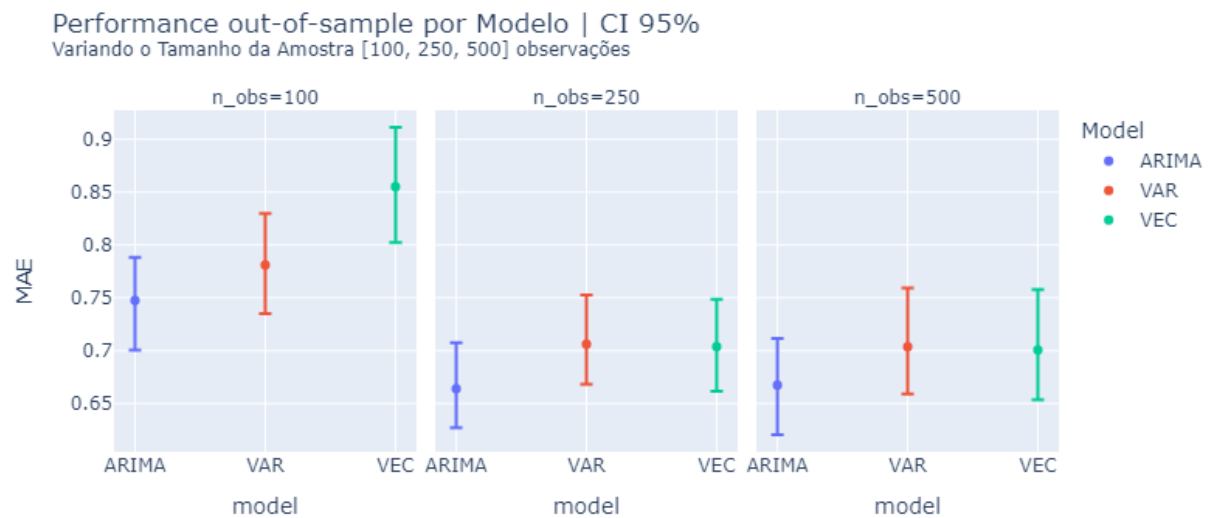


Figura 3 - Performance média variando o número de observações (MAE) . Fonte: Próprio Autor

Outro resultado sugerido pela Figura 3 é que, quanto maior o número de observações no algoritmo, menor é a média dos resíduos, ou seja, quanto maior o volume de observações utilizadas para estimação, menor é a distinção da capacidade preditiva entre modelos.

A Figura 4 mostra que as conclusões apresentadas na Figura 3 usando o MAE são as mesmas do MSE, com intervalos de confiança ligeiramente maiores e média superior. Portanto, decidiu-se não discutir os resultados segmentados pelas duas métricas e usar somente o MSE como referência, visto que todas as conclusões levaram a equivalência entre as métricas de resíduos.

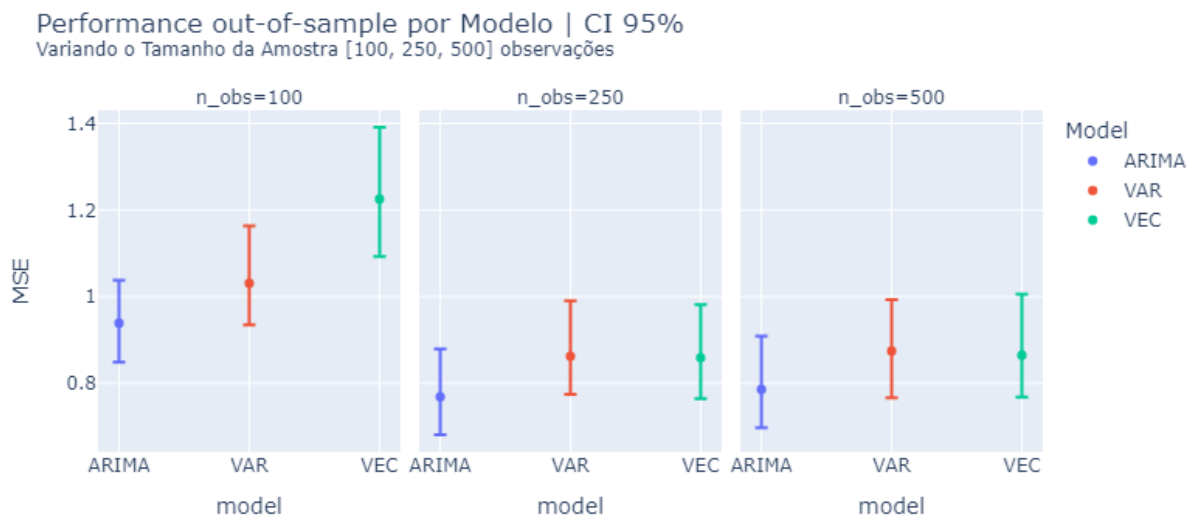


Figura 4 - Performance média variando o número de observações (MSE) . Fonte: Próprio Autor

Ao aumentar o tamanho da amostra de treinamento (de 85% para 95%), tem-se que todos os modelos convergem a uma performance equivalente dentro dos intervalos de confiança, conforme Figura 5. Mais uma vez, conforme esperado, quanto maior a base de treinamento, menor é o resíduo médio.

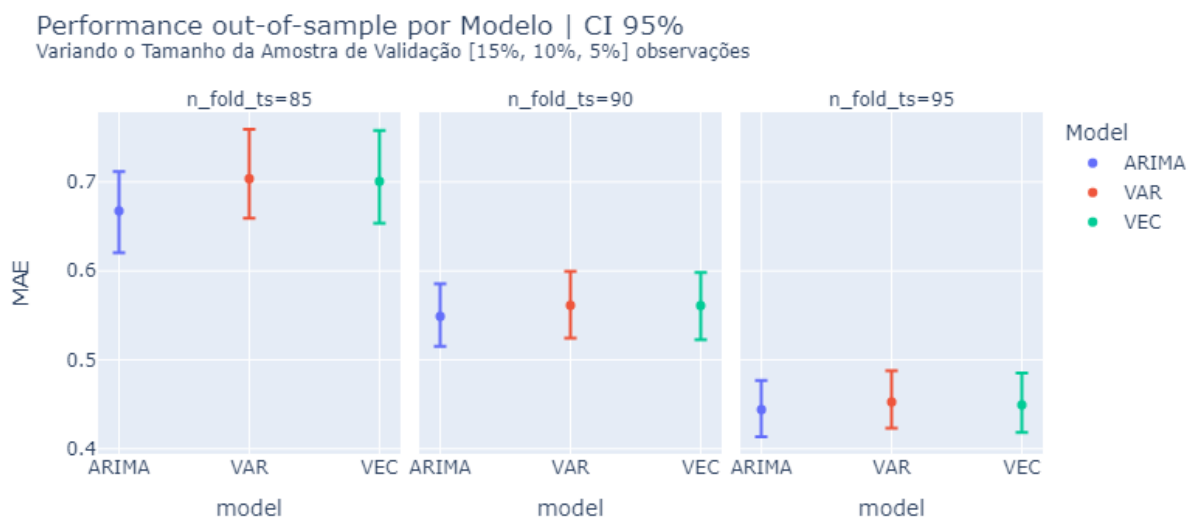


Figura 5 - Performance média variando o tamanho da amostra de validação (MAE) . Fonte: Próprio Autor

Não foram encontradas diferenças significativas na performance dos modelos ao formular o problema com duas ou três variáveis. A Figura 6 mostra que os resultados de performance obtidos nas simulações com 2 ou 3 variáveis são iguais e contidas num mesmo intervalo de confiança

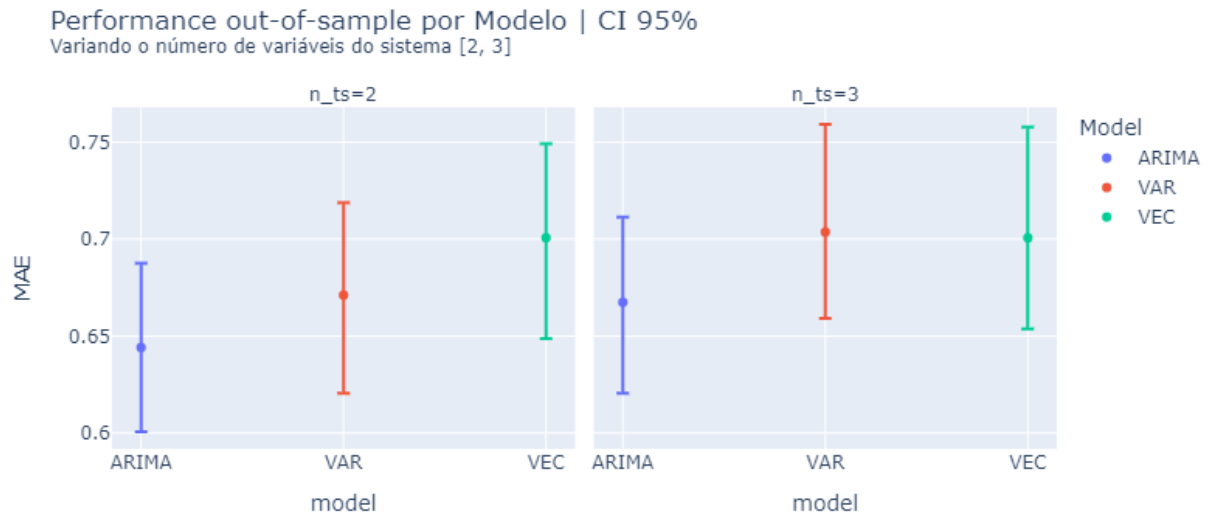


Figura 6 - Performance média variando o número de variáveis do sistema cointegrado (MAE) . Fonte: Próprio Autor

Ao aplicar uma suavização por média móvel de 12 períodos, os modelos VEC apresentam performance igual ou inferior aos modelos ARIMA e VAR conforme Figura 7. O fato observado mostra que ao utilizar uma média móvel as séries são submetidas a uma espécie de filtro que remove componentes de alta frequência e que afetam diretamente a média da performance dos modelos.

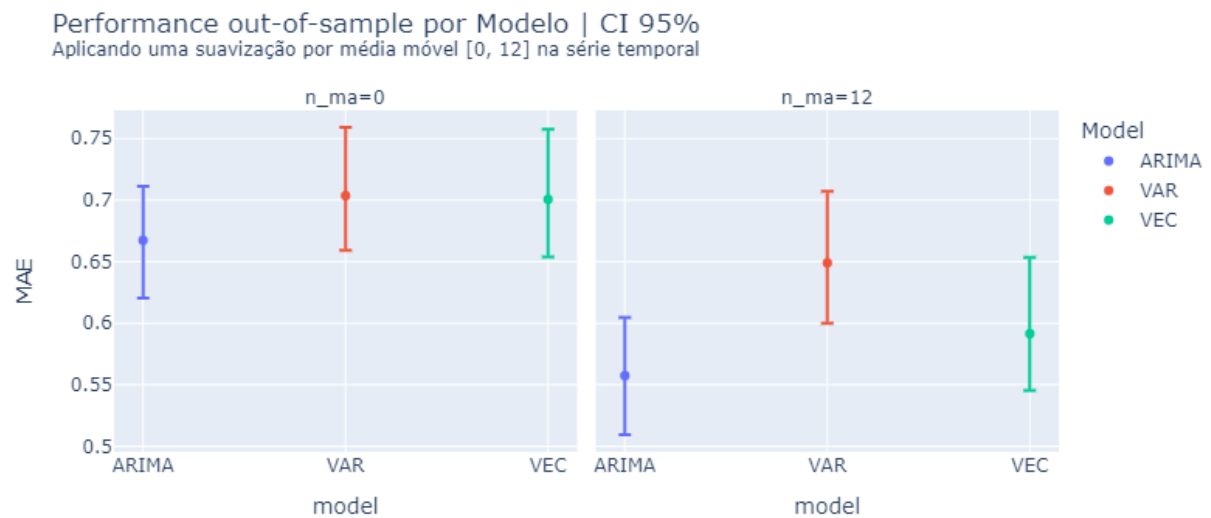


Figura 7 - Performance média aplicando uma média móvel 12 na série temporal (MAE) . Fonte: Próprio Autor

De uma forma geral, as médias de resíduos dos modelos são iguais ou equivalentes em quase todos os recortes de dados.

4.2. Teste de Diebold Mariano

O teste de Diebold Mariano foi executado para todas as combinações de modelos com níveis de confiança estatísticos de 95%. A hipótese nula testada é que a acurácia entre os modelos é estatisticamente igual em conjuntos de testes fora da amostra.

Foi observado que quanto maior o número de observações da base de dados, menor é a distinção de acurácia dos modelos. O número de observações parece influenciar no percentual de modelos com acurácia igual conforme Figura 8.

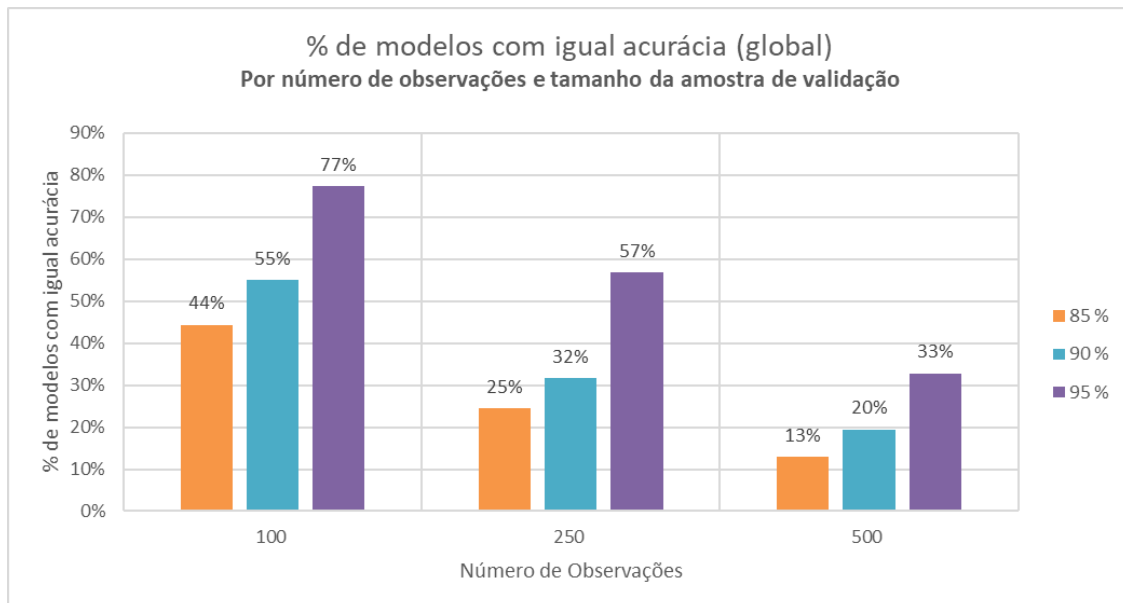


Figura 8 - Acurácia global dos Modelos, Diebold Mariano. Fonte: Próprio Autor

O comportamento observado na Figura 8 acontece de forma global, ou seja, para todas as combinações de modelo. Na Figura 9 foi observado o mesmo padrão ao comparar os modelos *VAR* vs *ARIMA* em uma mesma base de dados. Comportamento este também observado nas comparações entre *VEC* vs *ARIMA* (Figura 10) e *VEC* vs *VAR* (Figura 11).

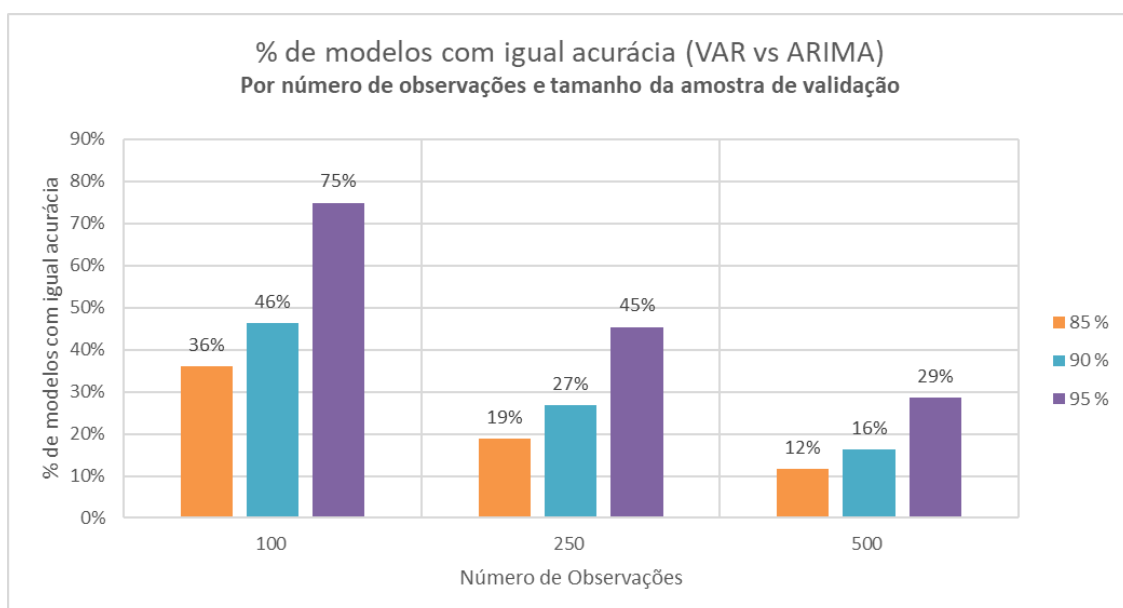


Figura 9 - Acurácia dos Modelos VAR vs ARIMA, Diebold Mariano. Fonte: Próprio Autor

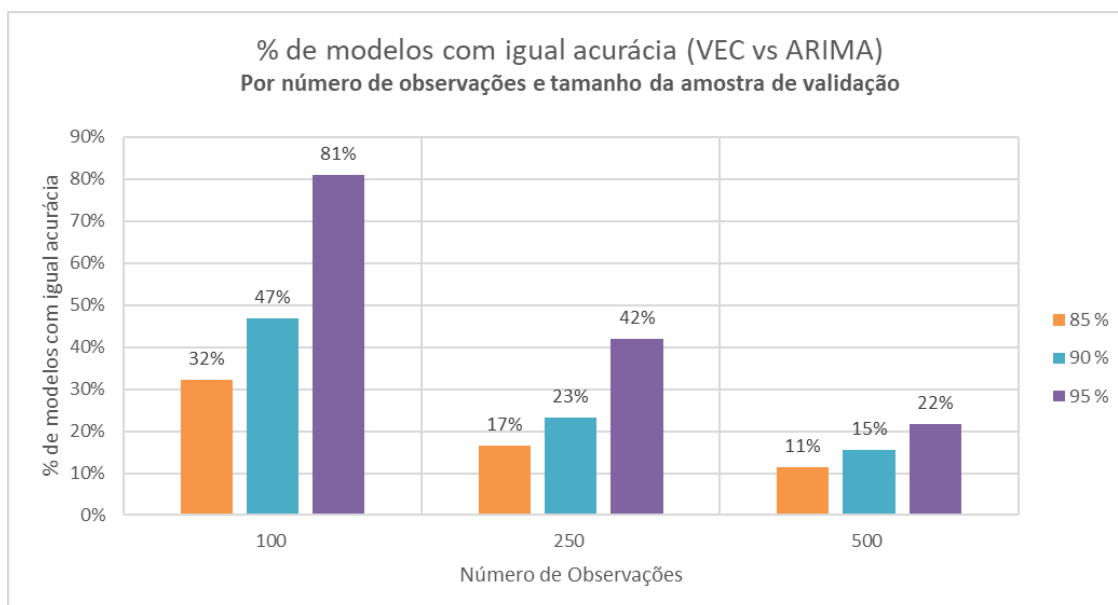


Figura 10 - Acurácia dos Modelos VEC vs ARIMA, Diebold Mariano. Fonte: Próprio Autor

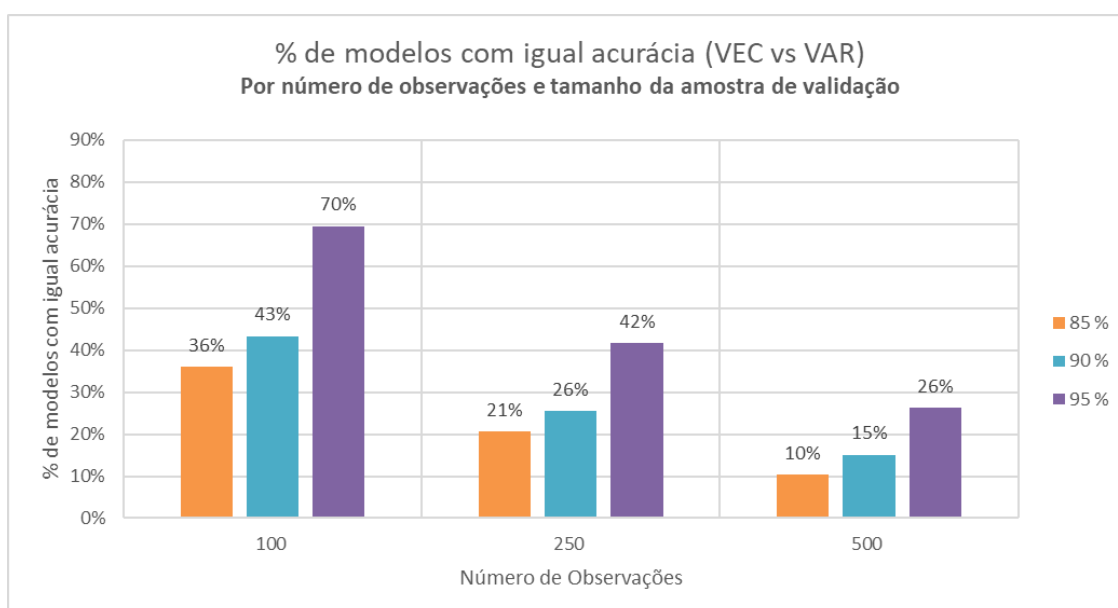


Figura 11 - Acurácia dos Modelos VEC vs VAR, Diebold Mariano. Fonte: Próprio Autor

Mais uma vez, alinhado aos resultados observados nas métricas de resíduos da sessão anterior ao variar os parâmetros de simulação, não foram observadas diferenças relevantes na acurácia dos modelos em conjuntos fora da amostra.

4.3. Teste de KS e ECDF`s

Ao avaliar os resultados do teste KS sobre os erros MAE, observa-se que para quase todas as combinações de parâmetros sem médias móveis não foi possível rejeitar, com 95% de confiança, a hipótese nula de que os resultados de performance dos modelos foram retirados de uma mesma amostra.

Ao observar as curvas de distribuição acumulada, tem-se uma sobreposição destas, indicando que não podemos rejeitar a hipótese que as curvas saíram de uma mesma população, ou seja, possuem performance similar.

A Figura 12 mostra as curvas de resíduos acumulados para os 3 modelos: ARIMA em azul, VAR em amarelo e VEC em vermelho. As três curvas estão sobrepostas para conjuntos de dados com 500 observações ou menos e 3 séries cointegradas.

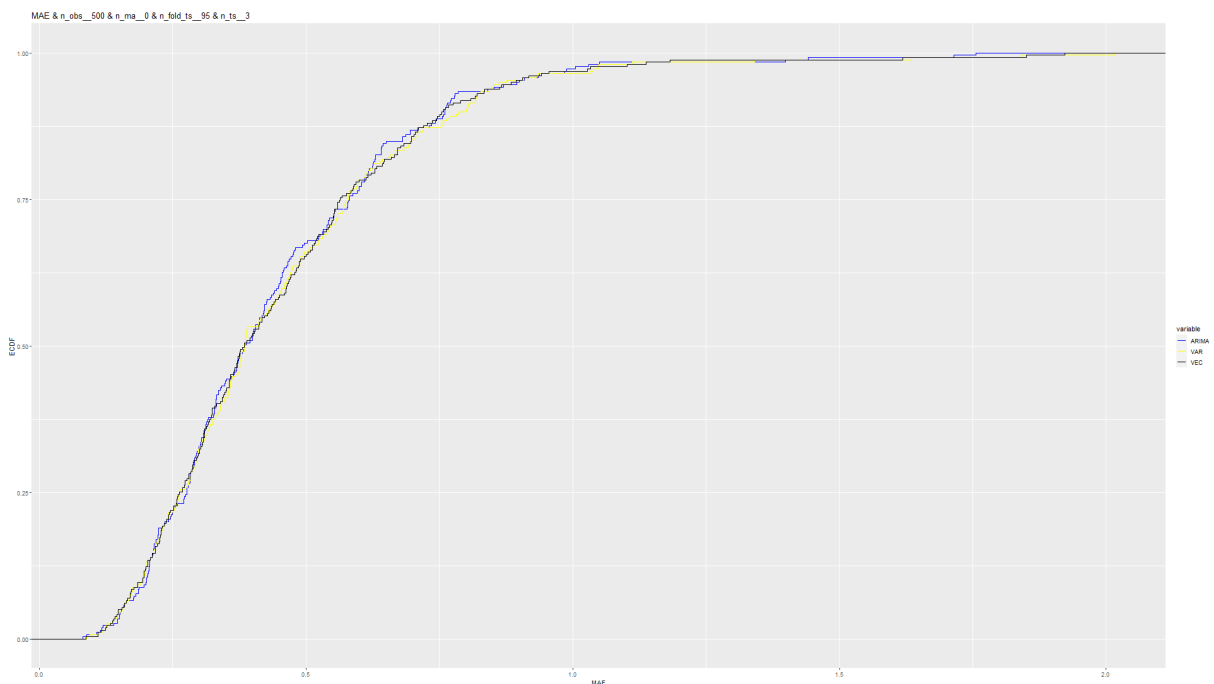


Figura 12 – Curva de Distribuição acumulada de resíduos MAE para conjuntos de dados com 500 observações e 3 séries cointegradas. . Fonte: Próprio Autor

O comportamento observado na Figura 12 acontece também na Figura 13 onde, ao invés de tratarmos de séries com 500 observações, tem-se séries mais curtas com 250 observações.

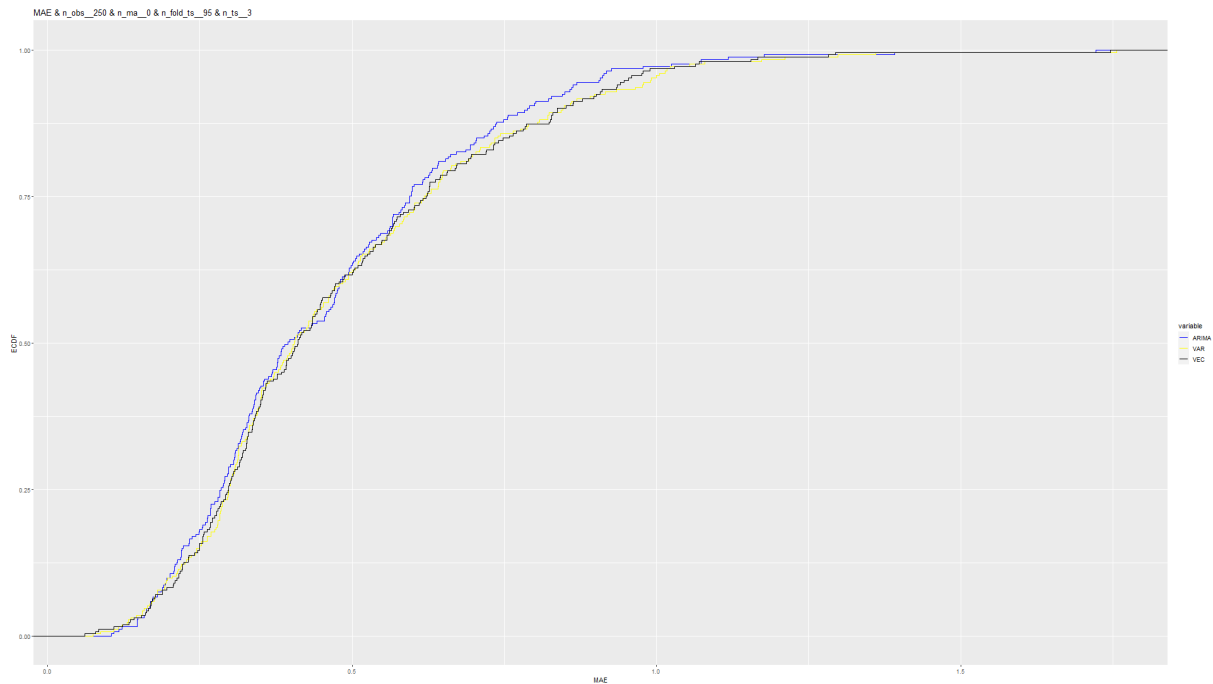


Figura 13 - Curva de Distribuição acumulada de resíduos MAE para conjuntos de dados com 250 observações e 3 séries cointegradas. . Fonte: Próprio Autor

A análise de KS, combinada aos gráficos de ECDF, está alinhada aos resultados anteriores que afirmam não existir uma diferença significativa de performance dos modelos nos conjuntos de dados testados.

Pode-se observar que em todas as métricas discutidas as simulações mostraram que a capacidade preditiva dos modelos em dados fora da amostra é, em média, equivalente aos valores dos intervalos de confiança a 95%.

5. Estudo de Caso Real

Com intuito de avaliar empiricamente as conclusões obtidas no estudo com simulações, foi realizado um experimento de modelagem com séries

temporais reais de câmbio (USD/BRL) e inflação acumulados em 12 meses (IPCA).

Os dados utilizados foram extraídos da plataforma financeira e site de notícias do mercado financeiro **Investing.com**. A série de inflação escolhida foi o [IPCA](#) e a de taxa de câmbio foi a [USD-BRL](#).

Ao observar a série de dados de IPCA vs USD-BRL em dois eixos verticais independentes e ao longo do tempo, é possível perceber que há uma relação de longo prazo entre elas até meados de 2017, quando realmente existe uma quebra estrutural, conforme Figura 14.

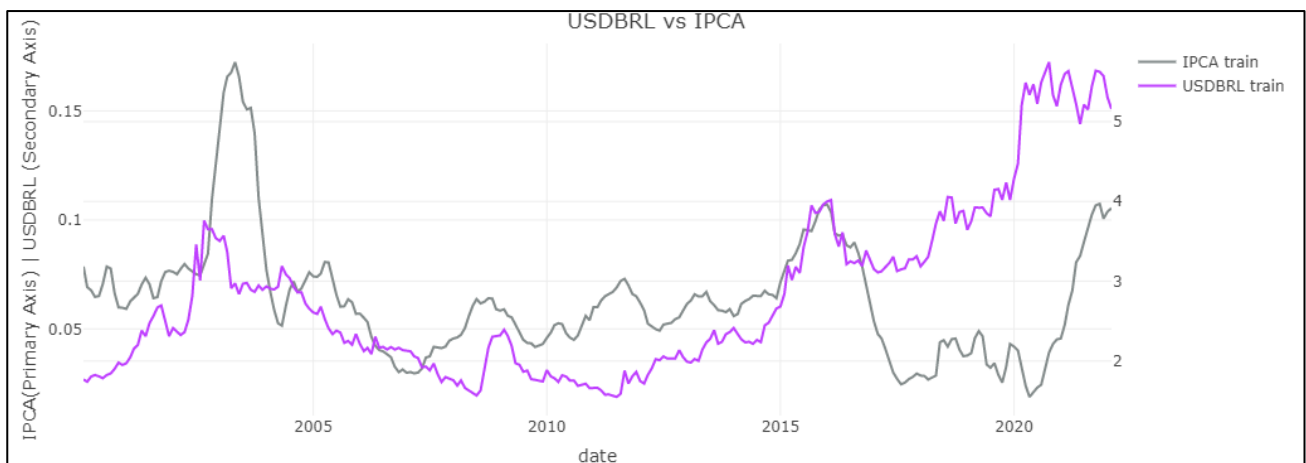


Figura 14 - USDBRL vs IPCA | Séries Temporais em 2 eixos verticais. Fonte: Próprio Autor

Dada essa mudança estrutural possivelmente motivada por crises econômicas e políticas, como por exemplo o impeachment da presidente Dilma Rousseff, foi definido que o estudo em questão iria utilizar os dados, como período para treinamento dos modelos, até dezembro de 2015, e os oito primeiros meses de 2016 seriam utilizados para validação, conforme Tabela 3.

Tabela 3 - Parâmetros para modelagem. Fonte: Próprio Autor

Parâmetros	Valor
Início período de Treino	2000-01-01
Fim do período de Treino	2015-12-01
Início período de Validação	2016-01-01
Fim do período de Validação	2016-08-01

Na Figura 15, as linhas contínuas representam os dados no período de treino do modelo, e os dados com marcadores redondos representam os dados do período de teste.

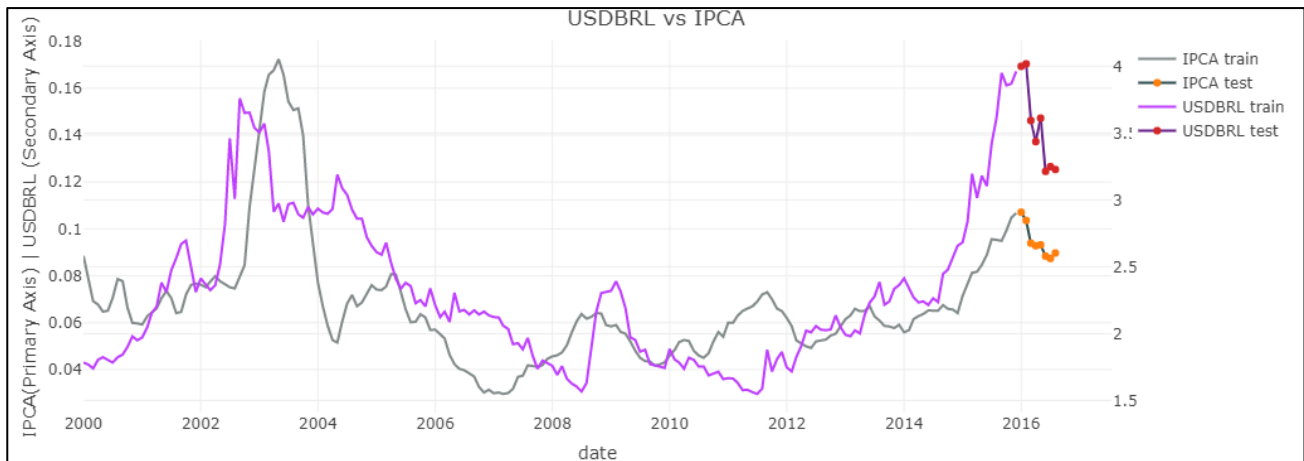


Figura 15 - USDBRL vs IPCA | Séries Temporais | Segmentados por Período de Treino e Teste. Fonte: Próprio Autor

Foi testada se as séries são $I(1)$, ou seja, estacionárias somente em primeiras diferenças. Tanto a série de inflação, quanto a taxa de câmbio são integradas de ordem 1 e estacionárias em primeiras diferenças, conforme Tabela 4.

Tabela 4 – ADF test. Fonte: Próprio Autor

Parâmetros	Augmented Dickey-Fuller Test
USDBRL	Dickey-Fuller = -1.1735, Lag order = 5, p-value = 0.9084, alternative hypothesis: stationary
IPCA	Dickey-Fuller = -2.8891, Lag order = 5, p-value = 0.2041, alternative hypothesis: stationary
Δ USDBRL	Dickey-Fuller = -4.5634, Lag order = 5, p-value = 0.01, alternative hypothesis: stationary
Δ IPCA	Dickey-Fuller = -5.0791, Lag order = 5, p-value = 0.01, alternative hypothesis: stationary

A partir da série temporal em primeiras diferenças do modelo, é possível validar visualmente o output do teste ADF, conforme Figura 16.

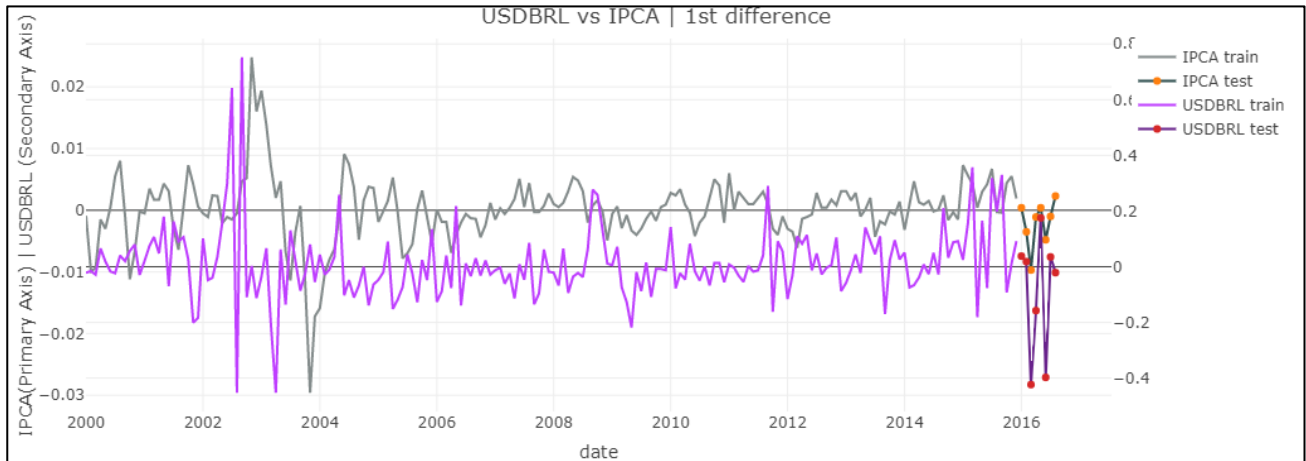


Figura 16 - $\Delta USDBRL$ vs $\Delta IPCA$ | Séries Temporais | Segmentados por Período de Treino e Teste. Fonte: Próprio Autor

A próxima etapa do algoritmo é checar se as séries são cointegradas, utilizando a estatística dos máximos autovalores de (Johansen, 1988). De acordo com os resultados do teste na Tabela 5, temos pelo menos um vetor de cointegração entre as séries de dados a 1% de significância.

Tabela 5 – Johansen-Procedure. Fonte: Próprio Autor

Parâmetros	Johansen-Procedure				
Values of test-statistic and critical values of test:		Test	10%	5%	1%
	$r \leq 1$	0.00	6.50	8.18	11.65
	$r = 0$	24.26	12.91	14.90	19.19
Eigenvectors, normalized to first column: (These are the cointegration relations)		IPCA.L2	USDBRL L2		
	IPCA.L2	1.0000	1.0000		
	USDBRL L2	-0.0403	0.1343		

A última etapa de seleção dos modelos é encontrar quais os conjuntos parâmetros que otimizam o critério bayesiano no período de treino do modelo. De acordo com a Tabela 6, os parâmetros que otimizaram o BIC foram de ordem 2 e diferenciação 1 para o modelo ARIMA, ordem 2 para o Δ VAR e ordem 1 para o VEC.

Tabela 6 – Modelos Escolhidos

Modelo	Especificação
ARIMA	ARIMA(2,1,0)
Δ VAR	Δ VAR(2)
VEC	VEC(1)

As métricas de erro MAE e MSE na comparação entre os dados observados e as previsões mostram que, para a modelagem da taxa de câmbio (USD-BRL) utilizando inflação IPCA, os modelos são equivalentes e apresentam níveis médios de erro em patamares iguais, conforme Figura 17 e Figura 18.

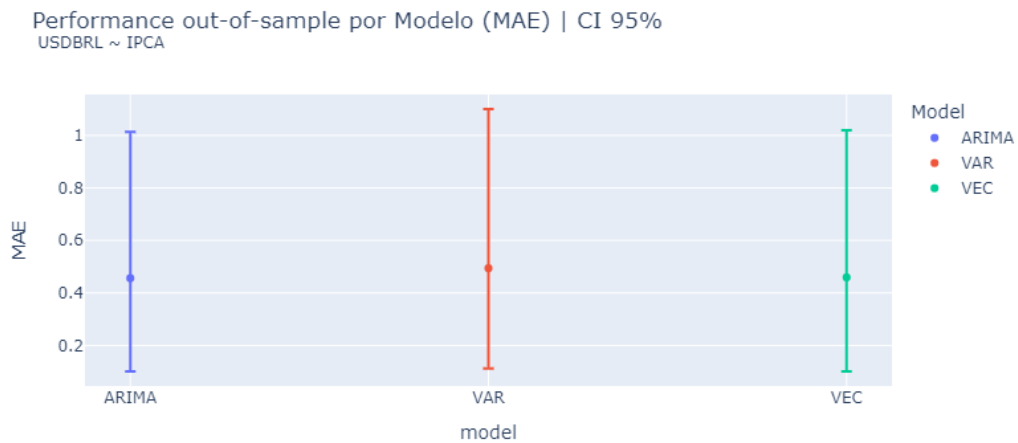


Figura 17 - Performance out-of-sample por Modelo (MAE) | USDBRL ~ IPCA. Fonte: Próprio Autor



Figura 18 - Performance out-of-sample por Modelo (MSE) | USDBRL ~ IPCA. Fonte: Próprio Autor

Portanto, no experimento com dados reais, a capacidade preditiva dos modelos foi similar, considerando os intervalos de confiança das métricas. Portanto houve um alinhamento entre as conclusões simuladas e com dados reais.

6. Conclusão

Este trabalho avaliou, via simulações de Monte Carlo, a capacidade preditiva do modelo univariado ARIMA em relação aos modelos multivariados VAR e VEC sob a condição de séries cointegradas. Para fornecer suporte empírico às conclusões obtidas via simulações, foi realizado um experimento de modelagem com séries temporais reais de câmbio (USD/BRL) e inflação acumulada em 12 meses (IPCA).

Os resultados mostraram que a capacidade preditiva dos modelos em dados fora da amostra de treinamento foi, em média, equivalente. Foi observado que o modelo VEC tem performance igual aos demais em diversos recortes de parâmetros. Era esperado que os modelos VEC performassem melhor em relação ao VAR e o ARIMA, a partir de expectativas criadas com a literatura econométrica, fato esse não observado. A perda de capacidade preditiva pode estar ligada à presença de erro de estimativa do parâmetro de vetor de cointegração.

Quanto maior o período de treinamento do modelo, menor é a média dos resíduos e modelos “treinados” com séries de dados longas (acima de 250 observações) não apresentaram resíduos estatisticamente distintos; os níveis de erro encontrados parecem convergir para uma mesma distribuição e média. O número de séries temporais no sistema não afetou a performance individual dos modelos. Quanto menor o número de observações da base de dados maior é a distinção de acurácia dos modelos fora da amostra de treinamento. A validação de modelos com MAE e MSE foram equivalentes; ambos convergiram para resultados de mesma interpretação.

Ao aplicar uma suavização por médias móveis de 12 períodos, os modelos VEC apresentaram performance igual ou inferior aos modelos ARIMA e VAR, fato explicado pela remoção de componentes da mais alta frequência que são modelados com maior precisão em modelos VEC.

No experimento com dados reais, a capacidade preditiva dos modelos foi similar dentro dos intervalos de confiança da média, alinhado a conclusão obtida via simulações. Contudo, os intervalos de confiança das estatísticas descritivas

alertam para uma necessidade maior de testes e de validações mais complexas para que os analistas/cientistas optem por uma estrutura em detrimento de outra. A simulação e toda validação apresentada reforçam e evidenciam a necessidade do uso de um conjunto grande de técnicas de validação para a seleção de modelos.

7. Referências Bibliográficas

ABDULLAH, L.; DWIVEDY, N. **ARIMA Model for Gold Bullion Coin. Selling Prices Forecasting**. International Journal of Advances in Applied Sciences (IJAAS), 1(4). 2012. p. 153-158.

AFANA, M. et al. **Artificial Neural Network for Forecasting Car Mileage per Gallon in the city**. Internacional Journal of Advanced Science and Technology, 124, p. 51-59, 2018. Disponível em: <http://dx.doi.org/10.14257/ijast.2018.124.05>.

ANGGRAENI, W.; BOGA, K.; MAHANANTO, F. **The Performance of ARIMAX Model and Vector Autoregressive (VAR) Model in Forecasting Strategic Commodity Price in Indonesia**. Procedia Computer Science, 124, p. 189-196, 2018. Disponível em: <http://doi.org/10.1016/j.procs.2017.12.146>.

ARAZ, O. M. et al. **Data Analytics for Operational Risk Management. Decision Sciences**, 51(6), p. 1316–1319, 2020. Disponível em: <https://doi.org/10.1111/dec.12443>.

BACHMEIER, L.J.; SWANSON, N.R. **Predicting inflation: Does the quantity theory help?**. Economic Inquiry, 43 (3), p. 570-585, 2005. Disponível em: <https://doi.org/10.1093/ei/cbi039>.

BOX, G.E.P.; JENKINS, G.M. **Time Series Analysis: Forecasting and Control**. Holden-Day, San Francisco, 1970.

CHEN, T.; XIE, L.; ZHANG, Y. **How does analysts' forecast quality relate to corporate investment efficiency?**. Journal of Corporate Finance, 43, p. 217-240, 2017. Disponível em: <https://doi.org/10.1016/j.jcorpfin.2016.12.010>.

DANIEL, W. Kolmogorov–Smirnov. **Applied Nonparametric Statistics**. 2. ed. Boston: PWS-Kent, 1990, p. 319–330. ISBN 978-0-534-91976-4.

Diebold, F. X. (2013). **Comparing Predictive Accuracy , Twenty Years Later : A Personal Perspective on the Use and Abuse of Diebold-Mariano Tests**. Comparing Model-Free Forecasts.

Diebold, F. X., & Mariano, R. S. (1995). **Comparing predictive accuracy**. Journal of Business and Economic Statistics, 13(3), 253–263. <https://doi.org/10.1080/07350015.1995.10524599>

Diebold, F. X., Rudebusch, G. D., & Borağan Aruoba, S. (2006). **The macroeconomy and the yield curve: A dynamic latent factor approach**. Journal of Econometrics, 131(1–2), 309–338. <https://doi.org/10.1016/j.jeconom.2005.01.011>

Lütkepohl, H. (1991). **Introduction to Multiple Time Series Analysis**. Springer Berlin Heidelberg. <https://doi.org/10.1007/978-3-662-02691-5>

Morettin, P. A. and Toloi, C. M. (2004). **Análise de Séries Temporais**. Edgard Blücher, São Paulo.

Enders, W. (2014) **Applied Econometric Time Series**. 4th Edition. John Wiley, New York.

Hatanaka, Michio (1996). **Time-Series-Based Econometrics: Unit Roots and Cointegration**. New York: Oxford University Press. pp. 48–49. ISBN 978-0-19-877353-5.

Mooney, C.Z. & Duval, R.D. (1993). **Bootstrapping: A nonparametric approach to statistical inference**. Sage University Paper series on Quantitative Applications in the Social Sciences, 07-095. Newbury Park, CA: Sage.

R Core Team (2016). **R: A language and environment for statistical computing**. R Foundation for Statistical Computing, Vienna. Available in: <https://www.R-project.org> (Accessed on Mar 10, 2022).

Zivot, E. & Wang, J. (2006). **Modeling Financial Time Series With S-Plus**. Springer Science + Business Media, Inc.

SCHWARZ, G (1978). **Estimating the dimensional of a model**. Annals of Statistics, Hayward, v.6, n.2, p.461-464, Mar.

Phillips, P.C.B. (1991). **Optimal Inference in Cointegrated Systems**. Econometrica, 59, 283-306

Engle, R. F., & Granger, C. W. J. (1987). **Co-Integration and Error Correction: Representation, Estimation, and Testing**. *Econometrica*, 55(2), 251. <https://doi.org/10.2307/1913236>

Fanchon, P., & Wendel, J. (1992). **Estimating VAR models under non-stationarity and cointegration: Alternative approaches for forecasting cattle prices**. *Applied Economics*, 24(2), 207–217. <https://doi.org/10.1080/00036849200000119>

Hargreaves, J. C., Annan, J. D., Edwards, N. R., & Marsh, R. (2004). **An efficient climate forecasting method using an intermediate complexity Earth System Model and the ensemble Kalman filter**. *Climate Dynamics*, 23(7–8), 745–760. <https://doi.org/10.1007/s00382-004-0471-4>

Hu, Z., Ge, Q., Li, S., & Xiong, M. (2020). **Artificial Intelligence Forecasting of Covid-19 in China**. *International Journal of Educational Excellence*, 6(1), 71–94. <https://doi.org/10.18562/ijee.054>

Ioannidis, J. P. A., Cripps, S., & Tanner, M. A. (2020). **Forecasting for COVID-19 has failed**. *International Journal of Forecasting*. <https://doi.org/10.1016/j.ijforecast.2020.08.004>

Johansen, S. (1988). **Statistical analysis of cointegration vectors**. *Journal of Economic Dynamics and Control*, 12(2–3), 231–254. [https://doi.org/10.1016/0165-1889\(88\)90041-3](https://doi.org/10.1016/0165-1889(88)90041-3)

Jones, S. A., Joy, M. P., & Pearson, J. (2002). **Forecasting demand of emergency care**. *Health Care Management Science*, 5(4), 297–305. <https://doi.org/10.1023/A:1020390425029>

Mohamed, J. (2020). **Time Series Modeling and Forecasting of Somaliland Consumer Price Index: A Comparison of ARIMA and Regression with ARIMA Errors**. 9(4), 143–153. <https://doi.org/10.11648/j.ajtas.20200904.18>

Ng, S. T., Fan, R. Y. C., & Wong, J. M. W. (2011). **An econometric model for forecasting private construction investment in Hong Kong**. *Construction Management and Economics*, 29(5), 519–534. <https://doi.org/10.1080/01446193.2011.570356>

Nyoni, T. (2018). **Modeling and Forecasting Naira / USD Exchange Rate In Nigeria: a Box - Jenkins ARIMA approach**. Mrpa, MPRA Paper No. 88622, 6–25. <https://mpra.ub.uni-muenchen.de/88622/>

Rogalski, R. J., & Vinso, J. D. (1977). **Stock Returns , Money Supply and the Direction of Causality**. American Finance Association, 32(4), 1017–1030.

Salisu, A. A., Ebuh, G. U., & Usman, N. (2020). **Revisiting oil-stock nexus during COVID-19 pandemic: Some preliminary results**. International Review of Economics and Finance, 69(June), 280–294. <https://doi.org/10.1016/j.iref.2020.06.023>

Sims, C. A. (1980). **Macroeconomics and Reality**. Econometrica, 48(1), 1. <https://doi.org/10.2307/1912017>

Søren, J. (1991). **Estimation and Hypothesis Testing of Cointegration Vectors in Gaussian Vector Autoregressive Models**. Econometrica, 59(6), 1551–1580. <https://doi.org/10.1017/CBO9781107415324.004>

Tehreem, H. S., Anser, M. K., Nassani, A. A., Abro, M. M. Q., & Zaman, K. (2020). **Impact of average temperature, energy demand, sectoral value added, and population growth on water resource quality and mortality rate: it is time to stop waiting around**. Environmental Science and Pollution Research. <https://doi.org/10.1007/s11356-020-09822-w>

Wangdi, K., Singhasivanon, P., Silawan, T., Lawpoolsri, S., White, N. J., & Kaewkungwal, J. (2010). **Development of temporal modelling for forecasting and prediction of malaria infections using time-series and ARIMAX analyses: A case study in endemic districts of Bhutan**. Malaria Journal, 9(1), 1–9. <https://doi.org/10.1186/1475-2875-9-251>

8. Repositório

Todos os scripts utilizados para confecção deste trabalho e resultados brutos estão armazenados no seguinte repositório do GitHub: https://github.com/ThiagueraBarao/master_degree.