



UNIVERSIDADE FEDERAL DE PERNAMBUCO  
CENTRO DE INFORMÁTICA  
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIAS DA COMPUTAÇÃO

FILIPPE COELHO DE LIMA DUARTE

**Um sistema híbrido baseado em combinação de preditores para previsão de vários passos à frente de séries temporais de taxas de mortalidade**

Recife

2024

FILIPPE COELHO DE LIMA DUARTE

**Um sistema híbrido baseado em combinação de preditores para previsão de vários passos à frente de séries temporais de taxas de mortalidade**

Tese apresentada ao Programa de Pós-graduação em Ciência da Computação do Centro de Informática da Universidade Federal de Pernambuco como requisito parcial para obtenção do grau de Doutor em Ciência da Computação.

**Área de Concentração:** Inteligência Computacional

**Orientador (a):** Paulo Salgado Gomes de Mattos Neto

**Coorientador (a):** Paulo Renato Alves Firmino

Recife

2024

.Catalogação de Publicação na Fonte. UFPE - Biblioteca Central

Duarte, Filipe Coelho de Lima.

Um sistema híbrido baseado em combinação de preditores para previsão de vários passos à frente de séries temporais de taxas de mortalidade / Filipe Coelho de Lima Duarte. - Recife, 2024. 137f.: il.

Tese (Doutorado) - Universidade Federal de Pernambuco, Centro de Informática, Programa de Pós-graduação em Ciência da Computação, 2024.

Orientação: Paulo Salgado Gomes de Mattos Neto.

Coorientação: Paulo Renato Alves Firmino.

Inclui referências e apêndices.

1. Sistemas híbridos; 2. Redes neurais artificiais; 3. Aprendizado profundo; 4. Previsão de múltiplos passos à frente; 5. Previsão de mortalidade. I. Mattos Neto, Paulo Salgado Gomes de. II. Firmino, Paulo Renato Alves. III. Título.

UFPE-Biblioteca Central

**Filipe Coelho de Lima Duarte**

**“Um sistema híbrido baseado em combinação de preditores  
para previsão de vários passos à frente de séries temporais de taxas  
de mortalidade”**

Tese de Doutorado apresentada ao Programa de Pós-Graduação em Ciência da Computação da Universidade Federal de Pernambuco, como requisito parcial para a obtenção do título de Doutor em Ciência da Computação. Área de Concentração: Inteligência Computacional

Aprovada em: 06/09/2024.

---

**Orientador: Prof. Dr. Paulo Salgado Gomes de Mattos Neto**

**BANCA EXAMINADORA**

---

Prof. Dr. Leandro Maciel Almeida  
Centro de Informática/UFPE

---

Prof. Dr. João Fausto Lorenzato de Oliveira  
Escola Politécnica de Pernambuco / UPE

---

Prof. Dr. Jorge Miguel Bravo  
Information Management School / Universidade Nova de Lisboa

---

Prof. Dr. Leonardo Nogueira Matos  
Departamento de Computação / UFS

---

Prof.Dr. Fernando Luiz Cyrino Oliveira  
Departamento de Engenharia Industrial / PUC-RJ

Dedico este trabalho à minha família que sempre me apoiou nesta jornada.

## **AGRADECIMENTOS**

Gostaria de expressar minha sincera gratidão a todos que contribuíram para o sucesso desta pesquisa e para a conclusão desta tese de doutorado.

Primeiramente, quero agradecer à minha família. Seu apoio incondicional, amor e paciência foram fundamentais para que eu pudesse me dedicar aos estudos e alcançar este objetivo. Agradeço especialmente à minha esposa e aos meus pais por estarem ao meu lado durante toda essa árdua jornada.

Ao meu orientador, Paulo Salgado Gomes de Mattos Neto, e ao meu coorientador, Paulo Renato Alves Firmino, sou imensamente grato. Pela paciência, orientação, conhecimento e incentivos, os quais foram essenciais para o desenvolvimento deste trabalho. Suas sugestões e correções foram valiosas e contribuíram significativamente para a qualidade e consecução desta tese.

Aos demais professores do Centro de Informática, agradeço por compartilharem seus conhecimentos e por proporcionarem um ambiente acadêmico estimulante. Suas disciplinas e debates enriqueceram minha formação.

Por fim, quero agradecer aos meus colegas da pós-graduação. As discussões em sala de aula, os grupos de estudo e as trocas de experiências foram enriquecedoras e fizeram parte dessa jornada acadêmica.

## RESUMO

Prever taxas de mortalidade para vários passos à frente é essencial em áreas como demografia, seguros e políticas públicas. Geralmente, os estudos de previsão de taxas de mortalidade têm se baseado na suposição de que os resíduos dos preditores são ruídos aleatórios. No entanto, problemas como má-especificação (i.e., sobreajuste e subajuste) ou o comportamento dinâmico do fenômeno temporal conduzem a modelos viesados ou imprecisos, deixando padrões nos resíduos ainda não modelados. Considerando esse problema, esta tese propõe um novo sistema híbrido para a previsão de vários passos à frente de taxas de mortalidade. O sistema proposto é composto de três etapas: (i) um modelo estatístico linear é usado para prever as taxas de mortalidade; (ii) um modelo de Aprendizado de Máquina (AM) é empregado para prever os padrões não-lineares dos resíduos de (i); e (iii) as estimativas da série e dos resíduos são combinadas por uma soma simples. Para cada conjunto de dados, o sistema seleciona em um *pool* de modelos de AM o mais promissor para prever os resíduos. O desempenho de previsão do sistema proposto foi avaliado por meio de uma análise experimental, a qual considerou quinze conjuntos de dados disponíveis na *Human Mortality Database* (HMD). Os dados englobaram as taxas de mortalidade da população total, masculina e feminina de cinco países: Austrália, Estados Unidos, França, Japão e Portugal, considerando cento e uma faixas etárias, desde o nascimento até os 100 anos de idade. A proposta utilizou o modelo Autorregressivo Integrado de Médias Móveis para prever as taxas de mortalidade e empregou um *pool* dos modelos Perceptron de Multicamadas (MLP), *Long Short-Term Memory* (LSTM) e *Neural basis expansion analysis for interpretable time series* (N-BEATS) para prever os resíduos do preditor linear. O sistema proposto foi comparado com alternativas da literatura, entre técnicas estatísticas tradicionais, modelos de AM e sistemas híbridos, utilizando *Mean Absolute Percentage Error* (MAPE) e *Mean Absolute Scaled Error* (MASE) como métricas de desempenho preditivo. A proposta alcançou uma média menor que 1,875% e 1,516 para o MAPE e MASE, respectivamente considerando todos os conjuntos de dados e dez horizontes de previsão. Ademais, os resultados sugerem que o sistema proposto é promissor, alcançando o menor erro de previsão conforme evidenciado pelo teste Nemenyi *Multiple comparisons with the best* (MCB). Perante o exposto, o sistema proposto pode contribuir, melhorando a acurácia de previsão das taxas de mortalidade, em atividades como projeção populacional, antecipação de necessidades na área da saúde, na busca pela solvência de seguradoras e fundos de pensão.

**Palavras-chaves:** Sistemas híbridos. Redes neurais artificiais. Aprendizado profundo. Previsão de múltiplos passos à frente. Previsão de mortalidade.

## ABSTRACT

Forecasting mortality rates for multiple steps ahead is essential in areas such as demography, insurance, and public policy. Generally, mortality rate forecasting studies have been based on the assumption that predictor residuals are random noise. However, issues such as misspecification (i.e., overfitting and underfitting) or the dynamic behavior of the temporal phenomenon lead to biased or imprecise models, leaving patterns in the residuals still unmodeled. Considering this problem, this thesis proposes a new hybrid system for multi-step ahead forecasting of mortality rates. The proposed system consists of three stages: (i) a linear statistical model is used to forecast mortality rates; (ii) a Machine Learning (ML) model is employed to predict non-linear patterns of residuals from (i); and (iii) the estimates of the series and residuals are combined by a simple sum. For each dataset, the system selects the most promising ML model from a pool to predict the residuals. The forecasting performance of the proposed system was evaluated through an experimental analysis, which considered fifteen datasets available in the Human Mortality Database (HMD). The data encompassed mortality rates for the total, male, and female populations of five countries: Australia, the United States, France, Japan, and Portugal, considering one hundred and one age groups, from birth to 100 years of age. The proposal used the Autoregressive Integrated Moving Average (ARIMA) model to forecast the mortality rate series and employed a pool of Multilayer Perceptron (MLP), Long Short-Term Memory (LSTM), and Neural basis expansion analysis for interpretable time series (N-BEATS) models to estimate the residuals of the linear predictor. The proposed system was compared with alternatives from the literature, including traditional statistical techniques, ML models, and hybrid systems, using Mean Absolute Percentage Error (MAPE) and Mean Absolute Scaled Error (MASE) as predictive performance metrics. The proposed system achieved an average lower than 1.875% and 1.516 for MAPE and MASE, respectively considering all datasets and ten forecast horizons. Furthermore, the results suggest that the proposal is promising, achieving the lowest forecast error as evidenced by the Nemenyi Multiple Comparisons with the Best (MCB) test. In light of the foregoing, by improving the accuracy of mortality rate forecasts, the proposed hybrid system can contribute in activities such as population projection, anticipation of healthcare needs, and ensuring the solvency of insurance companies and pension funds.

**Keywords:** Hybrid system. Artificial neural network. Deep Learning. Multi-step ahead fore-

casting. Mortality forecasting.

## LISTA DE FIGURAS

Figura 1 – Logaritmo natural das taxas de mortalidade da população francesa para as idades 0 a 100 anos entre os anos 1950 e 2019. . . . .	29
Figura 2 – Arquitetura do modelo Perceptron multicamadas (MLP) . . . . .	39
Figura 3 – Arquitetura do modelo LSTM. . . . .	41
Figura 4 – Arquitetura do modelo N-BEATS. . . . .	42
Figura 5 – Etapa de treinamento do sistema híbrido proposto utilizado nesta pesquisa. . . . .	53
Figura 6 – Etapas de teste do sistema híbrido proposto utilizado nesta pesquisa. . . . .	55
Figura 7 – Abordagem recursiva de previsão de múltiplos passos à frente usada na etapa de teste. . . . .	56
Figura 8 – Taxas de mortalidade para idades em escala natural e logarítmica da população francesa total. . . . .	61
Figura 9 – Logaritmo das taxas de mortalidade por idade da população masculina australiana. . . . .	62
Figura 10 – Logaritmo das taxas de mortalidade por idade da população feminina australiana. . . . .	62
Figura 11 – Logaritmo das taxas de mortalidade por idade da população total australiana. . . . .	63
Figura 12 – Logaritmo das taxas de mortalidade por idade da população masculina americana. . . . .	63
Figura 13 – Logaritmo das taxas de mortalidade por idade da população feminina americana. . . . .	64
Figura 14 – Logaritmo das taxas de mortalidade por idade da população total americana. . . . .	64
Figura 15 – Logaritmo das taxas de mortalidade por idade da população masculina francesa. . . . .	65
Figura 16 – Logaritmo das taxas de mortalidade por idade da população feminina francesa. . . . .	65
Figura 17 – Logaritmo das taxas de mortalidade por idade da população total francesa. . . . .	66
Figura 18 – Logaritmo das taxas de mortalidade por idade da população masculina japonesa. . . . .	66
Figura 19 – Logaritmo das taxas de mortalidade por idade da população feminina japonesa. . . . .	67
Figura 20 – Logaritmo das taxas de mortalidade por idade da população total japonesa. . . . .	67

Figura 21 – Logaritmo das taxas de mortalidade por idade da população masculina portuguesa. . . . .	68
Figura 22 – Logaritmo das taxas de mortalidade por idade da população feminina portuguesa. . . . .	68
Figura 23 – Logaritmo das taxas de mortalidade por idade da população total portuguesa.	69
Figura 24 – Distribuição dos <i>outliers</i> em relação ao MAPE para os modelos avaliados no experimento considerando o conjunto de teste . . . . .	79
Figura 25 – Efeito médio dos países, horizonte de previsão e gênero no MAPE para a curva de mortalidade entre 2010-2019. Modelo Lee-Carter, gênero feminino e país Austrália são os efeitos capturados pelo intercepto (MAPE médio de 3,0959%) que é comparável aos demais modelos, países e gêneros. . . . .	81
Figura 26 – Efeito médio do Sistema proposto e dos modelos monolíticos no MAPE para a curva de mortalidade entre 2010-2019. . . . .	82
Figura 27 – Efeito médio do Sistema proposto e dos demais modelos híbridos univariados de abordagem recursiva no MAPE para a curva de mortalidade entre 2010-2019. . . . .	82
Figura 28 – Efeito médio do Sistema proposto e dos modelos híbridos univariados e multivariados de abordagens MIMO e direta no MAPE para a curva de mortalidade entre 2010-2019. . . . .	83
Figura 29 – Boxplot do MAPE segregado por gênero, país e a categoria de se o modelo de previsão é multivariado. . . . .	84
Figura 30 – Boxplot do MASE segregado por gênero, país e a categoria de se o modelo de previsão é multivariado. . . . .	85
Figura 31 – MAPE (%) do Sistema proposto, ARIMA, ARIMA-LSTM e <i>Hybrid System for Mortality Forecasting</i> (HyS-MF) para a Austrália entre os anos 2010-2019	85
Figura 32 – MAPE (%) do Sistema proposto, ARIMA, ARIMA-LSTM e HyS-MF para os Estados Unidos entre os anos 2010-2019 . . . . .	86
Figura 33 – MAPE (%) do Sistema proposto, ARIMA, ARIMA-LSTM e HyS-MF para a França entre os anos 2010-2019 . . . . .	86
Figura 34 – MAPE (%) do Sistema proposto, ARIMA, ARIMA-LSTM e HyS-MF para o Japão entre os anos 2010-2019 . . . . .	86
Figura 35 – MAPE (%) do Sistema proposto, ARIMA, ARIMA-LSTM e HyS-MF para Portugal entre os anos 2010-2019 . . . . .	87

Figura 36 – MAPE (%) do Sistema proposto, LC e PLAT para a Austrália entre os anos 2010-2019 . . . . .	87
Figura 37 – MAPE (%) do Sistema proposto, LC e PLAT para os EUA entre os anos 2010-2019 . . . . .	88
Figura 38 – MAPE (%) do Sistema proposto, LC e PLAT para a França entre os anos 2010-2019 . . . . .	88
Figura 39 – MAPE (%) do Sistema proposto, LC e PLAT para o Japão entre os anos 2010-2019 . . . . .	88
Figura 40 – MAPE (%) do Sistema proposto, LC e PLAT para Portugal entre os anos 2010-2019 . . . . .	89
Figura 41 – MAPE (%) dos modelos univariados e modelo dos resíduos para a Austrália entre os anos 2010-2019 . . . . .	90
Figura 42 – MAPE (%) dos modelos univariados e modelo dos resíduos para os EUA entre os anos 2010-2019 . . . . .	90
Figura 43 – MAPE (%) dos modelos univariados e modelo dos resíduos para a França entre os anos 2010-2019 . . . . .	90
Figura 44 – MAPE (%) dos modelos univariados e modelo dos resíduos para o Japão entre os anos 2010-2019 . . . . .	91
Figura 45 – MAPE (%) dos modelos univariados e modelo dos resíduos para Portugal entre os anos 2010-2019 . . . . .	91
Figura 46 – Boxplot do Rank ordenado pela média do MAPE do Sistema proposto, sistemas híbridos recursivos, modelos monolíticos e multivariados recursivos para todos os conjuntos de dados. . . . .	92
Figura 47 – Boxplot do Rank ordenado pela média do MASE do Sistema proposto, sistemas híbridos recursivos, modelos monolíticos e multivariados recursivos para todos os conjuntos de dados. . . . .	92
Figura 48 – Teste de Nemenyi MCB que avalia a diferença entre os erros em absoluto de previsão do melhor modelo contra os demais. . . . .	95
Figura 49 – Curva de mortalidade na escala logarítmica e previsões pelo sistema proposto e ARIMA para a Austrália no ano de 2019. . . . .	103
Figura 50 – Curva de mortalidade na escala logarítmica e previsões pelo sistema proposto e ARIMA para os EUA no ano de 2019. . . . .	104

Figura 51 – Curva de mortalidade na escala logarítmica e previsões pelo sistema proposto e ARIMA para a França no ano de 2019. . . . .	105
Figura 52 – Curva de mortalidade na escala logarítmica e previsões pelo sistema proposto e ARIMA para o Japão no ano de 2019. . . . .	106
Figura 53 – Curva de mortalidade na escala logarítmica e previsões pelo sistema proposto e ARIMA para Portugal no ano de 2019. . . . .	107

## LISTA DE TABELAS

Tabela 1 – Valores dos hiper-parâmetros para os modelos não-lineares . . . . .	74
Tabela 2 – Seleção de modelo no sistema proposto para modelar os resíduos do ARIMA por conjunto de dados. . . . .	93
Tabela 3 – MAPE (%) para todos os conjuntos de dados, para o sistema proposto, modelos de abordagem recursiva e HyS-MF. Melhores valores em negrito. . .	96
Tabela 4 – MASE para todos os conjuntos de dados, para o sistema proposto, modelos de abordagem recursiva e HyS-MF. Melhores valores em negrito. . . . .	97
Tabela 5 – Diferença Percentual (DP) (%) do MAPE entre o modelo proposto e demais modelos para todos os conjuntos de dados. Destaques em negrito. . .	98
Tabela 6 – Diferença Percentual (DP) (%) do MASE entre o modelo proposto e demais modelos para todos os conjuntos de dados. Destaques em negrito. . . . .	99
Tabela 7 – Tempo médio de execução (s) do sistema proposto e demais modelos . . .	101

## LISTA DE ABREVIATURAS E SIGLAS

<b>AIC</b>	<i>Akaike Information Criterion</i>
<b>AM</b>	Aprendizado de Máquina
<b>ANN</b>	<i>Artificial Neural Networks</i>
<b>APC</b>	<i>Age-Period-Cohort</i>
<b>ARIMA</b>	<i>Autoregressive Integrated Moving Average</i>
<b>BMS</b>	Booth-Maindonald-Smith
<b>CBD</b>	Cairns-Blake-Dowd
<b>DP</b>	Diferença Percentual
<b>ETS</b>	<i>Exponential Smoothing</i>
<b>FDM</b>	<i>Functional Demographic Model</i>
<b>HMD</b>	<i>Human Mortality Database</i>
<b>HyS-MF</b>	<i>Hybrid System for Mortality Forecasting</i>
<b>LC</b>	Lee-Carter
<b>LSTM</b>	<i>Long Short-Term Memory</i>
<b>MAPE</b>	<i>Mean Absolute Percentage Error</i>
<b>MASE</b>	<i>Mean Absolute Scaled Error</i>
<b>MCB</b>	<i>Multiple comparisons with the best</i>
<b>MIMO</b>	<i>Multi-Input Multi-Output</i>
<b>MLP</b>	<i>Multilayer Perceptron</i>
<b>N-BEATS</b>	<i>Neural basis expansion analysis for interpretable time series</i>
<b>ReLU</b>	<i>Rectified Linear Unit</i>
<b>RMSE</b>	<i>Root Mean Squared Error</i>
<b>RNA</b>	Rede Neural Artificial
<b>RNAs</b>	Redes Neurais Artificiais
<b>SVR</b>	<i>Support Vector Regression</i>

## LISTA DE SÍMBOLOS

$\mathbf{a}_x$	Média do logaritmo natural da taxa de mortalidade para a idade $x$
$\beta_x$	Velocidade da variação da taxa de mortalidade para a idade $x$ em resposta às variações da tendência linear das taxas de mortalidade
$e_t$	Resíduos do modelo linear no tempo $t$
$e_{x,t}$	Série residual do conjunto de treinamento para idade $x$ no tempo $t$
$\mathbf{k}_t$	Tendência linear das taxas de mortalidade obtida por meio do primeiro componente principal para o ano $t$
$L_t$	Modelo linear estatístico no tempo $t$
$\hat{L}_{x,t}$	Modelo linear estatístico estimado para a idade $x$ no tempo $t$
$\ln$	Logaritmo natural
H	Quantidade de horizontes futuros do conjunto de teste
$\mathbf{m}_{x,t}$	Taxa central de mortalidade para a idade $x$ no ano $t$
MN	Modelagem não linear
ML	Modelagem Linear
$N_t$	Modelo não linear no tempo $t$
$\hat{N}_t$	Modelo não linear estimado no tempo $t$
$\hat{N}_{x,t}$	Estimativa do preditor não linear para a idade $x$ no tempo $t$
$z_t$	Série temporal no instante de tempo $t$
$\hat{z}_t$	Valor estimado para o sistema híbrido $z$ no tempo $t$
$\varepsilon_t$	Erro aleatório no tempo $t$
$a_t$	Termo do erro do ruído branco
$\phi(B)$	Operador autorregressivo estacionário

$\varphi(B)$	Operador autorregressivo generalizado
$\Delta^d$	Operador da diferença de ordem $d$
$\theta(B)$	Operador de média móvel
$d$	Grau da diferença
<b>W</b>	Matriz de pesos
<b>b</b>	Vector de <i>bias</i>
FC	Camada totalmente conectada - <i>Fully Connected</i>
<b>x</b>	Matriz dos dados de entrada
$\hat{\mathbf{x}}$	Estimativa da matriz dos dados de entrada
$\hat{\mathbf{z}}$	Estimativa da previsão final para todos os blocos residuais
ReLU	<i>Rectified Linear Unit</i>

## SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO</b>	<b>20</b>
1.1	OBJETIVOS	25
1.2	QUESTÕES DE PESQUISA	25
1.3	ESTRUTURA DO TRABALHO	27
<b>2</b>	<b>REVISÃO DA LITERATURA</b>	<b>28</b>
2.1	PREVISÃO DE MORTALIDADE	28
<b>2.1.1</b>	<b>Modelos para previsão de mortalidade</b>	<b>30</b>
2.1.1.1	<i>Modelos multivariados</i>	30
2.1.1.2	<i>Modelos univariados</i>	35
2.2	AUTORREGRESSIVO INTEGRADO DE MÉDIAS MÓVEIS	36
2.3	PERCEPTRON MULTICAMADAS	38
2.4	<i>LONG-SHORT TERM MEMORY</i>	40
2.5	N-BEATS	42
2.6	PREVISÃO DE SÉRIES TEMPORAIS CURTAS	43
2.7	PREVISÃO DE MÚLTIPLOS PASSOS À FRENTE	45
2.8	SISTEMAS HÍBRIDOS PARA PREVISÃO DE SÉRIES TEMPORAIS	47
2.9	CONSIDERAÇÕES DO CAPÍTULO	50
<b>3</b>	<b>SISTEMA PROPOSTO</b>	<b>51</b>
3.1	SISTEMA HÍBRIDO PROPOSTO PARA PREVISÃO DE VÁRIOS PASSOS À FRENTE	51
<b>3.1.1</b>	<b>Etapa de treinamento</b>	<b>53</b>
<b>3.1.2</b>	<b>Etapa de teste</b>	<b>55</b>
3.2	CONSIDERAÇÕES DO CAPÍTULO	57
<b>4</b>	<b>PROTOCOLO EXPERIMENTAL</b>	<b>58</b>
4.1	BASE DE DADOS	58
4.2	MÉTRICAS DE AVALIAÇÃO	70
4.3	PREVISÃO DAS TAXAS DE MORTALIDADE E ANÁLISE ESTATÍSTICA	71
4.4	CONSIDERAÇÕES CAPÍTULO	77
<b>5</b>	<b>RESULTADOS</b>	<b>78</b>
5.1	ANÁLISE DOS RESULTADOS DA REGRESSÃO	78

5.2	COMPARAÇÃO ENTRE ABORDAGENS UNIVARIADAS E MULTIVARIADAS . . . . .	84
5.3	COMPARAÇÃO DO DESEMPENHO DE PREVISÃO DAS CURVAS DE MORTALIDADE . . . . .	85
5.4	TEMPO DE EXECUÇÃO DOS ALGORITMOS . . . . .	100
5.5	ANÁLISE DE PLAUSIBILIDADE DAS PREVISÕES . . . . .	101
<b>6</b>	<b>CONCLUSÃO . . . . .</b>	<b>108</b>
6.1	CONSIDERAÇÕES FINAIS . . . . .	108
6.2	LIMITAÇÕES DA PESQUISA . . . . .	109
6.3	TRABALHOS FUTUROS . . . . .	110
6.4	ARTIGO PUBLICADO . . . . .	112
	<b>REFERÊNCIAS . . . . .</b>	<b>113</b>
	<b>APÊNDICE A – EFEITO PERCENTUAL MÉDIO NO MASE OBTIDO POR MEIO DA REGRESSÃO GAMA . . . . .</b>	<b>123</b>
	<b>APÊNDICE B – MASE PARA OS MODELOS ARIMA, ARIMA-LSTM, HYS-MF E PROPOSTO SEGREGADO POR CONJUNTO DE DADOS . . . . .</b>	<b>125</b>
	<b>APÊNDICE C – MASE PARA OS MODELOS PROPOSTO, LC E PLAT SEGREGADO POR CONJUNTO DE DADOS . . . . .</b>	<b>127</b>
	<b>APÊNDICE D – MASE PARA MODELOS UNIVARIADOS SEGREGADOS POR PAÍS, GÊNERO E MODELAGEM RESIDUAL . . . . .</b>	<b>129</b>
	<b>APÊNDICE E – PREVISÕES DO SISTEMA PROPOSTO . . . . .</b>	<b>131</b>

## 1 INTRODUÇÃO

A previsão da mortalidade humana é um campo de pesquisa ativo que atrai considerável atenção devido às suas aplicações em uma variedade de setores (Olivieri, 2001; Janssen, 2018; Perla et al., 2021). Na demografia, a precisão na previsão da taxa de mortalidade é fundamental para estimativas populacionais e expectativa de vida (Shi, 2021). No contexto governamental, essas previsões são cruciais para a formulação de políticas públicas nos planos de seguridade social, em iniciativas na saúde pública, e na formulação de políticas educacionais (Hyndman; Ullah, 2007; Vanella; Deschermeier; Wilke, 2020; Nigri; Levantesi; Aburto, 2022). Além disso, fundos de pensão e companhias de seguros utilizam essas previsões para tarifar contratos de seguros de vida, calcular as contribuições previdenciárias, e avaliar obrigações de longo prazo, influenciando assim estratégias de gestão de riscos (Olivieri, 2001; Giacometti et al., 2012; Mitchell et al., 2013; Bravo; Freitas, 2018; Janssen, 2018; Bravo, 2022). O avanço da economia financeira também levou ao desenvolvimento de novos produtos financeiros focados na gestão de riscos de mortalidade, como títulos indexados à longevidade e seus derivativos (Bravo, 2007; Bravo; Freitas, 2018; Levantesi; Nigri, 2020; Bravo, 2021b; Bravo, 2021a). Em cenários de pandemias globais, como a COVID-19, a previsão da mortalidade torna-se ainda mais crítica para a alocação de recursos de saúde (Hierro et al., 2020) e para a implementação de medidas de contenção (Zivkovic et al., 2021).

Companhias de seguros empregam modelos de previsão de mortalidade como ferramentas essenciais para estimar fluxos de caixa futuros e para gerenciar os riscos associados a contratos de seguros de vida e previdência (Olivieri, 2001; Mitchell et al., 2013). A precisão dessas previsões é de suma importância, pois ela afeta diretamente a eficiência operacional e financeira dessas organizações. Desenvolver métodos de projeção mais rigorosos não apenas aprimora a gestão de riscos, mas também contribui para a criação de produtos de seguro mais competitivos e economicamente viáveis. Isso ocorre porque uma modelagem mais precisa do risco de vida permite uma precificação mais justa e uma alocação de capital mais eficiente. Portanto, a busca por técnicas de previsão mais precisas é uma área de pesquisa ativa, cujos avanços podem ter implicações significativas para a indústria de seguros e para a sociedade como um todo (Atance; Debón; Navarro, 2020).

A dinâmica das taxas de mortalidade é um fenômeno complexo e não-linear que tem sido objeto de extenso estudo acadêmico (Mitchell et al., 2013; Wu; Wang, 2018; Lam; Wang,

2021). Os estudos, de uma forma geral, analisam séries temporais de taxas de mortalidade específicas por idade (ou faixa etária) de uma população, com cada série representando a evolução temporal da taxa de mortalidade para uma idade particular.

Nesse sentido, a literatura existente sobre o tema pode ser subdividida em duas abordagens principais, modelos multivariados e univariados (Feng; Shi, 2018). Segundo Feng e Shi (2018), os primeiros são concebidos para modelar múltiplas séries de taxas de mortalidade de uma ou mais populações empregando apenas um modelo de previsão, ao passo que os modelos univariados são utilizados para aprender os padrões temporais subjacentes a cada série de taxa de mortalidade de forma independente.

A abordagem multivariada, particularmente modelos derivados da estrutura Age-Period-Cohort (APC), é a mais prevalente (Hunt; Blake, 2021) e foi popularizada inicialmente pelo modelo Lee-Carter (LC) (Lee; Carter, 1992). Este modelo emprega técnica de fatoração matricial para reduzir a dimensionalidade dos dados, concentrando-se em componentes principais que podem ser projetados para o futuro. No entanto, frequentemente ele opera sob a suposição de que um único componente principal, geralmente modelado por um modelo estatístico linear, pode capturar a tendência temporal subjacente em todas as taxas de mortalidade. Esta suposição tem suas limitações, pois negligencia a dinâmica variação das taxas de mortalidade para diferentes faixas etárias ao longo do tempo (Mitchell et al., 2013). Além disso, ao assumir uma tendência temporal comum para todas as idades, a matriz de covariâncias das taxas de mortalidade pode superestimar a dependência temporal entre diferentes faixas etárias. Este erro de especificação pode, por sua vez, comprometer a acurácia dos preditores (Mitchell et al., 2013; Wu; Wang, 2018).

Algumas extensões subsequentes ao LC envolveram a utilização da distribuição etária das mortes e fixação do período de ajuste para capturar não-linearidades no componente temporal (Booth; Maindonald; Smith, 2002), o emprego de efeitos de coorte (Renshaw; Haberman, 2006), a utilização de múltiplos termos para as idades e período temporal por meio da análise funcional de dados (Hyndman; Ullah, 2007), a aplicação de modelos lineares generalizados (MLG)<sup>1</sup> (Hatzopoulos; Haberman, 2009), o exercício de modelos de AM (Hainaut, 2018; Nigri et al., 2019).

<sup>1</sup> MLG é uma técnica estatística de regressão que flexibiliza a escolha da distribuição do componente aleatório e da função de ligação entre a variável alvo (resposta) e o preditor linear. Ela tem sido aplicada em uma vasta gama de disciplinas uma vez que modela relações não-lineares. A referência básica para se aprofundar no tema é McCullagh e Nelder (1989) e sob uma perspectiva aplicada, sugere-se a leitura de Faraway (2016).

O debate entre modelos multivariados e univariados na previsão de taxas de mortalidade é uma questão central na literatura. Modelos multivariados, que são frequentemente favorecidos quando há uma correlação cruzada significativa entre diferentes faixas etárias, têm a vantagem de utilizar um conjunto mais amplo de informações (Preez; Witt, 2003; Feng; Shi, 2018). No entanto, estudos recentes (Feng; Shi, 2018; Wu; Wang, 2018; Lam; Wang, 2021) questionam a eficácia dessa abordagem, especialmente quando a correlação cruzada entre faixas etárias distantes é insignificante. Nesse contexto, modelos univariados emergem como uma alternativa atraente, capazes de modelar faixas etárias específicas sem a necessidade de considerar correlações cruzadas (Feng; Shi, 2018; Mitchell et al., 2013; Wu; Wang, 2018; Lam; Wang, 2021). Em relação aos modelos univariados, algumas metodologias têm sido estabelecidas como *benchmarks* na previsão da mortalidade, incluindo modelos *Autoregressive Integrated Moving Average* (ARIMA) (McNown; Rogers, 1989), *Exponential Smoothing* (ETS)<sup>2</sup> (Feng; Shi, 2018), e Regressão de Processos Gaussianos. Cada um desses modelos apresenta suas próprias vantagens e limitações. Por exemplo, enquanto modelos ARIMA são eficazes em capturar padrões lineares, eles falham em abordar componentes não-lineares. O modelo ETS, por outro lado, tem a limitação de dar menos peso a observações mais distantes e não considerar variações na dinâmica ao longo do tempo. A Regressão de Processos Gaussianos, embora flexível, é sensível à escolha das funções de covariância e assume uma distribuição normal dos dados ao longo do tempo.

Recentemente, algoritmos de Aprendizado de Máquina (AM)<sup>3</sup>, incluindo Redes Neurais Artificiais (RNAs)<sup>4</sup>, têm ganhado destaque na literatura para a previsão de taxas de mortalidade (Richman; Wüthrich, 2021; Hainaut, 2018; Perla et al., 2021; Bravo, 2021b). Esses modelos são promissores devido a sua capacidade de aprendizado, natureza orientada a dados e flexibilidade em modelar relações não-lineares. No entanto, a aplicação isolada de modelos de AM apresenta desafios (Zhang, 2003; Khashei; Bijari, 2011), incluindo a possibilidade de erros de especificação e problemas de sobre(sub)ajuste (de Mattos Neto; Cavalcanti; Madeiro, 2017; de Mattos Neto et al., 2021), corroborando com o teorema de Wolpert (1996), o qual

<sup>2</sup> ETS é uma família de métodos de previsão de séries temporais que produz previsões como médias ponderadas de observações passadas, atribuindo pesos que decaem exponencialmente à medida que as observações se tornam mais antigas (Hyndman et al., 2008).

<sup>3</sup> Segundo Mitchell et al. (1997), os algoritmos de Aprendizado de Máquina são aqueles que aprimoram seu desempenho automaticamente com a experiência. Considere consultar as obras de Mitchell et al. (1997), Bishop e Nasrabadi (2006), Goodfellow, Bengio e Courville (2016) para um melhor esclarecimento sobre AM.

<sup>4</sup> Para um maior aprofundamento no tema de redes neurais artificiais recomenda-se a leitura de Haykin (2009) e Goodfellow, Bengio e Courville (2016).

argumenta contra a existência de um algoritmo universalmente superior em todos os cenários.

Diante dessas dificuldades, e baseando-se na premissa de que as séries temporais frequentemente apresentam padrões lineares quanto não-lineares e que modelos não-lineares de AM não possuem a mesma capacidade de lidar com ambos os padrões supracitados simultaneamente, Zhang (2003) propôs um sistema híbrido que utilizou-se dos pontos fortes do modelo ARIMA para modelar e prever séries temporais e uma rede *Multilayer Perceptron* (MLP) para prever os resíduos daquele. Partindo dessa hipótese, esses sistemas têm sido utilizados em diversos contextos como finanças (Pai; Lin, 2005), epidemiologia (Chakraborty; Chattopadhyay; Ghosh, 2019; Morais; Gomes, 2022), economia (Dave et al., 2021), energia (Shi; Guo; Zheng, 2012; Fan et al., 2021), hidrologia (Su et al., 2024). Nesses contextos, os sistemas híbridos têm consistentemente superado os modelos individuais, sejam eles lineares ou não-lineares, em termos de acurácia de previsão. Aliás, estudos (Oliveira; Silva; de Mattos Neto, 2021; Santos Júnior et al., 2023; Medina; Oliveira, 2023) destacam que a combinação de preditores não apenas melhora a acurácia, mas também proporciona maior estabilidade, reduzindo o viés e variância das previsões quanto a incerteza de seleção de modelo.

Adicionalmente, a tarefa de prever taxas de mortalidade em vários passos à frente a partir de séries temporais curtas (i.e., aquelas com amostras pequenas ou poucas observações) acrescenta um nível adicional de dificuldade. De fato, a previsão de séries temporais curtas é um desafio complexo, particularmente dada a limitação da quantidade de padrões temporais disponíveis para o treinamento dos modelos. Evidências (Cerqueira; Torgo; Soares, 2022; Cruz-Nájera et al., 2022) apontam que modelos estatísticos produzem desempenho superior aos modelos de AM em séries temporais curtas. Além do mais, quando se considera um cenário de múltiplos passos à frente, as diferenças de desempenho entre os modelos estatísticos e de AM se tornam mais equilibradas com o aumento do tamanho da série (Cerqueira; Torgo; Soares, 2022). Por outro lado, foi observado que modelos híbridos podem produzir previsões mais acuradas que suas abordagens individuais em cenários de séries pequenas (Meng; Han; Hou, 2022; Duarte; de Mattos Neto; Firmino, 2024).

Várias estratégias têm sido propostas para abordar o problema de previsão de múltiplos passos à frente, incluindo abordagens recursivas, diretas e Multi-Input Multi-Output (MIMO) (Taieb et al., 2012). A abordagem recursiva emprega um modelo treinado para um passo à frente e o utiliza repetidamente, alimentando as previsões geradas como entrada para previsões subsequentes. A abordagem direta treina um modelo separado para cada horizonte de previsão, enquanto a MIMO considera todos os horizontes de previsão simultaneamente,

treinando um único modelo com múltiplas saídas (Bontempi, 2008; Hamzaçebi; Akay; Kutay, 2009; Taieb et al., 2012; Taieb, 2014; Ming et al., 2014). Entre essas, a abordagem recursiva é notável por sua simplicidade e eficácia preditiva, enquanto outras metodologias exigem mais dados e tempo de processamento, sem garantir melhorias significativas na acurácia da previsão (Taieb et al., 2012; Ming et al., 2014; Beyaztas; Shang, 2022).

Neste cenário, a implementação de um sistema híbrido que emprega uma seleção de modelos de AM de abordagem recursiva para a previsão de vários passos à frente de séries residuais de taxas de mortalidade emerge como uma solução interessante para os desafios anteriormente delineados. Desse modo, este estudo introduz um sistema híbrido que combina as forças do modelo estatístico (e.g., ARIMA), com alternativas de AM, incluindo Perceptron Multicamadas, LSTM (Hochreiter; Schmidhuber, 1997)<sup>5</sup> e N-BEATS (Oreshkin et al., 2019)<sup>6</sup>, que têm demonstrado desempenho superior em diversas aplicações de previsão de séries temporais (Zhang, 2003; Khashei; Bijari, 2011; Oliveira; Silva; de Mattos Neto, 2021; Oreshkin et al., 2019; Oreshkin et al., 2021; Ahmed; Hassan; Mstafa, 2022). O sistema proposto é estruturado em três fases distintas: (i) modelagem das séries temporais das taxas de mortalidade empregando o modelo estatístico linear, da qual se obtém as séries residuais; (ii) modelagem residual por meio da geração de um *pool* de modelos heterogêneos não-lineares de abordagem recursiva para previsão de múltiplos passos à frente, e (iii) composição da previsão final, derivada da soma aritmética da saída linear do ARIMA com a previsão gerada pelo modelo não-linear selecionado. Para cada conjunto de dados, o sistema proposto faz a seleção do melhor modelo AM do *pool* gerado em (ii) para prever os resíduos.

A eficácia do sistema híbrido foi validada em quinze conjuntos de dados cada qual considerando cento e uma séries de taxas de mortalidade disponíveis no repositório da HMD (Human Mortality Database, 2023), utilizando o MAPE e o MASE como métricas de desempenho. Os resultados, avaliados através do teste estatístico Nemenyi MCB, indicaram que o sistema híbrido proposto oferece melhorias estatisticamente significativas em relação às abordagens existentes. Diferentemente dos sistemas híbridos tradicionais, que selecionam o melhor modelo não-linear para cada série residual ou padrão isoladamente, o sistema proposto destaca-se por sua abordagem inovadora que busca o melhor modelo para todas as séries de uma dada população de forma simultânea. Essa estratégia demonstra ser fundamental para o aprimoramento

<sup>5</sup> *Long Short-Term Memory* (LSTM) pertence à família de redes neurais recorrentes que são especializadas em processar dados sequenciais. Consultar Goodfellow, Bengio e Courville (2016) para mais detalhes.

<sup>6</sup> N-BEATS combina arquiteturas de perceptron multicamadas e *residual nets* e foi projetado para realizar modelagem e previsão de séries temporais (Oreshkin et al., 2019; Oreshkin et al., 2021).

ramento da acurácia das previsões, especialmente considerando que, na maioria dos casos práticos, a avaliação do desempenho se baseia na análise da curva de mortalidade como um todo, integrando todas as idades. Ao levar em conta as relações entre as diferentes faixas etárias, o sistema proposto é capaz de identificar padrões e realizar previsões mais precisas.

Isto posto, o sistema proposto contribui para diversas áreas do conhecimento por meio do fornecimento de previsões acuradas de taxas de mortalidade, o que impacta positivamente no fornecimento de informações estratégicas para a tomada de decisão. Dessa forma, agentes públicos podem melhorar o processo de elaboração de políticas públicas na área de saúde, educação, sistemas de seguridade social. Ademais, entidades seguradoras e de fundos de pensão podem reduzir o risco de insolvência, calcular os prêmios de seguros de forma mais adequada, e fixar os valores das contribuições previdenciárias que consiga manter o equilíbrio atuarial e financeiro.

## 1.1 OBJETIVOS

O objetivo principal deste trabalho é propor um novo sistema híbrido para previsão de taxas de mortalidade em múltiplos passos à frente que combina um modelo estatístico linear com um modelo de AM não-linear para modelagem residual.

Entre os objetivos específicos deste trabalho, estão:

1. Desenvolver um sistema que realiza previsões em múltiplos passos à frente e combina modelos estatísticos lineares para prever as taxas de mortalidade e de aprendizado de máquina não-linear para prever os resíduos do preditor linear.
2. Aplicar o sistema proposto em dados reais de taxas de mortalidade.
3. Analisar o desempenho do sistema proposto e comparar com técnicas já existentes na literatura.

## 1.2 QUESTÕES DE PESQUISA

O presente trabalho é guiado pela seguinte questão de pesquisa:

- **Q1: Sistemas híbridos que combinam um modelo estatístico linear para prever as séries de taxas de mortalidade e um modelo de AM não-linear para prever**

---

**os resíduos do preditor linear em múltiplos passos à frente alcançam acurácia de previsão superior em relação aos modelos da literatura?**

A literatura de previsão de séries temporais evidencia que a combinação de modelos estatísticos lineares e de AM não lineares resulta em melhor acurácia preditiva em diversos conjuntos de dados. Tendo isso em vista, os sistemas híbridos desenvolvidos partem da ideia de modelar os padrões lineares por modelos lineares estatísticos e os resíduos, subjacentes ao modelo linear, a partir de um modelo não linear de AM. Nesse sentido, é importante investigar se os sistemas híbridos apresentam desempenho superior em relação às abordagens monolíticas para previsão de taxas de mortalidade em múltiplos passos à frente.

▪ **Q2: Qual abordagem de previsão para múltiplos passos à frente demonstra o melhor desempenho na previsão das séries residuais curtas?**

Diante de diversas abordagens de previsão de múltiplos passos à frente, e na presença de resultados mistos sobre qual delas seria a mais adequada para o cenário de modelagem residual de séries curtas, entende-se relevante investigar qual é a abordagem mais promissora. A identificação de uma abordagem que possa lidar efetivamente com as peculiaridades das séries residuais curtas, ao mesmo tempo em que maximiza o desempenho preditivo, contribuiria significativamente para o avanço das técnicas de previsão híbridas.

▪ **Q3: A seleção de modelo de AM não-linear na modelagem residual melhora a acurácia de previsão do sistema híbrido na previsão de taxas de mortalidade em vários passos à frente?**

Ao incorporar a seleção de modelo não-linear como uma etapa central da modelagem, o sistema proposto busca garantir que o modelo mais adequado seja escolhido, compreendendo toda a curva de mortalidade, no lugar de cada série das taxas de mortalidade. Em vez de procurar uma solução única que atenda a todos os cenários, essa estratégia reconhece a singularidade de cada população e busca o melhor modelo não linear para cada conjunto de dados. Ao fazer isso, a modelagem torna-se não apenas mais adaptável, mas também potencialmente mais precisa, garantindo que as estratégias mais apropriadas sejam empregadas para cada conjunto de dados, otimizando assim a acurácia preditiva.

Nesse contexto, levanta-se as hipóteses de pesquisa:

- Sistemas híbridos que combinam modelo estatístico linear na série temporal e modelo de AM não-linear na modelagem residual apresentam acurácia de previsão superior em relação às demais abordagens da literatura na previsão de taxas de mortalidade em vários passos à frente.
- A seleção de modelos não lineares sobre resíduos de preditores lineares podem melhorar a acurácia de previsões dos sistemas híbridos na previsão das taxas de mortalidade.

### 1.3 ESTRUTURA DO TRABALHO

A estrutura do restante do trabalho é composta por 5 capítulos descritos a seguir :

**Capítulo 2 - Revisão da literatura:** neste capítulo são apresentados os conceitos e as técnicas utilizadas na pesquisa, assim como os principais trabalhos relacionados.

**Capítulo 3 - Sistema proposto:** neste capítulo são apresentados os elementos constitutivos do sistema híbrido proposto objeto de estudo deste trabalho.

**Capítulo 4 - Protocolo experimental:** neste capítulo são apresentadas as etapas e metodologia dos experimentos realizados.

**Capítulo 5 - Resultados:** neste capítulo são apresentados os resultados obtidos dos experimentos realizados.

**Capítulo 6 - Conclusão:** neste capítulo são apresentadas as considerações finais, as limitações da pesquisa e os trabalhos futuros.

## 2 REVISÃO DA LITERATURA

Neste capítulo são apresentados os conceitos, as técnicas utilizadas, e os principais trabalhos relacionados à pesquisa. A Seção 2.1 apresenta a definição, as características da projeção da mortalidade, os modelos e as etapas para realizar a previsão de mortalidade. A Seção 2.1.1 apresenta os principais modelos para projetar a mortalidade. Na Seção 2.2 é apresentado o modelo *Autoregressive Integrated Moving Average* (ARIMA). Na Seção 2.3 é apresentado o modelo Perceptron multicamadas (MLP). Na Seção 2.4 é apresentado o modelo LSTM. Na Seção 2.5 é apresentado o modelo N-BEATS. Na Seção 2.6 são apresentados estudos que investigaram previsão em séries temporais curtas. Na Seção 2.7 são explicadas as principais abordagens de previsão de múltiplos passos à frente. Na Seção 2.8 o conceito e etapas de um Sistema híbrido para previsão de séries temporais. E por fim, na Seção 2.9 são apresentadas as considerações finais do capítulo.

### 2.1 PREVISÃO DE MORTALIDADE

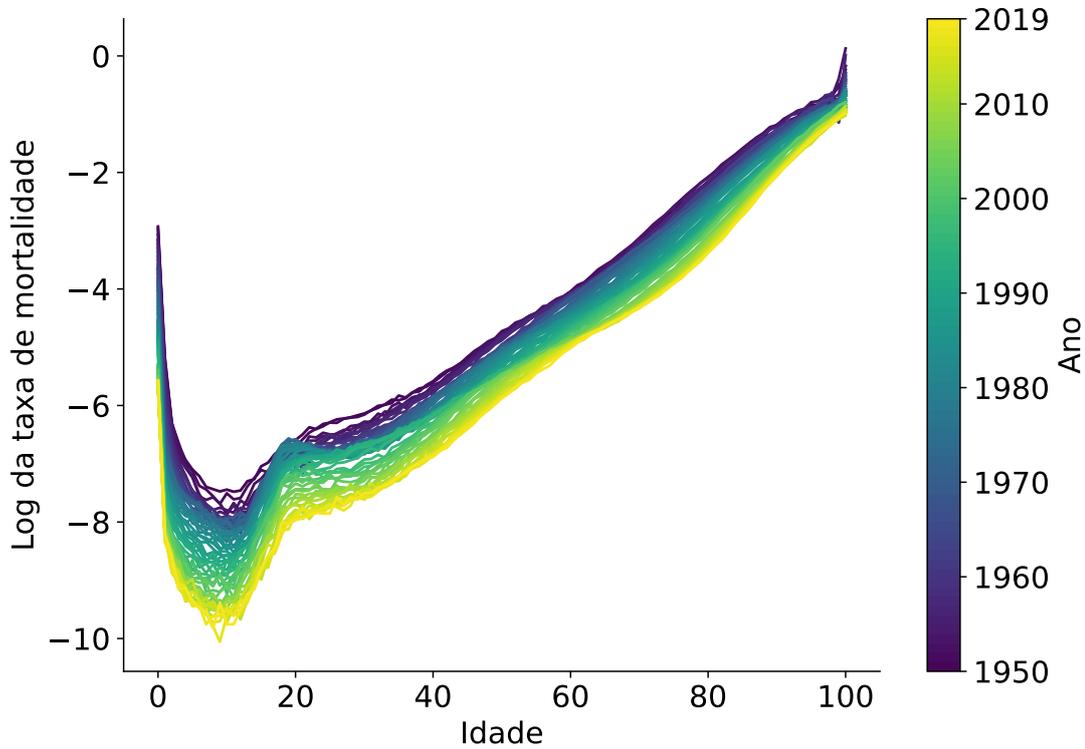
A previsão de mortalidade é usualmente o termo utilizado para representar o processo de prever as taxas de mortalidade humana. A taxa de mortalidade, denotada aqui como  $m_{x,t}$  ou  ${}_t m_x$ , representa a taxa central de morte para a idade  $x$  no ano  $t$ , ou seja, ela quantifica a quantidade de óbitos para uma determinada idade  $x$  em um período específico  $t$ . Segundo Bravo (2007), a taxa central de mortalidade é representada matematicamente por

$${}_t m_x = \frac{{}_t d_x}{{}_t E_x}, \quad (2.1)$$

em que  ${}_t d_x$  e  ${}_t E_x$  são o número de óbitos registrados e a quantidade de pessoas expostas ao risco para a idade  $x$  no ano  $t$ , respectivamente. A partir das taxas de mortalidade é possível construir as tábuas de mortalidade que, segundo Bravo (2007), é um instrumento tabular de análise demográfica o qual sintetiza uma coleção de funções que dão subsídios para se analisar a longevidade (i.e. expectativa de vida) e quantificar as probabilidades de falecimento e sobrevivência, por faixas etárias, de uma população. Essas tábuas são um dos pilares para a determinação dos preços e quantificação de riscos dos seguros de vida, e para o gerenciamento financeiro e atuarial dos fundos de pensões.

A título de exemplo, a Figura 1 ilustra a curva de mortalidade na escala logarítmica

Figura 1 – Logaritmo natural das taxas de mortalidade da população francesa para as idades 0 a 100 anos entre os anos 1950 e 2019.



Fonte: Elaborado pelo autor (2024).

da população francesa para os anos 1950 a 2019, a qual mostra a redução das taxas de mortalidade ao passo que os anos evoluem. Além de ser comum a diversos países, essa redução das taxas de mortalidade foi mais acentuada nas idades mais jovens, constituindo como um dos principais fatores para o aumento da expectativa de vida populacional (Aburto et al., 2020).

Prever a mortalidade humana é uma tarefa comumente empregada em diversos campos para atender a diversas finalidades específicas. Como exemplo dos mais comuns podem ser citados: atuarial (Olivieri, 2001; Booth; Maindonald; Smith, 2002; Mitchell et al., 2013; Janssen, 2018), demográfico (Janssen, 2018; Shi, 2021) e de saúde (Janssen, 2018).

Nesse sentido, busca-se realizar previsões acuradas sobre taxas de mortalidade humana tanto para curto quanto para o longo prazo; este que pode se estender por mais de 10 anos futuros. Os valores previstos são usados por entidades que gerenciam planos previdenciários e por instituições governamentais para gerenciar sistemas de seguridade social. É comum realizar previsões para 20 ou mais anos futuros tendo em vista a necessidade de avaliar e gerenciar riscos tanto quanto tarifar esses produtos, combinando-os com modelos financeiros que levam em conta o valor do dinheiro no tempo. Portanto, é fundamental que se selecione premissas adequadas sobre essas previsões para que seja possível desenvolver sistemas de seguridade e

previdenciários cada vez mais eficientes e com menor exposição a riscos de subscrição, ou seja, de assunção dos riscos assumidos (Olivieri, 2001). Seguindo essa linha de pensamento, busca-se projetar as taxas de mortalidade para todas as idades que compõem a população para uma determinada região; o que cobre, por sua vez, todas as idades cientificamente válidas, o que de forma agregada forma a curva de mortalidade humana (como ilustrado na Figura 1).

De uma forma geral, conforme Bell (1997), há duas formas de resolver o problema de previsão de taxas de mortalidade humana. A primeira e mais utilizada pela literatura diz respeito à modelagem multivariada de decomposição fatorial que é combinada com o emprego de um modelo linear estatístico extrapolativo para prever o primeiro e principal componente (Feng; Shi, 2018). Nesse sentido, é uma abordagem que simplifica o processo de especificação de modelos para cada série de taxa de mortalidade para as idades.

Por outro lado, a segunda abordagem, denominada univariada se baseia no emprego de modelos de previsão de séries temporais os quais são usados para cada idade de série de taxa de mortalidade (Bell, 1997; Feng; Shi, 2018). Após a realização do processo de previsão, é possível usar uma técnica para interpolar e preencher as previsões para idades omitidas caso seja necessário.

### **2.1.1 Modelos para previsão de mortalidade**

Os modelos de previsão de mortalidade usualmente são classificados em duas abordagens (Feng; Shi, 2018): multivariada e univariada. Essas abordagens vão ser detalhadas nas Seções 2.1.1.1 e 2.1.1.2, respectivamente.

#### *2.1.1.1 Modelos multivariados*

A abordagem de previsão de mortalidade multivariada é a mais comum segundo a literatura (Feng; Shi, 2018; Hunt; Blake, 2021). Na maioria dos casos (Booth; Maindonald; Smith, 2002; Feng; Shi, 2018; Hunt; Blake, 2021), ela tem como premissa um modelo demográfico o qual se respalda na determinação de fatores extraídos dos dados multivariados pautados em fundamentações demográficas, isto é, nas faixas-etárias das taxas de mortalidade, além de demais variáveis como diferenças em gênero, variáveis de controle como a presença de grupo de coorte entre outras.

Sob essa abordagem, os modelos mais estudados pela literatura seguem, portanto, a es-

estrutura denominada de *Age-Period-Cohort* (APC), da qual se desdobram diversos modelos de previsão de mortalidade (Hunt; Blake, 2021). A estrutura desse modelo é, segundo Hunt e Blake (2021), formalizada a partir da equação 2.2 a seguir,

$$\eta_{x,t} = \alpha_x + \sum_{i=1}^N \beta_x^{(i)} k_t^{(i)} + \beta_x^{(0)} \gamma_{t-x}, \quad (2.2)$$

em que  $x$  é a idade;  $t$  é o ano;  $\eta_{x,t}$  é a função de ligação que transforma a variável resposta (a qual pode ser alguma medida de mortalidade) em uma forma adequada para a modelagem e que a relaciona à estrutura proposta das preditoras;  $\alpha_x$  é a função estática da idade que captura o formato genérico da mortalidade entre todas as idades; o conjunto de  $N$  termos  $\beta_x^{(i)} k_t^{(i)}$  consiste na função do período  $k_t^{(i)}$  a qual determina a evolução das taxas de mortalidade ao longo do tempo, e na função da idade  $\beta_x^{(i)}$  que determina o padrão da variação da mortalidade entre as idades; o termo  $\beta_x^{(0)} \gamma_{t-x}$  é composto pelo efeito da coorte  $\gamma_{t-x}$ , o qual determina o efeito ao longo da vida específico de cada geração, denotado pela idade de nascimento, e a função da idade  $\beta_x^{(0)}$ , a qual varia o termo da coorte.

Sob essa estrutura, a literatura de previsão de mortalidade (Hunt; Blake, 2021) classifica os modelos multivariados em três famílias: LC (Lee; Carter, 1992) e seus derivados; Cairns-Blake-Dowd (CBD) (Cairns; Blake; Dowd, 2006) e extensões; e o modelo clássico APC (Hobcraft; Menken; Preston, 1982) e hibridização deste com as duas outras famílias.

O processo de determinação de um modelo baseado em fatores passa pela extração de um ou mais componentes principais dos dados os quais podem representar alguns comportamentos agregados, tais como a tendência temporal das taxas de mortalidade conforme a estrutura APC supracitada. Nesse sentido, é possível empregar um modelo linear extrapolativo para realizar as previsões desse componente. Essa fundamentação foi o ponto de partida para elaboração do modelo LC (Lee; Carter, 1992), o qual utiliza taxas de mortalidade em formato matricial onde as linhas contém as idades e as colunas representam os anos (tempo). O modelo LC se representa por meio da equação 2.3

$$\ln(\mathbf{m}_{x,t}) = \mathbf{a}_x + \beta_x \mathbf{k}_t + \varepsilon_{x,t}, \quad (2.3)$$

em que  $\mathbf{m}_{x,t}$  é a taxa central de mortalidade para a idade  $x$  no tempo  $t$ ,  $\mathbf{a}_x$  é a média de  $\ln(\mathbf{m}_x)$  para cada idade  $x$  durante todo o período de treinamento,  $\beta_x$  é o coeficiente que representa a velocidade de variação da taxa de mortalidade em resposta às variações de  $\mathbf{k}_t$ , este que, por sua vez, é o fator que captura a tendência linear das taxas de mortalidade. E

$\varepsilon_{x,t}$  é o termo de erro aleatório que foge ao modelo.

Após o modelo LC, estudos empregaram avanços na modelagem no sentido de melhorar o seu desempenho. Lee e Miller (2001), Booth, Maindonald e Smith (2002), Hyndman e Ullah (2007) implementaram abordagens diferentes mas com a fundamentação de que utilizar o conjunto de idades e mais variáveis poderia produzir resultados mais acurados. Outro modelo de previsão de mortalidade bastante discutido na literatura (Hunt; Blake, 2021) é o Cairns-Blake-Dowd (CBD) (Cairns; Blake; Dowd, 2006), o qual foi pensado para contornar o problema de as taxas de mortalidade projetadas estarem correlacionadas perfeitamente em modelos de termo de idade/período único. O modelo CBD é usualmente definido na forma da equação 2.4,

$$\text{logit}(q_{x,t}) = k_t^{(1)} + (x - \bar{x})k_t^{(2)}, \quad (2.4)$$

em que  $q_{x,t}$  é a probabilidade de falecimento da idade  $x$  durante o período  $t$ <sup>1</sup>;  $\bar{x}$  é a idade média no intervalo de dados utilizado;  $k_t^{(i)}$  o termo do período; e o uso da função de ligação logit.

A terceira família de modelos de mortalidade previamente discutida se respalda no modelo clássico APC de Hobcraft, Menken e Preston (1982). A forma original do modelo clássico APC (Hobcraft; Menken; Preston, 1982) para modelar as taxas de mortalidade segundo (Hunt; Blake, 2021) é dada como

$$\ln(\mu_{x,t}) = \alpha_x + k_t + \gamma_{t-x}, \quad (2.5)$$

em que  $\alpha_x$  representa o termo estático que captura o formato genérico da mortalidade entre todas as idades;  $k_t$  é o termo do período;  $\gamma_{t-x}$  é o efeito da coorte; e  $\mu_{x,t}$  representa a força de mortalidade (taxa instantânea de mortalidade) que é equivalente à taxa central de mortalidade  $m_{x,t}$ , quando se assume a hipótese de força de mortalidade constante em cada intervalo de tempo unitário (Bravo, 2007)<sup>2</sup>.

Segundo (Hunt; Blake, 2021), esse modelo clássico só foi introduzido na previsão de mortalidade no campo atuarial apenas com os trabalhos de Willets (2004), o qual fez uma análise do impacto da coorte na mortalidade. Após diversas extensões desses modelos, Plat (2009) propôs

<sup>1</sup> A probabilidade de falecimento, assumindo a hipótese de força de mortalidade constante é denotada por  $q_{x,t} = 1 - \exp(-\mu_{x,t})$  (Bravo, 2007; Hunt; Blake, 2021).

<sup>2</sup> O detalhamento matemático pode ser encontrado em Bravo (2007) e Hunt e Blake (2021).

um modelo que combina essa diversas abordagens multivariadas. A sua fórmula matemática está descrita a seguir,

$$\ln(\mu_{x,t}) = \alpha_x + k_t^{(1)} + k_t^{(2)}(\bar{x} - x) + k_t^{(3)}(\bar{x} - x)^+ + \gamma_{t-x}, \quad (2.6)$$

em que  $z^+ \equiv \max(z, 0)$ . Além disso, Plat (2009) também apresentou a versão simplificada desse modelo que se mostrou mais adequada para modelar idades mais elevadas. A versão simplificada não possui o termo  $(\bar{x} - x)^+$ .

Evidências apontam para o fato de eventos sociais, econômicos e mudanças climáticas impactarem as taxas de mortalidade humana (Makridakis; Hyndman; Petropoulos, 2020; Hunt; Blake, 2021). Entende-se essas variáveis como fatores de risco, pois suas variações impactam as taxas de mortalidade humana. Além dessas, existem os fatores associados à saúde humana tal como a presença de doenças como câncer, e também os hábitos de vida como tabagismo. Contudo, essas variáveis estão associadas a sistemas complexos que são difíceis de serem modelados, ao mesmo tempo em que a produção da previsão desses dados, etapa essencial na utilização de modelos multivariados, é mais uma etapa difícil e desafiadora que pode aumentar o erro do modelo de mortalidade.

Segundo Booth, Maindonald e Smith (2002), os possíveis efeitos positivos da utilização desse tipo de informação, isto é modelar adequadamente a mortalidade usando os fatores associados às variações das taxas de mortalidade, dificilmente serviriam para prevê-la em longo prazo, tendo em vista a falta de confiabilidade na obtenção das informações projetadas, tal como da dificuldade de observar correlação de longo prazo entre as taxas de mortalidade e os fatores de riscos supracitados.

No que diz respeito à utilização do aprendizado de máquina para previsão de mortalidade de forma multivariada, alguns estudos empregaram a combinação de técnicas. Richman e Wüthrich (2021) se sustentaram no aprendizado por representação <sup>3</sup>, tomando por base o modelo LC para desenvolver um Perceptron multicamadas MLP para múltiplas populações, o qual incluiu características como país, gênero e idade, e mostrou resultados competitivos frente às alternativas. Sob uma perspectiva semelhante, Hainaut (2018) desenvolveu uma modelagem híbrida que combinou uma rede neural artificial, um algoritmo genético <sup>4</sup> e o

<sup>3</sup> O aprendizado de representação é um conjunto de técnicas que visa aprender automaticamente a representação apropriada para o modelo (Bengio; Courville; Vincent, 2013).

<sup>4</sup> Os algoritmos genéticos são métodos da computação evolucionária (Eiben; Smith, 2015), inspirados no processo de evolução biológica, que podem ser utilizados para encontrar soluções de qualidade em problemas complexos (Mitchell et al., 1997). Para uma explicação mais detalhada sobre computação evolucionária e algoritmos genéticos consulte Eiben e Smith (2015) e Mitchell et al. (1997).

modelo de passeio aleatório. A rede neural foi usada para obter os componentes principais das taxas de mortalidades das idades entre 20 e 109 anos da França, Reino Unido e Estados Unidos, ao passo que o segundo foi empregado em conjunto com o algoritmo do gradiente descendente para otimizar os parâmetros daquela. Em seguida, o modelo de passeio aleatório gerou as previsões dos componentes principais. Numericamente, a proposta de Hainaut (2018) foi superior em termos de acurácia de previsão ao modelo LC. Nigri et al. (2019), por outro lado, utilizou uma rede neural artificial LSTM para prever a série temporal da tendência temporal extraída da fatoração do algoritmo do modelo LC. Os seus resultados mostraram que a rede LSTM obteve desempenho preditivo superior frente ao modelo ARIMA, que é comumente utilizado pela modelagem tradicional.

Posteriormente, observaram-se progressos no uso de modelos de aprendizado profundo para a previsão de taxas de mortalidade. Dentre os avanços, Perla et al. (2021) utilizaram a aprendizagem profunda combinando LSTM e redes neurais convolucionais (*Convolutional Neural Networks* - CNN)<sup>5</sup> num modelo global para prever as taxas de mortalidade de vários países. Eles mostraram que a sua abordagem obteve um desempenho preditivo superior ao modelo Lee-Carter em vários conjuntos de dados. Empregando redes neurais recorrentes aos dados da população portuguesa, Bravo (2021b) observou que a rede LSTM superou os modelos Lee-Carter e APC para ambos os sexos. Chen e Khaliq (2022) aplicaram o Lee-Carter, o Gated Recurrent Unit (GRU)<sup>6</sup>, o LSTM e o *Bidirectional*-LSTM à previsão da taxa de mortalidade das nove divisões censitárias da população dos EUA e verificaram que o Gated Recurrent Unit (GRU) apresentou a melhor acurácia de previsão. Recentemente, o *Transformer*<sup>7</sup> com o mecanismo de autoatenção, reconhecido por sua capacidade de processar eficientemente e com precisão linguagem natural, foi utilizado para prever taxas de mortalidade (Roshani; Izadi; Khaledi, 2022; Wang et al., 2024). Roshani, Izadi e Khaledi (2022) conduziram um experimento com onze países e ambos os gêneros. Eles observaram que o *Transformer* melhorou o desempenho de previsão em comparação com o ARIMA, o ETS e o LSTM em nove populações masculinas, mas na população feminina, o desempenho foi equivalente ao dos outros mode-

<sup>5</sup> As redes neurais convolucionais foram projetadas para processar dados estruturados em múltiplas *arrays* como por exemplo imagens coloridas. Eficazes em aprender automaticamente características espaciais hierárquicas, tornaram-se dominantes em diversas tarefas em visão computacional (LeCun; Bengio; Hinton, 2015).

<sup>6</sup> Gated Recurrent Unit é uma rede neural recorrente inspirada na rede LSTM (Cho et al., 2020), com uma arquitetura mais simples e resultados competitivos em processamento de linguagem natural e previsão de séries temporais (Cho et al., 2020; Yamak; Yujian; Gadosey, 2019).

<sup>7</sup> *Transformer* é uma arquitetura de rede neural profunda baseada no mecanismo atenção que aprende o contexto em dados sequenciais (Vaswani et al., 2017), e vem trazendo avanços significativos no processamento de linguagem natural.

los. Wang et al. (2024) investigaram o desempenho de várias arquiteturas de redes neurais artificiais na previsão da taxa de mortalidade de oito países e concluíram que o *Transformer* superou a CNN, o LSTM, a Rede Neural Recorrente (RNN) e o modelo Lee-Carter.

#### 2.1.1.2 Modelos univariados

A literatura apresenta diversos modelos univariados para prever taxas de mortalidade (Feng; Shi, 2018). Dentre eles podem ser citados estudos que propuseram e avaliaram a capacidade preditiva das técnicas ARIMA (McNown; Rogers, 1989), suavização exponencial (Feng; Shi, 2018), e Processos Gaussianos (Wu; Wang, 2018).

Essa abordagem parte da ideia de modelar cada taxa de mortalidade individualmente por um modelo de previsão e, em seguida, realizar a extrapolação para prever as taxas de mortalidade para todas as idades e diversos passos à frente. Os argumentos que fundamentam a escolha dessa abordagem em detrimento da multivariada se baseiam primordialmente na distribuição dos dados das taxas de mortalidade, pois para que a modelagem multivariada seja preterida seria necessário que as taxas de mortalidade apresentassem padrões similares ao longo do tempo entre as diversas idades (Feng; Shi, 2018).

Entretanto, isso não se verifica como foi observado por Wu e Wang (2018) em que as taxas de mortalidade entre idades vizinhas pode denotar fraca correlação cruzada entre si. Sendo assim, haveria necessidade de se modelar separadamente o comportamento das séries temporais das taxas de mortalidade para as idades.

O modelo proposto por Wu e Wang (2018) para prever as taxas de mortalidade da população francesa utilizou Processos Gaussianos. Eles testaram diversas funções de média e covariância, e propuseram duas especificamente, na qual a primeira é uma média ponderada, enquanto a de covariância é uma mistura espectral. O modelo proposto demonstrou desempenho superior no longo prazo, em termos da raiz quadrada do erro quadrático médio, em relação aos modelos clássicos da literatura como Lee e Miller (2001), Booth-Maindonald-Smith (BMS) (Booth; Maindonald; Smith, 2002), and *Functional Demographic Model* (FDM) (Hyndman; Ullah, 2007).

Diante das diversas abordagens empregadas para prever a mortalidade humana, Feng e Shi (2018) observaram que a perspectiva univariada de previsão de séries temporais produziria melhores resultados sobre a modelagem multivariada apontada na seção anterior, tais como Lee-Carter e suas extensões. A explicação dos autores partiu da observação empírica do fato

da correlação cruzada das taxas de mortalidade para as diversas idades demonstrarem insignificância estatística. Sendo assim, observaram comportamento heterogêneo entre as séries das taxas de mortalidade para diversos grupos etários.

Seguindo essa perspectiva, recentemente, Duarte, de Mattos Neto e Firmino (2024) propuseram um sistema híbrido (linear e não-linear), chamado HyS-MF, de abordagens recursiva e direta para prever taxas de mortalidade em vários passos à frente da população francesa. Utilizando ARIMA para modelar a série e N-BEATS para aprender os padrões e realizar previsões da série residual na forma direta, eles observaram que o HyS-MF auferiu o melhor desempenho de previsão em comparação a diversos modelos da estado da arte da literatura de previsão de séries temporais.

Tendo em vista os trabalhos discutidos, as evidências empíricas indicam que não existe um único método de previsão que apresente desempenho consistentemente superior em todos os conjuntos de dados e horizontes temporais. As abordagens principais para realizar previsões de taxas de mortalidade são multivariada e univariada. Embora menos popular, a abordagem univariada destaca-se por permitir a modelagem independente das dinâmicas de cada série temporal. Adicionalmente, tanto modelos estatísticos lineares quanto não lineares, incluindo os de Aprendizado de Máquina (AM), são empregados em ambas as abordagens. Os modelos de AM frequentemente demonstram maior acurácia de previsão, superando os modelos estatísticos tradicionais. Contudo, uma desvantagem dos modelos de AM é a tendência ao sobreajuste, o que pode comprometer sua capacidade de generalização para novos dados. Nesse sentido, amparando-se na abordagem univariada, a proposta desta tese opta por combinar a robustez e parcimônia do modelo estatístico linear com a adaptabilidade e acurácia de previsão aprimorada dos modelos não-lineares de AM para desenvolver um modelo de previsão de taxas de mortalidade para múltiplos passos à frente.

## 2.2 AUTORREGRESSIVO INTEGRADO DE MÉDIAS MÓVEIS

O formalismo autorregressivo Integrado de Médias Móveis (*Autoregressive Integrated Moving Average* - ARIMA) é um modelo estatístico linear utilizado no processo de análise e previsão de séries temporais. Esse modelo se fundamenta em uma formulação matemática na qual as observações e os termos de erro se relacionam de forma linear e autorregressiva ao longo do tempo.

Para utilizar o modelo ARIMA, aplica-se a metodologia Box e Jenkins (Box et al., 2015)

com a técnica de máxima verossimilhança <sup>8</sup> para estimar os parâmetros do modelo a partir dos dados observados. Isto posto, essa metodologia emprega um processo iterativo cíclico para elaboração do modelo, o qual, segundo Morettin e Toloí (2006), compreende quatro estágios. São eles:

1. Especificação: Uma classe geral de modelos é levada em consideração para a análise.
2. Identificação: Determinação dos valores apropriados para os componentes autorregressivos, de diferenciação e de médias móveis.
3. Estimação: Obtenção dos parâmetros do modelo identificado.
4. Verificação ou diagnóstico: Avaliação estatística dos resíduos se o modelo estimado é adequado para o seu fim.

Esse modelo é uma combinação de dois modelos chamados autorregressivo (*autoregressive* - AR) e média móveis (*moving average* - MA) dos termos de erros. Ainda, ele incorpora o componente de integração (*integrated* - I), que representa a operação inversa da diferença. Esta última é aplicada para remover tendências e promover a estacionariedade da série temporal, sendo um passo necessário na preparação dos dados para a modelagem ARIMA. De acordo com Morettin e Toloí (2006), o operador diferença  $\Delta$  aplicado a uma série temporal  $z_t$  pode ser descrito a partir da equação a seguir

$$\Delta z_t = z_t - z_{t-1}. \quad (2.7)$$

Além disso, esse modelo pode ser denotado como ARIMA( $p, d, q$ ), onde  $p$ ,  $d$  e  $q$  representam as ordens autorregressivas, diferenças e de médias móveis, respectivamente. A título de exemplo, conforme Box et al. (2015), o processo Autorregressivo (AR) de ordem  $p$ , representado por AR( $p$ ) é descrito formalmente como:

$$z_t = \phi_1 z_{t-1} + \phi_2 z_{t-2} + \phi_p z_{t-p} + a_t \quad (2.8)$$

no qual os  $p$  parâmetros  $\phi$  são diferentes de zero e  $a_t$  é o termo de erro com média zero ( $E[a_t] = 0$ ) e variância constante ( $Var[a_t] = \sigma_a^2$ ). Outra forma de representar matematicamente o processo AR( $p$ ) se dá pela utilização do operador autorregressivo como se segue:

$$(1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p) z_t = a_t \quad (2.9)$$

<sup>8</sup> A máxima verossimilhança é uma técnica estatística inferencial que visa encontrar estimadores (Casella; Berger, 2002). Nesse sentido, ela busca determinar os valores dos parâmetros que maximizam a função de verossimilhança, ou seja, que tornam os dados observados os mais prováveis possíveis sob o modelo assumido (Hyndman; Athanasopoulos, 2018).

ou

$$\phi(B)z_t = a_t \quad (2.10)$$

onde  $B$  é o operador de translação para o passado (e.g.  $Bz_t = z_{t-1}$ ,  $B^p z_t = z_{t-p}$ ,  $B = 1 - \Delta$ ).

Por outro lado, o processo de média móvel de ordem  $q$ ,  $MA(q)$ , pode ser escrito como:

$$z_t = a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \dots - \theta_q a_{t-q}, \quad (2.11)$$

no qual os  $q$  parâmetros  $\theta$  são diferentes de zero. Utilizando o operador de translação para o passado,  $B$ , o modelo  $MA(q)$  se equivale à equação

$$z_t = (1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q) a_t \quad (2.12)$$

ou

$$z_t = \theta(B) a_t. \quad (2.13)$$

Segundo Box et al. (2015), o modelo ARIMA é uma generalização do modelo *Autoregressive Moving Average* (ARMA), com o emprego do operador diferença, e pode ser descrito matematicamente a partir da Equação (2.14)

$$\varphi(B)z_t = \phi(B)\Delta^d z_t = \theta_0 + \theta(B)a_t, \quad (2.14)$$

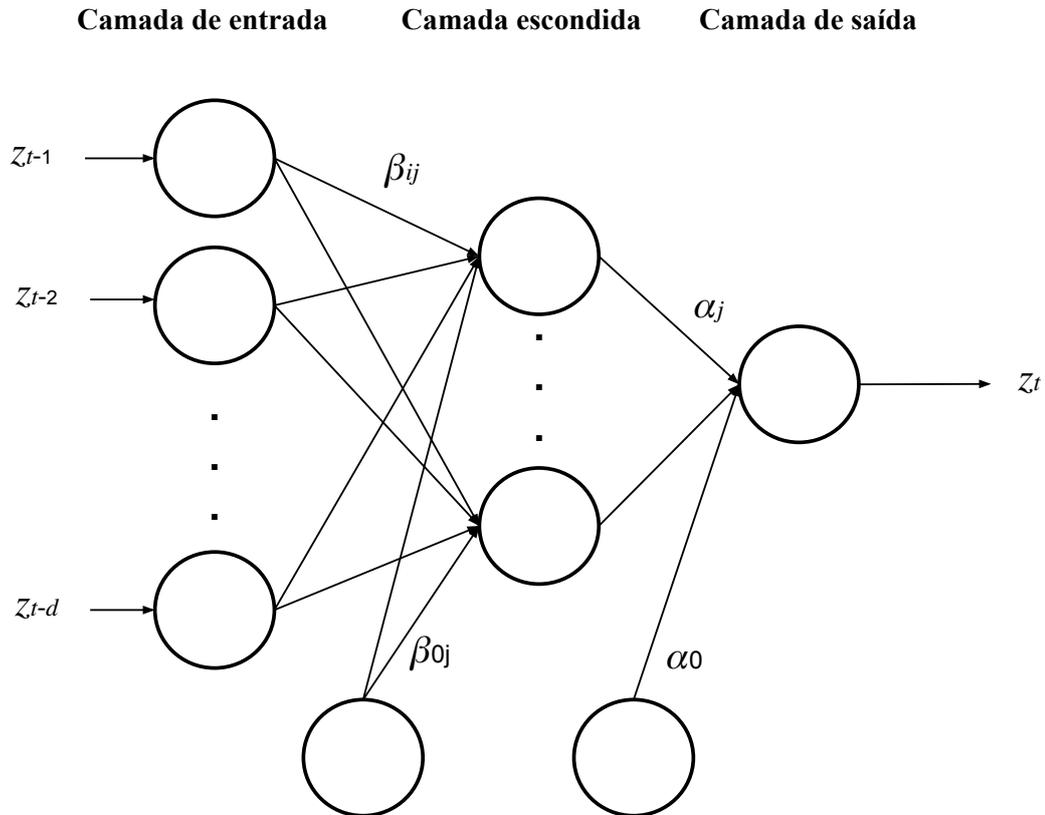
em que  $z_t$  representa a série temporal,  $a_t$  o termo de erro,  $\phi(B)$  o operador autorregressivo estacionário,  $\varphi(B)$  o operador autorregressivo generalizado com diferença de ordem  $d$ , e  $\theta(B)$  o operador de média móvel. Quando  $d = 0$ , o modelo pode ser representado por um processo estacionário ARMA.

A integração desses três componentes resulta em um modelo capaz de capturar uma ampla gama de dinâmicas temporais lineares, tornando o ARIMA uma ferramenta parcimoniosa e versátil na análise e previsão de séries temporais em diversos campos (Morettin; Tolo, 2006).

### 2.3 PERCEPTRON MULTICAMADAS

O Perceptron de multicamadas (MLP) é um modelo de Redes Neurais Artificiais (RNAs) que possui diversas unidades de processamento de informação, também chamadas de neurônios, distribuídas entre as camadas de entrada, intermediária(s) e saída. O processamento de informação da rede MLP se dá por meio da *feedforward* do sinal, ou seja, da propagação da informação da camada de entrada, passando pela(s) camada(s) intermediária(s) (ou escondida(s)) e finalizando na camada de saída (Haykin, 2009).

Figura 2 – Arquitetura do modelo Perceptron multicamadas (MLP)



Fonte: Elaborado pelo autor (2024).

A representação matemática de uma rede MLP aplicada ao problema de séries temporais segundo SANTOS JÚNIOR (2022) com apenas uma camada intermediária está denotada na Equação 2.15,

$$z_t = \alpha_0 + \sum_{j=1}^q \alpha_j g \left( \beta_{0j} + \sum_{i=1}^d \beta_{ij} z_{t-i} \right) + \varepsilon_t, \quad (2.15)$$

em que  $g(\cdot)$  é a função de ativação não-linear,  $\alpha$  e  $\beta$  são os pesos da rede,  $q$  a quantidade de neurônios na camada de saída,  $d$  a quantidade de defasagens temporais (*lags*) ou quantidade de neurônios na camada de entrada, e  $\varepsilon$  o termo de erro.

Nessa perspectiva, a camada de entrada recebe as informações defasadas da série temporal  $\{z_1, z_2, \dots, z_t\}$ . A ordem de defasagem denotada por  $d$ , determina o número de observações passadas que serão consideradas como entradas para a rede. Os valores defasados da série fornecem o contexto temporal necessário para a modelagem da série. As informações da camada de entrada são propagadas para a camada oculta, que consiste em  $q$  neurônios. Cada neurônio da camada oculta aplica uma função de ativação não-linear  $g(\cdot)$  à soma ponderada de suas entradas. A escolha da função de ativação (por exemplo, sigmoide, tangente hiperbólica

ou ReLU) pode influenciar o desempenho do modelo. A informação processada pela camada oculta é então transmitida à camada de saída. No contexto da Equação 2.15, esta camada aplica uma transformação linear aos sinais recebidos da camada oculta. A saída desta camada representa a previsão do modelo para o próximo valor da série temporal.

Essa arquitetura possibilita a captura de padrões temporais complexos e não-lineares intrínsecos à série temporal  $\{z_1, z_2, \dots, z_t\}$ . Esta estrutura, composta por camadas interconectadas por funções de ativação não-lineares, confere ao modelo a capacidade de representar e prever uma diversa gama de dinâmicas temporais (Zhang, 2003).

O aprendizado do modelo de redes neurais pode ser realizado por meio de algoritmo de retropropagação em conjunto com a técnica de otimização do gradiente descendente (Goodfellow; Bengio; Courville, 2016). Este mecanismo de aprendizado visa identificar e assimilar padrões no conjunto de dados por meio do ajuste iterativo dos pesos sinápticos. De acordo com Haykin (2009), o algoritmo opera da seguinte forma:

1. Inicialização: Os pesos do modelo são gerados aleatoriamente.
2. Propagação direta: Os dados de entrada são processados através das camadas da rede.
3. Cálculo do erro: A discrepância entre a saída da rede e o valor alvo é quantificada.
4. Retropropagação: O sinal de erro é propagado da camada de saída em direção à camada de entrada.
5. Ajuste dos pesos: O algoritmo do gradiente descendente utiliza as informações do erro propagado para atualizar os pesos da rede, minimizando a função de custo.
6. Iteração: Iterar do passo 2 ao 5 até que alguma condição de parada seja atendida.

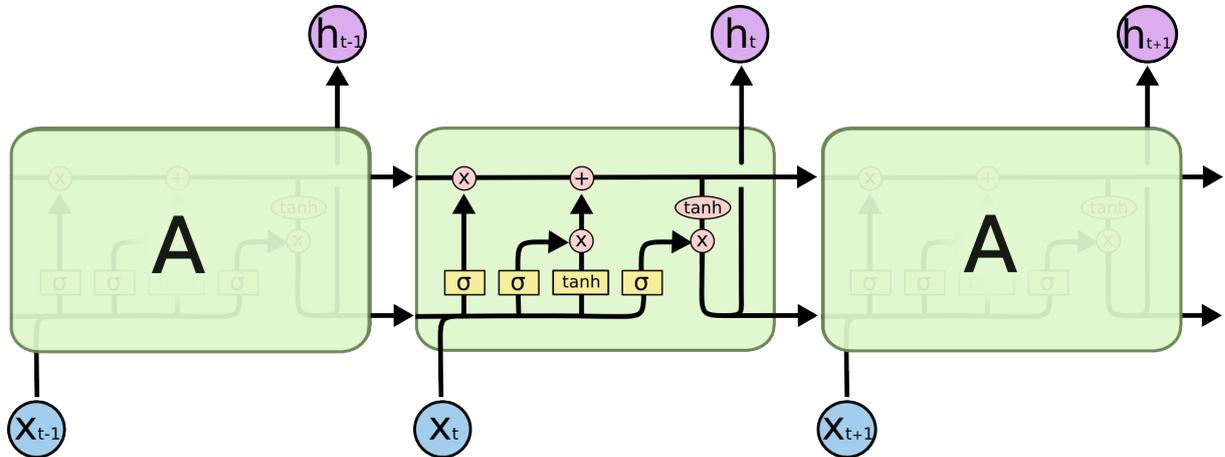
Este processo iterativo de ajuste permite que a rede refine progressivamente sua capacidade preditiva (Haykin, 2009), adaptando-se às características específicas da série temporal em análise.

## 2.4 LONG-SHORT TERM MEMORY

A rede neural *Long-short term memory* (LSTM) possui uma arquitetura de recorrência entre as conexões que visa capturar dependências temporais de curto e longo prazos de uma série

temporal (Hochreiter; Schmidhuber, 1997). A estrutura dessa rede pode ser melhor visualizada a partir da Figura 3 a seguir:

Figura 3 – Arquitetura do modelo LSTM.



Fonte: Olah (2015).

Matematicamente, considerando uma série temporal denotada por  $x_t$ , tem-se as equações de uma rede neural recorrente LSTM a seguir:

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f), \quad (2.16)$$

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i), \quad (2.17)$$

$$\tilde{C}_t = \tanh(W_C \cdot [h_{t-1}, x_t] + b_C), \quad (2.18)$$

$$C_t = f_t \cdot C_{t-1} + i_t \cdot \tilde{C}_t, \quad (2.19)$$

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o), \quad (2.20)$$

$$h_t = o_t \cdot \tanh(C_t), \quad (2.21)$$

em que  $f_t$  é o portão do esquecimento (*forget gate*),  $h_{t-1}$  é o estado oculto,  $i_t$  é o portão de entrada (*input gate*),  $C_t$  é o estado interno da célula de memória,  $\tilde{C}_t$  é o novo candidato à entrada no estado interno da célula de memória,  $o_t$  é o portão de saída (*output gate*) e  $h_t$  é a saída do modelo que também é transmitida para a próxima camada LSTM no tempo  $t + 1$ .

Com essa arquitetura, a rede neural possui um sistema adaptável, capaz de selecionar quais informações devem ser mantidas ao longo do tempo. Pensando de outra maneira, ela dispõe de uma capacidade de aprender o quê deve ser descartado do ponto de vista temporal, ou seja, o que deve ser taxado como irrelevante (Zhang et al., 2021). O elemento fundamental que permite a realização desse filtro é o portão (*gate*), o qual, por meio da função de ativação

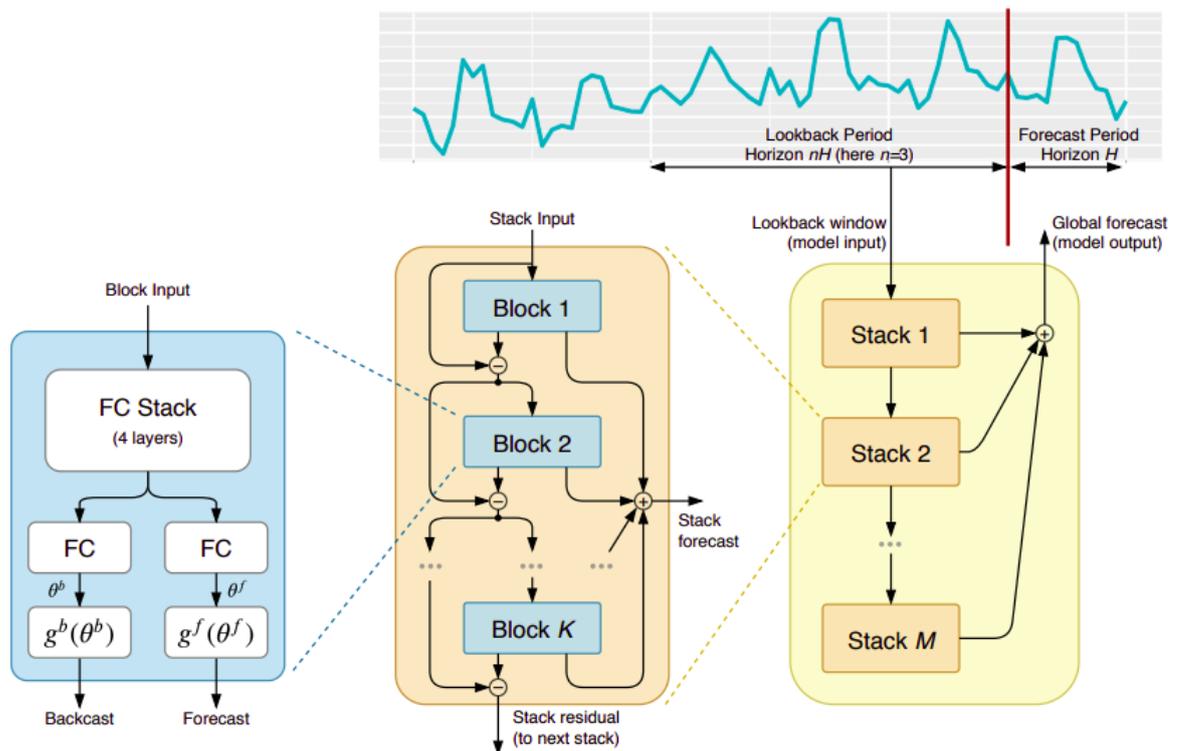
sigmoide, atribui o esquecimento ou a permanência do sinal dependendo resultado da saída do cálculo dessa função (i.e., resulta num valor entre 0 e 1, que é entendido como um espectro de quanto menor o valor, menos relevante é a informação).

Isso posto, a rede neural LSTM se mostra como uma das principais escolhas para a previsão de séries temporais em diversas áreas do conhecimento (Ahmed; Hassan; Mstafa, 2022), e também vem demonstrando resultados interessantes na previsão de taxas de mortalidade (Nigri et al., 2019; Perla et al., 2021; Bravo, 2021b; Chen; Khaliq, 2022; Lindholm; Palmborg, 2022).

## 2.5 N-BEATS

O modelo *Neural basis expansion analysis for interpretable time series forecasting* (N-BEATS) (Oreshkin et al., 2019) está enquadrado como um modelo de RNAs de arquitetura profunda, que objetiva ser simples e interpretável. A simplicidade surge do fato de ele não depender de características ou variáveis específicas de séries temporais, mas apenas é compreendido como uma regressão com sucessivas camadas escondidas totalmente conectadas por uma função não-linear, e dispendo de conexões residuais entre os blocos.

Figura 4 – Arquitetura do modelo N-BEATS.



Fonte: Oreshkin et al. (2019).

A arquitetura do modelo N-BEATS pode ser visualizada na Figura 4. O modelo N-BEATS

tem a característica de concatenar diversas camadas ocultas com a função de ativação *Rectified Linear Unit* (ReLU). Desse modo, são empregadas chamadas sucessivas entre esses múltiplos blocos com camadas ocultas com a transmissão da informação dos resíduos dos dados de entrada entre elas. Esses resíduos são calculados a partir dos dados de entrada e da modelagem do *backcast*, isto é, a estimação dos valores passados da série os quais podem ser considerados como os dados de entrada. Sendo assim, cada camada desse modelo possui duas saídas, *backcast* e *forecast*, sendo a primeira utilizada para modelagem dos resíduos passados relacionados aos dados de entradas, ao passo que segunda tem a intenção de realizar as previsões dos horizontes futuros (Oreshkin et al., 2021).

Por fim, todos os valores de saída das camadas (*forecast*) são somados, compondo a saída final que representa a previsão dos horizontes futuros. Segundo (Oreshkin et al., 2021), assume-se que existam  $R$  blocos residuais e que cada um deles possua  $L$  camadas escondidas. Ainda, considerando  $\mathbf{x}$  os dados de entrada da arquitetura, utilizando o sobrescrito para o bloco residual e a camada,  $r$  e  $l$ , respectivamente, e denotando a camada totalmente conectada com os pesos  $\mathbf{W}^{r,l}$  e *bias*  $\mathbf{b}^{r,l}$  como  $\text{FC}_{r,l}(\mathbf{h}^{r,l-1}) = \text{ReLU}(\mathbf{W}^{r,l}\mathbf{h}^{r,l-1} + \mathbf{b}^{r,l})$ , é possível descrever as operações do N-BEATS como sendo,

$$\begin{aligned} \mathbf{x}^r &= \text{ReLU}[\mathbf{x}^{r-1} - \hat{\mathbf{x}}^{r-1}], \\ \mathbf{h}^{r,1} &= \text{FC}_{r,1}(\mathbf{x}^r), \dots, \mathbf{h}^{r,L} = \text{FC}_{r,L}(\mathbf{h}^{r,L-1}), \\ \hat{\mathbf{x}}^r &= \mathbf{B}^r \hat{\mathbf{x}}^{r,L}, \\ \hat{\mathbf{z}}^r &= \mathbf{F}^r \mathbf{h}^{r,L}, \end{aligned} \tag{2.22}$$

em que  $\hat{\mathbf{x}}^0 \equiv 0$ ,  $\mathbf{x}^0 \equiv 0$ , a projeção matricial possua  $\mathbf{B}^r \in \mathbb{R}^{w \times d_h}$ ,  $\mathbf{F}^r \in \mathbb{R}^{H \times d_h}$ , e a previsão final é a soma das previsões para todos os blocos residuais, ou seja,  $\hat{\mathbf{z}} = \sum_r \hat{\mathbf{z}}^r$ .

Quanto ao desempenho de previsão, Oreshkin et al. (2019) mostraram que o modelo N-BEATS demonstrou acurácia preditiva superior a diversos modelos estatísticos e de aprendizado de máquina para os conjuntos de dados da competição M3 (Makridakis; Hibon, 2000) e M4 (Makridakis; Spiliotis; Assimakopoulos, 2020).

## 2.6 PREVISÃO DE SÉRIES TEMPORAIS CURTAS

Uma série temporal curta ou pequena é aquela composta por poucas unidades amostrais ou observações, isto é, que detém um tamanho amostral pequeno. A previsão de séries temporais

com essa característica (i.e. poucas observações históricas) configura-se como um desafio crucial em diversos domínios, como finanças, seguros, saúde, segurança pública etc. O tamanho amostral, a complexidade do modelo e do problema (i.e., grau de ruído contido nos dados) são fatores fundamentais que determinam a acurácia de previsão (Cerqueira; Torgo; Soares, 2022; Cruz-Nájera et al., 2022). Entretanto, não há um consenso sobre qual o tamanho amostral que caracterize uma série temporal como curta ou pequena, porém é possível traçar um paralelo com o que a literatura depreende acerca da acurácia de previsão em relação à quantidade de observações.

Sendo assim, quando os dados são escassos, é mais difícil aprender os padrões, modelar as séries temporais e realizar previsões acuradas (Thomakos et al., 2023). Sobretudo, essa dificuldade se intensifica em conjunto de dados muito pequenos, onde cada observação ganha mais importância em comparação com amostras maiores, podendo levar a previsões viesadas (Weigand; Lange; Rauschenberger, 2021).

Cerqueira, Torgo e Soares (2022) demonstraram a importância do tamanho do conjunto de dados em problemas de previsão de séries temporais. Seus experimentos revelaram que:

- A acurácia dos modelos de Aprendizado de Máquina (AM) melhorou com o aumento do tamanho das séries temporais.
- Modelos estatísticos superaram os modelos de AM quando as séries tinham menos de 60 observações.
- As reduções mais significativas nos erros de previsão ocorreram até 100 observações.
- Após 100 observações, os ganhos de desempenho foram mais graduais, com uma tendência à equalização entre modelos estatísticos e de AM.

Por sua vez, Cruz-Nájera et al. (2022) constataram desempenho de previsão superior dos modelos estatísticos lineares em relação aos modelos de AM não-lineares (e.g., *LightGBM* e *Artificial Neural Networks* (ANN)) em trinta e cinco séries muito curtas de crimes, as quais compreendiam entre quatorze e vinte e uma observações.

Ademais, no que tange ao desempenho de previsão de combinação de modelos em séries temporais curtas, mostrou-se que sistemas híbridos podem produzir resultados superiores aos modelos individuais (e.g., estatísticos e de AM) em previsão de carga (Meng; Han; Hou, 2022) e previsões de taxas de mortalidade (Duarte; de Mattos Neto; Firmino, 2024).

Estes resultados sugerem que a combinação de diferentes técnicas pode mitigar as limitações individuais de cada abordagem e reduzir incerteza de seleção de modelos, especialmente em cenários com dados limitados.

## 2.7 PREVISÃO DE MÚLTIPLOS PASSOS À FRENTE

A tarefa de previsão de múltiplos passos à frente é desafiadora (Sorjamaa et al., 2007; Hamzaçebi; Akay; Kutay, 2009; Taieb et al., 2012) e mais complexa que a tarefa de previsão de um passo à frente devido a alguns fatores como acumulação de erros e maior incerteza, que aumenta quanto mais longo for o horizonte futuro de previsão (Sorjamaa et al., 2007; Taieb et al., 2012). Dada uma série temporal  $\{z_1, z_2, \dots, z_t\}$  a qual compreende  $t$  observações, a previsão de múltiplos passos à frente visa prever as próximas  $H$  observações futuras,  $\{z_{t+1}, z_{t+1}, \dots, z_{t+H}\}$  (Taieb, 2014). Tradicionalmente (Taieb et al., 2012), a previsão de múltiplos passos à frente é empregada sob as seguintes abordagens:

- Recursiva ou iterativa
- Direta
- *Multi-Input Multi-Output* (MIMO)

A abordagem recursiva, segundo (Taieb et al., 2012), é aquela em que apenas um modelo  $f$  é treinado para prever um passo à frente por vez,

$$z_{t+1} = f(z_t, \dots, z_{t-d+1}) + \epsilon, \quad (2.23)$$

em que  $d$  é a quantidade de retardos temporais (*lags*).

Para se obter as previsões dos  $H$  horizontes, inicialmente o modelo  $f$  é empregado para prever o primeiro horizonte  $z_{t+1}$ . Em seguida, utiliza-se a previsão realizada como entrada para se obter o segundo horizonte de previsão utilizando o mesmo modelo  $f$ . Esse processo é contínuo até que se obtenha as previsões para todos os horizontes desejados. Matematicamente, considerando um modelo treinado  $\hat{f}$ , então as previsões são dadas por:

$$\hat{z}_{N+h} = \begin{cases} \hat{f}(z_N, \dots, z_{N+d-1}) & \text{se } h = 1 \\ \hat{f}(\hat{z}_{N+h-1}, \dots, \hat{z}_{N+1}, z_N, \dots, z_{N+d+h}) & \text{se } h \in \{2, \dots, d\} \\ \hat{f}(\hat{z}_{N+h-1}, \dots, \hat{z}_{N+h-d}) & \text{se } h \in \{d+1, \dots, H\} \end{cases} \quad (2.24)$$

Essa abordagem pode sofrer degradação na qualidade das previsões dependendo do grau de ruído contido nas previsões, o qual pode ser acumulado ao longo dos horizontes de previsão (Taieb et al., 2012).

A abordagem direta, segundo Hamzaçebi, Akay e Kutay (2009) e Taieb et al. (2012), consiste em realizar previsões independentes para cada horizonte. Desse modo, gera-se um modelo específico  $f_h$  para cada horizonte futuro de previsão de forma independente,

$$z_{t+h} = f_h(z_t, \dots, z_{t-d+1}) + \epsilon, \quad (2.25)$$

em que  $t \in \{d, \dots, N - H\}$ .

As previsões são obtidas utilizando os  $H$  modelos treinados  $\hat{f}_h$ ,

$$\hat{z}_{N+h} = \hat{f}_h(z_N, \dots, z_{N-d+1}). \quad (2.26)$$

Diferentemente da abordagem recursiva, a abordagem direta não manifesta o problema de acumulação de erros, entretanto ela trata as previsões de forma independente, o que pode incorrer numa perda de associação temporal entre as previsões (Taieb et al., 2012). Por sua vez, apresenta problemas para lidar com séries temporais menores.

A abordagem MIMO, segundo Taieb et al. (2012) e Bontempi (2008), considera a previsão de múltiplas saídas de uma única vez. Diferentemente ao mapeamento de uma saída comum às abordagens recursiva e direta, ela foi proposta no sentido de não desconsiderar, durante a etapa de modelagem e previsão, a dependência estocástica associada aos valores futuros (Bontempi, 2008; Taieb et al., 2012).

Essa abordagem, MIMO, treina um modelo  $F$  o qual aprende, de uma única vez, a associação entre os dados de entrada e as  $H$  horizontes de previsão,

$$[z_{t+H}, \dots, z_{t+1}] = F(z_t, \dots, z_{t+d-1}) + \epsilon, \quad (2.27)$$

em que  $t \in \{d, \dots, N - H\}$  e  $F : \mathbb{R}^d \Rightarrow \mathbb{R}^H$  é uma função vetorial, e  $\epsilon$  é um vetor de ruído branco com uma covariância que não é necessariamente diagonal (Taieb et al., 2012).

As previsões são obtidas de uma única vez por meio de um modelo  $\hat{F}$  de múltiplas saídas,

$$[\hat{z}_{t+H}, \dots, \hat{z}_{t+1}] = \hat{F}(z_N, \dots, z_{N-d+1}) \quad (2.28)$$

Sendo uma abordagem que busca reduzir os problemas das anteriores, a estratégia MIMO possui uma desvantagem, ou seja, ela restringe todos os horizontes de previsão à apenas um modelo de previsão (Taieb et al., 2012).

Perante o exposto, a previsão de múltiplos passos à frente em séries temporais é complexa e desafiadora, cada abordagem tradicional-recursiva, direta e MIMO-tem seus próprios méritos e limitações. Enquanto a abordagem recursiva é computacionalmente menos onerosa e útil para séries curtas, ela pode sofrer de acúmulo de erros, a direta evita tal acumulação mas ignora interdependências temporais. A MIMO tenta remediar algumas dessas falhas, mas restringe todos os horizontes, simultaneamente, a um único modelo e demanda mais dados que a recursiva. Essas complexidades tornam a busca por métodos mais eficazes e adaptáveis um campo de pesquisa em aberto.

## 2.8 SISTEMAS HÍBRIDOS PARA PREVISÃO DE SÉRIES TEMPORAIS

A combinação de modelos de previsão pode ser feita pelo emprego de máquinas de comitê (Mendes-Moreira et al., 2012), ou por meio de sistemas híbridos (Zhang, 2003), este por sua vez pode utilizar um modelo linear para prever os valores da série temporal e um não-linear para corrigir os erros do primeiro modelo. Dessa maneira, essa segunda abordagem assume que uma série temporal possui dois componentes, o linear e o não-linear. É partindo desse pressuposto que este trabalho foca na exploração de sistemas híbridos para previsão de vários passos à frente. De acordo com Hajirahimi e Khashei (2019), a razão pela qual os sistemas híbridos demonstram desempenho superior para previsão de séries temporais reside na capacidade deles reconhecerem padrões e modelar, por meio da redução do risco de ajustar inadequadamente um modelo, e também em simplificar o processo de seleção dos modelos adequados para tal tarefa.

Por sua vez, Zhang (2003) propôs um sistema híbrido  $z_t$  cuja fórmula matemática está representada na Equação 2.29 e que é composta por uma soma aritmética dos componentes linear e não-linear e um erro aleatório.

$$z_t = L_t + N_t + \varepsilon_t, \quad (2.29)$$

em que  $L_t$  é o preditor linear (e.g., ARIMA),  $N_t$  é o modelo não-linear (e.g., MLP) e  $\varepsilon_t$  é o erro aleatório remanescente, o qual não pode ser aprendido por  $N_t$ . Assim, para ajustar sistematicamente  $N_t$ , calcula-se  $e_t$ , que está expresso na Equação 2.30, como sendo os resíduos do modelo linear  $\hat{L}_t$  estimado para  $z_t$ ,

$$e_t = z_t - \hat{L}_t \quad (2.30)$$

Com base nisso, parte-se da ideia de que o modelo linear ARIMA para  $z_t$  não apresenta de autocorrelação linear em seus resíduos. Diante dessa premissa, Zhang (2003) aplicou uma rede neural artificial com arquitetura de Perceptron de múltiplas camadas para prever os resíduos do modelo ARIMA, conforme sugere a Equação 2.31,

$$\hat{N}_t = f(e_{t-1}, e_{t-2}, \dots, e_{t-n}) \quad (2.31)$$

em que  $f(\cdot)$  representa a modelagem não-linear realizada pela rede neural artificial para um conjunto de  $n$  valores defasados da série dos resíduos do modelo  $L_t$ .

Diante disso, Zhang (2003) enfatizou a necessidade de a modelagem do componente linear ser sistematicamente adequada, como forma de garantir que os erros do sistema híbrido sejam aleatórios, e não detenha autocorrelação linear. Por consequência, os valores previstos do sistema híbrido ( $\hat{z}_t$ ) são calculados pela combinação linear aditiva da previsão feita pelo modelo linear ( $\hat{L}_t$ ) e das previsões feitas pelo modelo não-linear  $\hat{N}_t$  como mostra a Equação 2.32.

$$\hat{z}_t = \hat{L}_t + \hat{N}_t \quad (2.32)$$

É importante frisar que o modelo híbrido supracitado proposto por Zhang (2003) foi desenvolvido para realizar previsões de um passo à frente.

Após Zhang (2003), alguns trabalhos seguiram sua arquitetura para desenvolver outros sistemas híbridos para previsão de séries temporais com foco em previsões de um passo à frente. Pai e Lin (2005) usaram ARIMA com uma *Support Vector Regression* (SVR) para prever preços de ações e observou maior acurácia para o híbrido. Chakraborty, Chattopadhyay e Ghosh (2019) combinaram de forma aditiva um modelo ARIMA com uma Rede Neural Autoregressiva para prever epidemias de dengue em múltiplos passos à frente. Os resultados indicaram que a hibridização apresenta maior precisão preditiva em comparação com modelos individuais. Analogamente, Moraes e Gomes (2022) combinou um modelo SARIMA <sup>9</sup> com uma MLP para prever incidência de casos de Covid-19 no mundo, e observou que o modelo híbrido mostrou-se superior às abordagens individuais, sobretudo em horizontes de sete e vinte oito dias à frente.

<sup>9</sup> *Seasonal Auto-Regressive Integrated Moving Average* (SARIMA) é uma extensão do modelo ARIMA que incorpora efeitos de sazonalidade.

Nessa mesma perspectiva, combinaram ARIMA e LSTM em problemas de previsão, tais como exportação na Indonésia (Dave et al., 2021), produção de gás e petróleo (Fan et al., 2021), seca no curto prazo (Xu et al., 2022). Panigrahi e Behera (2017) propuseram uma combinação de um ETS para prever a série temporal e uma MLP para prever os resíduos. Por sua vez, Oliveira, Silva e de Mattos Neto (2021) utilizaram seleção dinâmica de preditor na etapa de modelagem dos resíduos. Medina e Oliveira (2023) empregaram seleção de preditor tanto na etapa de modelagem residual como na combinação dos componentes lineares e não-lineares. Santos Júnior et al. (2023) usaram uma combinação de preditores na etapa de modelagem residual. Os resultados desses estudos apontaram que o sistemas híbridos demonstraram o melhor desempenho de previsão em comparação às abordagens individuais.

No que diz respeito à previsão de um passo e vários passos à frente, Babu e Reddy (2014) propuseram um sistema que combinou um ARIMA, uma RNAs e um filtro de média móvel. Esse sistema decompõe uma série temporal em dois componentes, alta e baixa volatilidade. Então, realiza primeiro modelagem com o ARIMA, em seguida a modelagem com uma Rede Neural Artificial (RNA), e por fim os combina com uma soma. Os resultados do estudo Babu e Reddy (2014) mostrou que seu sistema superou, em termos de acurácia de previsão, o sistema híbrido de Zhang (2003) para três conjuntos de dados em previsões de um passo e vários passos à frente. Ademais, destaca-se a quantidade de passos futuros considerados pelo estudo foi de no máximo cinco (5) observações futuras. Quando consideramos o escopo de previsão de mortalidade, a quantidade de passos futuros tende a necessitar de maiores valores, estabelecendo por exemplo pelo menos dez passos (Feng; Shi, 2018).

Seguindo essa perspectiva de pensamento, Shi, Guo e Zheng (2012) aplicou um sistema híbrido para prever vários passos à frente de velocidade de vento e geração de potência. Contudo, as conclusões observadas apontam para uma resposta não tão clara quanto à acurácia e o desempenho do sistema para a tarefa de previsão de múltiplos passos. Sendo assim, não conseguiram concluir se os sistemas híbridos sob a premissa de previsão de um passo à frente demonstraram superior desempenho preditivo em relação à modelagem de apenas único modelo.

Sob outro ponto de vista no que tange à maneira de hibridizar, Iftikhar et al. (2023) propuseram um modelo híbrido para prever infecções e óbitos relacionados ao vírus Mpox. Eles utilizaram o filtro Hodrick-Prescott para decompor a série em duas componentes: tendência e estocástica. Em seguida, combinaram modelos estatísticos e de aprendizado de máquina para fazer as previsões. Os resultados mostraram que a proposta superou os modelos individuais,

fornecendo previsões mais precisas. No que tange à previsão para múltiplos passos à frente de carga de curto prazo em cenários com poucos dados, (Meng; Han; Hou, 2022) propuseram um modelo híbrido que superou os modelos individuais tradicionais. Sua abordagem se deu pela decomposição da série temporal em componentes extrínsecos-variacionais e intrínsecos-estacionários, modeladas por *Gradient Boosting* e ARIMA, respectivamente.

Por fim, destaca-se o sistema híbrido recursivo-direto, HyS-MF (Duarte; de Mattos Neto; Firmino, 2024), desenvolvido para prever séries pequenas de taxas de mortalidade a partir da modelagem linear de abordagem recursiva na série temporal e não-linear de abordagem direta nos resíduos do preditor linear. Os autores mostraram que o sistema proposto se mostrou uma alternativa promissora no campo de previsão de taxas de mortalidade e sistemas híbridos, superando diversos modelos da literatura.

## 2.9 CONSIDERAÇÕES DO CAPÍTULO

Sob o contexto de sistemas híbridos para previsão de séries temporais, em contrapartida aos modelos apresentados e levando em consideração a ideia de combinar modelo linear e não-linear, este trabalho visa desenvolver um sistema híbrido cuja finalidade é combinar modelagens lineares e não-lineares para previsão de vários passos à frente que sejam capazes de melhorar as previsões frente à modelagem da literatura.

Especificamente, entende-se que há a necessidade de realizar o reconhecimento de padrões de longo prazo, atividade esta que pode demandar a geração de múltiplos modelos não-lineares para a modelagem residual. Ainda, considerando a natureza de curto prazo das séries temporais de mortalidade, é fundamental optar pela abordagem recursiva para previsões de múltiplos passos à frente, primordialmente pois ela é a menos complexa estatística e computacionalmente. Consequentemente, emprega-se a seleção de modelos não-lineares, levando em conta a dinâmica dos problemas observados em modelagem *data-driven*.

### 3 SISTEMA PROPOSTO

Este capítulo tem por objetivo apresentar os conceitos e etapas do sistema híbrido proposto para previsão de vários passos à frente em detalhes, e está dividido nas seguintes seções. Seção 3.1 apresenta o sistema híbrido proposto para previsão de vários passos à frente de taxas de mortalidade, e está subdividida em duas subseções. Seção 3.1.1 explica a etapa de treinamento, ao passo que Seção 3.1.2 detalha a etapa de teste. Seção 3.2, apresenta as considerações do capítulo.

#### 3.1 SISTEMA HÍBRIDO PROPOSTO PARA PREVISÃO DE VÁRIOS PASSOS À FRENTE

O sistema híbrido proposto neste trabalho visa realizar previsões de vários passos à frente. Como casos de estudo, serão considerados exercícios de previsão de taxas de mortalidade humana.

As Figuras 5 e 6 mostram as etapas de treinamento e teste do sistema híbrido proposto. O sistema proposto está dividido em duas etapas: Treinamento e Teste. A primeira etapa consiste no treinamento dos modelos lineares e não-lineares. O modelo ARIMA foi escolhido para modelar os padrões lineares das séries temporais de mortalidade, e *pool* com os modelos MLP, LSTM e N-BEATS foi selecionado para modelar e prever os padrões não-lineares das séries residuais resultantes da abordagem linear. As modelagens linear e não-linear utilizam a abordagem de previsão recursiva de múltiplos passos à frente (Taieb et al., 2012; Livieris; Pintelas, 2022), que desenvolve um modelo de previsão de um passo à frente para prever cada passo futuro (horizonte) de maneira iterativa, também chamada de recursiva. Nesse contexto, a simplicidade inerente à abordagem recursiva, combinada com sua capacidade de utilizar um conjunto de dados de treinamento maior, proporciona uma vantagem considerável (Ahani; Salari; Shadman, 2019), principalmente ao modelar séries temporais curtas. A modelagem linear e a modelagem não-linear são denotadas na Figura 5 por ML e MN, respectivamente.

Na fase de teste, ARIMA e o modelo não-linear preveem os novos dados não vistos. Assim, o modelo estatístico linear realiza a previsão das taxas de mortalidade, e os resíduos do preditor linear são previstos por meio do melhor preditor de AM não-linear selecionado de acordo com o conjunto de validação. Portanto, a previsão final é gerada a partir da combinação da previsão da série temporal com a estimativa da série residual.

A motivação para desenvolver o sistema proposto baseou-se na observação de que, embora os modelos lineares e não-lineares possuam pontos fortes, eles também apresentam limitações quando utilizados individualmente. A metodologia de Box & Jenkins (Box et al., 2015) assegura que o ARIMA, quando adequadamente ajustado à série temporal, seja capaz de capturar os padrões lineares, resultando em resíduos que contêm predominantemente padrões não-lineares e erros aleatórios. Ademais, a escolha de empregar um modelo estatístico linear, como o ARIMA, em vez de aplicar diretamente um modelo não-linear de AM à série temporal, reduz o risco de sobreajuste frequentemente associado aos modelos de AM (Zhang, 2003).

Por outro lado, modelos de AM como Perceptron Multicamadas (MLP), Long Short-Term Memory (LSTM) e N-BEATS são algoritmos capazes de capturar tanto padrões lineares quanto não-lineares. No entanto, as evidências empíricas (Zhang, 2003; Pai; Lin, 2005; Panigrahi; Behera, 2017; Rubio; Alba, 2022) apresentam resultados mistos quanto à eficácia desses modelos quando aplicados diretamente às séries temporais, ou seja, quando utilizados para modelar simultaneamente padrões lineares e não-lineares. Adicionalmente, cada um desses modelos possuem arquiteturas distintas, o que resulta em diferentes modos para aprender e representar os padrões das séries temporais. Sendo assim, esta diversidade de arquitetura implica que o problema de seleção de modelos persiste, pois nenhum modelo é universalmente superior em todos conjuntos de dados (Murphy, 2012; Silva et al., 2020). Paralelamente, estudos empíricos têm demonstrado que a combinação de modelos heterogêneos, em vez da utilização de modelos semelhantes, leva a um desempenho de previsão superior (Zhang, 2003; Mendes-Moreira et al., 2012; Hajirahimi; Khashei, 2019; Oliveira; Silva; de Mattos Neto, 2021; Duarte; de Mattos Neto; Firmino, 2024). Esta constatação sugere que a diversidade de modelos pode ser propícia para capturar diferentes aspectos das séries temporais.

O sistema híbrido proposto visa resolver os desafios mencionados anteriormente, levando em consideração três aspectos das séries de taxas de mortalidade: (i) são séries temporais curtas; (ii) podem ser descritas, em certa medida, por uma função denominada curva de mortalidade, que relaciona cada idade a uma taxa de mortalidade específica em um determinado unidade de tempo; e (iii) as previsões são realizadas para múltiplos passos à frente. Aproveitando os pontos fortes dos modelos supracitados e introduzindo uma inovação no processo de seleção de modelos, o sistema híbrido proposto opera da seguinte forma:

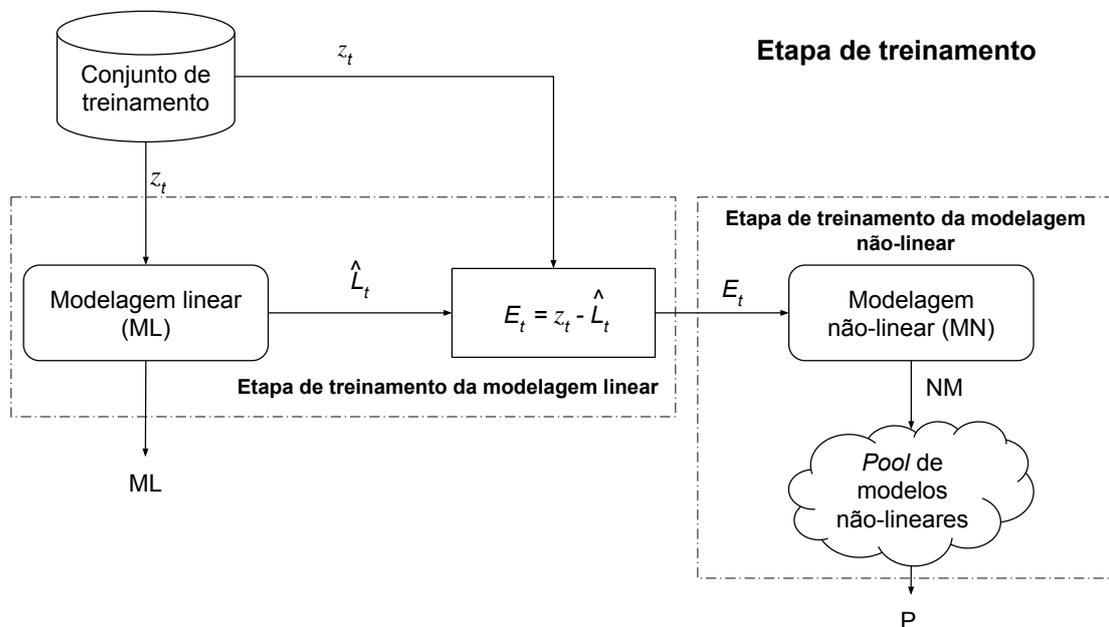
- O modelo ARIMA é empregado na modelagem dos padrões lineares das séries temporais.
- Um conjunto de modelos de AM é gerado para aprender os padrões não-lineares das

séries residuais usando uma abordagem recursiva em várias etapas.

- Em vez de selecionar o melhor modelo de AM para cada série residual individualmente, a proposta inova ao realizar um processo de seleção baseado na curva de mortalidade dos resíduos.

A curva de mortalidade dos resíduos é gerada por meio da interpolação *spline* cúbica das estimativas no conjunto de validação. Este processo de combinação introduz uma nova abordagem para a seleção de modelos, particularmente eficaz quando existe uma função capaz de descrever um comportamento comum entre as séries temporais em cada instante de tempo.

Figura 5 – Etapa de treinamento do sistema híbrido proposto utilizado nesta pesquisa.



Fonte: Elaborado pelo autor (2024).

### 3.1.1 Etapa de treinamento

O sistema proposto inicia o processo buscando o melhor preditor linear para cada série temporal do conjunto de treinamento  $z_t$ . Esse conjunto é composto por um subconjunto das idades de uma população. A escolha desse subconjunto está atrelada à capacidade de realizar treinamento e previsão de modelos para idades que representam maior variabilidade de uma curva de mortalidade humana. Segundo Wu e Wang (2018) a escolha de um grupo de vinte (20) idades seria suficiente para ajustar a curva de mortalidade (obter os valores da curva para as idades omitidas na etapa de treinamento) compreendida pelas idades de 0 até 100 anos.

Essa sistemática também reduz o tempo computacional dos experimentos, uma vez que não há necessidade de modelar as séries de todas as idades.

Por conseguinte, o sistema proposto nesta tese utiliza os fundamentos da metodologia Box & Jenkins (Box et al., 2015) para ajustar o modelo linear. Neste sentido, inicialmente, ele avalia a estacionariedade da série temporal, e caso essa condição não se verifique, o algoritmo aplica a operação da diferença. Em seguida, ele seleciona o preditor que assume normalidade e não mostra a autocorrelação linear relevante dos resíduos, e tem o valor mais baixo do *Akaike Information Criterion* (AIC), uma métrica que dá prioridade aos modelos com menos parâmetros (ou seja, os modelos parcimoniosos). O valor do AIC é calculado como:

$$\text{AIC} = 2k - 2 \ln(\hat{L}), \quad (3.1)$$

em que  $k$  é o número de parâmetros e  $\hat{L}$  é o valor da função de máxima verossimilhança subjacente ao modelo, face às suas estimativas de parâmetros e resíduos.

Após ter realizado o treinamento de um modelo linear para cada série temporal do subconjunto de idades para treinamento e calculado a sua estimativa  $\hat{L}_t$ , o sistema proposto calcula as séries residuais ( $E_t$ ) como a diferença entre os valores observados das séries temporais ( $z_t$ ) e as respectivas previsões como mostra a equação 3.2,

$$E_t = z_t - \hat{L}_t. \quad (3.2)$$

Após o treinamento do modelo linear, os resíduos são calculados especificamente para o conjunto de treinamento. Estes resíduos são então cronologicamente divididos em dois subconjuntos distintos, os quais são nomeados de subconjuntos de treinamento e validação de resíduos. O subconjunto de validação de resíduos consiste nas observações que ocorrem sequencialmente após o subconjunto de treinamento de resíduos.

Tendo em vista que os dados de mortalidade humana compreendem séries temporais de tamanho relativamente pequeno (i.e., em torno de 60 observações), opta-se pela abordagem recursiva de múltiplos passos à frente, também chamada de iterativa. Essa abordagem é a menos complexa do ponto de vista estatístico (i.e., mais parcimoniosa) e computacional, tal como tem demonstrado resultados competitivos e vem sendo empregada em diversos estudos (Beyaztas; Shang, 2022; Taieb et al., 2012).

Assim sendo, na etapa de modelagem de resíduos, um *pool* de modelos de AM não-lineares (MN) é projetado para modelagem residual ( $N_t$ ) com o objetivo de maximizar a acurácia de previsão de apenas um passo à frente, seguindo a abordagem recursiva. Sendo assim, cada

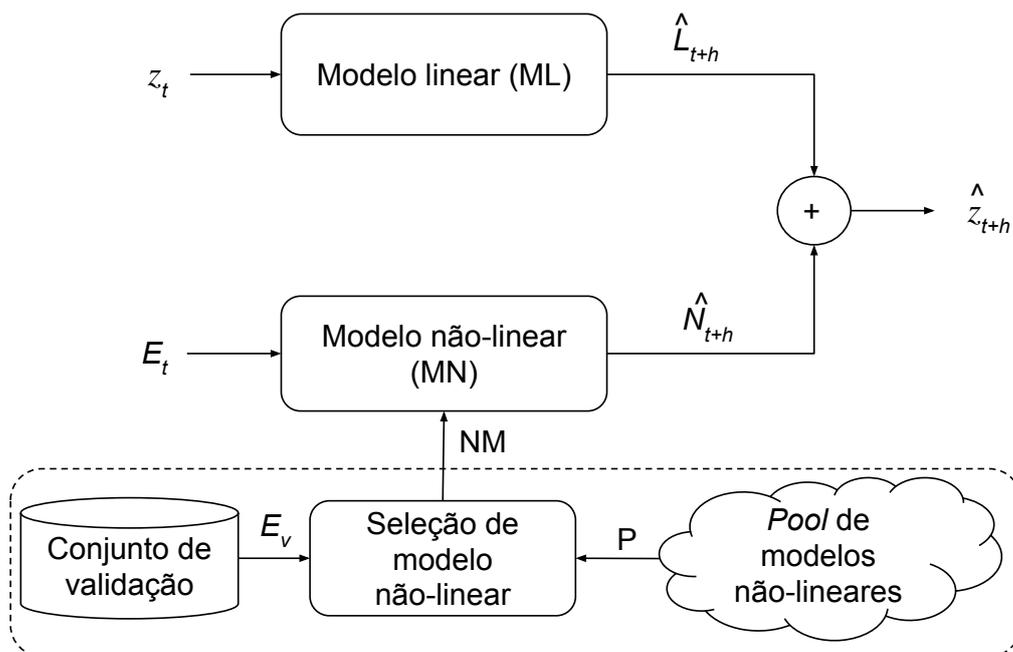
modelo de previsão (MN) realiza uma previsão por vez, e se retroalimenta das previsões realizadas (como dados de entrada) para obter as estimativas futuras até que se tenha todas as  $H$  estimativas. Os modelos MLP, LSTM (Hochreiter; Schmidhuber, 1997) e N-BEATS (Oreshkin et al., 2019) foram escolhidos por serem redes neurais de aprendizado de máquina capazes de aprender qualquer padrão modelável, tradicionalmente por serem escolhidas (exceto pela última) em sistemas híbridos, e dotadas resultados relevantes em termos de acurácia preditiva em uma vasta coleção de conjuntos de dados de séries temporais (Zhang; Patuwo; Hu, 1998; Richman; Wüthrich, 2021; Hajirahimi; Khashei, 2019; Nigri et al., 2019; Oreshkin et al., 2019; Oreshkin et al., 2021; Chen; Khaliq, 2022; Ahmed; Hassan; Mstafa, 2022).

Para o sistema proposto, foi considerado que o conjunto de teste tem  $H$  passos à frente. Para cada  $E_t$ , um *pool*  $P$  de modelos não-lineares é concebido para aprender os seus padrões.

Portanto, a Fase de Treino recebe o  $z_t$  como entrada e faz a Modelagem Linear e a Modelagem não-linear para produzir os Modelos de Treinamento, compreendendo um modelo linear e um *pool* de modelos não-lineares, para a etapa de Teste.

### 3.1.2 Etapa de teste

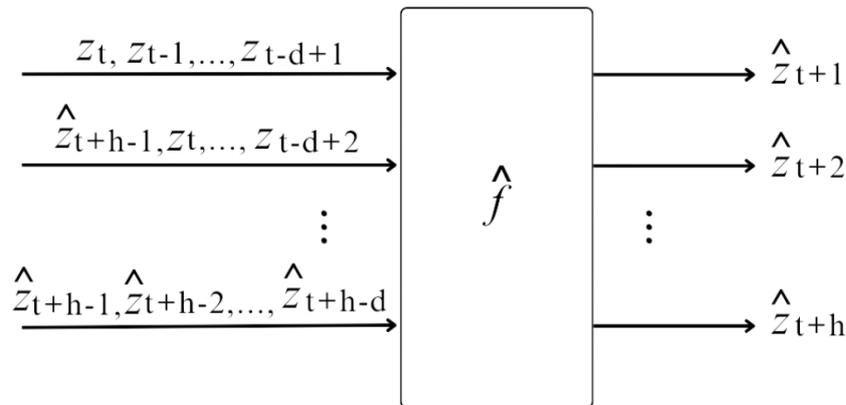
Figura 6 – Etapas de teste do sistema híbrido proposto utilizado nesta pesquisa.



Fonte: Elaborado pelo autor (2024).

A Figura 6 mostra a etapa de teste, que consiste na seleção de modelo não-linear e na

Figura 7 – Abordagem recursiva de previsão de múltiplos passos à frente usada na etapa de teste.



Fonte: Elaborado pelo autor (2024).

previsão das taxas de mortalidade e dos resíduos do preditor linear. Na fase de seleção do modelo, o conjunto de validação das séries residuais ( $E_v$ ) é usado. Sendo assim, conduzindo a modelagem não-linear (MN), as previsões para os dados de validação são geradas para modelos  $p$  denotados como  $P$ , que são então usados no processo de seleção de modelos. A seleção de modelos envolve: (i) aplicar a interpolação de *spline* cúbica na previsão da série residual para o subconjunto de idades previamente utilizadas para obter a previsão residual para todas as faixas etárias; (ii) avaliar a acurácia da previsão para o conjunto de validação com todas as idades, ou seja, selecionar o modelo não-linear (MN) que produz o menor erro de previsão para o resíduo do preditor linear.

Posteriormente, o sistema proposto utiliza os modelos treinados, ML e MN, para prever o conjunto de teste das taxas de mortalidade e dos resíduos do preditor linear, respectivamente. Neste caso, levando em conta o aspecto de previsão de vários passos à frente, os dados de entrada - série temporal ( $z_t$ ) e resíduos ( $E_t$ ) são utilizados para produzir a primeira previsão de ambos os modelos (ML e MN). Eles compreendem as observações temporais que antecedem o conjunto de teste. Tomando por base a abordagem recursiva (Equação 2.24 e Figura 7), do segundo horizonte em diante, as previsões geradas também retroalimentam os modelos do sistema proposto. Portanto, para cada idade  $x$ , a ML (parte linear) recebe  $z_t$  e prevê os  $\hat{L}_{t+h}$  para todos os horizontes temporais em múltiplos passos em frente de forma recursiva. A série residual do modelo linear ( $E_t$ ) do conjunto de treinamento é utilizada como entrada para o melhor preditor da MN para prever a parte não-linear  $\hat{N}_{t+h}$ . A partir da abordagem recursiva (similar ao ML), o MN realiza a previsão para cada  $h$  ( $h = 1, \dots, H$ ) horizonte futuro temporal.

---

Em seguida, para cada horizonte temporal futuro  $h$  onde  $h = 1, \dots, H$ , é aplicada uma soma para calcular o  $\hat{z}_{t+h}$  como mostrado na Equação 3.3.

$$\hat{z}_{t+h} = \hat{L}_{t+h} + \hat{N}_{t+h}. \quad (3.3)$$

## 3.2 CONSIDERAÇÕES DO CAPÍTULO

O sistema proposto tem como fundamento a combinação de um modelo linear e um não-linear para previsão de vários passos à frente, no qual o segundo é selecionado na etapa de teste a partir da avaliação do desempenho de previsão dos resíduos no conjunto de validação para todas as idades de uma dada população. A combinação é pautada na ideia de que uma série temporal é composta por dois componentes que se agregam por uma soma aritmética. Por meio disso, entende-se que a modelagem linear faz o papel de modelar os padrões lineares e, ao passo que o modelo não-linear realizaria o aprendizado dos padrões contidos nos desvios observados das séries em relação às previsões geradas pelo preditor linear.

## 4 PROTOCOLO EXPERIMENTAL

Este capítulo apresenta a metodologia dos experimentos deste trabalho para alcançar os objetivos da pesquisa, e está dividido nas seguintes seções. A Seção 4.1 apresenta as bases de dados utilizadas e as correspondentes funções de autocorrelação e autocorrelação parcial. A Seção 4.2 apresenta as métricas de avaliação. A Seção 4.3 detalha os procedimentos de aplicação e comparação do sistema proposto em relação à literatura. A Seção 4.4 apresenta as considerações do capítulo.

### 4.1 BASE DE DADOS

Os experimentos foram realizados com bases de dados de mortalidade de cinco populações disponíveis no *Human Mortality Database* (HMD) <sup>1</sup>. Utilizou-se a segregação por gênero masculina e feminina, mas também foi utilizada a população total, o que contabilizou, portanto, quinze bases de dados. Os países foram: Austrália, Estados Unidos (EUA), França, Japão e Portugal. Destaca-se que não foram utilizados dados brasileiros, pois as informações disponibilizadas pelo Instituto Brasileiro de Geografia e Estatística (IBGE) consistem em taxas de mortalidade por faixas etárias, não havendo dados disponíveis para cada idade específica. Inclusive, as séries históricas fornecida pelo IBGE iniciam-se apenas em 1991, o que representa um conjunto de dados reduzido para os propósitos deste estudo.

As bases de dados são compostas pela séries temporais das taxas brutas de mortalidade anuais para cada idade das populações supracitas para o período entre os anos 1950 e 2019, exceto para a Austrália a qual empregou o período de 1921 e 2019.

Os dados dos anos pré-1950 foram subtraídos tendo em vista que poderiam dificultar a análise dos experimentos desta pesquisa, uma vez que perfaziam períodos de guerras mundiais e de pandemia da gripe espanhola. Esses eventos, por sua vez produziram ruídos consideráveis para algumas regiões que modificaram sistematicamente e de elevada magnitude o comportamento das taxas de mortalidades das idades constitutivas do ciclo vital adulto. No caso de Portugal <sup>2</sup> embora o país não tenha participado da Segunda Guerra Mundial, optou-se por

<sup>1</sup> <https://www.mortality.org/>. Acessada em 18/01/2023.

<sup>2</sup> Há dados disponíveis a partir de 1940.

utilizar os dados portugueses a partir de 1950, visando a padronização do período de análise entre as diferentes populações estudadas. Quanto à Austrália, decidiu-se utilizar os dados desde 1921 com o objetivo de realizar uma comparação com um conjunto de dados mais amplo. Ademais, conforme será demonstrado adiante, o impacto da segunda guerra mundial nas taxas de mortalidade da Austrália foi irrelevante para os propósitos desta pesquisa. Esse período estendido pode evidenciar diferenças no desempenho de previsão em função do tamanho amostral disponível para treinamento. Dessa forma, a inclusão de um período mais extenso para a Austrália contribui para enriquecer a análise, sem comprometer a consistência e a comparabilidade dos resultados.

Além disso, foi adotada a transformação do logaritmo natural aos dados para garantir comparabilidade em relação aos demais trabalhos da literatura, e por conta que esse procedimento suaviza as séries temporais podendo melhorar o desempenho de modelos como ARIMA, notadamente quando a variância é estável (Lütkepohl; Xu, 2012). Ademais, essa transformação logarítmica restringe as previsões a manterem valores positivos mesmo na escala natural dos dados (Hyndman; Athanasopoulos, 2018).

A seguir estão expostos os gráficos de séries temporais utilizadas neste trabalho. Também, está apresentada a escolha de apenas um grupo de idades para modelagem pelo sistema proposto, o qual necessita de um subconjunto de idades da população.

Tomando por base o conjunto de dados da população francesa total, ou seja, considerando ambos os gêneros de forma agregada, detalha-se, a seguir, a análise das séries das taxas de mortalidade que pode se estender para as demais populações tendo em vista a semelhança entre os conjuntos de dados. A Figura 8 esboça o logaritmo das taxas de mortalidade em idades: (a) 0, 1, 5, 10, 15, 18 e 25; (b) 30, 31, 50, 51, 70 e 71 anos; e (c) 50, 51, e 52. A Figura 8 também mostra as taxas de mortalidade na escala natural em idades: (d) 0, 1, e 2; (e) 18, 25, 30; e (f) 90, 95 e 100. É possível notar que as séries de taxas de mortalidade têm geralmente uma tendência descendente ao longo do tempo. As séries com a redução mais significativa dos valores ocorreram na idade 0 (Figura 8a e 8d). Este comportamento pode ser explicado pelo progresso da medicina (Walk, 2017), saúde pública com a expansão do saneamento básico (Walk, 2017; Harris; Helgertz, 2019), e crescimento do rendimento (Kesztenbaum; Rosenthal, 2014).

Ainda, é possível verificar que a inclinação da série original para a idade 0 é mais acentuada que a da série da idade 1, evidenciando a importância de selecioná-las como séries para a modelagem do sistema proposto. Comparando com outras idades, percebe-se que o nível do

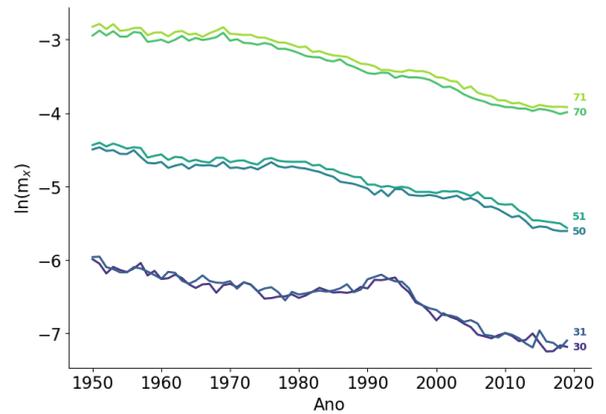
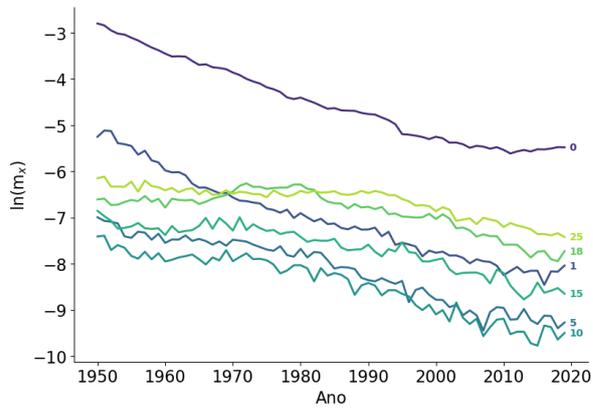
logaritmo natural da taxa de mortalidade para a idade 0 é mais elevado, significando que muitos recém-nascidos falecem assim que nascem. De uma forma geral, deduz-se que a taxa de mortalidade da população francesa decresceu ao longo dos anos detalhados nesta pesquisa. Contudo, esse decréscimo variou ao longo das idades, com um declínio mais lento nas idades mais elevadas em comparação às idades mais jovens. Isso mostra a relevância da diferença entre o comportamento das séries estudadas nesta pesquisa.

As Figuras 9-23 apresentam o logaritmo natural das taxas de mortalidade para o período de 1921 ou 1950 a 2019 dos países Austrália (iniciando em 1921), França, EUA, Japão e Portugal, segregado pelos gêneros masculino, feminino e para toda a população (total). Esses dados foram selecionados para modelagem através de sistemas híbridos, ARIMA e modelos individuais de AM. A partir dessas figuras, observou-se um comportamento similar entre os países, o que coaduna com a ideia dos padrões de mortalidade sistemáticos para a população humana (Milne, 2010). No entanto, Portugal apresentou maior variabilidade nas taxas de mortalidade para o grupo de idades entre 2 e 30 anos. De forma similar, o conjunto de dados femininos da Austrália, para idades entre 2 e 15 anos, também mostrou uma variabilidade maior em comparação com outros países. Essas variações podem ser explicadas por fatores exógenos que não estão no escopo desta pesquisa.

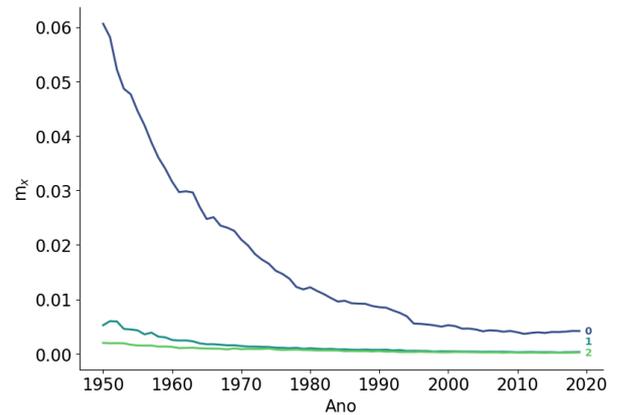
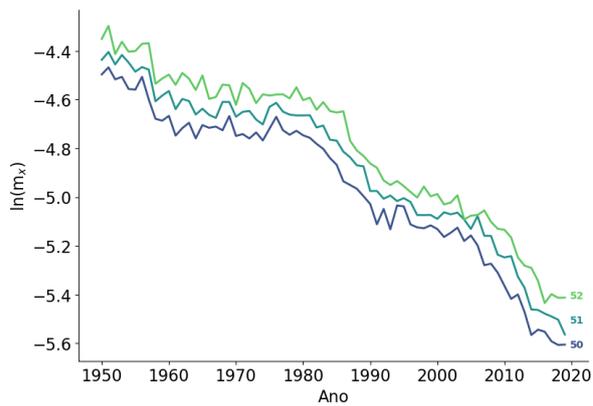
Em resumo, as taxas de mortalidade para idades próximas apresentam tendências decrescentes semelhantes ao longo do tempo. Isso indica que uma idade específica pode eficazmente representar um conjunto de idades com padrões temporais análogos. Assim, os métodos de previsão podem modelar esses agrupamentos de forma eficiente. Este ponto foi corroborado pelo estudo de Wu e Wang (2018), que demonstrou através de uma análise de sensibilidade que a curva de mortalidade, quando projetada para um subgrupo de idades e interpolada usando uma *spline* cúbica, assegura um ajuste preciso da curva para todas as idades.

Figura 8 – Taxas de mortalidade para idades em escala natural e logarítmica da população francesa total.

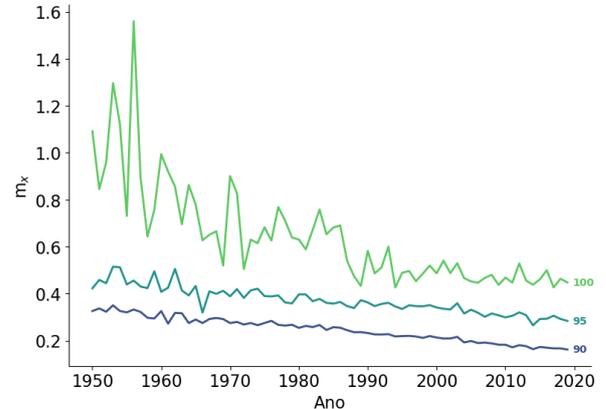
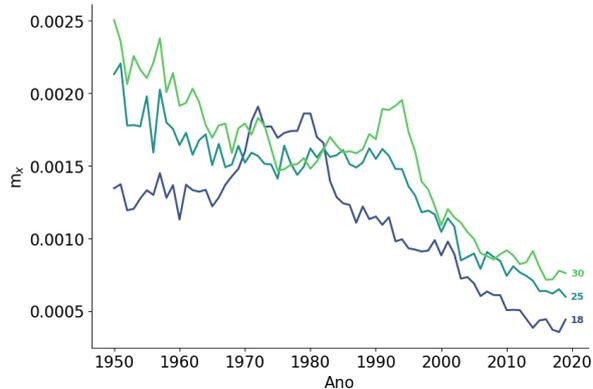
- (a) Logaritmo das taxas de mortalidade para as idades 0, 1, 5, 10, 15, 18 e 25 anos. (b) Logaritmo das taxas de mortalidade para as idades 30, 31, 50, 51, 70 e 71 anos.



- (c) Logaritmo das taxas de mortalidade para as idades 50, 51 e 52 anos. (d) Taxas de mortalidade para as idades 0, 1 e 2 anos.

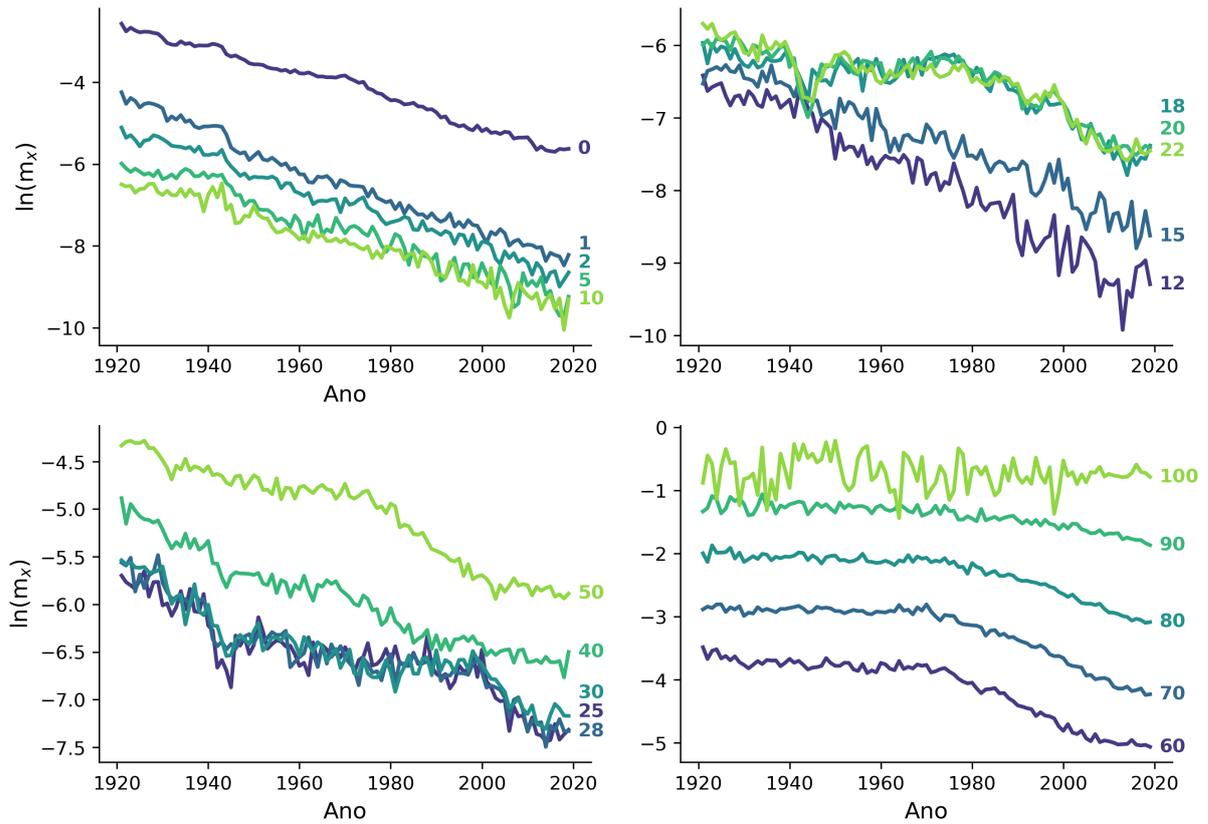


- (e) Taxas de mortalidade para as idades 18, 25 e 30 anos. (f) Taxas de mortalidade para as idades 90, 95 e 100 anos.



Fonte: Elaborado pelo autor (2024).

Figura 9 – Logaritmo das taxas de mortalidade por idade da população masculina australiana.



Fonte: Elaborado pelo autor (2024).

Figura 10 – Logaritmo das taxas de mortalidade por idade da população feminina australiana.

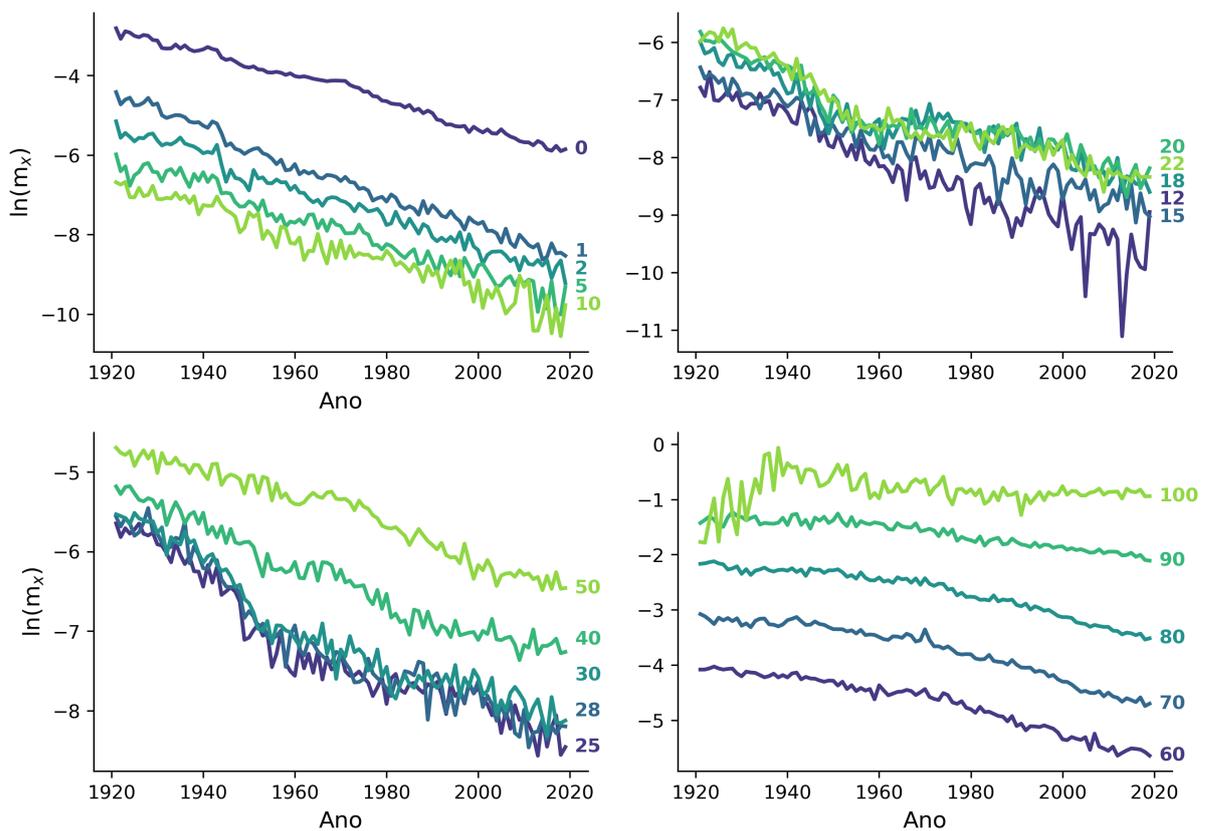


Figura 11 – Logaritmo das taxas de mortalidade por idade da população total australiana.

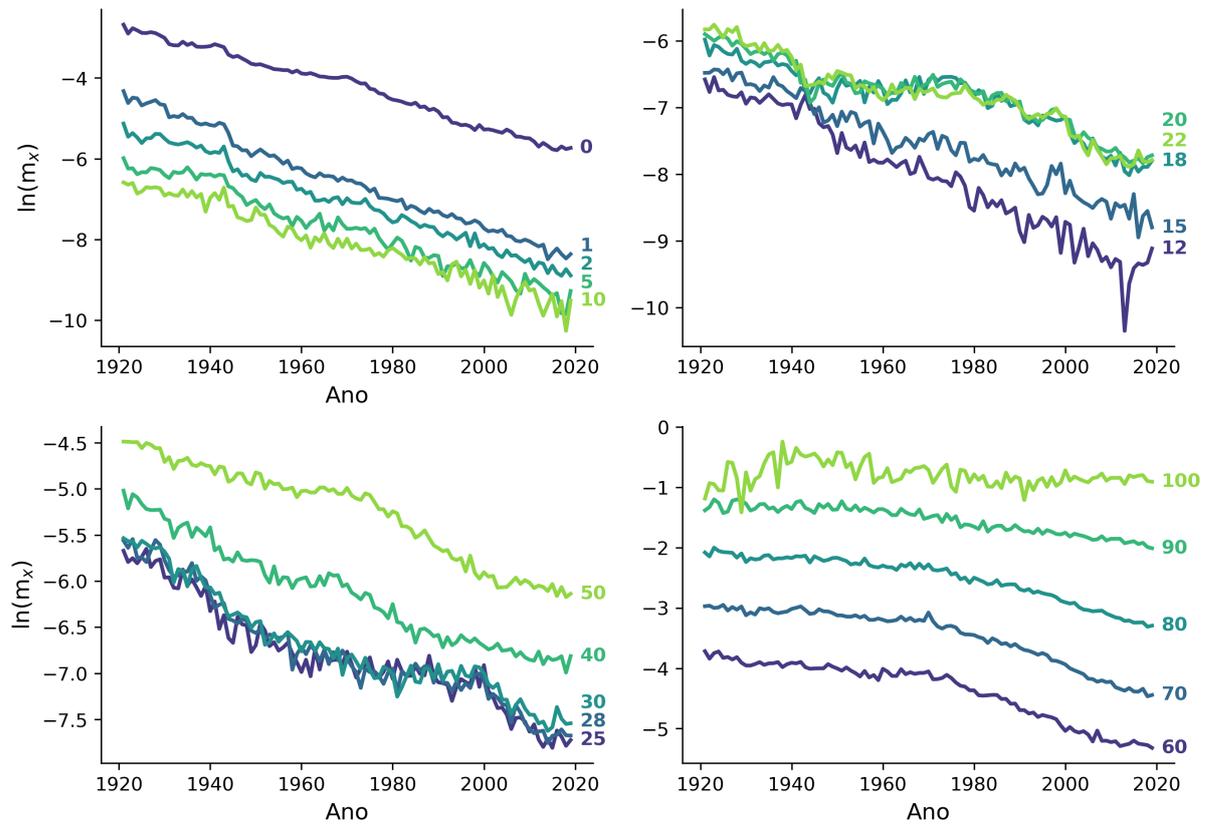


Figura 12 – Logaritmo das taxas de mortalidade por idade da população masculina americana.

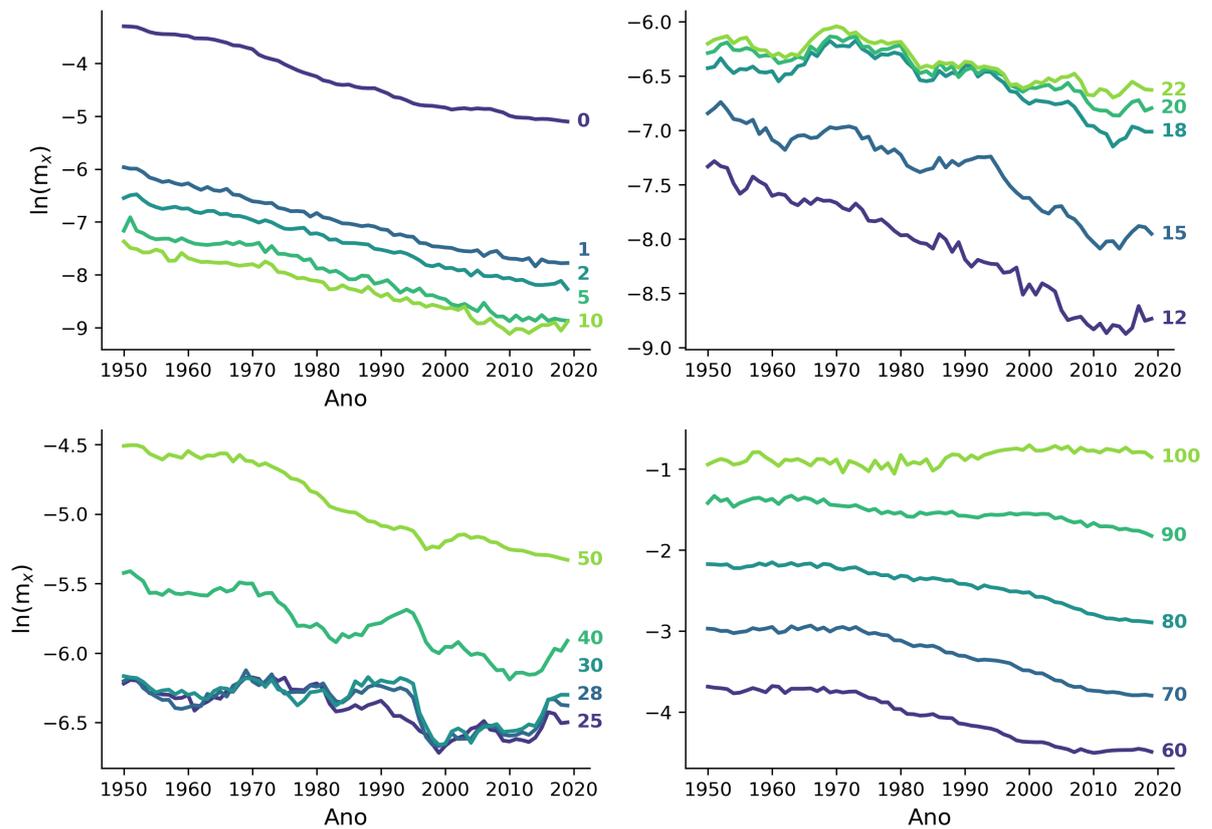


Figura 13 – Logaritmo das taxas de mortalidade por idade da população feminina americana.

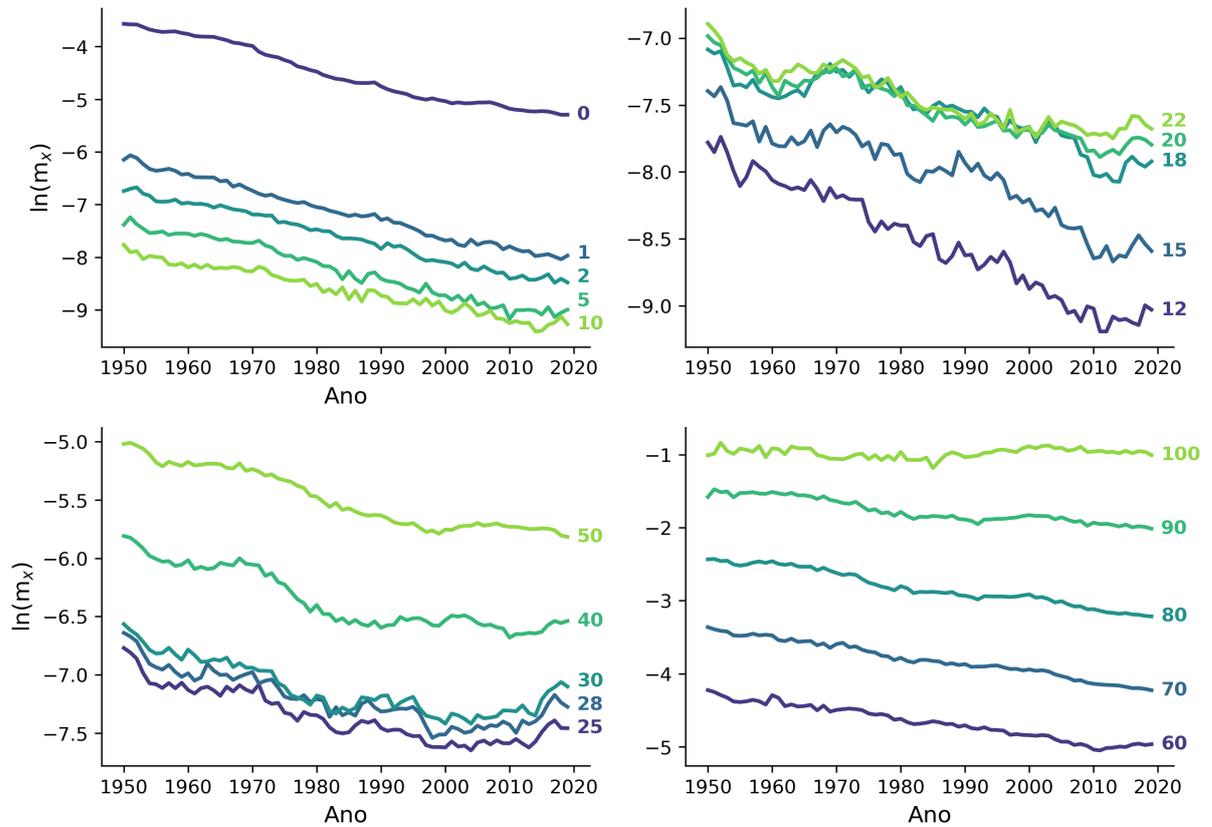


Figura 14 – Logaritmo das taxas de mortalidade por idade da população total americana.

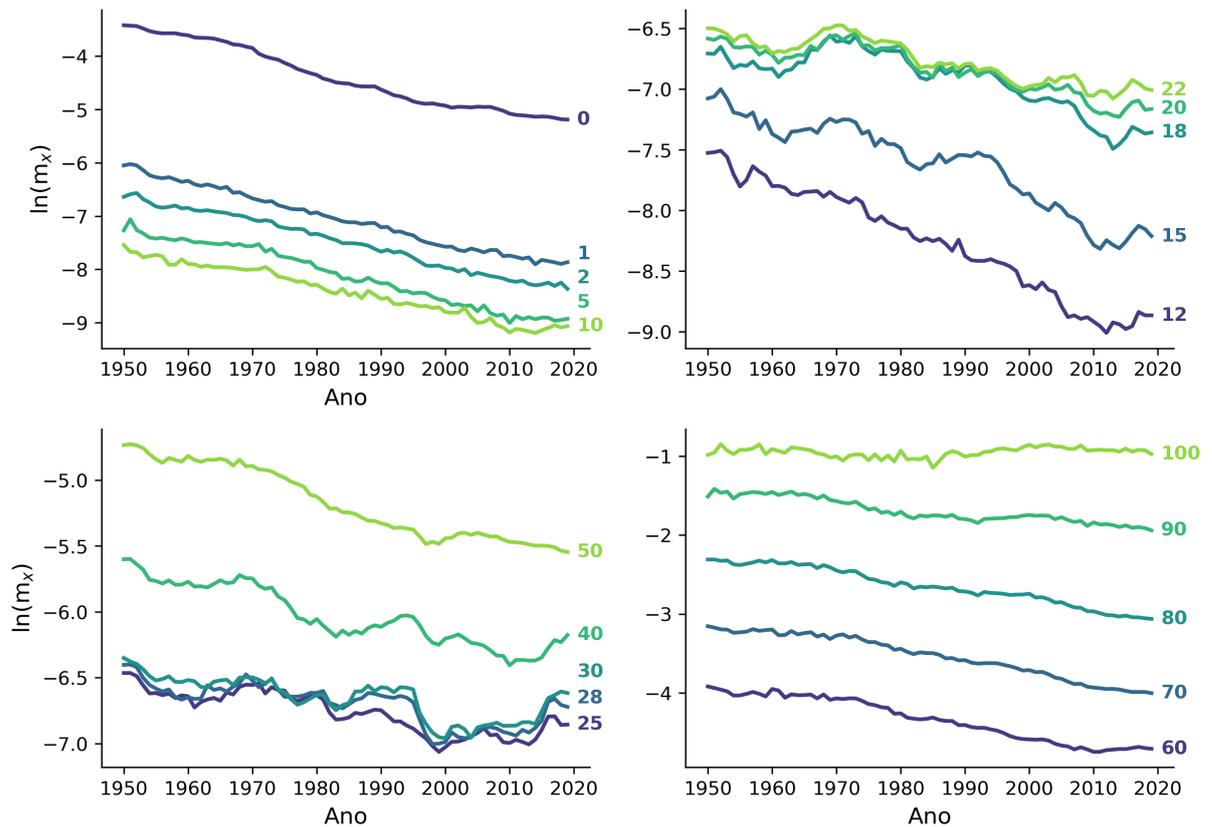


Figura 15 – Logaritmo das taxas de mortalidade por idade da população masculina francesa.

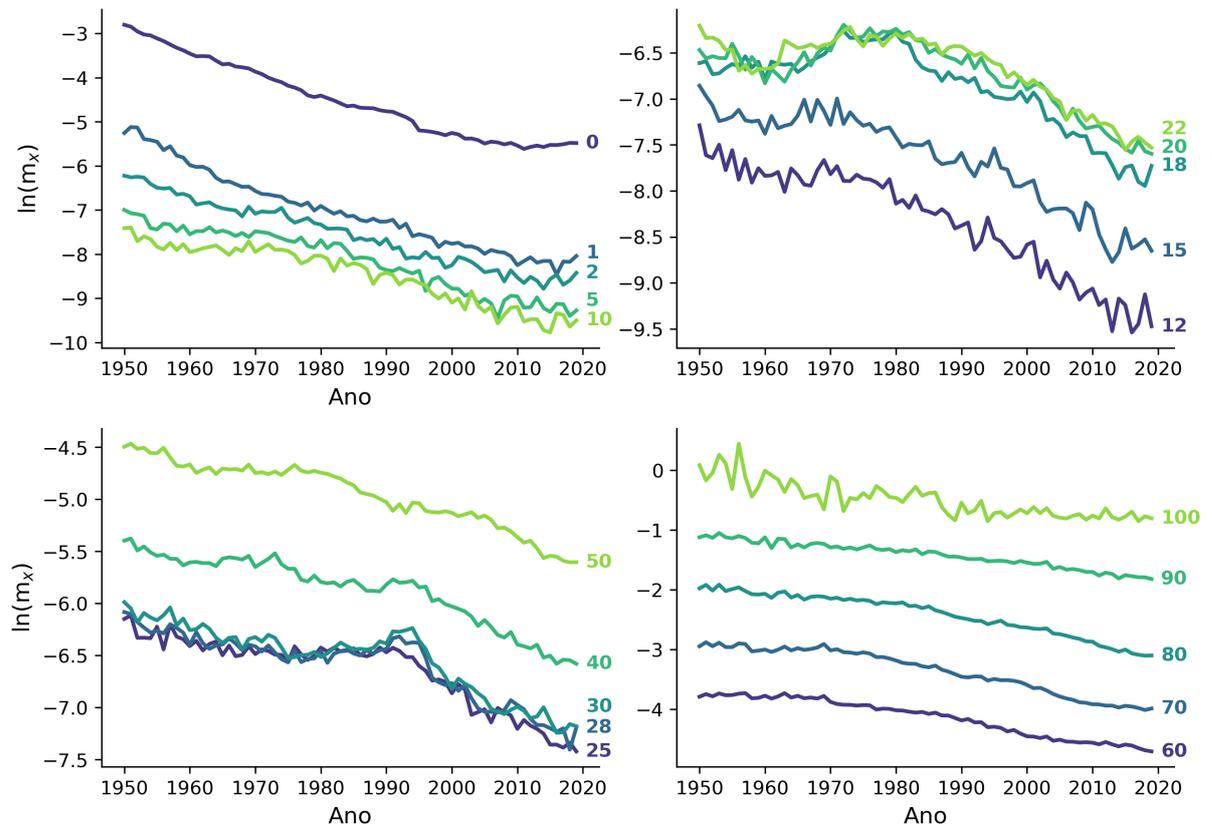


Figura 16 – Logaritmo das taxas de mortalidade por idade da população feminina francesa.

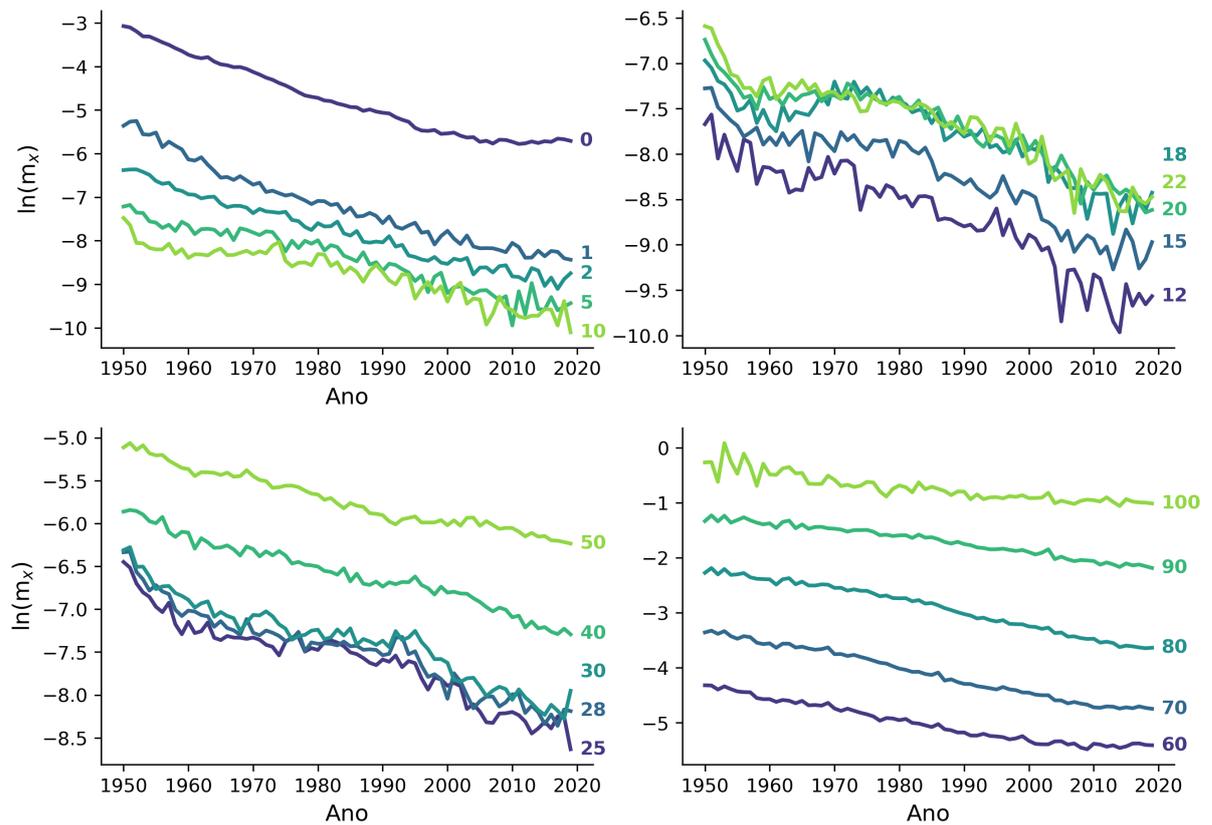


Figura 17 – Logaritmo das taxas de mortalidade por idade da população total francesa.

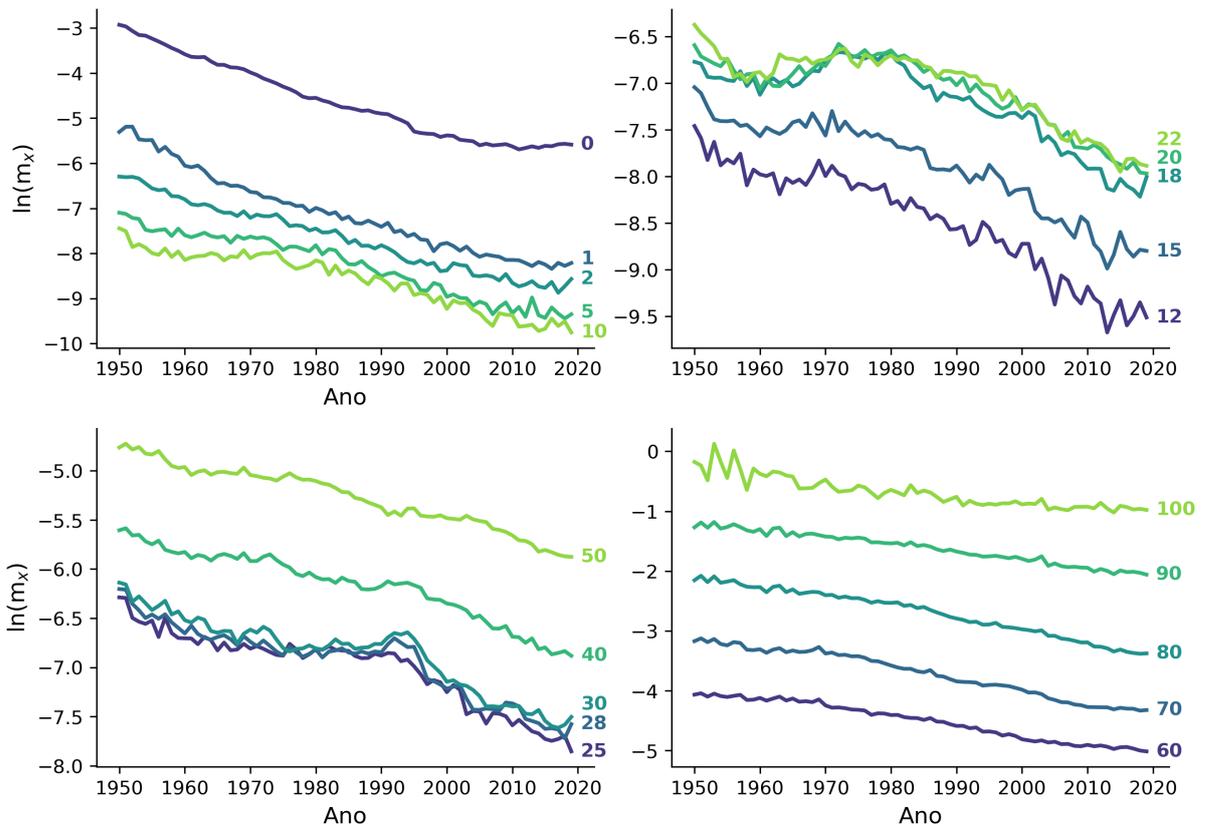


Figura 18 – Logaritmo das taxas de mortalidade por idade da população masculina japonesa.

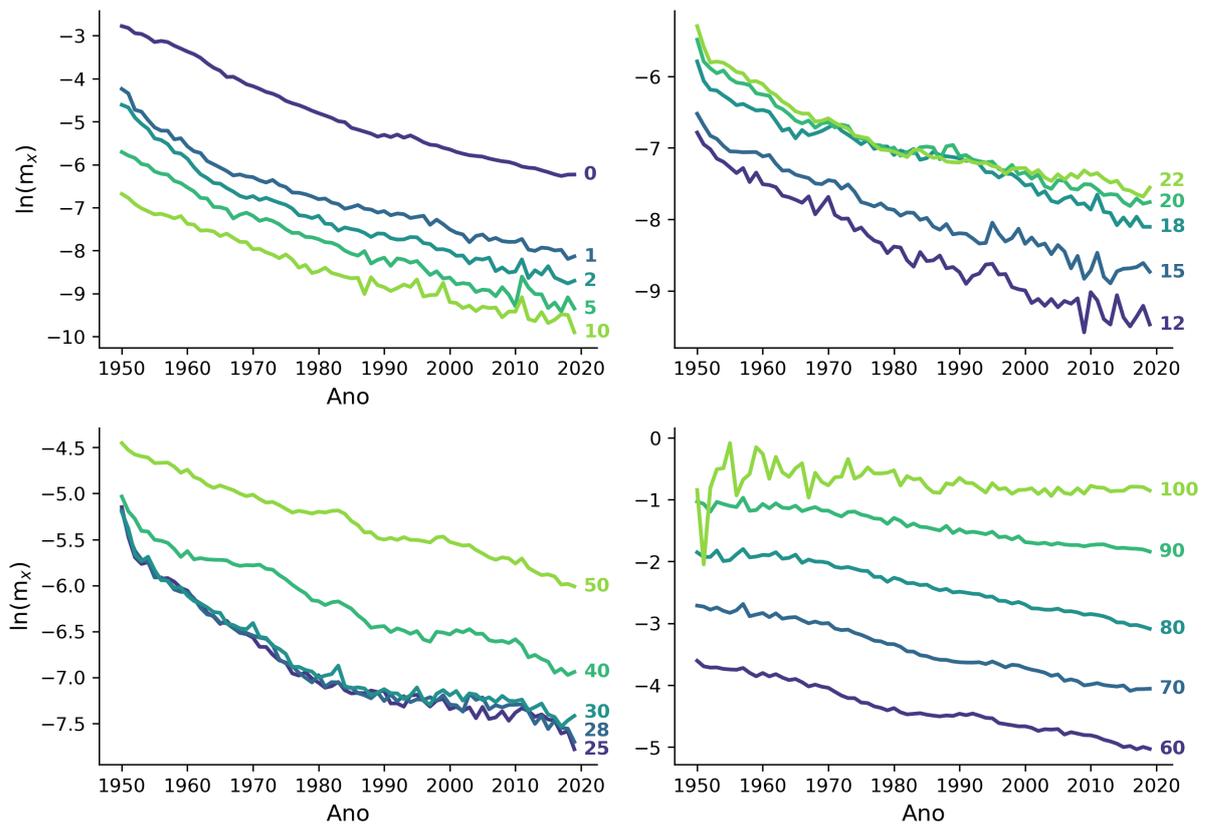


Figura 19 – Logaritmo das taxas de mortalidade por idade da população feminina japonesa.

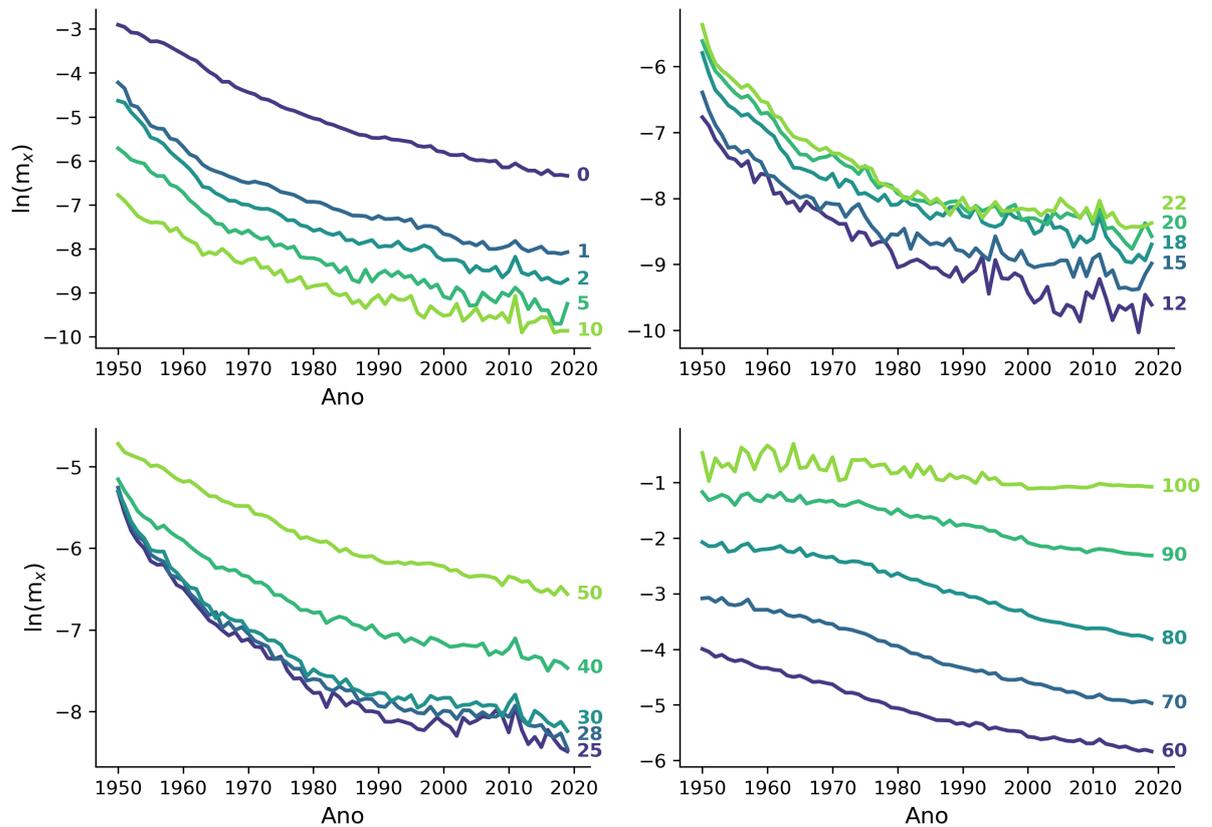


Figura 20 – Logaritmo das taxas de mortalidade por idade da população total japonesa.

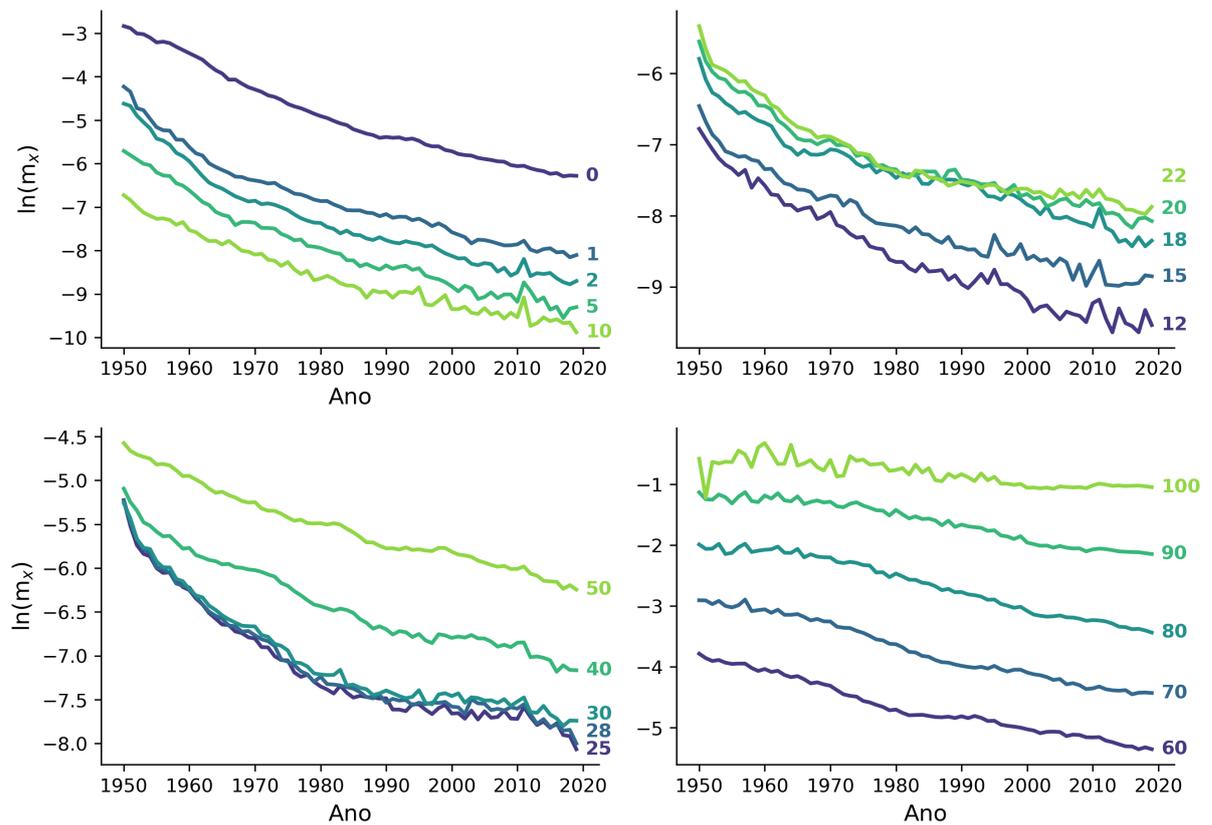


Figura 21 – Logaritmo das taxas de mortalidade por idade da população masculina portuguesa.

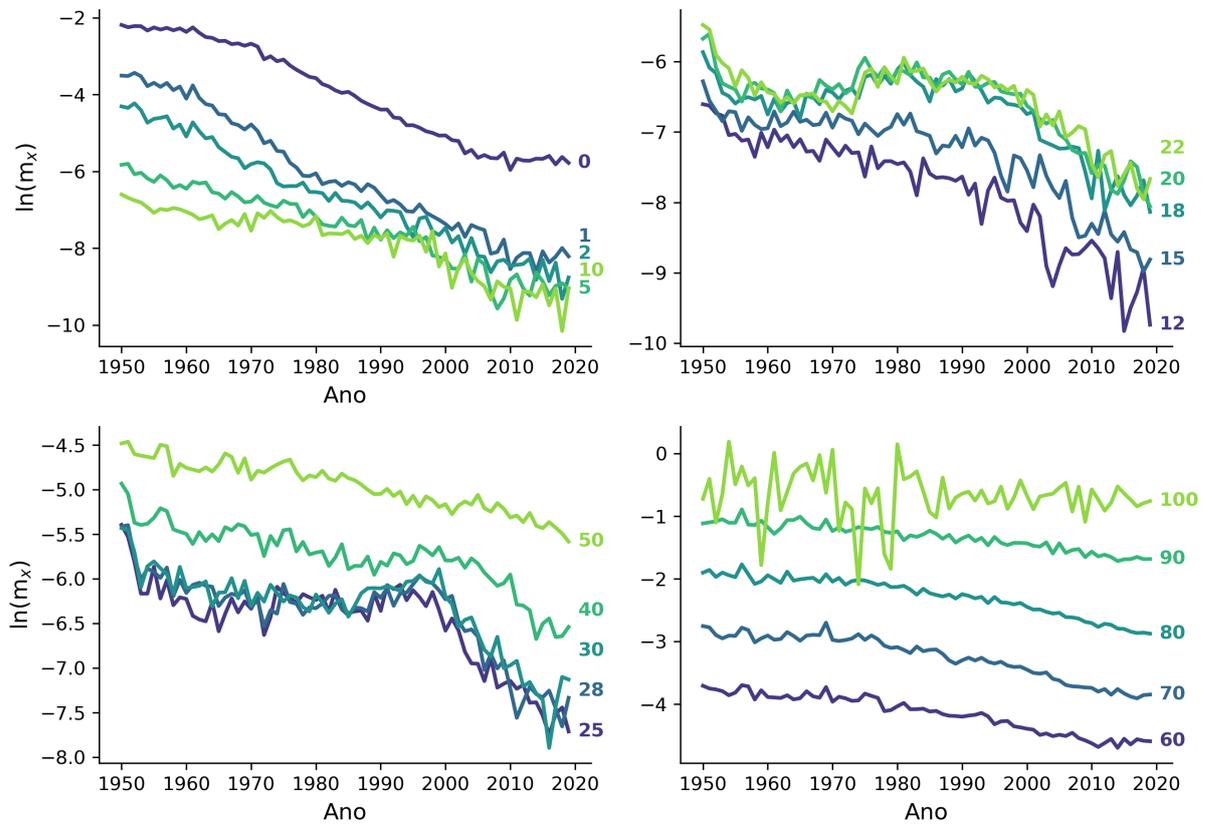


Figura 22 – Logaritmo das taxas de mortalidade por idade da população feminina portuguesa.

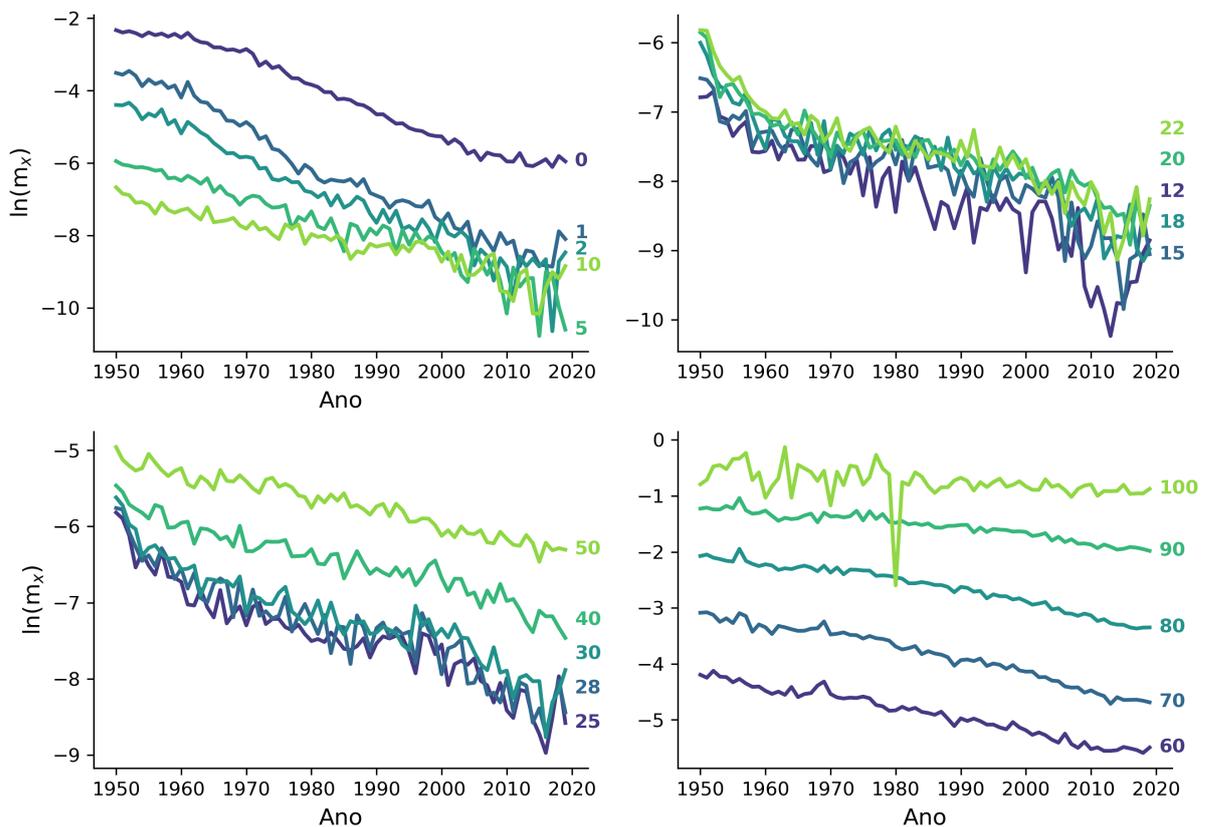
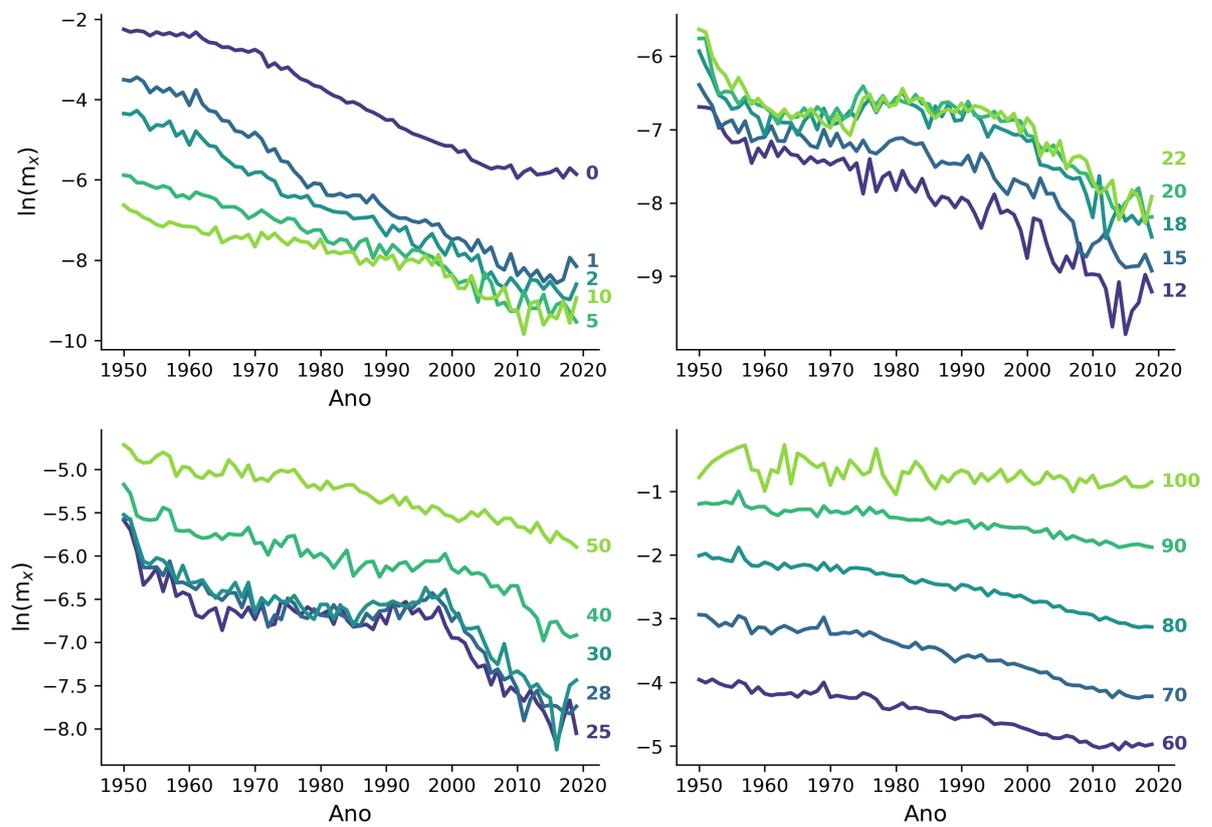


Figura 23 – Logaritmo das taxas de mortalidade por idade da população total portuguesa.



Fonte: Elaborado pelo autor (2024).

## 4.2 MÉTRICAS DE AVALIAÇÃO

Tradicionalmente, as métricas de erro em análises de séries temporais são calculadas para uma única série ao longo de um intervalo temporal específico. No entanto, este estudo adota uma abordagem mais abrangente, considerando múltiplas séries temporais simultaneamente. A metodologia de avaliação empregada neste trabalho apresenta as seguintes características:

1. Avaliação por Horizonte Temporal: Para cada horizonte de previsão  $h$ , os pontos de previsão englobam valores correspondentes a todas as idades (cada idade representa uma série temporal) consideradas no estudo.
2. Avaliação Multidimensional: O erro é calculado abrangendo todas as idades e horizontes temporais presentes no conjunto de teste.

As métricas MAPE e MASE foram utilizadas para avaliar o desempenho preditivo do sistema híbrido proposto. No que diz respeito à seleção do modelo não-linear na etapa SM do sistema proposto, optou-se pela métrica MAPE. Essas métricas são independentes da escala dos dados, e usualmente são escolhidas pelos trabalhos de previsão de séries temporais (Hyndman; Koehler, 2006; Cerqueira et al., 2017) tal como de previsão de mortalidade (Pascariu; Lenart; Canudas-Romo, 2019; Duarte; de Mattos Neto; Firmino, 2024), pois é por meio delas que se pode avaliar a acurácia em termos percentuais (MAPE) ou em relação ao random walk (MASE), o que facilita a comparação de modelos em diferentes conjuntos de dados.

O MAPE é calculado como sendo o percentual de erro médio da previsão e é independente da escala dos valores (Hyndman; Koehler, 2006). Por sua vez, ele denota um valor percentual em termos dos erros cometidos, assim como desconsidera os seus sentidos. O MAPE é interpretado de maneira que quanto menor o seu valor, melhor a acurácia de previsão. Desse modo, é uma margem de erro que pode ser usada para comparar desempenho preditivo em diversos conjuntos de dados, auxiliando o processo de tomada de decisão em múltiplos contextos, como por exemplo na avaliação de solvência de seguradoras. O valor do erro fornecido pelo MAPE é interpretado associadamente ao impacto financeiro gerado nas avaliações de riscos das operações de seguros. O MAPE é calculado por meio da Equação 4.1 descrita a seguir:

$$\text{MAPE} = \frac{1}{K} \sum_{i=1}^K \left| \frac{z_i - \hat{z}_i}{z_i} \right|, \quad (4.1)$$

em que  $K$  é o número total de observações do conjunto avaliado,  $z_i$  a  $i$ -ésima observação do conjunto, e  $\hat{z}_i$  é a previsão para  $z_i$  realizada pelo modelo.

O MASE é calculado como sendo o erro absoluto médio de previsão em comparação ao erro absoluto cometido pelo modelo *naïve* no conjunto de treinamento (Hyndman; Koehler, 2006). Entretanto, nesta pesquisa optou-se por utilizar o erro do modelo *naïve* no conjunto de teste, visto que as previsões são geradas em um cenário de múltiplos passos à frente. No que tange à lógica de interpretação para essa métrica, é possível empregar o mesmo raciocínio feito para o MAPE, ou seja, quanto mais próximo o valor do MASE de zero, maior a acurácia do modelo avaliado.

$$\text{MASE} = \frac{\frac{1}{K} \sum_{t=1}^K |z_t - \hat{z}_t|}{\frac{1}{K-1} \sum_{l=2}^K |z_l - z_{l-1}|}, \quad (4.2)$$

em que  $K$  é a quantidade de horizontes de tempo do conjunto de teste,  $z_t$  é o valor observado da série temporal no tempo  $t$ , e  $z_t - z_{t-1}$  é o erro do modelo *naïve* (i.e., *random-walk*) no conjunto de teste.

Ainda, o *Root Mean Squared Error* (RMSE) foi utilizado para otimizar os modelos de AM empregados no estudo. A fórmula matemática do RMSE pode ser vista a seguir na Equação 4.3:

$$\text{RMSE} = \sqrt{\frac{1}{K} \sum_{i=1}^K (z_i - \hat{z}_i)^2}. \quad (4.3)$$

em que,  $K$  é o número total de observações do conjunto de validação,  $z_i$ , a  $i$ -ésima observação, e  $\hat{z}_i$  é a  $i$ -ésima previsão realizada pelo modelo.

### 4.3 PREVISÃO DAS TAXAS DE MORTALIDADE E ANÁLISE ESTATÍSTICA

O sistema proposto foi aplicado aos dados das taxas de mortalidade dos quinze conjuntos de dados, e o desempenho de previsão foi avaliado com outros modelos univariados e multivariados da literatura.

Para os modelos univariados, o conjunto de dados de treinamento foi constituído pelo logaritmo natural das taxas de mortalidade de uma população para os seguintes grupos etários: 0, 1, 2, 5, 10, 12, 15, 18, 20, 22, 25, 28, 30, 40, 50, 60, 70, 80, 90, e 100 anos (Wu; Wang, 2018). Após a realização da previsão para o conjunto de teste desses modelos, conforme (Wu; Wang, 2018), foi gerada uma curva de mortalidade para as idades 0 a 100 anos considerando o conjunto de teste. A curva de mortalidade para um ano futuro foi gerada a partir das 20

idades supracitadas. A interpolação *spline* cúbica é utilizada para gerar a curva de mortalidade completa na escala logarítmica. A interpolação foi aplicada ao sistema proposto, ao sistema híbrido para múltiplos passos baseado em (Zhang, 2003), aos modelos ARIMA e aos modelos de AM. Para os modelos multivariados, os grupos etários são compostos pelas idades entre 0 a 100 anos. Optou-se por restringir o conjunto de dados até a idade de 100 anos para evitar comportamento errático para as idades maiores do que 100 anos (Lam; Wang, 2021), as quais compreendem a maior parte do banco de dados utilizado.

Os conjuntos de dados foram divididos em duas amostras (Wu; Wang, 2018; Lam; Wang, 2021). A primeira amostra, denominada conjunto de treinamento, compreende os dados entre os anos 1950 e 2009, exceto para o país Austrália, a qual optou-se pelo ano inicial de 1921. A segunda amostra, chamada conjunto de teste, é composta pelos valores entre 2010 e 2019.

Para os modelos MLP, LSTM e N-BEATS, a divisão de dados considerou os primeiros 80% da primeira amostra para treinamento e os 20% restantes para validação dos hiperparâmetros. Em todos os experimentos com modelos de AM, tanto a série temporal quanto a residual foram normalizadas no intervalo  $[0, 1]$ . A Equação 4.4 denota a normalização supracitada,

$$z_{norm_t} = \frac{z_t - z_{min}}{z_{max} - z_{min}}, \quad (4.4)$$

em que  $z_t$  é o valor observado da série temporal ou do resíduo do preditor linear sob estudo no momento  $t$ ,  $z_{min}$  é o valor mínimo da série temporal ou residual, e  $z_{max}$  é o valor máximo da série temporal ou residual.

Na fase de previsão para os dados de teste, o sistema híbrido proposto foi comparado com outros modelos da literatura: ARIMA, Lee-Carter (LC) (Lee; Carter, 1992), Plat (Plat, 2009), três sistemas híbridos, ARIMA-MLP (Zhang, 2003), ARIMA-LSTM (Xu et al., 2022) e ARIMA-N-BEATS com três tipos de abordagens multi-step: recursivo, MIMO e direta (considerando híbrido ARIMA-N-BEATS de abordagem direta como HyS-MF (Duarte; de Mattos Neto; Firmino, 2024)), seis modelos de AM monolíticos (MLP, LSTM, e N-BEATS) usando as abordagens recursiva (Atiya et al., 1999; Nigri et al., 2019; Beyaztas; Shang, 2022), direta (Atiya et al., 1999; Kline, 2004; Livieris; Pintelas, 2022; Oreshkin et al., 2019), e MIMO (Kline, 2004; Bontempi, 2008; Chandra; Goyal; Gupta, 2021; Livieris; Pintelas, 2022). Isto posto, ao todo foram avaliados 60 modelos de previsão, compondo modelos individuais (estatísticos, de AM e multivariados) e híbridos (univariados e multivariados).

Os modelos multivariados LC e PLAT foram executados com as bibliotecas StMoMo (Villegas; Kaishev; Millossovich, 2018) e forecast (Hyndman et al., 2024) da linguagem de

programação R<sup>3</sup>. O modelo ARIMA foi obtido a partir do algoritmo `auto_arima` implementado na biblioteca `pmdarima` (Smith et al., 2017–) na linguagem de programação Python, que é baseado no algoritmo de Hyndman-Khandakar (Hyndman; Khandakar, 2008)<sup>4</sup>. Além disso, a abordagem recursiva foi usada para prever os múltiplos passos à frente, e todos os sistemas híbridos deste estudo utilizaram a metodologia de modelagem linear acima mencionada.

As bibliotecas `TensorFlow` (Abadi et al., 2015) e `Keras` foram usados para treinar as redes neurais MLP, LSTM e N-BEATS. Dez execuções independentes foram realizadas para o MLP, LSTM e N-BEATS para equilibrar a considerável sensibilidade resultante da inicialização de pesos das RNAs. Para validação dos hiperparâmetros tomou-se como critério de desempenho o RMSE. A fim de otimizar os hiperparâmetros e selecionar a configuração de parâmetros do modelo não-linear, foi projetada a otimização Bayesiana (Mockus; Vytautas; Zilinskas, 1978), que é recomendada para problemas cujo tempo de execução é longo (Frazier, 2018; Snoek; Larochelle; Adams, 2012), baseada nos valores apresentados na Tabela 1, adaptado dos recentes trabalhos de previsão de sistemas híbridos (Santos Júnior; Oliveira; de Mattos Neto, 2019; Oliveira; Silva; de Mattos Neto, 2021). Os valores apresentados para 'Unidades na camada escondida' e 'Taxa de aprendizagem', especificamente [5, 100] e [1E-4, 1E-1], devem ser interpretados como intervalos fechados, onde o segundo foi amostrado em escala logarítmica. Para cada modelo de AM realizaram-se 10 ensaios, considerando que cada ensaio representa uma combinação de hiperparâmetros, e para cada ensaio executou-se independentemente a mesma configuração 5 vezes.

O sistema proposto foi comparado com abordagens como os modelos multivariados da literatura e suas versões hibridizadas (incorporando sistemas híbridos no lugar do modelo ARIMA). Foi investigada a capacidade preditiva dos modelos Lee-Carter (Lee; Carter, 1992) e Plat (Plat, 2009). Ressalta-se que esses modelos multivariados realizam previsão de múltiplos passos para toda a curva de mortalidade. O MAPE e o MASE foram calculados para toda a

<sup>3</sup> O código-fonte utilizado nesta pesquisa está disponível no repositório público em: <<https://github.com/filipecluarte/tese>>.

<sup>4</sup> Durante o experimento, avaliou-se a presença de dependência serial nos resíduos dos preditores gerados pelo ARIMA, por meio dos testes Brock-Dechert-Scheinkman (BDS) (Broock et al., 1996) e método *surrogate* com entropia de permutação (Zunino; Kulp, 2017). Os resultados não foram incorporados ao texto principal devido à sua natureza inconclusiva, tendo em vista que para uma quantidade significativa de casos onde se detectou evidências de padrão serial, a consequente modelagem dos resíduos deteriorou o sistema de previsão. O contrário também foi observado, ou seja, na ausência de evidências de dependência serial nas séries dos resíduos, a modelagem residual incorporada ao sistema híbrido produziu um sistema mais acurado. Uma das razões para esses resultados contraditórios são a quantidade considerável de ruído contida nos resíduos e o tamanho amostral insuficiente, o que pode ter comprometido a qualidade das inferências estatísticas. Em virtude disso, assumiu-se a hipótese de que os resíduos dos preditores do modelo ARIMA contém dependência serial não-linear.

Tabela 1 – Valores dos hiper-parâmetros para os modelos não-lineares

Modelo	Parâmetro	Valor
MLP	Número máximo de iterações	500
	Unidades na camada de entrada	2
	Unidades na camada escondida	[2, 100]
	Função de ativação	Tangente hiperbólica, ReLU
	Otimizador	Adam
	Taxa de aprendizagem	[1E-4, 1E-1]
LSTM	Número máximo de iterações	500
	Unidades na camada de entrada	2
	Unidades na camada escondida	[2, 100]
	Função de ativação	Tangente hiperbólica, sigmoidla e linear (recorrente)
	Otimizador	Adam
	Taxa de aprendizagem	[1E-4, 1E-1]
N-BEATS	Número máximo de iterações	500
	Unidades na camada de entrada	2
	Quantidade de <i>blocks</i> por <i>stack</i>	4
	Quantidade de camadas escondidas	[1, 2, 3, 4]
	Tipo da <i>stack</i>	Genérica
	Unidades na camada escondida	[2, 100]
	Otimizador	Adam
	Taxa de aprendizagem	[1E-4, 1E-1]

Fonte: Elaborada pelo autor (2024).

curva de mortalidade em cada horizonte de previsão do conjunto de teste.

Essa comparação foi realizada em três etapas: (i) análise de regressão<sup>5</sup> que avalia o efeito médio de cada modelo sobre o erro de previsão, (ii) comparação da acurácia de previsão (média e por rank) entre o sistema proposto e demais modelos para todos os horizontes de previsão de forma segregada por conjunto de dados, e (iii) aplicação de teste estatístico para quantificar se há diferenças no desempenho entre o sistema proposto e os melhores modelos de cada abordagem (i.e., estatística, híbrida e multivariada).

A comparação de todos os modelos de previsão foi avaliada por meio de um modelo de regressão linear generalizado<sup>6</sup> (McCullagh; Nelder, 1989), em que as observações dos dados

<sup>5</sup> A adoção de um modelo de regressão para análise dos resultados experimentais oferece, nesta pesquisa, uma metodologia estatística para avaliar o desempenho comparativo dos modelos de previsão. Assim, é possível quantificar as diferenças médias no erro de previsão entre os modelos, controlando simultaneamente por fatores exógenos potencialmente confundidores. Para uma compreensão aprofundada dos fundamentos teóricos e aplicações de modelos de regressão, recomenda-se consultar Dunn, Smyth et al. (2018), Gelman, Hill e Vehtari (2021) e Montgomery, Peck e Vining (2021).

<sup>6</sup> Um modelo linear generalizado é a generalização do modelo de regressão linear, que possui três componentes: variável dependente ( $y$ ) gerada a partir de uma distribuição da família exponencial, função de ligação ( $g(\cdot)$ ) e preditor linear ( $X\beta$ ) (McCullagh; Nelder, 1989). A equação do modelo pode ser descrita como:  $g(\mu) = X\beta$ .

são compostos pelas informações do MAPE ou MASE da curva de mortalidade para cada ano de previsão e modelo do conjunto de teste. A equação do modelo de regressão está denotada

$$\ln(MAPE)_i = \alpha + \sum_{j=1}^{m-1} \beta_j D_{modelo_j} + \gamma_1 Horizonte_i + \gamma_2 Total_i + \gamma_3 Masculino_i + \gamma_4 EUA_i + \gamma_5 França_i + \gamma_6 Japão_i + \gamma_7 Portugal_i + \varepsilon_i,$$

em que  $i$  é o índice da observação de um modelo de previsão; o logaritmo natural foi utilizado como a função de ligação para relacionar a variável dependente ( $MAPE$ ) e o preditor linear;  $\alpha$  é o intercepto;  $D_{modelo_j}$  é a variável dicotômica que assume o valor 1 para o  $j$ -ésimo modelo de previsão e 0 caso contrário,  $\beta_j$  é o coeficiente da variável  $D_{modelo_j}$ ;  $Horizonte_i$  é o horizonte de previsão (i.e., compreendido entre  $[0, \dots, 9]$ );  $Total_i$  é a variável dicotômica que assume o valor 1 no caso do gênero da  $i$ -ésima observação ser composta por ambos os sexos masculinos e feminino, e 0 caso contrário;  $Masculino_i$  é a variável dicotômica que assume o valor 1 no caso do gênero ser masculino e 0 no caso contrário;  $EUA_i$  é a variável dicotômica que assume o valor 1 no caso da população ser proveniente dos EUA e 0 no caso contrário;  $França_i$  é a variável dicotômica que assume o valor 1 no caso da população ser proveniente da França e 0 no caso contrário;  $Japão_i$  é a variável dicotômica que assume o valor 1 no caso da população ser proveniente do Japão e 0 no caso contrário;  $Portugal_i$  é a variável dicotômica que assume o valor 1 no caso da população ser proveniente do Portugal e 0 no caso contrário;  $\gamma$  é o coeficiente das variáveis de controle; e  $\varepsilon$  é o erro do modelo de regressão.

Essas variáveis de controle são essenciais para ajustar possíveis fatores confundidores que possam influenciar o  $MAPE$ , proporcionando uma especificação mais precisa do modelo. O controle sistemático dessas variáveis (e.g., país, grupo demográfico e ano) assegura que as diferenças observadas no  $MAPE$  sejam atribuíveis às características intrínsecas dos modelos de previsão, eliminando potenciais vieses decorrentes de variações demográficas ou temporais. Dessa forma, é possível determinar, do ponto de vista estatístico, qual modelo de previsão apresenta o menor Erro de previsão, medido pelo MAPE ou MASE.

Após a avaliação do modelo de regressão, realizou-se uma análise comparativa do desempenho na previsão da curva de mortalidade entre o sistema proposto, os modelos híbridos de abordagem recursiva, HyS-MF, ARIMA, LC e Plat. Essa análise compreendeu a avaliação (por meio da métrica MAPE e seu ranque) para todos os conjuntos de dados de forma agregada e isolada, mostrando os valores para todos horizontes do conjunto de teste. A Diferença Percentual (DP) das métricas de erro MAPE e MASE, para o conjunto de teste, conforme mostrado

na equação 4.5, foi utilizada para comparar a média da previsão da curva de mortalidade gerada pelos modelos em relação ao sistema proposto.

$$DP = 100 \cdot \frac{\text{Erro}_a - \text{Erro}_b}{\text{Erro}_a}, \quad (4.5)$$

no qual Erro é a métrica de erro,  $a$  corresponde ao modelo em comparação, e  $b$  é o sistema proposto.

O teste estatístico Nemenyi *multiple comparisons with the best* (MCB) (McDonald; Jr, 1967; Koning et al., 2005) foi aplicado aos valores dos erros absolutos para verificar se o desempenho do sistema proposto é estatisticamente similar (valor- $p \geq 0,05$ ) às alternativas encontradas na literatura. Desse modo, ele foi empregado para averiguar se o sistema proposto apresentou o menor erro, em absoluto, de previsão do ponto de vista estatístico em comparação a todos os demais modelos (i.e., levando em conta os dois melhores modelos de cada classe a partir dos resultados experimentais - monolíticos de AM, sistemas híbridos, multivariados), considerando todos os conjuntos de dados simultaneamente.

Para avaliar a eficiência computacional dos algoritmos de AM utilizados neste estudo, o tempo de execução foi considerado como uma métrica adicional além da acurácia das previsões. Para cada algoritmo, o tempo total médio de treinamento e teste foi calculado. A média do tempo foi obtida considerando tanto o treinamento quanto a geração de previsões para todos os conjuntos de dados analisados. Dentre os algoritmos avaliados, o tempo médio foi avaliado exclusivamente para os modelos univariados, devido ao seu maior custo computacional, uma vez que cada série temporal requer o desenvolvimento de um modelo independente. Em contraste, os modelos multivariados demonstram maior eficiência nesse contexto, uma vez que utilizam um único algoritmo para cada conjunto de dados, reduzindo consideravelmente o tempo de processamento.

Finalmente, as previsões para a curva de mortalidade geradas pelo sistema proposto foram avaliadas sob o aspecto da plausibilidade biológica, ou seja, uma análise dos padrões da curva ao longo das diferentes idades. Isso é crucial porque a mortalidade humana exibe um padrão conhecido como "J" (Gage, 1989; Milne, 2010; Richmond et al., 2021), e é essencial que as projeções reflitam esse padrão sistemático.

#### 4.4 CONSIDERAÇÕES CAPÍTULO

Neste capítulo foi apresentado o protocolo experimental desenvolvido neste trabalho. A pesquisa tá organizada em 3 etapas. Cada etapa busca alcançar um ou mais objetivos específicos da pesquisa. O sistema proposto utilizado como objeto de estudo é apresentado. Na primeira etapa, realizou-se a previsão das taxas de mortalidade para vários passos futuros para todos os modelos. Em seguida, comparou-se, por meio de um modelo estatístico de regressão, o desempenho preditivo do sistema proposto em relação a todos os modelos da literatura detalhados na seção. Na segunda análise, foi comparada a curva de mortalidade predita pelos modelos selecionados para cada conjunto de dados ao longo de todos os horizontes e também realizou-se teste estatístico a partir dos erros em absoluto. Na última etapa, aplicou-se teste estatístico para avaliar se o sistema proposto demonstrou melhor desempenho de previsão em comparação aos modelos com maior acurácia.

Ainda neste capítulo, foi apresentado o protocolo experimental. As séries temporais utilizadas foram descritas, assim como o pré-processamento e os valores percentuais para o conjunto de treinamento e de teste. Em seguida, foram apresentadas as métricas de avaliação, o RMSE foi utilizado para validação dos hiperparâmetros e as métricas MAPE e MASE foram empregadas para avaliar os modelos de previsão e responder as questões de pesquisa, além de ter a função de realizar futuras comparações com outros trabalhos da literatura. Em seguida, foi apresentada as configurações do sistema proposto, tais como os hiperparâmetros utilizados nas configurações dos modelos de AM e o método de otimização desses hiperparâmetros.

Por fim, foi apresentado o modelo de regressão e o teste estatístico os quais têm o objetivo de avaliar se o sistema proposto é superior estatisticamente às abordagens da literatura.

## 5 RESULTADOS

Este capítulo apresenta os resultados mais relevantes obtidos nesta pesquisa. O capítulo está organizado em quatro seções. A Seção 5.1 contém os resultados e a discussão do modelo de regressão estimado a partir dos dados experimentais observados. A seção 5.2 apresenta uma comparação do desempenho entre os modelos univariados e multivariados. A seção 5.3 mostra uma comparação do desempenho de previsão entre os sistema híbrido proposto, modelos híbridos, ARIMA, Lee-Carter (LC), Plat, modelos monolíticos de AM e multivariados híbridos. Na seção 5.5 é realizada uma avaliação e discussão da plausibilidade biológica das previsões taxas de mortalidade realizadas pelo sistema proposto.

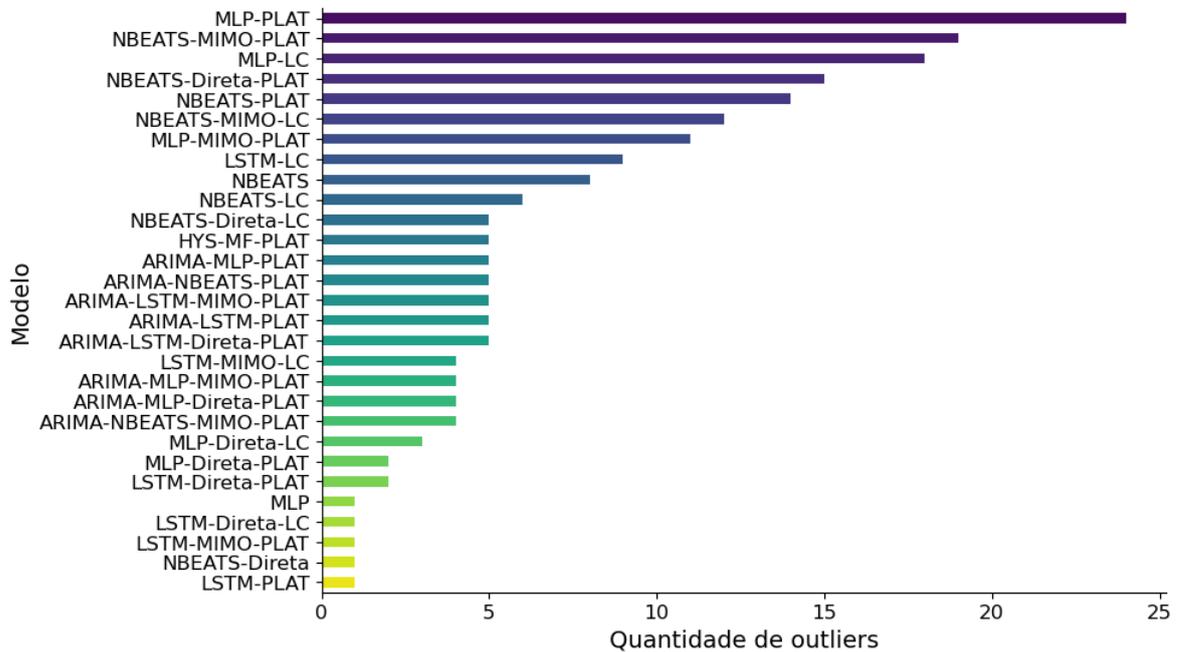
Os resultados experimentais estão divididos em quatro partes. Em primeiro lugar, analisam-se os resultados estatísticos do experimento por meio de um modelo linear generalizado de regressão (McCullagh; Nelder, 1989), o qual compara todos os modelos de previsão em relação ao modelo Lee-Carter, tal como avalia o impacto dos horizontes de previsão, dos países e do gênero no MAPE<sup>1</sup>. para as curvas de mortalidade e todos os anos do conjunto de teste. Em segundo lugar, compara-se o desempenho de previsão da curva de mortalidade para os quinze conjuntos de dados, considerando o ranque e a diferença percentual do MAPE e do MASE e testes estatísticos ponderando o melhor sistema híbrido em relação aos demais modelos de séries temporais da literatura. Em seguida, é apresentada uma avaliação da plausibilidade biológica das previsões das taxas de mortalidade a partir do comportamento das projeções em três etapas, isto é, no decrescimento inicial em  $J$  das primeiras idades; no súbito crescimento na idade adulto jovem; e na linearidade crescente da faixa etária idosa.

### 5.1 ANÁLISE DOS RESULTADOS DA REGRESSÃO

Para sintetizar os resultados experimentais e avaliar os efeitos de cada variável controlada do experimento, estimou-se um modelo linear generalizado de regressão gama, a partir da função de ligação logarítmica, tendo em vista à adequação da assimetria positiva da distribuição do MAPE. Para tanto, realizou-se uma remoção de *outliers* a partir da avaliação da amplitude interquartil (IQR = terceiro quartil - primeiro quartil), a qual determina que um ponto fora da

<sup>1</sup> Os resultados da regressão gama para o MASE foram semelhantes aos do MAPE e por isso decidiu-se por detalhar nesta seção apenas os dados do MAPE. Os resultados referentes ao MASE estão apresentados no Apêndice A

Figura 24 – Distribuição dos *outliers* em relação ao MAPE para os modelos avaliados no experimento considerando o conjunto de teste



Fonte: Dados da pesquisa (2024).

curva (*outlier*) seria aquele cujo valor estivesse a menor ou a maior de  $1,5 \cdot IQR$  em relação ao 1º e 3º quartis, respectivamente. Tendo em vista que a distribuição dos dados do MAPE é positivamente assimétrica, a aplicação do filtro de *outliers* removeu apenas valores a maior que 12,53 do MAPE, contabilizando 199 observações de um total de 9000<sup>2</sup>. Os *outliers* decorreram primordialmente dos modelos monolíticos de AM e dos multivariados e suas variações híbridas como mostra a Figura 24, a qual exibe a quantidade de observações consideradas como *outliers* para os modelos do conjunto de dados resultante do experimento.

A partir dessa situação, observou-se que o modelo de regressão gama apresentou ajuste adequado a partir da avaliação dos valores dos desvios dos resíduos. Para obter o valor do efeito de cada variável, aplicou-se a função exponencial e subtraiu-se pelo valor unitário. Ou seja, se considerarmos o coeficiente estimado para um modelo de regressão com função de ligação logarítmica uma variável qualquer por  $\beta$ , o efeito percentual na variável dependente (MAPE) é obtido por meio da equação a seguir,  $efeito = \exp(\beta) - 1$ . Desse modo, foi possível avaliar o efeito da variação percentual diretamente no MAPE para cada variável. Para o intercepto, bastou aplicar a função exponencial para se obter a média dos parâmetros de controle, tais como modelo Lee-Carter (LC), país Austrália e gênero feminino. O valor médio

<sup>2</sup> No que diz respeito aos dados computados para a métrica MASE, observaram-se 334 observações *outliers*.

do MAPE estimado dessas variáveis foi de 3,0959%<sup>3</sup>. Portanto, por padrão, essas foram as variáveis a serem comparadas com as demais variáveis do experimento.

A Figura 25 apresenta os efeitos percentuais das variáveis modelos, horizonte de previsão (i.e., 0, 1, ..., 9), país (Austrália, EUA, França, Japão e Portugal), gênero (Feminino, Masculino e Total) no MAPE (%) do experimento. Todas as variáveis do modelo de regressão foram estatisticamente significativas a 5%, exceto os modelos LC-ARIMA, MLP-MIMO, MLP-DIR. Da figura supracitada, observou-se que o país com menor MAPE médio foi os EUA, o qual apresentou uma redução de 26,6% do MAPE em comparação ao país Austrália<sup>4</sup>. Em contrapartida, o país com maiores erros de previsão do experimento foi Portugal, o qual auferiu um aumento de 29,2% no MAPE em relação à Austrália<sup>5</sup>. O horizonte de previsão apresentou efeito de 0,038, o que indica que a cada aumento unitário de horizonte de previsão, aumenta-se o MAPE médio em 3,8%, controlando as demais variáveis. Esse resultado é relevante para mostrar como quanto mais distante o horizonte de previsão, maior será o erro de previsão esperado mensurado pelo MAPE, evidenciando, portanto, a complexidade da tarefa de previsão de múltiplos passos à frente. O gênero total mostrou-se o mais fácil na tarefa de previsão de logaritmo das taxas de mortalidade, com redução percentual de 24% em comparação ao gênero feminino. Por outro lado, o gênero masculino mostrou-se o mais difícil, aumentando o MAPE em torno de 19,5% em relação ao feminino.

A Figura 26 mostra o efeito médio do sistema proposto e dos modelos monolíticos sobre o MAPE. O modelo que apresentou o melhor desempenho médio foi o Sistema proposto cujo efeito percentual no MAPE foi de -0,495, ou seja, em relação ao LC, ele demonstrou uma redução de 49,5% no valor do MAPE. O modelo ARIMA apresentou valor estimado de -0,452, ou seja, reduziu o MAPE em torno de 45,2% em relação ao modelo LC. Os demais modelos monolíticos apresentaram aumento médio no MAPE em comparação ao Lee-Carter. A Figura 27 demonstra os efeitos médios no MAPE em comparação ao LC dos sistemas híbridos de abordagem *multistep* recursiva. O ARIMA-LSTM mostrou o segundo melhor resultado de acurácia preditiva média, com efeito médio de -48%. Em terceiro lugar, pode-se verificar o ARIMA-MLP, o qual evidenciou efeito no MAPE de -47,7%. Em quarto lugar, destacou-se o ARIMA-NBEATS com efeito médio no MAPE de -46,6%. Os demais sistemas híbridos

<sup>3</sup> Enquanto que para o MASE foi de 1,7536.

<sup>4</sup> No modelo com o MASE, as variáveis não significativas estatisticamente foram LC-ARIMA, LSTM, MLP, N-BEATS, N-BEATS-DIR.

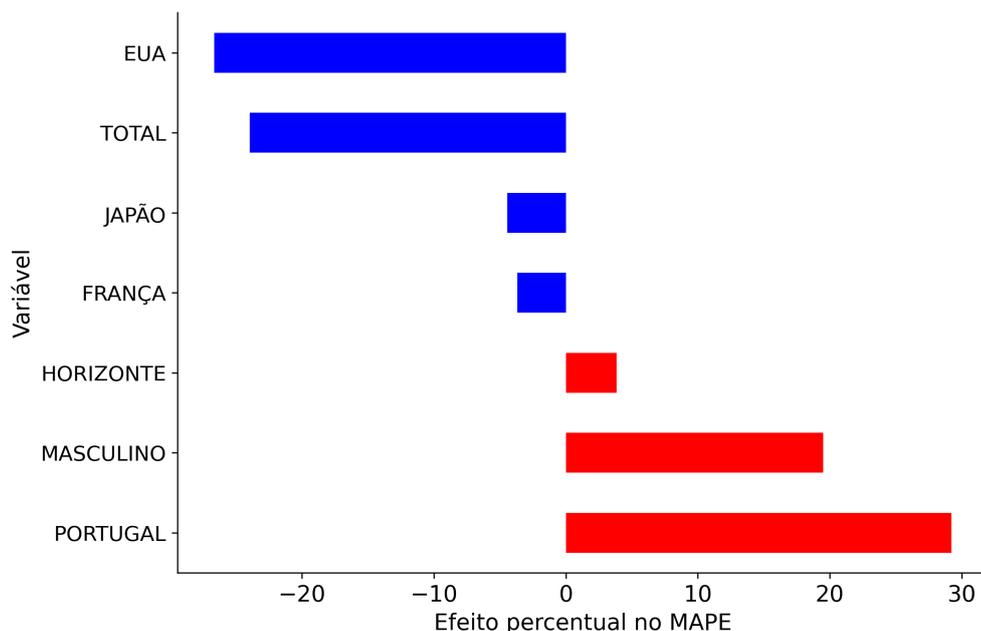
<sup>5</sup> Para o modelo com o MASE, observou-se uma notável diferença em relação ao efeito dos países na métrica de erro, uma vez que nessa medida todos os países, exceto Portugal, demonstraram um MASE superior ao da Austrália. Portugal, portanto, foi o país com o menor MASE, em média.

recursivos combinados com os modelos multivariados LC e PLAT apresentaram efeito de aumento no MAPE de pelo menos 59,8%, comparado ao LC.

A Figura 28 mostra o efeito médio no MAPE para os sistemas híbridos de abordagens MIMO e Direta. O proposto (HyS-MF) apresentou efeito médio estimado de -0,451. Ainda, notou-se que os modelos híbridos univariados ARIMA-LSTM-MIMO, ARIMA-LSTM-DIRECT demonstraram efeito equivalente de -46,6%, ao passo que ARIMA-MLP-MIMO e ARIMA-MLP-DIRECT exibiram efeitos de -46,4% e -45,7%, respectivamente. Dentre os sistemas híbridos univariados, o modelo com o menor efeito redutor no MAPE foi o ARIMA-NBEATS-MIMO. Os modelos multivariados híbridos de abordagem MIMO e Direta apresentaram MAPE médio superior ao modelo LC.

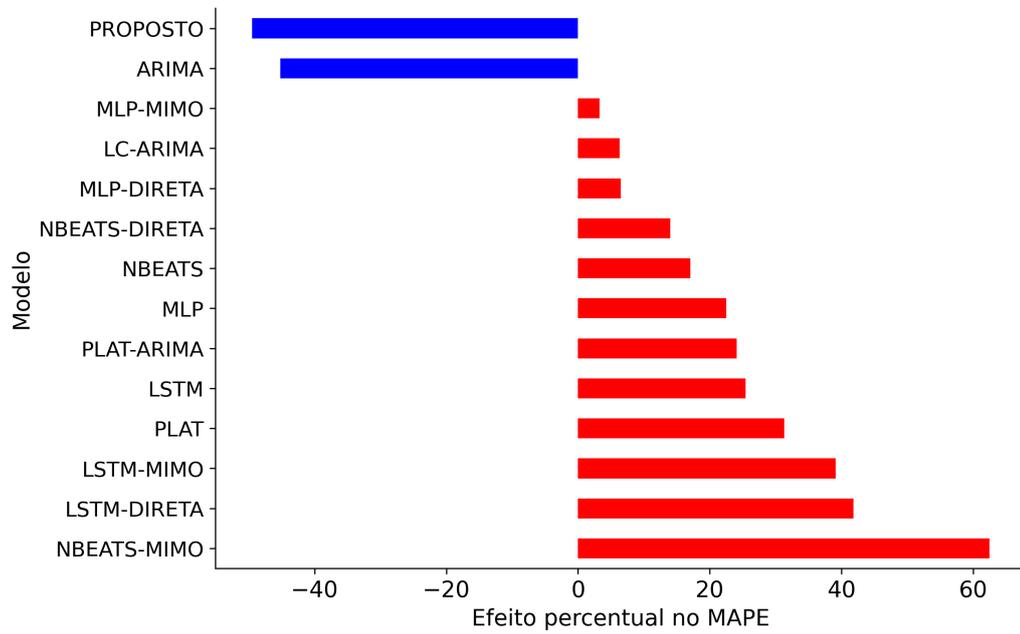
Diante disso, o resultado da regressão mostrou que os modelos monolíticos de AM e os sistemas híbridos aplicados aos modelos LC e Plat não demonstraram um ganho de acurácia de previsão em comparação com as abordagens da literatura. Ao contrário, notou-se que, para todos os casos, houve um aumento do efeito no MAPE (%), ou seja, uma piora na acurácia de previsão. Desse modo, é possível destacar que as tentativas de aprimoramento dos modelos multivariados clássicos da literatura falharam, pois não houve ganhos médios na capacidade preditiva para os modelos e sistemas híbridos empregados nesta pesquisa.

Figura 25 – Efeito médio dos países, horizonte de previsão e gênero no MAPE para a curva de mortalidade entre 2010-2019. Modelo Lee-Carter, gênero feminino e país Austrália são os efeitos capturados pelo intercepto (MAPE médio de 3,0959%) que é comparável aos demais modelos, países e gêneros.



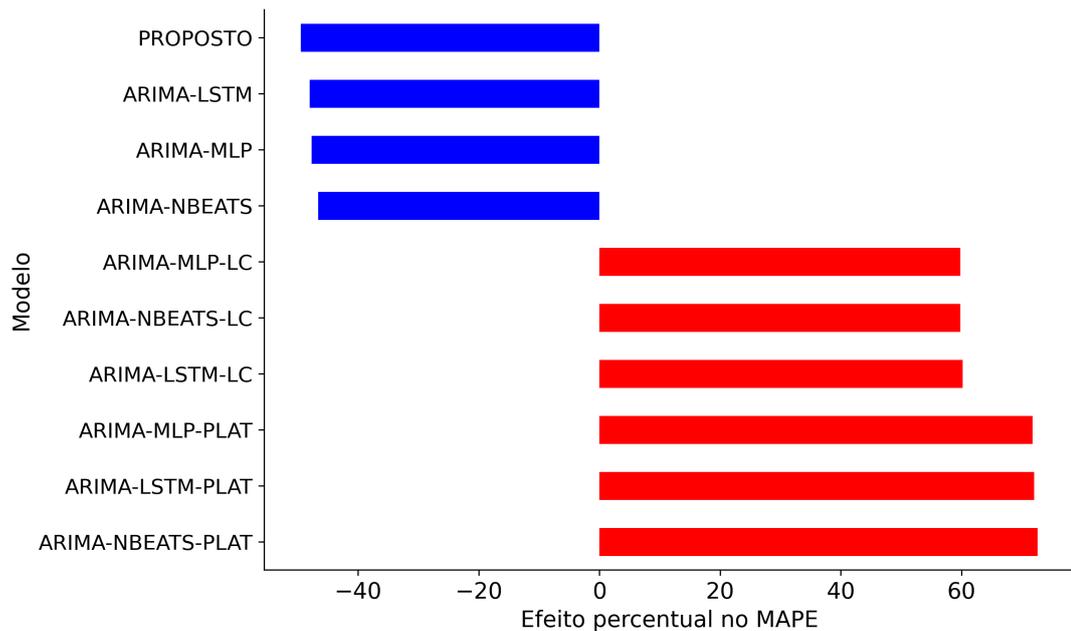
Fonte: Dados da pesquisa (2024).

Figura 26 – Efeito médio do Sistema proposto e dos modelos monolíticos no MAPE para a curva de mortalidade entre 2010-2019.



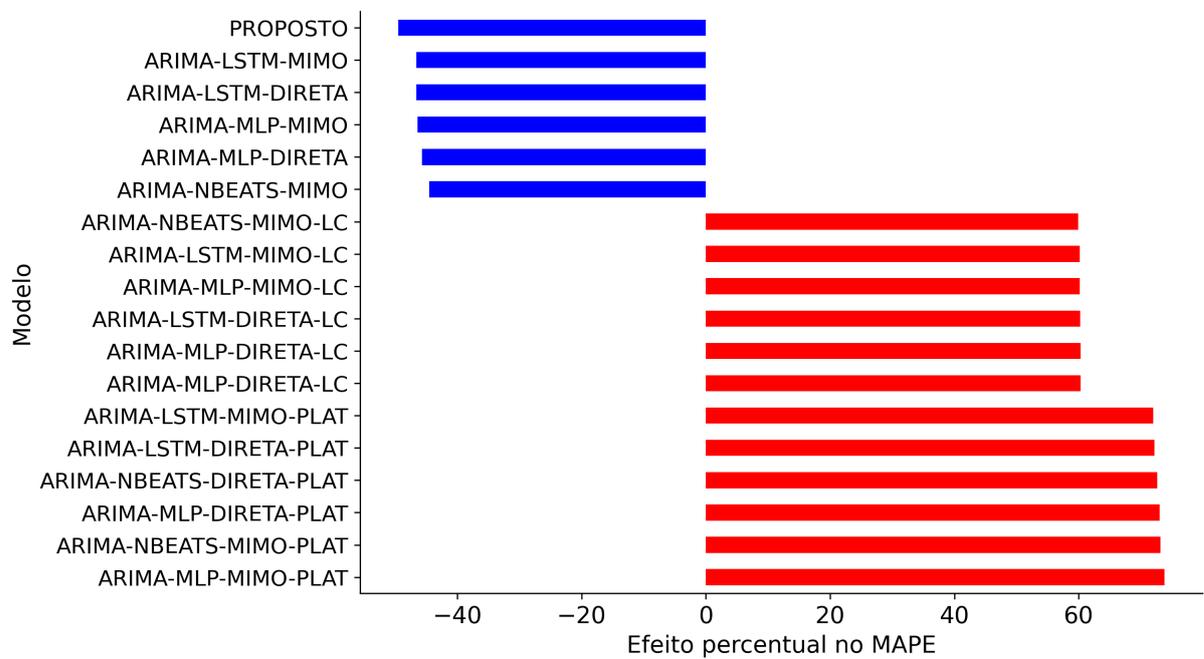
Fonte: Dados da pesquisa (2024).

Figura 27 – Efeito médio do Sistema proposto e dos demais modelos híbridos univariados de abordagem recursiva no MAPE para a curva de mortalidade entre 2010-2019.



Fonte: Dados da pesquisa (2024).

Figura 28 – Efeito médio do Sistema proposto e dos modelos híbridos univariados e multivariados de abordagens MIMO e direta no MAPE para a curva de mortalidade entre 2010-2019.

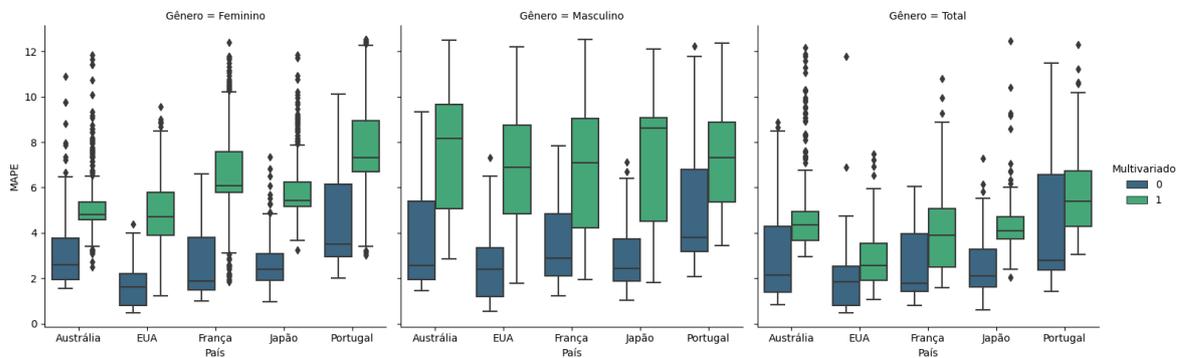


Fonte: Dados da pesquisa (2024).

## 5.2 COMPARAÇÃO ENTRE ABORDAGENS UNIVARIADAS E MULTIVARIADAS

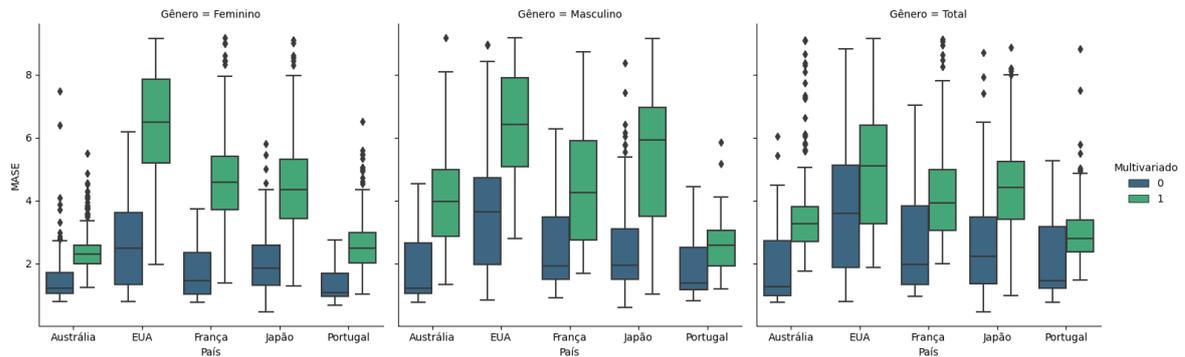
A Figura 29 exibe os boxplots do MAPE segregados por gênero, país e categoria do modelo (multivariado ou univariado), ao passo que a Figura 30 mostra os boxplots do MASE. A partir dessas figuras, nota-se que os modelos univariados alcançaram os menores quartis para o MAPE e MASE em todas as categorias de gênero, notadamente para as categorias masculino e feminino. Em relação aos países, percebeu-se a maior diferença entre os valores do 3º quartil da classe univariada em relação ao 1º quartil do multivariado das distribuições do MAPE e MASE nos EUA, exceto para o gênero total, onde a maior discrepância foi observada para o Japão. Diante desses achados, é possível inferir que, considerando os modelos de previsão, países e período de avaliação, os modelos univariados apresentam menores valores para o MAPE e MASE em comparação com os modelos multivariados clássicos e híbridos. Ademais, os resultados evidenciados acima estão em consonância com as evidências da literatura (Wu; Wang, 2018; Feng; Shi, 2018).

Figura 29 – Boxplot do MAPE segregado por gênero, país e a categoria de se o modelo de previsão é multivariado.



Fonte: Dados da pesquisa (2024).

Figura 30 – Boxplot do MASE segregado por gênero, país e a categoria de se o modelo de previsão é multi-variado.

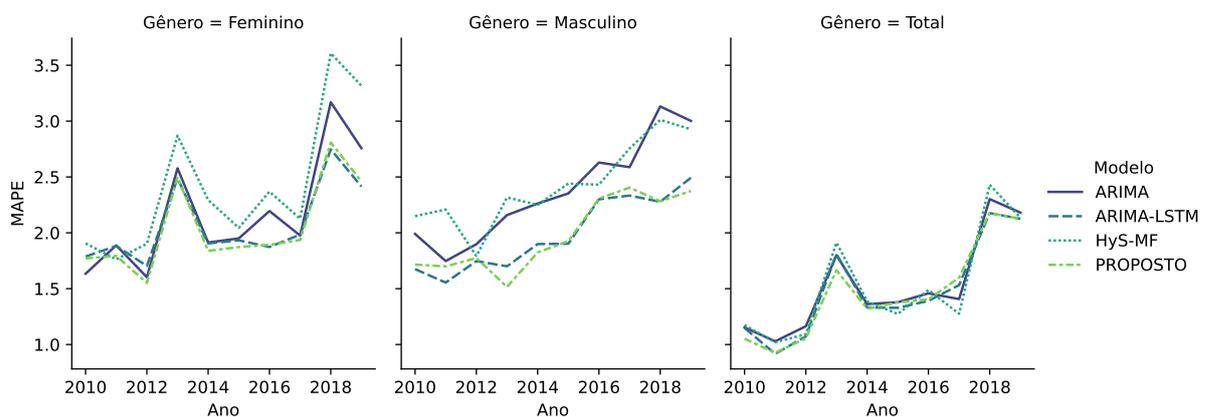


Fonte: Dados da pesquisa (2024).

### 5.3 COMPARAÇÃO DO DESEMPENHO DE PREVISÃO DAS CURVAS DE MORTALIDADE

As Figuras 31-35 apresentam o MAPE para cada horizonte de previsão<sup>6</sup>, demonstrando a evolução dessa métrica no tempo para o sistema proposto e os modelos ARIMA, ARIMA-LSTM e HyS-MF, segregado pelos conjuntos de dados, isto é por país e gênero<sup>7</sup>. Dessa figura, nota-se, na maior parte dos casos, uma tendência de crescimento do MAPE ao longo dos horizontes em todos os conjuntos de dados.

Figura 31 – MAPE (%) do Sistema proposto, ARIMA, ARIMA-LSTM e HyS-MF para a Austrália entre os anos 2010-2019

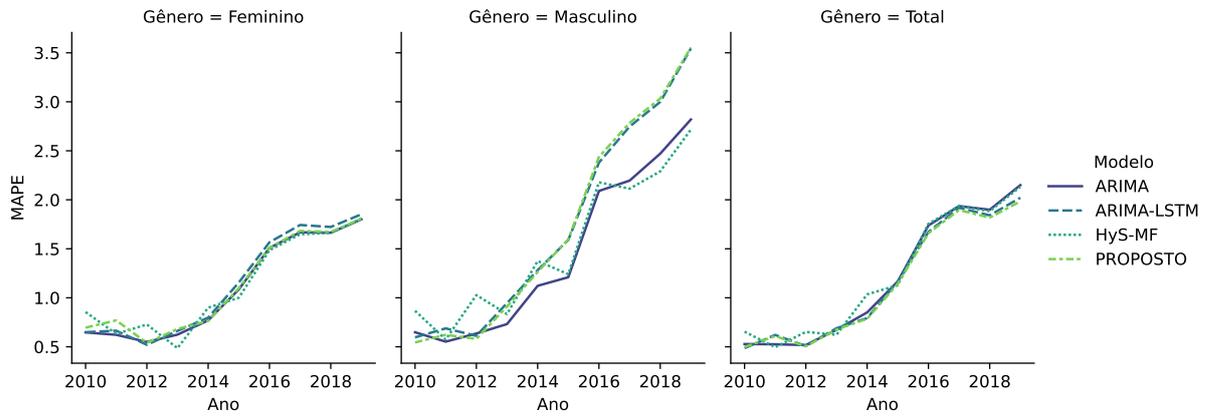


Fonte: Dados da pesquisa (2024).

<sup>6</sup> Nesta análise, levou-se em conta o valor do MAPE para todas as idades em cada horizonte de tempo do conjunto de teste.

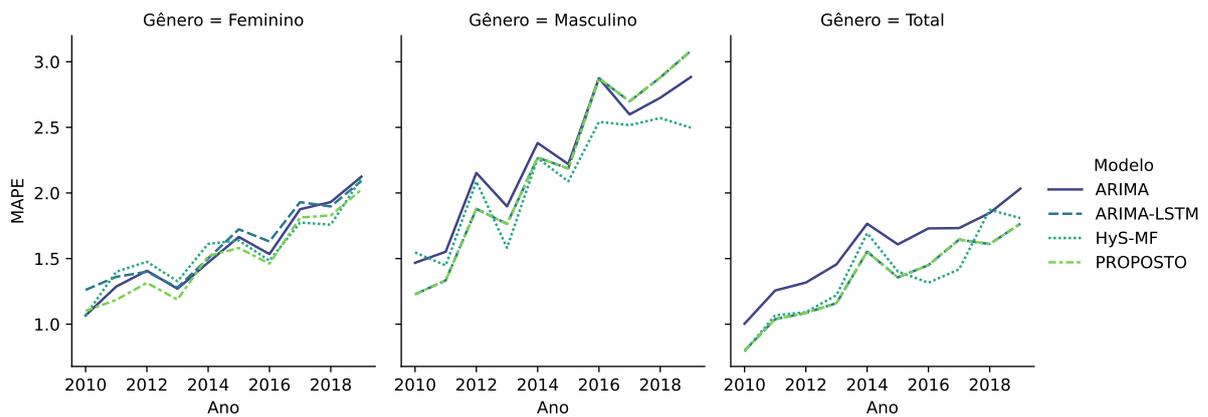
<sup>7</sup> No Apêndice B consta a Figura para o MASE.

Figura 32 – MAPE (%) do Sistema proposto, ARIMA, ARIMA-LSTM e HyS-MF para os Estados Unidos entre os anos 2010-2019



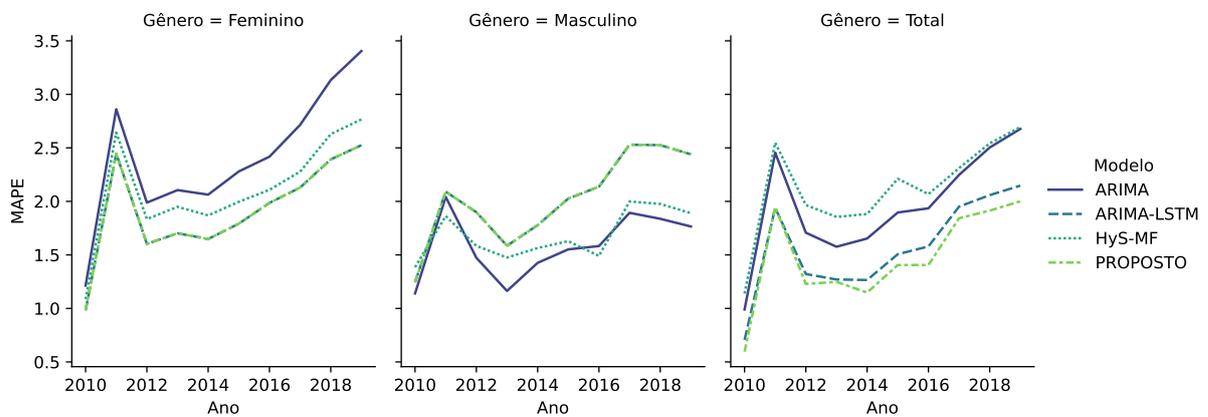
Fonte: Dados da pesquisa (2024).

Figura 33 – MAPE (%) do Sistema proposto, ARIMA, ARIMA-LSTM e HyS-MF para a França entre os anos 2010-2019



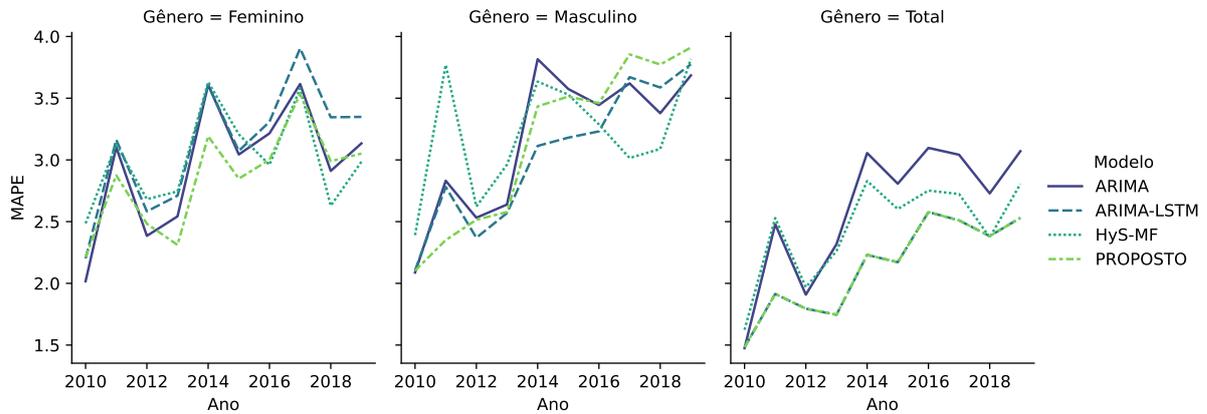
Fonte: Dados da pesquisa (2024).

Figura 34 – MAPE (%) do Sistema proposto, ARIMA, ARIMA-LSTM e HyS-MF para o Japão entre os anos 2010-2019



Fonte: Dados da pesquisa (2024).

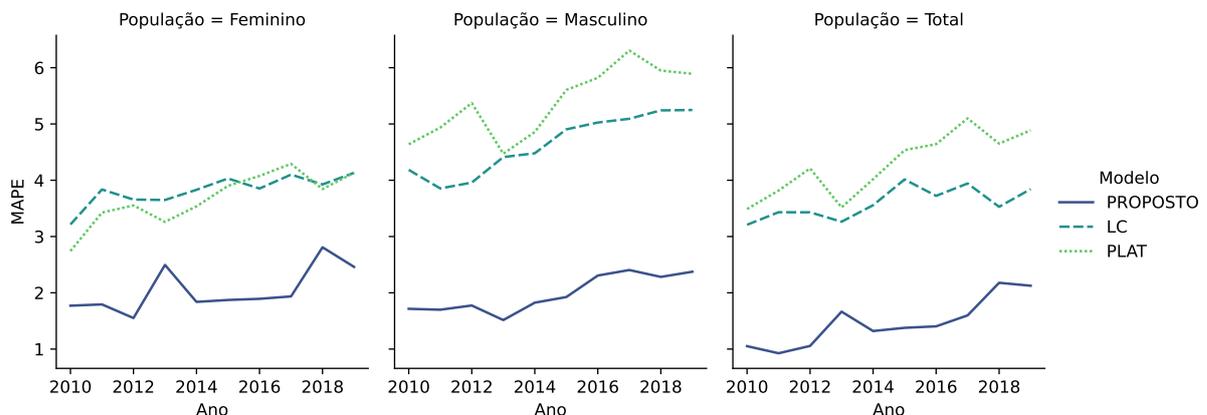
Figura 35 – MAPE (%) do Sistema proposto, ARIMA, ARIMA-LSTM e HyS-MF para Portugal entre os anos 2010-2019



Fonte: Dados da pesquisa (2024).

As Figuras 36-40 são semelhantes às Figuras 31-35, contudo aquelas apresentam os modelos LC e Plat em comparação ao sistema proposto<sup>8</sup>. É possível notar uma diferença no nível do MAPE para os modelos LC e Plat em comparação com o sistema proposto, no qual este modelo apresentou MAPE inferior em relação aos primeiros, exceto para França masculino e Portugal feminino. Nestes casos, o LC mostra nível do MAPE equivalente ao sistema proposto em alguns horizontes de previsão. Notadamente, na França masculino verificou-se que o MAPE entre o sistema proposto e o LC se aproximou mais contundentemente apenas nos últimos 4 horizontes de previsão. Em contrapartida, para Portugal feminino, percebeu-se que os MAPEs desses dois modelos foram mais próximos nos anos 2011, 2014 e 2017.

Figura 36 – MAPE (%) do Sistema proposto, LC e PLAT para a Austrália entre os anos 2010-2019

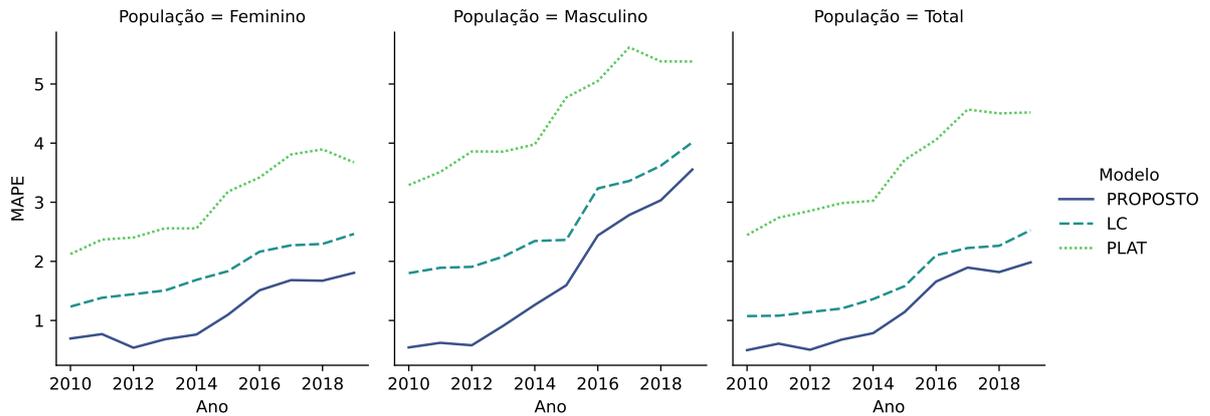


Fonte: Dados da pesquisa (2024).

As Figuras 41-45 mostram o MAPE ao longo dos horizontes de previsão para os modelos univariados ARIMA e sistemas híbridos por modelo não-linear (MLP, LSTM e N-BEATS) e

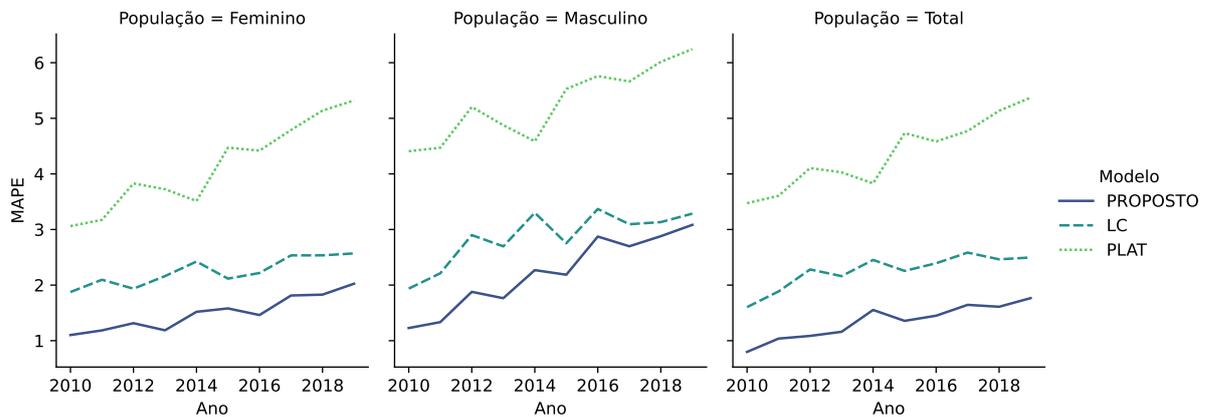
<sup>8</sup> Apêndice C exibe a figura com os dados do MASE.

Figura 37 – MAPE (%) do Sistema proposto, LC e PLAT para os EUA entre os anos 2010-2019



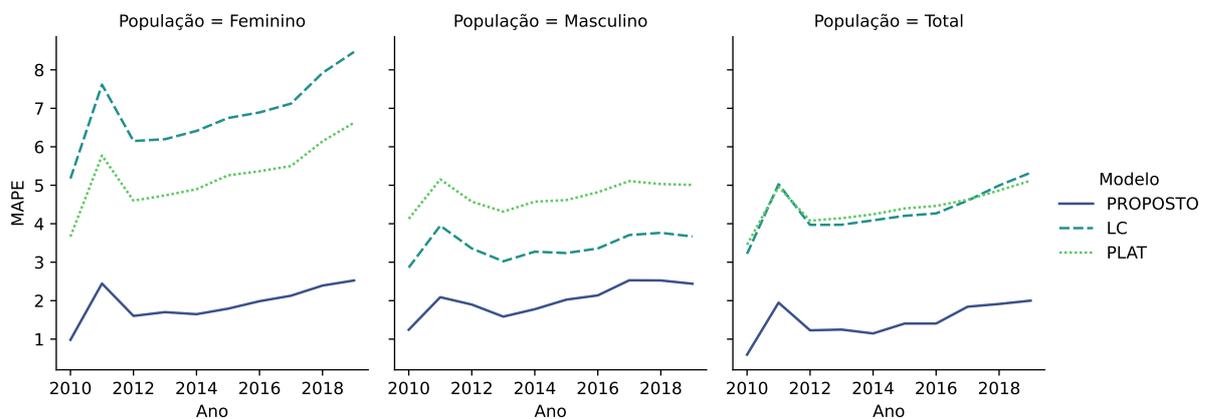
Fonte: Dados da pesquisa (2024).

Figura 38 – MAPE (%) do Sistema proposto, LC e PLAT para a França entre os anos 2010-2019



Fonte: Dados da pesquisa (2024).

Figura 39 – MAPE (%) do Sistema proposto, LC e PLAT para o Japão entre os anos 2010-2019

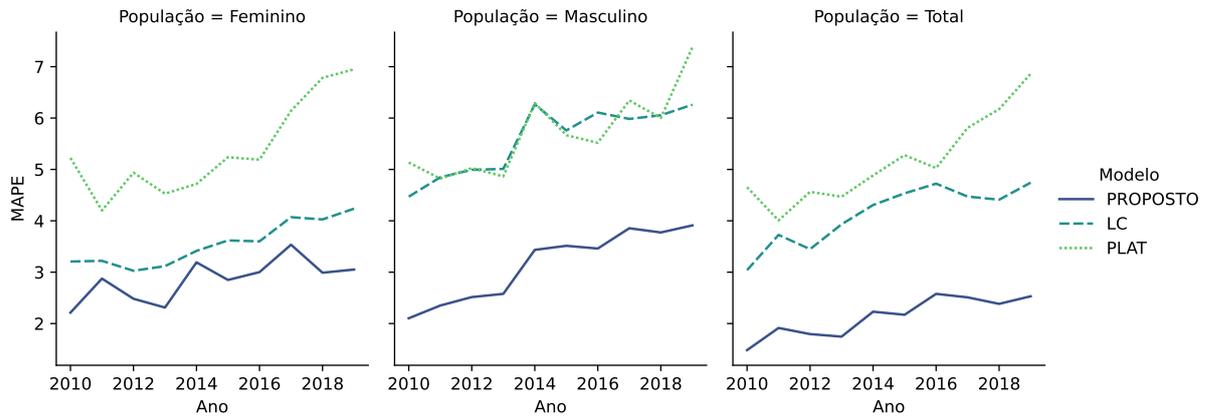


Fonte: Dados da pesquisa (2024).

segregados por conjuntos de dados<sup>9</sup>. Ou seja, os resultados estão dispostos de tal maneira que seja possível avaliar se a modelagem residual adicionada ao modelo ARIMA, independente

<sup>9</sup> O Apêndice D mostra a figura com os dados do MASE.

Figura 40 – MAPE (%) do Sistema proposto, LC e PLAT para Portugal entre os anos 2010-2019

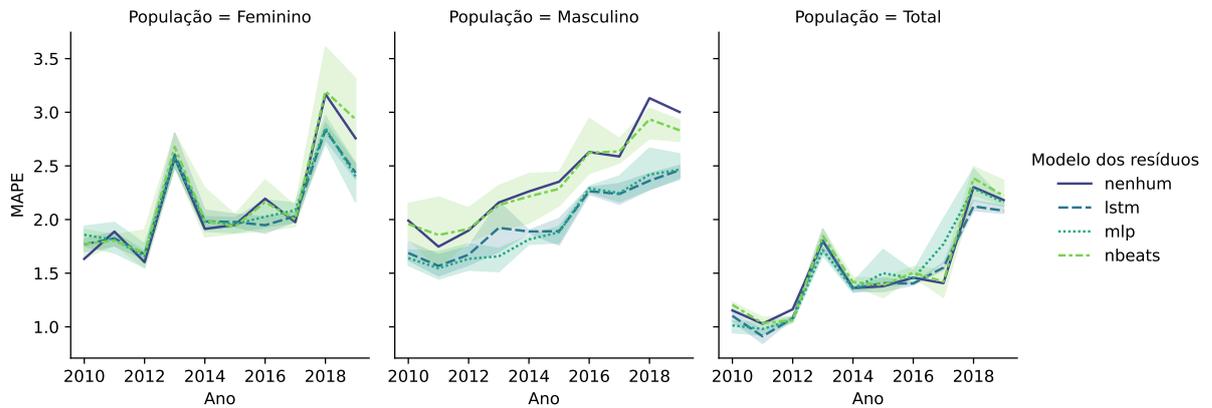


Fonte: Dados da pesquisa (2024).

de abordagem de múltiplos passos à frente, mostra redução no valor do MAPE. Também é possível observar se haveria diferenças no que tange a variabilidade dos resultados de acurácia preditiva quanto aos modelos não-lineares a partir da variação das abordagens *multi-step*. Pela figura supracitada, observou-se que a modelagem residual apresenta ganhos no desempenho preditivo para alguns conjuntos de dados, notadamente Austrália masculino, França masculino, Japão feminino e total, e Portugal masculino e total. Em se tratando de variação dos resultados por abordagem *multi-step*, notou-se que as modelagens residuais realizadas pelos modelos N-BEATS, principalmente, e em segundo lugar pelo LSTM foram menos robustas em relação à modelagem feita pelo modelo MLP. Em contrapartida, verificou-se por meio da regressão que os modelos híbridos univariados recursivos foram os modelos com melhores resultados. Desse modo, é possível deduzir que a escolha da abordagem de previsão múltiplos passos impacta mais os resultados do MAPE para o modelo N-BEATS e em menor intensidade pelo modelo LSTM. Esse resultado é relevante, pois demonstra que a escolha pela abordagem recursiva é assertiva quanto garante acurácia de previsão superior às alternativas. Além disso, essas evidências podem auxiliar a tomada de decisão ao passo que sugerem redução da complexidade da escolha de abordagem de múltiplos passos tal como pode reduzir o custo computacional, uma vez que a abordagem recursiva é aquela com menor tempo de processamento computacional.

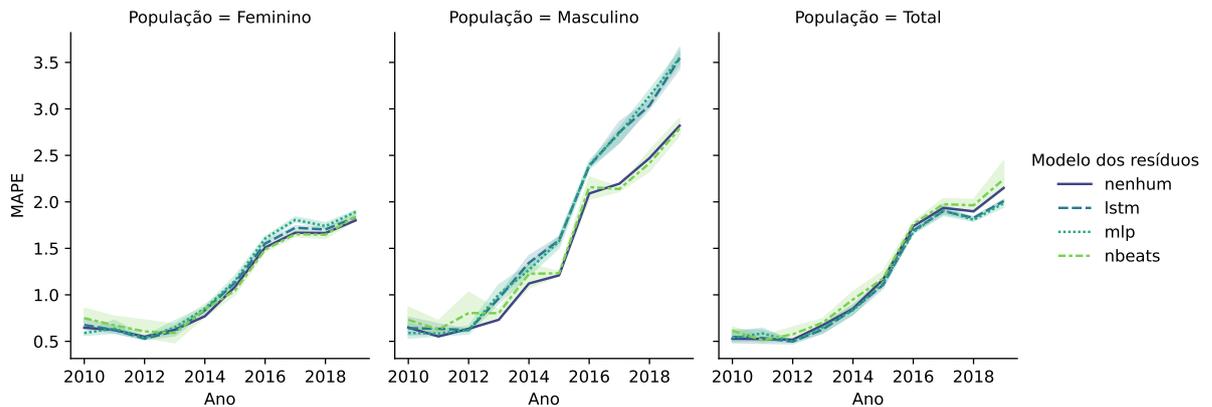
Ademais, foi realizada uma avaliação do ranque do MAPE e do MASE para os modelos selecionados e todos os conjuntos de dados. Figura 46 e Figura 47 mostram os *boxplots* com os dados do MAPE e MASE, respectivamente. Dessas imagens, notou-se que o sistema proposto exibiu o menor ranque, ou seja, o melhor resultado, tendo em vista que a mediana tal como valor máximo do ranque esteve abaixo dessas estatísticas em comparação aos demais modelos.

Figura 41 – MAPE (%) dos modelos univariados e modelo dos resíduos para a Austrália entre os anos 2010-2019



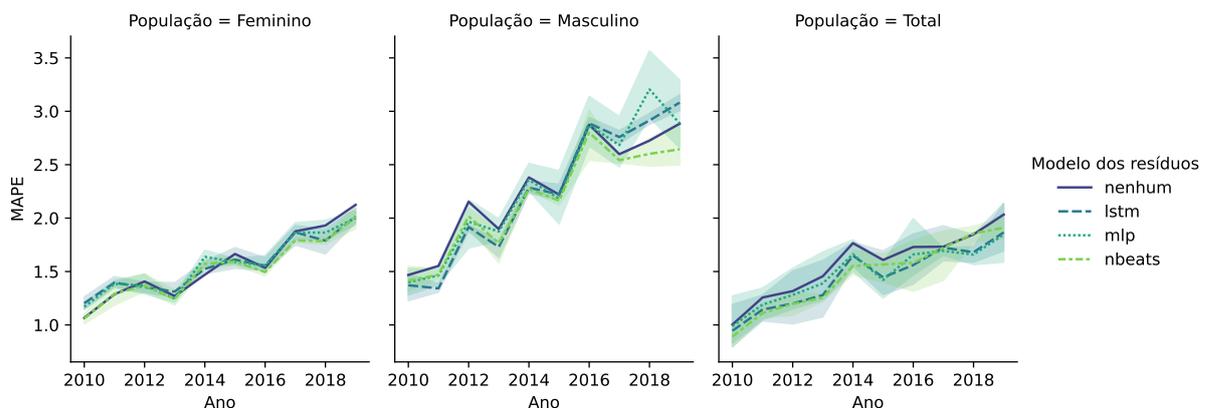
Fonte: Dados da pesquisa (2024).

Figura 42 – MAPE (%) dos modelos univariados e modelo dos resíduos para os EUA entre os anos 2010-2019



Fonte: Dados da pesquisa (2024).

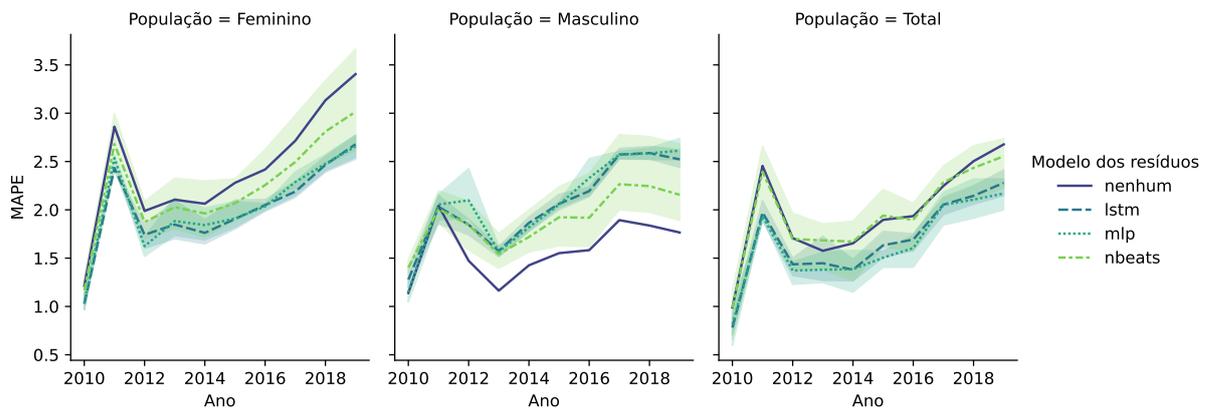
Figura 43 – MAPE (%) dos modelos univariados e modelo dos resíduos para a França entre os anos 2010-2019



Fonte: Dados da pesquisa (2024).

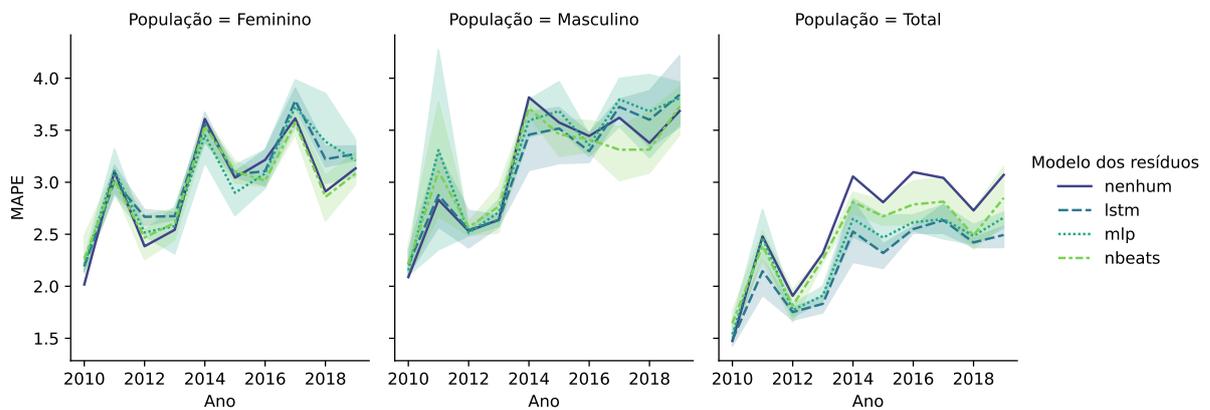
Ainda, considerando o MAPE, os modelos híbridos univariados ARIMA-LSTM, ARIMA-MLP, ARIMA-NBEATS e HyS-MF apresentaram os melhores desempenhos comparados à literatura, evidenciando ranque médio menor ou igual 5, dentre um conjunto de 25 modelos. Em relação

Figura 44 – MAPE (%) dos modelos univariados e modelo dos resíduos para o Japão entre os anos 2010-2019



Fonte: Dados da pesquisa (2024).

Figura 45 – MAPE (%) dos modelos univariados e modelo dos resíduos para Portugal entre os anos 2010-2019

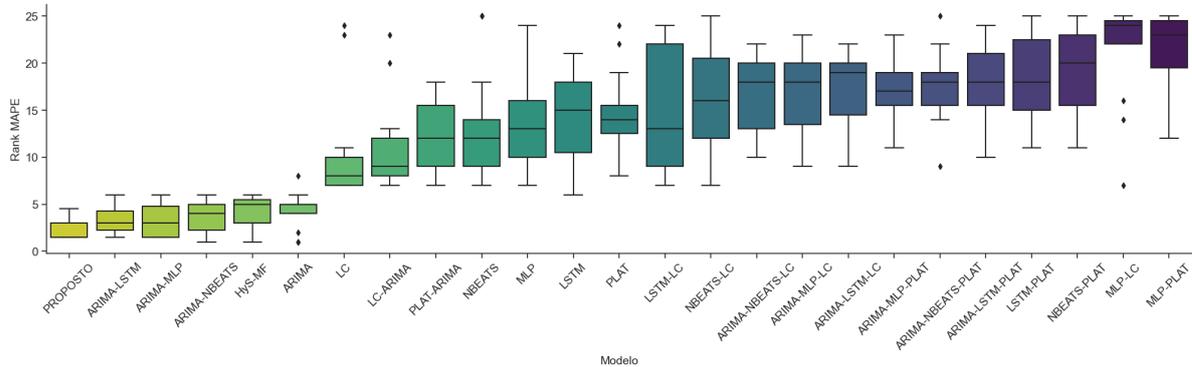


Fonte: Dados da pesquisa (2024).

ao MASE, dos quatro modelos híbridos destacados acima, apenas dois deles, ou seja, ARIMA-MLP e ARIMA-LSTM, obtiveram um ranque médio inferior ao ARIMA. Porém, quando se avalia a mediana do ranque, notou-se que o ARIMA-NBEATS, levou a um valor inferior ao ARIMA. No que diz respeito aos modelos multivariados, reparou-se que as versões clássicas tradicionais (*i.e.*, sem modificações), possuíram menores ranques quando equiparadas às suas versões hibridizadas ou ajustadas para os modelos de AM. Isto posto, esse resultado mostrou que optar pela modelagem híbrida univariada, preferencialmente com abordagem de múltiplos passos à frente recursiva no modelo não-linear, tendeu a produzir os melhores desempenhos de acurácia de previsão das taxas de mortalidade para os conjuntos de dados estudados. Também, vale destacar que a utilização da etapa de seleção de modelo não-linear (*i.e.*, incorporada no sistema proposto), promoveu um desempenho de previsão superior.

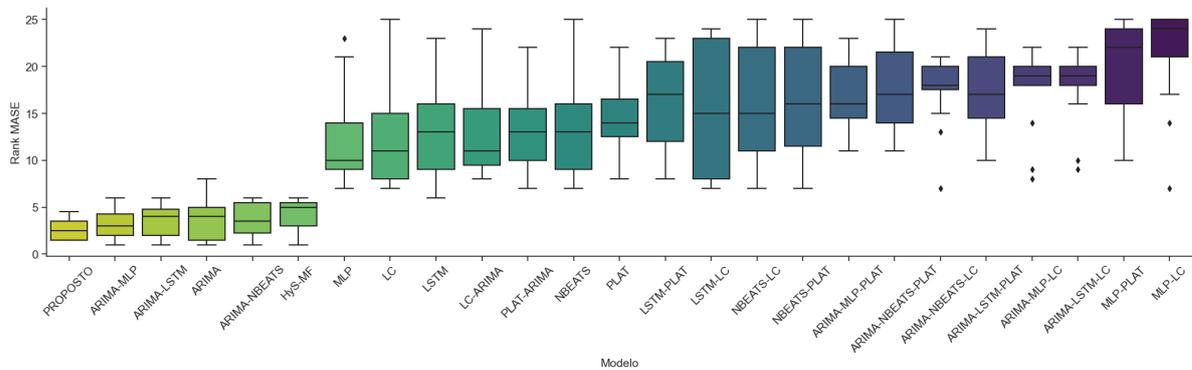
A Tabela 2 apresenta os modelos de previsão selecionados no conjunto de validação do sistema proposto para modelar os resíduos do ARIMA para os conjuntos de dados, segmentados

Figura 46 – Boxplot do Rank ordenado pela média do MAPE do Sistema proposto, sistemas híbridos recursivos, modelos monolíticos e multivariados recursivos para todos os conjuntos de dados.



Fonte: Dados da pesquisa (2024).

Figura 47 – Boxplot do Rank ordenado pela média do MASE do Sistema proposto, sistemas híbridos recursivos, modelos monolíticos e multivariados recursivos para todos os conjuntos de dados.



Fonte: Dados da pesquisa (2024).

por país e grupo demográfico (feminino, masculino e total). Dos quinze conjuntos de dados, N-BEATS foi selecionado três vezes para os conjuntos do gênero feminino da Austrália, EUA e França. MLP foi escolhido sete vezes, majoritariamente para os grupos masculinos e total. LSTM foi elegido cinco vezes, notadamente para França (masculino e total), Japão (feminino e masculino) e para a população total de Portugal. Esses padrões indicam uma adaptação dos modelos às especificidades de cada país e segmento demográfico, visando otimizar a acurácia das previsões conforme as características particulares dos dados analisados.

A Tabela 3 exibe o MAPE (%), considerando todas as idades (i.e., 0, 1, ..., 100), ou seja, as curvas de mortalidade para os 10 anos de teste, do sistema proposto, HyS-MF e modelos de abordagem recursiva, modelos selecionados para todos os conjuntos de dados avaliados. O modelo ARIMA apresentou o melhor desempenho de previsão para dois conjuntos de dados, EUA feminino e Japão masculino. O sistema proposto mostrou o melhor desempenho de previsão em 12 de 15 conjuntos de dados e também na média. Os sistemas híbridos com abordagem

Tabela 2 – Seleção de modelo no sistema proposto para modelar os resíduos do ARIMA por conjunto de dados.

Conjunto de dados		Modelo
Austrália	Feminino	N-BEATS
	Masculino	MLP
	Total	MLP
EUA	Feminino	N-BEATS
	Masculino	MLP
	Total	MLP
França	Feminino	N-BEATS
	Masculino	LSTM
	Total	LSTM
Japão	Feminino	LSTM
	Masculino	LSTM
	Total	MLP
Portugal	Feminino	MLP
	Masculino	MLP
	Total	LSTM

**Fonte:** Dados da pesquisa (2024).

recursiva na modelagem não-linear foram os modelos com os melhores desempenhos, como segue: ARIMA-LSTM, ARIMA-MLP e ARIMA-NBEATS obtiveram o menor MAPE em 3, 5 e 4 conjuntos de dados cada. O primeiro foi o melhor em França total, Japão feminino, e Portugal total; o segundo demonstrou o menor MAPE para Austrália masculino e total, EUA total, Japão total e Portugal feminino; ao passo que o terceiro modelo foi superior em Austrália feminino, EUA masculino, França feminino e Portugal masculino. Além disso, o HyS-MF obteve o melhor desempenho em um conjunto de dados, isto é, para a França masculino.

A Tabela 4 mostra o MASE considerando toda a curva de mortalidade para o conjunto de teste o qual compreende 10 horizontes de previsão para o sistema proposto, HyS-MF e modelos de abordagem recursiva. Esta tabela demonstra que o sistema proposto obteve os melhores resultados em média e em 7 de 15 conjunto de dados. Por sua vez, o ARIMA foi o modelo com o menor MASE em 4 conjunto de dados, ao passo que o HyS-MF foi o melhor em 2 dos 15.

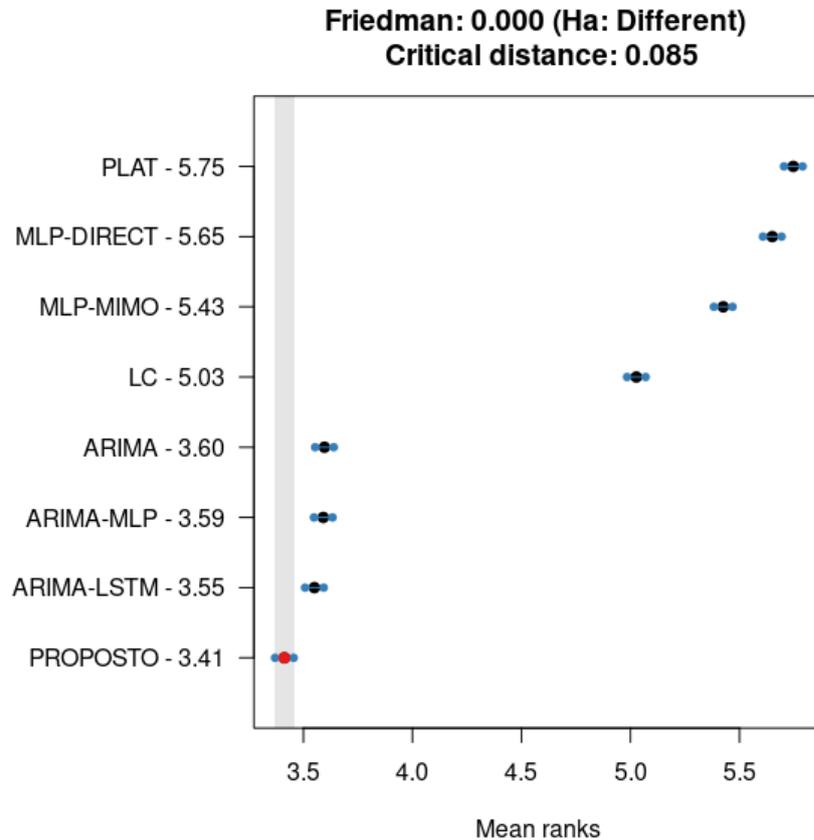
A Tabela 5 mostra a Diferença Percentual (DP) entre o sistema proposto (i.e., modelo com melhor desempenho geral) e os demais modelos avaliados nesta seção. Dessa tabela é possível verificar que no que diz respeito ao modelo ARIMA, o sistema proposto foi superior numericamente em dez dos quinze conjuntos de dados. Desses, a diferença percentual média foi de

5,8839%. A maior diferença percentual examinada entre esses modelos foi de 24,9641% para o conjunto de dados Japão total. Em comparação aos modelos tradicionais da literatura LC e PLAT, o sistema proposto foi superior em todos os conjuntos de dados, com DP de pelo menos 19,8101% e 44,4471%, respectivamente. As maiores diferenças percentuais observadas no experimento quanto aos modelos LC e Plat foram de 72,0463% e 69,1485%, respectivamente. Quando comparado aos modelos que compõem o *pool* do sistema proposto, ARIMA-LSTM, ARIMA-MLP e ARIMA-NBEATS, e desconsiderando os conjuntos de dados em que o modelo do pool foi selecionado, o sistema proposto auferiu o melhor resultado em 9 de 10, 8 de 8 e 11 de 12 casos, respectivamente. Sendo assim, nota-se que o sistema híbrido demonstra um ganho percentual relevante de acurácia preditiva pela métrica MAPE, que pode traduzir em considerável economia de recursos e redução de riscos para as entidades seguradoras e fundos de pensão, no caso de eles empregarem a abordagem híbrida proposta nesta pesquisa em detrimento dos modelos clássicos estatístico-demográficos. Os valores positivos apresentados na Tabela 5 representam uma melhoria do sistema proposto em relação aos outros métodos avaliados no experimento desta pesquisa, enquanto os valores negativos representam a situação contrária, ou seja, uma piora do sistema híbrido supracitado em comparação aos modelos investigados.

A Tabela 6 exibe a DP do sistema proposto em relação aos modelos de abordagem recursiva e ao HyS-MF considerando o MASE. Ela mostra que o sistema proposto demonstra um ganho de desempenho preditivo na maioria dos casos. Em relação ao ARIMA, o sistema proposto foi superior em 9 dos 15 conjuntos de dados. No que diz respeito ao HyS-MF, o sistema proposto apresentou ganho de desempenho em 11 de 15 casos. Quando comparado aos modelos que compõem o *pool* do sistema proposto, ARIMA-LSTM, ARIMA-MLP e ARIMA-NBEATS, e desconsiderando os conjuntos de dados em que o modelo do pool foi selecionado, o sistema proposto auferiu o melhor resultado em 8 de 10, 6 de 8 e 8 de 12 casos, respectivamente. Esse resultado mostrou que realizar a seleção de modelo considerando todas as idades de uma população melhorou o desempenho de previsão do sistema em comparação aos modelos híbridos individualmente.

A partir dessas análises e observadas algumas diferenças entre os modelos propostos e a literatura, foi realizado o teste estatístico de Nemenyi MCB. Este teste buscou avaliar o ranque dos resíduos em absoluto do melhor modelo contra os demais. A Figura 48 exibe os resultados do teste de Nemenyi para o sistema proposto e os dois melhores modelos em cada classe (*i.e.*, híbridos, lineares univariados, multivariados monolíticos e híbridos), o qual evidencia que o

Figura 48 – Teste de Nemenyi MCB que avalia a diferença entre os erros em absoluto de previsão do melhor modelo contra os demais.



Fonte: Dados da pesquisa (2024).

modelo com o menor ranque dos erros é o sistema proposto. Ainda, ela mostra que o sistema proposto é superior estatisticamente à literatura, a um nível de significância de 5%. Em outros termos, o sistema proposta alcança o menor erro de previsão em comparação com todos os modelos investigados nesse teste. Como destaque, o sistema proposto foi superior do ponto de vista do desempenho de previsão em comparação aos modelos ARIMA-LSTM, ARIMA-MLP, ARIMA-NBEATS, ARIMA, HyS-MF, monolíticos de AM e multivariados.

Perante o exposto, notou-se que o sistema proposto se consolidou como uma alternativa proveitosa e relevante perante a literatura para a previsão de taxas de mortalidade para vários passos à frente. Em princípio, é possível notar que o sistema proposto nesta tese resolveu, de forma mais acurada, o problema de prever taxas de mortalidade para múltiplos passos à frente. Outrossim, pensa-se ser possível aplicar a propositura em outras aplicações de previsão que envolvam cenários semelhantes aos destacados aqui, i.e., séries temporais curtas e vários passos à frente.

Tabela 3 – MAPE (%) para todos os conjuntos de dados, para o sistema proposto, modelos de abordagem recursiva e HyS-MF. Melhores valores em negrito.

Modelo	Austrália			EUA			França			Japão			Portugal			Média
	Feminino	Masculino	Total													
ARIMA	2,166	2,376	1,523	<b>1,094</b>	1,448	1,198	1,563	2,275	1,575	2,418	<b>1,587</b>	1,964	2,957	3,161	2,598	1,994
ARIMA-LSTM	2,072	1,988	1,483	1,133	1,738	1,172	1,607	<b>2,220</b>	<b>1,346</b>	1,920	2,026	1,575	3,125	<b>3,038</b>	<b>2,134</b>	1,905
ARIMA-MLP	2,185	<b>1,982</b>	<b>1,470</b>	1,174	1,733	1,398	1,525	2,418	1,398	1,949	2,034	<b>1,474</b>	<b>2,850</b>	3,149	2,299	1,920
ARIMA-NBEATS	<b>2,042</b>	2,197	1,607	1,122	<b>1,434</b>	1,250	<b>1,502</b>	2,222	1,508	1,998	2,236	1,629	2,851	2,997	2,597	1,946
HyS-MF	2,420	2,428	1,519	1,120	1,522	1,230	1,565	2,115	1,368	2,116	1,686	2,122	3,003	3,212	2,447	1,992
LC	3,823	4,640	3,594	1,828	2,661	1,656	2,247	2,868	2,258	6,870	3,419	4,367	3,554	5,575	4,134	3,566
LC-ARIMA	3,767	4,996	3,981	1,828	3,006	1,656	2,253	3,019	2,259	6,739	3,349	4,269	5,007	6,718	5,081	3,862
LC-ARIMA-LSTM	4,717	8,906	4,005	4,623	7,659	1,653	6,210	8,414	2,259	5,473	8,999	4,270	7,025	9,353	5,215	5,919
LC-ARIMA-MLP	4,703	8,902	4,025	4,645	7,692	1,642	6,243	8,373	2,260	5,471	8,923	4,148	7,021	9,431	5,118	5,906
LC-ARIMA-NBEATS	4,681	8,895	3,980	4,631	7,635	1,656	6,262	8,581	2,267	5,439	8,910	4,172	7,012	9,385	5,132	5,909
LC-LSTM	6,114	3,524	4,023	5,788	3,669	1,621	7,757	4,010	5,255	5,126	5,813	3,010	16,158	4,881	7,145	5,593
LC-MLP	5,592	15,033	9,174	3,569	18,037	1,575	8,471	9,150	5,304	6,413	9,654	4,589	7,093	10,672	8,249	8,172
LC-NBEATS	5,467	4,083	3,471	7,313	17,098	1,839	8,429	4,072	3,704	5,289	5,204	2,857	9,009	6,913	6,727	6,098
LSTM	4,647	6,323	5,496	1,832	3,464	2,401	3,832	6,114	4,627	2,139	3,863	2,541	7,532	7,776	7,202	4,653
MLP	6,417	5,595	5,725	1,575	4,061	2,208	3,635	4,038	2,841	4,665	4,696	3,308	5,262	9,321	5,879	4,615
NBEATS	31,597	5,988	3,918	1,618	2,262	14,347	3,409	5,284	3,723	2,281	3,960	3,873	5,130	6,517	6,234	6,676
PLAT	3,676	5,385	4,286	2,999	4,470	3,541	4,144	5,275	4,364	5,256	4,732	4,435	5,391	5,704	5,173	4,589
PLAT-ARIMA	3,372	5,501	4,123	2,885	4,116	3,180	3,764	5,608	4,585	4,638	4,386	4,112	5,115	5,459	4,519	4,358
PLAT-ARIMA-LSTM	4,755	9,010	4,162	4,157	9,140	3,186	5,848	12,098	4,644	5,146	9,001	4,168	7,209	8,249	4,570	6,356
PLAT-ARIMA-MLP	4,834	9,037	4,116	4,161	9,394	3,207	5,806	12,053	4,510	5,143	8,902	4,093	7,190	8,185	4,574	6,347
PLAT-ARIMA-NBEATS	4,780	8,699	4,095	4,157	9,330	3,320	5,940	12,201	4,593	5,152	8,862	4,053	7,210	8,829	4,567	6,386
PLAT-LSTM	5,536	7,836	4,619	4,603	5,533	3,640	8,201	5,044	4,119	7,145	5,032	3,673	11,849	7,278	7,422	6,102
PLAT-MLP	5,060	11,846	8,241	5,249	4,020	4,901	5,889	7,660	5,693	5,833	9,972	4,362	22,729	13,668	10,984	8,407
PLAT-NBEATS	4,966	698,159	6,438	5,579	5,097	3,265	5,833	8,059	20,006	5,836	4,317	3,685	9,965	6,196	8,045	53,030
PROPOSTO	<b>2,042</b>	<b>1,982</b>	<b>1,470</b>	1,122	1,733	<b>1,157</b>	<b>1,502</b>	<b>2,220</b>	<b>1,346</b>	<b>1,920</b>	2,026	<b>1,474</b>	<b>2,850</b>	3,149	<b>2,134</b>	<b>1,875</b>

Fonte: Dados da pesquisa (2024).

Tabela 4 – MASE para todos os conjuntos de dados, para o sistema proposto, modelos de abordagem recursiva e HyS-MF. Melhores valores em negrito.

Modelo	Austrália		EUA		França		Japão		Portugal		Média		
	Feminino	Masculino	Total	Feminino	Masculino	Total	Feminino	Masculino	Total	Feminino		Masculino	Total
ARIMA	1,059	1,112	1,033	<b>1,987</b>	<b>2,280</b>	2,709	<b>1,096</b>	1,495	1,489	1,967	0,904	1,159	1,225
ARIMA-LSTM	1,068	1,033	<b>0,994</b>	2,011	2,577	2,593	1,169	1,590	<b>1,393</b>	1,799	0,992	<b>1,074</b>	<b>1,114</b>
ARIMA-MLP	1,145	<b>1,015</b>	1,008	2,126	2,572	<b>2,553</b>	1,113	1,598	1,419	<b>1,742</b>	0,953	1,139	1,180
ARIMA-NBEATS	<b>1,034</b>	1,080	1,068	1,997	2,349	2,757	1,103	1,481	1,479	1,766	<b>0,891</b>	1,136	1,280
HyS-MF	1,162	1,131	1,043	1,991	2,345	2,736	1,125	<b>1,461</b>	1,438	2,022	0,969	1,180	1,210
LC	2,125	3,031	3,032	3,396	4,071	3,703	1,697	2,482	2,730	5,275	1,167	2,406	2,424
LC-ARIMA	2,083	3,508	3,438	3,396	4,838	3,703	1,700	2,523	2,725	5,172	1,647	2,730	2,837
LC-ARIMA-LSTM	2,343	4,915	3,460	7,750	11,616	3,695	4,867	4,630	2,724	5,173	2,455	2,891	2,890
LC-ARIMA-MLP	2,341	4,913	3,478	7,785	11,672	3,669	4,887	4,605	2,718	5,049	2,453	2,913	2,852
LC-ARIMA-NBEATS	2,339	4,908	3,437	7,762	11,575	3,704	4,898	4,732	2,703	5,067	2,449	2,900	2,857
LC-LSTM	2,749	1,614	3,476	9,683	5,651	3,285	5,795	2,432	5,556	3,235	4,566	1,974	3,499
LC-MLP	3,070	8,416	7,246	6,377	28,587	3,198	6,238	5,092	4,840	5,532	2,476	3,279	3,967
LC-NBEATS	2,524	2,041	2,382	12,290	26,806	3,729	6,212	2,531	4,139	3,410	2,863	2,263	3,329
LSTM	1,829	3,166	3,221	2,925	5,137	4,843	2,700	4,270	4,973	3,031	2,021	2,825	3,535
MLP	2,504	2,710	2,741	2,761	5,610	5,041	2,239	3,400	2,987	3,247	1,473	3,206	3,154
NBEATS	10,414	3,278	2,694	2,288	3,616	43,238	2,055	3,957	3,763	4,457	1,530	2,400	3,009
PLAT	1,653	2,728	3,005	4,858	6,370	7,085	2,998	3,952	4,791	4,675	1,814	2,232	2,950
PLAT-ARIMA	1,632	2,674	2,960	4,650	5,792	6,467	2,803	4,083	5,037	4,559	1,739	2,103	2,615
PLAT-ARIMA-LSTM	2,080	4,199	2,976	6,456	12,962	6,462	3,786	8,036	5,049	4,558	2,147	3,040	2,658
PLAT-ARIMA-MLP	2,114	4,215	2,950	6,453	13,199	6,482	3,762	8,022	4,955	4,547	2,149	3,167	2,678
PLAT-ARIMA-NBEATS	2,089	4,124	2,946	6,444	13,230	6,451	3,877	8,145	5,010	4,547	2,149	3,167	2,678
PLAT-LSTM	2,182	3,122	2,654	7,262	7,955	7,425	5,406	3,119	3,852	3,517	3,251	2,184	3,135
PLAT-MLP	1,926	5,083	5,018	8,566	6,807	7,904	3,486	4,479	4,973	5,199	4,721	3,563	4,685
PLAT-NBEATS	1,915	261,377	3,926	8,919	7,680	6,704	3,478	4,751	16,352	3,307	2,775	1,900	2,556
PROPOSTO	<b>1,034</b>	<b>1,015</b>	1,008	1,997	2,572	<b>2,553</b>	1,103	1,590	<b>1,393</b>	<b>1,742</b>	0,953	1,139	<b>1,114</b>

Fonte: Dados da pesquisa (2024).

Tabela 5 – Diferença Percentual (DP) (%) do MAPE entre o modelo proposto e demais modelos para todos os conjuntos de dados. Destaques em negrito.

Modelo	Austrália			EUA			França			Japão			Portugal		
	Feminino	Masculino	Total												
ARIMA	<b>5,730</b>	<b>16,569</b>	<b>3,485</b>	-2,562	-19,697	<b>3,440</b>	<b>3,924</b>	<b>2,453</b>	<b>14,518</b>	<b>20,579</b>	-27,605	<b>24,964</b>	<b>3,641</b>	<b>0,364</b>	<b>17,861</b>
ARIMA-LSTM	<b>1,428</b>	<b>0,307</b>	<b>0,859</b>	<b>1,010</b>	<b>0,317</b>	<b>1,240</b>	<b>6,574</b>	0,000	0,000	0,000	0,000	<b>6,403</b>	<b>8,796</b>	-3,652	0,000
ARIMA-MLP	<b>6,529</b>	0,000	0,000	<b>4,434</b>	0,000	0,000	<b>1,528</b>	<b>8,196</b>	<b>3,688</b>	<b>1,470</b>	<b>0,406</b>	0,000	0,000	0,000	<b>7,184</b>
ARIMA-NBEATS	0,000	<b>9,779</b>	<b>8,515</b>	0,000	-20,783	<b>7,467</b>	0,000	<b>0,121</b>	<b>10,713</b>	<b>3,887</b>	<b>9,393</b>	<b>9,506</b>	<b>0,036</b>	-5,074	<b>17,813</b>
HyS-MF	<b>15,612</b>	<b>18,370</b>	<b>3,189</b>	-0,133	-13,859	<b>5,903</b>	<b>4,040</b>	-4,963	<b>1,607</b>	<b>9,233</b>	-20,174	<b>30,537</b>	<b>5,106</b>	<b>1,951</b>	<b>12,780</b>
LC	<b>46,580</b>	<b>57,280</b>	<b>59,088</b>	<b>38,645</b>	<b>34,893</b>	<b>30,131</b>	<b>33,153</b>	<b>22,613</b>	<b>40,373</b>	<b>72,046</b>	<b>40,760</b>	<b>66,252</b>	<b>19,810</b>	<b>43,513</b>	<b>48,372</b>
LC-ARIMA	45,785	60,324	63,067	38,645	42,355	30,131	33,346	26,489	40,410	71,502	39,518	65,477	43,080	53,124	57,998
LC-ARIMA-LSTM	56,711	77,744	63,283	75,733	77,379	29,992	75,819	73,621	40,412	64,912	77,492	65,489	59,434	66,329	59,078
LC-ARIMA-MLP	56,577	77,733	63,466	75,851	77,474	29,529	75,945	73,491	40,418	64,898	77,300	64,467	59,412	66,608	58,306
LC-ARIMA-NBEATS	56,377	77,715	63,053	75,778	77,307	30,142	76,018	74,134	40,606	64,691	77,265	64,677	59,358	66,443	58,418
LC-LSTM	66,598	43,743	63,449	80,619	52,783	28,598	80,640	44,650	74,379	62,536	65,151	51,043	82,363	35,484	70,133
LC-MLP	63,481	86,814	83,972	68,569	90,394	26,545	82,273	75,743	74,620	70,056	79,019	67,883	59,822	70,490	74,131
LC-NBEATS	62,644	51,446	57,641	84,661	89,866	37,065	82,184	45,495	63,651	63,693	61,072	48,407	68,367	54,444	68,278
LSTM	56,054	68,651	73,248	38,777	49,989	51,809	60,809	63,697	70,907	10,217	47,561	41,994	62,167	59,503	70,368
MLP	68,178	64,570	74,317	28,784	57,339	47,601	58,691	45,036	52,620	58,837	56,862	55,444	45,848	66,213	63,702
NBEATS	93,537	66,895	62,471	30,669	23,413	91,935	55,947	57,991	63,842	15,817	48,849	61,949	44,454	51,676	65,770
PLAT	<b>44,447</b>	<b>63,187</b>	<b>65,697</b>	<b>62,592</b>	<b>61,241</b>	<b>67,324</b>	<b>63,761</b>	<b>57,928</b>	<b>69,149</b>	<b>63,463</b>	<b>57,197</b>	<b>66,768</b>	<b>47,134</b>	<b>44,793</b>	<b>58,744</b>
PLAT-ARIMA	39,432	63,965	64,340	61,121	57,900	63,614	60,104	60,421	70,634	58,592	53,817	64,158	44,286	42,307	52,775
PLAT-ARIMA-LSTM	57,057	78,000	64,676	73,011	81,044	63,679	74,319	81,653	71,011	62,679	77,495	64,638	60,468	61,824	53,303
PLAT-ARIMA-MLP	57,759	78,066	64,281	73,041	81,556	63,914	74,135	81,585	70,148	62,657	77,246	63,993	60,363	61,523	53,339
PLAT-ARIMA-NBEATS	57,276	77,212	64,091	73,012	81,430	65,150	74,718	81,808	70,688	62,724	77,142	63,633	60,475	64,332	53,271
PLAT-LSTM	63,112	74,702	68,170	75,630	68,688	68,208	81,689	55,994	67,317	73,123	59,743	59,875	75,950	56,729	71,248
PLAT-MLP	59,646	83,267	82,159	78,629	56,905	76,392	74,497	71,024	76,351	67,075	79,688	66,212	87,462	76,959	80,572
PLAT-NBEATS	58,879	99,716	77,162	79,891	66,010	64,561	74,255	72,460	93,270	67,094	53,079	60,007	71,402	49,171	73,474

Fonte: Dados da pesquisa (2024).

Tabela 6 – Diferença Percentual (DP) (%) do MASE entre o modelo proposto e demais modelos para todos os conjuntos de dados. Destaques em negrito.

Modelo	Austrália			EUA			França			Japão			Portugal		
	Feminino	Masculino	Total												
ARIMA	<b>2,346</b>	<b>8,696</b>	<b>2,412</b>	-0,497	-12,832	<b>5,782</b>	-0,664	-6,314	<b>6,469</b>	<b>18,384</b>	-22	<b>11,431</b>	-5,503	<b>1,742</b>	<b>9,081</b>
ARIMA-LSTM	<b>3,187</b>	<b>1,657</b>	-1,35	<b>0,696</b>	<b>0,173</b>	<b>1,549</b>	<b>5,611</b>	0	0	0	0	<b>3,148</b>	<b>3,858</b>	-6,056	0
ARIMA-MLP	<b>9,688</b>	0	0	<b>6,066</b>	0	0	<b>0,937</b>	<b>0,526</b>	<b>1,806</b>	-0,014	-0,701	0	0	0	<b>5,609</b>
ARIMA-NBEATS	0	<b>5,948</b>	5,625	0	-9,514	<b>7,402</b>	0	-7,319	<b>5,837</b>	<b>2,042</b>	<b>4,842</b>	-6,993	-0,252	<b>12,966</b>	
HyS-MF	<b>11,014</b>	<b>10,211</b>	3,429	-0,292	-9,685	<b>6,695</b>	<b>1,93</b>	-8,786	<b>3,131</b>	<b>4,512</b>	-21,963	<b>13,831</b>	<b>1,587</b>	<b>3,492</b>	<b>7,99</b>
LC	<b>51,316</b>	<b>66,499</b>	<b>66,77</b>	<b>41,202</b>	<b>36,812</b>	<b>31,064</b>	<b>35,017</b>	<b>35,958</b>	<b>48,967</b>	<b>74,44</b>	<b>43,74</b>	<b>66,973</b>	<b>18,327</b>	<b>52,673</b>	<b>54,053</b>
LC-ARIMA	50,335	71,057	70,688	41,202	46,828	31,064	35,11	36,987	48,885	74,022	42,651	66,32	42,121	58,3	60,743
LC-ARIMA-LSTM	55,855	79,339	70,876	74,236	77,855	30,909	77,336	65,666	48,85	63,317	71,954	66,328	61,176	60,61	61,468
LC-ARIMA-MLP	55,815	79,329	71,033	74,351	77,961	30,422	77,428	65,478	48,742	63,302	71,69	65,496	61,137	60,917	60,95
LC-ARIMA-NBEATS	55,774	79,312	70,679	74,274	77,775	31,073	77,48	66,407	48,463	63,024	71,64	65,618	61,073	60,735	61,024
LC-LSTM	62,379	37,079	71,015	79,378	54,478	22,283	80,966	34,63	74,926	58,628	52,955	46,149	79,12	42,32	68,175
LC-MLP	66,309	87,934	86,095	68,689	91,001	20,178	82,317	68,782	71,218	69,512	74,2	68,509	61,5	65,276	71,924
LC-NBEATS	59,013	50,25	57,697	83,753	90,404	31,537	82,242	37,189	66,337	58,777	47,108	48,921	66,701	49,682	66,545
LSTM	43,442	67,923	68,722	31,733	49,928	47,286	59,145	62,772	71,985	6,581	41,243	42,534	52,832	59,695	68,493
MLP	58,701	62,533	63,236	27,685	54,143	49,355	50,725	53,239	53,367	41,59	34,514	46,351	35,264	64,482	64,689
NBEATS	90,068	69,026	62,598	12,738	28,851	94,096	46,315	59,821	62,983	14,303	55,572	60,918	37,674	52,552	62,994
PLAT	<b>37,41</b>	<b>62,781</b>	<b>66,473</b>	<b>58,897</b>	<b>59,619</b>	<b>63,97</b>	<b>63,209</b>	<b>59,773</b>	<b>70,923</b>	<b>59,269</b>	<b>53,741</b>	<b>62,741</b>	<b>47,457</b>	<b>48,987</b>	<b>62,255</b>
PLAT-ARIMA	36,635	62,023	65,964	57,058	55,584	60,526	60,642	61,063	72,342	59,741	50,541	61,79	45,176	45,866	57,419
PLAT-ARIMA-LSTM	50,264	75,817	66,143	69,068	80,155	60,497	70,862	80,216	72,406	55,977	73,352	62,115	55,602	63,314	58,017
PLAT-ARIMA-MLP	51,061	75,907	65,848	69,058	80,51	60,617	70,677	80,183	71,883	55,418	73,209	61,778	55,596	62,549	58,108
PLAT-ARIMA-NBEATS	50,484	75,378	65,794	69,013	80,557	60,427	71,548	80,481	72,193	56,09	73,172	61,69	55,633	64,047	58,417
PLAT-LSTM	52,601	67,473	62,031	72,504	67,662	65,621	79,596	49,033	63,834	68,821	47,871	50,468	70,672	47,877	64,472
PLAT-MLP	46,292	80,023	79,921	76,69	62,207	67,703	68,354	64,508	71,988	48,806	73,036	66,493	79,807	68,043	76,227
PLAT-NBEATS	45,986	99,611	74,337	77,612	66,506	61,921	68,283	66,541	91,48	56,712	48,643	47,318	65,644	40,06	56,437

Fonte: Dados da pesquisa (2024).

#### 5.4 TEMPO DE EXECUÇÃO DOS ALGORITMOS

A eficiência computacional é um aspecto pertinente na aplicação prática de algoritmos de AM, especialmente em cenários que envolvem múltiplas séries temporais. Neste sentido, nesta seção são apresentados os resultados do tempo de execução dos algoritmos estudados. O tempo computacional foi mensurado em uma máquina com processador I7 da 9ª geração 2.60 GHz com 16 GB de memória RAM e GPU NVIDIA GeForce GTX 1660 Ti.

A Tabela 7 mostra o tempo médio de execução em segundos do sistema proposto e demais modelos avaliados nesta tese. Desse modo, dos modelos avaliados, o ARIMA apresentou tempo médio de 2,29 segundos para cada conjunto de dados. Por sua vez, o sistema proposto demonstrou um tempo médio de execução de 9.581,29 segundos. ARIMA-LSTM, que apresentou o segundo melhor desempenho em termos de acurácia de previsão, teve um tempo médio de 4.392,49 segundos. Embora o tempo médio de execução do sistema proposto tenha sido superior ao dos modelos individuais que o compõem, os ganhos obtidos em termos de acurácia de previsão foram significativamente maiores. Essa melhoria na previsões justifica o aumento no custo computacional, particularmente em aplicações onde a acurácia das previsões é indispensável e a atualização dos dados ocorre com baixa frequência, como nas taxas de mortalidade, que são atualizadas anualmente.

Ademais, o modelo com maior tempo de execução foi N-BEATS de abordagem direta, o qual levou em média 31.443 segundos para realizar treinamento e teste por conjunto de dados. Destaca-se também que a interpolação *spline* cúbica apresentou um tempo médio de execução de 0,0033 segundos, se mostrando materialmente insignificante quando comparado ao tempo de execução dos modelos de previsão. Sendo assim, a interpolação se mostra uma técnica eficiente para o sistema proposto e demais modelos univariados. Um aspecto interessante observado foi que, em alguns casos, a abordagem MIMO demonstrou um tempo médio de execução menor do que a abordagem recursiva, o que não era esperado. Esse resultado pode ser atribuído a dois fatores. Primeiramente, o processo de otimização na abordagem MIMO pode ter sido mais rápido a partir do critério de parada antecipada baseado no erro de validação. Em segundo lugar, a abordagem MIMO pode ter utilizado de forma mais eficiente os recursos computacionais, como as GPUs, ao realizar previsões em lotes para o conjunto de teste. Em contraste, a abordagem recursiva processa os dados de forma sequencial, o que tende a aumentar o tempo de execução.

Diante disso, a abordagem recursiva revelou-se como a mais econômica computacional-

Tabela 7 – Tempo médio de execução (s) do sistema proposto e demais modelos

Modelo	Tempo
ARIMA	2,29
ARIMA-LSTM	4.392,49
ARIMA-LSTM DIRETA	24.519,79
ARIMA-LSTM MIMO	3.921,29
ARIMA-MLP	2.272,27
ARIMA-MLP DIRETA	10.687,69
ARIMA-MLP MIMO	2.350,09
ARIMA-NBEATS	2.916,49
ARIMA-NBEATS DIRETA	27.820,09
ARIMA-NBEATS MIMO	3.259,09
LSTM	6.376,80
LSTM DIRETA	19.526,40
LSTM MIMO	1.867,50
MLP	2.670,60
MLP DIRETA	7.326,60
MLP MIMO	790,20
NBEATS	3.251,40
NBEATS DIRETA	31.443,00
NBEATS MIMO	2.913,60
PROPOSTO	9.581,29

**Fonte:** Dados da pesquisa (2024).

mente entre as estratégias avaliadas, representando uma porcentagem notadamente menor do tempo computacional total. Além de sua eficiência em termos de processamento, a abordagem recursiva também obteve a maior acurácia de previsão em média quando comparada às demais metodologias estudadas. Esses resultados indicam que a abordagem recursiva não apenas reduz o custo computacional, mas também proporciona previsões mais acuradas, tornando-a uma escolha oportuna para cenários que requerem múltiplas previsões à frente de taxas de mortalidade.

## 5.5 ANÁLISE DE PLAUSIBILIDADE DAS PREVISÕES

Esta seção tem como objetivo apresentar as previsões (Figuras 49a-53c) para a curva de mortalidade e avaliar sua plausibilidade. Em todos os casos analisados, abrangendo quinze conjuntos de dados, as previsões geradas pelo sistema proposto se alinham bem aos dados reais. As curvas de mortalidade na escala logarítmica, produzidas pelo sistema, exibem um comportamento decrescente nas idades mais jovens, seguido de um aumento que persiste até

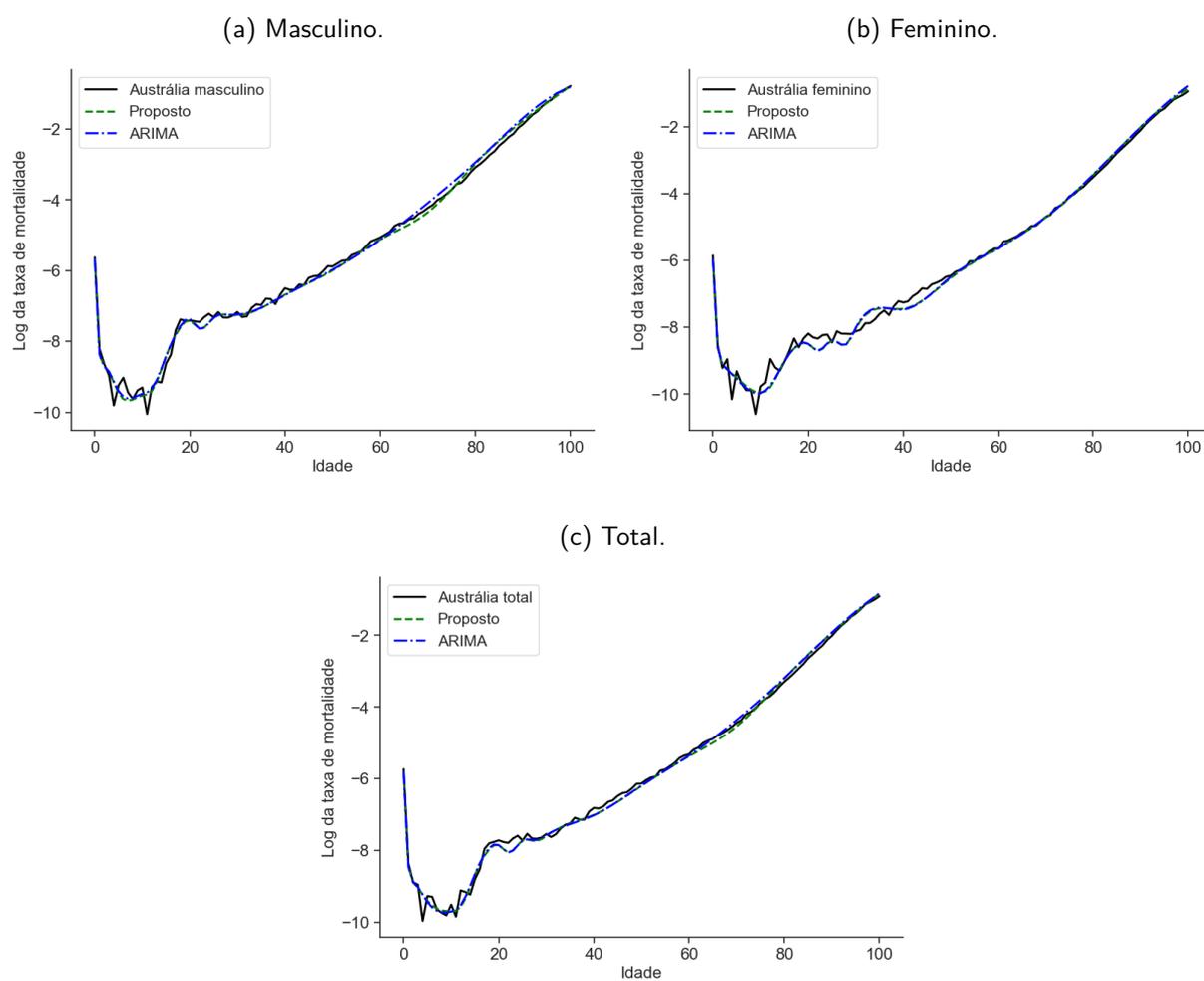
a fase adulta, conhecido como padrão "J". Além disso, o sistema híbrido proposto mostra uma monotonicidade na curva de mortalidade para as idades adultas (acima de 40 anos) e idosas.

Excetuando-se o conjunto de dados femininos do Japão (Figura 52b), tanto o sistema proposto quanto o modelo ARIMA geraram projeções mais suaves para a curva de mortalidade nas faixas etárias mais jovens. No caso específico dos dados do Japão, gênero feminino, a variabilidade das projeções foi mais intensa nas idades mais jovens, estabilizando-se por volta dos 40 anos. A partir dessa idade, as projeções tornaram-se consistentes com os valores reais da curva de mortalidade. Adicionalmente, observou-se uma subestimação sistemática nos valores da curva de mortalidade para as idades entre 20 e 40 anos nos conjuntos de dados dos EUA (Figuras 50b a 50c). Além das figuras apresentadas nesta seção, é possível ver no Anexo E as previsões das curvas de mortalidade geradas pelo sistema proposto para todos os conjuntos de dados e todos os anos do conjunto de teste (2010-2019).

O sistema proposto demonstrou consistência nas projeções de mortalidade para todos os conjuntos de dados avaliados, alinhando-se com os padrões observados na literatura (Bravo, 2007; Milne, 2010; Richmond et al., 2021). Em específico, as curvas de mortalidade exibiram uma queda inicial nas taxas para as faixas etárias mais jovens, corroborando com estudos anteriores que também observaram essa tendência (Richmond et al., 2021). Posteriormente, um aumento na mortalidade foi registrado, estabilizando-se em idades adultas e demonstrando um crescimento linear até as idades mais avançadas. Esse comportamento complexo, que vai da queda inicial à subsequente elevação e eventual estabilização e crescimento linear, é um padrão bem documentado e geralmente aceito no campo da demografia e estudos de mortalidade.

Em resumo, validou-se a eficácia do sistema híbrido proposto na previsão de taxas de mortalidade. As projeções foram não apenas estatisticamente robustas como se pode verificar na Seção 5.3, mas também plausíveis, aderindo aos padrões convencionalmente aceitos de mortalidade nas diversas fases da vida. Esse alinhamento com os padrões biológicos reforça a confiabilidade do sistema proposto e o torna uma ferramenta valiosa para pesquisas futuras e aplicações práticas em campos como demografia, saúde pública e atuária.

Figura 49 – Curva de mortalidade na escala logarítmica e previsões pelo sistema proposto e ARIMA para a Austrália no ano de 2019.



Fonte: Dados da pesquisa (2024).

Figura 50 – Curva de mortalidade na escala logarítmica e previsões pelo sistema proposto e ARIMA para os EUA no ano de 2019.

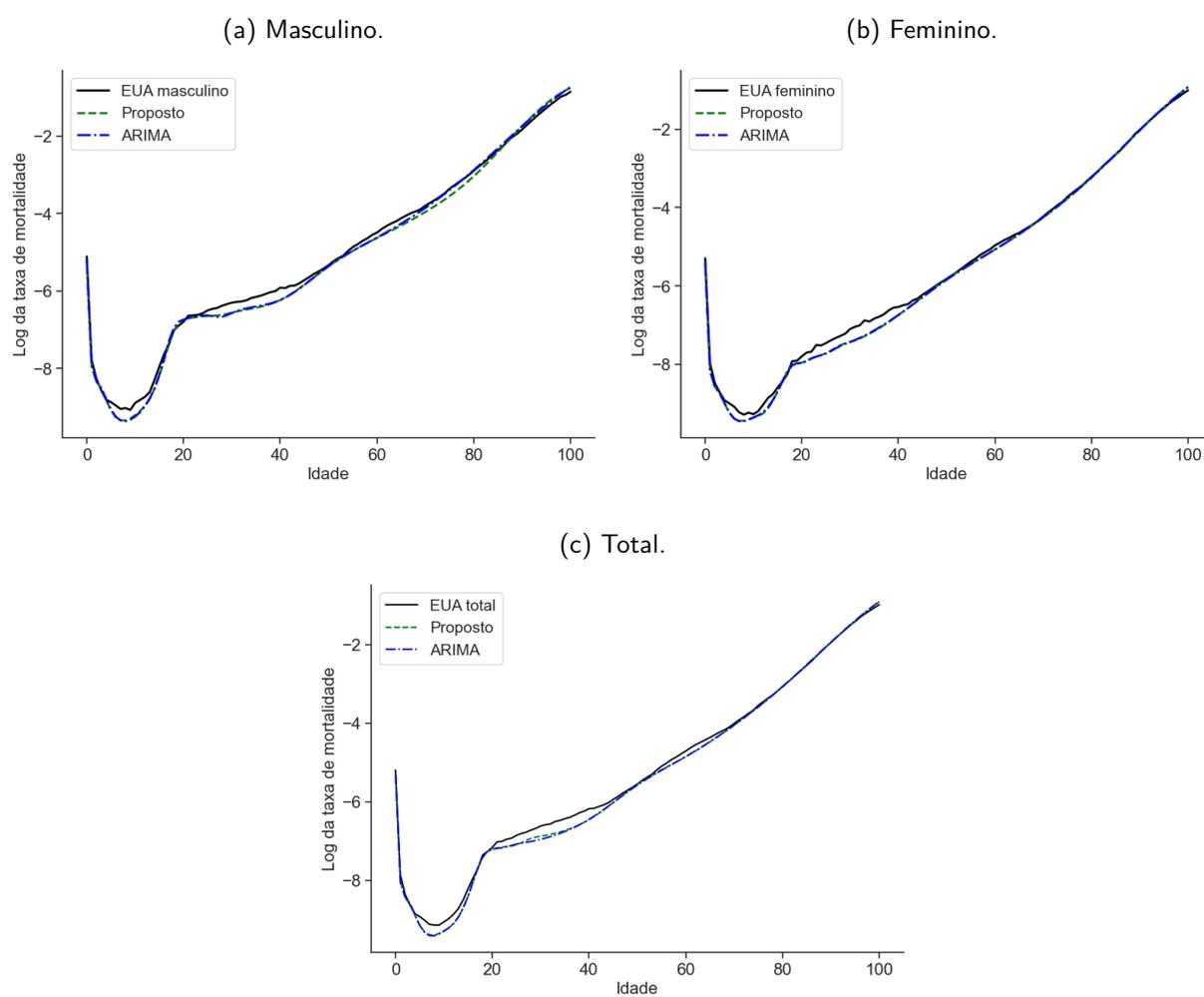


Figura 51 – Curva de mortalidade na escala logarítmica e previsões pelo sistema proposto e ARIMA para a França no ano de 2019.

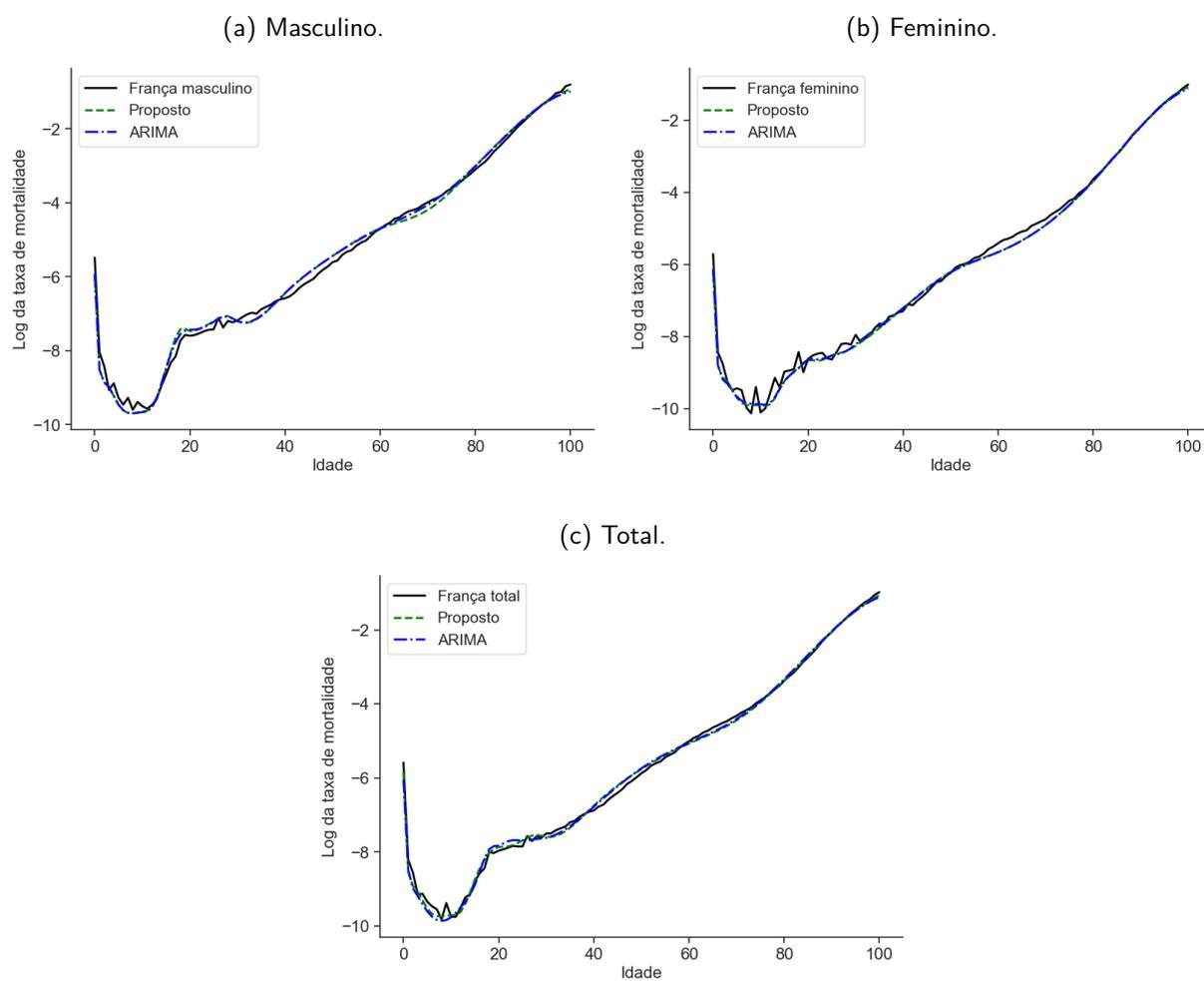


Figura 52 – Curva de mortalidade na escala logarítmica e previsões pelo sistema proposto e ARIMA para o Japão no ano de 2019.

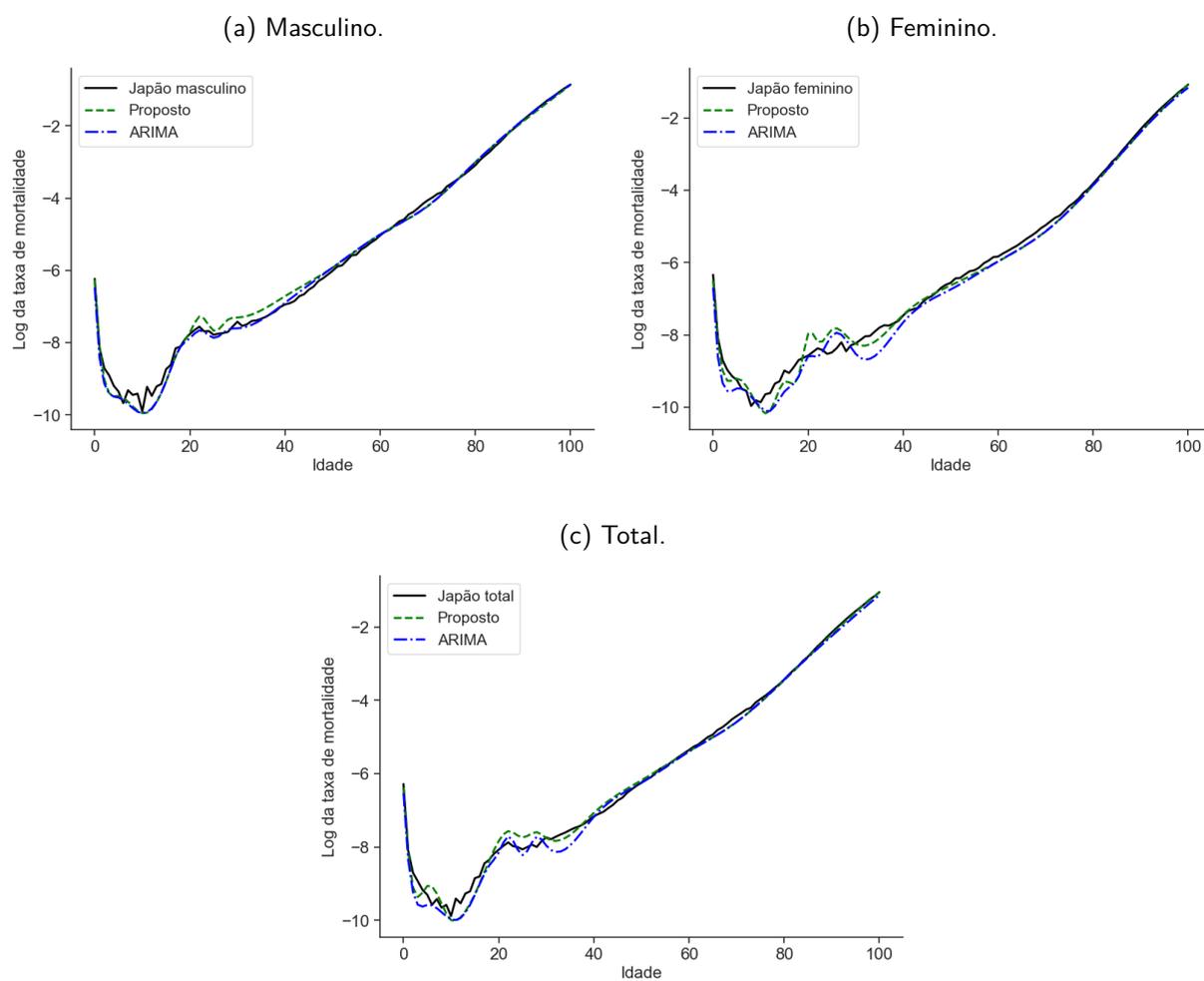
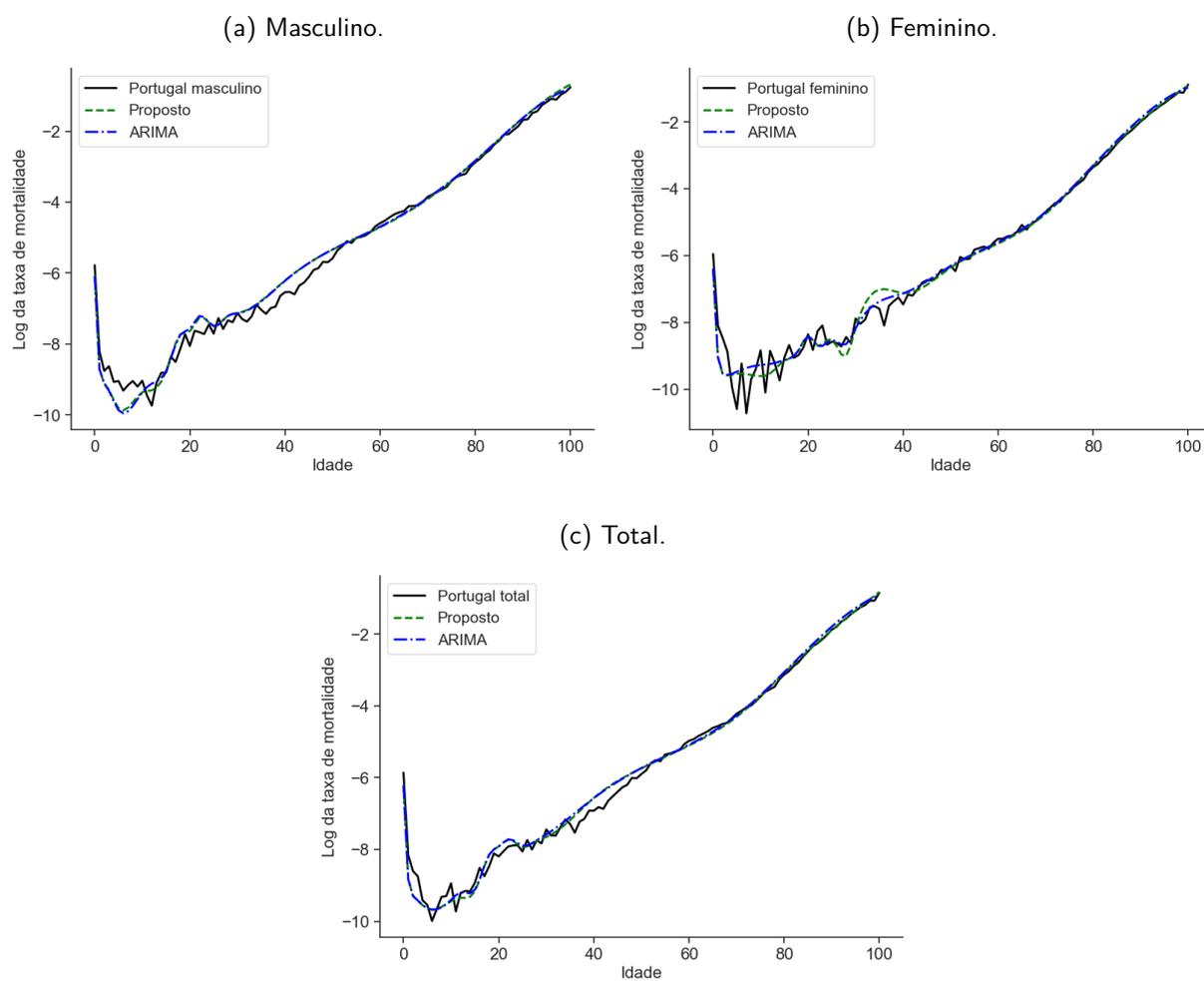


Figura 53 – Curva de mortalidade na escala logarítmica e previsões pelo sistema proposto e ARIMA para Portugal no ano de 2019.



## 6 CONCLUSÃO

Este capítulo apresenta a conclusão desta tese (6.1 Considerações finais), os principais achados da pesquisa e suas limitações metodológicas, com foco em sistemas híbridos e previsão de taxas de mortalidade e limitações (6.2). São também trazidas sugestões para próximas pesquisas (6.3 Trabalhos Futuros) envolvendo pontos de melhorias para o sistema proposto e elementos úteis para tomadores de decisão em diversas áreas. Por fim, é descrita a referência do artigo publicado fruto da pesquisa do doutorado na Seção 6.4.

### 6.1 CONSIDERAÇÕES FINAIS

Foi proposto um novo sistema híbrido para previsão de taxas de mortalidade humana. A modelagem da taxa de mortalidade é um desafio porque estas séries temporais apresentam não-linearidades, comportamento dinâmico, grupos etários com características distintas (Wu; Wang, 2018) e poucas observações. Portanto, o desenvolvimento de um sistema de previsão capaz de modelar padrões lineares e não lineares nas séries de taxas de mortalidade em um cenário de estimativa de longo prazo está no estado da arte.

O sistema proposto é uma combinação aditiva de modelo estatístico linear e modelo de AM não-linear, aproveitando as forças complementares de ambos. O modelo linear gera as estimativas para as séries das taxas de mortalidade, enquanto o modelo de AM não-linear é projetado para mapear a natureza dinâmica específica e não linear das séries residuais dos preditores lineares. O sistema híbrido proposto possui duas etapas, treinamento e teste. Na primeira etapa, também chamada de treinamento, utiliza-se a abordagem recursiva de previsão em múltiplos passos à frente para treinar um modelo estatístico linear e um *pool* de modelos não-lineares de AM. Na etapa de teste, o sistema seleciona o preditor não-linear mais acurado no conjunto de validação para todas as idades de um dado conjunto de dados, e gera a previsão final a partir da soma simples das estimativas da série e dos resíduos, aproveitando as forças complementares de ambos. O modelo ARIMA foi utilizado para modelar os padrões lineares e um *pool* de modelos (MLP, LSTM, N-BEATS) foi gerado para prever os resíduos daquele.

A proposta desta tese foi comparada com modelos monolíticos e híbridos da literatura. A avaliação experimental, realizada em quinze conjuntos de dados (Human Mortality Database, 2023), demonstrou que o sistema proposto obteve o melhor desempenho de previsão em

comparação aos modelos da literatura. Quando considerados o modelo linear generalizado e a aplicação do teste estatístico Nemenyi MCB, a proposta foi superior a todas as alternativas avaliadas. O mecanismo de seleção do modelo não-linear emergiu como o principal diferenciador da proposta, uma vez que o sistema proposto mostrou desempenho superior aos sistemas híbridos considerados como referência no processo de seleção de modelo.

Em suma, respondendo às questões de pesquisa desta tese:

- **Q.1:** Sistemas híbridos demonstraram desempenho superior em comparação com abordagens monolíticas;
- **Q.2:** Na modelagem residual, a abordagem recursiva de múltiplos passos à frente alcançou maior acurácia de previsão entre as abordagens avaliadas;
- **Q.3:** A incorporação de seleção de modelos não-lineares de AM na modelagem residual melhorou significativamente a acurácia do sistema híbrido na previsão de taxas de mortalidade em múltiplos horizontes temporais.

Essas constatações corroboram a eficácia do sistema híbrido proposto para previsão de séries curtas de taxas de mortalidade em múltiplos passos à frente, e destacam a relevância da abordagem recursiva e da seleção de modelo não-linear na modelagem residual.

Sendo assim, lista-se, a seguir, as contribuições desta tese: (i) Sistema híbrido para previsão de séries curtas de taxas de mortalidade em múltiplos passos à frente, que faz modelagem residual e combina preditores lineares e não-lineares; (ii) Na modelagem dos resíduos, utiliza-se a abordagem recursiva de previsão para vários passos à frente, a qual é mais eficiente em relação ao volume de dados; (iii) Sistema genérico que pode utilizar outros modelos lineares e não-lineares; (iv) Combinação de preditores que emprega a seleção do preditor não-linear mais promissor para prever os resíduos do preditor linear; (v) Análise estatística em quinze conjuntos de dados, dentre os quais cada um compreende cento e uma séries temporais, entre o sistema proposto, modelos individuais e sistemas híbridos.

## 6.2 LIMITAÇÕES DA PESQUISA

Uma das limitações do sistema proposto está na hipótese de que as séries temporais das taxas de mortalidade são formadas pela combinação aditiva de padrões lineares e não-lineares. Conseqüentemente, o emprego de um sistema híbrido que permita avaliar a melhor

forma de combinação dos supracitados padrões pode produzir previsões mais acuradas para algumas séries temporais de taxas de mortalidade. Outra limitação corresponde ao fato de se ter utilizado a seleção estática de preditor não-linear para produzir as estimativas dos resíduos do preditor linear. Isto posto, é possível que a seleção dinâmica de preditor pode melhorar o desempenho do sistema híbrido proposto.

Dentre as limitações do estudo realizado, destaca-se o fato de que os experimentos foram aplicados em séries temporais de tamanho amostral pequeno. Consequentemente, os resultados observados estão limitados às séries de taxas de mortalidade das populações avaliadas ou de outras populações com características similares. Outras limitações identificadas foram: (i) ausência de análise da incerteza da previsão; (ii) falta de análise de sensibilidade na escolha das idades para modelagem; (iii) não emprego da seleção dinâmica de preditores no processo de modelagem híbrida (Oliveira; Silva; de Mattos Neto, 2021); (iv) ausência de hibridização inteligente que busca a função mais adequada que otimize o desempenho da combinação (Santos Júnior; Oliveira; de Mattos Neto, 2019); (v) não realização de seleção do tamanho da janela das defasagens; (vi) não implementação de combinação de previsão por meio de componentes principais; (vii) não utilização do sistema proposto em uma aplicação do ramo de seguros, tais como no cálculo de anuidades, seguros em caso de morte e produtos financeiros indexados à longevidade humana; e (viii) falta de análise da complexidade computacional. Entretanto, no que diz respeito ao ponto (viii), i.e., avaliação do tempo de execução do algoritmo, este não se apresenta como um problema crítico para a aplicação aqui discutida, pois as previsões para as taxas de mortalidade são geradas anualmente e as séries são curtas.

### 6.3 TRABALHOS FUTUROS

Como pontos de melhoria do trabalho atual, pode-se destacar a investigação de outros modelos lineares, seleção da combinação de preditores, implementação de combinações não lineares do sistema híbrido (de Mattos Neto; Cavalcanti; Madeiro, 2017; Medina; Oliveira, 2023), combinação de sistemas híbridos com diferentes abordagens (e.g., resíduos multiplicativos, filtro de alta volatilidade, filtro de Hodrick–Prescott) (Firmino; de Mattos Neto; Ferreira, 2015; Babu; Reddy, 2014; Iftikhar et al., 2023), busca pelo tamanho da janela para os modelos não lineares das séries residuais, emprego de uma análise de sensibilidade da escolha das idades a serem modeladas pelo sistema híbrido, e a inclusão da incerteza de previsão utilizando a previsão conforme (Xu; Xie, 2023) para melhorar o processo decisório dos usuários

das previsões. Ainda, os estudos futuros abordarão a aplicação de uma hibridização entre algoritmo inteligente de otimização e modelos de AM (Jnr; Ziggah; Relvas, 2021; Zivkovic et al., 2021) para melhorar a acurácia da previsão em várias etapas (Liu; Loo; Pasupa, 2021; Niu et al., 2021). Investigação da interpretabilidade do modelo N-BEATS, no sentido de calibrar de forma mais eficaz o ajuste desse modelo, e entender como as previsões para as séries residuais estão sendo realizadas. Avaliação do desempenho do sistema proposto com séries de taxas de mortalidade por faixa etária no lugar de séries específicas por idade, haja vista que essa é a estrutura disponibilizada dos dados do Brasil. Averiguação do impacto do sistema proposto nas projeções demográficas populacionais. Também será analisada a acurácia de previsão do sistema proposto em conjuntos de dados com séries temporais curtas em outras áreas do conhecimento.

Além disso, vislumbra-se o desenvolvimento de um sistema inteligente mais sofisticado no contexto de previsão de taxas de mortalidade, auxiliando o processo decisório e capaz de ser eficazmente utilizado pela indústria. Isso incluiria a implementação de um *dashboard* interativo de suporte à tomada de decisão, fornecendo análises em tempo real e visões consolidadas das previsões. Adicionalmente, pretende-se integrar variáveis exógenas, como indicadores socioeconômicos e fatores ambientais, que possam influenciar as taxas de mortalidade. A incorporação dessas variáveis tem o potencial de tornar o sistema mais robusto e preciso, contribuindo para um avanço prático significativo na área de previsão de séries temporais em atuária, demografia, saúde pública, entre outras. Com base neste sistema preditivo inteligente, pode-se desenvolver mecanismos para calcular anuidades atuariais (i.e., pagamento de renda ou benefício em caso de sobrevivência ou falecimento) e precificar produtos financeiros vinculados à longevidade, como índices de longevidade, *survivor bond*, contratos *q-forward* e outros derivativos relacionados.

Outro produto futuro seria a aplicação de modelos hierárquicos que otimizem os parâmetros para cada subpopulação analisada. Dessa forma, um modelo global seria responsável por regular os parâmetros de modelos mais específicos, garantindo que as peculiaridades de cada grupo (e.g. gênero, região, extrato socioeconômico) sejam consideradas. Este enfoque hierárquico permitiria não apenas uma personalização mais profunda das previsões e consequentemente maior acurácia, mas também uma maior eficiência na alocação de recursos na saúde, previdência e seguros, ajudando gestores a tomar decisões mais fundamentadas e eficazes.

#### 6.4 ARTIGO PUBLICADO

O trabalho a seguir é resultado da pesquisa realizada durante o doutorado:

- "A hybrid recursive direct system for multi-step mortality rate forecasting" (Duarte; de Mattos Neto; Firmino, 2024).

Além desse trabalho, está em processo de elaboração mais um artigo o qual apresenta o sistema proposto nesta tese com os seus principais achados.

## REFERÊNCIAS

- ABADI, M.; AGARWAL, A.; BARHAM, P.; BREVDO, E.; CHEN, Z.; CITRO, C.; CORRADO, G. S.; DAVIS, A.; DEAN, J.; DEVIN, M.; GHEMAWAT, S.; GOODFELLOW, I.; HARP, A.; IRVING, G.; ISARD, M.; JIA, Y.; JOZEFOWICZ, R.; KAISER, L.; KUDLUR, M.; LEVENBERG, J.; MANÉ, D.; MONGA, R.; MOORE, S.; MURRAY, D.; OLAH, C.; SCHUSTER, M.; SHLENS, J.; STEINER, B.; SUTSKEVER, I.; TALWAR, K.; TUCKER, P.; VANHOUCHE, V.; VASUDEVAN, V.; VIÉGAS, F.; VINYALS, O.; WARDEN, P.; WATTENBERG, M.; WICKE, M.; YU, Y.; ZHENG, X. **TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems**. 2015. Software available from tensorflow.org. Disponível em: <<https://www.tensorflow.org/>>.
- ABURTO, J. M.; VILLAVICENCIO, F.; BASELLINI, U.; KJÆRGAARD, S.; VAUPEL, J. W. Dynamics of life expectancy and life span equality. **Proceedings of the National Academy of Sciences**, National Acad Sciences, v. 117, n. 10, p. 5250–5259, 2020.
- AHANI, I. K.; SALARI, M.; SHADMAN, A. Statistical models for multi-step-ahead forecasting of fine particulate matter in urban areas. **Atmospheric Pollution Research**, v. 10, n. 3, p. 689–700, 2019. ISSN 1309-1042.
- AHMED, D. M.; HASSAN, M. M.; MSTAFA, R. J. A review on deep sequential models for forecasting time series data. **Applied Computational Intelligence and Soft Computing**, Hindawi, v. 2022, 2022.
- ATANCE, D.; DEBÓN, A.; NAVARRO, E. A comparison of forecasting mortality models using resampling methods. **Mathematics**, Multidisciplinary Digital Publishing Institute, v. 8, n. 9, p. 1550, 2020.
- ATIYA, A.; EL-SHOURA, S.; SHAHEEN, S.; EL-SHERIF, M. A comparison between neural-network forecasting techniques-case study: river flow forecasting. **IEEE Transactions on Neural Networks**, v. 10, n. 2, p. 402–409, 1999.
- BABU, C. N.; REDDY, B. E. A moving-average filter based hybrid arima–ann model for forecasting time series data. **Applied Soft Computing**, Elsevier, v. 23, p. 27–38, 2014.
- BELL, W. R. Comparing and Assessing Time Series Methods for Forecasting Age-Specific Fertility and Mortality Rates. v. 13, n. 3, p. 279–303, 1997.
- BENGIO, Y.; COURVILLE, A.; VINCENT, P. Representation learning: A review and new perspectives. **IEEE transactions on pattern analysis and machine intelligence**, IEEE, v. 35, n. 8, p. 1798–1828, 2013.
- BEYAZTAS, U.; SHANG, H. Machine-learning-based functional time series forecasting: Application to age-specific mortality rates. **Forecasting**, MDPI, v. 4, n. 1, p. 394–408, 2022.
- BISHOP, C. M.; NASRABADI, N. M. **Pattern recognition and machine learning**. [S.l.]: Springer, 2006. v. 4.
- BONTEMPI, G. Long term time series prediction with multi-input multi-output local learning. **Proceedings of the 2nd European Symposium on Time Series Prediction (TSP), ESTSP08**, 01 2008.

BOOTH, H.; MAINDONALD, J.; SMITH, L. Applying Lee-Carter under conditions of variable mortality decline. **Population Studies**, v. 56, n. 3, p. 325–336, 2002.

BOX, G. E.; JENKINS, G. M.; REINSEL, G. C.; LJUNG, G. M. **Time series analysis: forecasting and control**. [S.l.]: John Wiley & Sons, 2015.

BRAVO, J. Tábuas de mortalidade contemporâneas e prospectivas: Modelos estocásticos, aplicações actuariais e cobertura do risco de longevidade. **Universidade de Évora, Évora**, 2007.

BRAVO, J. M. Forecasting longevity for financial applications: a first experiment with deep learning methods. In: SPRINGER. **Joint European Conference on Machine Learning and Knowledge Discovery in Databases**. [S.l.], 2021. p. 232–249.

BRAVO, J. M. Forecasting mortality rates with recurrent neural networks: A preliminary investigation using portuguese data. In: **CAPSI 2021 Proceedings**. [S.l.: s.n.], 2021.

BRAVO, J. M. Pricing participating longevity-linked life annuities: A bayesian model ensemble approach. **European Actuarial Journal**, Springer, v. 12, n. 1, p. 125–159, 2022.

BRAVO, J. M.; FREITAS, N. E. M. de. Valuation of longevity-linked life annuities. **Insurance: Mathematics and Economics**, Elsevier, v. 78, p. 212–229, 2018.

BROOCK, W. A.; SCHEINKMAN, J. A.; DECHERT, W. D.; LEBARON, B. A test for independence based on the correlation dimension. **Econometric Reviews**, Taylor & Francis, v. 15, n. 3, p. 197–235, 1996. Disponível em: <<https://doi.org/10.1080/07474939608800353>>.

CAIRNS, A. J.; BLAKE, D.; DOWD, K. A two-factor model for stochastic mortality with parameter uncertainty: theory and calibration. **Journal of Risk and Insurance**, Wiley Online Library, v. 73, n. 4, p. 687–718, 2006.

CASELLA, G.; BERGER, R. L. **Statistical inference**. [S.l.]: Duxbury Pacific Grove, CA, 2002. v. 2.

CERQUEIRA, V.; TORGO, L.; PINTO, F.; SOARES, C. Arbitrated ensemble for time series forecasting. In: SPRINGER. **Joint European Conference on Machine Learning and Knowledge Discovery in Databases**. [S.l.], 2017. p. 478–494.

CERQUEIRA, V.; TORGO, L.; SOARES, C. A case study comparing machine learning with statistical methods for time series forecasting: size matters. **Journal of Intelligent Information Systems**, Springer, v. 59, n. 2, p. 415–433, 2022.

CHAKRABORTY, T.; CHATTOPADHYAY, S.; GHOSH, I. Forecasting dengue epidemics using a hybrid methodology. **Physica A: Statistical Mechanics and its Applications**, Elsevier, v. 527, p. 121266, 2019.

CHANDRA, R.; GOYAL, S.; GUPTA, R. Evaluation of deep learning models for multi-step ahead time series prediction. **IEEE Access**, v. 9, p. 83105–83123, 2021.

CHEN, Y.; KHALIQ, A. Q. Comparative study of mortality rate prediction using data-driven recurrent neural networks and the lee-carter model. **Big Data and Cognitive Computing**, MDPI, v. 6, n. 4, p. 134, 2022.

- CHO, K.; MERRIENBOER, B. V.; GULCEHRE, C.; BAHDANAU, D.; BOUGARES, F.; SCHWENK, H.; BENGIO, Y. Learning phrase representations using rnn encoder-decoder for statistical machine translation. arxiv 2014. **arXiv preprint arXiv:1406.1078**, 2020.
- CRUZ-NájERA, M. A.; TREVIÑO-BERRONES, M. G.; PONCE-FLORES, M. P.; TERÁN-VILLANUEVA, J. D.; CASTÁN-ROCHA, J. A.; IBARRA-MARTÍNEZ, S.; SANTIAGO, A.; LARIA-MENCHACA, J. Short time series forecasting: Recommended methods and techniques. **Symmetry**, v. 14, n. 6, 2022. ISSN 2073-8994.
- DAVE, E.; LEONARDO, A.; JEANICE, M.; HANAFIAH, N. Forecasting indonesia exports using a hybrid model arima-lstm. **Procedia Computer Science**, Elsevier, v. 179, p. 480–487, 2021.
- DE MATTOS NETO, P. S.; OLIVEIRA, J. F. de; SANTOS JÚNIOR, D. S. de O.; SIQUEIRA, H. V.; MARINHO, M. H.; MADEIRO, F. An adaptive hybrid system using deep learning for wind speed forecasting. **Information Sciences**, v. 581, p. 495–514, 2021. ISSN 0020-0255.
- DE MATTOS NETO, P. S. G.; CAVALCANTI, G. D.; MADEIRO, F. Nonlinear combination method of forecasters applied to PM time series. **Pattern Recognition Letters**, v. 95, p. 65–72, 2017.
- DUARTE, F. C. d. L.; DE MATTOS NETO, P. S.; FIRMINO, P. R. A. A hybrid recursive-direct system for multi-step mortality rate forecasting. **The Journal of Supercomputing**, 2024.
- DUNN, P. K.; SMYTH, G. K. et al. **Generalized linear models with examples in R**. [S.l.]: Springer, 2018. v. 53.
- EIBEN, A. E.; SMITH, J. E. **Introduction to evolutionary computing**. [S.l.]: Springer, 2015.
- FAN, D.; SUN, H.; YAO, J.; ZHANG, K.; YAN, X.; SUN, Z. Well production forecasting based on arima-lstm model considering manual operations. **Energy**, Elsevier, v. 220, p. 119708, 2021.
- FARAWAY, J. J. **Extending the linear model with R: generalized linear, mixed effects and nonparametric regression models**. [S.l.]: Chapman and Hall/CRC, 2016.
- FENG, L.; SHI, Y. Forecasting mortality rates: multivariate or univariate models? **Journal of Population Research**, v. 35, n. 3, p. 289–318, sep 2018.
- FIRMINO, P. R. A.; DE MATTOS NETO, P. S.; FERREIRA, T. A. Error modeling approach to improve time series forecasters. **Neurocomputing**, Elsevier, v. 153, p. 242–254, 2015.
- FRAZIER, P. I. Bayesian optimization. **Recent Advances in Optimization and Modeling of Contemporary Problems**, p. 255–278, 2018.
- GAGE, T. B. Bio-mathematical approaches to the study of human variation in mortality. **American Journal of Physical Anthropology**, Wiley Online Library, v. 32, n. S10, p. 185–214, 1989.
- GELMAN, A.; HILL, J.; VEHTARI, A. **Regression and other stories**. [S.l.]: Cambridge University Press, 2021.

GIACOMETTI, R.; BERTOCCHI, M.; RACHEV, S. T.; FABOZZI, F. J. A comparison of the lee–carter model and ar–arch model for forecasting mortality rates. **Insurance: Mathematics and Economics**, v. 50, n. 1, p. 85–93, 2012. ISSN 0167-6687.

GOODFELLOW, I.; BENGIO, Y.; COURVILLE, A. **Deep Learning**. MIT Press, 2016. Disponível em: <<http://www.deeplearningbook.org>>.

HAINAUT, D. A neural-network analyzer for mortality forecast. **ASTIN Bulletin: The Journal of the IAA**, Cambridge University Press, v. 48, n. 2, p. 481–508, 2018.

HAJIRAHIMI, Z.; KHASHEI, M. Hybrid structures in time series modeling and forecasting: A review. **Engineering Applications of Artificial Intelligence**, Elsevier, v. 86, p. 83–106, 2019.

HAMZAÇEBİ, C.; AKAY, D.; KUTAY, F. Comparison of direct and iterative artificial neural network forecast approaches in multi-periodic time series forecasting. **Expert systems with applications**, Elsevier, v. 36, n. 2, p. 3839–3844, 2009.

HARRIS, B.; HELGERTZ, J. Urban sanitation and the decline of mortality. **The History of the Family**, Routledge, v. 24, n. 2, p. 207–226, 2019.

HATZOPOULOS, P.; HABERMAN, S. A parameterized approach to modeling and forecasting mortality. **Insurance: Mathematics and Economics**, Elsevier, v. 44, n. 1, p. 103–123, 2009.

HAYKIN, S. **Neural networks and learning machines, 3/E**. [S.I.]: Pearson Education India, 2009.

HIERRO, L. Á.; GARZÓN, A. J.; ATIENZA-MONTERO, P.; MÁRQUEZ, J. L. Predicting mortality for covid-19 in the us using the delayed elasticity method. **Scientific reports**, Nature Publishing Group, v. 10, n. 1, p. 1–6, 2020.

HOBBCRAFT, J.; MENKEN, J.; PRESTON, S. Age, period, and cohort effects in demography: a review. **Popul Index.**, 1982.

HOCHREITER, S.; SCHMIDHUBER, J. Long short-term memory. **Neural computation**, MIT press, v. 9, n. 8, p. 1735–1780, 1997.

HUMAN MORTALITY DATABASE. 2023. University of California, Berkeley (USA), and Max Plank Institute for Demographic Research (Germany). Available at <[www.mortality.org](http://www.mortality.org)>; accessed on 01/05/2023.

HUNT, A.; BLAKE, D. On the structure and classification of mortality models. **North American Actuarial Journal**, Taylor & Francis, v. 25, n. sup1, p. S215–S234, 2021.

HYNDMAN, R.; ATHANASOPOULOS, G.; BERGMEIR, C.; CACERES, G.; CHHAY, L.; O'HARA-WILD, M.; PETROPOULOS, F.; RAZBASH, S.; WANG, E.; YASMEEN, F. **forecast: Forecasting functions for time series and linear models**. [S.I.], 2024. R package version 8.23.0. Disponível em: <<https://pkg.robjhyndman.com/forecast/>>.

HYNDMAN, R.; KOEHLER, A. B.; ORD, J. K.; SNYDER, R. D. **Forecasting with exponential smoothing: the state space approach**. [S.I.]: Springer Science & Business Media, 2008.

- HYNDMAN, R. J.; ATHANASOPOULOS, G. **Forecasting: principles and practice**. [S.l.]: OTexts, 2018.
- HYNDMAN, R. J.; KHANDAKAR, Y. Automatic time series forecasting: the forecast package for r. **Journal of statistical software**, v. 27, n. 1, p. 1–22, 2008.
- HYNDMAN, R. J.; KOEHLER, A. B. Another look at measures of forecast accuracy. **International journal of forecasting**, Elsevier, v. 22, n. 4, p. 679–688, 2006.
- HYNDMAN, R. J.; ULLAH, M. S. Robust forecasting of mortality and fertility rates: A functional data approach. **Computational Statistics & Data Analysis**, v. 51, n. 10, p. 4942–4956, jun. 2007.
- IFTIKHAR, H.; DANİYAL, M.; QURESHI, M.; TAWIAH, K.; ANSAH, R. K.; AFRIYIE, J. K. A hybrid forecasting technique for infection and death from the mpox virus. **Digital Health**, SAGE Publications Sage UK: London, England, v. 9, p. 20552076231204748, 2023.
- JANSSEN, F. Advances in mortality forecasting: introduction. **Genus**, v. 74, n. 1, p. 21, s41118–018–0045–7, dez. 2018. ISSN 2035-5556. Disponível em: <<https://genus.springeropen.com/articles/10.1186/s41118-018-0045-7>>.
- JNR, E. O.-N.; ZIGGAH, Y. Y.; RELVAS, S. Hybrid ensemble intelligent model based on wavelet transform, swarm intelligence and artificial neural network for electricity demand forecasting. **Sustainable Cities and Society**, v. 66, p. 102679, 2021. ISSN 2210-6707.
- KESZTENBAUM, L.; ROSENTHAL, J.-L. **Income versus Sanitation; Mortality Decline in Paris, 1880-1914**. [S.l.], 2014. Disponível em: <<https://EconPapers.repec.org/RePEc:hal:psewpa:halshs-01018594>>.
- KHASHEI, M.; BIJARI, M. A novel hybridization of artificial neural networks and ARIMA models for time series forecasting. **Applied Soft Computing**, v. 11, n. 2, p. 2664–2675, 2011.
- KLINE, D. M. Methods for multi-step time series forecasting neural networks. In: ZHANG, G. P. (Ed.). **Neural networks in business forecasting**. Hershey, PA, USA: IGI Global, 2004. p. 226–250.
- KONING, A. J.; FRANSES, P. H.; HIBON, M.; STEKLER, H. O. The m3 competition: Statistical tests of the results. **International journal of forecasting**, Elsevier, v. 21, n. 3, p. 397–409, 2005.
- LAM, K. K.; WANG, B. Robust non-parametric mortality and fertility modelling and forecasting: Gaussian process regression approaches. **Forecasting**, Multidisciplinary Digital Publishing Institute, v. 3, n. 1, p. 207–227, 2021.
- LECUN, Y.; BENGIO, Y.; HINTON, G. Deep learning. **nature**, Nature Publishing Group UK London, v. 521, n. 7553, p. 436–444, 2015.
- LEE, R. D.; CARTER, L. C. Modeling and forecasting u. s. mortality. **Journal of the American Statistical Association**, v. 87, n. 419, p. 659–671, 1992.
- LEE, R. D.; MILLER, T. Evaluating the Performance of the Lee-Carter Method for Forecasting Mortality. **Demography**, v. 38, n. 4, p. 537–549, 2001.

- LEVANTESI, S.; NIGRI, A. A random forest algorithm to improve the lee–carter mortality forecasting: impact on q-forward. **Soft Computing**, Springer, v. 24, n. 12, p. 8553–8567, 2020.
- LINDHOLM, M.; PALMBORG, L. Efficient use of data for lstm mortality forecasting. **European Actuarial Journal**, Springer, p. 1–30, 2022.
- LIU, Z.; LOO, C. K.; PASUPA, K. A novel error-output recurrent two-layer extreme learning machine for multi-step time series prediction. **Sustainable Cities and Society**, v. 66, p. 102613, 2021. ISSN 2210-6707.
- LIVIERIS, I. E.; PINTELAS, P. A novel multi-step forecasting strategy for enhancing deep learning models' performance. **Neural Computing and Applications**, Springer, p. 1–18, 2022.
- LÜTKEPOHL, H.; XU, F. The role of the log transformation in forecasting economic variables. **Empirical Economics**, Springer, v. 42, n. 3, p. 619–638, 2012.
- MAKRIDAKIS, S.; HIBON, M. The m3-competition: results, conclusions and implications. **International journal of forecasting**, Elsevier, v. 16, n. 4, p. 451–476, 2000.
- MAKRIDAKIS, S.; HYNDMAN, R. J.; PETROPOULOS, F. Forecasting in social settings: The state of the art. **International Journal of Forecasting**, v. 36, n. 1, p. 15–28, jan 2020.
- MAKRIDAKIS, S.; SPILIOTIS, E.; ASSIMAKOPOULOS, V. The m4 competition: 100,000 time series and 61 forecasting methods. **International Journal of Forecasting**, Elsevier, v. 36, n. 1, p. 54–74, 2020.
- MCCULLAGH, P.; NELDER, J. A. Generalized linear models 2nd edition chapman and hall. **London, UK**, 1989.
- MCDONALD, B.; JR, W. T. Rank sum multiple comparisons in one-and two-way classifications. **Biometrika**, Oxford University Press, v. 54, n. 3-4, p. 487–497, 1967.
- MCNOWN, R.; ROGERS, A. Forecasting Mortality: A Parameterized Time Series Approach. **Demography**, v. 26, n. 4, nov. 1989. ISSN 00703370.
- MEDINA, M. C. C.; OLIVEIRA, J. F. L. de. A selective hybrid system for state-of-charge forecasting of lithium–ion batteries. **The Journal of Supercomputing**, Springer, v. 79, n. 14, p. 15623–15642, 2023.
- MENDES-MOREIRA, J.; SOARES, C.; JORGE, A. M.; SOUSA, J. F. D. Ensemble approaches for regression: A survey. **ACM Computing Surveys (CSUR)**, ACM, v. 45, n. 1, p. 10, 2012.
- MENG, H.; HAN, L.; HOU, L. An ensemble learning-based short-term load forecasting on small datasets. In: IEEE. **2022 IEEE 33rd Annual International Symposium on Personal, Indoor and Mobile Radio Communications (PIMRC)**. [S.l.], 2022. p. 346–350.
- MILNE, E. M. Dynamics of human mortality. **Experimental Gerontology**, Elsevier, v. 45, n. 3, p. 180–187, 2010.

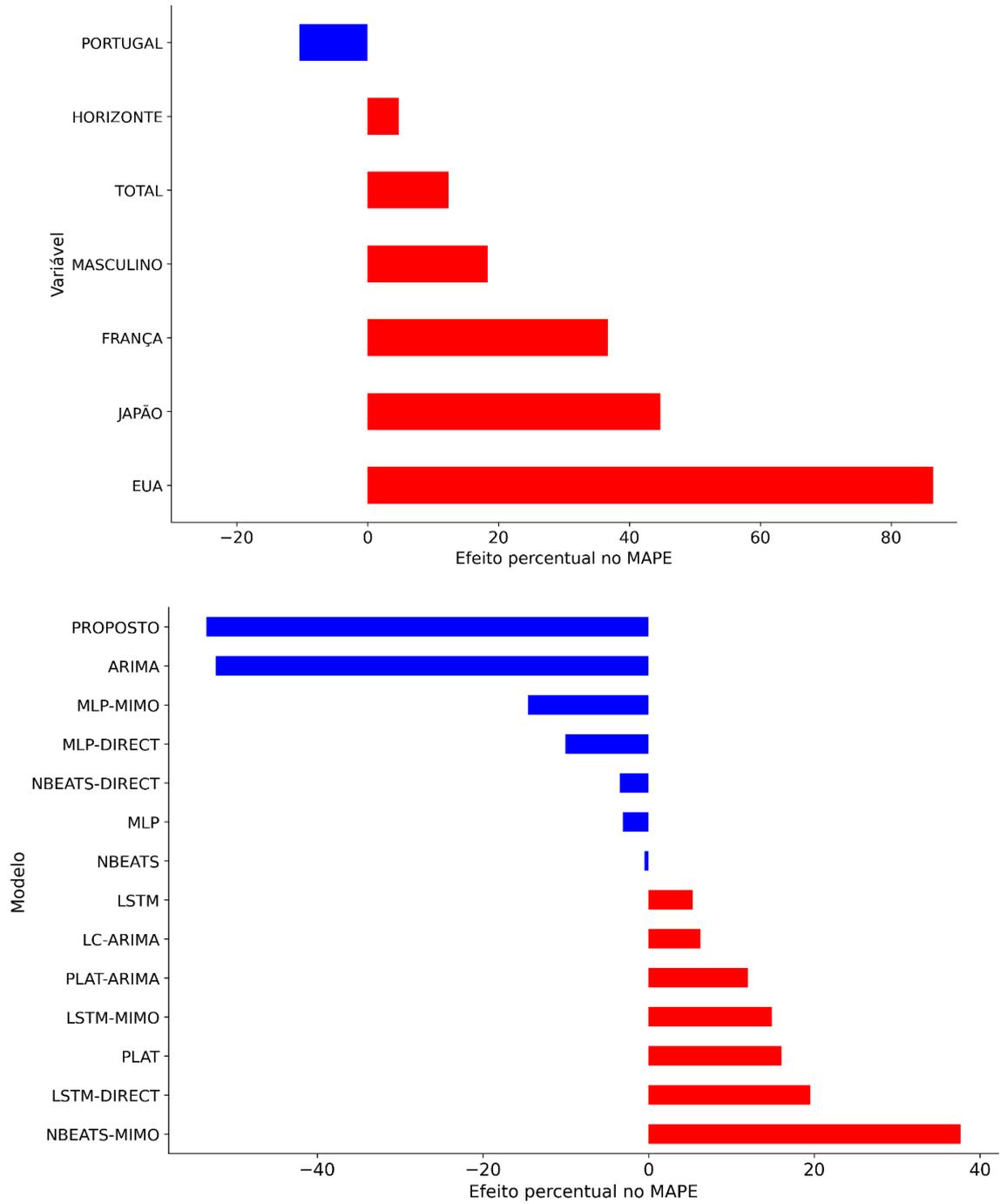
- MING, W.; BAO, Y.; HU, Z.; XIONG, T. et al. Multistep-ahead air passengers traffic prediction with hybrid arima-svms models. **The Scientific World Journal**, Hindawi, v. 2014, 2014.
- MITCHELL, D.; BROCKETT, P.; MENDOZA-ARRIAGA, R.; MUTHURAMAN, K. Modeling and forecasting mortality rates. **Insurance: Mathematics and economics**, Elsevier, v. 52, n. 2, p. 275–285, 2013.
- MITCHELL, T. M. et al. Machine learning. 1997. **Burr Ridge, IL: McGraw Hill**, v. 45, n. 37, 1997.
- MOCKUS, J.; VYTAUTAS, T.; ZILINSKAS, A. The application of bayesian methods for seeking the extremum. In: DIXON, L. C. W.; SZEGÖ, G. P. (Ed.). **Towards Global Optimization 2**. Amsterdam: North-Holland, 1978. p. 117–129.
- MONTGOMERY, D. C.; PECK, E. A.; VINING, G. G. **Introduction to linear regression analysis**. [S.l.]: John Wiley & Sons, 2021.
- MORAIS, L. R. de A.; GOMES, G. S. da S. Forecasting daily covid-19 cases in the world with a hybrid arima and neural network model. **Applied Soft Computing**, Elsevier, v. 126, p. 109315, 2022.
- MORETTIN, P. A.; TOLOI, C. M. **Análise de séries temporais: modelos lineares univariados**. 2. ed. [S.l.]: Editora Blucher, 2006.
- MURPHY, K. P. **Machine learning: a probabilistic perspective**. [S.l.]: MIT press, 2012.
- NIGRI, A.; LEVANTESI, S.; ABURTO, J. M. Leveraging deep neural networks to estimate age-specific mortality from life expectancy at birth. **Demographic Research**, JSTOR, v. 47, p. 199–232, 2022.
- NIGRI, A.; LEVANTESI, S.; MARINO, M.; SCOGNAMIGLIO, S.; PERLA, F. A deep learning integrated lee–carter model. **Risks**, Multidisciplinary Digital Publishing Institute, v. 7, n. 1, p. 33, 2019.
- NIU, W. jing; FENG, Z. kai; FENG, B. fei; XU, Y. shan; MIN, Y. wu. Parallel computing and swarm intelligence based artificial intelligence model for multi-step-ahead hydrological time series prediction. **Sustainable Cities and Society**, v. 66, p. 102686, 2021. ISSN 2210-6707.
- OLAH, C. **Understanding LSTM Networks**. 2015. Acessado em 15 de agosto de 2023. Disponível em: <<https://colah.github.io/posts/2015-08-Understanding-LSTMs/>>.
- OLIVEIRA, J. F. de; SILVA, E. G.; DE MATTOS NETO, P. S. A hybrid system based on dynamic selection for time series forecasting. **IEEE Transactions on Neural Networks and Learning Systems**, IEEE, 2021.
- OLIVIERI, A. Uncertainty in mortality projections: an actuarial perspective. **Insurance: mathematics and economics**, Elsevier, v. 29, n. 2, p. 231–245, 2001.
- ORESHKIN, B. N.; CARPOV, D.; CHAPADOS, N.; BENGIO, Y. N-beats: Neural basis expansion analysis for interpretable time series forecasting. **arXiv preprint arXiv:1905.10437**, 2019.

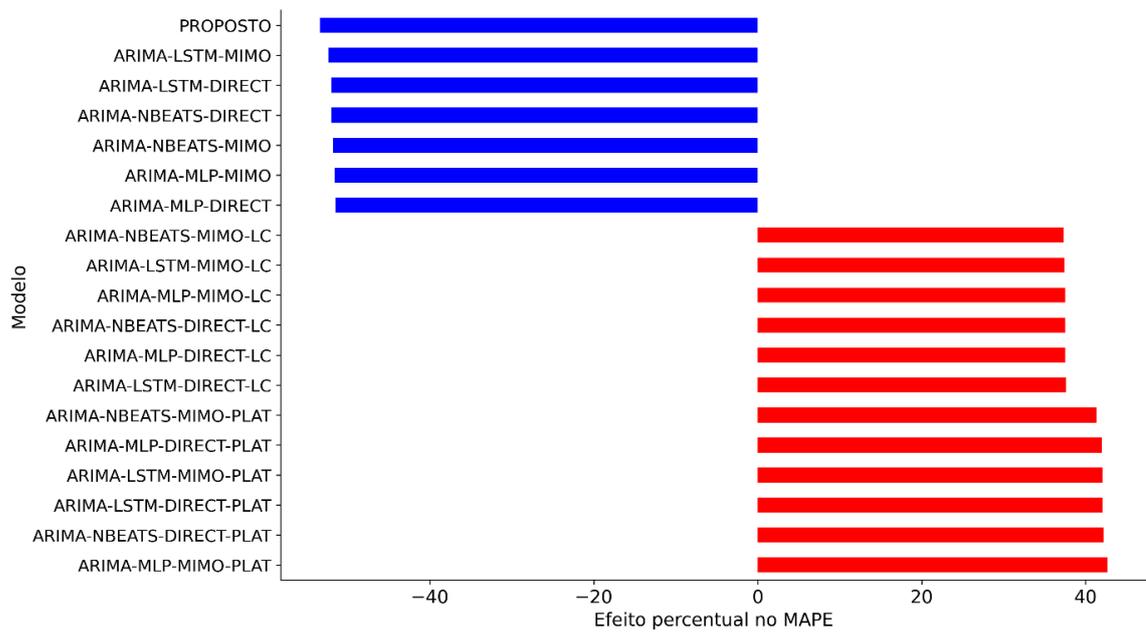
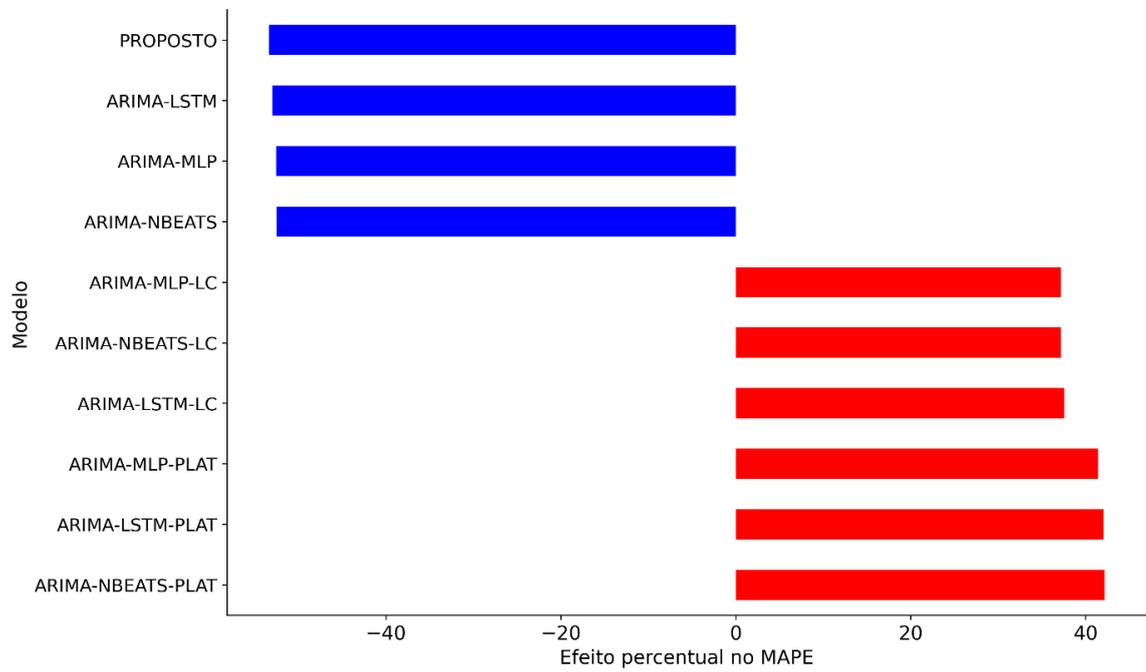
- ORESHKIN, B. N.; DUDEK, G.; PEŃKA, P.; TURKINA, E. N-beats neural network for mid-term electricity load forecasting. **Applied Energy**, v. 293, p. 116918, 2021. ISSN 0306-2619.
- PAI, P.-F.; LIN, C.-S. A hybrid ARIMA and support vector machines model in stock price forecasting. **Omega**, v. 33, n. 6, p. 497–505, 2005.
- PANIGRAHI, S.; BEHERA, H. S. A hybrid ets–ann model for time series forecasting. **Engineering applications of artificial intelligence**, Elsevier, v. 66, p. 49–59, 2017.
- PASCARIU, M. D.; LENART, A.; CANUDAS-ROMO, V. The maximum entropy mortality model: Forecasting mortality using statistical moments. **Scandinavian Actuarial Journal**, Taylor & Francis, v. 2019, n. 8, p. 661–685, 2019.
- PERLA, F.; RICHMAN, R.; SCOGNAMIGLIO, S.; WÜTHRICH, M. V. Time-series forecasting of mortality rates using deep learning. **Scandinavian Actuarial Journal**, Taylor & Francis, v. 2021, n. 7, p. 572–598, 2021.
- PLAT, R. On stochastic mortality modeling. **Insurance: Mathematics and Economics**, Elsevier, v. 45, n. 3, p. 393–404, 2009.
- PREEZ, J. D.; WITT, S. F. Univariate versus multivariate time series forecasting: an application to international tourism demand. **International Journal of Forecasting**, Elsevier, v. 19, n. 3, p. 435–451, 2003.
- RENSHAW, A.; HABERMAN, S. A cohort-based extension to the Lee–Carter model for mortality reduction factors. **Insurance: Mathematics and Economics**, v. 38, n. 3, p. 556–570, jun. 2006.
- RICHMAN, R.; WÜTHRICH, M. V. A neural network extension of the lee–carter model to multiple populations. **Annals of Actuarial Science**, Cambridge University Press, v. 15, n. 2, p. 346–366, 2021.
- RICHMOND, P.; ROEHNER, B. M.; IRANNEZHAD, A.; HUTZLER, S. Mortality: A physics perspective. **Physica A: Statistical Mechanics and its Applications**, Elsevier, v. 566, p. 125660, 2021.
- ROSHANI, A.; IZADI, M.; KHALEDI, B.-E. Transformer self-attention network for forecasting mortality rates. **Journal of the Iranian Statistical Society**, Iranian Statistical Society, v. 21, n. 1, p. 81–103, 2022.
- RUBIO, L.; ALBA, K. Forecasting selected colombian shares using a hybrid arima-svr model. **Mathematics**, MDPI, v. 10, n. 13, p. 2181, 2022.
- SANTOS JÚNIOR, D. S. d. O. **Método de Ensemble para correção de modelos ARIMA: uma abordagem de sistema híbrido para previsão de séries temporais**. Tese (Doutorado) — Universidade Federal de Pernambuco, 2022.
- SANTOS JÚNIOR, D. S. d. O.; DE MATTOS NETO, P. S.; OLIVEIRA, J. F. de; CAVALCANTI, G. D. A hybrid system based on ensemble learning to model residuals for time series forecasting. **Information Sciences**, Elsevier, v. 649, p. 119614, 2023.

- SANTOS JÚNIOR, D. S. d. O.; OLIVEIRA, J. F. d.; DE MATTOS NETO, P. S. G. An intelligent hybridization of ARIMA with machine learning models for time series forecasting. **Knowledge-Based Systems**, v. 175, p. 72–86, 2019.
- SHI, J.; GUO, J.; ZHENG, S. Evaluation of hybrid forecasting approaches for wind speed and power generation time series. **Renewable and Sustainable Energy Reviews**, Elsevier, v. 16, n. 5, p. 3471–3480, 2012.
- SHI, Y. Forecasting mortality rates with the penalized exponential smoothing state space model. **Journal of the Operational Research Society**, Taylor & Francis, p. 1–18, 2021.
- SILVA, E. G.; CAVALCANTI, G. D.; OLIVEIRA, J. F. L. de; NETO, P. S. de M. On the evaluation of dynamic selection parameters for time series forecasting. In: IEEE. **2020 International Joint Conference on Neural Networks (IJCNN)**. [S.l.], 2020. p. 1–7.
- SMITH, T. G. et al. **pmdarima: ARIMA estimators for Python**. 2017–. Disponível em: <<http://www.alkaline-ml.com/pmdarima>>.
- SNOEK, J.; LAROCHELLE, H.; ADAMS, R. P. Practical bayesian optimization of machine learning algorithms. In: PEREIRA, F.; BURGESS, C.; BOTTOU, L.; WEINBERGER, K. (Ed.). **Advances in Neural Information Processing Systems**. Curran Associates, Inc., 2012. v. 25. Disponível em: <[https://proceedings.neurips.cc/paper\\_files/paper/2012/file/05311655a15b75fab86956663e1819cd-Paper.pdf](https://proceedings.neurips.cc/paper_files/paper/2012/file/05311655a15b75fab86956663e1819cd-Paper.pdf)>.
- SORJAMAA, A.; HAO, J.; REYHANI, N.; JI, Y.; LENDASSE, A. Methodology for long-term prediction of time series. **Neurocomputing**, Elsevier, v. 70, n. 16-18, p. 2861–2869, 2007.
- SU, J.; LIN, Z.; XU, F.; FATHI, G.; ALNOWIBET, K. A. A hybrid model of arima and mlp with a grasshopper optimization algorithm for time series forecasting of water quality. **Scientific Reports**, Nature Publishing Group UK London, v. 14, n. 1, p. 23927, 2024.
- TAIEB, S. B. Machine learning strategies for multi-step-ahead time series forecasting. **Universit Libre de Bruxelles, Belgium**, p. 75–86, 2014.
- TAIEB, S. B.; BONTEMPI, G.; ATIYA, A. F.; SORJAMAA, A. A review and comparison of strategies for multi-step ahead time series forecasting based on the nn5 forecasting competition. **Expert systems with applications**, Elsevier, v. 39, n. 8, p. 7067–7083, 2012.
- THOMAKOS, D.; WOOD, G.; IOAKIMIDIS, M.; PAPAGIANNAKIS, G. Shots forecasting: Short time series forecasting for management research. **British Journal of Management**, Wiley Online Library, v. 34, n. 2, p. 539–554, 2023.
- VANELLA, P.; DESCHERMEIER, P.; WILKE, C. B. An overview of population projections—methodological concepts, international data availability, and use cases. **Forecasting**, MDPI, v. 2, n. 3, p. 346–363, 2020.
- VASWANI, A.; SHAZEER, N.; PARMAR, N.; USZKOREIT, J.; JONES, L.; GOMEZ, A. N.; KAISER, L. u.; POLOSUKHIN, I. Attention is all you need. In: GUYON, I.; LUXBURG, U. V.; BENGIO, S.; WALLACH, H.; FERGUS, R.; VISHWANATHAN, S.; GARNETT, R. (Ed.). **Advances in Neural Information Processing Systems**. Curran Associates, Inc., 2017. v. 30. Disponível em: <[https://proceedings.neurips.cc/paper\\_files/paper/2017/file/3f5ee243547dee91fbd053c1c4a845aa-Paper.pdf](https://proceedings.neurips.cc/paper_files/paper/2017/file/3f5ee243547dee91fbd053c1c4a845aa-Paper.pdf)>.

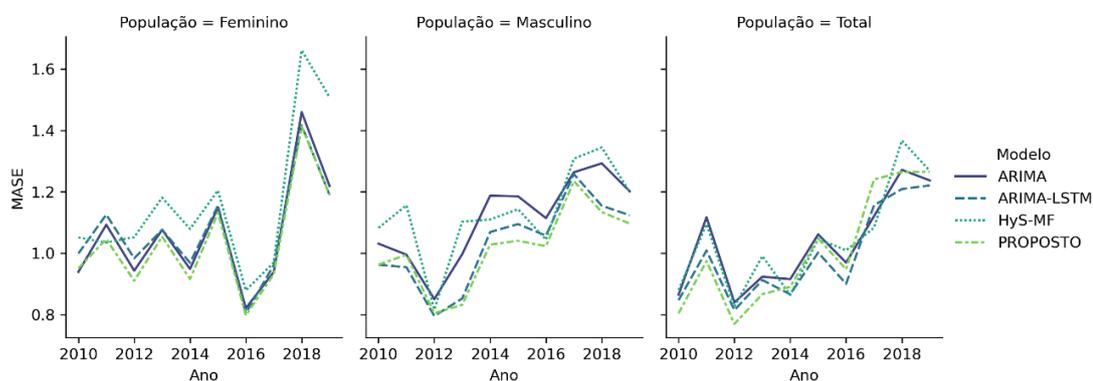
- VILLEGAS, A. M.; KAISHEV, V. K.; MILLOSOVICH, P. StMoMo: An R package for stochastic mortality modeling. **Journal of Statistical Software**, v. 84, n. 3, p. 1–38, 2018.
- WALK, F. van de. Infant mortality and the european demographic transition. In: **The decline of fertility in Europe**. [S.l.]: Princeton University Press, 2017. p. 201–233.
- WANG, J.; WEN, L.; XIAO, L.; WANG, C. Time-series forecasting of mortality rates using transformer. **Scandinavian Actuarial Journal**, Taylor & Francis, v. 2024, n. 2, p. 109–123, 2024.
- WEIGAND, A. C.; LANGE, D.; RAUSCHENBERGER, M. How can small data sets be clustered? Gesellschaft für Informatik eV, 2021.
- WILLETTS, R. C. The cohort effect: Insights and explanations. **British Actuarial Journal**, Cambridge University Press, v. 10, n. 4, p. 833–877, 2004.
- WOLPERT, D. H. The lack of a priori distinctions between learning algorithms. **Neural computation**, MIT Press, v. 8, n. 7, p. 1341–1390, 1996.
- WU, R.; WANG, B. Gaussian process regression method for forecasting of mortality rates. **Neurocomputing**, v. 316, p. 232–239, 2018.
- XU, C.; XIE, Y. Conformal prediction for time series. **IEEE Transactions on Pattern Analysis and Machine Intelligence**, IEEE, 2023.
- XU, D.; ZHANG, Q.; DING, Y.; ZHANG, D. Application of a hybrid arima-lstm model based on the spei for drought forecasting. **Environmental Science and Pollution Research**, Springer, v. 29, n. 3, p. 4128–4144, 2022.
- YAMAK, P. T.; YUJIAN, L.; GADOSEY, P. K. A comparison between arima, lstm, and gru for time series forecasting. In: **Proceedings of the 2019 2nd international conference on algorithms, computing and artificial intelligence**. [S.l.: s.n.], 2019. p. 49–55.
- ZHANG, A.; LIPTON, Z. C.; LI, M.; SMOLA, A. J. Dive into deep learning. **arXiv preprint arXiv:2106.11342**, 2021.
- ZHANG, G.; PATUWO, B. E.; HU, M. Y. Forecasting with artificial neural networks: The state of the art. **International journal of forecasting**, Elsevier, v. 14, n. 1, p. 35–62, 1998.
- ZHANG, G. P. Time series forecasting using a hybrid ARIMA and neural network model. **Neurocomputing**, Elsevier, v. 50, p. 159–175, 2003.
- ZIVKOVIC, M.; BACANIN, N.; VENKATACHALAM, K.; NAYYAR, A.; DJORDJEVIC, A.; STRUMBERGER, I.; AL-TURJMAN, F. Covid-19 cases prediction by using hybrid machine learning and beetle antennae search approach. **Sustainable Cities and Society**, v. 66, p. 102669, 2021. ISSN 2210-6707.
- ZUNINO, L.; KULP, C. W. Detecting nonlinearity in short and noisy time series using the permutation entropy. **Physics Letters A**, v. 381, n. 42, p. 3627–3635, 2017. ISSN 0375-9601. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0375960117308976>>.

## APÊNDICE A – EFEITO PERCENTUAL MÉDIO NO MASE OBTIDO POR MEIO DA REGRESSÃO GAMA

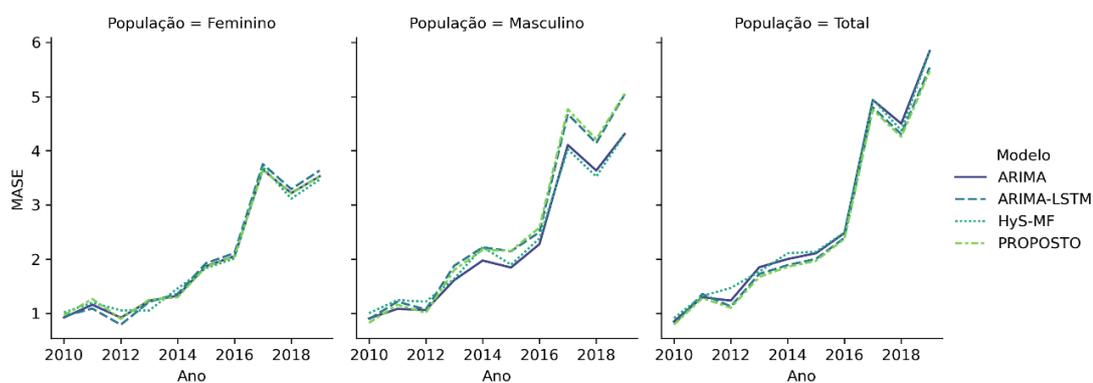




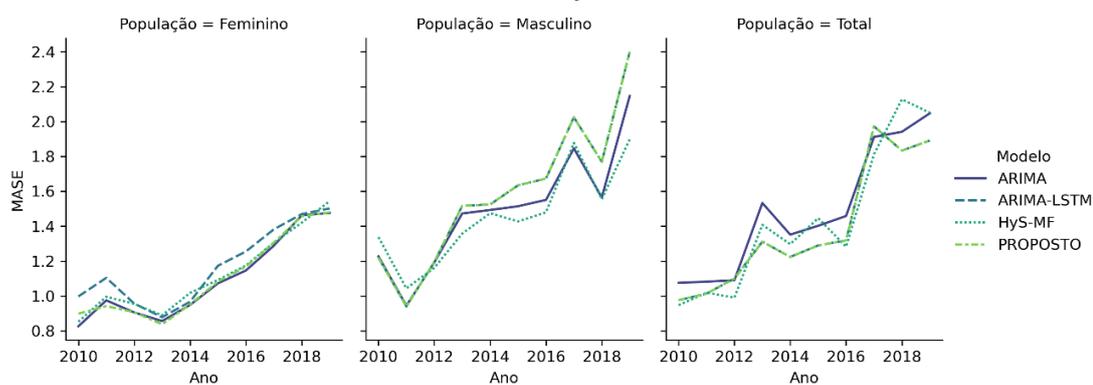
**APÊNDICE B – MASE PARA OS MODELOS ARIMA, ARIMA-LSTM, HYS-MF E PROPOSTO SEGREGADO POR CONJUNTO DE DADOS**  
**Austrália**



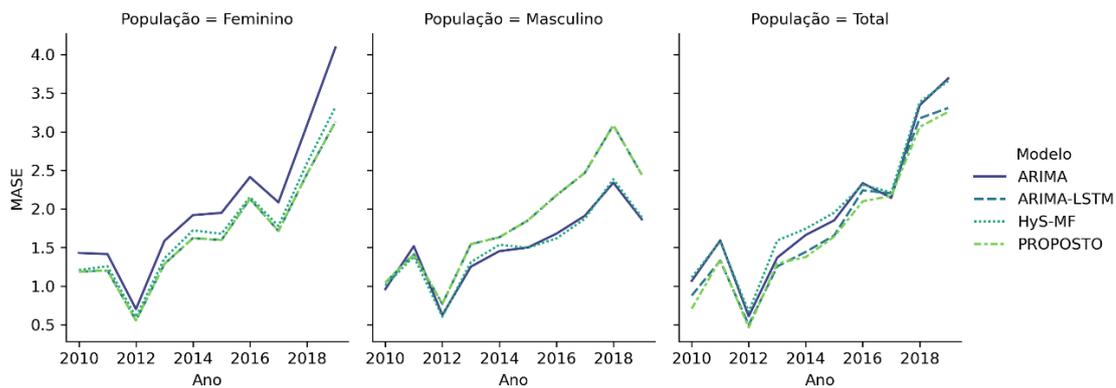
**Estados Unidos**



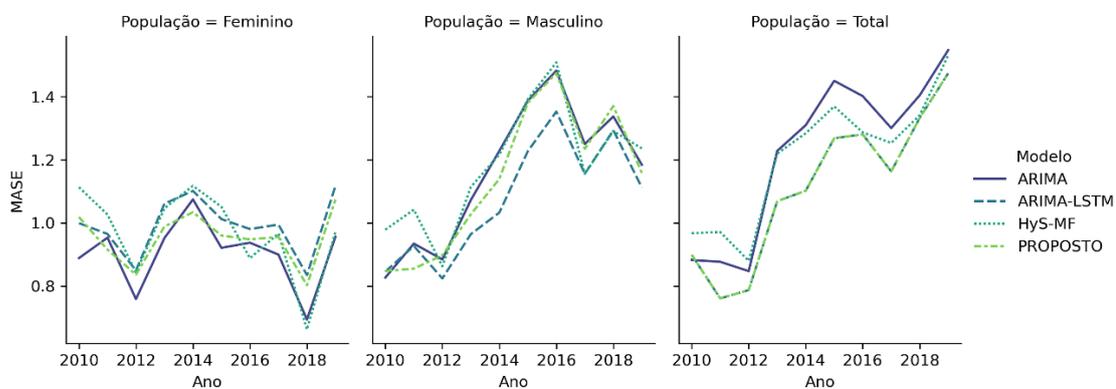
**França**



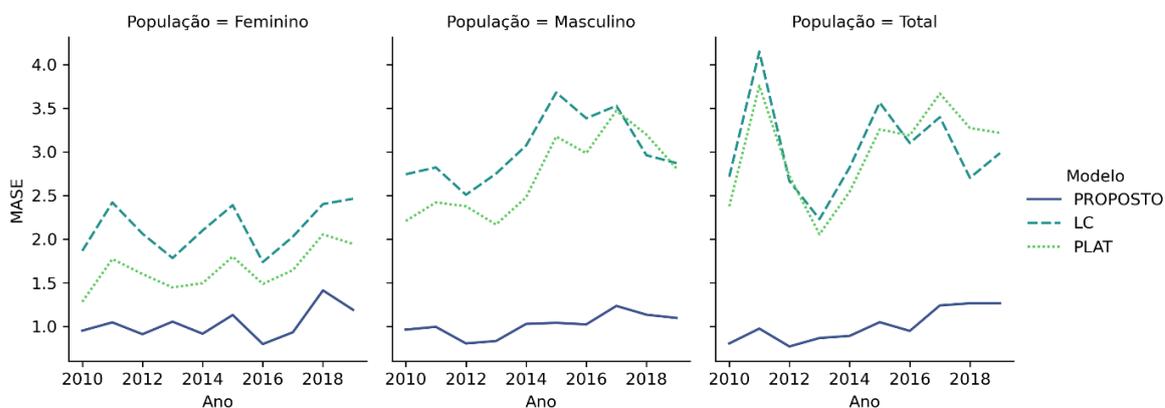
## Japão



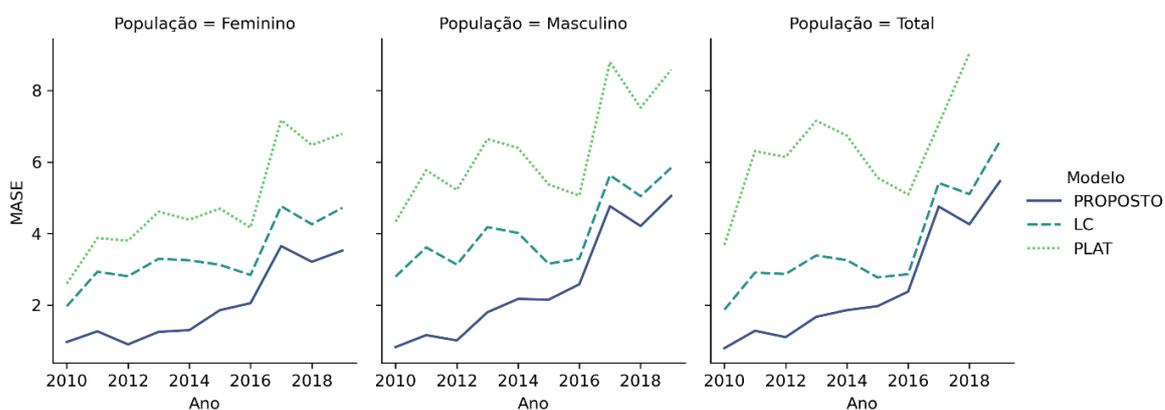
## Portugal



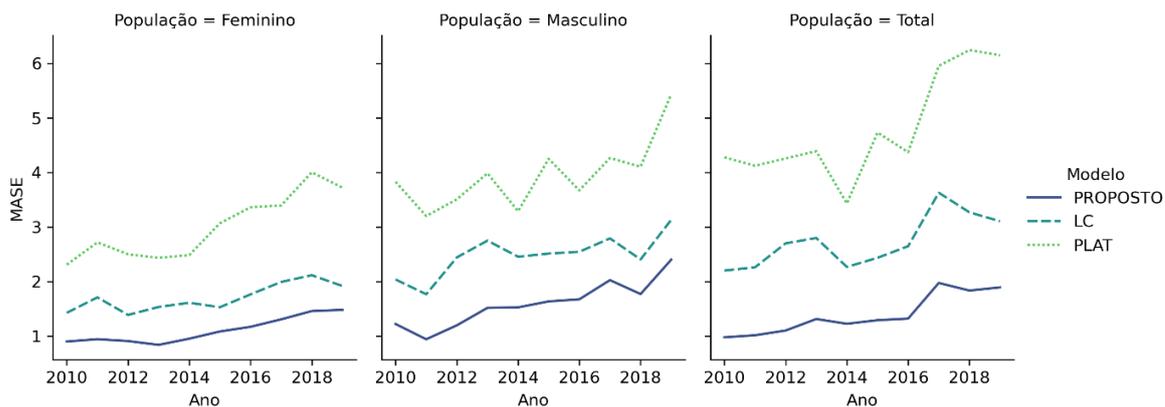
**APÊNDICE C – MASE PARA OS MODELOS PROPOSTO, LC E PLAT  
SEGREGADO POR CONJUNTO DE DADOS  
Austrália**



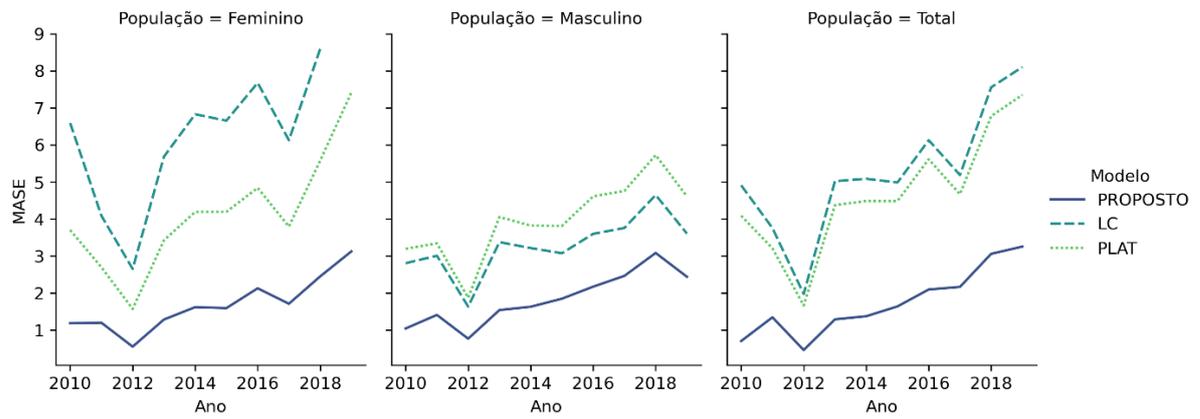
**Estados Unidos**



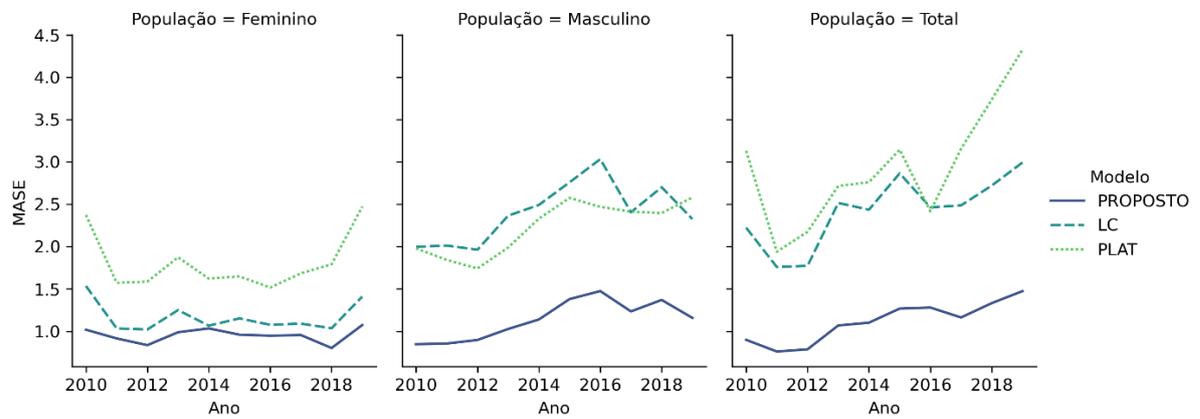
**França**



## Japão

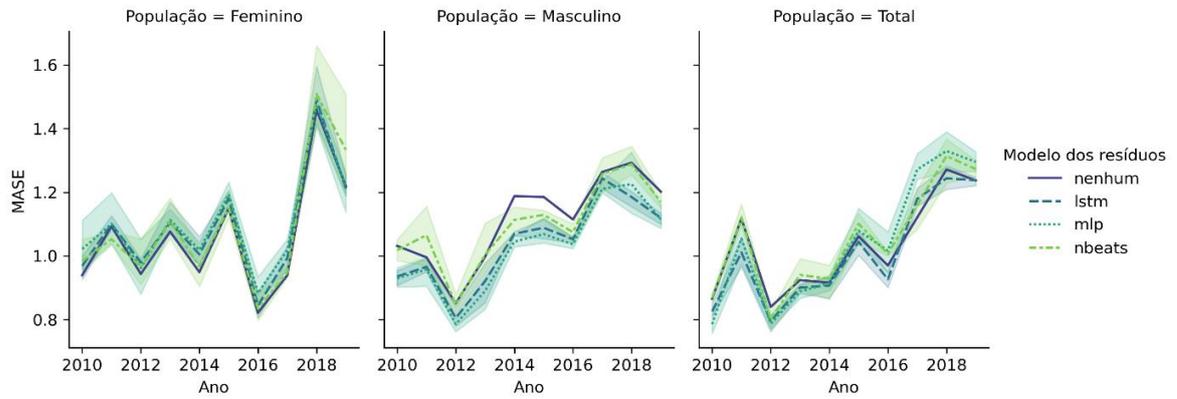


## Portugal

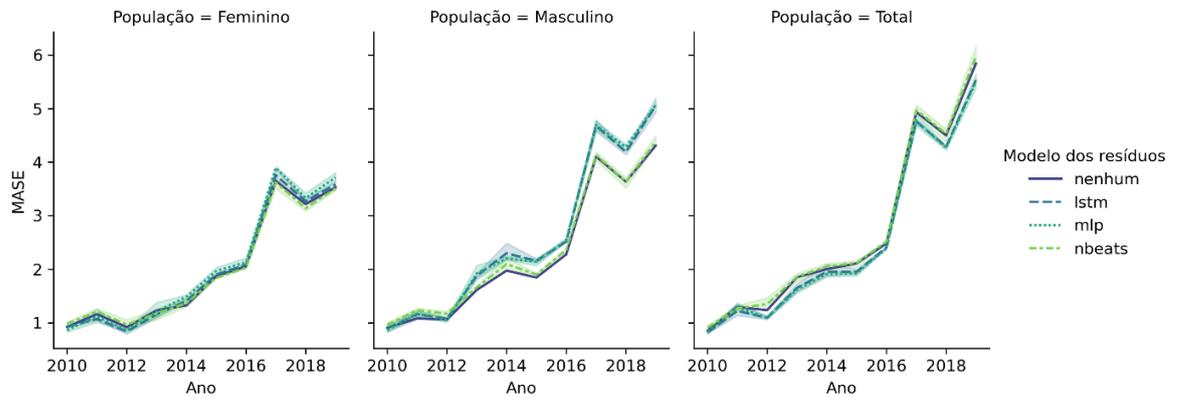


## APÊNDICE D – MASE PARA MODELOS UNIVARIADOS SEGREGADOS POR PAÍS, GÊNERO E MODELAGEM RESIDUAL

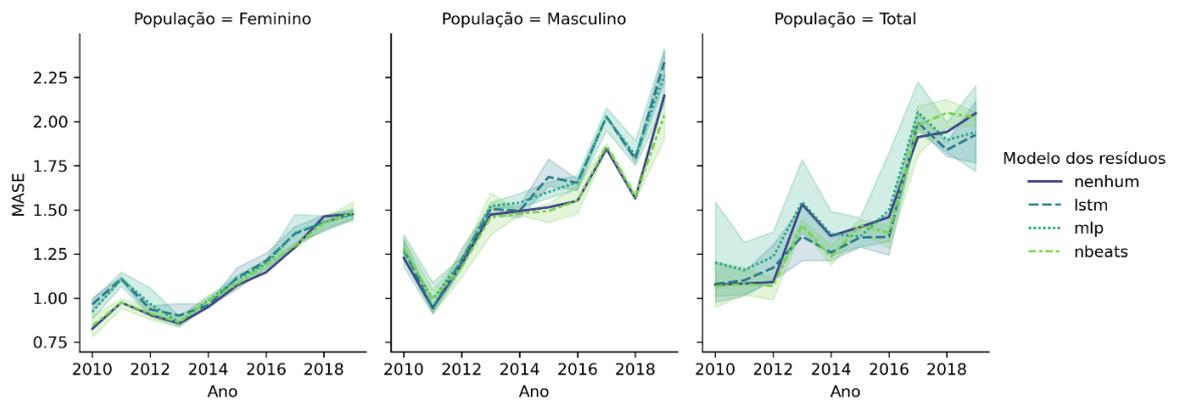
### Austrália

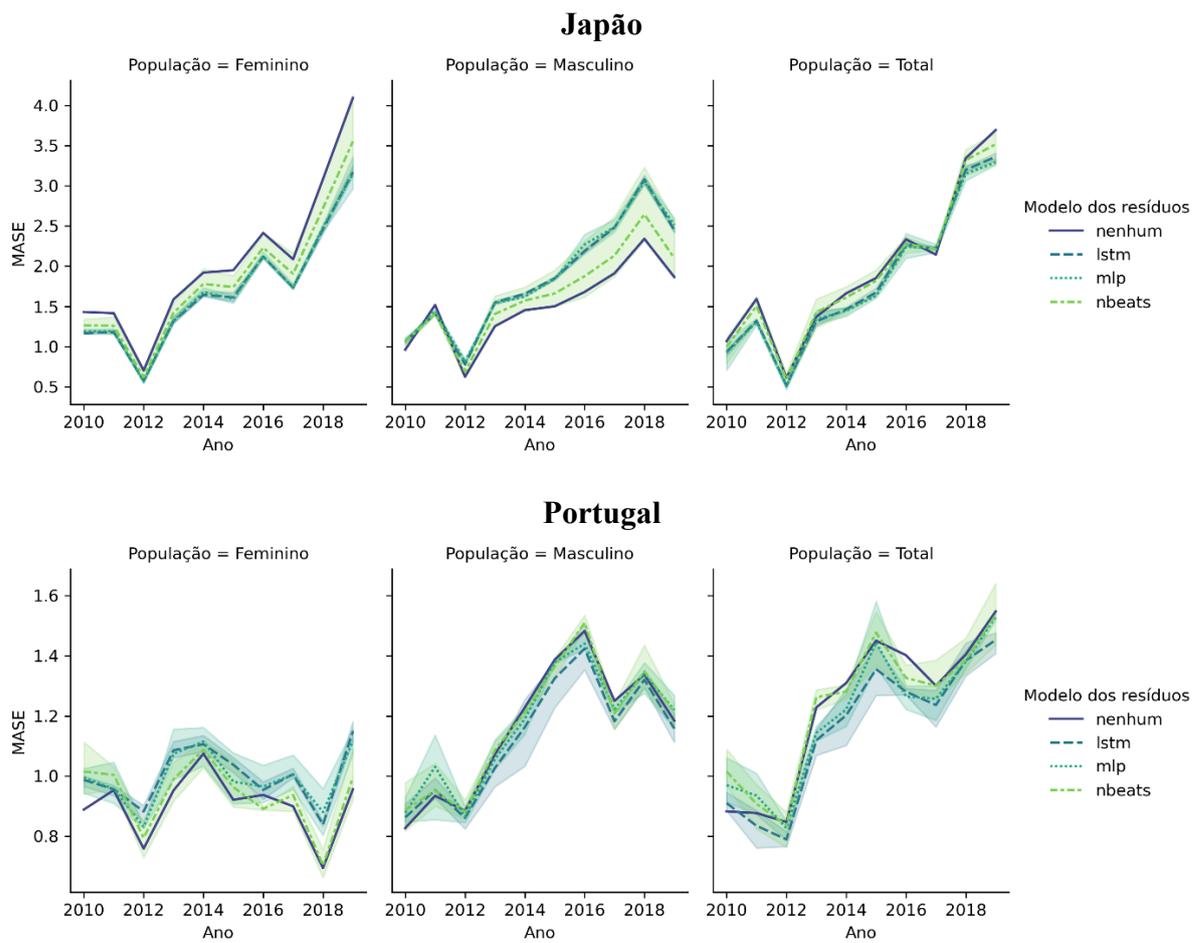


### Estados Unidos



### França





## APÊNDICE E – PREVISÕES DO SISTEMA PROPOSTO

