



UNIVERSIDADE FEDERAL DA PARAÍBA (UFPB)
CENTRO DE CIÊNCIAS SOCIAIS APLICADAS (CCSA)
DEPARTAMENTO DE FINANÇAS E CONTABILIDADE (DFC)
CURSO DE BACHARELADO EM CIÊNCIAS ATUARIAIS (CCA)

SABRINA DA SILVA LIMA

***DEEP LEARNING* NA GESTÃO DE RISCO DE LONGEVIDADE:**
O Modelo N-HITS nas Projeções de Mortalidade e Avaliação dos *q*-forwards

JOÃO PESSOA
2025

UNIVERSIDADE FEDERAL DA PARAÍBA (UFPB)
CURSO DE BACHARELADO EM CIÊNCIAS ATUARIAIS (CCA)

SABRINA DA SILVA LIMA

***DEEP LEARNING* NA GESTÃO DE RISCO DE LONGEVIDADE:**
O Modelo N-HITS nas Projeções de Mortalidade e Avaliação dos *q*-forwards

Trabalho de Conclusão de Curso
apresentado como requisito parcial à
obtenção do título de bacharel em
Ciências Atuariais pela Universidade
Federal da Paraíba.

Área de Concentração: Risco de
Longevidade.

Orientador: Prof. Dr. Filipe Coelho de
Lima Duarte.

JOÃO PESSOA
2025

Catálogo na publicação
Seção de Catalogação e Classificação

L732d Lima, Sabrina da Silva.

Deep Learning na gestão de risco de longevidade: o Modelo N-HITS nas projeções de mortalidade e avaliação dos q-forwards / Sabrina da Silva Lima. - João Pessoa, 2025.

55 f. : il.

Orientação: Filipe Coelho de Lima Duarte.
TCC (Graduação) - UFPB/CCSA.

1. Risco de longevidade. 2. q-forwards. 3. Lee-Carter. 4. N-HITS. 5. Deep learning. I. Duarte, Filipe Coelho de Lima. II. Título.

UFPB/CCSA

CDU 368(043)

SABRINA DA SILVA LIMA

DEEP LEARNING NA GESTÃO DE RISCO DE LONGEVIDADE:
O Modelo N-HITS nas Projeções de Mortalidade e Avaliação dos *q-forwards*

Trabalho de Conclusão de Curso como requisito parcial à obtenção do título de
bacharel em Ciências Atuariais pela Universidade Federal da Paraíba.

Aprovado em: 23/09/2025.

BANCA EXAMINADORA

Documento assinado digitalmente
 **FILIPÉ COELHO DE LIMA DUARTE**
Data: 06/10/2025 17:47:55-0300
Verifique em <https://validar.iti.gov.br>

Prof. Dr. Filipe Coelho de Lima Duarte (Orientador)
Universidade Federal da Paraíba (UFPB)

Documento assinado digitalmente
 **YURI MARTÍ SANTANA SANTOS**
Data: 08/10/2025 15:30:07-0300
Verifique em <https://validar.iti.gov.br>

Prof. Me. Yuri Martí Santana Santos
Universidade Federal da Paraíba (UFPB)

Documento assinado digitalmente
 **LUIZ CARLOS SANTOS JÚNIOR**
Data: 06/10/2025 17:54:16-0300
Verifique em <https://validar.iti.gov.br>

Prof. Dr. Luiz Carlos Santos Júnior
Universidade Federal da Paraíba (UFPB)

AGRADECIMENTOS

A vida é movimento, e cada mudança que enfrentamos se torna parte da nossa história. Nada é permanente, a não ser a certeza de que crescemos quando temos coragem de nos transformar. Este trabalho é fruto dessa jornada de mudanças constantes, de aprendizados e de pessoas que caminharam comigo, tornando o percurso mais leve e possível.

Agradeço primeiramente a Deus, pela presença constante e pelo amparo em todos os momentos em que pensei em desistir. Foi na fé que encontrei força para recomeçar e seguir adiante.

À minha mãe, pelo amor incondicional, apoio e dedicação que me sustentaram em cada etapa desta caminhada. Ao meu filho, Carlos Alberto, por ser fonte de inspiração diária e me mostrar que cada esforço vale a pena. E a mim mesma, pela resiliência e pela coragem de não desistir, mesmo diante das maiores dificuldades.

Aos professores Vera Lúcia, Filipe Coelho, Adriano Paixão e Luiz Carlos Júnior, pelo conhecimento transmitido, pelas orientações e pela paciência ao longo do percurso acadêmico. Vocês foram fundamentais na construção deste aprendizado.

Aos amigos que a universidade me presenteou — Marcos, Cléo, Jackson, Rodney e especialmente à Wandalúcia Neves —, minha eterna gratidão pelos momentos de parceria, estudo, risadas e apoio mútuo.

Este trabalho também é dedicado àqueles que acreditaram em mim mesmo quando eu duvidei. Cada palavra escrita aqui é reflexo de muitas mãos estendidas, olhares de incentivo e gestos de carinho.

Que este TCC não seja apenas um resultado acadêmico, mas também um símbolo da força da mudança, da beleza do progresso e da importância das pessoas que nos sustentam em cada etapa da caminhada.

“Na natureza nada se cria, nada se
perde, tudo se transforma”.

Antoine Laurent de Lavoisier

RESUMO

O risco de longevidade, associado à possibilidade de a população viver mais do que o inicialmente previsto, representa uma ameaça crescente para a solvência de seguradoras, fundos de pensão e sistemas previdenciários. Nesse contexto, derivativos indexados à mortalidade, como os *q-forwards*, surgem como instrumentos voltados à transferência e à gestão desse risco, cuja efetividade depende da precisão das projeções de mortalidade. Tradicionalmente, esses cálculos são realizados com modelos estatísticos, como o de Lee–Carter, que, embora comumente utilizados, apresentam limitações na captura de padrões não lineares. Este estudo tem como objetivo analisar o desempenho relativo do modelo *Neural Hierarchical Interpolation for Time Series Forecasting* (N-HITS), em comparação com o modelo estatístico Lee-Carter (LC) e o Perceptron Multicamadas (MLP), na previsão das taxas de mortalidade da França e da Austrália e na precificação dos *q-forwards*. Foram utilizados dados da França e da Austrália, obtidos do *Human Mortality Database* (1950–2018), divididos em períodos de treinamento (1950–2008) e teste (2009–2018). O desempenho foi avaliado pelas métricas MAPE e RMSE e pela aplicação sobre a precificação de *q-forwards*. Os resultados mostraram que o N-HITS superou consistentemente os modelos comparativos, reduzindo o MAPE em até 31% e o RMSE em até 43% na França, confirmando maior estabilidade preditiva também na Austrália, mesmo em cenários mais voláteis. Na precificação dos *q-forwards*, o modelo gerou preços mais consistentes e em torno de 5–7% superiores em idades críticas (70–80 anos), corrigindo a tendência do Lee–Carter de subprecificar o risco de longevidade. Conclui-se que o N-HITS amplia a precisão das projeções e melhora a eficácia dos derivativos indexados à mortalidade, oferecendo um avanço metodológico relevante para a gestão do risco de longevidade e contribuindo para a sustentabilidade financeira de instituições atuariais e previdenciárias.

Palavras-chave: Risco de longevidade; *q-forwards*; Lee-Carter; N-HITS; *Deep learning*.

ABSTRACT

Longevity risk, associated with the possibility that the population may live longer than initially expected, poses an increasing challenge to the solvency of insurance companies, pension funds, and social security systems. In this context, mortality-linked derivatives such as *q*-forwards emerge as instruments designed to transfer and manage this risk, whose effectiveness depends on the accuracy of mortality projections. Traditionally, these calculations are performed using statistical models such as the Lee–Carter model, which, although commonly used, show limitations in capturing nonlinear patterns. This study aims to analyze the relative performance of the Neural Hierarchical Interpolation for Time Series Forecasting (N-HITS) model, compared to the statistical Lee–Carter (LC) and the Multilayer Perceptron (MLP), in forecasting mortality rates for France and Australia and in pricing *q*-forwards. Mortality data from France and Australia were obtained from the Human Mortality Database (1950–2018), divided into training (1950–2008) and testing (2009–2018) periods. Model performance was evaluated using MAPE and RMSE metrics and through their application to *q*-forward pricing. The results indicate that N-HITS consistently outperformed the comparative models, reducing MAPE by up to 31% and RMSE by up to 43% in France, while maintaining stable predictive accuracy in Australia, even under more volatile conditions. In the pricing of *q*-forwards, the model generated more consistent values, approximately 5–7% higher at critical ages (70–80 years), correcting the Lee–Carter model's tendency to underprice longevity risk. It is concluded that N-HITS enhances the accuracy of mortality projections and improves the effectiveness of mortality-linked derivatives, offering a methodological contribution to longevity risk management and supporting the financial sustainability of actuarial and pension institutions.

Keywords: Longevity risk; *q*-forwards; Lee-Carter; N-HITS; Deep learning.

LISTA DE GRÁFICOS

Gráfico 1: MAPE/RMSE França e Austrália	31
Gráfico 2: MAPE/RMSE da França e Austrália (2)	32
Gráfico 3: Resumo visual - França	33
Gráfico 4: Resmo visual - Austrália	34

SUMÁRIO

1 INTRODUÇÃO	9
1.1 CONTEXTUALIZAÇÃO E PROBLEMA	9
1.2 OBJETIVOS	11
1.2.1 Objetivo geral	11
1.2.2 Objetivos específicos	11
1.3 JUSTIFICATIVA	11
1.4 ESTRUTURA DO TRABALHO	13
2 REVISÃO DE LITERATURA	14
2.1 LONGEVIDADE	14
2.2 RISCO DE LONGEVIDADE E GESTÃO DE RISCO DE LONGEVIDADE	15
2.3 MODELOS LEE-CARTER E <i>DEEP LEARNING</i> PARA PREVISÃO DE TAXAS DE MORTALIDADE	17
2.4 DEFINIÇÃO E APLICAÇÃO DO DERIVATIVO <i>Q-FORWARD</i>	19
2.5 DEFINIÇÃO E APLICAÇÃO DO MODELO N-HITS	22
2.6 ESTUDOS ANTERIORES	22
3 METODOLOGIA	23
3.1 ABORDAGEM GERAL	23
3.2 DADOS DA PESQUISA	23
3.2.1 Modelagem para a projeção das taxas de mortalidade	24
3.2.2. Avaliação do desempenho dos modelos preditores das taxas de mortalidade	25
3.2.3 Precificação dos <i>q</i> -forwards decorrentes dos modelos preditores das taxas de mortalidade	26
3.2.4 Impacto das previsões sobre a precificação dos <i>q</i> -forwards	27
3.2.4.1 Cenário de precificação para o <i>q</i> -forward	27
3.2.5 Ferramentas e Softwares Utilizados	28
4 ANÁLISE DOS RESULTADOS	30
4.1 ANÁLISE COMPARATIVA ENTRE OS MODELOS DE PROJEÇÃO DE MORTALIDADE	30
4.1.1 Análise tabelar dos modelos Lee-Carter, MLP e N-HITS.. Erro! Indicador não definido.	
4.2 ANÁLISE DO <i>Q-FORWARD</i>	39
5 CONCLUSÃO	42
5.1 CONSIDERAÇÕES FINAIS	42
5.2 LIMITAÇÕES DA PESQUISA	43
5.3 TRABALHOS FUTUROS	44
REFERÊNCIAS	45
APÊNDICE A	47
APÊNDICE B	48
APÊNDICE C	49
APÊNDICE D	50
APÊNDICE E	50
APÊNDICE F	53
APÊNDICE G	54

1 INTRODUÇÃO

O presente capítulo introdutório busca situar o leitor quanto ao contexto do estudo, destacando a relevância do risco de longevidade no cenário atuarial e financeiro, bem como a importância de utilizar-se de instrumentos promissores para a sua mitigação. A seguir, apresentam-se a contextualização e o problema de pesquisa, os objetivos que nortearam o trabalho, a justificativa que respalda a sua realização e, por fim, a estrutura geral adotada na investigação.

1.1 CONTEXTUALIZAÇÃO E PROBLEMA

A gestão de risco de longevidade é uma preocupação crescente em um mundo onde a expectativa de vida continua a aumentar, impactando significativamente setores como previdência, seguros e finanças (Bravo, 2021). Isso se deve à necessidade de equilibrar o impacto financeiro associado a populações que vivem mais do que o inicialmente previsto. Segundo Bravo (2021), a adoção de técnicas avançadas de projeção de mortalidade se torna essencial para a sustentabilidade desses setores, dado que erros nas estimativas de mortalidade podem resultar em graves consequências financeiras.

A utilização de derivativos indexados à mortalidade, como os *q-forwards*, é comumente reconhecida como uma estratégia eficaz para mitigar o risco de longevidade, que tem se tornado uma preocupação crescente nos setores de previdência e seguros. O *q-forward* é um derivativo estruturado em torno da previsão de taxas de mortalidade futuras, permitindo que uma das partes envolvidas se proteja contra a eventualidade de que a população viva mais tempo do que o inicialmente projetado. Isso é alcançado por meio de um contrato que estipula pagamentos em relação à diferença entre as taxas de mortalidade antecipadas e observadas (reais) (Cairns, Brown e Blake, 2006).

Esses derivativos são atraentes para os gestores de fundos de pensão e de seguradoras, pois ajudam a proteger tais entidades contra a exposição ao risco de longevidade, transferindo esse risco para outros. Como o impacto do próprio derivativo depende crucialmente da correção das expectativas de mortalidade, abordagens de modelagem sofisticadas, incluindo redes neurais e aprendizado profundo, estão sendo adotadas para uma antecipação de mortalidade mais precisa (Bravo, 2021).

No entanto, o valor dos derivativos indexados à mortalidade, como o *q-forward*, está diretamente relacionado à precisão das projeções de mortalidade, um desafio difícil que pode ser enfrentado por meio de análise de dados progressiva. Essas estimativas têm sido tradicionalmente realizadas usando os chamados modelos estatísticos e, em particular, o modelo Lee-Carter, um dos mais populares, duradouros e bem-sucedidos (Lee; Carter, 1992).

Contudo, embora eficazes em muitos casos, esses métodos têm fortes limitações (especialmente no caso de padrões não lineares e ao lidar com grandes volumes de dados complexos), uma situação que se torna mais comum nas projeções de mortalidade (Chen; Khaliq, 2021).

Neste contexto, o *deep learning* surge como uma abordagem mais promissora para a previsão de séries temporais, como as taxas de mortalidade. Diferentemente dos métodos tradicionais, que dependem diretamente da modelagem matemática por parte do analista, os algoritmos de *deep learning* são *data-driven*, ou seja, eles aprendem os padrões diretamente dos dados, o que permite a identificação de padrões não lineares. Isso elimina a necessidade de o analista formalizar regras específicas para descrever o comportamento da série temporal (Perla, 2020).

Porém, a transição de métodos estatísticos tradicionais para as técnicas de *deep learning* traz consigo desafios e oportunidades. Bravo (2021) destaca que enquanto as redes neurais recorrentes podem melhorar significativamente as previsões, elas também exigem grandes quantidades de dados e poder computacional.

Um exemplo dessa transição está no estudo de Levantesi e Nigri (2021), que aplicaram o algoritmo de *Random Forest* para melhorar as previsões feitas pelo modelo Lee-Carter. Este trabalho demonstrou que a aplicação de técnicas supervisionadas¹ de *machine learning* pode aumentar a precisão das previsões de mortalidade, o que provoca um impacto direto sobre a precificação de *q-forwards*. Ao melhorar a qualidade das previsões, os gestores de fundos de pensão e seguradoras podem usar esses derivativos de maneira mais eficaz para mitigar o risco de longevidade.

Existem diversas técnicas de *deep learning* desenvolvidas para prever séries

¹ Técnicas supervisionadas são métodos de aprendizado de máquina em que o modelo é treinado com um conjunto de dados rotulados, ou seja, em que as entradas estão associadas a saídas conhecidas. O objetivo é aprender um mapeamento ou função que, a partir das entradas, preveja corretamente as saídas para novos dados. Exemplos comuns incluem regressão linear e classificadores como redes neurais e árvores de decisão (James *et al.*, 2013).

temporais, como Redes Neurais Recorrentes (RNN) e *Long Short-Term Memory* (LSTM), que são comumente utilizadas por sua capacidade de lidar com dados sequenciais. Por outro lado, o modelo *Neural Hierarchical Interpolation for Time Series Forecasting* (N-HITS) se destaca por sua arquitetura promissora, que combina a interpolação hierárquica com blocos residuais.

Além disso, o N-HITS tem mostrado resultados competitivos em diversos conjuntos de dados, superando modelos tradicionais e outros algoritmos de *deep learning* em cenários que envolvem previsões de longo prazo (Challu *et al.*, 2022).

Diante disso, a pergunta central desta pesquisa é: Qual é o desempenho relativo do modelo N-HITS, quando comparado a Lee-Carter e Perceptron Multicamadas, na projeção de mortalidade e na precificação de *q-forwards*?

1.2 OBJETIVOS

1.2.1 Objetivo geral

O objetivo geral deste estudo é analisar o desempenho relativo do modelo N-HITS, em comparação com o Lee-Carter (LC) e o Perceptron Multicamadas (MLP), no processo de projeção das taxas de mortalidade e na consequente precificação dos *q-forwards*, como instrumento de gestão do risco de longevidade.

1.2.2 Objetivos específicos

- Projetar a mortalidade mediante o uso de modelos Lee-Carter e N-HITS;
- Avaliar a adequabilidade dos modelos utilizados;
- Analisar o impacto das diferentes projeções de mortalidade sobre a precificação do *q-forward*.

1.3 JUSTIFICATIVA

Este estudo se justifica pela necessidade de explorar e validar a aplicação dessas técnicas de *deep learning* na previsão de mortalidade, comparando a sua performance com as de outros métodos tradicionais. Além disso, estudar a precificação de derivativos *q-forwards* com base em previsões mais precisas pode

fornecer informações valiosas sobre como reduzir o risco inerente a esses instrumentos.

A integração de modelos de *deep learning* pode não apenas aprimorar a gestão de risco de longevidade, mas também contribuir para a estabilidade financeira e a eficiência dos mercados de seguro e previdência. Portanto, esta pesquisa tem relevância prática e acadêmica, oferecendo uma abordagem promissora para enfrentar um desafio contemporâneo significativo.

1.4 ESTRUTURA DO TRABALHO

O presente trabalho está estruturado em cinco capítulos principais. O primeiro capítulo apresenta a introdução, oferecendo uma contextualização detalhada do tema abordado. O segundo capítulo é dedicado ao referencial teórico, que fundamenta a pesquisa e explora os conceitos e metodologias relevantes. O terceiro capítulo descreve os procedimentos metodológicos adotados, detalhando as técnicas e as abordagens utilizadas na realização do estudo. O quarto capítulo apresenta o resultado das análises e métricas que foram aplicadas, acompanhadas da discussão de cada resultado. Por fim, o quinto capítulo apresenta a conclusão do trabalho, as limitações atuais e sugestões para trabalhos futuros.

2 REVISÃO DE LITERATURA

Este capítulo considera as ideias principais e os avanços recentes na área de gestão de risco de longevidade e previsão de mortalidade, assim como o uso de derivativos indexados à mortalidade, especificamente o *q-forward*.

2.1 LONGEVIDADE

A longevidade pode ser a característica de uma pessoa que vive muito tempo, indicando que a expectativa de vida aumentará com o tempo. A questão está diretamente ligada ao desenvolvimento da saúde, medicina, política, políticas socioeconômicas e públicas para o bem-estar (Bravo, 2021).

Na prática atuarial, longevidade refere-se à expectativa de vida dos indivíduos, isto é, ao tempo médio que uma pessoa ou grupo populacional espera viver com base em estatísticas demográficas e biométricas (Bravo, 2021).

Esse fenômeno tem se intensificado nas últimas décadas devido às melhorias dos tratamentos médicos e das políticas de saúde pública, além de uma maior conscientização sobre os hábitos de vida saudáveis (Lee; Carter, 1992). Como resultado, as instituições financeiras, as seguradoras e os fundos de pensão enfrentam desafios significativos na precificação de produtos e na gestão de passivos de longo prazo.

Dessa forma, observamos que a longevidade apresenta uma dualidade no contexto atuarial: por um lado, representa um avanço social e econômico, garantindo maior qualidade de vida; por outro, impõe grandes desafios para a sustentabilidade financeira de produtos previdenciários e seguritários (Bravo, 2021; Blake, Cairns e Dowd, 2006). A subestimação desse risco pode acarretar graves problemas financeiros, como o surgimento ou 14 quisição de déficits atuariais e o comprometimento da solvência das instituições (Chen e Khaliq, 2021). Portanto, compreender a dinâmica da longevidade e as suas implicações econômicas é essencial para desenvolver políticas e práticas capazes de mitigar os impactos negativos desse fenômeno.

2.2 RISCO DE LONGEVIDADE E GESTÃO DE RISCO DE LONGEVIDADE

O risco de longevidade ocorre quando uma população vive mais do que o inicialmente projetado, aumentando os custos associados aos benefícios pagos. Esse risco pode resultar no fracasso de instituições que precisam realizar pagamentos de pensões e benefícios de longo prazo (Blake, Cairns; Dowd, 2006). O risco de longevidade, como tal, tornou-se uma preocupação maior devido aos desenvolvimentos demográficos globais e ao aumento da expectativa de vida.

O surgimento do risco de longevidade como uma preocupação central está diretamente ligado às mudanças demográficas globais e ao aumento da expectativa de vida. Segundo dados das Nações Unidas (2022), a expectativa de vida ao nascer aumentou de forma significativa ao longo das últimas décadas, devido a avanços médicos, melhores condições de vida e uma redução das taxas de mortalidade infantil e em outras faixas etárias. Em 1950, a expectativa de vida global era de cerca de 47 anos, e em 2021 já ultrapassava os 72 anos.

Essa evolução demográfica cria desafios consideráveis para os regimes de previdência social e para os fundos de pensão, que agora enfrentam a necessidade de sustentar pagamentos de benefícios por períodos muito mais longos do que os previstos originalmente. Conforme discutido por autores como Bravo (2021) e Blake, Cairns e Dowd (2006), esse cenário tem pressionado as instituições a buscar instrumentos financeiros alternativos, como os derivativos de longevidade, para gerenciar esse risco de forma mais eficiente.

A gestão de risco de longevidade é um tema crítico, principalmente em setores como seguros, previdência privada e fundos de pensão, onde o aumento da expectativa de vida representa uma ameaça significativa à solvência a longo prazo (Bravo, 2021; Blake; Cairns; Dowd, 2006). Historicamente, ela tem sido feita, no Brasil, através da aquisição de produtos tradicionais e aquisição de fundos específicos; a tendência, no entanto, é de aquisição de derivativos de longevidade.

Assim, para cobrir esse risco, as instituições tiveram que contar com produtos tradicionais, como as anuidades de vida e os seguros de vida, que fornecem alguma forma de proteção contra o risco relacionado à longevidade. No entanto, com a crescente complexidade e o envelhecimento das populações, estamos testemunhando o surgimento de novos produtos financeiros mais sofisticados, que

permitem a transferência do risco de longevidade para os mercados de capitais (Bravo, 2021).

Além disso, no Brasil, algumas entidades de previdência complementar (EFPC) contam com fundos específicos para custear os riscos de morte, de invalidez e de sobrevivência, como é o caso da Funpresp, que oferece o Fundo de Cobertura de Benefícios Extraordinários (FCBE) para o risco de sobrevivência do assistido (Funpresp, 2019). Alternativamente, a EFPC pode contratar seguros específicos junto a sociedades seguradoras autorizadas a operar no país, visando cobrir planos de benefícios de risco, incluindo a sobrevivência, conforme previsto na Resolução CNPC Nº 47 de 01/10/2021 e na CNSP Nº 385 de 09/06/2020.

Nos últimos anos, contudo, o uso de derivativos de longevidade vem ganhando destaque como uma alternativa para mitigar o risco de longevidade. Segundo Bravo (2021), os derivativos de longevidade, como os *q-forwards*, *swaps* de longevidade e títulos de longevidade, têm se tornado ferramentas essenciais para permitir a transferência desse risco para o mercado financeiro.

Esses instrumentos oferecem uma forma de “securitizar” o risco de longevidade, permitindo que seguradoras e fundos de pensão se protejam contra flutuações inesperadas da expectativa de vida.

Os *q-forwards* são contratos futuros sobre a mortalidade projetada de uma determinada população. Eles permitem que uma das partes se proteja contra a possibilidade de que a população viva mais do que o esperado, trocando fluxos de caixa que variam de acordo com a taxa de mortalidade realizada (Bravo, 2021).

Outro derivativo proeminente são os *swaps* de longevidade, que são contratos financeiros nos quais o fluxo de dinheiro entre comprador e vendedor é baseado na expectativa de vida futura de uma população. Nesse contrato, a contraparte que aceita o risco de longevidade paga um fluxo fixo em troca de um fluxo que depende da sobrevivência de algum grupo (Coughlan *et al.*, 2007).

Essa ferramenta permite que o risco de longevidade seja transferido para aqueles investidores com apetite para assumir o risco em troca da geração de retornos financeiros (Blake *et al.*, 2013).

Por fim, os títulos de longevidade são outro tipo de derivativo que se baseia em índices de sobrevivência ou taxas de mortalidade. Esses títulos funcionam como uma forma de “seguro” contra o risco de longevidade, pagando aos investidores de acordo com as variações das taxas de mortalidade ou da longevidade de uma população. O

estudo de Blake *et al.* (2006) mostra que esses instrumentos oferecem uma forma eficiente de transferir o risco de longevidade para o mercado de capitais, reduzindo a pressão sobre os fundos de pensão.

Outro aspecto relevante na gestão do risco de longevidade é a verificação da aderência das tábuas biométricas utilizadas nas projeções atuariais. A Resolução CNPC nº 30/2018 (BRASIL, 2018), que revogou a Resolução CGPC nº 18/2006, determina que a utilização de qualquer tábua biométrica deve ser precedida de um estudo técnico que ateste a sua aderência às características da massa de participantes, o que implica a realização de testes estatísticos que avaliem se os óbitos estimados e observados, por idade, seguem a mesma distribuição probabilística.

Ademais, é possível realizar uma análise complementar, a avaliação do impacto das diferentes projeções de mortalidade sobre o fluxo de benefícios e sobre a *duration* dos planos de previdência, permitindo entender melhor como a escolha de tábuas ou dos modelos de previsão afeta a solvência de longo prazo e os custos associados à gestão dos planos.

2.3 MODELOS LEE-CARTER E *DEEP LEARNING* PARA PREVISÃO DE TAXAS DE MORTALIDADE

A previsão da mortalidade corresponde ao processo de estimar as taxas de mortalidade de uma população ao longo do tempo. A referida taxa, geralmente expressa como $m_{x,t}$, refere-se à taxa central de morte de indivíduos com idade x em um determinado ano t . Em termos práticos, essa medida indica a probabilidade de óbitos observada em uma faixa etária específica durante um período definido, sendo fundamental para a correta precificação de derivativos indexados à mortalidade.

De acordo com Bravo (2007), a taxa central de mortalidade pode ser formalmente calculada pela equação 1:

$$m_{x,t} = \frac{{}_t d_x}{{}_t E_x} \quad (1)$$

em que ${}_t d_x$ representa a quantidade de óbitos registrados para a idade x no ano t , e ${}_t E_x$ corresponde ao número de pessoas expostas ao risco nessa mesma idade e período.

Com base nas taxas de mortalidade, podem ser elaboradas as tábuas de mortalidade, que, conforme Bravo (2007), constituem ferramentas centrais da análise

demográfica. Essas tábuas sintetizam um conjunto de funções que permitem avaliar a longevidade populacional, estimar a expectativa de vida e calcular as probabilidades de sobrevivência e falecimento em diferentes idades.

No campo atuarial, representam um dos pilares para a precificação de seguros de vida e para o gerenciamento dos riscos e da sustentabilidade financeira de fundos de pensão.

Tradicionalmente, a literatura utiliza modelos estatísticos, como o modelo Lee-Carter (Lee; Carter, 1992), por sua simplicidade e robustez no ajuste de tendências de mortalidade ao longo do tempo. No entanto, o modelo Lee-Carter e outros métodos tradicionais, como o modelo Cairns-Blake-Dowd (CBD), apresentam limitações, especialmente ao lidar com grandes volumes de dados ou padrões não lineares nas taxas de mortalidade (CAIRNS; BLAKE; DOWD, 2006).

O modelo Lee-Carter (LC) é representado pela equação (2):

$$\ln(m_{x,t}) = a_x + \beta_x k_t + \varepsilon_{x,t}, \quad (2)$$

em que, $m_{x,t}$ corresponde à taxa central de mortalidade da idade x e no instante t ; o parâmetro a_x indica a média do logaritmo das taxas de mortalidade para cada idade ao longo do período analisado; β_x expressa a sensibilidade da mortalidade em relação às variações do índice temporal k_t , responsável por capturar a tendência geral da evolução da mortalidade; por fim, $\varepsilon_{x,t}$ representa o termo de erro aleatório, associado às flutuações não explicadas pelo modelo.

Como apontado por Chen e Khaliq (2021), as limitações dos métodos tradicionais na previsão de mortalidade se tornaram evidentes à medida que novos desafios surgiram com o aumento da complexidade dos dados disponíveis.

Entre os modelos de *deep learning* utilizados para prever a (taxa de) mortalidade, as redes neurais recorrentes (RNNs) e suas variantes, como as LSTMs (*Long Short-Term Memory*), têm se mostrado promissoras na previsão de séries temporais de mortalidade (Lindholm; Palmborg, 2022). As RNNs são especialmente eficazes na captura de dependências temporais de longo prazo, enquanto as LSTMs são capazes de lidar com o problema de gradientes desaparecendo, tornando-as mais adequadas para previsões de longo prazo (Perla, 2020).

Outro modelo promissor é o N-HITS (*Neural Hierarchical Interpolation for Time*

Series Forecasting), desenvolvido por Challu *et al.* (2022), que tem se destacado pela sua capacidade de capturar padrões hierárquicos nas séries temporais e fornecer previsões mais acuradas para taxas de mortalidade.

Pode ser revolucionário como os riscos de longevidade são abordados, com as técnicas de aprendizagem profunda sendo empregadas para a previsão de mortalidade, que se espera proporcionar um aumento substancial na precisão na precificação de derivativos, como os *q-forwards*.

Além desses modelos citados, outro destaque é o Perceptron Multicamadas (MLP). Este modelo é composto por neurônios artificiais organizados em três tipos de camadas: camada de entrada, uma ou três camadas ocultas e uma camada de saída. A informação viaja em uma direção (*feedforward*), e os sinais de entrada são processados sucessivamente pelas camadas ocultas, culminando na produção da resposta do modelo na saída (Haykin, 2009). Sua representação matemática, segundo Santos Júnior (2022), é dada por (3):

$$z_t = \alpha_0 + \sum_{j=1}^q \alpha_j g \left(\beta_{0j} + \sum_{i=1}^d \beta_{ij} z_{t-i} \right) + \varepsilon_t \quad (3)$$

Em que, $g(\cdot)$ representa a função de ativação não linear, enquanto α e β são os pesos ajustados pela rede durante o processo de treinamento; o parâmetro q corresponde ao número de neurônios presentes na camada oculta; e d indica o número de defasagens temporais (*lags*), ou seja, a quantidade de valores passados da série utilizados como entradas; o termo ε_t expressa o erro do modelo (Santos Júnior, 2022).

Nesse contexto, a camada de entrada utiliza observações históricas da série temporal $\{z_1, z_2, \dots, z_t\}$, sendo a ordem de defasagem d responsável por determinar o número de pontos anteriores considerados como informação. Essa estrutura fornece o contexto temporal necessário para a aprendizagem da rede (Haykin, 2009).

As informações propagam-se até a camada oculta, na qual cada neurônio aplica uma função de ativação não linear, como a sigmoide, a tangente hiperbólica ou a ReLU. A escolha dessa função é determinante para o desempenho da rede, uma vez que possibilita a introdução da não linearidade essencial para representar padrões complexos da série (Haykin, 2003).

Finalmente, a camada de saída combina os sinais recebidos da camada oculta por meio de uma transformação linear, resultando na previsão do próximo valor da série. Essa arquitetura de interconexão de múltiplas camadas dota a MLP com a capacidade de capturar os padrões temporais não lineares e representar as dinâmicas que não podem ser representadas de forma adequada por abordagens estatísticas clássicas (Zhang, 2003).

Uma vantagem fundamental do aprendizado profundo sobre as abordagens tradicionais é a sua capacidade de aprender diretamente dos dados, sem a necessidade de que o analista amostre o espaço de possíveis estruturas de modelo. Isso permite que os algoritmos identifiquem automaticamente padrões complexos e não lineares nos dados de mortalidade, proporcionando previsões mais precisas (Perla, 2020).

2.4 DEFINIÇÃO E APLICAÇÃO DO DERIVATIVO Q-FORWARD

O *q-forward* é um derivativo financeiro indexado à mortalidade, destinado à transferência e à gestão do risco de longevidade. Consiste em um contrato a termo no qual as partes acordam, para uma data futura, uma taxa de mortalidade previamente estabelecida. Conforme exposto por Coughlan *et al.* (2007), esse instrumento possibilita que seguradoras, fundos de pensão e demais instituições financeiras se protejam contra a hipótese de que os indivíduos vivam mais tempo do que o inicialmente projetado, mitigando, dessa forma, a exposição ao risco de longevidade.

O funcionamento do contrato ocorre da seguinte maneira: a parte compradora do *q-forward* compromete-se a pagar uma taxa de mortalidade fixada no início da operação e, em contrapartida, recebe a taxa de mortalidade efetivamente observada em determinado grupo populacional em um período futuro. Caso a mortalidade realizada se revele inferior à estimada, ou seja, se a população apresentar maior sobrevivência, o contrato gera um fluxo favorável à parte protegida, compensando o impacto financeiro decorrente dessa diferença.

Esse derivativo constitui um mecanismo de *hedge* de longevidade que se apoia em índices populacionais de mortalidade, não estando vinculado a carteiras específicas de segurados. Essa característica assegura maior clareza contratual e viabilidade operacional, reforçando sua utilidade como mecanismo de hedge de

longevidade. A literatura aponta que, ao transferir riscos relacionados à mortalidade e à sobrevivência, o *q-forward* fortalece a capacidade de seguradoras e fundos de pensão em manter equilíbrio financeiro no longo prazo (COUGHLAN *et al.*, 2007; BARRIEU; VERAART, 2014).

2.5 DEFINIÇÃO E APLICAÇÃO DO MODELO N-HITS

O N-HITS é um modelo de *deep learning* projetado especificamente para a previsão de séries temporais. Desenvolvido por Challu *et al.* (2022), o modelo diferencia-se por sua arquitetura hierárquica, que combina mecanismos de interpolação com blocos de redes neurais multicamadas (MLPs), permitindo capturar tanto padrões de curto quanto de longo prazo em séries de dados complexas.

A estrutura do N-HITS é composta por blocos sucessivos que processam a série em diferentes resoluções temporais. Cada bloco recebe como entrada uma janela de observações passadas, extrai representações não lineares por meio de regressões em camadas totalmente conectadas e gera coeficientes que, posteriormente, são interpolados para reconstruir a previsão em múltiplas escalas. O resultado final é obtido pela soma das previsões parciais de todos os blocos, o que possibilita ao modelo identificar tendências, ciclos e irregularidades que os métodos lineares tradicionais não capturam com a mesma eficiência.

Segundo Challu *et al.* (2022), a formulação do modelo pode ser representada de forma geral por:

$$\hat{y}_{t+1:t+H} = \sum_{l=1}^L I_l(f_l(x_{t-L+1:t}; \theta_l)) \quad (4)$$

Onde, $\hat{y}_{t+1:t+H}$ representa o vetor de previsões da série temporal para o horizonte futuro de $t + 1$ até $t + H$. O termo $x_{t-L+1:t}$ corresponde às observações passadas utilizadas como entrada do modelo, enquanto H indica o horizonte de previsão.

A função $f_l(\cdot)$ expressa a transformação realizada por cada bloco hierárquico, parametrizado por θ_l , e o operador $I_l(\cdot)$ realiza a interpolação das saídas desses blocos para o horizonte H .

2.6 ESTUDOS ANTERIORES

Seguem, no Quadro 1, os principais estudos sobre a previsão de mortalidade e a gestão do risco de longevidade, destacando seus objetivos, os métodos e os resultados.

Quadro 1: Estudos anteriores sobre previsão de mortalidade e gestão do risco de longevidade

Referência	Objetivo	Amostra	Método	Resultados
Blake, Cairns e Dowd (2006)	Avaliar a gestão de risco de longevidade com derivativos	Seguradoras e fundos de pensão	Modelos atuariais e derivativos financeiros	Concluiu que derivativos ajudam a mitigar riscos de longevidade
Lee e Carter (1992)	Modelar e prever a mortalidade nos EUA	Dados de mortalidade dos EUA	Modelo estatístico Lee-Carter	O modelo capturou tendências históricas, mas apresentou limitações em mudanças abruptas
Chen e Khaliq (2021)	Comparar redes neurais com o modelo Lee-Carter	Dados históricos de mortalidade	Redes neurais e Lee-Carter	RNNs demonstraram maior precisão em cenários com variabilidade acentuada
Challu et al. (2022)	Desenvolver o modelo N-HITS para previsão de séries temporais	Séries temporais complexas	<i>Deep learning</i> (N-HITS)	O modelo superou abordagens tradicionais em precisão e capacidade de generalização
Perla (2020)	Prever taxas de mortalidade com técnicas de <i>deep learning</i>	Taxas de mortalidade internacionais	Redes neurais profundas	Obteve maior precisão que métodos estatísticos convencionais
Lindholm e Palmborg (2022)	Explorar o uso eficiente de dados para previsão de mortalidade com LSTM	Séries temporais de mortalidade	Redes LSTM	Abordagem mais eficiente na captação de padrões dinâmicos

Fonte: Elaboração própria (2025)

Observa-se, a partir dos estudos anteriores, que embora existam contribuições relevantes da aplicação de métodos estatísticos (Lee-Carter, Cairns-Blake-Dowd) e técnicas de *machine learning* (Random Forest, LSTM) na previsão de mortalidade, há uma lacuna quanto à aplicação de arquiteturas mais recentes de *deep learning*, como o *N-HITS*. Assim, este trabalho diferencia-se por propor o uso desse modelo promissor em um contexto atuarial, avaliando o seu desempenho frente ao tradicional *modelo Lee-Carter* e analisando a sua aplicabilidade prática na precificação de *q-forwards*.

3 METODOLOGIA

A metodologia deste estudo foi elaborada para responder à questão central da pesquisa, que consiste em comparar o desempenho do modelo N-HITS em relação ao LC e MLPA partir de dados de mortalidade da Austrália e da França, foram utilizados os Modelos Lee-Carter, MPL e N-HITS para projetar a mortalidade.

Para isso, buscou-se analisar a capacidade desses modelos em projetar taxas de mortalidade da França e da Austrália e, a partir dessas projeções, avaliar seu impacto na precificação de derivativos indexados à mortalidade, com destaque para os *q-forwards*.

3.1 ABORDAGEM GERAL

A presente pesquisa caracteriza-se como aplicada, uma vez que buscou gerar conhecimento voltado à solução prática de problemas relacionados à gestão de risco de longevidade.

Quanto à abordagem, é classificada como quantitativa, pois utilizou métodos estatísticos e computacionais para análise de dados de mortalidade.

Em relação aos objetivos, enquadra-se como exploratória e descritiva, visto que investigou a aplicação de um modelo de *deep learning* ainda pouco explorado no contexto atuarial, descrevendo seus resultados comparativamente a um dos métodos tradicionais.

Quanto aos procedimentos técnicos, trata-se de uma pesquisa documental e experimental, baseada em dados secundários obtidos no *Human Mortality Database* e na implementação computacional dos modelos de previsão.

Por fim, o estudo adotou uma comparação entre métodos tradicionais de previsão de mortalidade e técnicas estado-da-arte de *deep learning*. A metodologia está dividida em três fases principais: coleta de dados, modelagem e avaliação de performance.

3.2 DADOS DA PESQUISA

Para avaliar a precisão dos derivativos indexados à mortalidade, como os *q-forwards*, foram utilizados dados de mortalidade da Austrália e da França — que foram

obtidos do *Human Mortality Database* (HMD), uma base confiável e utilizada com frequência em estudos de demografia e previsão de mortalidade —, abrangendo ambos os sexos, no período de 1950 a 2018.

A escolha desse recorte temporal segue abordagens extensivas adotadas na literatura atuarial, como os estudos de Cairns *et al.* (2006) e Shang, Booth e Hyndman (2011), onde é demonstrada a importância de séries históricas longas para capturar tendências e padrões de mortalidade ao longo do tempo, especialmente aqueles associados a melhorias nas condições de saúde, avanços médicos e políticas públicas.

Os dados foram divididos em dois conjuntos principais: a) Período de treinamento, de 1950 a 2008; b) Período de teste, de 2009 a 2018.

O recorte permite a separação clara entre treinamento e teste, essencial em modelagem preditiva. O período de 1950 a 2008 foi utilizado para o treinamento, garantindo que os modelos aprendessem com padrões históricos consolidados. O período de 2009 a 2018 foi reservado para o teste, oferecendo, assim, uma base recente, que não estava na calibração, para verificar a robustez e generalização dos modelos. Este método previne o sobreajuste e aumenta a confiabilidade das métricas de avaliação.

3.2.1 Modelos para a projeção das taxas de mortalidade

A modelagem das séries de taxas de mortalidade consistiu em aplicar três classes de modelos de previsão: tradicionais, redes neurais tradicionais e *Deep learning* (DL), conforme apresentado no capítulo anterior. Dentre os tradicionais, optou-se pelo modelo Lee-Carter (Lee; Carter, 1992) devido à sua aceitação na literatura atuarial e demográfica, além da sua capacidade de capturar tendências históricas de mortalidade de forma eficiente. Este modelo foi utilizado para prever as taxas de mortalidade com base em uma abordagem estatística clássica, servindo como parâmetro de comparação com o modelo de DL.

Dentre os modelos de redes neurais tradicionais, optou-se pelo uso do MLP, que em razão de sua utilização recorrente em problemas de classificação e previsão. Em relação aos modelos *Deep Learning*, optou-se pelo uso do N-HITS, cuja escolha se deu ao fato dele ser um modelo especializado na tarefa de previsão de séries temporais, possuindo a capacidade de modelar padrões não-lineares (Challu *et al.*, 2022).

3.2.2. Avaliação do desempenho dos modelos preditores das taxas de mortalidade

A precisão dos modelos foi avaliada com base em métricas de erro estabelecidas pela literatura, como o Erro Quadrático Médio (*Root Mean Squared Error* - RMSE) e o Erro Percentual Médio Absoluto (*Mean Absolute Percentage Error* - MAPE). Essas métricas são comumente utilizadas em estudos de previsão de séries temporais para medir a acurácia dos modelos, conforme destacado por Hyndman e Athanasopoulos (2018). O desempenho de previsão dos modelos foi testado em diferentes horizontes temporais, no período de 2009 a 2018, para verificar sua capacidade de prever a mortalidade em curto e longo prazo, tal como explorado por Bravo e Coelho (2021) em estudos de previsão de mortalidade.

O RMSE é uma medida que quantifica o erro médio das previsões do modelo em relação aos valores reais. A fórmula do RMSE é (equação 5):

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - p_i)^2} \quad (5)$$

n é o número de amostras; y_i é o valor observado para a amostra i ; p_i é o valor previsto pelo modelo para a amostra i .

O RMSE penaliza erros maiores de maneira mais significativa, o que o torna útil para detectar grandes desvios nas previsões. Quanto menor o valor do RMSE, maior a precisão do modelo de previsão (Hyndman; Koehler, 2006). Essa métrica é especialmente útil em casos onde os valores reais são variáveis em diferentes escalas, pois possibilita uma comparação direta de erro percentual, independente da magnitude dos valores previstos (Makridakis; Wheelwright ; Hyndman, 1998).

O *Mean Absolute Percentage Error* (MAPE) é uma métrica que calcula o erro percentual médio entre as previsões e os valores reais, oferecendo uma forma mais intuitiva de avaliar o desempenho preditivo. A fórmula do MAPE é (equação 6):

$$MAPE = \frac{1}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right| \times 100 \quad (6)$$

n é o número total de amostras; y_i é o valor real para a amostra i ; p_i é o valor previsto para amostra i .

O MAPE expressa o erro como uma porcentagem, facilitando a interpretação dos resultados e permitindo comparações entre diferentes modelos. Assim como no RMSE, quanto menor o valor do MAPE, maior a precisão do modelo.

3.2.3 Precificação dos q -forwards decorrentes dos modelos preditores das taxas de mortalidade

Para mensurar o desempenho dos q -forwards na mitigação do risco de longevidade, foi utilizado o modelo N-HITS em comparação aos modelos Lee-Carter e MLP. O modelo foi avaliado comparando suas previsões de taxas de mortalidade com os dados reais observados no período de 2010 a 2019, oriundos da base de dados *Human Mortality Database* (HMD). A análise visou avaliar a precisão das previsões e seu impacto na gestão do risco de longevidade por meio de métricas de erro.

A aplicação de técnicas de inteligência artificial permite que padrões complexos e não lineares sejam identificados nas séries temporais de mortalidade, aumentando a precisão do modelo e, por consequência, a eficiência dos q -forwards (Chen e Khaliq, 2021).

A precificação dos contratos q -forwards pode ser realizada a partir da abordagem do *Sharpe Ratio*, conforme proposto por Loeys *et al.* (2007). Nesse contexto, a taxa de mortalidade ajustada pelo risco é dada pela equação (7):

$$q_{s,T}^F = (1 - SR \cdot T \cdot \sigma q) q_{s,T}^E \quad (7)$$

$q_{s,T}^F$ é o preço do q -forward, ou seja, a taxa de mortalidade ajustada pelo risco (valor do contrato); $q_{s,T}^E$ é a taxa de mortalidade esperada, derivada do modelo atuarial ou da base de dados; SR é a Sharpe Ratio associado ao q -forward, utilizado como medida de prêmio de risco; T é o prazo de maturidade do contrato (tempo até o vencimento); σq é o desvio padrão histórico das variações relativas na taxa de mortalidade, definido como a equação (8):

$$\sigma q = \sqrt{VAR \left(\frac{\Delta q_{s,T}}{q_{s,T}} \right)} \quad (8)$$

Conforme destacado por Barrieu e Veraart (2016), a utilização dessa estrutura de precificação permite incorporar o risco de mortalidade de forma consistente, refletindo tanto a expectativa demográfica quanto as condições de mercado.

3.2.4 Impacto das previsões sobre a precificação dos *q-forwards*

Uma vez ajustados os modelos de previsão de mortalidade, essas projeções são utilizadas para precificar os *q-forwards*, seguindo uma abordagem baseada na metodologia de derivativos de mortalidade estabelecida por Levantesi e Nigri (2020).

No processo de precificação, as taxas de mortalidade previstas são comparadas com as taxas observadas durante o período de 2009-2018. Foi adotado um modelo de simulação de expectativa de vida com vários cenários para analisar como a precisão da previsão afeta o valor do derivativo. Isso é importante porque a incerteza em torno das previsões pode ter efeitos pronunciados no preço dos contratos, alterando os termos de transferência de risco de longevidade.

3.2.4.1 Cenário de Precificação para o *Q-forward*

Uma precificação tradicional de *q-forward* relacionada ao artigo está sujeita a um cenário de envelhecimento descrito pelos seguintes fatores para o propósito deste trabalho:

- Período de Avaliação: Previsões de doenças de 2009-2018 foram incluídas na análise, com dados de mortalidade específica por causas de 1950 a 2008 usados para calibrar o modelo.
- Populações Avaliadas: Para as duas populações de cada sexo da Austrália e França foram considerados os dados disponíveis até agora. Taxas de mortalidade estratificadas por idade foram incorporadas ao modelo para refletir os padrões específicos de idade de cada população.
- Simulação de Cenário: Os cenários que iremos modelar são os seguintes: diferenças na expectativa de vida ao nascer, simulando situações de aumento de 1 a 3 anos, de acordo com os artigos de base para esta metodologia.
- Precificação Baseada em *Mortality Forward Rates* (MFR): o valor do *q-forward* foi calculado a partir da taxa futura de mortalidade e o impacto da precisão da previsão será avaliado através das métricas de erro (RMSE e MAPE) já

descritas anteriormente.

Esse cenário simula as condições nas quais investidores e seguradoras negociam *q-forwards*, destacando como a incerteza das previsões de mortalidade influencia diretamente a precificação. Os resultados foram comparados com a metodologia apresentada por Levantesi e Nigri (2020), que demonstrou o desempenho de modelos de *Random Forest* e outros algoritmos de aprendizado de máquina na previsão de mortalidade.

Essa abordagem garante que a metodologia de precificação seja aplicada de forma ajustada ao contexto desta pesquisa, com simulações realistas e uma avaliação crítica do impacto das previsões no valor dos derivativos.

3.2.5 Ferramentas e softwares utilizados

Para implementar os modelos de previsão e conduzir a análise, todo o processo de previsão foi conduzido nos ambientes Python e R, que se mostrou adequado pela ampla disponibilidade de bibliotecas científicas e pela flexibilidade para integrar diferentes abordagens de modelagem. No ambiente Python foram implementados os 2 modelos—MLP e N-HITS — bem como os cálculos das métricas de avaliação, e o Lee-Carter foi implementado no R.

O modelo Lee–Carter foi estimado em R utilizando o pacote *Demography*. O MLP foi implementado com base em bibliotecas de *deep learning*, com camadas ocultas densas e funções de ativação não lineares. Essa estrutura buscou capturar padrões mais complexos de mortalidade, que não são contemplados por modelos lineares.

O N-HITS foi construído a partir do pacote *neuralforecast*, desenvolvido sobre o PyTorch. Trata-se de uma arquitetura especialmente projetada para previsões em horizontes mais longos, característica essencial no contexto atuarial.

Após a fase de treinamento e validação, os três modelos foram utilizados para gerar previsões anuais das taxas de mortalidade no período de 2009 a 2018, em diferentes idades selecionadas. Essas previsões foram então comparadas às observadas por meio de duas métricas principais: o MAPE e o RMSE.

Por fim, todos os resultados — previsões e métricas — foram exportados em arquivos “.csv”, servindo de base para as análises estatísticas, construção dos gráficos e elaboração dos quadros comparativos apresentados no capítulo de

resultados.

4 ANÁLISE DOS RESULTADOS

Este capítulo tem como objetivo expor, de forma sistemática, os desempenhos alcançados pelos diferentes métodos testados, destacando tanto as potencialidades quanto as suas limitações.

O foco principal deste estudo consiste em examinar comparativamente a acurácia dos modelos N-HITS, Lee-Carter (LC) e Perceptron Multicamadas (MLP) na previsão das taxas de mortalidade da França e da Austrália, buscando compreender como o desempenho relativo dessas abordagens influencia a precificação dos *q-forwards* e, por consequência, a gestão do risco de longevidade.

Para isso, foram utilizadas as métricas MAPE e RMSE, que permitem quantificar a qualidade das previsões sob diferentes perspectivas: o MAPE em termos relativos e o RMSE como medida de robustez contra grandes desvios.

4.1 ANÁLISE COMPARATIVA ENTRE OS MODELOS DE PROJEÇÃO DAS TAXAS DE MORTALIDADE

Inicialmente, foram elaborados *boxplots* para ilustrar a distribuição dos erros de previsão dos modelos avaliados. Os Gráficos 1 e 2 contemplam as métricas MAPE e RMSE para os países França e Austrália, permitindo observar não apenas as medianas dos erros, mas também a dispersão, a variabilidade e a ocorrência de valores extremos. Essa abordagem fornece uma visão mais clara sobre a estabilidade e a consistência de cada método.

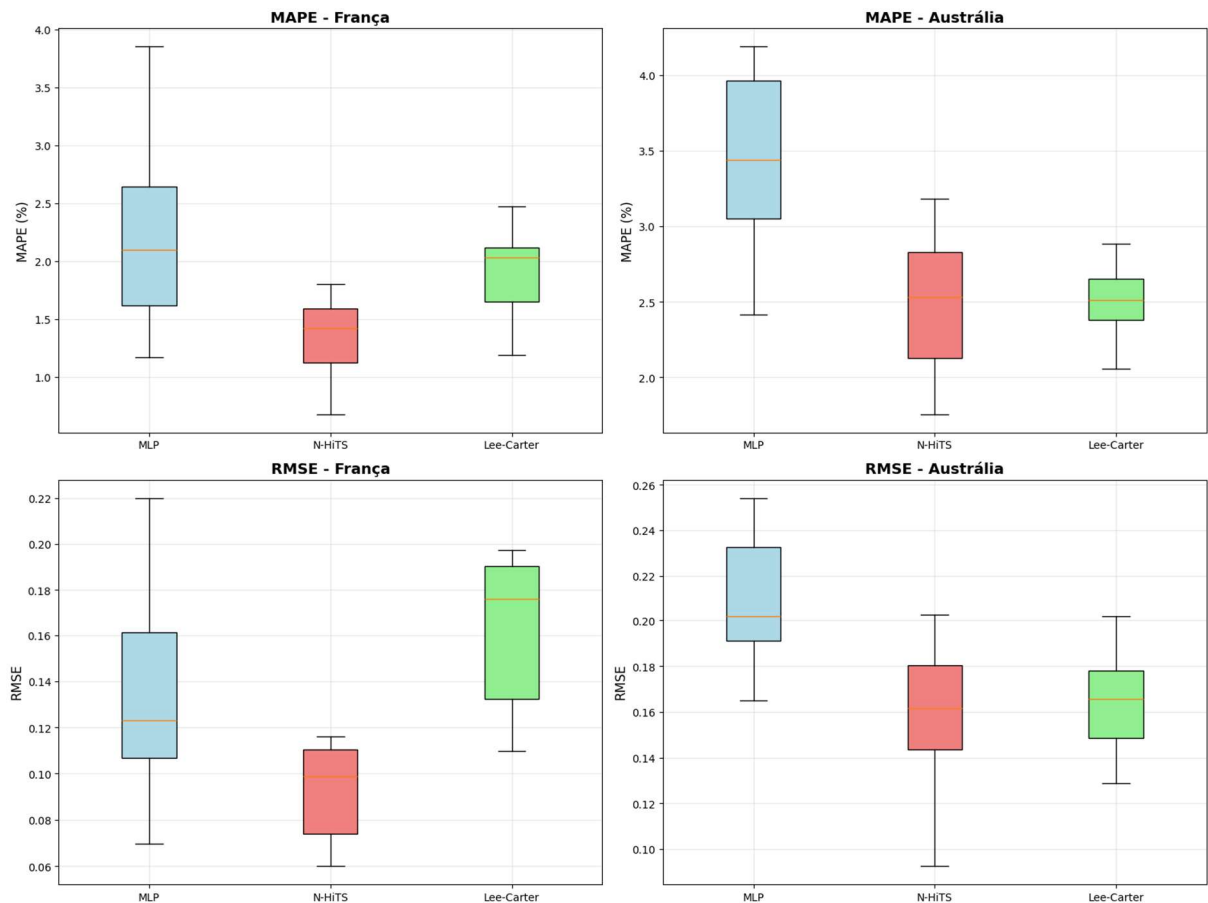
Referente ao MAPE e RMSE, observa-se nos Gráficos 1 e 2 que, na França, o modelo N-HITS obteve a menor mediana de erro e apresentou uma distribuição bastante compacta, sinalizando previsões mais estáveis e com menor variabilidade. O MLP, por sua vez, exibiu maior dispersão dos valores, evidenciando flutuações mais acentuadas em seu desempenho, enquanto o Lee-Carter apresentou resultados intermediários, mas ainda superiores aos obtidos pelo N-HITS.

Na Austrália, o padrão observado foi semelhante, embora as diferenças entre os modelos tenham se tornado mais evidentes. O MLP apresentou os maiores valores medianos e uma dispersão considerável, destacando a sua instabilidade no ajuste das séries temporais australianas. O N-HITS voltou a se destacar com valores reduzidos e distribuição mais concentrada, confirmando a sua robustez. O Lee-Carter, embora

tenha apresentado valores de MAPE mais próximos ao N-HITS, mostrou-se menos estável, reforçando as limitações inerentes aos modelos tradicionais em contextos mais complexos.

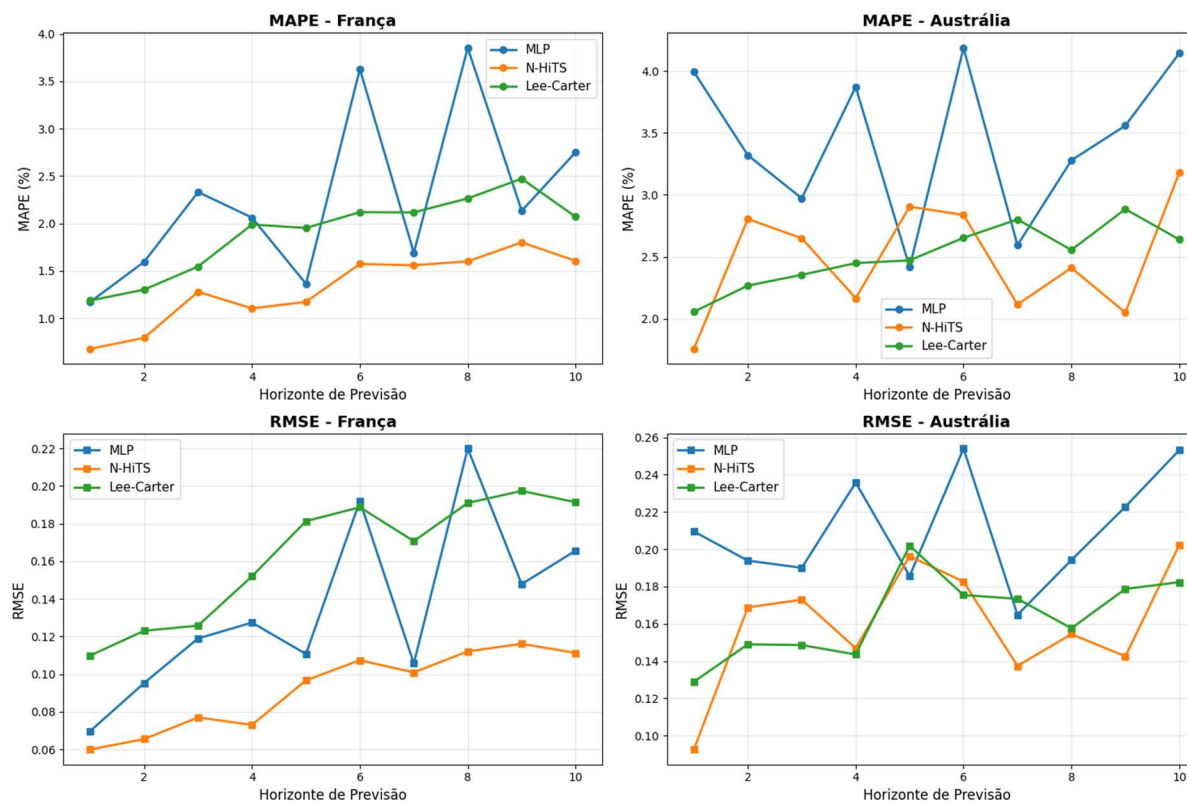
Nos gráficos correspondentes ao RMSE, verifica-se novamente a superioridade do N-HITS. Na França, o modelo apresentou os menores valores de erro e a menor dispersão, confirmando a sua consistência. O MLP apresentou desempenho intermediário, mas com maior amplitude interquartílica, enquanto o Lee-Carter apresentou os piores resultados em termos de RMSE, refletindo a sua limitação em capturar padrões mais sutis nas séries de mortalidade.

Gráfico 1: Métricas de erro (MAPE e RMSE) dos modelos analisados (MLP, N-HITS e Lee-Carter) para França e Austrália.



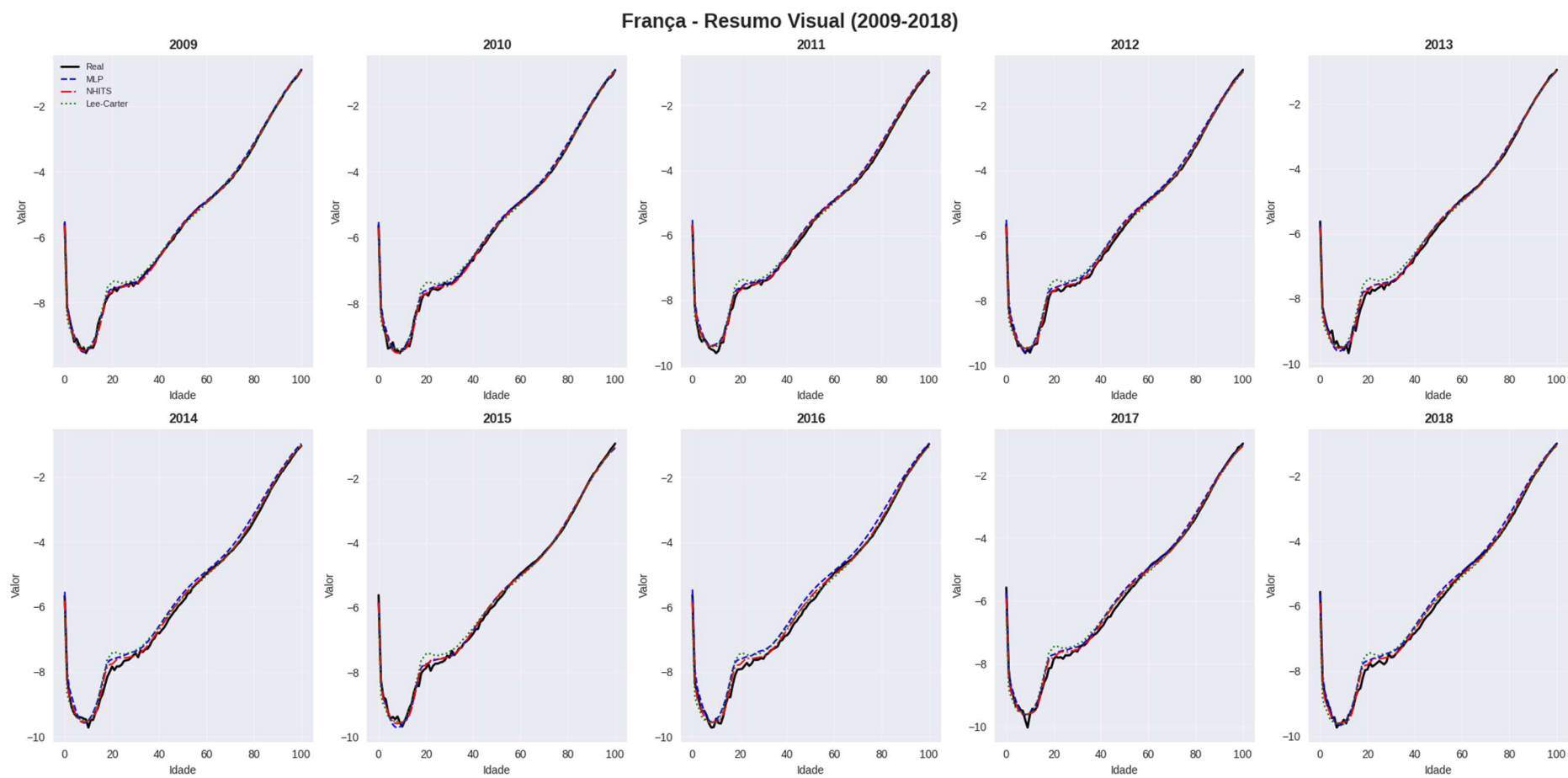
Fonte: Elaboração própria, a partir dos dados coletados (2025).

Gráfico 2: Métricas de erro (MAPE e RMSE) por horizonte de tempo dos modelos analisados (MLP, N-HITS e Lee-Carter) para França e Austrália.



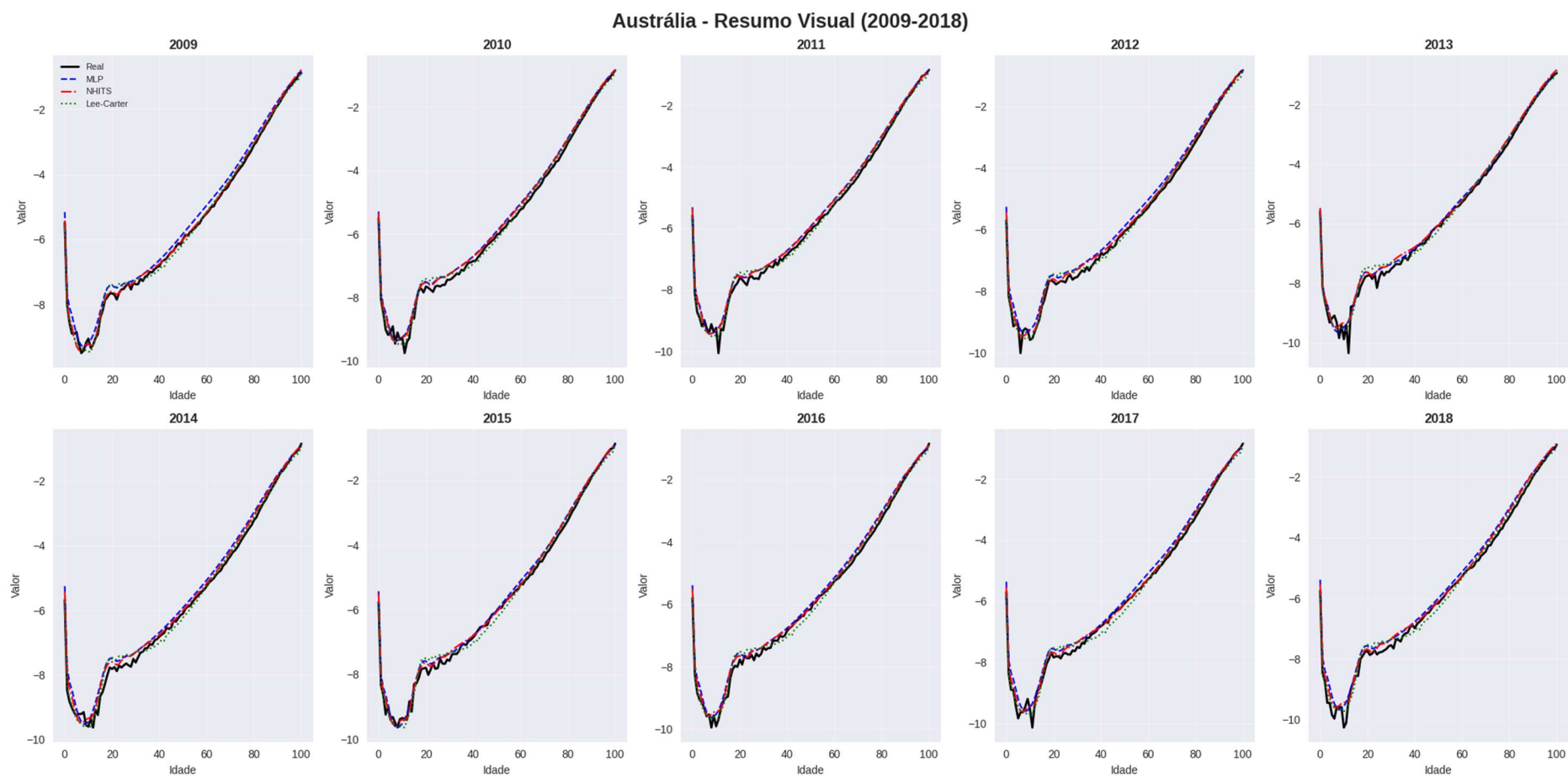
Fonte: Elaboração própria, a partir dos dados coletados (2025).

Gráfico 3: Curva de mortalidade e previsões para as idades de 0 a 100 dos modelos para conjunto de teste da França.



Fonte: Elaboração própria, a partir dos dados coletados (2025).

Gráfico 4: Curva de mortalidade e previsões para as idades de 0 a 100 dos modelos para conjunto de teste da Austrália.



Fonte: Elaboração própria, a partir dos dados coletados (2025).

Os gráficos 3 e 4 reúnem as séries temporais utilizadas neste estudo, referentes às previsões de mortalidade para a França e para a Austrália no período de 2009 a 2018, permitindo, assim, visualizar a aderência dos modelos N-HITS, MLP e Lee-Carter em relação aos valores reais observados, evidenciando diferenças de desempenho entre os modelos.

O N-HITS mostrou-se consistentemente superior, tanto em termos de média quanto de dispersão dos erros. Esse desempenho decorre de sua arquitetura hierárquica, que permite decompor a série em múltiplas escalas e capturar padrões temporais complexos (Challu *et al.*, 2022). O MLP, embora apresente flexibilidade para modelar relações não lineares, não incorpora diretamente a estrutura temporal, o que explica sua instabilidade (Haykin, 2009; James *et al.*, 2013). O Lee-Carter continua sendo um modelo válido, especialmente por sua simplicidade e interpretabilidade (Lee; Carter, 1992), contudo, carece de sofisticação para competir com técnicas modernas em contextos de maior variabilidade (Chen; Khaliq, 2021; Perla, 2020).

O que se esperava encontrar é que os erros de mortalidade relacionados a causas duplas são, pela natureza do caso, inevitáveis, mas que, assim como seria o caso em qualquer estudo empírico, erros desse tipo simplesmente se multiplicam naturalmente por duas razões – não por causa da vontade de alguém, ou porque podem ser corrigidos.

A primeira razão é que o horizonte de projeção é mais longo, e a segunda, porque mais incerteza está conectada com desenvolvimentos futuros. Esse tipo de comportamento é característico dos dados demográficos e, se for o caso, serve para ilustrar a necessidade de modelos que possam lidar com esse nível de complexidade (HYNDMAN; KOEHLER, 2006).

É, portanto, nesse ponto elevado que se destaca a superioridade do N-HITS, na sua maior estabilidade e realismo. Uma terceira razão pela qual tem sido bem-sucedido é que essa regra simples segue um caminho mais próximo do curso da evolução da mortalidade.

4.1.1 Análise tabelar dos modelos Lee-Carter, MLP e N-HITS

A Tabela 1 sintetiza os resultados das métricas de desempenho (MAPE e RMSE) para os três modelos avaliados – Lee-Carter, MLP e N-HITS – considerando

os dois países em estudo, França e Austrália.

Tabela 1: Estatísticas descritivas referentes ao MAPE e ao RMSE, por modelo de projeção das taxas de mortalidade (Lee-Carter, MLP e N-HITS).

Mo- delo	País	MAPE (%)				RMSE			
		Média	Mínimo	Máximo	Desvio	Média	Mínimo	Máximo	Desvio
Lee-Carter	Austrália	2,5131	2,0561	2,8849	0,2493	0,1639	0,1289	0,2019	0,0220
	França	1,9014	1,1865	2,4718	0,4198	0,1631	0,1099	0,1974	0,0329
MLP	Austrália	3,4336	2,4152	4,1870	0,6311	0,2104	0,1648	0,2540	0,0301
	França	2,2576	1,7107	3,8514	0,3199	0,1353	0,0698	0,2199	0,0461
N-HITS	Austrália	2,4873	1,7567	3,1800	0,4568	0,1597	0,0926	0,2023	0,0324
	França	1,3161	0,6772	1,7999	0,3760	0,0920	0,0600	0,1161	0,0211

Fonte: Elaboração própria, a partir dos dados coletados (2025).

As estatísticas descritivas (média, mínimo, máximo e desvio padrão) dispostas na tabela 1 foram baseados nos erros de previsão em todos os pontos de análise, ou seja, para cada idade prevista em cada ano do período de teste. O valor mínimo é o desvio mínimo entre valores reais e previstos, e o máximo é o máximo. O segundo painel à esquerda é o desvio padrão do termo de erro, que dá uma ideia da quantidade de variação da perturbação em torno da média: valores baixos sugerem melhores previsões, enquanto valores maiores indicam que perturbações um pouco maiores estão presentes.

Nesse sentido, a Tabela 1 apresenta uma perspectiva geral, em vez de uma específica do modelo, e descobre quão bem esses modelos funcionam na prática.

Na França, os resultados mostram claramente que o N-HITS atinge o menor MAPE médio (1,3161%) e o menor RMSE médio (0,0920). Além de apresentar a menor dispersão entre os valores mínimo e máximo. Esses resultados sugerem que o modelo foi capaz de reduzir o erro percentual, bem como ter boa confiabilidade em diferentes horizontes de previsão.

Esses resultados correspondem aos de outros estudos, trabalhando com séries temporais multiescala, nos quais o modelo N-HITS demonstrou ganhos em precisão e estabilidade em horizontes mais longos, capturando estruturas hierárquicas e padrões de iluminação complexos (Challu *et al.*, 2022), além de evidências empíricas de que os modelos de aprendizagem profunda superam as abordagens estatísticas clássicas na precisão de previsão de mortalidade (Perla, 2020).

Os resultados da MLP não foram tão sólidos, com MAPE médio de 2,2576% e RMSE de 0,1353. O modelo também apresentou uma maior dispersão, especialmente em RMSE (0,0461), indicando que algumas previsões eram muito mais instáveis que

outras. A dispersão observada é consistente com descobertas anteriores que enfatizaram a necessidade de adotar uma abordagem mais holística atravessando períodos. Como, por exemplo, empregando arquiteturas recorrentes, para calibrar erros em dados demográficos dinâmicos (Lindholm; Palmborg, 2022).

O Lee-Carter, por sua vez, apesar de ter um MAPE médio inferior (1,9014%) em relação à MLP, parecia ter um RMSE mais alto (0,1631), o que confirma que ele é limitado para capturar padrões mais complexos. No entanto, no geral, seu desempenho na França foi bastante adequado para um modelo tão simples.

Na Austrália, as distinções entre os modelos tornaram-se mais claras. Mais uma vez, o N-HITS também alcançou as melhores pontuações para ambas as medidas (MAPE e RMSE) e teve o menor mínimo da métrica (1,7567% de MAPE e 0,0926 de RMSE). Apesar de os resultados de Lee-Carter apresentados ali serem relativamente semelhantes (MAPE médio de 2,5131% e RMSE de 0,1639), nota-se a menor variabilidade nos resultados de N-HITS, ratificado como uma expressão de sua maior robustez.

Ao comparar diretamente o desempenho dos modelos, observa-se que na Austrália, a diferença entre o N-HITS e o Lee-Carter é marginal, com o N-HITS apresentando redução de 0,0258 p.p. no MAPE e 0,0042 no RMSE. Na França, o ganho de precisão é significativo: o N-HITS reduziu o erro percentual médio em 0,5853 p.p. e o RMSE em 0,0711, confirmando a sua superioridade na captura da dinâmica demográfica local.

Os resultados do modelo N-HITS, quando mais variabilidade demográfica é considerada, fornecem evidências mais fortes de que modelos neurais podem responder mais a não-linearidades e mudanças de regimes do que o modelo LC (Chen *et al.*; Khaliq, 2021). Apesar de suas limitações, o modelo clássico ainda possui propriedades competitivas em alguns critérios, como a sua capacidade de replicar a tendência de longo prazo (Lee; Carter, 1992).

A MLP, no entanto, exibiu um desempenho notavelmente inferior, com um MAPE médio de 3,4336% e valor de RMSE de 0,2104, e alta dispersão para ambos. Esses resultados indicam que à medida que a demografia se torna mais volátil, como no exemplo australiano, a MLP torna-se menos eficiente e estável, verificando que a MLP é limitada quando comparada a estruturas mais estendidas. O erro tende a aumentar com o aumento no horizonte de previsão.

Essa incerteza tende a se acumular ao longo do tempo, particularmente em

ambientes mais voláteis (Challu *et al.*, 2022). Essa estrutura, que vemos aqui, enfatiza a necessidade de arquiteturas que governem a proliferação de erros em múltiplas escalas de tempo.

Em suma, a evidência empírica deste estudo converge com os achados prévios: (i) modelos clássicos, como Lee-Carter, permanecem referências interpretáveis e úteis como linha de base (Lee; Carter, 1992); (ii) em cenários com maior variabilidade, redes neurais tendem a oferecer ganhos de acurácia e robustez (Chen; Khaliq, 2021; Perla, 2020; Challu *et al.*, 2022); e (iii) melhorias na qualidade preditiva têm implicações diretas para a precificação e a efetividade de instrumentos de transferência de risco de longevidade, como os derivativos indexados à mortalidade (Blake; Cairns; Dowd, 2006).

A Tabela 2 evidencia que o modelo N-HITS apresentou o melhor desempenho em todas as métricas, tanto na França quanto na Austrália.

Tabela 2: Melhor Modelo por Métrica.

País	Melhor Modelo	MAPE (%)	Melhor Modelo	RMSE
França	N-HITS	1,3161	N-HITS	0,0920
Austrália	N-HITS	2,4873	N-HITS	0,1597

Fonte: Elaboração própria, a partir dos dados coletados (2025).

Na França, N-HITS alcançou o menor erro percentual (MAPE = 1,3161%) e também a menor raiz do erro quadrático médio (RMSE = 0,0920). Isso demonstra não apenas elevada precisão, mas também grande consistência nas previsões, tornando-o o modelo mais confiável para a base francesa.

Este resultado é consistente com a descoberta de que estruturas hierárquicas de aprendizagem profunda, como o N-HITS, são capazes de aprender características temporais complexas de forma mais precisa e robusta em comparação com modelos estatísticos clássicos (Challu *et al.*, 2022; Perla, 2020).

Também verifica descobertas anteriores de que a redução tanto do erro percentual quanto da variabilidade é essencial em aplicações atuariais de longo prazo (Chen; Khaliq, 2021).

Na Austrália, embora os valores absolutos de erro tenham sido maiores (MAPE = 2,4873% e RMSE = 0,1597), o N-HITS manteve a liderança frente aos demais modelos. Esse resultado é relevante porque a série australiana apresenta maior volatilidade e heterogeneidade, e ainda assim o N-HITS conseguiu superar tanto o modelo clássico Lee-Carter quanto o MLP.

O fato de o N-HITS ter até mesmo uma vantagem neste cenário mais questionável é corroborado por estudos que apontam um sobreajuste do modelo Lee-Carter quando se trata de mudanças repentinas na mortalidade (Lee; Carter, 1992) e sustenta a melhor robustez das abordagens baseadas em redes neurais quando enfrentam heterogeneidade demográfica (Lindholm; Palmborg, 2022; Chen; Khaliq, 2021).

Além disso, os erros mais elevados na Austrália são consistentes com a maior incerteza das previsões em horizontes de previsão mais longos, que foi observada em estudos semelhantes (Levantesi; Nigri, 2020).

No geral, esses resultados são consistentes com a literatura que, por um lado, reconheceu o uso contínuo de modelos estatísticos clássicos como *benchmarks* interpretáveis (Lee; Carter, 1992) e, por outro lado, destacou o potencial dos métodos de *deep learning* para gerar previsões mais estáveis e robustas, com aplicações diretas na gestão do risco de longevidade e na precificação de derivativos indexados à mortalidade (Blake; Cairns; Dowd, 2006).

4.2 ANÁLISE DO Q-FORWARD

Os resultados da trajetória *q-forward* de 2009 a 2018, para Austrália e França, estão alinhados com as descobertas da literatura (LEVANTESI; NIGRI, 2020). As curvas de mortalidade estão diminuindo de maneira moderada, indicando um incremento da expectativa de vida, verificando-se de forma consistente até os 75–80 anos. Ainda, a volatilidade das taxas de mortalidade por idade ainda está sob controle, acelerando apenas nas idades extremas.

Isso apoia as descobertas empíricas observadas por Levantesi e Nigri (2020), de que o risco demográfico tende a aumentar em idades mais avançadas e, portanto, contratos projetados em blocos de idade e gênero são necessários para reduzir o risco de base.

Tabela 3: Preço do *q-forward* (em milhares) entre os modelos analisados.

Idades	Modelos	França	Austrália
60	LC	6,1381	4,8191
	MLP	7,1704	5,9138
	N-HITS	6,5617	5,1215
70	LC	12,3743	13,2969
	MLP	14,1965	15,1525
	N-HITS	12,7156	14,0737

80	LC	34,6917	41,1487
	MLP	40,5485	46,8741
	N-HITS	36,4847	43,9689

Fonte: Elaboração própria, a partir dos dados coletados (2025).

Na Austrália, os resultados da Tabela 1 indicaram que os modelos Lee–Carter e N-HITS apresentam desempenhos bastante semelhantes, com uma ligeira vantagem para o N-HITS. Esse padrão se refletiu na Tabela 3, onde os preços dos *q-forwards* obtidos pelo N-HITS foram mais elevados que os de Lee-Carter em todas as idades analisadas (60, 70 e 80 anos). Isso ocorre porque o modelo captou um nível maior de risco de longevidade, resultando em contratos de *hedge* mais caros. Do ponto de vista atuarial, tal comportamento implica que seguradoras pagariam valores mais altos para se proteger contra a incerteza, mas, em contrapartida, com base em projeções mais realistas.

Cabe observar que a diferença entre os preços derivados do LC e do N-HITS, embora relativamente pequena nas idades iniciais (60 anos), tende a se ampliar de maneira significativa nas idades mais avançadas (70 e 80 anos), tanto para a França quanto para a Austrália. Isso demonstra que o impacto do modelo na precificação é crescente conforme aumenta a incerteza demográfica.

Na França, por outro lado, a Tabela 2 revelou uma clara superioridade do N-HITS em termos de acurácia preditiva, tanto no MAPE quanto no RMSE. Essa melhora na previsão também se refletiu na Tabela 3, onde os preços dos *q-forwards* gerados pelo N-HITS ficaram consistentemente acima dos obtidos pelo Lee–Carter em todas as idades avaliadas. Nesse caso, o impacto é positivo para a gestão de risco de longevidade, pois o N-HITS corrige a tendência de subprecificação do Lee–Carter, proporcionando maior segurança e mitigando o risco de subestimativas que poderiam comprometer a solvência das instituições financeiras no longo prazo.

Nas comparações entre os modelos, o LC apresentou resultados mais suaves e estáveis, atuando como o *benchmark* clássico. Em contraste, o N-HITS consegue captar o padrão dinâmico não linear e as mudanças estruturais que caracterizam as séries de mortalidade. Mostrando, portanto, uma sensibilidade maior às recentes evoluções demográficas, embora com mais variabilidade em idades muito avançadas, podendo ser observada nos Apêndices A, B e C do presente trabalho.

O MLP exibe resultados moderados com um comportamento mais estocástico em alguns intervalos de idade. Outra descoberta relevante, reside no fator de ajuste

$(1 - SR \cdot T \cdot \sigma q)$, oscilando apenas em idades mais avançadas. Isto é consistente com o prêmio de risco implícito no contrato sendo parcimonioso e constante; consistente com Coughlan *et al.* (2007), que afirmam que os *q-forwards* são negociados com desconto em relação à taxa de mortalidade esperada apenas para recompensar um investidor que adquire um risco de longevidade.

Os resultados confirmam que o N-HITS superou o Lee–Carter no geral, reduzindo o MAPE em até 31% e o RMSE em 43%, na França, e produzindo preços de *q-forwards* mais consistentes. Em idades de 70 e 80 anos, o modelo N-HITS corrigiu a subprecificação do Lee-Carter, elevando os preços em cerca de 5–7% e oferecendo maior proteção atuarial.

Aos 60 anos, o N-HITS apresentou melhor equilíbrio em precificação, reduzindo erros de previsão nas métricas MAPE e RMSE em até 38 % em relação ao MLP e em torno de 3% em relação ao Lee–Carter. Além de gerar preços de *q-forward* 6,3% maiores que os do modelo clássico. Esses resultados são consistentes com o que foi observado por Challu *et al.* (2022), que destacam a capacidade do N-HITS de capturar padrões em múltiplas escalas temporais e controlar a propagação do erro em horizontes de previsão mais longos.

Portanto, embora seja a alternativa mais robusta, sua aplicação prática deve considerar a segmentação de portfólios por idade e gênero, conforme sugerem Coughlan *et al.* (2007).

5 CONCLUSÃO

Este capítulo apresenta a conclusão deste, respondendo à pergunta de pesquisa sobre o desempenho relativo do modelo N-HITS em comparação ao LC e MLP na previsão das taxas de mortalidade. Os resultados demonstraram que o N-HITS superou os demais modelos, oferecendo maior precisão na projeção das taxas de mortalidade e maior efetividade na precificação dos *q-forwards*, atendendo ao objetivo central da pesquisa.

Em seguida, são destacadas as considerações finais, principais limitações metodológicas e, por fim, são indicadas direções para trabalhos futuros, com vistas ao aprimoramento das aplicações em gestão do risco de longevidade.

5.1 CONSIDERAÇÕES FINAIS

A modelagem das taxas de mortalidade constitui um desafio metodológico, em virtude da natureza não linear dessas séries temporais, da presença de padrões dinâmicos complexos e das diferenças significativas entre grupos etários, além da limitação imposta pelo número relativamente reduzido de observações disponíveis em bases históricas (Bravo, 2021; Chen; Khaliq, 2021).

Nesse contexto, o desenvolvimento de modelos de previsão capazes de capturar tanto as estruturas lineares quanto os padrões não lineares e de lidar com os horizontes de longo prazo, representa uma necessidade científica e prática na área atuarial.

Este trabalho aplicou o modelo N-HITS como alternativa às técnicas tradicionais de previsão de mortalidade. Sua arquitetura hierárquica decompõe a série temporal em múltiplas escalas, captando padrões complexos e aumentando a acurácia preditiva. Para comparação, foram utilizados o modelo Lee–Carter, pela simplicidade e relevância histórica, e o Perceptron Multicamadas (MLP), como representante de redes neurais tradicionais.

O processo metodológico consistiu em duas etapas: (i) treinamento com dados de mortalidade da França e da Austrália (1950–2008) para ajuste dos modelos; e (ii) teste (2009–2018) para avaliar a capacidade preditiva. O desempenho foi medido por MAPE e RMSE, além de análises gráficas e por horizonte de previsão.

Os resultados experimentais mostraram que o N-HITS superou

consistentemente o Lee–Carter e o MLP, com menores erros e maior estabilidade, sobretudo na França, e robustez mesmo em séries mais voláteis, como na Austrália.

Essa superioridade evidencia o potencial do modelo em gerar previsões de mortalidade mais precisas e confiáveis. Na precificação dos *q-forwards*, o N-HITS produziu valores mais consistentes, com volatilidade controlada e fatores de ajuste estáveis, indicando que melhorias na acurácia preditiva se traduzem em instrumentos mais eficazes para a gestão do risco de longevidade.

Além disso, os resultados são consistentes com Levantesi & Nigri (2020), que interpreta o *q-forward* como um contrato de liquidação direta com base na qualidade das projeções de mortalidade. Ao provar que o *Deep Learning* reduz vieses e aumenta a estabilidade atuarial dos modelos associados, os resultados deste trabalho reforçam a consideração de que melhorias nos métodos podem tornar esses instrumentos de *hedge* mais eficientes. Essas conclusões reforçam o papel do N-HITS como uma ferramenta promissora no campo da previsão de mortalidade e destacam a sua relevância prática para a gestão atuarial e financeira de longo prazo.

Sendo assim, as principais contribuições desta pesquisa são sintetizadas da seguinte forma: (i) Aplicação inédita do N-HITS para previsão de taxas de mortalidade em contextos atuariais, estabelecendo comparações diretas com modelos clássicos (Lee-Carter) e redes neurais tradicionais (MLP); (ii) Evidência empírica de que o N-HITS supera o Lee–Carter e o MLP em termos de acurácia e estabilidade das previsões, apresentando reduções de até 31% no MAPE e 43% no RMSE em relação ao Lee–Carter, com destaque para a França, ainda que na Austrália os ganhos tenham sido mais modestos; (iii) Demonstração de que melhorias preditivas impactam diretamente no desempenho dos derivativos de longevidade, em especial os *q-forwards*, que apresentaram preços 5–7% mais altos em idades críticas (70–80 anos), corrigindo a tendência do Lee–Carter de subprecificar o risco de longevidade; (iv) Contribuição prática para o campo atuarial ao indicar que modelos baseados em *Deep Learning* podem ampliar a segurança e a sustentabilidade financeira de seguradoras, fundos de pensão e demais instituições expostas ao risco de longevidade.

5.2 LIMITAÇÕES DA PESQUISA

Apesar dos resultados alcançados, esta pesquisa apresenta algumas limitações que devem ser reconhecidas. Em primeiro lugar, os experimentos foram

conduzidos com base em dados de apenas dois países — França e Austrália —, o que restringe a generalização dos achados para outros contextos demográficos e socioeconômicos. Além disso, ainda que essas bases de dados sejam robustas e provenientes de fonte confiável, possíveis inconsistências de registro ou variações da qualidade dos dados podem influenciar os resultados obtidos.

Outra limitação refere-se à escolha dos modelos comparativos. Embora o estudo tenha incluído o Lee-Carter, referência clássica na literatura, e o Perceptron Multicamadas (MLP), representativo de arquiteturas neurais mais simples, outros modelos modernos de *Deep Learning*, como LSTM, GRU e *Transformers*, não foram considerados. Isso restringe o escopo da análise comparativa e abre espaço para investigações mais abrangentes.

Por fim, a análise concentrou-se na aplicação dos modelos para a previsão das taxas de mortalidade e o seu efeito sobre a precificação dos *q-forwards*. Outros derivativos de longevidade, como *longevity bonds* e *longevity swaps*, não foram explorados, o que limita a discussão sobre a aplicabilidade prática em instrumentos financeiros alternativos.

5.3 TRABALHOS FUTUROS

Diante das limitações apresentadas, sugerem-se alguns caminhos para pesquisas futuras. Uma primeira possibilidade é a ampliação do escopo geográfico, aplicando-se o modelo N-HITS a diferentes países e regiões, o que permitiria avaliar a sua robustez em contextos variados de mortalidade e longevidade.

Outro estudo seria a realização de uma comparação com outras arquiteturas de *Deep Learning*, incluindo modelos recorrentes (LSTM e GRU), redes baseadas em atenção (*Transformers*) e modelos híbridos que combinem métodos estatísticos e redes neurais. Essa abordagem poderia fornecer uma visão mais ampla sobre a posição relativa do N-HITS no estado da arte.

Adicionalmente, explorar a precificação de derivativos em carteiras reais de seguradoras e fundos de pensão, testando a aplicabilidade dos resultados em situações práticas. Além dos *q-forwards*, recomenda-se avaliar outros instrumentos financeiros ligados à longevidade, como *swaps* e *bonds*, ampliando o entendimento sobre o papel dos modelos preditivos no mercado financeiro.

REFERÊNCIAS

BARRIEU, P.; VERAART, A. E. D. Pricing longevity derivatives. *Insurance: Mathematics and Economics*, v. 55, p. 194-204, 2014.

BLAKE, D.; CAIRNS, A.; DOWD, K. *Living with mortality: longevity bonds and other mortality-linked securities*. *British Actuarial Journal*, v. 12, n. 1, p. 153-197, 2006.

BRASIL. Conselho Nacional de Previdência Complementar (CNPc). Resolução nº 47, de 1º de outubro de 2021. Dispõe sobre as regras e os parâmetros para o equacionamento de déficits em planos de benefícios. *Diário Oficial da União*, Brasília, DF, 4 out. 2021.

BRASIL. Conselho Nacional de Seguros Privados (CNSP). Resolução nº 385, de 9 de junho de 2020. Dispõe sobre regras e critérios para operação das coberturas de risco oferecidas em planos de previdência complementar. *Diário Oficial da União*, Brasília, DF, 12 jun. 2020.

BRASIL. Resolução CNPC nº 30, de 10 de outubro de 2018. Dispõe sobre as hipóteses de alteração dos planos de benefícios das entidades fechadas de previdência complementar e dá outras providências. *Diário Oficial da União*: Brasília, DF, 11 out. 2018.

BRAVO, Jorge M. *Forecasting mortality rates with Recurrent Neural Networks: A preliminary investigation using Portuguese data*. Universidade Nova de Lisboa, Université Paris-Dauphine PSL, MagIC & CEFAGE-UE, Portugal, 2021.

CHALLU, Cristian *et al.* N-HITS: *Neural Hierarchical Interpolation for Time Series Forecasting*. 2022.

CHEN, Yuan; KHALIQ, Abdul Q. M. *Comparative Study of Mortality Rate Prediction Using Data-Driven Recurrent Neural Networks and the Lee–Carter Model*. 2021.

COUGHLAN, G.; EPSTEIN, D.; SINHA, A.; HONIG, P. *q-forwards: Derivatives for Transferring Longevity and Mortality Risks*. Pension Advisory Group, JPMorgan Chase Bank NA, 2007. Disponível em: https://www.researchgate.net/publication/256109844_qForwards_Derivatives_for_Transferring_Longevity_and_Mortality_Risks. Acesso em: 10 set. 2025.

FUNPRESP. FCBE custeia riscos de morte, invalidez, aposentadorias especiais e

benefício de sobrevivência. 2019. Disponível em: <https://www.funpresp.com.br/fique-por-dentro/noticias/2019/fcbe-custeia-riscos-de-morte-e-invalidadez-aposentadorias-especiais-e-beneficio-de-sobrevivencia>.

HAYKIN, S. *Neural Networks and Learning Machines*. 3. ed. New Jersey: Prentice Hall, 2009.

HYNDMAN, R. J.; KOEHLER, A. B. *Another look at measures of forecast accuracy*. *International Journal of Forecasting*, v. 22, n. 4, p. 679-688, 2006.

JAMES, G.; WITTEN, D.; HASTIE, T.; TIBSHIRANI, R. *An introduction to statistical learning with applications in R*. Springer, 2013.

LEE, R. D.; CARTER, L. R. *Modeling and forecasting U.S. mortality*. *Journal of the American Statistical Association*, v. 87, n. 419, p. 659-671, 1992.

LEVANTESI, Susanna; NIGRI, Andrea. A random forest algorithm to improve the Lee–Carter mortality forecasting: impact on *q-forward*. *Soft Computing*, v. 24, p. 8553–8567, 2020.

LINDHOLM, M.; PALMBORG, L. *Efficient use of data for LSTM mortality forecasting*. 2022.

MAKRIDAKIS, S.; WHEELWRIGHT, S. C.; HYNDMAN, R. J. *Forecasting: methods and applications*. John Wiley & Sons, 1998.

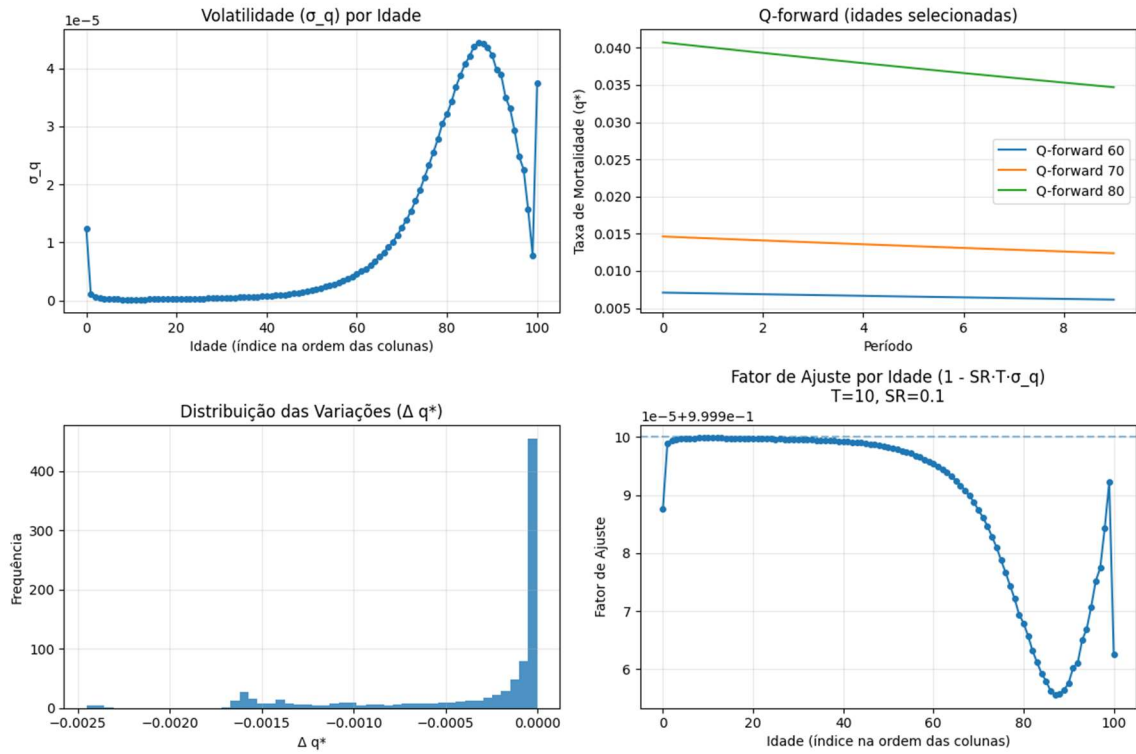
PERLA, Francesca. *Time-series forecasting of mortality rates using deep learning*. 2020.

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.

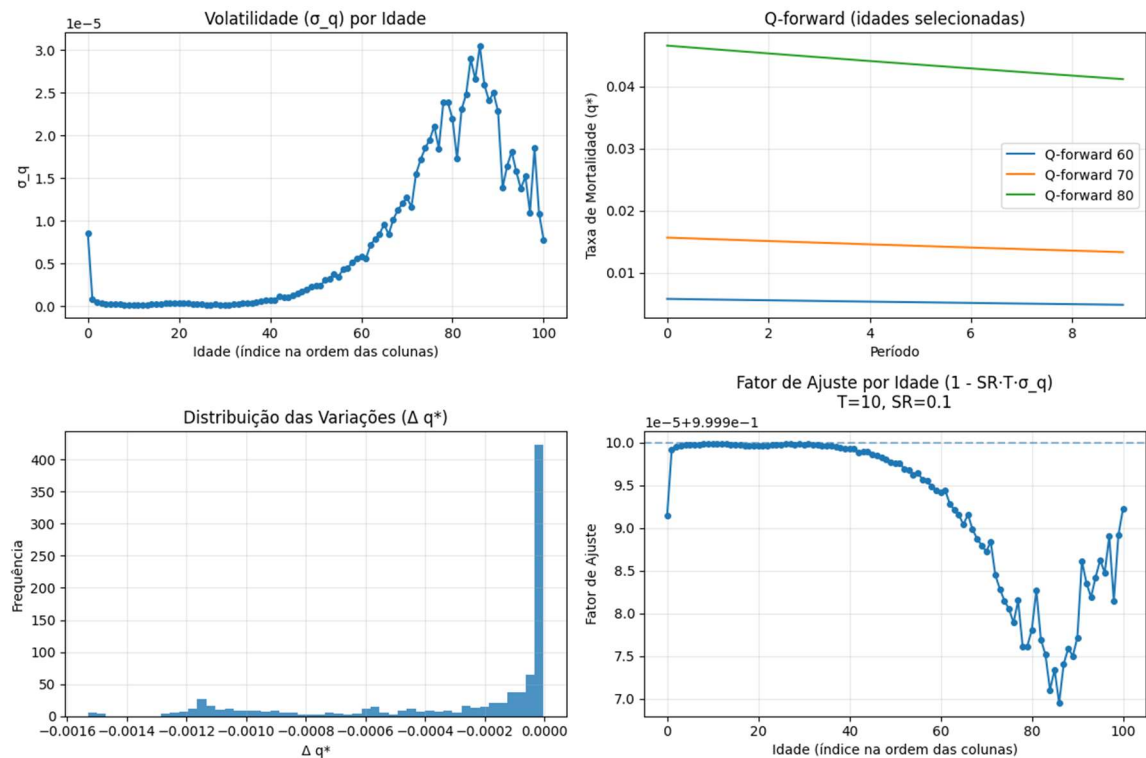
UNITED NATIONS. *World Population Prospects 2022*. 2022. Disponível em: <https://population.un.org/wpp/>. Acesso em: 19 ago. 2025.

APÊNDICE A – DESEMPENHO DO Q-FORWARD PARA O MODELO LEE-CARTER

França

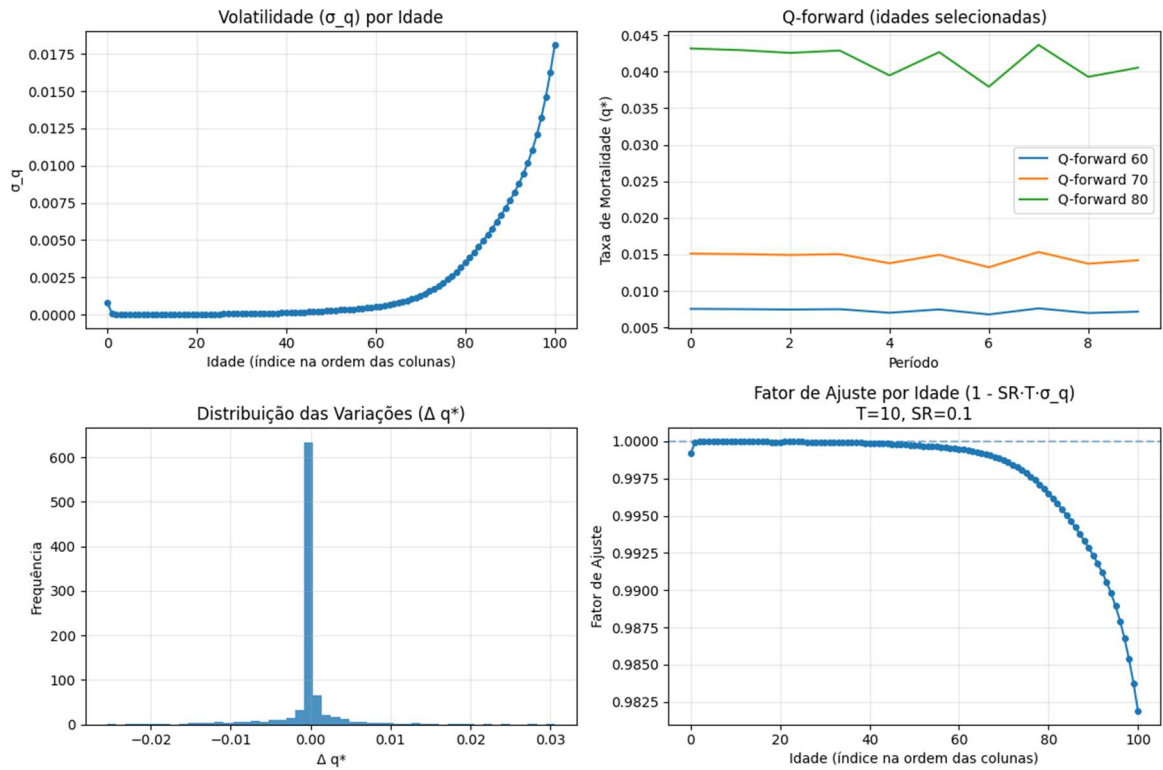


Austrália

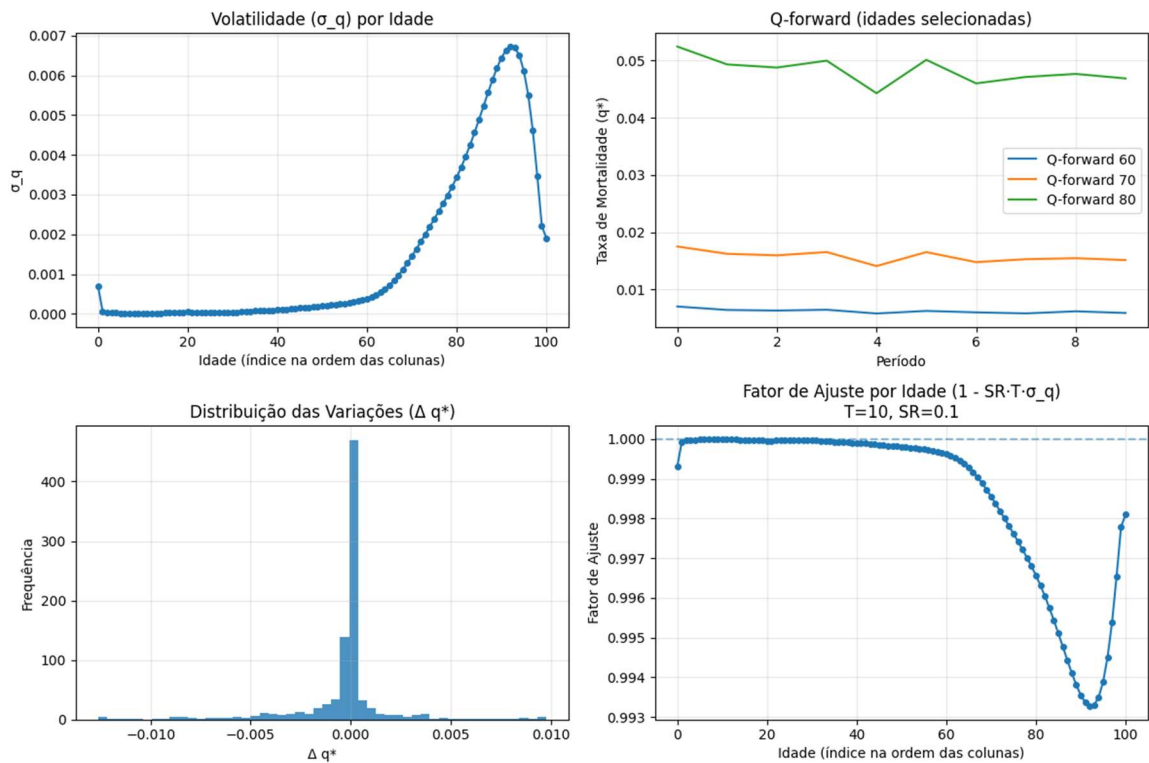


APÊNDICE B – DESEMPENHO DO Q-FORWARD PARA O MODELO MLP

França

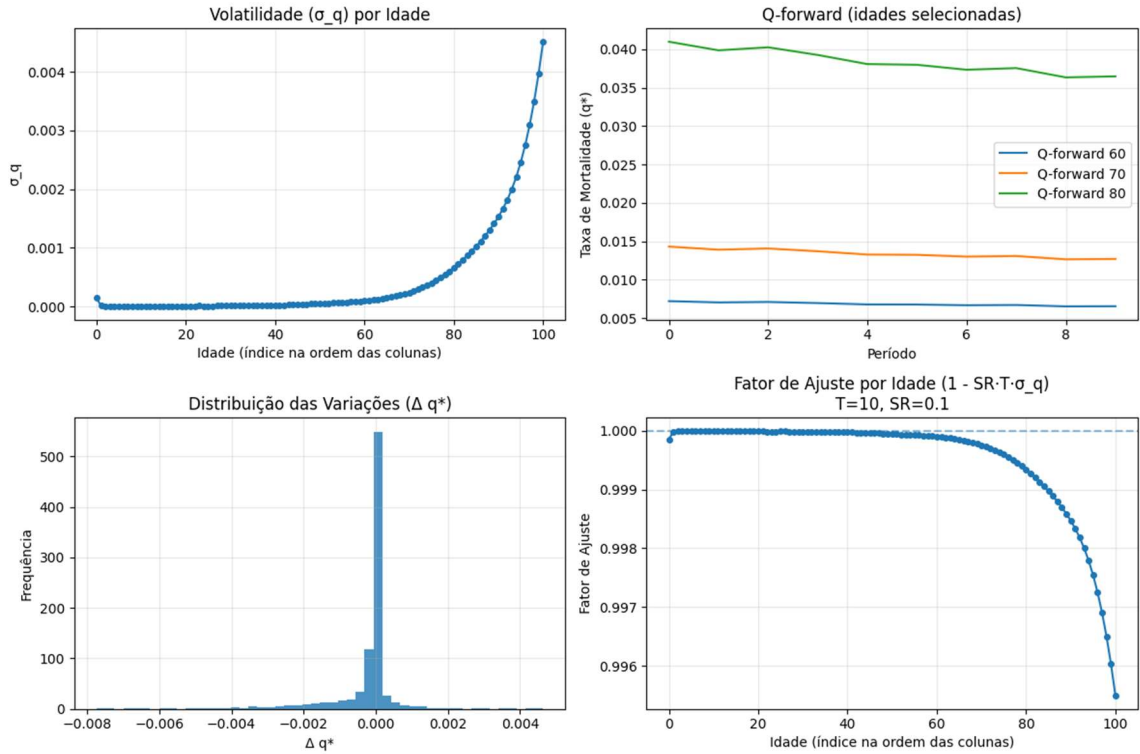


Austrália

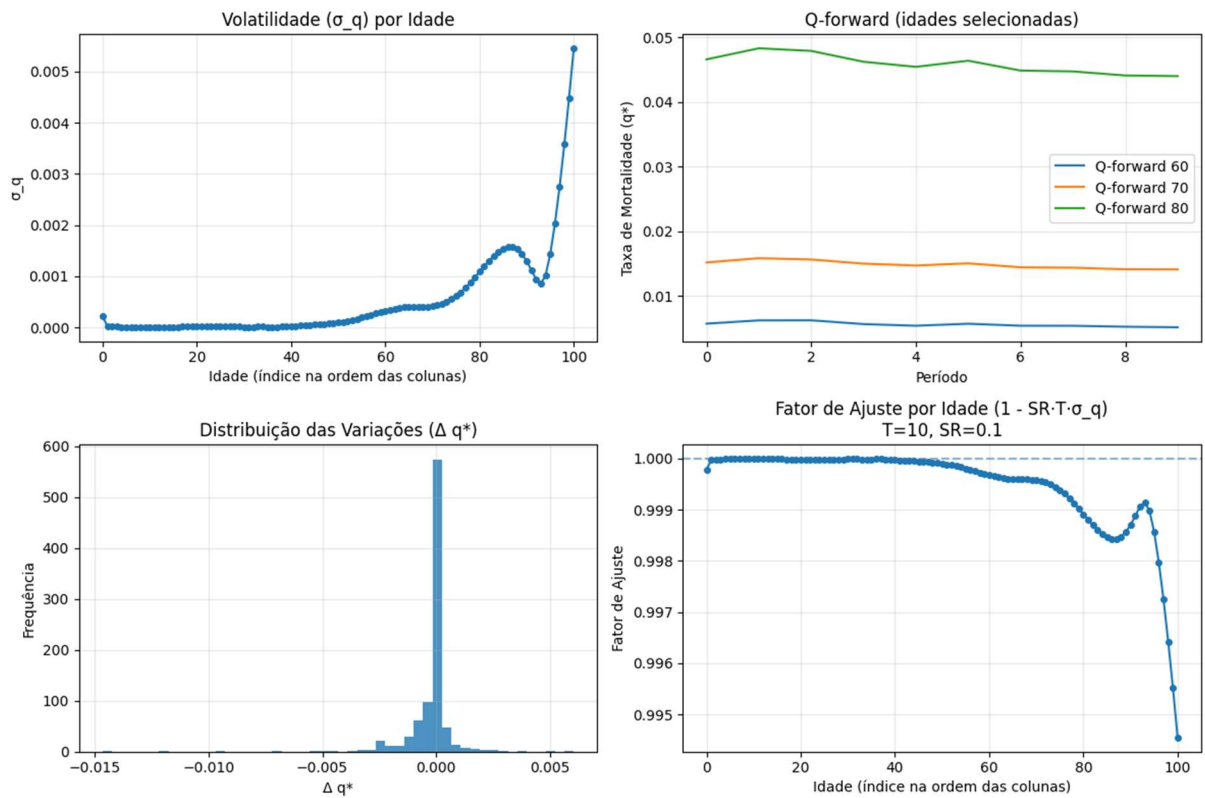


APÊNDICE C – DESEMPENHO DO Q-FORWARD PARA O MODELO N-HITS

França



Austrália



APÊNDICE D – SCRIPT EM R PARA O MODELO LEE-CARTER

```
# Script - Artigo Hybrid Systems for Mortality Forecasting

# lib
library(readr)
library(forecast)
library(demography)
library(lifecontingencies)
library(dplyr)

# exemplo para França (Austrália -> basta substituir o nome do arquivo na
importação e no saving)
french_mortality_data <- read.csv("france.csv")

train <- extract.years(french_mortality_data, years = 1950:2008)
test <- extract.years(french_mortality_data, years = 2009:2018)

# fitting a Lee-Carter
lee_carter <- lca(train, series = 'total', max.age = 100, years =
1950:2008, restype = "logrates")
fT<-forecast(lee_carter,h=10)
lee_carter_log_mortality_forecast <- log(fT$rate$total)
# saving
write_csv(data.frame(lc_forecast), 'lc.csv')
```

APÊNDICE E – SCRIPT EM PYTHON PARA O MODELO MLP

```
## MLP para Previsão de Séries Temporais
# Importação das bibliotecas
import pickle
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
from sklearn.neural_network import MLPRegressor
from sklearn.metrics import mean_squared_error as MSE
from utils import normalizar_serie, desnormalizar, split_sequence, divi-
sao_dados_temporais
from forecast_recursive import forecast
import optunity.metrics

def treinar_mlp(x_train, y_train, x_val, y_val, num_exec, execucao, idade):

    neuronios = [2, 5, 10, 15, 20, 100]
    func_activation = ['tanh', 'relu']
    alg_treinamento = 'adam'
    iteracoes = [1000]
    learning_rate = ['adaptive']
    qtd_lags_sel = len(x_train[0])

    rodadas_lista = []
    neuronios_lista = []
    func_ativacao_lista = []
    qtd_lag_lista = []
    iteracoes_lista = []
```

```

rmse_lista = []

for rodada in range(num_exec):
    melhor_mse = np.Inf
    melhor_neuronio = None
    melhor_modelo = None
    print(f'Rodada {rodada}')

    for neuronio in neuronios:
        for f_act in func_activation:
            for iteracao in iteracoes:
                for qtd_lag in range(1, x_train.shape[1] + 1):
                    mlp = MLPRegressor(hidden_layer_sizes=neuronio, ac-
tivation=f_act, solver=alg_treinamento, max_iter=iteracao, learn-
ing_rate=learning_rate[0], random_state=int(idade))
                    mlp.fit(x_train[:, -qtd_lag:], y_train)
                    predict_validation = mlp.predict(x_val[:, -
qtd_lag:])

                    novo_mse = np.sqrt(MSE(y_val, predict_validation))

                    rodadas_lista.append(rodada)
                    neuronios_lista.append(neuronio)
                    func_ativacao_lista.append(f_act)
                    qtd_lag_lista.append(qtd_lag)
                    iteracoes_lista.append(iteracao)
                    rmse_lista.append(novo_mse)

                    if novo_mse < melhor_mse:
                        melhor_mse = novo_mse
                        melhor_modelo = mlp
                        qtd_lags_sel = qtd_lag
                        melhor_neuronio = neuronio

    # salvar o melhor modelo da rodada
    print('Execução:', execucao, 'idade:', idade, 'rodada:', rodada,
'melhor configuração neuronios:', melhor_neuronio,
'qtd_lag:', qtd_lags_sel, 'RMSE:', melhor_mse)
    with open(f'Resultados - MLP - Artigo - Recursive/exec_{execu-
cao}_model_mlp_rodada_{rodada}_idade_{idade}.sav', 'wb') as file:
        pickle.dump(melhor_modelo, file)

    # salvar o dataframe com todos os resultados
    resultados_df = pd.DataFrame({'rodada': rodadas_lista, 'neuronios':
neuronios_lista, 'func_activation': func_ativacao_lista, 'qtd_lag':
qtd_lag_lista, 'iteracoes': iteracoes_lista, 'rmse': rmse_lista})
    resultados_df.to_csv(f'Resultados - MLP - Artigo - Recursive/exec_{exe-
cucao}_mlp_idade_{idade}_resultados_df.csv')
    # obter o melhor modelo
    melhor_mse_i = resultados_df['rmse'].argmin()
    with open(f'Resultados - MLP - Artigo - Recursive/exec_{execucao}_mo-
del_mlp_rodada_{rodada}_idade_{idade}.sav', 'rb') as file:
        melhor_modelo = pickle.load(file)
    # obter a melhor qtd de lags
    qtd_lags_sel = int(resultados_df.iloc[melhor_mse_i]['qtd_lag'])

    return melhor_modelo, qtd_lags_sel

def main(qtd_execucoes=1):
    # importando os dados

```

```

lnmx_series = pd.read_csv('lnmx_series.csv', sep = ',', index_col = 0)

# Separar os dados em treinamento e teste
# treinamento: 1816-1990
# teste: 1991-2010
treino = lnmx_series.loc[1950:2008,:]
teste = lnmx_series.loc[2009:2018, :]
lnmx_series = lnmx_series.loc[1950:2018,:]

# idades utilizadas
idades = treino.columns

for execucao in range(1, qtd_execucoes+1):
    print(f'Execução {execucao}')
    np.random.seed(execucao + 5)
    previsoes_MLP = np.zeros((10,20))

    for idade, i in zip(idades, range(0,len(idades)+1)):
        lnmx = lnmx_series[idade]
        treino_i = treino[idade]
        teste_i = teste[idade]

        # normalizando a série para MLP
        lnmx_norm = normalizar_serie(lnmx.values)
        treino_i_norm = lnmx_norm[:-10]

        # Separando os dados em X, y e utilizando 2 lags
        X, y = split_sequence(treino_i_norm, n_steps_in=2, n_steps_out
= 1)
        X_train, y_train, X_val, y_val = divisao_dados_temporais(X, y,
perc_treino = 0.8)

        # treinando o modelo
        MLP_model, lag_sel = treinar_mlp(X_train, y_train.ravel(),
X_val, y_val.ravel(), 10, execucao, idade)
        # previsão
        predictions = forecast(MLP_model, treino_i_norm, lag_sel, 10)

        # desnormalizar as previsões
        MLP_predict = desnormalizar(predictions, lnmx.values)

        # Salvando na matrix previsoes_MLP
        previsoes_MLP[:,i] = MLP_predict.values.reshape(10,)

        # transformar em série para realizar o plot
        MLP_predict.index = teste_i.index

    # Salvando os resultados
    previsoes_MLP = pd.DataFrame(previsoes_MLP, columns = idades)

    # Exportando os resultados para .csv
    previsoes_MLP.to_csv('Resultados - MLP - Artigo - Recursive/previ-
soes_MLP_recursive_serie.csv', index=None, header=True, encoding = 'utf-8')

if __name__ == '__main__':
    print('Iniciando o script MLP recursivo')
    main()
    print('FIM!')
```

APÊNDICE F – SCRIPT EM PYTHON PARA O MODELO N-HITS

```

import numpy as np
import pandas as pd
from ray import tune
from ray.tune.search.basic_variant import BasicVariantGenerator
from neuralforecast import NeuralForecast
from neuralforecast.auto import AutoNHITS
from neuralforecast.losses.pytorch import MSE

def main(qtd_execucoes=10):

    nhits_config = {
        "max_steps": 100, # Number of SGD steps
        "input_size": 2,
        "mlp_units": tune.choice([3 * [[64, 64]],
                                   3 * [[128, 128]]]),
        "random_seed": tune.randint(1, 10),
        "early_stop_patience_steps": 5
    }

    # importando os dados
    lnmx_series = pd.read_csv('lnmx_series.csv', sep = ',')

    # Separar os dados em treinamento e teste
    treino = lnmx_series[(lnmx_series.Year >= 1950) & (lnmx_series.Year <=
2008)].copy()
    teste = lnmx_series.loc[(lnmx_series.Year >= 2009) & (lnmx_series.Year
<= 2018)].copy()
    lnmx_series = lnmx_series.loc[(lnmx_series.Year >= 1950) & (lnmx_se-
ries.Year <= 2018)].copy()

    # idades utilizadas
    idades = lnmx_series.iloc[:, 1:].columns

    for execucao in range(1, qtd_execucoes+1):
        print(f'Execução: {execucao}')

        previsoes_nhits = np.zeros((10,20))

        # iteração para realizar o procedimento para cada idade
        for idade, i in zip(idades, range(0,len(idades)+1)):
            print("Idade:", idade)
            treino_i = treino[['Year', idade]].copy()
            treino_i['unique_id'] = 1.0
            treino_i.rename(columns={'Year': 'ds', idade:'y'}, in-
place=True)
            treino_i['ds'] = pd.to_datetime(treino_i.ds, format='%Y')

            teste_i = teste[['Year', idade]]
            teste_i['unique_id'] = 1.0
            teste_i.rename(columns={'Year': 'ds', idade:'y'}, inplace=True)
            teste_i['ds'] = pd.to_datetime(teste_i.ds, format='%Y')
            # treinamento
            model = AutoNHITS(h=10,
                              loss=MSE(),
                              config=nhits_config,
                              search_alg=BasicVariantGenerator(random_state=execucao),
                              backend='ray',
                              num_samples=2)

```

```

        nhits_model = NeuralForecast(models=[model], freq='Y', local_scaler_type='minmax')
        nhits_model.fit(treino_i, val_size=int(treino_i.shape[0] * 0.2))

        # previsao
        predictions = nhits_model.predict().reset_index().copy()

        # salvando na matriz previsoes error nhits
        previsoes_nhits[:,i] = predictions.iloc[:, -1].values.reshape(10,).copy()

        # Salvando os resultados
        previsoes_nhits = pd.DataFrame(previsoes_nhits, columns = idades)

        # Exportando os resultados para .csv
        previsoes_nhits.to_csv(f'Resultados - nhits - Artigo/previsoes_nhits_{execucao}.csv', index=None, header=True, encoding = 'utf-8')

if __name__ == '__main__':
    print('Iniciando o script NHiTS')
    main()
    print('FIM!')
```

APÊNDICE G – SCRIPT EM PYTHON PARA A INTERPOLAÇÃO SPLINE CÚBICA

```

import pandas as pd
import numpy as np

#leitura dos arquivos
arquivos = 'previsoes_mlp_australia_{}.csv'
arquivos_lista = []
for i in range(1, 11):
    df = pd.read_csv(arquivos.format(i))
    arquivos_lista.append(df)

# Média para cada horizonte de tempo e idade
medias_previsoes = []
# iterando os dataframes
for i in range(len(arquivos_lista)):
    idades_previsoes = []
    # para cada idade (coluna)
    for j in range(arquivos_lista[i].shape[1]):
        vetor = np.zeros(10)
        # para cada horizonte de tempo (linha)
        for t in range(arquivos_lista[i].shape[0]):
            vetor[t] = arquivos_lista[i].iloc[i,j]
            idades_previsoes.append(vetor.mean())
        medias_previsoes.append(idades_previsoes)

np.array(medias_previsoes).shape

df = pd.DataFrame(np.array(medias_previsoes),
columns=arquivos_lista[0].columns)
df.to_csv('previsoes_mlp_australia_medias.csv', index=False)

df.iloc[0,:].plot()

df
```

```

# As idades são os nomes das colunas após ler com header=0
x_idades = [int(float(col)) for col in dados.columns]
print("Idades extraídas das colunas:", x_idades)

# Ajustar jan_previsao para corresponder ao número de linhas de dados
jan_previsao_real = dados.shape[0] # Agora todas as linhas são dados
print("Janela de previsão real:", jan_previsao_real)

# Usar a função
curva_interp = interpolacao_curva(dados, x_idades,
jan_previsao=jan_previsao_real)

# Ver resultado
print("\nPrimeiras linhas da curva interpolada:")
print(curva_interp.head())
print("\nÚltimas linhas da curva interpolada:")
print(curva_interp.tail())
print(f"\nShape final da curva: {curva_interp.shape}")

# Salvar com idades como colunas (headers) de 0 a 100 - ATUALIZADA
def salvar_interpolacao_com_idades(curva_interp, nome_arquivo):
    # Transpor para ter previsões nas linhas e idades nas colunas
    dados_transpostos = curva_interp.T # 10 linhas (previsões) x 101
    colunas (idades)

    # Definir as idades como nomes das colunas
    dados_transpostos.columns = [str(i) for i in range(101)] # Colunas:
    '0', '1', '2', ..., '100'

    # Salvar com headers (idades como nomes das colunas)
    dados_transpostos.to_csv(nome_arquivo, index=False)
    print(f"Arquivo salvo como: {nome_arquivo}")
    print(f"Shape: {dados_transpostos.shape} (10 previsões x 101 idades)")
    print("As idades (0-100) são os headers das colunas")

Usar a função
salvar_interpolacao_com_idades(curva_interp,
'curvas_interpoladas_nhits_australia_completas.csv')

import matplotlib.pyplot as plt

# Plotar algumas curvas interpoladas
curva_interp.iloc[:, :3].plot(figsize=(10,6))
plt.title('Curvas Interpoladas de Mortalidade (3 primeiros horizontes de
previsão)')

plt.xlabel('Idade')
plt.ylabel('Taxa de mortalidade (log)') # Mais específico
plt.grid(True)
plt.legend(['Horizonte 1', 'Horizonte 2', 'Horizonte 3']) # Adicionar
legenda
plt.show()

```