



Universidade de Brasília

Instituto de Ciências Exatas  
Departamento de Ciência da Computação

# Implementação do Cálculo da Perda Esperada de Crédito para as Demonstrações Contábeis em IFRS 9 no Brasil

João Paulo Vieira Costa

Dissertação apresentada como requisito parcial para conclusão do  
Mestrado Profissional em Computação Aplicada

Orientador

Prof. Dr. João Carlos Félix Souza

Brasília  
2022

Ficha catalográfica elaborada automaticamente,  
com os dados fornecidos pelo(a) autor(a)

CC837i Costa, João Paulo Vieira  
Implementação do Cálculo da Perda Esperada de Crédito  
para as Demonstrações Contábeis em IFRS 9 no Brasil / João  
Paulo Vieira Costa; orientador João Carlos Félix Souza. --  
Brasília, 2022.  
118 p.

Dissertação (Mestrado - Mestrado Profissional em  
Computação Aplicada) -- Universidade de Brasília, 2022.

1. Risco de Crédito. 2. IFRS 9. 3. Risco de Modelo. 4.  
Gestão de Riscos. I. Souza, João Carlos Félix, orient. II.  
Título.



# Dedicatória

*Dedico este trabalho ao meu pai Vicente de Paulo de Souza Costa e à minha avó Eliotéria Vieira das Virgens, que já não podem compartilhar desta fase de amadurecimento pessoal e profissional comigo, mas com certeza serviram de apoio em todos os ensinamentos deixados.*

# Agradecimentos

Agradeço a minha família e amigos, pelo apoio e compreensão durante as ausências necessárias para realização dessa pesquisa. Em especial à Monique Ribeiro, que me acompanhou em todo o período de noites mal dormidas, eventos e momentos de lazer adiados, sendo o mais importante suporte nesta jornada.

Aos meus colegas, professores e ao meu orientador João Carlos Felix que compartilharam um período de muito aprendizado e autoria de trabalhos científicos.

Aos meus gestores do Banco do Brasil, por terem me dado o apoio necessário para realização da pesquisa e, principalmente, por me permitirem participar dos desafios relacionados aos temas deste trabalho.

Todas essas contribuições foram fundamentais para a construção deste trabalho e sem eles sua realização não seria possível.

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES), por meio do Acesso ao Portal de Periódicos.

# Resumo

Nas últimas décadas, os órgãos internacionais e reguladores locais apresentaram várias evoluções nas políticas para gerenciamento do risco de crédito e do capital necessário para cobertura de eventos de risco, com mecanismos desenvolvidos especialmente após a última crise financeira. O IFRS 9 é um padrão contábil utilizado internacionalmente em instituições financeiras e, dentre outras definições, traz regras e para mensuração da deterioração e perda esperada associada aos ativos de crédito das instituições financeiras e entrará em vigor no Brasil como padrão contábil em 2025. Este trabalho traz uma visão abrangente sobre os modelos utilizados para o cálculo da ECL e sua utilização. Para isso foram apresentados estudos de caso com aplicação de modelos tradicionais e de *machine learning* sobre os modelos de ajuste *lifetime* com modelos de Análise de Sobrevivência e modelos para projeção de NPL a partir das condições econômicas. Além disso, foi aplicado estudo de simulações para comparação do modelo vigente com o IFRS 9 em situações específicas, exemplificando que o uso de modelos bem ajustados é importante para evitar despesas de provisão elevadas que podem comprometer os resultados da instituição. Neste contexto, tendo em vista essa necessidade, aplicou-se metodologia para quantificação do risco de modelo associado aos modelos estudados, sintetizando os resultados obtidos no restante do trabalho na tomada de decisão sobre os melhores modelos a serem utilizados. Cada capítulo apresenta uma contribuição independente, mas trabalhos estão relacionados em seu uso para estimação da ECL no contexto do IFRS 9. Para a estimação da PD *lifetime*, o modelo RSF foi a opção de menor risco de modelo, considerando seu melhor desempenho em relação aos demais. Já sobre as projeções *forward-looking* da ECL, um modelo OLS apresentou melhores medidas de desempenho utilizando variáveis projetadas como insumo e por isso foi considerado apresentou menor risco.

**Palavras-chave:** Risco de Crédito, IFRS 9, Risco de Modelo, Gestão de Riscos

# Abstract

In recent decades, international bodies and local regulators have presented several evolution in policies for managing credit risk and the capital needed to cover risk events, with mechanisms developed especially after the last financial crisis. IFRS 9 is an accounting standard used internationally in financial institutions and, among other definitions, provides rules for measuring the deterioration and expected loss associated with the credit assets of financial institutions and will enter into force in Brazil as an accounting standard in 2025. This work provides a comprehensive view of the models used to calculate the ECL and its use. For this, case studies were presented with the application of traditional and *machine learning* models on the *lifetime* adjustment models with Survival Analysis models and models for NPL projection from the conditions economic. In addition, a simulation study was applied to compare the current model with the IFRS 9 in specific situations, exemplifying that the use of well-adjusted models is important to avoid high provisioning expenses that can compromise the institution's results. In this context, in view of this need, a methodology was applied to quantify the model risk associated with the models studied, summarizing the results obtained in the rest of the work in the decision making on the best models to be used. Each chapter presents an independent contribution, but works are related in its use for estimating ECL in the context of IFRS 9. For estimating PD *lifetime*, the model RSF was the option with the lowest model risk, considering its better performance compared to the others. As for the *forward-looking* projections of the ECL, a model OLS presented better performance measures using projected variables as input and therefore it was considered to have lower risk.

**Keywords:** Credit Risk, IFRS 9, Model Risk, Risk Management

# Sumário

<b>1</b>	<b>Introdução</b>	<b>1</b>
1.1	Demonstrações Financeiras em Instituições Financeiras . . . . .	2
1.2	Objetivo . . . . .	4
1.2.1	Objetivo Geral . . . . .	4
1.2.2	Objetivos Específicos . . . . .	4
<b>2</b>	<b>Referencial Teórico</b>	<b>6</b>
2.1	IFRS 9 . . . . .	6
2.1.1	Contexto Geral . . . . .	6
2.2	Implementação do Padrão Internacional IFRS 9 no Brasil . . . . .	7
2.3	Risco de Modelo . . . . .	8
2.3.1	Contexto Brasileiro . . . . .	9
<b>3</b>	<b>Metodologia</b>	<b>11</b>
3.1	Fontes de Risco de Modelo no Padrão IFRS 9 . . . . .	12
3.1.1	Ajuste da ECL ao Prazo . . . . .	13
3.1.2	Ajuste da ECL a Cenários <i>Forward-Looking</i> . . . . .	14
3.1.3	Estimativa da Perda Esperada . . . . .	15
3.2	Mensuração do Risco de Modelo . . . . .	16
3.2.1	Por que Priorizar Modelos? . . . . .	17
3.2.2	Mensuração do Risco de Modelo . . . . .	18
<b>4</b>	<b>Ajuste Lifetime</b>	<b>19</b>
4.1	Introduction . . . . .	19
4.2	Referencial Teórico . . . . .	20
4.2.1	Mineração de Dados e Modelos de PD . . . . .	20
4.2.2	Análise de Sobrevivência . . . . .	21
4.2.3	Modelo Cox <i>Proportional-Hazards</i> . . . . .	22
4.2.4	Modelos Baseados em Árvore . . . . .	22
4.2.5	<i>Survival Tree</i> . . . . .	23



4.2.6	<i>Random Survival Forest</i>	23
4.2.7	Trabalhos Correlatos	24
4.3	Metodologia	24
4.4	Resultados	26
4.4.1	<i>Survival Tree</i>	27
4.4.2	<i>Random Survival Forest</i>	29
4.4.3	Avaliação dos Modelos	31
4.5	Conclusão	32
<b>5</b>	<b>Ajuste Forward-Looking</b>	<b>34</b>
5.1	Introdução	34
5.2	Referencial Teórico	35
5.2.1	Modelagem Econométrica	35
5.2.2	Trabalhos Correlatos	36
5.2.3	Modelos de <i>Machine Learning</i>	37
5.3	Metodologia	39
5.4	Resultados	40
5.4.1	Dados Coletados	40
5.4.2	Análise Exploratória	44
5.4.3	Modelos Desenvolvidos	46
5.4.4	Modelos de Machine Learning	49
5.4.5	Avaliação dos Modelos	51
5.4.6	Conclusão	52
<b>6</b>	<b>Cálculo da Perda Esperada de Crédito</b>	<b>54</b>
6.1	Introdução	54
6.2	Referencial Teórico	55
6.3	Metodologia	57
6.4	Resultados	57
6.4.1	Dinâmica da carteira de crédito	57
6.4.2	Calibragem dos Parâmetros	59
6.4.3	Cálculos da ECL e PCLD da Carteira Simulada	60
6.4.4	Sensibilidade das Metodologias	63
6.4.5	Atualização Dinâmica dos Parâmetros de Risco	67
6.5	Conclusão	68
<b>7</b>	<b>Risco de Modelo</b>	<b>69</b>
7.1	Introdução	69

7.2	Referencial Teórico . . . . .	71
7.2.1	Gerenciamento do Risco de Modelo . . . . .	71
7.2.2	Análise de Decisão por Multicritérios (MCDA) . . . . .	74
7.3	Metodologia . . . . .	75
7.3.1	Método PROMETHEE . . . . .	75
7.4	Estudo de Caso . . . . .	77
7.4.1	Estruturação do problema . . . . .	77
7.4.2	Aplicação Numérica . . . . .	77
7.4.3	Análise de Sensibilidade . . . . .	78
7.5	Conclusões . . . . .	80
<b>8</b>	<b>Considerações Finais</b>	<b>82</b>
8.1	Limitações . . . . .	83
8.2	Contribuições . . . . .	84
	<b>Referências</b>	<b>85</b>
	<b>Apêndice</b>	<b>94</b>
<b>A</b>	<b>Crítérios e Escalas para o Risco de Modelo</b>	<b>95</b>
<b>B</b>	<b>Scripts para Simulação da Carteira de Crédito</b>	<b>97</b>

# Lista de Figuras

1.1	Implementação da IFRS 9 pelos respondentes da pesquisa <i>World Bank</i> . . .	3
3.1	Fluxo Metodológico. . . . .	11
3.2	CRISP-DM e Ciclo de Vida de Desenvolvimento de Software. . . . .	17
4.1	Importância das Variáveis. . . . .	27
4.2	<i>Survival Tree</i> para Probabilidade de Descumprimento. . . . .	29
4.3	Curvas de Sobrevivência dos Nós Finais. . . . .	30
5.1	Ajuste de Sazonalidade da taxa de desemprego. . . . .	43
5.2	Erros Absolutos da Projeção do PIB por Defasagem. . . . .	45
5.3	Taxas de Inadimplência das Unidades de Cortes Transversais. . . . .	47
5.4	SHAP Value Modelo Random Forest - Variável Modelo 4. . . . .	50
5.5	SHAP Value Modelo Random Forest Fin. . . . .	51
6.1	Simulação da Dinâmica de Atrasos. . . . .	58
6.2	Simulação da Dinâmica de Atrasos. . . . .	59
6.3	Exemplo de Contrato Simulado. . . . .	60
6.4	Taxas de Descumprimento Médias Simuladas por Modalidade. . . . .	62
6.5	Curvas de NPL, iPE e Iprov dos Dados Simulados. . . . .	63
6.6	Sensibilidade dos Cálculos a Variações da Inadimplência. . . . .	64
6.7	Impacto do Choque na atualização da ECL com Gatilho de Atualização. . .	65
6.8	Impacto do Choque nas Despesas de Provisão. . . . .	66
6.9	Impacto do Choque nas Despesas de Provisão com Atualização Dinâmica. .	67
7.1	Números <i>Fuzzy</i> Triangular para o Primeiro Critério. . . . .	78
7.2	Medida de Consistência. . . . .	79
7.3	Riscos de Modelo Atribuídos. . . . .	80

# Lista de Tabelas

3.1	Fontes de Risco de Modelo (IFRS 9) . . . . .	13
4.1	Resultados do Modelo <i>Survival Tree</i> . . . . .	28
4.2	Resultados do Modelo <i>Random Survival Forest</i> . . . . .	30
4.3	Comparação dos Modelos . . . . .	31
5.1	Taxas de Inadimplência por Unidade de Corte Transversal . . . . .	42
5.2	Variáveis Macroeconômicas para Modelagem . . . . .	42
5.3	Medidas de Erro da Projeção das Variáveis Explicativas (12 meses) . . . . .	44
5.4	P-valor dos Testes de Raiz Unitária das Variáveis Explicativas . . . . .	45
5.5	Modelos Desenvolvidos Utilizando Variáveis mais recentes . . . . .	48
5.6	Modelos Desenvolvidos Utilizando Defasagens Maiores . . . . .	48
5.7	MAPE <i>Random Forest</i> por Profundidade das Árvores . . . . .	49
5.8	Avaliação dos Modelos . . . . .	51
5.9	Teste T- <i>student</i> e Wicoxon Pareados dos Erros dos Modelos . . . . .	52
6.1	Classificações de Risco Resolução nº 2682 . . . . .	54
6.2	Dados Simulados x SFN - Jun/2021 . . . . .	61
7.1	Fontes de Risco de Modelo . . . . .	72
7.2	Classificação dos Modelos Desenvolvidos . . . . .	79

# Lista de Abreviaturas e Siglas

**ADF** Augmented Dickey-Fuller.

**AUROC** Area Under the Receiver Operating Characteristics.

**BCB** Banco Central do Brasil.

**BCBS** Basel Committee on Banking Supervision.

**BIS** Bank for International Settlements.

**CART** Classification and Regression Trees.

**CHF** Cumulative Hazard Function.

**CMN** Conselho Monetário Nacional.

**COSIF** Plano Contábil das Instituições do Sistema Financeiro Nacional.

**EAD** Exposure at Default.

**EBA** European Banking Authority.

**ECL** Expected Credit Loss.

**FED** Federal Reserve System.

**GAAP** Generally Accepted Accounting Principles.

**GMM** Generalized Method of Moments.

**IAS** International Accounting Standard.

**IASB** International Accounting Standards Board.

**IF** Instituição Financeira.

**IFRS** International Financial Reporting Standards.

**IRB** Internal Ratings-Based.

**KPSS** Kwiatkowski-Phillips-Schmidt-Shin.

**KS** Kolmogorov-Smirnov.

**LGD** Loss Given Default.

**MAE** Mean Absolute Error.

**MAPE** Mean Absolute Percentage Error.

**MCDA** Multiple-criteria decision analysis.

**NPL** Non-Performing Loans.

**OLS** Ordinary Least Squares.

**PCLD** Provisão para Créditos de Liquidação Duvidosa.

**PD** Probability of Default.

**PF** Pessoa Física.

**PH** Proporcional-Hazards.

**PIB** Produto Interno Bruto.

**PJ** Pessoa Jurídica.

**PP** Phillips-Perron.

**PROMETHEE** Preference Ranking Organization Method for Enrichment.

**RSF** Random Survival Forest.

**SFN** Sistema Financeiro Nacional.

**SGS** Sistema Gerenciador de Séries Temporais.

**SHAP** SHapley Additive exPlanations.

**TI** Tecnologia da Informação.

# Capítulo 1

## Introdução

A discussão sobre as melhorias na gestão do capital para cobertura de eventos de riscos nas instituições financeiras vem ocorrendo de forma contínua nas últimas décadas. Após a grande depressão de 1929, o *Bank for International Settlements* (BIS) foi fundado para promover a cooperação entre os bancos centrais em suas atribuições [1]. Na estrutura do BIS, o *Basel Committee on Banking Supervision* (BCBS) publicou, em 1988, o Acordo de Basileia I, com o objetivo de aprimorar a gestão de capital nas instituições financeiras [2].

Neste mesmo contexto, em 2004, o Segundo Acordo de Basileia II [3] foi publicado, trazendo avanços em relação à modelagem para quantificação do risco de crédito na gestão do capital. Na mesma época, o *Working Paper* 14 [4] com a apresentação boas práticas para gerenciamento dos modelos utilizados para gestão de riscos nas Instituições Financeiras (IFs). Mesmo com todos estes avanços e cooperação, a história recente mostrou a necessidade de constante evolução dos mecanismos associados à boa gestão dos ativos financeiros nas IFs.

Entre os anos de 2007-2008, a crise do *Subprime* causou uma das maiores recessões já observadas nas economias a um nível global. Os impactos observados naquele período tiveram origem no mercado de hipotecas dos EUA. Uma das causas mapeadas para esta crise foi o brusco crescimento do mercado de hipotecas vivenciado entre 2001 e 2006, sem o devido ajuste do risco de crédito às características dos tomadores e das operações contratadas naquele período. A atratividade deste mercado levou a um rápido crescimento dos preços dos ativos, seguido de uma deterioração das condições dos créditos, sem as proteções de liquidez adequadas para as instituições financeiras [5].

A partir deste evento, houve uma pressão para que as Instituições Financeiras e seus reguladores para uma melhor compreensão sobre o uso de modelos para quantificação de riscos, dando uma atenção especial para a avaliação se os modelos são suficientemente robustos sob mudanças nas premissas adotadas para seu uso [6]. Segundo Cohort [7], naquele momento, passou a ser entendida a importância do risco de modelo, conceito que

será discutido mais detalhadamente nos capítulos seguintes.

Após as lições aprendidas por conta da crise do *Subprime*, por atribuir a severidade da crise ao excesso de alavancagem em que operavam as instituições financeiras em alguns países, o Acordo de Basileia III [8] foi publicado com o objetivo de aprimorar a gestão do capital nas IFs, melhorar a transparência e a governança e gestão de riscos, inclusive com menções sobre a necessidade de salvaguardas adicionais para o risco de modelo.

Além da gestão do capital para cobertura de eventuais perdas, a precificação dos ativos nas IFs tem relação com as demonstrações financeiras e os resultados, declarados nos exercícios contábeis pelas IFs. Os diversos riscos envolvidos em relação aos ativos de IFs fazem com que a forma adequada de sua declaração em termos contábeis também seja objeto de constantes aprimoramentos nos últimos anos.

## 1.1 Demonstrações Financeiras em Instituições Financeiras

Tendo como função a intermediação financeira, as IFs que operam crédito, de modo geral, apresentam uma alavancagem elevada em relação a outros setores da economia. De um lado os bancos captam recursos, gerando passivos, e do outro lado os créditos concedidos são contabilizados como ativos da instituição. São evidentes os riscos associados à operação da intermediação bancária, como variações nos passivos administrados (risco de mercado), descasamentos entre os ativos e passivos (risco de liquidez) ou não cumprimento de obrigações pelos tomadores de créditos (risco de crédito) [9]. Dessa situação surge a necessidade de estabelecimento de capital mínimo para cobertura de quaisquer destes riscos.

Além da relação patrimonial dos bancos, e da gestão do capital nas IFs, os exercícios contábeis precisam considerar variações e riscos associados a estes ativos. Dessa forma, no mercado brasileiro, a Provisão para Créditos de Liquidação Duvidosa (PCLD) são passivos constituídos para cobertura de eventuais perdas associadas ao risco de crédito [10]. As despesas de PCLD, apresentadas nas demonstrações financeiras dos bancos, representam a variação da PCLD, acrescida dos valores contabilizados como prejuízo no exercício [11].

Assim como no contexto da gestão do capital, há discussões contínuas sobre as melhores práticas e os padrões adequados para a contabilidade bancária.

Em 2014, o *International Accounting Standards Board* (IASB) publicou o padrão *International Financial Reporting Standards* (IFRS) 9 [12], em substituição ao *International Accounting Standard* (IAS) 39 [13], tendo como principal mudança de paradigma a fundamentação nas perda esperada de crédito em substituição à perda incorrida [14]. A Figura



1.1 mostra o resultado de uma pesquisa realizada pelo Banco Mundial com os bancos centrais sobre o uso do padrão IFRS 9 [15].

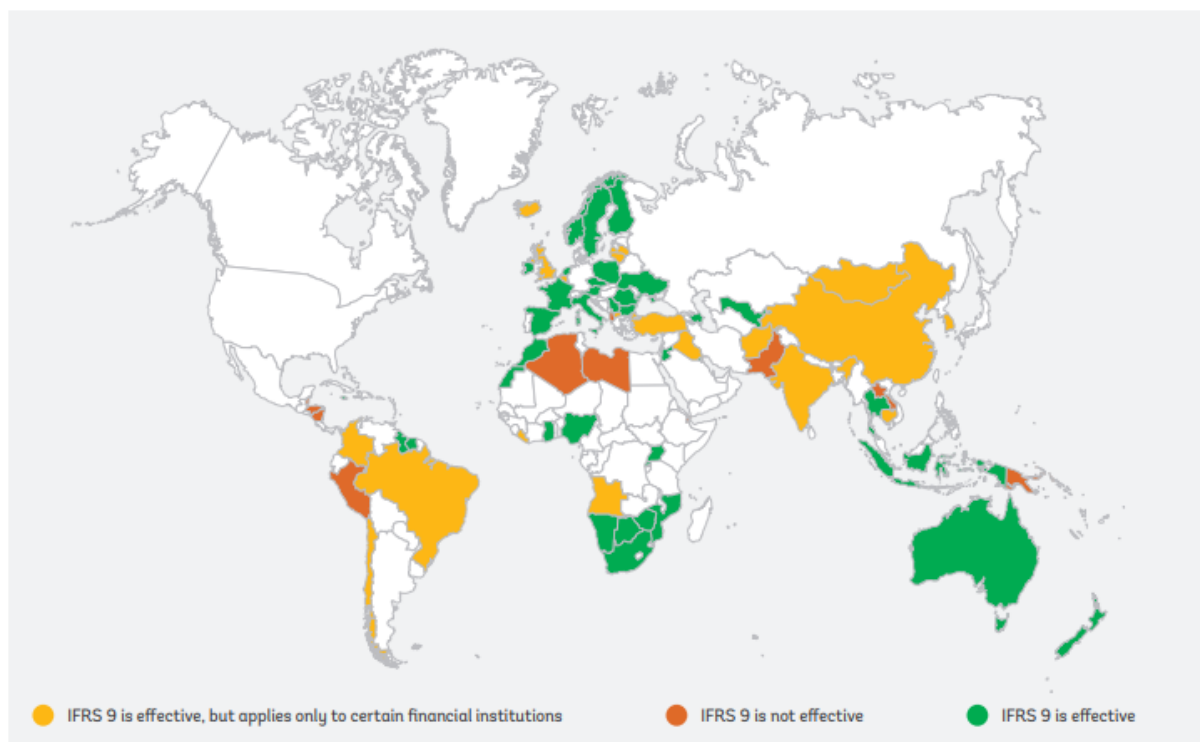


Figura 1.1: Implementação da IFRS 9 pelos respondentes da pesquisa *World Bank* (Fonte: *World Bank*[15]).

Ainda relativamente a pesquisa realizada pelo Banco Mundial em relação ao uso do IFRS 9, os reguladores foram questionados sobre os maiores desafios relacionados à implementação do padrão. Como resultado desse questionamento, o maior desafio, citado por 69% dos respondentes é a "modelagem do risco".

No Brasil, as normas contábeis tem fundamentam-se em uma junção entre a perda incorrida e a perda esperada [14], apresentada pelo Banco Central do Brasil (BCB) no *Generally Accepted Accounting Principles* (GAAP), ou COSIF, em que a PCLD é calculada pelas regras da Resolução CMN N° 2.682 [10]. No entanto, as demonstrações do padrão internacional IFRS 9 é apresentado para fins de reportes anuais, conforme Resolução CMN N° 4.818, de 29 de maio de 2020 [16].

Desde 2017, havia uma consulta pública do BCB [17] sobre a incorporação dos preceitos do IFRS 9 no COSIF, de modo que as IFs mantinham a expectativa da compatibilização do padrão internacional com o brasileiro. Enfim, em 25 de novembro de 2021, foi publicada a Resolução CMN N° 4.966 [18], com o seguinte contexto:

“Dispõe sobre os conceitos e os critérios contábeis aplicáveis a instrumentos financeiros, bem como para a designação e o reconhecimento das relações de proteção (contabilidade de *hedge*) pelas instituições financeiras e demais instituições autorizadas a funcionar pelo Banco Central do Brasil.”

Essa Resolução estabelece os critérios contábeis para as IFs autorizadas a funcionar pelo BCB e entra totalmente em vigor em 1 de janeiro de 2025, quando, entre outras, estará revogada a Resolução CMN N<sup>o</sup> 2.682.

## 1.2 Objetivo

Pela discussão apresentada nos tópicos anteriores, pode-se afirmar que a crise de 2007-2008 teve relação com a experiência em gerenciar capitais voláteis nas IFs, no gerenciamento do risco de crédito e seu impacto na liquidez [8] e no uso adequado de modelos para precificação de ativos [7]. Considerando, ainda, os desafios destacado pelos reguladores na modelagem de risco na implementação do padrão IFRS 9, fica evidenciada a relevância da continuidade das discussões a respeito do risco de modelo associado ao uso do padrão IFRS 9, tema ao qual se destina debater neste trabalho. Sob este cenário que se define os objetivos geral e específicos.

### 1.2.1 Objetivo Geral

Este trabalho tem como objetivo geral proporcionar uma visão geral dos modelos necessários para implementação do Cálculo da Perda Esperada de Crédito (*Expected Credit Loss* (ECL)) para as demonstrações financeiras em IFRS 9 no Brasil.

### 1.2.2 Objetivos Específicos

Para atingimento deste objetivo, serão necessários os seguintes objetivos específicos:

- Identificação dos principais Modelos relacionados às Demonstrações em IFRS 9;
- Levantamento das melhores práticas sobre o desenvolvimento dos modelos para as necessidades elencadas;
- Estudos de caso com aplicação de modelos utilizados utilizando dados e simulações alinhados às melhores práticas;
- Proposição de Metodologia para Avaliação do Risco de Modelo (quantificação);
- Aplicação da abordagem de mensuração do risco de modelo aos modelos utilizados para fins de IFRS 9.

Além desta introdução, este trabalho traz mais sete capítulos. No segundo capítulo, serão discutidos os conceitos e as contribuições apresentadas na literatura sobre os temas centrais deste trabalho, ou seja, o cálculo da perda esperada no padrão IFRS 9 e o risco de modelo. O terceiro capítulo aborda a metodologia da pesquisa aplicada neste trabalho. Do quarto ao sexto capítulo são apresentados estudos de casos independentes sobre os modelos mapeados, enquanto o sétimo capítulo apresenta a abordagem para o risco de modelo. Por fim, o oitavo capítulo apresenta as considerações finais sobre os resultados da pesquisa.

# Capítulo 2

## Referencial Teórico

Como foi discutido no capítulo anterior, o impacto causado pelo gerenciamento incorreto do uso de modelos pode gerar resultados severos para as instituições, podendo até produzir efeitos indesejados. Ou pelo menos pode-se afirmar que a boa gestão dos modelos utilizados pode minimizar efeitos negativos como os da crise de 2007-2008.

Em linha com este entendimento, considerando a abrangência do uso do padrão IFRS 9 nas IFs, e as estimativas provenientes do seu uso, o gerenciamento do risco de modelo se relaciona com o sucesso das melhorias esperadas pela sua adoção. Este capítulo apresenta discussões recentes sobre estes dois temas apresentadas na literatura e nas instituições responsáveis pela estabilidade financeira.

### 2.1 IFRS 9

#### 2.1.1 Contexto Geral

Desde a publicação do padrão IFRS 9 e, em especial, desde 2018, quando tornou-se mandatário em vários países, diversas contribuições na literatura discorrem sobre este tema. Alguns dos trabalhos trazem discussões sobre as diferenças e os impactos da transição do padrão IAS 39 para o IFRS 9 [19] [20].

Conforme descrito por Krüger [21], as provisões para perdas dos créditos, historicamente, eram baseadas na perda incorrida, o que significa que estariam relacionadas aos momentos de desaceleração da economia. Ainda segundo este mesmo autor, o foco da implementação do padrão IFRS 9 é justamente dar uma abordagem prospectiva a essas provisões, com o uso do cálculo da perda esperada de crédito. Nesse trabalho, os autores estudam os efeitos do ciclo econômico sobre as estimativas da perda esperada, e utilizou parâmetros de risco do Acordo de Basileia II, estabelecidos para o gerenciamento do capital para o cálculo da ECL conforme a equação 2.1 [3]:

$$ECL = PD \times LGD \times EAD \quad (2.1)$$

sendo,

- *Probability of Default* (PD), que é a probabilidade de um tomador se tornar inadimplente;
- *Loss Given Default* (LGD), a perda após a inadimplência; e
- *Exposure at Default* (EAD), tida como a exposição no momento da inadimplência.

Um estudo [22] utilizou as estimativas de perda esperada pelo padrão IFRS 9 e duas relações com os ciclos econômicos para testes de estresse, avaliando a sustentabilidade das IFs a as variações econômicas e seus efeitos nos créditos. Mechelli [23], em um estudo envolvendo 43 países que utilizam o padrão IFRS 9, avaliou os efeitos da qualidade da governança corporativa e das proteções a investidores a nível país sobre a relevância dos balanços patrimoniais seguindo o padrão IFRS 9. Neste estudo, os autores concluíram que o padrão IFRS 9 é mais relevante que o IAS 39 em situações de bons níveis de governança corporativa e de proteção aos investidores, bem como o oposto é verdade.

Georgiou [24] apresenta uma abordagem sobre a estimativa de cadeias de markov para estimativa do risco de crédito no contexto do padrão do IFRS 9, trazendo novas discussões sobre a forma de modelagem do risco de crédito atendendo ao padrão. Outros estudos recentes se comprometeram a avaliar os desafios do gerenciamento das estimativas e dos modelos utilizados para o cálculo da ECL no contexto da pandemia do COVID-19 [25] [26].

O que se pode notar é que o conceito de modelo, e conseqüentemente o risco de modelo, tem relação com as atividades relacionadas ao gerenciamento do risco de crédito. As alterações trazidas pelo IFRS 9, em sua substituição das perdas incorridas por perda esperada, com uma visão mais prospectiva, tem relação com o uso de modelos para estimativa dos efeitos dos fatores macroeconômicos sobre a deterioração dos créditos. Além disso, outros pontos do padrão têm efeitos no risco de modelo ao qual as IFs estão sujeitas, questão que será objeto de análise mais aprofundada nos capítulos seguintes.

## **2.2 Implementação do Padrão Internacional IFRS 9 no Brasil**

No Brasil, o padrão IFRS 9 tem sido reportado pelas IFs autorizadas a funcionar pelo BCB, desde que sejam companhias abertas, líderes de conglomerado prudencial enqua-

drados nos Segmentos 1, 2 ou 3, conforme Resolução 4.818 [16]. Estas demonstrações são consolidadas e anuais e não impactam a distribuição de resultados formal das instituições.

A partir de 1 de janeiro de 2025, conforme estabelecido na Resolução CMN Nº 4.966 [18], as regras vigentes serão substituídas pelas destacadas nesta Resolução, para compatibilização com o padrão internacional. Diferentemente do disposto na Resolução CMN Nº 2.682 [10], em que os níveis de provisionamento mínimos são estimulados pelo BCB de acordo com o atraso, a nova regulação contábil estabelece que devem ser estabelecidos grupos homogêneos de risco.

A partir destes grupos homogêneos de risco, a ECL deve ser calculada a partir do valor presente justo do ativo, probabilidade de um ativo financeiro se tornar problemático e da expectativa de recuperação dos ativos considerados problemáticos. Como pode ser notado, no caso brasileiro, a convergência entre os parâmetros de risco utilizados para gestão do capital e a forma de cálculo da nova regulação fica evidente.

Outras questões, que serão discutidas em mais detalhes nas seções seguintes, relativas a efeitos econômicos sobre o risco de crédito e forma de classificação dos ativos também são implementadas, com certo grau de liberdade para os métodos aplicados pelas IFs, desde que defensáveis e embasados com estudos internos.

Dessa forma, a necessidade de estimativas consistentes e confináveis nas IFs sobre os parâmetros de risco ganham ainda mais relevância, tendo em vista a possibilidade de perdas em decorrência do mau uso dos modelos associados a este objetivo.

## 2.3 Risco de Modelo

Conforme mencionado no capítulo anterior, há registro à necessidade de salvaguardas para o risco de modelo no Acordo de Basileia III [8], publicado em 2010, e nos Acordos anteriores o risco de modelo não é citado. Apesar disso, menções ao gerenciamento do risco de modelo em instituições financeiras na literatura tem registros anteriores.

Na década de 90, antes dos recentes avanços computacionais e da crise de 2007-2008, o *OCC-bulletin* 2000-16 [27] menciona a possibilidade de impactos severos na reputação e na lucratividade de bancos e várias outras formas de perdas, devido a erros nos modelos e refere esses eventos como risco de modelo. Outro trabalho [28] nesta mesma época traz uma discussão sobre os efeitos do risco de modelo nos preços, *hedge* e cálculo de capital, e define estrutura básica para gerenciamento do risco de modelo.

Este mesmo autor relaciona crises do mercado de capitais dos anos de 1970-1980 a falta de modelos e menciona que, naquele momento, a grande quantidade de modelos seria uma questão a ser gerenciada. As citações dessa época, tratam de modelos somente como técnicas estatísticas ou econométricas, porém, com as recentes evoluções do processamento

computacional e armazenamento de dados, popularizou-se o uso de modelos em diversas frentes, para além da gestão de risco, podendo-se relacionar o gerenciamento do risco de modelo a Inteligência Artificial [29] ou *Big Data* [30].

Em documento sobre gerenciamento do risco de modelo [31], o *Federal Reserve System* (FED), define modelo como: um método, sistema ou abordagem quantitativa que aplica teorias, técnicas e suposições estatísticas, econômicas, financeiras ou matemáticas para processar dados de entrada em estimativas quantitativas. Embora outras definições possam ser encontradas na literatura, este conceito será utilizado para este trabalho.

Conforme descrito pela *American Academy of Actuaries* [32], Risco de modelo pode ser definido como a possibilidade de perdas (econômicas, reputacionais, etc) decorrentes de falhas ou mal uso dos modelos. Na definição apresentada por Tunaru [33], o risco de modelo é "a incapacidade da estrutura estatística matemática proposta de refletir de maneira homogênea ao longo do tempo o objeto em análise", conceito que evidencia a necessidade de a estrutura de gerenciamento de risco de modelos ser um processo contínuo, não apenas premissas e regras na aprovação de uso de modelos.

### 2.3.1 Contexto Brasileiro

No contexto brasileiro, a Resolução CMN N<sup>o</sup> 4.557 [9], de 23 de fevereiro de 2017, trata das estruturas necessárias ao gerenciamento de riscos e capital nas instituições financeiras. Embora não haja, neste documento, menção específica do termo riscos de modelo, estabelece-se que:

“Os modelos para o gerenciamento de riscos, quando utilizados e relevantes, devem ser submetidos a avaliação periódica quanto: a) à adequação das metodologias utilizadas; e b) ao desempenho, incluindo a comparação entre perdas estimadas e observadas (*back-testing*).”

Dessa forma, mesmo sem a inclusão do risco de modelo como um dos riscos obrigatórios, a serem reconhecidos como risco relevante para as instituições financeiras, dos cinco maiores bancos brasileiros, quatro deles (Itaú Unibanco [34], Banco do Brasil [35], Bradesco [36] e Santander [37]) declararam o risco de modelo como um risco relevante em seus Relatórios de Gerenciamento de Risco dos anos de 2018 e 2019. No caso da Caixa Econômica Federal, mesmo sem declaração do risco de modelo como um risco relevante, foi especificada governança de gerenciamento dos modelos semelhantes aos demais bancos, com um inventário corporativo dos modelos utilizados, validação independente e monitoramento contínuo dos resultados do modelo. Neste contexto, ganha destaque a necessidade de priorização de modelos na realização dessas atividades, baseado em sua relevância e na mensuração do risco associado a cada modelo.

Este trabalho relaciona este contexto do gerenciamento do risco de modelo ao risco de crédito, indissociável das atividades das principais IFs do país e tradicionalmente mensurado pelo uso de dados históricos e modelos, considerando as demonstrações financeiras em IFRS 9. O capítulo seguinte descreve a forma de estruturação deste trabalho para relacionamento destes conceitos.



# Capítulo 3

## Metodologia

De modo geral, este trabalho se utiliza de abordagens quantitativas e qualitativas para solução de problemas reais, aplicáveis às IFs, no contexto do gerenciamento dos riscos associados à migração do procedimento de cálculo das PCLD do padrão da Resolução CMN N<sup>o</sup> 2.682 [10], utilizado desde 1999, para a Resolução CMN N<sup>o</sup> 4.966, cuja implementação deve ser concluída até 2025 [18]. Dessa forma, esta é uma pesquisa exploratória de natureza aplicada[38], tendo em vista que está centrado em proporcionar uma visão geral, aproximada acerca deste fato relevante para o funcionamento do mercado de crédito brasileiro.

O trabalho se divide em três partes principais, sendo elas: identificação e discussões sobre as fontes de risco de modelo no padrão IFRS 9 - *impairmant*; abordagem para priorização para gerenciamento dos riscos identificados; e, por fim, a associação dos dois conceitos para gerenciamento do risco de modelos associado aos modelos do IFRS 9. A Figura 3.1 ilustra as etapas seguidas para realização dessa pesquisa.

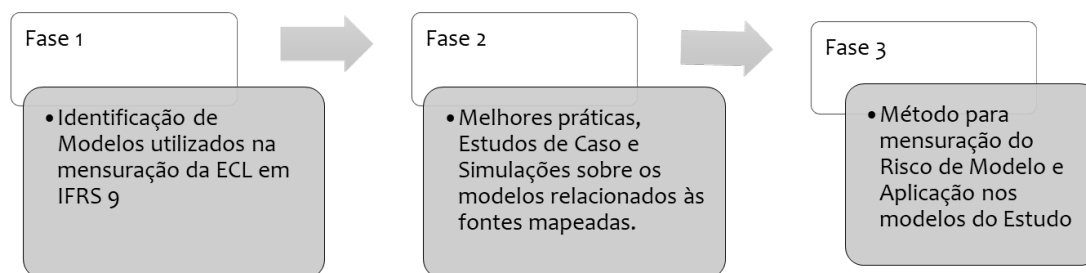


Figura 3.1: Fluxo Metodológico (Fonte: Elaborado Pelo Autor).

Em linhas gerais, a fase 2 da pesquisa trata sobre uma série de questões a respeito

da implementação do padrão IFRS 9 no Brasil. A discussão sobre estes pontos ganha relevância dada a iminência de implementação dos modelos a eles associados e da própria alteração da classificação e mensuração do risco de crédito dos ativos financeiros.

Este capítulo apresenta uma visão geral da pesquisa aplicada neste trabalho. No entanto, para cada uma dessas questões estudadas, há um capítulo que pode ser considerado como uma pesquisa individual, embora relacionada com as demais. Dessa forma, os objetivos específicos de cada pesquisa, e os procedimentos metodológicos, bem como a discussão sobre trabalhos correlatos serão apresentados nestes capítulos.

Os aspectos estudados nos capítulos seguintes derivam dos modelos identificados como necessários para implementação do padrão IFRS 9, tema tratado na seção seguinte.

### 3.1 Fontes de Risco de Modelo no Padrão IFRS 9

Conforme descrito no padrão IFRS 9 [12], as instituições devem reconhecer provisão para perdas esperadas de crédito em ativos financeiros em que os seus resultados financeiros dependam do recebimento de pagamentos de principal e juros em fluxos de caixas contratuais, o que inclui as carteiras de crédito. Neste contexto, o padrão especifica que, em cada data de reporte, as IFs devem aplicar os requerimentos de depreciação dos ativos ou redução do valor recuperável (*impairment*) para o reconhecimento e mensuração de uma provisão para perdas para ativos financeiros.

A tarefa de mensurar a perda esperada de crédito nasce da incerteza sobre os recebimentos dos pagamentos acordados nos contratos dos créditos. A incorporação dessas incertezas no cálculo pode ser feita de diferentes métodos pelas IFs, mas vale destacar que o conceito esta aplicação de métodos, a partir de dados de entrada, para se calcular estimativas quantitativas está perfeitamente alinhada ao conceito de modelo apresentado na seção anterior deste trabalho.

Além do conceito direto do cálculo da perda esperada, outros requerimentos de *impairment* do padrão têm relação com o risco de modelo. Como será apresentado em mais detalhes nos capítulos seguintes, esses requerimentos, geralmente, não são administrados como um modelo único de mensuração da ECL, mas sim como uma quantidade de mais abrangente de métodos associadas a eles. Dessa forma, foi feita uma análise de documento no padrão IFRS 9 [12] para identificação das fontes de risco de modelo nestes requerimentos. A Tabela 3.1 apresenta as fontes de risco de modelo identificadas nos requerimentos do cálculo da ECL no padrão IFRS 9.

Destaca-se que a publicação da Resolução CMN Nº 4.966 [18], em 25 de novembro de 2021, reforça o estabelecimento destes modelos, ao citar mais diretamente a necessidade de se calcular a ECL com base em um valor justo do crédito, da probabilidade do instrumento

Tabela 3.1: Fontes de Risco de Modelo (IFRS 9)

Fonte de Risco	Descrição	Parágrafos
Risco de inadimplência	A própria definição de ECL apresentada no padrão a descreve como a média das perdas ponderadas pelo risco de inadimplência.	Appendix A
Expectativa de recuperação e <i>write-off</i>	Se a IF não tiver perspectivas razoáveis de recuperar quaisquer fluxos de caixa adicionais do ativo financeiro, ela deve dar baixa do restante do ativo financeiro.	B5.4.9
ECL <i>lifetime</i> e <i>12-month</i> ECL	A IF deve estimar a ECL <i>lifetime</i> ou para o período de 12 meses para créditos, respectivamente, com ou sem aumentos significativos de risco.	5.5.3 - 5.5.5
Classificação dos Ativos	A definição de aumento significativo de risco prevê uma série de fatores a serem considerados para definir os aumentos significativos de risco, bem como o próprio conceito de inadimplência ( <i>default</i> ).	5.5.9 - 5.5.11
Cenários <i>forward-looking</i>	Na estimativa da ECL deve-se considerar também as informações macroeconômicas e como mudanças relevantes no cenário influenciam o risco de crédito.	B5.5.4 e B5.5.14

de crédito se tornar um ativo problemático e da expectativa de recuperação entre os ativos problemáticos. Para cada um dos modelos associados à esta implementação, serão apresentados estudos de caso, utilizando dados de operações de crédito e simulações, quando aplicável.

Além disso, será aplicada a metodologia de para priorização dos modelos de acordo com o risco de modelo a eles associados. Cada uma das subseções abaixo descreve brevemente a metodologia a ser aplicada para cada fonte de risco, sendo que, para cada uma delas, haverá um capítulo dedicada aos resultados obtidos.

### 3.1.1 Ajuste da ECL ao Prazo

Sobre os parâmetros de risco, vale destacar que as estimativas de PD consideram uma estimativa pelo período de 12 meses de acompanhamento dos portfólios de crédito. Isso significa que nos ativos classificados como aumento significativo de risco, em que há necessidade de acompanhamento por todo o período da operação (*lifetime*), o simples uso

do parâmetro dessa forma não é suficiente. Estimar a PD para diferentes períodos é a tarefa mais importante em relação à PD no contexto do IFRS 9 [39].

Além de modelos paramétricos e semi-paramétricos mais tradicionais para análise de sobrevivência utilizados nos trabalhos citados acima, o uso de modelos baseados em árvore é uma alternativa que envolve a segmentação do espaço das covariáveis em uma série de regiões mais simples, com o objetivo de fazer uma previsão para uma determinada observação. Esses modelos fornecem uma alternativa aos modelos lineares e aditivos, problemas de regressão linear e logística para problemas de classificação.

As mais conhecidas Árvores de Decisão podem apresentar vantagens quando comparadas a outros modelos de aprendizado de máquina para tarefas de classificação, como (i) a facilidade de explicação da relação entre variáveis independentes e dependentes, (ii) o processo lógico de criação dos nós que são mais semelhantes ao processo de tomada de decisão humana, (iii) as regras de classificação intuitivas transmitidas em figuras que representam as árvores de decisão etc.

Este contexto de modelos baseados em árvore também é aplicável a problemas Análise de Sobrevivência, permitindo a investigação de modelos de PD ao longo da vida da linha de crédito, conforme exigido pela regulamentação. Será apresentada a investigação de as árvores de sobrevivência e as árvores de sobrevivência aleatórias.

O Capítulo 4 aplicará um estudo de caso para modelagem da PD *lifetime* utilizando modelos de análise de sobrevivência baseados em árvores utilizando dados de uma financeira norte americana.

### 3.1.2 Ajuste da ECL a Cenários *Forward-Looking*

Outra necessidade associada às estimativas do IFRS 9 é a utilização de informações prospectivas, inclusive macroeconômicas e sua relação com o risco de crédito esperado para os próximos 12 meses. Mesmo antes da crise de 2007-2008 e das consequentes melhorias incorporadas aos mecanismos de gerenciamento de risco e quantificação do risco de crédito, a modelagem do *Non-Performing Loans* (NPL) já se mostrava uma forma de compreender o mercado de crédito. Chang [40], por exemplo, realizou um estudo sobre a relação entre o NPL brasileiro e a concentração de crédito. Usando uma abordagem por dados em painel, os autores identificaram um efeito significativo da concentração bancária no NPL.

Os dados utilizados neste trabalho foram obtidos a partir do Sistema Gerenciador de Séries Temporais [41], onde são disponibilizadas informações agregadas sobre o crédito no Brasil, além de informações sobre a atividade econômica, de política monetária, entre outros.

NPL utilizadas neste trabalho são compostas pelo saldo total de créditos em inadimplência superior a 90 dias sobre o total de crédito. As séries de NPL estão disponíveis,

mensalmente, para o período de março de 2011 a fevereiro de 2021. Já os dados de Produto Interno Bruto (PIB), presente na maior parte dos trabalhos de modelagem de NPL, são apresentados em periodicidade trimestral. Neste sentido, as estimações dos modelos e as projeções utilizaram bases trimestrais contemplando o período entre 2011Q2 e 2019Q4, para modelagem econométrica e de 2020Q1 a 2020Q4 para as projeções.

A partir do resultado do modelo desenvolvido, outro ponto de influência sobre o risco de modelo envolvido nestes cálculos é o efeito dos múltiplos cenários econômicos possíveis a serem utilizados. Apesar dos modelos serem construídos com informações observadas e prospectivas em comparação às carteiras de crédito, o momento do uso do modelo tem somente projeções econômicas disponíveis para os próximos 12 meses.

O Capítulo 5 apresentará estudos de caso com utilização de modelos para estimação dos efeitos de variáveis macroeconômicas na NPL. Além disso, será avaliado o efeito do uso de covariáveis projetadas no momento da predição, nos erros das estimativas.

### 3.1.3 Estimativa da Perda Esperada

A alteração do cálculo de provisão a partir dos níveis de risco da Resolução CMN Nº 2.682 [10] para a aplicação de grupos homogêneos de risco, para os quais devem ser calculados e acompanhados parâmetros de probabilidades de inadimplência e expectativa de recuperação [18] são as principais alterações em relação ao motor de cálculo das provisões de crédito. Estas alterações trazem impacto direto nos resultados e na alocação de capital das IFs. Diante dessas alterações, esta etapa do trabalho utiliza dados simulados para avaliação da dinâmica do cálculo e os impactos observados no resultado a partir delas.

A dinâmica dos atrasos dos créditos foram baseadas em estudo anterior aplicando a demonstrações financeiras em *Fintechs* [42], em que foi avaliado os impactos da PCLD nas demonstrações contábeis. De forma semelhante, este trabalho: mapeia a dinâmica das carteiras de crédito, implementa os cálculos dos parâmetros de risco e aplica as diferentes abordagens para cálculo das provisões.

### Dados Simulados

Na literatura sobre risco de crédito ou risco de modelo, o uso de simulações é comum para se ter maior controle sobre os variados cenários em que determinada abordagem pode ser aplicada. O BCBS, no *Working Paper* no 14 [4], apresenta um estudo em que são aplicadas simulações para gerar taxas de descumprimento anuais próximas à realidade, para comparar a efetividade de dois testes nas conclusões sobre o desempenho de modelos de PD.

Seitshiro [43] utiliza simulações de uma carteira de crédito em um trabalho para quantificar o risco de modelo ocasionado pela especificação errada de modelos de PD. Schwab [44] apresenta um estudo de caso, aplicado a empresas suíças, para explorar o impacto do setor financeiro sobre o crescimento de empresas de médio e pequeno porte. A partir da combinação de modelos *system dynamics* e *agent-based*, para descrever, respectivamente, comportamentos do sistema e dos agentes individuais, os autores concluem que os limites concedidos pelos bancos e as regulações bancárias na relação entre essas empresas e os bancos.

Para estudar as medidas de inadimplência do setor bancário brasileiro, um trabalho para discussões do BCB [45] utilizou e descreveu uma rotina de simulação de um sistema de contratação e atualização em uma carteira de crédito no mercado brasileiro.

O trabalho apresentado no Capítulo 6 utilizará as regras já mapeadas no exercício baseado em dinâmica de sistemas e simulações de Monte-Carlo para gerar banco de dados que seja suscetível a erros de modelos. Um estudo comparativo entre a abordagem vigente e a nova abordagem será apresentado de modo a verificar sua resposta a mudanças na carteira de crédito.

## 3.2 Mensuração do Risco de Modelo

Com essa crescente necessidade em se extrair cada vez mais valor a partir dos dados, além do desenvolvimento e certa popularização das técnicas em Ciência de Dados (*Machine Learning*, *Analytics*, Inteligência Artificial, dentre outras), muitas empresas têm implementado mudanças em sua estrutura com o objetivo de estabelecer governança de TI para este tipo de atividade. Os investimentos e movimentações declarados pelas IFs para os seus investidores indicam os temas considerados relevantes no setor.

O Banco do Brasil [46], por exemplo, anunciou ao mercado, em julho de 2019, a criação da Unidade de Inteligência Analítica, que foi criada com a missão de centralizar o acompanhamento das evoluções, inovações e desenvolvimento de técnicas, ferramentas e soluções com uso de Inteligência Analítica e de Inteligência Artificial na instituição. O Itaú Unibanco [34] criou o Centro de Excelência em *analytics* e o duplicou no ano seguinte a sua criação, com isso mais de 5 milhões de clientes foram impactados por projetos de *analytics* nessa época. Em 2018, o Bradesco informou que a plataforma de inteligência artificial chamada de Bia, desenvolvida em parceria com a IBM, em 12 meses de atendimento a correntistas já havia realizado mais de 22 milhões de interações com clientes [47].

Estes investimentos divulgados nos últimos anos pelas instituições financeiras mostram a relevância das atividades relacionadas ao desenvolvimento e uso de modelos para o

sucesso destas instituições. Como pode ser visto, foi intensificada a aplicação de recursos em tecnologias, treinamentos e contratação de profissionais qualificados para atividades deste contexto.

### 3.2.1 Por que Priorizar Modelos?

A condução dessas atividades nas IFs requer uma série de atividades por seus colaboradores e, dada a escassez dos recursos a serem alocados nestas atividades, deve ser tomada a decisão pela instituição sobre quais modelos devem ser priorizados na alocação destes recursos. Trazendo uma melhor percepção sobre os esforços relacionados à gestão dos modelos a Figura 3.2 ilustra a relação entre o ciclo de vida de desenvolvimento de *software* e um conceituado ciclo de vida de modelos analíticos [48] [49].

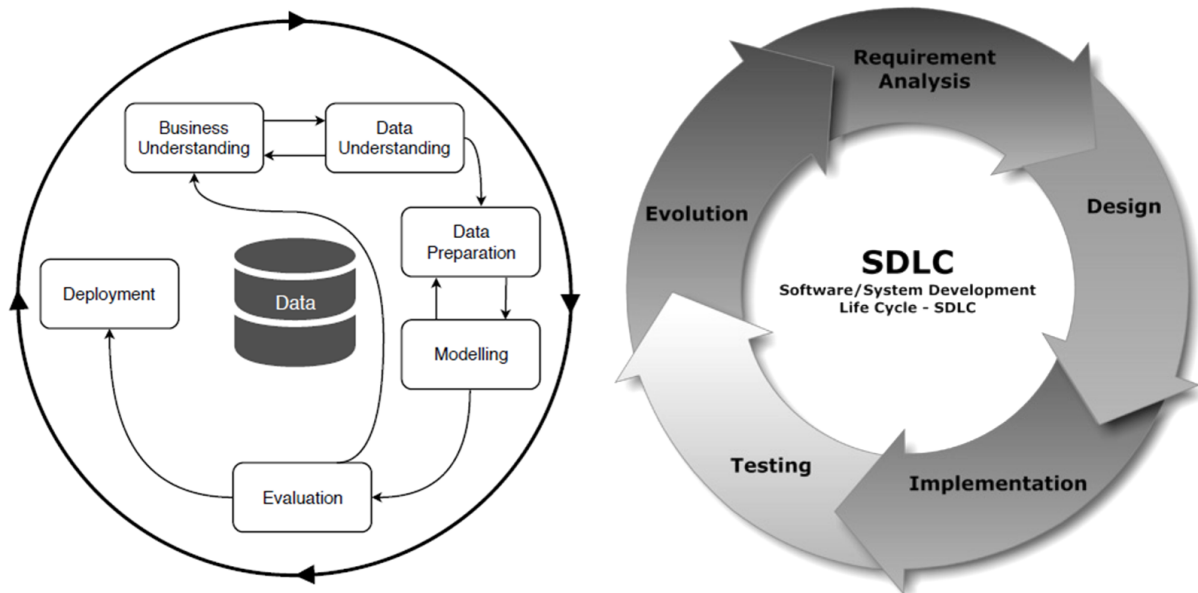


Figura 3.2: CRISP-DM e Ciclo de Vida de Desenvolvimento de Software (Fonte: Plumed [48], Lenhardt [49]).

Vários fatores podem ser considerados na decisão da instituição de priorizar a manutenção ou o desenvolvimento de um modelo analítico em detrimento de outros. Alinhamento à sua estratégia, requisitos regulatórios são exemplos. Diante do exposto nos capítulos anteriores, fica evidente a possibilidade de que os modelos implementados ocasionem perdas financeiras quando utilizados incorretamente, na concretização do evento de risco de modelo.

Pelo quadro apresentado, pode-se notar que uma instituição pode, por exemplo, ter a necessidade de desenvolver uma nova solução em inteligência artificial para melhoria da adequação do acesso ao crédito para seus clientes e, no mesmo momento, estar exposto

um elevado risco de modelo em suas estimativas de ECL das demonstrações financeiras em IFRS 9. Nessa situação, os processos podem concorrer por alocação de recursos de Tecnologia da Informação (TI) ou de cientistas de dados, trazendo a necessidade de priorização na alocação dos recursos pela instituição.

### 3.2.2 Mensuração do Risco de Modelo

Uma estrutura de gerenciamento de risco de modelo avançada contém, além dos aspectos tratados nesta seção, a existência de uma unidade de quantificação do risco de modelo [50]. No contexto da mensuração do risco de modelo, segundo o *American Academy of Actuaries* [32], a forma de mensuração do risco de modelo adotada por instituição pode ser qualitativa ou quantitativa, sendo descritas as seguintes abordagens:

**Scoring Model Risk:** Abordagem qualitativa e subjetiva para mensuração do risco de modelo;

**Quantificação do Risco de Modelo:** Geralmente, a mensuração quantitativa é feita como uma fração do risco operacional, desenvolvendo estimativas da frequência e severidade dos eventos de risco de modelo.

Embora haja metodologias quantitativas de mensuração do risco de modelo, nem sempre é possível quantificar matematicamente os riscos associados ao uso do modelo. Por este motivo, muitas vezes pode ser necessário optar por uma mensuração qualitativa do risco de modelos. Uma metodologia amplamente utilizada para processos de tomada de decisão subjetivos, conforme o caso dos *Scoring Model Risk* é a Análise de Decisão Multicritério.

No Capítulo 7, foram levantados os fatores e critérios a partir da literatura e trabalhos de consultorias, para priorização de modelos a partir da mensuração do risco de modelo. Além disso, como estudo de caso, foi aplicada a abordagem proposta sobre os modelos utilizados no cálculo da ECL no contexto da implementação da Resolução CMN 4.966 [18].

Considerando a utilização de todos estes requisitos quantitativos para mensuração da ECL no IFRS 9, os resultados obtidos em todos os capítulos se relacionam por serem utilizados para o mesmo fim. Nesta última aplicação do Capítulo 7, os resultados dos modelos obtidos nos demais capítulos são sintetizados numa mensuração do risco de modelo associado ao uso destes.



# Capítulo 4

## Ajuste Lifetime

### 4.1 Introduction

Ao divulgar o Acordo de Basileia II, o BCBS apresentou uma abordagem para as instituições financeiras medirem o capital necessário para fazer frente às perdas de risco de crédito por meio de parâmetros de risco estimados internamente. A partir desses parâmetros, as instituições financeiras podem calcular o valor da ECL, definida de acordo com a Equação 2.1[51].

Os modelos de PD visam identificar a probabilidade de um cliente ou contraparte em determinada operação não cumprir as cláusulas de um contrato de crédito. Os bancos constroem modelos de PD usando dados históricos sobre operações de crédito, características pessoais e comportamento de seus clientes. Em alguns modelos, as variáveis macroeconômicas também são usadas para modelagem de risco de crédito e podem trazer impacto significativo para os modelos, conforme descrito por [52].

As diretrizes do Basileia II sugerem amplamente que as exposições ao risco devem ser tanto tomadoras de risco quanto orientadas para transações e devem ser agrupadas em grupos de risco homogêneos [51], estabelecimento que é corroborado pelo CMN [18]. Seguindo essas diretrizes, as instituições financeiras aplicaram diferentes técnicas de modelagem para prever a PD.

A Resolução CMN N<sup>o</sup> 4.966 [18] estabelece que os ativos financeiros devem ser classificados quanto ao risco de crédito, em três estágios, sendo eles:

- Primeiro Estágio: Para ativos financeiros que não tenham apresentado aumento significativo de risco, deve ser calculada a perda esperada em 12 meses;
- Segundo Estágio: Para ativos financeiros que tenham apresentado aumento significativo no risco de se tornar um ativo problemático, em que a perda esperada deve ser calculada para todo o prazo do ativo *lifetime*; e

- Terceiro Estágio: Para ativos com problemas de recuperação de crédito ou inaplicáveis.

Quanto às operações do segundo estágio, tendo em vista a necessidade de ajustar as probabilidades de inadimplência ao período total da operação, as técnicas utilizadas para a sua estimativa podem necessitar de ajustes, em relação às técnicas de período fixo em 12 meses. Neste contexto, Vanek [39] destaca a necessidade de se estabelecer estimativas com múltiplos períodos, tendo em vista a necessidade de estimação para todo o prazo esperado do instrumento financeiro. Este autor elenca três metodologias relativamente populares encontradas na literatura para modelagem da PD, sendo elas os modelos Merton-*Type*, modelos de análise de sobrevivência e modelos de Markov.

Em novembro de 2021, o *European Banking Authority* (EBA) abriu um *paper* para discussão sobre o uso de modelos de *machine learning* para os modelos de *Internal Ratings-Based* (IRB), cujas metodologias são aproveitadas para os cálculos da metodologia de cálculo do IFRS 9. Estas discussões demonstram a necessidade de debate constante no meio acadêmico para subsidiar as melhores decisões nas instituições. Neste contexto, o trabalho deste capítulo tem o objetivo de avaliar o desempenho dos modelos de análise de sobrevivência baseados em árvore, em comparação aos modelos tradicionais de análise de sobrevivência e aos modelos de classificação para estimativa da PD. Espera-se contribuir para as discussões recentes sobre o uso de métodos de *machine learning* no risco de crédito e sobre o desempenho dos modelos de sobrevivência no contexto do IFRS 9.

Este capítulo está estruturado da seguinte forma. Na próxima seção, é apresentado o referencial teórico sobre o tema do estudo, destacando algumas das contribuições acadêmicas sobre o contexto deste trabalho. Em seguida, descreve-se os passos metodológicos seguidos para realização desta pesquisa. Por fim, são apresentados os resultados obtidos e as conclusões, destacando a contribuição e limitações do trabalho para a literatura acadêmica e para a prática de gerenciamento de risco de crédito.

## 4.2 Referencial Teórico

### 4.2.1 Mineração de Dados e Modelos de PD

A mineração de dados é definida como um processo de descoberta de padrões em data [53]. Técnicas de mineração de dados são usadas para resolver vários problemas que envolvem o uso de dados para padrões de descoberta. Problemas específicos relacionados à modelagem de PD representam desafios relevantes no setor bancário e podem ser investigados com o uso de mineração de dados.

Yeh [54] compara diferentes técnicas de mineração de dados para estimar a probabilidade de descumprimento, modelando o problema com o *default* como variável resposta binária. Neste trabalho, os autores comparam o desempenho de modelos de classificação como *K-Nearest Neighbors*, Regressão Logística, Análise Discriminante, *Naïve Bayes*, Redes Neurais Artificiais e Árvores de Decisão. Os resultados do estudo indicam que o modelo de Redes Neurais apresenta a melhores resultados do que as outras técnicas, incluindo as abordagens estatísticas tradicionais baseadas em Regressão Linear e Análise Discriminante. Sobre o trabalho destes autores, todos os métodos comparados são modelos utilizados para tarefas de classificação.

Para lidar com as limitações dos modelos de crédito tradicionais, que geralmente se concentram em uma PD em um determinado período, um mecanismo alternativo para estimar a PD *lifetime* é aplicar a Análise de Sobrevivência. Mais particularmente, o Análise de Sobrevivência leva em consideração o tempo até a ocorrência de um evento de interesse. No âmbito das operações de crédito, o caso de juros pode ser o default. Portanto, estas técnicas são úteis para analisar PD ao longo do prazo total da exposição de crédito [55].

## 4.2.2 Análise de Sobrevivência

Os métodos de Análise de Sobrevivência exploram o tempo até um evento em uma determinada população e, em comparação com outros modelos de classificação, como Análise Discriminante e Regressão Logística, adiciona o recurso de avaliação de probabilidade ao longo do tempo [55].

Portanto, a Análise de Sobrevivência pode ser usada para analisar se o tomador ficará inadimplente ou não, bem como a mudança na taxa de ocorrência da inadimplência até um determinado momento  $t$ :

$$h(t) = \lim_{\delta_t \rightarrow 0} \frac{(P(t \leq T < t + \delta_t | T \geq t))}{\delta_t} \quad (4.1)$$

onde  $T$  é uma variável aleatória associada à função de sobrevivência, especificada na equação 4.2.

$$S(t) = P(T \geq t) \quad (4.2)$$

Aplicando Análise de Sobrevivência ao crédito, Bellotti [55] realizou um estudo no qual usou o modelo Cox *Proportional-Hazards* (PH) para conduzir uma Análise de Sobrevivência usando uma amostra de transações de cartão de crédito fornecidas por um banco do Reino Unido. A técnica segue um modelo semi-paramétrico, onde o efeito das covariáveis é usado para ajustar uma distribuição de probabilidade desconhecida a uma distribuição

de probabilidade conhecida. Os resultados mostraram que a Análise de Sobrevida tem um desempenho competitivo em relação à Regressão Logística, com a vantagem de permitir estimar uma PD *lifetime*.

Adicionalmente, Durović [56] realizou um estudo sobre modelagem de PD, levando em consideração a implementação das diretrizes do IFRS 9. No estudo, os autores concluíram que a utilização da Análise de Sobrevida pode ser relevante para estimar a PD, considerando a estimativa obrigatória da exposição durante todo o período de vida da linha de crédito.

### 4.2.3 Modelo Cox *Proportional-Hazards*

O modelo de riscos proporcionais de Cox PH [57] é uma das abordagens mais tradicionais baseadas em técnicas de tempo até o evento. Ele assume que as taxas de falha de dois grupos são constantes, ou seja, seguem funções proporcionais.

Sejam  $p$  covariáveis e um vetor  $x = (x_1, \dots, x_p)$ , a função geral do modelo de Cox PH é dada conforme Equação 4.3:

$$\lambda(t) = \lambda_0(t)g(X\beta) \quad (4.3)$$

onde  $g(\cdot)$  é uma função não negativa.

Este modelo é composto por dois componentes: um  $\lambda_0(t)$  não paramétrico, que não é especificado; e um componente paramétrico que geralmente é usado como 4.4:

$$g(X\beta) = \exp(X\beta) = \exp(\beta_1x_1 + \dots + \beta_px_p) \quad (4.4)$$

onde  $\beta$  é um vetor  $p \times 1$  de parâmetros para cada covariável e a estimativa dos parâmetros é dada por máxima verossimilhança parcial [58]

### 4.2.4 Modelos Baseados em Árvore

Os modelos baseados em árvore [59] envolvem a segmentação do espaço das covariáveis em uma série de regiões mais simples com o objetivo de fazer uma previsão para uma dada observação. Esses modelos fornecem uma alternativa aos modelos lineares e aditivos, para problemas de classificação ou regressão. Nesse contexto, os modelos baseados em árvore são adequados para problemas de classificação e regressão em que existe um conjunto de variáveis explicativas e uma variável de resposta única.

As Árvores de Decisão apresentam muitas vantagens quando comparadas a outros modelos tradicionais de classificação e regressão, tais como (i) a facilidade de explicação da relação entre variáveis independentes e dependentes, (ii) o processo lógico para a criação

de nós que são mais semelhante ao processo de tomada de decisão humana, (iii) as regras de classificação intuitivas transmitidas em figuras que representam as árvores de decisão etc.

A intuição de Árvore de Decisão pode ser aplicada em um contexto de Análise de Sobrevivência, permitindo a investigação de modelos de PD ao longo da vida da linha de crédito, conforme exigido pelo regulamento. Neste estudo, investigamos árvores de sobrevivência e *Random Survival Forest* (RSF), que são técnicas de mineração de dados baseadas em aprendizado de máquina que são extensões de árvores de decisão.

### 4.2.5 *Survival Tree*

A *Survival Tree* é um método baseado em árvore em que uma regra de divisão é usada para agrupar indivíduos ou observações de suas covariáveis. Cada grupo é selecionado com base em seu comportamento de sobrevivência [60].

A *Survival Tree* tem sido aplicado em várias áreas. Por exemplo, [61] usou a técnica da *Survival Tree* para identificar os fatores de risco no diagnóstico de crianças com nefroblastoma. No estudo, foram identificadas características que influenciam o tempo de recidiva, malignidade ou morte do paciente e foi criada uma hierarquia de risco que permite a indicação de diferentes tratamentos dadas determinadas características.

Neste trabalho, foi construída a *Survival Tree* usando um critério de divisão com base na máxima verossimilhança, onde o ponto de corte escolhido maximiza a função de log-verossimilhança observada, descrita na equação 4.5 [60]:

$$ll(t) = \sum_{j=1}^K (n_{t1}(j) \ln(\hat{\pi}(j)) + n_{t0}(j) \ln(\hat{S}(j))) \quad (4.5)$$

onde  $\hat{\pi}(j)$  and  $\hat{S}(j)$  são os estimadores de máxima verossimilhança dos parâmetros  $\pi(j)$  e  $S(j)$ , que representam a probabilidade de eventos e as probabilidades de sobrevivência, respectivamente.

### 4.2.6 *Random Survival Forest*

A RSF, proposta por [62], usa métodos de *ensemble* em árvores de sobrevivência para obter a *Cumulative Hazard Function* (CHF). O RSF aproveita a randomização de duas maneiras; (i) desenhando aleatoriamente  $B$  amostras de *bootstrap* para cada árvore, e em cada nó de uma árvore, e (ii) selecionando aleatoriamente um subconjunto de variáveis para dividir. Portanto, o nó dividido é escolhido usando a variável candidata selecionada que maximiza as diferenças de sobrevivência nos nós filhos.

Ishwaran [62] apresenta os seguintes passos para estruturação de uma RSF:

1. Selecionar  $B$  amostras de *bootstrap*;
2. Construir uma *Survival Tree* para cada amostra. Selecione aleatoriamente  $p$  variáveis em cada nó. O ponto de corte maximiza a diferença de sobrevivência entre os nós filhos.
3. Aumente a árvore até que as restrições não sejam violadas.
4. Calcular o CHF para cada árvore e faça a média deles para obter o CHF do conjunto.
5. Com os dados fora da amostra, calcular o erro de predição para o conjunto CHF.

Bellini [63] considerou a RSF uma alternativa para modelar o tempo da vida de PD, assim como outras técnicas de análise de sobrevivência e aprendizado de máquina.

#### 4.2.7 Trabalhos Correlatos

Mageto [64] realizou um estudo comparativo entre a RSF e o modelo de Cox-PH na modelagem do risco de crédito, nesse trabalho, os autores verificaram melhor desempenho do modelo Cox-PH em relação ao RSF. Em estudo com objetivo semelhante, Frydman [65] verificou que o RSF é uma boa alternativa aos modelos tradicionais de análise de sobrevivência, tendo sido verificados melhor acurácia em dados de créditos para veículos.

Outros estudos [66][67] utilizaram variações de modelos baseados em árvores para estimação da PD, apresentando resultados melhores que os métodos mais comuns de Análise de Sobrevivência. Tais resultados apresentados em estudos recentes destacam a relevância do tema apresentado neste trabalho. Neste caso, analisando uma aplicação destes métodos para dados de refinanciamento de cartão de crédito.

### 4.3 Metodologia

Neste trabalho, apresenta-se um estudo de caso em uma pesquisa do tipo exploratória, com abordagem quantitativa por meio de modelos de análise de sobrevivência aplicados para modelagem da PD em dados de refinanciamento de cartão de crédito.

Neste trabalho, foram utilizadas técnicas que fornecem uma estimativa da PD ao longo do tempo. Considerando a vantagem dos modelos de *machine learning* em relação aos métodos tradicionais observada em pesquisas anteriores [54][65][66][67], este estudo enfoca o uso de métodos estatísticos não tradicionais, explorando diferentes técnicas de análise de sobrevivência. Em especial, além da análise de sobrevivência usual com base no modelo de risco proporcional de Cox, foram aplicados o *Survival Tree* e RSF para investigar a PD durante o tempo de vida de uma transação de crédito.

Foram utilizados dados de operações de refinanciamento de cartão de crédito de uma instituição financeira dos Estados Unidos da América. Mais especificamente, foi investigada a PD no refinanciamento de tomadores sobre suas dívidas de cartão de crédito, que pode ser considerada como um ativo com problemas de recuperação, dependendo das regras estipuladas na instituição financeira.

Os dados referem-se a operações contratadas no período de janeiro de 2014 a dezembro de 2018, além do seu histórico de pagamentos nas datas posteriores. Cada observação no banco de dados é um contrato de refinanciamento de cartão de crédito de um tomador com informações históricas. Foram desenvolvidos modelos com as 55.522 operações contratadas em 2014, sendo 40.100 operações com vencimento em 36 meses e 15.422 operações com prazo de 60 meses. O conjunto de dados foi dividido em 70% para treinamento e 30% para teste. A taxa de inadimplência dessa amostra foi de 14,96%. Os modelos foram validados em conjunto de dados fora da amostra e fora do tempo, em outra amostra com 102.025 operações de refinanciamento de cartão de crédito contratadas em 2015.

O banco de dados original incluía observações com 72 atributos, incluindo informações de contratos e tomadores de empréstimos, tais como: características das operações e dos indivíduos, histórico de pagamentos, situação de crédito etc. Algumas variáveis foram excluídas devido a (i) alta correlação com outras, (ii) pouco contraste entre seus níveis, (iii) muitos valores nulos, (iv) textos não informativos (como e-mails). Por fim, foram mantidas 25 variáveis explicativas e duas variáveis resposta: tempo e *default*.

O indicador variável de default foi construído considerando o somatório de 90 dias contados da data do último pagamento das operações inadimplentes. Os dados utilizados continham apenas a posição final de uma carteira de operações de crédito, com as datas dos últimos pagamentos. Dessa forma, a elaboração da data do descumprimento foi feita como a data do último pagamento, acrescida de 90 dias, quando atendeu ao conceito de ativo com problema de recuperação de crédito (Res. CMN 4.966 [18]). Por exemplo, se uma operação não liquidada na posição final, teve seu último pagamento em 01.06.2015, a data do default calculada seria em 31.08.2015.

Para as técnicas de análise de sobrevivência, este trabalho baseou-se nas seguintes premissas: (i) o evento de interesse na análise de sobrevivência é o *default*, (ii) as operações não *default* são consideradas censuradas à direita, (iii) todas as árvores foram controladas pelo percentual do total de operações nas quebras dos nós.

O banco de dados original incluía observações com 72 atributos, incluindo informações de contratos e tomadores de empréstimos, tais como: características das operações e dos indivíduos, histórico de pagamentos, situação de crédito etc. Algumas variáveis foram excluídas devido a (i) alta correlação com outras, (ii) pouco contraste entre seus níveis, (iii) muitos valores ausentes, (iv) textos não informativos (como e-mails). Por fim, foram

mantidas 25 variáveis explicativas e duas variáveis resposta: tempo e *default*.

Após a limpeza do banco de dados original, foi construído conjunto de dados contendo informações relevantes que permitem estudar a PD ao longo do tempo das operações de crédito. Foi utilizado o R Core Team [68], com o pacote **survival** [69], para o modelo Cox-PH e com os pacotes R **DStree** [70] e **randomForestSRC** [71] para os algoritmos *Survival Tree* e RSF, respectivamente.

## 4.4 Resultados

Várias configurações de técnicas foram usadas. Na busca por um modelo adequado e parcimonioso, primeiramente, foi processada uma *Survival Tree* para identificar covariáveis com medida de maior importância para selecionar as características para o modelo final. Cada quebra foi definida para conter pelo menos 1%, 5% ou 10% das operações, para mitigar problemas de *overfitting*. Todos os modelos levaram em consideração variáveis consideradas relevantes em função de sua relação econômica com a PD: prazo da operação, taxa de juros do contrato, grau de risco do tomador e receita anual do tomador.

Os resultados avaliados mostraram que as classificações de importância são as mesmas para os dois modelos (*Survival Tree* e RSF). A variável resposta, ou seja, a PD, é explicada principalmente pela taxa de juros e grau do tomador. A Figura 4.1 mostra a porcentagem da importância da variável calculada no Modelo de *Survival Tree* para recursos com pelo menos 1% da importância total. A importância relativa de um determinado preditor  $x$  é a soma das melhorias quadradas sobre todos os nós internos da árvore para a qual  $x$  foi escolhido como a variável de particionamento [59].

Na seleção das variáveis do modelo, foram considerados os valores da importância das variáveis mostradas na Figura 4.1. Assim, foram avaliados os impactos selecionando todas as variáveis ou por uma fração da importância total obtida do modelo *Survival Tree*.

Embora sejam baseados em técnicas de sobrevivência, os modelos foram comparados considerando um ponto de corte e utilizando indicadores de desempenho aplicáveis a modelos de classificação, à semelhança do trabalho de Andreeva [72]. Basicamente, foram comparados modelos com duas estatísticas de desempenho: (i) *Area Under the Receiver Operating Characteristics* (AUROC) que fornece uma medida de precisão de classificação; e (ii) a distância de Kolmogorov-Smirnov (KS) que compara as distribuições da PD calculada para operações inadimplentes e não inadimplentes.

Estes indicadores fornecem, em um determinado ponto do tempo em que as curvas de sobrevivência são calculadas, a capacidade de atribuir medidas diferentes entre os casos que apresentaram *default* e aqueles que não apresentaram. No caso do KS, quanto maior a distância entre as curvas de distribuição acumulada entre os dois grupos, melhor o



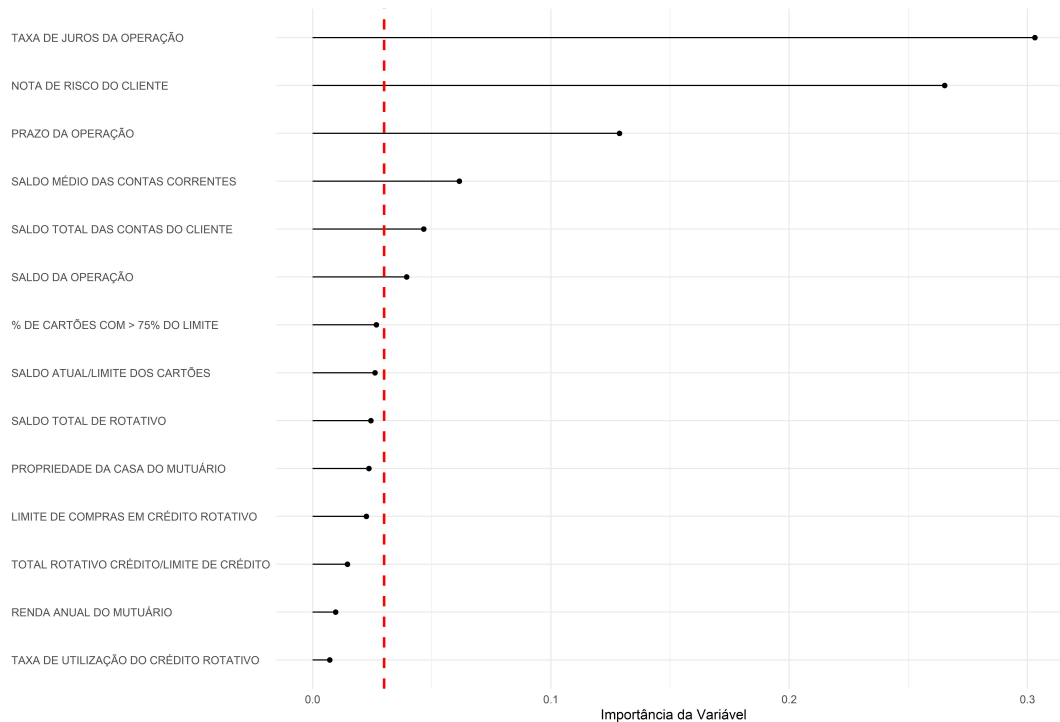


Figura 4.1: Importância das Variáveis (Fonte: Elaborado Pelo Autor).

modelo na sua função de identificar maus pagadores. Já a AUROC acrescenta a essa análise um sentido de direção, um modelo mal especificado poderia ter um indicador KS atribuindo pontuações mais elevadas a "bons pagadores", algo que não aconteceria na AUROC, por trazer uma medida baseada nas taxas de verdadeiros positivos em relação aos falsos negativos.

#### 4.4.1 *Survival Tree*

Na técnica de *Survival Tree*, foi avaliada a sensibilidade do desempenho do modelo em relação à inclusão das variáveis explicativas disponíveis. Foram seguidas três abordagens diferentes para selecionar características: (i) todas as variáveis independentes relevantes; (ii) seis variáveis independentes, que refletem 85% da métrica de importância total, e (iii) quatro variáveis, que refletem 76% da métrica de importância total. Além disso, foi comparado o desempenho do modelo sob diferentes observações mínimas de quebras dos nós da árvore. Para cada conjunto de variáveis, foram comparados modelos com 10%, 5% e 1% de observação nas quebras dos nós. A Tabela 4.1 descreve os indicadores de desempenho do modelo no conjunto de dados de teste.

De acordo com as métricas utilizadas neste estudo, não se nota ganho ao incluir as demais variáveis além das seis mais importantes. Dentre as configurações utilizadas,

Tabela 4.1: Resultados do Modelo *Survival Tree*

Variáveis	%obs nó	AUROC	KS
Full	10	0,6731	0,2541
	5	0,6730	0,2593
	1	0,6729	0,2565
6	10	0,6731	0,2541
	5	0,6731	0,2593
	1	0,6722	0,2565
4	10	0,6702	0,2469
	5	0,6681	0,2468
	1	0,6722	0,2565

entende-se que a configuração com 5% de observações nas quebras dos nós e seis variáveis foi a que apresentou melhor desempenho. Doravante, este melhor modelo selecionado será o único que para o qual serão apresentados os resultados. Devido à fácil interpretação do modelo *Survival Tree*, foram analisados os resultados da relevância das variáveis explicativas para explicar a probabilidade de descumprimento.

Para construção da árvore final, as quebras das árvores foram definidas com pelo menos 10% das observações. Essa medida foi necessária para evitar *overfitting* e melhorar os resultados fora da amostra de modelagem. Após aplicar a técnica *Survival Tree* levando em consideração as variáveis descritas na seção de preparação de dados e com as premissas detalhadas, foi obtido o resultado apresentado na Figura 4.2.

Em cada um dos nós apresentados na Figura 4.2 existe uma curva de sobrevivência e o algoritmo foi executado com base na definição de grupos com diferentes funções de sobrevivência. Dessa visualização é interessante notar a interpretação possível pelo modelo apresentado. A variável mais importante foi a taxa de juros, geralmente definida a partir de análises de créditos baseadas em informações mais precisas as quais a instituição teria sobre os clientes. Maiores taxas (acima de 12,92) tem as maiores PD, o que é esperado intuitivamente, tendo em vista as taxas calculadas para se obter retorno financeiro após cobertura da inadimplência prevista. Operações com prazos maiores têm maior PD no período completo, o que pode ser atribuído a um maior tempo de exposição ao risco de crédito. Nas últimas quebras, outras visualizações esperadas intuitivamente ocorrem, clientes com maiores saldos em conta corrente têm menor PD e clientes com notas de risco mais elevadas apresentam as maiores taxas entre os nós.

A Figura 4.3 mostra as curvas de sobrevivência de cada nó resultante da *Survival Tree*.

Grupos com diferentes curvas de sobrevivência foram separados. Cada curva pode ser entendida como a estimativa do modelo, não como uma probabilidade única em cada grupo, mas como a evolução do percentual de operações que entram em situação de descumprimento no tempo  $t$ .

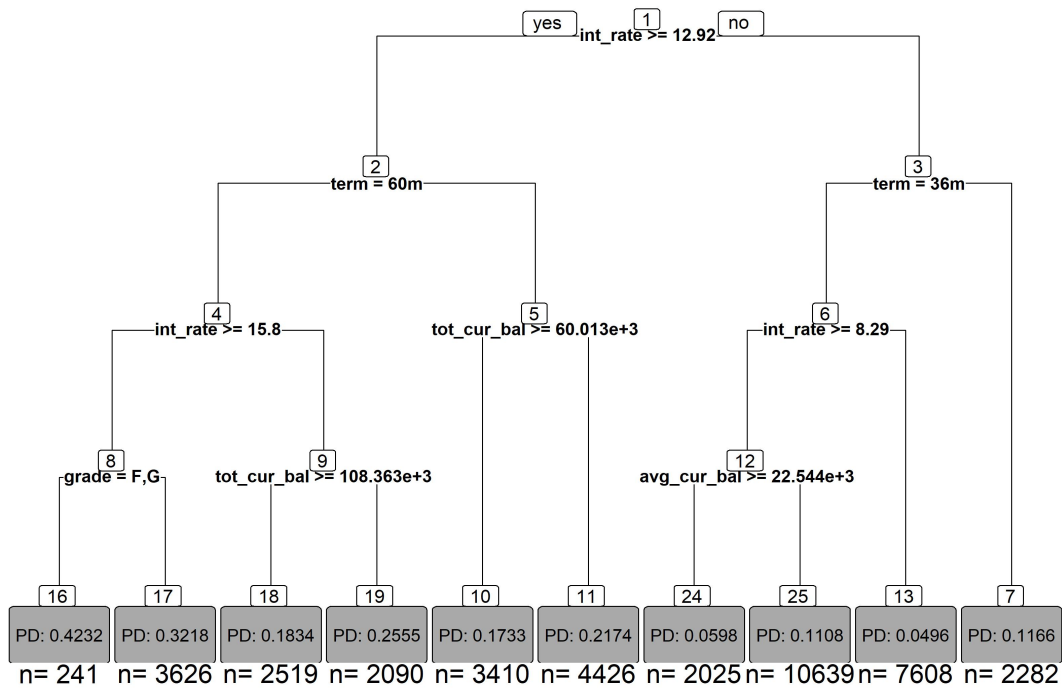


Figura 4.2: *Survival Tree* para Probabilidade de Descumprimento (Fonte: Elaborado Pelo Autor).

A facilidade de interpretação do modelo *Survival Tree* deve ser destacada como um ponto positivo no que se considera sobre os problemas de modelos de *machine learning* no risco de crédito. No entanto, sua estrutura simples pode resultar, eventualmente, em indicadores de desempenho piores que os algoritmos mais robustos, como RSF, também utilizado neste trabalho.

#### 4.4.2 *Random Survival Forest*

Para calcular o RSF, foram definidos diferentes parâmetros para medir o desempenho e avaliar a sensibilidade. Para cada árvore usada nas florestas, foi usada a mesma abordagem descrita anteriormente. A RSF foi executada em múltiplas configurações: com 1.000 árvores e todas as características e com 100, 500 e 1.000 árvores com as seis características mais importantes, selecionadas *n Survival Trees*. Os modelos que foram feitos sem controle de observações nos nós resultaram em *overfitting*, com bom desempenho apenas na amostra de treinamento.

A floresta foi executada em múltiplas configurações: com 1.000 árvores e todas as características e com 100, 500 e 1.000 árvores com as seis características mais importantes,

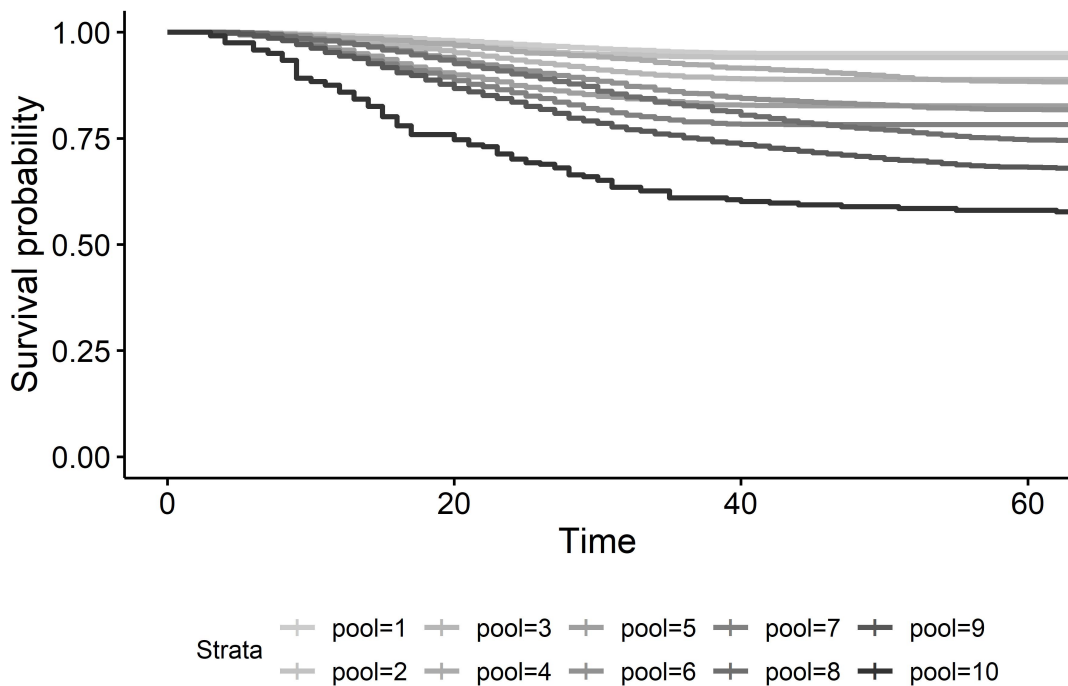


Figura 4.3: Curvas de Sobrevivência dos Nós Finais (Fonte: Elaborado Pelo Autor).

selecionadas no *Survival Tree*. Os modelos que foram feitos sem controle de observações nos nós finais resultaram em *overfitting*, com bom desempenho apenas na amostra de treinamento.

Por esse motivo, foram construídos modelos com árvores que tiveram pelo menos 5% e 10% nas quebras dos nós das árvores. Nestes modelos, cada nó de cada árvore tem pelo menos 10% das operações totais. A Tabela 4.2 mostra os resultados da RSF da base de teste.

Tabela 4.2: Resultados do Modelo *Random Survival Forest*

%obs quebra	Árvores	AUROC	KS
1	1000	0,6899	0,2732
	1000	0,6876	0,2670
5	500	0,6875	0,2697
	100	0,6872	0,2683
10	1000	0,6846	0,2627
	500	0,6845	0,2636
	100	0,6838	0,2633

Utilizando apenas as quatro covariáveis selecionadas, uma nova RSF utilizando os mesmos hiperparâmetros foi realizada para obter as estimativas do conjunto da curva de sobrevivência e probabilidade de ocorrência do evento.

Considera-se que os resultados do modelo com 1% de observações nas quebras dos nós a melhor alternativa, dadas as estatísticas analisadas. Este modelo servirá de base para validação e comparação com os demais métodos apresentados na subseção seguinte.

### 4.4.3 Avaliação dos Modelos

Foram comparadas a estimativa do parâmetro para os 12 meses de 2017 com a amostra de validação, com o objetivo de avaliar a aderência das estimativas em períodos menores do que o de desenvolvimento do modelo. A Tabela 4.3 mostra a comparação dos percentuais de inadimplência estimados pelo modelo *Survival Tree* e os observados ao longo dos 12 meses, pelo indicador KS para essas medidas.

Tabela 4.3: Comparação dos Modelos

PD Estimada	Modelo	AUROC	KS
<i>Lifetime</i>	<i>Survival Tree</i>	0,6792	0,2576
	<i>Survival Forest</i>	0,7039	0,2976
	Cox PH	0,7026	0,2955
	<i>Decision Tree</i>	0,6875	0,2868
	<i>Logistic Regression</i>	0,7021	0,2949
12 meses	<i>Survival Tree</i>	0,7106	0,3163
	<i>Survival Forest</i>	0,7274	0,3371
	Cox PH	0,7191	0,3278
	<i>Decision Tree</i>	0,6541	0,3015
	<i>Logistic Regression</i>	0,7283	0,3399

Foi identificado que os cinco modelos desenvolvidos para o PD *lifetime* possuem um poder discriminatório semelhante, considerando os valores AUROC da Tabela 4.3. O RSF é o modelo mais robusto em termos de discriminação, seguido do Modelo de Cox e da Regressão Logística. Da mesma forma, observou-se que na Estatística de KS, teste utilizado para avaliar o poder discriminatório dos modelos, o RSF também apresenta os melhores resultados, seguido pelo Modelo de Cox e pela Regressão Logística.

A capacidade de discriminar operações descumpridas e não descumpridas da *Survival Tree* menor do que os resultados da Árvore de Decisão. No entanto, as *Survival Tree* podem ser usadas para estimativas da PD ao longo do tempo, enquanto as Árvores de Decisão geram apenas uma probabilidade de descumprimento em um determinado momento.

De modo geral, a partir da análise dos indicadores de desempenho, as principais constatações a serem destacadas são: (i) os modelos de análise de sobrevivência tem desempenho semelhante aos modelos de classificação tradicionais, com a vantagem de permitir o ajuste das probabilidades a qualquer prazo de operação; (ii) em relação ao uso do RSF, verificou-

se ganho de desempenho em relação a métodos tradicionais, no entanto, deve-se avaliar o custo benefício do uso, tendo em vista a maior dificuldade dos efeitos das variáveis nas estimativas realizadas por este modelo.

## 4.5 Conclusão

O objetivo deste trabalho foi contribuir para as discussões recentes sobre avaliação do desempenho dos modelos de análise de sobrevivência baseados em árvore, em comparação aos modelos tradicionais de análise de sobrevivência e aos modelos de classificação para estimativa da PD.

Os resultados usando *Survival Tree* permitem a segmentação dos dados em grupos de risco com base no comportamento da curva de sobrevivência, e tem a vantagem de ter fácil interpretação, de forma gráfica e objetiva, das suas predições. Mais especificamente, o modelo implicou em sete grupos de risco, de 6,4% a 27,1% de inadimplência, com uma taxa média de inadimplência de 14,89%. Os resultados da aplicação do modelo em bases de teste apresentaram comportamento semelhante ao obtido no *dataset* de treinamento.

Apesar disso, o modelo apresentou indicadores de desempenho preditivo inferiores aos métodos tradicionais na estimativa *lifetime*. Com uma abordagem mais robusta e complementar, em comparação ao *Survival Tree*, o RSF apresentou melhor desempenho preditivo que métodos tradicionais, resultado alinhado aos obtidos em estudos recentes. No entanto, vale destacar que a técnica possui capacidade limitada da avaliação do sentido do efeito das covariáveis, devendo haver combinações com outros métodos ou limitar a análise à importância das variáveis.

Este trabalho possui algumas limitações importantes a serem destacadas. Os dados utilizados sobre o histórico de pagamentos dos créditos continham apenas a última posição da carteira, com datas de inadimplência. Portanto, se o mutuário apresentasse inadimplência e, posteriormente, quitasse a operação, este histórico de inadimplência não estaria disponível na base. Os resultados obtidos não permitem conclusão generalizada sobre a comparação dos métodos, estando estes limitados à metodologia seguida, ao período e aos ativos avaliados, servindo de contribuição para as discussões futuras sobre este tema.

Em estudos futuros, podem ser utilizadas técnicas que apresentaram bom desempenho em trabalhos recentes, como outras formas de utilização de modelos baseados em árvore, a exemplo do *Gradient Boosting* [67][66] e de variações das Redes Neurais para análise de sobrevivência, como aplicado por Blumenstock [73] no risco de crédito.

Por fim, este trabalho apresenta contribuições para a discussão recente sobre o uso de modelos de *machine learning* nos modelos utilizados na gestão do risco de crédito nos padrões mais atualizados de contabilidade bancária. Corrobora-se o entendimento de

que os métodos baseados em árvores para modelos de sobrevivência podem ser uma boa alternativa para modelagem da PD.

# Capítulo 5

## Ajuste Forward-Looking

### 5.1 Introdução

Em termos de acompanhamento de efeitos de variáveis macroeconômicas, um dos principais indicadores de risco de crédito utilizado é o *Non-Performing Loans* (NPL) [74], definido como a razão entre os saldos inadimplentes e a carteira total, segundo Vazquez [75]. Embora outros indicadores mais detalhados possam estar disponíveis nas instituições para indicar o risco de crédito, este trabalho utilizará os NPL por sua maior disponibilidade em dados públicos.

O padrão IFRS 9 destaca a necessidade de se considerar informações todas as informações prospectivas razoáveis, incluindo as macroeconômicas, na estimativa da ECL. Como abordado mais detalhadamente na próxima seção, as estimativas de efeitos macroeconômicos sobre os indicadores de inadimplência são feitas por meio de modelos preditivos.

Quando estes modelos são utilizados para projeções, as covariáveis utilizadas também são projetadas por modelos ou expectativas de mercado, que, no mercado brasileiro, passam por diversas revisões até a observação dos valores reais. Em situações extremas de mudanças de expectativas, entende-se que os erros das projeções podem impactar negativamente a estimativa da ECL, e possivelmente as decisões de uma Instituição Financeira.

Neste contexto, este trabalho tem o objetivo de estimar os efeitos de variáveis macroeconômicas sobre o NPL em instituições bancárias brasileiras e, a partir destas estimativas, avaliar se o uso de cenários aumenta os erros das projeções da NPL.

Além desta introdução, este trabalho é composto por quatro seções, sendo que na segunda seção há a revisão da literatura sobre o tema, na terceira seção apresenta-se o fluxo metodológico da pesquisa, na quarta seção, são destacados os resultados obtidos e, por fim, a quinta seção traz as considerações finais do estudo.



## 5.2 Referencial Teórico

Na Resolução CMN Nº 4.966 [18], que estabelece as regras para compatibilização do padrão IFRS 9 na contabilidade brasileira, fica estabelecido que a probabilidade de um instrumento financeiro caracterizar-se como descumprido, deve considerar, no mínimo o prazo do instrumento financeiro e, conforme descrito na Resolução:

“A situação econômica corrente e previsões razoáveis e justificáveis de eventuais alterações nas condições econômicas e de mercado que afetem o risco de crédito do instrumento, durante o seu prazo esperado..”

Tal citação da Resolução dá destaque à necessidade das IFs manterem estudos para aplicação de ajustes decorrentes de fatores macroeconômicos que impactam a probabilidade de um instrumento se tornar inadimplente e, conseqüentemente, a ECL.

Econometria é a denominação dada à avaliação empírica das relações econômicas, a partir da aplicação da estatística e da matemática a dados concretos [76]. Diversas técnicas de estudos destas relações são utilizadas no contexto do risco de crédito, conforme destacado nas subseções seguintes deste capítulo.

### 5.2.1 Modelagem Econométrica

De acordo com Hayashi [77], havendo um vetor de variáveis  $Y_{(n \times 1)}$ , uma matriz de covariáveis  $X_{(n \times K)}$  e uma matriz de parâmetros  $\beta_{(K \times n)}$ , um modelo clássico de Regressão Linear é descrito conforme Equação 5.1.

$$Y = X\beta + \epsilon \quad (5.1)$$

Ou seja, o modelo de regressão clássico, utilizado na econometria em suas diversas variações, ajusta uma matriz de  $K$  parâmetros  $\beta$  que melhor ajusta uma equação linear para explicar  $n$  observações da variável resposta  $Y$  em função de  $K$  covariáveis, ou variáveis explicativas  $X$ . O ajuste linear resultante dos parâmetros tem um vetor de erros  $\epsilon$ , que deve ser estritamente exógeno, ou seja  $E(\epsilon|X) = 0$  [77].

No contexto dos estudos econométricos, conforme destacado no trabalho de Hendry [78], no desenvolvimento de um modelo econométrico, para evitar erros de especificação, deve-se satisfazer os seguintes critérios:

- O modelo deve ser confirmado pelos dados, com previsões que tenham sentido lógico;
- Deve haver consistência com a teoria econômica;
- O modelo não pode ter correlações com seus resíduos;

- Os parâmetros do modelo devem ser estáveis;
- Os erros dos modelos devem ser ruídos brancos; e
- A modelagem deve incluir os modelos concorrentes, de modo que o modelo escolhido não pode ser menos adequado que outros modelos.

## Regressão para Dados em Painel

Há várias formas de dados disponíveis para análise empírica, por exemplo: cortes transversais, séries temporais, entre outros. Dentre estes, os dados em painel um mesmo corte transversal é acompanhado ao longo do tempo. Ou seja, os dados são caracterizados por uma dimensão espacial e uma temporal [76].

No estabelecimento da equação de uma regressão para dados em painel, considerando um modelo de NPL em função somente do Produto Interno Bruto (PIB) e da série defasada, por exemplo, poderia ser descrito a partir da Equação 5.2:

$$NPL_{it} = \beta_{0i} + \beta_1 NPL_{i(t-1)} + \beta_2 PIB_t \quad (5.2)$$

Ou seja, ajusta-se uma equação com interceptos ajustados para cada indivíduo, e os parâmetros dos efeitos das covariáveis seriam aplicados às variáveis no tempo. A diferenciação entre efeitos fixos e efeitos aleatórios se dá pela interpretação do parâmetro  $\beta_1$  como um valor fixo ou como uma variável aleatória que tem como média o valor do parâmetro. O teste de Hausman [79] é uma alternativa para verificar a existência de efeito aleatório em um modelo especificado.

Ainda segundo Gujarati [76], a unidade de corte transversal pode ser uma família, empresa ou estado, por exemplo. O conceito de grupo homogêneo de risco [18] em que são calculadas os parâmetros de risco pode ser interpretado como um corte transversal, motivo pelo qual as regressões para dados em painel tem sido amplamente utilizadas na literatura para modelagem de risco de crédito, conforme detalhado na subseção seguinte.

### 5.2.2 Trabalhos Correlatos

Bholat [80] realizaram um trabalho sobre os NPL no contexto do início do IFRS 9 e pontua que na sua implementação, por haver a necessidade de uma visão prospectiva da instituição, as provisões se tornam mais sensíveis às perspectivas macroeconômicas. Neste mesmo trabalho, os autores trazem a crise financeira de 2007-2009 como um bom exemplo dos efeitos da divergência de trajetória entre a NPL histórica e as projeções. Tal situação justifica a necessidade de uma visão mais prospectiva nas estimativas, no entanto

os efeitos positivos dessa aplicação dependem da qualidade das projeções e expectativas das instituições.

De acordo com Suárez e Serrano [81], na classificação de ativos do IFRS 9, os créditos são classificados como: estágio 1, que são os instrumentos financeiros cujo risco de crédito não tenha apresentado aumento significativo de risco; estágio 2, ativos com aumento significativo de risco de crédito sem apresentar inadimplência; e estágio 3, de créditos deteriorados, onde se enquadram os NPL.

Ainda segundo esses autores, a provisão do NPL não alteraria em relação à abordagem da perda incorrida, sendo a principal alteração o reconhecimento de perdas esperadas em exposições dos demais estágios. No estudo de Bolognesi [82] é apresentado o caso italiano sobre o custo de alavancagem envolvendo o NPL, com cálculo da ECL considerando os múltiplos cenários sobre vendas de ativos e recuperação nos NPL.

O ajuste de modelos econométricos para avaliação de impacto de variáveis macroeconômicas sobre o NPL tem variadas contribuições na literatura. O estudo de [75] realizou um trabalho com dados de instituições brasileiras reportadas ao BCB para aplicação em testes de estresse. Neste trabalho foram ajustadas regressões para dados em painel encontrando uma relação negativa entre o PIB e o NPL.

Um estudo [83] realizado com regressão para dados em painel sobre o NPL de 85 bancos da Itália, Grécia e Espanha, assim como no trabalho mencionado anteriormente, identificou uma relação negativa do PIB com os NPL, além de uma relação diretamente proporcional da taxa de desemprego com a NPL, bem como relações de outras variáveis sobre os resultados dos créditos.

Syed e Tripathi [84] verificaram efeitos significativos negativos do PIB e da inflação positivos da taxa de desemprego e das "*saving rates*" sobre os NPL em regressão aplicada a dados dos cinco países do bloco Brasil, Rússia, Índia, China e África do Sul (BRICS). Resultado semelhante ao apresentado por Ghosh, (2015) que verificou estas relações sobre os NPL em dados de todos os bancos comerciais de 50 estados dos Estados Unidos da América (EUA) e do Distrito de Colúmbia.

O estudo de Beck [85] com aplicação de modelos fixos e dinâmicos para dados em painel em dados de 75 países identificou efeitos significativos entre o crescimento real do PIB, preços de ações, taxa de câmbio e taxa de juros dos empréstimos sobre os NPL, além de efeito diretamente proporcional da defasagem da própria série de NPL, relação também corroborada em países da Zona do Euro [86].

### 5.2.3 Modelos de *Machine Learning*

Assim como na modelagem do RSF, e dos exemplos sobre o uso de *machine learning* na modelagem de risco de crédito, é comum que estes algoritmos de modelagem apresentem

indicadores de desempenho melhores que os métodos tradicionais. Tendo em vista a necessidade de estudos para justificar os impactos no risco de crédito [18] a questão da capacidade limitada de interpretação dos parâmetros é uma questão discutida sobre os modelos de risco de crédito.

Em novembro de 2021, o *European Banking Authority* (EBA) abriu um *paper* para discussão sobre o uso de modelos de *machine learning* para os modelos de IRB, cujas metodologias são aproveitadas para os cálculos da metodologia de cálculo do IFRS 9. Como desafios mais complexos estão relacionados a:

- a interpretação dos seus resultados;
- necessidade de assegurar a sua compreensão adequada pelas funções de gestão; e
- a justificativa os seus resultados perante os supervisores.

Diante da questão da interpretação limitada dos algoritmos mais avançados de *machine learning*, o trabalho de Lundberg [87] apresentou um *framework* unificado para interpretação de previsões destes modelos. O *SHapley Additive exPlanations* (SHAP) atribui a cada variável a alteração da predição esperada do modelo, condicionado ao valor da variável.

Sobre este assunto EBA elenca as técnicas utilizadas para explicar o motivo pelo qual um modelo produz cada estimativa. Destacam-se as seguintes técnicas amplamente utilizadas [88]:

- Ferramentas gráficas mostrando os efeitos das variáveis explicativas no modelo;
- Medidas de importância das variáveis para demonstrar o poder explicativo de cada variável;
- *Shapley values* para quantificar o impacto de cada variável no modelo;
- Técnicas como LIME, que fornecem aproximações simples do modelo nas proximidades de uma observação; e
- Explicações sobre como uma previsão específica do modelo poderia ser modificada alterando os valores das variáveis explicativas o mínimo possível.

Considerando a abrangência global das contribuições sobre tema, este trabalho traz contribuições sobre o NPL no mercado bancário brasileiro, além disso é apresentado um breve estudo de caso com aplicação de modelo de *machine learning* baseado em florestas aleatórias, com sua interpretação por meio do SHAP [87]. Na seção seguinte, são apresentados os procedimentos metodológicos utilizados para esta pesquisa.

## 5.3 Metodologia

A pesquisa deste capítulo é do tipo exploratória, com abordagem quantitativa para avaliação do impacto do uso de cenários nas previsões de taxas de clientes Pessoa Física em instituições bancárias a partir de regressões para dados em painel, adequada à composição dos dados utilizados, em que há duas dimensões (unidade transversal e tempo) [77]. Foram utilizados painéis de mínimos quadrados de efeitos fixos.

Foram utilizados dados agregados reportados pelas instituições financeiras e disponibilizados pelo BCB, no Painel de Operações de Crédito [89], sobre as carteiras de crédito, com a série temporal compreendendo o período entre junho de 2012 até setembro de 2021. Para construção das taxas de inadimplência trimestrais das unidades transversais, foram utilizadas as informações sobre a Carteira Ativa e Carteira Inadimplente Arrastada, com detalhamento sobre tipo de instituição (Bancário, Não Bancário e Cooperativas) e porte dos clientes.

As variáveis macroeconômicas foram obtidas pelo Sistema Gerenciador de Séries Temporais [90]. Além disso, foram extraídas, do Sistema de Expectativas de Mercado [91], o histórico da mediana das estimativas de variáveis macroeconômicas feitas pelos agentes do mercado. Essas informações foram utilizadas para avaliação do quanto o uso de cenários pode influenciar nas previsões do NPL.

A partir dessas fontes de informações, as seguintes etapas foram seguidas para elaboração deste trabalho:

- Extração e tratamento de dados sobre os NPL no mercado brasileiro;
- Extração e tratamento de séries temporais das variáveis macroeconômicas e projeções;
- Desenvolvimento de modelos econométricos para estimação do NPL com dados de 2012Q2 a 2020Q2; e
- Avaliação e comparação dos modelos entre:
  - dados macroeconômicos observados entre 2020Q3 e 2021Q2; e
  - dados projetados para o mesmo período em 2020Q2, que representam os dados disponíveis para aplicação no momento do uso do modelo.

Na comparação entre os modelos com uso de cenários ou com variáveis foram testadas as seguintes hipóteses para os modelos desenvolvidos:

$H_0$ : Os erros dos modelos não são influenciados pelo uso de projeções.

$H_a$ : Os erros dos modelos são maiores quando são utilizadas projeções de covariáveis.

Para extração e tratamento dos dados, foi utilizado o pacote Pandas [92] na linguagem Python [93] e para a desenvolvimento dos modelos econométricos, foi utilizado o software Eviews *University Edition* [94]. Em todos os testes de hipóteses foi considerado um nível de significância de 5%.

## 5.4 Resultados

Esta seção sumariza os principais resultados obtidos após tratamento dos dados e desenvolvimento dos modelos econométricos. Nas subseções a seguir serão descritos os procedimentos para extração e tratamento dos dados realizados, a análise exploratória das variáveis, os modelos desenvolvidos e a avaliação de desempenho dos modelos para projeção das taxas de inadimplência futuras.

### 5.4.1 Dados Coletados

Resumidamente, foram utilizadas três fontes de dados para elaboração deste trabalho, todas disponibilizadas pelo BCB, sendo elas:

- Dados de crédito do Painel de Operações de Crédito [89] com informações divulgadas sobre as carteiras de crédito de forma agregada;
- Variáveis macroeconômicas disponibilizadas por diversos órgãos públicos com consulta centralizada no Sistema Gerenciador de Séries Temporais (SGS) [90]; e
- Expectativas do mercado financeiro sobre as variáveis macroeconômicas, consultas no Sistema de Expectativas de Mercado [91].

Para cada uma das séries foram realizados procedimentos para extração e tratamento para utilização no desenvolvimento dos modelos.

#### Dados de Crédito

Foram extraídos 109 arquivos mensais com dados agregados sobre os créditos nas Instituições Financeiras com a situação do último dia útil de cada mês. Os arquivos contêm, dentre outras, informações sobre a carteira ativa, carteira inadimplente, detalhamento sobre cliente (Pessoa Física (PF) ou Pessoa Jurídica (PJ)), modalidade de crédito, Unidade da Federação (UF), ocupação dos clientes PF e porte dos clientes por renda.

Para construção das séries de inadimplência a serem utilizadas na estimação dos modelos, foram consideradas somente as séries de clientes PF de instituições bancárias. Não foi considerada distinção de tipo de indexação dos créditos nem de ocupação dos clientes e, tendo em vista o ajuste do modelo a partir de variáveis macroeconômicas nacionais, não foram consideradas informações sobre UF na definição das unidades de cortes transversais.

Além disso, alguns tratamentos foram feitos para melhoria do comportamento das séries das taxas de inadimplência, desde que este agrupamento apresentasse sentido negocial. Foram agrupadas as séries de clientes ‘Sem Rendimento’, ‘Até 1 salário-mínimo’ e ‘Mais de 1 a 2 salários-mínimos’. Também foram agrupadas as modalidades de cartão de crédito e CDC sem consignação.

A partir destes resultados, as carteiras ativas e carteiras inadimplentes foram somadas para as diferentes modalidades de crédito e portes de clientes, e as taxas de inadimplência foram calculadas como a razão entre o saldo inadimplente e a carteira ativa em cada combinação de modalidade de crédito e porte. Para construção do modelo de dados em painel, não foram consideradas as séries classificadas com o porte ‘PF - Indisponível’ por apresentar baixa quantidade e quebras de continuidade da série. Além disso, devido ao comportamento diferente das demais unidades.

A modalidade ‘PF - Rural e agroindustrial’ por apresentar movimentos em direções diferentes às demais modalidades e a modalidade ‘PF - Empréstimo com consignação em folha’ também não foi considerada por manter uma taxa relativamente constante ao longo do período. Como resultado, os dados foram modelados em 24 unidades de cortes transversais resultantes das combinações dos seis portes e quatro modalidades apresentadas na Tabela 5.1, com as respectivas taxas de inadimplência médias no período.

Destaca-se na Tabela 5.1 que há um padrão, em todos os portes, de ordenamento entre as modalidades em relação às taxas, sendo as menores taxas observadas no ‘Crédito Habitacional’, seguido por ‘Veículo’, ‘Cartão e Crédito sem Consignação’ e, por fim, ‘Outros Créditos’. Além disso, como é esperado intuitivamente, quanto maior a renda, menor as taxas de inadimplência, com uma única exceção observada na redução, seguida de aumento do crédito Veículo no porte de Mais de 10 a 20 salários-mínimos.

## **Variáveis Macroeconômicas**

Relativamente às variáveis macroeconômicas, os dados são disponibilizados em diferentes periodicidades e horizontes temporais no SGS. Foram extraídos dados sobre o PIB, Taxa de Câmbio, IPCA, Taxa Selic, População Desocupada e Força de Trabalho, A Tabela 5.2 mostra as variáveis disponíveis, as e as datas disponíveis no SGS.

As variáveis macroeconômicas foram consultadas por meio de *Application Programming Interface* (API) que acessam diretamente as séries temporais para análise. Foram

Tabela 5.1: Taxas de Inadimplência por Unidade de Corte Transversal

<b>Porte</b>	<b>Modalidade</b>	<b>Taxa de Inadimplência</b>
Até 2 salários-mínimos	Habitacional	0.021
	Veículos	0.046
	Cartão e CDC	0.106
	Outros créditos	0.130
Mais de 2 a 3 salários-mínimos	Habitacional	0.018
	Veículos	0.041
	Cartão e CDC	0.082
	Outros créditos	0.124
Mais de 3 a 5 salários-mínimos	Habitacional	0.015
	Veículos	0.035
	Cartão e CDC	0.067
	Outros créditos	0.113
Mais de 5 a 10 salários-mínimos	Habitacional	0.013
	Veículos	0.028
	Cartão e CDC	0.055
	Outros créditos	0.100
Mais de 10 a 20 salários-mínimos	Habitacional	0.010
	Veículos	0.023
	Cartão e CDC	0.041
	Outros créditos	0.081
Acima de 20 salários-mínimos	Habitacional	0.010
	Veículos	0.028
	Cartão e CDC	0.030
	Outros créditos	0.045

Tabela 5.2: Variáveis Macroeconômicas para Modelagem

<b>Série Temporal</b>	<b>Horizonte Temporal</b>
PIB	1995 a 2021
Taxa de Câmbio	1953 a 2021
População Desocupada	2012 a 2021
Força de Trabalho	2012 a 2021
IPCA	1980 a 2021
Taxa Selic	1986 a 2021

feitas transformações em algumas das variáveis. Para as taxas Selic e de IPCA, foram consideradas as taxas acumuladas de 12 meses em cada data de referência. Assim como no trabalho de Vazquez [75], o PIB foi considerado como a primeira diferença do logaritmo natural da série do PIB a preços de mercado dessazonalizado, com ajuste sazonal, disponibilizado no SGS, calculado conforme equação 5.3.



$$G(PIB) = \ln(PIB) - \ln(PIB_{-1}) \quad (5.3)$$

A taxa de desemprego foi calculada a partir da razão entre a População Desocupada e População na Força de Trabalho. Foi calculada como a taxa de desemprego com retirada do componente de sazonalidade, resultante da decomposição da taxa de desemprego utilizando o pacote statsmodels [95] na linguagem python. A Figura 5.1 mostra a série da taxa de desemprego e a taxa de desemprego dessazonalizada.

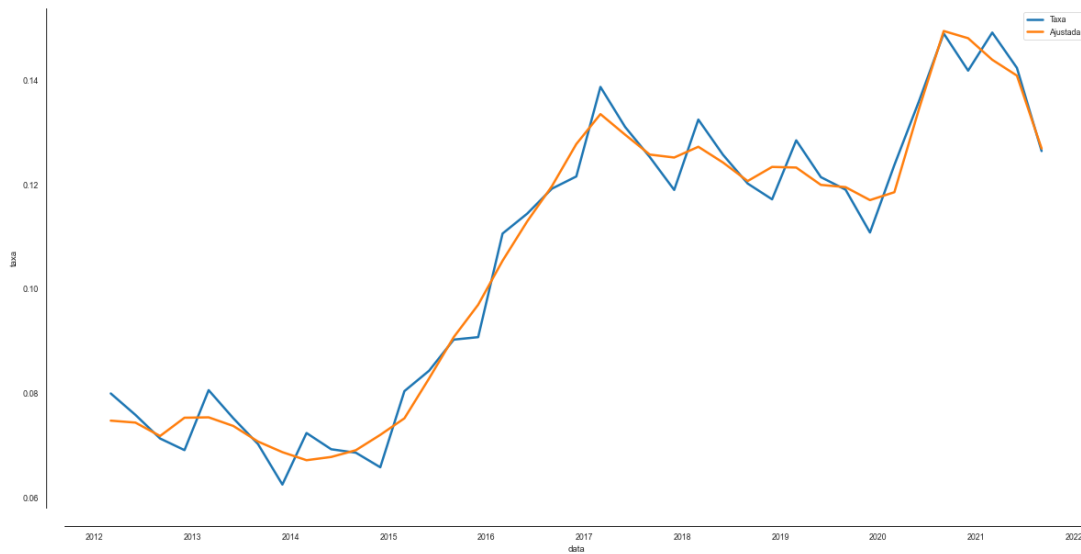


Figura 5.1: Ajuste de Sazonalidade da taxa de desemprego (Fonte: Elaborado Pelo Autor).

A partir da Figura 5.1, nota-se a existência de sazonalidade da série das taxas de desemprego. No desenvolvimento dos modelos, foi utilizada a série dessazonalizada para evitar possíveis inconsistências nos modelos.

### Expectativas das Variáveis Econômicas

Quando ao uso de cenários na estimativa da ECL no padrão IFRS 9, não há necessidade de simulações complexas ou elevado número de cenários para atendimento do padrão, sem o objetivo de se utilizar o pior ou o melhor cenário projetado [12]. Neste trabalho, considerou-se as medianas das projeções das instituições financeiras disponibilizadas no Sistema de Expectativas de Mercado.

No caso das projeções, os dados foram acessados via API, e foram construídas duas informações distintas:

- dados observados com projeções trimestrais a partir de setembro de 2020, utilizada para avaliar o desempenho preditivo do modelo com uso de informações disponíveis na data da projeção;
- base de expectativas de mercado para os próximos 12 meses, para avaliar a inclusão das expectativas de mercado como covariáveis.

Dessa forma, para construção das projeções, foi necessário transformar as variáveis para o mesmo padrão das variáveis macroeconômicas. As projeções do PIB são feitas como variações anuais e foram transformadas com aplicação das taxas aos valores observados e, posteriormente, aplicando a mesma função de crescimento da equação 5.3. Taxas Selic e IPCA foram tratadas da mesma forma que as variáveis macroeconômicas observadas, com valores acumulados dos últimos 12 meses, incluindo expectativas a partir de julho de 2020.

Para exemplificar as variações das expectativas conforme se aproxima a data da projeção, a Figuras 5.2 exemplifica os erros absolutos da estimativa (em comparação à última expectativa disponibilizada) relacionados aos dias de defasagem da estimativa para o PIB, com os diferentes trimestres apresentados nos quadrantes da figura.

Nota-se que as variações percentuais da estimativa reduzem conforme diminui a defasagem entre a variável observada. A Tabela 5.3 mostra as medidas de erro *Mean Absolute Percentage Error* (MAPE) e *Mean Absolute Error* (MAE), calculadas para a estimativa de um ano de defasagem para a última expectativa apresentada. Não há histórico de taxas de desemprego anteriores ao ano de 2021, impossibilitando esta análise para esta variável.

Tabela 5.3: Medidas de Erro da Projeção das Variáveis Explicativas (12 meses)

Variável	MAE	MAPE
Câmbio	0,7943	15,109%
PIB	4,0071	224,943%
IPCA	0,4110	54,906%
Selic	3,0000	80,449%

Destacam-se os erros elevados nas projeções do PIB (variação ano a ano), em especial por conta dos impactos decorrentes da pandemia do COVID-19, no primeiro trimestre de 2020. Essas variações devem ser fonte de avaliação para os agentes envolvidos nas projeções de taxa de inadimplência para cálculo da ECL.

## 5.4.2 Análise Exploratória

Foram aplicados os testes de estacionariedade de teste *Augmented Dickey-Fuller* (ADF) [96], *Phillips-Perron* (PP) [97], com a hipótese nula de raiz unitária, e *Kwiatkowski*

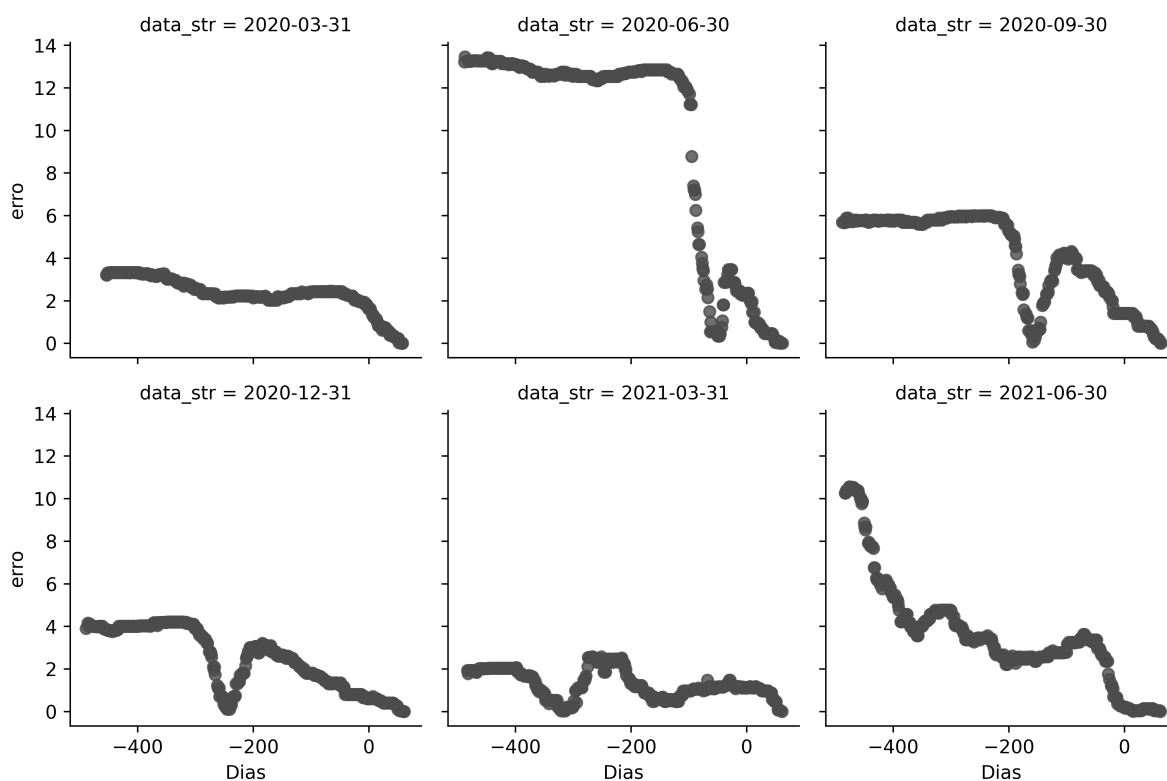


Figura 5.2: Erros Absolutos da Projeção do PIB por Defasagem (Fonte: Elaborado Pelo Autor).

Phillips-Schmidt-Shin (KPSS) [98], com hipótese nula de que a série é estacionária [77]. A Tabela 5.4 mostra os resultados dos testes:

Tabela 5.4: P-valor dos Testes de Raiz Unitária das Variáveis Explicativas

Variável	Tipo	ADF	PP	KPSS
G(PIB)	Nível	0,0001	0,0001	0,1000
Selic	Nível	0,0001	0,0037	0,0500
Selic	1ª Diferença	0,0001	0,0001	0,1000
Câmbio	Nível	0,0001	0,0001	0,1000
IPCA	Nível	0,0001	0,0012	0,1000
Tx. Desemprego	Nível	0,0667	0,9569	0,1000
Tx. Desemprego	1ª Diferença	0,0065	0,0073	0,1000

Em todos os testes realizados foram utilizados toda a série disponível, conforme apresentado na Tabela 5.4. Além disso, foram incluídos nos testes, sequencialmente, as componentes de intercepto e tendência nos testes e mantidos somente quando fosse existente efeito significativo. Nos testes apresentados, houve inclusão de efeito significativo de ten-

dência para a Selic em nível nos testes ADF e KPSS, no IPCA no teste KPSS e na taxa de desemprego em nível, nos testes ADF e KPSS.

Nos modelos estimados, as variáveis independentes foram em nível para as variáveis G(PIB), Câmbio e IPCA e, tendo em vista a estacionariedade e efeito de tendência observados em alguns dos testes, as variáveis Selic e Desemprego Dessazonalizada foram utilizadas como primeira diferença.

Para verificar as séries entre as unidades de cortes transversais, foram aplicados os testes Levin, Lin e Chu [99] e Im, Pesaran e Shin [100], rejeitando-se a hipótese nula de raiz unitária, exceto para as variáveis IPCA e câmbio. No caso do câmbio, tendo em vista que a série mais longa apresenta comportamento estacionário, que não é possível se obter os cortes transversais para todo o período, e sua presença em trabalhos relacionados [14] [85], optou-se por utilizá-la em modelos. Como o PIB foi considerado a preços de mercado, o IPCA não foi utilizado.

Sobre a variável dependente, a Figura 5.3 apresenta as séries das taxas de inadimplência para cada uma das unidades transversais, com os portes representados pelas cores e os quadrantes indicando as modalidades.

A partir destas séries, foram aplicadas regressões para dados em painel de efeitos fixos conforme procedimentos e resultados detalhados na próxima subseção.

### 5.4.3 Modelos Desenvolvidos

Após as transformações das variáveis, foram desenvolvidos modelos em que o sentido econômico das variáveis explicativas utilizadas fosse respeitado. Para todos os modelos, foi analisado graficamente a distribuição dos resíduos quanto aos padrões da distribuição normal, bem como sobre a correlação dos resíduos com os cortes transversais.

Tendo em vista que se trata de uma modelagem de uma taxa com variação de 0 a 1, utilizou-se a transformação logit nas taxas de inadimplência, conforme Equação 2 [75].

$$\text{Logit}(NPLt) = \log\left(\frac{NPLt}{1 - NPLt}\right) \quad (5.4)$$

No trabalho de [75], além do desenvolvimento de uma modelo de *Ordinary Least Squares* (OLS), foram ajustados modelos alternativos com a aplicação de *Generalized Method of Moments* (GMM), considerando o PIB uma variável exógena e o viés dos coeficientes associados à variável dependente defasada [75]. Assim como no trabalho de Vasquez [75] a aplicação de modelos GMM pode ser explorada em estudos futuros tomando como base os argumentos apresentados por esse autor. No entanto, vale destacar que no trabalho citado, não houve diferenças relevantes entre as conclusões obtidas com o modelo OLS e os modelos alternativos.

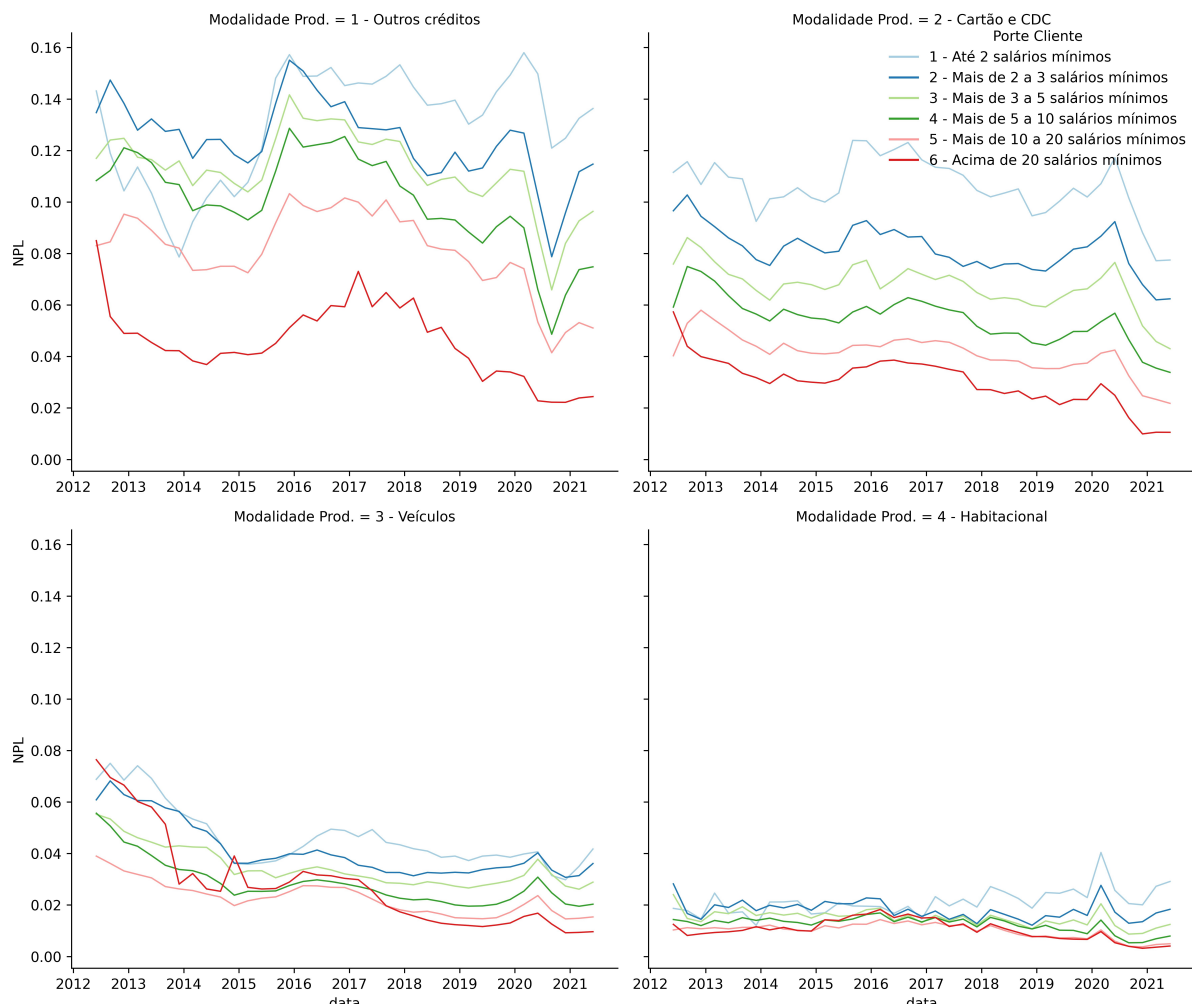


Figura 5.3: Taxas de Inadimplência das Unidades de Cortes Transversais (Fonte: Elaborado Pelo Autor).

Foram elaborados dois grupos de modelos, visando os testes posteriores de efeitos negativos dos erros dos cenários projetados sobre o desempenho dos modelos utilizados. Dessa forma, a Tabela 5.5 mostra os modelos desenvolvidos sem restrições de defasagem, que usariam projeções de mais longo prazo na estimativa.

De modo geral, em termos de explicação da variabilidade da taxa de inadimplência, os modelos apresentaram variações relevantes. Foi incluída uma variável indicadora (*dummy*) entre 2020Q2 e 2020Q4 para indicar os efeitos das medidas implementadas pelo BCB para controle da pandemia da COVID-19, com efeito significativo para redução do NPL.

A Tabela 5.6 apresenta um segundo grupo de modelos, em que se buscou estimativas que em um cenário de uso de modelo pudessem usar projeções de curso prazo ou somente variáveis observadas em  $t-12$ .

Tabela 5.5: Modelos Desenvolvidos Utilizando Variáveis mais recentes

Variável	Modelo1	p-valor	Modelo2	p-valor	Modelo3	p-valor	Modelo4	p-valor
$c$	-0,5143	0,0000	-0,5510	0,0000	-0,6086	0,0000	-0,5023	0,0000
$NPL_{(-1)}$	0,5348	0,0000	0,5374	0,0000	0,4410	0,0000	0,5494	0,0000
$NPL_{(-2)}$	0,3053	0,0000	0,3083	0,0000	0,3923	0,0000	0,3065	0,0000
$PIB_g$					-2,9535	0,0000		
$PIB_{g(-1)}$	-1,9166	0,0001	-1,2422	0,0277			-1,7249	0,0013
$PIB_{g(-2)}$	-1,6558	0,0004			-1,5331	0,0049		
$d(UR)_{(-1)}$			4,2370	0,0014				
$Dummy$	-0,1424	0,0000	-0,1580	0,0000	-0,3894	0,0000	-0,1569	0,0000
$Cambio$			0,0143	0,0303	0,0198	0,0076		
$Cambio_{-1}$							0,0124	0,0715
$d(Selic)_{(-1)}$							1,6978	0,0126
$d(Selic)_{(-4)}$					1,8677	0,0093		
$R^2_{ajst}$	0,9820		0,9821		0,9838		0,9819	

Tabela 5.6: Modelos Desenvolvidos Utilizando Defasagens Maiores

Variável	Modelo 5	p-valor	Modelo 6	p-valor
$c$	-1,3759	0,0000	-1,2293	0,0000
$NPL_{(-3)}$	0,3232	0,0000	0,4054	0,0000
$NPL_{(-4)}$	0,2536	0,0000	0,2007	0,0001
$PIB_{g(-4)}$	-2,1134	0,0044	-1,7331	0,0026
$PIB_e$			-0,0231	0,0000
$d(UR)_{(-4)}$	4,0201	0,0269		
$R^2$ Ajustado	0,9681		0,9702	

O Modelo 6 considera, além das variáveis observadas em  $t-12$ , a expectativa de crescimento econômico como covariável, com efeito estatisticamente significativo e negativo sobre os NPL. Como é esperado intuitivamente, a capacidade do modelo em explicar a variável dependente, representada pelo  $R^2$ , reduz, conforme limita-se o uso das defasagens. Nestes modelos a *dummy* não apresentou efeito estatisticamente significativo de redução do NPL.

Além dos modelos com defasagens maiores, uma forma de avaliar a possibilidade de mitigação do risco de modelo associado ao uso de projeções nas estimativas da NPL, o uso de algoritmos mais robustos de *machine learning*, conforme discussões mencionadas nas seções anteriores deste capítulo, pode ser uma alternativa interessante. A seção seguinte traz uma aplicação, com um exemplo de técnica de interpretação de importância e sentido dos efeitos das variáveis, semelhante às possibilidades pelas equações apresentadas nesta seção.

Tabela 5.7: MAPE *Random Forest* por Profundidade das Árvores

<i>Max Depth</i>	MAPE	
	Desenv.	<i>out-of-time</i>
10	1,0037	10,1322
8	1,2429	10,0914
6	1,7277	9,6922
5	2,0527	9,4859
4	2,3568	9,1308
3	2,8406	9,1792
2	5,1479	10,5775

#### 5.4.4 Modelos de Machine Learning

Conforme mencionado na subseção 4.2.4, os modelos baseados em árvore envolvem a segmentação das covariáveis em uma série de regiões mais simples com o objetivo de fazer uma previsão para uma dada observação. Neste contexto, os algoritmos de otimização aplicados podem apresentar soluções locais, a depender do passo inicial do algoritmo.

Neste contexto, as *Random forests* [101], é um algoritmo de *machine learning*, em que é realizado uma combinação de árvores CART, em que a variável preditora é do tipo numérica. Utilizando como base os mesmos sentidos esperados sobre as variáveis estudadas e seus efeitos sobre o NPL, foi avaliado modelos ajustados com critério de partição a partir do erro quadrático médio. Para cálculo dos modelos foi utilizada a linguagem Python, com o pacote Scikit-Learn [102].

De modo geral, utilizando todas as variáveis e mesma segmentação das curvas de NPL dos modelos anteriores, os modelos *Random Forest* apresentou ajuste apenas na amostra de treino, não apresentando bons resultados para projeção dinâmica em períodos futuros. A Tabela 5.7 mostra o resultado do MAPE no período de desenvolvimento e de teste do modelo, com poda das árvores.

Foram testados modelos com variações das quantidades de árvores, e alterações entre as variáveis consideradas, tendo em vista sua importância. Para redução do *overfitting* do modelo aos dados de treino, foi controlada a profundidade das árvores. Conforme apresentado na Tabela 5.7. Utilizando as mesmas variáveis do Modelo 4, da subseção anterior, mesmo controlado a profundidade em cinco quebras, os MAPE com variáveis observadas passou de 2,36%, no período de treino, para 9,09% no período de teste.

A Figura 5.4 mostra o gráfico agregado dos SHAP *values*, calculado para melhor compreensão dos efeitos de cada variável. O ordenamento e a amplitude do gráfico indicam a importância da variável na predição, as partes em vermelho são valores elevados da covariável e azuis os menores valores da covariável.

Na Figura 5.4, valores elevados da série defasada dos NPL impactam a predição com

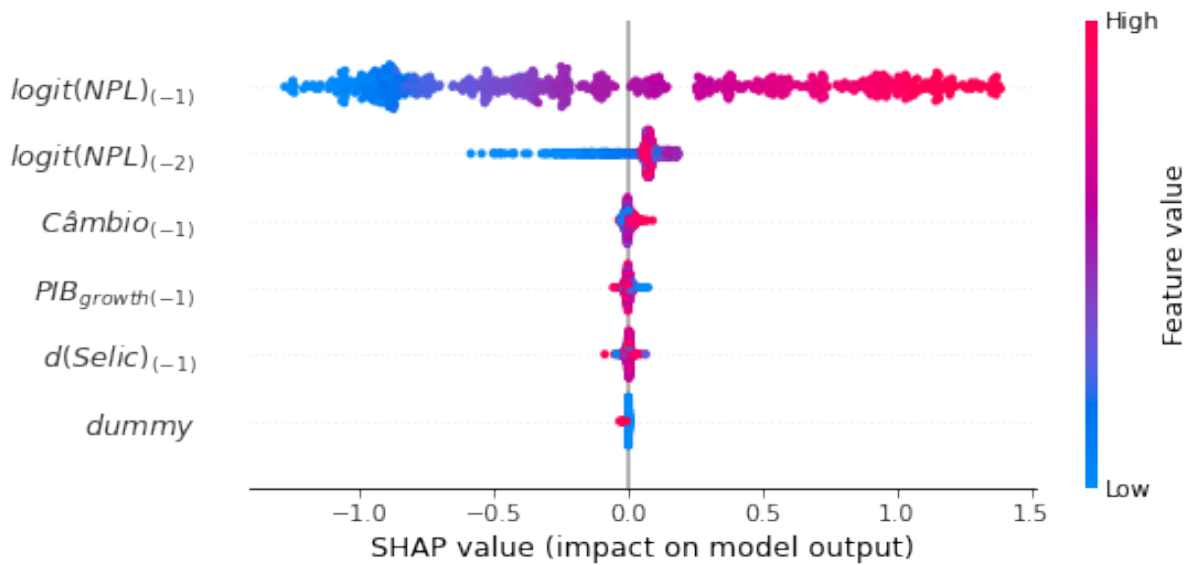


Figura 5.4: SHAP Value Modelo Random Forest - Variável Modelo 4 (Fonte: Elaborado Pelo Autor).

aumentos nos valores. Ao menos nas quatro variáveis mais importantes o sentido observado é o mesmo esperado pelo sentido econômico e o mesmo encontrado no Modelo 4. O efeito da Taxa Selic, não apresenta um padrão claro na análise gráfica e a *dummy* não apresentou impacto relevante nas projeções.

Considerando o *overfitting* do modelo ajustado e a concentração do efeito na defasagem da série, alinhado ainda ao objetivo de utilizar defasagens mais longas nas estimativas, uma alternativa foi o uso da série defasada em mais períodos, com o objetivo de avaliar se esta técnica de modelagem é uma alternativa para a redução do risco do uso de cenários projetados.

Após selecionar as variáveis de maior importância, obteve-se modelo com as variáveis  $\text{logit}(NPL_{(-4)})$ ,  $PIB_{growth}$ ,  $d(Tx_{Desmp.})$ ,  $Câmbio_{-2}$  e  $IPCA_{(-1)}$ . Neste contexto, com profundidade máxima das árvores do algoritmo em cinco níveis e regra de quebras baseadas no erro quadrático médio, observou-se na amostra de desenvolvimento MAPE de 3,26% e 8,19% no período de teste.

Vale destacar que o uso de quatro defasagens reduz a propagação dos erros provenientes da projeção dinâmica da série de NPL. A Figura 5.5 mostra o gráfico do SHAP para este modelo.

Utilizando as variáveis com maior importância do modelo, nota-se que o sentido econômico esperado pelas variáveis é observado mais claramente. Esta percepção serve de reforço às conclusões de estudos anteriores sobre os efeitos econômicos, bem como as equações destacadas nos modelos lineares desenvolvidos. No entanto, destaca-se que não



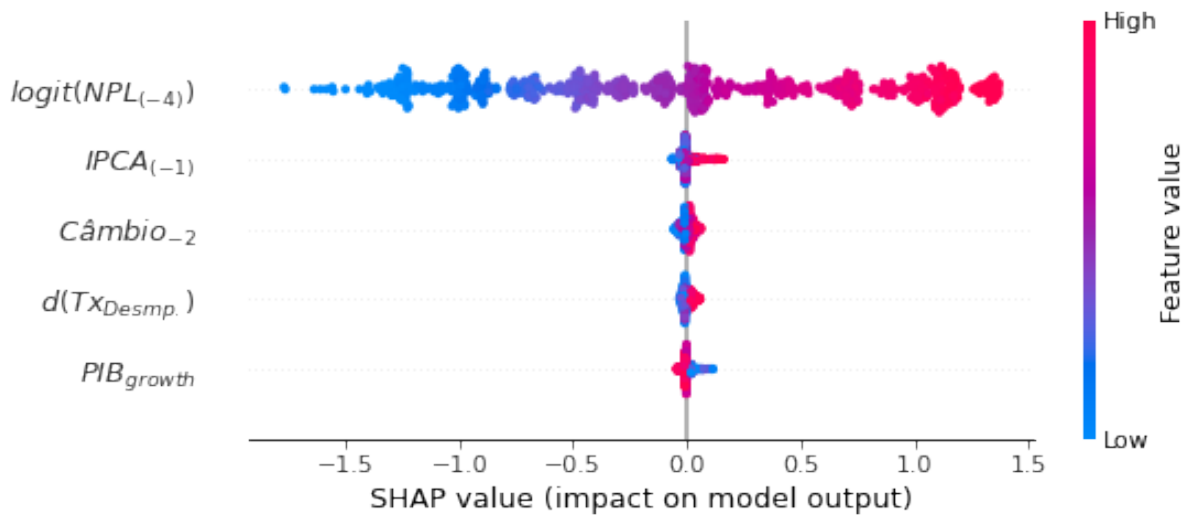


Figura 5.5: SHAP Value Modelo Random Forest Fin (Fonte: Elaborado Pelo Autor).

foi possível obter modelo que apresentasse desempenho melhor que os lineares utilizando este método de *machine learning* no período de teste da maneira como foi implementado.

#### 5.4.5 Avaliação dos Modelos

Para avaliação dos modelos, foi considerada a média erro percentual absoluto para comparação das taxas. Na Tabela 5.8, compara-se o desempenho dos modelos fora da amostra de desenvolvimento e com variáveis projetadas em 2020Q2.

Tabela 5.8: Avaliação dos Modelos

Modelo	MAPE			
	Desenv.	out-of-time	projetado	%dif
1	3,8064	7,7801	8,7130	11,99%
2	3,9918	7,5263	8,0542	7,01%
3	3,4379	8,2229	6,7040	-18,47%
4	3,8118	6,4023	6,6015	3,11%
5	3,9425	11,2325	-	-
6	3,7217	10,1737	-	-
7	3,2561	8,1900	8,0620	-1,56%

Dentre os modelos sem restrições de defasagens, o Modelo 3 teve menor taxa de erro na amostra de desenvolvimento, no entanto fica claro o *overfitting* nessa amostra, tendo em vista o maior crescimento na amostra fora do tempo, quando comparado aos demais. Justamente neste modelo, o efeito da utilização teve sentido contrário ao esperado intuitivamente, situação que pode ser atribuída à possíveis relações espúrias.

A 5.8 mostra os resultados de forma agregada em um indicador geral de erros do modelo. Para comparação considerando cada estimativa realizada nas unidades de cortes transversais e de tempo, foram aplicados os testes pareados de *T-student* e de Wilcoxon [103], sendo que o primeiro, pressupõe distribuição normal das amostras para comparação das médias e, o segundo foi utilizado como suporte em um teste não paramétrico de igualdade das medianas, ambos para avaliar se a diferença identificada tem significância estatística, testando-se a hipótese indicada na Seção 6 deste trabalho. A Tabela 5.9 apresenta os resultados dos dois testes.

Tabela 5.9: Teste *T-student* e Wilcoxon Pareados dos Erros dos Modelos

Modelo	<i>t-student</i>		Wilcoxon	
	Estatística	p-valor	Estatística	p-valor
1	-5,6376	0,0001	596,5	0,0001
2	-7,2947	0,0001	767,5	0,0001
3	4,8825	0,9999	3540,0	0,9999
4	-3,3333	0,0006	781,5	0,0014
7	1,8088	0,9632	764,0	0,9341

Para aplicação dos testes foi utilizado o pacote SciPy [104], do Python. Nos resultados das hipóteses testadas, se mantém o padrão identificado na 5.8, em que no Modelo 3 e no Randon Forest, não foi rejeitada a hipótese de aumento dos erros pelo uso de projeções das covariáveis nas projeções do NPL. Não houve diferenças nos resultados dos testes *T-student* ou Wilcoxon.

Nos demais modelos há aumento da taxa de erro entre 3,11% e 11,99%. Destaca-se que no modelo 4, que traz as variáveis PIB, Câmbio e Selic, houve melhor desempenho fora da amostra de desenvolvimento. Apesar dos aumentos das taxas de erro, os modelos alternativos não apresentaram desempenho satisfatório fora da amostra de desenvolvimento, indicando que as relações identificadas, apesar de coerentes com o sentido econômico, não mantém o efeito dos parâmetros no período testado. Vale destacar que a expectativa do PIB melhorou o desempenho do modelo tanto na avaliação do  $R^2$  ajustado, quanto no teste fora da amostra de desenvolvimento.

Portanto, dentre as possibilidades apresentadas neste trabalho, a melhor opção para uma estimativa realizada com as informações disponíveis sobre este público em 2020Q2 teria sido o uso de projeções no lugar das covariáveis.

## 5.4.6 Conclusão

O objetivo deste trabalho foi estimar os efeitos das variáveis macroeconômicas sobre o NPL nas instituições bancárias brasileiras para avaliação dos efeitos do uso de projeções utilizando cenários macroeconômicos sobre o desempenho dos modelos utilizados.

Foram estimados modelos econométricos para dados em painel utilizando dados do público PF segmentados por modalidade de crédito e porte de cliente. Conforme esperado intuitivamente e de acordo com os estudos disponíveis da literatura, nos diferentes modelos estimados há efeitos estatisticamente significativos e diretamente proporcionais da taxa de desemprego, taxa de câmbio, taxa Selic, além da própria série de NPL defasada sobre a NPL. O crescimento do PIB tem influência inversamente proporcional sobre a NPL.

Dos quatro modelos desenvolvidos, um dos modelos teve uma taxa de erro inferior com o uso de variáveis projetadas, no entanto este foi o modelo que teve aumento mais expressivo quando testado fora da amostra de desenvolvimento, o que pode explicar este resultado. Os três que apresentaram melhor desempenho no período de projeção apresentaram aumento, entre 3% e 12%, nas medidas de erro ao se utilizar variáveis projetadas disponíveis na data da estimativa, para o público e período estudados.

Apesar desse aumento observado, neste período, as projeções com modelos com uso de defasagens mais longas ou expectativas como covariável apresentaram taxas de erro mais elevadas. Dessa forma, entende-se que, apesar da influência negativa sobre o desempenho do modelo neste período, dentre as opções testadas, as melhores estimativas possíveis seriam com utilização de projeções.

Destaca-se que este resultado pode variar com a repetição deste estudo em outros períodos ou outros públicos, ou mesmo com particularidades de cada Instituição Financeira. Portanto, indica-se que este tipo de análise, sobre a confiabilidade das covariáveis utilizadas, faça parte do monitoramento dos modelos utilizados pelas instituições em cada uma de suas carteiras, de acordo com as metodologias aplicadas nas projeções, a fim de evitar decisões inadequadas sobre as distribuições de resultado ou do crédito.

Este trabalho tem algumas limitações. O tamanho da amostra de 31 períodos pode ser considerado pequena para o uso do método aplicado. No entanto, a avaliação de desempenho fora da amostra é uma forma de minimizar tal limitação. A aplicação de projeções da inadimplência, usualmente, deve considerar créditos não inadimplentes, dos estágios 1 e 2, no contexto do IFRS 9. No entanto, tal abordagem dependeria de dados mais detalhados sobre históricos detalhados de créditos que se tornam inadimplentes, além disso, entende-se que esta limitação não enfraquece as conclusões deste trabalho sobre o uso de cenários, considerando a vasta literatura associando cenários econômicos ao NPL.

Os resultados obtidos ficam limitados ao público e período estudados, o que pode ser objeto de estudos futuros para resultados mais abrangentes sobre o efeito do uso dos cenários, com maior disponibilidade de dados e ajustes de modelos e projeções em diferentes períodos.

# Capítulo 6

## Cálculo da Perda Esperada de Crédito

### 6.1 Introdução

Atualmente, no mercado brasileiro, a Resolução nº 2682, de 21 de dezembro de 1999 do Conselho Monetário Nacional [10], estabelece regras para as instituições financeiras classificarem suas operações de crédito em níveis e riscos. Esta classificação deve levar em consideração informações sobre o devedor, garantias e sobre a operação e, a partir desta categorização, Provisão para Créditos de Liquidação Duvidosa (PCLD).

São fixados critérios de revisão mensal das classificações de risco, em função dos atrasos das operações de crédito. Nestes critérios, ficam estabelecidos critérios de revisão mensal das classificações de risco, em função dos atrasos das operações de crédito. A Tabela 6.1 apresenta a relação dos níveis de riscos e os atrasos e percentual de provisionamento para cada um dele.

Tabela 6.1: Classificações de Risco Resolução nº 2682

Nível de Risco	Atraso	PCLD
AA	-	0,0%
A	0 a 14 dias	0,5%
B	15 a 30 dias	1,0%
C	31 a 60 dias	3,0%
D	61 a 90 dias	10,0%
E	91 a 120 dias	30,0%
F	121 a 150 dias	50,0%
G	151 a 180 dias	70,0%
H	mais de 180 dias	100,0%

Para todo o valor da carteira classificado com o nível de risco, deve-se constituir PCLD com o percentual apresentado. A Resolução nº 2682 estipula, ainda, que, após 180 dias classificada em nível de risco H, as operações de crédito podem ser contabilizadas como prejuízo.

As regras vigentes estão em período de transição para compatibilização com o Padrão Internacional [12], conforme disposição da Resolução CMN Nº 4.966 de 25 de novembro de 2021 [18], o que traz relevância sobre a compreensão das diferenças entre os cálculos de riscos inerentes a esta dinâmica do crédito que, segundo Wenzler [105], é fator preponderante para os quais é possível precifica-los de forma acurada, desde que se adote o modelo estrutural adequado. Com o uso de uma modelagem correta é possível a redução de riscos e incertezas, sendo fato necessário a ser levado em conta em toda análise acerca desta dinâmica.

Tendo em vista o disposto do Art. 78 da Resolução CMN Nº 4.966, as IFs devem divulgar notas explicativas nas demonstrações financeiras do exercício de 2024 com os impactos estimados pela implementação da nova regulação contábil.

Diante deste contexto de alteração das normas contábeis, o trabalho deste capítulo tem o objetivo de simular a dinâmica de uma carteira de crédito seus atrasos e provisionamento para comparar as abordagens de cálculo vigente (Resolução CMN Nº 2.682) e o padrão IFRS 9 (Resolução CMN Nº 4.966).

## 6.2 Referencial Teórico

Vários estudos realizados a partir do uso de técnicas de simulação para avaliações dos riscos associados ao mercado de crédito podem ser encontrados na literatura. Schwab [44] apresenta um estudo de caso, aplicado a empresas suíças, para explorar o impacto do setor financeiro sobre o crescimento de empresas de médio e pequeno porte. A partir da combinação de modelos *system dynamics* e *agent-based*, para descrever, respectivamente, comportamentos do sistema e dos agentes individuais, os autores concluem que os limites concedidos pelos bancos e as regulações bancárias na relação entre essas empresas e os bancos.

O estudo de Costa [42] foi realizado com objetivo de dar clareza sobre a dinâmica de lucratividade das *fintechs*, considerando as particularidades do gerenciamento de risco de crédito no mercado brasileiro. O sistema de simulações desse trabalho foi separado em três partes, sendo elas: simulação de aquisição, manutenção de clientes e seus custos associados; simulação do sistema de atrasos e recuperações nas operações dos clientes; e contabilização em que são considerados os resultados para o cálculo das principais receitas e despesas.

Relativamente à classificação das Operações de crédito por Nível de Risco e Provisão, o COSIF [106], descreve:

“As instituições financeiras e demais instituições autorizadas a funcionar pelo Banco Central do Brasil devem classificar as operações de crédito, em ordem crescente de risco, nos seguintes níveis: nível AA; nível A; nível B; nível C; nível D; nível E; nível F; nível G e nível H. Res 2682 art 1º I/IX.”

Além disso, destaca a forma de contabilização dos provisionamentos relativos às operações de crédito, destacando a forma de impacto dos provisionamentos no resultado financeiro das IFs como:

"A provisão para créditos de liquidação duvidosa deve ser constituída sobre o valor contábil dos créditos mediante registro a débito de DESPESAS DE PROVISÕES OPERACIONAIS e a crédito da adequada conta de provisão para operações de crédito. No caso de insuficiência, reajusta-se o saldo das contas de provisão a débito da conta de despesa.”

Sobre as despesas de provisão, conforme descrito trabalho de Caneca [11], baseado na Resolução Nº 2.682 [10] e nas notas explicativas dos bancos brasileiros, é calculada pela variação da provisão mais os valores contabilizados como prejuízo no período, conforme descrito na Equação 6.7.

$$DPCLD_{i,t} = \Delta A_{i,t} + \Delta B_{i,t} + \Delta C_{i,t} + \Delta D_{i,t} + \Delta E_{i,t} + \Delta F_{i,t} + \Delta G_{i,t} + \Delta H_{i,t} + PRJ_{i,t} \quad (6.1)$$

onde,  $DPCLD_{i,t}$  representa a despesa com PCLD no período,  $\Delta A_{i,t} - \Delta A_{i,t}$  são as variações da provisão em cada nível de risco e  $PRJ_{i,t}$  é o total de saldo contabilizado como prejuízo.

A Resolução CMN Nº 4.966 [18] caracteriza os grupos homogêneos de risco como:

"conjunto de instrumentos financeiros com características semelhantes que permitam a avaliação e a quantificação do risco de crédito de forma coletiva, considerando:

I - as características de risco de crédito da contraparte; II - as características de risco de crédito do instrumento, considerando a modalidade do instrumento e o tipo de garantias ou colaterais relacionados com o instrumento, quando existentes;”

Neste trabalho, alinhado a este conceito definiu-se os grupos homogêneos de risco a partir do porte da contraparte e da modalidade do instrumento financeiro.

Sobre a alocação dos grupos homogêneos de risco, a resolução Nº CMN 4.966 [18] destaca que a definição dos grupos deve ser revisada periodicamente e a alocação dos instrumentos nos grupos deve ser revisada mensalmente, devido aos atrasos e no mínimo anualmente nos demais casos.

## 6.3 Metodologia

O trabalho deste Capítulo caracteriza-se como um estudo de caso, utilizando técnicas de simulação para avaliar a dinâmica de provisionamento de créditos em IFs, aplicado ao contexto brasileiro. Para definição do sistema e das variáveis do modelo e avaliação de suas saídas, baseou-se nas variáveis do sistema utilizadas na etapa de atrasos e recuperações do sistema apresentado em estudos anteriores [42][45], além de dados dos dados de operações de crédito do SFN, obtidos conforme descrito no Capítulo 5.

As etapas seguidas para elaboração do trabalho apresentado neste capítulo foram:

- definição da dinâmica de carteiras de crédito a ser utilizada;
- calibragem dos parâmetros com valores do SFN;
- cálculos de parâmetros de risco para ECL [12] e PCLD [10].

A partir dos resultados obtidos após as etapas descritas, foi possível comparar o comportamento das metodologias em cenários de inadimplência próximos ao atual, além de avaliar a sua capacidade de cobertura para perdas de crédito sensibilidade das estimativas a choques na inadimplência. Para realizar os procedimentos descritos neste capítulo, foi utilizada a linguagem Python [93] e os dados simulados foram armazenados em banco de dados Postgre SQL.

## 6.4 Resultados

### 6.4.1 Dinâmica da carteira de crédito

Pode-se dividir o procedimento de simulação em quatro etapas: simulação da carteira de clientes; contratações diárias de operações de crédito; atualização da base de operações de crédito e cálculos dos indicadores. A Figura 6.1 mostra o modelo conceitual da simulação da carteira de crédito realizada para este trabalho.

Para simulação dos dados de crédito, como primeiro passo, foi simulada uma carteira de clientes com os mesmos seis portes designados no capítulo 5. A atribuição do porte do cliente seguiu uma distribuição polinomial com parâmetros 41,9% para o porte 1, 15,6% para o porte 2, 16,7% para o porte 3, 14,6% para o porte 4, 7,2% para o porte 5 e 4,1% para o porte 6. Utilizando a linguagem orientada a objetos Python [93], cada objeto da classe "cliente" simulado, recebeu os atributos: código do cliente, porte, probabilidade de inadimplência em 30 dias e um multiplicador para o valor da operação, sendo que os dois últimos atributos dependem do porte.

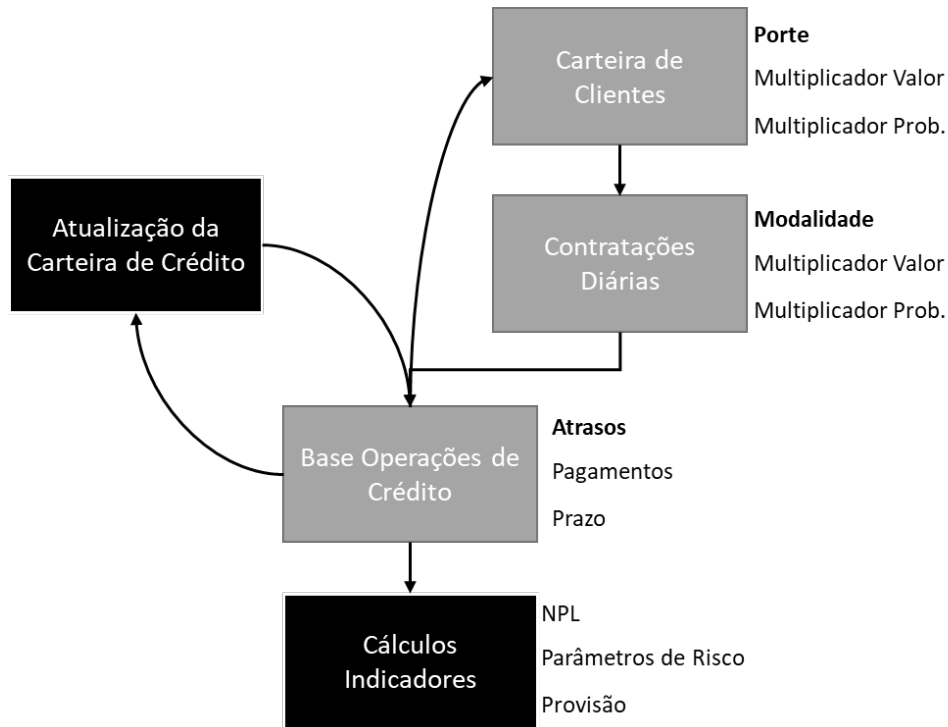


Figura 6.1: Simulação da Dinâmica de Atrasos (Fonte: Elaborado Pelo Autor).

A partir destes resultados, foram simulados objetos da classe "operação", com atributos, data de referência, cliente, modalidade, valor, probabilidade de não pagamento da operação, além de informações sobre pagamentos das parcelas. A atribuição da modalidade foi obtida a partir de uma distribuição Multinomial, a probabilidade de pagamento da operação foi obtida a partir da probabilidade do cliente e um multiplicador resultante da modalidade da operação. O valor da operação foi simulado a partir de uma distribuição Lognormal.

Foram definidas funções para contratações diárias de operações de crédito, dos clientes da carteira simulada em quantidade seguindo uma distribuição Poisson, com média de 100 contratações diárias. Posteriormente, foi construído procedimento para atualização mensal das operações contratadas e contratação de novas operações. De acordo com a situação de atraso da operação de crédito, foi definida probabilidade de pagamento ou agravamento dos atrasos a cada mês simulado. A Figura 6.2 ilustra as regras de migração das operações de crédito utilizadas nesse trabalho.

Os percentuais de migração para cada uma das situações de atraso foram calibrados com o objetivo de manter indicadores próximos aos do SFN, assunto discorrido na subseção seguinte.



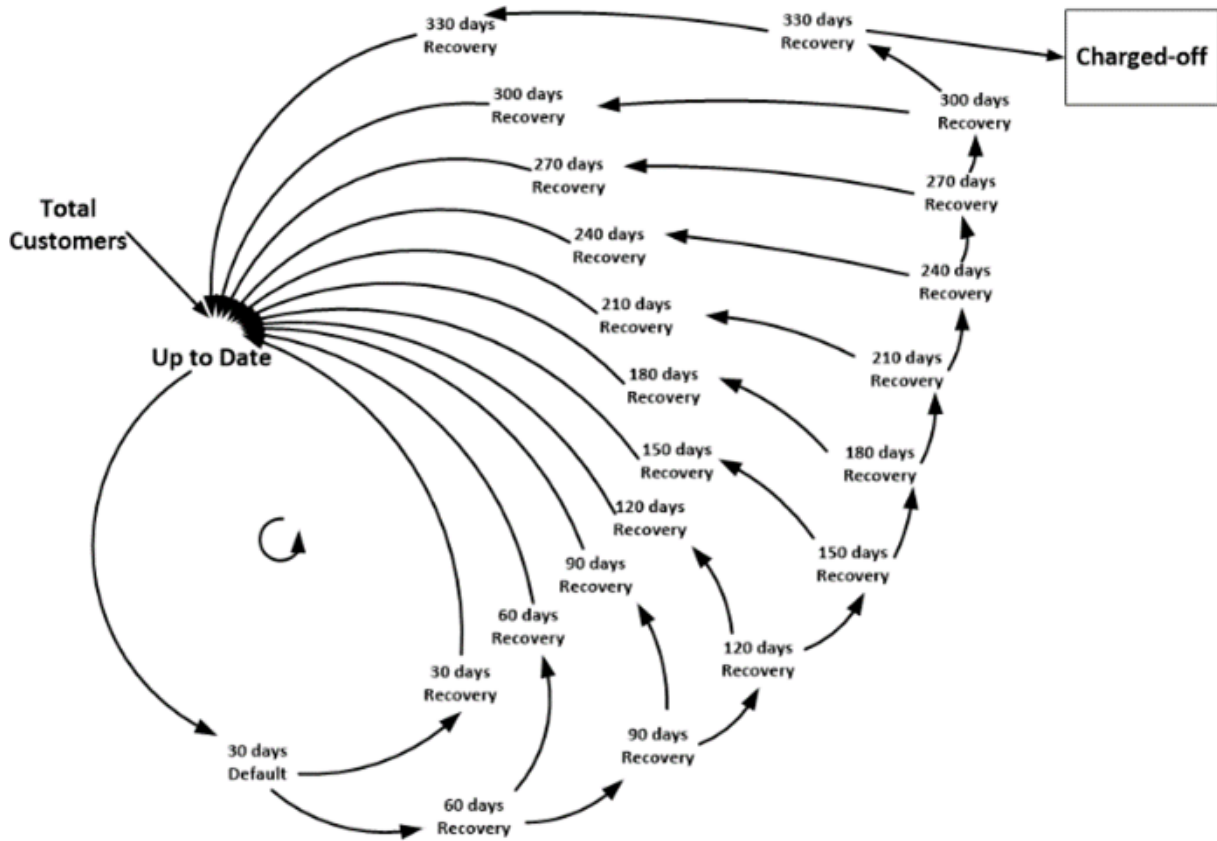


Figura 6.2: Simulação da Dinâmica de Atrasos (Fonte: Costa [42]).

### 6.4.2 Calibragem dos Parâmetros

Para calibragem dos parâmetros do sistema foram consideradas as características da carteira de crédito do SFN. Os dados foram obtidos a partir do Painel de Operações de Crédito [89] e segmentados conforme Tabela 5.1. As características calibradas foram: valor médio das operações de crédito, percentual de representatividade na carteira total e a NPL.

Considerando o modelo conceitual apresentado para a dinâmica de atualização da carteira de crédito, os parâmetros ajustáveis para alinhamento das características foram:

- probabilidade de atraso até 30 dias;
- matriz  $A_{13 \times 13}$  com probabilidades de atraso e recuperação para cada faixa de atraso de 30 dias;
- vetores  $p$  e  $m$  que representam multiplicadores de ajuste da probabilidade de atraso referentes, respectivamente, ao Porte do cliente e à Modalidade da operação;
- vetor  $r$  de multiplicador de recuperações de operações atrasadas devido à modalidade da operação; e

- vetores  $q$  e  $n$  que representam multiplicadores de ajuste do valor referentes, respectivamente, ao Porte do cliente e à Modalidade da operação;

No Apêndice B é apresentado o detalhamento dos códigos Python utilizados para os procedimentos de simulação com ajuste final dos parâmetros.

A partir destes procedimentos, foi possível construir dados detalhados de fluxos de pagamentos mensais de contratos de crédito das quatro modalidades de crédito, feitos por clientes de seis portes de renda, conforme segmentação utilizada no Capítulo 5. A Figura 6.3 apresenta um exemplo de um contrato simulado com suas atualizações mensais.

contrato character varying (20)	ref integer	prazo double precision	porte integer	modalidade integer	valor numeric	pgto numeric	atraso integer
00005000011	0	18	5	2	11987.57	0.00	0
00005000011	1	17	5	2	11376.41	731.03	0
00005000011	2	16	5	2	10759.15	731.03	0
00005000011	3	15	5	2	10135.72	731.03	0
00005000011	4	14	5	2	9506.05	731.03	0
00005000011	5	13	5	2	8870.08	731.03	0
00005000011	6	12	5	2	8227.76	731.03	0
00005000011	7	11	5	2	7579.01	731.03	0
00005000011	8	10	5	2	6923.77	731.03	0

Figura 6.3: Exemplo de Contrato Simulado (Fonte: Elaborado Pelo Autor).

Após ajustes dos percentuais de transição entre os estados de atraso, os resultados obtidos na simulação foram os apresentados na Tabela 6.2.

A Tabela 6.2 apresenta resultados de medidas dos créditos do SFN em comparação às obtidas pelas operações detalhadas produzidas pelas simulações. As informações comparadas foram a os valores médios das operações, o percentual do grupo na carteira e os NPL, todos calculados pelos segmentos e modalidades utilizados no Capítulo 5. Pela análise destes valores, nota-se que os valores obtidos pelas simulações são semelhantes aos observados no SFN, o que viabiliza a utilização destes dados para testes comparativos das abordagens de cálculo da ECL.

A partir das bases de operações de crédito geradas, foram calculados parâmetros de risco PD e LGD, além de aplicados os percentuais de provisionamento vigentes [10]. A subseção seguinte traz os resultados das comparações entre as metodologias na carteira simulada.

### 6.4.3 Cálculos da ECL e PCLD da Carteira Simulada

As simulações foram construídas com posições de uma carteira de crédito atualizada mensalmente, com um total de 200 meses simulados. Os dados simulados podem ser divididos

Tabela 6.2: Dados Simulados x SFN - Jun/2021

Mod.	Porte	Saldo Médio (R\$)		%Carteira		NPL	
		SFN	Sim.	SFN	Sim.	SFN	Sim.
1	1	1.331	1.385	6,29%	6,26%	13,63%	10,42%
	2	1.964	2.082	2,23%	2,29%	11,47%	8,15%
	3	2.587	2.770	2,31%	2,46%	9,64%	7,32%
	4	4.696	4.996	1,78%	2,15%	7,48%	7,11%
	5	8.472	8.244	0,70%	1,05%	5,11%	5,28%
	6	42.388	19.291	0,27%	0,49%	2,44%	2,43%
2	1	816	819	36,48%	33,92%	7,75%	6,49%
	2	1.232	1.257	12,70%	12,42%	6,24%	5,25%
	3	1.701	1.694	14,27%	12,95%	4,30%	4,17%
	4	2.763	3.041	10,22%	11,21%	3,38%	3,87%
	5	4.612	5.332	4,20%	5,08%	2,18%	3,20%
	6	10.619	11.218	2,24%	2,58%	1,05%	1,77%
3	1	14.742	9.774	0,98%	1,86%	4,18%	3,83%
	2	15.884	15.352	0,83%	0,67%	3,61%	3,89%
	3	20.897	22.070	1,00%	0,73%	2,89%	2,37%
	4	28.106	27.470	0,64%	0,63%	2,03%	2,23%
	5	36.850	29.479	0,21%	0,29%	1,54%	0,85%
	6	56.290	32.441	0,09%	0,13%	0,96%	1,01%
4	1	72.806	43.468	1,18%	1,21%	2,91%	2,51%
	2	88.415	64.911	0,39%	0,42%	1,83%	2,17%
	3	103.631	87.561	0,37%	0,48%	1,25%	2,67%
	4	140.946	152.899	0,34%	0,41%	0,79%	1,33%
	5	202.270	261.119	0,18%	0,18%	0,50%	1,09%
	6	346.176	355.180	0,09%	0,11%	0,41%	0,00%
Total		5391	5494	100,00%	100,00%	2,95%	2,88%

em três etapas consideradas neste estudo, sendo: a primeira fase de estabilização das taxas, decorrentes das taxas menores nos estados iniciais com operações recém contratadas; fase de choque na inadimplência, entre os meses de número 72 e 94, em que foi aplicada um choque de 15% nas taxas de atraso; e a terceira fase de estabilização, em que foi retirado choque e há uma estabilização da NPL.

Para cada período foram calculados os parâmetros de risco PD, como a média em período estipulado das taxas de descumprimento calculadas em safras de 12 meses de cada carteira.

$$TD_{t,i,j} = \frac{\sum D_{t,i,j}}{n_{t,i,j}} \quad (6.2)$$

Sendo  $i$  a modalidade da operações,  $j$  o porte do cliente,  $D_{i,j}$  variável indicadora de até 90 dias no período de 12 meses para o cliente do contrato na modalidade  $i$  e  $n$  a

quantidade de operações da modalidade  $i$  e porte  $j$ .

A Figura 6.4 mostras as taxas de descumprimento, calculadas em safras de 12 meses, para os períodos simulados.

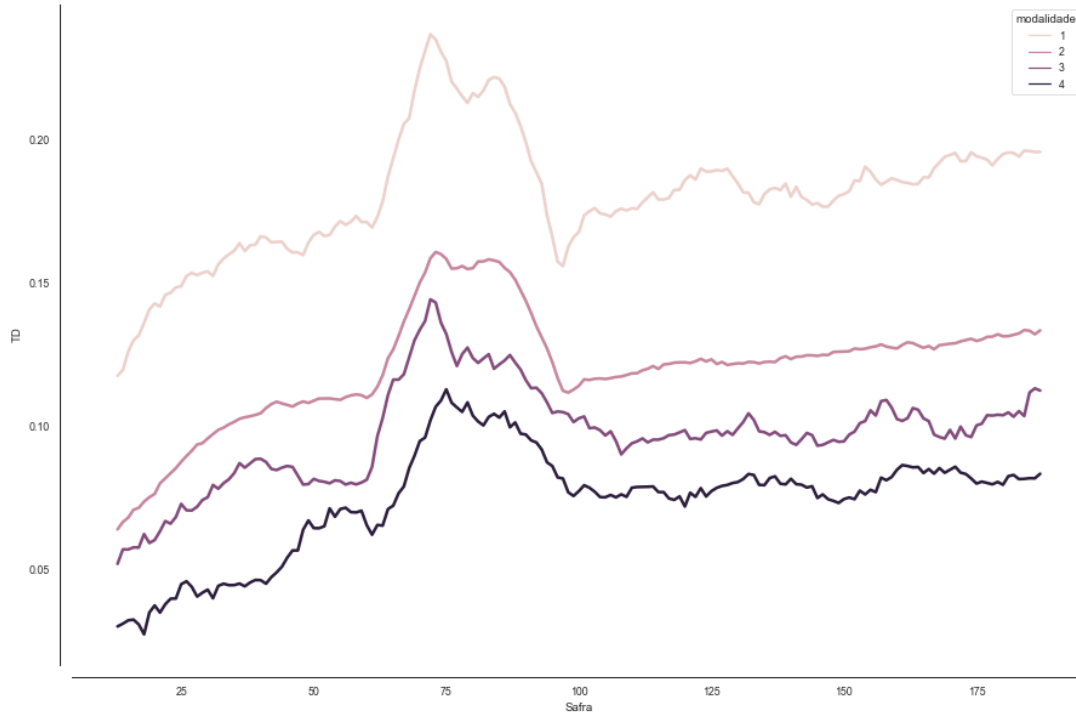


Figura 6.4: Taxas de Descumprimento Médias Simuladas por Modalidade (Fonte: Elaborado Pelo Autor).

A partir destas taxas foi calculada a  $PD_{i,j}$  para cada modalidade  $i$  e porte  $j$ , a partir da equação 6.2.

$$PD_{i,j} = \frac{\sum TD_{t,i,j}}{N} \quad (6.3)$$

Sendo  $N$  a quantidade dos períodos utilizados para cálculo da PD.

A LGD, foi calculada como a média das  $LGD_{t_0}$  para cada modalidade  $i$  conforme equação 6.4 [107].

$$LGD_{i,t_0} = 1 - \frac{\sum R_{i,j} - \sum P_{i,j}}{EAD_i} \quad (6.4)$$

sendo:  $EAD_i$  a exposição da modalidade  $i$  no momento do descumprimento,  $R_{i,j}$  os pagamentos recebidos para o ativo  $i$  em cada período  $j$  após o descumprimento e  $P_{i,j}$  os custos de recuperação, ambos corrigidos a valor presente no momento do descumprimento pela taxa efetiva de juros da simulação dos empréstimos.

Para ajuste *lifetime* da PD, foi utilizada o *Maturity adjustment*  $m$ , descrito em Basileia II [3], conforme Equações 6.5.

$$m = \frac{(1 + (M - 2,5) \times b)}{(1 - 1,5 \times b)} \quad (6.5)$$

sendo  $b$ :

$$b = (0,11852 - 0,05478 \times \ln(PD))^2 \quad (6.6)$$

e  $M$  o prazo, em anos, da operação de crédito.

A Figura 6.5 mostra a comparação das curvas dos índices de ECL e de Provisionamento baseado e da NPL nas regras da Resolução CMN 2.682.

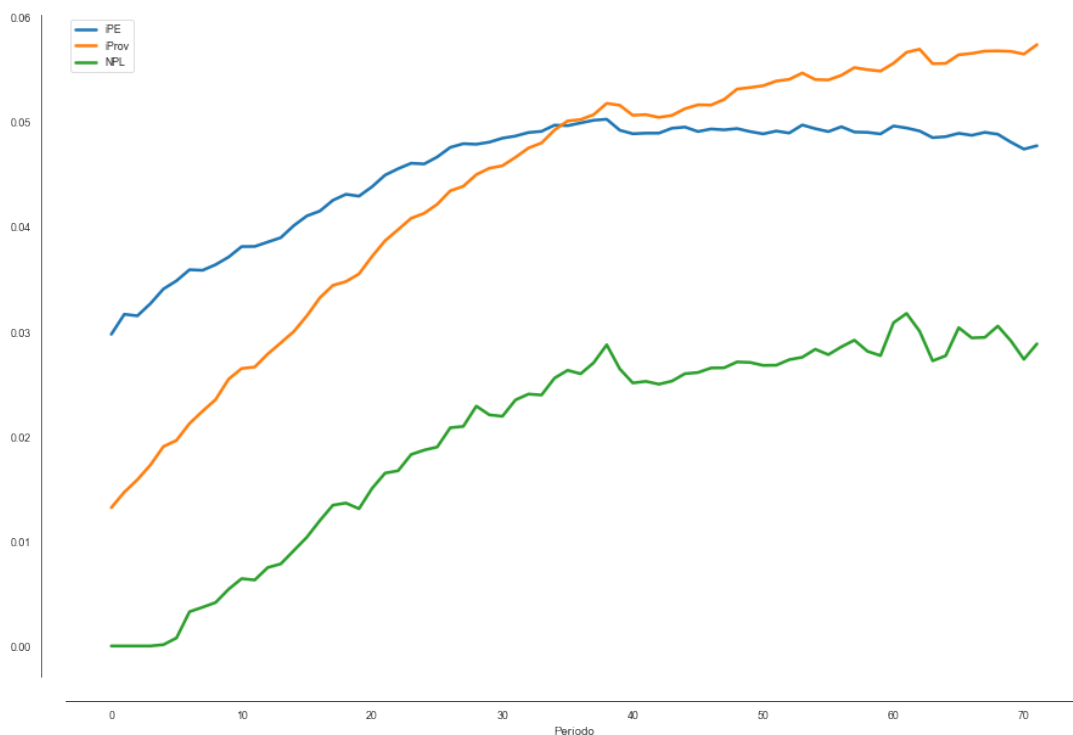


Figura 6.5: Curvas de NPL, iPE e Iprov dos Dados Simulados (Fonte: Elaborado Pelo Autor).

As curvas apresentadas na Figura 6.5 consideram apenas a primeira etapa de estabilização dos parâmetros simulados, a subseção seguinte traz a resposta das metodologias ao choque de 15% incluído entre os períodos 72 e 96 da simulação.

#### 6.4.4 Sensibilidade das Metodologias

Em ambas as Resoluções sobre os cálculos da provisão de crédito, admite-se adoção de critérios internos para cálculo das provisões em estágios iniciais, e ajustar estes critérios para

antecipação das perdas esperadas por mudanças no cenário econômico. As comparações realizadas neste capítulo não consideram as projeções e expectativas de perdas futuras, apenas a resposta da metodologia de cálculo implementada ao choque da inadimplência.

A Figura 6.6 apresenta a resposta das curvas dos indicadores de provisionamento e o indicador NPL.

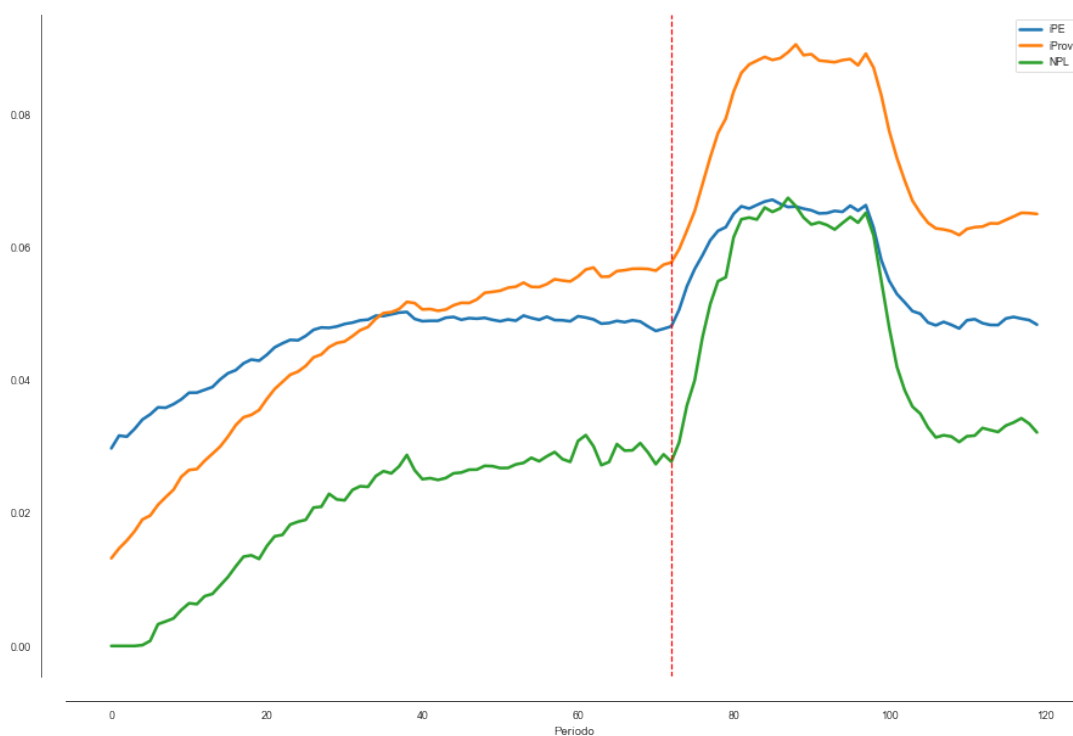


Figura 6.6: Sensibilidade dos Cálculos a Variações da Inadimplência (Fonte: Elaborado Pelo Autor).

Depreende-se da Figura 6.6 que as duas formas de cálculo respondem ao aumento da NPL. Para o caso da classificação em níveis de risco da Resolução 2.682, o aumento é decorrente dos aumentos dos percentuais de provisionamento resultante dos atrasos das operações. Já na aplicação dos parâmetros de risco, este aumento resulta, principalmente, da migração de operações para o terceiro estágio de classificação dos ativos, em que a PD recebe o valor de 100%.

A Resolução CMN 4.966 destaca a necessidade de revisão da alocação dos instrumentos financeiros nos grupos homogêneos de risco e da validação dos parâmetros utilizados no cálculo da ECL, no entanto, não fica especificado os prazos de recálculo dos parâmetros obtidos. Dessa forma, é possível que uma IF estabeleça gatilhos de atualização dos parâmetros, conforme os monitoramentos sejam realizados.

Dessa forma, considerando a defasagem de 12 meses necessária para apuração do parâmetro PD, por exemplo, a IF poderia avaliar a necessidade de revisão deste parâmetro

após o aumento de inadimplência. Para simulação deste cenário, calculou-se a atualização do parâmetro utilizando as taxas de descumprimento entre os períodos 72 e 84. A Figura 6.7 mostra a atualização dos valores de ECL com a atualização da PD a partir do período 96, em que seria possível sua apuração.

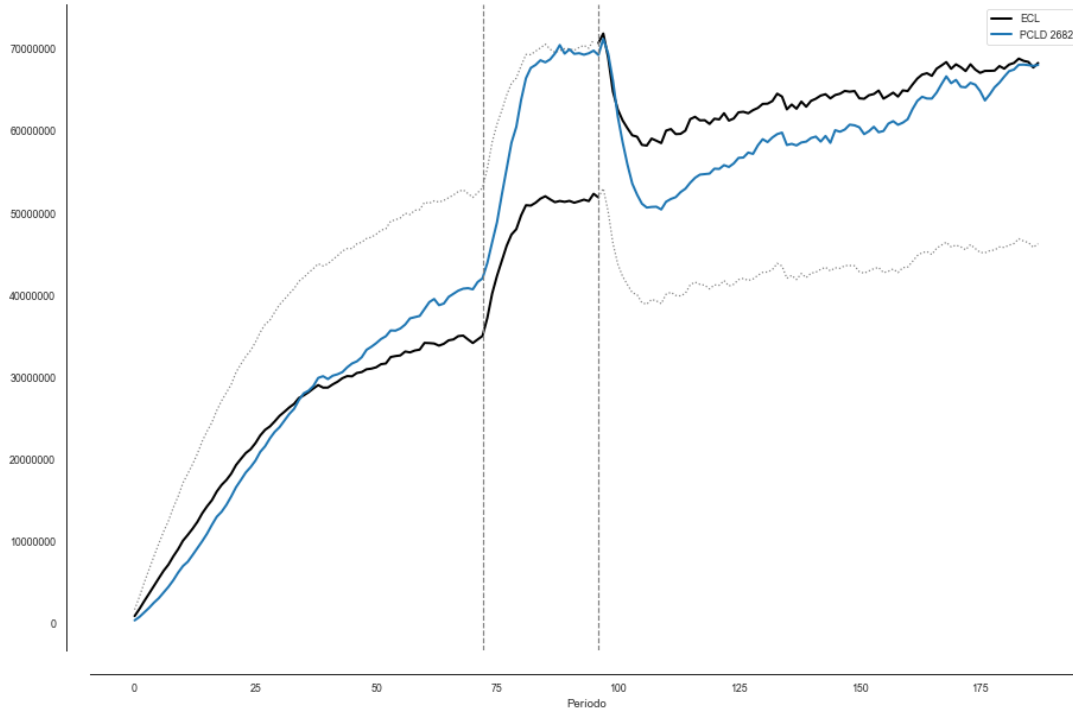


Figura 6.7: Impacto do Choque na atualização da ECL com Gatilho de Atualização (Fonte: Elaborado Pelo Autor).

Em um cenário como este, a variação resultante da atualização da PD elevaria a ECL a valores próximos aos da provisão calculada pelos níveis de provisionamento da Resolução 2.682, no entanto em apenas um período nas demonstrações contábeis.

Considerando o choque aplicado, esta variação de nível de provisionamento traria impacto financeiro à IF a partir das despesas de PCLD, conforme descritas por Caneca [11].

Denota-se da equação 6.7 que a despesa de provisão, independentemente da forma de cálculo do saldo de provisão total, pode ser obtida a partir da equação:

$$DPCLD_t = \Delta PE_{t,t-1} + PRJ_t \quad (6.7)$$

onde,  $DPCLD_{t,t-1}$  representa a despesa com PCLD no período,  $PE_t$  é a variação da provisão na carteira de crédito do período  $t$  em relação ao  $t-1$ , neste trabalho considerado trimestralmente, e  $PRJ_t$  é o total de saldo contabilizado como prejuízo no período.

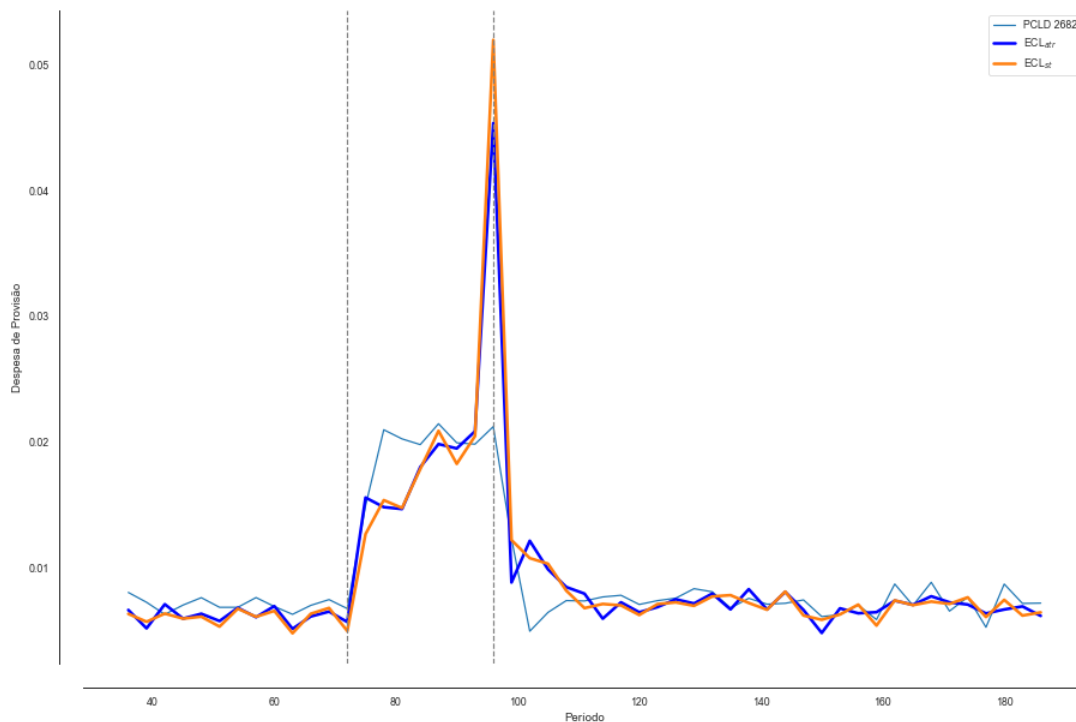


Figura 6.8: Impacto do Choque nas Despesas de Provisão (Fonte: Elaborado Pelo Autor).

A Figura 6.8 apresenta a evolução das despesas de provisão neste cenário.

Como pode ser observado pela Figura 6.8, o efeito da atualização da PD em um único período da série simulada tem como consequência uma despesa de PCLD elevada. Em situações como essas, os impactos para a IF podem variar desde a percepção negativa pelo mercado financeiro por uma demonstração de prejuízo, necessidade de uso do capital para cobertura de riscos ou mesmo problemas de liquidez resultantes da distribuição de resultados dos períodos de provisão reduzida.

Vale destacar que a exigência, atribuída pela Resolução CMN 4.966, de reavaliação dos grupos homogêneo de risco, mensalmente, em função da situação de atraso da operação, tende a mitigar esse efeito, considerando que a PD aumenta conforme avança o atraso da operação de crédito. Na Figura 6.8 são apresentadas duas linhas de ECL, sendo considerada a transição entre grupos homogêneos de risco em função do atraso e outra sem esta regra.

Nota-se que esta exigência reduz o impacto da despesa de provisão devido à não antecipação das perdas e atualização do parâmetro de risco. Apesar disso, observa-se um efeito relevante na situação apresentada na simulação.

Uma alternativa para mitigação deste evento de risco é a utilização dos modelos de ajuste *Forward looking*. No entanto, nem sempre todas as variáveis responsáveis pelo agravamento do risco de crédito estão nos modelos de projeção das taxas de inadimplência.



Portanto, a atualização dinâmica da PD é uma alternativa simples e que também pode mitigar este risco.

### 6.4.5 Atualização Dinâmica dos Parâmetros de Risco

A atualização dinâmica dos parâmetros de risco pode ser uma alternativa para estabelecer uma medida contemporânea da situação de risco de crédito das IFs. A Figura 6.9 apresenta o resultado das despesas de PCLD em considerando a atualização da última taxa de descumprimento disponível ( $t - 12$ ) e das médias móveis das últimas 12 e 24 taxas de descumprimentos disponíveis.

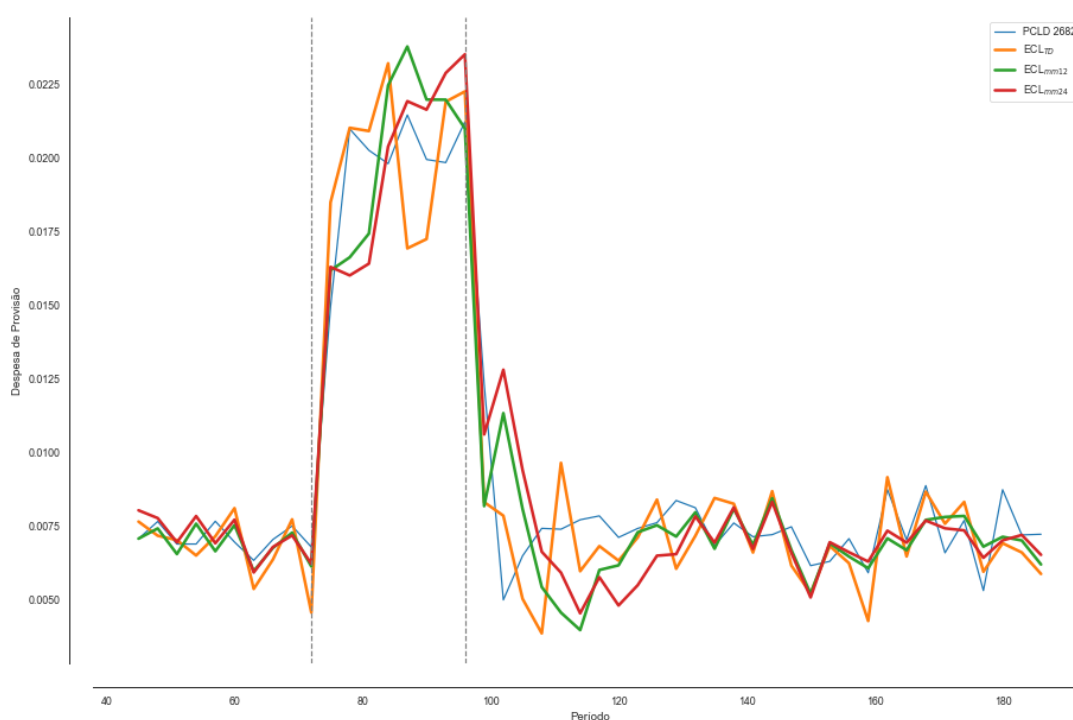


Figura 6.9: Impacto do Choque nas Despesas de Provisão com Atualização Dinâmica (Fonte: Elaborado Pelo Autor).

Observa-se que com essa aplicação, a perda esperada para risco de crédito é atualizada de modo a antecipar as perdas dos períodos, estabelecendo despesas de PCLD em níveis semelhantes aos obtidos pela metodologia da Resolução CMN 2.682 no cenário simulado.

A aplicação da média móvel é uma forma de minimização da variabilidade da ECL, estabilizando as despesas de PCLD nas demonstrações financeiras da IF. O coeficiente de variação despesas de PCLD das ECL com a taxa mais recente, média móvel de 12 meses e média móvel de 24 meses, respectivamente, foram de 0,43%, 0,37% e 0,39%. Ou seja,

apesar da inclusão da média móvel tender a estabilizar as despesas de provisão, períodos longos podem não representar uma situação atualizada do risco de crédito.

## 6.5 Conclusão

O uso de simulações para avaliação do cenário inerentes à atualização do cálculo da PCLD em IFs é uma boa prática para avaliação dos riscos associados a este contexto. Além disso, a situação atual, em que as principais IFs realizam reportes periódicos das demonstrações no padrão internacional é uma fase importante para avaliação destes riscos e da robustez da metodologia de cálculo.

O sistema de simulação desenvolvido neste trabalho mostrou consistência com as variáveis do SFN, está amparada na compreensão da dinâmica do crédito mapeada na literatura e têm detalhamento suficiente para cálculo dos parâmetros de risco e cálculo da provisão para risco de crédito nas diferentes metodologias.

As análises realizadas no cenário simulado, evidencia a relevância das validações dos parâmetros de risco utilizados na mensuração da ECL, tendo em vista os impactos negativos nas demonstrações financeiras possíveis em caso de não atualização destes parâmetros. Por este motivo, a supervisão dos órgãos competentes é importante no período de transição para o novo modelo e no decorrer do seu uso na contabilidade bancária, previsto para 2025 no Brasil.

Foi avaliado, neste cenário, o efeito de mitigação de grandes variações na despesa de PCLD pela exigência de reavaliação dos grupos homogêneos de risco pelo atraso. Em estudos futuros, é possível utilizar o mesmo sistema de simulações para avaliar impactos de outros fatores envolvidos na estimação da ECL, como: efeito do ajuste *lifetime*, efeitos de expectativas macroeconômicas equivocadas, entre outros.

Este estudo tem limitações, em especial, quanto a variedade de cenários simulados. O resultado é pontual em um cenário avaliado, mas a avaliação torna-se mais robusta para comparação das abordagens e dos possíveis riscos, caso diversos cenários sejam considerados, o que será explorado em estudos futuros.

# Capítulo 7

## Risco de Modelo

### 7.1 Introdução

No cenário atual do setor bancário brasileiro, com o avanço das *Fintechs* e do *Open Banking* o uso de dados e soluções analíticas na tomada de decisão e no desenvolvimento de soluções adequadas às exigências dos clientes torna-se fundamental para os bancos. Com isso, é cada vez mais frequente o uso de modelos analíticos para diferentes fins nas IFs. Ter uma política de gerenciamento do risco de modelo robusta, que significa assegurar que os modelos utilizados na instituição estão atendendo aos seus objetivos planejados, é parte fundamental para o sucesso no uso de soluções analíticas nessas instituições.

Embora a utilização de modelos possa trazer vários ganhos em diferentes aspectos, o uso incorreto de técnicas de modelagem, dos dados disponíveis ou mesmo aplicação incorreta do próprio modelo, pode acarretar perdas para as instituições. Neste contexto, conceitua-se o risco de modelo como a possibilidade destas perdas devido à erros no desenvolvimento e aplicação dos modelos [108].

O FED [31] descreve as principais definições e boas práticas relacionadas ao gerenciamento do risco de modelo em instituições financeiras, no documento *Supervisory Guidance on Model Risk Management*, doravante denominado OCC-11. Define-se que o gerenciamento do risco de modelo deve incluir boas práticas nas fases de desenvolvimento, implementação, uso e validação do modelo.

Embora não seja comum a existência de regulamentações específicas sobre o risco de modelo, como no caso do OCC-11 do FED, as grandes instituições financeiras declaram o risco de modelo como um risco relevante e estabelecem governança para utilização dos modelos em seus processos internos.

Por exemplo, o *Targeted Review of Internal Models* (TRIM) [109], publicado pelo *European Central Bank*, foi desenvolvido com o objetivo de reduzir a variabilidade na mensuração dos ativos das instituições financeiras ocasionada pelo uso inadequado de

modelos internos de mensuração de risco e capital. Neste documento, há várias menções ao risco de modelo e formas de mitigá-lo e controlá-lo e fica esclarecido que, apesar do foco do documento ser modelos internos, há expectativa de que seja gerenciado o risco associado a todos os modelos.

No contexto brasileiro, a Resolução N<sup>o</sup> 4.557 do CMN, de 23 de fevereiro de 2017 [9], trata das estruturas necessárias no gerenciamento de riscos e capital nas instituições financeiras. Embora não haja, neste documento, menção específica do termo risco de modelo, estabelece-se que: “Quanto aos modelos, a resolução n<sup>o</sup> 4557, indica que, quando utilizados, devem ser submetidos a avaliações periódicas quanto:

“a) à adequação das metodologias utilizadas; e b) ao desempenho, incluindo a comparação entre perdas estimadas e observadas (*back-testing*).”

Entre os cinco maiores bancos brasileiros em carteira de crédito, todos declaram uma estrutura de gerenciamento de risco de modelo, ou reconheceram o risco de modelo como um risco relevante para a instituição [34][35][36][37] em seus Relatórios de Gerenciamento de Risco dos anos de 2018 e 2019. Conforme descrito na Resolução CMN N<sup>o</sup> 4.557 e no OCC-11, a governança para modelos em instituições financeiras prevê uma validação independente e o monitoramento contínuo dos resultados dos modelos. Neste contexto, evidencia-se a necessidade de priorização de modelos na realização dessas atividades, baseado em sua relevância e na mensuração do risco associado a cada modelo.

Neste contexto, documentos divulgados pelas consultorias Deloitte [110] e KPMG [50] sobre o gerenciamento do risco de modelo em instituições financeiras incluem a definição de medida para quantificação do risco de modelo como uma boa prática de governança de modelos.

Este tipo de situação pode ser entendido como um problema de classificação, que pode ser resolvido pela aplicação de técnicas de Multiple-criteria decision analysis (MCDA). Embora não seja difundido o uso desta técnica aplicada à mensuração do risco de modelo, podem ser identificados trabalhos semelhantes em que esta técnica é aplicada para mensuração do risco em projetos.

Este trabalho apresenta uma abordagem para aplicação do MCDA para quantificação dos riscos associados aos modelos a ser utilizado em instituições financeiras para priorização dos processos estabelecidos na governança dos riscos de modelo. Este capítulo sintetiza todos os resultados obtidos neste trabalho, relacionando os riscos associados aos modelos utilizados neste trabalho para o cálculo da ECL pelos requisitos do IFRS 9. A partir desta mensuração é possível, dentro de cada objetivo associado aos modelos, proporcionar uma decisão a partir de critérios definidos, sobre os melhores modelos a serem utilizados neste contexto.

Além desta introdução, mais quatro seções compõem este trabalho. Na segunda seção são apresentados os referenciais teóricos que embasaram os estudos, na terceira seção é descrita a metodologia utilizada, a quarta seção contém a aplicação de um estudo de caso obtidos e a quinta apresenta as conclusões do estudo.

## 7.2 Referencial Teórico

Conforme descrito pela *American Academy of Actuaries* [32], Risco de modelo pode ser definido como a possibilidade de perdas (econômicas, reputacionais, etc) decorrentes de falhas ou mal uso dos modelos. Um conceito relevante no gerenciamento do risco de modelo é a definição de modelo. Deloitte [110] destaca a definição de modelo como parte da governança de modelos nas instituições financeiras, de modo que o conceito de modelo pode variar entre as instituições.

O OCC-11 define modelo como: um método, sistema ou abordagem quantitativa que aplica teorias, técnicas e suposições estatísticas, econômicas, financeiras ou matemáticas para processar dados de entrada em estimativas quantitativas. Embora outras definições possam ser encontradas na literatura, este conceito será utilizado para este trabalho.

### 7.2.1 Gerenciamento do Risco de Modelo

#### Fontes de Risco de Modelo

Os motivos pelos quais ocorre o risco de modelo variam na literatura pesquisada. No entanto, embora sejam categorizadas e agrupadas em diferentes formas, as fontes reconhecidas são semelhantes. OCC-11 considera duas razões pelas quais o risco de modelo pode ocorrer, sendo elas:

- O modelo pode ter erros em conceitos, simplificações, dados utilizados, entre outros, desde o seu desenvolvimento até a implementação, gerando cálculos errados em relação ao objetivo para o qual foi proposto; e
- O modelo pode ser utilizado incorretamente. Mesmo que o modelo tenha sido conduzido corretamente desde o desenvolvimento até a implementação, ele pode produzir estimativas imprecisas por alteração do público ao qual é aplicado ou alteração do ambiente em que opera.

Segundo Wu [111], as fontes do risco de modelo advêm da implementação e do uso incorretos de modelos de boa performance ou da implementação e uso corretos de modelos descalibrados. Em outros documentos descritos por consultorias relevantes no que se refere

ao gerenciamento do risco de modelo, os agrupamentos dos elementos em que se observa o risco de modelo variam. A Tabela 7.1 mostra um resumo dos diferentes agrupamentos:

Tabela 7.1: Fontes de Risco de Modelo

Documento	Desenvolvimento	Dados	Implementação	Uso	Documentação
Deloitte[110]	x		x	x	
KPMG[50]	x		x	x	
MSO [112]	x		x	x	x
EXL [113]	x	x	x	x	

Embora estes agrupamentos variem entre as referências utilizadas, há estabelecimentos de governança em todas elas. O OCC-11, por exemplo, trata da importância da documentação e sobre os dados dentro do tema desenvolvimento e implementação. Neste trabalho, os critérios considerados no gerenciamento do risco de modelos serão: desenvolvimento, documentação, dados, implementação e uso.

## Gestão do Risco de Modelos

Há conceitos definidos dentro da estrutura de gerenciamento do risco de modelo que são especificados na literatura. Os principais deles estão elencados abaixo:

**Ciclo de Vida:** Se refere às etapas de um modelo. Em cada etapa do ciclo de vida do modelo, há diferentes agentes envolvidos e diferentes responsabilidades, tendo como objetivo o gerenciamento adequado do risco de modelo. Hill [114] especifica as seguintes etapas do ciclo de vida de um modelo:

- Identificação e Iniciação;
- Desenvolvimento e Teste;
- Validação independente e Aprovação;
- Implementação e monitoramento;
- Manutenção e mudanças; e
- Aposentadoria do modelo.

**Inventário Corporativo de Modelos:** Documentação corporativa de todos os aspectos relevantes dos modelos em uso, em desenvolvimento ou em implementação na instituição. Um inventário de modelos deve descrever o motivo e o público para o qual o modelo foi desenvolvido, datas em que foi validado ou implementado (OCC-11). Conforme definido nos seus relatórios de gerenciamento de riscos do segundo trimestre de 2019, os cinco maiores bancos brasileiros declaram possuir um inventário corporativo de modelos.

Validação Independente: De acordo com o OCC-11, validação é o conjunto de procedimentos realizados para avaliar se o modelo está atingindo os resultados esperados. Conforme descrito no Art. 9º da Resolução CMN Nº 4.557 [9], quando os modelos utilizados nos bancos forem relevantes, estes devem ser avaliados quanto à robustez das metodologias utilizadas, bem como o seu desempenho por unidade que não seja responsável pelo seu desenvolvimento, nem pelo uso do modelo.

## Mensuração do Risco de Modelo

Conforme descrito pela KPMG [50] uma estrutura de gerenciamento de risco de modelo avançada contém, além dos aspectos tratados nesta seção, a existência de uma unidade de quantificação do risco de modelo. No contexto da mensuração do risco de modelo, segundo o *American Academy of Actuaries* [32], a forma de mensuração do risco de modelo adotada por instituição pode ser qualitativa ou quantitativa, sendo descritas as seguintes abordagens:

*Scoring Model Risk*: Abordagem qualitativa e subjetiva para mensuração do risco de modelo. Uma forma de mensurar o risco de modelo é a construção de um escore a partir dos atributos de risco listados na Tabela 7.1;

Quantificação do Risco de Modelo: Geralmente, a mensuração quantitativa é feita como uma fração do risco operacional, desenvolvendo estimativas da frequência e severidade dos eventos de risco de modelo.

Alguns trabalhos foram publicados para uma mensuração quantitativa do risco de modelo, estimando as perdas ocorridas devido ao uso de modelos em modelos de risco. Por exemplo, no trabalho de Danielsson [115], foi analisado o risco de modelo em modelos de risco de mercado. Em um exemplo associado ao contexto da implementação do cálculo da ECL na implementação do IFRS 9, o risco de modelo pode se materializar de várias formas. Sobre esse contexto, Seitshiro [43] aplicou uma abordagem quantitativa para mensurar o risco de modelo associado a erros de especificação em modelos de PD.

A quantificação mais detalhada, geralmente está alinhada a um modelo de uso específico, em que é possível quantificar as perdas possíveis pelo seu uso. No entanto, em um contexto mais abrangente sobre os modelos utilizados, esta análise de forma comparável entre todos os modelos inventariados na IF nem sempre é possível. Por isso, o uso de técnicas para avaliação da por meio de critérios, inclusive subjetivos, pode ser uma boa alternativa para atribuição de *Scoring Model Risk* no gerenciamento do risco de modelo.

Uma metodologia amplamente utilizada para processos de tomada de decisão subjetivos, para este tipo de problema é o MCDA. Mesmo não sendo encontrados trabalhos

sobre a aplicação do MCDA para mensuração do risco de modelo, alguns trabalhos com objetivos semelhantes estão disponíveis na literatura.

### 7.2.2 Análise de Decisão por Multicritérios (MCDA)

Belton [116] explicita que as diferentes escolhas ou cursos de ações a serem feitas, se torna um problema de tomadas de decisão com múltiplos critérios, quando existem vários desses padrões que entram em conflito substancialmente, assim levando em consideração que as decisões em nível corporativo, tanto nos setores públicos ou privado, geralmente envolvem considerar uma gama enorme de critérios, especialmente quando é preciso buscar o consenso entre grupos de interesses diferentes. Para esses autores, a subjetividade é inerente a toda tomada de decisão em particular na escolha de critérios para se referir nas decisões e no peso relativo dado a esses critérios, para tanto, o MCDA não dissipa essa subjetividade. A escolha do método para tomadas de decisão em múltiplos critérios é assertiva quando se concentra no contexto subjetivo de critérios para se embasar na tomada de decisão.

Brans [117], propuseram o método *Preference Ranking Organization Method for Enrichment* (PROMETHEE)I que fornece uma classificação parcial de ações e posteriormente desenvolveram a segunda versão desse método, o PROMETHEE II, que efetua uma classificação completa das ações. O método PROMETHEE proposto pelos referidos autores, se baseia no método de superação, onde a relação de dominância relacionada a um problema de multicritério é baseado na unanimidade dos pontos de vista. Esse método possui duas fases, que são: a construção de uma relação de superação e a exploração dessas relações para auxiliar o tomador de decisões. O PROMETHEE segue os seguintes princípios:

- Extensão da relação de critério;
- Relação valorizada de superação; e
- Exploração da relação de superação.

O PROMETHEE II fornece uma classificação completa que é agradável ao tomador de decisão, porém, algumas informações úteis sobre incompatibilidades são perdidas. Por fim os autores concluíram que este método fornece uma abordagem fácil, para resolver problemas de multicritério, considerando simultaneamente critérios estendidos e relações de superação.

Gul [118], em estudo sobre problemas de seleção de materiais, propuseram uma metodologia genérica de *fuzzy* MCDA, onde utiliza o método PROMETHEE com lógica *fuzzy* para resolução de problemas de seleção de materiais para um painel de instrumentos automotivos. O trabalho de Souza [119] aplicou o método PROMETHEE II para soluções



para perdas reais em sistemas de abastecimento de água, destacando a sua capacidade de hierarquizar as alternativas mais recomendadas.

Neste trabalho, as alternativas a serem hierarquizadas são os modelos utilizados nas instituições, baseando-se nos critérios destacados em consultorias relevantes. A priorização por meio do Risco de Modelo é relevante para definir alocação dos recursos da instituição na manutenção ou acompanhamento dos modelos envolvidos. A seção seguinte descreve os passos metodológicos para realização deste trabalho.

## 7.3 Metodologia

Este trabalho utiliza abordagem qualitativa e quantitativa para solução de problemas reais, caracterizando-se como uma pesquisa exploratória de natureza aplicada com uso de MCDA para quantificação do risco de modelo.

As etapas necessárias em aplicações do MCDA, as quais foram seguidas neste trabalho, envolvem as seguintes etapas, conforme descrito por Belton [116]: identificar e estruturar problema; construir modelo matemático; aplicar o modelo em um exemplo; e tomar a decisão.

### 7.3.1 Método PROMETHEE

Para Afful-Dadzie [120], o PROMETHEE é um método interessante na classificação e seleção de alternativas, levando em consideração sua robustez na comparação do desempenho das alternativas, considerando no *ranking* proposto. Para tanto existe uma versão desse método que utiliza números *fuzzy* para lidar com problemas incertos e subjetivos. O PROMETHEE proposto por Brans [117], foi ajustado para efetuar a comparação pareada de opções classificando-as na lógica *fuzzy*.

Digalwar [121] cita a popularidade no uso do algoritmo PROMETHEE, se deve às referidas vantagens:

- É um método de superação amigável;
- Foi aplicado com sucesso a problemas da vida real;
- Facilidade de aplicação, eficiência e interatividade, ele tem uma influência transparente de cada critério e peso na solução;
- Baseia-se na importância de uma diferença de desempenho entre duas soluções, que descreve melhor se uma solução deve ser preferida a outra;
- É um método de classificação bastante simples na concepção e aplicação em comparação com outros métodos para MCDA;

- Está bem adaptado a problemas em que um número finito de ações alternativas podem ser classificados; e
- Dados qualitativos e variáveis linguísticas podem ser facilmente tratados

Segundo Geldermann [122], todos os parâmetros usados têm uma significação real, para que o tomador de decisão possa interpretá-los imediatamente. Considerando a aplicação do PROMETHEE, segundo Brans [117], pode ser descrito conforme os passos abaixo:

**Primeiro** : Especificar para cada critério uma função de preferência generalizada.

**Segundo** : Definir um vetor contendo os pesos, que são uma medida para a importância relativa de cada critério. Se todos os critérios tiverem a mesma importância na opinião do tomador de decisão, todos os pesos poderão ser considerados iguais.

**Terceiro** : Definir todas as alternativas para cada relação de superação. Sendo assim o índice de preferência é uma medida da intensidade de preferência do tomador de decisão por uma alternativa em comparação com uma alternativa  $a_0$  para a consideração simultânea de todos os critérios. É basicamente uma média ponderada das funções de preferência.

**Quarto** : Como medida para a força das alternativas, a saída do fluxo é calculado, sendo o fluxo de saída compreendendo a soma dos valores dos arcos que deixam o nó, e portanto, produz uma medida do 'caráter de superação'.

**Quinto** : Como medida para a fragilidade das alternativas, a entrada o fluxo é calculado, medindo o "caractere ultrapassado"(analogicamente à saída do fluxo).

**Sexto** : Uma avaliação gráfica da relação de superação é derivada. Basicamente, quanto maior o fluxo de saída e menor o fluxo de entrada, melhor a ação. Este resultado é representado graficamente por uma pré-encomenda parcial (PROMETHEE I) ou por um pedido completo (PROMETHEE II).

Estruturando o problema desta forma, é possível aplicar método MCDA para mensurar o risco de modelo em um inventário corporativo de modelos, como pode ser observado no estudo de caso realizado, tema do qual trata a próxima seção.

Neste estudo de caso, os modelos utilizados foram alguns dos modelos utilizados para implementação do para estimativa da PD *lifetime* e para projeção das perdas esperadas a partir de informações macroeconômicas nos NPL, obtidos a partir dos trabalhos dos capítulos anteriores.

## 7.4 Estudo de Caso

Para elaboração de estudo de caso da abordagem proposta para mensuração do risco de modelo, foram utilizados oito modelos associados à gestão do risco de crédito em IFs. Neste caso os modelos tem a necessidade associada a requisitos da implementação do padrão contábil IFRS 9, tema discutido em mais detalhes nos capítulos anteriores deste trabalho.

### 7.4.1 Estruturação do problema

Este trabalho aplica a metodologia de apoio à decisão por multicritérios para uma mensuração qualitativa do risco de modelo. Os critérios considerados para a mensuração do risco de modelo utilizados neste trabalho foram os fatores de risco mais citados na literatura, sendo eles: desenvolvimento, implementação e uso. A partir destas três classes de critérios, uma instituição financeira pode classificar os modelos constantes do seu inventário corporativo em relação ao seu risco de modelo.

Para cada um dos fatores, desenvolvimento, implementação e uso, foi verificado as recomendações e as boas práticas nas referências sobre governança de risco de modelo pesquisadas, sendo elas: Deloitte [110], OCC-11 [31], KPMG [50], MSO [112]. Ao todo foram destacados 10 critérios para mensuração do risco de modelo, sendo cinco sobre a desenvolvimento, dois sobre a implementação e três sobre o uso do modelo.

No Apêndice A são apresentados os critérios e alternativas obtidos, baseado nos documentos que tratam sobre a governança de modelos em instituições financeiras, e suas respectivas escalas.

Além dos critérios sobre o risco de modelo, são apresentados, no Apêndice A, critérios inerentes ao impacto ou materialidade do risco de modelo associado. Como a abordagem foi aplicada a modelos de uso similar, a distinção entre as materialidades torna inexistente a distinção entre a maior parte das alternativas. Por este motivo, este trabalho limitou a aplicação numérica ao risco de modelo, sem aplicação na avaliação do impacto.

### 7.4.2 Aplicação Numérica

Segundo Digalwar [121] a maior parte das aplicações dos números *fuzzy* no algoritmo do PROMETHEE usam o *fuzzy Analytic Hierarchy Process* (AHP) para atribuir pesos aos critérios ou para os indicadores de performance. Este mesmo autor aplicou os números *fuzzy* para lidar com critérios em escalas qualitativas. Neste mesmo contexto, este trabalho utilizou números *fuzzy* para lidar com as escalas qualitativas dos critérios utilizados.

Geldermann [122] afirma que dos diversos modelos existentes no sentido de utilizar números *fuzzy* agregados nenhum pode ser considerado o melhor em qualquer circunstância. Neste trabalho, para atribuir escalas mais precisas nas escalas qualitativas, em cada um dos critérios não binários, foi aplicado método *Fuzzy* com *defuzzification* pela abordagem de *Centre of Area* (COA), conforme formulação demonstrada por este mesmo autor. A Figura 7.1 exemplifica os números *fuzzy* para a escala do critério 1 da classe de desenvolvimento.

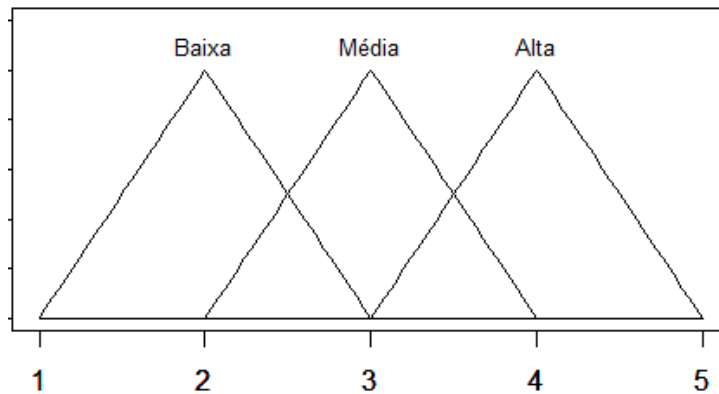


Figura 7.1: Números *Fuzzy* Triangular para o Primeiro Critério (Fonte: Elaborado Pelo Autor).

A partir dos números *fuzzy* após a *fuzzy* resultantes foram definidos os pesos para as escalas qualitativas. Como resultado do primeiro critério, aplicou-se os pesos Baixa 0,087; Média 0,1604; e Alta 0,7675. Este procedimento foi realizado para todos os critérios com escala qualitativa não binária.

A aplicação do método *Fuzzy* AHP e a definição dos pesos entre os critérios foram definidos pela consulta a tomadores de decisão na área de risco de modelo em instituição financeira brasileira de grande porte. Para auxiliar nos cálculos das matrizes de superação e dos fluxos de superação foi utilizado o software *Visual PROMETHEE Academic Edition*.

### 7.4.3 Análise de Sensibilidade

Uma medida de qualidade da abordagem por meio do método PROMETHEE é a relação dos autovetores entre as diferenças relativas calculadas pelos critérios. Considera-se [123] que valores acima de 70% indicam boa distinção entre as alternativas avaliadas. A Figura 7.2 mostra o resultado dessa análise para o modelo aplicado.

Nota-se que o modelo aplicado, com os critérios utilizados apresentam bons resultados para a comparação do risco de modelo entre as alternativas avaliadas.

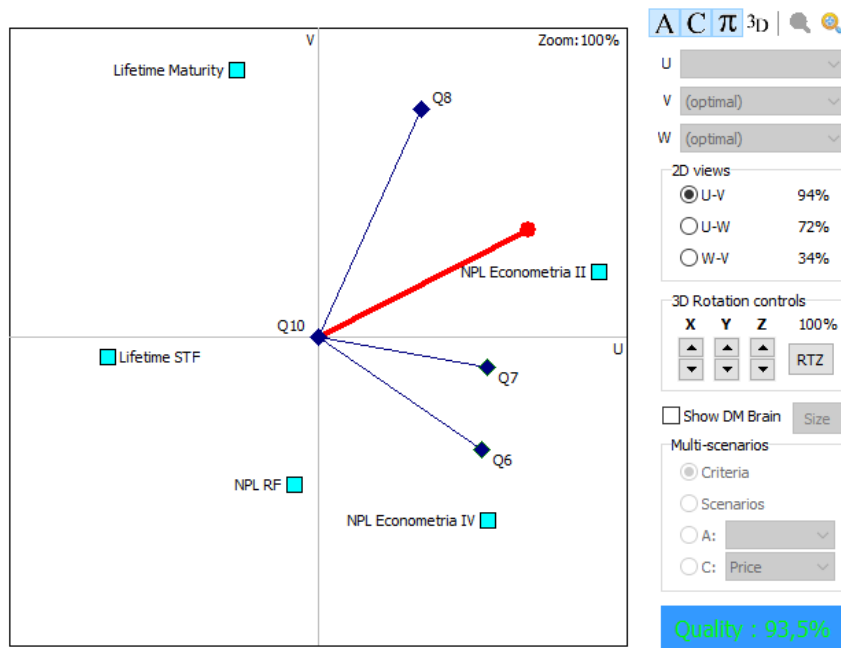


Figura 7.2: Medida de Consistência (Fonte: Visual PROMETHEE [123]).

A Tabela 7.2 mostra os resultados obtidos do: fluxo positivo de superação, que compara a força de um modelo som todos os outros quanto a todos os critérios; fluxo negativo de superação que compara as fraquezas dos modelos neste mesmo contexto; e o fluxo líquido que apresenta o ranqueamento dos modelos. Neste caso, os modelos estão classificados do maior para o menor risco, de modo que um maior fluxo líquido representa um maior risco de modelo.

Tabela 7.2: Classificação dos Modelos Desenvolvidos

Rank	Modelo	$\phi$	$\phi_+$	$\phi_-$
1	Lifetime Maturity	0,1186	0,3184	0,1998
2	Lifetime ST	0,0447	0,2524	0,2077
3	NPL Econometria 1	0,0259	0,2859	0,26
4	NPL Econometria 2	0,0259	0,2859	0,26
5	NPL RF	0,0105	0,207	0,1965
6	Lifetime Cox	-0,0213	0,1879	0,2092
7	Lifetime RSF	-0,0415	0,1931	0,2346
8	NPL Econometria 4	-0,1628	0,1639	0,3267

Como pode ser observado, considerando os critérios e pesos utilizados, o modelo de ajuste da PD para o prazo, a partir de uma equação fixa, utilizada nas simulações apresentadas no Capítulo 6 teve seu risco maior que todos os outros modelos. Além disso,

comparando os modelos para ajuste *lifetime*, dentre as alternativas testadas, os melhores indicadores de desempenho do Modelo RSF o destacaram como melhor opção.

Sobre os modelos de aplicação de efeitos econômicos sobre o NPL, o modelo econométrico IV foi considerado o de menor risco entre todos os modelos analisados. A Figura 7.3 traz uma visualização em que é possível avaliar os critérios pelos quais cada modelo recebeu cada *rank*.

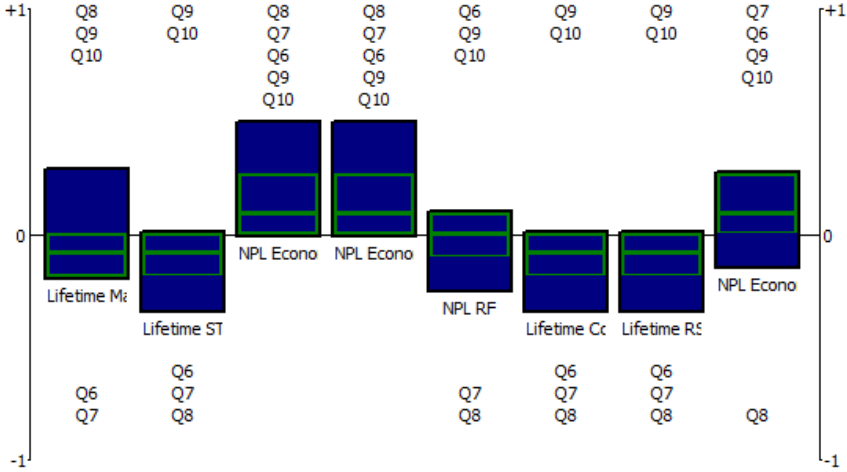


Figura 7.3: Riscos de Modelo Atribuídos (Fonte: Visual PROMETHEE [123]).

Basicamente, interpreta-se que o risco elevado sobre o uso da equação fixa de ajuste *lifetime* se deve principalmente à expectativa de descalibragem (Q8), que em comparação às alternativas tem valor mais elevado, que não compensam os aspectos positivos quanto à implementação (Q6 e Q7).

Vale destacar que o Modelo de previsão da NPL pelo modelo *random forest* apresentou vantagens de implementação em relação aos demais modelos econométricos. Isso se deve à maior facilidade de integração do modelo desenvolvido em Python [93] em comparação ao Eviews [94], na avaliação realizada.

## 7.5 Conclusões

O risco de modelo tem sido destacado pelas principais instituições financeiras brasileiras nos últimos anos, de modo que as discussões quanto às melhores formas de gerenciamento deste risco devem ser uma discussão constante tanto corporativamente, quanto na academia.

Tendo em vista os mais variados tipos de modelos que uma instituição financeira pode usar em sua gestão, um método para mensurar o risco de modelo pode ser complicado e

necessitar de informações qualitativas ou subjetivas para ser realizado. Neste contexto, o MCDA torna-se uma boa alternativa, considerando sua adequação em estudos desse tipo.

O PROMETHEE tem sido empregado como uma boa alternativa para problemas de mensuração que envolvam algum nível de subjetividade. Há trabalhos sobre gerenciamento de riscos, priorização de projetos entre outros temas. Dessa forma, apresenta-se como uma boa solução a ser utilizada nas instituições financeiras para mensuração do risco de modelo.

Os resultados obtidos no estudo de caso mostram que se pode obter bons resultados com a aplicação desta técnica, podendo ser utilizada tanto na seleção de modelos a serem implementados na instituição, ou para priorização de atividades de revisão ou retirada de uso de modelos com risco acima do padrão identificado.

Dado o entendimento que o risco de modelo é reconhecido e declarado como um risco relevante pelas instituições financeiras brasileiras e no mundo, e o entendimento de que uma medida de mensuração de risco de modelo é um artefato alinhado às melhores práticas de gerenciamento de riscos, entende-se que a aplicação proposta pode trazer ganhos para a melhoria da gestão de riscos neste segmento.

A aplicação em uma relação maior de modelos, para diferentes fins pode ser explorada em estudos futuros para avaliar a manutenção da consistência dos ordenamentos criados, incluindo além do risco de modelo o impacto percebido pela materialização do evento de risco de modelo em cada alternativa.

No contexto da aplicação deste modelo para os modelos do IFRS 9, conclui-se por este trabalho que, para a estimação da PD *lifetime*, discutida no Capítulo 4, o modelo RSF é a opção de menor risco de modelo, considerando seu melhor desempenho em relação aos demais. Já sobre as projeções *forward-looking* da ECL, do Capítulo 5, o modelo OLS de menores taxas de erro e utilizando variáveis projetadas como insumo apresentou menor risco, o uso de um modelo bem ajustado para prever as variações de maneira prospectiva seria uma boa alternativa para evitar problemas de variações acentuadas na despesa de provisão, apresentada no Capítulo 6.

# Capítulo 8

## Considerações Finais

Este trabalho teve o objetivo de proporcionar uma visão geral dos modelos necessários para implementação do Cálculo da Perda Esperada de Crédito (ECL) para as demonstrações financeiras em IFRS 9 no Brasil. Foram apresentadas aplicações independentes e discussões sobre os principais modelos utilizados para este fim, que são utilizados conjuntamente pelas instituições para obter estimativas mais precisas.

Considerando os objetivos específicos listados, foram identificados os principais modelos relacionados às demonstrações em IFRS 9, a partir da análise dos documentos sobre o padrão internacional e da sua implementação no mercado brasileiro. Basicamente, foram identificados, além dos modelos utilizados para gestão do capital PD, LGD e da exposição, os ajustes necessários devido a condições econômicas e para atualização da PD *lifetime*. Para cada um destes modelos, foi aplicado um estudo de caso e a discussão sobre o seu uso na gestão da ECL.

Sobre o ajuste *lifetime*, foi apresentada uma aplicação sobre uso de modelos de Análise de Sobrevivência com técnicas de *machine learning* baseadas em árvore. Tendo em vista a necessidade de variar o prazo da estimativa de uma probabilidade, a Análise de Sobrevivência mostrou-se uma boa alternativa sem perda de acurácia em relação aos métodos de classificação com a variação do período. Além disso, nota-se que o algoritmo mais robusto RSF mostrou melhores indicadores de ajuste do modelo.

A respeito do impacto na inadimplência por motivo das condições econômicas, foram aplicados modelos econométricos para projeção dos NPL e, a partir das projeções, foi avaliado o efeito sobre os erros do modelo, causados pelo uso de covariáveis projetadas. Nos modelos mais bem ajustados, verificou-se impacto em níveis estatisticamente significativos. No entanto, não foi possível identificar modelos concorrentes com desempenho preditivo melhor que os modelos com essa característica no período estudado.

Para os cálculos dos parâmetros de risco utilizados na estimativa da ECL, foi aplicado estudo comparativo, em base simulada, sobre a sensibilidade das formas de cálculo da



provisão vigente e o IFRS 9 com avaliação dos seus impactos sobre as demonstrações financeiras da IF. Com este trabalho foi possível exemplificar a importância da validação e monitoramento contínuo destes parâmetros, tendo em vista os impactos negativos que podem ser gerados pela não atualização em choques de inadimplência que não sejam reconhecidos pelos modelos de projeção.

No gerenciamento do risco de crédito, essas abordagens devem ser conduzidas conjuntamente. A carteira de ativos de crédito deve ser classificada em três estágios de agravamento de risco, os modelos para a estimativa *lifetime* devem ser utilizados para ajustar as operações com agravamento significativo do risco do ativo apresentar problemas de recuperação. De modo preventivo, os modelos a associação da situação econômica esperada com o risco de crédito é uma forma de evitar variações não previstas na carteira de crédito, o que pode levar a despesas de provisões elevadas, comprometendo os resultados e a liquidez das instituições.

Dado a quantidade elevada de modelos a serem geridos para uma boa gestão de riscos, foi proposta uma abordagem para mensuração do risco de modelo, para priorização dos modelos nos processos internos da instituição, que pode auxiliar na decisão de modelos de menores risco a partir de critérios definidos, bem como a necessidade de revisão dos modelos em uso. Para isso, foi utilizado o MCDA, com o método PROMETHEE utilizando números *fuzzy* para as avaliações qualitativas dos critérios levantados sobre desenvolvimento, implementação e uso dos modelos.

A partir da aplicação desta abordagem para mensuração do risco de modelo nos modelos utilizados no restante do trabalho, verificou-se bons indicadores de qualidade de ajustamento. Além disso, entende-se que houve coerência no ordenamento criado e nos efeitos de cada critério sobre a classificação de risco.

## 8.1 Limitações

A implementação do IFRS 9 no Brasil tem formalização recente e alguns pontos elencados podem ser melhores explorados para uma visão mais detalhada do contexto geral. Por exemplo, a questão do valor justo do ativo pode ser um objeto de definição por meio de modelos e estudos embasados. Apesar disso, este trabalho considerou em todos os capítulos o critério mínimo de valor justo, que é dado pelo saldo atualizado dos ativos. Do mesmo modo, o conceito de prazo esperado do ativo pode ser mais bem abordado, neste caso considerou-se apenas o prazo estipulado dos ativos financeiros.

As comparações dos modelos, tanto *lifetime* quanto os de estimação da NPL têm interpretação exploratória e não necessariamente generalizáveis, de modo que as conclusões ficam restritas ao público e período observados. Além disso, nos dois casos o uso de

modelos de *machine learning* utilizou uma quantidade limitada de técnicas, podendo ser melhor explorado em estudos futuros.

A aplicação da abordagem para quantificação do risco de modelos foi elaborada de modo a abranger todos os modelos utilizados na instituição. A aplicação somente nos modelos associados a um uso específico pode comprometer a avaliação da abordagem com a abrangência esperada.

## 8.2 Contribuições

Os objetivos inicialmente propostos foram atendidos a partir dos métodos aplicados, a pesquisa apresentada traz uma visão abrangente e contemporânea dos desafios que serão enfrentados nos próximos anos na migração do padrão vigente para o IFRS 9. Apesar do relacionamento entre os assuntos discutidos, cada capítulo apresentado traz uma aplicação independente em seus objetivos específicos.

# Referências

- [1] Bank for International Settlements: *History - foundation and crisis*. [https://www.bis.org/about/history\\_1foundation.htm](https://www.bis.org/about/history_1foundation.htm), acesso em 25/05/2021. 1
- [2] Basel Committee On Banking Supervision: *Basel I: International Convergence of Capital Measurement and Capital Standards*, 1988. 1
- [3] Basel Committee on Banking Supervision: *International Convergence of Capital Measurement and Capital Standards: A Revised Framework Comprehensive Version*. Número June. 2006, ISBN 9291316695. <http://www.bis.org/publ/bcbs128.pdf>. 1, 6, 63
- [4] Basel Committee on Banking Supervision: *Studies on the Validation of Internal Rating Systems*. Working Papers of the Basel Committee on Banking Supervision, (14):120, 2005. [http://www.bis.org/publ/bcbs\\_wp14.pdf](http://www.bis.org/publ/bcbs_wp14.pdf). 1, 15
- [5] Demyanyk, Yuliya e Otto Van Hemert: *Understanding the subprime mortgage crisis*. Review of Financial Studies, 24(6):1848–1880, 2011, ISSN 08939454. 1
- [6] Bernard, Carole e Steven Vanduffel: *A new approach to assessing model risk in high dimensions*. Journal of Banking and Finance, 58:166–178, 2015, ISSN 03784266. <http://dx.doi.org/10.1016/j.jbankfin.2015.03.007>. 1
- [7] Cohort, Pierre, Pierre Emmanuel Vehel e Frédéric Patras: *Toward model value-at-risk: Bespoke cdo tranches, a case study*. Journal of Risk Model Validation, 7:21–34, setembro 2013. 1, 4
- [8] Basel Committee On Banking Supervision: *Basel III: A global regulatory framework for more resilient banks and banking systems December 2010 (rev June 2011)*. 2010, ISBN 9291318590. <http://www.bis.org/publ/bcbs189.pdf>. 2, 4, 8
- [9] Conselho Monetário Nacional: *RESOLUÇÃO N<sup>o</sup> 4.557*. 2017. [https://www.bcb.gov.br/pre/normativos/busca/downloadNormativo.asp?arquivo=/Lists/Normativos/Attachments/50344/Res\\_4557\\_v1\\_0.pdf](https://www.bcb.gov.br/pre/normativos/busca/downloadNormativo.asp?arquivo=/Lists/Normativos/Attachments/50344/Res_4557_v1_0.pdf). 2, 9, 70, 73
- [10] *RESOLUÇÃO N<sup>o</sup> 2.682*. (1):pp. 36–54, 1999. 2, 3, 8, 11, 15, 54, 56, 57, 60
- [11] Caneca, Roberta Lira, Paulo Roberto e Barbosa Lustosa: *Ciclos Econômicos e Provisão Para Créditos de Liquidação Duvidosa nos Bancos Brasileiros*. Em XVII USP International Conference in Accounting, 2017. 2, 56, 65

- [12] International Accounting Standards Board: *Financial instruments*. (April 2001):241–259, 2014. 2, 12, 43, 55, 57
- [13] International Accounting Standards Board: *International Accounting Standard 39 Financial Instruments: Recognition and Measurement*, 2011. 2
- [14] Galdi, Fernando, André De Moura e Robson França: *Loan loss provision models in Brazilian banks*. Meditari Accountancy Research, 2021, ISSN 20493738. 2, 3, 46
- [15] Caruso, Ezio, Katia D’Hulster, Tatsiana Kliatskova e Juan Ortiz: *Accounting Provisioning Under the Expected Credit Loss Framework*. Relatório Técnico, World Bank Group, 2021. 3
- [16] Conselho Monetário Nacional: *RESOLUÇÃO Nº 4818, DE 29 DE MAIO DE 2020*. 2020. <https://www.in.gov.br/en/web/dou/-/resolucao-n-4.818-de-29-de-maio-de-2020-259635896>. 3, 8
- [17] Banco Central do Brasil: *EDITAL DE CONSULTA PÚBLICA 54/2017*, 2017. <https://www.bcb.gov.br/htms/EditaIConsultaPublica54.pdf>, acesso em 26/05/2021. 3
- [18] Conselho Monetário Nacional: *RESOLUÇÃO Nº 4966, DE 25 DE NOVEMBRO DE 2021*. 2021. <https://www.bcb.gov.br/estabilidadefinanceira/exibenormativo?tipo=Resolu%C3%A7%C3%A3o%20CMN&numero=4966>. 3, 8, 11, 12, 15, 18, 19, 25, 35, 36, 38, 55, 56
- [19] Müller, Viktoria: *Hedge Accounting and its Consequences on Portfolio Earnings—A Simulation Study*. Accounting in Europe, 17(2):204–237, 2020, ISSN 17449499. <https://doi.org/10.1080/17449480.2020.1775267>. 6
- [20] Groff, Maja Zaman e Barbara Mörec: *IFRS 9 transition effect on equity in a post bank recovery environment: the case of Slovenia*. Economic Research-Ekonomska Istrazivanja, 34(1):670–686, 2021, ISSN 1331677X. <https://doi.org/10.1080/1331677X.2020.1804425>. 6
- [21] Krüger, Steffen, Daniel Rösch e Harald Scheule: *The impact of loan loss provisioning on bank capital requirements*. Journal of Financial Stability, 36:114–129, 2018, ISSN 15723089. <https://doi.org/10.1016/j.jfs.2018.02.009>. 6
- [22] Apergis, Emmanuel, Iraklis Apergis e Nicholas Apergis: *A new macro stress testing approach for financial realignment in the Eurozone*. Journal of International Financial Markets, Institutions and Money, 61:52–80, 2019, ISSN 10424431. <https://doi.org/10.1016/j.intfin.2019.02.002>. 7
- [23] Mechelli, Alessandro e Riccardo Cimini: *The effect of corporate governance and investor protection environments on the value relevance of new accounting standards: the case of IFRS 9 and IAS 39*. Journal of Management and Governance, (0123456789), 2020, ISSN 1572963X. <https://doi.org/10.1007/s10997-020-09551-9>. 7

- [24] Georgiou, K., G. N. Domazakis, D. Pappas e A. N. Yannacopoulos: *Markov chain lumpability and applications to credit risk modelling in compliance with the International Financial Reporting Standard 9 framework*. European Journal of Operational Research, 292(3):1146–1164, 2021, ISSN 03772217. <https://doi.org/10.1016/j.ejor.2020.11.014>. 7
- [25] Barnoussi, Aziz el, Bryan Howieson e Ferdy van Beest: *Prudential Application of IFRS 9: (Un)Fair Reporting in COVID-19 Crisis for Banks Worldwide?! Australian Accounting Review*, 30(3):178–192, 2020, ISSN 18352561. 7
- [26] Neisen, Martin e Hermann Schulte-Mattler: *The effectiveness of IFRS 9 transitional provisions in limiting the potential impact of COVID-19 on banks*. Journal of Banking Regulation, (0123456789), 2021, ISSN 17502071. <https://doi.org/10.1057/s41261-021-00151-7>. 7
- [27] Banks, Comptroller of the Currency Administrator of National: *OCC 2000-16 OCC Bulletin*. páginas 1–9, 2000. 8
- [28] LHABITANT, FRANÇOIS SERGE: *Coping with model risk*. Em Lore, Marc e Lev Borodovsky (editores): *The Professional's Handbook of Financial Management*, páginas 415–440. Butterworth-Heinemann, 2000. 8
- [29] Brotcke, L.: *Modifying model risk management practice in the era of ai/ml*. Journal of Risk Management in Financial Institutions, 13(3):255–265, 2020. [www.scopus.com](http://www.scopus.com). 9
- [30] Ghavami, Peter: *Chapter 11 Why Data and Model Risk Management?*, páginas 143–150. De Gruyter, 2020. <https://doi.org/10.1515/9783110664065-011>. 9
- [31] FED: *Supervisory Guidance on Model Risk Management*. 2011:1–21, 2011. 9, 69, 77
- [32] American Academy of Actuaries: *Model Risk Management*. 2019. 9, 18, 71, 73
- [33] Tunaru, Radu: *Model Apocalypto*, página 157–184. Cambridge University Press, 2017. 9
- [34] Itaú Unibanco: *Riscos e Capital*. páginas 1–64, 2019. 9, 16, 70
- [35] Banco do Brasil: *Relatório de Gerenciamento de Riscos 3T19*. 2019. 9, 70
- [36] Bradesco: *Relatório de Gerenciamento de Riscos 3T19*. 9, 70
- [37] Santander: *Pillar 3 Disclosures*. 2018. 9, 70
- [38] Gil, Antônio Carlos: *Métodos e Técnicas de Pesquisa Social*. Atlas, 6th edição, 2008, ISBN 978-85-224-5142-5. 11
- [39] Vaněk, Tomáš e David Hampel: *The probability of default under ifrs 9: Multi-period estimation and macroeconomic forecast*. Acta Universitatis Agriculturae et Silviculturae Mendelianae Brunensis, 65(2):759–776, 2017, ISSN 12118516. 14, 20

- [40] Chang, E. J., S. M. Guerra, E. J.A. Lima e B. M. Tabak: *The stability-concentration relationship in the Brazilian banking system*. Journal of International Financial Markets, Institutions and Money, 18(4):388–397, 2008, ISSN 10424431. 14
- [41] Central Bank of Brazil: *Time Series Management System*. <https://www3.bcb.gov.br/sgspub/localizarseries/localizarSeries.do?method=prepararTelaLocalizarSeries>, acesso em 2021-03-28. 14
- [42] Costa, João Paulo Vieira, Cleber Mitchell de Lima, Newton Franklin Almeida, Ricardo Matos Chaim e João Carlos Félix Souza: *Model for Dynamics Credit Risk Characterization and Profit Inference in Credit Card Fintechs*. Em *WorldCIST 2021: Trends and Applications in Information Systems and Technologies*, páginas 411–421. 2021, ISBN 9783030726607. 15, 55, 57, 59
- [43] Seitshiro, M. B. e H. P. Mashele: *Quantification of model risk that is caused by model misspecification*. Journal of Applied Statistics, 2020, ISSN 13600532. <https://doi.org/10.1080/02664763.2020.1849055>. 16, 73
- [44] Schwab, Leila, Stefan Gold e Gerald Reiner: *Exploring financial sustainability of SMEs during periods of production growth: A simulation study*. International Journal of Production Economics, 212(October 2018):8–18, 2019, ISSN 09255273. <https://doi.org/10.1016/j.ijpe.2018.12.023>. 16, 55
- [45] Annibal, Clodoaldo Aparecido: *Inadimplência do Seto Bancário Brasileiro: uma avaliação de suas medidas*. Trabalhos para Discussão (BCB), (192):1–36, 2009. 16, 57
- [46] Banco do Brasil: *BB anuncia medidas para reforçar competitividade*, 2019. <https://www.bb.com.br/pbb/pagina-inicial/imprensa/n/59716/#/>, acesso em 2021-06-14. 16
- [47] Zmoginski, Felipe: *Inovação - Entender para Atender*, 2018. <https://noomis.febraban.org.br/temas/inovacao/entender-para-atender>, acesso em 14/06/2021. 16
- [48] Martinez-Plumed, Fernando, Lidia Contreras-Ochando, Cesar Ferri, Jose Hernandez Orallo, Meelis Kull, Nicolas Lachiche, Maria Jose Ramirez Quintana e Peter A. Flach: *CRISP-DM Twenty Years Later: From Data Mining Processes to Data Science Trajectories*. IEEE Transactions on Knowledge and Data Engineering, páginas 1–1, 2019, ISSN 1041-4347. 17
- [49] Lenhardt, W Christopher, Stanley Ahalt, Brian Blanton e Laura Christopher: *Data Management Lifecycle and Software Lifecycle Management in the Context of Conducting Science*. Journal of Open Research Software, 2(1):e15, 2014, ISSN 2049-9647. 17
- [50] KPMG: *Managing model risk: Perspectives from the Indian banking industry*. (March), 2016. 18, 70, 72, 73, 77

- [51] BCBS: *Basel II: International Convergence of Capital Measurement and Capital Standards: A Revised Framework - Comprehensive Version*. Número June. 2006, ISBN 9291316695. 19
- [52] Djeundje, Viani Biatat e Jonathan Crook: *Incorporating heterogeneity and macroeconomic variables into multi-state delinquency models for credit cards*. European Journal of Operational Research, 271(2):697–709, 2018. 19
- [53] Lefebvre-Ulrikson, Williams, G. Da Costa, L. Rigutti e I. Blum: *Data Mining*. 2016, ISBN 9780128047453. 20
- [54] Yeh, I. Cheng e Che hui Lien: *The comparisons of data mining techniques for the predictive accuracy of probability of default of credit card clients*. Expert Systems with Applications, 36(2 PART 1):2473–2480, 2009, ISSN 09574174. 21, 24
- [55] Bellotti, Tony e Jonathan Crook: *Credit scoring with macroeconomic variables using survival analysis*. Journal of the Operational Research Society, 60(12):1699–1707, 2008. 21
- [56] Durović, Andrija: *Macroeconomic Approach to Point in Time Probability of Default Modeling - IFRS 9 Challenges*. Journal of Central Banking Theory and Practice, 8(1):209–223, 2019, ISSN 23369205. 22
- [57] Cox, David R: *Regression models and life-tables*. Journal of the Royal Statistical Society: Series B (Methodological), 34(2):187–202, 1972. 22
- [58] Cox, David R: *Partial likelihood*. Biometrika, 62(2):269–276, 1975. 22
- [59] Breiman, L: *Algorithm cart*. Classification and Regression Trees. California Wadsworth International Group, Belmont, California, 1984. 22, 26
- [60] Bou-Hamad, Imad, Denis Larocque, Hatem Ben-Ameur, Louise C. Mâsse, Frank Vitaro e Richard E. Tremblay: *Discrete-time survival trees*. Canadian Journal of Statistics, 37(1):17–32, 2009, ISSN 03195724. 23
- [61] Cohn, Susan L., Andrew D.J. Pearson, Wendy B. London, Tom Monclair, Peter F. Ambros, Garrett M. Brodeur, Andreas Faldum, Barbara Hero, Tomoko Iehara, David Machin, Veronique Mosseri, Thorsten Simon, Alberto Garaventa, Victoria Castel e Katherine K. Matthay: *The International Neuroblastoma Risk Group (INRG) classification system: An INRG task force report*. Journal of Clinical Oncology, 27(2):289–297, 2009, ISSN 0732183X. 23
- [62] Ishwaran, Hemant, Udaya B Kogalur, Eugene H Blackstone, Michael S Lauer *et al.*: *Random survival forests*. The annals of applied statistics, 2(3):841–860, 2008. 23
- [63] Bellini, Tiziano: *Chapter 3 - lifetime pd*. Em Bellini, Tiziano (editor): *IFRS 9 and CECL Credit Risk Modelling and Validation*, páginas 91 – 153. Academic Press, 2019, ISBN 978-0-12-814940-9. <http://www.sciencedirect.com/science/article/pii/B9780128149409000116>. 24

- [64] Kwamboka Mageto, Dyana: *Modelling of Credit Risk: Random Forests versus Cox Proportional Hazard Regression*. American Journal of Theoretical and Applied Statistics, 4(4):247, 2015, ISSN 2326-8999. 24
- [65] Frydman, Halina e Anna Matuszyk: *Random survival forest for competing credit risks*. Journal of the Operational Research Society, 0(0):1–11, 2020, ISSN 14769360. <https://doi.org/10.1080/01605682.2020.1759385>. 24
- [66] Bai, Miaojun, Yan Zheng e Yun Shen: *Gradient boosting survival tree with applications in credit scoring*. Journal of the Operational Research Society, 0(0):1–17, 2021, ISSN 14769360. <https://doi.org/10.1080/01605682.2021.1919035>. 24, 32
- [67] Xia, Yufei, Lingyun He, Yinguo Li, Yating Fu e Yixin Xu: *A dynamic credit scoring model based on survival gradient boosting decision tree approach*. Technological and Economic Development of Economy, 27(1):96–119, 2021, ISSN 20294921. 24, 32
- [68] R Core Team: *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2019. <https://www.R-project.org/>. 26
- [69] Terry M. Therneau e Patricia M. Grambsch: *Modeling Survival Data: Extending the Cox Model*. Springer, New York, 2000, ISBN 0-387-98784-3. 26
- [70] Mayer, Peter, Denis Larocque e Matthias Schmid: *DStree: Recursive Partitioning for Discrete-Time Survival Trees*, 2014. <https://CRAN.R-project.org/package=DStree>, R package version 1.0. 26
- [71] Ishwaran, H., U.B. Kogalur, E.H. Blackstone e M.S. Lauer: *Random survival forests*. Ann. Appl. Statist., 2(3):841–860, 2008. <http://arXiv.org/abs/0811.1645v1>. 26
- [72] Andreeva, G: *European generic scoring models using survival analysis*. Journal of the Operational Research Society, 57(10):1180–1187, outubro 2006. <https://doi.org/10.1057/palgrave.jors.2602091>. 26
- [73] Blumenstock, Gabriel, Stefan Lessmann e Hsin Vonn Seow: *Deep learning for survival and competing risk modelling*. Journal of the Operational Research Society, 0(0):1–13, 2020, ISSN 14769360. <https://doi.org/10.1080/01605682.2020.1838960>. 32
- [74] Tran, Son Hung e Liem Thanh Nguyen: *Financial development, business cycle and bank risk in Southeast Asian countries*. Journal of Asian Finance, Economics and Business, 7(3):127–135, 2020, ISSN 22884645. 34
- [75] Vazquez, Francisco, Benjamin M. Tabak e Marcos Souto: *A macro stress test model of credit risk for the Brazilian banking sector*. Journal of Financial Stability, 8(2):69–83, 2012, ISSN 15723089. 34, 37, 42, 46
- [76] Gujarati, Damodar N. e Dawn C. Porter: *ECONOMETRIA BÁSICA*. AMGH Editora Ltda., <sup>a5</sup>ª edição, 2011, ISBN 9788580550511. 35, 36



- [77] Hayashi, Fumio: *Econometrics*. 2011, ISBN 9780691010182. 35, 39, 45
- [78] Hendry, David F. e Jean Francois Richard: *The Econometric Analysis of Economic Time Series*. International Statistical Review / Revue Internationale de Statistique, 51(2):111, 1983, ISSN 03067734. 35
- [79] Hahn, Jinyong e Jerry Hausman: *A NEW SPECIFICATION TEST FOR THE VALIDITY OF INSTRUMENTAL VARIABLES*. *Econometrica*, 70(1):163–189, 2002. 36
- [80] Bholat, David, Rosa M. Lastra, Sheri M. Markose, Andrea Miglionico e Kallol Sen: *Non-performing loans at the dawn of IFRS 9: Regulatory and accounting treatment of asset quality*. *Journal of Banking Regulation*, 19(1):33–54, 2018, ISSN 17502071. 36
- [81] Suárez, Javier e Antonio Sánchez Serrano: *Approaching Non-Performing Loans from a Macroprudential Angle*. Reports of the Advisory Scientific Committee, (7), 2018, ISSN 1556-5068. 37
- [82] Bolognesi, Enrica, Cristiana Compagno, Stefano Miani e Roberto Tasca: *Non-performing loans and the cost of deleveraging: The Italian experience*. *Journal of Accounting and Public Policy*, 39(6):106786, 2020, ISSN 18732070. <https://doi.org/10.1016/j.jaccpubpol.2020.106786>. 37
- [83] Messai, Ahlem Selma e Fathi Jouini: *Micro and macro determinants of non-performing loans*. *International Journal of Economics and Financial Issues*, 3(4):852–860, 2013, ISSN 21464138. 37
- [84] Syed, Aamir Aijaz e Ravindra Tripathi: *Non-performing loans in BRICS nations: Determinants and macroeconomic impact*. *Indian Journal of Finance*, 13(2):22–35, 2019, ISSN 09738711. 37
- [85] Beck, Roland, Petr Jakubik e Anamaria Piloiiu: *Key Determinants of Non-performing Loans: New Evidence from a Global Sample*. *Open Economies Review*, 26(3):525–550, 2015, ISSN 09237992. 37, 46
- [86] Anastasiou, Dimitrios, Helen Louri e Mike Tsionas: *Nonperforming loans in the euro area: Are core-periphery banking markets fragmented?* *International Journal of Finance and Economics*, 24(1):97–112, 2019, ISSN 10991158. 37
- [87] Lundberg, Scott M e Su in Lee: *A Unified Approach to Interpreting Model Predictions*. Em *31st Conference on Neural Information Processing Systems*, número Section 2, páginas 1–10, 2017. 38
- [88] EBA: *Report on Machine Learning in Credit Risk Eba Discussion Paper on Machine Learning for Irb*. (November), 2021. 38
- [89] Banco Central do Brasil: *SCR . data - Painel de Operações de Crédito Metodologia Conceitos Gráficos*, 2021. <https://www.bcb.gov.br/content/estabilidade/financeira/scr/scr.data/metodologia.pdf>, acesso em 2021-10-05. 39, 40, 59

- [90] Banco Central do Brasil: *Sistema Gerenciador de Séries Temporais*. <https://www3.bcb.gov.br/sgspub/localizarseries/localizarSeries.do?method=prepararTelaLocalizarSeries>, acesso em 2020-11-08. 39, 40
- [91] Banco Central do Brasil: *Sobre o Sistema de Expectativas de Mercado*. <https://www4.bcb.gov.br/pec/gci/port/expecmercado.asp?frame=1>, acesso em 2021-11-05. 39, 40
- [92] McKinney, Wes: *Data Structures for Statistical Computing in Python*. Em Walt, Stéfan van der e Jarrod Millman (editores): *Proceedings of the 9th Python in Science Conference*, páginas 56–61, 2010. 40
- [93] Van Rossum, Guido e Fred L Drake Jr: *Python reference manual*. Centrum voor Wiskunde en Informatica Amsterdam, 1995. 40, 57, 80
- [94] IHS Global Inc.: *EViews 11 University Edition*, 2020. [www.eviews.com](http://www.eviews.com). 40, 80
- [95] Seabold, Skipper e Josef Perktold: *statsmodels: Econometric and statistical modeling with python*. Em *9th Python in Science Conference*, 2010. 43
- [96] Fuller, Waiye: *Introduction to statistical time series*. JOHN WILEY SONS, UVC, 2nd edição, 1996, ISBN 0471552399. 44
- [97] Phillips, Peter C.B. e Pierre Perron: *Testing for a unit root in time series regression*. *Biometrika*, 75(2):335–346, 1988, ISSN 0006-3444. <https://doi.org/10.1093/biomet/75.2.335>. 44
- [98] Kwiatkowski, Denis, Peter C.B. Phillips, Peter Schmidt e Yongcheol Shin: *Testing the null hypothesis of stationarity against the alternative of a unit root. How sure are we that economic time series have a unit root?* *Journal of Econometrics*, 54(1-3):159–178, 1992, ISSN 03044076. 45
- [99] Levin, Andrew, Chien Fu Lin e Chia Shang James Chu: *Unit root tests in panel data: Asymptotic and finite-sample properties*. *Journal of Econometrics*, 108(1):1–24, 2002, ISSN 03044076. 46
- [100] Im, Kyung So, M. Hashem Pesaran e Yongcheol Shin: *Testing for unit roots in heterogeneous panels*. *Journal of Econometrics*, 115(1):53–74, 2003, ISSN 03044076. 46
- [101] Breiman, Leo: *Random Forests*. *Machine Learning*, 45:5–32, 2001. 49
- [102] Pedregosa, F., G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot e E. Duchesnay: *Scikit-learn: Machine learning in Python*. *Journal of Machine Learning Research*, 12:2825–2830, 2011. 49
- [103] Bussab, Wilton O e Pedro A Morettin: *Estatística Básica*. Saraiva, São Paulo, <sup>a</sup>6ª edição, 2010, ISBN 9788502081772. 52

- [104] Jones, Eric, Travis Oliphant, Pearu Peterson e Others: *SciPy: Open source scientific tools for Python*. <http://www.scipy.org/>. 52
- [105] Wenzler, Josef Stefan: *Dynamics of Global Credit Markets*. SSRN Electronic Journal, 2019, ISSN 1556-5068. 55
- [106] Banco Central do Brasil: *COSIF - Manual de Normas do Sistema Financeiro*. <https://www3.bcb.gov.br/aplica/cosif>, acesso em 26/05/2021. 56
- [107] Porto, Rogério F.: *A brief note on implied historical loss given default*. Journal of Credit Risk, 7(2):73–81, 2011, ISSN 17559723. 62
- [108] Collier, Zachary A e James H Lambert: *Principles and methods of model validation for model risk reduction*. Environment Systems and Decisions, 39(2):146–153, 2019, ISSN 2194-5411. <http://dx.doi.org/10.1007/s10669-019-09723-5>. 69
- [109] European Central Bank: *Targeted Review of Internal Models*. (April), 2021. 69
- [110] Deloitte ToucheTohmatsu Limited: *Model Risk Management Driving the value in modelling*. 2017. 70, 71, 72, 77
- [111] Wu, D. e D. L. Olson: *Enterprise risk management: Coping with model risk in a large bank*. Journal of the Operational Research Society, 61(2):179–190, 2010, ISSN 14769360. 71
- [112] Lamas, Luis, Carlos F. Gallejones, RUBen G. Moral, Juan G. Cascales e Javier Calvo: *Model Risk Management: Quantitative and qualitative aspects*. páginas 1–38, 2014. 72, 77
- [113] Exl, A N e White Paper: *AN EXL WHITE PAPER SR 11-07 Compliance in Model Risk Management*. 72
- [114] Hill, Jon R.: *The top 14 challenges for today's model risk managers: Has the time come to think about going beyond SR11-7?* Journal of Risk Management in Financial Institutions, 12(2):145–167, 2019, ISSN 17528895. 72
- [115] Danielsson, Jon, Kevin R. James, Marcela Valenzuela e Ilknur Zer: *Model risk of risk models*. Journal of Financial Stability, 23:79–91, 2016, ISSN 15723089. <http://dx.doi.org/10.1016/j.jfs.2016.02.002>. 73
- [116] Belton, Valerie e Theodor J Stewart: *Multiple Criteria Decision Analysis: An Integrated Approach*. Dordrecht, The Netherlands. Springer Science Business Media Dordrech, 2002, ISBN 9781461355823. 74, 75
- [117] Brans, J.P. e Ph Vincke: *A Preference Ranking Organisation Method*. Management Science, 31(6):647–656, 1985. 74, 75, 76
- [118] Gul, Muhammet, Erkan Celik, Alev Taskin Gumus e Ali Fuat Guneri: *A fuzzy logic based PROMETHEE method for material selection problems*. Beni-Suef University Journal of Basic and Applied Sciences, 7(1):68–79, 2018, ISSN 23148535. <https://doi.org/10.1016/j.bjbas.2017.07.002>. 74

- [119] Souza, Daniele, Saulo Bezerra, Andreia Oliveira e Sabrina Corrêa: *Aplicação do método Promethee II para gestão de perdas reais em sistemas de abastecimento de água*. Brazilian Journal of Development, páginas 3553–3566, 2020. 74
- [120] Afful-Dadzie, Eric, Stephen Nabareseh, Zuzana Komínková Oplatková e Peter Klimek: *Using fuzzy PROMETHEE to select countries for developmental aid*. Em *Intelligent Systems and Applications. Studies in Computational Intelligence*, volume 650, páginas 109–132. 2016. 75
- [121] Digalwar, Abhijeet K. e Prasanna A. Date: *Development of fuzzy PROMETHEE algorithm for the evaluation of Indian world-class manufacturing organisations*. International Journal of Services and Operations Management, 24(3):308–330, 2016, ISSN 17442389. 75, 77
- [122] Geldermann, Jutta, Thomas Spengler e Otto Rentz: *Fuzzy outranking for environmental assessment. Case study: Iron and steel making industry*. Fuzzy Sets and Systems, 115(1):45–65, 2000, ISSN 01650114. 76, 78
- [123] VPSolutions: *Visual PROMETHEE manual*. 78, 79, 80

## Apêndice A

# Critérios e Escalas para o Risco de Modelo

**Risco de Modelo**

Bloco	Número	Crítério	Escala
Desenvolvimento	1	Quanto à experiência dos analistas envolvidos na técnica de modelagem utilizada:	Alta Média Baixa
	2	Quanto ao suporte da técnica empregada na academia e na indústria para este fim:	Baixo Médio Alto Total
	3	Confiabilidade dos softwares e bibliotecas utilizadas para desenvolvimento do modelo	Alta Média Baixa
	4	Quanto ao atendimento dos pressupostos do modelo desenvolvido:	Não Aplica Total Parcial Não Atendidos
	5	Comparado a modelos para o mesmo uso, quanto o desempenho do modelo é	Não Aplica Pior Semelhante Melhor
Implementação	6	Foram utilizados dados externos ao banco na construção do modelo?	Sim Não
	7	Quando ao nível de intervenção do usuário para processamento das estimativas	Total Parcial Nenhum
Uso	8	Expectativa de descalibragem do modelo por mudança do cenário de desenvolvimento e uso	Alta Média Baixa
	9	A avaliação do desempenho do modelo em uso é feita por equipe independente do uso e do desenvolvimento do modelo?	Sim Não
	10	A tomada de decisão quanto à necessidade de revisão do modelo é independente das equipes de desenvolvimento e uso	Sim Não

**Impacto**

Bloco	Número	Crítério	Escala
Financeiro	1	Volume de perda direta possível pelo uso indevido do modelo	Muito baixo Baixo Médio Alto Muito Alto
Competitividade	2	Relevância do modelo para competitividade do banco no mercado	Muito baixo Baixo Médio Alto Muito Alto
Regulatório/Legal	3	Impacto do modelo quanto ao cumprimento de requisitos legais/regulatórios	Muito baixo Baixo Médio Alto Muito Alto
Impacto nos Riscos Relevantes	4	Impacto do no gerenciamento dos riscos relevantes da instituição	Muito baixo Baixo Médio Alto Muito Alto

## Apêndice B

# Scripts para Simulação da Carteira de Crédito

# Simulação Carteira de Crédito

December 29, 2021

```
[ ]: import numpy as np
import random
import pandas as pd
import numpy_financial as npf

import os
import psycopg2
import psycopg2.extras as ext
import datetime
```

```
[ ]: chave_prob = {1 : 2.00,
                 2: 1.30,
                 3: 1.00,
                 4: 0.90,
                 5: 0.50,
                 6: 0.25}

chave_valr = {1 : 1.00,
              2: 1.50,
              3: 2.00,
              4: 3.50,
              5: 5.65,
              6: 12.00}
```

```
[ ]: class cliente:
    def __init__(self, codigo):
        self.codigo = codigo
        self.porte = int(np.where(np.random.multinomial(1, [0.419, 0.156, 0.167,
↪0.146, 0.072, 0.041], size=1) == 1)[1][0] + 1)
        self.inad = max(0.02 + np.random.normal(0., 0.005), 0.0003) *
↪chave_prob[self.porte]
        self.mtlp = chave_valr[self.porte]
```



## 0.1 Simula base de clientes

```
[ ]: clientes = []

for i in range(3000000):
    clientes.append(cliente(str(i).zfill(7)))
```

```
[ ]: chave_mdld_rec = {1 : 1.000,
                      2: 0.950,
                      3: 0.900,
                      4: 0.750}

chave_atraso_rec = {1: 300,
                   2: 180,
                   3: 150,
                   4: 120}

chave_mdld_valor = {1 : 1.331,
                    2: 0.816,
                    3: 10.00,
                    4: 20.00}

chave_prazo =      {1 : 0,
                   2: 0,
                   3: 0,
                   4: 5}

chave_prt_inad =   {1 : 2.90,
                   2: 1.60,
                   3: 1.30,
                   4: 1.00}

chave_limit =      {1 : 60000,
                   2: 25000,
                   3: 70000,
                   4: 420000}
```

```
[ ]: import matplotlib.pyplot as plt

plt.hist(np.random.lognormal(7.9, 0.1, size=100000)*1.331)
```

```
[ ]: (np.random.lognormal(7.9, 0.1, size=100000)).mean()
```

```
[ ]: #[0.147, 0.779, 0.050, 0.024]
     #[1/4, 1/4, 1/4, 1/4]

class operacao:
    def __init__(self, contrato, ref, mes, dia, choque = 0.0):
```

```

        self.ref = ref
        self.cliente = random.sample(clientes, 1)[0]
        self.modalidade = int(np.where(np.random.multinomial(1, [0.147, 0.789, 0.
↪045, 0.019], size=1) == 1)[1][0] + 1)
        self.valor = min(((np.random.lognormal(7.9, 0.1, size=None)) *
↪chave_mdld_valor[self.modalidade] * self.cliente.mtlp), chave_limit[self.
↪modalidade])
        self.mtpl_inad = chave_prt_inad[self.modalidade]
        self.choque = choque
        self.p_inad = min(self.cliente.inad * self.mtpl_inad * (1 + self.choque)
↪, 0.80)
        self.prazo = int(np.random.uniform(low=1, high=10, size=None)) * 6 * (1
↪+ chave_prazo[self.modalidade])
        self.juros = 0.01
        self.mtpl_rec = chave_mdld_rec[self.modalidade]
        self.parcela = npf.pmt(self.juros, self.prazo, self.valor)
        self.multa = 0
        self.pagamento = 0
        self.contrato = contrato
        self.atraso = 0
        self.mes = mes
        self.dia = dia

```

### 0.1.1 Gerar Operações por dia e agregar na base com o mês

```

[ ]: def contrata(mes, ctr_dia, ref, choque):

    operacoes = []

    for d in range(1, 31):
        q = np.random.poisson(lam=ctr_dia, size=None)

        for i in range(q):
            operacoes.append(operacao(str(mes).zfill(3) + str(d).
↪zfill(2)+(str(i)).zfill(6), ref, mes, d, choque))

    return(operacoes)

```

### 0.1.2 Atualização do mês

#### e inclusão de novas contratações

```

[ ]: def atualiza_mes(anterior, ctr_dia=50, choque = 0.0):
    #print(len(anterior))
    atual = []

    for i in range(len(anterior)):

```

```

    atual.append(anterior[i])

    dia = 30 - anterior[i].dia

    #ajustar atraso para somar 30 dias se menor que 30 dia se não + 30

    p_atraso = np.random.binomial(size = None, n = 1, p = min(anterior[i].
↪p_inad * (1 + choque), 0.99999))

    if anterior[i].atraso < 30:
        temp_atraso = p_atraso * anterior[i].atraso + p_atraso * dia
    else:
        temp_atraso = p_atraso * anterior[i].atraso + p_atraso * 30

    if temp_atraso == 0:
        p = min(anterior[i].p_inad, 0.80)
    elif temp_atraso < 30:
        p = max(anterior[i].p_inad, (0.30 + 0.20 * (temp_atraso - 0)/30) *
↪chave_mdld_rec[atual[i].modalidade])
    elif temp_atraso < 60:
        p = max(anterior[i].p_inad, (0.50 + 0.10 * (temp_atraso - 30)/30) *
↪chave_mdld_rec[atual[i].modalidade])
    elif temp_atraso < 90:
        p = max(anterior[i].p_inad, 0.60 + 0.05 * (temp_atraso - 60)/30 *
↪chave_mdld_rec[atual[i].modalidade])
    elif temp_atraso < 120:
        p = 0.75 * chave_mdld_rec[atual[i].modalidade]
    elif temp_atraso < 150:
        p = 0.80 * chave_mdld_rec[atual[i].modalidade]
    elif temp_atraso < 180:
        p = 0.85 * chave_mdld_rec[atual[i].modalidade]
    elif temp_atraso < 270:
        p = 0.90
    elif temp_atraso < 300:
        p = 0.95
    elif temp_atraso < 330:
        p = 0.975
    elif temp_atraso < 360:
        p = 0.995
    else:
        p = 1.000

    atual[i].cliente.inad = anterior[i].cliente.inad * (p/anterior[i].p_inad)
    atual[i].p_inad = p
    atual[i].prazo = anterior[i].prazo - 1
    atual[i].ref = anterior[i].ref + 1

```

```

    atual[i].valor = max((anterior[i].valor) * (1 + anterior[i].juros) +
↳ ((anterior[i].parcela - atual[i].multa) * (1 - p_atraso)), 0)

    #if atual[i].modalidade == 4 and anterior[i].atraso > 90 and p_atraso ==
↳ 0:
        if anterior[i].atraso > chave_atraso_rec[atual[i].modalidade] and
↳ p_atraso == 0:
            atual[i].pagamento = atual[i].valor
            atual[i].valor = 0
            atual[i].prazo = 0
        else:
            atual[i].pagamento = (-anterior[i].parcela + anterior[i].multa) * (1
↳ - p_atraso)

            atual[i].atraso = temp_atraso

            atual[i].multa = - (anterior[i].parcela * p_atraso * (1 + anterior[i].
↳ juros)) + (anterior[i].multa * p_atraso)

            atual.extend(contrata(anterior[i].mes + 1, ctr_dia, anterior[i].ref, choque
↳ = choque))

    return(atual)

```

## 0.2 Bases geradas

```

[ ]: nmes = 200
mes_choque = [72, 96]
ichoque = 0.15
ctr_dia = 100

m = []
f = []

now = datetime.datetime.now()
print ("Início do processamento base 0 :", now.strftime("%Y-%m-%d %H:%M:%S"))

base0 = contrata(0, ctr_dia, 0, choque = 0.0)
linhas = len(base0)

conn = psycopg2.connect("host=localhost port=5432 dbname=Simulacred
↳ user=postgres password=12345678")
cur = conn.cursor(cursor_factory = ext.DictCursor)

```

```

for n in range(linhas):
    log = []
    sql = "INSERT INTO jpcosta.operacoes( ref, cliente, porte, contrato,
↳modalidade, prazo, valor, mes, dia, prob_opr, prob_cli, atraso, pgto ) VALUES
↳(%s, %s, %s, %s, %s, %s, %s, %s, %s, %s, %s, %s, %s) RETURNING *;"
    values = [base0[n].ref , base0[n].cliente.codigo, base0[n].cliente.porte,
↳base0[n].contrato, base0[n].modalidade, base0[n].prazo, base0[n].valor,
↳base0[n].mes, base0[n].dia, base0[n].p_inad, base0[n].cliente.inad, base0[n].
↳atraso, base0[n].pagamento]
    cur.execute(sql, values)
    conn.commit()
    results = cur.fetchall()
cur.close()
conn.close()

now = datetime.datetime.now()

print ("Fim do processamento base 0      :", now.strftime("%Y-%m-%d %H:%M:%S"))

for i in range(nmes):

    if ((i >= mes_choque[0]) & (i <= mes_choque[1])):
        texto = 'choque de ' + str(ichoque)
        choque = ichoque
    else:
        texto = 'sem choque'
        choque = 0.0

    now = datetime.datetime.now()
    print ("Início do processamento base", str(i + 1), "(", texto, "):", now.
↳strftime("%Y-%m-%d %H:%M:%S"))
    exec("base" + str(i+1) + " = atualiza_mes(" + "base" + str(i) + ", "+
↳str(ctr_dia) + ", choque = "+ str(choque) + ")",)
    exec("linhas = len(base" + str(i + 1) + ")")

    conn = psycopg2.connect("host=localhost port=5432 dbname=Simulacred
↳user=postgres password=12345678")
    cur = conn.cursor(cursor_factory = ext.DictCursor)
    exec("linhas = len(base" + str(i + 1) + ")")

    for n in range(linhas):
        exec("saldo = base" + str(i + 1) + "[n].valor")
        exec("dias = base" + str(i + 1) + "[n].atraso")
        exec("prazo = base" + str(i + 1) + "[n].prazo")
        if saldo < 1.0 and dias == 0 and prazo < 0:

```

```

        None
    else:
        sql = "INSERT INTO jpcosta.operacoes( ref, cliente, porte, contrato,
↳modalidade, prazo, valor, mes, dia, prob_opr, prob_cli, atraso, pgto ) VALUES
↳(%s, %s, %s, %s, %s, %s, %s, %s, %s, %s, %s, %s, %s) RETURNING *;"
        exec("values = [base"+ str(i+1) + "[n].ref " + ", base"+ str(i+1) +
↳"[n].cliente.codigo" + ", base"+ str(i+1) + "[n].cliente.porte" + ", base"+
↳str(i+1) + "[n].contrato" + ", base"+ str(i+1) + "[n].modalidade" + ", base"+
↳str(i+1) + "[n].prazo" + ", base"+ str(i+1) + "[n].valor" + ", base"+ str(i+1)
↳+ "[n].mes" +
        ", base"+ str(i+1) + "[n].dia" + ", base"+ str(i+1) + "[n].
↳p_inad" + ", base"+ str(i+1) + "[n].cliente.inad" + ", base"+ str(i+1) + "[n].
↳atraso" + ", base"+ str(i+1) + "[n].pagamento]", globals(), locals())
        cur.execute(sql, values)
        conn.commit()
        results = cur.fetchall()

        m.append(i)
        exec("f.append(" + "base" + str(i+1) + "[n].p_inad)")

    cur.close()
    conn.close()

    now = datetime.datetime.now()
    print ("Fim do processamento base", str(i + 1), " :", now.
↳strftime("%Y-%m-%d %H:%M:%S"))

```

```

[ ]: cur.close()
conn.close()

```