

Public Report - Módulo 3

Resumo da Sprint 2

Data: 17 a 29 de Agosto de 2025

Objetivo da Sprint

Ajustar e implementar um novo pipeline de simulação para estimar, em horizonte mensal, **(i)** a probabilidade de recessão em cada passo futuro e **(ii)** a probabilidade de ocorrer **pelo menos uma** recessão dentro desse horizonte. A abordagem combina um modelo Probit para o evento binário com um modelo VAR(1) para a dinâmica dos preditores, e usa Simulação de Monte Carlo para propagar incertezas.

Abordagem metodológica (visão geral)

- **Régressão Probit** — modela $P(\text{Recessão}_t = 1 | X_t) = \Phi(\beta_0 + X_t \beta)$. Coeficientes estimados por MLE via Newton-Raphson.
 - **VAR(1)** — modela a evolução temporal dos preditores: $X_t = c + A X_{t-1} + \varepsilon_t$. Parâmetros estimados por OLS.
 - **Monte Carlo** — amostragem de coeficientes Probit (para capturar incerteza em β) + simulação de muitos caminhos futuros dos preditores via VAR(1), cálculo das probabilidades de recessão em cada caminho e agregação estatística (médias, quantis, probabilidade de qualquer recessão).
-

Implementação (componentes e funções principais)

- **Funções utilitárias**
 - `std_norm_cdf`: CDF da normal padrão (usa `scipy.stats.norm.cdf` quando disponível; fallback para `erf`).
 - `_probit_mle`: MLE do Probit com Newton-Raphson — calcula log-verossimilhança, gradiente e Hessiana; adiciona pequenas regularizações e fallback via `pinv` quando necessário.
 - `fit_var1`: estima c , A e Σ (covariância das inovações) por OLS, com verificação mínima de observações.
 - `simulate_var1`: gera caminhos futuros de X usando decomposição de Cholesky com fallback por autovalores para garantir robustez numérica.
- **Função principal:** `monte_carlo_recession`
 - Valida entrada (coluna alvo e preditores).
 - Padroniza preditores (z-scores) para estabilidade numérica.
 - Ajusta Probit e VAR(1).
 - Amostra β a partir de Normal assintótica (`vcov`) e simula `n_sim` caminhos de preditores.
 - Para cada caminho calcula $P_t = \Phi(\beta_i' Z_t)$ e estima se ocorreu ao menos uma recessão (simulando Bernoulli por período).
 - Agrega resultados: média, percentis (5% e 95%), probabilidade de pelo menos uma recessão, estatísticas de β (média/desvio), e Efeitos Marginais Médios (AME) no ponto médio da amostra.
 - Implementa salvaguardas numéricas: jitter para matrizes quase singulares, limites em probabilidades, e alternativas para inversão de matrizes.

- **Exemplo mínimo** no bloco `if __name__ == "__main__":` que gera dados sintéticos, roda a simulação (horizonte 6, `n_sim` 2000) e imprime resultados exemplares.
-

Resultados entregues nesta Sprint

- Pipeline completo e modular para simulação probabilística de recessões, integrando Probit + VAR(1) + Monte Carlo.
 - Implementação robusta com tratamentos numéricos (jitter, fallback Cholesky/eigendecomposition, pinv).
 - Saída consolidada (objeto `MonteCarloOutput`) contendo:
 - `prob_mean`: probabilidade média de recessão por passo do horizonte.
 - `prob_q05` / `prob_q95`: intervalos (5% / 95%) por passo.
 - `prob_any_recession`: probabilidade estimada de pelo menos uma recessão no horizonte.
 - `beta_mean` / `beta_std`: resumo das amostras de β .
 - `ame_at_mean`: Efeitos Marginais Médios dos preditores.
 - `details`: metadados (convergência do Probit, iterações, parâmetros VAR1, médias/desvios das features, número de simulações, etc.).
 - Teste funcional com dados sintéticos demonstrando execução completa e retornos esperados.
-

Observações técnicas e limitações

- **Cálculo de "qualquer recessão"**: no código atual a ocorrência de pelo menos uma recessão por caminho é avaliada simulando Bernoulli por período usando `rng.uniform() < p_t`. alternativa determinística por caminho seria usar `1 - Π(1 - p_t)` (expectativa condicional). As duas abordagens têm interpretações distintas — a primeira simula realizations binárias por período, a segunda computa a probabilidade agregada condicionada aos p_t .
 - **Incerteza do VAR**: atualmente a incerteza nos parâmetros do VAR(1) não é explicitamente amostrada (somente a incerteza em β via `vcov` do Probit é capturada). Para capturar incerteza de dinâmica seria necessário bootstrap ou amostragem bayesiana dos parâmetros do VAR.
 - **Amostra e estabilidade**: VAR(1) exige tamanho mínimo de série (checagem já presente). Em amostras pequenas as estimativas de Σ e A podem ser ruidosas.
 - **Assunções do Probit**: observações independentes condicionais a X_t ; se existir autocorrelação no erro do Probit ou dependências não modeladas, estimativas podem ser viesadas.
 - **Performance**: `n_sim` grande aumenta custo computacional; a função foi escrita para ser vetorizada quando possível, mas pode-se parallelizar as simulações para ganho de velocidade.
-

Resumo da Sprint 3

Data: 01 a 12 de Setembro de 2025

Objetivo da Sprint

A terceira sprint teve como foco **comparar o modelo de Simulação de Monte Carlo com outros métodos probabilísticos**, avaliando desempenho, robustez e capacidade de capturar recessões econômicas. Além disso, foram exploradas **novas abordagens de modelagem e transformação de variáveis**, buscando melhorar a acurácia e estabilidade dos resultados.

Metodologia

Foram testados cinco modelos com diferentes naturezas estatísticas e temporais:

Modelo	Tipo	Evolução Temporal	Amostragem	Acurácia	Precisão	Recall
Régressão Múltipla	Estatístico Linear	Não	Não	N/A	N/A	N/A
Bootstrap Logístico	Estatístico Não Linear	Sim	Não	1.0	0.0	0.0
Cadeia de Markov	Probabilístico Temporal	Não	Sim	0.556	0.0	0.0
Threshold VAR (TVAR)	Estatístico Temporal	Não	Sim	1.0	1.0	1.0
Monte Carlo	Simulação Probabilística	Sim	Sim	0.875	1.0	0.875

Cada método foi avaliado por métricas de desempenho (acurácia, precisão, recall, F1-score, Brier Score e Log Loss), permitindo uma análise comparativa detalhada das capacidades de previsão de recessão.

Principais Resultados e Interpretação

Régressão Múltipla

- **R² (0.845):** Explica 84,5% da variância dos dados, bom desempenho em relações lineares.
- **Limitação:** Não captura dinâmica temporal nem mudanças de regime — inadequada para previsões econômicas complexas.

Bootstrap Logístico

- **Acurácia (1.0) e Precisão/Recall (0.0):** O modelo prevê apenas a classe dominante (não recessão).
- **Limitação:** Sofre com **desbalanceamento de classes**, falhando em identificar recessões reais.

Cadeia de Markov

- **Acurácia (0.556):** Desempenho próximo ao acaso.
- **Brier Score (0.253), Log Loss (1.113):** Indicam baixa calibração probabilística.
- **Limitação:** Suposição de transição de estados baseada apenas no período anterior é simplista para dados econômicos complexos.

Threshold VAR (TVAR)

- **Acurácia, Precisão, Recall, F1 (1.0):** Desempenho perfeito, capturando mudanças de regime.
- **Brier Score (0.354):** Boa calibração, mas sensível ao limiar escolhido e número de defasagens.
- **Limitação:** Pode superajustar (overfitting) em amostras pequenas.

Monte Carlo (modelo que está sendo usado)

- **Acurácia (0.875), Precisão (1.0), Recall (0.875), F1 (0.933):** Excelente equilíbrio entre identificação e precisão de recessões.
 - **Brier Score (0.173):** Probabilidades bem calibradas.
 - **Log Loss (0.538):** Confiança moderada e consistente nas previsões.
 - **Limitação:** Computacionalmente intensivo e sensível à qualidade dos dados de entrada.
-

Síntese Comparativa

- **Monte Carlo:** melhor equilíbrio entre precisão e recall; fornece previsões probabilísticas ricas e interpretáveis.
 - **Threshold VAR:** melhor desempenho geral em classificação binária, mas dependente de parâmetros.
 - **Cadeia de Markov:** estrutura temporal simples, mas com baixa capacidade preditiva.
 - **Bootstrap Logístico:** falha ao lidar com desbalanceamento de classes.
 - **Régressão Múltipla:** útil para explicação estática, mas não para previsão dinâmica.
-

Conclusões da Sprint 3

- O **modelo Monte Carlo** manteve-se robusto, apresentando **boas métricas probabilísticas e estabilidade** mesmo diante de novas variáveis e transformações.
- A **Cadeia de Markov** e o **Bootstrap Logístico** mostraram limitações em capturar recessões reais, reforçando a vantagem de métodos que consideram dependência temporal e incerteza multivariada.
- O **Threshold VAR** destacou-se em desempenho, mas requer validações adicionais para evitar sobreajuste.
- A normalização e reprocessamento das variáveis contribuíram para **melhorar a estabilidade dos modelos probabilísticos**.

A Sprint 3 consolidou a comparação empírica entre métodos probabilísticos, confirmado que o modelo de Simulação de Monte Carlo apresenta o melhor equilíbrio entre desempenho e interpretabilidade. O estudo reforça sua **vantagem para previsão de recessões**, especialmente por integrar dependência temporal e incerteza dos parâmetros. Embora o Threshold VAR tenha atingido desempenho máximo em métricas clássicas, o Monte Carlo mantém maior flexibilidade e robustez, justificando sua continuidade e refinamento no módulo seguinte.

Resumo da Sprint 4

Data: 14 a 26 de Setembro

Objetivo da Sprint

A quarta sprint teve como propósito **avaliar a resiliência e a generalização do modelo de Simulação de Monte Carlo** através de três abordagens principais:

1. **Validação cruzada (cross-validation)** — medir o poder discriminativo e a calibração das previsões.

2. **Stress testing** — testar a resposta do modelo sob cenários econômicos extremos (recessões severas, leves e booms).
3. **Rolling validation (validação temporal contínua)** — avaliar a estabilidade do modelo ao longo de diferentes janelas históricas, simulando previsões em série.

O objetivo final foi **detectar possíveis casos de overfitting** e **mensurar a consistência preditiva** do modelo em situações adversas.

Metodologia Aplicada

Três tipos de teste foram conduzidos com base no código do modelo `monte_carlo_recession` desenvolvido nas sprints anteriores:

◆ Validação Cruzada (TimeSeriesSplit)

- **Divisão:** 5 partições temporais.
- **Métricas utilizadas:**
 - **AUC (Area Under the Curve)** — mede a capacidade de distinguir recessões (classe positiva) de não recessões.
 - **Brier Score** — mede a calibração das probabilidades (quanto menor, melhor).

◆ Stress Testing

- Simulação de três cenários macroeconômicos extremos:
 - *Mild recession*: queda moderada de PIB e CLI.
 - *Severe recession*: queda acentuada de PIB, CLI e USSLIND.
 - *Boom*: forte crescimento econômico e melhora dos indicadores.
- Para cada cenário, foram geradas novas simulações do modelo Monte Carlo com 2000 iterações.

◆ Rolling Validation

- **Estrutura:** janela inicial de 120 observações, com previsões sucessivas horizonte-a-horizonte.
 - **Métricas:** AUC e Brier Score em cada janela, para medir estabilidade temporal.
-

Resultados Obtidos

Validação Cruzada

Métrica	Valor Médio	Interpretação
AUC	<i>nan</i>	O modelo não conseguiu gerar valores úteis de discriminação (problema de separabilidade entre classes).
Brier Score	0.423	Indica calibração moderada — as probabilidades estão relativamente distantes dos valores reais.

O desempenho foi limitado. A ausência de AUC mostra dificuldade em distinguir entre recessões e não recessões, enquanto o Brier Score aponta que as previsões, embora calibradas, ainda carecem de precisão.

Stress Testing

Cenário	Probabilidade de Recessão Estimada
Mild recession	1.000
Severe recession	1.000
Boom	1.000

O modelo reagiu de forma determinística, atribuindo 100% de probabilidade de recessão em todos os cenários — mesmo em períodos de expansão.

Essa resposta extrema pode indicar:

- **Sobreajuste (overfitting)** do modelo às condições históricas.
- **Falta de sensibilidade** a variações nas entradas em regimes econômicos diferentes.

Rolling Validation

Métrica	Valor Médio	Observações
AUC médio	<i>nan</i>	O modelo não apresentou capacidade discriminativa ao longo das janelas temporais.
Brier médio	0.434	Calibração levemente pior que na validação cruzada.
Brier_all	0.0 → 1.0	Forte variabilidade, indicando instabilidade temporal.

Em algumas janelas, o modelo apresentou previsões quase perfeitas, enquanto em outras, falhou completamente. Isso demonstra falta de consistência temporal — possivelmente por dependência de períodos específicos do treinamento ou variáveis altamente correlacionadas com o tempo.

Conclusões da Sprint 4

- **Validação cruzada:** desempenho limitado e ausência de AUC, indicando que o modelo **não discrimina bem os estados econômicos**.
- **Stress testing:** comportamento determinístico (100%), sugerindo **overfitting** e **baixa adaptabilidade a novos regimes**.
- **Rolling validation:** desempenho instável ao longo do tempo, com grande variação nos Brier Scores.

O modelo de Simulação de Monte Carlo mostrou boa estrutura probabilística, mas ainda não apresenta generalização satisfatória. Ele tende a reagir de forma extrema e pouco calibrada quando submetido a novos contextos econômicos ou cenários de stress. Os testes indicaram problemas de generalização e estabilidade, com o modelo apresentando respostas determinísticas sob stress e baixa discriminação temporal nas validações cruzadas e móveis. Apesar dessas limitações, o processo foi crucial para identificar fragilidades estruturais e orientar o refinamento metodológico necessário para a etapa final do projeto, que buscará aumentar a robustez, calibragem e sensibilidade do modelo frente a cenários econômicos distintos.

Resumo da Sprint 5

Data: 29 de Setembro a 09 de Outubro

Objetivo da Sprint

A Sprint 5 teve como principal objetivo **consolidar os resultados obtidos nas sprints anteriores e integrá-los formalmente ao Trabalho de Conclusão de Curso (TCC)**.

Nesta etapa, o foco foi a **sistematização do conteúdo técnico e analítico** desenvolvido nos módulos anteriores em seções estruturadas do documento final, apresentando de forma acadêmica:

1. A **estrutura prática dos modelos agregados à Simulação de Monte Carlo**;
2. A **comparação entre regressões (Logística e Probit)** e suas implicações no contexto econômico;
3. A **análise de contraprova dos modelos** frente à Simulação de Monte Carlo;
4. E a **apresentação detalhada dos resultados obtidos**, com métricas, validações e interpretações teóricas.

O objetivo geral desta sprint foi **transformar os resultados empíricos e experimentais do projeto em conteúdo científico documentado**, com rigor acadêmico e coesão metodológica.

Desenvolvimento e Conteúdo Produzido

Durante a Sprint 5, quatro novas seções foram redigidas e adicionadas ao TCC, baseadas no trabalho técnico realizado ao longo das Sprints 2, 3 e 4.

Conclusão da Sprint 5

A Sprint 5 representou o marco de **integração final do conteúdo técnico e teórico ao documento do TCC**, com destaque para:

- Estruturação das seções **4.0 a 5.0**, incorporando metodologia, comparações e resultados;
- Adaptação dos relatórios técnicos (Sprints 2–4) à linguagem acadêmica;
- Organização das métricas, tabelas e análises em formato científico;
- Consolidação das conclusões e recomendações para trabalhos futuros.

Esses avanços consolidam o modelo proposto como um **sistema probabilístico completo para previsão de recessões**, validado por múltiplas abordagens e com análise crítica das suas limitações e potencialidades.

A Sprint 5 marcou a **transição da fase experimental para a fase de documentação científica**.

O modelo de Simulação de Monte Carlo, inicialmente desenvolvido e testado em módulos independentes, foi formalmente descrito, justificado e comparado em termos de desempenho.

Com isso, o TCC passou a incorporar não apenas os resultados quantitativos, mas também a **fundamentação metodológica e teórica necessária** para consolidar a pesquisa como uma contribuição original no estudo da **imprevisibilidade das recessões econômicas e da aleatoriedade dos ciclos econômicos**.

Submissão para Congressos:

https://github.com/Inteli-College/2025-1A-T01-G18-INTERNO/tree/main/Modulo%203/Sprint%205/Submissao_Congressos

- **CNMAC-2025** - Congresso Nacional de Matemática Aplicada e Computacional
- **ICSE SEIP** - 47th International Conference on Software Engineering
- **ICSE 2026 NIER Track** - 48th International Conference on Software Engineering