



RIO

PUC

Dissertação de Mestrado

Framework de Apoio a Tomada de Decisão em um Portfólio de Comercialização de Gás Natural Baseado em Simulações de Monte Carlo

Nicholas De Marco Reckman

Pontifícia Universidade Católica do Rio de Janeiro
Centro Técnico Científico
Departamento de Engenharia Industrial

Rio de Janeiro 16 de Setembro, 2025



Pontifícia
Universidade
Católica do
Rio de Janeiro

Dissertação de Mestrado

Framework de Apoio a Tomada de Decisão em um Portfólio de Comercialização de Gás Natural Baseado em Simulações de Monte Carlo

Nicholas De Marco Reckman

Orientação: Professor Igor Tona Peres

Dissertação apresentada como requisito parcial para a obtenção do grau de Mestre em Logística (Opção Profissional) pelo programa de Pós-Graduação em Engenharia de Produção, do Departamento de Engenharia industrial da PUC-Rio.

Rio de Janeiro 16 de Setembro, 2025



Pontifícia
Universidade
Católica do
Rio de Janeiro

Framework de Apoio a Tomada de Decisão em um Portfólio de Comercialização de Gás Natural Baseado em Simulações de Monte Carlo

Nicholas De Marco Reckman

Dissertação apresentada como requisito parcial para a obtenção do grau de Mestre em Logística (Opção Profissional) pelo programa de Pós-Graduação em Engenharia de Produção, do Departamento de Engenharia industrial da PUC-Rio.

Professor Igor Tona Peres

Orientador

Departamento de Engenharia Industrial – PUC-Rio

Professor Fernando Luiz Cyrino Oliveira

Departamento de Engenharia Industrial – PUC-Rio

Professor Erick Meira de Oliveira

Financiadora de Estudos e Projetos – FINEP

Rio de Janeiro 16 de Setembro, 2025



Pontifícia
Universidade
Católica do
Rio de Janeiro

Todos os direitos reservados. A reprodução, total ou parcial, do trabalho é proibida sem autorização da universidade, da autora e do orientador.

Nicholas De Marco Reckman

Graduou-se em Engenharia de Produção pela PUC-RIO (Pontifícia Universidade Católica do Rio de Janeiro) em 2018. Desde então trabalha no mercado de energia, especializando-se na logística de Gás Natural, modelagem financeira e otimização de custos.

Ficha Catalográfica

Reckman, Nicholas De Marco

Framework de Apoio a Tomada de Decisão em um Portfólio de Comercialização de Gás Natural Baseado em Simulações de Monte Carlo / Nicholas De Marco Reckman; orientador: Igor Tona Peres. – 2025.

98 f: il. Color. ; 30 cm

Dissertação (mestrado) - Pontifícia Universidade Católica do Rio de Janeiro, Departamento de Engenharia industrial, 2025.

Inclui bibliografia

1. Engenharia Industrial – Teses. 2. Novo Mercado de Gás. 3. Simulação de Monte Carlo. 4. Modelagem Matemática. 5. Gás Natural. I. Peres, Igor Tona. II. Pontifícia Universidade Católica do Rio de Janeiro. Departamento de Engenharia industrial.

CDD: 658.5

Agradecimentos

Este trabalho é fruto de uma jornada desafiadora e enriquecedora, que só foi possível com o apoio e a dedicação de diversas pessoas, às quais expresso o meu mais profundo apreço.

Em primeiro lugar, agradeço ao meu orientador, Professor Igor Tona Peres, cuja orientação precisa, paciência e vasto conhecimento foram fundamentais para o desenvolvimento desta dissertação. Suas contribuições foram cruciais para que eu pudesse amadurecer no campo acadêmico.

Agradeço a minha namorada, Evelyn, que sempre esteve ao meu lado durante esse processo, me apoiando sempre com muito amor e palavras de carinho.

Agradeço também aos demais professores e colegas do programa do Mestrado Profissional em Logística, que sempre se mostraram dispostos a compartilhar conhecimentos e experiências.

Aos meus amigos, que estiveram ao meu lado nos momentos mais desafiadores, oferecendo apoio, compreensão e palavras de incentivo, meu muito obrigado.

A minha família, que sempre acreditou em mim e me proporcionou amor e suporte incondicional. Sem o apoio de meus pais, Johan e Fátima, e de minha irmã, Bianca, esta jornada não teria sido possível.

Por fim, agradeço a todas as pessoas que, de alguma forma, contribuíram para a conclusão deste trabalho. Cada um de vocês, tem a minha sincera e total gratidão.

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

Resumo

Reckman, Nicholas De Marco; Peres, Igor Tona (orientador). **Framework de Apoio a Tomada de Decisão em um Portfólio de Comercialização de Gás Natural Baseado em Simulações de Monte Carlo**. Rio de Janeiro, 2025. 98p. Dissertação de Mestrado – Departamento de Engenharia industrial, Pontifícia Universidade Católica do Rio de Janeiro.

A comercialização de gás natural é uma atividade complexa que envolve incertezas significativas, relacionadas tanto às condições de mercado quanto às variações na demanda e na oferta do produto. Com a implementação da Nova Lei do Gás no Brasil e a abertura gradual do mercado, surgiu a figura do consumidor livre de gás, trazendo novas dinâmicas comerciais e exigindo ferramentas avançadas para a gestão eficiente do portfólio de comercialização. Neste contexto, a presente dissertação propõe um framework de apoio à tomada de decisão para a gestão de portfólios de comercialização de gás natural, incluindo o desenvolvimento da modelagem matemática deste problema e da resolução do mesmo por meio de técnicas de simulação. O framework foi totalmente desenvolvido em Python para o escopo de operações diárias e visa auxiliar a figura do comercializador na otimização de suas estratégias de compra, venda e transporte de gás natural em níveis estratégicos e táticos. Analisando dados históricos de demanda de diversos consumidores e fornecedores de gás natural presentes em um dado portfólio, moldam-se padrões e curvas de distribuição que são replicadas com métodos estatísticos, simulando a operação diária de portfólio ao longo de um dado período. Além de simular perfis de demanda, o framework proposto também simula macro variáveis que afetam a operação e gestão do portfólio como a taxa de câmbio e balanceamento de saldo do mercado. Esse trabalho utiliza o caso específico de uma empresa comercializadora de gás natural, definida como um agente, fazendo previsões e tomando decisões sobre seu portfólio no ano operacional de 2026. Os resultados obtidos demonstram que a ferramenta é capaz de fornecer insights valiosos, identificando estratégias mais rentáveis e mitigando riscos associados às oscilações de demanda de gás natural.

Palavras-chave

Novo Mercado de Gás; Simulação de Monte Carlo; Modelagem Matemática; Gás Natural.

Abstract

Reckman, Nicholas De Marco; Peres, Igor Tona (Advisor). **Decision Support Framework for a Natural Gas Commercialization Portfolio Based on Monte Carlo Simulations**. Rio de Janeiro, 2025. 98p. Dissertação de Mestrado – Departamento de Engenharia industrial, Pontifícia Universidade Católica do Rio de Janeiro.

Natural gas trading is a complex activity that involves significant uncertainties related to both market conditions and fluctuations in product demand and supply. With the implementation of the New Gas Law in Brazil and the gradual opening of the market, the concept of the free gas consumer emerged, introducing new commercial dynamics and requiring advanced tools for efficient portfolio management. In this context, this dissertation proposes a decision support framework for managing natural gas trading portfolios, including the development of the mathematical model and its resolution through simulation techniques. The framework was entirely developed in Python for the scope of daily operations and aims to assist traders in optimizing their natural gas purchase, sale, and transportation strategies at strategic and tactical levels. By analyzing historical demand data from various consumers and natural gas suppliers within a given portfolio, patterns and distribution curves are shaped and replicated using statistical methods, simulating the daily portfolio operation over a given period. In addition to simulating demand profiles, the proposed framework also simulates macro variables that affect portfolio operations and management, such as exchange rates and market balance adjustments. This work uses the specific case of a natural gas trading company, defined as an agent, making forecasts and decisions about its portfolio in the 2026 operational year. The results obtained demonstrate that the tool can provide valuable insights, identifying more profitable strategies and mitigating risks associated with fluctuations in natural gas demand.

Keywords

New Gas Market; Monte Carlo Simulation; Mathematical Modeling; Natural Gas.

Sumário

1	Introdução	14
2	Referencial Teórico	17
2.1	O Mercado de Gás Natural e sua Transformação no Brasil	17
2.2	Gestão de Portfólios de Comercialização de Energia	18
2.3	Simulações de Monte Carlo na Comercialização de Gás Natural	19
2.4	Modelagem e Previsão de Séries Temporais	19
2.5	Frameworks de Apoio à Tomada de Decisão	20
3	Formulação do Problema	22
3.1	Contexto Operacional	22
3.2	Lacunas analíticas e de Modelagem	22
3.3	Desafios Computacionais e Estruturais	23
3.4	Pergunta de Pesquisa	24
4	Proposta	26
4.1	Visão Geral do Framework Composto	27
4.2	Estrutura de Dados e Parâmetros do Modelo	29
4.3	Simulação de Monte Carlo	34
4.4	Modelagem Matemática	38
4.5	Dados e arquivos de apoio	48
5	Análise e Resultados	49
5.1	Caso de Estudo	49
5.2	Implementação do Framework	53
5.3	Sensitividade do Modelo	63
5.4	Discussão dos Resultados	66
6	Conclusão	73
7	Referências bibliográficas	76
8	Apêndice	80
8.1	Artigo de complemento ao trabalho	80

Lista de figuras

Figura 4.1	Arquitetura geral do framework proposto	28
Figura 4.2	Fluxograma do tratamento e análise estatística das séries de demanda	33
Figura 4.3	Fluxograma do processo de simulação: escopo global e diário com variáveis de decisão e incertezas	36
Figura 4.4	Fluxograma diário do processo de programação de gás entre os agentes do sistema	37
Figura 5.1	Series temporais utilizadas como base para os consumidores no caso base	54
Figura 5.2	Índices de correlação entre os consumidores no caso base	55
Figura 5.3	Função de autocorrelação dos consumidores no caso base	56
Figura 5.4	Função de autocorrelação parcial dos consumidores no caso base	57
Figura 5.5	Fluxograma geral da simulação de Monte Carlo com sensibilidade das QDTC	65
Figura 5.6	Valor Presente Líquido (VPL) Unitário Médio em função da QDTC para cada zona de transporte.	67
Figura 5.7	Distribuições empíricas do VPL Unitário nas QDTC ótimas de cada zona de transporte.	69
Figura 5.8	Distribuição do VPL Unitário do portfólio com todas as QDTCs ótimas.	71

Lista de tabelas

Tabela 4.1	Principais distribuições de probabilidade consideradas e suas formas características	31
Tabela 4.2	Índices utilizados na modelagem matemática	39
Tabela 4.3	Valores considerados como parâmetros na modelagem e simulação	40
Tabela 4.4	Variáveis de tomada de decisão na modelagem e simulação	41
Tabela 5.1	Parâmetros contratuais dos consumidores no caso de estudo	49
Tabela 5.2	Parâmetros contratuais dos fornecedores no caso de estudo	50
Tabela 5.3	Parâmetros operacionais das operadoras de transporte	51
Tabela 5.4	Zonas de transporte utilizadas na simulação	52
Tabela 5.5	Parâmetros globais da simulação	53
Tabela 5.6	Resultados do teste ADF (p-valores) para as séries de consumo	58
Tabela 5.7	Resultados da calibração dos modelos ARMA para as séries de demanda e comercializadores.	59
Tabela 5.8	Resultados dos testes de estacionariedade aplicados aos resíduos dos modelos ARMA.	60
Tabela 5.9	Distribuições de probabilidade ajustadas aos resíduos das séries de demandas industriais.	61
Tabela 5.10	Tabelas de distribuição de probabilidades utilizadas na modelagem das demandas dos comercializadores.	62
Tabela 5.11	Variáveis de decisão sensibilizadas na simulação de Monte Carlo	64
Tabela 5.12	Resumo por zona: QDTC ótima, VPL unitário médio, intervalo de confiança, desvio-padrão e melhora percentual.	70
Tabela 5.13	Resumo cenário ótimo conjunto: VPL unitário médio, desvio-padrão e melhora percentual.	71

Lista de Abreviaturas

ADF	<i>Augmented Dickey–Fuller</i> (Teste de raiz unitária)
ARIMA	<i>AutoRegressive Integrated Moving Average</i>
BPMN	<i>Business Process Model and Notation</i>
CAPEX	<i>Capital Expenditure</i> (Investimentos)
CG	Capital de Giro
CO	Custos Operacionais
CPV	Custo dos Produtos Vendidos
CSLL	Contribuição Social sobre o Lucro Líquido
DA	Depreciação e Amortização
EALIM	Limite percentual de Excedente Autorizado
ECNU	Encargo de Capacidade Não Utilizada
ESEA	Encargo de Excedente Autorizado
ESENA	Encargo de Excedente Não Autorizado
FAC	Função de Autocorrelação (equivalente a ACF)
FC	Fluxo de Caixa
GNL	Gás Natural Liquefeito
IRPJ	Imposto de Renda da Pessoa Jurídica
JUR	Juros
KS	Teste de Kolmogorov–Smirnov
LAIR	Lucro Antes do Imposto de Renda
LAJIRDA	Lucro Antes de Juros, Imposto de Renda, Depreciação e Amortização
LIMDES	Limite de Desequilíbrio
LSMC	<i>Least Squares Monte Carlo</i>
MLE	<i>Maximum Likelihood Estimation</i> (Máxima Verossimilhança)
MSA	<i>Master Service Agreement</i> (Contrato Master de Serviço)
NTS	Nova Transportadora do Sudeste
PDES	Penalidade de Desequilíbrio

PM	Preço da Molécula
PT	Preço de Transporte
PVAR	Penalidade por Variação de Programação
QDC	Quantidade Diária Contratada
QDCNU	Quantidade Diária de Capacidade Não Utilizada
QDP	Quantidade Diária Programada
QDR	Quantidade Diária Realizada
QDTC	Quantidade Diária de Transporte Contratada
QEA	Excedente Autorizado (quantidade)
QENA	Excedente Não Autorizado (quantidade)
RF	Resultado Financeiro
RL	Resultado Líquido
RV	Receita de Vendas
SARIMA	<i>Seasonal ARIMA</i>
SDP	Saldo de Programação
SG&A	<i>Selling, General and Administrative</i> (Despesas Gerais e Administrativas)
SM	Suprimento Máximo
SOP	<i>Ship-or-Pay</i>
TARIFF	Tarifa de Transporte (parâmetro de zona)
TBG	Transportadora Brasileira Gasoduto Bolívia–Brasil
TOP	<i>Take-or-Pay</i>
VARLIM	Limite de variação de programação
VARMULT	Multiplicador da penalidade por variação
VPL	Valor Presente Líquido
WACC	<i>Weighted Average Cost of Capital</i> (Taxa de Desconto)

*Todos os modelos são incorretos, mas alguns
são úteis.*

George Box, *Estatístico*.

1

Introdução

O setor de gás natural tem assumido papel cada vez mais relevante no contexto energético mundial. Ele é frequentemente apontado como um combustível de transição, emitindo significativamente menos gases de efeito estufa do que o carvão e o óleo combustível, o que possibilita um alívio imediato à pegada de carbono global, ao mesmo tempo que suporta a integração de fontes renováveis intermitentes (S&P Global, 2024; CHEN et al., 2021).

No cenário brasileiro, esse tema ganha ainda mais relevância. O país possui reservas consideráveis de gás natural, demanda energética em expansão e um mercado ansioso por fontes alternativas e diversificação energética (Oxford Institute for Energy Studies, 2023). A promulgação da Nova Lei do Gás (Lei nº 14.134, de 8 de abril de 2021) estabeleceu um novo marco regulatório, promovendo o desmonte de monopólios verticais, a abertura à entrada de novos agentes e maior acesso à infraestrutura básica, como gasodutos e terminais de importação.

Essas transformações culminaram no surgimento do chamado Novo Mercado de Gás, que concede ao consumidor livre a possibilidade de negociar diretamente contratos de fornecimento, rompendo com a dependência histórica das distribuidoras locais na posição de intermediários obrigatórios na cadeia do gás. Contudo, essa liberdade traz um conjunto de complexidades adicionais: negociação de múltiplos contratos com diferentes modalidades entre múltiplas partes diferentes, desafios logísticos na alocação de capacidade de transporte e exposição à volatilidade dos preços no curto prazo e variáveis macroeconômicas como câmbio e política climática, todos aumentando substancialmente a incerteza na tomada de decisão.

Nesse contexto, a dinâmica de comercialização de gás envolve a interação entre um número crescente de partes, definidas como agentes, que deverão lidar com esses fatores de volatilidade e incerteza ao longo da cadeia de suprimento. Entre os principais, destacam-se:

- **Produtores e supridores:** responsáveis pela oferta da molécula de gás natural, oriunda tanto da produção doméstica quanto da importação via terminais de gás natural liquefeito (GNL) ou gasodutos internacionais.
- **Comercializadores:** agentes que intermedeiam contratos de compra e venda de gás, estruturando portfólios para capturar valor econômico e mitigar riscos contratuais e operacionais, estes podem também atuar na produção ou consumo final de gás.

- **Consumidores livres:** indústrias ou grandes usuários finais que, a partir da Lei nº 14.134/2021, podem negociar diretamente com fornecedores e comercializadores, rompendo a obrigatoriedade de contratação com distribuidoras locais.
- **Distribuidoras locais:** embora tenham perdido a exclusividade no fornecimento a grandes consumidores, permanecem responsáveis pelo atendimento ao mercado cativo e pela operação das redes de distribuição de baixa e média pressão em áreas de concessão designadas.
- **Operadores de transporte:** empresas concessionárias que gerenciam os gasodutos de transporte, garantindo acesso não discriminatório à infraestrutura e aplicando tarifas reguladas pela Agência Nacional do Petróleo, Gás Natural e Biocombustíveis (ANP).
- **Regulador:** responsável pela fiscalização, pela regulação tarifária e pela garantia da concorrência e transparência no setor.

A articulação entre esses e outros agentes torna o processo de comercialização particularmente complexo, uma vez que combina decisões de caráter contratual, logístico e regulatório, em um ambiente de elevada incerteza e competitividade.

Diante desse cenário, os agentes comercializadores desse novo mercado não podem se basear em médias históricas ou cenários determinísticos. É preciso recorrer a abordagens quantitativas robustas que capturem corretamente a dimensão estocástica dos fenômenos. Dentro deste contexto vêm o destaque a abordagem do uso de estudos e análises estocásticas como a simulação de Monte Carlo, capazes de gerar milhares de cenários alternativos com base em distribuições probabilísticas, oferecendo uma visão amplificada da incerteza envolvida, assim como a construção de modelos de análise e previsão de séries temporais, essenciais para estimar variáveis-chave como demanda e preços se baseando em padrões de dados históricos.

O uso da simulação estocástica de Monte Carlo fornece aos agentes do mercado métodos analíticos robustos para apoiar a tomada de decisão. Dessa forma, é possível avaliar simultaneamente o valor esperado das estratégias e a variabilidade associada ao risco. Essa abordagem permite responder a questões críticas, como: quais contratações maximizam o valor esperado? Qual o risco associado à alocação atual de capacidade de transporte nos gasodutos? E como eventos adversos externos afetam a rentabilidade e a segurança operacional?

Com base nesses fundamentos, esta dissertação tem como objetivo desenvolver um framework integrado para a análise de portfólios de comercialização de gás natural, considerando o contexto do Novo Mercado de Gás brasileiro

sob a perspectiva do agente comercializador. Esse framework analisa e modela toda a formulação matemática para ser aplicada dentro de uma simulação de Monte Carlo, oferecendo apoio analítico à tomada de decisão em cenários complexos e incertos do mercado de gás natural. Ele se diferencia por permitir a avaliação de contratos, transporte e demanda sob uma ótica probabilística, gerando resultados que subsidiam tanto decisões estratégicas quanto operacionais diárias.

Estrutura da dissertação

Para melhor organização do trabalho, a dissertação está estruturada da seguinte forma:

- **Capítulo 1 – Introdução:** apresenta o contexto do setor de gás natural no Brasil, os desafios do Novo Mercado de Gás e os objetivos gerais da pesquisa.
- **Capítulo 2 – Referencial Teórico:** fundamentação conceitual sobre os temas abordados na dissertação.
- **Capítulo 3 – Formulação do Problema:** delimitação do problema operacional e analítico, com definição das perguntas de pesquisa.
- **Capítulo 4 – Proposta:** descrição detalhada do framework metodológico, incluindo arquitetura, modelagem, simulação e implementação.
- **Capítulo 5 – Estudo de Caso e Resultados:** aplicação prática do framework em um agente comercializador real do Novo Mercado de Gás, com análise de resultados e sensibilidade.
- **Capítulo 6 – Conclusões:** síntese das contribuições acadêmicas e práticas, limitações do estudo e sugestões para pesquisas futuras.

2

Referencial Teórico

O presente capítulo reúne os principais fundamentos teóricos que sustentam a pesquisa, fornecendo a base conceitual e metodológica para a formulação e implementação do framework proposto. São apresentados os elementos centrais do setor de gás natural no Brasil, sua evolução regulatória e as implicações para a atuação dos agentes comercializadores. Em seguida, discutem-se conceitos de gestão de portfólios de energia, análise de risco e incerteza, bem como ferramentas estatísticas e computacionais relevantes.

A organização do capítulo visa não apenas contextualizar o leitor quanto às transformações estruturais do mercado, mas também destacar as contribuições da literatura em termos de modelagem estocástica, previsão de séries temporais e aplicação de simulações de Monte Carlo. Por fim, apresentam-se os principais frameworks e abordagens de apoio à tomada de decisão em ambientes caracterizados por elevada complexidade e volatilidade, preparando o terreno para a proposta metodológica desenvolvida nos capítulos subsequentes.

2.1

O Mercado de Gás Natural e sua Transformação no Brasil

O gás natural ocupa posição estratégica na matriz energética global por ser por ser amplamente reconhecido como um vetor de transição energética, dada sua menor intensidade de emissões de gases de efeito estufa em comparação com o carvão e o petróleo (International Energy Agency, 2022). Nos últimos anos, o Brasil passou por profundas transformações regulatórias no setor, impulsionadas pela promulgação da Nova Lei do Gás (Lei nº 14.134/2021), que buscou estabelecer um mercado mais competitivo, aberto e eficiente (Secretaria-Geral da União, 2021).

Historicamente, o mercado brasileiro de gás natural era caracterizado por forte concentração na Petrobras, tanto na produção quanto no transporte e distribuição (Agência Nacional do Petróleo, Gás Natural e Biocombustíveis (ANP), 2020). Esse arranjo resultava em barreiras à entrada de novos agentes e limitava a liquidez do mercado. Com a Nova Lei do Gás, buscou-se promover a desconcentração, estimulando a competição na infraestrutura de transporte e criando condições para a figura do consumidor livre, que pode negociar diretamente seu fornecimento com produtores e comercializadores, desvinculando-se das distribuidoras locais (Empresa de Pesquisa Energética, 2022).

Entretanto, ainda persistem obstáculos estruturais importantes — como

a concentração de mercado, atrasos regulatórios e falta de acesso transparente à infraestrutura — que dificultam a mudança efetiva para um mercado competitivo (Confederação Nacional da Indústria (CNI), 2025). A introdução do consumidor livre representa uma mudança estrutural na dinâmica de comercialização, impondo maior complexidade na gestão dos portfólios e na avaliação de risco dos agentes de mercado (CARDOSO et al., 2023).

Assim, a consolidação do Novo Mercado de Gás no Brasil exige dos comercializadores maior sofisticação analítica para lidar com volatilidade de preços, incertezas regulatórias e desafios logísticos. Nesse contexto, ferramentas baseadas em modelagem estocástica e simulações probabilísticas tornam-se fundamentais para apoiar a tomada de decisão desses agentes, tanto no nível estratégico quanto no operacional.

2.2

Gestão de Portfólios de Comercialização de Energia

A gestão de portfólios no contexto da comercialização de energia — especialmente nos mercados brasileiros de energia elétrica e gás natural — baseia-se em práticas quantitativas robustas para equilibrar risco e retorno. Modelos inspirados na Teoria Moderna de Portfólio (MPT), como otimização média-variância ou CVaR, são amplamente utilizados para suportar decisões contratuais frente à volatilidade dos preços e à variabilidade da demanda (MENEZES; LAUREANO, 2014).

Outros estudos exploram a aplicação de CVaR na otimização de portfólios de energia elétrica no Brasil, considerando o horizonte de recomendação de contratação de longo prazo e as peculiaridades do mercado específico do país (GUERREIRO, 2017). Isso evidencia o uso avançado de métricas de risco além do desvio padrão tradicional, com foco na minimização de perdas extremas.

Além disso, modelos acadêmicos de otimização estocástica aplicados à cadeia de valor do gás natural demonstram como agentes comerciais podem gerir simultaneamente contratos de longo prazo, mercado spot e capacidade de transporte para maximizar retornos e mitigar riscos operacionais (TOMAS-GARD et al., 2004). Em paralelo, na prática de mercado, comercializadores de gás natural liquefeito (GNL) estão implementando capacidades avançadas de trading e otimização de portfólio, incluindo arbitragem entre mercados e uso de contratos spot, para capturar oportunidades e melhorar a resiliência financeira (McKinsey & Company, 2021).

2.3

Simulações de Monte Carlo na Comercialização de Gás Natural

As simulações de Monte Carlo constituem um método amplamente utilizado para a avaliação quantitativa de risco e o apoio à tomada de decisão em estratégias de formação de portfólios e gestão contratual na comercialização de gás natural. No contexto da otimização de infraestrutura, como armazenamento, transporte e operações spot, os modelos baseados em Monte Carlo permitem simular milhares de cenários para preços de gás, volumes, restrições logísticas e premissas contratuais, possibilitando a análise probabilística de margens e flexibilidade de portfólio (JONG, 2015).

Além disso, métodos como Least Squares Monte Carlo (LSMC) têm sido aplicados com sucesso na precificação de contratos de armazenamento de gás, usando simulações de trajetórias de preço em conjunto com regressões para estimar o valor ótimo da operação sob condições estocásticas complexas (MALYSCHIEFF; TRAFALIS, 2017). Essa abordagem adota uma visão dinâmica e adaptativa, importante para que os agentes comerciais possam reagir a flutuações de mercado com agilidade e precisão.

Por fim, em cenários corporativos, simulações de Monte Carlo são amplamente utilizadas para estimar métricas financeiras como Valor Presente Líquido (VPL) e Value-at-Risk (VaR), fornecendo uma base estatística robusta para decisões de trade, hedge de contratos e avaliação de exposição de mercado (MURTHA, 1997).

2.4

Modelagem e Previsão de Séries Temporais

A previsão de séries temporais é fundamental para estimar tanto preços futuros quanto demandas. Modelos clássicos como ARIMA (AutoRegressive Integrated Moving Average) e suas variantes sazonais (SARIMA) constituem a base que alimenta estudos mais avançados sobre previsão e séries temporais (BOX et al., 2015).

Para preços de gás natural especificamente, abordagens mais sofisticadas empregam modelos probabilísticos baseados em *state-space*, capazes de lidar com caudas pesadas, heterocedasticidade condicional, e assimetria. Esses modelos incorporam variáveis exógenas como temperatura, níveis de estocagem e preços correlacionados (como carvão e petróleo), resultando em melhorias significativas no desempenho preditivo levando a reduções de até 13% no *Continuous Ranked Probability Score* (CRPS) em previsões day-ahead em comparação com benchmarks tradicionais (BERRISCH; ZIEL, 2020).

Um aspecto relevante é a integração entre modelos de previsão de séries temporais e simulações de Monte Carlo. As previsões estatísticas ou de aprendizado de máquina podem ser utilizadas para gerar distribuições condicionais de preços e demandas, que posteriormente alimentam o processo de simulação estocástica. Essa abordagem permite representar não apenas a média esperada da série, mas também a incerteza associada, propagando intervalos de confiança ou distribuições preditivas completas dentro dos cenários de risco (WERON, 2006).

Adicionalmente, é importante destacar que a presente dissertação é complementada por um artigo de autoria do próprio pesquisador, atualmente em processo de publicação e reproduzido integralmente na Seção 8 (Apêndice). Esse artigo aprofunda a discussão acerca da aplicação de técnicas de previsão de séries temporais no contexto da demanda de gás natural, explorando casos empíricos específicos e ampliando a gama de metodologias avaliadas em relação ao corpo principal da tese. Dessa forma, o material apresentado no apêndice não apenas expande os fundamentos teóricos aqui discutidos, como também fornece evidências adicionais da aplicabilidade prática das abordagens de previsão consideradas, fortalecendo a coerência entre a fundamentação teórica e a implementação do framework proposto (RECKMAN, 2025).

2.5

Frameworks de Apoio à Tomada de Decisão

Em ambientes de comercialização de gás natural, a tomada de decisão sob incerteza requer frameworks capazes de integrar, de forma coerente, técnicas de previsão, simulação e otimização estocástica. Um exemplo recente é o uso de políticas de controle estocástico multi-estágio aplicadas a redes de gás com flexibilidade de linepacking para lidar com incertezas em tempo real, mantendo a viabilidade das operações e reduzindo a variabilidade do sistema (DVORKIN et al., 2022).

Outra abordagem importante é o modelo de controle ótimo estocástico de inventário em redes de gás, que incorpora dinâmica da tubulação, restrições operacionais e métricas de risco no objetivo, permitindo comparar estratégias determinísticas, robustas e estocásticas e demonstrando claramente a vantagem do modelo estocástico em termos de volatilidade e desempenho econômico (ZAVALA, 2013).

Além disso, em sistemas integrados de eletricidade e gás, frameworks com otimização estocástica são capazes de quantificar o valor da incorporação de incertezas em demanda, oferta renovável, preços de combustíveis e emissões de carbono, o que ajuda a informar decisões estratégicas e garantir robustez nos

balanços entre diferentes setores energéticos (RIEPIN, 2020).

Por fim, um framework eficiente utilizado em energia combina a programação matemática com simulações de Monte Carlo, permitindo seleções de portfólios com avaliação robusta mesmo com múltiplas fontes de incerteza, distinguindo opções seguras, rejeitadas ou que requerem análise adicional, o que resulta em recomendações mais confiáveis para tomada de decisão (MAVROTAS, 2013).

3

Formulação do Problema

A gestão de portfólios de comercialização de gás natural em um ambiente regulatório em transformação, como o brasileiro após a promulgação da Nova Lei do Gás (Lei nº 14.134, de 8 de abril de 2021), demanda novas abordagens metodológicas para lidar com a crescente complexidade e incerteza dos processos de decisão. Esta seção apresenta uma formulação analítica do problema enfrentado pelos agentes comercializadores de gás natural, organizada em três principais dimensões complementares, o contexto operacional, as lacunas analíticas de modelagem e os desafios computacionais, com o objetivo de estabelecer as bases conceituais que orientarão o desenvolvimento do framework proposto.

3.1

Contexto Operacional

O Novo Mercado de Gás brasileiro estabelece um modelo mais aberto e competitivo, no qual consumidores livres podem negociar diretamente seus contratos com fornecedores, desintermediando as distribuidoras locais. Essa nova estrutura resulta em um aumento significativo do número de contratos e da complexidade dos fluxos de movimentação de gás, exigindo dos comercializadores decisões mais frequentes, estratégicas e integradas.

Além disso, as operações comerciais passam a ser mais sensíveis a fatores externos ao portfólio de comercialização como taxas de câmbio, variações de demanda setorial, penalidades contratuais por desequilíbrios, e tarifas reguladas de transporte. A gestão eficiente do portfólio requer, portanto, ferramentas que permitam analisar diferentes cenários de forma probabilística e dinâmica.

3.2

Lacunas analíticas e de Modelagem

A crescente sofisticação do setor de gás natural, impulsionada pela liberalização regulatória e pela intensificação da competição entre agentes, impõe desafios relevantes ao processo de tomada de decisão. Embora haja ampla disponibilidade de dados históricos e uma evolução contínua nas técnicas de modelagem preditiva, observa-se uma lacuna significativa no uso integrado desses recursos em ferramentas voltadas especificamente à gestão de portfólios de comercialização. A maioria das soluções ainda se apoia em métodos determinísticos ou em análises baseadas em médias históricas, que não cap-

turam adequadamente a volatilidade e a correlação entre as diversas variáveis envolvidas.

Nesse contexto, é fundamental desenvolver modelos que representem de forma estocástica as operações de comercialização de gás natural, considerando a incerteza associada à demanda, aos preços e aos fluxos contratuais. Tais modelos devem refletir a simultaneidade entre múltiplos contratos de suprimento e consumo, a flexibilidade operacional exigida pelo mercado e os mecanismos contratuais complexos envolvidos, como cláusulas de encargos de capacidade e penalidades por desequilíbrio. A ausência de ferramentas que considerem essas interações limita a capacidade do comercializador de estimar com precisão os impactos financeiros de suas decisões em diferentes cenários de mercado.

Além disso, as variações de curto e médio prazo na demanda e nas condições de fornecimento requerem modelos de previsão robustos, que consigam lidar com a sazonalidade, tendência e choques exógenos, aspectos frequentemente presentes em séries temporais do setor energético. O desafio é ampliar a capacidade preditiva com uso de métodos estatísticos e computacionais, como modelos ARIMA, SARIMA ou técnicas mais recentes de aprendizado de máquina, mantendo o equilíbrio entre precisão e interpretabilidade. A integração entre previsão de séries temporais e simulação de Monte Carlo, através da simulação de séries temporais, desponta como uma alternativa promissora para gerar cenários realistas e apoiar a avaliação de risco e retorno de diferentes configurações do portfólio.

3.3

Desafios Computacionais e Estruturais

A implementação prática das soluções propostas envolve desafios computacionais significativos. Nesse contexto, o desenvolvimento de um framework de apoio à decisão, entendido como uma estrutura computacional modular destinada a auxiliar o processo decisório sob condições de incerteza, aplicável à rotina operacional de comercializadores de gás natural, requer uma estrutura que seja ao mesmo tempo modular, escalável e eficiente em termos de processamento. Isso implica, primeiramente, na capacidade de integrar diversas fontes de dados heterogêneas, como históricos de consumo, preços de contratos, parâmetros de transporte e dados macroeconômicos, em um sistema unificado, apto a gerar insights acionáveis em tempo hábil.

O sistema deve permitir a geração automatizada de cenários estocásticos a partir de séries temporais previstas e simuladas com seus resíduos inclusos como variáveis aleatórias encaixadas em curvas de distribuição calibradas com base em dados reais. É fundamental que tais cenários reflitam as interdepen-

dências entre variáveis (por exemplo, entre demanda e preço ou entre consumo e penalidades logísticas), garantindo que as simulações reproduzam de forma fidedigna o comportamento do sistema em operação. A lógica de construção do modelo deve, portanto, estar fortemente alicerçada em testes estatísticos.

Outro aspecto importante refere-se à avaliação comparativa de estratégias. O framework deve permitir a simulação de múltiplas combinações de ações, como diferentes contratos de compra e venda, níveis de capacidade contratada no transporte ou estratégias de programação diária de volumes, e apresentar resultados consolidados com base em métricas financeiras reconhecidas, como o Lucro antes dos Juros, Impostos de Renda, Depreciação e Amortização (LAJIRDA), o Fluxo de Caixa (FC) e o Valor Presente Líquido (VPL), com uma atenção maior nesse estudo sobre a avaliação do VPL, essas métricas fornecem subsídios quantitativos para a escolha entre alternativas.

Adicionalmente, é desejável que a ferramenta seja desenvolvida em uma linguagem de programação amplamente adotada no meio técnico, como o Python, por sua versatilidade, facilidade de integração com bibliotecas estatísticas e visualização de dados, e pela possibilidade de customização e manutenção por equipes internas. Essa abordagem garante a adaptabilidade da solução aos diferentes perfis de comercializadores, bem como sua aplicabilidade em contextos reais de operação no Novo Mercado de Gás.

3.4

Pergunta de Pesquisa

Com base na contextualização apresentada, evidencia-se que o ambiente de comercialização de gás natural no Brasil demanda um novo paradigma de suporte à decisão. A transformação regulatória em curso, o aumento da competitividade e a complexidade das operações comerciais impõem aos agentes a necessidade de tomar decisões sob elevada incerteza, tanto no curto quanto no médio prazo. Nesse contexto, torna-se imprescindível o desenvolvimento de soluções analíticas que consigam representar fielmente as variáveis envolvidas, capturar seus comportamentos estocásticos e fornecer subsídios concretos para a construção de estratégias otimizadas e resilientes.

A formulação de modelos computacionais baseados em simulações de Monte Carlo se apresenta como uma abordagem promissora para enfrentar os desafios descritos. Tal técnica permite explorar a variabilidade dos cenários de mercado e antecipar os efeitos de decisões operacionais e contratuais em múltiplas condições de contorno. No entanto, a aplicação dessas metodologias exige uma estruturação cuidadosa, que integre dados históricos utilizados na calibração das simulações, regras de negócio que definem os limites operacionais

e contratuais, restrições operacionais representativas das condições reais de transporte e suprimento e métricas financeiras consistentes para avaliação de desempenho econômico.

Diante dessa complexidade, a presente dissertação se propõe a investigar uma questão central que orienta todo o desenvolvimento metodológico e técnico do trabalho. Essa pergunta busca refletir a interseção entre o problema prático enfrentado pelos agentes comercializadores e a lacuna de conhecimento identificada no domínio acadêmico e profissional da engenharia de produção aplicada ao setor energético.

Como desenvolver um framework de apoio à decisão que auxilie comercializadores de gás natural na definição de estratégias sob incertezas regulatórias, operacionais e econômicas?

Essa pergunta de pesquisa não apenas delimita o escopo da investigação, mas também orienta o desenho do framework proposto, guiando as decisões sobre quais variáveis considerar, quais técnicas utilizar e como validar a efetividade da ferramenta desenvolvida. Ao respondê-la, espera-se contribuir com uma solução prática e replicável que possa ser adotada por agentes reais de mercado, além de gerar conhecimento aplicável à modelagem de sistemas complexos sob incerteza, ampliando o repertório metodológico disponível na área de tomada de decisão em sistemas energéticos.

4

Proposta

A partir da formulação do problema apresentada no capítulo anterior, esta seção descreve em detalhes a estrutura metodológica e computacional proposta para apoiar a tomada de decisão na gestão de portfólios de comercialização de gás natural em ambientes de alta incerteza. O objetivo central é desenvolver um framework computacional que permita simular diferentes configurações operacionais e contratuais de comercialização de gás natural, avaliando os impactos econômicos decorrentes dessas combinações e fornecendo subsídios quantitativos à tomada de decisão sob condições de incerteza.

A proposta consiste em um framework desenvolvido integralmente em linguagem Python, estruturado para operar de forma modular e adaptável à realidade de comercializadores que atuam sob as diretrizes do Novo Mercado de Gás brasileiro. Esse framework integra os seguintes componentes principais:

1. **Modelagem estatística da demanda:** análise de séries temporais de consumo, aplicação de testes de auto-correlação (ACF) e auto-correlação parcial (PACF) nas séries, assim como testes de estacionariedade e de correlação entre as séries.
2. **Geração de cenários:** criação de múltiplas trajetórias possíveis de consumo e suprimento a partir das distribuições ajustadas, através de projeções das curvas de consumo avaliadas com um fator de erro estocástico baseado na distribuição do erro na fase de treinamento do modelo de previsão das séries.
3. **Simulação estocástica via Monte Carlo:** execução de múltiplas iterações sobre as curvas de demanda geradas para cada variação de uma variável de decisão no modelo de comercialização.
4. **Modelagem financeira:** consolidação dos resultados em métricas econômico-financeiras, com destaque para o Valor Presente Líquido (VPL).
5. **Integração e análise:** comparação entre estratégias de operação ou parâmetros contratuais sensibilizados nos cenários, permitindo avaliar o retorno esperado, a dispersão e a robustez dos resultados.

O modelo permite a consideração simultânea de múltiplos consumidores e supridores, respeitando as restrições contratuais e operacionais típicas do setor,

como cláusulas de take-or-pay e ship-or-pay, limites de transporte, e regras de balanceamento de rede. Ao longo deste capítulo, cada um dos componentes será apresentado em detalhes, com ênfase nos dados de entrada, fundamentos teóricos adotados, estrutura matemática e lógica computacional.

A seção também contempla a aplicação prática da metodologia a um estudo de caso representativo, utilizando dados operacionais projetados para o período de um mês de 31 dias. Com isso, busca-se não apenas validar o funcionamento do framework, mas também demonstrar sua aplicabilidade como ferramenta de apoio à decisão no contexto real de operação de agentes comercializadores de gás natural.

4.1

Visão Geral do Framework Composto

O framework desenvolvido nesta dissertação tem como objetivo descrever e operacionalizar uma estrutura computacional de apoio à decisão aplicada ao planejamento e à análise de portfólios de comercialização de gás natural, considerando as incertezas operacionais e econômicas que caracterizam esse setor. Diante de um ambiente de mercado cada vez mais volátil e fragmentado, a proposta busca auxiliar o comercializador na identificação de estratégias mais eficientes sob diferentes condições de consumo, fornecimento e estrutura contratual.

Concebido com base em princípios de modelagem estocástica e análise econômica, o framework permite simular o desempenho financeiro de um portfólio ao longo de um horizonte operacional pré-determinado, com granularidade diária. A estrutura geral do framework é apresentada na Figura 4.1, que resume o encadeamento lógico entre os módulos de entrada, processamento e análise.

No estado atual do desenvolvimento, o principal foco do componente estocástico está na demanda dos consumidores finais. Essas demandas são tratadas como previsões de modelos auto-regressivos de média móvel (ARMA) treinados nos dados históricos de cada consumidor, onde o fator de aleatoriedade é incluso na simulação de Monte Carlo através de distribuições aleatórias encaixadas nos resíduos de seus modelos ARMA treinados.

A simulação de Monte Carlo é utilizada para explorar a incerteza nas demandas de gás natural. Para cada valor sensibilizado da variável de decisão, que pode ser, por exemplo, o volume contratado com supridores ou a capacidade de transporte alocada, são gerados 100 cenários independentes do mesmo mês simulado, com resolução diária. Cada cenário representa uma possível realização futura da operação do portfólio, a partir das distribuições

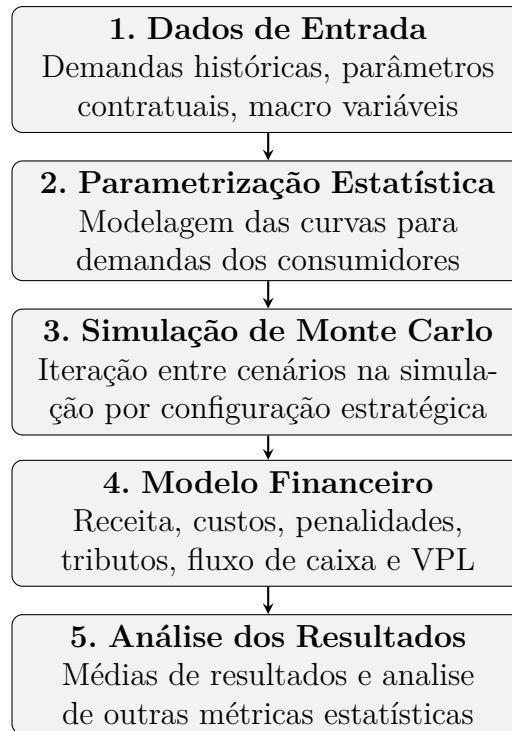


Figura 4.1: Arquitetura geral do framework proposto

de demanda dos consumidores.

Os resultados operacionais simulados são utilizados como entrada em um modelo de fluxo de caixa descontado, que consolida, para cada cenário, as receitas, os custos operacionais, as penalidades contratuais (como as cláusulas take-or-pay e ship-or-pay), os tributos incidentes e as variações no capital de giro. A métrica central de avaliação adotada é o Valor Presente Líquido (VPL), amplamente utilizado em análises financeiras por refletir o valor econômico agregado sob diferentes condições de incerteza.

Com isso, o framework permite ao comercializador comparar diferentes estratégias sob uma perspectiva quantitativa, considerando o desempenho médio do VPL como base, sua dispersão entre cenários e a sensibilidade do resultado a alterações na variável decisória. Embora o modelo, na versão atual, avalie exclusivamente o VPL como métrica de decisão, sua estrutura modular permite futuras extensões para incorporar indicadores adicionais de risco, como Value-at-Risk (VaR) ou Conditional Value-at-Risk (CVaR), ou ainda medidas operacionais específicas.

A próxima seção detalha os dados utilizados e a estrutura de parametrização das variáveis de entrada, com foco especial na modelagem das demandas e na representação estatística dos agentes do portfólio.

4.2

Estrutura de Dados e Parâmetros do Modelo

A estruturação adequada dos dados de entrada é um dos pilares fundamentais do framework proposto. A simulação estocástica da operação do portfólio exige a representação realista da variabilidade da demanda dos consumidores, que é a principal fonte de incerteza considerada nesta versão do modelo. Para isso, foram utilizadas séries históricas de consumo diário de gás natural, segregadas por ponto de entrega e por agente consumidor, com o objetivo de capturar os padrões estatísticos individuais de comportamento. As séries de consumo foram submetidas a testes estatísticos com o propósito de verificar suas propriedades fundamentais para a etapa de simulação.

Dentre essas análises, destacam-se os testes de estacionariedade e de autocorrelação. Para verificar a estacionariedade, foram aplicados testes formais como o Augmented Dickey-Fuller (ADF) (DICKEY; FULLER, 1979), que avaliam a presença de raiz unitária nas séries temporais. As hipóteses nulas desses testes são testadas e, se rejeitadas com significância estatística, indicarão que as séries não apresentam tendência persistente nem heterocedasticidade temporal relevante. Além disso, são calculadas as funções de autocorrelação (ACF) e autocorrelação parcial (PACF) (BOX et al., 2015) para séries temporais de consumo, a ausência de autocorrelação significativa reforça a suposição de independência temporal entre os dados ou não.

De acordo com os resultados destes testes estatísticos sobre os dados históricos de demanda, serão adequados e treinados modelos de previsão ARMA individuais para cada consumidor, após o treino destes modelos de previsão, os resíduos dos mesmos serão encaixados em distribuições de probabilidade aleatórias a serem utilizadas na simulação para replicar o fator estocástico dessas demandas.

Previamente à construções das curvas de consumo de cada consumidor que serão utilizadas na simulação, o framework realiza uma análise de correlação entre as séries históricas de demanda dos agentes do portfólio, com o objetivo de identificar possíveis interdependências e definir a estrutura conjunta de simulação. Essa análise é conduzida por meio do coeficiente de correlação de Pearson (PEARSON, 1895) (ou outro método, conforme aplicável), permitindo avaliar a intensidade e a direção da relação linear entre as variáveis de consumo. Esse processo visa identificar possíveis dependências lineares entre os perfis de consumo, o que pode ter implicações significativas na geração dos cenários estocásticos, especialmente no caso de consumidores cujo comportamento é sensível a fatores macroeconômicos ou operacionais comuns.

Para cada par de consumidores (x_i, x_j) , calcula-se o coeficiente de cor-

relção de Pearson (PEARSON, 1895) entre suas séries temporais de demanda diária. Esse coeficiente, denominado por ρ_{x_i, x_j} é dado por:

$$\rho_{x_i, x_j} = \frac{Cov(D_{X_i}, D_{X_j})}{\sigma_{D_{X_i}} \times \sigma_{D_{X_j}}} \quad (4-1)$$

Nesse calculo (4-1) D_{x_i} e D_{x_j} são os vetores de demanda diária dos consumidores x_i e x_j respectivamente, e $\sigma_{D_{x_i}}, \sigma_{D_{x_j}}$ são seus respectivos desvios padrão. Além da correlação linear, o framework permite a inspeção visual por meio de matrizes de dispersão (scatter plots) e heatmaps da matriz de correlação, auxiliando na detecção de padrões de dependência não triviais.

Com base nos valores obtidos, são adotadas as seguintes interpretações operacionais:

- Se $|\rho_{x_i, x_j}| < 0,3$, considera-se que não há correlação linear significativa;
- Para $0,3 \leq |\rho_{x_i, x_j}| < 0,7$, a correlação é considerada moderada;
- Valores de $|\rho_{x_i, x_j}| \geq 0,7$ indicam alta correlação e sugerem dependência estatística relevante.

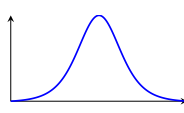
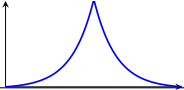
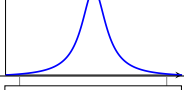
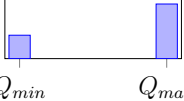
Nos casos em que a correlação entre consumidores for considerada significativa, o framework pode ser estendido para gerar amostras de demandas correlacionadas por meio de técnicas como transformações de Cholesky (GOLUB; LOAN, 2013) aplicadas à matriz de covariância, ou métodos baseados em cópulas (copula-based sampling). Entretanto, na versão atual deste trabalho, os coeficientes de correlação de Pearson observados entre os principais consumidores apresentaram valores inferiores a 0,7, não indicando evidências de dependência sistemática entre as séries. Dessa forma, as curvas de demanda foram tratadas como estatisticamente independentes, decisão metodologicamente consistente diante da ausência de correlação relevante e visando à simplificação computacional do processo de simulação.

Essa validação estatística da independência é fundamental para justificar o uso de simulações Monte Carlo com variáveis não correlacionadas e a modelagem de séries temporais utilizada para gerar os cenários na simulação. Com isso, evita-se a introdução de complexidade desnecessária sem evidência empírica que a justifique, mantendo o modelo robusto, interpretável e computacionalmente eficiente.

Após a validação da independência estatística entre os perfis de demanda, verificada por meio do teste de correlação, procede-se para a modelagem de séries temporais para cada consumidor industrial incluso na simulação, onde cada uma destas séries é treinada sobre um modelo ARMA de previsão (BOX et al.,

2015) com os resíduos do modelo treinado sendo associados a curvas de distribuições aleatórias que representem os aspectos estocásticos dessas demandas na simulação de Monte Carlo. Sobre estes resíduos serão aplicados testes estatísticos de aderência como Kolmogorov-Smirnov (SMIRNOV, 1948), levando também em consideração aspectos empíricos de aderência como comparação entre histogramas e densidades teóricas. Foram consideradas, prioritariamente, distribuições do tipo Normal, t-Student, Cauchy e Laplace para descrever os resíduos das séries modeladas, conforme pode-se observar na Tabela 4.1. Nos casos de modelagem de demandas de outros clientes comercializadores, esta será replicada de forma estacionária, com possíveis valores de consumo com probabilidades definidas, uma vez que estes consumidores não tem processos industriais definindo os seus volumes futuros e, conseqüentemente, não apresentam mudanças de suas propriedades estatísticas ao longo do tempo. Essa possibilidade de distribuição também esta descrita na Tabela 4.1.

Tabela 4.1: Principais distribuições de probabilidade consideradas e suas formas características

Distribuição	Parâmetros	Forma da distribuição
Normal	μ (média) σ (desvio padrão)	
t-Student	μ (localização) s (escala) ν (g.l.)	
Laplace	μ (localização) b (escala)	
Cauchy	x_0 (localização) γ (escala)	
Tabelar	$P(Q = Q_{max})$ Q_{min}, Q_{max}	

Além da incerteza inerente à demanda dos consumidores, o framework proposto incorpora uma importante dimensão operacional: o erro de programação dos volumes diários movimentados nas malhas de gasodutos de transporte. Esse erro representa a discrepância entre o volume programado por um agente comercializador e o volume efetivamente realizado no decorrer do dia em um determinado ponto da malha de gasodutos. No contexto do mercado de gás natural, essa discrepância pode ter impacto financeiro direto devido à aplicação de penalidades por desequilíbrio, encargos de variação de programação

e restrições de alocação logística, especialmente em ambientes regulados com contratos complexos entre os agentes transportadores e os comercializadores usuários do transporte.

De acordo com o que é observado nos dados históricos de operação do caso de estudo, o erro de programação se comporta e será tratado como uma variável aleatória, adicionada sobre o volume programado $QDP_{x,d}$ de um agente x em um determinado dia d . O erro, definido como $\epsilon_{x,d}$ é um fator relativo relacionado à programação e será considerado como uma distribuição normal com $\mu = 0$ e σ sendo sensibilizado como parâmetro para cada agente considerado na simulação.

A modelagem é embasada na observação da distribuição dos erros de programação, assim como sua média e desvio padrão. Adicionalmente, foram aplicados testes de normalidade como Kolmogorov-Smirnov (SMIRNOV, 1948) e Shapiro-Wilk (SHAPIRO; WILK, 1965) para avaliar o encaixe do erro em uma distribuição gaussiana.

É importante destacar que determinados agentes na simulação do caso de estudo deste trabalho, agentes consumidores e fornecedores de gás natural, não estão sujeitos a erros de variação de programação. Isso se dá devido ao fato que estes agentes entregam ou recebem gás de forma virtual dentro da malha de gasodutos, com o gás não se movimentando de fato na transação mas simplesmente tendo sua titularidade trocada entre as partes no sistema de transporte, eliminando as causas de erro. Esses agentes mantêm esse comportamento de transferência no duto porque, por sua vez, também são comercializadores de gás natural no mercado, e preferem manter o controle e a flexibilidade sobre a operação de transporte em seus próprios portfólios como revendedores do gás até seus respectivos clientes consumidores.

Outros dados inseridos no modelo como parâmetros incluem dados contratuais do portfólio em análise como volumes contratados, limites operacionais, tarifas de transporte e penalidades previstas. Esses parâmetros serão tratados como determinísticos e são aplicados diretamente na simulação financeira com base nos volumes simulados de consumo.

Variáveis macroeconômicas, como taxa de câmbio, não foram consideradas estocásticas nesta fase do modelo. No entanto, o framework foi projetado de forma modular para permitir a incorporação futura de tais variáveis, seja por simulação de séries temporais, distribuições empíricas ou projeções externas.

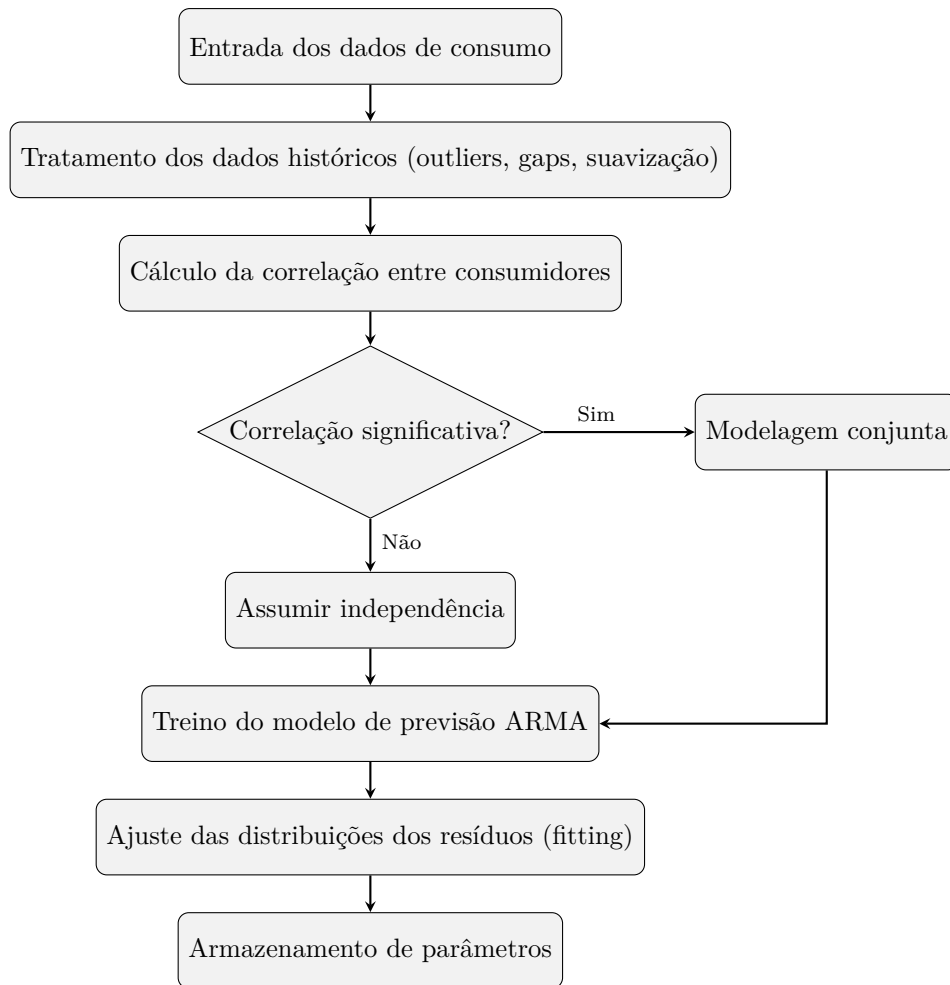


Figura 4.2: Fluxograma do tratamento e análise estatística das séries de demanda

Para consolidar o processo descrito nesta seção, a Figura 4.2 apresenta um fluxograma representando de forma resumida as etapas de tratamento, validação estatística e preparação das séries históricas de demanda dos consumidores a serem analisados na simulação do portfólio. O diagrama destaca os principais pontos de decisão, como a presença de correlação significativa entre agentes ou a verificação de estacionariedade, e os caminhos alternativos resultantes dessas análises. Esse procedimento estabelece a base estatística que sustenta a geração dos cenários estocásticos na próxima etapa do modelo, garantindo que os dados utilizados na simulação de Monte Carlo sejam consistentes, realistas e devidamente parametrizados.

Com os dados devidamente tratados, validados e parametrizados estatisticamente conforme descrito, o próximo passo do framework consiste na geração de cenários estocásticos por meio de simulação de Monte Carlo, conforme detalhado na Seção 4.3.

4.3

Simulação de Monte Carlo

A simulação de Monte Carlo constitui o módulo central de modelagem probabilística do framework proposto, sendo responsável pela geração de múltiplos cenários operacionais que incorporam as incertezas associadas à demanda dos consumidores e à execução contratual com os fornecedores. Esses cenários são utilizados para avaliar o desempenho financeiro de diferentes estratégias de contratação e operação sob condições que permitam replicar resultados realistas que deem apoio à tomada de decisão.

Para cada simulação, é considerado um horizonte de 31 dias consecutivos, com granularidade diária. Ainda que tanto os contratos de fornecimento e transporte de gás quanto os compromissos com os consumidores finais são majoritariamente estabelecidos em bases anuais, a decisão pela análise limitada a um mês se dá pelo fato de que não há mudanças na lógica de custo ao longo do ano e que há poucos dados históricos de consumo para se simular curvas de consumo previstas para horizontes maiores de previsão. A escolha do número de cenários simulados por estratégia não está fixada, sendo definida com base em critérios de estabilidade estatística dos resultados, especialmente da convergência dos valores médios e da variabilidade do Valor Presente Líquido (VPL) estimado. No caso estudado, utiliza-se uma quantidade de 100 iterações simuladas por cenário, mas esse valor pode ser ajustado conforme o nível de precisão desejado.

As curvas de consumo utilizadas nos cenários da simulação, são geradas a partir dos modelos preditivos ARMA treinados projetados para o período da simulação com a adição dos resíduos do modelo treinado colocados como distribuições aleatórias sorteados sobre as previsões para cada consumidor industrial, conforme descrito na Seção 4.2. A cada iteração, valores pseudoaleatórios são amostrados dessas distribuições que compõem os resíduos nas curvas simuladas, respeitando a independência entre as variáveis e entre os consumidores.

Além da variabilidade de demanda, são incluídos dois tipos de erro de programação de volumes de gás natural que adicionam um aspecto prevalente de incerteza sobre o modelo: (i) o erro de programação do consumidor, que representa a diferença entre o volume diário de entrega programado para o consumidor e o volume efetivamente consumido naquele dia, e (ii) o erro de entrega dos fornecedores, que modela a diferença entre o volume solicitado pelo comercializador e o volume efetivamente entregue pelos fornecedores. Ambos os erros são tratados como variáveis aleatórias Gaussianas centradas em zero e somadas sobre os volumes programados para cada dia.

É importante destacar que, na versão atual do framework, as variáveis de decisão modeladas concentram-se exclusivamente no nível operacional, com ênfase nos processos de contratação e operação da malha de gasodutos de transporte. Essa escolha se deve ao maior grau de complexidade e impacto financeiro envolvido na logística de transporte de gás natural, bem como ao potencial de otimização relevante nessa etapa. Em contrapartida, dimensões comerciais mais amplas, como a negociação de preços de venda com consumidores e fornecedores ou a definição de prazos contratuais, não são, por ora, modeladas como variáveis de decisão ou fontes de incerteza. Essas decisões permanecem exógenas ao modelo e são tratadas como parâmetros fixos no horizonte de análise, mas podem ser incorporadas em versões futuras do framework, à medida que se expandam os objetivos e a granularidade das análises pretendidas.

No contexto do setor de gás natural, as decisões de caráter estratégico estão associadas à contratação de capacidade de transporte de longo prazo nos pontos de entrada e saída da malha de gasodutos, definindo a base estrutural sobre a qual se apoiam as decisões operacionais de consumo e balanceamento. Estas decisões são tomadas antes da execução dos cenários e avaliadas por meio de varreduras parametrizadas: define-se um conjunto de valores possíveis para uma variável de decisão (por exemplo, contratar 1000 m³/dia, 2000 m³/dia, etc.) e simula-se um conjunto completo de cenários para cada valor, mantendo as mesmas seeds de geração aleatória. Essa abordagem permite comparar diretamente os resultados médios e a dispersão de desempenho entre alternativas de contratação, sob as mesmas condições de incerteza.

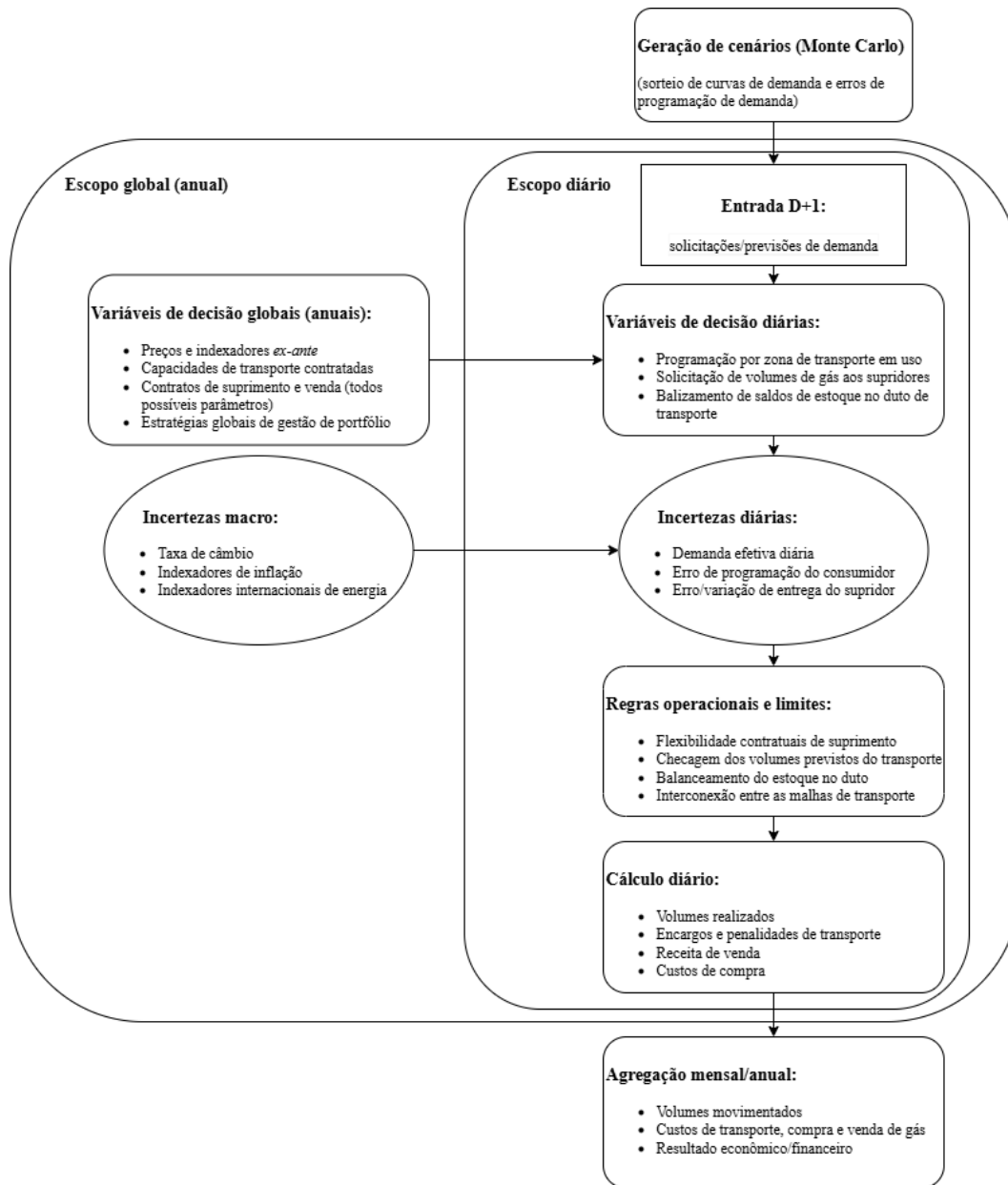


Figura 4.3: Fluxograma do processo de simulação: escopo global e diário com variáveis de decisão e incertezas

Já as decisões operacionais diárias, como o volume a ser solicitado de cada fornecedor, a alocação de gás por zona de transporte e a gestão de estoques de volumes nos dutos, como indicado na Figura 4.3, não são tratadas por varredura, mas sim por uma estratégia operacional pré-definida embutida no modelo para ser analisada. Essa estratégia pode ser uma lógica de decisão heurística, implementada em forma de fluxograma com base em regras como “manter igualdade entre consumo e suprimento diário”, “balancear o estoque diariamente” ou “não programar volumes de gás que causem penalidades sobre volumes excedentes no transporte”. Alternativamente, o modelo também está estruturado para permitir a substituição dessa lógica por métodos mais avan-

çados, como algoritmos baseados em redes neurais, que tomem decisões diárias com base nos estados observados do sistema naquele momento (capacidades de transporte contratadas, volumes solicitados pelos consumidores, estoques, parâmetros contratuais de suprimento, etc.).

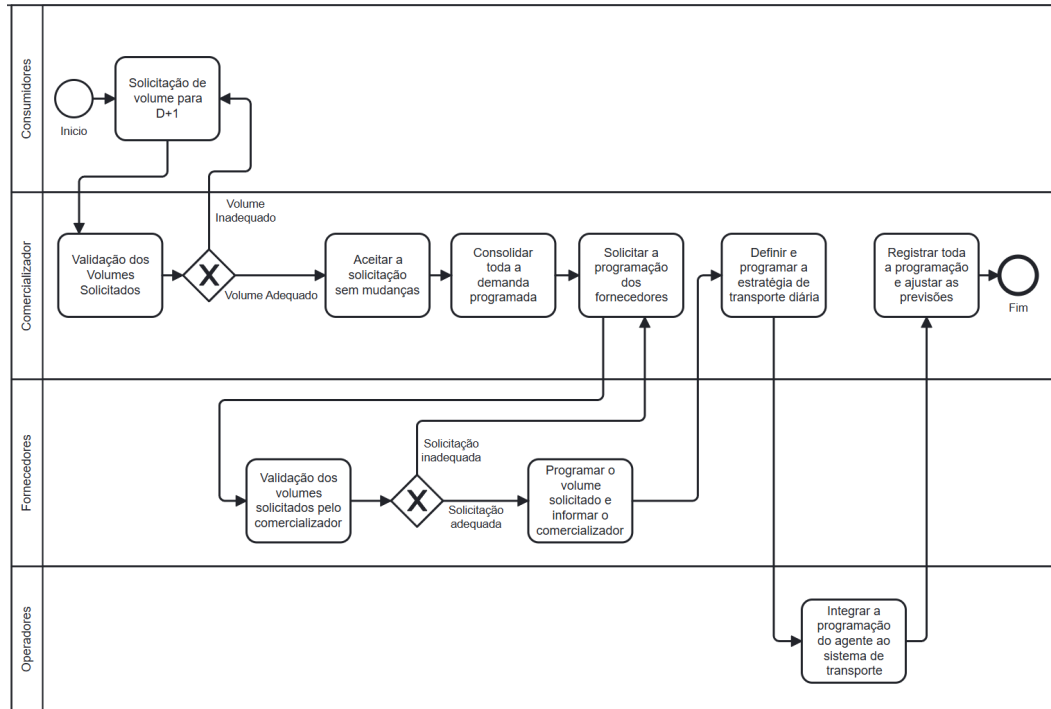


Figura 4.4: Fluxograma diário do processo de programação de gás entre os agentes do sistema

A Figura 4.4 ilustra o processo diário de programação de gás entre os agentes participantes do mercado, com base em uma notação de processos similar ao BPMN (Business Process Model and Notation). O fluxo inicia-se com a solicitação de volumes por parte dos consumidores para o dia seguinte (D+1), passando pela validação desses volumes pelo comercializador, consolidação das demandas, definição da estratégia logística e envio da programação para fornecedores e operadores de transporte. Ao longo do processo, há pontos de decisão nos quais os agentes avaliam a adequação dos volumes solicitados frente aos limites contratuais estabelecidos. As interações modeladas buscam capturar a lógica sequencial e as dependências operacionais envolvidas na execução dessa atividade central da cadeia de comercialização de gás natural.

Apesar da representação ordenada, é importante destacar que o processo real envolve maior complexidade e interações entre as partes que não são integralmente refletidas no modelo. Na simulação proposta, os critérios de aceitação ou rejeição de volumes programados por fornecedores consideram apenas a existência de excedentes contratuais, ou seja, solicitações além do teto acordado são recusadas. Por outro lado, no caso dos operadores de transporte,

adota-se a suposição de que todos os volumes programados são aceitos, ainda que penalidades financeiras sejam aplicadas quando se ultrapassam ou subdimensionam os limites operacionais definidos. Além disso, assume-se que a atuação do comercializador no sistema de transporte não interfere de forma significativa no despacho físico dos gasodutos, sendo tratado como um agente de impacto marginal. Em outras palavras, considera-se que sempre haverá capacidade disponível para acomodar as movimentações do comercializador, o que permite simplificar o modelo de fluxo nos gasodutos sem comprometer a representatividade dos resultados financeiros simulados.

Analisando o processo da simulação de Monte Carlo utilizada no framework, cada cenário completo gerado (31 dias) alimenta diretamente o modelo financeiro do framework, que calcula as métricas de desempenho da estratégia correspondente, com destaque para o VPL como a métrica de comparação final no caso de estudo. Ao final do processo de simulação, o conjunto de resultados permite avaliar não apenas o valor esperado das estratégias, mas também sua variabilidade, robustez e sensibilidade às variáveis contratuais e estratégias operacionais.

Essa integração entre simulação estocástica e avaliação financeira é o que permite ao comercializador simular múltiplas realidades possíveis e, com isso, tomar decisões com um certo nível de previsibilidade, maior embasamento quantitativo e controle sobre os riscos inerentes à sua operação.

4.4

Modelagem Matemática

A presente seção formaliza, por meio de uma estrutura matemática, os componentes econômicos do modelo proposto, estabelecendo os mecanismos de cálculo utilizados para avaliar o desempenho financeiro de cada estratégia analisada. A modelagem integra os fluxos de receita, custos operacionais, e volumes movimentados em um conjunto de equações que traduzem os eventos simulados na etapa estocástica em métricas quantificáveis.

Tabela 4.2: Índices utilizados na modelagem matemática

Símbolo	Descrição
t	Índice de tempo, utilizado nas equações com granularidade mensal.
t_0	Mês inicial do horizonte de simulação.
$t_{\max}$	Mês final do horizonte de simulação.
d	Dia específico dentro da simulação, usado em cálculos de operação diária.
x	Índice de consumidor ou supridor individual dentro do portfólio modelado.
S	Subgrupo do índice x , se referindo aos supridores.
C	Subgrupo do índice x , se referindo aos consumidores.
z	Zona de transporte, agrupamento de pontos na malha de dutos utilizados como entrada ou saída de gás.
o	Operador de transporte responsável pela malha ou trecho associado à zona z .

Para garantir a clareza na formulação das expressões matemáticas que compõem o modelo financeiro, adota-se a convenção de índices e notações descrita na Tabela 4.2.

Tabela 4.3: Valores considerados como parâmetros na modelagem e simulação

Símbolo	Descrição
$wacc$	Custo médio ponderado de capital da empresa (<i>Weighted average cost of capital</i>).
$IRPJ$	Taxa de Imposto de Renda sobre Pessoa Jurídica.
$CSLL$	Contribuição Social sobre o Lucro Líquido.
SGA_t	Despesas mensais de vendas, gerais e administrativas da empresa (<i>Selling, General and Administrative expenses</i>).
PM_x	Preço unitário de molécula, pode ser associado da mesma forma a uma compra ou venda de gás natural, a depender do papel definido para cada parte (de fornecedor ou de consumidor do gás).
QDC_x	Quantidade contratada de volume em metros cúbicos de gás natural por dia nos contratos de compra e venda de gás natural.
$\%TOP_x$	Percentual de obrigatoriedade mínima de consumo de gás natural em um contrato de compra e venda, em relação à QDC_x , para a aplicação do pagamento do PM_t , baseado no conceito de <i>Take-or-Pay</i> em contratos de gás, aplicado em base mensal.
PT_x	Preço unitário de transporte, pode ser associado da mesma forma a uma compra ou venda, associa ao preço do gás uma parcela para remunerar o custo de transporte do mesmo, neste caso sendo movimentado dentro da malha de gasodutos de transporte.
$\%SOP_x$	Percentual de obrigatoriedade mínima sobre o volume de gás natural (em relação à QDC_x) sobre o qual é cobrado o valor do PT_x , baseado no conceito de <i>Ship-or-Pay</i> em contratos de gás, aplicado em base diária.
ε_x	Desvio padrão do erro aplicado sobre a demanda real de cada consumidor de gás e seus respectivos volumes planejados para os mesmos períodos, representa essencialmente um erro de previsão de demanda em uma base D+1 para o próprio agente consumidor.
SM_x	Fator máximo de suprimento aplicável ao contrato de gás natural. Indica, em relação à QDC_x , o máximo de volume que a parte supridora naquele contrato está disposta a atender de demanda solicitada em base diária.
$TARIFF_z$	Tarifa unitária de transporte aplicada em base diária para o cálculo de encargos e penalidades de transporte em uma determinada zona de transporte de gás, de acordo com os volumes movimentados e capacidades contratadas naquela zona.
$\%EALIM_o$	Porcentagem limite para a movimentação de volumes excedentes autorizados para cada operador de transporte. Aplicada sobre o volume diário contratado em todas as zonas de transporte operadas pelo operador.
$VARMULT_o$	Fator multiplicador da tarifa de transporte que se é aplicada em volumes excessivos de erro de programação de transporte por operador, aplicada em base diária pra cada zona de transporte.
$VARLIM_o$	Fator multiplicador sobre o volume programado em transporte, que define o máximo de erro permitido com o volume realmente movimentado em uma determinada zona de transporte sem a aplicação de penalidades extras.
$LIMDES_o$	Percentual do volume contratado total com determinado operador de transporte que define o limite possível de desequilíbrio acumulado no duto ao fim do mês sem causar a aplicação de penalidade de desequilíbrio.

A Tabela 4.3 sintetiza os principais parâmetros considerados no processo de modelagem e simulação do framework proposto. Esses parâmetros abrangem aspectos econômico-financeiros e elementos contratuais e logísticos. Todos esses parâmetros serão contextualizados em mais detalhes no próximo trecho da formulação matemática.

Tabela 4.4: Variáveis de tomada de decisão na modelagem e simulação

Símbolo	Descrição
$QDTC_z$	Capacidade diária de transporte a ser contratada em cada zona de transporte utilizada. Irá embasar a maior parte dos encargos e penalidades envolvidos no transporte e valerá para todo o período anual de operação.
$QDP_{x,d}$	Quantidade de gás que será programada diariamente para cada supridor de gás, de forma a atender as volumes de gás solicitados ao agente comercializador por todos os seus clientes.
$QDP_{z,d}$	Quantidade de gás programada diariamente por zona de transporte em comunicação entre o comercializador e o operador de referência.

Finalmente, listam-se as variáveis de decisão aplicadas na modelagem e simulação, de acordo com a Tabela 4.4. Vale acrescentar que, é possível que outras variáveis possam ser analisadas e consideradas como variáveis de tomada de decisão, principalmente aquelas que são destacadas atualmente como parâmetros contratuais de compra e venda de gás natural, uma vez que estes contratos são completamente negociáveis entre as partes e não estão sujeitos a tantos controles regulatórios.

$$VPL = \sum_{t=t_0}^{t_{max}} \frac{FC_t}{(1 + wacc)^t} \quad (4-2)$$

O Valor Presente Líquido (VPL) é o principal indicador financeiro utilizado para avaliar a atratividade de determinadas estratégias de comercialização e gestão de portfólio de gás natural ao longo do tempo. Ele incorpora todos os fluxos de caixa projetados (FC_t), descontados pela taxa de custo de capital ($wacc$) (BREALEY; MYERS; ALLEN, 2020), permitindo comparar estratégias em termos de retorno ajustado ao risco. No modelo, o VPL será a métrica principal para a avaliação e comparação da performance das estratégias, porém, nada impede que o modelo seja utilizado para otimizar métricas mais localizadas no escopo da operação, maximizando ou minimizando outros valores mais específicas.

$$FC_t = RL_t + DA_t - JUR_t - CAPEX_t - \Delta CG_t \quad (4-3)$$

O valor resultante da equação (4-3) do fluxo de caixa (FC_t) consolida em cada período mensal: resultado líquido (RL_t), depreciação e amortização (DA_t), efeitos de juros (JUR_t), investimentos ($CAPEX_t$) e as variações no capital de giro (ΔCG_t). Esse fluxo serve como base para a análise de viabilidade econômica no modelo de Monte Carlo e alimentará a equação de valor presente líquido (4-2).

$$RL_t = LAIR_t - IR_t \quad (4-4)$$

$$IR_t = maior(LAIR_t, 0) \times (IRPJ + CSLL) \quad (4-5)$$

O resultado líquido (RL_t) é obtido a partir do lucro antes dos impostos ($LAIR_t$), reduzido da carga tributária (IR_t). Essa equação (4-4) evidencia a sensibilidade do resultado financeiro à tributação sobre o lucro. A equação (4-5) define o valor dos impostos de renda e contribuição social (IRPJ e CSLL) pagos sobre o lucro, definidos como 25% e 9% respectivamente. O uso do operador "maior" garante que, em caso de prejuízo ($LAIR_t < 0$), não haja incidência de imposto, alinhando-se à legislação fiscal vigente.

$$LAIR_t = LAJIRDA_t - DA_t + RF_t \quad (4-6)$$

O lucro antes dos impostos ($LAIR_t$) considera o lucro antes dos juros, imposto de renda, depreciação e amortização ($LAJIRDA$) ajustado pela depreciação e amortização (DA_t) e pelo resultado financeiro (RF_t). Esse ajuste permite avaliar a rentabilidade operacional antes dos encargos fiscais e financeiros, isolando o efeito das decisões de estrutura de capital.

$$LAJIRDA_t = RV_t - CPV_t - CO_t - SGA_t \quad (4-7)$$

O $LAJIRDA_t$ representa o resultado operacional do agente e é composto pela receita de venda de gás (RV_t) subtraída dos custos dos produtos vendidos (CPV_t), operacionais (CO_t) e administrativos (SGA_t). Essa medida é amplamente utilizada no setor energético para avaliar a eficiência operacional.

$$CPV_t = CGN_t + CT_t \quad (4-8)$$

O custo dos produtos vendidos (CPV_t) agrega os custos com a molécula

de gás natural (CGN_t) e os custos de transporte incorridos (CT_t). Essa decomposição reflete a estrutura típica de um portfólio de comercialização na malha de gasodutos, com dissociação clara entre commodity e logística.

$$CGN_t = \sum_{x \in S} MOL_{x,t} + TOP_{x,t} + TRANSP_{x,t} + SOP_{x,t} \quad (4-9)$$

onde : $S = \text{supridores}$

$$RV_t = \sum_{x \in C} MOL_{x,t} + TOP_{x,t} + TRANSP_{x,t} + SOP_{x,t} \quad (4-10)$$

onde : $C = \text{consumidores}$

O custo total de aquisição da molécula de gás (4-9) é dado pela soma dos componentes contratuais com cada supridor: preço da molécula ($MOL_{x,t}$), take-or-pay ($TOP_{x,t}$), custo de transporte associado a ser repassado ao consumidor ($TRANSP_{x,t}$) e a penalidade ship-or-pay ($SOP_{x,t}$). A inclusão de todos esses elementos permite capturar os diferentes tipos de contratos existentes no mercado. A receita de venda do gás (4-10) é estruturada de forma análoga à composição dos custos, adotando-se a hipótese de simetria contratual entre as operações de compra e venda. Assim, as condições contratuais aplicadas aos clientes do comercializador seguem o mesmo formato dos contratos de suprimento, diferenciando-se apenas pela inversão dos papéis de fornecedor e consumidor e pela aplicação dos respectivos termos de penalidades, transporte e capacidade não utilizada.

$$MOL_{x,t} = PM_x \times \sum_{d \in t} QDR_{x,d} \quad (4-11)$$

A parcela principal na composição da receita de venda do gás $MOL_{x,t}$ é o custo direto da molécula de gás adquirida de um fornecedor ou consumidor x , calculado como o produto entre o preço contratado por unidade energética (PM_x) e o volume efetivamente entregue ao longo do mês nesta mesma unidade. A estrutura da equação (4-11) captura variações mensais de parte da receita em função do consumo. A quantidade diária realizada ($QDR_{x,d}$), também presente na equação (4-11) define o volume de gás devidamente consumido diariamente pelo consumidor do supridor, essa variável de incerteza de demanda será utilizada em vários calculos de receitas e custos a partir deste ponto.

$$\begin{aligned}
TOP_{x,t} &= PM_x \times maior(TOPvol_{x,t} - \sum_{d \in t} QDR_{x,d}, 0) \\
\text{onde : } TOPvol_{x,t} &= \%TOP_x \times \sum_{d \in t} QDC_x
\end{aligned} \tag{4-12}$$

Embora a receita principal da venda esteja definida em (4-11), a natureza dos contratos firmes de suprimento de gás convém uma obrigatoriedade de consumo mínimo pela parte do consumidor, com qualquer diferença abaixo do mínimo de consumo sendo paga ao fim do mês para cumprir esta obrigatoriedade. Esse tipo de cobrança de um volume mínimo contratual é coloquialmente denominado "Take-or-Pay", exigindo pagamento mínimo ao fornecedor, mesmo sem consumo correspondente ($TOP_{x,t}$). A expressão (4-12) calcula a diferença entre o volume mínimo obrigatório ($TOPvol_{x,t}$) e o volume efetivamente consumido no período mensal de vigência da cláusula. O volume mínimo obrigatório por sua vez, é calculado através de uma porcentagem pré-determinada ($\%TOP_x$) sendo aplicada sobre o volume definido contratualmente entre as partes, a quantidade diária contratada (QDC_x). É importante distinguir os termos $TOP_{x,t}$ e $TOPvol_{x,t}$ presentes na equação (4-12), onde o ultimo destaca o volume de gás mínimo a ser pago por mensalmente, mesmo que não consumido, e o primeiro indica o montante financeiro desta diferença caso o efetivo consumido do mês seja menor que o volume mínimo.

$$TRANSP_{x,t} = PT_x \times \sum_{d \in t} QDR_{x,d} \tag{4-13}$$

Além dos valores referentes ao volume principal e à obrigatoriedade de consumo mínimo definidos em (4-11) e (4-12), também é necessário fazer o repasse dos custos de transporte do supridor particular ao volume do contrato e a jusante do ponto de entrega para o consumidor ($TRANSP_{x,t}$). Esse repasse é feito através de um preço unitário de transporte calculado, definido e aplicado pelo supridor (PT_x). Esse preço unitário de referência deve ser aplicado mensalmente sobre o volume consumido em (4-13), de forma análoga ao que ocorre sobre a receita sobre a molécula definida em (4-11).

$$SOP_{x,t} = PT_x \times \sum_{d \in t} maior(QDC_x \times \%SOP_x - QDR_{x,d}, 0) \tag{4-14}$$

De forma similar ao custo de "Take-or-Pay" definido em (4-12), custo de "Ship-or-Pay" cobra o consumidor pela capacidade de transporte subutilizada no contrato, dessa forma o supridor conseguirá cobrir os custos de transporte no

duto que tem obrigações de capacidade diárias. Em (4-14) calcula-se o volume ocioso a cada dia (se houver) e aplica o preço unitário de transporte contratado (PT_x) ao longo do mês, refletindo a ineficiência na alocação contratual entre agentes independentes no papel de consumidor e supridor.

$$QDR_{x,d} = maior(QDP_{x,d} \times (1 + \epsilon_x), 0) \quad (4-15)$$

onde : $QDP_{x,d} = menor(QDS_{x,d}, QDC_{x,d} \times SM_x)$

Esse volume é uma resultante do volume de fato programado entre as partes para o dia, adicionado de uma parcela multiplicativa de erro $(1 + \epsilon_x)$. Sobre o valor resultante é feita uma imposição que elimina a possibilidade que o erro simulado cause o valor realizado a se tornar negativo em casos raríssimos onde a variável de erro ϵ_x gere um valor randômico menor que menos um. O volume que o consumidor tem a liberdade para programar ao longo da sua operação por sua vez ($QDP_{x,d}$) também está limitado à um parâmetro contratual multiplicativo sobre a quantidade contratada, definido como o suprimento máximo (SM_x). Essas relações entre os volumes movimentados de gás refletem os limites operacionais e contratuais impostos pela malha de transporte ou pelo fornecedor para a oferta de volumes maiores dentro do mesmo contrato e a interação com o erro de programação. Ambos os volumes programados e os fatores de erro são variáveis pseudo-aleatórias que adicionam incerteza sobre o modelo estocástico.

$$CT_t = \sum_o (EST_{o,t} + ESEA_{o,t} + ESENA_{o,t} + ECNU_{o,t} + PVAR_{o,t} + PDES_{o,t}) \quad (4-16)$$

Dentro do portfólio de comercialização de gás do agente em estudo, todo supridor e consumidor está associado a um ponto específico dentro da malha de gasodutos onde este entrega ou recebe o gás, denominado como zona de transporte (z). Cada zona de transporte por sua vez esta associada a um operador específico (o) que é responsável por gerenciar e operar grandes concessões da malha de gasodutos de transporte interconectadas entre si (NTS, 2025; TBG, 2025). Os custos de transporte (CT_t) descritos em (4-16) englobam os encargos e penalidades associados à movimentação do gás na malha de gasodutos interligada, com a contratação sendo baseada na contratação de capacidade nas zonas de entrada e/ou saída. Esses custos refletem não apenas o valor contratado de capacidade (base tarifária), mas também os impactos financeiros por desvios operacionais do comercializador, como excesso de consumo ou subutilização. Esses componentes estão descritos nos contratos

de transporte como o MSA (Master Service Agreement) para cada operador de transporte, e a metodologia considera tanto eventos operacionais cotidianos quanto desvios sujeitos a penalização regulatória.

$$EST_{o,t} = \sum_{z \in o} \sum_{d \in t} QDEST_{z,d} \times TARIFF_z \quad (4-17)$$

$$QDEST_{z,d} = menor(QDTC_{z,d}, \sum_{x \in z} QDR_{x,d}) \quad (4-18)$$

O Encargo Base de Transporte ($EST_{o,t}$) representa o valor associado à movimentação dentro da faixa contratada de capacidade em uma zona de entrada ou saída. Ele é calculado com base na tarifa vigente ($TARIFF_z$) definida para a zona e a quantidade efetivamente utilizada, limitada pela capacidade contratada. Este componente corresponde à parte ordinária do custo de transporte, aplicada independentemente de penalizações. A quantidade diária de transporte contratada em cada zona ($QDTC_{z,d}$) é definida como uma variável de decisão no modelo, explorando a necessidade ótima de contratação de transporte tendo em vista toda a demanda de movimentação de gás empilhada associada a uma zona específica ($\sum_{x \in z} QDR_{x,d}$).

$$ESEA_{o,t} = \sum_{z \in o} \sum_{d \in t} QEA_{z,d} \times TARIFF_z \quad (4-19)$$

$$QEA_{z,d} = menor(EALIM_z, \sum_{x \in z} (QDR_{x,d}) - QDEST_{z,d}) \quad (4-20)$$

$$EALIM_z = QDTC_z \times \%EALIM_o \quad (4-21)$$

O Excedente Autorizado ($ESEA_{o,t}$) ocorre quando o transportador permite, sob certas condições operacionais, que volumes superiores à capacidade contratada sejam injetados ou retirados na malha. O contrato dos operadores prevê limites máximos ($EALIM_z$), proporcionalmente definidos por percentual de tolerância ($\%EALIM_o$) sobre a capacidade diária contratada ($QDTC_{z,d}$), cujo uso autorizado gera cobrança com a mesma base tarifária por zona. Essa faixa garante uma flexibilidade para volumes superiores sobre os contratos firmes de transporte.

$$ESEN_{o,t} = \sum_{z \in o} \sum_{d \in t} (QENA_{z,d} \times TARIFF_z \times 2) \quad (4-22)$$

$$QEN A_{z,t} = \sum_{x \in z} (QDR_{x,d}) - QEA_{z,d} - QDEST_{z,d} \quad (4-23)$$

Quando os volumes transportados excedem os limites contratados e não há autorização expressa para tal (i.e., não se qualificam como *ESEA*), o agente é penalizado com o Encargo de Excedente Não Autorizado (*ESENA_{o,t}*). O valor cobrado é o dobro da tarifa regular, reforçando o caráter punitivo da operação fora dos limites contratuais, conforme práticas estabelecidas pelas operadoras e autorizadas pela ANP.

$$ECNU_{o,t} = \sum_{z \in o} \sum_{d \in t} (QDCNU_{z,d} \times TARIFF_z) \quad (4-24)$$

$$QDCNU_{z,d} = \text{maior}(QDTC_{z,d} - \sum_{x \in z} QDR_{x,d}, 0) \quad (4-25)$$

O Encargo de Capacidade Não Utilizada (*ECNU_{o,t}*) reflete uma penalização indireta pela ineficiência na alocação contratual. Quando a capacidade reservada não é plenamente utilizada no dia, o valor é cobrado sobre o volume ocioso (*QDCNU_{z,d}*), incentivando o uso racional da infraestrutura e respeitando os princípios de otimização de rede previstos na legislação do Novo Mercado de Gás.

$$PVAR_{o,t} = \sum_{z \in o} \sum_{d \in t} (QVAR_{z,d} \times VARMULT_o \times TARIFF_z) \quad (4-26)$$

$$QVAR_{z,d} = \text{maior}(0, QDP_{z,d} \times VARLIM_o - |\sum_{x \in z} (QDR_{x,d}) - QDP_{z,d}|) \quad (4-27)$$

A Penalidade por Variação de Programação (*PVAR_{o,t}*) penaliza o desvio entre a programação declarada (*QDP_{z,d}*) e o volume efetivamente movimentado (*QDR_{x,d}*). Essa penalidade busca alinhar o planejamento com a execução operacional, fundamental em uma rede integrada onde alterações inesperadas afetam o balanceamento sistêmico. O contrato das operadoras impõe esse encargo com multiplicador (*VARMULT_o*) aplicado à tarifa da zona.

$$PDES_{o,t} = DES_{o,d} \times TARIFFMED_{o,d} \times 2 \quad (4-28)$$

onde : $d = \text{ultimo dia do mês } t$

$$DES_{o,d} = \text{maior}(0, |SDP_{o,d}| - QDTCMED_{o,d} \times LIMDES_o) \quad (4-29)$$

$$SDP_{o,d} = SDP_{o,d-1} + \left(\sum_{z \in (o \cap e)} QDR_{z,d} \right) - \left(\sum_{z \in (o \cap s)} QDR_{z,d} \right) \quad (4-30)$$

A penalidade de desequilíbrio ($PDES_{o,t}$) visa coibir o acúmulo de saldos residuais de gás ao longo do mês, os quais, se excederem o limite permitido ($LIMDES_o$), geram cobrança sobre o valor excedente no último dia útil do mês. Tal controle é essencial para garantir a previsibilidade e integridade do sistema de transporte, assegurando neutralidade operacional entre os diversos agentes.

Com a descrição dos elementos quantitativos que regem o modelo, conclui-se a construção da base que permite, a partir da simulação de Monte Carlo, a quantificação objetiva da performance de diferentes estratégias operacionais e contratuais. A modelagem matemática assegura coerência entre os resultados simulados e a realidade operacional e econômica de um comercializador de gás natural nesse ambiente de mercado, possibilitando a comparação entre alternativas de estratégias e de tomada de decisão. Essa estrutura será empregada, na próxima etapa, para a geração e análise dos resultados, permitindo a identificação das melhores estratégias através de métricas operacionais e econômicas diante das incertezas inerentes a operação.

4.5

Dados e arquivos de apoio

Todo o código-fonte desenvolvido para este trabalho foi implementado em Python (versão 3.11.9), utilizando bibliotecas como NumPy, Pandas, SciPy e Matplotlib para o processamento, análise estatística e visualização dos resultados. Os scripts, dados de entrada, resultados de simulação e documentação de uso estão disponíveis em repositório público acessível em:

LEAF - TESE - NR

5

Análise e Resultados

Este capítulo apresenta a aplicação do framework a um estudo de caso real, com o objetivo de avaliar sua capacidade de apoio a tomada de decisão em contextos reais de comercialização de gás natural.

Inicialmente, descreve-se o caso de estudo adotado, suas características operacionais, contratuais e o conjunto de dados utilizados. Em seguida, são detalhadas as etapas de implementação do framework, aplicadas na íntegra ao caso, com o intuito de validar sua funcionalidade prática e coerência metodológica. A seção seguinte explora a sensibilidade do modelo em relação a variáveis-chave, examinando o impacto de diferentes configurações sobre os resultados financeiros simulados. Por fim, discute-se criticamente o comportamento e os resultados observados nas simulações.

5.1

Caso de Estudo

Para validar a aplicabilidade do framework proposto e demonstrar sua capacidade de representar situações operacionais reais de comercialização de gás natural, esta seção apresenta um estudo de caso baseado em uma empresa comercializadora atuante no mercado brasileiro. De acordo com o que se é indicado na Tabela 5.1, o caso estudado compreende este agente comercializador operando um portfólio de 10 clientes e 2 fornecedores reais com contratos de suprimento de gás firmes ao longo de um mês de operação, compreendendo 31 dias. Parte destes clientes são indústrias que representam os consumidores finais do gás e parte são outros comercializadores que revenderão o gás para terceiros em sua parte.

Tabela 5.1: Parâmetros contratuais dos consumidores no caso de estudo

Consumidor	QDC (m ³ /dia)	ToP (%)	SoP (%)	Zona de Transporte	Flex. Sup. (%)	PM (R\$/m ³)	PT (R\$/m ³)
Comercializador A	1 360 000	80%	80%	TBG-DUTO	100%	1,6607	0,2209
Comercializador B	800 000	80%	100%	TBG-DUTO	100%	1,7857	0,2209
Comercializador C	60 000	90%	70%	TBG-DUTO	100%	1,7857	0,2209
Comercializador D	210 000	85%	100%	TBG-DUTO	105%	1,7972	0,2209
Indústria 1	50 000	85%	100%	NTS-MG1	110%	1,9587	0,4794
Indústria 2	224 000	90%	100%	TBG-SP2	105%	1,7857	0,3193
Indústria 3	66 000	80%	100%	TBG-SC2	110%	1,7857	0,3554
Indústria 4	52 000	90%	100%	NTS-SP3	105%	1,8003	0,3777
Indústria 5	97 000	90%	100%	TBG-SP2	105%	1,7857	0,3193
Indústria 6	34 000	95%	100%	TBG-SP2	105%	1,7857	0,3193

Os dados apresentados foram anonimizados para preservar a identidade comercial das empresas analisadas, sem prejuízo à representatividade estatística e operacional do portfólio. Cabe destacar que, nos casos em que o ponto de entrega ao consumidor coincide com a malha de transporte (isto é, quando o gás é recebido diretamente no duto como transferência de custódia), não há necessidade de contratação adicional de capacidade de transporte em uma zona de saída, com essa responsabilidade ficando com a parte compradora. Essa característica implica em um custo de transporte efetivamente inferior para esses consumidores, refletido na estrutura tarifária repassada a eles no modelo.

De forma análoga à estrutura adotada para os consumidores, os contratos de suprimento firmados com os fornecedores também seguem a mesma estrutura, contemplando os mesmos parâmetros. No entanto, diferentemente da relação de consumo, os papéis estão invertidos: neste caso, é o comercializador simulado quem solicita volumes diariamente, assumindo obrigações contratuais relacionadas ao consumo, enquanto os fornecedores desempenham o papel de agentes entregadores com limites operacionais e contratuais próprios.

Tabela 5.2: Parâmetros contratuais dos fornecedores no caso de estudo

Supridor	QDC (m ³ /dia)	ToP (%)	SoP (%)	Zona de Transporte	Flex. Sup. (%)	PM (R\$/m ³)	PT (R\$/m ³)
Supridor 1	2 800 000	70%	-	TBG-CORUMBA	100%	1,3474	-
Supridor 2	53 616	100%	-	NTS-DUTO	100%	1,6087	-

A Tabela 5.2 resume os principais parâmetros contratuais considerados para os dois fornecedores modelados no estudo de caso. No caso do Supridor 1, não há cobrança explícita de tarifa de transporte (PT), pois a responsabilidade pela contratação da capacidade na zona de entrada do gás recai integralmente sobre o comercializador simulado. Assim, o custo associado ao transporte desde o ponto de origem até a malha de entrega é assumido diretamente pelo agente comprador. Já o Supridor 2 opera sob uma lógica distinta: embora a entrega ocorra igualmente na malha de transporte, o próprio fornecedor é responsável pela alocação da capacidade de entrada, absorvendo esse encargo operacional. Consequentemente, o modelo não distingue explicitamente entre o preço da molécula (PM) e o custo de transporte (PT) nesse caso, pois entende-se que o valor de venda já incorpora, de forma agregada, ambos os componentes.

Com as definições dos contratos de fornecimento e de suprimento estabelecidas, seguimos para a caracterização das operadoras de transporte e das respectivas zonas em que se dividem os pontos de entrada e saída utilizados pelo comercializador. Essa estrutura é fundamental para o cálculo dos custos

logísticos e para se avaliar diretamente a estratégia de alocação e programação dos volumes ao longo da malha.

Tabela 5.3: Parâmetros operacionais das operadoras de transporte

Operadora	LIMDES (% da capacidade)	%EALIM (%)	VARLIM (%)	VARMULT (multiplicador)
NTS	5%	50%	50%	0,5
TBG	5%	50%	5%	2,0

As operadoras de transporte são os agentes responsáveis pela gestão física e contratual das malhas de gasodutos em que se realizam as movimentações de entrada e saída de gás natural. Embora existam outras operadoras de malhas de gasodutos de transporte no Brasil, o caso de estudo se limita às duas operadoras destacadas na Tabela 5.3, uma vez que não há nenhuma operação presente no caso que vá além das mesmas. Estas duas no caso são a Nova Transportadora do Sudeste (NTS) e a Transportadora Brasileira Gasoduto Bolívia-Brasil (TBG).

Seguindo com a Tabela 5.3, descrevem-se os parâmetros em consonância com a formulação matemática previamente estabelecida em razão da complexidade dos mesmos. O parâmetro do limite de desequilíbrio (LIMDES) representa o limite máximo de desbalanceamento do carregador antes que seja aplicada uma penalidade, expresso como percentual da capacidade total média de entrada e saída contratada na malha de acordo com o que foi descrito nas equações 4-28 e 4-29; valores inferiores a esse limite são considerados como operação equilibrada e não acarretam penalidades. O parâmetro percentual de excedente autorizado limite (%EALIM) define a margem percentual dentro da qual excedentes (volumes acima da capacidade contratada) ainda são aceitos operacionalmente ainda sem estarem sujeitos à tarifação adicional referenciada na equação 4-22. Já a variação limite (VARLIM) corresponde ao percentual limite aplicado sobre a programação diária antes da aplicação de uma penalidade por erro, que, por sua vez, é definida como a tarifa base de uma zona de transporte determinada multiplicada pelo último parâmetro atribuído ao operador (VARMULT), esse comportamento a esses dois parâmetros se demonstra nas equações 4-26 e 4-27.

Dando sequência, segue-se com a parametrização das zonas de transporte utilizadas no caso de estudo. No modelo proposto, essas zonas são utilizadas como base para o cálculo das tarifas aplicadas ao comercializador e para a verificação da conformidade das programações em relação aos contratos de capacidade.

Tabela 5.4: Zonas de transporte utilizadas na simulação

Zona	Operadora	Tipo de Zona	Tarifa de Transporte (R\$/m ³)
NTS-DUTO	NTS	E / S	–
NTS-MG	NTS	S	0,1332
NTS-SP	NTS	S	0,1252
NTS-REPLAN	NTS	INT	0,0221
TBG-DUTO	TBG	E / S	–
TBG-CORUMBA	TBG	E	0,2209
TBG-GASCAR	TBG	INT	0,0084
TBG-SP2	TBG	S	0,0984
TBG-SC2	TBG	S	0,1345

O parâmetro Tipo de zona classifica cada região como zona de entrada (E), saída (S) ou interconexão (INT), sendo algumas zonas classificadas como E / S por permitirem ambos os fluxos. A tarifa associada à zona representa o valor unitário aplicado à movimentação de gás naquela localidade, quando aplicável. A Tabela 5.4 resume as zonas utilizadas na simulação do caso base, com suas respectivas operadoras, classificações e tarifas unitárias.

Duas observações importantes devem ser feitas com relação à modelagem adotada. Primeiramente, as zonas classificadas como interconexão (INT), como NTS-REPLAN e TBG-GASCAR, representam pontos de interface entre redes de gasodutos de transporte geridas por operadores diferentes. Nesses casos, assume-se que o fluxo entre as redes ocorre de forma virtual entre os saldos para os carregadores, sendo a movimentação física do gás de fato resolvida entre as operadoras das malhas interconectadas. Isso acaba resultando numa programação perfeita para o usuário, ainda que este tenha que realizar as contratações de capacidade nas duas zonas que compõe a interconexão para que este possa operar entre as malhas. Em segundo lugar, as zonas denominadas DUTO (como NTS-DUTO e TBG-DUTO) não requerem contratação formal de capacidade nem geram encargos tarifários. Elas funcionam no modelo, como estruturas auxiliares para representar fluxos de gás entre diferentes agentes dentro de uma malha quando não há movimentação física do gás para uma zona de entrada ou saída específica, facilitando a parametrização e o controle lógico da programação do portfólio.

Além dos parâmetros específicos associados a consumidores, supridores, operadoras e zonas de transporte, o framework exige a definição de um conjunto de parâmetros globais que influenciam diretamente o comportamento da simulação e os resultados financeiros gerados. Esses parâmetros são utilizados em toda a estrutura do modelo, desde a geração dos cenários estocásticos até o cálculo final do Valor Presente Líquido (VPL). A Tabela 5.5 apresenta os principais valores definidos para a simulação do caso base, incluindo aspectos financeiros, fiscais e computacionais.

Tabela 5.5: Parâmetros globais da simulação

Parâmetro	Valor
Período da simulação	31 dias
Taxa de desconto (WACC)	15% a. a.
IRPJ	25%
CSLL	9%
Seed aleatória inicial	42
Fator de erro de programação	3%
Iterações de cenário (n)	100

Para esta primeira aplicação do framework, optou-se por restringir o escopo da modelagem econômico-financeira a elementos diretamente vinculados à gestão e operação do portfólio de gás natural. Dessa forma, certos componentes financeiros menos determinantes neste momento para a análise, foram deliberadamente excluídos da simulação. Entre os itens não considerados estão os juros (JUR), despesas com depreciação e amortização (DA), investimentos em ativos fixos (CAPEX), variações no capital de giro, custos operacionais fixos e despesas gerais e administrativas (SG&A). A exclusão desses elementos visa manter a transparência do modelo e concentrar a análise nos efeitos das decisões logísticas e contratuais, sem comprometer a coerência interna do cálculo do Valor Presente Líquido (VPL) nesse escopo de análise.

Sobre todo o caso de estudo se estabelece que as variáveis de decisão exploradas nessa instancia serão as quantidades diárias de transporte contratadas (QDTC) para cada zona em uso e as quantidades diárias programadas (QDP) para cada uma destas zonas ao longo da simulação, estas quais serão definidas automaticamente durante a simulação através de um algoritmo que replique uma estratégia básica de gestão operacional, que visa atender as demandas dentro dos limites de flexibilidade contratuais e balancear os saldos de portfólio na medida do possível.

5.2

Implementação do Framework

A aplicação do framework ao caso base inicia-se com o tratamento das séries históricas de consumo, que servirão de base para a caracterização estatística das incertezas de demanda. Os dados utilizados consistem em registros reais, com granularidade diária, coletados ao longo de 172 dias corridos desde o início do ano de 2025.

As medições estão expressas em metros cúbicos por dia (m^3/dia), mantendo unidade compatível com os contratos modelados e com os cálculos de volume, custo e receita realizados nas etapas posteriores do framework. Como destacado na Seção 5.1, os volumes totais analisados nestas series temporais

foram todos multiplicados por um mesmo fator somente conhecido pelo autor, para ajudar a manter o anonimato comercial destes agentes.

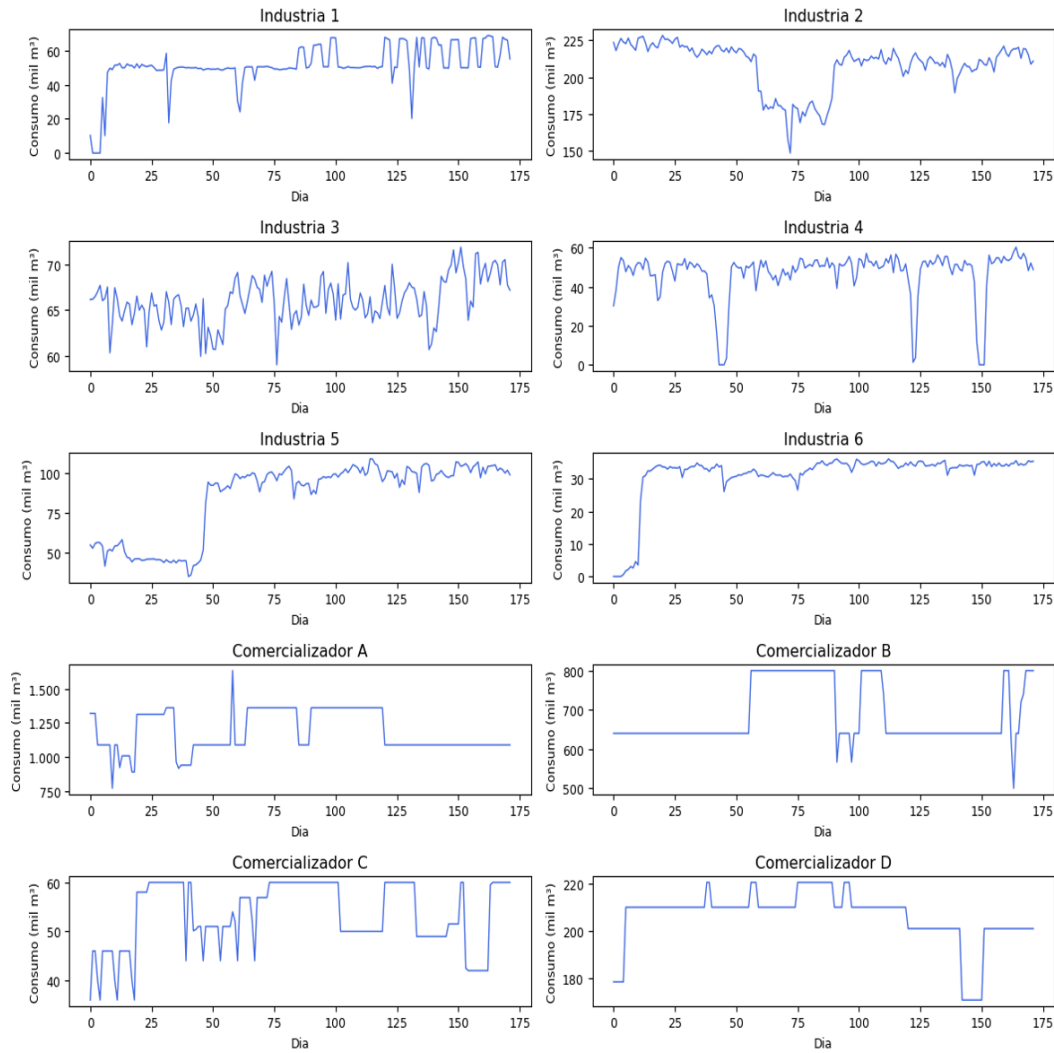


Figura 5.1: Series temporais utilizadas como base para os consumidores no caso base

Observando a Figura 5.1 é possível observar a clara diferença entre as curvas de consumo dos consumidores industriais, que tem uma curva de demanda muito mais orgânica e natural, e as curvas dos consumidores comercializadores, estes quais tem a demanda regida por parâmetros dentro de seus respectivos portfólios e operam programando em relação ao o que esta definido em seus parâmetros contratuais.

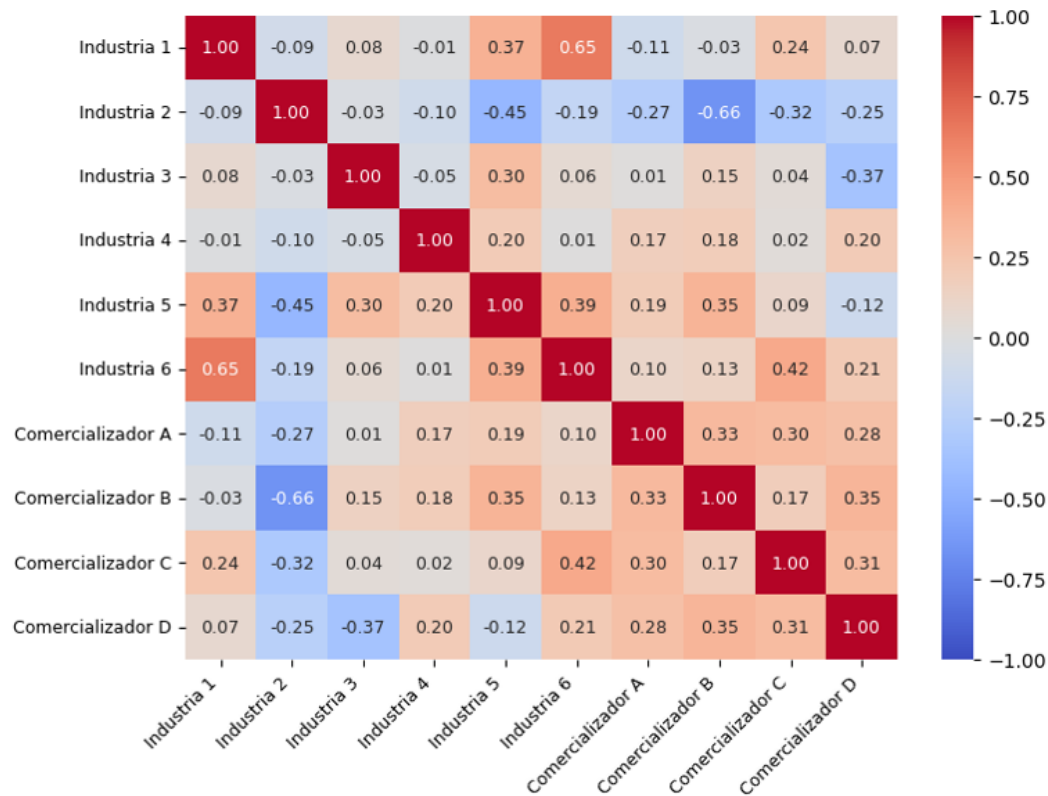


Figura 5.2: Índices de correlação entre os consumidores no caso base

Seguindo a lógica do framework, nenhum dos valores de correlação entre as séries (Figura 5.2) apresenta um valor absoluto superior acima de 0,7, com as relações mais altas ficando com um valor de 0,66, entre o Comercializador B e a Industria 2. Como nenhuma correlação ultrapassa o valor limítrofe definido na estrutura do framework para o uso de Modelagem conjunta das curvas de consumo, segue-se com a aplicação com todas as demandas sendo assumidas independentes entre si.

Após a análise de correlação entre os consumidores, o framework prossegue com a avaliação da estrutura temporal de cada série de consumo individual. Para isso, são utilizados dois instrumentos complementares: a função de autocorrelação (FAC), função de autocorrelação parcial (FACP) e o teste de Dickey-Fuller aumentado (ADF) (DICKY; FULLER, 1979). A FAC e a FACP fornecem uma visão gráfica da persistência temporal em diferentes defasagens, enquanto o ADF oferece uma estatística formal para testar a presença de raiz unitária, condição associada à não-estacionariedade.

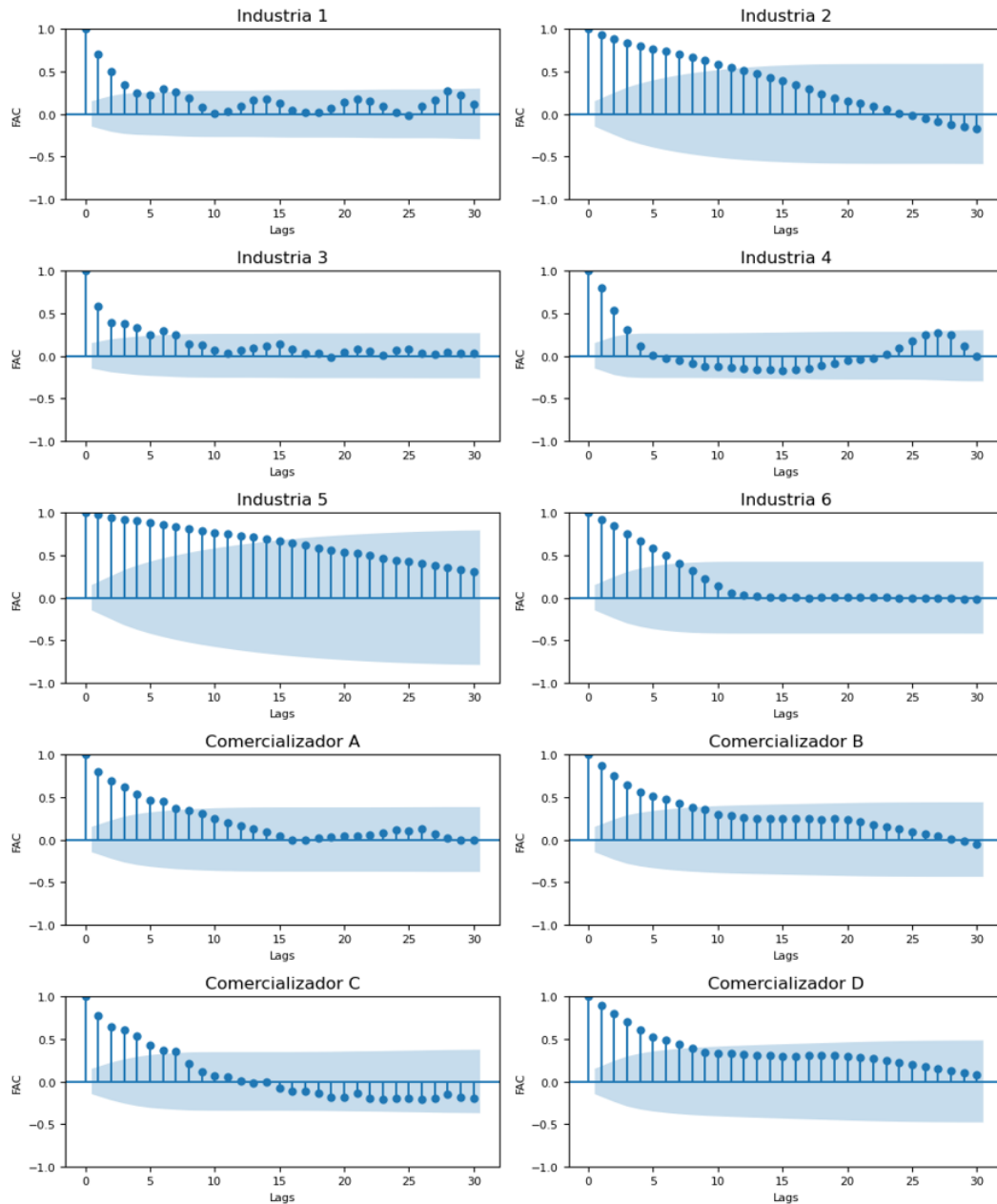


Figura 5.3: Função de autocorrelação dos consumidores no caso base

A Figura 5.3 apresenta os gráficos da FAC para as dez séries de consumo consideradas. Observa-se que, para a maior parte das séries, a autocorrelação decai rapidamente e oscila dentro dos limites de significância já a partir dos primeiros lags, ainda que algumas séries, notadamente as das Indústrias 2 e 5, apresentem um decaimento mais lento da autocorrelação.

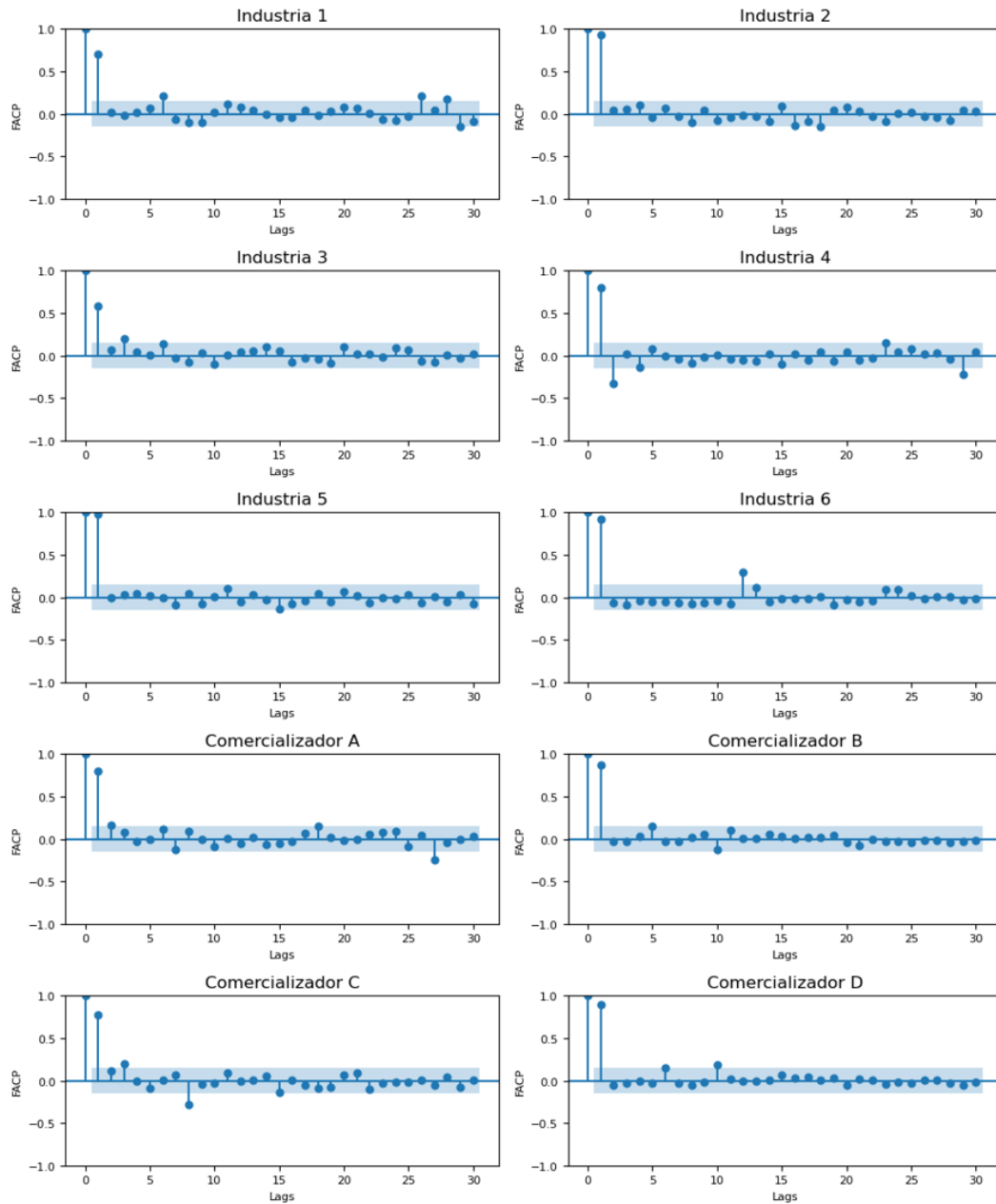


Figura 5.4: Função de autocorrelação parcial dos consumidores no caso base

A análise da Função de Autocorrelação Parcial (FACP) das séries temporais de consumo, ilustrada na Figura 5.4, evidencia que a maioria das séries, tanto das indústrias quanto dos comercializadores, apresenta um pico significativo no primeiro defasamento, seguido por valores próximos de zero nas demais defasagens. Esse comportamento indica que a dependência temporal direta ocorre predominantemente entre observações consecutivas, sugerindo a adequação de modelos autorregressivos de primeira ordem (AR(1)) para a maior parte das séries. Observa-se ainda que nenhuma das séries apresenta padrões oscilatórios ou múltiplos picos significativos, o que reforça a ausência de estruturas de autocorrelação parcial de ordem superior. Assim, as séries anali-

sadas demonstram uma dinâmica essencialmente de curta memória, compatível com processos estacionários simples e com a hipótese de que a influência das observações passadas se dissipa rapidamente.

Tabela 5.6: Resultados do teste ADF (p-valores) para as séries de consumo

Série	p-valor
Indústria 1	$2,9 \times 10^{-9}$
Indústria 2	0,1391
Indústria 3	0,0014
Indústria 4	$6,3 \times 10^{-7}$
Indústria 5	0,5057
Indústria 6	0,0004
Comercializador A	0,0086
Comercializador B	0,0216
Comercializador C	0,0078
Comercializador D	0,0147

Os resultados numéricos dos testes ADF, sintetizados na Tabela 5.6, corroboram a análise para a maioria das séries: oito delas apresentam p-valores inferiores a 0,05, permitindo rejeitar a hipótese nula de não-estacionariedade.

Em sequência buscam-se os parâmetros mais adequados nos modelos de previsão ARMA a serem treinados para cada série de demanda Industrial no caso base. A escolha dos parâmetros dos modelos foi realizada com base no critério de informação de Akaike (AIC) (AKAIKE, 1974), amplamente empregado na literatura para balancear a qualidade do ajuste e a complexidade do modelo. O AIC é definido como $AIC = 2k - 2\ln(\hat{L})$, em que k representa o número de parâmetros estimados e $\hat{L}$ é o valor máximo da função de verossimilhança do modelo. Assim, quanto menor o valor de AIC , melhor é o compromisso entre parcimônia e ajuste estatístico. Dessa forma, para cada série temporal de demanda, procedeu-se à estimação de uma grade de modelos $ARMA(p, q)$, variando-se os parâmetros de ordem p (número de termos autor-regressivos) e q (número de termos de médias móveis). O parâmetro p captura a dependência linear entre o valor atual e observações passadas da série, enquanto q representa a influência de choques ou ruídos anteriores sobre o valor presente. O modelo que apresentou o menor valor de AIC foi selecionado como o mais adequado para representar a estrutura temporal da respectiva demanda, garantindo que o ajuste mantivesse a consistência estatística sem incorrer em sobreparametrização.

Para limitar a complexidade computacional e evitar sobreajuste, as ordens autorregressiva (p) e de médias móveis (q) foram restringidas ao intervalo de $0 \leq p, q \leq 5$. Essa delimitação segue recomendações práticas da literatura de séries temporais (BOX et al., 2015), segundo as quais ordens

elevadas tendem a aumentar o risco de multicolinearidade entre termos e a reduzir a interpretabilidade dos coeficientes, sem ganhos significativos em termos de ajuste estatístico.

Tabela 5.7: Resultados da calibração dos modelos ARMA para as séries de demanda e comercializadores.

Série	Ordem ARMA (p, q)	AIC (full)	Status do Ajuste
Indústria 1	(5, 5)	3 433	OK
Indústria 2	(1, 5)	3 347	OK
Indústria 3	(4, 5)	3 005	OK
Indústria 4	(0, 5)	3 428	OK
Indústria 5	(2, 5)	3 298	OK
Indústria 6	(4, 5)	3 041	OK

Nota: Todos os modelos foram selecionados com base no menor valor do critério de informação de Akaike (AIC), considerando combinações das ordens autorregressiva (p) e de médias móveis (q) com diferenciação nula. O status “OK” indica convergência do processo de estimação e validade estatística dos parâmetros estimados.

Observa-se que, em todos os casos na Tabela 5.7, os modelos atingiram convergência e apresentaram valores de AIC consistentes com a estrutura temporal das respectivas séries, indicando um bom equilíbrio entre ajuste e parcimônia. A partir desses modelos, foram obtidos os resíduos, definidos como a diferença entre os valores observados e os valores ajustados pelo modelo, os quais representam a componente estocástica não explicada pela dinâmica autorregressiva e de médias móveis. O próximo passo metodológico consistiu em verificar se tais resíduos podem ser caracterizados como ruído branco estacionário e, em seguida, ajustar distribuições de probabilidade que melhor descrevam seu comportamento aleatório. Essa etapa é essencial para a posterior geração de cenários estocásticos e para a quantificação das incertezas associadas à previsão de demanda.

Após o ajuste dos modelos ARMA para cada série de demanda industrial, os resíduos resultantes foram analisados com o objetivo de verificar se apresentavam comportamento estacionário e ausência de autocorrelação, para que seja possível caracterizá-los como ruído branco ou não. Para isso, aplicaram-se os testes de raiz unitária de Dickey–Fuller aumentado (ADF) (DICKEY; FULLER, 1979) e o teste de estacionariedade de Kwiatkowski–Phillips–Schmidt–Shin (KPSS) (KWIATKOWSKI et al., 1992), ambos amplamente utilizados na literatura de séries temporais. O teste ADF possui

hipótese nula de não estacionariedade, de modo que valores de p inferiores a 0,05 indicam rejeição dessa hipótese e, portanto, estacionariedade da série. Em contraste, o teste KPSS adota hipótese nula de estacionariedade, sendo que valores de p inferiores a 0,05 sugerem não estacionariedade.

Tabela 5.8: Resultados dos testes de estacionariedade aplicados aos resíduos dos modelos ARMA.

Indústria	ADF p -valor	Estacionário (ADF)	KPSS p -valor	Estacionário (KPSS)
Indústria 1	0,0137	Verdadeiro	0,0100	Falso
Indústria 2	0,2343	Falso	0,0131	Falso
Indústria 3	0,0007	Verdadeiro	0,0100	Falso
Indústria 4	$1,77 \times 10^{-11}$	Verdadeiro	0,1000	Verdadeiro
Indústria 5	0,1849	Falso	0,0100	Falso
Indústria 6	0,0789	Falso	0,0100	Falso

A Tabela 5.8 apresenta os resultados obtidos para cada série de resíduos, incluindo as conclusões associadas a cada teste. Embora nem todos os resíduos dos modelos ARMA tenham apresentado resultados consistentes de estacionariedade nos testes de ADF e de KPSS, tal fato não inviabiliza a análise probabilística subsequente. Em séries empíricas de demanda, pequenas violações das condições estritas de estacionariedade são comuns, especialmente devido à presença de heterocedasticidade, eventos sazonais e choques exógenos de curta duração. Assim, ainda que os testes indiquem resultados mistos, considera-se que os resíduos estimados preservam adequadamente o comportamento estocástico remanescente após a remoção das componentes determinísticas. Nesse contexto, o ajuste de distribuições de probabilidade, não necessariamente centradas na média zero nem restritas à forma normal, permite caracterizar de maneira mais flexível a natureza aleatória dos resíduos, capturando assimetrias, curtoses e caudas pesadas que frequentemente ocorrem em dados de demanda industrial e comercial. Tal abordagem fornece uma representação estatística mais realista da incerteza, fundamental para a etapa de simulação estocástica e análise de risco associada às previsões de consumo.

O ajuste das distribuições de probabilidade é realizado com base no método da Máxima Verossimilhança (Maximum Likelihood Estimation – MLE) (CASELLA; BERGER, 2002), procedimento estatístico amplamente utilizado para inferência paramétrica. Esse método estima os parâmetros das distribuições candidatas a partir dos dados observados, maximizando a função de verossimilhança dos mesmos sob a hipótese de determinada distribuição. Essa abordagem é compatível com os principais tipos de distribuições considerados no framework e permite obter estimativas consistentes e eficientes para séries com variabilidade moderada.

Tabela 5.9: Distribuições de probabilidade ajustadas aos resíduos das séries de demandas industriais.

Indústria	Distribuição Ajustada	Parâmetros Estimados	p -valor (KS)
Indústria 1	t-Student	$(\nu = 2,19; \mu = 1\,630,95; \sigma = 5\,325,93)$	0,57
Indústria 2	t-Student	$(\nu = 3,15; \mu = 949,89; \sigma = 4\,236,49)$	0,36
Indústria 3	Laplace	$(\mu = 399,76; \sigma = 1\,693,02)$	0,74
Indústria 4	t-Student	$(\nu = 2,26; \mu = 1\,207,22; \sigma = 4\,383,78)$	0,77
Indústria 5	Cauchy	$(x_0 = 589,55; \gamma = 2\,469,88)$	0,33
Indústria 6	t-Student	$(\nu = 1,33; \mu = 671,19; \sigma = 839,81)$	0,30

Os resultados apresentados na Tabela 5.9 indicam que todas as distribuições ajustadas apresentaram p -valores do teste de Kolmogorov–Smirnov (KS) superiores a 0,05, o que significa que não há evidências estatísticas para rejeitar a hipótese de aderência entre os resíduos e as distribuições teóricas estimadas. Dessa forma, considera-se que os ajustes realizados foram adequados em todos os casos, representando de maneira satisfatória o comportamento estocástico das séries residuais.

Após o ajuste das distribuições de probabilidade para os resíduos das séries industriais, segue-se à modelagem das demandas dos clientes comercializadores no caso base. Nessa etapa, foram utilizadas tabelas de distribuição de probabilidades associadas a diferentes faixas de volume de consumo, permitindo representar de forma probabilística a variabilidade das demandas observadas em um destes consumidores. Essa abordagem possibilitou incorporar a incerteza inerente ao comportamento estacionário de consumo dos comercializadores.

Tabela 5.10: Tabelas de distribuição de probabilidades utilizadas na modelagem das demandas dos comercializadores.

Comercializador	Probabilidade (%)	Volume (m ³ /dia)
Comercializador A	54	1 088 000
	35	1 360 000
	8	1 312 000
	3	940 000
Comercializador B	69	640 000
	31	800 000
Comercializador C	44	60 000
	11	50 000
	8	48 960
	8	46 000
	8	51 000
	6	56 900
	6	52 000
	3	58 000
	3	44 000
	3	51 520
Comercializador D	58	210 000
	15	220 500
	14	201 000
	13	200 980

No caso dos comercializadores, descritos na Tabela 5.10, o padrão de consumo é tabelado e discreto, refletindo comportamentos contratuais rígidos ou programações comutadas entre valores fixos. Nessas situações, optou-se por representar os dados com distribuições discretas empíricas, com probabilidades associadas às observações recorrentes, capturando até 90% das probabilidades acumuladas para cada distribuição e depois normalizando estas tabelas de probabilidade para compreender um total de 100% dos cenários.

Com os parâmetros estatísticos devidamente ajustados para cada consumidor seja este industrial ou outro comercializador, o framework segue para a etapa de geração de cenários estocásticos via simulação de Monte Carlo. Nessa etapa, criam-se conjuntos sintéticos de demandas diárias ao longo do horizonte de planejamento. A abordagem adotada permite simular a variabilidade natural de consumo de forma controlada, respeitando as características empíricas

observadas em cada perfil de cliente.

A simulação de Monte Carlo é estruturada para refletir tanto as incertezas de consumo dos clientes quanto os desvios de programação relacionados à operação dos fornecedores e à execução logística do comercializador. A seguir, detalham-se os procedimentos utilizados para a geração dos cenários, o número de iterações adotado, as variações de variáveis decisórias exploradas e os componentes aleatórios adicionais incorporados à simulação, como erros de programação e flutuações operacionais.

5.3

Sensitividade do Modelo

A etapa seguinte do framework consiste na aplicação das distribuições de probabilidade ajustadas para a geração de cenários sintéticos de consumo ao longo do horizonte definido por meio de uma simulação de Monte Carlo. Para cada conjunto de variáveis decisórias testado, no caso, diferentes configurações de quantidades diárias de transporte contratadas (QDTCs), o modelo executa 100 iterações independentes, cada uma simulando um mês de 31 dias de operação.

A aleatoriedade dos cenários é controlada por um valor numérico chamado semente, fixada de forma incremental para cada iteração da simulação e que se re-estabelece ao se trocar qualquer uma das variáveis de decisão em avaliação. Com isso, assegura-se que, independentemente da variável de decisão em teste, os mesmos 100 cenários sejam utilizados para cada ponto de comparação, permitindo uma análise direta dos efeitos atribuíveis exclusivamente às variáveis sensibilizadas.

Os volumes diários de consumo dos clientes são sorteados com base nas distribuições ajustadas individualmente para cada consumidor. A esses volumes, somam-se os erros de programação associados à previsibilidade limitada de consumo e entrega, modelados como ruídos gaussiano centrado (distribuições normais) (JONDRAL, 2018). Essa modelagem permite representar de forma coerente a aleatoriedade tanto da demanda quanto da capacidade de resposta dos agentes envolvidos, sem perder o vínculo com as observações históricas.

Tabela 5.11: Variáveis de decisão sensibilizadas na simulação de Monte Carlo

Zona de Transporte	Faixa de QDTC m ³ /dia	Passo m ³ /dia	Tipo
NTS-MG	20 000 — 50 000	1 000	Saída
NTS-SP	20 000 — 52 000	1 000	Saída
TBG-SP2	200 000 — 355 000	5 000	Saída
TBG-SC2	30 000 — 66 000	1 000	Saída
TBG-CORUMBA	1 000 000 — 2 800 000	100 000	Entrada
TBG → NTS	20 000 — 60 000	1 000	Interconexão

Na simulação de Monte Carlo, as decisões estratégicas de contratação de capacidade de transporte (QDTC) são modeladas como variáveis de decisão sensíveis. A Tabela 5.11 apresenta as faixas de variação adotadas para cada zona de saída, ponto de entrada ou interconexão envolvida na operação do comercializador. Para cada combinação de valores dessas variáveis, o modelo executa 100 simulações de um mês, avaliando os impactos operacionais e financeiros decorrentes de diferentes níveis de contratação da infraestrutura de transporte oferecida pelo operador da rede.

Os valores definidos como padrão para as quantidades diárias de transporte contratadas (QDTCs) são estabelecidos como o limite superior do intervalo de sensibilização definido na Tabela 5.11, estes refletem as somas dos valores contratuais de cada consumidor e supridor, quando a realidade é que, quase sempre o volume contratual é colocado acima das demandas reais devido a uma mentalidade comercial de escassez.

Dado o volume elevado de simulações envolvidas na análise combinada de múltiplas variáveis decisórias, particularmente pela granularidade adotada em cada faixa, optou-se nesta primeira aplicação por explorar os intervalos de cada QDTC de forma independente. Ou seja, cada zona será inicialmente sensibilizada isoladamente no contexto do portfólio completo, mantendo-se as demais variáveis em seus valores de referência. Essa abordagem permite identificar, com menor custo computacional, os efeitos marginais de cada decisão sobre os resultados do modelo. Após a identificação dos pontos ótimos individuais, será possível aplicar uma análise complementar em que os valores ótimos de cada variável serão combinados e reavaliados conjuntamente, oferecendo uma visão integrada da configuração mais eficiente para o portfólio simulado.

É importante destacar que a sensibilização das variáveis de capacidade contratadas em cada zona devem ter impacto de custo majoritariamente localizado, isto é, a variação da quantidade de transporte alocada em uma zona específica não altera diretamente a estrutura global de custos de transporte do portfólio. O impacto financeiro permanece restrito aos encargos associados àquela zona em particular, sem repercussões sobre os custos ordinários das

demais. A única exceção a essa independência ocorre no caso das penalidades de desequilíbrio, cuja base de cálculo considera o saldo consolidado de capacidade contratada em toda a malha. Dessa forma, alterações isoladas em uma QDTC podem influenciar indiretamente o montante de penalidades incorridas, uma vez que estas derivam da interação sistêmica entre o portfólio completo e a rede de transporte.

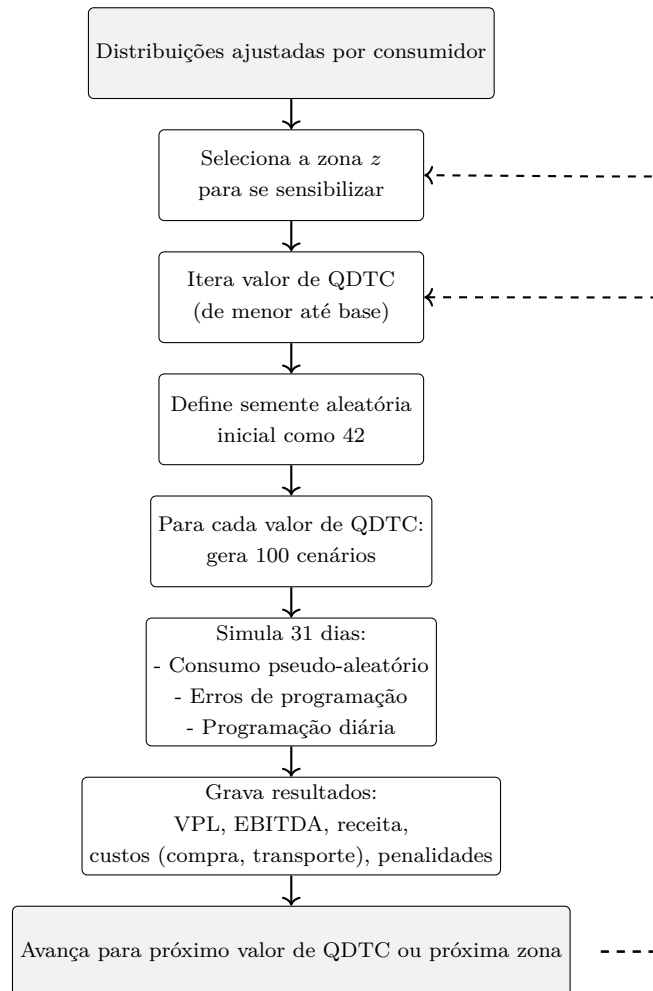


Figura 5.5: Fluxograma geral da simulação de Monte Carlo com sensibilização das QDTC

A lógica detalhada da simulação de Monte Carlo adotada neste trabalho está representada de forma esquemática na Figura 5.5. O diagrama ilustra o fluxo de dados e decisões que compõem a execução do modelo, incluindo a leitura das distribuições ajustadas, a sensibilização das variáveis de decisão (QDTCs), a geração de cenários com controle de aleatoriedade e a simulação diária do comportamento operacional do portfólio. Ao final de cada ciclo, os principais indicadores de desempenho econômico e operacional são registrados para posterior análise comparativa.

Concluída a etapa de simulação, a próxima seção apresenta e discute os principais resultados obtidos, explorando as relações entre as decisões estratégicas de contratação, os riscos operacionais simulados e os impactos sobre a rentabilidade e a eficiência do portfólio analisado.

5.4

Discussão dos Resultados

Nesta seção são analisados criticamente os resultados obtidos nas simulações, com ênfase no comportamento do Valor Presente Líquido (VPL) Unitário Médio em função das Quantidades Diárias de Transporte Contratadas (QDTC) para cada zona de transporte. A opção pelo uso da métrica do VPL Unitário não se deu pelo fácil acesso a valores comparativos de outros agentes comercializadores no mercado, mas sim porque essa métrica permite que estes resultados sejam equiparados com cenários onde o volume de gás natural sendo contratado é efetivamente comercializado.

O foco recai sobre a identificação dos pontos de contratação que maximizam o retorno econômico esperado, bem como sobre as implicações de variabilidade associadas a diferentes níveis de capacidade contratada.

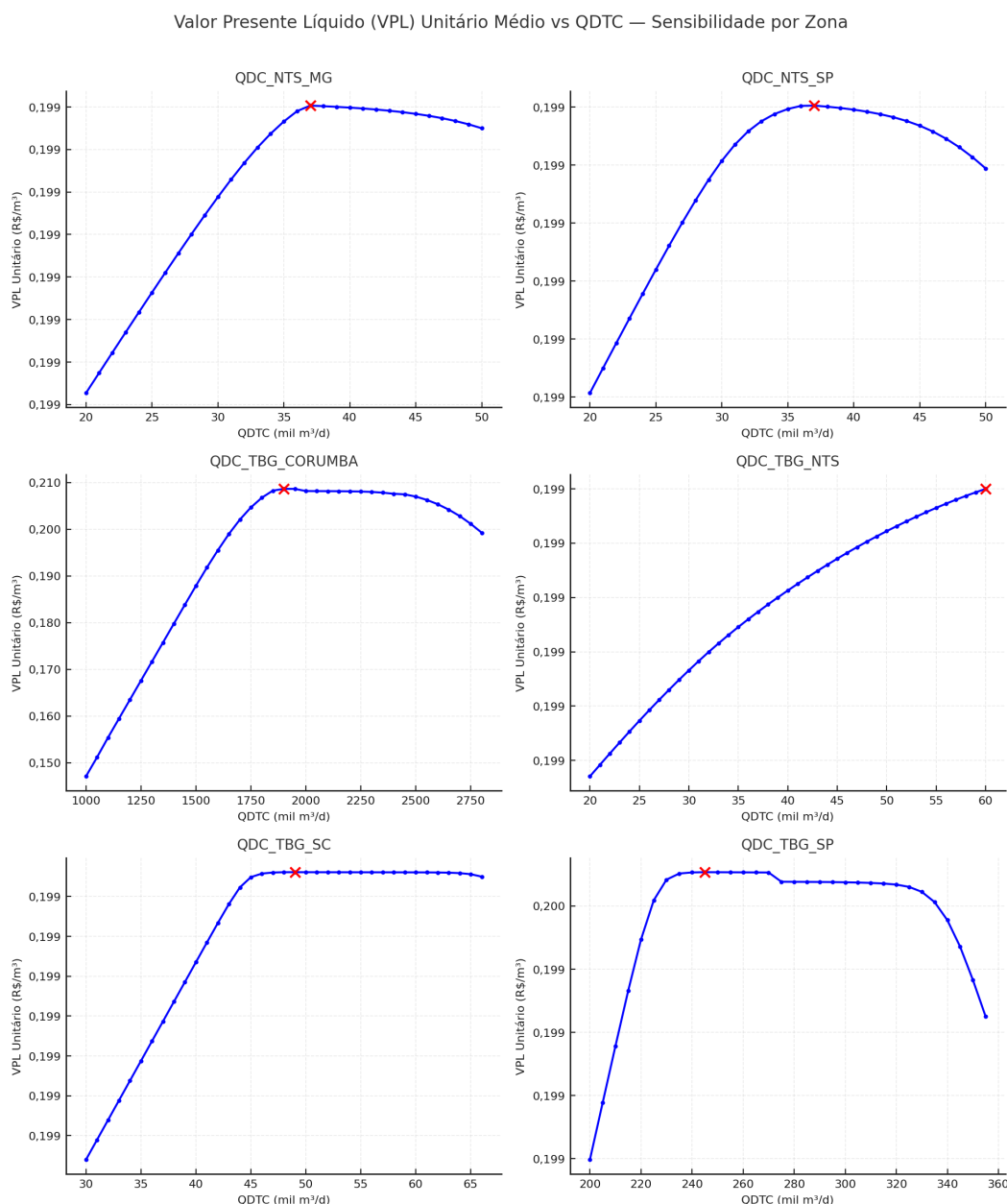


Figura 5.6: Valor Presente Líquido (VPL) Unitário Médio em função da QDTC para cada zona de transporte.

A Figura 5.6 sintetiza os resultados das seis zonas avaliadas, apresentando a trajetória do VPL Unitário Médio em função da QDTC (expressa em mil m³/d). Cada curva revela a existência de um ponto ótimo de contratação, marcado no gráfico, correspondente ao nível de QDTC que maximiza o VPL Unitário Médio na respectiva zona.

De maneira geral, observa-se que nas zonas NTS_MG e NTS_SP, os pontos ótimos situam-se em patamares médios de QDTC (em torno de 35 a 40 mil m³/d), evidenciando que a contratação excessiva tende a reduzir o retorno unitário devido ao custo associado à capacidade ociosa.

Observam-se melhorias de resultado em todas as zonas avaliadas, ainda assim a zona TBG_CORUMBA oferece o maior ganho possível na performance do portfólio entre o volume contratado padrão e o ótimo. Essa melhora de resultado faz sentido quando consideramos que TBG_CORUMBA é a zona que movimenta a maior parte do volume de gás do portfólio e, conseqüentemente, tem o maior impacto no resultado financeiro.

Observando a zona de interconexão TBG_NTS é claro que a contratação ótima se encontra além do intervalo de contratação possível, o que faz com que o maior valor do intervalo atual também seja o que traz os menores custos de transporte locais e maiores resultados financeiros para a operação do portfólio.

O comportamento identificado em todas as zonas de transporte é consistente com o trade-off esperado entre 0 subdimensionamento da QDTC, que aumenta a probabilidade de penalidades por excedentes não autorizados (ESENA) e reduz o retorno líquido; e o superdimensionamento da QDTC, que, embora reduza penalidades sobre volumes excedentes, implica custos fixos adicionais de capacidade não utilizada (ECNU), reduzindo o VPL Unitário.

Assim, o ponto ótimo de cada zona corresponde ao balanço entre estas duas forças. A diferença nos patamares ótimos reflete tanto a escala da zona quanto a variabilidade das demandas projetadas.

A análise dos pontos ótimos, contudo, não é suficiente por si só, uma vez que o Valor Presente Líquido Unitário em cada zona está sujeito a variabilidade decorrente das incertezas da demanda e da operação. Para aprofundar a compreensão desse aspecto, foram avaliadas as distribuições empíricas dos VPLs unitários resultantes das 100 iterações de Monte Carlo em cada cenário ótimo de QDTC por zona.

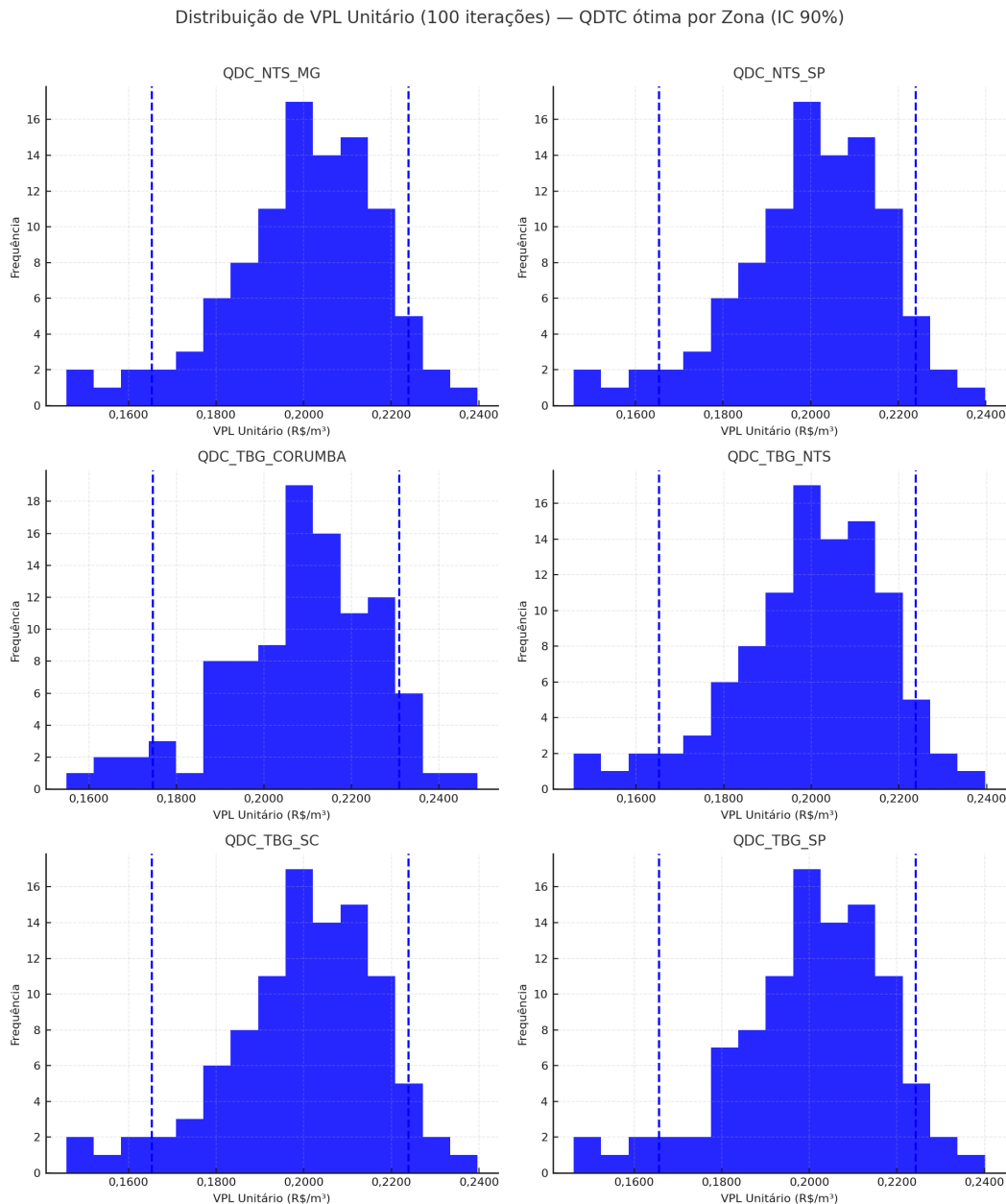


Figura 5.7: Distribuições empíricas do VPL Unitário nas QDTC ótimas de cada zona de transporte.

A Figura 5.7 apresenta os histogramas dessas distribuições, juntamente com os intervalos de confiança de 90% (percentis 5 e 95), permitindo identificar a dispersão dos resultados possíveis, para se avaliar a robustez ou fragilidade dos pontos de contratação considerados ótimos.

Em praticamente todas as zonas de saída, os intervalos de confiança observados são mais amplos do que os ganhos adicionais proporcionados pela escolha da QDTC ótima, o que evidencia que a variabilidade inerente ao portfólio como um todo pode se sobrepor ao benefício marginal da decisão ótima localizada para aquela determinada zona de saída. A única exceção é a zona

de TBG_CORUMBA, onde a amplitude do intervalo de confiança é proporcionalmente menor em relação ao ganho obtido, conferindo maior relevância estatística à decisão ótima. Essa diferença decorre do fato de o VPL representar um resultado consolidado do portfólio, de modo que zonas com menor representatividade exercem impactos mais limitados na métrica global. Em contrapartida, TBG_CORUMBA, por movimentar volumes significativamente maiores, apresenta influência mais significativa sobre o desempenho financeiro total, o que explica a maior distinção observada entre seu ponto ótimo e a faixa de incerteza associada.

Tabela 5.12: Resumo por zona: QDTC ótima, VPL unitário médio, intervalo de confiança, desvio-padrão e melhora percentual.

Zona de transporte	QDTC ótima (m ³ /d)	VPL unitário médio (R\$/m ³)	Intervalo de confiança (R\$/m ³)		Desvio- padrão (R\$/m ³)	Melhora (%)
			5% (inf.)	95% (sup.)		
NTS_MG	37 000	0,19930	0,16531	0,22392	0,017751	0,027%
NTS_SP	37 000	0,19940	0,16541	0,22404	0,017751	0,054%
TBG_CORUMBA	1 900 000	0,20863	0,17467	0,23102	0,017208	4,708%
TBG_NTS	60 000	0,19925	0,16528	0,22387	0,017751	–
TBG_SC2	49 000	0,19926	0,16529	0,22387	0,017751	0,006%
TBG_SP2	245 000	0,19971	0,16567	0,22430	0,017712	0,229%

A Tabela 5.12 evidencia que a otimização individual da contratação em cada zona resulta em ganhos financeiros marginais na maior parte dos casos, com exceção da zona TBG_CORUMBA, cuja representatividade no portfólio a torna mais crítica. Nesta zona, a diferença entre o cenário base e o ótimo implica em uma melhora média de 4,708% no VPL anual, valor significativamente superior ao observado nas demais zonas, onde os ganhos não ultrapassam frações percentuais. A partir deste ponto, a análise avança para a avaliação do cenário integrado, em que todas as QDTCs ótimas são adotadas simultaneamente, de modo a verificar o efeito conjunto sobre o desempenho financeiro consolidado do portfólio.

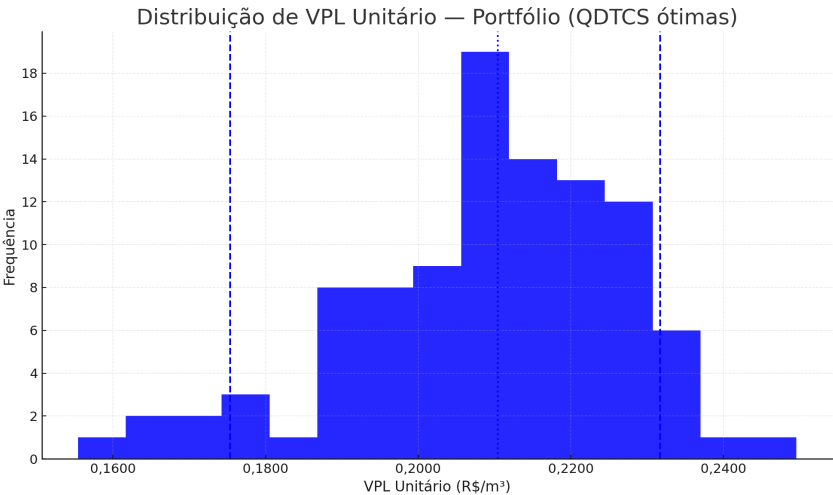


Figura 5.8: Distribuição do VPL Unitário do portfólio com todas as QDTCs ótimas.

A distribuição apresentada na Figura 5.8 revela que, ao adotar simultaneamente as QDTCs ótimas em todas as zonas de transporte, o portfólio apresenta um deslocamento positivo no valor esperado do VPL unitário, alcançando uma média de aproximadamente R\$0,209/m³. A consolidação de todas as decisões ótimas por zona de transporte reforça, que os efeitos isolados em zonas de menor representatividade se incrementam para gerar uma melhora anida superior na performance do portfólio.

Tabela 5.13: Resumo cenário ótimo conjunto: VPL unitário médio, desvio-padrão e melhora percentual.

Cenário onjunto	VPL unitário médio (R\$/m³)	Desvio-padrão (R\$/m³)	Melhora percentual (%)
QDTCs otimas	0,20928	0,017205	5,036%

Pela Tabela 5.13 pode-se perceber que a melhora da mudança simultânea para todas as QDTCs identificadas como ótimas é quase igual a soma de todos incrementos obtidos na análise por zona, com esta soma totalizando um valor de incremento no VPL unitário de 5,036%. Esse resultado faz sentido quando se lembra que o efeito das variáveis de decisão sobre as zonas de transporte afetam quase que exclusivamente seus centros de custo de forma isolada.

A melhoria do resultado é significativa quando consideramos que não houveram mudanças entre os parâmetros contratuais de compra e venda ou sobre os volumes de gás natural movimentados para os cenários simulados, com toda a melhoria do resultado tendo vindo somente do aspecto intermediário da operação considerado no custo do transporte do gás natural.

De qualquer forma, a aplicação do framework nesse caso de uso real conseguiu demonstrar de forma confiável o comportamento e resultado econômico

de cenários sensibilizados nas variáveis de decisão de contratação de transporte, as QDTCs. Dessa forma, o framework é bem sucedido na sua proposta inicial, independente da significância dos valores da melhora nos resultados, por já ter adereçado o ponto principal levantado na pergunta de pesquisa.

Cabe ressaltar que, até este ponto, a análise concentrou-se exclusivamente no comportamento do Valor Presente Líquido unitário como métrica principal de desempenho do portfólio. No entanto, o framework desenvolvido e a estrutura de simulação implementada registram simultaneamente diversas outras variáveis de interesse, tais como fluxos de caixa parciais, custos de transporte desagregados, penalidades por excedentes, além de indicadores de risco associados à variabilidade da demanda. Essas informações adicionais, embora não exploradas em profundidade nesta etapa, estão registradas, constituem um conjunto de dados que pode ser utilizado em análises futuras, seja para a subseção de resultados específicos, para a extensão deste estudo em direções complementares ou mesmo para o uso corporativo direto. Dessa forma, o modelo permanece viável para abordar questões de diferentes escopos na natureza da operação do portfólio de comercialização.

6

Conclusão

Este estudo teve como objetivo central propor, desenvolver e validar um framework quantitativo de apoio à decisão para portfólios de comercialização de gás natural, estruturado a partir da simulação de Monte Carlo. A motivação principal decorreu das transformações recentes no setor brasileiro, impulsionadas pelo Novo Mercado de Gás, que ampliaram a complexidade e a competitividade do ambiente, tornando insuficientes abordagens determinísticas tradicionais e evidenciando a necessidade de ferramentas capazes de lidar com incertezas inerentes ao processo de decisão cada vez mais complexo.

A formulação matemática do modelo também se mostrou adequada aos objetivos da pesquisa, pois conseguiu traduzir de maneira clara e consistente a lógica do problema decisório em restrições e funções de otimização compatíveis com a realidade do setor. Essa estruturação permitiu que o framework fosse ao mesmo tempo rigoroso, garantindo validade científica, e flexível, possibilitando a adaptação a diferentes configurações contratuais e contextos de mercado.

Os resultados obtidos ao longo do trabalho permitem responder de maneira positiva e fundamentada a essa questão. O framework desenvolvido demonstrou ser capaz de estruturar o processo decisório de forma coerente com a complexidade da operação e mercado, ao traduzir os elementos de incerteza oriundos das curvas de demanda no portfólio em distribuições de probabilidade aderentes que alimentam a simulação de Monte Carlo. Esse procedimento possibilitou a geração de cenários de consumo para medir indicadores de desempenho do portfólio contra a sensibilização de diversas possíveis variáveis de decisão, substituindo a visão simplificada de valores determinísticos por uma análise probabilística.

Além disso, a formulação matemática associada ao modelo mostrou-se essencial para viabilizar a operacionalização do framework, uma vez que permitiu formalizar o problema de decisão em termos de restrições contratuais, condições de transporte e alternativas de comercialização. Essa estruturação foi determinante para garantir que as simulações refletissem as opções reais disponíveis às empresas do setor, evitando abstrações excessivas ou desvios em relação à prática de mercado.

Assim, pode-se concluir que a integração entre formulação matemática, representação estocástica das variáveis críticas e simulação de Monte Carlo constitui uma resposta eficaz à pergunta de pesquisa. O framework proposto não apenas apoia a tomada de decisão em ambientes de elevada incerteza,

mas também amplia a qualidade da informação disponível para o agente comercializador, permitindo que escolhas estratégicas sejam embasadas em cenários probabilísticos, e não apenas em estimativas médias ou previsões pontuais.

No âmbito metodológico, o estudo apresenta avanços relevantes ao formalizar matematicamente um framework que alia a lógica da otimização contratual às propriedades da simulação estocástica. A formulação matemática utilizada mostrou-se adequada para representar a natureza operacional de um portfólio de comercialização de gás natural nesse mercado e o rigor científico do modelo. Essa contribuição metodológica é significativa porque demonstra que é possível traduzir problemas complexos de portfólio energético em estruturas de decisão robustas e adaptáveis, capazes de incorporar novas variáveis e cenários sem comprometer o modelo ou a análise estatística dos resultados.

Adicionalmente, o caráter replicável do framework constitui uma contribuição metodológica importante. O modelo pode ser aplicado a diferentes portfólios, ajustado para outras condições de consumo ou expandido para contemplar novos instrumentos de mercado. Essa flexibilidade abre caminho para sua utilização não apenas em pesquisas acadêmicas, mas também, e principalmente, em aplicações práticas neste mercado. Dessa forma, o trabalho oferece uma ponte entre a pesquisa e a prática corporativa, fortalecendo o diálogo entre teoria e aplicação.

Apesar das contribuições teóricas, metodológicas e práticas alcançadas, é importante reconhecer as limitações do estudo e as potenciais ameaças à validade dos resultados. A primeira delas está relacionada à base de dados utilizada para a calibração das distribuições probabilísticas. Embora os dados históricos tenham permitido representar de forma consistente a dinâmica dos preços e custos de transporte, a extensão temporal relativamente curta pode ter restringido a capacidade do modelo de capturar choques raros ou eventos extremos de baixa probabilidade, mas de alto impacto, comuns no setor energético.

A capacidade computacional extensa necessária para realizar a sensibilidade no caso de estudo também se demonstrou uma limitação, uma vez que restringe o dinamismo e flexibilidade na aplicação real do framework em base diária. Felizmente esse aspecto pode ser amenizado se trabalhando sobre a otimização do código em python e suas extensões ou na tradução para linguagens mais eficientes e rápidas.

A presente pesquisa abre espaço para diversas possibilidades de aprofundamento e expansão, tanto no campo metodológico quanto na aplicação prática. Em primeiro lugar, destaca-se a necessidade de aprimorar a base de

dados utilizada para calibração das séries de consumo modeladas. Estudos futuros poderiam incorporar séries temporais mais extensas, além de variáveis macroeconômicas, como câmbio e preços de commodities que afetam seriamente a performance de operações nestes ambientes. Como o código base e o framework como um todo existe agora já é possível aplicar essas variáveis, a questão mais crítica seria simplesmente como direcionar o estudo de caso e como modelar estas macro-séries temporais para a aplicação no modelo de simulação.

Outra linha promissora de investigação reside na integração do framework com modelos dinâmicos de formação de preços, permitindo capturar a evolução temporal das variáveis de mercado de maneira mais realista. A utilização de modelos de processos estocásticos não estacionários, bem como a inclusão de técnicas de aprendizado de máquina, principalmente quando se considera a avaliação de decisões diárias da operação do portfólio.

É importante destacar que se tem espaço para evolução futura do modelo avaliando modelos de previsão temporal mais avançados para simular as curvas de demanda dentro da simulação. Esses casos poderão ser tratados por meio de abordagens semelhantes às utilizadas no artigo *“Forecasting Natural Gas Consumption in a Cement Plant - A Case Study”*, publicado pelo autor pela revista científica Gestão e Produção. As séries de demanda modeladas com tais técnicas poderão ser posteriormente replicadas dentro da simulação de Monte Carlo do portfólio, ampliando a capacidade do framework em lidar com diferentes perfis estatísticos de consumo.

Além disso, futuros trabalhos poderiam expandir o escopo de aplicação do framework para portfólios híbridos de energia, integrando gás natural, com fontes geradoras como hidrogênio e biometano. Essa abordagem permitiria investigar sinergias e riscos cruzados entre diferentes mercados, aspecto cada vez mais relevante em um contexto de transição energética.

Em conjunto, essas propostas de pesquisas futuras reforçam o caráter dinâmico do tema abordado e evidenciam que o framework aqui desenvolvido representa não um ponto final, mas uma base flexível a partir da qual novos estudos podem avançar no entendimento e na gestão das incertezas no novo mercado livre de gás natural.

Referências bibliográficas

Agência Nacional do Petróleo, Gás Natural e Biocombustíveis (ANP). **Boletim Anual de Gás Natural 2020**. 2020. <https://www.gov.br/anp/pt-br/assuntos/dados-estatisticos/boletins-anp>. Acesso em: 18 ago. 2025.

AKAIKE, H. A new look at the statistical model identification. **IEEE Transactions on Automatic Control**, v. 19, n. 6, p. 716–723, 1974. Disponível em: <https://doi.org/10.1109/TAC.1974.1100705>.

BERRISCH, J.; ZIEL, F. Distributional modeling and forecasting of natural gas prices. **arXiv preprint**, 2020. Disponível em: <https://arxiv.org/abs/2010.06227>.

BOX, G. E. et al. **Time Series Analysis: Forecasting and Control**. [S.l.]: Wiley, 2015. ISBN 9781118675021.

BREALEY, R. A.; MYERS, S. C.; ALLEN, F. **Principles of Corporate Finance**. 13. ed. McGraw-Hill Education, 2020. ISBN 9781260013900. Disponível em: https://www.academia.edu/31292991/_Richard_A_Brealey_Stewart_C_Myers_Alan_J_Mar_BookFi_.

CARDOSO, J. V. M. et al. Os desafios do mercado de gás natural no brasil como potencial indutor da transição energética e da integração sul-americana. **Revista Tempo do Mundo**, v. 32, 2023. Disponível em: <https://www.ipea.gov.br/revistas/index.php/rtm/article/view/472>.

CASELLA, G.; BERGER, R. L. **Statistical Inference**. 2. ed. [S.l.]: Duxbury Press, 2002. ISBN 978-0534243128.

CHEN, P. et al. Natural gas as a key alternative energy source in sustainable renewable energy transition. **Frontiers in Energy Research**, 2021. Disponível em: <https://www.frontiersin.org/journals/energy-research/articles/10.3389/fenrg.2021.625023/full>.

Confederação Nacional da Indústria (CNI). **Gás Natural: uma avaliação da abertura do mercado brasileiro sob competência da União**. 2025. Relatório apresentado em seminário da CNI. Disponível em: <https://shorturl.at/TEOwy>.

DICKEY, D. A.; FULLER, W. A. Distribution of the estimators for autoregressive time series with a unit root. **Journal of the American Statistical Association**, Taylor & Francis, v. 74, n. 366, p. 427–431, 1979. Disponível em: <https://doi.org/10.1080/01621459.1979.10482531>.

DVORKIN, V. et al. Stochastic control and pricing for natural gas networks. **IEEE Transactions on Control of Network Systems**, v. 9, n. 1, p. 450–462, 2022. Disponível em: <https://doi.org/10.1109/TCNS.2021.3112764>.

Empresa de Pesquisa Energética. **Plano Indicativo de Gasodutos de Transporte 2022–2031**. Rio de Janeiro, 2022. Disponível em: <https://www.epe.gov.br/pt/publicacoes-dados-abertos/publicacoes/Plano-Indicativo-de-Gasodutos-de-Transporte>.

GOLUB, G. H.; LOAN, C. F. V. **Matrix Computations**. 4. ed. Johns Hopkins University Press, 2013. ISBN 978-1421407944. Disponível em: <https://jhupbooks.press.jhu.edu/title/matrix-computations>.

GUERREIRO, M. **Otimização de Portfólios de Comercialização de Energia no Brasil**. Tese (Doutorado) — FGV, 2017. Disponível em: https://repositorio.fgv.br/bitstream/handle/10438/18536/Tese_MGuerreiro_Versao%20final.pdf.

International Energy Agency. **World Energy Outlook 2022**. Paris: IEA, 2022. Disponível em: <https://www.iea.org/reports/world-energy-outlook-2022>.

JONDRAL, F. K. White gaussian noise – models for engineers. **Frequenz**, De Gruyter, v. 72, n. 5-6, p. 293–299, 2018. Disponível em: <https://doi.org/10.1515/freq-2017-0064>.

JONG, C. de. Pricing and optimization of natural gas storage in competitive natural gas markets. **Journal of Natural Gas Science and Engineering**, 2015. Disponível em: <https://www.kyos.com/optimization-natural-gas-storage/>.

KWIATKOWSKI, D. et al. 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**, v. 54, n. 1–3, p. 159–178, 1992. Disponível em: [https://doi.org/10.1016/0304-4076\(92\)90104-Y](https://doi.org/10.1016/0304-4076(92)90104-Y).

MALYSCHIEFF, A. M.; TRAFALIS, T. B. Natural gas storage valuation via least squares monte carlo and support vector machine regression. **Energy Systems**, v. 8, n. 4, p. 815–855, 2017. Disponível em: <https://link.springer.com/article/10.1007/s12667-017-0239-7>.

MAVROTAS, G. Combining mathematical programming and monte carlo simulation to deal with uncertainty in energy project portfolio selection. In: **Assessment and Simulation Tools for Sustainable Energy Systems**. Springer, 2013, (Green Energy and Technology, v. 129). p. 333–356. Disponível em: https://doi.org/10.1007/978-1-4471-5143-2_16.

McKinsey & Company. **LNG portfolio optimization: Putting the business model to the test**. 2021. Artigo da McKinsey & Company. Disponível em: <https://www.mckinsey.com/industries/oil-and-gas/our-insights/lng-portfolio-optimization-putting-the-business-model-to-the-test>.

MENEZES, M. C. P. P.; LAUREANO, F. H. G. C. Efficient power generating portfolio in brazil: Conciliating cost and risk under mean–variance theory. **Energy Economics**, 2014. Disponível em: <https://doi.org/10.1016/j.eneco.2013.07.012>.

MURTHA, J. A. Monte carlo simulation: Its status and future. **Journal of Petroleum Technology (JPT)**, 1997. Disponível em: <https://onepetro.org/JPT/article/49/04/361/108147/Monte-Carlo-Simulation-Its-Status-and-Future>.

NTS. **Contrato Master de Serviço de Transporte Firme de Gás Natural 2025 - NTS**. 2025. Documento institucional. Disponibilizado pela ANP. Acesso em: ago. 2025.

Oxford Institute for Energy Studies. **Liberalization: the key to unlocking natural gas potential in Brazil?** 2023. Disponível em: <https://www.oxfordenergy.org/publications/liberalization-the-key-to-unlocking-natural-gas-potential-in-brazil/>.

PEARSON, K. Notes on regression and inheritance in the case of two parents. **Proceedings of the Royal Society of London**, Royal Society, v. 58, p. 240–242, 1895. Disponível em: <https://doi.org/10.1098/rspl.1895.0041>.

RECKMAN, N. D. M. Forecasting natural gas consumption in a cement plant: A case study. 2025.

RIEPIN, I. Modelling uncertainty in coupled electricity and gas systems – is it worth the effort? **arXiv preprint**, 2020. Disponível em: <https://arxiv.org/abs/2008.07221>.

Secretaria-Geral da União. **Lei nº 14.134, de 8 de abril de 2021. Institui o Novo Mercado de Gás Natural**. 2021. https://www.planalto.gov.br/ccivil_03/_ato2019-2022/2021/lei/L14134.htm. Acesso em: 18 ago. 2025.

SHAPIRO, S. S.; WILK, M. B. An analysis of variance test for normality (complete samples). **Biometrika**, Oxford University Press, v. 52, n. 3/4, p. 591–611, 1965. Disponível em: <https://doi.org/10.1093/biomet/52.3-4.591>.

SMIRNOV, N. Table for estimating the goodness of fit of empirical distributions. **The Annals of Mathematical Statistics**, Institute of Mathematical Statistics, v. 19, n. 2, p. 279–281, 1948. Disponível em: <https://doi.org/10.1214/aoms/1177730256>.

S&P Global. The role of gas in the energy transition. 2024. Disponível em: <https://www.spglobal.com/en/research-insights/special-reports/look-forward/the-role-of-gas-in-the-energy-transition>.

TBG. **Contrato Master de Serviço de Transporte Firme de Gás Natural 2025 - TBG**. 2025. Documento institucional. Disponibilizado pela ANP. Acesso em: ago. 2025.

TOMASGARD, A. et al. Optimization models for the natural gas value chain. In: **Geometric Modelling, Numerical Simulation, and Optimization**. Springer, 2004. p. 521–558. Disponível em: https://link.springer.com/chapter/10.1007/978-3-540-68783-2_16.

WERON, R. **Modeling and Forecasting Electricity Loads and Prices: A Statistical Approach**. Wiley, 2006. Disponível em: <https://www.scirp.org/reference/referencespapers?referenceid=2210671>.

ZAVALA, V. M. **Stochastic Optimal Control Model for Natural Gas Network Operations**. [S.l.], 2013. Disponível em: <https://www.mcs.anl.gov/papers/P5023-1013.pdf>.

8

Apêndice

8.1

Artigo de complemento ao trabalho

Forecasting natural gas consumption in a cement plant – a case study

Previsão do consumo de gás natural em uma fábrica de cimento – estudo de caso

Nicholas De Marco Reckman¹ , Igor Tona Peres¹ 

¹ Pontifícia Universidade Católica do Rio de Janeiro – PUC-RIO, Departamento de Engenharia Industrial, Rio de Janeiro, RJ, Brasil. E-mail: reckman@aluno.puc-rio.br; igor.peres@puc-rio.br

How to cite: Reckman, N. M., & Peres, I. T. (2025). Forecasting natural gas consumption in a cement plant – a case study. *Gestão & Produção*, 32, e3825. <https://doi.org/10.1590/1806-9649-2025v32e3825>

Abstract: This study aims to investigate the forecasting of natural gas consumption in a cement plant in Rio de Janeiro, amidst the transition from the regulated to the open natural gas market, contextualized in the Brazil's "new gas law" policies. We use demand forecasting to potentially reduce gas acquisition, transportation, and storage costs in the transport pipeline grid. For this purpose, the company's historical daily demand data has been collected, as well as other exogenous data on related daily time series, all to be used in the training of univariate and multivariate forecasting models. The univariate models fitted in this study are the Naïve Method, Mean Method and Holt-Winters models. The multivariate forecasting models are dynamic regression and SARIMAX models, both of which link the target time series with the external-related ones. After training, these models are used to make predictions on the validation data period, and their accuracy is compared between themselves over five different accuracy metrics, thus highlighting the performance and suitability of each model. The results indicate that multivariate models present significantly better accuracy compared to univariate ones in the test period. The best results from the multivariate models present a MAPE value of 9.9% and an MPE of -0.3%.

Keywords: Demand forecasting; Natural gas; Time series; Open gas market.

Resumo: Este estudo tem como objetivo investigar a previsão do consumo de gás natural em uma fábrica de cimento no Rio de Janeiro, no contexto da transição do mercado regulado para o mercado livre de gás natural, conforme as diretrizes da "nova lei do gás" no Brasil. Utiliza-se a previsão da demanda como ferramenta potencial para reduzir os custos de aquisição, transporte e armazenamento de gás na malha de gasodutos. Para isso, foram coletados dados históricos diários de consumo da empresa, bem como outras séries temporais diárias relacionadas, a serem utilizadas no treinamento de modelos de previsão univariados e multivariados. Os modelos univariados ajustados neste estudo são o Método Ingênuo (Naïve), o Método da Média e o modelo de Holt-Winters. Os modelos multivariados de previsão são um modelo de regressão dinâmica e modelos SARIMAX, ambos conectando a série temporal alvo com variáveis externas relacionadas. Após o treinamento, os modelos são utilizados para realizar previsões no período

Received May 15, 2025 - Accepted July 29, 2025

Financial support: This work was supported by the Brazilian National Council for Scientific and Technological Development (CNPq) [309546/2025-7 to I.T.P.]; the Carlos Chagas Filho Foundation for Research Support in Rio de Janeiro State (FAPERJ) [E-26/210.858/2024 and E-26/204.540/2024 to I.T.P.]; the Coordination for the Improvement of Higher Education Personnel (CAPES); and the Pontifical Catholic University of Rio de Janeiro.



This is an Open Access article distributed under the terms of the Creative Commons Attribution license (<https://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

de validação, e sua acurácia é comparada entre si por meio de cinco métricas diferentes, evidenciando o desempenho e a adequação de cada modelo. Os resultados indicam que os modelos multivariados apresentam uma acurácia significativamente superior em comparação aos univariados no período de teste. Os melhores resultados entre os modelos multivariados apresentam um valor de MAPE de 9.9% e de MPE de -0.3%.

Palavras-chave: Previsão de consumo; Gás natural; Séries temporais; Mercado livre de gás.

1 Introduction

The enactment of the New Gas Law (Law No. 14.134 of April 8, 2021) represents a significant milestone for the Brazilian natural gas market. This law, designed to promote competitiveness, reduce costs, and attract new investments, holds the potential to transform the industry. It follows the "New Gas Market" program launched by the Federal Government in 2019, which aims to establish a more dynamic and integrated market aligned with international best practices (Brasil, 2021).

The new legislation aims to unbundle the market, allowing the entry of new players and facilitating access to essential infrastructure, such as pipelines. This change is crucial to improving the competitiveness of the sector and aligning with global trends in more open and diversified markets (Empresa de Pesquisa Energética, 2021).

These transformations are expected to increase the supply and reduce the costs of natural gas in Brazil. They also promise to foster investments in infrastructure and enhance the competitiveness of the national industry in various sectors such as petrochemicals, metallurgy, and thermoelectric generation (Brasil, 2021).

In this new market context, the predictability of gas demand is more crucial than ever for consumers. Before this market shift, natural gas consumers would not have had the option to be involved in natural gas logistics or upstream supply, only being relegated to a passive consumer role. As such, forecasting internal natural gas demand would not be as relevant a necessity due to this limited autonomy. In the new paradigm, accurate demand forecasting is not just a planning tool, but a strategic necessity. It ensures that enough gas is available to meet consumer needs without supply interruptions, and helps optimize the operation of distribution networks, minimizing operational costs and avoiding resource waste (Liu et al., 2023).

For this reason, it is of strategic interest to make demand forecasts based on consumer historical curves. Time series forecasting models can be used to do this, and they are trained on existing demand data and related time series.

This analysis will cover a case study of a cement factory in Rio de Janeiro consuming natural gas. The aim is to train forecasting models on a daily demand time series using historical data and other exogenous time series. The univariate models will be the Naïve Method, the Mean Method and the Holt-Winters. As for the multivariate models, one will be fitted through a dynamic regression approach, and the other will be based on a seasonal auto-regressive integrated moving-average model utilizing exogenous variables (SARIMAX). These models will be tested over the same validation period, and their results will be compared.

To make and evaluate these forecasts, firstly, this study will present an exploratory analysis of the historical daily demand data, including an autocorrelation analysis, looking for trends and seasonality in the series. Afterwards, the previously mentioned univariate forecasting models will be adjusted and fitted using the insights obtained from the exploratory analysis. A similar process is then conducted for the multivariate models, with a cross-correlation analysis being performed between the new added exogenous daily

time series and the historical demand series that will support the fitting of the dynamic regression and SARIMAX models. Five different accuracy metrics are then listed on their pros and cons and proposed to evaluate the performance of the forecasting models. All forecasting models will then be tested over a validation interval of the data corresponding to the latter 20% of the historical demand series, the forecasts are all made on a rolling manner on a D+1 basis, where each prediction is made in a daily basis and only for the next day. Finally, after constructing the testing predictions, their performance is compared to each other and some final conclusions are drawn.

2 Literature review

Natural gas demand forecasting has been extensively studied due to its critical role in energy planning, supply chain management, and economic decision-making. The literature has explored various forecasting methodologies, ranging from traditional statistical techniques to advanced machine learning models. This section presents an overview of key contributions in the field.

Early natural gas demand forecasting works predominantly relied on classical statistical models such as Autoregressive Integrated Moving Average (ARIMA) and its variations. For instance, Box & Pierce (1970) introduced the ARIMA framework, which has since become a cornerstone in time series forecasting. Studies such as Pindyck (1999) effectively applied ARIMA models to capture seasonality and economic trends in natural gas pricing and consumption.

Machine learning methods have gained prominence with the increasing data availability and computational power. Hastie et al. (2009) provided a comprehensive foundation on machine learning techniques, including regression trees, random forests, and support vector machines (SVM), which have been successfully applied in gas demand forecasting. For example, Hong (2009) demonstrated the superior performance of support vector regression (SVR) in predicting electric load demand compared to traditional methods.

More recent advancements incorporate deep learning architectures such as recurrent neural networks (RNNs) and long short-term memory (LSTM) networks into forecasting energy demands. Zhou et al. (2022) applied LSTM models to improve forecast accuracy for natural gas and electricity consumption in China, showing improved performance in capturing temporal dependencies.

Combining traditional statistical methods with machine learning models has emerged as a promising strategy. Zhang (2003) proposed a hybrid ARIMA-ANN model that leverages the strengths of both approaches. Such hybrid models have improved accuracy by simultaneously capturing linear and non-linear patterns. This study laid much of the groundwork for subsequent research on forecasting techniques utilizing Neural Networks and deep learning methods.

Soldo et al. (2014) has explored auto-regressive forecasting models to forecast day-ahead residential natural gas consumption, utilizing exogenous variables such as solar radiation. This approach is similar to the one proposed in this study in many ways.

In light of these studies, this paper focuses on applying classical forecasting methodologies in a more didactic way in an industrial setting on a day-ahead basis and utilizing exogenous explanatory variables. In this manner, it also aims to establish more scientific approaches to decision-making and forecasting in the developing new open gas market in Brazil, where agents are only beginning to take accountability for the management of natural gas supply and transportation and where the literature has not yet developed to real base cases.

3 Materials and methods

To carry out this study, a series of critical information and definitions must be outlined in stages according to the following descriptions.

3.1 Target time series analysis

Relevant data must be collected and processed for training and validating the time series forecasting models. The company provided a dataset with 972 observations containing historical daily time series from February 1, 2021, to September 30, 2023.

This dataset includes a time series of interest for this study, the Realized Daily Quantity (RDQ), which highlights the volume in cubic meters of natural gas consumed daily. This series will be forecasted in this study.

This series is divided between training and validation periods, representing 80% and 20% of the dataset's observations. The training period will be used to fit and adjust the forecasting models, while the validation period will be used to simulate forecasts from the trained models and comparatively measure their performance.

The company's RDQ (Figure 1) is analyzed. To train the forecasting models as accurately as possible, trends, seasonality, and abnormal behaviors in the target series must be detected.

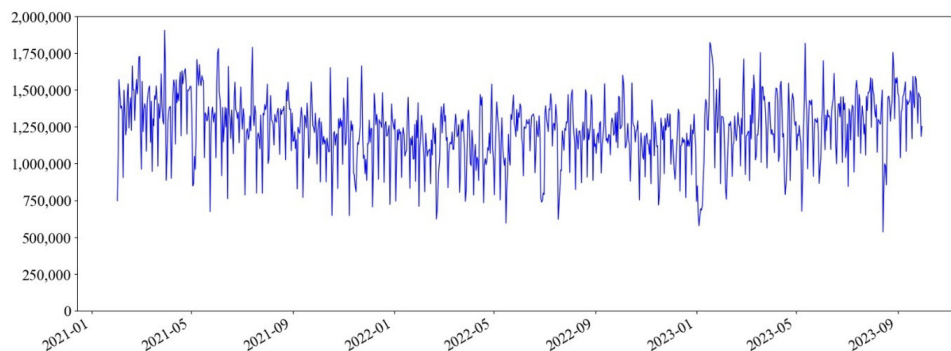


Figure 1. Historical natural gas demand curve (RDQ) [m³/day].

From the raw data, the average demand during the period is 1,233,061 m³/day, with a minimum of 534,887 m³/day and a maximum of 1,904,930 m³/day. However, demand on Mondays is much lower than on all other weekdays (Figure 2). This fact is consistent with the information available that the company's factory performs preventive maintenance on Mondays, temporarily and partially halting the kilns during these processes. The series also shows a standard deviation of 219,336 m³/day.

Additionally, an autocorrelation function (ACF) (Venables & Ripley, 2002) is performed to identify patterns of seasonality and trend in the series thoroughly. The ACF measures correlation values between the target series and its lags (previous time steps).

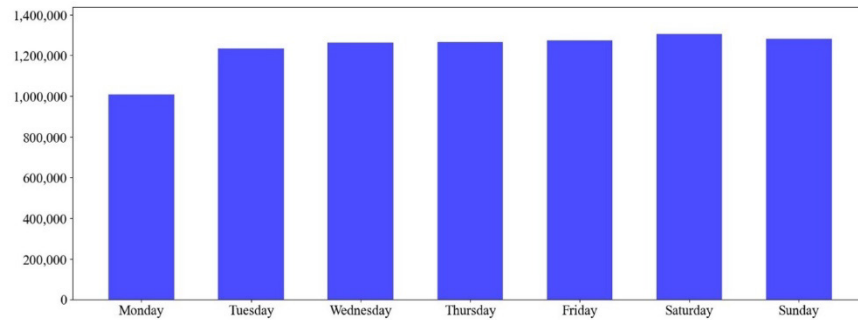


Figure 2. Average RDQ by weekday [m³/day].

The ACF result presents values above the significance threshold for values below lag 5 and multiples of lag 7 (Figure 3). These values in the ACF may indicate a slight growth trend and the presence of weekly seasonality, likely caused by the previously mentioned preventive maintenance events on Mondays.

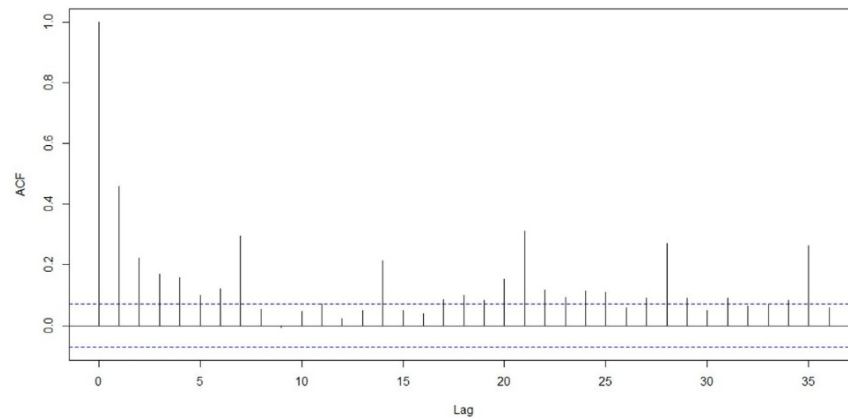


Figure 3. Autocorrelation function of the target series.

3.2 Univariate models

The univariate forecasting models are fitted and adjusted around the base time series analysis and ACF results. At this point in the study, three forecasting models are selected.

Firstly, the Naïve method, the forecast model follows the same value as the last observation in the time series, adhering to the following formulation (Hyndman & Athanasopoulos, 2018):

$$\hat{Z}_{T+h} = Z_T \quad (1)$$

where:

$$\begin{aligned} \hat{Z}_{T+h} &= \text{forecast for } Z_t \text{ made at time } T \text{ for horizon } h \\ Z_T &= \text{the last observation of a time series} \end{aligned}$$

In this case, the only relevant parameter for model adjustment is the definition of Z_T , which is equal to 1,006,760 NG m³/day for the last observation in the training period.

Consequently, the Naive model's forecasts for the validation period will start with this value for the first D+1 prediction and later will always replicate the last day's RDQ value for the next D+1 forecast cycle in accordance with Equation 1.

The Mean Method forecast model is also selected; here, the forecast will follow the average value of all observations of the target time series up to the forecasted observation (Hyndman & Athanasopoulos, 2018). Thus, it follows the following equation:

$$\hat{Z}_{T+h} = \frac{1}{T} \sum_{t=1}^T Z_t \quad (2)$$

In this method, the forecast for the first day of the validation period will equal the average of the observations in the training period, which in this case is 1,215,337 NG m³/day. As new D+1 forecasting cycles in the validation period are added and new previous days' RDQ values are observed, the average of the historical values that compose the D+1 forecast will be recalculated accordingly.

Holt-Winters is the last univariate forecasting method used. This method applies exponentially decreasing weights to past observations, producing forecasts that capture short-term trends and patterns by calculating a weighted average of past values, where weights decrease as data age.

There are three main variations of the exponential smoothing method: Simple Exponential Smoothing (SES), Holt's Model (Holt, 2004), and the Holt-Winters Model (Winters, 1960). In this case study, the ACF analysis (Figure 3) of the target time series indicated the presence of seasonality but not a trend, so the Holt-Winters model is chosen, as it is more suitable for time series with these seasonal characteristics (Winters, 1960). The equations for this model are as follows:

$$\hat{Z}_{T+h} = N_t + hT_t + S_{t+h-m(k+1)} \quad (3)$$

where:

$$N_t = \alpha(Z_t - S_{t-m}) + (1 - \alpha)(N_{t-1} + T_{t-1}) \quad (4)$$

$$T_t = \beta(N_t - N_{t-1}) + (1 - \beta)T_{t-1} \quad (5)$$

$$S_t = \gamma(Z_t - N_{t-1} - T_{t-1}) + (1 - \gamma)S_{t-m} \quad (6)$$

N_t = level of the series at time t

T_t = trend of the series at time t

S_t = seasonality of the series at time t

α = smoothing parameter for the level ($0 \leq \alpha \leq 1$)

β = smoothing parameter for the trend ($0 \leq \beta \leq 1$)

γ = smoothing parameter for the seasonality ($0 \leq \gamma \leq 1$)

$$k = \text{mod} \left(\frac{h-1}{m} \right) \quad (7)$$

h = forecast horizon

m = length of seasonal cycle

The model fitting process obtains the model's parameters α , β e γ by minimizing the sum of squared errors between the calibrated model and the target series values during the training period. In this case, the regression results indicate $\alpha = 0.19$, $\beta = 0.0$ and $\gamma = 1.0$, to be implemented in Equations 4-6. The regression process also detects no trend but perceives apparent seasonality with $m = 7$.

Testing the Holt-Winters fitted model for residuals autocorrelation in the training period (Figure 4) demonstrates that there are still significant values in the lag = 7 multiples, indicating that the fitted model does not fully capture the series' weekly seasonal behaviour.

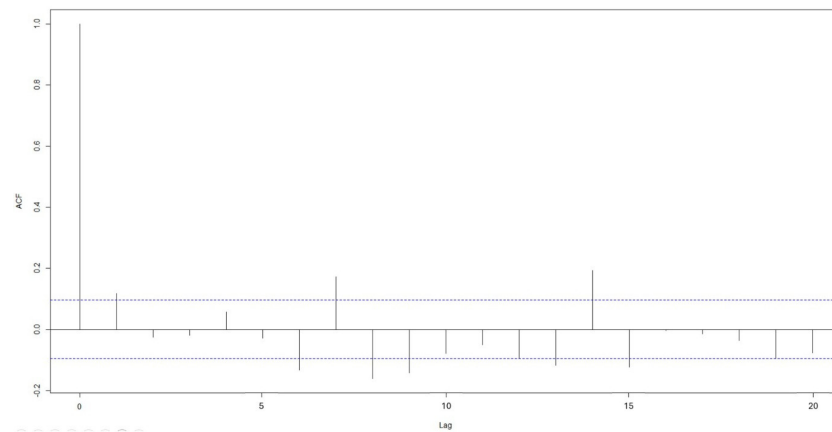


Figure 4. Autocorrelation function of Holt-Winters model residuals.

Running an Autoregressive Conditional Heteroskedasticity (ARCH) (Engle, 1982) test in the model's residuals returns a p-value of $3.8e-5$, rejecting the null hypothesis and demonstrating that the variance of residuals changes over time.

In sequence, the residuals are tested for normality, beyond analyzing the histogram of residuals visually for compatibility (Figure 5), a Jarque-Bera (Jarque & Bera, 1980) test is run over the residuals, resulting in a p-value of $2.2e-16$, indicating that the residuals do not fit into a normal distribution.

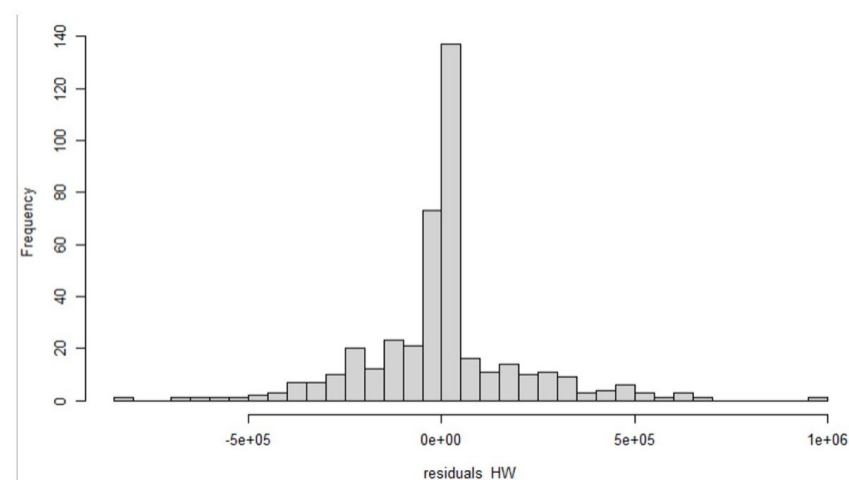


Figure 5. Histogram distribution of Holt-Winters model residuals.

3.3 Multivariate models

Beyond the univariate models, multivariate forecasting models for the target time series are developed using three additional daily time series as explanatory variables. The forecasting models to be trained are a Dynamic Regression and SARIMAX framework models. The exogenous series are as follows:

- Exchange Rate R\$ - US\$ (daily closing) FX [R\$/US\$]
- Brent Crude (daily futures) BRENT [US\$/bbl]
- Dow Jones U.S. Non-Ferrous Metals Index DJUSNF [-]

These additional time series (Investing.com, 2024a, 2024b, 2024c) were chosen based on the authors' technical and commercial knowledge and experience. As such, these choices rely on somewhat subjective perspectives. In the preprocessing analysis, no outliers needed to be removed from any time series, as all values are real and don't distort the overall analysis results.

To identify functional relationships between the explanatory and the target series and construct the multivariate forecasting models, the cross-correlation function (CCF) (Venables & Ripley, 2002) must be calculated between all exogenous variables and the target series. The CCF follows the same formulation as the autocorrelation function (ACF), but instead of calculating correlations between lags of the same series, it calculates correlations between lags of different time series. Three CCF analyses are performed, one for each explanatory series in the dataset.

Among these results, only the correlation values for negative lags between series will be relevant for the regression, as the forecast cannot benefit from future values of other series at the time of the D+1 forecast.

Overall, the CCFs show relatively low values, ranging from -0.30 to 0.15 (Figures 6-8), but they do present values beyond the significance threshold across all combinations.

For the multivariate model, it is shown that the most significant relationship between the lagged variables is in the RDQ – Brent relation.

For the dynamic regression model (Lütkepohl et al., 1999), the fitting process will be based on the results of the CCFs (Figures 6, 7, 8) and the ACF results for the target series (Figure 3). The fitted model must be structured by considering the lags that showed significant correlations between time series. The dynamic regression allows for including the exogenous variables' past values (only negative lags) to capture these significant correlations. The model follows the formulation:

$$RDQ_t = \alpha_0 + \sum_l \eta_l RDQ_{t-l} + \sum_k \alpha_k DJUSNF_{t-k} + \sum_j \beta_j BRENT_{t-j} + \sum_m \gamma_m FX_{t-m} + \epsilon_t \quad (8)$$

where:

RDQ_t = Realized Daily Quantity (target series)

$DJUSNF_t$ = Dow Jones U.S. Non – Ferrous Index

$BRENT_t$ = Brent Crude Daily Futures

FX_t = USD – BRL exchange rate (closing, selling)

α_0 = Intercept

η_l = lag coefficients of RDQ_t

$l \rightarrow$ index of target series lags with significant correlation values

α_k = lag coefficients of $DJUSNF_t$

$k \rightarrow$ index of significant cross lags between the target seiries and $DJUSNF$

β_j = lag coefficients of $BRENT_t$

$j \rightarrow$ index of significant cross lags between the target series and BRENT

$\gamma_m =$ lag coefficients of FX_t

$m \rightarrow$ index of significant cross lags between the target series and FX

$\epsilon_t =$ error term $\sim N(0, \sigma^2)$

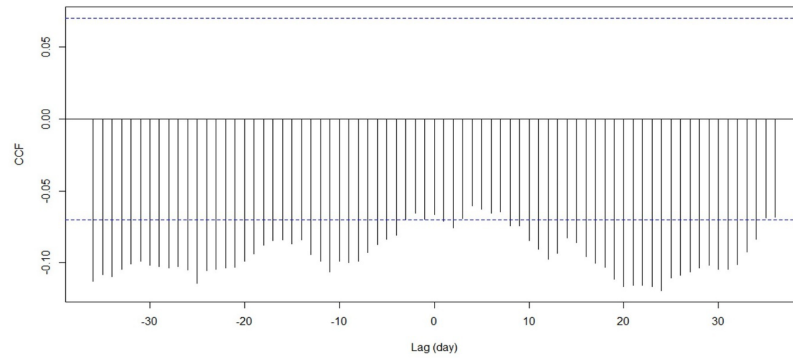


Figure 6. Cross-Correlation Function (RDQ – DJUSNF).

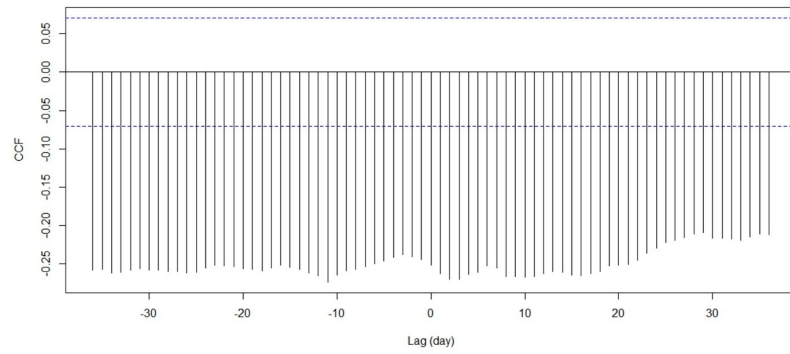


Figure 7. Cross-Correlation Function (RDQ – Brent).

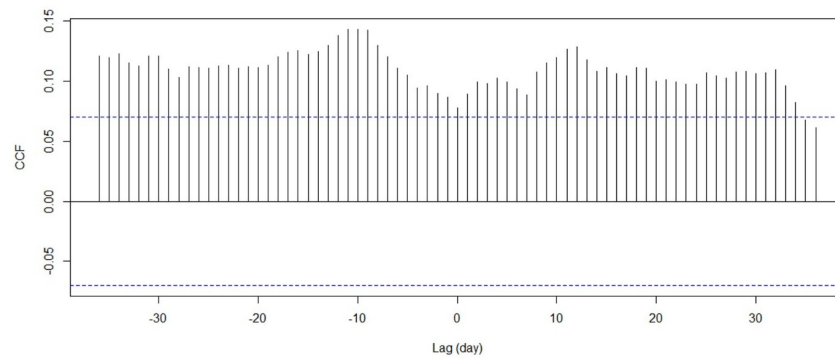


Figure 8. Cross Correlation Function (RDQ – FX).

Regarding the dynamic regression process itself, the initial regression to fit the model considers only the three most significant lag correlations for the target series ACF (Figure 3) and each CCF between variables as relevant intervals for the coefficients in the fitted model.

From this initial regression, the coefficients of the selected variable lags are adjusted to minimize the error in the training interval.

Once the coefficients are obtained, their significance levels are evaluated in the context of dynamic regression. If the p-value of at least one of the model coefficients is greater than 5% (which would indicate that the null hypothesis $H_0: \text{coef.} = 0$ cannot be rejected), the model is rejected, the variable with the least significant coefficient is removed, and the error minimization process is redone. This process is repeated until all the coefficients left in the model are significant.

After finishing the dynamic regression process, the results for the coefficients utilized in the multivariate model fitting are obtained and applied in Equation 8.

It is clear (Table 1) that the only significant relationship between different time series and the target series is given in the lagged BRENT correlation, which produces a coefficient in the multivariate model. This indicates a pattern over the training period where, when the price of natural gas increases (that being defined by the BRENT oil price), there would be a subsequent decrease in natural gas consumption.

Table 1. Parameters of the multivariate dynamic regression forecasting model.

Coefficient	Estimate	Std. Error	t-value	Pr (> t)
Intercept	5.630e+05	7.688e+04	7.323	6.15e-13
Lag (RDQ, -1)	4.121e-01	3.184e-02	12.946	< 2e-16
Lag (RDQ, -7)	2.213e-01	3.187e-02	6.944	8.15e-12
Lag (BRENT, -10)	-1.373e+03	4.535e+02	-3.028	0.00254

In this scenario, the model still presented an R^2 of 0.285, indicating that 28.5% of the changes in the target series are explained by variations in the selected lagged series, a value that is not significant.

Through the fitting process in the Dynamic Regression, the model's residuals for the training period still show significant correlation values for multiples of lag=7 (Figure 9), similar to the analysis of the residuals done in the Holt-Winters model.

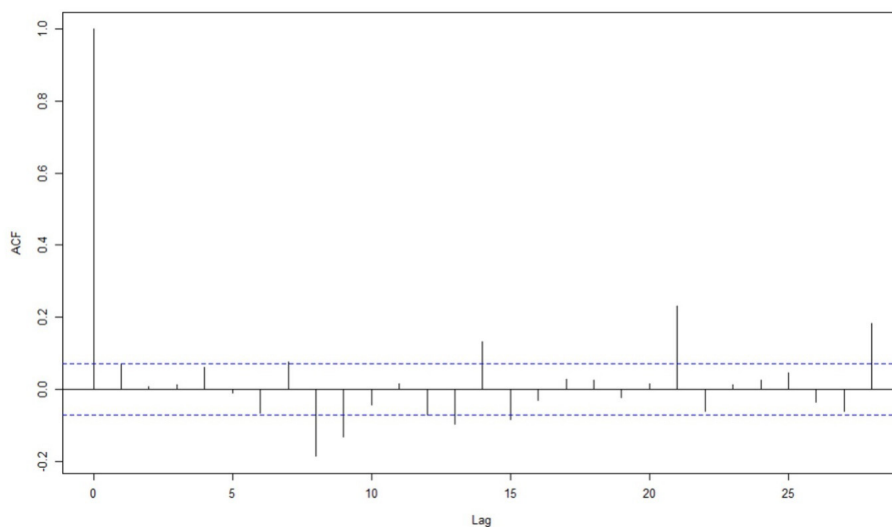


Figure 9. Autocorrelation function of the Dynamic Regression model residuals.

Checking the Dynamic Regression model's residuals for Homoskedasticity through the Breusch-Pagan (Breusch & Pagan, 1979) test, provides a p-value of $7.9e-5$, rejecting the null hypothesis of Homoskedasticity and indicating inconsistencies in the model's fit.

The residuals of the Dynamic regression model fit are tested for normality visually (Figure 10) and using the Jarque-Bera test. A p-value of $5.7e-7$ in the test indicates that, although these residuals are closer to normality than the ones in the Holt-Winters model, they still cannot be considered a normal distribution, leaving room for better fits in other forecasting models.

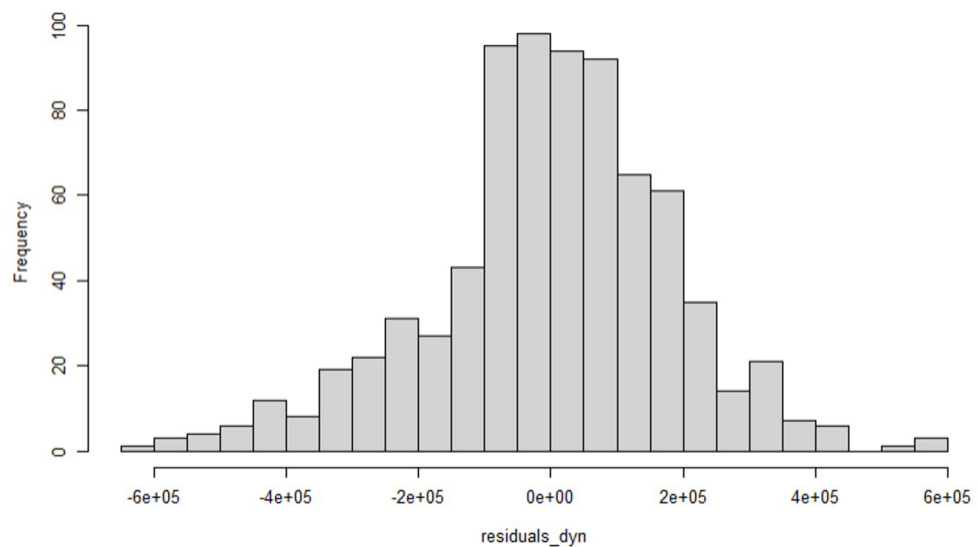


Figure 10. Histogram distribution of the Dynamic Regression model residuals.

As the final forecasting model considered in this study, the SARIMAX model has been chosen. This is a variation of the autoregressive integrated moving average (ARIMA) model that considers seasonality (S) and exogenous variables (X). The model is molded by the $SARIMAX(p, d, q)(P, D, Q, s)$ parameters.

In the fitting process for the SARIMAX model over the training data, multiple sets of different configurations of the $(p, d, q)(P, D, Q, s)$ parameters are tested, with the coefficient regression realized over each configuration. At each configuration, the Akaike Information Criterion (AIC) is calculated. It gives a metric balancing the model's fit with its own complexity to avoid overfitting. This becomes the primary metric for determining the model parameters. The configuration resulting in the lowest AIC value is the one utilized.

After the fitting process, the $SARIMAX(3, 1, 5)(2, 0, 2, 7)$ configuration is selected, with an AIC value of 21,123.1. Analyzing the autocorrelation function for the SARIMAX model's residuals (Figure 11), it becomes clear that this model is better fitted to capture the seasonality patterns than the previous regression-based models. This is likely due to the order 1 differentiation defined in the optimal model parameters.

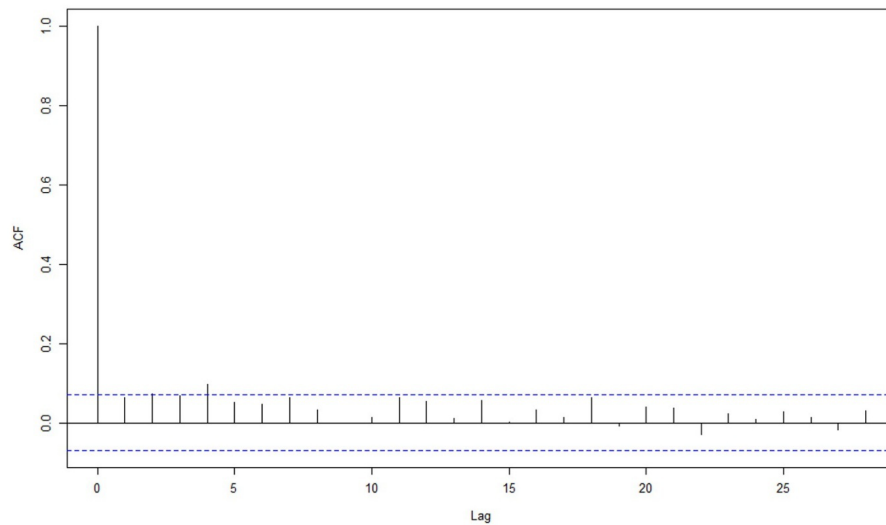


Figure 11. Autocorrelation function of the SARIMAX model residuals.

A Breusch-Pagan test is also conducted on the residuals to check their Homoskedasticity. This test results in a p-value of 0.00, rejecting the null hypothesis of Homoskedasticity and indicating that the model fit still has inconsistencies in the error distribution along the training interval.

Checking for normality in the SARIMAX model presents the best fit visually (Figure 12) but still returns a $2.2\text{e-}16$ p-value in the Jarque-Bera test, rejecting the hypothesis of normality.

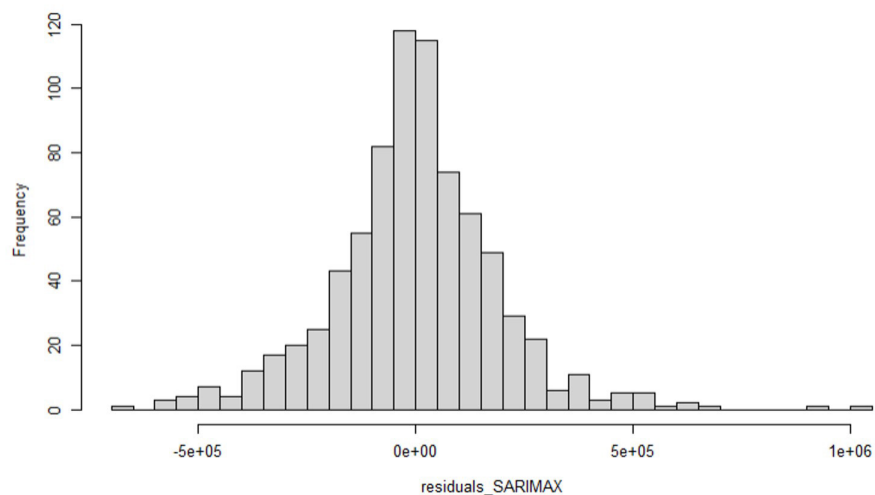


Figure 12. Histogram distribution of the SARIMAX model residuals.

3.4 Accuracy metrics

To measure the performance of the models' forecasts during the validation period, different error metrics will be used (Hyndman & Athanasopoulos, 2018). These include:

- ME – Mean Error

It is the average of the differences between the predicted and actual values. While it provides insight into systematic bias in the model, its primary limitation is that positive and negative errors can offset each other, potentially masking significant forecasting inaccuracies.

$$ME = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i) \quad (9)$$

- RMSE – Root Mean Squared Error

Calculated from the square root of the average squared differences between predicted and actual values. It gives greater weight to more significant errors, making it highly sensitive to outliers. This characteristic makes RMSE particularly suitable when significant deviations from the mean are critical to detect.

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (10)$$

- MAE – Mean Absolute Error

It represents the average of the absolute differences between predicted and actual values. Unlike RMSE, MAE treats all deviations equally, providing a balanced measure of forecast accuracy. However, compared to RMSE, it may underestimate the impact of large errors.

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (11)$$

- MPE – Mean Percentage Error

This metric calculates the average percentage errors, indicating the model's tendency to overestimate or underestimate values. Its primary drawback is sensitivity to extreme values, particularly when observed data points are near zero, which can distort the metric.

$$MPE = \frac{1}{n} \sum_{i=1}^n \frac{y_i - \hat{y}_i}{y_i} \quad (12)$$

- MAPE – Mean Absolute Percentage Error

It computes the average absolute percentage difference between predicted and actual values. This metric is intuitive and facilitates straightforward interpretation across different scales. However, like MPE, it becomes unreliable when dealing with values near zero, as it can produce disproportionately large errors.

$$MAPE = \frac{1}{n} \sum_{i=1}^n \frac{|y_i - \hat{y}_i|}{y_i} \quad (13)$$

Although other error metrics are included to add context to the results of this study, the MAPE will be the leading benchmark, as with series with all positive and significant values, it is still preferred as a factor of simplicity (Hyndman & Koehler, 2006).

3. Results

Initially, the univariate models—the Naive Model, the Mean Method Model, and the Holt-Winters Model—are forecasted through the validation period.

Visually (Figure 13.) the Holt-Winters fitted model performs best and follows the demand curve time series more closely.

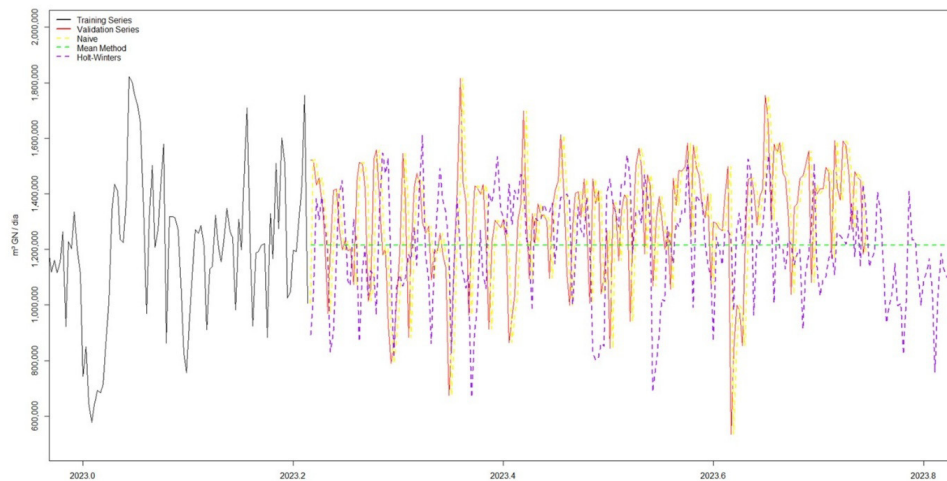


Figure 13. Univariate model rolling starts forecasts inside the validation period.

The multivariate model forecasts for the validation period are added at this point.

The perceived performance of the multivariate models (Figure 14), as it was with the Holt-Winters univariate model, follows the actual demand curve even closer.

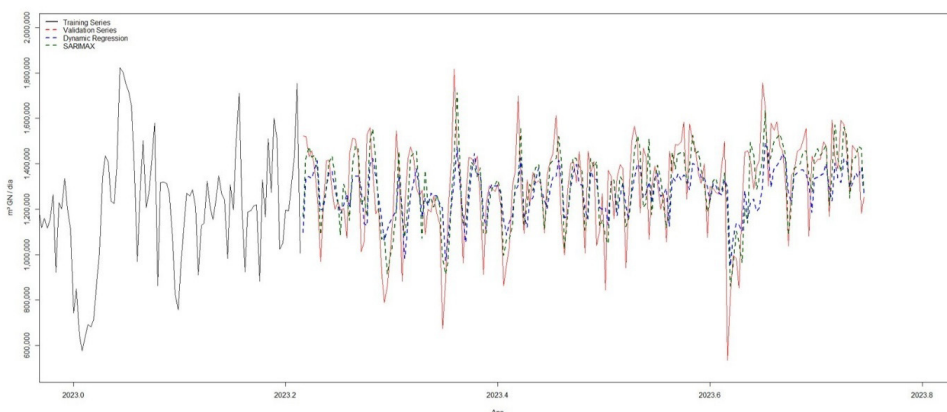


Figure 14. Multivariate models rolling start forecast for the validation period.

For an analytical and objective comparative evaluation of each model's performance, their predictions for the validation period are measured through the accuracy metrics previously described in the Materials and Methods section (Table 2).

Table 2. Accuracy metrics for all forecasting models in the validation period.

Model	ME	RMSE	MAE	MPE	MAPE
Naïve	1,267.5	219,027.8	163,073.8	-1.8%	13.9%
Mean Method	88,802.5	223,735.1	186,566.9	3.9%	14.9%
Holt-Winters	82,843.5	286,225.0	228,841.6	3.5%	18.4%
Dynamic Regression	27,068.8	181,299.1	141,811.4	-0.3%	11.8%
SARIMAX	562.8	158,905.6	119,930.9	-1.6%	9.9%

As the visual analysis revealed, the Multivariate Models present the best MAPE accuracy metrics among all proposed models in this study. Within the univariate models, there is one clear superior model, the Naïve model, presenting better accuracy and performance in all metrics.

4 Discussion

This study's results suggest that a multivariate dynamic regression approach offers a solid improvement over univariate methods for forecasting daily natural gas demand in an industrial context. Still, the model shows difficulties in capturing the seasonality patterns that are better perceived in the SARIMAX model, which is corroborated by the model's residuals correlation analysis and overall better accuracy scores in all metrics but the MPE. Multivariate models, especially those that incorporate external economic and market factors, capture complex interactions between time series variables more effectively, which aligns with findings from recent studies in natural gas forecasting. For example, machine learning models such as Long Short-Term Memory (LSTM) networks, which handle sequential dependencies, have shown enhanced predictive power in other markets with dynamic patterns. Integrating economic indicators like Brent crude prices can further improve forecasting accuracy, as these indicators frequently impact fuel-related demand fluctuations (Wei et al., 2019).

While the multivariate approach here benefits from dynamic regression's capacity to address lagged dependencies, some limitations arise due to low correlations between the explanatory series (Brent prices and currency exchange rates) and the demand series (RDQ). Studies using hybrid models like LSTM combined with feature selection methods (e.g., Principal Component Analysis) have shown promise in reducing model complexity while enhancing accuracy, even with varied external variables, as seen in complex energy forecasting tasks (Abbasimehr & Paki, 2022). Applying these techniques could further optimize the regression model by filtering out less relevant variables, potentially mitigating the noise introduced by weakly correlated factors.

On the other hand, the SARIMAX model creates a better link to the exogenous variables, even though its fitted parameters do not investigate specific lags between the variables. In general, the more flexible and expansive approach to the fitting method allowed the model to capture more difficult-to-see patterns in the target series by differentiating the series.

The impact of Brazil's New Gas Law (Law No. 14.134 of April 8, 2021) on natural gas demand forecasting highlights the increased importance of accurate demand modeling in a newly deregulated market. The law, aimed at enhancing competition and accessibility to infrastructure, underscores the need for reliable forecasts to manage demand fluctuations and cost efficiency.

In similar deregulated markets, forecasting accuracy is critical for balancing supply and demand, reducing costs, and enhancing system resilience. The forecast models analyzed in this study, while beneficial in their current form, could be enhanced with additional internal operational data to improve responsiveness to real-time market shifts, a strategy that is effective in industries where precise operational alignment is necessary (McKinsey, 2022).

5 Conclusion

The findings from this natural gas demand forecasting study suggest several avenues for future improvement and application. First, while the multivariate model utilized here provides accuracy gains over univariate methods, adopting machine learning techniques, such as Long Short-Term Memory (LSTM) networks and other neural networks, could enhance predictive capacity by capturing more complex, non-linear patterns in natural gas consumption data.

Although LSTM networks were not implemented in this study, their known ability to capture long-term dependencies and non-linear relationships in sequential data (Zhou et al., 2022) suggests that they could outperform both the dynamic regression and SARIMAX models used here, especially in contexts where seasonality and lag effects are difficult to model explicitly. Comparing results from future LSTM implementations with the benchmarks established by SARIMAX and dynamic regression could provide valuable insights into the trade-offs between model complexity, interpretability, and forecast accuracy in energy demand forecasting.

Moreover, this study underscores the need for a robust combination of historical demand data and high-correlation economic indicators, such as Brent crude prices and currency exchange rates, which, despite their initial limitations here, could yield stronger correlations with optimized selection techniques.

Additionally, the regulatory changes under Brazil's New Gas Law indicate a clear opportunity for demand forecasting to support cost-saving measures in natural gas procurement, storage, transport, and distribution. In deregulated markets, real-time and high-accuracy forecasts are increasingly critical, helping to balance supply with variable demand while minimizing financial and operational risks.

The methodology used here is an easily replicable template for forecasting in natural gas-consuming industries, especially in markets undergoing similar regulatory shifts. Future studies should explore incorporating operational-level data and non-linear, machine-learning-based models to forecast better the demand curves associated with each industry profile.

Statement on Data Availability

The authors hereby make available upon request all the data used to conduct this work as well as the data from the results obtained from such work.

References

- Abbasimehr, H., & Paki, R. (2022). Improving time series forecasting using LSTM and attention models. *Journal of Ambient Intelligence and Humanized Computing*, 13(1), 673-691. <http://doi.org/10.1007/s12652-020-02761-x>.

- Box, G. E. P., & Pierce, D. A. (1970). Distribution of residual autocorrelations in autoregressive-integrated moving average time series models. *Journal of the American Statistical Association*, 65(332), 1509-1526. <http://doi.org/10.1080/01621459.1970.10481180>.
- Brasil. Ministério de Minas e Energia - MME. (2021). *Manual de Boas Práticas Regulatórias do CMGN*. Ministério de Minas e Energia. Retrieved in 2024, September 16, from <https://www.gov.br/mme/pt-br/assuntos/secretarias/petroleo-gas-natural-e-biocombustiveis/novo-mercado-de-gas/cmgn/publicacoes/publicacoes>
- Breusch, T. S., & Pagan, A. R. (1979). A simple test for heteroscedasticity and random coefficient variation. *Econometrica*, 47(5), 1287-1294. <http://doi.org/10.2307/1911963>.
- Empresa de Pesquisa Energética – EPE. (2021). *Plano Decenal de Expansão de Energia 2031*. EPE. Retrieved in 2024, from <https://www.epe.gov.br/pt/publicacoes-dados-abertos/publicacoes/plano-decenal-de-expansao-de-energia-2031>
- Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the Variance of United Kingdom Inflation. *Econometrica*, 50(4), 987-1007. <http://doi.org/10.2307/1912773>.
- Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The elements of statistical learning*. Springer. <http://doi.org/10.1007/978-0-387-84858-7>
- Holt, C. C. (2004). Forecasting seasonals and trends by exponentially weighted moving averages. *International Journal of Forecasting*, 20(1), 5-10. <http://doi.org/10.1016/j.ijforecast.2003.09.015>.
- Hong, W. C. (2009). Electric load forecasting by support vector model. *Applied Mathematical Modelling*, 33(5), 2444-2454. <http://doi.org/10.1016/j.apm.2008.07.010>.
- Hyndman, R. J., & Athanasopoulos, G. (2018). *Forecasting: principles and practice*. Retrieved in 2024, September 16, from <https://research.monash.edu/en/publications/forecasting-principles-and-practice-2>
- Hyndman, R. J., & Koehler, A. B. (2006). Another look at measures of forecast accuracy. *International Journal of Forecasting*, 22(4), 679-688. <http://doi.org/10.1016/j.ijforecast.2006.03.001>.
- Investing.com. (2024a). *Dow Jones Nonferrous Metals – DJUSNF. DJ Nonferrous Metals Index Today*. Investing.com. Retrieved 16 September 2024, from <https://www.investing.com/indices/dj-nonferrous-metals>
- Investing.com. (2024b). *Histórico de Preços Petróleo Brent*. Investing.com. Retrieved in 2024, September 16, from <https://br.investing.com/commodities/brent-oil-historical-data>
- Investing.com. (2024c). *Preços Históricos USD/BRL*. Investing.com. Retrieved in 2024, September 16, from <https://br.investing.com/currencies/usd-brl-historical-data>
- Jarque, C. M., & Bera, A. K. (1980). Efficient tests for normality, homoscedasticity and serial independence of regression residuals. *Economics Letters*, 6(3), 255-259. [http://doi.org/10.1016/0165-1765\(80\)90024-5](http://doi.org/10.1016/0165-1765(80)90024-5).
- Liu, H., Liu, Y., Wang, C., Song, Y., Jiang, W., Li, C., Zhang, S., & Hong, B. (2023). Natural gas demand forecasting model based on LASSO and polynomial models and its application: a case study of China. *Energies*, 16(11), 11. <http://doi.org/10.3390/en16114268>.
- Lütkepohl, H., Teräsvirta, T., & Wolters, J. (1999). Investigating Stability and Linearity of a German M1 Money Demand Function. *Journal of Applied Econometrics*, 14(5), 511-525. [http://doi.org/10.1002/\(SICI\)1099-1255\(199909/10\)14:5<511::AID-JAE529>3.0.CO;2-C](http://doi.org/10.1002/(SICI)1099-1255(199909/10)14:5<511::AID-JAE529>3.0.CO;2-C).
- McKinsey. (2022). *Stronger forecasting in operations management—Even with weak data*. McKinsey. Retrieved in 2024, September 16, from <https://www.mckinsey.com/capabilities/operations/our-insights/ai-driven-operations-forecasting-in-data-light-environments>
- Pindyck, R. (1999). The long-run evolutions of energy prices. *The Energy Journal*, 20(2), 1-27. <http://doi.org/10.5547/ISSN0195-6574-EJ-Vol20-No2-1>.

- Soldo, B., Potočnik, P., Šimunović, G., Šarić, T., & Govekar, E. (2014). Improving the residential natural gas consumption forecasting models by using solar radiation. *Energy and Buildings*, 69, 498-506. <http://doi.org/10.1016/j.enbuild.2013.11.032>.
- Venables, B., & Ripley, B. (2002). *Modern applied statistics with S*. Cham: Springer. <http://doi.org/10.1007/978-0-387-21706-2>
- Wei, N., Li, C., Duan, J., Liu, J., & Zeng, F. (2019). Daily natural gas load forecasting based on a hybrid deep learning model. *Energies*, 12(2), 218. <http://doi.org/10.3390/en12020218>.
- Winters, P. R. (1960). Forecasting sales by exponentially weighted moving averages. *Management Science*, 6(3), 324-342. <http://doi.org/10.1287/mnsc.6.3.324>.
- Zhang, G. P. (2003). Time series forecasting using a hybrid ARIMA and neural network model. *Neurocomputing*, 50, 159-175. [http://doi.org/10.1016/S0925-2312\(01\)00702-0](http://doi.org/10.1016/S0925-2312(01)00702-0).
- Zhou, W., Pan, J., Tao, H., Ding, S., Chen, L., & Zhao, X. (2022). A novel grey seasonal model based on cycle accumulation generation for forecasting energy consumption in China. *Computers & Industrial Engineering*, 163, 107725. <http://doi.org/10.1016/j.cie.2021.107725>.

Editor-in-Chief

Pedro Munari

Author Contributions

Nicholas De Marco Reckman was responsible for Conceptualization, Data curation, Formal analysis, Investigation, Methodology, Software, Validation, Visualization, Writing - original draft and Writing - review & editing. Igor Tona Peres was responsible for Conceptualization, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Writing - original draft, Writing - review & editing.