



José Marcos Melo Mendes

**Previsão de demanda de importação de gás natural
liquefeito (GNL) no mercado Brasileiro**

Dissertação de Mestrado

Dissertação apresentada como requisito parcial para obtenção do grau de Mestre em Logística pelo Programa de Pós-Graduação em Engenharia de Produção (opção profissional) do Departamento de Engenharia Industrial da PUC-Rio.

Orientador: Prof. Reinaldo Castro Souza

Rio de Janeiro
Setembro de 2019



José Marcos Melo Mendes

**Previsão de demanda de importação de gás natural
liquefeito (GNL) no mercado Brasileiro**

Dissertação apresentada como requisito parcial para obtenção do grau de Mestre em Logística pelo Programa de Pós-Graduação em Engenharia de Produção (opção profissional) da PUC-Rio. Aprovada pela Comissão Examinadora abaixo.

Prof. Reinaldo Castro Souza

Orientador

Departamento de Engenharia Industrial - PUC-Rio

Prof. Antônio Márcio Tavares Thomé

Departamento de Engenharia Industrial - PUC-Rio

Prof. Soraida Aguilar Vargas

Departamento de Engenharia Industrial - PUC-Rio

Rio de Janeiro, 30 de setembro de 2019

Todos os direitos reservados. É proibida a reprodução total ou parcial do trabalho sem autorização do autor, do orientador e da universidade.

José Marcos Melo Mendes

Graduou-se em Engenharia Industrial Mecânica pelo Centro Federal de Tecnologia Celso Suckow da Fonseca (CEFET-RJ) em 2001. Atualmente trabalha na Petróleo Brasileiro S.A. (Petrobras), na área responsável pela gestão dos ativos de regaseificação de GNL.

Ficha Catalográfica

Mendes, José Marcos Melo

Previsão de demanda de importação de gás natural liquefeito (GNL) no mercado brasileiro / José Marcos Melo Mendes; orientador: Reinaldo Castro Souza. – 2019.

67 f.: il. color.; 30 cm

Dissertação (mestrado) – Pontifícia Universidade Católica do Rio de Janeiro, Departamento de Engenharia Industrial, 2019.

Inclui bibliografia

1. Engenharia Industrial - Teses. 2. GNL 3. Previsão. 4. Demanda térmica por gás natural. I. Souza, Reinaldo Castro. II. Pontifícia Universidade Católica do Rio de Janeiro. Departamento de Engenharia Industrial. III. Título.

CDD: 658.5

Agradecimentos

A Deus, por permitir que eu concluísse esta etapa da minha vida profissional.

À minha mãe por semear a importância dos estudos em nossa família, e por lutar muito por isso.

Ao meu orientador, Professor Reinaldo Castro Souza, pelo dom de simplificação de problemas. Por sua simplicidade.

À minha esposa e filhos por toda força e compreensão.

À pesquisadora Margarete do DEI da PUC-RJ que gentilmente rodou no R a função de transferência.

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

Resumo

Mendes, José Marcos Melo; Reinaldo Castro Souza (Orientador). **Previsão de demanda de importação de gás natural liquefeito (GNL) no mercado Brasileiro**. Rio de Janeiro, 2019. 67p. Dissertação de Mestrado - Departamento de Engenharia Industrial, Pontifícia Universidade Católica do Rio de Janeiro.

O gás natural liquefeito (GNL) tem importante papel na cadeia de suprimento de gás natural e energia elétrica no país. No Brasil, o produto importado visa prioritariamente o atendimento da demanda termelétrica para este insumo. O modelo de geração termelétrico é flexível e a característica principal deste modelo é a necessidade de se garantir uma oferta de combustível flexível para usinas que deverão suprir uma demanda termelétrica igualmente flexível. O objetivo deste trabalho é modelar e avaliar o desempenho de um modelo causal de regressão dinâmica em previsões de demandas de GNL, de longo prazo, para o atendimento das usinas termo-elétricas a gás natural e comparar a capacidade preditiva com modelos tradicionais univariados, vastamente discutidos na literatura, como o ARIMA e o amortecimento exponencial. Foram estabelecidos cenários da variável explicativa a fim de avaliar como tais cenários influenciam a demanda do GNL. O produto desenvolvido nessa dissertação mostrou a viabilidade dos modelos causais serem utilizados como ferramenta para a simulação da demanda de GNL considerando cenários das variáveis causais “What if analysis”.

Palavras-chave

GNL; previsão; demanda térmica por gás natural.

Abstract

Mendes, José Marcos Melo; Reinaldo Castro Souza (Advisor). **Forecasting Demand for Liquefied Natural Gas (LNG) in Brazilian Market.** Rio de Janeiro, 2019. 67p. Dissertação de Mestrado - Departamento de Engenharia Industrial, Pontifícia Universidade Católica do Rio de Janeiro.

Liquefied natural gas (LNG) plays an important role in the country's natural gas and electricity supply chain. In Brazil, the imported product primarily targets the supply of thermoelectric demand for this input. The thermoelectric generation model is flexible and the main characteristic of this model is the need to guarantee a flexible fuel supply for plants that must meet an equally flexible thermoelectric demand. The objective of this work is modeling and evaluate the performance of a causal dynamic regression model in long-term LNG demand forecasts, for supplying of natural gas thermoelectric plants and to compare the predictive capacity with traditional models widely discussed in the literature, such as the ARIMA and the dynamic regression model. Scenarios of the explanatory variable were established in order to evaluate how the explanatory variables influences LNG demand. The product developed in this dissertation showed the viability of causal models to be used as a tool for the simulation of LNG demand considering scenarios of the causal variables. "What if analysis".

Keywords

LNG; forecast; thermal demand for natural gas.

Sumário

1. Introdução	12
1.1. Motivação	13
1.2. Objetivo	14
1.3. Estrutura do trabalho	14
2. Mercado de Gás e Energia	15
2.1. Mercado de Gás Natural no Brasil	15
2.2. O Mercado de Gás Natural no Mundo e o GNL	24
2.3. Preço de Liquidação das Diferenças (PLD)	28
2.4. Térmicas a Gás Natural no Brasil	32
2.5. Terminais de importação e de Regaseificação no Brasil	35
3. Metodologia	38
3.1. Revisão Bibliográfica	38
3.2. Métodos de Previsão	42
3.2.1. Amortecimento exponencial	43
3.2.2. Modelagem auto-regressiva integrada de média móvel (ARIMA)	44
3.2.3. Modelos de Regressão Dinâmica	46
3.2.4. Modelos de Função de Transferência	49
3.2.5. Métricas de Desempenho	51
4. Resultados	53
4.1. Dados e Ajustes	53
4.2. Aplicação dos modelos Univariados	56
4.3. Aplicação dos modelos causais	57
4.4. Análise e comparação dos resultados	58
5. Conclusão	63
5.1. Trabalhos futuros	64
6. Referências bibliográficas	65

Lista de figuras

Figura 1 – Representação da Cadeia Logística do GNL.	12
Figura 2 – Oferta de gás natural ao mercado nacional.	16
Figura 3 – Segmentação da produção nacional.	17
Figura 4 – Previsão de Produção de Petróleo Nacional.	18
Figura 5 – Previsão produção bruta e líquida de gás natural.	18
Figura 6 – Rotas de Escoamento do Pré-sal.	20
Figura 7 – Importação Bolívia e volumes regaseificados.	20
Figura 8 – Demanda segmentada de gás natural.	22
Figura 9 – Evolução da Matriz de Geração de Eletricidade Brasileira.	22
Figura 10 – Capacidade e Processamento UPGNs em MMm3/dia em 2018.	23
Figura 11 – Volumes de GNL comercializados 1990-2018.	25
Figura 12 – Volumes Exportados de GNL e fatia do mercado em 2018.	25
Figura 13 – Volumes Importados de GNL e fatia do mercado em 2018.	26
Figura 14 – Preços de referência do gás natural.	27
Figura 15 – Problema de Decisão da Operação.	29
Figura 16 – Critério de planejamento da operação.	30
Figura 17 – Geração Térmica GN x Geração Térmica Total x Geração Total.	33
Figura 18 – ENA x MLT x CMO.	33
Figura 19 – Emissões por fontes de geração.	34
Figura 20 – Emissões de CO2 por fonte no setor de eletricidade nos EUA.	35
Figura 21 – Vista aérea do terminal de GNL do Pecém - Ceará	36
Figura 22 – Número dos métodos em artigos selecionados.	42
Figura 23 – Série Regas Total de janeiro de 2009 a janeiro de 2019.	54

Figura 24 – Volumes de Regas Totalizados x PLD NE.	55
Figura 25 – Função de Autocorrelação de Erros (FAC <i>Error</i>) – Modelo ARIMA (0,1,0).	57
Figura 26 – Volumes de Regaseificação Previstos Regressão Dinâmica em MMm ³ /dia.	60
Figura 27 – Volumes de Regaseificação Previstos Função de Transferência Box & Jenkins.	61

Lista de tabelas

Tabela 1 – Preços da Cadeia Produtiva de GNL.	28
Tabela 2 – Preços Vigentes PLD.	32
Tabela 3 – Terminais de GNL e capacidades de tancagem e regaseificação.	36
Tabela 4 – Números de artigos por palavra chave.	40
Tabela 5 -Números de artigos por palavra chave “natural gas”.	40
Tabela 6 – Desempenho <i>in sample</i> dos modelos univariados e de regressão dinâmica.	59
Tabela 7 – Desempenho <i>out-of-sample</i> dos modelos univariados e de regressão dinâmica.	59
Tabela 8 – Desempenho <i>out-of-sample</i> do modelo de Função de Transferência.	61
Tabela 9 – MAPE Previsões x Valores Reais – Fevereiro a Agosto de 2019.	62

Siglas

ANEEL – Agência Nacional de Energia Elétrica
ANP – Agência Nacional do Petróleo, Gás Natural e Biocombustíveis
ARIMA – Modelo Auto-Regressivo Integrado de Médias Móveis
Bm3 – bilhões de metros cúbicos
Bm3/ano – bilhões de metros cúbicos por ano
CCEE – Câmara de Comercialização de Energia Elétrica.
CMO – Custo Marginal de Operação
EPE – Empresa de Pesquisa Energética
EUA – Estados Unidos da América
FSRU – Floating Storage and Regaseification Unit
Gaspetro – Petrobras Gás S.A. (subsidiária da Petrobras)
GN – Gás Natural
GNL – Gás Natural Liquefeito
HH – Henry Hub, preço spot no ponto nodal (hub) na Luiziana (EUA)
LNG – Liquefied Natural Gas
MAS – Master Sales Agreements
MLT – Média de Longo Termo
MMA – Ministério do Meio Ambiente
MMBtu – Milhões de British Thermal Units
MME – Ministério das Minas e Energia
MMm³/dia – Milhões de metros cúbicos por dia
MTPA – Milhões de toneladas por ano
MW – Megawatt
NBP – preço spot no ponto nodal (hub) britânico
ONS – Operador Nacional do Sistema
PDE – Plano Decenal de Energia
PLD – Preço de Liquidação de Diferenças
UT – Usina Termoeletrica
TTF – preço spot no ponto nodal (hub) holandês
US\$ – dólares norte-americanos

Introdução

O mercado de gás natural no Brasil e no mundo está em pleno período de mudanças. Considerado como o combustível de transição para fontes mais limpas, a facilidade de transporte e alcance aos principais mercados, assim como o aumento da oferta e demanda são fatores que estimulam a indústria e tendem a estimular sua procura nos principais mercados. O gás natural liquefeito ou o GNL tem importante papel neste mercado pois possibilita a redução dos volumes e pressões necessários ao transporte e distribuição do produto. O GNL nada mais é que o gás natural liquefeito a temperaturas criogênicas, da ordem de menos 600°C e pressão atmosférica. O principal ganho é a redução do volume em 600 vezes quando liquefeito, ou seja 1 m³ de GNL, após regaseificado será aproximadamente 600 m³ de gás natural na forma gasosa. Esse modal de comercialização e transporte cresce ano a ano e movimentou em 2018, 316 milhões de toneladas (MTPA), ou 9,8% a mais que o ano anterior. Em grande parte o aumento está relacionado ao aumento da demanda nos principais mercados mundiais. (Fonte: IGU, 2019).

No Brasil, a presença na cadeia do GNL se dá apenas no elo de importação, através dos terminais de GNL flexíveis de regaseificação e pode ser representada pela Figura 1.

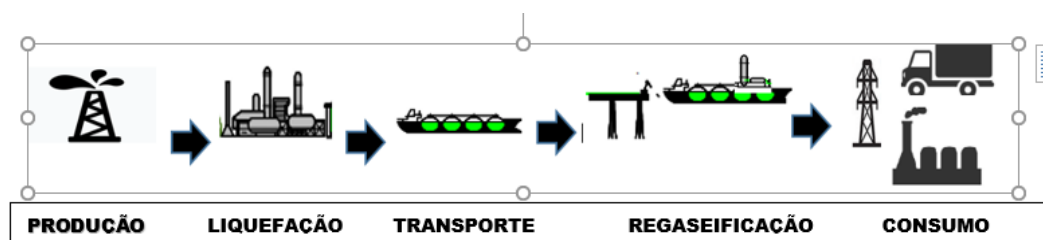


Figura 1: Representação da Cadeia Logística do GNL.
Fonte: Compilação Própria.

Os três terminais no Brasil tem a mesma filosofia de armazenamento e regaseificação em embarcações flutuantes do tipo FSRU (Floating, Storage, Regaseification Unit). Essas unidades flutuantes, dentre outras vantagens tem grande flexibilidade operacional e podem armazenar o GNL em tanques criogênicos até que a regaseificação seja necessária. Para o caso brasileiro, de uma

forma geral, a importação e regaseificação tradicionalmente flutua com a demanda de energia térmica que é fortemente dependente da geração hídrica. Uma breve abordagem sobre o mercado de gás no Brasil e no mundo, assim como a relação com o mercado de energia será realizada na seção 2.

Neste trabalho, foi dedicada atenção à previsão de regaseificação de GNL para um horizonte de até cinco anos. O objetivo é comparar os resultados de um modelo causal de regressão dinâmica com os resultados dos modelos univariados, ARIMA e amortecimento exponencial.

1.1

Motivação

A importância econômica do mercado de gás natural, em um momento de mudança estrutural no Brasil é um dos motivos para elaboração dessa pesquisa. Durante a redação deste trabalho está em curso a elaboração o programa do governo brasileiro, intitulado o “*Novo Mercado do Gás*” (Projeto de Lei 6407/13, que visa efetivar mudanças estruturais no mercado de forma a quebrar o monopólio da Petrobras em toda a cadeia do gás natural, atrair novos atores, e ampliar a concorrência.

Através de um novo conjunto de regramentos e incentivos legais de acesso e incentivo à ampliação da infraestrutura de Transporte, Processamento, Importação e escoamento da produção pretende-se aumentar a oferta de gás no país, seja através do pré-sal, seja através da importação e comercialização do GNL.

Vale ressaltar a assinatura do termo de cessação de conduta que a Petrobras assinou com o CADE (Conselho Administrativo de Defesa Econômica) em 8 de julho de 2019, que gerou uma série de obrigações que visam acelerar a abertura do mercado, dentre estas ações está a obrigação de não comprar mais gás dos outros produtores no E&P, desinvestir na participação nas distribuidoras de gás, negociar acesso ao processamento, desinvestir na capacidade de transporte e por fim arrendar até o final de 2020 terminal de GNL da Bahia de todos os Santos.

Por fim espera-se alcançar um mercado mais competitivo e que se espelha nas melhores práticas internacionais.

O cenário atual de mudanças e incertezas, demanda dentre outras ações estudos de previsão de investidores, governo, grandes consumidores e sociedade de uma forma geral interessados em dados que apoiem suas decisões.

O mercado de geração de energia é componente importante na demanda flexível de gás natural liquefeito e a previsão da demanda futura é importante ferramenta auxiliar ao apoio e decisão aos diversos agentes do setor.

1.2

Objetivo

O objetivo geral do trabalho é testar o desempenho dos modelos causais, via Regressão Dinâmica, em previsões de regaseificação de GNL, no horizonte de até 5 anos. A idéia é testar a variável explicativa PLD, variável de forte correlação entre os mercados de GNL e de energia.

A pesquisa tem como objetivo específico comparar a capacidade preditiva dos referidos modelos com métodos tradicionais discutidos amplamente na literatura, o ARIMA e o amortecimento exponencial. Para tanto, serão testados alguns cenários para a variável exógena e avaliados qual o comportamento da previsão de demanda de GNL.

1.3

Estrutura do trabalho

O trabalho está organizado a partir da seção 1 de Introdução ao GNL no Brasil. Em seguida, na seção 2, são apresentados alguns dados relevantes sobre o mercado de gás natural no Brasil e no mundo, sua relação com a demanda das térmicas a gás natural e consequentemente o PLD. Na seção 3 é apresentada a revisão da literatura, os esforços empreendidos nessa linha de pesquisa e os métodos de previsão e os parâmetros usados na aferição do desempenho das previsões do trabalho.

Ao longo da seção 4 é discutida a implementação dos modelos assim como seus resultados para a previsão de regaseificação de GNL. Por fim, na seção 5 há algumas conclusões e sugestão de oportunidades para futuros estudos.

2

Mercado de Gás e Energia

Pretende-se nessa seção abordar o mercado de gás natural, começando em 2.1 por uma breve análise do mercado no Brasil abordando os dados de oferta e demanda nacional, principais fontes de produção e consumo assim como alguns dados adicionais sobre a infra-estrutura nacional. Em seguida, será feita uma breve abordagem do mercado mundial de gás natural e GNL no mundo, e algumas semelhanças com o mercado Brasileiro. A seção 2.3 falará brevemente sobre o PLD, Preço de liquidação das diferenças, como a variável explicativa que será utilizado no modelo de previsão de importação de GNL. A seção seguinte falará sobre as térmicas a gás natural no Brasil. A seção 2.5 pretende mostrar os terminais de GNL no Brasil, suas principais características e capacidades assim como novos projetos e capacidades firmes para os próximos anos. Por fim, no momento da redação deste trabalho atravessa-se um processo de mudança estrutural no mercado de gás brasileiro e decidiu-se abordar brevemente as principais determinações governamentais, os principais objetivos pretendidos pelo governo e partes interessadas.

2.1

Mercado de Gás Natural no Brasil

A oferta de gás natural ao mercado brasileiro, é composta majoritariamente pela oferta de gás nacional associado à produção de petróleo, seguida da importação via gasoduto da Bolívia, e complementada por GNL regaseificado a fim de fechar o balanço de oferta e demanda nacional. Como pode ser visto na Figura 2, 76% do volume ofertado, ou 56,0 milhões m³/dia em janeiro de 2019, foi de origem nacional.

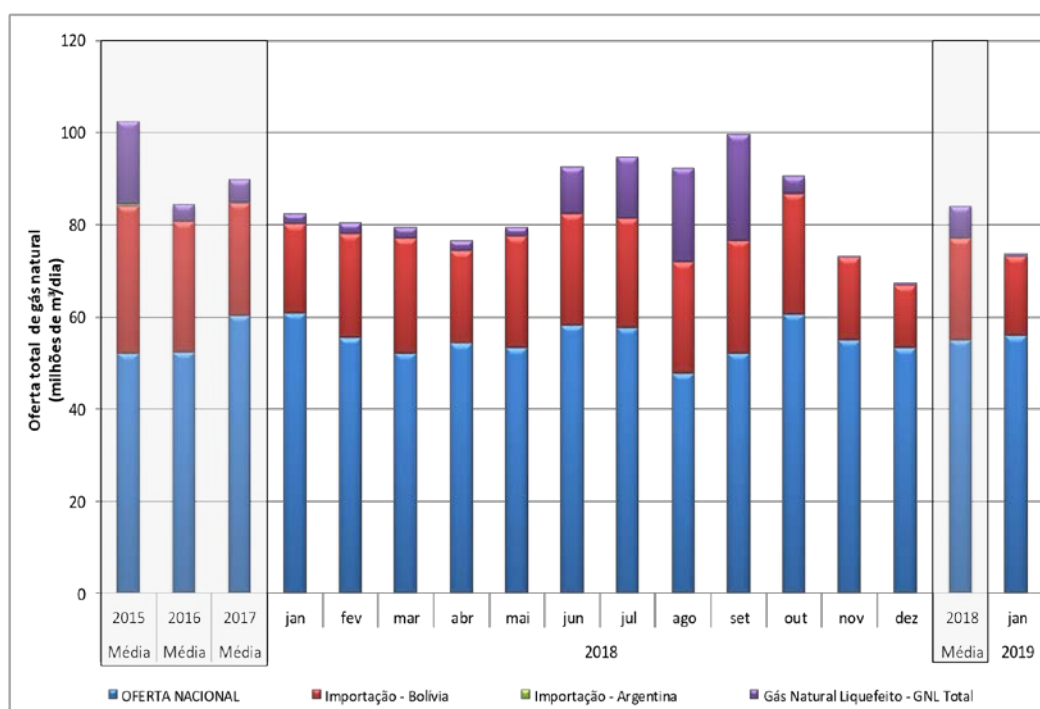


Figura 2: Oferta de gás natural ao mercado nacional.
Fonte: M.M.E (2019).

O volume de gás nacional que chegou ao mercado, em janeiro de 2019, representou 49,5% da produção nacional de 113 milhões de m³/dia, pois as outras parcelas da produção, tais como absorção em UPGNs, queima e perda, consumo nas unidades de exploração e produção - E&P e reinjeção são consumidas ao longo da cadeia como pode ser visto na Figura 3:

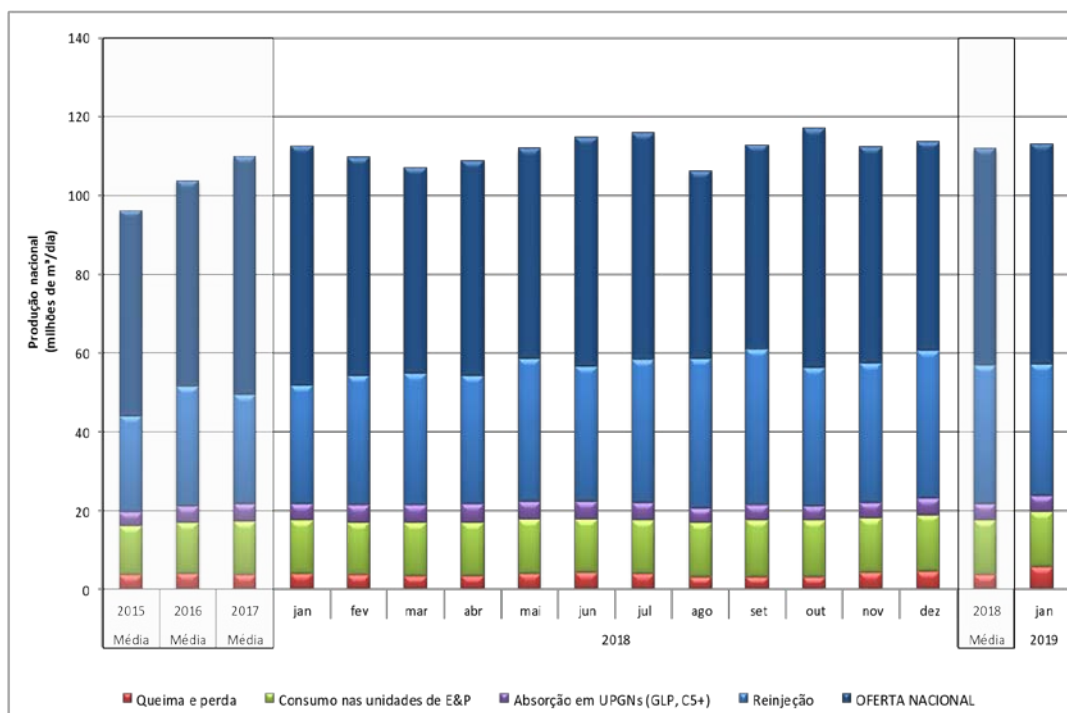


Figura 3: Segmentação da produção nacional.
Fonte: M.M.E (2019).

A produção de gás nacional tem grande parcela de origem em campos marítimos, e representou em janeiro de 2019, 82% da produção nacional. Fato relevante é a característica de gás associado à produção de óleo e representou em janeiro de 2019, 81,8% da produção nacional.

A projeção do plano decenal de energia de 2017, para o horizonte de 2018 a 2027, ou simplesmente PDE 2017, prevê o aumento dos volumes de produção, saltando do patamar de 2,6 milhões de barris por dia em 2017 para 5,1 milhões por dia em 2027, como mostra a Figura 4.

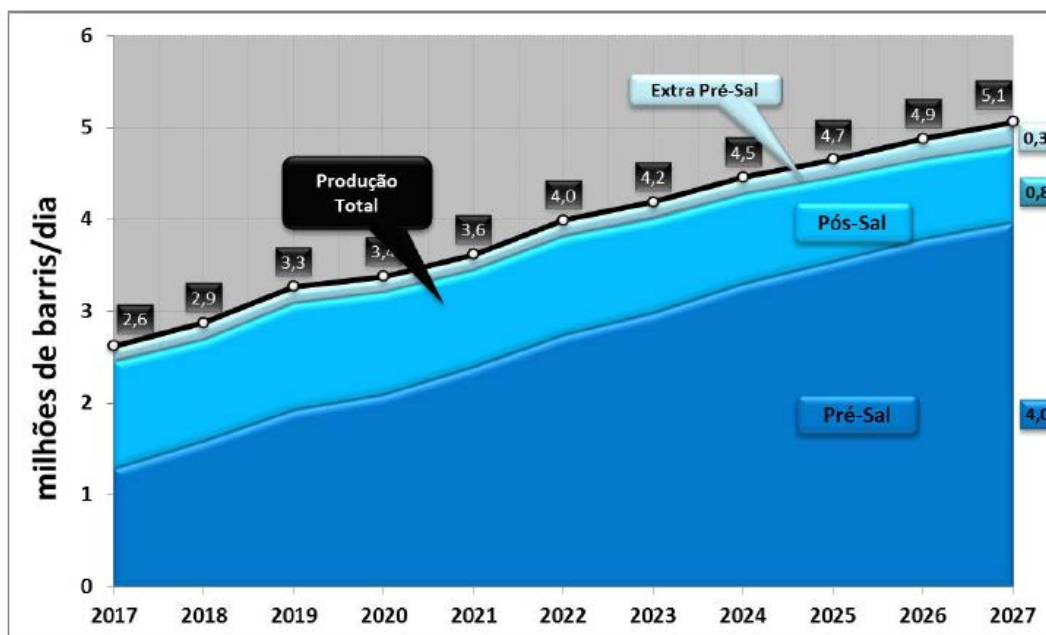


Figura 4: Previsão de Produção de Petróleo Nacional.
Fonte: PDE (2017).

Recente publicação do plano indicativo de processamento e escoamento (PIPE 2019), da EPE prevê que as produções bruta e líquida de gás natural, podem mais que duplicar no horizonte de 2030, conforme infográfico da Figura 5, abaixo:

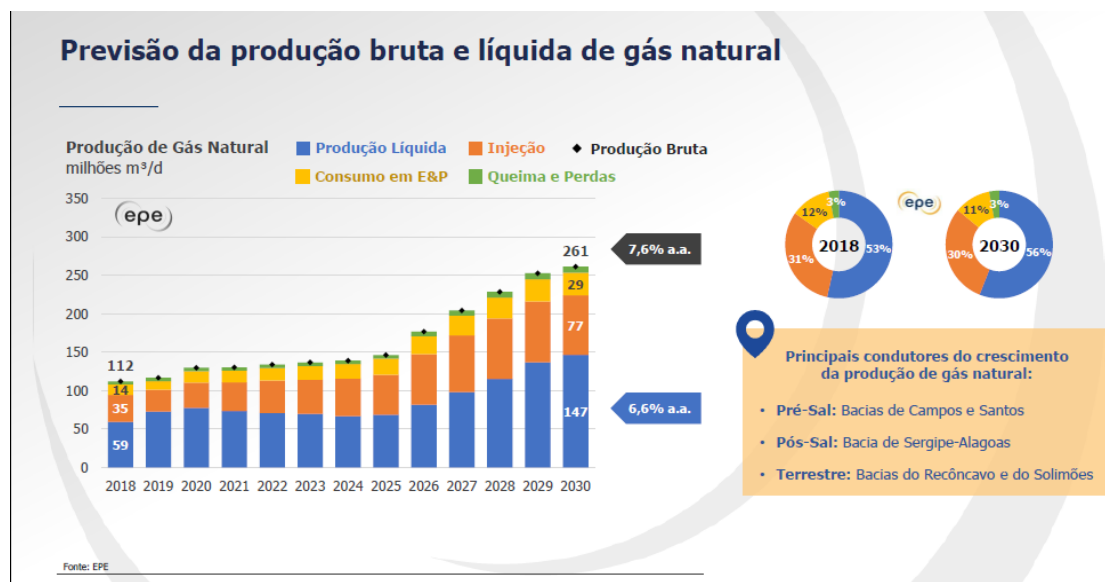


Figura 5: Previsão produção bruta e líquida de gás natural.
Fonte: EPE (2019).

Importante salientar que existem grandes incertezas quanto ao potencial de aumento de produção de gás nos próximos anos, principalmente devido aos elevados níveis de reinjeção que podem estar associados à produção de gás natural no pré-sal. Estimativas da EPE indicam que o teor de CO₂ no gás do pré-sal pode alcançar até 80%, a depender da localização do campo, e nessas condições as tecnologias de remoção apresentam alto custo e podem inviabilizar técnica-economicamente o aproveitamento comercial do gás de alguns campos do pré-sal.

Outro relevante desafio para o gás do pré-sal está relacionado à distância da costa e não menos importante às profundidades dos campos de produção. As estimativas dos estudos dos produtores e da EPE, indicam que 95% do gás está localizado a mais de 150 quilômetros da costa. Essas condições implicam em elevados custos de escoamento e carecem de infraestrutura para que se viabilize.

Atualmente a infraestrutura de gasodutos de escoamento existente e em construção refere-se às rotas 1, 2 e 3. A capacidade de escoamento deve chegar à 44MMm³/dia em 2020. O gasoduto rota 2 do pré-sal entrou em operação em 19 de janeiro de 2016 e possibilitou a injeção, naquela ocasião, de capacidade adicional de 13 MMm³/dia (MME, 2016). A Figura 6 mostra a representação das rotas existentes, assim como a de alguns estudos que não serão abordados aqui. Para mais informações sobre os dados e estudos da EPE para a infraestrutura de escoamento e processamento consultar o Plano Indicativo de Processamento e escoamento de Gás Natural, PIPE 2019 da EPE.

Principais gasodutos de escoamento em construção e em operação nos campos do pré-sal

Em operação:

Rota 1:
Bacia de Santos – Caraguatatuba/SP
Capacidade: 10 Mm³/d
(possibilidade de ampliação)

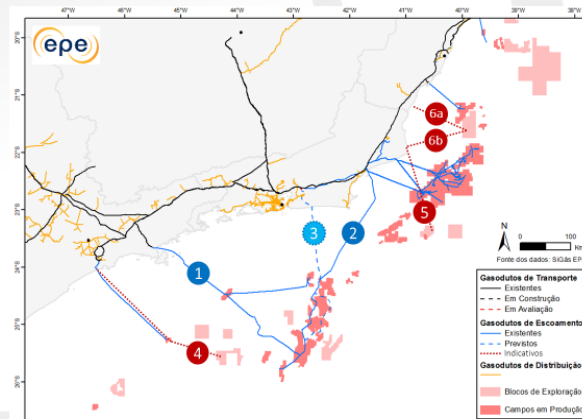
Rota 2:
Bacia de Santos – Cabiuñas/RJ
Capacidade: 16 Mm³/d
(autorizado para 20 Mm³/d)

Em construção:

Rota 3:
Bacia de Santos – Itaboraí/RJ
Capacidade: 18 Mm³/d
Previsão de conclusão: 2020
Investimento: R\$ 6 bilhões

Capacidade de escoamento em operação e em construção

44 Mm³/d



Empresa de Pesquisa Energética
Ministério de Minas e Energia



Figura 6: Rotas de Escoamento do Pré-sal.

Fonte: EPE (2019).

Com relação à importação de gás natural, o histórico demonstra que ela ocorreu basicamente por duas fontes, a saber: O gás natural boliviano, importado via gasoduto Bolívia-Brasil o GASBOL, e através dos três terminais de regaseificação da Petrobras, importado via cargas de GNL em navios metaneiros. Dado o objeto deste trabalho, um melhor detalhamento dos terminais de GNL será mostrado na seção do mercado de GNL, seção 2.1 deste trabalho. A Figura 7 mostra as ofertas de gás importado via gasoduto boliviano e por GNL regaseificado desde 2015.

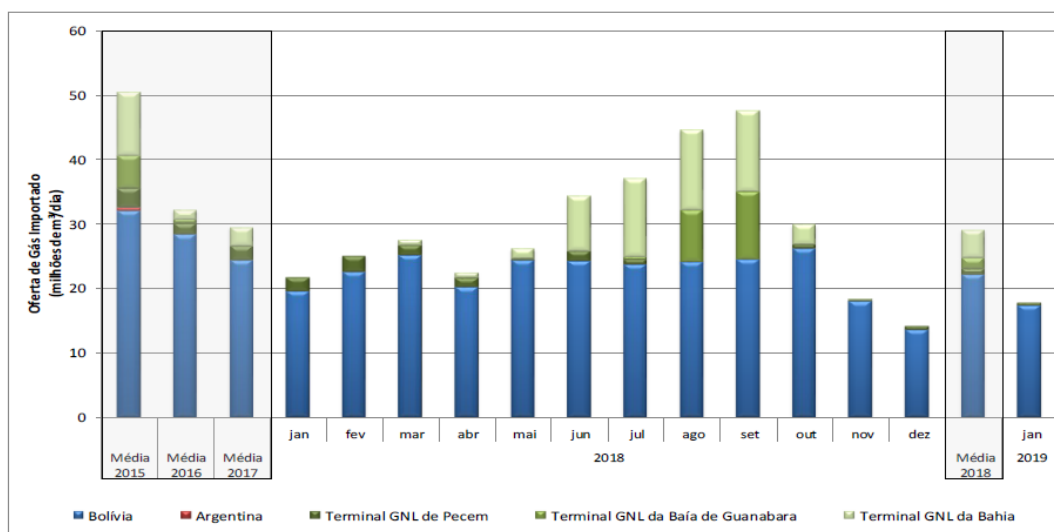


Figura 7: Importação Bolívia e volumes regaseificados.

Fonte: M.M.E. (2019).

Diferentemente do que ocorre na importação por gasoduto, onde o volume importado é considerado como oferta, na importação de GNL o volume importado não corresponde diretamente ao volume ofertado. No caso do GNL, é necessário considerar a possibilidade de armazenamento de parte da carga no navio regaseificador. Os volumes mostrados na Figura 7, de oferta de gás importado, referem-se aos volumes efetivamente regaseificados e injetados na malha de transporte para o período de referência.

Via de regra o gás nacional é prioritária, seguida do gás boliviano, dado os compromissos contratuais e menores preços históricos, seguida do GNL flexível, de forma a fechar ou complementar o balanço de oferta e demanda. Observamos também na Figura 7 a grande flutuação dos volumes mensais ofertados de GNL. O menor volume mensal de 2018 tendeu a zero e foi de 0,25 MMm³/dia em novembro e o maior volume do mesmo ano foi de 23,19 MM/m³ em setembro. Majoritariamente a variação explica-se pela demanda flexível de gás natural para o fim de geração de eletricidade nas termelétricas a gás natural.

A demanda de gás natural das distribuidoras é usualmente relacionada aos seguimentos de consumo, industrial, geração elétrica, automotivo, cogeração, residencial e comercial. Na Figura 8 pode-se observar que as maiores demandas estão nos segmentos industrial, de natureza mais estável, e do segmento de geração termelétrica com natureza mais flexível. No ano de 2018 a demanda média do seguimento industrial representou o volume médio mensal de 39,49 MMm³/dia e para a geração de eletricidade representou o volume mensal de 20,13 MMm³/dia. Apenas as duas fontes de consumo para o ano de 2018 representaram 84% do volume total consumido.

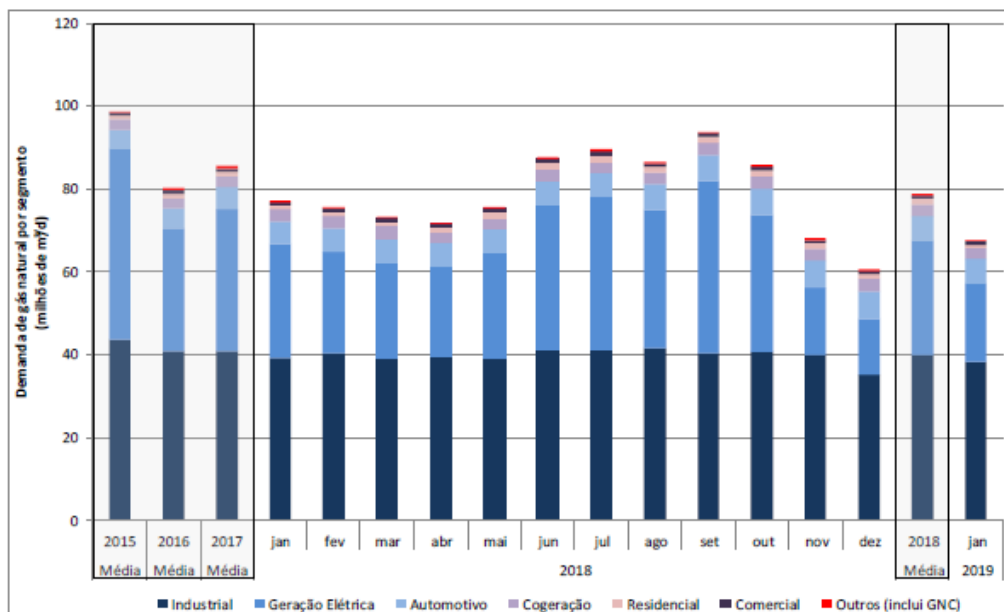


Figura 8: Demanda segmentada de gás natural.
Fonte: M.M.E. (2019).

A importância da capacidade instalada das térmicas a gás natural e outras fontes renováveis tem ganhado relevância ao longo dos anos com relação à complementaridade hidráulica e essa tendência tende a se ampliar para os próximos anos. A Figura 9 mostra a evolução da relevância da termoeletricidade na composição da geração de eletricidade brasileira nos últimos anos.

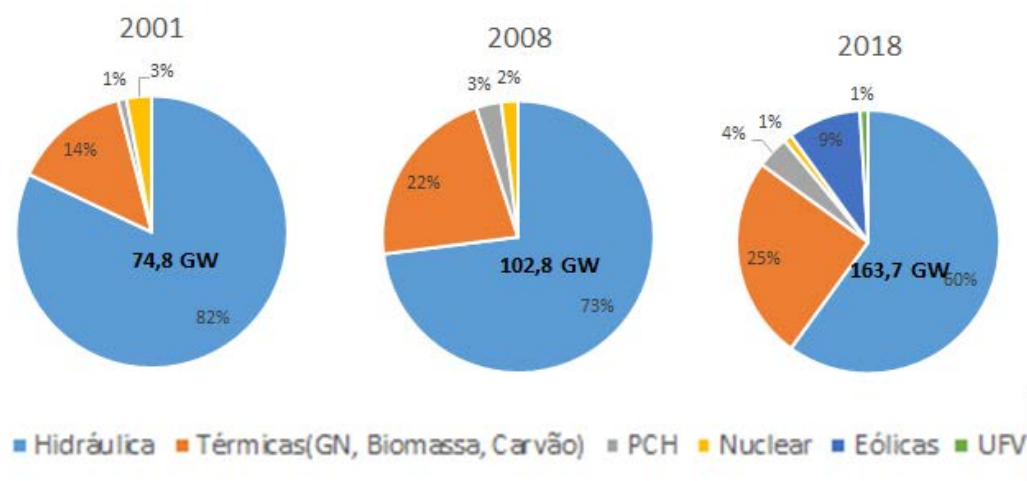


Figura 9: Evolução da Matriz de Geração de Eletricidade Brasileira.
Fonte: Adaptado de EPE (2019).

Com relação à infraestrutura de produção e movimentação de gás natural, em 2018, a malha dutoviária de transporte e distribuição correspondia respectivamente a 9.409 km (MME, 2019) e 34.649 km (Abegas, 2019). A

agência nacional de petróleo, gás natural e bicomcombustíveis, ANP, publicou em seu sítio eletrônico a lista de gasodutos de escoamento da produção (254 dutos, totalizando 4.650 km) e de gasodutos de transferência (5 dutos, totalizando 30 km).

Segundo o anuário estatístico de 2018 da ANP, as unidades de processamento de gás natural, ou UPGNS's, necessárias ao processamento e adequação do gás natural, assim como as suas capacidades e os volumes processados, são relacionadas na Figura 10.

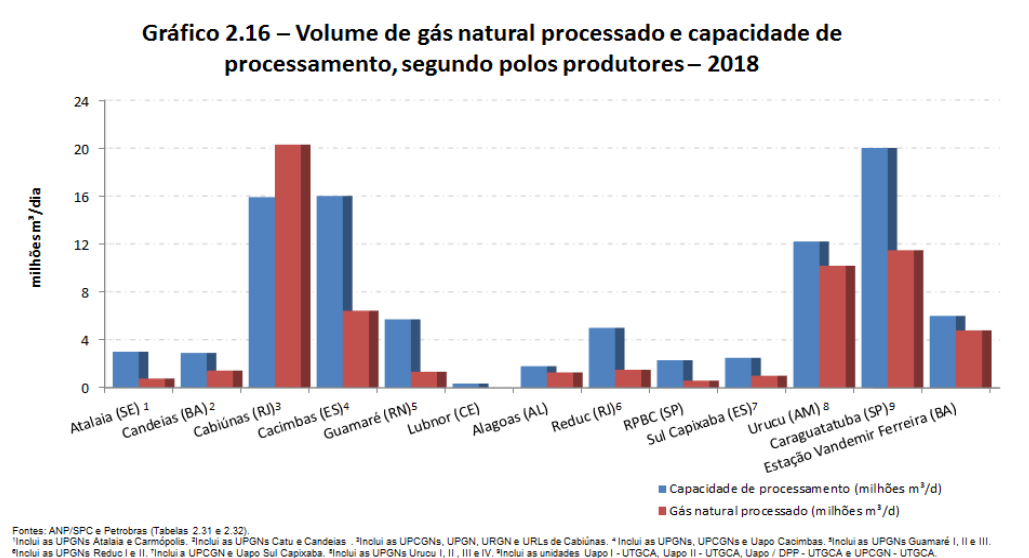


Figura 10: Capacidade e Processamento UPGNs em MMm3/dia em 2018.

Fonte: ANP (2019).

A capacidade de processamento total representou 93,65 MMm3/dia e volume processado foi de 61,07 MMm3/dia.

Foi mostrado até aqui um panorama da oferta e demanda no Brasil e a importância do gás natural assim como o perfil de produção e consumo. Uma breve visão da infraestrutura também foi citada. Nas próximas seções será abordado breve visão do mercado de GNL no mundo e no Brasil e das técnicas a gás natural, um dos principais destinos deste insumo.

2.2

O Mercado de Gás Natural no Mundo e o GNL

Em 2016 o consumo mundial de gás natural foi de 3,65 TCM (IEA, 2017), e no cenário mais provável da agência americana, a demanda alcançará 4,17 TCM em 2025 e 5,30 TCM em 2040. Em 2016 a maior demanda mundial do insumo foi realizada pelo mercado americano, seguida pela Europa e Eurásia. Ao final do período os EUA continuarão a ser o país mais demandante, seguido de China, Europa e Eurásia.

Apesar da tendência de crescimento da demanda em todos os mercados, esta tendência é mais expressiva nos países em desenvolvimento. Segundo as projeções do “World Energy Outlook, WEO (2017)” da agência americana de energia de 2017, IEA, a China quadruplicará a demanda de 200 BCM em 2016 para 800 BCM em 2040.

Tal qual o mercado Brasileiro os mercados mundiais também concentram a maior parte da demanda do gás natural, nos seguimentos Industrial e geração de energia. A média da demanda mundial entre os anos de 2000 e 2016 foi de 55 % para a geração de energia e 17% para a indústria (IEA, 2017).

O expressivo crescimento da oferta americana desde 2006 ocorreu a partir do shale gas, e de maior importador mundial de gás natural os EUA tendem a ser o maior exportador líquido até 2025 (IEA, 2017) e não há precedentes em tal crescimento na oferta no mercado mundial do energético.

O comércio mundial de gás natural ocorreu pelos modais de GNL e por gasodutos trans-nacionais, resumidamente os volumes comercializados em 2018 totalizaram 943,4 bcm, dos quais 431 bcm foram através do modal GNL. A taxa de crescimento do modal GNL em 2018 foi de 9,4% frente a apenas 0,4% do modal dutoviário. Isso mostra uma tendência de crescimento deste modal. Os volumes transacionados globalmente nesta modalidade vem crescendo ano-a-ano como mostra a Figura 11.

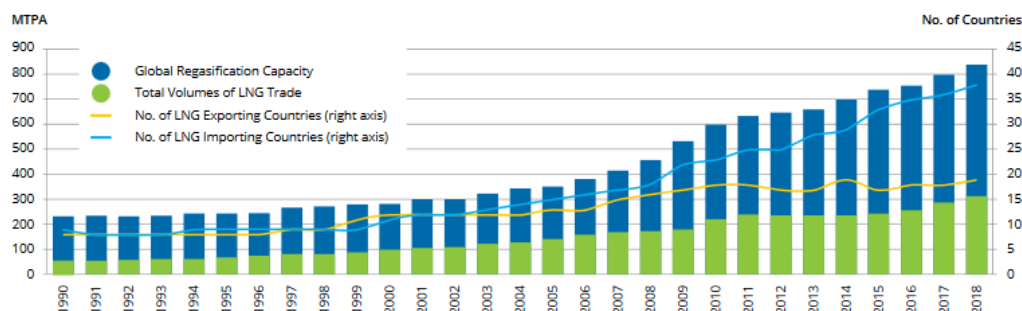


Figura 11: Volumes de GNL comercializados 1990-2018.
Fonte: IGU (2019).

A capacidade de exportação (liquefação) instalada global em 2018 representou 393MTPA (IGU, 2019) e o Qatar por mais de uma década permanece como maior exportador de GNL mundial (78,7 MTPA) respondendo por $\frac{1}{4}$ do volume comercializado, seguido por Austrália e Malásia. O quarto lugar em 2018 passou a ser ocupado pelo EUA. O volume total comercializado foi de 316,5 MTPA e a distribuição dos exportadores, assim como a sua participação no mercado podem ser vistos na Figura 12.

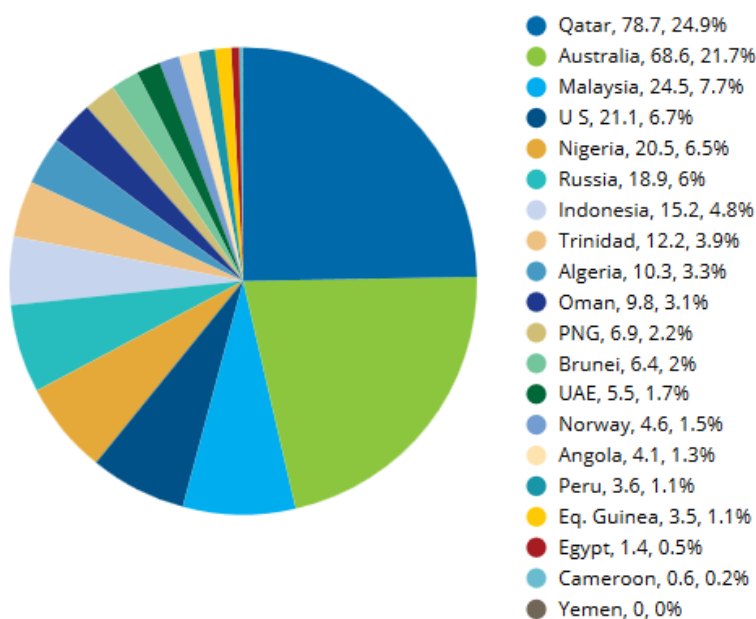


Figura 12: Volumes Exportados de GNL e fatia do mercado em 2018.
Fonte: IGU (2019).

Pelo lado da demanda, a capacidade instalada mundial de regaseificação do GNL em 2018 foi de 824 MTPA, dos quais 80,1 corresponderam a embarcações flutuantes de regaseificação e armazenamento de GNL, ou os chamados FSRU (Floating, Regaseification, Storage, Unit). Vale ressaltar que no

Brasil os três terminais existentes operam com este tipo de filosofia de regaseificação flutuante e a soma total da capacidade dos três terminais é de 47 MMm3/dia ou 13 MTPA (Fonte: MME). Maior detalhamento será dado na seção 2.5. A Figura 13 mostra os países importadores e a fatia de mercado por eles representado em 2018. Vale salientar que os países que importaram menos que 2 MTPA foram agregados em “outros”, dentre eles o Brasil, que importou em 2018, segundo dados do MME, 1,92 MTPA.

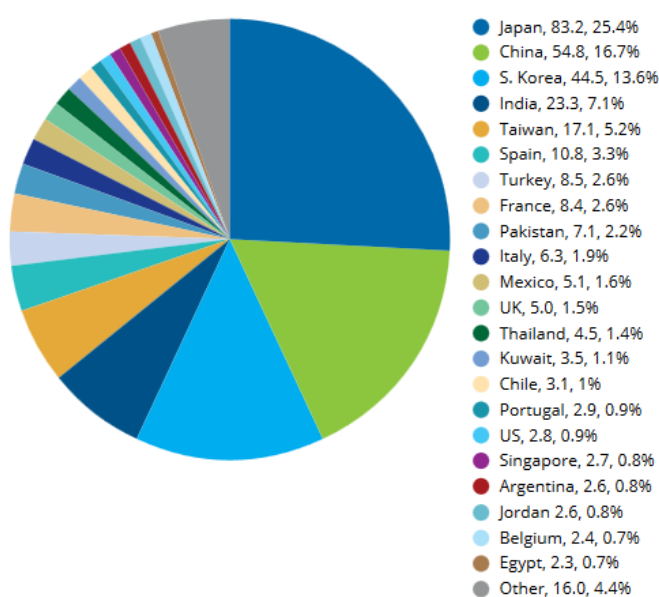


Figura 13: Volumes Importados de GNL e fatia do mercado em 2018.
Fonte: IHS, IGU (2019).

Nota: Outros – Polônia, Holanda, Brasil, Malásia, República Dominicana, EAU, Grécia, Bangladesh, Lituânia, Israel, Canada, Malta, Jamaica e Colômbia.

Em 2018 havia 19 países exportadores e 37 importadores (IGU, 2019), os principais pontos de movimentação e comercialização tendem a formar preços para as relações comerciais. Para o gás natural não há uma única referência de preço como o óleo combustível, ao contrário, há uma variedade de preços regionais, cada um com suas características, que gradualmente se tornam mais interligadas, devido ao movimento global de interação entre os mercados. O aumento do comércio global de GNL, devido à sua natural flexibilidade busca sempre a destinação mais vantajosa e conseqüentemente a movimentação entre os mercados.

A trajetória dos preços Norte americano (Henry-Hub), vem exercendo papel fundamental na formação de preços dos mercados globais. O Henry Hub é um ponto de entroncamento de gasodutos na Luisiana, EUA, onde o preço do gás é formado pela competição com as mais diversas fontes, ou “gas-to-gas competition”, i.e. é o preço que equilibra a oferta e demanda regional. A projeção dos preços nesse hub refletem uma tendência de aumento em sua oferta de shale gas e redução dos preços na sua produção nos EUA (IEA, 2017).

A trajetória dos preços, nos principais mercados é mostrada desde 2010 até janeiro de 2019, na Figura 14:

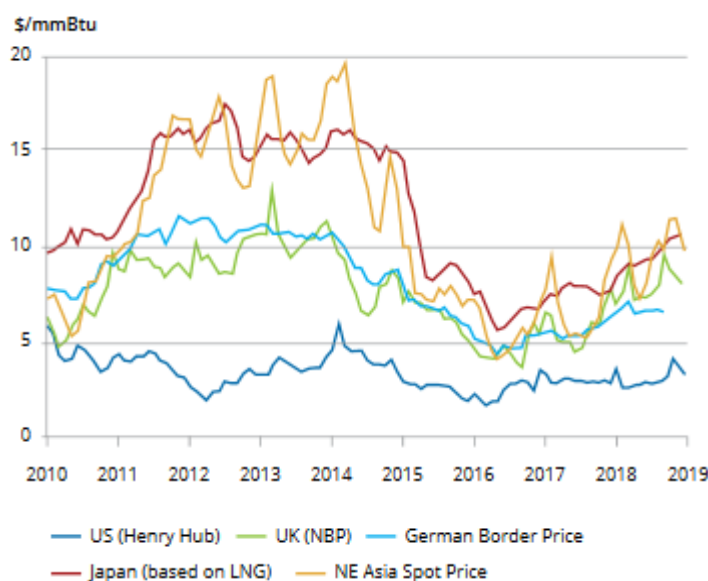


Figura 14: Preços de referência.
Fonte: IGU (2019).

Na avaliação da utilização do gás natural, os consumidores geralmente pesam as alternativas disponíveis e seus custos relativos ao gás. As considerações a depender da utilização são distintas, mas sem dúvida o preço tende a ser relevante. Vários países ao longo do tempo, principalmente onde havia abundância do energético subsidiaram os preços a patamares de \$1 a \$2/Mbtu e sem nenhuma surpresa a demanda subiu drasticamente em todos os setores. Entretanto países como Japão, Korea e China experimentaram aumento da demanda apesar de preços de \$10 - \$16/Mbtu nos últimos anos. Deste modo, não há uma única resposta para o aumento da procura pelo gás natural.

A história de liberalização dos mercados de gás da Alemanha, Reino Unido e EUA, nos mostra que nas últimas três décadas de dados, e após análise

estatística do padrão de comportamento a distintos níveis de preços, há uma tendência de aumento da demanda a preços inferiores ao patamar de \$6/Mbtu e decréscimo quando está acima de \$8/Mbtu (IEA, 2017).

Por fim os custos típicos, da cadeia produtiva do GNL estão expostos na Tabela 1, e observa-se que o custo de produção e liquefação tendem a ser os mais expressivos da cadeia:

Tabela 1: Preços da Cadeia Produtiva de GNL em US\$ / MMBTU.

	GNL EUA	GNL Nigéria
Custo de Produção	2,71	5,84 - 7,56
Liquefação	3	
Custo de Afretamento	0,76 - 1,42	Incluso
Margem do Supridor	1	Incluso
Regaseificação	0,8 - 2,5	0,8 - 2,5
GNL Regaseificado	8,3 - 10,6	6,64 - 10,06

Fonte: IEA (2017)

A intenção foi apresentar um panorama resumido e apontar ao longo do trabalho, algumas fontes e referências para a obtenção de dados e informações mais detalhadas. Os temas de contratos, formação de preços, tecnologias, equipamentos e tendências, dentre outros, não serão abordados por não serem escopo do trabalho.

2.3

Preço de Liquidação das Diferenças (PLD)

Em função da preponderância de usinas hidrelétricas no parque de geração brasileiro, são utilizados modelos matemáticos para o cálculo do PLD, que têm por objetivo encontrar a solução ótima de equilíbrio entre o benefício presente do uso da água e o benefício futuro de seu armazenamento, medido em termos da economia esperada dos combustíveis das usinas termelétricas (Fonte: ONS).

Por exemplo, se ocorrer expressivo despacho hidrelétrico, antes de período seco aumenta-se o risco de despachar térmicas de custo mais elevado no futuro, por outro lado, um despacho térmico anterior a um período úmido poderá provocar vertimentos, incorrendo em desperdício de energia. A Figura 15 representa a árvore de decisão deste problema operacional.

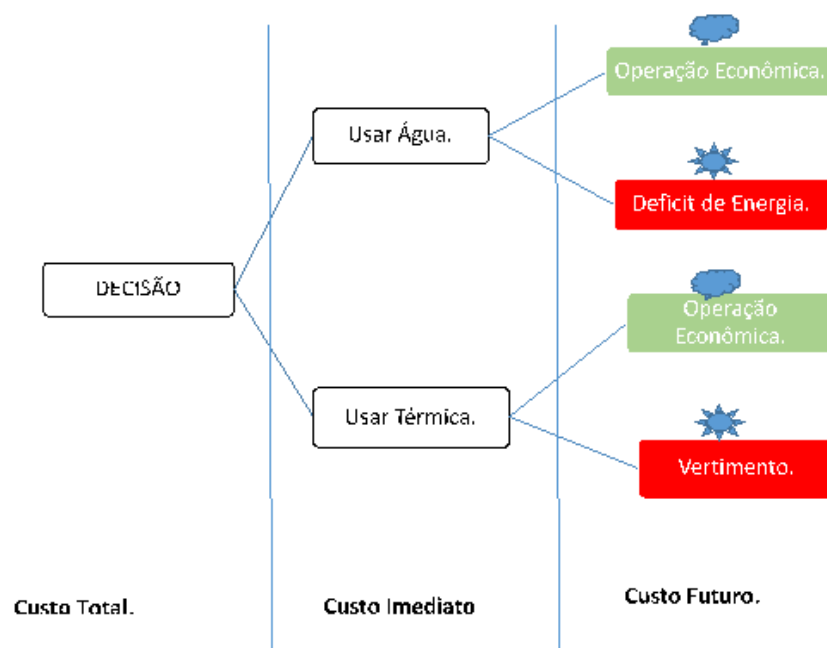


Figura 15: Problema de Decisão da Operação.
Adaptada da Fonte: ONS (2012).

A máxima utilização da energia hidrelétrica disponível em cada período é a premissa mais econômica, do ponto de vista imediato, pois minimiza os custos de combustível. No entanto, essa premissa resulta em maiores riscos de déficits futuros. Por sua vez, a máxima confiabilidade de fornecimento é obtida conservando o nível dos reservatórios o mais elevado possível, o que significa utilizar mais geração térmica e, portanto, aumento dos custos de operação. A Figura 16 representa o critério de planejamento da operação.

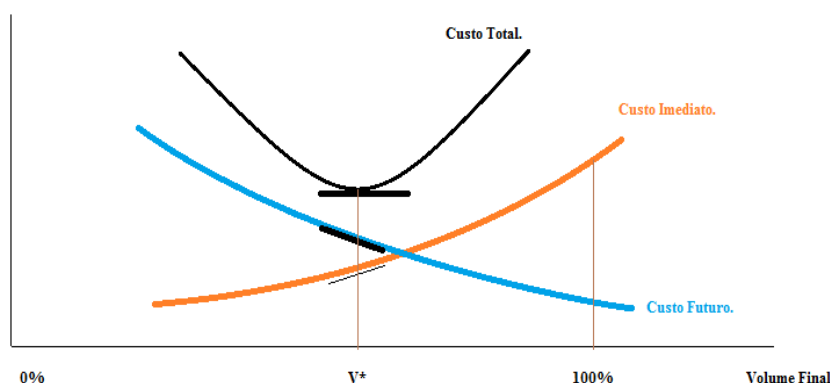


Figura 16: Critério de planejamento da operação.
Adaptada da Fonte: ONS (2012).

Com base nas condições hidrológicas, na demanda de energia, nos preços de combustível, no custo de déficit, na entrada de novos projetos e na disponibilidade de equipamentos de geração e transmissão, o modelo de precificação obtém o despacho (geração) ótimo para o período em estudo, definindo a geração hidráulica e a geração térmica para cada submercado. Como resultado desse processo são obtidos os Custos Marginais de Operação (CMO) para o período estudado, para cada patamar de carga e para cada submercado.

Em razão do atributo físico da energia elétrica necessitar de equilíbrio instantâneo entre demanda e oferta, a oferta prevista *ex ante* não necessariamente se iguala à demanda observada, fazendo-se necessário o equilíbrio instantâneo entre demanda e oferta em dois pontos: no fornecimento e na contabilização financeira. Com relação ao fornecimento, o ONS, responsável pelo planejamento da operação do sistema, centraliza o despacho das usinas geradoras por meio da agregação dos empreendimentos de geração e transmissão, e busca minimizar o custo da energia.

A contabilização financeira é realizada pela CCEE (Câmara de Comercialização de Energia), a qual contabiliza as diferenças entre o que foi consumido ou produzido e o que foi contratado. As diferenças positivas ou negativas são liquidadas no mercado de curto prazo (*spot*), valoradas ao preço de liquidação das diferenças (PLD), que é determinado semanalmente para cada patamar de carga e para cada submercado, tendo como base o CMO do subsistema. O PLD é limitado por um preço mínimo e um preço máximo. Os

intervalos de duração de cada patamar são determinados para cada mês de apuração pelo ONS e informados à CCEE.

Na CCEE são utilizados os mesmos modelos adotados pelo ONS para determinação da programação e despacho de geração do sistema, com as adaptações necessárias para refletir as condições de formação de preços.

No cálculo do PLD não são consideradas as restrições de transmissão internas a cada submercado e as usinas em testes, de forma que a energia comercializada seja tratada como igualmente disponível em todos os seus pontos de consumo e que, conseqüentemente, o preço seja único dentro de cada uma dessas regiões. No cálculo do preço são consideradas apenas as restrições de transmissão de energia entre os submercados (limites de intercâmbios).

O processo completo de cálculo do PLD - Preço de Liquidação das Diferenças consiste na utilização dos modelos computacionais NEWAVE e DECOMP, os quais produzem como resultado o Custo Marginal de Operação de cada submercado, respectivamente em base mensal e semanal.

Limites mínimo e máximo do PLD e Custo do déficit de energia

O PLD é limitado por valores mínimo e máximo de acordo com a Resolução Normativa ANEEL nº 633/14, com validade entre a primeira e a última semana operativa de preços do ano.

Já o custo do déficit de energia é entendido como o valor que se pode atribuir à insuficiência estrutural da oferta de energia elétrica, e é parâmetro de grande importância nos estudos de planejamento da expansão e programação da operação dos sistemas elétricos, bem como para o cálculo do PLD.

A divulgação do PLD é realizada semanalmente pela CCEE através de seu sítio como mostra a Tabela 2, e os valores mínimos e máximos vigentes nos três últimos anos são mostrados na Tabela 2.

Tabela 2: Preços Vigentes PLD

	PLD mínimo (R\$/ MWh)	PLD máximo (R\$/ MWh)	Custo do Déficit (R\$/ MWh)
2017	33,68	533,82	4650,00
2018	40,16	505,18	4596,31
2019	42,35	513,89	4944,89

Fonte: CCEE (2019).

2.4

Térmicas a Gás natural no Brasil

O parque termelétrico do país em operação é composto por 3.018 usinas e 42,36 GW de capacidade instalada e representa 24,49 % dos 166,69 GW da capacidade Total. (ANEEL, 2019).

O Parque termelétrico a gás natural é composto por 36 complexos de usinas, sendo 15 bicom bustíveis. A capacidade instalada representa 12,63 GW (Fonte: MME). Porém os dados de geração mostram que a geração das térmicas a gás natural representa em média 50% da geração total das térmicas do país. A Figura 17 mostra a geração total, a geração de energia de todas as térmicas e compara com a geração das térmicas a gás natural no período de janeiro de 2009 a dezembro de 2018.

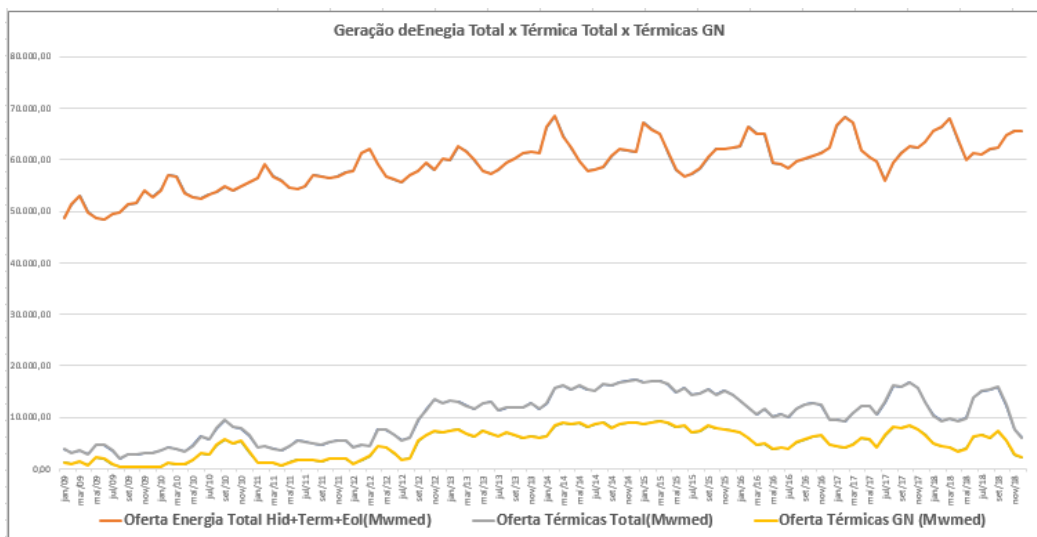


Figura 17: Geração Térmica GN x Geração Térmica Total x Geração Total. Adapta da Fonte: ONS (2019).

Adapta

A Figura 18, mostra desde abril de 2016 a influência da Energia Natural Afluente, ENA, sobre o preço da energia elétrica. Quando a ENA está abaixo da MLT (Média de Longo Termo), o Preço da Energia representado pelo CMO, tende a se elevar. E o contrário também é verdade, quando a ENA está maior que a MLT, o CMO tende a baixar.

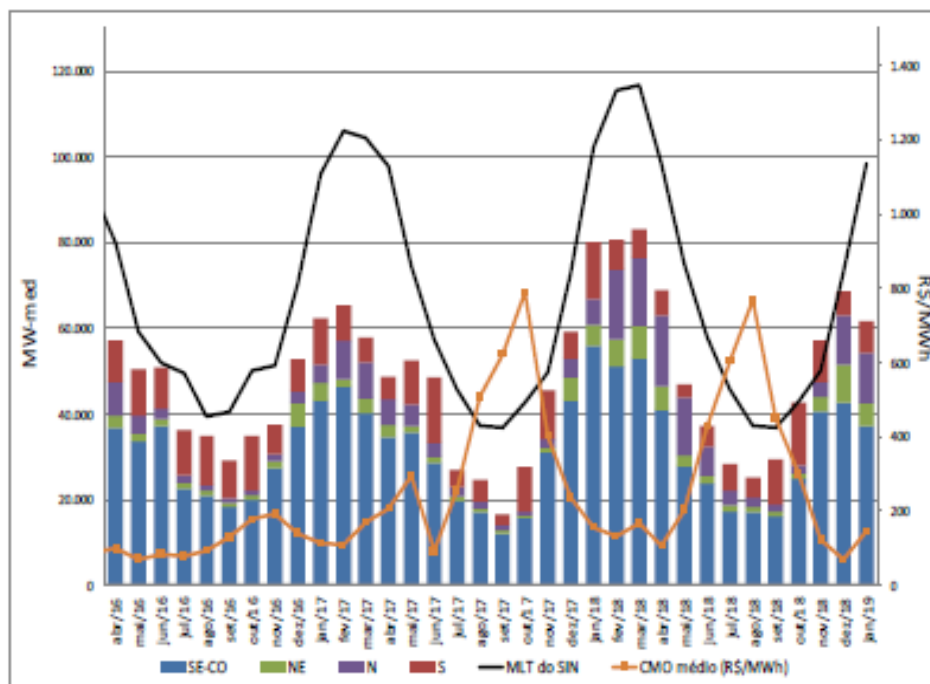


Figura 18: ENA x MLT x CMO. Fonte: ONS (2019).

Foi mostrado nas seções anteriores que a importância da termoeletricidade para o mercado de gás natural tem ganhado espaço ao longo dos anos e tende a se tornar mais expressiva para os próximos anos. Dentre outras razões há a característica de ser uma fonte menos poluente se a compararmos com outras fontes. Diversas referências internacionais indicam um movimento de transição energética para uma matriz mundial com emissões reduzidas em menor participação relativa de combustíveis fósseis. O papel do gás natural como possível combustível de transição se dá, principalmente, pela existência de infraestrutura já construída e amortizada em diversos países, além do baixo custo de adaptação das instalações industriais que utilizam fontes mais poluentes, como o óleo combustível, no Brasil, e o carvão, no restante do mundo (MIT, 2011).

Especificamente quanto à utilização como combustível para usinas termelétricas ou simplesmente UTEs, vale destacar que atualmente o gás natural desempenha um papel relevante de complementação para sazonalidades e intermitências no setor elétrico, e tem papel relevante como combustível de transição entre fontes mais poluentes e energias renováveis. A Figura 19 compara as emissões dos principais poluentes, e pode-se observar os particulados finos e os óxidos de enxofre são praticamente nulos nas emissões do gás natural. Observa-se que as emissões dos óxidos de nitrogênio NO e NO₂ (NO_x), são significativamente menores comparados às outras fontes. Vale ressaltar que algumas tecnologias de controle de emissões quando aplicadas podem reduzir ainda mais as emissões de Nox, como os controles pós combustão e o Dry Low Nox ou queima estagiada em câmaras de combustão de turbinas a gás. Importante que o NO_x do gás natural responde por menos de 10% das emissões globais que estão atribuídas em sua grande maioria ao óleo diesel no setor de transporte rodoviário.

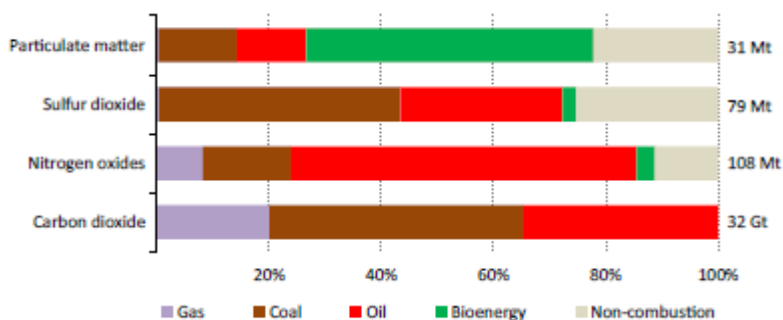


Figura 19: Emissões por fontes de geração.
Fonte: IEA(2017).

Políticas adequadas, preços atrativos e regulação entre outros temas, tendem a reduzir as emissões globalmente quando há a substituição por fontes menos poluentes. Exemplo ilustrado na Figura 20, mostra a redução de 25% nas emissões de CO₂ nos EUA através da substituição do carvão por gás natural na geração termelétrica entre os anos de 2006 e 2016.

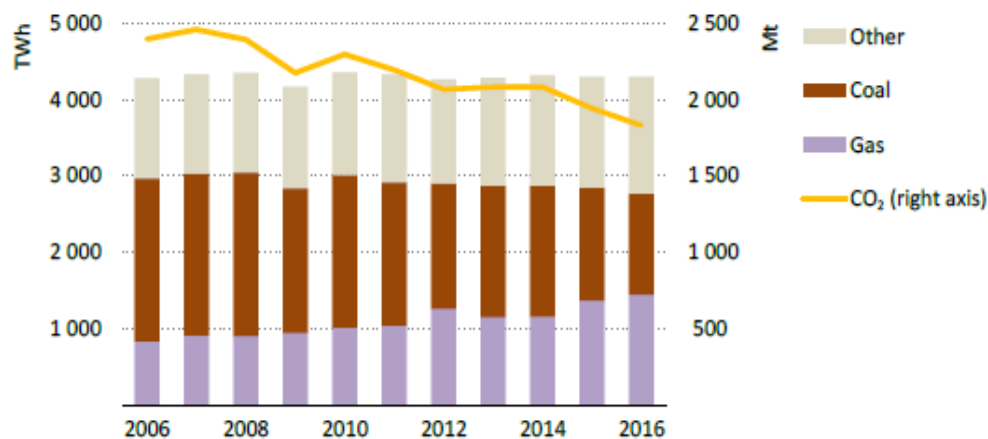


Figura 20: Emissões de CO₂ por fonte no setor de eletricidade nos EUA.
Fonte: IEA(2017).

A substituição gradativa do carvão por gás natural e a tendência de se buscar, tanto ciclos de operação mais eficientes e novas tecnologias disponíveis fazem com que o gás natural tenha importante papel na transição para outras fontes de energias mais limpas.

2.5

Terminais de importação e Regaseificação de GNL no Brasil

A importação de GNL teve início em janeiro 2009 no terminal do Pecém, no estado do Ceará. Com capacidade de regaseificação de 7MMm/dia através de FSRU, abastece basicamente gás natural para as termelétricas da região. Em abril do mesmo ano foi inaugurado o terminal da Baía da Guanabara com capacidade de regaseificação de 20MMm³/dia também opera com navio regaseificador. O último terminal inaugurado no país foi o terminal da Bahia de todos os Santos em Salvador, inaugurado em janeiro de 2014, também com capacidade de regaseificação de 20 MMm³/dia opera com navio regaseificador do tipo FSRU. Todos os terminais citados pertencem a Petrobras e são operados pela Transpetro. Os navios regaseificadores são afretados para o período contratual. Dada a

flexibilidade de movimentação dos navios e a natureza flexível da demanda, há dois navios em operação atualmente que atendem aos três terminais e resumidamente as capacidades tanto dos terminais como dos navios estão na Tabela 3:

Tabela 3: Terminais de GNL e capacidades de tancagem e regaseificação.

	Terminal de Pecém	Terminal da Baía da Guanabara	Terminal da Bahia
Capacidade de Regaseificação do Terminal (MMm ³ /dia)	7	20	20,00
Navio Regaseificador	Experience		Golar Winter
Capacidade de Regaseificação do Navio(MMm ³ /dia)	22,65		14,16
Capacidade de Armazenamento do Navio(m ³ de GNL)	173.000		138.000

Fonte: MME (2019).

A filosofia de operação é a importação através de navio carrier ou carregador que uma vez atracado ao terminal transfere a carga de GNL para o tanque do navio FSRU. A operação de regaseificação ocorre através da injeção do gás natural liquefeito já regaseificado na malha de gasodutos. A Figura 21 mostra a configuração típica da referida operação



Figura 21: Vista aérea do terminal de GNL do Pecém - Ceará.
Fonte: Petrobras (2019).

Atualmente há diversos projetos em estudo de novos terminais no Brasil, pelos mais distintos atores. A tendência é que a iniciativa privada assuma aos poucos a importação de GNL no Brasil. Os dados e estudos mais recentes podem ser obtidos no sítio da EPE. Vale a citação de dois projetos firmes de importação de GNL, já em execução, não só pela semelhança na filosofia de operação, mas também pelo modelo de negócio dedicado à operação termelétrica. Ambos os terminais, de Açu, no Rio de Janeiro e o de Sergipe tem em princípio o modelo “LNG to Power” o que significa dizer GNL exclusivamente dedicado às térmicas do projeto específico. Ambos os projetos a princípio, não tem conexão com a malha de gasodutos e são dedicados à exclusiva geração termelétrica.

O projeto Porto do Açu no Rio de Janeiro é previsto para entrar em operação em 2021. Basicamente é um FSRU afretado da BW com capacidade de tancagem de GNL de 174.000 m³, regas de 21 MMm³/dia, e suprirá uma central termelétrica de 1,3 GW de capacidade. É um consórcio formado pela Prumo, EIG, Siemens e BP (EPE, 2019).

O projeto de Sergipe, CELSE está previsto para entrar em operação em 2020, basicamente é um FSRU com capacidade de tancagem de GNL de 170.000 m³, e regas de 21 MMm³/dia dedicado a uma central termelétrica de 1,5 GW de capacidade (EPE, 2019).

3

Metodologia

O problema a ser tratado foi exposto dentro do contexto de importação de GNL para atendimento da demanda de gás natural flexível para as térmicas a gás natural no Brasil.

Neste capítulo, será apresentada uma revisão da literatura sobre previsão de demanda de importação de GNL a fim de atender às térmicas a gás natural, com o propósito de identificar os métodos já utilizados para previsão e as possíveis variáveis que explicam a referida demanda. Em seguida será apresentada breve descrição e formulação matemática dos métodos de previsão selecionados e finalmente as métricas para avaliação do desempenho dos modelos utilizados

3.1

Revisão da Bibliográfica

Esta revisão bibliográfica pretende identificar os principais métodos de previsão de demanda de gás natural liquefeito e identificar algumas variáveis explicativas utilizadas nos modelos causais. Como referência para a pesquisa na literatura, foi utilizado o artigo de Maçaira et al (2018) que fez uma revisão sistemática da literatura a fim de identificar os principais métodos de séries temporais que incorporam variáveis explicativas. Os métodos aplicados com maior frequência foram os modelos de regressão, seguido das redes neurais artificiais (ANN), modelos ARIMAX ou ARIMA com a incorporação da variável explicativa, SVM ou máquina de vetor de suporte e por fim os modelos estruturais. Alguns dos passos adotados serviram como orientação e serão abordados ao longo da pesquisa.

A base de dados selecionada foi a *Scopus web* por possibilitar a pesquisa de artigos das principais revistas de diversas áreas de conhecimento científico (Mongeon e Paul-Hus, 2016).

A consulta à base de dados utilizou filtros a fim de retornar apenas resultados em inglês e classificados como artigos e *reviews*.

A busca inicial por palavras-chave foi iniciada e algumas combinações de palavras foram propostas a fim de buscar artigos relevantes para a pesquisa. Alguns critérios de exclusão foram aplicados imediatamente que são:

- (i) Artigos que não tratassem especificamente da previsão de gás natural e gás natural liquefeito para fins de geração;
- (ii) Artigos que não tratassem do tema de séries temporais.

Na primeira rodada foram utilizadas as palavras-chave “ *liquefied natural gas* ” e “ *time series*” e retornou 9 artigos, dos quais sete foram eliminados por estarem relacionados à simulações e modelos de monitoramento de falha que não fazem referência aos métodos de previsões buscados.

No artigo de Cao et al. (2015), foram utilizados modelos ARIMA, a fim de prever a oferta dos principais supridores de gás para a China. E o artigo de Chae et al (2012) utiliza análise causal para atestar que o preço do gás natural influencia unidirecionalmente os preços da energia elétrica nos mercados de curto e longo prazo.

Na segunda rodada foram utilizadas as palavras-chave “ *LNG* ” e “ *time series*” e apenas um artigo foi selecionado. A maioria dos artigos tratava de assuntos como monitoramento e comparação de séries históricas para fins de simulação e análise de falhas em equipamentos, análises de investimento, dentre outros.

O estudo selecionado trata da previsão da demanda de GNL para suprir a co-geração de energia elétrica e calor durante 14 anos, e foi realizado utilizando redes neurais por Chung, et al. (2014). São utilizados modelos univariados ARIMA para prever a taxa de crescimento da população e da demanda energética no período analisado.

Com o objetivo de se buscar estudos causais, para o problema de previsão de GNL, aplicou-se a utilização das palavras chaves utilizadas por Maçaira et al (2018). Foram encontrados dois artigos como mostra a Tabela 4, e um dos estudos foi descartado por não estar relacionado a gás natural ou ao GNL.

Tabela 4: Números de artigos por palavra chave “Ing”.

Palavra-Chave	N ºArtigos
Ing AND "time series" AND "exogenous* variable"	2
Ing AND "time series" AND "explanat* variable"	0
Ing AND "time series" AND "input variable"	0
Ing AND "time series" AND "input variable"	0
Ing AND "time series" AND "predict variable"	0
Ing AND "time series" AND "dependent variable"	0
Ing AND "time series" AND "explicative variable"	0

O artigo de Ferkingstad (2011), aplicou a modelagem VAR (vetor auto-regressivo) e modelo multivariado de séries temporais de preços de energéticos nos mercados alemão e nórdico. Foi estimado um modelo causal para estimar a dinâmica dos preços em ambos os mercados e verificou-se que os mercados de energia destas duas regiões estão relacionados aos preços do gás natural. As variáveis exógenas utilizadas foram os níveis dos reservatórios nórdicos e a produção de energia eólica no mercado alemão.

Na rodada seguinte foi utilizada a palavra “natural gas” combinada com as mesmas palavras de busca anteriores. De forma condensada os resultados encontram-se na Tabela 5 e serão resumidos nos parágrafos a seguir.

Tabela 5: Números de artigos por palavra chave “natural gas”.

Palavra-Chave	N ºArtigos
"natural gas" AND "time series" AND "exogenous* variable"	2
"natural gas" AND "time series" AND "explanat* variable"	1
"natural gas" AND "time series" AND "input variable"	0
"natural gas" AND "time series" AND "input variable"	0
"natural gas" AND "time series" AND "predict variable"	0
"natural gas" AND "time series" AND "dependent variable"	0
"natural gas" AND "time series" AND "causal*model*"	0
"natural gas" AND "time series" AND "explicative variable"	0

Chen et al (2018) propôs um modelo auto-regressivo funcional utilizando temperaturas, vazões e nomações diárias, como variáveis exógenas (FARX), a fim fazer previsões de curto-prazo. Os resultados mostraram que a defasagem de 1 lagg da variável dependente é importante previsor e algumas variáveis explicativas como a temperatura podem se tornar insignificantes. A

Laib et al (2019), a fim de prever a demanda de gás natural para a geração de energia na Argélia utiliza redes neurais para as previsões não-lineares e monitora o comportamento da demanda, ajustando a demanda conforme modelos benchmark selecionados, dentre eles modelo de séries temporais sazonais com variáveis exógenas (SARIMAX). A variável considerada foi a temperatura diária.

Balitskiy, et al (2014), avalia a relação entre o crescimento do consumo de gás natural e o crescimento econômico na comunidade europeia. Foi criado um modelo multivariado que inclui como variáveis explanatórias, a formação bruta de capital fixo, e a força de trabalho. Utilizando testes de cointegração foi encontrada relação de longo-prazo com crescimento econômico, consumo de gás natural, trabalho e capital.

Com base nestes critérios de exclusão foram selecionados 7 artigos para a leitura e como última etapa foram inseridos mais 3 artigos considerados relevantes para a pesquisa, por tratarem de previsões de preços e demanda, utilizando modelos auto-regressivos onde o gás natural ou o GNL são relevantes.

O artigo de Aras e Aras (2004), estuda o consumo de gás natural residencial na Turquia. É proposta a divisão das previsões em um período quente e um não-quente, assim como modelos auto-regressivos distintos para os dois períodos. É investigado o comportamento dinâmico da demanda de gás e a temperatura ao longo dos dias. O autor verifica que os erros diminuíram com relação a um modelo de previsão anual único.

Uma modelagem utilizando ARIMAX para a previsão dos preços contratados de energia de longo prazo no mercado da Alemanha foi realizado por Flasz et al (2011). A modelagem foi feita utilizando variáveis com alta correlação, tais como os de preços de gás, carvão e petróleo.

Hönig (2019), a fim de realizar estudo técnico e econômico de plantas de liquefação de pequena escala no EUA utilizou modelos estatísticos ARIMA para prever os preços de GNL, energia e emissões.

A fim de condensar as modelagens utilizadas para os dez artigos é apresentado resumo na Figura 22.

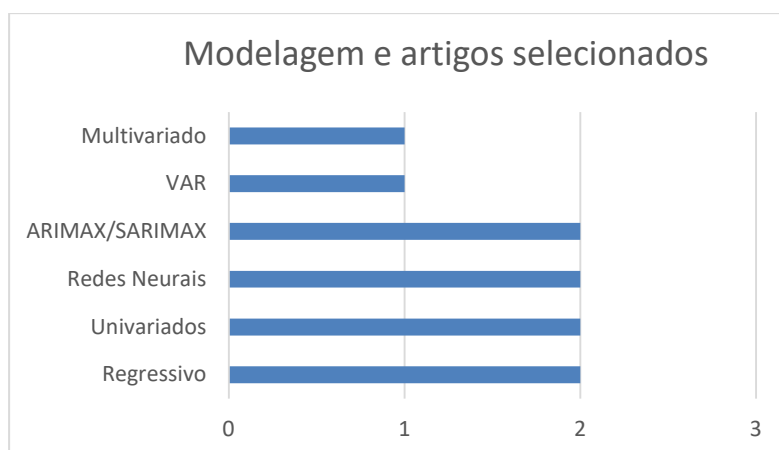


Figura 22: Número dos métodos em artigos selecionados.
Fonte: Compilação Própria.

3.2

Métodos de Previsão

Uma série temporal é um conjunto de observações de uma dada variável, ordenado segundo o parâmetro tempo, geralmente em intervalos equidistantes.

As técnicas de modelagem das séries temporais podem ser divididos essencialmente em univariadas e multivariadas (Maçaira et al, 2018).

Os modelos univariados utilizam apenas valores passados da própria série para se estimar os valores futuros. Existem diversos modelos tradicionais na literatura, tais como os modelos Box & Jenkins (Box & Jenkins, 1976) , e os modelos de amortecimento exponencial (Gardner Jr, 2006). Os modelos multivariados utilizam a influência de outras variáveis com capacidade explicativa no comportamento da série de saída para obter melhores resultados na representação da função de transferência.

Os modelos de amortecimento exponencial e os modelos Box & Jenkins são as duas abordagens mais utilizadas para previsão de séries temporais (Hyndman e Athanasopoulos, 2016).

Os modelos de regressão e ARIMA com incorporação de variável explicativa (ARIMAX), estão entre os métodos multivariados mais amplamente utilizados na literatura, como mostrou Maçaira et al (2018) que aplicando um rigoroso protocolo de revisão sistemática da literatura identificou os principais métodos de séries temporais que incorporam variáveis explicativas.

Serão apresentadas nesta seção breve descrição e a formulação dos modelos univariados de amortecimento exponencial, modelo Holt e Box & Jenkins, assim como os modelos causais de regressão dinâmica e função de transferência Box & Jenkins.

3.2.1

Amortecimento exponencial

As previsões produzidas usando métodos de amortecimento exponencial são geradas com base em médias ponderadas de observações passadas, com os pesos decaindo exponencialmente à medida que as observações se afastam do tempo que está sendo previsto. Por isso, quanto mais recente a observação, maior o peso associado, ou maior o impacto da observação para a previsão, enquanto que observações mais antigas possuem um peso menor. Essa estrutura gera previsões confiáveis rapidamente e para um amplo espectro de séries temporais, o que é uma grande vantagem e de grande importância para aplicações práticas.

A seleção do método é geralmente baseada no reconhecimento das componentes de tendência e sazonalidade da série temporal em estudo e como elas entram no método de amortecimento (de forma aditiva ou multiplicativa).

Para a série estudada foi aplicado a modelagem de Holt para séries sem sazonalidade.

A equação de Holt com tendência linear e sem amortecimento pode ser mostrada a seguir.

$$\begin{aligned}
 Z_T(\tau) &= \hat{a}_1(T) + \tau \hat{a}_2(T) \\
 \hat{a}_1(T) &= \alpha Z_T + (1 - \alpha)[\hat{a}_1(T - 1) + \hat{a}_2(T - 1)] \\
 \hat{a}_2(T) &= \beta[\hat{a}_1(T) - \hat{a}_1(T - 1)] + (1 - \beta)\hat{a}_2(T - 1) \quad (1)
 \end{aligned}$$

Onde:

$Z_T(\tau)$ é a previsão da série τ passos a frente de um instante genérico T , e α e β são hiperparâmetros do modelo. $\hat{a}_1(T)$ e $\hat{a}_2(T)$ são os estimadores do modelo.

A equação de Holt com tendência linear e com amortecimento pode ser mostrada a seguir:

$$\begin{aligned} Z_T(\tau) &= \hat{a}_1(T) + \sum_{i=1}^{\tau} \phi^i \hat{a}_2(T) \\ \hat{a}_1(T) &= \alpha Z_T + (1 - \alpha)[\hat{a}_1(T-1) + \phi \hat{a}_2(T-1)] \\ \hat{a}_2(T) &= \beta[\hat{a}_1(T) - \hat{a}_1(T-1)] + (1 - \beta) \phi \hat{a}_2(T-1) \end{aligned} \quad (2)$$

Onde

$\phi = \text{constante de amortecimento da tendência.}$

Os parâmetros $Z_T(\tau)$, $\hat{a}_1(T)$ e $\hat{a}_2(T)$ foram definidos na equação (1) acima.

A equação de Holt com tendência multiplicativa e com amortecimento pode ser mostrada a seguir:

$$\begin{aligned} Z_T(\tau) &= \hat{a}_1(T) \hat{a}_2 \hat{a}_2^{\sum_{i=1}^{\tau} \phi^i(T)} \\ \hat{a}_1(T) &= \alpha Z_T + (1 - \alpha)[\hat{a}_1(T-1) \hat{a}_2^{\phi}(T-1)] \\ \hat{a}_2(T) &= \beta[\hat{a}_1(T) / \hat{a}_1(T-1)] + (1 - \beta) \hat{a}_2^{\phi}(T-1) \end{aligned} \quad (3)$$

3.2.2

Modelagem auto-regressiva integrada de média móvel (ARIMA)

Os modelos ARIMA, também conhecidos como modelos de Box & Jenkins (2016) ou modelos auto regressivos integrado de médias móveis, fornecem outra abordagem para a previsão de séries temporais. Enquanto os modelos de amortecimento exponencial são baseados em uma descrição de tendência e sazonalidade nos dados, os modelos ARIMA visam descrever as autocorrelações nos dados.

Os modelos ARIMA têm como pressuposto que o processo a ser modelado é estacionário de segunda ordem, ou seja, a média e variância são constantes ao longo do tempo, e a correlação depende somente da defasagem entre as observações.

Se baseiam na estrutura da Função de Autocorrelação (FAC) e da Função de Autocorrelação Parcial (FACP) da série temporal aqui representada por Z_t , nas quais são identificados os parâmetros q (referente à média-móvel) e p (referente à parte auto regressiva) do modelo, respectivamente.

No modelo auto-regressivo (AR), a variável de interesse é modelada de acordo com uma combinação linear de seus valores passados e pelo ruído aleatório F_t , conforme segue abaixo:

$$Z_t = \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + \dots + \phi_p Z_{t-p} + F_t \quad (4)$$

onde os parâmetros ϕ_i são responsáveis por descrever como Z_t relaciona-se com o valor Z_{t-i} , sendo $i = 1, 2, \dots, p$.

No modelo de médias móveis (MA), ao invés de usar os valores passados da variável de previsão, utiliza-se os erros de previsão anteriores, ou seja, a série é modelada através da combinação dos ruídos F defasados.

$$Z_t = F_t + \theta_1 F_{t-1} + \theta_2 F_{t-2} + \dots + \theta_q F_{t-q} \quad (5)$$

onde os parâmetros θ_i descrevem como Z_t relaciona-se com o valor de F_{t-i} , onde $i = 1, 2, \dots, q$.

Uma série temporal estacionária é aquela cujas propriedades não dependem do tempo em que a série é observada. Assim, séries temporais com tendências, ou com sazonalidade, não são estacionárias - a tendência e a sazonalidade afetarão o valor da série temporal em tempos diferentes. Por outro lado, uma série de ruído branco é estacionária - não importa quando você a observa, ela deve ter a mesma aparência em qualquer momento (Hyndman e Athanasopoulos, 2016).

Quando a série temporal em estudo não é estacionária, primeiramente é necessário integrá-la até que a mesma se torne estacionária. Para tal, a série é transformada através de diferenças sucessivas. O número de diferenças necessárias para tornar a série de dados estacionários representa o valor do parâmetro de integração m . A primeira diferença pode ser representada por:

$$\nabla Z_t = Z_t - Z_{t-1}$$

onde $\nabla = (1 - B)$ é o operador diferença.

Surgem então os modelos ARIMA (p,d,q), que assumem a seguinte forma:

$$\phi(B)\nabla^d Z_t = \theta(B)F_t \quad (6)$$

Para assegurar a condição de estacionariedade de um conjunto de dados é possível utilizar-se de diversas ferramentas estatísticas, tais quais a estatística Q de Box-Pierce, Estatística de Ljung-Box e o teste Aumentado de Dick-Fuller (ADF).

Os processos encontrados na prática, além de raramente serem estacionários, apresentam muitas vezes componentes sazonais. Os modelos sazonais são conhecidos como SARIMA, ou ARIMA multiplicativo e têm uma formulação semelhante à dos modelos ARIMA:

$$\Phi(B^p)\phi(B)\nabla_p^D\nabla^d Z_t = \Theta(B^p)\theta(B)F_t \quad (7)$$

onde Φ e Θ são os parâmetros referentes à sazonalidade, e D o parâmetro que representa o total de diferenças sazonais necessárias para tornar a série estacionária na sazonalidade.

Maiores detalhes sobre os modelos ARIMA podem ser encontrados em Box & Jenkins (1976).

3.2.3

Modelos de Regressão Dinâmica

A regressão dinâmica é um modelo de regressão que combina a dinâmica de séries temporais e os efeitos de variáveis explicativas e tais modelos de regressão dinâmica devem ser usados quando existe uma estrutura de dependência entre a variável de interesse e variáveis causais e, ao mesmo tempo, quando a estrutura de correlação da série dependente indicar que não se pode supor a independência dos erros (Goodrich, 1989).

O termo “dinâmico” não quer dizer que os parâmetros do modelo evoluem no tempo, e está relacionado à estrutura de defasagem da série temporal e das variáveis causais do modelo. De forma bastante geral, a ideia dos modelos de Regressão Dinâmica é de que a inserção de variáveis causais são capazes de melhorar as previsões em relação à modelagem univariada.

Os modelos de regressão dinâmica, resumidamente podem ser descritos pela equação :

$$\varphi(B)y_t = \beta x_t + F_t \quad (8)$$

onde:

y_t = variável dependente no instante t

β = vetor de coeficientes das variáveis causais, que será estimado por mínimos quadrados

x_t = vetor de variáveis causais (exógenas) no instante t

F_t = ruído aleatório associado ao modelo, onde supomos que os F_t são independentes e identicamente distribuídos com densidade $N(0, \sigma^2)$ $\varphi(B)$ = polinômio autoregressivo de ordem p , isto é:

$$\varphi(B) = 1 - \varphi_1 B - \varphi_2 B^2 - \dots - \varphi_p B^p \quad (9)$$

sendo B o operador de atraso

A estrutura do modelo de regressão dinâmica permite considerar como elementos x_t tanto variáveis causais como também suas defasagens.

A generalização do modelo de regressão dinâmica, em que se admite uma estrutura de dependência dos erros é conhecido na literatura como modelo de regressão de Cochrane e Orcutt generalizados.

O modelo de regressão generalizado de Cochrane e Orcutt (1949) é dado por:

$$\varphi(B)y_t = \beta x_t + w_t \quad (10)$$

$$R(B)w_t = F_t \quad (11)$$

Onde $R(B)$ é um polinômio auto-regressivo.

Pode-se notar que a equação 10 tem a mesma forma da equação 2 da regressão dinâmica, entretanto os erros w_t apresentam uma estrutura AR dada pela equação 11.

Este modelo, dado pelas equações 10 e 11, ainda pode ser descrito em termos de uma única equação como a seguir. Note que da equação 10:

$$w_t = \varphi(B)y_t - \beta x_t \quad (12)$$

Substituindo esta última expressão na equação 11 tem-se que:

$$R(B) \cdot (\varphi(B)y_t - \beta x_t) = F_t \quad (13)$$

Desta última expressão nota-se que o modelo de regressão generalizado de Cochrane e Orcutt (1949) introduz defasagens tanto na variável dependente (y_t) quanto nas causais. A equação 13 indica também que a relação de causalidade entre y_t e x_t não é afetada pela introdução do polinômio autoregressivo $R(B)$.

Esta última equação pode ainda ser escrita em termos de novas variáveis

y_t^* e x_t^* dadas por:

$$y_t^* = R(B) \cdot y_t$$

$$x_t^* = R(B) \cdot x_t$$

onde $R(B)$ é chamado de “fator comum” e representa a estrutura de correlação presente no erro w_t .

A equação obtida então com estas novas variáveis é:

$$\varphi(B)y_t^* = \beta x_t^* + F_t \quad (14)$$

Logo, o modelo de regressão generalizado de Cochrane e Orcutt reduz-se ao modelo de regressão dinâmica usual ao se considerar as novas variáveis y_t^* e x_t^* . O modelo original de Cochrane e Orcutt tem como fator comum $R(B) = 1 - \alpha B$ e então a equação 13 reduz-se a:

$$\varphi(B) \cdot [y_t - \alpha y_{t-1}] = \beta \cdot [x_t - \alpha x_{t-1}] + F_t \quad (15)$$

A estratégia usualmente empregada para construir um modelo de regressão dinâmica é a estratégia *bottom-up*. Basicamente, parte-se de um modelo simples, com poucas variáveis, refina-o e inclui novas variáveis até encontrar um modelo apropriado. As previsões geradas por um modelo de regressão dinâmica dependem não só de valores passados da própria série, mas também dos valores previstos para as variáveis causais. Logo, para se obter as previsões da série y_t para $t+1$, $t+2$, $t+3$, ...; é necessário fornecer ao modelo os valores futuros do vetor de variáveis causais x_t . Esta condição possibilita a criação de cenários futuros para as variáveis causais, possibilitando analisar os possíveis valores futuros da variável de interesse (what if analysis).

Na escolha do modelo apropriado, deve-se levar em consideração não apenas a significância dos parâmetros, mas também sua consistência com a teoria econômica. Como exemplo, no modelo em estudo, a variação do PLD afeta o despacho termelétrico. O aumento do PLD, tende a aumentar o despacho das térmicas a gás natural, e a redução do PLD tende a reduzir a demanda de geração das térmicas a gás natural e a importação de GNL. Portanto na escolha de um modelo de regressão dinâmica, não é necessário apenas encontrar os parâmetros que tendem a ser adequados, mas também verificar a coerência dos coeficientes estimados.

3.2.4

Modelos de Função de Transferência

Os modelos Box e Jenkins univariados permitem fazer a análise - descrição e previsão - de uma série temporal através do conhecimento exclusivo dos seus valores. A existência de fenômenos aleatórios relacionados entre si conduz ao tratamento simultâneo de séries temporais correspondentes de forma que a informação contida em cada uma possa aperfeiçoar e enriquecer o estudo desses fenômenos (Box & Jenkins, 1976). O modelo de função de transferência é um modelo estatístico que descreve a relação entre uma variável de saída y e uma ou mais variáveis de entrada x com o objetivo de encontrar a combinação linear entre duas ou mais séries através de um modelo estocástico com o menor erro médio quadrático.

Sem perda de generalidade, consideremos o caso mais simples de uma única série de entrada X_t causando a saída Y_t . Suponhamos também que a série de entrada x_t e a série de saída y_t são estacionárias, ou foram “feitas” estacionárias através da aplicação apropriada de operadores diferenciais, simples (∇^d) ou sazonal (∇_s^d).

O modelo geral relacionando x_t e y_t é:

$$y_t = v_0 + v_1 x_{t-1} + v_2 x_{t-2} + \dots + \varepsilon_t \quad (16)$$

ou em termos do operador de atraso “B”:

$$y_t = \left(\frac{v_0 + v_1 B + v_2 B^2 + \dots}{v(B)} \right) x_t + \varepsilon_t \quad (17)$$

onde os pesos V_i ; $i = 0, 1, 2, \dots$ são as respostas a função impulso que indicam como a série X_t é “transferida” ou “reflete” na série de saída.

Analisando a formulação acima, Box & Jenkins constataram os seguintes problemas:

- modelo não-parcimonioso (número excessivo de pesos v_0, v_1, v_2, \dots);
- não leva em consideração um possível defasamento no efeito da variável de entrada - independência serial da série de saída.

Para contornar tais problemas propõem uma formulação parcimoniosa, considerando o polinômio de grau infinito em B da equação (17), i. e., $v(B)$, substituído por um quociente de dois polinômios de grau finito em B conforme mostrado abaixo:

$$y_t = \left(\frac{w_0 - w_1 B - \dots - w_s B^s}{1 - \delta_1 B - \dots - \delta_r B^r} \right) x_{t-b} + \varepsilon_t \quad (18)$$

Comparando as duas formulações:

$$(v_0 + v_1 B + v_2 B^2 + \dots) = \frac{(w_0 - w_1 B - \dots - w_s B^s)}{(1 - \delta_1 B - \dots - \delta_r B^r)} B^b \quad (19)$$

Para tornar a formulação dada pela equação (18) mais geral Box & Jenkins propõem a substituição do termo de erro ε_t (supostamente um ruído branco), por

uma componente de erro " η_t " que apresenta correlação serial, sendo assim possível sua modelagem por um ARMA (p, q). Desta forma, o modelo final é dado por:

$$y_t = \left(\frac{w_0 - w_1 B - \dots - w_s B^s}{1 - \delta_1 B - \dots - \delta_r B^r} \right) x_{t-b} + \eta_t \quad (20)$$

Onde: $\phi(B)\eta_t = \theta(B)a_t$;

a_t : Ruído Branco.

Para o caso de k séries de entrada temos:

$$y_t = \frac{w_i(B)}{\delta_1(B)} x_{t-b_i} + \eta_t ; i = 1, 2, \dots, k. \quad (21)$$

Onde: $\phi(B)\eta_t = \theta(B)a_t$;

a_t : Ruído Branco.

Para a obtenção da Função de Transferência, o procedimento segue a mesma sequência do caso univariado. Aqui também a identificação do modelo, que consiste na determinação dos valores (d, r, S, b, p, q), seguirá a sequência de identificação do modelo, estimação dos hiperparâmetros, e finalmente a previsão de Y_t , utilizando cenários da variável independente X_t .

3.2.5

Métricas de Desempenho

Nesta seção serão apresentadas as métricas que serão utilizadas para comparar o desempenho do modelo de regressão dinâmica com o desempenho dos modelos univariados ARIMA e amortecimento exponencial.

A métrica MAPE, do inglês *Mean Absolute Percentage Error*, ou Erro Médio Absoluto Percentual é uma das métricas mais utilizadas na literatura de métodos de previsão, pois fornece resultado em percentual.

O MAPE é calculado pela diferença entre os valores estimados e os valores reais. É usado tanto para avaliação *in-sample* como *out-of-sample*. A métrica é dada por:

$$MAPE = \frac{1}{N} \sum_{t=1}^n \left| \frac{Z_t - \hat{Z}_t}{Z_t} \right| \times 100 \quad (22)$$

sendo Z_t o valor da série temporal no período t , \hat{Z}_t o valor estimado pelo modelo para o período t e n o tamanho da amostra.

O BIC, do inglês *Bayesian Information Criteria*, desenvolvido por Schwarz (1978), é um critério de seleção de modelos entre uma classe de modelos paramétricos com diferentes números de parâmetros.

O critério de informação recompensa a qualidade do ajuste, medida pelo erro quadrático médio s , e penaliza por complexidade, ou seja, pelo número de parâmetros n . O BIC pode ser representado pela seguinte equação:

$$BIC = sT^{n/2T} \quad (23)$$

onde T representa o tamanho da amostra.

Um modelo específico é selecionado de uma família de modelos encontrando aquele que minimiza o BIC.

O R^2 é uma das formas de avaliar a qualidade do ajuste do modelo de regressões dinâmicas é através do coeficiente de determinação. Basicamente, este coeficiente indica quanto o modelo foi capaz de explicar os dados coletados, e é dado pela expressão:

$$R^2 = \frac{SQE}{SQT} = \frac{\sum_{t=1}^n (\hat{y}_t - \bar{y})^2}{\sum_{t=1}^n (y_t - \bar{y})^2} \quad (24)$$

Onde:

y_i é o valor observado

\hat{y}_i é o valor estimado (previsão) de y_i .

\bar{y} é a média das observações

SQT é a soma dos quadrados total.

SQE é a soma dos quadrados explicada.

É a razão entre a soma de quadrados da explicados da regressão e a soma de quadrados total. O R^2 varia entre 0 e 1, indicando, em percentagem, o quanto o modelo consegue explicar os valores observados. Quanto maior o R^2 , mais explicativo é o modelo, melhor ele se ajusta à amostra.

4

Resultados

Esta etapa do trabalho tem como objetivo apresentar os resultados das aplicações das abordagens de previsão apresentadas na seção 3. Está organizado em 4 seções, nas quais serão apresentadas, respectivamente, os dados, os ajustes e os resultados nos modelos univariados, os resultados do modelo com a variável explicativa, a comparação dos resultados da aplicação desses modelos utilizando-se as métricas de desempenho e por fim, uma análise de cenários.

4.1

Dados e Ajustes

No escopo deste trabalho foi considerado a série temporal mensal dos volumes de GNL regaseificados totais (Regas Total), em base mensal, divulgados no boletim de acompanhamento de indústria do gás natural divulgados pelo Ministério de Minas e Energia. Considerou-se no estudo os dados desde o início da operação dos terminais de GNL, em Janeiro/2009 até Janeiro de 2019 (115 observações).

A série mencionada representa os dados oficiais dos volumes regaseificados mensais, efetivamente fornecidos à malha de gás natural em MMm3/dia.

A Figura 23 representa a série temporal para o Regas Total de janeiro de 2009 a janeiro de 2019.

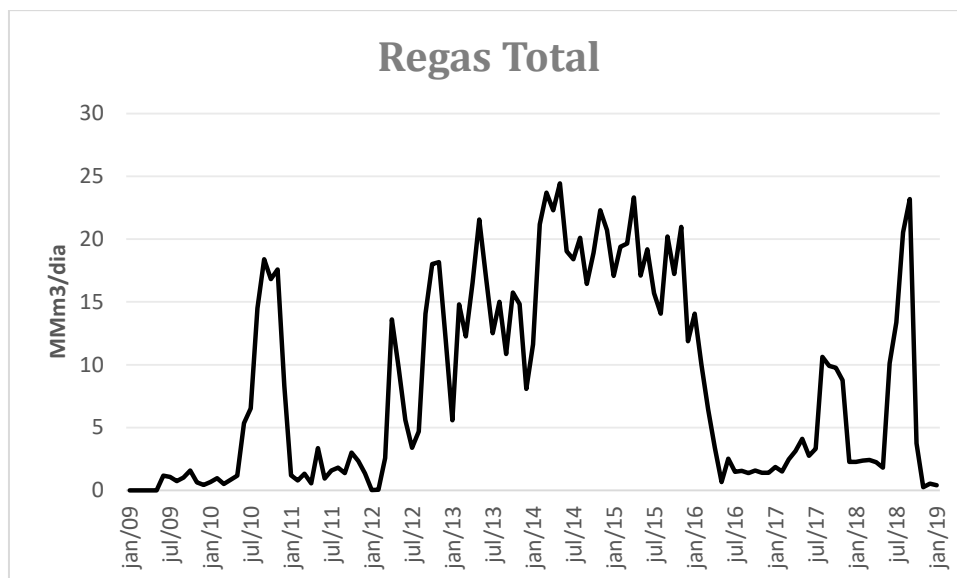


Figura 23: Gráfico da série Regas Total de janeiro de 2009 a janeiro de 2019.
Fonte: Compilação Própria.

Foi utilizado o *Forecast Pro for Windows* versão 3.50 (Stellwagen e Goodrich, 2011), para aplicação dos métodos de previsão univariados e regressão dinâmica.

A variável explicativa selecionada foi o PLD, preço de liquidação das diferenças abordada em 2.3. Os dados históricos foram coletados nos dados oficiais, públicos divulgados pela câmara de comercialização de energia, CCEE.

A variável explicativa PLD é a que melhor explica a demanda flexível por GNL importado, a fim de atender ao despacho termelétrico PLD e fechar o balanço de oferta e demanda de GN, como explicitado na seção 2. Adicionalmente foram testados modelos com as variáveis explicativas de preços internacionais de gás natural como o Henry Hub, entretanto a baixa correlação entre as variáveis apresentou resultados insatisfatórios e optou-se por não utilizá-las.

Após a normalização dos dados, tanto do Regas como do PLD, os gráficos foram traçados e a correlação mensurada. Foram feitas algumas simulações entre os PLDs por região e os volumes regaseificados por região e totalizado. A maior correlação foi observada entre o PLD do NE e o volume regaseificado total no Brasil que foi de 0,862. Entretanto observou-se comportamento adverso após janeiro de 2016 onde a correlação de ambas as variáveis modificou-se, para um patamar não esperado. A explicação para tal comportamento se deve ao fato dos novos volumes de gás natural, advindos do

pré-sal da bacia de campos e da entrada em operação, no mesmo período, do gasoduto rota 2. O gasoduto submarino de 401 km de extensão, tem capacidade para escoar 13 MMm3/dia de Santos (SP) até o terminal de tratamento de gás de Cabiúnas, o TECAB, em Macaé (RJ). A fim de ajustar a capacidade preditiva do modelo, adotou-se a mesma correlação de 0,862 para o período de janeiro de 2016 a janeiro de 2019, através de regressão linear. Uma vez encontrados os fatores α e β para o período anterior aplicou-se a seguinte formulação a fim de se estimar os novos pontos para o Regas ajustado.

$$\text{Regas ajustado} = \alpha + \beta * (\text{PLD})... \quad (25)$$

Onde:

Regas ajustado é a curva de regaseificação ajustada para a correlação de 0,862 de janeiro de 2016 a janeiro de 2019

$\alpha + \beta$ são os coeficientes da regressão.

Encontrados os novos pontos para o Regas ajustado, pôde-se traçar o gráfico de ambas as variáveis já normalizadas, que é apresentado na Figura 24. Em azul o gráfico a curva azul, do Regas totalizado sem o ajuste a partir de janeiro de 2016.

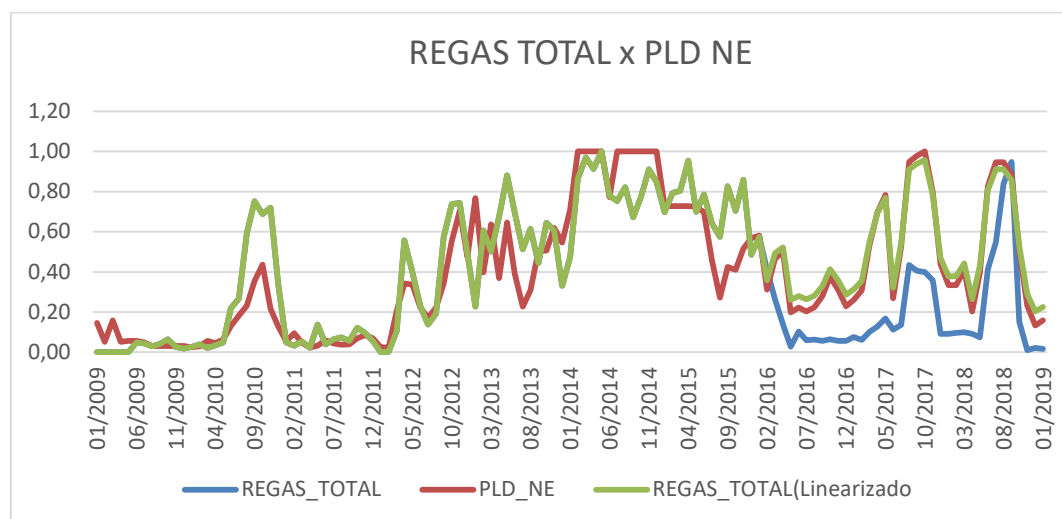


Figura 24: Volumes de Regas Totalizados x PLD NE.

Fonte: Compilação Própria.

4.2

Aplicação dos modelos univariados

Nesta seção serão apresentados os resultados obtidos a partir da aplicação dos métodos de amortecimento exponencial e modelos ARIMA.

Em especial, como visto no subitem 3.2.1, as previsões produzidas usando métodos de amortecimento exponencial são geradas com base em médias ponderadas de observações passadas, com os pesos decaindo exponencialmente à medida que as observações se afastam do tempo que está sendo previsto. A seleção do método é geralmente baseada no reconhecimento das componentes de tendência e sazonalidade da série temporal em estudo e como elas entram no método de amortecimento (de forma aditiva ou multiplicativa).

Para realizar as previsões tomando-se como método o amortecimento exponencial, foi utilizada uma abordagem a partir dos estados das componentes do método, conforme exposto subitem 3.2.1. A partir da avaliação dos métodos foi identificado um modelo sem tendência e sazonalidade aditiva, que melhor se ajustou à série temporal em estudo.

Por definição, o modelo linear ARIMA também é univariado, ou seja, vale-se da própria série temporal para realização de sua previsão.

Seguindo o procedimento de Box&Jenkins de identificação de um modelo através da visualização dos gráficos da Função de Autocorrelação e Autocorrelação Parcial, pôde-se perceber que, observando os *lags* não sazonais, verifica-se que a FAC a FACP não apresenta um corte significativos em nenhum dos lags

Identifica-se então a ordem $p = 0$ e após todos os diagnósticos e constatações, assim como realizadas as sobreposições, identificou-se um modelo ARIMA (0,1,0) como o mais adequado.

Foi verificado também se os erros gerados por este modelo não apresentam correlação, ou seja, se apresentam características de um ruído branco. Utilizou-se o teste de Ljung-Box, corroborado pelo gráfico da função de autocorrelação de erro, conforme Figura 39.

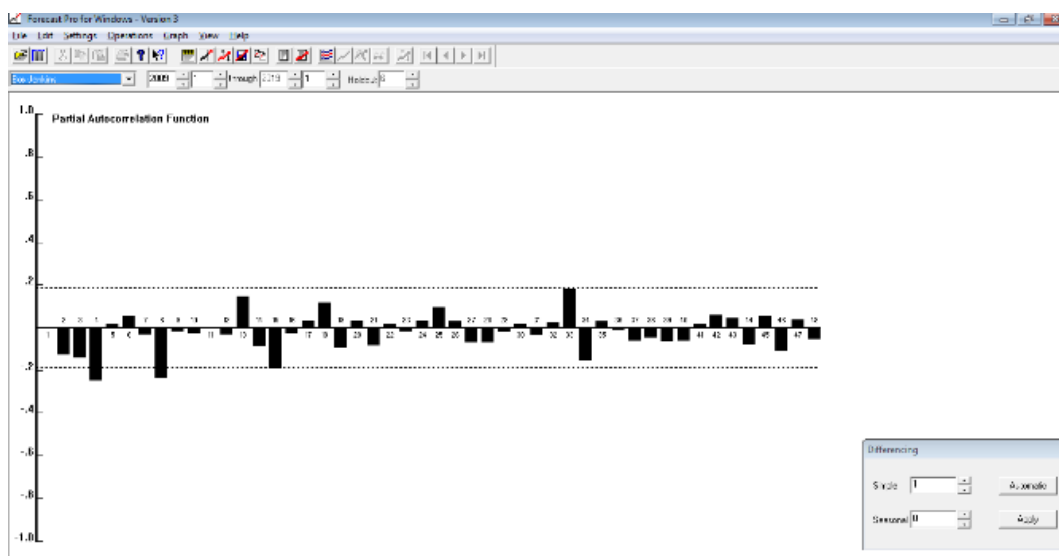


Figura 25: Função de Autocorrelação de Erros (FAC *Error*) – Modelo ARIMA(0,1,0).
Fonte: Compilação Própria.

4.3

Aplicação dos modelos causais

A Equação da Regressão Dinâmica e seus coeficientes encontrados são:

$$\text{Regas Total} = 0,57(\text{PLD}) + 0,46 \text{ RegasTotAjustado}[-1]. \quad (25)$$

O modelo de regressão encontrado para explicar a importação de GNL apresenta correlação positiva para a variável PLD. Variáveis como o preço tendem a ter relação imediata no consumo de qualquer insumo, entretanto para o caso estudado o preço Henry Hub, referência de preço do EUA muito utilizada no mercado de GNL, não apresentou bom resultado na simulação e foi descartada a sua utilização.

A variável explicativa, PLD é utilizada a fim de se determinar o despacho termo-hidráulico do próximo período operativo, como explicado na seção 2.3.

A previsão desta variável é usualmente gerada através do NEWAVE e fornecida para as áreas de estudos de uma grande empresa de petróleo a fim de tentar prever os próximos despachos e para estudos pelas mais diferentes áreas, como planejamento e marketing entre outras. A variável PLD é fornecida em patamares que vão desde 70% da Média de Longo Termo (MLT)* até 125% da MLT.

Para este estudo foi feita a avaliação “what if analysis” e foram avaliados os cenários pessimista, de menor incidência de chuvas, com PLD a 70% da MLT, o mais provável a 100% da MLT e o otimista, onde as vazões serão generosas a 125% da MLT.

Para a Função de transferência, Da equação (18), tem-se que:

$$Y_t = W_0.X_t + \phi Y_{t-1} + a_t, \text{ onde } X_t = \text{PLD}$$

4.4

Análise e comparação dos resultados

Após obter os resultados dos modelos uni-variados e causais, foi realizada uma análise comparativa dos resultados obtidos através destes modelos.

Com o objetivo de enriquecer a comparação entre os modelos, foram calculadas estatísticas que visam medir o desempenho preditivo dos mesmos tanto dentro da amostra (*in-sample*) quanto fora desta (*out-of-sample*). Para analisar o desempenho fora da amostra, foram utilizados os dados dos últimos 6 meses, ou seja, foram retirados da amostra os dados referentes aos meses de agosto de 2018 a janeiro de 2019. Feito isto, os parâmetros da equação são novamente estimados calculando-se, com origem em julho de 2018, a previsão para todo o período fora da amostra, ou seja, 6 previsões. Tem-se, neste momento, uma previsão para cada horizonte, ou seja, uma previsão para 1 mês à frente, 1 previsão para 2 meses à frente e assim sucessivamente. Depois, é feito um deslocamento da origem de previsão, ou seja, através do mesmo modelo estimado com dados até julho de 2018, desloca-se a origem para agosto de 2018, realizando-se as previsões para o período de setembro de 2018 a janeiro de 2019, ou seja, 5 previsões. Este processo se repete até que se obtenha apenas uma previsão fora da amostra. Ao final, tem-se 6 previsões para 1 mês à frente, 5 previsões para 2 meses à frente, 4 previsões para 3 meses à frente e assim sucessivamente, totalizando 21 previsões fora da amostra para diferentes horizontes.

Para comparar o desempenho dos modelos utilizados *in sampe* foi utilizada a métrica de desempenho MAPE (*Mean Absolute Percentage Error*), o BIC para

julgar os modelos do ponto de vista da sua quantidade de parâmetros e também o coeficiente de explicação (R^2 ajustado).

Os desempenhos *in sample* dos modelos univariados e do modelo de regressão dinâmica encontram-se na Tabela 6. Pode-se observar que o modelo de regressão dinâmica apresentou melhores resultados para as três métricas avaliadas, MAPE *in sample*, BIC e R^2 ajustado. O MAPE de 29,52% do modelo de regressão dinâmica indica que, dentro da amostra, o modelo apresenta erro de 29,52%. Por sua vez, o R^2 ajustado de 0,8355 indica que o modelo explica 83,55% da variação dos volumes regaseificados.

Tabela 6: Desempenho *in sample* dos modelos univariados e de regressão dinâmica.

	Amortecimento Exponencial	ARIMA	Regressão Dinâmica
MAPE	55,96%	43,44%	29,52%
BIC	16,36	16,54	13,07
R^2 Ajustado	74,45%	72,13%	83,55%
Ljung-Box (18)	p=0,4696	p=0,994	p=0,8363

O desempenho *out-of-sample* foi avaliado através do MAPE e MAPE acumulado. Além dos melhores resultados de desempenho *in sample* do modelo de regressão dinâmica, o modelo também demonstrou melhor desempenho na análise *out-of-sample*, conforme Tabela 7, quando comparado com os desempenhos dos modelos univariados.

Tabela 7: Desempenho *out-of-sample* dos modelos univariados e de regressão dinâmica.

MAPE out-of-sample											
Amortecimento No Trend/Add Season				ARIMA(0,1,0)				REGDIN x PLDNE100			
H	N	MAPE	MAPE acumulado	H	N	MAPE	MAPE acumulado	H	N	MAPE	MAPE acumulado
1	6	37	36,5	1	6	35	34,8	1	6	16	15,8
3	4	200	94,1	3	4	200	94,7	3	4	22	18,2
4	3	300	100	4	3	300	100	4	3	28	19,8
6	1	300	100	6	1	300	200	6	1	30,8	21,5

Nota: H = horizonte; N = Número de previsões

De uma forma geral o modelo de regressão dinâmica apresentou melhor desempenho fora da amostra.

As estatísticas acumuladas são importantes também, pois indicam se as previsões se deterioraram rapidamente com o horizonte de previsão. Neste caso, observa-se que o MAPE acumulado do modelo de regressão dinâmica é igual a 21,5%, indicando que as previsões continuaram consistentes, mesmo com um horizonte de previsão maior.

Ao final podemos então visualizar os volumes de regaseificação para os três cenários de demanda propostos, onde se prevê que choverá acima das médias de longo termo ou 125% da MLT, onde choverá pouco e teremos 70% da MLT e o cenário considerado mais provável de 100% da MLT. A Figura 26 apresenta os volumes realizados de regaseificação e os previstos para a regressão dinâmica.

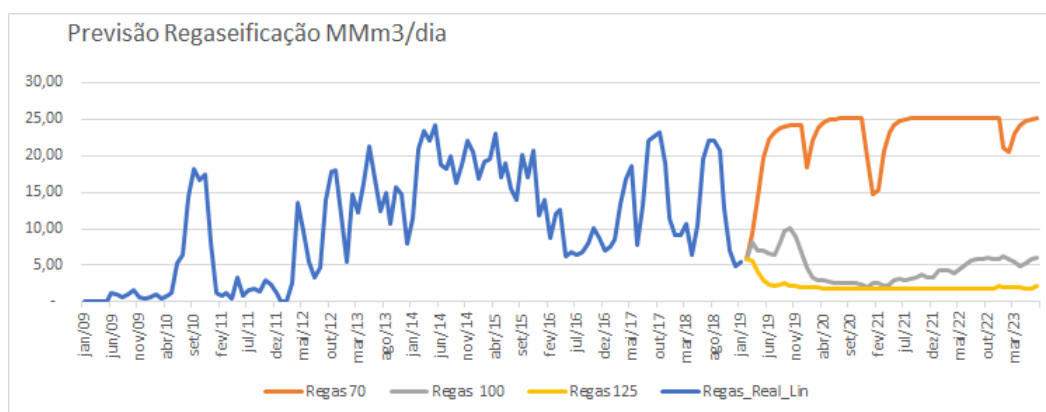


Figura 26: Volumes de Regaseificação Previstos Regressão Dinâmica em MMm³/dia.
Fonte: Compilação Própria.

O cenário de PLD a 70% da MLT tende a demandar o despacho máximo e por consequência os volumes de regaseificação máximos. Por outro lado para um cenário de muita chuva ou 125% da MLT a previsão de regaseificação do GNL para atendimento ao mercado térmico tende a zero em praticamente todo o período de previsão. Esse cenário de incerteza na previsão do PLD que tende a se repetir no período imediatamente posterior faz com que ferramentas e os modelos de previsão devam constantemente ser revisitados e ajustados, com a inserção dos dados mais atuais da variável explicativa, o PLD.

Adicionalmente, apenas para fins de teste e comparação, foi realizada uma rodada no R de uma função de transferência Box & Jenkins. Para tanto foram utilizadas os mesmos dados de PLD e Regaseificação para o modelo de regressão

dinâmica. Foi utilizado o ARIMAX.AR(3) sem defasagem da variável PLD e a plotagem dos resultados muito se assemelha a regressão dinâmica, como pode ser visto na Figura 27.

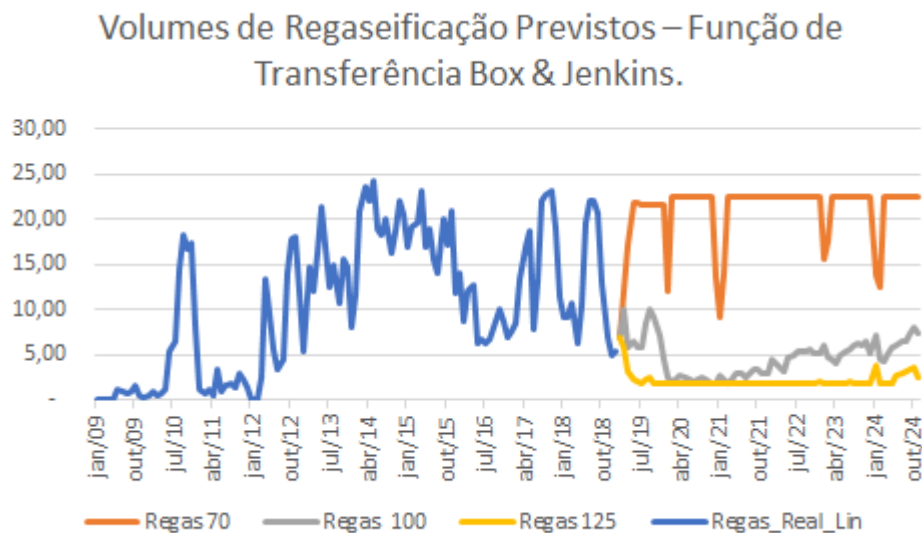


Figura 27: Volumes de Regaseificação Previstos Função de Transf. Box & Jenkins.
Fonte: Compilação Própria.

O MAPE in sample tendeu a infinito por conter zeros nos dados, e O MAPE acumulado out of sample do modelo encontrado foi muito similar ao encontrado na regressão dinâmica, como mostra a Tabela 8:

Tabela 8: Desempenho *out-of-sample* do modelo de Função de Transferência.

MAPE out-of-sample FT x PLDNE100			
H	N	MAPE	MAPE acumulado
1	6	6	15,5
2	5	8	20,8
3	4	9	24,7
4	3	9	26,4
5	2	8,4	26,1
6	1	8,1	25,1
Nota: H = horizonte; N = Número de previsões			

A fim de comparar as previsões dos modelos causais com os dados reais de regaseificação de fevereiro a agosto de 2019, já divulgados pelo MME, foi utilizado o MAPE como critério comparativo dos resultados obtidos pelos

modelos causais de regressão e função de transferência. para o período de 7 meses após a previsão. O resumo pode ser visto na Tabela 9.

Tabela 9: MAPE Previsões x Valores Reais – Fevereiro a Agosto de 2019.

MAPE Previsão X Real (%)	
REG DIN x PLD70	58,49
REG DIN x PLD_100	21,67
REG DIN x PLD_125	39,87
FT x PLD_70	60,47
FT x PLD_100	20,25
FT X PLD_125	41,28

Observa-se comportamento e números similares e a previsão com ligeiramente menor MAPE foi a utilizando o PLD de 100 % da MLT, no modelo de função de transferência. Observou-se que o critério, RMSE, ou a raiz quadrada do erro quadrático médio para a função de transferência de 0,11874, foi ligeiramente menor que a do modelo de regressão dinâmica, de 0,1254. Deste modo conclui-se que a capacidade preditiva do modelo de função de transferência foi ligeiramente superior à do modelo de regressão dinâmica.

Como a incerteza e os erros de previsão aumentam com o tempo conclui-se que as previsões devem ser refeitas periodicamente, tomando dados mais atuais que espelhem as vazões mais recentes. Sendo assim as previsões de curto prazo são mais recomendadas.

Conclusão

Como visto durante o desenvolvimento do trabalho, o gás natural liquefeito tem fundamental importância na cadeia de suprimento da indústria de gás natural e energia, e neste contexto, os volumes supridos por esta modalidade, de maneira flexível impactam e interessam aos mais diversos agentes da cadeia de suprimento e comercialização de gás e energia no Brasil.

A aplicação de modelos de previsão de importação de GNL, sem dúvida é ambiente fértil para novos estudos. Em função das mudanças e incertezas do novo mercado que se consolida neste momento, quem obtiver melhores ferramentas de previsão, pode melhor apoiar às decisões frente aos parceiros e concorrentes. Portanto este direcionamento ganha ainda mais motivação nas comunidades acadêmicas e empresariais.

No tocante à modelagem com objetivo específico de previsão de GNL, no mercado Brasileiro, os poucos trabalhos encontrados na literatura exploraram métodos estatísticos diversos, porém não foram encontrados, na pesquisa realizada, estudos envolvendo o modelos causais utilizando a variável explicativa PLD. Foram encontrados basicamente preços de derivados e temperaturas no verão e inverno. Isso é um diferencial deste trabalho em relação aos encontrados na literatura.

A incorporação de variáveis exógenas nos modelos de previsão demonstrou uma melhora na previsão de GNL se comparado aos modelos univariados, e o modelo de função de transferência mostrou ter capacidade preditiva ligeiramente superior ao modelo de regressão dinâmica.

Caso um novo cenário de preços e fontes de suprimento mais competitivos para o gás natural ora em construção se consolide, outras variáveis, tais como os preços do gás natural em centros de comercialização de referência podem se tornar significativos e combinados com a variável PLD, podem tornar as previsões mais assertivas.

5.1

Trabalhos futuros

O estudo desenvolvido, apesar de atingir o objetivo proposto, apresenta uma limitação por fazer referência à apenas uma variável explicativa. Porém outras variáveis econômicas como os preços em centros de referência nacionais e internacionais, tais como os mostrados na Figura 14 parecem fazer sentido e poderão ganhar peso caso o mercado ganhe maturidade e ofereça flexibilidade de suprimento aos carregadores e supridores de gás e energia. Neste cenário competitivo, onde haverá diversos supridores, os preços consequentemente tenderão a ser variáveis que farão sentido como explicação à importação de GNL.

Referências bibliográficas

ANEEL - Agência Nacional de Energia Elétrica), 2019. Disponível em < www.aneel.com.br >.

ANUÁRIO ESTATÍSTICO 2018, Agência Nacional do Petróleo (ANP). Disponível em < <http://www.anp.gov.br/publicacoes/anuario-estatistico/anuario-estatistico-2018> >.

ARAS, H., ARAS, N. **Forecasting residential natural gas demand**. Energy Sources, 2004, 26(5), pp. 463-472

BALITSKIY, S.; BILAN, Y.; STRIELKOWSKI, W. **Energy security and economic growth in the European union**. Journal of Security and Sustainability Issues, 2014, Volume 4, Issue 2, 2014, Pages 125-132.

BOX, G.E.P.; JENKINS, G.M.; REINSEL, G.C.; LJUNG, G.M. **Time Series Analysis: Forecasting and Control**, 1976. Fifth Edition. Wiley, 2016.

BP STATISTICAL REVIEW. Inglaterra, 2019. Disponível em < <https://www.bp.com/en/global/corporate/energy-economics/statistical-review-of-world-energy.html> >.

CCEE - Câmara de comercialização de Energia, 2019, PLD. Disponível em < <https://www.ccee.org.br> >.

CAO, F.; YIN, H.; SUN, R. **World natural gas export capacities and China's overseas investment options**. Natural Gas Industry, 2015, 35(11), pp. 113-118.

COCHRANE, D.; ORCUTT, G.H. **Application of Least Squares Regression to Relationships Containing Autocorrelated Error Terms**. Journal of American Statistical Association, 1949, 44, p. 32-61.

CHAE, Y.; KIM, M.; YOO, S.-H. **Does natural gas fuel price cause system marginal price, vice-versa, or neither? A causality analysis**. Energy, 2012, 47(1), pp. 199-204.

CHEN, Y.; CHUA, W.S.; KOCH, T.; **Forecasting day-ahead high-resolution natural-gas demand and supply in Germany**. Applied Energy, 2018, 228, pp. 1091-1110.

CHUNG, M.; PARK, H.-C.; COIMBRA, C.F.M. **Estimation of the building energy loads and LNG demand for a cogeneration-based community energy system: A case study in Korea**. Energy Conversion and Management, 2014, vol. 87, pp. 1010-1026.

FERKINGSTAD, E.; LØLAND, A.; WILHELMSEN, M. **Causal modeling and inference for electricity markets.** *Energy Economics*, 2011, 33(3), pp. 404-412.

FLASZA, R.; RIPPEL, M.; SOLC, J. **Modeling prices of electricity futures at EEX.** *World Academy of Science, Engineering and Technology*, 2011, 78, pp. 673-677.

GARDNER JR., E.S. **Exponential smoothing: The state of the art-Part II, 2006**, *International Journal of Forecasting*, 22(4), pp. 637-666.

GOODRICH, R.L. **Applied Statistical Forecasting.** Belmont, Business Forecast Systems, 1989.

HÖNIG, V.; PROCHAZKA, P.; OBERGRUBER, M.; SMUTKA, L.; KUCEROVÁ, V. **Economic and technological analysis of commercial LNG production in the EU.** *Energies*, 2019, 12(8), en12081565.

HYNDMAN, R.J.; ATHANASOPOULOS, G. **Forecasting: principles and practice.** 2016. Disponível em: <https://otexts.org/fpp/>. Acessado em maio de 2018.

IEA, **World Energy Outlook 2017.** Estados Unidos, 2019. Disponível em < <https://www.iea.org/weo2017/>>.

IGU **World LNG Report.** EUA, 2019. Disponível em < https://www.igu.org/sites/default/files/node-news_item-field_file/IGU%20Annual%20Report%202019_23%20loresfinal.pdf >.

LAIB, O.; KHADIR, M.T.; MIHAYLOVA, L. **Toward efficient energy systems based on natural gas consumption prediction with LSTM Recurrent Neural Networks .** *Energy*, vol. 177, 2019, pp. 530-542.

MAÇAIRA, P.M.; THOMÉ, A.M.T.; OLIVEIRA, F.L.C.; FERRER, A.L.C. **Time series analysis with explanatory variables: A systematic literature review.** *Environmental Modelling & Software*, 2018, p. 199-209

MME, **Boletim mensal de acompanhamento da indústria de gás natural.** Disponível em < <http://www.mme.gov.br/web/guest/secretarias/petroleo-gas-natural-e-combustiveis-renovaveis/publicacoes/boletim-mensal-de-acompanhamento-da-industria-de-gas-natural> >.

MIT, **Massachusetts Institute of Technology** The future of Natural Gas, 2011. Disponível em < <https://energy.mit.edu/wp-content/uploads/2011/06/MITEI-The-Future-of-Natural-Gas.pdf> >.

MONGEON, P.; PAUL-HUS, A. **The journal coverage of Web of Science and Scopus: a comparative analysis.** *Scientometrics*, 2016, vol. 106, Issue 1, p. 213228.

PETROBRAS, **Petróleo Brasileiro S.A.** Disponível em < <https://petrobras.com.br/pt/nossas-atividades/principais-operacoes/terminais-e-oleodutos/terminal-de-regaseificacao-de-pecem-gln.htm> >.

PDE 2017 - Plano Decenal de Energia 2017, Empresa de pesquisa Energética (EPE). Disponível em < <http://www.epe.gov.br/pt/publicacoes-dados-abertos/publicacoes/plano-decenal-de-expansao-de-energia-2027>>.

PIPE 2019 - Plano Indicativo de processamento e escoamento, **EPE**. Disponível em < [http://epe.gov.br/sites-pt/sala-de-imprensa/noticias/Documents/ PIPE%20 Apresenta%20a7%20a3o.pdf](http://epe.gov.br/sites-pt/sala-de-imprensa/noticias/Documents/PIPE%20Apresenta%20a7%20a3o.pdf) >.

R version 3.6.0 (2019-04-26) - "Planting of a Tree" Copyright (C) 2019
The R Foundation for Statistical Computing Platform: x86_64-w64-mingw32/x64 (64-bit).

SCHWARZ, G. **Estimating the dimension of a model**. The annals of statistics, 1978, vol. 6, Nº 2, p. 461-464.

STELLWAGEN, E.; GOODRICH, R. **Forecast Pro Statistical Reference Manual**. Belmont, MA: Business Forecast Systems, 2011.