

GUSTAVO MODENESI

**MODELO DE PREVISÃO DE DEMANDA DE GÁS
NATURAL**

Trabalho de Formatura apresentado à
Escola Politécnica da Universidade de
São Paulo para obtenção de Diploma
de Engenheiro de Produção.

São Paulo

2007

GUSTAVO MODENESI

**MODELO DE PREVISÃO DE DEMANDA DE GÁS
NATURAL**

Trabalho de Formatura apresentado à
Escola Politécnica da Universidade de
São Paulo para obtenção de Diploma
de Engenheiro de Produção.

Orientador: Miguel Cezar Santoro

São Paulo

2007

FICHA CATALOGRÁFICA

Modenesi, Gustavo

**Modelo de previsão de demanda de gás natural /
G. Modenesi. -- São Paulo, 2007.
163 p.**

**Trabalho de Formatura - Escola Politécnica da
Universidade
de São Paulo. Departamento de Engenharia de
Produção.**

**1.Gás natural – Brasil 2.Demanda (Previsão;
Modelos) I.Uni-
versidade de São Paulo. Escola Politécnica.
Departamento de Engenharia de Produção II.t.**

À minha família

AGRADECIMENTOS

Cabem, nesta oportunidade, meus mais sinceros agradecimentos àqueles que de alguma forma contribuíram não apenas para a elaboração deste trabalho de formatura, mas sim àquilo que ele representa: a conclusão de mais uma importante etapa em minha vida e o começo de mais uma longa jornada.

Aos meus pais, Cláudio Ricardo Modenesi e Eliza Takara Modenesi, pela educação que me proporcionaram, pela paciência nas horas difíceis e pelo estímulo que sempre me deram.

Ao meu professor orientador, Miguel Cezar Santoro, por aceitar e abraçar o tema proposto neste trabalho, e pelas inúmeras horas dispensadas nos ensinamentos, sugestões, encaminhamentos, revisões e conselhos.

À professora Linda Lee Ho, pela paciência, dedicação e enorme auxílio prestado, principalmente na parte estatística deste trabalho.

Aos meus amigos da graduação, pelo companheirismo, suporte e por tudo o que vivemos nesses anos de graduação.

À empresa onde estagiei durante o último ano da graduação e a seus funcionários com os quais tive contato direto, por terem oferecido acesso irrestrito aos recursos físicos e intelectuais vitais para a realização deste trabalho.

A todos os professores e funcionários da USP que tiveram participação no meu desenvolvimento acadêmico e pessoal.

RESUMO

O gás natural vem ganhando grande participação na matriz energética brasileira nos últimos anos, se tornando peça fundamental da economia e do desenvolvimento brasileiros.

Devido às características deste mercado, o aumento da oferta de gás natural somente pode ser realizado um longo período depois de tomada a decisão de fazê-lo, seja pelo tempo que se leva para se desenvolver campos de gás, para se construir gasodutos ou para se construir plantas de regaseificação, no caso do Gás Natural Liquefeito (este último a única opção de mais curto prazo para aumentar a oferta de gás, mas que possui limitações de volume).

Desta forma, mostra-se vital o conhecimento antecipado da demanda futura de gás natural, de forma que possibilite o levantamento, análise, escolha e execução de planos de aumento da oferta de gás.

Este trabalho objetiva desenvolver e discutir modelos de previsão da demanda brasileira de gás natural, primordialmente de curto a médio prazo (1 a 5 anos), os quais possam servir como base para que predições de longo prazo sejam realizadas por especialistas do setor.

Para tanto, utilizar-se-á os métodos de Holt, regressão explanatória e auto-regressão para cada uma das parcelas da demanda brasileira de gás natural, a saber: setor de Transportes, setor Energético, setor Industrial, setor Residencial e setor Comercial/Público.

Uma vez elaborados os mesmos, será realizada uma análise crítica destes frente a critérios como tamanho dos erros, número de dados incluídos no modelo, R^2 , amplitude dos intervalos de predição, respeito às hipóteses subjacentes ao modelo e alinhamento dos resultados do mesmo com outros.

ABSTRACT

Natural gas has been increasing its share in Brazilian energetic matrix in the last years, becoming a fundamental piece of Brazilian economy and development.

Due to some market characteristics, the increase of natural gas supply can only be realized a long period after its decision has been taken, may it be because of the time it takes to develop gas fields, to construct gas pipelines or to construct regaseification plants, in the Liquefied Natural gas case (this one the only short time option to increase gas supply, but that has volume restraints).

Therefore it is vital to know in advance the future demand of natural gas, in a way that it allows the elaboration, analysis, choice and execution of plans to increase supply.

This work aims to develop and discuss forecast models for Brazilian demand of natural gas, firstly for the short and medium term (1 to 5 years), which can work as references and bases for long term predictions to be made by sector specialists.

To accomplish that, it will be used Holt's, explanatory regression and auto-regression methods for each one of the portions of Brazilian natural gas demand: Transport sector, Energetic sector, Industrial sector, Residential sector and Commercial/Public sector.

Once elaborated the models they will be critically analyzed face to criteria such as size of the errors, number of data included in the model, R^2 , amplitude of prediction intervals, respect to models' subjacent hypothesis and alignment of outputs comparing to other models.

SUMÁRIO

1. Apresentação da empresa do estágio	21
2. O programa de estágio	25
2.1 Estudo sobre o mercado de usinas de geração térmica a gás.....	25
2.2 PMO da Prática de Energia	25
2.3 Auxílios pontuais em projetos e propostas de projeto	26
3. Objetivo do trabalho	27
4. Justificativa	29
5. O Mercado de Gás Natural.....	33
5.1 Principais utilizações	33
5.1.1 Industrial.....	34
5.1.2 GNV – Gás Natural Veicular.....	34
5.1.3 Residencial e Comercial	34
5.1.4 Geração Elétrica/Co-geração	34
5.2 Risco de Desabastecimento	34
6. Referencial Teórico	37
6.1 Visão Geral sobre previsões	37
6.2 Conceitos básicos relacionados a previsões	40
6.2.1 Estatísticas básicas.....	40
6.2.2 Padrões dos dados.....	44
6.2.3 Medidas de erros.....	45
6.2.4 Intervalos de previsão.....	48
6.2.5 Hipótese da Continuidade.....	50
6.3 Modelos de previsão	52
6.3.1 Previsão Quantitativa.....	52
6.3.2 Previsão Qualitativa.....	69
6.4 Utilização dos diferentes métodos de previsão.....	73
7. Modelo de previsão de demanda de gás natural.....	75
7.1 Setor de Transportes	78
7.1.1 Transportes – Método de Holt.....	78
7.1.2 Transportes – Regressão Explanatória	79
7.1.3 Transportes – Auto-regressão	82
7.1.4 Transportes – Comparação dos modelos	85
7.2 Setor Energético	89

7.2.1	Energético – Método de Holt.....	89
7.2.2	Energético – Regressão Explanatória	90
7.2.3	Energético – Auto-regressão	95
7.2.4	Energético - Comparação dos modelos.....	100
7.3	Setor Industrial.....	103
7.3.1	Setor Industrial – Holt.....	103
7.3.2	Setor Industrial – Regressão Explanatória	104
7.3.3	Setor Industrial – Auto-regressão	108
7.3.4	Industrial - Comparação dos modelos.....	113
7.4	Setor Residencial.....	116
7.4.1	Residencial – Método de Holt.....	116
7.4.2	Residencial – Regressão Explanatória	117
7.4.3	Residencial – Auto-regressão.....	122
7.4.4	Residencial – Comparação dos modelos.....	124
7.5	Setor Comercial/Público	127
7.5.1	Setor Comercial/Público – Método de Holt.....	127
7.5.2	Setor Comercial/Público – Regressão Explanatória	128
7.5.3	Setor Comercial/Público – Auto-regressão.....	133
7.5.4	Setor Comercial/Público – Comparação dos modelos.....	135
7.6	Total	139
8.	Conclusões	141
9.	Referências Bibliográficas.....	143
10.	Apêndices	145

LISTA DE FIGURAS

Figura 1: Organização matricial do BCG	21
Figura 2: Áreas de atuação do BCG	22
Figura 3: Visão geral dos modelos de previsão existentes	37
Figura 4: Processo de previsão	39
Figura 5: Decomposição de série temporal em suas parcelas	45
Figura 6: Escolha de variáveis para regressão múltipla	56
Figura 7: <i>Stepwise forward-with-a-backward-look regression</i>	57
Figura 8: Exemplo de decomposição de série temporal com sazonalidade aditiva.....	60
Figura 9: Exemplo de série temporal com sazonalidade multiplicativa.....	60
Figura 10: Passo-a-passo de decomposição aditiva com posterior previsão	61
Figura 11: Passo-a-passo de decomposição multiplicativa com posterior previsão.....	62
Figura 12: Metodologia Box-Jenkins para modelagem de séries temporais	67
Figura 13: Funcionamento do método Delphi.....	71
Figura 14: Possíveis variáveis explanatórias para demanda de transportes	79
Figura 15: Possíveis variáveis explanatórias para demanda energética	90
Figura 16: Possíveis variáveis explanatórias para demanda industrial.....	104
Figura 17: Possíveis variáveis explanatórias para a demanda do setor residencial.....	117
Figura 18: Possíveis variáveis explanatórias para a demanda do setor comercial/público	128
Figura 19: Dados históricos e projeção das variáveis explanatórias I.....	151
Figura 20: Dados históricos e projeção das variáveis explanatórias II.....	151

LISTA DE GRÁFICOS

Gráfico 1: Matriz energética brasileira – Evolução do consumo final por fonte	29
Gráfico 2: Evolução da participação do gás natural na matriz energética brasileira.....	30
Gráfico 3: Oferta de energia elétrica – Histórico e projeção	31
Gráfico 4: Participação por tipo na geração elétrica.....	31
Gráfico 5: Consumo de Gás Natural – Histórico e projeção	33
Gráfico 6: Oferta nacional de gás natural crescendo com importações	35
Gráfico 7: Tipos de correlação	41
Gráfico 8: Correlação e relacionamento não-linear de variáveis	42
Gráfico 9: Exemplo de gráfico ACF.....	44
Gráfico 10: Exemplificação dos tipos de erros.....	48
Gráfico 11: Exemplo de regressão linear	54
Gráfico 12: Ilustração do funcionamento do método Naive.....	58
Gráfico 13: Ilustração do funcionamento do método de Média Móvel.....	63
Gráfico 14: Consumo de Gás Natural por fonte	75
Gráfico 15: Consumo de Gás Natural por fonte – dados não empilhados	76
Gráfico 16: Modelo de Holt para demanda do Setor de Transportes	78
Gráfico 17: Resíduos Transporte vs Carros GNV I.....	80
Gráfico 18: ACF dos Resíduos Transporte vs Carros GNV I	81
Gráfico 19: Demanda de transportes em função de número de carros movidos a GNV.....	82
Gráfico 20: ACF da demanda do setor de Transportes	83
Gráfico 21: Resíduos auto-regressão Transporte.....	84
Gráfico 22: Previsão da demanda de transportes em função de auto-regressão.....	85
Gráfico 23: Comparação dos modelos - Transportes	86
Gráfico 24: Modelo de Holt para demanda do Setor Energético	89

Gráfico 25: Resíduos do Setor Energético – modelo explanatório	91
Gráfico 26: ACF dos resíduos - modelo explanatório setor energético	92
Gráfico 27: Série histórica de preços relativos gás natural/óleo combustível	93
Gráfico 28: Demanda do Setor Energético em função do modelo de regressão explanatória.	94
Gráfico 29: Demanda do Setor Energético em função do modelo de regressão explanatória com intervalos de confiança.....	95
Gráfico 30: ACF da demanda do Setor Energético.....	96
Gráfico 31: Setor energético – resíduos com <i>outlier</i>	97
Gráfico 32: Setor energético – resíduos sem <i>outlier</i>	98
Gráfico 33: Setor Energético – resíduos de segunda ordem	99
Gráfico 34: Demanda do setor energético em função do modelo de auto-regressão.....	100
Gráfico 35: Comparação dos modelos – Energético.....	101
Gráfico 36: Modelo de Holt para demanda do Setor Industrial.....	103
Gráfico 37: Resíduos Setor Industrial vs PIB	105
Gráfico 38: ACF dos resíduos Industrial vs PIB.....	106
Gráfico 39: Resíduos de ordem 2 da regressão explanatória do Setor Industrial	107
Gráfico 40: Demanda do setor industrial em função do modelo de regressão explanatória..	108
Gráfico 41: ACF da demanda do Setor Industrial.....	109
Gráfico 42: Setor industrial – resíduos de primeira ordem.....	110
Gráfico 43: Setor Industrial – resíduos de segunda ordem	111
Gráfico 44: Demanda do setor industrial em função do modelo de auto-regressão	112
Gráfico 45: Comparação dos modelos – Industrial.....	113
Gráfico 46: Modelo de Holt para demanda do Setor Residencial.....	116
Gráfico 47: Resíduos da regressão do Setor Residencial.....	118
Gráfico 48: Resíduos Residencial vs PIB	119
Gráfico 49: Setor Residencial - Resíduos de segunda ordem	120

Gráfico 50: Previsão da demanda residencial em função do PIB.....	121
Gráfico 51: ACF da demanda do setor Residencial	122
Gráfico 52: Resíduos da auto-regressão Residencial	123
Gráfico 53: Previsão da demanda Residencial em função de auto-regressão	124
Gráfico 54: Comparação dos modelos – Residencial	125
Gráfico 55: Modelo de Holt para demanda do Setor Comercial/Público.....	127
Gráfico 56: Resíduos Comercial/Público vs PIB	129
Gráfico 57: ACF comercial/público vs PIB	130
Gráfico 58: Resíduos de segunda ordem do setor Comercial/Público	131
Gráfico 59: Previsão da demanda comercial/público em função do modelo de regressão	132
Gráfico 60: ACF da demanda do Setor Comercial/Público	133
Gráfico 61: Resíduos do Setor Comercial/Público.....	134
Gráfico 62: Demanda do setor Comercial/Público em função do modelo de auto-regressão	135
Gráfico 63: Comparação dos modelos – Comercial/Público.....	136
Gráfico 64: Previsão de demanda total – método de Holt.....	139
Gráfico 65: Previsão de demanda total – Soma.....	140
Gráfico 66: Previsão de demanda total – Comparação.....	140

LISTA DE TABELAS

Tabela 1: <i>Lag</i> de uma série.....	43
Tabela 2: Exemplificação dos tipos de erros	47
Tabela 3: Exemplo de regressão linear.....	54
Tabela 4: Ilustração do funcionamento do método de média móvel.....	63
Tabela 5: Equivalência entre modelos de série temporal e modelos ARIMA.....	69
Tabela 6: Resumo de modelos e suas aplicações	74
Tabela 7: Análise das séries históricas de consumo por fonte de gás natural	76
Tabela 8: Quadro comparativo de modelos – Setor de Transportes.....	87
Tabela 9: Quadro comparativo de modelos – Setor Energético	102
Tabela 10: Quadro comparativo de modelos – Setor Industrial	114
Tabela 11: Quadro comparativo de modelos – Setor Residencial.....	126
Tabela 12: Quadro comparativo de modelos – Setor Comercial/Público	137
Tabela 13: Dados de consumo de gás natural por segmento.....	146
Tabela 14: Dados históricos das variáveis regressoras I	147
Tabela 15: Previsões das variáveis regressoras I.....	148
Tabela 16: Dados históricos das variáveis regressoras II.....	149
Tabela 17: Previsões das variáveis regressoras II	150
Tabela 18: Resultados da previsão do Setor de Transportes	155
Tabela 19: Resultados da previsão do Setor Energético.....	156
Tabela 20: Resultados da previsão do Setor Industrial	157
Tabela 21: Resultados da previsão do Setor Residencial	158
Tabela 22: Resultados da previsão do Setor Comercial/Público	159
Tabela 23: Resultados da previsão de demanda total	161

LISTA DE EQUAÇÕES

Equação 1: Correlação Linear	40
Equação 2: Autocorrelação.....	43
Equação 3: Erro de uma previsão	46
Equação 4: <i>Mean Error</i> (ME)	46
Equação 5: <i>Mean Absolute Error</i> (MAE)	46
Equação 6: <i>Mean Squared Error</i> (MSE).....	46
Equação 7: <i>Mean Percentage Error</i> (MPE)	46
Equação 8: <i>Mean Absolute Percentage Error</i> (MAPE)	47
Equação 9: Intervalo de previsão para próximo período	49
Equação 10: Cálculo do intervalo de confiança para regressão simples e múltipla	49
Equação 11: Cálculo do sigma da regressão simples	50
Equação 12: Cálculo do sigma da regressão múltipla	50
Equação 13: Forma geral da regressão simples.....	52
Equação 14: Forma linear da regressão simples.....	53
Equação 15: Forma geral da regressão múltipla.....	55
Equação 16: Forma linear da regressão múltipla	55
Equação 17: R^2 ajustado	56
Equação 18: Previsão Naive	58
Equação 19: Decomposição aditiva.....	59
Equação 20: Decomposição multiplicativa	59
Equação 21: Previsão por média simples	62
Equação 22: Single Exponential Smoothing (SES).....	65
Equação 23: Método de Holt.....	65
Equação 24: Forma geral da auto-regressão.....	66
Equação 25: Forma linear da auto-regressão.....	66

Equação 26: Diferenciação de nível um de uma série temporal	68
Equação 27: Diferenciação sazonal de uma série temporal	68
Equação 28: Forma geral dos modelos ARIMA	68
Equação 29: Demanda de transportes em função de modelo de regressão explanatória	79
Equação 30: Demanda de transportes em função de modelo de auto-regressão	83
Equação 31: Demanda do setor energético em função de modelo de regressão explanatória	91
Equação 32: Demanda do setor energético em função de modelo de auto-regressão I	97
Equação 33: Demanda do setor energético em função de modelo de auto-regressão II	98
Equação 34: Demanda do setor energético em função do modelo de auto-regressão III	99
Equação 35: Demanda industrial em função de modelo de regressão explanatória I	105
Equação 36: Demanda industrial em função de modelo de regressão explanatória II	106
Equação 37: Demanda industrial em função de modelo de regressão explanatória III	107
Equação 38: Demanda industrial em função de modelo de auto-regressão I	110
Equação 39: Demanda industrial em função de modelo de auto-regressão II	111
Equação 40: Demanda industrial em função de modelo de auto-regressão III	112
Equação 41: Demanda residencial em função de modelo de regressão explanatória I	117
Equação 42: Demanda residencial em função de modelo de regressão explanatória II	119
Equação 43: Demanda residencial em função de modelo de regressão explanatória III	120
Equação 44: Demanda residencial em função de modelo de auto-regressão	123
Equação 45: Demanda Comercial/Público em função de modelo de regressão explanatória I	129
Equação 46: Demanda Comercial/Público em função de modelo de regressão explanatória II	130
Equação 47: Demanda Comercial/Público em função de modelo de regressão explanatória III	131
Equação 48: Demanda Comercial/Público em função de modelo de auto-regressão	134
Equação 49: Previsão da demanda total	139

1. Apresentação da empresa do estágio

Fundada em 1963 por Bruce Henderson e existente no Brasil desde 1997, o Boston Consulting Group (BCG) é uma das maiores empresas mundiais de consultoria estratégica.

O BCG se organiza matricialmente através de áreas funcionais e áreas de expertise na indústria, conforme apresentado na figura 1:



Figura 1: Organização matricial do BCG

Mais detalhadamente, podem-se expandir as áreas funcionais em temas mais específicos, os quais se mostram mais facilmente compreensíveis, conforme apresentado na figura 2 abaixo:

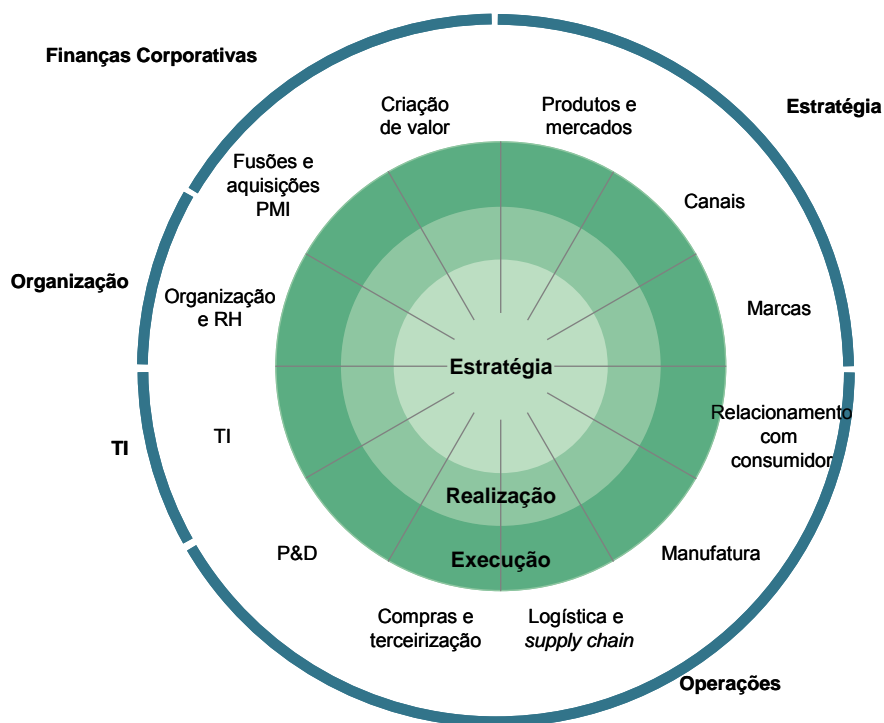


Figura 2: Áreas de atuação do BCG

De acordo com informações do site da empresa e do artigo sobre a mesma no Wikipedia, o BCG possui faturamento anual de cerca de U\$ 1,8 bi e 66 escritórios situados em 38 países ao redor do mundo (o único continente no qual o BCG não possui escritórios é a África). Nestes escritórios trabalham cerca de 4200 consultores (média de 64 consultores por escritório).

Conhecida mundialmente por seus serviços, ferramentas analíticas (matriz BCG de portfólio, Curva de Experiência, entre outros), publicações, e seus ex-membros, o BCG tem como missão ajudar corporações com grande relevância em seus segmentos de atuação a adquirir e sustentar uma vantagem estratégica de longo prazo. Como exemplos de clientes mundiais do BCG se pode citar: Whirlpool, Ford, Pfizer e IBM.

Além disso, no âmbito de responsabilidade social, o Boston Consulting Group realiza projetos sociais (*pro-bono*) mundialmente, trabalhando em parceria com governos, ONGs e outras iniciativas públicas. Entre eles podem-se citar: governo canadense, The Dallas Foundation e Harvard School Public Health.

Atualmente os principais concorrentes do BCG no mercado de consultoria estratégica são a McKinsey e Bain & Co.

Em 2007 o BCG foi considerado pela publicação Fortune como a oitava melhor empresa para se trabalhar nos EUA (tendo ficado em primeira na categoria de pequenas empresas).

O BCG trabalha sob a estrutura de projetos (organização clássica de empresas de consultoria), alocando uma equipe para cada cliente/projeto. Essas equipes são compostas, geralmente por um sócio da empresa, um gerente de projetos, um consultor e um ou dois analistas.

Tal estrutura de equipes é bastante flexível, podendo ser alterado tanto o número de pessoas por equipe como suas atribuições dentro dela (dentro das capacitações, nível de experiência de seus integrantes e demanda do projeto/cliente).

2. O programa de estágio

O programa de estágio iniciou-se em 15/01/07, com duração inicialmente prevista para 31/06/07, tendo sido posteriormente prorrogada para 31/12/07.

Ao longo do programa de estágio, desenvolveram-se três grupos principais de atividades:

2.1 Estudo sobre o mercado de usinas de geração térmica a gás

A principal atividade desenvolvida no estágio consistiu na elaboração de um estudo de mercado sobre usinas térmicas de geração a gás.

Para tanto, estudou-se o mercado de energia elétrica (histórico e projeções de oferta e demanda, estrutura de preços e estrutura competitiva), o mercado de gás natural (histórico e projeções de oferta e demanda, estrutura de preços, estrutura competitiva, rede de transporte e legislação) e o mercado de usinas térmicas em si (histórico de políticas de incentivo governamental, distribuição geográfica das usinas, vantagens e desvantagens competitivas frente a outras formas de geração elétrica, políticas de contratação de gás natural e de venda de energia elétrica).

O tema deste Trabalho de Formatura surgiu a partir de um ponto crítico encontrado durante a realização deste trabalho: “Qual será a demanda de gás natural no futuro, a médio e longo prazo? Haverá oferta suficiente para suprir esta demanda?”

O porquê de esta questão ser crítica será mais bem detalhado no item “Justificativa” deste mesmo relatório.

2.2 PMO da Prática de Energia

Conforme mencionado no item “Apresentação da empresa do estágio”, o BCG é organizado em práticas de expertise na indústria.

O programa de estágio foi realizado dentro da Prática de Energia, a qual engloba todas as iniciativas relacionadas a este tema que estão sendo desenvolvidas na empresa: projetos, propostas, estudos, publicações, etc.

Por se tratar de uma prática relativamente nova no escritório de São Paulo e que vem apresentando uma taxa de crescimento bastante acentuada (os valores não são divulgados neste relatório por questão de confidencialidade) a mesma carecia de uma maior estruturação.

Desta forma, uma das atividades do estágio consistiu em desenvolver e implementar um PMO (*Project Management Office*) para a Prática de Energia, de forma que se tivesse uma visão global do que estava sendo desenvolvido e que todas as iniciativas fossem melhor gerenciadas como um todo.

2.3 Auxílios pontuais em projetos e propostas de projeto

Ao longo do estágio o aluno também teve a oportunidade de ajudar na realização de projetos e propostas de projeto da empresa (a grande maioria ligada ao setor de Energia).

As atividades desempenhadas pelo aluno consistiam em realização de entrevistas, pesquisas na mídia, pesquisas em bases de dado do governo, análises de tendência, análises explanatórias, confecção de relatórios e elaboração de apresentações.

3. Objetivo do trabalho

Este trabalho objetiva desenvolver e discutir modelos de previsão da demanda brasileira de gás natural, primordialmente de curto a médio prazo (1 a 5 anos), os quais possam servir como base para que predições de longo prazo (aproximadamente de 15 a 20 anos) possam ser realizadas por especialistas do setor.

4. Justificativa

O gás natural é peça-chave na composição da matriz energética brasileira. Segundo dados do BEN (Balanço Energético Nacional, publicado anualmente pelo Ministério de Minas e Energia do Governo Federal) o gás natural é a fonte atualmente responsável por cerca de 9% de toda a energia consumida no Brasil (esses 9% incluem utilização industrial, GNV – Gás Natural Veicular, consumo residencial e consumo doméstico, não está incluído neste número a utilização do gás natural para geração de energia elétrica).

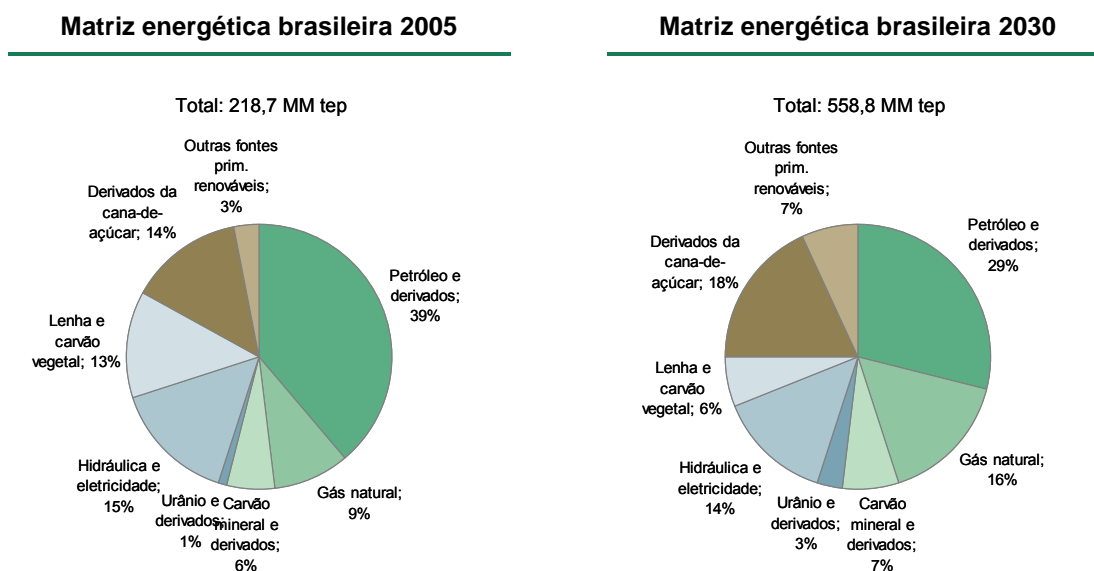


Gráfico 1: Matriz energética brasileira – Evolução do consumo final por fonte

Fonte: Balanço Energético Nacional 2006. Elaborado pelo autor

Além disso, o gás natural vem aumentando rapidamente sua participação na matriz energética brasileira, passando de uma participação de menos de 1% em 1980 para cerca de 16% em 2030, o que representa uma taxa anual de crescimento composto de 7% ao ano.

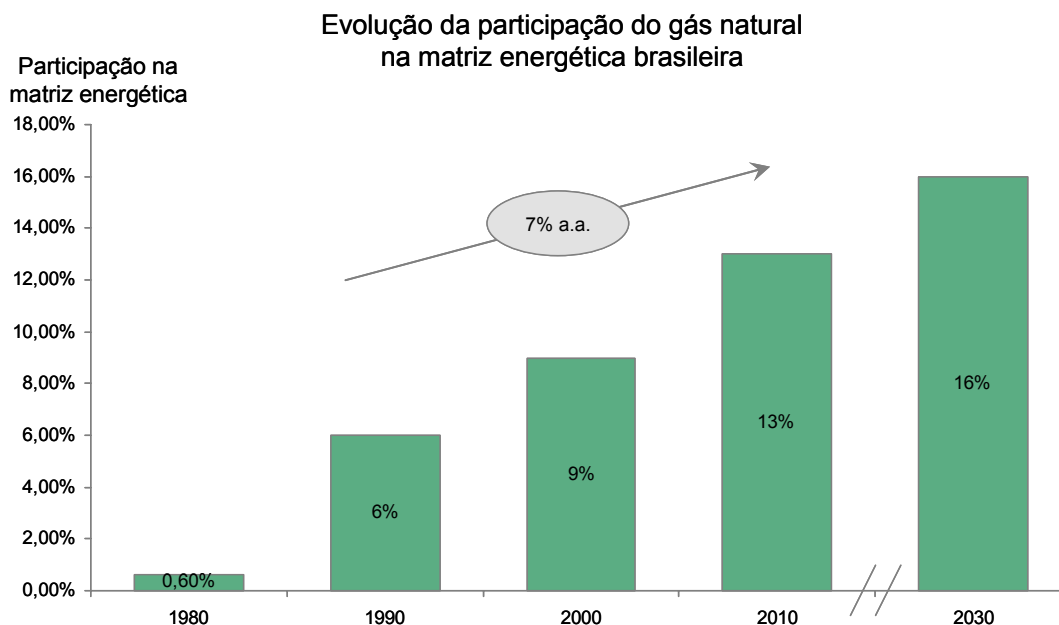


Gráfico 2: Evolução da participação do gás natural na matriz energética brasileira

Fonte: Plano Nacional de Expansão 2030. Elaborado pelo autor

Há de se acrescentar à importância do gás natural sua participação na geração elétrica nacional, a qual não está considerada nos gráficos acima.

A matriz de geração elétrica brasileira possui um caráter hidro-térmico. Nesta matriz, a grande maioria da energia gerada é de origem hidráulica e a geração térmica possui um caráter de complementaridade nos períodos de seca.

Oferta de energia elétrica — Histórico e Projeção

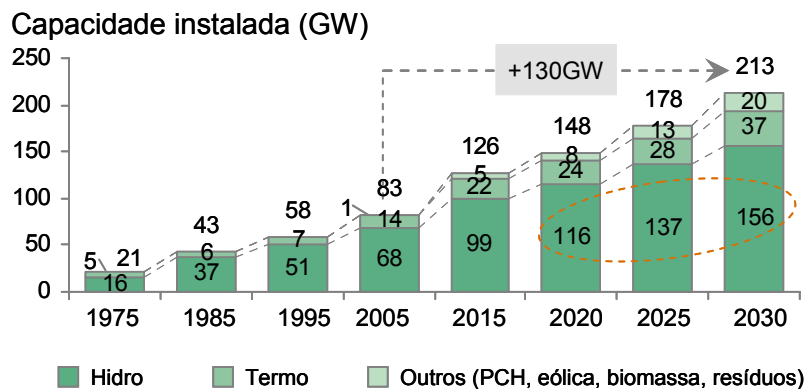


Gráfico 3: Oferta de energia elétrica – Histórico e projeção

Fonte: Plano Nacional de Expansão 2030. Elaborado pelo autor

Evolução mensal do fator de capacidade

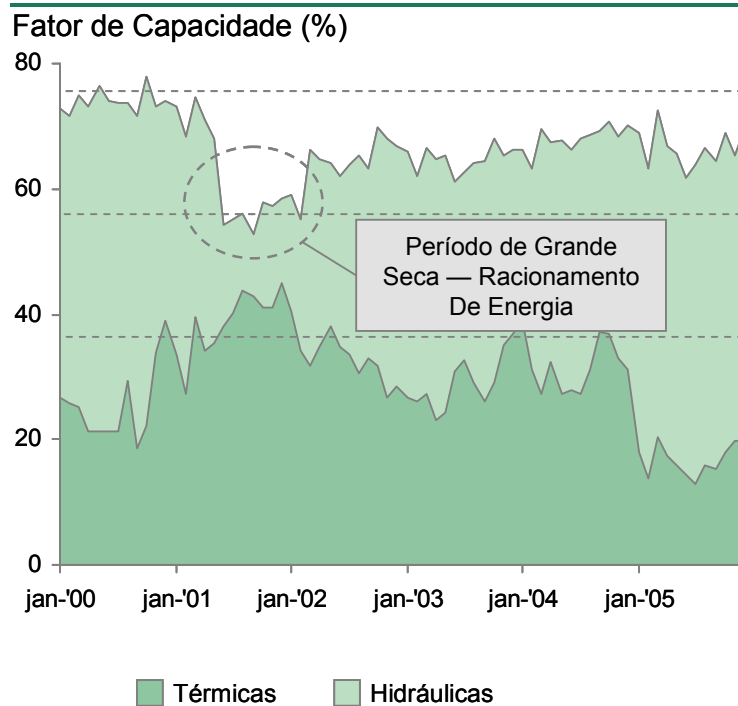


Gráfico 4: Participação por tipo na geração elétrica

Fonte: Empresa de Pesquisa Energética (EPE) e Ministério de Minas e Energia (MME). Elaborado pelo autor

Desta forma, a geração térmica (cuja principal componente com cerca de 50% da participação, de acordo com o Ministério de Minas e Energia, é a geração térmica a gás natural) possui uma importância estratégica para o parque gerador brasileiro na medida em que aumenta a segurança do mesmo contra possíveis “apagões” elétricos devido à volatilidade da geração hidráulica.

O tema gás natural também tem aparecido constantemente na mídia devido às complicações do fornecimento boliviano ao Brasil (cerca de 50% do gás atualmente consumido no Brasil é de origem boliviana) e do possível “apagão de gás natural” previsto por especialistas para o período de 2009 a 2010, no qual a demanda por gás natural será maior do que a oferta disponível (o que é ainda mais agravado pela incerteza em relação ao fornecimento boliviano para os próximos anos).

Desta forma, observa-se que o objetivo deste trabalho se mostra extremamente relevante na medida em que:

- Não existem previsões públicas de demanda de gás natural de médio e longo prazo (a maioria das previsões públicas tem horizonte até 2011~2012)
- Permite o planejamento da adequação da oferta de gás natural à demanda projetada (via desenvolvimento e exploração de campos nacionais de gás natural, importação por gasodutos ou importações via GNL – Gás Natural Liquefeito)
- É de grande interesse para empresas ligadas ao tema (indústrias consumidoras, distribuidoras, empresas de produtos substitutos como o GLP – Gás Liquefeito do Petróleo, geradoras de energia térmica e investidores).
- Trata de um tema de grande relevância para o desenvolvimento nacional.

5. O Mercado de Gás Natural

5.1 Principais utilizações

O gás natural no Brasil possui quatro principais utilizações, a saber: consumo industrial, consumo para geração elétrica de fonte térmica/co-geração, consumo como GNV e consumo residencial/comercial.

É apresentado a seguir o histórico recente do consumo de gás natural no Brasil por fonte, assim como a previsão para 2011 feita pela Petrobras:

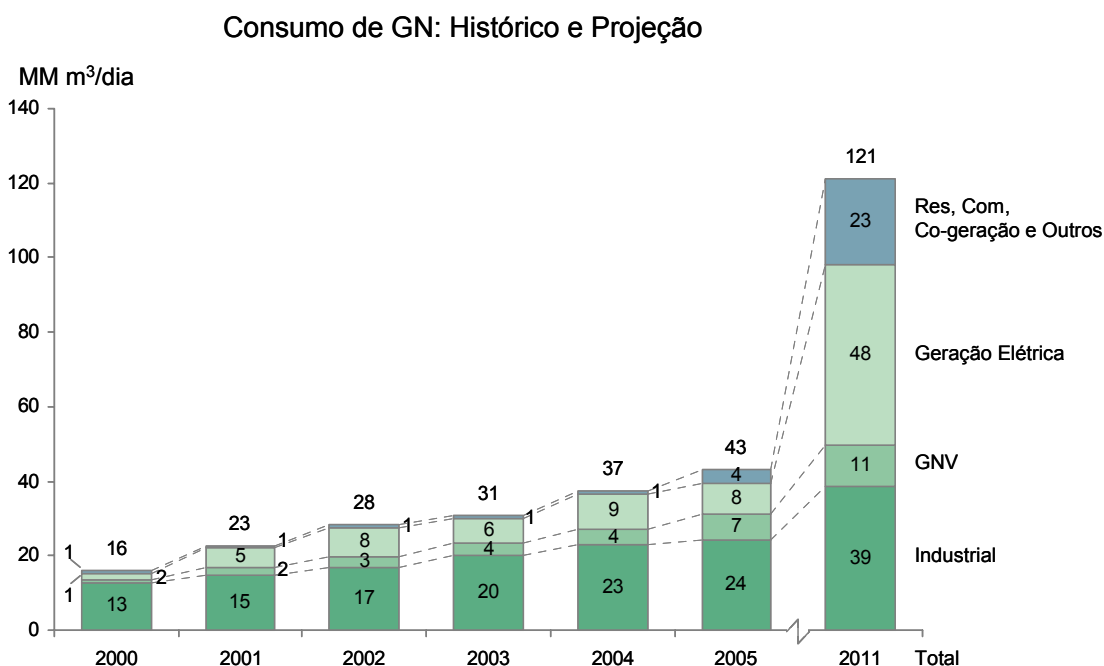


Gráfico 5: Consumo de Gás Natural – Histórico e projeção

Fonte: Plano de Negócios 2006-2011 Petrobras. Elaborado pelo autor

Observa-se no gráfico acima que a principal utilização do gás natural atualmente é o uso industrial, seguido pela geração elétrica e GNV.

Para um futuro próximo, espera-se um elevado crescimento da demanda de gás natural para geração térmica, tornando esta fonte a principal demandante de gás natural no Brasil. A demanda industrial continuará bastante alta e será seguida pela demanda residencial/comercial.

5.1.1 Industrial

O gás natural é utilizado nas indústrias como combustível para gerar calor. Atualmente alguns setores industriais no Brasil são bastante dependentes da oferta e dos preços do gás natural, como os setores industriais de fabricação de vidros e cerâmicas.

5.1.2 GNV – Gás Natural Veicular

O GNV (Gás Natural Veicular) é usado como combustíveis em automóveis. De aplicação bastante recente (foi introduzido no Brasil por volta da metade da década de 1990) tem apresentado recentemente um forte crescimento como opção economicamente atrativa de combustível.

5.1.3 Residencial e Comercial

Utilizado para gerar calor, concorre nesse mercado com o GLP (Gás Liquefeito do Petróleo). Tem apresentado um crescimento recente muito forte nesse mercado devido ao fato de seus preços ao consumidor serem mais competitivos que o GLP (uma vez realizados e amortizados os investimentos na rede de distribuição do gás).

5.1.4 Geração Elétrica/Co-geração

O gás natural também pode ser utilizado para a geração elétrica em usinas térmicas através de sua queima. Vale ressaltar que o principal tipo de usina térmica existente no Brasil (em capacidade instalada de potência) são as usinas térmicas a gás natural. Na indústria, o processo que aproveita o gás natural tanto para se gerar calor para o processo produtivo quanto para se gerar energia elétrica para a indústria é chamado de co-geração.

5.2 *Risco de Desabastecimento*

Um ponto preocupante em relação ao gás natural no Brasil atualmente é a estrutura de oferta de que se dispõe.

Oferta nacional de GN crescendo com importações

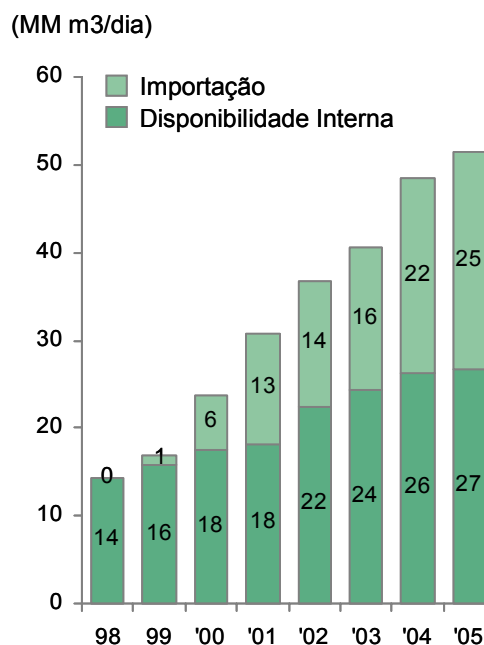


Gráfico 6: Oferta nacional de gás natural crescendo com importações

Fonte: Plano de Negócios 2006-2011 Petrobras. Elaborado pelo autor

Conforme o gráfico 6 ilustra, existe atualmente uma enorme dependência brasileira da importação boliviana de gás natural (cerca de 50% da oferta atual).

Tal fato se mostra preocupante devido à instabilidade política boliviana e ao histórico recente (nacionalização da produção do gás natural, greves em fábricas bolivianas de gás natural, renegociação de contratos com a Petrobras, etc.).

De acordo com informações divulgadas pela Petrobras, recentemente foi descoberto no Brasil o campo de Santos, enorme reservatório de gás natural com reserva de cerca de 400 bilhões de m³ de gás natural. (o que, por si só, seria suficiente para atender a demanda nacional de gás natural por cerca de 25 anos, caso se mantenha o nível de consumo atual de gás).

Porém, o desenvolvimento deste campo demanda alguns anos após sua descoberta. Dessa forma, o campo de Santos ainda não apresenta uma produção expressiva.

Apesar da existência do campo de Santos, é preciso se desenvolver desde já explorações em outros campos, visando encontrar novas fontes nacionais de gás natural, já que Santos, apesar de ser uma reserva muito grande para os padrões brasileiros deverá ser esgotada ainda nas próximas décadas.

De forma resumida, o cenário brasileiro possui um risco de déficit de gás natural no curto prazo bastante acentuado, devido ao crescimento não planejado da demanda e de problemas com seu maior fornecedor, a Bolívia.

No longo prazo, é muito difícil se discutir esse tema pelos seguintes motivos:

- Não existem previsões públicas de longo prazo da demanda de gás natural brasileira
- O contrato de fornecimento com a Bolívia vai até 2019. Após isso o mesmo deverá ser rediscutido e renegociado. Devido à instabilidade política na região, é difícil se prever se o mesmo será estendido ou não.
- A oferta nacional de gás natural depende da produção dos campos atualmente existentes e da descoberta de novos campos de gás.
- Existe a possibilidade de importação de Gás Natural Liquefeito (GNL) de outros países que não a Bolívia, os quais já estão previstos no curto prazo. Porém, ainda não há uma definição da Petrobras em relação à estratégia a ser adotada em relação ao GNL no médio e longo prazo (quantidade importada, fornecedores, locais de recebimento de carga, dentre outras variáveis).

Sendo assim, pode-se considerar o risco de desabastecimento de gás natural no longo prazo existente, porém atualmente imensurável.

6. Referencial Teórico

6.1 Visão Geral sobre previsões

Realizar previsões é uma das atividades gerenciais básicas de qualquer empresa, sendo utilizada na previsão futura de diversas variáveis, como: preços, demanda, oferta, vendas, lucro, etc.

Pode-se considerar como principal vantagem de se realizar uma previsão o fato de se conhecer, com certo grau de confiabilidade, o futuro com antecedência. Desta forma, podem-se planejar ações para responder a este futuro de forma a se minimizar perdas ou maximizar ganhos.

Um exemplo disso seria a realização de uma previsão de demanda e, posteriormente, de venda de determinado produto em determinado mercado, servindo como base para um Plano de Negócios do mesmo.

Nascido na década de 1950, o campo de previsões tem desenvolvido ferramentas para os mais diversos problemas e aplicações. Um panorama não-exaustivo das mesmas é apresentado na figura abaixo.

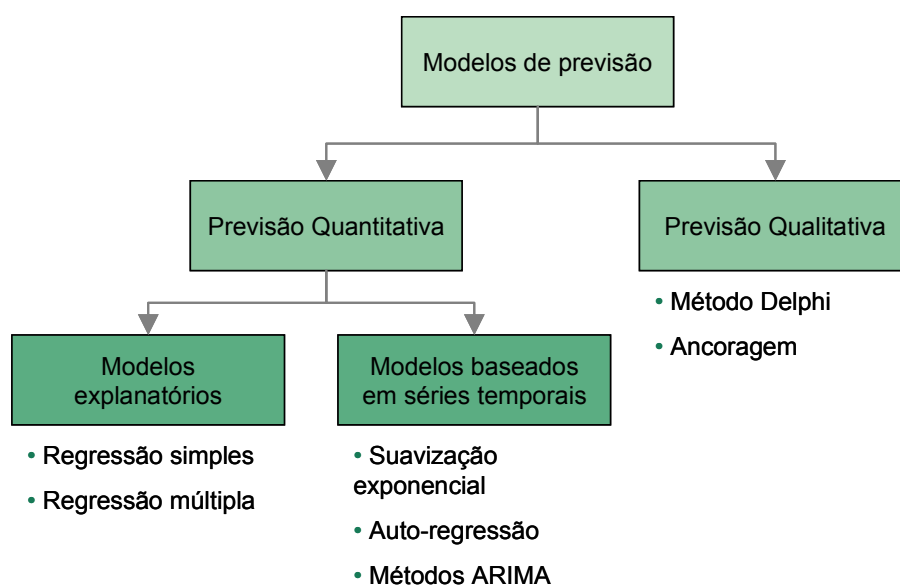


Figura 3: Visão geral dos modelos de previsão existentes

Conforme a figura acima ilustra, tem-se que as previsões podem ser feitas de duas formas básicas puras: quantitativamente ou qualitativamente.

Podem-se subdividir os modelos de previsão quantitativa em dois subgrupos: os modelos explanatórios e os modelos baseados em séries temporais.

De forma sucinta, os modelos explanatórios buscam prever os resultados futuros da variável de previsão através do *input* de valores futuros de variáveis relacionadas a esta num modelo matemático. Alguns exemplos de modelos de previsão explanatórios são a regressão simples, regressão múltipla e modelos econométricos (os quais fogem do escopo deste texto, não sendo assim aqui abordados).

Já os modelos baseados em séries temporais buscam compreender o comportamento passado de séries históricas e extrapolar o mesmo para o futuro, de forma a se obter uma previsão. Alguns exemplos de modelos de previsão baseados em séries temporais são os métodos de suavização exponencial, auto-regressão e os modelos ARIMA (*Autoregressive Integrated Moving Average*).

Previsão qualitativa é o nome que se dá a previsões que não são geradas diretamente de um modelo matemático, sendo geralmente elaboradas por especialistas (podendo contar com o auxílio ou não de modelos matemáticos). Um método bastante interessante de previsão qualitativa é o método DELPHI que, assim como os outros métodos acima mencionados serão mais bem detalhados posteriormente neste mesmo capítulo.

Vale ressaltar que o diagrama apresentado acima é uma classificação conceitual das diferentes formas de se fazer previsões existentes. É perfeitamente possível se fazer combinações de algumas das formas “puras” apresentadas anteriormente para se realizar a previsão de determinada variável num caso real (juntando-se, por exemplo, regressão simples com auto-regressão e ancoragem).

De acordo com Hanke (1998), o processo básico de previsão é ilustrado pelo seguinte fluxograma apresentado na figura a seguir:

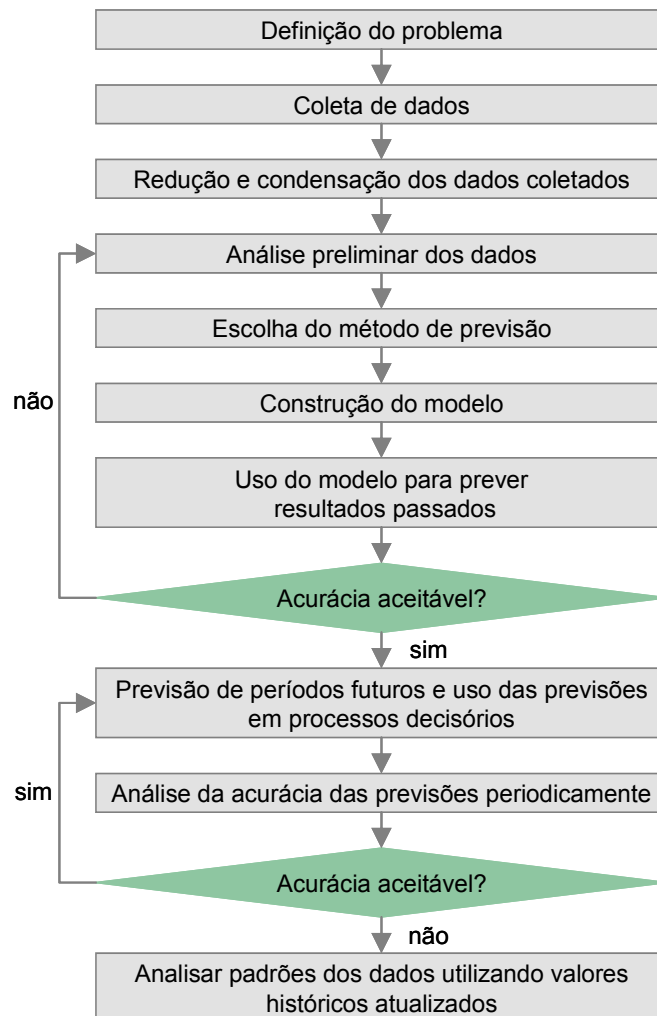


Figura 4: Processo de previsão

A seguir serão apresentados alguns conceitos básicos relacionados ao tema “Previsões”, serão mostradas algumas formas de se analisar a qualidade das previsões obtidas, se discorrerá mais profundamente sobre cada método de previsão citado neste item (quantitativos e qualitativos), e, finalmente, será apresentado um resumo dos métodos de previsão quantitativa apresentados neste texto, apresentando suas aplicabilidades específicas.

6.2 *Conceitos básicos relacionados a previsões*

6.2.1 Estatísticas básicas

Aqui serão apresentados alguns conceitos estatísticos básicos, cujo entendimento e compreensão são pré-requisitos para o adequado entendimento do trabalho que se segue.

Correlação

Correlação é uma estatística entre duas variáveis que indica o quanto as mesmas estão relacionadas entre si, ou seja, o quanto a variação de uma das variáveis afeta a variação de outra.

Vale destacar que isso não significa que as variáveis tenham alguma relação de causa e efeito entre si, podendo as mesmas ser relacionadas por uma terceira variável causal em comum ou até mesmo por mero acaso.

A correlação linear entre duas variáveis é calculada através da seguinte fórmula:

Equação 1: Correlação Linear

$$\text{Correl}(x, y) = \frac{\sum (x - \bar{x})(y - \bar{y})}{\sqrt{\sum (x - \bar{x})^2 \sum (y - \bar{y})^2}}$$

O gráfico a seguir ilustra os 3 tipos de correlação linear que podem ser encontrados entre duas variáveis: correlação positiva, correlação negativa e ausência de correlação.

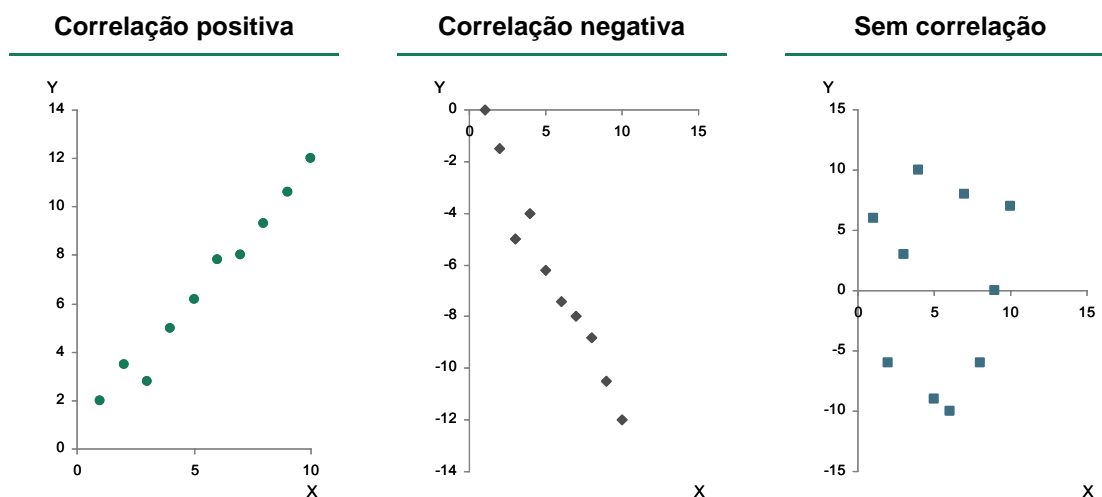


Gráfico 7: Tipos de correlação

Correlação positiva entre duas variáveis indica que quando uma aumenta, a outra também aumenta e quando uma diminui a outra também diminui. Correlação negativa entre duas variáveis quer dizer que quando uma aumenta a outra diminui e vice-versa. Já correlação nula (ou próxima de zero) representa que não há nenhuma relação estatística significativa entre as duas variáveis, sendo as mesmas independentes.

Vale ressaltar que duas variáveis podem ser correlacionadas de outra maneira que não a linear. Por isso, é sempre essencial se plotar um *scatter plot* das variáveis sendo estudadas antes de se iniciar o cálculo da correlação linear entre elas. O exemplo abaixo ilustra isso:

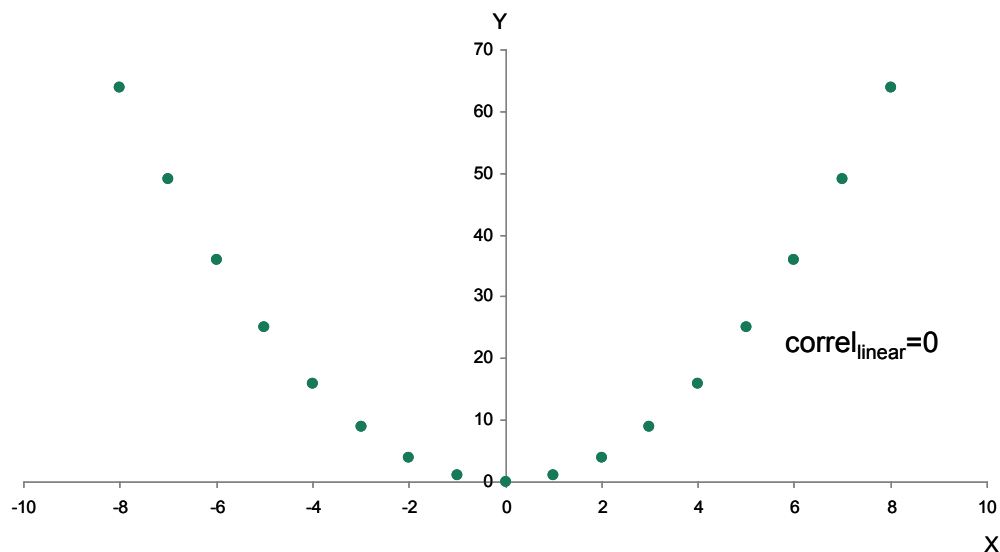


Gráfico 8: Correlação e relacionamento não-linear de variáveis

Como se pode observar, x e y possuem claramente um relacionamento quadrático. Porém, o cálculo da $\text{correl}_{\text{linear}}$ não demonstra isso, possuindo resultado nulo. Desta maneira, sempre é recomendado se plotar um *scatter plot* das variáveis sendo estudadas antes de se iniciar o estudo de correlações entre elas.

Autocorrelação

Autocorrelação é o análogo à correlação para o estudo de uma única variável.

Assim como a correlação estuda a relação entre duas variáveis, a autocorrelação (denominada pela letra r) estuda a relação entre uma série temporal e ela mesma, “atrasada” de um *lag* k .

A tabela abaixo ilustra o que representa o *lag* de uma série:

Tabela 1: *Lag* de uma série

Observação	Outcome	Outcome com lag 1	Outcome com lag 2
1	3	-	-
2	4	3	-
3	5	4	3
4	6	5	4
5	7	6	5
6	8	7	6
7	9	8	7
8	10	9	8
9	11	10	9
10	12	11	10

Desta forma, a autocorrelação de *lag* k é calculada da seguinte maneira:

Equação 2: Autocorrelação

$$r_k = \frac{\sum_{t=k+1}^n (Y_t - \bar{Y})(Y_{t-k} - \bar{Y})}{\sum_{t=1}^n (Y_t - \bar{Y})^2}$$

O *ACF plot* (do inglês *autocorrelation function*), também conhecido como correlograma, é uma representação gráfica que compila as diferentes autocorrelações de uma série temporal em seus diferentes *lags* em único gráfico.

O gráfico abaixo exemplifica isso:

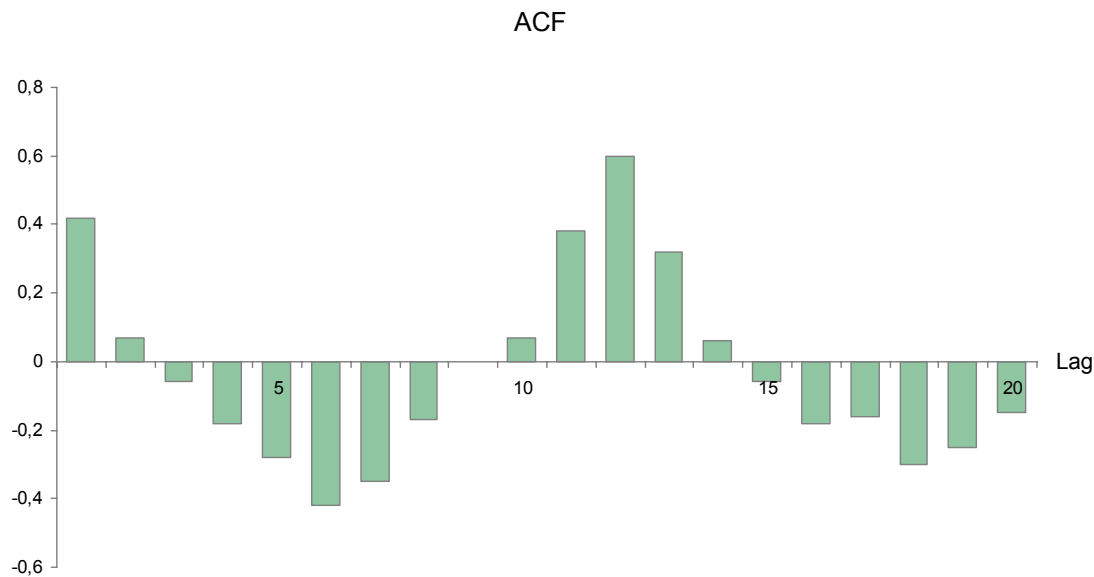


Gráfico 9: Exemplo de gráfico ACF

Desta forma, tem-se uma visualização rápida e prática das diferentes autocorrelações da série de uma só vez, facilitando-se a visualização de autocorrelações de maior valor, as quais indicam sazonalidade, o que é extremamente útil em diversos casos de estudo de séries temporais.

6.2.2 Padrões dos dados

Uma série de dados é composta, de forma genérica, de duas formas básicas de padrão: tendência-ciclicidade e sazonalidade, além de uma parcela de erro.

A figura abaixo ilustra a decomposição de uma série temporal em suas parcelas: tendência-ciclicidade, sazonalidade e erro (ou parte aleatória).

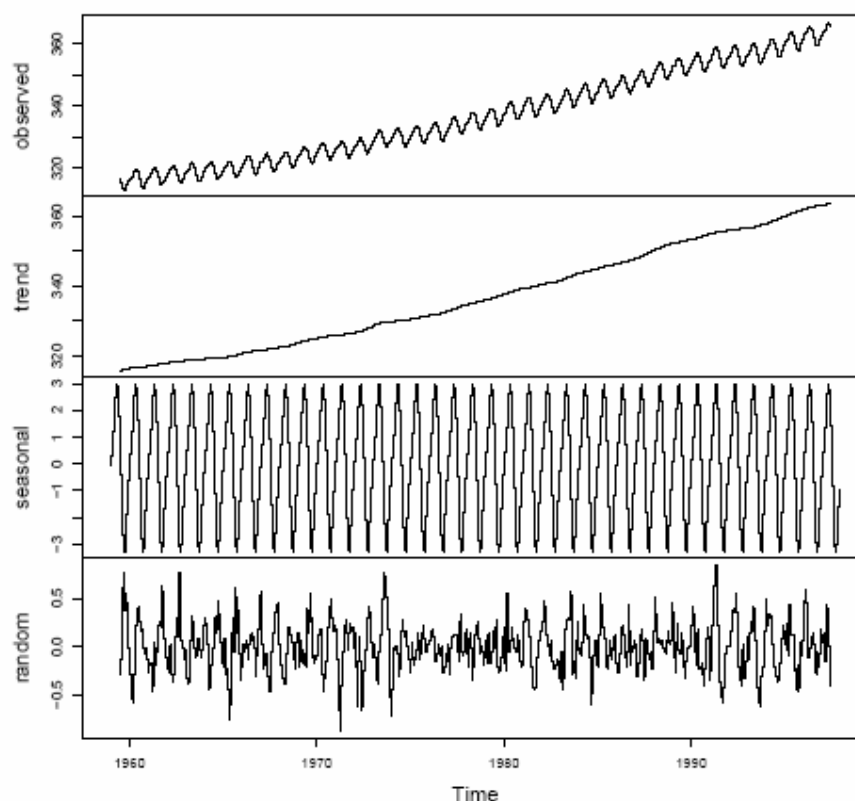


Figura 5: Decomposição de série temporal em suas parcelas

Fonte: Extraído de Makridakis (1998)

Tal fato é conceitualmente muito importante por duas principais razões.

A primeira delas é obter-se uma maior compreensão de uma série temporal, dividindo-a em suas partes básicas. Isso se mostra muito relevante em diversos problemas de previsão durante os estudos iniciais da série temporal que se deseja prever.

A segunda razão é que, compreender se uma série possui ou não tendência-ciclicidade e/ou sazonalidade é extremamente importante na hora de se escolher um modelo de previsão para a mesma, especialmente nos casos em que se está utilizando um método baseado em série temporal.

6.2.3 Medidas de erros

Existem diversas formas diferentes de se avaliar o erro de uma previsão.

Inicialmente, o erro de uma previsão é dado pela diferença entre o acontecimento real e a previsão realizada, conforme a fórmula abaixo:

Equação 3: Erro de uma previsão

$$e_t = Y_t - F_t$$

O erro mais simples de todos é denominado *Mean Error* (ME), que é calculado através da fórmula:

Equação 4: *Mean Error* (ME)

$$ME = \frac{1}{n} \sum e_t$$

Em seguida, tem-se um derivado do ME, que é o *Mean Absolute Error* (MAE), que utiliza os módulos dos erros, através da fórmula:

Equação 5: *Mean Absolute Error* (MAE)

$$MAE = \frac{1}{n} \sum |e_t|$$

Além destes dois, tem-se o muito utilizado *Mean Squared Error* (MSE), o qual utiliza a soma dos erros quadráticos:

Equação 6: *Mean Squared Error* (MSE)

$$MSE = \frac{1}{n} \sum (e_t)^2$$

Outro tipo de erro é o erro porcentual, calculado através do *Mean Percentage Error* (MPE), que é calculado através da fórmula:

Equação 7: *Mean Percentage Error* (MPE)

$$MPE = \frac{1}{n} \sum \left(\frac{Y_t - F_t}{Y_t} \right) \times 100$$

Analogamente ao *Mean Absolute Error* (MAE), existe também o *Mean Absolute Percentage Error* (MAPE), que segue a mesma lógica do MPE, mas utilizando os módulos dos erros percentuais no seu cálculo:

Equação 8: Mean Absolute Percentage Error (MAPE)

$$MAPE = \frac{1}{n} \sum \left| \frac{Y_t - F_t}{Y_t} \times 100 \right|$$

A seguir é apresentado um exemplo que mostra as diferenças de cálculo e significado de cada um dos erros.

Na tabela abaixo é apresentada uma série de 20 dados reais e suas respectivas previsões, assim como seus respectivos erros.

Tabela 2: Exemplificação dos tipos de erros

Período	Dado Real	Previsão	Erro	Erro absoluto	Erro Quadrático	Erro Porcentual	Erro Porcentual Absoluto
1	15,29	15	0,29	0,29	0,09	2%	2%
2	13,89	20	-6,11	6,11	37,36	-44%	44%
3	17,78	25	-7,22	7,22	52,2	-41%	41%
4	39,59	30	9,59	9,59	91,89	24%	24%
5	31,46	35	-3,54	3,54	12,57	-11%	11%
6	35,42	40	-4,58	4,58	20,96	-13%	13%
7	45,49	45	0,49	0,49	0,24	1%	1%
8	54,78	50	4,78	4,78	22,89	9%	9%
9	48,72	55	-6,28	6,28	39,42	-13%	13%
10	62,04	60	2,04	2,04	4,16	3%	3%
11	63,55	65	-1,45	1,45	2,1	-2%	2%
12	77,93	70	7,93	7,93	62,93	10%	10%
13	77,24	75	2,24	2,24	5,04	3%	3%
14	83,04	80	3,04	3,04	9,26	4%	4%
15	90,29	85	5,29	5,29	28,01	6%	6%
16	87,09	90	-2,91	2,91	8,5	-3%	3%
17	104,84	95	9,84	9,84	96,75	9%	9%
18	90,46	100	-9,54	9,54	91,02	-11%	11%
19	101,15	105	-3,85	3,85	14,8	-4%	4%
20	107,49	110	-2,51	2,51	6,31	-2%	2%

Plotando-se os gráficos reais e suas previsões obtêm-se o seguinte gráfico:

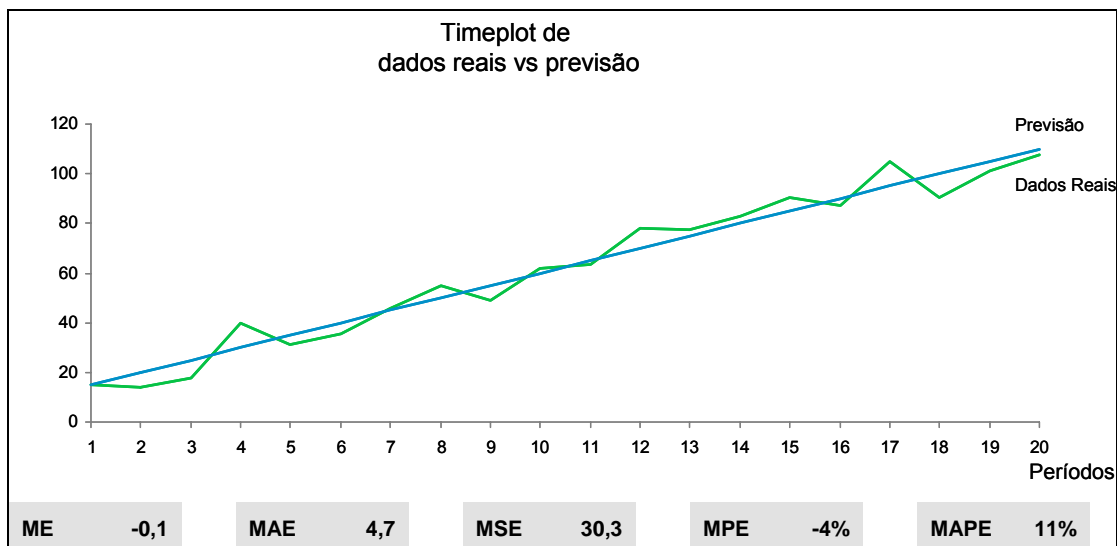


Gráfico 10: Exemplificação dos tipos de erros

No gráfico acima se pode observar as diferenças entre os dados reais e a previsão realizada (que, de uma forma geral, acompanha bastante bem os dados reais), assim como os respectivos erros da previsão – ME, MAE, MSE, MPE e MAPE.

6.2.4 Intervalos de previsão

Toda previsão, seja ela quantitativa ou qualitativa, é uma estimativa de um futuro incerto. Desta forma, ela sempre contém uma incerteza associada à mesma.

Por isso, muitas vezes o resultado de uma previsão não é apresentado como um número, mas sim como um intervalo de valores, dentro do qual o valor futuro da variável sendo prevista se encontrará com certo grau de confiabilidade.

Um exemplo disso seria, ao invés de dizer que a população brasileira em 2030 será de 238 milhões de pessoas, dizer que a população brasileira em 2030 estará entre 236 e 240 milhões de pessoas com 95% de certeza.

Além de deixar explícito o grau de incerteza associada à previsão, a utilização de intervalos também traz certo “alívio” psicológico para a mesma, porque tira dela o estigma de

estar “errada”, quando o resultado real não se iguala exatamente ao valor discreto que havia sido previsto.

O intervalo de previsão para uma previsão realizada para o próximo período é dado por:

Equação 9: Intervalo de previsão para próximo período

$$F_{n+1} \pm z\sqrt{MSE}$$

Onde F_{n+1} é a previsão para o próximo período, MSE é o *Mean Squared Error* e z é o valor associado ao nível de confiança que se deseja para o intervalo a ser determinado.

A tabela abaixo apresenta alguns valores de z associados aos seus respectivos níveis de confiança:

Tabela 3: Valores de z associados a níveis de confiança

z	Probabilidade	z	Probabilidade
0,674	0,5	1,96	0,95
0,842	0,6	2,054	0,96
1	0,68	2,326	0,98
1,036	0,7	2,576	0,99
1,15	0,75	2,807	0,995
1,282	0,8	3,09	0,998
1,645	0,9	3,29	0,999

Como se pode observar na tabela acima, quanto maior o intervalo de confiança desejado, ou seja, a certeza que se queira ter que o intervalo de previsão conterá o valor futuro da variável, maior o z e, conseqüentemente, maior será o intervalo de previsão.

Porém, o cálculo apresentado acima só serve para previsões baseadas em séries temporais para um período no futuro, já que o MSE é baseado em previsões com um período para o futuro.

No caso das regressões, tem-se as seguintes equações para intervalos de confiança:

Equação 10: Cálculo do intervalo de confiança para regressão simples e múltipla

$$F_n \pm z \times \sigma(Y_0)$$

No caso da regressão simples, o σ é calculado da seguinte maneira:

Equação 11: Cálculo do sigma da regressão simples

$$\sigma(Y_0) = \sigma_e \times \sqrt{1 + \frac{1}{n} + \frac{(X_0 - \bar{X})^2}{\sum_i (X_i - \bar{X})^2}}$$

Onde σ_e é o desvio-padrão dos erros, n o número de dados, \bar{X} a média dos X utilizados no modelo de regressão e X_0 o valor para o qual se quer calcular o intervalo de confiança.

No caso da regressão múltipla o σ é calculado através da seguinte equação:

Equação 12: Cálculo do sigma da regressão múltipla

$$\sigma(Y_0) = \sigma_e \times \sqrt{1 + c'(X'X)^{-1}c}$$

É importantíssimo ressaltar que as equações de intervalo de previsão aqui apresentadas possuem hipóteses por trás de sua formulação (idênticas às hipóteses de regressão), as quais são:

- Os erros (ou resíduos) possuem distribuição normal, com média zero e variância constante para todo x_i (ao longo de todo o espectro da variável explanatória)
- Quaisquer observações independentes não podem ser correlacionadas entre si

Desta forma, antes de usá-las deve-se verificar se as hipóteses por trás delas são válidas no problema sendo estudado.

6.2.5 Hipótese da Continuidade

Todos os métodos de previsão quantitativa, sejam eles explanatórios ou baseados em séries temporais, contém alguma espécie de Hipótese de Continuidade por trás deles.

Os métodos explanatórios de regressão, ao fazer uma previsão, assumem que os relacionamentos entre variáveis explanatórias e variável de previsão se mantêm constantes ao longo do tempo.

Já os métodos baseados em séries temporais assumem que o padrão da série (o qual estes métodos buscam, de alguma forma, extrapolar) será mantido no futuro.

Como se pode notar com alguma análise crítica, nem sempre estas hipóteses são válidas ou mesmo coerentes. Mudanças tecnológicas, políticas ou de mercado podem fazê-las ruir, acabando com a validade da previsão realizada. Desta forma, é sempre essencial ter em mente quais as hipóteses que estão por trás de um modelo e validá-las periodicamente para se averiguar se as mesmas ainda são válidas.

Os efeitos de grandes rupturas ou mudanças como, por exemplo, os impactos da informática na produtividade das tarefas administrativas não podem ser previstos de forma adequada por métodos quantitativos. Apenas os métodos qualitativos (previsões) são capazes de incorporar mudanças de tal magnitude às suas previsões.

6.3 Modelos de previsão

Nesta seção serão apresentados de forma resumida os principais modelos de previsão existentes, seu funcionamento e aplicações mais recomendadas.

Para um maior aprofundamento do tema recomenda-se a consulta à bibliografia indicada ao final deste trabalho.

6.3.1 Previsão Quantitativa

Serão aqui apresentados os métodos de previsão quantitativa, baseados em dados históricos e cálculos para se realizar previsões.

Os mesmos podem ser classificados em dois grupos principais: os métodos de previsão quantitativa explanatória, que visam prever o comportamento futuro de uma variável através de seu relacionamento histórico com *drivers*, e os métodos de previsão quantitativa baseados em séries temporais, os quais visam prever o comportamento futuro de uma variável estudando seu comportamento passado (isoladamente, sem relacioná-lo a nenhuma outra variável).

6.3.1.1 Previsão Quantitativa Explanatória

Serão aqui apresentados os dois principais métodos de previsão quantitativa explanatória: a regressão simples e a regressão múltipla.

6.3.1.1.1 Regressão Simples

Regressão simples é um método de previsão quantitativa explanatória que segue o seguinte formato:

Equação 13: Forma geral da regressão simples

$$Y = f(x)$$

Ou seja, a previsão será uma função do resultado futuro de outra variável, no caso o x .

Escrevendo a forma da regressão simples mais detalhadamente para o caso linear, tem-se:

Equação 14: Forma linear da regressão simples

$$Y_p = a + bx$$

Ou seja, Y_p , a previsão, é uma função linear de x . Já Y , o resultado real da variável no futuro seria Y_p mais um valor randômico de erro.

Para o cálculo dos coeficientes recomenda-se a leitura de MAKRIDAKIS, S. and S.C. WHEELWRIGHT. Forecasting – Methods and Applications, 3ª ed., Wiley, New York, 1998.

Vale destacar que praticamente qualquer software usado para se trabalhar dados, inclusive o MS Excel, realiza esses cálculos automaticamente.

A regressão simples é baseada na correlação linear (apresentada anteriormente neste trabalho). Ao elevarmos o coeficiente da correlação linear, R , ao quadrado, obtém-se o R^2 que indica a qualidade da regressão realizada. Este indicador varia de 0 a 1 e, quanto mais próximo de 1, melhor é a regressão sendo realizada.

Um ponto essencial de ser destacado são as hipóteses que estão por trás do modelo de regressão, as quais não sendo atendidas diminuem em muito o rigor matemático do modelo.

- Os erros (ou resíduos) possuem distribuição normal, com média zero e variância constante para todo x_i (ao longo de todo o espectro da variável explanatória)
- Quaisquer observações independentes não podem ser correlacionadas entre si

Estatisticamente pode-se avaliar a qualidade da regressão através da estatística t de cada coeficiente. De uma forma geral, considera-se que resultados de t , em módulo, maiores do que 2 são bons e que valores entre 1,5 e 2 são considerados aceitáveis (cerca de 85% a 95% de certeza que o coeficiente não é zero).

Segue um exemplo de regressão realizada no Excel, com base nos seguintes dados:

Tabela 3: Exemplo de regressão linear

x	y	x	y
1	3	16	39
2	10	17	50
3	7	18	35
4	17	19	45
5	12	20	44
6	19	21	47
7	12	22	54
8	23	23	43
9	18	24	55
10	24	25	50
11	25,6	26	62
12	34	27	59
13	25	28	50
14	28	29	62
15	37	30	70

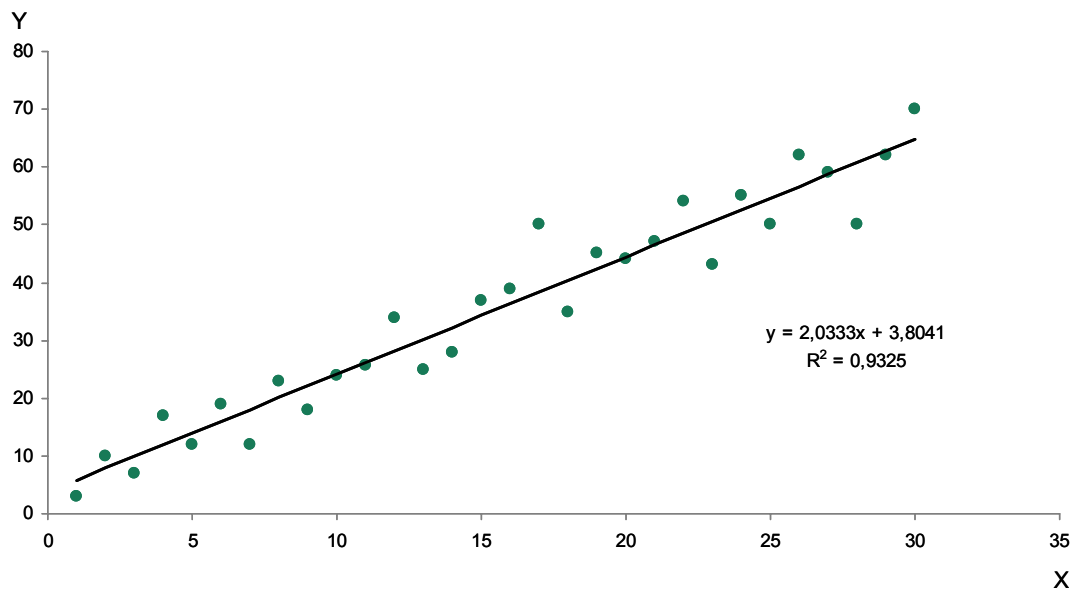


Gráfico 11: Exemplo de regressão linear

Vale ressaltar que, além da regressão linear, duas variáveis podem ter relacionamentos de diversos outros tipos, como: quadrática, cúbica, exponencial, logarítmica, etc. O cálculo do R e dos coeficientes para cada caso varia e não será abordado neste trabalho, porém os

mesmos são automaticamente calculados por praticamente qualquer software usado para se trabalhar dados estatisticamente.

6.3.1.1.2 Regressão múltipla

Regressão múltipla é um método de previsão quantitativa explanatória que segue o seguinte formato:

Equação 15: Forma geral da regressão múltipla

$$Y = f(x_1, x_2, x_3, \dots, x_n)$$

Ou seja, a previsão será uma função dos resultados futuro de outras variáveis, no caso os x_i .

Escrevendo a forma da regressão múltipla mais detalhadamente para o caso linear, tem-se:

Equação 16: Forma linear da regressão múltipla

$$Y_p = b_0 + b_1X_1 + b_2X_2 + \dots + b_nX_n$$

Ou seja, Y_p , a previsão, é uma função linear das variáveis explanatórias x_i . Já Y , o resultado real da variável no futuro seria Y_p mais um valor randômico de erro.

Assim como a regressão simples, a regressão múltipla também possui as hipóteses de que seus resíduos possuem distribuição normal com média zero e variância constante e que observações independentes não podem ser correlacionadas entre si. Caso estas hipóteses não se mostrem verdadeiras, o modelo perde muito de sua robustez matemática.

A qualidade da regressão também pode ser avaliada através do R^2 . Porém, no caso da regressão múltipla encontra-se um problema de que quanto mais variáveis forem adicionadas ao modelo, maior será o R^2 , isso porque maior será o nível de explicação do fenômeno.

Desta forma, deve-se utilizar uma variável diferente para avaliar adequadamente a qualidade da regressão sendo realizada, o R^2 ajustado.

Equação 17: R^2 ajustado

$$\overline{R^2} = 1 - (1 - R^2) \frac{(n - 1)}{(n - k - 1)}$$

Onde n é o número de observações e k é o número de variáveis explanatórias incluídas no modelo.

Assim, quanto maior for o R^2 ajustado encontrado, melhor é a qualidade da regressão realizada.

A seguir é apresentado um fluxograma sugerido por Makridakis (1998) para seleção de variáveis para o modelo de regressão múltipla:



Figura 6: Escolha de variáveis para regressão múltipla

Um dos métodos mais interessantes e recomendados para formar a *short list* de variáveis explanatórias é o *Stepwise Regression*. Existem 3 formas de fazê-lo (*stepwise forward regression*, *stepwise backward regression* e *stepwise forward-with-a-backward-look*

regression). Por ser o que, de acordo com Makridakis (1998), apresenta os melhores resultados, este trabalho utiliza o *Stepwise forward-with-a-backward-look regression*, cuja metodologia é apresentada abaixo.

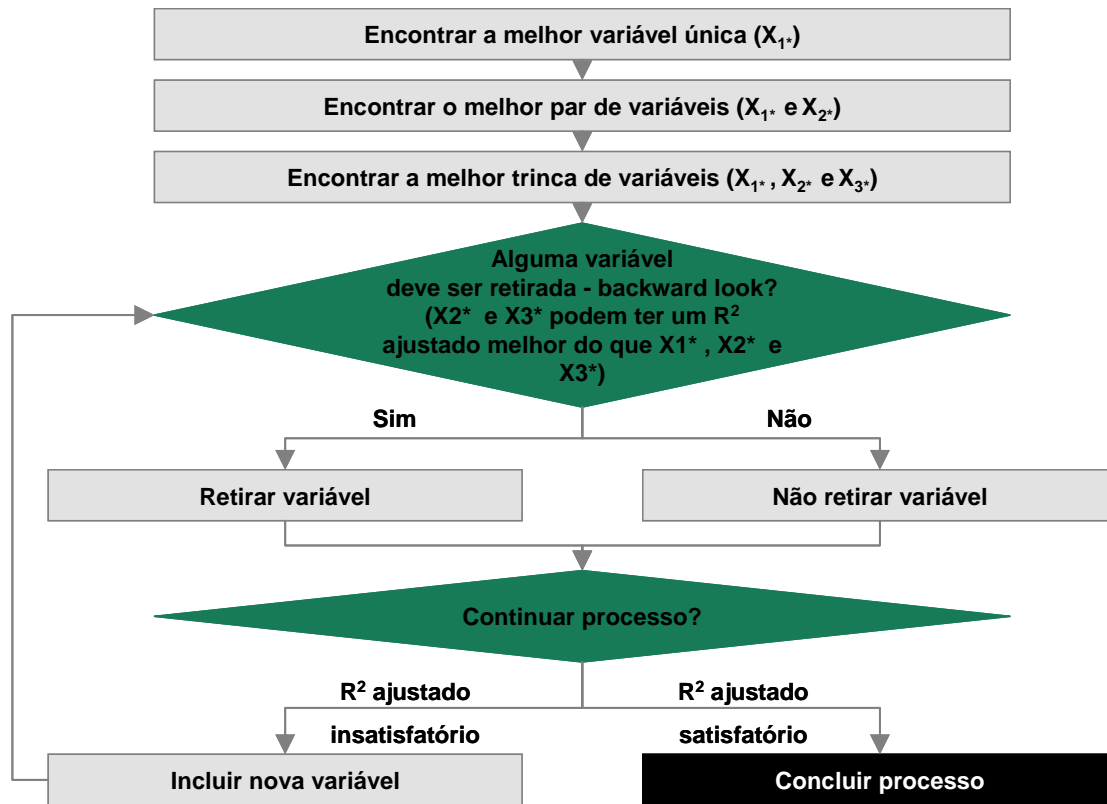


Figura 7: *Stepwise forward-with-a-backward-look regression*

Um grande problema encontrado em regressões múltiplas é o de multicolinearidade de variáveis explanatórias. Ou seja, quando tem-se uma variável explanatória com alta correlação com outra, ou mesmo com uma combinação linear de outras, a regressão fica inválida. Desta forma, é extremamente importante se avaliar, *a priori*, se as variáveis são independentes entre si.

6.3.1.2 Previsão Quantitativa baseada em séries temporais

Serão aqui apresentados os principais métodos de previsão quantitativa baseada em séries temporais: naive, estudo de séries temporais, média simples, média móvel modelos de suavização exponencial, auto-regressão e modelos ARIMA.

6.3.1.2.1 Naive

O método mais simples de previsão baseada em séries temporais é o chamado Naive (“Ingênuo”, numa tradução literal). Este método considera simplesmente que a estimativa para o próximo resultado é o último resultado ocorrido.

Em linguagem matemática tem-se que:

Equação 18: Previsão Naive

$$F_n = Y_{n-1}$$

Onde F_n é a previsão e Y_{n-1} é o último resultado ocorrido. Para ilustrar o funcionamento da Previsão Naive é apresentado um gráfico comparando tal previsão com o realizado.

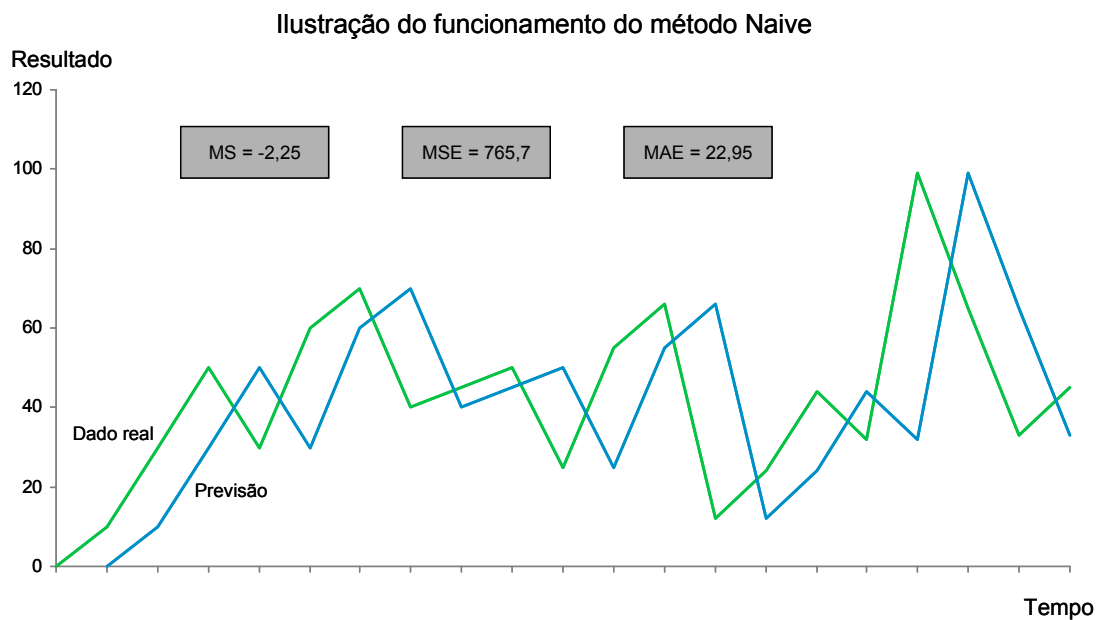


Gráfico 12: Ilustração do funcionamento do método Naive

Note-se que este método é extremamente simplista, possuindo assim baixa aplicabilidade. Além disso, ele não possui nenhum método ou forma de aprendizado com seus erros.

Desta forma serve basicamente como base comparativa de qualidade para outros modelos e como base para a Suavização Exponencial, a qual será apresentada mais adiante neste trabalho.

6.3.1.2.2 Decomposição de séries temporais

Conforme visto no item 6.2.3. uma série temporal é composta de três parcelas principais: tendência/ciclicidade, sazonalidade e erro. Desta forma, decompondo-se uma série temporal nestes três fatores podem-se realizar previsões para os dois primeiros (já que a parcela de erro não pode ser prevista por ser randômica) via qualquer outro método de previsão como auto-regressão ou suavização exponencial e depois juntá-los de forma a se obter uma previsão para a série como um todo.

Matematicamente, pode-se escrever o parágrafo acima da seguinte maneira:

Equação 19: Decomposição aditiva

$$Y_t = S_t + T_t + E_t$$

ou

Equação 20: Decomposição multiplicativa

$$Y_t = S_t \times T_t \times E_t$$

Nas duas equações acima é possível observar como, analiticamente, dividiu-se a série temporal em suas três componentes. A diferença entre as duas equações acima está em seu uso.

A primeira equação é mais apropriada para séries temporais em que a sazonalidade (ou o nível da sazonalidade), não depende do nível da série. Isso pode ser exemplificado pela figura abaixo.

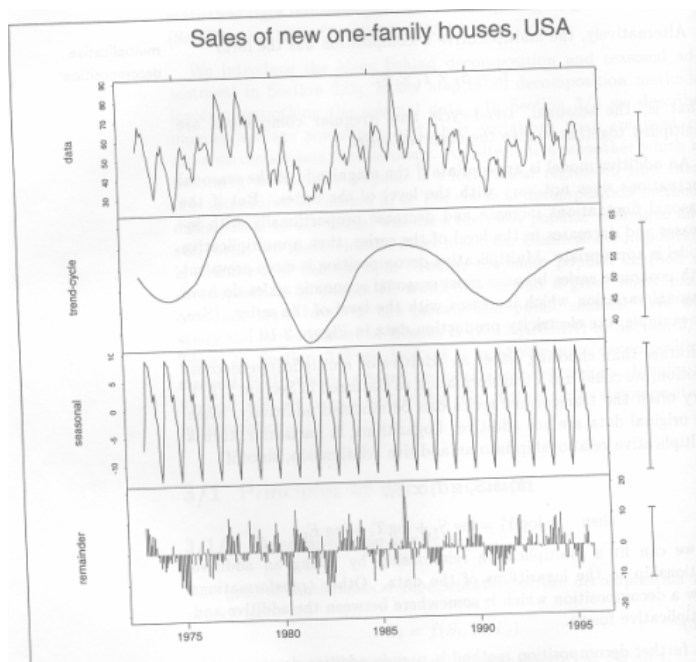


Figura 8: Exemplo de decomposição de série temporal com sazonalidade aditiva

Fonte: Extraído de Makridakis (1998)

Já a segunda equação é mais apropriada para séries temporais em que a sazonalidade (ou o nível da sazonalidade), depende do nível da série. Isso pode ser exemplificado pela figura abaixo, na qual a sazonalidade aumenta conforme o nível da série aumenta.

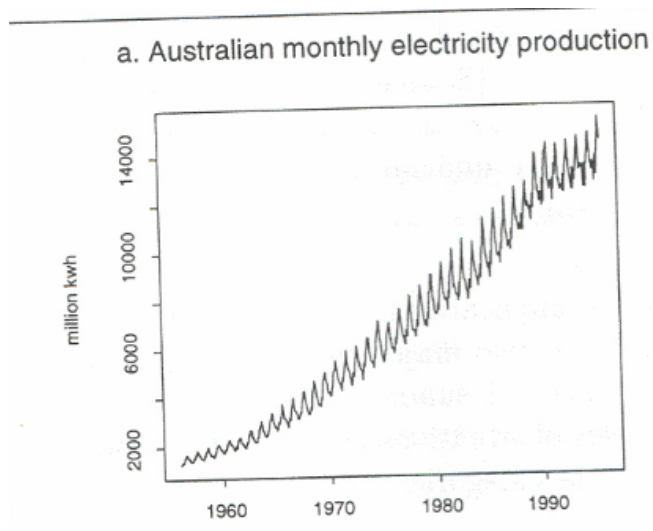


Figura 9: Exemplo de série temporal com sazonalidade multiplicativa

Fonte: Extraído de Makridakis (1998)

A figura a seguir apresenta o passo-a-passo para se realizar uma decomposição aditiva e uma posterior previsão com base nela:

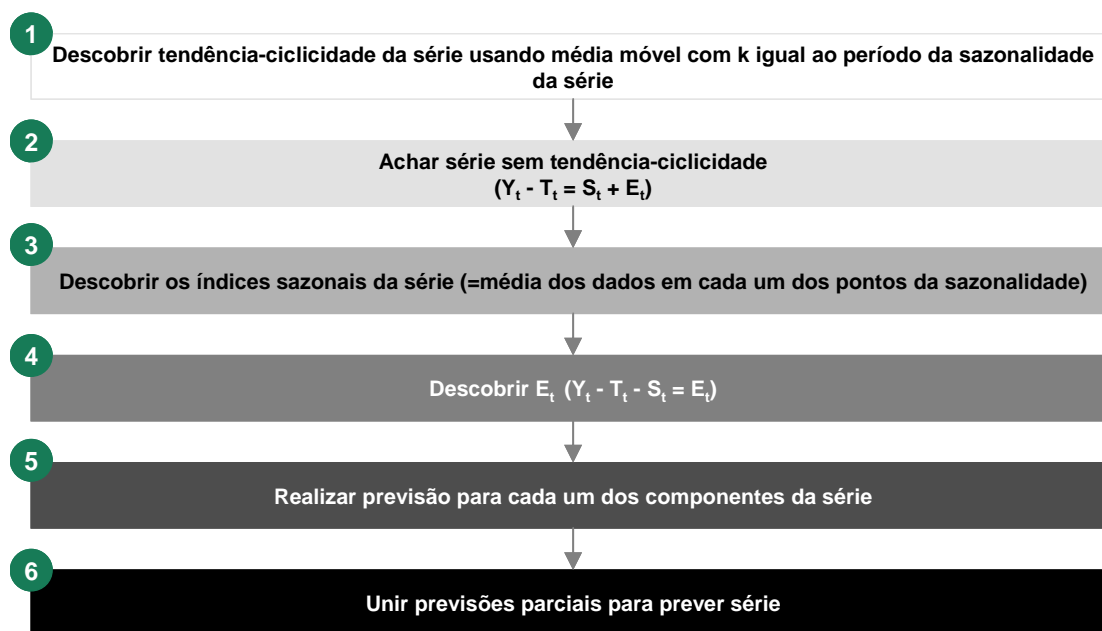


Figura 10: Passo-a-passo de decomposição aditiva com posterior previsão

Já a próxima figura apresenta o passo-a-passo para se realizar uma decomposição multiplicativa e uma posterior previsão com base nela:

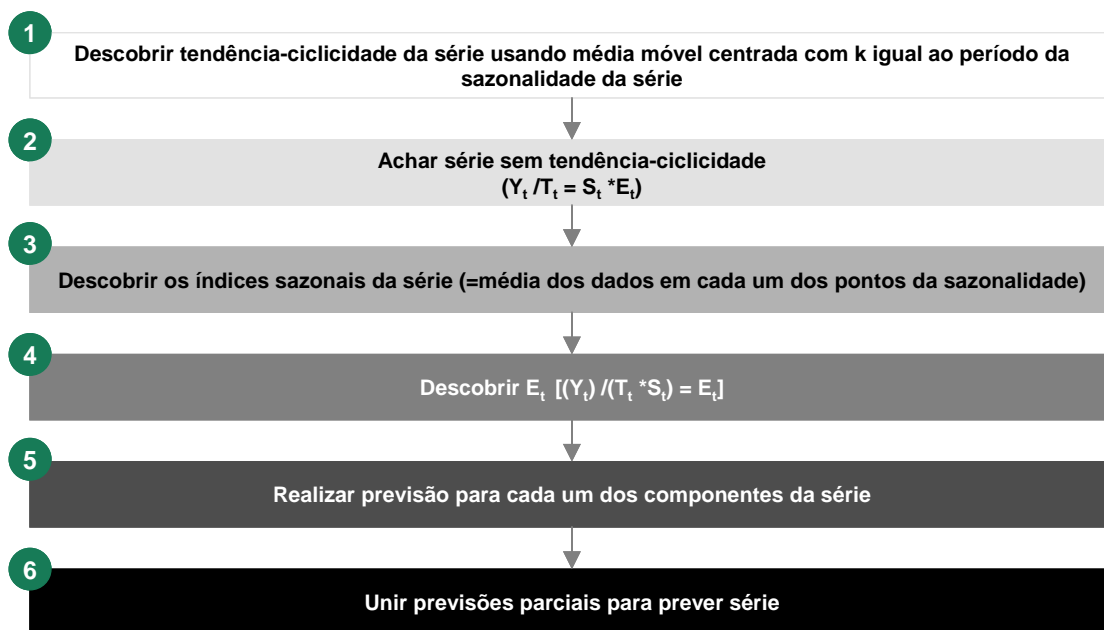


Figura 11: Passo-a-passo de decomposição multiplicativa com posterior previsão

6.3.1.2.3 Média Simples

Este método considera que a previsão para o próximo período é a média dos resultados dos períodos anteriores. Matematicamente, tem-se:

Equação 21: Previsão por média simples

$$F_{t+1} = \frac{1}{t} \sum_{i=1}^t Y_i$$

Porém, este método, por sua grande simplicidade, não apresenta bons resultados quando a série sendo prevista possui uma forte tendência/ciclicidade ou sazonalidade.

6.3.1.2.4 Média Móvel

Este método considera que a previsão para o período é a média dos últimos k períodos.

O funcionamento deste método é ilustrado na tabela abaixo (que possui os mesmos dados de Y_n que o exemplo da previsão Naive), onde $k = 3$.

Tabela 4: Ilustração do funcionamento do método de média móvel

Yn	Fn	Erro	Erro absoluto	Erro Quadrático
0				
10				
30	13	-17	17	278
50	30	-20	20	400
30	37	7	7	44
60	47	-13	13	178
70	53	-17	17	278
40	57	17	17	278
45	52	7	7	44
50	45	-5	5	25
25	40	15	15	225
55	43	-12	12	136
66	49	-17	17	300
12	44	32	32	1045
24	34	10	10	100
44	27	-17	17	300
32	33	1	1	2
99	58	-41	41	1654
65	65	0	0	0
33	66	33	33	1067
45	48	3	3	7

ME	-1,81
MSE	334,9
MAE	14,89

Ilustração do funcionamento do método de Média Móvel

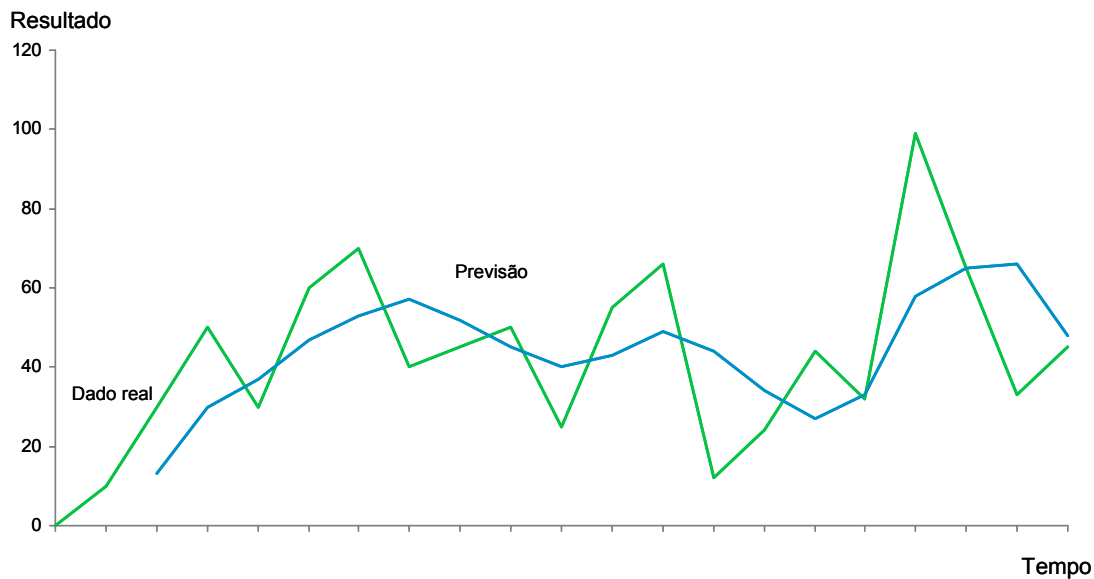


Gráfico 13: Ilustração do funcionamento do método de Média Móvel

Como se pode notar pelo gráfico, o método de média móvel suaviza as oscilações dos dados reais, diminuindo seus picos e vales. Quanto maior for a ordem da média móvel (o seu k), mais suavizada ficará essa curva (e quanto menor for o k , mais fortemente a previsão oscilará).

Pode-se fazer uma analogia de que o método de previsão “Ingênuo” é matematicamente o mesmo que um método de Média Móvel com $k = 1$.

Apesar de este método ser relativamente mais complexo que o de média simples, apresenta os mesmos problemas de utilização que este último, não sendo adequado para séries que possuam forte tendência/ciclicidade ou sazonalidade. Por outro lado ele consegue acompanhar “saltos” nas observações, apesar de fazê-lo com certo atraso.

6.3.1.2.5 Métodos de Suavização Exponencial

Os métodos de suavização exponencial são métodos quantitativos baseados em séries temporais geralmente utilizados para se realizar previsões de curto prazo (pelo fato de seus resultados irem se deteriorando conforme aumenta o horizonte de previsão) para uma grande quantidade de itens ao mesmo tempo (devido ao seu baixo consumo de memória computacional).

O funcionamento dos métodos de suavização exponencial é semelhante ao do baseado em média móvel, usando dados do passado para prever o futuro. A diferença entre estes é que os métodos de suavização exponencial ponderam as observações passadas dando maior peso às observações mais recentes, considerando que as mesmas são mais relevantes para explicar o futuro do que observações mais antigas.

Os três métodos principais de suavização exponencial são o Single Exponential Smoothing (SES), o método de Holt e o Método de Holt-Winters.

O método de Single Exponential Smoothing é o mais simples dos três, sendo mais recomendado para situações em que a variável de previsão não apresenta nem tendência/ciclicidade, nem sazonalidade.

Sua fórmula de cálculo é a seguinte:

Equação 22: Single Exponential Smoothing (SES)

$$F_{t+1} = \alpha Y_t + (1 - \alpha)F_t$$

Onde α pertence ao intervalo entre 0 e 1.

A equação acima diz que uma nova previsão é calculada com base na última observação e na última previsão ajustada pelo erro da mesma.

O termo α geralmente é determinado de forma que o mesmo minimize um indicador de erro da previsão, sendo este mais comumente o MSE (podendo ser também o MAPE, ME ou mesmo outro). Fazendo um paralelo entre média móvel e SES, tem-se que um α alto é semelhante a uma média móvel com ordem baixa, reagindo rapidamente a variações nas observações. Já um α baixo é semelhante a uma média móvel com ordem alta, se tornando bastante estável ao longo do tempo, demorando bastante para reagir a variações nas observações.

Para se iniciar as previsões, pode se adotar F_1 como Y_0 ou como a média das últimas n observações.

Já o método de Holt é mais adequado para lidar com variáveis de previsão que apresentem tendência/ciclicidade, mas não sazonalidade. Sua fórmula de cálculo é a seguinte:

Equação 23: Método de Holt

$$L_t = \alpha Y_t + (1 - \alpha)(L_{t-1} + b_{t-1})$$

$$b_t = \beta(L_t - L_{t-1}) + (1 - \beta)b_{t-1}$$

$$F_{t+m} = L_t + b_t m$$

Na equação acima a previsão, F_{t+m} , é baseada na estimativa do nível da série no instante t , L_t , mais a inclinação da reta, b_t , multiplicada pelo número de períodos m .

Os índices α e β podem variar entre 0 e 1 e devem ser encontrados de forma a minimizar o erro da previsão (via algoritmo de otimização de Pesquisa Operacional, cálculo diferencial ou “chutes” de pares ordenados).

Já o método de Holt-Winters é mais adequado para lidar com séries que apresentem tendência/ciclicidade e sazonalidade (seja ela aditiva ou multiplicativa). Devido à sua

complexidade e ao fato de não se aplicar ao problema sendo estudado neste Trabalho de Formatura, o mesmo não será mais detalhado neste relatório.

6.3.1.2.6 Auto-regressão

Auto-regressão é um método de previsão quantitativa baseado em séries temporais que segue o seguinte formato:

Equação 24: Forma geral da auto-regressão

$$Y_n = f(Y_{n-1}; Y_{n-2}; Y_{n-3} \dots)$$

Ou seja, a previsão do comportamento futuro será uma função dos resultados passados desta mesma série.

Escrevendo a forma da auto-regressão mais detalhadamente para o caso linear, tem-se:

Equação 25: Forma linear da auto-regressão

$$Y_n = a + b_1 \times Y_{n-1} + b_2 \times Y_{n-2} + b_3 \times Y_{n-3} + \dots$$

A auto-regressão também apresenta o indicador R^2 (no caso de auto-regressões de primeira ordem) ou R^2 -ajustado (no caso de auto-regressões de maior ordem), os quais indicam a qualidade da auto-regressão realizada. Além disso, a auto-regressão também possui as mesmas hipóteses da regressão sobre normalidade dos resíduos.

6.3.1.2.7 Modelos ARIMA

Modelos ARIMA são um conjunto de modelos de previsão qualitativos baseados em séries temporais, cujo nome é uma sigla que significa *Autoregressive/Integrated/Moving Average*. Ou seja, são modelos baseados em auto-regressões e/ou diferenciações (*integrated*) e/ou médias móveis.

Esse conjunto de modelos apresenta uma enorme quantidade de integrantes, mas todos eles são formados a partir dos três itens discriminados acima e possuem a mesma característica comum a todos os modelos de previsão baseados em série temporal, se mostrando, desta forma, mais apropriados para previsões de curto prazo.

Devido à sua grande versatilidade, tais modelos são de grande utilidade num problema de previsão. Porém, devido à grande complexidade da metodologia e de inúmeras variantes na mesma, será apresentada somente uma versão genérica simplificada da mesma.

Box e Jenkins (1970) apresentaram a seguinte metodologia para a modelagem geral de séries temporais:

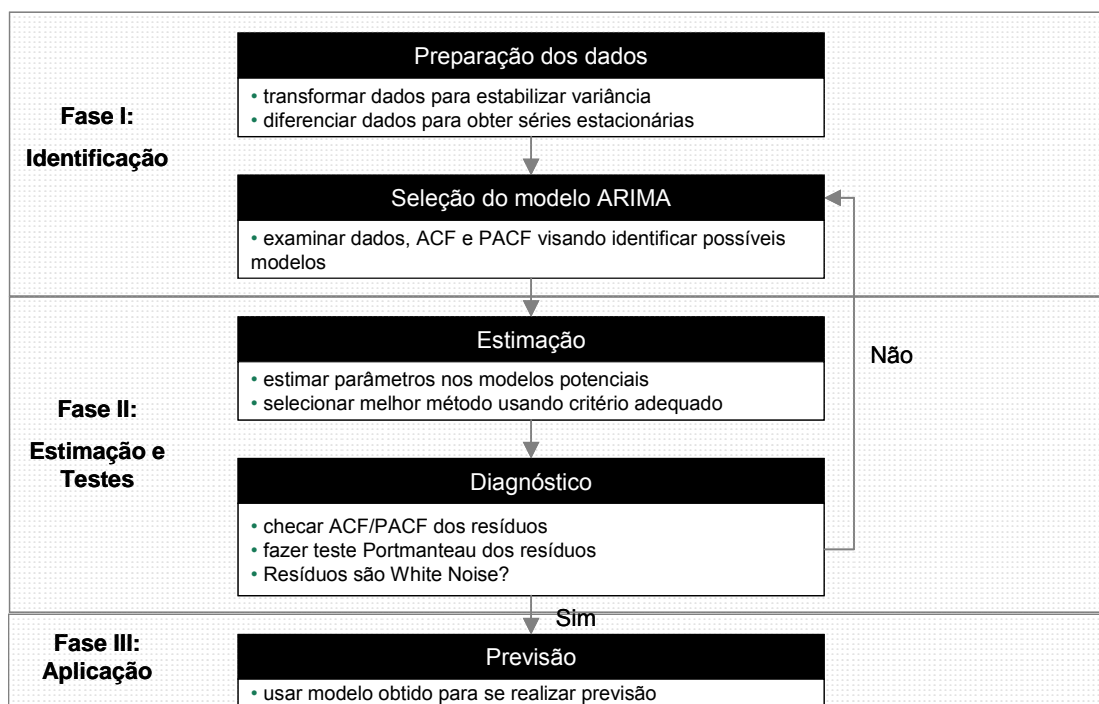


Figura 12: Metodologia Box-Jenkins para modelagem de séries temporais

Como a figura acima ilustra, a primeira etapa da modelagem de série temporal pela metodologia Box-Jenkins é preparar os dados transformando-os em uma série estacionária, tanto na média (sem tendência de crescimento ou declínio) e na variância (que deve se manter a mesma ao longo do tempo). Isso pode ser feito através de uso de médias móveis ou da diferenciação da série temporal:

Equação 26: Diferenciação de nível um de uma série temporal

$$Y'_t = Y_t - Y_{t-1}$$

A equação acima ilustra uma diferenciação de nível um, porém também existem diferenciações de maior nível (que possuem cálculos análogos ao acima apresentado) e também diferenciações sazonais, estas últimas representadas na equação abaixo:

Equação 27: Diferenciação sazonal de uma série temporal

$$Y'_t = Y_t - Y_{t-s}, \text{ onde } s \text{ representa o período da sazonalidade da série}$$

Uma vez tornada a série temporal estudada estacionária, utilizam-se as funções de ACF (*autocorrelation function*) e PACF (*partial autocorrelation function*, análogo ao ACF, mas que isola a atuação de cada *lag*) de modo a selecionar quais *lags* da série temporal devem ser usados para montar a equação geral de auto-regressão.

Feito isso, estima-se os coeficientes da equação de auto-regressão, verifica-se se os resultados estão de acordo com as hipóteses por trás da metodologia e obtém-se o modelo que representa a série temporal. De posse do modelo o mesmo é utilizado para realizar previsões sobre resultados futuros da série temporal.

Um modelo ARIMA possui a seguinte forma geral:

Equação 28: Forma geral dos modelos ARIMA

$$\text{ARIMA}(p, d, q)(P, D, Q)_s$$

Onde p representa a ordem da parte auto-regressiva do modelo utilizado, d representa o grau do primeiro nível de diferenciação utilizado e q representa a ordem da média móvel utilizada. Já as versões maiúsculas das letras representam o mesmo que estas para a parte sazonal da série temporal.

Desta forma, pode-se falar de modelos ARIMA (0,1,0) ou (2,1,0) ou mesmo de ARIMA (2,1,2)(1,0,0)_s.

Vale destacar que diversos modelos ARIMA são apenas outras formas de escrever outros modelos de séries temporais. A tabela abaixo ilustra a equivalência entre alguns modelos apresentados anteriormente neste relatório com modelos ARIMA.

Tabela 5: Equivalência entre modelos de série temporal e modelos ARIMA

ID	Modelo de série temporal	Modelo ARIMA
1	Smoothing	ARIMA (0,1,1)
2	Método de Holt	ARIMA (0,2,2)
3	aditivo	ARIMA (0,1,s+1)(0,1,0)s
4	Método de Helt-Winters multiplicativo	sem equivalências

6.3.2 Previsão Qualitativa

Previsão qualitativa é toda aquela que é resultado não-direto de um modelo matemático (o que não quer dizer que não seja baseada em dados e tendências históricas), baseada principalmente na experiência e conhecimento de um profissional ou de uma equipe de profissionais.

Ela é a mais recomendada para situações em que a “Hipótese da Continuidade” não se mostre válida ou coerente (visto que todos os métodos quantitativos se baseiam em tal hipótese), sejam tais situações previsões para o curto ou para o longo prazo.

6.3.2.1 Análise crítica das previsões qualitativas

De acordo com um estudo conduzido por Hogarth e Makridakis (1981), os usuários de previsões qualitativas se mostram, na maioria das vezes, insatisfeitos com a qualidade das previsões que utilizam e que, a acurácia das mesmas é, em geral, muito baixa.

Tal baixa qualidade das previsões qualitativas se deve a diversos motivos, dentre eles:

- Interesses pessoais do agente realizador da previsão, como, por exemplo, bater metas
- Influência do humor no resultado obtido
- Otimismo ou pessimismo inerente à pessoa

- Falta de interesse/comprometimento com o resultado
- Limitações da memória do agente realizador da previsão, que acaba por omitir ou minimizar os efeitos de algumas informações, fatos e acontecimentos relevantes para a previsão do futuro
- Inconsistência das previsões
- Não aprimoramento do método de previsão baseado em resultados e erros passados
- Falta de método para se realizar a previsão

Alguns dos itens acima apresentados são minimizados de forma razoavelmente fácil através de uma análise crítica do processo e resultado da análise feita. Já outros são bem mais difíceis de serem detectados e tratados.

De forma a minimizar diversos dos itens acima listados, será apresentado em seguida o Método Delphi, que, apesar de ser custoso, vem obtendo ótimos resultados no mundo empresarial.

6.3.2.2 Método Delphi

O método Delphi consiste, de forma resumida, em uma metodologia para se fazer previsões qualitativas.

Para se utilizar o método Delphi necessita-se de um coordenador do projeto/estudo e de uma equipe (interna ou externa à empresa) de especialistas no tema relacionado à variável para a qual se deseja realizar uma previsão qualitativa.

Inicia-se o método solicitando-se a cada um dos especialistas que gere uma previsão qualitativa (recomenda-se aqui que as mesmas sejam entregues em um formato uniforme, de forma a facilitar a próxima etapa do processo) e a envie ao coordenador do projeto/estudo.

Feito isso, cabe ao coordenador elaborar um resumo analítico das previsões elaboradas contendo dados como média das previsões, desvio-padrão das mesmas, hipóteses utilizadas em cada uma e quaisquer outras informações consideradas relevantes para a situação em questão.

Caso perceba-se através do resumo analítico que há um elevado grau de semelhança entre as previsões, conclui-se o processo de se elaborar a previsão através de uma reunião de fechamento do projeto/estudo.

Caso contrário, envia-se o resumo analítico aos especialistas e pede-se a eles que façam uma análise crítica do material e reelaborem suas previsões iniciais.

Com isso, segue-se com o ciclo do processo até se chegar a um grau de semelhança entre as previsões aceitável ou ao melhor resultado possível dentro do limite de tempo existente para a situação.

A figura abaixo ilustra de forma esquemática o funcionamento do método Delphi descrito anteriormente.

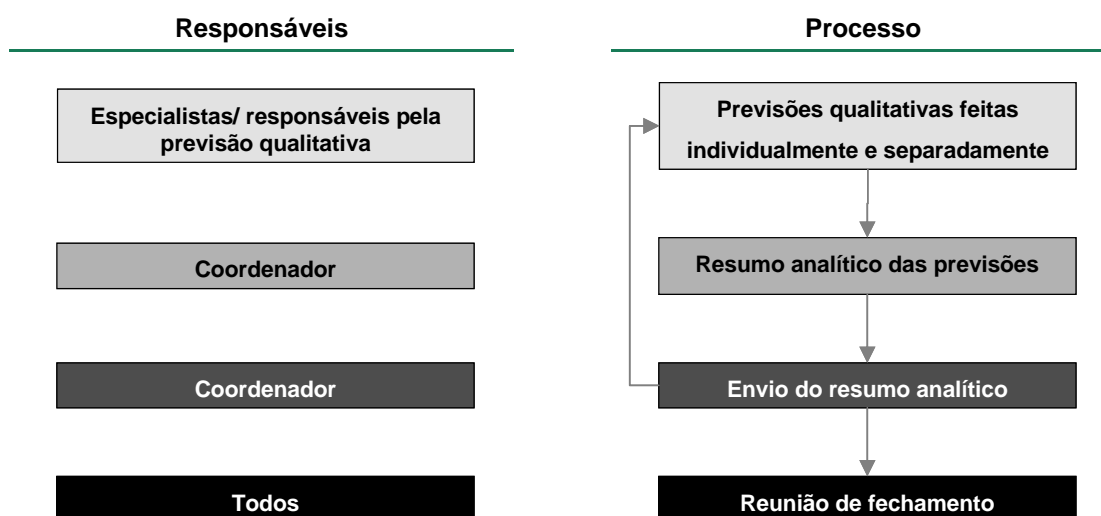


Figura 13: Funcionamento do método Delphi

6.3.2.3 Ancoragem

Ancoragem é uma tendência comportamental que se desenvolve quando alguém (especialmente a pessoa com maior cargo/poder dentro de uma reunião) apresenta uma previsão qualitativa.

Em muitos casos, mesmo quando a previsão apresentada é bastante divergente da opinião de outros presentes na reunião, a mesma se torna uma âncora e a previsão final adotada fica muito próxima da inicialmente apresentada.

Para evitar esse fenômeno comportamental, Hanke (1998) sugere que, antes de alguma reunião em que uma previsão terá papel central elabore-se uma previsão quantitativa. A mesma deverá, ainda antes da reunião, ser entregue a todos os participantes da mesma juntamente com os dados e hipóteses usados para chegar a tal previsão, além de uma explicação da metodologia utilizada na mesma.

É extremamente importante se deixar claro a todos os participantes que aquela previsão somente será válida caso não haja grandes mudanças no ambiente em que a empresa está inserida, o que é bastante improvável que ocorra. Desta forma, apenas a experiência dos participantes da reunião pode avaliar o impacto que tais mudanças podem ocasionar na variável de previsão.

Desta forma, pede-se que os mesmos listem as variáveis que eles acreditam que podem alterar o resultado da previsão e qual é a variação esperada por eles (isso deve ser feito através de formulários anônimos).

Após isso é feita uma compilação dos formulários recebidos e apresenta-se tal material na reunião, como material balizador para a discussão (pode-se considerar a “ancoragem” como uma variação do método Delphi apresentado anteriormente neste mesmo trabalho).

Um exemplo citado por Hanke (1998) de caso em que a ancoragem pode ser extremamente útil são reuniões de discussão de orçamento, as quais geralmente dependem centralmente do resultado futuro de uma variável para o qual cada participante da reunião possui uma expectativa bastante diferente.

6.4 Utilização dos diferentes métodos de previsão

Resumindo o que foi apresentado nesta Revisão Bibliográfica, nota-se que determinados tipos de métodos de previsão mostram-se mais adequados para determinados tipos de problemas em questão. As principais variáveis de escolha de um modelo para a solução de determinado problema são:

- O horizonte temporal para o qual se deseja fazer a previsão
- O(s) padrão(ões) encontrado(s) nos dados da série histórica da variável de previsão

A tabela abaixo apresenta a relação dos modelos apresentados nesta seção do relatório e as variáveis acima mencionadas, além de classificá-los em “Explanatórios”, “Série Temporal” ou “Qualitativos”.

Tabela 6: Resumo de modelos e suas aplicações

ID	Método	Padrão dos Dados	Horizonte Temporal	Tipo de Modelo
1	Regressão Simples	-	Intermediário e longo prazo	Explanatório
2	Regressão Múltipla	-	Intermediário e longo prazo	Explanatório
3	Naïve	Estacionário Com Tendência Sazonal	Curto prazo	Série Temporal
4	Média Simples	Estacionário	Curto prazo	Série Temporal
5	Média Móvel	Estacionário	Curto prazo	Série Temporal
6	Extrapolação	Estacionário Com Tendência Sazonal	Curto prazo	Série Temporal
7	Suavização Exponencial	Estacionário Com Tendência Sazonal	Curto prazo	Série Temporal
8	ARIMA	Estacionário Com Tendência Sazonal	Curto prazo	Série Temporal
9	Delphi	-	Intermediário e longo prazo	Qualitativo
10	Ancoragem	-	Intermediário e longo prazo	Qualitativo

Fonte: Hanke (1998)

7. Modelo de previsão de demanda de gás natural

Nesta seção serão desenvolvidos modelos buscando prever a demanda de gás natural brasileira para os próximos anos.

Inicialmente, tem-se que o consumo brasileiro de gás natural por fonte apresentou o seguinte perfil desde 1970:

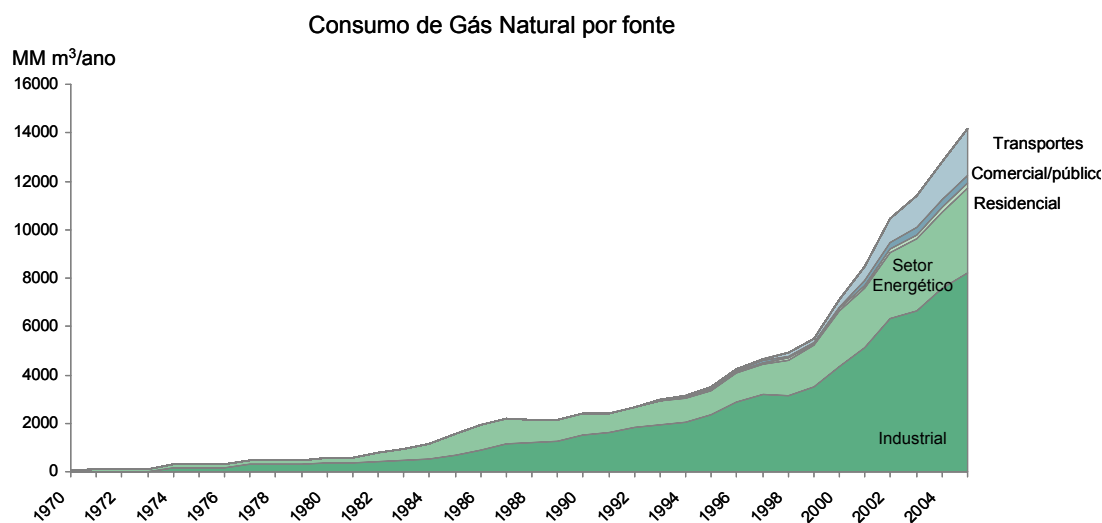


Gráfico 14: Consumo de Gás Natural por fonte

Fonte: Balanço Energético Nacional 2006. Elaborado pelo autor

Apresentando as séries históricas de volume de forma separada (não empilhadas) tem-se o seguinte gráfico, o qual talvez se mostre mais útil para a compreensão do comportamento de consumo de gás por cada fonte:

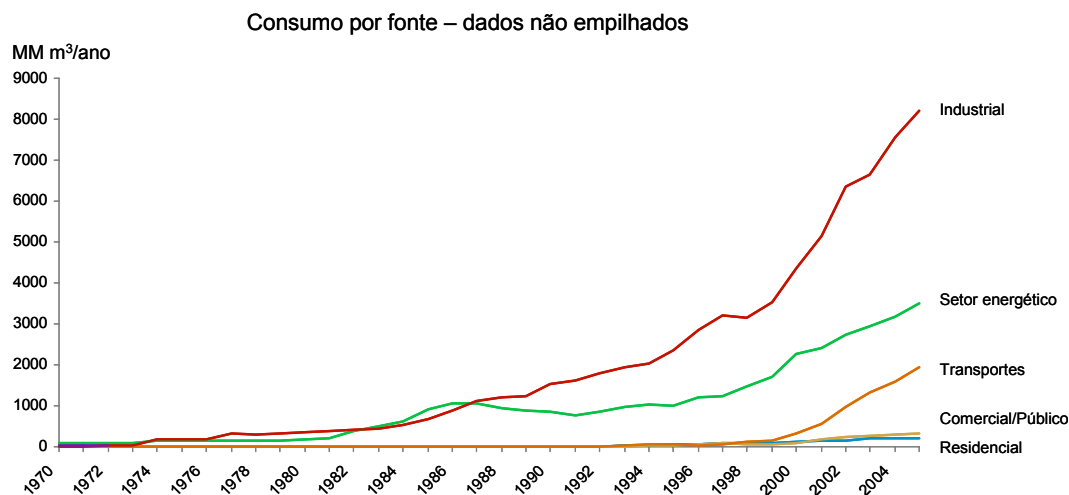


Gráfico 15: Consumo de Gás Natural por fonte – dados não empilhados

Fonte: Balanço Energético Nacional 2006. Elaborado pelo autor

Analisando-se de forma inicial o gráfico acima, chega-se às seguintes conclusões a respeito das curvas de cada segmento da demanda de gás natural:

Tabela 7: Análise das séries históricas de consumo por fonte de gás natural

Segmento de demanda	Número de dados utilizáveis	Perfil (visual) da curva	OBS
Setor Energético	36	Linear	
Transportes	17	Linear (a partir de 1999)	Consumo se torna expressivo a partir de 1999
Industrial	36	Exponencial ou linear	
Residencial	19	Linear ou exponencial	Perfil e amplitude das curvas, além do pequeno número de dados, dificulta identificação do perfil
Comercial/Público	19	Linear ou exponencial	Perfil e amplitude das curvas, além do pequeno número de dados, dificulta identificação do perfil

Vale destacar que entre 1970 e 2005 tem-se 36 pontos de dados, porém, se desconsiderou neste estudo os anos em que o consumo de gás em cada segmento foi zero (considerou-se que este consumo ainda não existia), de forma a não enviesar os modelos a

serem obtidos, Desta forma, nos casos dos segmentos de Transportes, Residencial e Comercial/Público o número de dados com os quais se trabalhará é inferior a 36.

A seguir serão desenvolvidos modelos de previsão quantitativa, tanto explanatórios como baseados em séries temporais, para cada uma das parcelas da demanda acima discriminadas.

7.1 Setor de Transportes

A seguir serão desenvolvidos diferentes modelos de previsão para a demanda do setor de transportes, tanto explanatórios como baseados em séries temporais.

7.1.1 Transportes – Método de Holt

Analizando inicialmente o gráfico com a série temporal da demanda do setor de transportes, observamos que a mesma possui uma tendência de crescimento e pouca ou nenhuma sazonalidade.

Desta forma, de acordo com a Revisão Bibliográfica deste trabalho, pode-se concluir que um modelo que se mostra adequado para a previsão destas séries é o modelo de Holt.

Utilizando-o para o consumo do setor de transportes obtém-se o seguinte gráfico:

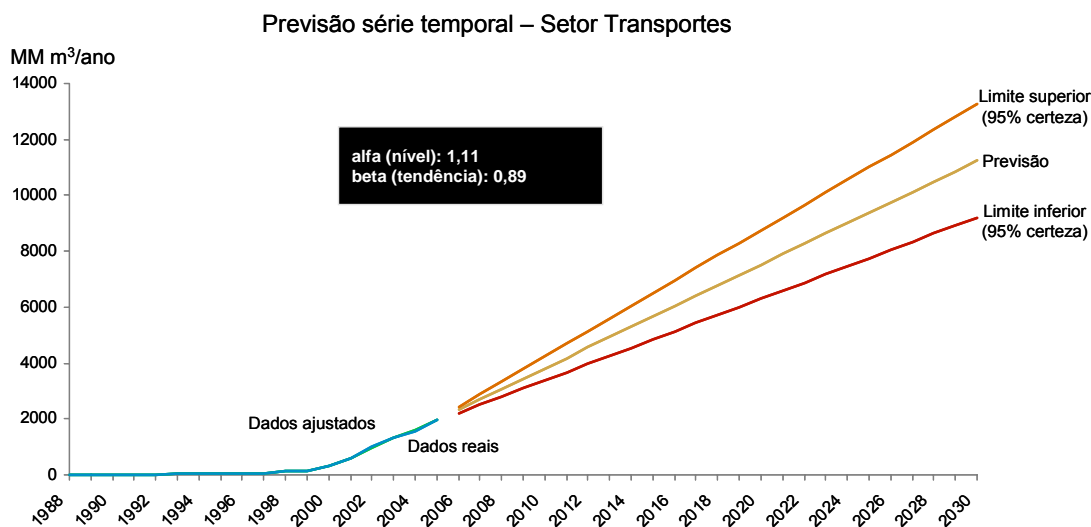


Gráfico 16: Modelo de Holt para demanda do Setor de Transportes

Nele pode-se observar uma previsão com tendência futura linear, de acordo com o funcionamento do modelo utilizado. Porém, vale ressaltar que, por considerar que o comportamento futuro de uma série temporal refletirá o seu passado, o método de Holt costuma obter resultados melhores no curto prazo do que no longo prazo.

7.1.2 Transportes – Regressão Explanatória

Buscando realizar uma previsão por regressão para o setor de Transportes, pode-se pensar nas seguintes variáveis que teriam algum relacionamento lógico com a mesma:

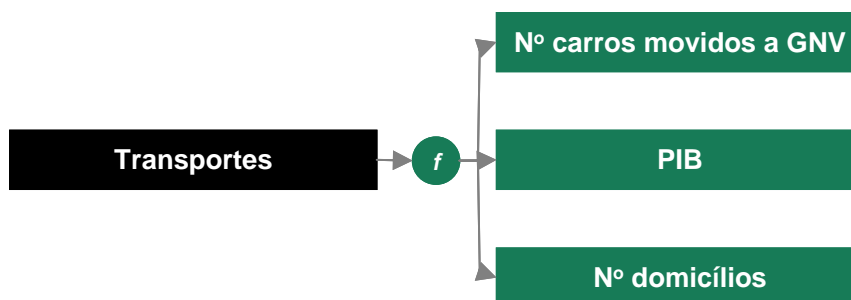


Figura 14: Possíveis variáveis explanatórias para demanda de transportes

Ou seja, considera-se que, de alguma forma, a demanda de gás do setor de transportes possa estar relacionada com o número de carros convertidos a GNV no Brasil, o PIB nacional ou o número de domicílios nacional. A primeira variável descrita deve, intuitivamente, ser fortemente relacionada com a demanda de gás do setor de transporte. Já as outras duas variáveis são “coringas”, relacionadas à aspectos macroeconômicos brasileiros, podendo ser utilizadas em diversos tipos de análises de regressão.

Rodando no Minitab® uma regressão *Stepwise forward-with-a-backward look* com essas três variáveis e os parâmetros de “Alpha-to-enter” e “Alpha-to-remove” de 15%, obtém-se a seguinte equação:

Equação 29: Demanda de transportes em função de modelo de regressão explanatória

$$y = 61,79 + 0,00185 \times \text{NúmeroDeCarrosGNV}$$

Tal equação possui um R^2 de 99,5%, resultado bastante elevado que pode ser justificado pelo baixo número de pontos que estão sendo incluídos na regressão (10). Além disso, todos os seus *t-stat* são maiores do que 2.

Analisando-se os resíduos que essa equação gera nos dados históricos da demanda de Transportes tem-se a seguinte distribuição:

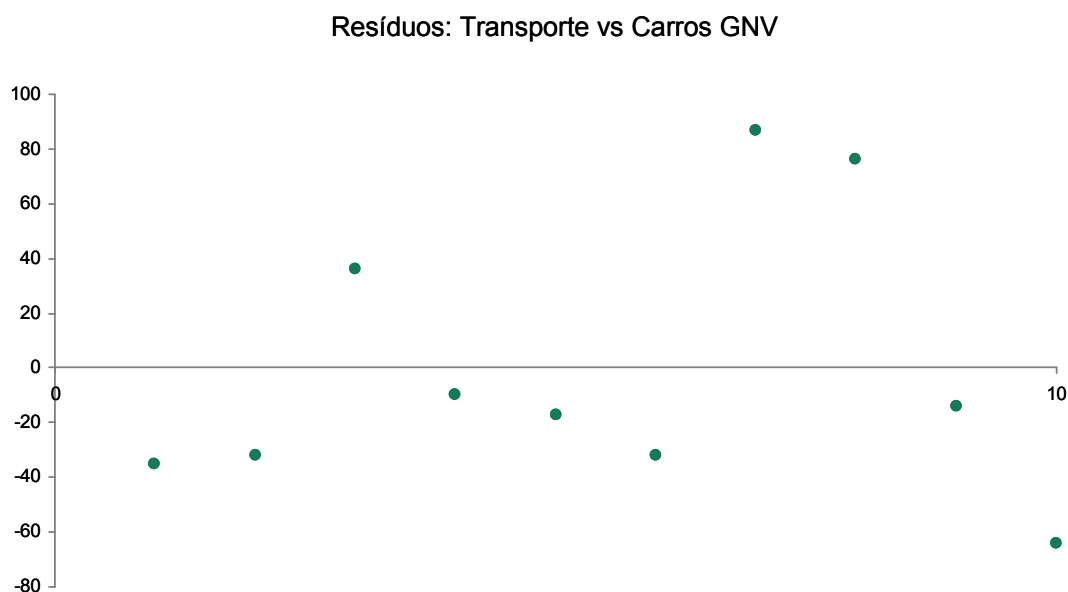


Gráfico 17: Resíduos Transporte vs Carros GNV I

O gráfico acima serve para verificar se a regressão realizada não viola nenhuma das hipóteses da regressão. Ou seja, busca-se que os resíduos tenham distribuição aleatória com média zero e variância constante ao longo do eixo x.

Aparentemente neste caso não se viola de forma abrupta as hipóteses acima descritas, apesar de se verificar um número muito alto de pontos abaixo do zero. Porém, como neste caso está se trabalhando com um número pequeno de dados (10) a análise acima fica prejudicada.

Outra análise que pode ser realizada é plotar o gráfico do ACF dos erros, para verificar se há alguma influência temporal relevante ao problema que não tenha sido incorporada ao modelo. O gráfico abaixo ilustra isso:

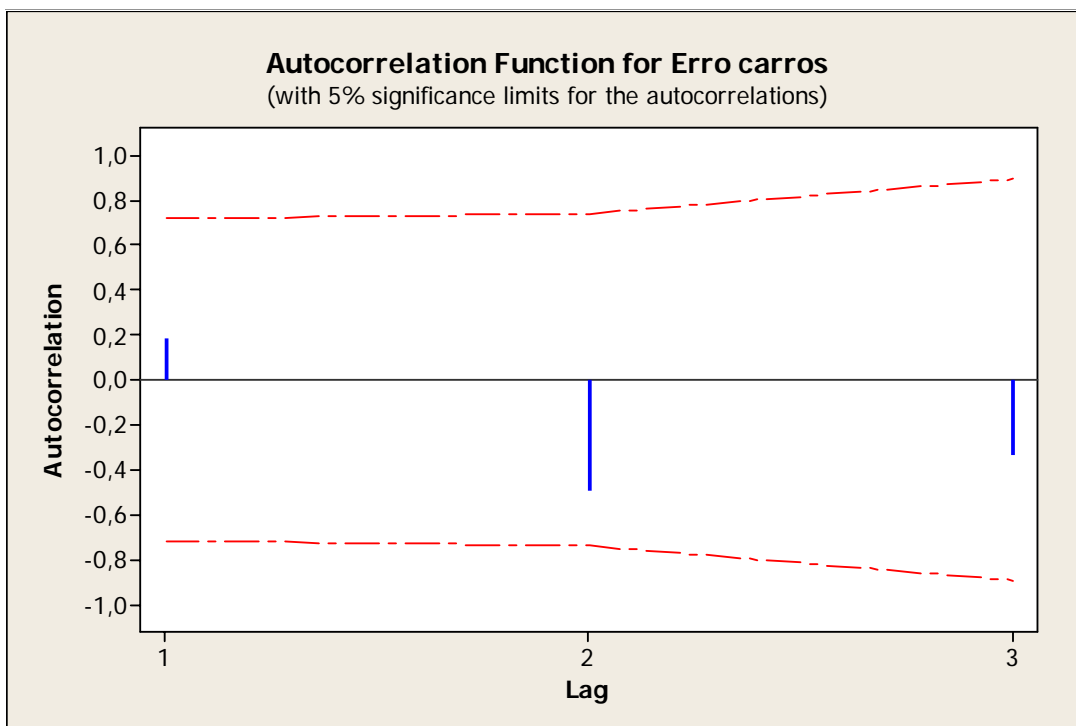


Gráfico 18: ACF dos Resíduos Transporte vs Carros GNV I

Como todos os ACFs ficaram dentro dos intervalos delimitados pelas linhas tracejadas pode-se concluir que não há nenhuma influência temporal relevante nos resíduos do modelo encontrado.

Baseado nas previsões para número de carros movidos a GNV no futuro obtém-se a seguinte previsão para a demanda de Transportes:

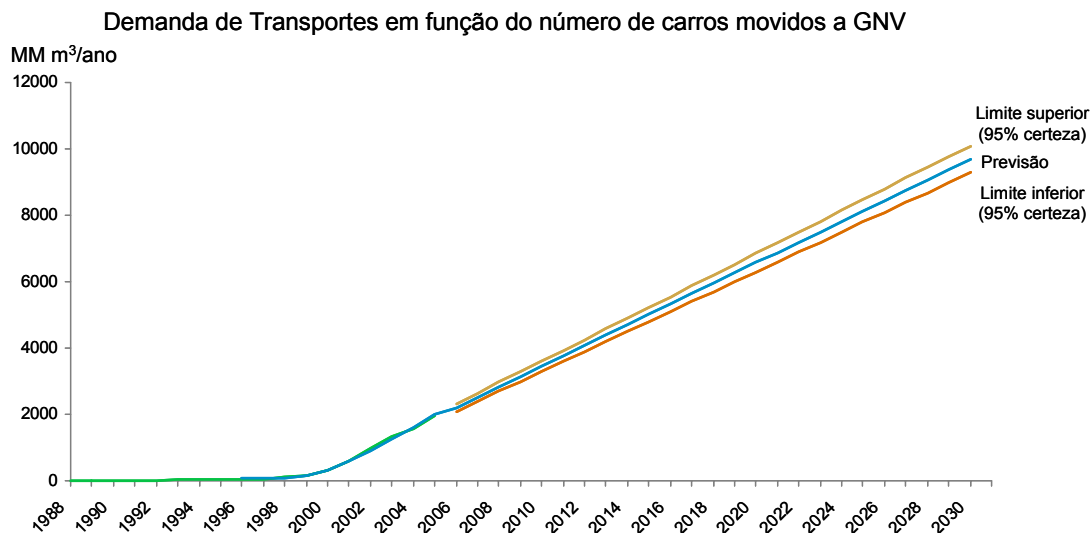


Gráfico 19: Demanda de transportes em função de número de carros movidos a GNV

Como se pode observar, através deste modelo projeta-se um crescimento linear para a demanda de Transportes no futuro. Além disso, nota-se que as curvas dos limites dos intervalos de confiança formam um ângulo bastante fechado entre si, o que denota uma aparente grande confiabilidade da previsão sendo realizada (considerando-se que as previsões de número de carros movidos a GNV não possuem erros, o que pode não ser verdadeiro já que as mesmas foram retiradas de pesquisas na mídia através da internet e podem ter sido elaboradas com pouco método e rigor).

Esta previsão considera que a relação descrita pela equação do modelo se manterá no futuro, ou seja, que a demanda de Transportes continuará seguindo a mesma proporcionalidade com o número de carros movidos a GNV, o que pode não ser verdade no futuro devido a eventuais mudanças no perfil da frota de carros brasileiros, aperfeiçoamento da tecnologia dos motores, ou outros fatores.

7.1.3 Transportes – Auto-regressão

Realizando-se a análise de ACFs diretamente para a série temporal de demanda do setor de transportes encontra-se o seguinte gráfico:

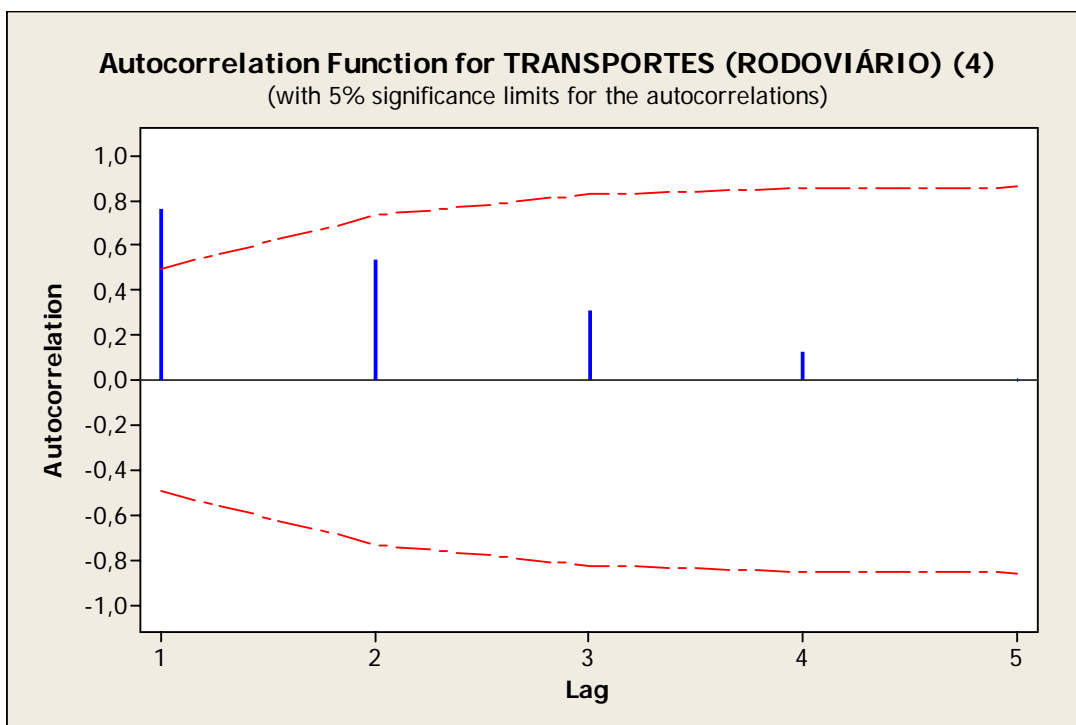


Gráfico 20: ACF da demanda do setor de Transportes

Tal gráfico indica que uma variável temporal deveria ser incluída na previsão de demanda do setor de transportes, o que nos leva a buscar um terceiro modelo.

Realizando-se uma auto-regressão da demanda do setor de Transportes, encontra-se a seguinte equação:

Equação 30: Demanda de transportes em função de modelo de auto-regressão

$$y = 26,39 + 1,846 \times \text{Lag1} - 0,80 \times \text{Lag2}$$

Tal regressão apresenta um R^2 -ajustado de 98,62, valor bastante elevado que pode ser justificado pelo baixo número de dados incluídos no modelo (apenas 16, que é o que a base histórica do BEN nos fornece). O *t-stat* encontrado dos coeficientes é maior do que 2, mas o *t-stat* da constante foi de 1,13. Por ser uma medida do nível inicial da série e influenciar pouco nos resultados finais, releva-se este problema.

Analisando-se os resíduos deste modelo encontra-se a seguinte distribuição:

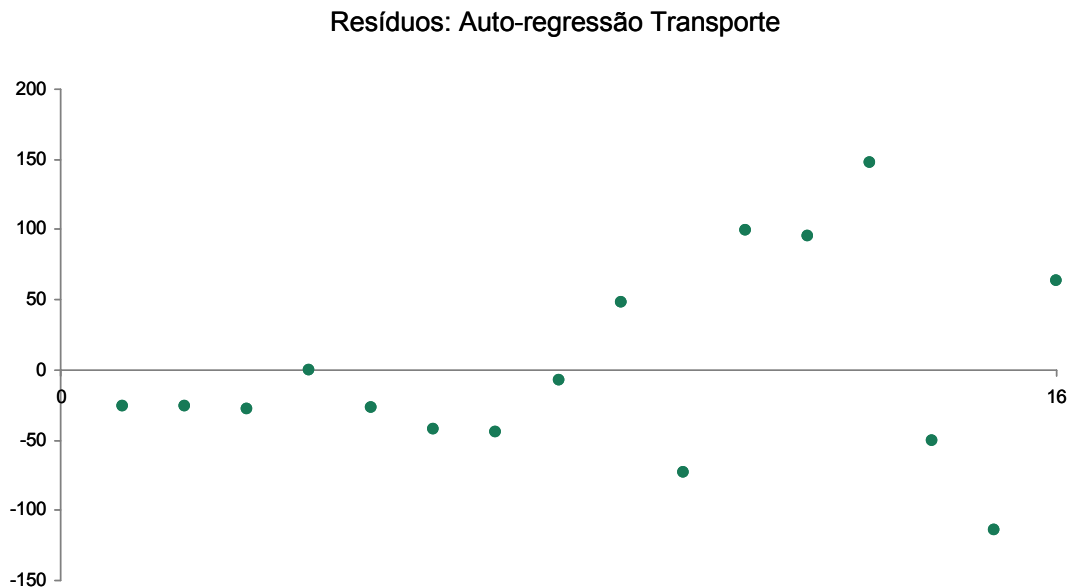


Gráfico 21: Resíduos auto-regressão Transporte

Como se pode observar neste caso, a série dos resíduos viola visivelmente a hipótese de homocedasticidade, não possuindo variância constante ao longo do eixo x (neste caso a variância aumenta conforme x aumenta), além de possuírem uma longa sequência abaixo do zero no início da série.

Realizando-se uma regressão *Stepwise-with-a-backward-look* com as variáveis explanatórias e os parâmetros anteriormente apresentados, observa-se que nenhuma nova variável é incluída no modelo.

A previsão obtida pelo mesmo é a seguinte:

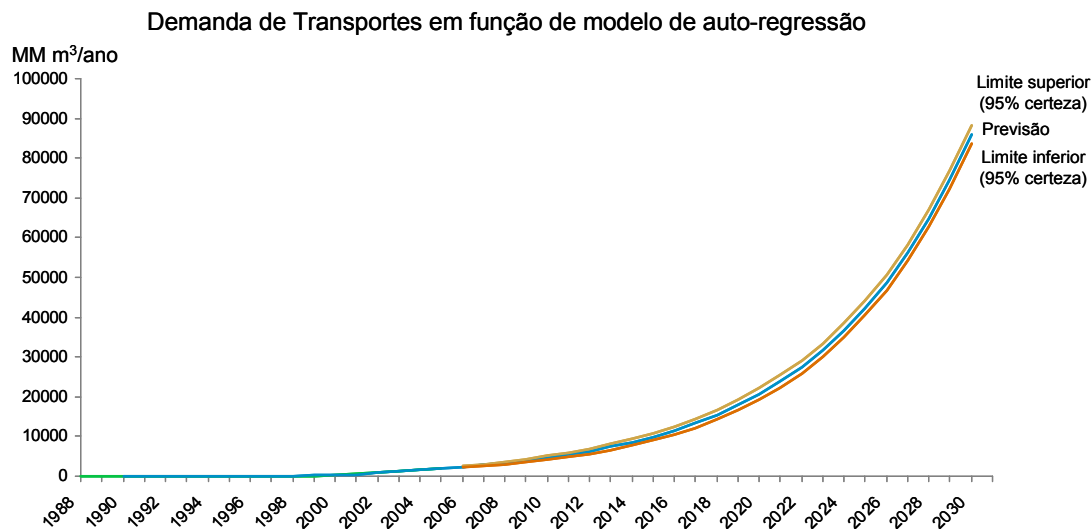


Gráfico 22: Previsão da demanda de transportes em função de auto-regressão

Neste caso prevê-se um crescimento exponencial para a demanda de gás natural do setor de Transportes.

Porém, tal modelo apesar de se mostrar apropriado para o curto prazo não se mostra necessariamente muito confiável para o longo prazo, já que a relação temporal que rege a série pode ser alterada no futuro.

Vale destacar que, devido à dificuldade matemática envolvida no cálculo de intervalos de predição para auto-regressões (dado que os mesmos podem ser facilmente auto-correlacionados), os intervalos de confiança da figura acima foram estimados. Calculou-se o intervalo de confiança da previsão um período à frente e depois se estimou os intervalos de confiança das outras previsões multiplicando-se a primeira previsão pela mesma taxa de crescimento verificada na previsão por Holt (a proporção com o método de Holt foi adotada pelo fato de ambos os modelos serem baseados em séries temporais).

7.1.4 Transportes – Comparação dos modelos

Comparando-se os três métodos de três previsões realizadas para a demanda de Transportes, obtém-se o seguinte resultado:

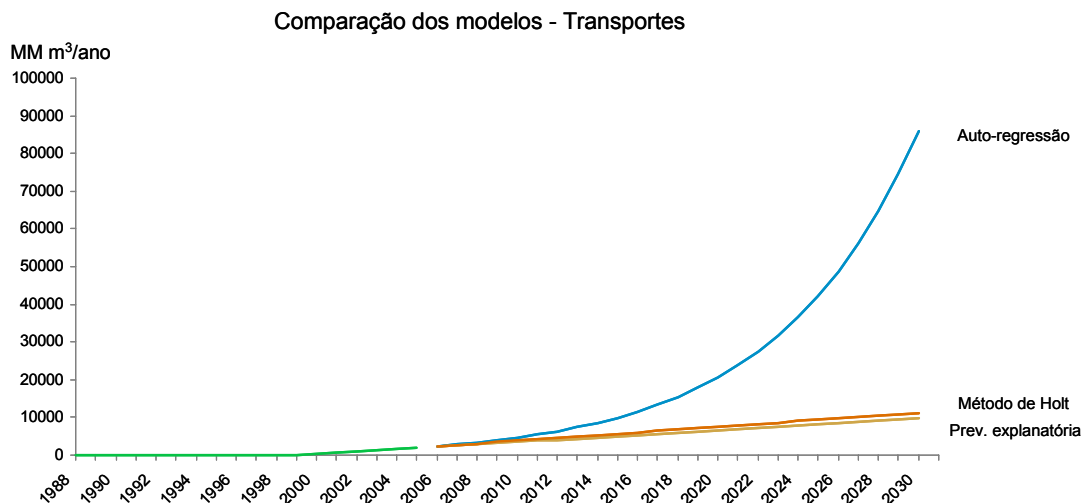


Gráfico 23: Comparação dos modelos - Transportes

Analisando-se o gráfico acima, observa-se inicialmente que os resultados obtidos pelos métodos de Holt e de regressão explanatória ficaram muito próximos um do outro. Além disso, pode-se observar que os resultados obtidos pelo modelo de auto-regressão se “descolam” dos demais, principalmente quando o horizonte de previsão é maior. Até por volta de 2011 os resultados das 3 previsões são próximos uns dos outros, o que nos leva a concluir que até este período foi encontrada uma faixa de resultados bastante confiáveis.

Montando-se um quadro comparativo dos modelos de previsão utilizados chega-se à seguinte tabela que compara diversos parâmetros, numéricos e subjetivos, diretos e indiretos, os quais fornecem alguma informação a respeito da qualidade das previsões obtidas.

Tabela 8: Quadro comparativo de modelos – Setor de Transportes

Setor Transportes	Regressão Explanatória	Auto-regressão	Holt
Número de dados históricos incluídos no modelo	10	16	18
Relação Intervalo de Confiança/Previsão em 2030	4%	3%	18%
R2 ou R2-ajustado encontrado	99,5%	98,6%	n/a
MSE	2.239,5	4.660,4	4.675,5
MAE	40,0	56,0	47,6
MAPE	23%	203%	31%
Problemas com resíduos de 1a ordem	*Número muito baixo de pontos abaixo de zero	*Não-homocedasticidade *Longa sequência abaixo do zero no início da série	n/a
Problemas com resíduos de 2a ordem	n/a	n/a	n/a
Método conceitualmente mais adequado para previsões de...	Curto/Médio prazo	Curto prazo	Curto prazo
Alinhamento de resultados com outros métodos?	Holt	Nenhum	Regr. Explanatória

Segue uma pequena explicação sobre cada um dos parâmetros acima elencados:

- **Número de dados históricos incluídos no modelo:** muito importante, dado que um modelo baseado num número menor de pontos possui menor poder de extrapolação do que um baseado num número maior de pontos.
- **Relação Intervalo de Confiança/Previsão em 2030:** este parâmetro é a razão entre o Limite Superior previsto em 2030 e a previsão base em 2030 e dá uma dimensão de quão precisa a previsão gerada pelo modelo se propõe a ser. Vale ressaltar que, como os dados estão representados percentualmente, previsões em patamares bastante elevados podem possuir resultados percentuais menores do que previsões em patamares mais baixos, mesmo possuindo uma amplitude absoluta maior.
- **R² ou R²-ajustado:** mostra o indicador de explicação (devidamente ajustado, se for o caso), encontrado para as regressões realizadas
- **MSE, MAE e MAPE:** Indicadores do erro entre o modelo e os dados históricos. As informações dos mesmos são complementares entre si
- **Problemas com resíduos:** análise dos maiores problemas encontrados durante as análises de resíduos das regressões

- **Método conceitualmente mais adequado para previsões de...:** indica qual o horizonte de previsão teórico mais adequado para o modelo
- **Alinhamento de resultados com outros métodos:** parâmetro subjetivo que ajuda a mostrar o quanto se pode ou não confiar em determinado modelo, através da validação de seus resultados por outros modelos

No caso da demanda de transportes, avalia-se que os modelos são aparentemente equilibrados entre si, com todos eles possuindo forças e fraquezas específicas. Desta forma, considera-se que o alinhamento dos resultados obtidos através dos métodos de Holt e de regressão explanatória possa torná-los mais robustos frente ao modelo de auto-regressão. Assim, recomenda-se a utilização dos resultados de um deles ou da média deles como previsão para a demanda do setor de transportes.

De toda forma, os modelos acima elaborados são mais apropriados para o curto ou médio prazo (entre 1 e 5 anos, aproximadamente), tornando seus resultados a longo prazo mais uma base para a realização de predições por especialistas do que um resultado a ser sem questionamentos.

7.2 Setor Energético

A seguir serão desenvolvidos diferentes modelos de previsão para a demanda do setor energético, tanto explanatórios como baseados em séries temporais.

7.2.1 Energético – Método de Holt

Analisando o gráfico com a série temporal da demanda do setor energético, observamos que o mesmo possui uma tendência de crescimento e pouca ou nenhuma sazonalidade.

Desta forma, de acordo com a Revisão Bibliográfica deste trabalho, pode-se concluir que um modelo que se mostra adequado para a previsão destas séries é o modelo de Holt.

Utilizando-o para o consumo do setor energético obtém-se o seguinte gráfico:

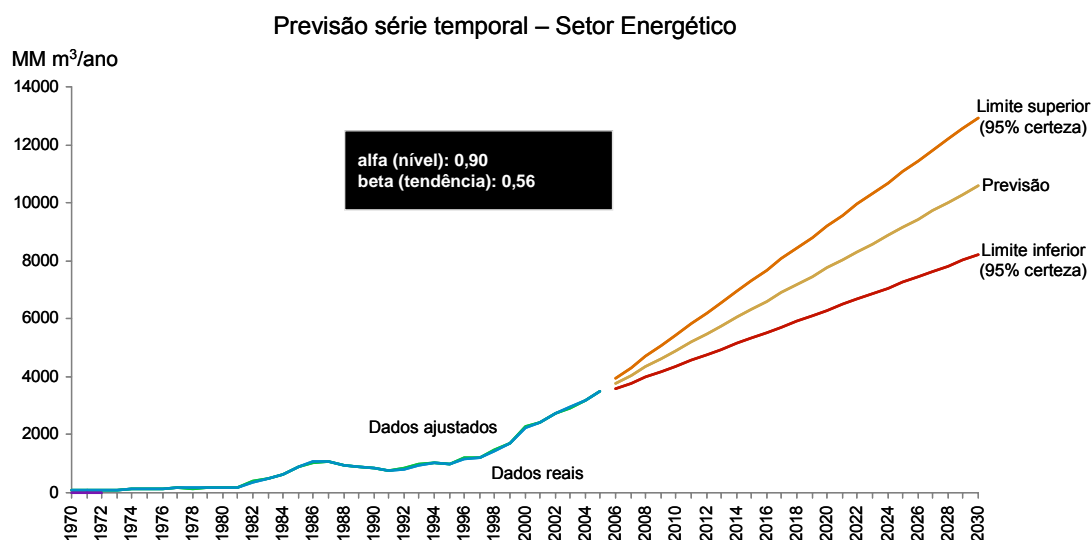


Gráfico 24: Modelo de Holt para demanda do Setor Energético

Nele pode-se observar uma previsão com tendência futura linear, de acordo com o funcionamento do modelo utilizado. Porém, vale ressaltar que, por considerar que o comportamento futuro de uma série temporal refletirá o seu passado, o método de Holt costuma obter resultados melhores no curto prazo do que no longo prazo.

7.2.2 Energético – Regressão Explanatória

Buscando-se realizar uma previsão por regressão para o setor Energético, pode-se pensar nas seguintes variáveis que teriam algum relacionamento lógico com a mesma:

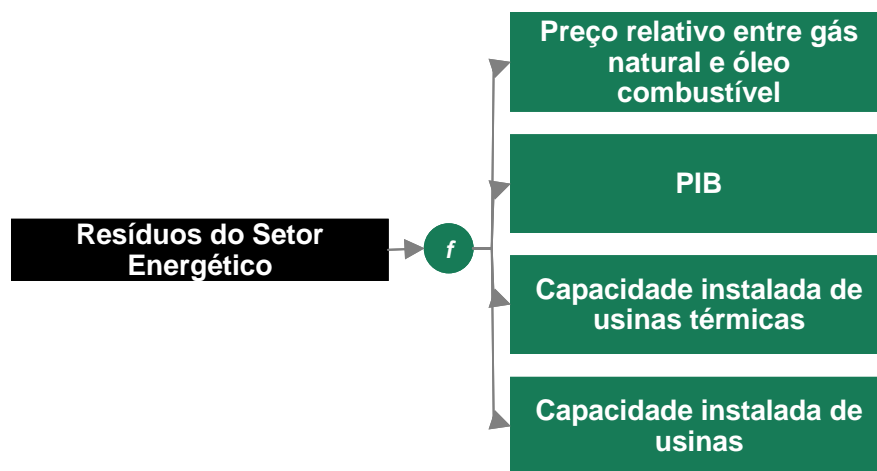


Figura 15: Possíveis variáveis explanatórias para demanda energética

Ou seja, considera-se que, de alguma forma, a demanda de gás do energético possa estar relacionada com:

- **Preço relativo entre gás natural e óleo combustível:** já que as usinas geradoras de energia podem migrar para o combustível que apresente uma melhor relação custo-combustível, impactando assim diretamente na demanda de gás deste setor
- **PIB:** já que quanto maior o PIB nacional, maior a demanda de energia elétrica, o que torna possivelmente maior a demanda de gás natural para gerar energia elétrica
- **Capacidade instalada de usinas térmicas:** já que quanto mais potência instalada de usinas térmicas, possivelmente maior é a demanda de gás natural para geração de energia
- **Capacidade instalada de usinas:** idem ao item anterior, só que sem discriminar o tipo de usina (que pode ser hidráulica ou térmica).

Rodando-se uma regressão *Stepwise-with-a-backward-look* com as possíveis variáveis explanatórias e a série histórica da demanda de gás natural do setor energético, chega-se à seguinte equação (a qual possui um R^2 -ajustado de 98% e todos os *t-stat* maiores do que 2):

Equação 31: Demanda do setor energético em função de modelo de regressão explanatória

$$y = 2870,6 + 14,3 \times \text{CapInstTérmica} + 16,1 \times \text{PIB} + 609 \times \text{Preço Relativo}$$

Note-se que a mesma diz que quanto maior o preço relativo do gás natural em relação ao óleo combustível, maior o consumo de gás das usinas térmicas. Uma possível inconsistência do modelo é que a lógica que se observa no mundo real é contrária a esta (nas usinas que possuem a opção de escolher seu combustível). Porém, dado que o resultado estatístico encontrado foi este, procede-se com a análise dos resultados.

Plotando-se os resíduos que essa equação gera obtém-se a seguinte distribuição:

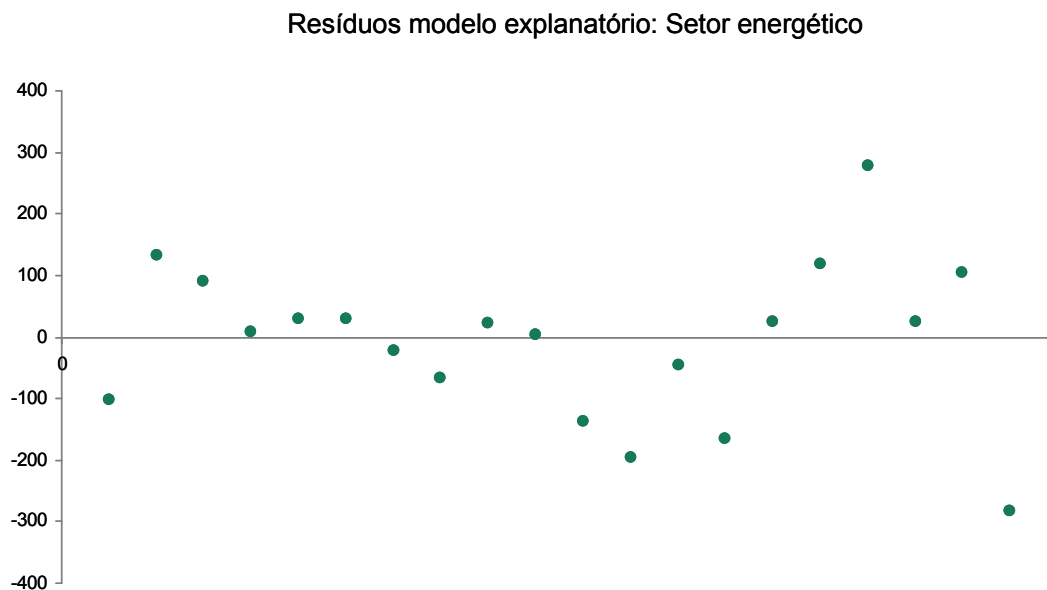


Gráfico 25: Resíduos do Setor Energético – modelo explanatório

Analisando-se os resíduos verifica-se que os mesmos aparentam ter algum grau de autocorrelação (devido a algumas seqüências de “subida” e “descida” da série), o qual não é

muito forte. Além disso, os mesmos desrespeitam a hipótese de homocedasticidade nos extremos do eixo x.

Outra análise que pode ser realizada é plotar o gráfico do ACF dos erros, para verificar se há alguma influência temporal relevante ao problema que não tenha sido incorporada ao modelo. O gráfico abaixo ilustra isso:

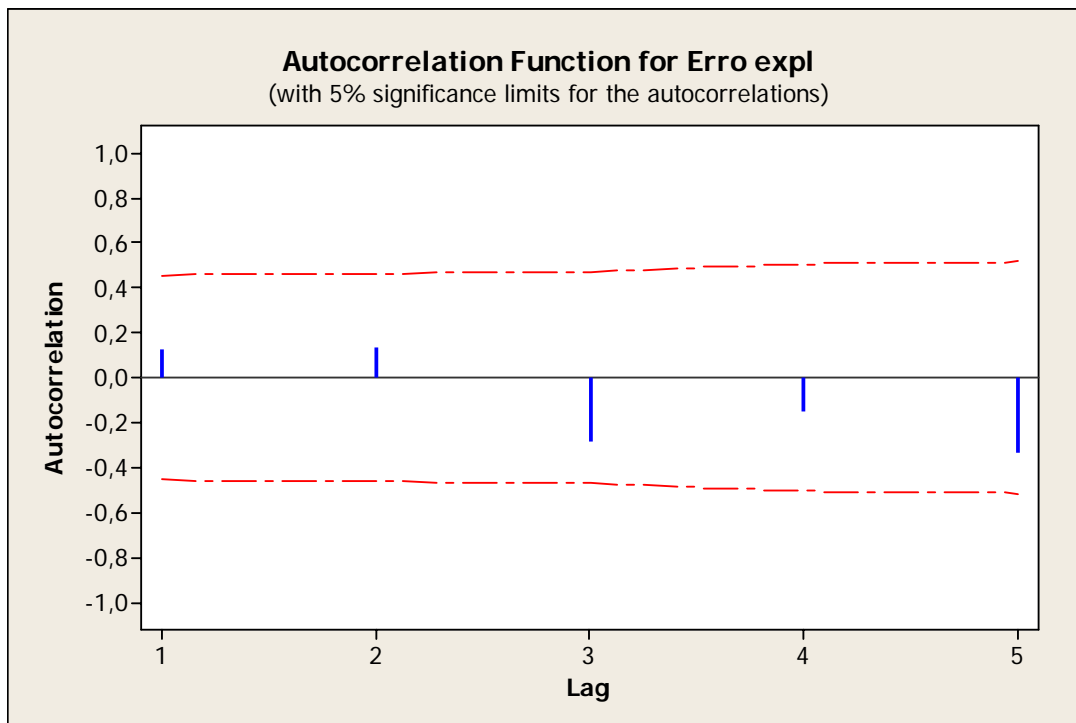


Gráfico 26: ACF dos resíduos - modelo explanatório setor energético

Como todos os ACFs ficaram dentro dos intervalos delimitados pelas linhas tracejadas pode-se concluir que não há nenhuma influência temporal relevante nos resíduos do modelo encontrado.

Para se proceder com a utilização do modelo para a realização da previsão, é necessário se possuir previsões de cada uma das variáveis de entrada do mesmo. Previsão do PIB e capacidade instalada térmica são baseadas em fontes externas de informação (vide Apêndice para maiores detalhes). Já a previsão do preço relativo do gás natural em relação ao óleo combustível deverá ser determinada neste trabalho.

Analisando a série histórica do preço relativo do gás natural em relação ao óleo combustível tem-se:

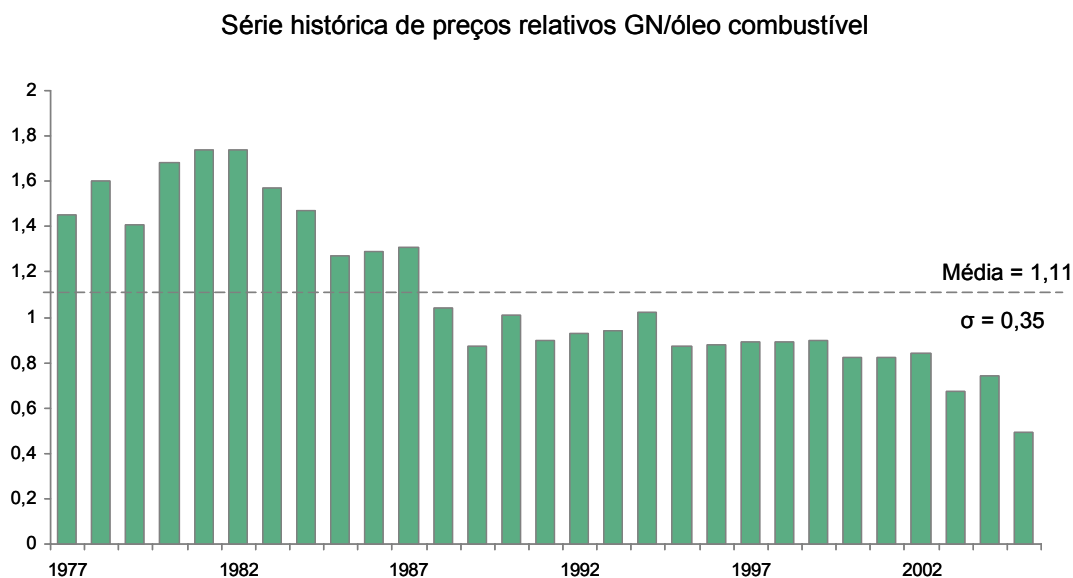


Gráfico 27: Série histórica de preços relativos gás natural/óleo combustível

Fonte: Balanço Energético Nacional 2006. Elaborado pelo autor

Pode-se observar no gráfico acima que, historicamente o preço relativo do gás natural em relação ao óleo combustível é de 1,11, com desvio-padrão de 0,35. Nota-se também que este preço relativo tem apresentado uma queda nos últimos anos, ou seja, o gás natural tem se tornado mais barato em relação ao óleo combustível.

Porém, de acordo com dados da Petrobras, com o advento da importação de GNL em vias de começar no Brasil, espera-se que tal preço relativo tenda a aumentar, visto que o GNL importado é mais caro que o gás natural de importação nacional e que o gás natural importado da Bolívia.

Desta forma, será adotado neste trabalho um valor médio de 1,11, constante para o futuro, de preço relativo do gás natural em relação ao óleo combustível, um valor para o cenário otimista de $1,11 + \sigma$ e um valor para o cenário pessimista de $1,11 - \sigma$, também constantes.

Desta forma, de posse de previsões de todas as variáveis de entrada do modelo, pode-se realizar a previsão da demanda de gás natural do setor energético com base no mesmo.

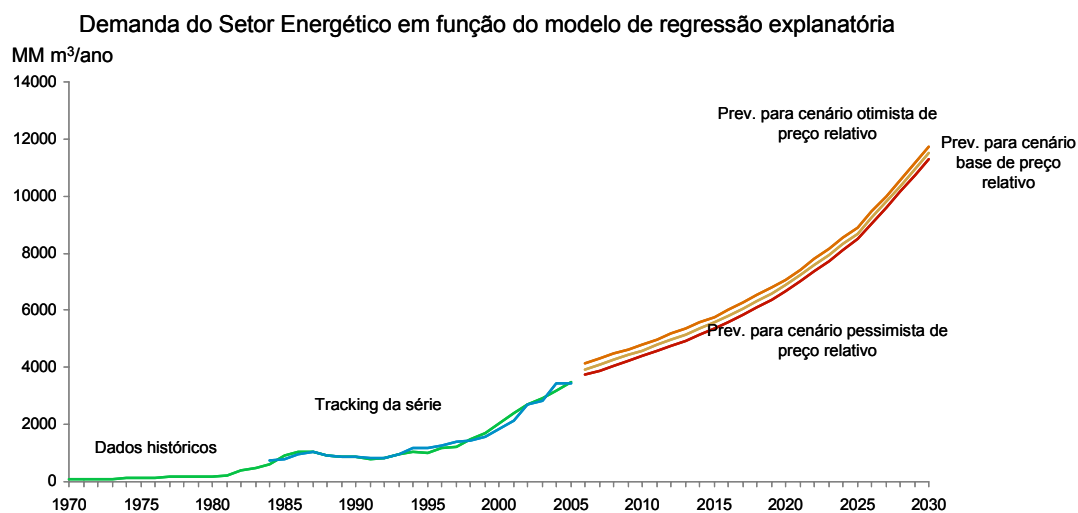


Gráfico 28: Demanda do Setor Energético em função do modelo de regressão explanatória

Como se pode observar, através deste modelo projeta-se um crescimento exponencial para a demanda do Setor Energético no futuro.

Além disso, percebe-se no gráfico acima que os diferentes cenários de preços relativos do gás natural influenciam pouco no resultado das previsões obtidas (o que minimiza o impacto das grandes simplificações realizadas na previsão futura do comportamento do mesmo).

Tendo em vista isso, pode-se elaborar o seguinte gráfico para o cenário base de preço relativo, contendo os intervalos de confiança para as previsões futuras de demanda do setor.

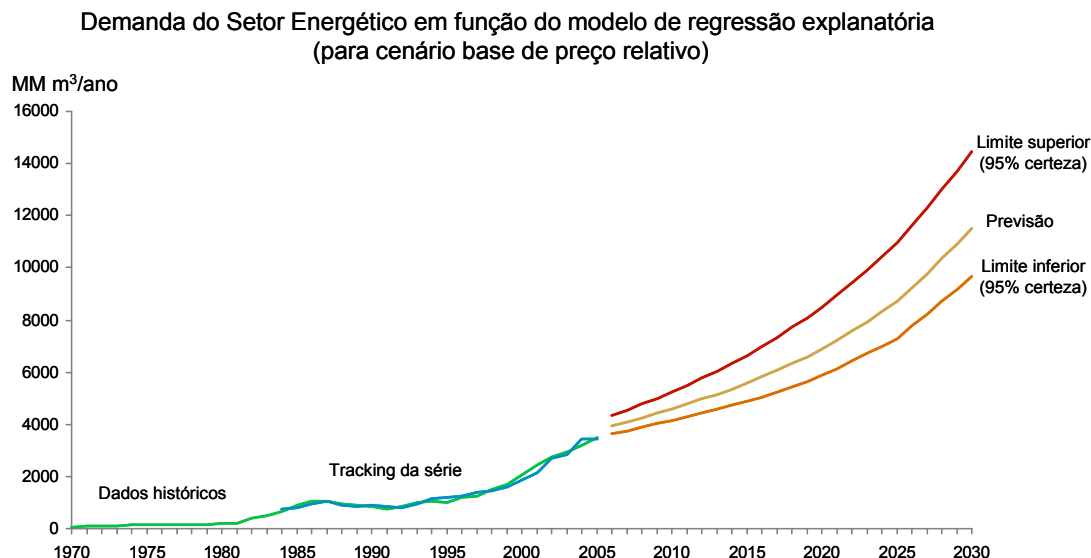


Gráfico 29: Demanda do Setor Energético em função do modelo de regressão explanatória com intervalos de confiança

Nota-se no gráfico acima o perfil exponencial da curva de previsão. Além disso, pode-se observar que curvas dos limites dos intervalos de confiança possuem uma “abertura” maior do que aquela verificada no caso do Setor de Transportes, denotando menor confiabilidade da previsão.

É importante ressaltar que esta previsão considera que a relação descrita pela equação do modelo será mantida no futuro, ou seja, que a demanda do setor Energético continuará seguindo a mesma proporcionalidade com PIB, capacidade instalada de usinas térmicas e preço relativo do gás natural em relação ao óleo combustível, o que pode não ser verdade no futuro.

Além disso, ela baseia suas previsões em previsões das variáveis de entrada, as quais, por si só, possuem erros (com maior destaque para a previsão do preço relativo entre gás natural e óleo combustível, cujo método de obtenção utilizado foi bastante simplificado).

7.2.3 Energético – Auto-regressão

Analisando-se agora a demanda do setor Energético de um ponto de vista temporal, percebe-se inicialmente que a mesma é imensamente influenciada pelo seu comportamento recente, conforme demonstrado pelo gráfico de ACF:

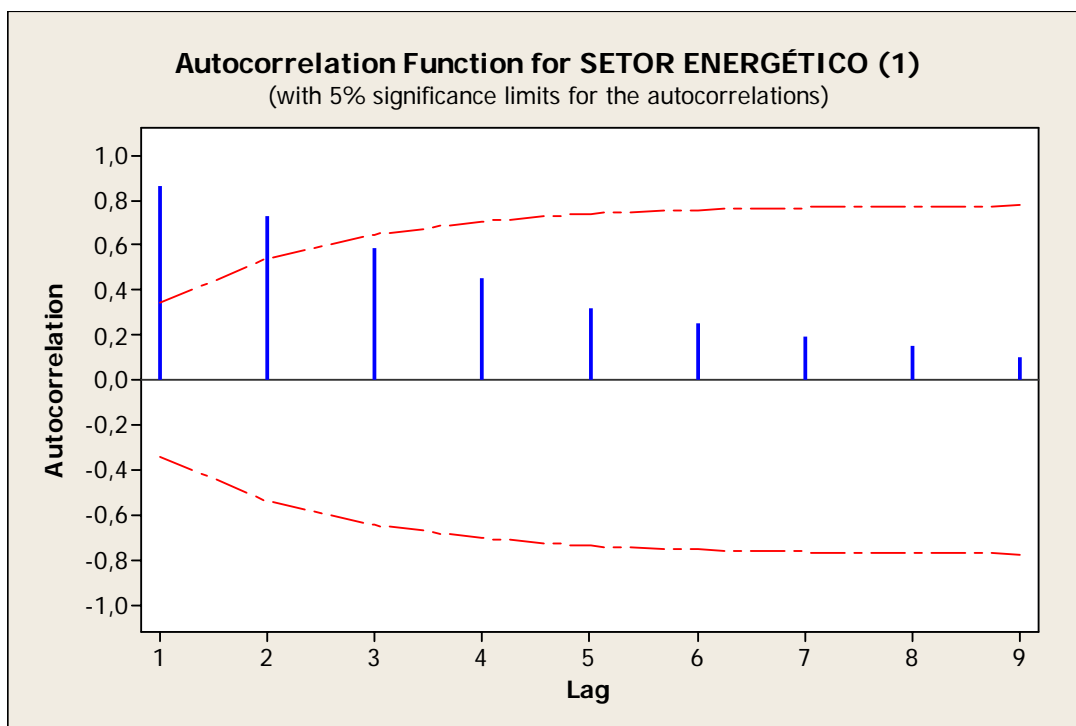


Gráfico 30: ACF da demanda do Setor Energético

Pelo gráfico acima, pode-se notar que a autocorrelação nos dois primeiros *lags* da série extrapolou os limites de controle. Desta forma, deve-se considerar a influência temporal na modelagem da série temporal.

Desta forma, buscará se explicar a série temporal de demanda de gás do setor energético através de uma equação de auto-regressão. Rodando-se um modelo de regressão Stepwise-with-a-backward-look com parâmetros de “*Alpha-to-enter*” e “*Alpha-to-remove*” de 0,15, sendo as possíveis variáveis explanatórias a própria série atrasada (*lagged*) de diversos períodos, obtém-se os seguintes resíduos:

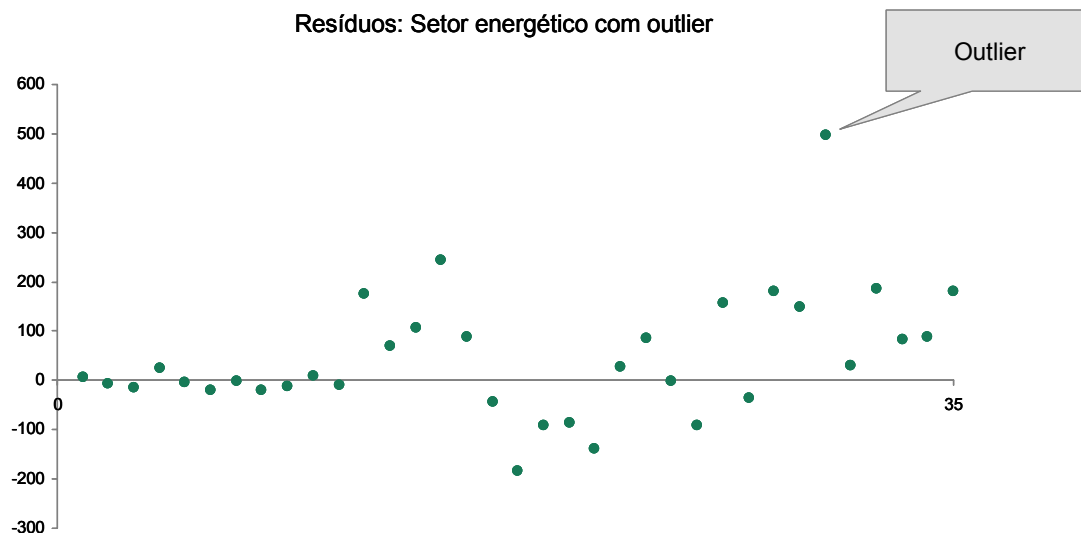


Gráfico 31: Setor energético – resíduos com outlier

Como se pode notar no gráfico acima, a base de dados utilizada possui um claro *outlier* cujo resíduo fica muito distante dos resultados previstos pelo modelo inicial.

Excluindo-se esse dado e rodando-se novamente o modelo encontra-se a seguinte equação:

Equação 32: Demanda do setor energético em função de modelo de auto-regressão I

$$y = 9,373 + 1,082 \times \text{Lag1} + \text{Erro}$$

Ou seja, o comportamento futuro da série depende do comportamento imediatamente passado da série de dados (R^2 de 98,96). O *t-stat* encontrado dos coeficientes é maior do que 2, mas o *t-stat* da constante foi de 0,33. Por ser uma medida do nível inicial da série e influenciar pouco nos resultados finais, releva-se este problema.

Analisando-se os resíduos do modelo, obtém-se o seguinte gráfico:

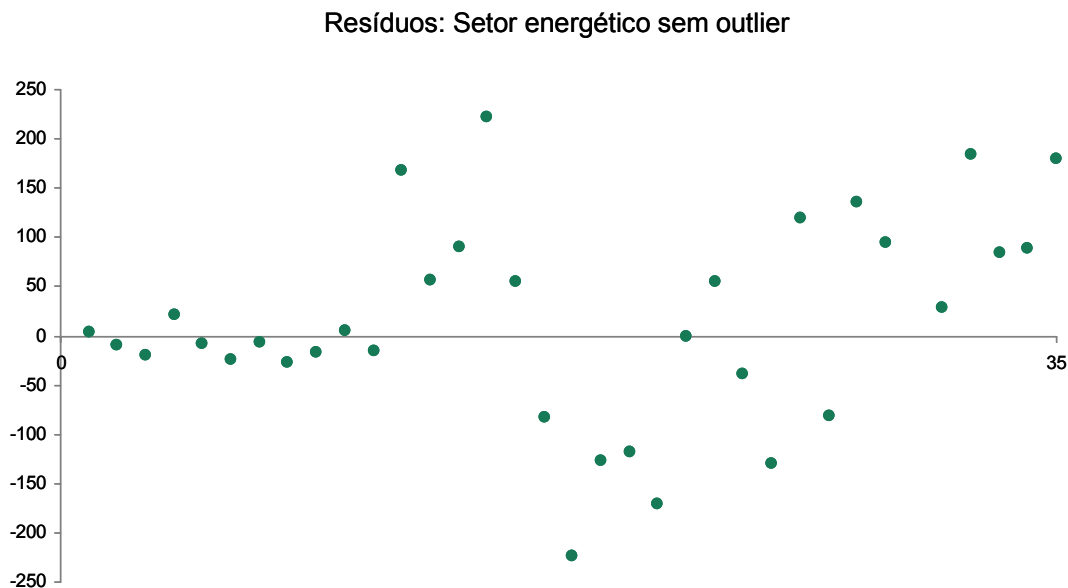


Gráfico 32: Setor energético – resíduos sem outlier

Pode-se observar que o mesmo não respeita completamente as hipóteses da regressão, já que no início os resíduos são muito próximos de zero, depois os mesmos se alternam entre positivos, negativos e positivos novamente.

Buscando-se aperfeiçoar o modelo anterior, pode-se realizar um *Stepwise-with-a-backward-look* tendo-se a série de resíduos como variável a ser explicada e como variáveis explanatórias as mesmas utilizadas no modelo de regressão explanatória do Setor Energético.

Desta forma, rodando-se a regressão, chega-se à seguinte equação:

Equação 33: Demanda do setor energético em função de modelo de auto-regressão II

$$\text{Erro} = 100 + 1,27 \times \text{CapacidadeInstaladaTérmica}$$

A qual possui um R^2 de 20,79% (ou seja, explica 20,79% do comportamento dos resíduos) e todos os *t-stat* maiores do que 1,5.

Analisando-se os resíduos (de segunda ordem) obtidos através desta regressão, encontra-se a seguinte distribuição:

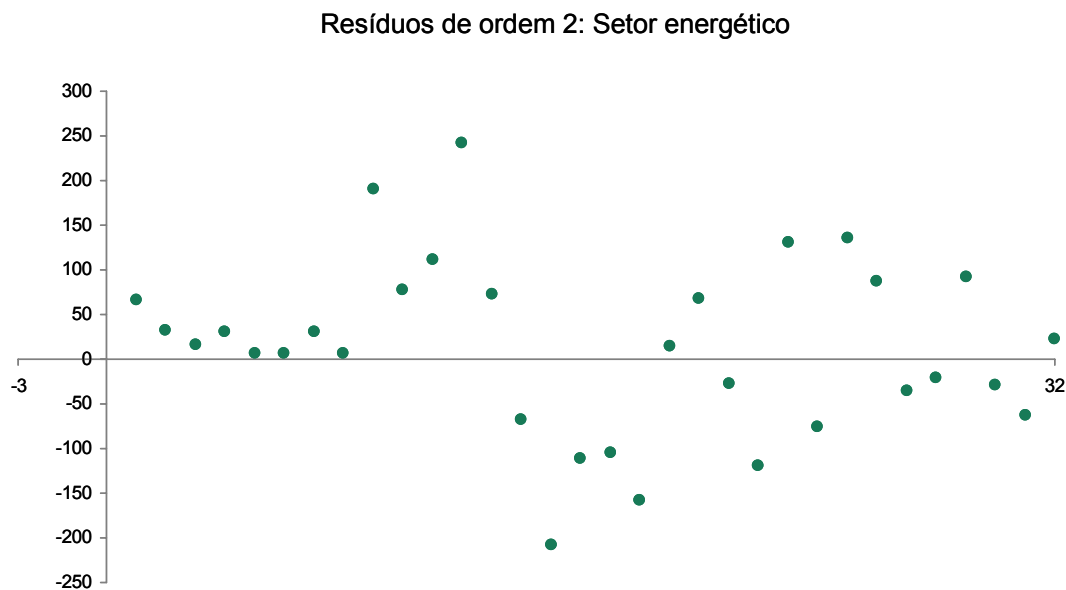


Gráfico 33: Setor Energético – resíduos de segunda ordem

Nota-se que estes resíduos apresentam problemas de ausência de homocedasticidade (principalmente no início da série) e padrão acima do zero também no início da série.

Porém, como a introdução de uma nova variável explanatória foi aceita no *Stepwise-with-a-backward-look*, considera-se que o resultado geral do modelo foi aprimorado.

Desta forma, trabalhando-se as duas equações parciais de regressão do setor energético, obtém-se:

Equação 34: Demanda do setor energético em função do modelo de auto-regressão III

$$y = 109,373 + 1,082 \times \text{Lag1} + 1,27 \times \text{CapInstTermica}$$

O qual fornece a seguinte previsão de demanda de gás para o Setor Energético:

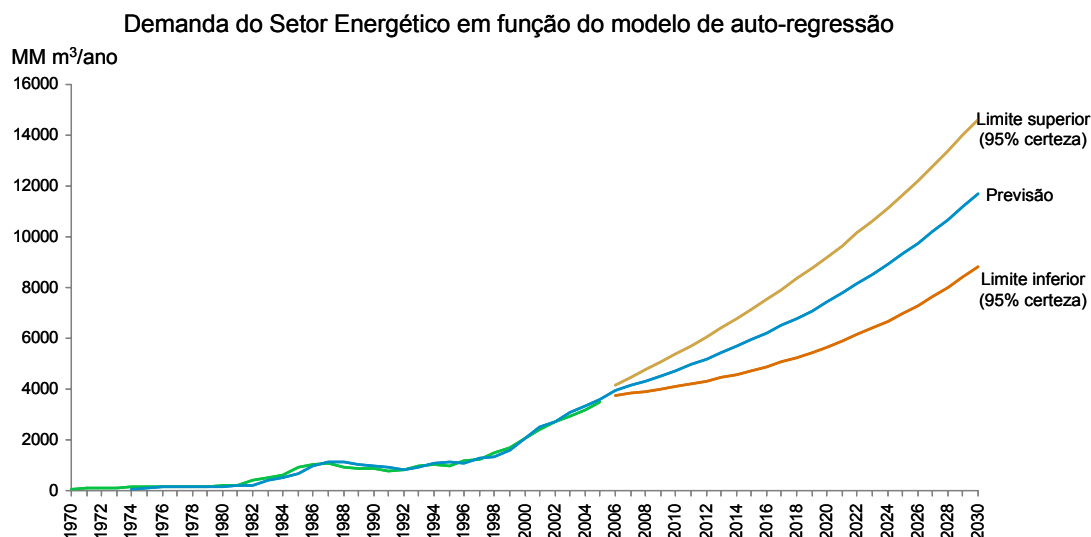


Gráfico 34: Demanda do setor energético em função do modelo de auto-regressão

Prevê-se um crescimento exponencial para a demanda de gás natural do setor de Energético com o modelo de auto-regressão, o qual, apesar de se mostrar apropriado para o curto prazo não se mostra necessariamente muito confiável para o longo prazo, já que a relação temporal que rege a série pode ser alterada no futuro.

Vale destacar que, devido à dificuldade matemática envolvida no cálculo de intervalos de predição para auto-regressões, os intervalos de confiança da figura acima foram estimados. Calculou-se o intervalo de confiança da previsão um período à frente e depois se estimou os intervalos de confiança das outras previsões multiplicando-se a primeira previsão pela mesma taxa de crescimento verificada na previsão por Holt (a proporção com o método de Holt foi adotada pelo fato de ambos os modelos serem baseados em séries temporais).

7.2.4 Energético - Comparação dos modelos

Comparando-se os métodos de previsão utilizados para a previsão da demanda de gás do Setor Energético, obtém-se o seguinte resultado:

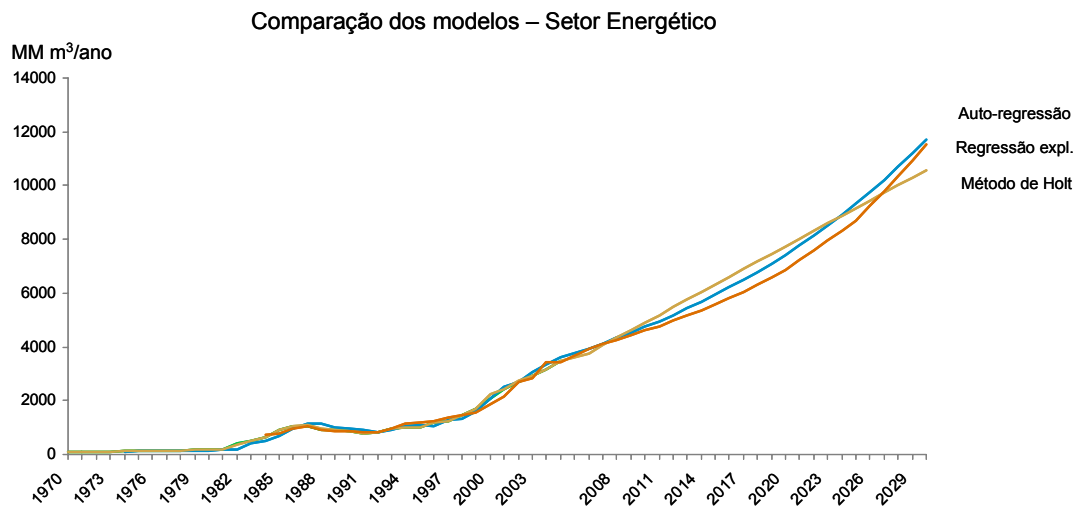


Gráfico 35: Comparação dos modelos – Energético

No gráfico acima, a título de simplificação, foram incluídos apenas os resultados do modelo de regressão explanatória para o cenário-base de preço relativo do gás natural em relação ao óleo combustível (já que os resultados dos diferentes cenários são bastante próximos entre si).

Analisando-se os gráficos acima se observa inicialmente que os resultados dos três modelos, cujas naturezas são bastante distintas entre si, encontram-se bastante próximos, tanto no curto quanto no médio e longo prazo.

Montando-se um quadro comparativo dos modelos de previsão utilizados chega-se à seguinte tabela:

Tabela 9: Quadro comparativo de modelos – Setor Energético¹

Setor Energético	Regressão Explanatória	Auto-regressão	Holt
Número de dados históricos incluídos no modelo	16	34	36
Relação Intervalo de Confiança/Previsão em 2030	26%	27%	22%
R2 ou R2-ajustado encontrado	98,0%	98,9%	n/a
MSE	15.007,2	10.656,7	13.622,0
MAE	87,8	76,9	77,5
MAPE	7%	11%	10%
Problemas com resíduos de 1a ordem	*Não-homocedasticidade *Resíduos auto-correlacionados	*Não-homocedasticidade	n/a
Problemas com resíduos de 2a ordem	n/a	*Não-homocedasticidade *Sequência acima do zero no início da série	n/a
Método conceitualmente mais adequado para previsões de...	Curto/Médio prazo	Curto prazo	Curto prazo
Alinhamento de resultados com outros métodos?	Auto-regressão e Holt	Regr. explanatória e Holt	Regr. explanatória e auto-regressão

Analisando-se individualmente cada modelo percebe-se que o de regressão explanatória possui seus resultados prejudicados pelo baixo número de dados históricos (somente 16) e pelo fato da série de resíduos não respeitar as hipóteses da regressão.

Já o modelo de Holt apresenta parâmetros da mesma ordem de grandeza que os outros modelos, não se destacando em nenhum aspecto.

O modelo de auto-regressão também sofre com problemas em seus resíduos, mas por outro lado possui um elevado número de dados históricos incluídos no modelo. Além disso, nota-se que durante a maior parte do intervalo de previsão tal modelo apresenta resultados intermediários entre os modelos de regressão explanatória e Holt.

Porém, na previsão da demanda do Setor Energético nota-se, acima de tudo, que apesar das forças e deficiências de cada modelo, a convergência gráfica dos 3 modelos ressalta a força dos resultados encontrados.

Desta forma, recomenda-se a utilização dos resultados do modelo de auto-regressão como previsão da demanda do Setor Energético.

¹ Nos casos de modelos mistos que misturam regressão explanatória e auto-regressão (ou vice-versa) o R^2 total é encontrado pela equação (R^2 total = $R^2_{\text{primeira ordem}} + (1 - R^2_{\text{primeira ordem}}) \times R^2_{\text{segunda ordem}}$)

7.3 Setor Industrial

A seguir serão desenvolvidos diferentes modelos de previsão explanatórios e baseados em séries temporais para a demanda do setor industrial.

7.3.1 Setor Industrial – Holt

Analisando o gráfico com a série temporal da demanda do setor industrial, observamos que o mesmo possui uma tendência de crescimento e pouca ou nenhuma sazonalidade.

Desta forma, de acordo com a Revisão Bibliográfica deste trabalho, pode-se concluir que um modelo que se mostra adequado para a previsão destas séries é o modelo de Holt.

Utilizando-o para o consumo do setor industrial obtém-se o seguinte gráfico:

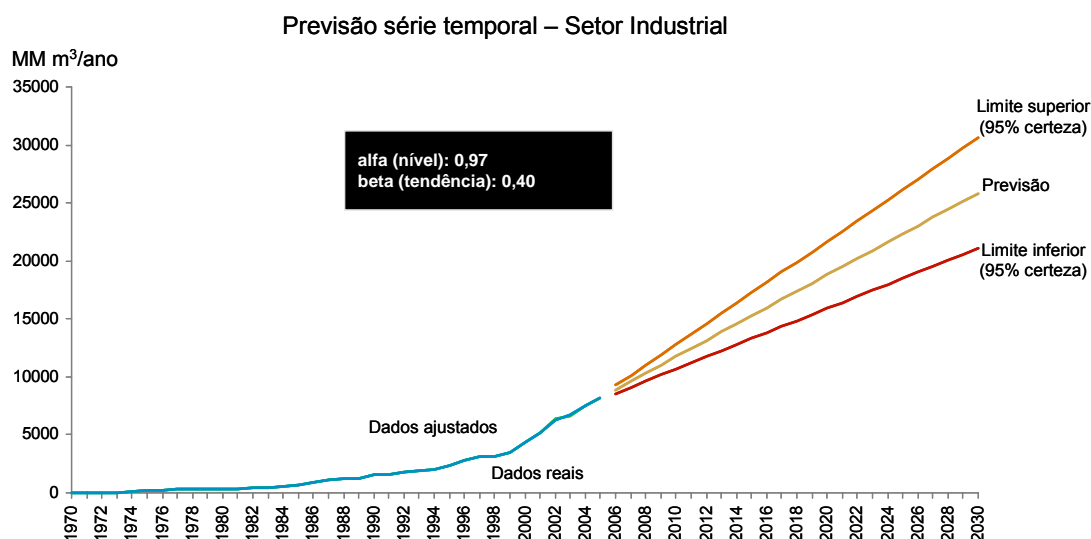


Gráfico 36: Modelo de Holt para demanda do Setor Industrial

Nele pode-se observar uma previsão com tendência futura linear, de acordo com o funcionamento do modelo utilizado. Porém, vale ressaltar que, por considerar que o comportamento futuro de uma série temporal refletirá o seu passado, o método de Holt costuma obter resultados melhores no curto prazo do que no longo prazo.

7.3.2 Setor Industrial – Regressão Explanatória

Buscando-se encontrar um modelo de regressão explanatória para a demanda de gás natural do Setor Industrial listou-se as seguintes variáveis como possíveis explanatórias:

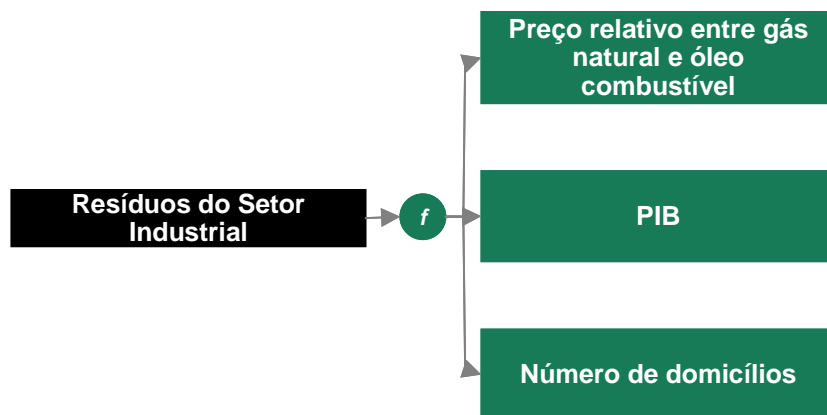


Figura 16: Possíveis variáveis explanatórias para demanda industrial

Ou seja, considera-se que, de alguma forma, a demanda de gás do setor industrial possa estar relacionada com:

- **Preço relativo entre gás natural e óleo combustível:** já que as indústrias podem migrar para o combustível que apresente uma melhor relação custo-benefício, impactando assim diretamente na demanda de gás deste setor (fenômeno facilmente detectável em notícias recentes da mídia frente ao aumento do preço do gás de origem boliviana)
- **PIB:** já que quanto maior o PIB nacional, possivelmente maior é o nível de operação industrial, o que acarretaria num possível maior consumo de gás
- **Número de domicílios:** acredita-se que, de alguma forma, o número de domicílios existentes no Brasil possa estar relacionado com a demanda industrial por gás natural

Rodando-se um *Stepwise-with-a-backward-look*, encontra-se a seguinte equação de regressão para a demanda de gás do setor Industrial (R^2 de 88,87% e todos os *t-stat* maiores do que 2):

Equação 35: Demanda industrial em função de modelo de regressão explanatória I

$$y = -11637 + 95,9 \times \text{PIB}$$

Analisando-se os resíduos deste modelo, encontramos a seguinte distribuição:

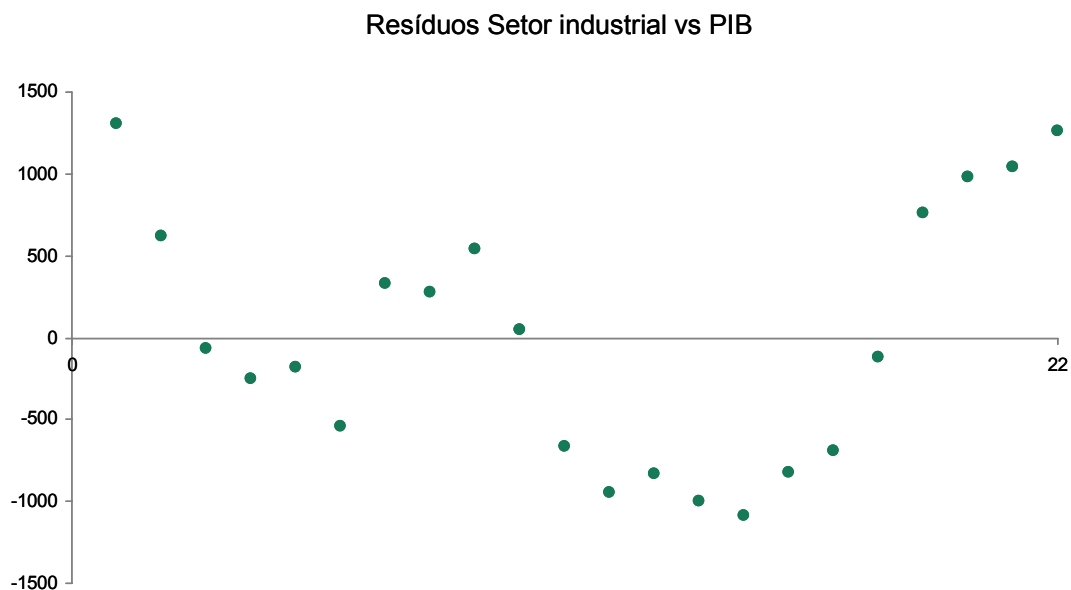


Gráfico 37: Resíduos Setor Industrial vs PIB

Como se pode notar, os resíduos acima violam a hipótese de normalidade por serem altamente autocorrelacionados. Além disso, os mesmos não se mostram homocedásticos ao longo do eixo x. Plotando-se o ACF dos erros obtém-se:

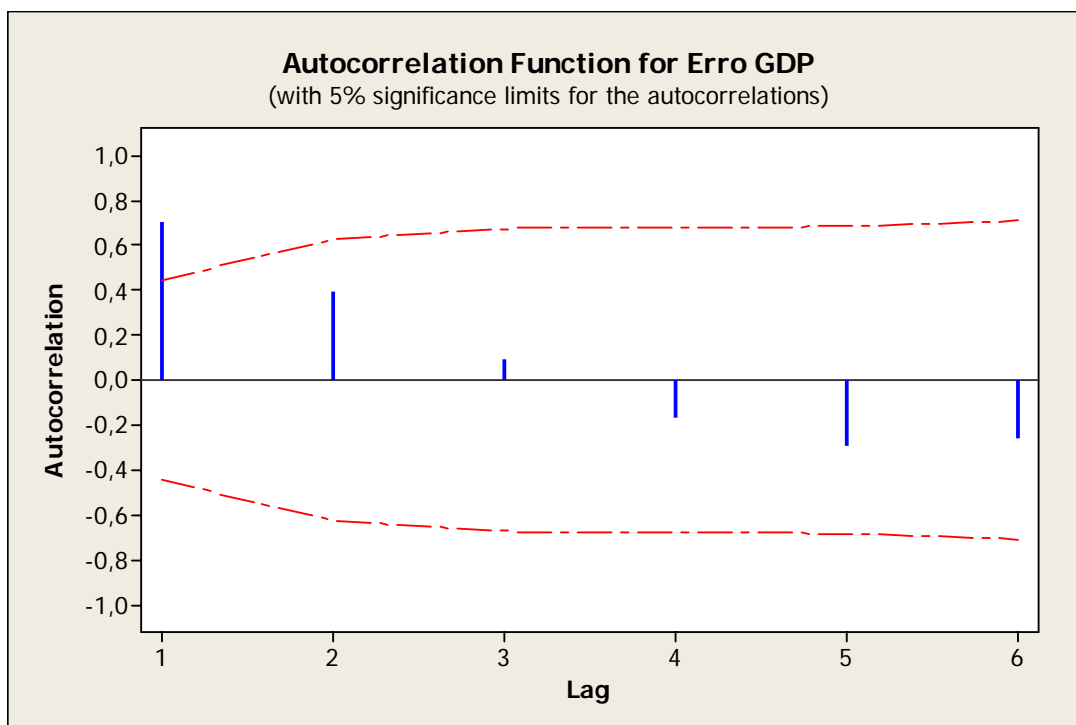


Gráfico 38: ACF dos resíduos Industrial vs PIB

Percebe-se que no *lag* 1 o ACF fica além dos limites permitidos. Isto leva a crer que a incorporação de uma variável temporal ao modelo anterior pode trazer resultados melhores do que os obtidos até então.

Desta forma, rodando-se um Stepwise forward-with-a-backward look dos resíduos com os *lags* da série de demanda industrial obtém-se a seguinte equação:

Equação 36: Demanda industrial em função de modelo de regressão explanatória II

$$\text{Resíduo} = -14,1 + 0,8 \times \text{Lag1}$$

Tal regressão apresenta um R^2 de 65,6% e considera-se que a mesma melhora significativamente os resultados obtidos, já que foi aceita pelo modelo de Stepwise forward-with-a-backward look. Porém o *t-stat* encontrado da constante da equação é de -1,28, valor inferior ao desejado mas que, por influir muito pouco nos resultados finais é relevado.

Assim, concluímos que a equação final de regressão do Setor Industrial fica sendo:

Equação 37: Demanda industrial em função de modelo de regressão explanatória III

$$y = -11651,1 + 95,9 \times \text{PIB} + 0,806 \times \text{Lag1}$$

Analisando-se os resíduos da equação acima observamos que os mesmos apresentam as mesmas características dos resíduos da demanda industrial em função do PIB. Porém o gráfico abaixo apresenta uma amplitude de dados muito menor do que os resíduos de ordem 1 derivados da regressão direta em relação ao PIB.

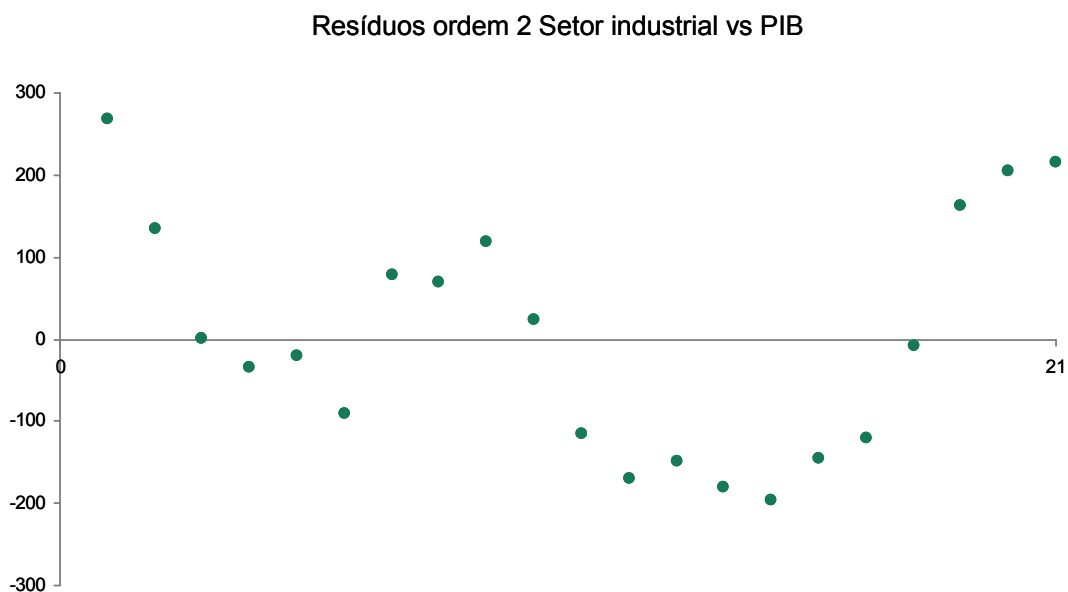


Gráfico 39: Resíduos de ordem 2 da regressão explanatória do Setor Industrial

Baseado nas previsões para número de PIB (vide Apêndice) no futuro obtém-se a seguinte previsão para a demanda de Industrial:

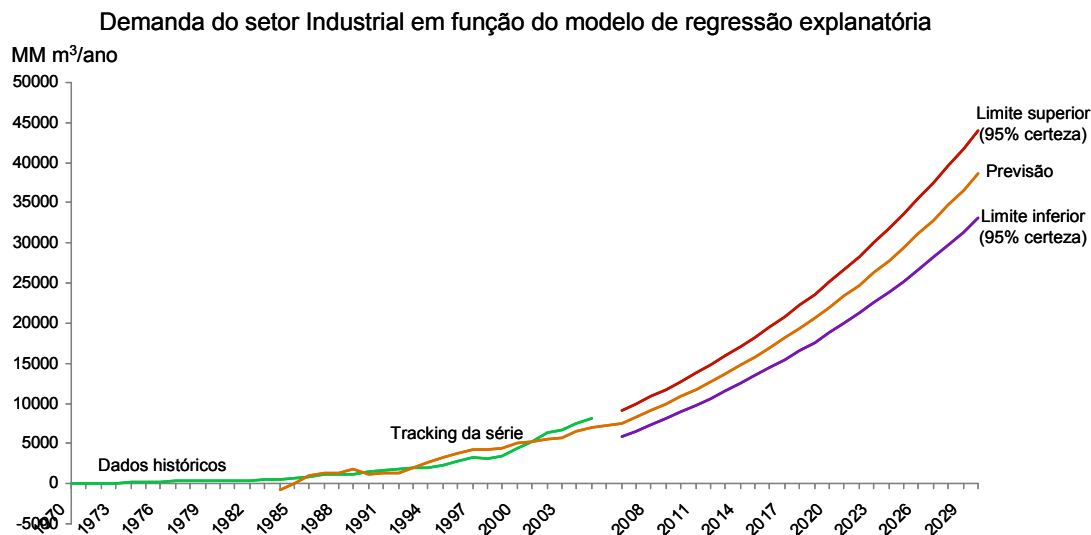


Gráfico 40: Demanda do setor industrial em função do modelo de regressão explanatória

Nota-se no gráfico acima o perfil exponencial da curva de previsão. Além disso, pode-se observar que curvas dos limites dos intervalos de confiança formam um ângulo entre si de abertura aparentemente intermediária.

7.3.3 Setor Industrial – Auto-regressão

Analisando-se agora a demanda do setor Industrial de um ponto de vista explanatório, percebe-se inicialmente que a mesma é imensamente influenciada pelo seu comportamento recente (assim como a demanda do Setor Energético), conforme demonstrado pelo gráfico de ACF:

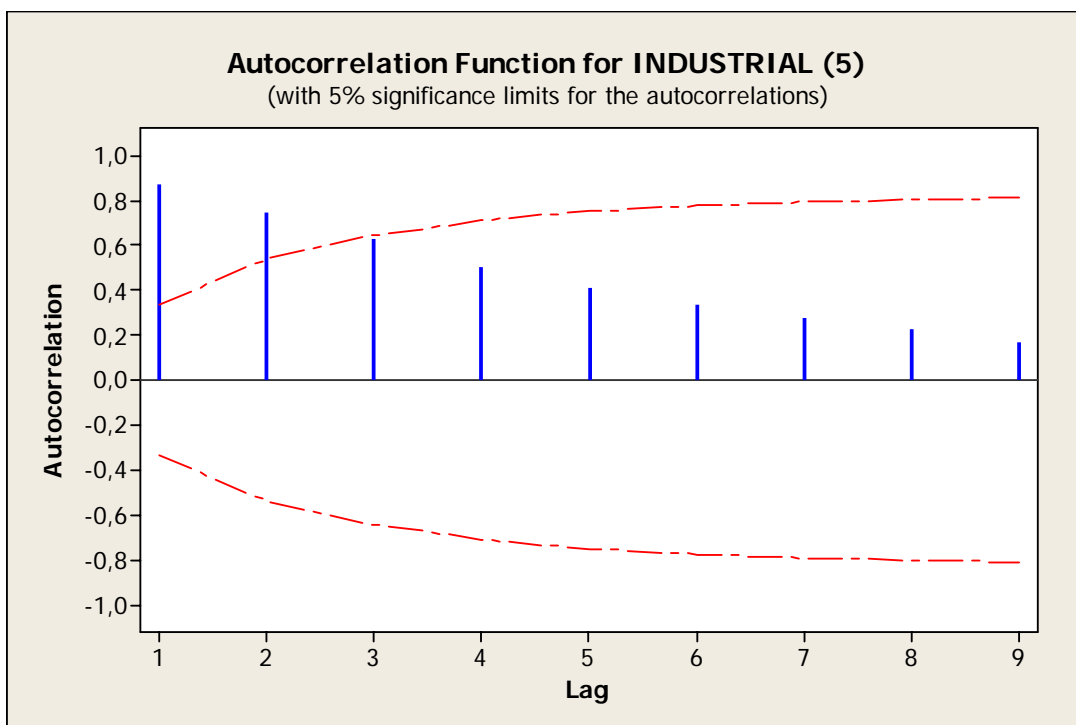


Gráfico 41: ACF da demanda do Setor Industrial

Pelo gráfico acima, pode-se notar que a autocorrelação nos dois primeiros *lags* da série extrapola os limites de controle. Desta forma, deve-se considerar a influência temporal na modelagem da série temporal.

Desta forma, inicialmente buscará se explicar a série temporal de demanda de gás do setor industrial através de uma equação de auto-regressão.

Rodando-se um modelo de regressão *Stepwise-with-a-backward-look* com parâmetros de “Alpha-to-enter” e “Alpha-to-remove” de 0,15, sendo as possíveis variáveis explanatórias a própria série atrasada (*lagged*) de diversos períodos, obtém-se o seguinte resultado de resíduos:



Gráfico 42: Setor industrial – resíduos de primeira ordem

Como se pode perceber pelo gráfico acima, os resíduos da regressão apresentam certo problema de homocedasticidade (variância aumenta conforme o x aumenta), além de seqüências muito grandes de resíduos abaixo do zero, principalmente no início da série.

A equação que representa o modelo é dada por:

Equação 38: Demanda industrial em função de modelo de auto-regressão I

$$y = 30,13 + 1,114 \times \text{Lag1}$$

Ou seja, o comportamento futuro da série depende do comportamento imediatamente passado da série de dados (R^2 de 99,26%). O t -stat encontrado do coeficiente é maior do que 2, mas o t -stat da constante foi de 0,66. Por ser uma medida do nível inicial da série e influenciar pouco nos resultados finais, releva-se este problema.

Buscando-se aperfeiçoar o modelo anterior, pode-se realizar uma Stepwise-with-a-backward-look tendo a série dos resíduos como variável a ser explicada e “Preço relativo entre gás natural e óleo combustível”, “PIB” e “Número de domicílios” como possíveis variáveis explanatórias.

Fazendo-se isso, chega-se à seguinte equação:

Equação 39: Demanda industrial em função de modelo de auto-regressão II

$$\text{Erro} = 404,5 + 360 \times \text{Preço Relativo}$$

A qual possui um R^2 de 5,88% (ou seja, explica 5,88% do comportamento dos resíduos). Apesar de ser um valor de R^2 baixo, considera-se que ele melhora estatisticamente a previsão sendo realizada, dado que a variável “Preço Relativo” foi aceita no Stepwise-with-a-backward-look (o qual poderia não deixar mais nenhuma variável entrar no modelo).

Analisando-se os resíduos (de segunda ordem) obtidos através desta regressão, encontra-se a seguinte distribuição:

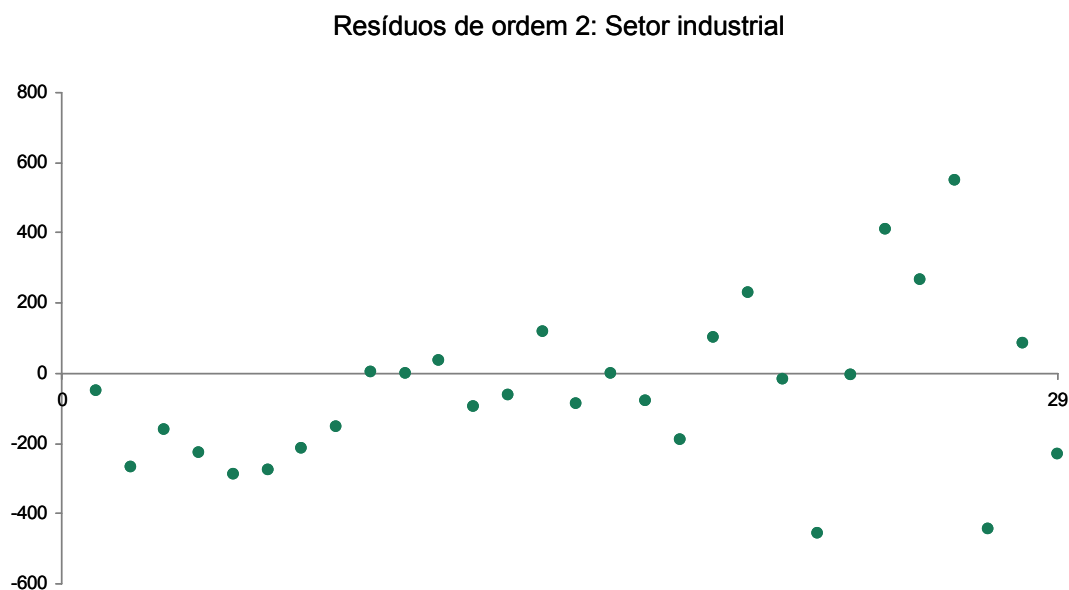


Gráfico 43: Setor Industrial – resíduos de segunda ordem

Nota-se que estes resíduos apresentam o mesmo problema de heterocedasticidade dos resíduos de primeira ordem, além de possuir uma grande série de pontos abaixo do zero no início da série.

Desta forma, trabalhando-se as duas equações parciais de regressão do setor energético, obtém-se:

Equação 40: Demanda industrial em função de modelo de auto-regressão III

$$y = 434,63 + 1,114 \times \text{Lag1} + 360 \times \text{Preço Relativo}$$

Note-se que, assim como no caso da demanda do setor industrial, preço relativo se mostrou positivamente correlacionado com a demanda do setor industrial, o que é um aparente contra-senso com a lógica subjacente a tal negócio. Porém, dado que este foi o resultado estatístico encontrado, procede-se com a análise do mesmo.

A equação acima fornece a seguinte previsão de demanda de gás para o Setor Industrial:

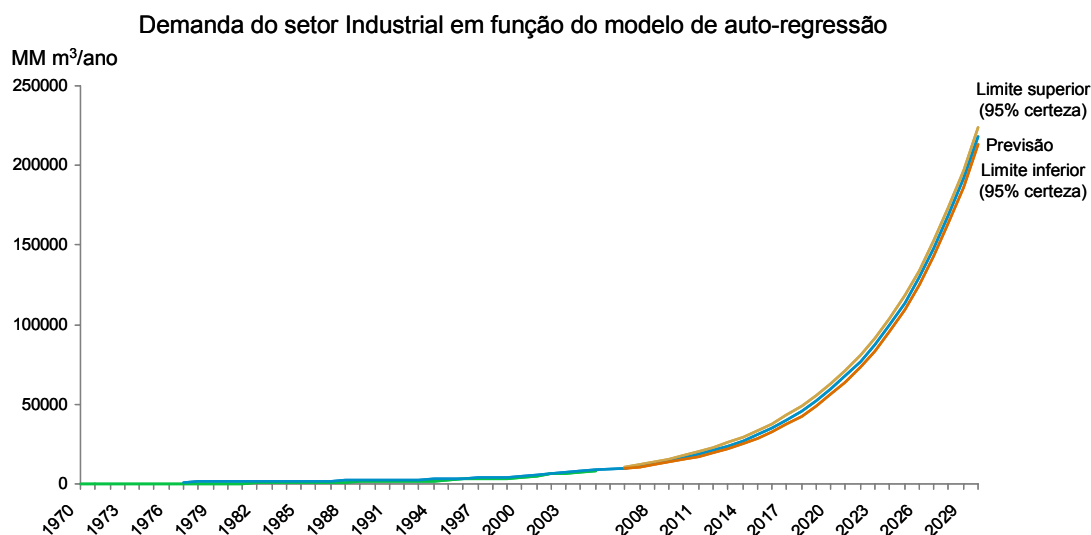


Gráfico 44: Demanda do setor industrial em função do modelo de auto-regressão

Neste caso prevê-se um crescimento exponencial para a demanda de gás natural do Setor Industrial.

Vale destacar que, devido à dificuldade matemática envolvida no cálculo de intervalos de predição para auto-regressões, os intervalos de confiança da figura acima foram estimados.

Calculou-se o intervalo de confiança da previsão um período à frente e depois se estimou os intervalos de confiança das outras previsões multiplicando-se a primeira previsão pela mesma taxa de crescimento verificada na previsão por Holt (a proporção com o método de Holt foi adotada pelo fato de ambos os modelos serem baseados em séries temporais).

7.3.4 Industrial - Comparação dos modelos

Comparando os resultados obtidos pelo Modelo de Holt, pelo modelo de regressão explanatória e pelo modelo de auto-regressão, chega-se ao seguinte gráfico:

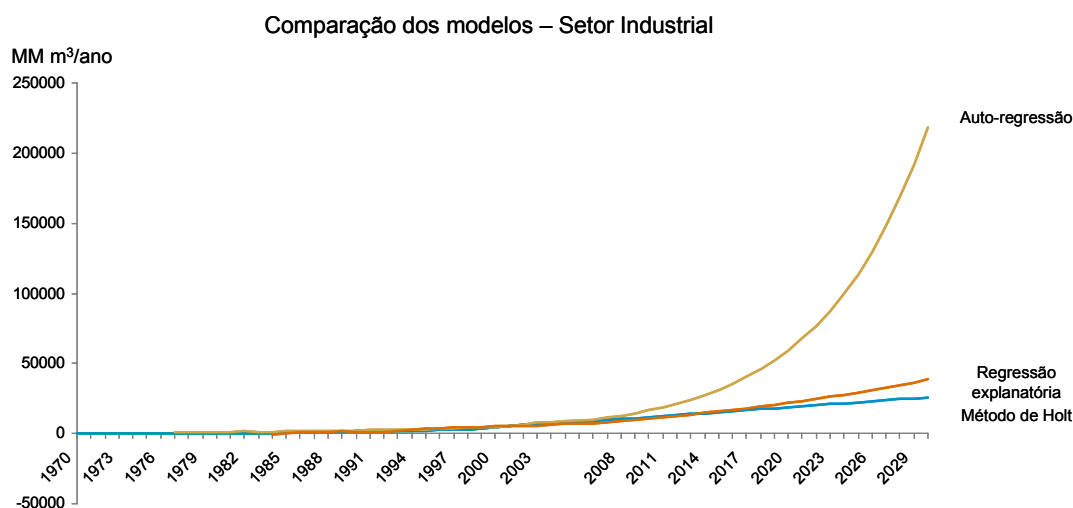


Gráfico 45: Comparação dos modelos – Industrial

A primeira observação que pode ser feita através do gráfico acima é de que o modelo baseado em regressão pelo PIB ficou, no longo prazo, muito mais próximo do modelo baseado no método de Holt do que no modelo baseado em auto-regressão, apesar do fato do primeiro possuir um perfil de crescimento exponencial e do segundo apresentar um perfil de crescimento linear.

Nota-se também que até por volta de 2010 as três séries apresentam resultados bastante próximos entre si. Após isso a série baseada em auto-regressão separa-se das demais com velocidade crescente.

Montando-se um quadro comparativo dos modelos de previsão utilizados chega-se à seguinte tabela:

Tabela 10: Quadro comparativo de modelos – Setor Industrial

Setor Industrial	Regressão Explanatória	Auto-regressão	Holt
Número de dados históricos incluídos no modelo	22	29	36
Relação Intervalo de Confiança/Previsão em 2030	14%	2%	18%
R2 ou R2-ajustado encontrado	88,9%	99,2%	n/a
MSE	571.549,4	40.918,0	45.170,3
MAE	653,5	137,0	138,9
MAPE	35%	23%	23%
Problemas com resíduos de 1a ordem	*Não-homocedasticidade *Resíduos auto-correlacionados	*Não-homocedasticidade *Sequência abaixo do zero no início da série	n/a
Problemas com resíduos de 2a ordem	*Não-homocedasticidade *Resíduos auto-correlacionados	*Não-homocedasticidade *Sequência abaixo do zero no início da série	n/a
Método conceitualmente mais adequado para previsões de...	Curto/Médio prazo	Curto prazo	Curto prazo
Alinhamento de resultados com outros métodos?	Holt	Nenhum	Regr. explanatória

Através da tabela acima se pode verificar que o modelo de regressão explanatória possui medidas de erros maiores do que os demais. Além disso, o mesmo é o que possui o menor número de dados históricos incluídos no modelo (devido a limitações de dados de PIB), além de problemas com seus resíduos. Por outro lado, ele apresenta grande alinhamento com os resultados obtidos por Holt e é, conceitualmente, o único modelo adequado para fazer previsões de médio prazo.

Já o modelo de auto-regressão possui uma excelente relação intervalo de confiança/previsão, além de um bom número de pontos incluídos no modelo. Por outro lado, também apresenta problemas com seus resíduos.

Já o método de Holt possui uma grande força que é o alinhamento de seus resultados com o modelo de regressão explanatória. Em relação aos outros parâmetros de comparação ele apresenta um resultado intermediário em comparação com os outros modelos obtidos.

Sendo assim, recomenda-se a utilização dos resultados dos modelos de regressão explanatória ou de Holt como previsão da demanda do setor industrial.

Desta forma, no caso da demanda do setor Industrial, pode-se considerar que os modelos de previsão aqui desenvolvidos apresentam bons resultados para o curto prazo (aproximadamente 5 anos), mas que seus resultados de longo prazo devem ser analisados frente à experiência de algum especialista da área de forma a validar qual(is) do(s) modelo(s) apresenta(m) resultados mais coerentes com a realidade do mercado.

7.4 Setor Residencial

A seguir serão desenvolvidos diferentes modelos de previsão explanatórios e baseados em séries temporais para a demanda do setor residencial.

7.4.1 Residencial – Método de Holt

Analisando-se inicialmente o gráfico com a série temporal da demanda do setor residencial, observamos que a mesma possui uma tendência de crescimento e pouca ou nenhuma sazonalidade.

Desta forma, de acordo com a Revisão Bibliográfica deste trabalho, pode-se concluir que um modelo que se mostra adequado para a previsão destas séries é o modelo de Holt.

Utilizando-o para o consumo do setor energético obtém-se o seguinte gráfico:

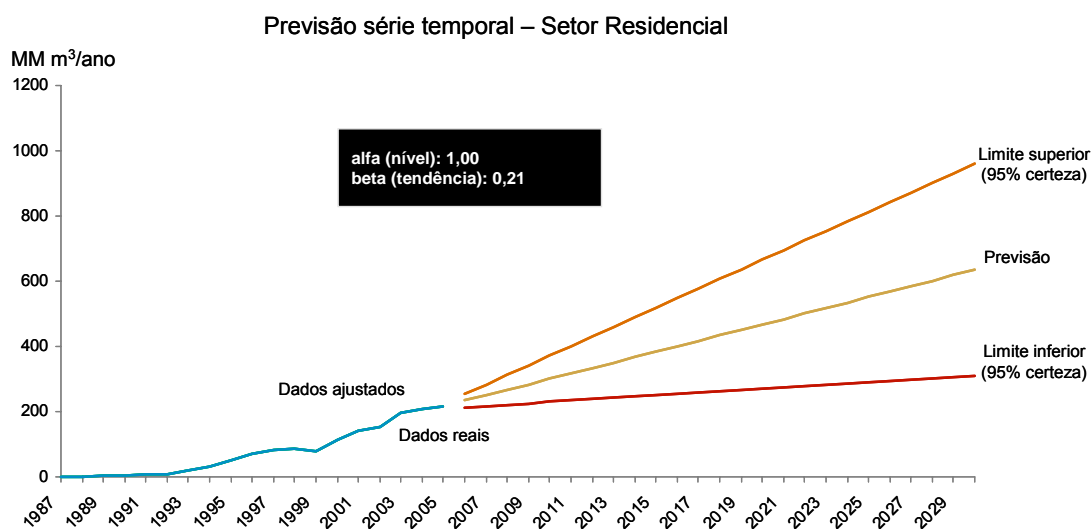


Gráfico 46: Modelo de Holt para demanda do Setor Residencial

Nele pode-se observar uma previsão com tendência futura linear, de acordo com o funcionamento do modelo utilizado. Porém, vale ressaltar que, por considerar que o comportamento futuro de uma série temporal refletirá o seu passado, o método de Holt costuma obter resultados melhores no curto prazo do que no longo prazo.

Além disso, pode-se, visualmente, perceber que as curvas do intervalo de confiança formam um ângulo bastante aberto entre si se comparado às previsões por método de Holt anteriormente realizadas. Isso ocorre devido a dois fatores: baixo número de dados históricos (somente 19) e ao fato da curva de demanda histórica do Setor Residencial dar um “soluço” entre 1997 e 2003.

7.4.2 Residencial – Regressão Explanatória

Buscando-se realizar uma previsão por regressão para o setor Residencial, pode-se pensar nas seguintes variáveis que teriam algum relacionamento lógico com a mesma:

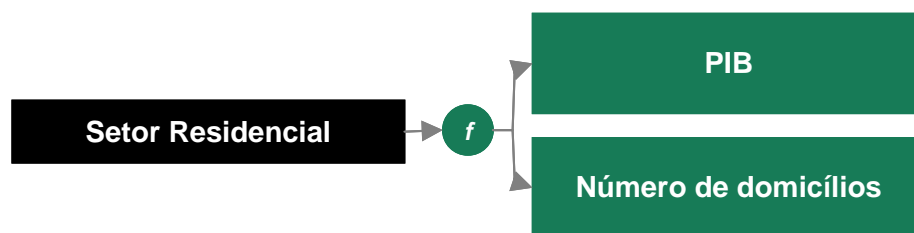


Figura 17: Possíveis variáveis explanatórias para a demanda do setor residencial

Ou seja, considera-se que, de alguma forma, a demanda de gás do setor residencial possa estar relacionada com o PIB e/ou o número de domicílios brasileiros. Note-se que o relacionamento entre número de domicílios e demanda do setor residencial é intuitiva (quanto mais domicílios, maior deve ser, de forma geral, a demanda do setor residencial). Já a variável PIB está sendo utilizada aqui como um “coringa”, já que é um indicador macroeconômico brasileiro de grande relevância que pode mostrar algum relacionamento estatístico com a variável em estudo.

Rodando no Minitab® uma regressão *Stepwise forward-with-a-backward look* com essas duas variáveis e os parâmetros de “Alpha-to-enter” e “Alpha-to-remove” de 15%, obtém-se a seguinte equação:

Equação 41: Demanda residencial em função de modelo de regressão explanatória I

$$y = 494,7 + 3,61 \times \text{PIB}$$

Como se pode observar, a regressão indicou que a relação entre demanda residencial e PIB é estatisticamente mais relevante do que demanda residencial e número de domicílios (não se pode usar PIB e número de domicílios numa mesma regressão múltipla por problemas de multicolinearidade entre estas duas variáveis).

A equação de regressão encontrada possui um R^2 de 94,9%, resultado bastante elevado, mas que é enfraquecido pelo baixo número de pontos utilizados (19) e pela análise dos resíduos da regressão. Além disso, todos os *t-stat* encontrados foram maiores do que 2.

Analisando-se os resíduos que essa equação gera obtém-se a seguinte distribuição:



Gráfico 47: Resíduos da regressão do Setor Residencial

Percebe-se claramente no gráfico acima que os resíduos são autocorrelacionados. Ou seja, o resultado de um resíduo depende do resultado do resíduo anterior (evidenciado pela distribuição com formato de senoide).

Outra análise que pode ser realizada é plotar o gráfico do ACF dos erros, para verificar se há alguma influência temporal relevante ao problema que não tenha sido incorporada ao modelo. O gráfico abaixo ilustra isso:

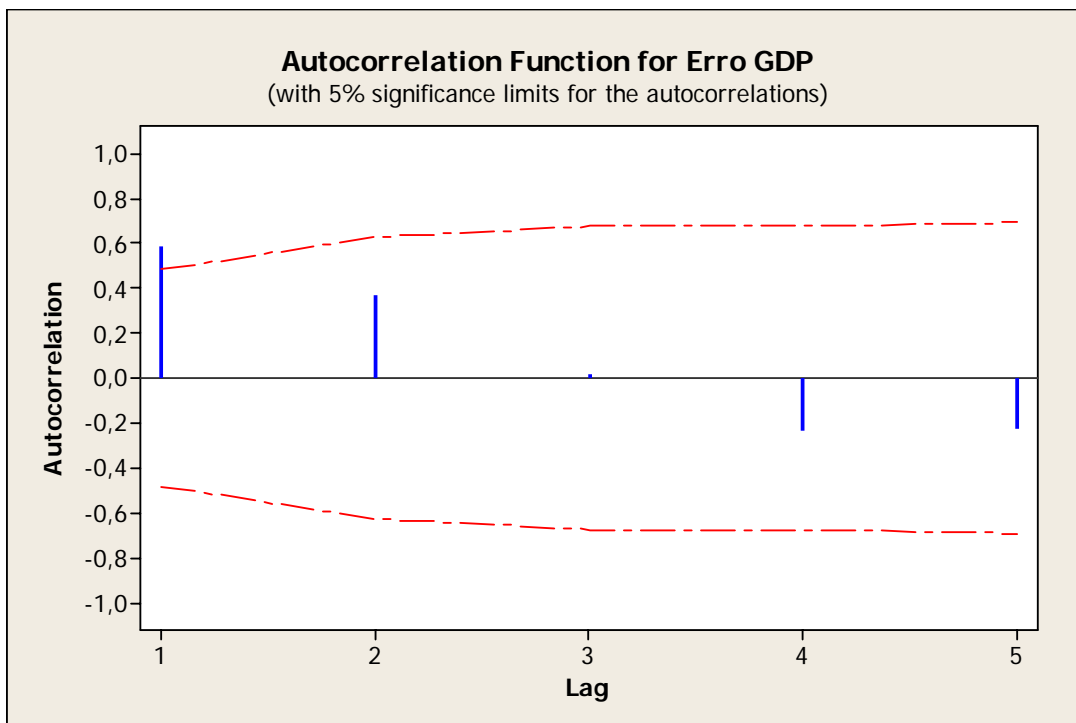


Gráfico 48: Resíduos Residencial vs PIB

Percebe-se que no lag 1 o ACF fica além dos limites permitidos. Isso leva a crer que a incorporação de uma variável temporal ao modelo anterior pode trazer resultados melhores do que os obtidos até então.

Desta forma, rodando-se um Stepwise forward-with-a-backward look dos resíduos com os lags da série de demanda residencial obtém-se a seguinte equação:

Equação 42: Demanda residencial em função de modelo de regressão explanatória II

$$\text{Resíduo} = 5,034 + 0,06 \times \text{Lag1}$$

Considera-se que apesar da regressão apresentar um R^2 de apenas 5,63% ela melhora significativamente os resultados obtidos, visto que foi aceita pelo modelo de *Stepwise forward-with-a-backward look*.

Assim, concluímos que a equação final de regressão do Setor Residencial fica sendo:

Equação 43: Demanda residencial em função de modelo de regressão explanatória III

$$y = 499,7 + 3,61 \times \text{PIB} + 0,06 \times \text{Lag1}$$

Analisando-se os resíduos da equação acima observamos que os mesmos apresentam as mesmas características dos resíduos da demanda residencial em função do PIB. Isso se deve ao fato da inclusão do *Lag1* na equação estar fornecendo um grau de explicação a mais muito pequeno (cerca de 5%).

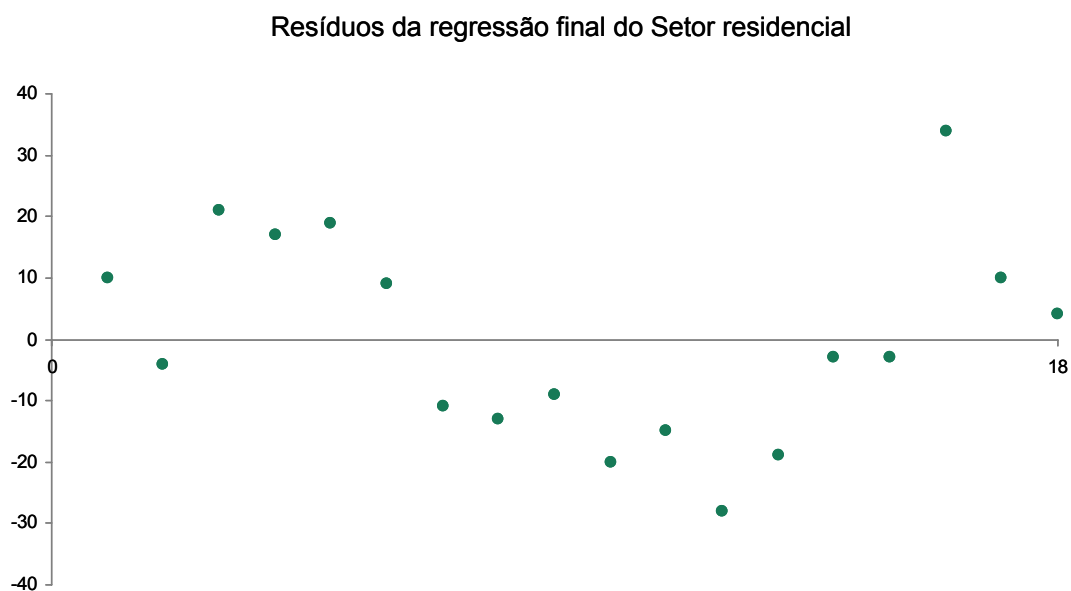


Gráfico 49: Setor Residencial - Resíduos de segunda ordem

Baseado nas previsões para número de PIB (vide Apêndice) no futuro obtém-se a seguinte previsão para a demanda Residencial:

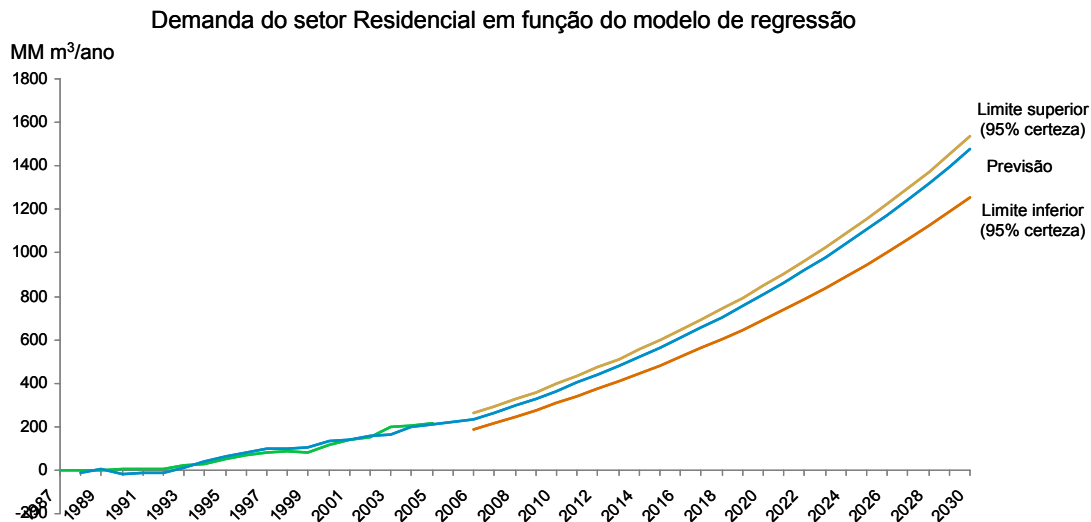


Gráfico 50: Previsão da demanda residencial em função do PIB

Como se pode observar, através deste modelo projeta-se um crescimento com perfil exponencial para a demanda Residencial no futuro.

Isso se deve na mesma estar ancorada no tamanho do PIB nacional, cuja previsão prevê um crescimento exponencial para o mesmo. Desta forma, como PIB e demanda Residencial estão linearmente relacionados pelo modelo obtido, a demanda residencial também apresenta este mesmo perfil de crescimento.

Um comentário que deve ser feito é a respeito da não-equidistância das curvas de intervalo de confiança em relação à curva de previsão. Isso se deve ao fato das curvas de intervalo de confiança terem sido calculadas com base na regressão de primeira ordem realizada (sem a incorporação da variável *Lag1* ao modelo, a qual gerou a regressão de segunda ordem), a título de simplificação dos cálculos. Como a variável *Lag1* apresenta pouca influência no modelo final, isso causa um impacto pequeno nas curvas de intervalo de confiança que, apesar de não estarem matematicamente precisas no gráfico acima, funcionam de forma muito prática para representar a incerteza associada ao modelo utilizado.

Vale ressaltar que esta previsão considera que a relação descrita pela equação do modelo se manterá no futuro, ou seja, que a demanda Residencial continuará seguindo a mesma proporcionalidade com o PIB brasileiro, o que pode não ser verdade no futuro (ou mesmo pode ter sido apenas uma coincidência estatística, já que se esperava que a demanda Residencial tivesse uma maior relação com número de domicílios, não com PIB). Além disso,

ela baseia suas previsões em previsões de PIB no futuro, as quais, por si só, possuem erros e incertezas associadas.

7.4.3 Residencial – Auto-regressão

Realizando-se a análise de ACFs diretamente para a série temporal de demanda do setor Residencial encontra-se o seguinte gráfico:

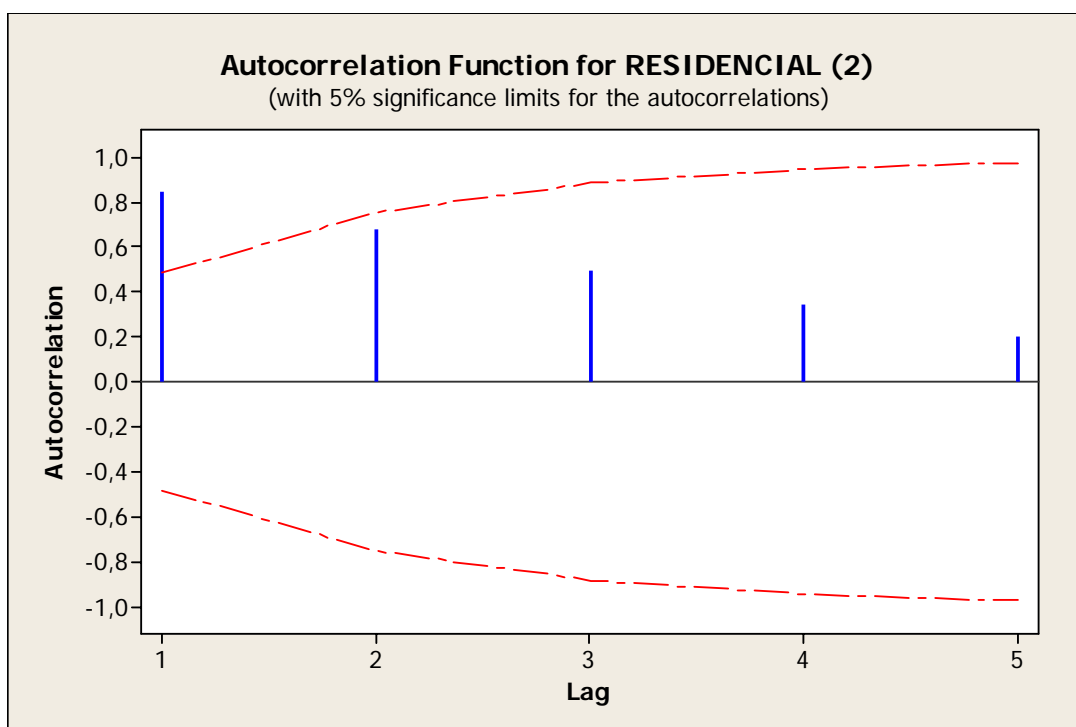


Gráfico 51: ACF da demanda do setor Residencial

Tal gráfico indica que uma variável temporal deveria ser incluída na previsão de demanda do setor residencial, o que nos leva a buscar um terceiro modelo.

Realizando-se uma auto-regressão da demanda do setor Residencial, encontra-se a seguinte equação:

Equação 44: Demanda residencial em função de modelo de auto-regressão

$$y = 8,007 + 1,065 \times \text{Lag1}$$

Tal regressão apresenta um R^2 -ajustado de 97,28%, valor bastante elevado que pode ser justificado pelo baixo número de dados incluídos no modelo (apenas 18, que é o que a base histórica do Balanço Energético Nacional – BEN - fornece). Além disso, todos os *t-stat* encontrados foram maiores do que 1,5.

Analisando-se os resíduos deste modelo encontra-se a seguinte distribuição:

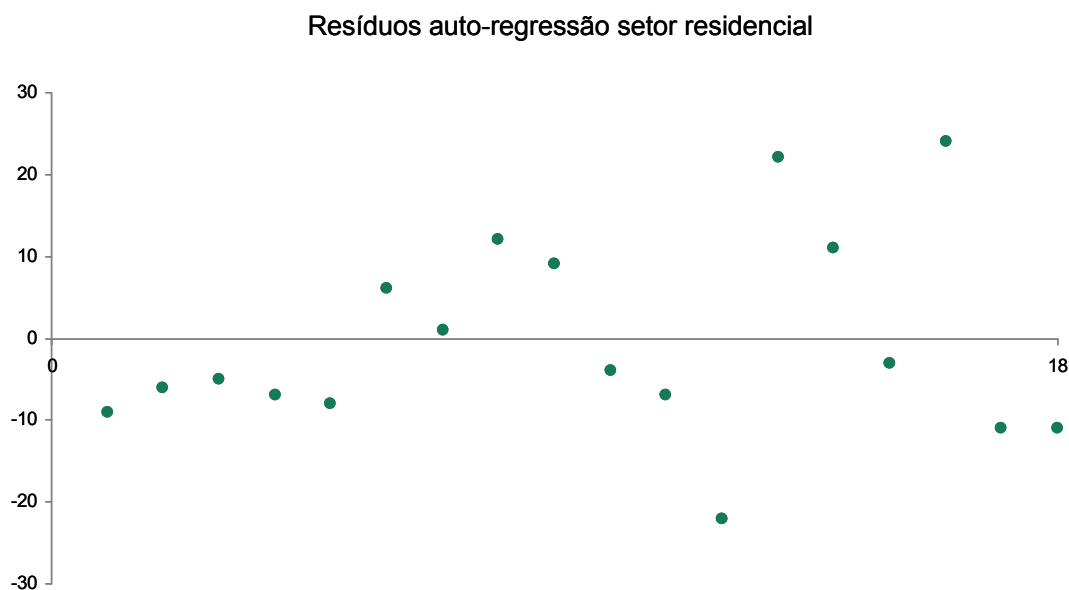


Gráfico 52: Resíduos da auto-regressão Residencial

Analisando-se o gráfico acima, percebe-se que o mesmo viola, de certa forma, a hipótese de homocedasticidade (apesar de tal violação não ser muito acintosa). Além disso, nota-se que a série oscila entre positivo e negativo a cada 3 a 5 pontos, o que viola a hipótese de normalidade dos resíduos.

Realizando-se uma regressão *Stepwise-with-a-backward-look* com as variáveis explanatórias e os parâmetros anteriormente apresentados, observa-se que nenhuma nova variável é incluída no modelo.

Desta forma, a previsão obtida pelo mesmo é a seguinte:

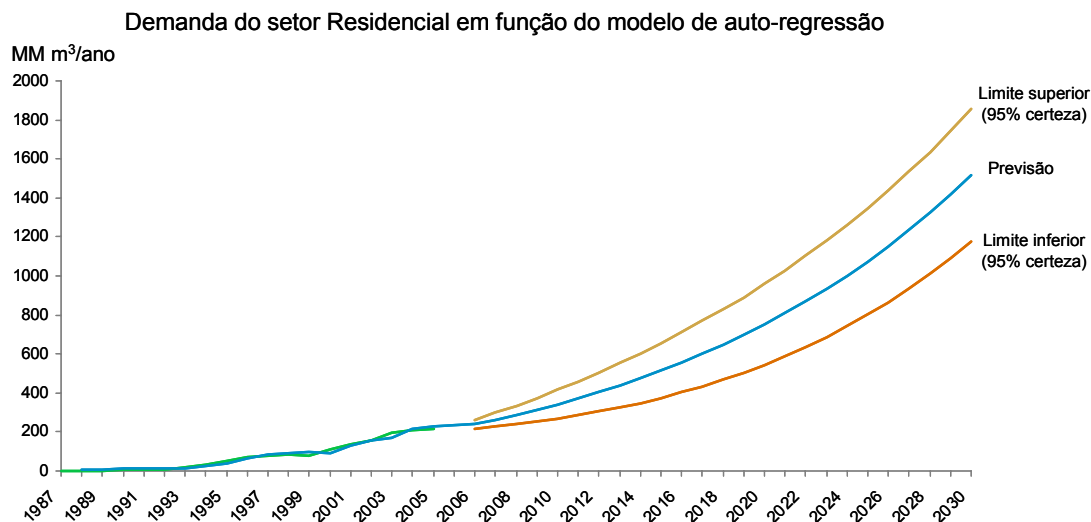


Gráfico 53: Previsão da demanda Residencial em função de auto-regressão

Neste caso prevê-se um crescimento exponencial para a demanda de gás natural do Setor Residencial.

Porém, esse modelo apesar de se mostrar apropriado para o curto prazo não se mostra necessariamente muito confiável para o longo prazo, já que a relação temporal que rege a série pode ser alterada no futuro.

Vale destacar que os intervalos de confiança da figura acima foram estimados devido à dificuldade matemática envolvida no cálculo de intervalos de predição para auto-regressões. Calculou-se o intervalo de confiança da previsão um período à frente e depois se estimou os intervalos de confiança das outras previsões multiplicando-se a primeira previsão pela mesma taxa de crescimento verificada na previsão por Holt (a proporção com o método de Holt foi adotada pelo fato de ambos os modelos serem baseados em séries temporais).

7.4.4 Residencial – Comparação dos modelos

Comparando-se os três métodos de três previsões realizadas para a demanda Residencial, obtém-se o seguinte resultado:

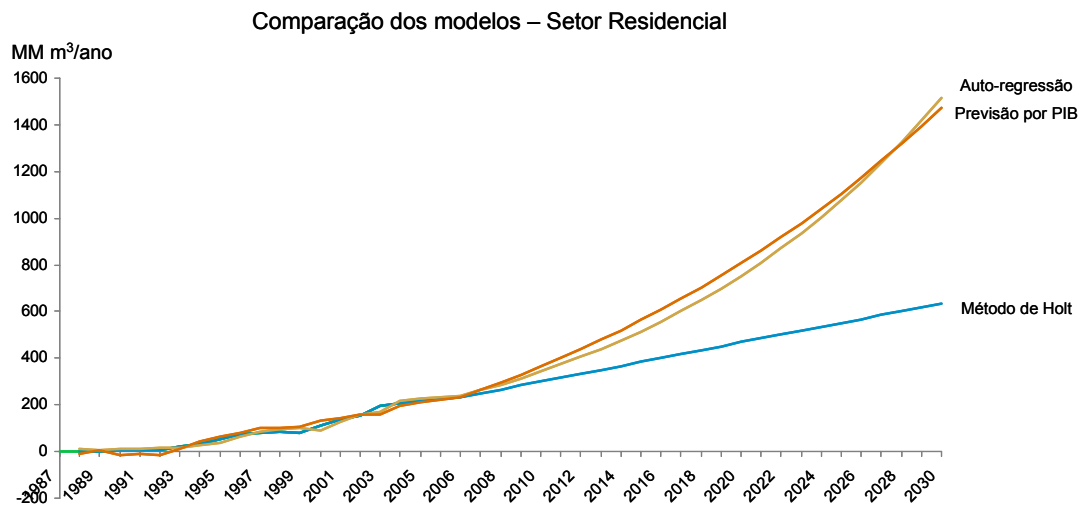


Gráfico 54: Comparação dos modelos – Residencial

Analisando-se o gráfico acima se observa inicialmente que os resultados obtidos pelos métodos de auto-regressão e regressão explanatória, dois métodos completamente distintos, ficaram muito próximos um do outro.

Além disso, pode-se perceber que até cerca de 2009 e 2010 as três previsões realizadas apresentam resultados bastante similares entre si. Após essa época as previsões realizadas pelo método de Holt se distanciam das demais, com uma distância que aumenta conforme o horizonte de previsão aumenta.

Montando-se um quadro comparativo dos modelos de previsão utilizados chega-se à seguinte tabela:

Tabela 11: Quadro comparativo de modelos – Setor Residencial

Setor Residencial	Regressão Explanatória	Auto-regressão	Holt
Número de dados históricos incluídos no modelo	19	18	19
Relação Intervalo de Confiança/Previsão em 2030	10%	22%	51%
R2 ou R2-ajustado encontrado	94,9%	97,3%	n/a
MSE	267,8	136,5	149,2
MAE	13,7	9,8	8,9
MAPE	110%	48%	38%
Problemas com resíduos de 1a ordem	*Resíduos auto-correlacionados	*Não-homocedasticidade *Resíduos auto-correlacionados	n/a
Problemas com resíduos de 2a ordem	*Resíduos auto-correlacionados	n/a	n/a
Método conceitualmente mais adequado para previsões de...	Curto/Médio prazo	Curto prazo	Curto prazo
Alinhamento de resultados com outros métodos?	Auto-regressão	Regressão explanatória	Nenhum

Percebe-se através da tabela acima que o modelo de regressão explanatória possui um número de dados aquém do número desejado (30), porém semelhante ao de seus pares. Além disso, tal modelo apresenta problemas com seus resíduos e erros mais elevados que os demais modelos. Por outro lado apresenta uma boa relação entre seu intervalo de confiança e previsão em 2030, é conceitualmente o mais adequado para a realização de previsões de médio prazo e apresenta resultados com grande alinhamento com os obtidos pelo modelo de auto-regressão, o que, de certa forma, os valida.

Já o modelo de auto-regressão também apresenta problemas com seus resíduos, além de uma relação entre intervalo de confiança e previsão em 2030 não mais do que razoável (comparando com o modelo de regressão explanatória). Conforme mencionado anteriormente, o alinhamento de seus resultados com o modelo de regressão explanatória é uma grande força do mesmo, já que o fato de modelos de naturezas tão distintas apresentarem resultados tão semelhantes é algo notável.

O modelo obtido através de Holt é o que apresenta os melhores resultados de erros, indicando sua grande aderência ao perfil dos dados históricos. Por outro lado, tal modelo apresenta uma relação intervalo de confiança e previsão em 2030 extremamente elevada e o perfil de seus resultados ficou dissonante em relação aos resultados de outros modelos.

Desta forma, recomenda-se a utilização dos resultados dos modelos de regressão explanatória ou auto-regressão como previsão da demanda do setor residencial.

7.5 Setor Comercial/Público

A seguir serão desenvolvidos diferentes modelos de previsão explanatórios e baseados em séries temporais para a demanda do setor comercial/público.

7.5.1 Setor Comercial/Público – Método de Holt

Analisando o gráfico com a série temporal da demanda do setor comercial/público, observamos que o mesmo possui uma tendência de crescimento e pouca ou nenhuma sazonalidade (apesar de apresentar um “soluço” entre 1997 e 2001).

Desta forma, de acordo com a Revisão Bibliográfica deste trabalho, pode-se concluir que um modelo que se mostra adequado para a previsão destas séries é o modelo de Holt.

Utilizando-o para o consumo do setor comercial/público obtém-se o seguinte gráfico:

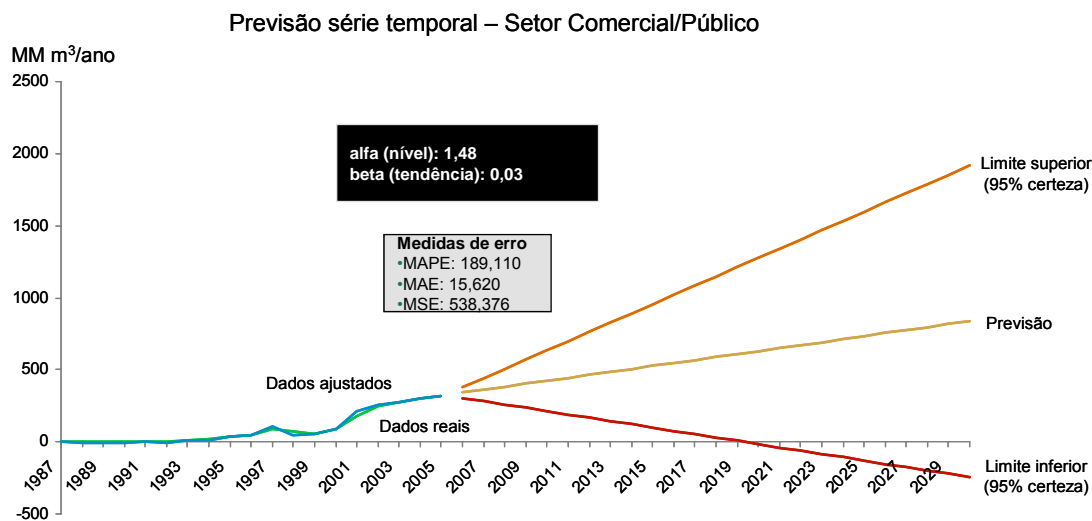


Gráfico 55: Modelo de Holt para demanda do Setor Comercial/Público

Nele pode-se observar uma previsão com tendência futura linear, de acordo com o funcionamento do modelo utilizado. Porém, vale ressaltar que, por considerar que o comportamento futuro de uma série temporal refletirá o seu passado, o método de Holt costuma obter resultados melhores no curto prazo do que no longo prazo.

Além disso, notamos os intervalos de previsão extremamente abertos, sendo que a previsão para 2030 com 95% de certeza fica entre zero (descarta-se aqui os resultados negativos obtidos devido à inconsistência de haver uma demanda negativa de gás natural) e aproximadamente 2000 milhões de metros cúbicos por ano. Isso se deve ao “solavanco” pelo qual a série passou entre 1997 e 2001. Ou seja, olhando-se para o passado da série e prevendo-se seu futuro com base nisso, tal “solavanco” aumenta muito a incerteza sobre o futuro da série histórica.

Outro problema que pode-se encontrar com a série histórica que enfraquece os resultados obtidos pelo método de Holt é o baixo número de pontos existentes (19). Além disso, os resultados extremamente baixos e relativamente constantes dos mesmos de 1987 até 1992 tornam estes pontos praticamente desprezíveis para o modelo porque tal série de dados vai contra a hipótese do modelo de Holt de crescimento linear da série.

7.5.2 Setor Comercial/Público – Regressão Explanatória

Buscando realizar uma previsão por regressão para o setor Comercial/Público, pode-se pensar nas seguintes variáveis que teriam algum relacionamento lógico com a mesma:

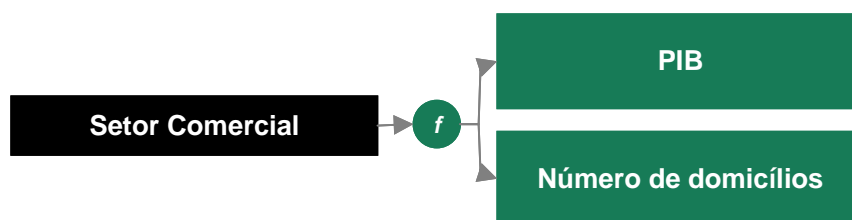


Figura 18: Possíveis variáveis explanatórias para a demanda do setor comercial/público

Ou seja, considera-se que, de alguma forma, a demanda de gás do setor comercial/público possa estar relacionada com o PIB e/ou o número de domicílios brasileiros.

Note-se que o relacionamento entre PIB e demanda do setor comercial/público é razoavelmente intuitiva (quanto maior a atividade econômica, maior deve ser a atividade comercial e maior deve ser, de forma geral, a demanda do setor comercial). Já a variável número de domicílios está sendo utilizada aqui como um “coringa”, já que é um indicador

macroeconômico brasileiro de grande relevância que pode mostrar algum relacionamento estatístico com a variável em estudo.

Rodando no Minitab® uma regressão *Stepwise forward-with-a-backward look* com essas três variáveis e os parâmetros de “*Alpha-to-enter*” e “*Alpha-to-remove*” de 15%, obtém-se a seguinte equação (a qual possui um R^2 de 82,82% e *t-stat* maiores do que 2):

Equação 45: Demanda Comercial/Público em função de modelo de regressão explanatória I

$$y = 712,7 + 5,08 \times \text{PIB}$$

Analizando-se os resíduos que essa equação gera nos dados históricos da demanda Comercial/Público tem-se a seguinte distribuição:

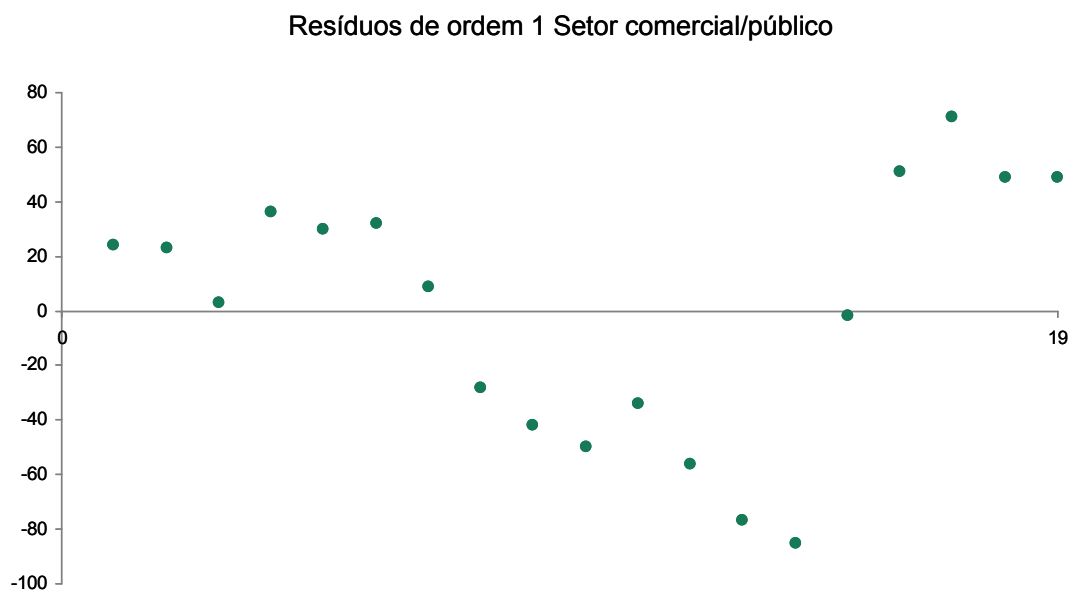


Gráfico 56: Resíduos Comercial/Público vs PIB

Nota-se no gráfico acima que os resíduos são autocorrelacionados e que os mesmos não são normalmente distribuídos.

Outra análise que pode ser realizada é plotar o gráfico do ACF dos erros, para verificar se há alguma influência temporal relevante ao problema que não tenha sido incorporada ao modelo. O gráfico abaixo ilustra isso:

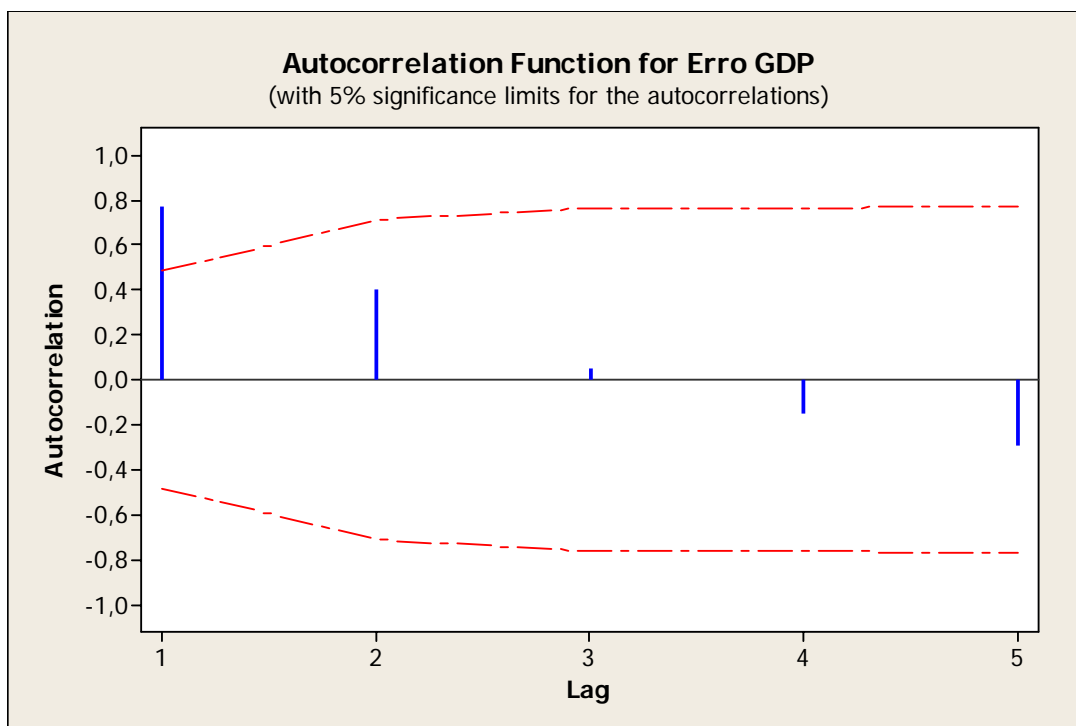


Gráfico 57: ACF comercial/público vs PIB

Tal gráfico indica que uma variável temporal deveria ser incluída na previsão de demanda do setor comercial/público, o que nos leva a buscar refinar o modelo de regressão anteriormente encontrado.

Utilizando-se um Stepwise forward-with-a-backward look dos resíduos com os lags da série encontramos a seguinte equação:

Equação 46: Demanda Comercial/Público em função de modelo de regressão explanatória II

$$\text{Resíduo} = 22,49 + 0,23 \times \text{Lag1}$$

Tal regressão apresenta um R^2 de 18,88%, ou seja, ela explica cerca de 20% do comportamento dos resíduos.

Desta forma, ficamos com a seguinte equação para o modelo de regressão da demanda comercial/pública:

Equação 47: Demanda Comercial/Público em função de modelo de regressão explanatória III

$$y = 73519 + 5,08 \times \text{PIB} + 0,23 \times \text{Lag1}$$

Analisando-se os resíduos de segunda ordem gerados por este modelo encontra-se a seguinte distribuição:

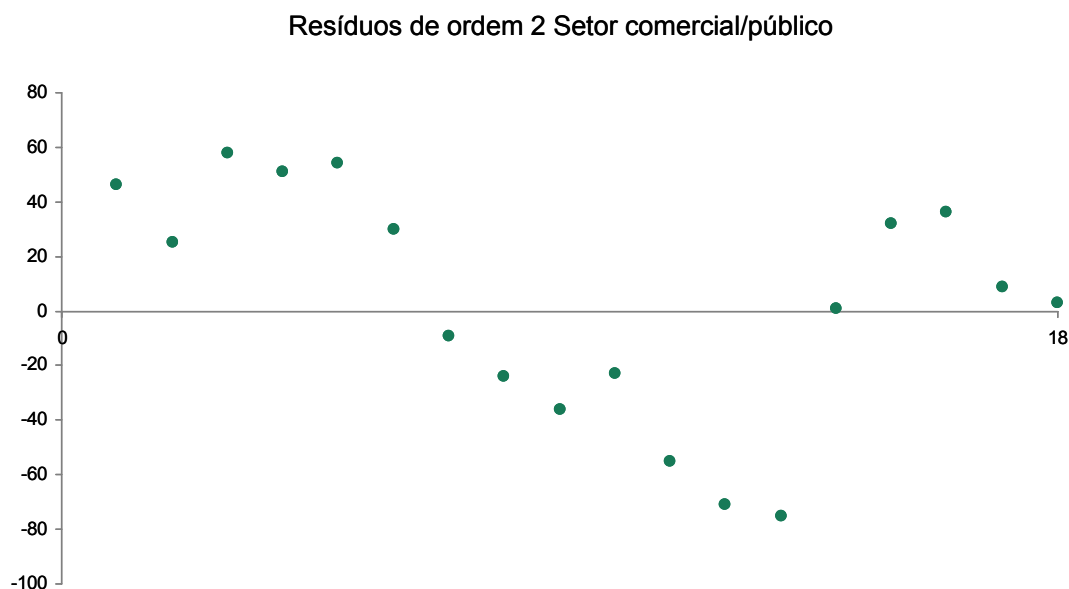


Gráfico 58: Resíduos de segunda ordem do setor Comercial/Público

Nota-se no gráfico acima que os resíduos são autocorrelacionados e que os mesmos não são normalmente distribuídos, assim como os resíduos de primeira ordem do modelo. Porém, são considerados estatisticamente melhores que os resíduos de primeira ordem, já que o Stepwise forward-with-a-backward look adicionou ao modelo original mais uma variável.

Desta forma, a previsão obtida pelo mesmo é a seguinte:

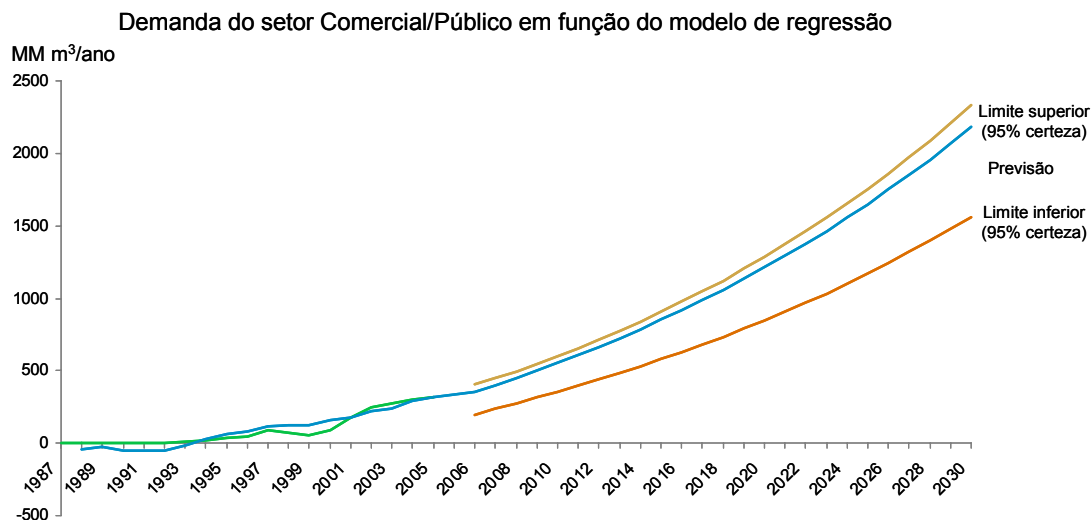


Gráfico 59: Previsão da demanda comercial/público em função do modelo de regressão

Como se pode observar, através deste modelo projeta-se um crescimento com perfil exponencial para a demanda Comercial/Pública no futuro.

Isto se deve ao fato da mesma estar ancorada no tamanho do PIB nacional, cuja previsão prevê um crescimento exponencial para o mesmo. Desta forma, como PIB e demanda Comercial/Pública estão linearmente relacionados pelo modelo obtido, a demanda residencial também apresenta este mesmo perfil de crescimento.

Um comentário que deve ser feito é a respeito da não-equidistância das curvas de intervalo de confiança em relação à curva de previsão. Isso se deve ao fato das curvas de intervalo de confiança terem sido calculadas com base na regressão de primeira ordem realizada (sem a incorporação da variável *Lag1* ao modelo, a qual gerou a regressão de segunda ordem), a título de simplificação dos cálculos. Como a variável *Lag1* apresenta pouca influência no modelo final, isso causa um impacto pequeno nas curvas de intervalo de confiança que, apesar de não estarem matematicamente precisas no gráfico acima, funcionam de forma muito prática para representar a incerteza associada ao modelo utilizado.

Vale ressaltar que esta previsão considera que a relação descrita pela equação do modelo se manterá no futuro, ou seja, que a demanda Residencial continuará seguindo a mesma proporcionalidade com o PIB brasileiro, o que pode não ser verdade no futuro. Além disso, ela baseia suas previsões em previsões de PIB no futuro, as quais, por si só, possuem erros e incertezas associadas.

7.5.3 Setor Comercial/Público – Auto-regressão

Analisando-se agora a demanda do setor Comercial/Público de um ponto de vista temporal, percebe-se inicialmente que a mesma é imensamente influenciada pelo seu comportamento recente, conforme demonstrado pelo gráfico de ACF:

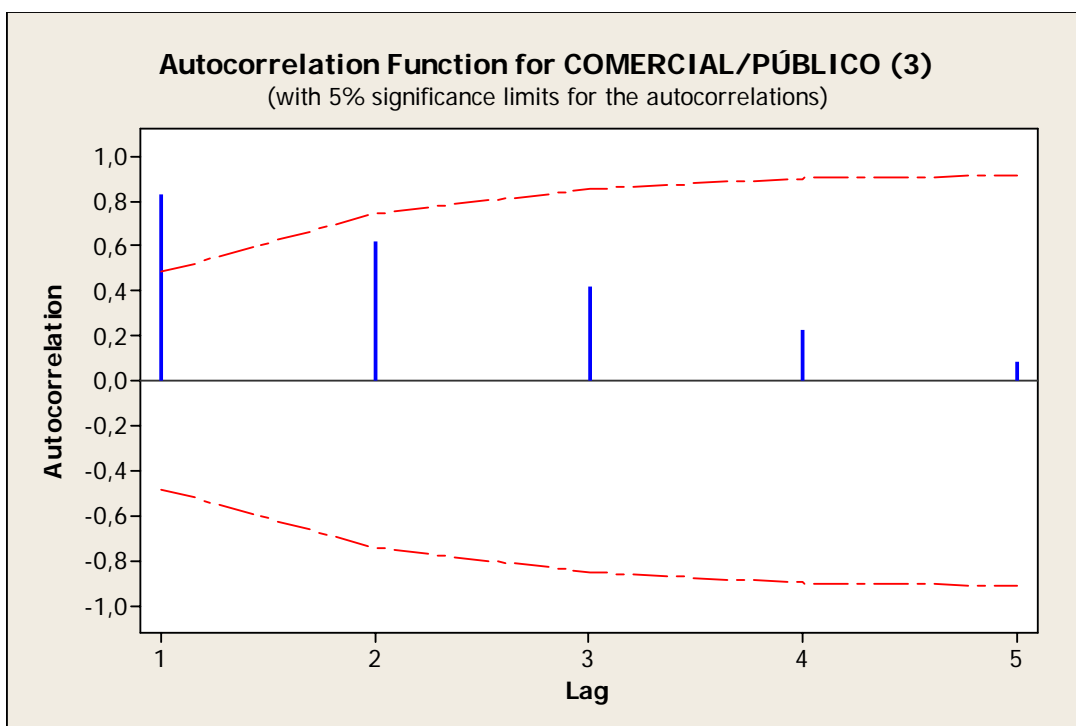


Gráfico 60: ACF da demanda do Setor Comercial/Público

Pelo gráfico acima, pode-se notar que a autocorrelação no primeiro lag da série extrapolou os limites de controle.

Desta forma, buscará se explicar a série temporal de demanda de gás do setor comercial/público através de uma equação de auto-regressão. Rodando-se um modelo de regressão *Stepwise-with-a-backward-look* com parâmetros de “Alpha-to-enter” e “Alpha-to-remove” de 0,15, sendo as possíveis variáveis explanatórias a própria série atrasada (*lagged*) de diversos períodos, obtém-se o seguinte resultado de resíduos:

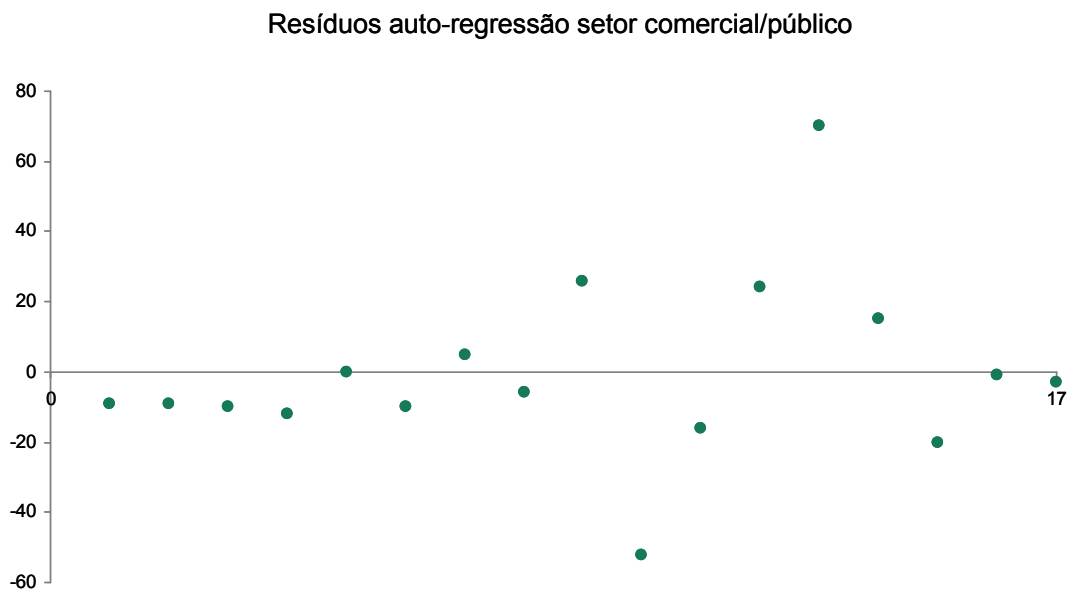


Gráfico 61: Resíduos do Setor Comercial/Público

Como se pode perceber pelo gráfico acima, os resíduos da regressão apresentam certo problema de homocedasticidade (variância aumenta conforme o x aumenta), além de seqüências muito grandes de resíduos abaixo do zero, principalmente no início da série.

A equação encontrada para o modelo é dada por:

Equação 48: Demanda Comercial/Público em função de modelo de auto-regressão

$$y = 10,16 + 1,463 \times \text{Lag1} - 0,45 \times \text{Lag2}$$

Ou seja, o comportamento futuro da série depende do comportamento passado da série de dados – últimos dois resultados (R^2 de 94,41%). O *t-stat* encontrado dos coeficientes é maior do que 2, mas o *t-stat* da constante foi de 1,17. Por ser uma medida do nível inicial da série e influenciar pouco nos resultados finais, releva-se este problema.

Tal equação fornece a seguinte previsão de demanda de gás para o Setor Comercial/Público:

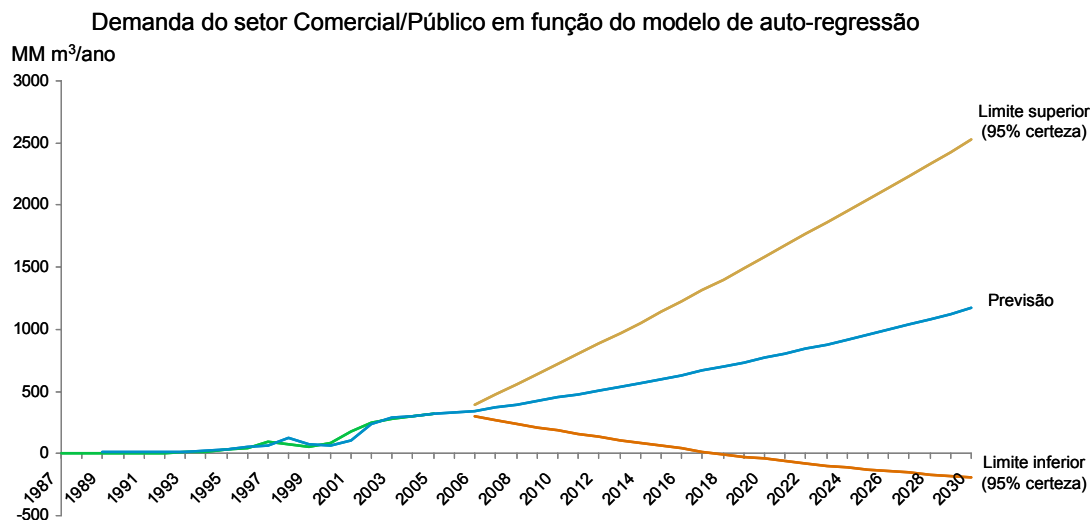


Gráfico 62: Demanda do setor Comercial/Público em função do modelo de auto-regressão

Neste caso prevê-se um crescimento exponencial para a demanda de gás natural do Setor Comercial/Público.

Apesar deste modelo se mostrar apropriado para o curto prazo, o mesmo não se mostra necessariamente muito confiável para o longo prazo, já que a relação temporal que rege a série pode ser alterada no futuro.

Vale destacar que os intervalos de confiança da figura acima foram estimados devido à dificuldade matemática envolvida no cálculo de intervalos de predição para auto-regressões. Calculou-se o intervalo de confiança da previsão um período à frente e depois se estimou os intervalos de confiança das outras previsões multiplicando-se a primeira previsão pela mesma taxa de crescimento verificada na previsão por Holt (a proporção com o método de Holt foi adotada pelo fato de ambos os modelos serem baseados em séries temporais).

7.5.4 Setor Comercial/Público – Comparação dos modelos

Comparando-se agora os resultados dos três modelos encontrados, chega-se ao seguinte gráfico:

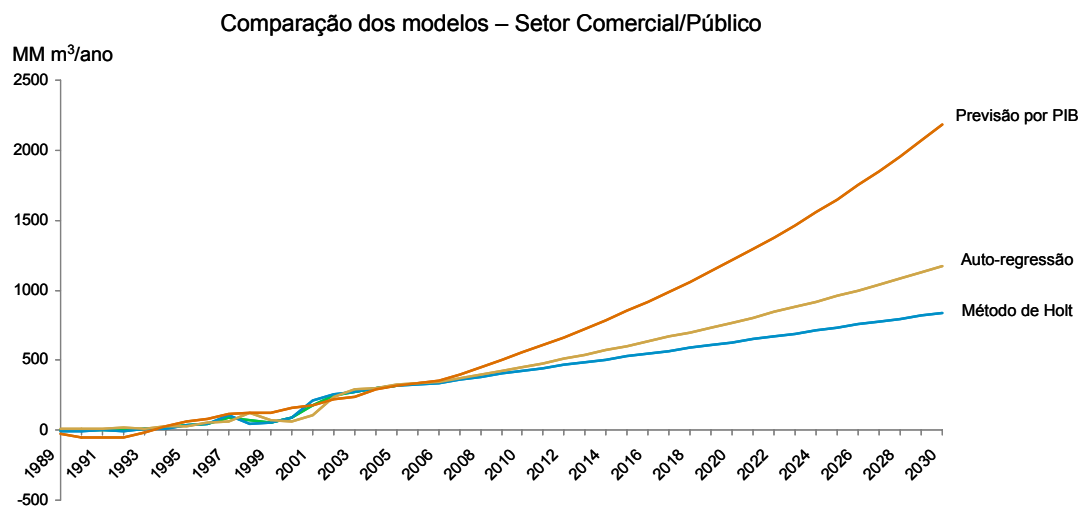


Gráfico 63: Comparação dos modelos – Comercial/Público

O primeiro ponto que se nota no gráfico acima é a considerável diferença entre os resultados dos três modelos encontrados (fato este que não ocorreu na previsão de nenhum dos outros setores). Enquanto a previsão do método de Holt (a qual possui intervalos de confiança bastante amplos, conforme visto anteriormente) segue um crescimento linear, as previsões dos modelos de regressão e auto-regressão seguem crescimentos exponenciais. Porém, ainda entre estas duas últimas, o crescimento exponencial da regressão é muito maior do que o crescimento da auto-regressão.

Além disso, todos os modelos possuem algumas fraquezas evidentes. O modelo de Holt possui intervalos de confiança muito abertos e os modelos baseados em regressão e auto-regressão desrespeitam muitas das hipóteses necessárias para que os mesmos sejam confiáveis.

Tais problemas/diferenças podem ser em grande parte atribuídos aos dados históricos que deram origem aos modelos. Além de escassos, o perfil dos mesmos possui um formato singular, com um “soluço” entre 1997 e 2001.

Montando-se um quadro comparativo dos modelos de previsão utilizados chega-se a:

Tabela 12: Quadro comparativo de modelos – Setor Comercial/Público

Setor Comercial/Público	Regressão Explanatória	Auto-regressão	Holt
Número de dados históricos incluídos no modelo	19	17	19
Relação Intervalo de Confiança/Previsão em 2030	20%	116%	129%
R2 ou R2-ajustado encontrado	82,8%	94,4%	n/a
MSE	2.065,0	603,0	538,4
MAE	39,0	17,0	15,6
MAPE	362%	123%	189%
Problemas com resíduos de 1a ordem	*Não-homocedasticidade *Resíduos auto-correlacionados	*Não-homocedasticidade *Sequência abaixo do zero no início da série	n/a
Problemas com resíduos de 2a ordem	*Não-homocedasticidade *Resíduos auto-correlacionados	n/a	n/a
Método conceitualmente mais adequado para previsões de...	Curto/Médio prazo	Curto prazo	Curto prazo
Alinhamento de resultados com outros métodos?	Nenhum	Relativamente próximo à Holt	Relativamente próximo à auto-regressão

Analisando-se a tabela acima, percebe-se que o modelo de regressão explanatória apresenta uma relação intervalo de confiança/previsão bem melhor que os outros modelos. Por outro lado, seus erros (MSE, MAPE e MAE) são bastante mais elevados, o que indica uma baixa aderência do modelo aos dados históricos (o que é reforçado pelo resultado de R^2 encontrado). Além disso, também se percebe que os resíduos encontrados apresentam problemas, não respeitando completamente as hipóteses do modelo de regressão linear.

O modelo de auto-regressão encontrado é prejudicado pela perda de dois dados históricos na confecção do mesmo (para gerar a série de *lag1* e *lag2*). Outros pontos fracos do mesmo são os problemas com seus resíduos, além de e sua alta relação entre intervalo de confiança e previsão em 2030. Um ponto forte que se observa neste modelo é o relativo alinhamento de seus resultados com os obtidos através do método de Holt o que serve, em parte, para validá-los.

Já o método de Holt encontrado prima pelo fato de ter sido gerado a partir de uma boa quantidade de dados em relação a seus pares (mas nem tantos dados quanto seria desejável) e por possuir baixas medidas de erros. Além disso, possui certo alinhamento de resultados com

os obtidos através do método de auto-regressão. Por outro lado, apresenta elevada relação entre intervalo de confiança e previsão em 2030.

De uma forma geral, pode-se considerar que, dentre todos os setores abordados neste trabalho, este é o que possui resultados de mais difícil interpretação e de menor confiança, problemas que são em grande parte frutos da série histórica com poucos dados e comportamento não muito alinhado com o dos modelos empregados (possuindo um solavanco entre 1997 e 2001). Com base em todas estas dificuldades, recomenda-se a utilização da média dos resultados dos três modelos como previsão para a demanda comercial/pública de gás natural.

Lembra-se novamente que qualquer modelo quantitativo de previsão possui algum tipo de Hipótese da Continuidade por trás do mesmo. Desta forma, qualquer previsão de longo prazo (por exemplo, até 2030) fica comprometida. Nestes casos, o modelo quantitativo pode ser um valioso *input* na elaboração de uma predição por algum especialista do setor.

7.6 Total

Utilizando-se o método de Holt para prever a demanda total futura de gás natural chega-se ao seguinte resultado:

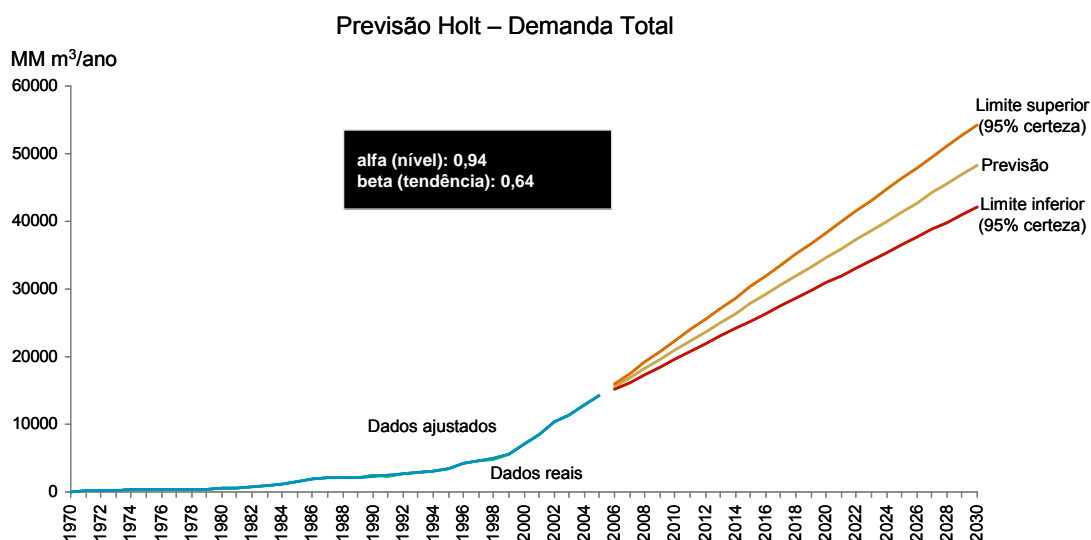


Gráfico 64: Previsão de demanda total – método de Holt

Por outro caminho, somando-se os resultados dos modelos das parcelas da demanda, também é possível se chegar a uma previsão da demanda total de gás. Com base nos modelos que apresentaram melhores resultados para cada parcela, pode-se encontrar a demanda total através de:

Equação 49: Previsão da demanda total

$$\text{Total}_t = \text{Transportes}_{\text{regr. expl.;t}} + \text{Energético}_{\text{auto regressão;t}} + \text{Industrial}_{\text{regr. expl.;t}} + \\ + \text{Residencial}_{\text{reg. expl;t}} + \text{Comercial / Público}_{\text{média(regr. expl.;auto regr.;Holt);t}}$$

O que gera os seguintes resultados:

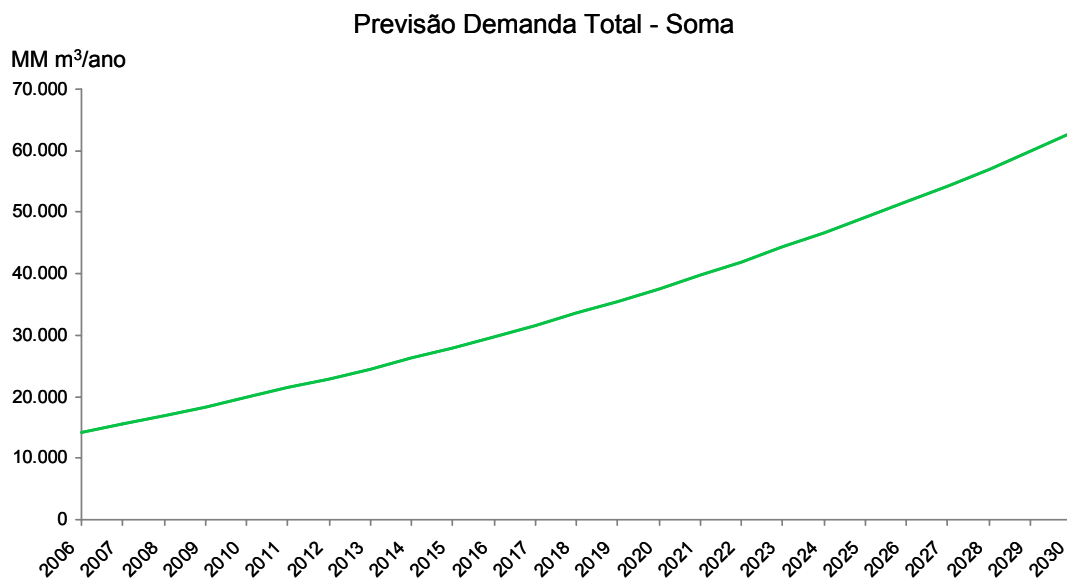


Gráfico 65: Previsão de demanda total – Soma

Comparando-se os dois resultados acima obtidos nota-se que, até cerca de 2016 ambos modelos apresentam resultados bastante próximos entre si. Após isso o modelo baseado em Holt segue sua trajetória de crescimento linear e é ultrapassado pelo modelo baseado nas somas das previsões de cada setor da demanda.

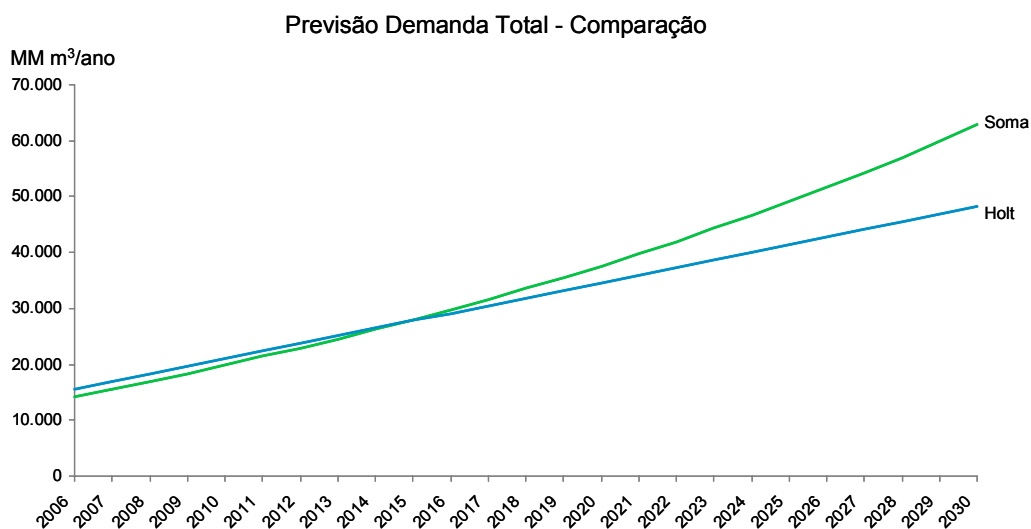


Gráfico 66: Previsão de demanda total – Comparação

8. Conclusões

O objetivo deste trabalho foi o de desenvolver e discutir modelos de previsão da demanda brasileira de gás natural.

Para tanto, estudou-se separadamente cada parcela da demanda de gás natural brasileira: Setor de Transportes, Setor Energético, Setor Industrial, Setor Residencial e Setor Comercial/Público. Para cada uma delas foi estudada sua série temporal e possíveis regressões com variáveis explanatórias.

Como resultado, para cada uma das parcelas da demanda desenvolveram-se 3 modelos distintos de previsão: através do método de Holt, através de regressões explanatórias e através de auto-regressões, cada um com suas forças e fraquezas, algumas específicas ao setor sendo estudado, outras inerentes ao funcionamento do modelo em si.

Feito isso, previu-se por meio de dois métodos distintos a demanda total de gás natural brasileira, comparando-se os resultados obtidos por meio destes métodos.

Um fator que comprometeu a qualidade dos resultados obtidos, principalmente nos casos dos setores Residencial e Comercial/Público foi a baixa disponibilidade de dados existentes em relação à demanda do setor. Nestes casos tal problema deve-se ao fato de se tratar de um mercado ainda novo, com pouca disponibilidade de dados históricos.

Em alguns casos de regressões explanatórias os resultados foram comprometidos pela ausência de séries históricas mais longas das variáveis explanatórias, como por exemplo, PIB. Caso tais séries sejam estendidas, é possível se aprimorar os resultados obtidos neste trabalho.

Deve-se destacar também que, devido à grande importância do gás natural para a economia brasileira, o mesmo pode ser alvo de políticas governamentais que influenciem de forma agressiva sua maior ou menor utilização (assim como ocorreu quando da construção do gasoduto Brasil-Bolívia). Tais políticas governamentais podem funcionar como grandes alavancas modificadoras da demanda, comprometendo de forma direta os resultados obtidos neste trabalho.

Outro ponto que merece destaque é o *trade-off* entre quantidade de previsões que foram realizadas (foram 5 setores, ao todo) e profundidade dos estudos deste trabalho. Caso se objetivasse prever a demanda de apenas um setor poderia ter-se despendido muito mais tempo e esforço no mesmo, gerando possivelmente um maior número de modelos de previsões possíveis e uma discussão mais profunda em torno dos mesmos.

De uma maneira geral obtiveram-se resultados bastante consistentes para previsões no curto e médio prazo. A consistência de resultados entre métodos com lógicas bastante diferentes é um ponto que merece destaque.

Porém, vale destacar que todos os modelos propostos apresentam algum tipo de hipótese da continuidade por trás de si, as quais podem se tornar inválidas no futuro, prejudicando a validade dos modelos desenvolvidos.

Desta forma, para realizações de previsões de longo prazo recomenda-se a utilização dos modelos desenvolvidos neste trabalho como base para a realização de predições por especialistas. Desta forma, os mesmos podem, à frente de seu conhecimento do setor, analisar criticamente os resultados e até mesmo o funcionamento dos modelos propostos, além de ponderar os impactos de grandes mudanças, como políticas governamentais, nos resultados obtidos.

9. Referências Bibliográficas

- BOX, G.E.P., G.M. JENKINS, and G.C. REINSELL. Time series analysis: Forecasting and control, 1ª ed., Holden-Day, São Francisco, 1970.
- DAWES, R.M., D.FAUST, and P.E. Meehl. Clinical versus actuarial judgement, Science, **243** 1668-1674, 1987
- HANKE, J.E., REITSCH, A.G. Business Forecasting, Prentice Hall, Upper Saddle River, New Jersey, 1998.
- HOGARTH, R. and S. Makridakis. Forecasting and planning: an evaluation, Management Science, **27** 115-138, 1981
- MAKRIDAKIS, S. and S.C. WHEELWRIGHT. Forecasting – Methods and Applications, 3ª ed., Wiley, New York, 1998.
- MORETTIN, P.A. and C. TOLOI. Análise de Séries Temporais, 1ª ed., Edgard Blucher, São Paulo, 2004.
- en.wikipedia.org/wiki/The_Boston_Consulting_Group
- oglobo.globo.com
- www.aneel.gov.br
- www.anp.gov.br
- www.bcg.com
- www.epe.gov.br
- www.estado.com.br
- www.folha.uol.com.br
- www.gasnet.com.br
- www.ibge.gov.br
- www.mme.gov.br
 - Plano Nacional de Expansão 2030
 - Plano Decenal de Expansão do Sistema Elétrico 2015

- Balanço Energético Nacional 2006 (ano base 2005)
- www.ons.gov.br
- www.petrobras.com.br
 - Plano de Negócios 2006-2011
 - Relações com o Investidor

10. Apêndices

Nesta seção são apresentadas as tabelas contendo os dados utilizados ao longo deste trabalho, assim como uma representação gráfica dos mesmos.

Inicialmente, tem-se a tabela de dados do consumo de gás natural por segmento da demanda.

Tabela 13: Dados de consumo de gás natural por segmento

Ano	Setor Energético	Residencial	Comercial/ Público	Transportes (rodoviário)	Industrial
1970	74	0	0	0	3
1971	93	0	0	0	12
1972	100	0	0	0	22
1973	98	0	0	0	23
1974	137	0	0	0	163
1975	149	0	0	0	173
1976	146	0	0	0	183
1977	160	0	0	0	312
1978	156	0	0	0	294
1979	161	0	0	0	311
1980	188	0	0	0	363
1981	197	0	0	0	381
1982	391	0	0	0	413
1983	489	0	0	0	449
1984	628	0	0	0	519
1985	911	0	0	0	680
1986	1.050	0	0	0	871
1987	1.062	1	1	0	1.131
1988	935	0	0	3	1.198
1989	894	2	1	2	1.246
1990	859	5	3	2	1.535
1991	768	6	4	2	1.617
1992	840	6	3	0	1.806
1993	974	20	13	25	1.947
1994	1.025	30	18	46	2.025
1995	989	52	36	49	2.353
1996	1.199	72	49	36	2.860
1997	1.226	81	92	47	3.194
1998	1.471	87	71	132	3.133
1999	1.696	79	57	159	3.517
2000	2.278	114	86	313	4.343
2001	2.419	140	180	572	5.141
2002	2.722	154	250	980	6.343
2003	2.938	196	275	1.328	6.658
2004	3.168	206	299	1.580	7.572
2005	3.500	217	321	1.945	8.209

Fonte: Balanço Energético Nacional 2006

Note-se que na tabela acima as células pintadas de verde claro são dados que não foram incluídos nas modelagens deste trabalho. Isso se deve ao fato de que se considerou que as mesmas representam dados de uma época em que o mercado em questão ainda não existia (ou que possuía porte desprezível).

Sendo assim, a utilização de tais dados apenas contaminaria os resultados dos modelos desenvolvidos ao longo deste trabalho.

Em seguida é apresentada a primeira parte da tabela contendo os dados históricos das variáveis explanatórias utilizadas ao longo deste trabalho.

Tabela 14: Dados históricos das variáveis regressoras I

Ano	População (MM hab)	PIB a preços constantes de 2005 (R\$*10 ¹¹)	Carros convertidos GNV
1970			
1971			
1972			
1973			
1974			
1975			
1976			
1977			
1978			
1979			
1980	119		
1981	121		
1982	124		
1983	127		
1984	130	113	
1985	133	122	
1986	136	131	
1987	139	136	
1988	141	136	
1989	144	140	
1990	147	134	
1991	149	135	
1992	152	135	
1993	154	141	
1994	156	149	
1995	159	156	
1996	161	160	4.800
1997	164	165	9.258
1998	166	165	18.658
1999	169	167	57.693
2000	171	174	144.917
2001	174	176	292.871
2002	176	180	449.435
2003	179	181	643.507
2004	182	189	827.958
2005	184	194	1.052.295

Fonte: IBGE, Gasnet e Mídia

A tabela em seguida apresenta as previsões ano-a-ano de cada uma das variáveis explanatórias apresentadas anteriormente. Note que todas as células pintadas de amarelo-claro representam previsões.

Tabela 15: Previsões das variáveis regressoras I

Ano	População (MM hab)	PIB a preços constantes de 2005 (R\$*10 ¹¹)	Carros convertidos GNV
2006	187	200	1.156.799
2007	189	208	1.325.443
2008	192	216	1.494.086
2009	194	225	1.662.730
2010	197	234	1.831.374
2011	199	244	2.000.018
2012	202	254	2.168.662
2013	204	264	2.337.306
2014	206	275	2.505.950
2015	208	287	2.674.594
2016	211	298	2.843.238
2017	213	310	3.011.882
2018	215	323	3.180.526
2019	217	336	3.349.169
2020	219	350	3.517.813
2021	221	365	3.686.457
2022	223	380	3.855.101
2023	225	395	4.023.745
2024	227	411	4.192.389
2025	229	428	4.361.033
2026	231	446	4.529.677
2027	233	464	4.698.321
2028	234	483	4.866.965
2029	236	503	5.035.609
2030	238	524	5.204.252

Fonte: IBGE, Gasnet e Mídia

A próxima tabela mostra os dados históricos do restante das variáveis explanatórias utilizadas ao longo deste trabalho.

Tabela 16: Dados históricos das variáveis regressoras II

Ano	Cap. inst total (MW)/100	Cap. inst térmica (MW)/100	Preço relativo GN/óleo comb.
1970			
1971			
1972			
1973			
1974	181,33	44,09	
1975	209,68	46,52	
1976	225,84	46,80	
1977	243,39	49,43	1,45
1978	269,72	53,07	1,60
1979	302,19	59,84	1,41
1980	334,72	58,23	1,68
1981	372,69	60,96	1,74
1982	393,46	61,90	1,74
1983	403,66	61,88	1,57
1984	410,96	61,73	1,47
1985	441,07	63,73	1,27
1986	449,53	65,10	1,29
1987	475,61	65,75	1,31
1988	495,75	66,90	1,04
1989	521,25	66,72	0,87
1990	530,50	68,35	1,01
1991	541,41	68,68	0,90
1992	550,49	66,83	0,93
1993	562,22	69,74	0,94
1994	576,29	70,51	1,02
1995	591,20	70,97	0,87
1996	608,01	70,25	0,88
1997	629,72	74,26	0,89
1998	652,09	77,93	0,89
1999	681,81	85,26	0,90
2000	737,12	106,42	0,82
2001	762,55	117,25	0,82
2002	824,58	151,40	0,84
2003	865,05	167,05	0,67
2004	907,33	197,27	0,74
2005	931,58	202,93	0,49

Fonte: Plano Nacional de Expansão 2030 e Balanço Energético Nacional 2006

Em seguida são apresentadas as previsões ano-a-ano das variáveis da tabela acima.

Tabela 17: Previsões das variáveis regressoras II

Ano	Cap. inst total (MW)/100	Cap. inst térmica (MW)/100	Preço relativo GN/óleo comb.
2006	964,42	204,63	1,11
2007	997,26	206,34	1,11
2008	1.030,10	208,05	1,11
2009	1.062,95	209,76	1,11
2010	1.095,79	211,46	1,11
2011	1.128,63	213,17	1,11
2012	1.161,47	214,88	1,11
2013	1.194,32	216,59	1,11
2014	1.227,16	218,29	1,11
2015	1.260,00	220,00	1,11
2016	1.304,00	224,00	1,11
2017	1.348,00	228,00	1,11
2018	1.392,00	232,00	1,11
2019	1.436,00	236,00	1,11
2020	1.480,00	240,00	1,11
2021	1.540,00	248,00	1,11
2022	1.600,00	256,00	1,11
2023	1.660,00	264,00	1,11
2024	1.720,00	272,00	1,11
2025	1.780,00	280,00	1,11
2026	1.850,00	298,00	1,11
2027	1.920,00	316,00	1,11
2028	1.990,00	334,00	1,11
2029	2.060,00	352,00	1,11
2030	2.130,00	370,00	1,11

Fonte: Plano Nacional de Expansão 2030 e Balanço Energético Nacional 2006

Abaixo são apresentadas as representações gráficas dos dados das tabelas de variáveis explanatórias. Vale destacar que as representações gráficas dos dados de consumo de gás natural por segmento da demanda foram anteriormente apresentadas no corpo deste mesmo relatório.

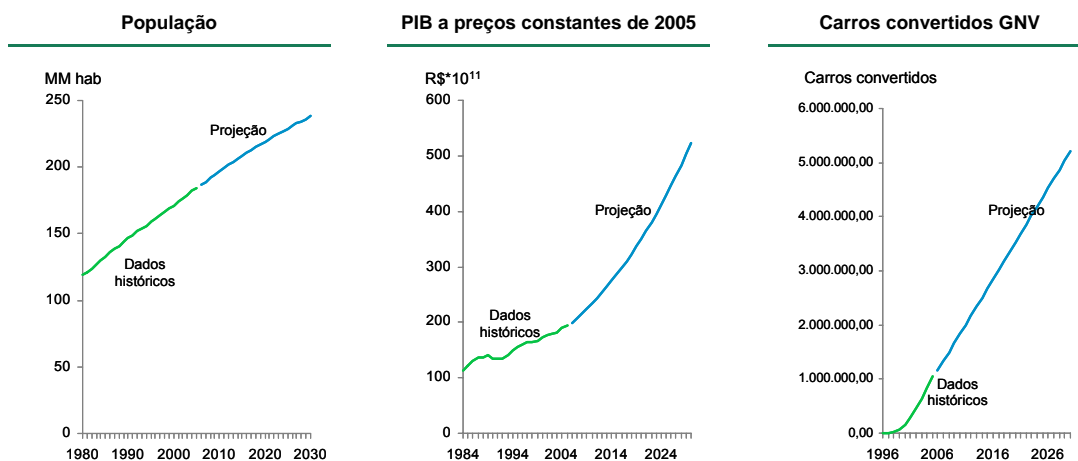


Figura 19: Dados históricos e projeção das variáveis explanatórias I

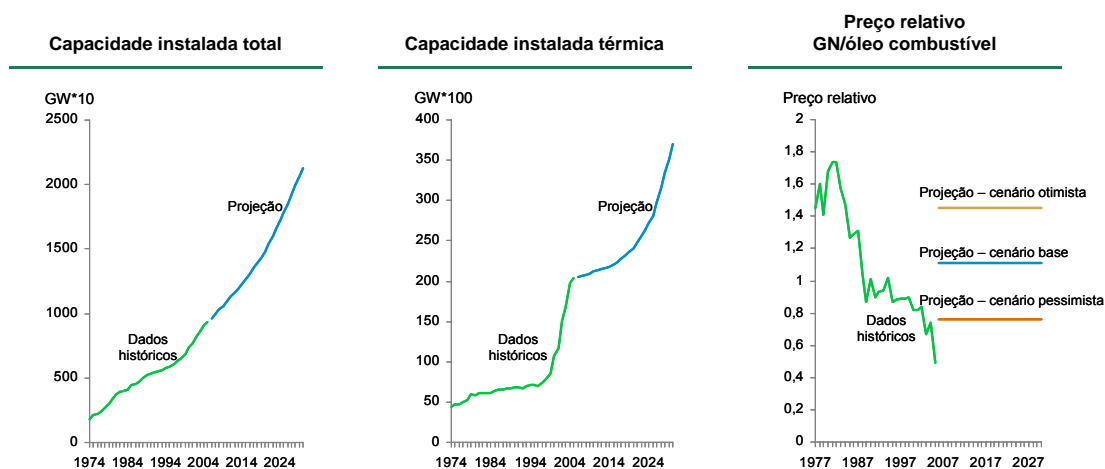


Figura 20: Dados históricos e projeção das variáveis explanatórias II

Analisando-se as duas figuras acima pode-se notar as seguintes características sobre cada variável:

- **População:** tem tanto histórico quanto projeção de crescimento linear. Assim sendo, variáveis linearmente correlacionadas com “População” devem apresentar o mesmo perfil histórico e de crescimento.

- **PIB a preços constantes de 2005:** apresenta um histórico de crescimento exponencial (apesar ter variado consideravelmente ao longo do tempo) e uma previsão de crescimento exponencial. Assim sendo, variáveis que sejam linearmente correlacionadas com PIB devem apresentar uma previsão de crescimento também exponencial. Aqui foi utilizada a série de dados de PIB a preços constantes a fim de retirar dos dados a influência da inflação.
- **Carros convertidos a GNV:** é um fenômeno relativamente recente no cenário brasileiro, tendo sido introduzido por volta de 1996. Desta forma, no início da série de dados verifica-se um crescimento exponencial abrupto dos carros convertidos. Já a previsão para o futuro é de que tal série mantenha um crescimento com taxa considerável, porém de perfil linear. Note-se que no caso desta variável foram encontradas apenas previsões para alguns anos, às quais foram divulgadas na mídia. Para os anos em que não se havia uma previsão utilizou-se uma interpolação linear de dados, já que os pontos que possuíam previsões atreladas delineavam um perfil linear entre si.
- **Capacidade instalada de usinas de geração total:** verifica-se que tanto os dados históricos quanto a previsão para esta série apresentam um perfil linear ou exponencial pouco acentuado. Desta forma, uma variável linearmente correlacionada com esta também apresentará o mesmo tipo de perfil.
- **Capacidade instalada térmica:** percebe-se que esta curva possui um perfil bastante peculiar. Em seu início observa-se um crescimento baixo, seguido por um grande “salto” ocorrido devido a programas de incentivo do governo durante a época de racionamento de energia (o qual ocorreu em 2001, mas que já era anunciado alguns anos antes disso), cujos efeitos se observam até hoje. Atualmente vive-se uma época de “esfriamento” do investimento em usinas térmicas, mas se prevê um novo grande crescimento no futuro devido ao novo risco de racionamento de energia que se prevê no curto prazo (por volta de 2010 ou 2011), além do esgotamento do potencial hídrico brasileiro no curto prazo. Desta forma, variáveis linearmente correlacionadas com esta devem apresentar um perfil futuro de crescimento exponencial acentuado.
- **Preço relativo GN/óleo combustível:** verifica-se que tal série de dados apresenta um perfil repleto de subidas e descidas, mas cuja macro-tendência é de decrescimento ao longo do tempo. Porém, a previsão de autoria própria realizada neste trabalho para o

futuro desta variável é a de que a mesma aumente no futuro, dado os recentes aumentos do gás boliviano noticiados recentemente, assim como o breve início da importação de GNL (gás natural liquefeito) para o abastecimento nacional, o qual possui preço mais elevado que o gás de produção nacional e importado da Bolívia. Como esta variável é de difícil previsão, optou-se por se realizar uma previsão constante para a mesma, porém, para três cenários possíveis: base, otimista e pessimista.

Em seguida são apresentados os resultados dos modelos de previsão desenvolvidos ao longo deste relatório.

Tabela 18: Resultados da previsão do Setor de Transportes

Setor de Transportes (em MM m3/ano)								
Holt			Regressão explanatória			Auto-regressão		
Previsão	IC acima (95%)	IC abaixo (95%)	Previsão	IC acima (95%)	IC abaixo (95%)	Previsão	IC acima (95%)	IC abaixo (95%)
2.328	2.445	2.211	2.202	2.319	2.084	2.352	2.486	2.218
2.699	2.891	2.508	2.514	2.640	2.388	2.813	3.033	2.593
3.070	3.340	2.801	2.826	2.961	2.691	3.337	3.646	3.028
3.442	3.790	3.093	3.138	3.282	2.993	3.936	4.336	3.537
3.813	4.241	3.385	3.450	3.605	3.295	4.623	5.114	4.132
4.184	4.691	3.677	3.762	3.928	3.596	5.412	5.994	4.830
4.555	5.142	3.968	4.074	4.251	3.896	6.318	6.992	5.644
4.926	5.593	4.259	4.386	4.575	4.197	7.360	8.126	6.595
5.297	6.044	4.551	4.698	4.899	4.497	8.559	9.416	7.702
5.669	6.495	4.842	5.010	5.223	4.797	9.938	10.886	8.989
6.040	6.946	5.133	5.322	5.547	5.096	11.524	12.565	10.484
6.411	7.397	5.425	5.634	5.872	5.396	13.350	14.482	12.218
6.782	7.849	5.716	5.946	6.196	5.695	15.451	16.675	14.228
7.153	8.300	6.007	6.258	6.521	5.995	17.870	19.185	16.554
7.525	8.751	6.298	6.570	6.846	6.294	20.652	22.060	19.245
7.896	9.202	6.589	6.882	7.171	6.593	23.855	25.354	22.356
8.267	9.653	6.881	7.194	7.496	6.892	27.541	29.132	25.950
8.638	10.104	7.172	7.506	7.821	7.191	31.783	33.466	30.100
9.009	10.556	7.463	7.818	8.146	7.490	36.665	38.440	34.890
9.380	11.007	7.754	8.130	8.471	7.788	42.284	44.150	40.417
9.752	11.458	8.045	8.442	8.796	8.087	48.750	50.708	46.791
10.123	11.909	8.336	8.754	9.121	8.386	56.192	58.242	54.141
10.494	12.360	8.628	9.066	9.447	8.685	64.757	66.899	62.614
10.865	12.812	8.919	9.378	9.772	8.983	74.614	76.848	72.380
11.236	13.263	9.210	9.690	10.097	9.282	85.958	88.284	83.632

Tabela 19: Resultados da previsão do Setor Energético

Setor Energético (em MM m3/ano)								
Holt			Regressão explanatória			Auto-regressão		
Previsão	IC acima (95%)	IC abaixo (95%)	Previsão	IC acima (95%)	IC abaixo (95%)	Previsão	IC acima (95%)	IC abaixo (95%)
3.774	3.964	3.584	3.942	4.347	3.622	3.956	4.159	3.754
4.058	4.329	3.786	4.098	4.555	3.754	4.140	4.446	3.833
4.341	4.699	3.983	4.260	4.776	3.886	4.331	4.746	3.916
4.624	5.071	4.178	4.427	5.008	4.019	4.531	5.057	4.004
4.908	5.444	4.371	4.600	5.251	4.154	4.739	5.378	4.101
5.191	5.818	4.564	4.779	5.506	4.291	4.957	5.709	4.206
5.475	6.192	4.757	4.964	5.773	4.430	5.185	6.050	4.320
5.758	6.567	4.949	5.157	6.051	4.573	5.423	6.401	4.445
6.041	6.941	5.142	5.355	6.340	4.719	5.671	6.763	4.580
6.325	7.316	5.334	5.562	6.642	4.870	5.931	7.136	4.726
6.608	7.691	5.526	5.808	6.985	5.058	6.205	7.523	4.886
6.892	8.066	5.717	6.062	7.341	5.251	6.491	7.923	5.059
7.175	8.441	5.909	6.324	7.709	5.449	6.790	8.336	5.244
7.459	8.816	6.101	6.595	8.091	5.653	7.102	8.761	5.442
7.742	9.191	6.293	6.874	8.487	5.862	7.427	9.201	5.654
8.025	9.566	6.484	7.220	8.948	6.137	7.773	9.660	5.885
8.309	9.941	6.676	7.575	9.424	6.418	8.133	10.134	6.132
8.592	10.317	6.868	7.940	9.915	6.705	8.509	10.624	6.394
8.876	10.692	7.059	8.315	10.422	6.999	8.902	11.131	6.673
9.159	11.067	7.251	8.701	10.946	7.301	9.312	11.655	6.969
9.442	11.442	7.443	9.241	11.613	7.759	9.752	12.209	7.295
9.726	11.817	7.634	9.793	12.299	8.224	10.212	12.782	7.641
10.009	12.193	7.826	10.356	13.003	8.698	10.690	13.375	8.006
10.293	12.568	8.018	10.933	13.727	9.180	11.190	13.988	8.391
10.576	12.943	8.209	11.522	14.471	9.671	11.711	14.623	8.798

Tabela 20: Resultados da previsão do Setor Industrial

Setor Industrial (em MM m3/ano)								
Holt			Regressão explanatória			Auto-regressão		
Previsão	IC acima (95%)	IC abaixo (95%)	Previsão	IC acima (95%)	IC abaixo (95%)	Previsão	IC acima (95%)	IC abaixo (95%)
8.917	9.258	8.577	7.502	9.152	5.852	9.977	10.373	9.580
9.623	10.131	9.114	8.287	9.986	6.588	11.291	11.883	10.698
10.328	11.013	9.643	9.104	10.860	7.347	12.788	13.587	11.990
11.033	11.900	10.167	9.954	11.778	8.130	14.496	15.505	13.487
11.739	12.787	10.690	10.839	12.739	8.940	16.442	17.664	15.220
12.444	13.676	11.212	11.761	13.746	9.776	18.661	20.096	17.226
13.149	14.565	11.734	12.720	14.799	10.641	21.191	22.840	19.541
13.855	15.455	12.255	13.719	15.902	11.536	24.074	25.939	22.210
14.560	16.345	12.775	14.758	17.054	12.462	27.362	29.441	25.282
15.266	17.235	13.296	15.840	18.259	13.422	31.109	33.404	28.814
15.971	18.126	13.816	16.967	19.518	14.416	35.382	37.892	32.871
16.676	19.016	14.337	18.140	20.832	15.448	40.252	42.978	37.526
17.382	19.906	14.857	19.361	22.204	16.518	45.804	48.746	42.863
18.087	20.797	15.377	20.632	23.635	17.628	52.134	55.291	48.977
18.792	21.688	15.897	21.955	25.129	18.780	59.350	62.723	55.977
19.498	22.578	16.417	23.332	26.686	19.977	67.576	71.165	63.987
20.203	23.469	16.937	24.766	28.310	21.221	76.953	80.758	73.148
20.909	24.360	17.457	26.258	30.004	22.512	87.644	91.665	83.623
21.614	25.250	17.977	27.812	31.768	23.855	99.831	104.068	95.594
22.319	26.141	18.497	29.429	33.608	25.251	113.724	118.177	109.271
23.025	27.032	19.017	31.113	35.524	26.701	129.563	134.232	124.894
23.730	27.923	19.537	32.866	37.521	28.210	147.618	152.504	142.733
24.435	28.814	20.057	34.690	39.602	29.779	168.202	173.303	163.101
25.141	29.705	20.577	36.590	41.769	31.410	191.667	196.985	186.350
25.846	30.595	21.097	38.567	44.026	33.107	218.418	223.951	212.885

Tabela 21: Resultados da previsão do Setor Residencial

Setor Residencial (em MM m3/ano)								
Holt			Regressão explanatória			Auto-regressão		
Previsão	IC acima (95%)	IC abaixo (95%)	Previsão	IC acima (95%)	IC abaixo (95%)	Previsão	IC acima (95%)	IC abaixo (95%)
234	255	212	234	263	189	239	262	216
250	284	217	265	294	217	262	298	227
267	313	221	297	326	246	288	335	240
284	342	225	330	360	276	314	375	253
300	371	230	365	396	308	343	417	269
317	401	234	402	433	340	373	460	286
334	430	238	440	472	373	405	506	305
351	459	242	479	512	408	440	553	326
367	489	246	520	554	444	476	603	349
384	518	250	563	598	481	515	655	375
401	548	254	608	644	520	557	710	403
417	577	258	655	692	561	601	768	434
434	607	262	703	742	603	648	828	468
451	636	266	754	794	646	698	891	504
468	665	270	807	849	691	751	958	544
484	695	274	862	905	738	808	1.028	588
501	724	278	919	964	787	869	1.102	635
518	754	282	979	1.026	838	933	1.180	686
534	783	286	1.041	1.090	891	1.002	1.262	742
551	813	290	1.106	1.157	946	1.075	1.348	802
568	842	294	1.174	1.226	1.004	1.153	1.439	866
585	872	298	1.245	1.298	1.063	1.236	1.536	936
601	901	301	1.318	1.374	1.125	1.324	1.637	1.011
618	931	305	1.395	1.452	1.190	1.418	1.745	1.091
635	960	309	1.475	1.534	1.257	1.518	1.858	1.178

Tabela 22: Resultados da previsão do Setor Comercial/Público

Setor Comercial/Público (em MM m3/ano)								
Holt			Regressão explanatória			Auto-regressão		
Previsão	IC acima (95%)	IC abaixo (95%)	Previsão	IC acima (95%)	IC abaixo (95%)	Previsão	IC acima (95%)	IC abaixo (95%)
340	379	302	352	404	199	345	393	296
361	442	280	399	449	236	370	472	268
382	506	257	449	497	275	396	553	240
402	570	235	500	547	315	424	635	213
423	634	212	553	600	355	452	717	186
444	698	189	608	656	398	480	800	160
465	763	166	665	714	441	509	884	135
485	827	144	725	775	486	539	969	110
506	891	121	787	839	532	570	1.054	86
527	955	98	851	905	581	601	1.140	63
547	1.019	75	918	975	630	633	1.227	40
568	1.083	52	988	1.047	682	666	1.314	18
589	1.148	30	1.060	1.123	735	700	1.403	-3
609	1.212	7	1.135	1.203	791	734	1.492	-23
630	1.276	-16	1.213	1.285	848	769	1.582	-43
651	1.340	-39	1.294	1.371	908	805	1.672	-61
671	1.404	-62	1.378	1.461	970	842	1.764	-79
692	1.468	-84	1.466	1.555	1.034	880	1.856	-96
713	1.533	-107	1.557	1.653	1.101	919	1.949	-112
733	1.597	-130	1.651	1.754	1.171	958	2.044	-128
754	1.661	-153	1.750	1.860	1.243	998	2.139	-142
775	1.725	-176	1.852	1.971	1.318	1.040	2.235	-155
795	1.789	-198	1.958	2.086	1.397	1.082	2.332	-168
816	1.853	-221	2.068	2.206	1.478	1.125	2.429	-179
837	1.918	-244	2.183	2.331	1.562	1.169	2.528	-190

Tabela 23: Resultados da previsão de demanda total

Ano	Total (em MM m3/ano)			Soma
	Holt			
	Previsão	IC acima (95%)	IC abaixo (95%)	Previsão
2006	15.550	16.004	15.096	14.240
2007	16.912	17.579	16.244	15.582
2008	18.273	19.165	17.380	16.966
2009	19.634	20.757	18.511	18.395
2010	20.995	22.350	19.640	19.869
2011	22.356	23.946	20.767	21.392
2012	23.717	25.542	21.893	22.965
2013	25.079	27.138	23.019	24.590
2014	26.440	28.735	24.145	26.269
2015	27.801	30.332	25.270	28.004
2016	29.162	31.929	26.395	29.801
2017	30.523	33.527	27.520	31.660
2018	31.885	35.125	28.644	33.582
2019	33.246	36.722	29.769	35.571
2020	34.607	38.320	30.894	37.629
2021	35.968	39.918	32.018	39.765
2022	37.329	41.516	33.143	41.975
2023	38.690	43.114	34.267	44.265
2024	40.052	44.712	35.391	46.635
2025	41.413	46.310	36.516	49.091
2026	42.774	47.908	37.640	51.648
2027	44.135	49.506	38.764	54.297
2028	45.496	51.104	39.889	57.043
2029	46.858	52.702	41.013	59.889
2030	48.219	54.300	42.137	62.839