

**UNIVERSIDADE FEDERAL DE MINAS GERAIS**  
**Escola de Engenharia**  
**Programa de Pós-Graduação em Engenharia Elétrica**

Vinicius Claudino Ferraz

**Previsão de Séries Temporais de Consumo Diário de Gás Natural no Brasil**

Belo Horizonte

2024

Vinicius Claudino Ferraz

**PREVISÃO DE SÉRIES TEMPORAIS DE CONSUMO DIÁRIO DE GÁS NATURAL  
NO BRASIL**

Dissertação apresentada ao Programa de Pós-Graduação em Engenharia Elétrica da Universidade Federal de Minas Gerais, como requisito parcial à obtenção do título de Mestre em Engenharia Elétrica.

Orientador: Prof. Dr. Walmir Matos Caminhas

Belo Horizonte  
2024

F381p

Ferraz, Vinicius Claudino.

Previsão de séries temporais de consumo diário de gás natural no Brasil [recurso eletrônico] / Vinicius Claudino Ferraz. - 2024.

1 recurso online (104 f. : il., color.) : pdf.

Orientador: Walmir Matos Caminhas.

Dissertação (mestrado) - Universidade Federal de Minas Gerais, Escola de Engenharia.

Apêndices: f. 101-104.

Bibliografia: f. 97-100.

1. Engenharia elétrica - Teses. 2. Gás natural - Teses. 3. Previsão - Teses. 4. Engenharia industrial - Teses. 5. Análise de séries temporais - Teses. 6. Combustíveis - Consumo - Brasil - Teses. I. Caminhas, Walmir Matos. II. Universidade Federal de Minas Gerais. Escola de Engenharia. III. Título.

CDU: 621.3(043)



UNIVERSIDADE FEDERAL DE MINAS GERAIS  
ESCOLA DE ENGENHARIA  
COLEGIADO DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA

"Previsão de Séries Temporais de Consumo Diário de Gás Natural no Brasil"

Vinicius Claudino Ferraz

Dissertação de Mestrado submetida à Banca Examinadora designada pelo Colegiado do Programa de Pós-Graduação em Engenharia Elétrica da Escola de Engenharia da Universidade Federal de Minas Gerais, como requisito para obtenção do grau de Mestre em Engenharia Elétrica.

Aprovada em 23 de agosto de 2024.

Por:

Prof. Dr. Walmir Matos Caminhas  
DELT (UFMG) - Orientador

Prof. Ph.D. Eduardo Mazoni Andrade Marçal Mendes  
DELT (UFMG)

Prof. Dr. Frederico Gadelha Guimarães  
DCC (UFMG)



Documento assinado eletronicamente por **Frederico Gadelha Guimaraes, Professor do Magistério Superior**, em 23/08/2024, às 15:15, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Walmir Matos Caminhas, Professor do Magistério Superior**, em 23/08/2024, às 15:15, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



Documento assinado eletronicamente por **Eduardo Mazoni Andrade Marçal Mendes, Professor do Magistério Superior**, em 23/08/2024, às 15:15, conforme horário oficial de Brasília, com fundamento no art. 5º do [Decreto nº 10.543, de 13 de novembro de 2020](#).



A autenticidade deste documento pode ser conferida no site

[https://sei.ufmg.br/sei/controlador\\_externo.php?](https://sei.ufmg.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0)

[acao=documento\\_conferir&id\\_orgao\\_acesso\\_externo=0](https://sei.ufmg.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0), informando o código verificador 3475111 e o código CRC 0368FAA6.

*A presente dissertação é dedicada aos irmãos J.,  
que assinam as psicografias a mim entregues.*

## **AGRADECIMENTOS**

Agradeço inicialmente aos meus Espíritos Guias, a quem imagino que estou dando muito trabalho.

À minha família, que é basicamente minha mãe Isa Borges Ferreira de Oliveira, presente, e tios de São Francisco/MG, primos, irmã e sobrinhos de Carmo da Mata/MG, às centenas, ainda que ausentes.

A todos(as) os(as) anônimos(as) que de alguma forma colocaram um tijolo imaterial nesta ponte, que no mínimo me deram apoio psíquico e moral.

Ao meu orientador, professor Walmir Matos Caminhas, por todo o apoio, amizade, competência e conhecimentos transmitidos.

Aos professores do Programa de Pós-Graduação em Engenharia Elétrica da Universidade Federal de Minas Gerais, pelo aprendizado durante as disciplinas.

À Fundação Universitária Mendes Pimentel, FUMP, a quem devo agradecimentos desde o primeiro período da graduação.

Ao professor Frederico Gadelha Guimarães (do Departamento de Ciência da Computação da UFMG) e à professora Patrícia de Oliveira e Lucas (do Instituto Federal do Norte de Minas Gerais, IFNMG, Campus Salinas), que tanto contribuíram neste projeto.

E aos nossos patrocinadores, uma vez que este trabalho foi parcialmente financiado pelo Convênio de PD&I entre a GASMIG e a UFMG. Isto é, o presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001.

“Cada ser humano é uma incógnita a ser equacionada por ele próprio.”

Joanna de Ângelis

p/ FRANCO, Divaldo P.  
in Autodescobrimento: Uma Busca Interior

## RESUMO

Este trabalho objetiva definir um modelo de previsão de curto prazo do consumo de Gás Natural (GN) no Brasil. O GN desempenha um papel crucial na matriz energética do Brasil, impactando diversos setores como indústria, geração de energia elétrica e aquecimento residencial. A previsão do consumo de gás natural a curto prazo é essencial para o cumprimento regulamentar e a eficiência econômica, especialmente para as empresas de distribuição. A complexidade do consumo de GN, influenciada por parâmetros econômicos, dados meteorológicos e fatores de calendário, apresenta desafios para previsões precisas. Visa-se preencher uma lacuna na literatura de estudos brasileiros e pesquisas limitadas sobre previsão no Brasil, em que está destacada a necessidade de estudos e modelos localizados, com os quais pretende-se contribuir. A metodologia é uma comparação de vários modelos como Regressão, ARIMA, *Random Forest*, Redes Neurais e Séries Temporais Nebulosas. O modelo proposto combina o Yamakawa *Neo-Fuzzy-Neuron* com o *Unscented Kalman Filter* (UKF) para previsão diária e semanal do consumo de GN. Os resultados experimentais destacam as melhorias do modelo em relação às abordagens existentes. Por fim, conclui-se que sempre o modelo adaptativo de Yamakawa é o mais eficiente, e que pode ser utilizado sempre que houver duas ou mais séries temporais, uma vez que cada uma delas avança independentemente, depois tudo se reúne em um UKF, que filtra as médias e os desvios-padrão.

Palavras-chave: gás natural; previsão; industrial; séries temporais; consumo.

## **ABSTRACT**

This work aims to define a short-term forecast model for Natural Gas (NG) consumption in Brazil. NG plays a crucial role in Brazil's energy matrix, impacting several sectors such as industry, electricity generation and residential heating. Forecasting short-term natural gas consumption is essential for regulatory compliance and economic efficiency, especially for distribution companies. The complexity of NG consumption, influenced by economic parameters, meteorological data and calendar factors, presents challenges for accurate predictions. The aim is to fill a gap in the literature of Brazilian studies and limited research on forecasting in Brazil, which highlights the need for localized studies and models, to which it's intended to contribute. The methodology is a comparison of several models such as Regression, ARIMA, Random Forest, Neural Networks and Fuzzy Time Series. The proposed model combines the Yamakawa Neo-Fuzzy-Neuron with the Unscented Kalman Filter (UKF) for daily and weekly prediction of NG consumption. The experimental results highlight the model's improvements over existing approaches. Finally, it is concluded that Yamakawa's adaptive model is always the most efficient, and that it can be used whenever there are two or more time series, since each of them advances independently, then everything comes together in a UKF, which filters the means and standard deviations.

Keywords: natural gas; forecasting; industrial; time series; consumption.

## LISTA DE FIGURAS

Figura 2.1 – Reservas Mundiais de Gás Natural. . . . .	25
Figura 2.2 – Maiores países consumidores de Gás Natural na América Sul e Central no ano de 2018. . . . .	25
Figura 2.3 – Fluxo do processo do <i>Unscented Kalman Filter</i> . . . . .	32
Figura 2.4 – Estrutura da rede NFN. . . . .	40
Figura 2.5 – Esquema de processamento do NFN. . . . .	41
Figura 2.6 – Funções de pertinência triangulares igualmente espaçadas. . . . .	42
Figura 3.1 – Visão 11h × 24h para os pontos de entrega A e B. . . . .	49
Figura 3.2 – Visão temporal para o ponto de entrega A. . . . .	49
Figura 3.3 – Visão temporal para o ponto de entrega B. . . . .	50
Figura 3.4 – Visão 11h × 24h para os pontos de entrega C e D. . . . .	50
Figura 3.5 – Visão temporal para o ponto de entrega C. . . . .	51
Figura 3.6 – Visão temporal para o ponto de entrega D. . . . .	51
Figura 3.7 – Visão 11h × 24h para o ponto de entrega E. . . . .	51
Figura 3.8 – Visão temporal para o ponto de entrega E. . . . .	52
Figura 3.9 – Fluxo de dias da semana para o ponto de entrega E. . . . .	53
Figura 3.10 – Autocorrelações para os pontos de entrega A, B, C e D. . . . .	54
Figura 3.11 – Autocorrelações para o ponto de entrega E. . . . .	55
Figura 3.12 – <i>Boxplots</i> para os pontos de entrega A e B. . . . .	55
Figura 3.13 – <i>Boxplots</i> para os pontos de entrega C, D e E. . . . .	56
Figura 3.14 – AICs em função do número de variáveis para os cinco pontos de entrega. . . . .	57
Figura 3.15 – Modelos Yamakawa-NFN combinados com modelos <i>Unscented Kalman Filter</i> . . . . .	58
Figura 3.16 – Modelo Yamakawa-NFN, cuja entrada é a vazão de Gás Natural de 0h às 10h59min e cuja saída é a vazão das 11h às 23h59min do mesmo dia. . . . .	59
Figura 3.17 – Funções de Pertinência Pontual e Intervalar. . . . .	59
Figura 4.1 – Previsão recursiva e especialista. . . . .	72
Figura 4.2 – Série temporal referente ao ano 2023 de consumo diário de gás natural das estações de recebimento B, C e D. . . . .	76
Figura 4.3 – Design dos experimentos para avaliação dos modelos de previsão diária de consumo de gás natural de 1, 4 e 7 passos à frente. HPO refere-se a otimização de hiperparâmetros. . . . .	78
Figura 4.4 – Diagrama de diferenças críticas (CD) dos resultados dos modelos NAIVE, ARIMA, WHOFTS, NFN, MLP, RF e LSTM para B. . . . .	80

Figura 4.5 – Diagrama de diferenças críticas (CD) dos resultados dos modelos NAIVE, ARIMA, WHOFTS, NFN, MLP, RF e LSTM para C. . . . .	81
Figura 4.6 – Diagrama de diferenças críticas (CD) dos resultados dos modelos NAIVE, ARIMA, WHOFTS, NFN, MLP, RF e LSTM para D. . . . .	82
Figura 4.7 – Acurácia dos modelos de previsão para estimativas com APes individuais menores que 5% para as estações de recebimento B, C e D. . . . .	83
Figura 4.8 – Dados semanais no treinamento — Visão $k \times (k + 1)$ e Visão temporal. . . . .	83
Figura 4.9 – Histogramas dos modelos Yamakawa-NFN vencedores para as variáveis semanais. . . . .	84
Figura 5.1 – Comparação entre dos dados reais, modelo sem adaptação e modelo adaptativo. . . . .	89
Figura 5.2 – Amostra comparativa da previsão pontual e intervalar do modelo Yamakawa Adaptativo para um horizonte de 7 dias. . . . .	91
Figura 5.3 – Resultados na validação para as variáveis semanais — valores reais vs. valores estimados nos horizontes 1 e 2 passos à frente. . . . .	93
Figura 5.4 – Resultados na validação para as variáveis semanais — valores reais vs. valores estimados nos horizontes 3 e 4 passos à frente. . . . .	93

## LISTA DE TABELAS

Tabela 2.1 – Modelos no cenário internacional. . . . .	30
Tabela 2.2 – Modelos no cenário do Brasil. . . . .	31
Tabela 2.3 – Percentuais de <i>fit</i> e <i>predict</i> . . . . .	45
Tabela 4.1 – Resultados do APE usando a abordagem Vespertina no ponto de distribuição A. . . . .	62
Tabela 4.2 – Resultados do APE usando a abordagem Vespertina no ponto de distribuição B. . . . .	62
Tabela 4.3 – Resultados do APE usando a abordagem Vespertina no ponto de distribuição C. . . . .	62
Tabela 4.4 – Resultados do APE usando a abordagem Vespertina no ponto de distribuição D. . . . .	63
Tabela 4.5 – Resultados do APE usando a abordagem Vespertina no ponto de distribuição E. . . . .	63
Tabela 4.6 – Resultados do APE usando a abordagem de previsão de um passo à frente no ponto de distribuição A. . . . .	64
Tabela 4.7 – Resultados do APE usando a abordagem de previsão de um passo à frente no ponto de distribuição B. . . . .	65
Tabela 4.8 – Resultados do APE usando a abordagem de previsão de um passo à frente no ponto de distribuição C. . . . .	66
Tabela 4.9 – Resultados do APE usando a abordagem de previsão de um passo à frente no ponto de distribuição D. . . . .	67
Tabela 4.10 – Resultados do APE usando a abordagem de previsão de um passo à frente no ponto de distribuição E. . . . .	68
Tabela 4.11 – Resultados do APE na segunda etapa para o ponto de distribuição A. . . . .	69
Tabela 4.12 – Resultados do APE na segunda etapa para o ponto de distribuição B. . . . .	69
Tabela 4.13 – Resultados do APE na segunda etapa para o ponto de distribuição C. . . . .	69
Tabela 4.14 – Resultados do APE na segunda etapa para o ponto de distribuição D. . . . .	69
Tabela 4.15 – Resultados do APE na segunda etapa para o ponto de distribuição E. . . . .	70
Tabela 4.16 – Resultados dos modelos Yamakawa-NFN adaptativos na quarta etapa. . . . .	70
Tabela 4.17 – Resultados dos modelos Yamakawa-NFN adaptativos após a fase 2 de testes. . . . .	71
Tabela 4.18 – Resultados na previsão de 7 passos à frente para o ponto de distribuição A. . . . .	73
Tabela 4.19 – Resultados na previsão de 7 passos à frente para o ponto de distribuição B. . . . .	73

Tabela 4.20–Resultados na previsão de 7 passos à frente para o ponto de distribuição C. . . . .	73
Tabela 4.21–Resultados na previsão de 7 passos à frente para o ponto de distribuição D. . . . .	74
Tabela 4.22–Resultados na previsão de 7 passos à frente para o ponto de distribuição E. . . . .	74
Tabela 4.23–Resultados na previsão para de 1 a 7 passos à frente após alteração nas bases de dados. . . . .	75
Tabela 4.24–Erro Percentual Absoluto Médio (MAPE) das previsões de 1, 4 e 7 passos à frente dos modelos NAIVE, ARIMA, WHOFTS, NFN, MLP, RF e LSTM para B. . . . .	78
Tabela 4.25–MAPE das previsões de 1, 4 e 7 passos à frente dos modelos NAIVE, ARIMA, WHOFTS, NFN, MLP, RF e LSTM para C. . . . .	79
Tabela 4.26–MAPE das previsões de 1, 4 e 7 passos à frente dos modelos NAIVE, ARIMA, WHOFTS, NFN, MLP, RF e LSTM para D. . . . .	79
Tabela 4.27–Resultados dos modelos ANFIS-SFC e ANFIS-FCM para variáveis semanais. . . . .	84
Tabela 4.28–Modelo Yamakawa-NFN com 1 a 5 variáveis semanais. . . . .	85
Tabela 5.1 – Resultados do APE usando a abordagem Vespertina nos cinco pontos de distribuição. . . . .	87
Tabela 5.2 – Resultados do APE usando a abordagem de previsão de um passo à frente nos 5 pontos de distribuição. . . . .	88

## LISTA DE ABREVIATURAS E SIGLAS

**AG** Algoritmo Genético

**ANFIS** *Adaptive Neuro-fuzzy Inference System*

**APE** Erro Percentual Absoluto

**ARIMA** *Autoregressive Integrated Moving Average*

**ARX** *Autoregressive-exogenous*

**CNN** *Convolutional Neural Network*

**DFCM** *Deep Fuzzy Cognitive Maps*

**DNN** *Deep Neural Network*

**FCM** *Fuzzy C-Means Clustering*

**FTS** *Fuzzy Time Series*

**GASMIG** Companhia de Gás de Minas Gerais

**GM** *Gray Model*

**GN** Gás Natural

**LSTM** *Long Short Term Memory*

**MAPE** Erro Percentual Absoluto Médio

**MLP** Perceptron Multicamadas

**MSE** *Mean Square Error*

**NFN** *Neo-Fuzzy Neuron*

**PIB** Produto Interno Bruto

**PWFTS** *Probabilistic Weighted Fuzzy Times Series*

**RF** *Random Forest*

**RL** Regressão Linear

**RNA** Redes Neurais Artificiais

**SFC** *Subtractive Fuzzy Clustering*

**SIGM** *Self-adapting Intelligent Gray Model*

**SVM** *Support Vector Machine*

**UKF** *Unscented Kalman Filter*

**WHOFTS** *Weighted High Order Fuzzy Times Series*

## LISTA DE SÍMBOLOS

$n$	número de variáveis de entrada
$x \in \mathbb{R}^n$	entrada, vista como um vetor coluna de $\mathcal{M}_{n \times 1}(\mathbb{R})$
$\ x\ $	norma euclideana = $\sqrt{x_1^2 + x_2^2 + \dots + x_n^2}$
$y, y^d \in \mathbb{R}^q$	saída real, ou de saída, e desejada
$\mu_A(x_i)$	grau de pertinência que $x_i \in A$
$[\underline{\mu}, \bar{\mu}]$	intervalo fechado de pertinência
$H$	total de amostras
$a_{ij}$	$j$ -ésimo valor modal de $x_i$
$m$	número de funções de pertinência de cada entrada
$w$	peso, em $\mathcal{M}_{n \times m}(\mathbb{R})$
$\alpha$	taxa de aprendizagem
$\hat{X}^{a-}$	estimativa a priori da matriz de estados aumentada
$\xi$	resíduo: diferença entre o valor real e o valor estimado
$\hat{X}^{x+}$	estimativa a posteriori da matriz de estados
$P^-, P^+$	estimativas da matriz de covariância a priori e a posteriori
$I_{5 \times 5}$	matriz identidade = $\text{diag}(1,1,1,1,1)$
$v_{11}[k]$	vazão no dia $k$ das 0h às 10h59min
$v_{24}[k]$	vazão no dia $k$ das 0h às 23h59min
$v_{13}[k]$	vazão no dia $k$ das 11h às 23h59min
$\nu[k]$	ruído branco referente ao dia $k$
$v_s[k]$	vazão na semana $k$
$M_a$	modelo auxiliar
$M_0$	modelo de previsão vespertina
$M_k$	modelo de previsão de $k$ dias à frente, $1 \leq k \leq 7$

## SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO</b>	<b>21</b>
1.1	Contexto e Motivação	21
1.2	Objetivos	21
1.2.1	Objetivo Geral	21
1.2.2	Objetivos Específicos	21
1.3	Organização do Texto	22
<b>2</b>	<b>REFERENCIAL TEÓRICO</b>	<b>24</b>
2.1	Apresentação Sobre o Gás Natural e Importância Para o Brasil	24
2.2	Modelos de Previsão de Médio e Longo Prazo	27
2.3	Modelos de curto prazo	28
2.4	Panorama dos Modelos de Análise	30
2.5	<i>Unscented Kalman Filter</i>	31
2.6	Parametrização de Cada Modelo	33
2.6.1	Regressão Linear (RL)	34
2.6.2	Média Móvel	34
2.6.3	<i>Naive</i>	34
2.6.4	Auto-regressivo integrado de médias móveis (ARIMA)	34
2.6.5	<i>Random Forest</i> (RF)	35
2.6.6	Redes Neurais Artificiais	36
2.6.7	ANFIS com <i>Fuzzy C-Means Clustering</i> (ANFIS-FCM)	36
2.6.8	ANFIS com <i>Subtractive Fuzzy Clustering</i> (ANFIS-SFC)	37
2.6.9	Suavização Exponencial	37
2.6.10	<i>First Order Conventional</i> (CFTS)	38
2.6.11	<i>First Order Weighted</i> (WFTS)	38
2.6.12	<i>Weighted High Order</i> (WHOFTS)	39
2.6.13	<i>High Order Probabilistic Weighted</i> (PWFTS)	39
2.7	Rede <i>Neo-Fuzzy-Neuron</i> (NFN)	40
2.7.1	Funções Triangulares	43
2.7.2	Implementação do Modelo de Yamakawa	43
2.7.3	Funções de Pertinência	44
2.7.4	Épocas	44
2.7.5	Percentuais de <i>Fit</i> e <i>Predict</i>	45
2.7.6	Modelo Invariante no Tempo vs Adaptativo	46
2.7.7	Mínimos e Máximos	46
2.7.8	Espaçamento	46

2.8	Métricas de Erro	46
<b>3</b>	<b>METODOLOGIA</b>	<b>48</b>
3.1	Análise Preliminar dos Dados	48
3.2	Escolha das Variáveis	52
3.2.1	A Sazonalidade no ponto de entrega E	52
3.3	Definição das Entradas e Saídas	53
3.3.1	Autocorrelações	54
3.3.2	<i>Boxplots</i>	55
3.3.3	O Critério de Akaike	56
3.4	Modelo Proposto	57
3.4.1	Modelo Intervalar	59
<b>4</b>	<b>EXPERIMENTOS</b>	<b>61</b>
4.1	Etapa Inicial	61
4.1.1	Previsão Vespertina	61
4.1.2	Previsão de Um Passo à Frente	63
4.2	Escolha de Três Modelos Invariantes no Tempo	68
4.3	Um Modelo Adaptativo	70
4.4	Fase de Testes	70
4.5	Previsão de Sete Dias à Frente	71
4.6	Otimização de Parâmetros	73
4.7	Homologação	74
4.8	Previsão de Múltiplos Passos à Frente	74
4.8.1	Metodologia	75
4.8.2	Pré-processamento das bases de dados	75
4.8.3	Configuração dos modelos	76
4.8.4	Experimentos	77
4.8.5	Resultados	78
4.9	Série Temporal Semanal	79
<b>5</b>	<b>DISCUSSÃO</b>	<b>86</b>
5.1	Etapa Inicial	86
5.1.1	Previsão Vespertina	86
5.1.2	Previsão de Um Passo à Frente	87
5.2	Escolha de Três Modelos Invariantes no Tempo	89
5.3	Um Modelo Adaptativo	89
5.4	Fase de Testes	90
5.5	Previsão de Sete Dias à Frente	90
5.6	Otimização de Parâmetros	90
5.7	Homologação	91
5.8	Previsão de Múltiplos Passos à Frente	91

5.9	Série Temporal Semanal	92
<b>6</b>	<b>CONSIDERAÇÕES FINAIS</b>	<b>94</b>
6.1	Trabalhos Futuros	95
	<b>REFERÊNCIAS</b>	<b>97</b>
<b>A</b>	<b>APÊNDICE — Códigos-Fonte</b>	<b>101</b>
A.1	Regressão Linear (RL)	101
A.2	Média Móvel	101
A.3	<i>Naive</i>	101
A.3.1	<i>Naive Sazonal</i>	101
A.4	Auto-regressivo integrado de médias móveis (ARIMA)	101
A.5	<i>Random Forest</i> (RF)	102
A.6	Redes Neurais Artificiais	102
A.7	ANFIS com <i>Fuzzy C-Means Clustering</i> (ANFIS-FCM)	103
A.8	ANFIS com <i>Subtractive Fuzzy Clustering</i> (ANFIS-SFC)	103
A.9	Suavização Exponencial	103
A.10	<i>First Order Conventional</i> (CFTS)	103
A.11	<i>First Order Weighted</i> (WFTS)	104
A.12	<i>Weighted High Order</i> (WHOFTS)	104
A.13	<i>High Order Probabilistic Weighted</i> (PWFTS)	104

# 1 INTRODUÇÃO

## 1.1 Contexto e Motivação

O Gás Natural (GN) é parte essencial da matriz energética brasileira, considerado uma fonte limpa em comparação ao carvão, por exemplo (IEA, 2022). A demanda tem aumentado e prever a necessidade e seu consumo fazem parte de um bom planejamento energético, que garantirá um desenvolvimento do país (GAO; SHAO, 2021). Apesar da importância da previsão de consumo de GN, há uma lacuna considerável de trabalhos na área.

Na legislação, costuma-se pagar multas sobre erro acima de uma faixa percentual de zero a 5%. Portanto, a motivação é economizar gastos.

## 1.2 Objetivos

### 1.2.1 Objetivo Geral

O objetivo deste trabalho é definir um padrão para previsão de consumo de GN em Minas Gerais, Brasil, nos horizontes diário e semanal, da empresa Companhia de Gás de Minas Gerais (GASMIG), previsão essa que poderá ser utilizada em outras empresas e países.

### 1.2.2 Objetivos Específicos

Os objetivos específicos deste trabalho são:

- A hipótese inicial é que o algoritmo de Takeshi Yamakawa (1992) é comparável. Objetiva-se provar essa hipótese.
- Comparar algoritmos comumente utilizados no estado da arte com o algoritmo de Yamakawa.
- Investiga-se a previsão de consumo de GN no Brasil com foco no curto prazo: previsão nas próximas 13 horas, previsão de um a sete passos diários à frente, além da previsão de um a quatro passos semanais à frente.
- É definido um modelo *baseline*, o qual deve ser melhorado.

### 1.3 Organização do Texto

O trabalho começa com uma revisão teórica, apresentada no Capítulo 2, que destaca o papel do GN na matriz energética brasileira e sua importância para o país, incluindo dados sobre reservas mundiais e consumo. São discutidos modelos de previsão de consumo de médio e longo prazo, assim como os de curto prazo, que são mais relevantes para o objetivo do trabalho. O modelo de previsão de Yamakawa (1992) é destacado por sua eficiência e simplicidade.

Explica-se que o GN é uma fonte de energia fóssil limpa e é crucial na matriz energética da Argentina (que foi a maior consumidora de GN na América do Sul e Central no ano de 2018, cf Figura 2.2), onde é usado principalmente para a indústria, geração de eletricidade e aquecimento doméstico (*BP Statistical Review of World Energy* — 2019). A indústria do GN é dividida em produção, processamento, transmissão, armazenamento e distribuição, com a regulação sendo feita pelo governo ou por meio de contratos de fornecimento (SOLDO et al., 2014).

De fato, a previsão precisa do consumo de GN é essencial para planejamento de curto prazo por empresas distribuidoras, devido a questões regulatórias e para evitar perdas econômicas (KHAN, 2015). A previsão é desafiadora devido à complexidade do consumo e aos diversos fatores que influenciam, como parâmetros econômicos, dados meteorológicos e calendário (LIU et al., 2021). A seguir, é feita uma introdução às redes *Neo-Fuzzy Neuron* (NFN) e ao *Unscented Kalman Filter* (UKF), que são os precursores do modelo que será proposto. Mais à frente, a parametrização de vários modelos é discutida, incluindo Regressão Linear, Média Móvel, *Naive*, *Autoregressive Integrated Moving Average* (ARIMA), *Random Forest* (RF), Redes Neurais Artificiais, *Adaptive Neuro-fuzzy Inference System* (ANFIS) com *Fuzzy C-Means Clustering* (FCM), ANFIS com *Subtractive Fuzzy Clustering* (SFC), Suavização Exponencial e modelos de série temporal nebulosa de alta ordem.

A seguir, para o modelo NFN, são discutidas funções de pertinência, épocas, percentuais de *fit* e *predict*, modelo invariante no tempo *versus* adaptativo, mínimos e máximos e espaçamento; e são definidas as métricas de erro que serão utilizadas.

No Capítulo 3, a metodologia deste trabalho é apresentada, incluindo a análise preliminar dos dados, a escolha das variáveis, a definição das entradas e saídas, as autocorrelações, os *boxplots* e o Critério de Akaike para seleção de variáveis.

Adiante é apresentado o modelo proposto, que combina modelos Yamakawa NFN com UKF para previsão do consumo diário e semanal de GN. É discutido um modelo intervalar para estimar intervalos de incerteza.

No Capítulo 4, os experimentos são descritos em etapas: (1) a etapa inicial

utiliza dados de 2018 a 2022 para treinamento e validação; (2) a escolha de três modelos invariantes no tempo, e de um modelo adaptativo; (3) duas etapas de testes; (4) a etapa de prever sete dias à frente; (5) a otimização dos parâmetros; (6) a homologação; (7) a previsão de múltiplos passos à frente (em que se descreve a metodologia, o pré-processamento das bases de dados, a configuração dos modelos, as métricas e os experimentos, com seus resultados) e a etapa (9), que é uma série temporal semanal.

No Capítulo 5, é feita a discussão e a análise dos dados; a explicação sobre o modelo de análise de curto prazo; e são apresentadas as considerações finais. Conclui-se, no Capítulo 6, que os modelos propostos apresentam melhorias em relação ao modelo *baseline*, especialmente em estimativas abaixo de 5% do Erro Percentual Absoluto (APE), faixa isenta do pagamento de multa.

## 2 REFERENCIAL TEÓRICO

Neste capítulo, apresentam-se os trabalhos que antecedem e servem de base para o desenvolvimento desta dissertação. Primeiramente, é importante destacar o papel do GN na matriz energética brasileira, sua previsão de uso e crescimento. Adiante, são mostrados os modelos de previsão de consumo de médio e longo prazo, para depois os modelos de curto prazo, que são mais relacionados ao objetivo deste trabalho. Por fim, é necessário destacar o modelo predictor de Yamakawa (1992) — escolhido por sua conhecida eficiência e simplicidade — que direciona as análises dos dados.

### 2.1 Apresentação Sobre o Gás Natural e Importância Para o Brasil

O GN é uma fonte de energia fóssil formada a partir da decomposição de matéria orgânica de animais e plantas pré-históricos. Ele é extraído por meio da perfuração de depósitos subterrâneos e pode ser encontrado tanto associado quanto não associado a depósitos de petróleo. Em comparação a outros combustíveis fósseis, como o carvão e o petróleo, o GN é considerado uma fonte de energia mais limpa, de acordo com o IEA (2022).

Gomes et al. (2014) ilustra (Figura 2.1) a configuração da exploração do GN no mundo, onde as reservas mundiais de GN totalizam 1,129 bilhão boe<sup>1</sup>, equivalendo a 95% das de Petróleo (1,188 bilhão boe). Os autores fazem referência à GASMIG, que é uma das 23 distribuidoras do país (GASNET, 2006).

É importante destacar que, segundo dados da GASMIG, 88,5% do GN é composto de metano (não confundir com **biogás** de compostagem). O GN é responsável por 12,2% da composição da matriz energética brasileira (levando em consideração produtos de cana, lenha e carvão vegetal, hidráulica, urânio, carvão mineral, GN, petróleo, outras). (MICHELENA, 2019)

A Argentina foi a maior consumidora de GN na América do Sul e Central no ano de 2018, produzindo 40 bilhões de metros cúbicos de GN em 2018, como mostra a Figura 2.2, sendo o 16º maior produtor do energético no mundo. Atualmente, o GN é a base da matriz energética primária da Argentina, onde é usado, principalmente, para apoiar sua indústria, gerar eletricidade e aquecimento para usuários domésticos.

A indústria do GN é dividida em produção e processamento, transmissão e armazenamento, e distribuição. Para a distribuição, o GN é transportado por dutos até

---

<sup>1</sup> Barril de Petróleo (Boe): unidade utilizada para comparar (converter) em equivalência térmica, uma quantidade de energia de barris de petróleo.

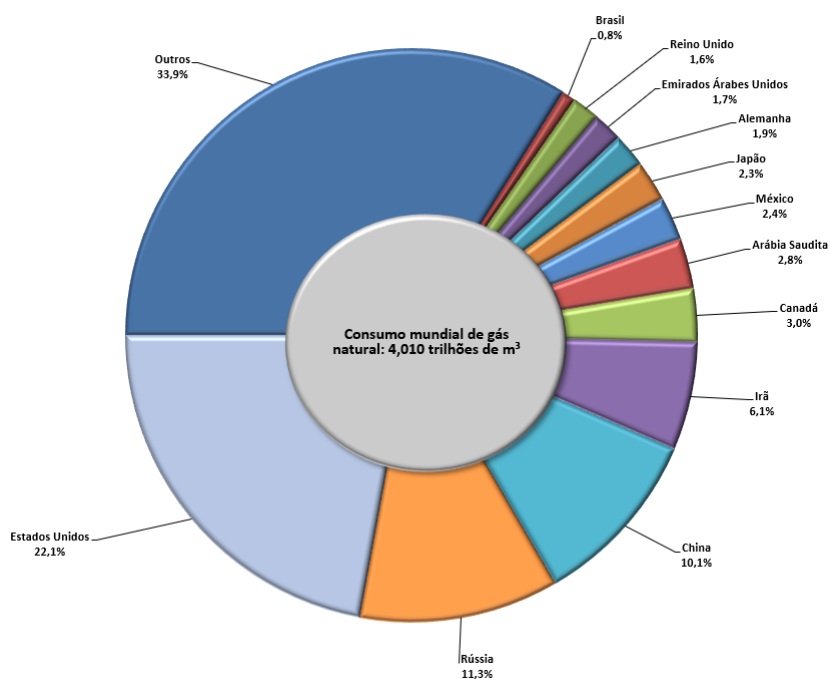


Figura 2.1 – Reservas Mundiais de Gás Natural.

Fonte: *Energy Institute, Statistical Review of World Energy 2024.*

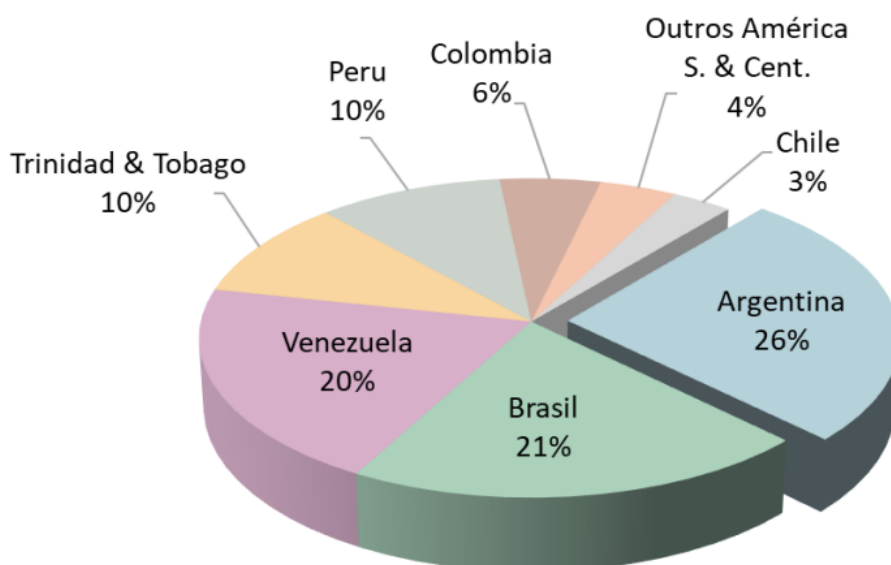


Figura 2.2 – Maiores países consumidores de Gás Natural na América Sul e Central no ano de 2018.

Fonte: *BP Statistical Review of World Energy — 2019.*

empresas distribuidoras, que fornecem para consumidores residenciais, comerciais e industriais. A regulação do consumo de GN é feita pelo governo ou por meio de contratos de fornecimento, que geralmente incluem penalidades para previsões de consumo com **erros fora de um intervalo** de tolerância (SOLDO et al., 2014).

Para evitar perdas econômicas, previsões precisas do consumo de GN são cruciais para empresas distribuidoras no planejamento de curto prazo. Além das questões regulatórias, a previsão de GN auxilia no planejamento de investimentos futuros em infraestrutura e políticas de gestão desse recurso (SHAIKH; JI, 2016).

Segundo Khan (2015), a natureza da demanda de energia e o conhecimento de seus determinantes são cruciais para previsões precisas das necessidades de energia. Por isso, a previsão do consumo de GN tem sido investigada em diferentes níveis (mundial, nacional, regional, por sistema de distribuição, por cliente individual, etc), diferentes horizontes de previsão (curto, médio e longo prazo) e diferentes fatores de influência (parâmetros econômicos, dados meteorológicos, dados históricos de consumo de GN, demográficos e de calendário) (LIU et al., 2021).

O consumo também varia de acordo com o tipo de consumidor. Consumidores industriais consomem a mesma quantidade de GN não importa qual seja o dia do mês ou a estação do ano. Consumidores residenciais que usam o gás para aquecimento, são caracterizados por ter um consumo significativamente menor no verão do que no inverno. Já os consumidores residenciais, que usam o gás apenas para cozinhar, são caracterizados pela variação diurna do consumo (SZOPLIK, 2015). Gao e Shao (2021) ainda acrescentam que as diferentes condições e estágios de desenvolvimento de um país também influenciam no seu consumo de GN.

A previsão de consumo de GN é uma tarefa desafiadora devido à sua complexidade e à influência dos vários fatores citados. Ao longo dos anos, são testados muitos modelos em diferentes cenários. Segundo Liu et al. (2021), a partir de 2006, com a conscientização sobre conservação de energia e proteção ambiental, o número de trabalhos na área aumentou e envolve modelos como: modelos de aprendizado de máquina gerais e de aprendizado profundo (*deep learning*), modelo combinado (*ensemble learning*) e modelos estatísticos tradicionais.

No Brasil, o GN ocupa a terceira posição na matriz energética nacional, ficando atrás apenas do petróleo e da cana-de-açúcar. Nos últimos dez anos, sua participação evoluiu de 5,4% para 13,3%, graças à sua utilização em termoeletricas e à extensão da malha dutoviária, que permitiu seu uso tanto pela indústria quanto em edificações residenciais, comerciais e públicas. O setor industrial é o maior consumidor de GN no país, respondendo por cerca de 34% do consumo total, e esse perfil tem se mantido relativamente constante desde 2000. Já no setor residencial, o consumo é de 11,5%, e sua utilização está relacionada à cocção de alimentos e aquecimento de água (EPE, 2022).

Apesar da importância do GN na matriz energética brasileira, especialmente para a indústria, a literatura tem uma grande lacuna, uma vez que, no Brasil, não é encontrado nenhum trabalho sobre previsão de curto prazo e apenas um de médio

prazo. Nesse trabalho, Santos (2019) aplica Regressão Linear (RL) para prever o consumo mensal de GN para quatro segmentos: (1) residencial e comercial, (em que o modelo usa taxa de câmbio), (2) automotivo, (3) industrial (em que usa indicador de produção industrial) e (4) geração de energia elétrica (em que usa energia natural afluyente). Todos os modelos usam como entrada o crescimento populacional, Produto Interno Bruto (PIB) e dados históricos de consumo.

## 2.2 Modelos de Previsão de Médio e Longo Prazo

Pesquisas que fazem previsões de longo e médio prazo trabalham com séries de dados anuais, trimestrais ou mensais. Normalmente, essas previsões utilizam uma variação maior de variáveis de entrada, como fatores econômicos, população, dados climáticos e dados históricos de consumo. Em Gao e Shao (2021), o PIB e a população são utilizados para treinar um modelo ANFIS para previsão anual de GN em quatro países do Oriente Médio. Fan, Wang e Hong (2018) utilizam apenas dados históricos de consumo para prever o consumo anual de GN na China usando um *ensemble learning* composto dos modelos *Gray Model* (GM) e *Self-adapting Intelligent Gray Model* (SIGM). Esses modelos são combinados através de uma média ponderada de suas previsões. Os pesos são ajustados por um Algoritmo Genético (AG).

Por sua vez, Sen, Gunay e Tunc (2019) utilizam dados históricos de população, PIB, inflação, preço do GN e a temperatura média do inverno para fazer previsões anuais na Turquia. Para isso, utilizam os modelos de regressão logarítmica, linear e não linear e análise de correlação para selecionar os valores defasados mais relevantes.

Já Beyca et al. (2019) apresentam um estudo sobre previsão mensal de GN em Istambul na Turquia. Como variáveis de entrada usa-se consumo de GN, temperatura, preço do GN e população. Os modelos de RL Multivariada, *Support Vector Machine* (SVM) e Redes Neurais Artificiais (RNA) são testados e comparados. O SVM apresenta melhores resultados. Meira, Oliveira e Menezes (2022) propõem prever o consumo mensal de GN de 18 economias europeias usando dados históricos de consumo. A abordagem seguida é um *ensemble* do tipo *Bagging* composto pelos modelos RNA, suavização exponencial e ARIMA. As previsões são combinadas usando regularização modificada M-Ridge e M-LASSO.

No Brasil, com relação a previsões de longo prazo, três trabalhos são publicados. Behrouznia et al. (2010) usam cinco modelos ANFIS para fazer previsões anuais para cinco países: Argentina, Brasil, Colômbia, Venezuela e Cuba. Os modelos são treinados usando dados históricos de população, PIB e consumo de GN. Costa et al. (2019) também usam as mesmas variáveis para fazer a previsão anual no Brasil com o modelo de estimativa dos mínimos quadrados robusto. O estudo apresenta um modelo

para previsão do consumo de GN de todos os setores (industrial, geração de energia elétrica, comercial e público, residencial e transporte) e dois modelos específicos para indústria e geração de energia elétrica, respectivamente.

Já em [Gracias, Lourenço e Rafikov \(2012\)](#), são avaliados modelos utilizados para estudo da dinâmica populacional para prever o consumo, a produção e a importação anual de GN no Brasil. Os autores concluem que o modelo logístico é o mais indicado para previsão de consumo, exponencial para previsão de produção e *Bertalanffy* para previsão de importação.

[Santos \(2019\)](#) aplica RL para prever o consumo mensal de GN para 4 segmentos no Brasil: (1) residencial e comercial, (2) automotivo, (3) industrial e (4) geração de energia elétrica. Todos os modelos usam como entrada o crescimento populacional, PIB e dados históricos de consumo. Além disso, o modelo (1) usa taxa de câmbio, (3) indicador de produção industrial e (4) energia natural afluyente.

### 2.3 Modelos de curto prazo

Previsões de curto prazo envolvem séries de dados semanais, diárias ou horárias e permitem que empresas de distribuição consigam equilibrar a oferta e a demanda de seus consumidores. Para essas previsões, utilizam-se principalmente dados climáticos, o tipo de data e dados históricos de consumo. [Ravnik e Hribersek \(2019\)](#) apresentam um estudo que cria perfis de consumo para realizar a previsão diária de consumo de GN de um passo à frente na Eslovênia. As previsões são feitas levando-se em consideração dados históricos de consumo e temperatura usando regressão logística. Para cada consumidor, é necessária a confecção de um modelo único e também um modelo para todos os consumidores. Também são desenvolvidos dois modelos para consumos influenciados e não influenciados pela temperatura. Além disso, cria-se um modelo para os dias da semana, um para fins de semana e outro para todos os dias. No total, são desenvolvidos oito modelos para representar cada perfil de consumo.

A previsão de consumo de GN de um passo à frente também é investigada ([SOLDO et al., 2014](#)). Nesse estudo, o objetivo é verificar a influência da radiação solar na previsão de dois sistemas de consumo: casa individual e distribuidora local. Os atributos de entrada usados são: histórico de consumo, temperatura, radiação e um índice futuro de consumo semanal de GN. São avaliados os modelos RNA, SVM, passeio aleatório, modelo de correlação de temperatura, regressão *stepwise* e um modelo linear *Autoregressive-exogenous* (ARX). Os autores concluem que o uso da radiação solar melhora a precisão das previsões e que o modelo ARX apresenta melhor desempenho.

Uma pesquisa de previsão de curto prazo de consumo de GN em 62 áreas operacionais de empresas de distribuição nos Estados Unidos é apresentado (MERKEL; POVINELLI; BROWN, 2018). Essas regiões contemplam áreas urbanas e rurais e climas diferentes. As variáveis de entrada incluem graus-dia de aquecimento, ponto de orvalho, graus-dia de resfriamento, dia da semana e dia do ano. Quatro modelos são comparados: RL, uma RNA e duas *Deep Neural Network* (DNN). Os resultados mostraram que o desempenho dos modelos dependem da área testada. Em alguns casos, a RL obtém melhor resultado que os modelos DNN.

É feito um estudo (ANAGNOSTIS; PAPAGEORGIOU; BOCHTIS, 2020) com RNA para previsão de consumo diário de GN em 15 cidades da Grécia. Avalia-se uma RNA treinada com dados históricos de consumo e temperatura, uma *Long Short Term Memory* (LSTM) treinada apenas com dados históricos de consumo e uma DNN com dados históricos de consumo, temperatura e indicadores de dia (mês, dia do mês, dia da semana e feriado). A série temporal é discretizada em dias, 365 pontos, depois 730, 1095 e 1460, de que é analisada a dinâmica através de cinco métricas de erro. Resultados apontam melhor desempenho para o modelo DNN.

Por sua vez, Du et al. (2022) apresentam uma abordagem híbrida de modelos *deep learning* para previsão de consumo diário de GN de várias regiões da China. A estrutura funciona da seguinte forma: um *autoencoder* é usado para extrair informações constantes e/ou dinâmicas da série histórica de consumo, uma LSTM é aplicada também à série de consumo para capturar o padrão de dependência de longo prazo e, por fim, uma camada de rede personalizada faz o acoplamento dessas informações para realizar a previsão. A abordagem proposta supera outros modelos como LSTM, um modelo híbrido de *Convolutional Neural Network* (CNN) e LSTM, SVM e árvore de decisão.

O estudo de Rehman et al. (2022) apresenta uma comparação de modelos *deep learning* e métodos estatísticos na previsão horária de consumo de GN em cinco locais na Espanha. As variáveis de entrada usadas são: dados históricos de consumo, indicador de feriados ao longo do ano e condições meteorológicas.

Em Andelkovic e Bajatovic (2020), as variáveis meteorológicas são utilizadas juntamente com o históricos de consumo para previsão horária de GN na cidade de Novi Sad na Sérvia. Os autores perceberam que as variáveis temperatura, radiação e umidade relativa tinham grande influência no consumo de GN, já que grande parcela é usada para aquecimento das residências na cidade. O modelo ANFIS é utilizado para fazer as previsões.

Outros três trabalhos de previsão horária fazem uso de técnicas de decomposição de séries temporais como forma de extrair mais informações dos dados originais. Em Svoboda, Kotik e Platos (2021), é feita uma comparação dos algoritmos

RF e Gradiente *Boosting*. Os autores utilizam várias variáveis para treinar os modelos, são elas: componente sazonal, de tendência e residual da série temporal de consumo de GN, indicador de feriados e dias que os antecedem, estações do ano, variáveis climáticas, preço do GN, indicador de dia do mês, dia da semana e mês.

Su et al. (2019) apresentam uma estrutura híbrida que usa a transformada de *Wavelet*, uma rede LSTM e um AG. A transformada de *Wavelet* é usada para decompor a série histórica de consumo de GN em componentes de alta e baixa frequência. Essas componentes são dadas como entrada para a LSTM, cuja arquitetura é selecionada usando um AG. Qiao et al. (2019) decompõe a série original de consumo de GN em várias IMFs (*Intrinsic Mode Function*) e um termo de tendência residual para treinar três tipos de RNA e um SVM.

## 2.4 Panorama dos Modelos de Análise

São revisados 22 estudos publicados entre 2010 e 2022 que abordam a previsão de consumo de GN. Entre esses estudos, três são revisões sistemáticas. O foco principal é compreender o panorama da previsão, destacando o horizonte temporal, os fatores que influenciam o consumo, os modelos utilizados e o nível de aplicação das previsões. A Tabela 2.1 apresenta um resumo dos trabalhos nesses termos.

Horizonte	Fatores de influência (Variáveis preditoras)	Modelos	Nível
Longo (anual)	PIB, População, Consumo, Inflação, Preço do GN.	ANFIS, Ensemble Learning, RL,	País.
Médio (mensal)	PIB, População, Consumo, Preço do GN, Variáveis climáticas.	Logística e Regressão Não linear, SVM, Perceptron Multicamadas (MLP), Deep Learning (LSTM, DNN),	País, Regional.
Curto (diário e horário)	Consumo, Variáveis climáticas, Indicadores de dia, uso de componentes da série original de consumo.	Métodos estatísticos, Abordagens Híbridas, RF, Gradiente Boosting.	Regional, Área de distribuição, Setor, Cliente individual.

Tabela 2.1 – Modelos no cenário internacional.

Observa-se que os trabalhos com horizonte de longo prazo concentram-se na previsão do consumo nacionalmente, considerando principalmente variáveis econômicas e o crescimento populacional. Já os estudos que investigam horizontes de médio prazo também incluem variáveis climáticas em suas análises.

Por outro lado, para horizontes de curto prazo, as previsões são realizadas em níveis mais específicos, chegando até mesmo ao consumo em residências individuais.<sup>2</sup> Nesses casos, as variáveis mais utilizadas são as relacionadas ao clima, indicadores diários e a decomposição da série temporal original do consumo.

No contexto brasileiro, as publicações analisadas se limitaram a previsões de longo e médio prazo. A Tabela 2.2 apresenta um resumo desses estudos. É importante observar que, em relação aos fatores que influenciam o consumo, os estudos brasileiros seguem o padrão de outros países, com exceção das variáveis climáticas. Isso ocorre porque o GN não é amplamente utilizado para aquecimento residencial no Brasil e tem sua principal aplicação na indústria.

Horizonte	Fatores de influência (Variáveis preditoras)	Modelos	Nível
Longo (anual)	PIB, População, Consumo.	ANFIS, RL, Mínimos Quadrados Robusto, Modelo logístico.	País e Setor.
Médio (mensal)	PIB, População, Consumo, Indicador de prod. industrial.		

Tabela 2.2 – Modelos no cenário do Brasil.

Outro aspecto relevante é que, em âmbito global, são avaliadas diversas variedades de modelos para a previsão de consumo de GN, incluindo modelos estatísticos, *machine learning* em geral, *deep learning*, *ensemble learning* e abordagens híbridas. No entanto, no Brasil, são testados poucos modelos.

Podemos concluir que existe uma lacuna significativa em estudos de previsão de GN no Brasil, especialmente em relação a horizontes de previsão de curto prazo, como diário e horário, e o uso de modelos de *machine learning*. Nesse sentido, é de extrema importância o desenvolvimento de pesquisas que contemplem o cenário brasileiro nas previsões de GN.

## 2.5 Unscented Kalman Filter

Aguirre (2007) expõe os Filtros de Kalman, uma ideia que remonta a Kalman (1960) e toda a teoria de observadores do espaço de estados. Por sua vez, Wan e Merwe (2001) apresentam o UKF. A Figura 2.3 mostra o fluxo do algoritmo que será utilizado, segundo Liu e Wang (2021).

<sup>2</sup> Neste trabalho, o consumo é previsto por cidade e/ou ponto de distribuição.

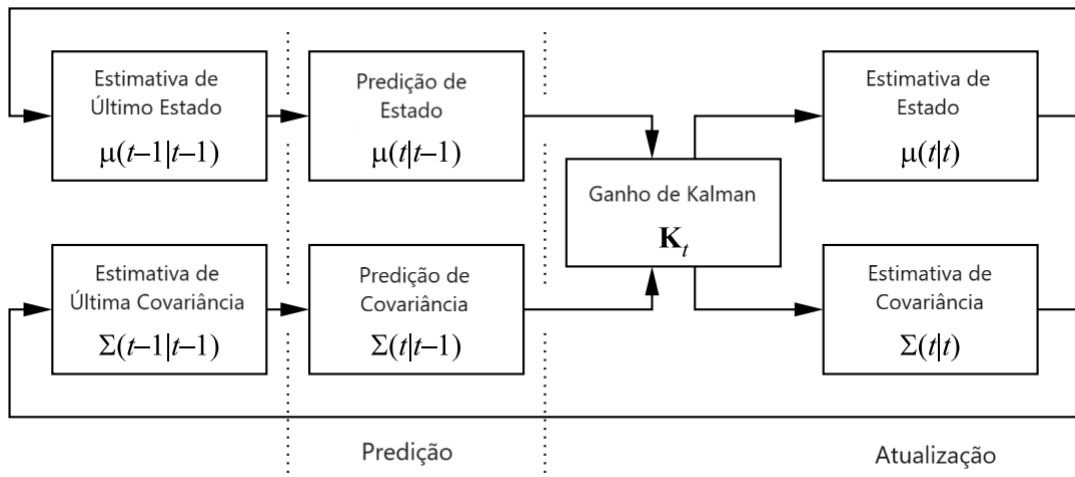


Figura 2.3 – Fluxo do processo do *Unscented Kalman Filter*.

Fonte: Liu e Wang (2021).

Existem um filtro de Kalman linear; um filtro de Kalman não linear, mas ainda linearizado pela matriz jacobiana; e o *Unscented Kalman Filter*, que não precisa das fórmulas explícitas das derivadas; ao invés disso, vale-se dos chamados pontos sigma.

Na fase inicial, que é executada uma única vez, inicializam-se as variáveis, tais quais os pesos  $w_c$  (pondera estados na Fase 3) e  $w_m$  (pondera momentos na Fase 4). Ambos os pesos são  $11 \times 1$ , isto é, matrizes de 11 linhas e 1 coluna.

Para nenhuma variável de entrada, 2 variáveis de estados e 1 variável de saída, a matriz aumentada  $P^a$  é inicializada com  $\text{diag}(100,100,10,10,1)$ , em que 100 ilustra a variância dos estados, 10 ilustra a covariância de ruídos de entrada entre si e 1 ilustra a variância do ruído de saída. E a quantidade de pontos sigma é inicializada com 11. Cada amostra é processada em seis fases:

#### • Fase 1

Na fase de criar pontos sigma, é selecionada uma população de pontos adequada, pela decomposição de Cholesky.  $L_{5 \times 5}$  funciona como se fosse a raiz quadrada de  $P^a$ , conforme expresso na Equação 2.1, em que  $x^a \in \mathbb{R}^5$ , inicialmente, são 2 estados concatenados com (0,0,0).

$$P^a = L^T L; X^a = [\hat{x}^a, \hat{x}^a + \gamma L, \hat{x}^a - \gamma L] \quad (2.1)$$

#### • Fase 2

Na fase de propagação, estimam-se 22 vezes os modelos Yamakawa (1992). Esta é a fase mais importante, em que os estados “passam para a frente”, conforme a

Equação 2.2, em que  $H_{1 \times 11}$  é uma constante.

$$\hat{X}_{ij}^{a-} = F(X_{ij}^a); \hat{Y}_i = H\hat{X}_i^a; 1 \leq i \leq 11; 1 \leq j \leq 2 \quad (2.2)$$

### • Fase 3

Na fase de atualizar momentos, são estimadas as covariâncias dos resíduos  $\xi_{1i} = X_i - \hat{X}_i$  e  $\xi_{2i} = Y_i - \hat{Y}_i$ . Elas são expressas na Equação 2.3, em que a matriz  $P^-$  é  $2 \times 2$ ,  $P_{yy}$  é escalar e  $P_{xy}$  é  $2 \times 1$ .

$$P^- = \sum_{i=1}^{11} w_c^i \xi_{1i} \xi_{1i}^\top; P_{yy} = \sum_{i=1}^{11} w_c^i \xi_{2i} \xi_{2i}^\top; P_{xy} = \sum_{i=1}^{11} w_c^i \xi_{1i} \xi_{2i}^\top \quad (2.3)$$

### • Fase 4

Nesta fase, é feita uma previsão a priori de  $\hat{X}^a[k]$  e  $\hat{Y}[k]$ , conforme a Equação 2.4. Podem-se entender tais previsões com suas variâncias como sendo uma variável aleatória gaussiana.

$$\hat{X}^{x-} = \sum_{i=1}^{11} w_m^i \hat{X}_i^x \in \mathbb{R}^2; \hat{Y}^- = \sum_{i=1}^{11} w_m^i Y_i \quad (2.4)$$

### • Fase 5

Na fase de assimilação, é feita uma previsão a posteriori de  $\hat{X}^a[k]$  e é calculada sua incerteza  $P^+$ , conforme a Equação 2.5, em que  $\eta$  é a denominada inovação, escalar, e  $K_{2 \times 1}$  é o ganho de Kalman. Esta é a fase que aguarda chegar o  $Y[k+1]$  real.

$$K = P_{xy} P_{yy}^{-1}; \eta = Y^{\text{real}} - \hat{Y}; \hat{X}^{x+} = \hat{X}^{x-} + K\eta; P^+ = P^- - K P_{yy} K^\top \quad (2.5)$$

### • Fase 6

Na fase de passar para a próxima amostra, garantimos que todos os autovalores de  $P$  são positivos. Se  $P$  não é definida positiva, então  $P = P + 10^{-5} I_{5 \times 5}$ .

## 2.6 Parametrização de Cada Modelo

Os principais hiperparâmetros dos modelos são otimizados usando *Grid Search* e os demais são mantidos em seus valores padrão (*default*) da biblioteca. Os experimentos utilizam MatLab e bibliotecas Python como a Scikit-learn<sup>3</sup>, e códigos próprios.

### 2.6.1 Regressão Linear (RL)

Este modelo<sup>4</sup> é semelhante e mais básico que o utilizado no Critério de Akaike (Seção 3.3.3). Para entradas  $x_1, x_2, \dots, x_n$ , e uma saída escalar  $y$ , a modelagem consiste em definir  $\theta \in \mathbb{R}^{n+1}$ , o vetor de parâmetros a serem determinados, tal que vale a Equação 2.6.

$$\hat{y} = \sum_{i=1}^n \theta_i x_i + \theta_0 \quad (2.6)$$

Deseja-se que a função objetivo  $J = \|y - \hat{y}\|$ , atinja o mínimo, de acordo com Aguirre (2007). Define-se cada resíduo como  $\xi_i = y_i - \hat{y}_i$ . Pela soma dos quadrados dos resíduos, pode-se determinar o quanto um modelo é linear (alta ou baixa linearidade).

### 2.6.2 Média Móvel

No modelo Média Móvel<sup>5</sup>, como o nome já diz, a média das informações dos últimos sete dias ( $k - 6$  a  $k$ ) é propagada para o dia  $k + 1$ , conforme a Equação 2.7.

$$\hat{y}[k + 1] = \frac{y[k] + y[k - 1] + y[k - 2] + y[k - 3] + y[k - 4] + y[k - 5] + y[k - 6]}{7} \quad (2.7)$$

### 2.6.3 Naive

O modelo *Naive*<sup>6</sup> é utilizado na previsão de 1 a 7 dias à frente, que é feita da seguinte maneira: os valores do dia  $k$  são replicados para os dias  $k + 1$  a  $k + 7$ .

Por sua vez, na versão sazonal<sup>7</sup>, a previsão é feita com as informações dos últimos sete dias ( $k - 6$  a  $k$ ), para então projetar os sete dias à frente.

<sup>3</sup> <https://scikit-learn.org/stable/>

<sup>4</sup> cf. Apêndice A.1

<sup>5</sup> cf. Apêndice A.2

<sup>6</sup> cf. Apêndice A.3

<sup>7</sup> cf. Apêndice A.3.1

### 2.6.4 Auto-regressivo integrado de médias móveis (ARIMA)

Define-se o sinal  $\Delta y[k] = y[k] - y[k - 1]$ , o ruído branco  $\nu[k]$  e os atrasos  $q^n y[k] = y[k + n]$ .

Conforme [Aguirre \(2007\)](#), este modelo<sup>8</sup> é dado por  $A(q)\Delta y[k] = C(q)\nu[k]$ , em que  $A$  e  $C$  são polinômios, de graus fixados  $\deg A(q) = n_y$  e  $\deg C(q) = n_\nu$ . Fatora-se  $A$ . A Equação 2.8 relembra fatoração de polinômios  $P(t)$ , cujas raízes são  $t_i$ .

$$a_2 t^2 + a_1 t + a_0 = a_2 (t - t_1)(t - t_2) = a_2 t_1 t_2 \left(1 - \frac{t}{t_1}\right) \left(1 - \frac{t}{t_2}\right) \quad (2.8)$$

$$\sum_{i=0}^n a_i t^i = a_n \prod_{i=1}^n (t - t_i) = (-1)^n a_n \prod_{i=1}^n t_i \left(1 - \frac{t}{t_i}\right)$$

No modelo **ARIMA**, exige-se que  $A(q) = (1 - \lambda_1 q^{-1})(1 - \lambda_2 q^{-1}) \cdots (1 - \lambda_{n_y} q^{-1})$ , com  $\lambda_1 = 1$ . Determinam-se os  $n_\nu + 1$  coeficientes de  $C$ , bem como  $\lambda_2, \lambda_3, \dots, \lambda_{n_y}$ . E otimizam-se os hiperparâmetros, utilizando a biblioteca **Pyramid**<sup>9</sup>.

### 2.6.5 Random Forest (RF)

[Svoboda, Kotik e Platos \(2021\)](#) mostram que o algoritmo **RF**<sup>10</sup> está na categoria de algoritmos baseados em árvore de decisão, que em geral são capazes de lidar com dinâmica não linear em séries temporais muito bem. São otimizados os parâmetros:

- `bootstrap` = `False` — toda a base de dados é utilizada para construir cada árvore;
- `max_depth` — a profundidade máxima da árvore;
- `max_features` — o número de características a considerar, a cada partição, ao procurar a melhor partição;
- `min_samples_leaf` — o número mínimo de amostras necessárias para estar em um nó de folha. Um ponto de particionamento em qualquer profundidade apenas é considerado se deixar pelo menos `min_samples_leaf` amostras de treino em cada um dos ramos esquerdo e direito. Isso pode dar um efeito de suavização do modelo durante a regressão;
- `min_samples_split` — o número mínimo de amostras requerido para particionar um nó interno;
- `n_estimators` — o número de árvores na floresta;

<sup>8</sup> cf. Apêndice [A.4](#)

<sup>9</sup> <https://github.com/alkaline-ml/pmdarima>

<sup>10</sup> cf. Apêndice [A.5](#)

- `cv = 3` — determina a estratégia de particionamento de validação cruzada. Especifica o número de partes em um *KFold* estratificado — que divide os dados em 3 partes consecutivas;
- `n_jobs` — número de tarefas a executar em paralelo. -1 significa utilizar todos os processadores.

### 2.6.6 Redes Neurais Artificiais

É utilizado o modelo<sup>11</sup> MLP<sup>12</sup> e são personalizados os parâmetros:

- `hidden_layer_sizes` — o *i*-ésimo elemento representa o número de neurônios na *i*-ésima camada escondida;
- `activation` — função de ativação para a camada escondida:
  - ‘logistic’, a função sigmoide logística, retorna  $f(x) = 1/(1 + \exp(-x))$ ;
  - ‘tanh’, a função tangente hiperbólica, retorna  $f(x) = \tanh x$ ;
  - ‘relu’, a função unidade linear retificada, retorna  $f(x) = \max\{0, x\}$ ;
- `alpha` — a força do termo da regularização L2, cujo termo é dividido pelo tamanho da amostra quando adicionado à perda (*loss*);
- `solver = ‘adam’` — refere-se a um otimizador baseado em gradiente estocástico proposto por Kingma, Diederik, e Jimmy Ba;
- `cv = 5` — divide os dados em 5 partes consecutivas;
- `scoring = Mean Square Error (MSE)` — estratégia para avaliar o desempenho do modelo durante a validação cruzada aplicada ao conjunto de teste;
- `random_state` — determina geração de número aleatório para inicialização de pesos, viés e amostragem em batelada;
- `n_iter_no_change = 200` — número máximo de épocas para fins de tolerância;
- `tol = 10-4` — tolerância para a otimização. Quando a perda ou o *score* não estiver melhorando por ao menos  $10^{-4}$  durante 200 iterações consecutivas, a convergência é considerada alcançada e o treinamento encerra;
- `max_iter = 500` — número máximo de iterações. O otimizador itera até a convergência (determinada por ‘tol’) ou este número de iterações. Note que isso determina o número de épocas (quantas vezes cada ponto de dados é utilizado), não o número de passos do gradiente.

<sup>11</sup> cf. Apêndice A.6

<sup>12</sup> [https://scikit-learn.org/stable/modules/generated/sklearn.neural\\_network.MLPRegressor.html](https://scikit-learn.org/stable/modules/generated/sklearn.neural_network.MLPRegressor.html)

### 2.6.7 ANFIS com *Fuzzy C-Means Clustering* (ANFIS-FCM)

A Clusterização C-Médias<sup>13</sup> é um método que permite a cada ponto amostral pertencer a *clusters* múltiplos, cada um com seu grau de pertinência. O algoritmo computa os centros dos *clusters* e valores de pertinência de modo a minimizar a função objetivo expressa na Equação 2.9. O MatLab<sup>14</sup> fornece esta função com valores a serem personalizados:

$$J = \sum_{i=1}^n \sum_{j=1}^c w_{ij}^m \|x_i - c_j\|^2, \text{ em que } w_{ij} = \frac{1}{\sum_{k=1}^c \left( \frac{\|x_i - c_j\|}{\|x_i - c_k\|} \right)^{2/(m-1)}} \quad (2.9)$$

- $n$  é o número de amostras;
- $c$  é o número de *clusters*, que fica melhor igual a 5;
- $m > 1$  é o expoente da matriz de partição *fuzzy* para controlar o grau de interceptação nebulosa. Essa sobreposição se refere a quão nebulosas as fronteiras entre os *clusters* estão, i.e., o número de amostras que têm pertinência significativa em mais de um *cluster*;
- o tipo de inferência é o Sugeno;
- $w_{ij}$  é o grau de pertinência do  $j$ -ésimo ponto no  $i$ -ésimo *cluster*. Para um ponto dado, a soma dos valores de pertinência para todos os *clusters* é igual a 1.

### 2.6.8 ANFIS com *Subtractive Fuzzy Clustering* (ANFIS-SFC)

O método de clusterização subtrativa<sup>15</sup> é uma extensão do método proposto por Yager e Filev (1994). Este método presume que cada amostra é o centro de um *cluster* em potencial. O algoritmo executa a seguinte rotina:

1. Calcular o grau de verossimilhança que cada ponto definiria o centro de um *cluster*, baseado na densidade dos dados da vizinhança;
2. Escolher o ponto com o maior potencial como sendo o primeiro centro de *cluster*;
3. Remover todos os pontos amostrais próximos ao primeiro centro de *cluster*. A proximidade é determinada pelo parâmetro *clusterInfluenceRange*, que fica melhor igual a 0,2;

<sup>13</sup> cf. Apêndice A.7

<sup>14</sup> <https://www.mathworks.com/help/fuzzy/fcm.html>

<sup>15</sup> cf. Apêndice A.8

4. Selecionar o ponto remanescente com o maior potencial, como sendo o próximo centro de *cluster*;
5. Repetir os passos 3 e 4 até que todos os dados estejam dentro da faixa de influência de algum centro de *cluster*.

### 2.6.9 Suavização Exponencial

Segundo [Meira, Oliveira e Menezes \(2022\)](#), métodos de suavização exponencial<sup>16</sup> são aqueles que atribuem pesos exponencialmente decrescentes para dados passados, isto é, as observações recentes são dados pesos maiores na previsão do que a observações mais antigas. Apesar de que as especificações básicas datam dos trabalhos de [Holt \(2004\)](#) e [Winters \(1960\)](#), estes métodos permanecem amplamente aplicados, principalmente devido a sua simplicidade e transparência, mas também devido a sua adaptabilidade a mudanças na série temporal ([GOODWIN et al., 2010](#)). Adicionalmente, suavização exponencial tem uma fundação teórica em modelagem de espaço de estados, que permite implementações diretas em pacotes estatísticos.

### 2.6.10 *First Order Conventional* (CFTS)

O principal hiperparâmetro deste modelo é o número de partições. São comparados 30, 35, 40, 45, 50, 55 e 60. Para otimizar os hiperparâmetros deste modelo<sup>17</sup>, é utilizado o *AG* desenvolvido por [Lucas, Silva e Guimarães \(2019\)](#), implementado na biblioteca *PyFTS*<sup>18</sup>. Desenvolvido inicialmente por [Chen \(1996\)](#), é composto de sete passos principais:

Passo 1: Definir o universo de discurso e os intervalos de observações.

Passo 2: Definir conjuntos nebulosos para as observações.

Passo 3: Fuzzificar as observações.

Passo 4: Estabelecer as FLRs (relações lógicas nebulosas).

Passo 5: Estabelecer relações nebulosas.

Passo 6: Prever.

Passo 7: Defuzzificar.

---

<sup>16</sup> cf. Apêndice [A.9](#)

<sup>17</sup> cf. Apêndice [A.10](#)

<sup>18</sup> <https://pyfts.github.io/pyFTS/>

### 2.6.11 *First Order Weighted (WFTS)*

Este modelo<sup>19</sup> é uma modificação do modelo anterior, proposta por Yu (2005). O passo 5 nos modelos convencionais é revisto para lidar com a questão da recorrência. O passo 6 continua o mesmo. O passo 7 é revisto para processar a matriz defuzzificada. Enquanto que o passo 8 é adicionado para lidar com a atribuição de pesos. O passo 9 é adicionado para calcular os resultados finais.

Passo 5: Estabelecer relações nebulosas (revisto).

Passo 6: Prever.

Passo 7: Defuzzificar (revisto).

Passo 8: Atribuir pesos.

Passo 9: Calcular os resultados.

### 2.6.12 *Weighted High Order (WHOFTS)*

Silva (2019) propõe dois métodos: o HOFTS (série temporal nebulosa de alta ordem) e o *Weighted High Order Fuzzy Times Series (WHOFTS)*<sup>20</sup>, que são métodos consensuais de séries temporais nebulosas. De acordo com o autor, o objetivo dessas propostas é facilitar a medição das características e precisão das séries temporais nebulosas. Sendo assim, o autor faz uma escolha de utilizar um método que representa o estado da arte da literatura de *Fuzzy Time Series (FTS)*, ao invés de usar vários métodos com pequenas melhorias.

De fato, é a partir do modelo de Chen (1996), que se desenvolvem novos modelos, que dominam o campo atualmente. Além de seu trabalho original de primeira ordem, muitas extensões são propostas que modificam vários aspectos do método, mudando a ordem, método de particionamento, fuzzificação, defuzzificação, introdução de transformações, etc.

Os métodos propostos incorporaram todos os hiperparâmetros explorados, mas a definição de seus melhores valores envolve mais métodos complexos de otimização.

Uma desvantagem comum dos modelos baseados em regras é que todos os conjuntos nebulosos no lado consequente das regras têm a mesma importância. Para corrigir isso, é comum adicionar pesos que indicam sua relevância na fase de defuzzificação.

---

<sup>19</sup> cf. Apêndice A.11

<sup>20</sup> cf. Apêndice A.12

### 2.6.13 High Order Probabilistic Weighted (PWFTS)

Nessa mesma tese, [Silva \(2019\)](#) implementa o *Probabilistic Weighted Fuzzy Times Series (PWFTS)*<sup>21</sup>, um novo método FTS para previsão pontual, intervalar e probabilística de um a muitos passos à frente. O método PWFTS intenciona produzir previsões lidando com duas fontes de incerteza: medições nebulosas e comportamento estocástico. A fuzzificação é induzida para fins de redução de dados, em um processo que lembra uma simples discretização. O comportamento estocástico é deduzido pela abordagem frequentista sobre a fuzzificação anterior.

## 2.7 Rede Neo-Fuzzy-Neuron (NFN)

O objetivo desta Seção é tratar do modelo NFN, que é uma arquitetura de rede não linear proposta por [Yamakawa \(1992\)](#), cuja estrutura é mostrada na Figura 2.4, em que as entradas são  $x \in \mathbb{R}^n$  e a saída é  $y \in \mathbb{R}$ , dada pela Equação 2.10.

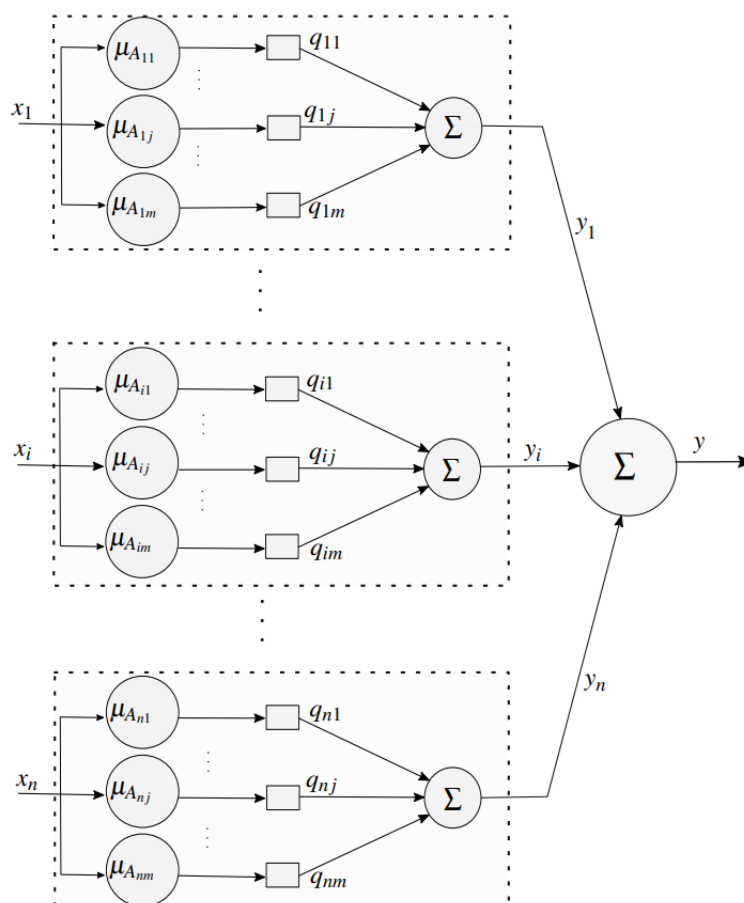


Figura 2.4 – Estrutura da rede NFN.

Fonte: [Souza et al. \(2020\)](#).

<sup>21</sup> cf. Apêndice A.13

$$y = f_1(x_1) + \dots + f_n(x_n) = \sum_{i=1}^n f_i(x_i) = \sum_{i=1}^n y_i(x_i). \quad (2.10)$$

Os valores  $y_i$  são calculados por meio deste conjunto de regras<sup>22</sup> nebulosas se-então:

$R_i^1$ : Se  $x_i$  é  $A_{i1}$ , então  $y_i$  é  $w_{i1}$ ;

$R_i^2$ : Se  $x_i$  é  $A_{i2}$ , então  $y_i$  é  $w_{i2}$ ;

⋮

$R_i^m$ : Se  $x_i$  é  $A_{im}$ , então  $y_i$  é  $w_{im}$ , em que  $A_{ij}$  são conjuntos nebulosos com funções de pertinência  $\mu_{ij}(x_i)$ . Assumindo funções complementares como ilustrado na Figura 2.5, o valor de  $f_i(x_i)$  é simplificado como na Equação 2.11.

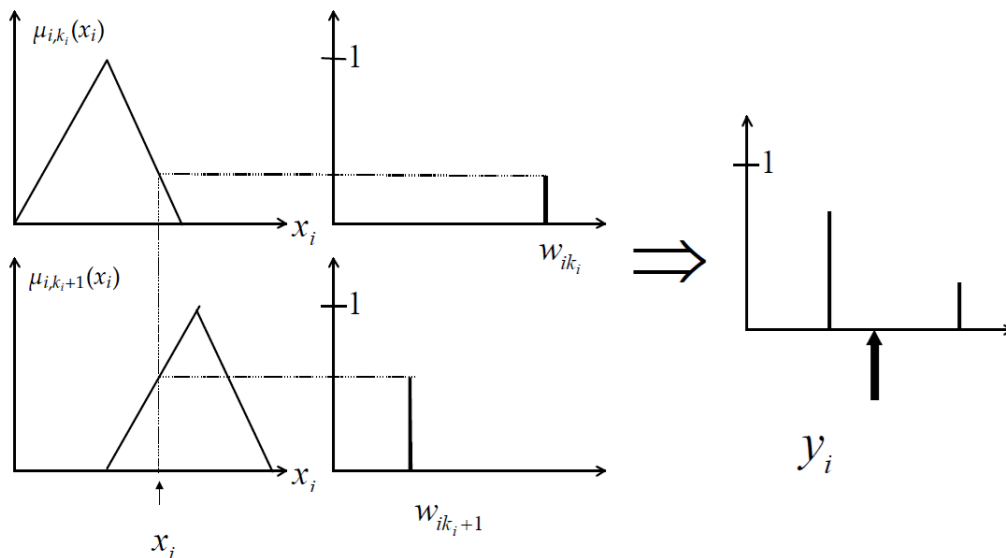


Figura 2.5 – Esquema de processamento do NFN.

Fonte: Caminhas e Gomide (2000).

$$f_i(x_i) = \frac{\sum_{j=1}^m \mu_{ij}(x_i) w_{ij}}{\sum_{j=1}^m \mu_{ij}(x_i)} = \frac{\mu_{i,k_i}(x_i) w_{i,k_i} + \mu_{i,k_i+1}(x_i) w_{i,k_i+1}}{\underbrace{\mu_{i,k_i}(x_i) + \mu_{i,k_i+1}(x_i)}_1}. \quad (2.11)$$

$$k_i = \min_j \{x_i - a_{i,j}\} \quad (2.12)$$

<sup>22</sup> É construído um conjunto de  $n$  regras fuzzy Takagi-Sugeno (TS) de ordem zero, acopladas às entradas e à saída.

As funções de pertinência triangulares uniformes são geradas a partir da quantidade de regras  $m$  desejada e do valor mínimo e máximo de  $x_i$ , sendo  $\min(x_i)$  o primeiro valor modal ( $a_{i1}$ ) e os demais são definidos por:

$$\Delta_i = \frac{\max(x_i) - \min(x_i)}{m - 1}, \quad (2.13)$$

$$a_{ij} = \min(x_i) + (j - 1)\Delta_i, \quad (2.14)$$

em que  $j$  é o índice da função de pertinência e  $\Delta_i$  o espaço entre elas, que as mantêm igualmente espaçadas, como ilustra a Figura 2.6. Para uma dada entrada  $x_i$ , exceto no máximo duas das  $m$  regras, indexadas por  $k_i$  e  $k_i + 1$ , todas as outras têm seu valor nulo. O valor do primeiro índice é dado pela Equação 2.12.

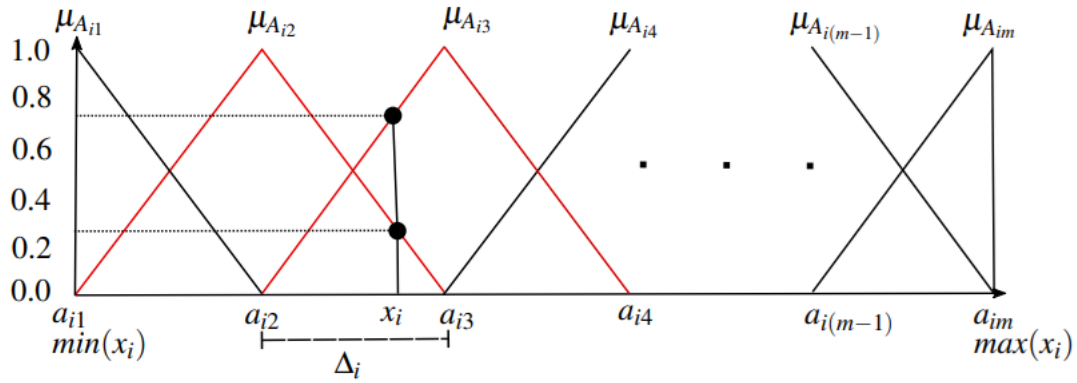


Figura 2.6 – Funções de pertinência triangulares igualmente espaçadas.

Fonte: (SOUZA et al., 2020).

De fato, a velocidade de processamento do NFN é maior, quando comparado com outros modelos de redes neurais, justamente porque apenas os pesos dos neurônios ativos precisam ser atualizados durante o aprendizado do NFN. (CAMINHAS; GOMIDE, 2000), em que os autores também comparam a complexidade computacional do NFN com a de uma RNA.

Nas Seções a seguir, será proposta uma personalização do NFN, a partir de uma otimização de uma métrica de erro.

Sejam  $(x_{t1}, \dots, x_{tn})$  os valores reais de entrada da  $t$ -ésima amostra. Se  $y_t^d$  é o valor de saída desejado e  $y_t$  a saída calculada, o MSE é dado pela soma sobre  $t$  de  $E_t$  na Equação 2.15.

$$E_t = \frac{1}{2}(y_t - y_t^d)^2 = E_t(w_{ij}) = E_t(\tilde{w}), \quad (2.15)$$

$$\tilde{w} = (w_{1,k_1}, w_{1,k_1+1}, \dots, w_{i,k_i}, w_{i,k_i+1}, \dots, w_{n,k_n}, w_{n,k_n+1}). \quad (2.16)$$

Repara-se que o **MSE** está em função quadrática de  $2n$  variáveis: os pesos. Logo, o mínimo local é também o global. Tais pesos são ajustados pelo método do gradiente, que possui uma taxa de aprendizado  $\alpha$ .

$$w_{i,k_i}^{j+1} = w_{i,k_i}^j - \alpha^j \frac{\partial E_t(w_{i,k_i}^j)}{\partial w_{i,k_i}^j}, \quad (2.17)$$

$$w_{i,k_i}^{j+1} = w_{i,k_i}^j - \alpha^j (y_t^j - y_t^d) \mu_{i,k_i}(x_{ti}). \quad (2.18)$$

$$w_{i,k_i+1}^{j+1} = w_{i,k_i+1}^j - \alpha^j (y_t^j - y_t^d) \mu_{i,k_i+1}(x_{ti}). \quad (2.19)$$

A soma dos erros é uma função não linear e sua convergência à solução ótima depende do ponto inicial (começando da origem), da direção de busca, da taxa de aprendizagem  $\alpha$  multiplicada por um percentual  $p_1$ . Portanto, os pesos são atualizados pelas Equações 2.18 e 2.19, em que  $j$  indexa épocas.

A taxa  $\alpha$  pode ser encontrada de forma numérica utilizando o método de busca unidimensional (BAZARAA; SHERALI; SHETTY, 1993). Neste trabalho, adota-se uma fórmula algebricamente fechada desenvolvida (CAMINHAS; GOMIDE, 2000) para alcançar **MSE** zero a cada passo de aprendizagem, que é dado pela Equação 2.20.

$$\alpha = \frac{p_1}{n \sum_{i=1} \mu_{i,k_i}(x_{ti})^2 + \mu_{i,k_i+1}(x_{ti})^2} \quad (2.20)$$

### 2.7.1 Funções Triangulares

Sejam  $(A, B, C) = (a_{ij}, a_{i,j+1}, a_{i,j+2})$ . Uma função  $\mu_{ij} : \mathbb{R} \rightarrow \mathbb{R}$  é triangular, quando pode ser escrita como na Equação 2.21.

$$\mu_{ij}(x) = \begin{cases} 0, & \text{se } x \in (-\infty, A] \\ \frac{x - A}{B - A}, & \text{se } x \in [A, B] \\ \frac{C - x}{C - B}, & \text{se } x \in [B, C] \\ 0, & \text{se } x \in [C, \infty) \end{cases}, \text{ cuja parte crescente é simplificada para:}$$

$\mu' : \mathbb{N} \rightarrow \mathbb{R}$

$$\mu'(k_i) = \begin{cases} 0, & \text{se } x \in (-\infty, a_{i,k_i}] \\ \frac{x - a_{i,k_i}}{a_{i,k_{i+1}} - a_{i,k_i}}, & \text{se } x \in [a_{i,k_i}, a_{i,k_{i+1}}] \\ 1, & \text{se } x \in [a_{i,k_{i+1}}, \infty) \end{cases} \quad (2.21)$$

### 2.7.2 Implementação do Modelo de Yamakawa

São implementados *fit* e *predict*, os quais, normalmente, fazem o modelo “se encaixar” nos melhores parâmetros, para depois predizerem com base nesses parâmetros.

*Fit* — é uma função executada quando o modelo é inicializado. Dadas as entradas  $x$ , e as saídas  $y$ , determina-se  $n =$  número de entradas ou variáveis;  $x_i$  inicial e  $x_i$  final, para cada  $i$ ; as amostras são permutadas aleatoriamente; a matriz de pesos  $w_{i,j}$  é inicializada com 1, para cada variável  $i$ ,  $1 \leq i \leq n$ , e cada função de pertinência  $j$ ,  $1 \leq j \leq m$ .

A matriz de valores modais  $a_{i,j}$  é inicializada com valores igualmente espaçados começando em  $x_i$  inicial e terminando em  $x_i$  final, para cada  $i$ .

Para cada época:

Para cada amostra  $k$ :

Para cada variável:

Encontrar  $k_i$  (pela Equação 2.12);

Determinar  $\mu'(k_i)$  (pela Equação 2.21);

Fim Para

Obter o  $\alpha$  ótimo (pela Equação 2.20);

Obter (Equação 2.11) o  $\hat{y} =$  saída emulada;

(cujo erro  $= y^d[k] - y[k]$  é minimizado, em módulo);

Para cada variável:

Cada peso  $w_{i,k_i}$  é atualizado com um percentual de  $\alpha$  ótimo vezes  $y[k] - y^d[k]$  vezes  $\mu'(k_i)$  (Equação 2.18);

Cada peso  $w_{i,k_{i+1}}$  é atualizado com um percentual de  $\alpha$  ótimo vezes  $\hat{y}[k] - y^d[k]$  vezes  $1 - \mu'(k_i)$  (Equação 2.19);

Fim Para

Fim Para

Fim Para

*Predict* — é uma função que utiliza um modelo para fazer uma predição. Com adaptação, esta função é praticamente idêntica à função *fit*, posta de lado apenas a inicialização.

### 2.7.3 Funções de Pertinência

As funções de pertinência expressam-se em um vetor  $m \in \mathbb{N}^n$ : para cada  $i$ ,  $1 \leq i \leq n$ , existe um  $m_i \in \mathbb{N}$ . Durante o trabalho, são feitos experimentos com diversos valores de funções de pertinência. Todavia, verifica-se que o número ideal de funções de pertinência são cinco, por ser aquele que minimiza o erro de validação.

### 2.7.4 Épocas

O número de épocas diz quantas varreduras são feitas no total de amostras. O ideal é uma única época, na atualização em tempo real. Durante o treinamento, são feitos experimentos com três épocas, no melhor caso.

### 2.7.5 Percentuais de *Fit* e *Predict*

São valores constantes, projetados para cada ponto de entrega, para cada horizonte 0, 1, 2, 3, 4, 5, 6 e 7 dias à frente, que são otimizados e mantidos fixos nos valores da Tabela 2.3. Valores para o algoritmo de Kohonen (1990) iguais a zero significam que mais preciso o modelo fica com funções de pertinência equidistantes. O valor de 0,236 no ponto B, por exemplo, significa que 23,6% do  $\alpha$ -ótimo foi a melhor constante encontrada.

São considerados valores para a série temporal  $v_{24}$ , isto é, o modelo  $M_0$ , para cada ponto de entrega, tanto do método do gradiente quanto do algoritmo de Kohonen (1990). A seguir, é feito o mesmo para a série temporal  $v_{11}$ , isto é, o modelo  $M_a$ . São considerados valores diferentes para *fit* e para *predict*, somente na previsão diária. Por fim, são exibidos acima os valores dos modelos  $M_1, M_2, \dots, M_7$ .

### 2.7.6 Modelo Invariante no Tempo vs Adaptativo

Diante das possibilidades, e dos diversos experimentos conduzidos, é necessário implementar um modelo adaptativo (parâmetros variantes no tempo), a fim de manter os parâmetros atualizados; porque a dinâmica do consumo evolui com o passar dos dias. Assim, um modelo invariante no tempo “caduca”, i.e., é depreciado, no sentido de que se fossem treinados novamente, os parâmetros ótimos seriam distintos.

### 2.7.7 Mínimos e Máximos

Como é feito um treinamento *online*, há chances de a entrada estar abaixo do mínimo ou acima do máximo. Para lidar com isso, implementamos a chamada

	Variável	Ponto A	Ponto B	Ponto C	Ponto D	Ponto E
<i>predict</i>	valores24	0,03	0,151	0,0174	0,0369	0,03
	valores24k	0,003	0,236	0,267	0,009	0
	valores11	0,03	0,138	0,13	0,0337	0,03
<i>fit</i>	valores11k	0,003	0,785	0,429	0	0
	fvalores24	0,0225	0,151	0,0174	0,0369	0,0445
	fvalores24k	0,011	0,236	0,267	0,009	0,064
Ambos	fvalores11	0,071	0,138	0,13	0,0337	0,139
	fvalores11k	0,009	0,785	0,429	0	0,795
	valores[0]	0,0684	0,115	0,0527	0,0669	0,153
	valores[1]	0,756	0,231	0,0576	0,115	0,186
	valores[2]	0,512	0,157	0,62	0,687	0,185
	valores[3]	0,59	0,137	0,0324	0,0668	0,156
	valores[4]	0,618	0,118	0,601	0,0858	0,175
	valores[5]	0,344	0,217	0,0292	0,0955	0,127
	valores[6]	0,676	0,14	0,833	0,035	0,0174
	valoresk[0]	0,07	0,869	0,847	0,011	0,799
	valoresk[1]	0,769	0,72	0,974	0,656	0,4
	valoresk[2]	0,389	0,879	0	0,81	0,687
	valoresk[3]	0,947	0	0,643	0,51	0,706
	valoresk[4]	0	0,854	0,493	0,469	0
	valoresk[5]	0,586	0	0,987	0,949	0
	valoresk[6]	0,443	0,826	0	0,013	0,083

Tabela 2.3 – Percentuais de *fit* e *predict*.

adaptação de contexto, conforme [Silva et al. \(2014\)](#), que mantém armazenados os intervalos [inicial, final] dinamicamente. Acrescentamos estas duas regras:

Se  $x_{ti} > x_i^{final}$ , então  $x_i^{final} = x_{ti}$ .

Se  $x_{ti} < x_i^{inicial}$ , então  $x_i^{inicial} = x_{ti}$ .

### 2.7.8 Espaçamento

Os valores modais, que no início são igualmente espaçados, também podem ser otimizados pelo chamado método de [Kohonen \(1990\)](#). Encontramos  $k_i = \min_j \{|x_{ti} - a^{ij}|\}$ , o índice da função de pertinência mais ativada por  $x_{ti}$ . Se  $2 \leq k_i \leq m - 1$ , então a  $k$ -ésima função de pertinência terá seu valor modal atualizado por um percentual  $p_2$  de  $\alpha$ -ótimo, da seguinte maneira:

$$a_{i,k_i} = a_{i,k_i} + p_2 \alpha (x_{ti} - a_{i,k_i}).$$

## 2.8 Métricas de Erro

Na legislação, costuma-se pagar multas sobre erro acima da faixa percentual de zero a 5%.

Por causa disso, as métricas de avaliação usadas são o Erro Percentual Absoluto (APE) e o MAPE (Equações 2.22 e 2.23) medidos para cada horizonte de previsão. Nessas equações,  $n$  é o número de amostras de teste,  $y$  são os valores reais e  $\hat{y}$  são os valores previstos pelos modelos.

Além disso, é calculada a acurácia dos modelos quando a previsão tem APE individual menor que 5%. Essa medida é empregada para verificar se os modelos avaliados geram previsões dentro da faixa aceitável de erro.

$$APE = \frac{|y - \hat{y}|}{y} \cdot 100\% \quad (2.22)$$

$$MAPE = \frac{1}{n} \sum_{i=1}^n APE_i \quad (2.23)$$

Neste capítulo, foi apresentado o panorama geral do referencial teórico, e mostra-se o modelo preferencial para conduzir a análise dos dados. Na Seção seguinte, será apresentada a metodologia.

### 3 METODOLOGIA

Neste capítulo, é apresentada a metodologia do trabalho.

Para dar sequência ao trabalho, é necessário apresentar a metodologia do trabalho. Primeiramente são apresentados os dados, as variáveis consideradas. Adiante é necessário explicar quais são as autocorrelações, bem como as entradas e as saídas. Por fim, mostram-se as parametrizações dos modelos.

O processamento das bases de dados evolui em nove etapas.

#### 3.1 Análise Preliminar dos Dados

O objetivo deste trabalho é avaliar o consumo de GN em curto prazo. Para tanto, são escolhidos cinco municípios, consumidores ou pontos de entrega (denominados de A, B, C, D e E), cujos dados fornecidos são milhares de amostras de três variáveis:

- data, apresentada no formato dia/mês/ano, com dois dígitos para o dia e o mês, e com quatro dígitos para o ano;
- vazão (consumo em metros cúbicos) de 0h às 10h59min, vista como entrada da previsão;
- vazão de 0h às 23h59min, vista como saída da previsão.

Esses são os dados iniciais. Combinou-se que eles seriam detalhados em etapas seguintes, uma vez que vazão é definida como consumo por unidade de tempo, sejam metros cúbicos por minuto, por exemplo.

Os dados são coletados entre 01/01/2018 e 23/01/2024, totalizando 2.214 dias. Formam um subconjunto do banco de dados real da GASMIG, por isso, como são sigilosos, é feita a normalização *Z-Score* (tomar todo o conjunto de dados, subtrair a média, em seguida dividir pelo desvio-padrão) somente na exibição, uma vez que todos os cálculos são realizados com os dados reais. Durante todo o trabalho, são analisados seis pontos de entrega, no entanto, um é excluído desta análise por conveniência. A seguir, é apresentado um panorama dos dados de cada localidade, considerando três variáveis.

### • Ponto A

O primeiro gráfico é  $11h \times 24h$  (Figura 3.1a) e evidencia uma dispersão com alta linearidade<sup>1</sup>. Já o segundo gráfico (Figura 3.2) apresenta as séries temporais no período de 01/01/2023 a 30/06/2023 (em preto), e 01/07/2023 a 31/12/2023 (em azul) no consumo de GN. As duas primeiras séries são geradas com o tratamento de dados pela média móvel diária (conforme será explicado na Seção 3.2), e as duas últimas séries são geradas tratando de maneiras diferentes para cada dia da semana (segunda-feira = 0, terça-feira = 1, quarta-feira = 2, quinta-feira = 3, sexta-feira = 4, sábado = 5 e domingo = 6).

Em todos os cinco pontos (A, B, C, D e E), ao final, são preferidas a terceira e a quarta séries temporais. Isto é, a dinâmica é melhor capturada com **sazonalidade** (com relação ao dia da semana) do que sem ela. Maiores detalhes na Seção 3.2.1.

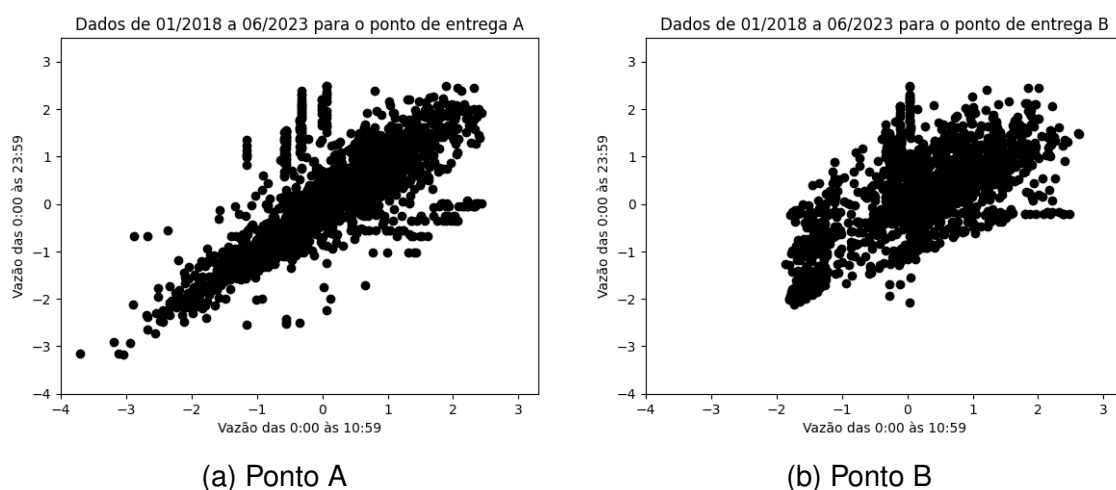


Figura 3.1 – Visão  $11h \times 24h$  para os pontos de entrega A e B.

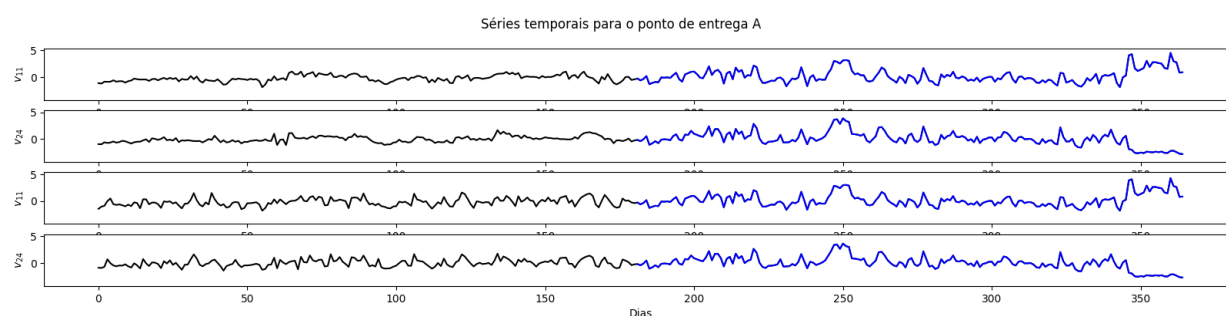


Figura 3.2 – Visão temporal para o ponto de entrega A.

<sup>1</sup> cf. Seção 2.6.1

**• Ponto B**

Considerando os desvios-padrão e os resíduos (cf. Seção 2.6.1), que são sigilosos, os gráficos evidenciarão dispersão máxima em B. Isso fica bem visualizado em gráficos 11h × 24h (Figura 3.1b) ou como séries temporais (Figura 3.3).

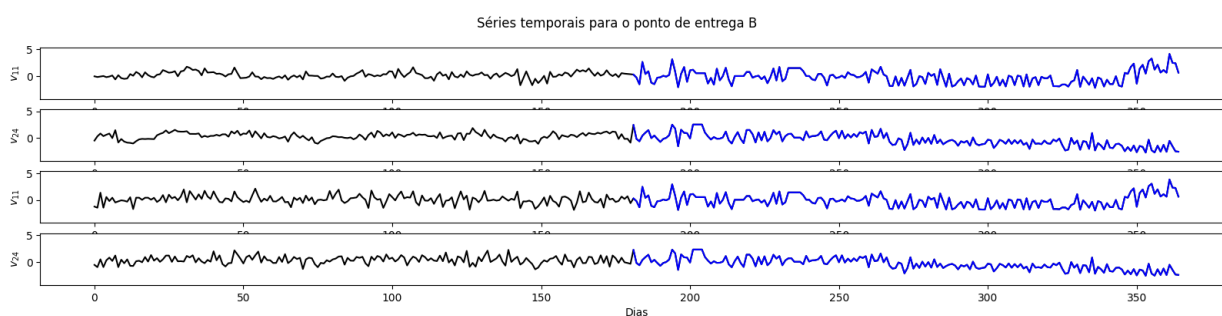


Figura 3.3 – Visão temporal para o ponto de entrega B.

**• Ponto C**

O desvio-padrão e/ou o resíduo (cf. Seção 2.6.1) em C é maior do que em A. Isso fica bem visualizado em gráficos 11h × 24h (Figura 3.4a) ou como séries temporais (Figura 3.5).

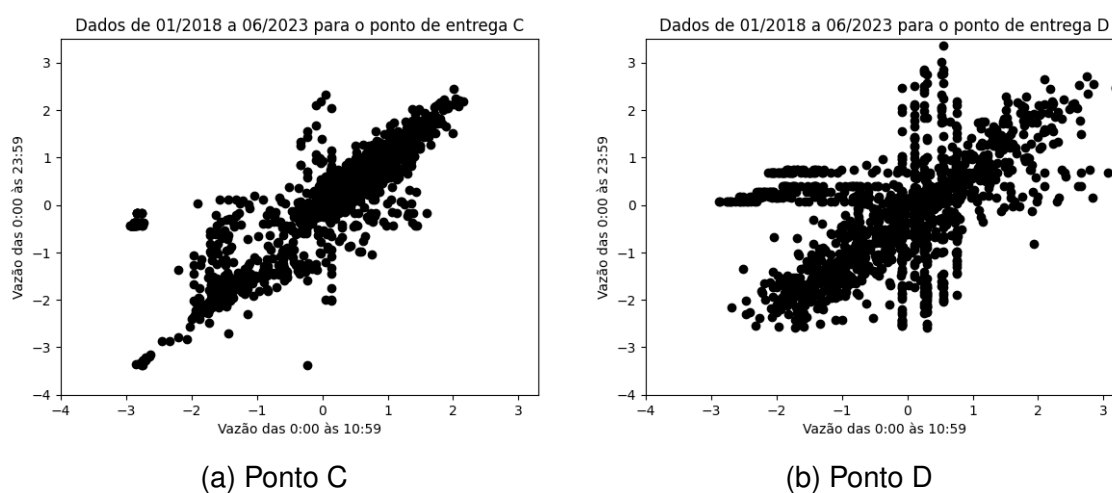


Figura 3.4 – Visão 11h × 24h para os pontos de entrega C e D.

**• Ponto D**

O desvio-padrão e/ou o resíduo (cf. Seção 2.6.1) em D é maior do que em C. Isso fica bem visualizado em gráficos 11h × 24h (Figura 3.4b) ou como séries temporais (Figura 3.6).

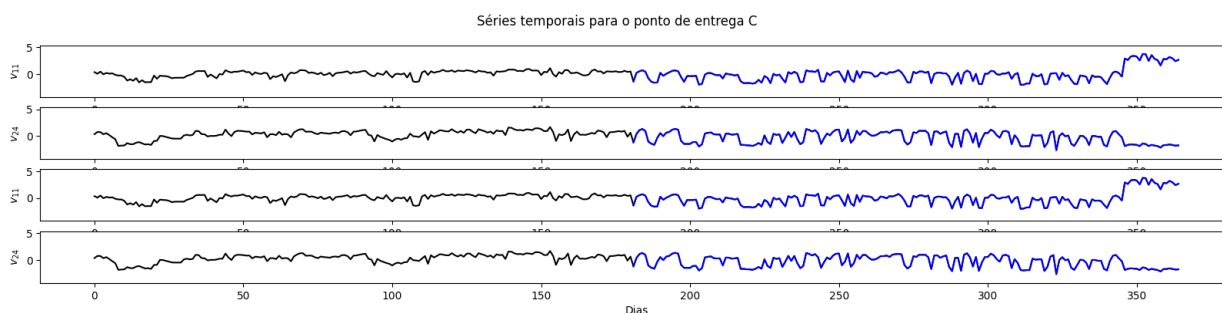


Figura 3.5 – Visão temporal para o ponto de entrega C.

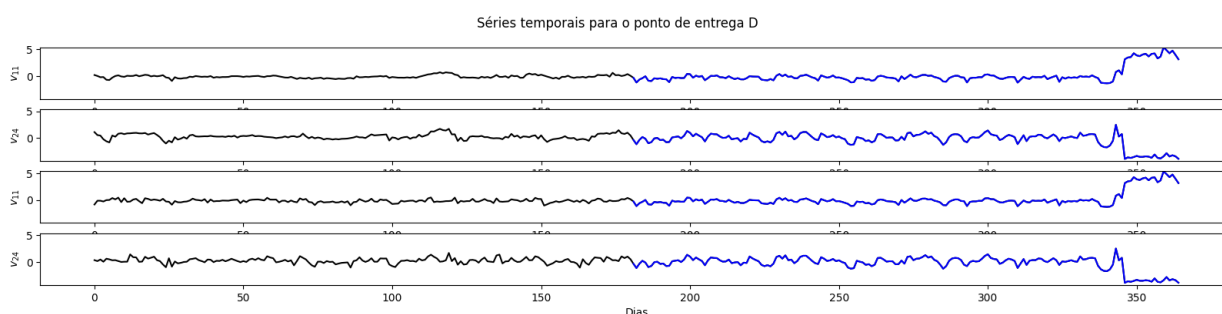
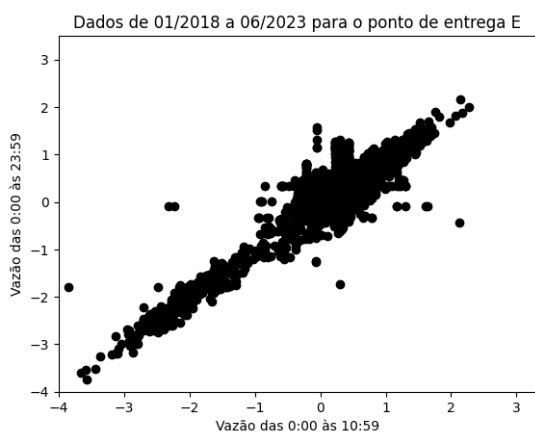


Figura 3.6 – Visão temporal para o ponto de entrega D.

#### • Ponto E

Considerando os desvios-padrão e os resíduos<sup>2</sup>, que são sigilosos, os gráficos evidenciariam dispersão mínima em E. Isso fica bem visualizado em gráficos  $11h \times 24h$  (Figura 3.7) ou como séries temporais (Figura 3.8). Maiores detalhes na Seção 3.2.1.

Figura 3.7 – Visão  $11h \times 24h$  para o ponto de entrega E.

<sup>2</sup> cf. Seção 2.6.1

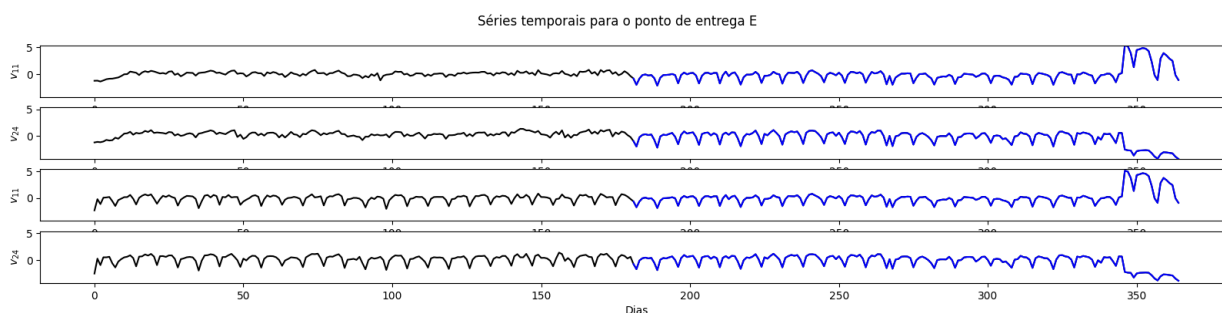


Figura 3.8 – Visão temporal para o ponto de entrega E.

## 3.2 Escolha das Variáveis

Um ponto crucial para o desenvolvimento deste trabalho é a decisão de quais seriam as variáveis utilizadas. Depois de várias considerações, as únicas variáveis utilizadas são as vazões de 0h às 10h59min, e a de 0h às 23h59min, além do dia da semana.

Disponibilizou-se também a data. Na etapa seguinte, é disponibilizada a leitura do consumo horário, e tornou-se possível tomar como dado o horário, o que configura 24 registros por dia. Mais adiante, a leitura é amostrada de 15 em 15 minutos, o que gerou 96 amostras por dia ( $4 \times 24$ ).

A análise inicia-se com a retirada dos *outliers*, que são os valores fora do intervalo  $[M - 3S, M + 3S]$ , em que  $M$  é a mediana e  $S$  é o desvio-padrão. Eles são substituídos pela média móvel. Coletam-se os sete valores de  $k - 4c$  até  $k + 3c$  (em que  $c = 1$  ou  $c = 7$ ) excluídos  $k$  e os NaN's<sup>3</sup>, divide-se a soma pela quantidade, desde que seja não nula. Para o caso de sete NaN's consecutivos, é utilizada a mediana global. Além disso, os dados inexistentes e os valores menores ou iguais a zero são substituídos pela média móvel.

### 3.2.1 A Sazonalidade no ponto de entrega E

Neste caso, na quarta etapa, o modelo atual se mostra pouco preciso no ponto de entrega E. Desse modo, foi necessário fazer uma adaptação considerando segunda-feira baseada em sexta-feira, e sábado em domingo, conforme ilustra a Figura 3.9. O que resultou em uma melhor acomodação e precisão dos dados.

<sup>3</sup> Not a Number

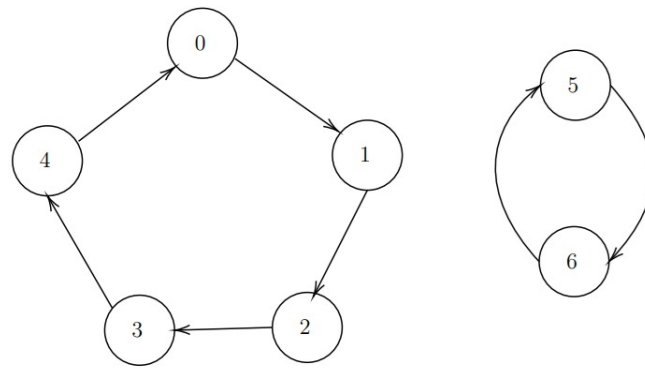


Figura 3.9 – Fluxo de dias da semana para o ponto de entrega E.  
Em que 5 = sábado e 6 = domingo.

### 3.3 Definição das Entradas e Saídas

A partir da etapa em que há amostras de 15 em 15 minutos, é possível escolher como agrupá-las. Quanto maior o número de variáveis, maior a complexidade, e piores ficariam os resultados.

Seja  $k$  a discretização do tempo contínuo em dias, isto é,  $k$  é o número de dias desde 01/01/2018. Para prever a saída  $\hat{v}_{24}[k]$ , algumas alternativas de entradas (que surgem naturalmente, por continuação da etapa inicial, em que é suficiente prever o consumo diário em função do consumo até as 10h59min — a empresa tem o costume de fazer suas previsões computacionais a partir das 11h da manhã de cada dia; e isso significa dizer que também constitui uma *baseline*) são:

- $v_{11}[k - 1], v_{11}[k - 2], \dots, v_{11}[k - a_1]$ , para algum  $a_1 \in \mathbb{N}$ ;
- incluir  $v_{11}[k]$ ;
- $v_{24}[k - 1], v_{24}[k - 2], \dots, v_{24}[k - a_2]$ , para algum  $a_2$ ;
- definindo a vazão vespertina  $v_{13} = v_{24} - v_{11}$ , incluir  $v_{13}[k - 1], v_{13}[k - 2], \dots, v_{13}[k - a_3]$ , para algum  $a_3$ ;
- ao invés disso, incluir  $v_{11}[k], v_{11}[k - 7], \dots, v_{11}[k - 7a_4]$ , para algum  $a_4$ ;
- incluir  $v_{24}[k - 7], v_{24}[k - 14], \dots, v_{24}[k - 7a_5]$ , para algum  $a_5$ ;
- incluir  $v_{13}[k - 7], v_{13}[k - 14], \dots, v_{13}[k - 7a_6]$ , para algum  $a_6$ .

A fim de determinar  $a_i$ , qualquer que seja  $i$ , uma das possibilidades é analisar as correlações entre as variáveis, e evitar o *overfitting*, que aparece à medida que é aumentada a complexidade dos modelos. Para encerrar essa escolha de entradas, são consideradas as visualizações — além das dispersões e das séries temporais — das autocorrelações e dos *boxplots*.

### 3.3.1 Autocorrelações

As Figuras 3.10 e 3.11 mostram a função de autocorrelação da vazão acumulada diária de cada ponto de distribuição. Observa-se que em todos os casos, não existe sazonalidade nos dados e que os valores passados mais recentes são mais significativos para prever a vazão acumulada diária.

As autocorrelações são calculadas por convolução, e é hachurado o intervalo de confiança. A fim de que não houvesse correlação, os valores deveriam estar dentro deste intervalo. Portanto, há alta correlação, que nos permitiria fazer um modelo considerando as entradas de zero a vinte dias atrás, o que seria um modelo complexo.

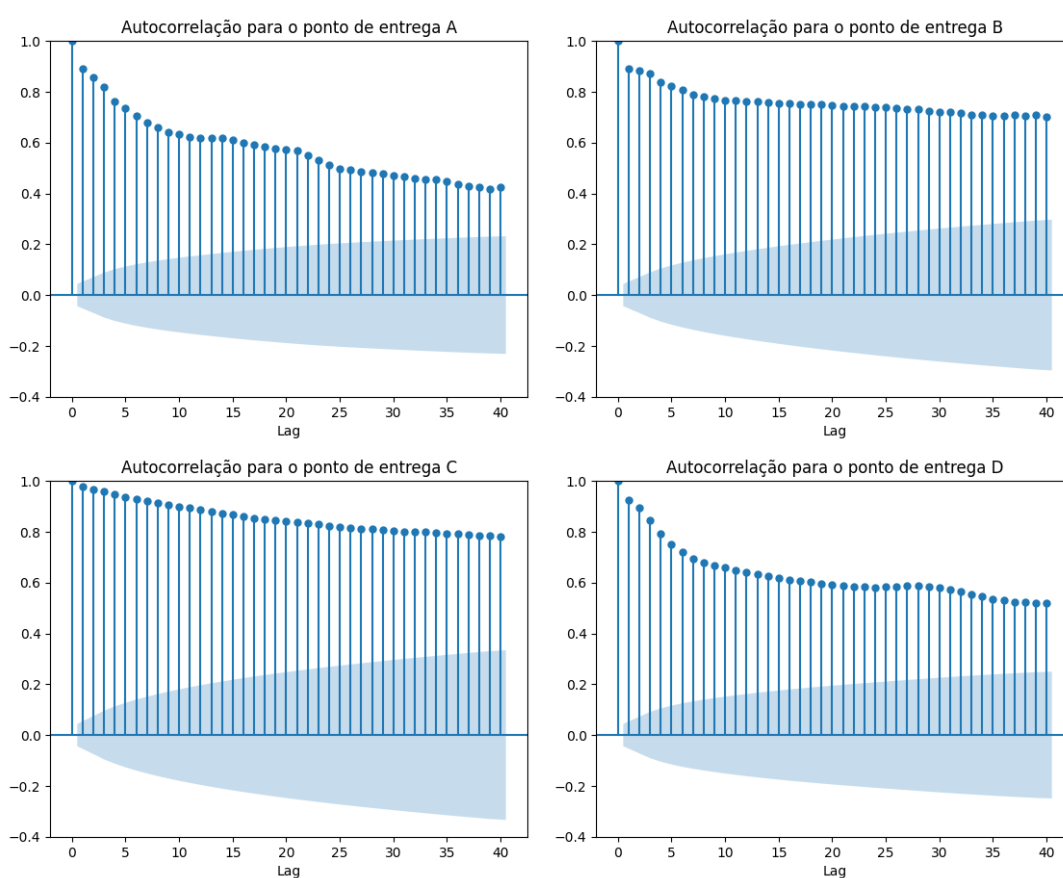


Figura 3.10 – Autocorrelações para os pontos de entrega A, B, C e D.  
No período 01/2018 a 06/2023.

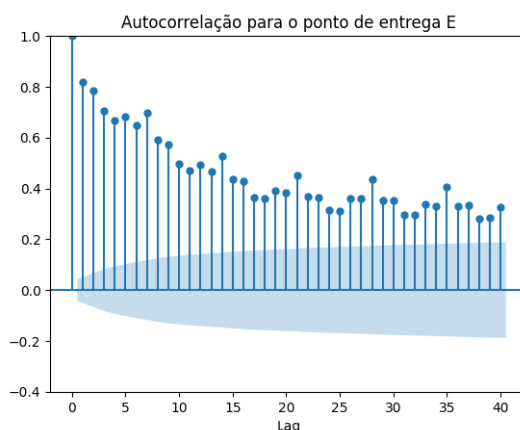


Figura 3.11 – Autocorrelações para o ponto de entrega E. No período 01/2018 a 06/2023.

### 3.3.2 Boxplots

Os *boxplots* das Figuras 3.12 e 3.13 apresentam a vazão acumulada diária agrupada por dia da semana. Observa-se que não há diferença no comportamento dos dados para diferentes dias da semana em todos os pontos de distribuição. Isso indica que não é necessário incluir um indicador de dia como variável explicativa para os modelos.

Essa sazonalidade começou a aparecer a partir de 01/07/2023, principalmente no ponto de entrega E.

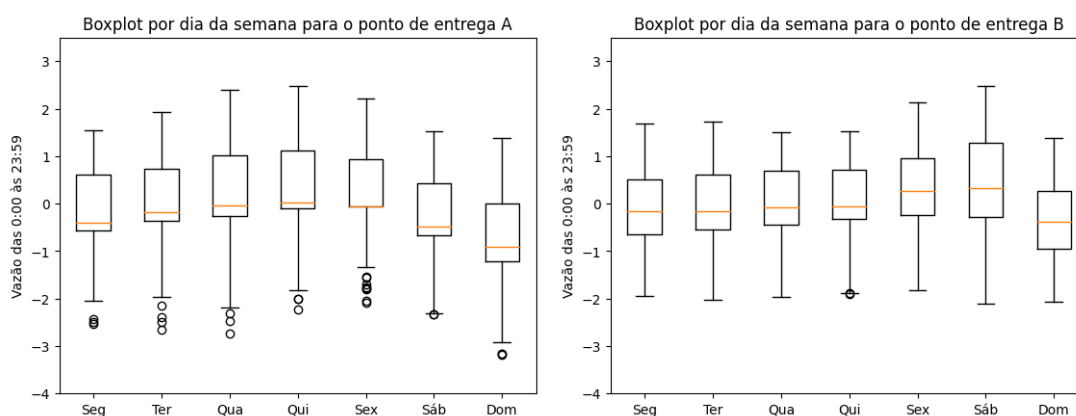


Figura 3.12 – *Boxplots* para os pontos de entrega A e B. No período 01/2018 a 06/2023.

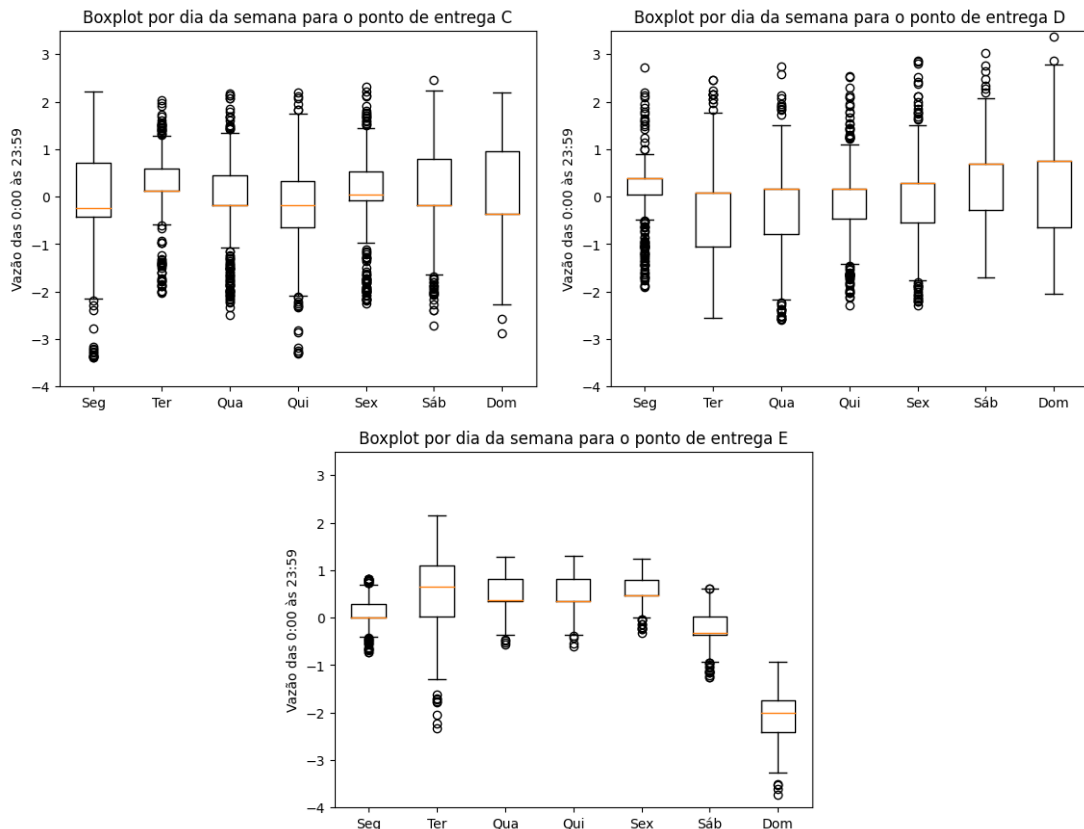


Figura 3.13 – *Boxplots* para os pontos de entrega C, D e E. No período 01/2018 a 06/2023.

### 3.3.3 O Critério de Akaike

Seja  $\theta \in \mathbb{R}^v$  uma matriz de regressores para  $\Psi \in \mathcal{M}_{h \times v}(\mathbb{R})$ , tal que:

$\hat{y} = \Psi\theta \Rightarrow \theta = (\Psi^\top \Psi)^{-1} \Psi^\top y$ . Então, define-se:

$$\text{AIC}(v) = 1 + \ln 2\pi + \ln \left[ \frac{1}{h} \sum_{i=1}^h \left( y_i - \sum_{j=1}^v \Psi_{ij} \theta_j \right)^2 \right] + \frac{2v}{h}.$$

Para selecionar a quantidade de variáveis de entrada, utiliza-se o critério de Akaike, descrito em [Aguirre \(2007\)](#). Tais variáveis são ordenadas daquela de maior importância à menos relevante. A seguir, calcula-se um número denominado *log-likelihood* ou representado pela sigla AIC (*Akaike Information Criterion*) em função das  $k$  primeiras variáveis e procura-se um ponto de mínimo, o que indicaria que acrescentar variáveis a partir daí seria piorar o modelo.

A Figura 3.14 é gerada fixando o modelo Yamakawa-NFN e começando com uma variável  $v_{11}[k]$  e acrescentando  $v_{11}[k - i]$ ,  $1 \leq i \leq 14$ . Cada vez que se incluía uma variável, calculava-se o AIC em função da nova base de dados. Verifica-se, por

exemplo que, para o ponto A, o AIC inicial e máximo é igual a 21,88 e o final e mínimo é de 21,78. Nota-se que, nos pontos de entrega C e E, o número ideal de variáveis é quatro, enquanto que no ponto de entrega D, o ideal são três. Porém, os gráficos dão a entender, à primeira vista, que nos pontos de entrega A e B, deveria-se utilizar até mesmo todas as 14 variáveis. No entanto, repara-se na taxa de variação em A:  $0,1/21,88 = 0,457\%$ ; em B:  $0,1/19,30 = 0,518\%$ ; em D:  $0,01/25,50 = 0,0392\%$ ; e em E:  $0,01/24,13 = 0,0414\%$ . Sendo a inclinação quase nula, isso nos permite escolher apenas uma variável.

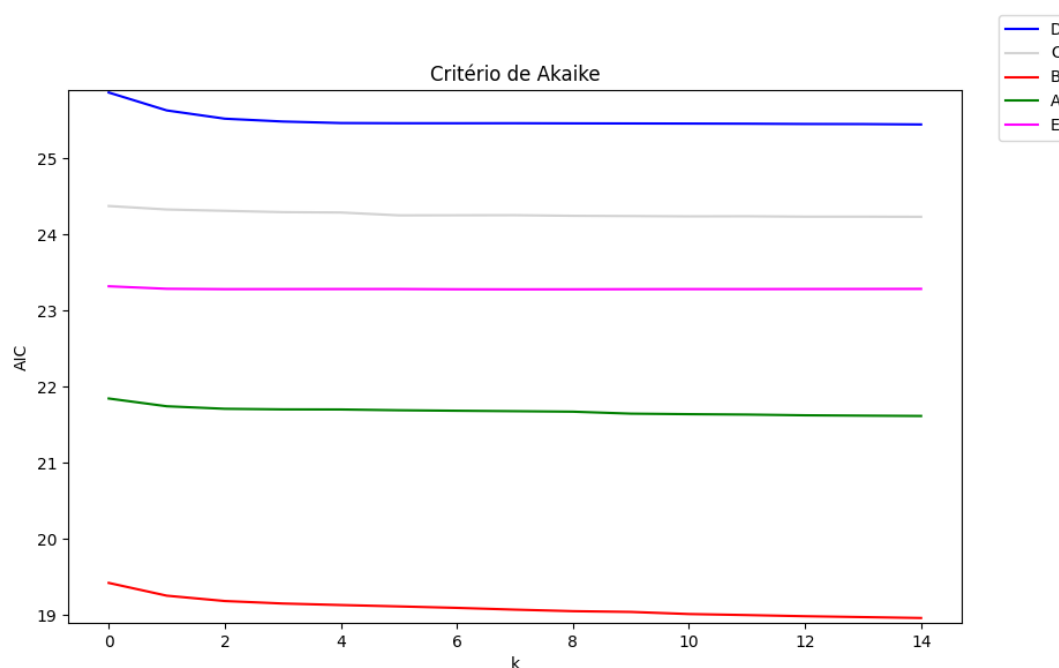


Figura 3.14 – AICs em função do número de variáveis para os cinco pontos de entrega. No período 01/2018 a 09/2022.

### 3.4 Modelo Proposto

A Figura 3.15 ilustra um modelo auxiliar  $M_a$  combinado com um modelo  $M_0$  para a previsão do consumo diário, bem como  $M_a$  combinado com  $M_n$  para previsão de  $n$  dias à frente, para cada  $n \in \{1, 2, \dots, 7\}$ .

Sejam  $v_{11}$  a vazão das 0h às 10h59min e  $v_{24}$  a vazão das 11h às 23h59min em um único dia. É definido um primeiro modelo Yamakawa de única entrada  $v_{11}[k]$  e única saída  $\hat{v}_{24}[k]$ . E um modelo auxiliar Yamakawa de única entrada  $v_{11}[k]$  e única saída  $\hat{v}_{11}[k+1]$ . Ambos são agrupados num terceiro e principal modelo UKF de 2 estados. Durante o treinamento, determinam-se os melhores pesos  $w_i$  que aproximam a função real  $v_{24} = f(v_{11})$ , não necessariamente linear.

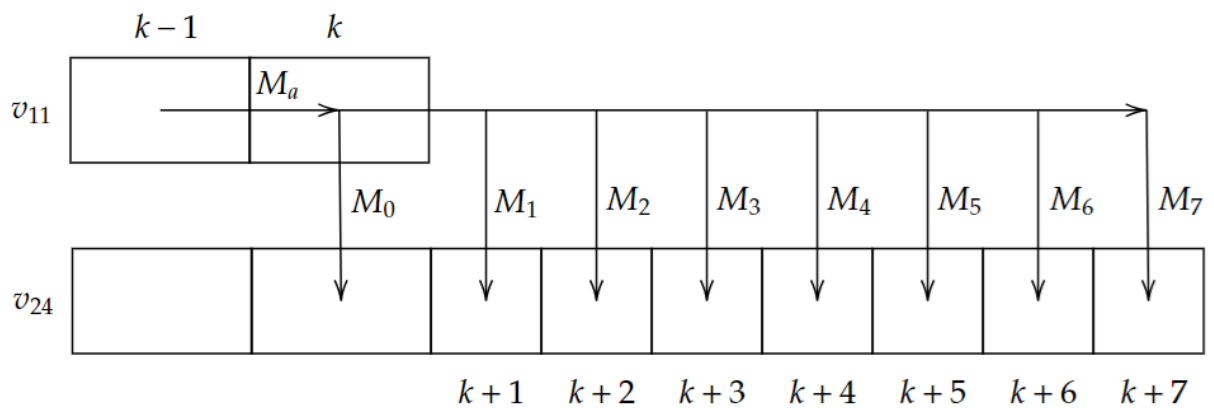


Figura 3.15 – Modelos Yamakawa-NFN combinados com modelos *Unscented Kalman Filter*.

Seis pontos de entrega deram origem a seis modelos **NFN**. São incluídos outros seis cuja entrada é  $v_{11}[k-1]$  e cuja saída é  $\hat{v}_{11}[k]$ . Para prever de um a sete dias à frente, são acrescentados modelos cuja entrada é  $v_{11}[k]$  e as saídas são  $\hat{v}_{24}[k+1]$ ,  $\hat{v}_{24}[k+2]$ ,  $\dots$ ,  $\hat{v}_{24}[k+7]$  (sem subtrair os valores da manhã, cf Seção 3.3), totalizando 54 modelos Yamakawa **NFN**, cada um com suas funções *anfis\_yamakawa* em Python, i.e., *fit* e *predict*. Todos os modelos são independentes entre si e todas as previsões são especialistas.

A validação é feita comparando com valores rotulados pela **GASMIG** e também pelo percentual de acerto. Consideramos que um acerto é uma saída com acurácia menor ou igual a 5% do valor desejado, pela métrica do **APE**.

Um dos pontos de entrega, para ficar melhor que o rotulado, é dividido em dois modelos, um para a faixa de segunda a sexta-feira; e outro para sábado  $\rightarrow$  domingo  $\rightarrow$  sábado seguinte. Ademais, na previsão semanal, é feita uma implementação intervalar, descrita na próxima Seção.

Para cada ponto de entrega, são comparados treinamento *online* ou em batelada. Da mesma forma, são feitos experimentos sem **UKF** versus com **UKF**. Todos os dias faz-se necessário entrar com os dados  $v_{24}$  do dia anterior e  $v_{11}$  do dia atual. Nesse momento, os novos modelos são salvos em arquivo.

No caso do modelo de  $n$  dias à frente, são adaptados os modelos referentes à saída  $v_{24}[k-1]$  em função da entrada  $v_{11}[k-1-n]$ . (Nesse caso, é adaptado também o modelo **UKF**.) Também são adaptados os modelos cuja saída é  $v_{11}[k]$  e cuja entrada é  $v_{11}[k-1]$ .

### 3.4.1 Modelo Intervalar

A figura 3.16 mostra um modelo Yamakawa NFN com uma única entrada escalar, uma única saída e cinco funções de pertinência triangulares, o que implica que somente duas regras ficam ativas de cada vez.

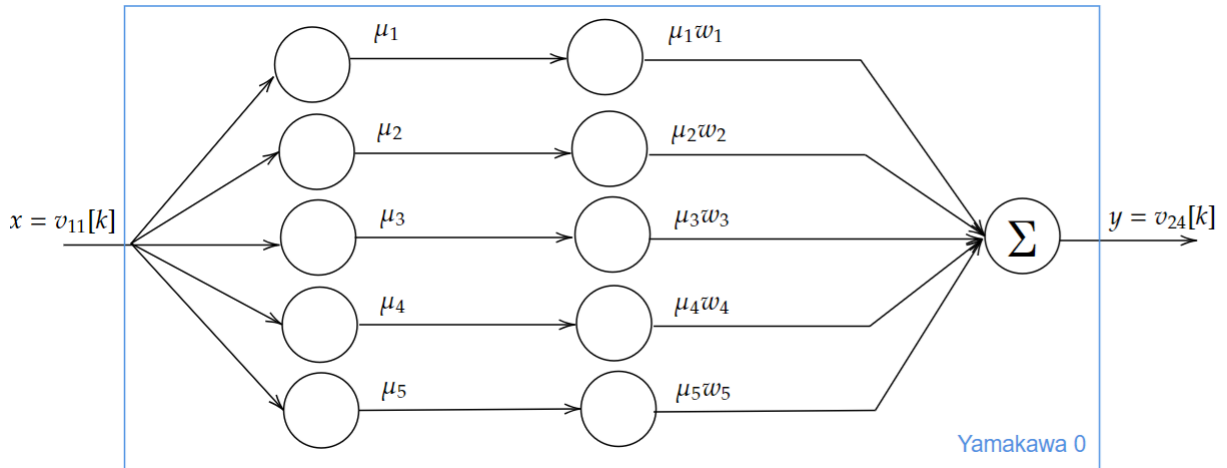


Figura 3.16 – Modelo Yamakawa-NFN, cuja entrada é a vazão de Gás Natural de 0h às 10h59min e cuja saída é a vazão das 11h às 23h59min do mesmo dia.

Sejam cinco conjuntos *fuzzy*  $A_i$ . Por exemplo, se  $v_{11}$  é  $A_1$  a 70% e  $v_{11}$  é  $A_2$  a 30%, então  $\hat{v}_{24}[k] = 0,7w_1 + 0,3w_2$ . Nossa proposta é dividir a faixa  $v_{11} \in [\text{mínimo}, \text{máximo}]$  em cinco intervalos de comprimentos iguais.

Quatro valores modais dividem esse intervalo. Os ideais não são equidistantes, e para otimizá-los implementamos o chamado método de *Kohonen* (1990), que utiliza outro percentual de  $\alpha$ -ótimo.

A validação é feita de forma análoga à Seção anterior.

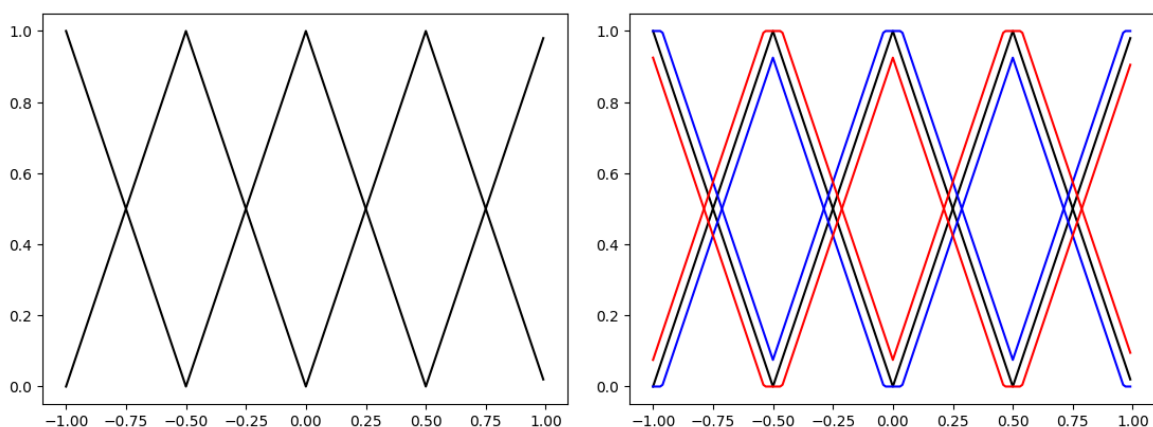


Figura 3.17 – Funções de Pertinência Pontual e Intervalar.

Na previsão semanal, a fim de estimar intervalos de incerteza, são definidos conjuntos *fuzzy* tipo 2, considerando  $[\mu_i - 0,075, \mu_i + 0,075]$  (Figura 3.17) e respeitando o mínimo de 0 e o máximo de 1 = 100%. Assim, são propagados para a fase seguinte  $[\underline{w}_i, \overline{w}_i]$  e, na saída,  $[\underline{\hat{y}}^i, \overline{\hat{y}}^i]$ . Finalmente, a previsão pontual passa a ser o ponto médio deste intervalo.

## 4 EXPERIMENTOS

Neste capítulo são apresentadas e descritas cada etapa de experimento. No total são nove, e cada uma compõe uma Seção do capítulo.

### 4.1 Etapa Inicial

Nesta etapa, são considerados dados de 01/01/2018 a 09/09/2022. O problema de estimativa de vazão é modelado de duas formas: (A) com o uso dos valores de vazão das 0h às 10h59min para estimar os valores de vazão das 11h às 23h59min e (B) com o uso dos valores passados para prever a vazão do dia seguinte.

Para efeito de comparação, define-se o modelo *baseline* como previsão do consumo de 0h às 23h59min com base no consumo de 0 às 10h59min, considerando-se um consumo uniforme durante o dia, o que equivale a uma regra de três simples:

$$\hat{v}_{24}[k] = \frac{24}{11} \times v_{11}[k].$$

Em ambas as abordagens, A e B, os modelos são treinados com 4 anos de dados (2018, 2019, 2020 e 2021) e avaliados com 252 registros de 2022. Nessa metodologia, todos os modelos também apresentam, no geral, melhora em relação ao modelo *baseline*: uma quantidade maior de dias abaixo de 5% do **APE**, em que não há multa, e um número menor de dias acima de 20%, que é a maior faixa de multa.

#### 4.1.1 Previsão Vespertina

As Tabelas 4.1, 4.2, 4.3, 4.4 e 4.5 — em que estão grifadas em azul as linhas que mostram as melhores previsões (neste caso, o maior valor para a faixa  $[0, 5)\%$  não gera empate) — evidenciam os dados e os tratamentos possíveis. A previsão é entendida como um problema de regressão com uma única entrada e uma única saída. São aplicados os modelos **RF**, **RL**, **MLP**, **ANFIS** com **SFC**, **ANFIS** com **FCM** e Yamakawa **NFN**. Repara-se que o **APE** é definido na Seção 2.8.

Modelo	APE				
	[0,5) %	[5,10)%	[10,15)%	[15,20)%	>20%
<i>baseline</i>	98	69	41	22	22
RF	106	59	49	23	15
RL	<b>108</b>	<b>65</b>	<b>39</b>	<b>26</b>	<b>14</b>
MLP	103	68	42	18	21
ANFIS-SFC	91	63	39	21	10
ANFIS-FCM	95	59	37	22	11
Yamakawa-NFN	106	68	42	23	13

Tabela 4.1 – Resultados do APE usando a abordagem Vespertina no ponto de distribuição A.

Modelo	APE				
	[0,5) %	[5,10)%	[10,15)%	[15,20)%	>20%
<i>baseline</i>	33	35	29	28	127
RF	42	41	41	44	84
RL	52	49	40	28	83
MLP	38	36	36	34	108
ANFIS-SFC	<b>56</b>	<b>26</b>	<b>36</b>	<b>40</b>	<b>66</b>
ANFIS-FCM	47	42	38	24	73
Yamakawa-NFN	57	36	44	35	80

Tabela 4.2 – Resultados do APE usando a abordagem Vespertina no ponto de distribuição B.

Modelo	APE				
	[0,5) %	[5,10)%	[10,15)%	[15,20)%	>20%
<i>baseline</i>	111	46	29	13	53
RF	148	45	22	9	28
RL	164	34	23	6	23
MLP	162	35	24	7	24
ANFIS-SFC	145	31	21	4	23
ANFIS-FCM	144	30	21	6	23
Yamakawa-NFN	<b>170</b>	<b>31</b>	<b>22</b>	<b>4</b>	<b>25</b>

Tabela 4.3 – Resultados do APE usando a abordagem Vespertina no ponto de distribuição C.

Modelo	APE				
	[0,5) %	[5,10)%	[10,15)%	[15,20)%	>20%
<i>baseline</i>	158	53	27	8	6
RF	174	61	9	5	3
RL	179	54	12	2	5
MLP	177	56	10	5	4
ANFIS-SFC	160	49	9	2	4
ANFIS-FCM	159	50	9	2	4
Yamakawa-NFN	<b>183</b>	<b>51</b>	<b>11</b>	<b>2</b>	<b>5</b>

Tabela 4.4 – Resultados do APE usando a abordagem Vespertina no ponto de distribuição D.

Modelo	APE				
	[0,5) %	[5,10)%	[10,15)%	[15,20)%	>20%
<i>baseline</i>	169	64	17	1	1
RF	198	43	8	2	1
RL	<b>200</b>	<b>38</b>	<b>10</b>	<b>4</b>	<b>0</b>
MLP	192	44	11	4	1
ANFIS-SFC	177	36	9	2	0
ANFIS-FCM	177	36	8	3	0
Yamakawa-NFN	199	38	12	3	0

Tabela 4.5 – Resultados do APE usando a abordagem Vespertina no ponto de distribuição E.

#### 4.1.2 Previsão de Um Passo à Frente

As Tabelas 4.6, 4.7, 4.8, 4.9 e 4.10 evidenciam os dados e os tratamentos possíveis. É avaliado o desempenho dos modelos ARIMA, Suavização Exponencial<sup>1</sup>, Média Móvel, RF (com diversas parametrizações), CFTS, WFTS, WHOFTS e PWFTS, em que é utilizada a variável vazão acumulada diária, todos eles comparados com o modelo *baseline*.

Também são comparados os modelos nebulosos ANFIS-SFC, ANFIS-FCM, Yamakawa-NFN, ANFIS-SFC, ANFIS-FCM versus Yamakawa-NFN, em que são utilizadas a vazão acumulada das 0h às 10h59min e a vazão acumulada diária.

Adiante, são comparados os modelos Yamakawa-NFN versus RF, utilizando apenas a vazão acumulada das 0h às 10h59min. Por fim, é incluído na comparação o modelo RF, utilizando apenas a vazão acumulada das 11h às 23h59min (com três parametrizações distintas).

<sup>1</sup> cf. Seção 2.6.9

Dado de entrada	Modelo	APE					
		Total	[0,5)	[5,10)	[10,15)	[15,20)	>20
	<i>baseline</i>	252	98	69	41	22	22
	<b>ARIMA</b>	252	101	66	43	26	16
	Suavização Exponencial	252	88	66	47	25	26
	Média Móvel	250	90	59	59	22	20
	<b>RF</b> ( $k-1$ )	251	80	60	52	33	26
	<b>RF</b> ( $k-1, k-2, k-3$ )	249	91	63	51	25	19
	<b>RF</b> ( $k-1, k-2, \dots, k-7$ )	245	95	56	49	27	18
	<b>RF</b> ( $k-7$ )	231	58	61	40	32	40
	<b>RF</b> ( $k-7, k-14$ )	220	59	51	56	33	21
	<b>RF</b> ( $k-7, k-14, k-21$ )	231	64	60	45	30	32
	<b>RF</b> ( $k-1, k-2, \dots, k-10$ )	242	96	59	38	36	13
	<b>RL</b> ( $k-1, k-2, \dots, k-10$ )	242	90	69	47	21	15
	<b>MLP</b> ( $k-1, k-2, \dots, k-10$ )	242	95	67	47	16	17
	<b>CFTS</b> ( $k-1$ )	251	74	67	52	36	22
	<b>WFTS</b> ( $k-1$ )	251	83	63	48	26	31
	<b>WHOFTS</b> ( $k-1$ )	251	76	75	43	33	24
	<b>PWFTS</b> ( $k-1$ )	251	76	75	48	26	26
	<b>ANFIS-SFC</b> ( $k, k-1, \dots, k-7$ )	224	99	67	37	12	9
	<b>ANFIS-FCM</b> ( $k, k-1, \dots, k-7$ )	224	100	71	32	14	7
	<b>Yamawawa-NFN</b> ( $k, k-1, \dots, k-7$ )	224	99	69	34	14	8
	<b>ANFIS-SFC</b> ( $k, k-7, k-14, k-21, k-28$ )	221	99	67	35	14	9
	<b>ANFIS-FCM</b> ( $k, k-7, k-14, k-21, k-28$ )	223	102	69	33	11	8
	<b>Yamakawa-NFN</b> ( $k, k-7, k-14, k-21, k-28$ )	<b>224</b>	<b>104</b>	<b>68</b>	<b>33</b>	<b>11</b>	<b>8</b>
	<b>Yamakawa-NFN</b> ( $k, k-1, k-2, \dots, k-5$ )	247	110	75	38	14	10
	<b>RF</b> ( $k, k-1, k-2, \dots, k-5$ )	247	97	78	44	16	12
	<b>RF</b> ( $k-1$ )	251	79	61	52	24	35
	<b>RF</b> ( $k-1, k-2, k-3$ )	249	77	75	45	28	24
	<b>RF</b> ( $k-1, k-2, \dots, k-7$ )	245	88	67	43	26	21

Tabela 4.6 – Resultados do APE usando a abordagem de previsão de um passo à frente no ponto de distribuição A.

Dado de entrada	Modelo	APE					
		Total	[0,5)	[5,10)	[10,15)	[15,20)	>20
	<i>baseline</i>	252	33	35	29	28	127
	<b>ARIMA</b>	252	62	61	37	30	62
	Suavização Exponencial	251	62	38	62	27	62
	Média Móvel	250	66	62	38	24	60
	<b>RF</b> ( $k-1$ )	251	44	51	37	29	90
	<b>RF</b> ( $k-1, k-2, k-3$ )	249	61	49	38	35	66
	<b>RF</b> ( $k-1, k-2, \dots, k-7$ )	245	55	51	39	29	71
	<b>RF</b> ( $k-7$ )	231	42	42	41	15	91
	<b>RF</b> ( $k-7, k-14$ )	231	44	46	32	34	75
	<b>RF</b> ( $k-7, k-14, k-21$ )	231	46	46	38	32	69
	<b>RF</b> ( $k-1, k-2, \dots, k-10$ )	242	49	47	47	37	62
	<b>RL</b> ( $k-1, k-2, \dots, k-10$ )	<b>243</b>	<b>66</b>	<b>56</b>	<b>32</b>	<b>27</b>	<b>61</b>
	<b>MLP</b> ( $k-1, k-2, \dots, k-10$ )	243	65	56	32	27	62
	CFTS ( $k-1$ )	251	42	46	45	29	89
	WFTS ( $k-1$ )	251	43	45	45	30	88
	<b>WHOFTS</b> ( $k-1$ )	251	43	45	41	30	92
	<b>PWFTS</b> ( $k-1$ )	251	43	49	44	27	88
	<b>ANFIS-SFC</b> ( $k, k-1, \dots, k-7$ )	224	55	41	38	43	47
	<b>ANFIS-FCM</b> ( $k, k-1, \dots, k-7$ )	224	51	51	43	33	46
	<b>Yamawawa-NFN</b> ( $k, k-1, \dots, k-7$ )	224	45	47	51	32	49
	<b>ANFIS-SFC</b> ( $k, k-7, k-14, k-21, k-28$ )	224	56	38	43	39	48
	<b>ANFIS-FCM</b> ( $k, k-7, k-14, k-21, k-28$ )	224	51	51	42	33	47
	<b>Yamakawa-NFN</b> ( $k, k-7, k-14, k-21, k-28$ )	224	46	45	47	38	48
	<b>Yamakawa-NFN</b> ( $k, k-1, k-2, \dots, k-5$ )	247	51	44	48	41	63
	<b>RF</b> ( $k, k-1, k-2, \dots, k-5$ )	247	47	49	52	43	56
	<b>RF</b> ( $k-1$ )	251	25	43	41	34	108
	<b>RF</b> ( $k-1, k-2, k-3$ )	249	57	37	50	43	62
	<b>RF</b> ( $k-1, k-2, \dots, k-7$ )	245	59	53	33	35	65

Tabela 4.7 – Resultados do APE usando a abordagem de previsão de um passo à frente no ponto de distribuição B.

Dado de entrada	Modelo	APE					
		Total	[0,5)	[5,10)	[10,15)	[15,20)	>20
	<i>baseline</i>	252	111	46	29	13	53
	<b>ARIMA</b>	252	130	48	25	10	39
	Suavização Exponencial	258	130	54	24	13	37
	Média Móvel	250	131	46	21	13	39
	<b>RF</b> ( $k-1$ )	251	92	79	31	11	38
	<b>RF</b> ( $k-1, k-2, k-3$ )	249	108	68	25	15	33
	<b>RF</b> ( $k-1, k-2, \dots, k-7$ )	245	115	60	27	9	34
	<b>RF</b> ( $k-7$ )	231	45	54	42	24	66
	<b>RF</b> ( $k-7, k-14$ )	228	50	45	45	25	66
	<b>RF</b> ( $k-7, k-14, k-21$ )	231	50	47	45	25	64
	<b>RF</b> ( $k-1, k-2, \dots, k-10$ )	242	92	54	40	25	31
	<b>RL</b> ( $k-1, k-2, \dots, k-10$ )	242	125	50	22	12	33
	<b>MLP</b> ( $k-1, k-2, \dots, k-10$ )	242	115	62	20	9	36
	<b>CFTS</b> ( $k-1$ )	251	95	52	29	23	52
	<b>WFTS</b> ( $k-1$ )	251	96	57	48	12	38
	<b>WHOFTS</b> ( $k-1$ )	251	100	68	36	12	35
	<b>PWFTS</b> ( $k-1$ )	251	94	77	34	10	36
Vazão acumulada das 0h às 10h59min + Vazão acumulada diária	<b>ANFIS-SFC</b> ( $k, k-1, \dots, k-7$ )	224	135	46	14	7	22
	<b>ANFIS-FCM</b> ( $k, k-1, \dots, k-7$ )	224	140	38	13	12	21
	<b>Yamawawa-NFN</b> ( $k, k-1, \dots, k-7$ )	224	126	46	9	7	36
	<b>ANFIS-SFC</b> ( $k, k-7, k-14, k-21, k-28$ )	224	138	43	13	6	24
	<b>ANFIS-FCM</b> ( $k, k-7, k-14, k-21, k-28$ )	224	140	38	13	11	22
	<b>Yamakawa-NFN</b> ( $k, k-7, k-14, k-21, k-28$ )	224	126	48	7	8	35
Vazão acumulada das 0h às 10h59min	<b>Yamakawa-NFN</b> ( $k, k-1, k-2, \dots, k-5$ )	247	139	66	12	7	23
	<b>RF</b> ( $k, k-1, k-2, \dots, k-5$ )	<b>247</b>	<b>155</b>	<b>43</b>	<b>21</b>	<b>7</b>	<b>21</b>
Vazão acumulada das 11 às 24h	<b>RF</b> ( $k-1$ )	251	101	64	28	17	41
	<b>RF</b> ( $k-1, k-2, k-3$ )	249	105	58	29	17	40
	<b>RF</b> ( $k-1, k-2, \dots, k-7$ )	245	110	55	23	16	41

Tabela 4.8 – Resultados do APE usando a abordagem de previsão de um passo à frente no ponto de distribuição C.

Dado de entrada	Modelo	APE					
		Total	[0,5)	[5,10)	[10,15)	[15,20)	>20
	<i>baseline</i>	252	158	53	27	8	6
	<b>ARIMA</b>	252	169	53	20	5	5
	Suavização Exponencial	252	169	49	22	7	5
	Média Móvel	250	163	51	23	7	6
	<b>RF</b> ( $k-1$ )	251	169	51	17	10	4
	<b>RF</b> ( $k-1, k-2, k-3$ )	249	165	51	25	4	4
	<b>RF</b> ( $k-1, k-2, \dots, k-7$ )	245	162	55	17	6	5
	<b>RF</b> ( $k-7$ )	245	103	67	39	24	12
	<b>RF</b> ( $k-7, k-14$ )	238	102	76	33	13	14
	<b>RF</b> ( $k-7, k-14, k-21$ )	231	105	66	38	11	11
	<b>RF</b> ( $k-1, k-2, \dots, k-10$ )	242	157	54	24	2	5
	<b>RL</b> ( $k-1, k-2, \dots, k-10$ )	242	167	49	20	3	3
	<b>MLP</b> ( $k-1, k-2, \dots, k-10$ )	242	157	50	24	7	4
	<b>CFTS</b> ( $k-1$ )	251	141	76	22	7	5
	<b>WFTS</b> ( $k-1$ )	251	167	51	20	9	4
	<b>WHOFTS</b> ( $k-1$ )	251	168	56	16	7	4
	<b>PWFTS</b> ( $k-1$ )	251	165	59	17	7	3
Vazão acumulada das 0h às 10h59min + Vazão acumulada diária	<b>ANFIS-SFC</b> ( $k, k-1, \dots, k-7$ )	224	160	48	12	2	2
	<b>ANFIS-FCM</b> ( $k, k-1, \dots, k-7$ )	224	160	52	7	4	1
	<b>Yamawawa-NFN</b> ( $k, k-1, \dots, k-7$ )	224	159	52	8	3	2
	<b>ANFIS-SFC</b> ( $k, k-7, k-14, k-21, k-28$ )	224	157	55	6	4	2
	<b>ANFIS-FCM</b> ( $k, k-7, k-14, k-21, k-28$ )	224	160	52	7	4	1
	<b>Yamakawa-NFN</b> ( $k, k-7, k-14, k-21, k-28$ )	<b>224</b>	<b>160</b>	<b>54</b>	<b>5</b>	<b>3</b>	<b>2</b>
Vazão acumulada das 0h às 10h59min	<b>Yamakawa-NFN</b> ( $k, k-1, k-2, \dots, k-5$ )	247	167	63	10	3	4
	<b>RF</b> ( $k, k-1, k-2, \dots, k-5$ )	247	174	56	9	5	3
Vazão acumulada das 11 às 24h	<b>RF</b> ( $k-1$ )	251	143	69	24	8	7
	<b>RF</b> ( $k-1, k-2, k-3$ )	246	153	55	23	10	5
	<b>RF</b> ( $k-1, k-2, \dots, k-7$ )	245	145	61	25	10	4

Tabela 4.9 – Resultados do APE usando a abordagem de previsão de um passo à frente no ponto de distribuição D.

Dado de entrada	Modelo	APE					
		Total	[0,5)	[5,10)	[10,15)	[15,20)	>20
	<i>baseline</i>	252	169	64	17	1	1
Vazão acumulada diária	<b>ARIMA</b>	252	188	54	8	0	2
	Suavização Exponencial	252	175	58	15	2	2
	Média Móvel	250	181	53	12	3	1
	<b>RF</b> ( $k-1$ )	251	166	71	12	1	1
	<b>RF</b> ( $k-1, k-2, k-3$ )	249	169	72	6	1	1
	<b>RF</b> ( $k-1, k-2, \dots, k-7$ )	245	176	58	10	0	1
	<b>RF</b> ( $k-7$ )	231	135	85	10	0	1
	<b>RF</b> ( $k-7, k-14$ )	231	135	85	10	0	1
	<b>RF</b> ( $k-7, k-14, k-21$ )	201	128	62	10	0	1
	<b>RF</b> ( $k-1, k-2, \dots, k-10$ )	242	183	54	3	1	1
	<b>RL</b> ( $k-1, k-2, \dots, k-10$ )	242	182	57	2	0	1
	<b>MLP</b> ( $k-1, k-2, \dots, k-10$ )	242	178	48	14	1	1
	<b>CFTS</b> ( $k-1$ )	251	136	93	20	1	1
	<b>WFTS</b> ( $k-1$ )	251	170	66	12	2	1
	<b>WHOFTS</b> ( $k-1$ )	251	164	74	12	0	1
<b>PWFTS</b> ( $k-1$ )	251	170	67	13	0	1	
Vazão acumulada das 0h às 10h59min + Vazão acumulada diária	<b>ANFIS-SFC</b> ( $k, k-1, \dots, k-7$ )	<b>224</b>	<b>180</b>	<b>36</b>	<b>5</b>	<b>2</b>	<b>1</b>
	<b>ANFIS-FCM</b> ( $k, k-1, \dots, k-7$ )	224	179	36	5	4	0
	<b>Yamawawa-NFN</b> ( $k, k-1, \dots, k-7$ )	224	178	36	6	4	0
	<b>ANFIS-SFC</b> ( $k, k-7, k-14, k-21, k-28$ )	224	180	35	6	3	0
	<b>ANFIS-FCM</b> ( $k, k-7, k-14, k-21, k-28$ )	224	180	35	5	4	0
	<b>Yamakawa-NFN</b> ( $k, k-7, k-14, k-21, k-28$ )	224	178	36	6	4	0
Vazão acumulada das 0h às 10h59min	<b>Yamakawa-NFN</b> ( $k, k-1, k-2, \dots, k-5$ )	247	196	36	12	3	0
	<b>RF</b> ( $k, k-1, k-2, \dots, k-5$ )	247	192	47	6	2	0
Vazão acumulada das 11 às 24h	<b>RF</b> ( $k-1$ )	251	159	73	16	1	2
	<b>RF</b> ( $k-1, k-2, k-3$ )	249	168	62	15	2	2
	<b>RF</b> ( $k-1, k-2, \dots, k-7$ )	245	170	61	11	1	2

Tabela 4.10 – Resultados do APE usando a abordagem de previsão de um passo à frente no ponto de distribuição E.

## 4.2 Escolha de Três Modelos Invariantes no Tempo

Na segunda etapa, são registrados 194.303 valores de vazão, a cada 15 minutos de 01/01/2018 a 17/07/2023. Além disso, os anos de 2019 e 2020 — em que o consumo muda de dinâmica, pois com a pandemia, há baixas vazões — são excluídos da base de dados de C, uma vez que com um rompimento de barragem na cidade referida, considera-se que não está representado o comportamento real desse ponto de distribuição. Após o processamento, são geradas duas bases de dados para cada ponto de distribuição: uma com a soma das vazões de 0h às 10h59min e outra com a soma das vazões de 0h às 23h59min. Com exceção de C, com 1.293, as bases de

dados ficam com 2.024 registros, dos quais 405 são utilizados para teste, totalizando 80% + 20%.

Depois da aplicação dos modelos na primeira etapa, é decidido usar três modelos invariantes no tempo, **RF<sup>2</sup>**, **RL** e Yamakawa-**NFN**, bem retratados nas Tabelas 4.11, 4.12, 4.13, 4.14 e 4.15 — em que estão grifadas em azul as colunas que mostram as melhores previsões para a faixa [0, 5)%. É claro que modelos variantes no tempo estavam previstos no cronograma — não nesta etapa.

Modelo	[0,5) %	[5,10)%	[10,15)%	[15,20)%	>20%	Total	Faixa 0 a 5%
<b>RL</b>	147	124	58	36	40	405	36,30 %
<b>RF</b>	155	110	66	33	41	405	<b>38,27 %</b>
Yamakawa	153	125	55	29	43	405	37,78 %

Tabela 4.11 – Resultados do APE na segunda etapa para o ponto de distribuição A.

Modelo	[0,5) %	[5,10)%	[10,15)%	[15,20)%	>20%	Total	Faixa 0 a 5%
<b>RL</b>	70	61	57	41	176	405	17,28 %
<b>RF</b>	90	71	51	47	146	405	<b>22,22 %</b>
Yamakawa	78	63	54	47	163	405	19,26 %

Tabela 4.12 – Resultados do APE na segunda etapa para o ponto de distribuição B.

Modelo	[0,5) %	[5,10)%	[10,15)%	[15,20)%	>20%	Total	Faixa 0 a 5%
<b>RL</b>	124	65	20	13	37	259	47,88 %
<b>RF</b>	128	59	24	11	37	259	<b>49,42 %</b>
Yamakawa	125	59	22	18	35	259	48,26 %

Tabela 4.13 – Resultados do APE na segunda etapa para o ponto de distribuição C.

Modelo	[0,5) %	[5,10)%	[10,15)%	[15,20)%	>20%	Total	Faixa 0 a 5%
<b>RL</b>	245	99	36	14	11	405	<b>60,49 %</b>
<b>RF</b>	243	98	33	18	13	405	60,00 %
Yamakawa	244	102	32	16	11	405	60,25 %

Tabela 4.14 – Resultados do APE na segunda etapa para o ponto de distribuição D.

<sup>2</sup> cf. Seção 2.6.5

Modelo	[0,5) %	[5,10)%	[10,15)%	[15,20)%	>20%	Total	Faixa 0 a 5%
RL	315	60	16	5	9	405	<b>77,78 %</b>
RF	312	62	17	5	9	405	77,04 %
Yamakawa	313	62	16	5	9	405	77,28 %

Tabela 4.15 – Resultados do APE na segunda etapa para o ponto de distribuição E.

### 4.3 Um Modelo Adaptativo

Ademais, na terceira etapa, os modelos apresentados na Seção anterior são implantados e testados durante o período de 01/07/2023 a 24/10/2023. Os modelos são alimentados com 115 amostras sem pré-processamento.

Para esses experimentos, são usadas duas bases de dados, apresentadas na primeira e na segunda séries temporais das Figuras 3.2, 3.3, 3.5, 3.6 e 3.8. A primeira, usada para escolha dos modelos e treinamento, contém 2.024 amostras de 01/01/2018 a 30/06/2023. A segunda contém 116 amostras sem pré-processamento referentes às datas 01/07/2023 a 24/10/2023. Esses dados são usados apenas para validação.

São desenvolvidos modelos adaptativos sem UKF e com ele. Em seguida, é gerada a Tabela 4.16, com o antes e o depois da implantação do modelo adaptativo, em sua primeira fase de testes.

Porcentagem de estimativas dentro da faixa de APE de 5%			
	Modelo Proposto Antes	Modelo Proposto Depois	Modelo <i>baseline</i>
A	57,75 %	<b>96,82 %</b>	59,48 %
B	34,48 %	<b>30,17 %</b>	15,52 %
C	13,79 %	<b>56,03 %</b>	53,45 %
D	97,41 %	<b>92,24 %</b>	79,31 %
E	79,31 %	<b>81,03 %</b>	62,07 %

Tabela 4.16 – Resultados dos modelos Yamakawa-NFN adaptativos na quarta etapa.

Repara-se que os resultados são ótimos em todos os pontos de distribuição.

### 4.4 Fase de Testes

Na quarta etapa, é feita a segunda fase de testes, que é referente aos dias 13/12/2023 a 23/01/2024. Os dados reais deste período são usados como teste e os dados coletados anteriormente são usados para treinar os modelos. Na fase de teste anterior, é observada uma mudança nas propriedades estatísticas dos dados pré-processados (01/2018 a 07/2023) e sem pré-processamento (07/2023 a 12/2023). Com o objetivo de melhorar os dados com pré-processamento, são feitos ajustes nesses

dados usando as propriedades dos dados sem pré-processamento. Consideram-se *outliers* valores maiores ou menores que o desvio-padrão referente ao dia da semana. Os *outliers* são substituídos pela média correspondente ao dia da semana.

Com essas novas bases de dados, que já são ilustradas para fins de comparação, realiza-se um novo experimento. Os modelos Yamakawa adaptativos são treinados com as séries históricas de vazão de 01/01/2018 a 12/12/2023 e validados em 40 registros referentes as datas 13/12/2023 a 23/01/2024. Devido à sazonalidade apresentada em E, é decidida a criação de dois modelos para essa base de dados: um modelo específico para semana, e outro para fim de semana.

Com a mudança em 30/06/2023 (mudança nas propriedades estatísticas), a sazonalidade passa a estar presente nos dados, e, conseqüentemente, faz-se necessário considerá-las. Por fim, é gerada a Tabela 4.17, com os resultados.

	Porcentagem de estimativas dentro da faixa de APE de 5%	
	Modelo Proposto	Modelo <i>baseline</i>
A	62,5 %	58,5 %
B	12,5 %	12,2 %
C	65,0 %	53,7 %
D	75,0 %	61,0 %
E (modelo único)	57,5 %	48,8 %
E (modelos específicos)	70,0 %	48,8 %

Tabela 4.17 – Resultados dos modelos Yamakawa-NFN adaptativos após a fase 2 de testes.

Repara-se novamente que os resultados são ótimos em todos os pontos de distribuição.

#### 4.5 Previsão de Sete Dias à Frente

São feitos experimentos com modelos UKF de vários estados. Conclui-se que a complexidade, quando aumentada, não compensa, pois os resultados não melhoram significativamente. Por isso, ao final, são mantidos os dois estados.

Para esses experimentos, são usadas duas bases de dados. A primeira, usada para escolha dos modelos e treinamento, é apresentada na Seção 3.3 e contém 2.024 amostras de 01/01/2018 a 30/06/2023. A segunda contém 116 amostras não manipuladas referentes às datas 01/07/2023 a 24/10/2023.

Os modelos *Naive*<sup>3</sup>, *Naive Sazonal*, ARIMA, SARIMA, WHOFTS e PWFTS são implementados com (1) previsão recursiva e Yamakawa Adaptativo com (2) previsão

<sup>3</sup> cf. Seção 2.6.3

especialista. A Figura 4.1 ilustra como essas metodologias funcionam. Em (1) temos apenas um modelo que faz previsão de um passo à frente. Para as previsões com horizonte maior ou igual a dois, o modelo é realimentado com suas previsões anteriores (percebe-se a previsão de  $\hat{y}[k+2]$  com 3 pontos reais e 1 ponto estimado; assim como a previsão de  $\hat{y}[k+3]$  com 2 pontos reais e 2 pontos estimados). A vantagem dessa metodologia é a facilidade de implementação e a escalabilidade das previsões. Porém, como para toda previsão existe uma incerteza associada, à medida que o horizonte de previsão aumenta, a incerteza dessas previsões vão sendo acumuladas, diminuindo a precisão do modelo.

Já em (2), para cada horizonte de previsão deve existir um modelo especialista. Nessa abordagem, o acúmulo de incertezas é minimizado; porém, devido ao número de modelos, é mais difícil de implementar e não é escalável a outros horizontes de previsão.

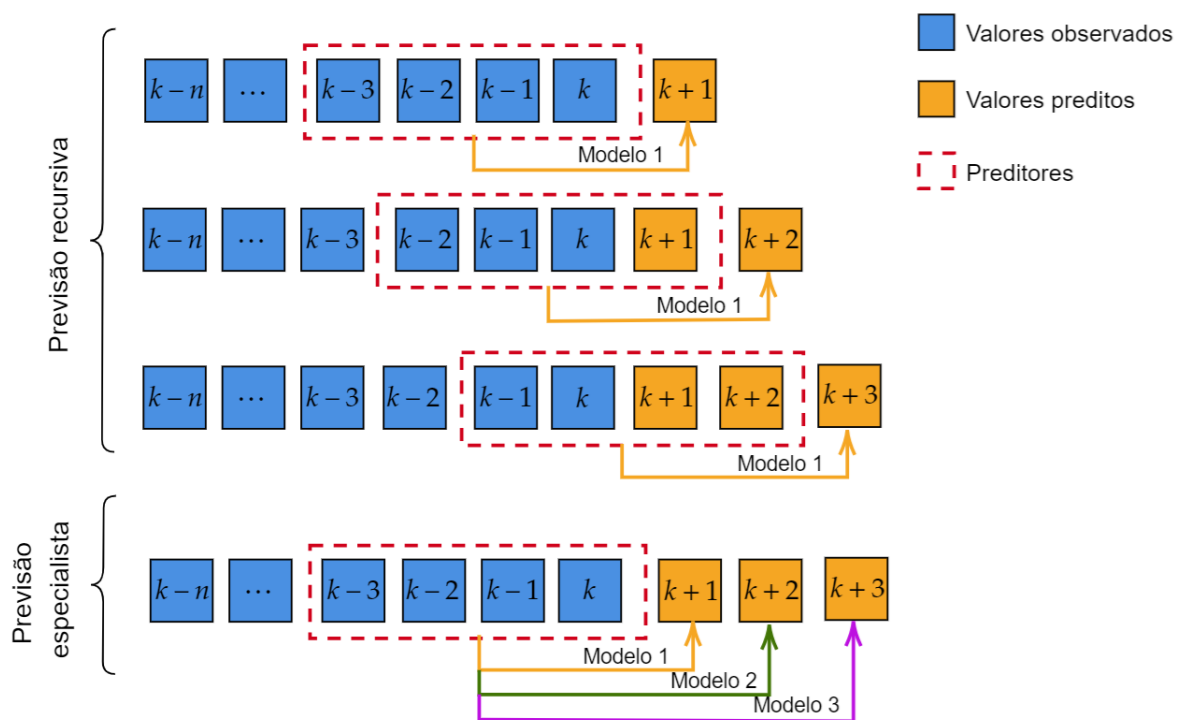


Figura 4.1 – Previsão recursiva e especialista.

Além disso, são feitos experimentos com previsões pontuais *versus* intervalares, disponíveis para permitir que o usuário consiga medir a incerteza do modelo e tomar melhores decisões.

Os resultados das previsões de 1 a 7 passos à frente são apresentados nas Tabelas 4.18, 4.19, 4.20, 4.21 e 4.22. Percebe-se que, nos pontos A e D, se sobressai o modelo mais à direita — Yamakawa adaptativo. Porém, somente no ponto C, o modelo *Naive* pode ser visto como melhor opção para 1 a 3 passos à frente. Por outro lado, é uma vantagem utilizar um único modelo para todas as previsões.

PASSOS	Percentual de APE na faixa de 0 a 5%						
	NAIVE	NAIVE SAZONAL	ARIMA	SARIMA	PWFTS	WHOFTS	UKF
1	20,00	11,30	19,13	19,13	20,87	16,52	<b>24,14</b>
2	14,91	12,28	<b>21,93</b>	16,67	16,67	16,67	20,69
3	11,50	13,27	15,93	12,39	14,16	18,59	<b>20,69</b>
4	<b>18,75</b>	8,93	12,50	8,93	16,07	16,07	16,38
5	14,41	9,01	9,91	12,61	11,71	10,81	<b>22,41</b>
6	10,00	9,01	12,73	16,36	7,27	9,09	<b>19,83</b>
7	11,01	12,84	14,68	11,93	13,76	16,51	<b>19,83</b>

Tabela 4.18 – Resultados na previsão de 7 passos à frente para o ponto de distribuição A.

PASSOS	Percentual de APE na faixa de 0 a 5%						
	NAIVE	NAIVE SAZONAL	ARIMA	SARIMA	PWFTS	WHOFTS	UKF
1	9,57	9,57	8,70	11,30	11,30	<b>13,91</b>	6,90
2	13,16	13,16	8,77	<b>14,04</b>	11,40	12,28	10,34
3	4,42	12,39	8,85	6,19	8,85	12,39	<b>13,79</b>
4	11,61	9,82	8,93	12,50	12,50	12,50	<b>12,93</b>
5	11,71	5,41	11,71	9,01	<b>11,71</b>	10,81	8,62
6	4,55	4,55	12,73	10,00	10,91	<b>13,64</b>	8,62
7	10,09	11,01	11,93	11,93	12,84	<b>14,68</b>	8,62

Tabela 4.19 – Resultados na previsão de 7 passos à frente para o ponto de distribuição B.

PASSOS	Percentual de APE na faixa de 0 a 5%						
	NAIVE	NAIVE SAZONAL	ARIMA	SARIMA	PWFTS	WHOFTS	UKF
1	<b>28,70</b>	14,78	7,83	11,30	13,91	12,17	21,55
2	<b>19,30</b>	14,91	7,02	8,77	13,16	11,40	13,79
3	<b>13,27</b>	11,50	13,27	13,27	12,39	7,96	12,93
4	13,39	16,07	9,82	7,14	9,82	6,25	<b>19,82</b>
5	11,71	14,41	8,11	5,41	12,61	9,01	<b>15,52</b>
6	11,82	10,00	<b>12,73</b>	10,00	9,09	9,01	5,17
7	15,60	13,76	6,42	8,26	11,93	14,68	<b>23,28</b>

Tabela 4.20 – Resultados na previsão de 7 passos à frente para o ponto de distribuição C.

#### 4.6 Otimização de Parâmetros

Nesta etapa, é feita a otimização dos parâmetros, tanto na previsão diária quanto nas previsões de 1 a 7 passos à frente. Buscam-se os percentuais de  $\alpha$ -ótimo, executando o algoritmo exaustivamente. Inicialmente, busca-se o primeiro algoritmo significativo de cada um dos percentuais de  $\alpha$ -ótimo nos modelos de Yamakawa, através de busca linear. De posse deste, é refinada a busca para o segundo e o terceiro algoritmos significativos.

PASSOS	Percentual de APE na faixa de 0 a 5%						
	NAIVE	NAIVE SAZONAL	ARIMA	SARIMA	PWFTS	WHOFTS	UKF
1	48,70	32,17	39,13	33,04	49,57	47,83	<b>56,90</b>
2	35,09	32,46	41,23	28,95	47,37	40,35	<b>52,59</b>
3	23,89	30,97	<b>44,25</b>	33,63	39,82	38,05	6,90
4	24,11	32,14	38,39	33,93	31,25	32,14	<b>49,14</b>
5	27,93	25,23	38,74	34,23	33,33	34,23	<b>50,86</b>
6	24,55	31,82	41,82	33,64	34,55	31,82	<b>50,00</b>
7	32,11	23,85	39,45	35,78	37,61	39,45	<b>45,69</b>

Tabela 4.21 – Resultados na previsão de 7 passos à frente para o ponto de distribuição D.

PASSOS	Percentual de APE na faixa de 0 a 5%						
	NAIVE	NAIVE SAZONAL	ARIMA	SARIMA	PWFTS	WHOFTS	UKF
1	<b>29,57</b>	47,83	28,70	26,96	31,30	31,30	17,24
2	28,95	20,18	28,95	<b>29,82</b>	26,32	21,05	26,72
3	18,58	23,01	21,24	18,58	27,43	26,55	<b>29,31</b>
4	23,21	18,75	20,54	20,54	25,00	25,89	<b>32,76</b>
5	25,23	18,02	<b>34,23</b>	29,73	28,83	30,63	28,45
6	34,55	27,27	30,91	28,18	36,36	<b>36,36</b>	31,03
7	50,46	20,18	<b>55,05</b>	51,38	31,19	32,11	33,62

Tabela 4.22 – Resultados na previsão de 7 passos à frente para o ponto de distribuição E.

## 4.7 Homologação

Devido às alterações realizadas nas bases de dados da Seção 4.4, é decidido executar novamente os testes para os modelos de previsão de 1 a 7 passos à frente. Os modelos adaptativos Yamakawa de 1 a 7 dias à frente são retreinados utilizando séries históricas de vazão do período de 01/01/2018 a 12/12/2023 e validados com 40 registros correspondentes às datas de 13/12/2023 a 23/01/2024.

Dado que com os resultados obtidos com modelos específicos para E, conforme apresentado na Seção 4.4, é demonstrado melhor desempenho, é decidido especializar também os modelos de previsão. Os resultados são sumarizados na Tabela 4.23 — em que estão grifadas em azul as células que mostram as melhores previsões (sendo que o desempate leva em consideração as faixas acima de 5% do APE que não estão disponíveis na Tabela).

## 4.8 Previsão de Múltiplos Passos à Frente

Essa etapa é detalhada no trabalho Ferraz et al. (2024), em que analisou-se o comportamento semestral (*fit* em um ano, *predict* em seis meses), do modelo Yamakawa (1992) em comparação com outros seis modelos.

Pontos	Percentual de APE na faixa de 0 a 5%						
	1	2	3	4	5	6	7
A sem UKF	<b>33,33</b>	14,29	9,52	<b>16,67</b>	23,81	21,43	<b>16,67</b>
A com UKF	30,95	<b>19,05</b>	<b>14,29</b>	14,29	<b>26,19</b>	<b>38,10</b>	7,14
B sem UKF	9,52	4,76	4,76	14,29	9,52	16,67	9,52
B com UKF	<b>9,52</b>	<b>16,67</b>	<b>9,52</b>	<b>14,29</b>	<b>9,52</b>	<b>16,67</b>	<b>9,52</b>
C sem UKF	9,52	4,76	<b>19,05</b>	4,76	23,81	9,52	<b>14,29</b>
C com UKF	<b>9,52</b>	<b>7,14</b>	14,29	<b>9,52</b>	<b>23,81</b>	<b>30,95</b>	9,52
D sem UKF	29,27	14,63	17,07	39,02	14,63	19,51	24,39
D com UKF	<b>31,70</b>	<b>36,59</b>	<b>19,51</b>	<b>39,02</b>	<b>39,02</b>	<b>19,51</b>	<b>24,39</b>
E sem UKF	<b>4,76</b>	4,76	2,38	4,76	7,14	<b>7,14</b>	4,76
E com UKF	0,00	<b>7,14</b>	<b>2,38</b>	<b>11,90</b>	<b>9,52</b>	4,76	<b>33,33</b>

Tabela 4.23 – Resultados na previsão para de 1 a 7 passos à frente após alteração nas bases de dados.

Sem UKF vs. com UKF.

#### 4.8.1 Metodologia

Nesta Seção é apresentada a metodologia usada no pré-processamento dos dados e na configuração dos modelos usados para prever as séries temporais diárias de consumo de GN de 1, 4 e 7 passos à frente. O estudo é realizado em três bases de dados referentes a três estações de recebimento de GN diferentes da GASMIG.

#### 4.8.2 Pré-processamento das bases de dados

As bases de dados utilizadas neste estudo abrangem um período de 6 anos (2018 a 2023) de consumo diário de três estações de recebimento de GN distintas, denominadas B, C e D. Os dados até junho de 2023 (B1) são registrados diretamente pelos sensores, porém apresentam algumas inconsistências, como o registro de vazões negativas e *outliers*. Já os registros de julho a dezembro de 2023 (B2) são realizados manualmente e não apresentam tais inconsistências, o que os torna uma referência confiável para a correção dos dados B1.

Para corrigir as inconsistências de B1, calculou-se a média ( $\mu_{B2}$ ) e o desvio-padrão ( $\sigma_{B2}$ ) para cada dia da semana usando B2 e em seguida substituiu-se os valores negativos e *outliers* de B1 pela média  $\mu_{B2}$ . São considerados *outliers* os valores maiores ou menores que  $\mu_{B2} \pm \sigma_{B2}$ . Uma outra opção seria buscar os valores de um ano atrás, opção essa que foi descartada dada a discretização em dias. Considerando as análises da Seção 3.3, não convém considerar  $k - 360$ .

A Figura 4.2 apresenta uma amostra normalizada<sup>4</sup> da série temporal de consumo diário referente ao ano 2023 para cada estação. Percebe-se que, por serem

<sup>4</sup> Z-Score, uma vez que os dados são sigilosos, todos os cálculos são feitos sem normalização, que está presente somente na exibição dos dados

estações que atendem consumidores industriais, nas três séries temporais, o consumo não varia independente da estação do ano e do dia. Esse padrão é visível também nos outros anos (2018 a 2022).

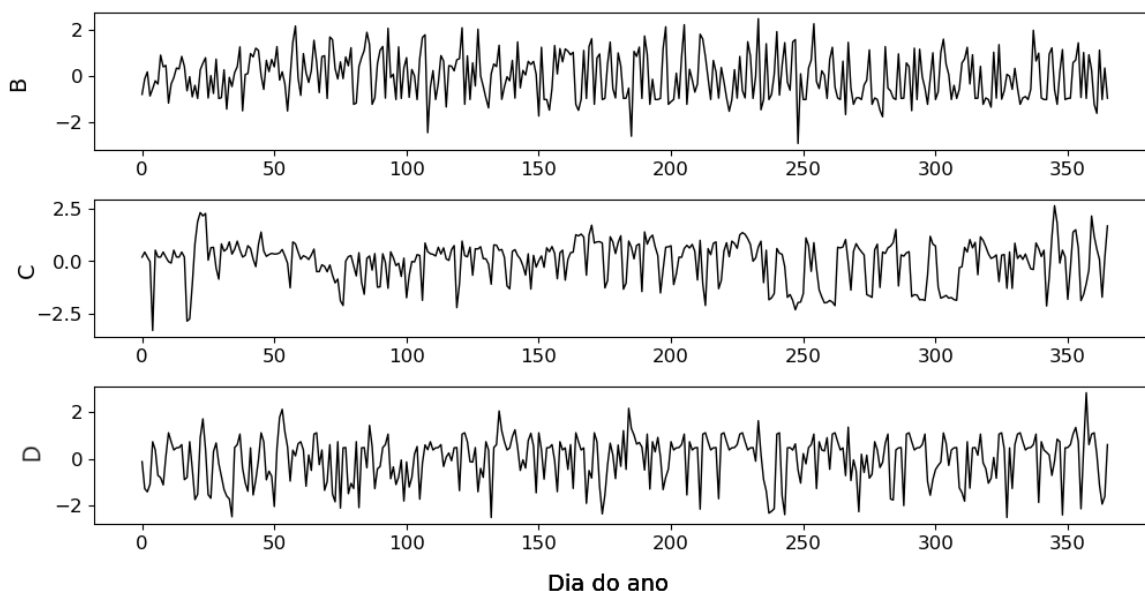


Figura 4.2 – Série temporal referente ao ano 2023 de consumo diário de gás natural das estações de recebimento B, C e D.

### 4.8.3 Configuração dos modelos

Os modelos comparados são o *NAIVE*, *ARIMA*, *WHOFTS*<sup>5</sup>, *NFN*, *MLP*, *RF* e *LSTM*. Para o *NAIVE*, *ARIMA* e *WHOFTS*, é empregada a abordagem de previsão recursiva. Para o *NFN*, *RF* e *MLP*, adotou-se um modelo especializado para cada horizonte de previsão, enquanto para o *LSTM*, é utilizada a técnica de previsão direta, ou seja, o modelo possui uma saída para cada horizonte de previsão.

Os hiperparâmetros dos modelos são otimizados utilizando um ano dos primeiros registros das bases de dados. Para a otimização dos hiperparâmetros do *ARIMA* utiliza-se a função `auto_arima` da biblioteca *Pyramid*<sup>6</sup>. Os resultados obtidos são os seguintes: para B,  $p = 1$ ,  $d = 1$  e  $q = 2$ ; para C,  $p = 3$ ,  $d = 1$  e  $q = 1$ ; e para D,  $p = 1$ ,  $d = 1$  e  $q = 1$ . Aqui,  $p$  representa a ordem do modelo,  $d$  é o grau de diferenciação e  $q$  é a ordem do modelo de média móvel.

Para otimizar os hiperparâmetros do *WHOFTS*, é utilizado o *AG* desenvolvido por *Lucas, Silva e Guimarães (2019)*, implementado na biblioteca *PyFTS*<sup>7</sup>. Os hiperparâmetros otimizados incluíram o número de partições ( $np$ ), método de

<sup>5</sup> cf. Seção 2.6.12

<sup>6</sup> <https://github.com/alkaline-ml/pmdarima>

<sup>7</sup> <https://pyfts.github.io/pyFTS/>

particionamento ( $mp$ ), função de pertinência ( $fp$ ), ordem ( $o$ ),  $\alpha$ -cut ( $ac$ ) e lags ( $l$ ). Os resultados obtidos são os seguintes:

- Para A:  $np = 25$ ,  $mp = \text{grid}$ ,  $fp = \text{triangular}$ ,  $o = 1$ ,  $ac = 0.2$  e  $l = 1$ .
- Para B:  $np = 12$ ,  $mp = \text{grid}$ ,  $fp = \text{triangular}$ ,  $o = 1$ ,  $ac = 0.19$  e  $l = 1$ .
- Para C:  $np = 35$ ,  $mp = \text{grid}$ ,  $fp = \text{triangular}$ ,  $o = 1$ ,  $ac = 0$  e  $l = 1$ .

Os hiperparâmetros do modelo LSTM são selecionados de maneira empírica, sendo eles: porcentagem de  $dropout = 0.02$ , normalização em lote, número de unidades = 67, número de camadas = 4 e número de entradas = 7.

Para o MLP, NFN e RF, é empregado o método de busca em grade (*grid search*). Em ambos os métodos, o número de entradas é definido como 7. Os hiperparâmetros otimizados para o MLP são o número de camadas intermediárias, resultando nos seguintes conjuntos para os casos B, C e D: (5, 20, 5), (10, 50, 10) e (10, 50, 10), respectivamente.

Para o NFN, é utilizada uma entrada com 5 funções de pertinência para todas as bases de dados.

Os hiperparâmetros otimizados do RF são os seguintes: o número de árvores de decisão ( $na$ ), o número máximo de características consideradas ao procurar a melhor divisão ( $mf$ ) e o número mínimo de amostras necessárias para estar em um nó folha ( $ms$ ). Os resultados obtidos são os seguintes:

- Para B:  $na = 10$ ,  $mf = \log_2 T$  e  $ms = 5$ .
- Para C:  $na = 5$ ,  $mf = \sqrt{T}$  e  $ms = 5$ .
- Para D:  $na = 10$ ,  $mf = \log_2 T$  e  $ms = 5$ .

Em que  $T$  é o número total de características.

#### 4.8.4 Experimentos

O objetivo dos experimentos é avaliar o desempenho de diferentes modelos na previsão diária de consumo de GN. A Figura 4.3 ilustra como os experimentos são realizados. Após o pré-processamento, as bases de dados são divididas em 10 janelas sobrepostas com deslocamento de 6 meses. É usado 1 ano de dados para treino e 6 meses para teste. Os modelos são treinados e avaliados em cada janela, observando os horizontes de previsão de 1, 4 e 7 dias à frente. A comparação dos modelos é feita usando a média do MAPE das 10 janelas de teste.

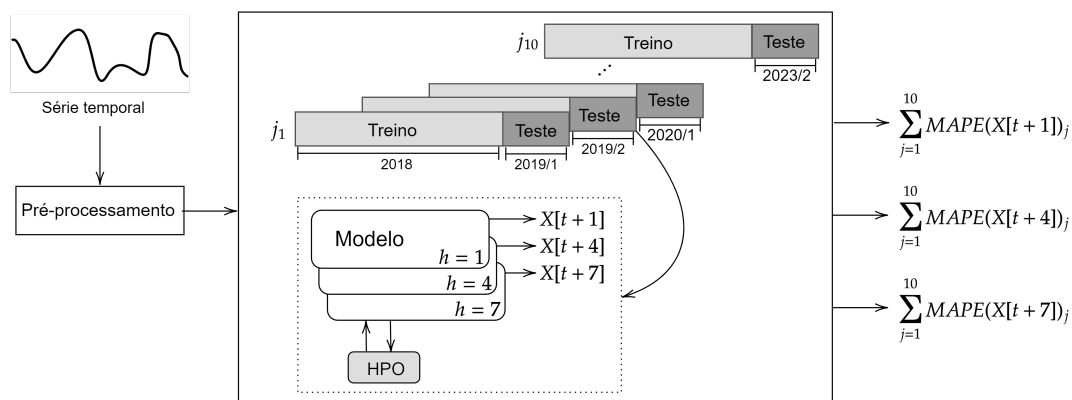


Figura 4.3 – Design dos experimentos para avaliação dos modelos de previsão diária de consumo de gás natural de 1, 4 e 7 passos à frente. HPO refere-se a otimização de hiperparâmetros.

Tabela 4.24 – MAPE das previsões de 1, 4 e 7 passos à frente dos modelos NAIVE, ARIMA, WHOFTS, NFN, MLP, RF e LSTM para B.

	1	4	7
NAIVE	30.49 ± 17.7	27.21 ± 12.10	24.89 ± 16.31
ARIMA	25.12 ± 11.6	23.47 ± 11.28	22.72 ± 7.9
WHOFTS	31.74 ± 22.85	27.27 ± 13.9	22.7 ± 9.9
NFN	28.18 ± 14.35	28.54 ± 17.12	27.82 ± 13.57
MLP	26.52 ± 9.66	26.83 ± 9.52	28.07 ± 12.97
RF	33.87 ± 31.95	33.09 ± 28.97	35.81 ± 36.55
LSTM	59.41 ± 17.72	59.21 ± 17.12	60.01 ± 15.72

#### 4.8.5 Resultados

Nesta Seção são apresentados os resultados dos experimentos que comparam os modelos NAIVE, ARIMA, WHOFTS, NFN, MLP, RF e LSTM.

As Tabelas 4.24, 4.25 e 4.26 apresentam a média mais ou menos o desvio-padrão ( $\mu \pm \sigma$ ) do MAPE<sup>8</sup> de 1, 4 e 7 passos à frente das 10 janelas de teste dos pontos B, C e D. Para verificar se existe diferença significativa nas médias dos resultados, os testes estatísticos de Friedman e o teste post-hoc de Nemenyi são aplicados para cada horizonte de previsão. Os testes são escolhidos após verificação, por meio do teste de Shapiro-Wilk, que existem amostras que não provêm de uma distribuição normal.

<sup>8</sup> cf. Seção 2.8

Tabela 4.25 – **MAPE** das previsões de 1, 4 e 7 passos à frente dos modelos NAIVE, ARIMA, WHOFTS, NFN, MLP, RF e LSTM para C.

	1	4	7
NAIVE	17.98 ± 17.27	18.83 ± 19.68	17.68 ± 18.35
ARIMA	18.50 ± 16.64	21.16 ± 16.78	17.05 ± 16.24
WHOFTS	19.97 ± 14.53	22.22 ± 15.26	19.61 ± 16.53
NFN	18.30 ± 8.80	18.80 ± 8.81	19.34 ± 7.06
MLP	20.33 ± 8.64	23.58 ± 10.7	23.47 ± 10.89
RF	21.39 ± 9.60	25.37 ± 12.66	24.43 ± 12.10
LSTM	96.76 ± 1.38	96.43 ± 1.58	96.22 ± 1.60

Tabela 4.26 – **MAPE** das previsões de 1, 4 e 7 passos à frente dos modelos NAIVE, ARIMA, WHOFTS, NFN, MLP, RF e LSTM para D.

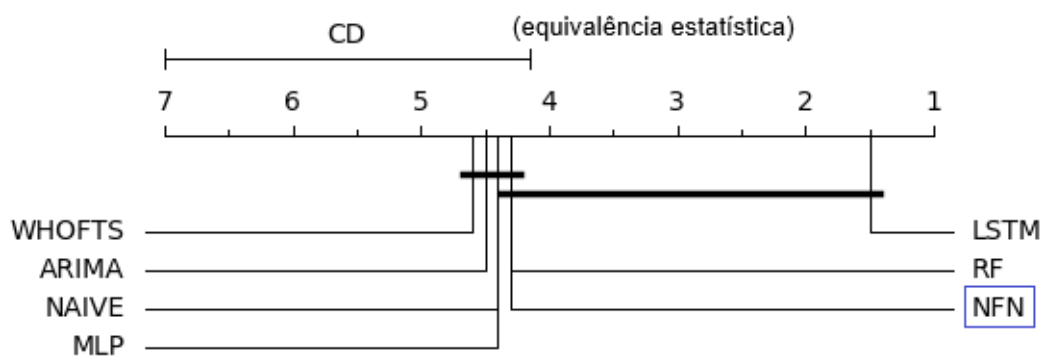
	1	4	7
NAIVE	6.09 ± 4.20	6.68 ± 2.92	6.54 ± 2.89
ARIMA	4.75 ± 2.30	5.40 ± 2.57	5.24 ± 1.57
WHOFTS	5.73 ± 1.42	5.73 ± 1.93	6.12 ± 1.29
NFN	5.77 ± 1.52	7.43 ± 4.75	6.51 ± 0.86
MLP	6.17 ± 0.96	6.54 ± 1.43	6.58 ± 1.14
RF	5.71 ± 0.85	6.44 ± 0.99	6.42 ± 1.05
LSTM	99.46 ± 0.18	99.47 ± 0.18	99.46 ± 0.19

Os resultados dos testes são apresentados em diagramas de diferenças críticas (DEMŠAR, 2006) nas Figuras 4.4, 4.5 e 4.6, para B, C e D, respectivamente; em que é grifada a posição relativa do NFN. Observando os três horizontes de previsão e as três bases de dados, os modelos NAIVE, ARIMA, WHOFTS, NFN, MLP e RF apresentam **médias estatisticamente iguais** (não há um melhor em absoluto), enquanto que o LSTM obtém desempenho abaixo dos demais.

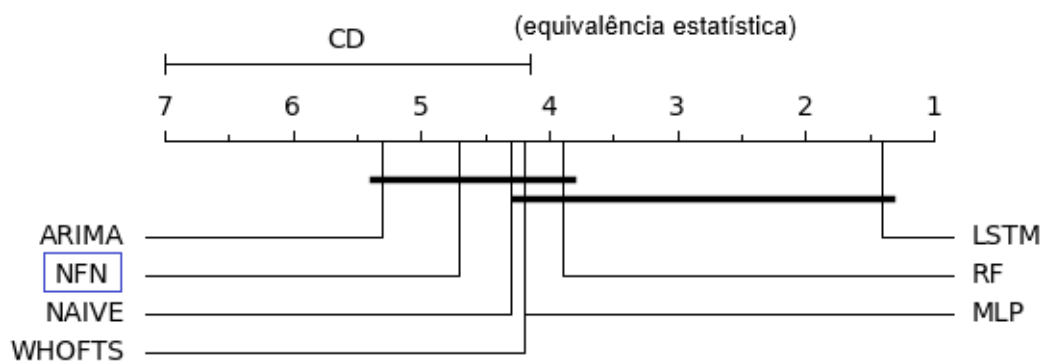
Por fim, a Figura 4.7 apresenta a acurácia dos modelos de previsão para estimativas com MAPE menores que 5% para as estações de recebimento B, C e D. Percebe-se que, para as bases de dados B e C, os modelos não conseguem atingir 50% de estimativas dentro da faixa de MAPE menor que 5%. Em D, com exceção do LSTM, todos os modelos atingem acurácia de 50%. O modelo LSTM apresenta acurácia 0 em C e D e muito próxima de 0 em B. Também é possível observar que os modelos ARIMA e NAIVE se sobressaem em relação aos demais em C e D.

#### 4.9 Série Temporal Semanal

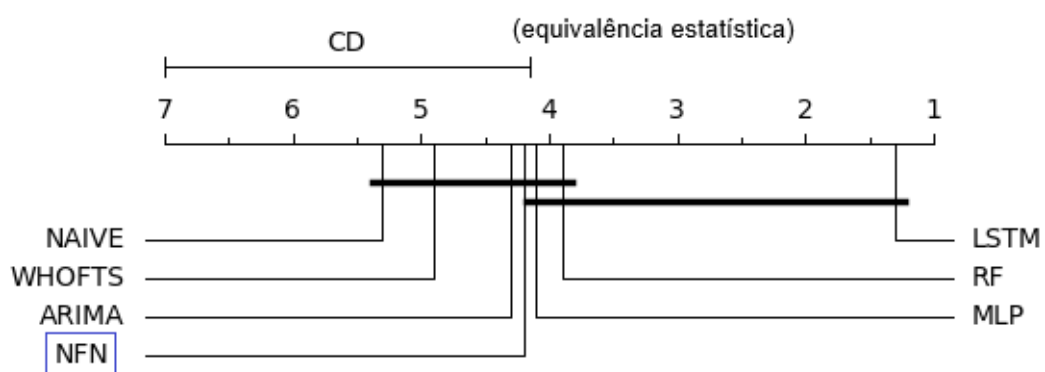
São somados todos os seis pontos de entrega, de segundas-feiras até domingos. Ambas as variáveis de entrada e saída são a soma do acumulado até as 23h59min. São gerados 287 pontos de treino: 01/01/2018 (uma segunda-feira) a 26/06/2023; e 32 pontos de validação: 03/07/2023 a 05/02/2024.



(a) Horizonte 1

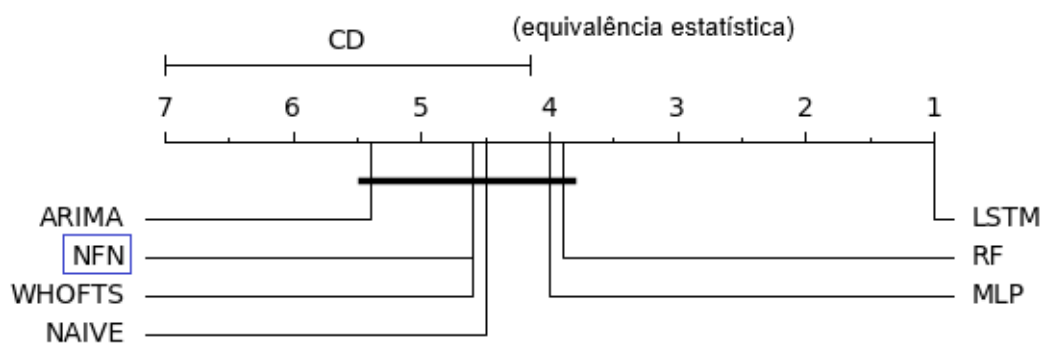


(b) Horizonte 4

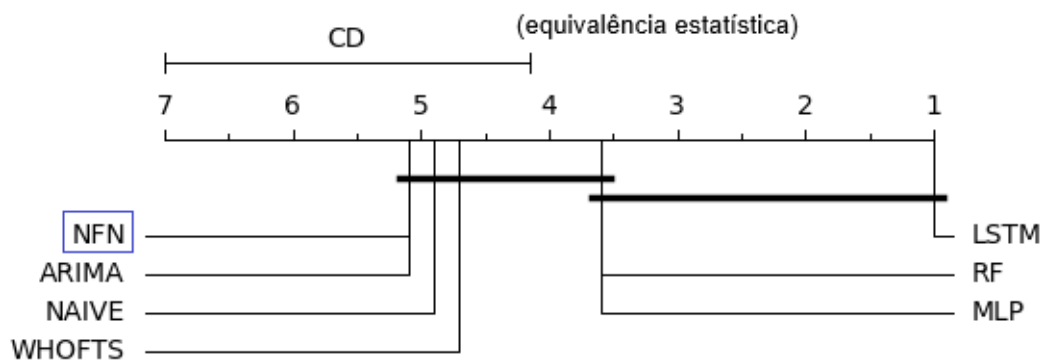


(c) Horizonte 7

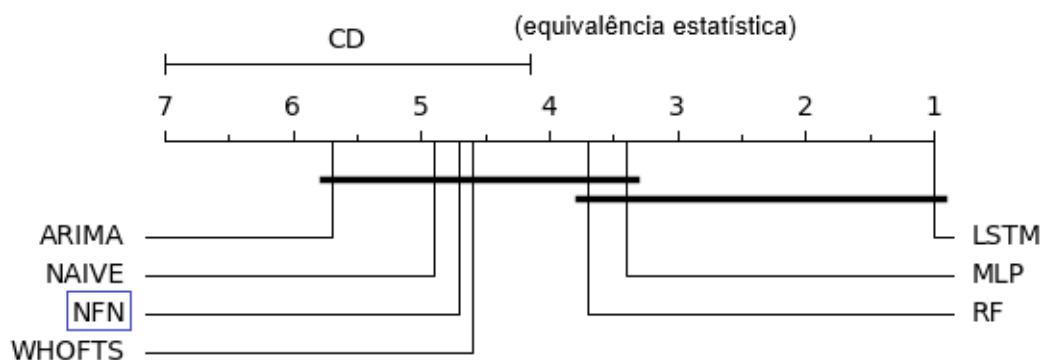
Figura 4.4 – Diagrama de diferenças críticas (CD) dos resultados dos modelos NAIVE, ARIMA, WHOFTS, NFN, MLP, RF e LSTM para B.



(a) Horizonte 1

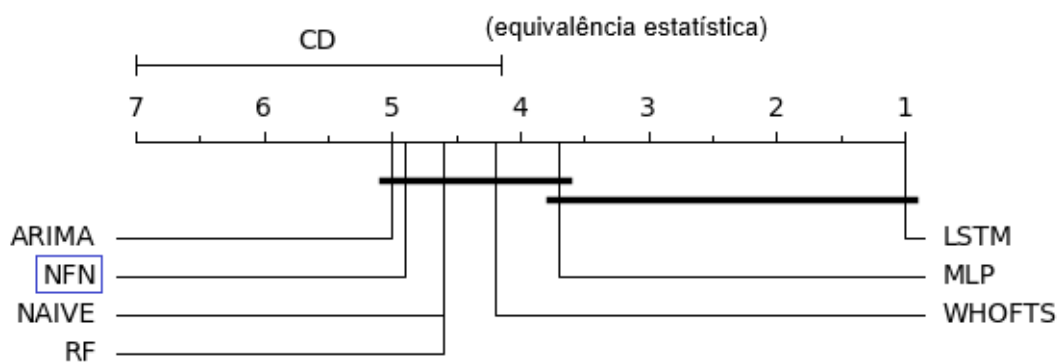


(b) Horizonte 4

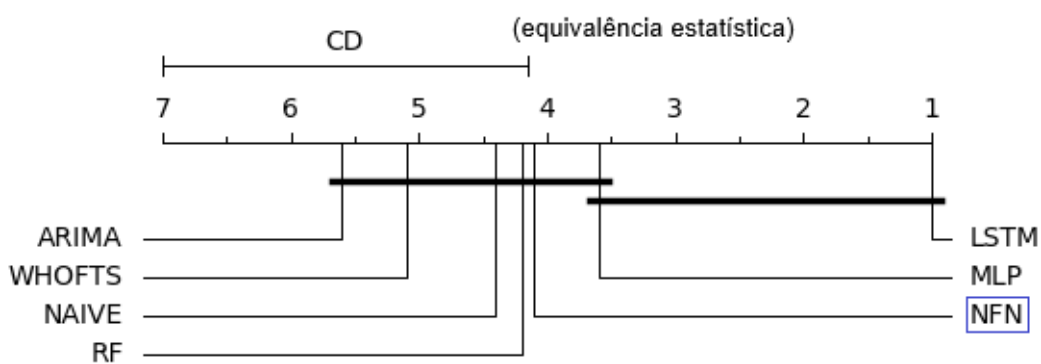


(c) Horizonte 7

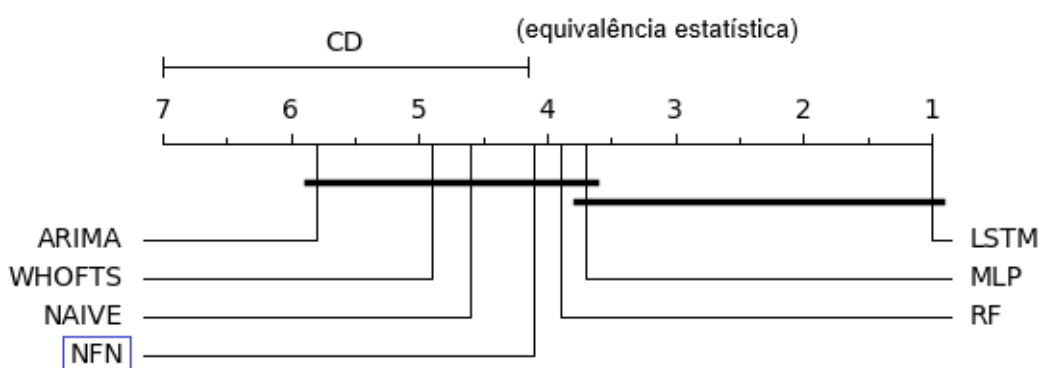
Figura 4.5 – Diagrama de diferenças críticas (CD) dos resultados dos modelos NAIVE, ARIMA, WHOFTS, NFN, MLP, RF e LSTM para C.



(a) Horizonte 1



(b) Horizonte 4



(c) Horizonte 7

Figura 4.6 – Diagrama de diferenças críticas (CD) dos resultados dos modelos NAIVE, ARIMA, WHOFTS, NFN, MLP, RF e LSTM para D.

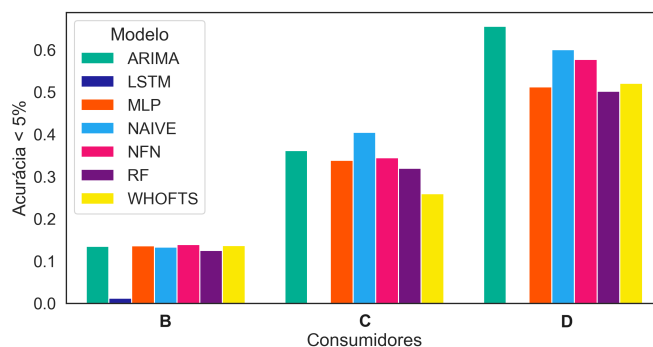


Figura 4.7 – Acurácia dos modelos de previsão para estimativas com APEs individuais menores que 5% para as estações de recebimento B, C e D.

Nesta etapa, são consideradas as seguintes variáveis: seja  $k$  a discretização do tempo contínuo em semanas, isto é,  $k$  é o número de semanas desde 01/01/2018. Pelo algoritmo Yamakawa-NFN básico, são testadas  $V = 1$  a 5 entradas (vazão acumulada de  $v_s[k - 1]$  até  $v_s[k - 5]$ ), em horizontes de 1 a 4 passos à frente (saídas escalares  $\hat{v}_s[k + 1]$  até  $\hat{v}_s[k + 4]$ ). São feitas comparações também com outros algoritmos nebulosos: SFC<sup>9</sup> e FCM.

A Figura 4.8 mostra a dispersão dos dados semanais, altamente próximos da aproximação por RL, e sua visão temporal.

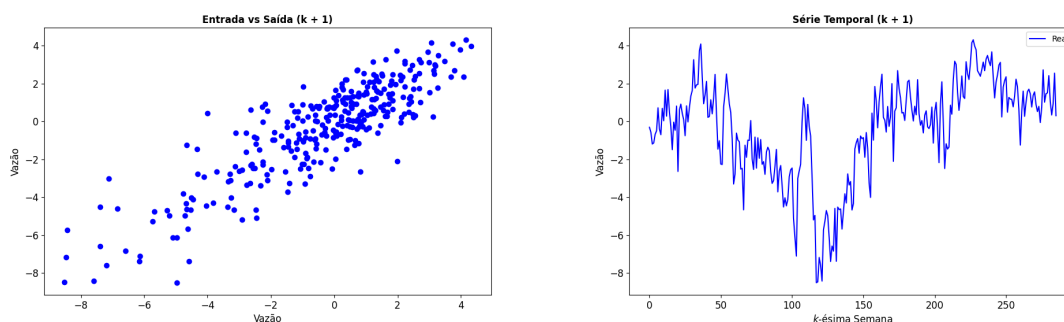


Figura 4.8 – Dados semanais no treinamento — Visão  $k \times (k + 1)$  e Visão temporal.

<sup>9</sup> cf. Seção 2.6.8

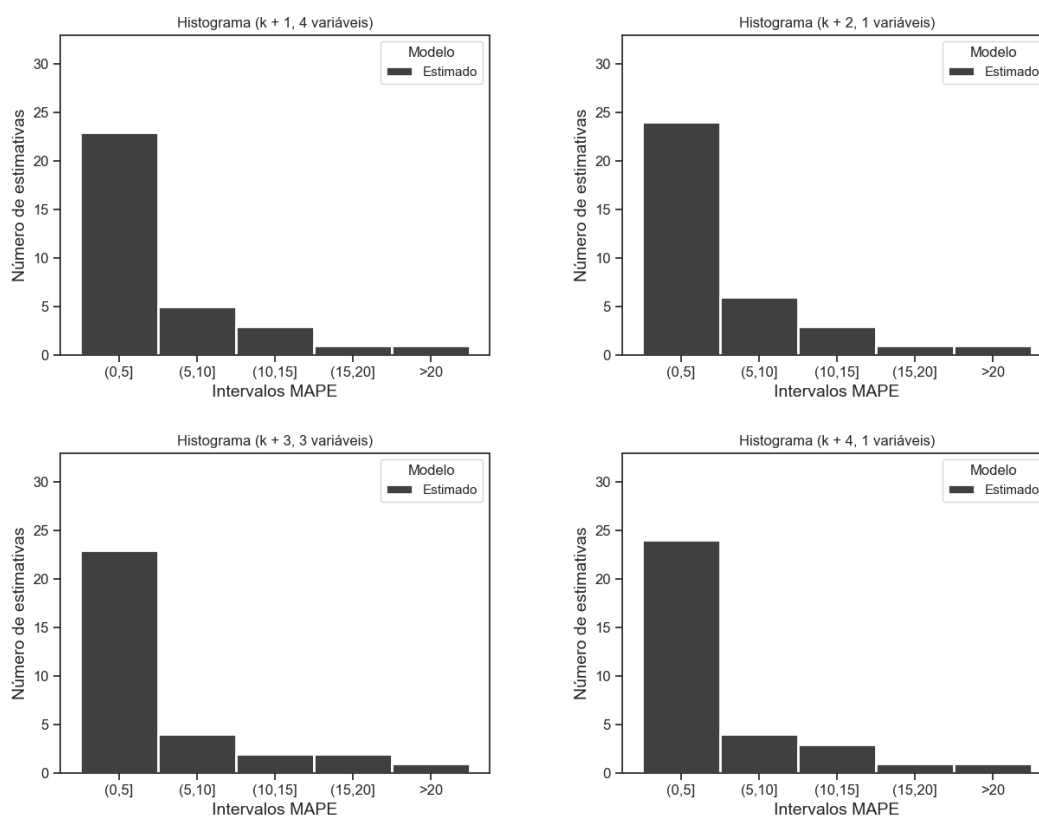


Figura 4.9 – Histogramas dos modelos Yamakawa-NFN vencedores para as variáveis semanais.

$k$	Algoritmo	[0,5)	[5,10)	[10, 15)	[15, 20)	>20	Acerto	Pontos
1	ANFIS-SFC	22	9	0	0	0	70,97%	31
1	ANFIS-FCM	23	7	1	0	0	74,19%	31
2	ANFIS-SFC	18	10	2	0	0	60,00%	30
2	ANFIS-FCM	19	9	2	0	0	63,33%	30
3	ANFIS-SFC	17	11	0	1	0	58,62%	29
3	ANFIS-FCM	15	13	0	1	0	51,72%	29
4	ANFIS-SFC	22	3	3	0	0	78,57%	28
4	ANFIS-FCM <sup>10</sup>	20	5	3	0	0	71,43%	28

Tabela 4.27 – Resultados dos modelos ANFIS-SFC e ANFIS-FCM para variáveis semanais.

<sup>10</sup> cf. Seção 2.6.7

Algoritmo	[0,5)	[5,10)	[10,15)	[15,20)	>= 20	% de Acerto	Pontos	Variáveis	KMais
Yamakawa	23	7	1	0	0	74,19	31	1	1
Yamakawa	23	5	2	0	0	76,67	30	2	1
Yamakawa	21	6	2	0	0	72,41	29	3	1
Yamakawa	22	4	2	0	0	78,57	28	4	1
Yamakawa	19	6	2	0	0	70,37	27	5	1
Yamakawa	23	5	2	0	0	76,67	30	1	2
Yamakawa	21	7	0	1	0	72,41	29	2	2
Yamakawa	21	5	2	0	0	75,00	28	3	2
Yamakawa	20	6	0	1	0	74,07	27	4	2
Yamakawa	18	6	1	1	0	69,23	26	5	2
Yamakawa	21	7	0	1	0	72,41	29	1	3
Yamakawa	21	5	2	0	0	75,00	28	2	3
Yamakawa	22	3	1	1	0	81,48	27	3	3
Yamakawa	18	6	2	0	0	69,23	26	4	3
Yamakawa	17	6	1	1	0	68,00	25	5	3
Yamakawa	23	3	2	0	0	82,14	28	1	4
Yamakawa	21	5	0	1	0	77,78	27	2	4
Yamakawa	19	5	1	1	0	73,08	26	3	4
Yamakawa	17	6	1	1	0	68,00	25	4	4
Yamakawa	15	7	1	1	0	62,50	24	5	4

Tabela 4.28 – Modelo Yamakawa-NFN com 1 a 5 variáveis semanais.

Por sua vez, a Figura 4.9 mostra os histogramas semanais. E a Tabela 4.27 mostra que o modelo NFN supera os outros dois algoritmos, nos quatro horizontes. Por fim, a Tabela 4.28 exhibe o desempenho do modelo NFN com diferentes números de variáveis.

## 5 DISCUSSÃO

Neste capítulo, a partir dos dados apresentados anteriormente são feitas a discussão e a análise dos dados.

### 5.1 Etapa Inicial

Esta é a primeira Etapa — de 01/01/2018 a 09/09/2022. Na escolha de variáveis, considerando-se a Seção 3.3, é decidido utilizar  $v_{11} \rightarrow v_{13}$  ao invés do modelo de previsão enviesado:  $v_{11} \rightarrow v_{24} = v_{11} + v_{13}$ .

Em ambas as abordagens, A e B, todos os modelos também apresentam, no geral, melhora da acurácia em relação ao modelo *baseline*: uma quantidade maior de dias abaixo de 5% do APE, em que não há multa, e um número menor de dias acima de 20%, que é a maior faixa de multa.

#### 5.1.1 Previsão Vespertina

Nos resultados para o ponto de distribuição A (Tabela 4.1), o desempenho dos modelos apresentam maior uniformidade, mas ainda conseguem melhorar as estimativas, quando comparadas com as do modelo *baseline*.

O ponto B (Tabela 4.2) obtém as estimativas mais baixas, mostrando que os dados são muito ruidosos e difíceis de ajustar. Mas, percebe-se que, principalmente os modelos ANFIS-SFC<sup>1</sup> e Yamakawa-NFN conseguem melhorar significativamente as estimativas em relação ao modelo *baseline*, aumentando o número de estimativas abaixo de 5% do APE, em que não há multa.

Já em C (Tabela 4.3), observa-se que o modelo Yamakawa-NFN se sobressai em relação aos demais nas estimativas abaixo de 5% do APE, em que não há multa. E também, todos os modelos reduzem pela metade as estimativas acima de 20% do APE em relação ao modelo *baseline*, que é a maior faixa de multa.

Ademais, em D (Tabela 4.4), observa-se que os modelos RL, RF, MLP e Yamakawa-NFN se sobressaem dos demais. Por fim, em E (Tabela 4.5), as estimativas acima de 20% do APE são reduzidas a zero e as estimativas abaixo de 5% do APE aumentam significativamente com todos os modelos avaliados. A Tabela 5.1 — em que estão grifadas em azul as células que mostram as melhores previsões (maior valor

<sup>1</sup> cf. Seção 2.6.8

para cada ponto de distribuição) — sintetiza os resultados. Nota-se que o **NFN** está em primeiro lugar em C e D, e em segundo lugar nos pontos A, B e E.

Modelo	Percentual de <b>APE</b> na faixa de 0 a 5%				
	Ponto A	Ponto B	Ponto C	Ponto D	Ponto E
<i>baseline</i>	38,89	13,10	44,05	62,70	67,06
<b>RF</b>	42,06	16,67	58,73	69,05	78,57
<b>RL</b>	<b>42,86</b>	20,63	65,60	71,03	<b>79,37</b>
<b>MLP</b>	40,87	15,08	64,29	70,24	76,19
<b>ANFIS-SFC</b>	40,63	<b>25,00</b>	64,73	71,43	79,02
<b>ANFIS-FCM</b>	42,41	20,98	64,29	70,98	79,02
<b>Yamakawa-NFN</b>	42,06	22,62	<b>67,46</b>	<b>72,62</b>	78,97

Tabela 5.1 – Resultados do APE usando a abordagem Vespertina nos cinco pontos de distribuição.

### 5.1.2 Previsão de Um Passo à Frente

No ponto A (Tabela 4.6), o uso da vazão acumulada das 0h às 10h59min melhora o desempenho dos modelos e reduz significativamente as estimativas maiores que 20% do **APE**. O modelo que apresenta melhor desempenho é o **Yamakawa-NFN** usando apenas a série de vazão acumulada das 0h às 10h59min.

No ponto B (Tabela 4.7), assim como nos resultados apresentados na metodologia Vespertina, todos os modelos melhoram significativamente as estimativas em relação ao da *baseline*. O modelo que apresenta melhor resultado é o modelo com estrutura mais simples, a Média Móvel<sup>2</sup>, seguido de outro modelo também simples, o **RL** com 9 entradas.

No ponto C (Tabela 4.8), percebe-se que o uso da vazão acumulada das 0h às 10h59min melhora o desempenho dos modelos e reduz significativamente as estimativas com **APE** maior que 20%, que é a maior faixa de multa. O modelo que apresenta melhor desempenho é o **RF** usando apenas a série de vazão acumulada das 0h às 10h59min.

Ademais, no ponto D (Tabela 4.9), usando como entrada a vazão acumulada diária, os modelos mais simples (com menos entradas) apresentam melhores resultados. Acrescentando a informação da vazão acumulada das 0h às 10h59min, o desempenho foi reduzido e não houve diferença em relação aos tipos ou quantidade de entradas. Já usando apenas a vazão acumulada das 0 às 11h, o modelo **RF** com 5 entradas obtém um aumento no número de estimativas abaixo de 5% do **APE**, em que não há multa.

<sup>2</sup> cf. Seção 2.6.2

Por fim, o ponto E (Tabela 4.10) obtém os melhores resultados usando apenas a série de vazão acumulada das 0h às 10h59min com os modelos Yamakawa-NFN e RF com 5 entradas. Percebe-se também que os modelos avaliados reduzem muito as estimativas com APE maior que 20%, que é a maior faixa de multa.

A Tabela 5.2 — em que estão grifadas em azul as células que mostram as melhores previsões (maior valor para cada ponto de distribuição) — resume os resultados. Vê-se que acrescentar variáveis no NFN diminui seu desempenho em comparação com a previsão vespertina. “Menos é mais.”

Dado de entrada	Modelo	APE A	APE B	APE C	APE D	APE E
	<i>baseline</i>	38,89	13,10	44,05	62,70	67,06
Vazão acumulada diária	ARIMA	40,08	24,60	51,59	67,06	74,60
	Suavização Exponencial	34,92	24,70	50,39	67,06	69,44
	Média Móvel	36,00	26,40	52,40	65,20	72,40
	RF ( $k-1$ )	31,87	17,53	36,65	67,33	66,14
	RF ( $k-1, k-2, k-3$ )	36,55	24,50	43,37	66,27	67,87
	RF ( $k-1, k-2, \dots, k-7$ )	38,78	22,45	46,94	66,12	71,84
	RF ( $k-7$ )	25,11	18,18	19,48	42,04	58,44
	RF ( $k-7, k-14$ )	26,81	19,05	21,93	42,86	58,44
	RF ( $k-7, k-14, k-21$ )	27,71	19,91	21,65	45,45	63,68
	RF ( $k-1, k-2, \dots, k-10$ )	39,67	20,25	38,02	64,88	75,62
	RL ( $k-1, k-2, \dots, k-10$ )	37,19	27,16	51,65	69,01	75,21
	MLP ( $k-1, k-2, \dots, k-10$ )	39,26	26,75	47,52	64,88	73,55
	CFTS ( $k-1$ )	29,48	16,73	37,85	56,18	54,18
	WFTS ( $k-1$ )	33,07	17,13	38,25	66,53	67,73
	WHOFTS ( $k-1$ )	30,28	17,13	39,84	66,93	65,34
PWFTS <sup>3</sup> ( $k-1$ )	30,28	17,13	37,45	65,74	67,73	
Vazão acumulada das 0h às 10h59min + Vazão acumulada diária	ANFIS-SFC ( $k, k-1, \dots, k-7$ )	44,20	24,55	60,27	71,43	80,36
	ANFIS-FCM ( $k, k-1, \dots, k-7$ )	44,64	22,77	62,50	71,43	79,91
	Yamawawa-NFN ( $k, k-1, \dots, k-7$ )	44,20	20,09	56,25	70,98	79,46
	ANFIS-SFC ( $k, k-7, k-14, k-21, k-28$ )	44,80	25,00	61,61	70,09	80,36
	ANFIS-FCM ( $k, k-7, k-14, k-21, k-28$ )	45,74	22,77	62,50	71,43	80,36
	Yamakawa-NFN ( $k, k-7, k-14, k-21, k-28$ )	46,43	20,54	56,25	71,43	79,46
Vazão acumulada das 0h às 10h59min	Yamakawa-NFN ( $k, k-1, k-2, \dots, k-5$ )	44,53	20,65	56,28	67,61	79,35
	RF ( $k, k-1, k-2, \dots, k-5$ )	39,27	19,03	62,75	70,45	77,73
Vazão acumulada das 11 às 23h59min	RF ( $k-1$ )	31,47	9,96	40,24	56,97	63,35
	RF ( $k-1, k-2, k-3$ )	30,92	22,89	42,17	62,20	67,47
	RF ( $k-1, k-2, \dots, k-7$ )	35,92	24,08	44,90	59,18	69,39

Tabela 5.2 – Resultados do APE usando a abordagem de previsão de um passo à frente nos 5 pontos de distribuição.

## 5.2 Escolha de Três Modelos Invariantes no Tempo

Descarta-se uma função de ativação “arco de tangente”, por incluir uma não-linearidade desnecessária. Com e sem ruído branco (uma variável aleatória gaussiana  $\sim N(0, 1)$ ), poucas são as diferenças utilizando NFN, o que significa que adicionar a variável ruído se mostra infrutífero.

Pelas Tabelas 4.11, 4.12, 4.13, 4.14 e 4.15, vê-se que os modelos RL, RF e Yamakawa-NFN melhoram as estimativas em relação ao modelo *baseline* e a maioria das estimativas está no intervalo de APE inferior a 5%, em que não há multa. Percebe-se também que o modelo NFN fica em segundo lugar em todos os cinco pontos de entrega.

Os resultados apresentados nesta Seção são usados para gerar os modelos definitivos de estimativa de vazão acumulada da 0h às 23h59min para cada ponto de distribuição. Escolhe-se o modelo Yamakawa-NFN para os pontos A, C, D e E; entretanto, é decidido utilizar o modelo RF somente para o ponto B. Esses resultados justificam a iniciativa de introduzir um modelo adaptativo, não só nos pesos e nos valores modais, mas inclusive com adaptação de contexto.

## 5.3 Um Modelo Adaptativo

Pela Tabela 4.16, com o antes e o depois da implantação do modelo adaptativo<sup>4</sup>, vemos melhora em A, C, E, e ligeira piora em B e D, mas ainda mais preciso que o modelo *baseline*. Por motivos sigilosos, resolveu-se implementar dois modelos, um especial para fins de semana. Como se depreende da Figura 5.1, durante a pandemia a dinâmica se altera e o modelo adaptativo se ajusta melhor aos dados reais do que o invariante no tempo.



Figura 5.1 – Comparação entre dos dados reais, modelo sem adaptação e modelo adaptativo.

No período 01/01/2018 a 31/12/2021. Ddados reais (linha azul);  
Modelo sem adaptação (vermelho) e modelo adaptativo (verde).

O modelo adaptativo utiliza precisão matemática da biblioteca Python mpmath<sup>5</sup>, que garante os mesmos números dos fontes MatLab. São testadas várias larguras

<sup>3</sup> cf. Seção 2.6.13

<sup>4</sup> variante no tempo

<sup>5</sup> <https://mpmath.org/doc/current/>

de intervalo, até obter a ótima faixa de  $\pm 7,5\%$ . Os resultados mostram que, com exceção do ponto E, o modelo adaptativo consegue se ajustar aos novos dados sem pré-processamento.

#### 5.4 Fase de Testes

Os resultados são apresentados na Tabela 4.17 e mostram que os modelos propostos obtêm melhor desempenho em todas as bases de dados. Além disso, essa etapa mostra que os resultados da terceira etapa são corroborados por dados reais, isto é, a prática confirmou a teoria — é a denominada *feature* de adaptabilidade.

É interessante notar-se que, em relação as estimativas abaixo da faixa de erro de 5% do APE, em que não há multa, o modelo *baseline* obtém um percentual maior em E, o qual é melhorado com modelos distintos para semana ou para fim de semana. Por isso, também conclui-se que o uso de modelos específicos ajustam-se melhor aos dados de E.

#### 5.5 Previsão de Sete Dias à Frente

Conforme visto nas Tabelas 4.18, 4.19, 4.20, 4.21 e 4.22, em 2 / 5 pontos de distribuição, o modelo Yamakawa Adaptativo consegue melhor desempenho em todos os horizontes de previsão, demonstrando sua capacidade de se adaptar aos novos dados de validação.

Além disso, essas Tabelas mostram que o UKF proposto fica em primeiro ou segundo lugar na maioria dos casos,  $22 / 35 = 62,9\%$ , enquanto que os outros 13 casos, todos os 6 modelos são comparáveis, ficando em último lugar o *Naive Sazonal*.

Com os resultados intervalares obtidos, é possível imprimir o gráfico da Figura 5.2, que apresenta a previsão pontual e intervalar de sete passos à frente desse modelo usando dados normalizados gerados pelo sistema no dia 12/02/2024, segunda-feira, para o ponto B.

#### 5.6 Otimização de Parâmetros

Escolher percentuais é relativamente fácil, pois são apenas dois números no quadrado  $[0,1] \times [0,1]$ , para cada modelo Yamakawa. Porém, com os recursos computacionais disponíveis, tornou-se um processo demorado.

Nesta etapa, é feita uma ampla pesquisa no estado da arte sobre *Deep Fuzzy Systems*; descartaram-se modelos complicados, como um *Deep Fuzzy Cognitive*

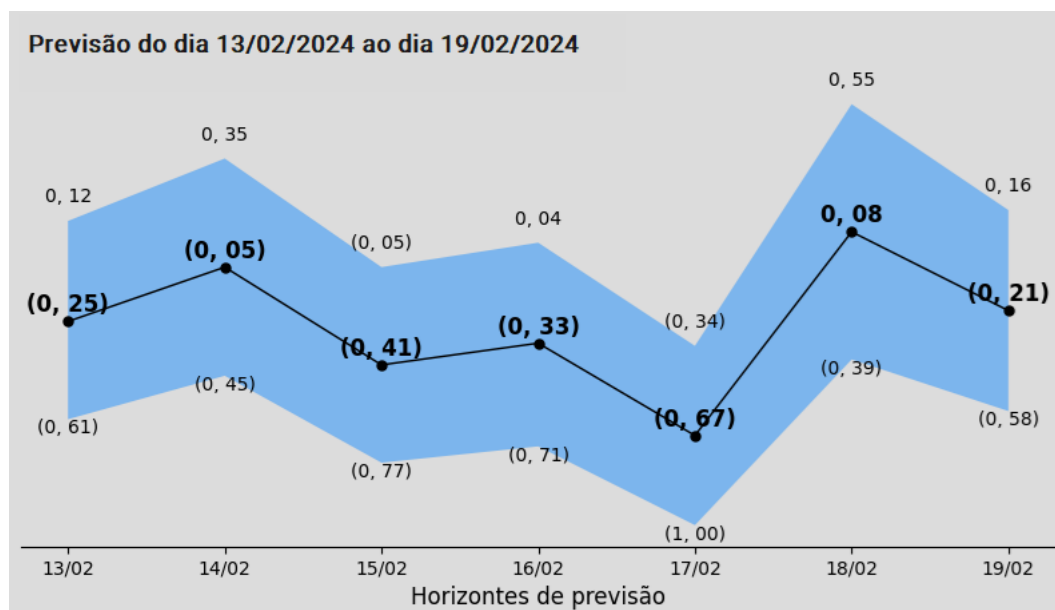


Figura 5.2 – Amostra comparativa da previsão pontual e intervalar do modelo Yamakawa Adaptativo para um horizonte de 7 dias. Previsão pontual (linha preta) vs. intervalar (faixa azul).

Maps (DFCM)<sup>6</sup> e um *ensemble* de dois DFCMs com um UKF.

## 5.7 Homologação

Conforme Tabela 4.23, vale a pena utilizar o UKF, pois os resultados são mais satisfatórios que sem o uso desse modelo, uma vez que apresentam mais acurácia, tornando a previsibilidade mais eficiente. O UKF ganha na maioria dos casos, e as outras faixas acima de 5%, 10%, 15% e 20% realizam os eventuais desempates. Tal modelo faz com que 2 séries independentes se tornem interdependentes, o que é a denominada *feature* de regularização de dados.

Os cinco pontos de entrega são validados pela GASMIG, e tudo está bem avaliado, como esperado, exceto o ponto D. Nesse ponto, não ocorre conforme previsto, e imaginamos que os valores de entrada são atípicos, isto é, uma demanda de GN a mais que o modelo entendeu como *outlier*.

## 5.8 Previsão de Múltiplos Passos à Frente

Essa etapa é detalhada no trabalho Ferraz et al. (2024), em que analisou-se a adaptabilidade (*fit* em um ano, *predict* em seis meses), do modelo Yamakawa (1992) em comparação com outros seis modelos.

<sup>6</sup> <https://github.com/FieldDoctor/DFCM>

O objetivo dos experimentos é avaliar diferentes modelos estatísticos e de aprendizado de máquina na previsão diária de volume de GN em três horizontes de previsão. Os resultados obtidos mostraram que modelos clássicos como o *NAIVE* e *ARIMA* apresentam resultados estatisticamente iguais aos modelos *WHOFTS*, *NFN*, *MLP* e *RF*. Esse resultado também é percebido em termos de acurácia, em que o *NAIVE* e *ARIMA* se sobressaem em duas das três bases de dados. O *LSTM* apresenta resultados bem inferiores aos demais modelos. Isso pode ser explicado por ser um modelo de aprendizado profundo, que depende de um conjunto grande de dados para treinamento.

Por fim, além dos experimentos mostrarem uma comparação entre modelos, é importante perceber que o problema de previsão de volume de GN não é um problema trivial, pois, no geral, os modelos apresentam baixa acurácia dentro da faixa aceitável de erro.

## Conclusão

O objetivo deste trabalho é conduzir um estudo de previsão diária de 1, 4 e 7 passos à frente para três diferentes estações de recebimento de GN da *GASMIG*. Para isso, são avaliados modelos clássicos, de aprendizado de máquina e de aprendizado profundo. Os resultados indicaram que os modelos de aprendizado de máquina apresentam desempenho semelhante aos modelos clássicos de previsão, como *NAIVE* e *ARIMA*.

## 5.9 Série Temporal Semanal

Verifica-se que o modelo *NFN* supera os modelos selecionados. Vê-se nas Figuras 5.3 e 5.4 — em que estão as visões temporais — que os dados estimados estão bem próximos dos reais (pois é considerada apenas a acurácia). Por outro lado, como são poucos os pontos de validação, as imagens revelam captação mais das tendências do que das dinâmicas.

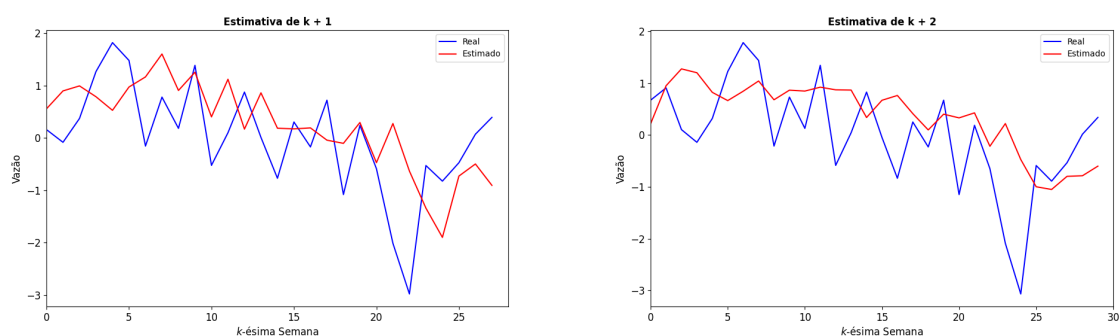


Figura 5.3 – Resultados na validação para as variáveis semanais — valores reais vs. valores estimados nos horizontes 1 e 2 passos à frente.

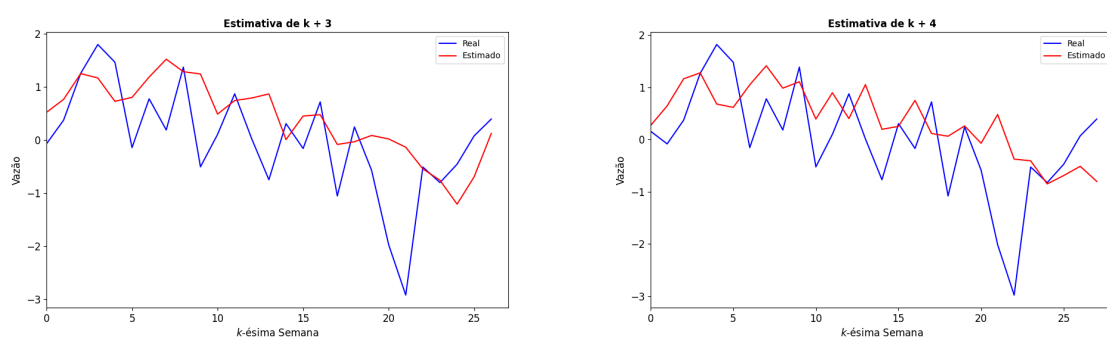


Figura 5.4 – Resultados na validação para as variáveis semanais — valores reais vs. valores estimados nos horizontes 3 e 4 passos à frente.

Nas séries semanais, o Yamakawa-NFN ganha nos quatro horizontes de previsão, pelas Tabelas 4.27 e 4.28,  $k+1$  com 4 variáveis a 78,57% de acerto;  $k+2$  com 1 variável a 76,67%;  $k+3$  com 3 variáveis a 81,48% e  $k+4$  com 1 variável a 82,14%. Os histogramas da Figura 4.9 detalham o desempenho do NFN nesses casos.

## 6 CONSIDERAÇÕES FINAIS

Foi feita uma revisão teórica que destaca o papel do GN na matriz energética brasileira e sua importância para o país, incluindo dados sobre reservas mundiais e consumo. Foram discutidos modelos de previsão de consumo de médio e longo prazo, assim como os de curto prazo, que são mais relevantes para o objetivo do trabalho.

Foi explicado que o GN é uma fonte de energia fóssil limpa e que a indústria do GN é dividida em produção, processamento, transmissão, armazenamento e distribuição, com a regulação sendo feita pelo governo ou por meio de contratos de fornecimento (SOLDO et al., 2014).

Também foi visto que a previsão precisa do consumo de GN é essencial para planejamento de curto prazo por empresas distribuidoras, devido a questões regulatórias e para evitar perdas econômicas (KHAN, 2015). A previsão é desafiadora devido à complexidade do consumo e aos diversos fatores que influenciam, como parâmetros econômicos, dados meteorológicos e calendário (LIU et al., 2021). A seguir, foi feita uma introdução às redes NFN e ao UKF, que são os precursores do modelo proposto. Mais à frente, a parametrização de vários modelos foi discutida, incluindo Regressão Linear, Média Móvel, Naive, ARIMA, RF, Redes Neurais Artificiais, ANFIS com FCM, ANFIS com SFC, Suavização Exponencial e modelos de série temporal nebulosa de alta ordem.

Esse estudo baseou-se nas seguintes metodologias: a análise preliminar dos dados, escolha das variáveis, definição das entradas e saídas, autocorrelações, *boxplots*, o critério de Akaike para seleção de variáveis e a parametrização de vários modelos.

Durante o desenvolvimento dessa dissertação, foram analisados os dados em cinco pontos de entrega. Primeiramente, foi feita comparação entre previsão vespertina e previsão do dia seguinte. Seguida da escolha de três modelos invariantes no tempo: RF, RL e Yamakawa-NFN. Em continuidade, foi escolhido um modelo adaptativo: UKF + NFN. Na fase de testes, foi aprimorado o tratamento dos dados e foram propostos modelos específicos, somente para o ponto E. A seguir, a previsão foi ampliada até sete dias à frente, concluindo-se que o modelo proposto é comparável com os da literatura. Por sua vez, otimizaram-se os parâmetros constantes de projeto, e na homologação, concluiu-se que com UKF é melhor do que sem UKF. Porém, ao analisar o comportamento anual, considerando-se dois semestres de treino e um semestre de validação, concluiu-se que sem UKF tornava-se melhor do que com ele. Por fim, concluiu-se que o modelo proposto é comparável também para séries temporais

semanais.

Ao analisar e comparar metodologias, concluiu-se que:

(1) Os modelos propostos apresentaram melhorias em relação ao modelo *baseline*, uma quantidade maior de dias abaixo de 5% do APE, em que não há multa, e um número menor de dias acima de 20% do APE, que é a maior faixa de multa, tornando possível economizar gastos.

(2) Dependendo das condições de treinamento e validação, vale a pena incluir o UKF no modelo.

(3) A menor quantidade de variáveis deixou o modelo com mais acurácia, por isso “menos é mais”.

(4) Com adaptação de valores modais (KOHONEN, 1990), é melhor do que com valores modais equidistantes.

(5) Com adaptação dos parâmetros, é melhor do que com modelo invariante no tempo.

(6) A previsão intervalar é melhor do que a pontual.

(7) Quando não se consideram as multas, ou seja, quando se altera a métrica de erro, mudam as conclusões.

(8) Prever a vazão de GN para consumidores industriais não é uma tarefa simples e há uma lacuna de pesquisa na exploração desse tema com as especificidades dos tipos de consumidores no Brasil. Outro desafio mencionado é a falta de disponibilidade de bases de dados públicas que possam auxiliar nesses estudos, dado o caráter empresarial dos dados envolvidos.

(9) Sempre que houver duas ou mais séries temporais, esta dissertação se aplica. Cada uma das séries temporais avança independentemente, depois tudo se reúne em um UKF, que filtra as médias e os desvios-padrão.

Para concluir, o algoritmo de Yamakawa (1992), apesar de sua simplicidade e seu baixo custo computacional, apresentou-se desempenho em média superior aos métodos comumente utilizados para previsão de séries temporais, de custo computacional superior; investigou-se a previsão de consumo de GN no Brasil com foco no curto prazo: previsão nas próximas 13 horas da mesma data, previsão de um a sete passos diários à frente, além da previsão de um a quatro passos semanais à frente. Assim, dado um modelo *baseline*, foi possível melhorá-lo consideravelmente.

## 6.1 Trabalhos Futuros

Considerando a importância do GN como matriz energética, e a relevância do tema, essa problemática não se encerra aqui. Há muitos desdobramentos possíveis, na abertura de perspectiva para trabalhos futuros que venham a preencher a lacuna existente na literatura. Nesse sentido, seria importante estudar:

(A) Primeiramente, é necessário otimizar o tratamento dos **outliers**, no momento em que eles chegam, isto é, no dia-a-dia. (B) Refinamento do UKF: Controle do traço da matriz  $P^a$  de covariâncias. (C) Refinamento ou estudo de um *ensemble* com o modelo *Naive*. (D) **Variáveis Mensais**: Previsão do somatório mensal, que seria uma extensão do semanal.

(E) **Sazonalidade**: Discussão sobre a necessidade de considerar a sazonalidade em modelos de previsão, especialmente para previsões de curto prazo. (F) **Escalabilidade e Adaptabilidade**: Exploração da escalabilidade e adaptabilidade do modelo proposto para diferentes horizontes e cenários de previsão. (G) **Otimização Multiobjetivo**: Custo mínimo e benefício máximo, ao mesmo tempo, são exemplos de duas funções objetivo.

## REFERÊNCIAS

- AGUIRRE, L. *Introdução à Identificação de Sistemas – Técnicas Lineares e Não-Lineares Aplicadas a Sistemas Reais*. Editora UFMG, 2007. ISBN 9788570415844. Disponível em: <https://books.google.com.br/books?id=f9lwE7Ph0fYC>.
- ANAGNOSTIS, A.; PAPAGEORGIOU, E.; BOCHTIS, D. Application of artificial neural networks for natural gas consumption forecasting. *Sustainability*, MDPI AG, v. 12, n. 16, p. 6409, ago. 2020. Disponível em: <https://doi.org/10.3390/su12166409>.
- ANDELKOVIC, A. S.; BAJATOVIC, D. Integration of weather forecast and artificial intelligence for a short-term city-scale natural gas consumption prediction. *Journal of Cleaner Production*, Elsevier BV, v. 266, p. 122096, set. 2020. Disponível em: <https://doi.org/10.1016/j.jclepro.2020.122096>.
- BAZARAA, M. S.; SHERALI, H. D.; SHETTY, C. *Nonlinear programming: Theory and algorithms*, John Wiley & Sons, Inc., New York, 1993.
- BEHROUZNIYA, A. et al. An adaptive network based fuzzy inference system-fuzzy data envelopment analysis for gas consumption forecasting and analysis: The case of south america. *IEEE*, jun. 2010. Disponível em: <https://doi.org/10.1109/icias.2010.5716160>.
- BEYCA, O. F. et al. Using machine learning tools for forecasting natural gas consumption in the province of istanbul. *Energy Economics*, North-Holland, v. 80, p. 937–949, 5 2019. ISSN 0140-9883.
- CAMINHAS, W.; GOMIDE, F. A fast learning algorithm for neofuzzy networks. *8th IPMU*, n. 1, p. 1784–1790, 2000.
- CHEN, S.-M. Forecasting enrollments based on fuzzy time series. *Fuzzy Sets and Systems*, Elsevier BV, v. 81, n. 3, p. 311–319, ago. 1996. Disponível em: [https://doi.org/10.1016/0165-0114\(95\)00220-0](https://doi.org/10.1016/0165-0114(95)00220-0).
- COSTA, O. L. V. et al. A robust least square approach for forecasting models: an application to brazil natural gas demand. *Energy Systems*, Springer Science and Business Media LLC, v. 11, n. 4, p. 1111–1135, ago. 2019. Disponível em: <https://doi.org/10.1007/s12667-019-00351-1>.
- DEMŠAR, J. Statistical comparisons of classifiers over multiple data sets. *The Journal of Machine learning research*, v. 7, n. 1, p. 1–30, 2006.
- DU, J. et al. A hybrid deep learning framework for predicting daily natural gas consumption. *Energy*, Elsevier BV, v. 257, p. 124689, out. 2022. Disponível em: <https://doi.org/10.1016/j.energy.2022.124689>.
- EPE. Atlas de eficiência energética do brasil. 2022. Disponível em: [https://www.epe.gov.br/sites-pt/publicacoes-dados-abertos/publicacoes/PublicacoesArquivos/publicacao-741/Atlas\\_Eficiencia\\_Energetica\\_Brasil\\_2022.pdf](https://www.epe.gov.br/sites-pt/publicacoes-dados-abertos/publicacoes/PublicacoesArquivos/publicacao-741/Atlas_Eficiencia_Energetica_Brasil_2022.pdf).

- FAN, G. F.; WANG, A.; HONG, W. C. Combining grey model and self-adapting intelligent grey model with genetic algorithm and annual share changes in natural gas demand forecasting. *Energies* 2018, Vol. 11, Page 1625, Multidisciplinary Digital Publishing Institute, v. 11, p. 1625, 6 2018. ISSN 1996-1073. Disponível em: <https://www.mdpi.com/1996-1073/11/7/1625/html><https://www.mdpi.com/1996-1073/11/7/1625>.
- FERRAZ, V. C. et al. Previsão de séries temporais de consumo diário de gás natural no brasil. Congresso Brasileiro de Automática (CBA). No prelo, 2024.
- GAO, F.; SHAO, X. Forecasting annual natural gas consumption via the application of a novel hybrid model. *Environmental Science and Pollution Research*, Springer Science and Business Media Deutschland GmbH, v. 28, p. 21411–21424, 5 2021. ISSN 16147499. Disponível em: <https://link.springer.com/article/10.1007/s11356-020-12275-w>.
- GASNET. Gasnet — o site do gás natural. 2006. Disponível em: <https://gasnet.com.br/>.
- GOMES, A. C. d. O. et al. Estudo da utilização do gás natural como fonte geradora de energia no brasil. 2014.
- GOODWIN, P. et al. The holt-winters approach to exponential smoothing: 50 years old and going strong. *Foresight*, Citeseer, v. 19, n. 19, p. 30–33, 2010.
- GRACIAS, A. C.; LOURENÇO, S. R.; RAFIKOV, M. Estimation of natural gas production, import and consumption in brazil based on three mathematical models. *Natural Resources*, Scientific Research Publishing, v. 3, n. 2, p. 42–47, 2012.
- HOLT, C. C. Forecasting seasonals and trends by exponentially weighted moving averages. *International journal of forecasting*, Elsevier, v. 20, n. 1, p. 5–10, 2004.
- IEA. *Natural gas explained*. [S.l.], 2022. Disponível em: <https://www.eia.gov/energyexplained/natural-gas/>. Disponível em: <https://www.eia.gov/energyexplained/natural-gas/>.
- KALMAN, R. E. A new approach to linear filtering and prediction problems. 1960.
- KHAN, M. A. Modelling and forecasting the demand for natural gas in pakistan. *Renewable and Sustainable Energy Reviews*, Elsevier BV, v. 49, p. 1145–1159, set. 2015. Disponível em: <https://doi.org/10.1016/j.rser.2015.04.154>.
- KOHONEN, T. Improved versions of learning vector quantization. In: *1990 IJCNN International Joint Conference on Neural Networks*. [S.l.: s.n.], 1990. p. 545–550 vol.1.
- LIU, H.; WANG, L. Latest developments of gesture recognition for human–robot collaboration. In: *Advanced Human-Robot Collaboration in Manufacturing*. [S.l.]: Springer, 2021. p. 43–68.
- LIU, J. et al. Natural gas consumption forecasting: A discussion on forecasting history and future challenges. *Journal of Natural Gas Science and Engineering*, Elsevier, v. 90, p. 103930, 6 2021. ISSN 1875-5100.

- LUCAS, P. d. O. e.; SILVA, P. C. d. L.; GUIMARÃES, F. G. Otimização Evolutiva de Hiperparâmetros para Modelos de Séries Temporais Nebulosas. In: *Anais do 14<sup>o</sup> Simpósio Brasileiro de Automação Inteligente*. Ouro Preto: [s.n.], 2019. p. 1 – 7.
- MEIRA, E.; OLIVEIRA, F. L. C.; MENEZES, L. M. de. Forecasting natural gas consumption using bagging and modified regularization techniques. *Energy Economics*, Elsevier B.V., v. 106, 2 2022. ISSN 01409883.
- MERKEL, G. D.; POVINELLI, R. J.; BROWN, R. H. Short-term load forecasting of natural gas with deep neural network regression †. *Energies 2018, Vol. 11, Page 2008*, Multidisciplinary Digital Publishing Institute, v. 11, p. 2008, 8 2018. ISSN 1996-1073. Disponível em: <https://www.mdpi.com/1996-1073/11/8/2008/htmhttps://www.mdpi.com/1996-1073/11/8/2008>.
- MICHELENA, B. D. Modelo temporal da oferta e demanda do gás natural liquefeito no brasil. Escola Politécnica, 2019.
- QIAO, W. et al. A novel hybrid prediction model for hourly gas consumption in supply side based on improved whale optimization algorithm and relevance vector machine. *IEEE Access*, Institute of Electrical and Electronics Engineers (IEEE), v. 7, p. 88218–88230, 2019. Disponível em: <https://doi.org/10.1109/access.2019.2918156>.
- RAVNIK, J.; HRIBERSEK, M. A method for natural gas forecasting and preliminary allocation based on unique standard natural gas consumption profiles. *Energy*, Pergamon, v. 180, p. 149–162, 8 2019. ISSN 0360-5442.
- REHMAN, A. et al. Assessment of deep learning and classical statistical methods on forecasting hourly natural gas demand at multiple sites in Spain. *Energy*, Elsevier BV, v. 244, p. 122562, abr. 2022. Disponível em: <https://doi.org/10.1016/j.energy.2021.122562>.
- SANTOS, B. M. d. Projeção de demanda de gás natural no mercado brasileiro utilizando método estatístico. 2019.
- SEN, D.; GUNAY, M. E.; TUNC, K. M. Forecasting annual natural gas consumption using socio-economic indicators for making future policies. *Energy*, Pergamon, v. 173, p. 1106–1118, 4 2019. ISSN 0360-5442.
- SHAIKH, F.; JI, Q. Forecasting natural gas demand in China: Logistic modelling analysis. *International Journal of Electrical Power & Energy Systems*, Elsevier, v. 77, p. 25–32, 5 2016. ISSN 0142-0615.
- SILVA, A. M. et al. A fast learning algorithm for evolving neo-fuzzy neuron. *Applied Soft Computing*, v. 14, p. 194–209, 2014. ISSN 1568-4946. Disponível em: <https://www.sciencedirect.com/science/article/pii/S1568494613001373>.
- SILVA, P. C. L. *Scalable Models for Probabilistic Forecasting with Fuzzy Time Series*. Tese (Doutorado) — UFMG, 09 2019. Disponível em: <http://doi.org/10.5281/zenodo.3374641>.
- SOLDO, B. et al. Improving the residential natural gas consumption forecasting models by using solar radiation. *Energy and Buildings*, Elsevier BV, v. 69, p. 498–506, fev. 2014. Disponível em: <https://doi.org/10.1016/j.enbuild.2013.11.032>.

SOUZA, J. H. de et al. Prognóstico de falhas utilizando neo-fuzzy-neuron com aprendizado on-line. Universidade Federal de Minas Gerais, 2020.

SU, H. et al. A hybrid hourly natural gas demand forecasting method based on the integration of wavelet transform and enhanced deep-RNN model. *Energy*, Elsevier BV, v. 178, p. 585–597, jul. 2019. Disponível em: <https://doi.org/10.1016/j.energy.2019.04.167>.

SVOBODA, R.; KOTIK, V.; PLATOS, J. Short-term natural gas consumption forecasting from long-term data collection. *Energy*, Elsevier BV, v. 218, p. 119430, mar. 2021. Disponível em: <https://doi.org/10.1016/j.energy.2020.119430>.

SZOPLIK, J. Forecasting of natural gas consumption with artificial neural networks. *Energy*, Pergamon, v. 85, p. 208–220, 6 2015. ISSN 0360-5442.

WAN, E. A.; MERWE, R. V. D. The unscented kalman filter. *Kalman filtering and neural networks*, Wiley Online Library, p. 221–280, 2001.

WINTERS, P. R. Forecasting sales by exponentially weighted moving averages. *Management science*, INFORMS, v. 6, n. 3, p. 324–342, 1960.

YAGER, R. R.; FILEV, D. P. Generation of fuzzy rules by mountain clustering. *Journal of Intelligent & Fuzzy Systems*, IOS Press, v. 2, n. 3, p. 209–219, 1994.

YAMAKAWA, T. A neo fuzzy neuron and its applications to system identification and prediction of the system behavior. In: *Proc. of the 2nd Int. Conf. on Fuzzy Logic & Neural Networks*. [S.l.: s.n.], 1992. p. 477–483.

YU, H.-K. Weighted fuzzy time series models for TAIEX forecasting. *Physica A: Statistical Mechanics and its Applications*, Elsevier BV, v. 349, n. 3-4, p. 609–624, abr. 2005. Disponível em: <https://doi.org/10.1016/j.physa.2004.11.006>.

## A APÊNDICE — Códigos-Fonte

Neste documento, são anexados trechos dos códigos-fonte.

### A.1 Regressão Linear (RL)

```
from sklearn.linear_model import LinearRegression
model = LinearRegression().fit(X_train, y_train)
```

### A.2 Média Móvel

```
yhat = df.rolling(window).mean()
if window != 1:
    yhat.dropna(inplace=True)
y_test = df.loc[window:df.shape[0]]
y_test = y_test.values.tolist()
yhat = yhat.values.tolist()
```

### A.3 Naive

```
forecast = np.zeros((0,7))
for i in range(0,diariot['vazao'].values.shape[0]):
forecast = np.vstack((forecast,
    diariot['vazao'].values[i] * np.ones((1,7))))
```

#### A.3.1 Naive Sazonal

```
def completar(x):
    x = np.array(x)
    while np.shape(x)[0] < 7:
        x = np.hstack((x[0], x))
    return np.reshape(x, (1,7))

forecast = np.zeros((0,7))
for i in range(0,diariot['vazao'].values.shape[0]):
forecast = np.vstack((forecast,
    completar(diariot['vazao'].values[max([i-6,0]):i+1])))
```

### A.4 Auto-regressivo integrado de médias móveis (ARIMA)

```
model = auto_arima(trainT, start_p=1, start_q=1, max_p=10, max_q=10, m=7,
```

```

    start_P=1, seasonal=True, %False,
    d=1, D=1, trace=True, error_action='ignore', stepwise=True)
history = [x for x in trainT]
yhat = np.zeros((testT.shape[0]))
tt = 0
model_fit = model.fit(history)
for t in range(len(testT)-1):
    yhat[t] = model_fit.predict(n_periods=1)
    history.append(testT[tt])
    model_fit = model.fit(history)
    tt = t + 1

```

## A.5 Random Forest (RF)

```

params = {
'bootstrap': [False], 'max_depth': [5,10,15,20],
'max_features': [3,5,7,9,10], 'min_samples_leaf': [2,3,4],
'min_samples_split': [4,5,6,7], 'n_estimators': [1,2,3,4,5,7,10]
}
model = GridSearchCV(estimator = ensemble.RandomForestRegressor(),
    param_grid = params, cv = 3, n_jobs = -1)
model.fit(X_train, y_train)

```

## A.6 Redes Neurais Artificiais

```

from sklearn.neural_network import MLPRegressor
def gridsearch_MLP(X, Y):
    estimator=MLPRegressor()
    param_grid = {'hidden_layer_sizes': [(5,5,1),(5,10,5,1), (10,10,1)],
        'activation': ['relu','tanh','logistic'], 'alpha': [0.0001, 0.05],
        'solver': ['adam']}
    gsc = GridSearchCV(estimator, param_grid,
        cv=5, scoring='neg_mean_squared_error', n_jobs=-1)
    grid_result = gsc.fit(X, Y)
    best_params = grid_result.best_params_
    return best_params

def fit_MLP(X_train, y_train):
    best_params = gridsearch_MLP(X_train, y_train)
    model = MLPRegressor(hidden_layer_sizes=
        best_params["hidden_layer_sizes"],
        activation =best_params["activation"],
        solver=best_params["solver"], random_state=1, max_iter=500,
        n_iter_no_change = 200).fit(X_train, y_train)
    return model

```

## A.7 ANFIS com *Fuzzy C-Means Clustering* (ANFIS-FCM)

```
options = genfisOptions('FCMClustering',
    'NumClusters', 5); % default = auto
    % 'Exponent', 10.0); % '2.0 (default)
    % 'FISType', "mamdani",... % 'sugeno' (default)
in_fis = genfis(xt,ydt,options);
ys=evalfis(in_fis,xv);
```

## A.8 ANFIS com *Subtractive Fuzzy Clustering* (ANFIS-SFC)

```
options = genfisOptions('SubtractiveClustering', % função do MatLab
    'ClusterInfluenceRange', 0.2); % default = 0.5
    % DataScale = "auto" (default)
    % SquashFactor = 1.25 (default)
    % AcceptRatio = 0.5 (default)
    % RejectRatio = 0.15 (default)
in_fis = genfis(xt,ydt,options);
ys=evalfis(in_fis,xv);
```

## A.9 Suavização Exponencial

```
history = [x for x in trainT]
yhat = np.zeros((testT.shape[0]))
tt = 0
level = 0.9
model_fit = SimpleExpSmoothing(history).fit(smoothing_level = level)
for t in range(len(testT)-1):
    yhat[t] = model_fit.forecast(1)
    history.append(testT[tt])
    model_fit = SimpleExpSmoothing(history).fit(smoothing_level = level)
    tt = t + 1
```

## A.10 *First Order Conventional* (CFTS)

```
from pyFTS.models import chen, yu
from pyFTS.partitioners import Grid
p = [30,35,40,45,50,55,60]
for i in p:
    partitioner = Grid.GridPartitioner(data=diario.values,npart=i)
    model1 = chen.ConventionalFTS(partitioner=partitioner)
    model1.fit(diario.values, save_model=True, order=1)
    forecasts = model1.predict(diario.values)
```

**A.11 First Order Weighted (WFTS)**

```
p = [30,35,40,45,50,55,60]
for i in p:
    partitioner = Grid.GridPartitioner(data=diario.values,npart=i)
    model1 = yu.WeightedFTS(partitioner=partitioner)
    model1.fit(diario.values, save_model=True, order=1)
    forecasts = model1.predict(diario.values)
```

**A.12 Weighted High Order (WHOFTS)**

```
model = pwfts.ProbabilisticWeightedFTS(partitioner=partitioner,lags=[1],
    alpha_cut=0.1, order =10)
model.fit(diario['vazao'].values, save_model=True,file_path='model_'+dataset)
```

**A.13 High Order Probabilistic Weighted (PWFTS)**

```
model = pwfts.ProbabilisticWeightedFTS(partitioner=partitioner,lags=[1],
    alpha_cut=0.1, order =10)
model.fit(diario['vazao'].values, save_model=True,file_path='model_PWFTS_'+dataset)
```