



CENTRO FEDERAL DE EDUCAÇÃO TECNOLÓGICA DE MINAS GERAIS
PROGRAMA DE PÓS-GRADUAÇÃO EM MODELAGEM MATEMÁTICA E COMPUTACIONAL

UMA ABORDAGEM *Fuzzy* EVOLUTIVA COM APRENDIZAGEM PARTICIPATIVA E FUNÇÕES GAUSSIANA MULTIVARIÁVEL

FERNANDA PEREIRA DOS SANTOS RODRIGUES

Orientador: Prof. Alisson Marques da Silva
CEFET-MG

BELO HORIZONTE
JANEIRO DE 2022

FERNANDA PEREIRA DOS SANTOS RODRIGUES

**UMA ABORDAGEM *Fuzzy* EVOLUTIVA COM
APRENDIZAGEM PARTICIPATIVA E FUNÇÕES
GAUSSIANA MULTIVARIÁVEL**

Dissertação apresentada ao Programa de Pós-Graduação em Modelagem Matemática e Computacional do Centro Federal de Educação Tecnológica de Minas Gerais, como requisito parcial para a obtenção do título de Mestre em Modelagem Matemática e Computacional.

Área de concentração: Modelagem Matemática e Computacional

Linha de pesquisa: Sistemas Inteligentes

Orientador: Prof. Alisson Marques da Silva
CEFET-MG

CENTRO FEDERAL DE EDUCAÇÃO TECNOLÓGICA DE MINAS GERAIS
PROGRAMA DE PÓS-GRADUAÇÃO EM MODELAGEM MATEMÁTICA E COMPUTACIONAL
BELO HORIZONTE
JANEIRO DE 2022

R696a Rodrigues, Fernanda Pereira dos Santos
Uma abordagem Fuzzy evolutiva com aprendizagem participativa e
funções gaussianas multivariáveis / Fernanda Pereira dos Santos Rodrigues. –
2022.
57 f.

Dissertação de mestrado apresentada ao Programa de Pós-Graduação em
Modelagem Matemática e Computacional.

Orientador: Alisson Marques da Silva.

Dissertação (mestrado) – Centro Federal de Educação Tecnológica de
Minas Gerais.

1. Sistemas difusos – Teses. 2. Computação evolutiva – Teses. 3. Teoria da
previsão – Teses. 4. Identificação do sistema – Teses. 5. Sistemas não lineares
– Teses. I. Silva, Alisson Marques da. III. Centro Federal de Educação
Tecnológica de Minas Gerais. IV. Título.

CDD 006.3



SERVIÇO PÚBLICO FEDERAL
MINISTÉRIO DA EDUCAÇÃO
CENTRO FEDERAL DE EDUCAÇÃO TECNOLÓGICA DE MINAS GERAIS
COORDENAÇÃO DO CURSO DE MESTRADO EM MODELAGEM MATEMÁTICA E COMPUTACIONAL

**“UMA ABORDAGEM FUZZY EVOLUTIVA COM APRENDIZAGEM
PARTICIPATIVA E FUNÇÕES GAUSSIANA MULTIVARIÁVEL”**

Dissertação de Mestrado apresentada por **Fernanda Pereira dos Santos Rodrigues**, em 15 de fevereiro de 2022, ao Programa de Pós-Graduação em Modelagem Matemática e Computacional do CEFET-MG, e aprovada pela banca examinadora constituída pelos professores:

Prof. Dr. Alisson Marques da Silva (Orientador)
Centro Federal de Educação Tecnológica de Minas Gerais

Prof. Dr. André Paim Lemos
Universidade Federal de Minas Gerais

Prof^ª. Dr^ª. Elisângela Martins de Sá
Centro Federal de Educação Tecnológica de Minas Gerais

Prof. Dr. Flávio Vinícius Cruzeiro Martins
Centro Federal de Educação Tecnológica de Minas Gerais

Visto e permitida a impressão,

Prof^ª. Dr^ª. Elizabeth Fialho Wanner
Presidenta do Colegiado do Programa de Pós-Graduação em
Modelagem Matemática e Computacional

Resumo

Este trabalho propõe uma abordagem *fuzzy* evolutiva construída sob um algoritmo de agrupamento recursivo não supervisionado com aprendizagem participativa e funções de pertinência Gaussiana multivariável. A partir dessa abordagem, são introduzidos quatro modelos que se diferenciam pelo método de atualização dos parâmetros do consequente das regras *fuzzy* e são denominados eFLS (*evolving Fuzzy with Multivariable Gaussian Participatory Learning and Weighted Recursive Least Squares*), eFMI (*evolving Fuzzy with Multivariable Gaussian Participatory Learning and Multi-Innovations Recursive Weighted Least Squares*), eFCE (*evolving Fuzzy with Multivariable Gaussian Participatory Learning and Recursive Maximum Correntropy*) e eFTLS (*evolving Fuzzy with Multivariable Gaussian Participatory Learning and Recursive Weighted Total Least Squares*). Na abordagem proposta, os grupos são criados usando uma medida de compatibilidade e um mecanismo de alerta. A medida de compatibilidade é calculada pela distância Euclidiana ou Mahalanobis de acordo com o número de amostras do grupo. O método de exclusão de grupos combina idade e população para excluir grupos inativos. Os grupos redundantes são mesclados se houver uma sobreposição notável entre dois grupos. Os parâmetros do consequente são atualizados por um algoritmo recursivo de mínimos quadrados ponderados, mínimos quadrados ponderados com multi-inovações, máxima correntropia ou mínimos quadrados totais ponderados. O desempenho dos modelos é avaliado e comparado com modelos alternativos do estado da arte na previsão de séries temporais e na identificação de sistemas não lineares. Os experimentos computacionais e as comparações sugerem que os modelos propostos possuem um desempenho comparável ou superior ao dos modelos alternativos.

Palavras-chave: Sistemas Evolutivos, Sistemas *Fuzzy*, Previsão, Identificação de Sistemas.

Abstract

This work introduces an evolving fuzzy approach constructed based on a non-supervised recursive clustering algorithm with participatory learning and multivariate Gaussian membership functions. From this approach, four models are introduced, called eFLS (evolving Fuzzy with Multivariable Gaussian Participatory Learning and Weighted Recursive Least Squares), eFMI (evolving Fuzzy with Multivariable Gaussian Participatory Learning and Multi-Innovations Recursive Weighted Least Squares), eFCE (evolving Fuzzy with Multivariable Gaussian Participatory Learning and Recursive Maximum Correntropy) and eFTLS (evolving Fuzzy with Multivariable Gaussian Participatory Learning and Recursive Weighted Total Least Squares). In the proposed approach, clusters are created using a compatibility measure and an alert mechanism. The compatibility measure is computed by Euclidian or Mahalanobis distance according to the number of samples in the cluster. Inactive clusters are excluded by an age and population based-method. Redundant clusters are merged whenever is a noticeable overlap between two clusters. The consequent parameters are updated by a recursive algorithm of Weighted Least Squares, Multi-Innovations Weighted Least Squares, Maximum Correntropy, or Weighted Total Least Squares. Finally, the performance of the models is evaluated and compared with alternative state-of-the-art models in time series forecasting and non-linear system identification problems. Computational experiments and comparisons suggest that the proposed models have superior or comparable performance to alternative models.

Keywords: Evolving Systems, Fuzzy Systems, Forecasting, System Identification.

Lista de Figuras

Figura 1 – Estrutura básica de um Sistema <i>Fuzzy</i>	7
Figura 2 – Fluxograma da abordagem proposta.	14
Figura 3 – Identificação da função de hipérbole 7-D pelo eFCE.	33
Figura 4 – Soma dos quadrados residuais na identificação da função de hipérbole - 7-D.	34
Figura 5 – Identificação da função de hipérbole 10-D pelo eFCE.	35
Figura 6 – Soma dos quadrados residuais na identificação da função de hipérbole - 10-D.	36
Figura 7 – Previsão da concentração de CO_2 pelo eFCE.	37
Figura 8 – Logaritmo da soma dos quadrados residuais na previsão da concentração de CO_2	38
Figura 9 – Previsão da produção de energia elétrica na CCPP pelo eFCE.	40
Figura 10 – Logaritmo da soma dos quadrados residuais na previsão da produção de energia elétrica na CCPP.	41
Figura 11 – Previsão de temperatura no Vale da Morte pelo eFCE.	43
Figura 12 – Soma dos quadrados residuais na previsão de temperatura no Vale da Morte.	44
Figura 13 – Previsão do preço dos títulos do Global 40 pelo eFMI.	45
Figura 14 – Logaritmo da soma dos quadrados residuais na previsão dos preços dos títulos do Global 40.	46
Figura 15 – Previsão do aluguel de bicicletas pelo eFMI.	47
Figura 16 – Soma dos quadrados residuais na previsão do aluguel de bicicletas.	48

Lista de Tabelas

Tabela 1 – Características dos modelos evolutivos.	12
Tabela 2 – Faixa e melhores valores dos parâmetros dos modelos.	30
Tabela 3 – Desempenho na identificação da função de hipérbole - 7-D.	33
Tabela 4 – Avaliação dos modelos propostos pelo teste MGN na identificação da função de hipérbole - 7-D.	34
Tabela 5 – Desempenho na identificação da função de hipérbole - 10-D.	35
Tabela 6 – Avaliação dos modelos propostos pelo teste MGN na identificação da função de hipérbole - 10-D.	36
Tabela 7 – Desempenho na previsão da concentração de CO_2	37
Tabela 8 – Avaliação dos modelos propostos pelo teste MGN na previsão da concentração de CO_2	39
Tabela 9 – Desempenho na previsão da produção de energia elétrica na CCPP.	40
Tabela 10 – Avaliação dos modelos propostos pelo teste MGN na previsão da produção de energia elétrica na CCPP.	41
Tabela 11 – Desempenho na previsão de temperatura no Vale da Morte.	42
Tabela 12 – Avaliação dos modelos propostos pelo teste MGN na previsão de temperatura no Vale da Morte.	43
Tabela 13 – Desempenho na previsão dos preços dos títulos do Global 40.	44
Tabela 14 – Avaliação dos modelos propostos pelo teste MGN na previsão dos preços dos títulos do Global 40.	46
Tabela 15 – Desempenho dos modelos na previsão do aluguel de bicicletas.	47
Tabela 16 – Avaliação dos modelos propostos pelo teste MGN na previsão de aluguel de bicicletas.	49

Lista de Algoritmos

Algoritmo 1 – Procedimento para criação e atualização de grupos e regras.	19
Algoritmo 2 – Procedimento para mesclagem de grupos e regras.	20
Algoritmo 3 – Procedimento para eliminação de grupos e regras.	21
Algoritmo 4 – Atualização dos parâmetros do consequente no eFLS.	22
Algoritmo 5 – Atualização dos parâmetros do consequente no eFMI.	24
Algoritmo 6 – Atualização dos parâmetros do consequente no eFCE.	25
Algoritmo 7 – Atualização dos parâmetros do consequente no eFTLS.	27
Algoritmo 8 – Pseudocódigo da abordagem proposta.	28

Lista de Símbolos

A	instância de tempo em que a regra foi ativada
a	índice de alerta
age	idade de um grupo/regra
B	função de pertinência Gaussiana multivariável
c	número de grupos/microgrupos e regras
D	distância da amostra atual ao centro do grupo
E	vetor de inovação
\bar{e}	resíduo do modelo
f	função não linear adequada para base de dados
g	ponto pelo qual o modelo de regressão passa
h	parâmetro do consequente
$\bar{\mathbf{h}}$	<i>features</i>
I	matriz identidade
i	índice da regra <i>fuzzy</i> , função de pertinência e grupo
i^*	índice do grupo com maior grau de compatibilidade
\bar{i}	índice do grupo com maior tempo de inatividade
j	índice das variáveis de entrada
L	matriz de ganho
k	índice da soma
\bar{k}	horizonte da previsão
l	variação do erro
m	número de variáveis de entrada
MGN	teste <i>Morgan-Granger-Newbold</i>

N	número de amostras
n	número de amostras atribuídas a um grupo
\bar{n}	defasagens para modelar a série
N_{max}	limiar para definição da medida de distância
NDEI	índice de erro não dimensional
P	matriz Hessiana ponderada inversa
p	probabilidade de sucesso
p_0	valor para iniciar a matriz Hessiana
$Q(\tau)$	matriz cujo os elementos da diagonal é 1 sobre cada elemento do vetor da função de pertinência normalizada das últimas v_r amostras até a amostra atual
R	regra <i>fuzzy</i>
r	valor de ocorrência
\hat{r}	vetor regressor da saída e das entradas da amostra atual
\bar{r}	vetor com a saída e as entradas da amostra atual
RMSE	raiz do erro quadrático médio
S	valor de ajuste
t	instância de tempo atual (passo)
u	menor autovetor da matriz de dispersão
V	variável aleatória
v_r	comprimento de inovação
x	vetor das variáveis de entrada de dimensão m
Y	vetor contendo as últimas saídas de v_r
y	saída desejada
\hat{y}	saída do modelo
y_i	consequente da i -ésima regra
z	número de limite de violações

μ	vetor com centro dos grupos
ν	vetor contendo média ponderada da saída e das entradas
Σ	matriz de dispersão (distância de Mahalanobis) ou matriz de identidade (distância Euclidiana)
τ	função de pertinência normalizada
γ	medida de compatibilidade
Γ_{max}	limite da medida de compatibilidade
α	nível de significância/intervalo de confiança
Γ_a	limite do mecanismo de alerta
φ	vetor com o primeiro elemento igual a 1 e os outros são os elementos do vetor de entrada
θ	vetor contendo as v_r amostras mais recentes
$\hat{\rho}_{\bar{e}}$	correlação de amostra
χ^2	distribuição Qui-quadrado
λ	taxa de aprendizagem
ω	tamanho da janela
ρ	parâmetro auxiliar para união dos grupos
σ	desvio padrão

Sumário

1 – Introdução	1
1.1 Motivação e Relevância	1
1.2 Caracterização do Problema	2
1.3 Objetivos	3
1.3.1 Objetivo Geral	3
1.3.2 Objetivos Específicos	3
1.4 Publicações	4
1.5 Contribuições da Dissertação	4
1.6 Organização do Trabalho	4
2 – Fundamentação Teórica	6
2.1 Fundamentos de Sistemas <i>Fuzzy</i>	6
2.2 Sistemas Evolutivos	7
2.3 Revisão da Literatura	8
3 – Abordagem Proposta	13
3.1 Introdução	13
3.2 Regras <i>Fuzzy</i>	15
3.3 Inicialização do Modelo e Cálculo da Saída	16
3.4 Criação e Atualização de Grupos e Regras	16
3.5 Mesclagem de Grupos e Regras	19
3.6 Eliminação de Grupos e Regras	20
3.7 Atualização dos Parâmetros do Consequente	21
3.7.1 eFLS - <i>evolving Fuzzy with Multivariable Gaussian Participatory Learning and Weighted Recursive Least Squares</i>	22
3.7.2 eFMI - <i>evolving Fuzzy with Multivariable Gaussian Participatory Learning and Multi-Innovations Recursive Weighted Least Squares</i>	22
3.7.3 eFCE - <i>evolving Fuzzy with Multivariable Gaussian Participatory Learning and Recursive Maximum Correntropy</i>	24
3.7.4 eFTLS - <i>evolving Fuzzy with Multivariable Gaussian Participatory Learning and Recursive Weighted Total Least Squares</i>	25
3.8 Parâmetros e Algoritmo da Abordagem Proposta	27
4 – Experimentos e Resultados	29
4.1 Metodologia	29
4.2 Identificação da Função da Hipérbole	32

4.2.1	Identificação da Função de Hipérbole - 7-D	32
4.2.2	Identificação da Função de Hipérbole - 10-D	34
4.3	Previsão da Concentração de CO_2	36
4.4	Previsão da Produção de Energia Elétrica em Usina de Ciclo Combinado - CCPP	39
4.5	Previsão de Temperatura no Vale da Morte	42
4.6	Previsão dos Preços dos Títulos da Global 40	43
4.7	Previsão de Aluguel de Bicicletas	45
5	– Conclusão	50
5.1	Trabalhos Futuros	50
	Referências	52

1 Introdução

Neste capítulo, apresenta-se uma descrição inicial do problema abordado por este trabalho juntamente com os desafios e as motivações que norteiam esta pesquisa. A Seção 1.1 expõe a motivação e relevância do trabalho e a Seção 1.2 descreve a caracterização do problema. A Seção 1.3 detalha os objetivos a serem alcançados com esta pesquisa. A Seção 1.4 destaca as publicações realizadas com base neste trabalho e a Seção 1.5 as principais contribuições. Por fim, conclui-se o capítulo com a Seção 1.6, na qual é descrito sucintamente o conteúdo de cada um dos capítulos.

1.1 Motivação e Relevância

Os sistemas evolutivos inicialmente tratados em [Kasabov e Song \(1999\)](#), [Angelov e Buswell \(2001\)](#), [Kasabov \(2001\)](#), [Kasabov e Song \(2002\)](#), [Angelov e Filev \(2004\)](#) são sistemas computacionais inteligentes capazes de resolver problemas em ambientes dinâmicos e não estacionários em vários domínios. Esses sistemas têm a capacidade de adaptar sua estrutura e parâmetros simultaneamente à medida que novas amostras são disponibilizadas a partir de um fluxo contínuo de dados ([ANGELOV, 2002](#)). Um sistema evolutivo atualiza sua estrutura e seus parâmetros conforme a demanda, dependendo das características do processo e das condições operacionais ([LUGHOFER, 2011](#)). Em resumo, as principais características dos sistemas evolutivos são sua estrutura flexível e seu aprendizado incremental e contínuo ([MACIEL; GOMIDE; BALLINI, 2012](#); [MACIEL et al., 2012](#)). Neste processo de aprendizagem, as amostras de um fluxo de dados são processadas apenas uma vez e então descartadas, diminuindo assim o consumo de memória ([LUGHOFER, 2013](#); [FERDAUS et al., 2019](#)). Alternativamente, pode-se dizer que o mecanismo de aprendizagem usa apenas a amostra de dados atual ([TUNG; QUEK; GUAN, 2013](#)).

Os componentes estruturais dos sistemas evolutivos podem, geralmente, ser baseados em redes neurais, regras *fuzzy* ou redes neuro-*fuzzy* ([ŠKRJANC et al., 2019b](#)). Os sistemas *fuzzy* evolutivos têm vantagens sobre redes neurais evolutivas porque são mais transparentes e explicáveis em termos linguísticos ([ANGELOV; FILEV; KASABOV, 2010b](#); [ŠKRJANC et al., 2019b](#)). A falta de interpretabilidade é uma grande desvantagem em diversas aplicações, especialmente quando a justificativa para a decisão do modelo é um requisito de confiança, como na medicina e finanças ([FERNANDEZ et al., 2019](#); [ARRIETA et al., 2020](#)).

Nos últimos anos, novas abordagens foram propostas gerando grandes avanços no âmbito dos sistemas *fuzzy* evolutivos e soluções práticas e persuasivas foram alcançadas

(ŠKRJANC et al., 2019b). No entanto, ainda existe uma demanda crescente para o desenvolvimento de sistemas *fuzzy* evolutivos, e as principais perspectivas estão relacionadas à adição, exclusão, divisão e união de grupos, neurônios, grânulos, folhas ou nuvens, a fim de garantir maior flexibilidade, caso ocorra mudança na dinâmica dos dados (LEITE; ŠKRJANC; GOMIDE, 2020; ŠKRJANC, 2020). Além disso, um desafio contínuo é a busca por algoritmos com boa precisão, alto grau de adaptabilidade e autonomia, e que sejam computacionalmente eficientes e interpretáveis (ŠKRJANC et al., 2019b).

Os sistemas *fuzzy* evolutivos têm sido utilizados com sucesso em diversas áreas. Exemplos de aplicações recentes em sistemas evolutivos incluem casas inteligentes e redes de sensores (ŠKRJANC et al., 2019b; ALIPERTI et al., 2019), sistemas de transporte (ANGELOV; FILEV; KASABOV, 2008; BLAŽIČ; ŠKRJANC, 2020), previsão e predição (LEITE; GOMIDE; ŠKRJANC, 2019; RONG; YANG; WONG, 2020), classificação (XYDEAS et al., 2006; ŠKRJANC et al., 2019a), identificação, detecção e diagnóstico de falhas (LUGHOFER; GUARDIOLA, 2008; LEITE; ŠKRJANC; GOMIDE, 2020), mineração de padrões sequenciais e reconhecimento de padrões (IGLESIAS et al., 2009; GARCIA et al., 2019), auto-localização e reconhecimento de pontos de referência (ZHOU; ANGELOV, 2007), para citar somente alguns. Sendo assim, o foco deste trabalho é no desenvolvimento de sistemas *fuzzy* evolutivos para tarefas regressão, mais especificamente, previsão de séries temporais e identificação de sistemas.

1.2 Caracterização do Problema

A previsão de séries temporais e a identificação de sistemas não lineares são tarefas desafiadoras (SUBHRAJIT; PRATAMA; SUNDARAM, 2019) e podem ser descritas como um problema não linear dinâmico e serem representadas por (LÜTKEPOHL, 2005):

$$\hat{y}^{t+\bar{k}} = f(\bar{\mathbf{h}}^t, \dots, \bar{\mathbf{h}}^{t-\bar{n}_h}, y^{t-1}, \dots, y^{t-\bar{n}_y}), \quad (1)$$

em que a predição no tempo $t + \bar{k}$ é realizada usando \bar{n}_y e \bar{n}_h defasagens para modelar a série, sendo t o passo atual, \bar{k} o horizonte da previsão, $\bar{\mathbf{h}}^t = [\bar{h}_1^t, \dots, \bar{h}_{\bar{n}_h}^t]$ as *features*¹, y a sequência de dados e f uma função não linear adequada para as observações anteriores. Observe que, em um contexto *online*, o fluxo de dados pode ser expresso como $S = y^1, y^2, \dots, y^t, \dots, y^\infty$ (FERDAUS et al., 2019).

Os desafios associados a modelagem de sistemas *online*, com dados variantes no tempo, têm as seguintes características:

- o fluxo de dados tem tamanho ilimitado, e as amostras chegam *online*, ocasionalmente, em tempo real (ŠKRJANC et al., 2019b);

¹ *Features* são dados existentes que se relacionam com a série temporal a ser aprendida)

- as amostras são processadas apenas uma vez e depois descartadas, ou seja, precisam ser aprendidas com uma única passagem nos dados (FERDAUS et al., 2019);
- o sistema deve lidar rapidamente com situações de variações nas condições ambientais do sistema, tais como *drift/shift* ou ambientes não estacionários (ŠKRJANC et al., 2019b; SUBHRAJIT; PRATAMA; SUNDARAM, 2019);
- grande quantidade de dados a serem processados (ŠKRJANC, 2020).

Nesse contexto, este trabalho sugere uma nova abordagem para sistemas *fuzzy* evolutivos para tarefas de regressão, como previsão e identificação de sistemas. A abordagem proposta é baseada em uma versão modificada do algoritmo de aprendizado participativo com funções de pertinência Gaussiana multivariável introduzido por Lemos, Caminhas e Gomide (2011). Ao contrário dos modelos anteriores, na abordagem proposta, a estimação dos grupos usa tanto a distância Euclidiana quanto a distância de Mahalanobis. Isso significa que a amostra é atribuída aos microgrupos (distância Euclidiana) ou aos grupos (distância de Mahalanobis). O uso das duas distâncias visa evitar o problema da singularidade ao calcular a inversa da matriz de dispersão do grupo quando este possui um pequeno número de amostras (ŠKRJANC, 2020). Um método de exclusão de grupos baseado no conceito de idade (LUGHOFFER; ANGELOV, 2011) e população (ANGELOV; FILEV, 2005) do grupo foi introduzido. A exclusão de regras está associada à capacidade do modelo esquecer conhecimentos antigos sempre que estes se tornam inúteis. Além disso, um novo método para a mesclagem de grupos é implementado baseado na sobreposição notável de um determinado par de grupos, visando um melhor desempenho para mesclar os grupos. Por fim, com base na abordagem proposta são introduzidos quatro modelos evolutivos que se diferenciam pelo algoritmo recursivo utilizado para atualização dos parâmetros do consequente: (i) Mínimos Quadrados Ponderados; (ii) Mínimos Quadrados Ponderados com Multi-inovações; (iii) Máxima Correntropia; (iv) Mínimos Quadrados Totais Ponderados.

1.3 Objetivos

1.3.1 Objetivo Geral

O objetivo geral deste trabalho é propor e implementar novos algoritmos *fuzzy* evolutivos para tarefas de previsão e identificação de sistemas.

1.3.2 Objetivos Específicos

Os objetivos específicos deste trabalho são:

- Revisar as metodologias para inclusão, mesclagem, remoção e atualização de grupos e regras em sistemas *fuzzy* evolutivos;

- Propor e implementar novos algoritmos *fuzzy* evolutivos que apresentem boa acurácia, alto grau de adaptabilidade e autonomia, e que sejam interpretáveis;
- Avaliar e comparar o desempenho dos modelos evolutivos propostos com o estado da arte em problemas de previsão e identificação de sistemas dinâmicos.

1.4 Publicações

O presente trabalho gerou as seguintes publicações:

- Rodrigues, F.; Silva, A.; Lemos, A.. **Evolving Fuzzy System with Multivariable Gaussian Participatory Learning and Recursive Maximum Correntropy** - eFCE. *FUZZY IEEE - International Conference on Fuzzy Systems*. 2021 - Artigo publicado.
- Rodrigues, F.; Silva, A.; Lemos, A.. **Um Preditor Fuzzy Evolutivo com Aprendizagem Participativa Gaussiana Multivariável e Máxima Correntropia** - eFCE. VI Congresso Brasileiro de Sistemas *Fuzzy* (VI CBSF). 2021 - Artigo publicado.
- Rodrigues, F.; Silva, A.; Lemos, A.. **An Evolving Fuzzy Predictor with Multivariable Gaussian Participatory Learning and Multi-Innovations Recursive Weighted Least Squares** - eFMI. *Evolving Systems*. 2022 - Artigo publicado.

1.5 Contribuições da Dissertação

As novidades e as principais contribuições deste trabalho são descritas a seguir:

- Descrição de uma nova abordagem *fuzzy* evolutiva cuja estrutura evolui com base em um algoritmo de agrupamento com aprendizado participativo, funções de pertinência Gaussiana multivariável e duas distâncias na estimativa dos grupos. A exclusão dos grupos é realizada utilizando o tempo de inatividade e o número de amostras associadas aos grupos. A mesclagem de grupos é realizada com base na sobreposição de dois grupos.
- Proposição de quatro algoritmos *fuzzy* evolutivos que se diferenciam pelo método de atualização dos parâmetros do consequente das regras *fuzzy*. Os algoritmos foram propostos para resolução de problemas gerais de regressão em ambientes não estacionários compreendendo condições ambientais voláteis. Destaca-se que os algoritmos também podem ser utilizados em ambientes com dados estáticos.

1.6 Organização do Trabalho

O texto é dividido em 5 capítulos:

- Este capítulo apresenta as motivações e os objetivos que tangenciam esta pesquisa.

- O Capítulo 2 trás uma breve revisão da literatura para lembrar as principais ideias, avanços, abordagens e lacunas no desenvolvimento de sistemas evolutivos;
- O Capítulo 3 detalha a abordagem proposta;
- O Capítulo 4 discorre sobre os experimentos computacionais realizados para avaliar e comparar o desempenho dos modelos propostos em relação aos algoritmos alternativos do estado da arte, considerando instâncias de previsão de séries temporais e problemas de identificação de sistemas não lineares;
- Por fim, o Capítulo 5 trata das considerações finais resumindo as contribuições do trabalho e listando sugestões para estudos futuros.

2 Fundamentação Teórica

Este capítulo discorre com os fundamentos de sistemas evolutivos e trás uma revisão da literatura que ilustra as principais abordagens. Na Seção 2.1 são discutidos os conceitos básicos dos sistemas *fuzzy*. A Seção 2.2 apresenta uma breve introdução aos sistemas evolutivos descrevendo suas principais características e as tendências no desenvolvimento desses sistemas. A Seção 2.3 trata de uma revisão da literatura e destaca algumas das relevantes abordagens em sistemas evolutivos.

2.1 Fundamentos de Sistemas *Fuzzy*

Os sistemas *fuzzy* são modelos matemáticos que se baseiam no conceito de lógica *fuzzy* introduzida por Zadeh (1965). Os sistemas *fuzzy* e suas variações têm sido amplamente utilizados em atividades nas quais são necessárias capacidade de raciocínio semelhantes à dos seres humanos. O conhecimento nos sistemas *fuzzy* é armazenado em um conjunto de regras *fuzzy*. As regras são suposições lógicas do tipo:

$$\text{SE } < \textit{antecedente} > \text{ ENTÃO } < \textit{consequente} >, \quad (2)$$

sendo que os antecedentes descrevem um conjunto de premissas, que podem ser satisfeitas totalmente ou parcialmente e são representados funções de pertinência (conjuntos *fuzzy*). O consequente é constituído de um conjunto de ações ou diagnósticos que são gerados quando as condições do antecedente são cumpridas. O consequente pode ser linguístico (MAMDANI; ASSILIAN, 1975) ou funcional (TAKAGI; SUGENO, 1985). Nos modelos linguísticos o consequente é representado por uma função de pertinência e nos funcionais por uma função de ordem zero ou de primeira ordem.

A Figura ilustra 1 ilustra a estrutura básica de um sistema *fuzzy* e suas principais etapas. Como pode ser visto, inicialmente as variáveis de entrada passam por um processo de fuzzificação. A fuzzificação determina o grau de pertinência de cada entrada no antecedente da regra, ou seja, ler e converte os dados, que normalmente são precisos (valores crisp), numa variável linguística usando as funções de pertinência armazenadas na base de conhecimento *fuzzy*. Após a fuzzificação, os dados são encaminhados para o mecanismo de inferência. Os valores fuzzificados são aplicados a um conjunto de regras de inferência que mapeiam as entradas fuzzificadas às regras, do tipo **SE-ENTÃO**, para gerar as saídas também fuzzificadas (FULCHER, 2008). Uma vez obtido o conjunto *fuzzy* de saída através do processo de inferência, é necessário fazer a defuzzificação do resultado, ou seja, é efetuada a conversão do conjunto *fuzzy* de saída em um valor crisp.

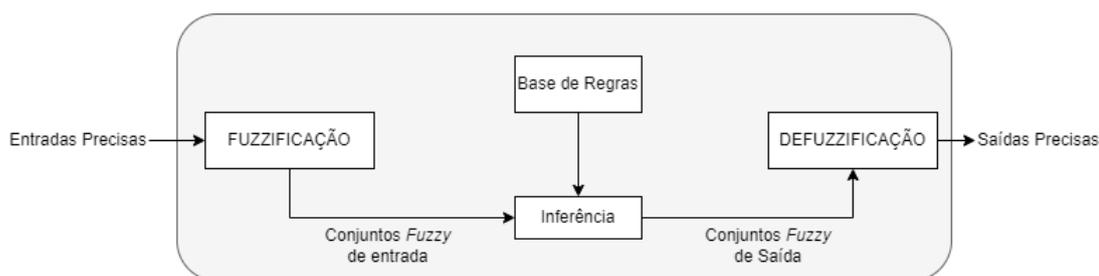


Figura 1 – Estrutura básica de um Sistema Fuzzy.

2.2 Sistemas Evolutivos

Os sistemas evolutivos são sistemas inteligentes, em geral, baseados em redes neurais, regras *fuzzy* ou redes neuro-*fuzzy* (ŠKRJANC et al., 2019b). A aplicação desses tipo de sistema tem crescido atualmente, devido há uma necessidade crescente de processar fluxos de dados para extrair conhecimento para modelar sistemas complexos. Geralmente, esses fluxos de dados são dinâmicos e não estacionários e requerem modelos que processem as amostras de dados de forma incremental, passo a passo e, eventualmente, em tempo real (ŠKRJANC, 2020). Os sistemas evolutivos têm a capacidade de desenvolver gradualmente sua estrutura simultaneamente com a atualização dos parâmetros a partir de um fluxo de dados.

É importante ressaltar, que o termo evolutivo não pode ser confundido com evolucionário, enquanto os sistemas evolutivos são capazes de ajustar sua estrutura e seus parâmetros em tempo real. Os sistemas evolucionários (algoritmo genético (GOLDBERG, 1989), programação genética (KOZA, 1992), dentre outros) baseiam-se no processo de evolução que ocorre em população de seres vivos e seus mecanismos de adaptação são inspirados no conceito de seleção, cruzamento e mutação de cromossomos.

As principais tendências em sistemas evolutivos estão no desenvolvimento de técnicas e metodologias para obtenção de sistemas com alto grau de adaptabilidade e autonomia, construídos a partir de um fluxo de dados coletado *online* (ANGELOV; ZHOU, 2008; ANGELOV; FILEV; KASABOV, 2010a; PRATAMA; ANAVATTI; LUGHOFER, 2013; ŠKRJANC et al., 2019b). A adaptabilidade e a autonomia são obtidas modificando os parâmetros e a estrutura do sistema à medida que ocorrem mudanças nos dados de entrada. Os modelos evolutivos se diferem pelo tipo de estrutura e pelo mecanismo adotado para realizar a evolução da estrutura, o que confere ao modelo um maior ou menor grau de flexibilidade, robustez e confiabilidade.

A base de regras de um sistema evolutivo pode ser construída incrementalmente por um algoritmo de agrupamento usando a organização espacial de variáveis de entrada e/ou saída (LE MOS; CAMINHAS; GOMIDE, 2010). O algoritmo de agrupamento processa os dados de entrada recursivamente e pode adicionar, excluir ou adaptar os grupos existentes

de forma dinâmica. Outras abordagens usam o erro de modelagem calculado recursivamente para construir a base de regras (WANG; ZENG; KEANE, 2010; SILVA et al., 2014). Alternativamente, pode-se dizer que nessas abordagens, a estrutura evolui com base na qualidade do modelo resultante.

2.3 Revisão da Literatura

A maioria dos sistemas evolutivos são baseados em regras de Takagi-Sugeno de ordem zero ou de primeira ordem (GARCIA et al., 2019) e os parâmetros do consequente são atualizados por mínimos quadrados recursivos, gradiente descendente ou versões modificadas desses algoritmos (ŠKRJANC et al., 2019b). O eTS (*evolving Takagi-Sugeno*), introduzido em Angelov (2002), é um marco importante no desenvolvimento de sistemas evolutivos. O eTS usa uma versão incremental do algoritmo de agrupamento subtrativo e regras funcionais de Takagi-Sugeno de primeira ordem. A base de regras evolui continuamente, adicionando novas regras e modificando regras e parâmetros existentes. A criação de novas regras ou a modificação das regras atuais é realizada pela avaliação recursiva do potencial das novas amostras. Os parâmetros do consequente das regras são atualizados por um algoritmo recursivo de mínimos quadrados.

Vários modelos evolutivos foram construídos sob a estrutura eTS. O Simpl_eTS (ANGELOV; FILEV, 2005) é uma versão rápida do eTS que visa simplificar os cálculos de potencial. O algoritmo combina o conceito de dispersão com uma medida de densidade de dados, que pode ser calculada com maior facilidade, resultando em um algoritmo semelhante e mais eficaz do ponto de vista computacional. O Simpl_eTS usa funções de pertinência Cauchy em vez das funções de pertinência Gaussianas usadas no eTS. Além disso, o Simpl_eTS apresenta um método para remover grupos/regras usando o conceito de população. Uma versão estendida do eTS foi proposta por Angelov e Zhou (2006) e denominada xTS (*eXtended Takagi-Sugeno*). Este modelo combina modelos Takagi-Sugeno de ordem zero e primeira ordem com um algoritmo de agrupamento incremental. No xTS, o raio de influência dos grupos é estimado recursivamente. Os grupos e regras são removidos pela proporção entre o conjunto de suporte, o número total de amostras e a idade dos grupos (ANGELOV; ZHOU, 2006). Um modelo flexível e autônomo chamado eTS+ foi proposto em Angelov, Filev e Kasabov (2010b) como uma versão aprimorada do eTS. No eTS+, os antecedentes e a estrutura da base de regras são atualizados usando os critérios de utilidade, idade, área de influência e densidade local. Este modelo não requer limites ou parâmetros definidos pelo usuário (ANGELOV, 2011).

O FLEXFIS (*Flexible Fuzzy Inference Systems*), detalhado em Lughofer (2008b), usa um algoritmo de agrupamento recursivo derivado de uma modificação da quantização vetorial (LUGHOFER, 2008a). Um novo grupo é criado apenas quando o número de amostras

associadas a este grupo excede um limite predefinido. Os parâmetros do consequente das regras são estimados por um algoritmo recursivo de mínimos quadrados ponderados. Em [Lughofer \(2012\)](#) uma versão melhorada do FLEXFIS foi introduzida, chamada FLEXFIS++. A principal melhoria apresentada no FLEXFIS++ é a atenuação da complexidade pela mesclagem e exclusão de regras devido ao mecanismo que reduz a dimensionalidade de maneira suave e gradual.

O aprendizado participativo foi proposto em [Yager \(1990\)](#) e modificado para uso em sistemas evolutivos no trabalho de [Silva, Gomide e Yager \(2005\)](#). Em [Lemos, Caminhas e Gomide \(2011\)](#) foi introduzido um novo modelo fuzzy evolutivo com aprendizado participativo chamado eMG (*evolving Multivariable Gaussian Fuzzy System*). O eMG usa funções de pertinência Gaussiana multivariável na representação dos grupos. O antecedente das regras é obtido pelo algoritmo de agrupamento com aprendizado participativo. Os parâmetros do consequente são atualizados por um algoritmo recursivo de mínimos quadrados ponderados. O algoritmo do eMG permite a inclusão de uma nova regra, a modificação dos parâmetros de uma regra existente e a mesclagem de regras redundantes. O eGAUSS+ demonstrado por [Škrjanc \(2020\)](#), assim como o eMG ([LEMOS; CAMINHAS; GOMIDE, 2011](#)), é baseado no conceito de aprendizagem participativa e usa funções de pertinência Gaussiana multivariável. Este algoritmo cria e adiciona novos grupos com base no conceito de tipicidade, calculado pela densidade Gaussiana não normalizada da amostra atual. Um novo grupo é criado sempre que o grau de tipicidade da amostra excede o valor limite para todos os grupos. Os grupos são mesclados utilizando o conceito de volume que, segundo os autores, diminui o custo computacional.

Alguns algoritmos *fuzzy* evolutivos têm uma estrutura granular como eOGS (*evolving Optimal Granular Systems*) ([LEITE; ŠKRJANC, 2019](#)), eFGP (*evolving Fuzzy Granular Predictor*) ([GARCIA et al., 2019](#)) e eIX (*evolving Internal-eXternal Fuzzy*) ([AGUIAR; LEITE, 2020](#)). O eOGS usa funções de pertinência Gaussiana para definir a granularidade do modelo e operar com formas hiper-retangulares nos grânulos. Os grânulos e regras são criados sempre que a amostra não está dentro do limite inferior e superior de nenhum grânulo. Para mesclar dois grânulos, obtém-se o menor valor da norma da diferença entre os pontos médios de todos os pares de grânulos. Se esse valor for menor que um limite (constante ou variante no tempo), os grânulos serão mesclados. A remoção do grânulo ocorre se o número de amostras que ativaram o grânulo for igual a zero após um determinado número de novas amostras.

O eFGP ([GARCIA et al., 2019](#)), por outro lado, possui funções de pertinência Trapezoidal e os parâmetros do consequente das regras são atualizados por um algoritmo recursivo de mínimos quadrados. Os grânulos e as regras são criados sempre que a nova amostra contém pelo menos um elemento fora da região de expansão de qualquer grânulo. Os grânulos são excluídos se não forem ativados por uma instância de tempo predefinida.

A mesclagem usa o método do casco convexo, em que a união dos grânulos é o menor conjunto convexo que contém os grânulos mesclados.

O eIX (*evolving Internal-eXternal Fuzzy*) (AGUIAR; LEITE, 2020) é um classificador evolutivo com uma estrutura granular de borda dupla e um algoritmo de treinamento não supervisionado. Este algoritmo usa funções de pertinência Trapezoidal e os grânulos são delimitados por hipercaixas internas e externas. Um novo grânulo é criado sempre que uma amostra não pertence a nenhuma região interna ou externa de um grânulo. A mesclagem ocorre se um par de grânulos estiver suficientemente próximo, ou seja, sempre que a norma da diferença em seus centros for menor que um valor predefinido entre 0 e 1. O eIX utiliza dois métodos de mesclagem, média ponderada e casco convexo. O método de mesclagem de casco convexo produz um grânulo mais espesso que encapsula todas as informações dos grânulos anteriores.

O ALMMo (*Autonomous Learning Multi-Model Systems*) é um sistema *fuzzy* evolutivo introduzido em Angelov, Gu e Príncipe (2018), que possui uma estrutura de agrupamento composta por uma estrutura sem forma ou limites denominada nuvem de dados. Novas regras são criadas com base na densidade local da nova amostra. Uma nova regra é adicionada sempre que a densidade local da nova amostra em cada nuvem de dados é maior do que as densidades dos dados nos protótipos das nuvens de dados existentes ou menor que a densidade mínima em qualquer um dos protótipos. Caso contrário, a nova amostra é atribuída à nuvem de dados mais próxima. A exclusão de regras é realizada com base nos conceitos de utilidade e importância. Inicialmente, as regras obsoletas são removidas usando um conceito de utilidade. Uma regra é removida se sua utilidade for menor que um parâmetro definido para monitorar a qualidade das regras *fuzzy*. Então, as regras menos importantes são excluídas sempre que a soma acumulada normalizada dos valores dos parâmetros for menor que o valor médio da soma acumulada normalizada de todas as regras.

Outro modelo baseado em nuvem de dados é o AutoCloud proposto por Bezerra et al. (2020). Sua estrutura evolui com base no conceito de tipicidade e excentricidade. No AutoCloud, uma nova nuvem de dados é criada se o novo centro de amostra for menor ou igual a um limite. Além disso, duas nuvens se fundem sempre que o número de amostras pertencentes a ambas as nuvens simultaneamente é maior do que o número de amostras que pertencem a elas separadamente.

O eNFN (*evolving Neo-Fuzzy-Neuron*) (SILVA et al., 2014) é uma rede *neuro-fuzzy* evolutiva que usa funções de pertinência Triangular e complementar. As regras são do tipo Takagi-Sugeno de ordem zero e podem ser adicionadas, excluídas ou ter seus parâmetros ajustados, dependendo do erro de modelagem calculado recursivamente. Para criar regras *fuzzy*, o valor médio do erro local das regras correspondentes às funções de pertinência mais ativas é calculado e comparado com o valor médio do erro de modelagem global.

Se o valor médio do erro local for maior do que o erro médio global, uma nova regra será adicionada. A exclusão de regras é baseada na idade da regra. Os parâmetros do consequente são atualizados por um algoritmo baseado no gradiente descendente com taxa de aprendizado ótima.

O eCauchy introduzido em Škrjanc et al. (2018) e em Škrjanc et al. (2019a) usa funções de pertinência Cauchy que permitem a criação de grupos com forma hiper elipsoidal. A criação de grupos é baseada no conceito de tipicidade. A tipicidade é calculada usando a densidade local do tipo Cauchy (ANGELOV; YAGER, 2011; BLAŽIČ; DOVŽAN; ŠKRJANC, 2014). Assim, um novo grupo é criado se o maior grau de tipicidade da amostra atual for menor que seu respectivo limite para todos os grupos. Por outro lado, uma regra será removida sempre que o número de amostras do grupo estiver abaixo de um limite predefinido. Quatro opções são apresentadas para mesclar os grupos: pela distância de Mahalanobis entre dois grupos; com base em graus de associação das amostras; pela correlação entre os níveis de disparo dos grupos; ou, no caso de modelos de regressão, a partir de informações dos modelos locais. Além disso, o eCauchy possui um mecanismo de divisão de grupo, no qual divide os grupos contendo um determinado número de amostras classificadas incorretamente.

A Tabela 1 resume as principais características dos modelos evolutivos descritos neste capítulo. A tabela ilustra as funções de pertinência e os métodos para adicionar, mesclar e excluir regras e/ou grupos usados pelos modelos. O algoritmo de aprendizado dos parâmetros do consequente também é apresentado.

Tabela 1 – Características dos modelos evolutivos.

Modelos	Função de Pertinência	Adição de regras e/ou grupos	Mesclagem de regras e/ou grupos	Remoção de regras e/ou grupos	Atualização Consequente	do
ALMMo (ANGELOV; GU; PRINCIPE, 2018)	Data clouds	Densidade da nova amostra	-	Medida de utilidade	FWRLS	
AutoCloud (BEZERRA et al., 2020)	Data clouds	Medida de excentricidade	O número de amostras sobrepostas	-	RLS	
eCauchy (ŠKRJANC et al., 2019a)	Cauchy	Distância entre uma amostra e os centros dos grupos	Distância Mahalanobis entre os grupos; correlação entre taxas de franja; a correlação entre os níveis de disparo dos grupos ou informações dos próprios modelos locais (modelos de regressão)	-	RLS	
eFGP (GARCIA et al., 2019)	Trapezoidal	Região de expansão de um grânulo	Grânulos vizinhos	Inatividade de grânulos e regras	RLS	
eGAUSS+ (ŠKRJANC, 2020)	Gaussiana	Tipicidade da amostra atual	Volume de grupos	-	RLS	
eIX (AGUIAR; LEITE, 2020)	Trapezoidal	Intervalo dos grânulos	Distância entre os centros das <i>beds</i>	-	Taxa de equilíbrio	
eMG (LEMOS; CAMINHAS; GOMIDE, 2011)	Gaussiana	Medida de compatibilidade da amostra atual	Grau de compatibilidade	-	RWLS	
eNFN (SILVA et al., 2014)	Triangular	Valor médio do erro local e erro médio global	-	Idade da regra	Gradiente descendente	
eOGS (LEITE; ŠKRJANC, 2019)	Gaussiana	Intervalo do grânulo	Grânulos vizinhos	Inatividade de grânulos e regras	RLS	
eTS (ANGELOV, 2002)	Gaussiana	Potencial da amostra atual	-	-	RLS	
eTS+ (ANGELOV; FILEV; KASABOV, 2010b)	Gaussiana	Potencial da amostra atual	-	Tamanho do conjunto de suporte do grupo	FWRLS	
FLEXFIS (LUGHOFER, 2008b)	Gaussiana	Distância entre uma amostra e os centros dos grupos	-	-	WRLS	
FLEXFIS+ (LUGHOFER, 2012)	Gaussiana	Distância entre uma amostra e os centros dos grupos	Índice de sobreposição	-	WRLS	
Simpl_eTS (ANGELOV; FILEV, 2005)	Cauchy	Dispersão da amostra atual	-	População de cada grupo	RLS	
xTS (ANGELOV; ZHOU, 2006)	Gaussiana	Potencial da amostra atual	-	Idade da regra	RLS	
eFMI (RODRIGUES; SILVA; LEMOS, 2022)	Gaussiana	Medida de compatibilidade da amostra atual	Sobreposição de grupos	Idade e população da regra/grupo	MI-WRLS	
eFCE (RODRIGUES; SILVA; LEMOS, 2021)	Gaussiana	Medida de compatibilidade da amostra atual	Sobreposição de grupos	Idade e população da regra/grupo	Máxima Correntropia	
eFLS (Proposto neste trabalho)	Gaussiana	Medida de compatibilidade da amostra atual	Sobreposição de grupos	Idade e população da regra/grupo	WRLS	
eFTLS (Proposto neste trabalho)	Gaussiana	Medida de compatibilidade da amostra atual	Sobreposição de grupos	Idade e população da regra/grupo	RWTLs	

3 Abordagem Proposta

A abordagem proposta neste trabalho foi construída empregando um algoritmo de agrupamento recursivo não supervisionado com aprendizagem participativa e funções de pertinência Gaussiana multivariável. São introduzidos quatro novos algoritmos, chamados, respectivamente, eFLS (*evolving Fuzzy with Multivariable Gaussian Participatory Learning and Weighted Recursive Least Squares*), eFMI (*evolving Fuzzy Multivariable Gaussian Participatory with Multi-Innovations Recursive Weighted Least Squares*), eFCE (*evolving Fuzzy with Multivariable Gaussian Participatory Learning and Recursive Maximum Correntropy*) e eFTLS (*evolving Fuzzy with Multivariable Gaussian Participatory Learning and Recursive Weighted Total Least Squares*). As próximas seções detalham a abordagem proposta. Mais especificamente, a Seção 3.1 introduz a abordagem proposta e sumariza suas principais etapas. A Seção 3.2 apresenta as regras e funções de pertinência. A Seção 3.3 ilustra como o primeiro grupo/regra é criado e destaca os cálculos para a obtenção da saída. Em seguida, a Seção 3.4 discorre sobre a metodologia para criar e atualizar grupos e regras. O método usado para a mesclagem dos grupos é descrito na Seção 3.5 e para eliminação na Seção 3.6. A Seção 3.7 ilustra os procedimentos a atualização dos parâmetros do consequente das regras. Finalmente, a Seção 3.8 descreve os parâmetros e apresenta o algoritmo da abordagem.

3.1 Introdução

A abordagem proposta pode, em uma única passagem, calcular a saída e atualizar sua estrutura e parâmetros simultaneamente. A estrutura evolui adicionando uma nova regra, mesclando ou excluindo uma regra existente. A Figura 2 mostra um fluxograma com as etapas da abordagem proposta e a seguir as principais etapas são sumarizadas:

- O modelo é inicializado com a primeira amostra e o primeiro grupo é criado centrado nesta amostra. A partir do centro do primeiro grupo é criado o antecedente da primeira regra e os parâmetros do consequente desta regra são definidos por um valor inicial pré-estabelecido. Esta etapa é realizada apenas uma vez.
- Em seguida, para cada amostra, verifique o número de amostras atribuídas a cada grupo. De acordo com o número de amostras dos grupos, a medida de compatibilidade usa a distância Euclidiana ou de Mahalanobis.
- Calcule o antecedente e o consequente das regras. Em seguida, obtenha a saída do modelo.
- Calcule a medida de compatibilidade e o índice de alerta. Encontre o grupo com a maior compatibilidade e decida se um novo grupo deve ser criado ou se o centro do

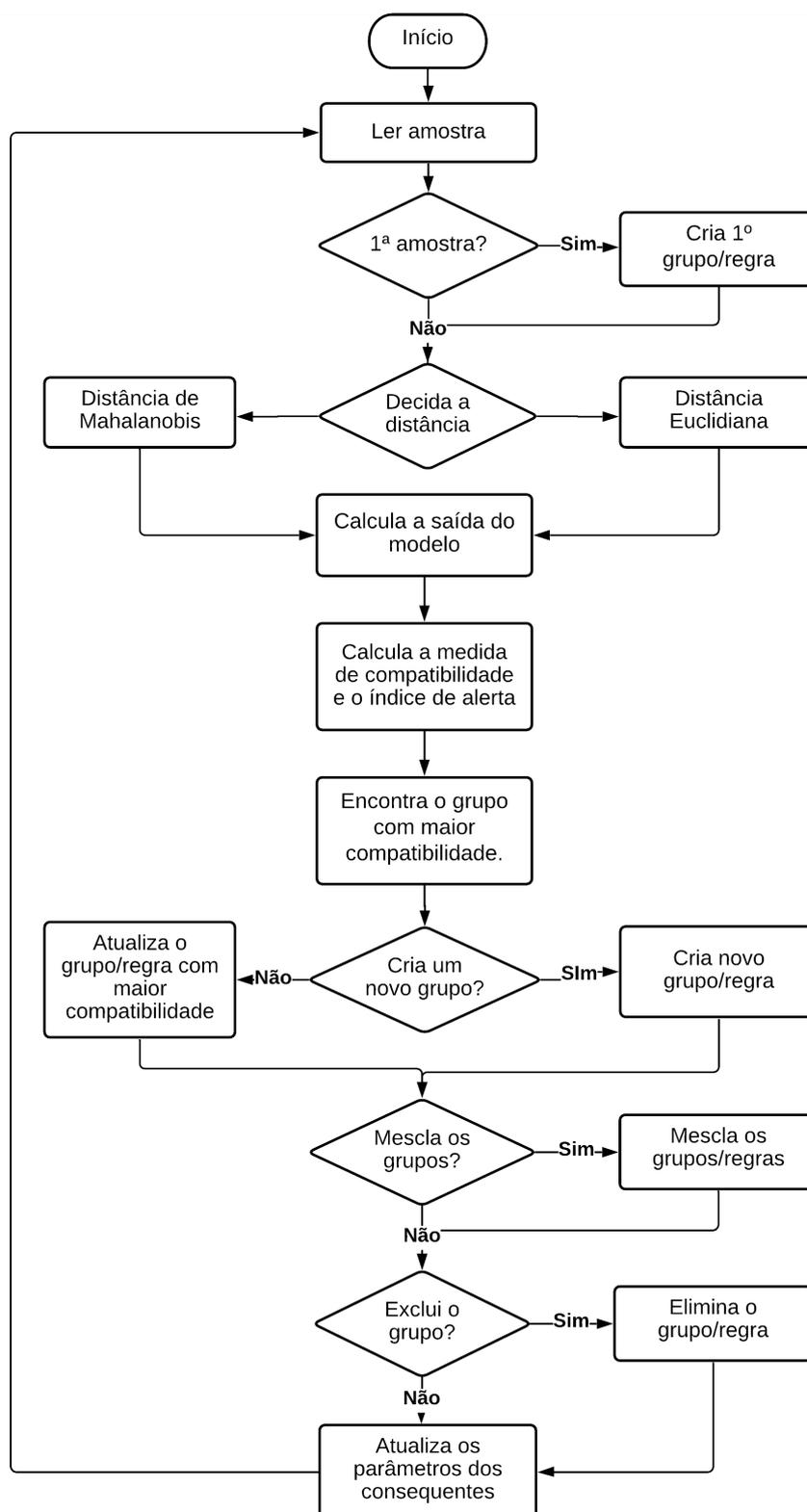


Figura 2 – Fluxograma da abordagem proposta.

grupo mais compatível deve ser atualizado.

- Verifique se há grupos redundantes e decida se eles devem ser mesclados.
- Encontre o grupo com maior tempo de inatividade e decida se ele deve ser removido.
- Atualize os parâmetros do consequente usando o algoritmo recursivo de mínimos quadrados ponderados, mínimos quadrados ponderados com multi-inovações, máxima correntropia ou mínimos quadrados totais ponderados.

3.2 Regras Fuzzy

A abordagem proposta usa regras *fuzzy* de Takagi-Sugeno de primeira ordem, cujos antecedentes são representados por funções de pertinência Gaussiana multivariável. As regras *fuzzy* são como se segue:

$$\begin{aligned}
 R_1 : \quad & \text{Se } x^t \text{ é } B_1 \text{ então } y_1^t = h_{10}^t + \sum_{j=1}^m h_{1j}^t x_j^t \\
 & \dots \\
 R_i : \quad & \text{Se } x^t \text{ é } B_i \text{ então } y_i^t = h_{i0}^t + \sum_{j=1}^m h_{ij}^t x_j^t \\
 & \dots \\
 R_c : \quad & \text{Se } x^t \text{ é } B_c \text{ então } y_c^t = h_{c0}^t + \sum_{j=1}^m h_{cj}^t x_j^t
 \end{aligned}$$

em que R_i é a i -ésima regra, i é o índice das regras *fuzzy* e grupos, x^t é a amostra de dados atual descrita como $[x_1^t \cdot \cdot \cdot x_j^t \cdot \cdot \cdot x_m^t]^T$, t é o passo atual, j é o índice das variáveis de entrada, m é o número de variáveis de entrada, B_i é uma função de pertinência Gaussiana multivariável com parâmetros extraídos do centro do grupo correspondente, y_i^t é o consequente da i -ésima regra (saída da regra), h_{i0}^t e h_{ij}^t são os parâmetros do consequente da i -ésima regra e c é o número de regras e grupos.

Uma função de pertinência Gaussiana multivariável pode ser descrita por:

$$f(x) = e^{-\frac{1}{2}(x^t - \mu_i^t)(\Sigma_i^t)^{-1}(x^t - \mu_i^t)^T}, \quad (3)$$

na qual μ é um vetor com o centro dos grupos (valor modal) definido como $[\mu_1^t \cdot \cdot \cdot \mu_i^t \cdot \cdot \cdot \mu_{c^t}^t]^T$, c^t é o número de grupos e microgrupos e Σ é uma matriz positiva definida simétrica $m \times m$. Na abordagem proposta, como nas versões anteriores, o número de grupos é igual ao número de regras *fuzzy*, ou seja, cada grupo representa uma regra. Os grupos são encontrados por um algoritmo de agrupamento que pode, a cada passo, criar um novo grupo, remover, mesclar ou atualizar os grupos existentes.

3.3 Inicialização do Modelo e Cálculo da Saída

Na abordagem proposta a primeira amostra é usada para criar o primeiro grupo, cujo centro é definido pelos valores da amostra. Em seguida, uma regra é criada com o valor modal do antecedente definido como o centro do grupo. Os parâmetros do consequente e a matriz de dispersão são inicializados como um valor pré-definido. A partir da segunda amostra, o algoritmo começa escolhendo a medida de distância.

A saída é obtida pela média ponderada das contribuições de cada regra, como segue:

$$\hat{y}^t = \sum_{i=1}^{c^t} \tau_i^t y_i^t, \quad (4)$$

na qual y_i^t é o consequente da i -ésima regra e τ_i^t são as funções de pertinência normalizadas calculadas no passo t por:

$$\tau_i^t = \frac{e^{D(x^t, \mu_i^t)}}{\sum_{i=1}^{c^t} e^{D(x^t, \mu_i^t)}}, \quad (5)$$

em que μ_i^t é o centro de i -ésimo grupo, \sum_i^t é uma matriz de dispersão usando a distância de Mahalanobis ou uma matriz de identidade aplicando a distância Euclidiana e onde $D(x^t, \mu_i^t)$ por ser descrito como:

$$D(x^t, \mu_i^t) = (x^t - \mu_i^t) \left(\sum_i^t \right)^{-1} (x^t - \mu_i^t)^T. \quad (6)$$

3.4 Criação e Atualização de Grupos e Regras

A estrutura dos grupos é atualizada à medida que cada nova amostra é apresentada. Este procedimento é realizado com base em uma medida de compatibilidade e um mecanismo de alerta. A medida de compatibilidade, $\gamma_i^t \in [0, 1]$, descreve a intensidade com que a nova amostra é compatível com cada um dos grupos. Em outras palavras, a medida de compatibilidade é uma função que calcula a compatibilidade entre a amostra atual x^t e o conhecimento atual do modelo representado pelos centros dos grupos existentes (YAGER, 1990; LEMOS; CAMINHAS; GOMIDE, 2011):

$$\gamma_i^t = f(x^t, \mu_i^t) = e^{-\frac{1}{2}D(x^t, \mu_i^t)}, \quad (7)$$

Seja n_i o número de amostras do i -ésimo grupo e N_{max} um parâmetro definido pelo usuário. Se n_i é menor que N_{max} , então os valores de $D(x^t, \mu_i^t)$ são calculados com base

na distância Euclidiana e \sum_i^t é uma matriz identidade $m \times m$. Caso contrário, os cálculos são feitos usando a distância de Mahalanobis, ou seja, \sum_i^t é a matriz de dispersão do grupo i no passo t (ŠKRJANC, 2020).

O limiar da medida de compatibilidade Γ_{max} para os cálculos da distância Euclidiana é obtido por:

$$\Gamma_{max} = e^{-(\min_i(\frac{\max(x_i) - \min(x_i)}{2\omega}))^2}, \quad (8)$$

em que ω é o tamanho da janela. Para a distância de Mahalanobis Γ_{max} é dado por:

$$\Gamma_{max} = e^{-\frac{1}{2}\chi_{m,\alpha}^2}, \quad (9)$$

na qual $\chi_{m,\alpha}^2$ é uma distribuição Qui-quadrado com m graus de liberdade e α é o intervalo de confiança unilateral. Depois de calcular a medida de compatibilidade entre a nova amostra e todos os grupos existentes, encontra-se o índice do grupo com o maior grau de compatibilidade i^* .

O alerta é um mecanismo utilizado para sinalizar quando a estrutura dos grupos não representa adequadamente o conhecimento atual do sistema e este precisa ser revisto (LEMOS; CAMINHAS; GOMIDE, 2011). Para cada nova amostra x^t o índice de alerta $a_i^t \in [0,1]$ é calculado para todos os grupos e estimado através da probabilidade cumulativa de V^t , isto é:

$$a_i^t = p(V^t < z), \quad (10)$$

em que $p(V^t = z)$ é uma distribuição binomial, na qual V^t , sendo definida como a somatória de uma sequência de variáveis aleatórias independentes e igualmente distribuídas geradas por uma distribuição de Bernoulli, com a mesma probabilidade de sucesso (α) (LEMOS; CAMINHAS; GOMIDE, 2011). A função de probabilidade é dada por:

$$p(V^t = z) = \frac{\omega!}{z!(\omega - z)!} \alpha^z (1 - \alpha)^{\omega - z}, \quad (11)$$

sendo $z = 0, \dots, \omega$, ω o número de observações e α o nível de significância. O nível de significância α depende do tamanho da janela ω e pode ser calculado automaticamente da seguinte forma (LEMOS; CAMINHAS; GOMIDE, 2011):

$$\alpha = \begin{cases} 0.01, & \text{se } \omega \geq 100 \\ 0.05, & \text{se } 20 \leq \omega < 100 \\ 0.1, & \text{se } 10 \leq \omega < 20. \end{cases} \quad (12)$$

O valor z_i^t é o número de violações do limiar e calculado por:

$$z_i^t = \begin{cases} \sum_{k=0}^{\omega-1} r_i^{t-k}, & t > \omega \\ 0, & \text{caso contrário,} \end{cases} \quad (13)$$

em que r_i^t é definido por:

$$r_i^t = \begin{cases} 0, & D_{(x^t, \mu_i^t)} < \chi_{m, \alpha}^2 \text{ (para distância Mahalanobis)} \\ 0, & D_{(x^t, \mu_i^t)} < \min_i \left(\frac{\max(x_i) - \min(x_i)}{2\omega} \right)^2 \text{ (para distância Euclidiana)} \\ 1, & \text{caso contrário.} \end{cases} \quad (14)$$

O limiar do índice de alerta Γ_a é usado para criar novos grupos e calculado por:

$$\Gamma_a = 1 - \frac{\alpha}{\omega}, \quad (15)$$

como a cada observação o ω é utilizado no cálculo do índice de alerta, é aplicada uma correção de comparações múltiplas, visando evitar a criação de falsos grupos.

Um novo grupo é criado se a medida de compatibilidade da amostra atual γ_i^t for menor que seu respectivo limiar Γ_{max} para todos os grupos e o índice de alerta $a_{i^*}^t$ do grupo com o maior grau de compatibilidade i^* é maior que seu respectivo limiar Γ_a . Mais precisamente, Se $\gamma_i^t < \Gamma_{max}, \forall i = 1, \dots, c^t$ e $a_{i^*}^t > \Gamma_a$ para $i^* = \max_i(\gamma_i^t)$, então um novo grupo é criado.

Após a criação, o número atual de grupos c^t é atualizado, o número de amostras do novo grupo é inicializado para $n_{c^t}^t = 1$, o centro do novo grupo é definido como $\mu_{c^t}^t = x^t$. Sua matriz de dispersão é inicializada $\sum_{c^t}^t = \sum_{init}$. Finalmente, os parâmetros do consequente $h_{c^t}^t$ são obtidos, como mostrado em (16), e a nova regra é criada.

$$h_{c^t}^t = \frac{\sum_{i=1}^{c^t} h_i^t \gamma_i^t}{\sum_{i=1}^{c^t} \gamma_i^t}. \quad (16)$$

Caso contrário, Se $\gamma_i^t > \Gamma_{max}, \forall i = 1, \dots, c^t$ e/ou $a_{i^*}^t < \Gamma_a$ para $i^* = \max_i(\gamma_i^t)$, então a amostra é atribuída para o grupo mais compatível e o centro deste grupo é atualizado usando:

$$\mu_{i^*}^{t+1} = \mu_{i^*}^t + S_{i^*}^t (x^t - \mu_{i^*}^t), \quad (17)$$

em que $S_{i^*}^t$ é calculado por:

$$S_{i^*}^t = \lambda (\gamma_{i^*}^t)^{1-a_{i^*}^t}, \quad (18)$$

na qual λ é a taxa de aprendizagem. Então, a matriz de dispersão do grupo é ajustada por (19), e os parâmetros do consequente são atualizados conforme descrito na Seção 3.7.

$$\sum_{i^*}^{t+1} = (1 - S_{i^*}^t) \left(\sum_{i^*}^t - S_{i^*}^t (x^t - \mu_{i^*}^t)(x^t - \mu_{i^*}^t)^T \right). \quad (19)$$

O Algoritmo 1 resume o método de criação e atualização dos grupos e regras.

Algoritmo 1: Procedimento para criação e atualização de grupos e regras.

```

if  $n_i^t < Nmax$  then
  | Calcular  $\Gamma_{max}$  - Eq. 8;
  |  $\sum_i^t =$  Matriz de identidade;
else
  | Calcular  $\Gamma_{max}$  - Eq. 9;
  |  $\sum_i^t =$  Matriz de dispersão;
end

Calcular  $\gamma_i^t$  para todos os grupos - Eq. 7 e 6;
Verificar se há violação de compatibilidade  $r^t$  - Eq. 14;
Verificar se a ocorrência está fora da janela definida  $z^t$  - Eq. 13;
Encontrar o índice do grupo com a maior compatibilidade  $i^*$ ;
Calcular  $a_{i^*}$  - Eq. 10;
Calcular  $\Gamma_a$  - Eq. 15;
if  $\gamma_i^t < \Gamma_{max}$  e  $a_{i^*}^t > \Gamma_a$  then
  | Atualizar o número de grupos  $c^t = c^t + 1$ ;
  | Criar o grupo  $c^t$ ;
  | Inicializar o número de amostras  $n_{c^t}^t = 1$ ;
  | Inicializar o centro  $\mu_{c^t}^t = x^t$ ;
  | Inicializar a matriz  $\sum_{c^t}^t = \sum_{init}$ ;
  | Inicializar os parâmetros dos consequentes  $h_{c^t}^t$  - Eq. 16;
  | Criar regra fuzzy  $c^t$ ;
  | Atualizar os índices dos grupos e regras;
else
  | Atualizar o grupo  $\mu_{i^*}^t$  - Eq. 17 e 18;
  | Atualizar a matriz de dispersão  $\sum_{i^*}^t$  - Eq. 19;
end

```

3.5 Mesclagem de Grupos e Regras

Na abordagem proposta, a mesclagem dos grupos ocorre sempre que um par de grupos se sobrepõe notavelmente (AGUIAR; LEITE, 2020). Mais formalmente, se dois grupos i^* e i estão suficientemente próximos, ou seja, quando seus centros, μ_{i^*} e μ_i , são tais que:

$$\|\mu_{i^*}^t - \mu_i^t\| \leq \rho \quad (i = 1 \dots c^t \text{ e } i \neq i^*), \quad (20)$$

em que ρ é o limiar para a mesclagem dos grupos. O centro do novo grupo $\mu_{i^* \cup i}^t$ é baseado no método da média ponderada e obtido como:

$$\mu_{i^* \cup i}^t = \mu_{i^*}^t - \frac{n_i^t}{n_{i^*}^t + n_i^t} (\mu_{i^*}^t - \mu_i^t), \quad (21)$$

onde n_i^t é o número de amostras do i -ésimo grupo, da mesma forma para $n_{i^*}^t$. O novo grupo $i^* \cup i$ está alinhado entre i^* e i , e sua posição dependerá do número de amostras dos grupos mesclados. A matriz de dispersão do grupo resultante é atualizada e definida por:

$$\sum_{i^* \cup i}^t = \frac{\sum_{i^*}^t + \sum_i^t}{2}. \quad (22)$$

Os parâmetros do consequente da regra recém-gerada são obtidos por:

$$h_{i^* \cup i}^t = \frac{h_{i^*}^t \gamma_{i^*}^t + h_i^t \gamma_i^t}{\gamma_{i^*}^t + \gamma_i^t}. \quad (23)$$

O número de amostras do novo grupo é obtido pela soma do número de amostras dos dois grupos mesclados, ou seja, $n_{i^* \cup i}^t = n_i^t + n_{i^*}^t$. Finalmente, o número de grupos e o índice são atualizados. O processo de mesclagem dos grupos é ilustrado no Algoritmo 2.

Algoritmo 2: Procedimento para mesclagem de grupos e regras.

```
// i* - índice do grupo atualizado ou criado
for i = 1...ct do
  if i* ≠ i e ||μi*t - μit|| < ρ then
    Calcular o centro μi* ∪ it - Eq. 21;
    Calcular a dispersão da matriz ∑i* ∪ it - Eq. 22;
    Calcular os parâmetros dos consequentes hi* ∪ it - Eq. 23;
    Calcular o número de amostras ni* ∪ it = nit + ni*t;
    Atualizar o número dos grupos ct = ct - 1;
    Atualizar os índices dos grupos e regras;
end
```

3.6 Eliminação de Grupos e Regras

A abordagem proposta para eliminação dos grupos é baseada no conceito de idade e população (ANGELOV; FILEV, 2005; LUGHOFER; ANGELOV, 2011). Neste trabalho, o conceito de idade é aplicado para definir o intervalo de tempo que um grupo permanece

sem ser ativado, ou seja, grau de pertinência diferente de zero. A idade de um grupo é obtida por:

$$idade_i^t = t - A_i, \quad (24)$$

onde i é o índice do grupo, A_i a instância de tempo que o i -ésimo grupo foi ativado e t é o passo atual. Para cada amostra x^t , encontre o índice do grupo inativo mais antigo \bar{i} . O grupo inativo mais antigo é excluído se

$$idade_{\bar{i}}^t > \omega, \quad (25)$$

em que $idade_{\bar{i}}^t$ é o tempo em que o grupo indexado por \bar{i} está inativo e ω é o tamanho da janela.

O outro método usado na eliminação de grupos é baseado na população, que representa o número de amostras n_i^t atribuídas a um grupo (ANGELOV; FILEV, 2005). A população do grupo é monitorada e se for inferior a 1% do total de amostras no instante t , o grupo é excluído, ou seja, se

$$\frac{n_{\bar{i}}^t}{t} < 0.01. \quad (26)$$

Para excluir um grupo, para cada nova amostra x_t encontre o grupo inativo mais antigo \bar{i} . O grupo indexado \bar{i} é excluído se

$$idade_{\bar{i}}^t > \omega \text{ e } \frac{n_{\bar{i}}^t}{t} < 0.01. \quad (27)$$

Depois de eliminar um grupo, o índice e o número de grupos são atualizados. A interseção dos dois mecanismos para a remoção do grupo garante que os grupos novos não sejam removidos. O Algoritmo 3 detalha o processo de exclusão do grupo.

Algoritmo 3: Procedimento para eliminação de grupos e regras.

Atualizar $idade_{i^*}$ - para o grupo atualizado ou criado - Eq. 24;

Encontrar \bar{i} - índice do grupo inativo mais antigo;

if ($idade_{\bar{i}}^t > \omega$) e $\left(\frac{n_{\bar{i}}^t}{t} < 0.01\right)$ **then**

 Excluir o grupo indexado por \bar{i} ;

 Excluir a regra indexada por \bar{i} ;

 Atualizar os índices dos grupos e regras;

 Atualizar o número de grupos $c^t = c^t - 1$;

3.7 Atualização dos Parâmetros do Consequente

Como descrito anteriormente, neste trabalho são propostos quatro modelos que se diferem pelo método de atualização dos parâmetros do consequente, a saber: eFLS (Seção

3.7.1), eFMI (Seção 3.7.2), eFCE (Seção 3.7.3) e eFTLS (Seção 3.7.4). Vale ressaltar que os algoritmos utilizam a mesma abordagem para a criação, atualização, mesclagem e exclusão de grupos.

3.7.1 eFLS - *evolving Fuzzy with Multivariable Gaussian Participatory Learning and Weighted Recursive Least Squares*

O procedimento para adaptação dos parâmetros do consequente do eFLS é realizado pelo algoritmo recursivo de mínimos quadrados ponderados (WRLS - *Weighted Recursive Least Squares*) (LUGHOFER, 2015). O WRLS busca maximizar a eficiência na estimação dos parâmetros dando a cada amostra a quantidade adequada de influência sobre a estimativa dos parâmetros. Salienta-se ainda que, pela função de otimização ser do tipo parábola, o WRLS consegue convergir em um ciclo de interação para o ótimo global (LUGHOFER, 2020). Porquanto, a fórmula para atualização dos parâmetros é dada por:

$$h_i^t = h_i^{t-1} + H_i^t x^{t-1} \tau_i(x^{t-1}) [y_i^{t-1} - ((x^{t-1})^T h_i^{t-1})] \quad (28)$$

em que H_i^t é obtido por:

$$H_i^t = H_i^{t-1} - \frac{\tau_i(x^{t-1}) H_i^{t-1} x^{t-1} (x^{t-1})^T H_i^{t-1}}{1 + (x^{t-1})^T H_i^{t-1} x^{t-1}}, \quad (29)$$

em que h_i^t é o parâmetro do consequente e H_i^t é a matriz da fórmula de atualização para regra i na iteração t . Os parâmetros do consequente do eFLS são inicializados com $h_i^0 = [y^0 \ 0 \ \dots \ 0]$ e $H^0 = \omega I_{m+1}$, onde ω é o tamanho da janela e I_{m+1} é uma matriz identidade $m + 1 \times m + 1$.

O Algoritmo 4 sumariza o procedimento para atualização dos parâmetros do consequente utilizando o algoritmo recursivo de mínimos quadrados ponderados.

Algoritmo 4: Atualização dos parâmetros do consequente no eFLS.

```

for  $i = 1 \dots c^t$  do
  | Calcular  $H_i^t$  - Eq. 29;
  | Atualizar  $h_i^t$  - Eq. 28;
end

```

3.7.2 eFMI - *evolving Fuzzy with Multivariable Gaussian Participatory Learning and Multi-Innovations Recursive Weighted Least Squares*

Os parâmetros do consequente das regras no eFMI são atualizados usando o um algoritmo recursivo de Mínimos Quadrados Ponderados com Multi-inovações (MI-WRLS -

Multi-Innovation Weighted Recursive Least Squares). De acordo com [Lughofer \(2020\)](#), o MI-WRLS torna-se uma opção para melhor convergência no caso de mudanças estruturais, que podem ser usadas de forma incremental e independentemente do método de atualização do antecedente escolhido. No MI-WRLS, o comprimento da inovação v_r é introduzido, o que orienta a quantidade de estimativas do vetor de inovação. Com isso, em cada interação, em comparação com o algoritmo RLS e ao WRLS, a precisão da estimação dos parâmetros é melhorada ([DING; CHEN; LI, 2007](#)). O vetor de inovação v_r representa os erros passados e são definidos por:

$$E_{(v_r,t)} = \begin{bmatrix} y^t - (\varphi_i^t)^T h_i^{t-1} \\ y^{t-1} - (\varphi_i^{t-1})^T h_i^{t-1} \\ \cdot \\ \cdot \\ \cdot \\ y^{t-v_r+1} - (\varphi_i^{t-v_r+1})^T h_i^{t-1} \end{bmatrix},$$

na qual $\varphi_i^t = [1 \ x_1^t \ x_2^t \ \dots \ x_m^t]^T$.

Para atualizar os parâmetros do consequente, inicialmente são definidos os vetores contendo as entradas $\theta_{i(v_r,t)}$ (30) e as saídas $Y_{i(v_r,t)}$ (31) de t a $t - v_r + 1$.

$$\theta_{i(v_r,t)} = [\varphi_i^t, \varphi_i^{t-1}, \dots, \varphi_i^{t-v_r+1}]. \quad (30)$$

$$Y_{i(v_r,t)} = [y_i^t, y_i^{t-1}, \dots, y_i^{t-v_r+1}]. \quad (31)$$

Então, uma matriz de ganho L_i^t é definida, e o inverso da matriz Hessiana ponderada P_i^t é atualizado.

$$L_i^t = P_i^{t-1} \theta_{i(v_r,t)} [Q(\tau_i^t) + \theta_{i(v_r,t)}^T P_i^{t-1} \theta_{i(v_r,t)}]^{-1}, \quad (32)$$

em que $Q(\tau_i^t)$ é calculado por:

$$Q(\tau_i^t) = \text{diag} \left(\left[\frac{1}{\tau_i^t}, \dots, \frac{1}{\tau_i^{t-v_r+1}} \right] \right), \quad (33)$$

e P_i^t é obtido por:

$$P_i^t = P_i^{t-1} - L_i^t \theta_{i(v_r,t)} P_i^{t-1}. \quad (34)$$

Finalmente, os parâmetros do consequente são atualizados por:

$$h_i^t = h_i^{t-1} + L_i^t [Y_{i(v_r,t)} - \theta_{i(v_r,t)} h_i^{t-1}]. \quad (35)$$

Os parâmetros do consequente são inicializados como $P_0 = p_0 I_m$ e $\theta_0 = \frac{1_m}{p_0}$, onde $p_0 = 10^6$ e 1_m é um vetor com dimensão m com todos os elementos iguais a 1 (DING; LIU; LIU, 2010). O Algoritmo 5 ilustra o procedimento para atualizar os parâmetros do consequente do eFMI.

Algoritmo 5: Atualização dos parâmetros do consequente no eFMI.

```

for  $i = 1 \dots c^t$  do
  Atualizar  $\varphi_i^t$ ;
  Atualizar  $\theta_i(v_r, t)$  - Eq. 30;
  Atualizar  $Y_i(v_r, t)$  - Eq. 31;
  Calcular  $Q(\tau_i^t)$  - Eq. 33;
  Calcular  $L_i^t$  - Eq. 32;
  Calcular  $P_i^t$  - Eq. 34;
  Atualizar  $h_i^t$  - Eq. 35;
end

```

3.7.3 eFCE - *evolving Fuzzy with Multivariable Gaussian Participatory Learning and Recursive Maximum Correntropy*

O procedimento para atualizar os parâmetros do consequente das regras do eFCE é baseado na Máxima Correntropia. Neste método minimiza-se o erro de saída maximizando a densidade de probabilidade do erro na fonte (BAO et al., 2018). As técnicas de ajuste de parâmetros normalmente se baseiam no critério MSE (*Mean Square Error*). Por outro lado, a adaptação dos parâmetros na Máxima Correntropia tem um fator de escala que é uma função exponencial do erro. Esta função visa reduzir a influência de possíveis discrepâncias ou níveis de ruído atípicos (LUGHOFER, 2020). Os parâmetros do consequente h_i^t são atualizados como segue:

$$h_i^t = h_i^{t-1} + \eta(x^t) H_i^t x^{t-1} \tau_i(x^{t-1}) [y_i^{t-1} - ((x^{t-1})^T H_i^{t-1} h_i^{t-1})], \quad (36)$$

em que H_i^t é obtido por:

$$H_i^t = H_i^{t-1} - \frac{\tau_i(x^{t-1}) H_i^{t-1} x^{t-1} (x^{t-1})^T H_i^{t-1}}{\eta(x^t) \tau_i(x^{t-1}) + (x^{t-1})^T H_i^{t-1} x^{t-1}}, \quad (37)$$

na qual o termo $\eta(x^t) = e^{-\frac{e_i^2}{2\sigma_r^2}}$, $e_i^t = (y^t - \hat{y}^t)$, $\sigma_r > 0$ é a largura do *kernel* e $\varrho \in [0,1]$ é um fator de esquecimento opcional.

O termo $\eta(x^t)$ é usado para dissipar *outliers* e atribui pesos menores a um erro maior durante o aprendizado. O parâmetro σ_r orienta o grau de dissipação dos erros maiores e é geralmente definido como uma fração da faixa de saída, definida principalmente entre 0,1 e 0,5 (LUGHOFER, 2020). Neste estudo, definiu $\sigma_r = 0,1$ e $\varrho = 1$ para todos os experimentos. Para a primeira amostra, os parâmetros do consequente são ajustados como $h_i^0 = [y^0 \ 0 \ \dots \ 0]$ e $H^0 = \omega I_{m+1}$, onde ω é o tamanho da janela e I_{m+1} é uma matriz identidade $m + 1 \times m + 1$.

O Algoritmo 6 sumariza o procedimento para atualização dos parâmetros do consequente utilizando a Máxima Correntropia.

Algoritmo 6: Atualização dos parâmetros do consequente no eFCE.

```

for  $i = 1 \dots c^t$  do
  Calcular  $e_i^t$ ;
  Calcular  $\eta(x^t)$ ;
  Calcular  $H_i^t$  - Eq. 37;
  Atualizar  $h_i^t$  - Eq. 36;
end

```

3.7.4 eFTLS - *evolving Fuzzy with Multivariable Gaussian Participatory Learning and Recursive Weighted Total Least Squares*

O eFTLS consiste na atualização dos parâmetros do consequente com base no algoritmo recursivo de mínimos quadrados totais ponderados (RWTLs - *Recursive Weighted Total Least Squares*) demonstrado em Davila (1994). Os algoritmos para atualização dos parâmetros do consequente descritos nas seções 3.7.1, 3.7.2 e 3.7.3 consideram para otimização o erro em relação a saída. No entanto, observa-se que na prática, podem ocorrer ruídos nos dados de entrada. Dessa forma, o RWTS propõe-se a remover tais anomalias, visando obter uma estimativa de parâmetros sem viés no caso das entradas e saídas ruidosas (LUGHOFER, 2020). Porquanto, a fórmula para atualização dos parâmetros é dada por:

$$h_i^t = \frac{g_i^t u_i^t - (u_{i1}^t x_1^t + \dots + u_{im}^t x_m^t)}{u_{i0}^t}, \quad (38)$$

na qual m é a dimensão do espaço de entrada, u_i^t é o menor autovetor da matriz de dispersão, g_i^t é um ponto pelo qual o modelo de regressão passa e $g_i^t u_i^t$ obtido por:

$$g_i^t u_i^t = u_{i0}^t y^t + u_{i1}^t x_1^t + \dots + u_{im}^t x_m^t. \quad (39)$$

Porém antes, o vetor regressor \hat{r}_i^t e o vetor da média ponderada $\nu_{r_i}^t$ da saída e das entradas da amostra atual são atualizados, ou seja,

$$\nu_{r_i}^t = \frac{(\bar{r}_i^t)^T \tau(x_i^t)}{1^t \tau(x_i^t)}, \quad (40)$$

e

$$\hat{r}_i^t = \bar{r}_i^t - \nu_{r_i}^t, \quad (41)$$

em que $\tau(x_i^t)$ são as funções de pertinências normalizadas (Eq. 5) e \bar{r}_i^t é obtido por:

$$\bar{r}_i^t = [y^t | x^t]. \quad (42)$$

Em seguida, a inversa da matriz Hessiana ponderada P_i é atualizada por:

$$P_i^t = I - L^{t-1} (\hat{r}_i^t)^T P_i^{t-1}, \quad (43)$$

com

$$L_i^{t-1} = \frac{P_i^{t-1} \hat{r}_i^t}{\frac{1}{\tau(x_i^t)} + (\hat{r}_i^t)^T P_i^{t-1} \hat{r}_i^t}. \quad (44)$$

Calcula o vetor normal unitário para o hiperplano afim da regra i ,

$$u_i^t = \frac{u_i^t}{\|u_i^t\|_2}, \quad (45)$$

$$u_i^t = P_i^t \left(\sum_i^t * u_i^t \right) \quad (46)$$

sendo \sum_i^t é uma matriz identidade usando a distância de Euclidiana. Caso contrário, \sum_i^t é a matriz de dispersão do grupo i . Os parâmetros do consequente são inicializados com $h_i^0 = [y^0 \ 0 \ \dots \ 0]$ e $H^0 = \omega I_{m+1}$, onde ω é o tamanho da janela e I_{m+1} é uma matriz identidade $m + 1 \times m + 1$.

O Algoritmo 7 ilustra o procedimento para atualizar os parâmetros do consequente do eFTLS.

Algoritmo 7: Atualização dos parâmetros do consequente no eFTLS.

```

for  $i = 1 \dots c^t$  do
  Atualizar  $\nu_{r_i}^t$ ; - Eq. 40
  Atualizar  $\hat{r}_i^t$ ; - Eq. 41
  Atualizar  $L_i^{t-1}$  - Eq. 44;
  Atualizar  $P_i^t$  - Eq. 43;
  Calcular  $u_i^t$  - Eq. 45 e Eq. 46;
  Calcular  $g_i^t u_i^t$ ; - Eq. 39;
  Atualizar  $h_i^t$  - Eq. 38;
end

```

3.8 Parâmetros e Algoritmo da Abordagem Proposta

Os modelos propostos possuem, respectivamente, seis e cinco parâmetros:

- λ é a taxa de aprendizagem usada na atualização do centro dos grupos. O valor de λ geralmente é definido entre 10^{-1} e 10^{-5} (LEMOS; CAMINHAS; GOMIDE, 2011).
- ρ é o parâmetro auxiliar para mesclagem dos grupos, com $\rho \in [0,1]$. Valores de ρ próximos de 0 significam que os grupos mesclados são os mais sobrepostos; valores próximos a 1, os grupos mais distantes (AGUIAR; LEITE, 2020). A prática sugere que $\rho \in [0,25, 0,45]$.
- ω é o tamanho da janela para o mecanismo de alerta e na eliminação dos grupos por idade. A prática sugere que o tamanho da janela ω fique na faixa de 5 a 200.
- \sum_{init} é a matriz de dispersão inicial usada para calcular a distância de Mahalanobis na criação de novos grupos. Os valores de \sum_{init} são geralmente ajustados para um valor baixo entre 10^{-1} e 10^{-4} (LEMOS; CAMINHAS; GOMIDE, 2011).
- N_{max} é o parâmetro que define o número de elementos pertencentes aos grupos que devem ser considerados para que a distância Euclidiana ou distância de Mahalanobis seja utilizada no cálculo da medida de compatibilidade (ŠKRJANC, 2020). Sugere-se $N_{max} \geq 2 * m$, com m o número de variáveis de entrada, evitando assim o problema de singularidade durante o cálculo da inversa da matriz de dispersão, quando é usado a distância de Mahalanobis. Por outro lado, usando $N_{max} = 0$, o modelo irá calcular apenas a distância de Mahalanobis.
- v_r é o comprimento da inovação, onde indica o número de erros passados que serão considerados para o cálculo dos parâmetros do consequente. Quando $v_r = 1$, o algoritmo MI-WRLS se reduz ao WRLS. Para a prática, $v_r \in [3, 20]$ é indicado. Esse parâmetro é utilizado, somente, para o eFMI.

Os procedimentos para cálculo da saída, evolução da estrutura e atualização dos parâmetros da abordagem proposta são sumarizados pelo Algoritmo 8.

Algoritmo 8: Pseudocódigo da abordagem proposta.

Entrada: $x^t, y^t, \rho, \omega, \lambda, \sum_{init}, N_{max}, v_r$
 Saída: \hat{y}^t
 Calcular α - Eq. 12;

for $t=1\dots$ **do**
 | Ler x^t ;
 | **if** $t == 1$ **then**
 | | Inicializar número de grupos $c^t = 1$;
 | | Criar primeiro grupo;
 | | Inicializar número de amostras $n_{c^t}^t = 1$;
 | | Inicializar centro $\mu_{c^t}^t = x^t$;
 | | Inicializar matriz $\sum_{c^t}^t = \sum_{init}$;
 | | Inicializar os parâmetros do consequente $h_{c^t}^t$;
 | | Criar primeira regra *fuzzy*;
 | **else**
 | | **for** $i=1\dots c^t$ **do**
 | | | **if** $n_i < N_{max}$ **then**
 | | | | $D(x^t, \mu_i^t)$ é a distância Euclidiana;
 | | | | **else**
 | | | | $D(x^t, \mu_i^t)$ é a distância de Mahalanobis;
 | | | | **end**
 | | | Calcular τ_i^t - Eq. 5;
 | | | Calcular y_t^i - Eq. 3;
 | | | **end**
 | | Calcular \hat{y}^t - Eq. 4;
 | | Ler y^t ;
 | | Criação e atualização dos grupos e regras - Alg. 1;
 | | Mesclagem de grupos e regras - Alg. 2;
 | | Eliminação de grupos e regras - Alg.3;
 | | Atualização dos parâmetros do consequente - Alg. 4 (eFLS), Alg. 5
 | | (eFMI), Alg. 6 (eFCE) e Alg. 7 (eFTLS);
 | **end**
end

4 Experimentos e Resultados

Este capítulo apresenta experimentos computacionais realizados para avaliar e comparar os modelos propostos com modelos alternativos do estado da arte. A Seção 4.1 apresenta a metodologia dos experimentos e as medidas de desempenho utilizadas. Os modelos propostos são avaliados na identificação de sistemas não lineares na Seção 4.2, na previsão da concentração de CO_2 na Seção 4.3, na previsão da produção de energia elétrica na Seção 4.4, na previsão de temperatura na Seção 4.5, na previsão dos preços dos títulos da Global 40 na Seção 4.6 e, por fim, na previsão de aluguel de bicicletas na Seção 4.7.

4.1 Metodologia

Os modelos propostos são avaliados em problemas de identificação de sistemas não lineares e previsão. Os resultados obtidos pelo eFCE, eFLS, eFMI e eFTLS são comparados com três sistemas evolutivos: eMG (LEMOS; CAMINHAS; GOMIDE, 2011), eNFN (SILVA et al., 2014) e eOGS (LEITE; ŠKRJANC, 2019), com o código em Matlab e disponibilizados pelos respectivos autores. Os modelos propostos foram desenvolvidos em Matlab.

Na literatura não existe uma padronização na construção de *workflows* para orientar o desenvolvimento, avaliação, teste, validação e comparação de desempenho de algoritmos em ambientes não estacionários (ŠKRJANC et al., 2019b). Assim, neste trabalho, optou-se por implementar uma metodologia na qual os experimentos computacionais simulam o processamento *online* como um fluxo de dados, ou seja, os parâmetros e a estrutura dos modelos evoluem para todas as amostras do conjunto de dados. Os melhores valores dos parâmetros dos modelos são obtidos por meio de uma busca exaustiva. Os valores dos parâmetros que obtém o menor erro são usados para avaliar a modelagem e o desempenho. Os melhores valores e a faixa de variação dos parâmetros dos modelos podem ser visualizados na Tabela 2. Os dados foram normalizados para o intervalo entre $[0,1]$ para todos os experimentos.

O desempenho dos modelos é avaliado para todas as amostras do conjunto de dados usando a Raiz do Erro Quadrado Médio (RMSE - *Root of the Mean Square Error*) e o Índice de Erro Não Dimensional (NDEI - *Non-Dimensional Error Index*), ou seja,

$$RMSE = \sqrt{\frac{\sum_{t=1}^N (y^t - \hat{y}^t)^2}{N}} \quad (47)$$

e

$$NDEI = \frac{RMSE}{\sigma(\hat{y}^t)}, \quad (48)$$

em que N é o número de amostras, y^t é a saída desejada no passo t , \hat{y}^t é a saída do modelo no passo t e σ é o desvio padrão da saída do sistema. O RMSE (47) calcula a média dos erros do modelo ao quadrado, dando menos importância para diferenças menores e mais importância para diferenças maiores. O NDEI (48) avalia o erro independente da escala dos dados e é obtido pela divisão do valor RMSE pelo desvio padrão (σ) da saída desejada.

Além do RMSE e NDEI, assim como em [Silva et al. \(2014\)](#), a Soma dos Quadrados Residuais (RSS - *Residual Sum of Squares*) foi empregada para analisar o desempenho dos modelos ao longo do tempo. O RSS (49) é calculado pela soma ao quadrado do resíduo para todas as amostras processadas até a etapa N .

$$RSS = \sum_{t=1}^N (y^t - \hat{y}^t)^2. \quad (49)$$

As medidas de erro são índices úteis para medir a precisão dos modelos, mas podem não considerar a superioridade de um modelo sobre outro em termos de significância estatística. O teste MGN (*Morgan-Granger-Newbold*) ([DIEBOLD; MARIANO, 2002](#)) é usado para avaliar modelos de predição e aplicado onde se assume que os erros dos modelos não são correlacionados. Neste trabalho o MGN será utilizado para comparar o desempenho dos modelos propostos entre si e contra os modelos alternativos. O cálculo das estatísticas do teste é descrito por:

$$MGN = \hat{\rho}_{\bar{e}} \sqrt{\frac{N-1}{1 - \hat{\rho}_{\bar{e}}^2}}, \quad (50)$$

em que $\hat{\rho}_{\bar{e}}$ é a correlação de amostra entre $l_1 = \bar{e}_1 + \bar{e}_2$ e $l_2 = \bar{e}_1 - \bar{e}_2$, onde \bar{e}_1 e \bar{e}_2 são os resíduos dos dois modelos. A estatística do teste MGN tem distribuição t com $N - 1$ graus de liberdade. A hipótese nula é equivalente à igualdade das duas variâncias do erro de previsão, ou seja, a correlação entre l_1 e l_2 será zero.

4.2 Identificação da Função da Hipérbole

Esta seção tem como objetivo analisar o comportamento dos modelos na identificação da função de hipérbole. Sabe-se que a hipérbole é o conjunto de pontos no plano cuja diferença, em valor absoluto, da distância entre dois pontos fixos (chamados focos) é constante. Como em Škrjanc (2020), a função a ser modelada é obtida por:

$$y^t = \frac{1}{0.1 + \frac{1}{m} \sum_{j=1}^m (1 - x_j^t)}, \quad (51)$$

onde m é o número de variáveis, t o passo atual e as entradas x^t são geradas aleatoriamente a partir de um intervalo $[0, 1]$. Para o experimento foram criadas 2900 amostras. O objetivo é identificar a hipérbole usando entradas aleatórias em um intervalo definido, de acordo com:

$$\hat{y}^t = f(x_1^t, \dots, x_j^t, \dots, x_m^t), \quad (52)$$

em que \hat{y}^t é a saída do modelo.

4.2.1 Identificação da Função de Hipérbole - 7-D

A Tabela 3 mostra o RMSE e o NDEI na identificação da função de hipérbole com 7 variáveis de entrada (7-D). As medidas de erro sugerem que os modelos propostos eFCE, eFLS e eFMI obtiveram, respectivamente, os 3 melhores resultados. Em seguida, o melhor desempenho foi alcançado pelo eMG, seguido pelo eNFN, eFTLS e eOGS. A Figura 3 ilustra os valores desejados e estimados pelo eFCE.

A soma dos quadrados residuais - RSS - dos modelos na identificação da função de hipérbole 7-D é apresentada na Figura 4. Os resultados visualizados na Figura 4 sugerem que a curva do erro do eFCE manteve-se inferior à dos outros algoritmos para todas as amostras. Observa-se também que o eFLS apresentou uma curva, para as primeiras amostras, pior que a do eFMI e similar a do eMG, porém ao longo do tempo o eFLS superou os modelos eFMI e eMG. Já o eFTLS apresentou a pior curva nas primeiras amostras, mas conseguiu, ao final do experimento, superar a curva do eOGS.

Tabela 3 – Desempenho na identificação da função de hipérbole - 7-D.

Modelos	RMSE	NDEI
eFCE	0.0391	0.2750
eFLS	0.0398	0.2804
eFMI	0.0427	0.3008
eMG	0.0452	0.3180
eNFN	0.0543	0.3822
eFTLS	0.0558	0.3927
eOGS	0.0962	0.6750

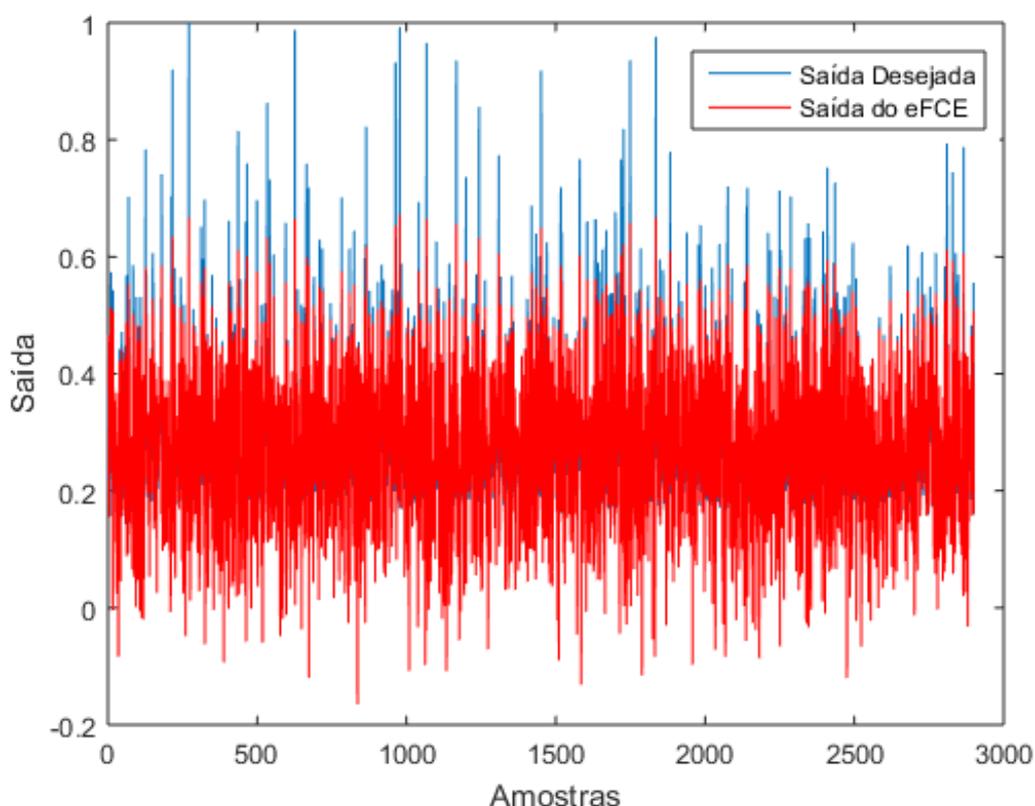


Figura 3 – Identificação da função de hipérbole 7-D pelo eFCE.

A Tabela 4 apresenta os resultados do teste MGN na comparação entre os pares de modelos na identificação da função hipérbole 7-D. A tabela mostra os valores da estatística do teste MGN e o p -valor correspondente, dado um nível de significância de 0,05. Assim, pode-se afirmar com base no MGN que o eFCE é estatisticamente superior ao eFMI, eFTLS, eMG, eNFN e eOGS. Entre o eFCE e o eFLS, conclui-se que não há diferença estatística entre os dois. Portanto, não se pode afirmar qual dos dois é de fato superior. Ao avaliar o eFLS pode-se afirmar que ele é estatisticamente superior ao eFMI, eFTLS, eMG, eNFN e eOGS. Considerando o eFMI pode-se concluir que ele é estatisticamente superior ao eFTLS, eMG, eNFN e eOGS. Além disso, que o eFMI é estatisticamente inferior ao eFLS e ao eFCE. Neste experimento o eFTLS foi estatisticamente superior ao eOGS e inferior aos

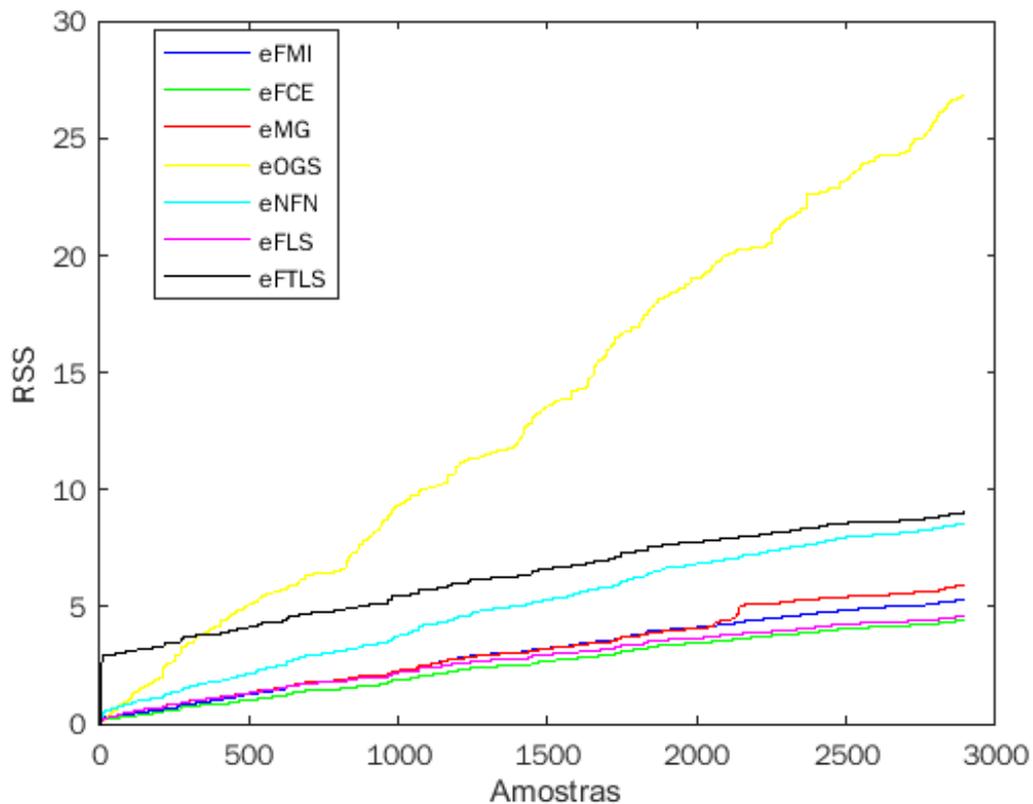


Figura 4 – Soma dos quadrados residuais na identificação da função de hipérbole - 7-D.

demais modelos.

Tabela 4 – Avaliação dos modelos propostos pelo teste MGN na identificação da função de hipérbole - 7-D.

Modelos	eFCE MGN/ <i>p</i> -valor	eFLS MGN/ <i>p</i> -valor	eFMI MGN/ <i>p</i> -valor	eFTLS MGN/ <i>p</i> -valor	eMG MGN/ <i>p</i> -valor	eNFN MGN/ <i>p</i> -valor	eOGS MGN/ <i>p</i> -valor
eFCE	—	0.32/0.37	8.34/0.00	29.53/0.00	12.01/0.00	23.88/0.00	56.09/0.00
eFLS	-0.32/0.63	—	7.40/0.00	27.42/0.00	12.99/0.00	22.59/0.00	55.68/0.00
eFMI	-8.34/1.00	-7.40/1.00	—	20.77/0.00	7.21/0.00	19.24/0.00	51.46/0.00
eFTLS	-29.53/1.00	-27.42/1.00	-20.77/1.00	—	-15.25/1.00	-1.67/0.95	32.01/0.00

4.2.2 Identificação da Função de Hipérbole - 10-D

A Tabela 5 ilustra o desempenho dos modelos pelo RMSE e NDEI na identificação da função de hipérbole 10-D. Como pode ser visto, o eFCE obteve o melhor desempenho, seguido pelo eFLS, eFMI, eMG, eNFN, eFTLS e eOGS. A saída desejada e a saída estimada pelo eFCE é ilustrada na Figura 5.

A Figura 6 ilustra a soma dos quadrados residuais - RSS - dos modelos a cada passo t na identificação da função hipérbole 10-D. Analisando a Figura 6, pode-se observar que a curva de erro do eFCE é menor para todas as amostras do experimento. O eFMI apresentou um desempenho inferior ao eFLS e eMG nas primeiras amostras, porém ao

Tabela 5 – Desempenho na identificação da função de hipérbole - 10-D.

Modelos	RMSE	NDEI
eFCE	0.0292	0.2302
eFLS	0.0304	0.2392
eFMI	0.0314	0.2476
eMG	0.0326	0.2565
eNFN	0.04165	0.3277
eFTLS	0.0468	0.3688
eOGS	0.0918	0.7165

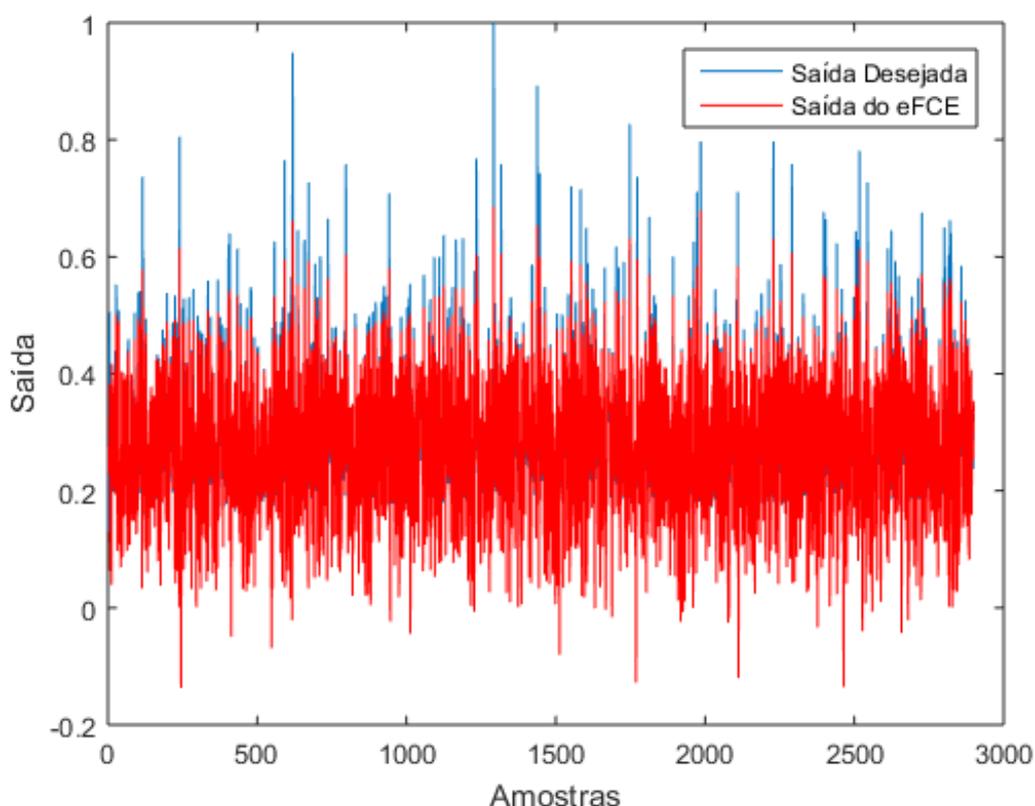


Figura 5 – Identificação da função de hipérbole 10-D pelo eFCE.

final da simulação superou o eMG. O eFTLS obteve o pior desempenho para as amostras iniciais, mas, em seguida, conseguiu superar o eOGS.

A Tabela 6 mostra as comparações entre pares de modelos usando o teste MGN na identificação de função de hipérbole com 10 variáveis de entrada, dado um nível de significância de 0,05. Os resultados mostram que o eFCE é estatisticamente superior a todos os modelos avaliados. Analisando os resultados do eFLS, constata-se que ele é superior estatisticamente aos modelos eFMI, eFTLS, eMG, eNFN e eOGS. Ao examinar os resultados do eFMI, conclui-se que ele é superior estatisticamente ao eFTLS, eMG, eNFN e eOGS, e inferior ao eFCE e eFLS. Por fim, o eFTLS é superior estatisticamente eOGS e

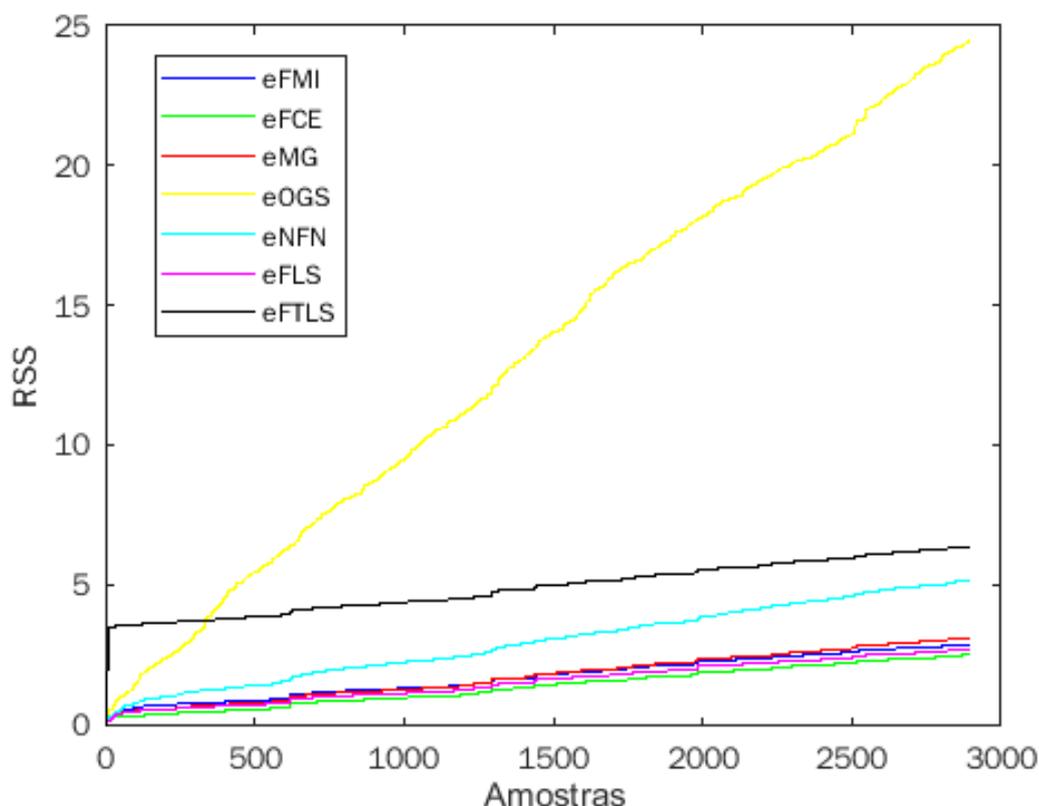


Figura 6 – Soma dos quadrados residuais na identificação da função de hipérbole - 10-D.

inferior aos demais.

Tabela 6 – Avaliação dos modelos propostos pelo teste MGN na identificação da função de hipérbole - 10-D.

Modelos	eFCE MGN/ <i>p</i> -valor	eFLS MGN/ <i>p</i> -valor	eFMI MGN/ <i>p</i> -valor	eFTLS MGN/ <i>p</i> -valor	eMG MGN/ <i>p</i> -valor	eNFN MGN/ <i>p</i> -valor	eOGS MGN/ <i>p</i> -valor
eFCE	—	1.72/0.05	4.71/0.00	40.86/0.00	9.00/0.00	25.62/0.00	76.69/0.00
eFLS	-1.72/0.95	—	11.58/0.00	35.58/0.00	13.76/0.00	25.03/0.00	76.09/0.00
eFMI	-4.71/1.00	-11.58/1.00	—	32.82/0.00	7.50/0.00	22.53/0.00	73.48/0.00
eFTLS	-40.87/1.00	-35.58/1.00	-32.82/1.00	—	-31.19/1.00	-6.61/1.00	40.17/0.00

4.3 Previsão da Concentração de CO_2

Esta seção considera o problema de concentração de CO_2 (dióxido de carbono) no forno a gás Box-Jenkins¹ (BOX; JENKINS, 1976). O conjunto de dados contém 296 amostras registradas, ao longo do tempo, da taxa de fluxo de gás no forno x^t e da concentração de CO_2 emitida y^t . O objetivo é prever o nível de CO_2 no passo atual t usando a observação anterior (y^{t-1}) e a taxa de fluxo de gás nos quatro passos anteriores (x^{t-1}, \dots, x^{t-4}), ou seja,

¹ <https://www.stat.purdue.edu/~chong/stat520/bjr-data/gas-furnace>

$$\hat{y}^t = f(y^{t-1}, x^{t-1}, x^{t-2}, x^{t-3}, x^{t-4}). \quad (53)$$

Os resultados experimentais obtidos na previsão da concentração de CO_2 são apresentados na Tabela 7. O melhor desempenho, considerando o RMSE e o NDEI, foi obtido pelo eFCE, seguido pelo eFMI, eFLS, eMG, eOGS, eNFN e eFTLS. Os valores desejados e estimados pelo eFCE para a previsão da concentração de CO_2 estão representados na Figura 7.

Tabela 7 – Desempenho na previsão da concentração de CO_2 .

Modelos	RMSE	NDEI
eFCE	0.0346	0.1601
eFMI	0.0366	0.1693
eFLS	0.0371	0.1716
eMG	0.0377	0.1744
eOGS	0.0602	0.2842
eNFN	0.0646	0.2999
eFTLS	0.5267	2.4341

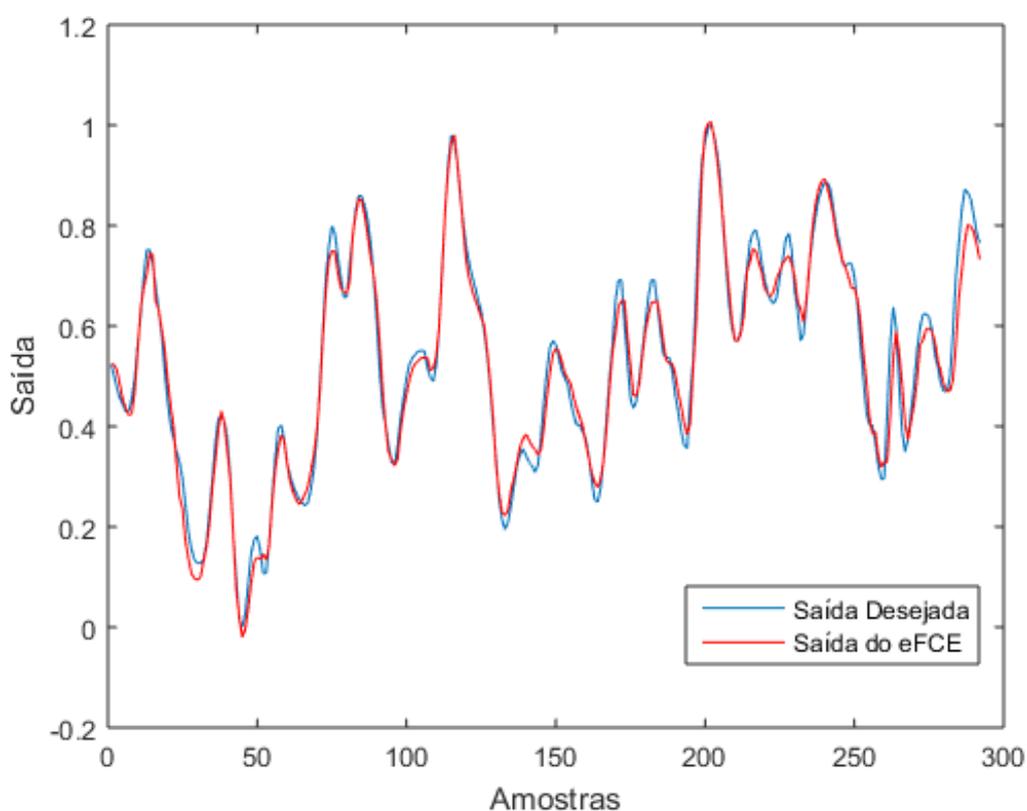


Figura 7 – Previsão da concentração de CO_2 pelo eFCE.

A Figura 8 ilustra o logaritmo da soma dos quadrados residuais - $\text{Log}(\text{RSS})$ - dos modelos na previsão da concentração de CO_2 . Por meio desta figura pode-se perceber

que o eFTLS obteve melhor desempenho até próximo da amostra duzentos. No entanto, após esse período o eFTLS obteve um erro elevado durante algumas amostras o que fez com que este modelo obtivesse o pior desempenho ao final do experimento. Além disso, nas amostras iniciais o eFMI obteve um desempenho inferior eMG, eFCE e eOGS. Porém, ao término do experimento o desempenho do eFMI foi similar ao do eMG, eFCE, eFCE, eFCE e superior ao do eOGS, eNFN e eFTLS. O eFCE foi o algoritmo que se manteve mais estável ao longo do experimento, sempre apresentando um bom desempenho. Vale ressaltar, o comportamento da curva de erro do eFTLS que apresentou um aumento significativo aproximadamente na amostra 200. Uma possibilidade seria o problema da singularidade ao calcular a inversa da matriz de dispersão do grupo quando este possui um pequeno número de amostras, pois o valor de $N_{max} = 9$, ou seja, não é maior ou igual a $2 * m$, com $m = 5$ o número de variáveis de entrada do experimento. No entanto, esse comportamento precisa melhor se investigado.

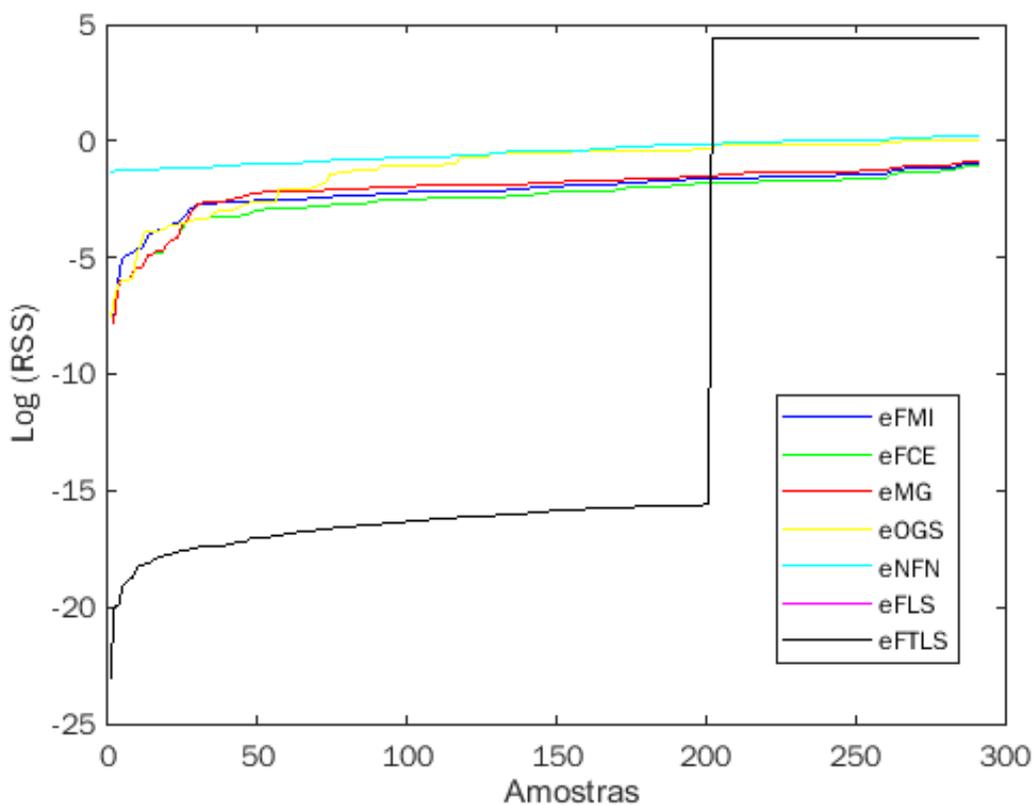


Figura 8 – Logaritmo da soma dos quadrados residuais na previsão da concentração de CO_2 .

Os resultados da comparação entre os modelos propostos e os modelos alternativos pelo teste MGN na previsão da concentração de CO_2 são apresentados na Tabela 8. Nesta análise considera-se um nível de significância de 0,05. Os resultados da mostram que o eFCE possui superioridade estatística sobre todos os demais modelos e que o eFCE é superior estatisticamente ao eNFN, eOGS e eFTLS. Comparando o eFCE em

relação ao eFMI e eMG, não se pode afirmar qual deles é de fato superior, ou seja, não há diferença estatística entre eles. Além disso, os resultados sugerem que o eFMI é estatisticamente superior ao eFTLS, eNFN e eOGS. Por fim, para este experimento, o eFTLS foi estatisticamente inferior aos demais modelos.

Tabela 8 – Avaliação dos modelos propostos pelo teste MGN na previsão da concentração de CO_2 .

Modelos	eFCE MGN/ <i>p</i> -valor	eFLS MGN/ <i>p</i> -valor	eFMI MGN/ <i>p</i> -valor	eFTLS MGN/ <i>p</i> -valor	eMG MGN/ <i>p</i> -valor	eNFN MGN/ <i>p</i> -valor	eOGS MGN/ <i>p</i> -valor
eFCE	—	3.72/0.00	4.70/0.00	131.43/0.00	3.90/0.00	12.94/0.00	11.22/0.00
eFLS	3.72/1.00	—	- 0.60/0.73	120.46/0.00	-0.85/ 0.44	10.36/0.00	9.12/0.00
eFMI	-4.70/1.00	0.60/0.27	—	122.51/0.00	0.61/0.27	11.27/0.00	9.56/0.00
eFTLS	-131.43/1.00	-120.46/1.00	-122.51/1.00	—	-120.69/1.00	-68.15/1.00	-74.52/1.00

4.4 Previsão da Produção de Energia Elétrica em Usina de Ciclo Combinado - CCGP

Esta seção considera a previsão da produção de energia elétrica, por hora, em plena carga de uma usina de ciclo combinado (CCGP-*Combined Cycle Power Plant*). O conjunto de dados² consiste em 9568 amostras coletadas em 674 dias diferentes entre 2006 e 2011, quando o CCGP foi configurado para funcionar em carga máxima. O objetivo é prever a saída de energia elétrica (y^t) usando como entrada quatro variáveis que influenciam o funcionamento do CCGP (KAYA; TUFEKCI, 2012; TUFEKCI, 2014): temperatura ambiente (x_1^t), pressão atmosférica (x_2^t), umidade relativa (x_3^t) e pressão do vapor de exaustão (x_4^t). O modelo para este conjunto de dados é definido por:

$$\hat{y}^t = f(x_1^t, x_2^t, x_3^t, x_4^t). \quad (54)$$

A Tabela 9 ilustra o RMSE e NDEI após a simulação da previsão da produção de energia elétrica na CCGP. Eles sugerem que o melhor desempenho foi alcançado pelo eFCE, seguido pelo eMG e eFMI. Os resultados obtidos por estes três modelos são comparáveis e superam os obtidos pelos demais modelos em mais de uma ordem de magnitude. A Figura 9 mostra a saída real e a saída estimada pelo eFCE.

A Figura 10 ilustra o logaritmo da soma dos quadrados residuais - Log(RSS) - na previsão da produção de energia elétrica. Como pode ser visto na figura o eFMI obteve o melhor desempenho nas amostras iniciais e depois foi superado pelo eFCE que obteve o melhor resultado ao final do experimento e pelo eMG que alcançou o segundo melhor desempenho. O eNFN obteve o quarto melhor desempenho neste experimento. Os modelos eFTLS e eFLS apresentaram ao longo do experimento um desempenho similar. Por fim, o algoritmo com o pior desempenho foi o eOGS.

² <https://archive.ics.uci.edu/ml/datasets/combined+cycle+power+plant>

Tabela 9 – Desempenho na previsão da produção de energia elétrica na CCPP.

Modelos	RMSE	NDEI
eFCE	0.0568	0.2514
eMG	0.0590	0.2610
eFMI	0.0626	0.2996
eNFN	0.0817	0.3533
eFTLS	0.1054	0.4321
eFLS	0.1058	0.4393
eOGS	0.1174	0.5257

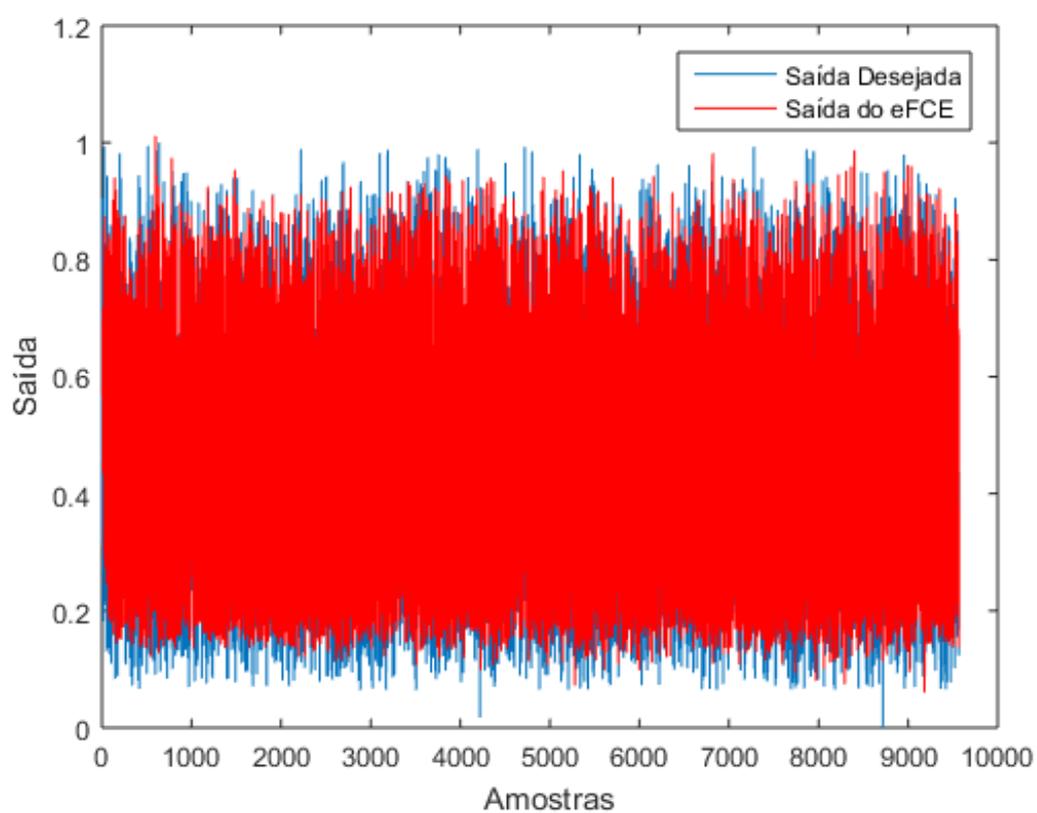


Figura 9 – Previsão da produção de energia elétrica na CCPP pelo eFCE.

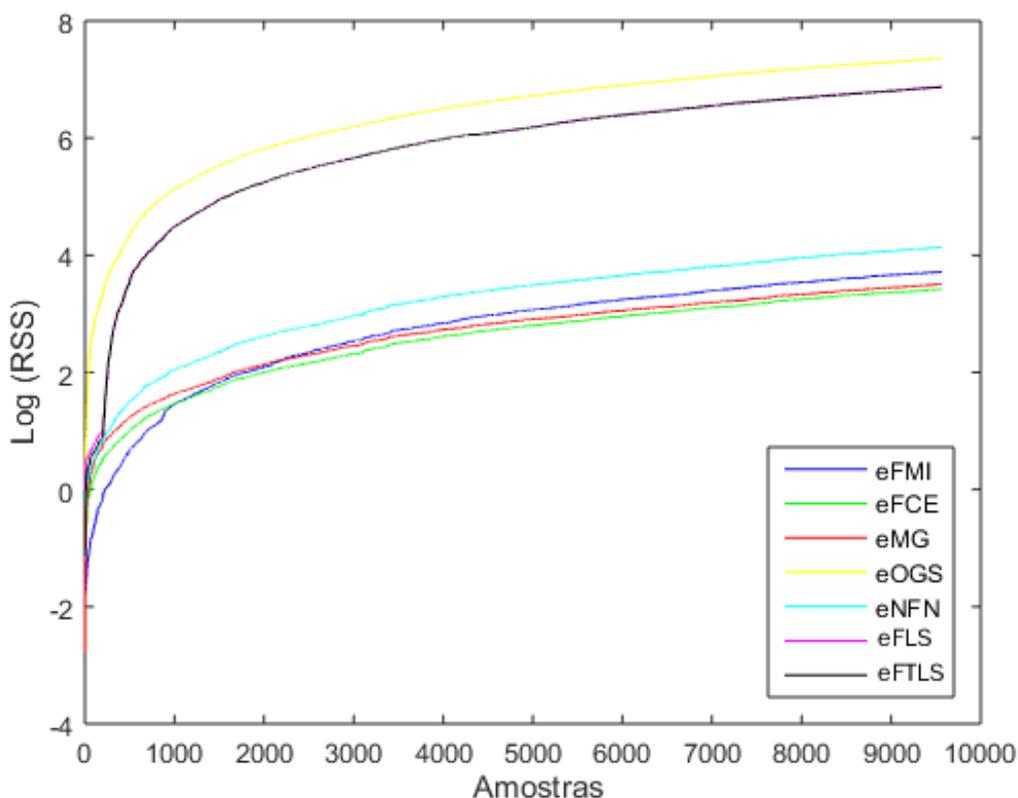


Figura 10 – Logaritmo da soma dos quadrados residuais na previsão da produção de energia elétrica na CCPP.

A Tabela 10 mostra os resultados das comparações entre os modelos propostos e os alternativos usando o teste MGN na previsão da produção de energia elétrica na CCPP, com 4 variáveis de entrada, dado um nível de significância de 0,05. Os resultados apresentados mostram que o eFCE é estatisticamente superior a todos os modelos avaliados e que o eFLS é estatisticamente superior ao eMG e ao eOGS, similar ao eFTLS e inferior aos demais. Pode-se concluir que o eFMI que é estatisticamente superior aos modelos eFLS, eFTLS, eNFN e eOGS e inferior ao eFCE e eMG. Em relação ao modelo eFTLS, constatou-se que ele é superior estatisticamente ao eOGS, similar ao eFLS e inferior aos demais modelos.

Tabela 10 – Avaliação dos modelos propostos pelo teste MGN na previsão da produção de energia elétrica na CCPP.

Modelos	eFCE MGN/p-valor	eFLS MGN/p-valor	eFMI MGN/p-valor	eFTLS MGN/p-valor	eMG MGN/p-valor	eNFN MGN/p-valor	eOGS MGN/p-valor
eFCE	—	154.65/0.00	4.76/0.00	154.02/0.00	3.90/0.00	14.00/0.00	281.94/0.00
eFLS	-154.65/1.00	—	-55.68/1.00	-0.68/0.65	72.69/0.00	-45.85/1.00	9.52/0.00
eFMI	-4.76/1.00	55.68/0.00	—	55.51/0.00	-3.89/1.00	8.27/0.00	179.56/0.00
eFTLS	-154.02/1.00	0.68/0.35	-55.51/1.00	—	-72.46/1.00	-45.15/1.00	9.12/0.00

4.5 Previsão de Temperatura no Vale da Morte

Nesta seção, avaliam-se os modelos na previsão da temperatura média no Vale da Morte³. No conjunto de dados meteorológicos do Vale da Morte são consideradas 1306 observações, com registro de temperatura média mensal de 1901 a 2009, em graus Celsius. O objetivo é prever a temperatura média mensal um passo à frente. Leite, Costa e Gomide (2012) e Garcia, Leite e Škrjanc (2020) sugerem utilizar os doze primeiros valores defasados da série como entradas. O modelo para este conjunto de dados é definido por:

$$\hat{y}^{t+1} = f(y^t, y^{t-1}, \dots, y^{t-11}, y^{t-11}). \quad (55)$$

A Tabela 11 ilustra os resultados obtidos na previsão de temperatura no Vale da Morte. O melhor desempenho em relação ao RMSE e NDEI foi obtido pelo eFCE, seguido pelo eFMI, eMG, eFTLS, eFLS, eNFN e eOGS. A Figura 11 ilustra a saída desejada e a estimada pelo eFCE na previsão de temperatura.

Tabela 11 – Desempenho na previsão de temperatura no Vale da Morte.

Modelos	RMSE	NDEI
eFCE	0.0508	0.2164
eFMI	0.0545	0.2221
eMG	0.0596	0.2307
eFTLS	0.0597	0.2389
eFLS	0.0608	0.2483
eNFN	0.0687	0.2659
eOGS	0.0785	0.3056

A soma dos quadrados residuais -RSS - a cada amostra t na previsão de temperatura é apresentada na Figura 12. Analisando a figura, pode-se observar que o melhor desempenho foi obtido pelo eFCE. O eFTLS e o eMG possuem um desempenho comparável ao longo do experimento. Destaca-se que nas primeiras amostras o eFMI apresentou um desempenho inferior ao do eFCE, eFLS, eFTLS e eMG, porém ao final amostragem conseguiu superar o eFTLS, eMG e eFLS. O eOGS apresentou o pior desempenho entre os modelos avaliados.

O teste estatístico MGN para os experimentos de previsão de temperatura no Vale da Morte é ilustrado na Tabela 12. Os resultados do teste MGN sugerem, com um nível de significância de 0,05, que o eFCE é estatisticamente superior aos modelos alternativos e propostos. A comparação também mostra que o desempenho do eFLS é estatisticamente superior ao do eNFN e do eOGS e que o do eFMI supera o do eFLS, eFTLS, eMG, eNFN e eOGS. Ademais, os resultados sugerem que o desempenho do eFTLS é superior ao do eFLS, eNFN e do eOGS e similar ao do eMG.

³ <https://www.nps.gov/deva/planyourvisit/weather.htm>

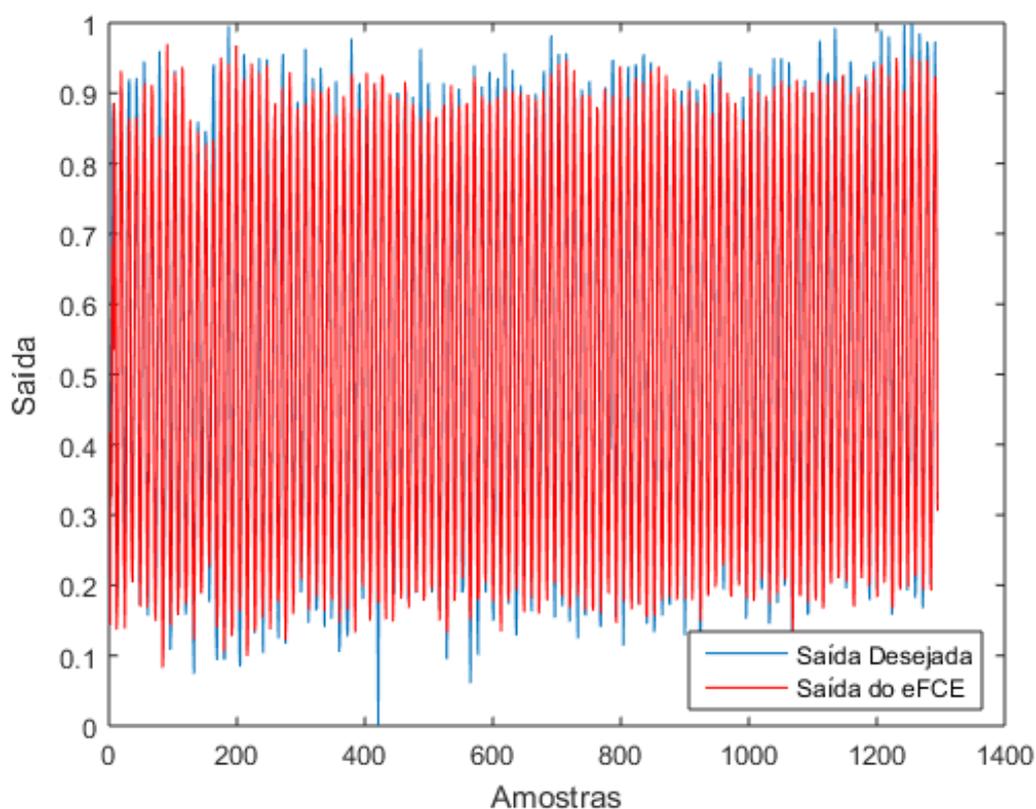


Figura 11 – Previsão de temperatura no Vale da Morte pelo eFCE.

Tabela 12 – Avaliação dos modelos propostos pelo teste MGN na previsão de temperatura no Vale da Morte.

Modelos	eFCE MGN/ <i>p</i> -valor	eFLS MGN/ <i>p</i> -valor	eFMI MGN/ <i>p</i> -valor	eFTLS MGN/ <i>p</i> -valor	eMG MGN/ <i>p</i> -valor	eNFN MGN/ <i>p</i> -valor	eOGS MGN/ <i>p</i> -valor
eFCE	—	12.35/0.00	3.63/0.00	8.95/0.00	7.72/0.00	21.32/0.00	41.40/0.00
eFLS	-12.35/1.00	—	-8.37/1.00	-4.61/1.00	-4.72/1.00	10.22/0.00	29.46/0.00
eFMI	-3.63/1.00	8.37/0.00	—	4.87/0.00	4.23/0.00	18.15/0.00	38.23/0.00
eFTLS	-8.95/1.00	4.61/0.00	-4.87/1.00	—	-0.79/0.78	14.16/0.00	37.03/0.00

4.6 Previsão dos Preços dos Títulos da Global 40

O Global 40 é o título brasileiro mais negociado. Esse título foi lançado em 2000 e expira em 2040. O movimento dos preços define a rentabilidade dos títulos. Portanto, prever o movimento dos preços é uma questão crucial para a tomada de decisão do investidor. O objetivo deste experimento é prever a direção dos preços dos títulos, pois os investidores acreditam que, quando os preços aumentam, a rentabilidade diminui, então você deve vender o título antes que isso aconteça. Se acontecer o contrário, será hora de comprar títulos. O conjunto de dados consiste em 1665 amostras extraídas dos preços diários dos títulos de 11 de agosto de 2000 a 16 de janeiro de 2007. Como em [Leite et al. \(2011\)](#), os dois primeiros valores defasados da série são usados como entrada para prever o preço, um passo à frente. O modelo para este conjunto de dados é definido por:

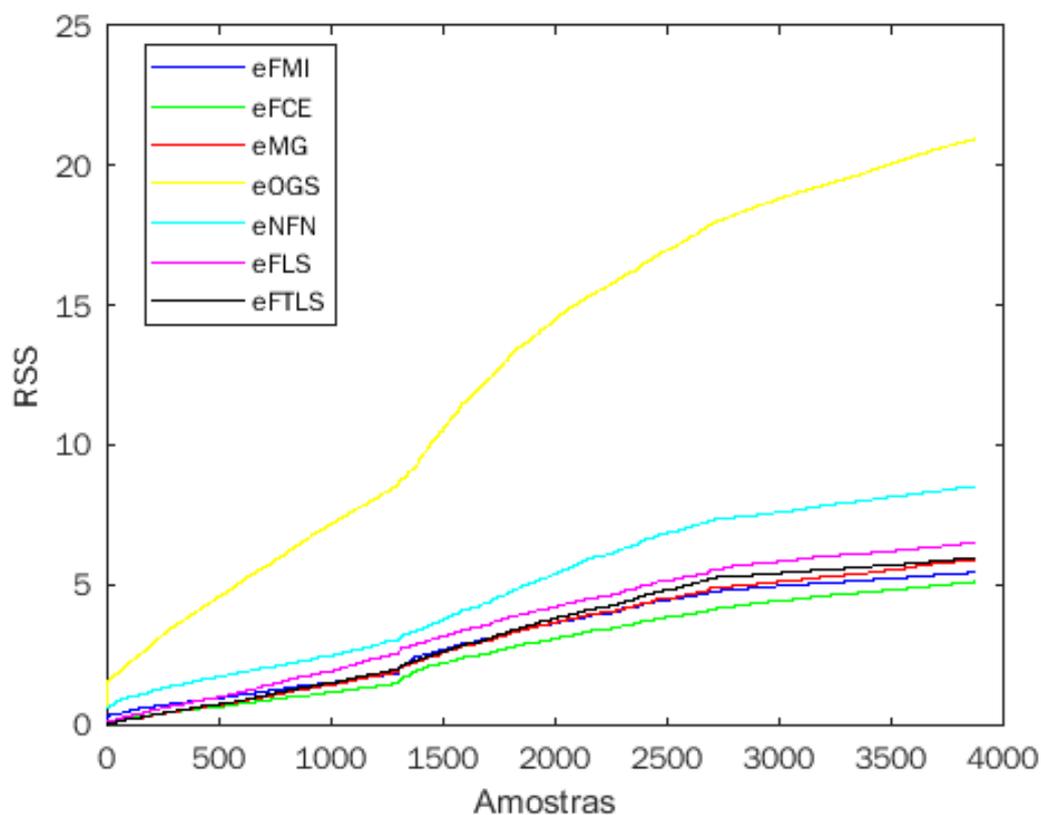


Figura 12 – Soma dos quadrados residuais na previsão de temperatura no Vale da Morte.

$$\hat{y}^{t+1} = f(y^t, y^{t-1}). \tag{56}$$

O desempenho dos modelos avaliado pelo RMSE e NDEI na previsão dos preços dos títulos da Global 40 está resumido na Tabela 13. Os melhores resultados foram alcançados pelos modelos eFMI, eFCE e eMG. Estes três modelos obtiveram resultados comparáveis e superaram os obtidos pelo eNFN, eFLS, eFTLS e eOGS. Os valores desejados e estimados pelo eFMI na previsão dos preços de títulos podem ser vistos na Figura 13.

Tabela 13 – Desempenho na previsão dos preços dos títulos do Global 40.

Modelos	RMSE	NDEI
eFMI	0.0113	0.0421
eFCE	0.0113	0.0422
eMG	0.0115	0.0428
eNFN	0.0159	0.0603
eFLS	0.0184	0.0683
eFTLS	0.0187	0.0694
eOGS	0.6341	2.3598

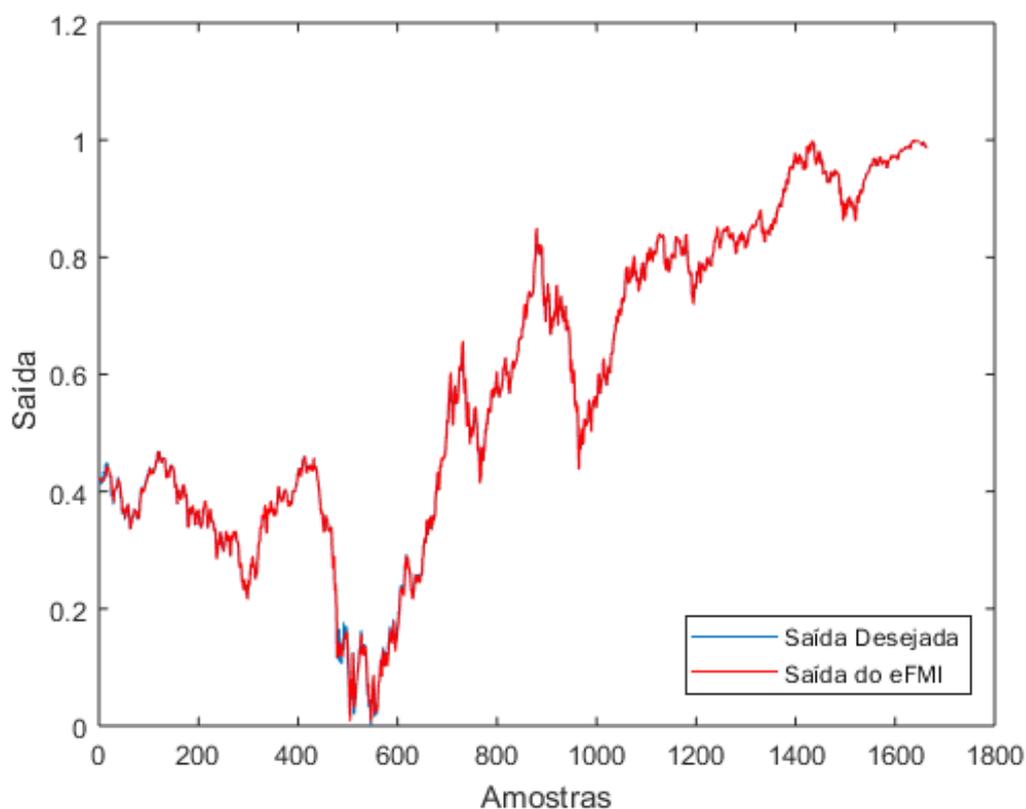


Figura 13 – Previsão do preço dos títulos do Global 40 pelo eFMI.

O logaritmo da soma dos quadrados residuais - Log (RSS) - dos modelos na previsão do preço dos títulos do Global 40 é apresentado na Figura 14. Os algoritmos eFMI, eFCE e eMG obtiveram os melhores resultados e apresentaram ao longo de todo o experimento um desempenho similar. O eNFN obteve o quarto melhor desempenho. Em seguida, os modelos eFLS e eFTLS apresentaram ao longo do experimento um desempenho comparável.

A Tabela 14 mostra os resultados comparativos na previsão dos preços dos títulos do Global 40 pelo teste MGN, dado um nível de significância de 0,05. Os resultados ilustram que o desempenho do eFCE é estatisticamente superior ao do eFLS, eFTLS, eNFN e eOGS e similar ao do eFMI e eMG. A comparação também mostra que o desempenho do eFLS é estatisticamente superior ao do eOGS e similar ao do eFTLS. Além disso, os resultados apresentados sugerem que o eFMI é estatisticamente superior ao eFLS, eFTLS, eNFN e eOGS. Ainda, percebe-se que o eFMI, eFCE e eMG são similares. Ainda se observa que o eFTLS é superior ao eOGS, similar ao eFLS e inferior aos demais.

4.7 Previsão de Aluguel de Bicicletas

Esta seção considera o problema de previsão do aluguel de bicicletas. Em sistemas de compartilhamento, os usuários podem alugar e devolver uma bicicleta em diferentes

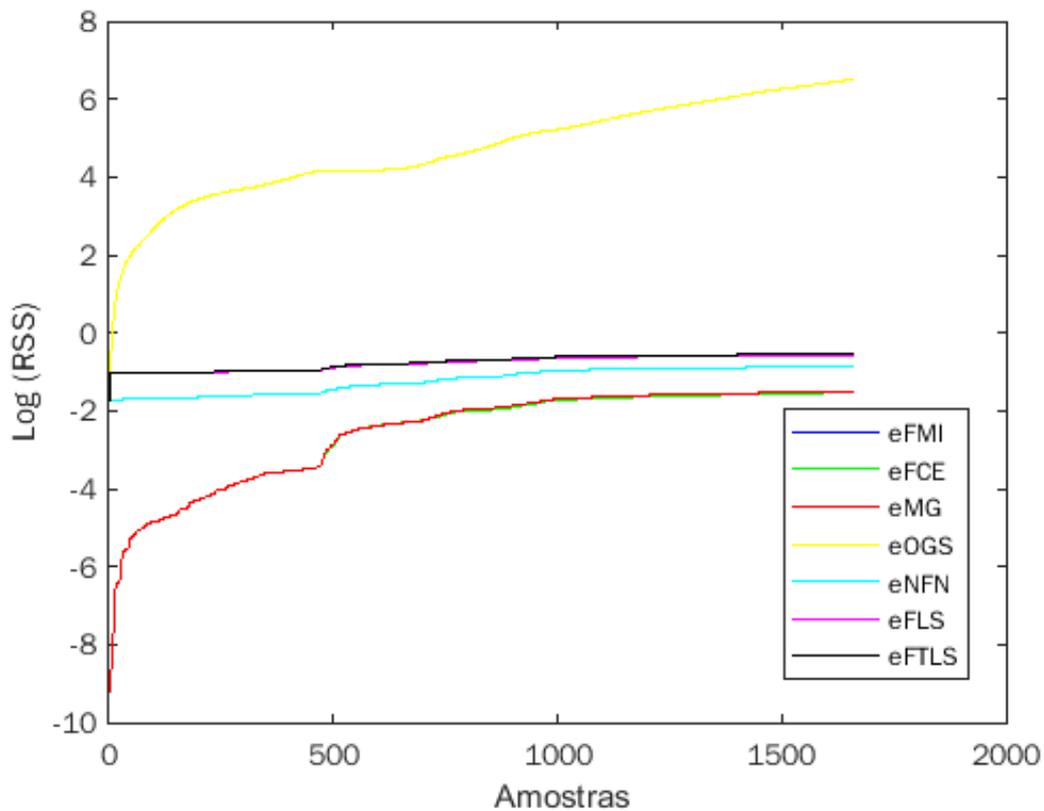


Figura 14 – Logaritmo da soma dos quadrados residuais na previsão dos preços dos títulos do Global 40.

Tabela 14 – Avaliação dos modelos propostos pelo teste MGN na previsão dos preços dos títulos do Global 40.

Modelos	eFCE MGN/ <i>p</i> -valor	eFLS MGN/ <i>p</i> -valor	eFMI MGN/ <i>p</i> -valor	eFTLS MGN/ <i>p</i> -valor	eMG MGN/ <i>p</i> -valor	eNFN MGN/ <i>p</i> -valor	eOGS MGN/ <i>p</i> -valor
eFCE	—	25.24/0.00	-0.32/0.53	26.12/0.00	0.81/0.41	18.76/0.00	481.94/0.00
eFLS	-25.24/1.00	—	-25.24/1.00	3.28/0.44	-24.38/1.00	-9.25/1.00	296.54/0.00
eFMI	0.32/0.47	25.24/0.00	—	26.12/0.00	0.81/0.40	18.76/0.00	481.94/0.00
eFTLS	-26.12/1.00	-3.28/0.56	-26.12/1.00	—	-25.25/1.00	-10.03/1.00	292.07/0.00

locais da cidade. O conjunto de dados Capital Bike Sharing ⁴ contém 731 amostras, que são informações obtidas durante 2 anos. Os dados desse sistema de empréstimo de bicicletas foram coletados em Washington D.C. e estão disponíveis no UCI Machine Learning Repository (UCI, 2020). O objetivo é prever o número de bicicletas alugadas usando 9 variáveis de entrada: estação (x_1^t); mês (x_2^t); feriado (x_3^t); dia da semana (x_4^t); situação climática (x_5^t); temperatura (x_6^t); temperatura aparente (x_7^t); umidade do ar (x_8^t); e velocidade do vento (x_9^t) (GARCIA; LEITE; ŠKRJANC, 2020). O modelo para este conjunto de dados é descrito por:

$$\hat{y}^t = f(x_1^t, x_2^t, \dots, x_8^t, x_9^t). \tag{57}$$

⁴ <https://archive.ics.uci.edu/ml/datasets/Bike+Sharing+Dataset>

Os resultados de RMSE e NDEI obtidos na previsão do compartilhamento de bicicletas são exibidos na Tabela 15. O eFMI obteve o melhor desempenho seguido pelo eNFN, eFTLS, eMG, eFCE, eOGS e eFLS. Os resultados alcançados pelo eFMI e eNFN são comparáveis e superam em pelo menos uma ordem de magnitude os dos demais modelos. A Figura 15 ilustra os valores desejados e os estimados pelo eFMI na previsão do aluguel de bicicletas.

Tabela 15 – Desempenho dos modelos na previsão do aluguel de bicicletas.

Modelos	RMSE	NDEI
eFMI	0.0953	0.4275
eNFN	0.0966	0.4334
eFTLS	0.1068	0.4794
eMG	0.1096	0.4916
eFCE	0.1284	0.5761
eOGS	0.1376	0.5936
eFLS	0.1583	0.7094

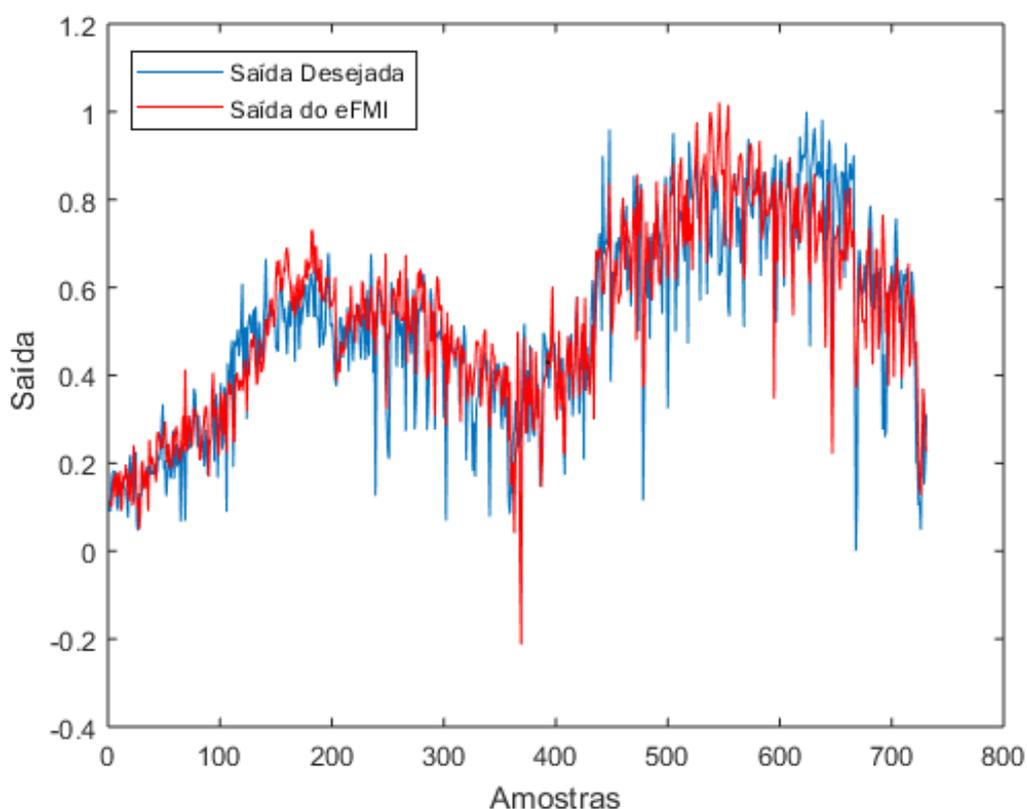


Figura 15 – Previsão do aluguel de bicicletas pelo eFMI.

A soma dos quadrados residuais - RSS - dos modelos para cada passo t é mostrado na Figura 16 para a previsão do aluguel de bicicletas. Analisando a Figura 16, pode-se observar, no final da previsão, que a curva de erro do eFMI é comparável com a do eNFN,

mas superior ao eFTLS. Ainda, os resultados visualizados na Figura 16 sugerem que, nas primeiras amostras, o eFCE mostrou uma boa performance, porém durante o restante da simulação a curva de erro piorou e ficou menor apenas em relação ao eOGS e eFLS. O eFTLS, nas primeiras amostras, apresentou a pior curva de erro, mas finalizou o experimento obtendo resultados que superaram o eFCE, eOGS e eFLS.

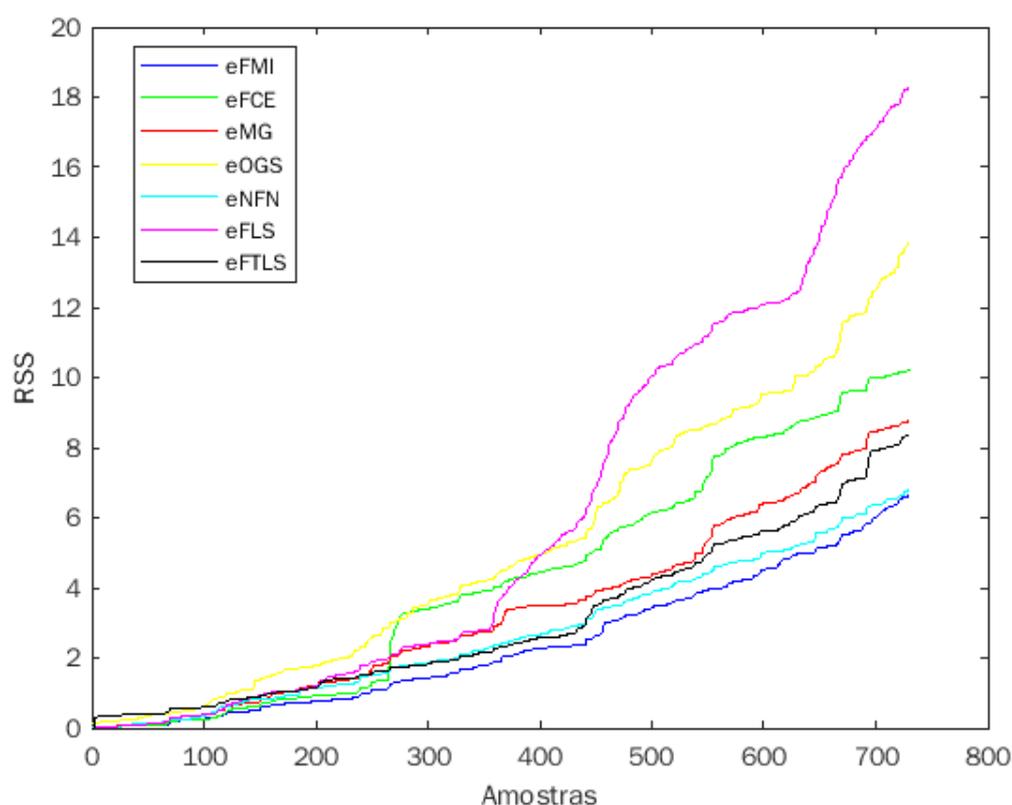


Figura 16 – Soma dos quadrados residuais na previsão do aluguel de bicicletas.

Os resultados da comparação entre os modelos propostos e alternativos pelo teste MGN na previsão do aluguel de bicicletas são descritos na Tabela 16, com nível de significância de 0,05. A comparação mostra que o eFCE possui desempenho estatisticamente superior ao eFLS e eOGS, inferior aos demais modelos. Os resultados sugerem que o eFLS é, para este experimento, estatisticamente inferior a todos os modelos avaliados. O resultado alcançado pelo eFMI mostrou-se estatisticamente superior ao eFCE, eFLS, eFTLS, eMG e eOGS, porém similar ao eNFN. Por fim, os resultados ilustram que o eFTLS é estatisticamente superior ao eFCE, eFLS e eOGS, similar ao eMG e inferior ao eFMI e eNFN.

Tabela 16 – Avaliação dos modelos propostos pelo teste MGN na previsão de aluguel de bicicletas.

Modelos	eFCE MGN/ <i>p</i> -valor	eFLS MGN/ <i>p</i> -valor	eFMI MGN/ <i>p</i> -valor	eFTLS MGN/ <i>p</i> -valor	eMG MGN/ <i>p</i> -valor	eNFN MGN/ <i>p</i> -valor	eOGS MGN/ <i>p</i> -valor
eFCE	—	8.17/0.00	-6.73/1.00	-3.58/1.00	-2.99/1.00	-5.97/1.00	4.40/0.00
eFLS	-8.17/1.00	—	-13.79/1.00	-11.90/1.00	-11.5/1.00	-13.04/1.00	-2.34/0.99
eFMI	6.73/0.00	13.79/0.00	—	3.58/0.00	4.32/0.00	0.43/0.34	11.12/0.00
eFTLS	3.58/0.00	11.90/0.00	-3.58/1.00	—	0.77/0.22	-3.11/0.99	7.74/0.00

5 Conclusão

Este trabalho introduziu uma nova abordagem para a construção de modelos *fuzzy* evolutivos alicerçada sob um algoritmo de agrupamento participativo e funções de pertinência Gaussiana multivariável. A estrutura da abordagem proposta evolui pela inclusão, exclusão, união e atualização de grupos e regras. A criação de um novo grupo/regra utiliza uma medida de similaridade baseada na distância Euclidiana para microgrupos e na distância de Mahalanobis para grupos. O uso de duas distâncias resolve o problema de calcular o inverso da matriz de dispersão de um grupo com poucas amostras. A exclusão de grupos e regras inspira-se no conceito de idade e população para excluir grupos inativos ou com pouca representatividade. O mecanismo de mesclagem é baseado na sobreposição notável de dois grupos. A partir dessa abordagem, foram propostos quatro modelos, que se diferem pela abordagem utilizada para atualização dos parâmetros do consequente, são eles: o eFLS que utiliza o algoritmo recursivo de mínimos quadrados ponderados; o eFMI que implementa o algoritmo de mínimos quadrados recursivos ponderados com multi-inovações; o eFCE que emprega o algoritmo recursivo de máxima correntropia; e o eFTLS que é baseado no algoritmo recursivo de mínimos quadrados totais ponderados.

O desempenho dos modelos propostos foi avaliado e comparado com modelos evolutivos do estado da arte usando instâncias de previsão de séries temporais e de identificação de sistemas não lineares. Os resultados computacionais indicam que os modelos propostos têm desempenho superior ou comparável ao dos modelos alternativos. Além disso, é possível observar que, o modelo eFCE obteve o melhor performance, em 5 dos 7 experimentos realizados, em comparação aos demais modelos propostos e alternativos. O modelo eFMI, comparando com os modelos alternativos e propostos, apresentou melhor desempenho em 2 dos 7 experimentos. Isto sugere que, para os experimentos realizados neste trabalho, o algoritmo recursivo de Máxima Correntropia possui melhor desempenho na atualização de parâmetros do consequente, em relação ao modelos propostos. Destaca-se que o algoritmo de atualização dos parâmetros do consequente, que apresentou o melhor performance, tem como objetivo reduzir a influência de possíveis discrepâncias ou níveis de ruídos atípicos.

5.1 Trabalhos Futuros

Com base nos experimentos realizados e nas conclusões extraídas, sugere-se como propostas de trabalhos futuros:

- abordar técnicas de seleção automática de parâmetros dos modelos propostos, tornando-os mais autônomos;

- realizar a análise de complexidade e, se necessário, investigar mecanismos para reduzir a complexidade dos algoritmos a fim de reduzir seu custo computacional;
- avaliar o uso de novos algoritmos para ajustar os parâmetros do consequente, visando ter boa acurácia;
- generalizar os modelos para problemas com múltiplas saídas e para tarefas de classificação, tornando-se aplicáveis para uma gama maior de tarefas de previsão, identificação de sistemas e classificação.

Referências

AGUIAR, C.; LEITE, D. Unsupervised Fuzzy eIX: Evolving Internal-eXternal Fuzzy Clustering. In: **Proceedings of the IEEE Conference on Evolving and Adaptive Intelligent Systems (EAIS)**. Bari, Italy: IEEE, 2020. p. 1–8. Citado 5 vezes nas páginas 9, 10, 12, 19 e 27.

ALIPERTI, A. et al. A Fuzzy Density-based Clustering Algorithm for Streaming Data. In: **Proceedings of the IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)**. New Orleans, LA, USA: IEEE, 2019. p. 1–6. Citado na página 2.

ANGELOV, P. **Evolving rule-based models: a tool for design of flexible adaptive systems**. Heidelberg, Berlin: Springer-Verlag, 2002. v. 92. Citado 3 vezes nas páginas 1, 8 e 12.

ANGELOV, P. Fuzzily Connected Multimodel Systems Evolving Autonomously From Data Streams. **IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)**, v. 41, n. 4, p. 898–910, August 2011. ISSN 1083-4419. Citado na página 8.

ANGELOV, P.; BUSWELL, R. Evolving rule-based models: A tool for intelligent adaptation. In: **Proceedings of the 9th Joint IFSA World Congress and 20th NAFIPS International Conference**. Vancouver, BC, Canada: IEEE, 2001. v. 2, p. 1062–1067. ISBN 0-7803-7078-3. Citado na página 1.

ANGELOV, P.; FILEV, D. An approach to online identification of Takagi-Sugeno fuzzy models. **IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)**, v. 34, n. 1, p. 484–498, February 2004. ISSN 1083-4419. Citado na página 1.

ANGELOV, P.; FILEV, D. Simpl_eTS: a simplified method for learning evolving Takagi-Sugeno fuzzy models. In: **Proceedings of the 14th IEEE International Conference on Fuzzy Systems, FUZZ-IEEE '05**. Reno, NV, USA: IEEE, 2005. p. 1068–1073. Citado 5 vezes nas páginas 3, 8, 12, 20 e 21.

ANGELOV, P.; FILEV, D.; KASABOV, N. Guest Editorial Evolving Fuzzy Systems-Preface to the Special Section. **IEEE Transactions on Fuzzy Systems**, v. 16, n. 6, p. 1390–1392, December 2008. Citado na página 2.

ANGELOV, P.; FILEV, D.; KASABOV, N. **Evolving Intelligent Systems: Methods and Applications**. 1. ed. Hoboken, NJ, USA: Wiley & IEEE Press, 2010. Citado na página 7.

ANGELOV, P.; FILEV, D. P.; KASABOV, N. Evolving Takagi-Sugeno Fuzzy Systems from Streaming Data (eTS+). In: **Evolving Intelligent Systems: Methodology and Applications**. New York, USA: IEEE, 2010. p. 21–50. Citado 3 vezes nas páginas 1, 8 e 12.

ANGELOV, P.; GU, X.; PRÍNCIPE, J. Autonomous Learning Multimodel Systems From Data Streams. **IEEE Transactions on Fuzzy Systems**, v. 26, n. 4, p. 2213–2224, November 2018. Citado 2 vezes nas páginas 10 e 12.

ANGELOV, P.; YAGER, R. Simplified fuzzy rule-based systems using non-parametric antecedents and relative data density. In: **Proceedings of the Workshop on Evolving and Adaptive Intelligent Systems (EAIS)**. Paris, France: IEEE, 2011. p. 62–69. Citado na página 11.

- ANGELOV, P.; ZHOU, X. Evolving Fuzzy Systems from Data Streams in Real-Time. In: **Proceedings of the International Symposium on Evolving Fuzzy Systems**. Ambelside, UK: IEEE, 2006. p. 29–35. Citado 2 vezes nas páginas 8 e 12.
- ANGELOV, P.; ZHOU, X. On line learning fuzzy rule-based system structure from data streams. In: **Proceedings of the IEEE International Conference on Fuzzy Systems (IEEE World Congress on Computational Intelligence)**. Hong Kong, China: IEEE, 2008. p. 915–922. ISBN 978-1-4244-1818-3. Citado na página 7.
- ARRIETA, A. B. et al. Explainable Artificial Intelligence (XAI): Concepts, taxonomies, opportunities and challenges toward responsible AI. **Information Fusion**, v. 58, p. 82–115, June 2020. Citado na página 1.
- BAO, R. et al. Correntropy-based evolving fuzzy neural system. **IEEE Transactions on Fuzzy Systems**, v. 26, n. 3, p. 1324–1338, 2018. Citado na página 24.
- BEZERRA, C. et al. An Evolving Approach to Data Streams Clustering Based on Typicality and Eccentricity Data Analytics. **Information Sciences**, v. 518, p. 13–28, May 2020. Citado 2 vezes nas páginas 10 e 12.
- BLAŽIČ, S.; DOVŽAN, D.; ŠKRJANC, I. Cloud-based identification of an evolving system with supervisory mechanisms. In: **Proceedings of the International Symposium on Intelligent Control (ISIC)**. Juan Les Pins, France: IEEE, 2014. p. 1906–1911. Citado na página 11.
- BLAŽIČ, S.; ŠKRJANC, I. Incremental Fuzzy C-Regression Clustering From Streaming Data for Local-Model-Network Identification. **IEEE Transactions on Fuzzy Systems**, v. 28, n. 4, p. 758–767, 2020. ISSN 1941-0034. Citado na página 2.
- BOX, G.; JENKINS, G. **Time Series Analysis: Forecasting and Control**. 2nd edição. San Francisco, US: Holden-Day, 1976. Citado na página 36.
- DAVILA, C. An efficient recursive total least squares algorithm for fir adaptive filtering. **IEEE Transactions on Signal Processing**, v. 42, n. 2, p. 268–280, 1994. Citado na página 25.
- DIEBOLD, F.; MARIANO, R. Comparing Predictive Accuracy. **Journal of Business & Economic Statistics**, v. 20, n. 1, p. 134–144, 2002. Citado na página 31.
- DING, F.; CHEN, H.; LI, M. Multi-innovation least squares identification methods based on the auxiliary model for miso systems. **Applied Mathematics and Computation**, Elsevier Science Inc., USA, v. 187, n. 2, p. 658–668, apr 2007. Citado na página 23.
- DING, F.; LIU, P. X.; LIU, G. Multiinnovation Least-Squares Identification for System Modeling. **IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)**, v. 40, n. 3, p. 767–778, October 2010. Citado na página 24.
- FERDAUS, M. M. et al. Palm: An incremental construction of hyperplanes for data stream regression. **IEEE Transactions on Fuzzy Systems**, v. 27, n. 11, p. 2115–2129, 2019. Citado 3 vezes nas páginas 1, 2 e 3.
- FERNANDEZ, A. et al. Evolutionary Fuzzy Systems for Explainable Artificial Intelligence: Why, When, What for, and Where to? **IEEE Computational Intelligence Magazine**, v. 14, n. 1, p. 69–81, February 2019. Citado na página 1.

FULCHER, J. Computational intelligence: a compendium. In: **Computational Intelligence: An Introduction**. Berlin: Springer, 2008. v. 115, p. 3–78. ISBN 978-3-540-78292-6. Citado na página 6.

GARCIA, C. et al. Evolvable fuzzy systems from data streams with missing values: With application to temporal pattern recognition and cryptocurrency prediction. **Pattern Recognition Letters**, v. 128, p. 278–282, December 2019. Citado 4 vezes nas páginas 2, 8, 9 e 12.

GARCIA, C.; LEITE, D.; ŠKRJANC, I. Incremental missing-data imputation for evolving fuzzy granular prediction. **IEEE Transactions on Fuzzy Systems**, v. 28, n. 10, p. 2348–2362, 2020. Citado 2 vezes nas páginas 42 e 46.

GOLDBERG, D. E. **Genetic Algorithms in Search, Optimization, and Machine Learning**. New York: Addison-Wesley, 1989. Citado na página 7.

IGLESIAS, J. et al. Modelling evolving user behaviours. In: **Proceedings of the IEEE Workshop on Evolving and Self-Developing Intelligent Systems**. Nashville, TN, USA: IEEE, 2009. p. 16–23. ISBN 978-1-4244-2754-3. Citado na página 2.

KASABOV, N. Evolving fuzzy neural networks for supervised/unsupervised online knowledge-based learning. **IEEE Transactions on Systems, Man and Cybernetics, Part B (Cybernetics)**, v. 31, n. 6, p. 902–918, 2001. ISSN 10834419. Citado na página 1.

KASABOV, N.; SONG, Q. Dynamic evolving fuzzy neural networks with "m-out-of-n" activation nodes for on-line adaptive systems. **Technical report TR99/04, Department of information science, University of Otago**, Department of Information Science - University of Otago, Dunedin, New Zealand, p. 1–29, March 1999. Citado na página 1.

KASABOV, N.; SONG, Q. DENFIS: Dynamic Evolving Neural-Fuzzy Inference System and Its Application for Time-Series Prediction. **IEEE Transactions on Fuzzy Systems**, v. 10, n. 2, p. 144–154, 2002. Citado na página 1.

KAYA, H.; TUFEKCI, P. Local and global learning methods for predicting power of a combined gas & steam turbine. In: **Proceedings of the International Conference on Emerging Trends in Computer and Electronics Engineering - ICETCEE'12**. Dubai: IEEE, 2012. p. 13–18. Citado na página 39.

KOZA, J. R. **Genetic Programming: On the Programming of Computers by Means of Natural Selection**. Cambridge, MA, USA: MIT Press, 1992. Citado na página 7.

LEITE, D.; COSTA, P.; GOMIDE, F. Interval Approach for Evolving Granular System Modeling. In: SAYED-MOUCHAWEH, M.; LUGHOFFER, E. (Ed.). **Learning in Non-Stationary Environments**. New York, NY: Springer, 2012. p. 271–300. Citado na página 42.

LEITE, D. et al. Fuzzy granular evolving modeling for time series prediction. In: **2011 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE 2011)**. Taipei, Taiwan: IEEE, 2011. p. 2794–2801. Citado na página 43.

LEITE, D.; GOMIDE, F.; ŠKRJANC, I. Multiobjective Optimization of Fully Autonomous Evolving Fuzzy Granular Models. In: **Proceedings of the IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)**. New Orleans, LA, USA: IEEE, 2019. p. 1–7. Citado na página 2.

LEITE, D.; ŠKRJANC, I. Ensemble of Evolving Optimal Granular Experts, OWA Aggregation, and Time Series Prediction. **Information Sciences**, v. 504, p. 95–112, December 2019. Citado 3 vezes nas páginas [9](#), [12](#) e [29](#).

LEITE, D.; ŠKRJANC, I.; GOMIDE, F. An overview on evolving systems and learning from stream data. **Evolving Systems**, v. 11, p. 181–198, March 2020. Citado na página [2](#).

LEMOS, A.; CAMINHAS, W.; GOMIDE, F. Fuzzy Multivariable Gaussian Evolving Approach for Fault Detection and Diagnosis. In: **Computational Intelligence for Knowledge-Based Systems Design. IPMU 2010**. Heidelberg, Berlin: Springer, 2010. v. 6178, p. 360–369. Citado na página [7](#).

LEMOS, A.; CAMINHAS, W.; GOMIDE, F. Multivariable Gaussian Evolving Fuzzy Modeling System. **IEEE Transactions on Fuzzy Systems**, v. 19, n. 1, p. 91–104, October 2011. ISSN 1063-6706. Citado 7 vezes nas páginas [3](#), [9](#), [12](#), [16](#), [17](#), [27](#) e [29](#).

LUGHOFER, E. Extensions of vector quantization for incremental clustering. **Pattern Recognition**, v. 41, n. 3, p. 995–1011, March 2008. Citado na página [8](#).

LUGHOFER, E. FLEXFIS: A Robust Incremental Learning Approach for Evolving Takagi–Sugeno Fuzzy Models. **IEEE Transactions on Fuzzy Systems**, v. 16, n. 6, p. 1393–1410, 2008. Citado 2 vezes nas páginas [8](#) e [12](#).

LUGHOFER, E. **Evolving Fuzzy Systems: Methodologies, Advanced Concepts and Applications**. 1. ed. New York: Springer-Verlag, 2011. 45-87 p. Citado na página [1](#).

LUGHOFER, E. Flexible evolving fuzzy inference systems from data streams (FLEXFIS++). In: Sayed-Mouchaweh, M.; Lughofer, E. (Ed.). **Learning in Non-Stationary Environments**. New York, NY: Springer, 2012. p. 205–245. Citado 2 vezes nas páginas [9](#) e [12](#).

LUGHOFER, E. On-line assurance of interpretability criteria in evolving fuzzy systems - Achievements, new concepts and open issues. **Information Sciences**, v. 251, p. 22–46, 2013. ISSN 0020-0255. Citado na página [1](#).

LUGHOFER, E. Evolving fuzzy systems: Fundamentals, reliability, interpretability, useability and applications. In: **Proceedings of the 7th International Joint Conference on Computational Intelligence (IJCCI)**. Lisbon, Portugal: IEEE, 2015. v. 1, p. 11–11. Citado na página [22](#).

LUGHOFER, E. Variants of Recursive Consequent Parameters Learning in Evolving Neuro-Fuzzy Systems. In: **Proceedings of the IEEE Conference on Evolving and Adaptive Intelligent Systems (EAIS)**. Bari, Italy: IEEE, 2020. p. 1–10. Citado 4 vezes nas páginas [22](#), [23](#), [24](#) e [25](#).

LUGHOFER, E.; ANGELOV, P. Handling drifts and shifts in on-line data streams with evolving fuzzy systems. **Applied Soft Computing**, v. 11, n. 2, p. 2057–2068, March 2011. Citado 2 vezes nas páginas [3](#) e [20](#).

LUGHOFER, E.; GUARDIOLA, C. Applying evolving fuzzy models with adaptive local error bars to on-line fault detection. In: **Proceedings of the 3rd International Workshop on Genetic and Evolving Systems**. Witten-Bommerholz, Germany: IEEE, 2008. p. 35–40. Citado na página [2](#).

LÜTKEPOHL, H. **New Introduction to Multiple Time Series Analysis**. Berlin: Springer Science Business Media, 2005. Citado na página [2](#).

MACIEL, L.; GOMIDE, F.; BALLINI, R. An enhanced approach for evolving participatory learning fuzzy modeling. In: **Proceedings of the IEEE Conference on Evolving and Adaptive Intelligent Systems**. Madrid, Spain: IEEE, 2012. p. 23–28. Citado na página [1](#).

MACIEL, L. et al. Evolving fuzzy systems for pricing fixed income options. **Evolving Systems**, v. 3, p. 5–18, 2012. ISSN 1868-6478. Citado na página [1](#).

MAMDANI, E.; ASSILIAN, S. An experiment in linguistic synthesis with a fuzzy logic controller. **International Journal of Man-Machine Studies**, v. 7, n. 1, p. 1–13, 1975. ISSN 0020-7373. Citado na página [6](#).

PRATAMA, M.; ANAVATTI, S. G.; LUGHOFER, E. Evolving fuzzy rule-based classifier based on GENEFIS. In: **Proceedings of the IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)**. Hyderabad, India: IEEE, 2013. p. 1–8. Citado na página [7](#).

RODRIGUES, F. P. S.; SILVA, A. M.; LEMOS, A. P. Evolving fuzzy system with multivariable gaussian participatory learning and recursive maximum correntropy - efce. In: **Proceedings of the IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)**. Luxembourg, Luxembourg: IEEE, 2021. v. 1, n. 1, p. 1–6. Citado na página [12](#).

RODRIGUES, F. P. S.; SILVA, A. M.; LEMOS, A. P. Evolving fuzzy predictor with multivariable Gaussian participatory learning and multi-innovations recursive weighted least squares: eFMI. **Evolving Systems**, n. 1, p. 1–36, 2022. Citado na página [12](#).

RONG, H.-J.; YANG, Z.-X.; WONG, P. K. Robust and Noise-Insensitive Recursive Maximum Correntropy-Based Evolving Fuzzy System. **IEEE Transactions on Fuzzy Systems**, v. 28, n. 9, p. 2277–2284, 2020. ISSN 1941-0034. Citado na página [2](#).

SILVA, A. et al. A fast learning algorithm for evolving neo-fuzzy neuron. **Applied Soft Computing**, v. 14, p. 194–209, 2014. ISSN 1568-4946. Citado 5 vezes nas páginas [8](#), [10](#), [12](#), [29](#) e [31](#).

SILVA, L.; GOMIDE, F.; YAGER, R. Participatory learning in fuzzy clustering. In: **Proceedings of the 14th IEEE International Conference on Fuzzy Systems**. Reno, NV, USA: IEEE, 2005. p. 857–861. Citado na página [9](#).

SUBHRAJIT, S.; PRATAMA, M.; SUNDARAM, S. A novel spatio-temporal fuzzy inference system (spatfis) and its stability analysis. **Information Sciences**, v. 505, p. 84–99, 2019. Citado 2 vezes nas páginas [2](#) e [3](#).

TAKAGI, T.; SUGENO, M. Fuzzy identification of systems and its applications to modeling and control. **IEEE Transactions on Systems, Man, and Cybernetics**, SMC-15, n. 1, p. 116–132, 1985. Citado na página [6](#).

TUFEKCI, P. Prediction of full load electrical power output of a base load operated combined cycle power plant using machine learning methods. **International Journal of Electrical Power & Energy Systems**, v. 60, p. 126–140, 09 2014. Citado na página [39](#).

TUNG, S.; QUEK, C.; GUAN, C. eT2FIS: An Evolving Type-2 Neural Fuzzy Inference System. **Information Sciences**, v. 220, p. 124–148, 2013. ISSN 0020-0255. Citado na página [1](#).

UCI. **UC Irvine Machine Learning Repository**. 2020. <https://archive.ics.uci.edu/ml/index.php>. Citado na página 46.

WANG, D.; ZENG, X.; KEANE, J. A structure evolving learning method for fuzzy systems. **Evolving Systems**, v. 1, p. 83–95, 2010. ISSN 1868-6478. Citado na página 8.

XYDEAS, C. et al. Advances in classification of EEG signals via evolving fuzzy classifiers and dependant multiple HMMs . **Computers in Biology and Medicine**, v. 36, n. 10, p. 1064–1083, October 2006. ISSN 0010-4825. Citado na página 2.

YAGER, R. R. A model of participatory learning. **IEEE Transactions on Systems, Man, and Cybernetics**, v. 20, n. 5, p. 1229–1234, Sep/Oct 1990. Citado 2 vezes nas páginas 9 e 16.

ZADEH, L. Fuzzy sets. **Information and Control**, Elsevier BV, v. 8, n. 3, p. 338–353, jun 1965. Citado na página 6.

ZHOU, X.; ANGELOV, P. Autonomous visual self-localization in completely unknown environment using evolving fuzzy rule-based classifier. In: **Proceedings of the IEEE Symposium on Computational Intelligence in Security and Defense Applications**. Honolulu, HI, USA: IEEE, 2007. p. 131–138. Citado na página 2.

ŠKRJANC, I. Cluster-volume based merging concept for incrementally evolving fuzzy Gaussian clustering - eGAUSS+. **IEEE Transactions on Fuzzy Systems**, v. 28, n. 9, p. 2222–2231, July 2020. Citado 8 vezes nas páginas 2, 3, 7, 9, 12, 17, 27 e 32.

ŠKRJANC, I. et al. Inner Matrix Norms in Evolving Cauchy Possibilistic Clustering for Classification and Regression from Data Streams. **Information Sciences**, v. 478, p. 540–563, April 2019. Citado 3 vezes nas páginas 2, 11 e 12.

ŠKRJANC, I. et al. Evolving fuzzy and neuro-fuzzy approaches in clustering, regression, identification, and classification: A survey. **Information Sciences**, v. 490, p. 344–368, July 2019. Citado 6 vezes nas páginas 1, 2, 3, 7, 8 e 29.

ŠKRJANC, I. et al. Large-Scale Cyber Attacks Monitoring using Evolving Cauchy Possibilistic Clustering. **Applied Soft Computing**, v. 62, p. 592–601, January 2018. Citado na página 11.