
Novo Método de Identificação Fuzzy para Sistemas MISO e suas Aplicações

Jefferson Beethoven Martins



UNIVERSIDADE FEDERAL DE UBERLÂNDIA
FACULDADE DE ENGENHARIA ELÉTRICA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA

Uberlândia
2020

Jefferson Beethoven Martins

**Novo Método de Identificação Fuzzy para
Sistemas MISO e suas Aplicações**

Tese de doutorado apresentada ao Programa de Pós-graduação da Faculdade de Engenharia Elétrica da Universidade Federal de Uberlândia como parte dos requisitos para a obtenção do título de Doutor em Engenharia Elétrica.

Área de concentração: Inteligência Artificial

Orientador: Keiji Yamanaka

Coorientador: Ana Maria Amarillo Bertone

Uberlândia

2020

Dados Internacionais de Catalogação na Publicação (CIP)
Sistema de Bibliotecas da UFU, MG, Brasil.

M386n	Martins, Jefferson Beethoven, 1985-
2020	Novo método de identificação Fuzzy para sistemas MISO e suas aplicações [recurso eletrônico] / Jefferson Beethoven Martins. - 2020.
	Orientador: Keiji Yamanaka.
	Coorientadora: Ana Maria Amarillo Bertone.
	Tese (Doutorado) - Universidade Federal de Uberlândia, Programa de Pós-Graduação em Engenharia Elétrica.
	Disponível em: http://doi.org/10.14393/ufu.te.2020.3007
	Inclui bibliografia.
	Inclui ilustrações.
	1. Engenharia elétrica. I. Yamanaka, Keiji, (Orient.). II. Bertone, Ana Maria Amarillo, (Coorient.). III. Universidade Federal de Uberlândia. Programa de Pós-Graduação em Engenharia Elétrica. IV. Título.

CDU: 621.3

Rejâne Maria da Silva – CRB6/1925



ATA DE DEFESA - PÓS-GRADUAÇÃO

Programa de Pós-Graduação em:	Engenharia Elétrica				
Defesa de:	Tese de Doutorado, 271, PPGEELT				
Data:	Dezessete de julho de dois mil e vinte	Hora de início:	09:00	Hora de encerramento:	12:30
Matrícula do Discente:	11423EEL009				
Nome do Discente:	Jefferson Beethoven Martins				
Título do Trabalho:	Novo método de identificação fuzzy para sistemas MISO e suas aplicações.				
Área de concentração:	Processamento da informação				
Linha de pesquisa:	Inteligência artificial				
Projeto de Pesquisa de vinculação:					

Reuniu-se por meio de Videoconferência, a Banca Examinadora, designada pelo Colegiado do Programa de Pós-graduação em Engenharia Elétrica, assim composta: Professores Doutores: Ana Maria Amarillo Bertone - FAMAT/UFU; Rosana da Motta Jafelice - FAMAT/UFU; Igor Santos Peretta - FEELT/UFU; Wesley Pacheco Calixto - IFG; Edilberto Pereira Teixeira - UNIUBE; Keiji Yamanaka - FEELT/UFU; orientador(a) do(a) candidato(a).

Iniciando os trabalhos o(a) presidente da mesa, Dr(a). Keiji Yamanaka, apresentou a Comissão Examinadora e o candidato(a), agradeceu a presença do público, e concedeu ao Discente a palavra para a exposição do seu trabalho. A duração da apresentação do Discente e o tempo de arguição e resposta foram conforme as normas do Programa.

A seguir o senhor(a) presidente concedeu a palavra, pela ordem sucessivamente, aos(as) examinadores(as), que passaram a arguir o(a) candidato(a). Ultimada a arguição, que se desenvolveu dentro dos termos regimentais, a Banca, em sessão secreta, atribuiu o resultado final, considerando o(a) candidato(a):

Aprovado(a).

Esta defesa faz parte dos requisitos necessários à obtenção do título de Doutor.

O competente diploma será expedido após cumprimento dos demais requisitos, conforme as normas do Programa, a legislação pertinente e a regulamentação interna da UFU.

Nada mais havendo a tratar foram encerrados os trabalhos. Foi lavrada a presente ata que após lida e achada conforme foi assinada pela Banca Examinadora.



Documento assinado eletronicamente por **Wesley Pacheco Calixto, Usuário Externo**, em 17/07/2020, às 12:54, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por **Keiji Yamanaka, Professor(a) do Magistério Superior**, em 17/07/2020, às 13:38, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por **Rosana Sueli da Motta Jafelice, Professor(a) do Magistério Superior**, em 17/07/2020, às 20:29, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por **Igor Santos Peretta, Professor(a) do Magistério Superior**, em 18/07/2020, às 06:14, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por **Ana Maria Amarillo Bertone, Professor(a) do Magistério Superior**, em 20/07/2020, às 12:50, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por **Edilberto Pereira Teixeira, Usuário Externo**, em 22/07/2020, às 10:57, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



A autenticidade deste documento pode ser conferida no site https://www.sei.ufu.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador **2098933** e o código CRC **0CC195E7**.

Agradecimentos

Agradeço primeiramente a Deus pelo dom da vida. Agradeço ao meu orientador, Keiji Yamanaka, e coorientadora, Ana Maria Amarillo Bertone, pelas grandes instruções ministradas. Eles me ensinaram o que é pesquisar, o que é ciência e como a vida humana pode ser transformada através das contribuições da academia. Agradeço a minha amada esposa Mara Dalila e minha pequena filha Helena pela compreensão e apoio incondicionais. Agradeço aos meus pais e irmãs pelo amparo em todos os momentos.

Resumo

Técnicas de identificação de sistemas são essenciais para o conhecimento de fenômenos naturais e processos de diferentes origens. O principal intuito da presente pesquisa tem sido construir, teoricamente e computacionalmente, uma abordagem para lidar com sistemas de dados com várias entradas e uma saída. Duas etapas fuzzy distintas e bem definidas estão presentes no processo de identificação utilizado neste estudo: clusterização dos dados e uma inferência do tipo Takagi-Sugeno-Kang. A principal contribuição desta pesquisa é a construção dos antecedentes e dos consequentes do sistema de inferência nos quais são utilizados métodos inéditos. Uma forma de validar esta nova metodologia tem sido através de simulações com diferentes bases de dados. Os experimentos envolveram comparações com outras técnicas consolidadas como: Nonlinear Auto-Regressive with eXogenous inputs, Hammerstein-Wiener e Multilayer-Perceptron (redes neurais artificiais). A codificação tem sido feita sem a utilização de toolboxes, exceto para os testes que têm sido feitos com a rotina Compare do software Matlab, a qual utiliza medida de acurácia denominada raiz média quadrática do erro normalizada (Normalized Root Mean Square em inglês). Ao final, e como resultado das comparações, tem havido ganho de precisão na resposta e queda no esforço computacional. Como aplicação da nova técnica, têm sido realizados testes de predição de dados, com resultados promissores, além de estruturar uma nova metodologia para a interpretação de sinais provenientes de exames médicos como, por exemplo, eletroencefalogramas. Futuros trabalhos incluem a introdução da teoria dos conjuntos fuzzy do tipo 2, extensão da teoria dos conjuntos fuzzy, na dinâmica da inferência, com o objetivo de tornar ainda melhor, o desempenho da metodologia proposta.

Palavras-chave: Identificação Fuzzy de sistemas MISO, Clusterização Fuzzy, Predição.

Abstract

System identification techniques are essential for the knowledge of natural phenomena and processes from different sources. The main purpose of this research has been to build, theoretically and computationally, an approach to deal with data systems with multiple inputs and an output. Two distinct and well-defined steps are present in the identification process used in this study: data clustering and a Takagi-Sugeno-Kang type inference. The main contribution of this research is the construction of the antecedents and the consequences of the inference system in which new methods are used. One way to validate this new methodology has been through simulations with different databases. The experiments involved comparisons with other consolidated techniques such as: Nonlinear Auto-Regressive with eXogenous inputs, Hammerstein-Wiener and Multilayer-Perceptron (artificial neural networks). The coding has been done without the use of toolboxes, except for the tests, it has been done with the Compare routine of the Matlab software, which uses an accuracy measure called Normalized Root Mean Square (English). As a result of the comparisons, there has been a gain in precision in the response and a decrease in the computational effort. As an application of the new technique, data prediction tests have been carried out, with promising results, in addition to structuring a new methodology for the interpretation of signals coming from medical examinations, such as electroencephalograms. Future works include the introduction of the type-2 fuzzy set theory, an extension of the fuzzy set theory, to the inference dynamic with the objective of making the performance of the proposed methodology even better.

Keywords: Fuzzy identification of MISO systems, Fuzzy clustering, Prediction.

Lista de ilustrações

Figura 1 – Elementos presentes em um sistema de inferência fuzzy.	28
Figura 2 – Representação de um clustering crisp referente aos conjuntos Q e W.	32
Figura 3 – Representação de um clustering fuzzy referente aos conjuntos Q e W.	32
Figura 4 – Conjunto fuzzy do tipo 2 intervalar.	38
Figura 5 – Exemplo de conjunto gaussiano fuzzy do tipo 2 intervalar.	39
Figura 6 – Sistema baseado em regras fuzzy do tipo 2.	40
Figura 7 – Projeção do α -nível da matriz de pertinências sobre o eixo i	42
Figura 8 – Fluxograma do processo de identificação fuzzy de sistemas proposto.	45
Figura 9 – Estrutura geral de um modelo Hammerstein-Wiener.	50
Figura 10 – Diagrama geral de uma Rede Neural Artificial.	51
Figura 11 – Comparação do desempenho da identificação dos dados <i>Thermic</i>	55
Figura 12 – Comparação do desempenho da identificação dos dados <i>Exchanger</i>	56
Figura 13 – Comparação do desempenho da identificação dos dados <i>Glassfurnace</i>	57
Figura 14 – Comparação do desempenho da identificação dos dados <i>Hidiga</i>	59
Figura 15 – Gráfico de trinta meses de predição da base de dados de energy.	63

Lista de tabelas

Tabela 1 – Pertinência dos dados da Figura 2 referentes aos conjuntos Q e W.	32
Tabela 2 – Pertinência dos dados da Figura 3 referentes aos conjuntos Q e W.	32
Tabela 3 – Tabela comparativa do desempenho dos dados <i>Thermic</i>	53
Tabela 4 – Tabela comparativa do desempenho dos dados <i>Exchanger</i>	54
Tabela 5 – Tabela comparativa do desempenho dos dados <i>Glassfurnace</i>	57
Tabela 6 – Tabela comparativa do desempenho dos dados <i>Hidiga</i>	58

Sumário

1	INTRODUÇÃO	19
1.1	Organização do Manuscrito da Tese de Doutorado	20
2	REVISÃO BIBLIOGRÁFICA	21
2.1	Alicerce Bibliográfico e Estado da Arte	21
3	FUNDAMENTOS TEÓRICOS	25
3.1	Teoria dos Conjuntos Fuzzy	26
3.2	Modelo de Regressão Multivariada	29
3.3	Clusterização Fuzzy de Dados	31
3.4	Ferramentas para Trabalhos Futuros	38
4	O SISTEMA PROPOSTO	41
4.1	A Inferência TSK Proposta	41
4.2	A Busca de um Código Próprio	45
4.3	Família de Modelos Apropriados para a Abordagem Proposta .	46
5	MATERIAIS E MÉTODOS	49
5.1	Estudo Comparativo de Validação	49
6	RESULTADOS OBTIDOS	53
6.1	Predição Fuzzy de Sistemas MISO	61
7	CONCLUSÃO	65
7.1	Contribuições	66
7.2	Trabalhos futuros	67
	REFERÊNCIAS	69

APÊNDICES 75

APÊNDICE A – TRABALHOS PRÓPRIOS RELACIONADOS .	77
A.1 CNMAC 2016	78
A.2 Congresso Brasileiro de Sistemas Fuzzy 2016	85
A.3 Trends in Applied and Computational Mathematics 2017	97
A.4 Encontro de Biomatemática 2018	112
A.5 Apresentação na Universidade of Colorado at Denver - USA . .	117
APÊNDICE B – PRINCIPAL RESULTADO PUBLICADO: IEEE LATIN AMERICA - 2019	119

CAPÍTULO 1

Introdução

As técnicas de identificação fuzzy de sistemas são vastamente utilizadas e possuem reconhecida eficácia. O objetivo inicial desta pesquisa foi construir, teoricamente e computacionalmente, um método de identificação de sistemas do tipo múltipla-entrada saída-única (multiple-input single-output em inglês, MISO, sua abreviatura). Além disso, foram acrescentadas aplicações em várias famílias de problemas de engenharia.

Da citada proposta inicial surgiu o novo método, o qual é a junção de um procedimento de clusterização fuzzy e um procedimento de inferência fuzzy do tipo Takagi-Sugeno-Kang (TAKAGI; SUGENO, 1985; SUGENO; KANG, 1988), normalmente chamado de TSK. Após a estruturação teórica foram feitas várias simulações computacionais comparativas com outras técnicas clássicas de identificação de sistemas que comprovam, de fato, duas vantagens da nova abordagem: ganho de precisão e maior velocidade de resposta, em outras palavras, menor esforço computacional.

As metodologias escolhidas para comparação foram: NARX, (Nonlinear Auto-Regressive with eXogenous inputs) (VILLARREAL-LÓPEZ, 2019), Hammerstein-Wiener (WILLS et al., 2013) e o método Multi-Layer-Perceptron, algoritmo Backpropagation de Redes Neurais Artificiais (CASTRO; BRAGA, 2013).

Em todas as simulações, tomou-se uma parte da base de dados para treinamento e o restante para validação. Cada base de dados teve uma percentagem de validação que é especificada nas tabelas de comparação do desempenho das diferentes metodologias. Além disso, utilizou-se a medida “FIT”, que é empregada pela rotina Compare do Matlab e se refere à medida Normalized Root Mean Square (NRMSE) ou raiz média quadrática do erro normalizada, detalhada no Capítulo 5 e também explicada no Apêndice B.

A novidade obtida nesta pesquisa é a redução da informação do sistema MISO para a entrada temporal, detalhe substancial para alcançar menor esforço computacional do processo e maior precisão da resposta. Esta inovação é explanadameticulosamente no Capítulo 4.

Foram realizadas diversas simulações computacionais através de código próprio com o objetivo de validar a metodologia. Os dados empregados são de variadas origens:

benchmarks disponíveis em portais de grandes universidades ao redor do mundo (LEUVEN, 2017), dados de simulações próprias e dados coletados para as simulações, cedidos de forma exclusiva por empresas como a Usina Hidrelétrica de Igarapava-SP; além de outro grupo levantado junto a uma companhia privada local. Algumas simulações incluíram validações com o método “cross validation” (SANTOS et al., 2009), conhecido como k-fold, o que permite um melhor treinamento dos dados a serem identificados.

As aplicações da nova metodologia são feitas a uma família ampla de problemas envolvendo sistemas MISO, entre cujas variáveis de entrada há a temporal acompanhada de um número finito de atributos, como por exemplo, temperatura, umidade, entre outros. Esta estrutura, de clusterização e inferência fuzzy, pode ser aplicada a problemas específicos que envolvam identificação da dinâmica do modelo como controle, predição, entre outras possibilidades. Nesse sentido, este estudo concentrou esforços na área de predição na qual foram feitas algumas simulações com modelos reais.

Além disso, foi feito um estudo para séries temporais, entrada e saída única (single-input single-output, SISO, sua abreviatura), abrangendo o estudo analítico dos dados produzidos por um electroencefalograma (EEG).

Vários trabalhos foram compilados durante o desenvolvimento desta pesquisa, os quais serão enumerados no Apêndice A, assim como aquele que justifica esta defesa, apresentado em sua totalidade no Apêndice B.

1.1 Organização do Manuscrito da Tese de Doutorado

O texto está organizado da seguinte maneira: na Introdução são encontrados o contexto, que justifica a pesquisa, e os objetivos. No Capítulo 2, Revisão Bibliográfica, encontra-se o alicerce bibliográfico e o estado da arte da área escolhida. No Capítulo 3, Fundamentos Teóricos, são apresentados os fundamentos teóricos matemáticos de conjuntos fuzzy, regressão multivariada e clusterização fuzzy. O Capítulo 4, O Sistema Proposto, apresenta o método de inferência fuzzy e outros aspectos da metodologia em questão. No Capítulo 5, Materiais e Métodos, estão detalhados os experimentos e os detalhes substanciais para validar o novo método. O Capítulo 6, Resultados Obtidos, elenca todos os aspectos relevantes das comparações com outras técnicas, além de apresentar a principal aplicação do método proposto: a predição. O Capítulo 7, Conclusão, apresenta o caminho percorrido desde o início da pesquisa, salientando os entraves e aprendizados. No Apêndice A são enumerados artigos próprios, relacionados ao presente assunto, bem como as explanações com as novidades propostas e seus respectivos algoritmos. Os diversos testes efetuados resultaram no trabalho publicado e detalhado no Apêndice B, artigo principal, pré-requisito para a conclusão do programa de doutorado desta faculdade.

CAPÍTULO 2

Revisão bibliográfica

2.1 Alicerce Bibliográfico e Estado da Arte

Em relação ao alicerce bibliográfico e partindo do pressuposto que a identificação fuzzy de sistemas MISO constitui a ferramenta central para aplicações realizadas nesta pesquisa, são enumerados os principais artigos que sustentam as conclusões alcançadas.

O processo de identificação fuzzy utilizado segue as mesmas etapas sugeridas por Babuska (BABUSKA, 1999), no que diz respeito à clusterização seguida da inferência fuzzy do tipo Takagi-Sugeno (TAKAGI; SUGENO, 1985). Diante das várias técnicas fuzzy utilizadas para clusterização, o fuzzy C-means (BEZDEK, 1981) e Gustafson-Kessel (GUSTAFSON; KESSEL, 1979), também conhecido como GK, são os mais utilizados devido aos seus desempenhos (SERIR; RAMASSO; ZERHOUNI, 2012). O algoritmo de clusterização utilizado no presente trabalho é o GK em consequência da sua reconhecida qualidade baseada na distância de Mahalanobis.

Outros trabalhos contribuíram de forma decisiva para o desenvolvimento deste estudo e são citados a seguir:

O professor Robert Babuska, da Universidade de Delft-Holanda, efetuou diversos trabalhos na área de identificação fuzzy de sistemas. Estes trabalhos geraram numerosos artigos (ABONIY; BABUSKA; SZEIFERT, 2002), livros (BABUSKA, 1999) e um toolbox para Matlab (BABUSKA et al., 2018). Os citados estudos foram usados como referência durante muitos meses, os quais trouxeram grandes ganhos e aprendizados nos primeiros passos aqui elencados.

A área de identificação fuzzy atualmente tem focado em controle de sistemas (AGUIRRE, 2007); entretanto, neste estudo, optou-se pela comprovação da eficiência do novo método comparando-o a técnicas clássicas de identificação de sistemas.

Encorajados pelos resultados positivos a favor da nova metodologia, em termos de identificação de sistemas, têm sido feitos testes introdutórios na área de predição de dados, baseando-se nas informações sobre a dinâmica do passado. Técnicas similares podem ser encontradas na literatura (VORONIN; PARTANEN, 2013),(DOVZAN; SKRJANC,

2016), (MACIEL; GOMIDE; BALLINI, 2015) e (PEREIRA et al., 2014). Vários autores abordam outros métodos de predição que tem influenciado e contribuído para este estudo e que revisar-se-á seguidamente:

Plamen Angelov, professor da Universidade de Lancaster, importante autor que tem produzido notáveis artigos na área de clusterização on-line. Seus trabalhos foram decisivos no desenvolvimento dos estudos aqui apresentados. Em seu artigo (ANGELOV; FILEV, 2004) é apresentada uma abordagem para aprendizado on-line a partir do modelo do tipo TSK. Ele foi baseado em uma nova técnica que combina aprendizado supervisionado e não-supervisionado bem como a atualização recursiva do modelo do tipo TSK; novas regras e a revisão das regras antigas são feitas através do processo recursivo. Essa técnica é chamada de “Evolving Takagi-Sugeno Model”. O conceito foi aplicado em situações reais e se apresenta promissor para identificação da dinâmica de sistemas, controle (ANDONOVSKI et al., 2017), predição, extração de conhecimento, entre outros. Segundo (ANGELOV; FILEV, 2004) os modelos do tipo TSK se tornaram poderosos e largamente utilizados em problemas de engenharia, principalmente para problemas relacionados a controle. Até 2004, os modelos de predição fuzzy não eram adaptativos, o que impossibilitava sua aplicação on-line.

O artigo de autoria de Janos Abonyi (ABONIY et al., 2001) afirma que os modelos de identificação fuzzy de sistemas são adequados para a tarefa, entretanto podem se tornar complexos devido às características dos dados. Para solucionar o problema do excesso de regras, busca-se eliminar a redundância e criar um modelo mais “enxuto”. O trabalho propõe um novo método de agrupamento por via do algoritmo de clusterização, separação das variáveis consequentes mais relevantes do TSK por um método de mínimos quadrados ortonormais e, por fim, uma seleção dos antecedentes com base no critério de separabilidade de Fisher.

Murad (HARASHEH, 2016) desenvolveu um modelo que realiza a predição de preços de energia elétrica no mercado italiano. De acordo com ele, a predição dos preços de eletricidade se tornou um elemento crucial para a tomada de decisões públicas e privadas em todo o mundo. Diante desta situação, Murad apresenta um arquétipo de predição para o preço de eletricidade no mercado italiano, no qual foram utilizadas técnicas de inteligência artificial, redes neurais e árvores de regressão ensacadas. O modelo foi feito para prever o preço da energia no segundo semestre de 2015, entretanto, pode ser usado para outros períodos.

Dejan Dovzan, Universidade de Ljubljana, tem desenvolvido vários trabalhos nas seguintes áreas: sistemas fuzzy evoluídos, detecção de faltas e controle preditivo. Em seu artigo (DOVZAN; SKRJANC, 2016) é apresentada uma técnica para predição do consumo de energia elétrica. Destaca-se a importância desta predição devido a sua ligação aos lucros de empresas. Em (DOVZAN; SKRJANC, 2016) afirma-se que a estrutura do modelo Takagi-Sugeno é poderosa para lidar com dados altamente não-lineares.

Em termos de estado da arte, a área em questão ainda está em intensa produção acadêmica, sendo aplicada em diferentes âmbitos do conhecimento ao redor do mundo. Para ilustrar, são citados alguns recentes exemplos:

F. H. Fernandez et al. desenvolveu uma pesquisa que investiga o comportamento dos conjuntos fuzzy nos antecedentes durante a execução da inferência de Takagi-Sugeno (FERNÁNDEZ; KERRE; JIMÉNEZ, 2016). A observação é feita durante a identificação da dinâmica de sistemas não-lineares variantes no tempo. O ponto central do artigo é a determinação dos parâmetros do antecedente devido às características dos dados. Por fim, o artigo apresenta uma abordagem para gerar os parâmetros da inferência de Takagi-Sugeno, demonstrando a possibilidade de usar conjuntos fuzzy do tipo “interval-valued”.

Outro importante artigo que utiliza o fuzzy C-means é apresentado em (ZHOU; YANGA; SHAO, 2017), o qual expõe um processo de busca de padrões de consumo de lares que utilizam energia elétrica.

Em (ESPITIA et al., 2019) é exposta uma nova técnica neuro-fuzzy adaptativa para sistemas de controle fuzzy, aplicada a um ambiente de geração e consumo de energia distribuídos. A técnica fuzzy é aplicada em duas etapas: em primeiro lugar é feita a identificação da dinâmica do sistema e em seguida, é construída a estratégia de controle.

Mona Alghamdi e Plamen Angelov fizeram um estudo para prever a produtividade de uma lavoura de soja (ALGHAMDI et al., 2019). Além disso, identificaram os recursos mais impactantes nesta cultura e empregaram regras fuzzy auto-organizáveis.

Zhang Qinlia e Chen Yua, propõem um algoritmo Takagi-Sugeno para a filtragem de partículas (QINLIA; YUA, 2019). Esta técnica utiliza uma clusterização fuzzy C-regression e modela sistemas com ruídos não gaussianos. Um importante fator é que os parâmetros do consequente de Takagi-Sugeno foram alcançados através de clusterização fuzzy C-regression e pela filtragem de Kalman.

Em (BEZERRA et al., 2020) é exposto um algoritmo utilizado para detectar anomalias em fluxos de dados. O AutoCloud, nome da técnica, é on-line, recursivo e não precisa de treinamento ou conhecimento sobre os dados. Cada cluster é chamado de nuvem e cada dado pode pertencer a várias nuvens devido ao uso de conceitos fuzzy.

Os artigos citados anteriormente, junto a outros relacionados à predição, trouxeram a justificativa para este estudo. A presente abordagem comporta a influência de todos os autores citados, especialmente de Robert Babuska e Plamen Angelov. Em seguida, são apresentados os Fundamentos Teóricos, Capítulo 3, os quais são a sustentação do proposto estudo.

CAPÍTULO 3

Fundamentos teóricos

A modelagem é uma área vista como uma associação de um completo conhecimento da natureza, do comportamento do sistema e de um tratamento matemático apropriado que conduz a um modelo utilizável (AGUIRRE, 1996). A identificação de sistemas aborda a construção de modelos com base em dados experimentais; área que permite a criação de representações que simulam a dinâmica de um determinado processo ou fenômeno. O desenvolvimento de modelos matemáticos de sistemas reais é um tema importante em muitas disciplinas de engenharia e ciências, sendo usados para simulações, análises e entendimento dos mecanismos fundamentais de sistemas (AVELAR, 2012).

Nesta vasta área de estudo, os principais problemas estão ligados a modelos de fenômenos reais que, em sua maioria, apresentam não-linearidades. Existem várias técnicas lineares que alcançam modelos apropriados, entretanto o seu uso industrial, por exemplo, é restrito. Devido à presença de distúrbios, variáveis desconhecidas e complexa dinâmica de determinados conjuntos de dados, os modelos lineares podem não ser adequados. Na medida em que as representações lineares são substituídas em algumas aplicações por suas correspondentes não-lineares, torna-se possível analisar e reproduzir fenômenos e comportamentos dinâmicos mais complexos (AGUIRRE, 2007). Diversas são as técnicas desenvolvidas para o propósito citado; cada uma possui suas vantagens, desvantagens e peculiaridades. Diante disso, as abordagens com a teoria de conjuntos fuzzy se apresentam como opções promissoras (ANGELOV; FILEV, 2004; ALGHAMDI et al., 2019; BEZERRA et al., 2020).

Existem inúmeras técnicas de modelagem, que podem ser divididas em três grandes grupos (AGUIRRE, 2007):

1. Caixa-branca: a modelagem é feita através do conhecimento pleno dos aspectos envolvidos no fenômeno a ser estudado. A física e a matemática são utilizadas para que, por exemplo, equações diferenciais sejam concebidas, entre outras ferramentas que interpretam o modelo. O principal problema relacionado a esta abordagem é que nem sempre é possível ter todas as informações necessárias.

2. Caixa-cinza: há poucas informações além das entradas e saídas, o que permite, por exemplo, tomar decisões mais apropriadas durante a escolha de parâmetros da modelagem.
3. Caixa-preta: não há informações sobre as características do processo, ou seja, apenas as entradas e as saídas são conhecidas. O desafio das técnicas caixa-preta é alcançar um modelo aceitável na ausência de informações a priori.

Dentro desta terceira categoria têm surgido novas técnicas relacionadas ao assunto. O fato dos humanos serem frequentemente incitados a conduzir tarefas complexas sob considerável incerteza, tem estimulado a pesquisa de padrões alternativos de modelagem e controle (LOTUFO; LUCENA, 2011). Por esse motivo, são introduzidas as chamadas metodologias “inteligentes”, que empregam técnicas motivadas por sistemas biológicos e inteligência humana para desenvolver modelos e controladores aplicados a sistemas dinâmicos. Estas técnicas exploram métodos de representação alternativa, usando linguagem natural, regras, modelos qualitativos, entre outros. Modelagem e controle fuzzy são exemplos de técnicas que fazem uso do conhecimento humano e de processos dedutivos.

Entre sistemas inteligentes destaca-se a identificação de sistemas fuzzy. Neste contexto, um procedimento comum é a divisão da identificação fuzzy em duas etapas distintas: clusterização e inferência. Assim, a teoria dos conjuntos fuzzy constitui uma das principais ferramentas deste tipo de sistema inteligente, cujos conceitos básicos serão revisados na seguinte Seção, 3.1.

3.1 Teoria dos Conjuntos Fuzzy

No ano 1965, na Universidade da Califórnia em Berkley, Lotfali Askar-Zadeh publicou o primeiro trabalho sobre a teoria dos conjuntos fuzzy. O objetivo era permitir a representação de ideias vagas ou imprecisas (JAFELICE; BARROS; BASSANEZI, 2012).

Um *conjunto fuzzy* F do conjunto universo \mathcal{U} é o gráfico de uma função μ , denominada função de pertinência do conjunto F . A função μ associa a cada elemento x de \mathcal{U} um número $\mu_F(x)$, entre zero e um. O valor μ_F é chamado de grau de pertinência de x a F (JAFELICE, 2003). Assim, o conjunto fuzzy F pode ser escrito matematicamente como

$$F = \{(x, \mu_F(x)), x \in \mathcal{U}, \mu(x) \in [0, 1]\}$$

Os valores $\mu_F(x) = 1$ e $\mu_F(x) = 0$ indicam, respectivamente, a pertinência plena e a não pertinência do elemento x a F .

Um conjunto clássico A de \mathcal{U} pode ser associado a um conjunto fuzzy através da função característica de A , como função de pertinência, isto é,

$$\mu_A(x) = 1, \text{ se } x \in A; \mu_A(x) = 0, \text{ se } x \notin A;$$

Quando o conjunto A é unitário, é chamado de conjunto *crisp*, que também pode ser associado a um número real (JAFELICE, 2003). Outras definições necessárias são:

1. Para o número real $\alpha \in (0, 1]$, define-se o α -nível do conjunto fuzzy A como

$$[A]^\alpha = \{x \in \mathcal{U} : \mu_A(x) \geq \alpha\}.$$

2. O *suporte* de um conjunto fuzzy A é formado por todos os elementos de \mathcal{U} que têm grau de pertinência diferente de zero em A e denota-se por $\text{supp}(A)$. O *nível zero* do conjunto A é o fecho topológico do suporte de A .
3. Um conjunto fuzzy N é chamado *número fuzzy* quando o conjunto universo \mathcal{U} , em que N está definido, é o conjunto dos números reais \mathbb{R} e verifica-se que existe $x \in \mathcal{U}$ tal que $\mu_N(x) = 1$, $[N]^\alpha$ é um intervalo fechado, para todo $\alpha \in (0, 1]$ e o suporte de N é limitado.

Outro conceito importante da teoria dos conjuntos fuzzy é a *norma triangular* (t-norma) (JAFELICE; BARROS; BASSANEZI, 2012). Uma t-norma é uma operação binária : $[0, 1] \times [0, 1] \rightarrow [0, 1]$ tal que, para todo $x, y, z, w \in [0, 1]$, as seguintes propriedades são satisfeitas:

1. Comutatividade: $x * y = y * x$.
2. Associatividade: $(x * y) * z = x * (y * z)$.
3. Monotonicidade: se $x \leq y, w \leq z$, então $x * w \leq y * z$.
4. Condições de contorno: $x * 0 = 0$ e $x * 1 = x$.

Em continuação, descreve-se sucintamente o conceito de *sistema de inferência fuzzy* (Fuzzy Inference System - FIS em inglês). A literatura descreve vários tipos de sistemas de inferência, sendo os mais utilizados nas engenharias, o método de Mamdani (JAFELICE; BARROS; BASSANEZI, 2012) e o método de Sugeno. A primeira proposta de uma estrutura de inferência baseada na teoria de conjuntos fuzzy deste tipo, foi feita por Takagi e Sugeno em 1985 (TAKAGI; SUGENO, 1985) e, seguidamente, por Sugeno e Kang em 1988 (SUGENO; KANG, 1988). Esta estrutura é denominada sistema de inferência de Sugeno ou simplesmente modelo TSK. Estes modelos funcionam como bons aproximadores para sistemas representados através de suas relações de entrada e saída (black-box). Os dois modelos fuzzy são baseados na utilização de uma base de regras condicionais de inferência. A diferença está na relação das entradas e as saídas do processo.

Em geral, a estrutura de um modelo de inferência é composta pelos seguintes elementos:

1. Fuzzificação, na qual entradas precisas são transformadas em conjuntos fuzzy de entrada.

2. Regras, elaboradas por meio da opinião de um especialista (Método de Mamdani) ou por métodos numéricos, como no caso do modelo TSK.
3. Máquina de Inferência Fuzzy, a qual, no modelo Mamdani, retorna conjuntos fuzzy de saída, enquanto o modelo TSK utiliza funções afins.
4. Defuzzificação, que é o processo de obter-se uma saída precisa, a partir da entrada precisa do processo.

A Figura 1 apresenta o processo geral de construção de uma inferência fuzzy. Na

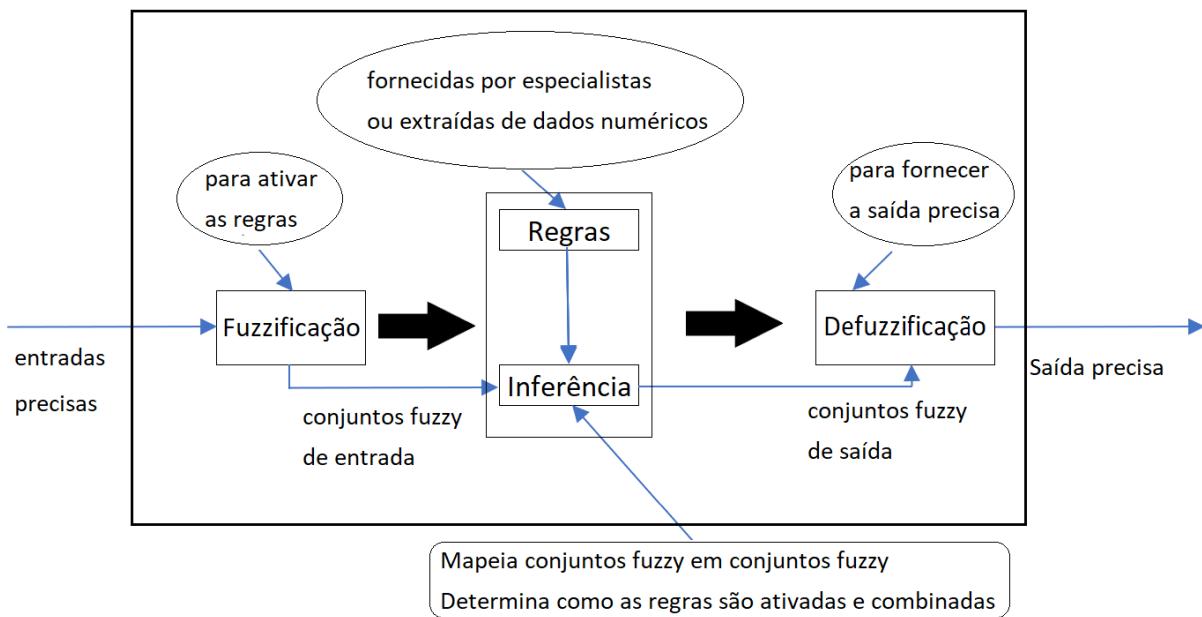


Figura 1 – Elementos presentes em um sistema de inferência fuzzy.

Fonte: Autores, 2020.

saída da inferência temos um conjunto fuzzy para o qual se escolhe um método para defuzzificar, obtendo assim uma saída precisa, que pode ser um número real ou um vetor que a represente. O método mais comum de defuzzificação é conhecido como *Centro de Gravidade* ou *Centroide*. Este método é a média ponderada pelos pesos $\mu_A(z_i)$, que indica o grau de compatibilidade do valor z_i com o conceito modelado pelo conjunto fuzzy A . Assim, o centroide de A tem como fórmula

$$\text{centroide de } A = \frac{\sum_{i=1}^N z_i \mu_A(z_i)}{\sum_{i=1}^N \mu_A(z_i)} \quad (1)$$

Nota-se que, particularmente para o modelo TSK, uma regra genérica, R_k , vem dada pela expressão

$$\text{Se } x^1 \in A_1^k \text{ e } x^2 \in A_2^k \dots \text{ e } x^m \in A_m^k \text{ então } y^k = \phi_k(x^1, x^2, \dots, x^m), \quad (2)$$

em que A_i^k são os conjuntos fuzzy antecedentes da regra R_k e ϕ_k uma função afim.

A partir da definição das regras, a defuzzificação é realizada através de uma t – norma, *, que consolida o número

$$\omega_k = \mu_{A_1^k} * \mu_{A_2^k} * \dots * \mu_{A_m^k},$$

o qual representa o peso da regra no sistema. De fato, supondo que r representa o número de regras, o valor defuzzificado obtido pelo método TSK, vem dado por

$$\hat{y} = \frac{\sum_{k=1}^r \omega_k \phi_k(x^1, x^2, \dots, x^m)}{\sum_{k=1}^r \omega_k}. \quad (3)$$

Uma maneira de obter-se uma expressão da função ϕ_k é utilizando método de regressão linear multivariado. A seguir, na Seção 3.2, far-se-á uma sucinta explanação deste método.

3.2 Modelo de Regressão Multivariada

Na regressão linear múltipla (STEINBRUCH; WINTERLE, 1987) assume-se que existe uma relação linear entre uma variável Y (a variável dependente) e k variáveis independentes, x_j , $j = 1, \dots, k$. Um conjunto de dados pode ser escrito na seguinte forma:

x_1	x_2	\dots	x_k	Y
x_{11}	x_{12}	\dots	x_{1k}	y_1
x_{21}	x_{22}	\dots	x_{2k}	y_2
\vdots	\vdots	\vdots	\vdots	\vdots
x_{n1}	x_{n2}	\dots	x_{nk}	y_n

Estão representados, na tabela, os valores de k variáveis independentes e os valores da variável resposta, depois de efetuada uma determinada experiência para uma amostra de tamanho n .

Um modelo de regressão linear múltipla descreve uma relação entre as Y variáveis independentes, x_j , e a variável dependente, Y , da seguinte maneira

$$Y = \theta_0 \mathbf{u} + \theta_1 x_1 + \theta_2 x_2 + \dots + \theta_k x_k + \varepsilon, \quad (4)$$

em que, $\mathbf{u} = (1, 1, \dots, 1)$ representa os parâmetros reais θ_j , $j = 0, \dots, k$, os quais são os coeficientes de regressão e ε o erro de ajuste. Supõe-se que o modelo (4) permite descrever a i -ésima resposta y_i , valendo a igualdade

$$y_i = \theta_0 + \theta_1 x_{i1} + \theta_2 x_{i2} + \dots + \theta_k x_{ik} + \varepsilon_i = \theta_0 + \sum_{j=1}^k \theta_j x_{ij} + \varepsilon_i, \quad i = 1, 2, \dots, n, \quad (5)$$

ou seja, cada observação $(x_{i1}, x_{i2}, \dots, x_{ik}, y_i)$ satisfaz a equação (5). Para que o ajuste deste modelo de regressão proceda, utiliza-se a notação matricial já que o modelo (4) é um sistema de n equações que pode ser representado da seguinte forma

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\theta} + \boldsymbol{\varepsilon} \quad (6)$$

em que

$$\mathbf{Y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}, \quad \mathbf{X} = \begin{bmatrix} 1 & x_{11} & x_{12} & \cdots & x_{1k} \\ 1 & x_{21} & x_{22} & \cdots & x_{2k} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & x_{n1} & x_{n2} & \cdots & x_{nk} \end{bmatrix}, \quad \boldsymbol{\theta} = \begin{bmatrix} \theta_0 \\ \theta_1 \\ \vdots \\ \theta_k, \end{bmatrix} \quad \text{e} \quad \boldsymbol{\varepsilon} = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n, \end{bmatrix}$$

O intuito seguinte é encontrar o vector $\hat{\boldsymbol{\theta}}$ que minimize a soma de quadrados do erro, para o qual reescreve-se a equação (6) da forma

$$\boldsymbol{\varepsilon} = \mathbf{Y} - \mathbf{X}\boldsymbol{\theta}.$$

Observando que $\mathbf{Y}^T \mathbf{X} \boldsymbol{\theta}$ é um número real, então o seu transposto é também um número real com o mesmo valor, ou seja,

$$\mathbf{Y}^T \mathbf{X} \boldsymbol{\theta} = (\mathbf{Y}^T \mathbf{X} \boldsymbol{\theta})^T = \boldsymbol{\theta}^T \mathbf{X}^T \mathbf{Y}.$$

Assim, obtém-se

$$\begin{aligned} |\boldsymbol{\varepsilon}|^2 &= \sum_{i=1}^n \varepsilon_i^2 = \boldsymbol{\varepsilon}^T \boldsymbol{\varepsilon} \\ &= (\mathbf{Y} - \mathbf{X} \boldsymbol{\theta})^T (\mathbf{Y} - \mathbf{X} \boldsymbol{\theta}) \\ &= (\mathbf{Y}^T - (\mathbf{X} \boldsymbol{\theta})^T)(\mathbf{Y} - \mathbf{X} \boldsymbol{\theta}) \\ &= \mathbf{Y}^T \mathbf{Y} - \mathbf{Y}^T \mathbf{X} \boldsymbol{\theta} - (\mathbf{X} \boldsymbol{\theta})^T \mathbf{Y} + (\mathbf{X} \boldsymbol{\theta})^T (\mathbf{X} \boldsymbol{\theta}) \\ &= \mathbf{Y}^T \mathbf{Y} - \mathbf{Y}^T \mathbf{X} \boldsymbol{\theta} - \boldsymbol{\theta}^T \mathbf{X}^T \mathbf{Y} + \boldsymbol{\theta}^T \mathbf{X}^T (\mathbf{X} \boldsymbol{\theta}) \\ &= \mathbf{Y}^T \mathbf{Y} - 2\boldsymbol{\theta}^T \mathbf{X}^T \mathbf{Y} + \boldsymbol{\theta}^T \mathbf{X}^T \mathbf{X} \boldsymbol{\theta} \end{aligned}$$

Dessa forma, o vetor solução da equação (6) será o que resolve a equação

$$\frac{\partial |\boldsymbol{\varepsilon}|^2}{\partial \boldsymbol{\theta}} = 0,$$

que é equivalente a

$$-2\mathbf{X}^T \mathbf{Y} + 2\mathbf{X}^T \mathbf{X} \boldsymbol{\theta} = 0,$$

ou ainda equivalente a

$$\mathbf{X}^T \mathbf{X} \boldsymbol{\theta} = \mathbf{X}^T \mathbf{Y}$$

de onde obtém-se a solução para o ajuste da tabela de dados

$$\hat{\boldsymbol{\theta}} = (\mathbf{X}^T \mathbf{X})^+ \mathbf{X}^T \mathbf{Y}, \quad (7)$$

na qual a notação A^+ denota a matriz de Penrose da matriz A .

De fato, e completando esta explanação relacionada ao importante conceito de matriz pseudo-inversa, considera-se um sistema linear $A\gamma = y$, em que A é uma matriz $m \times (2k+1)$, com $m > 2k + 1$, e em que γ e y são vetores-colunas de comprimento m . Supondo que o sistema é compatível, é possível ter tanto uma única solução e, como consequência, a matriz A é inversível; ou é possível existir infinitos vetores γ que satisfazem o sistema linear $A\gamma = y$. Desse conjunto solução, é preferida a escolha daquele que tenha a menor norma euclidiana, ou seja, o desejo é o de resolver o problema de minimização

$$\hat{\gamma} = \min\{\|\gamma\|, \text{ } C \text{ solução de } A\gamma = y\} \quad (8)$$

em que $\|\cdot\|$ é a norma euclidiana de \mathbb{R}^m . Moore (MOORE, 1920) e Penrose (PENROSE, 1955) mostraram que o problema de minimização (8) tem uma única solução obtida através da aplicação da pseudo-inversa da matriz A , denominada A^+ , sendo portanto a solução $\hat{\gamma} = A^+ Y$. Nos casos em que a matriz A é inversível, tem-se que $A^{-1} = A^+$. Estas afirmações são as utilizadas na abordagem de regressão apresentada anteriormente.

Outro importante aspecto na modelagem desta pesquisa é que a inferência TSK está relacionada com a construção dos seus antecedentes que é realizada, na metodologia proposta, via clusterização fuzzy. Uma revisão deste conceito é feita a seguir na Seção 3.3.

3.3 Clusterização Fuzzy de Dados

Técnicas de clusterização são, em sua maioria, métodos não supervisionados que podem ser usados para organizar grupos a partir de um critério de similaridade matemática. Os algoritmos de clusterização possuem a capacidade de revelar estruturas subjacentes e suas aplicações são várias, como: processamento de imagens, reconhecimento de padrões, identificação de sistemas, entre outros. Do ponto de vista da teoria fuzzy, há, entre outras categorias, dois critérios mais utilizados de classificação dos elementos de um conjunto (BEZDEK, 1981):

1. O clustering crisp, no qual cada elemento pertence somente a um cluster.
2. O clustering fuzzy, o qual permite que um elemento pertença a vários clusters simultaneamente, com diferentes graus de pertinência.

Na Figura 2 há dois conjuntos de elementos: $Q = \{A, B, C, D, E, F\}$ e $W = \{G, H, I, J, K, L, M\}$ que têm suas pertinências apresentadas na Tabela 1.

Dada uma clusterização A_1, A_2, \dots, A_c do conjunto Z , as características do clustering crisp são:

1. A união de todos os clusters é igual ao conjunto de todos os dados, ou seja, $\bigcup_{i=1}^c A_i = Z$.

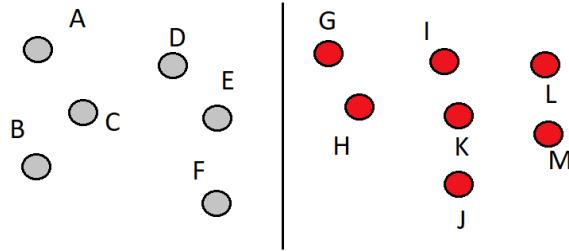


Figura 2 – Representação de um clustering crisp referente aos conjuntos Q e W.

Fonte: Autores, 2020.

Tabela 1 – Pertinência dos dados da Figura 2 referentes aos conjuntos Q e W.

A	B	C	D	E	F	G	H	I	J	K	L	M
1	1	1	1	1	1	0	0	0	0	0	0	0
0	0	0	0	0	0	1	1	1	1	1	1	1

2. A intersecção de dois clusters quaisquer é vazia, ou seja, $A_i \cap A_j = \emptyset$, $1 \leq i \neq j \leq c$.
3. Não se pode ter um cluster sem pelo menos um elemento, ou seja, $A_i \neq \emptyset$, $1 \leq i \leq c$.

A Figura 3 ilustra os mesmos elementos da Figura 2; entretanto é representada uma gradatividade das pertinências, desde as cores mais escuras até as cores mais claras nos elementos de cada conjunto, indicando o grau de pertinência ao cluster. Os respectivos graus de pertinência são detalhados na Tabela 2.

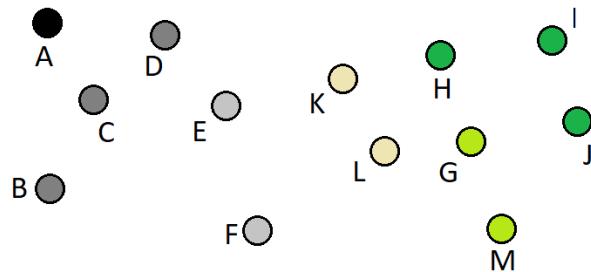


Figura 3 – Representação de um clustering fuzzy referente aos conjuntos Q e W.

Fonte: Autores, 2020.

Tabela 2 – Pertinência dos dados da Figura 3 referentes aos conjuntos Q e W.

A	B	C	D	E	F	G	H	I	J	K	L	M
1	0.8	0.8	0.8	0.7	0.7	0.3	0.2	0	0.2	0.6	0.6	0.3
0	0.2	0.2	0.2	0.3	0.3	0.7	0.8	1	0.8	0.4	0.4	0.7

O diferencial da clusterização fuzzy é o modo como os dados são particionados: cada elemento possui certo grau de pertinência em relação a cada cluster, isto é, valores entre 0

e 1 definem “o quanto” cada elemento se relaciona a determinado cluster. Tal conceito faz com que cada elemento pertença a todos os conjuntos, entretanto a “intensidade” de sua relação é definida pelo grau de pertinência. De fato, sendo $Z = \{z_1, z_2, \dots, z_i, \dots, z_N\}$ um conjunto de dados pertencente ao universo U dos conjuntos fuzzy C_1, C_2, \dots, C_c , denominados clusters da clusterização fuzzy, denota-se por μ_{ij} , $1 \leq i \leq N$ e $1 \leq j \leq c$, a pertinência do elemento z_i ao cluster C_j . Assim, as características da clusterização fuzzy são:

1. Cada elemento z_i tem uma pertinência $0 \leq \mu_{ij} \leq 1$ referida a cada cluster C_j ;
2. O somatório de todas as pertinências de um elemento em relação a cada cluster deve ser igual a 1, ou seja, $\sum_{j=1}^c \mu_{ij} = 1$ para cada $i = 1, 2, \dots, N$;
3. A soma das pertinências de todos os z_i de um cluster C_j não pode ser nula, ou seja, $\sum_{i=1}^N \mu_{ij} > 0$, $j = 1, \dots, c$.

No caso, $Z \subset \mathbb{R}^n$, $n \in \mathbb{N}$ é um conjunto dados

$$z_i = (i, x_{i1}, x_{i2}, \dots, x_{i(n-1)}, z_{in}),$$

em que z_{in} representa a saída dos dados que depende dos atributos (entradas) $x_{i1}, x_{i2}, \dots, x_{i(n-1)}$ e do indexador $i \in \mathbb{N}$, no qual será aplicado o método de clusterização de vetores de \mathbb{R}^n , contendo N amostras de uma observação. Quando o sistema analisado é dinâmico-autônomo (leva implícita a variável temporal) então o primeiro atributo x_{i1} assume o papel de controlador do sistema. Sem perda de generalidade, assumimos que a primeira entrada $i \in \mathbb{N}$ é o controlador do sistema na metodologia.

Associada a Z há, no processo de clusterização, a chamada matriz de pertinência dada por

$$U = \begin{bmatrix} \mu_{11} & \dots & \mu_{1j} & \dots & \mu_{1c} \\ \mu_{21} & \dots & \mu_{2j} & \dots & \mu_{2c} \\ \vdots & \vdots & \ddots & & \vdots \\ \mu_{N1} & \dots & \mu_{Nj} & \dots & \mu_{Nc} \end{bmatrix}, \quad (9)$$

em que μ_{ij} é a pertinência do ponto z_i ao cluster C_j .

Uma forma de similaridade muito usada, que atende aos objetivos geométricos, é a distância definida por

$$d_{ij, M_j}^2 = (\mathbf{z}_i - \mathbf{v}_j)^T \mathbf{M}_j (\mathbf{z}_i - \mathbf{v}_j), \quad 1 \leq j \leq c, \quad 1 \leq i \leq N, \quad (10)$$

em que M_j é uma matriz simétrica definida positiva e $\mathbf{v}_j \in \mathbb{R}^n$ são os chamados centros de clusters C_j . Quanto menor a distância do elemento ao centro de cluster, o grau de pertinência é maior, sendo 1 quando ambos coincidem.

De acordo com a álgebra linear, dada uma matriz $M_{n \times n}$ simétrica e definida positiva, pode-se definir em \mathbb{R}^n um produto interno (STEINBRUCH; WINTERLE, 1987) pela fórmula

$$\langle v, w \rangle = v \cdot M \cdot w^T, \quad (11)$$

em que $v, w \in \mathbb{R}^n$. A norma ao quadrado induzida por (11) é, consequentemente, da forma

$$\|v\|^2 = \langle v, v \rangle,$$

e, por conseguinte, a distância induzida em \mathbb{R}^n por essa norma é dada por

$$d(v, w) = \|v - w\| = (v - w) \cdot M \cdot (v - w)^T. \quad (12)$$

Nota-se que a distância euclidiana corresponde à matriz identidade em \mathbb{R}^n , ou seja $M = I$.

Voltando à clusterização fuzzy, várias metodologias são bem conhecidas e utilizadas na área das engenharias. Um dos primeiros métodos se tornou conhecido como Fuzzy C-means, o qual foi desenvolvido por James Dunn (DUNN, 1974) e aprimorado por James Bezdek (BEZDEK, 1981).

Após o surgimento do algoritmo fuzzy C-means, emerge uma numerosa família de algoritmos de clusterização que empregam novos meios de realizar o processo de agrupamento (clustering). O algoritmo de Gustafson-Kessel (GUSTAFSON; KESSEL, 1979) (GK) faz parte do conjunto de melhorias realizadas a partir do fuzzy C-means. Este algoritmo é uma poderosa técnica com uma vasta gama de aplicações. Sua principal característica é a adaptação geométrica do formato do cluster estimado a partir da matriz M que se baseia na distância de Mahalanobis (MAESSCHALCK, 2000). Dessa forma, a cada iteração, k , a estrutura geométrica dos clusters muda, enquanto o algoritmo C-means usa a mesma distância em todas iterações.

O algoritmo GK tem duas alternativas de início:

1. Construir uma matriz de pertinência, em geral escolhida em forma randômica (ABONIY; BABUSKA; SZEIFERT, 2002).
2. Escolher os centros de clusters entre os dados Z , também em forma randômica.

Neste estudo notou-se que a primeira opção pode levar a um esforço computacional maior, particularmente quando há um número grande de dados. Assim, foi adotada a segunda opção como início do algoritmo.

Procura-se, a seguir, uma relação entre a distância de Mahalanobis e o grau de pertinência dos pontos amostrados, formulando-se assim um problema de otimização

$$\text{Mínimo } \sum_{i=1}^c \sum_{j=1}^N \mu_{ij}^2 d_{ij, M_j}^2, \quad (13)$$

para alguma matriz M_j simétrica definida positiva e restrita às condições:

$$\mathbf{R}_1 : 0 \leq \mu_{ij} \leq 1, i = 1, \dots, c, j = 1, \dots, N$$

$$\mathbf{R}_2 : \sum_{j=1}^N \mu_{ij} > 0, i = 1, \dots, c; \quad (14)$$

$$\mathbf{R}_3 : \sum_{i=1}^c \mu_{ij} = 1, j = 1, \dots, N.$$

que são as características de uma clusterização fuzzy.

Resolvendo o problema (13)-(14) através do método dos multiplicadores de Lagrange, encontra-se:

$$\mu_{ij} = \frac{1}{\sum_{k=1}^c \left(\frac{d_{ij, M_j}^2}{d_{ik, M_j}^2} \right)}, i = 1, 2, \dots, N, j = 1, 2, \dots, c \quad (15)$$

ou seja, determina-se uma relação entre a distância d_{ij, M_j} e os graus de pertinência. A Eq. (15) será utilizada para construir a matriz U nas diferentes iterações. No passo seguinte, calcula-se a chamada matriz de covariância:

$$\mathbf{F}_j = \frac{\sum_{i=1}^N \mu_{ij}^2 (\mathbf{z}_i - \mathbf{v}_j)(\mathbf{z}_i - \mathbf{v}_j)^T}{\sum_{i=1}^N \mu_{ij}^2},$$

a qual permite uma atualização da matriz M_j por meio da fórmula

$$\mathbf{M}_j = \det(\mathbf{F}_j)^{\frac{1}{n}} \mathbf{F}_j^{-1},$$

sendo \mathbf{F}_j^{-1} a matriz inversa de \mathbf{F}_j , definindo assim uma nova distância para cada iteração.

Finalmente, os centros dos clusters são calculados da forma

$$v_j = \frac{\sum_{i=1}^N \mu_{ji}^m z_i}{\sum_{i=1}^N \mu_{ji}^m}. \quad (16)$$

É checado o valor do erro na iteração, o qual é calculado a partir da norma do máximo da diferença entre a matrizes de pertinências antes e após a alteração de M_j , ou seja

$$\text{erro} = \max_{\substack{i \in [1, N], \\ j \in [1, c]}} |U_{ij}^k - U_{ij}^{k+1}|.$$

Define-se um critério de parada do algoritmo dado por $\text{erro} \leq \varepsilon$, ε é um número pre-determinado.

Resumindo, um pseudocódigo do algoritmo de clusterização é apresentado no Algoritmo 1, considerando a base de dados denominada Energy.

Algoritmo 1 Clusterização

```

1:  $Z \leftarrow \text{energy}$ 
2:  $n \leftarrow \text{colunas}(Z)$  //  $\text{colunas}(Z)$  é o número de colunas de  $Z$ 
3:  $t \leftarrow Z(:, 1)$  //  $Z(:, 1)$  significa todas as linhas de  $Z$ 
Require:  $c$  // valor obtido por uma rotina de otimização
4:  $m \leftarrow 2$ 
5:  $tol \leftarrow 1$ 
6:  $N \leftarrow \text{linhas}(Z)$  //  $\text{linhas}(Z)$  é o número de linhas de  $Z$ 
7:  $N1 \leftarrow \text{uns}(N, 1)$  //  $\text{uns}$  significa vetor de entradas iguais a 1
8:  $n1 \leftarrow \text{uns}(n, 1)$ 
9:  $c1 \leftarrow \text{uns}(1, c)$ 
10:  $U \leftarrow \text{zeros}(N, c)$  //  $\text{zeros}$  significa vetor de entradas iguais a 0
11:  $d \leftarrow U$ 
12:  $v \leftarrow []$  //  $[]$  é um arranjo de elementos (vetores, matrizes, funções)
13:  $e1 \leftarrow \text{arredonda}(N/c)$ 
14: for  $j = 1$  to  $c$  do
15:   for  $i = 1$  to  $n$  do
16:      $V(j, i) \leftarrow Z(e1 * j, i)$ 
17:   end for
18: end for
19: for  $j = 1$  to  $c$  do
20:    $ZV \leftarrow Z - N1 * V(j, :)$ 
21:    $d(:, j) \leftarrow \text{soma}((ZV \cdot^2)^T)^T$  //  $\cdot^2$  significa elevar cada entrada ao expoente dado
22: end for
23:  $d \leftarrow (d) \cdot^{(-1/(m-1))}$ 
24:  $U0 \leftarrow (d \cdot /(\text{soma}(d^T)^T * c1))$  //  $\cdot /$  significa dividir cada entrada ao expoente dado
25: while  $\text{max}(\text{max}(U0 - U)) > tol$  do
26:    $U \leftarrow U0$ 
27:    $Um \leftarrow U \cdot^m$ 
28:    $SomaU \leftarrow \text{soma}(Um)$ 
29:    $V \leftarrow (Um^T * Z) \cdot / (n1 * SomaU)^T$ 
30:   for  $j = 1$  to  $c$  do
31:      $ZV \leftarrow Z - N1 * V(j, :)$ 
32:      $f \leftarrow n1 * Um(:, j)^T \cdot^* ZV^T * ZV / SomaU(1, j)$  //  $\cdot^*$  significa multiplicar
        cada entrada ao expoente dado
33:      $d(:, j) \leftarrow \text{soma}(ZV * (\text{det}(f)(1/n) * \text{Penrose}(f)) \cdot^* ZV, 2)$  //  $\text{Penrose}(A)$ 
        é a pseudo-inversa de A
34:   end for
35:    $d \leftarrow (d) \cdot^{\wedge}(-1/(m-1))$ 
36:    $U0 \leftarrow (d \cdot /(\text{soma}(d^T)^T * c1))$ 
37: end while

```

Na metodologia apresentada neste estudo, após a etapa de clusterização descrita nesta seção, segue a construção de um sistema de inferência fuzzy que difere do tradicional por seus elementos inovadores. Os citados elementos são retratados detalhadamente no Capítulo 4. Além disso, são apresentados conceitos que sustentam os possíveis trabalhos futuros para o estudo em voga. Diante desta necessidade apresenta-se a Seção 3.4.

3.4 Ferramentas para Trabalhos Futuros

Os parâmetros α e β são escolhidos através de um algoritmo de otimização, no qual são estabelecidos vários valores e busca-se aquele que minimiza o erro. Uma outra forma mais consistente para a escolha destes parâmetros é considerá-los como números fuzzy do tipo 2 intervalares.

De fato, em 1975, Zadeh introduziu os conjuntos fuzzy tipo 2 (ZADEH, 1975), o que fez com que conjuntos fuzzy, como descritos na Seção 3.1, se tornassem conhecidos como conjuntos Fuzzy do tipo 1. Após algumas décadas, os trabalhos de Karnik e Mendel (KARNIK; MENDEL, 1998; MENDEL, 2007) aparecem como os primeiros a apresentar a teoria completa do SBRF (sistema baseado em regras fuzzy) do tipo 2, incluindo as operações, o redutor de tipo e métodos de defuzzificação. O desempenho do SBRF do tipo 2 mostrou uma maior capacidade de trabalhar com incertezas e seu auge atual é devido a essa importante característica.

Alguns dos conceitos básicos desta extensão da teoria fuzzy serão expostos a seguir: dado um conjunto universo X , um conjunto fuzzy \tilde{A} do tipo 2 em X é o gráfico da função $\mu_{\tilde{A}}$ definida em $X \times [0, 1]$ em que $\mu_{\tilde{A}}(x, u) \in [0, 1]$, denominada função de pertinência de \tilde{A} . Em outras palavras, o conjunto fuzzy do tipo 2 de X (\tilde{A}) é dado por:

$$\tilde{A} = \{((x, u), \mu_{\tilde{A}}(x, u)) | (x, u) \in X \times [0, 1], \mu_{\tilde{A}}(x, u) \in [0, 1]\}.$$

Quando a função de pertinência é constante e igual a 1, ou seja, $\mu_{\tilde{A}}(x, u) = 1$ então o conjunto chama-se conjunto *fuzzy do tipo 2 intervalar*. Na Figura 4 é apresentado um exemplo de conjunto fuzzy do tipo 2 intervalar.

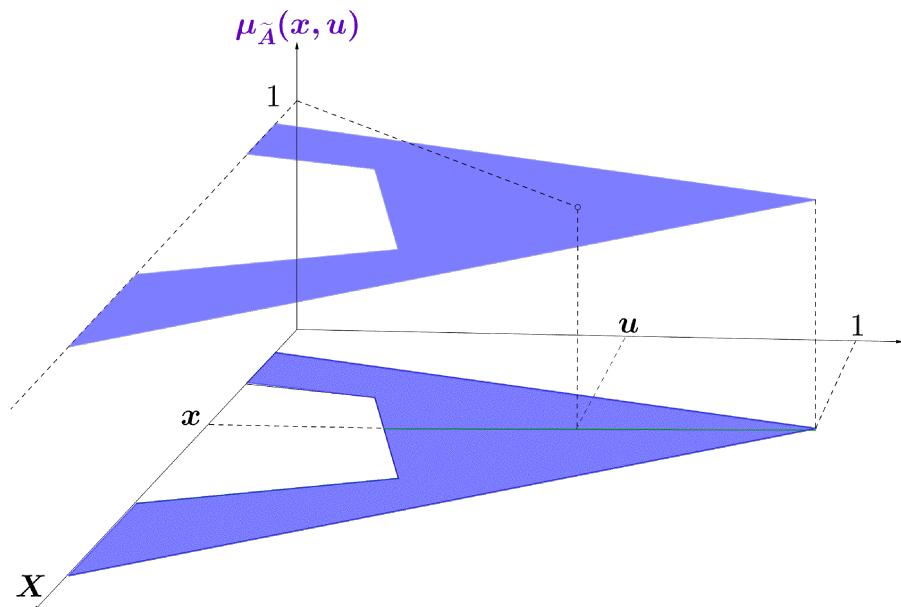


Figura 4 – Conjunto fuzzy do tipo 2 intervalar.

Fonte: Autores, 2020.

A seguir, são apresentadas funções de pertinência primárias do tipo 1 superior e inferior do conjunto fuzzy tipo 2 intervalar, denotadas por $\bar{\mu}_{\tilde{A}}(x)$, $\underline{\mu}_{\tilde{A}}(x)$, $x \in X$, respectivamente, sendo definidas por:

$$\bar{\mu}_{\tilde{A}}(x) = \sup\{u|u \in [0, 1], \mu_{\tilde{A}}(x, u) > 0\}, \quad (17)$$

$$\underline{\mu}_{\tilde{A}}(x) = \inf\{u|u \in [0, 1], \mu_{\tilde{A}}(x, u) > 0\}. \quad (18)$$

O conjunto J_x é, por definição, o conjunto dos pares (x, u) do domínio de \tilde{A} de pertinência positiva, ou seja,

$$J_x = \{(x, u)|u \in [\underline{\mu}_{\tilde{A}}(x), \bar{\mu}_{\tilde{A}}(x)]\}.$$

A região, limitada pelas funções de pertinência primárias inferior e superior, é conhecida como a “pegada da incerteza” (Footprint Of Uncertain, FOU), ou seja

$$FOU(\tilde{A}) = \{(x, u)|x \in X \text{ e } u \in [\underline{\mu}_{\tilde{A}}(x), \bar{\mu}_{\tilde{A}}(x)]\}, \quad (19)$$

em que $\bar{\mu}_{\tilde{A}}(x)$ e $\underline{\mu}_{\tilde{A}}(x)$ são definidas em (17) e (18), respectivamente. Nota-se que um conjunto fuzzy do tipo 2 intervalar está totalmente definido pela FOU. Assim, o gráfico utilizado é de duas dimensões; a Figura 5 ilustra um conjunto fuzzy do tipo 2 intervalar gaussiano, definido completamente pela FOU.

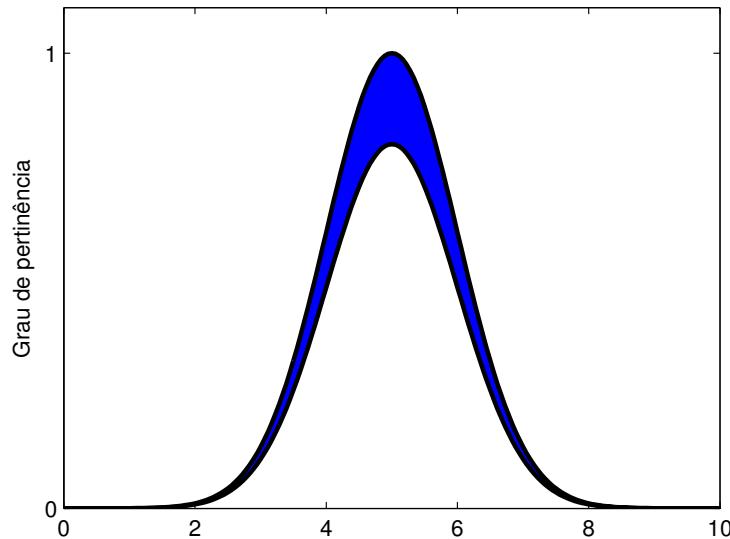


Figura 5 – Exemplo de conjunto gaussiano fuzzy do tipo 2 intervalar.

Fonte: Autores, 2020.

O esquema de um SBRF do tipo 2 intervalar é apresentado na Figura 6. O SBRF do tipo 2 é composto por cinco componentes: fuzzificador, inferência, base de regras, redutor

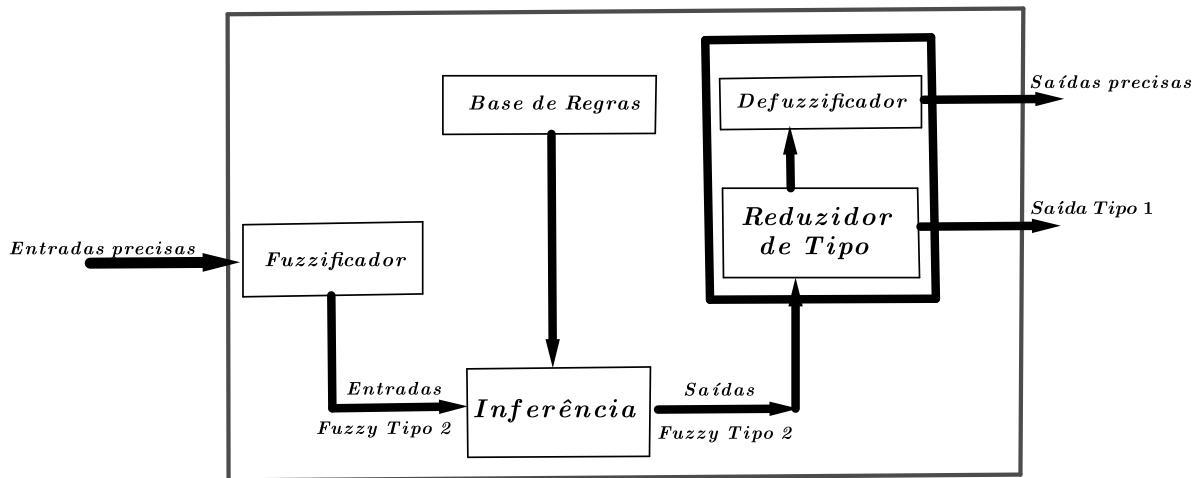


Figura 6 – Sistema baseado em regras fuzzy do tipo 2.

Fonte: Autores, 2020.

do tipo e defuzzificador. Este sistema é composto por, no mínimo, um conjunto fuzzy do tipo 2 presente em um dos antecedentes que compõe uma das regras que formam o sistema.

Karnik e Mendel mostraram que existe o mínimo, c_m , e o máximo, c_M , dos centroides, equação (1), dos conjuntos fuzzy do tipo 1 incluídos na FOU. Assim, criou-se o algoritmo de Karnik-Mendel que é o processo conhecido como *Reducor de Tipo* dentro do SBRF do tipo 2. Desta etapa pode-se obter um conjunto fuzzy do tipo 1 como saída do sistema ou uma saída defuzzificada dada pela média de c_m e c_M . Mediante os citados conceitos, apresenta-se o Capítulo 4 que detalha a inferência TSK proposta.

CAPÍTULO 4

O sistema proposto

4.1 A Inferência TSK Proposta

Nesta etapa é que reside a inovação da metodologia proposta nesta pesquisa, na qual é construído um sistema de inferência fuzzy do tipo Sugeno. A primeira proposta de uma estrutura de inferência baseada na teoria de conjuntos fuzzy, desse tipo, foi feita por Takagi e Sugeno em 1985 (TAKAGI; SUGENO, 1985) e, em seguida, por Sugeno e Kang em 1988 (SUGENO; KANG, 1988). Esse tipo de inferência fuzzy é conhecido, na atualidade, como Takagi-Sugeno-Kang (TSK).

Como consequência da clusterização, são obtidos quatro elementos fundamentais para a construção da inferência TSK:

1. Os centros \mathbf{v}_j de clusters que serão os centros das funções de pertinências do tipo gaussiana, correspondentes aos antecedentes da inferência.
2. Matriz de pertinências U .
3. As projeções dos α -níveis sobre o primeiro eixo i (controlador da indexação) correspondente aos antecedentes, A_j , dados por

$$[A_j]^\alpha = \{i \in \mathbb{N}, \mu_{ij}(z_i) \geq \alpha\}, \quad j = 1, \dots, c,$$

em que o α é determinado por um processo de otimização.

4. O desvio padrão da função do tipo gaussiana é dado por $\sigma_j = \beta(\max A_j - \min A_j)$, β determinado pelo mesmo processo de otimização.

Nota-se que os conjuntos A_j representam uma projeção dos valores de pertinência mais altos sobre o primeiro eixo i de coordenadas correspondente a

- Uma variável temporal (t) caso o modelo seja evolutivo no tempo.

- Uma variável de controle escolhida dentro de uma faixa adequada para cada modelo. Esta variável é fornecida em modelos nos quais a saída do fenômeno é conhecida ou controlada.
- Um eixo de valores inteiros que representa a posição (*i*) do vetor z_i na lista de dados Z para modelos não-evolutivos sem informação de controle de saída.

Na Figura 7 apresenta-se o processo de “projetar” as pertinências maiores da matriz de pertinências U sobre o primeiro eixo t .

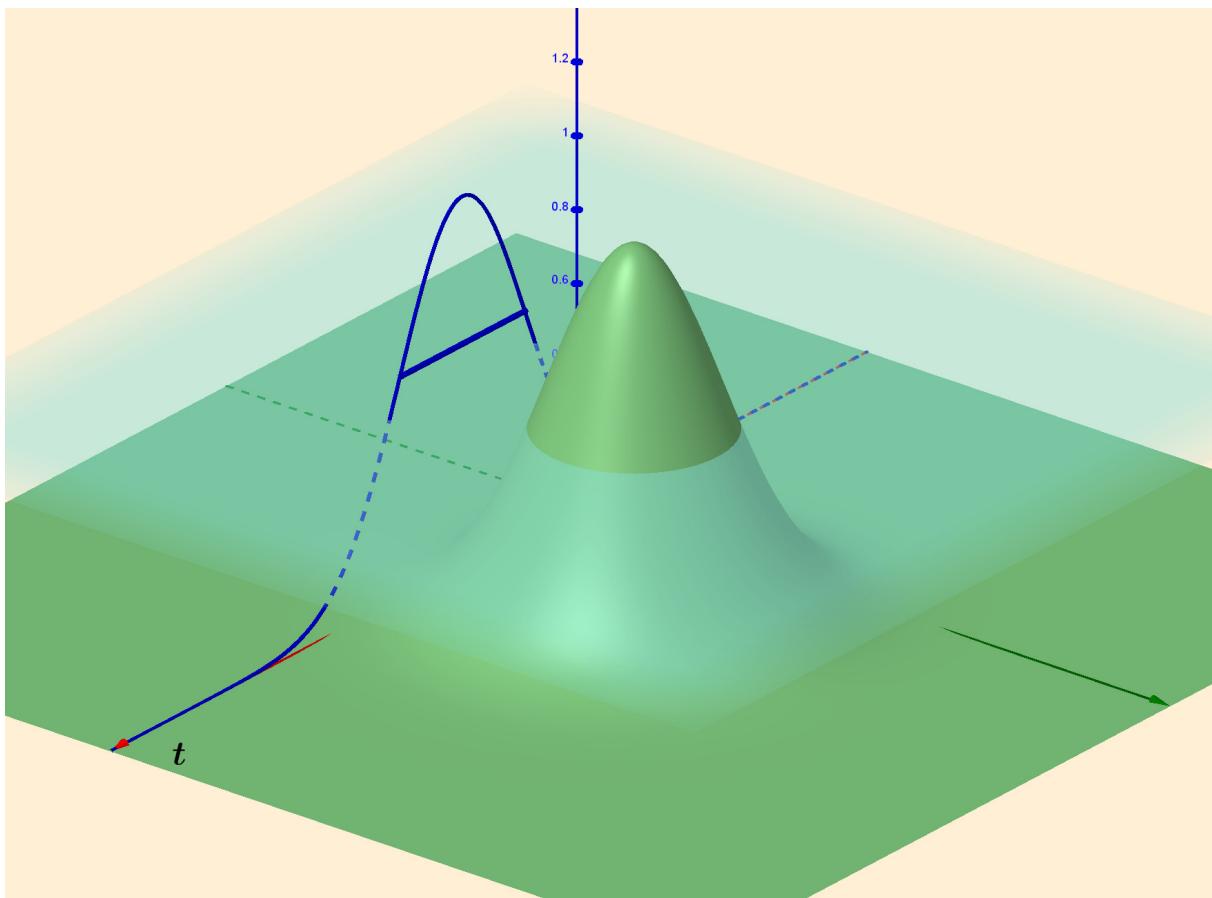


Figura 7 – Projeção do α -nível da matriz de pertinências sobre o eixo *i*.

Fonte: Autores, 2020.

Assim, ficam definidos os antecedentes da inferência sendo

$$\varrho_j(x) = \exp\left(-\frac{(x - v_j)^2}{2\sigma_j^2}\right), \quad j = 1, \dots, c, \quad (20)$$

Consequentes da inferência: regressão multivariada local incorporando os atributos.

$$z_{in} = \theta_{j0} + \theta_{j1}i + \sum_{k=2}^n \theta_{jk}x_{i(k-1)} + \delta_{ji}, \quad i = 1, 2, \dots, N,$$

em que $\boldsymbol{\theta}_j = (\theta_{j0}, \theta_{j1}, \dots, \theta_{jn})$, são as incógnitas e $\boldsymbol{\delta}_j = (\delta_{j1}, \delta_{j2}, \dots, \delta_{jN})$, é o erro $\boldsymbol{\delta}_j$ a minimizar. Salienta-se que a regressão linear multivariada permite a representação de todos atributos no processo de defuzzificação.

Representando o primeiro eixo de coordenadas por i , as regras fuzzy são então estabelecidas por: Obtido o vector $\boldsymbol{\theta}_j = (\theta_{j0}, \theta_{j1}, \dots, \theta_{jn})$ são estabelecidas as regras, em número igual ao número de clusters. De fato, $(x_0, x_1, x_2, \dots, x_{n-1})$, $x_0 \in [1, N]$, é definido

Regra^j: Se x_0 é A_j então $Y^j(x_0) = \theta_{j0} + \theta_{j1}x_0 + \sum_{k=2}^n \theta_{jk}x_{k-1}$.

O número de regras é definido pelo número de clusters, o qual é alcançado através do mesmo processo de otimização citado para α e β .

Por último, é feita a defuzzificação, da forma clássica, calculando a média ponderada como segue

$$y(z_{i1}, z_{i2}, \dots, z_{i(n-1)}) = \frac{\sum_{j=1}^c \varrho_j(i) \cdot \mathbf{f}(i)}{\sum_{j=1}^c \varrho_j(i)}, \quad i = 1, 2, \dots, N. \quad (21)$$

Nota-se que expressão (21) é uma *fórmula explícita contínua* que ajusta os dados experimentais. Finalmente, o pseudocódigo (Algoritmo 2), e o fluxograma de identificação (Figura 8) são apresentados.

Algoritmo 2 Identificação

```

1:  $Z \leftarrow energy$ 
2:  $n \leftarrow colunas(Z)$ 
3:  $t \leftarrow Z(:, 1)$ 
Require:  $Clusterização(Z)$  // sub-rotina como a do Algoritmo 1
4:  $u \leftarrow U0$ 
5:  $univ \leftarrow []$ 
6: for  $j = 1$  to  $c$  do
7:    $univ(j) \leftarrow U(:, j) > \alpha$  //  $\alpha$  é determinado por uma sub-rotina de otimização
8: end for
9:  $empties \leftarrow univ = \emptyset$ 
10:  $univ(empties) \leftarrow []$ 
11:  $c \leftarrow comprimento(univ)$ 
12:  $sigma \leftarrow []$ 
13: for  $j = 1$  to  $c$  do
14:    $sigma(1, j) \leftarrow \beta * (max(univj) - min(univj))$  //  $\beta$  é determinado pela mesma
      sub-rotina de otimização que  $\alpha$ 
15: end for
16:  $Y \leftarrow []$ 
17: for  $j = 1$  to  $c$  do
18:    $y(j, :) \leftarrow gaussiana(sigma(1, j), V(j))$  // gaussiana de centro  $v_j$  e desvio padrão
       $\sigma$ 
19: end for
20:  $X \leftarrow [Z(:, 1 : n - 1), N1]$  //  $[A, B]$  é a concatenação das matrizes  $A$  e  $B$ 
21:  $n2 \leftarrow uns(n, 1)$ 
22:  $normemb \leftarrow []$ 
23:  $tkout \leftarrow []$ 
24: for  $j = 1$  to  $c$  do
25:    $normemb(j, :) \leftarrow y(j, :) \cdot /soma(y(j, :))$ 
26:    $tkoutj \leftarrow Penrose((X^T \cdot^* (n2 * normemb(j, :))) * X) * (X^T \cdot^* (n2 * normemb(j, :))) * Z(:, n)$ 
27: end for
28:  $num \leftarrow []$ 
29: for  $j = 1$  to  $c$  do
30:    $num(j, :) \leftarrow (tkout(j, 1) * t + tkout(j, 2) \cdot^* y(j, :)^T)$  //  $tkout(j, k)$  é a entrada  $k$  do
      arranjo  $tkout$ 
31: end for
32:  $defu \leftarrow soma(num) \cdot /soma(y)$ 

```

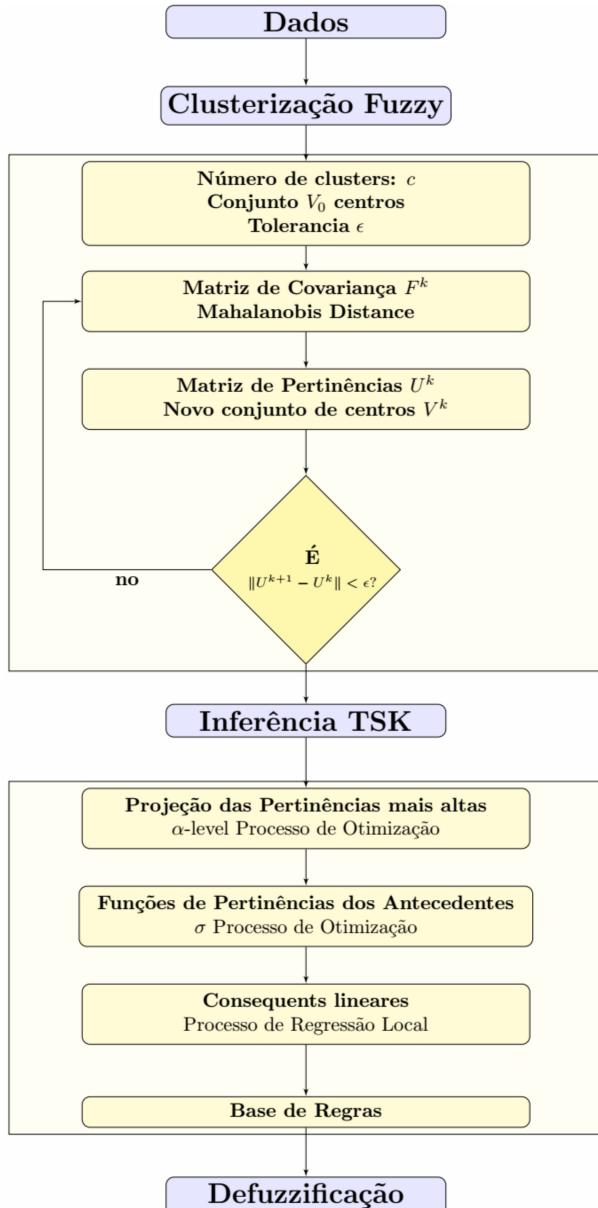


Figura 8 – Fluxograma do processo de identificação fuzzy de sistemas proposto.

Fonte: Autores, 2020.

Como etapa intermediária entre o desfecho da parte teórica da nova metodologia e validação da mesma, houve uma análise cuidadosa dos softwares computacionais que iriam ser mais eficientes na implementação da mesma. A investigação resultou na formulação de um código próprio que é descrita na Seção 4.2, continuação do assunto abordado.

4.2 A Busca de um Código Próprio

Nos testes de identificação com uma célula de combustível a hidrogênio (AVELAR, 2012), surgiu a necessidade de aprofundamento no processo de clusterização que apresenta como o cerne: o algoritmo de Gustafson-Kessel. A sua capacidade foi testada e apresentada

através do trabalho apresentado no IV Congresso Brasileiro de Sistemas Fuzzy 2016 (CBSF) (BERTONE; CUNHA; MARTINS, 2016). A principal contribuição desse estudo foi entender a importância da obtenção de gráficos de qualidade para o entendimento da geometria dos clusters e seus distintos níveis de pertinência.

Como apresentado nesse trabalho, várias plataformas foram testadas e comparadas; alguns testes foram realizados através da linguagem de programação Python, a qual permite certas facilidades devido a sua grande quantidade de bibliotecas. Além disso, a qualidade gráfica das figuras foi um diferencial como apresentado no trabalho apresentado no IV CBSF-2016 (artigo exposto na sua totalidade no Apêndice A).

Outrossim, salienta-se que nesse período foi feito o desenvolvimento completo de um código de clusterização, no qual não foram utilizados toolboxes ou quaisquer elementos que possam “encapsular” e obscurecer o procedimento. Após os testes com as linguagens Python (BANIN, 2018), C (SCHILDIT, 1997), Octave (OCTAVE, 2016) e Matlab (MATLAB, 2016), optou-se pelo Matlab.

Em busca de aplicar esta nova metodologia foram escolhidos os tipos de problemas a serem analisados por esta técnica, cujas características são elencadas na Seção 4.3.

4.3 Família de Modelos Apropriados para a Abordagem Proposta

Os modelos apropriados para a abordagem proposta são:

□ Gerais

1. A saída do modelo depende de vários “atributos” $\{x_i\}_{i=1}^{i=n-1} n \geq 2$. Atributos são elementos que identificam a dinâmica em determinado instante, ou que contribui para o valor de saída do modelo. Por exemplo, para um modelo de energia, a temperatura, a umidade ambiente, podem ser considerados como atributos.
2. A saída responde com um comportamento “similar” frente à mesma combinação de atributos; ou seja, a saída pode ser descrita por uma função explícita contínua f :

$$saída_k = f(x_{k1}, \dots, x_{k(n-1)}).$$

□ Particulares:

1. Série temporal, entrada temporal acrescida de atributos ou SISO: o modelo depende de uma variável de entrada t , escolhida uma medida de tempo (minutos, horas, dias). Nesse caso o indexador $i = t$ é o eixo de projeção. Ou uma única entrada que é o eixo de projeção.

2. Não sendo SISO ou controlado pelo tempo, a projeção dos valores superiores de pertinência aos cluster é considerada sobre os valores de um dos atributos. Por exemplo, daquele atributo que funcione como controlador do modelo.

Modelos de engenharia que apresentam essas propriedades são bem reconhecidos pela função f . Caso contrário, falhas evidentes devem ocorrer na validação.

Uma vez achada a família de problemas, estabelecida uma codificação confiável e um software mais apropriado para os objetivos da metodologia, foram feitas inúmeras simulações com o intuito de validá-la. O Capítulo 5 apresenta e explica algumas das simulações efetuadas.

CAPÍTULO 5

Materiais e métodos

5.1 Estudo Comparativo de Validação

As simulações começaram com a célula de combustível a hidrogênio (AVELAR, 2012) que deram as primeiras dicas da acurácia e menor esforço computacional da nova metodologia; principal contribuição deste estudo. Estes resultados bem sucedidos incentivaram a buscar, em benchmarks mundiais (LEUVEN, 2017), novos conjuntos de dados, entre os quais foram escolhidos dados do tipo MISO, alguns com fortes características de “linearidade por partes” e outros, com muito “ruído”.

Os testes foram feitos através da rotina Compare do Matlab, a qual recebe dois grupos por parâmetro: o modelo e os dados medidos. Esta ferramenta simula o resultado de um arquétipo de sistema dinâmico e sobrepõe essa resposta aos dados medidos. Tal fato torna a rotina Compare um referencial em testes para que uma nova técnica seja avaliada, pois uma comparação é feita com outras técnicas consagradas. O gráfico também apresenta a medida raiz média quadrática do erro normalizada (normalized root mean square error - NRMSE, sua sigla em inglês) entre a resposta simulada e os dados medidos. Esta rotina deve ser utilizada quando se deseja validar um modelo selecionado, sendo possível empregar a comparação com modelos e dados no domínio do tempo ou no domínio da frequência. A citada ferramenta também pode ser usada para previsões, de forma que as entradas são usadas para projetar a resposta futura. A Equação do NRMSE é apresentada na sequência, na qual *dados* é um elemento de Z , *aprox* é o resultado da identificação, *média* é a média aritmética dos dados amostrados e $\|\cdot\|$ é a norma Euclidiana do espaço real.

$$NRMSE = 100 \cdot \left(1 - \frac{\|dados - aprox\|^2}{\|dados - média(dados)\|^2} \right) \quad (22)$$

As técnicas Nonlinear Auto-Regressive with eXogenous inputs (VILLARREAL-LÓPEZ, 2019), Hammerstein-Wiener (FERREIRA, 2017) e Multilayer Perceptron (CASTRO; BRAGA, 2013) foram utilizadas para comparação, pois estão fortemente presentes na literatura. A primeira citada, também conhecida como NARX, é uma representação

caixa-preta não-linear, largamente utilizada para identificação de sistemas (VILLARREAL-LÓPEZ, 2019). Estruturas dinâmicas de modelos não-lineares diferem entre si a partir do vetor de regressores, $\varphi(k)$, empregado; os modelos NARX se apresentam conforme:

$$\begin{aligned}\varphi(k) &= [\varphi_1(k)\varphi_2(k)\cdots\varphi_\epsilon(k)]^T \\ &= [y(k-1)\cdots y(k-n_y) \quad u(k-\tau_d-1)\cdots u(k-\tau_d-n_u)]^T\end{aligned}$$

onde $u(k-i)$ e $y(k-j)$ representam a entrada e a saída medidas no processo nos instantes amostrados $k-i$ e $k-j$; ϵ é o número de regressores utilizados no modelo $\epsilon = n_y + n_u$ e τ_d representa o tempo morto.

O modelo de Hammerstein-Wiener é uma estrutura de aproximação de sistemas com atuadores não-lineares, ou que possuam outros efeitos não-lineares, que podem ocorrer à entrada do sistema (FERREIRA, 2017); a Figura 9 apresenta a sua estrutura. O sinal de entrada, u_k , é aplicado ao bloco não-linear $F(\cdot)$; em seguida obtém-se um sinal intermediário, v_k , que é u_k alterado pela não-linearidade estática. Por fim, o sinal passa pelo bloco G_q , com um modelo dinâmico linear e é alcançada a saída y_k .

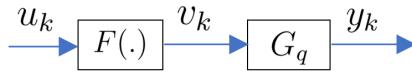


Figura 9 – Estrutura geral de um modelo Hammerstein-Wiener.

Fonte: Ferreira, 2017.

O sistema de equações que descreve este modelo é:

$$\begin{aligned}v(k) &= f(u(k)) \\ y(k) &= G(q)v(k)\end{aligned}$$

O Multilayer Perceptron (MLP) é uma das vertentes das Redes Neurais Artificiais (RNA), as quais têm sido muito utilizadas na identificação de sistemas dinâmicos nos últimos 30 anos (CASTRO; BRAGA, 2013; MANZAN, 2016). As Redes Neurais Artificiais foram criadas, a priori, para aproximação funcional e reconhecimento de padrões. Com a evolução desta técnica foram surgindo novas aplicações, sendo as suas duas principais características: a arquitetura e o algoritmo de treinamento (PAIVA, 1999). A principal peculiaridade das redes MLP é a presença de uma ou várias camadas intermediárias de neurônios, localizadas entre as entradas e as saídas (SILVA; SPATTI; FLAUZINO, 2010). Além das duas aplicações já citadas, as redes Perceptron de múltiplas camadas são utilizadas para:

- ❑ Identificação e controle de processos.
- ❑ Previsão de séries temporais.

□ Otimização de sistemas.

As redes MLP são treinadas de maneira supervisionadas e possuem arquitetura feed-forward, ou seja, os dados são inseridos na camada de entrada e propagam em direção à(s) camada(s) escondida(s) e saída. A Figura 10 apresenta o diagrama geral de uma Rede Neural Artificial, na qual X_1, X_2 e X_n representam as entradas; F_1 e F_m representam a camada intermediária, também conhecida como escondida, e $Y(k)$ a saída.

Os detalhes referentes aos testes efetuados são apresentados no Capítulo 6, Resultados Obtidos, o qual discute as peculiaridades de cada base de dados e de cada técnica empregada.

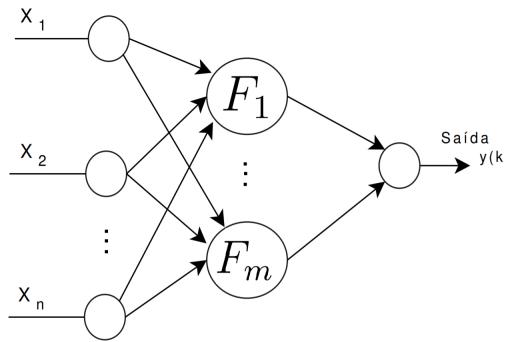


Figura 10 – Diagrama geral de uma Rede Neural Artificial.

Fonte: Autores, 2020.

CAPÍTULO 6

Resultados Obtidos

As comparações efetuadas revelam que a técnica proposta apresenta a melhor medida FIT em todos os testes apresentados. Além disso, o esforço computacional, medido em tempo de processamento, nem sempre é o menor, entretanto é o mais viável. Há métodos que apresentaram tempo de processamento inferior em alguns testes, entretanto o seu FIT também foi mais baixo.

Dito isso, são apresentadas as simulações selecionadas para esta seção, cada uma com um nome simbólico para a base de dados. Os testes de redes neurais artificiais foram feitos com o auxílio do professor José Ricardo Gonçalvez Manzan, especialista no assunto. Salienta-se que o tempo medido foi obtido na etapa de construção do modelo (treinamento).

Simulação “Thermic”: base de dados disponibilizada por Favoreel K. U. Leuven, departamento de engenharia elétrica da ESAT/SISTA na Bélgica. Os dados representam a densidade do fluxo de calor através de uma parede. As entradas são o instante e a temperatura interna da parede. A saída é o fluxo de calor através da parede. São 1680 dados (MISO) com características ruidosas, o que foi a motivação para o emprego neste trabalho; fato que pode ser confirmado através da Figura 11.

A Tabela 3 compara os desempenhos da identificação pelas diferentes metodologias e o tempo de execução da identificação. O maior FIT foi obtido através do emprego da metodologia proposta neste trabalho, 95,53%.

Tabela 3 – Tabela comparativa do desempenho dos dados *Thermic*.

Metodologia	FIT	tempo de execução (em segundos)	Validação 20%
Identificação fuzzy	95,53%	12,224	
NARX	45,58%	1,234	
Redes Neurais Artificiais	17,21%	175,766	
Hammerstein-Wiener	15,5%	5,631	

A Figura 11 apresenta o resultado comparativo da identificação. A metodologia

proposta neste trabalho é identificada como *defudata*, Narx como *narxdata*, Multilayer Percetron como *nnetdata* e Hammerstein-Wiener como *hammweidata*. O modelo *nnetdata* foi gerado a partir dos seguintes parâmetros:

- 80% dos dados foram usados para treinamento e 20% para validação.
- As entradas foram normalizadas.
- A rotina empregada foi *newff*, a qual utilizou uma função de ativação tangente hiperbólica com gradiente descendente (SILVA; SPATTI; FLAUZINO, 2010), momentum (0.9) e aprendizado adaptativo.
- Taxa de aprendizagem igual a 0.01.
- O erro máximo tolerado foi igual a 0.001.

A função *idnlarx* foi empregada para os testes referentes à técnica Narx, a qual cria um modelo ARX não-linear junto ao seu vetor de regressores. A função *idnlhw* foi empregada para os testes referentes à técnica Hammerstein-Wiener, a qual cria um modelo HW com as ordens especificadas (Nb, Nf e Nk) e usa funções lineares por partes como estimadores de não-linearidade de entrada e saída. Onde Nb é o número de zeros mais 1, Nf é o número de polos e Nk é o atraso.

Simulação “Exchanger”: base de dados disponibilizada por Sergio Bittanti, do Politecnico di Milano, Dipartimento di Elettronica e Informazione. O processo é formado por um trocador de calor de vapor saturado de líquido, em que a água é aquecida por vapor saturado pressurizado através de um tubo de cobre. A variável de saída é a temperatura do líquido de saída, sendo 4000 dados (MISO) com características ruidosas, o que foi a motivação para o emprego neste trabalho; fato que pode ser confirmado através da Figura 12.

A Tabela 4 compara os desempenhos da identificação pelas diferentes metodologias e o tempo de execução da identificação. O maior FIT foi obtido através do emprego da metodologia proposta neste trabalho, 80,61%.

Tabela 4 – Tabela comparativa do desempenho dos dados *Exchanger*.

Metodologia	FIT	tempo de execução (em segundos)	Validação 15%
Identificação fuzzy	80,61%	5,364	
Redes Neurais Artificiais	45,7%	48,82	
Hammerstein-Wiener	30,7%	5,553	
NARX	26,46%	1,206	

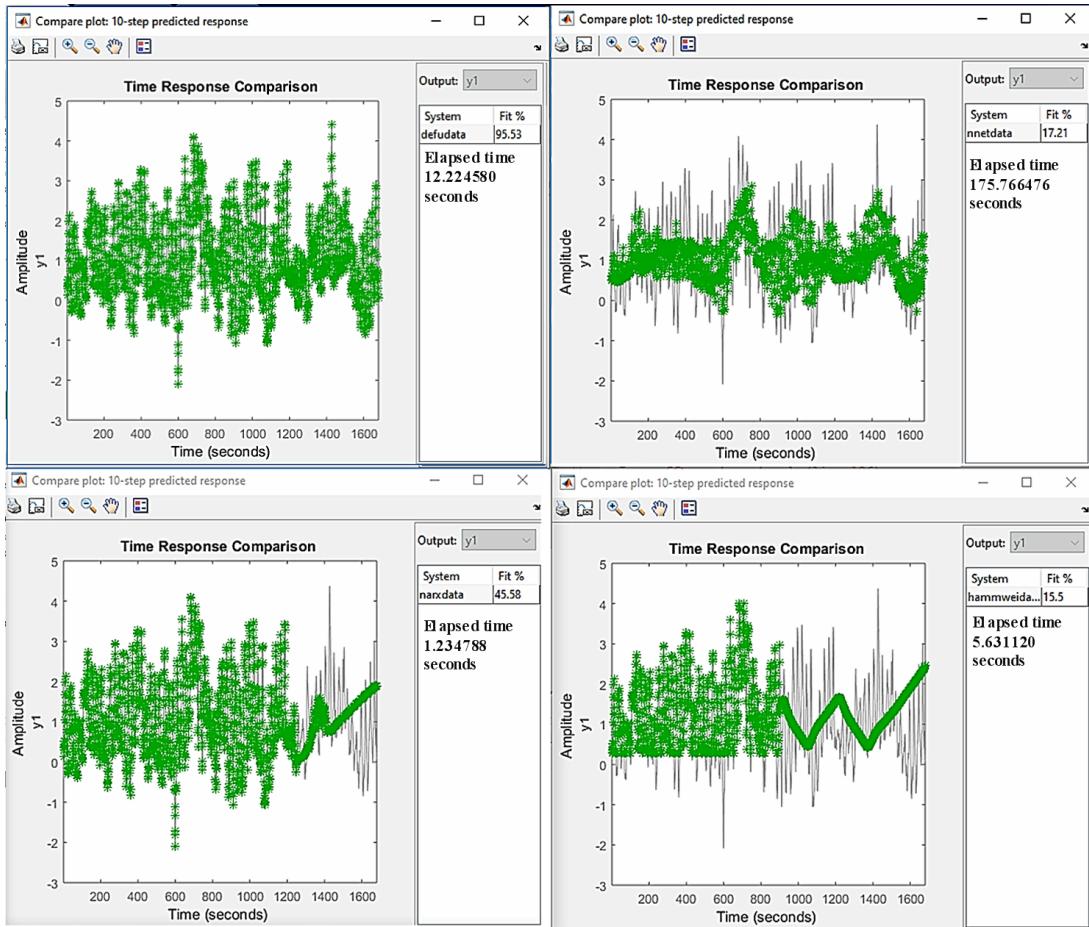


Figura 11 – Comparação do desempenho da identificação dos dados *Thermic*.

Fonte: Autores, 2020.

A Figura 12 apresenta o resultado comparativo da identificação. A metodologia proposta neste trabalho é identificada como *defudata*, Narx como *narxdata*, Multilayer Percetron como *nnetdata* e Hammerstein-Wiener como *hammweida*.

O modelo *nnetdata* foi gerado a partir dos seguintes parâmetros:

- 85% dos dados foram usados para treinamento e 15% para validação.
- As entradas foram normalizadas.
- A rotina empregada foi *newff*, a qual utilizou uma função de ativação tangente hiperbólica com gradiente descendente, momentum (0.9) e aprendizado adaptativo.
- Taxa de aprendizagem igual a 0.01.
- O erro máximo tolerado foi igual a 0.00087.

A função *idnlarx* foi empregada para os testes referentes à técnica Narx, a qual cria um modelo ARX não-linear junto ao seu vetor de regressores. A função *idnlhw* foi empregada para os testes referentes à técnica Hammerstein-Wiener, a qual cria um modelo HW com

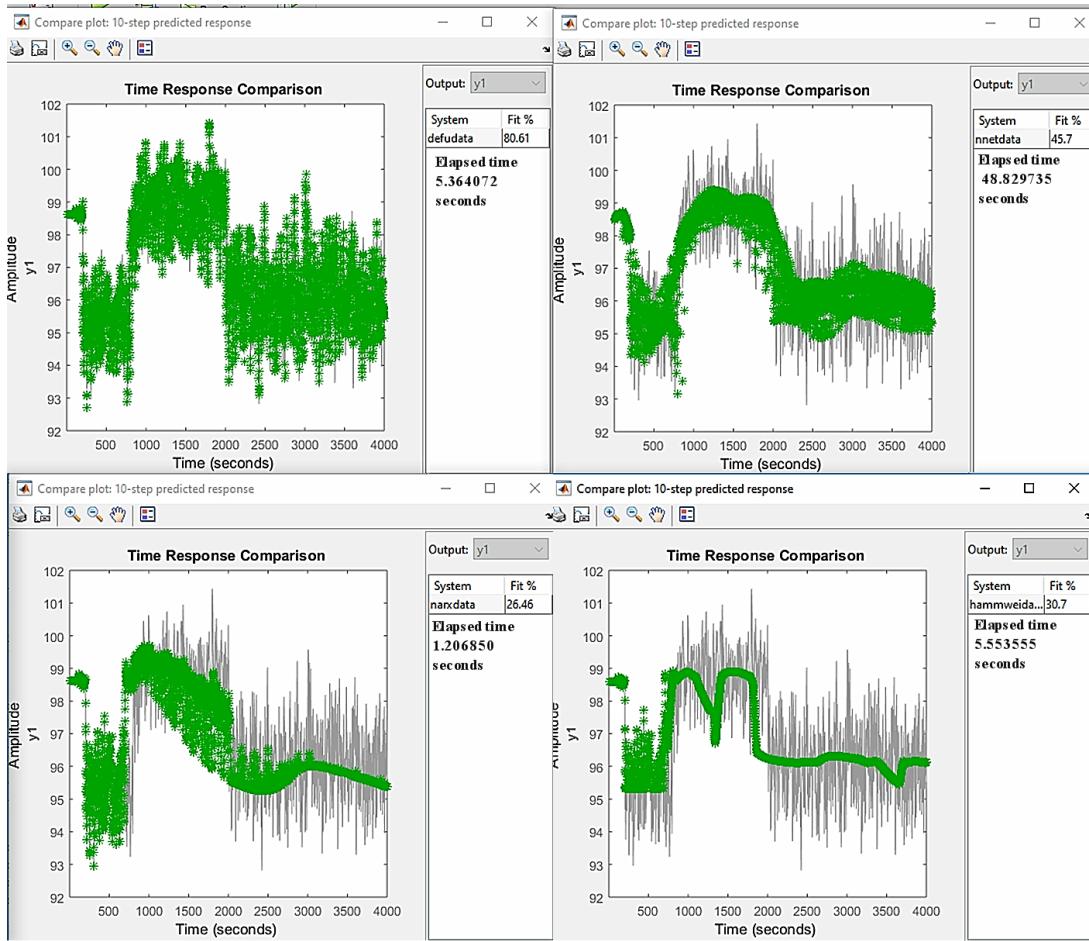


Figura 12 – Comparaçāo do desempenho da identificaçāo dos dados *Exchanger*.

Fonte: Autores, 2020.

as ordens especificadas (N_b , N_f e N_k) e usa funções lineares por partes como estimadores de não-linearidade de entrada e saída.

Simulação “Glassfurnace”: base de dados disponibilizada por Peter Van Overschee, K. U. Leuven ESAT - SISTA. Os dados foram retirados de um forno de vidro (Philips) com 1247 amostras. O sistema é (MISO), em que as entradas representam o instante e aquecimento e a saída, a temperatura medida pelo sensor. Esta base de dados possui características ruidosas, o que foi a motivação para o emprego neste trabalho; fato que pode ser confirmado através da Figura 13.

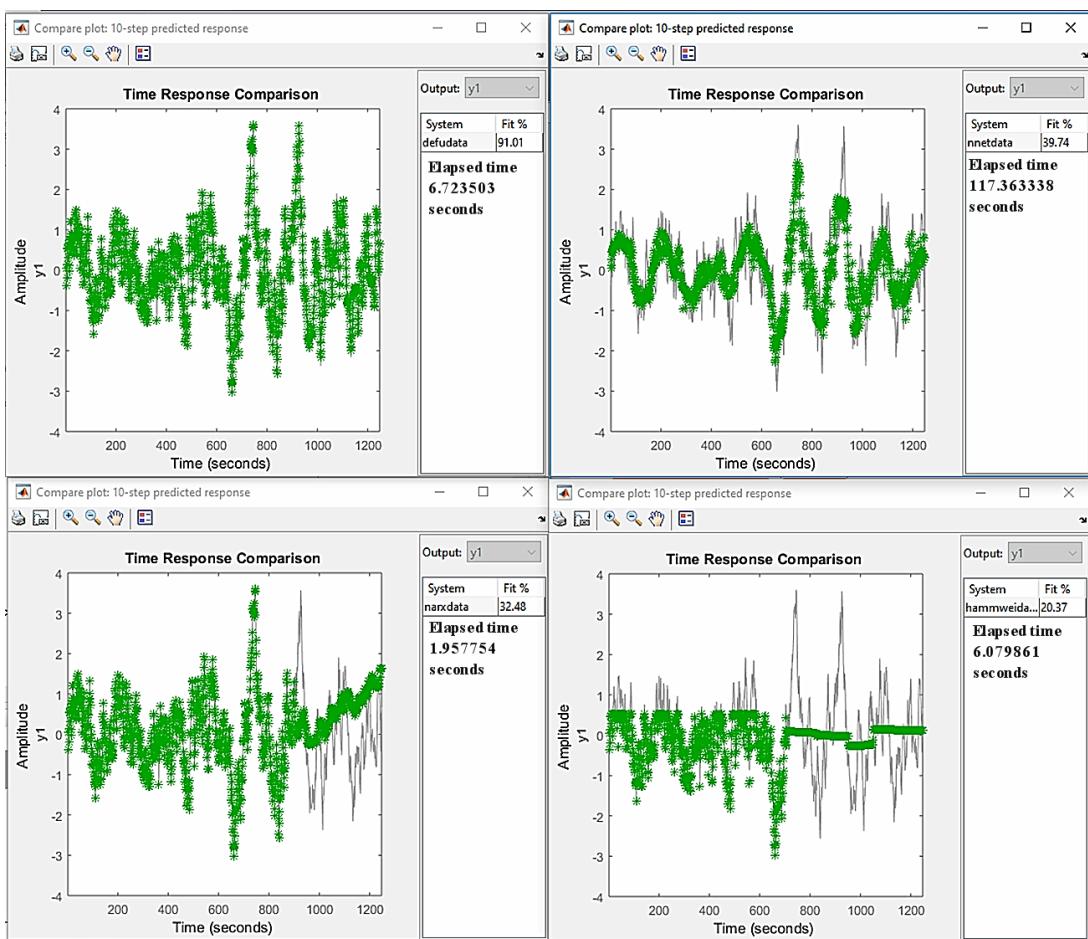
A Tabela 5 compara os desempenhos da identificação pelas diferentes metodologias e o tempo de execução da identificação. O maior FIT foi obtido através do emprego da metodologia proposta neste trabalho, 91,01%.

A Figura 13 apresenta o resultado comparativo da identificação. A metodologia proposta neste trabalho é identificada como *defudata*, Narx como *narxdata*, Multilayer Percetron como *nnetdata* e Hammerstein-Wiener como *hammweidata*.

No que diz respeito à comparação, o modelo *nnetdata* foi gerado a partir dos seguintes parâmetros:

Tabela 5 – Tabela comparativa do desempenho dos dados *Glassfurnace*.

Metodologia	FIT	tempo de execução (em segundos)	Validação 20%
Identificação fuzzy	91,01%	6,723	
Redes Neurais Artificiais	39,74%	117,363	
NARX	32,48%	1,957	
Hammerstein-Wiener	20,37%	6,079	

Figura 13 – Comparaçāo do desempenho da identificaçāo dos dados *Glassfurnace*.

Fonte: Autores, 2020.

- 80% dos dados foram usados para treinamento e 20% para validação.
- As entradas foram normalizadas.
- A rotina empregada foi *newff*, a qual utilizou uma função de ativação tangente hiperbólica com gradiente descendente, momentum (0.9) e aprendizado adaptativo.
- Taxa de aprendizagem igual a 0.001.
- O erro máximo tolerado foi igual a 0.000364.

A função *idnlarx* foi empregada para os testes referentes à técnica Narx, a qual cria um modelo ARX não-linear junto ao seu vetor de regressores. A função *idnlhw* foi empregada para os testes referentes à técnica Hammerstein-Wiener, a qual cria um modelo HW com as ordens especificadas (Nb, Nf e Nk) e usa funções lineares por partes como estimadores de não-linearidade de entrada e saída.

Simulação “Hidiga”: base de dados cedida pela Usina Hidrelétrica de Igarapava (MISO).

A Tabela 6 compara os desempenhos da identificação pelas diferentes metodologias e o tempo de execução da identificação. O maior FIT foi obtido através do emprego da metodologia proposta neste trabalho, 96,81%.

Tabela 6 – Tabela comparativa do desempenho dos dados *Hidiga*.

Metodologia	FIT	tempo de execução (em segundos)	Validação 15%
Identificação fuzzy	96,81%	5,330	
Redes Neurais Artificiais	82,32%	192,735	
NARX	36,12%	1,227	
Hammerstein-Wiener	32,39%	1,827	

A Figura 14 apresenta o resultado comparativo da identificação. A metodologia proposta neste trabalho é identificada como *defudata*, Narx como *narxdata*, Multilayer Percetron como *nnetdata* e Hammerstein-Wiener como *hammweidata*.

No que diz respeito à comparação, o modelo *nnetdata* foi gerado a partir dos seguintes parâmetros:

- ❑ 85% dos dados foram usados para treinamento e 15% para validação.
- ❑ As entradas foram normalizadas.
- ❑ A rotina empregada foi *newff*, a qual utilizou uma função de ativação tangente hiperbólica com gradiente descendente, momentum (0.9) e aprendizado adaptativo.
- ❑ Taxa de aprendizagem igual a 0.001.
- ❑ O erro máximo tolerado foi igual a 0.0001.

A função *idnlarx* foi empregada para os testes referentes à técnica Narx, a qual cria um modelo ARX não-linear junto ao seu vetor de regressores. A função *idnlhw* foi empregada para os testes referentes à técnica Hammerstein-Wiener, a qual cria um modelo HW com as ordens especificadas (Nb, Nf e Nk) e usa funções lineares por partes como estimadores de não-linearidade de entrada e saída.

Os dados coletados na Usina Hidrelétrica de Igarapava-SP apresentam características altamente desafiadoras. O levantamento foi feito através de um visita técnica “in loco”,

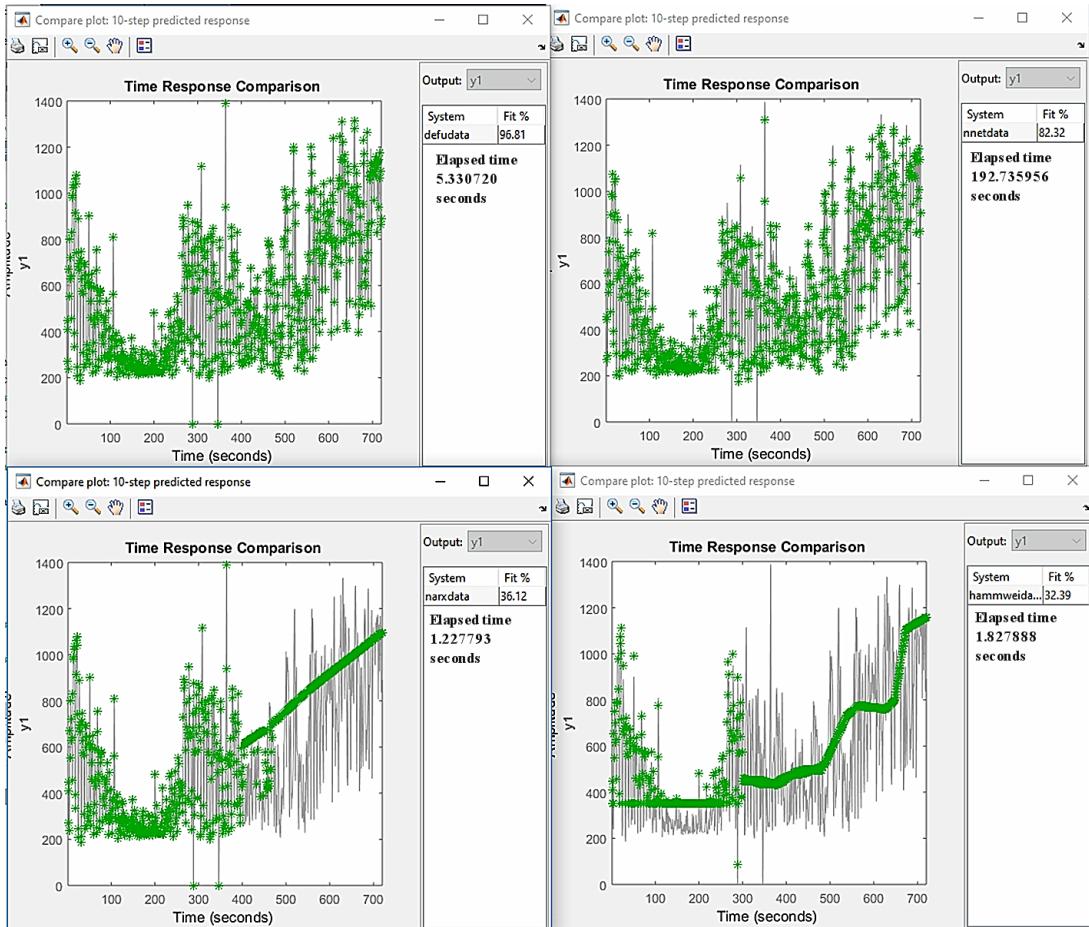


Figura 14 – Comparaçāo do desempenho da identificaçāo dos dados *Hidiga*.

Fonte: Autores, 2020.

onde todos os atributos envolvidos foram explicados detalhadamente pelo engenheiro eletricista responsável.

Simulações “Hydrocell”: Esta base de dados foi gentilmente cedida por Avelar (AVELAR, 2012). São 38300 dados (MISO) com característica “linear por partes”, o que foi a motivação para o emprego neste trabalho. O funcionamento de uma célula a combustível de hidrogênio com um conversor elevador *full-bridge* e um inversor monofásico PWM senoidal é apresentado neste trabalho que, além de criar um software, simula o sistema real. O modelo proposto foi desenvolvido para uma célula de 1,2 KW e o processo de identificação de sistemas utilizado é dos mínimos quadrados. O supracitado trabalho (AVELAR, 2012) desenvolveu um modelo para simular a célula a combustível de hidrogênio ligada à rede elétrica, levando em consideração o efeito da temperatura gerado internamente. As principais características do software simulador são:

- ❑ A célula é capaz de gerar tensão entre 20 V e 50 V (contínua).
- ❑ A tensão varia com a corrente drenada em seus terminais, sendo elevada a 380V antes de ser aplicada no módulo inversor.

- ❑ Todos os dados são armazenados em um microcomputador.
- ❑ As características da planta são diferentes para subida e descida de temperatura, pois o aquecimento é mais rápido que o resfriamento.
- ❑ A célula a combustível de hidrogênio de membrana polimérica (PEMFC Nexa, fabricada pela Ballard) é testada em laboratório para levantamento estático e dinâmico para levantamento do seu modelo.
- ❑ Toda a estrutura da célula a combustível é simulada através do software PSIM.
- ❑ O modelo encontrado representa a dinâmica da planta (corrente-temperatura).

Esta base de dados foi a primeira a ser utilizada nos testes. Sua característica não é essencialmente desafiadora, entretanto foi importante para a maturação da técnica inovadora proposta neste trabalho.

No que diz respeito à comparação, o modelo *nnetdata* foi gerado a partir dos seguintes parâmetros:

- ❑ 70% dos dados foram usados para treinamento e 30% para validação.
- ❑ As entradas foram normalizadas.
- ❑ A rotina empregada foi *newff*, a qual utilizou uma função de ativação tangente hiperbólica com gradiente descendente, momentum (0.9) e aprendizado adaptativo.
- ❑ Taxa de aprendizagem igual a 0.001.
- ❑ O erro máximo tolerado foi igual a 0.0001.

A função *idnlarx* foi empregada para os testes referentes à técnica Narx, a qual cria um modelo ARX não-linear junto ao seu vetor de regressores. A função *idnlhw* foi empregada para os testes referentes à técnica Hammerstein-Wiener, a qual cria um modelo HW com as ordens especificadas (Nb, Nf e Nk) e usa funções lineares por partes como estimadores de não-linearidade de entrada e saída.

Simulações “energy”: base de dados que retrata uma série temporal na qual a entrada representa o instante em meses; a saída representa a energia gerada na região sudeste e centro-oeste do Brasil de 2000 a 2016. São 204 valores medidos em GWh e disponibilizados no site do Operador Nacional do Sistema Elétrico (ELéTRICO, 2015). Estes dados foram utilizados para testar a capacidade da técnica proposta neste estudo frente a séries temporais.

Salienta-se que a maturidade atingida no desenvolvimento da experimentação que deu como fruto o trabalho (BERTONE; MARTINS, 2017b), assim como os estudos comparativos com medida de acurácia da rotina Compare, foram decisivos para a concepção

e melhoramento do processo de predição posteriormente desenvolvido. A citada predição foi a principal aplicação resultante do novo método de identificação de sistemas proposto neste estudo, sendo explicada na Seção 6.1.

6.1 Predição Fuzzy de Sistemas MISO

A predição fuzzy de sistemas de várias entradas e uma saída foi uma vertente que surgiu e se consolidou durante o desenvolvimento do presente trabalho. Novamente, uma família adequada de processos preditivos pode ser manejada através da técnica de identificação de sistemas proposta. De fato, os fenômenos a serem previstos estão condicionados à presença de um comportamento repetido ou similar em resposta a certos atributos que produzem o valor de saída. Fenômenos aleatórios podem não ser bons candidatos para a aplicação do método.

Para que a predição ocorra é necessário realizar a identificação fuzzy de dados Z , do passado, do fenômeno a ser predito, ou seja, recolher as observações até o instante t anterior à predição. Nessa pré-análise é obtida uma função explícita contínua f que relaciona o instante t e os atributos com a saída correspondente, de forma que

$$\text{saída}(t) = f(t, \text{atributos}(t)).$$

No instante $t + 1$, o conjunto de dados Z é atualizado com um novo dado coletado z_{t+1} . Assim, é obtido o conjunto $Z = [Z, z_{t+1}]$ para o qual se repete a dinâmica da predição. Uma "janela" é construída para que apenas parte dos dados sejam utilizados na criação do modelo. Dados "muito antigos" podem piorar a qualidade do processo; a cada iteração a "janela" caminha e mantém somente dados mais recentes para a construção do arquétipo.

Um pseudocódigo do algoritmo de predição é apresentado no Algoritmo 3:

A base de dados energy (ELéTRICO, 2015) tem sido utilizada para validar a predição como aplicação da metodologia. Para efetuar a simulação, dividiu-se o conjunto amostrado em duas categorias: uma que representa o passado, P , e outra que seria predita, F . Para que a dinâmica seja assimilada, F é utilizada. O conjunto de dados P constitui os primeiros 104 valores do conjunto *energy*. O processo de identificação de sistemas ocorre no intervalo $[t_1, t_P]$, no qual t_1 é o primeiro mês coletado e t_P é o último mês coletado. A predição do instante t_{P+1} ocorre segundo descrito no algoritmo 3. A medida de acurácia utilizada para esta simulação é dada por

$$P = 1 - \left| \frac{\text{valor real} - \text{valor predito}}{\text{valor real}} \right|$$

A Tabela 6.1 apresenta a acurácia para trinta meses de predição simuladas.

Algoritmo 3 Predição

```

1:  $Z \leftarrow energy$  // dados provenientes de informações do passado
2:  $n \leftarrow \text{colunas}(Z)$ 
Require:  $p$ 
3: while  $k \leq p$  do
4:    $N \leftarrow \text{linhas}(Z)$ 
Require:  $Clusterização(Z)$ ,  $Identificação(Z)$  // sub-rotina como no Algoritmo 2
5:    $t \leftarrow [1, 2, \dots, i]$ 
6:    $num \leftarrow []$ 
7:   for  $j = 1$  to  $c$  do
8:      $num(j, :) \leftarrow (tkout(j, 1) * t^T + tkout(j, 2) * y(j, :)^T$ 
9:   end for
10:   $defu \leftarrow \text{soma}(num) / \text{soma}(y)$  // fórmula explícita contínua da predição
11:   $predito \leftarrow \text{soma}(num) / \text{soma}(y(:, N + 1))$ 
Require:  $z_{N+1}$  // observação atualizada do experimento
12:   $Z \leftarrow [Z, z_{N+1}]$ 
13:   $k \leftarrow k + 1$ 
14: end while

```

Resultado da acurácia (P)
para 30 meses de predição simulada

Mês	P	Mês	P
1	0,95	2	0,9
3	0,95	4	0,99
5	0,97	6	0,97
7	0,95	8	0,97
9	0,96	10	0,99
11	0,96	12	0,97
13	0,96	14	0,86
15	0,97	16	0,99
17	0,96	18	0,97
19	0,95	20	0,98
21	0,93	22	0,98
23	0,96	24	0,98
25	0,96	26	0,92
27	0,99	28	0,99
29	0,99	30	0,95

A Figura 15 apresenta o resultado de 30 meses de predição. Os pontos azuis são os 30 valores reais dos meses e os pontos em cor magenta são os valores previstos. As bandas verdes representam os limites de 0,90 da medida m , o FIT total alcançado foi 0,99081.

As bases *energy* e *hydrocell* são também descritas detalhadamente no Apêndice B, o qual apresenta o principal trabalho publicado.

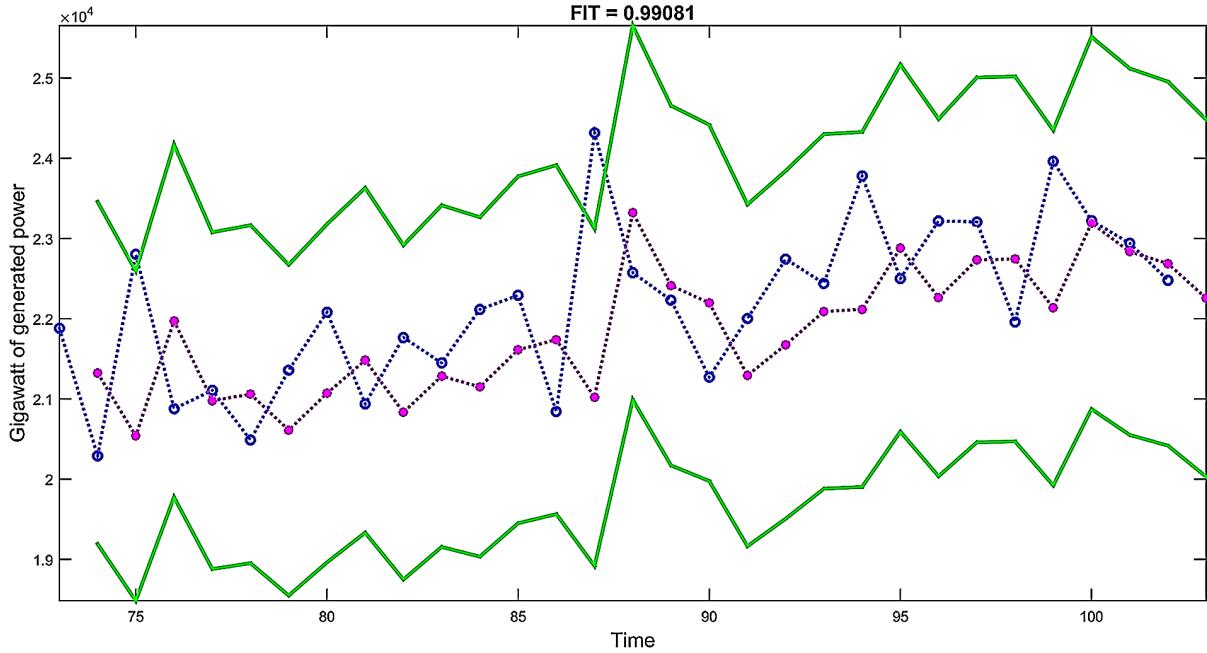


Figura 15 – Gráfico de trinta meses de predição da base de dados de energy.

Fonte: Autores, 2020.

As principais análises levantadas são:

- ❑ Thermic é uma base de dados de complexa modelagem e a técnica proposta neste trabalho alcançou FIT igual a 95,53% em 12,22 segundos, o melhor desempenho. A técnica nnetdata, a qual se saiu melhor nos demais testes, apresentou FIT igual a 17,21% e tempo necessário para gerar o modelo igual a 175,76 segundos. A técnica HW completou sua tarefa em apenas 5,63 segundos, entretanto seu FIT é igual a 5,63%, baixo. Para esta base de dados, a técnica naxdata (FIT igual a 45,58% e tempo de processamento igual a 1,23 segundos) apresentou resultado melhor que HW e nnetdata.
- ❑ Exchanger é uma base de dados “ruidosa” e a técnica proposta neste trabalho alcançou FIT 80,61% em 5,36 segundos, o melhor desempenho. Todas as outras técnicas apresentaram FIT abaixo de 46%.
- ❑ Glasfurnace é uma base de dados de difícil modelagem e a técnica proposta neste trabalho alcançou FIT 91,01% em 6,72 segundos, o melhor desempenho. Todas as outras técnicas apresentaram FIT abaixo de 40%, além disso, a técnica nnetdata necessitou 117,36 segundos para completar sua execução.
- ❑ Hidiga é um base de dados com muitos atributos (entradas), o que provoca grande esforço computacional para alcançar um modelo. A técnica com menor tempo de processamento foi narxdata, entretanto seu FIT foi baixo (36,12%). A técnica

nnetdata apresentou um FIT alto, 82,32%, entretanto seu tempo de processamento foi alto 192,73 segundos. A técnica proposta neste trabalho alcançou o FIT 96,81% em 5,33 segundos, o melhor desempenho.

Por fim, o Capítulo 7, Conclusão, apresenta os entraves, os desafios superados e as propostas para trabalhos futuros.

CAPÍTULO 7

Conclusão

O caminho percorrido permitiu a aquisição de salutar experiência. O seu desenvolvimento cronológico destaca o primeiro artigo (BERTONE et al., 2016) que mostrou o quanto a identificação fuzzy de sistemas é robusta para lidar com sistemas não-lineares e do tipo black-box. O software Fuzzy Modeling and Identification Toolbox de (BABUSKA et al., 2018) foi essencial para o início da pesquisa. Porém, uma averiguação bibliográfica cuidadosa, proporcionou ideias norteadoras que auxiliaram e influenciaram a escolha de ferramentas mais adequadas. No trabalho (BERTONE; CUNHA; MARTINS, 2016) foram obtidas conclusões que explanaram, especificamente, o processo de clusterização através do algoritmo de Gustafson-Kessel. A clusterização é parte importante do procedimento de identificação de sistemas fuzzy e, consequentemente, da predição, campo em que se pretendia aplicar a nova metodologia. Passos relevantes no esclarecimento teórico e tomadas de decisões na área computacional foram dados durante o período de desenvolvimento deste estudo.

Em (BERTONE; MARTINS, 2017a) foi utilizada a simulação de uma célula de combustível a hidrogênio, em termos de identificação e validação do tipo k -fold; estudo que fez com que as dificuldades da abordagem fuzzy se tornassem claras e, principalmente, a falta de documentação adequada para o desenvolvimento da mesma. Nesta etapa efetuou-se a entrega de uma fórmula explícita contínua e foi obtida a compreensão do processo de clusterização sob o ponto de vista geométrico.

O artigo (MARTINS; BERTONE, 2018) apresenta uma aplicação do método estudado para a clusterização de sinais de eletroencefalograma (EEG); passo inicial, mas satisfatório em uma área ligada ao estudo de várias doenças como: epilepsia, distúrbios do sono, tumores cerebrais, entre outros. Através deste trabalho, aplicado a séries temporais, notou-se que com uma clusterização de poucos clusters, a curva obtida continha uma informação valiosa para uma pré-análise do EEG, elemento crucial para estudos desta natureza.

O sucesso na identificação de uma série temporal incentivou a aplicação da metodologia a benchmarks. Foi nessa etapa do estudo que entendeu-se que a técnica poderia ser

usada com promissores resultados para sistemas de múltiplas entradas e uma saída e, como consequência, para predição de dados on-line, âmbito novo e de grande impacto nas aplicações de engenharia. Neste propósito, buscou-se a cessão de dados da Usina Hidrelétrica de Igarapava-SP, os quais foram testados com o objetivo de predizer a geração diária de energia. Os resultados foram avaliados em termos de acerto estatístico e aplicados para outros dados provenientes de benchmarks do consumo de energia brasileiro.

Dessa fase surge a inovadora técnica de identificação fuzzy com dois aspectos relevantes: baixo esforço computacional e considerável taxa de acerto a partir de medidas exigentes de acurácia. Como consequência da teoria desenvolvida, o código foi feito sem o uso de qualquer toolbox. A aplicação mais aprofundada desta nova metodologia foi na área de predição on-line, campo de grande atividade em todo o mundo.

Por fim, foi aceito o artigo (BERTONE; YAMANAKA; MARTINS, 2019) que apresenta a técnica de identificação fuzzy de sistemas de maneira completa e aplicada à predição; o que abre possibilidades para diferentes trabalhos futuros. Diante disso, o objetivo da pesquisa foi alcançado: apresentar um novo método fuzzy para identificação de sistemas com várias entradas e uma saída.

7.1 Contribuições

A principal contribuição desta pesquisa é a apresentação de um novo método de identificação de sistemas com comprovada eficiência, tanto na precisão, quanto no baixo esforço computacional. Para checar estas vantagens foram feitas numerosas simulações com dados provenientes de fontes próprias, benchmarks mundiais, dados cedidos de forma exclusiva (UHE Igarapava-SP) e uma empresa local de Uberlândia. Neste estudo, as aplicações da metodologia se reduziram a simulações iniciais na área de predição. Quando o modelo apresenta um certo padrão de comportamento, a aplicação da metodologia mostrou resultados promissores.

Diante disso, cita-se os trabalhos apresentados em congressos e em revistas qualificadas nas áreas de matemática e engenharias IV:

Apresentações em congressos:

1. Identificação Fuzzy da Dinâmica de um Sistema de uma Célula a Combustível de Hidrogênio. Proceeding Series of the Brazilian Society of Computational and Applied Mathematics, Vol 5, Número 1, Gramado 2016. DOI: 10.5540/03.2017.005.01.0189
Link: Proceeding of SBMAC - Gramado - 2016.
2. Algoritmo de Gustafson-Kessel na Clusterização de Dados de uma Célula a Combustível de Hidrogênio. Proceedings do Quarto Congresso Brasileiro de Sistemas Fuzzy

(IV CBSF), Campinas, p. 21-32, Campinas 2016.

Link: Book of Abstracts IV CBSF - 2016.

3. Novel method based in fuzzy clustering for EEG signal analysis. Boletim Digital do II Encontro de Biomatemática, p. 71-74, Unicamp 2018.

Link: II Encontro de Biomatemática - ENCBIOMAT - 2018.

4. Novel Fuzzy System Identification: Comparative Study and Application for Data Forecasting. Apresentação na Universidade de Colorado, Denver, Colorado, Estados Unidos da América, 2018.

Artigos publicados:

1. Black-Box Fuzzy Identification of a Nonlinear Hydrogen Fuel Cell Model. Trends in Applied and Computational Mathematics, Brazilian Society of Computational and Applied Mathematics, p. 405-418, 2017. Doi: 10.5540/tema.2017.018.03.0405

Link: TEMA - 2017.

2. Novel Fuzzy System Identification: Comparative Study and Application for Data Forecasting, IEEE Latin America Transactions, Vol. 17, p. 1793-1799, 2019.

Link: IEEE Latin America Transactions - 2019.

7.2 Trabalhos futuros

A nova metodologia, desenvolvida neste estudo, ainda possui fatores em sua estrutura que podem ser melhorados. Entre os aspectos teóricos e computacionais está a construção dos antecedentes da inferência de Sugeno que são funções do tipo gaussiana (veja Capítulo 4) que dependem de dois parâmetros, α (nível considerado para a projeção) e β , associado ao desvio padrão σ da função gaussiana. Esses parâmetros são determinados através de um processo de otimização que devolve o valor “preferido” da modelagem. O código que estabelece essa preferência repete a dinâmica da projeção, tentando encontrar o mínimo do máximo absoluto da distância dos pontos z_i projetados e a função gaussiana construída com os valores dos parâmetros $\alpha \in [\alpha_0, \alpha_1]$ e $\beta \in [\beta_0, \beta_1]$.

O código próprio é eficiente a partir das conclusões das simulações efetuadas, mas ainda existem aspectos que podem ser melhorados relacionados ao esforço computacional. Para o artigo publicado na revista IEEE Latin América, utilizou-se também uma modificação do algoritmo de Guo (GUO, 2011), que, em seu desempenho, ofereceu praticamente os mesmos resultados dos adquiridos com a codificação própria. Notou-se, porém, que pensando os parâmetros α e β como números fuzzy do tipo 2 intervalares como definidos na Seção 3.4; o resultado pode ser consideravelmente superior como tem sido comprovado através de

simulações de modelos estudados recentemente. Como consequência, o sistema de inferência seria do tipo 2 intervalar, com antecedentes constituídos por funções gaussianas do tipo 2 intervalares. Essa modificação substituirá o algoritmo de otimização que determina α e β e apresenta promissor potencial para melhoria na metodologia. Outrossim, há o interesse teórico que pode chamar a atenção da comunidade matemática.

Outro aspecto desta investigação a ser melhorado é a sua aplicação a outros processos de engenharia. Como já destacado, foi iniciado um estudo na área da engenharia biomédica com resultados também auspiciosos. Assim sendo, há o intuito de reiniciar o projeto, assim como sua extensão para processos que clamam por métodos de identificação de dinâmica de dados.

Em recentes modelagens feitas através do emprego da técnica proposta nesse trabalho, tem-se incorporado mais uma família de modelos. O modelo tem como característica possuir alguma informação de controle da saída, por exemplo, seu domínio ser D . Neste caso o indexador-controlador pode ser um vetor randômico do domínio D onde são projetados os altos valores de pertinência. As primeiras simulações com um modelo do tipo parecem promissoras.

Referências

ABONIY, J.; BABUSKA, R.; SZEIFERT, F. Modified gath-geva fuzzy clustering for identification of takagi-sugeno fuzzy models. **IEEE Trans. on Systems, Man, and Cybernetics**, v. 5, n. 1, p. 612–621, 2002. Disponível em: <<https://doi.org/10.1109/TSMCB.2002.1033180>>.

ABONIY, J. et al. Compact ts-fuzzy models through clustering and ols plus fis model reduction. **10th IEEE International Conference on Fuzzy Systems**, 2001.

AGUIRRE, L. A tutorial introduction to nonlinear dynamics and chaos part ii: Modeling and control. **Controle e Automação**, v. 1, n. 1, p. 50 – 66, 1996.

AGUIRRE, L. A. **Introdução à identificação de sistemas**. 3. ed. Belo Horizonte - MG: Editora UFMG, 2007.

ALGHAMDI, M. et al. Self-organising and self-learning model for soybean yield prediction. **Sixth International Conference on Social Networks Analysis, Management and Security (SNAMS)**, 2019. Disponível em: <<https://doi.org/10.1109/SNAMS.2019.8931888>>.

ANDONOVSKI, G. et al. A practical implementation of robust evolving cloud-based controller with normalized data space for heat-exchanger plant. **Applied soft computing - Elsevier**, v. 48, p. 29–38, 2017. Disponível em: <<https://doi.org/10.1016/j.asoc.2016.05.036>>.

ANGELOV, P.; FILEV, D. An approach to online identification of takagi-sugeno fuzzy models. **IEEE Transactions on Systems, Man, and Cybernetics**, v. 34, n. 1, 2004. Disponível em: <<https://doi.org/10.1109/TSMCB.2003.817053>>.

AVELAR, H. **Estudo e desenvolvimento de um sistema de energia baseado em célula a combustível para injeção de potência na rede elétrica**. Tese (Doutorado) — Faculdade de Engenharia Elétrica, UFU, Uberlândia, Brasil, 2012.

BABUSKA, R. **Fuzzy Algorithms For Control**. 2. ed. Delft - Netherlands: Kluwer Academic Publishers, 1999.

BABUSKA, R. et al. **Fuzzy Modeling and Identification Toolbox**. 2018.

BANIN, S. L. **Python 3 - Conceitos e Aplicações - Uma Abordagem Didática**. 1. ed. [S.l.]: Editora Érica, 2018.

BERTONE, A. M. A.; CUNHA, M. J.; MARTINS, J. B. Algoritmo de gustafson-kessel na clusterização de dados de uma célula a combustível de hidrogênio. In: . Congresso Brasileiro de Sistemas Fuzzy - 2016, 2016. Disponível em: <https://www.ime.unicamp.br/~cbsf4/Papers_IVCBSF/BookOfAbstractsIVCBSF.pdf>.

BERTONE, A. M. A. et al. Identification of the dynamics of a hydrogen fuel cell system through fuzzy logic. In: . Proceeding Series of the Brazilian Society of Computational and Applied Mathematics, 2016. Disponível em: <<https://proceedings.sbmac.org.br/sbmac/issue/view/7/showToc>>.

BERTONE, A. M. A.; MARTINS, J. B. Black-box fuzzy identification of a nonlinear hydrogen fuel cell model. **TEMA - Trends in Applied and Computational Mathematics**, Brazilian Society of Computational and Applied Mathematics, 2017. Disponível em: <<https://doi.org/10.5540/tema.2017.018.03.405>>.

_____. Novel fuzzy identification method and a proof-of-concept case study to benchmark. In: . [S.I.]: Congresso Nacional de Matemática Aplicada e Computacional - CNMAC 2017, 2017.

BERTONE, A. M. A.; YAMANAKA, K.; MARTINS, J. B. Novel fuzzy system identification: Comparative study and application for data forecasting. **IEEE Latin America Transactions**, v. 17, p. 1793–1799, 2019. Disponível em: <<https://doi.org/10.1109/TLA.2019.8986416>>.

BEZDEK, J. **Pattern Recognition with Fuzzy Objective Function Algorithms**. Plenum Press, 1981. Disponível em: <<https://doi.org/10.1007/978-1-4757-0450-1>>.

BEZERRA, C. G. et al. An evolving approach to data streams clustering based on typicality and eccentricity data analytics. **Information Science**, n. 518, p. 13 – 28, 2020. Disponível em: <<https://doi.org/10.1016/j.ins.2019.12.022>>.

CASTRO, C. L.; BRAGA, A. P. Neural networks and learning systems, ieee transactions on. **Novel cost-sensitive approach to improve the multilayer perceptron performance on imbalanced data**, n. 6, p. 888–899, 2013. Disponível em: <<https://doi.org/10.1109/TNNLS.2013.2246188>>.

DOVZAN, D.; SKRJANC, I. Evolving fuzzy model for short-term prediction of energy consumption profiles. **Evolving and Adaptive Intelligent Systems (EAIS), 2016 IEEE Conference on**, 2016. Disponível em: <<https://doi.org/10.1109/EAIS.2016.7502498>>.

DUNN, J. C. Well-separated clusters and optimal fuzzy partitions. **Journal of Cybernetics**, Taylor and Francis, v. 4, n. 1, p. 95–104, 1974. Disponível em: <<https://doi.org/10.1080/01969727408546059>>.

ELÉTRICO, O. N. do S. **Geração de energia no centro-oeste e sudeste de 2000 a 2016**. 2015.

ESPITIA, H. E. et al. Proposal of an adaptive neurofuzzy system to control flow power in distributed generation systems. **Hindawi Complexity**, 2019. Disponível em: <<https://doi.org/10.1155/2019/1610898>>.

- FERNÁNDEZ, F. H.; KERRE, E. E.; JIMÉNEZ, B. L. M. A global fuzzy model for non linear systems using interval valued fuzzy sets. **Revista de Ingería Electrónica Automática y Comunicaciones**, v. 37, p. 50–57, 2016.
- FERREIRA, F. E. **Identificação de Modelos de Hammerstein e Wiener com Curva Estática Intervalar**. Dissertação (Mestrado) — Universidade Federal de Ouro Preto, Ouro Preto - MG - Brazil, 2017.
- GUO, H. A simple algorithm for fitting a gaussian function. **IEEE Signal Processing Magazine**, v. 28, p. 134–137, 2011. Disponível em: <<https://doi.org/10.1109/MSP.2011.941846>>.
- GUSTAFSON, D.; KESSEL, W. Fuzzy clustering with a fuzzy covariance matrix. In: . San Diego, USA: [s.n.], 1979. v. 1, n. 1, p. 761,766.
- HARASHEH, M. Forecasting wholesale electricity prices with artificial intelligence models: The italian case. **Sciprints**, 2016. Disponível em: <<https://doi.org/10.20944/preprints201607.0001.v1>>.
- JAFELICE, R. **Modelagem Fuzzy para Dinâmica de Transferência de Soropositivos para HIV em Doença Plenamente Manifesta**. Tese (Doutorado) — FEEC, Universidade Estadual de Campinas, Campinas, Brasil, 2003.
- JAFELICE, R. M.; BARROS, L. C.; BASSANEZI, R. C. **Teoria dos Conjuntos Fuzzy**. 2^a. ed. [S.l.]: SBMAC, 2012. v. 17.
- KARNIK, N. N.; MENDEL, J. M. Introduction to type-2 fuzzy logic systems. **Fuzzy Systems Proceedings, IEEE World Congress on Computational Intelligence**, v. 2, p. 915–920, 1998.
- LEUVEN, D. of E. E. E. K. **Daisy. Database for the Identification of Systems**. 2017.
- LOTUFO, F. A.; LUCENA, S. E. Identificação de sistemas complexos utilizando lógica nebulosa. **Conferência Brasileira de dinâmica, controle e aplicações**, v. 8, n. 1, 2011. Disponível em: <<https://doi.org/10.5540/DINCON.2011.001.1.0171>>.
- MACIEL, L.; GOMIDE, F.; BALLINI, R. Risk management using evolving possibilistic fuzzy modeling. **XV Encontro Brasileiro de Finanças**, 2015. Disponível em: <<https://doi.org/10.1109/FUZZ-IEEE.2015.7338039>>.
- MAESSCHALCK, D. J.-R. D. M. R. J. D. The mahalanobis distance. **Chemometrics and Intelligent Laboratory Systems**, 2000. Disponível em: <[https://doi.org/10.1016/S0169-7439\(99\)00047-7](https://doi.org/10.1016/S0169-7439(99)00047-7)>.
- MANZAN, J. R. G. **Análise de Desempenho de Redes Neurais Artificiais do Tipo Multilayer Perceptron por meio do distanciamento dos pontos do espaço de saída**. Tese (Doutorado) — Faculdade de Engenharia Elétrica, UFU, Uberlândia, Brasil, 2016.
- MARTINS, J. B.; BERTONE, A. M. A. Novel method based in fuzzy clustering for eeg signal analysis. In: . [S.l.]: Encontro de Biomatemática - 2018, 2018.
- MATLAB. **Matlab**. 2016.

- MENDEL, J. M. "type-2 set and systems: An overview". **Supplement to IEEE Computacional Intelligence Magazine**, v. 2, p. 20–29, 2007. Disponível em: <<https://doi.org/10.1109/MCI.2007.357235>>.
- MOORE, E. H. On the reciprocal of the general algebraic matrix. **Bulletin of the American Mathematical Society**, 1920.
- OCTAVE, G. **GNU Octave**. 2016.
- PAIVA, R. P. P. C. e. **Identificação neuro-difusa, aspectos de interpretabilidade**. Dissertação (Mestrado) — Faculdade de Ciências e Tecnologia Universidade de Coimbra, Coimbra, Portugal, 1999.
- PENROSE, R. A generalized inverse for matrices. **Proceedings of the Cambridge Philosophical Society**, 1955. Disponível em: <<https://doi.org/10.1017/S0305004100030401>>.
- PEREIRA, R. et al. Fuzzy clustering applied to a demand response model in a smart grid contingency scenario. **Power Electronics, Electrical Drives, Automation and Motion (SPEEDAM), 2014 International Symposium on**, 2014. Disponível em: <<https://doi.org/10.1109/SPEEDAM.2014.6872122>>.
- QINLIA, Z.; YUA, C. An improved generalized fuzzy model based on epanechnikov quadratic kernel and its application to nonlinear system identification. **Automatic Control and Computer Sciences**,, v. 53, p. 12–21, 2019. Disponível em: <<https://doi.org/10.3103/S0146411619010103>>.
- SANTOS, L. D. M. dos et al. Procedimentos de validação cruzada em mineração de dados para ambiente de computação paralela. **Escola Regional de Alto Desempenho Arquiteturas Multicore**, 2009.
- SCHILDT, H. **C: completo e total**. 3. ed. [S.l.]: Pearson, 1997.
- SERIR, L.; RAMASSO, E.; ZERHOUNI, N. Evidential evolving gustafson–kessel algorithm for online data streams partitioning using belief function theory. **International Journal of Approximate Reasoning**, v. 53, p. 747–768, 2012. Disponível em: <<https://doi.org/10.1016/j.ijar.2012.01.009>>.
- SILVA, I. N.; SPATTI, D. H.; FLAUZINO, R. A. **Redes Neurais Artificiais**. 1. ed. [S.l.]: Artliber Editora, 2010.
- STEINBRUCH, A.; WINTERLE, P. **Álgebra Linear**. 1. ed. [S.l.]: Pearson Universidades, 1987.
- SUGENO, M.; KANG, G. T. Structure identification on fuzzy model. **Fuzzy Sets and Systems**, v. 28, p. 329–346, 1988. Disponível em: <[https://doi.org/10.1016/0165-0114\(88\)90113-3](https://doi.org/10.1016/0165-0114(88)90113-3)>.
- TAKAGI, T.; SUGENO, M. Fuzzy identification of systems and its applications to modeling and control. **IEEE Transactions on Systems, Man, and Cybernetics**, v. 15, n. 1, p. 116–132, 1985. Disponível em: <<https://doi.org/10.1109/TSMC.1985.6313399>>.

- VILLARREAL-LÓPEZ, E. A new scheme for multiple fault detection and isolation for rotational mechatronic systems, by means of analytical redundancy and adaptive filtering. **Revista DYNA - Universidad Nacional de Colombia**, v. 86, p. 40–48, 2019. Disponível em: <<https://doi.org/10.15446/dyna.v86n209.73363>>.
- VORONIN, S.; PARTANEN, J. Price forecasting in the day-ahead energy market by an iterative method with separate normal price and price spike frameworks. **Energies**, v. 6, 2013. Disponível em: <<https://doi.org/10.3390/en6115897>>.
- WILLS, A. et al. Identification of hammerstein–wiener models. **Automatica**, v. 49, n. 1, p. 70 – 81, 2013. ISSN 0005-1098. Disponível em: <<https://doi.org/10.1016/j.automatica.2012.09.018>>.
- ZADEH, L. A. The concept of a linguistic variable and its application to approximate reasoning-1. **Information Science**, v. 8, p. 199–249, 1975. Disponível em: <[https://doi.org/10.1016/0020-0255\(75\)90036-5](https://doi.org/10.1016/0020-0255(75)90036-5)>.
- ZHOU, K.; YANGA, S.; SHAO, Z. Household monthly electricity consumption pattern mining: A fuzzy clustering-based model and a case study. **Journal of Cleaner Production - Elsevier**, v. 141, p. 900–908, 2017. Disponível em: <<https://doi.org/10.1016/j.jclepro.2016.09.165>>.

Apêndices

APÊNDICE A

Trabalhos próprios relacionados

Os trabalhos próprios relacionados ao presente estudo são detalhados neste capítulo. Eles incluem trabalhos apresentados em congressos e publicados em revistas, expostos na sua forma original e de maneira cronológica:

- 1** Identificação Fuzzy da Dinâmica de um Sistema de uma Célula a Combustível de Hidrogênio. Proceeding Series of the Brazilian Society of Computational and Applied Mathematics, Vol 5, Número 1, Gramado 2016.

- 2** Algoritmo de Gustafson-Kessel na Clusterização de Dados de uma Célula a Combustível de Hidrogênio. Proceedings do Quarto Congresso Brasileiro de Sistemas Fuzzy (IV CBSF), Campinas, p. 21-32, Campinas 2016.

- 3** Black-Box Fuzzy Identification of a Nonlinear Hydrogen Fuel Cell Model. Trends in Applied and Computational Mathematics, Brazilian Society of Computational and Applied Mathematics, p. 405-418, 2017.

- 4** Novel method based in fuzzy clustering for EEG signalanalysis. Boletim Digital do II Encontro de Biomatemática, p. 71-74, Unicamp 2018.

- 5** Novel Fuzzy System Identification: Comparative Study and Application for Data Forecasting. Apresentação na Universidade de Colorado, Denver, Colorado, Estados Unidos da América, 2018.

A.1 CNMAC 2016

Proceeding Series of the Brazilian Society of Applied and Computational Mathematics, Vol. 5, N. 1, 2017.

Trabalho apresentado no CNMAC, Gramado - RS, 2016.

Proceeding Series of the Brazilian Society of Computational and Applied Mathematics

Identificação Fuzzy da Dinâmica de um Sistema de uma Célula a Combustível de Hidrogênio

Ana Maria A. Bertone¹
 Jefferson Beethoven Martins²
 Márcio José Cunha³
 Darizon Alves de Andrade⁴
 Luis Paulo Fagundes⁵
 Faculdade de Matemática, FAMAT, UFU

Resumo: A identificação de sistemas dinâmicos é uma área, em franca expansão, que utiliza inúmeras ferramentas matemáticas para alcançar seus objetivos. Uma das formas de identificação é a técnica nebulosa (fuzzy), a qual foi utilizada para obter uma aproximação de dados reais gerados por um software simulador de uma célula a combustível de hidrogênio. O processo de identificação do sistema dinâmico é feito através de uma clusterização fuzzy dos dados de entrada, utilizando o algoritmo de Gustafson-Kessel, para imediatamente aplicar o método de inferência de Takagi e Sugeno (TS). Os testes de identificação e validação obtêm excelentes resultados em termos de “variance accounted for” (VAF), medida utilizada de proximidade estatística. Os resultados comprovam que a técnica difusa é uma promissora ferramenta para identificação de sistemas dinâmicos não-lineares.

Palavras-chave: Identificação de sistemas, método de Takagi e Sugeno, fuzzy clustering, células a combustível de hidrogênio.

1 Introdução

Fontes alternativas de energia são buscadas devido ao impacto que as tradicionais (carvão e petróleo) geram no meio ambiente. Uma das alternativas em estudo a nível mundial é o hidrogênio; já existem aplicações em funcionamento, entretanto os principais problemas de implantação se relacionam ao seu caráter explosivo e alto custo de pesquisas. O funcionamento de uma célula a combustível de hidrogênio com um conversor elevador *full-brigde* e um inversor monofásico PWM senoidal é mostrado no trabalho de Avelar [1] que, além de criar um software, simula o sistema real. O modelo proposto foi desenvolvido para uma célula de 1,2 KW e o processo de identificação de sistemas utilizado é dos mínimos quadrados.

¹anamaria@famat.ufu.br

²jefferson@ifm.edu.br

³cunhamjg@gmail.com

⁴darizon@ufu.br

⁵lpfagundeseai@gmail.com

O supracitado trabalho [1] desenvolveu um modelo para simular a célula a combustível de hidrogênio ligada à rede elétrica, levando em consideração o efeito da temperatura gerado internamente. O software gerado por esta pesquisa foi utilizado no nosso trabalho, simulando a planta real. Citando algumas das principais características do software simulador, temos que, a célula é capaz de gerar tensão entre 20 V e 50 V (contínua) e a tensão varia com a corrente drenada em seus terminais, sendo elevada a 380V antes de ser aplicada no módulo inversor. Coletados os dados estes são armazenados em um microcomputador. Salienta-se que as características da planta são diferentes para subida e descida de temperatura, pois o aquecimento é mais rápido que o resfriamento. A célula a combustível de membrana polimérica (PEMFC Nixa, fabricada pela Ballard) foi testada em laboratório para levantamento estático e dinâmico, sendo que toda a estrutura da célula a combustível é simulada através do software PSIM. Destaca-se que o modelo encontrado representa, de forma excelente, a dinâmica da planta (corrente-temperatura). Um esquema do processo aplicado por Avelar [1] é mostrado na Figura 1.

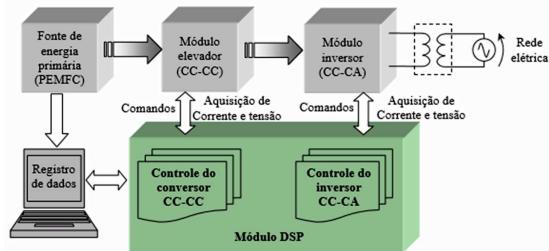


Figura 1: Fonte alternativa de energia baseada em célula a combustível (hidrogênio) [1].

Uma foto da montagem do laboratório onde é feito o recolhimento dos dados é mostrada na Figura 2.



Figura 2: O laboratório para levantamento do comportamento estático e dinâmico da célula a combustível [1].

Os dados empregados na nossa proposta são gerados através do simulador descrito [1]. A estrutura de dados é do tipo SISO *single input - single output*, ou seja, apenas uma entrada e uma saída. São 38.300 dados de entrada (corrente) com seus 38.300 dados de saída (temperatura) correspondentes. Os valores de corrente variam entre 24,9034A e 1,24136A; os valores de temperatura variam entre 54,835°C e 0°C. Os dados de entrada são gerados com o intuito de percorrer o maior espectro de frequência possível, sendo a taxa de amostragem de 0,2 segundos.

2 Metodologia

A identificação de sistemas é o estudo de técnicas matemáticas que permitem buscar um modelo que reproduza, em parte, um sistema dinâmico real [2]. Os sistemas são, basicamente, divididos em lineares e não-lineares, e os métodos de análise são conhecidos como caixa-branca (modelagem pela física), caixa-cinza (modelagem com entradas, saídas e informações adicionais sobre o sistema) e modelagem caixa-preta (apenas entradas e saídas, sem nenhuma informação adicional). A identificação caixa-cinza é um estudo relativamente novo e se apresenta como um campo para novas abordagens. Esta modelagem foi utilizada na pesquisa que deu origem ao software simulador [1]. Os testes provaram uma excelente eficácia do programa como mostrado na Figura 3.

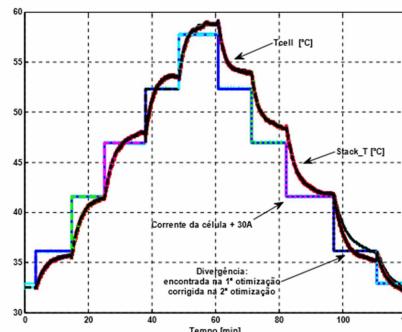


Figura 3: Comparação entre a temperatura da célula medida ($Stack_T$) e a simulada (T_{cell}) com modelo digitalizado, utilizando os parâmetros variáveis com a subida e descida da corrente, após otimizações do modelo. Fonte [1]

Existem várias maneiras de efetuar a aproximação de sistemas não-lineares. A principal diferença entre eles é a característica de tratar o sistema em forma global ou local [3]. Os globais descrevem o sistema através de funções que relacionam as variáveis de forma não-linear. Exemplos desse tipo de identificação é o método de estrutura **Narx** (Nonlinear **A**uto**R**egressive with **e**Xogenous input) [2]. As aproximações locais tentam dividir o sistema em subsistemas mais simples, em geral de tipo linear [4].

A lógica fuzzy tem demonstrado bom desempenho diante de difíceis tarefas de identificação e controle presentes, por exemplo, na indústria. Recomendada para a sistemas complexos, com múltiplas entradas e saídas, não-lineares e quando há poucas informações sobre o processo, tem sido utilizada em todo o mundo através da norma IEC 61131-7, se consolidando no mercado, por exemplo, nos controladores industriais [5].

As técnicas de modelagem fuzzy podem ser classificadas como de aproximação local, pois particionam o processo em diversas regiões fuzzy. Para cada região no espaço de entrada, uma regra é definida e esta determina, pela sua vez, a saída do modelo. Uma estrutura de regras é estabelecida e depende do mecanismo de inferência utilizado. Uns dos mais usados para identificação de sistemas, com grande complexidade e de natureza fortemente não-linear, é o do Takagi e Sugeno (TS) [6].

Na identificação do sistema deste trabalho, o modelo é construído a partir dos dados numéricos através do uso da técnica de clustering fuzzy, ou seja, a formação de grupos de dados de entrada por um certo critério de proximidade. Reunidos os dados em subsistemas locais, se aplica a inferência de TS.

Matematicamente, a formulação do fuzzy clustering procede, dado um banco de dados

$$z_k = [z_{1k}, z_{2k}, \dots, z_{nk}]^T \in \mathbb{R}^n, \quad k = 1, \dots, N$$

a encontrar uma matriz de partição fuzzy

$$U = \begin{bmatrix} \mu_{11} & \dots & \mu_{1k} & \dots & \mu_{1N} \\ \vdots & \dots & \vdots & \dots & \vdots \\ \mu_{c1} & \dots & \mu_{ck} & \dots & \mu_{cN} \end{bmatrix}$$

e os centros de cada cluster $V = \{v_1, v_2, \dots, v_c\}$, $v_i \in \mathbb{R}^n$, usando um método de otimização. O número c , números de clusters, é muito importante aos efeitos da identificação fuzzy dos sistemas não-lineares. Esta escolha é validada dependendo dos dados e do algoritmo utilizado.

O problema de otimização é resolvido dependendo da distância, d , entre os pontos do banco de dados e os pontos de V . Assim, a função objetivo do problema de otimização vem dada por

$$J(Z, V, U, A) = \sum_{i=1}^c \sum_{j=1}^N \mu_{i,j}^m d^2(z_j, v_i), \quad (1)$$

sujeita às restrições

$$\begin{aligned} 0 \leq \mu_{i,j} \leq 1, \quad i = 1, \dots, c, \quad j = 1, \dots, N & \text{ (restrição no grau de pertinência);} \\ 0 < \sum_{j=1}^N \mu_{i,j} < 1, \quad i = 1, \dots, c & \text{ (restrição do cluster não vazio);} \\ \sum_{i=1}^c \mu_{i,j} = 1, \quad j = 1, \dots, N & \text{ (restrição do grau de pertinência total)} \end{aligned} \quad (2)$$

onde m é conhecido como o parâmetro de fuzzificação e em geral é tomado igual a 2 [3], como é o caso do presente estudo.

Existem vários algoritmos para resolver o problema de otimização (1)-(2). Na presente análise é utilizado o algoritmo de Gustafson e Kessel (GK) [7], onde a distância é entre pontos é a induzida pela norma

$$d_{i,k,A_i}^2 = (z_k - v_i)^T A_i (z_k - v_i)$$

Esta distância é ajustada a cada iteração do algoritmo, o que representa uma adaptação topológica local da estrutura dos dados. A validação da escolha do número de clusters e do parâmetro de fuzzificação escolhido é feito através de uma medida de validação. Também existem várias medidas desse tipo sendo que, a utilizada neste trabalho, é a do Coeficiente de Partição (CP) e Entropia da Partição (EP) [8]. A validação positiva do clustering vem dada quando o índice de validade CP está perto de 1 e EP perto de 0.

O modelo fuzzy afim de TS consiste em regras, R_i , $i = 1, 2, \dots, K$, associadas a determinadas K variáveis, com a seguinte estrutura:

$$R_i : \text{Se } \underbrace{\mathbf{x} \text{ é } A_i}_{\text{antecedente}} \text{ então } \underbrace{y_i = \mathbf{a}_i^T \mathbf{x} + b_i}_{\text{consequente}},$$

onde $\mathbf{x} \in \mathbb{R}^p$, para algum $p \in \mathbb{N}$, A_i é um conjunto fuzzy caracterizado pela sua função de pertinência $\mu_A(x) \in [0, 1]$, $y_i \in \mathbb{R}$ é a saída da regra i -ésima do sistema, \mathbf{a}_i é um vetor paramétrico e $b_i \in \mathbb{R}$. Assim, dadas as saídas y_i se obtém a saída global do modelo TS utilizando a média fuzzy com peso $y = \frac{\sum_{i=1}^K \beta_i(\mathbf{x}) y_i}{e \sum_{i=1}^K \beta_i(\mathbf{x})}$, com $\beta_i(\mathbf{x}) = \mu_{A_i}(\mathbf{x})$, sendo que o grau de pertinência do antecedente é calculado por uma operação fuzzy, por exemplo o produto, dos componentes de \mathbf{x} . Portanto, denotando $\gamma_i(\mathbf{x}) = \frac{\beta_i(\mathbf{x})}{\sum_{i=1}^K \beta_i(\mathbf{x})}$, a saída global do sistema de inferência TS vem dado por

$$y = \left(\sum_{i=1}^K \gamma_i(\mathbf{x}) \mathbf{a}_i^T \right) \mathbf{x} + \sum_{i=1}^K \gamma_i(\mathbf{x}) b_i.$$

Reunidas as metodologias fuzzy (de clustering e inferência), após a aquisição dos dados, entradas e saídas são selecionadas. O algoritmo usado para a obtenção do modelo fuzzy segue as etapas mostradas na Figura 4

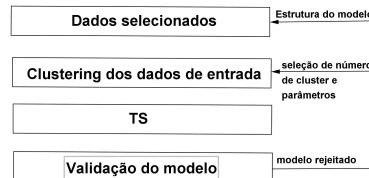


Figura 4: Algoritmo de identificação usando o clustering e TS.

O desempenho do modelo é avaliado pela medida de proximidade estatística da variância VAF (Variance Accounted For), que é dada pela expressão

$$\frac{1 - \text{variância(saída real)} - \text{variância(saída do modelo)}}{\text{variância(saída real)}} \times 100.$$

Quanto mais próximo estiver o resultado dos 100%, melhor será a qualidade do modelo.

3 Simulações numéricas

No primeiro passo do algoritmo da Figura 4 o número de clusters selecionado é de 23, validado pelos índices $CP=0.9436$ e $EP=0.1293$. O clustering dos dados de entrada obtido mediante o algoritmo GK é mostrado na Figura 5. Os pontos em azul representam os dados de entrada do modelo agrupados em clusters com formato elipsoidal e, as diferentes cores, representando os graus de pertinência.

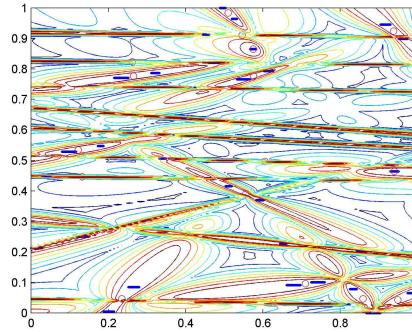


Figura 5: Clustering dos dados de entrada.

Para iniciar o processo de modelagem são selecionados os 25.000 primeiros valores de entrada e saída, com a finalidade de identificação (busca do modelo) e os 13.300 últimos valores são utilizados para validação (confirmação da eficiência do modelo). A Figura 6 mostra o gráfico gerado pelo modelo fuzzy de clusterização TS com todos os dados Figura 6(a) e no período de validação Figura 6(b) e seus respectivos VAF.

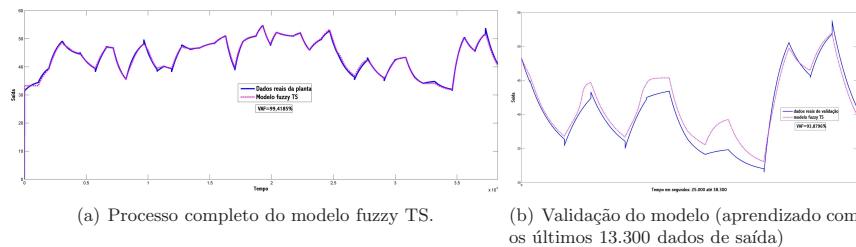


Figura 6: O modelo fuzzy TS em versão completa (a) e no trecho de validação (b).

4 Conclusão

A técnica nebulosa (fuzzy) é utilizada com excelente desempenho na identificação da dinâmica de um sistema não-linear proposto com dados reais. Com os resultados apresentados comprova-se a eficiência da técnica, abrindo inúmeras possibilidades de aplicação a diversos outros problemas reais. O desempenho da identificação via a lógica difusa, que combina o clustering, também difuso, dos dados de entrada com o método de inferência difuso de Takagi e Sugeno, alcançou 99,75% quando o mesmo grupo de dados é utilizado para a identificação e validação. Ao separar dois grupos de dados a porcentagem alcançada é de 94%, sendo os primeiros 25.000 dados para a identificação e 13.300 dados para validação. Diante do magnífico resultado é notório que o aprendizado da dinâmica do sistema ocorre de forma efetiva.

Referências

- [1] H.J. Avelar. *Estudo e desenvolvimento de um sistema de energia baseado em célula a combustível para injeção de potência na rede elétrica*. PhD thesis, Faculdade de Engenharia Elétrica, Universidade Federal de Uberlândia, Uberlândia, Brasil, 2012.
- [2] L. A. Aguirre. *Introdução à Identificação de Sistemas*. Editora UFMG, Belo Horizonte, 3 edition, 2007.
- [3] R. Babuska. *Fuzzy Modeling for Control*. Kluwer Academic Publishers, Boston, USA, 1998.
- [4] T. A. Johansen e B. A. Foss. Semi-empirical modeling of non-linear dynamic systems through identification of operating regimes and local models. *Modeling Identification and Control*, 16(4):213 – 232, 1995.
- [5] M. G. Simões e I.S. Shaw. *Controle e Modelagem Fuzzy*. Editora Edgar Blücher Ltda/FAPESP, 2 edition, 2007.
- [6] T. Takagi e M. Sugeno. Fuzzy identification of systems and its applications to modeling and control. *IEEE Transactions on Systems, Man, and Cybernetics*, 15(1):116 – 132, 1985.
- [7] D. Gustafson e W. Kessel. Fuzzy clustering with a fuzzy covariance matrix. In *Proc. IEEE CDC*, pages 761 – 766, San Diego, CA, USA, 1979.
- [8] M-W. Lee e K-C Kwak. Context-based Gustafson and Kessel clustering with information granules. *International Journal of Emerging Technology and Advanced Engineering*, 4(11):76 – 80, 2014.

A.2 Congresso Brasileiro de Sistemas Fuzzy 2016

Quarto Congresso Brasileiro de Sistemas Fuzzy (IV CBSF)
16–18 de Novembro de 2016, Campinas – SP, Brasil.

Algoritmo de Gustafson-Kessel na Clusterização de Dados de uma Célula a Combustível de Hidrogênio

Ana Maria A. Bertone^{*(1)} Jefferson Beethoven Martins ^{**(2)} Márcio José da Cunha ^{***(2)}

(1) FAMAT, (2) FEELT, Universidade Federal de Uberlândia, Av. João Naves de Ávila, 2121
Uberlândia, MG 38.408-902, Brazil

Resumo O algoritmo de clusterização fuzzy de Gustafson-Kessel é aplicado a um banco de dados numérico, simulado a partir de uma célula real de combustível a hidrogênio. Este estudo constitui a etapa inicial de um projeto mais amplo, que envolve a identificação, também fuzzy, de sistemas dinâmicos não lineares. Nesta parte do estudo é feita a implementação do algoritmo de clusterização no software livre Python (Spyder). O objetivo é testar a eficiência do procedimento, através da interface de desenvolvimento, diante de um grande número de dados presentes no experimento. Como validação da escolha do número de clusters são utilizados simultaneamente os índices de participação e entropia.

Palavras-chave Fuzzy clustering; Gustafson-Kessel; Célula a combustível de Hidrogênio

1 Introdução

A clusterização é uma técnica matemática utilizada para agrupar dados segundo um critério, ou seja, similaridade matemática. Esta técnica é utilizada, principalmente, em mineração de dados (data mining). Os dados utilizados em um processo de clusterização podem pertencer a diferentes contextos como: histórico faturamento de uma empresa, números sobre a evolução de uma doença em determinado local ou, até mesmo, valores de pressão e temperatura utilizados em um equipamento industrial controlado. A diversidade das aplicações da técnica fuzzy de clustering pode ser vista em [1], em que Costa et al. caracterizam morfológicamente e estimam a diversidade genética de pimentas presentes no estado do Amazonas. O método utilizado é o hierárquico das médias das distâncias UPGMA (Unweighted Pair-Group Method Using an Arithmetic Average). Em [2] Rizzo et al., mostram um mapeamento digital de solos, utilizando o método

^{*} anamaria@famat.ufu.br

^{**} jefferson@ifm.edu.br

^{***} mjcunha@eletrica.ufu.br

fuzzy de C-Means [3]. Em [4], Alves et al., os cento e sessenta bairros da cidade de Rio de Janeiro são agrupados pelo algoritmo fuzzy C-means, para obter importantes informações para a gestão municipal. Finalmente, citando apenas trabalhos recentes feitos a nível nacional, [5] Veras et al. propõem um algoritmo que detecta os exsudatos (depósitos de gordura) através do tratamento de imagens. O estudo utiliza fuzzy clustering e técnicas de morfologia matemática.

Em 1979 Donald E. Gustafson e William C. Kessel publicaram o artigo [6] apresentando um algoritmo modificado do tradicional fuzzy C-Means [7]. Foi a partir dessa data que o algoritmo passou a ser conhecido como Gustafson-Kessel (GK). A principal diferença em relação ao fuzzy C-Means é a substituição da distância euclidiana, de forma que a matriz de covariância detecta as formas geométricas com mais precisão, independentes de cada cluster, isto a cada passo do algoritmo.

Em [8], Avelar apresenta o projeto de uma fonte alternativa de energia, com o funcionamento de uma célula a combustível de hidrogênio, através de um conversor elevador full-bridge e um inversor monofásico PWM senoidal. O modelo proposto em [8] foi implementado para uma célula de 1,2 KW e o processo de identificação da dinâmica de sistemas é o método dos mínimos quadrados ordinários. Uma das principais contribuições do trabalho de Avelar foi representar a dinâmica da temperatura interna da célula em função da corrente pois, outros simuladores, normalmente, não o fazem. Devido sua alta confiabilidade, o trabalho de [8] foi empregado para gerar os dados utilizados neste estudo. Por meio deste método de clusterização fuzzy, um grande número de dados numéricos gerados são agrupados [3]. Como ponto de partida, é desenvolvido o algoritmo de GK na plataforma Anaconda (Continuum Analytics) que usa a linguagem Python, a qual permite o uso de inúmeras bibliotecas por meio da importação de rotinas. Esta linguagem de desenvolvimento é livre, de fácil obtenção e com grande suporte de fóruns e sites da área.

O Spyder (Python) é um ambiente que facilita a programação devido a sua interface e quantidade de bibliotecas. De alta inteligibilidade, ele possui robusto debugger que torna mais rápido o trabalho do programador, testando, em tempo de execução, de forma muito parecida com o Matlab. Junto ao Anaconda Python, o Spyder se torna muito útil para programação científica. O cálculo de matrizes transpostas, multiplicação entre matrizes e geração de figuras são apenas algumas das facilidades apresentadas por essa linguagem de programação e ambiente de desenvolvimento integrados. O Spyder é uma IDE compatível com diferentes plataformas como MS Windows e Linux, possuindo simples instalação e grande número de informações a seu respeito (suporte amplo e facilitado).

Um importante aspecto a ser ressaltado é a escolha do número de clusters para o uso do algoritmo GK. Neste trabalho foram considerados simultaneamente os índices de partição e de entropia, que calculam a nebulosidade da clusterização.

Este trabalho é organizado em cinco seções. A Seção 1 é a introdução a este estudo. A seção 2 explica a origem dos dados utilizados neste estudo (particularidades da célula a combustível de hidrogênio). Na seção 3 é explicado o

conceito de clusterização fuzzy e suas propriedades. A seção 4 mostra os passos do algoritmo de Gustafson-Kessel no ambiente Python, as simulações realizadas, elucidando os resultados. A seção 5 são feitas as conclusões e são traçadas futuras ideias a ser desenvolvidas, ligadas a este estudo.

2 Simulador de células a combustível de hidrogênio

Existem várias fontes de energia, todavia as alternativas são buscadas devido ao impacto que as tradicionais (carvão e petróleo) geram no meio ambiente. Uma das possibilidades em estudo a nível mundial é o hidrogênio; já existem aplicações em funcionamento, entretanto os principais problemas de implantação se relacionam ao seu caráter explosivo e alto custo de pesquisas. O funcionamento de uma célula a combustível de hidrogênio com um conversor elevador “full-bridge” e um inversor monofásico PWM senoidal é mostrado no trabalho de Avelar [8] que, além de criar um software, simula o sistema real com fidelidade. O modelo proposto foi desenvolvido para uma célula de 1,2 KW e o processo de identificação de sistemas utilizado é dos mínimos quadrado ordinários.

O supracitado trabalho [8] desenvolveu um modelo para simular a célula a combustível de hidrogênio ligada à rede elétrica, levando em consideração o efeito da temperatura gerado internamente. O software gerado por esta pesquisa foi utilizado no nosso trabalho, simulando a planta real. Ao citar algumas das principais características do software simulador, observa-se que a célula é capaz de gerar tensão entre 20 V e 50 V (contínua) e a tensão varia com a corrente drenada em seus terminais, sendo elevada a 380V antes de ser aplicada no módulo inversor. Todos os dados coletados são armazenados em um microcomputador. Salienta-se que as características da planta são diferentes para subida e descida de temperatura, pois o resfriamento é mais lento que o aquecimento. A célula a combustível de membrana polimérica (PEMFC Naxa, fabricada pela Ballard) foi testada em laboratório para levantamento estático e dinâmico, sendo toda a estrutura da célula a combustível, simulada através do software PSIM. Um esquema do processo aplicado por Avelar [8] é mostrado na Figura 1

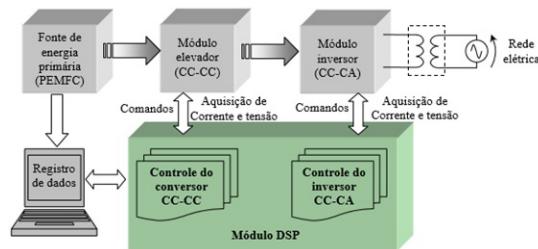


Figura 1. Fonte alternativa de energia baseada em célula a combustível (hidrogênio) [8].

Destaca-se que o modelo encontrado representa, de forma excelente, a dinâmica da planta (corrente-temperatura). Uma foto da montagem do laboratório[8] onde é feito o recolhimento dos dados é mostrada na Figura 2.



Figura 2. Laboratório para levantamento do comportamento estático e dinâmico da célula a combustível [8].

Os dados empregados na nossa proposta são gerados através do simulador descrito [8]. A estrutura de dados é do tipo SISO (single input - single output em inglês), ou seja, apenas uma entrada e uma saída. São 38.300 dados de entrada (corrente) com seus 38.300 dados de saída (temperatura) correspondentes. Os dados de entrada são gerados com o intuito de percorrer o maior espectro de frequência possível, sendo a taxa de amostragem de 0,2 segundos, os dados de saída são gerados através do modelo de Avelar[8].

3 Clusterização Fuzzy

Clusterização é a classificação não-supervisionada de dados, formando agrupamentos ou clusters. Esta metodologia representa uma das principais etapas de processos de análise de dados. A análise de clusters tem como objetivo a organização de padrões, representados matematicamente na forma de vetores ou pontos em um espaço multidimensional em clusters, de acordo com alguma medida de similaridade. A principal característica da classificação não-supervisionada, é agrupar um conjunto de padrões não-rotulados em clusters que possuam alguma propriedade em comum.

Distinguimos em um processo de clusterização as seguintes componentes:

1. representação dos valores que envolvem as definições do número, tipo e modo de apresentação das características de cada padrão;

2. definição de uma medida de similaridade, que é uma função de distância definida entre pares de dados ou padrões.
3. clusterização de acordo com a medida escolhida; se a função é determinística ou binária, ou seja, um padrão pertence ou não-pertence a um dado grupo, sendo uma clusterização crisp ou hard. Caso a medida seja uma função de pertinência fuzzy, então um padrão pode apresentar graus de pertinência em relação aos grupos.
4. validação do resultado que geralmente se recorre a critérios de otimalidade, muitas vezes definidos de forma subjetiva [10].

O método de clusterização fuzzy clustering permite que os elementos do banco de dados pertençam a vários grupos simultaneamente com diferentes graus de pertinências. Assim, o conjunto de dados é particionado em um número c de conjuntos fuzzy, onde c é o número total de clusters.

As principais características da clusterização fuzzy são:

- Cada elemento deve ter um conjunto de pertinências que expressa o “quanto ele pertence” a cada cluster.
- O somatório de todas as pertinências de um elemento é igual a um.

Concretamente, a formulação do fuzzy clustering a partir de um banco de dados de uma única entrada com n elementos é

$$z = (z_1, z_2, \dots, z_n) \in \mathbb{R}^n,$$

sendo necessário encontrar uma matriz de pertinência fuzzy

$$U = \begin{bmatrix} \mu_{11} & \dots & \mu_{1k} & \dots & \mu_{1n} \\ \mu_{21} & \dots & \mu_{2k} & \dots & \mu_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \mu_{1c} & \dots & \mu_{1k} & \dots & \mu_{cn} \end{bmatrix}$$

e os centros de cada cluster

$$V = \{v_1, v_2, \dots, v_c\}, \quad v_i \in \mathbb{R},$$

usando um método de otimização.

O algoritmo de GK estende a clássica clusterização fuzzy C-Means utilizando uma atualização da distância, com o objetivo de detetar clusters com diferentes estruturas geométricas [6]. Cada cluster possui sua própria geometria induzida por uma matriz M^i , determinando a distância ao centro do cluster como sendo

$$d_{i,j,M_i}^2 = (z_j - v_i)M^i(z_j - v_i)^T, \quad i = 1, \dots, c, \quad j = 1, \dots, n. \quad (1)$$

Sendo $M = (M^1, \dots, M^c)$ a c -upla de matrizes correspondente a cada cluster, o funcional objetivo do algoritmo de GK [6] é dado por

$$J(X, U, V, M) = \sum_{i=1}^c \sum_{j=1}^n \mu_{i,j}^m d_{i,j,M_i}^2,$$

sujeito às restrições

$$0 \leq \mu_{i,j} \leq 1, i = 1, \dots, c, j = 1, \dots, n \text{ (restrição do grau de pertinência);}$$

$$0 < \sum_{j=1}^n \mu_{i,j} < 1, i = 1, \dots, c \text{ (restrição do cluster não vazio);} \quad (2)$$

$$\sum_{i=1}^c \mu_{i,j} = 1, j = 1, \dots, n \text{ (restrição do grau de pertinência total),}$$

em que m é conhecido como o parâmetro de fuzzificação e, em geral, é tomado igual a 2 [3], como é o caso do presente estudo. Além disso, cada matriz M_i tem como restrição

$$\det(M_i) = \text{determinate de } A_i < \rho_i, \rho_i > 0, \text{ para todo } i.$$

Assim, M_i é obtida usando o método dos multiplicadores de Lagrange dado pela seguinte equação

$$M_i = \rho_i \det(F_i)^{\frac{1}{N}} F_i^{-1}, \quad (3)$$

em que N é a dimensão do espaço do cluster e F_i é a matriz de covariância do i -ésimo cluster, definida por

$$F_i = \frac{\sum_{j=1}^N \mu_{ij}^m (z_j - v_i)(z_j - v_i)^T}{\sum_{j=1}^N \mu_{ij}^m}. \quad (4)$$

Finalmente, os centros dos clusters são calculados através da forma [3]

$$v_i = \frac{\sum_{j=1}^N \mu_{ij}^m z_j}{\sum_{j=1}^N \mu_{ij}^m}. \quad (5)$$

A validação da escolha do número de clusters é equivalente a encontrar o número ideal de clusters a partir de determinado conjunto de dados. Vários índices de validade têm sido estudados para determinar o número ótimo de clusters. Bezdek [7] propõe dois índices:

1. Coeficiente de Partição (CP) que mede o número de superposições entre clusters e definido por

$$CP(c) = \frac{1}{N} \sum_{i=c}^c \sum_{j=1}^n \mu_{ij}^2.$$

Temos que $\frac{1}{c} \leq CP(c) \leq 1$ e o valor ótimo c_0 encontrado em um intervalo de escolha $I = [c_{min}, c_{max}] \cap \mathbb{Z}$, \mathbb{Z} conjunto dos números inteiros positivos, se calcula como

$$c_0 = \max_{c \in I} CP(c).$$

2. Índice de Classificação de Entropia (CE) que mede a “nebulosidade” (fuzziness em inglês) do cluster pela fórmula

$$CE(c) = -\frac{1}{N} \sum_{i=c}^c \sum_{j=1}^n \mu_{ij} \log(\mu_{ij}).$$

O valor ótimo c_0 é encontrado em um conjunto de escolha $I = [c_{min}, c_{max}]$ como sendo

$$c_P = \max_{c \in I} PC(c).$$

O número de clusters, c , é muito importante devido aos efeitos, por exemplo, da identificação fuzzy dos sistemas não-lineares. No nosso estudo é escolhido o número ótimo combinando os índices CP e CE simultaneamente em dois casos: para valores de c no de escolha $I = [2, 6] \cap \mathbb{Z}$, para mostrar em forma clara, no gráfico obtido, a estrutura geométrica dos clusters. O outro valor foi escolhido em $I = [20, 25] \cap \mathbb{Z}$, com o objetivo de ser utilizado na identificação do sistema dinâmico que relaciona os dados da célula de oxigênio.

4 O algoritmo de Gustafson-Kessel na interface Python Spyder

Descrevemos nesta seção o passo-a-passo do algoritmo de GK, mostrando em cada passo os resultados da programação no Python para o caso de $c = 2$.

Passo 1: Toma-se como valor de fuzzificação $m = 2$. Seja Z o conjunto de dados e $\{c_{ini} \mid i = 1, \dots, c\} \subset Z$. A forma de escolher os centros não foi aleatória, com o objetivo de comparar a performance do algoritmo em termos dos índices PC e CE. Assim, foram escolhidos dois elementos dos dados de forma “espalhada”, como mostra a tabela da Figura 3

		0	1
0	6.27710	0.00000	
1	24.90340	54.11760	

Figura 3. Coordenadas dos centros de cluster escolhidos: a primeira coluna são as abscissas.

Passo 2. Calcula-se a matriz de partição $U0 = (\mu_{ij})$, que contém as pertinências de cada dado ao clusters. Parte da matriz é mostrada na tabela da Figura 4

U0 - NumPy array

	0	1
17274	0.11561	0.88439
17275	0.11515	0.88485
17276	0.11468	0.88532
17277	0.11422	0.88578
17278	0.11377	0.88623
17279	0.11331	0.88669
17280	0.11286	0.88714
17281	0.11240	0.88760
17282	0.11195	0.88805
17283	0.11150	0.88850
17284	0.11106	0.88894
17285	0.11061	0.88939
17286	0.11017	0.88983

Figura 4. Matriz de partição $U0$.

Passo 3. Calcula-se a matriz de covariância Cov_i de cada cluster, dado pela fórmula (4). As primeiras matrizes de covariância obtidas pelo algoritmo é mostrada na Figura 5.

Cov - List (2 elements)

Index	Type	Size	Value
0	float64	(2, 2)	array([[55.70744182, 25.07117375], [25.07117375, 32.54794913]])
1	float64	(2, 2)	array([[55.70744182, 25.07117375], [25.07117375, 32.54794913]])

Figura 5. As duas matrizes de covariância de cada cluster. Cada “array” corresponde à matriz de um cluster, 0 ou 1, ou seja, primeiro e segundo cluster da linguagem Python.

Passo 4. Calcula-se o centro, v_i , de cada cluster, através do cálculo da fórmula (5).

Obtém-se dois primeiros centros de cluster cujas coordenadas são mostradas na tabela da Figura 6

V1 - NumPy array		
	0	1
0	5.17029	38.70991
1	17.88010	47.31395

Figura 6. Coordenadas dos novos centros de cluster: a primeira coluna são as abscissas.

Passo 5. Calcula-se a matriz M_i que induz a distância e a geometria de cada cluster, usando a fórmula (3). No primeiro loop do algoritmo obtém-se a matriz M_i mostrado na Figura 7.

M - List (2 elements)			
Index	Type	Size	Value
0	float64	(2, 2)	array([[1.05912397, -0.10967064], [-0.10967064, 0.95553276]])
1	float64	(2, 2)	array([[0.84318932, -0.27963945], [-0.27963945, 1.27871428]])

Figura 7. As matrizes M_i que induzem a norma de cada cluster. Cada “array” corresponde a um cluster, 0 ou 1.

Passo 6. Calcula-se a distância de cada centro de cluster em relação a cada elemento, utilizando a distância induzida pela matriz M_i , obtendo a matriz das distâncias, $d1$, algumas de cujas entradas são ilustradas na tabela da Figura 8

d1 - NumPy array							
	6355	6356	6357	6358	6359	6360	6361
0	215.977	216.046	216.115	216.185	216.255	216.324	216.395
1	9.590	9.537	9.485	9.432	9.379	9.327	9.275

Figura 8. Algumas das entradas da nova matriz das distâncias $d1$.

Passo 7. Actualiza-se a matriz de pertinência $U0$ pela matriz U , calculando

$$\mu_{ij}^2 = \frac{1}{\sum_{k=1}^c \left(\frac{(d_{ij}^2)^2}{(d_{kj}^2)^2} \right)^{\frac{1}{m-1}}}, \quad i = 1, \dots, c, \quad j = 1 \dots n.$$

Passo 8. Calcula-se o erro e o primeiro ciclo é terminado. No nosso estudo calculamos o erro por meio da norma do máximo, ou seja,

$$\text{erro} = \max \|\mu_{ij}^1 - \mu_{ij}^2\|,$$

em que μ_{ij}^k representam as entradas das matrizes de partição em voltas de loop de ordem k e seguinte. Neste passo 8 do primeiro loop do algoritmo, o erro obtido é de 0,639188829.

Passo 9. Caso o erro seja maior que o pré-estabelecido, se deve voltar ao Passo 2. Neste caso a meta pre-estabelecida, como teste, é do erro ser menor que 0,05.

Uma vez terminado o algoritmo, foram calculados os índices de validação CP e CE, que tornaram os valores descritos na Tabela 1.

	2	3	4	5	6
CP	0.8078	0.7433	0.6822	0.6380	0.6385
CE	0.3201	0.4659	0.6057	0.7314	0.7447

	20	21	22	23	24	25
CP	0.6520	0.7226	0.66331	0.6798	0.7407	0.6721
CE	0.8264	0.6715	0.8194	0.7877	0.6225	0.8124

Tabela 1. Tabela de valores dos índices CP e CE retornados após a finalização de cada algoritmo com erro de saída 0,01 e mesmo centros iniciais de clusters da tabela da Figura 3. Em destaque os maiores valores de CP e os menores valores de CE simultâneos.

Com a finalidade de construir o gráfico das duas clusterizações escolhidas, são calculados os autovalores e autovetores das matrizes de covariança, que informam sobre a estrutura geométrica do cluster. De fato, o eixo principal da elipse de nível 1 do cluster, tem a direção do autovetor da matriz de covariança do cluster, correspondente ao maior autovalor, sendo a raiz quadrada desse autovalor a medida do semieixo maior. A direção perpendicular é dado pelo outro autovetor e a medida do eixo menor é a raiz quadrada do autovalor correspondente.

Na Figura 9 é mostrado a clusterização de dois clusters e na Figura 10 a de 24 clusters.

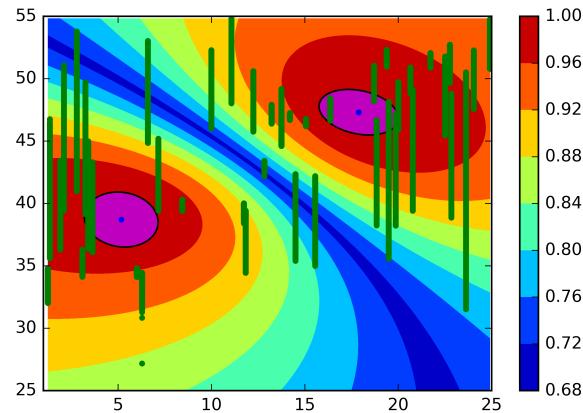


Figura 9. Partição fuzzy dos dados da célula de oxigênio. Os pontos em cor verde representam os pares de atributos (entrada, saída). As elipses preenchidas de cor magenta representam a geometria do cluster de, exatamente, nível 1. As outras regiões representam níveis de pertinência menor a 1, como indica a escala ao lado.

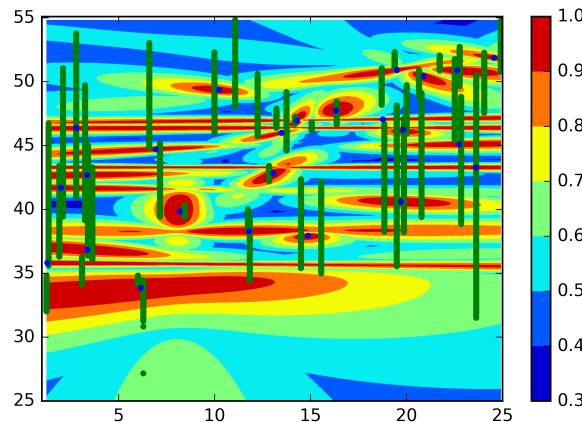


Figura 10. A partição fuzzy dos dados da célula de oxigênio, em cor verde, por 24 clusters com as regiões de nível coloridas de acordo com a escala ao lado.

5 Conclusões

O algoritmo de clusterização de Gustafson e Kessel funciona com sucesso para o grupo de dados utilizados na presente pesquisa. Isto motiva essa primeira etapa de um projeto maior, cujo objetivo é construir, utilizando um software livre, um aplicativo de propósito geral. A complexidade desse aplicativo iria abranger identificação de sistemas não lineares do tipo single input - single output (só uma saída e uma entrada), até multiple input - multiple output, ou seja, múltiplas entradas e múltiplas saídas. Para testar este algoritmo Gustafson Kessel são utilizados, neste estudo, apenas uma entrada, com 38300 valores, e seus correspondentes dados de saída, obtidos em laboratório em períodos amostrais de 0,2 segundos.

Os desafios são grandes, mas possibilidades também são. Clusterizar, identificar e validar modelos através de métodos caixa-preta, como identificação de sistemas através de lógica fuzzy, é um promissor campo de pesquisa, sendo este o foco dos autores do presente artigo.

Referências

1. Costa, L.V., Bentes, J.L.S., Lopes, M.T.G., Alves, S.R.M., Viana Júnior, J.M.: Caracterização de acessos de pimentas do Amazonas. Horticultura Brasileira, 33, 290 – 298 (2015)
2. Rizzo, R., Demattê, J.A.M., Lacerda, M.P.C.: Espectros VIS-NIR do Solo e Fuzzy K-Médias Aplicados na Delimitação de Unidades de Mapeamento de Solos em Topossequências. R. Bras. Ci. Solo, 39, 1533 – 1543 (2015)
3. Abonyi, J., Babuska, R., Szeifert, F.: Modified Gath-Geva fuzzy clustering for identification of Takagi-Sugeno fuzzy models. IEEE Trans. on Systems, Man, and Cybernetics, 32(5), 612 – 621 (2002)
4. Alves, B.D.S., Barbosa, M.T.S., Caffarena, E.R., da Silva, A.S.: Caracterização do envelhecimento populacional no município do Rio de Janeiro: contribuições para políticas públicas sustentáveis. Saúde Colet, Rio de Janeiro, 24(1), 63 – 69 (2016)
5. Veras, R.M.S., Medeiros, F.N.S., Araújo, F.H.D., Santana, A.M., Silva, R.R.V.: Detecção de exsudatos em imagens de retina por técnicas de morfologia matemática e agrupamento nebuloso. Rev. Bras. Eng. Biom., 29(1), 45 – 56 (2013)
6. Gustafson, D.E., Kessel, W.C.: Fuzzy clustering with fuzzy covariance matrix. Proceedings of the IEEE Control and Decision Conference, San Diego, .761 – 766 (1979)
7. Bezdek, J., Keller, J., Krishnapuram, R., Pal, N.: Fuzzy models and algorithms for pattern recognition and image processing, Springer (2005).
8. Avelar, H.J.: Estudo e desenvolvimento de um sistema de energia baseado em célula a combustível para injeção de potência na rede elétrica.: Tese de Doutorado Faculdade de Engenharia Elétrica, Universidade Federal de Uberlândia (2012).
9. Eclipse Foundation
<http://www.eclipse.org> Último acesso Maio (2016)
10. Jain, A.K., Dubes, R.C.: Algorithms for Clustering Data, Prentice Hall, (1988)

A.3 Trends in Applied and Computational Mathematics 2017

Tema

Tendências em Matemática Aplicada e Computacional, 18, N. 3 (2017), 405-418
 © 2017 Sociedade Brasileira de Matemática Aplicada e Computacional
www.scielo.br/tema
 doi: 10.5540/tema.2017.018.03.0405

Black-Box Fuzzy Identification of a Nonlinear Hydrogen Fuel Cell Model

A.M.A. BERTONE^{1*}, J.B. MARTINS² and K. YAMANAKA²

Received on November 15, 2016 / Accepted on August 28, 2017

ABSTRACT. A fuzzy identification of the system's dynamic is developed with a data generated by a hydrogen fuel cell simulator. The data obtained is single input/single output, without having previous knowledge of the system model, and showing nonlinear behavior. The choice of the fuzzy method for identification is based on those particular data features, and the malleability of the mathematical fuzzy technique. The objective of the fuzzy identification is to reach an analytic formula for a better understanding of the cause-effect relationships of the data, followed by its validation. The dynamic system identification process is performed using fuzzy clustering through the Gustafson and Kessel algorithm, followed by a Takagi and Sugeno fuzzy inference method. The k -fold technique, is the cross validation tool, used to confirm the lack of data over-training. The novelty of this approach covers mathematical and engineering features that makes this study interdisciplinary. For the mathematical contribution, there is a three-dimensional graphic interpretation of the data clustering geometry, obtained through own code computer simulations. Concerning to the engineering context, the novelty is based on the use of the fuzzy approach to the hydrogen fuel cell. Both contributions have no precedent in the literature. The results of the fuzzy identification show high reliability in terms of cross validation, making the fuzzy approach a promising tool for black-box identification. Combining this technique with others will provide powerful instrument for industrial problems.

Keywords: hydrogen fuel cell, fuzzy clustering, identification of dynamical systems, Takagi Sugeno inference method.

1 INTRODUCTION

A fuel cell uses the chemical energy of hydrogen or another fuel to cleanly and efficiently produce electricity. The first fuel cells were invented in 1838 [19]. A century later, NASA used fuel cells for programs searching to generate power for satellites and space capsules. Since then, fuel cells have been used in many other applications, being unique in terms of the variety of their potential applications. They can provide power for systems as large as a utility power station, and as small as a laptop computer. Besides, fuel cells have lower emissions than combustion engines, there are no carbon dioxide emissions, as a consequence, and no air pollutants [23]. These motivations

*Corresponding author: Ana Maria Amarillo Bertone – E-mail: anamaria@famat.ufu.br

¹Faculdade de Matemática – FAMAT, Universidade Federal de Uberlândia, 38408-902 Uberlândia, MG, Brasil.

²Faculdade de Engenharia Elétrica, Universidade Federal de Uberlândia, 38408-902 Uberlândia, MG, Brasil.
 E-mails: jefferson@iftm.edu.br; keiji@ufu.br

have encouraged Avelar [2] to study the functioning of a hydrogen fuel cell with a full-bridge elevator converter, and a single-phase inverter sinusoidal Pulse Width Modulation, manipulation of the pulse width of a square wave to control the power of a particular device. The high reliability of Avelar's software, allows us to recreate the data, used for identification in this study. The details of Avelar's simulation software is done in Section 2.

A system identification is a mathematical technique that seeks a model that reproduces, in part, the real dynamic of the data [1]. Among the analysis of linear and nonlinear systems can be included: White box, for physics modeling; gray box, for models with inputs, outputs, and additional information about the system; and black box, only composed by inputs and outputs, without any additional information. This study refers to the latest type, having as only knowledge the input and output data depending on time, besides its nonlinear behavior. There are several methods for the approach of nonlinear systems, among which could be highlighted: neural networks [21], and functional analysis [10], initially conceived by Wiener [22]. The main difference between these methods is the way that the system is considered: global or local [4]. An example of such identification method is the structure **NARX** (Nonlinear AutoRegressive with eXogenous input) [1, 13]. Local approaches attempt to divide the system into simpler subsystems, generally of the linear type. Fuzzy modeling techniques can be classified as local approximation.

The fuzzy identification used in this research is based in the fuzzy clustering of the data, combined with a Takagi and Sugeno fuzzy inference method. In 1979 Donald E. Gustafson and William C. Kessel published the article [9] presenting a modified Fuzzy C-Means traditional algorithm [5]. Nowadays this algorithm has become known as Gustafson-Kessel (GK), being the basis of a family of fuzzy clustering algorithms as described in [4, 7, 11, 14, 16, 18].

From the mathematical point of view, at the clustering stage of identification, this study has explored the geometry of the clustering qualitatively, performing graphics that are not common in the literature. At the identification's stage of fuzzy inference, the main objective is to obtain an analytic formula that is faithful to the dynamics of the hydrogen fuel cell. Besides, it applies the rigorous k -fold method to test the accuracy of the raised model. These two stages, besides their novelties, show the interdisciplinary aspect of this research, which is one of the main contributions. Another innovation is the fuzzy approach for the hydrogen fuel cell model identification. The procedure of these two stages are sketched in Section 3.

The cross-validation method known as k -fold is used for $k = 4$. This method yields 80.35% of accuracy, based in the Variance Accounted For (VAF) index, a result that demonstrates the ability of the analytic formula obtained to describe the hydrogen fuel cell system dynamic. The codification of the fuzzy clustering and fuzzy identification is developed in C language. The graphics are obtained through the connection between C and the Spyder-Python platform, open source integrated development environment. The graphics can be checked in Section 3 To complete the description of the organization of the work, in section 3, the concept of fuzzy clustering and the Takagi-Sugeno fuzzy inference method are delineated.

The organization of the study finishes with the conclusion expounded in Section 4.

2 HYDROGEN FUEL CELL SIMULATOR

The functioning of a hydrogen fuel cell with an elevator converter full-bridge, and a single-phase inverter sinusoidal PWM, is shown in the Avelar's work [2] that has created an efficient tool for the study of this type of model. The proposed model was developed for a cell of 1.2 kW and the system identification process used is least squares.

Avelar [2] has developed a model to simulate the hydrogen fuel cell connected to the grid, taking into account the effect of internal temperature. This software is used in to provide the system output, simulating the real plant operation. The main features of the simulator software are:

- the cell is capable to generate voltage between 20V and 50V (continuous).
- the voltage varies with the current drawn at its terminals being raised to 380V before being applied to the inverter module.
- all data is stored in a microcomputer.
- the plant characteristics are different to rise and drop as the heating is faster than cooling.
- a cell polymeric membrane fuel (PEMFC Nexa, manufactured by Ballard) is tested in the laboratory for static and dynamic survey to survey the model.
- all the fuel cell structure is simulated using PSIM® software, an electronic circuit simulation software package, designed specifically for use in power electronics, and motor drive simulations.
- the model presents reasonable performance to represent the plant's dynamics (current-temperature).

A process scheme applied by Avelar [2] is shown in Figure 1, and the data from this study are generated by the simulator described in [2].

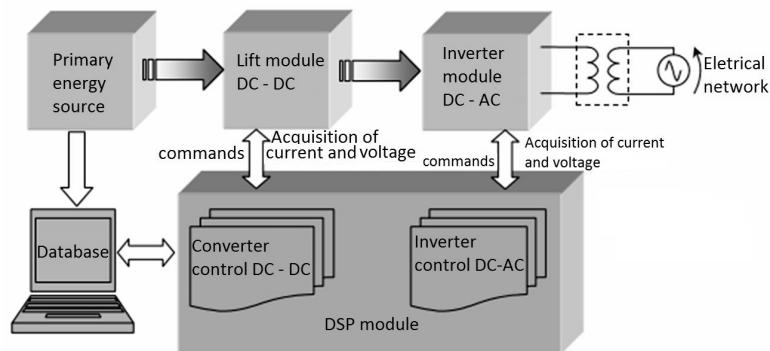


Figure 1: Alternative energy source based on fuel cell (hydrogen). Source: [2].

The type structure is SISO (single input/single output), or only one input and one output. 38,300 are input data (current) with their 38,300 output data (temperature) matching. The current values range from 24.9034A to 1.24136A, which A stands for Ampère unit; temperature values ranging from 0°C to 54.835°C. Input data has been generated in order to traverse the largest possible frequency range, with sampling rate equal 0.2 seconds. The laboratory where the collection is made [2] is shown in Figure 2.



Figure 2: The laboratory where has been obtained the fuel cell hydrogen data. Source: [2].

The gray-box identification studies seek to show important information at the moment of perform the estimation parameters; this modeling is used in this research using data software simulator [2]. It is important to emphasize the efficiency of the simulator demonstrated by tests which, in comparative performances, is shown in Figure 3.

3 RESULTS OF THE FUZZY IDENTIFICATION METHOD FOR THE HYDROGEN FUEL CELL DYNAMIC SYSTEM MODEL

As part of the fuzzy identification technique, the clustering of the data is performed. In general, clustering is a classification of non-supervised of data, forming groups of elements called clusters. The objective of this technique is the organization of the patterns represented by vector, or points, in the multidimensional space, according with a similarity measure. In the process of clustering we have the following components:

1. definition of a similarity between elements. In this case a distance induced by a symmetric matrix, positive is defined to the center of the cluster;
2. defining the criteria for the element belonging or not to the cluster. This measure could be either crisp (classic partition) or fuzzy. In the former case, the measure is a membership function and the pattern is defined by the degree of membership of the element to the cluster.

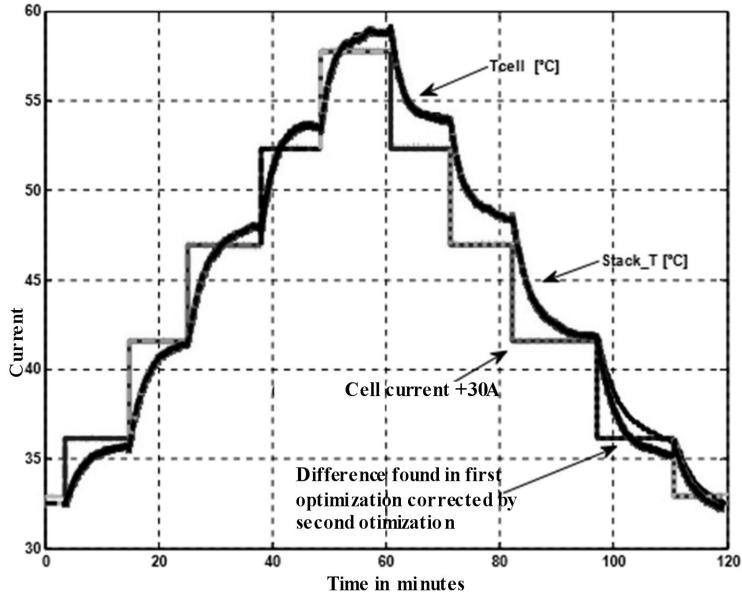


Figure 3: Comparison between the cell temperature measurement (Stack_T), and simulated (T_cell) with digitalized model. Variable parameters are used with the rise and fall of the current after the optimization model. Source: [2].

The fuzzy clustering allows that the elements of the data belong to different groups simultaneously with different membership degrees. Finally, a number c of fuzzy sets is obtained. Other requirements of the clustering process are

- Each element has to have a membership degree of each cluster;
- The sum of all membership degrees of the element to each cluster is exactly one.

Mathematically describing the process of our study, consider the vector

$$z_j = (z_{j1}, z_{j2}, z_{j3}) \in \mathbb{R}^3, j = 1, \dots, N,$$

where z_{j1} represents the input of the system, with z_{j2} corresponding to the output at the time z_{j3} , and $N = 38.300$. In the first step of clustering through the GK algorithm, consider the following elements:

- The Euclidean distance as a measure of similarity;
- A set, V_{ini} , of center of cluster which are chosen from the set $\{z_j\}_{j=1}^{38.300}$ in a way that the centers are well distributed among the data.

With those elements as a starting point, the next purpose of the clustering process is to find the entries of the matrix

$$U = \begin{bmatrix} \mu_{11} & \dots & \mu_{1k} & \dots & \mu_{1N} \\ \mu_{21} & \dots & \mu_{2k} & \dots & \mu_{2N} \\ \vdots & \vdots & \ddots & & \vdots \\ \mu_{c1} & \dots & \mu_{ck} & \dots & \mu_{cN} \end{bmatrix}.$$

For this end, it is formulate an optimization problem given by

$$\text{minimize} \sum_{i=1}^c \sum_{j=1}^N \mu_{i,j}^m d_{i,j}^2 \quad (3.1)$$

where v_i is an element of V for $i = 1, \dots, c$,

$$d_{i,j}^2 = d^2(z_j, v_i) = (z_j - v_i) I (z_j - v_i)^T, \quad i = 1, \dots, c, \quad j = 1, \dots, N, \quad (3.2)$$

and I is the identity matrix. The optimization problem has the following constrains:

$$\begin{aligned} 0 \leq \mu_{i,j} \leq 1, \quad i = 1, \dots, c, \quad j = 1, \dots, N \quad (\text{degree of membership}); \\ \sum_{j=1}^N \mu_{i,j} > 0, \quad i = 1, \dots, c \quad (\text{non emptiness of the cluster}); \\ \sum_{i=1}^c \mu_{i,j} = 1, \quad j = 1, \dots, N \quad (\text{total degree of membership}), \end{aligned} \quad (3.3)$$

where m is known as the fuzzy parameter of the clustering [4]. The number m determines how much fuzziness is wanted for the clustering. It has been taken for this study $m = 2$.

Once the optimization problem (3.1) is solved using the Lagrange Multipliers, the membership function μ_i is obtained for each cluster, which, at each element of the data z_k , is given by

$$\mu_i(z_k) = \left(\frac{d_{i,k}^2}{\sum_{j=1}^c d_{j,k}^2} \right)^{-\frac{1}{m}}, \quad i = 1, \dots, c, \quad k = 1, \dots, N. \quad (3.4)$$

As a convenient choice, the new clusters' centers are chosen as the center of mass of the membership function μ_i^m given by

$$v_i = \frac{\sum_{j=1}^N \mu_{ij}^m z_j}{\sum_{j=1}^N \mu_{ij}^m}. \quad (3.5)$$

Therefore, a set $V = \{v_1, v_2, \dots, v_c\}$, has been constructed $v_i \in \mathbb{R}^3$, that represents the new clusters' centers. Next, considering for each $i = 1, \dots, c$ the matrix A_i , $3 \times N$ and B_i , $N \times 3$, given by

$$A_i = \frac{\mu_{ij}^m}{\sum_{j=1}^N \mu_{ij}^m} (z_j - v_i)^T, \quad B_i = (z_j - v_i),$$

we define the matrix 3×3 given by $F_i = A_i \cdot B_i$, where \cdot is operator product of matrix. The matrix F_i is a 3×3 which entries represent the variance of the memberships of the elements of the cluster i through μ_{ij}^m . This matrix is known as the covariance matrix of the cluster i . Finally, the matrix that induces the new distance in the GK process is defined by

$$M_i = \det(F_i)^{\frac{1}{3}} F_i^{-1}, \quad (3.6)$$

that determines the new geometry features of the cluster with center v_i . From the matrix M_i , the updated distance measure is defined as

$$d_{i,j}^2 = (z_j - v_i) M^i (z_j - v_i)^T, \quad i = 1, \dots, c, \quad j = 1, \dots, N. \quad (3.7)$$

The procedure can be repeated, ending when an established criteria for stop, s , is set, and the error

$$\text{error} = \max \left\| U_{ij}^k - U_{ij}^{k+1} \right\|,$$

is less than s , where U_{ij}^k represents the entries of the memberships degrees of the data z_j to the cluster i at the step k of the GK algorithm.

After the fuzzy clustering process is finished, the fuzzy inference of Takagi and Sugeno (TS) [20]. An example of the fuzzy identification method is shown in Figure 4, supposing two clusters, C_1 and C_2 , obtained by the clustering process.

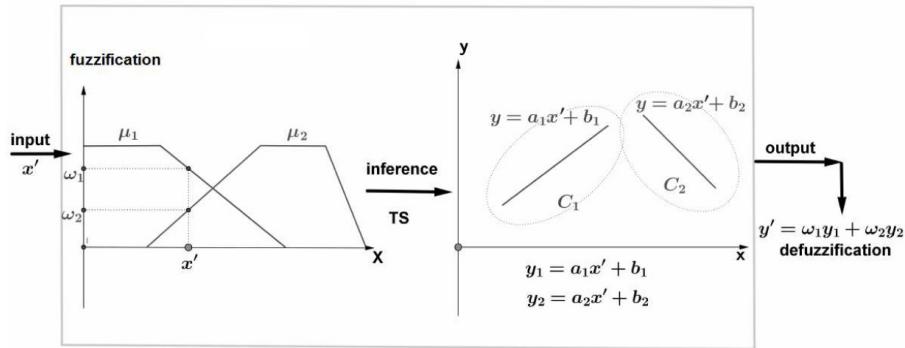


Figure 4: Identification algorithm using clustering and TS for two clusters. Source: by the authors.

Projecting the entries of the matrix U in a plane with the values of the entry as elements in the edge x , a membership function is obtained for each cluster, which is named μ_1 and μ_2 respectively for C_1 and C_2 . x' is an input element, associated to the values $\mu_1(x') = \omega_1$ and $\mu_2(x') = \omega_2$. On the other hand, the clustering process is determined by the edge in the direction of the entry data for each cluster, for instance $y_1 = a_1x + b_1$ for cluster C_1 , and $y_2 = a_2x + b_2$ for cluster C_2 . The rules established by the Takagi-Sugeno inference method are for this example

$$\text{if } (x', \mu_i(x')) \text{ then } y_i(x') = a_i x' + b_i, i = 1, 2.$$

Finally, the defuzzified output is given by the sum giving by $\sum_{i=1}^2 \omega_i y_i(x')$, which is a convex combination of the values $y_i(x')$ by the total degree of membership restriction in (3.3).

The model performance is evaluated by the value in percentage of variance (VAF, Variance Accounted For) which is given by the expression

$$\left(1 - \frac{\text{variance(real output)} - \text{variance(model output)}}{\text{variance(real output)}} \right) \times 100. \quad (3.8)$$

The closer the value (3.8) is to 100%, the better is the performance of the approximation. The decision of the best number of clusters to perform the GK algorithm, is based on the best overall VAF. Once the decision of the number of clusters is made, and a dynamic formula is obtained, next step is to test the efficiency of the approach, through the cross-validation, so call k -fold method [6]. The main objective of this method is to verify the model behavior on multiple data groups, to be used for training and testing. In this work a 4-fold is applied for the validation of the data. The procedure is given as follows:

- The data group is divided into four parts, group A, B, C, and D, composing by 9,575 inputs each, from the total of 38,300 data inputs;
- In the first step, the data groups B, C, and D are together for training, and the group A for validation. The VAF rate is calculated corresponding to this simulation. This number is calculated VAF1;
- A new arrangement of the groups A, B, C, and D is obtained, making the circulate arrangement of the letters, so that, in the second step, the group D is the validation data and the groups A, B, and C are the training data. The calculated validation rate is VAF2;
- By repeating the rearranging of the groups by circulation, the validation rates VAF3 and VAF4 are obtained;
- The arithmetic average $\frac{\text{VAF1} + \text{VAF2} + \text{VAF3} + \text{VAF4}}{4}$ is the validation rate for the model.

An illustration of the method is shown in Figure 5.

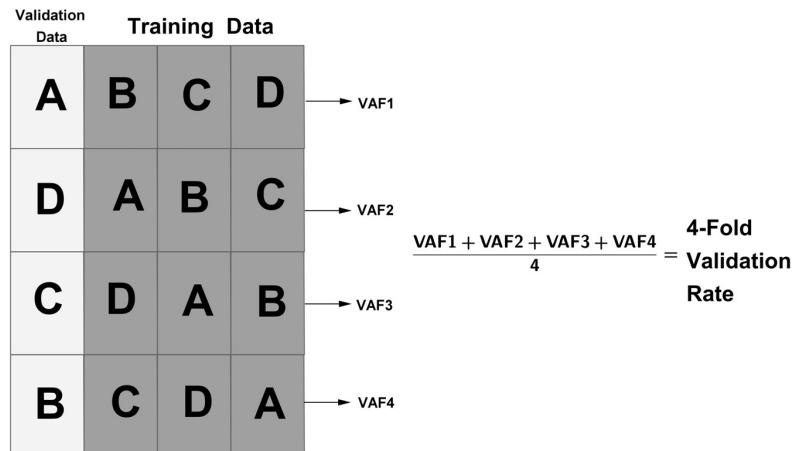


Figure 5: The procedure of the 4-fold validation. Source: by the authors.

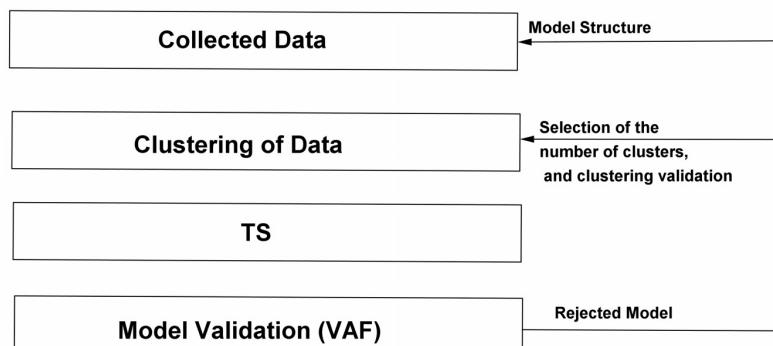


Figure 6: Complete process involving fuzzy clustering and Takagi-Sugeno inference. Source: by the authors.

Summarizing the process, the entire algorithm to obtain the fuzzy identification of the model, as has been detailed in this section, follows the steps shown in Figure 6.

The fuzzy clustering, through the GK algorithm is performed for 20 clusters, numbers are chosen in the range [3,30], due to the higher overall VAF=85.87%. It is shown in Figure 7 the three-dimensional graphic of the data (input,output, time), and the centers of the 20 clusters.

One of the clusters and its level of membership degree is shown in Figure 8. Each layer is a surface correspond to a degree of membership. For instance, the outside darkest surface is the level 0.2; the darkest surface in the center is the level 1.

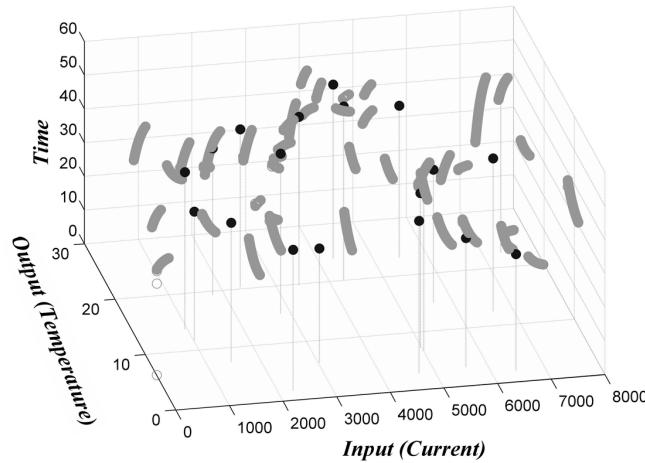


Figure 7: The graphic of the data (input, output, time) together with the centers of the 20 clusters.
Source: by the authors.

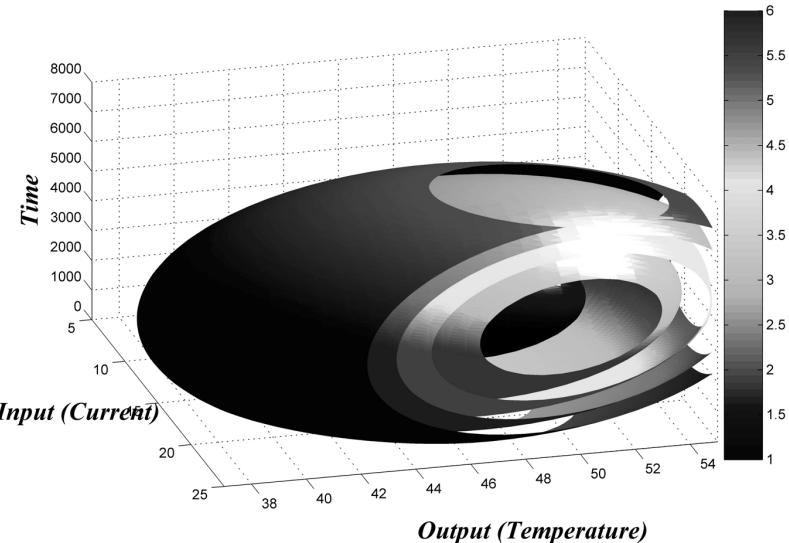


Figure 8: The graphic of first cluster constructed with the GK algorithm. The minor ellipsoid in black corresponds to the membership degree 1. The scale on the right side specifies the corresponding membership degree of the ellipsoid surface. Source: by the authors.

The 20 clusters projection on the plane (input,output) is shown in Figure 9. The different gray colors indicate the membership degree of the points to the cluster.

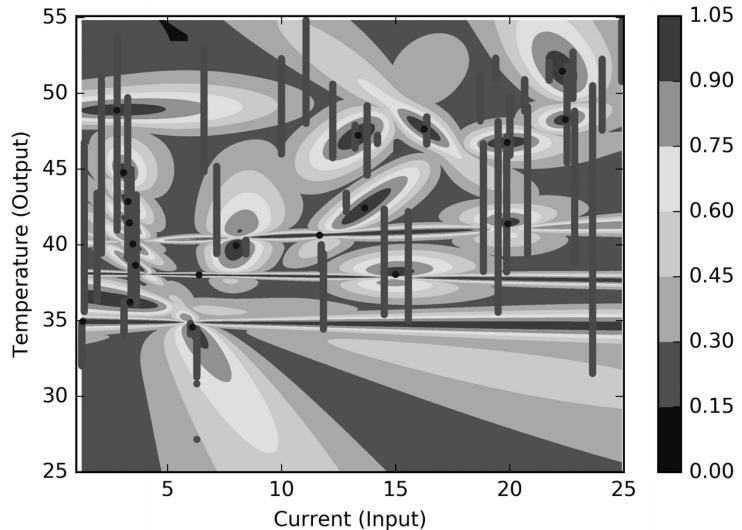


Figure 9: The result of the fuzzy clustering of the (input, output) data obtained through the GK algorithm with 20 clusters. Source: by the authors.

Once made the system identification through the algorithm, developed in C, a cross validation technique is applied. The results of the 4-fold validation technique is shown in Table 1.

Table 1: The results of the 4-fold validation technique. Source: by the authors.

VAF1	VAF2	VAF3	VAF4
85.87%	84.15%	70.68%	80.69%

The accuracy rate for the fuzzy identification procedures is the arithmetic average

$$\frac{VAF1 + VAF2 + VAF3 + VAF4}{4} = 80.35\%.$$

The fuzzy identification result compared with the real data is shown in Figure 10.

4 CONCLUSION

In order to identify simulated data of a hydrogen fuel cell dynamics, a fuzzy technique is applied. Through a fuzzy clustering, followed by a Takagi and Sugeno inference method, is generated a model that recreates the dynamics of the single input/single output system, and returns an analytic expression for the relationship between the entries of the data.

It is important to notice that in this type of system identification, black box type, there is no previous information about the model. Furthermore, the behavior that the data show is strongly

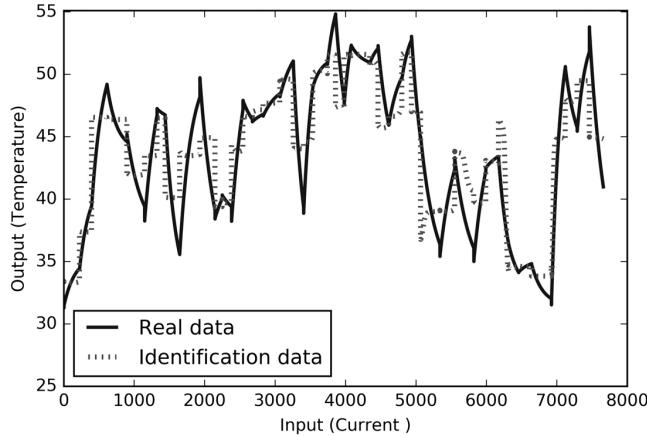


Figure 10: The graphic of the real data and the identification data corresponding to the same time. Source: by the authors.

nonlinear. Fuzzy technique has the proper pliability for these type of models, which allows to reach faithful formulas, and with capacity to describe the real phenomenon. The identification process that has been applied in this work is divided into three main parts: the fuzzy clustering, a Takagi and Sugeno inference method, and a cross validation of the type k -fold. The chosen validation, with $k = 4$, allows the elimination of over training, which reached the value VAF up to 80.35%.

It is important to highlight, as a mathematical contribution of this work, the three-dimensional graphic, which is an interpretation of the geometry of the data clustering. The engineering relevance of the study rests in the use of fuzzy identification in data from a hydrogen fuel cell. The two novel aspects of this proposal are, to the best of our knowledge, uncommon in the literature.

Fuzzy methodology has shown good performance for identification tasks, validated by the k -fold technique, specially in cases when there is little information about the process, as is the case of this study. The improvement on the codification, for better performance in accuracy and computer speed, is a research currently in progress. Furthermore, a future survey is among the purposes, involving a comparative study with other techniques and its combinations applied to same and different data.

RESUMO. Uma identificação fuzzy da dinâmica de um sistema é desenvolvida com dados gerados por um simulador de uma célula de combustível de hidrogênio. Os dados obtidos são do tipo de entrada única – saída única, sem conhecimento prévio do modelo do sistema, e mostrando um comportamento não-linear. A escolha do método fuzzy para identificação baseia-se nessas características específicas dos dados e na maleabilidade da técnica mate-

mática fuzzy. O objetivo da identificação fuzzy é chegar a uma fórmula analítica para uma melhor compreensão do relacionamento causa – efeito dos dados, seguido de sua validação. O processo dinâmico de identificação do sistema é realizado usando clusterização fuzzy através do algoritmo Gustafson e Kessel, seguido por um método de inferência fuzzy de Takagi e Sugeno. A técnica k -fold, é a ferramenta de validação cruzada, usada para evitar o excesso de treinamento dos dados. A novidade da abordagem proposta abrange características matemáticas e de engenharia que faz este estudo de caráter interdisciplinar. Como contribuição matemática, existe uma interpretação gráfica tridimensional da geometria de agrupamento de dados, obtida através de simulações com codificação própria. Relativo ao contexto de engenharia, a novidade é baseada no uso da abordagem fuzzy para a simulação da célula de combustível de hidrogênio. Ambas as contribuições não têm precedentes na literatura. Os resultados da identificação fuzzy mostram alta confiabilidade em termos de validação cruzada, tornando a abordagem fuzzy uma ferramenta promissora para identificação de modelos de caixa preta. Combinar esta técnica com outras já estabelecidas irá fornecer um instrumento poderoso para problemas industriais.

Palavras-chave: Célula de Hidrogênio, clusterização fuzzy, identificação de sistemas dinâmicos, método de Takagi Sugeno.

REFERENCES

- [1] L.A. Aguirre. *Introduction to System Identification*. UFMG publisher, Belo Horizonte, (2007).
- [2] H.J. Avelar. *Study and development of an energy system based in a hydrogen fuel cell for the power injection in electrical net*. PhD Thesis Faculty of Electrical Engineering, FEELT, Federal University of Uberlandia, MG, Brazil, (2012).
- [3] J. Abonyi, R. Babuska & F. Szeifert. Modified Gath-Geva fuzzy clustering for identification of Takagi-Sugeno fuzzy models. *IEEE Trans. on Systems, Man, and Cybernetics*, **32**(5) (2002), 612–621.
- [4] R. Babuska. *Neuro-Fuzzy Methods for Modeling and Identification*. Springer Verlag, Heidelberg, (2002).
- [5] J.C. Bezdek. *Pattern Recognition with Fuzzy Objective Function Algorithms*. Plenum Press, New York, (1981).
- [6] I.G. Costa, F.A.T. de Carvalho & M.C.P. de Souto. Comparative analysis of clustering methods for gene expression time course data. *Genetics and Molecular Biology*, **27**(4) (2004), 623–631.
- [7] D. Dovzan, V. Logar & I. Skrjanc. Implementation of an Evolving Fuzzy Model (eFuMo) in a Monitoring System for a Waste-Water Treatment Process. *IEEE Transactions on Fuzzy Systems*, **23**(5) (2015), 1761–1776.
- [8] J.C. Dunn. A fuzzy relative of the ISODATA process and its use in detecting compact well-separated clusters. *J. Cybern*, **3**(3) (1974), 32–57.
- [9] D.E. Gustafson & W.C. Kessel. Fuzzy clustering with fuzzy covariance matrix. In: *Proceedings of the IEEE Control and Decision Conference*, (1979), 761–766.

- [10] M. Franz & B. Schölkopf. A unifying view of Wiener and Volterra theory and polynomial kernel regression. *Neural Computation MIT Press Journals*, **18**(12) (2006), 3097–3118.
- [11] R. Hyde, P. Angelov & A.R. Mackenzie. Fully online clustering of evolving data streams into arbitrarily shaped clusters. *Elsevier Information Sciences*, **382-383** (2017), 96–114.
- [12] A.K. Jain & R.C. Dubes. Algorithms for Clustering Data. Prentice Hall, Englewood Cliffs, (1988).
- [13] L. Ljung. System Identification Theory for the User. Prentice Hall, Englewood Cliffs, (1987).
- [14] L. Maciel, F. Gomide & R. Ballini. Risk management using evolving possibilistic fuzzy modeling. In: Preprint submitted to XV Encontro Brasileiro de Finanças, Mackenzie Presbyterian University, (2015).
- [15] J.B. MacQueen. Some Methods for classification and Analysis of Multivariate Observations. In: Proceedings of 5th Berkeley Symposium on Mathematical Statistics and Probability, University of California Press, (1967), 281–297.
- [16] R. Pereira, V.M.F. Mendes, J. Figueiredo, J. Martins & J.C. Quadrado. Fuzzy clustering applied to a Demand Response Model in a Smart Grid Contingency Scenario. In: International Symposium on Power Electronics, Electrical Drives, Automation and Motion (SPEEDAM), (2014).
- [17] E. Ruspini. Numerical methods for fuzzy clustering. *Information Sciences* **2** (1970), 319–350.
- [18] S. Shafieezadeh-A & A. Kalhor. Evolving Takagi-Sugeno Model Based on Online Gustafson-Kessel Algorithm and Kernel Recursive Least Square Method. In: IEEE Conference on Evolving and Adaptive Intelligent Systems (EAIS), (2014).
- [19] C.F. Schoenbein. Lecture of 13 March 1839. *Berichte der Verhandlungen der naturforschenden Gesellschaft in Basel*, **4** (1839), 52–55.
- [20] T. Takagi & M. Sugeno. Fuzzy Identification of Systems and Its Applications to Modeling and Control. *IEEE Transactions on Systems, Man, and Cybernetics*, **1** (1985), 116–132.
- [21] T.A. Tutunji. Parametric system identification using neural networks. *Applied Soft Computing – Elsevier*, **47** (2016), 251–261.
- [22] N. Wiener. Nonlinear Problems in Random Theory. Wiley and MIT Press, Massachusetts, (1958).
- [23] Office of Energy Efficiency & Renewable Energy. Acess in November 2016: <http://energy.gov/eere/fuelcells/fuel-cells>.

A.4 Encontro de Biomatemática 2018

Novel method based in fuzzy clustering for EEG signal analysis

Jefferson Beethoven Martins¹, Ana Maria Amarillo Bertone²,

¹Faculdade de Engenharia Elétrica, UFU, Uberlândia, MG

²Faculdade de Matemática, UFU, Uberlândia, MG

Introduction

The signals from an **ElectroEncephaloGram**, EEG for short, are well known for being highly subjective and may appear at random in the time scale. These are the main reasons for extracting and analyzing EEG signal parameters using computers as more efficient in diagnostics. With this motivation, nonlinear methods have been proposed to extract parameters for analysis and classification of time series signals. The Higuchi fractal dimension (HFD) parameter [2] measures the complexity, and the chaotic nature of the time series signals. The sample entropy (SampEn) is a statistical parameter to quantify its predictability [4].

New methods have been proposed for the analysis and simplification of the original time series captured, for instance, by the electroencephalography procedure. Among others, the **Empirical Mode Decomposition** (EMD), developed by Huang et al. [1], has been applied to nonlinear and non stationary signal analysis. The EMD method objective to break down the signal in **Intrinsic Mode Functions** (IMF), without leaving the time domain.

A novel method is proposed in this study with the objective of extracting a single IMF function from the **Original Times Series** (OTS). This function, that we denote as ESF (extracted simplified function), matches the HFD and the SampEn parameters of the OTS, with a small absolute difference error as a validation of the method. A fuzzy clustering based in Gustafson and Kessel algorithm [3] is applied to the time series interpreted as a time-output system, that divides the temporal universe in subsets containing the highest degree of memberships information, obtained from the clustering procedure. From there, and using the Hilbert –Huang transform (HHT), the amplitude modulation bandwidth (BAM) and frequency modulation bandwidth (BFM) are extracted from analytic representation of the single IMF. Furthermore, a new procedure of fuzzy clustering is performed with the

²amabertone@ufu.br

Novel method based in fuzzy clustering for EEG signal analysis

Jefferson Beethoven Martins¹, Ana Maria Amarillo Bertone²,

¹Faculdade de Engenharia Elétrica, UFU, Uberlândia, MG

²Faculdade de Matemática, UFU, Uberlândia, MG

Introduction

The signals from an ElectroEncephaloGram, EEG for short, are well known for being highly subjective and may appear at random in the time scale. These are the main reasons for extracting and analyzing EEG signal parameters using computers as more efficient in diagnostics. With this motivation, nonlinear methods have been proposed to extract parameters for analysis and classification of time series signals. The Higuchi fractal dimension (HFD) parameter [2] measures the complexity, and the chaotic nature of the time series signals. The sample entropy (SampEn) is a statistical parameter to quantify its predictability [4].

New methods have been proposed for the analysis and simplification of the original time series captured, for instance, by the electroencephalography procedure. Among others, the Empirical Mode Decomposition (EMD), developed by Huang et al. [1], has been applied to nonlinear and non stationary signal analysis. The EMD method objective to break down the signal in Intrinsic Mode Functions (IMF), without leaving the time domain.

A ~~novel~~ method is proposed in this study with the objective of extracting a single IMF function from the Original Times Series (OTS). This function, that we denote as ESF (extracted simplified function), matches the HFD and the SampEn parameters of the OTS, with a small absolute difference error as a validation of the method. A fuzzy clustering based in Gustafson and Kessel algorithm [3] is applied to the time series interpreted as a time-output system, that divides the temporal universe in subsets containing the highest degree of memberships information, obtained from the clustering procedure. From there, and using the Hilbert –Huang transform (HHT), the amplitude modulation bandwidth (BAM) and frequency modulation bandwidth (BFM) are extracted from analytic representation of the single IMF. Furthermore, a new procedure of fuzzy clustering is perform with the

²amabertone@ufu.br

parameters obtained to conclude an effective classification of EEG signals in three types: normal, not normal or unknown.

Methods and results

The OTS used to apply the new method is a real data extracted from [5]. The results of the computational elapsed time (CET) for the fuzzy clustering process (FCP), for the calculation of the HFD and SampEn parameters of each data (OTS and ESF), the final results of this calculations, and the validation absolute error (VAE), are described in Table 1.

Table 6.2: Informative table of the results in simulations.

	CET (sec)	HFD	SampEn	VAE
FCP	2.052			
Original time series	0.053	1.41617	1.41616	0.00001
Simplify time series	2.720	0.6215	0.6217	0.0002

The graphic of the OTS together with the ESF obtained by the fuzzy clustering method, are shown at the top of Figure 1, and in the same Figure 1, both curves are shown in different plots.

EncBiomat

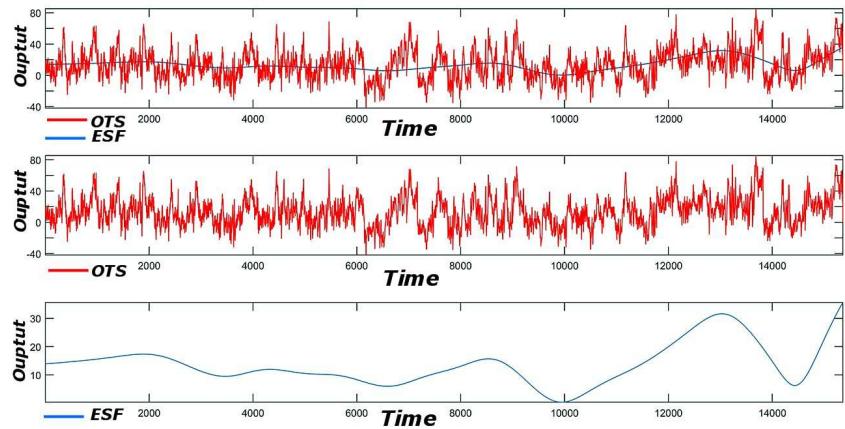


Figure 6.6: Top: OTS (red output) and ESF (blue output), middle plot: OTS; bottom plot: ESF.

Next step is to apply the HHT transform follows by the new application of the FCP for the classification of the EEG signal. This second step of the research is already complete. Due to the small space of this abstract it will be presented at the congress communication.

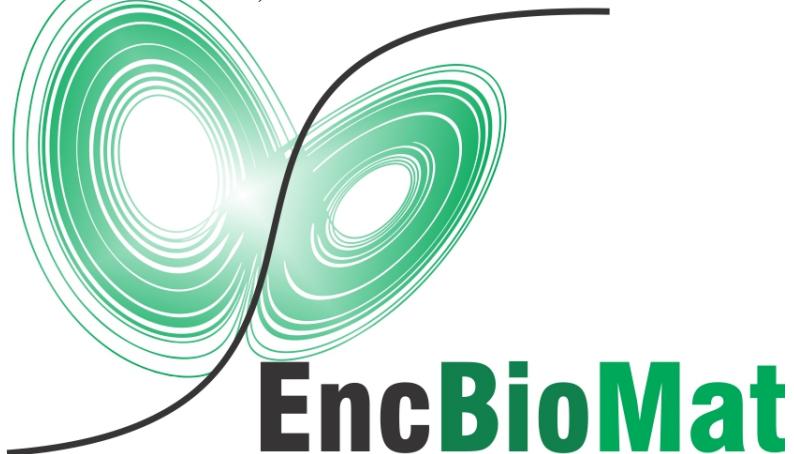
Conclusions

Promising results encourage to improve the technique of extracting a single intrinsic mode function from fuzzy unsupervised learning technique, in order to analyze and submit to the whole procedure of classification of EEG signals through the Hilbert - Huang transform. A new procedure of fuzzy clustering, is finally perform with the parameters obtained concluding in an effective classification of EEG signals in normal, not normal or unknown.

Referências

- [1] N. E., Huang, Z. Shen, S. R. Long, M. C. Wu, H. H. Shih, Q. Zheng, Yen, N. C.; Tung, C. C.; Liu, H. H. (1998). The Empirical Mode Decomposition and the Hilbert Spectrum for Nonlinear and Nonstationary Time Series Analysis. In *Proceedings of the Royal Society of London A.* 454: 903–995, 1998.

- [2] T. Higuchi. Relationship between the fractal dimension and the power law index for a time series: A numerical investigation. *Physica D*, 46, 254-264, 1990.
- [3] D. E. Gustafson and W. C. Kessel, Fuzzy clustering with fuzzy covariance matrix. In *Proceedings of the IEEE Control and Decision Conference*, San Diego, USA, 1979.
- [4] J. S. Richman, J. R. Moorman. Physiological time-series analysis using approximate entropy and sample entropy. *American Journal of Physiology. Heart and Circulatory Physiology*. 278 (6): 2039–2049, 2000.
- [5] R. Q. Quiroga, M. Schuermann. EEG, ERP and single cell recordings database. Available in: <http://https://vis.caltech.edu/rodri/data.htm>. Access in: March 11, 2017.



A.5 Apresentação na Universidade of Colorado at Denver - USA



University of Colorado
Denver

Mathematical and Statistical Sciences
Seminars on Fuzzy Theory

Chairman: Prof. Weldon Lodwick

Novel Fuzzy System Identification: Comparative Study and Application for Data Forecasting

Ana Maria A. Bertone and Jefferson Beethoven Martins
Federal University of Uberlandia



APÊNDICE B

Principal Resultado Publicado: IEEE Latin America - 2019

IEEE.org | IEEE Xplore Digital Library | IEEE-SA | IEEE Spectrum | More Sites

IEEE Xplore®
Digital Library

Institutional Sign In

Browse ▾ My Settings ▾ Get Help ▾ Subscribe

All Enter keywords or phrases (Note: Searches metadata only by default. A search for 'smart grid' = 'smart AND grid')

Journals & Magazines > IEEE Latin America Transactions > Volume: 17 Issue: 11

Novel Fuzzy System Identification: Comparative Study and Application for Data Forecasting

Publisher: IEEE Cite This

3 Author(s) Jefferson Beethoven Martins  ; Ana Maria A. Bertone ; Keiji Yamanaka View All Authors

PDF  R  E  C   

Abstract	Abstract: A novel fuzzy clustering technique for system identification based on the Takagi-Sugeno fuzzy inference is proposed. In this paper, the identification algorithm differs from others found in the literature, in the manner that the inference's antecedents and consequents are built. Gaussian fuzzy sets with support on the cluster's ordinal set of a determined α -level are unidimensional antecedents that have as a consequent an affine function, which recovers the attributes of the collected data. Two simulations examples are performed to compare to the new method: the comparisons are done with well-known algorithms in terms of the normalized root mean square goodness of the fit measure, and the computational speed in the identification process. For that purpose, Matlab System Identification toolbox is used in order to consolidate a unique source of comparison. The method has been tested in the prediction field, as part of a project supporting Brazilian local companies.
Authors	
Keywords	

Published in: IEEE Latin America Transactions (Volume: 17 , Issue: 11 , November 2019)

Page(s): 1793 - 1799 DOI: 10.1109/TLA.2019.8986416

Date of Publication: November 2019  Publisher: IEEE

Electronic ISSN: 1548-0992

Novel Fuzzy System Identification: Comparative Study and Application for Data Forecasting

J. Martins, A. Bertone, and K. Yamanaka

Abstract—A novel fuzzy clustering technique for system identification based on the Takagi-Sugeno fuzzy inference is proposed. In this paper, the identification algorithm differs from others found in the literature, in the manner that the inference's antecedents and consequents are built. Gaussian fuzzy sets with support on the cluster's ordinal set of a determined α -level are unidimensional antecedents that have as a consequent an affine function, which recovers the attributes of the collected data. Two simulations examples are performed to compare to the new method: the comparisons are done with well-known algorithms in terms of the normalized root mean square goodness of the fit measure, and the computational speed in the identification process. For that purpose, Matlab System Identification toolbox is used in order to consolidate a unique source of comparison. The method has been tested in the prediction field, as part of a project supporting Brazilian local companies.

Index Terms—Fuzzy clustering, System identification, Takagi-Sugeno inference.

I. INTRODUÇÃO

A IDENTIFICAÇÃO de sistemas não-lineares é uma tarefa importante e desafiadora [11]. Numerosos problemas reais são oriundos de processos não-lineares, cuja modelagem exige a utilização de sistemas adaptativos capazes de reproduzir o processo de forma flexível e robusta.

A identificação *fuzzy* de sistemas é empregada nesta pesquisa por meio de clusterização *fuzzy* e, em seguida, é aplicada a inferência de Takagi-Sugeno, dois processos distintos e bem definidos. Relacionado a este estudo, o modelo de Takagi-Sugeno (T-S) [23-25] é combinado com técnicas de clusterização *fuzzy* e tem se mostrado eficiente para lidar com sistemas complexos [1,3,12-17,26]. O método de clusterização escolhido é o algoritmo de Gustafson e Kessel (G-K) [11], o qual estende o conhecido *fuzzy C-means*, através do uso de uma distância adaptativa. O algoritmo G-K é a base para outros algoritmos de agrupamento, como descrito em [5,11].

Este trabalho foi desenvolvido nas faculdades de Engenharia Elétrica e Matemática da UFU.

São os autores: J. B. Martins, Faculdade de Engenharia Elétrica - Universidade Federal de Uberlândia, Uberlândia, Minas Gerais, Brasil. jefferson@ifm.edu.br; A. M. A. Bertone, Faculdade de Matemática - Universidade Federal de Uberlândia, Uberlândia, Minas Gerais, Brasil. amabertone@ufu.br e K. Yamanaka, Faculdade de Engenharia Elétrica - Universidade Federal de Uberlândia, Uberlândia, Minas Gerais, Brasil. keiji@ufu.br.

Os algoritmos de clusterização têm o potencial de evidenciar estruturas subjacentes, razão pela qual são explorados nas pesquisas de análise de dados [1,8, 13,14,22].

O processo de construção de um modelo visa determinar quantos parâmetros serão necessários para atingir uma capacidade representativa do fenômeno estudado. De forma geral, quanto menor o número de parâmetros, mais valorizada é a modelagem. Além disso, deve ser considerado também o esforço computacional requerido para a obtenção do modelo. Todavia, o que se percebe é que o segundo elemento citado não tem sido levado em conta na maioria dos resultados encontrados na literatura.

Um dos objetivos deste estudo é equalizar esses dois aspectos importantes, buscando uma obtenção de uma modelagem matemática simplificada em um tempo computacional aceitável.

O processo se inicia através do algoritmo G-K que é executado para determinar o centro dos clusters e a matriz de pertinências. Em seguida, conjuntos *fuzzy* que representam os antecedentes de T-S, são construídos através de uma inédita projeção das pertinências mais altas sobre o eixo ordinal dos dados amostrados. Este passo do algoritmo é feito com o objetivo de evitar a utilização das clássicas funções Gaussianas multivariadas. Posteriormente, os consequentes da inferência recuperam na sua fórmula afim, os atributos dos dados amostrados, representados pelas variáveis de entrada do experimento analisado. A finalização é efetuada com a defuzificação do sistema de inferência T-S, otimizada pelo método dos mínimos quadrados.

Por fim, alguns exemplos são apresentados para ilustrar a viabilidade e eficácia do método proposto. De fato, a presente técnica de clusterização *fuzzy* do tipo T-S (abreviada por CFTS) é aplicada a dois grupos de dados com distintas quantidades de amostragem e características, com o objetivo de testá-las e compará-las em diferentes cenários. Renomadas estruturas de representação de modelos como o *nonlinear autoregressive moving average with exogenous inputs* (NARMAX) [7] e o método de Hammerstein e Wiener [28], são utilizados para confrontar as particularidades de cada ferramenta. A razão de escolher essas duas técnicas consagradas é que esses métodos fornecem uma fórmula analítica para a modelagem.

Muitas outras comparações têm sido feitas ao longo desta pesquisa que, por razões de espaço desta apresentação, não estão sendo discutidas neste trabalho. Estas outras comparações realizadas incluiram dois tipos de métodos de redes neurais e

dados extraídos de *benchmarks*, com o objetivo de validar a técnica em diferentes tipos de sistemas, desde séries temporais até sistemas de várias entradas e várias saídas.

Como meio de comparação utiliza-se o *toolbox* de identificação de sistemas do Matlab® [18], o qual possui a rotina *compare*, que emprega a raiz quadrada do erro médio quadrático normalizado (*normalized root mean square error* em inglês) (NRMSE) como medida de acurácia. A escolha da ferramenta *compare* foi feita com o ideal de consolidar um único meio de comparação.

Apesar do uso do Matlab, todo o processo do método proposto é desenvolvido com código próprio que proporciona independência e permite melhorias futuras. Como pesquisa ainda em curso, a nova ferramenta foi aplicada na área de predição, com o objetivo de empregar a técnica proposta e testar sua eficiência na área. De fato, recentemente alguns métodos presentes na literatura [2,8,9,20,21], se baseiam em técnicas de redes neurais ou *fuzzy* para fazer previsões em diversas áreas, como: preço de ações, clima, energia, entre outros. Por exemplo, Voronim [27] apresenta um método que prevê preços e picos normais no mercado de eletricidade; Dovzan [9] propõe um modelo *fuzzy* para empresas que fornecem eletricidade. O método apresentado neste trabalho é aplicado a um processo de predição de energia [19].

De maneira sucinta, as principais contribuições desta pesquisa são a inovação na construção do modelo de antecedentes e consequentes de T-S, ganho no esforço computacional do procedimento completo, código próprio que permite evolução da técnica e a possibilidade comprovada de ser aplicada à predição, entre outras áreas. Além dessas vantagens, a metodologia está sendo utilizada como ferramenta de predição em projetos multidisciplinares, envolvendo companhias brasileiras privadas especializadas em venda online.

A organização do trabalho segue a seguinte estrutura: após uma breve revisão sobre clusterização *fuzzy* e os fundamentos de inferência T-S, apresenta-se o algoritmo para construir o conjunto de regras “se-então”, assunto desenvolvido na Seção 2. Simulações de comparação com dois métodos de identificação de sistemas são apresentadas na Seção 3. Uma aplicação do método à predição é desenvolvida na Seção 4. Na Seção 5 conclui-se o artigo e sugere-se questões para pesquisas futuras.

II. ALICERCE MATEMÁTICO

Esta seção fornece uma breve revisão sobre as ferramentas necessárias para respaldar a nova metodologia elaborada neste estudo.

A. Algoritmo de Clusterização de G-K

Seja um conjunto Z de dados amostrados cujos elementos são os vetores linha z_i da forma:

$$z_i = (i, x_{i1}, x_{i2}, \dots, x_{in}; output_i)^T, \quad (1)$$

$i = 1, 2, \dots, N$, N é o número total de dados amostrados. O procedimento de clusterização inicia-se resolvendo um problema de otimização cuja função objetivo é dada por:

$$K(Z; U, V, M) = \sum_{j=1}^c \sum_{i=1}^N \mu_{ji}^m d_{ji, M^j}^2(z_i, v_j), \quad (2)$$

em que m é conhecido como o parâmetro de fuzificação [4]; acompanhada das restrições relativas ao agrupamento *fuzzy*:

$$\begin{aligned} 0 \leq \mu_{ji} \leq 1, & j = 1, \dots, c, \quad i = 1, \dots, N \\ \sum_{j=1}^c \mu_{ji} &= 1, \quad i = 1, \dots, N, \end{aligned} \quad (3)$$

em que μ_{ji} é o grau de pertinência de z_i em relação ao cluster C_j e c é o número de clusters. Observa-se que a primeira restrição está relacionada ao grau de pertinência do ponto em relação a cada cluster e a última restrição é a soma total de todos os graus de pertinência relacionados a um mesmo ponto. Na função objetivo, U representa a matriz de pertinência:

$$U = \begin{bmatrix} \mu_{11} & \dots & \mu_{1i} & \dots & \mu_{1N} \\ \mu_{21} & \dots & \mu_{2i} & \dots & \mu_{2N} \\ \vdots & & \vdots & & \vdots \\ \mu_{j1} & \dots & \mu_{ji} & \dots & \mu_{jN} \\ \vdots & & \vdots & & \vdots \\ \mu_{c1} & \dots & \mu_{ci} & \dots & \mu_{cN} \end{bmatrix}, \quad (4)$$

V representa o conjunto de centros de clusters e M representa a geometria do cluster, definindo a distância entre os pontos de Z e os elementos de V da maneira subsequente:

$$d_{ji}^2 = (z_i - v_j) M^j (z_i - v_j)^T, \quad j = 1, \dots, c, \quad i = 1, \dots, N. \quad (5)$$

Os pontos estacionários da função objetivo da Eq. (2) sob as restrições (3) podem ser encontrados usando o método dos multiplicadores de Lagrange. Efetivamente, em [1] mostra-se que se $m > 1$, então U é um extremo de K se e somente se:

$$\mu_{ji} = \left(\frac{d_{ij}^2}{\sum_{j=1}^c d_{ij}^2} \right)^{\frac{1}{m}}, \quad j = 1, \dots, c, \quad i = 1, \dots, N, \quad (6)$$

e a fórmula para os centros de cluster é:

$$v_j = \frac{\sum_{i=1}^N \mu_{ji}^m z_i}{\sum_{i=1}^N \mu_{ji}^m}. \quad (7)$$

Os multiplicadores de Lagrange proveem mais informações como a expressão de M^j sendo:

$$M^j = \det(F_j)^{\frac{1}{n}} F_j^{-1}, \quad j = 1, \dots, c, \quad (8)$$

em que F_j é a matriz de covariância do j -ésimo cluster, definida por:

$$F_j = (\sum_{i=1}^N \mu_{ji}^m)^{-1} \cdot \sum_{i=1}^N \mu_{ji}^m (z_i - v_j) \cdot (z_i - v_j)^T. \quad (9)$$

Para iniciar o algoritmo de G-K, considera-se $m=2$ [1] e um erro de parada e . A seguir é construído o conjunto V , de centros iniciais de clusters, tomando valores do próprio conjunto de dados Z , igualmente espaçados. Na sequência, é realizado o cálculo das distâncias entre o conjunto de dados e o centro de cada cluster, através da distância Euclidiana, ou seja M^j é a matriz identidade I , para todo j . Da Eq. (6), obtém-se a matriz U_0 e, como consequência, os novos centros de clusters definidos pela fórmula da Eq. (7). A partir daí, obtém-se a nova geometria do cluster determinado pela matriz M^j da Eq. (8) – Eq. (9). Ao finalizar, é obtida a nova matriz de pertinências U voltando à Eq. (4). As etapas do algoritmo G-K são retomadas no fluxograma mostrado na Figura 3.

B. Modelo Fuzzy T-S

O sistema de inferência T-S é baseado em regras “se-então”, em que o antecedente é um conjunto fuzzy e o consequente é uma função afim. Das informações das entradas da matriz U da Eq. (4) são construídos os antecedentes do sistema de inferência T-S, A_j , considerando o seguinte conjunto:

$$O_j = \{(i, \mu_{ji})\}_{i=1}^N, j = 1, \dots, c \quad (10)$$

o qual é ajustado por uma função do tipo Gaussiana. É salutar observar que O_j é composto por pares, cujo primeiro componente, i , é o lugar que ocupa o ponto z_i no conjunto Z (ordinal). O segundo componente do par é o grau de pertinência de z_i em relação ao cluster j . Considerando o universo $I = \{1, 2, \dots, N\}$ e o número α , é construído o nível- α do conjunto fuzzy A_j , ou seja, o conjunto definido por:

$$[A_j]^\alpha = \{i \in I, \mu_{ji}(z_i) \geq \alpha\}, \text{ para } 0 < \alpha \leq 1, \quad (11)$$

e para $\alpha = 0$ o fecho topológico do conjunto suporte de A_j definido por $\{i \in I, \mu_{ji}(z_i) > 0\}$.

Para melhor entendimento geométrico do processo de obtenção dos antecedentes do sistema de inferência, visualiza-se na Fig. 1 a construção de um plano à altura α que determina a projeção do nível α da função de pertinência multivariada no plano vertical, determinado pelo eixo ordinal e o eixo das pertinências.

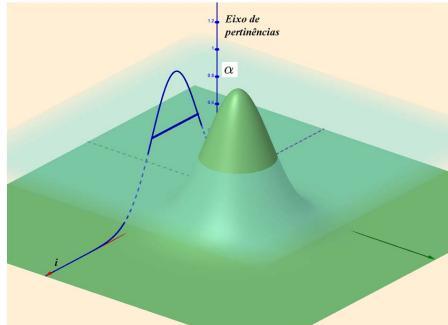


Fig. 1. O plano, à altura α , corta a função de pertinência e é feita a projeção sobre o plano vertical determinado pelo eixo temporal e o eixo das pertinências.

Essa projeção é feita em forma computacional reunindo no conjunto O_j as primeiras coordenadas dos vetores z_i que possuem pertinência maior ou igual a um número α no corresponde cluster j . Esse conjunto está representado pelo segmento azul no plano $i\mu_{ji}$, que será a base para a construção da função gaussiana antecedente. O valor de α é determinado através de um algoritmo (modificado) de ajuste dos dados a uma função do tipo Gaussiana [10] O_j a uma curva gaussiana [10], obtendo a fórmula analítica:

$$\mu_{A_j}(x) = \exp\left(-\frac{(x-v_j)^2}{2\sigma^2}\right), j = 1, \dots, c, x \in \mathbb{R} \quad (12)$$

Na Fig. 2 é mostrado o ajuste de um dos conjuntos O_j , em uma das simulações da Seção 3.

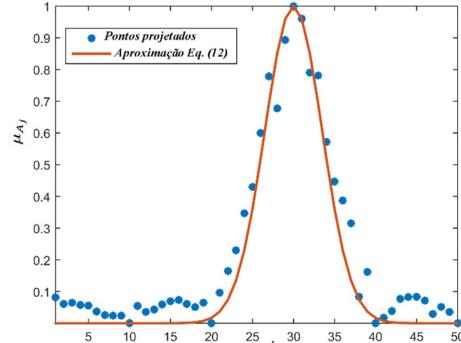


Fig. 2. Conjunto de dados O_3 representado pelos pontos de cor azul para o exemplo 2 da Seção 3, ajustado pela função Gaussiana dada na Eq. (12).

As regras fuzzy são estabelecidas a seguir:

Regraj: Se i é A_j então

$$y^j(i) = a_{i0} + a_{i1}i + a_{i2}x_{i1} + \dots + a_{i(n+1)}x_{in}, i \in I, \quad (13)$$

em que $x_{i1}, x_{i2}, \dots, x_{in}$ são os atributos dos dados amostrados, $Regraj$, $j = 1, \dots, c$ é a j -ésima regra fuzzy, c é o número de regras, y^j a saída do j -ésimo sistema linear do modelo fuzzy. Para calcular $[1 \ a_{i0} \ a_{i1} \ \dots \ a_{i(n+1)}]^T$, é utilizado o método dos mínimos quadrados para minimizar:

$$E = \min_i \left\{ \frac{1}{2} \| \hat{y}(i) - \text{saída}(i) \|^2 \right\} \quad (14)$$

no qual:

$$\hat{y}(i) = \frac{\sum_{j=1}^c \mu_{A_j}(i) \cdot y^j(i)}{\sum_{j=1}^c \mu_{A_j}(i)}, \quad i \in I. \quad (15)$$

O resultado final do modelo fuzzy é alcançado pela fórmula analítica:

$$\hat{y}(x) = \frac{\sum_{j=1}^c \mu_{A_j}(x) \cdot \hat{y}(x)}{\sum_{j=1}^c \mu_{A_j}(x)}, \quad x \in \mathbb{R}. \quad (16)$$

Em resumo, o algoritmo é descrito no fluxograma da Fig. 3.

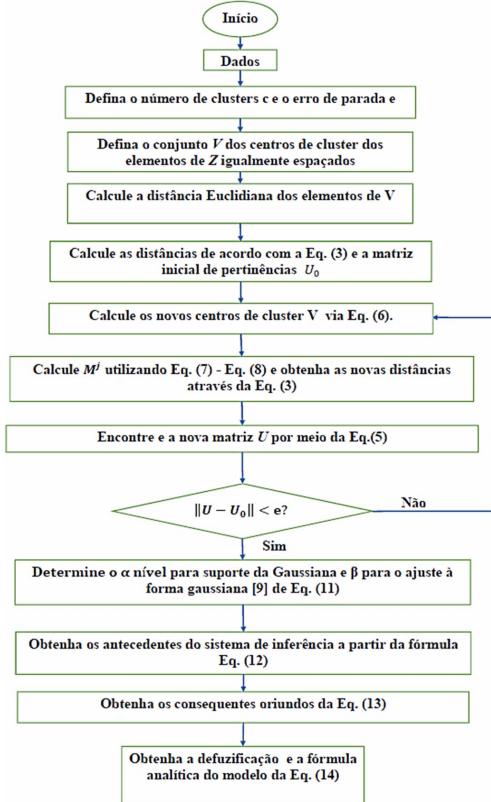


Fig. 3. Fluxograma do método de identificação proposto.

III. COMPARAÇÃO COM OUTRAS METODOLOGIAS

Nesta seção são mostrados os resultados das simulações feitas no Matlab em que se compararam a nova técnica com as estruturas de representação de modelos *NARMAX* e de Hammerstein-Wiener.

A. Simulações

A simulação é efetuada empregando a rotina *compare*, que retorna a resposta simulada de um modelo de sistema dinâmico. A rotina também exibe o gráfico da identificação e a medida de acurácia *NRMSE*. A tabela de comparação é formulada em dois viéses: a *NRMSE* e o Tempo de Execução Computacional (TEC). As comparações foram feitas através de código próprio, alcançado em uma máquina única de comparação e tempo medido em segundos. O *NRMSE* é definido como:

$$NRMSE = 100 \cdot \left(1 - \frac{\|dados - approx\|^2}{\|dados - média(dados)\|^2} \right), \quad (17)$$

em que *dados* é um elemento de *Z*, *approx* é o resultado da identificação, *média* é a média aritmética dos dados amostrados

e $\|\cdot\|$ é a norma Euclidiana do espaço real. O valor do *NRMSE* obtido representa a percentagem de acerto na identificação.

Dois tipos de base de dados são utilizados para comparar a técnica apresentada com as duas outras metodologias:

- (1) Uma série temporal, chamada neste estudo de *energy*, com informações reais de geração de energia do Brasil [19];
- (2) dados MISO coletados a partir de uma célula a combustível de hidrogênio simulada [4].

Para a construção do modelo Auto-Regressivo utilizou-se as funções *idnlarx* e *nlarx*. Para o modelo Hammerstein-Wiener utilizou-se as funções *idnlhw* e *nlhw*. Como o desempenho das três técnicas depende do número de parâmetros utilizados na modelagem, escolheu-se o mesmo número para todas as técnicas. No caso (1) o número de parâmetros é de 34 e no caso (2) o número de parâmetros é de 41.

(1) O banco de dados *energy* apresenta uma série temporal, na qual a entrada é representada pelo mês e a saída é a energia gerada nas regiões sudeste e centro-oeste do Brasil de 2000 a 2016. O Operador Nacional do Sistema Elétrico, entidade vinculada ao governo brasileiro, fornece, neste conjunto de dados, 204 valores medidos em GWh [19].

Na Tabela I são destacadas a acurácia na medida *NRMSE* e o TEC dos testes comparativos. O gráfico com o resultado alcançado através do método *NARMAX* é mostrado na Fig. 4. Na Fig. 5 é exibido a resposta do método de Hammerstein e Wiener e na Fig. 6 a da metodologia *fuzzy*.

TABELA I
TABELA COMPARATIVA PARA O BANCO DE DADOS ENERGY

NARMAX		HW		CFTS	
NRMSE	TEC	NRMSE	TEC	NRMSE	TEC
67,129%	0,974 s	64,143%	4,122 s	75,715%	0,0225 s

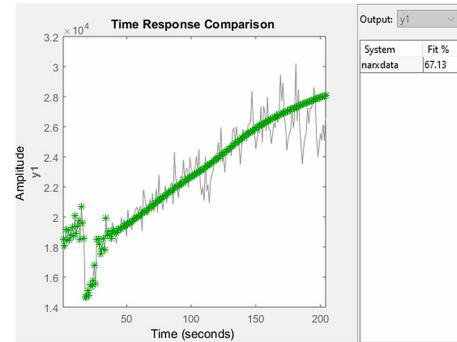


Fig. 4. Gráfico e desempenho do *NRMSE* demonstrados através da rotina *compare*, referente ao banco de dados *energy* sob a ação da metodologia *NARMAX*, de nome *narxdata* expresso na figura. No quadro superior à direita, com o nome *Fit %* é mostrado o *NRMSE* da simulação.

(2) O banco de dados que, neste estudo foi chamado de *hydrocell*, surgiu a partir do trabalho de Avelar [4]. O *hydrocell*, simula uma célula de combustível a hidrogênio que considera o efeito da temperatura interna. Toda a estrutura da célula de combustível foi simulada no software PSIM® [4]. Neste

ambiente a célula é capaz de gerar tensão contínua entre 20V e 50V; a tensão varia conforme a corrente consumida nos seus terminais, sendo elevada para 380V antes de ser aplicada ao módulo inversor. As características da planta são diferentes diante do aumento e da queda de temperatura, pois o aquecimento é mais rápido que o resfriamento. A célula é do tipo *PEMFC Nexa*, fabricada pela Ballard, sendo testada em laboratório para levantamento estático e dinâmico, propiciando assim o modelo obtido [4].

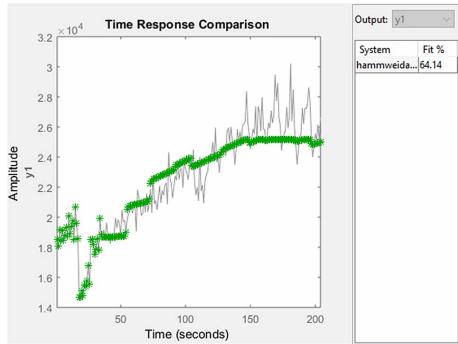


Fig. 5. Gráfico construído no *compare*, referente ao *energy* utilizando o método de Hammerstein e Wiener, de nome *hammweidata* expresso na figura.

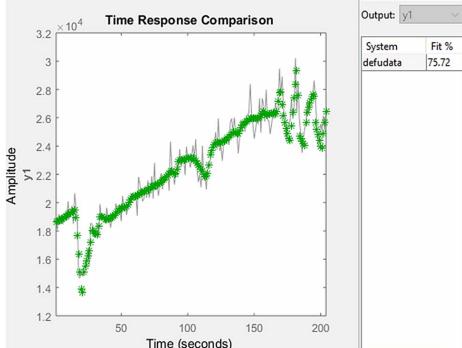


Fig. 6. Gráfico construído no *compare*, referente ao *energy* utilizando a técnica proposta, de nome *defudata* expresso na figura.

A modelagem clássica foi explorada com o toolbox de identificação *fuzzy* de Babuska [5], resultando em um trabalho de identificação primária publicado em [6]. A partir desse trabalho, que incluiu simulações com a técnica *n-fold* [6], foi proposto como objetivo a melhoria do *toolbox* de identificação de Babuska, principalmente em termos do esforço computacional, ganhando também em acurácia.

Na Tabela II são destacadas as médias *fit* e *tc*, como descritas anteriormente, dos testes comparativos. Os gráficos com os resultados alcançados através do *toolbox* de identificação do Matlab para o método *NARMAX* são mostrados na Fig. 7. Na Fig. 8 é mostrado o resultado do método de Hammerstein e Wiener e na Fig. 9 o da metodologia proposta neste estudo.

TABELA II
TABELA COMPARATIVA PARA O BANCO DE DADOS
HYDROCELL – EXTRAÍDA DO MATLAB

NARMAX		HW		CFTS	
NRMSE	TEC	NRMSE	TEC	NRMSE	TEC
11,113%	3,858 s	53,382%	12,79 s	92,346%	0,8289 s

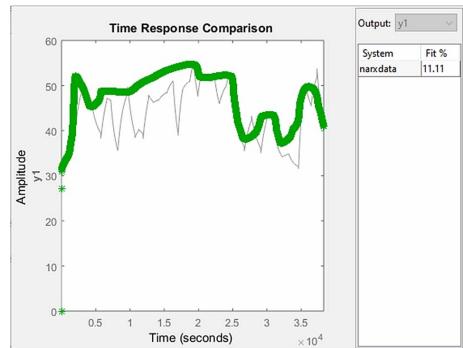


Fig. 7. Gráfico e desempenho do *NRMSE* demonstrados através da rotina *compare*, referente ao banco de dados *hydrocell* sob a ação da metodologia *NARMAX*, de nome *naridata* expresso na figura. No quadro superior à direita, com o nome *Fit %* é mostrado o *NRMSE* da simulação.

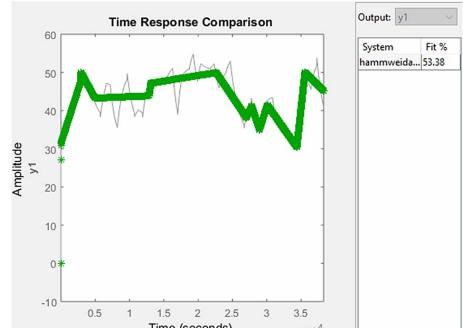


Fig. 8. Gráfico construído no *compare*, referente ao *hydrocell*, utilizando o método de Hammerstein e Wiener, de nome *hammweidata* expresso na figura.

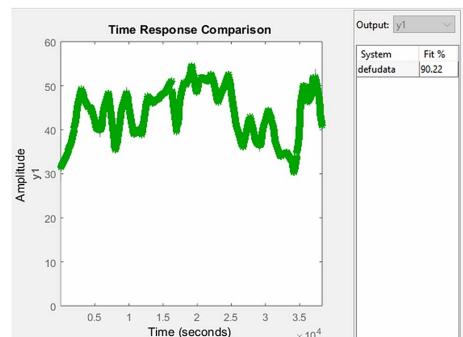


Fig. 9. Gráfico construído no *compare*, referente ao *hydrocell*, utilizando o método de Hammerstein e Wiener, de nome *defudata* expresso na figura.

Destaca-se que, em termos de qualidade, baseando-se nas medidas NRMSE e TEC, a metodologia proposta neste trabalho apresentou o melhor desempenho em todos os casos.

IV. APLICAÇÃO PARA PREDIÇÃO

O banco de dados *energy* é utilizado para predição como aplicação do novo método proposto neste estudo. Os dados são divididos em *passado*, compostos pelos primeiros 104 dados, e o restante dos dados é considerado *futuro*, a fim de validar os dados previstos em confronto com os dados reais. Para que a predição ocorra é necessário realizar a identificação fuzzy dos dados do *passado*, com o objetivo de detectar o padrão a ser previsto. De fato, a partir dos dados do *passado* e usando a fórmula analítica da Equação (18), $f(t)$ é definido como padrão do intervalo $[t_1, t_p]$, em que t_1 é a primeira semana e t_p a última semana coletada. A predição no instante t_{p+1} , é o valor de $f(t_{p+1})$. A seguinte predição é feita incorporando o dado real da semana t_{p+1} , a fim de realimentar o processo com um novo passado e continuar a iteração. A medida de acurácia da predição utilizada neste experimento [18] é dada por:

$$P = 1 - \frac{|\text{valor real} - \text{valor predito}|}{\text{valor real}}. \quad (18)$$

A Fig. 10 mostra o resultado de 30 semanas de previsão, com os detalhes na Tabela III.

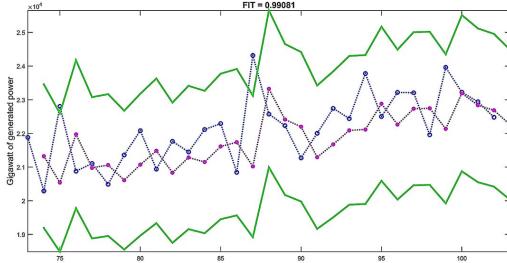


Fig. 10. Gráfico de trinta semanas de previsão do banco de dados de *energy*. Os pontos azuis são os 30 valores reais das semanas e os pontos em cor magenta são os valores previstos. As bandas verdes representam os limites de 0,90 da medida P da fórmula (18).

TABELA III
RESULTADOS DA ACURÁCIA (P) PARA 30 SEMANAS DE
PREVISÃO SEMANAL

Semana	P	Semana	P
1	0,95	2	0,9
3	0,95	4	0,99
5	0,97	6	0,97
7	0,95	8	0,97
9	0,96	10	0,99
11	0,96	12	0,97
13	0,96	14	0,86
15	0,97	16	0,99
17	0,96	18	0,97
19	0,95	20	0,98
21	0,93	22	0,98
23	0,96	24	0,98
25	0,96	26	0,92
27	0,99	28	0,99
29	0,99	30	0,95

Nota-se que a precisão, em geral, da tabela em que, em 30 semanas, só aparece um dado abaixo de 0,90 de acerto no sentido da medida P. Semanas seguidas de alto acerto na medida P correspondem a semanas de dados estáveis do *energy*. Salienta-se que este não tem sido o único experimento de previsão efetuado ao longo desta pesquisa, sendo testado em outros bancos de dados coletados pelos autores e provenientes de *benchmarks*.

V. CONCLUSÃO

A estrutura da identificação de sistemas através de modelos do tipo Takagi-Sugeno é discutida nesta pesquisa, todavia há a introdução de um novo método para a construção dos antecedentes e dos consequentes das regras fuzzy.

A partir do método *fuzzy* de clusterização de Gustafson e Kessel, extrai-se a informação da amplitude de um α -nível correspondente a cada cluster. Diante disso, uma função Gaussiana unidimensional é definida como antecedente da regra, como projeção da correspondente Gaussiana multivariável. O consequente é uma função afim que incorpora os atributos dos dados amostrados, recuperando a informação qualitativa da modelagem na saída do sistema de inferência. Produto da defusificação do modelo de Takagi-Sugeno, surge a fórmula analítica que interpreta quantitativamente o fenômeno estudado.

A nova abordagem tem sido testada por meio de dados coletados pelos autores e por dados extraídos de uma companhia brasileira pública do setor de energia, comparada com duas estruturas de representação de modelos que também fornecem fórmulas analíticas.

Como resultado da comparação, a técnica demonstrou melhor desempenho em todos os experimentos, em função de duas medidas comparativas: a raiz quadrada do erro médio quadrático normalizado e o esforço computacional.

Além disso, o método *fuzzy* concebido nesta pesquisa é atrativo para outras áreas como, por exemplo, predição de dados. A experimentação com resultados convincentes abre ramificações para trabalhos futuros na área de predição; assim como em outras em que a ferramenta de identificação de sistemas é imprescindível.

REFERÊNCIAS

- [1] J. Abonyi, R. Babuska and F. Szeifert, *Modified Gath-Geva fuzzy clustering for identification of Takagi-Sugeno fuzzy models*, IEEE Trans. on Systems, Man, and Cybernetics, pp. 612 – 621, 2002.
- [2] A.Y. Alanis, *Electricity Prices Forecasting using Artificial Neural Networks*, IEEE Latin American Transactions, Vol 16, No 1, pp. 105 – 111, 2018.
- [3] P. P. Angelov, D. P. Filev, *An approach to online identification of Takagi-Sugeno fuzzy models*, IEEE Transactions on Systems, Man, and Cybernetics – Part B: Cybernetics, Vol. 34, No.1, pp. 484 – 498, 2004.
- [4] H.J. Avelar, *Estudo e desenvolvimento de um sistema de energia baseado em célula a combustível para injeção de potência na rede elétrica*, Tese de doutorado da Faculdade de Engenharia Elétrica, Universidade Federal de Uberlândia, MG, Brazil, 2012.
- [5] R. Babuska, *Neuro-Fuzzy Methods for Modeling and Identification*, Springer Verlag, Heidelberg, 2002.
- [6] A. M. A. Bertone, J. B. Martins and K. Yamanaka, *Black-box fuzzy identification of a nonlinear hydrogen fuel cell model*, Journal Trends in Applied and Computational Mathematics, Vol. 18, No.13, 2017.

- [7] S. A. Billings, *Nonlinear system identification: NARMAX methods in the time, frequency, and spatio-temporal domains*. Hoboken, New Jersey: John Wiley & Sons Ltd, 2013.
- [8] S.M. Chen, K. Tanuwijaya, *Multivariate fuzzy forecasting based on fuzzy time series and automatic clustering techniques*, Expert Systems with Applications, Vol. 38, pp. 10594 – 10605, 2011.
- [9] D. Dovzan, V. Logar, I. Skrjanc, *Implementation of an Evolving Fuzzy Model (eFuMo) in a Monitoring System for a Waste-Water Treatment Process*, IEEE Transactions on Fuzzy Systems, Vol. 23, No.5, pp. 1761 – 1776, 2015.
- [10] H. Guo, *A Simple Algorithm for Fitting a Gaussian Function* IEEE Signal Processing Magazine Vol 28, No 5, pp. 134-137, 2011.
- [11] D.E. Gustafson, W.C. Kessel, *Fuzzy clustering with fuzzy covariance matrix*, in Proceedings of the IEEE Control and Decision Conference, pp. 761 – 766, 1979.
- [12] D. Kukolj, E. Levi, *Identification of complex systems based on neural and Takagi-Sugeno fuzzy model*, IEEE Transactions on Systems, Man, and Cybernetics - Part B: Cybernetics, Vol. 34, pp. 272 – 282, 2004.
- [13] C.C. Kung, J.Y. Su, *Affine Takagi-Sugeno fuzzy modeling algorithm by fuzzy c-regression models clustering with a novel cluster validity criterion*, IET Control Theory, Vol. 1 (5), pp. 1255 – 1265, 2007.
- [14] L. Ledo, M. R. Delgado, J. V. de Oliveira, *Synthesis of Probabilistic Fuzzy Classifiers Using GK Clustering and Bayesian Estimation*, IEEE Latin American Transactions, Vol. 15, No. 3, 2017.
- [15] C. Li, J. Zhou, B.F.P. Kou, J. Xiao, *T-S fuzzy model identification Based on a Novel Hyperplane-Shaped Membership Function*, IEEE Trans. Fuzzy Syst., Vol. 25 (5), pp. 305 – 317, 2017.
- [16] X. Liu, H. Fang, Z. Chen, *A novel cost function based on decomposing least square support vector machine for Takagi-Sugeno fuzzy system identification*, IET Control Theory Appl., Vol. 8, pp. 338 – 347, 2014.
- [17] L. Ljung, *System Identification Theory for the User*, Prentice Hall, Englewood Cliffs, 1987.
- [18] MATLAB 2016a, *The MathWorks*, Inc., Natick, Massachusetts, United States. Retrieved October 2019 from <https://se.mathworks.com/help/ident/ref/compare.html>
- [19] National Operator of the Electrical System, Brazil, <http://www.ons.org.br>, Accessed in 2016.
- [20] H. R. O. Rocha, L. J. Silvestre, W.C. Celeste, D.J. C. Coura, L. O. Rigo Jr, *Forecast of Distributed Electrical Generation System Capacity Based on Seasonal Micro Generators using ELM and PSO*, IEEE Latin America Transactions, Vol16, No 4, pp. 1136 -1141, 2018.
- [21] J. J. Rubio, I. Elias, D. R. Cruz, J. Pacheco, G. J. Gutierrez, A. Zacarias, *A Fuzzy Algorithm for the Prediction of Future Data*, IEEE Latin America Transactions, Vol. 15, No. 8, 2017.
- [22] Ruspini, *Numerical methods for fuzzy clustering*, Inf. Sci., Vol. 2, pp. 319 – 350, 1970.
- [23] M. Sugeno, K. Tanaka, *Successive identification of a fuzzy model and its applications to prediction of a complex system*, Fuzzy Sets and Systems, pp. 315 – 334, 1991.
- [24] M. Sugeno, T. Yasukawa, *A fuzzy-logic-based approach to qualitative modeling*, IEEE Trans. Fuzzy Systems, Vol. 1, pp. 7 – 25, 1993.
- [25] T. Takagi, M. Sugeno, *Fuzzy identification of systems and its applications to modeling and control*, IEEE Transactions on Systems, Man, and Cybernetics, vol. 15 pp. 116 – 132, 1985.
- [26] S.H. Tsai, Y.W. Chen, *A novel identification method for Takagi-Sugeno fuzzy model*, Fuzzy Sets and Systems, Vol. 338, pp. 117 – 135, 2018.
- [27] Sergey Voronin, Jarmo Partanen, *Price Forecasting in the Day-Ahead Energy Market by an Iterative Method with Separate Normal Price and Price Spike Frameworks*, Energies, Vol. 6, pp. 5897 – 5920, 2013.
- [28] A. Wills, T. B. Schon, L. Ljung, B. Ninness, *Identification n of Hammerstein – Wiener models*, Automatica, Vol. 49, No. 1, pp. 7 – 81, 2013.



Jefferson Beethoven Martins é engenheiro de computação pela Universidade de Uberaba (2008) e mestre em Inovações Tecnológicas pela Universidade Federal do Triângulo Mineiro (2013). Atualmente é aluno de doutorado da Universidade Federal de Uberlândia e professor no Instituto Federal do Triângulo Mineiro. Suas principais áreas de interesse são identificação de sistemas e programação para Android.



Ana Maria Amarillo Bertone é graduada em Educação Media pelo Instituto de Professores Artigas (1982), Montevideu-Uruguay, possui mestrado em Matemática Aplicada pela Universidade Federal do Rio de Janeiro (1991) e doutorado em Matemática pela Universidade de Brasília (1998). Atualmente é professora da Faculdade de matemática da Universidade Federal de Uberlândia. Sua principal área de interesse é Teoria dos conjuntos Fuzzy.



Keiji Yamanaka é graduado em engenharia elétrica pela Universidade Federal de Uberlândia (1980) e mestre em Engenharia Elétrica (1992) pela mesma instituição. Possui Doutorado em Engenharia Elétrica e de Computação pela Nagoya Institute Of Technology, Japão (1999). Atualmente é professor titular e atua no curso de Engenharia de Computação da Universidade Federal de Uberlândia. Suas principais áreas de interesse são Inteligência Computacional como: redes neurais artificiais, computação evolutiva, e em reconhecimento de padrões.
