

**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. Bertone<sup>1</sup>  
Jefferson Beethoven Martins<sup>2</sup>  
Márcio José Cunha<sup>3</sup>  
Darizon Alves de Andrade<sup>4</sup>  
Luis Paulo Fagundes<sup>5</sup>  
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-bridge* 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.

---

<sup>1</sup>anamaria@famat.ufu.br

<sup>2</sup>jefferson@iftm.edu.br

<sup>3</sup>cunhamjg@gmail.com

<sup>4</sup>darizon@ufu.br

<sup>5</sup>lpfagundeseai@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 Nexa, 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.

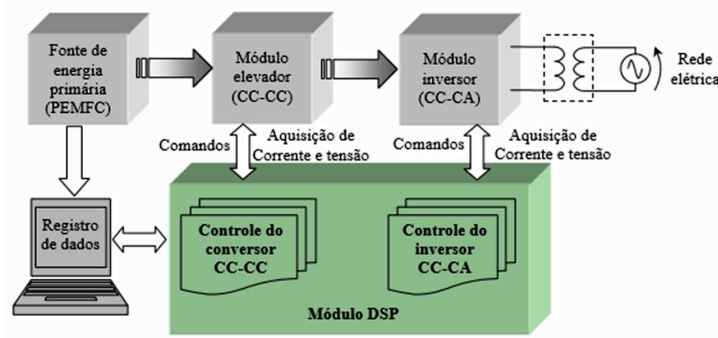


Figura 1: Fonte alternativa de energia baseada em célula a combustível (hidrogênio) [1].

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.

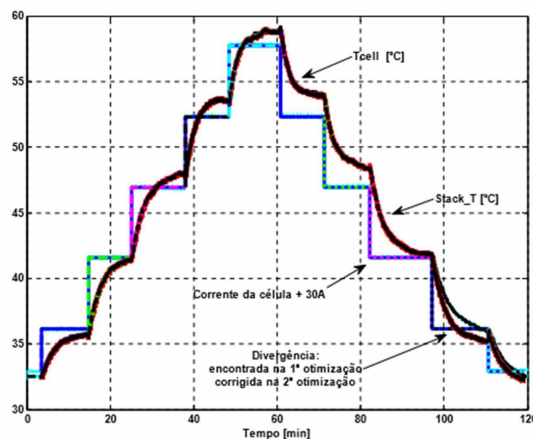


Figura 3: Comparação entre a temperatura da célula medida (Stack\_T) e a simulada (T\_cell) com modelo digitalizado, utilizando os parâmetros variáveis com a subida e descida da corrente, após otimizações do modelo. Fonte [1]

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** (**N**onlinear **A**uto**R**egressive with **eX**ogenous 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 e 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  $\mu_{A_i}(x) \in [0, 1]$ ,  $y_i \in \mathbb{R}$  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  $y = \frac{\sum_{i=1}^K \beta_i(\mathbf{x}) y_i}{\sum_{i=1}^K \beta_i(\mathbf{x})}$ , 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

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

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

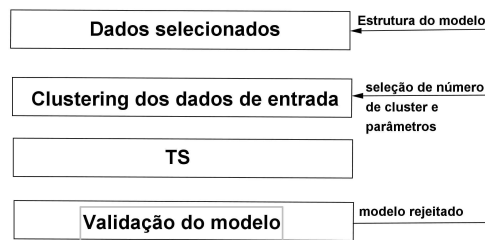


Figura 4: Algoritmo de identificação usando o clustering e TS.

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.

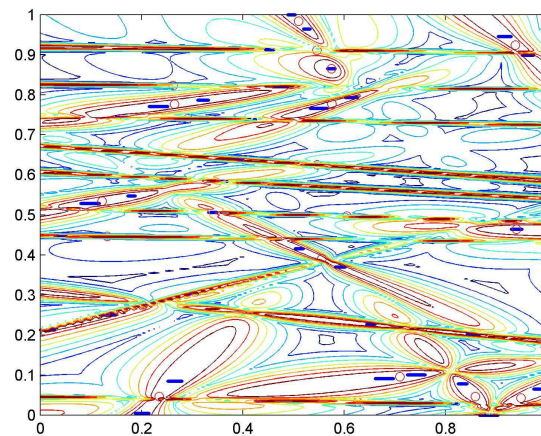
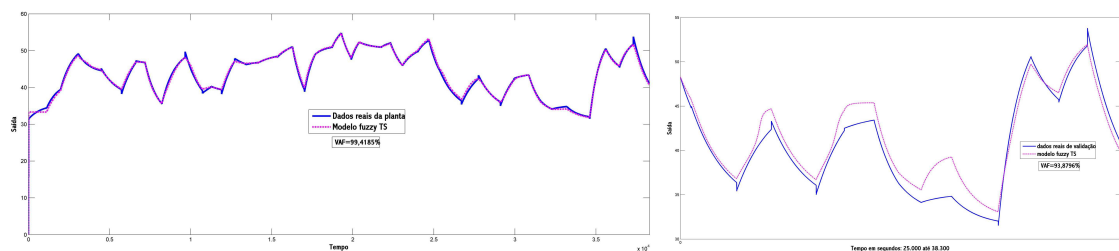


Figura 5: Clustering dos dados de entrada.

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.



(a) Processo completo do modelo fuzzy TS.

(b) Validação do modelo (aprendizado com os últimos 13.300 dados de saída)

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.