

Multivariable Evolving Fuzzy Modeling Approach with Time Varying Order State Space

Antonio Barros Feitosa Júnior¹, Ginalber Luiz de Oliveira Serra²

¹ Universidade Federal do Maranhão

²Instituto Federal de Educação, Ciência e Tecnologia do Maranhão

antonio.barros@discente.ufma.br, ginalber@ifma.edu.br

Abstract. *In this paper, a methodology for the identification of nonlinear multivariable dynamic systems via an evolving fuzzy model, is presented. The evolving fuzzy model obtained is able to adapt its structure autonomously, according to the data flow. The local submodels, in the consequent of the fuzzy rules, in the state space, are able to adjust their order as a function of time, for each fuzzy rule, independently. Furthermore, this methodology aims to reduce the number of algorithm parameters to be specified during its initialization. The proposed methodology was applied for the identification of an industrial evaporator, whose results demonstrate its applicability.*

Resumo. *Neste trabalho, é apresentada uma metodologia para a identificação de sistemas dinâmicos multivariável não lineares via modelo nebuloso evolutivo. O modelo evolutivo obtido é capaz de adaptar sua estrutura de forma autônoma, de acordo com o fluxo de dados. Os submodelos locais, no consequente das regras nebulosas, no espaço de estados, são capazes de ajustar sua ordem em função do tempo, para cada regra nebulosa, independentemente. Além disso, esta metodologia visa reduzir o número de parâmetros do algoritmo a serem especificados durante sua inicialização. A metodologia proposta foi aplicada para identificação de um evaporador industrial, cujos resultados demonstram sua aplicabilidade.*

1. Introdução

A maioria dos processos na indústria é caracterizada por um comportamento não linear e variante no tempo, o que traz a necessidade da obtenção de modelos cada vez mais eficazes. Tais modelos devem ser capazes de representar características do sistema como incertezas, não linearidades, perturbações, variação temporal dos parâmetros do processo, etc [Babuška 2003]. Entretanto, mudanças no comportamento ou nas condições operacionais do sistema podem degradar o desempenho do modelo obtido. Enquanto pequenas mudanças nos parâmetros do sistema podem ser entendidas como uma forma de incerteza e podem ser tratadas adequadamente usando mecanismos de estimação de parâmetros, mudanças na estrutura do sistema requerem um nível mais alto de adaptação [Škrjanc et al. 2019]. Os sistemas evolutivos são sistemas inteligentes adaptativos que, diferentemente dos sistemas adaptativos, atualizam sua estrutura e parâmetros simultaneamente através do fluxo de dados. Enquanto sistemas adaptativos atuam principalmente com estimativa de parâmetros, sistemas evolutivos se beneficiam do aprendizado com a

experiência, herança, mudança gradual e geração de conhecimento a partir de fluxos de dados [Gomide 2017].

A principal contribuição da metodologia proposta é a identificação de um modelo nebuloso evolutivo no espaço de estados capaz de evoluir sua estrutura de forma *on-line*, aumentando ou reduzindo o número de regras para melhor representar o fluxo de dados. Além disso, os parâmetros do consequente são estimados recursivamente através dos parâmetros de Markov do sistema e a Realização de Auto-Sistema. Como os consequentes das regras são modelos no espaço de estados, a metodologia proposta pode facilmente ser aplicada a sistemas com uma entrada e uma saída e sistemas com múltiplas entradas e múltiplas saídas. Ainda, a ordem do submodelo no consequente de cada regra é ajustada de forma automática, não necessitando que essa seja fixada *a priori*. Por fim, como os modelos no espaço de estados para cada regra são definidos pelo Algoritmo ERA (do inglês *Eigensystem Realization Algorithm*), é garantido que, para cada regra, o modelo é de realização mínima, permitindo que o mesmo possua a menor ordem possível, garantindo controlabilidade e observabilidade, além de transparência e interpretabilidade.

As principais características da metodologia proposta são descritas a seguir:

- **Flexibilidade do modelo.** O modelo obtido através desta metodologia se adéqua em tempo real ao fluxo de dados, variando o número de regras nebulosas, adaptando os parâmetros do consequente e até mesmo variando a ordem do modelo para cada regra. Tamanha flexibilidade é desejável, pois, na prática, sistemas dinâmicos reais são não lineares, variantes no tempo ou mesmo estão inseridos em ambientes não estacionários [Pires 2018];
- **Otimização de custo computacional.** Nesta proposta, o algoritmo OKID (do inglês *Observer/Kalman Filter Identification*) é executado de forma recursiva, não necessitando do armazenamento e processamento em batelada de dados de entrada e saída do sistema real, reduzindo significativamente a necessidade de alocação de memória e tempo de execução. Por se tratar de uma realização mínima, o consequente para cada regra apresenta as menores dimensões entre todos os modelos realizáveis, que possuem as mesmas relações de entrada-saída para o sistema dinâmico real, o que também reduz o custo computacional. Além disso, a definição automática das dimensões das matrizes de Henkel utilizadas durante a realização de auto-sistema, permite o uso otimizado de todos os parâmetros de Markov identificados.

2. Estrutura do Modelo Dinâmico

A i -ésima regra do modelo nebuloso evolutivo da planta a ser identificado é dada por:

$$R^i : \mathbf{SE}(\mathbf{z}_k \sim \xi_k^i) \text{ ENTÃO } \begin{cases} \mathbf{x}_{k+1}^i &= \mathbf{A}_k^i \mathbf{x}_k^i + \mathbf{B}_k^i \mathbf{u}_k \\ \mathbf{y}_k^i &= \mathbf{C}_k^i \mathbf{x}_k^i + \mathbf{D}_k^i \mathbf{u}_k \end{cases} \quad i = 1, 2, \dots, L, \quad (1)$$

onde L é o número de regras, $\mathbf{z}_k = [\mathbf{u}_k^T \ \mathbf{y}_k^T]^T \in \mathbb{R}^{r+q}$ é a amostra de dados no instante k , ξ_k^i representa a i -ésima nuvem de dados, \sim indica que a associação *fuzzy* é expressa linguisticamente como "está associado a", $\mathbf{A}_k^i \in \mathbb{R}^{n_i \times n_i}$, $\mathbf{B}_k^i \in \mathbb{R}^{n_i \times r}$, $\mathbf{C}_k^i \in \mathbb{R}^{q \times n_i}$, $\mathbf{D}_k^i \in \mathbb{R}^{q \times r}$, $\mathbf{x}_k^i = [x_k^1 \ x_k^2 \ \dots \ x_k^{n_i}] \in \mathbb{R}^{n_i}$, $\mathbf{u}_k = [u_k^1 \ u_k^2 \ \dots \ u_k^r] \in \mathbb{R}^r$ e $\mathbf{y}_k^i = [y_k^1 \ y_k^2 \ \dots \ y_k^q] \in \mathbb{R}^q$ são, respectivamente, matriz de estados, matriz de entrada, matriz de

saída, matriz de transmissão direta, vetor de estados, vetor de entrada e vetor de saída do modelo linear local para a i -ésima regra. n_i é a ordem do modelo para a i -ésima regra.

O antecedente do sistema é representado por uma nuvem de dados que descreve um determinado subconjunto de todo o conjunto de dados. Assim, a abordagem proposta substitui as funções de pertinência escalares por uma função não paramétrica que é representada pelas densidades locais [Yang et al. 2021]. Considerando-se o tipo de distância euclidiana, neste artigo, a densidade local da i -ésima nuvem de dados é definida da seguinte forma [Angelov et al. 2017, Yang et al. 2021]:

$$\gamma_k^i = \frac{1}{1 + \frac{\|\mathbf{z}_k - \mathbf{\Upsilon}_k^i\|^2}{\Xi_k^i - \|\mathbf{\Upsilon}_k^i\|^2}} \quad (2)$$

onde $\mathbf{\Upsilon}_k^i$ e Ξ_k^i são, respectivamente, o valor médio e o produto escalar das amostras de dados dentro da i -ésima nuvem de dados. A densidade local normalizada para cada nuvem de dados é dada por:

$$\lambda_k^i = \frac{\gamma_k^i}{\sum_{i=1}^L \gamma_k^i}, \quad i = 1, 2, \dots, L. \quad (3)$$

e saída do modelo nebuloso é dada por

$$\tilde{\mathbf{y}}_k = \sum_{i=1}^L \lambda^i \mathbf{y}_k^i \quad (4)$$

3. Estimação do Antecedente

A estimação do antecedente é feita utilizando o conceito proposto inicialmente por [Angelov and Yager 2012], no qual foi introduzido um tipo simplificado de sistema baseado em regras nebulosas, chamado AnYa, que oferece uma maneira de definir o antecedente das regras, sem definir as funções de pertinência por variável de maneira explícita. As proposições do antecedente das regras nebulosas são formadas sobre nuvens de dados, que são livres de parâmetros e reduzem os problemas de definição e ajuste de funções de pertinência [Angelov et al. 2013]. O sistema proposto inicia com zero nuvens de dados e forma suas nuvens de dados a partir do fluxo de dados *online*. A primeira amostra de dados é usada para construir a primeira nuvem de dados e é designada como seu ponto focal ξ_1 . À medida que as novas amostras de dados são adquiridas, mais nuvens de dados são formadas, desde que sejam atendidos dois critérios. O primeiro critério relaciona a densidade global da i -ésima nuvem de dados ($\gamma_k^{i(G)}$) e a densidade global da amostra de dados (Υ_k) no instante k , conforme a equação (5).

$$\left(\Upsilon_k - \gamma_k^{i(G)} > 0 \right) \text{ OU } \left(\Upsilon_k - \gamma_k^{i(G)} < 0 \right); \quad \forall i \in 1, \dots, L \quad (5)$$

com

$$\Upsilon_k = \frac{1}{1 + \frac{\|\mathbf{z}_k - \mathbf{\Upsilon}_k^G\|^2}{\Xi_k^G - \|\mathbf{\Upsilon}_k^G\|^2}} \quad (6)$$

e

$$\gamma_k^{i(G)} = \frac{1}{1 + \frac{\|\mathbf{\Upsilon}_k^i - \mathbf{\Upsilon}_k^G\|^2}{\Xi_k^G - \|\mathbf{\Upsilon}_k^G\|^2}}, \quad (7)$$

sendo $\mathbf{\Upsilon}_k^G$ e Ξ_k^G , respectivamente, a média global e o produto escalar global de todas as amostras de dados observadas:

$$\mathbf{\Upsilon}_k^G = \frac{k-1}{k} \mathbf{\Upsilon}_{k-1}^G + \frac{1}{k} \mathbf{z}_k, \quad \mathbf{\Upsilon}_1^G = \mathbf{z}_1 \quad (8)$$

e

$$\Xi_k^G = \frac{k-1}{k} \Xi_{k-1}^G + \frac{1}{k} \|\mathbf{z}_k\|^2, \quad \Xi_1^G = \|\mathbf{z}_1\|^2. \quad (9)$$

O segundo critério é a informação de distância, definido como

$$\zeta_{ki} > \rho_k^i; \quad \forall i \in 1, \dots, L, \quad (10)$$

onde ζ_{ki} representa a distância entre a amostra atual \mathbf{z}_k e o ponto focal da i -ésima nuvem de dados ξ_i

$$\zeta_{ki} = \|\mathbf{z}_k - \xi_i\|. \quad (11)$$

ρ_k^i representa a propagação da nuvem de dados e é atualizado da seguinte maneira:

$$\rho_k^i = \frac{1}{2} (\rho_{k-1}^i + \varphi_k^i), \quad \rho_0^i = 1, \quad (12)$$

e φ_k^i representa a dispersão local da i -ésima nuvem de dados sobre o espaço de dados no k -ésimo instante de tempo e é expresso como

$$\varphi_k^i = \sqrt{\Xi_k^i - \|\mathbf{\Upsilon}_k^i\|^2}. \quad (13)$$

Quando ambos os critérios, representados pelas equações (5) e (10), são satisfeitos, uma nova nuvem de dados é formada e a nova amostra é definida como seu ponto focal, ou seja,

$$\xi_{L+1} = \mathbf{z}_k, \quad (14)$$

e, caso contrário, o ponto focal da nuvem de dados mais próxima é atualizado pela nova amostra de dados, como segue:

$$\xi_f = \mathbf{z}_k; \quad f = \arg \min_{i=1}^L \|\mathbf{z}_k - \xi_i\|. \quad (15)$$

Durante o aprendizado, a distância entre duas densidades locais das nuvens de dados pode se tornar baixa. Nesse caso, a regra redundante precisa ser removida. A distância normalizada dos dados das nuvens duplicadas é expressa por:

$$d_k^i = \frac{D_k^{ij}}{\sum_{j=1}^L D_k^{ij}}, \quad D_k^{ij} = \|\mathbf{\Upsilon}_k^i - \mathbf{\Upsilon}_k^j\|, \quad (16)$$

sendo i e j definidos como $\arg \min_{i=1}^L \min_{j=1}^{L, i \neq j} \|\mathbb{T}_k^i - \mathbb{T}_k^j\|$. A i -ésima nuvem de dados é removida se a distância normalizada d_k^i ficar abaixo de um valor limite d_e .

Regras obsoletas são removidas seguindo o critério de utilidade, dado por:

$$U_k^i = \frac{\sum_{l=1}^k \lambda^l}{k - I^{i*}}, \quad i = [1, L], \quad t = 2, 3, \dots, \quad (17)$$

onde I^{i*} é o instante em que a i -ésima regra nebulosa foi criada. Regras que satisfazem a seguinte condição são removidas:

$$\text{SE } U_k^i < \eta \quad \text{ENTÃO } L \leftarrow L - 1, \quad \eta = [0,03; 0,1]. \quad (18)$$

4. Estimação do Consequente

O consequente de cada regra nebulosa é formado pelo submodelo linear no espaço de estados, dado por:

$$\mathbf{x}_{k+1}|^i = \mathbf{A}_k^i \mathbf{x}_k|^i + \mathbf{B}_k^i \mathbf{u}_k \quad (19a)$$

$$\mathbf{y}_k|^i = \mathbf{C}_k^i \mathbf{x}_k|^i + \mathbf{D}_k^i \mathbf{u}_k. \quad (19b)$$

A estimação dos parâmetros do consequente inicia-se com uma versão recursiva nebulosa do algoritmo OKID, a fim de estimar os parâmetros de Markov nebulosos para cada regra nebulosa. Adicionando e subtraindo o fator $\mathbf{G}^i \mathbf{y}_k|^i$ ao lado direito da equação (19a), tem-se:

$$\begin{aligned} \mathbf{x}_{k+1}|^i &= \mathbf{A}_k^i \mathbf{x}_k|^i + \mathbf{B}_k^i \mathbf{u}_k + \mathbf{G}^i \mathbf{y}_k|^i - \mathbf{G}^i \mathbf{y}_k|^i \\ &= [\mathbf{A}^i + \mathbf{G}^i \mathbf{C}^i] \mathbf{x}_k|^i + [\mathbf{B}^i + \mathbf{G}^i \mathbf{D}^i] \mathbf{u}_k - \mathbf{G}^i \mathbf{y}_k|^i \\ &= \bar{\mathbf{A}}^i \mathbf{x}_k|^i + \bar{\mathbf{B}}^i \boldsymbol{\nu}_k|^i, \end{aligned} \quad (20)$$

onde

$$\bar{\mathbf{A}}^i = [\mathbf{A}^i + \mathbf{G}^i \mathbf{C}^i], \quad (21)$$

$$\bar{\mathbf{B}}^i = [\mathbf{B}^i + \mathbf{G}^i \mathbf{D}^i, -\mathbf{G}^i], \quad (22)$$

$$\boldsymbol{\nu}_k|^i = \begin{bmatrix} \mathbf{u}_k \\ \mathbf{y}_k|^i \end{bmatrix}, \quad (23)$$

e $\mathbf{G}^i \in \mathbb{R}^{n \times q}$ é uma matriz escolhida para tornar a matriz $\bar{\mathbf{A}}^i$ estável o quanto desejado. A solução da equação (20) é dada por:

$$\mathbf{x}_k|^i = \sum_{j=1}^k (\bar{\mathbf{A}}^i)^{j-1} \bar{\mathbf{B}}^i \boldsymbol{\nu}_{k-j}|^i. \quad (24)$$

Substituindo-se a equação (24) em (19b), tem-se:

$$\mathbf{y}_k|^i = \mathbf{C} \bar{\mathbf{A}}^k \mathbf{x}_0 + \sum_{j=1}^k \mathbf{C} (\bar{\mathbf{A}}^i)^{j-1} \bar{\mathbf{B}}^i \boldsymbol{\nu}_{k-j}|^i + \mathbf{D}^i \mathbf{u}_k. \quad (25)$$

Para um valor suficientemente grande de p ($k > p$), tem-se $(\bar{\mathbf{A}}^i)^p \approx \mathbf{0}$. A equação (25) pode então ser reescrita da seguinte forma:

$$\begin{aligned} \mathbf{y}_k|^i &= \sum_{j=1}^p \mathbf{C}^i (\bar{\mathbf{A}}^i)^{j-1} \bar{\mathbf{B}}^i \boldsymbol{\nu}_{k-j}|^i + \mathbf{D}^i \mathbf{u}_k \\ &= \sum_{j=1}^p \bar{\mathbf{M}}_j^i \boldsymbol{\nu}_{k-j}|^i + \mathbf{D}^i \mathbf{u}_k, \end{aligned} \quad (26)$$

onde $\bar{\mathbf{M}}_j^i$ é o j -ésimo parâmetro de Markov do observador para o sistema linear do conseqüente da i -ésima regra nebulosa.

A equação (26) tem a seguinte representação matricial [Torres and Serra 2018]:

$$\mathbf{y}_k|^i = \boldsymbol{\theta}_k^i \boldsymbol{\phi}_k^i \quad (27)$$

onde $\boldsymbol{\theta}_k^i = [\mathbf{D}_k^i \ \bar{\mathbf{M}}_{k_1}^i \ \cdots \ \bar{\mathbf{M}}_{k_p}^i]$, com o subíndice k indicando que $\boldsymbol{\theta}_k^i$ é estimada utilizando os dados obtidos até o k -ésimo instante de tempo, e $\boldsymbol{\phi}_k^i = [\mathbf{u}_k^T \ (\mathbf{v}_{k-1}|^i)^T \ \cdots \ (\mathbf{v}_{k-p}|^i)^T]^T$. Em batelada, tem-se:

$$\mathbf{Y}_k|^i = \boldsymbol{\theta}_k^i \boldsymbol{\Phi}_k^i, \quad (28)$$

onde

$$\mathbf{Y}_k|^i = [\mathbf{y}_{p+1}|^i \ \mathbf{y}_{p+2}|^i \ \cdots \ \mathbf{y}_k|^i] \quad (29)$$

e

$$\boldsymbol{\Phi}_k^i = [\boldsymbol{\phi}_{p+1}^i \ \boldsymbol{\phi}_{p+2}^i \ \cdots \ \boldsymbol{\phi}_k^i]. \quad (30)$$

Considerando-se a equação (4), a saída do modelo nebuloso é dada, em batelada, por:

$$\tilde{\mathbf{Y}}_k = \sum_{i=1}^L \boldsymbol{\theta}_k^i \boldsymbol{\Phi}_k^i \boldsymbol{\Lambda}_k^i, \quad (31)$$

onde

$$\boldsymbol{\Lambda}^i = \begin{bmatrix} \lambda_{p+1}^i & 0 & \cdots & 0 \\ 0 & \lambda_{p+2}^i & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_k^i \end{bmatrix} \quad (32)$$

é a matriz diagonal de ponderação dos valores de pertinência na i -ésima regra, e

$$\tilde{\mathbf{Y}}_k = [\mathbf{y}_{p+1} \ \mathbf{y}_{p+2} \ \cdots \ \mathbf{y}_k] \quad (33)$$

é o vetor de saída.

A solução do problema de mínimos quadrados é feita utilizando a abordagem local para garantir a interpretabilidade dos modelos nebulosos obtidos [Torres 2018], como segue:

$$\boldsymbol{\theta}_k^i = \tilde{\mathbf{Y}}_k \boldsymbol{\Lambda}_k^i (\boldsymbol{\Phi}_k^i)^T \left[\boldsymbol{\Phi}_k^i \boldsymbol{\Lambda}_k^i (\boldsymbol{\Phi}_k^i)^T \right]^{-1}. \quad (34)$$

Com a chegada de uma nova amostra de dados, adiciona-se os termos \mathbf{u}_{k+1} e \mathbf{y}_{k+1} em (34). Dessa forma, tem-se:

$$\boldsymbol{\theta}_{k+1}^i = \tilde{\mathbf{Y}}_{k+1}^i \boldsymbol{\Lambda}_{k+1}^i (\boldsymbol{\Phi}_{k+1}^i)^T \left[\boldsymbol{\Phi}_{k+1}^i \boldsymbol{\Lambda}_{k+1}^i (\boldsymbol{\Phi}_{k+1}^i)^T \right]^{-1}, \quad (35)$$

onde $\tilde{\mathbf{Y}}_{k+1}^i = [\tilde{\mathbf{Y}}_k^i \quad \mathbf{y}_{k+1}]$ e $\boldsymbol{\Phi}_{k+1}^i = [\boldsymbol{\Phi}_k^i \quad \boldsymbol{\phi}_{k+1}^i]$.

A solução da equação (35) requer a inversão de uma matriz de dimensões elevadas, o que pode ser computacionalmente dispendioso e, além disso, induzir erros numéricos [Wu et al. 2015]. Para contornar este problema, a equação (35) pode ser calculada pelo método dos mínimos quadrados recursivos, sendo a matriz de covariância dada por:

$$\begin{aligned} \mathbf{P}_{k+1}^i &= \left[\boldsymbol{\Phi}_{k+1}^i \boldsymbol{\Lambda}_{k+1}^i (\boldsymbol{\Phi}_{k+1}^i)^T \right]^{-1} \\ &= \left[\boldsymbol{\Phi}_k^i \boldsymbol{\Lambda}_k^i (\boldsymbol{\Phi}_k^i)^T + \boldsymbol{\phi}_{k+1}^i \lambda_{k+1}^i (\boldsymbol{\phi}_{k+1}^i)^T \right]^{-1}. \end{aligned} \quad (36)$$

Usando o lema da inversão de matrizes, a equação (36) pode ser reescrita como:

$$\mathbf{P}_{k+1}^i = \mathbf{P}_k^i \left[\mathbf{I} - \frac{\boldsymbol{\phi}_{k+1}^i (\boldsymbol{\phi}_{k+1}^i)^T \mathbf{P}_k^i}{(\lambda_{k+1}^i)^{-1} + (\boldsymbol{\phi}_{k+1}^i)^T \mathbf{P}_k^i \boldsymbol{\phi}_{k+1}^i} \right]. \quad (37)$$

Substituindo-se (37) em (35), tem-se:

$$\begin{aligned} \boldsymbol{\theta}_{k+1}^i &= \tilde{\mathbf{Y}}_{k+1}^i \boldsymbol{\Lambda}_{k+1}^i (\boldsymbol{\Phi}_{k+1}^i)^T \left[\boldsymbol{\Phi}_{k+1}^i \boldsymbol{\Lambda}_{k+1}^i (\boldsymbol{\Phi}_{k+1}^i)^T \right]^{-1} \\ &= \boldsymbol{\theta}_k^i + [\mathbf{y}_{k+1} - \boldsymbol{\theta}_k^i \boldsymbol{\phi}_{k+1}^i] \frac{(\boldsymbol{\phi}_{k+1}^i)^T \mathbf{P}_k^i}{(\lambda_{k+1}^i)^{-1} + (\boldsymbol{\phi}_{k+1}^i)^T \mathbf{P}_k^i \boldsymbol{\phi}_{k+1}^i}. \end{aligned} \quad (38)$$

Os parâmetros de Markov do observador podem então ser obtidos, recursivamente, reformulando-se as equações (37) e (38), da seguinte forma:

$$\mathbf{Kw}_k^i = \frac{(\boldsymbol{\phi}_{k+1}^i)^T \mathbf{P}_k^i}{(\lambda_{k+1}^i)^{-1} + (\boldsymbol{\phi}_{k+1}^i)^T \mathbf{P}_k^i \boldsymbol{\phi}_{k+1}^i} \quad (39)$$

$$\boldsymbol{\theta}_{k+1}^i = \boldsymbol{\theta}_k^i + [\mathbf{y}_{k+1} - \boldsymbol{\theta}_k^i \boldsymbol{\phi}_{k+1}^i] \mathbf{Kw}_k^i \quad (40)$$

$$\mathbf{P}_{k+1}^i = \mathbf{P}_k^i [\mathbf{I} - \boldsymbol{\phi}_{k+1}^i \mathbf{Kw}_k^i] \quad (41)$$

A partir da estimação dos parâmetros de Markov do observador, pode-se obter os parâmetros de Markov nebulosos do sistema, para a i -ésima regra, por meio do particionamento dos parâmetros de Markov do observador, da seguinte maneira:

$$\boldsymbol{\theta}_k^i = [\bar{\mathbf{M}}_{-1}^i \quad \bar{\mathbf{M}}_0^i \quad \bar{\mathbf{M}}_1^i \quad \cdots \quad \bar{\mathbf{M}}_{p-1}^i] \quad (42)$$

onde

$$\begin{aligned} \bar{\mathbf{M}}_j^i &= \mathbf{C}_k^i (\bar{\mathbf{A}}_k^i)^j \bar{\mathbf{B}}_k^i \\ &= \left[\mathbf{C}_k^i (\mathbf{A}_k^i + \mathbf{G}_k^i \mathbf{C}_k^i)^j (\mathbf{B}_k^i + \mathbf{G}_k^i \mathbf{D}_k^i), \quad -\mathbf{C}_k^i (\mathbf{A}_k^i + \mathbf{G}_k^i \mathbf{C}_k^i)^j \mathbf{G}_k^i \right] \\ &\triangleq [\bar{\mathbf{M}}_j^{i(1)}, \quad \bar{\mathbf{M}}_j^{i(2)}]; \quad j = 0, 1, 2, \dots \end{aligned} \quad (43)$$

e

$$\bar{\mathbf{M}}_{-1}^i = \mathbf{D}_k^i. \quad (44)$$

Os parâmetros de Markov nebulosos do sistema para a i -ésima regra são então obtidos através da equação a seguir:

$$\begin{aligned} \mathbf{M}_0^i &= \mathbf{D}_k^i \\ \mathbf{M}_{j+1}^i &= \bar{\mathbf{M}}_j^{i(1)} + \sum_{s=0}^{j-1} \bar{\mathbf{M}}_s^{i(2)} \mathbf{M}_{j-s-1}^i + \bar{\mathbf{M}}_s^{i(2)} \mathbf{D}_k^i, \quad j = 0, 1, 2, \dots \end{aligned} \quad (45)$$

Uma vez obtidos os parâmetros de Markov nebulosos do sistema, é feita a realização de auto-sistema para cada regra. Para tanto, inicialmente, é formada uma matriz de Hankel para cada regra, utilizando os parâmetros de Markov nebulosos obtidos anteriormente, conforme segue:

$$\mathbf{H}_0^i = \begin{bmatrix} \mathbf{M}_1^i & \mathbf{M}_2^i & \cdots & \mathbf{M}_\gamma^i \\ \mathbf{M}_2^i & \mathbf{M}_3^i & \cdots & \mathbf{M}_{\gamma+1}^i \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{M}_\beta^i & \mathbf{M}_{\beta+1}^i & \cdots & \mathbf{M}_{\beta+\gamma-1}^i \end{bmatrix}. \quad (46)$$

onde a matriz \mathbf{H}_0^i possui dimensões $\gamma r \times \beta q$, onde r é o número de entradas do sistema, q é o número de saídas do sistema, γ e β são inteiros de forma que $\gamma r \geq \beta q$.

No intuito de reduzir o número de parâmetros a serem especificados durante a implementação da proposta, os parâmetros p , β e γ são determinados de forma automática, através das equações a seguir:

$$p = \text{ceil} \left(\frac{5\hat{n}}{r+q} \right) (r+q) \quad (47)$$

$$\beta = \text{ceil} \left(\frac{5\hat{n}}{r+q} \right) r \quad (48)$$

$$\gamma = \text{ceil} \left(\frac{5\hat{n}}{r+q} \right) q, \quad (49)$$

onde o operador $\text{ceil}(\cdot)$ arredonda o elemento entre parênteses para o número inteiro mais próximo maior ou igual a esse elemento.

As equações (47), (48) e (49) permitem que os parâmetros p , β e γ sejam determinados automaticamente, tendo o usuário que determinar apenas uma estimativa \hat{n} da ordem do sistema. Vale ressaltar que a estimativa \hat{n} não é o limite máximo da ordem do modelo que pode ser identificado. De fato, a maior ordem do modelo que pode ser identificada é $n_{max} = \beta q$, que é dada por:

$$n_{max} = \text{ceil} \left(\frac{5n}{r+q} \right) r q. \quad (50)$$

Uma vez formada a matriz \mathbf{H}_0^i , é feita sua decomposição em valores singulares, como segue:

$$\mathbf{H}_0^i = \mathbf{R}^i \mathbf{\Sigma}^i \mathbf{S}^{iT}, \quad (51)$$

com

$$\Sigma^i = \begin{bmatrix} \Sigma_n^i & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}, \quad (52)$$

na qual $\mathbf{0}$ são matrizes nulas com dimensões apropriadas e

$$\Sigma_n^i = \text{diag} [\sigma_1^i, \sigma_2^i, \dots, \sigma_j^i, \sigma_{j+1}^i, \dots, \sigma_n^i], \quad (53)$$

sendo $\sigma_1^i \geq \sigma_2^i \geq \dots \geq \sigma_j^i \geq \sigma_{j+1}^i \geq \dots \geq \sigma_n^i \geq 0$.

A matriz Σ_n^i , para cada regra, é então truncada de modo a conter apenas os n_i primeiros valores singulares suficientes para satisfazerem a seguinte equação:

$$\frac{\sum_{a=1}^{n_i} \sigma_a^i}{n} \geq \epsilon, \quad 0 < \epsilon \leq 1. \quad (54)$$

Por conseguinte, formam-se as matrizes $\mathbf{R}_{n_i}^i$ e $\mathbf{S}_{n_i}^i$ a partir das primeiras n_i colunas de \mathbf{R}^i e \mathbf{S}^i , respectivamente. Então, tem-se

$$\mathbf{H}_0^i = \mathbf{R}_{n_i}^i \Sigma_{n_i}^i \mathbf{S}_{n_i}^{i T}. \quad (55)$$

Em seguida, são obtidas as matrizes \mathbf{P}_β^i e \mathbf{Q}_γ^i , dadas por

$$\mathbf{P}_\beta^i = \mathbf{R}_{n_i}^i \Sigma_{n_i}^{i 1/2}, \quad (56a)$$

$$\mathbf{Q}_\gamma^i = \Sigma_{n_i}^{i 1/2} \mathbf{S}_{n_i}^{i T}. \quad (56b)$$

Então, as matrizes \mathbf{B}_k^i e \mathbf{C}_k^i são obtidas como segue:

$$\mathbf{B}_k^i = \text{as } r \text{ primeiras colunas de (56b)}, \quad (57a)$$

$$\mathbf{C}_k^i = \text{as } q \text{ primeiras linhas de (56a)}. \quad (57b)$$

Para determinar a matriz \mathbf{A}_k^i , é formada a matriz de Hankel generalizada \mathbf{H}_1^i da seguinte forma:

$$\mathbf{H}_1^i = \begin{bmatrix} \mathbf{M}_2^i & \mathbf{M}_3^i & \dots & \mathbf{M}_{\gamma+1}^i \\ \mathbf{M}_3^i & \mathbf{M}_4^i & \dots & \mathbf{M}_{\gamma+2}^i \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{M}_{\beta+1}^i & \mathbf{M}_{\beta+2}^i & \dots & \mathbf{M}_p^i \end{bmatrix}. \quad (58)$$

A matriz \mathbf{A}_k^i é, então, obtida a partir da seguinte equação:

$$\mathbf{A}_k^i = \Sigma_{n_i}^{i -1/2} \mathbf{R}_{n_i}^{i T} \mathbf{H}_1^i \mathbf{S}_{n_i}^i \Sigma_n^{i -1/2}. \quad (59)$$

5. Resultados Experimentais

Nesta seção serão apresentados os principais resultados experimentais, referentes à implementação da metodologia proposta, a fim de ilustrar sua eficiência e aplicabilidade.

5.1. Evaporador Industrial

Considera-se, nos resultados experimentais a seguir, a identificação de um evaporador industrial de quatro estágios, utilizado para reduzir o teor de água de um determinado produto [Jafari et al. 2014, Zhu et al. 1994]. As entradas deste sistema são: fluxo de alimentação para o primeiro estágio do evaporador (u^1), fluxo de vapor para o primeiro estágio do evaporador (u^2) e fluxo de água de resfriamento (u^3). As saídas deste sistema são: teor de matéria seca (y^1), fluxo do produto final (y^2) e temperatura do produto final (y^3). Os dados experimentais foram obtidos em [Van Overschee and De Moor 1994]. Foram utilizados os seguintes parâmetros durante a inicialização do algoritmo de identificação: $d_e = 0,1$, $\eta = 0,03$, $\hat{n} = 4$ e $\epsilon = 0,8$. Utilizou-se 6000 amostras de dados, sendo as amostras 3000 a 6000 utilizadas na etapa de validação. Ao término da identificação do evaporador industrial, o modelo nebuloso evolutivo possui 5 regras e a ordem final para todos os submodelos locais no conseqüente das regras nebulosas, no espaço de estados, é igual a 8, como pode ser verificado na figura 1. Os resultados da predição das saídas através do modelo identificado, no o intervalo [3000,6000] do conjunto de dados são mostrados nas figuras 2 a 4.

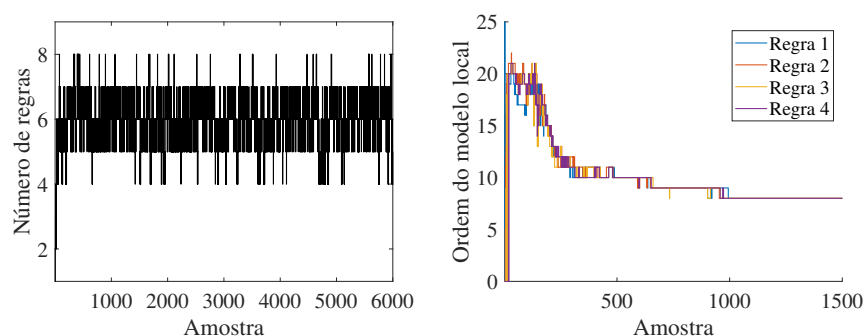


Figura 1. Evolução do número de regras e ordem do modelo local

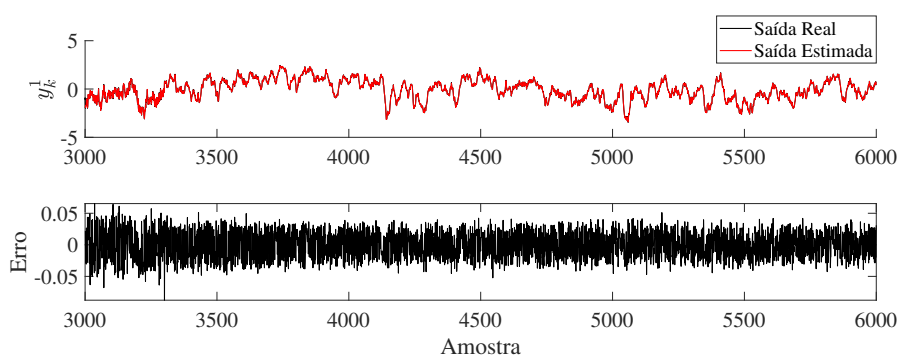


Figura 2. Saída y^1 real e estimada

A tabela 1 apresenta uma comparação entre os resultados obtidos nesta proposta, os resultados obtidos em [Torres 2018] e os resultados obtidos em [Jafari et al. 2014].

6. Conclusão

Neste artigo foi proposta uma metodologia para identificação de sistemas dinâmicos via modelo *fuzzy* evolutivo no espaço de estados com ordem variável, para lidar com

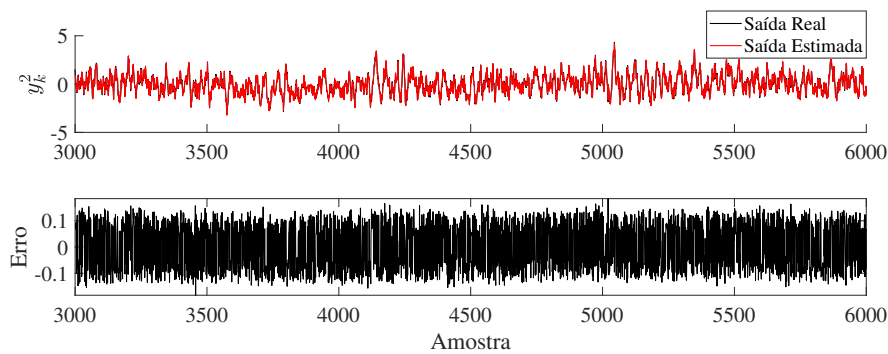


Figura 3. Saída y^2 real e estimada

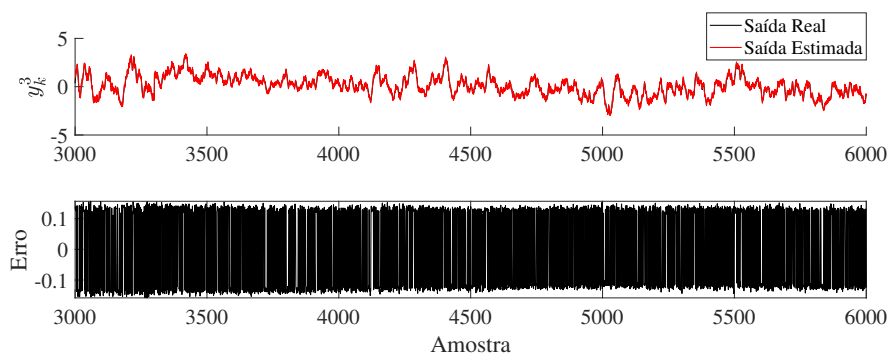


Figura 4. Saída y^3 real e estimada

dinâmicas complexas, não lineares e variantes no tempo. O modelo nebuloso evolutivo é capaz de alterar sua estrutura de forma autônoma de acordo com o fluxo de dados. A realização mínima dos submodelos nos consequentes das regras nebulosas garante a redução de sua complexidade computacional. Além disso, a ordem do submodelo local no espaço de estados é ajustada, independentemente para cada regra nebulosa, a partir da dinâmica inerente ao fluido dos dados experimentais. A metodologia proposta apresenta sua eficiência e aplicabilidade a partir da identificação de um evaporador industrial.

Referências

- Angelov, P., Gu, X., and Kangin, D. (2017). Empirical data analytics. *International Journal of Intelligent Systems*, 32(12):1261–1284.
- Angelov, P. and Yager, R. (2012). A new type of simplified fuzzy rule-based system,. *International Journal of General Systems*, 41(2):163–185.
- Angelov, P., Škrjanc, I., and Blažič, S. (2013). Robust evolving cloud-based controller for a hydraulic plant. In *2013 IEEE Conference on Evolving and Adaptive Intelligent Systems (EAIS)*, pages 1–8.
- Babuška, R. (2003). *Neuro-Fuzzy Methods for Modeling and Identification*, pages 161–186. Physica-Verlag HD, Heidelberg.
- Gomide, F. (2017). *Evolving Granular Neural Networks From Data Streams*, pages 1–9. John Wiley & Sons, Ltd.
- Jafari, M., Salimifard, M., and Dehghani, M. (2014). Identification of multivariable nonlinear systems in the presence of colored noises using iterative hierarchical least squares

Tabela 1. Comparação dos resultados obtidos através da metodologia proposta com os resultados obtidos em [Torres 2018]¹ e [Jafari et al. 2014]². Os melhores resultados são apresentados em negrito.

Métrica	Variável	Metodologia				
		Proposta	F-ERA ¹	Hammerstein ²	Wiener I ²	Wiener II ²
MSE	y^1	0,0005	0,0108	0,0234	0,0781	0,0233
	y^2	0,0086	0,0212	0,1001	0,2153	0,1261
	y^3	0,0170	0,0109	0,0204	0,0461	0,0172
FIT (%)	y^1	99,9534	90,0872	85,4225	73,3457	85,4391
	y^2	99,1077	85,1431	67,7202	52,6547	63,7737
	y^3	98,3089	89,5791	85,7438	78,5727	86,9015

algorithm. *ISA Transactions*, 53(4):1243–1252. Disturbance Estimation and Mitigation.

Pires, D. S. (2018). *Metodologia para modelagem de filtros de Kalman nebulosos evolutivos de realização mínima no espaço de estados via dados experimentais*. PhD thesis, Universidade Federal do Maranhão, Centro de Ciências Exatas e Tecnologia, São Luís, MA.

Torres, L. M. M. (2018). *Metodologia nebulosa para identificação recursiva no espaço de estados baseado em agrupamento evolutivo de dados*. Master's thesis, Universidade Federal do Maranhão, Centro de Ciências Exatas e Tecnologia, São Luís, MA.

Torres, L. M. M. and Serra, G. L. O. (2018). A novel approach for online multivariable evolving fuzzy modeling from experimental data. In *2018 IEEE International Conference on Fuzzy Systems (FUZZ-IEEE)*, pages 1–6.

Van Overschee, P. and De Moor, B. (1994). Data from an industrial evaporator. In *DaISy: Database for the Identification of Systems*, KU Leuven, Belgium. Department of Electrical Engineering, ESAT/STADIUS.

Wu, C.-Y., Tsai, J.-H., Guo, S.-M., Shieh, L.-S., Canelon, J., Ebrahimzadeh, F., and Wang, L. (2015). A novel on-line observer/kalman filter identification method and its application to input-constrained active fault-tolerant tracker design for unknown stochastic systems. *Journal of the Franklin Institute*, 352(3):1119–1151.

Yang, Z.-X., Rong, H.-J., Wong, P. K., Angelov, P., Yang, Z.-X., and Wang, H. (2021). Self-evolving data cloud-based pid-like controller for nonlinear uncertain systems. *IEEE Transactions on Industrial Electronics*, 68(5):4508–4518.

Zhu, Y., Overschee, P. V., Moor, B. D., and Ljung, L. (1994). Comparison of three classes of identification methods. In Blanke, M. and Söderström, T., editors, *Proceedings of the of SYSID '94*, volume 1, pages 175 – 180, Copenhagen, Denmark. Danish Automation Society.

Škrjanc, I., Iglesias, J. A., Sanchis, A., Leite, D., Lughofer, E., and Gomide, F. (2019). Evolving fuzzy and neuro-fuzzy approaches in clustering, regression, identification, and classification: A survey. *Information Sciences*, 490:344–368.