

# METODOLOGIA EVOLUTIVA EM ESPAÇO DE COMPONENTES NÃO-OBSERVÁVEIS PARA PREVISÃO DE SÉRIES TEMPORAIS MULTIVARIÁVEIS

SELMO EDUARDO RODRIGUES JÚNIOR\*, GINALBER LUIZ DE OLIVEIRA SERRA†

\*Av. Newton Belo, s/n, Vila Maria, CEP: 65900-000  
Instituto Federal de Educação, Ciência e Tecnologia  
Imperatriz, Maranhão, Brasil

†Av. Getúlio Vargas, 04, Monte Castelo, CEP: 65030-005  
Instituto Federal de Educação, Ciência e Tecnologia  
São Luís, Maranhão, Brasil

selmo.junior@ifma.edu.br, ginalber@ifma.edu.br

**Abstract**— This paper presents the mathematical formulation of an evolving neuro-fuzzy Takagi-Sugeno network, which is able to modify its structure composed by a fuzzy rules basis to forecast multivariate time series. The unobservable components are extracted from the multivariate time series by Multivariate Recursive Analysis Singular Spectral method proposed in this paper. These components are the inputs of the neuro-fuzzy network. Thus, the forecasting is performed for each component separately and, therefore, these various results are summed to obtain the future values of the multivariate time series. In turn, the consequences of the fuzzy rules are represented by linear models in state space, where the states correspond to the unobservable components themselves. The neuro-fuzzy network was applied to predict the behavior of a multivariate time series with medical data of Photoplethysmography, Blood Pressure and Electrocardiogram.

**Keywords**— Multivariate Time Series, Forecasting, Unobservable Components, Neuro-Fuzzy Takagi-Sugeno Network, Evolving Systems.

**Resumo**— Esse artigo apresenta a formulação matemática de uma rede *neuro-fuzzy Takagi-Sugeno* evolutiva, que é capaz de modificar sua estrutura composta por uma base regras *fuzzy* visando a previsão de séries temporais multivariáveis. As componentes não-observáveis, que são padrões extraídos das séries multivariáveis pelo método Análise Espectral Singular Recursiva Multivariável proposto nessa pesquisa, são as entradas da rede *neuro-fuzzy*. Sendo assim, a previsão é realizada para cada componente separadamente e, por conseguinte, soma-se estes vários resultados para obter os valores futuros da série temporal multivariável. Por sua vez, os consequentes das regras *fuzzy* são representados por modelos lineares no espaço de estados, sendo que os estados correspondem às próprias componentes não-observáveis. A rede *neuro-fuzzy* foi aplicada para prever o comportamento de uma série temporal multivariável com dados médicos de Fotopletismografia, Pressão Arterial e Eletrocardiograma.

**Palavras-chave**— Série Temporal Multivariável, Previsão, Componentes Não-Observáveis, Rede *Neuro-Fuzzy Takagi-Sugeno*, Sistemas Evolutivos.

## 1 Introdução

Devido sobretudo ao avanço tecnológico e ao aumento da complexidade dos sistemas em geral, há disponível uma grande quantidade de dados brutos em diversas formatações que são gerados a uma velocidade cada vez maior. Esse grande volume, a variedade e a velocidade dos dados são os atributos que dão suporte ao conceito de *Big Data*. O *Big Data* é bastante desafiador para as técnicas de aprendizagem de máquina, já que precisam lidar com as peculiaridades para extração de conhecimento em grandes volumes de dados [1, 2]. As séries temporais também estão incluídas neste cenário, pois existem inúmeras disponíveis nas mais diversas áreas, cuja análise pode fornecer conhecimentos úteis para os especialistas na tomada de decisões. À vista disso, esse artigo abrange séries temporais multivariáveis, ou seja, várias séries temporais de um mesmo domínio simultaneamente.

Uma das tarefas mais importantes envolvendo séries temporais é a sua previsão, cujo resultado é essencial para que os especialistas façam proje-

ções e realizem estratégias preventivas [3,4]. Diversos métodos que se baseiam em técnicas de inteligência computacional vêm sendo aplicados para previsão de séries temporais [2]. Por exemplo, um método híbrido para previsão de séries não-lineares e caóticas baseado em um modelo *neuro-fuzzy* e no método de Análise Espectral Singular otimizado pelo algoritmo de enxame de partículas é desenvolvido em [5]. Já em [6], uma nova metodologia que considera as componentes não-observáveis em uma estrutura *neuro-fuzzy Takagi-Sugeno* evolutiva é proposta para previsão de séries temporais sazonais e univariáveis. Entretanto, a previsão de séries temporais multivariáveis é um problema que envolve dados mais complexos, não-lineares e não-estacionários. Dessa forma, alguns trabalhos focam no desenvolvimento de metodologias para prever séries com essas complexidades. Um modelo de previsão para obter os valores futuros de séries temporais multivariáveis geofísicas é proposto em [7]. Em [8], é apresentada uma máquina de aprendizagem extrema melhorada para a previsão sequencial de séries temporais multivariáveis. Por sua vez, em [9], os autores propõem

um método de previsão e discutem o desempenho de previsão de séries temporais multivariáveis em comparação com séries temporais univariáveis, com aplicação em sistemas de trânsito.

Sendo assim, o objetivo desse trabalho é realizar a previsão de séries temporais multivariáveis através de uma estrutura *neuro-fuzzy* evolutiva. Por ser evolutiva, essa rede é capaz de modificar sua estrutura de acordo com o comportamento dos dados, ou seja, a quantidade de regras *fuzzy* pode aumentar ou ser reduzida. A fim de efetuar a previsão, a rede *neuro-fuzzy* usa as componentes não-observáveis, ou seja, padrões ocultos contidos nos dados que são extraídos do conjunto de séries temporais. A extração dessas componentes é efetuada através da Análise Espectral Singular Recursiva Multivariável, técnica apresentada nesta pesquisa. As componentes de uma série temporal específica são consideradas para efetuar a previsão de todas as séries temporais abordadas pelo método, portanto, diz-se que a rede *neuro-fuzzy* realiza a previsão das séries multivariáveis de forma acoplada. As principais contribuições dessa pesquisa são: a previsão de séries temporais multivariáveis usando uma rede *neuro-fuzzy* evolutiva, o uso das componentes não-observáveis das séries temporais de forma acoplada no espaço de estados para previsão em múltiplos passos a frente e o método para extração de componentes não-observáveis denominado Análise Espectral Singular Recursiva Multivariável (AES-RM).

## 2 Metodologia para Previsão de Séries Temporais Multivariáveis

A metodologia usada nesse artigo considera uma estrutura *neuro-fuzzy Takagi-Sugeno* para prever séries temporais multivariáveis. Esse método possui dois estágios de operação: o estágio de treinamento e o estágio de previsão. No primeiro estágio, as observações das séries temporais no instante  $k$  estão disponíveis. Sendo assim, essa rede apresenta uma característica *online* e evolutiva no treinamento, ou seja, a quantidade de regras *fuzzy* pode aumentar ou diminuir dependendo do comportamento dos dados que chegam a cada instante [10]. As entradas dessa rede *neuro-fuzzy* são componentes não-observáveis extraídas dos dados das séries temporais através do método AES-RM.

Em seguida, um algoritmo de agrupamento evolutivo considera os dados de componentes não-observáveis para formular os antecedentes das regras *fuzzy*. Ademais, as proposições consequentes das regras *fuzzy* são equações no espaço de estados, onde os estados são as próprias componentes. Portanto, os dados dos componentes são usados para estimar as matrizes de estado dos consequentes pelo método de mínimos quadrados recursivos ponderados *fuzzy* [10]. A rede é treinada para realizar o mapeamento entre os dados das compo-

nentes não-observáveis no instante  $k - 1$  para o instante  $k$  um passo a frente.

Por outro lado, não há dados da série temporal multivariável disponíveis no estágio de previsão. Neste caso, a rede *neuro-fuzzy* não evolui e não atualiza os consequentes de regras *fuzzy*, isto é, mantém sua estrutura de forma fixa. Como não há dados disponíveis para extrair as componentes não-observáveis pelo método AES-RM, os estados ou componentes obtidos das equações de estado no instante anterior são realimentados para a entrada da rede. Com isso, a previsão é executada várias etapas a frente, com horizonte de previsão  $h$ . Portanto, antes de efetivamente prever os dados das séries originais, a rede realiza a previsão das componentes de forma separada. A partir destas, obtém-se a previsão das séries temporais originais.

### 2.1 Método de Análise Espectral Singular Recursiva Multivariável (AES-RM)

A extração das componentes não-observáveis das séries temporais multivariáveis é efetuada pelo método proposto AES-RM, inspirado nos trabalhos de [11] e [6]. Este método possui 5 sub-etapas que serão descritas a seguir.

#### 2.1.1 Condições Iniciais

As condições iniciais do AES-RM começa com a determinação de três parâmetros:  $W$ , quantidade de amostras da série temporal para as condições iniciais;  $L$ , dimensão da matriz de covariância; e  $n$ , quantidade de componentes não-observáveis extraídas de cada série temporal. Após isso, um procedimento batelada é iniciado usando as primeiras  $W$  amostras da série temporal multivariável  $\mathbf{y}_k$ , sendo que  $\mathbf{y}_k = [y_k^1 \ y_k^2 \ \dots \ y_k^m]$ , para  $m$  a quantidade de séries temporais e  $k$  um instante qualquer. Um intervalo com dados das séries temporais é formado como segue:

$$\Phi = \{\mathbf{y}_k \in \mathbb{R}^{1 \times m} | k = 1, 2, \dots, W\}. \quad (1)$$

Usando-se os dados desse intervalo  $\Phi$ , a matriz trajetória  $\mathbf{M}$  pode ser então formulada por:

$$\mathbf{M} = \begin{bmatrix} y_1^1 & y_2^1 & \dots & y_{K_W}^1 & \dots & y_{K_W}^m \\ y_2^1 & y_3^1 & \dots & y_{K_W+1}^1 & \dots & y_{K_W+1}^m \\ \vdots & \vdots & \vdots & \vdots & \dots & \vdots \\ y_L^1 & y_{L+1}^1 & \dots & y_W^1 & \dots & y_W^m \end{bmatrix},$$

$$\text{sendo } K_W = W - L + 1 \text{ e } \mathbf{M} \in \mathbb{R}^{L \times (K_W \cdot m)}. \quad (2)$$

O valor de  $K_W$  é o número de colunas para cada série temporal na matriz de trajetória após  $W$  amostras de  $\mathbf{y}_k$ . Logo, a matriz de covariância inicial considerando a matriz trajetória  $\mathbf{M}$  é calculada da seguinte forma:

$$\mathbf{S}_W = \mathbf{M}\mathbf{M}^T, \quad \mathbf{S}_W \in \mathbb{R}^{L \times L}. \quad (3)$$

### 2.1.2 Atualização Recursiva da Matriz de Covariância

Após a obtenção de  $\mathbf{S}_W$ , o procedimento batelada nas condições iniciais é finalizado. A partir desse momento, a matriz  $\mathbf{S}_W$  passa a ser atualizada recursivamente. O método AES-RM armazena toda informação histórica da série temporal na matriz de covariância  $S_k$ . Para cada observação  $\mathbf{y}_k$  nos instantes  $k = W + 1, W + 2, \dots$ , o valor  $K_k$  é incrementado em uma unidade. Para atualizar recursivamente a matriz de covariância, é necessário obter a seguinte matriz *a priori*:

$$\mathbf{Q}_k = \begin{bmatrix} y_{K_k}^1 & y_{K_k}^2 & \cdots & y_{K_k}^m \\ y_{K_k+1}^1 & y_{K_k+1}^2 & \cdots & y_{K_k+1}^m \\ \vdots & \vdots & \cdots & \vdots \\ y_k^1 & y_k^2 & \cdots & y_k^m \end{bmatrix}_{L \times m}. \quad (4)$$

Assim, a atualização recursiva de  $\mathbf{S}_k$  é dada pela expressão:

$$\mathbf{S}_k = \mathbf{S}_{k-1} + \mathbf{Q}_k \mathbf{Q}_k^T, \quad \mathbf{S}_k \in \mathbb{R}^{L \times L}, \quad (5)$$

onde  $\mathbf{Q}_k \in \mathbb{R}^{L \times m}$ .

### 2.1.3 Decomposição em Valores Singulares

Em seguida, usa-se a Decomposição em Valores Singulares (*Singular Value Decomposition*, SVD) para obtenção dos valores singulares da matriz  $\mathbf{S}_k$ . Os  $L$  autovalores são calculados em cada instante, ou seja,  $\rho_k^1 \geq \rho_k^2 \geq \rho_k^3 \geq \dots \geq \rho_k^L \geq 0$ . O sistema de autovetores ortonormais associados a cada autovalor acima são representados por  $\mathbf{p}_k^1, \mathbf{p}_k^2, \mathbf{p}_k^3, \dots, \mathbf{p}_k^L$ . À vista disso, a decomposição de  $\mathbf{y}_k$  é:

$$\mathbf{y}_k = \boldsymbol{\theta}_k^1 + \boldsymbol{\theta}_k^2 + \dots + \boldsymbol{\theta}_k^L, \quad \text{para}$$

$$\boldsymbol{\theta}_k^l = \tilde{p}_k^l \mathbf{Q}_k^T \mathbf{p}_k^l, \quad l = 1, \dots, L \text{ e } \boldsymbol{\theta}_k^l \in \mathbb{R}^{1 \times m}, \quad (6)$$

onde  $\tilde{p}_k^l$  é o elemento correspondente a última posição do autovetor  $\mathbf{p}_k^l$ .

### 2.1.4 Agrupamento do AES-RM

Por conseguinte, agrupa-se o conjunto de índices  $\{1, \dots, L\}$  em  $n$  subconjuntos disjuntos de índices  $I_1, I_2, \dots, I_n$ , sendo que  $n$  é o número de componentes não-observáveis definidos nas condições iniciais. Sendo assim, seja  $I = i_1, \dots, i_q$  um conjunto genérico de índices. O termo associado a esse subconjunto genérico é  $\theta_k^I = \theta_k^{I_1} + \theta_k^{I_2} + \dots + \theta_k^{I_q}$ . Portanto, para os  $n$  subconjuntos:

$$\mathbf{y}_k = \boldsymbol{\theta}_k^{I_1} + \boldsymbol{\theta}_k^{I_2} + \dots + \boldsymbol{\theta}_k^{I_n}, \quad (7)$$

onde  $\boldsymbol{\theta}_k^{I_j} \in \mathbb{R}^{1 \times m}$ , para  $j = 1, 2, \dots, n$ .

### 2.1.5 Determinação das Componentes Não-Observáveis

Finalmente, a extração das componentes não-observáveis é determinada por:

$$\mathbf{c}_k^j = \boldsymbol{\theta}_k^{I_j}, \quad \text{onde } \mathbf{c}_k^j \in \mathbb{R}^{1 \times m}, \quad \text{para } j = 1, \dots, n, \quad (8)$$

logo:

$$\mathbf{y}_k = \mathbf{c}_k^1 + \mathbf{c}_k^2 + \dots + \mathbf{c}_k^n, \quad \mathbf{c}_k^j = [c_k^{1,j} \ c_k^{2,j} \ \dots \ c_k^{m,j}], \quad (9)$$

sendo que  $c_k^{l,j}$  ( $l = 1, \dots, m$  e  $j = 1, \dots, n$ ) representa a  $j$ -ésima componente da  $l$ -ésima série temporal em  $\mathbf{y}_k$ . A soma dessas componentes é exatamente igual ao termo correspondente da série temporal multivariável.

### 2.2 Formulação dos Antecedentes das Regras Fuzzy por Agrupamento Evolutivo

Após a extração das componentes não-observáveis, a formulação dos antecedentes das regras *fuzzy* deve ser iniciada. Seja  $n$  a quantidade de componentes total para todas as  $m$  séries temporais contidas em  $\mathbf{y}_k$ . Essa etapa realiza o agrupamento do espaço de dados das componentes, sendo que cada grupo ou *cluster* está associado a um submodelo representado por uma regra *fuzzy* [10]. Esse agrupamento é evolutivo, ou seja, a quantidade de *clusters* ou regras *fuzzy*  $R$  pode aumentar ou diminuir conforme o comportamento dos dados das componentes não-observáveis extraídas. Para cada instante  $k$ , o vetor  $\mathbf{d}_k$  de dados é gerado usando todas as componentes das  $m$  séries conforme a seguir:

$$\mathbf{d}_k = [c_{k-1}^{1,1} \ c_{k-1}^{1,2} \ \dots \ c_{k-1}^{1,n} \ c_{k-1}^{2,1} \ \dots \ c_{k-1}^{2,n} \ \dots \ c_{k-1}^{m,1} \ \dots \ c_{k-1}^{m,n} \ y_k^1 \ y_k^2 \ \dots \ y_k^m],$$

$$\text{onde } \mathbf{d}_k \in \mathbb{R}^{1 \times N}, \quad N = (n + 1) \cdot m. \quad (10)$$

É necessário padronizar o vetor  $\mathbf{d}_k$  para que os dados estejam em intervalos comparáveis [12]. O procedimento de padronização de cada elemento do vetor  $\mathbf{d}_k$  é representado por:

$$\hat{d}_k^j = \frac{d_k^j - \bar{d}_k^j}{\delta_k^j},$$

$$\text{for } j = 1, \dots, N \text{ and } k = W + 2, W + 3, \dots, \quad (11)$$

sendo que  $\hat{d}_k^j$  é o valor padronizado do elemento  $d_k^j$ . A média  $\bar{d}_k^j$  e o desvio padrão  $\delta_k^j$  são atualizados pelas seguintes expressões:

$$\bar{d}_k^j = \frac{k-1}{k} \bar{d}_{k-1}^j + \frac{1}{k} d_k^j,$$

$$\delta_k^j = \sqrt{\frac{k-1}{k} (\delta_{k-1}^j)^2 + \frac{1}{k} \|\mathbf{d}_k - \mathbf{d}_k\|^2},$$

$$\text{para } j = 1, \dots, N \text{ e } k = W + 2, W + 3, \dots. \quad (12)$$

Além disso, as condições iniciais são:

$$\begin{aligned} \bar{d}_k^j &= 0, \quad j = 1, \dots, N; \quad k = W + 1, \\ \delta_k^j &= 1, \quad j = 1, \dots, N; \quad k = W + 1, \\ \hat{d}_k^j &= 0, \quad j = 1, \dots, N; \quad k = W + 1. \end{aligned} \quad (13)$$

onde  $W + 1$  é o primeiro instante após as condições iniciais do método AES-RM. O agrupamento evolutivo, baseado na técnica *eClustering+* [12], avalia se cada vetor  $\hat{\mathbf{d}}_k$  que chega a todo instante é representativo o suficiente para formar o centro de um novo *cluster*. Caso seja afirmativo,  $\hat{\mathbf{f}}_i = \hat{\mathbf{d}}_k$ , sendo que  $\hat{\mathbf{f}}_i$  é o centro padronizado do *cluster*  $i$ , para  $i = 1, 2, \dots, R$ .

O critério utilizado para gerar novos *clusters* é baseado na densidade dos vetores  $\mathbf{d}_k$ , ou seja:

$$D_k(\hat{\mathbf{d}}_k) > \max_{i=1}^R D_k(\hat{\mathbf{f}}_{i*}) \text{ OU } D_k(\hat{\mathbf{d}}_k) < \min_{i=1}^R D_k(\hat{\mathbf{f}}_{i*}) \quad (14)$$

onde  $D_k(\hat{\mathbf{d}}_k)$  é a densidade do vetor padronizado  $\hat{\mathbf{d}}_k$ ,  $\hat{\mathbf{f}}_{i*}$  corresponde ao ponto central do *cluster*  $i$  e  $D_k(\hat{\mathbf{f}}_{i*})$ , por sua vez, é a densidade do ponto central do *cluster*, que é calculado recursivamente como:

$$D_k(\hat{\mathbf{f}}_{i*}) = \frac{k-1}{k-1 + (k-2) \left( \frac{1}{D_{k-1}(\hat{\mathbf{f}}_{i*})} - 1 \right) + \Upsilon_k}$$

onde  $\Upsilon_k = \sum_{j=1}^N (\hat{d}_k^j - \hat{d}_{k-1}^j)$ .

(15)

Já a densidade  $D_k(\hat{\mathbf{d}}_k)$  do vetor padronizado de dados é obtida conforme a seguir

$$D_k(\hat{\mathbf{d}}_k) = \frac{k-1}{(k-1) \left( \sum_{j=1}^N (\hat{d}_k^j)^2 + 1 \right) + b_k - 2 \sum_{j=1}^N \hat{d}_k^j g_k^j},$$

para  $k = W + 2, W + 3, \dots$

onde  $D_{W+1}(\hat{\mathbf{d}}_{W+1}) = 1$  (condição inicial),

$$b_k = b_{k-1} + \sum_{j=1}^N (\hat{d}_k^j)^2, \quad b_{W+1} = 0,$$

$$g_k^j = g_{k-1}^j + \hat{d}_k^j, \quad g_{W+1}^j = 0, \quad \text{para } j = 1, 2, \dots, N, \quad (16)$$

As funções de pertinência dessa metodologia proposta são gaussianas. Logo, considerando cada entrada ou componente não-observável no vetor de dados  $\hat{\mathbf{d}}_k$  ( $\hat{d}_k^j$ , para  $j = 1, 2, \dots, n$ ), a função de pertinência é dada por:

$$\mu_i^j(\hat{d}_k^j) = \exp \left( -\frac{(\hat{f}_{i*}^j - \hat{d}_k^j)^2}{2(\sigma_{k,i}^j)^2} \right), \quad (17)$$

para  $i = 1, 2, \dots, R$  e  $j = 1, 2, \dots, M$ ,

onde  $\sigma_{k,i}^j$  representa o raio do *cluster*  $i$  no eixo de entrada  $j$  para o instante  $k$ , e  $M = n \cdot m$ . Esse valor pode ser recursivamente atualizado, como citam [6, 10].

A contribuição  $\lambda_k^i$  de cada regra *fuzzy*  $i$  na saída final, também conhecida como grau de ativação normalizado, varia entre 0 e 1 e é calculada

considerando as funções de pertinência. Dessa forma, usando o operador *fuzzy* norma- $t$ :

$$\begin{aligned} \tau_k^i &= t \left\{ \mu_i^j(\hat{d}_k^j) \right\} = \prod_{j=1}^n \mu_i^j(\hat{d}_k^j) \\ \lambda_k^i &= \frac{\tau_k^i}{\sum_{i=1}^R \tau_k^i}, \quad \text{para } i = 1, \dots, R. \end{aligned} \quad (18)$$

### 2.3 Estimação das Matrizes de Estado dos Consequentes

A estrutura de uma regra *fuzzy* é representada por:

$$\begin{aligned} \text{SE } (\hat{c}_{k-1}^{1,1} \sim \hat{f}_{i*}^{1,1}) \text{ E } \dots \text{ E } (\hat{c}_{k-1}^{m,n} \sim \hat{f}_{i*}^{m,n}) \\ \text{ENTÃO } \mathbf{x}_k = \mathbf{A}_k^i \mathbf{x}_{k-1} + \mathbf{B}_k^i \mathbf{u}_{k-s}, \end{aligned} \quad (19)$$

para  $\mathbf{A}_k^i \in \mathbb{R}^{n \times n}$ ,  $\mathbf{B}_k^i \in \mathbb{R}^{n \times n}$ , e:

$$\begin{aligned} \mathbf{x}_k &= [c_k^{1,1} \ c_k^{1,2} \ \dots \ c_k^{1,n} \ c_k^{2,1} \ \dots \ c_k^{m,n}]^T; \\ \mathbf{x}_{k-1} &= [c_{k-1}^{1,1} \ c_{k-1}^{1,2} \ \dots \ c_{k-1}^{1,n} \ c_{k-1}^{2,1} \ \dots \ c_{k-1}^{m,n}]^T; \\ \mathbf{u}_{k-s} &= [c_{k-s}^{1,1} \ c_{k-s}^{1,2} \ \dots \ c_{k-s}^{1,n} \ c_{k-s}^{2,1} \ \dots \ c_{k-s}^{m,n}]^T. \end{aligned} \quad (20)$$

sendo que as componentes não-observáveis são os estados. O vetor  $\mathbf{u}_{k-s}$  usa os dados das componentes no instante  $k-s$ , com  $s$  sendo o período sazonal da série temporal multivariável determinado pelos especialistas. Para estimar as matrizes de estado, usa-se o método dos Mínimos Quadrados Recursivos Ponderados *Fuzzy* Local [10] para cada regra *fuzzy*, considerando o grau de ativação normalizado  $\lambda_k^i$  como fator de ponderação. Portanto, a estimação dessas matrizes que consideram a dinâmica de todas as  $n$  componentes é realizada da seguinte forma:

$$\begin{aligned} \mathbf{x}_k &= \mathbf{A}_k^i \mathbf{x}_{k-1} + \mathbf{B}_k^i \mathbf{u}_{k-s} \\ \mathbf{x}_k &= [\mathbf{A}_k^i \ \mathbf{B}_k^i] \begin{bmatrix} \mathbf{x}_{k-1} \\ \mathbf{u}_{k-s} \end{bmatrix} \longrightarrow \mathbf{x}_k = \Delta_k^i \mathbf{v}^T. \end{aligned} \quad (21)$$

Logo, a estimação recursiva para cada regra *fuzzy*  $i = 1, 2, \dots, R$  é dada por:

$$\begin{aligned} \mathbf{G}_k^i &= \frac{\mathbf{P}_{k-1}^i \mathbf{v}^T}{\mathbf{v} \lambda_k^i \mathbf{P}_{k-1}^i \mathbf{v}^T + 1}, \\ \Delta_k^i &= \Delta_{k-1}^i + \lambda_k^i \mathbf{G}_k^i (\mathbf{x}_k^T - \mathbf{v} \Delta_{k-1}^i), \\ \mathbf{P}_k^i &= \mathbf{P}_{k-1}^i - \lambda_k^i \mathbf{G}_k^i \mathbf{v} \mathbf{P}_{k-1}^i, \\ \text{para } \mathbf{G}_k^i &\in \mathbb{R}^{2M \times 1}, \Delta_k^i \in \mathbb{R}^{2M \times M}, \mathbf{P}_k^i \in \mathbb{R}^{2M \times 2M}. \end{aligned} \quad (22)$$

### 2.4 Saída da Rede Neuro-Fuzzy

A saída de cada regra *fuzzy* é representada por  $\tilde{\mathbf{x}}_k^i$ . A partir dessas saídas dos submodelos locais, é possível obter a saída global  $\tilde{\mathbf{x}}_k$ , que corresponde a previsão de cada componente separadamente com base nas componentes no instante  $k-1$  (Equação 10). É importante ressaltar que esse vetor se

diferencia de  $\mathbf{x}_k$ , sendo que neste último os valores das componentes são conhecidos e usados no treinamento do método. Assim sendo, os estados globais são calculados por:

$$\tilde{\mathbf{x}}_k = \sum_{i=1}^R \lambda_k^i \tilde{\mathbf{x}}_k^i. \quad (23)$$

Por fim, a rede *neuro-fuzzy* reconstrói cada série temporal original no instante  $k$  ao somar os valores das componentes não-observáveis correspondentes a cada série  $y_k^l$ . Usando o vetor  $\tilde{\mathbf{x}}_k$ , a previsão de cada série temporal  $\tilde{y}_k^l$  é representada por:

$$\tilde{y}_k^l = \tilde{c}_k^{l,1} + \tilde{c}_k^{l,2} + \dots + \tilde{c}_k^{l,n}, \text{ para } l = 1, \dots, m. \quad (24)$$

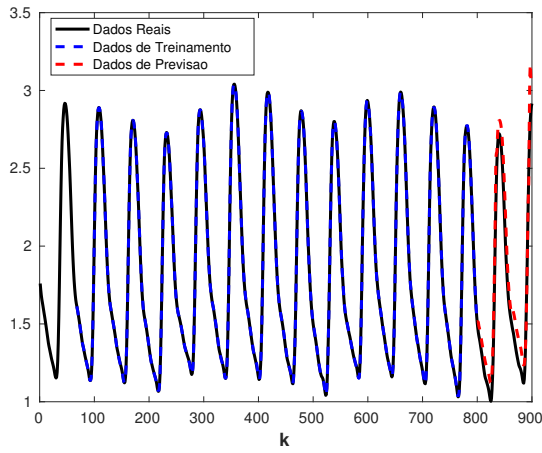


Figura 1: Série Temporal de Fotopletismografia.

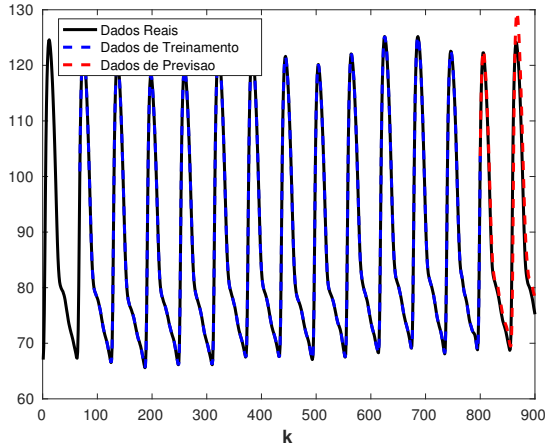


Figura 2: Série Temporal de Pressão Arterial.

### 3 Resultados Experimentais

Para avaliar a rede *neuro-fuzzy* proposta, um estudo de caso real envolvendo uma série multivariável  $\mathbf{y}_k$  constituída por 3 séries temporais reais

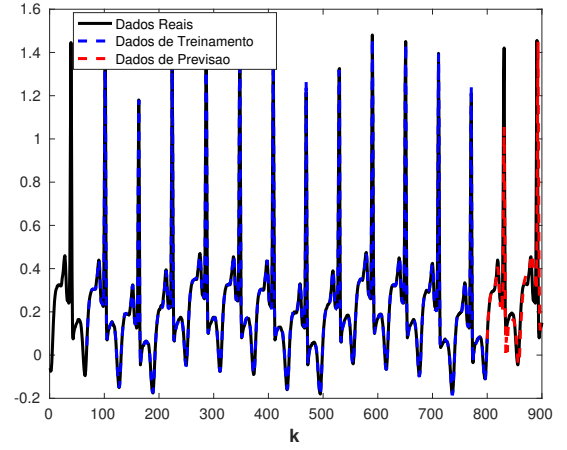


Figura 3: Série Temporal de Eletrocardiograma.

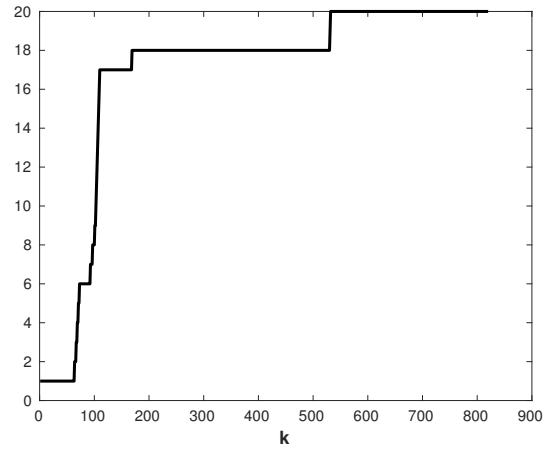


Figura 4: Quantidade de Regras *Fuzzy* ou *Clusters*.

$(y_k^1, y_k^2, y_k^3)$  foi considerada. Estas 3 séries temporais são os sinais de Fotopletismografia, Pressão Arterial e Eletrocardiograma. O monitoramento em tempo-real e a projeção dessas variáveis é importante para o acompanhamento das condições do paciente pois permite a adoção de ações preventivas [14]. O período sazonal delas é aproximadamente  $s = 63$  amostras. O objetivo desse experimento é realizar a previsão  $h = 80$  passos a frente. Em primeiro lugar, foram definidos os parâmetros iniciais do método:  $W = 67$ ,  $L = 8$  e  $n = 8$ , ou seja, 8 componentes extraídas de cada série, totalizando  $n = 24$  componentes não-observáveis.

Os resultados da previsão podem ser visualizados na Figura 1. Em preto, os dados reais da série estão representados. A parte inicial em preto são os dados usados no intervalo  $\Phi$  (Eq. 1). Em azul, os dados gerados pela rede no treinamento são exibidos, enquanto a previsão de  $h = 80$  passos a frente está em vermelho. A Figura 1.(d) exibe a mudança na quantidade de regras *fuzzy* nesse experimento, finalizando com  $R = 20$  regras *fuzzy*

Tabela 1: Valor do RMSE para Treinamento (RMSE\_T) e Previsão (RMSE\_P) de cada Série Temporal

| Série              | RMSE_T | RMSE_P |
|--------------------|--------|--------|
| Fotoplestimografia | 0.0040 | 0.1326 |
| Pressão Arterial   | 0.0675 | 4.1151 |
| Eletrocardiograma  | 0.0079 | 0.1888 |

ou *clusters*.

Para avaliar os resultados estatisticamente, a métrica RMSE (*Root Mean Square Error*) foi adotada neste trabalho, cuja formulação matemática é:

$$RMSE = \sqrt{\frac{(\tilde{y}_k - y_k)^2}{N}} \quad (25)$$

sendo que  $\tilde{y}_k$  é a observação resultante do método proposto,  $y_k$  corresponde aos dados reais das séries temporais e  $N$  a quantidade de observações consideradas. Na Tabela 1, os resultados para treinamento (RMSE\_T) e previsão (RMSE\_P) da metodologia proposta podem ser visualizados para cada série temporal. Percebe-se que pelos resultados estatísticos, a rede *neuro-fuzzy* se apresenta como uma boa alternativa para previsão de séries multivariáveis e, além disso, é capaz de adaptar sua estrutura de acordo com o comportamento dos dados.

#### 4 Conclusões

A rede *neuro-fuzzy* tem como principais vantagens a evolução de sua estrutura formada por regras *fuzzy*, estimação recursiva das componentes não-observáveis e a previsão de séries temporais multivariáveis. Os submodelos de cada regra *fuzzy* consideram todos os dados das componentes extraídas através das equações no espaço de estados, representando a dinâmica multivariável de forma acoplada. Portanto, os padrões ou componentes de uma série fornecem dados importantes para prever as componentes de outras séries temporais. O grande desafio é ajustar os parâmetros corretos para todas as séries temporais consideradas, já que um ajuste pode ser benéfico para uma série e prejudicar a previsão de outra. Algoritmos de otimização multiobjetivo podem ser utilizados para seleção desses parâmetros em trabalhos futuros. Aplicar a Decomposição em Valores Singulares para cada iteração tem um custo computacional elevado, cuja pesquisa visa reduzi-lo em trabalhos futuros. Busca-se também comparar essa metodologia multivariável proposta com outras técnicas na literatura. Além disso, deseja-se expandir a extração evolutiva de conhecimentos para outras aplicações em *Big Data*.

#### Referências

- [1] A. Gandomi and M. Haider, Beyond the hype: Big data concepts, methods, and analytics, *International Journal of Information Management*, Vol. 35, pp. 137-144, (2015), doi: 10.1016/j.ijinfomgt.2014.10.007.
- [2] L. Zhou et. al, Machine learning on big data: Opportunities and challenges, *Neurocomputing*, Vol. 237, pp. 350-361, (2017), doi: 10.1016/j.neucom.2017.01.026 .
- [3] J. Armstrong. Principles of forecasting: a handbook for researchers and practitioners. Norwell: Kluwer Academic Publishers, 2001.
- [4] P. Young, Recursive Estimation and Time-Series Analysis: An Introduction for the Student and Practitioner, 2nd Edition, Springer, 2011.
- [5] M. Abdollahzade et. al., A new hybrid enhanced local linear neuro-fuzzy model based on the optimized singular spectrum analysis and its application for nonlinear and chaotic time series forecasting, *Information Sciences*, Vol. 295, pp. 107-125, (2015), doi: 10.1016/j.ins.2014.09.002.
- [6] S. E. Rodrigues Júnior and G. L. O. Serra, A novel intelligent approach for state space evolving forecasting of seasonal time series, *Engineering Applications of Artificial Intelligence*, Vol. 64, pp. 272-285, (2017), doi: 10.1016/j.engappai.2017.06.016.
- [7] S. Pravilovic et. al., Using multiple time series analysis for geosensor data forecasting, *Information Sciences*, Vol. 380, pp. 31-52, (2017), doi: 10.1016/j.ins.2016.11.001.
- [8] X. Wang and M. Han, Improved extreme learning machine for multivariate time series online sequential prediction, *Engineering Applications of Artificial Intelligence*, Vol. 40, pp. 28-36, (2015), doi: 10.1016/j.engappai.2014.12.013.
- [9] Y. Yin and P. Shang, Forecasting traffic time series with multivariate predicting method, *Applied Mathematics and Computation*, Vol. 291, pp. 266-278, (2016), doi: 10.1016/j.amc.2016.07.017.
- [10] E. Lughofer, Evolving Fuzzy Systems: Methodologies, Advanced Concepts and Applications, Springer Berlin Heidelberg (2011), doi: 10.1007/978-3-642-18087-3.
- [11] N. Golyandina, V. Nekrutkin, A. Zhigljavsky, Analysis of Time Series Structure: SSA and Related Techniques, Chapman & Hall/CRC, New York-London, 2001.

- [12] P. Angelov. Autonomous Learning Systems: From Data Streams to Knowledge in Real-time. John Wiley & Sons, Lancaster University, United Kingdom, 2013.
- [13] K. Alligood, T. Sauer, J. Yorke, Chaos: An Introduction to Dynamical Systems, Springer Berlin Heidelberg (1996).
- [14] M. Kachuee, M. Kiani, H. Mohammadzade, M. Shabany. Cuff-Less High-Accuracy Calibration-Free Blood Pressure Estimation Using Pulse Transit Time, IEEE International Symposium on Circuits and Systems (IS-CAS'15), 2015.