Determinação da Estrutura de Modelos PWARX com Base no Erro Dinâmico de Predição

Pedro Augusto da Silva Braga* Samuel Carlos Pessoa Oliveira* Samir Angelo Milani Martins*

* GCOM - Grupo de Controle e Modelagem Universidade Federal de São João del-Rei Praça Frei Orlando, 170 Centro. São João del-Rei, MG, Brasil. (e-mails: pedro.asb@hotmail.com, samuelcpoliveira@gmail.com, martins@ufsj.edu.br).

Abstract: The present work proposes a complete methodology based on the problem of identification of non-linear systems, using PieceWise affine AutoRegressive eXogenous models. The proposed approach is based on giving an exclusive attention to the regressors that compose the vector of regressors of this class of models. Thus, a real system composed of a DC-DC *buck* converter was identified and the results showed that a greater representativeness of the system is directly associated with which regressors to use in that model class.

Resumo: O presente trabalho propõe uma metodologia completa com base no problema de identificação de sistemas não lineares utilizando modelos autoregressivos com entrada exógena afim por partes, denominados como modelos PWARX (*Piece Wise affine AutoRegressive eXogenous*). A abordagem proposta tem como base dar uma atenção exclusiva aos regressores que compõem o vetor de regressores dessa classe de modelos. Sendo assim, foi identificado um sistema real composto por um conversor CC-CC *buck* e os resultados mostraram que uma maior representatividade do sistema está diretamente associada a quais regressores utilizar nessa classe de modelo.

Keywords: System identification; PWARX model; switched models; black box models; DC-DC *buck* converter.

Palavras-chaves: Identificação de sistemas; modelo PWARX; modelos chaveados; modelos caixa-preta; conversor CC-CC *buck*.

1. INTRODUÇÃO

A identificação de um sistema dinâmico tem como base a construção de modelos matemáticos por meio de dados medidos. Dessa forma, pouco ou até mesmo nenhum conhecimento prévio do sistema é necessário. Um modelo matemático pode ser entendido como uma abstração de um sistema real representado por meio de equações (Aguirre, 2015).

Nos últimos anos, tem-se observado um empenho crescente em técnicas de identificação de sistemas com modelos afim chaveados (do inglês - *switched affine models*)(Du et al., 2018; Zwart, 2019; Hojjatinia et al., 2019; Fey et al., 2020) e modelos afim por partes (do inglês - *piecewise affine models*) (Barbosa et al., 2018; Schirrer et al., 2018; Lassoued e Abderrahim, 2019; Kersting e Buss, 2019; Du et al., 2020). Isto se deve ao fato de que identificar um modelo global que consiga cobrir diversas situações pode torná-lo muito complexo. Assim, uma alternativa para representar todo o sistema consiste na utilização de diferentes modelos locais para cada situação. Dessa forma, os modelos locais podem ser simplificados e ainda assim é possível fazer previsões precisas dentro de seus domínios (Roll, 2003). Os modelos auto-regressivos com entrada exógena afim por partes (PWARX - *PieceWise affine AutoRegressive eXogenous*) (Fantuzzi et al., 2002; Bemporad et al., 2003) pertencem a uma classe em que o espaço de regressores é dividido em regiões poliédricas e para cada uma com um respectivo submodelo afim (modelo ARX com regressor constante).

Trabalhos presentes na literatura com a identificação dessa classe de modelo ainda carecem de atenção quanto à escolha dos regressores que irão compor o vetor de regressores, como exemplos, (Nakada et al., 2005; Sun et al., 2018; Lassoued e Abderrahim, 2019; Du et al., 2020). O presente trabalho tem como foco a escolha de diferentes regressores nesse tipo de representação. Para isso, utilizando a taxa de redução de erro (ERR - *Error Reduction Ratio*), apresentando uma metodologia completa para essa classe de modelos e, assim, fazendo uso do critério de informação de Akaike (AIC - *Akaike Information Criteria*) e das técnicas da teoria de aprendizagem de máquinas, como, k-means e máquina de vetores de suporte (SVM - do inglês Support Vector Machines).

É feita uma comparação da técnica ao selecionar os regressores com base na ERR bem como ao adotar a ordem padrão utilizada na literatura. Com isso, os resultados mostraram que é possível obter modelos que apresentam de modo mais adequado o sistema quando a atenção está voltada para os regressores do modelo.

O restante do trabalho é dividido da seguinte forma: a seção 2 traz os conceitos preliminares necessários ao entendimento do trabalho desenvolvido. A metodologia e os resultados são apresentados nas próximas duas seções, 3 e 4, respectivamente. Por fim, a conclusão contendo algumas sugestões de possíveis trabalhos futuros está na seção 5.

2. CONCEITOS PRELIMINARES

2.1 Modelo PWARX

O modelo PWARX é definido por (Fantuzzi et al., 2002; Bemporad et al., 2003):

$$y(k) = \begin{cases} \theta_1^T \begin{bmatrix} x(k) \\ 1 \end{bmatrix} + e(k), & \text{se } x(k) \in \mathcal{X}_1, \\ \vdots & (1) \\ \theta_s^T \begin{bmatrix} x(k) \\ 1 \end{bmatrix} + e(k), & \text{se } x(k) \in \mathcal{X}_s. \end{cases}$$

para k = 1, 2, ..., N, sendo N o número de amostras do sistema. Os vetores $u(k) \in \mathbb{R}^m$, $y(k) \in \mathbb{R}^p$ e $e(k) \in \mathbb{R}^p$ são a entrada, saída e o ruído, por essa ordem. O vetor de regressores x(k) é definido por:

$$x(k) = [y^{T}(k-1) \dots y^{T}(k-n_{y})$$
(2)
$$u^{T}(k-1) \dots u^{T}(k-n_{u})]^{T} \in \mathbb{R}^{n},$$

no qual $n = pn_y + mn_u$ com os inteiros não negativos n_y e n_u que representam os atrasos escolhidos para saída e entrada do sistema, respectivamente. $\mathcal{X} \subseteq \mathbb{R}^n$ é o espaço de regressores e, \mathcal{X}_i , com $i = 1, 2, \ldots, s$ representam as partições poliédricas convexas de \mathcal{X} . Cada partição poliédrica \mathcal{X}_i satisfaz as condições: $\mathcal{X}_i \neq \emptyset, \forall i \in \{1, 2, \ldots, s\}$, $\mathcal{X}_i \cap \mathcal{X}_j = \emptyset \; \forall i, j \in \{1, 2, \ldots, s\} \in \bigcup_{i=1}^s \mathcal{X}_i = \mathcal{X}$. O número de submodelos é dado por $s. \; \theta_i \in \mathbb{R}^{(n+1) \times p}$ para $i = 1, 2, \ldots, s$ são os parâmetros a serem estimados.

2.2 Taxa de Redução de Erro - ERR

A ERR é uma técnica que tem como base o erro dinâmico de predição um passo à frente. Dessa forma, ela associa cada regressor candidato na composição do modelo a um índice correspondente à contribuição na explicação da variância dos dados de saída do sistema, sendo que esta é definida por:

$$\sigma_{\xi}^{2} = \lim_{N \to \infty} \frac{1}{N} \left[\langle \mathbf{y}, \mathbf{y} \rangle - \sum_{i=1}^{n_{\theta}} g_{i}^{2} \langle \mathbf{w}_{i}, \mathbf{w}_{i} \rangle \right], \qquad (3)$$

em que g_i indica cada elemento do vetor de parâmetros; \mathbf{w}_i o regressor ortogonal *i* incluso (associado a um regressor $\boldsymbol{\psi}_i$ candidato ao modelo); \mathbf{y} representa a série temporal dos dados de saída de tamanho N; e n_{θ} é o número total de regressores candidatos para compor o modelo.

Para cada termo incluído no modelo, temos como consequência a diminuição da variância de um fator de $\frac{1}{N}g_i^2 \langle \mathbf{w}_i, \mathbf{w}_i \rangle$. A ERR, devido à inclusão de um i-ésimo regressor, é definida por (Korenberg et al., 1988):

$$[ERR]_i = \frac{\hat{g}_i^2 \langle \mathbf{w}_i, \mathbf{w}_i \rangle}{\langle \mathbf{y}, \mathbf{y} \rangle},\tag{4}$$

em que \hat{g}_i se refere à estimativa dos parâmetros que podem ser calculados por:

$$\hat{g}_i = \frac{\langle \mathbf{w}_i, \mathbf{y} \rangle}{\langle \mathbf{w}_i, \mathbf{w}_i \rangle}, \quad i = 1, \dots, n_{\theta}.$$
 (5)

Dessa maneira, quanto maior a ERR de um determinado regressor i, espera-se uma maior representatividade do modelo, caso ele seja utilizado. Maiores detalhes estão disponíveis em (Aguirre, 2015).

2.3 Critério de Informação de Akaike - AIC

O AIC é uma abordagem bem conhecida para quantificar a relação entre a qualidade de ajuste e a complexidade dos modelos, sendo utilizado para determinar o número de regressores do modelo, de modo que ele represente de forma satisfatória o sistema. Em se tratando dessa técnica com modelo PWARX, o número de termos e regiões de um modelo deve minimizar a seguinte função (Akaike, 1974):

$$AIC(n_{\theta},s) = N \ln[\sigma_{\xi}^2(n_{\theta},s)] + 2n_{\theta}s, \qquad (6)$$

em que N representa o tamanho da série temporal dos dados de identificação, σ_{ξ}^2 é a variância do erro de modelagem, ou seja, o erro de predição um passo à frente; n_{θ} o número de regressos do espaço de regressores mais o termo afim; e s o número de regiões do modelo. O produto $n_{\theta}s$ em $AIC(n_{\theta},s)$ representa o grau de liberdade do modelo, ou seja, o número de parâmetros livres e servem para penalizar o critério (para cada parâmetro inserido uma penalidade de 2).

2.4 K-means

O k-means é um algoritmo que divide N observações dentre k agrupamentos. Seja x_i com $i = 1, \ldots, N$ as amostras que se deseja separar pelos C_j com $j = 1, \ldots, k$ agrupamentos; o procedimento consiste em realizar os agrupamentos dos dados de modo que minimize a seguinte função:

$$J(\mathcal{C}) = \sum_{j=1}^{k} \sum_{x_i \in \mathcal{C}_j} \|x_i - \mu_j\|^2,$$
 (7)

sendo $C = \{C_1, \ldots, C_k\}$ e μ_j são os centros de gravidade (*centroids*). O procedimento é feito de acordo com seguintes passos (Jain e Dubes, 1988):

- (1) Selecionar aleatoriamente os valores dos k centroides μ_i e repetir os passos 2 e 3 até o algoritmo convergir.
- (2) Calcular para cada amostra a sua distância com os centroides e atribuí-la ao centroide mais próximo, formando os k agrupamentos.

(3) Calcular as novas posições para os *centroides* tomando o ponto médio dos objetos pertencentes aos *centroides* do passo anterior.

2.5 Máquina de vetores de suporte

A máquina de vetores de suporte (SVM) serve para particionar os dados a fim de localizar o melhor hiperplano \mathcal{H}_{ij} que separa os pontos de dados de um agrupamento C_i dos de outro agrupamento C_j .

Para a SVM, o melhor hiperplano é aquele com a maior margem entre os dois agrupamentos, desse modo, deseja-se maximizar essa margem de separação entre os agrupamentos. A margem é a largura máxima entre os hiperplanos de separação paralelos ao hiperplano que não possui pontos de dados internos. Os vetores de suporte (do inglês *support vectors*) são os pontos de dados mais próximos dos hiperplanos de separação.

O objetivo é calcular o hiperplano $\mathcal{H}_{ij} = \{x : a_{ij}^T x + b_{ij} = 0\}$ no qual $a_{ij} \in \mathbb{R}^n$ e $b_{ij} \in \mathbb{R}$ que maximiza a margem $\{x : -1 \leq a_{ij}^T x + b_{ij} \leq 1\}$ a qual é igual a $2/||a_{ij}||$, de modo que:

$$\begin{cases} a_{ij}^T x(k) + b_{ij} \ge 1 & \forall k \in \mathcal{C}_i, \\ a_{ij}^T x(k) + b_{ij} \le -1 & \forall k \in \mathcal{C}_j. \end{cases}$$

Para o caso em que os dados não são linearmente separáveis é introduzido os escalares não negativos $\nu_{ik} \in \nu_{jk}$, de forma que (Vapnik, 1998):

$$\begin{cases} a_{ij}^T x(k) + b_{ij} \ge 1 - \nu_{ik} & \forall k \in \mathcal{C}_i, \\ a_{ij}^T x(k) + b_{ij} \le -(1 - \nu_{jk}) & \forall k \in \mathcal{C}_j. \end{cases}$$

Essas variáveis $\nu_{ik} \in \nu_{jk}$ correspondem a uma medida de violação dos dados x(k) da classificação linear $a_{ij}^T x + b_{ij} \ge 1$ e $a_{ij}^T x + b_{ij} \le -1$, as quais se desejam minimizar. A fim de considerar o compromisso entre largura da margem e da classificação incorreta, é resolvido o seguinte problema de programação quadrática:

minimizar
$$||a_{ij}||^2 + \gamma \sum_{h \in \{i,j\}, k \in \mathcal{C}_h} \nu_{hk}$$

sujeito a $a_{ij}^T x(k) + b_{ij} \ge 1 - \nu_{ik}, \forall k \in \mathcal{C}_i,$ (8)
 $a_{ij}^T x(k) + b_{ij} \le -(1 - \nu_{jk}), \forall k \in \mathcal{C}_j,$
 $\nu_{ik} \ge 0, \nu_{jk} \ge 0,$

em que $\gamma > 0$ é o parâmetro de regularização o qual impõe um peso diferente para treinamento em relação à generalização, determinado empiricamente.

A SVM é um método classificado como aprendizagem supervisionada pelo fato de necessitar de uma préclassificação nos dados para poder fazer a divisão deles em um determinado hiperplano. Maiores informações sobre essa técnica podem ser encontradas em (Boser et al., 1992; Vapnik, 1995).

2.6 Método dos mínimos quadrados

A estimativa dos parâmetros $\hat{\theta}_i$ de uma região poliédrica \mathcal{X}_i é feita pelo método dos mínimos quadrados, apresentado pela seguinte equação:

$$\hat{\theta}_i = (X_i^T X_i)^{-1} X_i^T Y_i, \tag{9}$$

Para cada $i = 1, 2, \ldots, s$, e

$$X_{i} = [\bar{x}(k_{i1}) \ \bar{x}(k_{i2}) \dots \ \bar{x}(k_{iN_{i}})]^{T},$$

$$Y_{i} = [y(k_{i1}) \ y(k_{i2}) \dots \ y(k_{iN_{i}})]^{T},$$

$$\mathcal{C}_{i} = \{k_{i1}, \ k_{i2}, \dots, \ k_{iN_{i}}\},$$

no qual N_i denota a cardinalidade de C_i , em que $N_i \ge n+1$ para que seja possível estimar $\hat{\theta}_i$ pelos dados de C_i e, evidentemente, $\sum_{i=1}^{s} N_i = N$.

O método dos mínimos quadrados tem como objetivo encontrar $\hat{\theta}_i$ que minimize a seguinte função:

$$\hat{\theta}_i = \underset{\theta_i}{\operatorname{arg\,min}} \sum_{k \in \mathcal{C}_i} \left\| y(k) - \theta_i^T \bar{x}(k) \right\|^2, \ \bar{x}(k) = \begin{bmatrix} x(k) \\ 1 \end{bmatrix}.$$
(10)

2.7 Validação do modelo

A etapa de validação do modelo é responsável por testar a qualidade do modelo identificado. Essa etapa pode envolver diferentes tipos de testes de modo que avalie a relação do modelo com os dados observados. É necessário ressaltar que a ideia principal é nunca aceitar um modelo como verdadeiro ou correto, mas sim descartar aqueles que são obviamente incorretos, existindo, dessa forma, um ingrediente subjetivo nessa etapa.

Inspeções visuais, como gráficos e medições numéricas de ajuste, podem ser usadas para decidir se o modelo é bom o suficiente. Um indicador adequado é o RMSE (do inglês - *Root Mean Squared Error*), expresso na seguinte forma:

$$RMSE = \frac{\sqrt{\sum_{k=1}^{N} [y(k) - \hat{y}(k)]^2}}{\sqrt{\sum_{k=1}^{N} [y(k) - \overline{y}]^2}},$$
(11)

sendo $\hat{y}(k)$ a simulação livre do sinal e \overline{y} o valor médio do sinal medido y(k). Esse índice mede o erro em uma unidade de medida coerente com os dados reais. Assim, modelos representativos são aqueles que apresentam esse índice normalizado com o valor menor que a unidade. Quando isso ocorre, significa que, em média, o erro demonstrado pelo modelo é menor que o erro dado pela média da série temporal.

Outro estudo para avaliar a qualidade de um modelo consiste na análise de resíduos. O vetor de resíduos é definido da seguinte forma: $\boldsymbol{\xi} = \mathbf{y} - \Psi \hat{\boldsymbol{\theta}}$, em que \mathbf{y} são os dados de identificação utilizados para estimar $\hat{\boldsymbol{\theta}}$ e construir a matriz de regressores Ψ . O primeiro teste consiste em avaliar se o vetor de resíduos possui comportamento de ruído branco. Caso isso ocorra, o modelo consegue explicar

de forma linear um passo à frente tudo aquilo que era possível ser modelado. Ao assumir ruído branco presente nos dados, a condição desejável é expressa da seguinte maneira:

$$r_{\xi} = E[\xi(k-\tau)\xi(k)] = \delta(\tau), \qquad (12)$$

em que $E[\cdot]$ indica a esperança matemática e $\delta(\tau)$ é o impulso unitário. O segundo teste da análise de resíduos consiste em avaliar se $\boldsymbol{\xi}$ e a entrada \mathbf{u} , utilizada para construir o modelo, não estão correlacionados. Caso isso ocorra, o vetor de erros cometido pelo modelo não depende do sinal de entrada. Esta condição é expressa:

$$r_{u\xi} = E[u(k-\tau)\xi(k)] = 0 \ \forall \tau.$$

$$(13)$$

Outros tipos de testes de correlação lineares e não lineares que podem ser feitos são apresentados em (Aguirre, 2015).

3. METODOLOGIA

A primeira parte do processo de identificação com modelos PWARX, proposto no presente trabalho, consiste em coletar os dados de um determinado sistema, os quais são divididos em 2 parcelas: a primeira para identificação do modelo e a segunda para realizar o processo de validação.

3.1 Identificação de modelos PWARX

A presente seção descreve o processo de identificação considerando sistemas com uma entrada e uma saída (SISO - Single-Input and Single-Output). Para isso, devem ser escolhidos os atrasos $n_y \in n_u$ para definir os regressores candidatos para compor x(k) e também devem ser escolhidos os limites $s_{min} \in s_{max}$ de modo que o número de regiões do modelo $s \in [s_{min}, s_{max}]$.

O próximo passo consiste em classificar os regressores utilizando a ERR (Equação 4) de modo hierárquico e então a ordem dos três primeiros é modificada de modo a conter o termo afim (regressor constante de modelos ARX) e os regressores $y(k - \tilde{n}_y)$ e $u(k - \tilde{n}_u)$ em que \tilde{n}_y e \tilde{n}_u são os atrasos dos regressores melhores classificados pela ERR dos dados de saída e entrada do sistema. Após a definição desses três primeiros regressores a ordem dos demais segue a ERR.

Para cada diferente valor de s e n_{θ} é aplicado o critério de informação de Akaike (Equação 6), de modo que $n_{\theta} \geq 3$, a fim de que cada modelo seja PWARX, com termo constante, possuindo efeito de memória dinâmica e influência externa (entrada exógena). O cálculo de $AIC(n_{\theta},s)$ envolve a construção do modelo. Inicialmente. os dados do vetor de regressores x(k) são agrupados em s agrupamentos usando o k-means. Em seguida, aplica-se a SVM para encontrar os hiperplanos que fazem as fronteiras entre as regiões do espaço de regressores \mathcal{X} . Após este procedimento, os parâmetros de cada região são estimados via método dos mínimos quadrados (Equação 9). Para casos em que o número de agrupamentos for maior do que dois, a separação das regiões com a SVM é feita utilizando a abordagem "um contra um", proposta inicialmente por (Knerr et al., 1990). Aplicações desse método podem ser encontradas em (Hsu e Lin, 2002).

Assim os modelos são classificados de forma categórica do menor valor de $AIC(n_{\theta},s)$ para o maior. É selecionado aquele modelo que minimiza essa função. Enfim, para validar o modelo, é utilizada a segunda parcela dos dados e é avaliado o desempenho do estimador, com os testes das Equações 12 e 13 e o desempenho dinâmico pelo índice do RMSE (Equação 11). A Figura 1 apresenta o fluxograma do processo descrito.



Figura 1. Diagrama em blocos do processo de identificação de modelos PWARX.

4. RESULTADOS

O processo descrito na metodologia foi aplicado em um sistema real composto por um conversor CC-CC *buck*, em que os dados foram originalmente coletados em (Aguirre et al., 2000). Os dados do sistema foram divididos em duas parcelas iguais, para identificação e validação, sendo escolhidos $n_y = 8$, $n_u = 8$, $s_{min} = 2$ e $s_{max} = 4$.

Com o intuito de fazer a comparação do método proposto, foram computados dois AIC. No primeiro $(AIC_{\rm err})$ os regressores foram classificados na ordem de importância pela ERR; já no segundo $(AIC_{\rm ref})$, os regressores foram classificados na ordem: y(k-1), u(k-1), y(k-2), u(k-2), $u(k-n_y)$, $u(k-n_u)$, padrão adotado na literatura com a identificação de modelos PWARX. As Tabelas 1 e 2 apresentam os cinco modelos que mais minimizam $AIC_{\rm err}$ e $AIC_{\rm ref}$, nessa ordem, os quais foram computados os seus respectivos RMSE utilizando os dados de validação.

Os resultados dos cinco modelos com menores $AIC_{\rm err}(n_{\theta},s)$ variaram entre $0,202 \leq RMSE \leq 0,333$, enquanto os de $AIC_{\rm ref}(n_{\theta},s)$ variaram entre $0,310 \leq RMSE \leq 0,839$. Isso mostra que os modelos gerados com a classificação da ERR são mais representativos do que os modelos com ordem

Tabela 1. AIC dos modelos PWARX com regressores classificados pela ERR.

$AIC_{\rm err}(n_{\theta},s)$	$n_{ heta}$	s	RMSE
-268,932	8	2	0,274
-266,362	9	2	0,273
-263,883	12	2	0,333
-260,292	5	2	0,202
-254,081	10	2	0,323

Tabela 2. AIC dos modelos PWARX com regressores classificados na ordem da literatura.

$AIC_{ref}(n_{\theta},s)$	$n_{ heta}$	s	RMSE
-277,462	8	2	$0,\!356$
-277,139	9	2	$0,\!556$
-271,847	6	2	0,310
-271,266	8	3	0,839
-271,256	12	2	$0,\!390$

na literatura, principalmente ao comparar $AIC_{\rm err}(8,2)$ com $AIC_{\rm ref}(8,2)$ e, $AIC_{\rm err}(9,2)$ com $AIC_{\rm ref}(9,2)$, os quais possuem o mesmo número de n_{θ} e s e apresentam uma diferença de 0,082 e 0,283 respectivamente, no valor do RMSE.

A equação do modelo que mais minimiza $AIC_{err}(n_{\theta},s)$, ou seja, $AIC_{err}(8,2)$, é apresentada conforme a Equação 14:

$$y(k) = \hat{\theta}_i^T \begin{bmatrix} x(k) \\ 1 \end{bmatrix}, \text{ se } x(k) \in \mathcal{X}_i, i = 1, 2, \qquad (14)$$

em que:

$$\begin{aligned} x(k) = & [y^T(k-1) \ u^T(k-5) \ y^T(k-2) \ u^T(k-1) \\ & y^T(k-8) \ u^T(k-7) \ y^T(k-3)]^T, \end{aligned}$$

e:

$$\hat{\theta}_1 = \begin{bmatrix} 1,010 \ 0,067 \ - \ 0,175 \ - \ 3,308 \ 0,013 \ 0,535 \\ - \ 0,129 \ 10,182 \end{bmatrix}^T, \\ \hat{\theta}_2 = \begin{bmatrix} 0,707 \ 1,940 \ 0,367 \ - \ 2,255 \ 0,047 \ - \ 0,064 \\ - \ 0,414 \ 5,014 \end{bmatrix}^T,$$

O hiperplano de separação das regiões do modelo é:

$$\mathcal{H}_{12} = \{x: \hat{h}_{12}^T \bar{x} = 0\}, \hat{h}_{12}^T = [a_{12}^T b_{12}],$$

em que:

$$\hat{h}_{12} = [-0.678\ 0.072\ -1.384\ -0.151\ 0.375\ 0.236\ -3.777\ 71.205]^T,$$

A análise de resíduos do modelo 14 foi feita (Equações 12 e 13) e são apresentadas nas Figuras 2 e 3. O intervalo de confiança adotado foi de 95% (linhas pontilhadas horizontais nas figuras). Ou seja, existe uma confiança de 95% na análise de resíduos de que o valor da autocorrelação $\hat{r}_{\xi}(\tau)$ e correlação cruzada $\hat{r}_{\xi u}(\tau)$ sejam nulos para qualquer atraso, quando este permanecer dentro da faixa de confiança. Nesse sentido, caso isso ocorra, o valor é estatisticamente zero. A Figura 4 apresenta a comparação



Figura 2. Função de autocorrelação do vetor de resíduos do modelo $AIC_{\rm err}(8,2)$.



Figura 3. Função de correlação cruzada da entrada do sistema com o vetor de resíduos e do modelo $AIC_{\rm err}(8,2)$.



Figura 4. Curva dinâmica do conversor CC-CC *buck* e do modelo $AIC_{\rm err}(8,2)$.

da simulação livre do modelo com os dados de saída do sistema.

Em ambos os casos, o modelo passa pelos testes da análise de resíduos, já que o comportamento do resíduo $\xi(k)$ é similar ao de um ruído branco - este também não possui correlação com a entrada u(k) do sistema. Conforme a Figura 4, o modelo consegue acompanhar de forma satisfatória o sistema utilizando basicamente dois modelos ARX que formam o PWARX.

5. CONCLUSÕES

Nesse trabalho foi apresentada uma abordagem para determinação de estrutura de modelo PWARX em que existe uma atenção aos regressores que compõem o modelo, além do número de regiões que o espaço de regressores será divido. Dessa forma, a estrutura do modelo é determinada por meio do AIC com auxílio da ERR e de técnicas da teoria da aprendizagem, como a SVM e k-means. Assim, os resultados dos modelos identificados se mostraram promissores, sendo possível identificar um sistema não linear com determinada precisão, a partir de modelos lineares locais.

Como propostas futuras, a substituição da ERR, que leva em consideração um passo à frente, pode ser substituída por uma outra, como o SRR (do inglês - *Simulation error Reduction Ratio*) (Piroddi e Spinelli, 2003) a qual utiliza a simulação livre do modelo. A generalização do método pode ser feita utilizando sistemas de múltiplas saídas e outra possibilidade é a determinação da estrutura de modelos PWARX com métodos heurísticos.

AGRADECIMENTOS

Os autores agradecem à FAPEMIG, CNPq, INERGE, CAPES e à Universidade Federal de São Jõao del-Rei pelo apoio.

REFERÊNCIAS

- Aguirre, L.A. (2015). Introdução à identificação de sistemas; técnicas lineares e não lineares: Teoria e aplicação. *Belo Horizonte*.
- Aguirre, L.A., Donoso-Garcia, P.F., e Santos-Filho, R. (2000). Use of a priori information in the identification of global nonlinear models-a case study using a buck converter. *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications*, 47(7), 1081– 1085.
- Akaike, H. (1974). A new look at the statistical model identification. *IEEE transactions on automatic control*, 19(6), 716–723.
- Barbosa, B.H., Aguirre, L.A., e Braga, A.P. (2018). Piecewise affine identification of a hydraulic pumping system using evolutionary computation. *IET Control The*ory & Applications, 13(9), 1394–1403.
- Bemporad, A., Garulli, A., Paoletti, S., e Vicino, A. (2003). A greedy approach to identification of piecewise affine models. In *International Workshop on Hybrid Systems: Computation and Control*, 97–112. Springer.
- Boser, B.E., Guyon, I.M., e Vapnik, V.N. (1992). A training algorithm for optimal margin classifiers. In *Proceedings of the fifth annual workshop on Computational learning theory*, 144–152.

- Du, Y., Liu, F., Qiu, J., e Buss, M. (2020). A semisupervised learning approach for identification of piecewise affine systems. *IEEE Transactions on Circuits* and Systems I: Regular Papers.
- Du, Z., Balzano, L., e Ozay, N. (2018). A robust algorithm for online switched system identification. *IFAC-PapersOnLine*, 51(15), 293–298.
- Fantuzzi, C., Simani, S., Beghelli, S., e Rovatti, R. (2002). Identification of piecewise affine models in noisy environment. *International Journal of Control*, 75(18), 1472– 1485.
- Fey, A., Thiele, G., e Krüger, J. (2020). Sparse identification for non-linear dynamics applied to hysteresiscontrolled systems. arXiv preprint arXiv:2003.07465.
- Hojjatinia, S., Lagoa, C.M., e Dabbene, F. (2019). Identification of switched autoregressive and switched autoregressive exogenous systems from large noisy data sets. arXiv preprint arXiv:1908.11713.
- Hsu, C.W. e Lin, C.J. (2002). A comparison of methods for multiclass support vector machines. *IEEE transactions* on Neural Networks, 13(2), 415–425.
- Jain, A.K. e Dubes, R.C. (1988). Algorithms for clustering data. Prentice-Hall, Inc.
- Kersting, S. e Buss, M. (2019). Recursive estimation in piecewise affine systems using parameter identifiers and concurrent learning. *International Journal of Control*, 92(6), 1264–1281.
- Knerr, S., Personnaz, L., e Dreyfus, G. (1990). Single-layer learning revisited: a stepwise procedure for building and training a neural network. In *Neurocomputing*, 41–50. Springer.
- Korenberg, M., Billings, S.A., Liu, Y., e McIlroy, P. (1988). Orthogonal parameter estimation algorithm for nonlinear stochastic systems. *International Journal of Control*, 48(1), 193–210.
- Lassoued, Z. e Abderrahim, K. (2019). Identification and control of nonlinear systems using piecewise autoregressive exogenous models. *Transactions of the Institute of Measurement and Control*, 41(14), 4050–4062.
- Nakada, H., Takaba, K., e Katayama, T. (2005). Identification of piecewise affine systems based on statistical clustering technique. *Automatica*, 41(5), 905–913.
- Piroddi, L. e Spinelli, W. (2003). Structure selection for polynomial narx models based on simulation error minimization. *IFAC Proceedings Volumes*, 36(16), 357– 362.
- Roll, J. (2003). Local and piecewise affine approaches to system identification. Univ.
- Schirrer, A., Mayrhofer, J., e Jakubek, S. (2018). Accurate piecewise-affine time-discrete modeling of hysteresiscoupled system dynamics. *IFAC-PapersOnLine*, 51(15), 156–161.
- Sun, X., Cai, Y., Wang, S., Xu, X., e Chen, L. (2018). Piecewise affine identification of tire longitudinal properties for autonomous driving control based on data-driven. *IEEE Access*, 6, 47424–47432.
- Vapnik, V. (1995). The nature of statistical learning theory. Springer science & business media.
- Vapnik, V. (1998). Statistical learning theory wiley. New York, 1.
- Zwart, J. (2019). Sparsity based hybrid system identification using a sat solver.