Algoritmos Multi-Objetivos para Detecção de Estruturas em Modelos NARX utilizando técnicas PEM e SEM

Henrique Carvalho de Castro *
Bruno Henrique Groenner Barbosa **

* Programa de Pós-graduação em Engenharia de Sistemas e Automação, Universidade Federal de Lavras, MG, (e-mail: henriquec.castro@outlook.com).

** Departamento de Automática, Universidade Federal de Lavras, MG, (e-mail: brunohb@ufla.br)

Abstract: This paper explores the structure selection problem in NARX models excited by non-ideal input signals, or with poor excitation properties, in output-error problems. It is proposed to use Multi-objective Genetic Algorithms together with Extended Least Squares algorithm. Two different approaches are used for these algorithms: the prediction error minimization and the free run simulation error minimization, both using the reduction of the amount of selected regressors as a second objective. The results show that these approaches have equivalent performances and are superior to the traditional FROE technique's performance in smooth input problems. While these techniques find the ideal model from 75 to 100% of executions in Monte Carlo simulations, the FROE technique finds from 0 to 14%. Moreover, due to its lower computational cost, it is advised to use the genetic approach with prediction error minimization approach.

Resumo: Este trabalho explora o problema de seleção de estrutura em modelos NARX excitados com sinais de entrada não-ideais, ou com propriedades de excitação pobres, em problemas de erro na saída. É proposto o uso de Algoritmos Genéticos Multi-objetivos em conjunto com o algoritmo de Mínimos Quadrados Estendido. Duas abordagens diferentes são usadas para estes algoritmos: a minimização do erro de predição e a minimização do erro de simulação livre, ambas utilizando a redução do número de regressores selecionados como segundo objetivo. Os resultados mostram que estas abordagens apresentam desempenhos equivalentes e que são superiores ao desempenho da técnica tradicional FROE nos problemas com entrada suave. Enquanto essas técnicas encontram o modelo ideal de 75 a 100% das execuções em simulações Monte Carlo, a técnica FROE encontra de 0 a 14%. Ademais, pelo seu menor custo computacional, aconselha-se a abordagem genética com minimização do erro de predição.

Keywords: System Identification; NARX models; Genetic algorithms; Prediction error; Simulation error

Palavras-chaves: Identificação de sistemas; Modelos NARX; Algoritmos genéticos; Erro de predição; Erro de simulação

1. INTRODUÇÃO

Construir um modelo matemático a partir dos dados de entrada e saída de um sistema é um processo conhecido como Identificação de Sistemas. Esse modelo permite o entendimento das características do sistema e a predição de seu comportamento. O sistema modelado pode ser um processo de manufatura, um sistema de produção de energia limpa, sistemas que envolvem o estudo de propriedades mecânicas e térmicas do grafeno ou nanotubos de carbono ou mercado de ações e fenômenos climáticos, entre outros (Garg et al., 2017).

De modo geral, a maior parte dos sistemas reais de interesse são não-lineares (Pope and Rayner, 1994). Nesse sentido, modelos NARMAX (do inglês *Nonlinear AutoRegressive Moving Average with eXogenous varia-*

bles) (Leontaritis and Billings, 1985) são de bastante interesse na área de Identificação de Sistemas devido à sua flexibilidade e capacidade de representação. Diversos algoritmos são propostos na literatura para modelos NARX/NARMAX, em maior parte baseados em técnicas PEM (do inglês *Prediction Error Minimization*) para estimação de parâmetros, o que é computacionalmente conveniente dada a característica linear-nos-parâmetros do modelo (Falsone et al., 2015). Porém, a principal dificuldade enfrentada por esses algoritmos é a seleção da estrutura apropriada. O ponto-chave na seleção de estrutura é escolher uma estrutura de modelo que seja tão simples quanto possível, mas também suficientemente complexa para capturar a dinâmica subjacente aos dados (Aguirre and Letellier, 2009).

Um dos critérios mais utilizados para seleção de estrutura é o *Error Reduction Ratio* (ERR) (Billings et al., 1989), que avalia a significância de um termo do modelo quanto a sua habilidade em explicar a variância dos dados de saída, sendo considerando um algoritmo de minimização do erro um-passo-à-frente (PEM). Um algoritmo baseado no critério ERR é o FROE (do inglês *Forward-Regression Orthogonal Estimator*) (Billings et al., 1989).

No entanto, Piroddi and Spinelli (2003) discutem as limitações de algoritmos FROE. É mostrado que na presença de determinados ruídos ou certas características de sinais de entrada, o FROE pode encontrar modelos incorretos ou redundantes, sendo considerado uma técnica de busca local, com grande possibilidade de encontrar soluções sub-ótimas (Falsone et al., 2015). Tais modelos podem se tornar extremamente imprecisos e até instáveis. Neste sentido, trabalhos recentes têm mostrado melhores desempenhos de modelos obtidos a partir da redução do erro de simulação livre - SEM (do inglês *Simulation Error Minimization*), ou erro de simulação, em certas condições (Milanese and Novara, 2005; Connally et al., 2007; Piroddi, 2008; Aguirre et al., 2010; Barbosa et al., 2011; Ribeiro and Aguirre, 2018; Falsone et al., 2015; Retes and Aguirre, 2019).

Considerando a estimação de parâmetros, Aguirre et al. (2010) mostraram que técnicas do tipo SEM são superiores às PEM em casos de problemas com erro na saída e até mesmo em problemas de erros-nas-variáveis, pois as abordagens PEM dependem de um ajuste adequado de modelo de ruído para a obtenção de bons resultados, enquanto que na abordagem SEM a identificação de tal modelo não é necessária. Da mesma forma, na detecção da estrutura correta para um conjunto de dados, Piroddi and Spinelli (2003); Falsone et al. (2015) mostraram que técnicas SEM são aconselháveis. Esse fato ficou bastante claro quando a excitação do sistema é realizada por entradas com características pobres de excitação. Falsone et al. (2015) apresentaram um método probabilístico chamado RaMSS (*Randomized Model Structure Selection*) onde os erros de simulação e predição são utilizados ao mesmo tempo na definição da estrutura correta de um modelo. Esses autores mostraram que o uso do erro de simulação livre é importante para descartar modelos que aparentemente possuem bom desempenho mas que não captam a dinâmica do sistema.

A obtenção de modelos por SEM é uma tarefa mais difícil pelo fato do espaço de soluções ser mais complexo como mostrado em (Connally et al., 2007), além de ocasionar em um maior custo computacional. Métodos alternativos para a solução do problema de seleção de estrutura podem ser derivados dos Algoritmos Genéticos (GA) (Holland, 1975; Goldberg and Holland, 1988) como em (Li and Jeon, 1993), (Rodriguez-Vazquez et al., 2004), (Madar et al., 2005) e (Chen et al., 2007).

Dessa forma, inspirado por Aguirre et al. (2010) que fizeram uma comparação entre as técnicas PEM e SEM na estimação de parâmetros de modelos NARX por meio de GA, este trabalho pretende analisar o desempenho dessas técnicas como funções custo em um Algoritmo Genético multiobjetivo, NSGA-II (Deb et al., 2002), para seleção de estrutura de modelos NARX, comparando com o desempenho do algoritmo FROE, em sistemas simulados

excitados com entradas não ideais. Isso posto, é proposta uma abordagem genética bi-objetiva, em que além do erro de predição (ou de simulação) é minimizado o número de termos do modelo.

O artigo está organizado da seguinte forma. Na próxima seção é apresentada uma fundamentação teórica. Na Seção III é descrita a abordagem proposta. Os resultados obtidos a partir de sistemas simulados são apresentados na Seção IV. Por fim, as conclusões são escritas na Seção V.

2. FUNDAMENTAÇÃO TEÓRICA

Modelagem matemática é a área do conhecimento que estuda formas de desenvolver e implementar modelos matemáticos de sistemas dinâmicos reais. Os modelos podem ser obtidos por diversas formas, podendo ser classificadas como (Sjöberg et al., 1995; Aguirre, 2004):

- *modelagem caixa-branca*: os modelos são obtidos por equações físicas ou químicas que regem o sistema. Deve-se conhecer o sistema profundamente;
- *modelagem (ou identificação) caixa-cinza*: o conhecimento prévio do sistema auxilia as etapas do problema de identificação;
- *modelagem (ou identificação) caixa-preta*: nenhum conhecimento prévio do sistema é necessário. A identificação do processo é realizada por meio de sinais de entrada e saída do sistema, obtidos em testes sobre o processo.

Considerando a identificação caixa-preta, suas principais etapas são: testes dinâmicos e coleta de dados; escolha da representação matemática; determinação da estrutura do modelo e estimação dos parâmetros.

2.1 Modelos polinomiais NARMAX

Nos modelos NARMAX, o valor atual da saída do sistema é obtida a partir dos valores de entrada e saída atrasados. Assim, um polinômio NARMAX (Leontaritis and Billings, 1985) é representado pela seguinte equação:

$$y(k) = f^l(y(k-1), \dots, y(k-n_y), u(k-1), \dots, u(k-n_u), e(k-1) \dots e(k-n_e)) + e(k), \quad (1)$$

em que f é uma função polinomial não-linear de grau l , $y(k)$, $u(k)$ e $e(k)$ são a saída, a entrada e o ruído, respectivamente, e n_y , n_u e n_e são os respectivos atrasos máximos. Modelos NARX são caracterizados pela ausência dos termos de ruído.

Importante observar que o número de termos presentes no modelo aumenta exponencialmente com o incremento da ordem de não-linearidade e o máximo de atrasos das saídas e entradas. Segundo (Aguirre and Billings, 1995), um modelo sobreparametrizado pode acarretar em instabilidade numérica na estimação dos parâmetros, esforço computacional desnecessário e representação de dinâmicas não existentes no sistema real. Torna-se necessário, portanto, selecionar as variáveis (regressores) que melhor representam o sistema, etapa chamada de detecção ou seleção de estrutura.

2.2 Forward-Regression Orthogonal Estimator

No algoritmo FROE, a estrutura do modelo é incrementada de forma iterativa até que seja atingida determinada precisão de predição um-passo-à-frente. Os parâmetros são estimados por meio do algoritmo de Mínimos Quadrados ortogonais. Técnicas de ortogonalização são utilizadas de forma que, a cada passo, a significância de cada candidato a regressor possa ser avaliada separadamente por meio do ERR:

$$[ERR]_j = \frac{\hat{g}_j^2 \sum_{i=1}^N w_j^2(i)}{\sum_{i=1}^N y^2(i)} \quad (2)$$

em que w_j é o j -ésimo regressor ortogonal auxiliar e $\hat{g}_j$ seu correspondente parâmetro estimado. Os regressores com maior valor de ERR são adicionados ao modelo (mais detalhes em Chen et al. (1989)).

2.3 Algoritmos Genéticos Multiobjetivos

Algoritmos Genéticos (GA) (Holland, 1975; Goldberg and Holland, 1988) são modelos computacionais capazes de resolver problemas de otimização. Eles são baseados em metáforas de processos naturais para simular a evolução das espécies. Uma população de indivíduos, também chamados de cromossomos, passa por processos de seleção, mutação e reprodução, a fim de evoluir seus genótipos.

Duas questões de extrema importância em se tratando de GA são: como representar matematicamente esses genótipos e como os avaliar. Em sua forma mais simples, os GA possuem representação binária, ou seja, um indivíduo é representado por um vetor binário no qual cada bit representa um gene do indivíduo (cromossomo). Por exemplo, um indivíduo I que possui 5 genes poderia ser representado por $I = [1 \ 0 \ 0 \ 1 \ 0]$. A avaliação dos indivíduos é feita por meio de uma função custo. Esta deve ser escolhida de forma criteriosa, contendo todo conhecimento que se tem do problema, incluindo suas restrições. A função custo utiliza as informações codificadas no cromossomo e retorna um valor numérico que é a medida da qualidade desse cromossomo.

Problemas nos quais a solução é avaliada por várias funções custo necessitam de algoritmos multiobjetivos para serem otimizados. Nesse caso, dois conceitos devem ser esclarecidos: *dominância* e *conjunto Pareto-ótimo*. Em um problema no qual pretende-se maximizar todos os objetivos, uma solução $f(A) = \{f_1(A), f_2(A), \dots, f_n(A)\}$ é dita dominante em relação a $f(B)$ se cada coordenada da avaliação de A é maior ou igual a respectiva coordenada de B , havendo pelo menos uma coordenada de A que seja maior que a respectiva coordenada de B (Eiben et al., 2003):

$$A \succ B \Rightarrow \forall i \in \{1, 2, \dots, n\} f_i(A) \geq f_i(B), \text{ e} \quad (3) \\ \exists i \in \{1, 2, \dots, n\} f_i(A) > f_i(B),$$

em que o símbolo $\succ$ indica dominância.

O conjunto Pareto-ótimo é composto por todas as soluções não-dominadas. São aquelas soluções em que sua qualidade não pode ser incrementada em nenhum objetivo sem afetar negativamente um dos outros (mais detalhes em Eiben et al. (2003); Linden (2008)).

2.4 Mínimos Quadrados Estendido

Modelos lineares-nos-parâmetros, como o NARX polinomial, podem ter seus parâmetros estimados diretamente pelo estimador Mínimos Quadrados (MQ) da seguinte forma:

$$\hat{\theta}_{MQ} = [\Psi^T \Psi]^{-1} \Psi^T y \quad (4)$$

em que Ψ é a matriz de regressores, y é o vetor de dados de saída e $\hat{\theta}_{MQ}$ os parâmetros estimados por MQ.

Nos casos em que há ruído na saída, o MQ obtém estimativas polarizadas, sendo necessária a identificação de um modelo de ruído, transformando o modelo NARX em um NARMAX. No entanto, o modelo NARMAX não é linear-nos-parâmetros. Um método derivado do MQ pode ser utilizado a fim de obter os parâmetros de um modelo NARMAX, qual seja, o algoritmo dos Mínimos Quadrados Estendido (MQE) (Young, 1968; Aguirre, 2004). Nesse caso, considere que o resíduo ($\xi = y - \Psi \hat{\theta}_{MQ}$) da identificação possa ser modelado como:

$$\xi(k) = c_i v(k-i) + v(k), \quad (5)$$

em que $v(k)$ é ruído branco e c_i um parâmetro do modelo de ruído. Adiciona-se os termos $v(k-1)$ à matriz de regressores:

$$\Psi^* = \begin{bmatrix} \Psi & \begin{matrix} v(k-1) \\ v(k) \\ v(k+1) \\ \vdots \\ v(k+N-2) \end{matrix} \end{bmatrix}, \quad (6)$$

chamada de matriz estendida e define-se um novo vetor de parâmetros $\theta^* = [\theta \ c_i]$ a ser estimado pelo MQ. Como $v(k)$ é desconhecido, deve-se utilizar um processo iterativo, sendo o vetor de resíduos calculado em cada iteração do método, implementado portanto o MQE.

3. METODOLOGIA

O primeiro passo para otimizar a escolha de regressores do modelo NARX polinomial por meio dos AG é determinar sua representação cromossomial. Na abordagem proposta, foi selecionada a representação binária, sendo que cada elemento do cromossomo indica o uso ou não do regressor no modelo. Portanto, o tamanho do indivíduo é definido pelo número de atrasos máximos na entrada (n_u), na saída (n_y) e pelo grau de não-linearidade (l) permitidos. Nos exemplos estudados, o espaço de busca será restrito em todas as combinações de regressores possíveis com $n_u = n_y = l = 2$, o que totaliza 15 candidatos a regressores. Por exemplo, um sistema $S(k) = y(k-1) + u(k-1) + y(k-1)^2$ tem representação $S' = 101010000000000$.

Uma vez definida a representação, deve-se escolher a função custo. Dois objetivos são buscados neste trabalho: a minimização do erro do modelo e a minimização da quantidade de regressores (N_r) selecionados. O erro utilizado é o erro quadrático normalizado:

$$SSE_{norm} = \sqrt{\frac{\sum_{i=1}^N (y(i) - \hat{y}(i))^2}{\sum_{i=1}^N y(i)^2}}, \quad (7)$$

em que N é a quantidade de amostras de dados, y a saída do sistema e $\hat{y}$ a estimativa do modelo, podendo

ser predição um-passo-à-frente (usada na abordagem GA-PEM) ou simulação-livre (usada na abordagem GA-SEM). Os parâmetros do modelo são estimados por MQE, com modelo de ruído definido por:

$$e(k) = c_1 e(k-1) + c_2 e(k-2), \quad (8)$$

sendo que o número de atrasos máximo do modelo de ruído n_e foi definido como 2 pois este é o máximo atraso utilizado para a variável y nos exemplos estudados. O aumento de n_e apenas aumentaria o custo computacional sem melhoria de desempenho.

A busca pela melhor estrutura é realizada para dois sistemas simulados com dois tipos de entradas: um sinal de entrada aleatório com distribuição normal de média 0 e variância 1 (persistentemente excitante, considerado uma entrada ideal), $u(t) = WGN(0, 1)$; e um sinal de entrada suave (com propriedade de excitação ruim), $u(t) = WGN(0, 0,25)$ que foi processado por um filtro passa-baixa com polos em 0,95 e 0,9. Os sistemas simulados são apresentados nas seguintes equações de erro na saída:

$$\begin{aligned} S_1 : w(k) &= 0,8 \cdot u(k-1)^2 + 1,2 \cdot w(k-1) \\ &\quad - 0,9 \cdot w(k-2) - 0,2 \\ y(k) &= w(k) + e(k) \\ S_2 : w(k) &= 0,75 \cdot w(k-2) + 0,25 \cdot u(k-1) \\ &\quad - 0,20 \cdot w(k-2)u(k-1) \\ y(k) &= w(k) + e(k) \end{aligned} \quad (9)$$

em que $e(k)$ é um ruído gaussiano de média 0 e variância 0,02.

Para minimizar N_r e SSE_{norm} é utilizado o algoritmo NSGA-II que retorna uma solução Pareto-ótima. Assim, é selecionado desse conjunto o modelo que possui o mesmo número de regressores que o sistema simulado em estudo (4 para o sistema S_1 e 3 para o sistema S_2). O importante agora é observar se o modelo ideal está presente no conjunto Pareto. Entretanto, em uma abordagem prática deverá ser implementado um decisor para selecionar dentre as soluções Pareto-ótimas aquela mais adequada, como apresentado por Barroso et al. (2007). O critério de parada do algoritmo FROE é estabelecido também de acordo com o tamanho dos modelos estudados.

Nos dois exemplos, são geradas 200 amostras para serem utilizadas como base no processo de otimização da estrutura do modelo.

O critério escolhido para determinar um sucesso é o algoritmo selecionar exatamente os regressores pertencentes aos sistemas. Com simulações diferentes são realizadas para cada tipo de entrada em cada um dos sistemas. São comparados os sucessos entre a seleção de estrutura por GA-PEM, GA-SEM e FROE. É realizada uma análise estatística dos parâmetros estimados por MQE nas soluções consideradas sucesso e uma comparação da distribuição de seleção de variáveis nos diferentes métodos.

4. RESULTADOS

A Tabela 1 apresenta a taxa de acerto dos algoritmos para as diferentes entradas. Nota-se que, para a entrada ideal $u(t) = WGN(0,1)$, GA-PEM, GA-SEM e o FROE têm taxas de acertos relativamente semelhantes, entre 97

e 100%. O problema se torna evidente quando analisado o número de acertos dos algoritmos para a entrada suave. Nesse caso, GA-PEM e GA-SEM mantêm suas performances em 100% de acerto no sistema S_1 , e diminuem para 79% e 75% no sistema S_2 , respectivamente. Já o método FROE reduz de 99% para 14% no sistema S_1 e de 97% para 0% no sistema S_2 . Na mesma tabela é possível observar que o custo computacional da abordagem GA-SEM é bastante superior em relação às demais técnicas. Enquanto FROE e GA-PEM consomem um tempo médio de 0,02s e 0,95s, respectivamente, por execução Monte Carlo, GA-SEM consome 7,01s.

Difícilmente se sabe mensurar se a entrada de um sistema é ideal, ou o quão ideal ela é, em se tratando de sistemas reais. Além disso, quanto maior a quantidade de dados, maior será a diferença entre os tempos de execução de GA-PEM e GA-SEM. Com essas informações, o algoritmo GA-PEM se mostra uma excelente escolha por sua maior taxa de acertos e menor custo computacional.

Tabela 1. Taxa média de acerto em 100 execuções de Monte Carlo

Número de acertos com Entrada Ideal			
Sistema	GA-PEM	GA-SEM	FROE
S_1	100%	99%	99%
S_2	100%	100%	97%
Número de acertos com Entrada Suave			
S_1	100%	100%	14%
S_2	79%	75%	0%
Tempo Médio(s)	0,95	7,01	0,02

As tabelas 2 e 3 apresentam a distribuição da seleção de variáveis realizada pelos algoritmos no problema de seleção de estrutura com entrada suave. Nelas são mostradas a frequência com a qual cada variável foi selecionada para o modelo. Observa-se que, para o sistema S_1 , as técnicas GA-PEM e GA-SEM selecionam todas as variáveis presentes no modelo ideal em 100% das execuções. Já FROE concentra grande parte de seu erro selecionando a variável $u(k-2)^2$ (selecionada em 63% das execuções). Já para o sistema S_2 , Os métodos GA-PEM e GA-SEM não apresentam concentração de erro em nenhuma variável específica, ao passo que FROE seleciona erroneamente em 96% das execuções a variável $y(k-2)u(k-2)$. Esses fenômenos podem ser explicados pela alta correlação entre as variáveis $u(k-1)$ e $u(k-2)$, o que os tornam igualmente significantes frente ao algoritmo de decisão do método FROE (deveriam ser selecionadas $u(k-1)^2$ para o sistema S_1 e $y(k-2)u(k-1)$ para o sistema S_2).

As Tabelas 4 e 5 apresentam a média e desvio padrão da estimação dos parâmetros por MQE de todos os modelos corretos encontrados por meio do GA com entrada suave. Nota-se que a entrada com propriedade de excitação pobre não afeta a estimação dos parâmetros do modelo NARX pelo MQE, os quais apresentam médias iguais aos valores ideais com elevada precisão.

5. CONCLUSÃO

Este trabalho explora o problema de detecção de estruturas em modelos NARX excitados por entradas suaves, que

se caracterizam por possuírem propriedades de excitação pobres. Tal entrada faz com que métodos tradicionais de seleção de estrutura, como o FROE, não consigam encontrar o modelo correto para o sistema. É proposto o uso de algoritmos de busca multi-objetivos (NSGA-II) para resolver o problema de seleção, tendo como objetivos a

Tabela 2. Distribuição da seleção de variáveis do sistema S_1 com entrada suave

Variáveis presentes no sistema			
Variáveis	GA-PEM	GA-SEM	FROE
<i>constante*</i>	100%	100%	100%
$y(k-1)$	100%	100%	55%
$y(k-2)$	100%	100%	93%
$u(k-1)^2$	100%	100%	56%
Outras variáveis			
$u(k-1)$	-	-	-
$u(k-2)$	-	-	-
$y(k-1)^2$	-	-	-
$y(k-1)y(k-2)$	-	-	-
$y(k-1)u(k-1)$	-	-	-
$y(k-1)u(k-2)$	-	-	-
$y(k-2)^2$	-	-	20%
$y(k-2)u(k-1)$	-	-	-
$y(k-2)u(k-2)$	-	-	6%
$u(k-1)u(k-2)$	-	-	7%
$u(k-2)^2$	-	-	63%

*FROE não pesquisa pelo termo constante, ele é adicionado forçadamente ao sistema S_1 .

Tabela 3. Distribuição da seleção de variáveis do sistema S_2 com entrada suave

Variáveis presentes no sistema			
Variáveis	GA-PEM	GA-SEM	FROE
$y(k-2)$	100%	82%	94%
$u(k-1)$	79%	82%	-
$y(k-2)u(k-1)$	100%	79%	83%
Outras variáveis			
<i>constante*</i>	1%	8%	-
$y(k-1)$	1%	5%	4%
$u(k-2)$	6%	11%	-
$y(k-1)^2$	-	5%	1%
$y(k-1)y(k-2)$	-	-	4%
$y(k-1)u(k-1)$	-	1%	-
$y(k-1)u(k-2)$	1%	3%	1%
$y(k-2)^2$	-	2%	17%
$y(k-2)u(k-2)$	-	4%	96%
$u(k-1)^2$	5%	9%	-
$u(k-1)u(k-2)$	6%	3%	-
$u(k-2)^2$	1%	6%	-

*FROE não pesquisa pelo termo constante.

Tabela 4. Parâmetros estimados por MQE - S_1

Parâmetro	Valor Real	Média	Desvio Padrão
θ_1	0,80	0,80	$9,30 \cdot 10^{-05}$
θ_2	1,20	1,20	$8,10 \cdot 10^{-05}$
θ_3	-0,90	-0,90	$6,50 \cdot 10^{-05}$
θ_4	-0,20	-0,20	$1,46 \cdot 10^{-03}$

Tabela 5. Parâmetros estimados por MQE - S_2

Parâmetro	Valor Real	Média	Desvio Padrão
θ_1	0,75	0,75	$1,18 \cdot 10^{-03}$
θ_2	0,25	0,25	$6,58 \cdot 10^{-04}$
θ_3	-0,20	-0,20	$4,20 \cdot 10^{-04}$

redução do erro do modelo e da quantidade de variáveis selecionadas (N_r). Duas funções custo são comparadas: minimização do erro de predição (PEM) e minimização do erro de simulação livre (SEM).

Os algoritmos GA-PEM, GA-SEM e FROE apresentaram resultados semelhantes na seleção de estrutura do modelo NARX quando o sistema em estudo é excitado com entrada ideal. Porém, quando o sistema é excitado com entrada suave, FROE perde drasticamente sua eficiência (de 97 – 99% para 0 – 14%), enquanto GA-PEM e GA-SEM diminuem seu sucesso apenas para um dos sistemas simulados (passando de 100% para 79% e 75%, respectivamente). O uso da técnica SEM fez com que o tempo de processamento do algoritmo fosse aproximadamente 7 vezes superior ao tempo de processamento do algoritmo com o uso da técnica PEM e aproximadamente 350 vezes superior ao do algoritmo FROE.

Verificou-se que, em se tratando do problema de seleção de estrutura por meio de GA, o uso de técnica SEM não apresentou melhores resultados que o uso de PEM, o que acontece no problema de estimação de parâmetros. Além disso, mesmo que o método FROE possua um bom desempenho nos casos em que a entrada do sistema é ideal e apresentando um custo computacional menor que os outros método, ele não seria recomendado. Dificilmente se sabe se uma entrada em um sistema real é ideal ou não. O algoritmo GA-PEM é, portanto, o mais confiável, por apresentar as melhores taxas de sucesso e baixo custo computacional comparado ao algoritmo GA-SEM.

Uma análise da distribuição da seleção de variáveis permitiu entender o problema causado pela entrada suave na seleção de estrutura pelo método FROE, explicado pela correlação entre os termos $u(i-1)$ e $u(i-2)$. Por fim, comprova-se que a entrada suave não influencia na estimação dos parâmetros pelo método dos mínimos quadrados estendido (MQE).

Este trabalho limitou-se a um espaço de busca pequeno, porém, em muitos casos, o alto grau de não-linearidade e elevados atrasos máximos fazem com que as matrizes de regressores sejam exageradamente grandes, o que torna o custo computacional muito elevado. Trabalhos futuros pretendem continuar explorando técnicas PEM e SEM na seleção de estruturas em modelos NARX com sinais mais ruidosos, além de buscar alternativas, como os modelos NARX-neurais.

REFERÊNCIAS

- Aguirre, L.A. (2004). *Introdução à Identificação de Sistemas - Técnicas Lineares e Não-Lineares Aplicadas a Sistemas Reais*. Editora UFMG, 2ª edition.
- Aguirre, L.A., Barbosa, B.H., and Braga, A.P. (2010). Prediction and simulation errors in parameter estimation for nonlinear systems. *Mechanical Systems and Signal Processing*, 24(8), 2855–2867.
- Aguirre, L.A. and Letellier, C. (2009). Modeling nonlinear dynamics and chaos: a review. *Mathematical Problems in Engineering*, 2009.
- Aguirre, L.A. and Billings, S.A. (1995). Dynamical effects of overparametrization in nonlinear models. *Physica D: Nonlinear Phenomena*, 80(1-2), 26–40.

- Barbosa, B.H.G., Aguirre, L.A., Martinez, C.B., and Braga, A.P. (2011). Black and gray-box identification of a hydraulic pumping system. *Control Systems Technology, IEEE Transactions on*, 19(2), 398–406.
- Barroso, M.S.F., Takahashi, R.H.C., and Aguirre, L.A. (2007). Multi-objective parameter estimation via minimal correlation criterion. *Journal of Process Control*, 17(4), 321–332.
- Billings, S., Chen, S., and Korenberg, M. (1989). Identification of mimo non-linear systems using a forward-regression orthogonal estimator. *International journal of control*, 49(6), 2157–2189.
- Chen, Q., Worden, K., Peng, P., and Leung, A. (2007). Genetic algorithm with an improved fitness function for (n) arx modelling. *Mechanical Systems and Signal Processing*, 21(2), 994–1007.
- Chen, S., Billings, S.A., and Luo, W. (1989). Orthogonal least squares methods and their application to non-linear system identification. *International Journal of control*, 50(5), 1873–1896.
- Connally, P., Li, K., and Irwing, G.W. (2007). Prediction- and simulation-error based perceptron training: Solution space analysis and a novel combined training scheme. *Neurocomputing*, 70(4-6), 819–827.
- Deb, K., Pratap, A., Agarwal, S., and Meyarivan, T. (2002). A fast and elitist multiobjective genetic algorithm: Nsga-ii. *IEEE Transactions on Evolutionary Computation*, 6(2), 182–197.
- Eiben, A.E., Smith, J.E., et al. (2003). *Introduction to evolutionary computing*, volume 53. Springer.
- Falsone, A., Piroddi, L., and Prandini, M. (2015). A randomized algorithm for nonlinear model structure selection. *Automatica*, 60, 227–238.
- Garg, A., Tai, K., and Panda, B. (2017). System identification: Survey on modeling methods and models. In *Artificial Intelligence and Evolutionary Computations in Engineering Systems*, 607–615. Springer.
- Goldberg, D.E. and Holland, J.H. (1988). Genetic algorithms and machine learning. *Machine learning*, 3(2), 95–99.
- Holland, J.H. (1975). Adaptation in natural and artificial systems. *The University of Michigan Press*.
- Leontaritis, I. and Billings, S.A. (1985). Input-output parametric models for non-linear systems part i: deterministic non-linear systems. *International journal of control*, 41(2), 303–328.
- Li, C.J. and Jeon, Y. (1993). Genetic algorithm in identifying non linear auto regressive with exogenous input models for non linear systems. In *1993 American Control Conference*, 2305–2309. IEEE.
- Linden, R. (2008). *Algoritmos genéticos*. Brasport, 3ª edition.
- Madar, J., Abonyi, J., and Szeifert, F. (2005). Genetic programming for the identification of nonlinear input - output models. *Industrial & Engineering Chemistry Research*, 44(9), 3178–3186.
- Milanese, M. and Novara, C. (2005). Model quality in identification of nonlinear systems. *IEEE Transactions on Automatic Control*, 50(10), 1606–1611.
- Piroddi, L. (2008). Simulation error minimization methods for NARX model identification. *Int. J. Modeling, Identification and Control*, 3(4), 392–403.
- Piroddi, L. and Spinelli, W. (2003). An identification algorithm for polynomial narx models based on simulation error minimization. *International Journal of Control*, 76(17), 1767–1781.
- Pope, K.J. and Rayner, P.J. (1994). Non-linear system identification using bayesian inference. In *Proceedings of ICASSP'94. IEEE International Conference on Acoustics, Speech and Signal Processing*, volume 4, IV–457. IEEE.
- Retes, P.F.L. and Aguirre, L.A. (2019). Narmax model identification using a randomised approach. *International Journal of Modelling, Identification and Control*, 31(3).
- Ribeiro, A.H. and Aguirre, L.A. (2018). “parallel training considered harmful?”: Comparing series-parallel and parallel feedforward network training. *Neurocomputing*, 316, 222 – 231.
- Rodriguez-Vazquez, K., Fonseca, C.M., and Fleming, P.J. (2004). Identifying the structure of nonlinear dynamic systems using multiobjective genetic programming. *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans*, 34(4), 531–545.
- Sjöberg, J., Zhang, Q., Ljung, L., Beneviste, A., Delyon, B., Glorennec, P., Hjalmarsson, H., and Juditsky, A. (1995). *Non-linear black-box modeling in system identification: A unified overview*. Automatica.
- Young, P.C. (1968). The use of linear regression and related procedures for the identification of dynamic processes. In *Seventh Symposium on Adaptive Processes*, 53–53. IEEE.