DETERMINAÇÃO DO NÚMERO DE TERMOS DE MODELOS POLINOMIAIS: O CRITÉRIO DE INFORMAÇÃO DE AKAIKE MULTIOBJETIVO

Samir Angelo Milani Martins*, Alípio Monteiro Barbosa†, Erivelton Geraldo Nepomuceno*

*GCoM – Grupo de Controle e Modelagem, Departamento de Engenharia Elétrica
Universidade Federal de São João del-Rei, Praça Frei Orlando 170 - Centro, 36307-352
São João del-Rei, Minas Gerais, Brasil

†Programa de Pós-Graduação em Engenharia Elétrica
Universidade Federal de Minas Gerais - Av. Antônio Carlos 6627, 31270-901
Belo Horizonte, MG, Brasil

Emails: `martins@ufsj.edu.br`, `alipiomonteiro@yahoo.com.br`, `nepomuceno@ufsj.edu.br`

Abstract— This paper presents a multiobjective formulation, based on the Akaike information criterion in order to assist the process of determining the structure of mathematical models of real systems. It is proposed the Multiobjective Akaike Information Criterion (MAIC), which considers in the modeling additional features to dynamic system information. This approach is employed in the modeling of a DC-DC *Buck* converter, where static (theoretical) and dynamic (experimental) data are considered aiming to obtain a polynomial NARX model for the system.

Keywords— System identification, multiobjective optimization, structure selection, information criterion, NARX models.

Resumo— O presente trabalho apresenta uma formulação multiobjetivo, baseado no critério de informação de Akaike, a fim de auxiliar o procedimento de determinação de estruturas de modelos matemáticos de sistemas reais. É proposto o critério de informação de Akaike multiobjetivo (MAIC), o qual considera no processo de modelagem características adicionais à informação dinâmica do sistema. Tal abordagem é empregada na modelagem de um conversor CC-CC *Buck*, em que são considerados dados estáticos (teóricos) e dinâmicos (experimentais) para obtenção de um modelo NARX polinomial para o sistema.

Palavras-chave— Identificação de sistemas, otimização multiobjetivo, seleção de estrutura, critério de informação, modelos NARX.

1 Introdução

A utilização de informação auxiliar na determinação de estruturas de modelos polinomiais NARX é uma forma eficiente de obter modelos estáveis e representativos. Geralmente, a identificação de sistemas utiliza apenas dados de entrada e saída, deixando para a etapa de validação a verificação se o modelo é estável e representa características desejadas e atende a índices pré-determinados. A partir de trabalhos de Johansen (1996) e colaboradores, a área de identificação de sistemas começou a se preocupar com a possibilidade de utilizar informações auxiliares no processo de modelagem. Dessa preocupação, surge uma nova técnica, chamada de Identificação Multiobjetivo de Sistemas (IMS) (Martins et al., 2011; Nepomuceno et al., 2007; Barroso et al., 2007). Inicialmente, o uso de informações auxiliares se dá exclusivamente na consolidada etapa de estimação dos parâmetros do modelo. Recentemente, (Martins et al., 2013) propuseram a taxa de redução de erro multiobjetivo, técnica capaz de quantificar a contribuição de cada regressor à luz de um critério multiobjetivo. Os resultados apresentados foram promissores, permitindo que estruturas fossem escolhidas levando-se em conta características dinâmicas e estáticas.

Com o objetivo de continuar a pesquisa realizada em (Martins et al., 2013), este trabalho propoe o desenvolvimento do critério de informação de Akaike multiobjetivo (MAIC, do inglês *Multiobjective Akaike Information Criterion*), para auxílio na determinação de estruturas de modelos NARX polinomiais. Tal técnica é vista como uma expansão multiobjetivo do critério de informação de Akaike, a qual utiliza informações além daquelas imbuídas nos dados dinâmicos para se estimar um número *ideal* de regressores de um modelo.

O restante do artigo está organizado da seguinte forma. Na seção 2 são abordados os conceitos preliminares . A seção 3 apresenta a metodologia para o desenvolvimento do MAIC e consequente obtenção dos resultados. A análise e discussão dos resultados são tratados na seção 4. Na seção 5 é apresentada a conclusão, além de propostas para futuras pesquisas.

2 Propriedades dos modelos NARX

2.1 Modelos NARX

Modelos NARX (Billings, 1980) descrevem sistemas não-lineares por meio de equações de diferença, relacionando a saída atual por meio da combinação das saídas e entradas passadas. Esse tipo de modelo pode ser utilizado, entre outras aplicações, para problemas de controle, cujo principal objetivo é encontrar uma descrição simples para o sistema. Em particular, o modelo NARX polinomial pode ser representado como:

$$\begin{aligned} y(k) &= F^{\ell}[y(k-1), \cdots, y(k-n_y), \\ & \quad u(k-1), \cdots, u(k-n_u), \\ & \quad e(k-1), \cdots, e(k-n_e)] + \Xi(k), \end{aligned} \quad (1)$$

em que $y(k)$ é a saída, $u(k)$ é a entrada exógena, e $e(k)$ é o sinal de ruído. $\Xi(k)$ representa o erro de predição. n_y, n_u, e n_e são as ordens da saída, da entrada exógena e do ruído, geralmente modelado por um processo de média móvel. Neste trabalho, os parâmetros do modelo foram estimados via mínimos quadrados estendidos, processo no qual o ruído é modelado por um processo média móvel (MA), estimado pelo resíduo. Tal estimador foi utilizado visando evitar a polarização na estimação. A função F^{ℓ} pode representar uma grande variedade de funções, incluindo funções lineares. Neste artigo, F^{ℓ} é restrita a funções polinomiais, lineares ou não.

2.2 *Determinação da Estrutura via ERR*

A taxa de redução de erro - ERR (do inglês *Error Reduction Ratio*) é uma das técnicas utilizadas para detecção de estruturas de modelos NARX polinomiais. Tal técnica baseia-se no erro dinâmico de predição, quando visto um passo à frente, associando cada termo candidato à um índice correspondente à contribuição na explicação da variância dos dados de saída do sistema. A variância dos dados de saída pode ser definida como:

$$\mathrm{var}[\xi(k)] = \lim_{N\to\infty} \frac{1}{N}\left[y^T y - \sum_{i=1}^{N} g_i^2 \Omega_i^T \Omega_i\right], \quad (2)$$

em que g_i indica cada elemento do vetor de parâmetros g, Ω_i representa os regressores do modelo ortogonal associado e y representa a série temporal dos dados de saída e N o tamanho da série. O modelo ortogonal associado pode ser obtido por meio de transformações de Householder do modelo original.

Caso nenhum termo fosse acrescentado ao modelo, a variância $\xi(k)$ seria exatamente igual ao erro quadrático de saída ($N = 0$). A cada termo acrescentado ao modelo, a variância decresce um valor de $\frac{1}{N}(g_i^2 w_i^T w_i)$, sendo w_i o regressor incluído e g_i o respectivo parâmetro ortogonal associado (Aguirre, 2007). O fator de redução da variância devido a inclusão de um dado termo pode ser normalizada em relação ao erro quadrático médio do sinal de saída. Deste modo, tem-se a ERR como sendo:

$$[\mathrm{ERR}] = \frac{(g_i^2 w_i^T w_i)}{y^T y}. \quad (3)$$

A taxa de redução de erro pode ser utilizada para quantificar cada regressor, ou seja, qual a parcela de cada regressor na contribuição da redução da variância dos resíduos. Contudo, acredita-se que o número de termos que devem ser inseridos no modelo não possa ser determinado por esta técnica, sendo necessário a utilização de uma técnica auxiliar como o critério de informação de Akaike (AIC).

2.3 *O Critério de Informação de Akaike - AIC*

Alguns dos procedimentos que podem ser utilizados para a determinação do número de termos que devem ser inclusos no modelo são os critérios de informação. Especificamente, o critério de informação de Akaike fornece o número de regressores que devem ser inclusos no modelo, de forma que o mesmo represente de forma adequada o sistema. Quando aplicada em conjunto com a ERR (que classifica os termos de modo hierárquico), pode fornecer um modelo, baseado em dados dinâmicos, para representar um sistema.

De acordo com esta técnica, o número de termos de um modelo deve minimizar a função custo AIC:

$$\mathrm{AIC} = N \log(\mathrm{var}[\xi(k)]) + 2n_{\theta}, \quad (4)$$

em que N é o tamanho da série temporal dos dados de identificação, n_{θ} o número de regressores candidatos ao modelo e $\mathrm{var}[\xi(k)]$ é a variância do erro dinâmico de modelagem (resíduos ou erro de predição um passo a frente). O resultado obtido pelo AIC pode ser visto como um número *ótimo* de termos que devem ser inseridos no modelo do sistema (Barbosa et al., 2010; Nepomuceno, 2002).

Porém, ressalta-se como grande fator limitante das técnicas (ERR e AIC) a utilização exclusiva de dados dinâmicos e das informações ali contidas, para a quantificação da contribuição de cada regressor e do número de regressores a serem utilizados na composição do modelo final. Conjectura-se que a inclusão de demais informações acerca do sistema, ao se estimar o modelo, permita a obtenção de modelos mais globais e representativos, válidos em uma faixa de operação mais ampla.

Ademais, o desenvolvimento de técnicas capazes de considerar diferentes características do sistema pode ser também justificada pelo fato de existirem sistemas em que não se pode coletar dados com informações dinâmicas suficientes para compor um modelo. Ademais, muitas vezes os dados estão contaminados por ruídos com relação sinal-ruído muito baixa. Assim sendo, busca-se alternativas, de modo a possibilitar a obtenção de modelos de sistemas, utilizando outras informações que não dinâmicas.

3 Critério de Informação de Akaike Multiobjetivo

Uma das principais características de modelos NARX polinomiais é a sua relativa facilidade de extração de informações (ganho estático, curva estática, pontos fixos) acerca do sistema, estimadas pelo modelo. Nepomuceno et al. (2007) como curva estática, localização dos pontos fixos e outras informações podem ser estimadas por modelos NARX polinomiais, em função da multiplicação de uma matriz de regressores pelo vetor de parâmetros. Nesse sentido, considerando a $q-$ésima informação y_q [1], estimada por $\hat{y_q}$, o erro de estimação (resíduo) é dado por:

$$\xi_q = y_q - \hat{y_q}. \tag{5}$$

O AIC, por sua vez, considera que existe apenas uma informação a ser considerada no modelo (informação dinâmica). Dessa forma, busca estimar um número de termos *ótimo* a ser utilizado no processo de modelagem. Nesse índice, tem-se que a primeira parcela da equação 4 é reduzida, mediante a inclusão de regressores. Em contrapartida, a segunda parcela cresce significativamente (para o AIC, multiplicada por 2) com o aumento do número de termos. Assim, o AIC fornece um número de termos a serem inseridos no modelo com compromisso simultâneo entre complexidade e redução da variância do resíduo.

De forma análoga, buscou-se o desenvolvimento de um critério que considerasse, além dos resíduos dinâmicos, resíduos referente às diferentes características do processo modelado. É definido dessa forma o critério de informação de Akaike multiobjetivo (MAIC - *Multiobjective Akaike Information Criterion*), expansão do critério de informação de Akaike, como sendo:

$$\begin{aligned} \mathrm{MAIC}(n_\theta) &= w_1 N_1 \ln(\mathrm{var}\{\xi_1\}) + \cdots + \\ & \quad w_z N_z \ln(\mathrm{var}\{\xi_z\}) + 2n_\theta, \\ &= \sum_{i=1}^{z} (w_i N_i \ln(\mathrm{var}\{\xi_i\})) + 2n_\theta, \end{aligned} \tag{6}$$

sendo $\sum_{i=1}^{z} w_i = 1$ os pesos associados à cada informação, z o número de características do processo consideradas no processo de modelagem, N_i o tamanho do $i-$ésimo vetor de resíduos e n_θ o número de termos incluídos. Importante notar, que para o caso em que se considera apenas informação dinâmica, tem-se $w_1 = 1$ e a Equação 6 se torna equivalente a Equação 4.

Do ponto de vista multiobjetivo, o número *ideal* de termos se dá quando o índice MAIC é minimizado. Isso significa que, com tal número de regressores, há um compromisso mútuo ao minimizar a variância dos resíduos considerados e ao escolher o modelo mais simples possível, para aquele conjunto de ponderações utilizado.

O fator 2, que pode ser alterado, foi mantido, uma vez que o MAIC é a expansão multiobjetivo do AIC. É importante destacar que um conjunto de soluções eficientes podem ser obtidas, variando o valor de w_i. Ademais, o método permite, com a variação dos pesos, obter uma faixa de valores na qual o tamanho do modelo deve ser inserido, de modo a obter modelos representativos, considerando diferentes aspectos do processo modelado.

4 Resultados

O MAIC será aplicado neste trabalho na determinação do número de termos de um modelo polinomial NARX para dois objetivos. Inicialmente, para ilustrar a aplicação do MAIC a Figura 1 ilustra um processo de modelagem em que o MAIC foi estimado 5 vezes, considerando 5 conjuntos de pesos distintos. A princípio, suponha que sejam considerados dados dinâmicos e estáticos no processo de modelagem. Nesse caso, para um conjunto de pesos igualmente espaçados:

$$W = [W_{\mathrm{D}}, W_{\mathrm{S}}] = \begin{bmatrix} 1 & 0 \\ 0{,}75 & 0{,}25 \\ 0{,}50 & 0{,}50 \\ 0{,}25 & 0{,}75 \\ 0 & 1 \end{bmatrix}. \tag{7}$$

Cada linha da matriz W, quando aplicada à equação 6, fornece um número de termos que minimiza o índice MAIC (Figura 1, à esquerda). No caso bi-objetivo considerado, cada solução pode ser mapeada em uma fronteira Pareto-ótima, em que os funcionais F_1 e F_2 são relativos aos erros quadráticos dinâmico e estático, respectivamente.

Para o caso hipotético considerado, um modelo com um número de regressores entre $2 \leq n_\theta \leq 9$ seria *ótimo* de acordo com as informações utilizadas para identificação. Sendo assim, o procedimento de seleção de estruturas se segue, em que se escolhe utilizando alguma técnica apropriada (ERR, por exemplo) quais regressores serão inseridos no processo.

Após ilustrar o uso do MAIC, passa-se para aplicação da proposta para uma sistema real. Um conversor CC-CC buck, excitado por um sinal PRBS, foi utilizado para identificação e verificação da técnica proposta. Tal sistema foi escolhido por apresentar característica estática afim e dinâmica não-linear. Isso faz com que a modelagem, de um ponto de vista global, considerando diferentes aspectos e características do sistema, se torne complexa. Os regressores foram selecionados por meio da ERR, sendo o número de regressores de cada modelo estimado pelo MAIC.

[1] Por exemplo, dados dinâmicos, curva estática, ganho estático e localização dos pontos fixos.

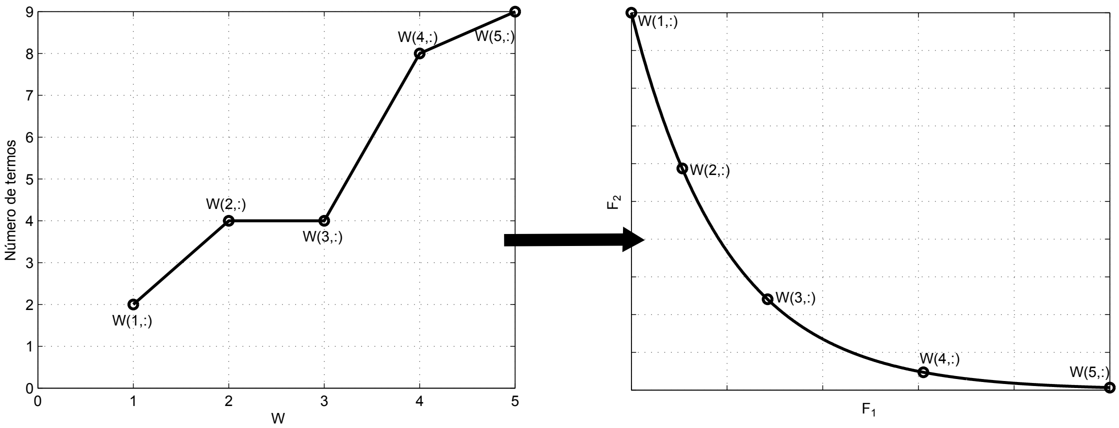

Figura 1: Representação gráfica do MAIC. O gráfico à esquerda mostra a solução mínima do MAIC para um conjunto de ponderações. À direita, o mapeamento em sua respectiva curva Pareto-ótimo.

Considerou-se para a presente aplicação a incorporação de informações estáticas e dinâmicas no modelo. Portanto, o MAIC (expresso pela equação 6) é definido para o caso bi-objetivo, como:

$$\begin{aligned} \mathrm{MAIC}(n_\theta) &= w_{\mathrm{D}} N_{\mathrm{D}} \ln(\mathrm{var}\{\xi_{\mathrm{D}}\}) + \\ & \quad w_{\mathrm{S}} N_{\mathrm{S}} \ln(\mathrm{var}\{\xi_{\mathrm{S}}\}) + 2n_\theta, \end{aligned} \tag{8}$$

em que o subíndice D se refere à informação dinâmica e S se refere à informação estática.

A fim de avaliar o comportamento da técnica em diferentes pontos de operação, escolheu-se pesos linearmente espaçados para w_{D} e w_{S}, sendo tais valores apresentados na equação 7.

Do critério de informação de Akaike mono-objetivo, considera-se o ponto de dobra, ou o valor de mínimo, como sendo uma boa estimativa para o número de termos do modelo. A Figura 2 apresenta os valores de mínimos para as curvas da Figura 3, sendo o primeiro ponto da curva referente ao pares de pesos $\langle w_S = 1, w_D = 0\rangle$, até o último ponto, com 9 regressores, para os pesos $\langle w_S = 0, w_D = 1\rangle$.

Como pode-se observar, para o conjunto de pesos investigado obtém-se modelos com 1, 4 e 9 regressores, os quais minimizam o índice MAIC, para diferentes combinações de pesos. Um aspecto bastante interessante é que tal ferramenta fornece, baseado na variações dos pesos, uma faixa na qual o tamanho do modelo deve ser inserido, de modo a compor um modelo fiel ao sistema. No presente caso, um modelo representativo, do ponto de vista estático e dinâmico, seriam modelos com, no máximo, 9 regressores.

Portanto, três modelos foram selecionados a partir do MAIC e ERR. Os modelos com 1, 4 e 9 regressores, são definidos a seguir.

$$y_1(k) \quad = \quad 0{,}9998y(k-1)$$

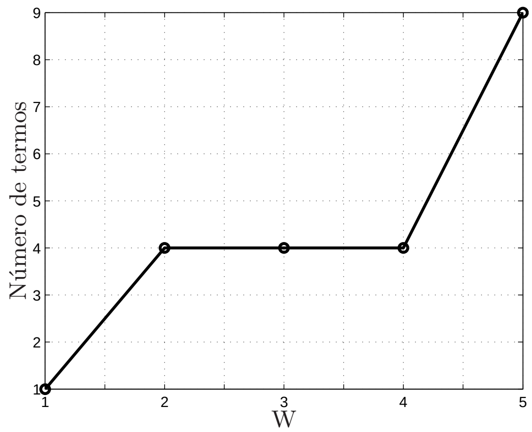

Figura 2: Solução mínima do MAIC para as ponderações apresentadas na Figura 3.

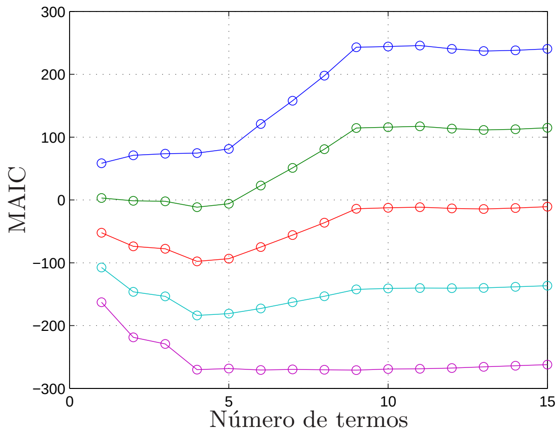

Figura 3: Valores do MAIC para um conjunto de ponderações. De baixo para cima, as ponderações estáticas e dinâmicas: $[W_D, W_S]$ = [(1;0), (0,75;0,25), (0,50;0,50), (0,25;0,75), (0;1)].

$$
\begin{aligned}
y_2(k) &= 2{,}2574y(k-1) - 0{,}6369y(k-2) \\
&\quad -0{,}0331y(k-1)^2 - 0{,}4200u(k-1)^2
\end{aligned}
$$

$$
\begin{aligned}
y_3(k) &= -1{,}3299y(k-1) - 0{,}5648y(k-2) \\
&\quad +0{,}2568y(k-1)^2 + 24{,}0242u(k-1)^2 \\
&\quad +0{,}2285y(k-2)^2 \\
&\quad -0{,}4548y(k-2)y(k-1) \\
&\quad +170{,}9235 + 0{,}7591u(k-1)y(k-1) \\
&\quad -125{,}4085u(k-1)
\end{aligned}
$$

As Figuras 4 e 5 ilustram os comportamentos dinâmico e estático dos modelos selecionados. O modelo y_1 apresentou um elevado erro para simulação dinâmica, porém o ajuste estático foi preciso. Por outro lado, o modelo y_3 apresentou um bom ajuste dinâmico e uma aproximação estática ruim. Os dois modelos, y_1 e y_3, representam os extremos da curva Pareto (ver Figura 1). Uma desempenho intermediário foi obtido com o modelo y_2, modelo de quatro termos.

O desempenho dos modelos são quantificados pelos índices RMSE (estático e dinâmico)[2] e custo quadrático dinâmico e estático. Os valores de tais índices, ilustrados nas Figuras 6 e 7, mostram claramente que pode existir um modelo intermediário, que represente, simultaneamente, características dinâmicas e estáticas de um sistema.

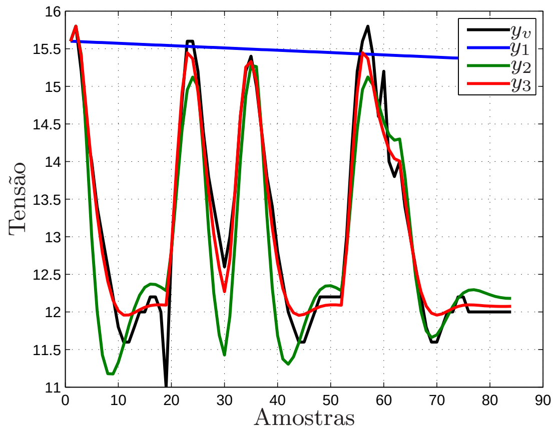

Figura 4: Simulação livres. y_v são os dados de validação, y_1 a resposta do modelo com um termo, y_2 o modelo com quatro termos e y_3 o modelo com nove termos.

5 Conclusão

O presente trabalho apresentou uma formulação para o critério de informação de Akaike multiobjetivo, MAIC. A formulação proposta é uma

[2]Para detalhes matemáticos dos índices, veja (Martins et al., 2012)

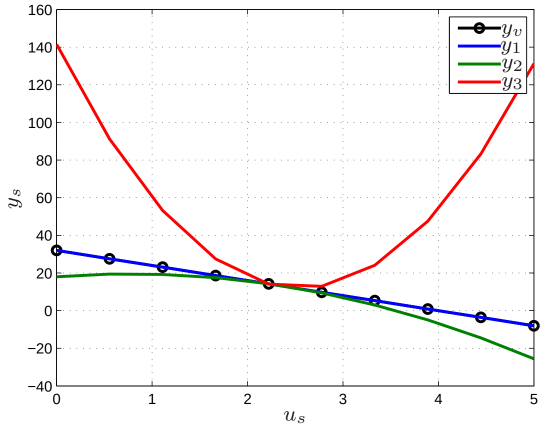

Figura 5: Resposta estática. y_v são os dados de validação, y_1 a resposta do modelo com um termo, y_2 o modelo com quatro termos e y_3 o modelo com nove termos.

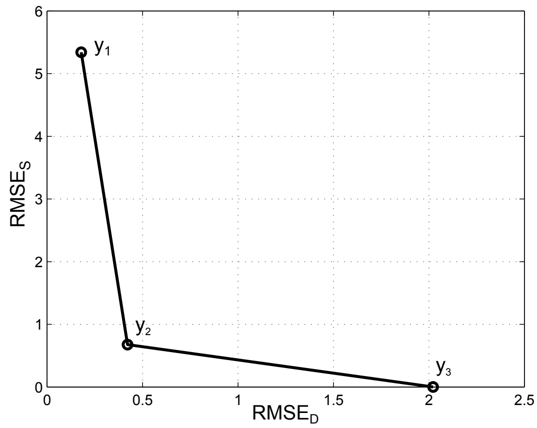

Figura 6: Índices RMSE estático e dinâmico dos modelos y_1, y_2 e y_3.

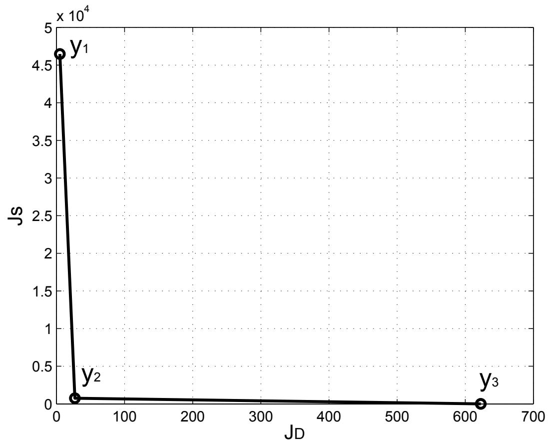

Figura 7: Custos quadráticos estático e dinâmico dos modelos.

extensão multiobjetivo do AIC e permite um auxílio adicional na etapa da seleção de estruturas de modelos polinomiais.

A proposta foi ilustrada por meio de um exemplo experimental, em que dois objetivos foram considerados: minimizar simultaneamente erros dinâmico e estático. Incorporar a informação da curva estática fez com que o modelo fosse melhor representado. Ou seja, a um baixo custo no desempenho dinâmico, obteve-se um modelo com um característica estática melhor ajustada.

Uma vez que não se conhece o modelo original do sistema, não é possível afirmar que um conjunto de regressores são verdadeiros ou não. Contudo, pode-se concluir, que a possibilidade de incorporar outras informações é um benefício adicional e pode proporcionar modelos mais gerais.

Algumas questões ainda devem ser investigadas, dentre elas, destaca-se duas possibilidades. A primeira consiste em verificar a melhor forma de distribuir o peso de cada objetivo e a segunda, definir uma metodologia para tomada de decisão quando o critério de informação multiobjetivo sugere o uso de mais de um modelo.

Agradecimentos

Os autores agradecem o Professor Luis Aguirre pelas sugestões e a Fapemig, o CNPq e a Capes, pelo apoio financeiro.

Referências

Aguirre, L. A. (2007). *Introdução à Identificação de Sistemas: Técnicas lineares e não-lineares aplicadas a sistemas reais*, Editora da UFMG. 3^a edição.

Barbosa, A. M., Takahashi, R. H. C. e Aguirre, L. A. (2010). Técnicas de otimização bi-objetivo aplicadas a problemas de determinação de estrutura em modelos polinomiais NARX, *XVIII Congresso Brasileiro de Automática*, Bonito – MS, pp. 3189–3196.

Barroso, M. F. S., Takahashi, R. e Aguirre, L. (2007). Multi-objective parameter estimation via minimal correlation criterion, *Journal of Process Control* **17**(4): 321–332.

Billings, S. A. (1980). Identification of nonlinear systems - a survey, *IEE Proceedings-D Control Theory and Applications* **127**(6): 272–285.

Johansen, T. A. (1996). Identification of non-linear systems using empirical data and prior knowledge - an optimization approach, *Automatica* **32**(3): 337–356.

Martins, S. A. M., Nepomuceno, E. G. e Barroso, M. F. S. (2012). Detecção de Estruturas de Modelos NARMAX Polinomiais: a Taxa de Redução de Erro Multi-objetivo (MERR), *Anais do XIX Congresso Brasileiro de Automática. Campina Grande, PB.*

Martins, S. A. M., Nepomuceno, E. G. e Barroso, M. F. S. (2013). Improved Structure Detection for Polynomial NARX Models Using a Multiobjective Error Reduction Ratio, *Journal of Control, Automation and Electrical Systems* . Aceito para publicação.

Martins, S. A. M., Nepomuceno, E. G. e Figueiredo, J. P. (2011). Detecção de estruturas de modelos narmax polinomiais: uma abordagem inteligente multi-objetivo., *Anais do X Simpósio Brasileiro de Automação Inteligente*, pp. 320–325.

Nepomuceno, E. G. (2002). *Identificação multiobjetivo de sistemas não-lineares*, Dissertação de Mestrado, PPGEE, Universidade Federal de Minas Gerais, Belo Horizonte, Brasil.

Nepomuceno, E. G., Takahashi, R. H. C. e Aguirre, L. A. (2007). Multiobjective parameter estimation for non-linear systems: affine information and least-squares formulation, *International Journal of Control* **80**(6): 863–871.