

# Combinatorial Network of Dynamic Models: A Method to Improve Bad-Quality Models

L. da Costa, S. Milani, and G. Vasconcelos

**Abstract**—This work proposes a combinatorial network of dynamic models in order to combine models that doesn't approximate well a system, to obtain an output that have a better performance considering the system behaviour. The network can be useful in situations where a good model cannot be obtained from data, such when there is a bad signal noise ratio in identification data-set or when a specific input cannot be applied to generate identification data. To do this, we present two different approaches: an analytical one and a numerical method, both combining a weighted sum of the bad quality models. The method is tested on models obtained through a multi-objective system identification procedure, and from models obtained through an interval system identification procedure. The combined model has improved the performance in the validation indexes analyzed, reaching a reduction up to 65% in the RMSE index, 95% in the MSE index of the static curve, 87% in the energy of the residues vector and a reduction of 21% in the auto-correlation energy of the residues vector.

**Index Terms**—Combinatorial network of dynamic models, NARX polynomial, System identification, Combined models.

## I. INTRODUÇÃO

**I**DENTIFICAÇÃO de sistemas é a área de estudo responsável pela obtenção modelos matemáticos para sistemas experimentais. Um tipo de representação bastante utilizada é a representação NARX (*Nonlinear AutoRegressive with eXogenous input*) [1] das quais se destaca sua variante polinomial [2] bastante popular por suas características, como ser função suave, possuir modelos que representam o ruído, possuir uma interpretação direta e a disponibilidade de procedimentos de seleção de estrutura e ajuste de parâmetros eficientes.

Nem sempre é possível obter modelos adequados, seja por dificuldades na coleta de dados, mal condicionamento numérico ou por alta complexidade do sistema. Uma solução é a inserção de conhecimento prévio como em [3] onde é utilizado o *Multiobjective Error Reduction Ratio* MERR para obter um conjunto de modelos que minimizam objetivos estático e dinâmico. Procedimentos de identificação multiobjetivo de forma geral resultam em uma família de modelos que minimizam os múltiplos objetivos, sendo este conjunto denominado Pareto-Ótimo. Neste cenário é necessário escolher entre os modelos identificados, qual será utilizado para a referida aplicação.

Outras técnicas de identificação de sistemas que resultam em um conjunto de modelos podem ser encontradas na literatura como é o caso de [4], em que a aritmética

intervalar é empregada na estimação de parâmetros de modelos polinomiais.

Este trabalho propõe uma rede combinatória de modelos dinâmicos CNDM (*Combinatorial Network of Dinamic Models*) que combina múltiplos modelos de forma paralela. A abordagem é voltada para os casos onde existem restrições no processo de identificação e bons modelos não podem ser obtidos. A metodologia propõe a utilização de modelos com baixa representatividade do sistema, com intuito de gerar uma única saída que possua melhor desempenho ao representar suas características. São propostas duas abordagens, uma analítica e outra numérica, sendo a segunda baseada no algoritmo gradiente descendente.

Na literatura podem ser vistas técnicas onde mais de um modelo são combinados, como em [5], onde é proposto o conjunto estendido de modelos ou *extended model set*. Nesta abordagem é utilizada uma mistura de funções de base em um mesmo modelo NARMAX para a identificação de sistemas bruscamente não lineares. O modelo resultante pode então ser considerado a combinação de modelos NARMAX com funções de base distintas. Esta abordagem difere da metodologia proposta uma vez que não utiliza a combinação de modelos com a mesma função de base como por exemplo múltiplos modelos polinomiais. A técnica proposta por este trabalho combina modelos NARMAX previamente obtidos, independentemente de suas funções de base.

Outra técnica que se relaciona com a metodologia desenvolvida é a computação de reservatório ou RC [6] (do inglês *Reservoir Computing*). RC é um tipo de modelo computacional capaz de modelar uma vasta quantidade de sistemas não lineares [7]. Se trata de uma rede composta por uma camada de entrada, um reservatório e uma camada de saída. A camada de entrada é mapeada por meio de funções não lineares em um espaço de alta dimensão chamado de reservatório que pode ser visto como uma expansão não linear e de alta dimensão da entrada. Por fim os pesos lineares da camada de saída são ajustados, sendo a saída do modelo dada pela soma dos estados internos do reservatório ponderada pelos pesos da camada de saída [8]. Sua principal diferença para a metodologia proposta está na dimensão e estrutura do reservatório, uma vez que na RC o reservatório é composto por um grande número de estados e sua estrutura exata não é essencial e muitas vezes aleatória [9]. Por outro lado, a CNDM combina um número reduzido de estados internos que possuem as estruturas definidas por procedimentos de identificação de sistemas.

Como estudo de caso foram utilizados modelos pertencentes a um conjunto Pareto-Ótimo obtidos por meio do MERR disponível em [3]. Também foram utilizadas representações pertencentes a um conjunto de modelos obtidos por meio da

L. Andrade, S. Martins and G. Amaral was with the Department of Electrical Engineering, Universidade Federal de São João del Rei, São João del Rei, MG, BR e-mail: martins@ufsj.edu.br.

identificação de sistemas intervalar disponível em [4].

O restante do trabalho é dividido da seguinte forma: A Seção 2 trás conceitos prévios suficientes para o entendimento do trabalho. A Seção 3 apresenta a metodologia adotada sendo seguida pelos resultados na Seção 4. Por fim a Seção 5 apresenta as conclusões do trabalho, bem como perspectivas de trabalhos futuros.

## II. DEFINIÇÕES

Esta Seção apresenta um conjunto de definições que auxiliam o entendimento do presente trabalho.

**Definição 2.1 (Modelos NARX Polinomiais):** Os modelos NARX (*Nonlinear AutoRegressive with eXogenous input*) polinomiais são representações auto regressivas representadas como em (1):

$$y_k = F^\ell[y_{k-1}, \dots, y_{k-n_y}, u_{k-1}, \dots, u_{k-n_u}] + e_k, \quad (1)$$

em que  $y_k$  e  $u_k$  são a saída e a entrada respectivamente,  $n_y$  e  $n_u$  os máximos atrasos correspondentes, e  $e_k$  o resíduo.  $F^\ell[\cdot]$  é uma função polinomial de grau  $\ell$  [10].

**Definição 2.2 (Gradiente Descendente):** O Gradiente Descendente é um algoritmo de otimização iterativo utilizado para minimizar funções custo, como o erro quadrático de predição que pode ser representado pela Equação (2):

$$J = (X\omega - y)^T(X\omega - y), \quad (2)$$

em que  $J$  é o erro quadrático,  $X$  é a matriz de regressão,  $y$  a saída esperada e  $\omega$  os parâmetros. A matriz  $\omega$  é atualizada a um passo  $\alpha$  na direção do gradiente da função custo em relação aos parâmetros ( $\frac{\partial J}{\partial \omega}$ ) como mostrado na Equação (3).

$$\Delta\omega = \alpha(y - \bar{y})^T X, \quad (3)$$

em que  $\bar{y}$  é a predição do modelo a cada iteração.

**Definição 2.3 (Mínimos Quadrados Intervalares):** Os mínimos quadrados intervalares [4] insere intervalos no método de mínimos quadrados, com o intuito de manusear erros numéricos. Desta forma os parâmetros ajustados são um intervalo de valores. Os mínimos quadrados intervalares pode ser calculado como:

$$\hat{\theta}_I = [\Psi_I^T \Psi_I]^{-1} \Psi_I^T y, \quad (4)$$

em que  $\hat{\theta}_I = [\hat{\theta}, \bar{\theta}]$ , é o intervalo de coeficientes calculados  $\Psi_{I,i,j} = [\Psi_{i,j}, \bar{\Psi}_{i,j}]$  é a matriz de regressores intervalar, em que  $i$  e  $j$  são os índices da linha e da coluna da matriz  $\Psi_I$ , e  $y$  é o vetor de dados.

**Definição 2.4 (Raiz do Erro Quadrático Médio):** O RMSE é um índice utilizado para avaliar o desempenho dinâmico de modelos matemáticos, comparando o erro cometido pelo modelo com o erro médio para a série temporal. O RMSE que pode ser calculado como em (5):

$$\text{RMSE} = \frac{\sqrt{\sum_{k=1}^N (y_k - \hat{y}_k)^2}}{\sqrt{\sum_{k=1}^N (y_k - \bar{y}_k)^2}}, \quad (5)$$

em que  $\hat{y}_k$  é a saída do modelo,  $y_k$  a saída real do sistema e  $\bar{y}_k$  sua média em  $N$  amostras [11].

**Definição 2.5 (Erro Quadrático Médio):** O MSE é um índice que calcula a média dos erros quadráticos entre valores de uma predição e os valores reais a serem preditos, e pode ser calculado como em (6) [11]:

$$\text{MSE} = \frac{1}{N} \sum_{k=1}^N (y_k - \hat{y}_k)^2, \quad (6)$$

**Definição 2.6 (Modelo de má Qualidade):** Define-se um modelo de má qualidade como uma representação matemática que não atenda a um requisito pré definido para uma dada aplicação.

**Observação 2.1:** No contexto deste trabalho serão considerados modelos de má qualidade aqueles que apresentarem  $\text{RMSE} \geq 0,25$ .

**Definição 2.7 (Correlação Cruzada Entre Vetor de Resíduos e Entrada):** Função capaz de verificar se os erros cometidos pelo modelo são correlacionados com a entrada, indicando que alguma característica do sistema experimental não foi explicado pelo modelo. Calcula-se como (7):

$$r_{u\xi}(k) = \lim_{N \rightarrow \infty} \frac{1}{2N+1} \sum_{i=-n}^n u(i)\xi(i+k), \quad (7)$$

em que  $\xi$  é o vetor de resíduos,  $u$  é o vetor de entrada e  $k$  indica o tempo discreto.

**Definição 2.8 (Auto Correlação do Vetor de Resíduos):** Função capaz de verificar se os erros cometidos pelo modelo possuem características aleatórias. Calcula-se como 8:

$$r_{\xi\xi}(k) = \lim_{N \rightarrow \infty} \frac{1}{2N+1} \sum_{i=-n}^n \xi(i)\xi(i+k). \quad (8)$$

**Definição 2.9 (Energia de um sinal):** A energia de um sinal  $x$  é uma medida de seu tamanho e pode ser calculada como (9):

$$E_x = \int_{-\infty}^{\infty} x^2(t)dt, \quad (9)$$

para  $E_x$  sendo a energia do sinal  $x$ .

## III. METODOLOGIA

Este trabalho propõe um método que combina a saída de modelos paralelamente, como pode ser visto na Figura 1, de forma que as saídas destes modelos possam compor os estados internos da nova representação. Os modelos podem ser obtidos por meio de procedimentos de identificação multi-objetivo, identificação intervalar ou quais quer outros métodos de identificação sendo o intuito a melhoria de modelos.

Para aplicar a metodologia proposta, os modelos de má qualidade utilizados devem ser estáveis. Além disso o problema deve estar condicionado de forma que a resposta real do sistema deve ser limitada na região que compreende as respostas dos submodelos utilizados. Esta condição pode ser testada por meio de (10):

$$y_{1_{min}} \geq y_{max} \geq y_{min} \geq y_{2_{max}}, \quad (10)$$

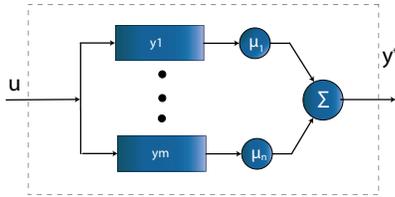


Fig. 1. CNDM sendo  $y^*$  o modelo combinado,  $y_i$  para  $i = 1, \dots, m$  os submodelos utilizados e  $\mu_i$  para  $i = 1, \dots, m$  seus respectivos pesos.  $u$  é a entrada aplicada.

sendo  $y_{1max}$  o máximo valor de  $y_1$ ,  $y_{2min}$  o mínimo valor de  $y_2$ .  $y_{max}$  e  $y_{min}$  são os valores máximo e mínimo de  $y$ .

Não satisfeita a condição (10), o problema pode ser condicionado calculando as grandezas (11) e (12) e utilizando no restante do procedimento os modelos transformados  $y_1^*$  (13) e  $y_2^*$  (14).

$$c_1 = y_{max} + y_{1min} \quad (11)$$

$$c_2 = y_{min} + y_{2max} \quad (12)$$

$$y_1^*(k) = y_1(k) + c_1 \quad (13)$$

$$y_2^*(k) = y_2(k) - c_2 \quad (14)$$

Ao somar os valores das constantes (11) e (12) às saídas dos modelos, estas são deslocadas para cima e para baixo no eixo das abscissas, garantindo a condição (10). Caso sejam utilizados mais de dois modelos, as grandezas (11) e (12) podem ser calculadas para os  $m$  submodelos dois a dois e seus respectivos modelos compostos devem ser utilizados no procedimento.

As subseções seguintes apresentam uma abordagem analítica e uma numérica para a definição dos pesos  $\mu_1$  a  $\mu_n$  que melhor ajuste a CNDM aos dados.

#### A. Método Analítico

Esta subseção apresenta uma solução analítica para o cômputo dos pesos da rede, sendo esta abordagem restrita à utilização de 2 submodelos. A saída da CNDM (Figura 1)  $y^*$  é então uma soma ponderada da saída dos modelos  $y_1$  e  $y_2$  pelos pesos  $\mu_1$  e  $\mu_2$  como em (15).

$$y^* = \mu_1 y_1 + \mu_2 y_2. \quad (15)$$

Assumindo um problema convexo,  $\mu_2 = 1 - \mu_1$ , a saída  $y^*$  pode ser representada como em (16).

$$y^* = \mu_1 y_1 + (1 - \mu_1) y_2. \quad (16)$$

Deseja-se que a saída da CNDM seja a mais próxima possível da resposta do sistema, ou  $y^* \approx y$ . Desta forma, abandonando o subscrito 1 em  $\mu_1$ :

$$y \approx \mu y_1 + (1 - \mu) y_2. \quad (17)$$

Assumindo que para uma entrada limitada o modelo produza uma saída limitada, é possível calcular as áreas de  $y$ ,  $y_1$  e  $y_2$ . Integrando os dois lados da Equação (17):

$$\int_0^n y dt \approx \int_0^n \mu y_1 + (1 - \mu) y_2 dt. \quad (18)$$

Reagrupando os termos em (18) e isolando o peso  $\mu$  pode-se então encontrar uma estimativa do valor  $\mu$  que faça com que  $y^* \approx y$ .

$$\mu \approx \frac{\int_0^n y(t) dt - \int_0^n y_2(t) dt}{\int_0^n y_1(t) dt - \int_0^n y_2(t) dt}. \quad (19)$$

Ou no domínio do tempo discreto, aproximando a integral por um somatório:

$$\mu \approx \frac{\sum_{k=0}^n y_k - \sum_{k=0}^n y_{2k}}{\sum_{k=0}^n y_{1k} - \sum_{k=0}^n y_{2k}}. \quad (20)$$

Um resumo do procedimento pode ser visto no Algoritmo 1.

---



---

#### 1 Algoritmo 1: Rede Combinatória de Modelos

2  $y_{max} \leftarrow \max(y_k)$

3  $y_{min} \leftarrow \min(y_k)$

4  $y_{1min} \leftarrow \min(y_{1k})$

5  $y_{2max} \leftarrow \max(y_{2k})$

6 **if**  $!(y_{1min} \geq y_{max} \geq y_{min} \geq y_{2max})$  **then**

/\* Se o problema não estiver bem condicionado \*/

7  $c_1 \leftarrow y_{max} + y_{1min}$

8  $c_2 \leftarrow y_{min} + y_{2max}$

9  $y_{1k}^* \leftarrow y_{1k} + c_1$

10  $y_{2k}^* \leftarrow y_{2k} - c_2$

11  $y_{1k} \leftarrow y_{1k}^*$

12  $y_{2k} \leftarrow y_{2k}^*$  // prosseguir com modelos transformados

13  $\mu \leftarrow \frac{\sum_{k=0}^n y_k - \sum_{k=0}^n y_{2k}}{\sum_{k=0}^n y_{1k} - \sum_{k=0}^n y_{2k}}$

---

#### B. Método Numérico

O método numérico computa os pesos da rede para um número arbitrário de modelos. Para tal é implementado um adaptador linear com o auxílio do algoritmo gradiente descendente, como pode ser visto na Figura 2. O gradiente descendente foi utilizado por ser capaz de lidar com a função custo minimizada e por não envolver o cálculo de matrizes pseudo inversas, cálculo requerido na utilização do método dos mínimos quadrados.

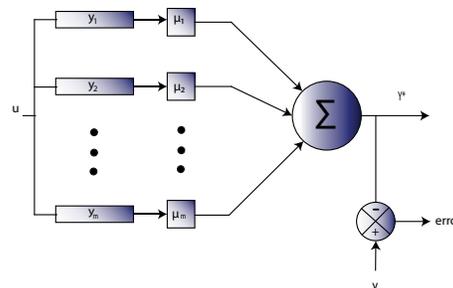


Fig. 2. Adaptador Polinomial. Em que  $y_1 \dots y_n$  são os submodelos utilizados na implementação,  $\mu_1 \dots \mu_n$  seus respectivos pesos.  $y^*$  é o modelo composto  $y$  é a resposta real do sistema e  $\xi$  é o sinal de erro definido por  $\xi = y^* - y$ .

Adotando notação matricial, a saída do adaptador linear é calculada como (21):

$$y^* = \bar{\Gamma}\bar{\mu}, \quad (21)$$

em que  $\bar{\Gamma}$  é uma matriz cujas colunas são as séries temporais correspondentes aos modelos  $y_1 \dots y_m$ , e  $\bar{\mu}$  é um vetor cujas linhas representam os pesos de cada um dos  $m$  submodelo.

O erro quadrático médio pode então ser calculado como em (22) e seu gradiente com relação aos pesos como em (23).

$$EQM = (\bar{\Gamma}\bar{\mu} - y)^T (\bar{\Gamma}\bar{\mu} - y) / n, \quad (22)$$

$$\Delta\bar{\mu} = \alpha(y - \bar{\Gamma}\bar{\mu})^T \bar{\Gamma}, \quad (23)$$

em que  $EQM$  é o erro quadrático médio,  $n$  o tamanho da série temporal e  $\alpha$  é o passo de aprendizagem para o algoritmo do gradiente descendente.

As Equações (21), (22) e (23) são computadas iterativamente até se atingir uma tolerância pré estabelecida para o erro quadrático médio. Um resumo do procedimento pode ser visto no Algoritmo 2.

---

### 1 Algoritmo 2: Rede Combinatória de Modelos

```

2  $\bar{y} \leftarrow [y_1, y_2, \dots, y_m]$  // Matriz contendo
   submodelos
3  $\bar{\mu} \leftarrow rand(1, m)$  // Inicia vetor de pesos
   aleatoriamente
4  $tol \leftarrow tolerancia$ 
5  $\alpha \leftarrow Passo\ de\ aprendizagem$ 
6  $y^* = \bar{y}\bar{\mu}$ 
7  $EQM = (\bar{y}\bar{\mu} - y)^T (\bar{y}\bar{\mu} - y) / n$ 
8 while  $tol \geq EQM$  do
9   |
10  |  $y^* = \bar{y}\bar{\mu}$ 
11  |  $EQM = (\bar{y}\bar{\mu} - y)^T (\bar{y}\bar{\mu} - y) / n$ 
12  |  $\Delta\bar{\mu} = \alpha(y - \bar{y}\bar{\mu})^T \bar{y}$ 
13  |  $\bar{\mu} = \bar{\mu} + \Delta\bar{\mu}$ 

```

---

### C. Estudos de Caso

Foram utilizados dois sistemas de teste. O primeiro se trata de um sistema puramente simulado, sendo um modelo NARX polinomial fornecido em [12] e representado por (24).

$$y(k) = 0,5y(k-1) + 0,8u(k-2) + u(k-1)^2 - 0,05y(k-2)^2 + 0,5. \quad (24)$$

O sinal de entrada foi um ruído branco Gaussiano com  $\bar{u} = 0$  e  $\sigma^2 = 1$ . Foram utilizadas 630 amostras, sendo 500 para identificação e 130 para a validação.

O sistema (24) foi identificado em [3] utilizando o MERR, sendo minimizados os objetivos estático e dinâmico. Os modelos utilizados são vistos em (25) e (26).

$$y_1(k) = 0,5821y_1(k-1) + 0,8184u(k-2) + 0,9459u(k-1)^2 - 0,3209y_1(k-2) - 0,0292y_1(k-2)y_1(k-1) + 0,8348, \quad (25)$$

$$y_2(k) = 0,7019y_2(k-1) + 0,8018u(k-2) - 0,0695y_2(k-2)y_2(k-1) + 0,9786u(k-1)^2 + 0,0042u(k-1)y_2(k-1) - 0,0812u(k-1)y_2(k-2). \quad (26)$$

Estes modelos se enquadram na definição de modelos de má qualidade. Deste modo, melhorias devem ser realizadas para que os modelos possam ser utilizados na representação do sistema.

O segundo exemplo estudado é um sistema eletromecânico composto por um motor e um gerador identificado em [4]. A estrutura foi selecionada pela taxa de redução de erro em conjunto com o critério de informação de Akaike. Para o ajuste dos parâmetros foi utilizado os mínimos quadrados intervalar. Neste trabalho foram utilizados o modelo do limite inferior e o modelo do limite superior representados pelos parâmetros  $\underline{\theta}$  e  $\bar{\theta}$  como os modelos  $y_1$  e  $y_2$ . A título de comparação, o modelo do ponto médio do intervalo ( $y_{MMQi}$ ) foi utilizado. A estrutura do modelo utilizado é vista em (27), e seus parâmetros em (28). Foram utilizadas 8320 amostras, 50 % para identificação e 50% para validação.

$$\Psi = \begin{bmatrix} y(k-1) & y(k-2) & u(k-1) & u(k-1)y(k-1) \\ u(k-2)u(k-1)y(k-2) & u(k-2)y(k-1) \\ u(k-2)y(k-2) & y(k-2)^2 \end{bmatrix}, \quad (27)$$

$$\hat{\theta} = \begin{bmatrix} [ 1, 86983289867678; & 1, 87052216754597] \\ [-0, 87500271106274; & -0, 87433604737629] \\ [ 0, 00877520297552; & 0, 07522496598523] \\ [-0, 07754127207970; & -0, 07522496598523] \\ [ 0, 01156721443224; & 0, 01159157820069] \\ [ 0, 06736048557893; & 0, 06969424138646] \\ [-0, 09177956625574; & -0, 08882860964415] \\ [ 0, 07814928825610; & 0, 08111662045620] \\ [ 0, 00273140158559; & 0, 00275676329139] \end{bmatrix}. \quad (28)$$

## IV. RESULTADOS

Os resultados são divididos de acordo com a abordagem utilizada, e subdivididos de acordo com o estudo de caso.

### A. Abordagem Analítica

1) *Exemplo 1 - Identificação Multiobjetivo:* O procedimento foi implementado utilizando o sistema (24) e os modelos (25) e (26). Um conjunto de 50 amostras da simulação livre da CNDM, e dos submodelos utilizados são vistos na parte de cima da Figura 3. Na parte inferior da Figura 3 são vistas as curvas estáticas para os mesmos modelos. Os pesos

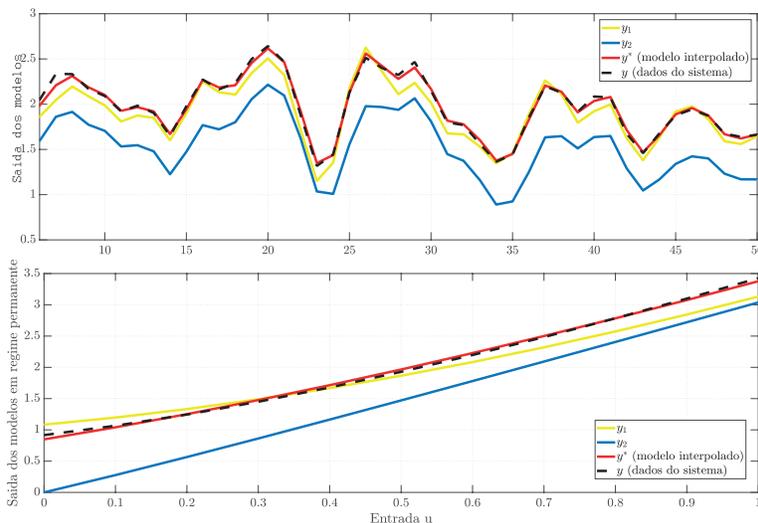


Fig. 3. Simulação livre dos modelos do Exemplo 1 para o método Analítico. Acima: Os 50 primeiros pontos de simulação dos modelos  $y_1$  e  $y_2$  juntamente com a rede  $y^*$  e os dados do sistema  $y$ ; Abaixo: Curvas estáticas para os modelos  $y_1$ ,  $y_2$ ,  $y^*$  e para o sistema  $y$ .

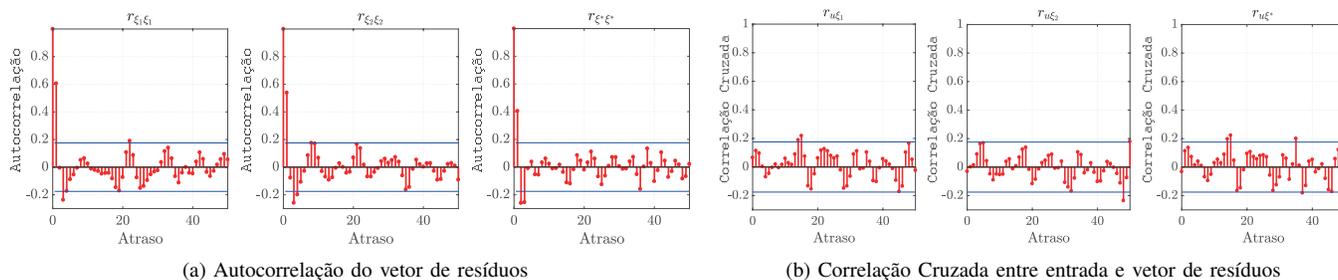


Fig. 4. Funções de Autocorrelação do vetor de resíduos e Correlação cruzada entre vetor de resíduos e entrada para o Exemplo 1 na aplicação do método analítico.

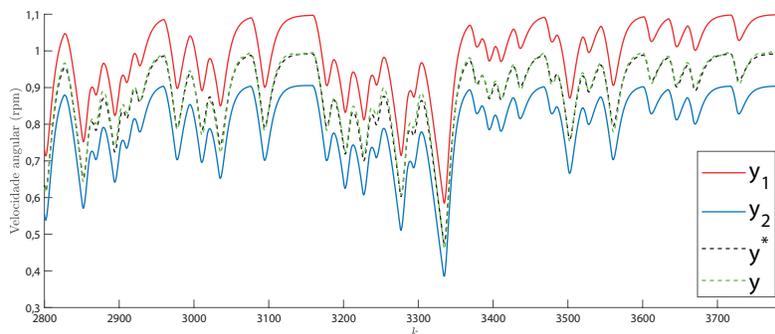


Fig. 5. Simulação livre dos modelos  $y_1$  e  $y_2$  juntamente com a rede  $y^*$  e os dados do sistema  $y$  para o Exemplo 2 utilizando o método analítico.

encontrados pela abordagem analítica foram  $\mu_1 = 0,5076$  e  $\mu_2 = 0,4924$ .

O desempenho dinâmico da CNDM foi avaliado a partir do índice RMSE e por meio da análise dos resíduos, comparando os índices computados com aqueles obtidos individualmente para os modelos utilizados na rede. Calculou-se a autocorrelação do vetor de resíduos, vista na Figura 4a. É importante notar que a função de autocorrelação dos resíduos  $r_{\xi^* \xi^*}$  não excede os limites de confiança de 95%,

indicando que os resíduos possuem características de um sinal aleatório. A correlação cruzada entre o vetor de resíduos e o vetor de entrada pode ser vista na Figura 4b. Nota-se que a função de correlação cruzada  $r_{u \xi^*}$  também não excede o limite de confiança, indicando que os resíduos não dependem da entrada aplicada. Para quantificar esta análise e comparar os resultados da CNDM com o dos submodelos utilizados, foram calculadas as energias do vetor de resíduos e das funções de autocorrelação e correlação cruzada. Para avaliar

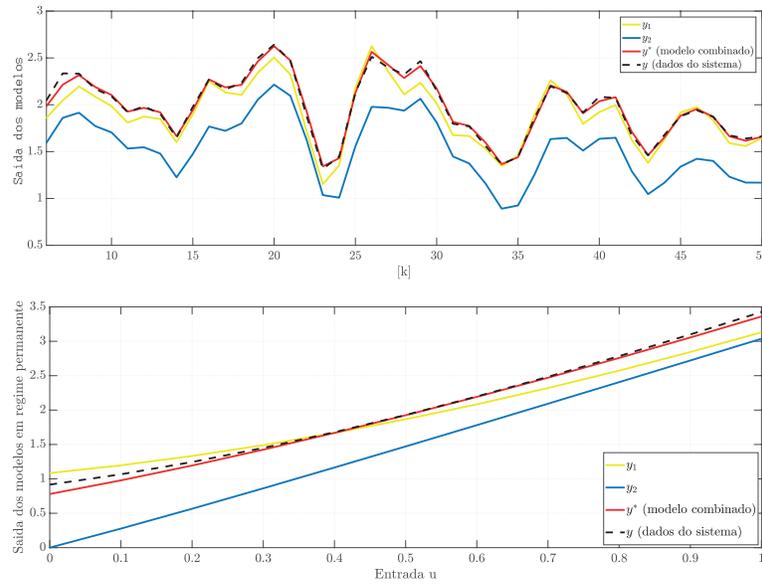


Fig. 6. Simulação livre dos modelos do Exemplo 1 para o método numérico. Acima: Os 50 primeiros pontos de simulação dos modelos  $y_1$  e  $y_2$  juntamente com CNDM  $y^*$  e os dados do sistema  $y$ ; Abaixo: Curvas estáticas para os modelos  $y_1$ ,  $y_2$ ,  $y^*$  e para o sistema  $y$ .

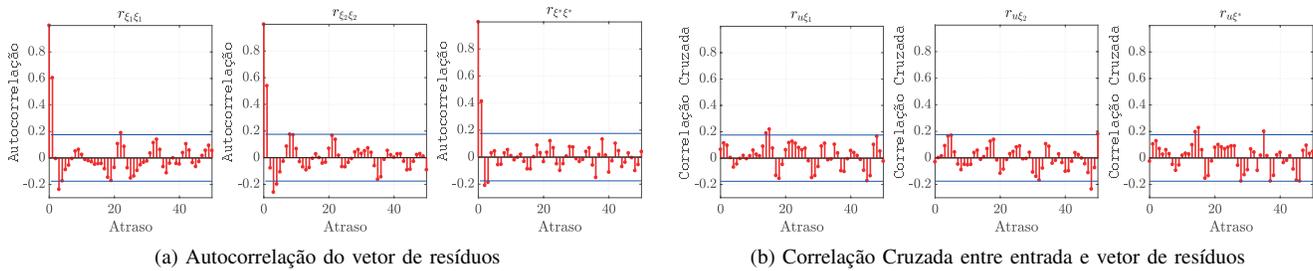


Fig. 7. Funções de Autocorrelação do vetor de resíduos e Correlação cruzada entre vetor de resíduos e entrada para o Exemplo 1 na aplicação do método numérico.

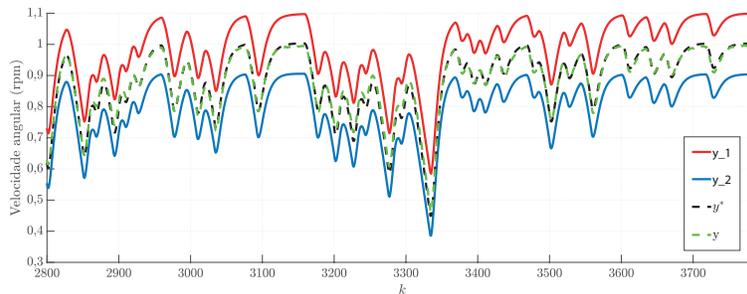


Fig. 8. Simulação livre dos modelos  $y_1$  e  $y_2$  juntamente com a rede  $y^*$  e os dados do sistema  $y$  para o Exemplo 2 utilizando o método numérico.

o comportamento estático da rede calculou-se o índice MSE da curva estática para os modelos. A Tabela I apresenta um comparativo entre o desempenho dos modelos.

A Tabela I mostra uma diminuição de cerca de 65% no índice RMSE e uma redução de 95% no índice estático MSE se comparado com os submodelos utilizados em sua obtenção. Ao analisar apenas o comportamento dinâmico (Figura 3), a saída dos modelos podem aparentar semelhança. Entretanto, é importante frisar que os modelos foram originalmente obtidos por meio de um procedimento multiobjetivo de identificação

de sistemas, minimizando-se os custos quadráticos estático e dinâmico, e que o modelo resultante obteve melhor representatividade do sistema em ambas características, destacando-se uma expressiva melhora no comportamento estático. A CNDM obteve também uma redução de 87% na energia do vetor de resíduos  $E_\xi$  e uma redução de 21% na energia da autocorrelação do vetor de resíduos  $E_{r_{\xi\xi}}$ .

É possível notar também que, mesmo os submodelos utilizados possuindo estruturas diferentes daquela pertencente ao sistema a ser identificado, a rede obteve uma boa representa-

TABELA I

COMPARAÇÃO DE DESEMPENHO DOS MODELOS - EXEMPLO 1, ABORDAGEM ANALÍTICA.  $E_{\xi}$  É A ENERGIA DO VETOR DE RESÍDUOS,  $E_{r_{\xi\xi}}$  É A ENERGIA DA AUTOCORRELAÇÃO DO VETOR DE RESÍDUOS E  $E_{r_{\xi u}}$  É A CORRELAÇÃO CRUZADA ENTRE O VETOR DE RESÍDUOS E A ENTRADA

Modelo	RMSE	$E_{\xi}$	$E_{r_{\xi\xi}}$	$E_{r_{\xi u}}$	MSE
$y_1$	0,3071	1,8008	2,0916	1,4353	0,0268
$y_2$	1,2949	26,829	2,1868	1,6116	0,3188
$y^*$	0,1086	0,2361	1,6652	1,4701	0,0012

tividade do sistema, alcançando uma melhora significativa nos índices de validação analisados.

2) *Exemplo 2 - Identificação de Sistemas Intervalar:* A metodologia analítica também foi aplicada utilizando os modelos em (27) e (28). O modelo do extremo superior foi escolhido como  $y_1$  e o modelo do extremo inferior como  $y_2$ . A resposta é comparada com o modelo do ponto médio do intervalo de parâmetros ( $y_{MMQi}$ ). Uma amostra de 990 pontos da resposta da CNDM  $y^*$  em comparação com as respostas dos modelos  $y_1$  e  $y_2$  pode ser vista na Figura 5.

Foi realizada uma análise a partir do índice RMSE, comparando com os submodelos utilizados para composição da CNDM e com o modelo do ponto médio ( $y_{MMQi}$ ) originalmente utilizado como modelo referência. A comparação foi realizada com esse modelo por se tratar da escolha mais óbvia na necessidade de selecionar um modelo dentre o intervalo identificado. Os resultados são apresentados na Tabela II. A CNDM  $y^*$  apresentou uma redução de 79% do índice RMSE se comparado com os modelos  $y_1$  e  $y_2$  utilizados para obtê-lo, e uma redução de 11 % se comparado com o modelo do ponto médio do intervalo de parâmetros. Neste contexto, devido a natureza do procedimento de identificação utilizado, é natural que a respostas dos modelos se assemelhem. Porém, é importante notar que a rede combinatoria obteve um melhor desempenho do que a escolha óbvia no intervalo de parâmetros.

TABELA II

COMPARAÇÃO DE DESEMPENHO DOS MODELOS - ABORDAGEM ANALÍTICA, IDENTIFICAÇÃO DE SISTEMAS INTERVALAR.  $y_{MMQi}$  É O MODELO DO PONTO MÉDIO DO INTERVALO DE PARÂMETROS

Modelo	RMSE	$E_{r_{\xi\xi}}$	$E_{r_{\xi u}}$
$y_1$	0,6930	17,56	1,69
$y_2$	0,5766	20,95	1,65
$y_{MMQi}$	0,1336	13,67	1,73
$y^*$	0,1190	13,62	1,77

**B. Abordagem Numérica**

Em todos os casos da abordagem numérica foi utilizado um passo de aprendizagem  $\alpha = 0.0001$  e uma tolerância  $tol = 0.00001$ ,

1) *Exemplo 1 - Identificação Multiobjetivo:* A metodologia numérica foi aplicada aos dados do sistema (24) e modelos (25) e (26). Na parte de cima da Figura 6 pode ser observada a simulação livre da CNDM bem como dos modelos utilizados na combinação. Abaixo na mesma imagem (Figura 6) podem ser vistas as curvas estáticas para os mesmos modelos.

Neste contexto os pesos encontrados foram  $\mu_1 = 0,510$  e  $\mu_2 = 0,503$ .

Para avaliar a CNDM foi utilizado o índice RMSE e a análise de resíduos. A autocorrelação do vetor de resíduos  $r_{\xi\xi}$  pode ser vista na Figura 7a. A correlação cruzada entre o vetor de resíduos e o vetor de entrada  $r_{u\xi}$  pode ser visto na Figura 7b. É possível notar que, como na abordagem analítica as funções de correlação ficaram dentro dos limites de confiança de 95%, cabendo aqui as mesmas observações realizadas. A análise é quantificada por meio das energias do vetor de resíduos e das funções de autocorrelação e correlação cruzada. A Tabela III apresenta um comparativo entre o desempenho da rede e dos submodelos.

Por meio da Tabela III nota-se uma diminuição de cerca de 65% no índice RMSE e uma redução de 87% no índice estático MSE em relação aos submodelos da rede, indicando que a CNDM obteve um melhor desempenho em ambos objetivos minimizados pelo MERR. A CNDM obteve também uma redução de 87% na energia do vetor de resíduos e uma redução de 23% na energia da autocorrelação do vetor de resíduos.

TABELA III

COMPARAÇÃO DE DESEMPENHO DOS MODELOS - ABORDAGEM NUMÉRICA, IDENTIFICAÇÃO MULTI OBJETIVO.  $E_{\xi}$  É A ENERGIA DO VETOR DE RESÍDUOS,  $E_{r_{\xi\xi}}$  É A ENERGIA DA AUTOCORRELAÇÃO DO VETOR DE RESÍDUOS E  $E_{r_{\xi u}}$  É A CORRELAÇÃO CRUZADA ENTRE O VETOR DE RESÍDUOS E A ENTRADA

Modelo	RMSE	$E_{\xi}$	$E_{r_{\xi\xi}}$	$E_{r_{\xi u}}$	MSE
$y_1$	0,3071	1,8008	2,0916	1,4353	0,0268
$y_2$	1,2949	26,829	2,1868	1,6116	0,3188
$y^*$	0,1070	0,2292	1,5959	1,4509	0,0034

2) *Exemplo 2 - Identificação de Sistemas Intervalar:* A resposta da CNDM  $y^*$  em comparação com as respostas dos modelos  $y_1$ ,  $y_2$  e  $y_{MMQi}$  pode ser vista na Figura 8.

TABELA IV

COMPARAÇÃO DE DESEMPENHO DOS MODELOS - ABORDAGEM NUMÉRICA, IDENTIFICAÇÃO DE SISTEMAS INTERVALAR.  $y_{MMQi}$  É O MODELO DO PONTO MÉDIO DO INTERVALO DE PARÂMETROS

Modelo	RMSE	$E_{r_{\xi\xi}}$	$E_{r_{\xi u}}$
$y_1$	0,6930	17,56	1,69
$y_2$	0,5766	20,95	1,65
$y_{MMQi}$	0,1336	13,67	1,73
$y^*$	0,1080	13,60	1,65

Novamente a avaliação da CNDM foi realizada pelo índice RMSE e pela análise de resíduos. A comparação dos resultados é vista na Tabela IV. A CNDM  $y^*$  apresentou uma redução de 83% do índice RMSE se comparado com os modelos  $y_1$  e  $y_2$  utilizados para obtê-la, e uma redução de 28% se comparado com o modelo do ponto médio de parâmetros  $y_{MMQi}$ , indicando que o desempenho da rede foi superior ao do modelo que representa a escolha óbvia dentro do intervalo de parâmetros.

Ambos os métodos se mostraram eficazes em obter modelos com boa representatividade do sistema. As duas abordagens resultaram em modelos semelhantes, sendo a maior diferença apresentada pelo método analítico que obteve uma maior redução no índice estático MSE para o Exemplo 1.

## V. CONCLUSÃO

Este trabalho apresentou a Rede Combinatória de Modelos Dinâmicos CNDM, capaz de combinar modelos paralelamente, de forma a obter representações com melhor desempenho. Abordagens correlatas são vistas na literatura como o conjunto estendido de modelos que combina múltiplas funções de base em um único modelo NARMAX. Contudo, tal técnica não é capaz de combinar modelos polinomiais previamente identificados como no procedimento proposto por este trabalho. Outra técnica relacionada é a computação de reservatório ou *reservoir computing*, que combina uma grande quantidade de estados internos do chamado reservatório para o cômputo da resposta final. Tal abordagem se diferencia da metodologia desenvolvida principalmente pela quantidade de estados internos do modelo, uma vez que o reservatório possui uma alta dimensão. Outra diferença se dá pela estrutura do reservatório que não passa por uma seleção e muitas vezes é aleatória enquanto na metodologia introduzida os estados internos do modelo são selecionados como modelos previamente identificados.

Foram realizados dois estudos de caso sendo o primeiro no contexto da Identificação Multiobjetivo. Neste sentido, foram combinados modelos que minimizam os objetivos estático e dinâmico. A rede foi capaz de obter um melhor desempenho em ambos os objetivos, apresentando redução expressiva nos índices analisados.

No segundo estudo de caso, modelos obtidos a partir do mínimos quadrados intervalares foram combinados. A CNDM obteve melhor representatividade do que os submodelos utilizados e do que o modelo que representa a escolha óbvia ao selecionar apenas um modelo do conjunto.

Como trabalho futuro pretende-se expandir a abordagem para casos onde os pesos da rede sejam funções. Pretende-se também estudar a composição de novas funções custo para o problema e investigar diferentes técnicas para minimiza-las.

## AGRADECIMENTOS

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001. Os autores também agradecem à FAPEMIG, CNPq, e à Universidade Federal de São João del Rei pelo apoio.

## REFERÊNCIAS

- [1] S. Chen and S. Billings, "Representations of non-linear systems: The NARMAX model," *International Journal of Control*, vol. 49, no. 3, pp. 1013–1032, 1989.
- [2] L. A. Aguirre, G. G. Rodrigues, and J. C. R. F., "Identificação de sistemas não-lineares utilizando modelos NARMAX polinomiais - uma revisão e novos resultados," *SBA Controle & Automação*, vol. 9, no. 2, pp. 90–106, 1998.
- [3] S. A. M. Martins, E. G. Nepomuceno, and M. F. S. Barroso, "Improved structure detection for polynomial NARX models using a multiobjective error reduction ratio," *Journal of Control, Automation and Electrical Systems*, vol. 24, no. 6, pp. 764–772, 2013.
- [4] M. L. C. Peixoto, M. T. R. Matos, W. R. L. Júnior, S. A. Martins, and E. G. Nepomuceno, "Identificação de Sistemas Dinâmicos com Aritmética Intervalar," *Simpósio Brasileiro de Automação Inteligente*, pp. 471–476, 2017.

- [5] S. Chen and S. A. Billings, "Extended model set, global data and threshold model identification of severely non-linear systems," *International Journal of Control*, vol. 50, no. 5, pp. 1897–1923, 1989.
- [6] M. Lukoševičius, H. Jaeger, and B. Schrauwen, "Reservoir computing trends," *KI-Künstliche Intelligenz*, vol. 26, no. 4, pp. 365–371, 2012.
- [7] L. Gonon and J.-P. Ortega, "Reservoir computing universality with stochastic inputs," *IEEE transactions on neural networks and learning systems*, 2019.
- [8] F. Hadaeghi and H. Jaeger, "Computing optimal discrete readout weights in reservoir computing is np-hard," *Neurocomputing*, vol. 338, pp. 233–236, 2019.
- [9] Y. Paquot, F. Duport, A. Smerieri, J. Dambre, B. Schrauwen, M. Haelterman, and S. Massar, "Optoelectronic reservoir computing," *Scientific reports*, vol. 2, p. 287, 2012.
- [10] I. J. Leontaritis and S. A. Billings, "Input-output parametric models for non-linear systems Part I: Deterministic non-linear systems," *International Journal of Control*, vol. 41, no. 2, pp. 303–328, 1985.
- [11] L. Aguirre, *Introdução à Identificação de Sistemas – Técnicas Lineares e Não-Lineares Aplicadas a Sistemas Reais*. Editora UFMG, 2015.
- [12] L. Piroddi and W. Spinelli, "An identification algorithm for polynomial NARX models based on simulation error minimization," *International Journal of Control*, vol. 76, no. 17, pp. 1767–1781, 2003.



**Luan Pascoal** Received BSc in Electrical Industrial Engineering from the Federal University of São João Del-Rei (2018). Currently is a M.Sc. degree student in Electrical Engineering from the Federal University of São João del-Rei. His research interests are in nonlinear system identification, machine learning and data science.



**Samir Angelo** received the BSc and MSc of engineering degrees from Federal University of São João del-Rei and the DSc in Electrical Engineering from the Federal University of Minas Gerais. He is an adjunct professor since may 2013 at Electrical Engineering Department from Federal University of São João del-Rei. His recent research interests include nonlinear system identification, circuit and systems control, hysteresis modelling and control.



**Gleison Fransoares** Received BSc in Electrical Industrial Engineering from the Federal University of São João Del-Rei (2000), a MSc degree in Electrical Engineering from the Federal University of Minas Gerais (2001) and DSc in Electrical Engineering from the Federal University of Minas Gerais (2006). Currently is an associate professor at the Federal University of São João Del-Rei. He has experience in Mathematics, with emphasis on nonlinear dynamical Systems.