

Trabalho de Conclusão de Curso

Identificação de sistemas usando NARX com termos transcendentes

de Jonh Lucas Alves da Silva

Orientado por Prof. Dr. Ícaro Bezerra Queiroz de Araújo

> Universidade Federal de Alagoas Instituto de computação Maceió, Alagoas 2024

UNIVERSIDADE FEDERAL DE ALAGOAS INSTITUTO DE COMPUTAÇÃO

IDENTIFICAÇÃO DE SISTEMAS USANDO NARX COM TERMOS TRANSCENDENTES

Trabalho de Conclusão de Curso submetido ao Instituto de Computação da Universidade Federal de Alagoas como requisito parcial para a obtenção do grau de Engenheiro de Computação.

JONH LUCAS ALVES DA SILVA

Orientador: Prof. Dr. Ícaro Bezerra Queiroz de Araújo

Banca Avaliadora:

João Raphael Souza Martins Prof. Dr., IC-UFAL Jobson de Araújo Nascimento Prof. Me., IC-UFAL

Maceió, Alagoas Novembro de 2024

Catalogação na Fonte Universidade Federal de Alagoas Biblioteca Central

Divisão de Tratamento Técnico

Bibliotecário: Cláudio Albuquerque Reis - CRB-4 - 1753

S586i Silva, Jonh Lucas Alves da.

Identificação de sistemas usando NARX com termos transcedentes / John Lucas Alves da Silva. $-\,2024.$

48 f.: il.

Orientadora: Ícaro Bezerra Queiroz de Araújo.

Monografia (Trabalho de Conclusão de Curso em Engenharia da Computação) — Universidade Federal de Alagoas. Instituto de Computação. Maceió, 2024.

Bibliografia. f. 47-48.

1. Sistemas não-lineares. 2. Engenharia da computação. I. Título.

CDU: 004.436.2

JONH LUCAS ALVES DA SILVA

Identificação de sistemas usando NARX com termos transcendentes

Trabalho de Conclusão de Curso submetido ao Instituto de Computação da Universidade Federal de Alagoas como requisito parcial para a obtenção do grau de Engenheiro de Computação.

Banca Examinadora:

Orientador: Prof. Dr. Ícaro Bezerra Queiroz de Araújo Instituto de Computação Universidade Federal de Alagoas

Examinador: Prof. Me. Jobson de Araújo Nascimento Instituto de Computação Universidade Federal de Alagoas

Examinador: Prof. Dr. João Raphael Souza Martins Instituto de Computação Universidade Federal de Alagoas

AGRADECIMENTOS

Primeiramente, gostaria de agradecer aos meus pais, Givanildo e Simone, por todo apoio e incentivo que me deram. Vocês são minha maior fonte de motivação e orientação. Estendo os agradecimentos às minhas irmãs, que apesar da implicância, foram grandes apoiadoras ao longo de toda jornada, e aos demais familiares pelo carinho e suporte prestado, em especial durante a realização deste trabalho.

Ao meu orientador, Prof. Ícaro, agradeço por todo o empenho e disponibilidade. Sua participação não começou a partir da escrita, mas, sim, na idealização, a oferta da disciplina de identificação de sistemas foi a principal responsável por despertar o interesse pelo tema.

Sou grato a todos os amigos que foram companhias indispensáveis durante toda a vida acadêmica. Aos demais professores e a universidade, obrigado por todo conhecimento adquirido e pela formação profissional. Em resumo, meu sincero agradecimento a todos que contribuíram para a realização deste trabalho.

RESUMO

Os modelos lineares foram amplamente descritos na literatura, possuindo métodos de análise e controle consolidados. Esses modelos são caracterizados pela simplicidade, eficiência, estabilidade e robustez, contudo essa classe não é capaz representar com fidelidade sistemas mais complexos. Além disso, o desenvolvimento de novas representações matemáticas, o avanço do poder computacional e o aumento da demanda por estimações mais precisas motivam o estudo de modelos não lineares. Este trabalho aborda a identificação offline de sistemas por meio de uma abordagem caixa cinza. Onde o modelo NARX polinomial foi alterado para que novos regressores gerados a partir das entradas e saídas fossem adicionados. Sua representação permaneceu linear nos parâmetros, então o método de mínimos quadrados foi utilizado para estimação. Enquanto, seleção de estrutura foi realizada a partir dos algoritmos SEMP e FRP. Os casos de estudo abordaram as medições sobre um sistema de tanques conectados, um benchmark de motores acoplados à pólia e um exemplo numérico. Por fim, os modelos obtidos foram simulados, validados e comparados com sua contraparte polinomial, a fim de verificar o ganho vindo da aplicação de funções não lineares para criação de novos regressores.

Palavras-chave: sistemas não lineares, NARX, MMQ, seleção de estrutura.

ABSTRACT

Linear models have been widely described in the literature and have consolidated analysis and control methods. These models are characterized by their simplicity, efficiency, stability and robustness, but this class is not capable of faithfully representing more complex systems. In addition, the development of new mathematical representations, advances in computing power and the increased demand for more accurate estimates have motivated the study of non-linear models. This work addresses the offline identification of systems using a gray box approach. Where the polynomial NARX model was altered so that new regressors generated from the inputs and outputs were added. Its representation remained linear in parameters, so the least squares method was used for estimation. Meanwhile, structure selection was carried out using the SEMP and FRP algorithms. The case studies covered measurements on a system of connected tanks, a benchmark of motors coupled to the pulley and a numerical example. Finally, the models obtained were simulated, validated and compared with their polynomial counterpart, in order to verify the gain from applying non-linear functions to create new regressors.

Keywords: non-linear systems, NARX, MMQ, structure selection.

LISTA DE ILUSTRAÇÕES

Figura 1 – Identificação de processos offline	13
Figura 2 — Etapas do procedimento de identificação	14
Figura 3 – (a) Modelo de Hammerstein, sua generalização multivariáveis, (b) re-	
presentação de Kortmann e Unbehauen,(c) NARX MISO e (d) NARX	
proposto	17
Figura 4 – Representação detalhada do (a) NARX SISO e (b) do modelo com as	
alterações propostas	18
Figura 5 – Esquema de tanques acoplados	25
Figura 6 – Etapas da identificação do modelo proposto	27
Figura 7 – Sistema de tanques	29
Figura 8 – Sinais coletados	30
Figura 9 – Diagrama esquemático	30
Figura 10 – Sinais de estimação CE8, sendo o primeiro a entrada e o segundo a saída	ı 31
Figura 11 – Diagrama esquemático Ball and Beam	32
Figura 12 – Sinais de estimação Ball and Beam	33
Figura 13 – Simulações do tanque superior para a (a) estimação e (b) validação	35
Figura 14 – Simulações do tanque inferior para a (a) estimação e (b) validação. $$	36
Figura 15 – Simulações livres sobre o conjunto de validação do tanque 1 $\ \ldots \ \ldots$	37
Figura 16 – Sinais de validação do tanque 2	38
Figura 17 – Simulação livre sobre o conjunto de estimação e validação	40
Figura 18 – Simulação livre durante a estimação e validação	40
Figura 19 – Simulação livre da velocidade da bola para a estimação e validação. $$. $$.	42
Figura 20 – Comparação entre as simulações do modelo proposto e polinomial tanto	
para o conjunto de estimação quanto para o de validação	44
Figura 21 – Comparação da posição estimada com o sinal original	45

LISTA DE TABELAS

Tabela 1 –	Crescimento do número de regressores para um sistema SISO com ape-	
	nas uma função não linear aplicada.	19
Tabela 2 –	Parâmetros numéricos para simulação	32
Tabela 3 –	Métricas de ambas as simulações referentes ao tanque superior	34
Tabela 4 –	Métricas de ambas as simulações referentes ao tanque inferior \dots	36
Tabela 5 –	Comparação entre o RMSE de ambos os modelos referentes ao tanque 1	37
Tabela 6 –	Comparação entre o MAPE de ambos os modelos referentes ao tanque 1	38
Tabela 7 –	Métricas de erro do tanque inferior durante a validação	38
Tabela 8 –	Métricas de ambas as simulações referentes aos motores acoplados. $$. $$.	36
Tabela 9 –	Comparação entre o RMSE de ambos os modelos	41
Tabela 10 –	Comparação entre o MAPE de ambos os modelos	41
Tabela 11 –	Comparação entre o EAT de ambos os modelos	41
Tabela 12 –	Métricas de ambas as simulações referentes aos motores acoplados	43
Tabela 13 –	Comparação entre o RMSE de ambos os modelos	43
Tabela 14 –	Comparação entre o MAPE de ambos os modelos	43

LISTA DE ABREVIATURAS E SIGLAS

ARX Autoregressive with Exogenous Inputs

ARMAX Autoregressive Moving Average with Exogenous Inputs

NARX Nonlinear Autoregressive with Exogenous Inputs

NARMAX Nonlinear Autoregressive Moving Average with Exogenous Inputs

SEMP Simulation Error Minimization with Pruning

SRR Simulation error Reduction Ratio

PRBS Pseudorandom binary sequence

LIT Linear e invariante no tempo

RMSE Root Mean Squared Error

MAPE Mean Absolute Percentage Error

SISO Single-input and single-output

MIMO Multiple-input and multiple-output

ERR Error reduction ratio

SRR Simulation Error Reduction Ratio

MSSE erro quadrático médio simulado

MMQ Método de mínimos quadrados

SUMÁRIO

1	INTRODUÇÃO
1.1	Justificativa
1.2	Objetivos
1.2.1	Objetivo geral
1.2.2	Objetivos específicos
2	FUNDAMENTAÇÃO TEÓRICA
2.1	Identificação de sistemas
2.2	Representação NAR(MA)X
2.2.1	Modelo NARX Racional
2.2.2	Modelo NARX Polinomial
2.3	NARX Proposto
2.4	Estimação de parâmetros
2.4.1	Estimador de Mínimos Quadrados
2.5	Seleção de estrutura
2.5.1	SRR
2.5.2	Simulation Error Minimization with Pruning - SEMP
2.6	Validação
2.6.1	Métricas de erro
2.7	Tanques acoplados
3	METODOLOGIA
3.1	Processo de identificação
3.2	Coleção de funções
3.3	Casos de estudo
3.3.1	Sistema de tanques acoplados
3.3.2	Motores acoplados a uma polia
3.3.3	Ball and Beam
4	RESULTADOS
4.1	Tanques acoplados
4.1.1	Tanque superior
4.1.2	Tanque inferior
4.1.3	Comparativo
4.2	Coupled Electric Drives
4.2.1	Modelo proposto

4.2.2	Comparativo	40
4.3	Exemplo numérico	42
4.3.1	Modelo proposto	42
4.3.2	Comparativo	43
5	CONCLUSÃO	46
5.1	Melhorias futuras	46
	REFERÊNCIAS	47

1 INTRODUÇÃO

Em diversas ocasiões replicar o comportamento de uma planta ou processo poder ser necessário, independente de sua motivação, descobrir uma relação de causa e efeito entre variáveis independentes pode sem útil. Vários fenômenos da natureza, processos industriais ou eventos biológicos podem ser expressos por um conjunto de equações matemáticas, que pode ser útil para tarefas de controle, detecção de defeitos ou otimização (COELHO; COELHO, 2015).

Reconhecida sua importância, existem diferentes aspectos a serem considerados sobre a modelagem. Um deles é opacidade, a seguir duas classificações usuais:

- Modelagem caixa branca: também conhecida como modelagem física, baseia-se nas leis e princípios da física que regem o processo. A representação matemática é legível e sua estrutura é conhecida previamente.
- Modelagem caixa preta: é uma abordagem experimental, não requer conhecimento prévio do sistema. Baseia-se apenas nas medições para identificação, ou seja, reflete apenas as propriedades presentes nos dados de estimação. Sua estrutura é opaca, a relação entre as entradas e saídas nem sempre é interpretável.

A modelagem caixa cinza é uma técnica intermediária entre as duas que busca combinar as vantagens de ambas. O conhecimento prévio é incorporado na obtenção de modelos caixa preta, que geralmente são mais fáceis de desenvolver, a fim de melhorar os resultados (AGUIRRE, 2015; COELHO; COELHO, 2015; SCHOUKENS; LJUNG, 2019).

Os modelos lineares foram amplamente documentados pela teoria de sistemas e controle clássica. No entanto, segundo Aguirre todo sistema real em certa medida possui dinâmicas não lineares. As representações não lineares são capazes de descrever comportamentos mais complexos, como bifurcação, ciclos limite, zona morta, histerese, etc. Algumas delas são a série de Volterra, modelos de Hammerstein-Weinner, representações NAR(MA)X e redes neurais artificiais (AGUIRRE, 2015).

Trabalhos como CHEN e BILLINGS (1989), KORENBERG S. A. BILLINGS e McILROY (1988) e AGUIRRE, RODRIGUES e JÁCOME (1998) balizaram o conhecimento sobre os modelos NAR(MA)X. Essa família de modelos é muito popular e vem sendo aplicada com sucesso na indústria (SCHOUKENS; LJUNG, 2019). Ela é caracterizada pela flexibilidade, eficiência, fácil interpretação e pela simulação simples. Mais recentemente, outros trabalhos também abordaram o tema, como Kadochnikova et al. (2023), Schoukens e Tiels (2017) e Deng et al. (2015). Segundo CHEN e BILLINGS (1989) as estruturas mais comuns são a racional e a polinomial, porém estudos sobre outras formas de representação vem sendo feitos como o Mozaryn (2023) que utiliza um modelo NARX

com a estrutura de uma rede neural recorrente. Portanto, vários conceitos continuam em aberto e podem ser explorados.

1.1 Justificativa

Como visto anteriormente, os modelos NAR(MA)X polinomiais possuem diversas qualidades. A linearidade nos parâmetros facilita sua estimação, porém torna-o fraco não linearmente. Dificultando identificação de dinâmicas não-lineares altamente complexas (AGUIRRE; RODRIGUES; JÁCOME, 1998). Este trabalho tem como proposta adicionar funções não lineares ao conjunto dos regressores candidatos, buscando amenizar esse problema.

Uma das etapas para identificação de sistemas mencionado por COELHO e COELHO (2015) é composta pela realização dos ensaios e pela coleta dados. No entanto, nem sempre a execução das duas tarefas é possível, então detalhes importantes dos sinais estão fora do controle do executante. O tempo de amostragem, a duração do ensaio e a seleção das entradas não são definidos por quem estimará o sistema, restando-lo apenas a opção de manipular os dados disponíveis da melhor maneira possível. Dada essa condição, a adição de funções não lineares estáticas é uma alternativa para extração informação adicional dos sinais, o que é muito bem-vindo em situações com dados restritos. Além disso, outros trabalhos também debruçaram-se sobre o uso de blocos não lineares para estimação. Alguns deles são Deng et al. (2015), Li et al. (2024) e Schoukens e Tiels (2017), cada um com uma abordagem distinta. Este trabalho é apenas mais uma contribuição para o estudo tema.

A escolha da representação matemática é outra etapa importante, a qual deve conseguir descrever o sistema corretamente. O estudo aqui proposto buscará verificar a expansão da capacidade de representação do modelo. Vale ressaltar que a melhoria na identificação não é garantida, pois as adições não necessariamente serão relevantes e se selecionadas incorretamente resultariam no aparecimento de dinâmicas espúrias. Portanto, busca-se encontrar modelos com desempenho melhor ou igual à forma polinomial, logo a implementação foi feita de modo que é possível remover as não linearidades estáticas e retornar a forma clássica, alterando apenas alguns parâmetros.

1.2 Objetivos

1.2.1 Objetivo geral

Desenvolver uma rotina para estimação de modelos não lineares que sejam robustos e eficientes. Enquanto, mantém-se uma estrutura de transparente, possibilitando uma fácil interpretação, e uma forma simples de seleção e simulação. Com a adição de termos não algébricos, busca-se um ganho de desempenho.

1.2.2 Objetivos específicos

- Realizar a identificação offline de sistemas.
- Implementar algoritmos clássicos para seleção de estrutura.
- Utilizar método baseado no EMQ para estimar os parâmetros.
- Usar predições livres para validar os modelos.
- Verificar o ganho ao aplicar funções não lineares aos regressores candidatos.

2 FUNDAMENTAÇÃO TEÓRICA

2.1 Identificação de sistemas

A identificação offline consiste em excitar um sistema a partir de uma entrada, armazenar os sinais de entrada e saída e em outro momento processar os dados para estimação e avaliar o modelo, vide a Figura 1 (COELHO; COELHO, 2015). Realizar mais de um ensaio é uma boa prática, pois um conjunto de dados deve ser usado para identificação e os restantes para avaliação.

entrada saída processo disco computador modelo

Figura 1 – Identificação de processos offline

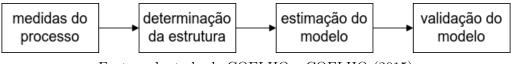
Fonte: adaptado de COELHO e COELHO (2015)

Segundo COELHO e COELHO (2015), o processo de identificação é composto pelas 4 etapas vistas na Figura 2. Apesar de seu foco estar sobre sistemas lineares, os conceitos apresentados por ele podem ser aproveitados aqui. Ao passo que, Aguirre (2015) enuncia 5 etapas, sendo 4 delas análogas as citadas anteriormente, são elas:

- coleta de dados. Idealmente, os sinais de excitação e o tempo de amostragem seriam escolhidos e, em seguida, a(s) entrada(s) e saída(s) armazenadas, contudo, geralmente, a única opção é utilizar os dados disponibilizados por outros ou em condições arbitrárias de operação. O uso do material de terceiros reduz consideravelmente controle sobre os dados, mas se as amostras tiverem conteúdo adequado, não comprometerão sua finalidade. Esta etapas é a análoga à medição do processo;
- escolha da representação matemática. Existem diversos critérios que norteiam a
 escolha, entre eles estão domínio, linearidade, memória e opacidade, todavia o importante é que a representação escolhida deve ser capaz de descrever a dinâmica do
 sistema. O foco deste trabalho está sobre modelos NARX/NARMAX;

- determinação de estrutura. Após definir a topologia do modelo, deve-se selecionar o melhor conjunto de elementos para representar o sistema. Em redes neurais, esta etapa corresponde a definição dos hiperparâmetros ou a escolha do grau, ordem e dos regressores candidatos quando trata-se de NAR(MA)X;
- estimação de parâmetros. Parte da escolha do algoritmo que geralmente resolve um problema de otimização. Quando a representação é linear nos parâmetros, o MMQ ou derivados são aplicados;
- validação. Averígua a qualidade do modelo, comparando os dados preditos com as observações. Métricas de erro podem ser adotadas para quantificar o desempenho modelo. No entanto, sua aprovação depende do cumprimento de critérios previamente definidos;

Figura 2 – Etapas do procedimento de identificação



Fonte: adaptado de COELHO e COELHO (2015)

A aquisição dos dados será descrita em 3, enquanto as demais etapas serão abordas nas seções seguintes.

2.2 Representação NAR(MA)X

A estrutura de um modelo NARX monovariável com período de amostragem normalizado é:

$$y(k) = F[y(k-1), ..., y(k-n_y), u(k-d), ..., u(k-n_u)] + e(k)$$
(2.1)

enquanto o modelo NARMAX é dado por

$$y(k) = F[y(k-1), ..., y(k-n_y), u(k-d), ..., u(k-n_u), e(k-1), ..., e(k-n_e)] + e(k)$$
(2.2)

em que:

- O instante $k \in [1, ..., N]$ sendo N o número de amostras
- F é uma função não linear
- y(k) é o sinal da saída no instante k.
- u(k) é o sinal da entrada no instante k.

- e(k) é o ruído aditivo do sistema no instante k.
- n_y , n_u e n_e os maiores atrasos referentes a y, u e e.
- \bullet d é o tempo morto ou atraso do sistema.

O desafio para estimação desses modelos reside em encontrar uma função F que represente o sistema corretamente. As aproximações mais comuns para função são representações polinomiais e racionais (AGUIRRE; RODRIGUES; JÁCOME, 1998; CHEN; BILLINGS, 1989). A equação 2.1 pode ser representada na forma em encurtada

$$y = F(y, u) + e$$

indo além, extrapolando a equação para o caso multivariável a forma $y_n = F_n(\mathbf{Y}, \mathbf{U}) + \mathbf{e}_n$ é obtida, onde $\mathbf{Y} = [\mathbf{y}_1, \dots, \mathbf{y}_n]$ e $\mathbf{U} = [\mathbf{u}_1, \dots, \mathbf{u}_m]$.

2.2.1 Modelo NARX Racional

Considerando que a função F é uma razão de funções, a e b, o sistema pode ser escrito da seguinte forma:

$$y(k) = \frac{a(y(k-1), ..., y(k-n_y), u(k-d), ..., u(k-n_u))}{b(1, y(k-1), ..., y(k-n_y), u(k-d), ..., u(k-n_u))} + e(k)$$
(2.3)

geralmente a e b são definidos como polinômios de grau finito (CHEN; BILLINGS, 1989). Logo, reescrevendo a equação anterior, temos:

$$y(k) = \frac{\sum_{i} a_{i} \prod_{j=1}^{n_{y}} y(k-j) \prod_{r=1}^{n_{u}} u(k-r)}{1 + \sum_{i} b_{i} \prod_{j=1}^{d_{y}} y(k-j) \prod_{r=1}^{d_{u}} u(k-r)} + e(k)$$
(2.4)

sendo a_i e b_i os contantes referentes aos parâmetros. Abaixo, um exemplo arbitrário

$$y(k) = \frac{a_0 + a_1 y(k-1) + a_2 u(k-1) + a_5 u(k-1)^2}{b_0 + b_1 y(k-1) + b_2 u(k-1)}$$
(2.5)

A partir da Equação anterior, algumas observações podem ser feitas: o modelo não é linear nos parâmetros e os polinômios do numerador e denominador são independentes. Note que os regressores u(k-1) e y(k-1) aparecem em ambos os termos. Dado a sua flexibilidade, alguns sistemas podem ser melhor aproximados por essa classe de modelos, por exemplo sinais com sobressaltos, em contrapartida o NARX racional é mais sensível a presença de ruído. Além disso, a não linearidade nos parâmetros impossibilita aplicação direta de técnicas baseadas em MMQ (AGUIRRE, 2015).

2.2.2 Modelo NARX Polinomial

Quando uma função polinomial de grau ℓ é atribuída a F, sua estrutura é dada por:

$$y(k) = \theta_0 + \sum_{i_1=1}^n \theta_{i_1} x_{i_1}(k) + \sum_{i_1=1}^n \sum_{i_2=i_1}^n \theta_{i_1 i_2} x_{i_1}(k) \cdot x_{i_2}(k) + \dots + \sum_{i_1=1}^n \dots \sum_{i_l=i_{l-1}}^n \theta_{i_1 \dots i_l} x_{i_1}(k) \dots x_{i_l}(k) + e(k)$$

$$(2.6)$$

em que $n = n_y + n_u - d + 1$, θ_i é uma constante e os valores de x_i são dados de forma que $x_1(k) = y(k-1)$, $x_2(k) = y(k-2)$, ..., $x_{n_y}(k) = y(k-n_y)$, $x_{n_y+1}(k) = u(k-d)$, ..., $x_n(k) = u(k-n_u)$ (AGUIRRE; RODRIGUES; JÁCOME, 1998). Vale ressaltar que os polinomiais são lineares nos parâmetros, assim podendo ser escritos na forma matricial vista abaixo.

$$\mathbf{y} = \mathbf{\Psi}^T \boldsymbol{\theta} + \boldsymbol{\xi} \tag{2.7}$$

onde $\Psi = [\Psi_1, ..., \Psi_j]^T$, $\theta = [\theta_1, ..., \theta_j]$ e $\boldsymbol{\xi} = [\xi_1, ..., \xi_j]$. De forma que uma determinada predição é dada por

$$y(k) = \psi^{T}(k-1)\boldsymbol{\theta} + \xi(k)$$
(2.8)

O exemplo abaixo é obtido aplicando os parâmetros $l=2,\,n_y=1$ e $n_u=1$ a equação 2.6,

$$y(k) = \theta_1 + \theta_2 y(k-1) + \theta_3 u(k-1) + \theta_4 y(k-1)^2 + \theta_5 y(k-1)u(k-1) + \theta_6 u(k-1)^2 + e(k)$$
(2.9)

Note que apesar da quantidade pequena de atrasos e apenas grau 2, o número de termos já é elevado. A quantidade de parâmetros, n_{θ} , de um modelo multivariável com mesmo número de atrasos para entradas e saídas é dado por

$$n_{\theta} = 1 + \sum_{i=1}^{\ell} n_i, \quad n_i = \begin{cases} 1, & \text{se } i = 0\\ \frac{n_{(i-1)} \cdot (n_y n_{out} + n_u n_{in} + i - 1)}{i}, & \text{se } i > 0 \end{cases}$$
 (2.10)

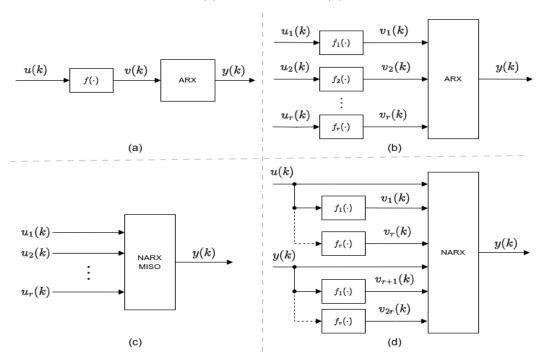
onde n_{out} e n_{in} são o número de saídas e entradas. A Equação 2.10 é uma extrapolação da equação presente em KORENBERG S. A. BILLINGS e McILROY (1988), que representa o caso monovariável, para múltiplas variáveis. Considerando os valores do exemplo anterior, tem-se $n_{\theta} = 6$, conforme visto na Equação 2.9.

2.3 NARX Proposto

A proposta é construir uma função não polinomial com novos regressores, x_i , obtidos a partir de y e u, ou seja, F não será apenas uma combinação algébrica dos atrasos das entradas e saídas. Como visto na Figura 4, os regressores adicionais vêm da aplicação de funções não lineares, f, aos sinais já conhecidos, isso aumenta a capacidade de representação do modelo, possibilitando a aproximação de dinâmicas mais complexas.

A Figura 3 apresenta um comparativo de diagramas, sendo 3.a, 3.b e 3.c estruturas bem documentadas na literatura e 3.d o foco deste trabalho. Mais especificamente, os modelos de Hammerstein são constituídos por um bloco não linear estático em cascata com um bloco linear dinâmico e invariante no tempo (AGUIRRE; RIBEIRO, 2014). Enquanto, o NARX MISO é uma extrapolação da Equação 2.1 para múltiplas entradas (AGUIRRE, 2015), herdando todas as características descritas na seção anterior. O modelo proposto, 3.d, pode ser aproximado ao visto em 3.c, se um exercício de abstração for realizado: imaginando que as variáveis intermediárias, $v_i(k)$, são entradas independentes de u(k) e y(k), um modelo com múltiplas entradas em cascata com um bloco NARX seria obtido, assim como no segundo. Em outras palavras, a proposta é semelhante a um sistema MISO, em que o bloco NARX recebe múltiplos entradas, mas aqui os sinais adicionais não são independentes de y e u.

Figura 3 – (a) Modelo de Hammerstein, sua generalização multivariáveis, (b) representação de Kortmann e Unbehauen,(c) NARX MISO e (d) NARX proposto.



Retomando a Figura 4, o modelo da esquerda mostra caso SISO detalhadamente, já o da direita possui blocos f_i adicionais e o bloco dinâmico NARX, que representa um polinômio usando u, y, v e w, respectivamente a entrada, a saída e dois sinais provindos a partir dos dois anteriores. Uma generalização possível incorpora múltiplas entradas e diferentes funções não lineares, além de uma única f, aplicadas aos atrasos. Assim, cada entrada poderia estar relacionada a mais de uma variável intermediária. A Equação 2.11 exemplifica uma situação específica em que $f(a) = \sin(a)$, nela F depende de 4 variáveis, o que é pouco perto do que pode ser alcançado.

$$y = F^{\ell}(u, v, w, y) = F^{\ell}(u, \sin(u), \sin(y), y)$$
(2.11)

Caso fosse expandida, assim como na Equação 2.9, o número de termo cresceria consideravelmente. Vale ressaltar que a Equação 2.11, apesar das não linearidades, pode escrita de forma matricial. A quantidade de regressores, n_{θ} , para um caso arbitrário é dada por

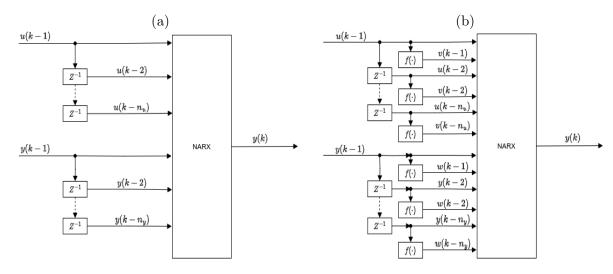
$$n_{\theta} = 1 + \sum_{i=1}^{\ell} n_{i},$$

$$n_{i} = \begin{cases} 1, & \text{se } i = 0\\ \frac{n_{(i-1)} \cdot (n_{y} n_{out} + n_{u} n_{in} + n_{v} n_{in} n_{fu} + n_{w} n_{out} n_{fy} + i - 1)}{i}, & \text{se } i > 0 \end{cases}$$

$$(2.12)$$

sendo n_y , n_u , n_v e n_w os atrasos respectivos a saída, entrada e não lineares aplicadas a ambas, ademais, n_{in} e n_{out} são número de entradas e saídas para um caso MIMO. Por fim, n_{fu} e n_{fy} são os números de funções aplicadas. Observe que para simplificar o cálculo presume-se que os atrasos são compartilhados entre os regressores do mesmo tipo.

Figura 4 – Representação detalhada do (a) NARX SISO e (b) do modelo com as alterações propostas.



O incremento desses novos fatores acarretam no crescimento progressivo do número de regressores. Aumentando custo computacional e dificultando a seleção de estrutura. A adição de funções não lineares vem acompanhada de novos desafios: a função pode ser indefinida no domínio dado e a independência linear dos regressores não é garantida. Em Aguirre (2015, p. 455), um problema mais complexos é discutido; não existe uma solução definitiva que consiga inferir quais os regressores devem ser gerados, a fim obter-se um resultado satisfatório, apenas observando o conjunto de dados. Aplicando ao contexto dessa subseção, não há como determinar quais são as melhores funções f apenas olhando os dados. Além disso, o desafio principal permanece: determinar F de modo a representar o sistema corretamente.

A Tabela 1 vem da aplicação da Equação 2.12 em diferentes parâmetros. É fácil notar que o número de regressores cresce rapidamente com aumento dos atrasos e do grau do

polinômio. Logo, manipular esse extenso conjunto de dados não é viável. Por isso, a aplicação correta da técnica de seleção de estrutura é tão importante, ela deve escolher o melhor conjunto para representar a dinâmica a partir de uma quantidade massiva de regressores candidatos.

Tabela 1 – Crescimento do número de regressores para um sistema SISO com apenas uma função não linear aplicada.

ℓ	n_y	n_u	n_v	n_w	n_{θ}
1	1	1	1	1	5
2	1	1	1	1	15
3	1	1	1	1	35
1	2	2	1	1	7
2	2	2	1	1	28
3	2	2	1	1	84
1	4	4	4	4	17
2	4	4	4	4	153
3	4	4	4	4	969

2.4 Estimação de parâmetros

Existem diversos métodos para a estimação de parâmetros de modelos não lineares registrados na literatura. Contudo, dada a representação matemática escolhida, NARX polinomial multivariável, essa tarefa resume-se a encontrar os valores θ da Equação 2.7. Assim como o NARX polinomial, o modelo proposto é linear nos parâmetros, que podem ser estimados pelo método de mínimos quadrados (AGUIRRE; RODRIGUES; JÁCOME, 1998; ARAÚJO, 2019).

2.4.1 Estimador de Mínimos Quadrados

O MMQ é um método clássico de estimação que serviu de base para o desenvolvimento de diversos outros. Foi criado Gauss para realização de estudo astronômicos, mas com o tempo diversos outros campos apropriaram-se dele (COELHO; COELHO, 2015).

Partindo da função escalar, vista abaixo, em que $f(\mathbf{x}, \boldsymbol{\theta}) : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$

$$y = f(\mathbf{x}, \boldsymbol{\theta}) \tag{2.13}$$

O método busca estimar os valores de θ a partir dos conjuntos de medições $\{y_1, \ldots, y_n\}$ e $\{x_1, \ldots, x_n\}$. Para isso, algumas considerações devem ser feitas:

- A função f e o vetor θ são os mesmo para todas as observações.
- A Equação 2.13 deve poder ser reescrita como

$$y = \mathbf{x}^T \boldsymbol{\theta} \tag{2.14}$$

• O número de observações estritamente necessário para estimação, geralmente, é $n = dim(\theta)$.

Levando em conta as considerações anteriores a Equação 2.13 pode ser escrita na seguinte forma:

$$\mathbf{y} = X\boldsymbol{\theta} \tag{2.15}$$

em que $X \in \mathbb{R}^{n \times n}$, $\boldsymbol{\theta} \in \mathbb{R}^{n \times 1}$ e $\mathbf{y} \in \mathbb{R}^n$. Realizando uma manipulação algébrica, o vetor dos parâmetros pode ser determinado a partir da Equação abaixo, desde que X não seja singular.

$$\boldsymbol{\theta} = X^{-1}\mathbf{y} \tag{2.16}$$

Dado o erro inerente a medição, ruído, um número de medições maior que mínimo é necessário. Assim, considerando que o número de observações, N, é maior que n, a equação anterior pode ser escrita na forma matricial

$$\mathbf{y} = X\hat{\boldsymbol{\theta}} + \boldsymbol{\xi} \tag{2.17}$$

sendo que $X \in \mathbb{R}^{N \times n}$, $\hat{\boldsymbol{\theta}} \in \mathbb{R}^{n \times 1}$, $\boldsymbol{\xi} \in \mathbb{R}^{N \times 1}$ e $\mathbf{y} \in \mathbb{R}^{N \times 1}$. Note que o acento em $\hat{\boldsymbol{\theta}}$ indica que os parâmetros não são os valores reais, mas, sim, estimados, consequentemente levando ao aparecimento do resíduo, $\boldsymbol{\xi}$ (AGUIRRE, 2015).

Contextualizando as informações anteriores com o restante deste trabalho. O desafio para determinar F^{ℓ} é análogo a encontrar f. Quanto a 1ª consideração, o modelo NARX é invariante no tempo, então f e θ são constantes. Em relação a segunda, tanto o modelo polinomial quanto o proposto são lineares nos parâmetros, aceitando a representação 2.14.

Como N > n, a matriz X não é mais quadrada, logo não é inversível. Por isso, o conjunto de operações abaixo deve ser aplicada na Equação 2.15 para obtenção da matriz pseudoinversa.

 \bullet Multiplicação de ambos os lados da equação por X^T

$$X^T \mathbf{y} = X^T X \boldsymbol{\theta} \tag{2.18}$$

 $\bullet\,$ A X^TX é quadrada, então, se não for singular, obtêm-se

$$\boldsymbol{\theta} = [X^T X]^{-1} X^T \mathbf{y} \tag{2.19}$$

Observando a Equação percebe-se que quanto melhor a estimativa de $\hat{\boldsymbol{\theta}}$, menor o valor de $\boldsymbol{\xi}$. Se o índice de qualidade, erro quadrático, for definido como $J_{MQ} = \|\boldsymbol{\xi}\|^2$, o $\hat{\boldsymbol{\theta}}$ que minimiza o erro é dado por

$$\hat{\boldsymbol{\theta}} = arg_{\boldsymbol{\theta}} min(J_{MQ})$$

$$= [X^T X]^{-1} X^T \mathbf{y}$$
(2.20)

Os cálculos necessários para obtenção da equação acima foram omitidos, mas foram amplamente relatados na literatura. O importante é perceber que a solução que minimiza o erro quadrático é igual a obtida utilizando a matriz de pseudoinversa. É importante garantir que X^TX seja inversível, ou seja, que a condição

$$det\{X^TX\} \neq 0 \tag{2.21}$$

seja satisfeita. Em relação ao modelo proposto, os regressores candidatos devem ser linearmente independentes. Além disso, os parâmetros estimados somente convergirão para os valores verdadeiros se o ruído for branco e erro de estimação for estaticamente independente dos regressores (AGUIRRE, 2015; COELHO; COELHO, 2015).

2.5 Seleção de estrutura

Quando o MMQ é aplicado, considera-se que todas as variáveis independentes devem ser incorporados. No entanto, como visto em Tabela 1, uma grande quantidade regressores é gerado por parâmetros configuráveis relativamente pequenos. Esta etapa busca descobrir quais deles devem ser mantidos. Destaca-se que manipular extensas matrizes tem um alto custo computacional, mesmo ignorando-o, outro problema permanece; definir um ponto de corte para aproveitamento dos regressores (AGUIRRE, 2015).

Existem diversos algoritmos para seleção de estrutura em modelos NARX e diferentes índices associados. Por exemplo, o método clássico de Gram-Schimdt (CGS), o modificado (MGS) e o forward regression orthogonal least squares (FROLS) utilizam a taxa de redução de erro (ERR). Enquanto, o simulation error minimization with pruning (SEMP) utiliza taxa de redução de erro por simulação (SRR). Cada um performa melhor em condições específicas e tem custos de computação associados. Isso posto, o SEMP não necessita realizar operações de ortogonalização, diferentemente dos demais citados, e é particularmente eficiente quando sistema real pode ser representado fielmente pela família de modelo NARX (PIRODDI; SPINELLI, 2003; AGUIRRE, 2015; BILLINGS, 2013).

2.5.1 SRR

Segundo Piroddi e Spinelli (2003), um critério de seleção deve possuir as seguintes propriedades:

- A cada iteração, o termo que possui o maior valor deve ser um dos regressores que compõem o sistema.
- A cada iteração, os regressores menos relevantes devem ter valores menores do que os regressores corretos.
- Os termos correspondentes aos regressores corretos devem ter valores comparáveis do critério.

O critério simulation error reduction ratio (SRR) é caracterizado pela diminuição do erro médio quadrado simulado (MSSE) ao adicionar o regressor j ao modelo.

$$[SRR]_j = \frac{MSSE(M_i) - MSSE(M_{i-1})}{\frac{1}{N} \sum_{t=1}^{N} y^2(t)}$$
(2.22)

onde M_i é o modelo obtido na i-ésima iteração e M_{i+1} o modelo candidato obtido na iteração subsequente, resultante da inclusão do j-ésimo regressor. Esse critério atende mais rigorosamente as propriedades citadas anteriormente do que o ERR. Vale ressaltar MSSE refere-se a simulação livre, o uso do preditor de um passe à frente levaria a

$$[ERR]_{j} = \frac{MSPE(M_{i}) - MSPE(M_{i-1})}{\frac{1}{N} \sum_{t=1}^{N} y^{2}(t)},$$
(2.23)

sendo MSPE o erro quadrado médio predito (PIRODDI; SPINELLI, 2003).

2.5.2 Simulation Error Minimization with Pruning - SEMP

O algoritmo FROLS e o simulation Error Minimization (SEM) apresentam algumas similaridades, ambos realizam a redução do EMQ para estimação e selecionam um novo regressor com base no valor de um índice, respectivamente ERR e SRR. No entanto, nem sempre um regressor permanece relevante com passar das iteração, por isso a poda é necessária, cujo processo avalia o quanto cada termo contribui para modelo e define sua permanência. Adicionando-a ao SEM, tem-se o SEMP que é caracterizado por gerar modelos compactos e robustos. Sua execução é composta pelos seguintes passos:

- 1. Inicialização: $\Pi_{in} = \emptyset$ e $J_{S,old} = \infty$
- 2. Forward regression:

Faça $\Pi_{out} = \Phi \setminus \Pi_{in}$

 $\forall \varphi_k(t) \in \Pi_{out}$, faça $\Pi_k = \Pi_{in} \cup \{\varphi_k(t)\}$, estime o modelo por meio de $\theta_k = emq(y, \Pi_k)$ e guarde $J_S(\theta_k, \Pi_k)$.

Faça $w = argmin_k[J_S(\theta_k, \Pi_k)]$

Se $J_S(\theta_w, \Pi_w) < J_{S,old}$, então $\Pi_{in} = \Pi_{in} \cup \{\varphi_w(t)\}$ e $J_{S,old} = J_S(\theta_{in}, \Pi_{in})$, senão termine.

3. Poda:

 $\forall \varphi_k(t) \in \Pi_{in}$, faça $\Pi_k = \Pi_{in} \setminus \{\varphi_k(t)\}$, estime o modelo $\theta_k = emq(y, \Pi_k)$ e guarde $J_S(\theta_k, \Pi_k)$.

Faça $w = argmin_k[J_S(\theta_k, \Pi_k)].$

Se $J_S(\theta_w, \Pi_w) < J_{S,old}$, então $\Pi_{in} = \Pi_{in} \setminus \{\varphi_w(t)\}$ e vá para etapa 3, senão $J_{S,old} = J_S(\theta_{in}, \Pi_{in})$ e vá para 2.

sendo Φ o conjunto dos regressores candidatos, φ um regressor, Π_{in} e Π_{out} os regressores incluídos e excluídos do modelo, a função $J_S = MSSE$ e emq o estimador de mínimos

quadrados. Em outras palavras, a cada iteração, o regressor que melhor contribui para desempenho é adicionado ao modelo. O acréscimo pode ocasionar a perda de relevância de um antigo regressor, o qual deve ser removido. Processo repete-se o número de vezes estipulado ou até não ocorrerem mudanças no modelo. Note que índice adotado foi o MSSE, caso o MSPE o substitui-se, o algoritmo forward-regression with pruning (FRP) seria obtido, ocasionando uma diminuição severa do tempo de computação, em contrapartida seu desempenho pode ser inferior ao SEMP (PIRODDI; SPINELLI, 2003; BONIN; SEGHEZZA; PIRODDI, 2010). Levando em conta a consideração anterior, o FRP pode ser usado em um momento inicial redução do espaço de busca e posteriormente o SEMP determinaria o modelo definitivo.

2.6 Validação

Após selecionar a estrutura e estimar os parâmetros, é necessário definir o quão bem o modelo representa o sistema. Dividir os dados, de modo que um conjunto utilizado na identificação não seja usado durante a validação, é um cuidado importante a se tomar. Perceba que o modelo pode estar sobreajustado aos dados usados na estimação e não representar bem a dinâmica em outras ocasiões, por isso essa medida busca garantir sua capacidade de generalização.

Neste trabalho, serão abordados dos tipos de simulação: a predição de um passo à frente e a predição livre ou de infinitos passos à frente. Supondo que F^{ℓ} tenha sido determinada, as predições são dadas, respectivamente, por

$$\hat{y} = F^{\ell}(y, u) \tag{2.24}$$

$$\hat{y} = F^{\ell}(\hat{y}, u) \tag{2.25}$$

o acento em \hat{y} detona que trata-se de um valor estimado. O passo à frente indica que a predição do instante k utiliza apenas observações de y anteriores a ele. Por outro lado, os infinitos passos à frente indicam que a predição atual utiliza as próprias predições em instantes anteriores para calcular um nova. Vale ressaltar que essa simulação não é totalmente desvinculada das observações, pois os valores iniciais são calculados por meio delas (AGUIRRE, 2015).

Supondo que o sistema pode ser escrito na forma $\hat{y} = \psi^T \hat{\boldsymbol{\theta}}$, sua estrutura é linear e com atrasos n_y e n_u , os regressores teriam a seguinte forma:

$$\psi^{T} = [y(k-1) \dots y(k-n_y) \ u(k-1) \dots u(k-n_u)]$$

$$\psi^{T} = [\hat{y}(k-1) \dots \hat{y}(k-n_y) \ u(k-1) \dots u(k-n_u)]$$

a primeira refere-se ao preditor de um passo à frente, já a segunda a de infinitos passos à frente. Retomando o exemplo de 2.11 sua predição livre seria

$$\hat{y}(k) = \theta_1 + \theta_2 \hat{y}(k-1) + \theta_3 u(k-1) + \theta_4 \sin[\hat{y}(k-1)] + \theta_5 \sin[u(k-1)]$$
(2.26)

note que o parâmetro da função seno também é acentuado, logo estimado. Os seus valores preditos têm o seguinte andamento

$$\hat{y}(1) = [y(0) \ u(0) \ \sin y(0) \ \sin u(0)] \hat{\boldsymbol{\theta}}
\hat{y}(2) = [\hat{y}(1) \ u(1) \ \sin \hat{y}(1) \ \sin u(1)] \hat{\boldsymbol{\theta}}
\vdots
\hat{y}(k) = [\hat{y}(k-1) \ u(k-1) \ \sin \hat{y}(k-1) \ \sin u(k-1)] \hat{\boldsymbol{\theta}},$$
(2.27)

portanto algumas observações servem para estabelecer a condição inicial do sistema simulado.

2.6.1 Métricas de erro

Após definir o tipo de simulação a ser executada, surge a necessidade mensurar a performance do modelo. Isto é, um índice, que informe o quando os dados preditos condizem com os amostrados, precisa ser obtido para aferir a qualidade da estimação.

O erro médio absoluto, MAE, é uma medida comumente utilizada que calculada como

$$MAE = \frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2,$$
(2.28)

sendo n o número de amostra, y os valores medidos e \hat{y} os preditos. Quanto menor o valor obtido, mais próximo as predições estão dos valores reais. Essa métrica dá o mesmo peso para cada amostra do erro. Enquanto, a raiz do erro quadrático médio, RMSE, é dada por

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2}$$
 (2.29)

Assim como MAE, o RMSE tem a mesma unidade de medida da variável e seu valor é inversamente proporcional a qualidade da estimação. Por outro lado, as instâncias de erro com maior valor absoluto ou *outliers* têm maior impacto sobre o valor final, diferente da anterior que é não é tão afetado. Segundo Chai e Draxler (2014), o MAE é recomendado para erros com uma distribuição uniforme, ao passo que o RMSE é mais adequado para distribuições normais.

Por sua vez, o erro absoluto percentual médio, MAPE, é dado por

$$MAPE = \frac{1}{n} \sum_{i=1}^{n} \left| \frac{y_i - \hat{y_i}}{y_i} \right|,$$
 (2.30)

sendo n o número de amostra, y os valores medidos e \hat{y} os preditos. O valor de y_i no denominador, claramente, não pode ser 0 e também os valores $y_i < 1$ são um problema, pois podem gerar grande erros percentuais, cujo impacto tornaria o valor final sem sentido. A desconsideração de valores atipicamente altos durante o cálculo pode ser uma solução possível (MAKRIDAKIS, 1993).

2.7 Tanques acoplados

O conhecimento sobre mecânica dos fluidos será de grande valia para o entendimento do funcionamento de um dos casos de estudo. As equações aqui apresentadas servirão como informação auxiliar para determinação da topologia do modelo de tanques. Para isso, alguns conceitos da hidrodinâmica serão apresentados abaixo.

Partindo da lei de conservação de massa, sua forma integral pode ser escrita na forma reduzida

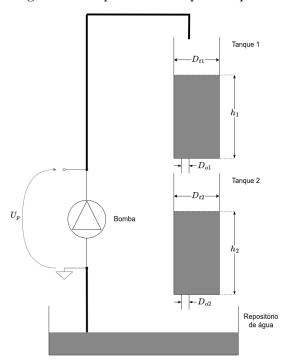
$$\frac{d}{dt} \int_{VC} \rho \, d\forall + \int_{SC} \rho(\vec{V} \cdot \vec{n}) \, dA = 0 \tag{2.31}$$

sendo VC volume de controle, SC superfície de controle, ρ a massa específica, \forall volume e \vec{V} o vetor velocidade. Considerando que o escoamento incompressível (massa específica constante), o vetor velocidade é normal à área e o perfil de velocidade nas seções transversais é uniforme, a equação pode ser escrita como

$$\frac{d\forall}{dt} + \sum_{i=1}^{n} (V_i A_i)_{sai} - \sum_{j=1}^{n} (V_j A_j)_{entra} = 0$$
 (2.32)

onde i=1,...,n são as superfícies de controle com fluxo de saída e j=1,...,m são as superfícies de controle com fluxo de entrada. Ou seja, nesse caso a variação do volume interno é a diferença entre o volume que entra e o que sai pelas superfícies de controle (FRAGOSO JR; LINS; BLENINGER, 2021).

Figura 5 – Esquema de tanques acoplados.



O esquema visto na Figura 5 representa o sistema presente no laboratório. Se a

equação 2.32 for aplicada nesse contexto, o seguinte conjunto de equações é obtido:

$$\frac{d\forall_1}{dt} + q_2 - q_1 = 0
\frac{d\forall_2}{dt} + q_3 - q_2 = 0$$
(2.33)

sendo q_1 , q_2 e q_3 , respectivamente, a vazão volumétrica fornecida pela bomba, a que deixa o tanque superior em direção ao inferior e, por fim, a que deixa o tanque inferior para o repositório.

Como os tanques são tubulares é possível escrever $\frac{d\forall}{dt} = A\frac{dh}{dt}$ e em seguida manipular a equação anterior para obter

$$\frac{dh_1}{dt} = \frac{q_1 - q_2}{A_{t1}}
\frac{dh_2}{dt} = \frac{q_2 - q_3}{A_{t2}}$$
(2.34)

sendo A_{t1} e A_{t2} as áreas das seções tubulares e h_1 e h_2 as alturas dos respectivos tanques. A vazão q_1 pode ser descrita de acordo com as especificações da bomba, enquanto a aplicação da equação de Bernoulli nas demais vazões resulta em $q_2 = V_1 a_1 = a_1 \sqrt{2gh_1}$ e $q_3 = a_2 \sqrt{2gh_2}$, onde g é a constante da força gravitacional (FOX; MCDONALD; PRITCHARD, 2014). Dado que no exemplo $a_1 = a_2$ e $A_{t1} = A_{t2}$, denote $\alpha = \frac{a_1}{A_{t1}} \sqrt{2g} = \frac{a_2}{A_{t2}} \sqrt{2g}$. Resultando no conjunto equações diferenciais abaixo.

$$\frac{dh_1}{dt} = -\alpha \cdot \sqrt{h_1} + \frac{K_m}{A_{t1}} \cdot u$$

$$\frac{dh_2}{dt} = \alpha \cdot \sqrt{h_1} - \alpha \cdot \sqrt{h_2} \tag{2.35}$$

sendo K_m a constante de fluxo da bomba e u a tensão de entrada (APKARIAN; LACHE-RAY; ABDOSSALAMI, 2013). Por fim, usando a aproximação implícita de Euler com passo de integração T_s , a equação anterior torna-se

$$h_1(t_{n+1}) = h_1(t_n) + T_s \cdot \left(-\alpha \cdot \sqrt{h_1(t_n)} + \frac{K_m}{A_{t_1}} \cdot u(t_n)\right)$$

$$h_2(t_{n+1}) = h_2(t_n) + T_s \cdot \left(\alpha \cdot \sqrt{h_1(t_n)} - \alpha \cdot \sqrt{h_2(t_n)}\right)$$
(2.36)

3 METODOLOGIA

3.1 Processo de identificação

As etapas descritas na seção 2.1 terão sua aplicação detalhada a seguir. Partindo da Figura 6, nota-se a ausência da coleta de dados, a qual será descrita caso a caso em seções posteriores. Inicialmente, os parâmetros configuráveis são requisitados, ℓ , n_y , f_1 e etc., cabe ao usuário defini-los. O grau e os atrasos recebem um número inteiro positivo cada, enquanto as funções são selecionadas a partir de uma lista de números naturais, onde 0 representa ausência e os demais valores a quantidade de atrasos. No exemplo abaixo, todos foram reunidos em dicionário.

Código 1 – Dicionário de parâmetros configuráveis em Python

O uso de critérios de informação pode facilitar a descoberta dos valores, mas dada a não linearidade, seu resultado pode ser inutilizado. O ajuste manual, apesar de incerto, é uma alternativa.

GRAU (n_y, n_u, d) $(y, v, u, \ldots, u^{\ell})$ REPRESENTAÇÃO REVISÃO DE $(y, v, u, ..., u^t)$ INÍCIO ATRASOS SIMBÓLICA LINEARIDADE $(f_1, ..., f_n)$ FUNCŌES NÃO LINEARES $\hat{y}(k) = X^T(k-1)\hat{\theta}$ REGRESSORES SELEÇÃO DE VÁLIDO? MODELO CANDIDATOS ESTRUTURA ENTRADA

Figura 6 – Etapas da identificação do modelo proposto

Como visto na seção 2.3, acrescentar os regressores não lineares, a primeira vista, não indica, mas a independência linear dos regressores pode ser comprometida. Em outras palavras, a matriz ψ tornaria-se singular, impossibilitando a estimação dos parâmetros. A representação simbólica realiza o produtório dos termos, vale ressaltar que apenas símbolo são manipulados, não números. Por exemplo, os parâmetros $\ell=2$, $n_y=1$, $n_u=1$, d=0

e $f_1(x) = \sqrt{x}$ com apenas um atraso sobre y geraram

$$\psi^{T} = \left[\underbrace{y_{1} \ u_{1} \sqrt{y_{1}}}_{\ell=1} \ \underbrace{y_{1}^{2} \ y_{1} u_{1} \ y_{1} \sqrt{y_{1}} \ u_{1}^{2} \ u_{1} \sqrt{y_{1}} \ (\sqrt{y_{1}})^{2}}_{\ell=2} \right], \tag{3.1}$$

em que o 1 sobrescrito representa um atraso, essa notação foi adotada para encurtar a escrita da equação. Percebe-se que $(\sqrt{y_1})^2 = y_1$, o que quebra a restrição citada acima. A revisão de linearidade utiliza recursos da biblioteca Sympy para simplificar as expressões e em seguida remover as redundâncias. Esse tipo de dependência linear é evidente, por isso foi facilmente removida. No entanto, dependências menos óbvias não são captadas de forma tão simples. O usuário deve auxiliar na revisão dos termos antes de prosseguir para próxima etapa, adicionando regras para remove-los e inspecionando-os.

A geração dos regressores candidatos consiste valorar as equações simbólicas e agrupalas em um matriz. Aqui os sinais armazenados se fazem presentes. Note que novos problemas surgem: não necessariamente o domínio da função coincide com os sinais e a função aplicada pode causar apenas uma transformação linear. O primeiro problema é exemplificado pela \sqrt{x} quando um dos sinais tem valores negativos. O segundo, quando a função f(x) = |x| e x > 0 em todas as amostras. Mais uma responsabilidade é atribuída ao usuário; analisar os sinais antes de selecionar as funções.

A etapa seguinte engloba a seleção de estrutura e a estimação de parâmetros. Os algoritmos utilizados serão SEMP, FRP e o MMQ. Por fim, a validação irá simular o modelo com um conjunto de dados específico e quantificar o erro. Todo o processo será repetido até um modelo satisfatório ser encontrado. Como os critérios de informação não tiveram contribuição relevantes, o ajuste será manual. Os parâmetros e os índices serão armazenados a cada ciclo e o modelo que possuir menores valores será prestigiado.

3.2 Coleção de funções

A fim de auxiliar o enfrentamento de dois problema citados na 2.3; a busca por F^{ℓ} e a escolha dos regressores candidatos que devem ser gerados, um conjunto de funções não lineares foi disponibilizado para geração de termos não lineares. A coleção inicial é a seguinte

$$[sen(x), \cos(x), \log(x), \tanh(x), sign(x), \tanh(2x), \tanh(5x), \tanh(10x), \\ \tanh(0.5x), \sqrt{x}, |x|]$$
 (3.2)

sendo sign a função sinal. Como o modelo é linear nos parâmetros não é possível estimar a constante v em $\tanh(v\cdot x)$, por isso diferente valores de v foram predefinidos. Novamente, vale ressaltar que durante a execução deve-se verificar o domínio da função e a independência linear dos termos. O mesmo vale para quando uma nova função for adicionada ao conjunto.

A aplicação de todas as funções ao mesmo tempo, gerando uma quantidade massiva de regressores, ou a adição funções à esmo não são práticas viáveis, tanto a definição de parâmetros configuráveis quanto a incorporação de novas funções devem guiar-se pelo conhecimento prévio do sistema.

3.3 Casos de estudo

O primeiro exemplo consiste em um conjunto de dados coletados de um sistema de tanques presente no laboratório de controle da universidade, dada a facilidade de acesso aos dados será caso mais extenso. As entradas foram definidas arbitrariamente e não houve calibração ou tratamento posterior a coleta dos dados, ou seja, o comportamento momentâneo do sensor e dos demais componentes será incorporado a dinâmica do modelo. O segundo caso será um benchmark presente em um repositório focado identificação, STADIUS's Identification Database. Por último, os sinais gerados computacionalmente descrevem um exemplo popular na teoria de controle.

3.3.1 Sistema de tanques acoplados

O sistema consiste em dois tanques dispostos verticalmente, chamaremos o tanque superior de tanque 1 e o inferior de tanque 2. O tanque 1 é alimentado por uma bomba hidráulica de e uma mangueira que leva o fluido até o fundo do recipiente. Ambos os tanques possuem abertura no fundo, que derramam o líquido no recipiente logo abaixo. Ambas as aberturas não são reguláveis, adicionalmente o tanque 1 possui uma válvula que leva fluido diretamente a cuba sem passar pelo segundo tanque.



Figura 7 – Sistema de tanques

Fonte: Quanser

Os tanques possuem altura total de 30 cm e o diâmetro interno de 4.45 cm. Enquanto, a bomba pode realizar recalque ou sucção, sendo sua tensão de alimentação limitada a 12V e -12V, respectivas aos modos de operação citados. Foram coletados os sinais referentes

ao nível de líquido de cada recipiente, provenientes dos sensores presentes no fundo dos tubos. Além disso, o sinal de tensão aplicado foi agrupado com os citados anteriormente para formar um conjunto de dados, vide a Figura 8.

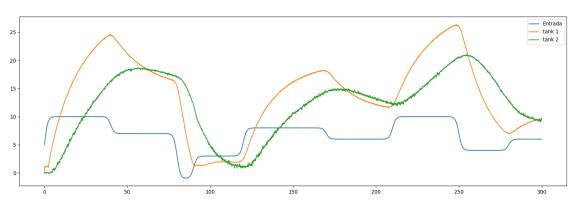


Figura 8 – Sinais coletados

3.3.2 Motores acoplados a uma polia

O sistema CE8 coupled electric drives consiste em dois motores que acionam uma polia através de uma correia flexível. O que pode ser visto na Figura 9. A polia é sustentada por uma mola, ou seja sua posição é móvel e apresenta uma dinâmica levemente amortecida. Enquanto, os motores têm posição fixa e são controlados individualmente. Levando a um controle simultâneo tanto da velocidade quanto da tensão da correia (WIGREN; SCHOUKENS, 2017).

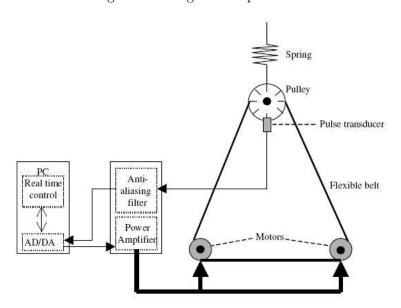


Figura 9 – Diagrama esquemático

Fonte: Wigren e Schoukens (2017)

O sinal de entrada é a tensão aplicada aos motores e a saída representa a velocidade angular da polia, que é medida por um contador de pulsos. Este sensor é insensível ao

sentido rotação, ou seja as medições sempre possuem sinal positivo. Os ensaios divergem principalmente pelo tipo de entrada aplicada: PRBS ou distribuição uniforme. Wigren e Schoukens (2017) argumentam que para identificação de sistemas não lineares tanto a frequência quanto amplitude do sinal são importantes. Por isso, os resultados vindos do segundo tipo foram apreciados por esse trabalho, trata-se de dois conjuntos de dados composto pelas duplas (u1, z11) e (u2, z12). A primeira foi utilizada para estimação e a segunda para validação.

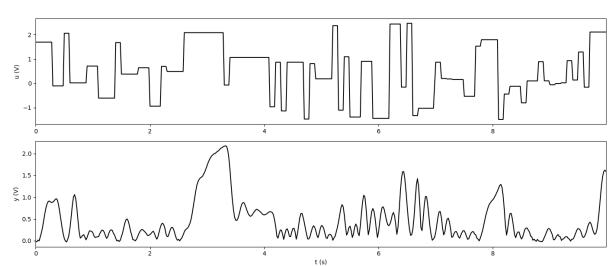


Figura 10 – Sinais de estimação CE8, sendo o primeiro a entrada e o segundo a saída

A figura acima contém a dupla (u1, z11), ambos os sinais representam tensões, sendo a primeira aplicada nos motores e a segunda lida pelo sensor de velocidade. A coleta durou 10s e o período de amostragem foi de 20ms. O desafio para identificação reside em representar a dinâmica gerada pela constante de tempo para o acionamento dos motores em conjunto com a ação da mola e do filtro analógico de passa baixa (WIGREN; SCHOUKENS, 2017).

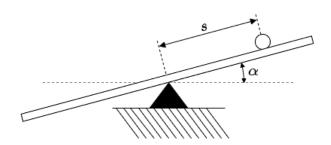
3.3.3 Ball and Beam

O último caso aborda sinal gerados computacionalmente para um sistema *Ball and Beam*, que comumente descrito por

$$\left(\frac{J}{r^2} + m\right)\ddot{s} - ms\dot{\alpha}^2 + mg\sin\alpha = 0 \tag{3.3}$$

onde α e s são, respectivamente, o ângulo da barra (rad) e a posição da bola (m) (GEMBALCZYK; DOMOGAłA; LEśNIOWSKI, 2023). A Figura 11 apresenta o sistema, destacando as variáveis capturadas.

Figura 11 – Diagrama esquemático Ball and Beam



Entretanto, a modelagem anterior desconsidera vários efeitos presentes no processo real, porém as equações abaixo oferecem uma descrição mais complexa do sistema

$$N = m(g\cos\alpha - r\dot{\alpha}^2 + \ddot{\alpha}s + 2\dot{s}\dot{\alpha}) \tag{3.4}$$

$$F_{at} = -sgn(\dot{s})\mu N \tag{3.5}$$

$$m\ddot{s} = F_{at} - mg\sin\alpha + m\dot{\alpha}^2 s + m\ddot{\alpha}r \tag{3.6}$$

onde N é força normal que age sobre a bola e F_{at} é força de atrito (COLÓN et al., 2014). Os parâmetros abaixo são baseados nos valores presentes em Gembalczyk, Domogała e Leśniowski (2023).

Tabela 2 – Parâmetros numéricos para simulação.

Parâmetro	Símbolo	Valor	Unidade
Momento de inércia	J	4.948×10^{-6}	$kg \cdot m^2$
Raio da bola	r	0.0118	m
Massa da bola	m	0.071	kg
Aceleração gravitacional	g	9.81	m/s^2
Coeficiente de atrito	μ	0.003	

A entrada, α , foi determinada arbitrariamente, a partir de uma distribuição uniforme, vista em 12.a. As alterações realizadas buscavam amortecer o movimento angular, atender os pontos discutidos em 3.3.2 e evitar instabilidades. O resultado pode ser observado na Figura 12.b. Para produção do sinal de saída s, utilizou-se a aproximação Euler para integração discreta, adotando um intervalo $t_i=0.01s$ entre as amostras de α e s. O processo é descrito pelo seguinte conjunto de equações discretas:

$$N[k] = m(g\cos\alpha[k-1] - r\dot{\alpha}^2[k-1] + \ddot{\alpha}[k-1]s[k-1] + 2\dot{s}[k-1]\dot{\alpha}[k-1])$$
 (3.7)

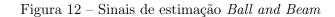
$$F[k] = -\mu \operatorname{sgn}(\dot{s}[k-1]) N[k] \tag{3.8}$$

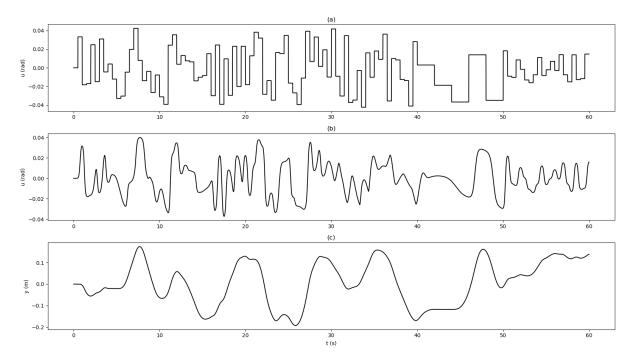
$$\ddot{s}[k] = a F[k] - b \sin(\alpha [k-1]) + c s[k-1] \dot{\alpha}^2 [k-1] + d \ddot{\alpha} [k-1]$$
(3.9)

$$\dot{s}[k] = \dot{s}[k-1] + t_i \cdot \ddot{s}[k] \tag{3.10}$$

$$s[k] = s[k-1] + t_i \cdot \dot{s}[k] + e[k] \tag{3.11}$$

sendo as constantes $a=\frac{r^2}{J+m\,r^2},\;b=\frac{mgr^2}{J+mr^2},\;c=\frac{mr^2}{J+mr^2},\;d=\frac{mr^3}{J+mr^2},\;t_i$ é o passo de integração e a variável e o ruído aditivo a saída. O sinal s pode ser visto na Figura 12.c.





4 RESULTADOS

Nesta seção, cada estudo de caso será composto pela apresentação da estrutura selecionada, das métricas de desempenho e, por fim, a comparação com modelos NARX polinomial.

4.1 Tanques acoplados

Nesta seção, denominou-se de y_1 a altura di líquido no tanque superior, y_2 a altura no tanque inferior e u a tensão aplicada na bomba.

4.1.1 Tangue superior

O modelo a seguir foi obtido a partir do conjunto de regressores candidatos gerado pelos parâmetros $\ell = 1$, $n_{y1} = 2$, $n_{y2} = 0$, d = 8 e $n_u = 12$, que foram obtidos manualmente por experimentação.

$$y_1(k) = 0.02485 \cdot u(k-12) + 1.00029 \cdot y_1(k-1) - 0.04768 \cdot \sqrt{y_1(k-1)}$$
 (4.1)

A raiz quadrada foi a única não linearidade selecionada e foi aplicada apenas nas saídas. Como as alturas são sempre positivas, a raiz sobre o regressor $y_1(k-1)$ não trará empecilhos a estimação, pois a função é definida e inversível no intervalo dado, 0 a 30. Assim, as condições discutidas na seção 3.1 foram atendidas. Além disso, conciliando com o conhecimento a priori do sistema, observa-se a similaridade do modelo obtido com a Equação 2.36. O que foi o fator determinante para definir o fim do ajuste manual.

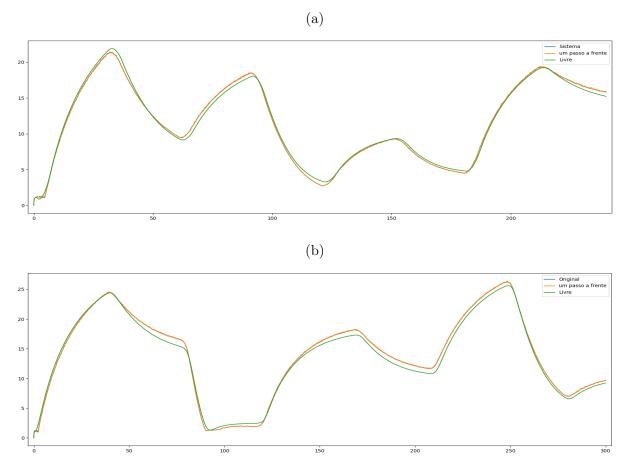
Tabela 3 – Métricas de ambas as simulações referentes ao tanque superior

Simulação	MAPE		RMSE	
Silifulação	Estimação	Validação	Estimação	Validação
Um passo a frente	0.36%	0.34%	0.0402	0.0397
Livre	3.95%	6.13%	0.3779	0.6498

Posteriormente, os valores apresentados na tabela acima serão comparados com os obtidos em um modelo NARX polinomial, de modo a verificar o ganho vindo da adição de não linearidades ao modelo.

Em ambos os gráficos da Figura 13, as predições apresentam grande similaridade, de tal forma que a curva azul, referente a saída real do sistema, está quase sobreposta pela laranja. O que é refletido pelos valores baixos de RMSE e MAPE referentes a simulação de um passo a frente. Por outro lado, as simulações livres possuem diferenças facilmente observáveis em relação ao sinal original, numericamente quantificadas pelos valores maiores vistos na última linha da Tabela 3, contudo as curvas estão bem ajustadas

Figura 13 – Simulações do tanque superior para a (a) estimação e (b) validação.



mesmo com o desempenho inferior. Nota-se que os maiores valores estão concentrados na simulação livre durante a validação, isso se deve ao maior rigor presente nessa etapa. Onde os dados não foram vistos pelo modelo durante sua identificação e os valores produzidos não usam amostras do sistema real.

4.1.2 Tanque inferior

Analogamente ao processo descrito na seção anterior, usando $\ell = 1$, $n_{y1} = 2$, $n_{y2} = 2$ e $n_u = 0$, o seguinte modelo para o tanque inferior foi encontrado

$$y_2(k) = 0.66020 \cdot y_2(k-1) + 0.33451 \cdot y_2(k-2) + 0.05483 \cdot \sqrt{y_1(k-1)} - 0.02333 \cdot \sqrt{y_2(k-1)} - 0.052261$$

$$(4.2)$$

O melhor modelo encontrado, descrito na Equação 4.2, tem os menores valores possíveis atribuídos aos parâmetros l e n_u . O valor $n_u = 0$ detona que conhecendo altura do tanque superior, o modelo não depende diretamente da entrada. Enquanto, $\ell = 1$ diz que uma combinação linear dos regressores candidatos foi suficiente para representar a altura do tanque 2.

Novamente, as curvas azuis e laranjas estão sobrepostas, assim sendo impossível distinguilas, vide os valores baixos referentes a simulação de um passo a frente presentes na Tabela

(a)

16 Original
10 Um passo a frente
11 Original
12 Original
13 Original
14 Um passo a frente
15 Original
16 Original
17 Original
18 Original
19 Original
19 Original
19 Original
19 Original
10 Original

Figura 14 – Simulações do tanque inferior para a (a) estimação e (b) validação.

20 - Original Um passo a frente Simulação livre

15 - 5 - 0 - 50 - 100 - 150 - 200 - 250 - 300

4. Em comparação com os sinais de saída medidos no tanque 1, a presença do ruído é perceptível no tanque 2, dando o aspecto de linhas mais grossas as curvas azuis. Isso se deve a maior turbulência do líquido no tanque inferior, que cai livremente em direção ao fundo ou choca-se com volume presente. Os distúrbios na medição aparentaram ter uma distribuição não gaussiana, pois os parâmetros estimados estavam polarizados, por isso um filtro passa baixa foi aplicado ao sinal de estimação. Entretanto, o ruído foi mantido tanto na representação gráfica quanto no cálculo das métricas, buscou-se manter a equidade de condições para comparação dos modelos, pois o polinomial performou melhor no sinal não filtrado.

Tabela 4 – Métricas de ambas as simulações referentes ao tanque inferior

Simulação	MAPE		RMSE	
Silifulação	Estimação	Validação	Estimação	Validação
Um passo a frente	2.595%	1.936%	0.1178	0.1315
Livre	3.034%	3.786%	0.1378	0.3574

Dada a pouca relevância vista na apresentação das simulações de um passo a frente, as curvas serão omitidas das figuras nas seções seguintes. Entretanto, as métricas de erro

sobre elas serão mantidas por contribuírem para avaliação numérica dos resultados.

4.1.3 Comparativo

Seguindo o mesmo procedimento, agora para um modelo polinomial os seguintes modelos foram encontrados:

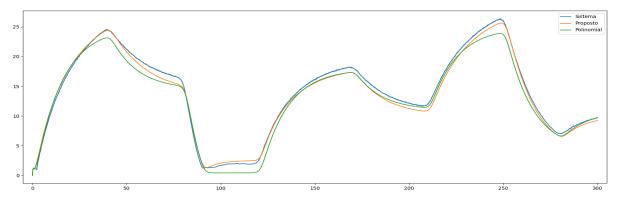
$$y_1(k) = 0.99289 \cdot y_1(k-1) + 0.02485 \cdot u(k-12) - 0.07151$$
 (4.3)

$$y_2(k) = 0.65144 \cdot y_2(k-1) + 0.3344 \cdot y_2(k-2) + 0.01255 \cdot y_1(k-1) - 0.00018 \cdot y_1(k-1)^2 + 0.00025 \cdot y_2(k-1)^2$$
(4.4)

tendo usado $\ell=1, n_{y1}=2, n_{y2}=0, d=8$ e $n_u=12$ para o tanque superior. Enquanto no inferior, os parâmetros foram $\ell=2, n_{y1}=1, n_{y2}=2, d=0$ e $n_u=2$.

O modelo visto na Equação 4.3 é de grau 1 assim como na Equação 4.1, ambas as representações apresentaram dinâmicas espúrias para graus mais elevados. Apesar da semelhança na estrutura, o modelo proposto teve um desempenho melhor que NARX polinomial, observe na Figura 15. Particularmente, quando a altura atingiu valores maiores que vinte ou próximos a zero, onde a curva laranja está mais ajustada as medições.

Figura 15 – Simulações livres sobre o conjunto de validação do tanque 1



Na Tabela 5, os valores são muito parecidos com suas contrapartes durante a predição de um passo a frente, porém os valores de RMSE da simulação livre durante a validação são bem diferentes. O menor valor referente ao modelo proposto indica sua maior capacidade de generalização, visto que desempenha melhor sobre um novo conjunto de dados.

Tabela 5 – Comparação entre o RMSE de ambos os modelos referentes ao tanque 1

Simulação	Polinomial		Proposto	
Silifulação	Estimação	Validação	Estimação	Validação
Um passo a frente	0.040	0.040	0.040	0.039
Livre	0.398	1.131	0.377	0.649

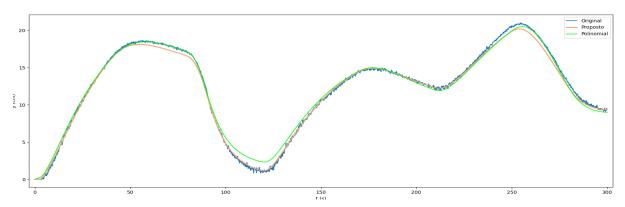
Quantitativamente, a superioridade do modelo proposto pode ser vista na disparidade de amplitude do MAPE sobre o conjunto de validação, 35.893% e 6.136%. Por outro lado, os valores da predição de um passo a frente são muito parecidos entre si, não tendo grande valia para análise, por isso sua representação gráfica foi novamente omitida.

Tabela 6 – Comparação entre o MAPE de ambos os modelos referentes ao tanque 1

Simulação	Polinomial		Proposto	
Silifulação	Estimação	Validação	Estimação	Validação
Um passo a frente	0.357%	0.367%	0.360%	0.341%
Livre	4.541%	35.893%	3.693%	6.136%

Mudando o foco para o tanque inferior. Em comparação a Equação 4.2, o modelo polinomial, Equação 4.4, diferencia-se por ter regressores de grau 2, mas continua desconsiderando a entrada u. Por sua vez, a análise da Figura 16 intui que, mais uma vez, o modelo proposto desempenhou melhor do que seu concorrente.

Figura 16 – Sinais de validação do tanque 2



Corroborando com a intuição anterior, as métricas referentes a validação mostram que o modelo proposto reproduz melhor a dinâmica do sistema. Entretanto, o modelo polinomial não tem comportamento discrepante na maior parte do tempo, suas falhas concentram-se no período de descarga do líquido, entre 100s e 150s.

Tabela 7 – Métricas de erro do tanque inferior durante a validação

Métrica Polinor		1		
Metrica	Estimação	Validação	Estimação	Validação
RMSE	0.2648	0.5103	0.1378	0.3574
MAPE	7.244%	11.412%	3.034%	3.786%

As métricas da predição de um passo a frente calculadas foram muito semelhantes, fato recorrente nas tabelas anteriores, em razão disso foram omitidas na tabela acima. Nos casos de estudo seguintes, quando forem relevantes, serão expostas.

4.2Coupled Electric Drives

Livre

4.2.1Modelo proposto

O modelo a seguir foi obtido a partir do conjunto de regressores candidatos gerado pelos parâmetros $\ell = 3$, $n_y = 10$, $n_u = 10$ e d = 2, que foram ajustados manualmente por experimentação. Diferente da seção anterior, o algoritmo usado para seleção de estrutura foi o FRP, pois o número de regressores candidatos gerados é muito alto e, como visto abaixo, a quantidade de selecionados também.

$$y_m(k) = -0.34066 \cdot y_m(k-3) + 0.01868 \cdot u(k-1) \cdot y_m(k-6) + 0.03369 \cdot u(k-4)$$

$$+ 0.02055 \cdot u(k-6) + 0.02805 \cdot u(k-5) \cdot y_m(k-10) \cdot tanh(2 \cdot y_m(k-1))$$

$$+ 0.00813 \cdot u(k-3) \cdot u(k-9)^2 + 0.02115 \cdot u(k-7) + 1.09045 \cdot y_m(k-1)$$

$$+ 0.041 \cdot y_m(k-10) + 0.00914 \cdot u(k-5)^2 \cdot tanh(2 \cdot y_m(k-1))$$

$$(4.5)$$

Lembrando que a saída definitiva, y, é dada por $y(k) = |y_m(k)|$, visto que o sensor de velocidade é insensível ao sentido de rotação.

Na tabela 8, é possível observar que os valores do MAPE para o conjunto de validação são melhores que os referentes a estimação, contrariando o comportamento visto no caso anterior, que é o esperado. Os valores são muito discrepantes em relação ao que foi apresentado até aqui, em especial o 78.201%. Por outro lado, as grandezas do RMSE não apresentam anomalias evidentes, observa-se que a ordem de grandeza é similar para todos e os valores referentes a estimação sempre são menores do que os de validação. Ressaltando a necessidade de verificar o porquê do baixo desempenho da métrica MAPE.

abela 8 – Métricas de ambas as simulações referentes aos motores acoplados					
Simulação	MAPE		RMSE		
Simulação	Estimação	Validação	Estimação	Validação	
Um passo a frente	34.317%	24.500%	0.0335	0.0528	

78.201%

s.

29.081%

0.0686

0.1068

Analisando mais a fundo, foi observado que o baixo desempenho do modelo está concentrado nos valores de saídas com baixa amplitude. Por exemplo, quando valor esperado é 0.02 e o previsto é 0.2, o erro percentual absoluto (APE) nessa amostra é 900%. Aliado a baixa quantidade total de amostras, casos como esse tem grande impacto sobre o valor final. Por isso, o resultado obtido na métrica MAPE não implica que a identificação foi inválida. Como visto na Figura 17, o comportamento do sinal original é bem descrito, apesar da estimação ser mais oscilatória que o sistema real.

Note que curva da simulação de um passo a frente foi omitida, já que sua representação não traria benefícios a análise gráfica, pois estaria quase sobreposta a curva azul, analogamente ao que foi visto nas subseções 4.1.1 e 4.1.2. Usando a mesma justificativa, a curva será ignorada no restante do trabalho.

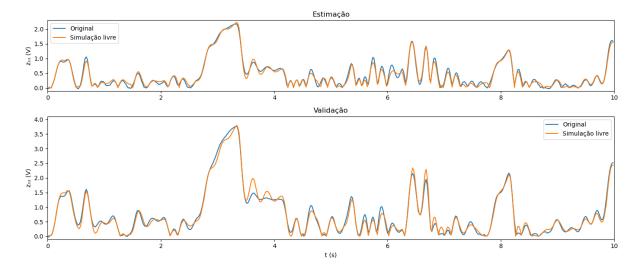


Figura 17 – Simulação livre sobre o conjunto de estimação e validação.

4.2.2 Comparativo

O modelo a seguir foi obtido a partir do conjunto de regressores candidatos gerado pelos parâmetros $\ell=3,\,n_y=10,\,n_u=10$ e d=0, que foram ajustados manualmente por experimentação.

$$y_m(k) = 1.16519 \cdot y_m(k-1) + 0.00923 \cdot u(k-3) \cdot u^2(k-9) - 0.35792 \cdot y_m(k-3)$$

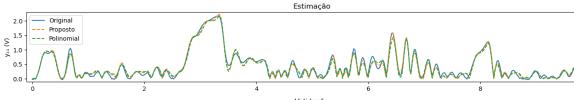
$$+ 0.03193 \cdot u(k-4) + 0.01202 \cdot u(k-5) \cdot u(k-6) \cdot y_m(k-10)$$

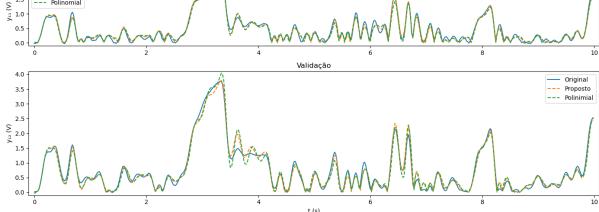
$$+ 0.05274 \cdot y_m(k-9) + 0.02903 \cdot u(k-6)$$

$$(4.6)$$

Similarmente a Equação 4.5, termos de grau 3, muitos de regressores e atrasos foram selecionados.

Figura 18 – Simulação livre durante a estimação e validação.





Graficamente, o comportamento também é semelhante, os dois modelos são subamortecidos, mas destoam do sinal original. Observável por volta dos 4s, quando a entrada assume um valor contínuo durante um período prolongado que causaria acomodação do sinal de saída. O modelo polinomial possui alguns sobressaltos, indicando um desempenho pior. Ademais, não há razão para caracterizar um dos modelos como superior.

Analisando a tabela abaixo, nota-se a melhor desempenho geral do modelo proposto em relação ao polinomial. Ressaltam-se os valores da simulação livre para conjunto de validação que descrevem um teste mais rigoroso.

Tabela 9 – Comparação entre o RMSE de ambos os modelos.

Simulação	Polinomial		Proposto	
Silifulação	Estimação	Validação	Estimação	Validação
Um passo a frente	0.0375	0.0538	0.0335	0.0528
Livre	0.0805	0.1339	0.0686	0.1068

O problema mencionado anteriormente permanece, ambos os modelos apresentam valores altos de MAPE, mostrando a baixa performance na faixa próxima de zero. O conjunto possui poucas amostras e muitas têm baixa magnitude, logo removê-las do cálculo não é uma alternativa. Por isso, a nova métrica foi adotada, veja a Tabela 11, analisando-a percebe-se que o modelo proposto mantêm o desempenho superior. No entanto,

Tabela 10 – Comparação entre o MAPE de ambos os modelos.

Simulação	Polinomial		Proposto	
Silifulação	Estimação	Validação	Estimação	Validação
Um passo a frente	34.895%	27.151%	34.317%	24.500%
Livre	92.463%	34.786%	78.201%	29.081%

Tabela 11 – Comparação entre o EAT de ambos os modelos.

Simulação	Polinomial		Proposto	
Simulação	Estimação	Validação	Estimação	Validação
Um passo a frente	12.9314	17.1897	11.4015	15.8938
Livre	31.8347	49.0200	26.9585	38.5658

Mesmo com muitos regressores selecionados, respectivamente 10 e 7, e grau 3, a dinâmica composta pela ação da mola em conjunto com a correia elástica não foi perfeitamente modelada, assim atestando a complexidade da não linearidade presente no sistema. Em suma, a ligeira superioridade nos resultados quantitativos não é uma vantagem significativa para prestigiar 4.5 e desconsiderar 4.6. A escolha deve ponderar qual atende melhor as suas necessidades, decidindo entre uma maior a acurácia e a simplicidade do cálculo.

4.3 Exemplo numérico

4.3.1 Modelo proposto

Um dos algoritmos presentes em Overschee e Moor (1996) realiza a identificação de sistema Ball and Beam. Apesar da representação matemática escolhida ser diferente, algumas práticas lá aplicadas podem ser replicadas neste trabalho: a filtragem do ruído e a utilização da velocidade, invés da posição, como saída. Dado que o sinal foi gerado artificialmente, basta suprimir o erro aditivo da saída durante cálculo. Já a velocidade foi obtida a partir da aproximação de Euler da derivada. Logo, a variável a ser estimada é $\dot{y}(k) \approx \frac{dy}{dt}(k)$. Diante disso, modelo a seguir foi obtido a partir do conjunto de regressores candidatos gerado pelos parâmetros $\ell=2,\,n_y=8,\,n_u=8,\,d=2,\,5$ atrasos para a função cos e 1 para sgn, que foram definidos por experimentação.

$$\dot{y}(k) = -0.09919 \cdot u(k-4) + 0.04885 \cdot u(k-8) \cdot \cos(\dot{y}(k-1)) + 0.79278 \cdot \dot{y}(k-1)$$

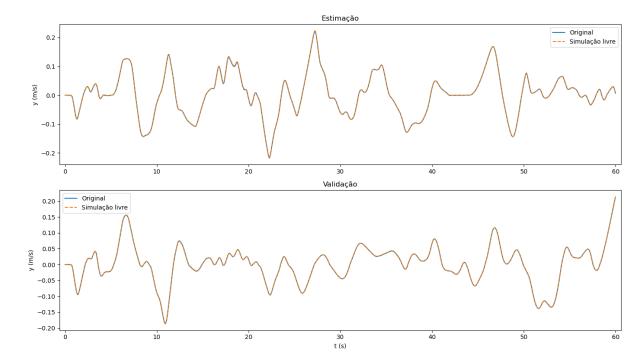
$$-0.03404 \cdot \dot{y}(k-8) \cdot \cos(u(k-5)) - 0.00015 \cdot \cos(u(k-2)) \cdot sgn(\dot{y}(k-1))$$

$$+0.40497 \cdot \dot{y}(k-2) - 0.09626 \cdot \dot{y}(k-3) - 0.06745 \cdot \dot{y}(k-4)$$

$$(4.7)$$

Observando a Figura 19, percebe-se que o modelo estimado está bem ajustado a curva em ambos casos. O que corrobora com valores vistos na Tabela 12, todas as métricas têm baixa magnitude. Lembrando que \dot{y} não é a saída definitiva. Posteriormente, ela será apresentada, possibilitando perceber o impacto que os erros de predição, mesmo que pequenos, causam no resultado final.

Figura 19 – Simulação livre da velocidade da bola para a estimação e validação.



Simulação	MAPE		RMSE	
Simulação	Estimação	Validação	Estimação	Validação
Um passo a frente	0.1529%	0.0457%	$7.90 \cdot 10^{-5}$	$7.08 \cdot 10^{-5}$
Livre	1.3675%	0.1225%	$7.64 \cdot 10^{-4}$	$5.62 \cdot 10^{-4}$

Tabela 12 – Métricas de ambas as simulações referentes aos motores acoplados.

4.3.2 Comparativo

O modelo a seguir foi obtido a partir do conjunto de regressores candidatos gerado pelos parâmetros $\ell = 3$, $n_y = 6$, $n_u = 7$ e d = 1, que foram ajustados manualmente por experimentação.

$$\dot{y}(k) = 1.16838 \cdot \dot{y}(k-1) - 0.16927 \cdot \dot{y}(k-6) - 0.04291 \cdot u(k-2) + 0.03349 \cdot u(k-7)
- 1.90190 \cdot u(k-2) \cdot \dot{y}(k-1)^2 + 27.96621 \cdot \dot{y}(k-1)^2 \cdot \dot{y}(k-2)
- 27.96489 \cdot \dot{y}(k-1)^3$$
(4.8)

Analisando a Tabela 13, nota-se que o modelo polinomial possui valores maiores em todas as ocasiões. Em especial, o RMSE da simulação livre que é mais de 6 vezes maior durante a estimação e mais de 9 vezes na validação. Indicando um desempenho inferior ao modelo proposto.

Tabela 13 – Comparação entre o RMSE de ambos os modelos.

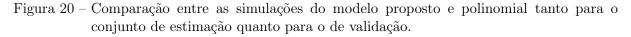
Simulação	Polinomial		Proposto	
Silifulação	Estimação	Validação	Estimação	Validação
Um passo a frente	$9.35 \cdot 10^{-5}$	$8.78 \cdot 10^{-5}$	$7.90 \cdot 10^{-5}$	$7.08 \cdot 10^{-5}$
Livre	$4.86 \cdot 10^{-3}$	$5.14 \cdot 10^{-3}$	$7.64 \cdot 10^{-4}$	$5.62 \cdot 10^{-4}$

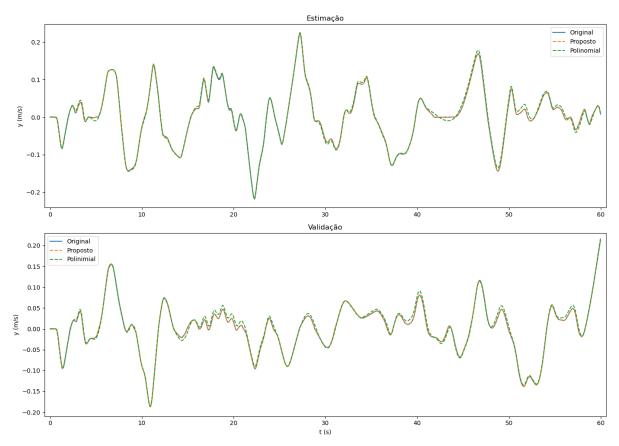
Dados os valores baixos do MAPE, todos inferiores a 2% e presentes na Tabela 14, espera-se que não ocorram grandes discrepâncias entre a curva gerada e o sinal original.

Tabela 14 – Comparação entre o MAPE de ambos os modelos.

Simulação	Polinomial		Proposto	
Silifulação	Estimação	Validação	Estimação	Validação
Um passo a frente	0.9306%	0.8265%	0.1529%	0.0457%
Livre	1.5378%	0.5769%	1.3675%	0.1225%

Analisando a Figura 20, conclui-se que as considerações anteriores são verdadeiras: o modelo proposto tem o melhor desempenho e ambos não apresentam valores atípicos. Verifica-se que o modelo polinomial em certos momentos aproxima bem o sistema, quase sobrepondo curva original, porém quando próximo a zero diferença é nítida. Na estimação, as falhas tornam-se evidentes após os 40s. Já para validação, o desvio no comportamento ocorre antes, a partir dos 14s as diferença podem ser notadas.





Como anunciado, a saída real será apresentada para averiguação do desempenho verdadeiro do modelo, as curvas foram obtidos através da integração discreta de \dot{y} . O modelo polinomial apresentou pequenos desvios que eram compensados com tempo, mas na reta final não há compensação e erro é acumulado. Enquanto na validação, as trajetórias começam divergir por volta dos 14s e não voltam coincidir, vide a Figura 21. Finalizando, o desempenho do modelo proposto mostrou-se realmente superior ao polinomial, sendo capaz de representar a dinâmica com mais fidelidade.

Estimação Original
Proposto
Polinomial y (m) -0.2 30 Validação 10 20 40 50 60 0.3 Original
Proposto
Polinimial 0.2 y (m) 0.0 -0.1 -0.2 30 t (s) 10 20 40 50

Figura 21 – Comparação da posição estimada com o sinal original.

5 CONCLUSÃO

Este trabalho apresentou a identificação de sistema usando modelos da família NARX. Escolheu-se uma representação linear nos parâmetros, sendo utilizado o SEMP para seleção de estrutura e o MMQ para estimação parâmetros. Em seguida, os modelos foram validados e comparados com sua contraparte polinomial.

O desafio de adicionar informação ao modelo e descobrir quais regressores candidatos devem ser gerados foi enfrentado por meio da aplicação de uma coleção de funções não lineares predefinida. Quanto a determinação de F^{ℓ} foi adotada uma representação com funções transcendentes, não polinomial, aliada a algoritmos clássicos de seleção de estrutura.

Foram três casos de estudo: tanques acoplados, motores acoplados à polia e ball and beam sendo, respectivamente, obtidos por coleta, repositório de terceiros e simulação computacional. No primeiro, a proposta teve resultados positivos em ambos os tanques. No segundo, o resultado foi levemente superior, mas uma vantagem clara não foi estabelecida. Ambos tiveram um número considerável de regressores selecionados e grau 3. Considerando a maior simplicidade para o cálculo, o modelo polinomial poderia ser admitido sem perda de desempenho evidentes. O último exigiu uma mudança de estratégia que o tornou mais sensível a ruído, mas como os sinais foram gerados computacionalmente, não houve grande empecilho para removê-lo. O desempenho do modelo proposto foi superior, em especial na validação, onde a discrepância foi nítida.

Os problemas levantados anteriormente, seção 2.3, não foram definitivamente solucionados; um grande esforço ainda é requerido durante o processo, no entanto, a proposta demostrou bons resultados em todos os casos de estudo, performando melhor quando aliado ao conhecimento prévio. Portanto, o uso dos termos não lineares mostrou-se uma boa adição para tarefa de identificação e uma alternativa viável ao NARX polinomial em certas situações. Uma boa prática seria estimar ambos os modelos e verificar se dada a situação a adição se justifica.

5.1 Melhorias futuras

- Usar a representação NARMAX: tornaria o mais resistente a presença de ruído, em contrapartida aumentaria o já alto número de regressores candidatos.
- Diferentes algoritmos de seleção de estrutura: cada seletor performa melhor em situações específicas.
- Critérios de informação: a aplicação pode levar a resultados mais robustos e facilitar o ajuste.

REFERÊNCIAS

- AGUIRRE, L. Introdução à Identificação de Sistemas. 4. ed. Belo Horizonte: Editora UFMG, 2015.
- AGUIRRE, L.; RIBEIRO, A. H. Relações estáticas de modelos narx miso e sua representação de hammerstein. In: . [S.l.: s.n.], 2014.
- AGUIRRE, L.; RODRIGUES, G.; JÁCOME, C. Identificação de sistemas não lineares utilizando modelos narmax polinomiais—uma revisão e novos resultados. **SBA Controle & Automação**, v. 9, n. 2, p. 90–106, 1998.
- APKARIAN, J.; LACHERAY, H.; ABDOSSALAMI, A. INSTRUCTOR WORKBOOK: Coupled Tanks Experiment for MATLAB/Simulink Users. Markham, Ontario, 2013.
- ARAÚJO, Í. B. Q. Maximização da correntropia por simulação com poda aplicado a detecção de estruturas e estimação de parâmetros de modelos NARX. Tese (Doutorado em Engenharia Elétrica e de Computação) Centro de Tecnologia, Universidade Federal do Rio Grande do Norte, Natal, Julho 2019. Disponível em https://repositorio.ufrn.br/handle/123456789/27738.
- BILLINGS, S. Nonlinear system identification: Narmax methods in the time, frequency, and spatio-temporal domains. In: _____. Wiley, 2013. ISBN 9781118535554. Disponível em: https://books.google.com.br/books?id=SaQ2AAAAQBAJ.
- BONIN, M.; SEGHEZZA, V.; PIRODDI, L. Narx model selection based on simulation error minimisation and lasso. **Control Theory & Applications, IET**, v. 4, p. 1157 1168, 08 2010.
- CHAI, T.; DRAXLER, R. R. Root mean square error (rmse) or mean absolute error (mae)? arguments against avoiding rmse in the literature. **Geoscientific Model Development**, v. 7, n. 3, p. 1247–1250, 2014. Disponível em: https://gmd.copernicus.org/articles/7/1247/2014/.
- CHEN, S.; BILLINGS, S. Representations of non-linear systems: the narmax model. **International Journal of Control**, v. 49, n. 3, p. 1013–1032, 1989.
- COELHO, A. A. R.; COELHO, L. d. S. Identificação de sistemas dinâmicos lineares. 2. ed. Florianópolis: editora UFSC, 2015.
- COLÓN, D. et al. Modeling and analysis of a ball and beam system including impacts and dry friction. In: [S.l.: s.n.], 2014. v. 8.
- DENG, L. et al. Narmax model based pseudo-hammerstein identification for rate-dependent hysteresis. **5th International Conference on Intelligent Control and Information Processing, ICICIP 2014 Proceedings**, v. 2015, p. 155–162, 01 2015.
- FOX, R. W.; MCDONALD, A. T.; PRITCHARD, P. J. Introdução à Mecânica dos Fluidos. 8. ed. [S.l.]: LTC, 2014. 312-323 p.

- FRAGOSO JR, C.; LINS, R.; BLENINGER, T. Mecânica dos Fluidos para Engenharia e Ciências Ambientais. [S.l.: s.n.], 2021. 194-204 p. ISBN 978-85-88686-46-5.
- GEMBALCZYK, G.; DOMOGAłA, P.; LEŚNIOWSKI, K. Underactuated ball and beam system a comparative study. **Actuators**, v. 12, n. 59, 2023.
- KADOCHNIKOVA, A. et al. Integrated identification of the nonlinear autoregressive models with exogenous inputs (narx) for engineering systems design. **IEEE**Transactions on Control Systems Technology, v. 31, n. 1, p. 394–401, 2023.
- KORENBERG S. A. BILLINGS, Y. P. L. M.; MCILROY, P. J. Orthogonal parameter estimation algorithm for non-linear stochastic systems. **International Journal of Control**, Taylor & Francis, v. 48, n. 1, p. 193–210, 1988. Disponível em: https://doi.org/10.1080/00207178808906169.
- LI, F. et al. Identification methodology for mimo hammerstein nonlinear model with process noise. Transactions of the Institute of Measurement and Control, 10 2024.
- MAKRIDAKIS, S. Accuracy measures: theoretical and practical concerns. **International Journal of Forecasting**, v. 9, n. 4, p. 527–529, 1993. ISSN 0169-2070. Disponível em: https://www.sciencedirect.com/science/article/pii/0169207093900793.
- MITREA, C.; LEE, C.; WU, Z. A comparison between neural networks and traditional forecasting methods: A case study. **International Journal of Engineering Business Management**, v. 1, 09 2009.
- MOZARYN, J. Narx recurrent neural network model of the graphene-based electronic skin sensors with hysteretic behaviour. In: [S.l.: s.n.], 2023.
- OVERSCHEE, P. V.; MOOR, B. D. **Subspace Identification for Linear Systems**. [S.l.]: MATLAB Central File Exchange, 1996. Disponível em: https://www.mathworks.com/matlabcentral/fileexchange/2290-subspace-identification-for-linear-systems. Acesso em: 12 de out. de 2024.
- PIRODDI, L.; SPINELLI, W. An identification algorithm for polynomial narx models based on simulation error minimization. **International Journal of Control**, v. 76, n. 17, p. 1767–1781, 2003.
- SCHOUKENS, J.; LJUNG, L. Nonlinear system identification: A user-oriented road map. **IEEE Control Systems Magazine**, v. 39, n. 6, p. 28–99, 2019.
- SCHOUKENS, M.; TIELS, K. Identification of block-oriented nonlinear systems starting from linear approximations: A survey. **Automatica**, v. 85, p. 272–292, 2017. ISSN 0005-1098. Disponível em: https://www.sciencedirect.com/science/article/pii/S0005109817303990.
- WIGREN, T.; SCHOUKENS, M. Coupled electric drives data set and reference models. **Department of Information Technology, Uppsala University**, 2017.