



Universidade Federal de Ouro Preto  
Instituto de Ciências Exatas e Aplicadas  
Departamento de Engenharia Elétrica



## **Trabalho de Conclusão de Curso**

# **Identificação de modelos de nível para uma planta didática SMAR PD3-F**

**Rafael Ferreira Félix**

João Monlevade, MG  
2024

**Rafael Ferreira Félix**

**Identificação de modelos de nível para uma  
planta didática SMAR PD3-F**

Trabalho de Conclusão de curso apresentado à Universidade Federal de Ouro Preto como parte dos requisitos para obtenção do Título de Bacharel em Engenharia Elétrica pelo Instituto de Ciências Exatas e Aplicadas da Universidade Federal de Ouro Preto.  
Orientadora: Prof.<sup>a</sup> Wendy Yadira Eras Herrera

**Universidade Federal de Ouro Preto  
João Monlevade  
2024**

## SISBIN - SISTEMA DE BIBLIOTECAS E INFORMAÇÃO

F316i Felix, Rafael Ferreira.  
Identificação de modelos de nível para uma planta didática SMAR PD3-  
F. [manuscrito] / Rafael Ferreira Felix. - 2024.  
57 f.: il.: color., tab..

Orientadora: Profa. Dra. Wendy Herrera.  
Monografia (Bacharelado). Universidade Federal de Ouro Preto.  
Instituto de Ciências Exatas e Aplicadas. Graduação em Engenharia  
Elétrica .

1. Indústria - Planta didática. 2. Indústrias - Processos - Medidas de  
segurança. 3. Modelos lineares (Estatística). 4. Modelos matemáticos. 5.  
Processos - Desempenho. I. Herrera, Wendy. II. Universidade Federal de  
Ouro Preto. III. Título.

CDU 519.8

Bibliotecário(a) Responsável: Flavia Reis - CRB6-2431



## FOLHA DE APROVAÇÃO

**Rafael Ferreira Félix**

### Identificação de modelos para uma planta didática de nível SMAR PD3-F

Monografia apresentada ao Curso de Engenharia Elétrica da Universidade Federal de Ouro Preto como requisito parcial para obtenção do título de bacharel em Engenharia Elétrica

Aprovada em 16 de dezembro de 2024

#### Membros da banca

Dra. Wendy Yadira Eras Herrera - Orientadora - Universidade Federal de Ouro Preto  
Dr. Renan Fernandes Bastos - Coordenador do COEE - Universidade Federal de Ouro Preto

Wendy Yadira Eras Herrera, orientadora do trabalho, aprovou a versão final e autorizou seu depósito na Biblioteca Digital de Trabalhos de Conclusão de Curso da UFOP em 20/12/2024



Documento assinado eletronicamente por **Wendy Yadira Eras Herrera, PROFESSOR DE MAGISTERIO SUPERIOR**, em 20/12/2024, às 09:32, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



A autenticidade deste documento pode ser conferida no site [http://sei.ufop.br/sei/controlador\\_externo.php?acao=documento\\_conferir&id\\_orgao\\_acesso\\_externo=0](http://sei.ufop.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0), informando o código verificador **0833317** e o código CRC **488C44D9**.

# Agradecimentos

Agradeço a todos os amigos, colegas e professores que contribuíram pela minha formação, me tornando um ser humano melhor e mais preparado para o meio de trabalho.

Agradeço de forma especial a minha orientadora Wendy Yadira, o professor Marcelo Moreira, aos meus colegas Marphiza e Victor pelo conhecimento e tempo fornecidos, tornando esse trabalho de conclusão de curso possível. Agradeço também a minha mãe, Vanuza Ferreira por me encorajar a correr atrás dos meus sonhos e por sempre me fornecer suporte e incentivos em momentos de dificuldade.

*"Perfection is not attainable, but if we chase perfection we can catch excellence."*

*– Vince Lombardi*

# Resumo

O setor industrial apresenta diversas variáveis de processos de extrema importância que devem ser medidas e monitoradas, visto que estas estão diretamente relacionadas ao desempenho e segurança de processos. Visando aproximar alunos e docentes dos processos industriais, a *SMAR Technology Company* e outras empresas desenvolveram plantas didáticas que possuem a capacidade de simular processos industriais permitindo o estudo do comportamento de algumas variáveis de processos, como nível, temperatura, pressão, entre outras. Por meio da modelagem matemática é possível obter modelos matemáticos que descrevam a dinâmica destas variáveis de processos, permitindo compreender e avaliar o comportamento da variável estudada. Neste trabalho, investiga-se o problema de modelagem matemática para descrever a dinâmica da variável de nível da planta didática SMAR PD3-F. A metodologia proposta utiliza técnicas de identificação de sistema para modelos do tipo caixa preta. Os ensaios foram realizados utilizando a planta didática SMAR PD3-F. A partir de uma dada configuração da planta didática são coletados dados referentes ao sinal de entrada, a vazão de entrada, e o sinal de saída, o nível no tanque de aquecimento. A partir de ensaios experimentais, são obtidos modelos lineares ARX e não lineares NARX para representar a variável de nível da planta. Os resultados obtidos sugerem que o modelo NARX apresenta o melhor desempenho em descrever a dinâmica da variável de nível da planta didática, alcançando valores do índice RMSE de 0,197, enquanto, o modelo ARX apresenta um desempenho muito próximo, possuindo o índice RMSE de 0,216. Dessa forma, conclui-se que o processo de identificação de sistemas é um método promissor para obter de modelos lineares e não lineares que descrevam a dinâmica das variáveis envolvidas nos processos industriais.

**Palavras-chave:** Modelagem matemática, Identificação de sistemas, Planta didática SMAR PD3-F, ARX, NARX.

# Abstract

The industrial sector involves several critical process variables that must be measured and monitored, as they are directly related to process performance and safety. To bring students and faculty closer to industrial processes, SMAR Technology Company and other companies have developed educational plants capable of simulating industrial processes, allowing the study of the behavior of certain process variables, such as level, temperature, and pressure. Through mathematical modeling, it is possible to obtain mathematical models that describe the dynamics of these process variables, allowing to understand and evaluate the behavior of the studied variable. This study investigates the mathematical modeling problem to describe the dynamics of the level variable in the SMAR PD3-F educational plant. The proposed methodology uses system identification techniques for black-box models. The experiments were conducted using the SMAR PD3-F educational plant. Based on a given configuration of the educational plant, data were collected regarding the input signal (input flow rate) and the output signal (level in the heating tank). From experimental tests, linear ARX and non-linear NARX models were obtained to represent the level variable of the plant. The results suggest that the NARX model performed best in describing the dynamics of the level variable in the educational plant, achieving an RMSE (Root Mean Square Error) of 0.197, while the ARX model showed very close performance with an RMSE of 0.216. Thus, it is concluded that system identification is a promising method for obtaining linear and non-linear models that describe the dynamics of variables involved in industrial processes.

**Keywords:** Mathematical modeling, Systems identification, Didactic plant SMAR PD3-F, ARX, NARX.



# Lista de ilustrações

Figura 1 – Planta didática SMAR e suas dimensões. . . . .	5
Figura 2 – Posicionador de válvula inteligente (FY-31). . . . .	6
Figura 3 – Transmissor de pressão utilizado para medir e transmitir valores de nível e vazão. . . . .	7
Figura 4 – Bomba hidráulica trifásica modelo HYDRO-BLOC P500T. . . . .	8
Figura 5 – IHM do inversor de Frequência CFW08. . . . .	8
Figura 6 – Esquemático da planta didática para a malha 31 mostrando os equipamentos utilizados e a direção do fluxo da água. . . . .	9
Figura 7 – Painel da planta didática SMAR PD3-F. . . . .	11
Figura 8 – Topologia da planta didática SMAR PD3-F. . . . .	13
Figura 9 – Fluxograma de comparação das etapas do método de identificação de sistemas para os modelos lineares e não lineares. . . . .	16
Figura 10 – Diagrama de comparação das etapas utilizadas no processo identificação de sistemas para os modelos lineares e não lineares. . . . .	29
Figura 11 – Diagrama de fluxo de água para o protocolo experimental que utiliza a válvula pneumática e o motor monofásico. . . . .	30
Figura 12 – Sinal de nível do tanque: (a) abertura de 30%, (b) abertura de 30%. . .	31
Figura 13 – Diagrama do fluxo de água para o protocolo experimental que utiliza o inversor de frequência e a bomba trifásica. . . . .	32
Figura 14 – Sinal de (a) referência do sinal seno, (b) vazão de entrada, (c) nível do tanque em cada ponto de operação 75% (vermelho), 60% (azul), 45% (verde) e 30% (magenta). . . . .	34
Figura 15 – Sinal de vazão de entrada para o ponto de operação de 45%. (a) Dados modelagem. (b) Dados Validação. . . . .	37
Figura 16 – Sinal da variável de nível para o ponto de operação de 45%. (a) Dados modelagem. (b) Dados Validação. . . . .	37
Figura 17 – Critério de Informação Akaike (AIC) para o ponto de operação de 45%. . . . .	38
Figura 18 – Comparação entre a simulação livre (vermelho) e os dados de validação (azul) para o modelo ARX (4.4) no ponto de operação de 45%. . . . .	39
Figura 19 – Comparação entre a simulação livre (vermelho) e os dados de validação (azul) para o modelo NARX (4.7) no ponto de operação de 45%. . . . .	41
Figura 20 – Comparação da simulação livre dos modelos: ARX (vermelho) e NARX (azul), e os dados de validação (preto) para o ponto de operação de 45%. . . . .	41
Figura 21 – Diagrama P&I - malhas de controle e dispositivos. . . . .	45

# Lista de Siglas

<b>AIC</b>	Critério de Informação Akaike
<b>AR</b>	Modelos AutoRegressivos (do inglês, <i>AutoRegressive</i> )
<b>ARMA</b>	Modelos AutoRegressivos com Média Móvel (do inglês, <i>AutoRegressive with Moving Average</i> )
<b>ARMAX</b>	Modelos AutoRegressivos com Média Móvel e entradas eXógenas (do inglês, <i>AutoRegressive Moving Average with eXogenous inputs</i> )
<b>ARX</b>	Modelos AutoRegressivos com entradas eXógenas (do inglês, <i>AutoRegressive with eXogenous inputs</i> )
<b>ERR</b>	Método da Taxa de Redução de Erro
<b>GH</b>	Golub-Householder
<b>ICEA</b>	Instituto de Ciência Exatas e Aplicadas
<b>IHM</b>	Interface Homem Máquina
<b>MQ</b>	Mínimos Quadrados
<b>MQO</b>	Mínimos Quadrados Ortogonal
<b>NARMAX</b>	Modelos Não lineares AutoRegressivos com Média Móvel e entradas eXógenas (do inglês, <i>Nonlinear AutoRegressive Moving Average with eXogenous inputs</i> )
<b>NARX</b>	Modelos Não lineares AutoRegressivos com entradas eXógenas (do inglês, <i>AutoRegressive with eXogenous inputs</i> )
<b>OPC</b>	Comunicação em Plataforma Aberta (do inglês, <i>Open Platform Communication</i> )
<b>PID</b>	Proporcional Integrador Derivativo
<b>PI</b>	Proporcional Integrador
<b>RMSE</b>	Raíz Quadrada do Erro Médio (do inglês, <i>Root Mean Squared Error</i> )

# Sumário

<b>1</b>	<b>INTRODUÇÃO</b>	<b>1</b>
<b>1.1</b>	<b>Motivação e Justificativa</b>	<b>1</b>
<b>1.2</b>	<b>Estado da Arte</b>	<b>3</b>
<b>1.3</b>	<b>Objetivos</b>	<b>4</b>
<b>1.4</b>	<b>Estrutura do trabalho</b>	<b>4</b>
<b>2</b>	<b>PLANTA DIDÁTICA SMAR</b>	<b>5</b>
<b>2.1</b>	<b>Descrição da planta didática SMAR</b>	<b>5</b>
<b>2.2</b>	<b>Equipamentos da Planta Didática SMAR PD3-F</b>	<b>6</b>
2.2.1	Posicionador Inteligente de Válvulas (FY)	6
2.2.2	Transmissor indicador de nível (LIT) e Transmissor indicador de vazão (FIT)	6
2.2.3	Bombas Hidráulicas	7
2.2.4	Inversor de Frequência	8
<b>2.3</b>	<b>Características da Planta Didática SMAR PD3-F</b>	<b>9</b>
2.3.1	Circuito Hidráulico	9
2.3.2	Circuito Pneumático	10
2.3.3	Circuito Elétrico	10
2.3.4	Sistema de Comunicação	11
<b>2.4</b>	<b>Considerações Parciais</b>	<b>12</b>
<b>3</b>	<b>METODOLOGIA</b>	<b>14</b>
<b>3.1</b>	<b>Identificação de Sistemas</b>	<b>14</b>
<b>3.2</b>	<b>Modelos Lineares</b>	<b>16</b>
3.2.1	Teste dinâmicos e coleta de dados	16
3.2.2	Escolha da representação matemática	17
3.2.3	Determinação da estrutura do modelo	17
3.2.4	Estimação de parâmetros	18
3.2.5	Validação de modelos	19
3.2.5.1	A Simulação Livre	19
3.2.5.2	O índice de erro médio quadrático (RMSE)	20
<b>3.3</b>	<b>Modelos Não Linear</b>	<b>20</b>
<b>3.4</b>	<b>Determinação da estrutura e estimação dos parâmetros</b>	<b>23</b>
3.4.1	O método de Golub-Householder (GH)	24
3.4.1.1	A transformação de Householder	25
<b>3.5</b>	<b>Taxa de Redução de Erro (ERR)</b>	<b>26</b>
<b>3.6</b>	<b>Algoritmo GH</b>	<b>27</b>

<b>3.7</b>	<b>Considerações Parciais</b> . . . . .	<b>28</b>
<b>4</b>	<b>RESULTADOS</b> . . . . .	<b>30</b>
<b>4.1</b>	<b>Motor Monofásico</b> . . . . .	<b>30</b>
<b>4.2</b>	<b>Motor Trifásico</b> . . . . .	<b>32</b>
4.2.1	Sinal seno . . . . .	33
4.2.2	Identificação de sistema para todos os pontos de operação . . . . .	35
4.2.3	Identificação de Sistema para o Ponto de Operação 45% . . . . .	36
4.2.3.1	Sistema Linear . . . . .	38
4.2.3.2	Sistema Não Linear . . . . .	39
4.2.4	Comparação entre os Modelos ARX e NARX para o Ponto de Operação de 45% . . . . .	41
<b>4.3</b>	<b>Considerações Parciais</b> . . . . .	<b>41</b>
<b>5</b>	<b>CONCLUSÕES E PROPOSTAS DE CONTINUIDADE</b> . . . . .	<b>43</b>
<b>5.1</b>	<b>Considerações Finais</b> . . . . .	<b>43</b>
<b>5.2</b>	<b>Propostas de Continuidade do TCC</b> . . . . .	<b>44</b>
<b>A</b>	<b>DIAGRAMA P&amp;I</b> . . . . .	<b>45</b>
	<b>REFERÊNCIAS</b> . . . . .	<b>46</b>

# 1 Introdução

## 1.1 Motivação e Justificativa

No cotidiano nos deparamos com vários tipos de sistemas estáticos e dinâmicos. Sistemas estáticos são aqueles onde as variáveis e propriedades do sistema não mudam ao longo do tempo, ou seja, o comportamento das variáveis de entrada e saída estão diretamente interligadas, onde mudanças nas variáveis de entrada geram mudanças imediatas nas variáveis de saída. Os sistemas dinâmicos são sistemas em que seu comportamento evolui no tempo, ou seja, as variáveis de entrada e saída são independentes até que em dado momento as variáveis atuais dependam de variáveis anteriores.

Buscando compreender o comportamento de sistemas dinâmicos, a modelagem matemática é utilizada como tentativa de descrever matematicamente estes sistemas. Assim, a partir das diversas técnicas utilizadas na literatura, como a modelagem caixa branca, a modelagem caixa cinza e a modelagem caixa preta (AGUIRRE, 2004), são obtidas representações lineares e não lineares.

A modelagem de caixa branca é empregada quando é viável obter o modelo do sistema diretamente das leis físicas ou das características do processo como, por exemplo, sistema massa-mola e circuitos RLC (MOMBELLO; PORTILLO; FLÓREZ, 2020), onde o comportamento do sistema pode ser descrito com precisão graças ao entendimento claro das interações físicas envolvidas. A modelagem caixa preta é utilizada quando não é possível obter um modelo analisando a física do sistema e tem-se disponível somente os sinais de entrada e saída do sistema, como, por exemplo, a modelagem caixa preta do sistema de água subglacial proposta por Murray e Clarke (1995). Na modelagem caixa cinza ocorre uma combinação das técnicas utilizadas nas modelagens caixa branca e caixa preta, como, por exemplo, a modelagem caixa cinza para um circuito RC proposto por Pereira e Munareto (2019).

A modelagem caixa preta é bastante utilizada na literatura por não ser necessária a compreensão da física do sistema abordado, tornando seu uso mais simples quando comparado com a modelagem caixa branca e cinza. Ao utilizar a modelagem caixa preta pode ser empregado o processo de identificação de sistemas para a obtenção de modelos lineares e não lineares. Ao lidar com modelos lineares, este processo possui cinco etapas: testes dinâmicos e coleta de dados, escolha da representação matemática, determinação da estrutura do modelo, estimação dos parâmetros e validação de modelos (AGUIRRE, 2004). Para os modelos não lineares as etapas de testes dinâmicos e coleta de dados, escolha da representação matemática e validação dos modelos permanecem as mesmas utilizadas para as representações lineares, porém durante as etapas de determinação da

estrutura do modelo e estimação dos parâmetros são empregados métodos diferentes dos utilizados para as representações lineares, tanto que dependendo dos métodos utilizados, estas etapas podem ocorrer de forma simultânea.

Desta forma, a modelagem matemática permite descrever vários sistemas dinâmicos ou processos industriais, cujo comportamento evolui ao longo do tempo. Visando aproximar estudantes e docentes aos processos industriais, empresas como a SMAR *technology company* desenvolvem plantas didáticas que possuem a capacidade de simular processos industriais que possibilitam o estudo do comportamento de variáveis destes processos industriais.

A planta didática SMAR PD3-F é um sistema dinâmico que apresenta várias malhas de controle, nas quais é possível monitorar as variáveis de nível, temperatura, pressão e vazão. Entretanto, com o desgaste dos equipamentos e instrumentos, além da falta de manutenção, as medições realizadas pelos instrumentos podem perder sua precisão, assim tornando a modelagem matemática uma opção para obter um modelo que represente a variável que deseja ser estudada, para compreender e validar seu comportamento.

Existem vários estudos na literatura utilizando as plantas didáticas como Lopes et al. (2012), VERLY et al. (2019), entre outros. Por exemplo, no estudo desenvolvido por Lopes et al. (2012) é utilizado o processo de identificação de sistemas dinâmico assimétrico baseando em redes neurais para obter modelos que representem a variável de nível de uma planta didática a partir do treinamento das redes neurais. Na investigação realizada por VERLY et al. (2019) é utilizada a modelagem caixa preta e cinza visando obter modelos lineares e não lineares que representem a variável de nível da planta didática. Assim, são utilizadas as representações Modelos AutoRegressivos com entradas eXógenas (do inglês, *AutoRegressive with eXogenous inputs*) (ARX) para representar modelos lineares e os modelos Modelos Não lineares AutoRegressivos com entradas eXógenas (do inglês, *AutoRegressive with eXogenous inputs*) (NARX) e de Hammerstein e de Wiener para representações não lineares.

Neste trabalho, investiga-se o problema de modelagem matemática para descrever o comportamento da variável de nível do tanque da planta didática SMAR PD3-F. A metodologia proposta consiste na utilização do processo de identificação baseado na modelagem caixa preta para obter modelos que caracterizem a variável de nível a partir de dados coletados na planta didática SMAR PD3-F localizada no laboratório de Controle e Automação do Instituto de Ciência Exatas e Aplicadas (ICEA) em João Monlevade.

O primeiro protocolo experimental considerado neste trabalho utiliza uma configuração da planta didática com a bomba monofásica e a válvula pneumática. A bomba monofásica é responsável por bombear a água presente no reservatório de água até o tanque de medição de nível e a abertura da válvula pneumática estabelece os pontos de operações desejados. Ao realizar alguns testes dinâmicos na configuração da planta didática supracitada, percebeu-se que os dados coletados possuíam medições diferentes da

variável de nível, sendo mantida a mesma configuração da planta didática. Dessa forma, constatou-se que houve uma diminuição na precisão dos instrumentos de medição devido ao desgaste dos equipamentos e por falta de manutenções regulares.

Buscando contornar o problema de precisão dos equipamentos, foi considerado um novo protocolo experimental que não utilize a válvula pneumática. A nova configuração utiliza o inversor de frequência junto com a bomba trifásica para bombear a água do reservatório de água para o tanque de aquecimento. Para este protocolo experimental foram escolhidos 4 pontos de operação de 75%, 60%, 45% e 30% da variável de nível do tanque de aquecimento, visando trabalhar em diferentes faixas de operação que permitam analisar a dinâmica da variável de nível entre o tanque parcialmente vazio e parcialmente cheio. Foram considerados sinais de referência do tipo senoidal para cada ponto de operação, variando somente a amplitude do sinal senoidal de um ponto de operação para outro. Assim, a partir do processo de identificação de sistemas e dos sinais de entrada,  $u_k$ , e saída,  $y_k$ , foram escolhidas representações lineares ARX e não lineares NARX para obter modelos matemáticos que descrevam a variável de nível. Para validar os modelos identificados são utilizadas duas métricas: a simulação livre e o índice Raíz Quadrada do Erro Médio (do inglês, *Root Mean Squared Error*) (RMSE).

## 1.2 Estado da Arte

VERLY et al. (2019) utilizam o processo de identificação de modelos da planta didática SMAR PD3-F para descrever a variável de nível. Nesse trabalho foram utilizadas técnicas de modelagens do tipo caixa preta e caixa cinza para sistemas dinâmicos lineares e não lineares. Para a escolha da representação matemática são escolhidos os modelos ARX, NARX e o modelo *Hammerstein* e *Wiener*. Os modelos do tipo ARX são utilizados para representar a parcela linear dos modelos, e os modelos NARX conseguem representar os modelos em sua parcela não linear. Durante a determinação da estrutura dos modelos produzidos, os autores explicitam a importância da escolha dos polos, zeros e do atraso puro de tempo. Dessa forma, para determinar o número de termos candidatos, são utilizados dois métodos o AIC e a Método da Taxa de Redução de Erro (ERR). Com a classificação dos termos candidatos significativos para a obtenção de cada modelo, os parâmetros que os acompanham são obtidos utilizando algoritmos derivados do estimador de Mínimos Quadrados (MQ). Por fim, a validação dos modelos encontrados mostram que os modelos não lineares apresentaram uma melhor resposta que o modelo linear.

Lopes et al. (2012) emprega o processo de identificação de sistemas para uma planta didática SMAR utilizando redes neurais artificiais por meio de uma arquitetura *feed-forward*, realizando o treinamento da rede neural de forma que seja possível modelar um sistema de controle de nível da planta didática. Para realizar o monitoramento e envio de comandos para a planta didática, os autores utilizam o software *Matlab* em

conjunto com o protocolo Comunicação em Plataforma Aberta (do inglês, *Open Platform Communication*) (OPC), onde são aplicadas entradas em degrau ao modelo neural no *Matlab* visando observar a resposta da planta didática relacionada a variável de nível.

### 1.3 Objetivos

O objetivo deste trabalho é investigar o problema de modelagem matemática da variável de nível do tanque de uma planta didática de nível SMAR PD3-F. Para atingir o presente objetivo, foram definidos os seguintes objetivos específicos:

1. Estudar o funcionamento da planta didática de nível SMAR PD3-F.
2. Obter modelos matemáticos lineares e não lineares que caracterizam a dinâmica da variável de nível.
3. Aplicar duas métricas, a simulação livre e o índice RMSE, para quantificar a qualidade dos modelos matemáticos obtidos que descrevem o comportamento da variável de nível da planta.

### 1.4 Estrutura do trabalho

O presente trabalho encontra-se dividido em 5 capítulos organizados a seguir.

No Capítulo 1, aborda-se a introdução do tema de estudo descrevendo a motivação e a justificativa e o estado da arte, abrangendo o problema de modelagem matemática da variável de nível da planta didática SMAR PD3-F e os objetivos principais deste trabalho.

No Capítulo 2 realiza-se uma descrição da planta didática SMAR PD3-F apresentando as características da planta, instrumentos, sistemas de comunicações visando compreender o comportamento e funcionamento da planta.

O Capítulo 3 descreve a metodologia empregada neste trabalho. Dessa forma, é apresentado o processo de identificação de sistemas baseado na modelagem caixa preta utilizada neste trabalho para a obtenção de modelos lineares e não lineares.

No Capítulo 4, apresentam-se os protocolos experimentais e os resultados obtidos nas etapas de modelagem matemática da variável de nível da planta didática SMAR PD3-F.

Por fim, no Capítulo 5, têm-se as considerações finais deste trabalho e sugestões de trabalhos futuros.



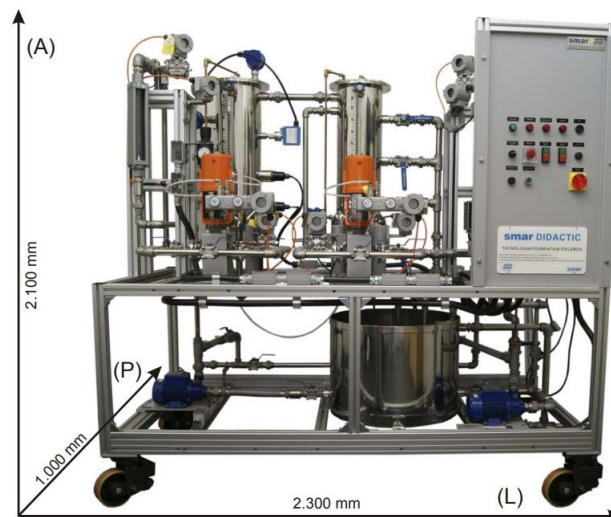
## 2 Planta Didática SMAR

Nas seções desse capítulo será apresentada a planta didática SMAR na totalidade, descrevendo sua composição em *hardware* e *software*. O *hardware* da planta apresenta três tipos de ligações: hidráulica, pneumática e elétrica. Por meio destas ligações é descrito o funcionamento dos equipamentos e instrumentos, além do comportamento de fluxo da água na planta. A composição de *software* da planta demonstra como a rede de comunicação *FOUNDATION Fieldbus*, o *software System302*, o protocolo OPC e o *software Matlab* estabelecem uma comunicação possibilitando realizar a comunicação com a planta didática e seus equipamentos permitindo fornecer comandos e receber os dados coletados.

### 2.1 Descrição da planta didática SMAR

A planta didática SMAR PD3-F fabricada pela SMAR *Technology Company* mostrada na Figura 1 foi desenvolvida para simular o comportamento de variáveis de pressão, temperatura e nível em processos industriais (SMAR, 2015). Essa planta apresenta um sistema compacto que possui processos de controle e supervisão, possibilitando uma reprodução fiel das características encontradas no campo profissional.

Figura 1 – Planta didática SMAR e suas dimensões.



Fonte: Retirado de SMAR (2015).

A partir de uma combinação específica das válvulas manuais e pneumáticas, a planta didática PD3-F possui duas malhas de controle em sua composição, a malha de controle 31 e 32. A malha de controle 31 é responsável por controlar a variável de nível e a temperatura do tanque de aquecimento enquanto que a malha de controle 32 é utilizada

para realizar a mistura de fluidos no tanque de mistura. Além das malhas de controle previamente fornecidas pelo fabricante, é possível criar novas malhas de controle realizando uma reconfiguração dos dispositivos presentes na planta conforme os objetivos desejados.

## 2.2 Equipamentos da Planta Didática SMAR PD3-F

A planta didática em sua composição apresenta equipamentos que possuem funções específicas para seu funcionamento. Assim, visando compreender como a planta didática SMAR PD3-F trabalha é necessário conhecer os equipamentos que a compõem. Nesta seção, são apresentados os equipamentos utilizados neste trabalho.

### 2.2.1 Posicionador Inteligente de Válvulas (FY)

Os posicionadores inteligentes de válvulas (FY) são válvulas pneumáticas utilizadas na planta didática. Estes equipamentos são utilizados na automação de sistemas industriais para controlar a posição e o movimento de válvulas, sendo projetados para receber um sinal de controle elétrico ou pneumático e converter esse sinal em movimento físico (SMAR, 2015). Esse tipo de posicionador geralmente possui recursos avançados de controle e monitoramento, dentre estes destaca-se o monitoramento remoto do equipamento que permite que os dados do posicionador possam ser acessados e analisados a partir de uma localização remota, facilitando a supervisão e o gerenciamento de sistemas.

Figura 2 – Posicionador de válvula inteligente (FY-31).



Fonte: SMAR *Technology Company*.

### 2.2.2 Transmissor indicador de nível (LIT) e Transmissor indicador de vazão (FIT)

O transmissor indicador de nível (LIT) é um dispositivo utilizado para realizar a medição do nível de líquidos ou sólidos em tanques, reservatórios, silos e outros recipientes

enquanto o transmissor indicador de vazão (FIT) é utilizado para medir o fluxo de líquidos, gases ou vapor em um determinado sistema. Esses equipamentos desempenham um papel crucial em muitos setores industriais, como petróleo e gás, química, alimentos e bebidas, tratamento de água e muitos outros, onde a medição precisa do nível e vazão dos fluidos é essencial. A planta didática conta com transmissores de pressão do tipo LD302D para realizar a medição e transmissão dos valores de nível e vazão, conforme mostrado na Figura 3.

Ambos equipamentos LIT e FIT são compostos por um indicador e um transmissor. O indicador de nível é uma parte visível do equipamento que permite verificar visualmente os valores medidos para o nível e a vazão no tanque. O transmissor de nível é responsável por converter os valores obtidos em um sinal elétrico padronizado, como 4-20 mA, que pode ser facilmente interpretado e utilizado por sistemas de controle e monitoramento (SMAR, 2015). Vale ressaltar que a planta didática SMAR PD3-F possui apenas a medição de nível no tanque de aquecimento e para ambos os tanques, somente é realizada a medição da vazão de entrada.

Figura 3 – Transmissor de pressão utilizado para medir e transmitir valores de nível e vazão.



Fonte: SMAR *Technology Company*.

### 2.2.3 Bombas Hidráulicas

As bombas hidráulicas são responsáveis por realizar a circulação de água pela planta. Para a realização desse processo a planta conta com uma bomba monofásica (modelo HYDRO-BLOC P500) e uma bomba trifásica (modelo HYDRO-BLOC P500T). Ambas as bombas presentes no sistema conseguem bombear água para qualquer parte do sistema, porém é necessário customizar o sistema de forma que o fluxo de líquido vá para o tanque desejado. Quando é realizada uma comparação entre as bombas, a Figura 4 mostra uma bomba trifásica que possui uma ampla gama de funcionamento, visto que esta pode ser ajustada para trabalhar com diferentes frequências com o auxílio do inversor de frequência da planta didática (KRISHNAN, 2001).

Figura 4 – Bomba hidráulica trifásica modelo HYDRO-BLOC P500T.



Fonte: SMAR *Technology Company*.

#### 2.2.4 Inversor de Frequência

O inversor de frequência é um dispositivo eletrônico projetado para controlar a velocidade de motores elétricos de corrente alternada. Eles desempenham um papel fundamental em otimizar o consumo de energia, melhorar a eficiência operacional e permitir o controle preciso da velocidade em várias aplicações industriais (KRISHNAN, 2001). A planta didática SMAR PD3-F conta com o inversor da fabricante WEG do tipo CFW08, o qual possui uma Interface Homem Máquina (IHM) utilizada para ajustar seus parâmetros, conforme mostrado na Figura 5.

Figura 5 – IHM do inversor de Frequência CFW08.



Fonte: Do Autor.

## 2.3 Características da Planta Didática SMAR PD3-F

A planta didática SMAR PD3-F possui três circuitos: pneumático, elétrico e hidráulico que conectam os equipamentos do sistema. Além disso, uma rede de comunicação é necessária para realizar a comunicação entre os diversos equipamentos e instrumentos na planta. A seguir apresentam-se essas características.

### 2.3.1 Circuito Hidráulico

Na Figura 6 mostra-se o circuito hidráulico composto por equipamentos responsáveis por direcionar o fluxo de água na planta, veja Tabela 1. A partir da malha de controle escolhida o fluxo de água se difere, porém, em todas as malhas de controle o fluxo de água se inicia ao sair do reservatório de água (RA) que tem a água bombeada pela bomba monofásica (BM) ou pela bomba trifásica (BT). Para controlar a bomba trifásica é utilizado o inversor de frequência (IF), este que possibilita controlar o modo de operação da bomba resultando em diferentes vazões de entrada, ou seja, diferentes quantidades de água bombeada do reservatório de água (RA) para o sistema. Dessa forma, a partir das bombas presentes na planta é possível direcionar o fluxo de água de acordo com uma combinação de válvulas manuais (VM) e pneumáticas (FY) para o tanque de aquecimento (T1) ou tanque de mistura (T2) conforme o objetivo desejado. O tanque de aquecimento (T1) consegue aquecer a água e realizar o controle de nível, já o tanque de mistura (T2) é utilizado para realizar a mistura de água proveniente do tanque de aquecimento (T1). No apêndice A é representado o diagrama P&I da planta didática SMAR PD3-F.

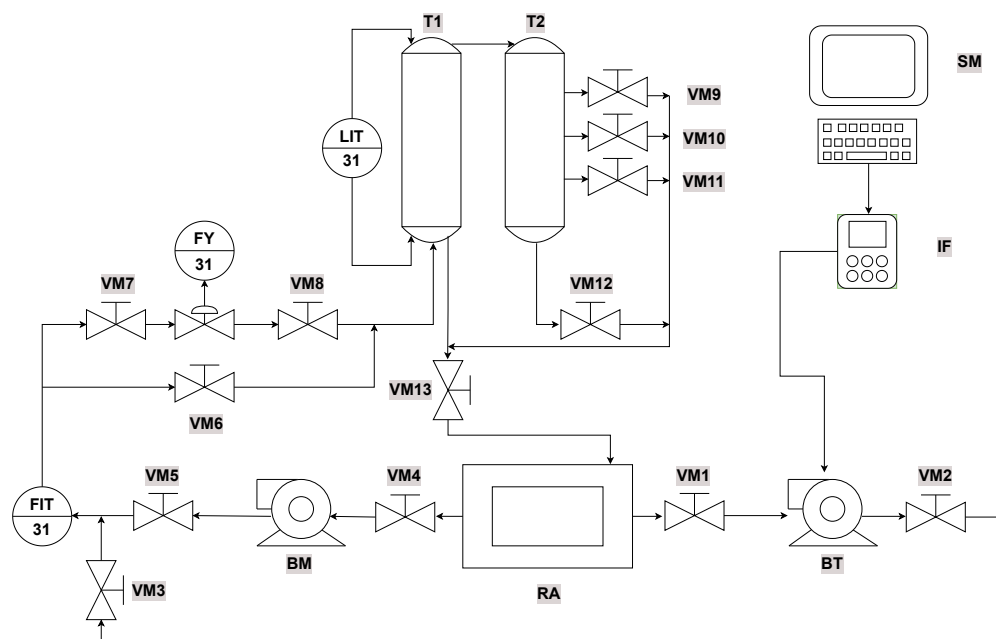


Figura 6 – Esquemático da planta didática para a malha 31 mostrando os equipamentos utilizados e a direção do fluxo da água.

Tabela 1 – Sigla dos componentes presentes na planta didática.

<b>Sigla</b>	<b>Equipamento</b>
VM <i>i</i>	Válvulas Manuais para $i = 1$ até 13.
FY	Válvula Pneumática
BM	Bomba Monofásica
BT	Bomba Trifásica
RA	Reservatório de Água
T1	Tanque De Aquecimento
T2	Tanque De Mistura
FIT	Transmissor Indicador De Vazão
LIT	Transmissor Indicador De Nível
IF	Inversor De Frequência
SM	Software MATLAB

Para a medição da vazão de entrada, utiliza-se o transmissor indicador de vazão de entrada (FIT). Para realizar a medição do nível no tanque de aquecimento (T1) é utilizado o transmissor indicador de nível (LIT). Este instrumento utiliza da relação entre a diferença de pressão, possibilitando encontrar o nível no tanque pela seguinte expressão

$$\Delta p = p_1 - p_2 = h\rho g, \quad (2.1)$$

em que  $\Delta p$  indica a diferença de pressão no tanque,  $p_1$  é a pressão exercida na parte inferior do tanque,  $p_2$  é a pressão na parte superior do tanque,  $h$  é o nível do tanque,  $\rho$  é a densidade do fluido e  $g$  é a aceleração gravitacional. Ao isolar o nível  $h$ , tem-se

$$h = \frac{\Delta p}{\rho g}. \quad (2.2)$$

### 2.3.2 Circuito Pneumático

O circuito pneumático utilizam a rede de ar comprimido da planta didática para comutar as válvulas pneumáticas a partir dos posicionadores inteligente de válvulas (FY), possibilitando abrir ou fechar as válvulas pneumáticas conforme a abertura desejada. A alimentação do conjunto de válvulas pneumáticas/posicionador inteligente deve apresentar uma pressão entre 4,0 bar e no máximo 7,0 visto que pressões abaixo desses valores comprometem o funcionamento do conjunto e valores acima podem danificar o equipamento (SMAR, 2015).

### 2.3.3 Circuito Elétrico

O circuito elétrico é responsável pela alimentação de todos os equipamentos elétricos da planta a partir do painel de controle. Na Figura 7 mostra-se o painel de controle que permite energizar a planta didática e seus equipamentos, monitorar o estado e realizar acionamentos manuais de alguns equipamentos.

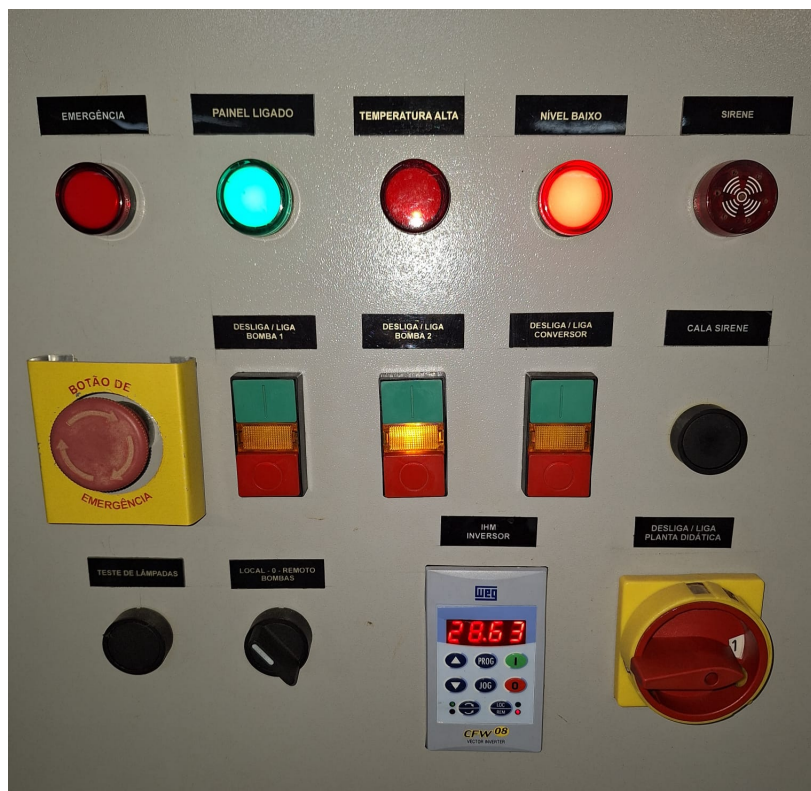


Figura 7 – Painel da planta didática SMAR PD3-F.

No painel se encontra quatro estados: emergência, painel ligado, temperatura alta e nível baixo. Ao ligar a planta pela botoeira de duas posições, a indicação visual de painel ligado irá acender. O estado de temperatura alta está relacionado a temperatura no tanque de aquecimento, visto que este tanque conta com uma resistência utilizada para aquecer a água neste tanque. O indicativo de nível baixo acende quando a quantidade de água presente no tanque de aquecimento está baixa, simultaneamente a sirene é acionada, sendo necessário aumentar a quantidade de água no tanque de aquecimento ou utilizar o botão calar sirene para desabilitar o aviso sonoro. Para validar se todos os leds de avisos visuais estão funcionando, o painel conta com o botão teste de lâmpadas, o qual aciona todas as indicações visuais. Além disso, a planta possibilita alterar o acionamento das bombas para que estas possam ser acionadas localmente pelas botoeiras do painel ou de forma remota. Também, é possível acionar o conversor e o inversor de frequência.

### 2.3.4 Sistema de Comunicação

Para possibilitar a comunicação dos equipamentos presentes na planta de forma que estes trabalhem em sincronia é necessário um sistema de comunicação. Os sistemas de comunicação são um conjunto de tecnologias e equipamentos utilizados para facilitar a comunicação e troca de informações entre diferentes partes de uma planta industrial (VITTURI; ZUNINO; SAUTER, 2019). Esses sistemas são essenciais para o controle e monitoramento dos processos industriais, permitindo que os operadores possam visualizar

e analisar dados em tempo real, detectar problemas e tomar decisões rápidas e precisas para otimizar a produção (KUROSE; ROSS, 2007).

A planta didática SMAR utiliza a rede *FOUNDATION Fieldbus* para realizar a comunicação entre os dispositivos de campo, interconectando equipamentos, instrumentos e outros dispositivos eletrônicos estes apresentando sinais de entrada/saída compartilhando o mesmo barramento de comunicação.

O protocolo *FOUNDATION Fieldbus* foi desenvolvido para realizar a integração de diferentes tipos de equipamentos e sistemas, permitindo uma maior flexibilidade e adaptabilidade na configuração e operação dos processos. Além disso, o protocolo possui uma estrutura aberta e padronizada, o que significa que diferentes fabricantes podem produzir equipamentos compatíveis com o protocolo, facilitando a interoperabilidade entre os sistemas (VERHAPPEN; PEREIRA, 2008).

Na Figura 8 mostra-se a topologia de rede da planta didática utilizando a rede *FOUNDATION Fieldbus*. A partir da topologia de rede é possível perceber como estão conectados os elementos que compõem a planta.

O switch presente na planta didática realizada a interligação do computador, com o rack da plataforma de controle e automação de processos por meio da rede Ethernet. A plataforma de controle e automação de processos possibilita a comunicação e aquisição de dados do processo com instrumentos e atuadores que utilizam o protocolo *FOUNDATION Fieldbus* e PROFIBUS. A plataforma de controle é composta pelo rack DFI-1 e DFI-2, sendo utilizado neste trabalho o rack DFI-2 onde destacasse o módulo DF63 utilizado como controlador multifuncional com funções *gateway* entre redes *FOUNDATION Fieldbus* HSE independentes responsáveis por realizar a comunicação com o computador utilizando cabos de rede Ethernet. As redes *FOUNDATION Fieldbus* H1 recebem/enviam dados para os instrumentos de campo pelo terminador de barramento Fieldbus (BT302) da planta que utiliza cabo de par trançado blindado projetado para suportar tanto a comunicação de dados quanto a alimentação dos dispositivos em campo na mesma fiação. Já o cabo FT-4 é utilizado para realizar a ligação do conversor tiristorizado (TY-31) onde uma resistência básica à propagação de chamas é exigida (SMAR, 2015).

## 2.4 Considerações Parciais

Neste capítulo, foram apresentados os equipamentos utilizados na planta didática SMAR PD3-F, as características dos seus circuitos (pneumático, elétrico e hidráulico) e o sistema de comunicação da planta. A partir dessas informações, é possível compreender como os equipamentos e instrumentos da planta didática se comportam e como o sistema de comunicação se interliga por meio do switch e da plataforma de controle e automação de processos. Assim, ao compreender a planta didática na totalidade é possível entender de forma mais clara a utilidade de cada equipamento para realizar o monitoramento e



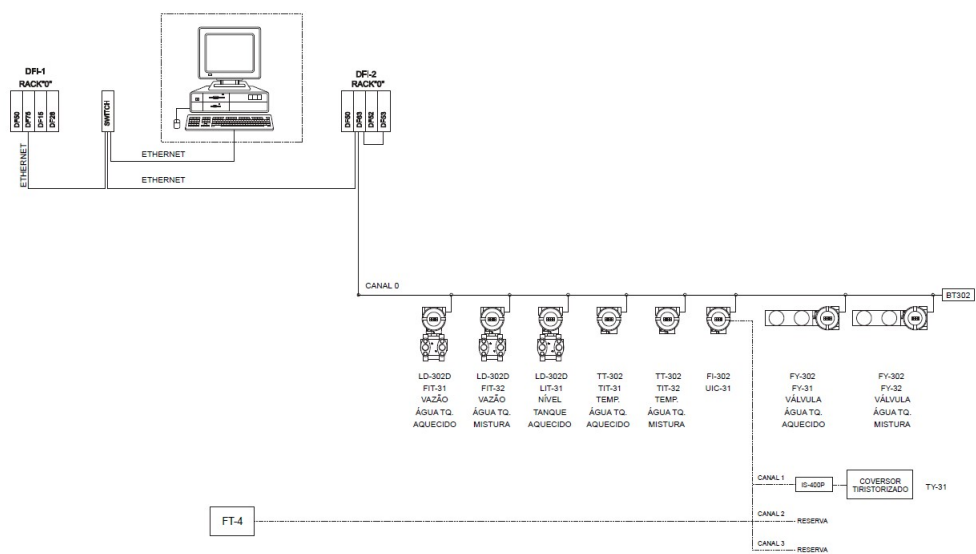


Figura 8 – Topologia da planta didática SMAR PD3-F.

estudo da variável de nível.

## 3 Metodologia

Neste capítulo, na seção 3.1 é apresentada as cinco etapas do processo de identificação de sistemas. A partir deste processo de identificação de sistemas são mostrados os métodos utilizados para obter o modelo linear (seção 3.2) e modelo não linear (seção 3.3) que caracterizem o nível de água no tanque de aquecimento da planta didática SMAR.

### 3.1 Identificação de Sistemas

Os modelos matemáticos podem ser obtidos por meio diferentes técnicas presentes na literatura, neste trabalho foi selecionada a modelagem caixa preta utilizando dados de entrada e saída coletados na planta didática SMAR. Para a obtenção de modelos matemáticos é utilizado o processo de identificação de sistemas descrito em Aguirre (2004). Este procedimento consiste em cinco etapas para se obter um modelo matemático de um sistema dinâmico. A seguir, descreve-se o processo de identificação de sistemas.

#### 1. Testes dinâmicos e coleta de dados

A primeira etapa consiste na obtenção de dados do sistema. Esses dados são obtidos a partir da operação padrão do sistema, e de testes realizados para observar o comportamento do sistema de forma dinâmica. Com isso é necessário saber como escolher e coletar os sinais. A escolha dos dados de entrada e saída considera a relação dinâmica das variáveis presentes no sistema, de forma que estas possuam uma correlação significativa, assim justificando o modelo produzido.

#### 2. Escolha da representação matemática a ser usada

A escolha da representação matemática depende da aplicação, pois se deve atentar aos dados disponíveis para representar os modelos e os padrões estipulados pelo projetista para representar o sistema abordado. Dessa forma, é necessário compreender que existem dois tipos de modelos: os modelos lineares e modelos não lineares. As representações mais conhecidas e utilizadas na literatura para os modelos lineares e não lineares são os modelos que fazem o uso de dados de entrada e saída para modelar o sistema.

#### 3. Determinação da estrutura do modelo

A determinação da estrutura do modelo linear consiste em escolher a quantidade de polos e zeros, e o atraso puro. A determinação dos polos e zeros de modelos lineares se dá pela escolha da ordem do modelo. A ordem do modelo é de suma importância, visto que é responsável por expressar a melhor representação do modelo. Caso a

escolha da ordem seja menor que a ordem efetiva do sistema físico, a representação do sistema não caracteriza o sistema. Com a escolha de uma ordem que seja muito maior que a do sistema físico, a estimação dos parâmetros serão mal condicionados, de forma a representar a dinâmica do sistema de forma errônea. Para determinar a ordem dos modelos lineares são utilizados métodos que fornecem uma estimativa da quantidade de termos que devem compor um determinado modelo, assim, a partir dessa estimativa tem-se a ordem que o modelo deve possuir. Apesar da indicação fornecida por estes métodos, deve-se validar por meio de testes se a ordem fornecida proporciona a melhor representação. Para os modelos não lineares, a obtenção da estrutura do modelo se torna uma tarefa um pouco mais complexa, visto que com o aumento do grau de não linearidade do modelo e dos polos e zeros do modelo, a quantidade de termos candidatos para compor o modelo aumenta consideravelmente, sendo assim necessário utilizar técnicas para selecionar os termos mais significativos para serem utilizados no modelo.

#### 4. Estimação de parâmetros

A estimação de parâmetros consiste em obter os parâmetros do modelo que representam o sistema. Por exemplo, no modelo a seguir, os parâmetros  $\theta_1$  e  $\theta_2$  são os parâmetros a serem obtidos.

$$y_k = \theta_1 y_{k-1} + \theta_2 u_{k-1}. \quad (3.1)$$

#### 5. Validação do modelo

Após a obtenção do modelo identificado, é necessário realizar a validação deste modelo. Nesta etapa, verifica-se se o modelo identificado descreve o comportamento do sistema dinâmico real. Durante a etapa de validação alguns cuidados devem ser observados para a validação ser realizada de forma efetiva. A característica necessária para que o modelo seja validado e eficiente, é que este seja generalizado, de forma que apresente o mesmo comportamento para diferentes dados. Com isso, durante a validação os dados utilizados devem ser diferentes dos usados para a obtenção do modelo.

Ao utilizar o processo de identificação de modelos para os sistemas lineares e não lineares, percebe-se que estes apresentam algumas semelhanças e diferenças em algumas etapas, conforme mostrado na Figura 9. Os modelos lineares e não lineares são similares nas etapas de testes dinâmicos e coletas de dados, na escolha da representação e na validação do modelo. Porém, durante as etapas de determinação da estrutura do modelo e estimação dos parâmetros encontra-se a diferença entre estes modelos. Para os modelos lineares em cada uma das etapas é utilizado um método específico, enquanto para os modelos não lineares essas etapas são realizadas simultaneamente utilizando uma combinação de métodos.

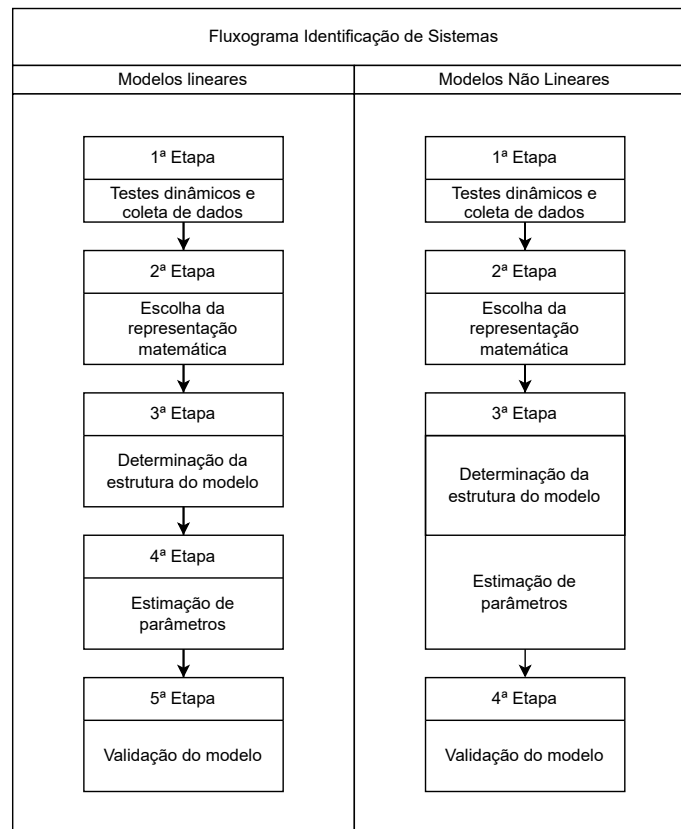


Figura 9 – Fluxograma de comparação das etapas do método de identificação de sistemas para os modelos lineares e não lineares.

## 3.2 Modelos Lineares

Os modelos lineares são uma classe de modelos amplamente utilizados para descrever a relação entre variáveis de forma linear. Eles são fundamentados na suposição de que a relação entre as variáveis pode ser aproximada por uma combinação linear das variáveis independentes (AGUIRRE, 2004).

Em um modelo linear, a variável dependente é expressa como uma combinação linear das variáveis independentes, ponderadas por coeficientes. O objetivo é estimar esses coeficientes de forma que o modelo se ajuste aos dados observados da melhor maneira possível. A seguir será apresentada a metodologia utilizada na identificação de sistema para modelos lineares, evidenciando os métodos utilizados em cada uma das cinco etapas.

### 3.2.1 Teste dinâmicos e coleta de dados

A partir do conhecimento do fluxo de água da planta didática apresentado no Capítulo 2, o processo de testes dinâmicos e coleta de dados foi realizado observando o comportamento da vazão de entrada como variável de entrada, e o nível no tanque de

aquecimento como variável de saída. Para os testes realizados foram desenvolvidos protocolos experimentais utilizando diferentes configurações, sendo empregada uma bomba monofásica e uma bomba trifásica em cada configuração.

Durante a etapa de testes dinâmicos e coleta de dados é necessário escolher os sinais de entrada utilizados para excitar um sistema, visto que esse sinais devem possuir a capacidade de excitar a dinâmica do sistema. Assim, para este trabalho é utilizado o sinal do tipo seno.

O sinal seno foi escolhido por ser um sinal de referência bastante utilizado para avaliar a estabilidade de sistemas e estimar parâmetros de sistemas (LJUNG, 1998). Esse sinal de referência seno é dado por

$$y_k = V_p \sin(2\pi f_0 dt) + \text{offset}, \quad (3.2)$$

em que  $V_p$  é o valor de pico,  $f_0$  é a frequência,  $dt$  é o período de amostragem e  $\text{offset}$  é o valor de  $\text{offset}$  adicionado ao sinal.

### 3.2.2 Escolha da representação matemática

A modelagem matemática da planta de nível SMAR PD3-F é pouco tratada na literatura. Ao utilizar estruturas de representação linear a literatura apresenta alguns modelos polinomiais como Modelos AutoRegressivos (do inglês, *AutoRegressive*) (AR), ARX, Modelos AutoRegressivos com Média Móvel (do inglês, *AutoRegressive with Moving Average*) (ARMA) e Modelos AutoRegressivos com Média Móvel e entradas eXógenas (do inglês, *AutoRegressive Moving Average with eXogenous inputs*) (ARMAX). Dentre esses modelos, utiliza-se o modelo ARX para caracterizar a variável de nível na planta didática SMAR PD3-F, visto que este modelo é bastante utilizado em trabalhos acadêmicos por fornecer uma boa compreensão do comportamento da variável observada (DELCROIX et al., 2021). O modelo ARX é descrito por

$$y_k = a_1 y_{k-1} + \dots + a_p y_{k-p} + b_1 u_{k-1} + \dots + b_q u_{k-q} + e_k, \quad (3.3)$$

em que  $y_k$  é o sinal de saída,  $u_{k-n}$  é o sinal de entrada,  $\theta^T = [a_1 \dots a_p \ b_1 \dots b_p]$  é o vetor de parâmetros,  $p$  e  $q$  são, respectivamente, a ordem dos termos autorregressivos da saída e entrada,  $e_k$  é o erro de modelagem.

### 3.2.3 Determinação da estrutura do modelo

A ordem de um modelo está relacionada ao número de parâmetros dos sinais de entrada e saída utilizados para descrever o comportamento do sistema. Existem vários métodos para determinar o número de termos de um modelo, como AIC (AKAIKE, 1974), o critério de Rissanen (STEINBERG; GASSER; FRANKE, 1985), o critério de informação Bayesiano (MEHRA, 1972), entre outros. Neste trabalho, utiliza-se o AIC

para determinar a ordem do modelo da planta de nível SMAR PD3-F, este método foi escolhido por sua simplicidade e por comumente ser utilizado no processo de seleção de ordem de modelos lineares (AGUIRRE, 2004).

O AIC é uma medida de avaliação de modelos estatísticos que consideram a qualidade do ajuste do modelo e sua complexidade. Foi proposto Akaike (1974) e é amplamente utilizado em ciência de dados, econometria e outras áreas relacionadas à modelagem estatística. O AIC indica a singularidade gerada quando se aumenta a ordem de um sistema, ou seja, representa a variação dos resíduos devido à inclusão de novos parâmetros. O AIC pode ser representado por

$$AIC(n_\theta) = N \ln[\sigma_{erro}^2(n_\theta)] + 2n_\theta, \quad (3.4)$$

em que  $N$  é o número de dados,  $\sigma_{erro}^2(n_\theta)$  é a variância dos resíduos e  $n_\theta$  é o número de parâmetros presentes no modelo. O erro de polarização ou erro sistemático do sistema ( $\sigma_{erro}^2(n_\theta)$ ) diminui a medida que  $n_\theta$  aumenta. Percebe-se que a partir de um certo ponto a adição de mais termos não irá influenciar na diminuição da variação dos resíduos, assim é necessário observar as duas parcelas presentes na equação (3.4). A primeira parcela indica a diminuição na variância dos resíduos devido à inclusão de um novo termo e a segunda parcela indica o custo da inclusão de cada termo. Assim, quando a adição de um novo termo apresentar uma diminuição na variação dos resíduos menor que a penalização por incluir este termo, chega-se a conclusão que este termo não deve estar presente no modelo.

### 3.2.4 Estimação de parâmetros

Para a estimação dos parâmetros é utilizado o método MQ que minimiza a soma do quadrado das diferenças entre os dados e a saída estimada pelo modelo (AGUIRRE, 2004). A estimação dos parâmetros utilizando MQ é dada por

$$y = \Psi\hat{\theta} + \xi, \quad (3.5)$$

em que  $\hat{\theta}$  é um vetor de parâmetros estimados,  $\Psi$  é a matriz de regressores,  $y$  é o vetor de medições. O vetor  $\xi$  representa os erros gerados quando aplicamos  $y$  como  $\Psi\hat{\theta}$ , ou seja, são os resíduos.

Buscando uma representação de  $\hat{\theta}$  que consiga reduzir  $\xi$  foi definido o somatório do quadrado dos erros da seguinte forma:

$$J_{MQ} = \sum_{i=1}^N \zeta(i)^2 = (y - \Psi\hat{\theta})^T (y - \Psi\hat{\theta}), \quad (3.6)$$

A equação (3.6) representa a função de custo que indica o índice que possibilita quantificar a qualidade do ajuste de  $\Psi\hat{\theta}$  em relação ao dados do vetor  $y$ . Visando encontrar o valor de  $\hat{\theta}$  que possibilita minimizar a função custo (3.6) tem-se

$$\hat{\theta}_{MQ} = (\Psi^T \Psi)^{-1} \Psi^T y, \quad (3.7)$$

em que  $(\Psi^T \Psi)^{-1} \Psi^T$  representa a matriz pseudoinversa. A matriz pseudoinversa é utilizada em sistemas sobredeterminados, ou seja, em situações que a  $\Psi$  não é quadrada. Dessa forma, a partir de (3.7) é possível encontrar  $\hat{\theta}$  mesmo que ao utilizar uma matriz  $\Psi$  não quadrada.

### 3.2.5 Validação de modelos

Finalizando o processo de identificação de sistemas, a última etapa consiste na validação do modelo identificado. A partir da etapa de validação, é possível verificar se o modelo identificado consegue descrever o comportamento da variável medida. Nesse trabalho, são utilizadas duas métricas: a simulação livre e o índice RMSE para quantificar a qualidade do modelo obtido. Essas duas métricas de validação são discutidas a seguir.

#### 3.2.5.1 A Simulação Livre

O processo de simulação livre (ou predição de infinitos passos a frente) é um modelo de validação que utiliza predições passadas para compor o vetor de regressores de forma que seja possível continuar realizando o processo de predição. Dessa forma, a baixo apresenta-se um exemplo retirado de (AGUIRRE, 2004).

**Example 1.** Supondo que um determinado modelo tenha o seguinte vetor de regressores

$$\psi_{yu,k-1}^Y = \begin{bmatrix} y_{k-1} & y_{k-2} & u_{k-1} & u_{k-2} \end{bmatrix}, \quad (3.8)$$

Para esse tipo de simulação, sempre é necessário iniciar o modelo com o uso de valores medidos, assim consideramos que

$$\begin{aligned} y_1 &= y_{k-1}, \\ y_2 &= y_{k-2}, \\ u_1 &= u_{k-1}, \\ u_2 &= u_{k-2}. \end{aligned}$$

Assim, é possível representar a predição do próximo passo

$$\hat{y}_3 = \psi_{yu,k-1}^Y \hat{\theta} = \begin{bmatrix} y_2 & y_1 & u_2 & u_1 \end{bmatrix} \hat{\theta}. \quad (3.9)$$

A realização de infinitas predições a frente é encontrada utilizando os passos anteriores

$$\hat{y}_4 = \psi_{\hat{y}u,k-1}^Y \hat{\theta} = \begin{bmatrix} \hat{y}_3 & y_2 & u_3 & u_2 \end{bmatrix} \hat{\theta}, \quad (3.10)$$

$$\hat{y}_5 = \psi_{\hat{y}u,k-1}^Y \hat{\theta} = \begin{bmatrix} \hat{y}_4 & \hat{y}_3 & u_4 & u_3 \end{bmatrix} \hat{\theta}. \quad (3.11)$$

△

Neste trabalho, utiliza-se a simulação livre como uma forma de realizar testes para observar se o modelo identificado consegue descrever a dinâmica do sistema real.

### 3.2.5.2 O índice de erro médio quadrático (RMSE)

Outra métrica utilizada para validar um modelo consiste em calcular o índice RMSE. O índice RMSE representa o valor médio quadrático do erro de predição do modelo normalizado pelo valor médio quadrático do erro de predição de um preditor trivial. O índice RMSE é dado por

$$RMSE = \frac{\sqrt{\sum_{k=1}^N (y_k - \hat{y}_k)^2}}{\sqrt{\sum_{k=1}^N (y_k - \bar{y})^2}}, \quad (3.12)$$

em que  $\hat{y}_k$  é a simulação livre do sinal e  $\bar{y}$  é o valor médio do sinal medido  $y_k$ , sendo que a média é calculada na janela de identificação. A partir da equação (3.12), percebe-se que se RMSE for próximo de um, significa que o desempenho na predição obtido a partir do modelo identificado não é significativamente melhor que o preditor trivial. O alto valor do índice RMSE indica que existe a possibilidade de que o modelo identificado não representa o comportamento dinâmico do processo.

## 3.3 Modelos Não Linear

Como grande parte dos sistemas dinâmicos apresentam características não lineares, para uma representação mais fiel dos sistemas são utilizadas representações não lineares. Ao utilizar representações não lineares haverá um aumento da complexidade dos algoritmos utilizados para sua obtenção, visto que ao trabalhar com modelos não lineares a quantidade de termos candidatos aumenta consideravelmente, sendo assim necessário utilizar métodos que consigam selecionar quais termos possuem maior relevância em descrever a variável de estudo abordada. Dessa forma, existem métodos que tratam as etapas de determinação da estrutura do modelo e a estimação de parâmetros isoladamente, como também existem métodos em que estas etapas são executadas simultaneamente. Neste trabalho serão utilizados métodos para realizar a execução destas etapas de forma simultânea.

Como os modelos NARX, Modelos Não lineares AutoRegressivos com Média Móvel e entradas eXógenas (do inglês, *Nonlinear AutoRegressive Moving Average with eXogenous inputs*) (NARMAX), modelos de Hammerstein e de Wiener e outras (AGUIRRE, 2004). Para este trabalho foi utilizada a representação NARX, por ser amplamente utilizada em trabalhos acadêmicos demonstrando um desempenho bastante positivo como, por exemplo, em (JÚNIOR, 2006). A representação do modelo NARX é descrito por

$$y_k = F^l [y_{k-1}, \dots, y_{k-n_y}, u_{k-1}, \dots, u_{k-n_u}], \quad (3.13)$$

em que  $F^l[\cdot]$  consiste em uma representação polinomial dos dados de entrada  $u(k)$  e saída  $y(k)$  que possui o grau de não linearidade  $l$ . Desta forma (3.13) pode ser representada por somatórios dos termos variando o grau de linearidade de  $1 \leq m \leq l$ . A partir da



variação do grau de linearidade, cada termo relacionado aos dados de entrada e de saída apresenta um fator de grau, ou seja, os dados de entrada apresentarão um fator de grau  $(m - p)$  e os dados de saída um fator de grau  $p$ . Além disso, os dados de entrada e saída são multiplicados pelo parâmetro  $c_{p,m-p}(n_1, \dots, n_m)$  obtendo a seguinte expressão

$$y_k = \sum_{m=0}^l \sum_{p=0}^m \sum_{n_1, n_m}^{n_y, n_u} c_{p,m-p}(n_1, \dots, n_m) \prod_{i=1}^p y_{k-n_i} \prod_{i=p+1}^m u_{k-n_i}, \quad (3.14)$$

sendo que

$$\sum_{n_1, n_m}^{n_y, n_u} = \sum_{n_1=1}^{n_y} \dots \sum_{n_m=1}^{n_u}, \quad (3.15)$$

em que o limite superior varia dependendo do tipo de fator que está sendo analisado, ou seja, caso o fator seja do tipo  $y_{k-n_i}$  o limite superior será  $n_y$ , caso o fator seja do tipo  $u_{k-n_i}$  o limite superior será  $n_u$ . Visando elucidar (3.14), a seguir apresenta-se um exemplo 10.3.1 retirado de Aguirre (2004).

**Example 2.** A expansão de uma função  $F^l[\cdot]$  de grau de linearidade  $l = 2$  é dada por

$$\begin{aligned} y_k = & c_{0,0} + \sum_{n_1=1}^{n_y} c_{1,0}(n_1) y_{k-n_1} + \sum_{n_1=1}^{n_u} c_{0,1}(n_1) u_{k-n_1} \\ & + \sum_{n_1=1}^{n_y} \sum_{n_2=1}^{n_y} c_{2,0}(n_1, n_2) y_{k-n_1} y_{k-n_2} \\ & + \sum_{n_1=1}^{n_y} \sum_{n_2=1}^{n_u} c_{1,1}(n_1, n_2) y_{k-n_1} u_{k-n_2} \\ & + \sum_{n_1=1}^{n_u} \sum_{n_2=1}^{n_u} c_{0,2}(n_1, n_2) u_{k-n_1} u_{k-n_2}. \end{aligned} \quad (3.16)$$

△

Dessa forma, visando demonstrar a variação da quantidade de regressores ao utilizar valores fixos para os máximos atrasos de entrada  $n_u$  e saída  $n_y$ , variando somente o grau de linearidade  $l$ , apresenta-se os dois seguintes exemplos.

**Example 3.** Para os máximos atrasos, tem-se  $n_y = n_u = 2$  e o grau de linearidade  $l=1$ , a expansão da função não linear é dada por

$$\begin{aligned} y_k = & c_{0,0} + \sum_{n_1=1}^{n_y=2} c_{1,0}(n_1) y_{k-n_1} + \sum_{n_1=1}^{n_u=2} c_{0,1}(n_1) u_{k-n_1}, \\ = & c_{0,0} + c_{1,0}(1) y_{k-1} + c_{1,0}(2) y_{k-2} + c_{0,1}(1) u_{k-1} + c_{0,1}(2) u_{k-2}. \end{aligned} \quad (3.17)$$

A partir da Equação 3.17 são obtidas as informações apresentadas na Tabela 2. O regressor constante  $c_{0,0}$  é desconsiderado, visto que este não é necessário para explicar a dinâmica do sistema, pois este não apresenta uma relação direta com os dados de entrada e saída do modelo. Dessa forma, o modelo apresenta quatro regressores para descrever a dinâmica do sistema.

△

Tabela 2 – Regressores e parâmetros para um modelo com os máximos atrasos  $n_y = n_u = 2$  e o grau de linearidade  $l=1$ .

Quantidade de regressores	Regressores	Parâmetros
1	$y_{k-1}$	$c_{1,0}(1)$
2	$y_{k-2}$	$c_{1,0}(2)$
3	$u_{k-1}$	$c_{0,1}(1)$
4	$u_{k-2}$	$c_{0,1}(2)$

**Example 4.** Para os máximos atrasos, tem-se  $n_y = n_u = 2$  e o grau de linearidade  $l=2$ , a expansão da função não linear é dada por

$$\begin{aligned}
y_k &= c_{0,0} + \sum_{n_1=1}^{n_y=2} c_{1,0}(n_1)y_{k-n_1} + \sum_{n_1=1}^{n_u=2} c_{0,1}(n_1)u_{k-n_1} \\
&+ \sum_{n_1=1}^{n_y=2} \sum_{n_2=1}^{n_y=2} c_{2,0}(n_1, n_2)y_{k-n_1}y_{k-n_2} \\
&+ \sum_{n_1=1}^{n_y=2} \sum_{n_2=1}^{n_u=2} c_{1,1}(n_1, n_2)y_{k-n_1}u_{k-n_2} \\
&+ \sum_{n_1=1}^{n_u=2} \sum_{n_2=1}^{n_u=2} c_{0,2}(n_1, n_2)u_{k-n_1}u_{k-n_2}. \\
&= c_{0,0} + c_{1,0}(1)y_{k-1} + c_{1,0}(2)y_{k-2} + c_{0,1}(1)u_{k-1} + c_{0,1}(2)u_{k-2} \\
&+ c_{2,0}(1, 1)y_{k-1}y_{k-1} + c_{2,0}(2, 1)y_{k-2}y_{k-1} + c_{2,0}(2, 2)y_{k-2}y_{k-2} \\
&+ c_{1,1}(1, 1)y_{k-1}u_{k-1} + c_{1,1}(2, 1)y_{k-2}u_{k-1} + c_{1,1}(2, 2)y_{k-2}u_{k-2} \\
&+ c_{0,2}(1, 1)u_{k-1}u_{k-1} + c_{0,2}(2, 1)u_{k-2}u_{k-1} + c_{0,2}(1, 2)u_{k-1}u_{k-2} + c_{0,2}(2, 2)u_{k-2}u_{k-2},
\end{aligned} \tag{3.18}$$

Ao trabalhar com o grau de linearidade  $l=2$ , percebe-se que o modelo possui regressores não lineares, apresentando regressores que realizam multiplicações entre si, aumentando consideravelmente a quantidade de regressores. Na Tabela 3 são demonstradas todos os possíveis regressores e seus parâmetros. Ao desconsiderar o regressor constante, o modelo apresenta quatorze regressores para descrever a dinâmica do sistema.

△

A partir dos exemplos 3 e 4, percebe-se que o aumento do grau de linearidade  $l$  do modelo e dos máximos atrasos com relação aos dados de entrada  $n_u$  e saída  $n_y$  proporcionam um grande aumento dos termos candidatos. Dessa forma, ao trabalhar com modelos não lineares é necessário escolher quais os regressores devem ser utilizados, visto que nem todos apresentam uma contribuição significativa para descrever a dinâmica de um sistema. Na próxima seção serão apresentados os métodos utilizados para realizar a determinação da estrutura e estimação dos parâmetros para os modelos não lineares.

Tabela 3 – Regressores e parâmetros para um modelo com os máximos atrasos  $n_y = n_u = 2$  e o grau de linearidade  $l=2$ .

Quantidade de regressores	Regressores	Parâmetros
1	constante	$c_{0,0}$
2	$y_{k-1}$	$c_{1,0}(1)$
3	$y_{k-2}$	$c_{1,0}(2)$
4	$u_{k-1}$	$c_{0,1}(1)$
5	$u_{k-2}$	$c_{0,1}(2)$
6	$y_{k-1}y_{k-1}$	$c_{2,0}(1,1)$
7	$y_{k-2}y_{k-1}$	$c_{2,0}(2,1)$
8	$y_{k-2}y_{k-2}$	$c_{2,0}(2,2)$
9	$u_{k-1}u_{k-1}$	$c_{1,1}(1,1)$
10	$u_{k-2}u_{k-1}$	$c_{1,1}(2,1)$
11	$u_{k-2}u_{k-2}$	$c_{1,1}(2,2)$
12	$y_{k-1}u_{k-1}$	$c_{0,2}(1,1)$
13	$y_{k-2}u_{k-1}$	$c_{0,2}(2,1)$
14	$y_{k-1}u_{k-2}$	$c_{0,2}(1,2)$
15	$y_{k-2}u_{k-2}$	$c_{0,2}(2,2)$

### 3.4 Determinação da estrutura e estimação dos parâmetros

Para esta etapa do processo de identificação de sistemas é necessária uma maior atenção para sistemas não lineares. A partir da decomposição explicitada pelos exemplos 3 e 4 é demonstrado que ao aumentar o grau de não linearidade e/ou os máximos atrasos do sistema  $n_y$  e  $n_u$  maior será a quantidade de termos candidatos para representar o modelo. Para o modelo do exemplo 4 tem-se o seguinte vetor de regressores que possui medições  $k - 1$ .

$$\psi_{k-1} = \begin{bmatrix} y_{k-1} & y_{k-2} & u_{k-1} & u_{k-2} & y_{k-1}^2 & y_{k-1}y_{k-2} \\ y_{k-2}^2 & u_{k-1}^2 & u_{k-1}u_{k-2} & u_{k-2}^2 & y_{k-1}u_{k-1} & y_{k-2}u_{k-1} & y_{k-1}u_{k-2} & y_{k-2}u_{k-2} \end{bmatrix}^T. \quad (3.19)$$

Visando obter os parâmetros do modelo apresentado é utilizada a seguinte representação

$$y_k = \psi_{k-1}^T \hat{\theta} + \xi_k. \quad (3.20)$$

em que  $\psi_{k-1}$  representa o vetor de regressores,  $\xi_k$  é o vetor de resíduos os quais pode incluir medições até  $k - 1$  e  $\hat{\theta}$  é o vetor de parâmetros. Ao aplicar (3.20) em (3.19) obtém-se a matriz de regressores  $\psi$ . A partir da matriz de regressores  $\psi$  é possível perceber se o método clássico MQ consegue estimar os parâmetros do modelo, considerando que a matriz de regressores possua poucas colunas e os regressores dependam apenas de valores dos sinais de entrada  $u_k$  e saída  $y_k$ . Caso a matriz de regressores possua muitas colunas, isto ocorre na identificação de sistema não lineares, e a estimativa dos parâmetros apresente uma baixa qualidade, é necessário utilizar outras alternativas como os métodos de Mínimos Quadrados Ortogonal (MQO) que necessitam que a matriz de re-

gressores seja ortogonal. Visando condicionar a matriz de regressores considerando sua ortogonalidade, na literatura existem métodos bastante utilizados, como o método clássico de Gram-Schmidt (CGS), método modificado de Gram-Schmidt (MGS), método de Golub-Householder Golub-Householder (GH), entre outros (AGUIRRE, 2004).

Neste trabalho é utilizado o método de GH para realizar a obtenção da estrutura do modelo enquanto simultaneamente a taxa ERR pondera quais são os regressores mais significativos para compor o modelo.

### 3.4.1 O método de Golub-Householder (GH)

O método de Golub-Householder (GH) consiste em um método baseado nos MQO para a obtenção dos parâmetros para modelos não lineares. Este método opera sobre (3.21) alterando matrizes e vetores para se obter o vetor de parâmetros  $\hat{\theta}$ .

$$y(k) = \psi^T \hat{\theta} + \xi(k). \quad (3.21)$$

Inicialmente o modelo (3.21) é multiplicado por uma matriz  $Q \in R^{n_\theta \times n_\theta}$ . Dessa forma, tem-se

$$\begin{aligned} Qy(k) &= Q\psi^T \hat{\theta} + Q\xi(k), \\ y^*(k) &= (\psi)^{*T} \hat{\theta} + \xi^*(k). \end{aligned} \quad (3.22)$$

A função custo de (3.22) é representada por

$$J_{MQ}^* = \sum_{i=1}^N \xi^*(i)^2 = \xi^{*T} \xi^*,$$

De forma semelhante igual realizado para os sistemas lineares (3.6), a função custo pode ser expressa

$$\begin{aligned} J_{MQ}^* &= (y^* - \Psi^* \hat{\theta})^T (y^* - \Psi^* \hat{\theta}) \\ J_{MQ}^* &= y^{*T} y^* - y^{*T} \Psi^* \hat{\theta} - \hat{\theta}^T \Psi^{*T} y^* + \hat{\theta}^T \Psi^{*T} \Psi^* \hat{\theta} \\ J_{MQ}^* &= y^T Q^T Q y - y^T Q^T Q \Psi \hat{\theta} - \hat{\theta}^T \Psi^T Q^T Q y + \hat{\theta}^T \Psi^T Q^T Q \Psi \hat{\theta} \\ J_{MQ}^* &= (y - \Psi \hat{\theta})^T Q^T Q (y - \Psi \hat{\theta}). \end{aligned} \quad (3.23)$$

Ao realizar a comparação da função custo dos modelos lineares (3.6) com (3.23), percebe-se que para que estas sejam iguais  $J_{MQ}^* = J_{MQ}$  é necessário que  $Q^T Q = I$ , ou seja, a matriz  $Q$  deve ser ortonormal e a matriz  $\Psi^*$  considere que

$$\Psi^* = Q\Psi = \begin{bmatrix} V \\ 0 \end{bmatrix}, \quad (3.24)$$

sendo  $V$  uma matriz triangular superior,  $V \in R^{n_\theta \times n_\theta}$ . Desta forma,

$$\begin{aligned}\Psi^{*T}\Psi^* &= \begin{bmatrix} V^T & 0^T \end{bmatrix} \begin{bmatrix} V \\ 0 \end{bmatrix}, \\ &= V^T V \\ &= \Psi^T Q^T Q \Psi \\ &= \Psi^T \Psi.\end{aligned}\tag{3.25}$$

Quando  $J_{MQ}^*$  é reescrita, obtém-se

$$\begin{aligned}J_{MQ}^* &= \left( \begin{bmatrix} \mathbf{y}_1^* \\ \mathbf{y}_2^* \end{bmatrix} - \begin{bmatrix} V \\ 0 \end{bmatrix} \hat{\theta} \right)^T \left( \begin{bmatrix} \mathbf{y}_1^* \\ \mathbf{y}_2^* \end{bmatrix} - \begin{bmatrix} V \\ 0 \end{bmatrix} \hat{\theta} \right) \\ &= \left( \begin{bmatrix} \mathbf{y}_1^{*T} & \mathbf{y}_2^{*T} \end{bmatrix} - \hat{\theta}^T \begin{bmatrix} V^T & 0 \end{bmatrix} \right) \left( \begin{bmatrix} \mathbf{y}_1^* \\ \mathbf{y}_2^* \end{bmatrix} - \begin{bmatrix} V \\ 0 \end{bmatrix} \hat{\theta} \right) \\ &= \mathbf{y}_1^{*T} \mathbf{y}_1^* + \mathbf{y}_2^{*T} \mathbf{y}_2^* - \hat{\theta}^T V^T \mathbf{y}_1^* - \mathbf{y}_1^{*T} V \hat{\theta} + \hat{\theta}^T V^T V \hat{\theta} \\ &= (\mathbf{y}_1^* - V \hat{\theta})^T (\mathbf{y}_1^* - V \hat{\theta}) + \mathbf{y}_2^{*T} \mathbf{y}_2^*.\end{aligned}\tag{3.26}$$

A partir de (3.26) percebe-se que ao substituir  $y^*$  por um conjunto de dois vetores  $\mathbf{y}_1^*$  e  $\mathbf{y}_2^*$  e substituir  $\Psi^*$  pela consideração (3.24) tem-se  $y_1^* = V \hat{\theta}$  quando a função de custo é minimizada resultando em  $J_{MQ}^* = \mathbf{y}_2^{*T} \mathbf{y}_2^*$ . Dessa forma, como o objetivo do uso do método é encontrar os parâmetros para o modelo e sabendo que  $V$  é uma matriz triangular superior, os parâmetros podem ser encontrados pela seguinte expressão

$$\hat{\theta}_{MQ} = V^{-1} \mathbf{y}_1^*.\tag{3.27}$$

Para determinar os parâmetros dos modelos é necessário possuir a matriz  $V$ , porém para obtê-la é necessário antes possuir a matriz  $Q$  obtida pela transformação de Householder da matriz de regressores. Assim, na próxima seção é demonstrado a transformação Householder.

#### 3.4.1.1 A transformação de Householder

A transformação Householder tem por objetivo obter a matriz  $Q$ , sendo que esta matriz necessita que as seguintes imposições sejam satisfeitas

1. Ser ortonormal, ou seja,  $Q^T Q = I$ ;
2. Satisfazer (3.24).

Para realizar o procedimento é inicialmente definida a matriz estendida  $\tilde{\Psi} = [\Psi \ y] \in R^{N \times (n+1)}$ . Dessa forma, a matriz  $Q$  é obtida ao encontrar todas as  $n$  transformações Householder,  $H^{(i)}$ , de forma que juntas componham a matriz  $Q$  como visto logo abaixo

$$Q = H^{(n)} H^{(n-1)} \dots H^{(1)},\tag{3.28}$$

em que  $Q\tilde{\Psi} = \tilde{\Psi}^n$ , sendo  $\tilde{\Psi}^n$

$$\tilde{\Psi}^{(n)} = \begin{bmatrix} V & \mathbf{y}_1^* \\ 0 & \mathbf{y}_2^* \end{bmatrix}, \quad (3.29)$$

em que  $y_1^*$ ,  $y_2^*$  e  $V$  são considerações demonstradas e utilizadas para se obter (3.26). Para obter uma matriz de transformação Householder, tem-se

$$H^{(i)} = I - \beta^{(i)} \mathbf{v}^{(i)} \mathbf{v}^{(i)T}, i = 1, \dots, n, \quad (3.30)$$

em que  $v^{(i)} \in R^N$  e esse é o vetor HouseHolder dado por

$$\mathbf{v}^{(i)} = \begin{bmatrix} v_1^i \\ v_2^i \\ \vdots \\ v_N^i \end{bmatrix}, \quad (3.31)$$

$$v_j^i = \begin{cases} 0, & j < i; \\ \tilde{\psi}_{ii}^{(i-1)} + \text{sgn}[\tilde{\psi}_{ii}^{(i-1)}] \sigma^{(i)}, & j = i; \\ \tilde{\psi}_{ji}^{(i-1)}, & j > i, \end{cases} \quad (3.32)$$

em que  $\tilde{\psi}_{ji}^{(i-1)}$  é o  $(ji)$ -ésimo elemento da matriz estendida  $\tilde{\Psi}^{(i-1)}$  e  $\text{sgn}[\ ]$  retorna o resultado da função sinal para o atual elemento. Os escalares  $\beta^{(i)}$  e  $\sigma^{(i)}$  são dados por

$$\beta^{(i)} = \frac{1}{\sigma^{(i)}(\sigma^{(i)} + |\tilde{\psi}_{ii}^{(i-1)}|)} \quad (3.33)$$

$$\sigma^{(i)} = \sqrt{\sum_{j=i}^N (\tilde{\psi}_{ji}^{(i-1)})^2}. \quad (3.34)$$

O primeiro passo do algoritmo utiliza  $\tilde{\Psi}^{(0)} = \tilde{\Psi}$ , porém para os demais passos tem-se

$$\begin{aligned} \tilde{\Psi}^i &= H^{(i)} \tilde{\Psi}^{(i-1)}, \\ &= [I - \beta^{(i)} \mathbf{v}^{(i)} \mathbf{v}^{(i)T}] \tilde{\Psi}^{(i-1)}, \\ &= \tilde{\Psi}^{(i-1)} - \mathbf{v}^{(i)} \beta^{(i)} [\mathbf{v}^{(i)}]^T \tilde{\Psi}^{(i-1)}, i = 1, \dots, n. \end{aligned} \quad (3.35)$$

Dessa forma, ao utilizar (3.35) e as definições (3.32), (3.33) e (3.34) é possível obter  $\tilde{\Psi}^n$  visando se encontrar a matriz  $Q$  e conseqüentemente a matriz  $V$ , assim possibilitando encontrar os parâmetros do modelo por (3.27).

### 3.5 Taxa de Redução de Erro (ERR)

A partir da problemática que ao trabalhar com modelos não lineares a quantidade de termos candidatos para compor o modelo aumenta significativamente, o algoritmo ERR

é utilizado para determinar a estrutura de modelos não lineares, ou seja, esse critério é capaz de ajudar a selecionar quais regressores devem ser incluídos em um modelo, dado um conjunto de regressores candidatos selecionando os termos candidatos que possuem o maior ERR dentre os termos candidatos. Para isso é necessário calcular a variância do erro de modelagem  $\xi(k)$ , sendo esta variância dada por

$$\text{Var}\xi(k) = \lim_{N \rightarrow \infty} \frac{1}{N} \left[ y^T y - \sum_{i=1}^n g_i^2 w_i^T w_i \right], \quad (3.36)$$

ao qual  $N$  corresponde ao número de termos incluídos ao modelo,  $y$  o vetor dos dados de saída,  $g_i$  o vetor de parâmetros e  $w_i$  os regressores ortogonais em relação aos dados. Analisando (3.36) percebe-se que o modelo altera a variância  $\xi(k)$  ao serem acrescentados novos termos ao modelo, sendo  $g_i^2 w_i^T w_i$  a parte responsável por explicar a adição de novos termos. A taxa de redução de erro ao adicionar  $i$ -ésimos regressores, que pode ser expressa por

$$[ERR]_i = \frac{g_i^2 w_i^T w_i}{y^T y}. \quad (3.37)$$

Vale ressaltar que (3.37) é limitada, pois esta necessita de valores da estimativa dos parâmetros  $g_i$  calculados anteriormente. Como para os modelos não lineares deseja-se escolher os regressores antes de estimar os parâmetros, (3.37) sofre alterações para que em base ortogonais os parâmetros possam ser calculados ao mesmo tempo, em que se escolhe os regressores candidatos a compor o modelo. Desta forma, para determinar  $\hat{g}_i$  para um modelo em base ortogonal,

$$g_i = \frac{w_i y}{w_i^T w_i}, i = 1, \dots, n_\theta. \quad (3.38)$$

Assim, este método de obtenção da taxa de redução de erro pode ser utilizado quando se trata de um modelo representado em base ortogonal.

## 3.6 Algoritmo GH

Durante o processo de identificação de sistemas para modelos não lineares, a estimação dos parâmetros e a obtenção da estrutura do modelo é realizada ao mesmo tempo, dessa forma, neste trabalho é utilizado o algoritmo de GH que utiliza a combinação dos métodos de GH e da taxa ERR de forma retornar o modelo utilizando os termos candidatos mais significativos para descrever a variável de nível. Ao iniciar o algoritmo de GH é criada a matriz estendida

$$\tilde{\Psi}^{(0)} = [\Psi \ y] = [\psi_1^0 \ \dots \ \psi_{n_\theta}^0 \ y^0]. \quad (3.39)$$

A matriz estendida pode ser descrita em função de  $V$  e  $\mathbf{y}$  como descrito em (3.29). Assim, após  $k - 1$  transformações de Householder (3.30) aplicadas a matriz obtém-se

$$\tilde{\Psi}^{(k-1)} = \begin{bmatrix} V_{k-1} & & & & \\ 0 & \psi_j^{(k-1)} & \dots & \psi_{n_\theta}^{(k-1)} & \mathbf{y}^{(k-1)} \end{bmatrix} \quad (3.40)$$

em que 0 representa uma matriz  $(N - k + 1) \times (k - 1)$ . A função custo para um modelo que contem os  $k - 1$  parâmetros com cada respectivo regressor é dada por

$$J_{MQ}^{*(k-1)} = \sum_{i=k}^N \left( y_i^{(k-1)} \right)^2. \quad (3.41)$$

Ao realizar a  $k$ -ésima transformação Householder e esta é aplicada a  $\tilde{\Psi}^{(k-1)}$  tem-se  $\tilde{\Psi}^{(k)}$ , assim sendo possível obter a função custo com caso um novo regressor seja adicionado ao modelo sendo esta representada por

$$\begin{aligned} J_{MQ}^{*(k-1)} &= \sum_{i=k}^N \left( y_i^{(k-1)} \right)^2 - \left( y_i^{(k)} \right)^2 \\ &= \sum_{i=k+1}^N \left( y_i^{(k)} \right)^2. \end{aligned} \quad (3.42)$$

Dessa forma, o algoritmo visa determinar o  $k$ -ésimo regressor que ao ser incluído no modelo resulta no menor valor para a função custo encontrada dentre as colunas  $\psi_j^{(k-1)}$  ...  $\psi_{n_\theta}^{(k-1)}$ . Para determinar tal regressor é inicialmente necessário determinar  $a_j^{(k)}$  e  $b_j^{(k)}$  dados por

$$a_j^{(k)} = \sum_{i=k}^N \left( \tilde{\psi}_{ij}^{(k-1)} \right)^2, j = k, \dots, n_\theta, \quad (3.43)$$

$$b_j^{(k)} = \sum_{i=k}^N \tilde{\psi}_{ij}^{(k-1)} y_i^{(k-1)}, j = k, \dots, n_\theta. \quad (3.44)$$

Visando quantificar o grau de importância ao adicionar um regressor ao modelo é utilizada a taxa de redução de erro ERR obtida por

$$\begin{aligned} [ERR]_i^{(k)} &= \frac{\left( y_j^{(k)} \right)^2}{\langle \mathbf{y}, \mathbf{y} \rangle} = \frac{\left( \hat{g}_j^{(k)} \right)^2 a_j^{(k)}}{\langle \mathbf{y}, \mathbf{y} \rangle} = \left( \frac{\langle \mathbf{w}_j, \mathbf{y} \rangle}{\langle \mathbf{w}_j, \mathbf{w}_j \rangle} \right)^2 \frac{a_j^{(k)}}{\langle \mathbf{y}, \mathbf{y} \rangle} \\ &= \left( \frac{b_j^{(k)}}{a_j^{(k)}} \right)^2 \frac{a_j^{(k)}}{\langle \mathbf{y}, \mathbf{y} \rangle} = \frac{\left( b_j^{(k)} \right)^2}{a_j^{(k)} \langle \mathbf{y}, \mathbf{y} \rangle}. \end{aligned} \quad (3.45)$$

Dessa forma, a partir de (3.45) determina-se a estrutura do modelo ao mesmo tempo que se estima os parâmetros  $\hat{g}_j^{(k)}$  (3.38), calculando a taxa ERR para cada regressor que ainda não foi incluído ao modelo. Ao fim de todas as iterações tem-se o ranqueamento dos regressores que apresentam maior importância para descrever o modelo desejado, sendo possível determinar a estrutura mais eficiente para o modelo com os termos mais significativos.

### 3.7 Considerações Parciais

Neste capítulo é apresentado o processo de identificação de sistemas descrito por Aguirre (2004), demonstrando as etapas necessárias para a obtenção de modelos lineares



e não lineares. Para os modelos não lineares é demonstrada a etapa que difere do processo de identificação para modelos lineares, evidenciando como é determinada a estrutura e estimado os parâmetros do modelo. Dessa forma, a partir do diagrama demonstrado na Figura 10 demonstrasse as etapas em que ambos os modelos são semelhantes e em que parte esses se diferem.

SEQUENCIA DAS ETAPAS	ETAPA	SISTEMA LINEAR	SISTEMA NÃO LINEAR
1ª	Teste dinâmicos e coleta e dados	✓	✓
2ª	Escolha da representação matemática	✓	✓
3ª	Determinação da estrutura do modelo	✓	✓
4ª	Estimação dos parâmetros	✓	
5ª	Validação do modelo	✓	✓

Figura 10 – Diagrama de comparação das etapas utilizadas no processo identificação de sistemas para os modelos lineares e não lineares.



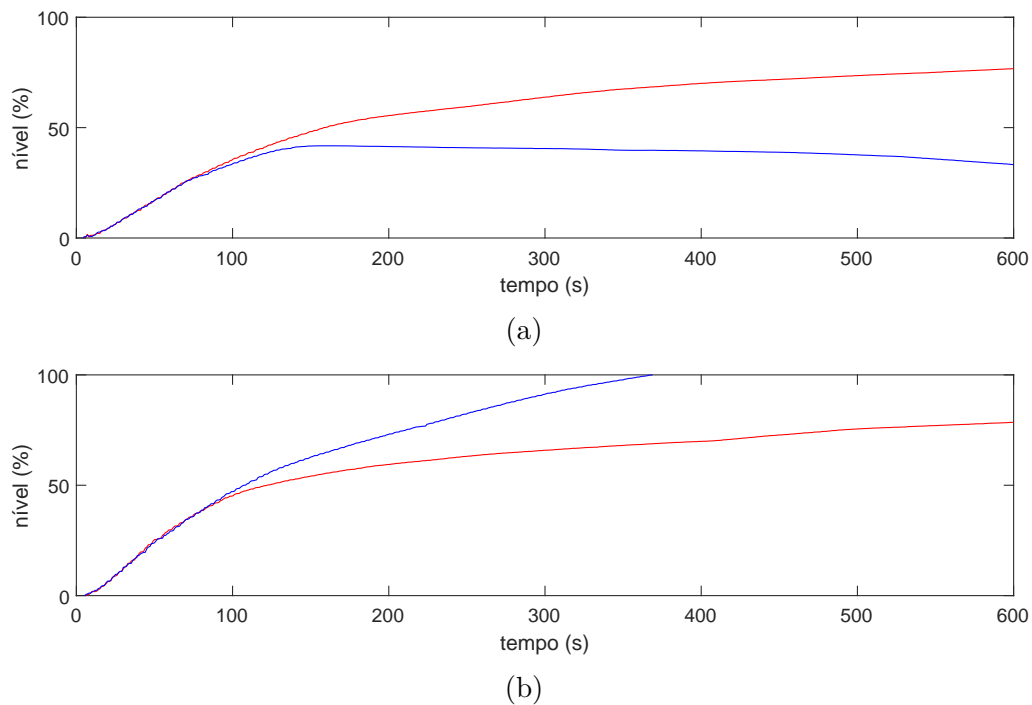


Figura 12 – Sinal de nível do tanque: (a) abertura de 30%, (b) abertura de 30%.

coletados para estes ensaios são mostrados na Figura 12. No sinal do nível do tanque com abertura de 30% (Figura 12a), observa-se que um ensaio (vermelho) apresenta um nível final próximo a 78% do nível do tanque e o outro ensaio (azul) indica um nível de aproximadamente 40%. Para as coletas de dados do nível do tanque com a abertura de 35% (Figura 12b), um ensaio (vermelho) apresenta um nível final de aproximadamente 80% enquanto o outro ensaio (azul) alcançou o nível de 100% do tanque. É possível perceber que em ambos os ensaios realizados encontrasse dificuldade na repetibilidade dos ensaios, visto que para uma mesma configuração e abertura da válvula pneumática são obtidas coletas com valores extremamente diferentes para a variável de nível. Assim, levantasse a hipótese que por falta de manutenções regulares nos equipamentos e nas tubulações da planta didática, alguns equipamentos como, por exemplo, as válvulas pneumáticas, podem não estar funcionando corretamente, resultando nos dados coletados.

Buscando contornar o problema de repetibilidade encontrado, foi desenvolvido um novo protocolo experimental onde os testes dinâmicos não utilizam a abertura da válvula pneumática como atuador para estabelecer os pontos de operações desejados. Dessa forma, foi proposto um protocolo experimental já utilizado no laboratório de Controle e Automação do ICEA. Esse protocolo utiliza a bomba trifásica em conjunto com o inversor de frequência para bombear a água presente no reservatório de água para o tanque de aquecimento a partir de um sinal de referência considerado.

## 4.2 Motor Trifásico

Para a configuração utilizando o motor trifásico todas as válvulas manuais (VM) utilizadas e a válvula pneumática (FY) estão completamente abertas. Assim, o fluxo de água sai do reservatório de água (RA) que a partir da válvula manual (VM1) vai para a bomba trifásica que bombeia a água passando pelas válvulas manuais (VM2) e (VM3) até se deparar com o primeiro instrumento, o transmissor indicador de vazão (FIT) que obtém e indica a atual vazão de entrada no sistema. Assim, após o instrumento (FIT) o fluxo passa pelas válvulas manuais (VM6, VM7, VM8) e pela válvula pneumática (FY-31) para chegar ao tanque de aquecimento (T1). No tanque de aquecimento o transmissor indicador de nível (LIT) é utilizado para realizar a medição do nível de água no tanque de aquecimento por meio da variação de pressão descrita anteriormente. Por fim, o fluxo de água que sai do tanque e retorna para o reservatório de água (RA).

A definição dos parâmetros utilizados nos ensaios é realizado pelo computador que possui o software *Matlab* (SM) que utiliza o protocolo de comunicação OPC para realizar a comunicação com a planta didática e seus componentes. Desta forma, é gerado um sinal de referência no software *Matlab* que é aplicado no inversor de frequência (IF) fazendo que a bomba trifásica (BT) mantenha uma vazão de entrada dada pelo sinal referência. Na Figura 13 mostra-se a composição da planta didática utilizada neste protocolo, destacando em verde os equipamentos utilizados e a direção do fluxo de água.

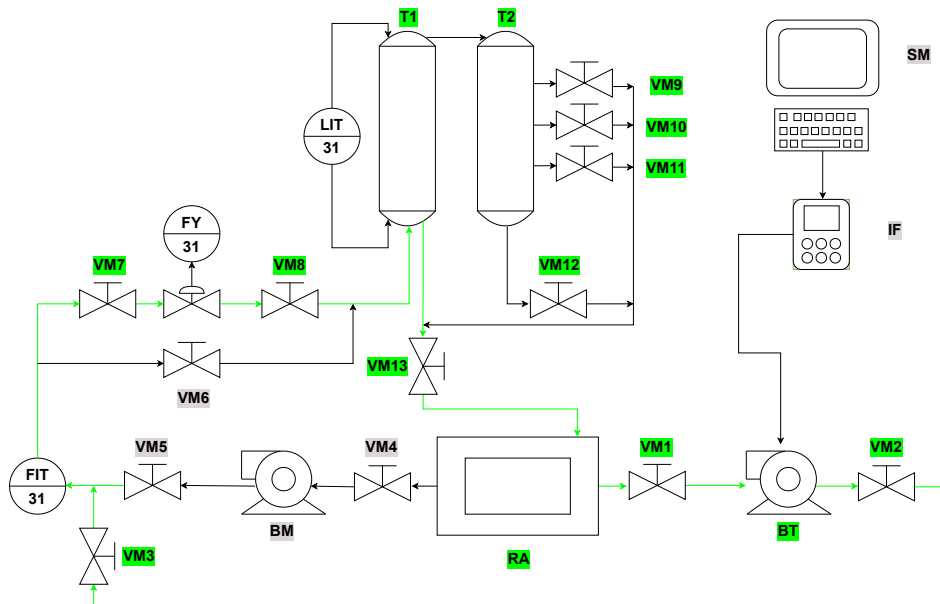


Figura 13 – Diagrama do fluxo de água para o protocolo experimental que utiliza o inversor de frequência e a bomba trifásica. Este protocolo contorna os problemas de repetibilidade na coleta de dados.

Neste trabalho, utiliza-se o sinal de referência seno para excitar a planta didática para coletar dados da variável de nível. Para os ensaios realizados foram utilizados quatro

pontos de operação, sendo os pontos de operação 75%, 60%, 45% e 30% do nível do tanque de aquecimento da planta didática. Os dados coletados são utilizados no processo de identificação de sistema com o intuito de obter modelos matemáticos que representem a variável de nível, gerando modelos lineares ARX (3.3) e não lineares NARX (3.13) para cada ponto de operação. Estes pontos de operação foram escolhidos visando trabalhar com diferentes faixas de operação que capturem o comportamento do tanque em condições que este esteja entre parcialmente cheio (75%) e parcialmente vazio (30%). A seguir, é analisado o comportamento da planta didática e os modelos obtidos utilizando o processo de identificação de sistemas.

### 4.2.1 Sinal seno

O sinal seno (3.2) foi utilizado para gerar os sinais de referência. Dessa forma, foi gerado um sinal de referência para ser aplicado pelo inversor de frequência para controlar a bomba trifásica para cada ponto de operação. Os valores utilizados para os parâmetros do sinal seno (3.2) são fixos como

$$y_k = V_p \sin(2\pi(0,005)(5)) + \text{offset}, \quad (4.1)$$

em que a frequência  $f_0$  e a taxa de amostragem  $dt$  escolhida possuem estes valores para que o sistema consiga coletar e mostrar as variações do sinal ao longo do tempo, de forma que não se perca nenhuma parte da dinâmica do sinal. Para determinar o valor de pico  $V_p$  e o *offset* do sinal foi necessário escolher as amplitudes máximas ( $a_{max}$ ) e mínimas ( $a_{min}$ ) do sinal senoidal. O *offset* inserido ao sinal tem por objetivo que sempre sejam utilizados valores positivos para o sinal de referência. Assim, por meio das amplitudes é possível encontrar o valor de pico  $V_p$  e o *offset* do sinal como

$$V_p = \frac{a_{min} + a_{max}}{2}; \quad (4.2)$$

$$\text{offset} = \frac{a_{min} - a_{max}}{2}; \quad (4.3)$$

Para obter cada sinal de referência foram realizados testes visando encontrar a amplitude máxima ( $a_{max}$ ) para cada um dos quatro pontos de operação desejados. Para os pontos de operação foi alterado somente a amplitude máxima ( $a_{max}$ ) da senoide, mantendo fixo os demais parâmetros. Na Tabela 4 mostram-se as amplitudes máximas e mínimas utilizadas para gerar o sinal de referência para cada ponto de operação. Dessa forma, a partir de (4.1) e da Tabela 4, foi construído o sinal de referência para cada ponto de operação, conforme mostrado na Figura 14a. Ao analisar os sinais de referência, percebe-se que para alterar os pontos de operação somente é alterada a amplitude do sinal senoidal.

Assim, a partir de cada sinal de referência foram realizadas coletas de dados obtendo a vazão de entrada e o nível no tanque para os quatro pontos de operação. Ao

Tabela 4 – Amplitude máxima ( $a_{max}$ ) e mínima ( $a_{min}$ ) utilizada para gerar o sinal de referência em cada ponto de operação.

Ponto de operação (%)	75	60	45	30
$a_{max}$	70	65	60	55
$a_{min}$	40	40	40	40

aplicar o sinal de entrada senoidal, é esperado que o sinal de saída apresente um comportamento com características senoidais. Os sinais de referência, vazão de entrada e o nível encontrado em cada ponto de operação são mostrados respectivamente na Figura 14a, Figura 14b e Figura 14c.

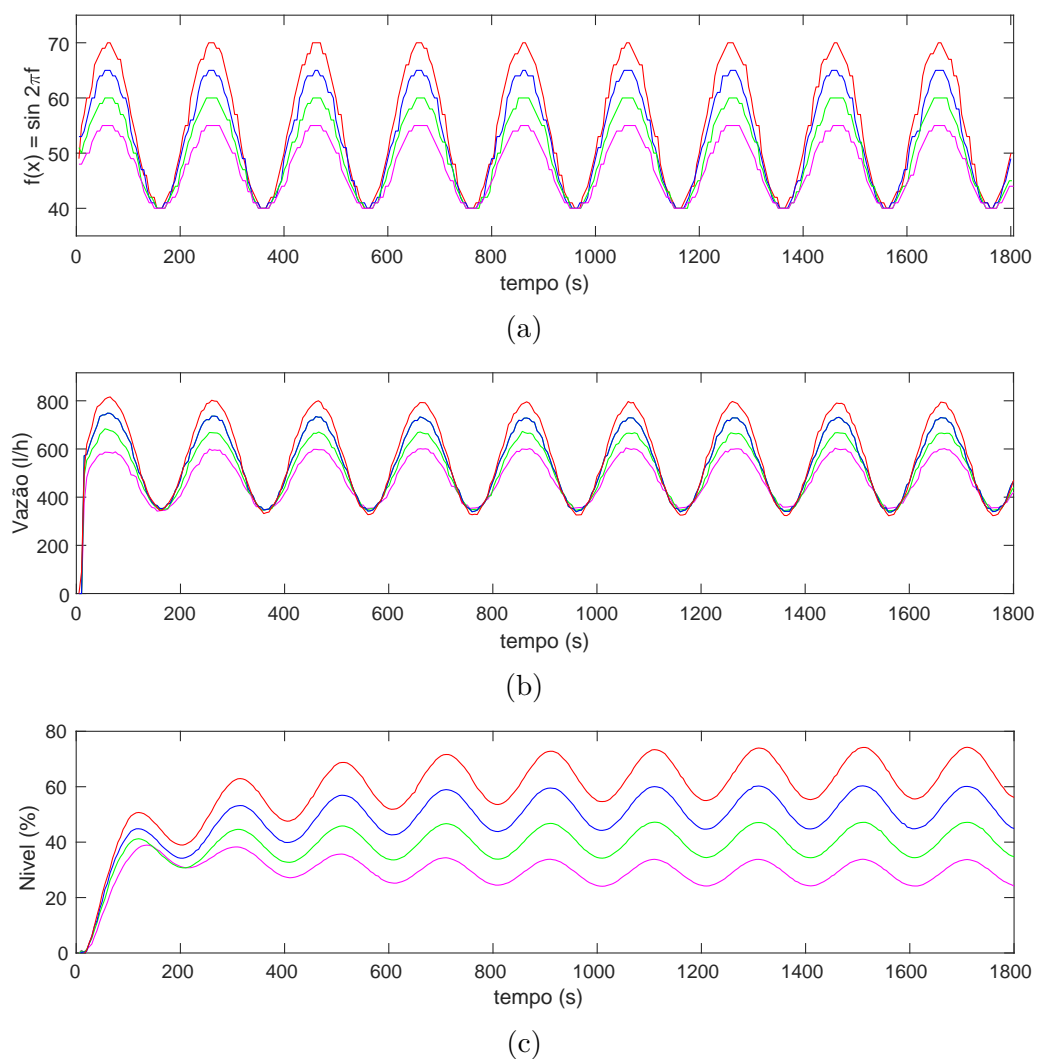


Figura 14 – Sinal de (a) referência do sinal seno, (b) vazão de entrada, (c) nível do tanque em cada ponto de operação 75% (vermelho), 60% (azul), 45% (verde) e 30% (magenta).

Ao analisar a vazão de entrada na Figura 14b, percebe-se a semelhança dos sinais de entrada com os sinais de referência gerados para cada ponto de operação, conforme mostrado na Figura 14a. Assim, ao comparar ambas as figuras, percebe-se que o sinal de

referência para cada ponto de operação conseguiu reproduzir o comportamento desejado para a vazão de entrada.

Ao analisar os sinais da variável de nível na Figura 14c, percebe-se que a planta didática no ponto de operação de 75% (vermelho) e 60% (azul) apresenta um comportamento com características semelhantes para a variável de nível, visto que esta cresce durante grande parte do tempo de ensaio até encontrar seu valor máximo ao final do ensaio. Entretanto, ao observar o ponto de operação de 45% (verde) percebe-se que após os 300 segundos iniciais o nível se encontra bastante próximo ao valor encontrado ao final do ensaio, havendo apenas um pequeno crescimento do nível durante o restante do ensaio. Ao comparar o ponto de operação de 30% (magenta) com os demais pontos, percebe-se que neste ponto o comportamento da variável de nível possui um comportamento diferente dos demais, ou seja, após o primeiro ciclo de oscilação da senoide, o nível do tanque começa a decrescer até encontrar seu estado estacionário.

Dessa forma, ao analisar os quatro pontos de operação percebe-se que os pontos de operação que apresentaram uma vazão de entrada com valores maiores ou iguais a 700 litros/hora (vermelho, azul e verde) apresentaram um comportamento onde o nível do tanque após o primeiro pico da senoide tende a crescer seu valor até o fim do ensaio, sendo que quanto maior a vazão de entrada, maior é o nível encontrado. Por outro lado, para valores de vazão de entrada menores ou iguais a 600 litros/hora (magenta) o comportamento da variável é inverso, ou seja, o nível no tanque diminui após o primeiro pico da senoide. Assim, observa-se que a vazão de entrada está diretamente relacionada ao comportamento do nível do tanque, demonstrando que após o primeiro pico da senoide, valores de vazão maiores ou próximos a 700 litros/hora fornecem um comportamento crescente do nível do tanque e valores próximos a 600 litros/hora um comportamento decrescente.

#### 4.2.2 Identificação de sistema para todos os pontos de operação

Dessa forma, com os dados coletados são obtidos modelos visando representar o comportamento da variável de nível para cada ponto de operação utilizando o processo de identificação de sistemas que contem cinco etapas. Para o processo de identificação de sistemas a vazão de entrada é utilizada como dado de entrada e o nível é utilizado como variável de saída. Neste trabalho, escolhe-se o modelo ARX para representar os modelos lineares e o modelo NARX para representar os modelos não lineares. O uso de representações lineares e não lineares tem por objetivo gerar modelos para a variável de nível, possibilitando compreender se ao utilizar representações não lineares são obtidos modelos que conseguem descrever melhor a variável de nível do que os modelos obtidos pelas representações lineares.

Para os modelos lineares após a escolha da representação matemática é realizada a determinação da estrutura do modelo utilizando o critério AIC (3.4) visando obter a quantidade de regressores necessários para compor o modelo e são posteriormente esti-

dados os parâmetros dos regressores do modelo utilizando o método MQ (3.7). Para os modelos não lineares a etapa de determinação da estrutura do modelo e estimação dos parâmetros é realizada simultaneamente. Essa etapa consiste em utilizar o algoritmo de GH, este que utiliza o método de GH para determinar a estrutura do modelo e o método ERR para estimar os parâmetros do modelo, assim possibilitando selecionar os regressores mais significativos para representar o modelo. Por fim é realizada a validação do modelo utilizando a simulação livre e o índice RMSE. Na Tabela 5, observa-se que de uma maneira geral o ponto de operação de 45% apresenta melhores resultados, pois possui o menor índice RMSE para os modelos ARX (3.3) e NARX (3.13) considerando dois ou três regressores. O processo de identificação de sistemas foi realizado para todos os pontos de operação, porém somente é apresentado o procedimento para o modelo de 45% que apresentou os melhores resultados do índice RMSE.

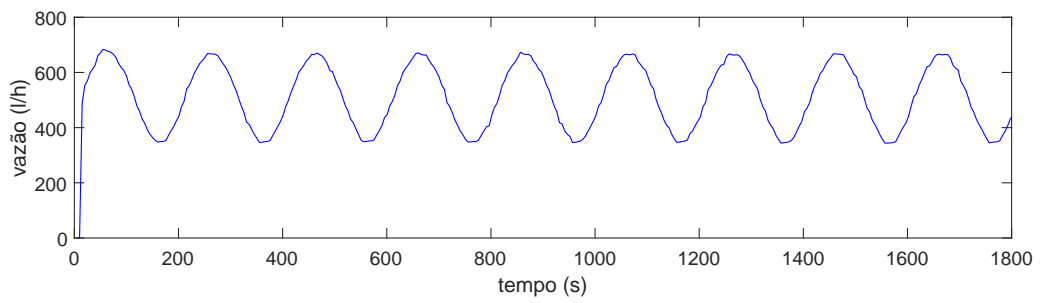
#### 4.2.3 Identificação de Sistema para o Ponto de Operação 45%

Considera-se o ponto de operação de 45% para obter modelos linear e não linear. Nas Figura 15 e Figura 16, mostram-se os dados de vazão de entrada e o nível utilizados no processo de identificação de sistemas, sendo estes divididos entre os dados de modelagem e o dados de validação. Os dados de modelagem são os dados utilizados para obter o modelo, enquanto os dados de validação são utilizados para validar o modelo identificado.

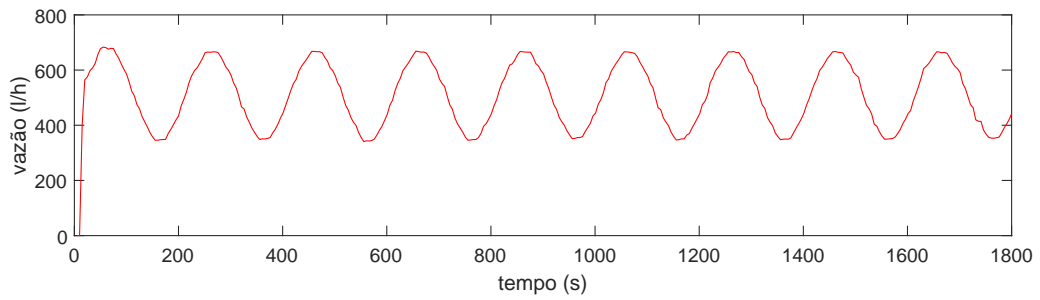
Tabela 5 – Índice RMSE para os modelos ARX e NARX com o sinal de referência senoidal para modelos utilizando dois e três regressores.

<b>Bomba Trifásica - Sinal Senoidal - RMSE</b>					
<b>Ponto de Operação (%)</b>	75	60	45	30	<b>Número de Regressores</b>
<b>Nível (%)</b>	74,66	60,55	47,20	38,71	
<b>ARX</b>	0,289	0,275	0,216	0,579	2
	1,000	1,000	1,000	1,000	3
<b>NARX</b>	0,201	0,201	0,318	1,000	2
	0,774	0,214	0,197	1,000	3



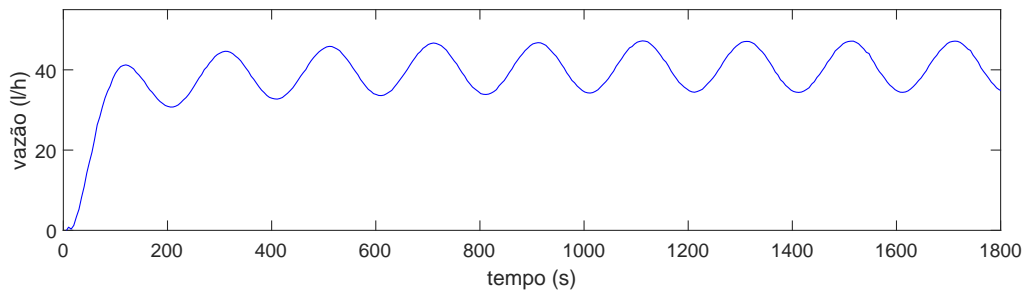


(a)

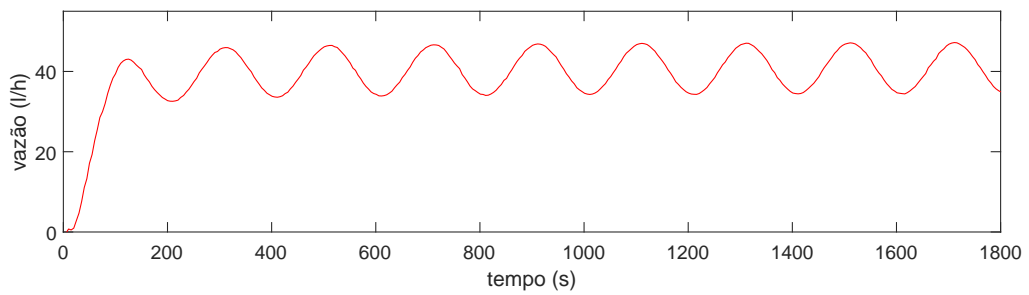


(b)

Figura 15 – Sinal de vazão de entrada para o ponto de operação de 45%. (a) Dados modelagem. (b) Dados Validação.



(a)



(b)

Figura 16 – Sinal da variável de nível para o ponto de operação de 45%. (a) Dados modelagem. (b) Dados Validação.

### 4.2.3.1 Sistema Linear

A partir dos dados coletados utiliza-se o processo de identificação de sistemas para modelos lineares. Para representar a variável de nível, os modelos lineares obtidos são compostos por dois ou três regressores. A quantidade de regressores foi obtida a partir do uso do critério AIC, sendo obtidos modelos utilizando dois e três regressores visando comparar seus desempenhos.

Como apresentado anteriormente subseção 4.2.2, o modelo ARX foi escolhido para representar o modelo linear. Visando determinar a ordem a estrutura do modelo é utilizado o critério AIC que nos fornece a ordem do modelo, ou seja, a quantidade de regressores  $u_k$  e  $y_k$  necessários para representar o modelo. Ao analisar a curva AIC é necessário observar o primeiro joelho da curva. O joelho da curva é o ponto onde o índice AIC deixa de ter uma mudança significativa ao adicionar novos regressores, assim supõem-se que esse seja o ponto onde deve-se parar de inserir regressores. Na Figura 17, observa-se que o primeiro joelho da curva se encontra com dois ou três regressores. Dessa forma, ao compor o modelo com dois regressores, o modelo será composto por dois regressor relacionados aos dados de entrada ( $u_k$ ) e dois aos dados de saída ( $y_k$ ). Ao serem realizados testes, percebeu-se que ao reduzir a ordem do modelo para um, também encontrasse um modelo capaz de descrever a dinâmica do sistema.

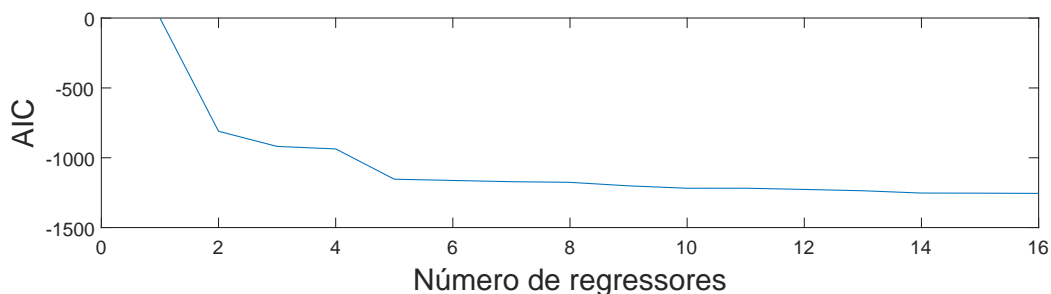


Figura 17 – AIC para o ponto de operação de 45%.

Para a estimação de parâmetros é utilizado o método MQ para estimar os parâmetros do modelo. Assim o modelo identificado é dado por

$$y_k = 9,30e^{-01}y_{k-1} + 5,55e^{-03}u_{k-1}. \quad (4.4)$$

Para validar o modelo (4.4) é utilizada a simulação livre e o índice RMSE. Na Figura 18 mostra-se a simulação livre, observa-se que no intervalo de tempo de 0 a 100 segundos é o intervalo de tempo em que a simulação livre melhor se assemelha aos dados de validação, sendo este o intervalo de tempo até se alcançar o primeiro pico da variável de nível. No intervalo de tempo de 100 a 200 segundos, observa-se que o modelo apresenta dificuldade em acompanhar os dados de validação do nível no primeiro pico de oscilação. Após os 200 segundos, a simulação livre permanece próxima dos dados de validação.

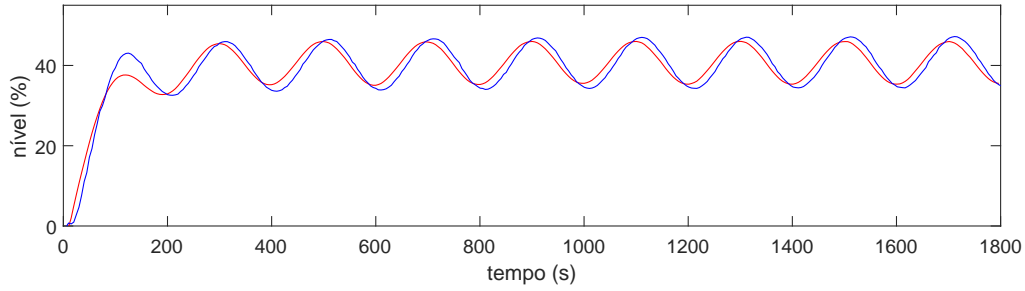


Figura 18 – Comparação entre a simulação livre (vermelho) e os dados de validação (azul) para o modelo ARX (4.4) no ponto de operação de 45%.

A segunda métrica utilizada é o índice RMSE que alcançou o valor de 0,216. Dessa forma, observa-se que o modelo identificado descreve o comportamento da variável de nível da planta, devido ao fato que quanto mais próximo o índice RMSE é de zero, melhor é a sua capacidade de descrever o comportamento da variável de interesse.

#### 4.2.3.2 Sistema Não Linear

O processo de identificação para modelos não lineares apresenta uma diferença quando comparado com os modelos lineares. A Figura 9 mostra que as etapas de determinação da estrutura e estimação dos parâmetros nos modelos não lineares são realizadas de forma simultânea.

Para os modelos obtidos foram utilizados os máximos atrasos para os dados de entrada ( $n_u$ ) e de saída ( $n_y$ ) igual a dois, além de utilizar o grau de não linearidade igual a  $l = 2$ . Dessa forma, para esses parâmetros existem 14 possíveis combinações de regressores, conforme mostrado na Tabela 6.

Tabela 6 – Possíveis combinações de regressores para os máximos atrasos  $n_u = 2$ ,  $n_y = 2$  e grau de linearidade  $l = 2$ .

Regressores modelo NARX.			
1. $y_{k-1}$	2. $y_{k-2}$	3. $u_{k-1}$	4. $u_{k-2}$
5. $u_{k-1}u_{k-1}$	6. $u_{k-1}u_{k-2}$	7. $u_{k-2}u_{k-2}$	8. $y_{k-1}u_{k-1}$
9. $y_{k-1}u_{k-2}$	10. $y_{k-2}u_{k-1}$	11. $y_{k-2}u_{k-2}$	12. $y_{k-1}y_{k-1}$
13. $y_{k-1}y_{k-2}$	14. $y_{k-2}y_{k-2}$		

A partir da Tabela 6, percebe-se que a quantidade de regressores aumenta consideravelmente conforme o aumento dos máximos atrasos de  $n_u$ ,  $n_y$  e do grau de não linearidade  $l$ . Dessa forma, para determinar a estrutura do modelo (selecionar os regressores mais significativos entre os quatorze regressores) e estimar os parâmetros é utilizado o algoritmo de GH descrito na seção 3.6.

O algoritmo de GH utiliza a matriz de regressores referente a todos os regressores mostrados na Tabela 6 como

$$\Psi = \begin{bmatrix} y_{k-1} & y_{k-2} & u_{k-1} & u_{k-2} & \dots & y_{k-1}y_{k-2} & y_{k-2}y_{k-2} \end{bmatrix} \quad (4.5)$$

por meio de (3.6) é criada a matriz de regressores estendida  $\tilde{\Psi}^{(0)}$  dada por

$$\tilde{\Psi}^{(0)} = \begin{bmatrix} y_{k-1} & y_{k-2} & u_{k-1} & u_{k-2} & \dots & y_{k-1}y_{k-2} & y_{k-2}u_{k-2} & y \end{bmatrix}. \quad (4.6)$$

A partir da matriz estendida, é realizada a primeira iteração do algoritmo. Dessa forma, é realizada a primeira transformação Householder (3.30) e em seguida, é calculada a taxa ERR (3.45) para cada uma das colunas e por fim, é realizada a permutação da coluna com a maior taxa ERR com a coluna da atual iteração. Após realizadas todas as iterações obtêm-se todos os regressores ordenados pela taxa ERR. Na Tabela 7 mostram-se os regressores ordenados de forma decrescente conforme a taxa ERR com seu correspondente parâmetro estimado.

Tabela 7 – Regressores ordenados de forma decrescente conforme a taxa ERR com seu respectivo parâmetro estimado.

Regressores	Parâmetro Estimado	ERR
1. $y_{k-1}$	0,6822	$4,5386e^{-04}$
2. $u_{k-1}u_{k-2}$	$2,5830e^{-06}$	$6,2154e^{-06}$
3. $y_{k-1}u_{k-1}$	$1,3365e^{-06}$	$2,4119e^{-06}$
4. $u_{k-1}$	0,0032	$1,5845e^{-06}$
5. $u_{k-2}u_{k-2}$	$-6,1136e^{-06}$	$1,1023e^{-06}$
6. $y_{k-1}y_{k-1}$	0,1658	$1,0437e^{-06}$
7. $y_{k-2}$	0,2084	$5,3018e^{-07}$
8. $y_{k-2}u_{k-2}$	-0,0949	$3,1935e^{-07}$
9. $u_{k-2}$	0,0025	$1,3590e^{-07}$
10. $y_{k-1}y_{k-2}$	-0,0714	$1,3354e^{-07}$
11. $u_{k-1}u_{k-1}$	$-2,9426e^{-06}$	$7,5758e^{-08}$
12. $y_{k-1}u_{k-2}$	0,0023	$5,3501e^{-08}$
13. $y_{k-2}u_{k-1}$	-0,0022	$4,1343e^{-08}$
14. $y_{k-2}u_{k-1}$	0,0047	$3,1224e^{-08}$

A partir dos regressores encontrados pelo algoritmo de GH, foram considerados modelos contendo dois e três regressores. Dessa forma, o modelo que apresentou o melhor desempenho em representar a variável de nível da planta didática no ponto de operação de 45% foi o modelo contendo três regressores dado por

$$y_k = 9,70e^{-01}y_{k-1} + 8,53e^{-06}u_{k-1}u_{k-2} - 5,30e^{-05}y_{k-1}u_{k-1} \quad (4.7)$$

Para validar o modelo (4.7) é utilizada a simulação livre e o índice RMSE. Ao comparar a simulação livre do modelo NARX com os dados de validação, observa-se que durante o intervalo de 100 a 200 segundos o modelo não consegue alcançar o nível no primeiro pico de subida apresentado pelos dados de validação, porém já no primeiro pico de descida o modelo já acompanhar os dados de validação.

O índice RMSE do modelo não linear apresentou o valor de 0,197, mostrando que apesar do modelo nos 200 segundos iniciais não conseguir acompanhar os dados de validação, após os 200 segundos o modelo obtido consegue descrever a variável de nível.

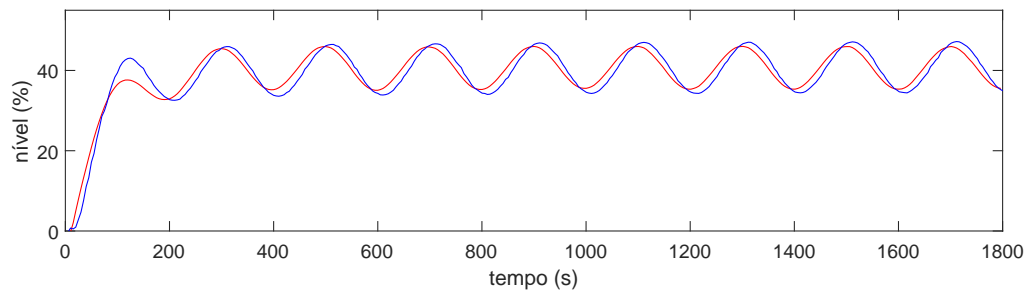


Figura 19 – Comparação entre a simulação livre (vermelho) e os dados de validação (azul) para o modelo NARX (4.7) no ponto de operação de 45%.

#### 4.2.4 Comparação entre os Modelos ARX e NARX para o Ponto de Operação de 45%

A partir dos resultados encontrados na Figura 20 é possível comparar os modelos ARX (4.4) e NARX (4.7) verificando que os modelos encontrados possuem um comportamento muito próximo. Ambos os modelos entre o intervalo de 100 a 200 segundos iniciais de ensaio apresentam dificuldade em acompanhar o primeiro pico da variável de nível, e após os 200 segundos, ambos modelos descrevem o comportamento da variável de nível.

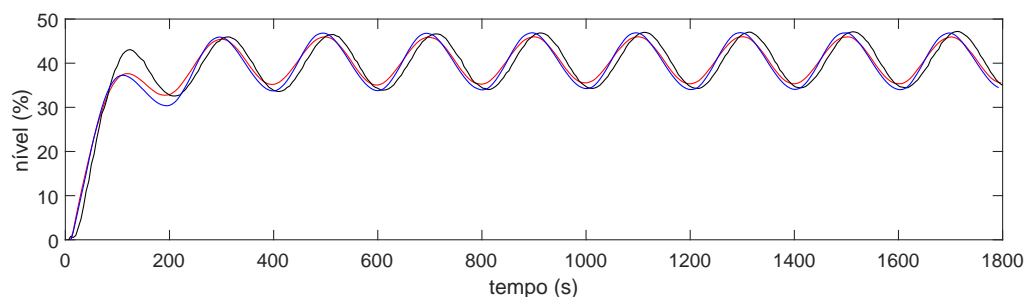


Figura 20 – Comparação da simulação livre dos modelos: ARX (vermelho) e NARX (azul), e os dados de validação (preto) para o ponto de operação de 45%.

Para mensurar qual dos modelos obtidos apresenta o melhor desempenho é utilizado o índice RMSE, conforme mostrado na Tabela 5. Observa-se que o  $RMSE = 0,197$  indica que o modelo NARX descreve de forma mais aproximada a variável de nível. Entretanto, destaca-se que a diferença de precisão encontrada entres o modelo ARX e NARX, é de 1,9%. Dessa forma, é necessário analisar a relação entre a precisão e a metodologia utilizada para obter um determinado modelo matemático que atinja o objetivo desejado.

### 4.3 Considerações Parciais

Neste capítulo, é descrito o protocolo experimental considerado neste trabalho. A partir desse protocolo foram realizados testes dinâmicos para obter os dados da vazão de entrada e da variável de nível do tanque, possibilitando obter modelos ARX e NARX que

descrevem a variável de nível do tanque por meio do processo de identificação de sistemas. Neste trabalho, considera-se o ponto de operação de 45% em virtude de apresentar o melhor desempenho em descrever a variável de nível conforme o índice RMSE comparado com os outros pontos de operação. Os resultados mostram que o modelo NARX obteve o melhor desempenho em descrever a variável de nível com uma precisão maior que 2% em comparação com o modelo ARX.

# 5 Conclusões e Propostas de Continuidade

## 5.1 Considerações Finais

Neste trabalho, investiga-se o problema de modelagem matemática para descrever o comportamento da variável de nível da planta didática SMAR PD3-F. A metodologia proposta consiste na utilização de técnicas de identificação de sistemas baseado na modelagem caixa preta para obter modelos ARX e NARX que caracterizam a dinâmica do nível da planta.

Na planta didática SMAR PD3-F, o protocolo experimental considera o uso do inversor de frequência da planta em conjunto com a bomba trifásica. Para que a partir de um sinal de referência, o inversor de frequência controle a bomba trifásica, fazendo com que esta gere uma vazão de entrada com características deste sinal. Para este protocolo foram definidos quatro pontos de operação da variável de nível do tanque, sendo estes 75%, 60%, 45% e 30%. Para cada um destes pontos de operação foi gerado um sinal de referência senoidal, que fornece características senoidais nos sinais de vazão de entrada e nível do tanque. Por fim, para cada um dos pontos de operação, foram coletados os dados das vazões de entrada e níveis do tanque. Esses dados foram utilizados, respectivamente, como dados de entrada,  $u_k$ , e dados de saída,  $y_k$ , para o processo de identificação de sistemas.

A partir dos dados coletados foram obtidos modelos ARX e NARX para representar a variável de nível da planta didática PD3-F para cada um dos quatro pontos de operação. Entretanto, foi apresentado somente o processo de identificação de sistemas para o ponto de operação de 45%, pois os modelos obtidos mostram melhor desempenho quando comparado com os outros pontos de operação.

Para a obtenção do modelo linear ARX é utilizado o critério AIC para determinar um modelo de segunda ordem. A partir de testes realizados, percebeu-se que ao utilizar somente um regressor de entrada e um regressor de saída, o modelo linear descreve a variável de nível, cujos parâmetros são estimados utilizando o método MQ. Para realizar a validação do modelo identificado, são utilizadas a simulação livre e o índice RMSE alcançando o valor de 0,216.

No processo de identificação para obter um modelo não linear NARX, a etapa de determinação da estrutura do modelo e estimação dos parâmetros são realizadas simultaneamente pelo algoritmo GH. Este algoritmo utiliza o método GH e a taxa ERR para determinar quais são os regressores que possuem maior relevância para descrever a variável de nível. Dessa forma, foi obtido um modelo que possui três regressores, sendo um regressor relacionado aos dados de saída, um regressor que contém a combinação dos

entre dados de entrada e outro regressor que possui a combinação de dados de entrada e saída. A validação do modelo obtido é realizada utilizando a simulação livre e o índice RMSE igual a 0,197.

Neste trabalho, foram obtidos modelos lineares e não lineares que caracterizam a dinâmica da variável de nível da planta didática SMAR PD3-F. Ao comparar ambos os modelos ARX e NARX, é possível perceber que o modelo NARX apresenta um melhor desempenho, porém o modelo ARX possui um desempenho muito próximo, tornando-o uma opção bastante viável, visto que para obter um modelo de representação linear o nível de complexidade para sua obtenção é menor quando comparado ao se obter um modelo não linear. Assim, o processo de identificação de sistemas do tipo caixa preta mostra-se como um método promissor para a modelagem matemática em processos industriais.

## 5.2 Propostas de Continuidade do TCC

Ao longo do trabalho realizado, alguns pontos foram levantados como propostas de continuidade a serem investigadas. Entre elas, pode-se destacar:

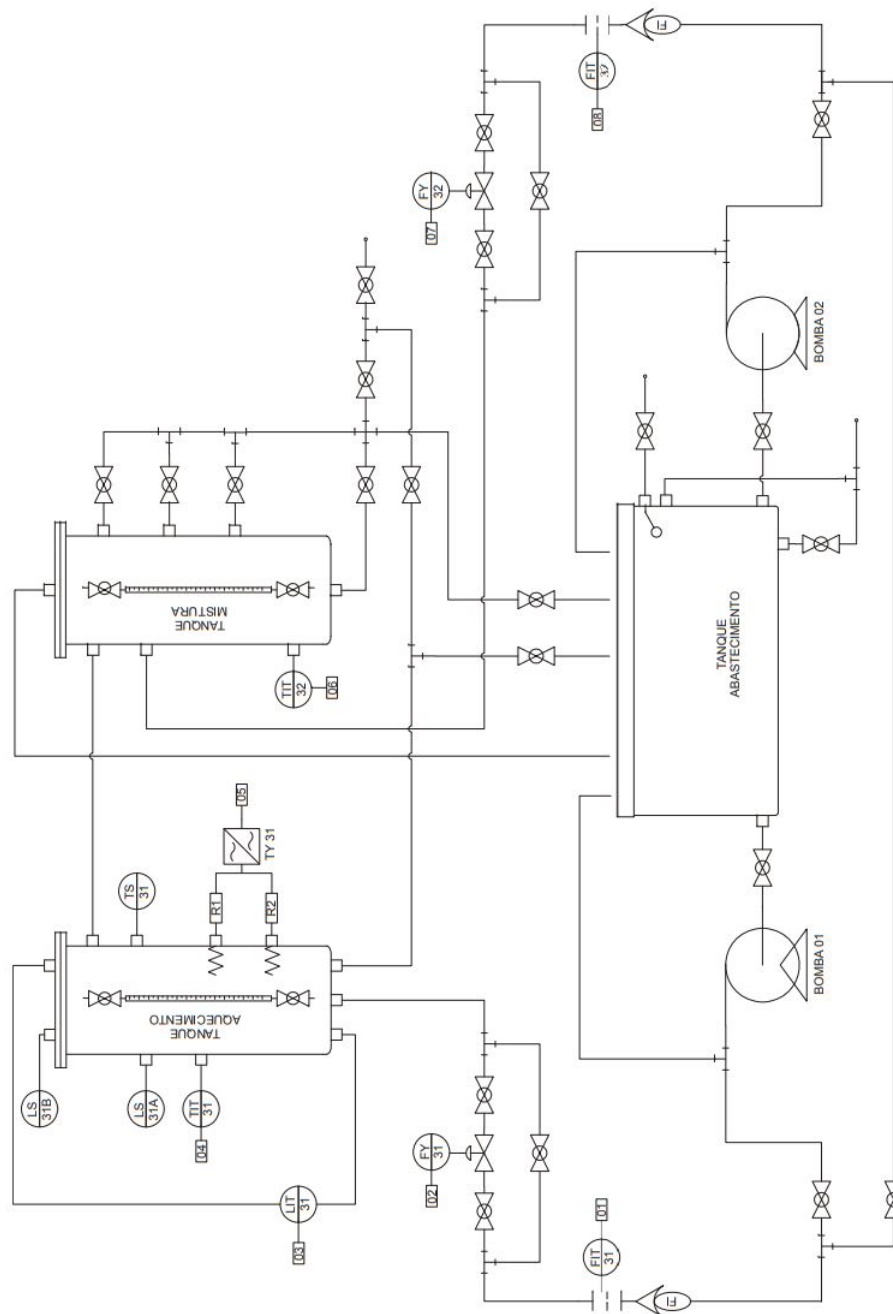
- A abordagem proposta utiliza a metodologia caixa preta para obter modelos ARX e NARX que descrevam a variável de nível do tanque considerando a aplicação de uma entrada com características senoidais. Como tentativa de melhorar os modelos obtidos, propõe-se o emprego da modelagem caixa cinza.
- Buscando realizar uma representação diferente de modelos não lineares, propõe-se utilizar os Modelos de Hammerstein e de Wiener. Esses modelos são comumente utilizados para realizar a representação de sistemas dinâmicos complexos, sendo bastante úteis para representar sistemas que exibem não linearidade em diferentes partes da dinâmica.



# A Diagrama P&I

O diagrama P&I da planta didática SMAR PD3-F apresenta as malhas 31 e 32 presentes em sua composição e os equipamentos alocados em cada malha.

Figura 21 – Diagrama P&I - malhas de controle e dispositivos.



Fonte: Retirado de SMAR (2015).

# Referências

- AGUIRRE, L. A. *Introdução à identificação de sistemas—Técnicas lineares e não-lineares aplicadas a sistemas reais*. [S.l.]: Editora UFMG, 2004. 1, 14, 16, 18, 19, 20, 21, 24, 28
- AKAIKE, H. A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, v. 19, n. 6, p. 716–723, 1974. 17, 18
- DELCROIX, B. et al. Autoregressive neural networks with exogenous variables for indoor temperature prediction in buildings. In: SPRINGER. *Building Simulation*. [S.l.], 2021. v. 14, p. 165–178. 17
- JÚNIOR, J. M. P. d. M. Redes neurais dinâmicas para predição e modelagem não-linear de séries temporais. 2006. 20
- KRISHNAN, R. Electric motor drives: modeling, analysis, and control. (*No Title*), 2001. 7, 8
- KUROSE, J. F.; ROSS, K. W. Computer networking: A top-down approach edition. *Addison Wesley*, 2007. 12
- LJUNG, L. System identification. In: *Signal analysis and prediction*. [S.l.]: Springer, 1998. p. 163–173. 17
- LOPES, J. C. S. et al. Identificação de um sistema dinâmico assimétrico por redes neurais artificiais. 2012. 2, 3
- MEHRA, R. Approaches to adaptive filtering. *IEEE Transactions on automatic control*, IEEE, v. 17, n. 5, p. 693–698, 1972. 17
- MOMBELLO, E. E.; PORTILLO, Á.; FLÓREZ, G. A. D. New state-space white-box transformer model for the calculation of electromagnetic transients. *IEEE Transactions on Power Delivery*, IEEE, v. 36, n. 5, p. 2615–2624, 2020. 1
- MURRAY, T.; CLARKE, G. K. Black-box modeling of the subglacial water system. *Journal of Geophysical Research: Solid Earth*, Wiley Online Library, v. 100, n. B6, p. 10231–10245, 1995. 1
- PEREIRA, A. P.; MUNARETO, S. D. S. Modelagem caixa cinza por analogia a um circuito rc. In: *Congresso Brasileiro de Automática-CBA*. [S.l.: s.n.], 2019. v. 1, n. 1. 1
- SMAR. *Plantas Didáticas PD3-F: Manual de instruções, operação e manutenção: Plantas didáticas. 3 ed.* 2015. Disponível em: <[www.smar.com.br](http://www.smar.com.br)>. 5, 6, 7, 10, 12, 45
- STEINBERG, H.-W.; GASSER, T.; FRANKE, J. Fitting autoregressive models to eeg time series: An empirical comparison of estimates of the order. *IEEE transactions on acoustics, speech, and signal processing*, IEEE, v. 33, n. 1, p. 143–150, 1985. 17
- VERHAPPEN, I.; PEREIRA, A. *Foundation Fieldbus*. [S.l.]: ISA, 2008. 12
- VERLY, A. et al. Identificação de modelos para uma planta de nível didática smar pd3-f. In: *Congresso Brasileiro de Automática-CBA*. [S.l.: s.n.], 2019. v. 1, n. 1. 2, 3

---

VITTURI, S.; ZUNINO, C.; SAUTER, T. Industrial communication systems and their future challenges: Next-generation ethernet, iiot, and 5g. *Proceedings of the IEEE*, v. 107, n. 6, p. 944–961, 2019. 11