



Universidade Federal de São João del-Rei
PPGEL - Programa de Pós-Graduação em Engenharia Elétrica

Determinação de Estruturas de Modelos Neurais

Luciana Maria Margoti

Dissertação submetida à banca examinadora designada pelo Colegiado do Programa de Pós-Graduação em Engenharia Elétrica da Universidade Federal de São João del-Rei, como parte dos requisitos necessários à obtenção do grau de Mestre em Engenharia Elétrica.

Orientador: Prof. Dr. Gleison Fransoares Vasconcelos Amaral

Coorientador: Prof. Dr. Márcio Falcão Santos Barroso

São João del-Rei, 30 de Junho de 2011

O Senhor nosso Deus nos segura pela mão e
nos diz: Não temas, Eu te ajudo."

(Isa 41, 13)

Resumo

Neste trabalho, utiliza-se diferentes sistemas, com conjuntos de dados de características diferentes, para a determinação de estrutura de redes neurais artificiais do tipo MLP (*perceptron* de múltiplas camadas). A partir dos sistemas analisados, verifica-se como as características dos dados utilizados na identificação do modelo podem ser úteis na escolha da estrutura da rede neural. Para testar esta hipótese, utilizou-se três diferentes sistemas reais: um aquecedor térmico de dissipação variável, que possui dados estáticos e dinâmicos coletados em uma planta, sendo que os dados dinâmicos excursionam toda faixa de operação dos dados estáticos; um conversor *buck* CC-CC, em que a aquisição dos dados dinâmicos foi realizada para uma pequena faixa de operação da curva estática (teórica); e uma planta de neutralização de pH, cujos dados dinâmicos também estão contidos numa faixa restrita da curva estática, sendo que esta última possui uma característica de saturação, que tem sido motivadora para diversos trabalhos na literatura. Os resultados obtidos, no geral, mostraram a capacidade das redes neurais em generalizar a informação aprendida. Algumas características específicas de cada conjunto de dados forneceram informações importantes, que permitiram testar e validar diferentes arquiteturas de redes neurais. O processo de obtenção de modelos partiu da identificação caixa preta, em que foram testadas redes neurais com características diferentes para conexões entre os regressores de entrada e de saída e os neurônios da camada oculta. Na identificação caixa cinza foram utilizados modelos de blocos interconectados, *Hammerstein* e *Wiener*, e diferentes arquiteturas para impor restrições aos pesos da rede, garantindo o comportamento do sistema em estado estacionário.

Abstract

This work use different systems, with different characteristics, to restrict the topology of the MLP (multilayer perceptron) neural network. From different systems, it was verified how the characteristics of the data used for model identification may be useful to chose the topology of the neural network. To test this hypothesis, three practical systems were used. A electrical heater, which has static and dynamic data collected in a plant. A dc-dc buck converter, which the dynamic data is in a small operation range in a converter steady-state voltage relation. And the pH neutralization process wich the a dynamic linear data represent the dynamics of the process, in a plant, the other hand, dynamic data is in a small operation range in a static tritration curve. Conditions under which the pH process attracted great attention. The results showed the ability of neural networks and the incorporating prior knowledge, into the structure selection procedure, was used to restrict the differents topologies of network. The process of identification started from the black box identification, differents topologies were tested for parameter estimation. In the gray box identification were used Wiener and Hammerstein models, and differents topologies to impose restrictions, guaranteeing the wide operation range in a steady-state.

Agradecimentos

A Deus, por sempre, em todos os momentos de minha vida, ter me segurado pela mão.

Aos meus orientadores, Gleison Fransoares Vasconcelos Amaral e Márcio Falcão Santos Barroso, pelo que me ensinaram, pela ajuda e grande paciência. À vocês cabe o mérito deste trabalho, que para muitos pode parecer pequeno, mas que para mim foi um grande aprendizado e descoberta.

À minha família, pai e mãe, Antônio e Ana Maria, e irmãs, Lucilene e Graziela, que me proporcionaram experiências singulares, sem as quais eu não teria me tornado o que sou hoje. E pela compreensão, carinho e ajuda em todos os momentos. Com vocês, em qualquer lugar e à qualquer hora.

À minha avó, Ana, que mesmo não estando mais entre nós, e mesmo não entendendo porque uma menina fazia um curso de Engenharia Elétrica, sempre sentiu-se orgulhosa ao dizer que eu era uma Engenheira.

Ao meu namorado, Vinícius, pelo amor, amizade, compreensão e respeito em todos os momentos. Mesmo naqueles em que a ausência era grande e a paciência pequena. Te amo!

Aos amigos do mestrado. Sem eles, grande parte do que está apresentado neste trabalho, não teria se tornado possível. Muito obrigada pela grande ajuda, dicas, risos e aprendizado.

Aos amigos, a família que Deus me permitiu escolher. Pela ajuda, solidariedade, incentivo e carinho.

À UFSJ - Universidade Federal de São João del-Rei pelo apoio financeiro, que possibilitou a conclusão deste trabalho.

Conteúdo

Resumo	ii
Abstract	iii
Agradecimentos	iv
Sumário	v
Lista de Tabelas	vii
Lista de Figuras	viii
1 Introdução	1
1.1 Motivação	2
1.2 Objetivo	3
1.3 Apresentação do trabalho	3
2 Identificação de Sistemas	4
2.1 Procedimentos para identificação de sistemas	4
2.2 Representações de sistemas não lineares	5
2.2.1 Modelos ARX, NARX e NARMAX	5
2.2.2 Modelos de Hammerstein e Wiener	6
2.2.3 Redes Neurais	8
2.3 Validação dos Modelos	13
3 Descrição dos Sistemas Modelados	15
3.1 Aquecedor com Dissipação Variável	15
3.2 Conversor <i>Buck</i>	18
3.3 Planta de Neutralização de pH	20
4 Estudo de Casos	
Identificação Caixa Preta	24
4.1 Aquecedor com Dissipação Variável	25
4.2 Conversor <i>Buck</i> - CC-CC	28
4.3 Planta de Neutralização de pH	33
5 Estudo de Casos:	
Identificação Caixa Cinza	38
5.1 Modelos de Wiener e Hammerstein	38
5.1.1 Aquecedor com Dissipação Variável	38

5.1.2	Conversor <i>Buck</i> - CC-CC	42
5.1.3	Planta de Neutralização de pH	46
5.2	Treinamento de Redes Neurais com Restrição	50
5.2.1	Aquecedor com Dissipação Variável	51
5.2.2	Conversor <i>Buck</i> CC-CC	58
5.2.3	Planta de Neutralização de pH	65
6	Conclusão	73
6.1	Sugestão para trabalhos futuros	74
	Referências Bibliográficas	78

Lista de Tabelas

4.1	Estruturas utilizadas de redes neurais - Aquecedor	25
4.2	Índices de validação para identificação caixa preta do sistema Aquecedor, para cada uma das estruturas utilizadas	27
4.3	Estrutura para treinamento das redes neurais - conversor <i>buck</i> CC-CC	28
4.4	Índices de validação para identificação caixa preta do sistema conversor <i>buck</i> , para cada uma das estruturas utilizadas	32
4.5	Estrutura para treinamento das redes neurais - pH	33
4.6	Índices de validação para identificação caixa preta da planta de neutralização de pH, para cada uma das estruturas utilizadas	36
5.1	Índices de validação para identificação do modelo de Wiener e Hammerstein neurais do Aquecedor de Dissipação Variável	42
5.2	Índices de validação para identificação dos modelos de Wiener e Hammerstein neural do conversor <i>buck</i> - CC-CC	46
5.3	Índices de validação para identificação do modelo de Wiener neural da planta de neutralização de pH	50
5.4	Arquiteturas para obtenção dos pesos estáticos (restrições) - Aquecedor	51
5.5	Estruturas utilizadas para o treinamento dinâmico das redes neurais com imposição de restrições - Aquecedor	52
5.6	Índices de validação para identificação caixa cinza, com imposição de restrições, do sistema Aquecedor.	57
5.7	Arquiteturas para obtenção dos pesos estáticos (restrições) - conversor <i>buck</i> CC-CC	58
5.8	Estruturas utilizadas para o treinamento dinâmico das redes neurais com imposição de restrições - conversor <i>buck</i> CC-CC	60
5.9	Índices de validação para identificação caixa cinza, com imposição de restrições, do sistema conversor <i>buck</i> CC-CC.	64
5.10	Arquiteturas para obtenção dos pesos estáticos (restrições) - planta de neutralização de pH	65
5.11	Estruturas utilizadas para o treinamento dinâmico das redes neurais com imposição de restrições - planta de neutralização de pH	67
5.12	Índices de validação para identificação caixa cinza, com imposição de restrições, planta de neutralização de pH	71

Lista de Figuras

2.1	Modelos de blocos interconectados. $f(\cdot)$ representa a função estática não linear e $g(\cdot)$ representa a função dinâmica linear. Em (a) Modelo de Hammerstein e em (b) Modelo de Wiener.	6
2.2	Modelo de um neurônio artificial.	9
2.3	Rede neural com uma camada oculta.	10
2.4	Função de Limiar.	10
2.5	Função linear por partes.	11
2.6	Função sigmóide (tangente hiperbólico).	11
2.7	Estruturas propostas por Narendra and Parthasarathy (1990).	12
3.1	Diagrama de blocos do sistema (Aquecedor).	15
3.2	Característica estática obtida para o sistema com o ventilador ligado.	16
3.3	Dados de entrada, em pu.	17
3.4	Dados de saída, em pu.	17
3.5	Conversor estático <i>buck</i>	18
3.6	Característica estática do conversor <i>buck</i>	19
3.7	Sinal de entrada do sistema.	19
3.8	Sinal de saída do sistema.	20
3.9	Esquema do processo de neutralização de pH.	21
3.10	Característica estática.	22
3.11	Entrada do sistema de neutralização de pH.	22
3.12	Saída dinâmica do sistema de neutralização de pH.	23
4.1	(a) Dados dinâmicos: resposta da rede 1 em linha tracejada (- -) e saída do sistema em linha contínua (-). (b) Característica estática: reprodução da característica estática da rede 1 em linha tracejada (- -) e característica estática do sistema em linha contínua (-).	26
4.2	(a) Dados dinâmicos: resposta da rede 2 em linha tracejada (- -) e saída do sistema em linha contínua (-). (b) Característica estática: reprodução da característica estática da rede 2 em linha tracejada (- -) e característica estática do sistema em linha contínua (-).	26
4.3	(a) Dados dinâmicos: resposta da rede 3 em linha tracejada (- -) e saída do sistema em linha contínua (-). (b) Característica estática: reprodução da característica estática da rede 3 em linha tracejada (- -) e característica estática do sistema em linha contínua (-).	27
4.4	(a) Dados dinâmicos: resposta da rede 4 em linha tracejada (- -) e saída do sistema em linha contínua (-). (b) Característica estática: reprodução da característica estática da rede 4 em linha tracejada (- -) e característica estática do sistema em linha contínua (-).	27
4.5	(a) Dados dinâmicos: resposta da rede 1 em linha tracejada (- -) e saída do sistema em linha contínua (-). (b) Característica estática: reprodução da característica estática da rede 1 em linha tracejada (- -) e característica estática do sistema em linha contínua (-).	31

4.6	(a) Dados dinâmicos: resposta da rede 2 em linha tracejada (- -) e saída do sistema em linha contínua (-). (b) Característica estática: reprodução da característica estática da rede 2 em linha tracejada (- -) e característica estática do sistema em linha contínua (-).	31
4.7	(a) Dados dinâmicos: resposta da rede 3 em linha tracejada (- -) e saída do sistema em linha contínua (-). (b) Característica estática: reprodução da característica estática da rede 3 em linha tracejada (- -) e característica estática do sistema em linha contínua (-).	32
4.8	(a) Dados dinâmicos: resposta da rede 4 em linha tracejada (- -) e saída do sistema em linha contínua (-). (b) Característica estática: reprodução da característica estática da rede 4 em linha tracejada (- -) e característica estática do sistema em linha contínua (-).	32
4.9	(a) Dados dinâmicos: resposta da rede 1 em linha tracejada (- -) e saída do sistema em linha contínua (-). (b) Característica estática: reprodução da característica estática da rede 1 em linha tracejada (- -) e característica estática do sistema em linha contínua (-).	35
4.10	(a) Dados dinâmicos: resposta da rede 2 em linha tracejada (- -) e saída do sistema em linha contínua (-). (b) Característica estática: reprodução da característica estática da rede 2 em linha tracejada (- -) e característica estática do sistema em linha contínua (-).	35
4.11	(a) Dados dinâmicos: resposta da rede 3 em linha tracejada (- -) e saída do sistema em linha contínua (-). (b) Característica estática: reprodução da característica estática da rede 3 em linha tracejada (- -) e característica estática do sistema em linha contínua (-).	36
4.12	(a) Dados dinâmicos: resposta da rede 4 em linha tracejada (- -) e saída do sistema em linha contínua (-). (b) Característica estática: reprodução da característica estática da rede 4 em linha tracejada (- -) e característica estática do sistema em linha contínua (-).	36
5.1	Modelos de blocos interconectados. $f(\cdot)$ representa a função estática não linear e $g(\cdot)$ representa a função dinâmica linear. Em (a) Modelo de Hammerstein e em (b) Modelo de Wiener.	39
5.2	Validação da característica estática do sistema. Característica estática do sistema(-). Característica estática resultante do ajuste por redes neurais (- -).	39
5.3	Função inversa da característica estática. Característica do sistema(-). Característica resultante do ajuste por redes neurais (- -).	40
5.4	Validação do modelo Wiener Neural para o aquecedor de dissipação variável. Dados de validação (-). Saída da rede(- -).	40
5.5	Ajuste da característica estática. Característica do sistema(-). Característica resultante do ajuste por redes neurais (- -).	41
5.6	Validação do modelo Hammerstein Neural para o aquecedor de dissipação variável. Dados de validação (-). Saída da rede(- -).	41
5.7	Validação da característica estática do sistema. Característica estática do sistema(-). Característica estática resultante do ajuste por redes neurais (- -).	42
5.8	Função inversa da característica estática. Característica do sistema(-). Característica resultante do ajuste por redes neurais (- -).	43
5.9	Validação do modelo Hammerstein Neural para o conversor <i>buck</i> CC-CC. Característica do sistema(-). Característica resultante do ajuste por redes neurais (- -).	44
5.10	Ajuste da característica estática. Característica do sistema(-). Característica resultante do ajuste por redes neurais (- -).	45
5.11	Validação do modelo Hammerstein Neural para o conversor <i>buck</i> CC-CC. Característica do sistema(-). Característica resultante do ajuste por redes neurais (- -).	46
5.12	Validação da característica estática do sistema. Característica estática do sistema(-). Característica estática resultante do ajuste por redes neurais (- -).	47

5.13	Função inversa da característica estática. Característica do sistema(-). Característica resultante do ajuste por redes neurais (- -).	47
5.14	Validação do modelo Wiener Neural para a planta de neutralização de pH. Característica do sistema(-). Característica resultante do ajuste por redes neurais (- -).	48
5.15	Ajuste da característica estática. Característica do sistema(-). Característica resultante do ajuste por redes neurais (- -).	49
5.16	Representação do modelo Hammerstein Neural para a planta de neutralização de pH. Característica do sistema(-). Característica resultante do ajuste por redes neurais (- -).	49
5.17	Característica estática: reprodução da característica estática da rede em linha tracejada (- -) e característica estática do sistema em linha contínua (-). (a) Rede 1. (b) Rede 2. (c) Rede 3. (d) Rede 4.	52
5.18	Característica dinâmica: reprodução da característica dinâmica da rede em linha tracejada (- -) e característica dinâmica do sistema em linha contínua (-). (a) Rede 1. (b) Rede 2. (c) Rede 3. (d) Rede 4.	56
5.19	Característica estática: reprodução da característica estática da rede em linha tracejada (- -) e característica estática do sistema em linha contínua (-). (a) Rede 1. (b) Rede 2. (c) Rede 3. (d) Rede 4.	57
5.20	Característica estática: reprodução da característica estática da rede em linha tracejada (- -) e característica estática do sistema em linha contínua (-). (a) Rede 1. (b) Rede 2. (c) Rede 3. (d) Rede 4.	59
5.21	Característica dinâmica: reprodução da característica dinâmica da rede em linha tracejada (- -) e característica dinâmica do sistema em linha contínua (-). (a) Rede 1. (b) Rede 2. (c) Rede 3. (d) Rede 4.	63
5.22	Característica estática: reprodução da característica estática da rede em linha tracejada (- -) e característica estática do sistema em linha contínua (-). (a) Rede 1. (b) Rede 2. (c) Rede 3. (d) Rede 4.	64
5.23	Característica estática: reprodução da característica estática da rede em linha tracejada (- -) e característica estática do sistema em linha contínua (-). (a) Rede 1. (b) Rede 2. (c) Rede 3. (d) Rede 4.	66
5.24	Característica dinâmica: reprodução da característica dinâmica da rede em linha tracejada (- -) e característica dinâmica do sistema em linha contínua (-). (a) Rede 1. (b) Rede 2. (c) Rede 3. (d) Rede 4.	70
5.25	Característica estática: reprodução da característica estática da rede em linha tracejada (- -) e característica estática do sistema em linha contínua (-). (a) Rede 1. (b) Rede 2. (c) Rede 3. (d) Rede 4.	71

Capítulo 1

Introdução

Ao aplicar técnicas de identificação de sistemas, busca-se modelos matemáticos onde pouca ou nenhuma informação, além dos dados de entrada e saída, é conhecida. No caso de somente as informações de entrada e saída do sistema estarem disponíveis, os modelos são conhecidos por caixa preta ou modelo empírico (Corrêa and Aguirre, 2004). Modelos obtidos do conhecimento de todas as equações matemáticas que o descrevem são conhecidos como caixa branca. Neste tipo de identificação, todos os termos da estrutura do modelo e seus parâmetros possuem significado físico. Finalmente, qualquer modelo que não esteja em nenhum destes extremos é obtido por técnicas de identificação caixa cinza (Aguirre, 2007; Corrêa and Aguirre, 2004). Este tipo de identificação apresenta algumas vantagens em relação às outras, pois consegue obter modelos com menor número de parâmetros, maior capacidade de reproduzir características fora dos dados de identificação, maior robustez e maior adequação para o desenvolvimento de sistemas de controle (Corrêa and Aguirre, 2004). Com a identificação caixa cinza, incorporando-se informação auxiliar, ou seja, não somente os dados de entrada e saída, é possível obter um modelo mais global. Neste trabalho, entende-se por modelo mais global aquele que representa a dinâmica e a estática do sistema em uma ampla faixa de operação (Funkquist, 1997).

No processo de identificação é impossível agregar ao modelo todas as características do sistema. Por isso, vale ressaltar que não existe o modelo do sistema, mas sim uma família de modelos com características e desempenhos variados (Aguirre, 2007). Em cada aplicação, um modelo deve ser capaz de reproduzir as características necessárias para o estudo em questão (Margoti et al., 2010). No caso de sistemas não lineares, maior é a dificuldade de se escolher uma estrutura de modelo adequada, uma vez que, devido ao grande número de possibilidades de estruturas para um dado objetivo de modelagem, corre-se o risco de escolher uma estrutura mais complexa do que a requerida (Aguirre, 1997). Nesse contexto, o conhecimento de informação auxiliar contribui significativamente para a melhoria da qualidade do modelo obtido (Abonyi et al., 2001).

No processo de identificação e modelagem, atenção especial é dada a sistemas que apresentam curva estática não-linear, em espaços bidimensionais (Aguirre, 1997). A curva estática não linear fornece indícios do tipo de estrutura que deve ser utilizada para determinado modelo. A partir disto, sistemas com características semelhantes de curva estática podem ser modelados por estruturas semelhantes, ainda que os sistemas sejam completamente diferentes.

Um modelo que represente bem as características em estado estacionário tem, em geral, uma melhor capacidade de representação do sistema real (Barroso et al., 2002). Além disso, é possível acrescentar, como restrição no processo de identificação, a característica de estado estacionário de forma a se obter um modelo que apresente melhor desempenho (Barroso and Nepomuceno, 2004; Amaral, 2001).

1.1 Motivação

Alguns trabalhos na literatura mostram a importância da característica estática do sistema na escolha da estrutura da representação (Barroso and Nepomuceno, 2004; Amaral, 2001; Aguirre, 1997). Por outro lado, pouco se pesquisa a respeito da utilização da característica estática na escolha da representação a ser utilizada.

Diversos são os tipos de representações existentes e utilizadas nos processos de identificação de sistemas. Um deles são as redes neurais (Narendra and Parthasarathy, 1990; Amaral, 2001; Leyva et al., 1997; Kamat et al., 2005), que conseguem aproximar funções não lineares satisfatoriamente. As redes possuem a capacidade de generalização e conseguem resolver problemas complexos (de grande escala) que são atualmente intratáveis (Haykin, 2008). Por meio de seu treinamento, baseado em algoritmos com objetivos de minimização de erros (*back-propagation*) estima-se o erro da camada de saída e o realimenta nas camadas intermediárias da rede (Braga et al., 2007; Haykin, 2008), de forma que, a cada iteração, a função custo, ou seja, a função do erro médio quadrático entre dados coletados e dados da saída da rede, seja minimizado. Algumas variações do algoritmo *back-propagation* são citados em Braga et al. (2007)), tais como *QuickProp* e *RProp*.

Os modelos NARX (*Non-linear AutoRegressive with eXogenous inputs*) e NARMAX (*Non-linear AutoRegressive Moving Average with eXogenous inputs*) são utilizados para representar não linearidades dinâmicas e são modelos baseados em equações de diferenças. Estes modelos discretos explicam o valor de saída $y(t)$ em função de valores prévios dos sinais de saída e entrada (Aguirre, 2007).

Um outro tipo de representação são os modelos de Hammerstein e Wiener, que propõem uma abordagem através da conexão de blocos de acordo com as relações entre a dinâmica e a estática da entrada e saída do sistema. A modelagem de sistemas não lineares utilizando redes neurais em blocos interconectados foi inicialmente proposta por Narendra and Parthasarathy (1990). O modelo de Hammerstein apresenta blocos, em cascata, representando a estática não linear e a dinâmica linear e o modelo de Wiener possui o bloco referente a dinâmica linear conectado ao bloco que representa a estática não linear (Coelho et al., 2002; Bai and Reyland Jr., 2008). Estes modelos de blocos interconectados permitem que o bloco de estática não linear seja constituído por diferentes representações de modelos. Assim pode-se definir modelos de blocos interconectados neurais, polinomiais (Coelho et al., 2002; Bai and Reyland Jr., 2008; Jia et al., 2005; Smith et al., 2007; Kalafatis et al., 2005).

Desta maneira, as representações aqui descritas são utilizadas para obtenção de modelos de três diferentes tipos de sistemas: um eletrônico, um térmico e um químico. A escolha destes sistemas deve-se a diferença entre suas características estáticas e dinâmicas, pretendendo-se assim, buscar a relação existente entre não linearidade e tipo de representação a ser utilizada.

A identificação e modelagem do sistema térmico (aquecedor) foi proposta, inicialmente, em Cassini (1999). A escolha deste sistema baseia-se no fato de se tratar de dados reais, sendo os dados dinâmicos coletados em toda a faixa de operação dos dados que representam a característica estática do sistema e ter sido amplamente estudado pelo grupo MACSIN¹. Além de Cassini (1999), que propôs modelos polinomiais NARMAX, por meio de estimação recursiva, Amaral (2001) utilizou os mesmos dados coletados para demonstrar os resultados obtidos por meio do treinamento de redes neurais. Em Amaral (2001) foram abordadas a identificação caixa preta, com técnicas para poda de redes, e caixa cinza, usando treinamento com restrição para obtenção dos modelos.

O sistema eletrônico é um conversor CC-CC *buck*. Esse sistema já foi modelado por diversos autores. Aguirre et al. (2000) propuseram uma comparação da superioridade da utilização de informação auxiliar para modelagem deste sistema; Alonge et al. (2007) utilizaram o modelo de blocos interconectados de Hammerstein, com identificação caixa preta e Leyva et al. (1997) utilizou redes neurais no processo de modelagem. Nestes trabalhos, os resultados foram satisfatórios e em Alonge

¹Modelagem Análise e Controle de Sistemas Não-Lineares - <http://www.cpdee.ufmg.br/MACSIN/>

et al. (2007), onde foi utilizada a identificação caixa preta, o processo se mostrou satisfatório para uma faixa de operação. A partir destes resultados, pretende-se verificar o comportamento de representações de blocos interconectados.

O sistema químico, que consiste em uma planta de neutralização de pH, é um processo comum e muito importante em indústrias químicas e de biotecnologia e que desafia sua identificação, modelagem e controle por ser um problema com estática não linear bastante complexa e que varia sensivelmente sob pequenas mudanças nas condições do processo (Yoo et al., 2004). Por sua relevância e complexidade, muitos pesquisadores, tais como Campos (2007); Syafiea et al. (2007); Vale et al. (2008); Mahmoodi et al. (2009); Henson and Seborg (1994); Kalafatis et al. (2005); Smith et al. (2007); Tan et al. (2005); Gómez et al. (2004) entre outros, vem estudando e desenvolvendo modelos, a partir de diferentes abordagens. De acordo com as características estáticas e dinâmicas do processo, a grande maioria dos estudos desenvolvidos propõe a utilização dos modelos de blocos interconectados, como a configuração dos modelos de Hammerstein e Wiener. Constatou-se por meio de resultados, nos estudos que foram desenvolvidos, que o processo de pH é modelado para uma faixa de operação muito específica e, também, modelos que representam bem a estática não linear, não vislumbram totalmente a modelagem da dinâmica linear, ou vice versa. Nota-se também que grande parte dos autores utilizam o processo de identificação caixa preta.

1.2 Objetivo

As representações objeto de estudo desse trabalho são redes neurais perceptron de múltiplas camadas (MLP - *Multilayer Perceptron*) e modelos de blocos interconectados (Hammerstein e Wiener). A partir de três estudos de caso, pretende-se avaliar como as características estáticas do sistema podem auxiliar na escolha do tipo de representação. Para isso os sistemas serão modelados a partir de procedimento caixa preta e caixa cinza.

1.3 Apresentação do trabalho

Este trabalho foi desenvolvido como se segue:

No Capítulo 2 são apresentados conceitos sobre identificação de sistemas e sobre a representação de sistemas não lineares. Neste contexto, serão descritas as identificações caixa preta e caixa cinza, com a utilização de redes neurais. Realizando uma descrição mais detalhada da utilização de redes no processo de modelagem, demonstrando como a utilização de informação auxiliar pode melhorar os resultados obtidos, por meio do treinamento com restrição e dos modelos por blocos interconectados.

O Capítulo 3 apresenta os três sistemas a serem estudados e modelados, bem como suas características estáticas e dinâmicas. Os estudos de caso serão realizados em um sistema térmico (aquecedor), sistema eletrônico (conversor *buck*) e um sistema químico (planta de neutralização de pH).

No Capítulo 4 são apresentados os resultados da identificação caixa preta dos três sistemas estudados.

Por sua vez, o Capítulo 5 trará os modelos obtidos, para cada um dos sistemas, por meio da utilização dos blocos interconectados de Hammerstein e Wiener e por meio do treinamento com restrição.

No último capítulo deste trabalho, os resultados obtidos são avaliados e sugere-se algumas propostas para trabalhos futuros.

Capítulo 2

Identificação de Sistemas

Diversas são as possibilidades para obtenção de modelos de um sistema, seja pela física do processo (modelagem caixa branca) ou por meio de técnicas de identificação (modelagem caixa preta e caixa cinza), com o objetivo de descrever as relações existentes entre as variáveis do processo. A modelagem caixa branca requer que sejam conhecidas as leis físicas do processo, exigindo conhecimento aprofundado do sistema. Porém o levantamento e formulação matemática de todos os fenômenos que afetam o comportamento de um sistema é tarefa extremamente complexa. Desta maneira, torna-se impossível que o modelo reproduza exatamente o comportamento do sistema original (Coelho, 2001). Já na modelagem por meio de identificação, são utilizadas técnicas em que se encontra modelos de sistemas por meio de dados experimentais, não exigindo, contudo, o conhecimento prévio do processo (Cassini, 1999), modelagem caixa preta (Sjöberg et al., 1995; Choi et al., 1998). Caso haja conhecimento de alguma parte do processo, esta pode ser utilizada de maneira a facilitar o processo de identificação, introduzida como informação auxiliar, modelagem caixa cinza (Alonge et al., 2007; Amaral, 2001; Aguirre et al., 2000).

A identificação de sistemas é um processo alternativo. Com poucas informações a respeito de um sistema, por exemplo dados de entrada e saída, busca-se um modelo capaz de explicar a relação existente entre os dados, ou pelo menos aproximar esta relação. Os procedimentos utilizados para identificação de um sistema serão descritos a seguir.

2.1 Procedimentos para identificação de sistemas

Segundo Aguirre (2007), as principais etapas de um processo de identificação são:

1. Testes dinâmicos e coleta de dados - a proposta da identificação de sistemas é buscar uma relação existente entre dados disponíveis de um sistema. Portanto, é necessário obter tais dados com a finalidade de se obter modelos. Neste trabalho, serão utilizados dados obtidos dos sistemas descritos no Capítulo 3.
2. Escolha da representação matemática a ser utilizada - diversos são os tipos de representações existentes. Neste trabalho emprega-se o uso de Redes Neurais Artificiais e, também, os Blocos Interconectados de Hammerstein e Wiener como modelos do sistema. Os modelos resultantes serão do tipo NARX (*Non-linear AutoRegressive with exogenous inputs*).
3. Determinação da estrutura do modelo - Para o caso das Redes Neurais Artificiais, a escolha da estrutura significa definir o número de neurônios a ser utilizado, o número de camadas, o tipo de função de ativação, as conexões necessárias e quais regressores utilizar.

4. Estimação de parâmetros - esta etapa é realizada por meio de algoritmos específicos para cada representação escolhida. Mínimos Quadrados é um exemplo de método para obtenção de parâmetros dos modelos, por exemplo. Os modelos neurais deste trabalho foram obtidos por meio do algoritmo de Levenberg-Marquardt, para minimização do erro de treinamento.
5. Validação do modelo - após a obtenção dos modelos, é necessário verificar se eles incorporam as características de interesse do sistema original. Sendo esta uma etapa subjetiva, pois o resultado da validação dependerá da aplicação desejada ao modelo obtido.

2.2 Representações de sistemas não lineares

Diversos são os tipos de representação para sistemas não lineares. Contudo, a utilização de cada uma delas, depende, não só da finalidade do modelo obtido, como também da ferramenta disponível para utilização.

Destacam-se neste trabalho a utilização de redes neurais (Braga et al., 2007; Haykin, 2008; Amaral, 2001), para identificação caixa preta e caixa cinza. Obtendo-se modelos caixa cinza, a partir das representações de blocos interconectados de Hammerstein e Wiener e pelo treinamento de redes neurais com restrição. Tais modelos serão descritos por equações NARX e NARMAX (Barroso, 2001; Norquay et al., 1998; Mahmoodi et al., 2009; Coelho, 2001). A obtenção de modelos por meio de redes neurais artificiais será detalhada na Seção 2.2.3.

2.2.1 Modelos ARX, NARX e NARMAX

O modelo ARX (*AutoRegressive with eXogenous inputs*) é um tipo de representação utilizada para sistemas lineares discretos, que relacionam variáveis de entrada e saída do sistema. Podendo ser escrito na forma da Equação (2.1),

$$A(q)y(k) = B(q)u(k) + e(k), \quad (2.1)$$

sendo $A(q)$ e $B(q)$ polinômios arbitrários, $u(k)$ e $y(k)$ vetores de entrada e saída do sistema no instante k , respectivamente e $e(k)$ representa o ruído.

Os modelos NARX (*Nonlinear AutoRegressive with eXogenous inputs*), são modelos discretos no tempo, que explicam o valor da saída $y(k)$ em função de valores prévios dos sinais de entrada e saída, representado de maneira geral pela Equação (2.2).

$$y(k) = f[y(k-1), \dots, y(k-n_y), u(k-\tau_d), \dots, u(k-n_u)], \quad (2.2)$$

sendo que n_y , n_u e τ_d são, respectivamente, máximos atrasos em y e em u e atraso puro de tempo.

A fim de evitar a polarização dos parâmetros que compõem o modelo, podem ser incluídos termos de ruído. Quando estes termos de ruído são incluídos no modelo NARX, ele passa a ser chamado de NARMAX (*Non-linear AutoRegressive Moving Average with eXogenous inputs*). Os modelos NARMAX também relacionam a não linearidade existente entre o par entrada-saída passadas para a saída atual (Sjöberg et al., 1995). De maneira geral, o modelo NARMAX pode ser representado pela Equação (2.3).

$$y(k) = f[y(k-1), \dots, y(k-n_y), u(k-\tau_d), \dots, u(k-n_u), e(k), e(k-1), \dots, e(k-n_c)]. \quad (2.3)$$

Em que $e(k)$ representa os efeitos que não podem ser bem representados pela função f e n_c é o maior atraso do modelo do ruído.

2.2.2 Modelos de Hammerstein e Wiener

Os modelos de Hammerstein e Wiener (Crama and Schoukens, 2004; Figueroa et al., 2008; Wang and Ding, 2008; Biagiola and Figueroa, 2009; Santos et al., 2010; Margoti et al., 2010) são representações de sistemas baseados em dois blocos interconectados. De forma que um dos blocos representa a dinâmica linear e o outro, a estática não linear (Coelho, 2001). O que diferencia estas duas representações é a posição de cada um destes blocos. Como pode ser visto na Figura 2.1 nos blocos de Hammerstein (Figura 2.1(a)), o bloco estático precede o bloco que representa a dinâmica linear. Já no modelo de Wiener (Figura 2.1(b)), a dinâmica linear precede a estática não linear.

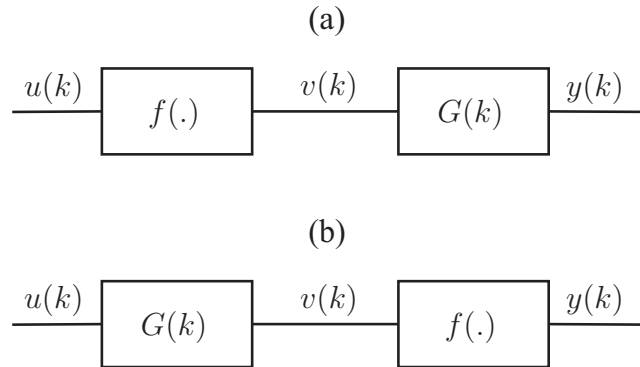


Figura 2.1: Modelos de blocos interconectados. $f(\cdot)$ representa a função estática não linear e $g(\cdot)$ representa a função dinâmica linear. Em (a) Modelo de Hammerstein e em (b) Modelo de Wiener.

O bloco estático não linear é representado por uma função estática não linear, denotada por $f(\cdot)$. O bloco dinâmico é representado por um modelo dinâmico, representado por $G(k)$. Para representar a dinâmica linear destes modelos, tipicamente escolhe-se o modelo ARX (*autoregressive model with exogenous variables*) (Norquay et al., 1998).

Entre os modelos não lineares, os modelos de blocos interconectados são estruturas eficientes em modelagem não linear (Arefi et al., 2008). A maior dificuldade desta representação é a obtenção do sinal intermediário aos blocos, fazendo-se necessária a utilização de uma informação prévia e/ou uma restrição para estimar os parâmetros da estrutura detectada (Pearson and Pottman, 2000). O sinal $v(k)$, que representa o sinal intermediário, não possui significado físico (Arefi et al., 2008) e não está disponível.

Para a obtenção do sinal intermediário $v(\hat{k})$ (sinal estimado) no modelo de Wiener, a função estática deve ser inversível para que os parâmetros do modelo dinâmico linear possam ser estimados. Desta maneira, $v(\hat{k})$ pode ser estimado através da inversa da função $f(\cdot)$.

$$y(k) = f(\hat{v}(k)), \quad (2.4)$$

$$\hat{v}(k) = g(y(k)), \quad (2.5)$$

Sendo $\hat{v}(k)$ a inversa da função $y(k)$.

De acordo com Coelho (2001), o bloco dinâmico linear, que será representado por um modelo ARX, com o par de entrada e saída sendo $u(k)$ e $v(\hat{k})$, respectivamente, pode ser escrito como mostra a Equação (2.6):

$$v(k) = \sum_{j=1}^{n_v} \theta_j v(k-j) + \sum_{i=1}^{n_u} \sigma_i u(k-i). \quad (2.6)$$

Substituindo-se a Equação (2.4) em (2.6):

$$y(k) = f \left(\sum_{j=1}^{n_v} \theta_j v(k-j) + \sum_{i=1}^{n_u} \sigma_i u(k-i) \right). \quad (2.7)$$

Considerando um atraso j a Equação (2.5) ficará:

$$v(k-j) = g(y(k-j)). \quad (2.8)$$

Substituindo-se, então, a Equação (2.8) em (2.7), o modelo polinomial da representação de Wiener será obtida com relação aos sinais de entrada, $u(k)$, e saída, $y(k)$, do sistema:

$$y(k) = f \left(\sum_{j=1}^{n_y} \theta_j g(y(k-j)) + \sum_{i=1}^{n_u} \sigma_i u(k-i) \right). \quad (2.9)$$

Conclui-se então que, o modelo de Wiener pode conter termos cruzados do tipo $y(k-j)^m u(k-i)^{m-p}$, sendo i e j os atrasos de cada um destes termos e m e $(m-p)$ o grau de não linearidade dos termos obtidos. Porém não pode conter termos cruzados entre termos de mesma natureza, entre termos de entradas ou entre os termos de saída, como se segue (Coelho, 2001):

1. $u(k-i_1)^p u(k-i_2)^{m-p}$, para $i_1 \neq i_2$ e $1 \leq p \leq m$;
2. $y(k-i_1)^p y(k-i_2)^{m-p}$, para $j_1 \neq j_2$ e $1 \leq p \leq m$;
3. $y(k-i)^p y(k-j)^{m-p}$, para $i \neq j$ e $1 \leq p \leq m$.

No modelo de Hammerstein, o sinal intermediário, $v(k)$, é obtido pelo mapeamento do sinal de entrada, $u(k)$, por meio da função $f(\cdot)$, em que:

$$\hat{v}(k) = f(u(k)), \quad (2.10)$$

ou na forma normal:

$$\hat{v}(k-i) = f(u(k-i)), i = 1 \dots n_u, \quad (2.11)$$

em que n_u representa o máximo atraso da saída.

O modelo ARX é obtido utilizando o par de entrada e saída, $v(k)$ e $y(k)$, respectivamente, do bloco dinâmico linear, tal que:

$$y(k) = \sum_{j=1}^{n_y} \theta_j y(k-j) + \sum_{i=1}^{n_v} \sigma_i v(k-i), \quad (2.12)$$

sendo:

n_y - o atraso máximo da saída do modelo ARX;

n_v - o atraso máximo da entrada do modelo ARX;

θ_j - parâmetros relacionados aos regressores de saída do modelo ARX;

σ_i - parâmetros relacionados aos regressores de entrada do modelo ARX.

Substituindo-se (2.10) em (2.12):

$$y(k) = \sum_{j=1}^{n_y} \theta_j y(k-j) + \sum_{i=1}^{n_v} \sigma_i f(u(k-i)). \quad (2.13)$$

De acordo com a Equação (2.13), cada termo de ordem m , tal que $0 \leq m \leq l$, contém um fator de ordem 1 em $y(k - j)$ e um fator de ordem m em $u(k - i)$, sendo que l é o grau de não linearidade que pode ter o polinômio obtido do modelo de Hammerstein. Desta maneira:

1. a não linearidade atua somente nos regressores da entrada;
2. não há termos do tipo $u(k - i)^{m-q}u(k - i_1)^m$, com $i \neq i_1$, $0 \leq q \leq m$ e $(m - q) \leq l$.

Sendo assim, a não linearidade do modelo de Hammerstein não atua em termo com atrasos diferentes.

Ainda de acordo com as conclusões do trabalho de Coelho (2001), o modelo de Wiener pode ser utilizado para sistemas do tipo SISO (*single input single output*), enquanto que o modelo de Hammerstein admite sistema que possuem mais de uma entrada porém somente uma saída, MISO (*multiple input single output*), ambos em estado estacionário.

O conhecimento prévio da curva estática, para cada um dos pontos de operação de interesse, seja por meio do conhecimento teórico da curva característica em estado estacionário do sistema em estudo, é uma das maneiras de se obter as funções $f(\cdot)$ e $g(\cdot)$ *a priori*. Desta maneira, as estimativas de $\hat{v}(k)$ se tornam mais fáceis, uma vez que as funções podem ser obtidas por ajuste de curva estática aos dados estáticos de entrada e saída do sistema. A única restrição é a existência da função $g(\cdot)$ (Santos et al., 2010).

2.2.3 Redes Neurais

Redes Neurais Artificiais (RNA's) são modelos matemáticos compostos por neurônios, unidades de processamento, que calculam algumas funções, normalmente não lineares. Estas redes tem capacidade de aprender por meio de exemplos e de generalizar a informação aprendida. São compostas por neurônios interconectados que tentam reproduzir as funções das redes de neurônios biológicas (Braga et al., 2007), buscando implementar seu comportamento funcional e sua dinâmica. O primeiro modelo artificial de um neurônio biológico surgiu na década de 40, porém foi na década de 80 que houve um grande interesse pelas RNAs na comunidade internacional (Braga et al., 2007), por consequência do avanço tecnológico.

A capacidade computacional das redes neurais deve-se, primeiramente, à sua estrutura paralelamente distribuída e também à sua capacidade de generalização (Haykin, 2008). Sendo um tipo de representação não linear, a rede neural é muito útil caso o sinal de entrada seja inerentemente não linear.

As RNAs são motivadas pela analogia com o cérebro humano. Desta maneira, as redes neurais são compostas geralmente de:

- neurônios, unidades centrais de processamento de informação;
- um conjunto de sinapses ou elos de conexões, caracterizadas pelas conexões (pesos sinápticos) entre sinal de entrada e neurônios, ou entre neurônios de diferentes camadas. Estes pesos podem assumir valores tanto negativos quanto positivos;
- somador, que pondera os sinais de entrada pelas sinapses dos neurônios; e,
- função de ativação, que restringe o intervalo de amplitude do sinal de saída;
- *bias*, representado por b , que aplicado externamente tem o efeito de aumentar ou diminuir a entrada da função de ativação.

A saída de um único neurônio com n entradas é descrita pela Equação (2.14):

$$y = f \left(\sum_{j=1}^n w_j x_j + b \right), \quad (2.14)$$

sendo que b (bias) e w_j são constantes e f é chamada de função de ativação. A função de ativação é responsável por gerar a saída do neurônio a partir da soma das entradas ponderadas pelos pesos e o termo de polarização (bias), w_j . Neurônios individuais possuem capacidade computacional limitada. Porém, um conjunto de neurônios artificiais, conectados em forma de uma rede, são capazes de resolver problemas complexos (Braga et al., 2007; Haykin, 2008).

A Figura 2.2 mostra o modelo de um neurônio.

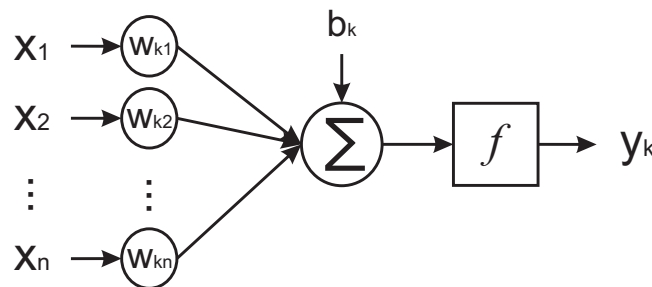


Figura 2.2: Modelo de um neurônio artificial.

sendo x_1, x_2, \dots, x_n os sinais de entrada; $w_{k1}, w_{k2}, \dots, w_{kn}$ os pesos sinápticos do neurônio, em que o índice k indica o neurônio em questão e o segundo índice, o terminal de entrada; b_k representa bias; f é a função de ativação; e y_k é o sinal de saída do neurônio.

De forma análoga, a Equação (2.15) descreve uma rede neural com uma única camada oculta (Figura 2.3).

$$y(k) = f_0 \left(b_0 + \sum_{j=1}^{N_j} w_j^0 f_j \left(b_j + \sum_{i=1}^{N_a} w_{ji}^1 u(k-i) \sum_{v=1}^{N_b} w_{jv}^1 y(k-v) \right) \right), \quad (2.15)$$

sendo:

y_k - saída da rede;

$u(k-i)$ - i -ésimo atraso da entrada $u(k)$;

$y(k-v)$ - v -ésimo atraso da saída $y(k)$;

w_{ji}^1 peso da conexão entre a i -ésima entrada e o j -ésimo neurônio da camada oculta;

w_{jv}^1 peso da conexão entre a v -ésima entrada e o j -ésimo neurônio da camada oculta;

w_j^0 peso da conexão entre a j -ésima neurônio da camada oculta e o neurônio da camada de saída;

b - termo de polarização (bias), b_j para os neurônios da camada oculta e b_0 para o neurônio da camada de saída;

N_j - número de neurônios da camada oculta;

N_a e N_b - número máximo de atrasos da entrada e saída respectivamente;

f_0 - função de ativação da camada de saída;

f_j - função de ativação da camada oculta.

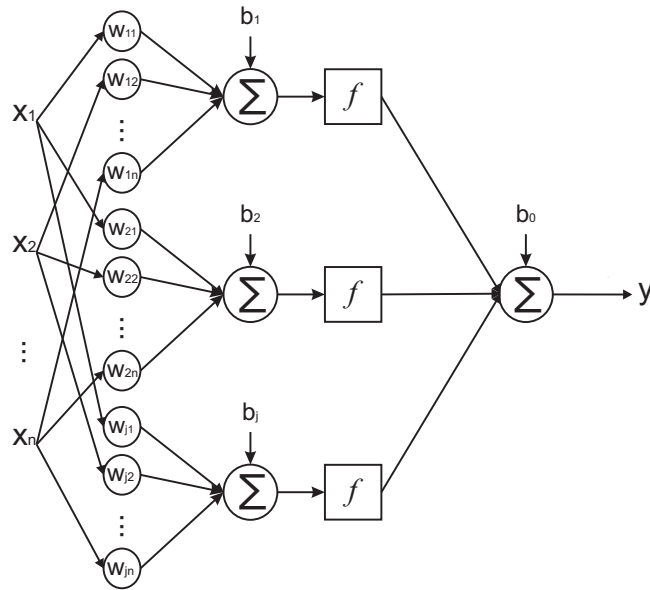


Figura 2.3: Rede neural com uma camada oculta.

Funções de ativação

Os três tipos básicos de função de ativação (Haykin, 2008), representada por f na Figura 2.2, são:

- Limiar (Figura 2.4). Definida por:

$$f = \begin{cases} 1 & \text{se } v \geq 0, \\ 0 & \text{se } v < 0. \end{cases}$$

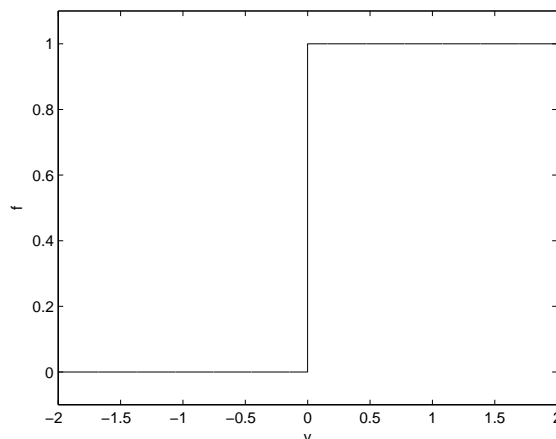


Figura 2.4: Função de Limiar.

- Linear por partes (Figura 2.5). Definida por:

$$f = \begin{cases} 1 & \text{se } v \geq +\frac{1}{2}, \\ v & \text{se } +\frac{1}{2} > v > -\frac{1}{2}, \\ 0 & \text{se } v \leq -\frac{1}{2}. \end{cases}$$

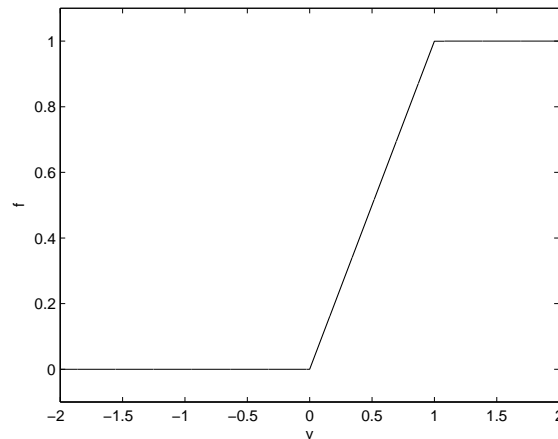


Figura 2.5: Função linear por partes.

- Sigmóide (sigmoidal) - tangente hiperbólico (Figura 2.6). Definida por:

$$f = \frac{1 - e^{-x}}{1 + e^{-x}}.$$

sendo a o parâmetro de inclinação da função sigmóide.

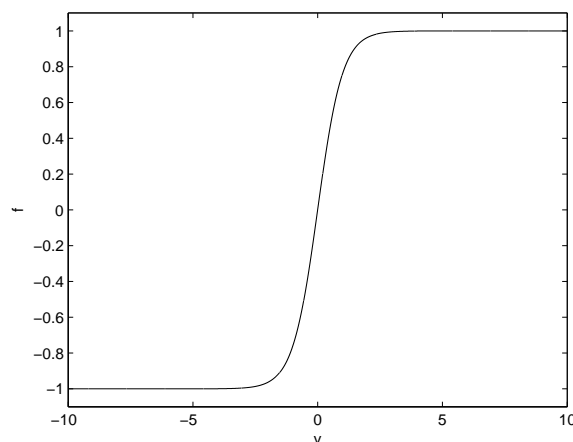


Figura 2.6: Função sigmóide (tangente hiperbólico).

Aprendizado da rede neural

Aprendizado é o processo pelo qual os parâmetros livres de uma rede neural são ajustados por meio de uma forma continuada de estímulo pelo ambiente externo, sendo o tipo específico de aprendizado definido pela maneira particular como ocorrem os parâmetros livres (Braga et al., 2007; Haykin, 2008). Um exemplo de critério para treinamento de uma rede é o erro médio quadrático da resposta da rede em relação à massa de dados fornecida. Por meio do treinamento, o aprendizado ocorre à medida que o erro médio quadrático diminui.

O algoritmo *backpropagation* (retropropagação) permite treinar redes com múltiplas camadas (redes MLP - *Multilayer Perceptron*). As redes MLP possuem função de ativação sigmoidais em sua

camada oculta e permitem aproximar qualquer função contínua, com uma camada oculta, ou qualquer função, com duas camadas (Cybenko, 1989).

Por se apresentar como um processo lento em problemas mais complexos, para treinamento de redes, o algoritmo de retropropagação possui algumas variações que, desde sua criação, foram propostas para acelerar o tempo de treinamento das redes neurais e melhorar o desempenho. Uma destas variações é o algoritmo *Levenberg-Marquardt* que é uma aproximação do método de quasi-Newton (Fun and Hagan, 1996).

Este algoritmo propõe o ajuste dos pesos por meio da matriz hessiana, de acordo com a Equação (2.16).

$$\Delta w_{ji}(t) = - [\nabla^2 E(w_{ji}(t)) + \mu I]^{-1} \nabla E(w_{ji}(t)), \quad (2.16)$$

sendo $\nabla^2 E(w_{ji}(t))$ a matriz Hessiana e $\nabla E(w_{ji}(t))$ o gradiente (Braga et al., 2007).

Estruturas de redes neurais

No trabalho pioneiro de Narendra and Parthasarathy (1990) foram propostas quatro estruturas para redes neurais, como pode ser visto na Figura 2.7. Nestas topologias, u e y representam, respectivamente, os regressores de entrada e saída, e as funções representadas por F são funções não lineares.

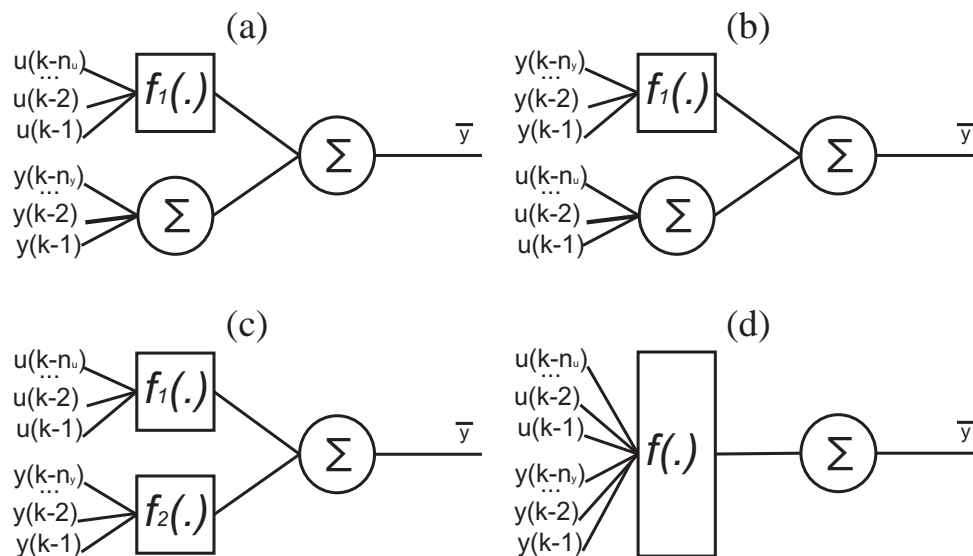


Figura 2.7: Estruturas propostas por Narendra and Parthasarathy (1990).

A partir das estruturas propostas por Narendra and Parthasarathy (1990) foram levantadas as restrições necessárias para o treinamento proposto no Capítulo 5, Seção 5.2.

Imposição de restrições às redes neurais

Em Amaral (2001) foi proposta a imposição de restrição ao treinamento da rede neural para o sistema do aquecedor de dissipação variável. Em seu trabalho foi utilizada a estrutura (a) da Figura 2.7, proposta por Narendra and Parthasarathy (1990). O treinamento de uma rede neural com imposição de restrição pode ser realizado como se segue:

- A partir da determinação da estrutura de treinamento para a característica estática do sistema, obtém-se o modelo que o representa;

Dado como exemplo, na forma vetorial, a Equação (2.17):

$$\bar{y} = w_2 (\text{tansig}(w_1 \bar{u} + b_1)) + b_2. \quad (2.17)$$

Sendo:

w_1 e w_2 - matrizes que representam os pesos das camadas de entrada e saída, respectivamente;
 b_1 e b_2 - vetores que representam o termo *bias*, das respectivas camadas.

- Reescreve-se a equação dinâmica em sua forma estática, de maneira que todos os regressores de entrada e saída sejam substituídos por \bar{u} e \bar{y} , respectivamente. De maneira que $x(k-1) = x(k-2) = \dots = x(k-n) = \bar{x}$ e $y(k-1) = y(k-2) = \dots = y(k-n) = \bar{y}$.
- Utiliza-se os pesos obtidos na curva estática para treinar a rede dinâmica obtida como um problema de otimização. De maneira que o erro médio quadrático, entre os dados do sistema e a saída da rede, seja minimizado.
- Valida-se o modelo tanto do ponto de vista estático quanto dinâmico.

2.3 Validação dos Modelos

Conforme descrito no início deste capítulo, a validação do modelo é uma etapa importante no processo de identificação, uma vez que, com ela, será verificado se o modelo incorpora as características do sistema, para a finalidade que foi desenvolvido. Para tanto, são propostos os seguintes passos para validação dos modelos:

1. simulação infinitos passos à frente (simulação livre), para a característica dinâmica do sistema em estudo. Para tanto, são utilizados dados diferentes daqueles utilizados para obtenção do modelo (Sjöberg et al., 1995);
2. capacidade do modelo em reproduzir a característica estática;
3. comparação entre a predição do modelo com o sinal de saída, por meio do índice RMSE (*root mean square error*),

O erro RMSE é dado pela equação (2.18):

$$RMSE = \left\{ \frac{1}{N} \sum_{k=1}^N [\hat{y}_i(k) - y_i(k)]^2 \right\}^{\frac{1}{2}}, \quad (2.18)$$

sendo $\hat{y}_i(k)$ a saída do modelo na predição infinitos passos à frente e $y_i(k)$ o valor real.

4. inspeção visual, numa validação qualitativa, da curva estática do sistema e a do modelo obtido;
5. análise da correlação do resíduo, η , na identificação um passo à frente. Sendo que, para a análise dos dados dinâmicos, este coeficiente é considerado satisfatório se, entre o sinal de saída simulada e o ruído, ele for, em módulo, mais próximo de zero. Para a característica estática, o valor deste coeficiente de correlação deve estar, em módulo, entre 0 e 1, e quanto maior for a correlação entre a saída estática do sistema e a saída estática simulada pelo modelo, indica um bom ajuste desta característica (Coelho, 2001).

Para os casos de validação por meio da função de correlação, este índice foi aplicado aos resíduos de identificação infinitos passos à frente (Aguirre et al., 2004). Neste caso, deseja-se que os resíduos de identificação sejam ruído branco, ou seja, que a sua autocorrelação seja estatisticamente zero. Desta maneira, é possível conjecturar que o modelo foi capaz de explicar os dados dinâmicos sem incorporar dinâmicas espúrias (Santos et al., 2010). Já no ajuste estático, quanto maior for a correlação entre o sinal de saída do sistema e o sinal estimado do modelo (saída), indicará melhor qualidade no ajuste desta característica.

O índice de correlação entre dois sinais quaisquer pode ser definido pela equação (2.19).

$$r_{xy} = \frac{\sigma_{xy}}{\sigma_x \sigma_y}, \quad (2.19)$$

sendo σ_{xy} a covariância entre x e y e σ_x e σ_y os desvios padrão de x e y , respectivamente.

Capítulo 3

Descrição dos Sistemas Modelados

Neste capítulo serão apresentadas as descrições e características de cada um dos três sistemas identificados. Em todos os casos, os testes foram realizados em laboratório ou de forma analítica, anteriormente, por seus idealizadores. Coube, a este trabalho, a utilização dos dados e obtenção de modelos baseados em identificação caixa preta e caixa cinza, com a utilização de redes neurais.

3.1 Aquecedor com Dissipação Variável

A identificação e modelagem do sistema térmico (aquecedor) foi proposta em Cassini (1999). A Figura 3.1 apresenta o diagrama que representa o sistema do aquecedor, a ser identificado.

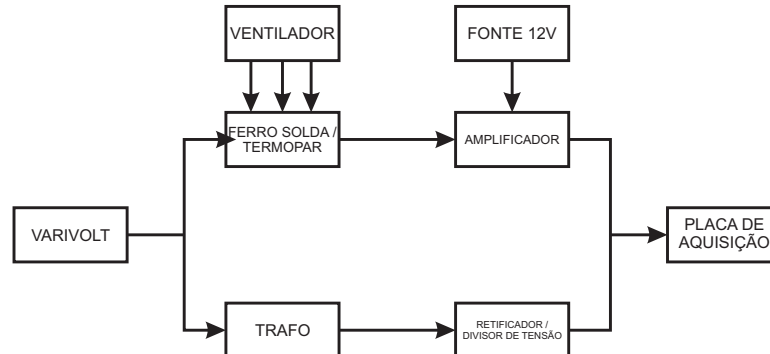


Figura 3.1: Diagrama de blocos do sistema (Aquecedor).

A entrada do sistema é tensão elétrica no divisor de tensão e a saída, tensão elétrica do circuito amplificador. A faixa de variação do divisor de tensão vai de 0 a 127 V, de forma proporcional à uma variação, na entrada do sistema de 0 a 5 V. A placa de aquisição de dados utilizada por Cassini (1999) é do tipo PLC 711s.

O conjunto transformador-retificador-divisor de tensão objetiva condicionar o sinal de entrada do sistema de forma a se obter um outro sinal na faixa de medição da placa. Portanto, o sinal lido pela placa não é o sinal de tensão aplicado na entrada do sistema, mas é representativo do mesmo (Cassini, 1999).

O varivolt foi utilizado para variar a tensão de entrada à faixa de 0 a 136V. Que, ao passar pelo transformador, foi reduzida à faixa de 0 a 18V, para ser retificada. Foi ainda utilizado um ventilador, alimentado em 127V, para resfriar o ferro de solda, a fim de variar o ganho estático do sistema.

Teste Estático

Realizado para o levantamento da curva estática do sistema. Consistiu em aplicar um valor fixo na entrada do sistema e, aguardando que o sistema atingisse a estabilidade, fosse medido o valor da saída.

O teste realizado por Cassini (1999) teve duração de três horas e as grandezas estão indicadas em pu, sendo que 1 pu corresponde a 136V na tensão de entrada e a 998,51° C no ferro de solda. O resultado deste teste pode ser verificado na Figura 3.2.

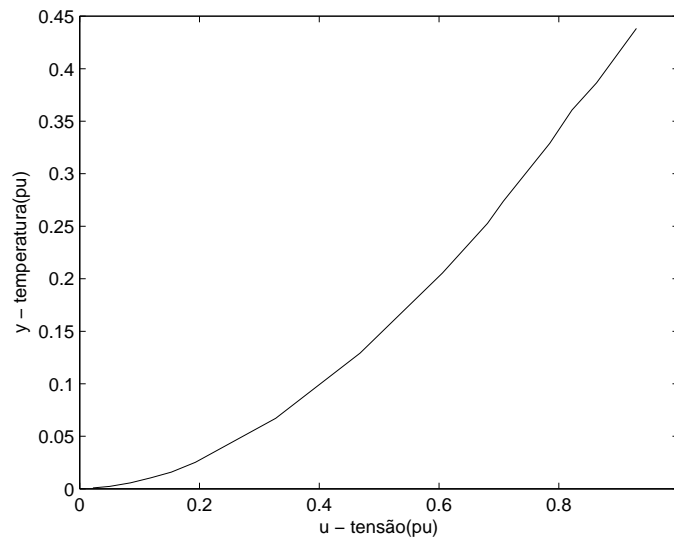


Figura 3.2: Característica estática obtida para o sistema com o ventilador ligado.

Teste Dinâmico

O teste dinâmico foi realizado por um período de quatro horas e os dados, coletados com uma taxa de amostragem de 6s, um total de 2520 pontos. Cada patamar do sinal de excitação do sistema foi mantido por 1 min gerando 10 amostras por patamar, aproximadamente.

A característica dinâmica dos dados de entrada e saída do sistema podem ser verificadas nas Figuras 3.3 e 3.4, respectivamente. Desta massa de dados, metade serão utilizadas no processo de identificação do modelo e a outra metade, para validação do mesmo.

Os dados dinâmicos foram coletados em toda a faixa de operação da característica estática coletada deste sistema.

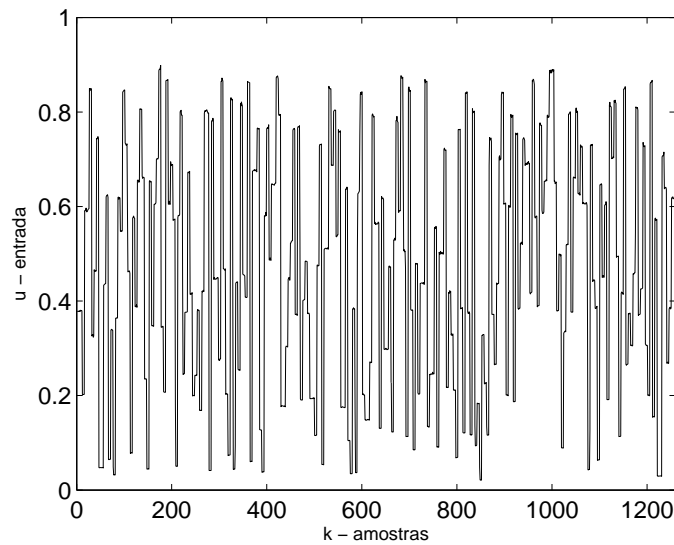


Figura 3.3: Dados de entrada, em pu.

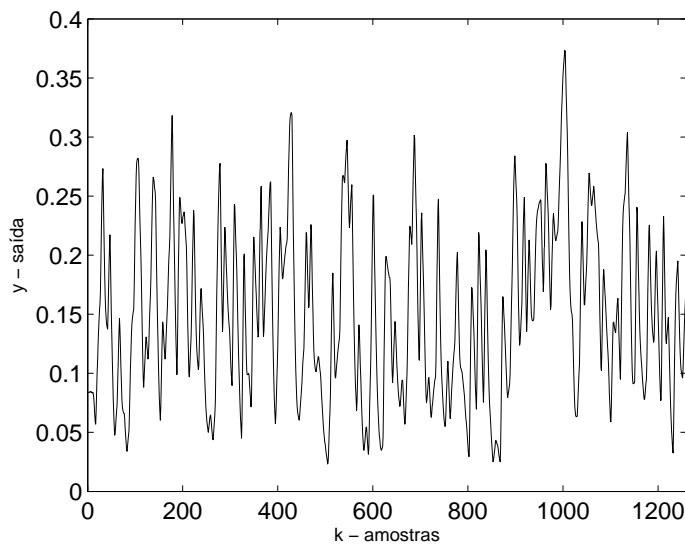


Figura 3.4: Dados de saída, em pu.

Resultados da literatura

Em Cassini (1999), que utilizou estimação recursiva para obtenção dos parâmetros dos modelos polinomiais NARMAX, obteve-se índices de validação RMSE entre 0,2363 e 0,2512 para a mesma massa de dados que será utilizada no presente trabalho. A variação deste índice de validação ocorreu devido à variação do grau de não linearidade dos polinômios obtidos, que partiu de polinômios lineares, que não apresentaram resultado satisfatório, e do número de atrasos de entrada e saída, durante o processo de obtenção dos modelos.

Amaral (2001) propôs a identificação deste mesmo processo utilizando o treinamento de redes neurais, por meio das abordagens caixa preta e caixa cinza, treinamento com restrição. Os índices RMSE obtidos no treinamento caixa preta para os dados dinâmicos, devido às técnicas de poda adotadas em seu trabalho, ficaram entre 0,0141 e 0,0223. No treinamento com restrição, utilizando uma

das estruturas propostas em Narendra and Parthasarathy (1990), o mesmo índice de validação foi 0,0196.

Por meio da obtenção de modelos baseados nos blocos interconectados de Hammerstein e Wiener, Coelho (2001) obteve índices de validação RMSE, para os dados dinâmicos, respectivamente, 0,0933 e 0,2816. Conseguindo índices de correlação para a característica estática em 0,9990, no modelo de Hammerstein e 0,9990, no de Wiener.

3.2 Conversor *Buck*

A Figura 3.5 mostra a estrutura de um conversor CC-CC *buck*, operando em modo contínuo. Durante os testes a fonte CC de alimentação foi mantida constante, a 24V. A razão cíclica é definida como a proporção de tempo em que a chave está ligada em relação ao período de operação, T , ou seja, $D=T_{ligado}/T$. A razão cíclica foi variada usando-se técnicas de modulação por largura de pulso (PWM) a uma taxa de $1/T=33\text{kHz}$, via circuito integrado LM3524 (Aguirre, 2007).

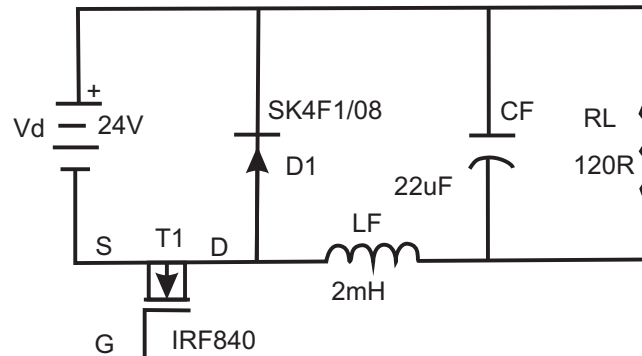


Figura 3.5: Conversor estático *buck*.

Quando a razão cíclica está perto da unidade, a corrente através do indutor e a tensão na carga, v_0 , aumentam, pois a fonte v_d energiza a malha formada por ela, o capacitor e o indutor. Quando $D \rightarrow 0$, a tensão de saída, v_0 diminui segundo um regime dinâmico diferente, evidenciando, aspectos não lineares do conversor.

Característica Estática

Qualquer que seja ao modelo obtido para o conversor CC-CC *buck*, é desejável que ele incorpore a relação estática do conversor, dada por (Aguirre, 2007):

$$\begin{aligned}
 v_0 &= D v_d, \\
 v_0 &= (1 - D') v_d = \left(1 - \frac{\bar{u} - 1}{3}\right) v_d, \\
 v_0 &= \frac{4v_d}{3} - \frac{v_d}{3} \bar{u},
 \end{aligned} \tag{3.1}$$

sendo v_0 a tensão na carga (RL), v_d a tensão constante de alimentação e \bar{u} o valor em estado estacionário da entrada do modelo $u(k)$ (Barroso, 2001) e,

$$D' = \frac{T_{desligado}}{T}, \tag{3.2}$$

em que $T = T_{ligado} + T_{desligado}$. Aplicando-se os valores de \bar{u} , entre 0 e 4V (pontos de operação do sistema), à Equação (3.1), obtêm-se a curva estática do conversor *buck* (Figura 3.6).

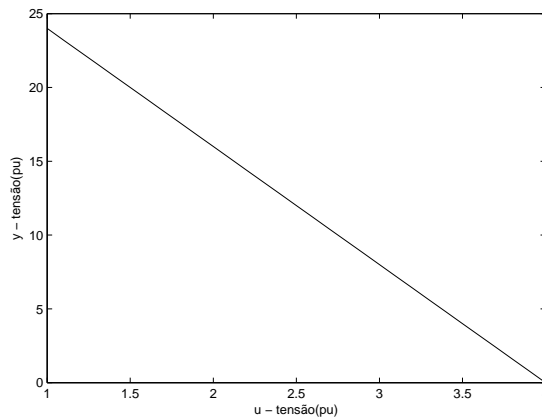


Figura 3.6: Característica estática do conversor *buck*.

Testes Dinâmicos

Para excitar a dinâmica do sistema, foi utilizado um sinal PRBS, gerado em um microcomputador e aplicado ao conversor utilizando um conversor D/A. Devido a testes preliminares, o tempo de amostragem utilizado foi $T_s = 10\mu s$. As Figuras 3.7 e 3.8 mostram os sinais obtidos.

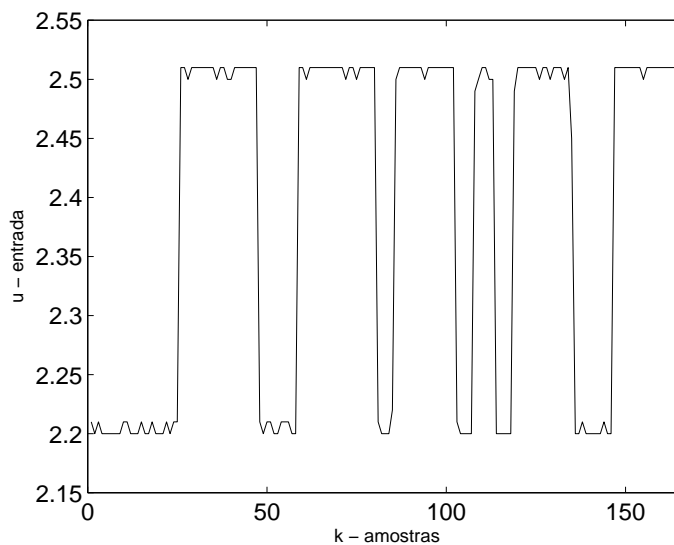


Figura 3.7: Sinal de entrada do sistema.

Os dados dinâmicos obtidos para identificação deste sistema estão contidos em uma pequena faixa de operação da característica estática. Destes dados, metade será utilizada na identificação do modelo e a outra metade, na validação do mesmo.

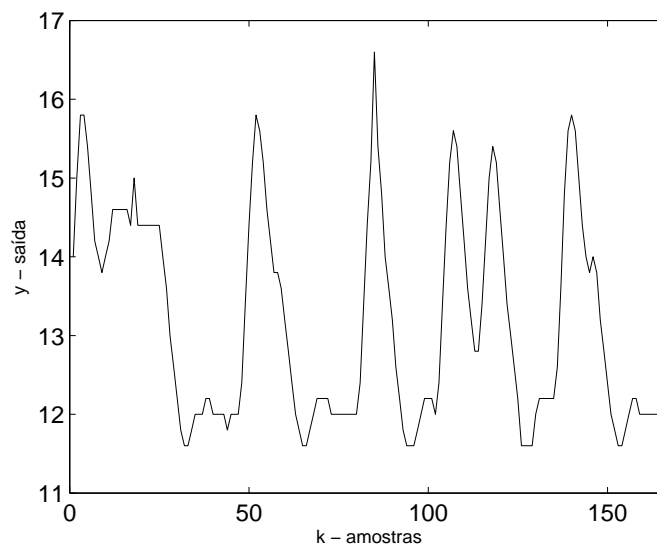


Figura 3.8: Sinal de saída do sistema.

Resultados da literatura

O conversor *buck* CC-CC foi também estudado por Coelho (2001), e na obtenção do modelos por blocos interconectados obteve índices para validação do modelo de Hammerstein para os dados dinâmicos, $RMSE$ 0,4756, e correlação para a característica estática de 1. Para o modelo de Wiener obteve o índice $RMSE$ de 0,4693 e correlação, para a característica estática, 1.

Em Barroso (2001), a obtenção de modelos para este mesmo sistema baseou-se na otimização mono-objetivo, utilizando informação auxiliar, com os métodos dos mínimos quadrados (MQ), mínimos quadrados recursivos (MQR) e algoritmo elipsoidal para estimação dos parâmetros. Obtidos os índices de validação, $RMSE$, para característica estática e dinâmica, respectivamente, 2,4792 e 0,5612, para o MQ; 0,0028 e 0,6750, para o MQR; e 0,1187 e 0,8288, para o elipsoidal.

3.3 Planta de Neutralização de pH

O processo de neutralização de pH, observado neste trabalho, se resume na adição de três soluções, ácido, base e tampão, a um recipiente chamado reator. A vazão de base é definida como variável de entrada e o nível de acidez do pH é a variável de saída. A Figura 3.9 apresenta o esquema da planta de neutralização de pH, adaptado de Henrique et al. (2000), esta planta encontra-se instalada nas dependências do Centro Universitário do Leste de Minas Gerais - UNILESTE - onde foram realizados os testes e coleta de dados. TQ1 é o tanque de ácido (HNO_3), TQ2 de solução tampão ($NaHCO_3$), TQ3 de base ($NaOH$) e TQ4 é o coletor da mistura. B1, B2 e B3 são as bombas dosadoras de cada solução. A mistura das soluções ocorre no reator. As vazões de tampão e ácido são mantidas constantes, assim como o volume no reator.

As Equações (3.3), (3.4) e (3.5) descrevem a dinâmica dentro do reator.

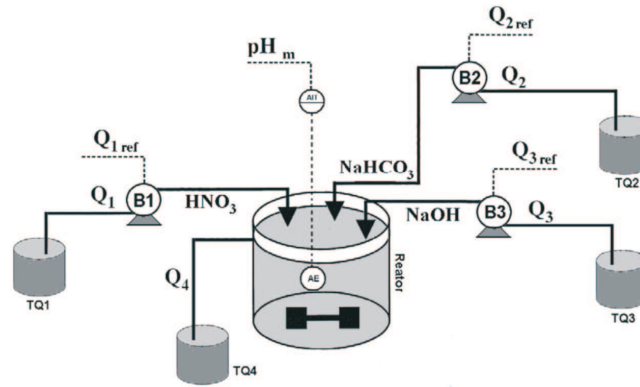


Figura 3.9: Esquema do processo de neutralização de pH.

$$\frac{dh_r}{dt} = \frac{Q_1 + Q_2 + Q_3 - c\sqrt{h_r - h_{r0}}}{A_r}, \quad (3.3)$$

$$\frac{dW_a}{dt} = \frac{Q_1(W_{a1} - W_a) + Q_2(W_{a2} - W_a) + Q_3(W_{a3} - W_a)}{A_r \cdot h_r}, \quad (3.4)$$

$$\frac{dW_b}{dt} = \frac{Q_1(W_{b1} - W_b) + Q_2(W_{b2} - W_b) + Q_3(W_{b3} - W_b)}{A_r \cdot h_r}, \quad (3.5)$$

sendo Q_1 , Q_2 e Q_3 as vazões dos tanques na planta; c a constante da válvula de saída utilizada na planta; $(h_r - h_{r0})$ a diferença de nível no tanque; A_r é a área transversal do reator. W_a e W_b são dados em mol/cm^3 , chamadas de espécies invariantes da concentração, em que se realiza o balanço das equações químicas da reação.

Característica Estática

A característica estática (Figura 3.10) é obtida por meio do modelo fenomenológico que descreve a curva de titulação de pH, proposta em Campos (2007), que apresenta resultados semelhantes à planta, em escala de laboratório. Para determinar o estado estacionário do sistema (Campos, 2007), a Equação (3.4) passa a ser representada por:

$$\frac{dW_a}{dt} = \frac{Q_1(W_{a1} - W_a) + Q_2(W_{a2} - W_a) + Q_3(W_{a3} - W_a)}{A_r \cdot h_r} = 0, \quad (3.6)$$

$$W_a = \frac{Q_1W_{a1} + Q_2W_{a2} + Q_3W_{a3}}{Q_1 + Q_2 + Q_3}. \quad (3.7)$$

De forma análoga:

$$W_b = \frac{Q_1W_{b1} + Q_2W_{b2} + Q_3W_{b3}}{Q_1 + Q_2 + Q_3}. \quad (3.8)$$

De acordo com as concentrações de W_a e W_b , calcula-se a concentração hidrogênica pela Equação (3.9)

$$X^4 + (K_{a1} - W_a)X^3 + (K_{a1}K_{a2} - K_{a1}W_a - K_w - K_{a1}W_b)X^2 - (K_{a1}K_w - K_{a1}K_{a2}W_a + 2K_{a1}K_{a2}W_b)X - K_{a1}K_{a2}K_w = 0, \quad (3.9)$$

sendo, as variáveis K_{ai} , $i=1, 2$, relações de equilíbrio químico e $[H^+] = X$. O pH será determinado pela Equação (3.10).

$$pH = -\log_{10}[X]. \quad (3.10)$$

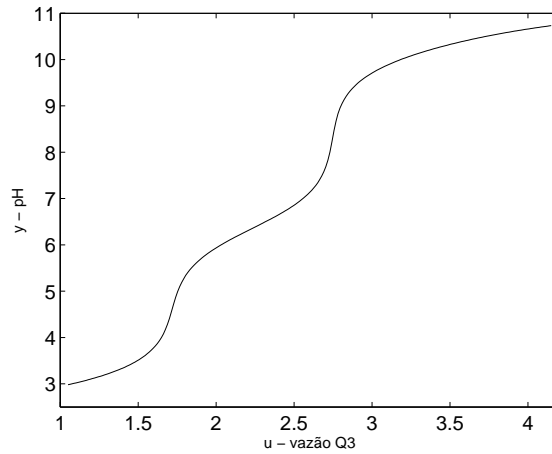


Figura 3.10: Característica estática.

Testes Dinâmicos

Os dados dinâmicos (Figuras 3.11 e 3.12) foram coletados na planta representada na Figura 3.9, que encontra-se instalada nas dependências do Centro Universitário do Leste de Minas Gerais - UNILESTE - onde foram realizados os testes e coleta de dados.

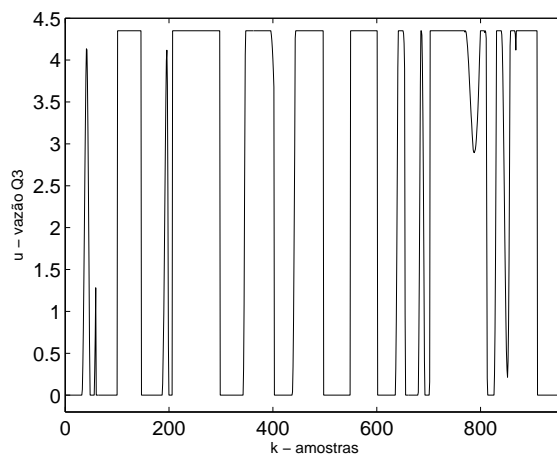


Figura 3.11: Entrada do sistema de neutralização de pH.

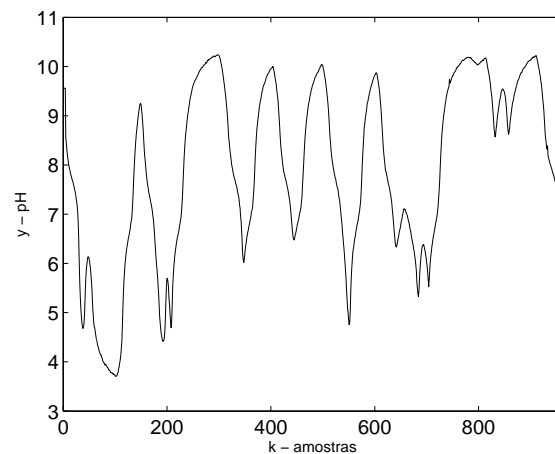


Figura 3.12: Saída dinâmica do sistema de neutralização de pH.

Resultados da literatura

Resultados recentes obtidos e comparados no trabalho Margoti et al. (2010), para a planta de neutralização de pH foram em Campos (2007), o índice RMSE, para validação do modelo dinâmico, 0,6058 e para correlação da característica estática, 0,7180, com modelo NARX racional. Já em Lima et al. (2008), os mesmos índices, RMSE (dinâmica) e correlação (estática), foram respectivamente 0,3457 e 0,9874, por meio da estimação bi-objetivo e considerando uma faixa de operação específica dos dados.

Capítulo 4

Estudo de Casos

Identificação Caixa Preta

Para a identificação caixa preta dos sistemas propostos no Capítulo 3 será realizado o treinamento das redes utilizando a rotina *nnarmax2.m* (Norgaard, 2000), que implementa o algoritmo de *Levenberg-Marquardt*. Os critérios de parada adotados para cada treinamento, tais como número de épocas de treinamento, valor máximo do erro médio quadrático, foram escolhidos aleatoriamente e serão descritos na seção de cada um dos sistemas.

Os modelos obtidos neste capítulo, por meio do treinamento da redes, foram aqueles capazes de reproduzir a característica dinâmica de cada um dos sistemas estudados. Foi testada, também, a capacidade do modelo dinâmico obtido reproduzir a característica estática de cada um dos sistemas. A validação do treinamento dinâmico foi realizada com dados da simulação da rede infinitos passos à frente. Esta simulação foi realizada por meio da função *nnsimul.m*. Na reprodução da característica estática do modelo foi utilizada a mesma função, de maneira que:

- cada valor de entrada da característica estática (u) é aplicado à rede dinâmica treinada, até que a saída desta rede fique constante;
- a verificação da estabilidade foi realizada por meio da diferença de 4 saídas passadas. Se estas diferenças estiverem menores que a tolerância considerada (erro médio quadrático), a saída é considerada constante;
- os valores são guardados em forma de pares ordenados (u, \hat{y}) , que, graficamente, representarão a característica estática reproduzida pela rede obtida.

Foram testadas as quatro estruturas propostas por Narendra and Parthasarathy (1990), expostas em 2.2.3, a fim de verificar qual arquitetura melhor representa os problemas tratados neste trabalho. Foi escolhida sempre a que apresentava melhor desempenho, levando-se em consideração a utilização de menor número de parâmetros.

Como exposto anteriormente, na estrutura referente à Figura 2.7(a), os atrasos dos termos em u (entrada) são conectados nos neurônios não lineares da rede, enquanto que, os atrasos referentes aos termos de saída, y , são conectados somente no neurônio linear. Para a segunda estrutura proposta (Figura 2.7(b)), os regressores referentes à saída são conectados aos neurônios não lineares e os regressores de saída, ao neurônio linear. A Figura 2.7(c) mostra a terceira estrutura proposta e apresenta os atrasos de entrada e saída conectando-se em neurônios com função de ativação não linear distintos. E, por fim, na quarta estrutura apresentada (Figura 2.7(d)), tem-se uma rede completamente conectada, regressores de entrada e saída conectados em neurônios não lineares.

4.1 Aquecedor com Dissipação Variável

Para a identificação caixa preta utilizando o sistema Aquecedor, foi realizada a normalização dos dados, dividindo-os por 5. Dos dados dinâmicos, metade foi utilizada no processo de identificação e a outra metade, para validação.

Os critérios de parada utilizados para os treinamentos das redes neurais para o aquecedor, na identificação caixa preta, foram:

1. número máximo de épocas de treinamento igual à 200;
2. erro médio quadrático 10^{-7} .
3. para outros parâmetros foram tomados os valores padrão (*default*) do pacote de redes neurais.

A descrição das redes utilizadas para cada uma das estruturas propostas por Narendra and Parthasarathy (1990) encontra-se na Tabela 4.1. As colunas referentes ao número de neurônios lineares e número de neurônios não lineares representam os neurônios da camada oculta. Para a camada de saída, em todas as estruturas, foi utilizado um neurônio linear.

Tabela 4.1: Estruturas utilizadas de redes neurais - Aquecedor

Rede	Estrutura Figura 2.7	Num. de neurônios não lineares	Num. de neurônios lineares	Máximo atraso da entrada	Máximo atraso da saída	Máximo atraso termos de resíduo
1	(a)	2	1	2	2	6
2	(b)	2	1	2	2	6
3	(c)	3	-	2	2	6
4	(d)	3	-	2	2	6

As Equações (4.1), (4.2), (4.3) e (4.4) apresentam os modelos obtidos para cada um dos treinamentos apresentados na Tabela 4.1, respectivamente para as redes 1, 2, 3 e 4.

$$\begin{aligned} \hat{y}(k) = & 1,188 y(k-1) - 0,2939 y(k-2) + 1,336 - \\ & -1,473 \operatorname{tansig}(0,1711 u(k-1) + 0,03108 u(k-2) + 0,4643) + \\ & +1,390 \operatorname{tansig}(0,1997 u(k-1) + 0,03543 u(k-2) - 0,5509), \end{aligned} \quad (4.1)$$

$$\begin{aligned} \hat{y}(k) = & 11,71 \operatorname{tansig}(0,2328 y(k-1) - 0,1383 y(k-2) + 0,2017) - \\ & -3,619 \operatorname{tansig}(4,382 y(k-1) - 3,907 y(k-2) + 1,814) + \\ & +0,03978 u(k-1) + 0,004711 u(k-2) + 1,095, \end{aligned} \quad (4.2)$$

$$\begin{aligned} \hat{y}(k) = & 2,920 \operatorname{tansig}(0,4304 y(k-1) - 0,1141 y(k-2) + 0,1156) + \\ & +2,041 \operatorname{tansig}(2,506 u(k-1) + 0,1045 u(k-2) - 3,476) - \\ & -0,9372 \operatorname{tansig}(2,916 u(k-1) + 0,1008 u(k-2) - 3,463) + 0,7652, \end{aligned} \quad (4.3)$$

$$\begin{aligned} \hat{y}(k) = & 1,675 \operatorname{tansig}(0,5992y(k-1) - 0,06268y(k-2) + 0,1794u(k-1) + \\ & + 0,01652u(k-2) - 1,185) - 1,274 \operatorname{tansig}(0,3477y(k-1) - 0,2675y(k-2) + \\ & + 0,2627u(k-1) + 0,04071u(k-2) + 0,4958) + 0,9181 \operatorname{tansig}(1,522y(k-1) - \\ & - 0,7095y(k-2) + 0,1957u(k-1) + 0,03660u(k-2) + 0,2443) + 1,754. \end{aligned} \quad (4.4)$$

As figuras a seguir mostram a validação qualitativa dos modelos obtidos. Em 4.1(a) verifica-se a validação do resultado do modelo obtido para os dados dinâmicos, de acordo com treinamento da rede 1, apresentada na Tabela 4.1 e a Figura 4.1(b), a reprodução da característica estática obtida por este modelo. Analogamente, as Figuras 4.2, 4.3, 4.4 apresentam as validações qualitativas dos modelos referentes às redes 2, 3 e 4, da Tabela 4.1.

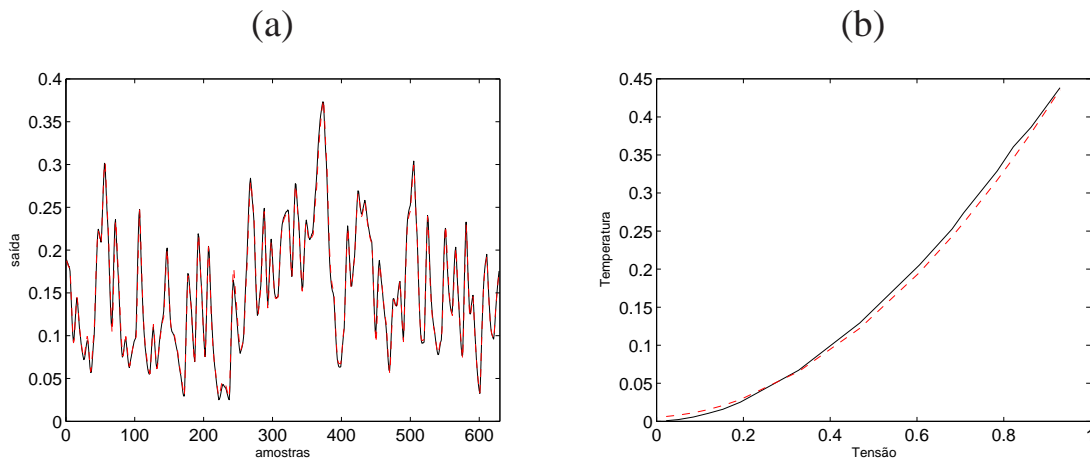


Figura 4.1: (a) Dados dinâmicos: resposta da rede 1 em linha tracejada (- -) e saída do sistema em linha contínua (-). (b) Característica estática: reprodução da característica estática da rede 1 em linha tracejada (- -) e característica estática do sistema em linha contínua (-).

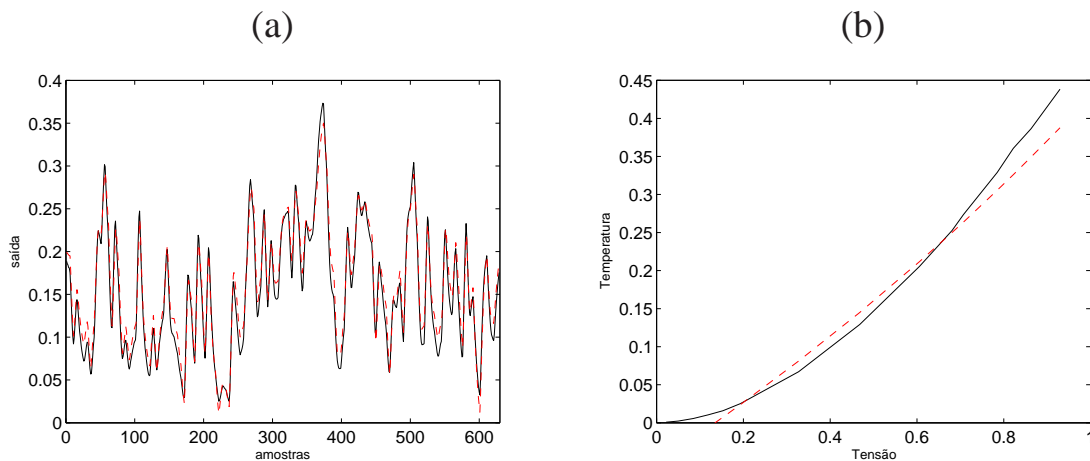


Figura 4.2: (a) Dados dinâmicos: resposta da rede 2 em linha tracejada (- -) e saída do sistema em linha contínua (-). (b) Característica estática: reprodução da característica estática da rede 2 em linha tracejada (- -) e característica estática do sistema em linha contínua (-).

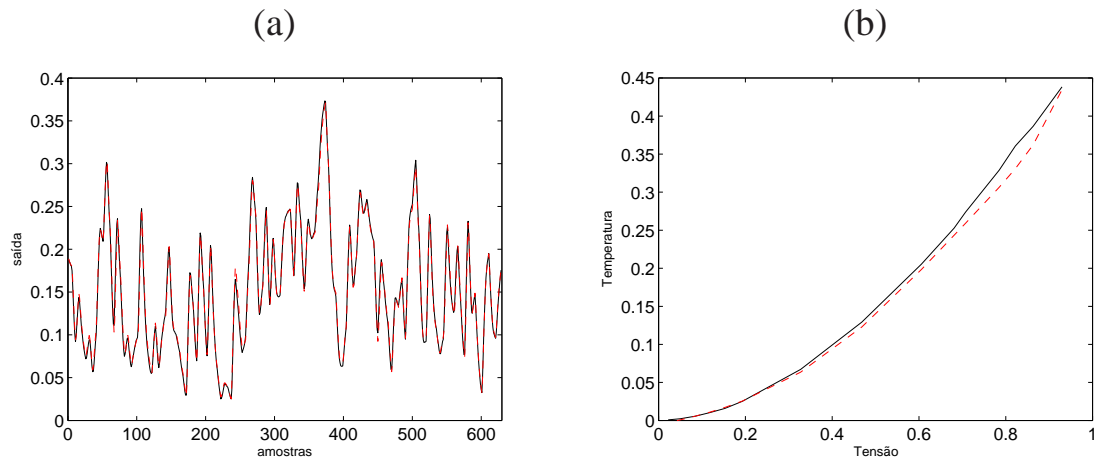


Figura 4.3: (a) Dados dinâmicos: resposta da rede 3 em linha tracejada (- -) e saída do sistema em linha contínua (-). (b) Característica estática: reprodução da característica estática da rede 3 em linha tracejada (- -) e característica estática do sistema em linha contínua (-).

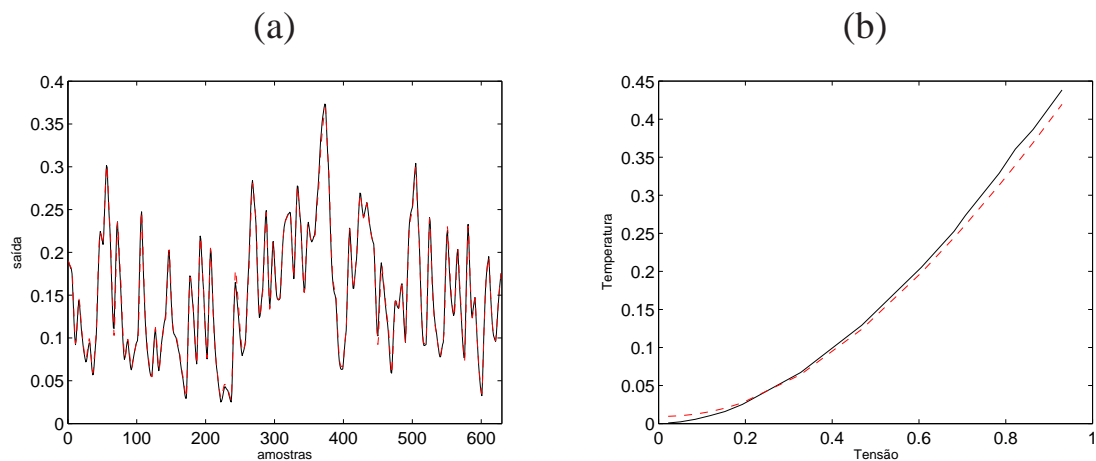


Figura 4.4: (a) Dados dinâmicos: resposta da rede 4 em linha tracejada (- -) e saída do sistema em linha contínua (-). (b) Característica estática: reprodução da característica estática da rede 4 em linha tracejada (- -) e característica estática do sistema em linha contínua (-).

Os índices de validação para os dados dinâmicos e estáticos são apresentados na Tabela 4.2.

Tabela 4.2: Índices de validação para identificação caixa preta do sistema Aquecedor, para cada uma das estruturas utilizadas

Rede	Estrutura Figura 2.7	$RMSE_{\text{dinâmico}}$	$\eta_{\text{dinâmico}}$	$\eta_{\text{estático}}$
1	(a)	0,1975	0,1406	0,9996
2	(b)	0,2584	0,2742	0,9916
3	(c)	0,1984	0,1689	0,9991
4	(d)	0,1994	0,1456	0,9999

Discussão sobre os resultados

De acordo com a Tabela 4.2, que mostra os índices de validação para os treinamentos caixa preta do sistema aquecedor, verifica-se que as redes apresentaram resultados dentro dos níveis estabelecidos pelos índices. A rede 2, no que se refere à reprodução da característica estática, percebe-se uma baixa correlação, que pode ser verificado qualitativamente pela Figura 4.3. Para as demais redes, os índices de correlação estático, para reprodução desta característica, apresentaram melhor desempenho, e, concomitante com a análise dos índices dinâmicos e das respectivas figuras, nota-se que as redes 1 e 4 possuem melhores índices de validação para a característica dinâmica.

Conforme previsto em Cassini (1999); Barroso (2001), a não linearidade dos dados deste sistema está em u , o que explica os melhores índices obtidos em redes que a não linearidade também está para os dados de entrada u . A rede 4 em particular, por ser uma rede completamente conectada, tem a garantia de ser um aproximador universal conforme exposto no Capítulo 2, Seção 2.2.3, porém a rede 1, que apresenta não linearidade para os dados de entrada, é uma rede mais simplificada que tem relação com a característica do sistema.

4.2 Conversor *Buck* - CC-CC

À partir dos dados disponíveis do sistema conversor *buck* - CC-CC foi realizada a identificação caixa preta, observando as quatro estruturas diferentes propostas no trabalho de Narendra and Parthasarathy (1990), com a finalidade de modelar o comportamento do deste sistema. Inicialmente, os dados dinâmicos foram decimados com um fator $\Delta = 12$, como proposto no processo de tratamento dos dados em Aguirre (2007). Destes dados, metade foi utilizada no processo de identificação do modelo e a outra metade, para validação.

Os critérios de parada utilizados para estes treinamentos foram:

1. número máximo de épocas de treinamento igual à 200;
2. erro médio quadrático 10^{-7} .

A descrição das redes utilizadas, para cada uma das estruturas propostas por Narendra and Parthasarathy (1990), encontra-se na Tabela 4.3. As colunas referentes ao número de neurônios lineares e número de neurônios não lineares representam os neurônios da camada oculta. Para a camada de saída, em todas as estruturas, foi utilizado 1 neurônio linear.

Tabela 4.3: Estrutura para treinamento das redes neurais - conversor *buck* CC-CC

-	Estrutura Figura 2.7	Núm. de neurônios não lineares	Núm. de neurônios lineares	Máximo atraso da entrada	Máximo atraso da saída	Máximo atraso termos de resíduo
1	(a)	4	1	4	6	3
2	(b)	7	1	7	9	5
3	(c)	6	-	5	9	7
4	(d)	5	-	4	7	3

Os modelos obtidos para cada um dos treinamentos apresentados na Tabela 4.1 estão representados nas Equações (4.5), (4.6), (4.7) e (4.8), respectivamente para as redes 1, 2, 3 e 4.

$$\begin{aligned}
\hat{y}(k) = & 0,8458 y(k-1) + 0,6060 y(k-2) - 0,4571 y(k-3) - 0,2987 y(k-4) + \\
& + 0,04512 y(k-5) + 0,1635 y(k-6) + 1,701 + 1,726 \operatorname{tansig}(-0,7055 u(k-1)) + \\
& + 0,04500 u(k-2) + 0,3697 u(k-3) + 0,3392 u(k-4) + 0,6731) \\
& + 0,5948 \operatorname{tansig}(-0,4884 u(k-1) - 0,05873 u(k-2) + 0,3526 u(k-3)) - \\
& + 0,06118 u(k-4) + 0,3733) + 1,821 \operatorname{tansig}(-1,861 u(k-1) - 0,2735 u(k-2)) + \\
& + 1,242 u(k-3) + 0,3579 u(k-4) + 0,3864) + 0,4454 \operatorname{tansig}(0,03951 u(k-1) - \\
& - 0,1791 u(k-2) + 0,2341 u(k-3) - 0,3352 u(k-4) + 0,1959). \tag{4.5}
\end{aligned}$$

$$\begin{aligned}
\hat{y}(k) = & -0,7115 \operatorname{tansig}(0,1830 y(k-1) + 0,2123 y(k-2) - 0,1912 y(k-3)) - \\
& - 0,03534 y(k-4) - 0,1713 y(k-5) - 0,007390 y(k-6) + 0,01582 y(k-7) - \\
& - 0,1841 y(k-8) - 0,2424 y(k-9) - 0,2333) - \\
& - 1,319 \operatorname{tansig}(0,1998 y(k-1) - 0,09208 y(k-2) - 0,04484 y(k-3)) - \\
& - 0,1144 y(k-4) - 0,4021 y(k-5) + 0,1329 y(k-6) + 0,008645 y(k-7) + \\
& + 0,1250 y(k-8) + 0,1213 y(k-9) + 0,09064) - \\
& - 1,043 \operatorname{tansig}(0,1540 y(k-1) + 0,1256 y(k-2) + 0,02799 y(k-3)) - \\
& - 0,2238 y(k-4) - 0,1060 y(k-5) - 0,1035 y(k-6) - 0,02774 y(k-7) - \\
& - 0,2040 y(k-8) - 0,1338 y(k-9) - 0,2399) - \\
& - 0,8146 \operatorname{tansig}(0,09848 y(k-1) - 0,2104 y(k-2) - 0,05250 y(k-3)) - \\
& - 0,008324 y(k-4) - 0,05258 y(k-5) - 0,09732 y(k-6) - 0,1814 y(k-7) - \\
& - 0,05708 y(k-8) - 0,1106 y(k-9) - 0,02138) + \\
& + 0,8602 \operatorname{tansig}(0,09380 y(k-1) - 0,2090 y(k-2) + 0,1906 y(k-3)) + \\
& + 0,03134 y(k-4) + 0,04704 y(k-5) + 0,2515 y(k-6) + 0,2513 y(k-7) + \\
& + 0,2458 y(k-8) - 0,1576 y(k-9) - 0,2059) - \\
& - 0,8597 \operatorname{tansig}(0,1730 y(k-1) - 0,3398 y(k-2) - 0,2734 y(k-3)) - \\
& - 0,1177 y(k-4) - 0,1171 y(k-5) - 0,02815 y(k-6) + 0,1771 y(k-7) + \\
& + 0,1903 y(k-8) + 0,3052 y(k-9) + 0,02258) + \\
& + 0,8741 \operatorname{tansig}(0,06906 y(k-1) + 0,2118 y(k-2) + 0,1839 y(k-3)) + \\
& + 0,01520 y(k-4) + 0,01141 y(k-5) + 0,04351 y(k-6) - 0,1375 y(k-7) + \\
& + 0,1536 y(k-8) + 0,1897 y(k-9) + 0,1047) - 1,008 u(k-1) - \\
& - 1,403 u(k-2) - 0,2634 u(k-3) + 0,8397 u(k-4) + 1,807 u(k-5) + \\
& + 1,283 u(k-6) + 0,5807 u(k-7) + 3,326, \tag{4.6}
\end{aligned}$$

$$\begin{aligned}
\hat{y}(k) = & 4,419 \operatorname{tansig}(0,3908 y(k-1) + 0,2207 y(k-2) + \\
& + 0,3123 y(k-3) + 0,1967 y(k-4) + 0,1879 y(k-5) + 0,5398 y(k-6) + \\
& + 0,5322 y(k-7) + 0,3455 y(k-8) + 0,3018 y(k-9) + 0,1524) + \\
& + 0,6831 \operatorname{tansig}(0,1151 y(k-1) + 0,1990 y(k-2) - \\
& - 0,01560 y(k-3) + 0,2632 y(k-4) + 0,3045 y(k-5) - 0,1042 y(k-6) - \\
& - 0,08922 y(k-7) + 0,3009 y(k-8) + 0,01546 y(k-9) + 0,1946) + \\
& + 4,436 \operatorname{tansig}(0,1895 y(k-1) + 0,1967 y(k-2) + \\
& + 0,1410 y(k-3) - 0,01936 y(k-4) + 0,2473 y(k-5) + 0,1435 y(k-6) + \\
& + 0,06649 y(k-7) + 0,04503 y(k-8) + 0,2051 y(k-9) - 0,003006) + \\
& + 4,197 \operatorname{tansig}(0,3835 u(k-1) + 0,7434 u(k-2) + \\
& + 0,9253 u(k-3) + 0,8948 u(k-4) + 0,7149 u(k-5) + 0,3663) - \\
& - 14,73 \operatorname{tansig}(0,1643 u(k-1) + 0,1491 u(k-2) + \\
& + 0,1216 u(k-3) + 0,08924 u(k-4) + 0,03452 u(k-5) - 0,9964) - \\
& - 0,5306 \operatorname{tansig}(0,6301 u(k-1) + 0,6611 u(k-2) + \\
& + 0,7952 u(k-3) + 0,3832 u(k-4) + 0,6227 u(k-5) + 0,09093) + 4,638, \quad (4.7)
\end{aligned}$$

$$\begin{aligned}
\hat{y}(k) = & 3,370 \operatorname{tansig}(0,5752 y(k-1) + 0,1539 y(k-2) + 0,4957 y(k-3) + \\
& + 0,6141 y(k-4) + 0,3236 y(k-5) + 0,5729 y(k-6) + 0,2521 y(k-7) + \\
& + 0,001169 u(k-1) - 0,06928 u(k-2) + 0,1088 u(k-3) + 0,1402 u(k-4) + \\
& + 0,1206) + 5,220 \operatorname{tansig}(0,3720 y(k-1) - 0,2825 y(k-2) + \\
& + 0,1507 y(k-3) - 0,08601 y(k-4) - 0,005716 y(k-5) + 0,04423 y(k-6) - \\
& - 0,01594 y(k-7) - 0,4909 u(k-1) + 0,4359 u(k-2) - 0,2726 u(k-3) + \\
& + 0,1668 u(k-4) - 1,904) + 3,437 \operatorname{tansig}(0,2414 y(k-1) + \\
& + 0,1708 y(k-2) + 0,04981 y(k-3) - 0,003943 y(k-4) + 0,1212 y(k-5) - \\
& - 0,08514 y(k-6) + 0,2414 y(k-1) + 0,1708 y(k-2) + 0,04981 y(k-3) - \\
& - 0,003943 y(k-4) + 0,1212 y(k-5) - 0,08514 y(k-6) - 0,06284) - \\
& - 0,9288 \operatorname{tansig}(0,2483 y(k-1) - 0,07087 y(k-2) - 0,05023 y(k-3) + \\
& + 0,1763 y(k-4) + 0,2694 y(k-5) + 0,2126 y(k-6) - 0,08085 y(k-7) - \\
& - 0,2046 u(k-1) + 0,1683 u(k-2) - 0,2009 u(k-3) - 0,1382 u(k-4) + \\
& + 0,09382) + 3,616 \operatorname{tansig}(0,1931 y(k-1) + 0,1565 y(k-2) - \\
& - 0,005640 y(k-3) + 0,1630 y(k-4) + 0,07150 y(k-5) + 0,2469 y(k-6) - \\
& - 0,1115 y(k-7) - 0,04804 u(k-1) + 0,2327 u(k-2) + 0,1255 u(k-3) + \\
& + 0,02348 u(k-4) + 0,07894) + 3,489. \quad (4.8)
\end{aligned}$$

Nas Figuras 4.5, 4.6, 4.7, 4.8 são apresentadas as simulações dos modelos obtidos nas Equações (4.5), (4.6), (4.7) e (4.8), respectivamente, e a comparação com as curvas características do sistema. Em (a) a comparação e validação para os dados dinâmicos e em (b), a reprodução da característica estática.

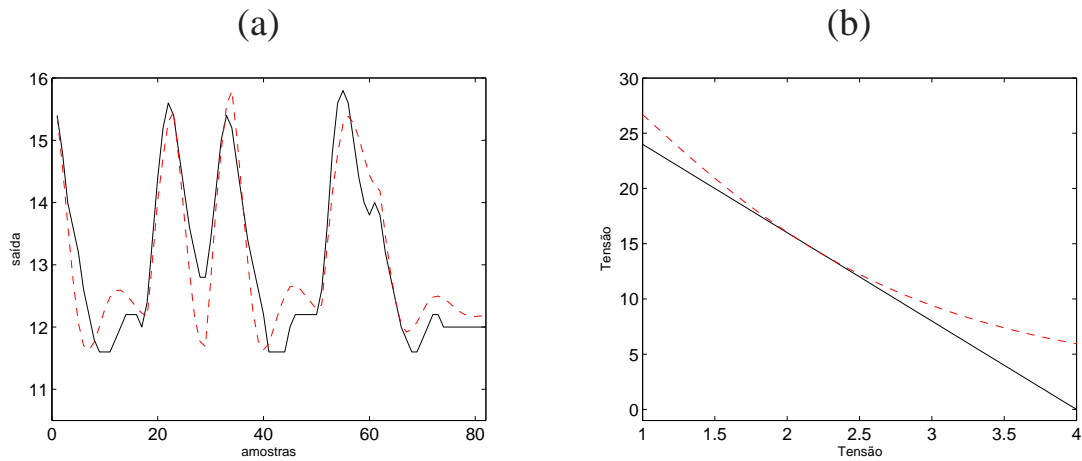


Figura 4.5: (a) Dados dinâmicos: resposta da rede 1 em linha tracejada (- -) e saída do sistema em linha contínua (-). (b) Característica estática: reprodução da característica estática da rede 1 em linha tracejada (- -) e característica estática do sistema em linha contínua (-).

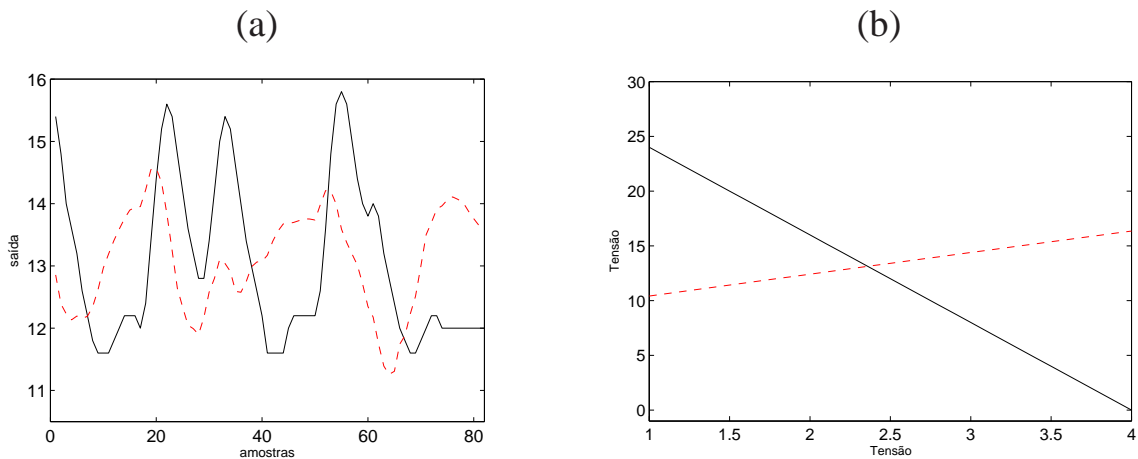


Figura 4.6: (a) Dados dinâmicos: resposta da rede 2 em linha tracejada (- -) e saída do sistema em linha contínua (-). (b) Característica estática: reprodução da característica estática da rede 2 em linha tracejada (- -) e característica estática do sistema em linha contínua (-).

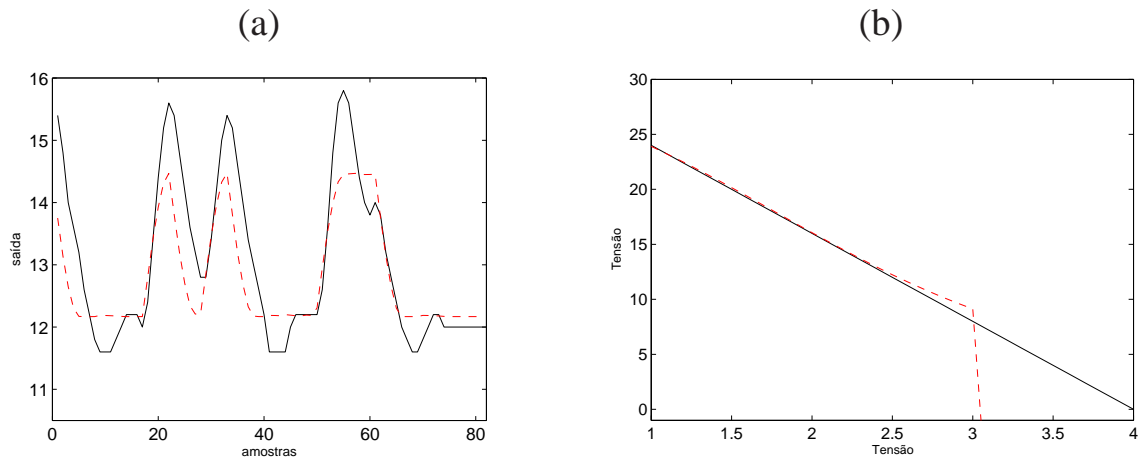


Figura 4.7: (a) Dados dinâmicos: resposta da rede 3 em linha tracejada (- -) e saída do sistema em linha contínua (-). (b) Característica estática: reprodução da característica estática da rede 3 em linha tracejada (- -) e característica estática do sistema em linha contínua (-).

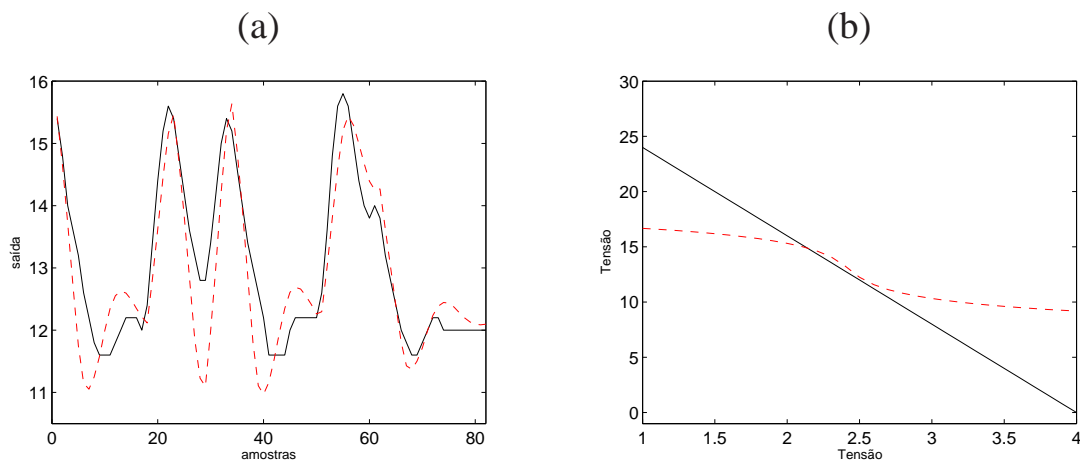


Figura 4.8: (a) Dados dinâmicos: resposta da rede 4 em linha tracejada (- -) e saída do sistema em linha contínua (-). (b) Característica estática: reprodução da característica estática da rede 4 em linha tracejada (- -) e característica estática do sistema em linha contínua (-).

Os índices de validação para os dados dinâmicos e estáticos são apresentados na Tabela 4.4.

Tabela 4.4: Índices de validação para identificação caixa preta do sistema conversor *buck*, para cada uma das estruturas utilizadas

	Estrutura Figura 2.7	$RMSE_{\text{dinâmico}}$	$\eta_{\text{dinâmico}}$	$\eta_{\text{estático}}$
1	(a)	0,3915	0,3546	0,9775
2	(b)	1,1235	0,8209	-1,000
3	(c)	0,5039	0,8010	0,9386
4	(d)	0,5078	0,2314	0,9698

Tabela 4.5: Estrutura para treinamento das redes neurais - pH

-	Estrutura Figura 2.7	Núm. de neurônios não lineares	Núm. de neurônios lineares	Máximo atraso da entrada	Máximo atraso da saída	Máximo atraso termos de resíduo
1	(a)	2	1	4	5	6
2	(b)	3	1	2	4	4
3	(c)	4	-	2	5	4
4	(d)	3	-	4	7	4

Discussão sobre os resultados

Na identificação caixa preta do conversor *buck*, pela Tabela 4.3, verifica-se que as estruturas necessárias para obtenção dos índices mostrados na Tabela 4.4 foram redes com grande número de neurônios e atrasos dos termos de entrada e saída, em comparação com as redes obtidas para os outros sistemas. Uma suposição para tal seria o fato de que os dados dinâmicos não contemplam toda a faixa de operação dos dados estáticos, o que torna necessária uma rede com arquitetura mais complexa.

Nas redes 1 e 4, além de índices melhores, percebe-se nas Figuras 4.5, 4.8, que dentro da faixa de operação dos dados dinâmicos, estas conseguem reproduzir a estática do sistema, o que não acontece com as redes 2 e 3, como mostrado nas Figuras 4.6, 4.7. A identificação caixa preta, à partir dos dados do conversor *buck*, pode ser um indicativo de qual a melhor estrutura para treinamento e obtenção de modelos.

4.3 Planta de Neutralização de pH

A partir dos dados disponíveis do sistema do pH, obtidos conforme a descrição apresentada na Seção 3.3, foi realizada a identificação caixa preta. De forma análoga aos outros sistema, realizando treinamentos de redes com diferentes estruturas, de acordo com o trabalho de Narendra and Parthasarathy (1990). Para a realização destes treinamentos, metade dos dados dinâmicos foram utilizados na identificação e a outra metade, na validação do modelo.

Assim como os dos outros sistemas, os critérios de parada utilizados para estes treinamentos foram:

1. número máximo de épocas de treinamento igual à 200;
2. erro médio quadrático 10^{-7} .

A descrição das estruturas de redes utilizadas para obtenção dos modelos encontram-se na Tabela 4.5. As colunas referentes ao número de neurônios lineares e número de neurônios não lineares representam os neurônios da camada oculta. Para a camada de saída, em todas as estruturas, foi utilizado 1 neurônio linear.

As Equações (4.9), (4.10), (4.11) e (4.12) apresentam os modelos obtidos para os dados dinâmicos, de acordo com treinamento da rede de estrutura apresentada na Tabela 4.5.

$$\begin{aligned}
\hat{y}(k) = & 1,318 y(k-1) - 0,4412 y(k-2) - 0,2062 y(k-3) + 0,6437 y(k-4) - \\
& - 0,3353 y(k-5) + 0,1735 + 0,5702 \text{tansig}(0,06479 u(k-1) - \\
& - 0,1214 u(k-2) + 0,6187 u(k-3) - 0,6647 u(k-4) - 0,08807) - \\
& - 0,5468 \text{tansig}(0,06562 u(k-1) - 0,1791 u(k-2) + 0,5911 u(k-3) - \\
& - 0,6686 u(k-4) + 0,1030), \tag{4.9}
\end{aligned}$$

$$\begin{aligned}
\hat{y}(k) = & 6,468 \text{tansig}(0,9316 y(k-1) - 1,272 y(k-2) + 0,1825 y(k-3) + \\
& + 0,2701 y(k-4) - 0,7705) + 0,4571 \text{tansig}(0,1769 y(k-1) + \\
& + 0,4625 y(k-2) + 0,1948 y(k-3) + 0,08851 y(k-4) + 0,06503) - \\
& - 4,021 \text{tansig}(1,182 y(k-1) - 2,019 y(k-2) + 0,3132 y(k-3) + \\
& + 0,4513 y(k-4) + 0,4934) + 0,002698 u(k-1) + \\
& + 0,02207 u(k-2) + 6,332, \tag{4.10}
\end{aligned}$$

$$\begin{aligned}
\hat{y}(k) = & 8,066 \text{tansig}(0,05473 y(k-1) + 0,1753 y(k-2) - 0,05210 y(k-3) - \\
& - 0,1018 y(k-4) + 0,05103 y(k-5) - 0,8951) + \\
& + 3,973 \text{tansig}(0,03043 u(k-1) - 0,02184 u(k-2) + 0,5978) - \\
& - 0,6657 \text{tansig}(0,3084 u(k-1) - 0,02253 u(k-2) - 0,8638) + \\
& + 1,353 \text{tansig}(0,08574 u(k-1) + 0,05868 u(k-2) - 0,03476) + 4,410, \tag{4.11}
\end{aligned}$$

$$\begin{aligned}
\hat{y}(k) = & 0,7261 \text{tansig}(1,724 y(k-1) + 1,449 y(k-2) + 1,490 y(k-3) + \\
& + 1,624 y(k-4) + 1,186 y(k-5) + 1,880 y(k-6) + 1,153 y(k-7) + \\
& + 1,454 u(k-1) + 1,557 u(k-2) + 1,725 u(k-3) + 1,464 u(k-4) + 0,7975) + \\
& + 3,247 \text{tansig}(2,343 y(k-1) + 2,182 y(k-2) + 1,971 y(k-3) + \\
& + 2,032 y(k-4) + 2,152 y(k-5) + 1,881 y(k-6) + 2,377 y(k-7) + \\
& + 0,8056 u(k-1) + 0,5937 u(k-2) + 0,9127 u(k-3) + 0,1837 u(k-4) + 0,06573) + \\
& + 11,60 \text{tansig}(0,1498 y(k-1) - 0,1035 y(k-2) + 0,04383 y(k-3) - \\
& - 0,002967 y(k-4) + 0,01747 y(k-5) - 0,02694 y(k-6) + 0,009094 y(k-7) + \\
& + 0,0008186 u(k-1) + 0,001039 u(k-2) + 0,0007118 u(k-3) + \\
& + 0,00006406 u(k-4) - 0,6114) + 3,006. \tag{4.12}
\end{aligned}$$

As Figuras 4.9, 4.10, 4.11, 4.12, (a), mostram a validação dos resultados dos modelos obtidos para os dados dinâmicos, após o treinamento das redes apresentadas na Tabela 4.5 e em (b), as reproduções da característica estática obtidas pelos modelos.

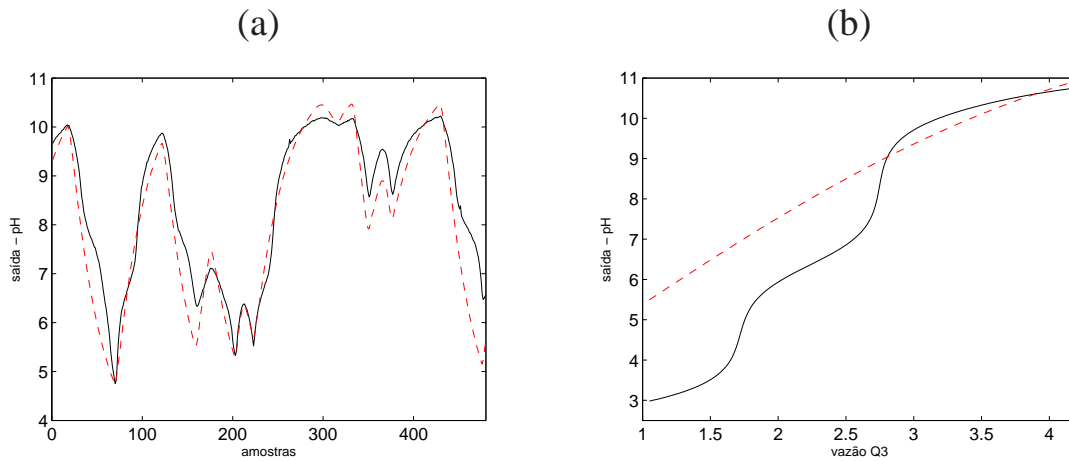


Figura 4.9: (a) Dados dinâmicos: resposta da rede 1 em linha tracejada (- -) e saída do sistema em linha contínua (-). (b) Característica estática: reprodução da característica estática da rede 1 em linha tracejada (- -) e característica estática do sistema em linha contínua (-).

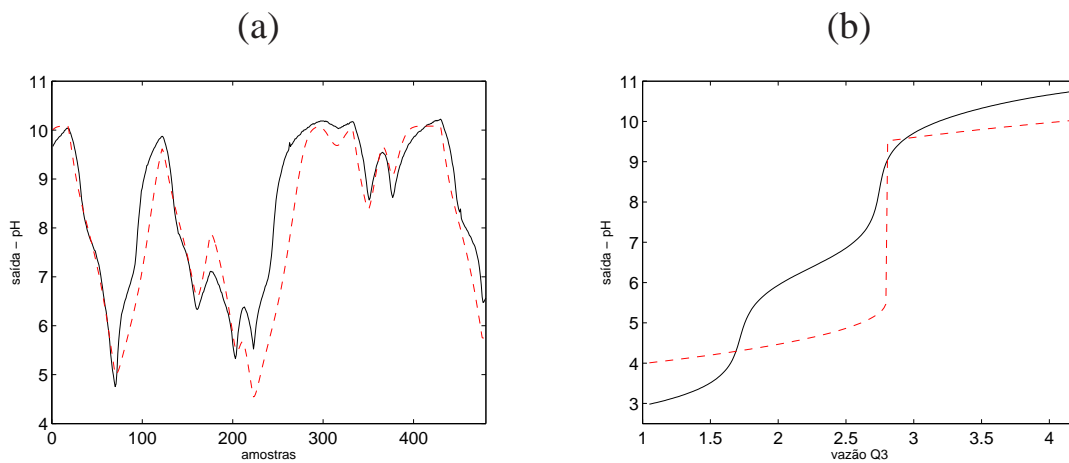


Figura 4.10: (a) Dados dinâmicos: resposta da rede 2 em linha tracejada (- -) e saída do sistema em linha contínua (-). (b) Característica estática: reprodução da característica estática da rede 2 em linha tracejada (- -) e característica estática do sistema em linha contínua (-).

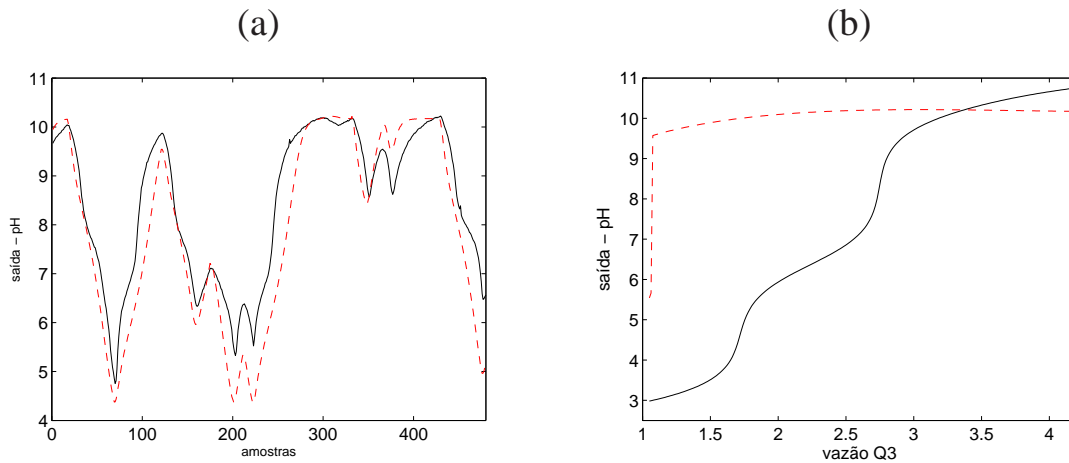


Figura 4.11: (a) Dados dinâmicos: resposta da rede 3 em linha tracejada (- -) e saída do sistema em linha contínua (-). (b) Característica estática: reprodução da característica estática da rede 3 em linha tracejada (- -) e característica estática do sistema em linha contínua (-).

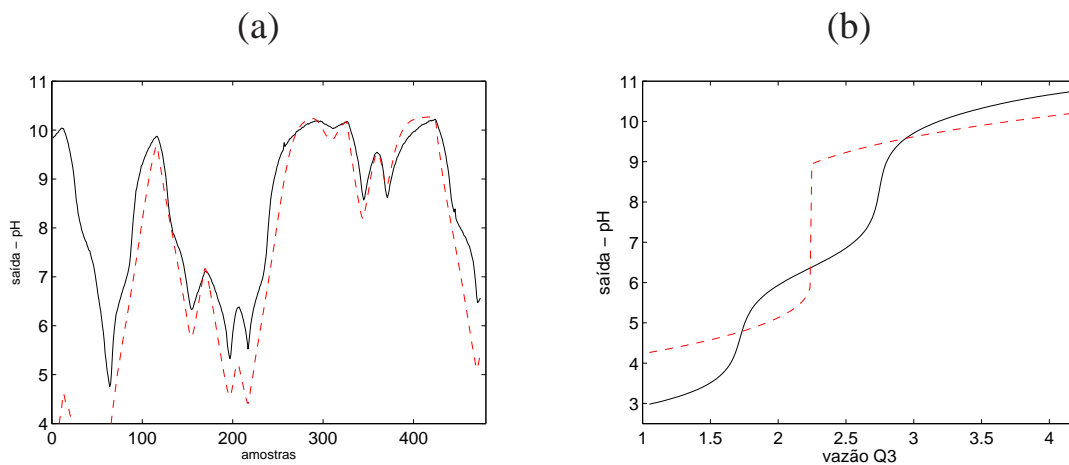


Figura 4.12: (a) Dados dinâmicos: resposta da rede 4 em linha tracejada (- -) e saída do sistema em linha contínua (-). (b) Característica estática: reprodução da característica estática da rede 4 em linha tracejada (- -) e característica estática do sistema em linha contínua (-).

Os índices de validação para os dados dinâmicos e estáticos são apresentados na Tabela 4.6.

Tabela 4.6: Índices de validação para identificação caixa preta da planta de neutralização de pH, para cada uma das estruturas utilizadas

Rede	Estrutura Figura 2.7	$RMSE_{\text{dinâmico}}$	$\eta_{\text{dinâmico}}$	$\eta_{\text{estático}}$
1	(a)	0,4108	0,2017	0,9861
2	(b)	0,4824	0,0013	0,9279
3	(c)	0,5203	0,4671	0,4464
4	(d)	0,9964	0,1193	0,9221

Discussão sobre os resultados

As redes 1 e 4 apresentam melhores índices de validação, tanto dinâmicos quanto estáticos. Entretanto, pela característica dos dados, em que os dados dinâmicos não abrangem toda a faixa de operação da característica estática, verifica-se pelas Figuras 4.9 e 4.9 que as redes não representam bem a característica estática.

Inicialmente, no Capítulo 3, Seção 3.3, verifica-se que o sinal de entrada para obtenção da característica dinâmica foi um sinal PRBS, com amplitude fixa, que excita o sistema, em apenas 1 ponto de operação. Por outro lado, a característica estática do sistema não está completamente contida na massa de dados dinâmicos, o que explica a dificuldade da reprodução desta característica após o treinamento dinâmico.

Capítulo 5

Estudo de Casos: Identificação Caixa Cinza

5.1 Modelos de Wiener e Hammerstein

São apresentados os modelos obtidos para cada um dos sistemas apresentados no Capítulo 3, por meio das representações de blocos interconectados, Wiener e Hammerstein (Seção 5.1).

Na identificação destes modelos, utilizou-se redes neurais com estruturas e critérios de treinamento escolhidos por tentativa e erro. Os treinamentos foram realizados utilizando o *toolbox* para identificação de sistemas dinâmicos (*Neural network based system identification*) (Norgaard, 2000), em que o método de Levenberg-Marquardt é utilizado para minimização do erro médio quadrático. Para os ajustes de característica estática dos sistemas foi utilizado o *toolbox* padrão de Redes Neurais, do *software Matlab*.

Para a obtenção do modelo de Wiener (Figura 5.1(b)), após o ajuste da característica estática, é necessário obter a inversa desta função. Por meio dos dados de entrada (\bar{u}) e saída (\bar{y}) da característica estática, estima-se a função inversa do modelo, simulando, neste ponto, o par entrada-saída como sendo, \bar{y} e \bar{u} , respectivamente. Com a inversa da estática, aplica-se, na entrada desta função, a saída dinâmica do sistema, a fim de simular o sinal intermediário $v(k)$, que, conforme exposto anteriormente, não possui significado físico. Assim, torna-se possível simular o modelo ARX (bloco dinâmico linear), com o par entrada-saída sendo u e $v(k)$, respectivamente.

Para obtenção do modelo de Hammerstein (Figura 5.1(a)), ajusta-se a característica estática por meio dos dados estáticos. Neste modelo não é necessário obter uma função inversa, porém é necessário aplicar a entrada \bar{u} , dos dados dinâmicos, ao modelo obtido para a característica estática e, simular a saída deste bloco, que será o sinal $v(k)$. Para a obtenção do modelo ARX, do bloco referente à dinâmica linear, utiliza-se o par entrada-saída $v(k)$ - $y(k)$, respectivamente.

5.1.1 Aquecedor com Dissipação Variável

Modelo de Wiener

O ajuste neural para característica estática utilizou uma rede com 2 neurônios na camada oculta, com função de ativação tangente hiperbólica e 1 neurônio, com função de ativação linear, na saída da rede. A Figura 5.2 mostra a validação qualitativa do ajuste realizado nesta representação e a Equação (5.1) descreve o modelo que ajusta esta característica.

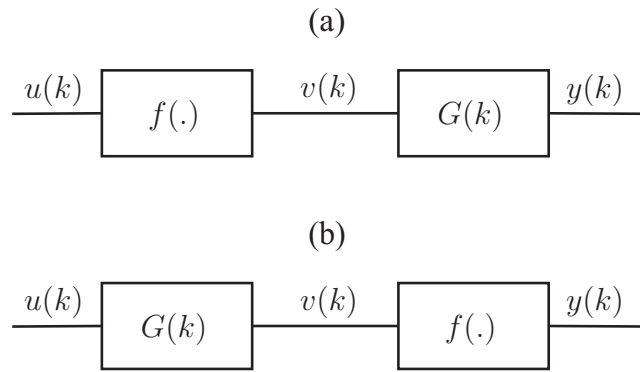


Figura 5.1: Modelos de blocos interconectados. $f(\cdot)$ representa a função estática não linear e $g(\cdot)$ representa a função dinâmica linear. Em (a) Modelo de Hammerstein e em (b) Modelo de Wiener.

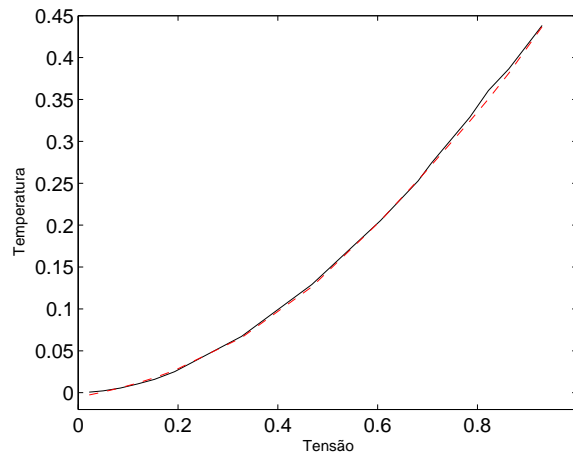


Figura 5.2: Validação da característica estática do sistema. Característica estática do sistema(-). Característica estática resultante do ajuste por redes neurais (- -).

$$\begin{aligned} \bar{y} = & 0,2563 \operatorname{tansig}(3,606 \bar{u} - 4,008) + \\ & + 1,353 \operatorname{tansig}(0,9923 \bar{u} - 0,4865) + 0,4594. \end{aligned} \quad (5.1)$$

Para a estimação da função inversa do modelo de Wiener, treinou-se uma rede, com 2 neurônios na camada oculta, com função de ativação tangente hiperbólica (*tansig*) e 1 neurônio na camada de saída, com função de ativação linear. A Figura 5.3 apresenta o ajuste da função inversa obtida, cujo modelo está descrito na Equação (5.2).

Com a função inversa foi obtido o sinal intermediário $v(k)$ e treinou-se uma rede neural dinâmica, para obtenção do modelo ARX polinomial, contendo 1 neurônio linear na camada oculta e 1 neurônio linear na camada de saída. Os atrasos máximos dos regressores de entrada, saída e termos de resíduos foram, respectivamente, 2, 5 e 3. A Equação (5.3) descreve o modelo neural de Wiener e a Figura 5.4, sua representação. A Tabela 5.1 contém os índices de validação para este modelo obtido.

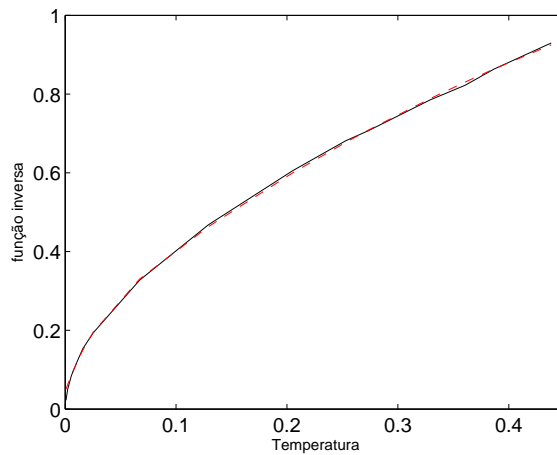


Figura 5.3: Função inversa da característica estática. Característica do sistema(-). Característica resultante do ajuste por redes neurais (- -).

$$\bar{v} = 11,08 \operatorname{tansig}(0,2175 y + 1,155) + 4,146 \operatorname{tansig}(4,399 y + 6,011) - 12,90. \quad (5.2)$$

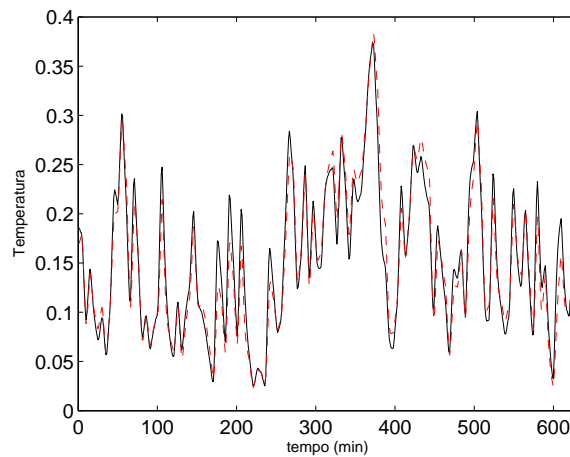


Figura 5.4: Validação do modelo Wiener Neural para o aquecedor de dissipação variável. Dados de validação (-). Saída da rede(- -).

$$\begin{aligned} \hat{v}(k) = & 1,509 v(k-1) - 0,7094 v(k-2) + 0,2042 v(k-3) - 0,1186 v(k-4) + \\ & + 0,05726 v(k-5) + 0,07499 u(k-1) - 0,01461 u(k-2). \end{aligned} \quad (5.3)$$

Modelo de Hammerstein

Para ajuste da característica estática do aquecedor de dissipação variável, foi treinada uma rede neural com 2 neurônios com função de ativação tangente hiperbólica (*tansig*), com critérios de parada sendo o número máximo de épocas igual à 100 ou o erro médio quadrático, igual à 0,0001. A Figura 5.5 mostra o resultado, validação qualitativa, deste treinamento. A Equação (5.4) descreve o modelo.

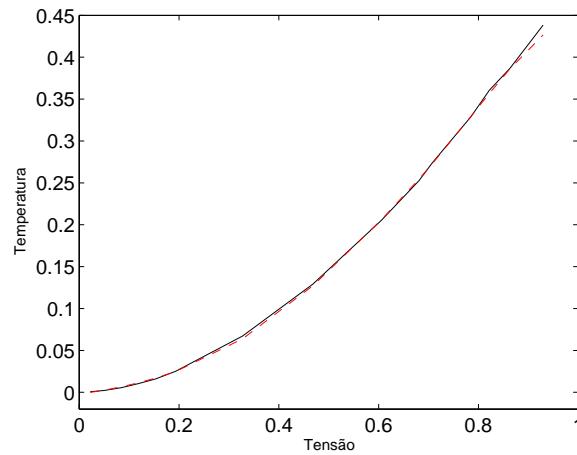


Figura 5.5: Ajuste da característica estática. Característica do sistema(-). Característica resultante do ajuste por redes neurais (- -).

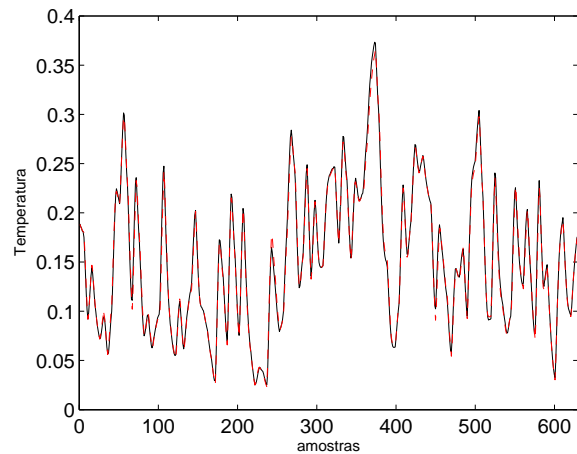


Figura 5.6: Validação do modelo Hammerstein Neural para o aquecedor de dissipação variável. Dados de validação (-). Saída da rede(- -).

$$\begin{aligned} \bar{y} = & 0,4494 \operatorname{tansig}(2,332 \bar{u} - 1,822) + \\ & + 0,7371 \operatorname{tansig}(1,366 \bar{u} - 0,1188) + 0,1103. \end{aligned} \quad (5.4)$$

Após o ajuste da característica estática, procedeu-se com a identificação do modelo ARX, para o bloco dinâmico linear. Utilizando somente neurônios lineares na rede, para obter um modelo ARX e regressores para entrada, saída e termos de resíduos, respectivamente, 2, 3 e 3. A Equação (5.5) mostra o modelo Hammerstein neural obtido para o aquecedor e a Figura 5.6 o representa. E na Tabela 5.1 são apresentados os índices para validação quantitativa do modelo estimado.

$$\begin{aligned} \hat{y}(k) = & 0,5936 y(k-1) + 0,4493 y(k-2) - 0,2040 y(k-3) + \\ & + 0,09003 v(k-1) + 0,06497 v(k-2). \end{aligned} \quad (5.5)$$

Resultados

A Tabela 5.1 mostra os índices de validação dos modelos de Wiener e Hammerstein neurais para o Aquecedor.

Tabela 5.1: Índices de validação para identificação do modelo de Wiener e Hammerstein neurais do Aquecedor de Dissipação Variável

Modelo	$RMSE_{dinâmico}$	$\eta_{dinâmico}$	$\eta_{estático}$
Wiener	0,2536	0,0276	0,9999
Hammerstein	0,0814	0,2937	0,9998

Discussão dos resultados

Por meio dos índices de validação percebe-se que para o sistema em estudo o modelo de Wiener apresentou melhores índices. Isto pode ser verificado tanto pelo índice de correlação dinâmica do modelo de Wiener, quanto pelo índice RMSE do modelo de Hammerstein, que por estar bem próximo de 0 indica a incorporação de ruído ao modelo. Por meio da análise do índice de correlação estático verifica-se que tanto o modelo de Hammerstein quanto o de Wiener, conseguem reproduzir tal característica.

5.1.2 Conversor *Buck* - CC-CC

Modelo de Wiener

A rede neural para ajuste da característica estática utilizou uma rede com 3 neurônios na camada oculta, com função de ativação tangente hiperbólica (*tansig*) e 1 neurônio, com função de ativação linear, na saída da rede. A Figura 5.7 mostra o ajuste realizado nesta representação e a Equação (5.6) mostra o modelo que ajusta esta característica.

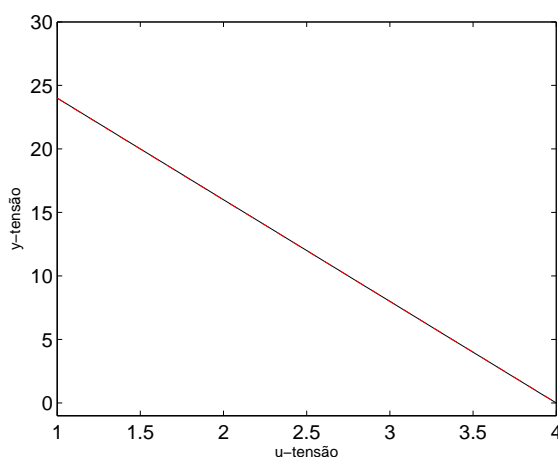


Figura 5.7: Validação da característica estática do sistema. Característica estática do sistema(-). Característica estática resultante do ajuste por redes neurais (- -).

$$\begin{aligned} \bar{y} = & -1,844 \operatorname{tansig}(0,4050 \bar{u} - 0,7346) - 2,032 \operatorname{tansig}(0,3851 \bar{u} + 0,6206) - \\ & - 0,004752 \operatorname{tansig}(12,60 \bar{u} + 13,20) - 0,02859. \end{aligned} \quad (5.6)$$

Treinou-se uma rede com 3 neurônios na camada oculta, com função de ativação tangente hiperbólica (*tansig*) e 1 neurônio na camada de saída, com função de ativação linear, para a estimação da função inversa do modelo de Wiener. A Figura 5.8 apresenta o ajuste da função inversa obtida, cujo modelo está descrito na Equação (5.7).

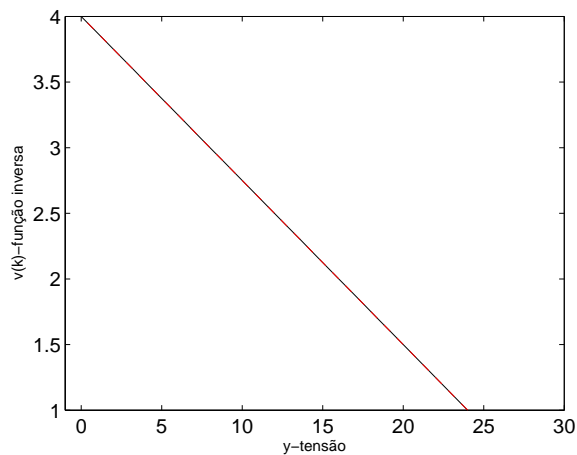


Figura 5.8: Função inversa da característica estática. Característica do sistema(-). Característica resultante do ajuste por redes neurais (- -).

$$\begin{aligned} \bar{v} = & 0,1535 \operatorname{tansig}(6,402 y - 13,07) - 8,745 \operatorname{tansig}(0,1120 y - 0,2102) - \\ & - 0,4174 \operatorname{tansig}(0,4284 y + 1,113) - 1,322. \end{aligned} \quad (5.7)$$

O sinal intermediário \bar{v} foi estimado por meio da função inversa da característica estática. Desta maneira, foi treinada uma rede neural dinâmica, para obtenção do modelo ARX, contendo 1 neurônio linear na camada oculta e 1 neurônio linear na camada de saída. Os atrasos máximos dos regressores de entrada, saída e termos de resíduos foram, respectivamente, 6, 9 e 8. A Figura 5.9 representa o modelo obtido e a Equação (5.8) o descreve. A Tabela 5.2 contém os índices de validação para este modelo obtido.

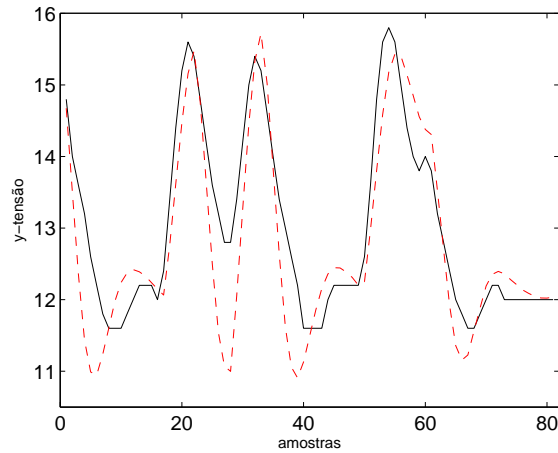


Figura 5.9: Validação do modelo Hammerstein Neural para o conversor *buck* CC-CC. Característica do sistema(-). Característica resultante do ajuste por redes neurais (- -).

$$\begin{aligned}
 \hat{v}(k) = & 1,088 v(k-1) - 0,2558 v(k-2) - 0,3169 v(k-3) - 0,06058 v(k-4) + \\
 & + 0,07876 v(k-5) - 0,04194 v(k-6) + 0,1394 v(k-7) - 0,2146 v(k-8) + \\
 & + 0,04754 v(k-9) + 0,3169 u(k-1) + 0,008468 u(k-2) + 0,03735 u(k-3) + \\
 & + 0,07832 u(k-4) + 0,04241 u(k-5) + 0,04768 u(k-6). \quad (5.8)
 \end{aligned}$$

Modelo de Hammerstein

Na obtenção do modelo de Hammerstein, foi treinada, inicialmente, uma rede neural para obtenção do ajuste da característica estática do sistema. Os parâmetros escolhidos para este treinamento foram 3 neurônios na camada oculta, com função de ativação tangente hiperbólica (*tansig*), e 1 neurônio, com função de ativação linear, na camada de saída. O modelo obtido para a característica estática está representado na Figura 5.10 e descrito na Equação (5.9).

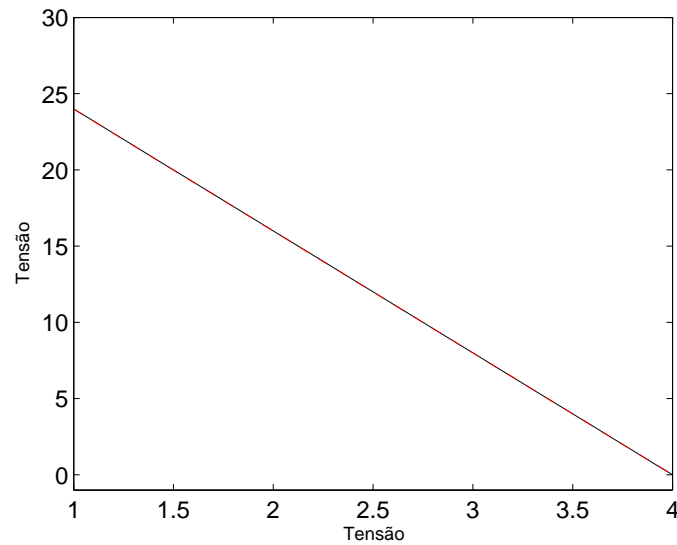


Figura 5.10: Ajuste da característica estática. Característica do sistema(-). Característica resultante do ajuste por redes neurais (- -).

$$\begin{aligned} \bar{y} = & -0,9622 \operatorname{tansig}(0,5132 \bar{u} - 0,9379) - 0,0002103 \operatorname{tansig}(9,864 \bar{u} + 0,1122) - \\ & - 3,338 \operatorname{tansig}(0,2761 \bar{u} + 0,4291) + 0,6434. \end{aligned} \quad (5.9)$$

Para a identificação do bloco linear, foi utilizada uma rede com 1 neurônio na camada oculta e 1 na camada de saída. Os atrasos máximos dos regressores de entrada, saída e termos de resíduos foram, respectivamente, 2, 4 e 3. O modelo ARX obtido está representado na Equação (5.10) e validado qualitativamente na Figura 5.11.

$$\begin{aligned} \hat{y}(k) = & 0,7580 y(k-1) + 0,2142 y(k-2) - 0,3091 y(k-3) - 0,08989 y(k-4) + \\ & + 0,3368 v(k-1) + 0,09629 v(k-2). \end{aligned} \quad (5.10)$$

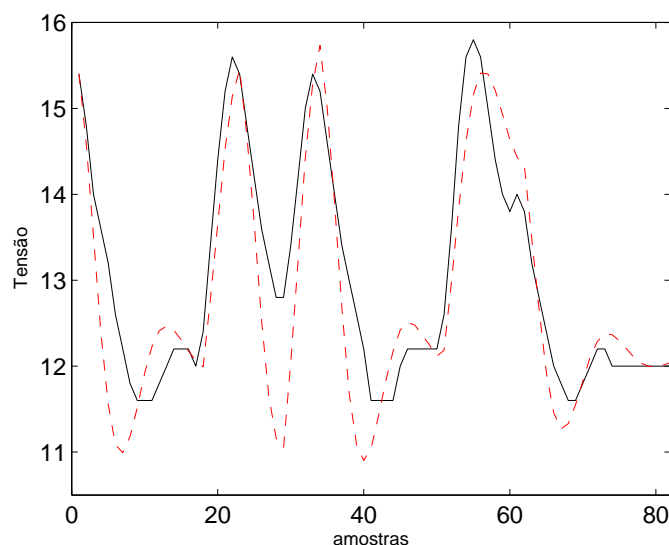


Figura 5.11: Validação do modelo Hammerstein Neural para o conversor *buck* CC-CC. Característica do sistema(-). Característica resultante do ajuste por redes neurais (- -).

Resultados

A Tabela 5.2 mostra os índices de validação dos modelos de Wiener e Hammerstein neurais para o conversor *buck* CC-CC.

Tabela 5.2: Índices de validação para identificação dos modelos de Wiener e Hammerstein neural do conversor *buck* - CC-CC

Modelo	$RMSE_{\text{dinâmico}}$	$\eta_{\text{dinâmico}}$	$\eta_{\text{estático}}$
Wiener	0,5223	0,1519	1,0000
Hammerstein	0,5002	0,1341	1,0000

Discussão dos resultados

Para o conversor *buck* CC-CC tanto os blocos de Wiener quanto os blocos de Hammerstein foram capazes de modelar satisfatoriamente o sistema. Isso pode ser verificado por meio dos índices de validação apresentados na Tabela 5.2. É possível perceber também que, ainda que os dados dinâmicos estejam restritos à uma pequena faixa de operação da curva estática, os modelos baseados em blocos interconectados conseguem reproduzir a característica estática em todos os pontos de operação. Observou-se que, para a massa de dados utilizadas na obtenção do modelo, os blocos de Wiener e Hammerstein tiveram resultados semelhantes. Uma hipótese para tais resultados seria a pouca informação contida nos dados utilizados.

5.1.3 Planta de Neutralização de pH

Modelo de Wiener

O ajuste neural para característica estática utilizou uma rede com 3 neurônios na camada oculta, com função de ativação tangente hiperbólica (*tansig*) e 1 neurônio, com função de ativação linear, na saída da rede. A Figura 5.12 mostra a validação qualitativa do ajuste realizado nesta representação e a Equação (5.11) descreve o modelo que ajusta esta característica.

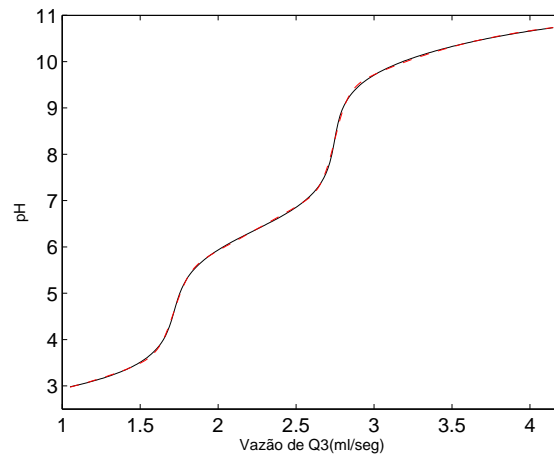


Figura 5.12: Validação da característica estática do sistema. Característica estática do sistema(-). Característica estática resultante do ajuste por redes neurais (- -).

$$\begin{aligned} \bar{y} = & 0,6516 \operatorname{tansig}(1,163 \bar{u} + 0,2247) + 0,2608 \operatorname{tansig}(17,90 \bar{u} - 1,717) + \\ & + 0,2137 \operatorname{tansig}(15,74 \bar{u} + 8,952) - 0,04866. \end{aligned} \quad (5.11)$$

Para a estimação da função inversa do modelo de Wiener, treinou-se uma rede, com 3 neurônios na camada oculta, com função de ativação tangente hiperbólica (*tansig*) e 1 neurônio na camada de saída, com função de ativação linear. A Figura 5.13 apresenta o ajuste da função inversa obtida, cujo modelo está descrito na Equação (5.12).

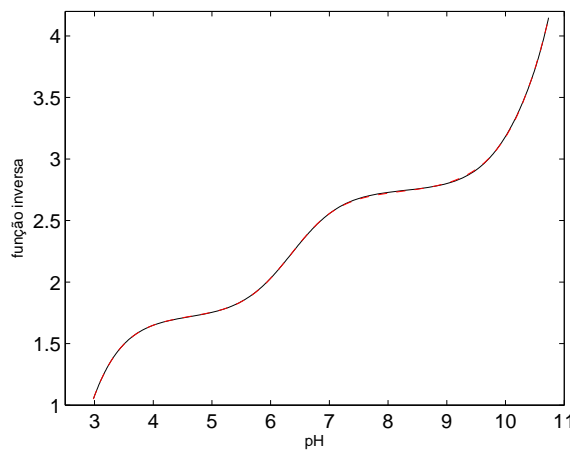


Figura 5.13: Função inversa da característica estática. Característica do sistema(-). Característica resultante do ajuste por redes neurais (- -).

$$\begin{aligned} \hat{v} = & 1,844 \operatorname{tansig}(3,547 y - 4,106) + 0,3240 \operatorname{tansig}(4,617 y + 0,6080) + \\ & + 2,220 \operatorname{tansig}(4,239 y + 5,348) - 0,6160. \end{aligned} \quad (5.12)$$

Após a obtenção da função inversa, e obtido o sinal intermediário \hat{v} , foi treinada uma rede neural dinâmica, para obtenção do modelo ARX, contendo 1 neurônio linear na camada oculta e 1 neurônio linear na camada de saída. Os atrasos máximos dos regressores de entrada, saída e termos de resíduos foram, respectivamente, 2, 3 e 2. A Equação (5.13) descreve o modelo neural de Wiener e a Figura 5.14, sua representação. A Tabela 5.3 contém os índices de validação para este modelo obtido.

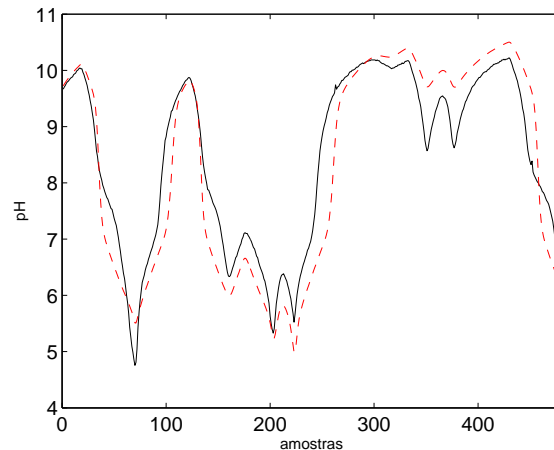


Figura 5.14: Validação do modelo Wiener Neural para a planta de neutralização de pH. Característica do sistema(-). Característica resultante do ajuste por redes neurais (- -).

$$\begin{aligned} \hat{y}(k) = & 1,940 y(k-1) - 1,199 y(k-2) + 0,2558 y(k-3) + \\ & + 0,001000 v(k-1) + 0,002802 v(k-2). \end{aligned} \quad (5.13)$$

Modelo de Hammerstein

Para a obtenção do modelo de Hammerstein Neural, para a planta de neutralização de pH, foi treinada, inicialmente, uma rede neural para obtenção do ajuste da característica estática do sistema. Os parâmetros escolhidos para este treinamento foram 3 neurônios na camada oculta, com função de ativação tangente hiperbólica (*tansig*), e 1 neurônio, com função de ativação linear, na camada de saída. O modelo obtido para a característica estática está representado na Figura 5.15 e descrito na Equação (5.14).

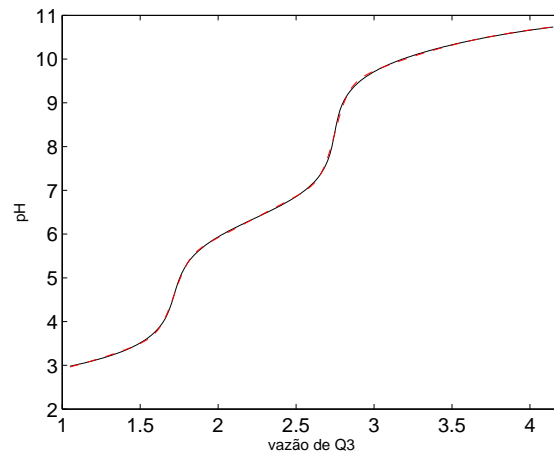


Figura 5.15: Ajuste da característica estática. Característica do sistema(-). Característica resultante do ajuste por redes neurais (- -).

$$\begin{aligned} \bar{y} = & 0,6625 \operatorname{tansig}(1,140 \bar{u} + 0,2432) + 0,2641 \operatorname{tansig}(17,37 \bar{u} - 1,679) + \\ & + 0,2097 \operatorname{tansig}(16,19 \bar{u} + 9,193) - 0,05717. \end{aligned} \quad (5.14)$$

Para a identificação do modelo ARX, dos blocos interconectados de Hammerstein, foi utilizada uma rede com neurônios lineares, 1 na camada e 1 na camada de saída. Os atrasos máximos dos regressores de entrada, saída e termos de resíduos foram, respectivamente, 3, 5 e 5. O modelo ARX obtido está representado na Equação (5.15) e validado qualitativamente na Figura 5.16.

$$\begin{aligned} \hat{y}(k) = & 1,440 y(k-1) - 0,4132 y(k-2) - 0,3245 y(k-3) + \\ & + 0,5308 y(k-4) - 0,2456 y(k-5) + 0,002907 v(k-1) + \\ & + 0,007004 v(k-2) + 0,003934 v(k-3). \end{aligned} \quad (5.15)$$

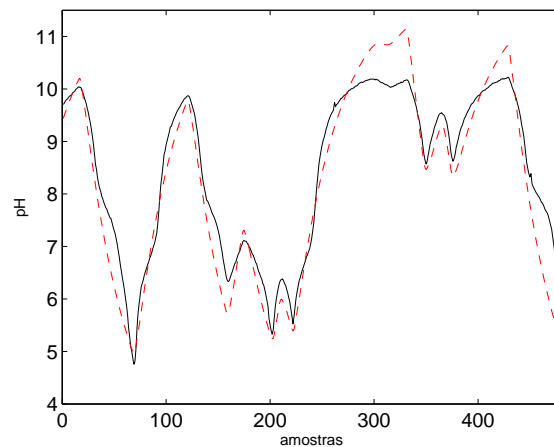


Figura 5.16: Representação do modelo Hammerstein Neural para a planta de neutralização de pH. Característica do sistema(-). Característica resultante do ajuste por redes neurais (- -).

Resultados

A Tabela 5.3 apresenta os índices para validação quantitativa do modelo estimado.

Tabela 5.3: Índices de validação para identificação do modelo de Wiener neural da planta de neutralização de pH

Modelo	$RMSE_{\text{dinâmico}}$	$\eta_{\text{dinâmico}}$	$\eta_{\text{estático}}$
Wiener	0,3606	0,3324	1,0000
Hammerstein	0,2939	0,4606	1,0000

Discussão dos resultados

A identificação do sistema de neutralização de pH apresentou melhores índices nos blocos interconectados de Wiener. Isto pode ser verificado na Tabela 5.3 em que os índices de $\eta_{\text{estático}}$ são idênticos nos dois modelos, porém o índice $\eta_{\text{dinâmico}}$ é menor no modelo de Wiener, que é indicativo de menor incorporação de ruído. Embora o índice $RMSE$ do modelo de Hammerstein apresente-se menor ele indica que, neste modelo, foi incorporado ruído na obtenção dos parâmetros do modelo identificado.

5.2 Treinamento de Redes Neurais com Restrição

As restrições impostas ao treinamento das redes neurais foram baseadas nas estruturas propostas por Narendra and Parthasarathy (1990), descritas em 2.2.3. As redes estáticas, que foram utilizadas para obtenção dos pesos das restrições, foram obtidas por meio do *toolbox* padrão de Redes Neurais do *software Matlab*. Após a obtenção destes, utilizou-se a função *fmincon*, do *toolbox* de Otimização (*Matlab*). A partir desta função foi possível buscar os pesos para a rede dinâmica que, além de atender às restrições impostas pelos pesos estático, minimizaram o erro médio quadrático (função custo), entre o sinal de saída do sistema e a saída da rede. Para a validação dinâmica das redes foi utilizada a função *nmarx* ((Norgaard, 2000)), que simula a rede infinitos passos à frente.

Com a obtenção dos pesos para a rede dinâmica, prosseguiu-se com a verificação da capacidade daquela rede de reproduzir a característica estática, o que comprovou a não violação do conjunto de restrições. Procedeu-se então com a validação destes modelos. A validação dinâmica foi realizada com base nos mesmos critério seguidos para os modelos anteriores, em que metade dos dados foram utilizados para identificação e a outra metade para validação.

Os critérios utilizados para os treinamentos à partir dos dados estáticos foram:

1. número máximo de épocas de treinamento igual à 1000;
2. erro médio quadrático 10^{-4} ;
3. função de ativação não linear - tangente hiperbólica (*tansig*);
4. para outros parâmetros foram tomados os valores padrão (*default*) do pacote de redes neurais.

Nas subseções seguintes, são descritas as estruturas de redes utilizadas para cada treinamento, bem como os modelos obtidos e os respectivos índices de validação.

5.2.1 Aquecedor com Dissipação Variável

A Tabela 5.4 mostra as arquiteturas de redes utilizadas para a obtenção dos pesos estáticos, para imposição das restrições, de cada uma das estruturas propostas por Narendra and Parthasarathy (1990).

Tabela 5.4: Arquiteturas para obtenção dos pesos estáticos (restrições) - Aquecedor

Rede	Estrutura mostrada na Figura 2.7	Neurônios não lineares conectados na entrada (\bar{u})	Neurônios lineares conectados na saída (\bar{y})
1	(a)	4	-
2	(b)	-	4
3	(c)	5	5
4	(d)	6	6

As Equações (5.16), (5.17), (5.18) e (5.19) apresentam os modelos obtidos para cada um dos treinamentos apresentados na Tabela 5.4, respectivamente para as redes 1, 2, 3 e 4.

$$\begin{aligned} \bar{y} = & 5,800 \operatorname{tansig}(0,6051 \bar{u} - 0,4789) - 0,6536 \operatorname{tansig}(0,04059 \bar{u} - 0,1392) - \\ & - 2,120 \operatorname{tansig}(0,08448 \bar{u} - 0,3501) - 5,263 \operatorname{tansig}(0,5105 \bar{u} - 0,1968) + \\ & + 0,7548. \end{aligned} \quad (5.16)$$

$$\begin{aligned} \bar{y} = & 0,9411 \operatorname{tansig}(1,035 \bar{y} - 0,004416) + 1,950 \operatorname{tansig}(0,4227 \bar{y} + 0,01494) - \\ & - 0,3680 \operatorname{tansig}(0,1470 \bar{y} - 0,003018) - 0,6361 \operatorname{tansig}(1,345 \bar{y} + 0,04577) + \\ & + 0,02912 \bar{u} + 0,01096. \end{aligned} \quad (5.17)$$

$$\begin{aligned} \bar{y} = & 0,1048 \operatorname{tansig}(0,2189 \bar{u} - 0,001210) + 2,731 \operatorname{tansig}(0,08292 \bar{u} + 0,03783) + \\ & + 0,4113 \operatorname{tansig}(0,04058 \bar{u} + 0,01248) - 1,307 \operatorname{tansig}(0,09171 \bar{u} - 0,03288) - \\ & - 0,3962 \operatorname{tansig}(0,3439 \bar{u} + 0,01505) - 0,2731 \operatorname{tansig}(0,8853 \bar{y} + 0,1224) + \\ & + 0,5191 \operatorname{tansig}(0,1769 \bar{y} + 0,01476) + 0,3460 \operatorname{tansig}(0,4213 \bar{y} - 0,007138) + \\ & + 0,8226 \operatorname{tansig}(0,9473 \bar{y} - 0,1797) + 0,2326 \operatorname{tansig}(0,8075 \bar{y} - 0,01307) + \\ & + 0,03178. \end{aligned} \quad (5.18)$$

$$\begin{aligned} \bar{y} = & 0,3941 \operatorname{tansig}(0,1693 \bar{y} + 0,002492 \bar{u} - 0,00006709) + 0,9457 \operatorname{tansig}(0,3151 \bar{y} + \\ & + 0,005530 \bar{u} - 0,0008586) + 0,8358 \operatorname{tansig}(0,3569 \bar{y} + 0,004834 \bar{u} - 0,0005859) - \\ & - 0,01870 \operatorname{tansig}(0,03423 \bar{y} - 0,0002162 \bar{u} + 0,00002911) + 0,9134 \operatorname{tansig}(0,3186 \bar{y} + \\ & + 0,005797 \bar{u} - 0,0003450) + 0,1246 \operatorname{tansig}(0,07582 \bar{y} + 0,0007006 \bar{u} - 0,00004171) + \\ & + 0,000001370. \end{aligned} \quad (5.19)$$

As Figuras a seguir mostram a validação qualitativa dos modelos estáticos obtidos para cada uma das redes descritas na Tabela 5.4. A Figura 5.17, (a), (b), (c) e (d) representam os modelos descritos pelas Equações (5.16), (5.17), (5.18) e (5.19), respectivamente.

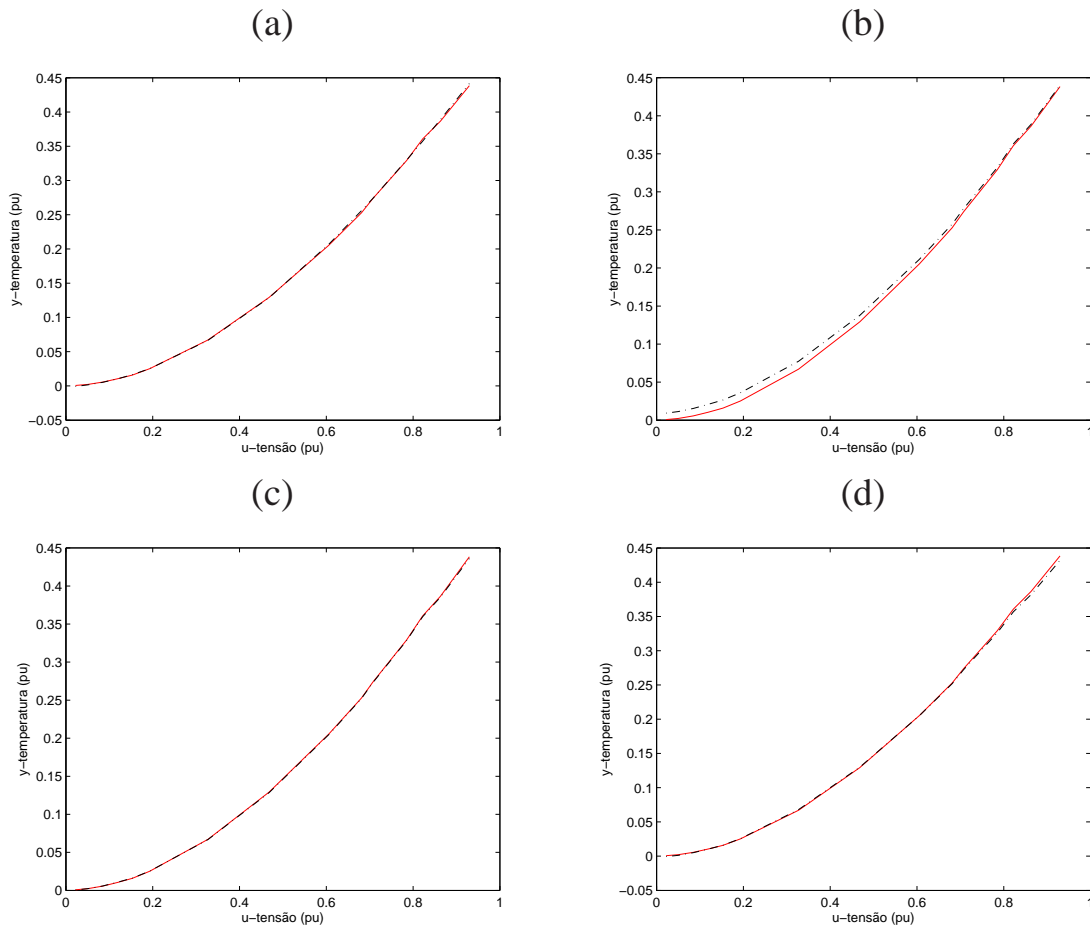


Figura 5.17: Característica estática: reprodução da característica estática da rede em linha tracejada (- -) e característica estática do sistema em linha contínua (-). (a) Rede 1. (b) Rede 2. (c) Rede 3. (d) Rede 4.

A Tabela 5.5 mostra as arquiteturas de redes utilizadas para o treinamento dinâmico, impondo as restrições das redes estáticas.

Tabela 5.5: Estruturas utilizadas para o treinamento dinâmico das redes neurais com imposição de restrições - Aquecedor

Rede	Estrutura mostrada na Figura 2.7	Neurônios não lineares conectados aos regressores de entrada	Neurônios não lineares conectados aos regressores de saída	Num. de neurônios lineares	Máximo atraso da entrada	Máximo atraso da saída
1	(a)	4	-	1	4	5
2	(b)	-	4	1	7	8
3	(c)	5	5	-	5	6
4	(d)	6	6	-	8	9

As Equações (5.20), (5.21), (5.22) e (5.23) apresentam os modelos obtidos para cada um dos treinamentos apresentados na Tabela 5.5, respectivamente para as redes 1, 2, 3 e 4.

$$\begin{aligned}
\hat{y}(k) = & 1,512y(k-1) - 0,1788y(k-2) - 0,4339y(k-3) - 0,03470y(k-4) + \\
& + 0,1151y(k-5) + 0,01497 + 0,1150 \operatorname{tansig}(0,8556u(k-1) + 1,337u(k-2) - \\
& - 1,381u(k-3) - 0,2069u(k-4) - 0,4789) - 0,01296 \operatorname{tansig}(5,749u(k-1) - \\
& - 26,72u(k-2) + 16,70u(k-3) + 4,315u(k-4) - 0,1392) + \\
& + 0,04204 \operatorname{tansig}(1,415u(k-1) - 1,086u(k-2) - 0,5366u(k-3) + 0,1230u(k-4) + \\
& + 0,3501) - 0,1044 \operatorname{tansig}(0,5413u(k-1) + 1,686u(k-2) - 1,482u(k-3) - \\
& - 0,2347u(k-4) - 0,1968). \tag{5.20}
\end{aligned}$$

$$\begin{aligned}
\hat{y}(k) = & 0,9411 \operatorname{tansig}(0,4884y(k-1) - 0,4825y(k-2) - 0,3969y(k-3) - \\
& - 0,07209y(k-4) + 0,8741y(k-5) + 0,3943y(k-6) - -0,9385y(k-7) + \\
& + 1,168y(k-8) - 0,004416) + 1,950 \operatorname{tansig}(0,5873y(k-1) - 0,4620y(k-2) + \\
& + 0,5024y(k-3) - 0,05544y(k-4) - 0,1857y(k-5) + 0,01884y(k-6) - \\
& - 0,07385y(k-7) + 0,09116y(k-8) + 0,01494) + 0,3680 \operatorname{tansig}(1,607y(k-1) + \\
& + 0,3481y(k-2) - 1,212y(k-3) - 0,5786y(k-4) - 0,3596y(k-5) + \\
& + 0,4503y(k-6) + 1,206y(k-7) - 1,608y(k-8) + 0,003018) - \\
& - 0,6361 \operatorname{tansig}(1,124y(k-1) - 1,226y(k-2) - 0,1930y(k-3) + \\
& + 0,2271y(k-4) + 0,7442y(k-5) + 0,5214y(k-6) - 1,391y(k-7) + \\
& + 1,538y(k-8) + 0,04577) + 0,03471u(k-1) - 0,008505u(k-2) - \\
& - 0,004266u(k-3) - 0,007988u(k-4) + 0,0003871u(k-5) + 0,01140u(k-6) + \\
& + 0,003381u(k-7) + 0,01096. \tag{5.21}
\end{aligned}$$

$$\begin{aligned}
\hat{y}(k) = & -0,2731 \operatorname{tansig}(1,023y(k-1) + 0,2801y(k-2) - 0,6957y(k-3) - \\
& - 1,038y(k-4) + 0,6137y(k-5) + 0,7026y(k-6) + 0,1224) - \\
& - 0,5191 \operatorname{tansig}(0,3510y(k-1) - 1,567y(k-2) - 0,5421y(k-3) + \\
& - 0,8556y(k-4) - 0,4786y(k-5) + 1,205y(k-6) - 0,01476) + \\
& + 0,3460 \operatorname{tansig}(1,603y(k-1) - 1,371y(k-2) - 0,8931y(k-3) - \\
& - 0,7447y(k-4) + 3,971y(k-5) - 2,144y(k-6) - 0,007138) + \\
& + 0,8226 \operatorname{tansig}(1,157y(k-1) - 0,7111y(k-2) + 0,2708y(k-3) - \\
& - 0,2235y(k-4) - 1,180y(k-5) + 1,635y(k-6) - 0,1797) + \\
& + 0,2326 \operatorname{tansig}(1,232y(k-1) - 1,084y(k-2) - 0,3369y(k-3) + \\
& + 1,267y(k-4) - 0,6341y(k-5) + 0,3642y(k-6) - 0,01307) - \\
& - 0,1048 \operatorname{tansig}(0,1762u(k-1) - 0,4097u(k-2) - 0,6163u(k-3) + \\
& + 1,295u(k-4) - 0,6644u(k-5) + 0,001210) + 2,731 \operatorname{tansig}(0,1048u(k-1) - \\
& - 0,09490u(k-2) + 0,1118u(k-3) - 0,1213u(k-4) + 0,08259u(k-5) + \\
& + 0,03783) + 0,4113 \operatorname{tansig}(0,4067u(k-1) - 0,6642u(k-2) + \\
& + 0,6741u(k-3) - 0,4853u(k-4) + 0,1093u(k-5) + 0,01248) - \\
& - 1,307 \operatorname{tansig}(0,2880u(k-1) - 0,4104u(k-2) + 0,4694u(k-3) - \\
& - 0,4818u(k-4) + 0,2264u(k-5) - 0,03288) - \\
& - 0,3962 \operatorname{tansig}(0,09990u(k-1) + 0,1491u(k-2) + 0,01901u(k-3) - \\
& - 0,03382u(k-4) + 0,1097u(k-5) + 0,01505) + 0,03178.
\end{aligned} \tag{5.22}$$

$$\begin{aligned}
\hat{y}(k) = & -0,00002507 - 0,9457 \operatorname{tansig}(-2,433y(k-1) + 376,8y(k-2) + \\
& + 55,85y(k-3) - 152,1y(k-4) - 372,3y(k-5) - 328,4y(k-6) - \\
& - 2,288y(k-7) + 369,7y(k-8) + 54,89y(k-9) - 723,9u(k-1) + \\
& + 296,8u(k-2) + 1461,0u(k-3) + 475,8u(k-4) - 181,0u(k-5) - \\
& - 882,5u(k-6) - 742,7u(k-7) + 296,4u(k-8) + 0,0008586) + \\
& + 0,8358 \operatorname{tansig}(11,59y(k-1) + 226,8y(k-2) + 183,9y(k-3) - \\
& - 134,9y(k-4) - 498,5y(k-5) - 190,8y(k-6) + 7,963y(k-7) + \\
& + 218,1y(k-8) + 176,2y(k-9) - 1181,0u(k-1) - 178,1u(k-2) + \\
& + 483,8u(k-3) + 1188,0u(k-4) + 1048,0u(k-5) + 4,208u(k-6) - \\
& - 1186,0u(k-7) - 178,8u(k-8) - 0,0005859) - \\
& - 0,01870 \operatorname{tansig}(-2,420y(k-1) + 352,6y(k-2) + 52,15y(k-3) - \\
& - 142,0y(k-4) - 347,4y(k-5) - 306,5y(k-6) - 2,258y(k-7) + \\
& + 344,7y(k-8) + 51,08y(k-9) - 107,7u(k-1) - 806,7u(k-2) - \\
& - 667,4u(k-3) + 368,0u(k-4) + 1549,0u(k-5) + 549,5u(k-6) - \\
& - 99,03u(k-7) - 785,8u(k-8) + 0,00002911) - \\
& - 0,9134 \operatorname{tansig}(10,29y(k-1) + 211,4y(k-2) + 171,3y(k-3) - \\
& - 126,6y(k-4) - 466,4y(k-5) - 178,8y(k-6) + 7,664y(k-7) + \\
& + 205,1y(k-8) + 165,8y(k-9) + 734,9u(k-1) - 282,5u(k-2) - \\
& - 1443,0u(k-3) - 460,9u(k-4) + 174,6u(k-5) + 846,9u(k-6) + \\
& + 713,0u(k-7) - 283,0u(k-8) + 0,0003450) + \\
& + 0,1246 \operatorname{tansig}(-1,947y(k-1) + 331,7y(k-2) + 49,33y(k-3) - \\
& - 134,1y(k-4) - 328,4y(k-5) - 289,7y(k-6) - 1,842y(k-7) + \\
& + 326,5y(k-8) + 48,63y(k-9) + 1128,0u(k-1) + 166,1u(k-2) - \\
& - 455,8u(k-3) - 1114,0u(k-4) - 982,5u(k-5) - 8,052u(k-6) + \\
& + 1104,0u(k-7) + 162,8u(k-8) - 0,00004171) .
\end{aligned} \tag{5.23}$$

As figuras a seguir mostram a validação qualitativa dos modelos dinâmicos obtidos após a imposição das restrições. A Figura 5.18, (a), (b), (c) e (d) representa os modelos descritos pelas Equações (5.16), (5.17), (5.18) e (5.19), respectivamente.

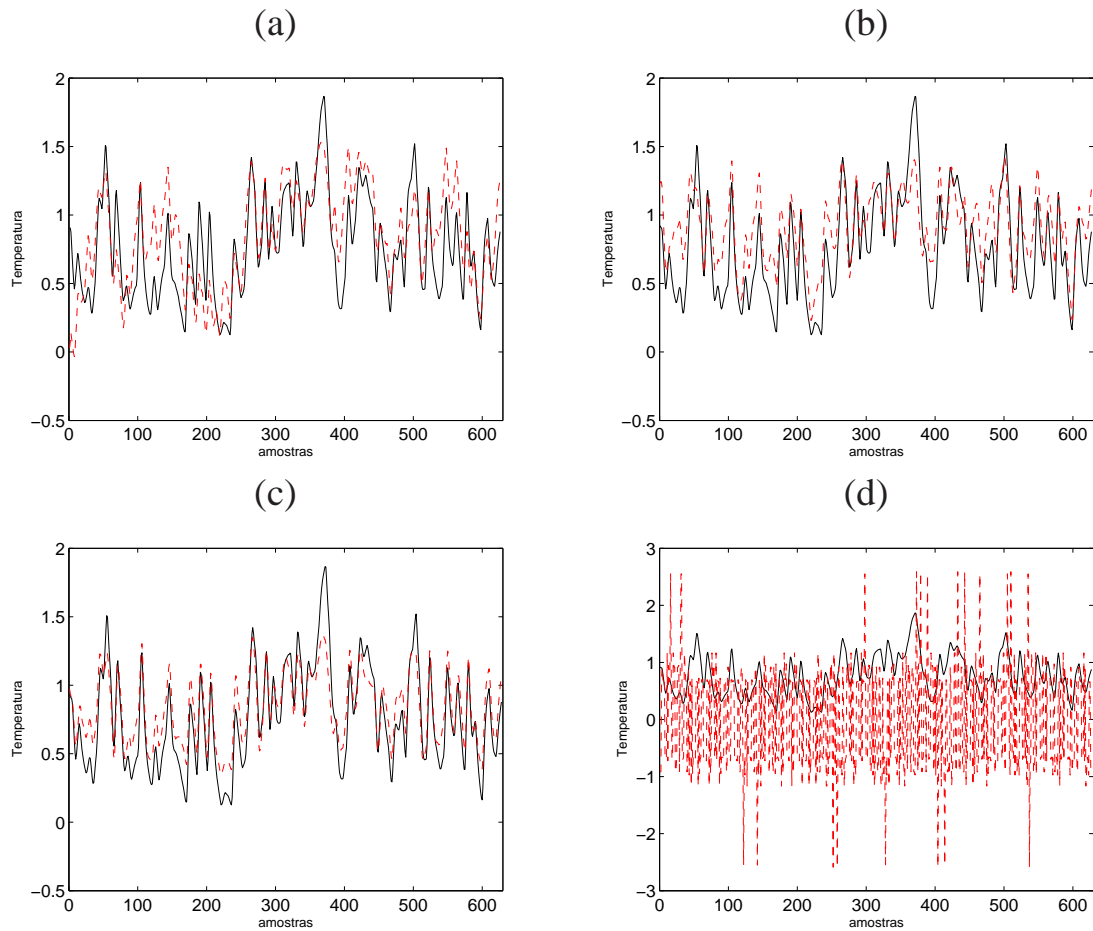


Figura 5.18: Característica dinâmica: reprodução da característica dinâmica da rede em linha tracejada (- -) e característica dinâmica do sistema em linha contínua (-). (a) Rede 1. (b) Rede 2. (c) Rede 3. (d) Rede 4.

Partindo do modelo dinâmico de cada uma das estruturas propostas, verificou-se a capacidade de reprodução da respectiva característica estática. A Figura 5.19, (a), (b), (c) e (d), mostra tal reprodução

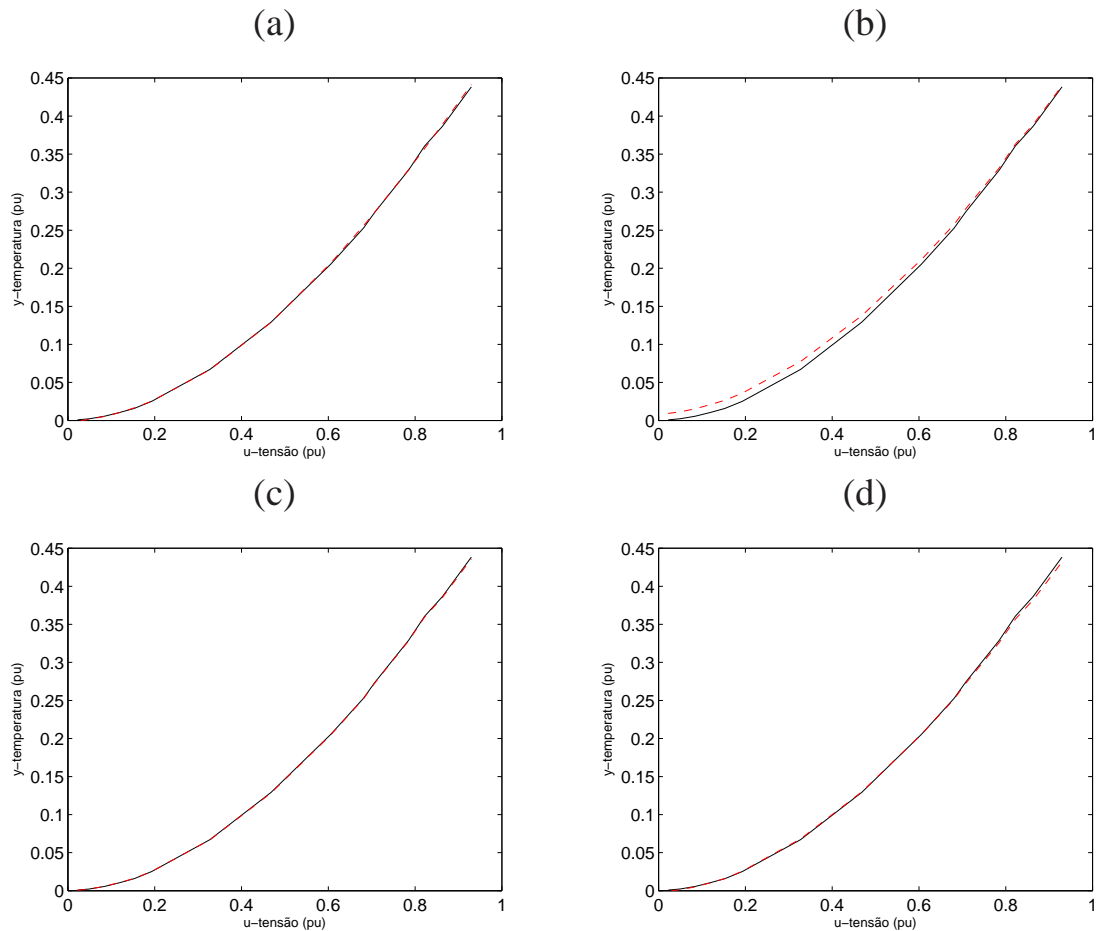


Figura 5.19: Característica estática: reprodução da característica estática da rede em linha tracejada (- -) e característica estática do sistema em linha contínua (-). (a) Rede 1. (b) Rede 2. (c) Rede 3. (d) Rede 4.

A Tabela 5.6 mostra os índices de validação para cada uma das estruturas de treinamento utilizadas.

Tabela 5.6: Índices de validação para identificação caixa cinza, com imposição de restrições, do sistema Aquecedor.

Rede	Estrutura Figura 2.7	$RMSE_{\text{dinâmico}}$	$\eta_{\text{dinâmico}}$	$\eta_{\text{estático}}$
1	(a)	0,7601	0,3729	1,0000
2	(b)	0,8710	0,7236	1,0000
3	(c)	0,6740	0,7950	1,0000
4	(d)	1,0234	0,3199	1,0000

Discussão dos resultados

Conforme pode ser verificado na Tabela 5.6, a rede 1 apresentou melhor resultado, comparando-se os índices de validação de todas as estruturas utilizadas na imposição de restrições. Nas redes 2 e 3, observou-se que os modelos incorporaram ruído, analisando os índices de correlação dinâmicos. A rede 4 não conseguiu modelar o sistema pois seu índice $RMSE$ foi superior à unidade. De acordo com os treinamentos obtidos para o caixa preta, utilizando as mesmas estruturas para as

conexões nas redes, esperava-se que a rede 4 apresentasse resultado mais significativo. Porém por esta ser uma rede completamente conectada, a imposição de restrição limita muito o ajuste dos pesos. Tal característica pode ser verificada pelo aumento dos valores absolutos dos pesos na rede em questão.

5.2.2 Conversor *Buck* CC-CC

A Tabela 5.7 descreve as arquiteturas de redes utilizadas para a obtenção dos pesos estáticos, para imposição das restrições, de cada uma das estruturas propostas por Narendra and Parthasarathy (1990).

Tabela 5.7: Arquiteturas para obtenção dos pesos estáticos (restrições) - conversor *buck* CC-CC

Rede	Estrutura mostrada na Figura 2.7	Neurônios não lineares conectados na entrada (\bar{u})	Neurônios lineares conectados na saída (\bar{y})
1	(a)	6	-
2	(b)	-	4
3	(c)	4	4
4	(d)	9	9

As Equações (5.24), (5.25), (5.26) e (5.27) apresentam os modelos obtidos para cada um dos treinamentos apresentados na Tabela 5.7, respectivamente para as redes 1, 2, 3 e 4.

$$\begin{aligned} \bar{y} = & -9,045 \operatorname{tansig}(0,02248 \bar{u} + 0,08163) - 41,26 \operatorname{tansig}(0,03222 \bar{u} - 0,02740) - \\ & - 53,53 \operatorname{tansig}(0,05209 \bar{u} - 0,1244) - 33,05 \operatorname{tansig}(0,02966 \bar{u} - 0,001764) - \\ & - 92,64 \operatorname{tansig}(0,02782 \bar{u} + 0,03375) - 5,560 \operatorname{tansig}(0,03000 \bar{u} - 0,02487) + \\ & + 27,93. \end{aligned} \quad (5.24)$$

$$\begin{aligned} \bar{y} = & 7,237 \operatorname{tansig}(1,701 \bar{y} + 1,697) + 5,694 \operatorname{tansig}(2,301 \bar{y} + 0,6665) + \\ & + 6,979 \operatorname{tansig}(0,6562 \bar{y} + 2,609) - 2,160 \operatorname{tansig}(2,672 \bar{y} - 0,3489) - \\ & - 8,0 \bar{u} + 14,25. \end{aligned} \quad (5.25)$$

$$\begin{aligned} \bar{y} = & 3,579 \operatorname{tansig}(4,576 \bar{u} + 7,775) - 13,86 \operatorname{tansig}(0,4630 \bar{u} - 0,2268) + \\ & + 0,9916 \operatorname{tansig}(5,220 \bar{u} + 6,073) + 1,991 \operatorname{tansig}(2,150 \bar{u} + 4,073) + \\ & + 3,602 \operatorname{tansig}(0,7686 \bar{y} + 4,590) + 3,943 \operatorname{tansig}(0,1154 \bar{y} + 0,4038) + \\ & + 9,496 \operatorname{tansig}(0,06611 \bar{y} - 0,5065) - 0,8118 \operatorname{tansig}(0,4383 \bar{y} + 3,367) + \\ & + 6,399. \end{aligned} \quad (5.26)$$

$$\begin{aligned}
\bar{y} = & -1,180 \operatorname{tansig}(6,242 \bar{y} + 1,249 \bar{u} + 0,5080) + 3,168 \operatorname{tansig}(0,1665 \bar{y} - \\
& - 0,1551 \bar{u} - 0,03409) + 2,273 \operatorname{tansig}(0,1486 \bar{y} + 0,1327 \bar{u} + 0,03705) + \\
& + 4,968 \operatorname{tansig}(0,08085 \bar{y} - 0,6280 \bar{u} - 0,1540) + 3,399 \operatorname{tansig}(1,072 \bar{y} + \\
& + 0,6020 \bar{u} + 0,1828) + 5,101 \operatorname{tansig}(0,3234 \bar{y} + 0,5711 \bar{u} + 0,1524) + \\
& + 2,540 \operatorname{tansig}(0,08098 \bar{y} - 1,419 \bar{u} - 0,3505) + 1,283 \operatorname{tansig}(0,2475 \bar{y} + \\
& + 0,09598 \bar{u} + 0,03086) + 2,582 \operatorname{tansig}(0,1385 \bar{y} - 0,4380 \bar{u} - 0,1060) + \\
& + 2,861.
\end{aligned} \tag{5.27}$$

As figuras a seguir mostram a validação qualitativa dos modelos estáticos obtidos para cada uma das redes descritas na Tabela 5.7. A Figura 5.20, (a), (b), (c) e (d) representam os modelos descritos pelas Equações (5.24), (5.25), (5.26) e (5.27), respectivamente.

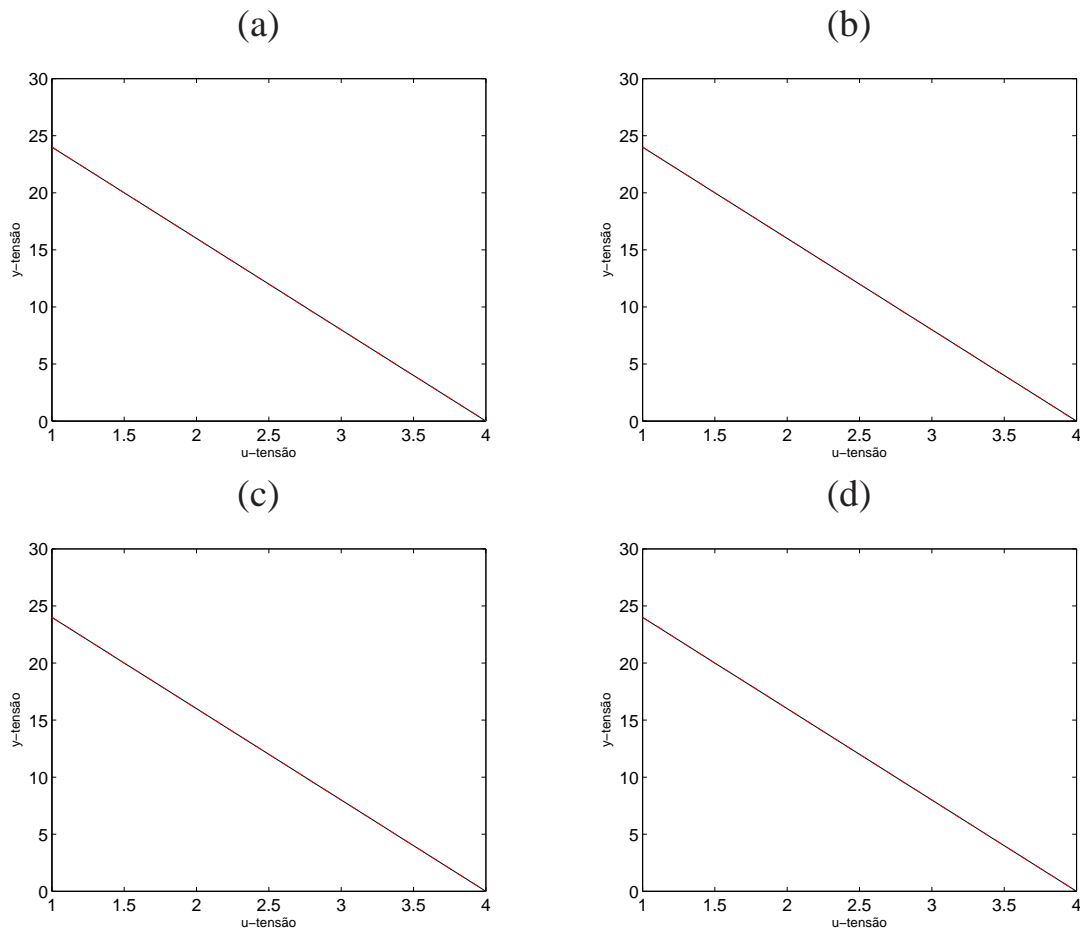


Figura 5.20: Característica estática: reprodução da característica estática da rede em linha tracejada (- -) e característica estática do sistema em linha contínua (-). (a) Rede 1. (b) Rede 2. (c) Rede 3. (d) Rede 4.

A Tabela 5.8 mostra as arquiteturas de redes utilizadas para o treinamento dinâmico, impondo as restrições das redes estáticas.

Tabela 5.8: Estruturas utilizadas para o treinamento dinâmico das redes neurais com imposição de restrições - conversor *buck* CC-CC

Rede	Estrutura mostrada na Figura 2.7	Neurônios não lineares conectados aos regressores de entrada	Neurônios não lineares conectados aos regressores de saída	Núm. de neurônios lineares	Máximo atraso da entrada	Máximo atraso da saída
1	(a)	6	-	1	1	2
2	(b)	-	4	1	9	10
3	(c)	4	4	-	5	6
4	(d)	9	9	-	10	11

As Equações (5.28), (5.29), (5.30) e (5.32) apresentam os modelos obtidos para cada um dos treinamentos apresentados na Tabela 4.3, respectivamente para as redes 1, 2, 3 e 4.

$$\begin{aligned}
\hat{y}(k) = & 1,337y(k-1) - 0,6042y(k-2) + 7,468 - 2,418 \operatorname{tansig}(0,02248u(k-1) - \\
& - 0,08163) - 11,03 \operatorname{tansig}(0,03222u(k-1) - 0,02740) - \\
& - 14,31 \operatorname{tansig}(0,05209u(k-1) - 0,1244) - 8,836 \operatorname{tansig}(0,02966u(k-1) - \\
& - 0,001764) - 24,77 \operatorname{tansig}(0,02782u(k-1) + 0,03375) - \\
& - 1,486 \operatorname{tansig}(0,03000u(k-1) - 0,02487). \tag{5.28}
\end{aligned}$$

$$\begin{aligned}
\hat{y}(k) = & 7,237 \operatorname{tansig}(0,1702y(k-1) + 0,1702y(k-2) + 0,1703y(k-3) + \\
& + 0,1702y(k-4) + 0,1702y(k-5) + 0,1701y(k-6) + 0,1700y(k-7) + \\
& + 0,1700y(k-8) + 0,1700y(k-9) + 0,1700y(k-10) + 1,697) + \\
& + 5,694 \operatorname{tansig}(0,2301y(k-1) + 0,2300y(k-2) + 0,2301y(k-3) + \\
& + 0,2301y(k-4) + 0,2300y(k-5) + 0,2301y(k-6) + 0,2301y(k-7) + \\
& + 0,2301y(k-8) + 0,2301y(k-9) + 0,2301y(k-10) + 0,6665) + \\
& + 6,979 \operatorname{tansig}(0,06561y(k-1) + 0,06567y(k-2) + 0,06567y(k-3) + \\
& + 0,06567y(k-4) + 0,06561y(k-5) + 0,06567y(k-6) + 0,06561y(k-7) + \\
& + 0,06561y(k-8) + 0,06554y(k-9) + 0,06555y(k-10) + 2,609) - \\
& - 2,160 \operatorname{tansig}(0,2673y(k-1) + 0,2672y(k-2) + 0,2673y(k-3) + \\
& + 0,2673y(k-4) + 0,2673y(k-5) + 0,2672y(k-6) + 0,2671y(k-7) + \\
& + 0,2671y(k-8) + 0,2671y(k-9) + 0,2671y(k-10) - 0,3489) - \\
& - 4,040u(k-1) - 2,430u(k-2) - 2,013u(k-3) - 0,8393u(k-4) - \\
& - 0,3167u(k-5) + 0,04460u(k-6) + 0,7003u(k-7) + 0,8192u(k-8) + \\
& + 0,07473u(k-9) + 14,25. \tag{5.29}
\end{aligned}$$

$$\begin{aligned}
\hat{y}(k) = & 3,602 \operatorname{tansig}(0,1280y(k-1) + 0,1282y(k-2) + 0,1282y(k-3) + \\
& + 0,1282y(k-4) + 0,1280y(k-5) + 0,1281y(k-6) + 4,590) + \\
& + 3,943 \operatorname{tansig}(0,4815y(k-1) - 0,8763y(k-2) + 0,5799y(k-3) - \\
& - 0,1100y(k-4) - 0,07630y(k-5) + 0,1166y(k-6) + 0,4038) + \\
& + 9,496 \operatorname{tansig}(0,06780y(k-1) + 0,06100y(k-2) - 0,04182y(k-3) - \\
& - 0,02290y(k-4) - 0,01257y(k-5) + 0,01460y(k-6) - 0,5065) - \\
& - 0,8118 \operatorname{tansig}(0,07295y(k-1) + 0,07299y(k-2) + 0,07326y(k-3) + \\
& + 0,07313y(k-4) + 0,07292y(k-5) + 0,07309y(k-6) + 3,367) + \\
& + 3,579 \operatorname{tansig}(0,9151u(k-1) + 0,9151u(k-2) + 0,9152u(k-3) + \\
& + 0,9151u(k-4) + 0,9152u(k-5) + 7,775) - \\
& - 13,86 \operatorname{tansig}(0,4114u(k-1) + 0,09542u(k-2) - 0,01216u(k-3) + \\
& + 0,01951u(k-4) - 0,05121u(k-5) - 0,2268) + \\
& + 0,9916 \operatorname{tansig}(1,044u(k-1) + 1,044u(k-2) + 1,044u(k-3) + \\
& + 1,044u(k-4) + 1,044u(k-5) + 6,073) + 1,991 \operatorname{tansig}(0,4299u(k-1) + \\
& + 0,4299u(k-2) + 0,4301u(k-3) + 0,4301u(k-4) + 0,4302u(k-5) + \\
& + 4,073) + 6,399.
\end{aligned} \tag{5.30}$$

$$\begin{aligned}
\hat{y}(k) = & 2, 204 - 3, 168 \operatorname{tansig} (35330, 0y(k - 1) - 41460, 0y(k - 2) - \\
& - 24600, 0y(k - 3) + 64600, 0y(k - 4) - 13430, 0y(k - 5) + 24670, 0y(k - 6) - \\
& - 16680, 0y(k - 7) + 929, 6y(k - 8) - 23240, 0y(k - 9) + 35300, 0y(k - 10) - \\
& - 41410, 0y(k - 11) - 3887, 0u(k - 1) + 6684, 0u(k - 2) - 7170, 0u(k - 3) - \\
& - 4128, 0u(k - 4) + 11970, 0u(k - 5) - 2112, 0u(k - 6) + 4761, 0u(k - 7) - \\
& - 2699, 0u(k - 8) + 475, 9u(k - 9) - 3890, 0u(k - 10) + 0, 03409) - \\
& - 2, 273 \operatorname{tansig} (28760, 0y(k - 1) - 60360, 0y(k - 2) + \\
& + 17600, 0y(k - 3) - 20460, 0y(k - 4) + 20850, 0y(k - 5) + 3266, 0y(k - 6) + \\
& + 7440, 0y(k - 7) - 31120, 0y(k - 8) + 45620, 0y(k - 9) + 28770, 0y(k - 10) - \\
& - 60370, 0y(k - 11) - 5630, 0u(k - 1) + 8231, 0u(k - 2) + 5188, 0u(k - 3) - \\
& - 10910, 0u(k - 4) + 3171, 0u(k - 5) - 3706, 0u(k - 6) + 3758, 0u(k - 7) + \\
& + 582, 2u(k - 8) + 4952, 0u(k - 9) - 5633, 0u(k - 10) - 0, 03705) + \\
& + 4, 968 \operatorname{tansig} (14990, 0y(k - 1) - 23090, 0y(k - 2) + 18240, 0y(k - 3) + \\
& + 657, 2y(k - 4) + 24860, 0y(k - 5) - 33770, 0y(k - 6) + 43060, 0y(k - 7) + \\
& + 26190, 0y(k - 8) - 63050, 0y(k - 9) + 15010, 0y(k - 10) - 23100, 0y(k - 11) + \\
& + 6852, 0u(k - 1) + 3806, 0u(k - 2) - 12310, 0u(k - 3) + 1788, 0u(k - 4) - \\
& - 5094, 0u(k - 5) + 2375, 0u(k - 6) - 806, 2u(k - 7) + 3559, 0u(k - 8) - \\
& - 7015, 0u(k - 9) + 6842, 0u(k - 10) - 0, 1540) + \\
& + 3, 399 \operatorname{tansig} (15790, 0y(k - 1) - 1803, 0y(k - 2) + 22440, 0y(k - 3) - \\
& - 36290, 0y(k - 4) + 40680, 0y(k - 5) + 23780, 0y(k - 6) - 65630, 0y(k - 7) + \\
& + 12580, 0y(k - 8) - 25610, 0y(k - 9) + 15840, 0y(k - 10) - 1776, 0y(k - 11) + \\
& + 4107, 0u(k - 1) - 11990, 0u(k - 2) + 2091, 0u(k - 3) - 4784, 0u(k - 4) + \\
& + 2678, 0u(k - 5) - 500, 3u(k - 6) + 3860, 0u(k - 7) - 6703, 0u(k - 8) + \\
& + 7140, 0u(k - 9) + 4101, 0u(k - 10) + 0, 1828) - \\
& - 5, 101 \operatorname{tansig} (25040, 0y(k - 1) - 33860, 0y(k - 2) + 43320, 0y(k - 3) + \\
& + 26380, 0y(k - 4) - 63280, 0y(k - 5) + 15150, 0y(k - 6) - 23140, 0y(k - 7) + \\
& + 18420, 0y(k - 8) + 758, 0y(k - 9) + 25150, 0y(k - 10) - 33940, 0y(k - 11) - \\
& - 10370, 0u(k - 1) + 3696, 0u(k - 2) - 3172, 0u(k - 3) + 4282, 0u(k - 4) + \\
& + 1107, 0u(k - 5) + 5463, 0u(k - 6) - 5089, 0u(k - 7) + 8740, 0u(k - 8) + \\
& + 5704, 0u(k - 9) - 10360, 0u(k - 10) - 0, 1524) + \\
& + 2, 540 \operatorname{tansig} (36280, 0y(k - 1) + 19280, 0y(k - 2) - 70670, 0y(k - 3) + \\
& + 8010, 0y(k - 4) - 30410, 0y(k - 5) + 11290, 0y(k - 6) - 6427, 0y(k - 7) + \\
& + 18050, 0y(k - 8) - 41240, 0y(k - 9) + 36450, 0y(k - 10) + 19400, 0y(k - 11) + \\
& + 2287, 0u(k - 1) - 4574, 0u(k - 2) + 2873, 0u(k - 3) - 298, 9u(k - 4) + \\
& + 4052, 0u(k - 5) - 6488, 0u(k - 6) + 7325, 0u(k - 7) + 4293, 0u(k - 8) - \\
& - 11750, 0u(k - 9) + 2283, 0u(k - 10) - 0, 3505) + \\
& + 1, 283 \operatorname{tansig} (60110, 0y(k - 1) - 18830, 0y(k - 2) + 19710, 0y(k - 3) - \\
& - 22120, 0y(k - 4) - 4340, 0y(k - 5) - 28870, 0y(k - 6) + 30560, 0y(k - 7) - \\
& - 47330, 0y(k - 8) - 30230, 0y(k - 9) + 60240, 0y(k - 10) - 18900, 0y(k - 11) + \\
& + 3883, 0u(k - 1) - 3555, 0u(k - 2) - 386, 6u(k - 3) - 4733, 0u(k - 4) +
\end{aligned}$$

$$\begin{aligned}
& + 5796,0u(k-5) - 8002,0u(k-6) - 4973,0u(k-7) + 11050,0u(k-8) - \\
& - 2965,0u(k-9) + 3880,0u(k-10) + 0,03086) - \\
& - 2,582 \operatorname{tansig}(23740,0y(k-1) - 18200,0y(k-2) - 365,1y(k-3) - \\
& - 24930,0y(k-4) + 34580,0y(k-5) - 43410,0y(k-6) - 26290,0y(k-7) + \\
& + 64310,0y(k-8) - 14940,0y(k-9) + 23760,0y(k-10) - 18240,0y(k-11) - \\
& - 2809,0u(k-1) + 356,2u(k-2) - 3984,0u(k-3) + 6530,0u(k-4) - \\
& - 7249,0u(k-5) - 4224,0u(k-6) + 11780,0u(k-7) - 2219,0u(k-8) + \\
& + 4618,0u(k-9) - 2802,0u(k-10) + 0,1060). \tag{5.32}
\end{aligned}$$

As figuras a seguir mostram a validação qualitativa dos modelos dinâmicos obtidos após a imposição das restrições. A Figura 5.21, (a), (b), (c) e (d) representa os modelos descritos pelas Equações (5.24), (5.25), (5.26) e (5.27), respectivamente.

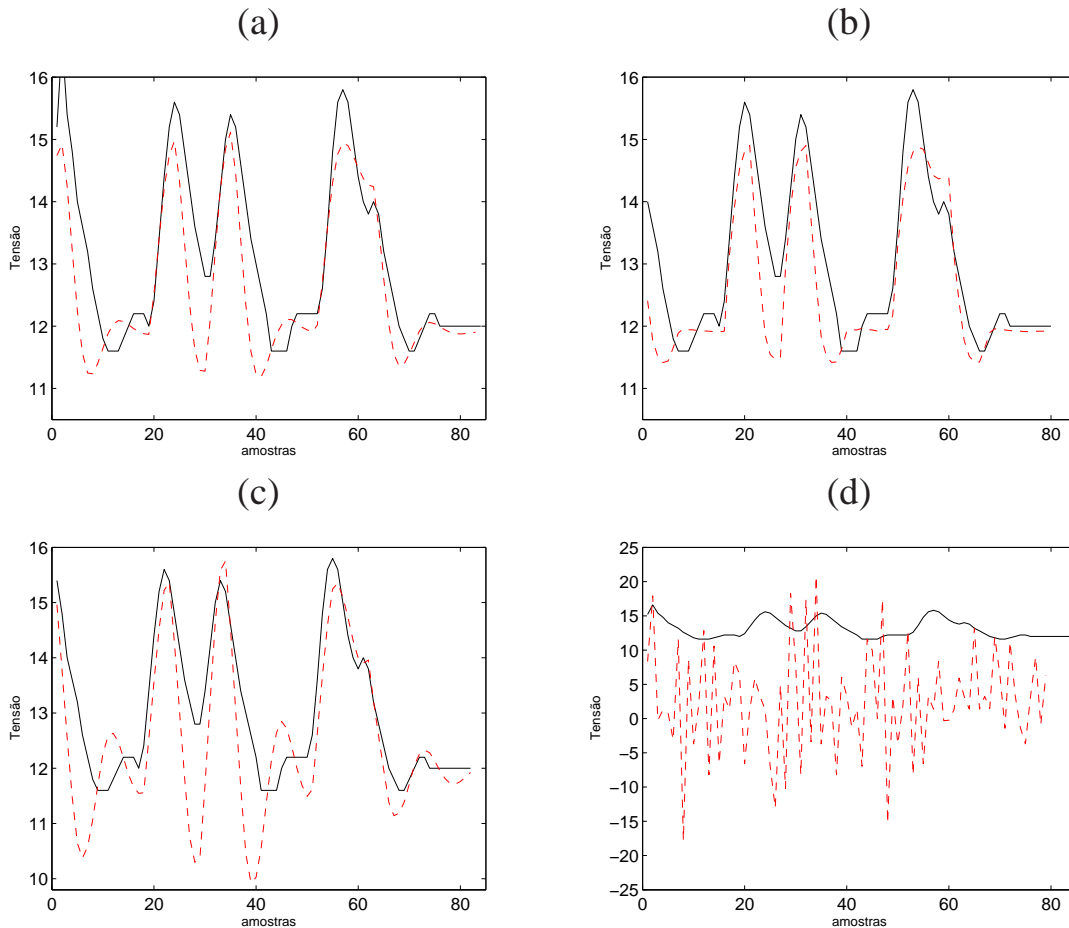


Figura 5.21: Característica dinâmica: reprodução da característica dinâmica da rede em linha tracejada (- -) e característica dinâmica do sistema em linha contínua (-). (a) Rede 1. (b) Rede 2. (c) Rede 3. (d) Rede 4.

Partindo do modelo dinâmico de cada uma das estruturas propostas, verificou-se a capacidade de reprodução da respectiva característica estática. A Figura 5.22, (a), (b), (c) e (d), mostra tal reprodução

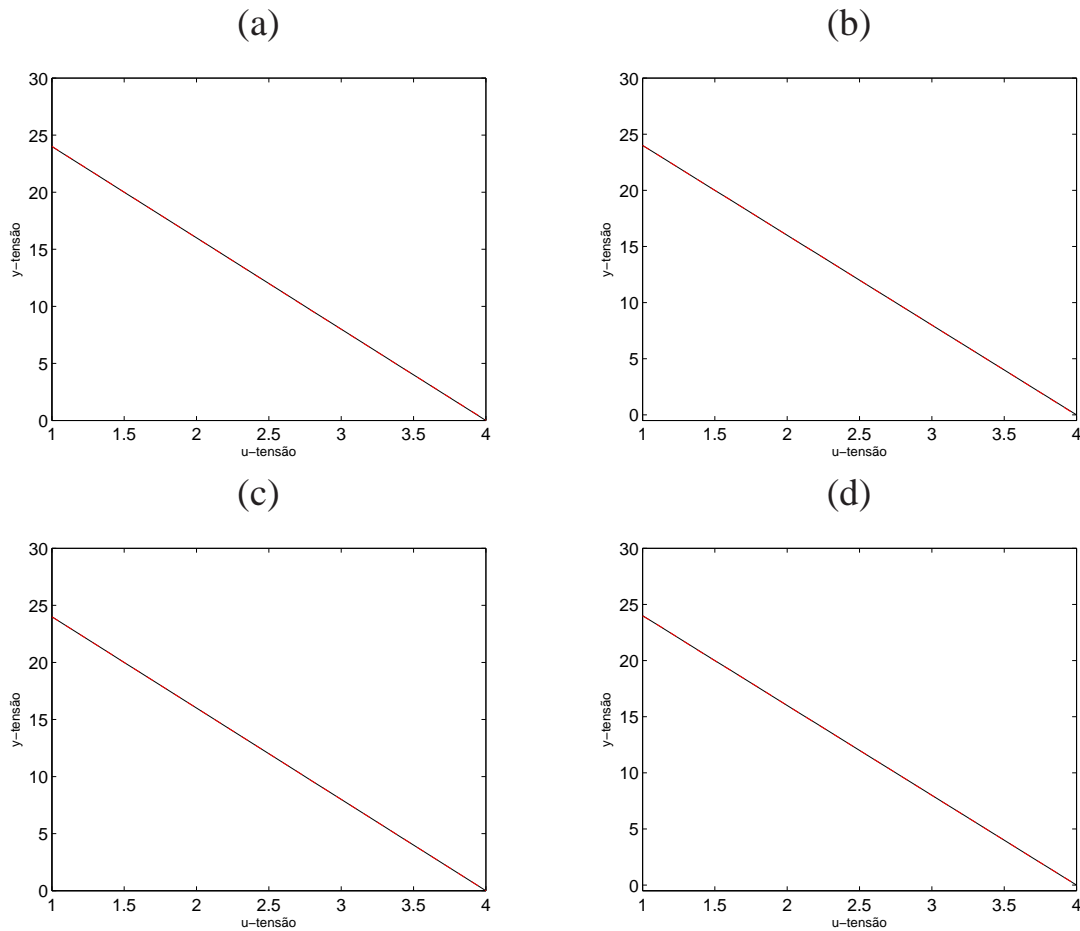


Figura 5.22: Característica estática: reprodução da característica estática da rede em linha tracejada (- -) e característica estática do sistema em linha contínua (-). (a) Rede 1. (b) Rede 2. (c) Rede 3. (d) Rede 4.

A Tabela 5.9 mostra os índices de validação para cada uma das estruturas de treinamento utilizadas.

Tabela 5.9: Índices de validação para identificação caixa cinza, com imposição de restrições, do sistema conversor *buck* CC-CC.

Rede	Estrutura Figura 2.7	$RMSE_{\text{dinâmico}}$	$\eta_{\text{dinâmico}}$	$\eta_{\text{estático}}$
1	(a)	0,9663	0,3629	1,0000
2	(b)	0,5857	0,3892	1,0000
3	(c)	0,6705	0,1432	1,0000
4	(d)	1,0021	0,1183	1,0000

Discussão dos resultados

A Tabela 5.9 demonstra que a rede 1 apresentou melhor resultado e menor incorporação de ruído ao modelo. A rede 2 teve um índice de correlação dinâmica um pouco maior que a rede 1, porém o $RMSE$ apresentou um resultado inferior. A rede 3, apesar de manter índices de correlação dinâmica mais baixo, numa análise visual na Figura 5.21(c) é possível perceber que este modelo não atinge os máximos e mínimos locais do sistema. A rede 4, de forma análoga ao sistema do aquecedor de

dissipação variável, além de não modelar o sistema, os pesos sinápticos tiveram valores muito grandes. Diversas tentativas foram feitas para tentar obter um modelo satisfatório, porém foi verificado que, mesmo aumentando cada vez mais o número de atrasos e neurônios, os resultados permaneciam sempre com a mesma característica.

5.2.3 Planta de Neutralização de pH

A Tabela 5.10 descreve as arquiteturas de redes utilizadas para a obtenção dos pesos estáticos, para imposição das restrições, de cada uma das estruturas propostas por Narendra and Parthasarathy (1990).

Tabela 5.10: Arquiteturas para obtenção dos pesos estáticos (restrições) - planta de neutralização de pH

Rede	Estrutura mostrada na Figura 2.7	Neurônios não lineares conectados na entrada (\bar{u})	Neurônios lineares conectados na saída (\bar{y})
1	(a)	6	-
2	(b)	-	3
3	(c)	7	5
4	(d)	8	8

As Equações (5.33), (5.34), (5.35) e (5.36) apresentam os modelos obtidos para cada um dos treinamentos apresentados na Tabela 5.10, respectivamente para as redes 1, 2, 3 e 4.

$$\begin{aligned} \bar{y} = & -70,58 \operatorname{tansig}(0,2895 \bar{u} + 0,4797) - 26,24 \operatorname{tansig}(20,48 \bar{u} - 56,57) + \\ & + 34,23 \operatorname{tansig}(0,3719 \bar{u} - 0,2060) + 0,7873 \operatorname{tansig}(10,59 \bar{u} - 18,24) + \\ & + 26,74 \operatorname{tansig}(20,51 \bar{u} - 56,64) + 0,6168 \operatorname{tansig}(5,936 \bar{u} - 16,39) + \\ & + 44,86. \end{aligned} \quad (5.33)$$

$$\begin{aligned} \bar{y} = & -2,704 \operatorname{tansig}(0,5260y(k-1) + 0,5257y(k-2) + 0,5257y(k-3) + 0,7932) - \\ & - 2,191 \operatorname{tansig}(0,7746y(k-1) + 0,7746y(k-2) + 0,7741y(k-3) - 0,3061) + \\ & + 14,59 \operatorname{tansig}(0,08473y(k-1) - 0,01384y(k-2) - 0,002137y(k-3) - 0,4586) + \\ & + 0,003573u(k-1) + 0,04420u(k-2) + 11,46. \end{aligned} \quad (5.34)$$

$$\begin{aligned} \bar{y} = & -1,782 \operatorname{tansig}(0,02353 \bar{u} - 0,6290) + 2,356 \operatorname{tansig}(0,03714 \bar{u} + 0,8523) - \\ & - 0,3236 \operatorname{tansig}(0,6012 \bar{u} - 0,2340) + 0,3131 \operatorname{tansig}(0,1367 \bar{u} + 0,1704) + \\ & + 2,448 \operatorname{tansig}(0,1731 \bar{u} + 0,9017) + 3,541 \operatorname{tansig}(0,2511 \bar{y} - 2,327) + \\ & + 1,752 \operatorname{tansig}(0,7938 \bar{y} - 0,06988) + 2,508 \operatorname{tansig}(0,3078 \bar{y} - 1,122) - \\ & - 2,218 \operatorname{tansig}(6,939 \bar{y} - 0,2520) - 0,3657 \operatorname{tansig}(1,369 \bar{y} + 0,03666) - \\ & - 0,5299 \operatorname{tansig}(2,021 \bar{y} + 0,2099) + 2,466 \operatorname{tansig}(4,667 \bar{y} + 0,4542) + \\ & + 1,107. \end{aligned} \quad (5.35)$$

$$\begin{aligned}
\bar{y} = & 1,516 \operatorname{tansig}(0,07117 \bar{y} - 0,001124 \bar{u} + 0,1555) + 3,300 \operatorname{tansig}(0,1397 \bar{y} + \\
& + 0,02724 \bar{u} - 1,525) + 1,703 \operatorname{tansig}(0,06401 \bar{y} + 0,01755 \bar{u} + 0,1526) + \\
& + 1,543 \operatorname{tansig}(0,06990 \bar{y} + 0,002308 \bar{u} + 0,1546) + 1,955 \operatorname{tansig}(0,07862 \bar{y} - \\
& - 0,02016 \bar{u} + 0,09346) + 1,898 \operatorname{tansig}(0,07818 \bar{y} - 0,01935 \bar{u} + 0,1014) + \\
& + 1,871 \operatorname{tansig}(0,08114 \bar{y} - 0,02576 \bar{u} + 0,1170) + \\
& + 2,236 \operatorname{tansig}(0,06378 \bar{y} + 0,01649 \bar{u} + 0,1782) + 1,366.
\end{aligned} \tag{5.36}$$

As figuras a seguir mostram a validação qualitativa dos modelos estáticos obtidos para cada uma das redes descritas na Tabela 5.10. A Figura 5.23, (a), (b), (c) e (d) representam os modelos descritos pelas Equações (5.33), (5.34), (5.35) e (5.36), respectivamente.

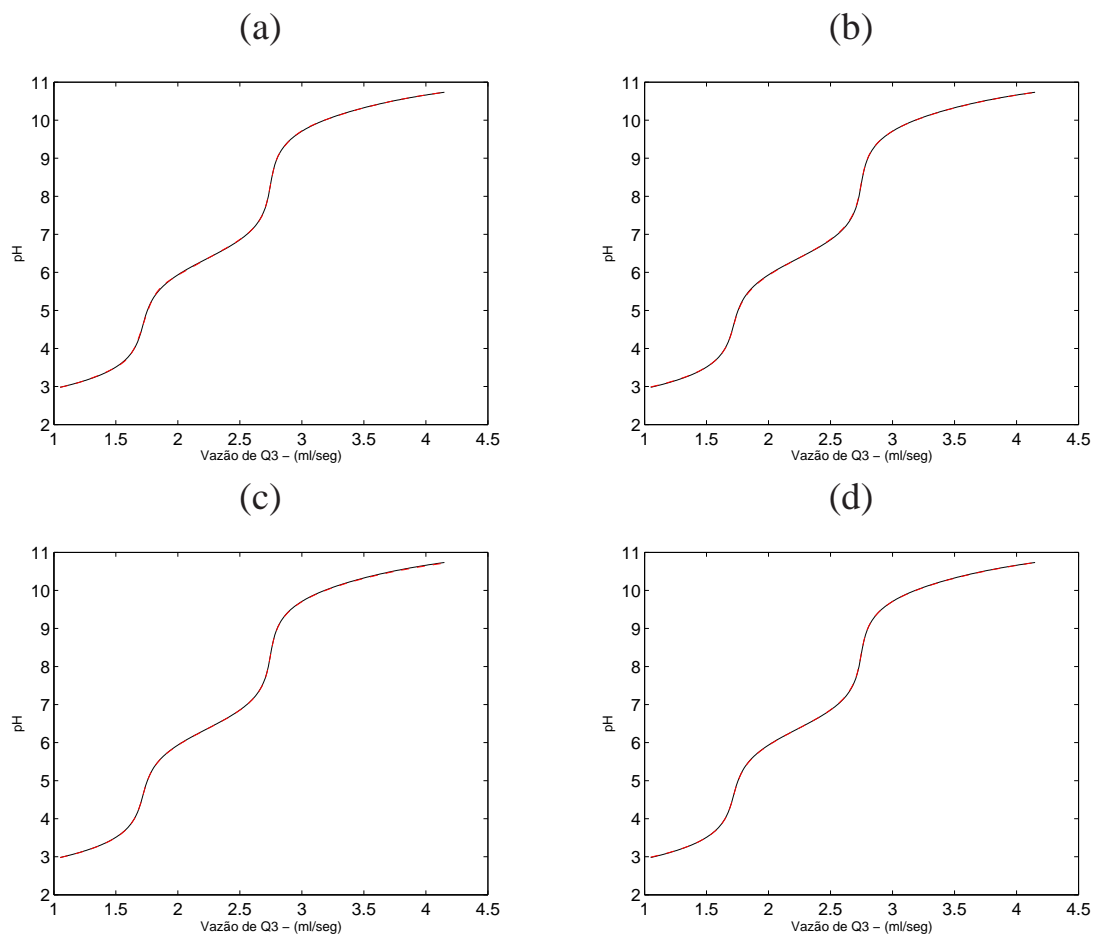


Figura 5.23: Característica estática: reprodução da característica estática da rede em linha tracejada (- -) e característica estática do sistema em linha contínua (-). (a) Rede 1. (b) Rede 2. (c) Rede 3. (d) Rede 4.

A Tabela 5.11 mostra as arquiteturas de redes utilizadas para o treinamento dinâmico, impondo as restrições das redes estáticas.

Tabela 5.11: Estruturas utilizadas para o treinamento dinâmico das redes neurais com imposição de restrições - planta de neutralização de pH

Rede	Estrutura mostrada na Figura 2.7	Neurônios não lineares conectados aos regressores de entrada	Neurônios não lineares conectados aos regressores de saída	Núm. de neurônios lineares	Máximo atraso da entrada	Máximo atraso da saída
1	(a)	6	-	1	1	2
2	(b)	-	3	1	2	3
3	(c)	7	5	-	10	11
4	(d)	8	8	-	7	8

As Equações (5.37), (5.38), (5.39) e (5.40) apresentam os modelos obtidos para cada um dos treinamentos apresentados na Tabela 4.5, respectivamente para as redes 1, 2, 3 e 4.

$$\begin{aligned}
\hat{y}(k) = & 1,241y(k-1) - 0,2608y(k-2) + 0,8711 - 1,370 \operatorname{tansig}(0,2895u(k-1) + \\
& + 0,4797) - 0,5094 \operatorname{tansig}(20,48u(k-1) - 56,57) + \\
& + 0,6647 \operatorname{tansig}(0,3719u(k-1) - 0,2060) + 0,01529 \operatorname{tansig}(10,59u(k-1) - \\
& - 18,24) + 0,5192 \operatorname{tansig}(20,51u(k-1) - 56,64) + 0,01198 \operatorname{tansig}(5,936u(k-1) - \\
& - 16,39). \tag{5.37}
\end{aligned}$$

$$\begin{aligned}
\hat{y}(k) = & -2,704 \operatorname{tansig}(0,5260y(k-1) + 0,5257y(k-2) + 0,5257y(k-3) + 0,7932) - \\
& - 2,191 \operatorname{tansig}(0,7746y(k-1) + 0,7746y(k-2) + 0,7741y(k-3) - 0,3061) + \\
& + 14,59 \operatorname{tansig}(0,08473y(k-1) - 0,01384y(k-2) - 0,002137y(k-3) - 0,4586) + \\
& + 0,003573u(k-1) + 0,04420u(k-2) + 11,46. \tag{5.38}
\end{aligned}$$

$$\begin{aligned}
\hat{y}(k) = & -3,541 \operatorname{tansig}(-204,4y(k-1) - 262,9y(k-2) - 160,6y(k-3) - \\
& - 167,8y(k-4) - 68,64y(k-5) - 7,600y(k-6) + 149,2y(k-7) + \\
& + 118,9y(k-8) + 214,7y(k-9) + 253,7y(k-10) + 135,2y(k-11) + 2,327) + \\
& + 1,752 \operatorname{tansig}(-10,23y(k-1) + 12,36y(k-2) - 4,723y(k-3) + 4,369y(k-4) - \\
& - 14,05y(k-5) + 25,61y(k-6) - 21,70y(k-7) + 1,696y(k-8) + 11,91y(k-9) - \\
& - 13,65y(k-10) + 9,197y(k-11) - 0,06988) + 2,508 \operatorname{tansig}(-16,99y(k-1) - \\
& - 56,66y(k-2) + 65,98y(k-3) - 46,20y(k-4) + 20,85y(k-5) - 30,26y(k-6) + \\
& + 33,01y(k-7) - 26,70y(k-8) - 10,82y(k-9) + 53,39y(k-10) + 14,70y(k-11) - \\
& - 1,122) + 2,218 \operatorname{tansig}(-4,751y(k-1) - 4,109y(k-2) - 3,350y(k-3) - \\
& - 2,501y(k-4) - 1,623y(k-5) - 0,6525y(k-6) + 0,3558y(k-7) + 1,153y(k-8) + \\
& + 2,016y(k-9) + 2,812y(k-10) + 3,711y(k-11) + 0,2520) - \\
& - 0,3657 \operatorname{tansig}(-60,97y(k-1) - \\
& - 61,02y(k-2) - 62,01y(k-3) - 58,04y(k-4) - 48,05y(k-5) - 27,68y(k-6) + \\
& + 5,758y(k-7) + 29,11y(k-8) + 56,18y(k-9) + 95,81y(k-10) + 132,3y(k-11) + \\
& + 0,03666) - 0,5299 \operatorname{tansig}(0,07742y(k-1) + 0,09515y(k-2) + 0,1302y(k-3) + \\
& + 0,1292y(k-4) + 0,1633y(k-5) + 0,1608y(k-6) + 0,2305y(k-7) + \\
& + 0,2465y(k-8) + 0,2277y(k-9) + 0,2629y(k-10) + 0,2975y(k-11) + 0,2099) - \\
& - 2,466 \operatorname{tansig}(-2,003y(k-1) - 1,761y(k-2) - 1,449y(k-3) - \\
& - 1,138y(k-4) - 0,7924y(k-5) - 0,3942y(k-6) - 0,06577y(k-7) + \\
& + 0,2455y(k-8) + 0,5741y(k-9) + 0,8684y(k-10) + 1,249y(k-11) - 0,4542) + \\
& + 1,782 \operatorname{tansig}(-176,4u(k-1) - 163,4u(k-2) - 121,3u(k-3) - \\
& - 65,90u(k-4) - 5,357u(k-5) + 52,47u(k-6) + 100,1u(k-7) + 128,9u(k-8) + \\
& + 132,3u(k-9) + 118,6u(k-10) + 0,6290) + 2,356 \operatorname{tansig}(-131,5u(k-1) - \\
& - 138,5u(k-2) - 108,3u(k-3) - 60,39u(k-4) - 4,590u(k-5) + 50,30u(k-6) + \\
& + 93,60u(k-7) + 113,6u(k-8) + 102,6u(k-9) + 83,27u(k-10) + 0,8523) - \\
& - 0,3236 \operatorname{tansig}(-58,98u(k-1) - 28,41u(k-2) - 20,73u(k-3) - \\
& - 13,39u(k-4) + 0,01279u(k-5) + 3,326u(k-6) + 6,996u(k-7) + 10,29u(k-8) + \\
& + 12,97u(k-9) + 88,50u(k-10) - 0,2340) - 0,3131 \operatorname{tansig}(-16,27u(k-1) + \\
& + 5,861u(k-2) - 11,61u(k-3) - 9,763u(k-4) - 7,234u(k-5) - \\
& - 15,76u(k-6) + 0,4012u(k-7) + 17,91u(k-8) + 25,36u(k-9) + \\
& + 10,97u(k-10) - 0,1704) - 2,448 \operatorname{tansig}(-235,1u(k-1) - 177,4u(k-2) - \\
& - 112,1u(k-3) - 68,31u(k-4) - 33,75u(k-5) + 1,861u(k-6) + \\
& + 55,99u(k-7) + 134,2u(k-8) + 202,4u(k-9) + 232,1u(k-10) - \\
& - 0,9017) + 1,107.
\end{aligned} \tag{5.39}$$

$$\begin{aligned}
\hat{y}(k) = & 1,604 + 3,300 \operatorname{tansig}(-27410, 0y(k-1) + 28630, 0y(k-2) - 32430, 0y(k-3) - \\
& - 28140, 0y(k-4) - 39910, 0y(k-5) + 64400, 0y(k-6) - 37400, 0y(k-7) + \\
& + 72270, 0y(k-8) + 20720, 0u(k-1) - 10170, 0u(k-2) + 7253, 0u(k-3) - \\
& - 11730, 0u(k-4) - 10400, 0u(k-5) - 14050, 0u(k-6) + 18370, 0u(k-7) - \\
& - 1,525) - 1,703 \operatorname{tansig}(-27380, 0y(k-1) + 28600, 0y(k-2) - \\
& - 32410, 0y(k-3) - 28120, 0y(k-4) - 39880, 0y(k-5) + 64340, 0y(k-6) - \\
& - 37370, 0y(k-7) + 72200, 0y(k-8) - 8727, 0u(k-1) + 25370, 0u(k-2) - \\
& - 5689, 0u(k-3) + 11830, 0u(k-4) - 7260, 0u(k-5) - 5919, 0u(k-6) - \\
& - 9597, 0u(k-7) - 0,1526) - \\
& - 1,543 \operatorname{tansig}(-27360, 0y(k-1) + 28580, 0y(k-2) - 32370, 0y(k-3) - \\
& - 28090, 0y(k-4) - 39840, 0y(k-5) + 64280, 0y(k-6) - 37330, 0y(k-7) + \\
& + 72130, 0y(k-8) + 18370, 0u(k-1) - 13450, 0u(k-2) + 20830, 0u(k-3) - \\
& - 10390, 0u(k-4) + 7215, 0u(k-5) - 11970, 0u(k-6) - 10620, 0u(k-7) - \\
& - 0,1546) - 1,955 \operatorname{tansig}(-27330, 0y(k-1) + 28550, 0y(k-2) - \\
& - 32350, 0y(k-3) - 28070, 0y(k-4) - 39800, 0y(k-5) + 64230, 0y(k-6) - \\
& - 37300, 0y(k-7) + 72070, 0y(k-8) - 13790, 0u(k-1) + 18980, 0u(k-2) - \\
& - 13000, 0u(k-3) + 21450, 0u(k-4) - 9914, 0u(k-5) + 7765, 0u(k-6) - \\
& - 11500, 0u(k-7) - 0,09346) + 1,898 \operatorname{tansig}(-27310, 0y(k-1) + \\
& + 28530, 0y(k-2) - 32320, 0y(k-3) - 28040, 0y(k-4) - 39770, 0y(k-5) + \\
& + 64170, 0y(k-6) - 37260, 0y(k-7) + 72010, 0y(k-8) - 10340, 0u(k-1) - \\
& - 14050, 0u(k-2) + 18860, 0u(k-3) - 13260, 0u(k-4) + 21340, 0u(k-5) - \\
& - 10150, 0u(k-6) + 7594, 0u(k-7) + 0,1014) - \\
& - 1,871 \operatorname{tansig}(-27280, 0y(k-1) + 28500, 0y(k-2) - 32290, 0y(k-3) - \\
& - 28010, 0y(k-4) - 39730, 0y(k-5) + 64100, 0y(k-6) - 37230, 0y(k-7) + \\
& + 71940, 0y(k-8) - 8971, 0u(k-1) - 7612, 0u(k-2) - 11340, 0u(k-3) + \\
& + 21700, 0u(k-4) - 10540, 0u(k-5) + 24190, 0u(k-6) - 7420, 0u(k-7) - \\
& - 0,1170) + 2,236 \operatorname{tansig}(-27250, 0y(k-1) + 28470, 0y(k-2) - 32250, 0y(k-3) - \\
& - 27990, 0y(k-4) - 39690, 0y(k-5) + 64040, 0y(k-6) - 37190, 0y(k-7) + \\
& + 71860, 0y(k-8) + 7852, 0u(k-1) - 11550, 0u(k-2) - 10190, 0u(k-3) - \\
& - 13920, 0u(k-4) + 19220, 0u(k-5) - 13130, 0u(k-6) + 21720, 0u(k-7) + \\
& + 0,1782).
\end{aligned} \tag{5.40}$$

As figuras a seguir mostram a validação qualitativa dos modelos dinâmicos obtidos após a imposição das restrições. A Figura 5.24, (a), (b), (c) e (d) representa os modelos descritos pelas Equações (5.33), (5.34), (5.35) e (5.36), respectivamente.

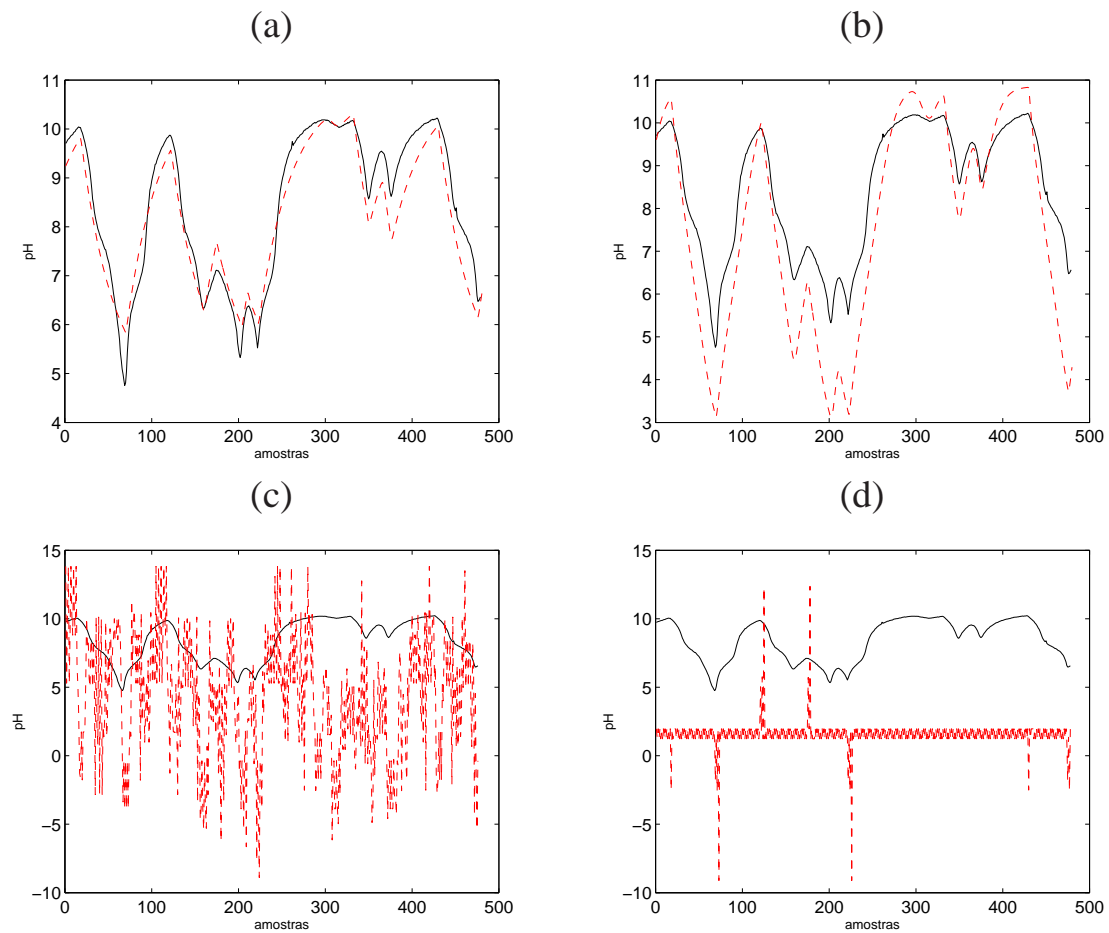


Figura 5.24: Característica dinâmica: reprodução da característica dinâmica da rede em linha tracejada (- -) e característica dinâmica do sistema em linha contínua (-). (a) Rede 1. (b) Rede 2. (c) Rede 3. (d) Rede 4.

Partindo do modelo dinâmico de cada uma das estruturas propostas, verificou-se a capacidade de reprodução da respectiva característica estática. A Figura 5.25, (a), (b), (c) e (d), mostra tal reprodução

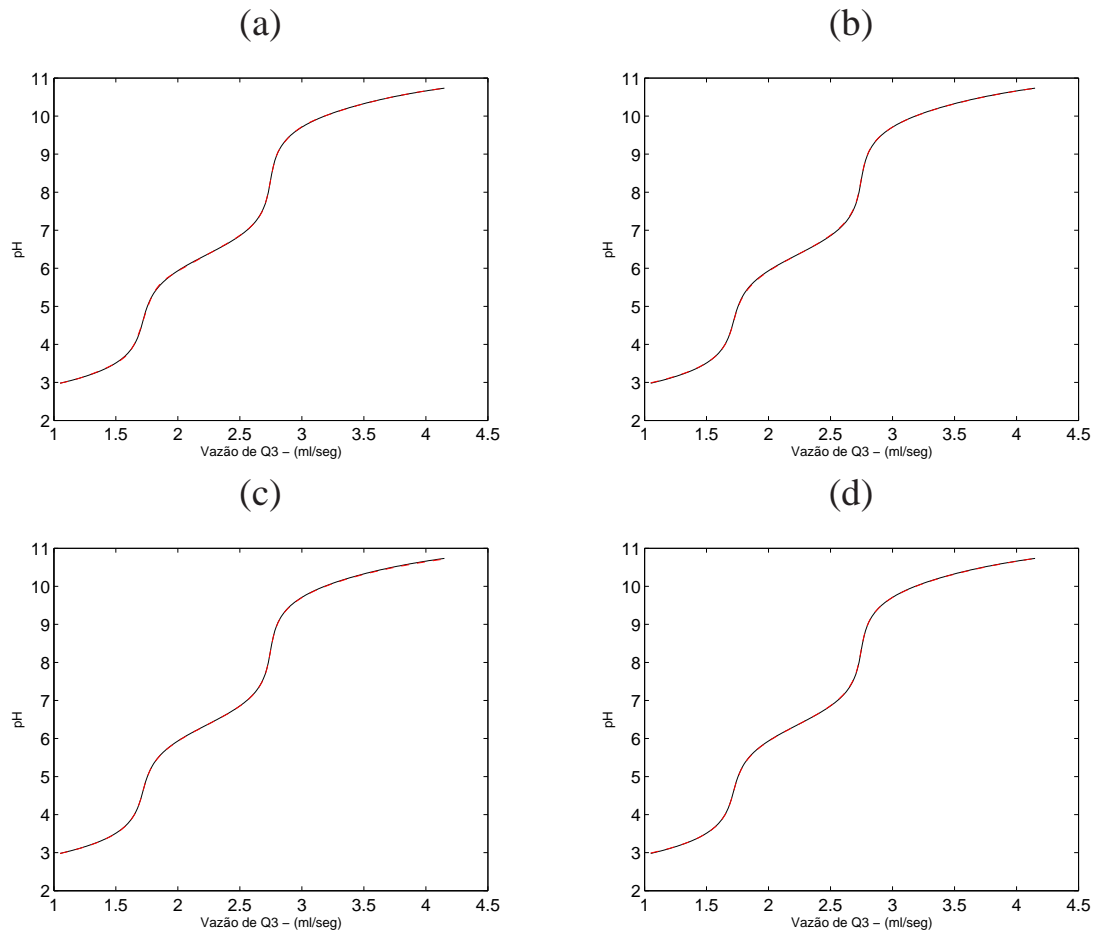


Figura 5.25: Característica estática: reprodução da característica estática da rede em linha tracejada (- -) e característica estática do sistema em linha contínua (-). (a) Rede 1. (b) Rede 2. (c) Rede 3. (d) Rede 4.

A Tabela 5.12 mostra os índices de validação para cada uma das estruturas de treinamento utilizadas.

Tabela 5.12: Índices de validação para identificação caixa cinza, com imposição de restrições, planta de neutralização de pH

Rede	Estrutura Figura 2.7	$RMSE_{\text{dinâmico}}$	$\eta_{\text{dinâmico}}$	$\eta_{\text{estático}}$
1	(a)	0,4121	0,5418	1,0000
2	(b)	0,5080	0,8673	1,0000
3	(c)	1,0087	0,2123	1,0000
4	(d)	1,0192	0,7584	1,0000

Discussão dos resultados

De acordo com os resultados obtidos para o treinamento com restrição do sistema de neutralização de pH, mostrados na Tabela 5.12 é possível verificar que a rede 1 apresentou melhores resultados, índices de validação. Pelo índice de correlação dinâmico ($\eta_{\text{dinâmico}}$) da rede 2, verifica-se que, após a identificação, não restou somente ruído branco nos dados. Por meio de uma inspeção visual da Figura 5.24(b) é possível verificar, ainda, que a rede 2 erra na aproximação dos máximos e mínimos locais

do sistema. Nas redes 3 e 4 verifica-se que, mesmo aumentando os números de neurônios e atrasos dos regressores de entrada e saída do sistema, estas estruturas não modelam o sistema.

Capítulo 6

Conclusão

No processo de identificação, utilizando redes neurais, foi possível verificar com este trabalho que a incorporação de informação auxiliar colabora com determinação de estruturas das redes neurais. Os resultados dos treinamentos caixa preta mostraram-se pouco satisfatórios principalmente para os sistemas em que a característica dinâmica descrevia uma pequena faixa de operação dos dados estáticos. Para o sistema do aquecedor de dissipação variável, que o sistema apresentava característica dinâmica em toda a faixa de operação da estática, a identificação caixa preta apresentou resultados satisfatórios.

Na identificação utilizando os modelo de Wiener e Hammerstein observou-se que os sistemas apresentaram melhores resultados para o modelo que a não linearidade estava presente em u . No caso do conversor *buck* CC-CC, em que os índices de validação dos dois modelos ficaram bem próximos. Conjectura-se que os dados possuem pouca informação dinâmica, o que dificultou a tomada de decisão entre os dois modelos testados. Um nova coleta de dados, contemplando mais pontos e excitando o sistema dentro de uma faixa ampla de operação, poderia facilitar a identificação e escolha de uma melhor estrutura para treinamento.

Nos treinamentos de redes neurais com restrição observou-se bons resultados porém, a rede completamente conectada, que é um aproximador universal, demonstrou dificuldades no ajuste dos pesos à partir das restrições. É possível que o algoritmo de treinamento utilizado não tenha sido capaz de fazer o ajuste satisfatoriamente. As tentativas de permitir um maior "grau de liberdade" à rede, aumentando números de neurônios e número de atrasos aos regressores de entrada e saída, não modificaram os resultados. Observou-se que, como obtido nos modelos de Hammerstein e Wiener, as estruturas para treinamento com restrição que apresentaram resultados dentro dos índices estabelecidos, foram aquelas que a não linearidade estava presente em u , devido às características de cada sistema.

Com a utilização da informação estática, os modelos obtidos, as estruturas que continham não linearidade em u , foram mais eficazes. Por outro lado, a imposição de restrição, principalmente naquelas onde a rede é completamente conectada, por restringir mais a variação dos pesos, os modelos obtidos não apresentaram bons índices de validação. É necessário levar em consideração que a massa de dados de dois dos sistemas estudados, conversor *buck* CC-CC e planta de neutralização de pH, não contempla toda a característica estática do sistema, o que pode ter prejudicado a obtenção dos modelos.

Para as diferentes estruturas da identificação caixa preta e caixa cinza, treinamento com restrição, aquelas que continham um neurônio linear na camada oculta da rede apresentaram melhores resultados. Supondo-se assim que, a presença de um neurônio linear na rede possa facilitar a busca por um melhor modelo.

6.1 Sugestão para trabalhos futuros

- Analisar outras características que podem funcionar como indicadores de escolha de estruturas. Como por exemplo, a análise da derivada primeira da curva estática dos sistemas;
- Verificar outros algoritmos para treinamento com restrições, de acordo com as características do sistemas;
- Fazer um levantamento de outras possibilidades para imposição de restrições e escolha de funções de ativação não lineares diferentes, de acordo com as curvas características do sistema;
- Implementar rotinas que permitam entrada de dados, escolha de estruturas, escolha do algoritmo de treinamento, escolha da imposição de restrição de maneira mais generalizada;
- Outra possibilidade, não testada neste trabalho, foram as variações de funções de ativação não lineares e técnica de poda (*pruning*) das redes neurais, a fim de otimizar o desempenho destas e diminuir o número de termos do modelo, eliminando termos desnecessários ao modelo (Aguirre, 2007; Amaral, 2001);
- Reproduzir as plantas que originaram os dados utilizados neste trabalho, a fim de coletar dados reais em uma ampla faixa de operação;
- A partir dos modelos obtidos, extrapolar os dados de treinamento e validação para além da faixa de operação do sistema.

Bibliografia

- Abonyi, J., Babuska, R., and Szeifert, F. (2001). Fuzzy Modeling with Multivariate Membership Functions: Gray-box identification and control design. *IEEE Transactions on Systems, Man, and Cybernetics*.
- Aguirre, L. (1997). Recovering map static nonlinearities from chaotic data using dynamical models. *Physica D: Nonlinear Phenomena*, 100(1-2):41 – 57.
- Aguirre, L. (2007). *Introdução à Identificação de Sistemas: técnicas lineares e não-lineares aplicadas a sistemas reais*. Ed. UFMG. 3ª edição.
- Aguirre, L., Barroso, M., Saldanha, R., and Mendes, E. (2004). Imposing steady-state performance on identified nonlinear polynomial models by means of constrained parameter estimation. *IEE Proceedings- Control Theory and Applications*, 151(2):174–179.
- Aguirre, L., Garcia, P., and Filho, R. (2000). Use of a priori information in the identification of global nonlinear models - a case study using a buck converter. *IEEE Transactions on Circuits and Systems-I: Fundamental Theory and Applications*, 47(7):1081–1085.
- Alonge, F., D’Ippolito, F., Raimondi, F., and Tumminaro, S. (2007). Nonlinear modeling of dc/dc converters using the Hammerstein’s approach. *IEEE Transactions on Power Electronics*, 22(4):1210–1221.
- Amaral, G. (2001). Uso de redes neurais e conhecimento a priori na identificação de sistemas dinâmicos não lineares. Master’s thesis, Universidade Federal de Minas Gerais.
- Arefi, M., Montazeri, A., Poshtan, J., and Jahed-Motlagh, M. (2008). Wiener-neural identification and predictive control of a more realistic plug-flow tubular reactor. *Chemical Engineering Journal*, 138:274–282.
- Bai, E. and Reyland Jr., J. (2008). Towards identification of Wiener systems with the least amount of a priori information on the nonlinearity. *Automatica*, (44):910–919.
- Barroso, M. (2001). Métodos de otimização mono-objetivo aplicadas a identificação caixa-cinza de sistemas não-lineares. Master’s thesis, Dissertação de Mestrado PPGEE. Universidade Federal de Minas Gerais.
- Barroso, M. and Nepomuceno, E. (2004). Multiobjective Identification of a Buck DC-DC Converter. *XV Congresso Brasileiro de Automática*.
- Barroso, M., Saldanha, R., and Aguirre, L. (2002). Comparação de Métodos Mono-objetivos em Identificação Caixa-cinza. *Tendências Em Matemática Aplicada e Computacional*, 3(2):43–52.

- Biagiola, S. and Figueroa, J. (2009). Wiener and Hammerstein uncertain models identification. *Mathematics and Computers in Simulation*, 79:3296–3313.
- Braga, A., Carvalho, A., and Ludermir, T. (2007). *Redes neurais artificiais: teoria e aplicações*. LTC, Rio de Janeiro. 2^a edição.
- Campos, R. (2007). Projeto e construção de planta piloto de neutralização de pH e proposta de metodologia para incorporação de informações auxiliares na identificação Narx Racional. Master's thesis, Centro Universitário do Leste de Minas Gerais - Unileste, Cel. Fabriciano.
- Cassini, C. (1999). Estimação recursiva de características estáticas não-lineares utilizando modelos polinomiais NARMAX. Master's thesis, Universidade Federal de Minas Gerais.
- Choi, J., Cho, B., VanLandingham, H., Mok, H., and Song, J. (1998). System identification of power converters based on a black-box approach. *IEEE Transactions on Circuits and Systems-I*, 45(11):1148–1158.
- Coelho, M. (2001). Modelos de Hammerstein e de Wiener: conexões com modelos NARX e sua aplicação em identificação de sistemas não-lineares. Master's thesis, Universidade Federal de Minas Gerais.
- Coelho, M., Aguirre, L., and Corrêa, M. (2002). Metodologia para Representação de Modelos NARX Polinomiais na Forma de Hammerstein e Wiener. *Tendências em Matemática Aplicada e Computacional*, (1):71–80.
- Corrêa, M. and Aguirre, L. (2004). Identificação não-linear caixa-cinza: Uma revisão e novos resultados. *Revista Controle e Automação*, 15(2):109–126.
- Crama, P. and Schoukens, J. (2004). Hammerstein - Wiener system estimator initialization. *Automatica*, 40:1543–1550.
- Cybenko, G. (1989). Approximation by superpositions of a sigmoid function. *Mathematics of Control*, 2:303–314.
- Figueroa, J., Biagiola, S., and Agamennoni, O. (2008). An approach for identification of uncertain Wiener systems. *Mathematical and Computer Modelling*, 48:305–315.
- Fun, M. and Hagan, M. (1996). Levenberg-Marquardt training for modular networks. *Proceedings of the IEEE International Conference on Neural Networks*, 1:468–473.
- Funkquist, J. (1997). Grey-Box Identification of a Continuous Digester: a distributed-parameter process. *Control Eng. Practice*, (5(7)):913–919.
- Gómez, J., Jutan, A., and Baeyens, E. (2004). Wiener model identification and predictive control of a ph neutralisation process. *IEE Proc.-Control Theory Appl.*, 151(3):329–338.
- Haykin, S. (2001, reimpressão 2008). *Redes Neurais: princípios e prática*. Bookman, 2^l edition.
- Henrique, H., Lima, E., and Seborg, D. (2000). Model structure determination in neural network models. *Chemical Engineering Science*, 55:5457–5469.
- Henson, M. and Seborg, D. (1994). Adaptive nonlinear control of a pH neutralization process. *IEEE Transactions on Control Systems Technology*, 2(3):169–182.

- Jia, L., Chiu, M., and Ge, S. (2005). A noniterative neuro-fuzzy based identification method for Hammerstein processes. *Journal of Process Control*, (15):749–761.
- Kalafatis, A., Wang, L., and Cluett, W. R. (2005). Linearizing feedforward feedback control of pH processes based on the Wiener model. *Journal of Process Control*, (15):103–112.
- Kamat, S., Diwanji, V., Smith, J., and Madhavan, K. (2005). Modeling of pH process using recurrent neural network and wavenet. *IEEE International Conference on Computational Intelligence for Measurement Systems and Applications*, pages 209–214.
- Leyva, R., Salamero, L., Jammes, B., Marpinad, J., and Guinjoan, F. (1997). Identification and control of power converters by means of neural networks. *IEEE Transactions on Circuits and Systems-I: Fundamental Theory and Applications*, 44(8):735–742.
- Lima, C., Barroso, M., Campos, R., Mendes, E., and Corrêa, M. (2008). Utilização de informação a priori e algoritmo biobjetivo para identificação de um processo de neutralização de pH em malha fechada. *Congresso Brasileiro de Automática*.
- Mahmoodi, S., Poshtan, J., Jahed-Motlagh, M., and Montazeri, A. (2009). Nonlinear model predictive control of a pH neutralization process based on Wiener-Laguerre model. *Chemical Engineering Journal*, (146):328–337.
- Margoti, L., Campos, A., Santos, A., Milagres, N., Amaral, G., and Barroso, M. (2010). Identificação de uma planta de neutralização de pH utilizando a representação de Wiener. *XVIII Congresso Brasileiro de Automática*, pages 2039–2046.
- Narendra, K. and Parthasarathy, K. (1990). Identification and control of dynamical systems using neural network. *IEEE Transactions on Neural Networks*, 1(1):4–27.
- Norgaard, M. (2000). Neural network based system identification - TOOLBOX. Technical Report 00-E-891, Technical University of Denmark.
- Norquay, S., Palazoglu, A., and Romagnoli, J. (1998). Model predictive control based on Wiener models. *Chemical Engineering Science*, 53:75–84.
- Pearson, R. and Pottman, M. (2000). Gray-box identification of block-oriented nonlinear models. *Journal of Process Control*, 10:301–315.
- Santos, A., Milagres, N., Campos, A., Margoti, L., Amaral, G., and Barroso, M. (2010). Aplicação de representações em blocos interconectados em identificação caixa-cinza de sistemas dinâmicos não lineares. *XVIII Congresso Brasileiro de Automática*, pages 4224–4230.
- Sjöberg, J., Zhang, Q., Ljung, L., Benveniste, A., Delyon, B., Glorennec, P., Hjalmarsson, H., and Juditsky, A. (1995). Nonlinear black-box modeling in system identification: a unified overview. *Automatica*, 31(12):1691–1724.
- Smith, J., Kamat, S., and Madhavan, K. (2007). Modeling of pH process using wavenet based Hammerstein model. *Journal of Process Control*, (17):551–561.
- Syafiea, S., Tadeoa, F., and Martinezb, E. (2007). Model-free learning control of neutralization processes using reinforcement learning. *Engineering Applications of Artificial Intelligence*, (20):767–782.

- Tan, W., Lu, F., Loh, A., and Tan, K. (2005). Modeling and control of a pilot pH plant using genetic algorithm. *Engineering Applications of Artificial Intelligence*, (18):485–494.
- Vale, M., Fonseca, D., Maitelli, A., and Araújo, F. (2008). Controle adaptativo por modelo de referência aplicado em uma planta de neutralização de pH. *VII Conferência Internacional de Aplicações Industriais*.
- Wang, D. and Ding, F. (2008). Extended stochastic gradient identification algorithms for Hammerstein-Wiener ARMAX systems. *Computers and Mathematics with Applications*, 56:3157–3164.
- Yoo, A., Lee, T., and Yang, D. (2004). Experimental simultaneous state and parameter identification of a pH neutralization process based on an extended kalman filter. *Korean Journal Chemical Eng.*, 21(4):753–760.