

Universidade Federal de Minas Gerais  
Centro de Pesquisa e Desenvolvimento em Engenharia Elétrica  
Programa de Pós-Graduação em Engenharia Elétrica

# **Estimação Recursiva de Características Estáticas Não Lineares Utilizando Modelos Polinomiais NARMAX**

Cecília Carabetti da Silveira Cassini

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

Orientador: Luis Antônio Aguirre

30 de abril de 1999

# Resumo

Diversos sistemas dinâmicos em Engenharia possuem características estáticas não lineares bem definidas. Tais características definem o ganho do sistema que, nesses casos, não é constante e muitas vezes varia ao longo do tempo.

O objetivo deste trabalho é a estimação recursiva das características estáticas não lineares de um sistema real utilizando modelos polinomiais NARMAX de forma a permitir quantificar mudanças temporais em tais características.

A fim de testar algumas idéias desenvolvidas neste trabalho, utilizou-se um sistema térmico simples. A característica estática não linear deste sistema pode ser facilmente obtida através de testes estáticos e pode ser utilizada para validar os resultados.

Primeiramente, foram identificados modelos com termos lineares, quadráticos e cúbicos, utilizando as seguintes técnicas para detecção de estrutura de modelos: análise de agrupamentos de termos, algoritmo baseado num critério de redução de erro e critérios de informação. Os modelos identificados foram validados com base em testes estatísticos, dinâmicos e na capacidade de recuperação da curva estática, ou seja, foram considerados aqueles modelos que eram válidos tanto dinamicamente quanto estaticamente.

A seguir, os modelos obtidos anteriormente foram comparados e alguns foram selecionados para realizar a identificação recursiva. Os parâmetros do modelo e a característica estática não linear do sistema foram estimados utilizando algumas das estruturas obtidas anteriormente. Através de um teste onde a característica estática do sistema foi forçada a mudar, foi verificado que o procedimento proposto neste trabalho foi capaz de acompanhar tal mudança recursivamente. Acredita-se que esse tipo de informação possa ser útil em problemas de controle e monitorização de sistemas dinâmicos.

# Abstract

Many real dynamical systems have well-defined nonlinear static characteristic curves. The system gain is, in general, determined by such curves and varies both with the operating point and in time.

The main objective of this work is to recursively estimate nonlinear static characteristic curves of a real system using polynomial NARMAX models. In this way it is hoped that temporal variations of such curves can be quantified.

In order to test some of the ideas developed in this work, a simple thermal system was used. The nonlinear static characteristic curves of this system can be easily obtained from static testing thus suggesting its use to validate the results.

Firstly, models with linear, quadratic and cubic terms were identified. In order to determine the model structures the following complimentary approaches were used: term cluster analysis, the Error Reduction Ratio (ERR) criterion and information criteria. The identified models were validated statistically and dynamically. Another aspect that was taken into account in choosing the best models was how well the nonlinear static characteristic was represented in each model. Ten models were described and compared using these criteria.

Secondly, a few model structures, from the models obtained in the first step, were chosen to perform the recursive identification. Hence the model parameters and respective nonlinear static characteristic curves were estimated using a recursive algorithm. A test was performed on the thermal system during which the nonlinear static characteristic curve was changed. The resulting data were used together with the estimation procedures defined previously and the results suggest that it was possible to recursively follow the changes in the nonlinear static characteristic curve. It is believed that this type of information might be useful in control problems and supervision of nonlinear dynamical systems.

# Agradecimentos

A Deus por tudo.

A meus pais pela presença constante.

Ao professor Luis Aguirre pela orientação, sabedoria, incentivo e confiança dedicada aos alunos.

Ao professor Walmir pela colaboração em empréstimos de equipamentos.

Aos colegas do grupo MACSIN (Modelagem, Análise e Controle de Sistemas Não Lineares) e do CPDEE (Centro de Pesquisa e Desenvolvimento em Engenharia Elétrica) pelo auxílio e amizade.

O apoio financeiro da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) e da Pró-Reitoria de Pesquisa (PRPq) dessa Universidade foram de vital importância para a conclusão desse trabalho.

# Conteúdo

<b>Resumo</b>	<b>i</b>
<b>Abstract</b>	<b>ii</b>
<b>Agradecimentos</b>	<b>iii</b>
<b>Nomenclatura</b>	<b>xii</b>
<b>1 Introdução</b>	<b>1</b>
1.1 Identificação de Sistemas Não Lineares . . . . .	3
1.2 Apresentação do Trabalho . . . . .	4
<b>2 Identificação de Sistemas</b>	<b>6</b>
2.1 Introdução . . . . .	6
2.2 Identificação . . . . .	7
2.3 Experimentação . . . . .	9
2.4 Detecção de Não Linearidades . . . . .	11
2.4.1 Testes no Domínio do Tempo . . . . .	11
2.4.2 Método de Correlação Cruzada Não Linear . . . . .	11
2.4.3 Método de Autocorrelação Não Linear . . . . .	12
2.5 Representações para Sistemas Não Lineares . . . . .	13
2.6 Modelo NARMAX . . . . .	13
2.7 Estimação de Parâmetros . . . . .	16
2.8 Regressores Ortogonais . . . . .	17
2.8.1 Método Clássico de Ortogonalização de Gram-Schmidt	18
2.8.2 Transformação de Householder . . . . .	20
2.9 Detecção de Estrutura . . . . .	21
2.9.1 Agrupamentos de Termos . . . . .	22
2.9.2 A Taxa de Redução de Erro . . . . .	24
2.9.3 Critérios de Informação . . . . .	26

2.9.4	Composição dos Métodos Citados para Detectar a Estrutura do Modelo . . . . .	27
2.10	Redes <i>Multi-Layer Perceptron</i> . . . . .	28
2.11	Outras Representações para Sistemas Não Lineares . . . . .	33
2.12	Validação do Modelo . . . . .	35
2.12.1	Validação Estatística . . . . .	35
2.12.2	Validação Dinâmica . . . . .	37
2.13	Algoritmo Recursivo com Fator de Esquecimento . . . . .	38
2.14	Comentários . . . . .	41
<b>3</b>	<b>Identificação em Batelada</b>	<b>43</b>
3.1	Introdução . . . . .	43
3.2	Descrição do Sistema . . . . .	43
3.3	Testes . . . . .	46
3.3.1	Testes em Degrau . . . . .	46
3.3.2	Testes Estáticos . . . . .	46
3.3.3	Testes Dinâmicos . . . . .	48
3.4	Detecção de Não Linearidades . . . . .	48
3.5	Taxa de Amostragem . . . . .	54
3.6	Análise de Agrupamentos de Termos e Coeficientes de Agrupamentos . . . . .	57
3.7	Uso de Critérios de Informação . . . . .	63
3.8	Modelos Identificados em Batelada . . . . .	64
3.8.1	Modelo Linear . . . . .	65
3.8.2	Modelo com Grau de Não Linearidade Dois . . . . .	69
3.8.3	Modelo com Grau de Não Linearidade Três . . . . .	77
3.8.4	Comparação entre os Modelos Apresentados . . . . .	84
3.9	Conclusões . . . . .	89
<b>4</b>	<b>Estimação de Não Linearidades Estáticas</b>	<b>91</b>
4.1	Introdução . . . . .	91
4.2	Pontos Fixos em Sistemas Autônomos . . . . .	92
4.3	Pontos Fixos em Sistemas Não Autônomos . . . . .	94
4.4	Estabilidade de Pontos Fixos . . . . .	95
4.5	Estimação de Não Linearidades Estáticas Utilizando Modelos Polinomiais . . . . .	97
4.5.1	Estimação de Características Estáticas a partir de Modelos Lineares . . . . .	97
4.5.2	Estimação de Características Estáticas a partir de Modelos Quadráticos . . . . .	98

4.5.3	Estimação de Características Estáticas a partir de Modelos Cúbicos . . . . .	104
4.6	Comparação entre as Características Estáticas . . . . .	111
4.7	Conclusões . . . . .	112
<b>5</b>	<b>Identificação Recursiva</b> . . . . .	<b>114</b>
5.1	Introdução . . . . .	114
5.2	Teste Dinâmico Variável . . . . .	114
5.3	Identificação Recursiva . . . . .	116
5.3.1	Modelos com Um Ponto Fixo . . . . .	116
5.3.2	Modelos com Dois Pontos Fixos . . . . .	125
5.4	Conclusões . . . . .	131
	<b>Conclusão</b> . . . . .	<b>132</b>
	<b>Bibliografia</b> . . . . .	<b>134</b>

# Lista de Figuras

i

2.1	Estrutura genérica de uma rede <i>Feedforward</i> . . . . .	28
2.2	Diagrama de blocos ilustrando a estimação recursiva. . . . .	40
3.1	Sistema . . . . .	44
3.2	Diagrama de blocos do sistema. . . . .	45
3.3	Característica estática – Teste 02. . . . .	47
3.4	Característica estática – Teste 07. . . . .	47
3.5	Característica estática – Teste 08. . . . .	48
3.6	Massa de dados $\langle din2 \rangle$ . . . . .	49
3.7	Massa de dados $\langle din3 \rangle$ . . . . .	50
3.8	Autocorrelação do sinal de entrada de $\langle din2 \rangle$ . . . . .	51
3.9	Autocorrelação do sinal de entrada de $\langle din3 \rangle$ . . . . .	51
3.10	Autocorrelação não linear do sinal de saída de $\langle din2 \rangle$ . . . . .	52
3.11	Autocorrelação não linear do sinal de saída de $\langle din3 \rangle$ . . . . .	53
3.12	Correlação cruzada não linear do sinal de saída da massa de dados $\langle din2 \rangle$ . . . . .	53
3.13	Correlação cruzada não linear do sinal de saída da massa de dados $\langle din3 \rangle$ . . . . .	54
3.14	Autocorrelação linear do sinal de saída de $\langle din2 \rangle$ utilizada para determinar a taxa de amostragem. . . . .	55
3.15	Autocorrelação não linear do sinal de saída de $\langle din2 \rangle$ utilizada para determinar a taxa de amostragem. . . . .	55
3.16	Autocorrelação linear do sinal de saída de $\langle din3 \rangle$ utilizada para determinar a taxa de amostragem. . . . .	56
3.17	Autocorrelação não linear do sinal de saída de $\langle din3 \rangle$ utilizada para determinar a taxa de amostragem. . . . .	56
3.18	Massa de dados $\langle din2 \rangle$ dizimada. . . . .	58
3.19	Massa de dados $\langle din3 \rangle$ dizimada. . . . .	59

3.20	Variações dos coeficientes de agrupamentos para $\langle din2 \rangle$ . . .	61
3.21	Variações dos coeficientes de agrupamentos para $\langle din3 \rangle$ . . .	62
3.22	Critério AIC - massa de dados $\langle din2 \rangle$ . . . . .	63
3.23	Critério AIC - massa de dados $\langle din3 \rangle$ . . . . .	64
3.24	Validação dinâmica para o modelo linear (3.6). . . . .	65
3.25	Validação estatística para o modelo linear (3.6). . . . .	66
3.26	Validação dinâmica para o modelo linear (3.7). . . . .	67
3.27	Validação estatística para o modelo linear (3.7). . . . .	68
3.28	Validação dinâmica para o modelo não linear (3.8). . . . .	69
3.29	Validação estatística para o modelo não linear (3.8). . . . .	70
3.30	Validação dinâmica para o modelo não linear (3.9). . . . .	71
3.31	Validação estatística para o modelo não linear (3.9). . . . .	72
3.32	Validação dinâmica para o modelo não linear (3.10). . . . .	73
3.33	Validação estatística para o modelo não linear (3.10). . . . .	74
3.34	Validação dinâmica para o modelo não linear (3.11). . . . .	75
3.35	Validação estatística para o modelo não linear (3.11). . . . .	76
3.36	Validação dinâmica para o modelo não linear (3.12). . . . .	78
3.37	Validação estatística para o modelo não linear (3.12). . . . .	79
3.38	Validação dinâmica para o modelo não linear (3.13). . . . .	80
3.39	Validação estatística para o modelo não linear (3.13). . . . .	81
3.40	Validação dinâmica para o modelo não linear (3.14). . . . .	82
3.41	Validação estatística para o modelo não linear (3.14). . . . .	83
3.42	Validação dinâmica para o modelo não linear (3.15). . . . .	84
3.43	Validação estatística para o modelo não linear (3.15). . . . .	85
4.1	Característica estática obtida a partir do modelo (3.6). . . . .	98
4.2	Característica estática obtida a partir do modelo (3.7). . . . .	99
4.3	Característica estática obtida a partir do modelo (3.8) em relação ao ponto fixo 1. . . . .	100
4.4	Característica estática obtida a partir do modelo (3.8) em relação ao ponto fixo 2. . . . .	100
4.5	Característica estática obtida a partir do modelo (3.9) para o ponto fixo 1. . . . .	101
4.6	Característica estática obtida a partir do modelo (3.9) para o ponto fixo 2. . . . .	101
4.7	Característica estática obtida a partir do modelo (3.10). . . . .	103
4.8	Característica estática obtida a partir do modelo (3.11). . . . .	103
4.9	Característica estática obtida a partir do modelo (3.12) para o ponto fixo 1. . . . .	105

4.10	Característica estática obtida a partir do modelo (3.12) para o ponto fixo 2. . . . .	105
4.11	Característica estática obtida a partir do modelo (3.13) para o ponto fixo 1. . . . .	106
4.12	Característica estática obtida a partir do modelo (3.13) para o ponto fixo 2. . . . .	107
4.13	Característica estática obtida a partir do modelo (3.13) para o ponto fixo 3. . . . .	107
4.14	Característica estática obtida a partir do modelo (3.14) para o ponto fixo 1. . . . .	108
4.15	Característica estática obtida a partir do modelo (3.14) para o ponto fixo 2. . . . .	109
4.16	Característica estática obtida a partir do modelo (3.15) para o ponto fixo 1. . . . .	109
4.17	Característica estática obtida a partir do modelo (3.15) para o ponto fixo 2. . . . .	110
5.1	Massa de dados $\langle din4 \rangle$ . . . . .	115
5.2	Autocorrelação do sinal de entrada de $\langle din4 \rangle$ . . . . .	116
5.3	Autocorrelação linear (2.1) do sinal de saída de $\langle din4 \rangle$ utilizada para determinar a taxa de amostragem. . . . .	117
5.4	Autocorrelação não linear (2.2) do sinal de saída de $\langle din4 \rangle$ utilizada para determinar a taxa de amostragem. . . . .	118
5.5	Variações dos coeficientes de agrupamentos de (5.1). . . . .	120
5.6	Variações dos parâmetros da estrutura (5.1). . . . .	121
5.7	Vista A - Característica estática obtida a partir da estrutura (5.1) utilizando algoritmo recursivo. . . . .	122
5.8	Vista B - Característica estática obtida a partir da estrutura (5.1) utilizando algoritmo recursivo. . . . .	123
5.9	Comparação entre a característica estática obtida a partir da estrutura (5.1) e a partir dos testes estáticos. . . . .	124
5.10	Variações dos coeficientes de agrupamentos da estrutura (5.2). . . . .	126
5.11	Variações dos parâmetros do modelo recursivo (5.2). . . . .	127
5.12	Vista A - Característica estática obtida para o ponto fixo estável, utilizando algoritmo recursivo a partir da estrutura (5.2). . . . .	128
5.13	Vista B - Característica estática obtida para o ponto fixo estável, utilizando algoritmo recursivo a partir da estrutura (5.2). . . . .	129

5.14	Comparação entre a característica estática obtida a partir da estrutura (5.2) e a partir dos testes estáticos. . . . .	130
------	--	-----

# Lista de Tabelas

i

3.1	Tabela comparativa entre os modelos apresentados em relação à ordem do modelo, atraso máximo em $u$ , grau de não linearidade, número de termos de processo, número de termos candidatos e validações estatística e dinâmica. . . . .	87
3.2	Tabela comparativa entre os modelos apresentados ilustrando os agrupamentos candidatos ao modelo e os selecionados. . .	88
4.1	Tabela comparativa entre os modelos obtidos no capítulo anterior, em relação ao número de pontos fixos e eficiência na estimação da característica estática não linear. . . . .	111
5.1	Inicialização dos parâmetros do modelo (5.1). . . . .	119
5.2	Inicialização dos parâmetros do modelo (5.2). . . . .	125

# Nomenclatura

## Siglas

AIC	<i>Akaike Information Criterion</i>
ARMAX	<i>AutoRegressive Moving Average with eXogenous inputs</i>
HNM	<i>Hybrid Neural Model</i>
NARMAX	<i>Nonlinear AutoRegressive Moving Average with eXogenous inputs</i>
NARX	<i>Nonlinear AutoRegressive with eXogenous inputs</i>
RBF	<i>Radial Basis Function</i>
RMSE	<i>Root Mean Square Error</i>
RNA	<i>Artificial Neural Network</i>

## Variáveis

$A$	Matriz triangular superior utilizada no processo de ortogonalização de Gram-Schmidt
$b_j^h$	Termo de polarização do $j$ -ésimo neurônio da camada intermediária de uma RNA <i>Feedforward</i>
$b_i^o$	Termo de polarização do $i$ -ésimo neurônio da camada de saída de uma RNA <i>Feedforward</i>
$D$	Matriz diagonal utilizada no processo de ortogonalização de Gram-Schmidt
$e(k)$	Ruído presente no sistema
$f(\cdot)$	Função não linear
$\mathbf{g}$	Vetor de parâmetros auxiliar utilizado no processo de ortogonalização de Gram-Schmidt
$\mathbf{G}_k$	Matriz de ganho para ajuste de parâmetros no algoritmo recursivo

$J$	Função custo
$l$	Grau de não linearidade de um modelo polinomial NARMAX
$M$	Número de termos candidatos a um modelo polinomial NARMAX
$m_l$	Grau de não linearidade de termos de um modelo polinomial NARMAX
$n$	Número de nodos da camada intermediária de uma RNA <i>Feedforward</i>
$N$	Número de dados experimentais
$n_e$	Atraso máximo nos termos de ruído
$n_n$	Número de termos de ruído
$n_p$	Número de termos de processo
$n_r$	Número de regressores candidatos a um modelo polinomial NARX
$N_r$	Número de entradas aplicadas ao $j$ -ésimo neurônio da camada intermediária de uma RNA <i>Feedforward</i>
$n_u$	Atraso máximo nos termos referentes à entrada do sistema
$n_y$	Atraso máximo nos termos referentes à saída do sistema
$n_\theta$	Número de parâmetros de um modelo polinomial NARMAX
$p$	Ordem do fator em $y$ no termo de grau $m_l$
$\mathbf{P}$	Matriz de regressores ortogonal
$\mathbf{P}_k$	Matriz de covariância do estimador para a iteração $k$
$\hat{\mathbf{P}}$	Matriz de regressores modificada utilizada na transformação de Householder
$\mathbf{R}$	Matriz triangular superior utilizada na transformação de Householder
$T_s$	Período de amostragem utilizado para detecção de estrutura e estimação de parâmetros
$T_{ss}$	Período de amostragem utilizado para coleta de dados experimentais
$\mathbf{u}$	Vetor composto pelos valores medidos da variável de entrada
$v_j$	Soma ponderada para o $j$ -ésimo neurônio da camada intermediária
$v_k$	Inovação na iteração $k$
$w_{ji}^h$	Peso entre $i$ -ésimo nodo da camada de entrada e $j$ -ésimo nodo da camada intermediária de uma RNA

	<i>Feedforward</i>
$w_j^o$	Peso entre nodos da camada intermediária e da camada de saída de uma RNA <i>Feedforward</i>
$w_i$	Regressor auxiliar utilizado no processo de ortogonalização de Gram-Schmidt
$x_i$	Regressores do modelo NARMAX/ entradas da RNA
$\mathbf{y}(k)$	Vetor composto pelos valores das variáveis de saída no instante $k$
$\mathbf{y}$	Vetor composto pelos valores medidos da variável de saída
$\bar{y}$	Localização de ponto fixo
$\hat{y}(k)$	Predição de um passo à frente no instante $k$
$\Theta$	Vetor de parâmetros a estimar
$\hat{\theta}_{p,m_l}(n_y, n_u)$	Parâmetro estimado, onde os subíndices indicam o tipo de termo e os argumentos os respectivos atrasos
$\xi(k)$	Resíduo no instante $k$
$\Sigma_{y^p u^{m_l-p}}$	Coefficiente do agrupamento $\Omega_{y^p u^{m_l-p}}$
$\tau_m$	Constante utilizada para determinar o período de amostragem $T_s$
$\phi$	Valor crítico utilizado para determinar o número de graus de liberdade para o critério AIC
$\phi_{y'y'}$	Função de autocorrelação linear do sinal de saída
$\phi_{y2'y2'}$	Função de autocorrelação não linear do sinal de saída
$\Omega$	Agrupamento de termos do mesmo tipo

## Operadores

$E[\cdot]$	Esperança matemática
$f^m$	Derivada de ordem $m$ da função $f$
$\sigma^2(\cdot) = \text{Var}(\cdot)$	Variância

# Capítulo 1

## Introdução

A modelagem de sistemas é uma tarefa de vital importância para o desenvolvimento da ciência e da tecnologia. Modelos de sistemas dinâmicos têm aplicações nas mais diversas áreas, tais como: engenharia, economia, medicina, ecologia e em algumas áreas sociais. A utilização de um modelo para representar um determinado sistema, pode estar relacionada com os mais diversos objetivos, tais como: compreender certas dinâmicas do processo estudado, prever o comportamento do sistema sob diversas condições de operação, analisar e projetar controladores, estimar variáveis do processo que não podem ser medidas diretamente, otimizar o comportamento do sistema, permitir detecção eficiente de falhas no sistema, permitir o estudo do sistema em regiões de operação dispendiosas ou problemáticas no sistema real, permitindo um treinamento de operação seguro e eficiente (Matko et al., 1992).

Dentre os tipos de modelos utilizados para representar sistemas físicos, pode-se citar os modelos físicos, os mentais, os gráficos e os matemáticos ou analíticos. Os modelos físicos podem ser protótipos ou modelos análogos, por exemplo, ratos são “modelos” análogos aos humanos, podendo, em algumas situações, ser utilizados para testar novos medicamentos. Estes tipos de modelos são, na maioria dos casos, obtidos a longo prazo e com custo elevado. Os modelos mentais são de carácter intuitivo, existindo, portanto, no cérebro humano. Tais modelos são imprecisos, além de dificultarem a comunicação ou passagem de informação a outras pessoas. Os modelos gráficos têm como desvantagem o fato de ser extremamente difícil obter informações do sistema através dos mesmos. Por outro lado, os modelos matemáticos são fáceis de serem manipulados e seu custo é relativamente baixo. Além de apresentar outras vantagens como concisão e ausência de ambiguidade na

interpretação. Desta forma, os modelos matemáticos são os mais comumente utilizados (Matko et al., 1992).

Um modelo matemático pode ser definido como um mapeamento das relações das variáveis físicas do sistema em uma estrutura matemática correspondente. Quando o mapeamento só considera o sistema em regime permanente, o modelo tem um carácter estático e é representado por equações algébricas. Por outro lado, quando o mapeamento inclui a resposta transitória, o modelo matemático é dito dinâmico, sendo descrito por equações diferenciais (Matko et al., 1992) ou equações de diferenças.

Existem diversas maneiras de modelar um sistema. Pode-se citar a modelagem pela física do processo e a modelagem através de técnicas de identificação de sistemas (Norton, 1986; Ljung, 1987; Söderström and Stoica, 1989). A modelagem pela física consiste na determinação de equações baseadas na física do processo, exigindo um conhecimento detalhado do mesmo, o que torna difícil algumas vezes sua aplicação. A identificação, por outro lado, determina o modelo através de dados de entrada e saída do sistema, não exigindo conhecimento prévio do mesmo (identificação caixa preta) (Sjöberg et al., 1995), embora tal conhecimento possa ser utilizado, quando disponível, para melhorar a qualidade dos modelos identificados (identificação caixa cinza) (Lindskog and Ljung, 1994).

Um aspecto importante da modelagem através da física do processo é que a mesma estrutura de modelo pode ser utilizada para sistemas similares, apenas se reestimando os parâmetros do modelo (Matko et al., 1992). A identificação, por outro lado, permite a modelagem de sistemas complexos ou até mesmo, de sistemas não claramente definidos (Norton, 1986).

A fim de conseguir um modelo que melhor atenda aos objetivos iniciais, os dois tipos de modelagem podem ser realizados em conjunto, ou seja, pode-se utilizar conhecimento *a priori* do processo para auxiliar na determinação do modelo, caracterizando um procedimento de identificação em caixa cinza. Há vários trabalhos na literatura utilizando tal procedimento, podendo ser mencionados alguns. Em (Cubillos et al., 1996) utilizou-se uma estrutura de RNA (Rede Neural Artificial) acoplada às equações de balanço de transferência de massa e energia para identificar um modelo para o sistema. O desempenho obtido com a identificação em caixa cinza, utilizando uma estrutura híbrida foi comparado com o obtido através da identificação em caixa preta, utilizando somente a RNA. Os resultados obtidos com o modelo híbrido foram superiores aos obtidos com a RNA em caixa preta. Em (Jácome, 1996) utilizou-se conhecimento *a priori* da forma das não linearidades estáticas de interesse no sistema, para auxiliar no processo de detecção de estrutura de modelos polino-

miais NARMAX (*Nonlinear AutoRegressive Moving Average with eXogenous inputs*) (Chen and Billings, 1989), obtendo desta forma, modelos mais concisos e representativos do sistema considerado. De forma similar, (Hernández and Arkun, 1993) também utilizou a característica estática não linear como forma de auxiliar na validação dos modelos do sistema.

Outra maneira de combinar técnicas de identificação com conhecimento *a priori* do sistema, consiste em determinar uma estrutura de modelo através de leis físicas que regem o mesmo e, posteriormente, estimar os seus parâmetros (Cubillos et al., 1996; Cubillos and Lima, 1997).

## 1.1 Identificação de Sistemas Não Lineares

A maior parte dos sistemas reais são não lineares por natureza. Se a faixa de variações das variáveis do sistema não for ampla, então o sistema pode ser linearizado dentro de uma faixa de variação relativamente pequena das variáveis (Ogata, 1993), ou seja, em alguns casos, é possível representar sistemas não lineares através de modelos lineares por partes. No entanto, em geral a utilização de modelos não lineares é mais adequada, já que tais modelos incorporam as não linearidades existentes no sistema. Dentre as representações não lineares mais comumente utilizadas, pode-se citar modelos híbridos (Haber and Unbehauen, 1990), RNAs e RBFs (*Radial Basis Function*) (Haykin, 1994; Braga et al., 1998), HNM (*Hybrid Neural Model*) (Cubillos and Lima, 1997), modelos polinomiais contínuos (Gouesbet and Letellier, 1994) e modelos NARMAX (Chen and Billings, 1989).

Muitos sistemas reais são variantes no tempo. Nestes casos, quando deseja-se realizar um controle das variáveis do processo, é recomendável fazê-lo *on-line* (Åström and Wittenmark, 1995). Além disso, alguns sistemas, como por exemplo, os envolvendo válvulas, apresentam falhas constantes que não são facilmente detectadas, exigindo a necessidade de serem realizados testes estáticos específicos. No entanto, tais testes demandam tempo, além de nem sempre poderem ser efetuados em sistemas reais (Gruhn et al., 1998). Sendo assim, dentre as inúmeras aplicações da identificação *on-line*, realizada através da utilização de algoritmos recursivos, pode-se citar as seguintes: detecção de falhas em equipamentos e controle adaptativo. No primeiro caso, um controle do tipo MBC (*Model Based Control*) pode ser implementado, desde que se tenha em mãos uma estrutura de modelo capaz de representar a dinâmica do sistema em questão (Ljung, 1994). No segundo caso, por sua vez, pode-se utilizar uma estrutura de modelo pré-definida, estimando-se os parâmetros do modelo recursivamente. Pode-se a partir daí, utilizar

conceitos de pontos fixos e agrupamentos de termos para estimar as características estáticas com base nos modelos, sem a necessidade da realização de novos testes na planta.

Além de aplicações como em processos de detecção de falhas, o conhecimento da característica estática não linear pode ser, em princípio, utilizado no próprio controle ou para análise do comportamento do sistema (Hernández and Arkun, 1993; Pröll and Karim, 1994).

## 1.2 Apresentação do Trabalho

Este trabalho tem como objetivo estudar a estimação da característica estática não linear de um sistema, que sofre variação num determinado tempo. Para alcançar isso, pretende-se utilizar modelos capazes de representar a dinâmica do sistema e algoritmos recursivos.

Como o sistema exibe dinâmica não linear, é recomendável a utilização de um tipo de representação para sistemas não lineares. Além disso, como o objetivo principal do trabalho é a estimação da característica estática não linear, bem como suas alterações ao longo do tempo, é desejável utilizar um tipo de representação a partir da qual seja possível extrair tais informações. Conforme será mencionado no capítulo 2 e confirmado nos capítulos 4 e 5, os modelos polinomiais NARMAX apresentaram-se bastante satisfatórios.

O capítulo 2 proporciona ao leitor uma visão geral do processo de identificação que será utilizado neste trabalho. São descritos os procedimentos de identificação em batelada e identificação recursiva. Na identificação em batelada são descritas as seguintes fases: experimentação, detecção de não linearidades, detecção de estrutura, estimação de parâmetros e validações dinâmica e estatística dos modelos apresentados. A identificação recursiva engloba a fase de experimentação e estimação de parâmetros. Neste trabalho, não foram realizadas alterações na estrutura do modelo recursivamente.

O capítulo 3 apresenta o sistema utilizado, os resultados obtidos nos testes estáticos e dinâmicos, o procedimento de detecção de estrutura, estimação de parâmetros e validação estatística e dinâmica de alguns modelos lineares, quadráticos e cúbicos, de acordo com o definido no capítulo anterior para a identificação em batelada. Uma preocupação neste capítulo foi a de verificar como certos agrupamentos de termos afetam a característica estática. Além disso, desejava-se determinar uma estrutura que permitisse recuperar a característica estática *de forma analítica*.

No capítulo 4 é realizada a estimação da característica estática não linear do sistema, tendo em vista os modelos obtidos no capítulo 3 e utilizando

conceitos de pontos fixos e agrupamentos de termos. Neste capítulo as características estimadas são comparadas às obtidas através dos testes estáticos, com o intuito de validar os modelos do ponto de vista estático.

No capítulo 5, utilizando algumas das estruturas de modelo definidas no capítulo 3, realizou-se a estimação recursiva dos parâmetros dos modelos, bem como da característica estática não linear. Para isso, foi realizado um teste dinâmico no sistema a ser identificado. Durante o teste, foram alteradas as condições de operação do sistema de tal maneira a produzir variações na característica estática não linear. Usando-se o procedimento descrito neste trabalho, foi possível estimar recursivamente a referida característica não linear. Acredita-se que a capacidade de monitorar em tempo real características estáticas de sistemas e *ser capaz de quantificar tais características algebricamente*, pode ser importante na monitorização e controle de processos reais.

O capítulo seguinte apresenta as conclusões deste trabalho, juntamente com algumas propostas de continuação.

## Capítulo 2

# Identificação de Sistemas

### 2.1 Introdução

Na interação de um indivíduo e um determinado sistema é necessário o conhecimento das variáveis do sistema, bem como da maneira com a qual as mesmas se relacionam, ou seja, é necessária a obtenção de um modelo para representar o sistema em questão (Ljung, 1987).

Existem diversos tipos de modelos. Pode-se citar os modelos físicos, os mentais, os gráficos e os matemáticos ou analíticos. A escolha do tipo de modelo para representar o sistema dependerá das características estáticas e/ou dinâmicas do sistema, bem como do propósito para qual o modelo está sendo determinado. Os modelos matemáticos, no entanto, permitem aplicações mais avançadas, sendo os mais utilizados, seja na engenharia, biologia, medicina, economia ou outras áreas (Ljung, 1987).

A modelagem matemática de sistemas dinâmicos é uma tarefa de vital importância para o desenvolvimento da ciência e da tecnologia. Na engenharia, os modelos matemáticos permitem analisar e prever o comportamento de um sistema sob diversas condições de operação, e ajustar o desempenho do mesmo, caso os resultados obtidos não sejam satisfatórios. Além disto, com base no modelo pode-se estudar certas características estáticas de elementos de interesse na planta, como saturações, histereses e ganhos variáveis.

É importante salientar que um modelo matemático para um dado sistema não é único (Ogata, 1993). Deve-se observar, em primeiro lugar, se o modelo escolhido é capaz de explicar os dados medidos satisfatoriamente, podendo ser usado para explicar e compreender os fenômenos observados no sistema real. Em segundo lugar, é fundamental que o modelo seja capaz de atender ao objetivo final para o qual foi obtido. Em alguns casos, o objetivo principal

da modelagem de um sistema físico é auxiliar ou tornar mais eficiente o projeto de controladores ou acionadores. Em outros casos, o próprio modelo é o objetivo principal, para por exemplo, estudar e analisar um determinado processo, ou para prever dados futuros.

Existem diversas maneiras de modelar um sistema, tais como a modelagem pela física do processo e a modelagem através de técnicas de identificação (Ljung, 1987; Norton, 1986; Söderström and Stoica, 1989). A modelagem pela física do processo é uma abordagem analítica realizada através de conhecimento de leis físicas do processo, exigindo um conhecimento detalhado do mesmo, o que torna, em alguns casos, difícil a sua aplicação. A identificação, por outro lado, é uma abordagem experimental, que consiste no desenvolvimento de técnicas para a estimação de modelos de sistemas dinâmicos através de dados experimentais, não exigindo conhecimento prévio do processo, embora tal conhecimento possa ser utilizado se disponível.

Uma comparação pode ser feita entre a modelagem através de leis físicas do processo e a identificação de sistemas. Uma vantagem da modelagem matemática é que ela permite visualizar características intrínsecas ao sistema, sendo que, na maioria dos casos, todos os parâmetros do modelo têm significado físico. A desvantagem deste tipo de modelagem é que a determinação do modelo pode ser uma tarefa bastante difícil ou viável somente a longo prazo. Em casos em que o processo a ser modelado é muito complexo, a obtenção de modelos usando apenas conhecimento físico do mesmo é praticamente impossível. Nestes casos, é mais apropriado utilizar técnicas de identificação de sistemas. No entanto, a modelagem baseada na física do processo pode ser combinada com métodos de identificação (Åström and Wittenmark, 1990), para auxiliar na detecção da estrutura ou apenas para a determinação de coeficientes de um modelo, proporcionando desta forma, uma modelagem mais ágil e eficiente. Uma situação muito comum na prática é quando se tem a estrutura do modelo, mas os parâmetros do mesmo são variáveis, ou quando alguns deles não são conhecidos, devendo ser estimados.

## 2.2 Identificação

A identificação pode ser paramétrica ou não paramétrica. A identificação paramétrica é caracterizada por um vetor de parâmetros  $\Theta$  e o modelo correspondente por  $M(\Theta)$ . A não paramétrica é aquela realizada através de funções gráficas, tais como os métodos de resposta ao impulso, resposta ao degrau e diagramas de resposta em frequência.

A identificação não paramétrica é, em alguns casos, mais fácil de ser re-

alizada, mas é menos precisa (Söderström and Stoica, 1989). O que se faz, normalmente, é utilizar a mesma para determinar algumas características do sistema, tais como tempo morto, constante de tempo e ganho e, posteriormente, utilizar um método paramétrico para determinar o modelo do sistema.

A identificação de sistemas é um processo que exige o envolvimento do modelador (Ljung, 1991). O modelador deve analisar quais variáveis do sistema são relevantes para a modelagem e se a estrutura do modelo escolhida é adequada (Söderström and Stoica, 1989), caso contrário, o mesmo deve tomar as decisões necessárias para corrigir o problema.

Pode-se citar os seguintes passos no processo de identificação de sistemas:

1. Experimentação;
2. Detecção de não linearidades no sistema;
3. Determinação da estrutura do modelo;
4. Estimação dos parâmetros do modelo;
5. Validação do modelo encontrado.

O primeiro passo consiste na determinação da taxa de amostragem, no projeto de sinais de excitação e na coleta de dados experimentais do sistema. O segundo passo consiste na verificação de prováveis não linearidades no sistema. O terceiro passo consiste na determinação da estrutura do modelo e o seguinte, no uso de um algoritmo para estimação dos parâmetros do mesmo. Estas duas últimas fases podem ser realizadas separadamente ou em conjunto. E, finalmente, utilização de procedimentos que permitam verificar se o modelo escolhido é uma representação válida do sistema. Em caso contrário, retorna-se ao primeiro passo e realiza-se uma modificação na estrutura.

A identificação de sistemas é de extrema importância no projeto de controladores, uma vez que todas as técnicas para análise e projeto dos mesmos são baseadas na disponibilidade de modelos apropriados do processo dinâmico. Assim, o desenvolvimento do controle digital tem sido acompanhado por um desenvolvimento análogo de métodos de identificação de sistemas (Åström and Wittenmark, 1990). Além disso, no caso de controle adaptativo, onde os parâmetros devem ser estimados *on-line*, um método de identificação pode ser usado para encontrar o modelo, bem como estimar os parâmetros do mesmo. Uma vez que há diversas aplicações possíveis para um modelo, a sua identificação deve ser feita sempre com o objetivo final do mesmo em mente (Stoica et al., 1986).

## 2.3 Experimentação

A experimentação é caracterizada pela determinação do sinal de entrada, intervalo de amostragem e coleta de dados para estimação dos parâmetros do modelo (Söderström and Stoica, 1989). Em primeira instância determina-se o sinal de entrada e o intervalo de amostragem, e em seguida executam-se os testes para coleta dos dados.

Esta fase é de extrema importância para a identificação de um sistema, uma vez que os resultados posteriores dependerão da qualidade e montante de informação contida nos dados coletados (Billings and Fadzil, 1985).

Dados coletados durante o processo de operação normal podem ser usados para a identificação do sistema. No entanto, a presença de malhas de realimentação pode levar a problemas. Portanto, é preferível, sempre que possível, excitar o processo com sinais externos. Tais sinais devem ser capazes de excitar toda a dinâmica do sistema durante o processo de identificação. Por exemplo, para um sistema não linear, isto significa excitar as não linearidades de interesse no sistema (Billings and Fadzil, 1985).

Quando for utilizado o algoritmo original de mínimos quadrados ortogonal, a identificação de modelos não lineares deve ser realizada com os dados experimentais originais, ou seja, sem remover a média dos sinais de entrada e saída. Caso contrário, o modelo se tornará dependente da entrada e seus respectivos parâmetros se tornarão função das características estatísticas da mesma, ou seja, tal modelo será válido somente para sinais com características semelhantes (Billings and Fadzil, 1985). É importante salientar que tal problema pode ser contornado se realizada uma normalização dos sinais, conforme abordado em (Mendes and Billings, 1996).

A escolha do período de amostragem deve ser realizada com cautela, pois afetará diretamente a determinação da estrutura do modelo, estimação de parâmetros e, por consequência, a qualidade do modelo final. Uma taxa de amostragem lenta tende a tornar os dados adjacentes não correlacionados, o que impossibilita a determinação de um modelo para o processo. Por outro lado, se os dados forem amostrados com taxa muito alta, estes tendem a ser altamente correlacionados, causando problemas como mal condicionamento da matrix de regressores e dificuldade na determinação da estrutura do modelo (Billings and Aguirre, 1995).

Em vista disso, são realizados testes em degrau para estimar a constante de tempo predominante do sistema, e a partir daí, determina-se um período de amostragem  $T_{ss}$  para a coleta de dados experimentais. No entanto, como os dados coletados com esta taxa de amostragem são, normalmente, superamostrados, deve-se definir um novo período de amostragem  $T_s$  que será

utilizado para a identificação.

Um dos procedimentos utilizados para auxiliar na escolha da taxa de amostragem consiste em determinar, em primeira instância, as funções de autocorrelação linear  $\phi_{y'y'}$  (2.1) e não linear  $\phi_{y^2'y^2'}$  (2.2) da saída do sistema:

$$\phi_{y'y'}(\tau) = E[(y(k) - \overline{y(k)})(y(k - \tau) - \overline{y(k)})], \quad (2.1)$$

$$\phi_{y^2'y^2'}(\tau) = E[(y^2(k) - \overline{y^2(k)})(y^2(k - \tau) - \overline{y^2(k)})], \quad (2.2)$$

onde  $\overline{y(k)}$  e  $\overline{y^2(k)}$  representam os valores médios e o apóstrofe neste caso indica média removida. Com base nas funções de correlação acima, pode-se definir a seguinte constante:

$$\tau_m = \min\{\tau_{y'}, \tau_{y^2'}\}, \quad (2.3)$$

onde  $\tau_{y'}$  é o instante do primeiro mínimo de  $\phi_{y'y'}$  e  $\tau_{y^2'}$  é o instante do primeiro mínimo de  $\phi_{y^2'y^2'}$ .

Desta forma, o período de amostragem pode ser determinado através da seguinte relação (Aguirre, 1995):

$$\frac{\tau_m}{20} \leq T_s \leq \frac{\tau_m}{10}, \quad (2.4)$$

em alguns casos o limite superior pode ser  $\tau_m/5$  e o inferior  $\tau_m/25$ . Desta forma, se  $T_{ss}$  estiver dentro do intervalo proposto pela equação (2.4), então os dados experimentais originais podem ser utilizados. Se, no entanto,  $T_{ss}$  for menor que o limite inferior sugerido, os dados devem ser dizimados a fim de satisfazer (2.4). Se  $T_{ss}$  for maior que o limite superior um novo teste deve ser realizado.

Em grande parte dos sistemas o primeiro mínimo de  $\phi_{y^2'y^2'}$  ocorre em valores menores de  $\tau$  do que o primeiro mínimo de  $\phi_{y'y'}$ , ou seja, um período de amostragem menor é requerido se utilizada a Equação (2.4) do que se utilizada somente a autocorrelação linear da saída (Aguirre, 1994a). Colocando de outra forma, se os dados forem amostrados levando em conta apenas a função de autocorrelação linear, certas dinâmicas não lineares podem ser subamostradas e desta forma, não estarão contidas nos dados experimentais.

## 2.4 Detecção de Não Linearidades

Nos primeiros estágios da identificação de um processo é importante determinar, com base nos dados, se o mesmo exhibe características não lineares. Em caso afirmativo, um modelo não linear deverá ser usado para representar o sistema (Billings and Fadzil, 1985; Billings and Voon, 1983; Haber, 1985).

Os testes para detecção de não linearidades devem ser independentes do tipo de representação, bem como da ordem do sistema, uma vez que tais testes serão realizados anteriormente à detecção de estrutura, exigindo a necessidade de testes não paramétricos. Existem diversos testes para detectar não linearidades no sistema, tais como: testes no domínio do tempo, testes em regime permanente, testes do valor médio da saída, testes no domínio da frequência e testes de correlação (Haber, 1985).

Os fatores que influenciam na escolha do teste são simplicidade e eficiência do mesmo, a possibilidade da planta sair de sua operação normal, e se o modelo é contínuo ou discreto. Com base nestes fatores, os testes de correlação e os testes no domínio do tempo são mais eficientes se comparados aos demais.

### 2.4.1 Testes no Domínio do Tempo

Tais testes consistem em excitar o sistema com um sinal de entrada  $u_1(k)$ , obtendo na saída o sinal  $y_1(k)$  correspondente. Repetir o procedimento para um sinal de entrada  $u_2(k) = \gamma u_1(k)$ , onde  $\gamma$  é uma constante, e obtendo na saída  $y_2(k)$ .

Medir a saída  $y_0$  ou termo constante, para uma entrada nula. Subtrair este valor de  $y_1(k)$  e  $y_2(k)$  e encontrar a seguinte razão:

$$\delta(k) = \frac{y_2(k) - y_0}{y_1(k) - y_0} \quad (2.5)$$

O sistema será linear se  $\delta(k)$  for constante e igual  $\gamma$  (Haber, 1985).

### 2.4.2 Método de Correlação Cruzada Não Linear

Como os sistemas reais usualmente contêm um termo constante, a fim de eliminar o seu efeito, os testes de correlação são aplicados à entrada e saída normalizadas:

$$u'(k) = [u(k) - \overline{u(k)}] / \sigma^2(u(k)) \quad (2.6)$$

$$y'(k) = [y(k) - \overline{y(k)}] / \sigma^2(y(k)) , \quad (2.7)$$

onde  $\sigma^2(u(k))$  e  $\sigma^2(y(k))$ , são as variâncias da entrada e saída do sistema, respectivamente, e  $\overline{u(k)}$  e  $\overline{y(k)}$  correspondem aos respectivos valores médios.

Pode-se também definir a seguinte variável:

$$x(k) = [u'^2(k) - \overline{u'^2(k)}] / \sigma^2(u'^2(k)) . \quad (2.8)$$

Seja a seguinte função de correlação (Haber, 1985):

$$\phi_{xy'}(\tau) = E[x(t + \tau)y'(t)] \quad (2.9)$$

donde tem-se que:

$$\phi_{xy'}(\tau) = 0 \quad \forall \tau , \quad (2.10)$$

se o sistema for linear.

### 2.4.3 Método de Autocorrelação Não Linear

Seja a seguinte função de correlação (Haber, 1985):

$$\phi_{vy'}(\tau) = E[v(t + \tau)y'(t)] , \quad (2.11)$$

onde  $v = [y'^2 - \overline{y'^2(k)}] / \sigma(y'^2(k))$ . Se  $\phi_{vy'}(\tau)$  é igual a zero  $\forall \tau$ , então o sistema é linear.

É importante salientar que os testes de correlação, representados pelas equações (2.9 e 2.11), podem ser realizados utilizando-se a mesma massa de dados de identificação, ao contrário do teste no domínio do tempo, que exige a execução de um outro teste, embora o mesmo seja relativamente simples.

Por outro lado, os testes de correlação para detecção de não linearidades só serão válidos quando o sistema for excitado por um sinal aleatório ou pseudo-aleatório, caso contrário, é impossível determinar se os resultados encontrados são provenientes de correlação entre ruído presente nos dados, ou se é decorrente de não linearidades presentes na dinâmica do sistema (Billings and Fadzil, 1985; Haber, 1985).

Em (Billings and Fadzil, 1985; Billings and Voon, 1983; Haber, 1985) podem ser encontradas descrições mais detalhadas dos testes de detecção de não linearidades.

## 2.5 Representações para Sistemas Não Lineares

Uma vez determinada a presença de não linearidades no sistema, escolhe-se uma representação matemática. Embora, em alguns casos, seja possível representar sistemas com modelos lineares por partes, em torno de pontos de operação pré-definidos, em geral, um modelo não linear é mais adequado, uma vez que incorpora as não linearidades intrínsecas ao sistema original.

Existem diversas representações para modelagem de sistemas não lineares, dentre elas, pode-se citar as séries de Volterra e Wiener, os modelos por blocos orientados, polinomiais, bilineares, semi-lineares com parâmetros variáveis, lineares por partes (Billings, 1980; Chen and Billings, 1989; Haber and Keviczky, 1985; Haber and Unbehauen, 1990) e redes neurais artificiais (Braga et al., 1998; Haykin, 1994).

Como cada algoritmo é melhor para representar um tipo de sistema de acordo com o tipo de não linearidade presente nos dados (Çinar, 1995), não se pode dizer, de imediato, que uma das abordagens seja mais adequada, mesmo porque, para algumas representações, não existe uma forma sistemática para determinar a estrutura “ideal” de um dado sistema (Ronco and Gawthrop, 1997). A escolha do tipo de modelo também depende da existência ou não de informação *a priori*, do propósito para o qual o modelo está sendo utilizado e da complexidade do sistema (Billings, 1980).

## 2.6 Modelo NARMAX

Para sistemas lineares discretos utilizam-se modelos de equações lineares a diferenças, que relacionam variáveis de entrada e saída do sistema, pode-se citar os modelos ARX (*AutoRegressive with eXogenous inputs*), ARMAX (*AutoRegressive Moving Average with eXogenous inputs*), dentre outros. Uma descrição análoga para sistemas não lineares discretos são os modelos NARMAX (*Nonlinear AutoRegressive Moving Average with eXogenous inputs*), que realizam um mapeamento não linear das entradas e saídas passadas para a saída atual (Sjöberg et al., 1995). Tais modelos são capazes de representar uma ampla classe de sistemas não lineares e ao contrário do que ocorre com as Séries de Volterra e Wiener (Billings, 1980; Haber and Keviczky, 1985), evitam a necessidade de grande número de parâmetros (Chen and Billings, 1989).

Seja o seguinte modelo geral multivariável:

$$\begin{aligned} \mathbf{y}(k) = & f(\mathbf{y}(k-1), \dots, \mathbf{y}(k-n_y), \mathbf{u}(k-1), \dots, \mathbf{u}(k-n_u), \\ & \mathbf{e}(k-1), \dots, \mathbf{e}(k-n_e)) + \mathbf{e}(k) \end{aligned} \quad (2.12)$$

onde

$$\mathbf{y}(k) = \begin{bmatrix} y_1(k) \\ y_2(k) \\ \vdots \\ y_m(k) \end{bmatrix}, \mathbf{u}(k) = \begin{bmatrix} u_1(k) \\ u_2(k) \\ \vdots \\ u_r(k) \end{bmatrix}, \mathbf{e}(k) = \begin{bmatrix} e_1(k) \\ e_2(k) \\ \vdots \\ e_m(k) \end{bmatrix} \quad (2.13)$$

são os vetores de saída, entrada e ruído, respectivamente,  $m$  é o número de saídas,  $r$  o número de entradas,  $n_y$ ,  $n_u$  e  $n_e$  são os atrasos máximos para a saída, entrada e ruído, respectivamente,  $e(k)$  é ruído branco e  $f(\cdot)$  é qualquer função vetorial não linear. A representação (2.12) é referida como um modelo entrada-saída multivariável NARMAX.

A equação (2.12) pode ser decomposta em  $m$  equações escalares, como se segue:

$$\begin{aligned} y_i(k) = & f_i(y_1(k-1), \dots, y_1(k-n_y), \dots, y_m(k-1), \dots, y_m(k-n_y), \\ & u_1(k-1), \dots, u_1(k-n_u), \dots, u_r(k-1), \dots, u_r(k-n_u), \\ & e_1(k-1), \dots, e_1(k-n_e), \dots, e_m(k-1), \dots, e_m(k-n_e)) + \\ & e_i(k), \quad i = 1, \dots, m. \end{aligned} \quad (2.14)$$

A fim de melhorar a flexibilidade da estrutura do modelo os atrasos máximos para cada entrada, saída e ruído podem ser diferentes para cada subsistema:

$$\begin{aligned} y_i(k) = & f_i(y_1(k-1), \dots, y_1(k-n_{y1}^i), \dots, y_m(k-1), \dots, y_m(k-n_{ym}^i), \\ & u_1(k-1), \dots, u_1(k-n_{u1}^i), \dots, u_r(k-1), \dots, u_r(k-n_{ur}^i), \\ & e_1(k-1), \dots, e_1(k-n_{e1}^i), \dots, e_m(k-1), \dots, e_m(k-n_{em}^i)) + \\ & e_i(k), \quad i = 1, \dots, m. \end{aligned} \quad (2.15)$$

A utilização da representação (2.15) para a identificação de um sistema, exige que a mesma seja parametrizada. Expansões polinomiais de  $f_i(\cdot)$  são normalmente satisfatórias, embora não sejam a única escolha.

Expandindo  $f(\cdot)$  como um polinômio de grau  $l$  (Chen et al., 1989), tem-se:

$$\begin{aligned} y_i(k) = & \theta_0^i + \sum_{i_1=1}^{M_r} \theta_{i_1}^i x_{i_1}(k) + \sum_{i_1=1}^{M_r} \sum_{i_2=i_1}^{M_r} \theta_{i_1 i_2}^i x_{i_1}(k) x_{i_2}(k) + \dots \\ & + \sum_{i_1=1}^{M_r} \dots \sum_{i_l=i_{l-1}}^{M_r} \theta_{i_1 \dots i_l}^i x_{i_1}(k) \dots x_{i_l}(k) + e_i(k), \quad i = 1, \dots, m, \end{aligned} \quad (2.16)$$

onde  $\theta_i$  são os parâmetros,  $x_i$  os monômios<sup>1</sup>,  $y_i$  a saída,  $e_i$  é ruído branco e  $M_r$  é igual a:

$$M_r = m \times (n_y + n_e) + r \times n_u . \quad (2.17)$$

O número de termos candidatos ao modelo polinomial NARMAX é dado por:

$$\left. \begin{aligned} M &= \sum_{j=0}^l n_{ij}, & n_{i0} &= 1 \\ n_{ij} &= \frac{n_{ij-1} \left[ \sum_{k=1}^m n_{yk}^i + n_{ek}^i + \sum_{k=1}^r n_{uk}^{i+j-1} \right]}{j}, & j &= 1, \dots, l \end{aligned} \right\} \quad (2.18)$$

Percebe-se que o número de termos candidatos cresce rapidamente com o número de atrasos, grau de não linearidade e com o número de entradas e saídas do modelo. Sendo assim, se todos os regressores forem incluídos no modelo, a estimação dos parâmetros da expansão polinomial (2.16), implicará, geralmente, em mal condicionamento da matriz de regressores. Desta forma, anteriormente ou paralelamente à estimação de parâmetros, deve-se utilizar um método para determinar quais termos, dentre os candidatos, devem ser incluídos no modelo (Billings and Fadzil, 1985). Tal procedimento é denominado detecção de estrutura.

Pode-se considerar ainda, o caso particular de um modelo polinomial NARMAX que não tenha modelo de ruído:

$$\begin{aligned} y_i(k) &= f_i(y_1(k-1), \dots, y_1(k-n_y), \dots, y_m(k-1), \dots, y_m(k-n_y) \\ &\quad u_1(k-1), \dots, u_1(k-n_u), \dots, u_m(k-1), \dots, u_m(k-n_u)) \\ &\quad + e_i(k), \quad i = 1, \dots, m . \end{aligned} \quad (2.19)$$

Este modelo é conhecido como modelo NARX (*Nonlinear AutoRegressive with exogenous inputs*). O número de monômios, a partir dos quais os regressores candidatos serão formados, é igual a:

$$n_r = m \times n_y + r \times n_u . \quad (2.20)$$

Tais monômios são funções das entradas e saídas do sistema, bem como de seus respectivos atrasos:

---

<sup>1</sup>As diversas combinações dos monômios, às vezes chamadas de multinômios, constituem regressores do modelo.

$$\left. \begin{aligned} x_1(k) &= y_1(k-1), \dots, x_{m \times n_y}(k) = y_m(k-n_y) \\ x_{m \times n_y+1}(k) &= u_1(k-1), \dots, x_{n_r}(k) = u_r(k-n_u) \end{aligned} \right\} \quad (2.21)$$

Como os modelos do processo e do ruído podem, na prática, ser desacoplados (Billings et al., 1989), primeiro modela-se o processo utilizando o modelo NARX e posteriormente, adicionam-se termos de ruído, caso os testes de validação indiquem que os resíduos não são brancos.

## 2.7 Estimação de Parâmetros

A estimação de parâmetros pode ser vista como um problema de otimização, onde o melhor modelo é aquele que minimiza uma determinada função custo.

A Equação (2.16) pode ser representada pela seguinte equação de regressão linear:

$$y(k) = \sum_{i=1}^M p_i(k)\theta_i + e(k), \quad k = 1, \dots, N \quad (2.22)$$

onde  $y(k)$  é a saída do sistema,  $p_i(k)$  são os regressores,  $\theta_i$  são os parâmetros a serem estimados e  $e(k)$  representa o ruído presente nos dados, dinâmica não modelada, etc.

Expandindo a Equação (2.22) para  $N$  situações, tem-se:

$$\begin{aligned} y(1) &= \mathbf{p}^T(1)\boldsymbol{\Theta} + e(1) \\ y(2) &= \mathbf{p}^T(2)\boldsymbol{\Theta} + e(2) \\ &\vdots \\ y(N) &= \mathbf{p}^T(N)\boldsymbol{\Theta} + e(N), \end{aligned}$$

onde

$$\boldsymbol{\Theta} = [\theta_1 \theta_2 \dots \theta_M]^T$$

e

$$\mathbf{p}^T(k) = [p_1(k) p_2(k) \dots p_M(k)] .$$

Desprezando-se o termo de ruído, o sistema de equações pode ser representado na forma matricial:

$$\mathbf{y} = \mathbf{P} \boldsymbol{\Theta} , \quad (2.23)$$

onde

$$\mathbf{y} = \begin{bmatrix} y(1) \\ \vdots \\ y(N) \end{bmatrix}, \quad \mathbf{P} = \begin{bmatrix} \mathbf{p}^T(1) \\ \vdots \\ \mathbf{p}^T(N) \end{bmatrix} \quad (2.24)$$

A identificação permite encontrar um modelo que melhor represente a dinâmica do sistema, de acordo com uma determinada função custo. No entanto, sempre ocorre uma diferença entre a predição de um passo à frente e o valor medido. Este erro é denominado resíduo e pode ser expresso por:

$$\xi(k) = y(k) - \mathbf{p}^T \hat{\Theta} = y(k) - y(\hat{k}) . \quad (2.25)$$

Como na maioria dos casos reais há presença de ruído nos dados, utiliza-se um número de amostras maior que o número de regressores. Desta forma, obtém-se uma matriz de regressores  $\mathbf{P}$  não quadrada, o que impede que o vetor de parâmetros  $\Theta$  seja determinado diretamente através da equação (2.23). Define-se então, um algoritmo para estimar  $\Theta$  que minimiza a seguinte função custo:

$$J = \frac{1}{2} \sum_{k=1}^N \xi^2(k) = \frac{1}{2} \xi^T \xi = \frac{1}{2} \|\xi\|^2, \quad (2.26)$$

onde  $\|\cdot\|$  indica norma euclidiana.

A solução que minimiza a equação (2.26) é o estimador de mínimos quadrados definido por:

$$\hat{\Theta} = (\mathbf{P}^T \mathbf{P})^{-1} \mathbf{P}^T \mathbf{y} . \quad (2.27)$$

## 2.8 Regressores Ortogonais

A estimação direta dos parâmetros do modelo envolvendo todos os termos candidatos, geralmente resulta em mal condicionamento da matriz de regressores. Sendo assim, técnicas de ortogonalização podem ser utilizadas a fim de contornar o problema. A equação (2.22) pode ser escrita na forma que se segue:

$$\mathbf{y} = \mathbf{P} \hat{\Theta} + \Xi , \quad (2.28)$$

onde

$$\mathbf{y} = \begin{bmatrix} y(1) \\ \vdots \\ y(N) \end{bmatrix}, \quad \mathbf{P} = \begin{bmatrix} \mathbf{p}^T(1) \\ \vdots \\ \mathbf{p}^T(N) \end{bmatrix}, \quad \hat{\Theta} = \begin{bmatrix} \hat{\theta}_1 \\ \vdots \\ \hat{\theta}_M \end{bmatrix}, \quad \Xi = \begin{bmatrix} \xi(1) \\ \vdots \\ \xi(N) \end{bmatrix}, \quad (2.29)$$

e

$$\mathbf{p}^T(k) = [p_1(k), \dots, p_M(k)] . \quad (2.30)$$

As seguintes condições, em relação à equação (2.28), devem ser satisfeitas:

1.  $\xi$  é ruído branco. Então, no caso polinomial,  $\xi$  é não correlacionado com  $p_i(k)$ ,  $i = 1, \dots, M$ .
2. Os processos a serem identificados são ergódicos.
3.  $\mathbf{P}^T\mathbf{P}$  é definida positiva.

A primeira condição é para garantir estimadores não polarizados. A segunda é introduzida para que a média de realizações seja equivalente à média temporal. A terceira condição, por sua vez, vem evitar que a matriz normal  $\mathbf{P}^T\mathbf{P}$  seja singular ou mal condicionada<sup>2</sup>. Se, por outro lado, a terceira condição não for satisfeita, métodos numéricos podem ser utilizados na tentativa de contornar o problema, ou seja, pode-se dizer que esta não é uma condição rígida. No entanto, neste trabalho, todas as condições deverão ser satisfeitas.

### 2.8.1 Método Clássico de Ortogonalização de Gram-Schmidt

Assumindo que a terceira condição tenha sido satisfeita, pode-se decompor  $\mathbf{P}^T\mathbf{P}$  em:

$$\mathbf{P}^T\mathbf{P} = \mathbf{A}^T\mathbf{D}\mathbf{A} , \quad (2.31)$$

onde  $\mathbf{A}$  é uma matriz triangular superior,  $M \times M$ , com os elementos da diagonal principal iguais a um:

$$\mathbf{A} = \begin{bmatrix} 1 & \alpha_{12} & \alpha_{13} & \cdots & \alpha_{1M} \\ 0 & 1 & \alpha_{23} & \cdots & \alpha_{2M} \\ & & & \vdots & \\ 0 & 0 & 0 & 1 & \alpha_{M-1M} \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \quad (2.32)$$

e  $\mathbf{D}$  é uma matriz diagonal,  $M \times M$ , com elementos positivos. Então:

$$\mathbf{y} = \mathbf{P}(\mathbf{A}^{-1}\mathbf{A})\Theta + \Xi \quad (2.33)$$

---

<sup>2</sup>Se  $\mathbf{P}^T\mathbf{P}$  é definida positiva, a sua forma diagonal será composta somente por elementos positivos e, desta forma, seu determinante será maior que zero, ou seja, se  $\mathbf{P}^T\mathbf{P}$  é definida positiva então é também uma matriz não singular (Lipschutz, 1991).

ou

$$\mathbf{y} = \mathbf{W}\mathbf{g} + \boldsymbol{\Xi} \quad (2.34)$$

onde

$$\left. \begin{aligned} \mathbf{W} &= \mathbf{P}\mathbf{A}^{-1} \\ \mathbf{g} &= \mathbf{A}\boldsymbol{\Theta} \end{aligned} \right\} \quad (2.35)$$

A matriz  $\mathbf{W}$  é definida por:

$$\mathbf{W} = \begin{bmatrix} w_1(1) & \cdots & w_M(1) \\ \vdots & & \vdots \\ w_1(N) & \cdots & w_M(N) \end{bmatrix} \quad (2.36)$$

onde  $w_i$  são os regressores auxiliares e  $g_i$  são os parâmetros estimados utilizando-se tais regressores. O processo de estimação de parâmetros utilizando-se os novos regressores é realizado como se segue;

- Passo 1:  $\alpha_{11} = 1$ ,  $w_1(k) = p_1(k)$  e  $\hat{g}_1 = \frac{\sum_{k=1}^N w_1(k)y(k)}{\sum_{k=1}^N w_1^2(k)}$
- Passo  $j$ :  $j = 2, \dots, M$ :  $\alpha_{ij} = \frac{\sum_{k=1}^N w_i(k)p_j(k)}{\sum_{k=1}^N w_i^2(k)}$ ,  $i = 1, \dots, j-1$ ,  $\alpha_{jj} = 1$   
 $w_j(k) = p_j(k) - \sum_{i=1}^{j-1} \alpha_{ij}w_i(k)$ , e  $\hat{g}_j = \frac{\sum_{k=1}^N w_j(k)y(k)}{\sum_{k=1}^N w_j^2(k)}$

- Calcular  $\hat{\theta}_i$  utilizando as seguintes equações:

$$\left. \begin{aligned} \hat{\theta}_M &= \hat{g}_M \\ \hat{\theta}_i &= \hat{g}_i - \sum_{j=i+1}^M \alpha_{ij}\hat{\theta}_j, \quad i = M-1, \dots, 1 \end{aligned} \right\} \quad (2.37)$$

Algumas propriedades do estimador ortogonal são (Billings et al., 1989):

- Os regressores auxiliares,  $w_i$ , são ortogonais, ou seja:  $\sum_{k=1}^N w_i(k)w_j(k) = 0$ ,  $i \neq j$ ;

- O vetor de parâmetros  $\hat{\Theta}$  é equivalente ao obtido através da equação normal  $(\mathbf{P}^T \mathbf{P}) \Theta = \mathbf{P}^T \mathbf{y}$ , o que era esperado pois o algoritmo ortogonal é apenas outra forma de solucionar a equação normal;
- A proporção da variância da saída explicada por  $w_i$  é:

$$[err]_i = \frac{g_i^2 \sum_{k=1}^N w_i^2(k)}{\sum_{k=1}^N y^2(k)}, \quad i = 1, \dots, M, \quad (2.38)$$

que será detalhada na seção (2.9.2).

O erro quadrático médio de predição (MSPE - *Mean Square Prediction Error*) é máximo quando não existe nenhum termo no modelo, ou seja, quando  $n_\theta = n_p + n_n = 0$ , sendo igual ao valor quadrático médio do próprio sinal. A inserção do  $i$ -ésimo termo  $g_i w_i$  ao modelo, ocasiona redução no MSPE de um fator igual a  $g_i^2 w_i^2(k)$  (Billings et al., 1989). Desta forma, quanto maior o fator  $g_i^2 w_i^2(k)$ , maior a proporção da variância da saída explicada pelo referido termo e portanto, maior é a importância do termo para o modelo.

### 2.8.2 Transformação de Householder

Um outro método de ortogonalização é a transformação de Householder (Chen, 1970). Seja a matriz de regressores  $\mathbf{P}$  e o vetor de saída  $\mathbf{y}$  (2.24). Definindo a seguinte matriz de regressores modificada:

$$\tilde{\mathbf{P}}^{(0)} = [\mathbf{P} : \mathbf{y}] = [\mathbf{p}_1 \cdots \mathbf{p}_M : \mathbf{y}] = [\tilde{\mathbf{p}}_1^{(0)} \cdots \tilde{\mathbf{p}}_M^{(0)} : \mathbf{y}^{(0)}], \quad (2.39)$$

a  $k$ -ésima transformação de Householder é definida como (Chen et al., 1989):

$$\left. \begin{aligned} \sigma^{(k)} &= \left( \sum_{i=k}^N (\tilde{p}_{ik}^{(k-1)})^2 \right)^{1/2} \\ \beta^{(k)} &= \frac{1}{\sigma^{(k)} (\sigma^{(k)} + |\tilde{p}_{kk}^{(k-1)}|)} \\ v_i^{(k)} &= \begin{cases} 0, & \text{para } i < k \\ \tilde{p}_{kk}^{(k-1)} + \text{sgn}[\tilde{p}_{kk}^{(k-1)}] \sigma^{(k)}, & \text{para } i = k \\ \tilde{p}_{ik}^{(k-1)}, & \text{para } i > k \end{cases} \end{aligned} \right\} \quad (2.40)$$

e

$$\tilde{\mathbf{P}}^{(k)} = \tilde{\mathbf{P}}^{(k-1)} - \mathbf{v}^{(k)}(\beta^{(k)}(\mathbf{v}^{(k)})^T \tilde{\mathbf{P}}^{(k-1)}), \quad k = 1, \dots, M. \quad (2.41)$$

Desta forma, no instante seguinte à  $k$ -ésima iteração, as colunas da matriz modificada  $\tilde{P}$  de ordem inferior a  $k$ , serão ortogonais entre si. Obtém-se a seguinte matriz triangular superior  $\tilde{P}$  de dimensão  $N \times (M + 1)$ :

$$\tilde{\mathbf{P}}^{(M)} = \begin{bmatrix} \mathbf{R} & \mathbf{y}_1 \\ 0 & \mathbf{y}_2 \end{bmatrix}, \quad (2.42)$$

onde  $\mathbf{R}$  é uma matriz triangular superior  $M \times M$ ,  $\mathbf{y}_1 = y(1 : M)$  e  $\mathbf{y}_2 = y(M + 1 : N)$ .

Considerando a seguinte equação matricial (Chen et al., 1989):

$$\mathbf{R}\boldsymbol{\Theta} = \mathbf{y}_1, \quad (2.43)$$

os parâmetros podem ser estimados por:

$$\hat{\boldsymbol{\Theta}} = \mathbf{R}^{-1} \mathbf{y}_1. \quad (2.44)$$

A soma do quadrado dos resíduos, por sua vez, será dada por:

$$\|\mathbf{y}_2\|^2. \quad (2.45)$$

Observa-se que no processo de ortogonalização através da transformação de Householder, ao contrário do processo de ortogonalização de Gram-Schmidt, não há a necessidade da utilização de um vetor de parâmetros auxiliar.

## 2.9 Detecção de Estrutura

Um problema chave em identificação de sistemas é encontrar uma estrutura adequada, através da qual um “bom” modelo pode ser estimado.

Quando se estimam os parâmetros do modelo através de um método de mínimos quadrados, minimiza-se a função custo descrita pela equação (2.26). No entanto, não é possível garantir, em primeiro lugar, se o modelo é fiel aos dados experimentais e em segundo, se o estimador é polarizado ou não. Tais

inconvenientes só podem ser contornados se a *estrutura escolhida for adequada*, ou seja, se a mesma é capaz de representar o sistema em questão.

Infelizmente, não é obvio qual é a melhor estrutura para representar um sistema. Sabe-se, no entanto, que um modelo com mais termos que o necessário, geralmente provoca, na melhor das hipóteses, efeitos espúrios e na pior, instabilidade numérica (Aguirre, 1994a). Além do número de termos do modelo, efeitos espúrios também podem ser ocasionados se o modelo estiver sobreparametrizado em relação ao número de pontos fixos e à ordem do mesmo (Mendes and Billings, 1998). Sabe-se ainda que, um modelo muito flexível causa generalização pobre (Reed, 1993), além de que modelos menores têm maior velocidade de processamento e maior simplicidade.

Embora a determinação da estrutura do modelo seja essencial, esta ainda é a tarefa mais árdua da identificação de sistemas não lineares, em vista da grande quantidade de modelos possíveis e da inexistência de uma regra que permita escolher, de imediato, tal estrutura. Por esta razão este ainda é um problema pesquisado, embora exista grande número de trabalhos na literatura a respeito. No caso de modelos NARMAX, usam-se critérios de informação (Leontaritis and Billings, 1987; Kortmann et al., 1988), ERR (*Error Reduction Ratio*) (Billings et al., 1989; Chen et al., 1989) e agrupamentos de termos (Aguirre, 1994a; Aguirre and Billings, 1995a; Aguirre and Mendes, 1996; Aguirre, 1994b), dentre outros, para escolher a melhor estrutura que seja capaz de representar o sistema. Tais critérios serão descritos a seguir.

### 2.9.1 Agrupamentos de Termos

Expandindo o modelo NARX (2.19) na forma polinomial, com termos com grau de não linearidade variando entre  $1 \leq m_l \leq l$ , cada termo de grau  $m_l$  pode conter um fator em  $y(k - n_i)$  de ordem  $p$ , um fator em  $u(k - n_i)$  de ordem  $(m_l - p)$  e um coeficiente  $\theta_{p, m_l - p}(n_1, \dots, n_{m_l})$ :

$$y(k) = \sum_{m_l=0}^l \sum_{p=0}^{m_l} \sum_{n_1, n_{m_l}}^{n_y, n_u} \theta_{p, m_l - p}(n_1, \dots, n_{m_l}) \prod_{i=1}^p y(k - n_i) \prod_{i=p+1}^{m_l} u(k - n_i) \quad (2.46)$$

onde

$$\sum_{n_1, n_{m_l}}^{n_y, n_u} = \sum_{n_1=1}^{n_y} \cdots \sum_{n_{m_l}=1}^{n_u} . \quad (2.47)$$

Por exemplo, expandindo a equação (2.46) para grau de não linearidade  $l$  igual a dois:

$$\begin{aligned}
y(k) = & \theta_{0,0} + \sum_{n_1=1}^{n_y} \theta_{1,0}(n_1)y(k-n_1) + \sum_{n_1=1}^{n_u} \theta_{0,1}(n_1)u(k-n_1) + \\
& \sum_{n_1=1}^{n_y} \sum_{n_2=1}^{n_u} \theta_{2,0}(n_1, n_2)y(k-n_1)y(k-n_2) + \\
& \sum_{n_1=1}^{n_y} \sum_{n_2=1}^{n_u} \theta_{1,1}(n_1, n_2)y(k-n_1)u(k-n_2) + \\
& \sum_{n_1=1}^{n_y} \sum_{n_2=1}^{n_u} \theta_{0,2}(n_1, n_2)u(k-n_1)u(k-n_2) . \tag{2.48}
\end{aligned}$$

É importante salientar que a massa de dados composta por  $u(k)$  e  $y(k)$  foi obtida através dos dados contínuos  $u(t)$  e  $y(t)$  com taxa de amostragem igual a  $1/T_s$  e que os coeficientes dos termos dependem do período de amostragem  $T_s$ .

Se o período de amostragem for pequeno de tal forma que:

$$\left. \begin{aligned}
y(k-1) &\approx y(k-2) \approx \dots \approx y(k-n_y) \\
u(k-1) &\approx u(k-2) \approx \dots \approx u(k-n_u)
\end{aligned} \right\} ,$$

a equação (2.46) pode ser escrita na forma que se segue:

$$y(k) \approx \sum_{n_1, n_{m_l}}^{n_y, n_u} \theta_{p, m_l-p}(n_1, \dots, n_{m_l}) \sum_{m_l=0}^l \sum_{p=0}^{m_l} y(k-n_i)^p u(k-n_i)^{m_l-p} \tag{2.49}$$

As constantes  $\sum_{n_1, n_{m_l}}^{n_y, n_u} \theta_{p, m_l-p}(n_1, \dots, n_{m_l})$  na equação acima são os coeficientes dos agrupamentos de termos  $\Omega_{y^p u^{m_l-p}}$ , que contém termos da forma  $y(k-i)^p u(k-j)^{m_l-p}$  para  $m_l = 0, \dots, l$  e  $p = 0, \dots, m_l$ . Tais coeficientes são denominados coeficientes de agrupamentos, sendo representados por  $\sum_{y^p u^{m_l-p}}$  (Aguirre and Billings, 1995a).

Então, o conjunto de termos candidatos para o modelo NARX é constituído pela união de todos os agrupamentos possíveis, onde:

$$\begin{aligned}
\text{agrupamentos possíveis} &= \bigcup_{p=0, \dots, m_l; m_l=0, \dots, l} \Omega_{y^p u^{m_l-p}} \\
&= \text{constante} \cup \Omega_y \cup \Omega_u \cup \Omega_{y^2} \cup \Omega_{yu} \\
&\quad \cup \Omega_{u^2} \cup \dots \tag{2.50}
\end{aligned}$$

É importante salientar que quando  $T_s \rightarrow 0$ , os termos dentro de cada agrupamento se tornarão equivalentes e seus respectivos coeficientes tenderão a zero, exceto  $\Sigma_y$ , que assumirá valor unitário (Aguirre and Billings, 1995a), ou seja:

$$\lim_{T_s \rightarrow 0} \Sigma_y = 1 \quad (2.51)$$

e

$$\lim_{T_s \rightarrow 0} \Sigma_{y^p} u^{m-p} = 0 \text{ para os demais agrupamentos.} \quad (2.52)$$

Então um agrupamento de termos é definido como um conjunto de termos do mesmo tipo e sendo seus coeficientes definidos como a soma dos coeficientes de todos os termos do agrupamento. Esta abordagem baseia-se no fato de que dois ou mais termos desnecessários podem ter coeficientes com amplitudes significativas, mas que se anulam ou possuem valores irrelevantes se comparados com os demais agrupamentos, caracterizando um agrupamento espúrio. Tais agrupamentos além de aumentar o número de termos candidatos, tendem a gerar dinâmicas não existentes no sistema original diminuindo, portanto, a confiabilidade do modelo.

Esta abordagem é diferente à proposta por (Kadtke et al., 1993), que apresenta um método de detecção baseado na eliminação de coeficientes *dos termos* que possuem valor inferior a um certo limiar, método denominado *zeroing and refitting*. Nessa abordagem, após a eliminação dos coeficientes, reestimam-se os parâmetros do modelo reduzido, sendo baseado na idéia de que termos espúrios têm coeficientes insignificantes.

Um problema introduzido pelo *zeroing and refitting* é que os coeficientes dependem do período de amostragem e da existência de ruído nos dados, além de que termos espúrios podem apresentar, em alguns casos, valores com ordem de grandeza relevante ou até mesmo superior a dos efetivos, ou ainda, os termos efetivos podem ter coeficientes pequenos (Aguirre, 1994a).

### 2.9.2 A Taxa de Redução de Erro

Métodos de ortogonalização da matriz de regressores, como por exemplo, o método clássico de ortogonalização de Gram-Schmidt (Billings et al., 1989) ou a transformação de Householder (Chen et al., 1989), descritos na seção (2.8), podem ser utilizados para auxiliar no processo de detecção de estrutura. Tais métodos, além de eliminar os problemas de mal condicionamento numérico,

permitem analisar a relevância de cada termo para o modelo, verificando a porção da variância da saída explicada pelo mesmo.

Utilizando a propriedade dos regressores ortogonais definida pela equação (2.38) e o método clássico de Gram Schmidt, introduzido anteriormente, pode-se definir um critério para detecção de estrutura denominado ERR (*Error Reduction Ratio*) que é dado pela seguinte expressão:

$$[err]_i = \frac{g_i^2 w_i^T w_i}{y^T y}, \quad 1 \leq i \leq M, \quad (2.53)$$

onde  $w_i$  e  $g_i$  são os regressores ortogonais e os parâmetros estimados, respectivamente,  $y$  é a saída e  $M$  é o número de termos possíveis para o modelo, segundo o critério adotado. Quanto maior o índice ERR, maior é a importância do termo para o modelo. Assim, os termos que apresentarem um índice ERR muito baixo, não são incluídos no modelo. Desta forma, este critério evita mal condicionamento numérico já que termos linearmente dependentes ou irrelevantes são eliminados.

O índice  $err_i$  do  $i$ -ésimo termo,  $w_i$ , indica a proporção do valor médio quadrático da saída,  $\mathbf{y}^T \mathbf{y}$ , explicada por esse termo. Consequentemente,

$$\sum_{i=1}^{n_\theta} err_i = 1, \quad (2.54)$$

e desta forma, os termos irrelevantes para o modelo, contribuirão com um percentual desprezível para o índice  $err$ , podendo ser eliminados.

Após a seleção dos termos relevantes para o modelo, calculam-se os parâmetros originais  $\theta_i$ , caso se esteja usando o método de Gram-Schmidt, ou seja, este critério permite ao mesmo tempo selecionar a estrutura e estimar os parâmetros do modelo.

De forma similar, pode-se determinar a estrutura através da transformação de Householder (Chen et al., 1989). Definindo:

$$(y_j^i)^2 = \left( \frac{b_j^{(k)}}{a_j^{(k)}} \right)^2 \text{ para } j = k, \dots, M, \quad (2.55)$$

onde  $a_j^{(k)}$  e  $b_j^{(k)}$  são inicializados da seguinte maneira:

$$(a_j^{(1)})^2 = \sum_{i=1}^N (\tilde{p}_{ij}^{(0)})^2,$$

$$b_j^{(1)} = \sum_{i=1}^N \tilde{p}_{ij}^{(0)} y_i(0), \quad j = 1, \dots, M \quad (2.56)$$

Calculando  $\tilde{\mathbf{P}}^{(k)}$  de acordo com as Equações (2.40) e (2.41),  $(a_j^{(k+1)})^2$  e  $b_j^{(k+1)}$  para  $j = k + 1, \dots, M$  podem ser determinados por:

$$\begin{aligned} (a_j^{(k+1)})^2 &= (a_j^{(k)})^2 - (\tilde{p}_{kj}^{(k)})^2, \\ b_j^{(k+1)} &= b_j^{(k)} - \tilde{p}_{kj}^{(k)} y_k^{(k)}, \quad j = k + 1, \dots, M \end{aligned} \quad (2.57)$$

Os regressores selecionados para o modelo serão aqueles que possuírem maiores valores de  $(y_j^{(i)})^2$ . Desta forma, o fator  $(y_j^{(i)})^2 / \langle \mathbf{y}, \mathbf{y} \rangle$  corresponde ao índice  $[err]_i$  calculado pelo método de ortogonalização de Gram-Schmidt.

É importante salientar que, a transformação de Householder elimina a necessidade de se estimarem os parâmetros durante a detecção de estrutura como ocorre no método clássico de Gram-Schmidt.

### 2.9.3 Critérios de Informação

O AIC (*Akaike's Information Criterion*) (Akaike, 1974) é um critério estatístico, que verifica a redução na variância dos resíduos à medida que termos são incorporados ao modelo, mas levando em consideração o princípio da parsimônia<sup>3</sup>.

Sabendo-se que o nível de significância do critério AIC é dado pelo nível de significância da distribuição  $\chi^2$  com  $s$  graus de liberdade<sup>4</sup>, e valor crítico  $\phi(s) = 2s$ . O AIC é definido através da seguinte expressão:

$$AIC(\phi) = N \log[\text{Var}(\xi)] + n_\theta \phi, \quad (2.58)$$

onde  $N$  é o número de dados de treinamento,  $\text{Var}(\xi)$  é a variância dos resíduos,  $n_\theta$  é o número de termos do modelo,  $\phi$  é o valor crítico da distribuição  $\chi^2$  com um grau de liberdade para um dado nível de significância.

<sup>3</sup>O princípio da parsimônia afirma que, entre dois ou mais modelos candidatos eficazes em relação à representação da dinâmica do sistema, deve-se escolher o modelo com o menor número de parâmetros independentes (Söderström and Stoica, 1989). Tal princípio permite a escolha de um modelo mais fiel às características do sistema.

<sup>4</sup>O número de graus de liberdade corresponde à diferença entre o número de parâmetros de dois modelos a serem comparados e o nível de significância corresponde à probabilidade de selecionar-se um modelo com  $s$  termos a mais do que o necessário.

Os valores que apresentam melhores resultados são AIC(2) e AIC(4) (Leontaritis and Billings, 1987) e portanto, os mais utilizados.

O primeiro termo é relativo à variância dos resíduos e o segundo segue o princípio da parsimônia, que penaliza a complexidade do modelo (Leontaritis and Billings, 1987). Desta forma, um modelo que inclui todos os termos candidatos faz com que o primeiro termo da Equação (2.58) seja mínimo, uma vez que a variância dos resíduos diminui com o aumento do número de parâmetros. Por outro lado, o número de termos candidatos  $M$  sendo elevado, ocasionará valor máximo no segundo termo da equação. Desta forma, este critério reduz as chances de se determinar um modelo sobreparametrizado.

Alguns outros critérios de informação tais como FPE (*Final Prediction Error Technique*), LILC (*Khinchin's Law of Iterated Logarithm Criterion*) e BIC (*Bayesian Information Criterion*) (Kortmann et al., 1988; Leontaritis and Billings, 1987) também podem ser usados.

É importante salientar que, a eficiência dos critérios de informação dependerá das características do processo (Kortmann et al., 1988), ou seja, tais critérios indicam se um modelo é adequado do ponto de vista estatístico, ou seja, não há garantias de que o modelo selecionado será dinamicamente eficiente. Por outro lado, a complexidade de um modelo selecionado com base nos testes estatísticos é, geralmente, próxima da complexidade de bons modelos escolhidos de acordo com sua dinâmica. Desta forma, quando o sistema a ser identificado for não linear, tais critérios podem ser utilizados para reduzir o número de modelos candidatos (Aguirre, 1994a).

Além disso, deve-se ter em mente que o número de termos sugerido pelo critério AIC dependerá da escolha dos termos a serem incluídos no modelo, ou seja, do grau de não linearidade, dos atrasos máximos e da ordem de importância definida pelo algoritmo baseado no ERR. Desta forma, alterando-se estas condições obtêm-se resultados diferentes para o critério AIC.

#### 2.9.4 Composição dos Métodos Citados para Detectar a Estrutura do Modelo

Em relação à detecção de estrutura de um modelo NARMAX, pode-se considerar duas categorias: detecção através de técnicas de crescimento da estrutura e detecção através de técnicas de eliminação. A abordagem através de eliminação é mais susceptível a problemas numéricos do que através de técnicas de construção. Por outro lado, na presença de ruído, as técnicas construtivas tendem a selecionar termos desnecessários. Desta forma, o ideal é então, utilizar ambas técnicas. A construtiva é usada para evitar problemas numéricos e modelos muito complexos e a eliminativa é usada para descartar

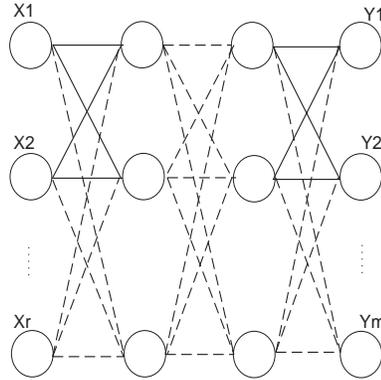


Figura 2.1: Estrutura genérica de uma rede *Feedforward*. Note que os termos de polarização não estão indicados.

termos desnecessários do modelo (Aguirre, 1994a). Sendo assim, as técnicas de agrupamentos de termos são utilizadas para selecionar os agrupamentos efetivos. Este procedimento reduz o número de termos candidatos para o modelo e aumentam as chances de se obter um modelo que melhor represente a dinâmica do sistema original. Em seguida, utiliza-se o critério da taxa de redução de erro (ERR) para determinar a ordem de importância dos termos dos agrupamentos selecionados e o AIC para determinar o número de termos necessário, permitindo selecionar quantos serão escolhidos de acordo com a ordem de importância indicada pelo índice ERR.

## 2.10 Redes *Multi-Layer Perceptron*

A rede *Feedforward* indicada na Figura (2.1), *Multi-Layer Perceptron* (MLP) (Braga et al., 1998; Haykin, 1994), é constituída de unidades de processamento elementares, denominadas neurônios. Nesta estrutura, as saídas dos neurônios de uma determinada camada são conectadas às entradas da camada seguinte, através de conexões sinápticas (pesos).

A camada de entrada não executa nenhuma transformação, apenas passa os sinais de entrada para os neurônios da camada seguinte. Em identificação, o número de neurônios da camada de saída depende do número de variáveis de saída do sistema. O número de neurônios da camada de entrada é determinado pelo número de variáveis de entrada e saída do sistema e por seus

atrasos máximos,  $n_u$  e  $n_y$ <sup>5</sup> e o número de neurônios das camadas intermediárias depende da complexidade dos dados (complexidade do sistema). Este tipo de estrutura *feedforward* com as variáveis do processo atrasadas na entrada da rede é denominado *Time Delay Neural Network* - TDNN (Lang and Hinton, 1988).

Sabe-se que, Redes Neurais Artificiais com uma camada interna, são capazes de aproximar com certa precisão, qualquer função contínua, quando não se faz restrições sobre o número de nodos ou sobre a ordem de grandeza dos pesos (Cybenko, 1989). Sendo assim, serão consideradas redes com apenas uma camada intermediária, contendo  $n$  nodos nesta camada,  $r$  entradas e  $m$  saídas.

A saída do  $j$ -ésimo neurônio da camada intermediária, no instante  $k$  é:

$$y_j(k) = f_j\left(\sum_{i=1}^{N_r} w_{ji}^h(k)x_i(k) + b_j^h(k)\right) , \quad (2.59)$$

onde  $x_i$  é a  $i$ -ésima entrada do nodo,  $w_{ji}^h$  são as conexões sinápticas entre um neurônio da camada de entrada e outro da camada intermediária,  $b_j^h$  são termos de polarização,  $y_j$  é a saída do neurônio  $j$ ,  $N_r$  é o número de entradas aplicadas ao neurônio e  $f_j$  é uma função não linear, denominada função de ativação, sendo sigmóides e tangentes hiperbólicas as mais utilizadas. A função de ativação da camada intermediária será designada somente por  $f$  indicando que o mesmo tipo de função será escolhido para todos os neurônios desta camada.

A saída do  $i$ -ésimo neurônio da camada de saída é:

$$y_i(k) = f_i\left(\sum_{j=1}^n w_j^o(k)f\left(\sum_{i=1}^{N_r} w_{ji}^h(k)x_i(k) + b_j^h(k)\right) + b_i^o(k)\right) , \quad (2.60)$$

onde  $w_j^o$  são os pesos da camada de saída,  $b_i^o$  são os termos de polarização para esta camada e  $n$  é o número de neurônios da camada intermediária.

Em identificação, pode-se considerar a camada de saída como sendo linear (Billings and Chen, 1992; Häck and Köhne, 1996) e não utilizar termos de polarização nesta camada (Billings and Chen, 1992), obtendo-se na saída da rede:

$$y_i(k) = \sum_{j=1}^n w_j^o f\left(\sum_{i=1}^{N_r} w_{ji}^h(k)x_i(k) + b_j^h(k)\right) . \quad (2.61)$$

---

<sup>5</sup>O número de entradas da rede, aqui considerada, será correspondente ao número  $n_r$  de regressores (2.20) de um modelo NARX.

Pode-se expressar a equação acima também por:

$$y_i(k) = \sum_{j=1}^n w_j^o f(v_j) , \quad (2.62)$$

onde

$$v_j = \sum_{i=1}^{N_r} w_{ji}^h(k) x_i(k) + b_j^h(k) . \quad (2.63)$$

A fim de analisar e comparar a estrutura interna da rede com o modelo polinomial NARMAX, a mesma será decomposta em Série de Taylor (Häck and Köhne, 1996).

Seja  $f(v_j) = \tanh(v_j)$  analítica na vizinhança de um ponto  $a$ . Decompondo  $f(v_j)$  em Série de Taylor (Kreyszig, 1993), tem-se:

$$f(v_j) = \sum_{m=0}^{\infty} \frac{f^{(m)}(a)}{m!} (v_j - a)^m . \quad (2.64)$$

Assumindo  $a = 0$  tem-se:

$$\begin{aligned} f(v_j) = & \frac{f(0)}{0!} + \frac{f'(0)}{1!} v_j + \frac{f''(0)}{2!} v_j^2 + \frac{f'''(0)}{3!} v_j^3 + \\ & \frac{f^{IV}(0)}{4!} v_j^4 + \frac{f^V(0)}{5!} v_j^5 + \dots . \end{aligned} \quad (2.65)$$

Calculando as derivadas da função  $f(v_j) = \tanh(v_j)$ :

$$\begin{aligned} f'(v_j) &= \operatorname{sech}^2(v_j) \\ f''(v_j) &= -2\operatorname{sech}^2(v_j)\tanh(v_j) \\ f'''(v_j) &= -2\operatorname{sech}^4(v_j) + 4\operatorname{sech}^2(v_j)\tanh^2(v_j) \\ f^{IV}(v_j) &= 16\operatorname{sech}^4(v_j)\tanh(v_j) - 8\operatorname{sech}^2(v_j)\tanh^3(v_j) \\ f^V(v_j) &= 16\operatorname{sech}^6(v_j) - 88\operatorname{sech}^4(v_j)\tanh^2(v_j) + 16\operatorname{sech}^2(v_j)\tanh^4(v_j) \\ &\vdots \end{aligned} \quad (2.66)$$

No ponto  $a = 0$ , as derivadas indicadas em (2.66) assumem os seguintes valores:

$$\begin{aligned} f'(0) &= 1 \\ f''(0) &= 0 \\ f'''(0) &= -2 \\ f^{IV}(0) &= 0 \\ f^V(0) &= 16 . \end{aligned} \quad (2.67)$$

Substituindo (2.67) em (2.65), a saída do  $j$ -ésimo nodo da camada intermediária será:

$$y_j(k) = \tanh(v_j) = v_j - \frac{1}{3}v_j^3 + \frac{2}{15}v_j^5 + \dots \quad (2.68)$$

A saída da rede contendo uma camada intermediária com  $n$  nodos, realizando função de ativação tangente hiperbólica,  $r$  entradas e uma saída com função de ativação linear, será:

$$y(k) = w_1^o \tanh(v_1) + \dots + w_n^o \tanh(v_n) \quad (2.69)$$

Substituindo a equação (2.68) na (2.69), tem-se:

$$y(k) = w_1^o [v_1 - \frac{1}{3}v_1^3 + \frac{2}{15}v_1^5 + \dots] + \dots + w_n^o [v_n - \frac{1}{3}v_n^3 + \frac{2}{15}v_n^5 + \dots] \quad (2.70)$$

Decompondo  $v_j$  de acordo com a equação (2.63), tem-se uma expressão para a saída da rede:

$$\begin{aligned} y(k) = & w_1^o [(w_{11}^h x_1 + \dots + w_{1r}^h x_r + b_1^h) - \frac{1}{3}(w_{11}^h x_1 + \dots + w_{1r}^h x_r + b_1^h)^3 + \\ & \frac{2}{15}(w_{11}^h x_1 + \dots + w_{1r}^h x_r + b_1^h)^5 + \dots] + \dots + w_n^o [(w_{n1}^h x_1 + \dots + w_{nr}^h x_r + b_n^h) - \\ & \frac{1}{3}(w_{n1}^h x_1 + \dots + w_{nr}^h x_r + b_n^h)^3 + \frac{2}{15}(w_{n1}^h x_1 + \dots + w_{nr}^h x_r + b_n^h)^5 + \dots] \quad (2.71) \end{aligned}$$

A fim de ilustrar a Equação (2.71), considere uma rede com um nodo na camada intermediária, duas entradas e uma saída. Utilizando a equação (2.71) para este caso, tem-se:

$$\begin{aligned} y(k) = & w_1^o [w_{11}^h x_1 + w_{12}^h x_2 + b_1^h - \frac{1}{3}((w_{11}^h x_1)^3 + (w_{12}^h x_2)^3 + (b_1^h)^3 + \\ & 3(w_{11}^h x_1)^2 w_{12}^h x_2 + 3(w_{12}^h x_2)^2 w_{11}^h x_1 + 3(b_1^h)^2 w_{11}^h x_1 + \\ & 3(b_1^h)^2 w_{12}^h x_2 + 3b_1^h (w_{11}^h x_1)^2 + 3b_1^h (w_{12}^h x_2)^2 + \\ & 6b_1^h w_{12}^h x_2 w_{11}^h x_1) + \dots] \quad (2.72) \end{aligned}$$

Supondo que as entradas  $x_1$  e  $x_2$  da rede sejam compostas por regressores originários de agrupamentos distintos e agrupando-se os termos referentes aos mesmos agrupamentos, tem-se:

$$\begin{aligned} y(k) = & [w_1^o b_1 - \frac{1}{3}w_1^o (b_1^h)^3] + x_1 [w_1^o w_{11}^h (1 - (b_1^h)^2)] + x_2 [w_1^o w_{12}^h (1 - (b_1^h)^2)] + \\ & x_1^2 [-b_1^h w_1^o (w_{11}^h)^2] + x_2^2 [-b_1^h w_1^o (w_{12}^h)^2] + x_1 x_2 [-2b_1^h w_1^o w_{11}^h w_{12}^h] + \\ & x_1^3 [-\frac{1}{3}w_1^o (w_{11}^h)^3] + x_2^3 [-\frac{1}{3}w_1^o (w_{12}^h)^3] + x_1^2 x_2 [-w_1^o (w_{11}^h)^2 w_{12}^h] + \\ & x_1 x_2^2 [-w_1^o (w_{12}^h)^2 w_{11}^h] + \dots \quad (2.73) \end{aligned}$$

Renomeando os coeficientes de cada termo:

$$y(k) = a_0 + a_1x_1 + a_2x_2 + a_3x_1^2 + a_4x_2^2 + a_5x_1x_2 + a_6x_1^3 + a_7x_2^3 + a_8x_1x_2^2 + a_9x_1^2x_2 + \dots \quad (2.74)$$

Se as entradas da RNA forem sinais dependentes das variáveis de entrada e saída do sistema, bem como de seus respectivos atrasos, o modelo representado em (2.74) pode ser visto como um polinomial NARX<sup>6</sup>, onde o número de regressores  $n_r$  do modelo polinomial NARX equivale ao número de entradas da rede, enquanto que, o grau de não linearidade da rede expandida na forma polinomial, dependerá da função polinomial utilizada para aproximar a função de ativação.

Quando o objetivo da identificação é recuperar não linearidades estáticas provenientes do sistema original, deve-se escolher os regressores e agrupamentos adequados, ou seja, aqueles que contém informação necessária para tornar viável o propósito mencionado. Nestes casos, pode-se usar conhecimento *a priori* para auxiliar na determinação da estrutura do modelo (Aguirre et al., 1998b).

Observando a Equação (2.72) percebe-se que os regressores lineares, quadráticos ou cúbicos existem devido a um só nodo, ou seja, sua presença se deve à função de ativação e resulta em *combinações diferentes dos mesmos pesos*.

Supondo que fosse desejável eliminar um agrupamento, a eliminação de qualquer que seja o agrupamento, levaria de contrapartida a eliminação de algum outro. Por exemplo, se o agrupamento  $\Omega x_1^2$  é espúrio, ou impede de alguma forma a recuperação da não linearidade estática desejada<sup>7</sup>, o mesmo deveria ser eliminado do conjunto de agrupamentos possíveis para o modelo. No entanto, esta atitude levaria a rede, através de um processo de detecção de estrutura, por exemplo, *pruning* (Reed, 1993), a eliminar também a entrada  $x_1$  da rede, uma vez que ambos os agrupamentos dependem do peso  $w_{11}^h$ . Observando agora as Equações (2.73) e (2.74), percebe-se que o agrupamento  $\Omega x_1^2$  será eliminado se seu respectivo coeficiente  $a_3 = -b_1^h w_1^o (w_{11}^h)^2$  for zero. No entanto, os coeficientes dos demais agrupamentos que envolvem a entrada  $x_1$  da rede, também dependem do peso  $w_{11}^h$ , como por exemplo o coeficiente do agrupamento  $\Omega x_1$  é igual a  $a_1 = w_1^o w_{11}^h (1 - (b_1^h)^2)$ . Desta

<sup>6</sup>Embora o modelo seja não linear nos parâmetros, pois os parâmetros  $a_i$  são funções de vários pesos da rede, o modelo, de certa forma, é similar ao polinomial NARX.

<sup>7</sup>Cada agrupamento é responsável por representar um tipo de não linearidade presente nos dados.

forma, a eliminação de  $\Omega x_1^2$  ocasionada pela eliminação de  $w_{11}^h$ <sup>8</sup> acarretaria também a eliminação dos seguintes agrupamentos:  $\Omega x_1$ ,  $\Omega x_1 x_2$ ,  $\Omega x_1^3$ ,  $\Omega x_1 x_2^2$  e  $\Omega x_1^2 x_2$  (se considerada a série truncada para grau de não linearidade igual a três, pois caso contrário, outros agrupamentos poderiam ser eliminados).

Sendo assim, embora tanto redes MLPs quanto modelos polinomiais NARMAX apresentem resultados confiáveis na determinação de um modelo dinâmico para o sistema (Billings and Chen, 1992; Cassini et al., 1998), na recuperação analítica de não linearidades estáticas, através de conceitos de agrupamentos de termos e pontos fixos, modelos polinomiais (Pröll and Karim, 1994) e racionais (Aguirre et al., 1999) NARX parecem ser mais adequados.

## 2.11 Outras Representações para Sistemas Não Lineares

- Séries de Volterra e Wiener mapeiam entradas passadas na saída atual, resultando em um número excessivo de parâmetros (Chen and Billings, 1989), tornando inviável a sua utilização para sistemas de ordem superior a dois (Billings, 1980). Séries de Volterra podem ser representadas como se segue:

$$y(k) = \sum_{i=0}^{\infty} h_{x0/ki} (u(1), \dots, u(k)), \quad (2.75)$$

onde  $h_{x0/ki}$  são *kernels* de Volterra, que correspondem aos parâmetros a serem estimados. Para condições iniciais nulas,  $h_{x0/ki}$  é substituído por  $h_{ki}$ .

- Modelos por Blocos Orientados são aqueles representados por conexões entre sistemas dinâmicos lineares e elementos estáticos não lineares (Billings, 1980), ou seja, tais modelos exigem conhecimento *a priori* de características estáticas não lineares (Haber and Unbehauen, 1990). O que difere do procedimento realizado por modelos polinomiais NARMAX, pois embora conhecimento *a priori* possa ser utilizado para aux-

---

<sup>8</sup>O peso  $w_{11}^h$  foi aqui escolhido para exemplificar a questão, uma vez que os demais pesos acarretam maiores problemas. A eliminação do peso  $w_1^o$  seria equivalente a eliminar o modelo em questão. A eliminação do termo de polarização  $b_1^h$  vem por sua vez, interferir em uma maior quantidade de agrupamentos, inclusive alguns envolvendo a entrada  $x_2$  e o termo constante. Sendo assim, não foram consideradas tais possibilidades.

iliar na determinação da estrutura do modelo, esta não é uma condição estritamente necessária (Aguirre et al., 1998b).

Várias estruturas para sistemas não lineares podem ser consideradas como casos particulares do modelo NARMAX (2.12):

- Modelos Lineares: Caso um sistema seja operado próximo ao ponto de equilíbrio, um modelo linear pode ser utilizado para aproximar  $f(\cdot)$ . No entanto, o uso de tais representações é bastante restrito, já que é um modelo polinomial de grau um e na prática funções polinomiais de grau superior podem ser requeridas, afim de aproximar com maior precisão o sistema em questão (Chen and Billings, 1989).
- Modelos Lineares por Partes: São aqueles que separam os dados em várias faixas de operação, permitindo a utilização de modelos lineares (Çinar, 1995). A dificuldade de tais modelos, na prática, deve-se à necessidade da determinação de uma vizinhança, para a qual o sistema será considerado linear (Aguirre, 1994a), além de não fornecer, em muitos casos, a precisão necessária, sendo sua utilização bastante restrita (Chen and Billings, 1989).
- Modelos Bilineares têm como forma geral:

$$y(k) = a_0 + \sum_{i=1}^{n_y} a_i y(k-i) + \sum_{i=1}^{n_u} b_i u(k-i) + \sum_{i=1}^{n_y} \sum_{j=1}^{n_u} c_{ij} y(k-i)u(k-j) \quad (2.76)$$

Percebe-se que a representação acima é um caso particular do NARMAX (Chen and Billings, 1989), para grau de não linearidade igual a dois.

- Modelos Racionais têm a seguinte forma:

$$y(k) = \frac{b(y(k-1), \dots, y(k-n_y), u(k-1), \dots, u(k-n_u))}{a(y(k-1), \dots, y(k-n_y), u(k-1), \dots, u(k-n_u))}, \quad (2.77)$$

onde  $a(\cdot)$  e  $b(\cdot)$  são polinômios de grau finito (Chen and Billings, 1989).

- Modelos *Output-Affine*: Consiste em uma função,  $f(\cdot)$ , finitamente realizável e polinomial limitada, satisfazendo a seguinte equação (Chen and Billings, 1989):

$$y(k) = \sum_{i=1}^{n_y} \frac{a_i(u(k-1), \dots, u(k-n_y))}{a_0(u(k-1), \dots, u(k-n_y))} y(k-i) + \frac{a_{n_y+1}(u(k-1), \dots, u(k-n_y))}{a_0(u(k-1), \dots, u(k-n_y))} \quad (2.78)$$

Assim, modelos lineares, bilineares, lineares por partes, racionais, *output-affine* e polinomiais são casos particulares do NARMAX (2.12). Os modelos lineares e bilineares são casos particulares do polinomial NARMAX, enquanto que os lineares por partes apresentam alguns inconvenientes já citados anteriormente. Além disso, os modelos polinomiais NARMAX apresentam vantagem em relação aos modelos racionais e aos modelos *output-affine*, uma vez que o primeiro é linear nos parâmetros (Chen and Billings, 1989). Devido a estas razões o polinomial NARMAX será utilizado neste trabalho.

## 2.12 Validação do Modelo

A validação é o procedimento final da identificação de sistemas, sendo efetuada após a determinação da estrutura e estimação de parâmetros e tem como objetivo checar se o modelo é capaz de representar a dinâmica do sistema em questão e se o modelo é não polarizado. Um modelo polarizado pode ser capaz de prever a resposta do sistema quando usa-se a mesma massa de dados da estimação, mas o mesmo pode não ocorrer para uma massa de dados diferente (Billings and Tao, 1991).

Uma estrutura sobreparametrizada resulta em cálculos computacionais desnecessários tanto para estimar os parâmetros, quanto para utilizar o modelo estimado. Além disso, um modelo sobreparametrizado é geralmente impreciso (Åström and Wittenmark, 1990).

Na procura de um modelo “ideal” devem ser observadas as seguintes questões: se o modelo encontrado é flexível o suficiente e se o modelo é desnecessariamente complexo, ou colocando de outra forma, se o modelo é confiável do ponto de vista do propósito desejado.

### 2.12.1 Validação Estatística

Quando um método de mínimos quadrados é utilizado para estimar os parâmetros de uma determinada estrutura, obtém-se um modelo com valores mínimos para a função custo (Equação (2.26)), mas no entanto, se a estrutura escolhida for inadequada, o modelo será ineficiente em relação à representação

de dinâmicas provenientes do sistema original. Se os resíduos são autocorrelacionados ou correlacionados com qualquer dos regressores, o estimador será polarizado. A fim de detectar estes problemas, usam-se critérios estatísticos como testes de correlação e distribuição Q-quadrada. Neste trabalho serão utilizados apenas os testes de correlação.

Para sistemas lineares, as seguintes funções de correlação podem ser utilizadas:

$$\begin{aligned}\phi_{\xi\xi}(\tau) &= \delta(\tau), \\ \phi_{u\xi}(\tau) &= 0 \quad \forall \tau.\end{aligned}\tag{2.79}$$

No entanto, tais funções de correlação falham, na maioria das vezes, quando o sistema a ser identificado é não linear. Torna-se, portanto, necessária a utilização de funções de correlação não lineares (Billings and Fadzil, 1985; Billings and Tao, 1991; Billings and Voon, 1983; Billings and Zhu, 1994a), tais como:

$$\begin{aligned}\phi_{\xi\xi^2}(\tau) &= 0 \quad \forall \tau, \\ \phi_{\xi\xi u}(\tau) &= 0 \quad \forall \tau, \\ \phi_{u^2\xi}(\tau) &= 0 \quad \forall \tau, \\ \phi_{u^2\xi^2}(\tau) &= 0 \quad \forall \tau, \\ \phi_{\hat{y}\xi\xi^2}(\tau) &= 0 \quad \forall \tau, \\ \phi_{\hat{y}\xi u^2}(\tau) &= 0 \quad \forall \tau, \\ \phi_{\xi^2\xi^2}(\tau) &= \delta(\tau).\end{aligned}\tag{2.80}$$

Na prática, as esperanças matemáticas são estimadas, uma vez que  $N$  tem valor finito. Sendo assim, ao invés de verificar se as funções de correlação são nulas, estabelecem-se intervalos de confiança, nos quais tais funções devem estar limitadas. Estes intervalos são definidos por  $\pm 1,96/\sqrt{N}$ , onde  $N$  é o número de dados.

Então, no caso de sistemas não lineares, se qualquer das funções (2.79 e 2.80) estiver fora do intervalo de confiança diz-se que de alguma forma o modelo não é satisfatório. No entanto, não se pode afirmar qual termo está ausente, ou se é um termo linear ou não, nem mesmo se a falha está no modelo do ruído ou do processo, uma vez que as funções de correlação não lineares são capazes de detectar tanto problemas em termos lineares quanto em não lineares (Billings and Fadzil, 1985; Billings and Voon, 1983; Billings and Zhu, 1994a).

Em (Billings and Fadzil, 1985; Billings and Tao, 1991; Billings and Voon, 1983; Billings and Zhu, 1994b; Stoica et al., 1986) podem ser encontradas descrições mais detalhadas das funções de correlação utilizadas para validar estatisticamente o modelo e em (Billings and Zhu, 1994a) os testes de validação estatística utilizando funções de correlação são estendidos para sistemas MIMO (*Multivariable Input Multivariable Output*).

Na aplicação do modelo para predição, ou seja, utilizando-se uma massa de dados diferente da usada na estimação, pode-se através dos parâmetros calculados, estimar o valor da saída, predizendo um passo à frente, recalculando os resíduos e determinando a função de autocorrelação desses últimos. Isto permite, então escolher o modelo que apresentar menores valores dos resíduos de validação. Este procedimento pode ser justificado pelo fato do método de mínimos quadrados minimizar a função custo (equação 2.26), considerando apenas os resíduos de estimação (Stoica et al., 1986).

O critério de validação estatística garante, em princípio, apenas que não existem correlações entre os resíduos, ou entre resíduos e a entrada e assim por diante, ou seja, os resíduos são imprevisíveis e o modelo é adequado do ponto de vista estatístico.

No entanto, os modelos validados estatisticamente podem não apresentar o mesmo comportamento dinâmico que os sistemas originais (Aguirre and Billings, 1994). Desta forma é usado um outro critério de validação, a fim de verificar a fidelidade do modelo considerado, denominado validação dinâmica.

### 2.12.2 Validação Dinâmica

É entendido por validação dinâmica a comparação do comportamento dinâmico do sistema e do modelo identificado. Um exemplo são as predições do modelo, que para um modelo linear são dadas por:

$$\begin{aligned} \hat{y}_k = & +\hat{\theta}_{1,0}(1)\hat{y}_{k-1} + \cdots + \hat{\theta}_{1,0}(n_y)\hat{y}_{k-n_y} + \hat{\theta}_{0,1}(1)u_{k-1} + \cdots \\ & +\hat{\theta}_{0,1}(n_u)u_{k-n_u}. \end{aligned} \quad (2.81)$$

A recursão é inicializada com amostras da saída experimental desde  $y_1$  até  $y_{n_y}$ , e a partir daí são utilizados apenas valores da saída simulada.

Características dinâmicas provenientes do sistema original não recuperadas pelo modelo, bem como regimes espúrios apresentados pelo mesmo, causados por deficiências na estrutura do modelo, pobre estimação de parâmetros ou de tempo morto, são problemas facilmente detectados (porém não diagnosticados) através deste teste, uma vez que os erros são cumulativos.

Em alguns casos podem ser usados expoentes de Lyapunov, diagramas de bifurcação, mapas e seções de Poincaré e dimensão de correlação (Aguirre and Billings, 1994).

## 2.13 Algoritmo Recursivo com Fator de Esquecimento

Enquanto os algoritmos de estimação em batelada processam todas as amostras da massa de dados simultaneamente, obtendo uma única estimativa dos parâmetros do modelo, os métodos de estimação recursiva processam os dados experimentais seqüencialmente, atualizando os parâmetros do modelo a cada período de amostragem (Ljung, 1987).

Tanto a identificação em batelada, quanto a recursiva são largamente utilizadas. A escolha do método depende do tipo de problema e do sistema a ser identificado.

Para sistemas invariantes no tempo ambas abordagens podem ser usadas. Por outro lado, quando o sistema é variante no tempo, a estimação recursiva é, na maioria dos casos, mais adequada. Isto ocorre porque se processa cada amostra de uma vez e, embora os valores anteriores tenham influência no cálculo atual, a quantidade de informação anterior utilizada na estimação atual pode ser ponderada. Desta forma, pode-se atribuir maior peso às amostras atuais do que às passadas, proporcionando uma estimativa mais confiável aos dados experimentais.

Em muitas aplicações reais, é desejável ter um modelo disponível *on-line*. No caso de sistemas invariantes no tempo, pode ser utilizado para detecção de falhas. No caso de sistemas variantes no tempo, a utilização da identificação *on-line* é mais ampla, como por exemplo, predição e controle em tempo real. Sendo assim, nas aplicações *on-line*, pode-se utilizar um algoritmo recursivo para atualizar os parâmetros do modelo de acordo com as variações no sistema, à medida que os dados são disponibilizados (Ljung, 1987).

Desta forma, a identificação em batelada é utilizada quando a estimativa só é requerida uma vez, ou é única para toda a massa de dados (Norton, 1986), sendo assim, sua aplicação mais comum é para estimação *off-line* de parâmetros. Ao passo que, a identificação recursiva pode ser utilizada tanto *on-line*, quanto *off-line*, para sistemas variantes ou invariantes no tempo, uma vez que permite a atualização dos parâmetros do modelo. No caso monovariável, a identificação recursiva tem ainda como vantagem a não necessidade de cálculo da inversa da matriz normal, já que a estimação é realizada passo

a passo.

Embora a estimação recursiva seja muito utilizada para encontrar modelos capazes de representar sistemas variantes no tempo, a detecção da estrutura destes modelos através de algoritmos recursivos é muitas vezes difícil e desnecessária (Pröll and Karim, 1994). Desta forma, é muito comum escolher uma estrutura de modelo utilizando a identificação em batelada e estimar os parâmetros do mesmo de forma recursiva.

Algumas aplicações da estimação recursiva são (Norton, 1986):

- Identificação *on-line* para por exemplo, predição e controle em tempo real,
- Identificação *off-line*, para sistemas variantes ou invariantes no tempo. No caso de sistema variantes no tempo pode-se acompanhar as variações da dinâmica do sistema, através da inserção de um fator de esquecimento no algoritmo.

Uma descrição do processo de identificação recursiva está indicada na Figura 2.2.

A matriz de covariância do estimador,  $\mathbf{P}_k$ <sup>9</sup>, é atualizada de acordo com a equação (2.82), e indica o quão confiável é a estimativa no instante  $k$  (Norton, 1986).

$$\mathbf{P}_k = \mathbf{P}_{k-1} - \frac{\mathbf{P}_{k-1} \mathbf{X}_k \mathbf{X}_k^T \mathbf{P}_{k-1}}{1 + \mathbf{X}_k^T \mathbf{P}_{k-1} \mathbf{X}_k} . \quad (2.82)$$

Os parâmetros do modelo serão ajustados como se segue:

$$\hat{\Theta}_k = \hat{\Theta}_{k-1} + \mathbf{G}_k v_k , \quad (2.83)$$

onde  $\mathbf{G}_k = \mathbf{P}_k \mathbf{X}_k$  e  $v_k = (y_k - \mathbf{X}_k^T \hat{\Theta}_{k-1})$ , são o ganho e a inovação, respectivamente.

A saída estimada será:

$$\hat{y}_k = \mathbf{X}_k^T \hat{\Theta}_k . \quad (2.84)$$

As equações (2.82) e (2.83) não conseguem, na maioria dos casos, acompanhar variações no sistema ao longo do tempo. O problema é que o ganho  $\mathbf{G}_k$  torna-se pequeno à medida que o tempo aumenta. Isto ocorre porque

---

<sup>9</sup>Nesta seção  $\mathbf{P}_k$  indica matriz de covariância do estimador na iteração  $k$ .

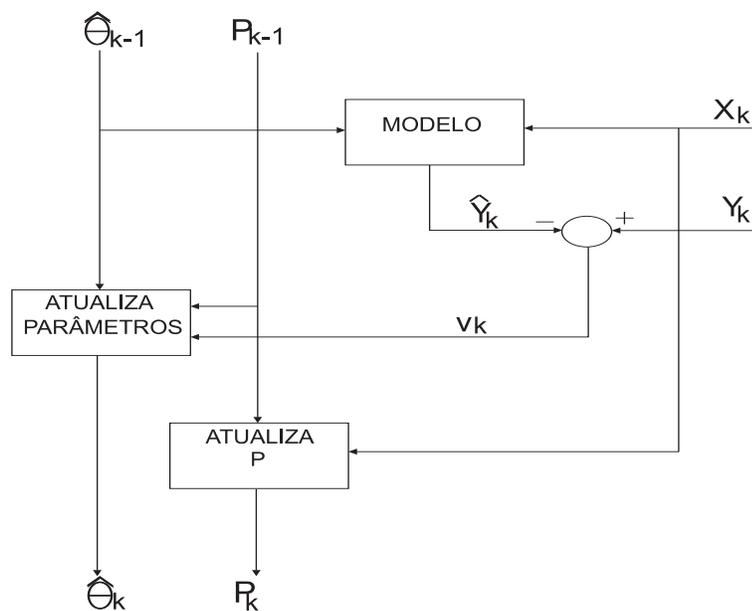


Figura 2.2: Diagrama de blocos ilustrando o processo de ajuste de parâmetros na estimação recursiva.  $P_{k-1}$  é a matriz de covariância de  $\hat{\Theta}_{k-1}$ ,  $X_k$  é o vetor de regressão no instante  $k$ ,  $y_k$  é a saída medida,  $\hat{y}_k$  é a saída estimada e  $v_k$  é a inovação.

a matriz de covariância dos parâmetros  $\mathbf{P}_k$  diminui indicando um aumento de “confiança” nas estimativas com o passar do tempo. Desta forma, alterações ocorridas após um longo intervalo de tempo quase não são percebidas pelo estimador, pois mesmo que  $(y_k - \mathbf{X}_k^T \hat{\boldsymbol{\Theta}}_{k-1})$  seja grande,  $\mathbf{G}_k$  será pequeno. Deve-se portanto, se necessário, alterar  $\mathbf{P}_k$  de forma a aumentar o ganho  $\mathbf{G}_k$ , permitindo assim perceber desvios nos valores dos parâmetros estimados decorrentes de variações de dinâmica (Norton, 1986). Uma forma de alterar  $\mathbf{P}_k$  é atribuir um maior peso às estimativas recentes do que às passadas. Modificando a matriz de covariância  $\mathbf{P}_k$  e o ganho  $\mathbf{G}_k$ , obtem-se:

$$\mathbf{P}_k = \mathbf{P}_{k-1} - \frac{w_k \mathbf{P}_{k-1} \mathbf{X}_k \mathbf{X}_k^T \mathbf{P}_{k-1}}{1 + w_k \mathbf{X}_k^T \mathbf{P}_{k-1} \mathbf{X}_k}, \quad (2.85)$$

e

$$\mathbf{G}_k = w_k \mathbf{P}_k \mathbf{X}_k. \quad (2.86)$$

Substituindo  $\mathbf{P}_k$  por  $\mathbf{P}'_k = w_k \mathbf{P}_k$  e incluindo um fator de esquecimento  $\beta_k = w_{k-1}/w_k$  em (2.85), tem-se:

$$\mathbf{P}'_k = \frac{1}{\beta_k} \left( \mathbf{P}'_{k-1} - \frac{\mathbf{P}'_{k-1} \mathbf{X}_k \mathbf{X}_k^T \mathbf{P}'_{k-1}}{\beta_k + \mathbf{X}_k^T \mathbf{P}'_{k-1} \mathbf{X}_k} \right), \quad (2.87)$$

onde  $\beta_k$  é um número que indica o montante de informação relevante existente na última amostra (Ljung, 1987), sendo denominado o fator de esquecimento e que normalmente, varia entre 0,95 e 0,99 (Norton, 1986). Neste caso, o ganho  $\mathbf{G}_k$  indicado na equação (2.83) será igual a  $\mathbf{G}_k = \mathbf{P}'_k \mathbf{X}_k$ .

Na maioria dos casos, a matriz de covariância é inicializada com valores altos (Åström and Wittenmark, 1995; Ljung, 1987), indicando baixa confiança nos parâmetros iniciais. No entanto, como na estimação recursiva é permitido utilizar informação *a priori* para inicializar os vetores de parâmetros, pode-se estimar os parâmetros utilizando um método de mínimos quadrados em batelada e utilizar tais valores para inicializar a recursão (Norton, 1986). Isso evita, em alguns casos, convergência lenta dos parâmetros estimados.

## 2.14 Comentários

Sistemas não lineares são, na maioria dos casos, melhor representados por modelos não lineares, por incorporar as não linearidades intrínsecas aos sistemas originais, bem como permitir a reprodução e análise de fenômenos que

ocorrem nos referidos sistemas, como por exemplo, bifurcação e caos (Aguirre and Billings, 1995a).

Mesmo quando o objetivo do modelo é controlar um sistema não linear em sua região de operação linear, ainda constitui vantagem determinar um modelo não linear, pois o mesmo será válido para uma faixa de operação mais ampla (Billings and Fadzil, 1985).

A escolha de uma estrutura para representar um determinado sistema está intimamente ligada ao objetivo final do modelo. Sistemas que apresentam características estáticas não lineares, são necessariamente não lineares. Contudo, a utilização de um modelo não linear para representar a dinâmica do sistema dependerá das características dinâmicas do mesmo, como por exemplo, variações das constantes de tempo. Pode-se dizer então que, tais sistemas podem ser representados por modelos lineares ou não, dependendo das características do sistema, do grau de confiabilidade exigido para o modelo, e do propósito para o qual o modelo foi encontrado. Se o objetivo é encontrar um modelo que melhor represente a dinâmica do sistema original, um modelo linear pode, em alguns casos, ser utilizado. Se no entanto, o objetivo é recuperar as não linearidades estáticas presentes no sistema, um modelo não linear é mais adequado.

## Capítulo 3

# Identificação em Batelada

### 3.1 Introdução

Neste capítulo será feita a identificação de um sistema onde deseja-se modelar a relação entre a tensão elétrica aplicada a um ferro de solda e a temperatura correspondente, de forma a tornar viável a estimação da característica estática não linear do sistema. Além disso, deseja-se que o modelo escolhido seja satisfatório dinamicamente, ou seja, o modelo deve ser capaz de representar a dinâmica do sistema e estimar a característica estática não linear.

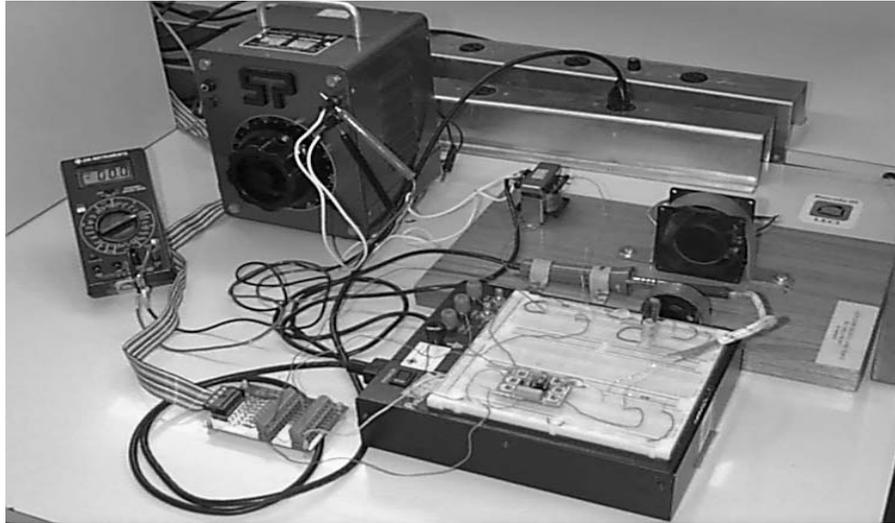
O capítulo está organizado como se segue: na seção 3.2 é feita a descrição do sistema a ser identificado, na seção 3.3 são apresentados os resultados dos testes realizados para identificar o sistema, na seção 3.4 são realizados testes para detecção de não linearidades nos dados, na seção 3.5 define-se a taxa de amostragem para detecção de estrutura e estimação de parâmetros, na seção 3.5 é feita a análise de agrupamentos de termos e na seguinte utiliza-se um critério de informação a fim de ajudar na detecção da estrutura do modelo. Finalmente, são apresentados alguns modelos para representar o sistema.

### 3.2 Descrição do Sistema

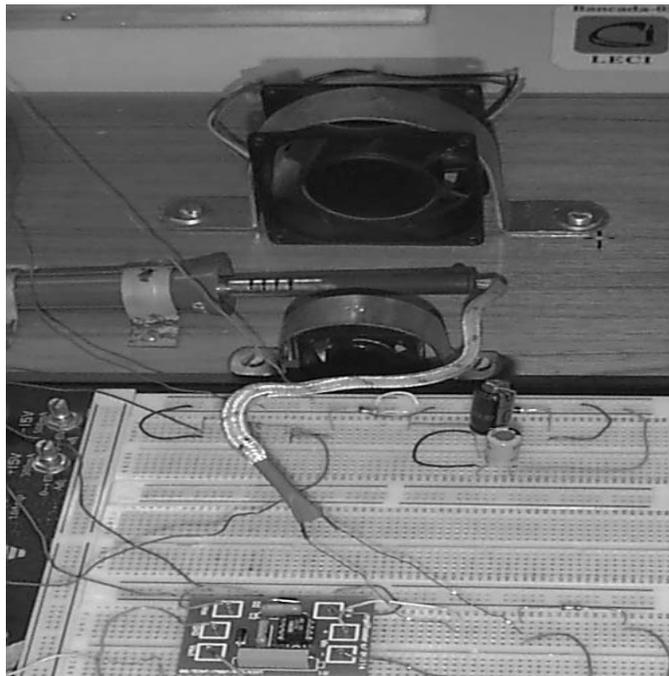
O sistema utilizado é mostrado na Figura 3.1 e seu diagrama de blocos na Figura 3.2. A entrada e saída do sistema são, respectivamente, a tensão elétrica no divisor de tensão<sup>1</sup> e a tensão de saída do circuito amplificador. Ambos sinais foram coletados por uma placa de aquisição de dados (PCL 711s).

---

<sup>1</sup>O divisor de tensão é de tal forma que uma variação de 0–136V no varivolt corresponda a uma variação de 0 – 5V na entrada do sistema.



(a)



(b)

Figura 3.1: Sistema. (a) Detalhe 1 (b) Detalhe 2.

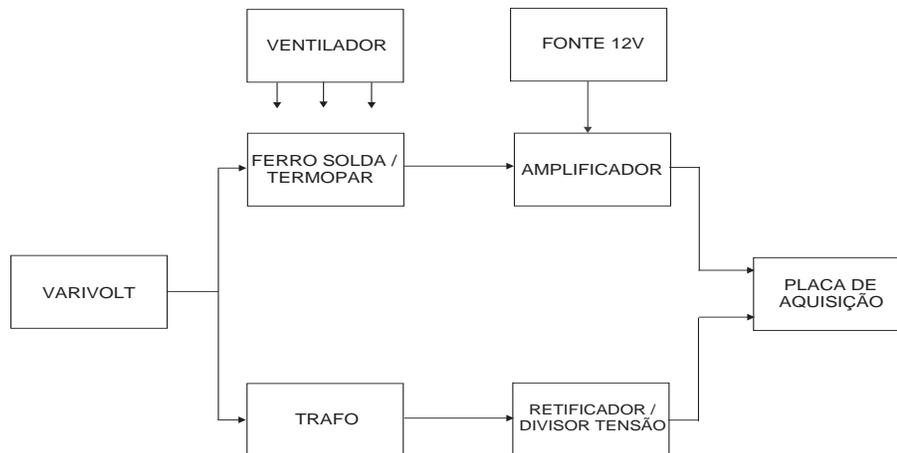


Figura 3.2: Diagrama de Blocos do Sistema.

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<sup>2</sup>. Portanto, o sinal lido pela placa não é o sinal de tensão aplicado na entrada do sistema, mas é representativo do mesmo.

Usando-se um varivolt aplica-se uma tensão na faixa de 0 – 136V no ferro de solda. Esta tensão é reduzida pelo transformador passando a estar na faixa de 0 – 18V, que será retificada. A tensão contínua é aplicada ao divisor de tensão que irá ajustá-la a uma faixa de 0 – 5V.

As variações de temperatura ocorridas no ferro de solda, decorrentes das respectivas variações de tensão na sua entrada, são medidas pelo termopar acoplado ao ferro de solda. As variações de tensão nos terminais do termopar são amplificadas usando-se um amplificador de instrumentação para uma faixa de 0 – 4V e enviadas para a placa de aquisição.

O ventilador alimentado em 127V é utilizado para resfriar o ferro de solda, variando o ganho estático do sistema, conforme explicado com mais detalhes no restante do capítulo.

<sup>2</sup>Limites de tensão para a placa de aquisição utilizada: 0 – 5V ou 0 – 10V. Sendo selecionada a faixa 0 – 5V.

### 3.3 Testes

A fim de coletar os dados para a identificação, foram realizados os seguintes testes com o ventilador ligado e desligado:

- Testes Preliminares em Degrau;
- Testes Estáticos;
- Testes Dinâmicos.

#### 3.3.1 Testes em Degrau

Os testes em degrau têm por finalidade estimar as constantes de tempo dominantes do sistema. Para o degrau positivo obteve-se um valor igual a 1,81 min e para o degrau negativo o obtido foi 3,33 min. Desta forma, os testes estáticos serão realizados com degraus positivos, aguardando um período de 11 min para permitir que o sistema estabilize.

#### 3.3.2 Testes Estáticos

Os testes estáticos fornecem a curva estática, curva em regime permanente, do sistema. Dentre os testes realizados foram selecionados apenas três, sendo dois com o ventilador desligado e um com o mesmo ligado. As características estáticas obtidas estão representadas pelas Figuras 3.3, 3.4 e 3.5, onde as grandezas estão indicadas em pu. Sendo 1pu correspondente a 136V na tensão de entrada e a 998,51°C na temperatura do ferro de solda.

Cada teste teve duração total de três horas. Para cada ponto no gráfico<sup>3</sup>, variou-se a tensão no varivolt e, após estabilizada a temperatura, realizou-se a medição da tensão de saída do amplificador.

O ganho do sistema depende da temperatura ambiente. Uma variação de apenas 4 (quatro) graus Celsius na mesma provocou variações sensíveis na temperatura do ferro de solda. Isso pode ser observado comparando-se as Figuras 3.3 e 3.4. Sendo a tensão de entrada igual a 0,92pu, a temperatura do ferro de solda correspondente é igual a 0,72pu para temperatura ambiente igual a 23°C e 0,76pu para temperatura ambiente igual a 27°C. Este fato vem incentivar a idéia de estimação recursiva dos parâmetros do modelo, pois desta forma o modelo será capaz de acompanhar as alterações ocorridas no ambiente.

---

<sup>3</sup>Vale salientar que os dados apresentados são discretos, sendo que estes foram ligados no gráfico para facilitar a sua visualização.

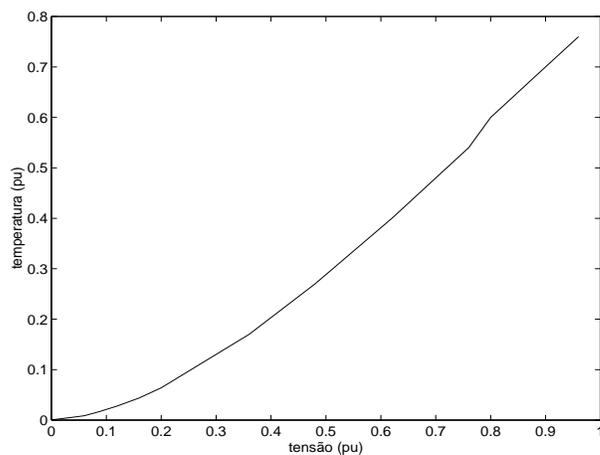


Figura 3.3: Característica estática obtida para o sistema com o ventilador desligado e temperatura ambiente igual a 23°C – Teste 02.

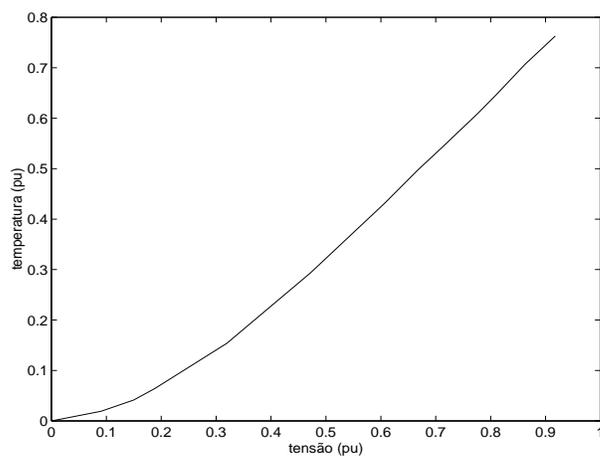


Figura 3.4: Característica estática obtida para o sistema com o ventilador desligado e temperatura ambiente igual a 27°C – Teste 07.

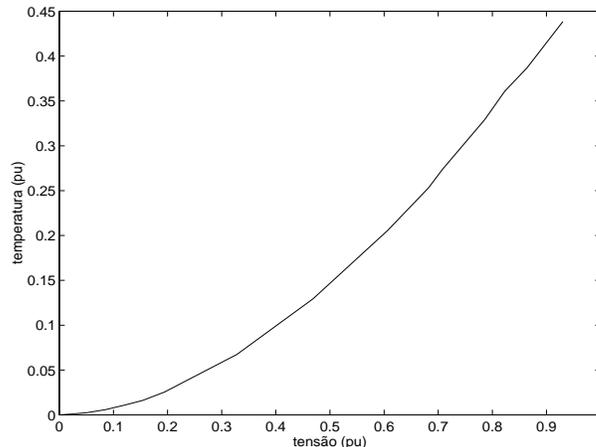


Figura 3.5: Característica estática obtida para o sistema com o ventilador ligado e temperatura ambiente igual a  $26^{\circ}\text{C}$  – Teste 08.

### 3.3.3 Testes Dinâmicos

Os testes dinâmicos têm por finalidade fornecer a massa de dados que será utilizada na identificação do modelo. Foram realizados dois testes dinâmicos, o primeiro com o ventilador desligado e o segundo com o mesmo alimentado em 127V. Trechos das respectivas massas de dados estão indicadas nas Figuras 3.6 e 3.7. Tais testes foram realizados a uma temperatura ambiente de 24 graus Celsius por um período de quatro horas, com taxa de amostragem igual a 6s, resultando em um total de 2510 pontos para o primeiro teste e 2520 para o segundo, persistindo para ambos os casos um número aproximado de 250 patamares de 1 min (aproximadamente 10 amostras por patamar).

As autocorrelações dos sinais de entrada indicadas nas Figuras 3.8 e 3.9 correspondem aos testes dinâmicos com o ventilador desligado e ligado, respectivamente. Tais figuras revelam o caráter aleatório dos sinais usados. A rigor, para que tais sinais tivessem funções de autocorrelação tipicamente de sinais aleatórios, eles precisariam ser dizimados de um fator de 10.

## 3.4 Detecção de Não Linearidades

Como já mencionado anteriormente, um dos primeiros passos para a identificação de um sistema, é verificar se o mesmo exibe ou não dinâmica não linear, o que sugerirá a escolha de um modelo não linear. Desta forma, pode-se uti-

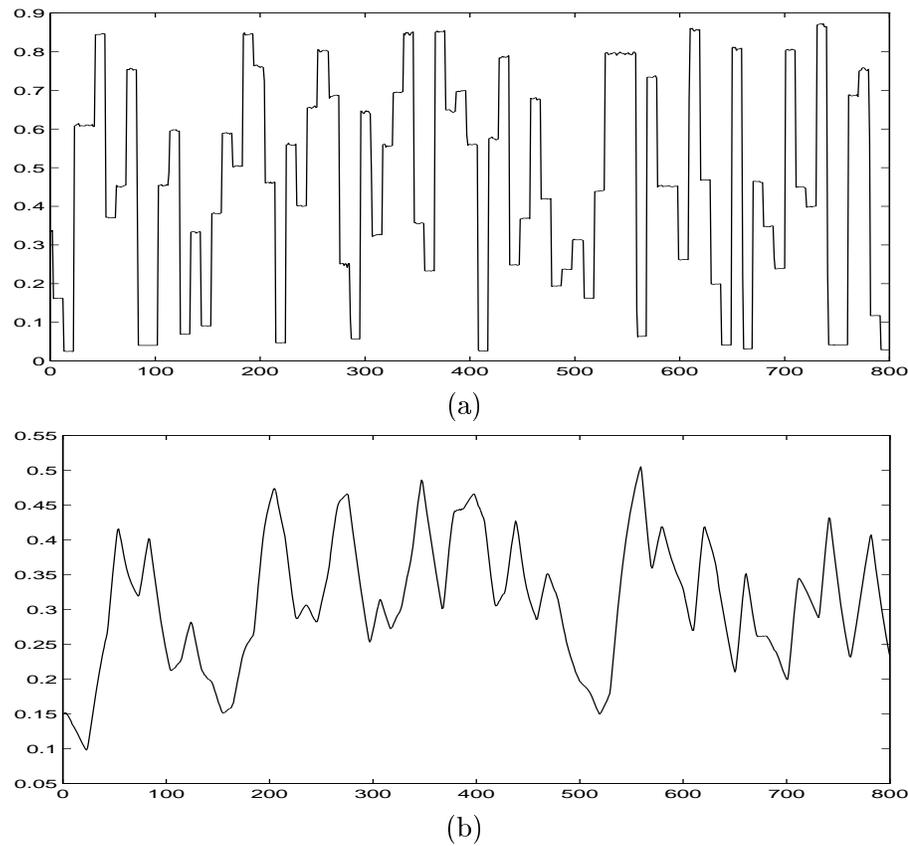


Figura 3.6: Janela de 1 a 800 amostras da massa de dados  $\langle din2 \rangle$ . (a) Dados de entrada (em pu) (b) Dados de saída (em pu) – ventilador desligado.

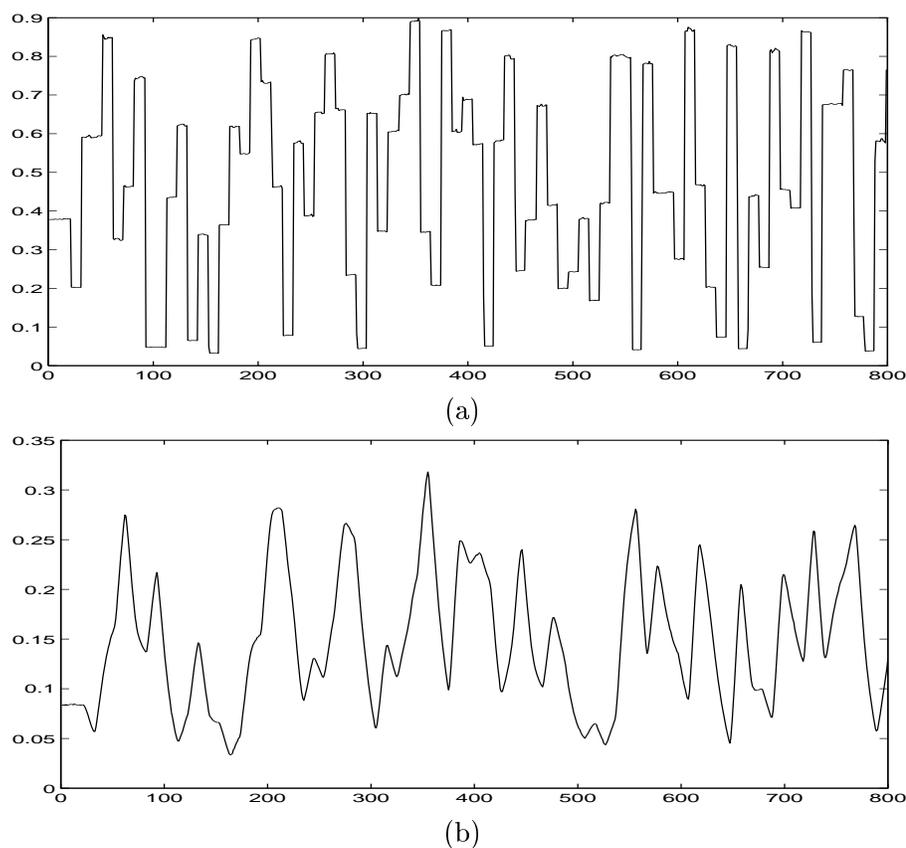


Figura 3.7: Janela de 1 a 800 amostras da massa de dados  $\langle \text{din3} \rangle$ . (a) Dados de entrada (em pu) (b) Dados de saída (em pu) – ventilador alimentado em 127V.

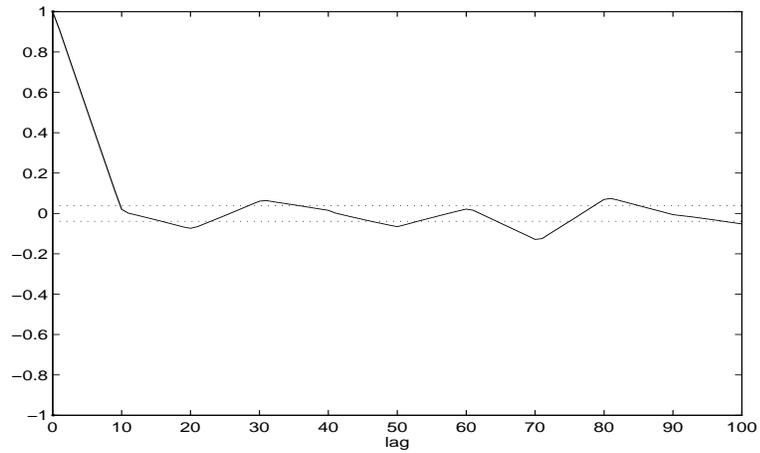


Figura 3.8: Autocorrelação do sinal de entrada da massa de dados  $\langle din2 \rangle$ .

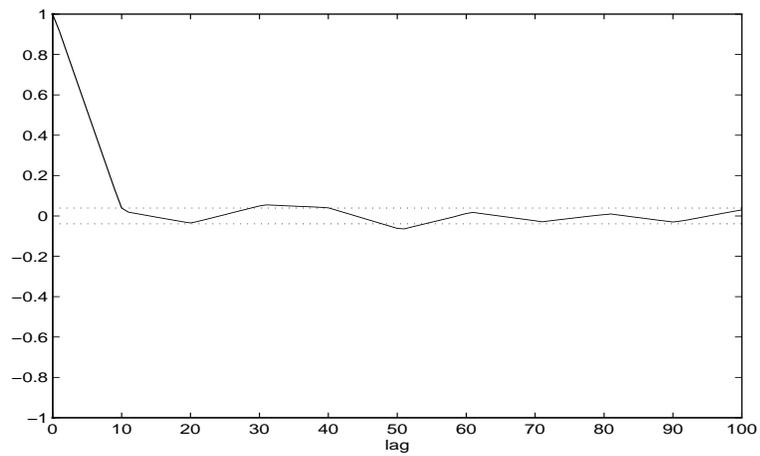


Figura 3.9: Autocorrelação do sinal de entrada da massa de dados  $\langle din3 \rangle$ .

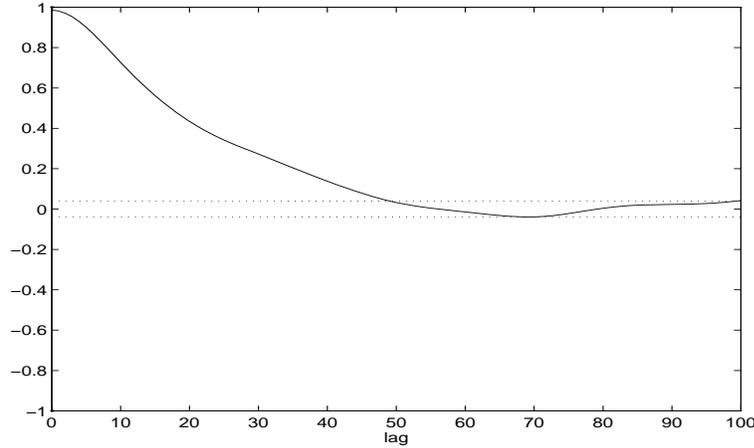


Figura 3.10: Autocorrelação não linear (3.2) do sinal de saída para verificar a existência de possíveis não linearidades na massa de dados  $\langle din2 \rangle$ .

lizar as funções de correlação definidas na seção 2.4 para detectar presença de não linearidades nos dados experimentais.

Seja a seguinte função de correlação (Haber, 1985):

$$\phi_{xy'}(\tau) = E[x(t + \tau)y'(t)] \quad (3.1)$$

onde  $x(k) = [u'^2(k) - \overline{u'^2(k)}]/\sigma^2(u'^2(k))$ . Se o sistema for linear  $\phi_{xy'}(\tau) = 0, \forall \tau$ .

De forma similar, a função de correlação (Haber, 1985) indicada em (3.2) também pode ser utilizada para detectar presença de não linearidades nos dados:

$$\phi_{vy'}(\tau) = E[v(t + \tau)y'(t)] \quad (3.2)$$

onde  $v = [y'^2 - \overline{y'^2(k)}]/\sigma(y'^2(k))$ . Neste caso, o sistema será linear se  $\phi_{vy'}(\tau)$  for igual a zero  $\forall \tau$ .

Observando agora as Figuras 3.10 e 3.12, para a massa de dados  $\langle din2 \rangle$ , e as Figuras 3.11 e 3.13, para massa de dados  $\langle din3 \rangle$ , e conforme descrito em (Billings and Voon, 1983; Haber, 1985), os valores fora do intervalo de confiança indicam a presença de não linearidades nos dados.

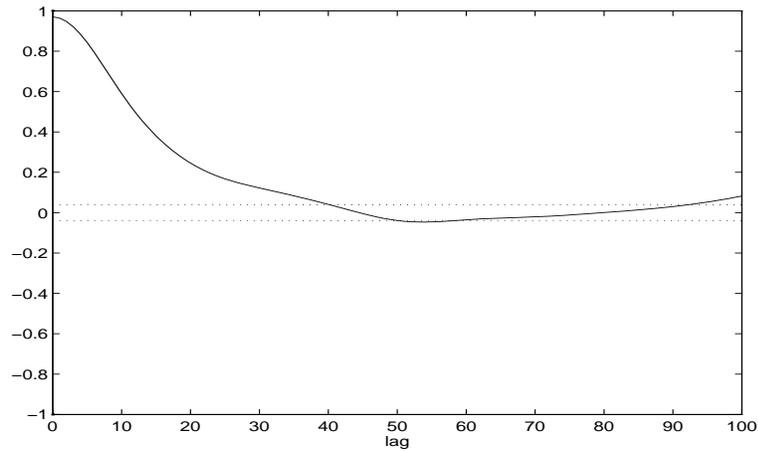


Figura 3.11: Autocorrelação não linear (3.2) do sinal de saída para verificar a existência de possíveis não linearidades na massa de dados  $\langle din3 \rangle$ .

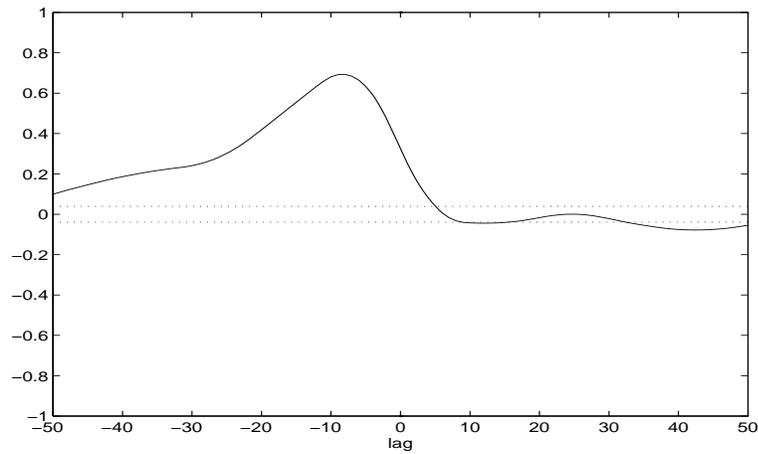


Figura 3.12: Correlação cruzada não linear (3.1) entre o sinal de saída e o de entrada para verificar a existência de possíveis não linearidades na massa de dados  $\langle din2 \rangle$ .

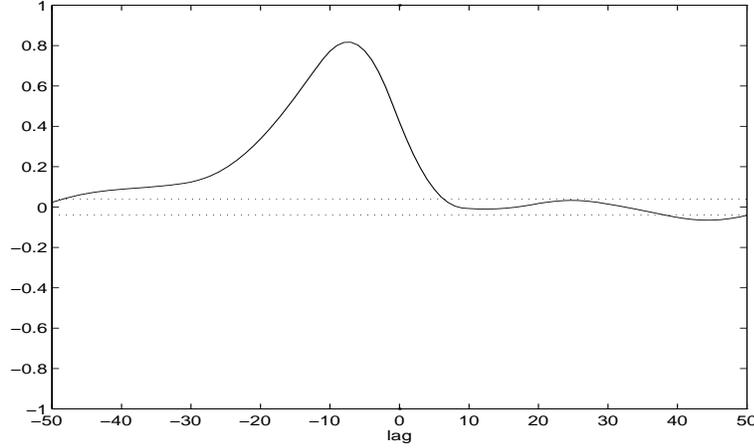


Figura 3.13: Correlação cruzada não linear (3.1) entre o sinal de saída e o de entrada para verificar a existência de possíveis não linearidades na massa de dados  $\langle \text{din3} \rangle$ .

### 3.5 Taxa de Amostragem

A determinação da taxa de amostragem para a detecção de estrutura e estimação de parâmetros é realizada com base nos testes de correlação linear e não linear (Aguirre, 1995). Tais funções estão indicadas nas Figuras 3.14 e 3.15 para a massa de dados  $\langle \text{din2} \rangle$  e nas Figuras 3.16 e 3.17 para a massa de dados  $\langle \text{din3} \rangle$ .

Sendo o instante do primeiro mínimo para a massa  $\langle \text{din2} \rangle$  igual a  $\tau_m \approx 70 * 6 = 420s$ . A taxa de amostragem de trabalho<sup>4</sup> deve ser:

$$\frac{\tau_m}{25} \leq T_s \leq \frac{\tau_m}{5} \Rightarrow 16,8 \leq T_s \leq 84 . \quad (3.3)$$

Para a massa  $\langle \text{din3} \rangle$  o instante do primeiro mínimo igual a  $\tau_m \approx 53,33 * 6 = 320s$ . A taxa de amostragem de trabalho deve ser:

$$\frac{\tau_m}{25} \leq T_s \leq \frac{\tau_m}{5} \Rightarrow 12,8 \leq T_s \leq 64 . \quad (3.4)$$

Para fins práticos será usada a mesma taxa de amostragem tanto para a massa  $\langle \text{din2} \rangle$  quanto para a  $\langle \text{din3} \rangle$ , que será igual a 12s. No entanto,

<sup>4</sup>Taxa de amostragem de trabalho refere-se à utilizada para detecção de estrutura e estimação dos parâmetros do modelo.

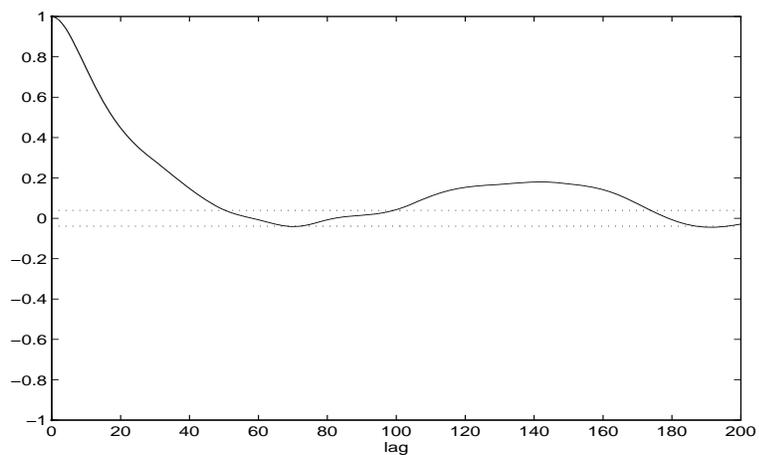


Figura 3.14: Autocorrelação linear (2.1) do sinal de saída da massa de dados  $\langle din2 \rangle$  utilizada para determinar a taxa de amostragem.

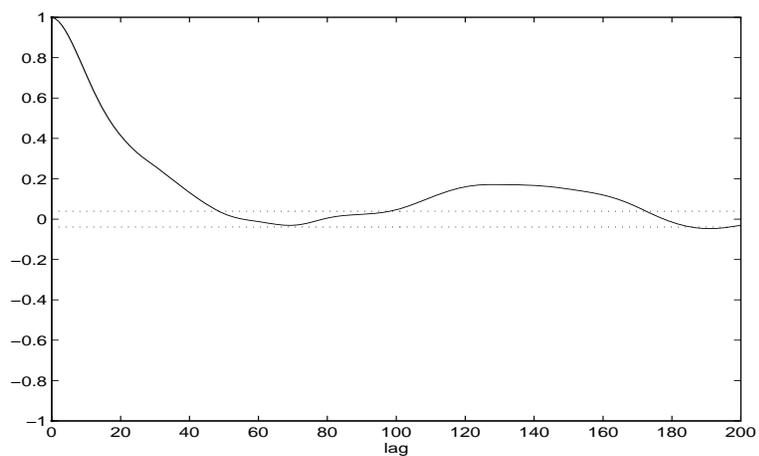


Figura 3.15: Autocorrelação não linear (2.2) do sinal de saída da massa de dados  $\langle din2 \rangle$  utilizada para determinar a taxa de amostragem.

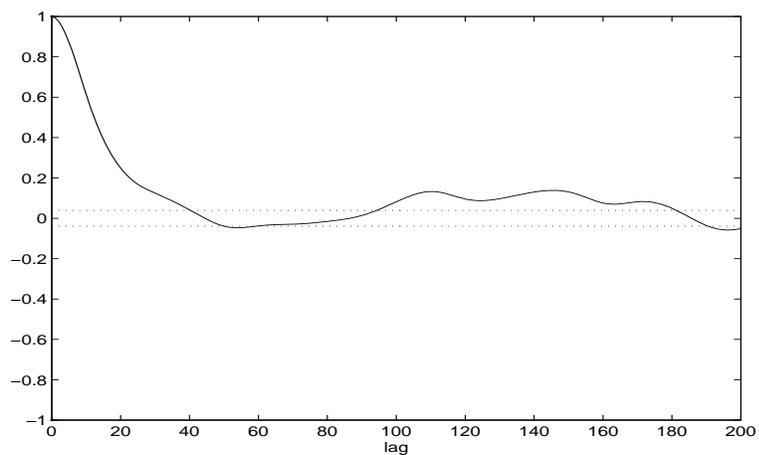


Figura 3.16: Autocorrelação linear (2.1) do sinal de saída da massa de dados  $\langle din3 \rangle$  utilizada para determinar a taxa de amostragem.

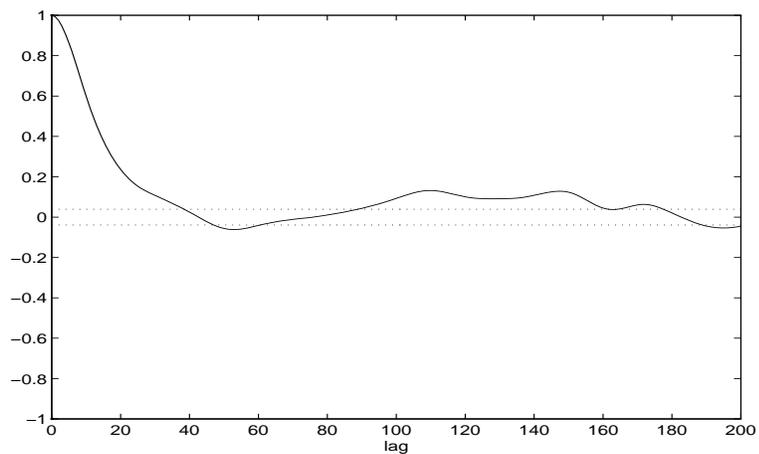


Figura 3.17: Autocorrelação não linear (2.2) do sinal de saída da massa de dados  $\langle din3 \rangle$  utilizada para determinar a taxa de amostragem.

é importante salientar que, procedendo desta maneira, a massa  $\langle din2 \rangle$  ainda estará um pouco superamostrada.

Baseando-se na taxa de amostragem definida acima, os dados de entrada e saída foram dizimados, passando agora para uma taxa de amostragem de 12s, equivalente a cinco amostras por patamar, o que resulta em 1254 pontos para identificação. As massas de dados dizimadas para os testes dinâmicos com o ventilador desligado e ligado estão representadas nas Figuras 3.18 e 3.19, respectivamente.

### 3.6 Análise de Agrupamentos de Termos e Coeficientes de Agrupamentos

Um agrupamento de termos é definido como o conjunto de todos os agrupamentos do mesmo tipo e seus coeficientes como a soma de todos os coeficientes do agrupamento (Aguirre, 1994b).

A análise de agrupamentos de termos permite, em alguns casos, detectar agrupamentos espúrios. Um agrupamento espúrio não deve ser incluído no modelo, uma vez que tal agrupamento tende a aumentar a complexidade do modelo, podendo causar problemas como mal condicionamento numérico, além de gerar, em muitos casos, comportamentos espúrios não existentes no sistema original (Aguirre and Billings, 1995a).

Após terem sido determinados os agrupamentos candidatos ao modelo, de acordo com a ordem e grau de não linearidade pré-definidos, são traçados gráficos que permitirão fazer a análise de agrupamentos de termos, detectando agrupamentos possivelmente espúrios. Tais gráficos ilustram as variações dos coeficientes de agrupamentos de termos à medida que o número de termos do modelo é alterado.

Com base nos gráficos de coeficientes de agrupamentos em função do número de termos, identifica-se um agrupamento espúrio quando:

- seu coeficiente possui valor irrelevante comparado aos demais,
- quando o mesmo oscila em torno de zero à medida que o número de termos aumenta,
- quando o surgimento do referido agrupamento ocorre somente para um número elevado de termos.

A primeira condição deve-se ao fato de que alguns termos podem ter coeficientes significantes, mas cancelarem-se entre si e a segunda que um coeficiente de agrupamento oscilatório revela inconsistência dos parâmetros

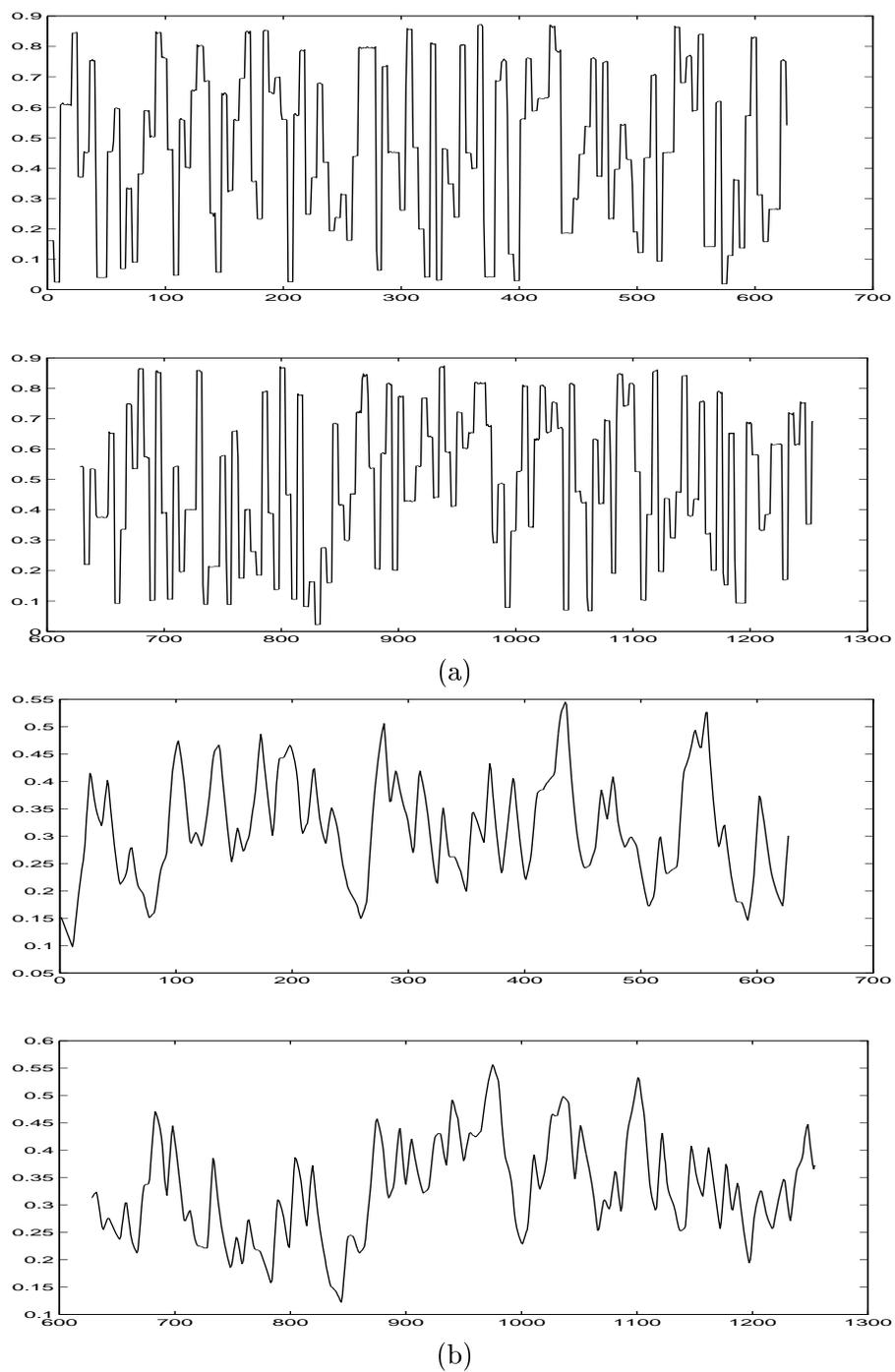
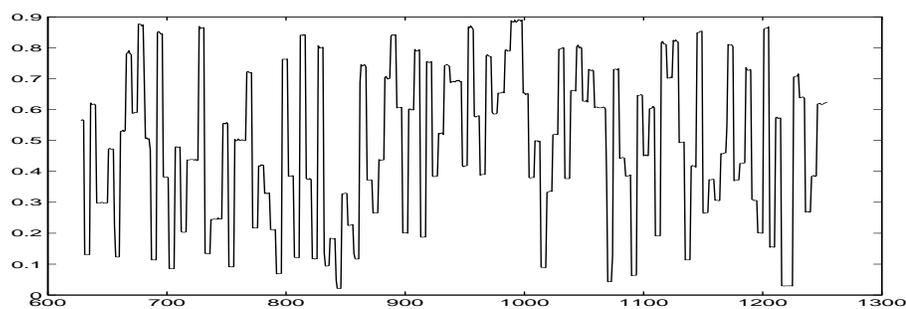
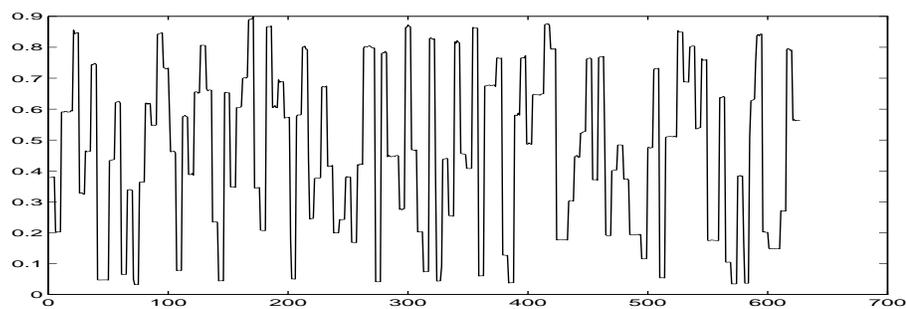
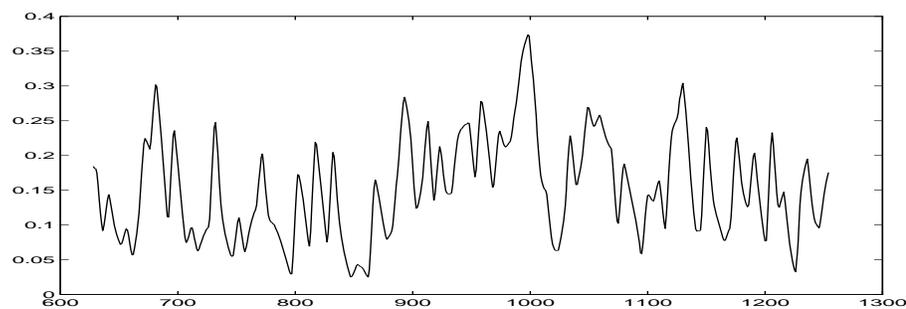
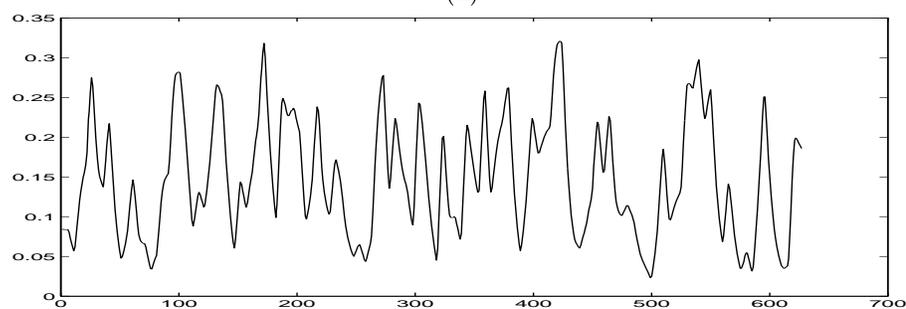


Figura 3.18: Massa de dados  $\langle din2 \rangle$  dizimada de 2. (a) Dados de entrada (em pu) (b) Dados de saída (em pu).



(a)



(b)

Figura 3.19: Massa de dados  $\langle din3 \rangle$  dizimada de 2. (a) Dados de entrada (em pu) (b) Dados de saída (em pu).

estimados. A terceira, por sua vez, vem evitar problemas de mal condicionamento numérico, já que as chances de ocorrerem tais problemas aumentam à medida que o número de termos aumenta (Aguirre, 1994b).

A análise de agrupamentos de termos tanto para a massa de dados  $\langle \text{din2} \rangle$ , quanto para  $\langle \text{din3} \rangle$ , Figuras 3.20 e 3.21, foi realizada após definido o fundo de escala que é igual a cinco, evitando prováveis erros de análise referentes à ordem dos agrupamentos.

Análises dos agrupamentos de termos das massas de dados  $\langle \text{din2} \rangle$ , Figura 3.20-*a, b, e*,  $\langle \text{din3} \rangle$ , Figura 3.21-*a, b, e*, sugerem que os agrupamentos  $\Omega_y$ ,  $\Omega_{y^2}$ ,  $\Omega_{u^2}$  sejam efetivos, uma vez que os módulos de seus respectivos coeficientes não são desprezíveis se comparados aos demais, além de que não ocorre troca de sinal para tais coeficientes. No entanto, se observadas as Figuras 3.20-*j* e 3.21-*j*, percebe-se que o termo constante só surge após o 18<sup>o</sup> e 32<sup>o</sup> termo, respectivamente, podendo ser considerado espúrio.

Análise da Figura 3.20-*c, h* sugerem que os agrupamentos  $\Omega_{y^3}$  e  $\Omega_{u^2y}$  sejam espúrios, uma vez que só aparecem após o 19<sup>o</sup> e 25<sup>o</sup> termo, respectivamente.

Analisando agora as variações dos coeficientes de agrupamentos para a massa de dados  $\langle \text{din3} \rangle$ , Figura 3.21-*d, i*, observa-se que o agrupamento  $\Omega_u$  só surge após o 17<sup>o</sup> termo e o agrupamento  $\Omega_{uy^2}$  após o 25<sup>o</sup> termo. Desta forma, a análise de agrupamentos sugere que, no caso da massa de dados  $\langle \text{din3} \rangle$ , os agrupamentos  $\Omega_u$  e  $\Omega_{uy^2}$  sejam espúrios.

Com base nas análises de agrupamentos descritas aqui, verificou-se que alguns agrupamentos, tais como  $\Omega_{y^3}$  e  $\Omega_{u^2y}$ , podem ser espúrios para uma massa de dados e efetivos para a outra. Desta forma, é fácil verificar que a importância de um determinado agrupamento está relacionada com a informação contida nos dados, ou seja, podem existir dinâmicas no sistema não excitadas pela entrada utilizada por um determinado teste que, no entanto, estão presentes em uma outra massa de dados. Isto pode acarretar em resultados diferentes na análise de agrupamento de termos, com a consequência de se ter um modelo com estrutura um pouco mais complexa do que o necessário.

Observa-se também que a análise de agrupamentos de termos irá auxiliar na escolha do grau de não linearidade, ou seja, se todos os agrupamentos de um mesmo grau forem considerados espúrios, têm-se uma indicação de que o grau de não linearidade escolhido não é adequado.

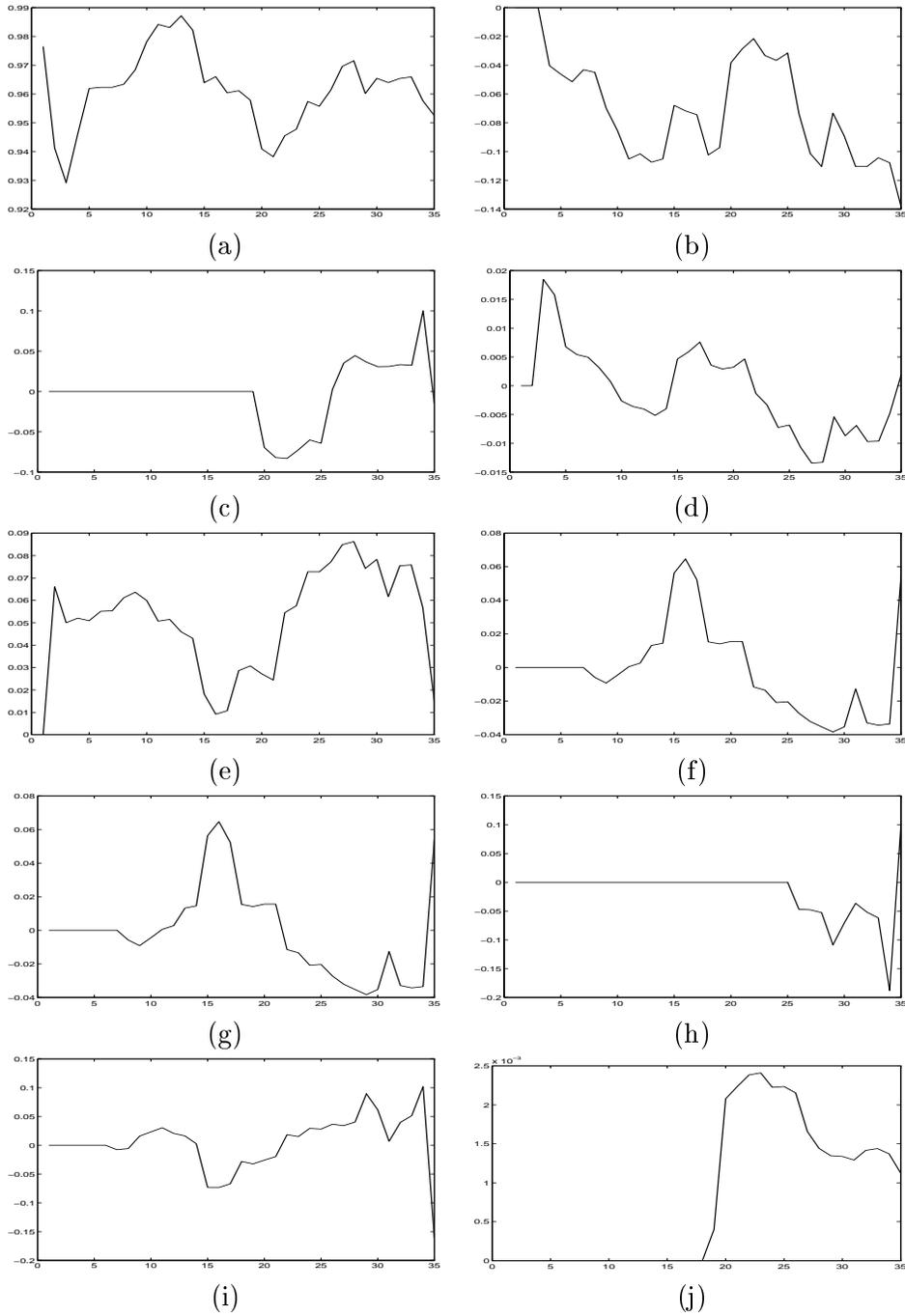


Figura 3.20: Coeficientes de agrupamentos em função do número de termos do modelo para a massa de dados  $\langle din2 \rangle$ : (a)  $\Sigma_y$  (b)  $\Sigma_{y^2}$  (c)  $\Sigma_{y^3}$  (d)  $\Sigma_u$  (e)  $\Sigma_{u^2}$  (f)  $\Sigma_{u^3}$  (g)  $\Sigma_{uy}$  (h)  $\Sigma_{u^2y}$  (i)  $\Sigma_{uy^2}$  (j)  $\Sigma_0$  (O valor máximo para este coeficiente é 0,0025).

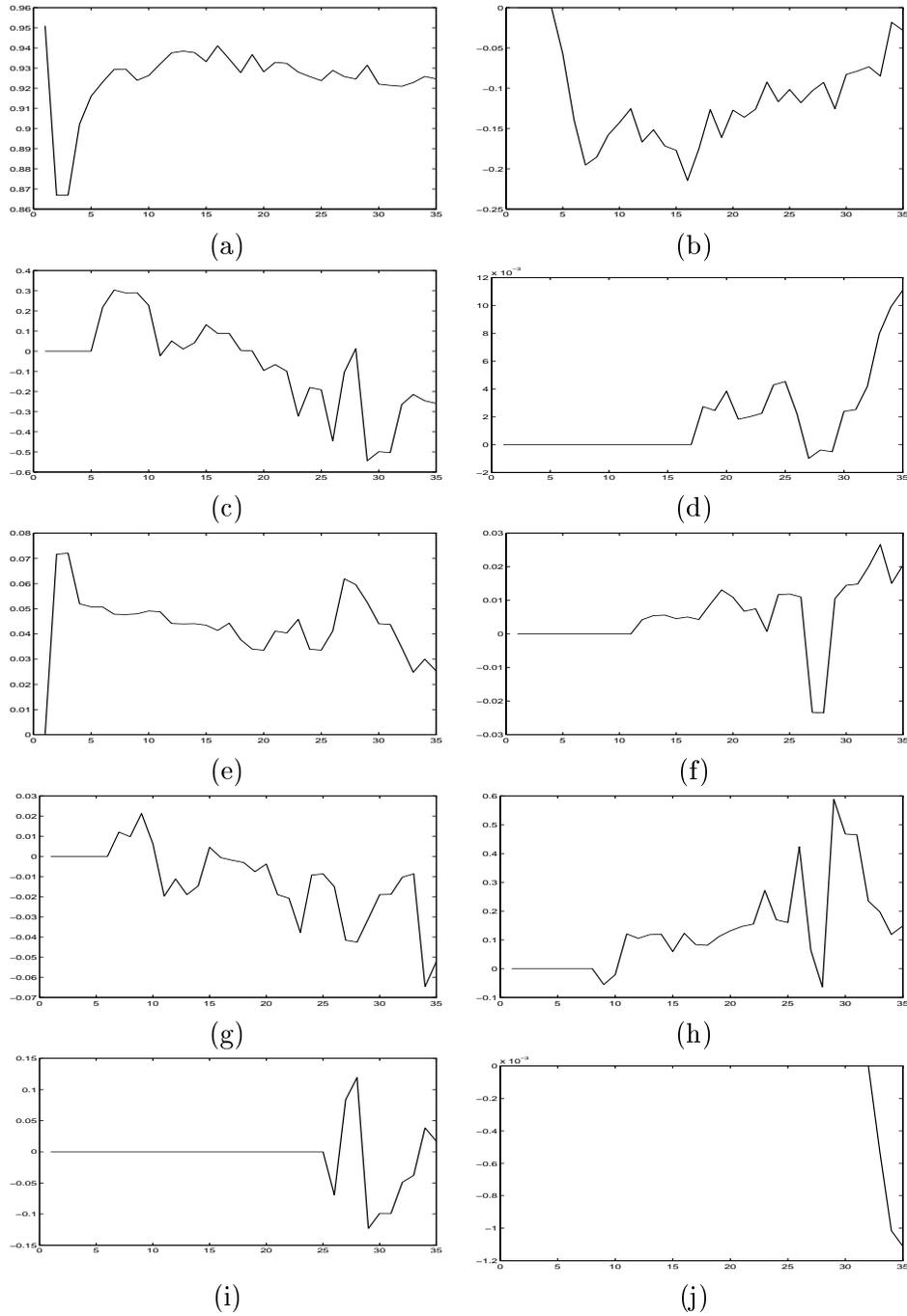


Figura 3.21: Coeficientes de agrupamentos em função do número de termos do modelo para a massa de dados  $\langle din3 \rangle$ : (a)  $\Sigma_y$  (b)  $\Sigma_{y^2}$  (c)  $\Sigma_{y^3}$  (d)  $\Sigma_u$  (e)  $\Sigma_{u^2}$  (f)  $\Sigma_{u^3}$  (g)  $\Sigma_{uy}$  (h)  $\Sigma_{u^2y}$  (i)  $\Sigma_{uy^2}$  (j)  $\Sigma_0$  (O valor máximo para este coeficiente é  $-0,0012$ ).

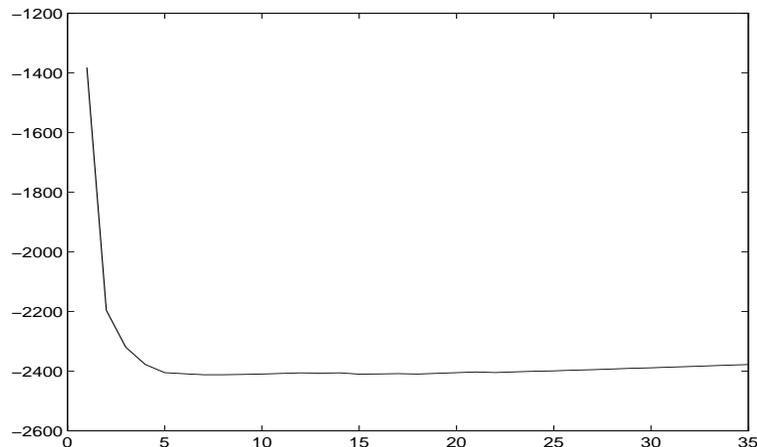


Figura 3.22: Critério AIC - massa de dados  $\langle din2 \rangle$ .

### 3.7 Uso de Critérios de Informação

O critério de informação utilizado foi o  $AIC(2)$ , que é dado por:

$$AIC(2) = N \log[\text{Var}(\xi)] + 2n_{\theta} , \quad (3.5)$$

onde  $\text{Var}(\xi)$  corresponde à variância dos resíduos,  $N$  é o número de dados utilizado para identificar o modelo, sendo, neste caso, igual a 627 e  $n_{\theta}$  corresponde ao número de termos do modelo<sup>5</sup>, que foi variado de 1 a 35.

Desta forma, após definir grau de não linearidade igual a três, atrasos máximos na entrada e na saída iguais a dois e cinco termos de ruído, foram estimados os parâmetros do modelo utilizando o método de mínimos quadrados ortogonal, e através do critério baseado no algoritmo ERR, foram selecionados automaticamente os termos que mais explicavam a variância da saída, obtendo então as Figuras 3.22 e 3.23. É importante, portanto, mencionar que o número de termos escolhido pelo critério AIC dependerá da seleção realizada pelo algoritmo ERR.

O  $AIC(2)$  sugere um modelo com aproximadamente cinco termos, já que tais casos apresentam valores mínimos para as Figuras 3.22 e 3.23. É importante salientar que, critérios de informação têm por finalidade *auxiliar na*

<sup>5</sup>O grau de não linearidade utilizado foi três, que resulta em trinta e cinco termos candidatos.

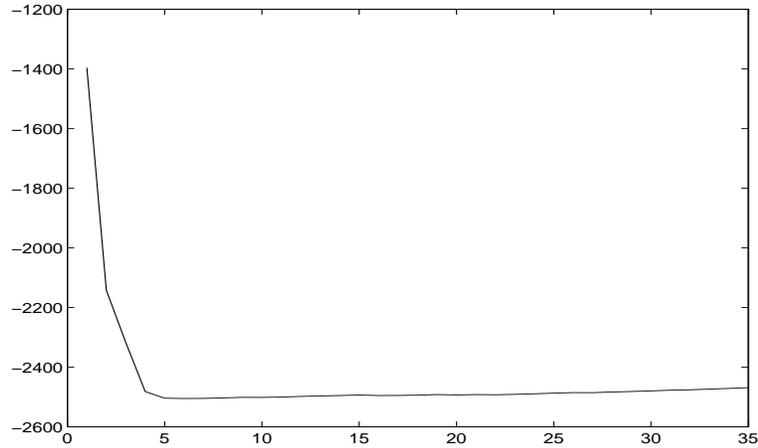


Figura 3.23: Critério AIC - massa de dados  $\langle din3 \rangle$ .

*detecção da estrutura do modelo*, ou seja, há casos em que um número maior ou menor de termos é requerido para atender ao propósito desejado.

### 3.8 Modelos Identificados em Batelada

Foram utilizadas as massas de dados dizimadas indicadas nas Figuras 3.18 e 3.19 para identificar modelos para o sistema. Sendo que os 627 primeiros pontos foram utilizados para encontrar os modelos e os pontos restantes para validação dos mesmos.

Assim procedendo, uma vez escolhido o grau de não linearidade  $l$  foram determinados modelos com o número de termos *sugerido* pelo critério AIC(2). Desta forma, selecionou-se modelos com número de termos  $n_p$  em torno de cinco<sup>6</sup> e por tentativa, determinou-se os atrasos máximos,  $n_y$ ,  $n_u$  e  $n_e$ . Além disso, foram utilizados cinco termos lineares de ruído, de forma a minimizar a polarização dos parâmetros do modelo.

Escolhidos então, os atrasos máximos, o grau de não linearidade, o número de termos de processo e de ruído para compor o modelo, o algoritmo baseado no ERR, apresentado no capítulo 2, selecionou automaticamente os termos que melhor explicavam a variância da saída.

<sup>6</sup>É importante salientar que critérios de informação, como já mencionado, tem por finalidade auxiliar no procedimento de determinação da complexidade do modelo. No entanto, em alguns casos, um número de termos maior é requerido.

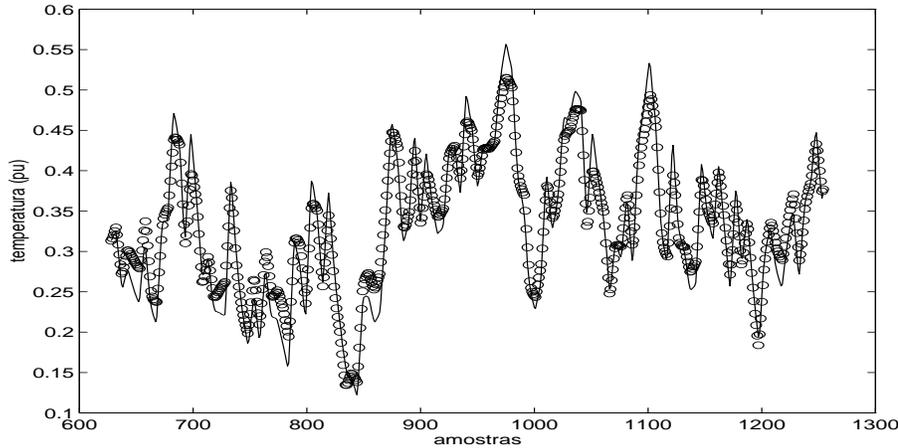


Figura 3.24: Validação dinâmica para o modelo linear (3.6) identificado com a massa de dados  $\langle din2 \rangle$ . (—) dados (o) previsões do modelo (3.6).

Dentre os modelos obtidos, foram selecionados os que apresentaram melhores resultados na validação dinâmica, bem como na recuperação da característica estática não linear, que será discutida no capítulo 4. Foram selecionados modelos ARMAX e modelos NARMAX com grau de não linearidade dois e três. Para cada caso obteve-se um modelo para a massa de dados  $\langle din2 \rangle$  e outro para a  $\langle din3 \rangle$ , perfazendo um total de 10 modelos.

### 3.8.1 Modelo Linear

O modelo ARMAX (*AutoRegressive Moving Average with eXogenous inputs*) com  $n_y = 2$ ,  $n_u = 2$ ,  $n_e = 5$ ,  $l = 1$ ,  $n_p = 4$ ,  $n_n = 5$  obtido para a massa de dados  $\langle din2 \rangle$ , após ter sido eliminado o termo constante  $\Omega_0$  do conjunto dos termos candidatos, é representado na equação (3.6).

$$y(k) = 1,3817y(k-1) + 0,0411u(k-1) - 0,4296y(k-2) - 0,0077u(k-2) + \sum_{i=1}^{n_e} \hat{\theta}_i \xi(k-i) + \xi(k) . \quad (3.6)$$

Da mesma forma, o modelo ARMAX com  $n_y = 2$ ,  $n_u = 2$ ,  $n_e = 5$ ,  $l = 1$ ,  $n_p = 4$ ,  $n_n = 5$  obtido para a massa de dados  $\langle din3 \rangle$ , após ter sido eliminado o termo constante  $\Omega_0$  do conjunto dos termos candidatos, é representado pela equação (3.7).

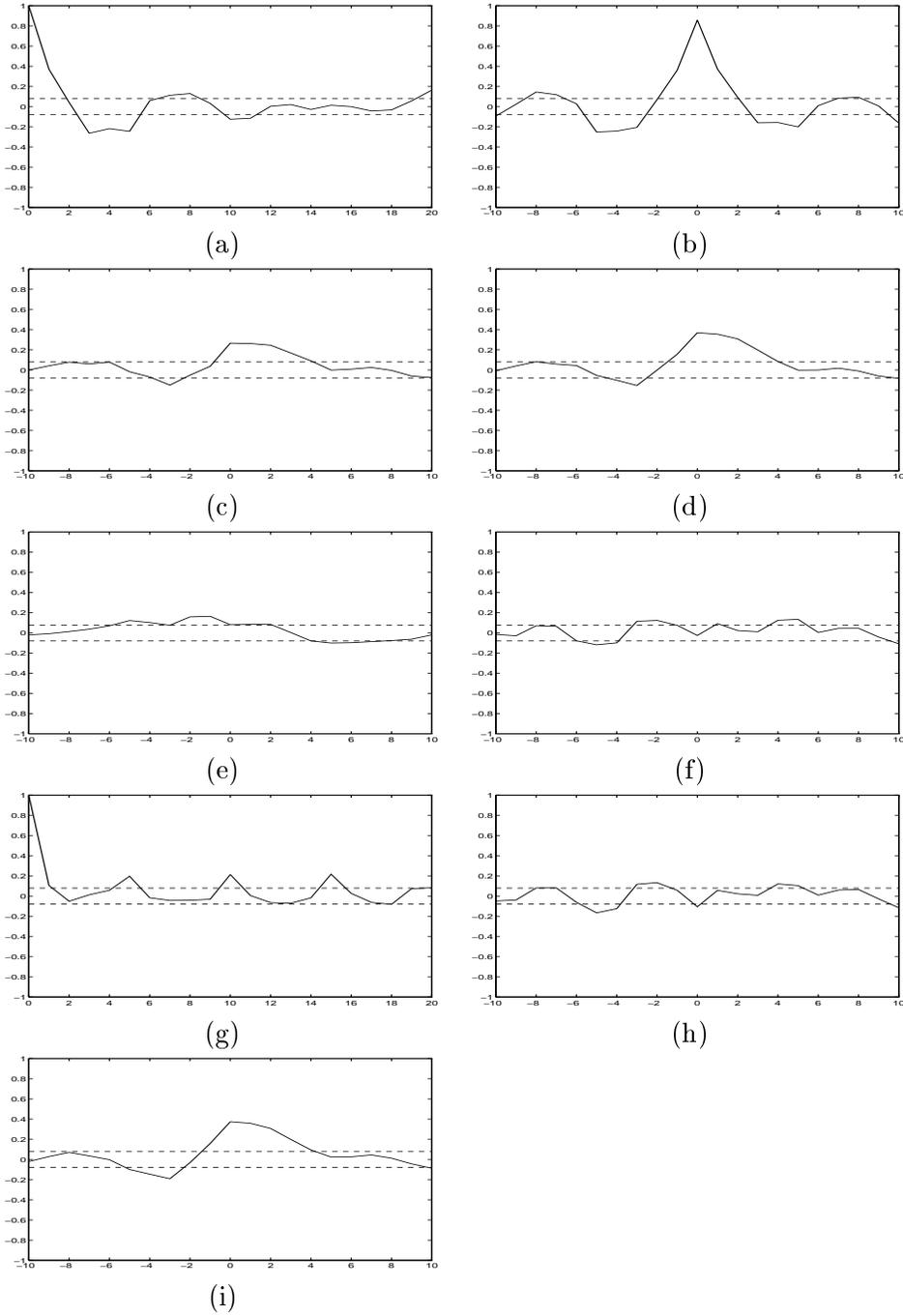


Figura 3.25: Validação estatística para o modelo linear (3.6) e massa de dados  $\langle din2 \rangle$ . (a)  $\phi_{e'e'}$  (b)  $\phi_{(ue)'}e'$  (c)  $\phi_{u'e'}$  (d)  $\phi_{u'^2e'}$  (e)  $\phi_{u'^2e'^2}$  (f)  $\phi_{e'^2e'}$  (g)  $\phi_{e'^2e'^2}$  (h)  $\phi_{e'^2(ye)'}'$  (i)  $\phi_{u'^2(ye)'}'$ .

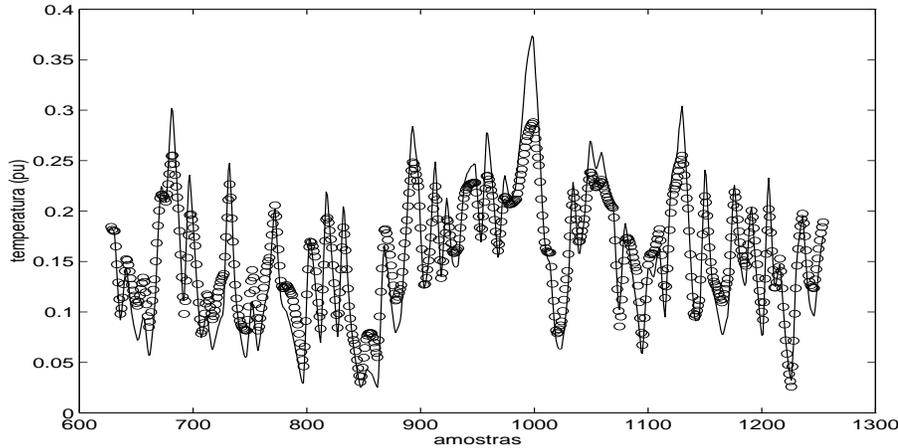


Figura 3.26: Validação dinâmica para o modelo linear (3.7) identificado com a massa de dados  $\langle din3 \rangle$ . (–) dados (o) previsões do modelo (3.7).

$$y(k) = 1,4575 y(k-1) + 0,0359 u(k-1) - 0,5374 y(k-2) - 0,0096 u(k-2) + \sum_{i=1}^{n_e} \hat{\theta}_i \xi(k-i) + \xi(k) . \quad (3.7)$$

Observa-se através das Figuras 3.24 e 3.26, que as previsões feitas com os modelos ARMAX (3.6) e (3.7) são menos precisas nos pontos extremos da curva. Observando as variações dos coeficientes de agrupamentos na Figuras 3.20 e 3.21 percebe-se que agrupamentos do tipo  $\Omega_{y^2}$  e  $\Omega_{u^2}$  parecem ser importantes. Por definição, tais agrupamentos não estão presentes em um modelo ARMAX. Podendo ser esta uma das razões da deficiência de previsão em alguns trechos da curva.

A validação estatística do modelo (3.6), que está indicada na Figura 3.25, é marginalmente pior do que para o modelo (3.7), que está indicada na Figura 3.27. Isto pode ter sido ocasionado pelas autocorrelações nos atrasos 20 e 30 da entrada do sistema para a massa de dados  $\langle din2 \rangle$ , como pode ser observado na Figura 3.8. O gráfico de autocorrelação foi traçado para a taxa de amostragem utilizada para a aquisição de dados, que é igual a 6s. Para o atraso 20, temos que o instante de tempo correspondente é igual a 120s. Como a taxa de amostragem de trabalho utilizada foi 12s, temos que o atraso 20 na Figura 3.8, corresponde ao atraso 10 nos gráficos de validação estatística. De

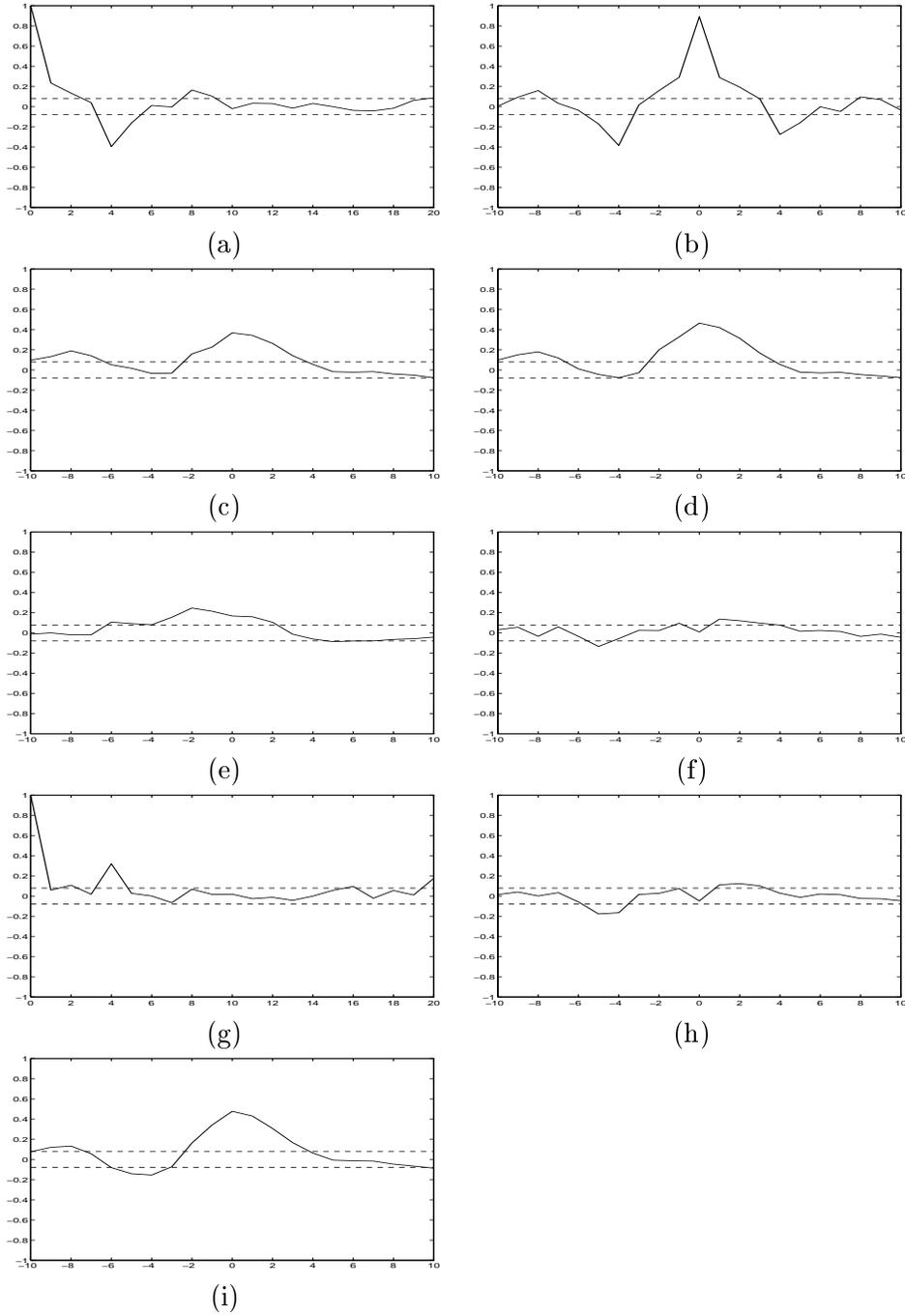


Figura 3.27: Validação estatística para o modelo linear (3.7) e massa de dados  $\langle din3 \rangle$ . (a)  $\phi_{e'e'}$  (b)  $\phi_{(ue)'}e'$  (c)  $\phi_{u'e'}$  (d)  $\phi_{u'^2e'}$  (e)  $\phi_{u'^2e'^2}$  (f)  $\phi_{e'^2e'}$  (g)  $\phi_{e'^2e'^2}$  (h)  $\phi_{e'^2(ye)'}'$  (i)  $\phi_{u'^2(ye)'}'$ .

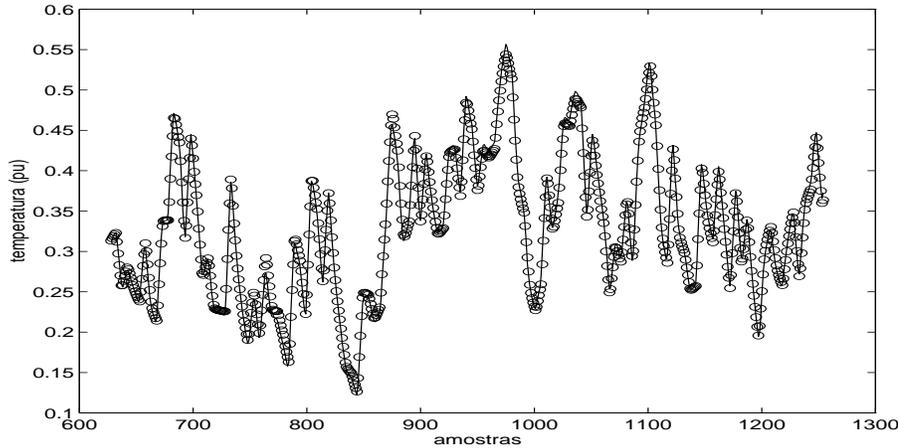


Figura 3.28: Validação dinâmica para o modelo não linear (3.8) identificado com a massa de dados  $\langle din2 \rangle$ . (–) dados (o) predições do modelo (3.8).

forma similar, o atraso 30 corresponde ao atraso 15, ou seja, deve-se achar o atraso correspondente considerando que as massas de dados, Figuras 3.6 e 3.7, foram dizimadas de 2. Desta forma, se observadas as letras  $a, b, f, h, i$  da Figura 3.25 nos atrasos 10 e 15, vemos que se apresentam correlacionadas. O mesmo não ocorre para a Figura 3.27, já que a entrada do sistema para a massa de dados  $\langle din3 \rangle$  não se apresenta correlacionada para tais atrasos.

### 3.8.2 Modelo com Grau de Não Linearidade Dois

O modelo NARMAX (*Nonlinear Autoregressive Moving Average with exogenous inputs*) com  $n_y = 2$ ,  $n_u = 2$ ,  $n_e = 5$ ,  $l = 2$ ,  $n_p = 7$ ,  $n_n = 5$ , identificado para a massa de dados  $\langle din2 \rangle$ , obtido com todos os agrupamentos considerados dentro do conjunto dos candidatos, é igual a:

$$\begin{aligned}
 y(k) = & 1,1740 y(k-1) + 0,0461 u(k-1)^2 + 0,0045 u(k-2) \\
 & + 0,0581 y(k-2)^2 - 0,2098 y(k-2) + 0,0087 u(k-2) u(k-1) \\
 & - 0,1104 y(k-1)^2 + \sum_{i=1}^{n_e} \hat{\theta}_i \xi(k-i) + \xi(k) .
 \end{aligned} \quad (3.8)$$

Da mesma forma, o modelo NARMAX com  $n_y = 2$ ,  $n_u = 2$ ,  $n_e = 5$ ,  $l = 2$ ,  $n_p = 7$ ,  $n_n = 5$ , identificado para a massa de dados  $\langle din3 \rangle$ , obtido

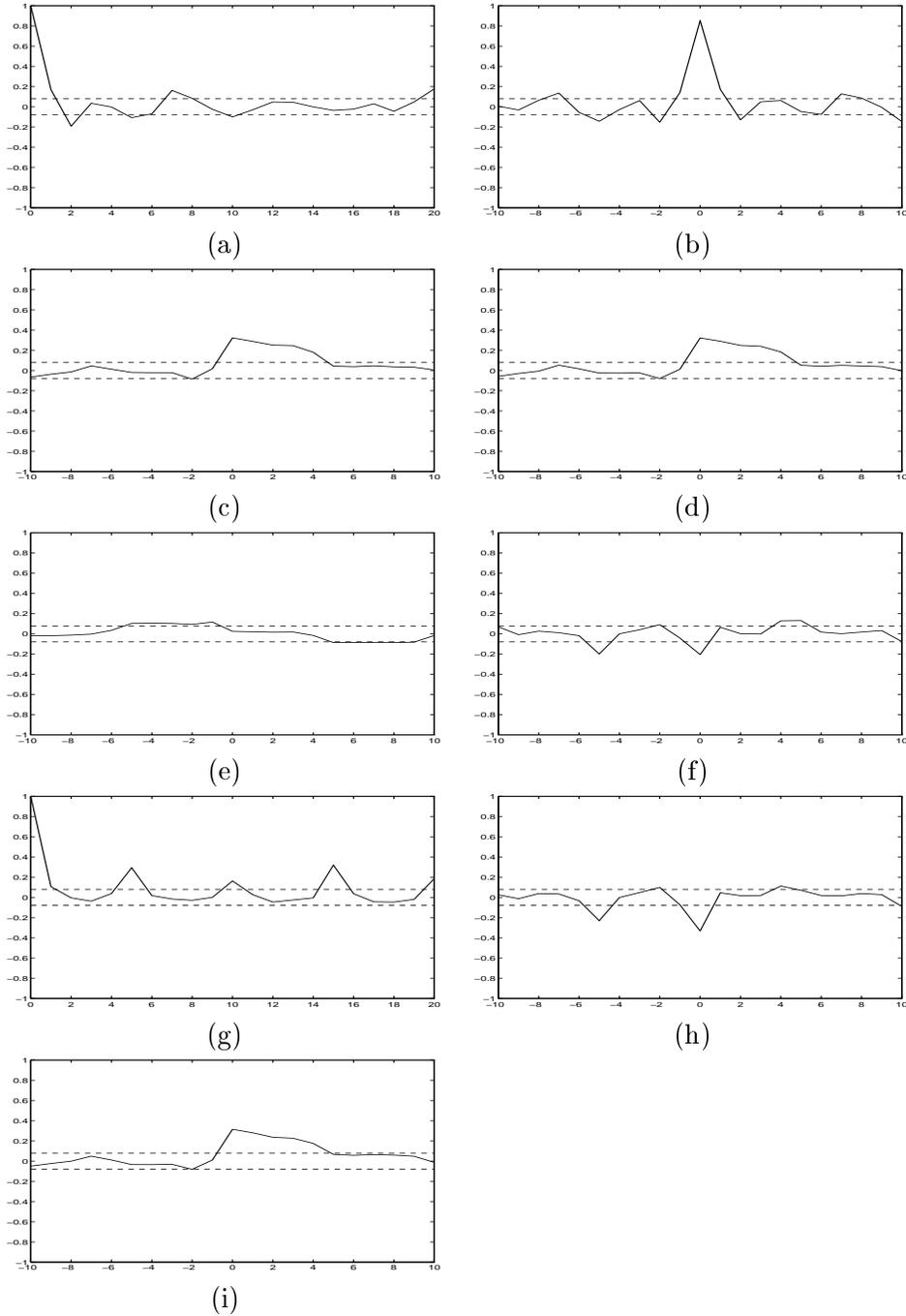


Figura 3.29: Validação estatística para o modelo não linear (3.8) obtido através da massa de dados  $\langle din2 \rangle$ . (a)  $\phi_{e'e'}$  (b)  $\phi_{(ue)'}e'$  (c)  $\phi_{u'e'}$  (d)  $\phi_{u'^2e'}$  (e)  $\phi_{u'^2e'^2}$  (f)  $\phi_{e'^2e'}$  (g)  $\phi_{e'^2e'^2}$  (h)  $\phi_{e'^2(ye)'}$  (i)  $\phi_{u'^2(ye)'}$ .

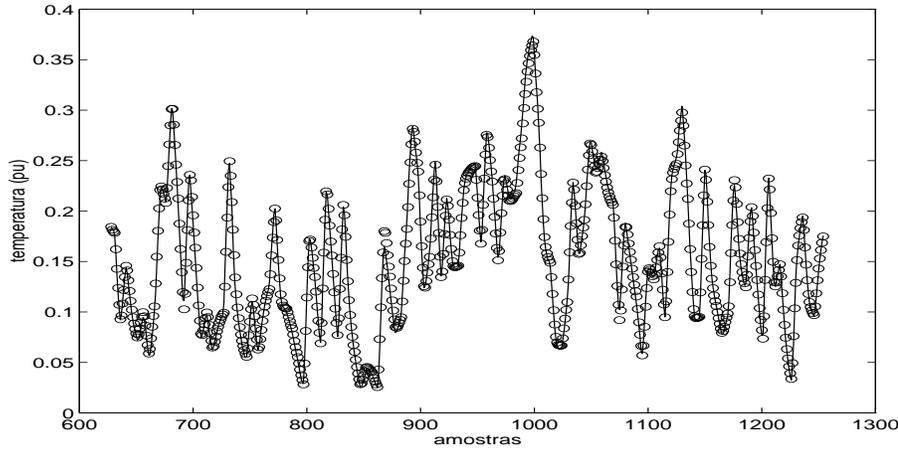


Figura 3.30: Validação dinâmica para o modelo não linear (3.9) identificado com a massa de dados  $\langle din3 \rangle$ . (–) dados (o) previsões do modelo (3.9).

com todos os agrupamentos considerados dentro do conjunto dos candidatos, é:

$$\begin{aligned}
 y(k) = & 1,2482 y(k-1) + 0,0106 u(k-2) u(k-1) + 0,0373 u(k-1)^2 \\
 & - 0,3381 y(k-2) - 0,0191 y(k-1)^2 + 0,0030 u(k-1) \\
 & - 0,0091 u(k-2) y(k-1) + \sum_{i=1}^{n_e} \hat{\theta}_i \xi(k-i) + \xi(k) . \quad (3.9)
 \end{aligned}$$

As validações estatísticas dos modelos (3.8) e (3.9), que estão indicadas nas Figuras 3.29 e 3.31, respectivamente, não apresentam alterações consideráveis se comparadas às obtidas pelos modelos lineares.

No entanto, as validações dinâmicas representadas pelas Figuras 3.28 e 3.30, indicam que tais modelos são capazes de representar a dinâmica do sistema satisfatoriamente, mesmo nos picos. Isto pode ser confirmado observando a Tabela 3.1.

Conforme discutido na análise de agrupamentos de termos, os agrupamentos  $\Omega_u$  e  $\Omega_0$  parecem ser espúrios, sendo portanto, desnecessários ao modelo. Desta forma, tais agrupamentos foram excluídos do conjunto dos agrupamentos candidatos, de forma a facilitar a determinação de um modelo para o sistema.

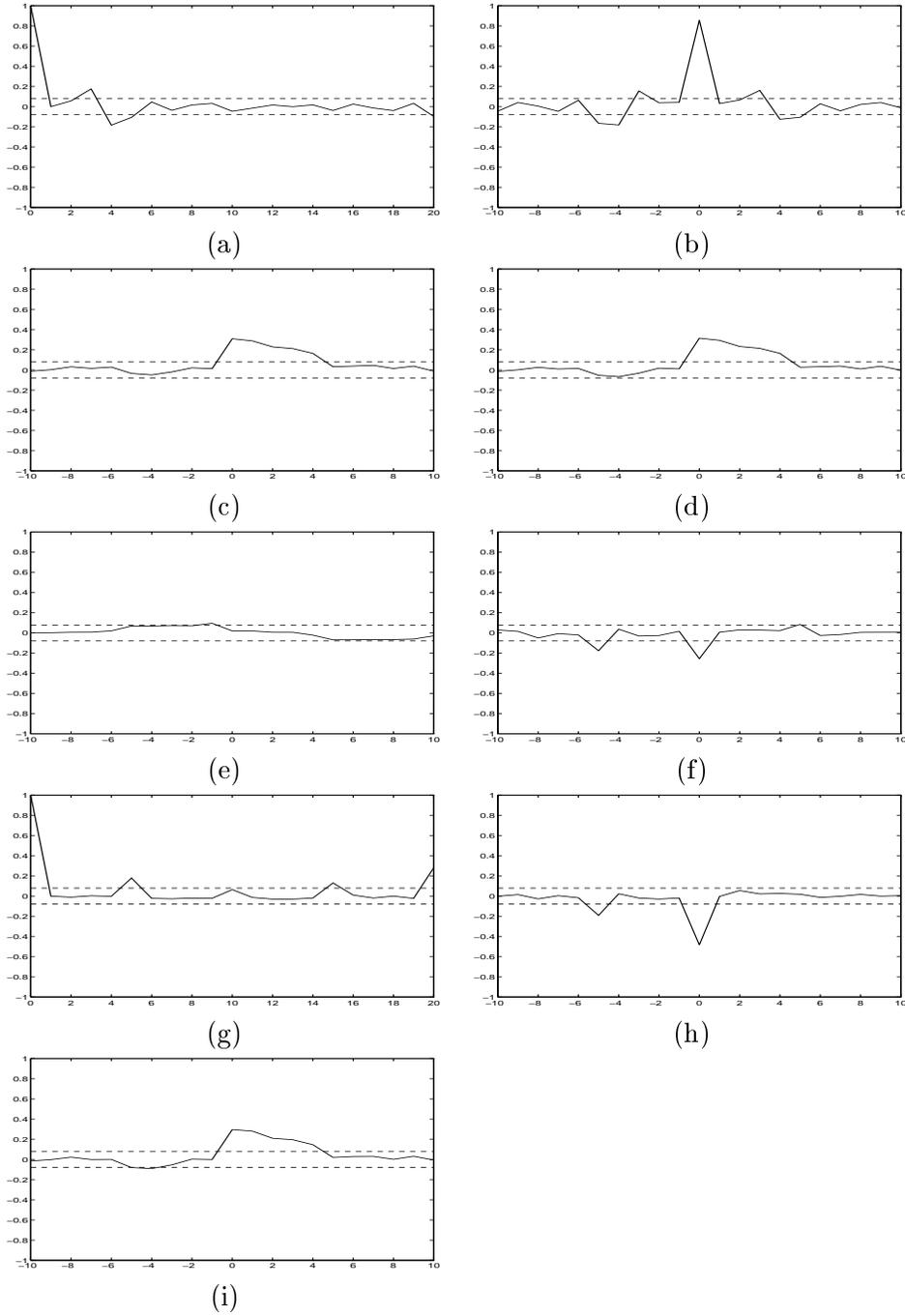


Figura 3.31: Validação estatística para o modelo não linear (3.9) obtido através da massa de dados  $\langle din3 \rangle$ . (a)  $\phi_{e'e'}$  (b)  $\phi_{(ue)'}e'$  (c)  $\phi_{u'e'}$  (d)  $\phi_{u'^2e'}$  (e)  $\phi_{u'^2e'^2}$  (f)  $\phi_{e'^2e'}$  (g)  $\phi_{e'^2e'^2}$  (h)  $\phi_{e'^2(ye)'}'$  (i)  $\phi_{u'^2(ye)'}'$ .

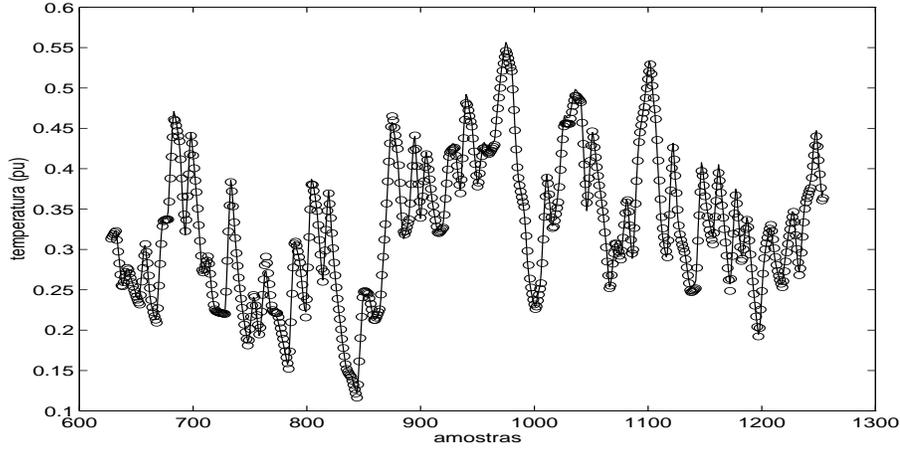


Figura 3.32: Validação dinâmica para o modelo não linear (3.10) identificado com a massa de dados  $\langle din2 \rangle$ . (—) dados (o) previsões do modelo (3.10).

Além disso, como será visto no próximo capítulo, modelos sem o agrupamento  $\Omega_{y^2}$  facilitam a recuperação da característica estática não linear. Desta forma, tais agrupamentos também foram excluídos do conjunto dos agrupamentos candidatos.

O modelo NARMAX com  $n_y = 2$ ,  $n_u = 2$ ,  $n_e = 5$ ,  $l = 2$ ,  $n_p = 7$ ,  $n_n = 5$  identificado para a massa de dados  $\langle din2 \rangle$ , com os agrupamentos  $\Omega_{y^2}$ ,  $\Omega_0$  e  $\Omega_u$  excluídos do conjunto dos agrupamentos candidatos, é:

$$\begin{aligned}
 y(k) = & 1,3920 y(k-1) + 0,0454 u(k-1)^2 - 0,4235 y(k-2) \\
 & - 0,4388 u(k-2) y(k-1) + 0,3756 u(k-2) y(k-2) \\
 & + 0,0218 u(k-2)^2 + 0,0097 u(k-2) u(k-1) \\
 & + \sum_{i=1}^{n_e} \hat{\theta}_i \xi(k-i) + \xi(k) .
 \end{aligned} \tag{3.10}$$

Da mesma forma, o modelo NARMAX com  $n_y = 2$ ,  $n_u = 2$ ,  $n_e = 5$ ,  $l = 2$ ,  $n_p = 7$ ,  $n_n = 5$  identificado para a massa de dados  $\langle din3 \rangle$ , com os agrupamentos  $\Omega_{y^2}$ ,  $\Omega_0$  e  $\Omega_u$  excluídos do conjunto dos agrupamentos candidatos, é:

$$y(k) = 1,2929 y(k-1) + 0,0101 u(k-2) u(k-1) + 0,0407 u(k-1)^2$$

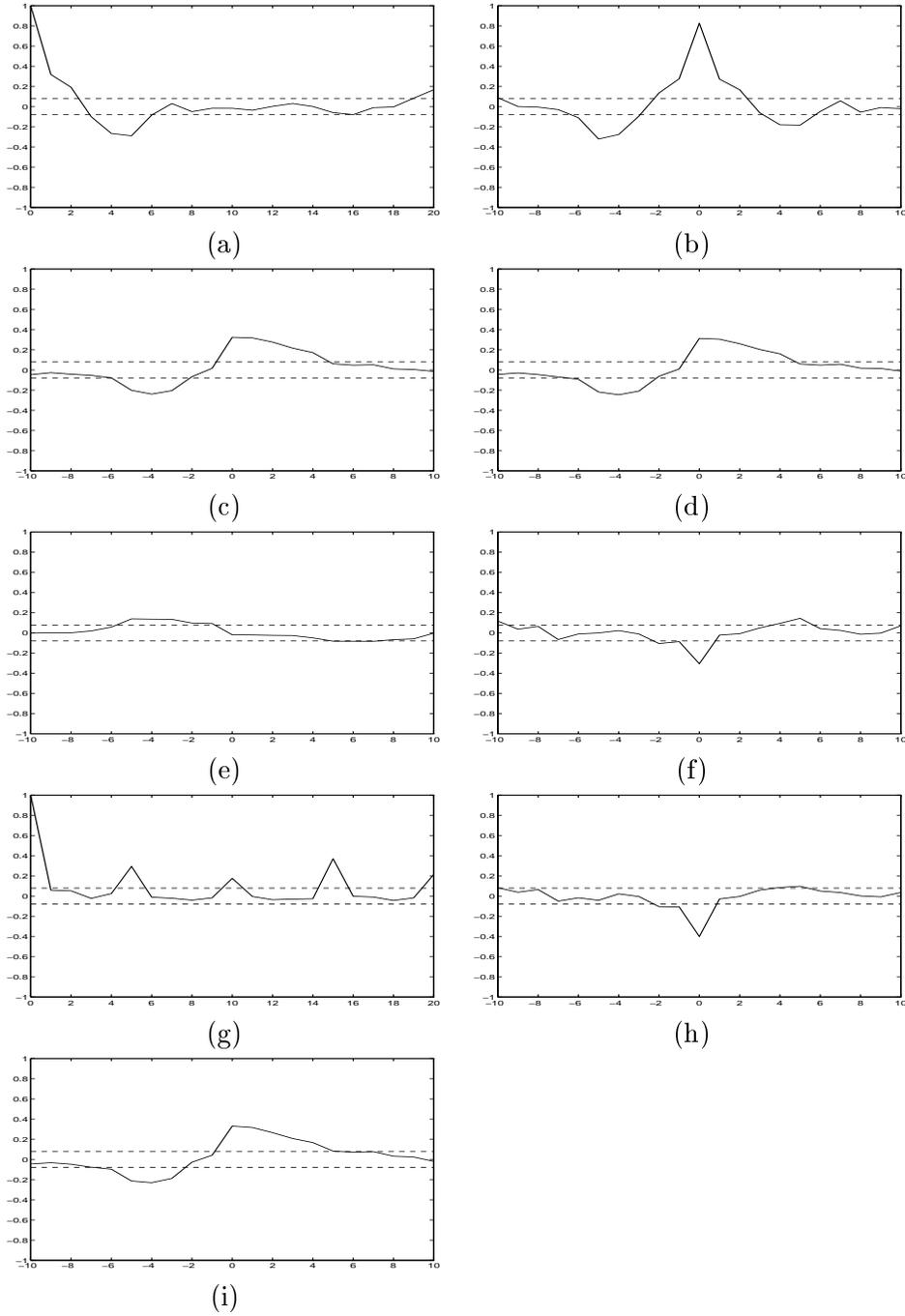


Figura 3.33: Validação estatística para o modelo não linear (3.10) e massa de dados  $\langle din2 \rangle$ . (a)  $\phi_{e'e'}$  (b)  $\phi_{(ue)'}e'$  (c)  $\phi_{u'e'}$  (d)  $\phi_{u'^2e'}$  (e)  $\phi_{u'^2e'^2}$  (f)  $\phi_{e'^2e'}$  (g)  $\phi_{e'^2e'^2}$  (h)  $\phi_{e'^2(ye)'}'$  (i)  $\phi_{u'^2(ye)'}'$ .

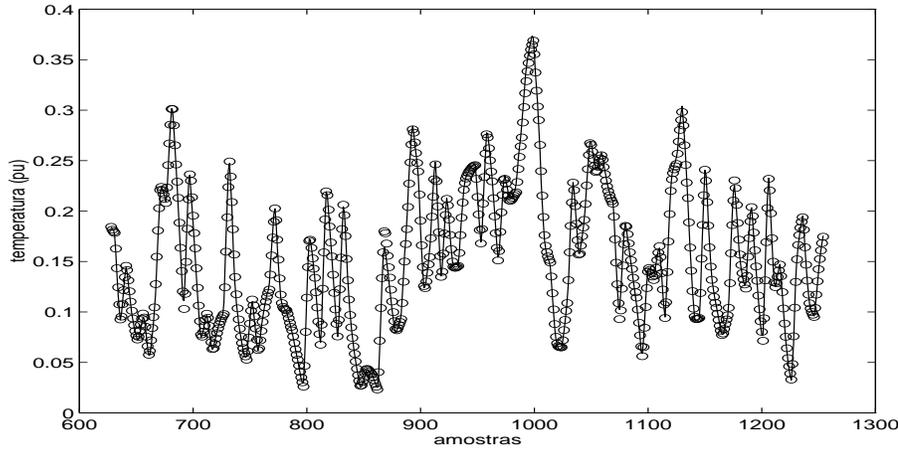


Figura 3.34: Validação dinâmica para o modelo não linear (3.11) identificado com a massa de dados  $\langle din3 \rangle$ . (—) dados (o) previsões do modelo (3.11).

$$\begin{aligned}
 & -0,3779 y(k-2) - 0,1280 u(k-2) y(k-1) \\
 & + 0,0957 u(k-2) y(k-2) + 0,0051 u(k-2)^2 \\
 & + \sum_{i=1}^{n_e} \hat{\theta}_i \xi(k-i) + \xi(k) .
 \end{aligned} \tag{3.11}$$

As validações estatísticas para os modelos (3.10) e (3.11), que estão representadas nas Figuras 3.33 e Figura 3.35, ainda indicam que os resíduos são correlacionados.

Por outro lado, as Figuras 3.32 e 3.34 referentes aos modelos (3.10) e (3.11), respectivamente, indicam que os mesmos são eficazes para representar a dinâmica do sistema a ser identificado, satisfazendo, portanto, a primeira condição exigida para os modelos identificados.

Desta forma, comparando-se os modelos (3.8) e (3.9), obtidos considerando todos os agrupamentos possíveis no conjunto dos candidatos, com os modelos (3.10) e (3.11), obtidos após terem sido eliminados os agrupamentos  $\Omega_0$ ,  $\Omega_u$  e  $\Omega_{y^2}$  do conjunto dos candidatos, não se observam diferenças significativas no que diz respeito à capacidade de previsão (observar Tabela 3.1). No entanto, é importante mencionar que os agrupamentos  $\Omega_0$  e  $\Omega_u$  foram eliminados, conforme sugerido pela análise de agrupamentos de termos, com o objetivo de melhorar a capacidade de recuperação da característica estática não linear. Ao passo que, o agrupamento  $\Omega_{y^2}$  foi excluído dos candidatos com o objetivo de facilitar a recuperação da característica estática. A re-

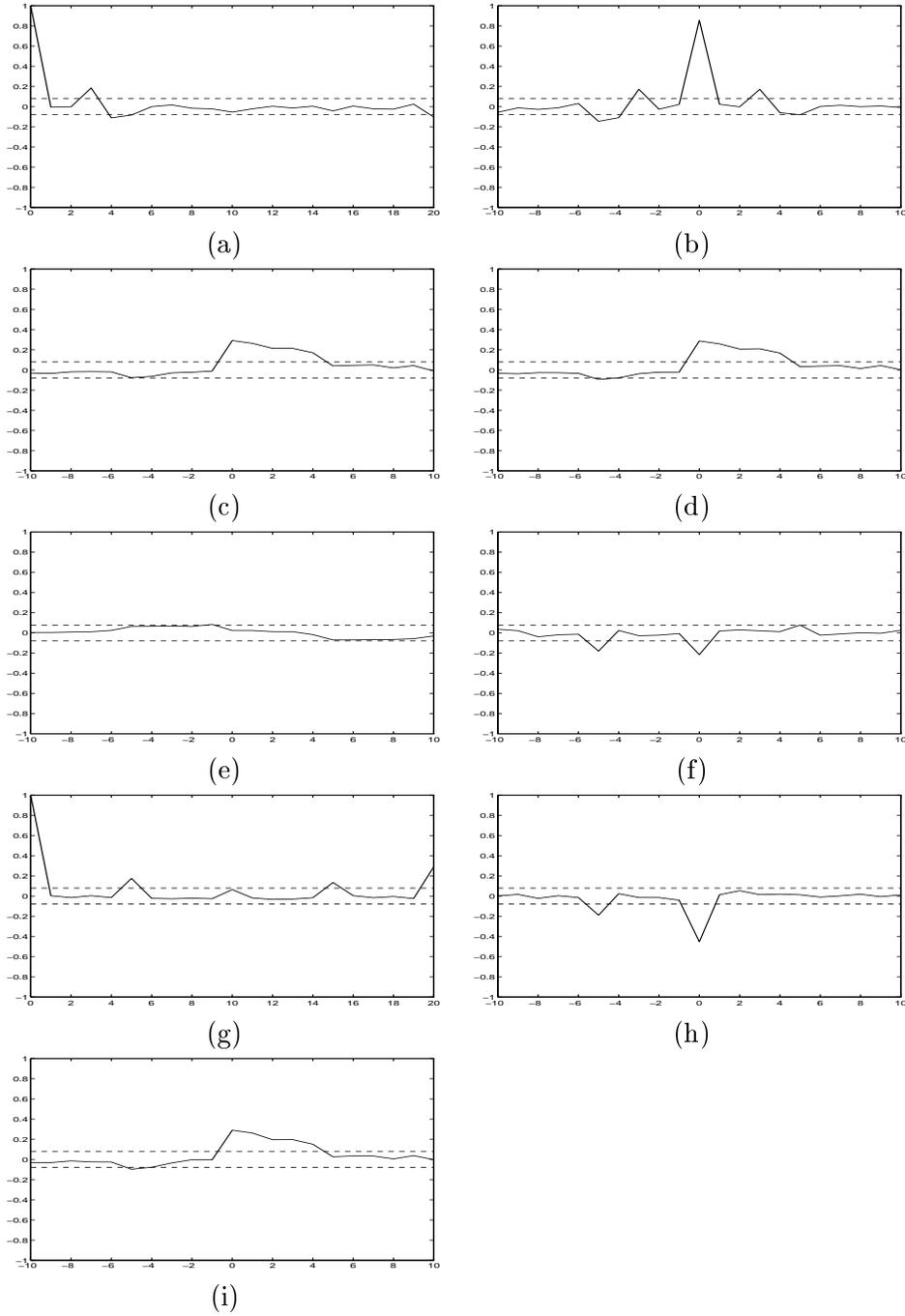


Figura 3.35: Validação estatística para o modelo não linear (3.11) obtido através da massa de dados  $\langle din3 \rangle$ . (a)  $\phi_{e'e'}$  (b)  $\phi_{(ue)'}e'$  (c)  $\phi_{u'e'}$  (d)  $\phi_{u'^2e'}$  (e)  $\phi_{u'^2e'^2}$  (f)  $\phi_{e'^2e'}$  (g)  $\phi_{e'^2e'^2}$  (h)  $\phi_{e'^2(ye)'}$  (i)  $\phi_{u'^2(ye)'}$ .

cuperação de tal característica a partir de tais modelos será discutida no próximo capítulo.

### 3.8.3 Modelo com Grau de Não Linearidade Três

O modelo NARMAX com  $n_y = 2$ ,  $n_u = 2$ ,  $n_e = 5$ ,  $l = 3$ ,  $n_p = 15$ ,  $n_n = 5$  identificado para a massa de dados  $\langle \text{din2} \rangle$ , está representado pela equação (3.12). Todos os agrupamentos foram considerados no conjunto dos agrupamentos candidatos. O algoritmo baseado no *ERR* não selecionou o agrupamento  $\Omega_{y^3}$  para esta massa de dados. Isto pode ser justificado observando a Figura 3.20-c, que sugere que o agrupamento  $\Omega_{y^3}$  é espúrio.

$$\begin{aligned}
y(k) = & 1,1204 y(k-1) + 0,0395 u(k-1)^2 + 0,0047 u(k-2) \\
& + 3,2253 y(k-2)^2 - 0,1564 y(k-2) - 0,0213 u(k-2) u(k-1) \\
& + 0,0087 u(k-2)^2 y(k-1) + 0,0098 u(k-2) u(k-1)^2 \\
& + 3,1842 y(k-1)^2 - 6,4774 y(k-2) y(k-1) \\
& - 0,6074 u(k-1) y(k-2) - 1,2103 u(k-2) u(k-1) y(k-1) \\
& + 1,1283 u(k-2) u(k-1) y(k-2) + 0,6673 u(k-1) y(k-1) \\
& + 0,0465 u(k-2)^2 u(k-1) + \sum_{i=1}^{n_e} \hat{\theta}_i \xi(k-i) + \xi(k) . \quad (3.12)
\end{aligned}$$

Da mesma forma, o modelo NARMAX com  $n_y = 2$ ,  $n_u = 2$ ,  $n_e = 5$ ,  $l = 3$ ,  $n_p = 15$ ,  $n_n = 5$  identificado para a massa de dados  $\langle \text{din3} \rangle$ , com todos os agrupamentos compondo o conjunto dos candidatos, é:

$$\begin{aligned}
y(k) = & 1,3307 y(k-1) + 0,0120 u(k-2) u(k-1) + 0,0315 u(k-1)^2 \\
& - 0,3975 y(k-2) - 0,1770 y(k-1)^2 + 0,4793 y(k-2)^3 \\
& - 0,0250 u(k-1) y(k-2) - 0,0600 u(k-2) y(k-1) \\
& - 1,0543 u(k-1) y(k-1)^2 + 0,5037 u(k-2) y(k-1)^2 \\
& - 0,3481 y(k-1)^3 + 0,0046 u(k-1)^3 \\
& + 0,8866 u(k-1) y(k-2) y(k-1) - 0,2768 u(k-2) y(k-2)^2 \\
& + 0,0896 u(k-1) y(k-1) \\
& + \sum_{i=1}^{n_e} \hat{\theta}_i \xi(k-i) + \xi(k) . \quad (3.13)
\end{aligned}$$

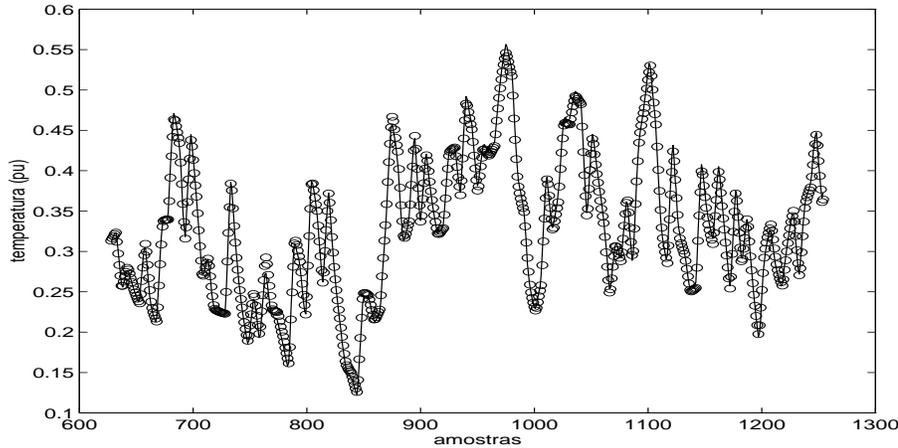


Figura 3.36: Validação dinâmica para o modelo não linear (3.12) com grau de não linearidade igual a três identificado com a massa de dados  $\langle din2 \rangle$ . (—) dados (o) predições do modelo (3.12).

As validações estatísticas para os modelos (3.12) e (3.13) estão indicadas nas Figuras 3.37 e 3.39, respectivamente. Ao passo que, as validações dinâmicas estão representadas nas Figuras 3.36 e 3.38,

As características estáticas<sup>7</sup> obtidas para os modelos (3.12) e (3.13) foram melhores do que as obtidas para os modelos anteriores. No entanto, no caso do modelo (3.13) a recuperação analítica da característica estática não linear é mais difícil devido à inclusão do agrupamento  $\Omega_{y3}$ . Desta forma, para que se obtenha características satisfatórias tanto para a massa de dados  $\langle din2 \rangle$ , quanto para a  $\langle din3 \rangle$ , utilizando-se o mesmo conjunto de agrupamentos candidatos, e obtendo modelos compostos pelos mesmos agrupamentos<sup>8</sup>, algumas alterações devem ser feitas.

Observando a Figura 3.20 percebe-se que o agrupamento  $\Omega_{y3}$  só aparece após o 20<sup>o</sup> termo. Desta forma, baseando na análise de agrupamentos da massa de dados  $\langle din2 \rangle$ , pode-se concluir que o agrupamento  $\Omega_{y3}$  é um agrupamento espúrio. Isso é confirmado pelo fato do algoritmo baseado no

<sup>7</sup>A recuperação da característica estática será discutida no próximo capítulo.

<sup>8</sup>Na estimação recursiva que será realizada neste trabalho, só serão estimados os parâmetros do modelo. Desta forma, uma estrutura deve ser definida através da identificação em batelada. É desejável que os modelos identificados para as massas de dados  $\langle din2 \rangle$  e  $\langle din3 \rangle$  sejam compostos pelos mesmos tipos de agrupamentos, restando apenas, determinar a estrutura com base nestes agrupamentos.

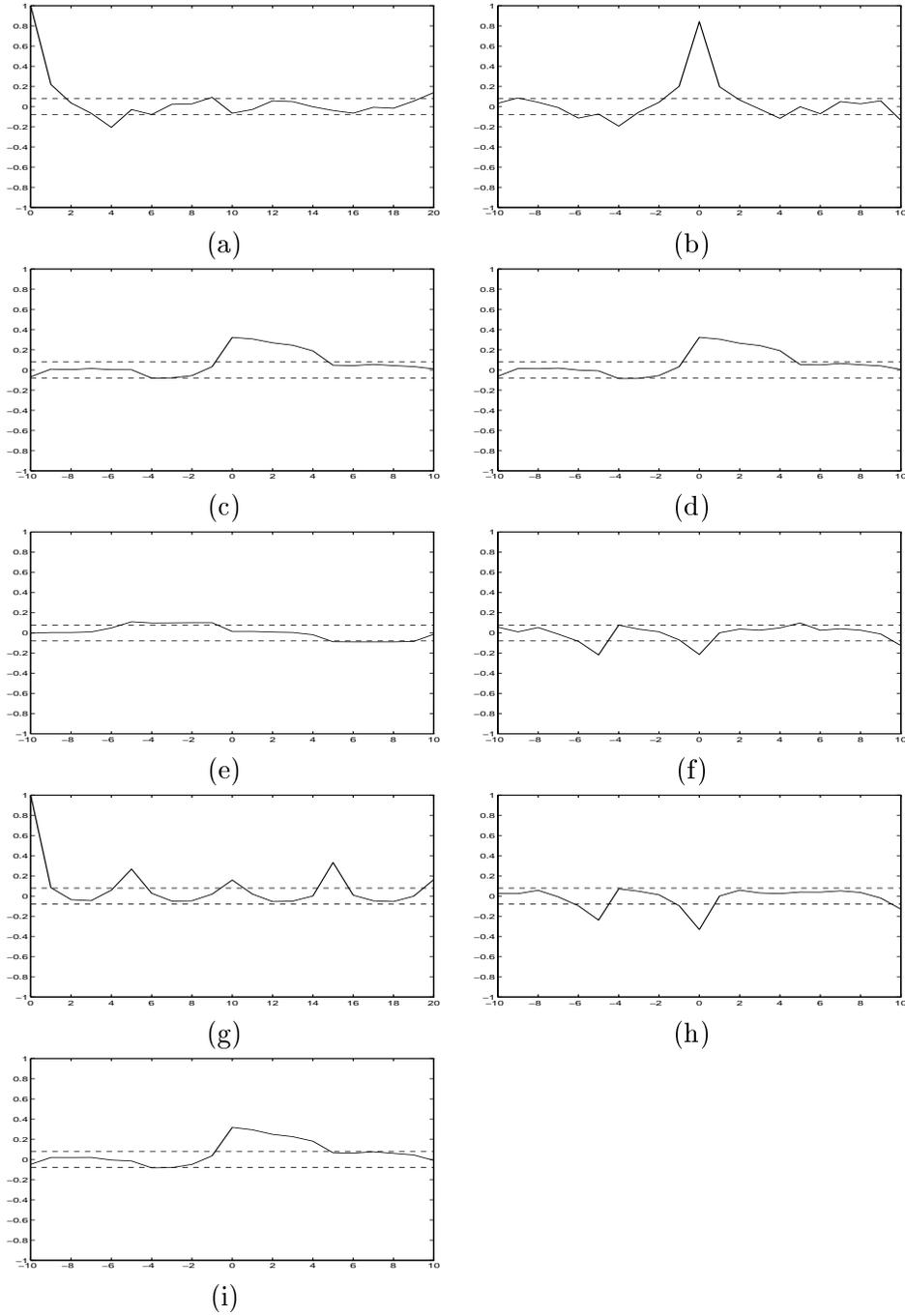


Figura 3.37: Validação estatística para o modelo não linear (3.12) obtido através da massa de dados  $\langle din2 \rangle$ . (a)  $\phi_{e'e'}$  (b)  $\phi_{(ue)'}e'$  (c)  $\phi_{u'e'}$  (d)  $\phi_{u'^2e'}$  (e)  $\phi_{u'^2e'^2}$  (f)  $\phi_{e'^2e'}$  (g)  $\phi_{e'^2e'^2}$  (h)  $\phi_{e'^2(ye)'}$  (i)  $\phi_{u'^2(ye)'}$ .

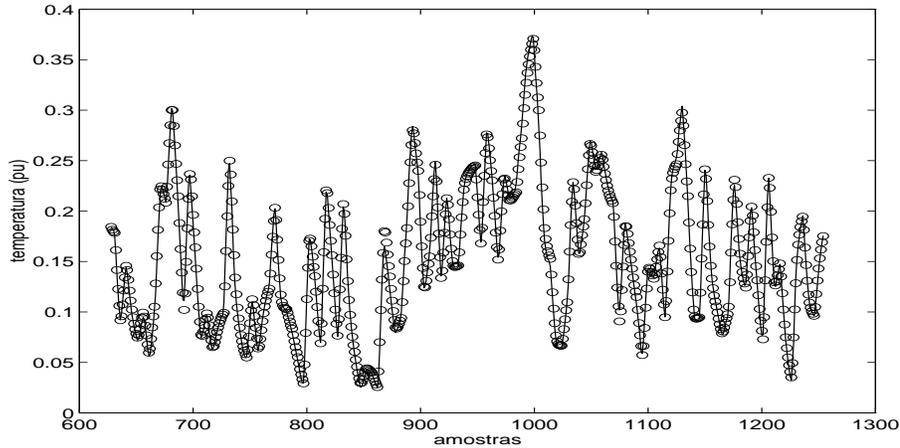


Figura 3.38: Validação dinâmica para o modelo não linear (3.13) identificado com a massa de dados  $\langle din3 \rangle$ . (–) dados (o) previsões do modelo (3.13).

$ERR$  não ter selecionado tal agrupamento para o modelo (3.12).

Por outro lado, a Figura 3.21 não sugere que o agrupamento  $\Omega_{y^3}$  seja espúrio. De fato,  $\Omega_{y^3}$  é selecionado pelo  $ERR$  para compor o modelo (3.13). No entanto, a recuperação analítica da característica estática a partir de um modelo contendo este agrupamento se torna uma tarefa mais complexa, como será visto no capítulo seguinte.

Além da dificuldade citada acima, a inclusão do agrupamento  $\Omega_{y^3}$  no modelo dificulta a escolha de uma estrutura única para a estimação recursiva, já que para a massa de dados  $\langle din2 \rangle$  este agrupamento não foi selecionado pelo algoritmo  $ERR$ .

De forma a obter estruturas compostas pelos mesmos agrupamentos tanto para a massa de dados  $\langle din2 \rangle$ , quanto para a  $\langle din3 \rangle$ , eliminaram-se, do conjunto de agrupamentos candidatos, os agrupamentos  $\Omega_u$ ,  $\Omega_0$ ,  $\Omega_{y^3}$  e  $\Omega_{u^2y}$ , que foram considerados espúrios pela análise de agrupamentos de termos<sup>9</sup>.

Os modelos NARMAX com  $n_y = 2$ ,  $n_u = 2$ ,  $n_e = 5$ ,  $l = 3$ ,  $n_p = 15$ ,  $n_n = 5$  e com os agrupamentos  $\Omega_u$ ,  $\Omega_0$ ,  $\Omega_{y^3}$  e  $\Omega_{u^2y}$  excluídos do conjunto dos agrupamentos candidatos, identificados para as massas de dados  $\langle din2 \rangle$  e  $\langle din3 \rangle$ , estão representados nas equações (3.14) e (3.15), respectivamente.

<sup>9</sup>O agrupamento  $\Omega_{u^2y}$ , apesar de ter sido considerado espúrio pela análise de agrupamentos de termos, mostrou-se importante para o modelo, pois sua eliminação resultou na degradação da característica estática estimada a partir da massa de dados  $\langle din3 \rangle$ .

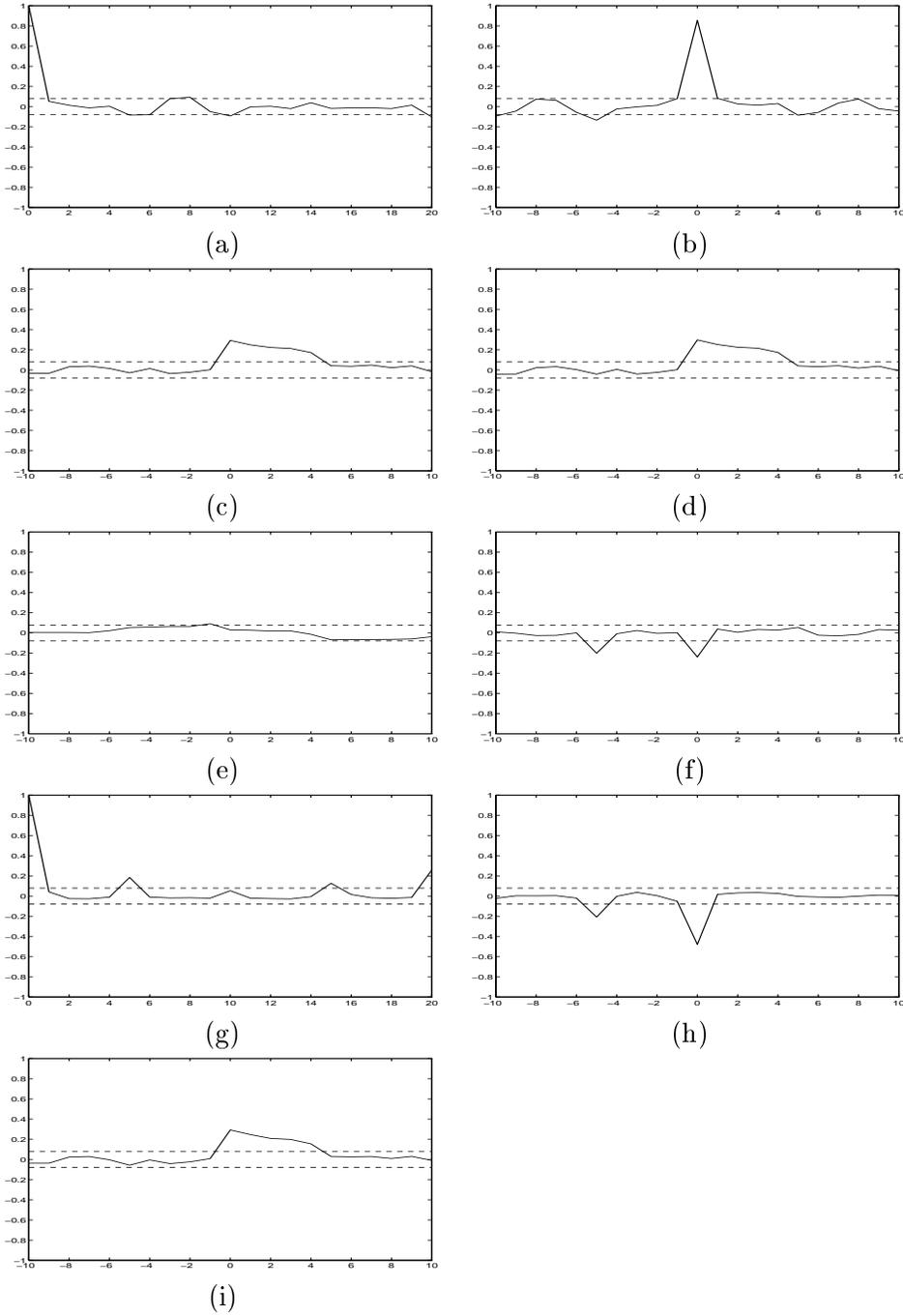


Figura 3.39: Validação estatística para o modelo não linear (3.13) obtido através da massa de dados  $\langle din3 \rangle$ . (a)  $\phi_{e'e'}$  (b)  $\phi_{(ue)'}e'$  (c)  $\phi_{u'e'}$  (d)  $\phi_{u'^2e'}$  (e)  $\phi_{u'^2e'^2}$  (f)  $\phi_{e'^2e'}$  (g)  $\phi_{e'^2e'^2}$  (h)  $\phi_{e'^2(ye)'}$  (i)  $\phi_{u'^2(ye)'}$ .

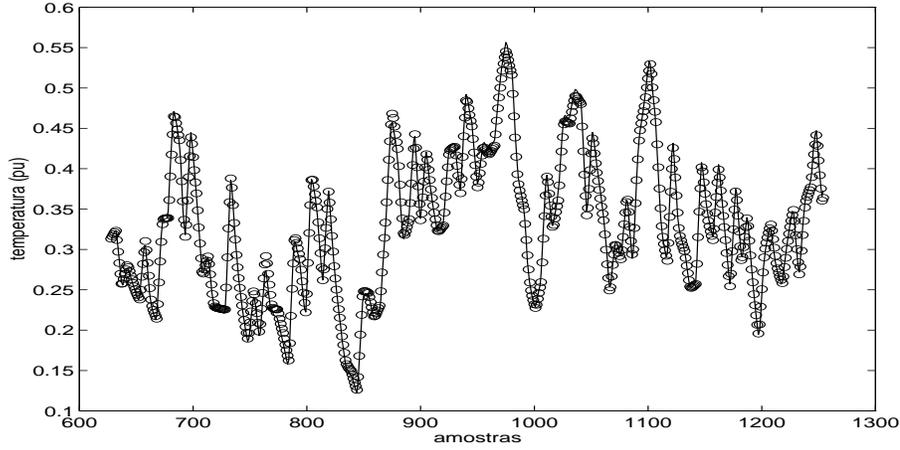


Figura 3.40: Validação dinâmica para o modelo não linear (3.14) identificado com a massa de dados  $\langle din2 \rangle$  e com os agrupamentos  $\Omega_u$ ,  $\Omega_0$ ,  $\Omega_{y^3}$  e  $\Omega_{u^2y}$  excluídos dos agrupamentos candidatos. (—) dados (o) previsões do modelo (3.14).

$$\begin{aligned}
 y(k) = & 1,3040 y(k-1) + 0,0389 u(k-1)^2 - 0,3303 y(k-2) \\
 & - 0,1913 y(k-1)^2 + 0,0179 u(k-2) u(k-1) \\
 & - 0,0034 u(k-2) u(k-1)^2 - 0,3396 u(k-2) y(k-1) \\
 & + 0,3125 u(k-2) y(k-2) + 0,0244 u(k-2)^3 \\
 & + 0,1513 u(k-1) y(k-1) - 0,0167 u(k-2)^2 u(k-1) \\
 & - 0,1305 u(k-1) y(k-2) + 0,0027 u(k-1)^3 + 0,1262 y(k-2)^2 \\
 & + 0,0032 u(k-1) y(k-1)^2 + \sum_{i=1}^{n_e} \hat{\theta}_i \xi(k-i) + \xi(k) . \quad (3.14)
 \end{aligned}$$

$$\begin{aligned}
 y(k) = & 1,3698 y(k-1) + 0,0090 u(k-2) u(k-1) + 0,0331 u(k-1)^2 \\
 & - 0,4400 y(k-2) - 0,9072 y(k-1)^2 + 0,7568 y(k-2) y(k-1) \\
 & - 1,7758 u(k-1) y(k-2)^2 + 4,4094 u(k-1) y(k-2) y(k-1) \\
 & - 0,0559 u(k-2) y(k-2) + 0,8038 u(k-2) y(k-1)^2 \\
 & - 0,0712 u(k-1) y(k-2) - 2,7957 u(k-1) y(k-1)^2 \\
 & + 0,0035 u(k-1)^3 + 0,1408 u(k-1) y(k-1)
 \end{aligned}$$

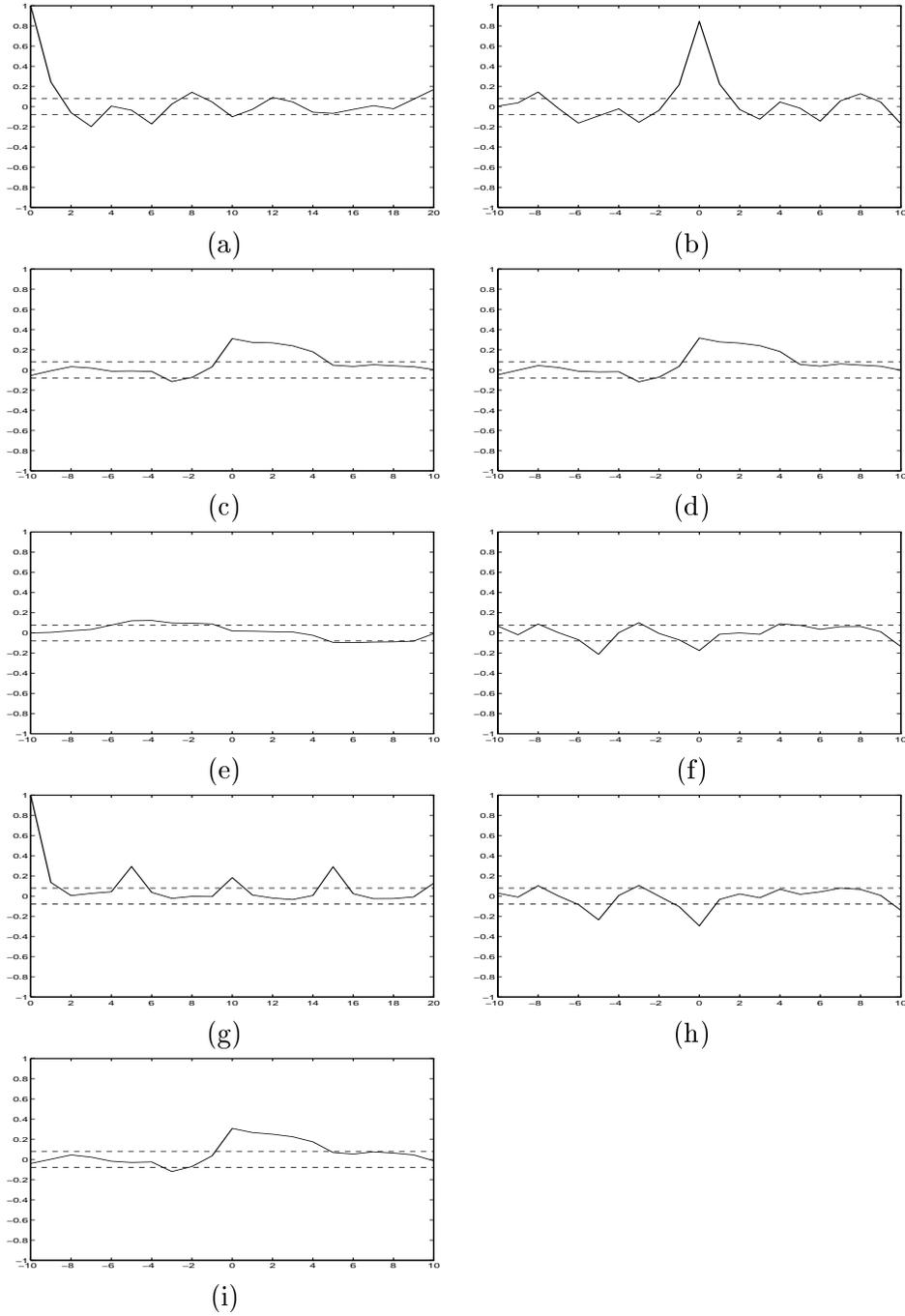


Figura 3.41: Validação estatística para o modelo não linear (3.14) obtido através da massa de dados  $\langle din2 \rangle$  e com os agrupamentos  $\Omega_u$ ,  $\Omega_o$ ,  $\Omega_{y^3}$  e  $\Omega_{u^2y}$  excluídos dos agrupamentos candidatos. (a)  $\phi_{e'e'}$  (b)  $\phi_{(ue)'}e'$  (c)  $\phi_{u'e'}$  (d)  $\phi_{u'^2e'}$  (e)  $\phi_{u'^2e'^2}$  (f)  $\phi_{e'^2e'}$  (g)  $\phi_{e'^2e'^2}$  (h)  $\phi_{e'^2(ye)'}$  (i)  $\phi_{u'^2(ye)'}$ .

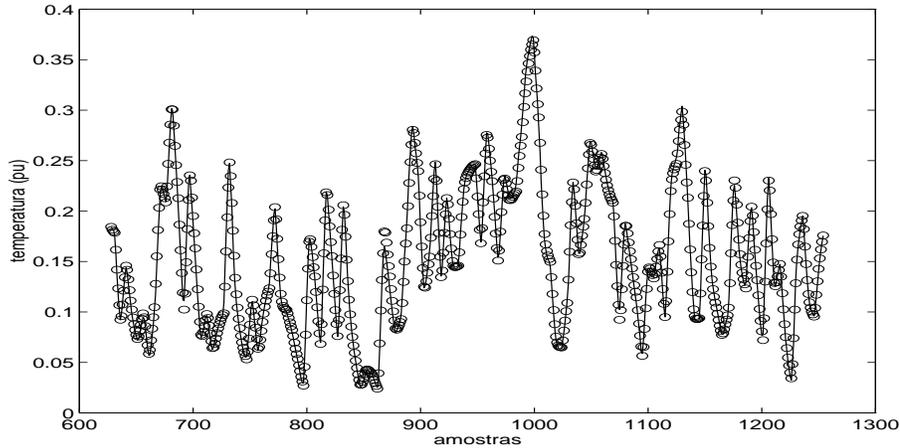


Figura 3.42: Validação dinâmica para o modelo não linear (3.15) identificado com a massa de dados  $\langle \text{din3} \rangle$  e com os agrupamentos  $\Omega_u$ ,  $\Omega_0$ ,  $\Omega_{y^3}$  e  $\Omega_{u^2y}$  excluídos dos agrupamentos candidatos. (—) dados (o) previsões do modelo (3.15).

$$\begin{aligned}
 & -0,5649 u(k-2) y(k-2) y(k-1) \\
 & + \sum_{i=1}^{n_e} \hat{\theta}_i \xi(k-i) + \xi(k) .
 \end{aligned} \tag{3.15}$$

As validações estatísticas para os modelos (3.14) e (3.15), que estão indicadas nas Figuras 3.41 e 3.43, apresentam resultados semelhantes aos dos modelos (3.12) e (3.13), como pode ser observado na Tabela 3.1.

Os modelos (3.14) e (3.15), como os demais modelos NARMAX aqui considerados, são capazes de representar a dinâmica do sistema, conforme indicado pelas validações dinâmicas representadas nas Figuras 3.40 e 3.42. Além disso, apesar de não possuírem a mesma estrutura, os modelos (3.14) e (3.15) são constituídos pelos mesmos tipos de agrupamentos, podendo ser um ponto de partida para a estimação recursiva.

### 3.8.4 Comparação entre os Modelos Apresentados

Observando na Tabela 3.1 os resultados referentes à validação estatística, percebe-se que os resultados piores são para os modelos obtidos através da massa de dados  $\langle \text{din2} \rangle$ , possivelmente, devido às autocorrelações nos atrasos 20 e 30 da entrada (Figura 3.8). Além disso, a superamostragem dos

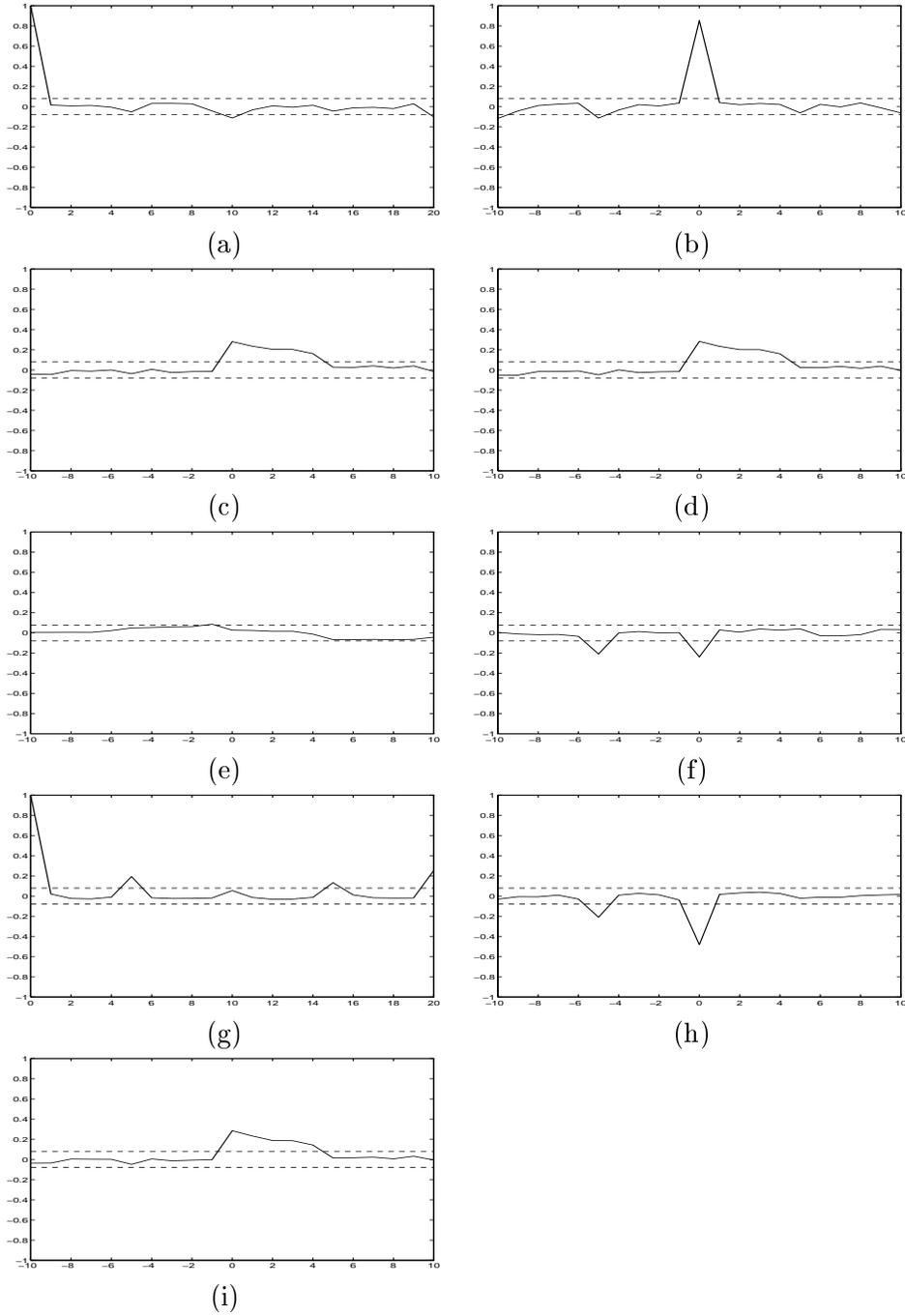


Figura 3.43: Validação estatística para o modelo não linear (3.15) identificado com a massa de dados  $\langle din3 \rangle$  e com os agrupamentos  $\Omega_u$ ,  $\Omega_o$ ,  $\Omega_{y^3}$  e  $\Omega_{u^2y}$  excluídos dos agrupamentos candidatos. (a)  $\phi_{e'e'}$  (b)  $\phi_{(ue)'}e'$  (c)  $\phi_{u'e'}$  (d)  $\phi_{u'^2e'}$  (e)  $\phi_{u'^2e'^2}$  (f)  $\phi_{e'^2e'}$  (g)  $\phi_{e'^2e'^2}$  (h)  $\phi_{e'^2(ye)'}$  (i)  $\phi_{u'^2(ye)'}$ .

dados também influencia na autocorrelação dos resíduos e correlação cruzada entre resíduos e regressores.

No entanto, a validação baseada nas funções de correlação é um critério estatístico, que verifica a hipótese de que os resíduos sejam brancos, mas que nem sempre garante que o modelo seja eficaz dinamicamente. Sendo assim, estes pequenos problemas detectados pelas funções de correlação serão desconsiderados no que diz respeito à escolha do modelo.

O objetivo é apenas verificar se o modelo é capaz de representar a dinâmica do sistema e de recuperar a característica estática não linear, o que será discutido no próximo capítulo.

Observando agora os resultados referentes à validação dinâmica percebe-se que, exceto para os modelos ARMAX, os resultados obtidos foram igualmente satisfatórios.

Vale mencionar que, além dos modelos ARMAX apresentados, foram simulados outros modelos variando-se os atrasos máximos da entrada e saída do sistema. No entanto, os resultados obtidos não foram melhores dos que os aqui apresentados.

No caso de modelos não lineares com grau de não linearidade três, foram utilizados quinze termos, para que as características estáticas obtidas tanto para a massa de dados  $\langle \text{din2} \rangle$ , quanto para a  $\langle \text{din3} \rangle$ , fossem satisfatórias. Além disso, esta foi a condição na qual os modelos com grau de não linearidade três, para ambas as massas de dados, possuem os mesmos tipos de agrupamentos.

A Tabela 3.2 apresenta os agrupamentos candidatos ao modelo e os selecionados. Conforme apresentado neste capítulo, no caso dos modelos não lineares, uma vez definido o grau de não linearidade  $l$ , foram determinados modelos incluindo todos os agrupamentos dentro do conjunto dos candidatos. Em seguida, foram eliminados aqueles agrupamentos considerados espúrios de acordo com a análise de agrupamentos de termos e o agrupamento  $\Omega_{y^l}$ , de forma a facilitar a recuperação da característica estática, que será discutida no próximo capítulo.

Observando os agrupamentos selecionados para os modelos (3.10), (3.11) e (3.14), (3.15) na Tabela 3.2, percebe-se que, a eliminação dos agrupamentos espúrios e do agrupamento  $\Omega_{y^l}$ , resultou na obtenção de modelos compostos pelos mesmos tipos de agrupamentos para ambas massas de dados, o que facilitará a determinação da estrutura para a estimação recursiva.

Na identificação em batelada, feita neste capítulo, as seguintes condições foram consideradas como requisitos na escolha de modelos<sup>10</sup>:

<sup>10</sup>Tais requisitos provavelmente seriam diferentes em outras aplicações. Pode-se citar, por

Tabela 3.1: Tabela comparativa entre os modelos apresentados em relação à ordem do modelo ( $n_y$ ), atraso máximo em  $u$  ( $n_u$ ), grau de não linearidade ( $l$ ), número de termos de processo ( $n_p$ ), número de termos candidatos ( $M$ ) e validações estatística e dinâmica do modelo considerado. A validação estatística foi analisada com base no número de pontos fora do intervalo de confiança para os atrasos de  $-5$  a  $5$  ou de  $0$  a  $5$ . O índice *RMSE* (*Root Mean Square Error*) verifica a capacidade de predição do modelo em relação ao modelo linear da referida massa de dados, sendo dado por:

$$RMSE = \left[ \sum_{i=1}^N (y_i - \hat{y}_i)^2 \right]^{\frac{1}{2}} / \left[ \sum_{i=1}^N (y_i - \hat{y}_{iL})^2 \right]^{\frac{1}{2}}.$$

Modelo	Massa Dados	$n_y$	$n_u$	$l$	$n_p$	$M$	Val. Est.	<i>RMSE</i>
(3.6)	< <i>din2</i> >	2	2	1	4	4	42	1
(3.7)	< <i>din3</i> >	2	2	1	4	4	36	1
(3.8)	< <i>din2</i> >	2	2	2	7	15	32	0,3557
(3.9)	< <i>din3</i> >	2	2	2	7	15	31	0,2431
(3.10)	< <i>din2</i> >	2	2	2	7	9	41	0,3657
(3.11)	< <i>din3</i> >	2	2	2	7	9	26	0,2363
(3.12)	< <i>din2</i> >	2	2	3	15	35	26	0,3369
(3.13)	< <i>din3</i> >	2	2	3	15	35	18	0,2512
(3.14)	< <i>din2</i> >	2	2	3	15	22	28	0,3396
(3.15)	< <i>din3</i> >	2	2	3	15	22	19	0,2365

Tabela 3.2: Tabela comparativa entre os modelos apresentados ilustrando os agrupamentos candidatos ao modelo e os selecionados.

Modelo	$l$	Agrup. Candidatos	Agrup. Selecionados
(3.6)	1	Todos, exceto $\Omega_0$	$\Omega_y, \Omega_u$
(3.7)	1	Todos, exceto $\Omega_0$	$\Omega_y, \Omega_u$
(3.8)	2	Todos	$\Omega_y, \Omega_u, \Omega_{u^2}, \Omega_{y^2}$
(3.9)	2	Todos	$\Omega_y, \Omega_u, \Omega_{u^2}, \Omega_{y^2}, \Omega_{uy}$
(3.10)	2	Todos, exceto $\Omega_u, \Omega_0, \Omega_{y^2}$	$\Omega_y, \Omega_{uy}, \Omega_{u^2}$
(3.11)	2	Todos, exceto $\Omega_u, \Omega_0, \Omega_{y^2}$	$\Omega_y, \Omega_{uy}, \Omega_{u^2}$
(3.12)	3	Todos	$\Omega_y, \Omega_u, \Omega_{u^2}, \Omega_{y^2}, \Omega_{uy},$ $\Omega_{u^3}, \Omega_{u^2y}$
(3.13)	3	Todos	$\Omega_y, \Omega_{u^2}, \Omega_{y^2}, \Omega_{uy}, \Omega_{u^3},$ $\Omega_{y^3}, \Omega_{uy^2}$
(3.14)	3	Todos, exceto $\Omega_0, \Omega_u, \Omega_{y^3}, \Omega_{u^2y}$	$\Omega_y, \Omega_{u^2}, \Omega_{y^2}, \Omega_{uy},$ $\Omega_{u^3}, \Omega_{uy^2}$
(3.15)	3	Todos, exceto $\Omega_0, \Omega_u, \Omega_{y^3}, \Omega_{u^2y}$	$\Omega_y, \Omega_{u^2}, \Omega_{y^2}, \Omega_{uy},$ $\Omega_{u^3}, \Omega_{uy^2}$

- O modelo deve ser capaz de representar a dinâmica do sistema, e
- O modelo deve ser capaz de recuperar a característica estática não linear, o que será discutido no próximo capítulo.

No entanto, visando a estimação recursiva, além do citado acima, os modelos identificados para as massas de dados  $\langle \text{din2} \rangle$  e  $\langle \text{din3} \rangle$ , devem ser analisados em conjunto e devem ser observadas as seguintes questões:

- A obtenção de modelos para as massas de dados  $\langle \text{din2} \rangle$  e  $\langle \text{din3} \rangle$  com os mesmos tipos de agrupamentos, facilita a escolha da estrutura para a estimação recursiva;
- É desejável que a estrutura de modelo escolhida, seja eficaz na recuperação da característica estática, tanto para a massa de dados  $\langle \text{din2} \rangle$ , quanto para  $\langle \text{din3} \rangle$ . Espera-se que a estrutura escolhida seja válida também para a massa de dados que será utilizada na estimação recursiva.

### 3.9 Conclusões

O critério  $AIC(2)$  sugere um modelo com número de termos em torno de cinco. Os modelos obtidos para grau de não linearidade igual a dois estão em torno deste valor. No entanto, para modelos com grau de não linearidade igual a três, modelos com número de termos em torno de cinco, mostraram-se insatisfatórios. Foram simulados modelos com 10 termos. No entanto, embora a recuperação da característica estática, que será discutida no capítulo seguinte, tenha sido satisfatória, os modelos possuíam agrupamentos diferentes, o que dificultaria a estimação recursiva. Modelos com quinze termos foram satisfatórios tanto em relação à validação dinâmica quanto à recuperação da característica estática não linear. Além disso, apresentaram os mesmos tipos de agrupamentos para as massas de dados  $\langle \text{din2} \rangle$  e  $\langle \text{din3} \rangle$ .

Com base nas validações dinâmicas obtidas para os modelos lineares para as massas de dados  $\langle \text{din2} \rangle$  e  $\langle \text{din3} \rangle$ , pode-se dizer que modelos ARMAX são insuficientes para prever a saída do sistema para faixas mais distantes do ponto de operação no qual o teste foi feito. Uma das razões pode ser que, segundo a análise de agrupamentos de termos, os agrupamentos  $\Omega_{y^2}$  e  $\Omega_{u^2}$

---

exemplo, a predição em sistemas reais. Para esta aplicação não é necessário o conhecimento da característica estática não linear do sistema.

parecem ser efetivos, e no entanto, não constam dos agrupamentos candidatos de um modelo ARMAX.

Os modelos NARMAX com grau de não linearidade igual a três não apresentam vantagens relevantes, do ponto de vista da dinâmica do sistema, em relação aos modelos com grau de não linearidade igual a dois. No entanto, deve-se ter em mente que o objetivo final do modelo é a estimação das não linearidades existentes no sistema original.

A análise de agrupamentos de termos aqui utilizada tem por finalidade aumentar a confiabilidade do modelo no que diz respeito à recuperação das características estáticas não lineares. Além disso, facilitará a escolha de uma estrutura para a estimação recursiva. Isto pode ser verificado observando que os modelos NARMAX (3.10) e (3.11), (3.14) e (3.15) possuem os mesmos tipos de agrupamentos, o que não ocorre para os demais casos.

## Capítulo 4

# Estimação de Não Linearidades Estáticas

### 4.1 Introdução

Dentre as representações para sistemas com características estáticas não lineares, os modelos por blocos orientados são frequentemente utilizados (Çinar, 1995). Tais modelos são também denominados modelos híbridos, por serem compostos por submodelos dinâmicos lineares e estáticos não lineares. Neste caso, a identificação de um sistema não linear se restringe à identificação de submodelos lineares e utilização de conhecimento *a priori* da característica estática não linear.

A característica estática não linear que será utilizada nos modelos híbridos, pode ser obtida através de testes estáticos, ou por uma aproximação polinomial, onde fixa-se a estrutura e estimam-se os parâmetros.

Os submodelos lineares dos modelos híbridos, por sua vez, podem ser determinados através de técnicas de correlação associadas à estimação de *kernels* de Volterra (Billings and Fakhouri, 1982). Além disso, pode-se também utilizar técnicas não paramétricas, como por exemplo, métodos no domínio da frequência, respostas ao degrau, ou ainda, as técnicas de correlação descritas em (Billings and Fakhouri, 1982) estendidas a modelos discretos (Haber and Unbehauen, 1990).

Desta forma, embora modelos híbridos apresentem a vantagem de fornecer uma visão da forma das não linearidades presentes no sistema, este tipo de representação apresenta alguns inconvenientes:

- a identificação dos submodelos lineares utilizando *kernels* de Volterra é, normalmente, difícil;

- técnicas não paramétricas nem sempre são amigáveis;
- para sistemas nos quais a característica estática varia ao longo do tempo, a utilização deste tipo de representação para estimação recursiva dos parâmetros do modelo, exigiria conhecimento *a priori* da característica estática não linear em todos os instantes de tempo.

No entanto, quando utilizam-se modelos polinomiais NARMAX, o procedimento para a estimação da característica estática não linear é diferente. Embora conhecimento *a priori* possa ser utilizado para auxiliar na determinação da estrutura do modelo, esta não é uma condição estritamente necessária (Aguirre et al., 1998a; Aguirre et al., 1998b). Desta forma, para um sistema variante no tempo, pode-se utilizar informação *a priori* sobre as não linearidades estáticas de interesse para um determinado instante de tempo e as demais curvas podem ser estimadas através do modelo obtido.

Sendo assim, se a estrutura escolhida para o modelo for adequada (Jácome, 1996), é possível estimar, de forma eficiente e simples, as características estáticas de interesse utilizando os conceitos de agrupamentos de termos descritos no capítulo 2, e de pontos fixos, que serão descritos a seguir. Desta forma, quando o objetivo final é estimar as não linearidades estáticas presentes no sistema original, modelos polinomiais NARMAX parecem ser bastante adequados.

O capítulo está organizado como se segue: a seção 4.2 apresenta definição e localização de pontos fixos em sistemas autônomos, a seção 4.3 estende os conceitos de pontos fixos para sistemas não autônomos e a seção 4.4 discute estabilidade dos pontos fixos. Finalmente são apresentadas as características estáticas obtidas através dos modelos discutidos no capítulo 3.

## 4.2 Pontos Fixos em Sistemas Autônomos

Os pontos fixos de um mapeamento discreto<sup>1</sup> são definidos como os pontos nos quais  $y(k) = y(k + i)$ ,  $i \in \mathcal{Z}$  e correspondem, na maioria dos casos, a pontos de partida para a análise de sistemas não lineares (Wiggins, 1990).

Enquanto modelos lineares possuem apenas um ponto fixo, o número de pontos fixos de um modelo polinomial NARMAX depende do grau de não linearidade dos regressores da saída e pode ser determinado através da análise de agrupamentos de termos e coeficientes de agrupamentos (Aguirre and Mendes, 1996).

---

<sup>1</sup>No caso de modelos contínuos, os pontos fixos são, por definição, as soluções de  $\dot{x} = f(x) = 0$  (Thompson and Stewart, 1986).

Um modelo autônomo com grau de não linearidade  $l$  pode ser analisado em regime estático em função de agrupamentos de termos e coeficientes de agrupamentos, da seguinte forma:

$$y(k) \approx \theta_{0,0} + y(k) \sum_{n_1=1}^{n_y} \theta_{1,0}(n_1) + y(k)^2 \sum_{n_1, n_2}^{n_y, n_y} \theta_{2,0}(n_1, n_2) + \dots \\ + y(k)^l \sum_{n_1, n_l}^{n_y, n_y} \theta_{l,0}(n_1, \dots, n_l) . \quad (4.1)$$

Eliminando o argumento  $k$ , a equação (4.1) pode ser descrita como se segue:

$$\Sigma_{y^l} y^l + \dots + \Sigma_{y^2} y^2 + (\Sigma_y - 1)y + \Sigma_0 = 0 , \quad (4.2)$$

onde  $\Sigma_0 = \theta_{0,0}$  é uma constante, e  $\Sigma_x$  é o coeficiente do agrupamento  $\Omega_x$ .

Os pontos fixos do modelo correspondem às raízes de (4.2). Desta forma, para um modelo polinomial autônomo com grau de não linearidade  $l$ , o número de pontos fixos é igual a  $l$  se  $\Sigma_{y^l} \neq 0$ .

Se, no entanto,  $\Sigma_0 = 0$ , a equação (4.2) pode ser expressa na forma que se segue:

$$[\Sigma_{y^l} y^{l-1} + \dots + \Sigma_{y^2} y + (\Sigma_y - 1)]y = 0 . \quad (4.3)$$

Neste caso, o modelo agrupado na forma da equação (4.3) possui um ponto fixo trivial e  $l - 1$  pontos fixos não triviais.

A localização dos pontos fixos de um modelo polinomial de um sistema autônomo, pode ser descrita em termos de coeficientes de agrupamentos  $\Sigma_0, \Sigma_{y^i}, i = 1, \dots, l$  na forma apresentada pela equação (4.2) (Aguirre and Mendes, 1996).

A localização do ponto fixo de um modelo linear representado na forma da equação (4.4):

$$(\Sigma_y - 1)y + \Sigma_0 = 0 , \quad (4.4)$$

é dada por:

$$\bar{y} = \frac{\Sigma_0}{1 - \Sigma_y} . \quad (4.5)$$

De forma similar, a localização dos pontos fixos de um modelo quadrático, cuja representação agrupada é:

$$\Sigma_{y^2} y^2 + (\Sigma_y - 1)y + \Sigma_0 = 0 \quad , \quad (4.6)$$

pode ser expressa da seguinte forma:

$$\bar{y}_{1,2} = \frac{1 - \Sigma_y \pm \sqrt{\Delta}}{2\Sigma_{y^2}} \quad , \quad (4.7)$$

onde  $\Delta = (\Sigma_y - 1)^2 - 4\Sigma_{y^2}\Sigma_0$ . É importante salientar que, se  $\Sigma_{y^2} = 0$ , existirá somente um ponto fixo dado pela equação (4.5).

### 4.3 Pontos Fixos em Sistemas Não Autônomos

Estendendo a equação (4.1) para sistemas não autônomos, tem-se:

$$y(k) \approx \sum_{n_1, n_{m_l}}^{n_y, n_u} \theta_{p, m_l - p}(n_1, \dots, n_{m_l}) \sum_{m_i=0}^l \sum_{p=0}^{m_i} y(k)^p u(k)^{m_i - p} \quad , \quad (4.8)$$

onde  $m_l$  corresponde ao grau de não linearidade dos termos e que varia na faixa de  $1 \leq m_l \leq l$ . Cada termo de grau  $m_l$  pode conter um fator em  $y(k)$  de ordem  $p$ , um fator em  $u(k)$  de ordem  $(m_l - p)$  e um coeficiente  $\theta_{p, m_l - p}(n_1, \dots, n_{m_l})$ .

A equação (4.8) pode ser reescrita como:

$$\begin{aligned} \Sigma_{y^l} y^l + \sum_{m_i=l-1}^l [\Sigma_{u^{m_i-(l-1)} y^{l-1}} u^{m_i-(l-1)}] y^{l-1} + \dots \\ + \sum_{m_i=p}^l [\Sigma_{u^{m_i-p} y^p} u^{m_i-p}] y^p + \dots + \sum_{m_i=1}^l [\Sigma_{u^{m_i-1} y} u^{m_i-1} - 1] y + \\ \sum_{m_i=1}^l \Sigma_{u^{m_i}} u^{m_i} + \Sigma_0 = 0 \quad . \end{aligned} \quad (4.9)$$

O procedimento para determinação do número de pontos fixos em sistemas não autônomos é idêntico ao procedimento para seu correspondente autônomo. No entanto, a localização dos mesmos apresenta algumas diferenças.

Para um modelo linear representado na forma que se segue:

$$(\Sigma_y - 1)y + \Sigma_u u + \Sigma_0 = 0 . \quad (4.10)$$

O ponto fixo é dado por:

$$\bar{y} = \frac{\Sigma_0 + \Sigma_u u}{1 - \Sigma_y} . \quad (4.11)$$

Da mesma forma, para um modelo quadrático na forma da equação (4.12):

$$\Sigma_{y^2} y^2 + (\Sigma_y - 1 + \Sigma_{uy} u)y + \Sigma_{u^2} u^2 + \Sigma_u u + \Sigma_0 = 0 . \quad (4.12)$$

Os pontos fixos são dados por:

$$\bar{y}_{1,2} = \frac{1 - \Sigma_y - \Sigma_{uy} u \pm \sqrt{\Delta}}{2\Sigma_{y^2}} , \quad (4.13)$$

onde  $\Delta = (\Sigma_y - 1 + \Sigma_{uy} u)^2 - 4\Sigma_{y^2}(\Sigma_{u^2} u^2 + \Sigma_u u + \Sigma_0)$ .

Se, no entanto,  $\Sigma_{y^2} = 0$ , existirá somente um ponto fixo determinado pela equação (4.14):

$$\bar{y} = \frac{\Sigma_0 + \Sigma_u u + \Sigma_{u^2} u^2}{1 - \Sigma_y - \Sigma_{uy} u} . \quad (4.14)$$

Para modelos com variáveis de entrada, a localização dos pontos fixos dependerá do valor correspondente da entrada do sistema, observar equações (4.11), (4.13) e (4.14), caracterizando uma “curva” estática. Desta forma, pode-se estimar a característica estática para um sistema não autônomo a partir de um modelo dinâmico utilizando-se os conceitos de agrupamentos de termos e pontos fixos.

## 4.4 Estabilidade de Pontos Fixos

Seja um mapeamento  $f$  e um ponto fixo  $p$  deste mapeamento. Se todos os pontos localizados na vizinhança  $N_\epsilon(p)^2$  de  $p$  são atraídos para  $p$ , então  $p$  é um atrator. Se, no entanto, os pontos situados na vizinhança de  $p$  são repelidos de  $p$ , então  $p$  é um repulsor (Alligood et al., 1997).

---

<sup>2</sup> $N_\epsilon(p)$  é o intervalo de números  $\{x \in \mathcal{R} : |x - p| < \epsilon\}$ , onde  $\epsilon$  é positivo.

Pontos fixos são também denominados pontos de equilíbrio e sua estabilidade influi na estabilidade do sistema. Se um ponto fixo é um atrator então ele é um ponto de equilíbrio estável. Ao passo que, se um ponto fixo for um repulsor, então tal ponto corresponde a um ponto de equilíbrio instável (Thompson and Stewart, 1986).

Considerando somente a parte autônoma<sup>3</sup> do modelo NARX<sup>4</sup> mono-variável,  $y(k) = f(y(k-i))$   $i = 1, \dots, n_y$ , e definindo o seguinte sistema matricial:

$$\begin{bmatrix} y(k-n_y+1) \\ y(k-n_y+2) \\ \vdots \\ y(k) \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ \frac{f_1}{y(k-n_y)} & \frac{f_2}{y(k-n_y+1)} & \frac{f_3}{y(k-n_y+2)} & \cdots & \frac{f_{n_y}}{y(k-1)} \end{bmatrix} \begin{bmatrix} y(k-n_y) \\ y(k-n_y+1) \\ \vdots \\ y(k-1) \end{bmatrix}, \quad (4.15)$$

onde  $\sum_{i=1}^{n_y} f_i = f$  e sendo  $f_i/y(k-i) = 0$  caso o fator  $y(k-i)$  não faça parte do mapeamento  $f(\cdot)$ . O Jacobiano de  $f$  é definido como:

$$Df = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \\ \frac{\partial f}{\partial y(k-n_y)} & \frac{\partial f}{\partial y(k-n_y+1)} & \frac{\partial f}{\partial y(k-n_y+2)} & \cdots & \frac{\partial f}{\partial y(k-1)} \end{bmatrix} \quad (4.16)$$

A estabilidade dos pontos fixos do mapeamento  $f$  pode ser analisada determinando-se o Jacobiano definido por (4.16) em cada ponto fixo e calcu-

<sup>3</sup>O primeiro passo para a análise de estabilidade de um sistema não autônomo é a determinação dos pontos fixos de seu equivalente autônomo. Isto é feito considerando  $u(k-i) = 0$ ,  $i = 1, \dots, n_u$  (Aguirre and Billings, 1995b).

<sup>4</sup>Somente a parte determinística do modelo NARMAX é utilizada para analisar a estabilidade dos pontos fixos. A parte estocástica tem por finalidade evitar que os parâmetros estimados sejam polarizados (Aguirre and Billings, 1995b), sendo utilizada neste trabalho, nos processos de detecção de estrutura e estimação de parâmetros e desconsiderada para os demais fins.

lando os seus respectivos autovalores. Caso os autovalores sejam estáveis<sup>5</sup>, o ponto fixo é um atrator, caso contrário o mesmo é um repulsor (Aguirre and Billings, 1995b). Se existirem pelo menos um autovalor instável e um estável relacionados ao mesmo ponto fixo, tal ponto é denominado de sela.

## 4.5 Estimação de Não Linearidades Estáticas Utilizando Modelos Polinomiais

As características estáticas não lineares serão estimadas com base nos modelos encontrados no capítulo 3 e na localização dos pontos fixos para modelos não autônomos, conforme descrito na seção 4.3.

### 4.5.1 Estimação de Características Estáticas a partir de Modelos Lineares

Os modelos ARMAX (3.6) e (3.7) agrupados estão representados nas equações (4.17) e (4.18), respectivamente.

$$y = 0,9521 y + 0,0334 u. \quad (4.17)$$

$$y = 0,9201 y + 0,0263 u. \quad (4.18)$$

De acordo com a equação (4.11) a característica estática obtida para o modelo (4.17) é dada por:

$$\bar{y} = \frac{0,0334 u}{0,0479}. \quad (4.19)$$

Da mesma forma, a característica estática obtida para o modelo (4.18) é:

$$\bar{y} = \frac{0,0263 u}{0,0799}. \quad (4.20)$$

As características estáticas indicadas nas Figuras 3.3 e 3.4, bem como a estimada através da equação (4.19), estão representadas na Figura 4.1. Assim

---

<sup>5</sup>Um autovalor de um sistema discreto é estável se estiver localizado dentro do círculo de raio unitário centrado na origem no plano  $\mathcal{Z}$ .

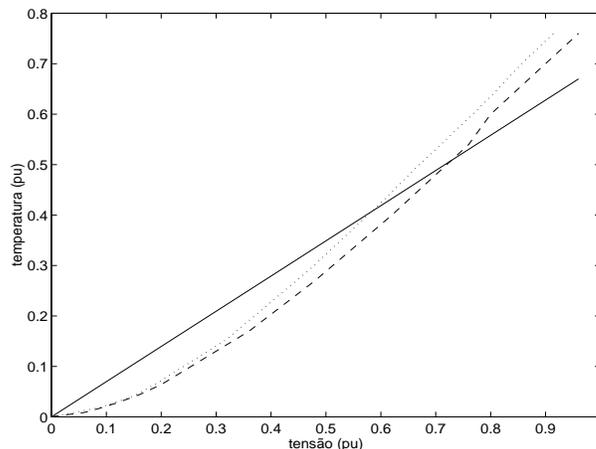


Figura 4.1: (—) Característica estática obtida a partir do modelo linear (3.6) encontrado com a massa de dados  $\langle din2 \rangle$ , (- -) Figura 3.3, (···) Figura 3.4.

como, a característica estimada através da equação (4.20) está representada na Figura 4.2. Como esperado, as características estáticas dos modelos lineares correspondem a uma “média” da característica estática não linear do sistema.

Desta forma, baseando-se nas características estáticas obtidas, Figuras 4.1 e 4.2, bem como nas validações dinâmicas, Figuras 3.24 e 3.26, pode-se concluir que modelos ARMAX são insatisfatórios para descrever as não linearidades presentes no sistema.

#### 4.5.2 Estimação de Características Estáticas a partir de Modelos Quadráticos

Expressando o modelo NARMAX (3.8) em termos de agrupamentos e coeficientes de agrupamentos, tem-se:

$$y = 0,9642 y - 0,0523 y^2 + 0,0045 u + 0,0548 u^2 . \quad (4.21)$$

Conforme equação (4.13) a característica estática para o modelo (4.21) será:

$$\bar{y}_{1,2} = \frac{-0,0358 \pm \sqrt{\Delta}}{0,1046} , \quad (4.22)$$

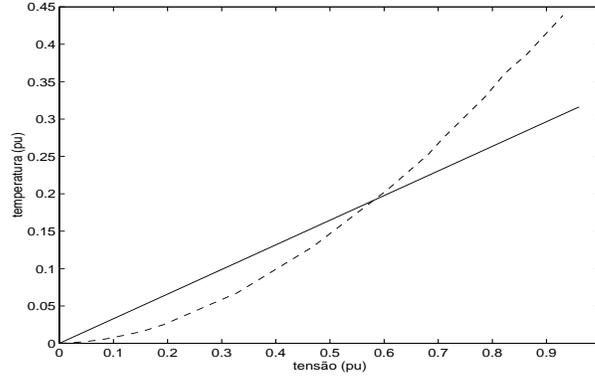


Figura 4.2: (—) Característica estática obtida a partir do modelo linear (3.7) encontrado com a massa de dados  $\langle \text{din3} \rangle$ , (---) Figura 3.5.

onde  $\Delta = 0,0013 + 0,2092(0,0045 u + 0,0548 u^2)$ .

Observa-se que o modelo agrupado (4.21) possui dois pontos fixos, o que resulta em duas curvas estáticas distintas. As características estáticas para os pontos fixos instável e estável estão representadas nas Figuras 4.3 e 4.4, juntamente com as características indicadas nas Figuras 3.3 e 3.4. No entanto, como um dos pontos fixos é instável, somente a característica estática não linear estimada para o ponto fixo estável deve ser considerada.

Da mesma forma, o modelo NARMAX (3.9) em termos de agrupamentos e coeficientes de agrupamentos é dado por:

$$y = 0,9101 y - 0,0191 y^2 + 0,0030 u + 0,0479 u^2 - 0,0091 uy \quad , \quad (4.23)$$

e a expressão da sua característica estática será igual a:

$$\overline{y}_{1,2} = \frac{-(0,0899 + 0,0091 u) \pm \sqrt{\Delta}}{0,0382} \quad , \quad (4.24)$$

onde  $\Delta = (0,0899 + 0,0091 u)^2 + 0,0764(0,0030 u + 0,0479 u^2)$ .

As características estáticas não lineares obtidas para a massa de dados  $\langle \text{din3} \rangle$ , para os pontos fixos instável e estável da equação (4.23), estão representadas nas Figuras 4.5 e 4.6.

Com base nas características estimadas, pode-se afirmar que, em relação aos modelos ARMAX, os modelos NARMAX com grau de não linearidade

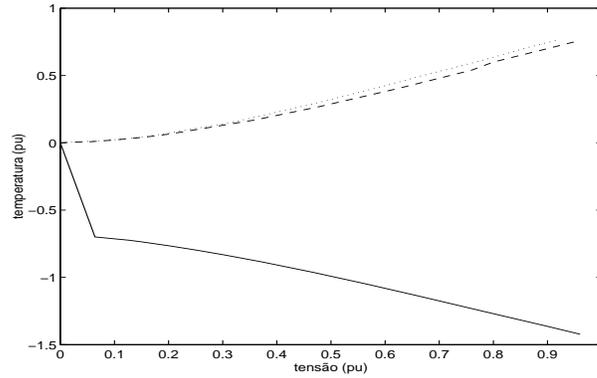


Figura 4.3: (—) Característica estática obtida a partir do modelo NARMAX (3.8) em relação ao ponto fixo 1 encontrado com a massa de dados  $\langle din2 \rangle$ , (--) Figura 3.3, ( $\cdots$ ) Figura 3.4.

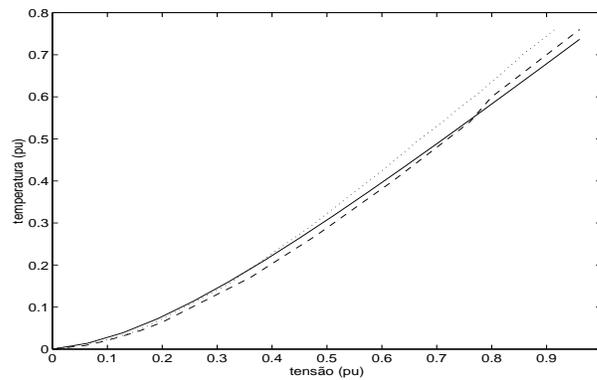


Figura 4.4: (—) Característica estática obtida a partir do modelo NARMAX (3.8) em relação ao ponto fixo 2 encontrado com a massa de dados  $\langle din2 \rangle$ , (--) Figura 3.3, ( $\cdots$ ) Figura 3.4.

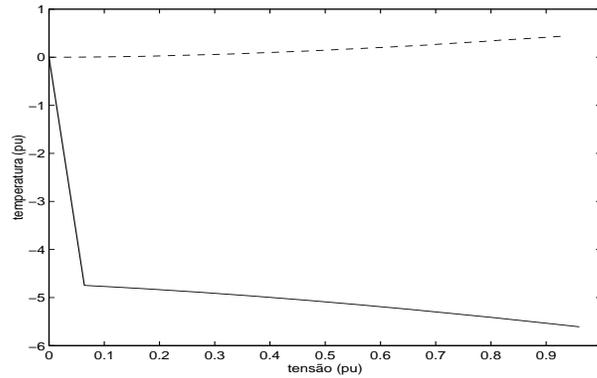


Figura 4.5: (—) Característica estática obtida para o ponto fixo 1 a partir do modelo NARMAX (3.9) encontrado com a massa de dados  $\langle din3 \rangle$ , (---) Figura 3.5.

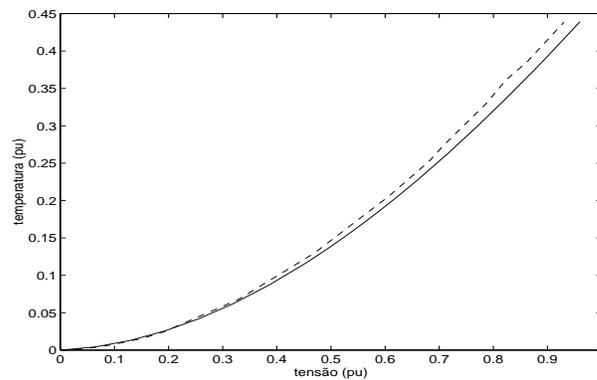


Figura 4.6: (—) Característica estática obtida para o ponto fixo 2 a partir do modelo NARMAX (3.9) encontrado com a massa de dados  $\langle din3 \rangle$ , (---) Figura 3.5.

igual a dois são muito mais eficientes no que diz respeito à estimação das não linearidades estáticas contidas nos dados experimentais.

Como mencionado no capítulo 3, os agrupamentos  $\Omega_0$  e  $\Omega_u$  foram excluídos dos agrupamentos candidatos, conforme sugerido pela análise de agrupamentos de termos. Observou-se que modelos sem os agrupamentos mencionados proporcionavam a estimação de uma característica estática mais fiel aos dados.

Além disso, a fim de verificar se modelos com apenas um ponto fixo são capazes de representar a dinâmica do sistema e estimar a característica estática, o agrupamento  $\Omega_{y^2}$  também foi excluído do conjunto dos candidatos, obtendo-se os modelos agrupados (4.25) e (4.26) referentes aos modelos (3.10) e (3.11), respectivamente.

$$y = 0,9685 y + 0,0769 u^2 - 0,0632 uy \quad . \quad (4.25)$$

$$y = 0,9150 y + 0,0559 u^2 - 0,0323 uy \quad . \quad (4.26)$$

De acordo com a equação (4.14) a característica estática para o modelo (4.25) é dada por:

$$\bar{y} = \frac{0,0769 u^2}{0,0315 + 0,0632 u} \quad , \quad (4.27)$$

e para o modelo (4.26) é:

$$\bar{y} = \frac{0,0559 u^2}{0,0850 + 0,0323 u} \quad . \quad (4.28)$$

As características estáticas obtidas através das equações (4.27) e (4.28) estão representadas nas Figuras 4.7 e 4.8, respectivamente.

Observando a Figura 4.7 percebe-se que o modelo (3.10) aproxima com mais eficiência a característica descrita na Figura 3.3 do que a descrita na Figura 3.4. Isto ocorre pois há diferença entre a temperatura ambiente dos dois testes. Este fato, como já foi dito, vem tornar mais evidente a necessidade de uma estimação recursiva dos parâmetros dos modelos e, por consequência, das não linearidades estáticas.

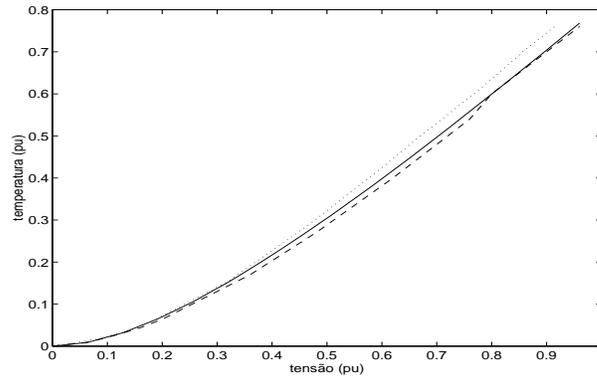


Figura 4.7: (—) Característica estática obtida a partir do modelo NARMAX (3.10) encontrado com a massa de dados  $\langle din2 \rangle$ , (---) Figura 3.3, ( $\cdots$ ) Figura 3.4.

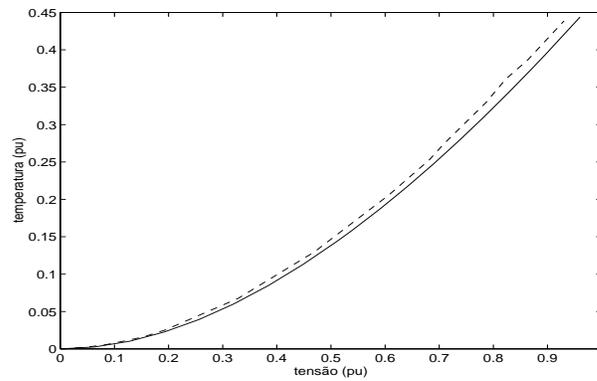


Figura 4.8: (—) Característica estática obtida a partir do modelo NARMAX (3.11) encontrado com a massa de dados  $\langle din3 \rangle$ , (---) Figura 3.5.

### 4.5.3 Estimação de Características Estáticas a partir de Modelos Cúbicos

Em regime estático o modelo NARMAX (3.12) pode ser descrito na seguinte forma:

$$y = 0,9640 y - 0,0679 y^2 + 0,0047 u + 0,0182 u^2 + 0,0563 u^3 + 0,0599 uy - 0,0733 u^2 y. \quad (4.29)$$

A característica estática pode ser obtida, resolvendo a equação (4.29) como:

$$\bar{y}_{1,2} = \frac{-(0,0360 - 0,0599 u + 0,0733 u^2) \pm \sqrt{\Delta}}{0,1358}, \quad (4.30)$$

onde  $\Delta = (0,0360 - 0,0599 u + 0,0733 u^2)^2 + 0,2716(0,0047 u + 0,0182 u^2 + 0,0563 u^3)$ .

Observa-se que para a massa de dados  $\langle din2 \rangle$  o algoritmo *ERR* não selecionou o agrupamento  $\Omega_{y^3}$  e portanto, o modelo só possui dois pontos fixos.

A característica estática recuperada através do ponto fixo um, Figura 4.9, do modelo (4.30) é insatisfatória já que este ponto fixo é instável. Da mesma forma, a característica recuperada considerando o ponto fixo estável está indicada na Figura 4.10.

Da mesma forma, o regime estático do modelo NARMAX (3.13) pode ser descrito como:

$$y = 0,9332 y - 0,1770 y^2 + 0,1312 y^3 + 0,0435 u^2 + 0,0046 u^3 + 0,0045 uy + 0,0592 uy^2. \quad (4.31)$$

Como a inserção do agrupamento  $\Omega_{y^3}$  ocasionou o aparecimento de um terceiro ponto fixo, a recuperação da característica estática foi realizada numericamente. Equações que permitem determinar a característica estática não linear para  $l = 3$  de forma analítica podem ser encontradas em (Aguirre and Mendes, 1996). A característica estática que é similar à do sistema está representada na Figura 4.12. Enquanto que, as curvas obtidas para os demais pontos fixos estão representadas nas Figuras 4.11 e 4.13.

Observa-se que a característica estática obtida através do modelo (3.13) aproxima-se muito bem da obtida no teste estático 08 (Figura 3.5) discutido

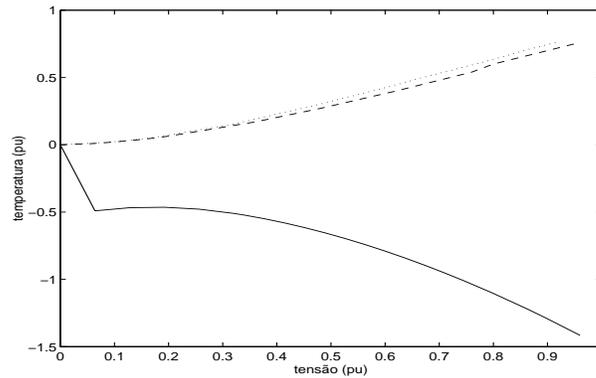


Figura 4.9: (—) Característica estática obtida para o ponto fixo 1 a partir do modelo NARMAX (3.12) com grau de não linearidade três e dois pontos fixos encontrado com a massa de dados  $\langle din2 \rangle$ , (--) Figura 3.3 ( $\cdots$ ) Figura 3.4.

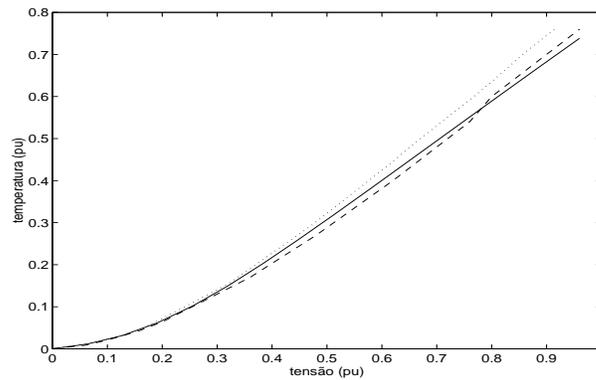


Figura 4.10: (—) Característica estática obtida para o ponto fixo 2 a partir do modelo NARMAX (3.12) com grau de não linearidade três e dois pontos fixos encontrado com a massa de dados  $\langle din2 \rangle$ , (--) Figura 3.3, ( $\cdots$ ) Figura 3.4.

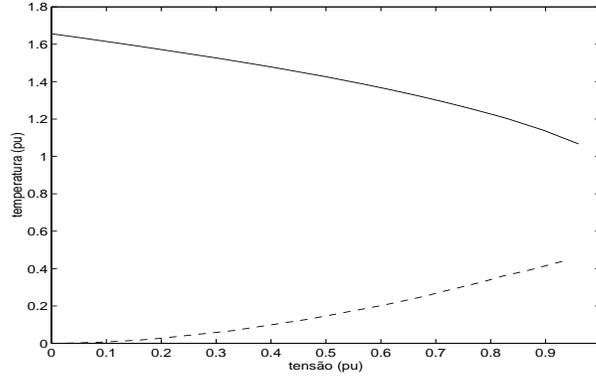


Figura 4.11: (—) Característica estática obtida para o ponto fixo 1 a partir do modelo NARMAX (3.13) encontrado com a massa de dados  $\langle \text{din3} \rangle$ , (- -) Figura 3.5.

na seção 3.3. No entanto, a inclusão do terceiro ponto fixo torna a recuperação analítica da característica estática significativamente mais difícil. De forma a eliminar o terceiro ponto fixo, o agrupamento  $\Omega_{y^3}$  será eliminado do conjunto dos agrupamentos candidatos.

Além disso, para que se possa verificar melhores resultados nas aproximações da curva estática, os agrupamentos  $\Omega_0$ ,  $\Omega_u$ ,  $\Omega_{y^3}$  e  $\Omega_{u^2y}$  foram excluídos dos agrupamentos candidatos, por terem sido considerados espúrios pela análise de agrupamentos de termos, discutida na seção (3.5).

Expressando os modelos NARMAX (3.14) e (3.15) em termos de agrupamentos e coeficientes de agrupamentos, obtem-se as equações (4.32) e (4.33) em regime permanente.

$$y = 0,9737y - 0,0651y^2 + 0,0568u^2 + 0,0070u^3 - 0,0063uy + 0,0032uy^2 . \quad (4.32)$$

$$y = 0,9298y - 0,1504y^2 + 0,0421u^2 + 0,0035u^3 + 0,0136uy + 0,0768uy^2 . \quad (4.33)$$

A localização das raízes de (4.32) e (4.33) é dada por (4.34) e (4.35), respectivamente.

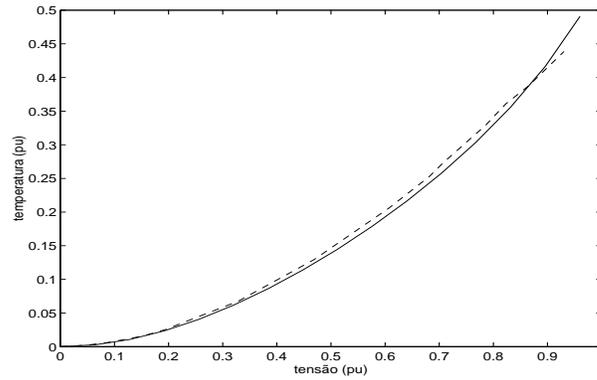


Figura 4.12: (—) Característica estática obtida para o ponto fixo 2 a partir do modelo NARMAX (3.13) encontrado com a massa de dados  $\langle din3 \rangle$ , (---) Figura 3.5.

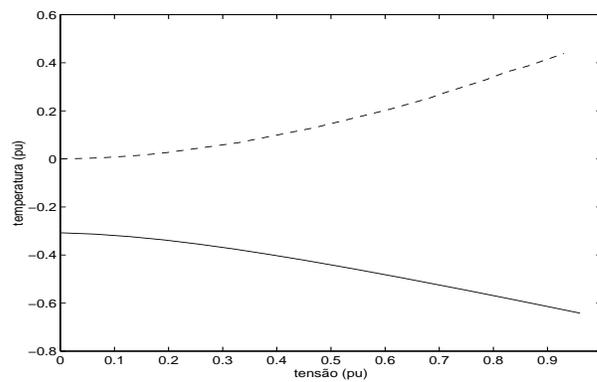


Figura 4.13: (—) Característica estática obtida para o ponto fixo 3 a partir do modelo NARMAX (3.13) encontrado com a massa de dados  $\langle din3 \rangle$ , (---) Figura 3.5.

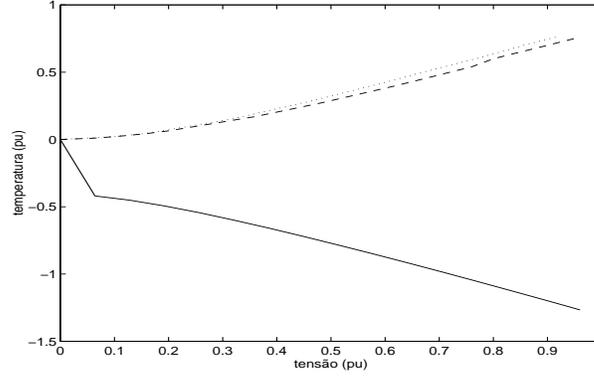


Figura 4.14: (—) Característica estática obtida para o ponto fixo 1 a partir do modelo NARMAX (3.14) com grau de não linearidade três e dois pontos fixos, tendo-se excluído os agrupamentos  $\Omega_0$ ,  $\Omega_u$ ,  $\Omega_{y^3}$  e  $\Omega_{u^2y}$ , encontrado com a massa de dados  $\langle din2 \rangle$ , (---) Figura 3.3, ( $\cdots$ ) Figura 3.4.

$$\bar{y}_{1,2} = \frac{-(0,0263 + 0,0063 u) \pm \sqrt{\Delta}}{2(0,0651 - 0,0032 u)}, \quad (4.34)$$

onde  $\Delta = (0,0263 + 0,0063 u)^2 + 4(0,0651 - 0,0032 u)(0,0568 u^2 + 0,0070 u^3)$ .

$$\bar{y}_{1,2} = \frac{-(0,0702 - 0,0136 u) \pm \sqrt{\Delta}}{2(0,1504 - 0,0768 u)}, \quad (4.35)$$

onde  $\Delta = (0,0702 - 0,0136 u)^2 + 4(0,1504 - 0,0768 u)(0,0421 u^2 + 0,0035 u^3)$ .

As características estimadas para os pontos fixos estáveis das equações (4.34) e (4.35), estão representadas nas Figuras 4.15 e 4.17, respectivamente. Ao passo que, as recuperadas através dos pontos fixos instáveis estão indicadas nas Figuras 4.14 e 4.16.

Como os agrupamentos selecionados para os modelos (3.14) e (3.15) são equivalentes e a estimação das curvas estáticas foram satisfatórias para os dois casos, Figuras 4.15 e 4.17<sup>6</sup>, poderá ser escolhida uma estrutura composta por estes agrupamentos para recuperar a característica estática na estimação recursiva<sup>7</sup>.

<sup>6</sup>A Figura 3.4 não será utilizada para fins de escolha da estrutura do modelo, já que a diferença na temperatura ambiente entre o teste  $\langle din2 \rangle$  e o teste estático 07 são relevantes, impedindo qualquer que seja o modelo de aproximar com eficiência a curva.

<sup>7</sup>Na estimação recursiva não serão realizadas alterações no que diz respeito à estrutura

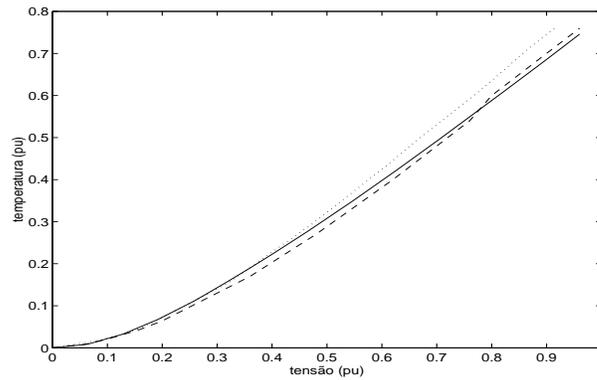


Figura 4.15: (—) Característica estática obtida para o ponto fixo 2 a partir do modelo NARMAX (3.14) com grau de não linearidade três e dois pontos fixos, tendo-se excluído os agrupamentos  $\Omega_0$ ,  $\Omega_u$ ,  $\Omega_{y^3}$  e  $\Omega_{u^2y}$ , encontrado com a massa de dados  $\langle din2 \rangle$ , (--) Figura 3.3, ( $\cdots$ ) Figura 3.4.

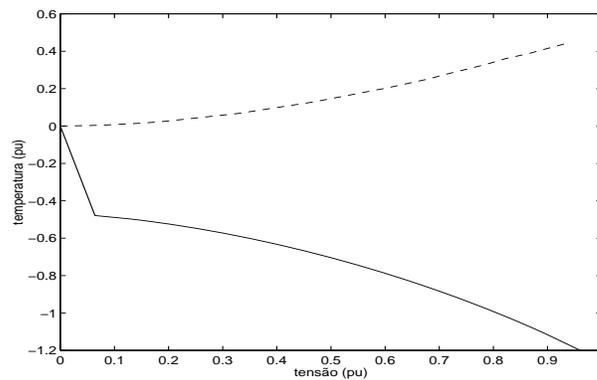


Figura 4.16: (—) Característica estática obtida para o ponto fixo 1 a partir do modelo NARMAX (3.15) com grau de não linearidade três e dois pontos fixos, tendo-se excluído os agrupamentos  $\Omega_0$ ,  $\Omega_u$ ,  $\Omega_{y^3}$  e  $\Omega_{u^2y}$ , encontrado com a massa de dados  $\langle din3 \rangle$ , (--) Figura 3.5.

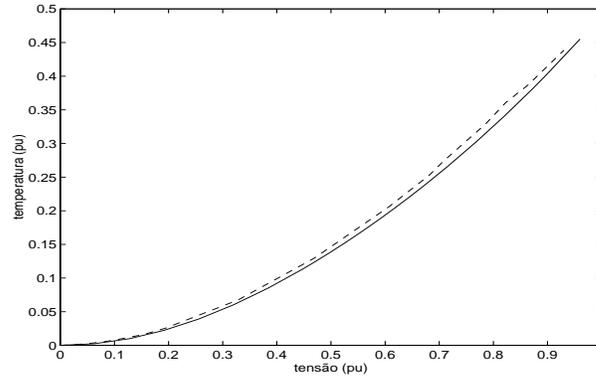


Figura 4.17: (—) Característica estática obtida para o ponto fixo 2 a partir do modelo NARMAX (3.15) com grau de não linearidade três e dois pontos fixos, tendo-se excluído os agrupamentos  $\Omega_0$ ,  $\Omega_u$ ,  $\Omega_{y^3}$  e  $\Omega_{u^2y}$ , encontrado com a massa de dados  $\langle din3 \rangle$ , (---) Figura 3.5.

---

do modelo, só serão variados os parâmetros do mesmo.

Tabela 4.1: Tabela comparativa entre os modelos obtidos no capítulo anterior, em relação ao número de pontos fixos e eficiência na estimação da característica estática não linear. O índice  $\overline{RMSE}$  (*Root Mean Square Error*) verifica a capacidade de recuperação da característica estática não linear do modelo em relação ao modelo linear da referida massa de dados, sendo dado por:  $\overline{RMSE} = [\sum_{i=1}^{N_e} (y_i(\infty) - \bar{y}_i)^2]^{\frac{1}{2}} / [\sum_{i=1}^{N_e} (y_i(\infty) - \bar{y}_{iL})^2]^{\frac{1}{2}}$ , sendo  $y_i(\infty), i = 1, \dots, N_e$  os  $N_e$  valores obtidos nos testes estáticos e  $\bar{y}_i$  os valores em regime permanente obtidos a partir dos modelos. As características obtidas pelos modelos foram comparadas com a Figura 3.3, no caso da massa de dados  $\langle din2 \rangle$  e com a Figura 3.5, no caso da massa de dados  $\langle din3 \rangle$ .

Modelo	Massa Dados	Pontos Fixos	$\overline{RMSE}$
(3.6)	$\langle din2 \rangle$	1	1
(3.7)	$\langle din3 \rangle$	1	1
(3.8)	$\langle din2 \rangle$	2	0,2594
(3.9)	$\langle din3 \rangle$	2	0,2203
(3.10)	$\langle din2 \rangle$	1	0,1804
(3.11)	$\langle din3 \rangle$	1	0,1926
(3.12)	$\langle din2 \rangle$	2	0,2205
(3.13)	$\langle din3 \rangle$	3	0,1387
(3.14)	$\langle din2 \rangle$	2	0,2130
(3.15)	$\langle din3 \rangle$	2	0,1515

## 4.6 Comparação entre as Características Estáticas

A fim de comparar os modelos em relação ao número de pontos fixos e capacidade de recuperação da característica estática não linear será observada a Tabela 4.1.

Na tentativa de avaliar se os modelos (3.10) e (3.11) são mais eficientes do que os modelos (3.8) e (3.9), serão observados os índices  $\overline{RMSE}$  e serão comparadas as Figuras 4.4 e 4.6 com as Figuras 4.7 e 4.8. Nota-se que, as características estáticas obtidas a partir dos modelos (3.10) e (3.11) foram melhores do que as obtidas para os modelos (3.8) e (3.9).

Desta forma, conforme mencionado anteriormente, a eliminação dos agrupamentos espúrios  $\Omega_0, \Omega_u$ , neste caso, resultou na melhora da aproximação da característica estática obtida através dos modelos. Enquanto que, a elimi-

nação do agrupamento  $\Omega_{y^2}$ , facilitou a recuperação analítica da característica estática, já que os modelos só possuem um ponto fixo. Estes resultados estão de acordo com a análise de agrupamentos de termos discutida na seção 3.5.

Sendo assim, embora os modelos (3.10) e (3.11), bem como os modelos (3.8) e (3.9), sejam válidos quanto à representação da dinâmica do sistema, a eliminação de um segundo ponto fixo, no presente caso, pode ser considerada uma vantagem, já que torna a recuperação da característica estática mais fácil.

Observando agora a Tabela 3.2, nota-se que os modelos NARMAX (3.10) e (3.11) com um ponto fixo, assim como os modelos NARMAX (3.14) e (3.15) com dois pontos fixos, possuem os mesmos tipos de agrupamentos, tanto para a massa de dados  $\langle \text{din2} \rangle$ , quanto para a  $\langle \text{din3} \rangle$ . Como este é um fator importante para a identificação recursiva realizada posteriormente, estes modelos serão brevemente comparados.

De acordo com a Tabela 4.1, tem-se que a característica estática estimada com o modelo (3.10) aproxima-se mais da obtida no teste estático 02 do que a estimada com o modelo (3.14). No entanto, a característica obtida com o modelo (3.15) representa melhor a obtida no teste estático 08, do que a obtida pelo modelo (3.11).

Além disso, os modelos (3.10) e (3.11) possuem sete termos, a mesma estrutura e somente um ponto fixo. Enquanto que, os modelos (3.14) e (3.15) possuem quinze termos, dois pontos fixos e são compostos pelos mesmos tipos de agrupamentos, mas não exatamente pela mesma estrutura. Estes constituem em fatores importantes na escolha de modelos para estimar a característica estática de forma recursiva, conforme detalhado no próximo capítulo.

## 4.7 Conclusões

Modelos lineares são insatisfatórios para representar o sistema em questão e incapazes de recuperar as não linearidades estáticas presentes nos dados. Desta forma, com base nos resultados, pode-se dizer que um modelo será satisfatório no que diz respeito à estimação da característica estática não linear, se o mesmo for também satisfatório do ponto de vista dinâmico.

No entanto, um modelo eficaz dinamicamente nem sempre o é em relação à estimação das não linearidades estáticas, a não ser que os regressores escolhidos sejam adequados. Este fato foi observado através dos modelos aqui considerados, uma vez que os modelos não lineares são aproximadamente equivalentes para representar a dinâmica do sistema, mas o mesmo não ocorre

em relação à estimação das características estáticas.

Desta forma, pode-se dizer que uma estimação eficiente das não linearidades estáticas contidas nos dados, requer um modelo válido dinamicamente, já que depende das estimativas de seus coeficientes, ou seja, se um modelo é dinamicamente inviável, dificilmente, será possível retirar informação do mesmo a respeito de suas características estáticas não lineares.

Através dos resultados aqui obtidos, foi possível observar que um maior número de pontos fixos dificulta na recuperação da característica estática não linear. Sendo assim, reduziu-se o número de pontos fixos, eliminando os agrupamentos  $\Omega_{y^i}$   $i > 1$  e manteve-se constante o grau de não linearidade dos modelos.

É importante salientar que, em alguns casos, é necessário um maior número de pontos fixos para se alcançar o objetivo do modelo, conforme ocorrido em (Hernández and Arkun, 1993), onde era necessário que o modelo tivesse três pontos fixos para reproduzir a curva estática do sistema.

Neste capítulo, as características estáticas dos modelos identificados previamente foram analisadas e comparadas. O resultado foi a escolha de uma estrutura única para fazer a estimação recursiva de parâmetros e de características estáticas, conforme detalhado no próximo capítulo.

## Capítulo 5

# Identificação Recursiva

### 5.1 Introdução

Neste capítulo será utilizada uma massa de dados que incorpora variação no sistema em um certo instante de tempo para identificar modelos capazes de verificar tal variação. Para isto utilizaram-se algumas das estruturas de modelos obtidas no capítulo 3 e estimaram-se recursivamente os parâmetros do modelo.

Além disso, como a característica estática do sistema em estudo, varia com as condições do ambiente e com a presença de distúrbio, uma estimação recursiva de tal característica é desejável. Por esta razão, para a massa de dados mencionada acima, utilizou-se a identificação recursiva não só para a estimação dos parâmetros do modelo, mas também para uma atualização das curvas estáticas não lineares obtidas.

O capítulo está organizado como se segue: na seção 5.2 são apresentadas as especificações do teste dinâmico englobando variações no sistema, bem como as respectivas massas de dados. Na seção 5.3 é realizada a estimação recursiva de parâmetros e da característica estática não linear, com base em algumas estruturas definidas no capítulo 3. Finalmente algumas considerações são feitas com base nos modelos e características obtidos.

### 5.2 Teste Dinâmico Variável

Foi realizado um teste dinâmico para coleta dos dados que foram utilizados na identificação recursiva. Foram coletados 1190 pontos, sendo que, durante a coleta das 544 primeiras observações, o ventilador estava desligado. As demais observações foram feitas com o ventilador ligado.

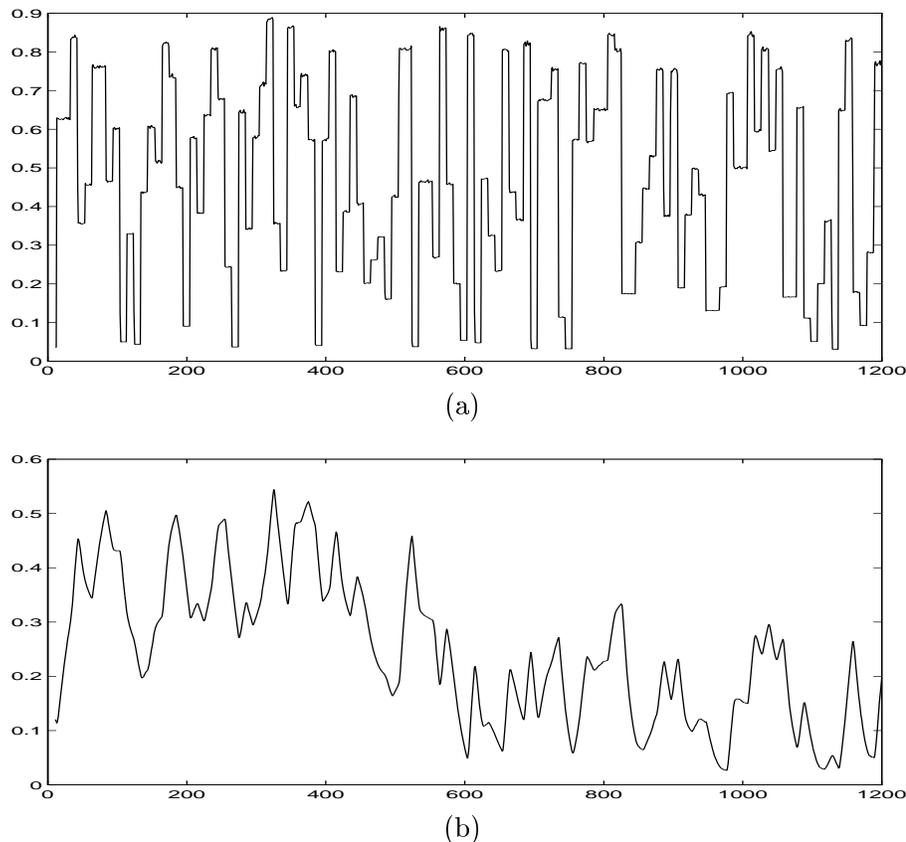


Figura 5.1: Massa de dados incluindo variação no sistema  $\langle din4 \rangle$ . (a) Dados de entrada (em pu) (b) Dados de saída (em pu).

A massa de dados completa está indicada na Figura 5.1. A entrada utilizada foi um sinal de características aleatórias, observar a função de autocorrelação mostrada na Figura 5.2, com patamares de 1 min e período de amostragem igual a 6s. O teste teve duração de 2 horas e foi realizado a uma temperatura ambiente de 24°C.

As autocorrelações lineares e não lineares da saída estão indicadas nas Figuras (5.3) e (5.4), respectivamente. Como o sinal de saída possui valor médio diferente para as janelas de 1–544 e de 545–1190 amostras, as funções de correlação devem ser traçadas separadamente. Semelhante ao realizado anteriormente, observando os gráficos, os dados foram dizimados de dois antes da estimação recursiva dos parâmetros do modelo.

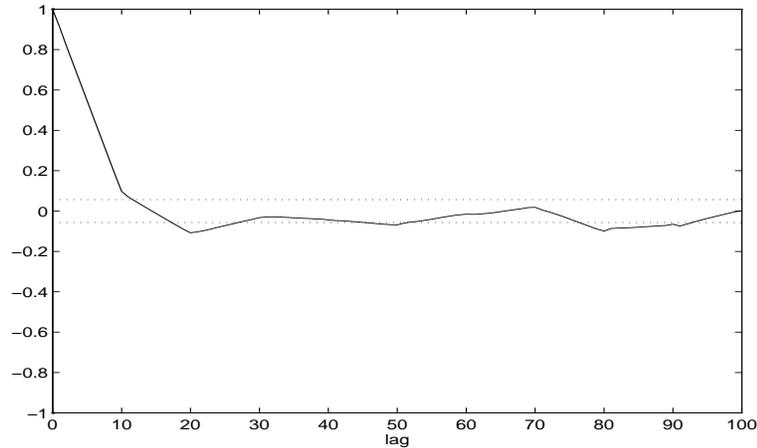


Figura 5.2: Autocorrelação do sinal de entrada da massa de dados  $\langle din4 \rangle$ .

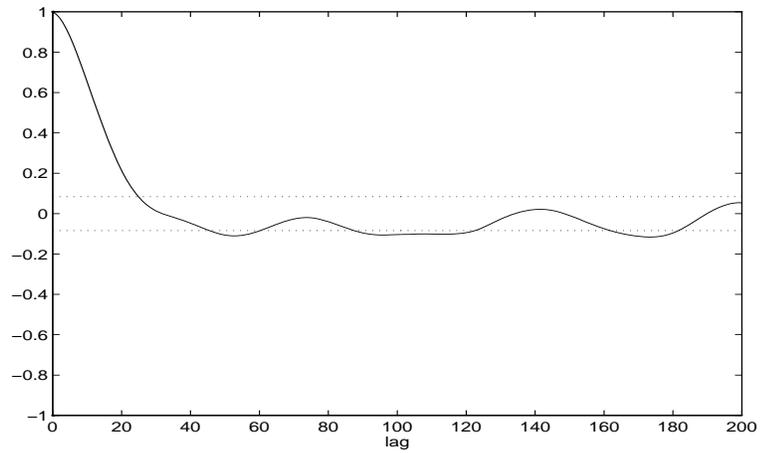
### 5.3 Identificação Recursiva

Os modelos descritos no capítulo 3 foram obtidos através da identificação em batelada iterativa. No entanto, como a massa de dados  $\langle din4 \rangle$  incorpora variação no sistema em um certo instante de tempo, deve-se utilizar um método de estimação de parâmetros que seja capaz de acompanhar tais variações, de forma a proporcionar uma estimativa mais fiel aos dados experimentais. Desta forma, será utilizado o algoritmo recursivo, equações (2.83) e (2.87), com fator de esquecimento igual a 0,99.

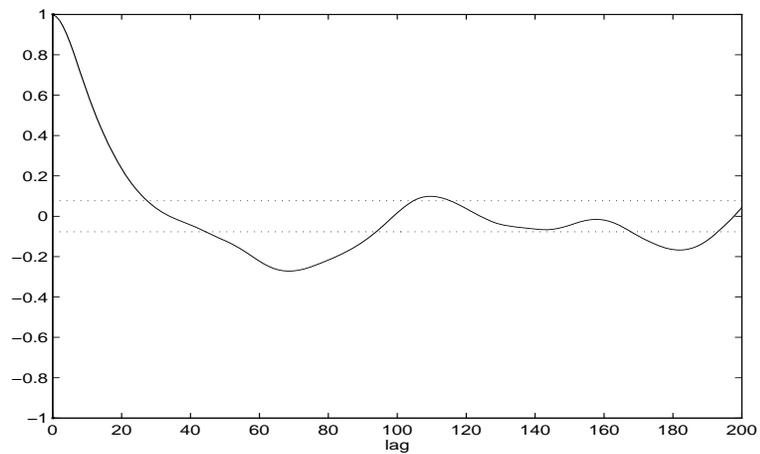
Como somente os parâmetros do modelo serão estimados recursivamente, serão utilizadas algumas das estruturas obtidas através da identificação em batelada. Com base nos resultados obtidos nos capítulos 3 e 4, somente dois casos serão considerados: modelos com termos quadráticos com um ponto fixo e modelos com termos cúbicos com dois pontos fixos.

#### 5.3.1 Modelos com Um Ponto Fixo

Na tentativa de verificar se modelos quadráticos com um ponto fixo são satisfatórios quanto à estimação da característica estática de forma recursiva, considerou-se a estrutura do modelo (3.10) com sete termos e atrasos máximos igual a dois, obtido após terem sido excluídos os agrupamentos  $\Omega_0$ ,  $\Omega_u$  e  $\Omega_{y^2}$  dos candidatos. Assim procedendo, a seguinte estrutura será utilizada para obter modelos de forma recursiva:

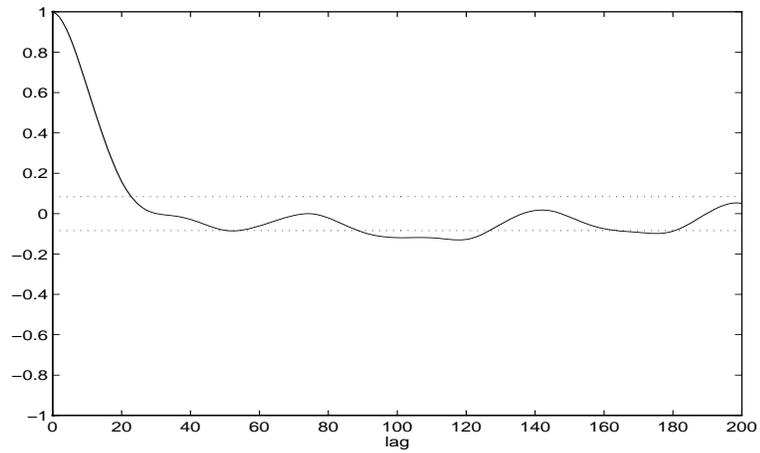


(a)

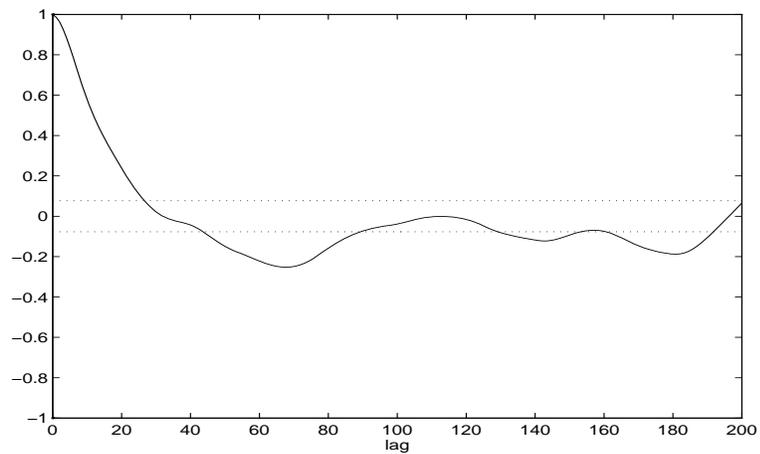


(b)

Figura 5.3: Autocorrelação linear (2.1) do sinal de saída da massa de dados  $\langle din4 \rangle$  utilizada para determinar a taxa de amostragem. (a) Janela de 1 – 544 amostras (b) Janela de 545 – 1190 amostras



(a)



(b)

Figura 5.4: Autocorrelação não linear (2.2) do sinal de saída da massa de dados  $\langle din4 \rangle$  utilizada para determinar a taxa de amostragem. (a) Janela de 1 – 544 amostras (b) Janela de 545 – 1190 amostras

Tabela 5.1: Inicialização dos parâmetros do modelo (5.1).

$\hat{\theta}_{1,0}(1)$	$\hat{\theta}_{0,2}(1, 1)$	$\hat{\theta}_{1,0}(2)$	$\hat{\theta}_{1,1}(1, 2)$	$\hat{\theta}_{1,1}(2, 2)$	$\hat{\theta}_{0,2}(2, 2)$	$\hat{\theta}_{0,2}(2, 1)$
1,3920	0,0454	-0,4235	-0,4388	0,3756	0,0218	0,0097

$$\begin{aligned}
y(k) = & \hat{\theta}_{1,0}(1)y(k-1) + \hat{\theta}_{0,2}(1, 1)u(k-1)^2 + \hat{\theta}_{1,0}(2)y(k-2) \\
& + \hat{\theta}_{1,1}(1, 2)u(k-2)y(k-1) + \hat{\theta}_{1,1}(2, 2)u(k-2)y(k-2) \\
& + \hat{\theta}_{0,2}(2, 2)u(k-2)^2 + \hat{\theta}_{0,2}(2, 1)u(k-2)u(k-1) \\
& + \sum_{i=1}^{n_e} \hat{\theta}_i \xi(k-i) + \xi(k) .
\end{aligned} \tag{5.1}$$

Os parâmetros  $\hat{\theta}_{p, m_i - p}(n_y, n_u)^1$  da equação (5.1) serão estimados através do algoritmo recursivo com fator de esquecimento  $\beta = 0,99$ .

Inicializou-se o vetor de parâmetros com os valores calculados através da identificação em batelada, utilizando a massa de dados  $\langle \text{din2} \rangle$ , que estão indicados na Tabela 5.1. Ao passo que, a matriz de covariância inicial foi igual a  $\mathbf{P}_0 = 1000\mathbf{P}_\mathbf{I}$ , onde  $\mathbf{P}_\mathbf{I}$  é uma matriz quadrada unitária de dimensão  $(n_p + n_n) \times (n_p + n_n)$ .

É importante salientar que na identificação recursiva, descrita neste trabalho, trabalha-se com uma estrutura de modelo determinada a partir da identificação em batelada, pois desta forma somente os parâmetros do mesmo devem ser corrigidos com o passar do tempo. No entanto, a inicialização dos parâmetros não precisa necessariamente ser proveniente de resultados anteriores. Este procedimento foi aqui adotado com o intuito de facilitar a convergência do algoritmo, determinando assim um modelo mais confiável aos dados experimentais.

As variações dos coeficientes de agrupamentos ao longo do tempo estão indicadas na Figura 5.5 e as variações dos parâmetros na Figura 5.6. Observando a Figura 5.5-a nota-se que o coeficiente de agrupamento  $\Omega_y$  não apresenta alterações relevantes, indicando que o mesmo não é bruscamente alterado mesmo com o sistema sofrendo a variação mencionada.

De forma similar ao realizado no capítulo 3, utilizaram-se os conceitos de pontos fixos e agrupamentos de termos para estimar as características

<sup>1</sup>Parâmetro estimado, onde os subíndices  $m_i$  e  $p$  indicam o tipo de termo e os argumentos os respectivos atrasos.

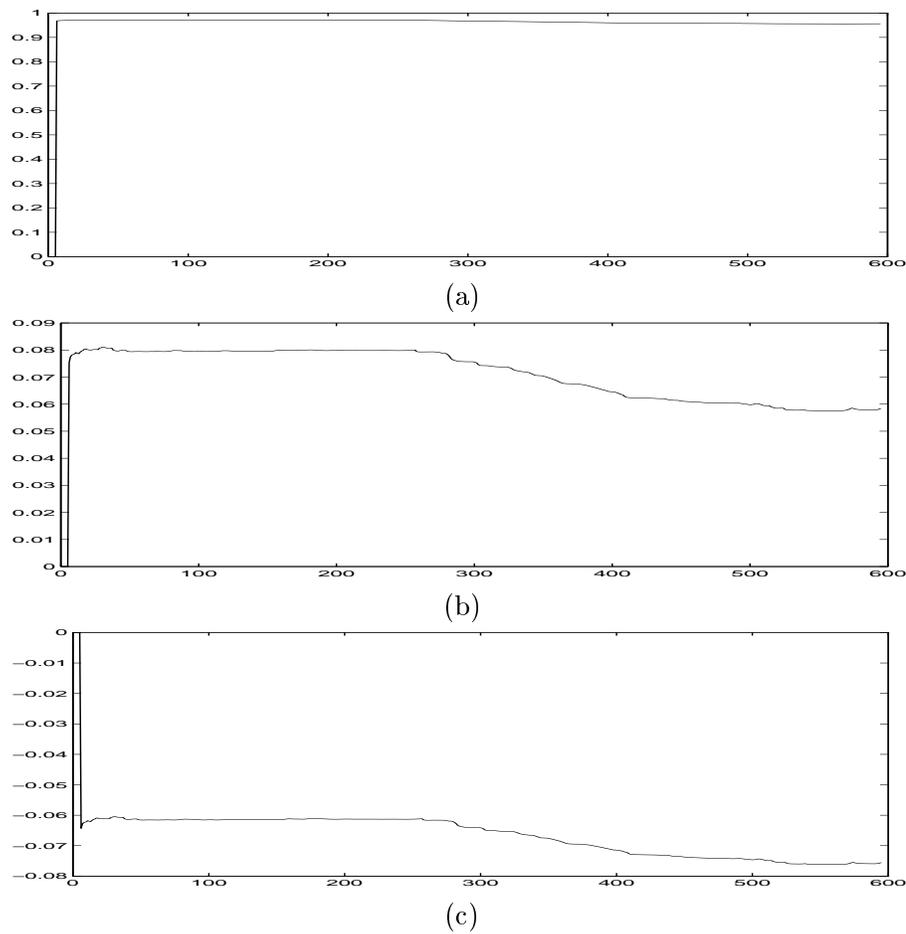


Figura 5.5: Variações dos coeficientes de agrupamentos da estrutura (5.1) ao longo do tempo. (a)  $\Sigma_y$  (b)  $\Sigma_{u^2}$  (c)  $\Sigma_{uy}$ .

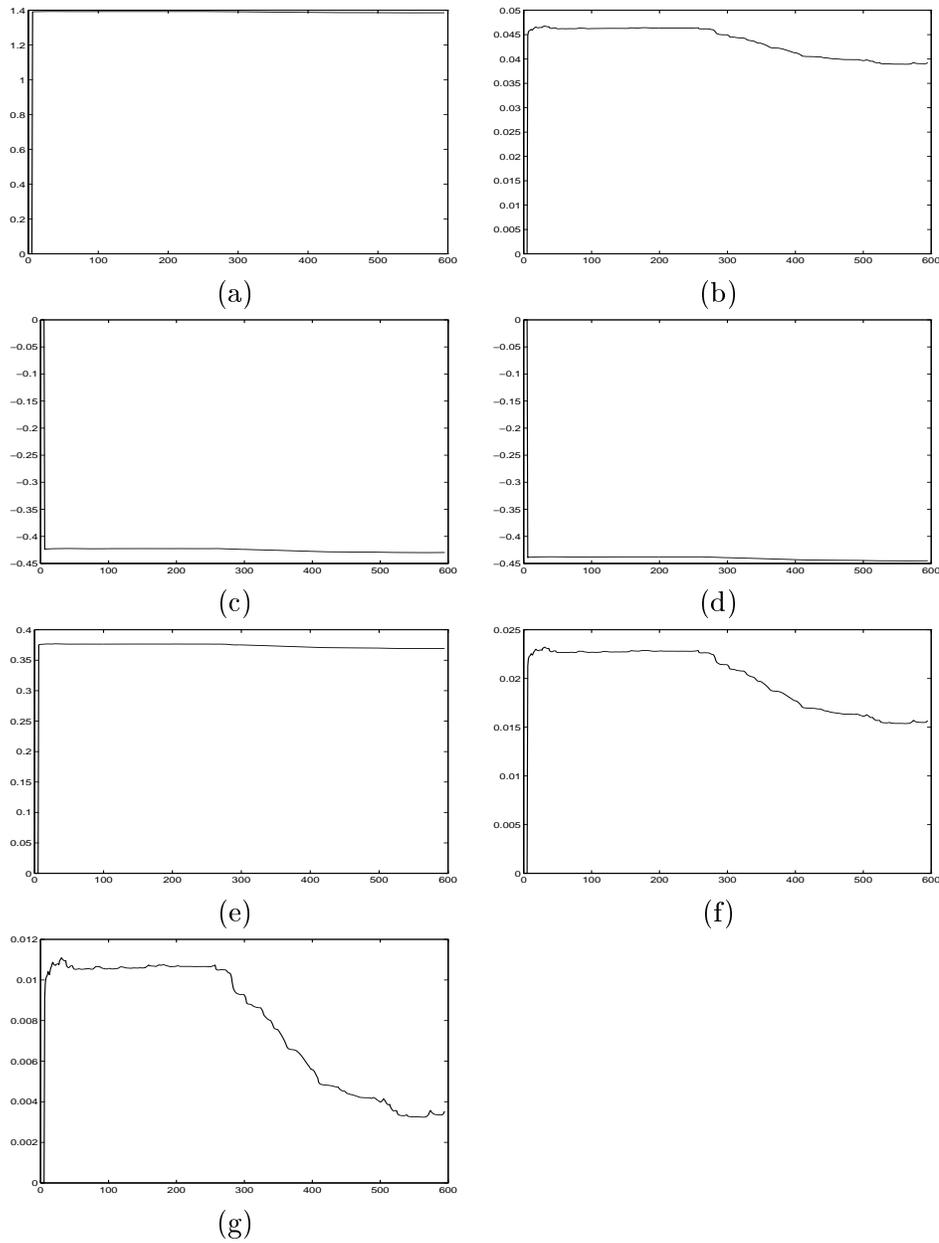


Figura 5.6: Variações dos parâmetros da estrutura (5.1) ao longo do tempo. (a)  $\hat{\theta}_{1,0}(1)$  (b)  $\hat{\theta}_{0,2}(1, 1)$  (c)  $\hat{\theta}_{1,0}(2)$  (d)  $\hat{\theta}_{1,1}(1, 2)$  (e)  $\hat{\theta}_{1,1}(2, 2)$  (f)  $\hat{\theta}_{0,2}(2, 2)$  (g)  $\hat{\theta}_{0,2}(2, 1)$ .

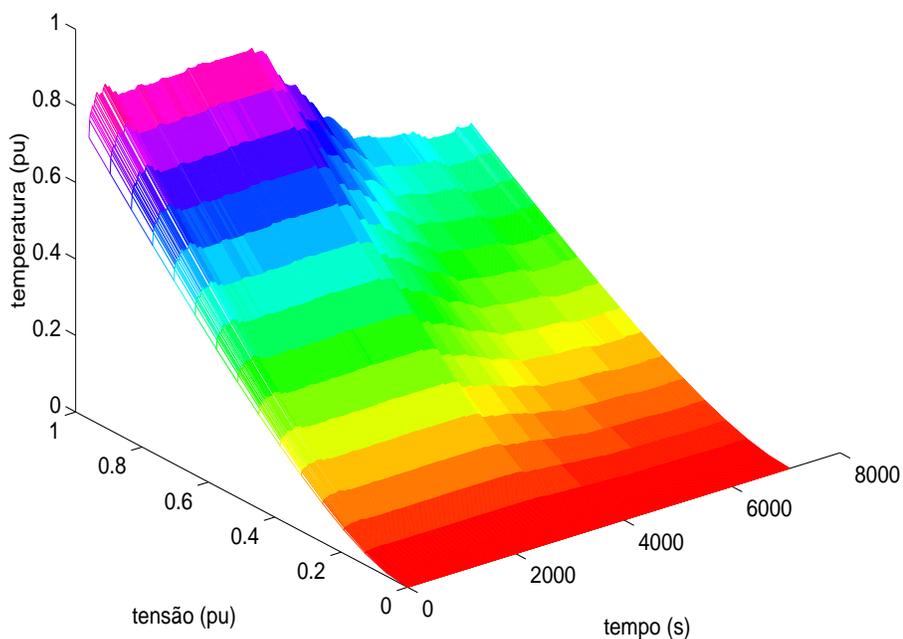


Figura 5.7: Vista A - Característica estática obtida a partir da estrutura de modelo quadrática (5.1) utilizando algoritmo recursivo.

estáticas a partir dos modelos. No entanto, como foi realizado recursivamente obteve-se uma característica para cada iteração do algoritmo de estimação de parâmetros. As vistas A e B da recursão da característica estática não linear recuperada através da estrutura de modelo (5.1), estão indicadas nas Figuras 5.7 e 5.8, respectivamente.

A Figura 5.9 compara as curvas obtidas através da estrutura (5.1) nos instantes de tempo  $t_1 = 2400s$  e  $t_2 = 6600s$  (corte realizado nas Figuras 5.7 e 5.8) com as Figuras 3.3 e 3.5. Percebe-se que as características estáticas obtidas recursivamente são similares às obtidas nos testes estáticos. Desta forma, pode-se dizer que a estrutura (5.1) resulta em modelos válidos em relação à curva estática do sistema.

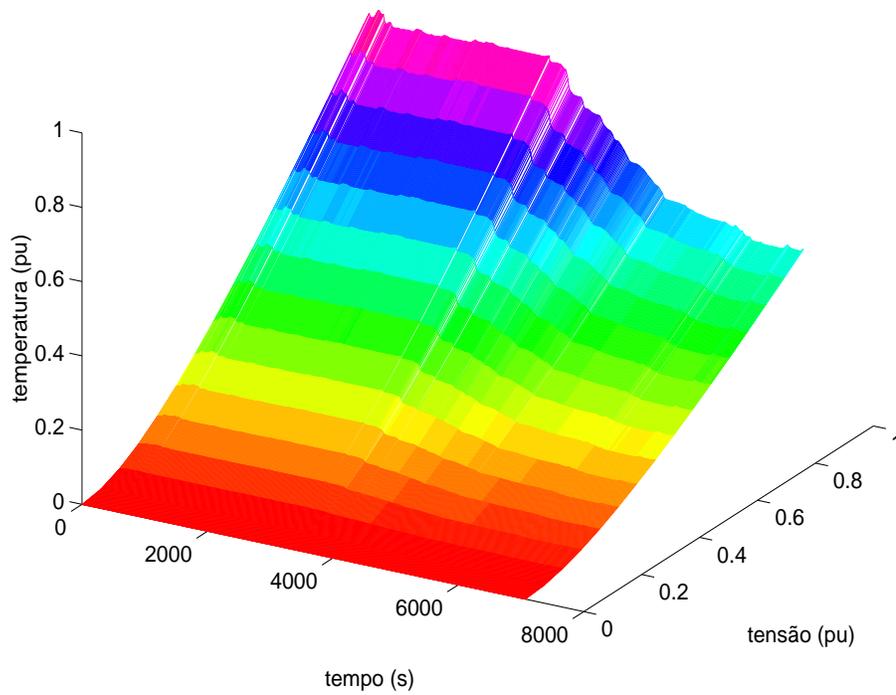


Figura 5.8: Vista B - Característica estática obtida a partir da estrutura de modelo quadrática (5.1) utilizando algoritmo recursivo.

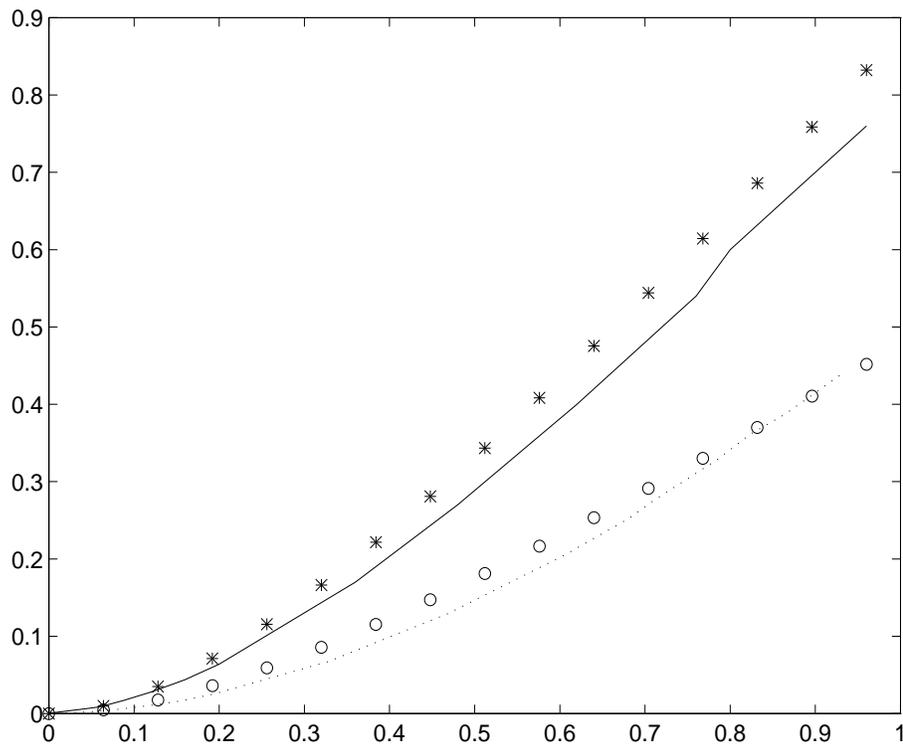


Figura 5.9: Comparação entre a característica estática obtida a partir da estrutura (5.1) e a partir dos testes estáticos. (—) Figura 3.3 (···) Figura 3.5 (\*) Característica estática no instante  $t_1$  (o) Característica estática no instante  $t_2$ .

Tabela 5.2: Inicialização dos parâmetros do modelo (5.2).

$\hat{\theta}_{1,0}(1)$	$\hat{\theta}_{0,2}(1, 1)$	$\hat{\theta}_{1,0}(2)$	$\hat{\theta}_{2,0}(1, 1)$	$\hat{\theta}_{0,2}(2, 1)$	$\hat{\theta}_{0,3}(2, 1, 1)$
1,3040	0,0389	-0,3303	-0,1913	0,0179	-0,0034
$\hat{\theta}_{1,1}(1, 2)$	$\hat{\theta}_{1,1}(2, 2)$	$\hat{\theta}_{0,3}(2, 2, 2)$	$\hat{\theta}_{1,1}(1, 1)$	$\hat{\theta}_{0,3}(2, 2, 1)$	$\hat{\theta}_{1,1}(2, 1)$
-0,3396	0,3125	0,0244	0,1513	-0,0167	-0,1305
$\hat{\theta}_{0,3}(1, 1, 1)$	$\hat{\theta}_{2,0}(2, 2)$	$\hat{\theta}_{2,1}(1, 1, 1)$			
0,0027	0,1262	0,0032			

### 5.3.2 Modelos com Dois Pontos Fixos

Utilizando a massa de dados  $\langle \text{din2} \rangle^2$  encontrou-se através da identificação em batelada a estrutura do modelo (3.14). Tal estrutura será utilizada para a identificação recursiva, devendo ser estimados os parâmetros  $\hat{\theta}_{p, m_i - p}(n_y, n_u)$  do seguinte modelo:

$$\begin{aligned}
y(k) = & \hat{\theta}_{1,0}(1) y(k-1) + \hat{\theta}_{0,2}(1, 1) u(k-1)^2 + \hat{\theta}_{1,0}(2) y(k-2) \\
& + \hat{\theta}_{2,0}(1, 1) y(k-1)^2 + \hat{\theta}_{0,2}(2, 1) u(k-2) u(k-1) \\
& + \hat{\theta}_{0,3}(2, 1, 1) u(k-2) u(k-1)^2 + \hat{\theta}_{1,1}(1, 2) u(k-2) y(k-1) \\
& + \hat{\theta}_{1,1}(2, 2) u(k-2) y(k-2) + \hat{\theta}_{0,3}(2, 2, 2) u(k-2)^3 \\
& + \hat{\theta}_{1,1}(1, 1) u(k-1) y(k-1) + \hat{\theta}_{0,3}(2, 2, 1) u(k-2)^2 u(k-1) \\
& + \hat{\theta}_{1,1}(2, 1) u(k-1) y(k-2) + \hat{\theta}_{0,3}(1, 1, 1) u(k-1)^3 \\
& + \hat{\theta}_{2,0}(2, 2) y(k-2)^2 + \hat{\theta}_{2,1}(1, 1, 1) u(k-1) y(k-1)^2 \\
& + \sum_{i=1}^{n_e} \hat{\theta}_i \xi(k-i) + \xi(k) .
\end{aligned} \tag{5.2}$$

Os parâmetros iniciais correspondem aos coeficientes do modelo (3.14) e estão indicados na Tabela 5.2.

As variações dos coeficientes de agrupamentos ao longo do tempo estão indicadas na Figura 5.10 e as variações dos parâmetros na Figura 5.11.

A característica estática estimada através do algoritmo recursivo, utilizando a estrutura definida em (5.2) para o ponto fixo estável, está indicada

<sup>2</sup>A estrutura identificada em batelada para a massa de dados  $\langle \text{din3} \rangle$  apresentou resultados piores em relação à estimação da curva estática de forma recursiva, não sendo, portanto, aqui considerada.

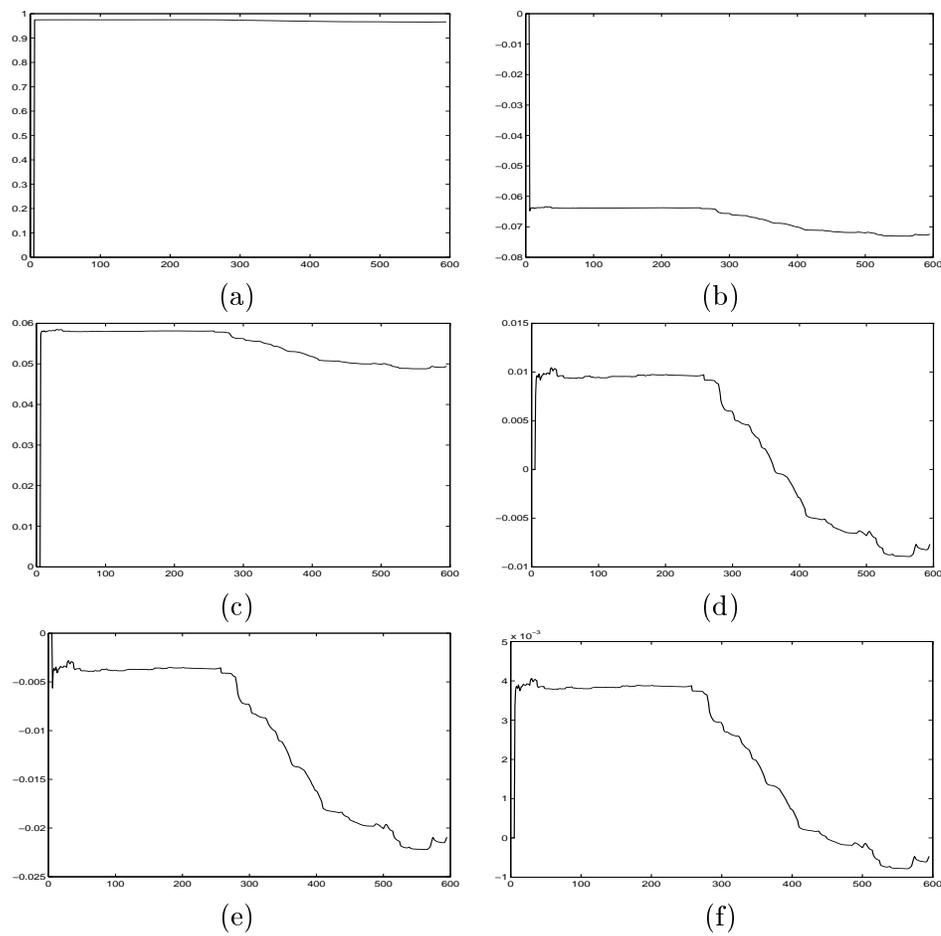


Figura 5.10: Variações dos coeficientes de agrupamentos da estrutura (5.2) ao longo do tempo. (a)  $\Sigma_y$  (b)  $\Sigma_{y^2}$  (c)  $\Sigma_{u^2}$  (d)  $\Sigma_{u^3}$  (e)  $\Sigma_{uy}$  (f)  $\Sigma_{uy^2}$ .

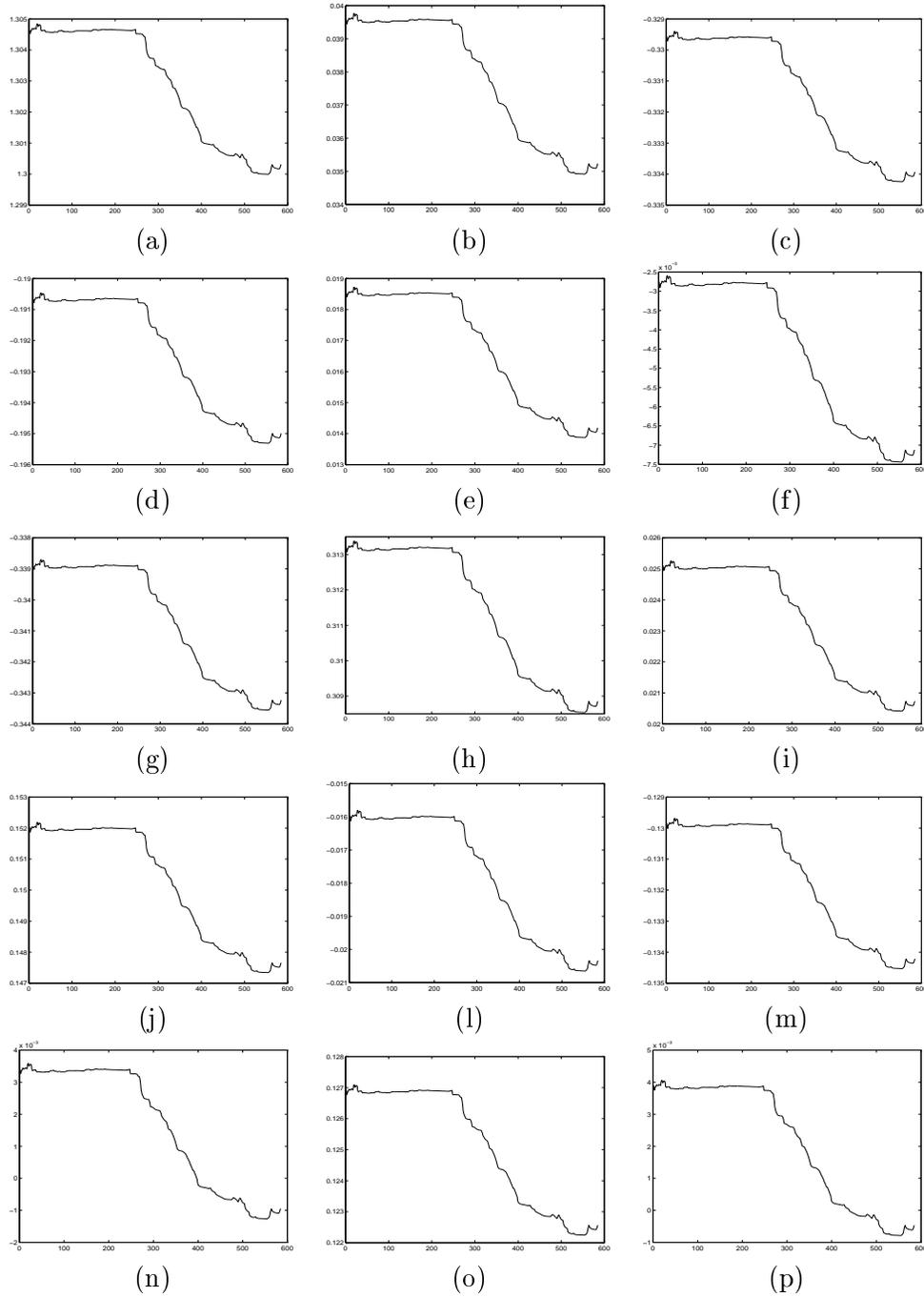


Figura 5.11: Variações dos parâmetros do modelo recursivo (5.2) ao longo do tempo. (a)  $\hat{\theta}_{1,0}(1)$  (b)  $\hat{\theta}_{0,2}(1,1)$  (c)  $\hat{\theta}_{1,0}(2)$  (d)  $\hat{\theta}_{2,0}(1,1)$  (e)  $\hat{\theta}_{0,2}(2,1)$  (f)  $\hat{\theta}_{0,3}(2,1,1)$  (g)  $\hat{\theta}_{1,1}(1,2)$  (h)  $\hat{\theta}_{1,1}(2,2)$  (i)  $\hat{\theta}_{0,3}(2,2,2)$  (j)  $\hat{\theta}_{1,1}(1,1)$  (l)  $\hat{\theta}_{0,3}(2,2,1)$  (m)  $\hat{\theta}_{1,1}(2,1)$  (n)  $\hat{\theta}_{0,3}(1,1,1)$  (o)  $\hat{\theta}_{2,0}(2,2)$  (p)  $\hat{\theta}_{2,1}(1,1,1)$ .

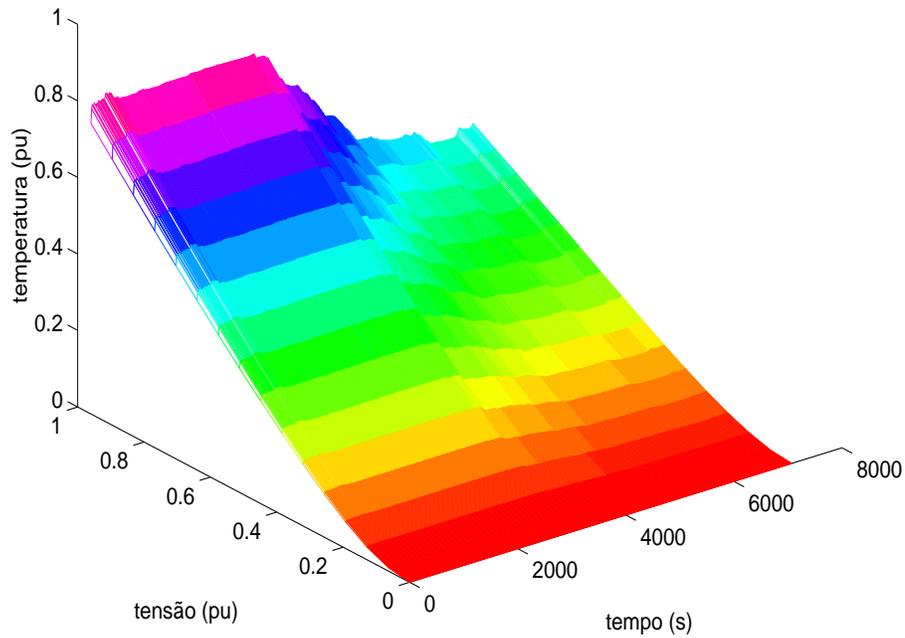


Figura 5.12: Vista A - Característica estática obtida para o ponto fixo estável, utilizando algoritmo recursivo a partir da estrutura de modelo cúbica (5.2).

nas Figuras 5.12 e 5.13.

Assim como o modelo quadrático, o cúbico apresentou-se satisfatório em relação à estimação da característica estática recursivamente (observar Figura 5.14).

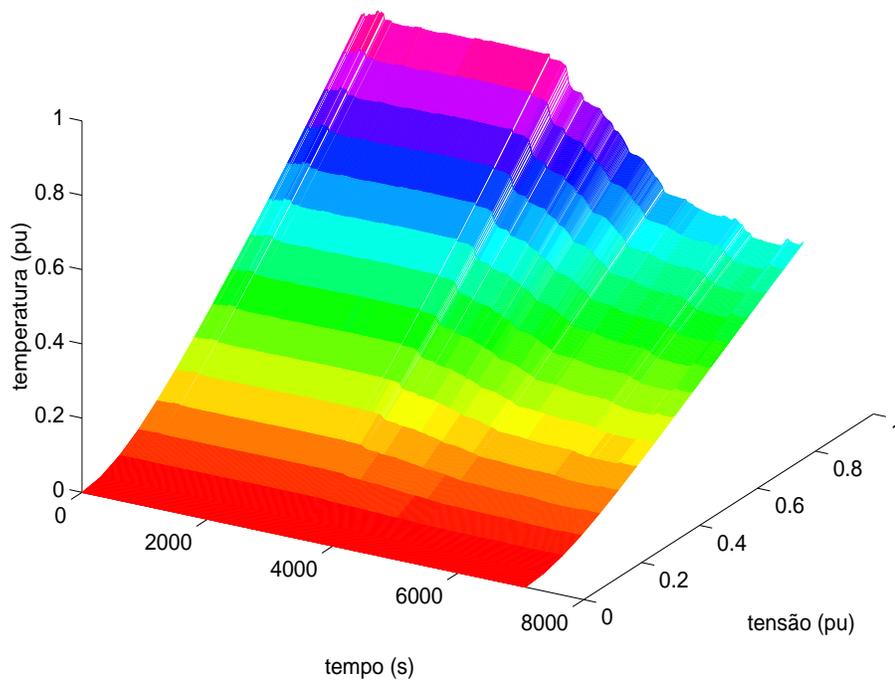


Figura 5.13: Vista B - Característica estática obtida para o ponto fixo estável, utilizando algoritmo recursivo a partir da estrutura de modelo cúbica (5.2).

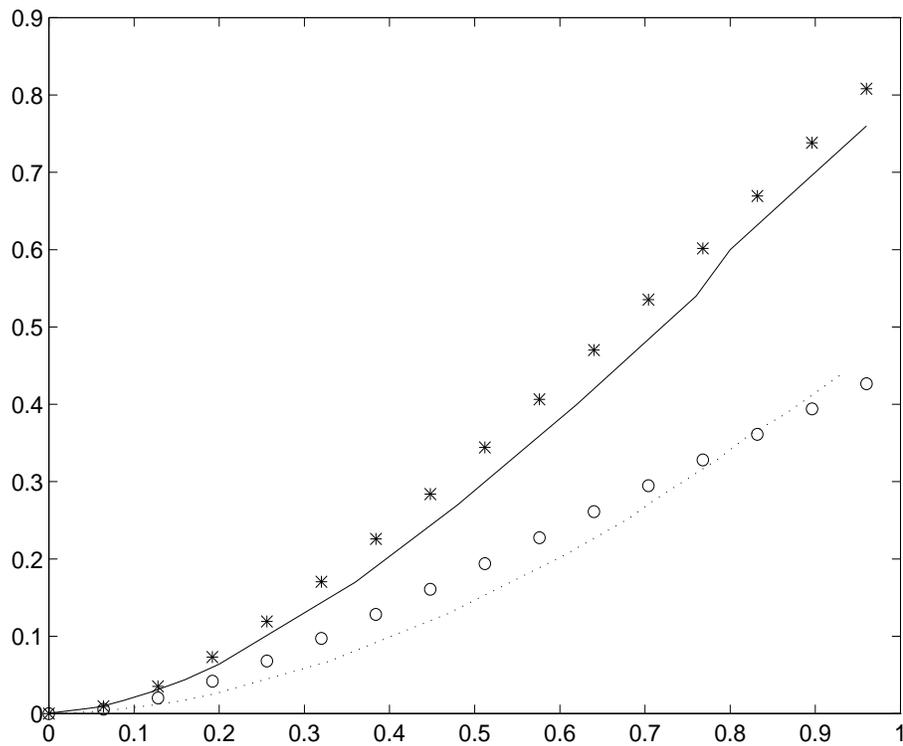


Figura 5.14: Comparação entre a característica estática obtida a partir da estrutura (5.2) e a partir dos testes estáticos. (—) Figura 3.3 (···) Figura 3.5 (\*) Característica estática no instante  $t_1$  (o) Característica estática no instante  $t_2$ .

## 5.4 Conclusões

Comparando-se as características estáticas não lineares obtidas de forma recursiva através das estruturas com um (Figuras 5.7 e 5.8) e dois pontos fixos (Figuras 5.12 e 5.13), respectivamente, observa-se que ambas estruturas são eficazes no que diz respeito à estimação da característica estática não linear. Além disso, acredita-se que os modelos obtidos recursivamente também são dinamicamente válidos, já que foram utilizadas as mesmas estruturas de modelos validados anteriormente.

No entanto, a fim de evitar sobreparametrização, sempre que possível, deve-se utilizar modelos mais simples. Sendo assim, é desejável a obtenção de modelos com menor número de termos, menor grau de não linearidade e menor número de pontos fixos. Desta forma, os modelos com termos quadráticos com um ponto fixo são preferíveis aos modelos com termos cúbicos com dois pontos fixos. Além disso, deve-se ter em mente que a estrutura com um ponto fixo torna a estimação da característica estática mais fácil, já que existirá somente uma curva possível.

# Conclusão

A determinação de modelos é uma tarefa de grande importância para análise e controle de sistemas. No entanto, como sistemas reais são, em geral, variantes no tempo, é desejável a utilização de técnicas capazes de acompanhar tais variações. Desta forma, um procedimento de identificação associado a técnicas recursivas de estimação de parâmetros torna-se bastante adequado.

Este trabalho identificou modelos capazes de representar tanto a dinâmica, quanto a curva estática de sistemas não lineares. A partir dos modelos identificados é possível recuperar a característica estática não linear do sistema.

Foram identificados modelos polinomiais NARMAX através do procedimento de estimação em batelada. Após terem sido determinados modelos capazes de representar a dinâmica do sistema, eles foram analisados em regime permanente, o que resultou em equações algébricas da característica estática.

Dentre os modelos considerados dinamicamente e estaticamente válidos, identificados através do procedimento de estimação em batelada, selecionaram-se algumas estruturas que se apresentaram mais viáveis do ponto de vista da identificação recursiva. Utilizando então, uma outra massa de dados, na qual foram registradas variações do sistema no decorrer do tempo, estimaram-se recursivamente os parâmetros dos modelos selecionados, bem como as características estáticas não lineares a partir da análise em regime estacionário.

Com relação à estimação das características estáticas, o uso de modelos conforme discutido neste trabalho, apresentou certas vantagens em relação aos testes estáticos convencionais. Em primeiro lugar, este procedimento permite a utilização de uma mesma massa de dados tanto para determinar uma estrutura de modelo, quanto para determinar a forma da característica estática. Sendo assim, numa situação real, a característica estática poderia ser estimada a partir de dados dinâmicos resultantes de um teste relativamente simples, para o qual não seria necessário interromper a operação do processo. Em segundo lugar, pode-se identificar uma estrutura de modelo *off-*

*line* que será utilizada para predição e estimação *on-line* das características estáticas. Acredita-se que tais informações podem ser utilizadas para auxiliar em problemas de controle e monitorização de sistemas variantes no tempo.

Algumas propostas para continuação deste trabalho são:

- Estimação *on-line* da característica estática. Isto requeriria a implementação do algoritmo recursivo descrito neste trabalho por um programa executável em tempo real.
- Investigar como as informações obtidas sobre as características estáticas não lineares poderiam ser usadas para auxiliar em problemas de controle.
- Utilização de fator de esquecimento variável para verificar se é possível acompanhar variações do sistema de forma mais rápida.
- A aplicação das técnicas investigadas neste trabalho a outros sistemas, como por exemplo, máquinas elétricas.

# Bibliografia

- Aguirre, L. A. (1994a). Some remarks on structure selection for nonlinear models. *Int. J. Bifurcation and Chaos*, 4(6):1707–1714.
- Aguirre, L. A. (1994b). Term clustering and the order selection of linear continuous systems. *Journal of The Franklin Institute*, 331B(4):403–415.
- Aguirre, L. A. (1995). A nonlinear correlation function for selecting the delay time in dynamical reconstruction. *Physics Letters A*, 203:88–94.
- Aguirre, L. A. and Billings, S. A. (1994). Validating identified nonlinear models with chaotic dynamics. *Int. J. Bifurcation and Chaos*, 4(1):109–125.
- Aguirre, L. A. and Billings, S. A. (1995a). Improved structure selection for nonlinear models based on term clustering. *International J. Control*, 62(3):569–587.
- Aguirre, L. A. and Billings, S. A. (1995b). Retrieving dynamical invariants from chaotic data using NARMAX models. *Int. J. Bifurcation and Chaos*, 5(2):449–474.
- Aguirre, L. A., Cassini, C. C. S., Rodrigues, G. G., and Jácome, C. R. F. (1998a). Recovering static nonlinearities using NARX polynomial models. In *Proceedings of XII Brazilian Automatic Control Conference - XII CBA*, volume IV, pages 1431–1435, Uberlândia, M.G., Brazil.
- Aguirre, L. A., Correa, M. V., and Mendes, E. M. A. M. (1999). Using rational narmax models to recover static nonlinearities. Accepted to the invited session on Laguerre functions, European Control Conference, Germany.
- Aguirre, L. A. and Mendes, E. M. (1996). Global nonlinear polynomial models: Structure, term clusters and fixed points. *Int. J. Bifurcation and Chaos*, 6(2):279–294.

- Aguirre, L. A., Rodrigues, G. G., and Jácome, C. (1998b). “Identificação de sistemas não lineares utilizando modelos NARMAX polinomiais — uma revisão e novos resultados”. *Controle & Automação*, 9(2):90–106.
- Akaike, H. (1974). A new look at the statistical model identification. *IEEE Trans. Automat. Contr.*, 19(6):716–723.
- Alligood, K. T., Sauer, T. D., and Yorke, J. A. (1997). *CHAOS An Introduction to Dynamical Systems*. Springer-Verlag New York, Inc.
- Åström, K. J. and Wittenmark, B. (1990). *Computer Controlled Systems - Theory and Design*. Prentice-Hall. Second Edition.
- Åström, K. J. and Wittenmark, B. (1995). *Adaptive Control*. Addison-Wesley Publishing Company. Second Edition.
- Billings, S. A. (1980). Identification of nonlinear systems — a survey. *IEE Proceedings Pt. D*, 127(6):272–285.
- Billings, S. A. and Aguirre, L. A. (1995). Effects of the sampling time on the dynamics and identification of nonlinear models. *Int. J. Bifurcation and Chaos*, 5(6):1541–1556.
- Billings, S. A. and Chen, S. (1992). Neural networks and system identification. In Warwick, K., Irwing, G. W., and Hunt, K. J., editors, *Neural Networks for Systems and Control*, chapter 9, pages 181–205. Peter Peregrinus, London.
- Billings, S. A., Chen, S., and Korenberg, M. J. (1989). Identification of MIMO nonlinear systems using a forward-regression orthogonal estimator. *International J. Control*, 49(6):2157–2189.
- Billings, S. A. and Fadzil, M. B. (1985). The practical identification of systems with nonlinearities. In *Proc. 7th IFAC Symposium on Identification and System Parameter Estimation*, pages 155–160, York, U. K.
- Billings, S. A. and Fakhouri, S. Y. (1982). Identification of systems containing linear dynamic and static nonlinear elements. *Automatica*, 18(1):15–26.
- Billings, S. A. and Tao, Q. H. (1991). Model validity tests for nonlinear signal processing applications. *International J. Control*, 54(1):157–199.
- Billings, S. A. and Voon, W. S. F. (1983). Structure detection and model validity tests in the identification of nonlinear systems. *IEE Proceedings Pt. D*, 130(4):193–199.

- Billings, S. A. and Zhu, Q. M. (1994a). Model validation tests for multivariable nonlinear models including neural networks. Technical Report 484, Department of Automatic Control and Systems Engineering, University of Sheffield.
- Billings, S. A. and Zhu, Q. M. (1994b). Nonlinear model validation using correlation tests. *International J. Control*, 60(6):1107–1120.
- Braga, A. P., Carvalho, A. P. L. F., and Ludermir, T. B. (1998). *Fundamentos de Redes Neurais Artificiais*. XI Escola Brasileira de Computação.
- Cassini, C. C. S., Braga, A. P., and Aguirre, L. A. (1998). Um estudo comparativo entre modelos polinomiais NARMAX e redes neurais artificiais aplicados à identificação de sistemas. In *V Simpósio Brasileiro de Redes Neurais, Belo Horizonte, M. G.*, pages 237–240.
- Çinar, A. (1995). Nonlinear time series models for multivariable dynamic processes. *Chemometrics and Intelligent Laboratory Systems*, 30:147–158.
- Chen, C. (1970). *Linear System Theory and Design*. Holt, Rinehart and Winston.
- Chen, S. and Billings, S. A. (1989). Representations of nonlinear systems: the NARMAX model. *International J. Control*, 49(3):1013–1032.
- Chen, S., Billings, S. A., and Luo, W. (1989). Orthogonal least squares methods and their application to nonlinear system identification. *International J. Control*, 50(5):1873–1896.
- Cubillos, F. A., Alvarez, P. I., Pinto, J. C., and Lima, E. L. (1996). Hybrid-neural modeling for particulate solid drying processes. *Powder Technology*, 87:153–160.
- Cubillos, F. A. and Lima, E. L. (1997). Identification and optimizing control of a rougher flotation circuit using an adaptable hybrid-neural model. *Minerals Engineering*, 10(7):707–721.
- Cybenko, G. (1989). Approximation by superpositions of a sigmoidal function. *Mathematics of Control, Signal, and Systems*, 2:303–314.
- Gouesbet, G. and Letellier, C. (1994). Global vector field reconstruction by using a multivariate polynomial  $L - 2$  approximation on nets. *pre*, 49(6):4955–4972.

- Gruhn, P., Pittman, J., Wiley, S., and LeBlanc, T. (1998). Increase plant safety with online valve testing. In *Tech*, pages 39–43.
- Haber, R. (1985). Nonlinearity tests for dynamic processes. In *IFAC Symposium on Identification and System Parameter Estimation*, pages 409–413.
- Haber, R. and Keviczky, L. (1985). Identification of ‘linear’ systems having signal-dependent parameters. *International J. Systems Sci.*, 16(7):869–884.
- Haber, R. and Unbehauen, H. (1990). Structure identification of nonlinear dynamic systems — A survey on input/output approaches. *Automatica*, 26(4):651–677.
- Häck, M. and Köhne, M. (1996). Estimation of wastewater process parameters using neural networks. *Wat. Sci. Tech.*, 33(1):101–115.
- Haykin, S. (1994). *Neural Networks - A Comprehensive Foundation*. Prentice-Hall.
- Hernández, E. and Arkun, Y. (1993). Control of nonlinear systems using polynomial ARMA models. *AIChE Journal*, 39(3):446–460. Process Systems Engineering.
- Jácome, C. R. F. (1996). *Uso de Conhecimento Prévio na Identificação de Modelos Polinômiais NARMAX*. Dissertação de Mestrado, PPGEE, Universidade Federal de Minas Gerais, Belo Horizonte, Brazil.
- Kadtke, J. B., Brush, J., and Holzfuss, J. (1993). Global dynamical equations and lyapunov exponents from noisy chaotic time series. *International Journal of Bifurcation and Chaos*, 3(3):607–616.
- Kortmann, M., Janiszowski, K., and Unbehauen, H. (1988). Application and comparison of different identification schemes under industrial conditions. *International J. Control*, 48(6):2275–2296.
- Kreyszig, E. (1993). *Advanced Engineering Mathematics*. John Wiley. 7th Edition.
- Lang, K. J. and Hinton, J. E. (1988). The development of the time-delay neural networks architecture for speech recognition. Technical Report CMU-CS-88-152, Carnegie-Mellon University.

- Leontaritis, I. J. and Billings, S. A. (1987). Model selection and validation methods for non-linear systems. *International J. Control*, 45(1):311–341.
- Lindskog, P. and Ljung, L. (1994). Tools for semi-physical modeling. Technical Report LiTH-ISY-R-1599, Department of Electrical Engineering Linköping University.
- Lipschutz, S. (1991). *Algebra Linear*. Schaum McGraw-Hill.
- Ljung, L. (1987). *System Identification - Theory for the User*. Prentice-Hall, New Jersey.
- Ljung, L. (1991). *System Identification for Use with Matlab - User's Guide*. The MathWorks, Inc.
- Ljung, L. (1994). System identification in a MIC perspective. *Modeling, Identification and Control*, 15(3):153–159.
- Matko, D., Zupančič, B., and Karba, R. (1992). *Simulation and Modelling of Continuous Systems - A Case Study Approach*. Prentice Hall International (UK).
- Mendes, E. M. A. M. and Billings, S. A. (1996). Important issues on system identification using algorithms for structure selection. In 12<sup>th</sup> *Conference on Robotics and Factories to the Future*, Middlesex University, England, U. K.
- Mendes, E. M. A. M. and Billings, S. A. (1998). On overparametrization of nonlinear discrete systems. *Int. J. Bifurcation and Chaos*, 8(3):535–556.
- Norton (1986). *An Introduction to Identification*. Academic Press.
- Ogata, K. (1993). *Engenharia de Controle Moderno*. Prentice-Hall do Brasil. 2edição.
- Pröll, T. and Karim, M. N. (1994). Model-predictive pH control using real-time NARX approach. *AIChE Journal*, 40(2):269–282. Process Systems Engineering.
- Reed, R. (1993). Pruning algorithms - a survey. *IEEE Transactions on Neural Networks*, 4(5):740–747.
- Ronco, E. and Gawthrop, P. J. (1997). Neural networks for modelling and control. Technical Report CSC97008, Centre for System and Control Department of Mechanical Engineering - University of Glasgow.

- Sjöberg, J., Zhang, Q., Ljung, L., Benveniste, A., Deylon, B., Glorennec, P., Hjalmarsson, H., and Juditsky, A. (1995). Nonlinear black-box modeling in system identification: a unified overview. Technical report.
- Söderström, T. and Stoica, P. (1989). *System Identification*. Prentice Hall International (UK) Ltd.
- Stoica, P., Janssen, P., and Söderström, T. (1986). Model-structure selection by cross-validation. *International J. Control*, 43(6):1841–1878.
- Thompson, J. M. T. and Stewart, H. B. (1986). *Nonlinear Dynamics and Chaos*. John Wiley & Sons Ltd.
- Wiggins, S. (1990). *Introduction to Applied Nonlinear Dynamical Systems and Chaos*. Springer-Verlag, New York.