

Laboratório de Modelagem, Análise e Controle de Sistemas Não-Lineares

Departamento de Engenharia Eletrônica

Universidade Federal de Minas Gerais

Av. Antônio Carlos 6627, 31270-901 Belo Horizonte, MG Brasil



Utilização de Auto-Consistência como Ferramenta Auxiliar na Seleção de Estrutura de Modelos NARX Polinomiais

Marcela Andrade Alves

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.

Orientadores: Dr. Luis Antonio Aguirre
Dr. Marcelo Vieira Corrêa

Belo Horizonte, 17 de dezembro de 2009

Dedicatória

*À minha família, que me proporcionou a
realização deste trabalho, e a todos que
me acompanharam nesta jornada.*

Agradecimentos

A Deus, primeiramente, por esta oportunidade, pelo dom do discernimento e por se mostrar presente nos momentos mais difíceis desta caminhada. A Ele dedico toda minha gratidão por abençoar, guiar e iluminar os meus passos nesta importante etapa da minha vida.

Agradeço, de maneira especial, à minha família, pela compreensão e pelo incondicional apoio, e aos meus familiares, por todo incentivo, sem os quais não seria possível a realização deste trabalho. Agradeço também à minha prima Rose, por ter me recebido em sua casa.

Ao Carlos Renato, pelo amor, compreensão e companheirismo dedicados ao longo de todos estes anos e por ter estado ao meu lado em mais esta importante conquista.

Agradeço ao orientador Luis Antonio Aguirre, pela inestimável dedicação demonstrada durante todo o mestrado e pelas valiosas críticas e sugestões, que contribuíram não somente para a realização deste trabalho, mas também para meu crescimento e minha formação. Muito obrigada pela paciência e comprometimento sempre dedicados!

Gostaria de demonstrar a minha imensa gratidão e o meu sincero reconhecimento àquele que tem me acompanhado desde a graduação. Ao orientador Marcelo Vieira Corrêa, o meu muitíssimo obrigada pelo acolhimento, pelas palavras de incentivo e encorajamento e por todo apoio e confiança dedicados durante todo esse tempo.

Ao professor Eduardo Mazoni Mendes, por toda atenção, disponibilidade e prestatividade demonstradas ao longo de todo o mestrado.

Ao professor Hani Camille Yehia, pelo apoio e compreensão.

A todos os demais professores que contribuíram para minha formação durante o mestrado: Leonardo Antônio Borges Tôrres, José Luiz Silvino e Carlos Andrey Maia.

Aos colegas do CPDEE, pelo companheirismo.

Aos colegas do grupo MACSIN, especialmente ao Davidson Lafitte Firmo, pela prestatividade, e à Mara Cristina Coelho, pela amizade e convivência.

Ao Felipe de Brito Freitas e à Raquel Regis Ribeiro, por toda prestatividade e amizade.

À Patrícia Kennedy, por toda colaboração.

Ao pesquisador Michael Small, pela atenção prestada às minhas solicitações.

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), pelo apoio financeiro.

Epígrafe

“Dê-me, Senhor, agudeza para entender, capacidade para reter, método e faculdade para aprender, sutileza para interpretar, graça e abundância para falar. Dê-me, Senhor, acerto ao começar, direção ao progredir e perfeição ao concluir.”

Santo Tomás de Aquino

Resumo

A escolha da estrutura de modelos a serem utilizados para representar os regimes dinâmicos descritos pelos dados é um problema crucial na identificação de sistemas. O índice ERR (*error reduction ratio*), que se baseia na redução do erro de predição (de um passo à frente), apesar de ser um critério amplamente usado na seleção de estrutura, pode escolher termos incorretos ou redundantes em condições não ideais de identificação, ou seja, quando os dados disponíveis não são adequados (superamostrados ou ruidosos) ou quando o sinal de entrada é relativamente lento. Por sua vez, o critério SRR (*simulation error reduction ratio*), diferentemente do ERR, pode ser eficiente em condições não ideais de identificação, além de resultar em modelos mais compactos e, portanto, mais robustos. Entretanto, tal critério, que se baseia na redução do erro de simulação (predição livre), requer um esforço computacional significativamente grande. Um critério baseado na redução do erro de predição de dois passos à frente (ERR_2) é proposto neste trabalho a fim de ser empregado nos casos em que o sinal de entrada é relativamente lento. Para isso, são investigados três casos com dados simulados e dois com dados experimentais de sistemas reais. Resultados aqui apresentados mostram que a utilização de auto-consistência entre os critérios ERR e ERR_2 pode auxiliar na seleção de estrutura de modelos NARX polinomiais.

Abstract

The structure selection of models used to represent dynamics described by data is a crucial problem in system identification. The ERR index (error reduction ratio), which is based on the prediction error minimization (one step ahead), despite being a criterion widely used in structure selection, can choose incorrect or redundant terms in non ideal identification conditions, that is, when the available data are not suitable (oversampled or noisy) or when the input signal is relatively slow. On the other hand, the SRR criterion (simulation error reduction ratio), differently from the ERR, may be effective in non ideal identification conditions. Moreover, SRR yields more compact models that are, therefore, more robust. However, such criterion, which is based on the simulation error minimization (free-run prediction), requires a significantly large computational effort. Thus, in this work, a criterion based on the prediction error of two steps ahead minimization (ERR_2) is proposed to be applied on the cases in which the input signal is relatively slow. To accomplish that, were investigated five cases studies: three with simulated data and two cases with experimental data from real systems. The results presented here show that the use of self-consistency between the criteria ERR and ERR_2 can assist in the selection of structure of polynomial NARX models.

Sumário

Resumo	ix
Abstract	xi
Lista de Tabelas	xvii
Lista de Figuras	xxi
Lista de Símbolos	xxvii
Lista de Acrônimos	xxix
1 Introdução	1
1.1 Objetivo	6
1.2 Apresentação do Trabalho	6
2 Identificação de Sistemas	7
2.1 Introdução	7
2.2 Etapas de Identificação	9
2.2.1 Coleta de Dados	9
2.2.2 Escolha da Representação Matemática	10
2.2.3 Seleção da Estrutura do Modelo	12
2.2.4 Estimação dos Parâmetros do Modelo	13

2.2.5	Validação do Modelo	15
2.3	Considerações Finais	17
3	Seleção de Estrutura de Modelos Não-Lineares	19
3.1	Introdução	19
3.2	Critérios de Informação	20
3.3	Estimação dos Máximos Atrasos	22
3.4	Critério ERR	24
3.4.1	Condições Não Ideais de Identificação	26
3.5	Critério SRR	27
3.6	Agrupamentos de Termos e Coeficientes de Agrupamentos	28
3.7	Considerações Finais	31
4	Predição de Dois Passos à Frente	33
4.1	Introdução	33
4.2	Predição de Dois Passos à Frente	34
4.3	Relação entre Modelos de Um e Dois Passos à Frente . . .	36
4.4	Obtenção de \mathcal{M}_{1PF} a partir de \mathcal{M}_{2PF}	46
4.5	Metodologia Preliminar	55
4.6	Considerações Finais	59
5	Estudo de Casos Simulados	61
5.1	Introdução	61
5.2	Estudo de Casos	61
5.2.1	Primeiro Caso	62
5.2.2	Segundo Caso	83
5.2.3	Terceiro Caso	99
5.3	Metodologia Revisada	108
5.4	Considerações Finais	112

6	Estudo de Casos Experimentais	113
6.1	Introdução	113
6.2	Estudo de Casos	113
6.2.1	Primeiro Caso	115
6.2.2	Segundo Caso	125
6.3	Considerações Finais	135
7	Conclusões e Propostas Futuras	137
	Referências Bibliográficas	141
A	Rotinas Computacionais	153
A.1	apress	153
A.2	fnn_modificado	154
A.3	modelo_moda	156
A.4	my_aic	157
A.5	odccf	158
A.6	orthregkp	159
A.7	predicao	160
B	Funções de Correlação	163

Lista de Tabelas

4.1	Agrupamentos de termos presentes nos modelos de um e dois passos à frente apresentados nos Exemplos 4.3.1, 4.3.2 e 4.3.3.	40
4.2	Termos primários e termos secundários dos modelos de dois passos à frente apresentados nos Exemplos 4.3.1 e 4.3.2.	42
4.3	Relação entre os termos primários de um modelo de dois passos à frente e os termos de seu respectivo modelo de um passo à frente.	43
4.4	Relação entre os termos do modelo (4.18) e os possíveis termos de seu respectivo modelo de um passo à frente.	47
4.5	Relação entre os termos do modelo (4.20) e os possíveis termos de seu respectivo modelo de um passo à frente.	49
4.6	Relação entre os termos do modelo (4.22) e os possíveis termos de seu respectivo modelo de um passo à frente (parte 1).	51
4.7	Relação entre os termos do modelo (4.22) e os possíveis termos de seu respectivo modelo de um passo à frente (parte 2).	52
4.8	Relação entre os termos do modelo (4.22) e os possíveis termos de seu respectivo modelo de um passo à frente (parte 3).	53
5.1	Modelo moda obtido pelo critério ERR_1 a partir dos dados apresentados na Figura 5.2.	66

5.2	Modelo moda obtido pelo critério ERR_2 a partir dos dados apresentados na Figura 5.2.	69
5.3	Modelo moda obtido pelo critério ERR_1 a partir dos dados apresentados na Figura 5.8.	73
5.4	Modelo moda obtido pelo critério ERR_2 a partir dos dados apresentados na Figura 5.8.	75
5.5	Relação entre os termos do modelo da Tabela 5.4 e os possíveis termos de seu respectivo modelo de um passo à frente (parte 1).	76
5.6	Relação entre os termos do modelo da Tabela 5.4 e os possíveis termos de seu respectivo modelo de um passo à frente (parte 2).	77
5.7	Modelo obtido pelo critério ERR_1 a partir dos dados apresentados na Figura 5.8 após a eliminação do termo $y(k-2)$	79
5.8	Modelo final obtido pelo critério ERR_1 a partir dos dados apresentados na Figura 5.8.	81
5.9	Modelos obtidos em Piroddi e Spinelli (2003a) para o sistema \mathcal{S}_1 a partir de uma entrada AR.	82
5.10	Modelos obtidos em Wei e Billings (2008) para o sistema \mathcal{S}_1 a partir de uma entrada AR.	83
5.11	Modelo moda obtido pelo critério ERR_1 a partir dos dados apresentados na Figura 5.16.	87
5.12	Modelo moda obtido pelo critério ERR_2 a partir dos dados apresentados na Figura 5.16.	89
5.13	Modelo obtido pelo critério ERR_1 a partir dos dados apresentados na Figura 5.16 após a eliminação do termo $y(k-1)^2$	90
5.14	Relação entre os termos do modelo da Tabela 5.12 e os possíveis termos de seu respectivo modelo de um passo à frente.	91

5.15	Novo modelo moda obtido pelo critério ERR_1 a partir dos dados apresentados na Figura 5.16.	93
5.16	Novo modelo moda obtido pelo critério ERR_2 a partir dos dados apresentados na Figura 5.16.	96
5.17	Relação entre os termos do modelo da Tabela 5.16 e os possíveis termos de seu respectivo modelo de um passo à frente (parte 1).	97
5.18	Relação entre os termos do modelo da Tabela 5.16 e os possíveis termos de seu respectivo modelo de um passo à frente (parte 2).	98
5.19	Modelo moda obtido pelo critério ERR_1 a partir dos dados apresentados na Figura 5.26.	102
5.20	Modelo moda obtido pelo critério ERR_2 a partir dos dados apresentados na Figura 5.26.	104
5.21	Relação entre os termos do modelo da Tabela 5.20 e os possíveis termos de seu respectivo modelo de um passo à frente (parte 1).	105
5.22	Relação entre os termos do modelo da Tabela 5.20 e os possíveis termos de seu respectivo modelo de um passo à frente (parte 2).	106
5.23	Modelo obtido pelo critério ERR_1 a partir dos dados apresentados na Figura 5.26 após a eliminação do termo $y(k-2)^2$. . .	107

Lista de Figuras

4.1	Fluxograma que representa os passos utilizados na metodologia adotada neste trabalho.	56
5.1	Representação de uma massa de dados dividida em 500 janelas contendo 500 amostras cada.	62
5.2	Dados do sistema \mathcal{S}_1 para uma entrada aleatória.	63
5.3	Resultado gráfico da aplicação da função FAC ao sinal de entrada apresentado na Figura 5.2.	64
5.4	Resultado gráfico da aplicação do critério de informação de Akaike ao modelo moda obtido pelo critério ERR_1 a partir dos dados apresentados na Figura 5.2.	65
5.5	Resultado gráfico da aplicação da estatística APRESS ao modelo moda obtido pelo critério ERR_1 a partir dos dados apresentados na Figura 5.2.	66
5.6	Resultado gráfico da aplicação do critério de informação de Akaike ao modelo moda obtido pelo critério ERR_2 a partir dos dados apresentados na Figura 5.2.	67
5.7	Resultado gráfico da aplicação da estatística APRESS ao modelo moda obtido pelo critério ERR_2 a partir dos dados apresentados na Figura 5.2.	68

5.8	Dados do sistema \mathcal{S}_1 para uma entrada AR.	69
5.9	Resultado gráfico da aplicação da função FAC ao sinal de entrada apresentado na Figura 5.8.	70
5.10	Resultado gráfico da aplicação do critério de informação de Akaike ao modelo moda obtido pelo critério ERR_1 a partir dos dados apresentados na Figura 5.8.	71
5.11	Resultado gráfico da aplicação da estatística APRESS ao modelo moda obtido pelo critério ERR_1 a partir dos dados apresentados na Figura 5.8.	72
5.12	Resultado gráfico da aplicação do critério de informação de Akaike ao modelo moda obtido pelo critério ERR_2 a partir dos dados apresentados na Figura 5.8.	74
5.13	Resultado gráfico da aplicação da estatística APRESS ao modelo moda obtido pelo critério ERR_2 a partir dos dados apresentados na Figura 5.8.	75
5.14	Resultado gráfico da aplicação do critério de informação de Akaike ao modelo obtido pelo critério ERR_1 a partir dos dados apresentados na Figura 5.8 após a eliminação do termo $y(k - 2)$	80
5.15	Resultado gráfico da aplicação da estatística APRESS ao modelo obtido pelo critério ERR_1 a partir dos dados apresentados na Figura 5.8 após a eliminação do termo $y(k - 2)$	81
5.16	Dados do sistema \mathcal{S}_2 para uma entrada AR.	84
5.17	Resultado gráfico da aplicação da função FAC ao sinal de entrada apresentado na Figura 5.16.	85

5.18	Resultado gráfico da aplicação do critério de informação de Akaike ao modelo moda obtido pelo critério ERR_1 a partir dos dados apresentados na Figura 5.16.	86
5.19	Resultado gráfico da aplicação da estatística APRESS ao modelo moda obtido pelo critério ERR_1 a partir dos dados apresentados na Figura 5.16.	87
5.20	Resultado gráfico da aplicação do critério de informação de Akaike ao modelo moda obtido pelo critério ERR_2 a partir dos dados apresentados na Figura 5.16.	88
5.21	Resultado gráfico da aplicação da estatística APRESS ao modelo moda obtido pelo critério ERR_2 a partir dos dados apresentados na Figura 5.16.	89
5.22	Resultado gráfico da aplicação do critério de informação de Akaike ao novo modelo moda obtido pelo critério ERR_1 a partir dos dados apresentados na Figura 5.16.	92
5.23	Resultado gráfico da aplicação da estatística APRESS ao novo modelo moda obtido pelo critério ERR_1 a partir dos dados apresentados na Figura 5.16.	93
5.24	Resultado gráfico da aplicação do critério de informação de Akaike ao novo modelo moda obtido pelo critério ERR_2 a partir dos dados apresentados na Figura 5.16.	94
5.25	Resultado gráfico da aplicação da estatística APRESS ao novo modelo moda obtido pelo critério ERR_2 a partir dos dados apresentados na Figura 5.16.	95
5.26	Dados do sistema \mathcal{S}_3 para uma entrada AR.	99
5.27	Resultado gráfico da aplicação da função FAC ao sinal de entrada apresentado na Figura 5.26.	100

5.28	Resultado gráfico da aplicação do critério de informação de Akaike ao modelo moda obtido pelo critério ERR_1 a partir dos dados apresentados na Figura 5.26.	101
5.29	Resultado gráfico da aplicação da estatística APRESS ao modelo moda obtido pelo critério ERR_1 a partir dos dados apresentados na Figura 5.26.	102
5.30	Resultado gráfico da aplicação do critério de informação de Akaike ao modelo moda obtido pelo critério ERR_2 a partir dos dados apresentados na Figura 5.26.	103
5.31	Resultado gráfico da aplicação da estatística APRESS ao modelo moda obtido pelo critério ERR_2 a partir dos dados apresentados na Figura 5.26.	104
6.1	Representação de uma massa de dados dividida em 51 janelas contendo 150 amostras cada.	114
6.2	Dados utilizados na identificação de um forno a gás.	115
6.3	Resultado gráfico da aplicação da função FAC ao sinal de entrada apresentado na Figura 6.2.	116
6.4	Resultado gráfico da aplicação do critério de informação de Akaike ao modelo moda obtido pelo critério ERR_1 a partir dos dados apresentados na Figura 6.2.	117
6.5	Resultado gráfico da aplicação da estatística APRESS ao modelo moda obtido pelo critério ERR_1 a partir dos dados apresentados na Figura 6.2.	118
6.6	Validação dinâmica do modelo (6.1).	119
6.7	Resultado gráfico da aplicação do critério de informação de Akaike ao modelo moda obtido pelo critério ERR_2 a partir dos dados apresentados na Figura 6.2.	120

6.8	Resultado gráfico da aplicação da estatística APRESS ao modelo moda obtido pelo critério ERR_2 a partir dos dados apresentados na Figura 6.2.	121
6.9	Validação dinâmica do modelo (6.4).	122
6.10	Comparação dos valores de RMSE correspondentes a modelos formados por termos presentes no modelo (6.1).	124
6.11	Dados utilizados na identificação de um pequeno aquecedor elétrico.	125
6.12	Resultado gráfico da aplicação da função FAC ao sinal de entrada apresentado na Figura 6.11.	126
6.13	Resultado gráfico da aplicação do critério de informação de Akaike ao modelo moda obtido pelo critério ERR_1 a partir dos dados apresentados na Figura 6.11.	127
6.14	Resultado gráfico da aplicação da estatística APRESS ao modelo moda obtido pelo critério ERR_1 a partir dos dados apresentados na Figura 6.11.	128
6.15	Validação dinâmica do modelo (6.5).	129
6.16	Resultado gráfico da aplicação do critério de informação de Akaike ao modelo moda obtido pelo critério ERR_2 a partir dos dados apresentados na Figura 6.11.	131
6.17	Resultado gráfico da aplicação da estatística APRESS ao modelo moda obtido pelo critério ERR_2 a partir dos dados apresentados na Figura 6.11.	132
6.18	Validação dinâmica do modelo (6.8).	133
6.19	Validação dinâmica do modelo (6.9).	134
B.1	Resultado gráfico da aplicação da função FCC aos dados do sistema (B.7).	167

B.2	Resultado gráfico da aplicação da função FCC de ordem elevada aos dados do sistema (B.7).	167
B.3	Resultado gráfico da aplicação da função ODCCF aos dados do sistema (B.7).	168
B.4	Resultado gráfico da aplicação da função FCC aos dados do sistema (B.8).	169
B.5	Resultado gráfico da aplicação da função FCC de ordem elevada aos dados do sistema (B.8).	170
B.6	Resultado gráfico da aplicação da função ODCCF aos dados do sistema (B.8).	170
B.7	Resultado gráfico da aplicação da função FCC aos dados do sistema (B.9).	172
B.8	Resultado gráfico da aplicação da função FCC de ordem elevada aos dados do sistema (B.9).	172
B.9	Resultado gráfico da aplicação da função ODCCF aos dados do sistema (B.9).	173

Lista de Símbolos

$u(k)$	sinal de entrada;
$y(k)$	sinal de saída;
$x(k)$	sinal de entrada ou sinal de saída;
$e(k)$	sinal de ruído;
$\nu(k)$	ruído branco;
$\xi(k)$	resíduo;
n_u	máximo atraso no sinal de entrada;
n_y	máximo atraso no sinal de saída;
n_e	máximo atraso no sinal de ruído;
ℓ	grau de não-linearidade;
τ_d	tempo morto;
$\boldsymbol{\psi}(k-1)$	vetor de regressores (pode conter observações até o instante $(k-1)$);
Ψ	matriz de regressores;
θ	parâmetro;
n_θ	número de parâmetros de um modelo;
n	número de termos de um modelo;
N	número de amostras;
τ	atraso;
$r(\tau)$	função de correlação (FAC ou FCC);

$\rho(\tau)$	função de correlação (ODACF ou ODCCF);
R	valor limiar;
ϵ^u	máximo valor absoluto do ruído presente no sinal de entrada;
ϵ^y	máximo valor absoluto do ruído presente no sinal de saída;
\mathcal{M}_{1PF}	modelo de um passo à frente;
\mathcal{M}_{2PF}	modelo de dois passos à frente;
a, b, c, d, e	atrasos presentes nos regressores;
\mathcal{S}	sistema;
$\min[\cdot]$	valor mínimo;
$\max[\cdot]$	valor máximo;
$\ \cdot\ _2$	norma euclidiana;
T	transposição;
$\hat{\cdot}$	valor estimado.

Lista de Acrônimos

AIC	critério de informação de Akaike (<i>Akaike information criterion</i>);
APRESS	<i>adjustable prediction error sum of squares</i> ;
AR	auto-regressivo;
BIC	critério de informação de Bayes (<i>Bayesian information criterion</i>);
EMQ	estendido de mínimos quadrados;
ERR	taxa de redução de erro (<i>error reduction ratio</i>);
ERR ₁	critério que se baseia na redução do erro de predição de um passo à frente (corresponde ao próprio ERR);
ERR ₂	critério que se baseia na redução do erro de predição de dois passos à frente;
FAC	função de autocorrelação;
FCC	função de correlação cruzada;
FIR	resposta finita ao impulso;
FNN	falsos vizinhos mais próximos (<i>false nearest neighbors</i>);
FPE	erro final de predição (<i>final prediction error</i>);
FROE	<i>forward-regression orthogonal estimator</i> ;
FRP	<i>forward-regression with pruning</i> ;

IFOS	<i>integrated forward orthogonal search;</i>
MDL	comprimento mínimo de descrição (<i>minimum description length</i>);
MQ	mínimos quadrados;
NARMAX	modelo não-linear auto-regressivo de média móvel com entradas exógenas (<i>nonlinear autoregressive moving average model with exogenous inputs</i>);
NARX	modelo não-linear auto-regressivo com entradas exógenas (<i>nonlinear autoregressive model with exogenous inputs</i>);
ODACF	<i>omni-directional auto-correlation function;</i>
ODCCF	<i>omni-directional cross-correlation function;</i>
PEM	<i>prediction error minimization;</i>
PRESS	<i>predicted residual sum of squares;</i>
RMSE	raiz do erro médio quadrático (<i>root mean squared error</i>);
SEM	<i>simulation error minimization;</i>
SEMP	<i>simulation error minimization with pruning;</i>
SRR	critério que se baseia na redução do erro de simulação (predição livre) (<i>simulation error reduction ratio</i>);
TSEMP	<i>two-stage simulation error minimization with pruning.</i>

Introdução

“A sabedoria é filha da experiência.”

Leonardo da Vinci

A modelagem de sistemas é de vital importância para o desenvolvimento da ciência e da tecnologia. A evolução do conhecimento científico ao longo dos séculos vem se apoiando em regras e leis matemáticas que descrevem fenômenos físicos observados na natureza. A modelagem permitiu estudar sistemas minúsculos, como o átomo, e sistemas imensos, como o sistema solar.

Representar, por meio de modelos matemáticos, sistemas e fenômenos observados sempre foi um desafio. Modelagem matemática é a área do conhecimento que estuda maneiras de desenvolver e implementar modelos matemáticos de sistemas reais. Tais modelos têm sido utilizados ao longo da história para os mais diversos fins, tais como projeto de sistemas de monitorização e controle, predição, estimação de estados, simulação e treinamento.

Durante muito tempo, modelos lineares foram utilizados para representar sistemas reais, que são, em sua maioria, não-lineares. Isso ocorreu em virtude das restrições teóricas e computacionais e das facilidades encontradas na análise e no projeto de sistemas lineares. Entretanto, um modelo linear aproxima a dinâmica de um sistema não-linear apenas na vizinhança de um ponto de operação (Fiedler-Ferrara e Prado, 1994; Ogata, 2003). Portanto, a representatividade de um modelo linear para um sistema não-linear é limitada. A utilização de modelos lineares impede que uma gama

de comportamentos dinâmicos apresentados pelos sistemas reais seja reproduzida, tais como ciclos limite, bifurcações, dinâmicas quasi-periódicas e caos (Aguirre, 1994b). Com o desenvolvimento tecnológico, principalmente na área computacional, tornou-se possível o desenvolvimento de técnicas de modelagem não-linear.

A escolha de qual representação deve ser utilizada na modelagem de sistemas não-lineares ainda é uma questão que parece não ter uma resposta definitiva. Há uma grande variedade de representações não-lineares, tais como redes neurais artificiais (Elsner, 1992; Henrique et al., 1998; Braga et al., 2000; Amaral, 2001; Haykin, 2001), funções de base radial (Casdagli, 1989; Alves, 2004), *wavelets* (Graps, 1995; Billings e Coca, 1999), séries de Volterra (Volterra, 1930; Billings, 1980), modelos de blocos interconectados (Wiener, 1958; Wigren, 1993; Patwardhan et al., 1998; Pearson e Pottmann, 2000; Coelho, 2002), modelos polinomiais e racionais (Chen e Billings, 1989; Haber e Unbehauen, 1990; Foss e Johansen, 1992; Billings e Chen, 1989; Billings e Zhu, 1991; Zhu e Billings, 1991, 1993; Corrêa, 1997, 2001; Campos, 2007). Independente do tipo de representação usada, um problema comum a todas elas é a determinação da estrutura.

A seleção de estrutura representa uma etapa extremamente importante no processo de identificação de sistemas. Essa importância encontra-se no fato de as características dinâmicas e estáticas do sistema estarem intimamente ligadas à sua estrutura.

A escolha da ordem é uma questão muito importante na determinação da estrutura tanto de modelos lineares quanto de não-lineares (Rhodes e Morari, 1997a; Porcaro et al., 2009). Se a ordem do modelo for muito menor que a ordem do sistema real, sua dinâmica não poderá ser reproduzida pelo modelo. Por outro lado, se a ordem do modelo for muito maior que a necessária, diversos problemas numéricos e dinâmicos poderão ocorrer (Aguirre e Billings, 1995a). Por exemplo, o método de falsos vizinhos (FNN), apesar de ter sido desenvolvido originalmente por Kennel et al. (1992) a fim de determinar a menor dimensão de imersão, também pode ser empregado na estimação da ordem de modelos (Rhodes e Morari, 1995, 1997a; Bomberger e Seborg, 1998).

As associações existentes entre as variáveis envolvidas em um sistema, que são normalmente detectadas por funções de correlação, podem conter informações relevantes a respeito dos atrasos a serem considerados nos regressores empregados na identificação de tal sistema, facilitando, assim,

a etapa de seleção de estrutura. As funções de correlação linear, tais como a função de autocorrelação (FAC) e a função de correlação cruzada (FCC), são normalmente utilizadas quando se deseja analisar associações lineares entre os sinais. Por esse motivo, elas são amplamente usadas na identificação de sistemas lineares, principalmente na etapa de validação de modelos. Entretanto, essas funções podem fornecer resultados incorretos se efeitos não-lineares estiverem presentes nos dados (Zhang et al., 2007). Além disso, as funções de correlação linear não são capazes de detectar associações não-lineares entre sinais.

Vários métodos, tais como testes de correlação multi-dimensional (Billings e Zhu, 1994; Aguirre, 1997; Maddess et al., 2004), FCC de ordem elevada (Aguirre, 1995; Zhu e Billings, 1997) e testes de correlação multi-direcional (Mao e Billings, 2000), foram desenvolvidos com o intuito de detectar associações não-lineares. Entretanto, a maioria dos testes de correlação não-linear não é tão simples quanto os testes de correlação linear (Billings e Voon, 1983). A fim de detectar correlações lineares e não-lineares, uma série de FCCs de primeira ordem, denominada ODCCFs (*omni-directional cross-correlation functions*) normalizadas, foi proposta em Zhang et al. (2007).

Em Luna et al. (2006), os atrasos presentes nos regressores de modelos de séries temporais foram determinados empregando o critério de informação mútua parcial.

Em se tratando de modelos polinomiais, um dos fatores que contribui para aumentar a complexidade da seleção de estrutura na identificação de sistemas não-lineares é o fato de que o número de termos candidatos cresce rapidamente com o aumento do grau de não-linearidade e dos máximos atrasos nos sinais de entrada e saída do modelo. O conjunto de termos candidatos pode ser reduzido em função de conhecimento *a priori* do sistema (Aguirre et al., 2000). Entretanto, nem sempre há informação *a priori* que pode ser utilizada na determinação da estrutura do modelo. Além disso, mesmo quando há tal conhecimento, não é óbvio como ele pode ser empregado. Deve-se ressaltar que, se o número de termos do modelo for maior que o necessário, resultando em um modelo sobreparametrizado (modelo com um número excessivo de termos), seus termos redundantes, além de acarretarem grande tempo computacional e mal condicionamento numérico, podem fazer com que o modelo apresente regimes dinâmicos espúrios (Aguirre e Billings, 1995a; Piroddi e Spinelli, 2003a).

Diversos métodos podem, então, ser usados a fim de se determinar o número de termos de um modelo. Entre tais métodos, destacam-se os critérios de informação, tais como o critério do erro final de predição (FPE) (Akaike, 1970), o critério de informação de Akaike (AIC) (Akaike, 1974), o critério de informação de Bayes (BIC) (Schwarz, 1978) e o critério de comprimento mínimo de descrição (MDL) (Rissanen, 1978). Além dos critérios de informação, um outro método estatístico também pode ser utilizado na estimação do número de termos de um modelo. Trata-se da estatística APRESS (*adjustable prediction error sum of squares*) (Billings e Wei, 2008), que é derivada da estatística PRESS (*predicted error sum of squares*) (Allen, 1974), um tipo de validação cruzada (Stone, 1974) bastante utilizado.

Como uma alternativa na seleção de estrutura, Kadtko et al. (1993) apresentaram um método bastante simples, que consiste em eliminar os termos do modelo que apresentam coeficientes muito menores que os demais, o que é, em geral, ineficiente no caso de sistemas não-lineares, especialmente quando há ruído presente nos dados (Aguirre, 1994a). Uma técnica complementar, baseada nos conceitos de agrupamentos de termos e coeficientes de agrupamentos, foi apresentada em Aguirre e Billings (1995b). Tal técnica auxilia na identificação de termos espúrios, gerando modelos concisos e dinamicamente válidos. Em Aguirre e Mendes (1996), é mostrado como o número e a localização de pontos fixos de modelos polinomiais podem ser determinados a partir de seus agrupamentos de termos e de seus respectivos coeficientes de agrupamentos e como tal informação pode ser utilizada na seleção de estrutura de modelos polinomiais não-lineares.

A programação genética (Koza, 1992) é um método de otimização simbólica que pode ser usado na seleção de estrutura de modelos não-lineares (Gray et al., 1998; Metenidis et al., 2004; Rodríguez-Vázquez et al., 2004; Madár et al., 2005; Beligiannis et al., 2005). Apesar de sua aplicabilidade prática, a programação genética pode causar alguns problemas computacionais como, por exemplo, a geração de um número excessivo de parâmetros. Um método de otimização multi-objetivo baseado na programação genética e destinado à seleção de estrutura de modelos NARMAX polinomiais foi proposto em Rodríguez-Vázquez et al. (2004).

Na tentativa de encontrar uma solução para o problema da seleção de estrutura de modelos não-lineares, um algoritmo baseado na correlação

quadrada e na informação mútua foi proposto em Wei e Billings (2008). Com esse mesmo propósito, Mendes e Billings (2001) desenvolveram um método alternativo que combina o uso de um estimador ortogonal com o de um procedimento similar aos do tipo *all subset selection*, que, segundo os autores, pode ser empregado com sucesso até mesmo no caso de sistemas caóticos.

Um critério amplamente usado na seleção de estrutura é o ERR (*error reduction ratio*) (Billings et al., 1989), que se baseia na redução do erro de predição (de um passo à frente). Apesar de sua ampla aplicabilidade, foi mostrado, em Piroddi e Spinelli (2003a,c), que o ERR pode escolher termos incorretos ou redundantes em condições não ideais de identificação, ou seja, quando os dados disponíveis não são adequados (superamostrados ou ruidosos) ou quando o sinal de entrada é relativamente lento.

Um novo critério de seleção de estrutura foi apresentado em Piroddi e Spinelli (2003a). Tal critério, denominado SRR (*simulation error reduction ratio*), se baseia na redução do erro de simulação (predição livre). Em Piroddi e Spinelli (2003d), foi mostrado que o SRR, diferentemente do ERR, pode ser eficiente em condições não ideais de identificação, além de resultar em modelos mais compactos e, portanto, mais robustos. Entretanto, tal critério apresenta algumas desvantagens, tais como: (i) esforço computacional significativamente grande, (ii) inviabilidade em identificação de sistemas caóticos em virtude de sua extrema sensibilidade a condições iniciais (Piroddi, 2008) e (iii) dificuldade presente na aplicação em modelos de séries temporais (modelos sem entrada).

Em virtude disso, é proposto, neste trabalho, um critério intermediário entre o ERR e o SRR, ou seja, um critério baseado na redução do erro de predição de k passos à frente a fim de ser utilizado como uma ferramenta auxiliar na seleção de estrutura de modelos NARX polinomiais. Para isso, escolheu-se $k = 2$, dando origem, assim, a um novo critério de seleção de estrutura, que, por sua vez, se baseia na redução do erro de predição de dois passos à frente (ERR_2). A fim de identificar e eliminar possíveis termos espúrios selecionados pelo ERR, é empregada, neste trabalho, auto-consistência entre os critérios ERR e ERR_2 . Um termo é considerado auto-consistente quando é selecionado tanto pelo ERR quanto pelo ERR_2 , sendo necessário, para isso, considerar a correspondência entre as estruturas geradas por cada um desses critérios, já que a estrutura obtida a partir da predição de dois passos à frente corresponde a um modelo de

dois passos. Para isso, é analisado, neste trabalho, um caso específico de condição não ideal de identificação, caso no qual o sinal de entrada é relativamente lento.

1.1 Objetivo

A razão deste trabalho consiste em desenvolver um critério baseado na redução do erro de predição de dois passos à frente (ERR_2) a fim de auxiliar na seleção de estrutura de modelos NARX polinomiais nos casos em que o sinal de entrada utilizado no processo de identificação é relativamente lento.

1.2 Apresentação do Trabalho

No Capítulo 2, são descritas, de forma sucinta, as principais etapas do processo de identificação de sistemas. O Capítulo 3 apresenta os métodos de seleção de estrutura usados e/ou citados ao longo deste trabalho. No Capítulo 4, é proposto um novo método de seleção de estrutura baseado na redução do erro de predição de dois passos à frente. No Capítulo 5, tal método é aplicado a dados simulados e, no Capítulo 6, a dados experimentais de sistemas reais. As conclusões e as propostas futuras são apresentadas no Capítulo 7. As rotinas computacionais desenvolvidas ao longo deste trabalho se encontram no Apêndice A. Por fim, é apresentado, no Apêndice B, um estudo que avalia o uso da função de correlação na seleção de estrutura.

Identificação de Sistemas

“Um pouco de ciência nos afasta de Deus. Muito, nos aproxima.”

Louis Pasteur

2.1 Introdução

A tentativa de explicar ou de reproduzir comportamentos de sistemas físicos é algo que há muito tempo desperta o interesse de pesquisadores. Com o desenvolvimento dos processos industriais e a necessidade de controlá-los, tornou-se necessário desenvolver modelos que reproduzissem suas características estáticas e também dinâmicas. 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, otimizar o comportamento do sistema, analisar e projetar controladores, permitir detecção eficiente de falhas no sistema, estimar variáveis do processo que não podem ser medidas diretamente, permitir o estudo do sistema em regiões de operação despendiosas ou problemáticas no sistema real, permitindo um treinamento de operação seguro e eficiente (Matko et al., 1992).

A obtenção de um modelo que reproduza exatamente o comportamento de um sistema é algo impossível. Na prática, os modelos se aproximam satisfatoriamente do comportamento do sistema para uma determinada aplicação específica e uma faixa de operação limitada. Portanto, não existe

um único modelo para um determinado sistema, mas diversos modelos para o mesmo.

O comportamento de um sistema pode ser representado por diferentes tipos de modelos. A escolha do tipo de modelo depende das características estáticas e/ou dinâmicas do sistema, bem como do propósito para qual o modelo é determinado. Os modelos matemáticos, no entanto, permitem aplicações mais avançadas, sendo, por esse motivo, os mais utilizados, seja na engenharia, biologia, medicina, economia ou em outras áreas (Ljung, 1999).

Uma das formas de se obter modelos matemáticos é a *modelagem caixa branca*¹. Essa técnica se baseia nas informações sobre o sistema a ser modelado e nas leis físicas que descrevem os fenômenos envolvidos. Uma importante característica da modelagem fenomenológica é o significado físico dos modelos obtidos. Como desvantagem dessa técnica, pode-se citar a sua inviabilidade no caso de sistemas grandes e complexos. Os modelos obtidos são normalmente representados por equações diferenciais no domínio do tempo ou por funções complexas no domínio de Laplace.

*Identificação de sistemas*² é uma área do conhecimento que estuda maneiras de modelar e analisar sistemas a partir de observações, ou seja, de dados coletados diretamente dos sistemas (Ljung, 1999). Uma grande vantagem apresentada pela modelagem caixa preta é a relativa facilidade de obtenção de modelos em virtude do pouco ou nenhum conhecimento prévio do sistema exigido por essa técnica. Por outro lado, a ausência do significado físico dos modelos obtidos e o número excessivo de parâmetros são exemplos de desvantagens desse tipo de modelagem.

A *modelagem caixa cinza*, também conhecida como *identificação caixa cinza*, busca combinar as vantagens dos procedimentos de identificação caixa preta e caixa branca. Nesse caso, tanto dados de entrada e saída obtidos do sistema quanto informação auxiliar são usados no processo de identificação.

Neste capítulo, é apresentada uma breve discussão referente às principais etapas de identificação de sistemas e às ferramentas utilizadas ao longo deste trabalho.

¹Também conhecida como *modelagem pela física* ou *natureza do processo* ou ainda *modelagem fenomenológica* ou *conceitual*.

²Também conhecida como *modelagem* (ou *identificação*) *caixa preta* ou *modelagem empírica*.

2.2 Etapas de Identificação

As principais etapas de um problema de identificação de sistemas são:

- coleta de dados;
- escolha da representação matemática;
- seleção da estrutura do modelo;
- estimação dos parâmetros do modelo;
- validação do modelo.

Essas etapas são empregadas tanto na identificação de sistemas lineares quanto de sistemas não-lineares. Tais etapas são abordadas de forma sucinta nas seções seguintes. Maiores detalhes podem ser encontrados em Aguirre (2004).

2.2.1 Coleta de Dados

Durante o processo de identificação, o sinal de excitação apresenta duas importantes funções. Uma delas é excitar as características dinâmicas e estáticas do sistema em toda a faixa de frequência de interesse. A outra está relacionada ao perfil de amplitudes desse sinal, que é responsável pela excitação das não-linearidades presentes no sistema. As características não excitadas não podem ser modeladas em virtude de não estarem contidas nos dados coletados. Portanto, o sinal de excitação deve ser persistentemente excitante.

Uma etapa importante no processo de identificação é a escolha do tempo ou período de amostragem, que corresponde ao intervalo entre duas amostras. Na prática, a frequência de amostragem é normalmente escolhida entre 5 e 10 vezes a maior frequência de interesse.

Um sinal pode ser considerado subamostrado quando o tempo de amostragem utilizado na coleta dos dados é muito grande. Por outro lado, quando o tempo de amostragem é muito pequeno, tem-se um sinal superamostrado. Como descrito por Billings e Aguirre (1995), um sinal superamostrado provoca instabilidade numérica e elevado esforço computacional devido ao mal condicionamento da matriz de regressores, dificultando a determinação da estrutura do modelo. Nesse caso, o sinal deve

ser decimado a fim de que se torne devidamente amostrado. Para isso, é necessário verificar o grau de correlação entre as observações adjacentes do sinal. Um procedimento que pode ser empregado na escolha do melhor tempo de amostragem foi proposto em Aguirre (1995).

2.2.2 Escolha da Representação Matemática

Há uma grande variedade de representações não-lineares que podem ser utilizadas na identificação de sistemas. A escolha do tipo de representação depende principalmente da finalidade do modelo e das ferramentas disponíveis para sua obtenção.

Na prática, as representações são selecionadas de acordo com a conveniência do projetista, a complexidade da aplicação ou a possibilidade de interligação com outros conceitos. Por exemplo, a maior parte das técnicas de controle é baseada em modelos lineares devido, basicamente, a dois fatores. O primeiro refere-se ao grande volume de ferramentas de análise e síntese estruturadas sobre a teoria de sistemas lineares. O segundo motivo reside no fato de que a realimentação reduz os efeitos não-lineares inerentes ao processo controlado.

Entre as representações não-lineares, destacam-se as redes neurais artificiais (Elsner, 1992; Henrique et al., 1998; Braga et al., 2000; Amaral, 2001; Haykin, 2001), que são modelos matemáticos inspirados no cérebro humano. Elas constituem um sistema de processamento paralelo e distribuído, composto de unidades de processamento simples (neurônios) que têm a capacidade de armazenar conhecimento experimental e torná-lo disponível para o uso. De uma maneira geral, as redes neurais são caracterizadas por sua topologia, pelas características computacionais de seus neurônios e pelo algoritmo de treinamento. As funções de base radial (Casdagli, 1989; Alves, 2004), que são casos particulares de redes neurais, constituem um outro tipo de representação não-linear.

Um outro tipo de representação não-linear são as *wavelets* (Graps, 1995; Billings e Coca, 1999). *Wavelets* são funções matemáticas que ampliam intervalos de dados, separando-os em diferentes componentes de frequência, permitindo a análise de cada componente em sua escala correspondente.

A série de Volterra (Volterra, 1930; Billings, 1980), apesar de sua ampla aplicabilidade na representação de sistemas não-lineares, requer um

número muito grande de parâmetros mesmo para sistemas pouco não-lineares. Isso ocorre em virtude de a saída ser explicada apenas em função da entrada.

Entre as representações não-lineares, podem-se destacar também os modelos de blocos interconectados (Wiener, 1958; Wigren, 1993; Patwardhan et al., 1998; Pearson e Pottmann, 2000; Coelho, 2002), compostos por um modelo dinâmico linear em cascata com uma função estática não-linear. A disposição de seus blocos define modelos com comportamentos diferentes. Quando a função estática precede o modelo dinâmico, tem-se o modelo de Hammerstein. Por outro lado, quando o modelo dinâmico precede a função estática, tem-se o modelo de Wiener. Tanto no modelo de Hammerstein quanto no de Wiener, não se tem acesso ao sinal intermediário aos blocos. Isso representa uma dificuldade na identificação de tais modelos. Uma possível solução é a utilização de informação *a priori* e/ou alguma restrição como, por exemplo, o conhecimento da curva estática.

Um outro tipo de representação não-linear são os modelos polinomiais e racionais (Chen e Billings, 1989; Haber e Unbehauen, 1990; Foss e Johansen, 1992; Billings e Chen, 1989; Billings e Zhu, 1991; Zhu e Billings, 1991, 1993; Corrêa, 1997, 2001; Campos, 2007).

Na área de identificação de sistemas não-lineares, o modelo NARMAX (*nonlinear autoregressive moving average model with exogenous inputs*) polinomial (Leontaritis e Billings, 1985a,b) tem atraído considerável interesse nos últimos anos (Korenberg et al., 1987; Leontaritis e Billings, 1987; Billings et al., 1989; Aguirre e Billings, 1995b; Jácome, 1996; Rodrigues, 1996; Mao e Billings, 1997; Aguirre e Jácome, 1998; Aguirre et al., 1998; Cassini, 1999; Palumbo e Piroddi, 2001; Coelho, 2002; Previdi, 2002; Piroddi e Spinelli, 2003a,d,c; Rodríguez-Vázquez et al., 2004; Spinelli et al., 2006; Iglesias et al., 2007; Jazayeri et al., 2007; Piroddi, 2008).

Em particular, os modelos polinomiais têm sido freqüentemente empregados por serem lineares nos parâmetros, no caso do modelo NARX (*nonlinear autoregressive model with exogenous inputs*), ou pseudo-lineares, no caso do modelo NARMAX, cujos parâmetros podem ser estimados por simples algoritmos da família dos mínimos quadrados (MQ).

O modelo NARMAX polinomial é um modelo discreto que explica o valor de saída em função de valores prévios dos sinais de saída $y(k)$, de entrada $u(k)$ e de ruído $e(k)$. O modelo NARMAX polinomial é representado

como (Leontaritis e Billings, 1985a,b):

$$y(k) = F^\ell[y(k-1), \dots, y(k-n_y), u(k-\tau_d), \dots, u(k-\tau_d-n_u+1), e(k-1), \dots, e(k-n_e)] + e(k), \quad (2.1)$$

sendo que $e(k)$ indica todos os efeitos que não podem ser bem representados por $F^\ell[\cdot]$, que é uma função polinomial com grau de não-linearidade $\ell \in \mathbb{N}$. Os termos τ_d , n_y , n_u e n_e representam, respectivamente, o tempo morto e os máximos atrasos em y , em u e em e .

A parte determinística da equação (2.1) pode ser expandida como o somatório de termos com graus de não-linearidade variando na faixa $1 \leq m \leq \ell$. Assim, cada termo de grau m poderá conter um fator de grau p do tipo $y(k-i)$ e um fator de grau $(m-p)$ do tipo $u(k-i)$ sendo multiplicado por um parâmetro representado por $c_{p,m-p}(n_1, \dots, n_m)$. Matematicamente, tem-se (Peyton-Jones e Billings, 1989):

$$y(k) = \sum_{m=0}^{\ell} \sum_{p=0}^m \sum_{n_1, n_m}^{n_y, n_u} c_{p,m-p}(n_1, \dots, n_m) \prod_{i=1}^p y(k-n_i) \prod_{i=p+1}^m u(k-n_i), \quad (2.2)$$

sendo que

$$\sum_{n_1, n_m}^{n_y, n_u} \equiv \sum_{n_1=1}^{n_y} \cdots \sum_{n_m=1}^{n_u}. \quad (2.3)$$

Ressalta-se que o limite superior será n_y se o somatório se referir a fatores do tipo $y(k-n_i)$. Por outro lado, se o somatório se referir a fatores do tipo $u(k-n_i)$, tal limite será n_u .

2.2.3 Seleção da Estrutura do Modelo

A seleção de estrutura constitui um problema crucial no processo de identificação de sistemas principalmente no caso de sistemas não-lineares. Na identificação de sistemas lineares, em se tratando de modelos polinômiais, o número de termos candidatos cresce linearmente com a ordem do modelo. Nesse caso, a seleção da estrutura do modelo se restringe, basicamente, à escolha do número de pólos e zeros e à determinação do atraso

puro de tempo (Aguirre, 2004). Se a ordem do modelo for maior que a ordem efetiva do sistema, os termos redundantes do modelo induzirão cancelamentos de pólos e zeros em sua função de transferência. Em van den Boom e van den Enden (1974), tal fato é usado na determinação da ordem de modelos lineares.

Entretanto, na identificação de sistemas não-lineares, ainda se tratando de modelos polinomiais, o número de termos candidatos cresce rapidamente com o aumento do grau de não-linearidade ℓ e dos máximos atrasos n_u , n_y e n_e (ver equação (2.1)), dificultando, assim, a escolha de termos apropriados, que é fundamental para a obtenção de um modelo compacto e robusto. É importante ressaltar que, se o número de termos do modelo for maior que o necessário, resultando em um modelo sobreparametrizado (modelo com um número excessivo de termos), seus termos redundantes, além de acarretarem grande tempo computacional e mal condicionamento numérico, podem fazer com que o modelo apresente regimes dinâmicos espúrios (Aguirre e Billings, 1995a; Piroddi e Spinelli, 2003a) ou, até mesmo, se torne instável. Isso faz com que o processo de seleção de estrutura seja muito mais complexo no caso de modelos não-lineares do que no de modelos lineares.

Portanto, enquanto termos redundantes em modelos lineares levam ao cancelamento de pólos e zeros em suas funções de transferência, tais termos destroem a dinâmica original do sistema quando presentes em modelos não-lineares. Assim, torna-se importante desenvolver procedimentos que permitam selecionar os termos realmente necessários na reprodução da dinâmica do sistema original. No Capítulo 3, serão apresentados os métodos de seleção de estrutura usados e/ou citados ao longo deste trabalho.

2.2.4 Estimação dos Parâmetros do Modelo

Após determinar a estrutura do modelo, deve-se estimar seus parâmetros utilizando os dados obtidos em experimentação (dados de identificação). Os algoritmos normalmente empregados na estimação de parâmetros de modelos polinomiais são derivados do estimador de mínimos quadrados.

A fim de estimar os parâmetros da equação (2.2), é conveniente colocar o modelo da seguinte forma:

$$\begin{aligned} y(k) &= \boldsymbol{\psi}^T(k-1)\hat{\boldsymbol{\theta}} + \xi(k) \\ &= \hat{y}(k) + \xi(k), \end{aligned} \tag{2.4}$$

sendo que k indica o instante considerado, $\boldsymbol{\psi}(k-1)$ corresponde ao vetor de regressores, que pode conter observações até o instante $(k-1)$, $\hat{\boldsymbol{\theta}}$ representa o vetor de parâmetros estimados³ e $\xi(k)$, o erro cometido pelo modelo ao tentar explicar $y(k)$ como $\boldsymbol{\psi}^T(k-1)\hat{\boldsymbol{\theta}}$.

É importante notar que, como o vetor de regressores é formado por observações tomadas até o instante $(k-1)$, $\hat{y}(k)$ corresponde à saída do modelo predita para o instante seguinte, ou seja, k . Por essa razão, tal predição é denominada predição de um passo à frente. Portanto, $\hat{y}(k)$ é a predição de um passo à frente de $y(k)$, sendo que $\boldsymbol{\psi}^T(k-1)\hat{\boldsymbol{\theta}}$ representa o preditor de um passo à frente. Em virtude disso, $\xi(k)$ é também conhecido como erro de predição de um passo à frente ou, simplesmente, erro de predição⁴.

Tomando a equação (2.4) ao longo do conjunto de dados e representando o conjunto de equações resultante em forma matricial, tem-se:

$$\mathbf{y} = \Psi\hat{\boldsymbol{\theta}} + \boldsymbol{\xi}, \quad (2.5)$$

sendo que Ψ é a matriz de regressores. É importante destacar que $\xi(k)$, que representa o erro cometido pelo modelo ao tentar explicar $y(k)$, não deve ser confundido com $e(k)$, que corrompe $y(k)$.

Isolando $\boldsymbol{\xi}$ na equação (2.5), tem-se:

$$\boldsymbol{\xi} = \mathbf{y} - \Psi\hat{\boldsymbol{\theta}}. \quad (2.6)$$

No estimador de mínimos quadrados, os parâmetros são estimados minimizando a seguinte função de custo:

$$J_{\text{MQ}} = \sum_{i=1}^N \xi(i)^2 = \boldsymbol{\xi}^T \boldsymbol{\xi}. \quad (2.7)$$

O estimador de mínimos quadrados pode ser expresso por:

$$\hat{\boldsymbol{\theta}}_{\text{MQ}} = (\Psi^T \Psi)^{-1} \Psi^T \mathbf{y}. \quad (2.8)$$

Sabe-se que, em certas situações, o estimador de mínimos quadrados (MQ) se torna polarizado. Uma possível solução para esse problema con-

³O símbolo ($\hat{}$) sobre as variáveis faz referência a valores estimados.

⁴ $\xi(k)$ é também conhecido como resíduo ou erro de modelagem.

siste em estender a matriz de regressores Ψ incluindo regressores de ruído a fim de que a parte modelável do erro seja modelada e o resíduo ξ seja branco e, portanto, não correlacionado com a matriz estendida de regressores. Para isso, o resíduo é utilizado como estimativa do ruído usado na formação da nova matriz de regressores. Esse procedimento refere-se ao estimador estendido de mínimos quadrados (EMQ).

Assim como o EMQ, existem outros estimadores derivados do MQ, tais como o estimador generalizado de mínimos quadrados (GMQ), o estimador de mínimos quadrados totais (MQT) e o estimador de mínimos quadrados com restrições (MQR). Recentemente, um estimador multi-objetivo que utiliza tanto dados dinâmicos quanto estáticos foi proposto em Barroso et al. (2007). Em Corrêa et al. (2002), o uso de informação *a priori* foi incorporado na estimação de parâmetros utilizando, para isso, o algoritmo elipsóide (Bland et al., 1981).

2.2.5 Validação do Modelo

A validação de um modelo é a etapa final do processo de identificação de sistemas. É por meio dos procedimentos de validação que se verifica a capacidade que um modelo tem de representar as características de interesse do sistema original. Um modelo será considerado válido se incorporar as características do sistema que são fundamentais para uma determinada aplicação. A escolha da ferramenta a ser empregada na verificação da qualidade de um modelo dependerá do propósito para o qual tal modelo foi projetado. Curiosamente, pode-se dizer que poucos são os trabalhos presentes na literatura que tratam exclusivamente do problema de validação de modelos.

Uma das vertentes da validação de modelos dinâmicos prevê o uso de funções de correlação a fim de detectar alguma possível dinâmica não modelada nos resíduos. No caso de sistemas lineares, deve ser verificado se os resíduos são brancos e não correlacionados com a entrada. Isso pode ser feito calculando-se as funções de autocorrelação dos resíduos e de correlação cruzada entre os resíduos e a entrada (Box e Jenkins, 1976; Ljung, 1999). Entretanto, esses testes não são suficientes no caso de sistemas não-lineares (Billings e Voon, 1983, 1986; Billings e Tao, 1991).

Assim, Billings e Voon (1986) propuseram um conjunto de funções de correlação destinado à validação de modelos não-lineares. Entretanto,

é importante ressaltar que tal método garante apenas que não existam correlações não modeladas nos resíduos. Portanto, os modelos validados podem não apresentar o mesmo comportamento dinâmico dos sistemas originais (Aguirre e Billings, 1994, 1995a). Por se basear em critérios estatísticos, esse tipo de validação é denominado validação estatística.

Uma nova função de correlação, denominada ODCCF⁵ (*omni-directional cross-correlation function*), foi proposta em Zhang et al. (2007). Em Zhu et al. (2007), tal função foi empregada na validação estatística de modelos. Nesse mesmo trabalho, foi realizado um estudo comparativo entre o conjunto de funções de correlação proposto em Billings e Voon (1986) e a função de correlação ODCCF. Em tal estudo, a função proposta em Zhang et al. (2007) se mostrou eficiente em casos nos quais o método de Billings e Voon falhou.

A validação dinâmica, por sua vez, visa verificar se um modelo é eficiente na recuperação de certas características dinâmicas do sistema. Uma das técnicas utilizadas para validar um modelo dinamicamente é a predição livre⁶, na qual o modelo é simulado indefinidamente reutilizando predições passadas. Para isso, é necessário que o modelo seja inicializado com dados medidos (dados de validação). Diversos índices podem ser empregados a fim de quantificar o desempenho do modelo na simulação livre (Fair, 1986). Um critério muito utilizado é o índice RMSE (*root mean squared error*), que é dado por:

$$\text{RMSE} = \frac{\sqrt{\sum_{k=1}^N (y(k) - \hat{y}(k))^2}}{\sqrt{\sum_{k=1}^N (y(k) - \bar{y})^2}}, \quad (2.9)$$

sendo que $\hat{y}(k)$ é a predição livre do sinal e \bar{y} é o valor médio do sinal medido $y(k)$. É importante ressaltar que, quanto menor for o valor do RMSE, maior será a eficiência do modelo na recuperação da característica dinâmica do sistema.

Na validação dinâmica de sistemas caóticos, são normalmente utilizados invariantes dinâmicos, tais como: expoentes de Lyapunov, dimensão de correlação, diagramas de bifurcação, mapas e seções de Poincaré e projeções bidimensionais de atratores. Tais invariantes podem ser facilmente encontrados na literatura (Abarbanel et al., 1993; Aguirre e Billings, 1994;

⁵A função ODCCF será apresentada no Apêndice B.

⁶Também conhecida como predição de infinitos passos à frente ou simulação livre.

Fiedler-Ferrara e Prado, 1994; Monteiro, 2002).

No caso da validação estática, o que se verifica é a capacidade de um modelo de recuperar a característica estática do sistema. Característica estática é a relação entre a saída e a entrada de um sistema em estado estacionário. Apesar de o uso desse tipo de validação não ser um procedimento muito comum, a validação estática tem sido recentemente empregada (Hernández e Arkun, 1993; Corrêa, 2001; Coelho, 2002; Campos, 2007).

2.3 Considerações Finais

Neste capítulo, foram descritas, de forma sucinta, as principais etapas do processo de identificação de sistemas. Os métodos de seleção de estrutura usados e/ou citados ao longo deste trabalho serão apresentados no Capítulo 3.

Seleção de Estrutura de Modelos Não-Lineares

“A dúvida é o princípio da sabedoria.”

Aristóteles

3.1 Introdução

A escolha da estrutura a ser utilizada para representar os regimes dinâmicos descritos pelos dados é um problema crucial na identificação de sistemas. Embora vários critérios relacionados com a seleção de regressores em modelos NARX/NARMAX polinomiais tenham sido propostos (Korenberg et al., 1987; Leontaritis e Billings, 1987; Billings et al., 1989; Aguirre e Billings, 1995b; Mao e Billings, 1997; Mendes e Billings, 2001; Piroddi e Spinelli, 2003a; Rodríguez-Vázquez et al., 2004; Zhang et al., 2007; Wei e Billings, 2008), a seleção de estrutura continua sendo um problema sem solução definitiva.

Neste capítulo, são apresentados os métodos de seleção de estrutura usados e/ou citados ao longo deste trabalho. Entre eles, encontram-se o critério de informação de Akaike (AIC) (Akaike, 1974) e a estatística APRESS (Billings e Wei, 2008), utilizados na determinação do número de termos de um modelo, e o método de falsos vizinhos (FNN), que, apesar de ter sido originalmente proposto a fim de determinar a menor dimensão de imersão (Kennel et al., 1992), também pode ser empregado na estimação dos máximos atrasos de um modelo (Rhodes e Morari, 1995, 1997a;

Bomberger e Seborg, 1998). Além desses métodos, também são apresentados o critério ERR (Billings et al., 1989), que se baseia na redução do erro de predição (de um passo à frente) e é amplamente usado na seleção de estrutura de modelos, e o critério SRR, baseado na redução do erro de simulação (predição livre) e recentemente proposto em Piroddi e Spinelli (2003a). Por fim, são apresentados os conceitos de agrupamentos de termos e coeficientes de agrupamentos, que também podem ser aplicados na seleção de estrutura de modelos.

A função de correlação é uma ferramenta que apresenta várias aplicações na área de identificação de sistemas. Tal função, além de ser utilizada, por exemplo, na determinação do tempo de amostragem, na estimação do tempo morto de um processo, na verificação de redundância de informações entre as variáveis de um sistema e na validação estatística, também pode ser empregada na seleção de estrutura, auxiliando na determinação dos atrasos presentes nos regressores de um modelo. Apesar de funções de correlação não terem sido usadas neste trabalho durante a etapa de seleção de estrutura, foi realizado um estudo comparativo entre algumas funções de correlação a fim de avaliar seus desempenhos quando empregadas na seleção de estrutura de modelos não-lineares. Tal estudo será apresentado no Apêndice B deste trabalho.

3.2 Critérios de Informação

Diversos métodos podem ser usados a fim de se determinar o número de termos de um modelo. Entre tais métodos, destacam-se os critérios de informação, tais como o critério do erro final de predição (FPE) (Akaike, 1970), o critério de informação de Akaike (AIC) (Akaike, 1974), o critério de informação de Bayes (BIC) (Schwarz, 1978) e o critério de comprimento mínimo de descrição (MDL) (Rissanen, 1978). Em Nakamura et al. (2006), foi realizado um estudo comparativo entre alguns dos critérios de informação na modelagem de séries temporais.

De acordo com o critério de informação de Akaike, o número ideal de termos em um modelo deve minimizar a seguinte função (Akaike, 1974):

$$\text{AIC}(n) = N \ln(\text{Var}[\xi(k)]) + 2n, \quad (3.1)$$

sendo que N corresponde ao número de amostras, $\text{Var}[\xi(k)]$, à variância

do resíduo $\xi(k)$ e n , ao número de termos do modelo.

À medida que termos são incluídos no modelo, o número de graus de liberdade aumenta, permitindo um ajuste mais exato aos dados. Assim, $Var[\xi(k)]$ diminui à medida que n aumenta. Entretanto, a partir de um determinado momento, a redução na variância do resíduo devido à inclusão de um novo termo passa a ser insignificante e, conseqüentemente, não justifica a inclusão de tal termo. Assim sendo, a primeira parcela da equação (3.1) quantifica a redução na variância do resíduo resultante da inclusão de um termo, ao passo que a segunda parcela penaliza a inclusão de cada termo.

É importante ressaltar que, como o critério de informação de Akaike é fundamentalmente estatístico, não se pode afirmar que o modelo com o número de termos selecionado pelo AIC seja capaz de reproduzir as propriedades dinâmicas do sistema original (Aguirre e Billings, 1994; Aguirre et al., 2006).

Assim como o AIC, outros critérios de informação (Gooijer et al., 1985) foram desenvolvidos no contexto de sistemas lineares. Contudo, tais critérios fornecem resultados consistentes para muitos sistemas não-lineares (Aguirre, 1994a). O critério de informação de Akaike foi implementado na função `my_aic`, que se encontra no Apêndice A.4 deste trabalho.

Além dos critérios de informação, um outro método estatístico também pode ser utilizado na estimação do número de termos de um modelo. Trata-se da estatística APRESS (*adjustable prediction error sum of squares*) (Billings e Wei, 2008), que é derivada da estatística PRESS¹ (*predicted residual sum of squares*) (Allen, 1974), um tipo de validação cruzada (Stone, 1974) bastante utilizado. Considerando novamente que N corresponde ao número de amostras e n , ao número de termos do modelo, a estatística APRESS pode ser definida da seguinte forma (Billings e Wei, 2008):

$$\text{APRESS}(n) = \left(\frac{1}{1 - C(n, \alpha)/N} \right)^2 \text{MSE}(n), \quad (3.2)$$

sendo que $\text{MSE}(n) = (1/N) \sum_{k=1}^N [y(k) - \hat{y}(k)]^2$ e $C(n, \alpha) = n\alpha$. Neste trabalho, utilizou-se $\alpha = 0, 1, \dots, 10$. Vale a pena observar que, para

¹Também conhecida como LOO (*leave-one-out*).

$\alpha = 0$, APRESS=MSE. A estatística APRESS foi implementada na função `apress`, que se encontra no Apêndice A.1 deste trabalho.

3.3 Estimação dos Máximos Atrasos

O método de falsos vizinhos (FNN) foi originalmente desenvolvido por Kennel et al. (1992) a fim de determinar a menor dimensão de imersão, ou seja, o número mínimo de coordenadas de atraso necessário na reconstrução do espaço de fases a partir de uma única série temporal. O FNN também pode ser empregado na estimação dos máximos atrasos nos sinais de entrada n_u e saída n_y de um modelo (Rhodes e Morari, 1995, 1997a; Bomberger e Seborg, 1998).

No método de falsos vizinhos, além de nenhuma estrutura específica de modelo ser assumida, os máximos atrasos são estimados diretamente a partir dos dados de entrada e saída. Entretanto, tal método apresenta como desvantagem o esforço computacional, que cresce rapidamente com o aumento do número de amostras e da dimensão do modelo. A fim de resolver esse problema, Feil et al. (2004) propuseram um algoritmo de agrupamento *fuzzy* baseado no método de falsos vizinhos. Em Bomberger e Seborg (1998), além do FNN, o método de números de Lipschitz (He e Asada, 1993) também foi empregado na estimação dos máximos atrasos dos modelos.

O algoritmo FNN empregado em modelos não autônomos² apresenta os seguintes passos (Rhodes e Morari, 1997a):

1. Forme o seguinte vetor de regressores:

$$\psi_{n_y, n_u}(k) = [y(k - \tau), \dots, y(k - n_y\tau), u(k - \tau), \dots, u(k - n_u\tau)]^T, \quad (3.3)$$

sendo que τ , n_u e n_y representam, respectivamente, o atraso de tempo e os máximos atrasos nos sinais de entrada $u(k)$ e saída $y(k)$.

²Um modelo é considerado não autônomo se ele depende explicitamente do tempo. No contexto de modelagem, um modelo com pelo menos uma entrada será denominado não autônomo.

2. Para um dado vetor de regressores $\boldsymbol{\psi}_{n_y, n_u}(k)$, encontre um outro vetor de regressores $\boldsymbol{\psi}_{n_y, n_u}(j)$ no conjunto de dados tal que a seguinte distância

$$d = \left\| \boldsymbol{\psi}_{n_y, n_u}(k) - \boldsymbol{\psi}_{n_y, n_u}(j) \right\|_2 \quad (3.4)$$

seja minimizada.

3. Determine se a seguinte expressão

$$\frac{|y(k) - y(j)|}{\left\| \boldsymbol{\psi}_{n_y, n_u}(k) - \boldsymbol{\psi}_{n_y, n_u}(j) \right\|_2} \leq R \quad (3.5)$$

é verdadeira ou falsa. Se a equação (3.5) for verdadeira, o vizinho será considerado verdadeiro. Caso contrário, será um falso vizinho. Deve-se ressaltar que R é um valor limiar escolhido previamente.

4. Repita os passos anteriores para todos os valores de k . Em seguida, calcule a porcentagem de falsos vizinhos presentes no conjunto de dados.
5. Repita os passos anteriores aumentando os valores de n_y e n_u até que a porcentagem de falsos vizinhos chegue a zero ou a algum valor aceitavelmente pequeno.

É importante ressaltar que o algoritmo FNN é sensível à escolha de R . Nos trabalhos nos quais o método de falsos vizinhos é abordado, parece não haver um consenso em relação à escolha apropriada de R . Enquanto Kennel et al. (1992) sugeriram que R fosse escolhido de forma que $R \geq 10$, outros autores preferiram limitar R a certas faixas de valores, tais como $10 \leq R \leq 15$ (Rhodes e Morari, 1997a) e $10 \leq R \leq 50$ (Bomberger e Seborg, 1998). Por outro lado, Abarbanel et al. (1993) recomendaram limitar R a um único valor ($R = 17,3$). Entretanto, acredita-se que a escolha de um único valor de R que seja adequado a todos os conjuntos de dados seja algo improvável de ocorrer.

Sabe-se que o FNN é eficiente quando aplicado a dados livres de ruído. Entretanto, um estudo realizado em Rhodes e Morari (1997b) mostrou que, para séries temporais contaminadas com ruído, a porcentagem de falsos

vizinhos para uma determinada dimensão de imersão cresce à medida que se aumenta o número de amostras. Uma solução para esse problema, proposta nesse mesmo trabalho, foi estendida em Rhodes e Morari (1997a) para o caso de modelos não autônomos.

A solução encontrada sugere que a equação (3.5) seja substituída por (Rhodes e Morari, 1997a):

$$\frac{|y(k) - y(j)|}{\left\| \boldsymbol{\psi}_{n_y, n_u}(k) - \boldsymbol{\psi}_{n_y, n_u}(j) \right\|_2} \leq R + \frac{2R\sqrt{n_y\epsilon^y + n_u\epsilon^u} + 2\epsilon^y}{\left\| \boldsymbol{\psi}_{n_y, n_u}(k) - \boldsymbol{\psi}_{n_y, n_u}(j) \right\|_2}, \quad (3.6)$$

sendo que ϵ^u e ϵ^y representam, respectivamente, o máximo valor absoluto do ruído presente nos sinais de entrada e saída. Em Rhodes e Morari (1996), foi observado que a simples inclusão dos termos ϵ^u e ϵ^y aumenta a eficiência do algoritmo até mesmo quando se utilizam valores muito menores que os reais valores de ϵ^u e ϵ^y . Esse novo método de falsos vizinhos, que considera o ruído presente nos dados, foi implementado na função `fnn_modificado`, que se encontra no Apêndice A.2 deste trabalho.

Durante o processo de identificação dos dados utilizados neste trabalho, foi possível perceber a subjetividade do método de falsos vizinhos em virtude de sua sensibilidade à escolha de R , ϵ^y e τ . Portanto, apesar das promessas dos autores, verificou-se que, nos casos analisados, o FNN não se mostrou útil na estimação dos máximos atrasos. Em virtude disso, o uso de tal método não será apresentado neste trabalho.

3.4 Critério ERR

Um critério amplamente usado na seleção de estrutura de modelos é o ERR (*error reduction ratio*) (Billings et al., 1989). Esse critério, que se baseia na redução do erro de predição (de um passo à frente), avalia a importância dos termos do modelo em função de sua habilidade de explicar a variância do sinal de saída.

A fim de definir a taxa de redução de erro (ERR), o modelo da equação (2.4) será representado da seguinte forma:

$$\begin{aligned}
y(k) &= \boldsymbol{\psi}^T(k-1)\hat{\boldsymbol{\theta}} + \xi(k) \\
&= \sum_{i=1}^{n_\theta} \hat{\theta}_i \psi_i(k-1) + \xi(k),
\end{aligned} \tag{3.7}$$

sendo que n_θ corresponde ao número de parâmetros.

Analogamente, o modelo representado em base ortogonal pode ser expresso por:

$$y(k) = \sum_{i=1}^{n_\theta} \hat{g}_i w_i(k-1) + \xi(k), \tag{3.8}$$

sendo que \hat{g}_i são os parâmetros estimados e w_i , os regressores ortogonais sobre os dados. Vale ressaltar que, na ausência de regressores, ou seja, quando $n_\theta = 0$, o erro de predição corresponde ao próprio sinal de saída $y(k)$.

A redução no valor da variância do resíduo, que ocorre à medida que novos termos são incluídos no modelo, pode ser normalizada em relação à variância total do sinal de saída. Logo, a taxa de redução de erro devido à inclusão do i -ésimo regressor no modelo é definida como (Chen et al., 1989):

$$[ERR]_i = \frac{\hat{g}_i^2 \langle \mathbf{w}_i, \mathbf{w}_i \rangle}{\langle \mathbf{y}, \mathbf{y} \rangle}, \tag{3.9}$$

sendo que o ERR indica a porção da variância da saída explicada pela inclusão de um novo termo no modelo. O operador $\langle \cdot \rangle$ representa o produto interno de dois vetores³. Ressalta-se que, normalmente, a escolha dos termos candidatos ao modelo é feita considerando aqueles que apresentam os maiores valores de ERR.

Os algoritmos de identificação que se baseiam na redução do erro de predição (resíduo) são conhecidos como PEM (*prediction error minimization*). Nesses algoritmos, o ERR de cada termo pode variar enormemente dependendo da ordem na qual eles são incluídos no modelo. Além disso, se um termo essencial à construção de um modelo não estiver presente no conjunto de termos candidatos, sua ausência será compensada por vários

³Assim, em relação à equação (3.9), $\langle \mathbf{w}_i, \mathbf{w}_i \rangle = \mathbf{w}_i^T \mathbf{w}_i$ e $\langle \mathbf{y}, \mathbf{y} \rangle = \mathbf{y}^T \mathbf{y}$.

termos espúrios. Isso faz com que os modelos obtidos pelos algoritmos PEM sejam, em alguns casos, imprecisos ou, até mesmo, instáveis na predição de longo alcance ou na simulação (predição livre). Em Piroddi e Spinelli (2003a,c), foi mostrado que, em condições não ideais de identificação⁴, ou seja, quando os dados disponíveis não são adequados (superamostrados ou ruidosos) ou quando o sinal de entrada é relativamente lento, os algoritmos PEM podem selecionar termos incorretos ou redundantes.

Um critério baseado nos conceitos de agrupamentos de termos e coeficientes de agrupamentos pode ser usado na prática para reduzir o tamanho dos modelos selecionados pelo ERR. Segundo tal critério, termos pertencentes a um mesmo agrupamento podem ser eliminados se o coeficiente de agrupamento (a soma dos parâmetros associados a todos os termos de um agrupamento) for muito menor que os coeficientes dos termos que compõem tal agrupamento. Entretanto, pesquisas mais recentes têm destacado que há casos nos quais a presença desses termos é necessária (Aguirre e Jácome, 1998). Piroddi e Spinelli (2003a) mostram que a eliminação de termos importantes afeta a precisão do modelo na simulação. A Seção 3.6 apresenta os conceitos de agrupamentos de termos e coeficientes de agrupamentos aplicados na seleção de estrutura de modelos não-lineares.

3.4.1 Condições Não Ideais de Identificação

O uso de dados contaminados por ruído pode ser considerado uma condição não ideal de identificação, pois, como a maioria dos algoritmos empregados na seleção de estrutura tenta obter um modelo que represente fielmente as informações contidas nos dados sem qualquer conhecimento da estrutura real do modelo, é provável que termos espúrios sejam selecionados quando os dados se encontrarem corrompidos em virtude da presença de ruído. O algoritmo ideal deveria, então, ser capaz de obter um modelo que minimizasse os efeitos dos termos espúrios a fim de que as dinâmicas fundamentais contidas nos dados fossem recuperadas pelo modelo identificado (Wei e Billings, 2008).

O uso de um tempo de amostragem muito pequeno também pode prejudicar a seleção da estrutura de um modelo por duas razões. A primeira está relacionada com o fato de que um sinal superamostrado provoca ins-

⁴As condições não ideais de identificação serão discutidas com mais detalhes na Seção 3.4.1.

tabilidade numérica devido ao mal condicionamento da matriz de regressores (Billings e Aguirre, 1995). A segunda está relacionada com o fato de que um tempo de amostragem muito pequeno pode fazer com que alguns termos candidatos se tornem indistinguíveis. Assim, tem-se que $y(k) \approx y(k-1)$. Conseqüentemente, o valor de ERR associado ao termo $y(k-1)$ tende a um. Como $y(k-1)$ é quase sempre o primeiro termo a ser selecionado pelo critério ERR, independente da estrutura real do modelo, e como a soma dos valores de ERR dos termos selecionados não pode ser maior que um, todos os demais termos do modelo apresentam valores de ERR muito pequenos, dificultando, assim, a etapa de seleção de estrutura (Piroddi e Spinelli, 2003a; Wei e Billings, 2008). Portanto, pode-se afirmar que o uso de dados superamostrados também pode ser considerado uma condição não ideal de identificação.

Um sinal de entrada de baixa frequência, quando aplicado a um sistema, freqüentemente gera um sinal de saída com dinâmica lenta. Isso pode fazer com que $y(k) \approx y(k-1)$. Conseqüentemente, assim como no caso de dados superamostrados, o valor de ERR associado ao termo $y(k-1)$ se torna muito próximo de um. Assim, $y(k-1)$ é quase sempre selecionado como o primeiro termo, sem se importar se tal termo pertence ou não à estrutura real do modelo. Isso mostra que o tipo de entrada usada no processo de identificação pode afetar profundamente a seleção da estrutura de um modelo, independente de qual algoritmo seja empregado (Piroddi e Spinelli, 2003a; Wei e Billings, 2008).

3.5 Critério SRR

Ao contrário do ERR, o critério SRR (*simulation error reduction ratio*) (Piroddi e Spinelli, 2003a,b,d; Spinelli et al., 2006; Piroddi, 2008), por sua vez, se baseia na redução do erro de simulação (predição livre).

Os algoritmos de identificação que se baseiam na redução do erro de simulação são conhecidos como SEM (*simulation error minimization*). Em Piroddi e Spinelli (2003d), foi mostrado que esses algoritmos, ao contrário dos algoritmos PEM, podem ser eficientes em condições não ideais de identificação. Além disso, tais algoritmos resultam em modelos mais compactos e, portanto, mais robustos. Isso faz com que os algoritmos SEM sejam mais apropriados para predição de longo alcance ou simulação.

Entretanto, os algoritmos SEM requerem um esforço computacional significativamente grande. Além disso, tais algoritmos não são viáveis em identificação de sistemas caóticos em virtude de sua extrema sensibilidade a condições iniciais (Piroddi, 2008). Uma outra desvantagem desses algoritmos é a dificuldade presente na aplicação em modelos de séries temporais (modelos sem entrada).

Muitos algoritmos que combinam estimação de parâmetros e seleção de estrutura de modelos em um único procedimento de otimização têm sido propostos na literatura (Leontaritis e Billings, 1987; Aguirre e Billings, 1995b; Mao e Billings, 1997). Um dos algoritmos mais populares e eficientes dessa categoria é o FROE (*forward-regression orthogonal estimator*) (Billings et al., 1989), que se baseia na redução do erro de predição. Em Piroddi e Spinelli (2003a), o critério ERR foi substituído pelo SRR no algoritmo FROE, transformando-se, assim, em um algoritmo do tipo SEM.

De fato, a importância real dos termos pode mudar durante o processo de construção do modelo. Isso acontece tanto com o critério ERR quanto com o SRR. Em virtude disso, Piroddi e Spinelli (2003a) propuseram a eliminação dos termos que não contribuem significativamente para a qualidade do modelo. Esse procedimento de eliminação é denominado *pruning*. A introdução do procedimento *pruning* em um algoritmo de identificação tende a gerar modelos mais compactos e, portanto, mais robustos.

Quando esse procedimento de eliminação é aplicado a um algoritmo que se baseia na redução do erro de predição, obtém-se o FRP (*forward-regression with pruning*). Da mesma forma, quando esse procedimento é aplicado a um algoritmo que se baseia na redução do erro de simulação, obtém-se o SEMP (*simulation error minimization with pruning*) (Piroddi e Spinelli, 2003a). Em Spinelli et al. (2006), um algoritmo de identificação, denominado TSEMP (*two-stage simulation error minimization with pruning*), foi proposto especificamente para lidar com dados superamostrados, que constitui um caso específico de condição não ideal de identificação.

3.6 Agrupamentos de Termos e Coeficientes de Agrupamentos

Os conceitos de agrupamentos de termos e coeficientes de agrupamentos são adequados para determinar o número, a localização e a simetria de

pontos fixos⁵ dos modelos, que representa uma informação útil na seleção da estrutura de modelos não-lineares em situações em que há algum conhecimento *a priori*, como, por exemplo, em problemas de recuperação de características estáticas do sistema original (Aguirre, 2004).

Se o tempo de amostragem for muito pequeno, tal que

$$\begin{aligned} y(k-1) &\approx y(k-2) \approx \dots \approx y(k-n_y) \\ u(k-1) &\approx u(k-2) \approx \dots \approx u(k-n_u), \end{aligned} \quad (3.10)$$

a equação (2.2) poderá ser reescrita da seguinte forma (Aguirre e Billings, 1995b):

$$y(k) \approx \sum_{n_1, n_m}^{n_y, n_u} c_{p, m-p}(n_1, \dots, n_m) \sum_{m=0}^{\ell} \sum_{p=0}^m y(k-1)^p u(k-1)^{m-p}. \quad (3.11)$$

Definição 3.6.1 *As constantes $\sum_{n_1, n_m}^{n_y, n_u} c_{p, m-p}(n_1, \dots, n_m)$ são os coeficientes dos agrupamentos de termos $\Omega_{y^p u^{m-p}}$, que contêm termos da forma $y(k-i)^p u(k-j)^{m-p}$ para $m = 0, \dots, \ell$ e $p = 0, \dots, m$. Tais coeficientes são chamados de coeficientes de agrupamentos e são representados por $\sum_{y^p u^{m-p}}$ (Aguirre e Billings, 1995b).*

Exemplo 3.6.1 ⁶

O modelo

$$\begin{aligned} y(k) = & 2,1579y(k-1) - 1,3203y(k-2) + 0,16239y(k-3) + \\ & 0,22480 \times 10^{-3}y(k-3)^3 - 0,48196 \times 10^{-2}y(k-1)^3 + \\ & 0,19463 \times 10^{-2}u(k-2) + 0,34160 \times 10^{-3}u(k-1) + \\ & 0,35230 \times 10^{-2}y(k-1)^2y(k-2) - \\ & 0,12162 \times 10^{-2}y(k-1)y(k-2)y(k-3) \end{aligned} \quad (3.12)$$

pode ser representado, conforme a equação (2.2), com

⁵Também conhecidos como pontos de equilíbrio.

⁶Aguirre e Billings, 1995b.

$$\begin{array}{ll}
c_{1,0}(1) = 2,1579 & c_{1,0}(2) = -1,3203 \\
c_{1,0}(3) = 0,16239 & c_{3,0}(3,3,3) = 0,22480 \times 10^{-3} \\
c_{3,0}(1,1,1) = -0,48196 \times 10^{-2} & c_{0,1}(2) = 0,19463 \times 10^{-2} \\
c_{0,1}(1) = 0,34160 \times 10^{-3} & c_{3,0}(1,1,2) = 0,35230 \times 10^{-2} \\
c_{3,0}(1,2,3) = -0,12162 \times 10^{-2} & c_{p,m-p}(\cdot) = 0.
\end{array}$$

Os coeficientes de agrupamentos do modelo (3.12) são:

$$\begin{aligned}
\sum_y &= c_{1,0}(1) + c_{1,0}(2) + c_{1,0}(3) = 0,99999 \\
\sum_{y^3} &= c_{3,0}(1,1,1) + c_{3,0}(1,1,2) + c_{3,0}(1,2,3) + c_{3,0}(3,3,3) = -2,2880 \times 10^{-3} \\
\sum_u &= c_{0,1}(1) + c_{0,1}(2) = 2,2879 \times 10^{-3},
\end{aligned}$$

que correspondem, respectivamente, aos agrupamentos Ω_y , Ω_{y^3} e Ω_u . \square

Portanto, um agrupamento da forma $\Omega_{y^p u^{m-p}}$ é um conjunto de termos do tipo $y(k-i)^p u(k-j)^{m-p}$ para $m = 0, \dots, \ell$ e $p = 0, \dots, m$, e seu respectivo coeficiente, $\sum_{y^p u^{m-p}}$, é o somatório dos coeficientes de todos os termos no modelo que pertencem ao referido agrupamento (Aguirre, 2004).

Vale a pena ressaltar que, conforme as aproximações consideradas na equação (3.10), os agrupamentos de termos e seus coeficientes de agrupamentos são definidos para um modelo em regime permanente com entradas constantes.

Os conceitos de agrupamentos de termos e coeficientes de agrupamentos podem ser utilizados na seleção de estrutura de modelos não-lineares auxiliando na identificação de termos espúrios, gerando modelos concisos e dinamicamente válidos.

Para isso, basta estimar uma família de modelos com um número crescente de termos escolhidos por um determinado critério de seleção. Em seguida, traçam-se gráficos que mostram a variação dos coeficientes de agrupamentos em função do número de termos dos modelos. Com base nesses gráficos, um agrupamento pode ser considerado espúrio quando:

- seu surgimento ocorre apenas para um número elevado de termos;
- seus coeficientes apresentam valores irrelevantes quando comparados aos demais;

- seus coeficientes oscilam em torno de zero à medida que o número de termos aumenta;
- seus coeficientes tendem a zero;
- seus coeficientes trocam de sinal, apresentando valores positivos e negativos.

Procedendo dessa forma, o conjunto de termos candidatos pode ser significativamente reduzido. Isso, além de representar uma economia de tempo de computação, diminui as chances de obter modelos com características dinâmicas espúrias (Aguirre, 2004).

Jácome (1996) mostrou que modelos polinomiais não-lineares identificados a partir do conjunto de agrupamentos efetivos (agrupamentos não espúrios) têm melhores chances de reproduzir as características dinâmicas do sistema original, bem como certas características estáticas. Além disso, tais modelos tendem a ser mais robustos aos efeitos prejudiciais da sobreparametrização (Aguirre e Billings, 1995a; Mendes e Billings, 1998).

É importante mencionar que, em alguns sistemas, o fato de o coeficiente de um determinado agrupamento de termos assumir valores muito pequenos não implica necessariamente que tal agrupamento seja espúrio (Aguirre e Jácome, 1998). Maiores detalhes podem ser encontrados em Aguirre et al. (2002).

3.7 Considerações Finais

Neste capítulo, foram apresentados os métodos de seleção de estrutura usados e/ou citados ao longo deste trabalho. Entretanto, ressalta-se que cada um deles apresenta alguma desvantagem ou limitação. No próximo capítulo, um critério baseado na redução do erro de predição de dois passos à frente é proposto a fim de auxiliar na seleção de estrutura de modelos NARX polinomiais.

Predição de Dois Passos à Frente

“A mente que se abre a uma nova idéia jamais voltará ao seu tamanho original.”

Albert Einstein

4.1 Introdução

Conforme mencionado no capítulo anterior, o critério ERR, que se baseia na redução do erro de predição (de um passo à frente), pode selecionar termos incorretos ou redundantes em condições não ideais de identificação. Em virtude disso, é proposto, neste trabalho, um critério baseado na redução do erro de predição de k passos à frente a fim de ser utilizado como uma ferramenta auxiliar na seleção de estrutura de modelos NARX polinomiais nos casos em que o sinal de entrada é relativamente lento. Para isso, escolheu-se $k = 2$, dando origem, assim, a um critério baseado na redução do erro de predição de dois passos à frente (ERR_2). A fim de identificar e eliminar possíveis termos espúrios selecionados pelo ERR, é empregada, neste trabalho, auto-consistência entre os critérios ERR e ERR_2 , sendo que um termo é considerado auto-consistente quando é selecionado tanto pelo ERR quanto pelo ERR_2 . Como a estrutura obtida pelo ERR_2 corresponde a um modelo de dois passos, é preciso, então, que os modelos obtidos por tal critério sejam representados na forma de modelos de um passo à frente. Para isso, realizou-se um estudo a partir do qual

foi possível determinar as relações entre os termos dos modelos de um e dois passos à frente. Exemplos apresentados neste capítulo mostram como essas relações podem ser empregadas na obtenção de um modelo de um passo à frente a partir de seu respectivo modelo de dois passos.

4.2 Predição de Dois Passos à Frente

Nesta seção, é apresentada a técnica de predição de dois passos à frente proposta neste trabalho. A fim de simplificar o procedimento aqui descrito, será utilizado, no exemplo apresentado a seguir, um modelo linear.

Para isso, considere o seguinte modelo (de um passo à frente):

$$y(k) = \theta_1 y(k-1) + \theta_2 y(k-2) + \theta_3 y(k-3) + \theta_4 u(k-1) + \theta_5 u(k-2), \quad (4.1)$$

cujos vetor de regressores é representado por:

$$\boldsymbol{\psi}(k-1) = [y(k-1) \quad y(k-2) \quad y(k-3) \quad u(k-1) \quad u(k-2)]^T. \quad (4.2)$$

Como a equação (4.1) se trata de um modelo de um passo à frente, verifica-se que seu vetor de regressores, representado pela equação (4.2), é formado por observações tomadas até o instante $(k-1)$. Para que o modelo (4.1) seja representado na forma de um modelo de dois passos à frente, tal modelo é, primeiramente, reescrito da seguinte forma:

$$y(k+1) = \theta_1 y(k) + \theta_2 y(k-1) + \theta_3 y(k-2) + \theta_4 u(k) + \theta_5 u(k-1). \quad (4.3)$$

Em seguida, substituindo a equação (4.1) em (4.3), chega-se ao seguinte modelo de dois passos à frente:

$$\begin{aligned} y(k+1) &= \theta_1 [\theta_1 y(k-1) + \theta_2 y(k-2) + \theta_3 y(k-3) + \theta_4 u(k-1) + \\ &\quad \theta_5 u(k-2)] + \theta_2 y(k-1) + \theta_3 y(k-2) + \theta_4 u(k) + \\ &\quad \theta_5 u(k-1) \\ &= \theta_1^2 y(k-1) + \theta_1 \theta_2 y(k-2) + \theta_1 \theta_3 y(k-3) + \theta_1 \theta_4 u(k-1) + \\ &\quad \theta_1 \theta_5 u(k-2) + \theta_2 y(k-1) + \theta_3 y(k-2) + \theta_4 u(k) + \\ &\quad \theta_5 u(k-1) \\ &= (\theta_1^2 + \theta_2) y(k-1) + (\theta_1 \theta_2 + \theta_3) y(k-2) + \theta_1 \theta_3 y(k-3) + \\ &\quad \theta_4 u(k) + (\theta_1 \theta_4 + \theta_5) u(k-1) + \theta_1 \theta_5 u(k-2). \end{aligned} \quad (4.4)$$

É importante ressaltar que, da mesma forma que $y(k)$ foi substituído pelo modelo (4.1) na equação (4.3), resultando no modelo de dois passos à frente apresentado em (4.4), $y(k-1)$ e $y(k-2)$ também podem ser substituídos em (4.3). Assim, substituindo-se k por $(k-1)$ em (4.1), tem-se:

$$y(k-1) = \theta_1 y(k-2) + \theta_2 y(k-3) + \theta_3 y(k-4) + \theta_4 u(k-2) + \theta_5 u(k-3). \quad (4.5)$$

Da mesma forma, substituindo-se k por $(k-2)$ em (4.1), chega-se a:

$$y(k-2) = \theta_1 y(k-3) + \theta_2 y(k-4) + \theta_3 y(k-5) + \theta_4 u(k-3) + \theta_5 u(k-4). \quad (4.6)$$

Em seguida, substituindo as equações (4.1), (4.5) e (4.6) em (4.3), chega-se ao seguinte modelo de dois passos à frente:

$$\begin{aligned} y(k+1) &= \theta_1 y(k) + \theta_2 y(k-1) + \theta_3 y(k-2) + \theta_4 u(k) + \theta_5 u(k-1) \\ &= \theta_1 [\theta_1 y(k-1) + \theta_2 y(k-2) + \theta_3 y(k-3) + \theta_4 u(k-1) + \\ &\quad \theta_5 u(k-2)] + \theta_2 [\theta_1 y(k-2) + \theta_2 y(k-3) + \theta_3 y(k-4) + \\ &\quad \theta_4 u(k-2) + \theta_5 u(k-3)] + \theta_3 [\theta_1 y(k-3) + \theta_2 y(k-4) + \\ &\quad \theta_3 y(k-5) + \theta_4 u(k-3) + \theta_5 u(k-4)] + \theta_4 u(k) + \\ &\quad \theta_5 u(k-1) \\ &= \theta_1^2 y(k-1) + 2\theta_1 \theta_2 y(k-2) + (2\theta_1 \theta_3 + \theta_2^2) y(k-3) + \\ &\quad 2\theta_2 \theta_3 y(k-4) + \theta_3^2 y(k-5) + \theta_4 u(k) + \\ &\quad (\theta_1 \theta_4 + \theta_5) u(k-1) + (\theta_1 \theta_5 + \theta_2 \theta_4) u(k-2) + \\ &\quad (\theta_2 \theta_5 + \theta_3 \theta_4) u(k-3) + \theta_3 \theta_5 u(k-4). \end{aligned} \quad (4.7)$$

Analisando-se os modelos (4.4) e (4.7), verifica-se que o termo mais adiantado de y (no caso, $y(k-1)$) encontra-se atrasado em relação a $y(k+1)$ em dois intervalos de amostragem, o que caracteriza um modelo de dois passos à frente. Assim, tanto o modelo (4.4) quanto o (4.7) correspondem ao modelo (4.1) representado na predição de dois passos à frente. Entretanto, verifica-se que o modelo (4.7) é desnecessariamente grande. Se o processo de iteração for feito continuamente, o tamanho do modelo crescerá infinitamente. Modelos como o (4.7), construídos a partir da iteração de um modelo menor, são denominados modelos degenerados (Judd e Nakamura, 2006; Nakamura e Small, 2006; Nakamura et al., 2006). A

fim de evitar um aumento desnecessário no tamanho dos modelos de dois passos à frente, estes serão representados, neste trabalho, utilizando a menor estrutura possível. Portanto, o modelo considerado neste caso para representar o modelo (4.1) na predição de dois passos à frente será o da equação (4.4).

Comparando-se os modelos (4.1) e (4.4), verifica-se que o modelo de dois passos à frente, além de ser formado por todos os termos presentes no modelo de um passo, também é composto pelo termo $u(k)$. Para que tal termo possa ser selecionado na predição de dois passos à frente, é necessário que ele esteja presente no conjunto de termos candidatos. Como, na prática, não é possível saber se o modelo de dois passos à frente apresenta algum termo formado pelo regressor $u(k)$, todos os possíveis termos compostos por tal regressor devem, então, ser incluídos no conjunto de termos candidatos a fim de que possam ser selecionados na predição de dois passos à frente.

O resíduo gerado na predição de dois passos à frente, que consiste, obviamente, no erro de predição de dois passos à frente, ao substituir o erro de predição (de um passo à frente) no critério ERR, dá origem a um novo critério de seleção de estrutura. Tal critério será denominado ERR_2 , pois se baseia na redução do erro de predição de dois passos à frente. Como o ERR é baseado na redução do erro de predição (de um passo à frente), tal critério será aqui denominado ERR_1 .

O procedimento necessário para se empregar a predição de dois passos à frente foi implementado no algoritmo de identificação `orthregkp`, originalmente criado por Eduardo Mendes. Tal algoritmo, que pode ser utilizado na predição de k passos à frente (sendo $k = 1, 2, 3, \dots$), se encontra no Apêndice A.6 deste trabalho.

4.3 Relação entre Modelos de Um e Dois Passos à Frente

Nesta seção, pretende-se estabelecer uma relação entre os termos de um modelo de dois passos e os termos de seu respectivo modelo de um passo à frente. Para isso, serão considerados, inicialmente, os seguintes exemplos:

Exemplo 4.3.1

Considere o seguinte modelo de um passo à frente:

$$y(k) = \theta_1 y(k-1) + \theta_2 y(k-2)^2 + \theta_3 u(k-1)u(k-2) + \theta_4 u(k-1) + \theta_5, \quad (4.8)$$

que, na predição de dois passos à frente, pode ser representado da seguinte forma:

$$y(k+1) = \theta_1 y(k) + \theta_2 y(k-1)^2 + \theta_3 u(k)u(k-1) + \theta_4 u(k) + \theta_5. \quad (4.9)$$

Substituindo a equação (4.8) em (4.9), tem-se:

$$\begin{aligned} y(k+1) &= \theta_1[\theta_1 y(k-1) + \theta_2 y(k-2)^2 + \theta_3 u(k-1)u(k-2) + \\ &\quad \theta_4 u(k-1) + \theta_5] + \theta_2 y(k-1)^2 + \theta_3 u(k)u(k-1) + \\ &\quad \theta_4 u(k) + \theta_5 \\ &= \theta_1^2 y(k-1) + \theta_1 \theta_2 y(k-2)^2 + \theta_1 \theta_3 u(k-1)u(k-2) + \\ &\quad \theta_1 \theta_4 u(k-1) + \theta_1 \theta_5 + \theta_2 y(k-1)^2 + \theta_3 u(k)u(k-1) + \\ &\quad \theta_4 u(k) + \theta_5. \end{aligned} \quad (4.10)$$

Exemplo 4.3.2

Considere o seguinte modelo de um passo à frente:

$$\begin{aligned} y(k) &= \theta_1 y(k-1)u(k-2)^2 + \theta_2 u(k-1)u(k-3) + \theta_3 u(k-2) + \\ &\quad \theta_4 y(k-2)u(k-1) + \theta_5, \end{aligned} \quad (4.11)$$

que, na predição de dois passos à frente, pode ser representado da seguinte forma:

$$\begin{aligned} y(k+1) &= \theta_1 y(k)u(k-1)^2 + \theta_2 u(k)u(k-2) + \theta_3 u(k-1) + \\ &\quad \theta_4 y(k-1)u(k) + \theta_5. \end{aligned} \quad (4.12)$$

Substituindo a equação (4.11) em (4.12), tem-se:

$$\begin{aligned}
y(k+1) &= \theta_1[\theta_1 y(k-1)u(k-2)^2 + \theta_2 u(k-1)u(k-3) + \theta_3 u(k-2) + \\
&\quad \theta_4 y(k-2)u(k-1) + \theta_5]u(k-1)^2 + \theta_2 u(k)u(k-2) + \\
&\quad \theta_3 u(k-1) + \theta_4 y(k-1)u(k) + \theta_5 \\
&= \theta_1^2 y(k-1)u(k-2)^2 u(k-1)^2 + \\
&\quad \theta_1 \theta_2 u(k-1)u(k-3)u(k-1)^2 + \theta_1 \theta_3 u(k-2)u(k-1)^2 + \\
&\quad \theta_1 \theta_4 y(k-2)u(k-1)u(k-1)^2 + \theta_1 \theta_5 u(k-1)^2 + \\
&\quad \theta_2 u(k)u(k-2) + \theta_3 u(k-1) + \theta_4 y(k-1)u(k) + \theta_5. \quad (4.13)
\end{aligned}$$

Exemplo 4.3.3

Considere o seguinte modelo de um passo à frente:

$$y(k) = \theta_1 y(k-1)^2 + \theta_2 u(k-1)u(k-2) + \theta_3 y(k-2)^2, \quad (4.14)$$

que, na predição de dois passos à frente, pode ser representado da seguinte forma:

$$y(k+1) = \theta_1 y(k)^2 + \theta_2 u(k)u(k-1) + \theta_3 y(k-1)^2. \quad (4.15)$$

Substituindo a equação (4.14) em (4.15), tem-se:

$$\begin{aligned}
y(k+1) &= \theta_1[\theta_1 y(k-1)^2 + \theta_2 u(k-1)u(k-2) + \theta_3 y(k-2)^2]^2 + \\
&\quad \theta_2 u(k)u(k-1) + \theta_3 y(k-1)^2 \\
&= \theta_1[\theta_1^2 y(k-1)^4 + 2\theta_1 \theta_2 y(k-1)^2 u(k-1)u(k-2) + \\
&\quad 2\theta_1 \theta_3 y(k-1)^2 y(k-2)^2 + \theta_2^2 u(k-1)^2 u(k-2)^2 + \\
&\quad 2\theta_2 \theta_3 u(k-1)u(k-2)y(k-2)^2 + \theta_3^2 y(k-2)^4] + \\
&\quad \theta_2 u(k)u(k-1) + \theta_3 y(k-1)^2 \\
&= \theta_1^3 y(k-1)^4 + 2\theta_1^2 \theta_2 y(k-1)^2 u(k-1)u(k-2) + \\
&\quad 2\theta_1^2 \theta_3 y(k-1)^2 y(k-2)^2 + \theta_1 \theta_2^2 u(k-1)^2 u(k-2)^2 + \\
&\quad 2\theta_1 \theta_2 \theta_3 u(k-1)u(k-2)y(k-2)^2 + \theta_1 \theta_3^2 y(k-2)^4 + \\
&\quad \theta_2 u(k)u(k-1) + \theta_3 y(k-1)^2. \quad (4.16)
\end{aligned}$$

A partir dos Exemplos 4.3.1, 4.3.2 e 4.3.3, é possível generalizar que:

- o grau de não-linearidade de um modelo de dois passos à frente pode ser maior que ou igual ao grau de não-linearidade de seu respectivo modelo de um passo.

O grau de não-linearidade de um modelo de dois passos à frente é igual ao de seu respectivo modelo de um passo quando, em tal modelo, o regressor $y(k-1)$ se encontra em um termo linear (Exemplo 4.3.1). Entretanto, quando tal regressor se encontra, em um modelo de um passo à frente, multiplicado por algum regressor de entrada ou algum outro regressor de saída ou apresenta grau de multiplicidade maior que um, o grau de não-linearidade do modelo de dois passos à frente é maior que o de seu respectivo modelo de um passo (Exemplos 4.3.2 e 4.3.3);

- os máximos atrasos nos sinais de entrada n_u e saída n_y de um modelo de dois passos à frente correspondem aos mesmos máximos atrasos de seu respectivo modelo de um passo.

Isso acontece porque o regressor $y(k-1)$, ao se transformar em $y(k)$ na predição de dois passos à frente, transfere os máximos atrasos do modelo de um passo para o seu respectivo modelo de dois passos à frente;

- alguns agrupamentos de termos presentes em um modelo de um passo à frente podem não estar contidos em seu respectivo modelo de dois passos.

Quando o regressor $y(k-1)$, em um modelo de um passo à frente, se encontra em um termo linear, todos os agrupamentos de termos desse modelo encontram-se presentes em seu respectivo modelo de dois passos. Vale a pena ressaltar que, nesse caso, os agrupamentos de termos de um modelo de um passo à frente são idênticos aos de seu respectivo modelo de dois passos (Exemplo 4.3.1). Por outro lado, quando o regressor $y(k-1)$ se encontra multiplicado por algum regressor de entrada ou algum outro regressor de saída ou apresenta grau de multiplicidade maior que um, todos os agrupamentos de termos de um modelo de um passo à frente podem estar presentes (Exemplo 4.3.3) ou não (Exemplo 4.3.2) em seu respectivo modelo

de dois passos. A partir da Tabela 4.1, é possível verificar, por exemplo, que o agrupamento Ω_{u^2y} , presente no modelo de um passo do Exemplo 4.3.2, não se encontra no modelo de dois passos à frente;

- presença/ausência de termo constante em modelos de um passo à frente implica presença/ausência de termo constante em modelos de dois passos e vice-versa.

Como pode ser visto nos Exemplos 4.3.1 e 4.3.2, o termo constante se encontra presente tanto nos modelos de um passo (equações (4.8) e (4.11)) quanto nos de dois passos à frente (equações (4.10) e (4.13)). O Exemplo 4.3.3, por sua vez, mostra que os modelos de um (equação (4.14)) e dois passos à frente (equação (4.16)) não apresentam termo constante.

Tabela 4.1: Agrupamentos de termos presentes nos modelos de um e dois passos à frente apresentados nos Exemplos 4.3.1, 4.3.2 e 4.3.3.

	Modelo de um passo	Modelo de dois passos
Exemplo 4.3.1	$\Omega_y, \Omega_{y^2}, \Omega_{u^2}, \Omega_u, \Omega_0$	$\Omega_y, \Omega_{y^2}, \Omega_{u^2}, \Omega_u, \Omega_0$
Exemplo 4.3.2	$\Omega_{u^2y}, \Omega_{u^2}, \Omega_u, \Omega_{uy}, \Omega_0$	$\Omega_{u^4y}, \Omega_{u^4}, \Omega_{u^3}, \Omega_{u^3y}, \Omega_{u^2}, \Omega_u, \Omega_{uy}, \Omega_0$
Exemplo 4.3.3	$\Omega_{y^2}, \Omega_{u^2}$	$\Omega_{y^4}, \Omega_{u^2y^2}, \Omega_{u^4}, \Omega_{u^2}, \Omega_{y^2}$

Além disso, é possível perceber que cada um dos modelos se expande de uma maneira diferente na predição de dois passos à frente. Isso depende de como o regressor $y(k-1)$ aparece no modelo de um passo à frente, ou seja, se tal regressor se encontra em um termo linear ($\theta_1 y(k-1)$ - Exemplo 4.3.1) ou multiplicado por algum regressor de entrada ou algum outro regressor de saída ($\theta_1 y(k-1)u(k-2)^2$ - Exemplo 4.3.2) ou, no caso mais complexo, com grau de multiplicidade maior que um ($\theta_1 y(k-1)^2$ - Exemplo 4.3.3).

Observando esses exemplos, verifica-se que modelos de dois passos podem ser facilmente obtidos a partir de modelos de um passo à frente. Entretanto, o contrário, ou seja, a obtenção de um modelo de um passo a partir de seu respectivo modelo de dois passos à frente, não é uma tarefa tão simples assim. Em virtude disso, diversos modelos foram analisados a

fim de se tentar encontrar a relação existente entre os termos dos modelos de um e dois passos à frente. Devido à complexidade presente na expansão do modelo do Exemplo 4.3.3 na predição de dois passos à frente, este trabalho se limitou a lidar apenas com modelos de um passo à frente nos quais o regressor $y(k-1)$ apresenta grau de multiplicidade igual a um.

A fim de facilitar o estudo das relações entre os termos dos modelos de um e dois passos à frente, foi necessário classificar os termos de um modelo de dois passos em dois tipos: *termos primários* e *termos secundários*. Para compreender a definição desses termos, é preciso lembrar que, ao representar um modelo de um passo à frente na forma de um modelo de dois passos, é necessário, primeiramente, que todos os seus termos sejam adiantados em um intervalo de amostragem. Assim, o regressor $y(k-1)$ se transforma em $y(k)$ que, em seguida, é substituído pelo modelo de um passo à frente. Dessa forma, todos os termos do modelo de um passo são multiplicados pelo(s) regressor(es) que acompanha(m) $y(k)$. No caso do Exemplo 4.3.2, todos os termos do modelo (4.11) são multiplicados pelo regressor $u(k-1)^2$. Os termos de um modelo de dois passos à frente gerados dessa forma serão denominados *termos primários*. Ressalta-se que, quando o regressor $y(k-1)$ se encontra em um termo linear no modelo de um passo à frente, os termos primários de seu respectivo modelo de dois passos correspondem aos próprios termos do modelo de um passo à frente (Exemplo 4.3.1), pois, nesse caso, $y(k)$ não se encontra multiplicado por nenhum outro regressor. Todos os demais termos adiantados, ou seja, todos os termos nos quais não consta o regressor $y(k)$, formam os termos restantes do modelo de dois passos à frente, que serão denominados *termos secundários*. Portanto, os termos de um modelo de dois passos à frente gerados pela substituição do regressor $y(k)$ pelo seu respectivo modelo de um passo corresponderão aos termos primários. Por outro lado, todos os demais termos presentes no modelo de dois passos à frente corresponderão aos termos secundários. A Tabela 4.2 apresenta os termos primários e os termos secundários dos modelos de dois passos à frente mostrados nos Exemplos 4.3.1 e 4.3.2.

Em relação à Tabela 4.2, um fato deve ser ressaltado. No modelo (4.10), o termo constante $\theta_1\theta_5$ pode ser considerado um termo primário, ao passo que a constante θ_5 , um termo secundário. Isso mostra que um termo constante em um modelo de dois passos à frente pode ser tanto primário quanto secundário. O modelo (4.13), por sua vez, mostra que a

constante θ_5 pode ser considerada um termo secundário. Assim, é possível afirmar que a presença de termo constante em modelos de um passo à frente sempre implicará termo constante do tipo secundário em modelos de dois passos e, conseqüentemente, todo termo constante (secundário ou não) em modelos de dois passos à frente implicará termo constante em modelos de um passo. Entretanto, quando as constantes do modelo (4.10) são somadas ($\theta_1\theta_5 + \theta_5$), tem-se uma única constante, que não é um termo puramente primário nem puramente secundário. Dessa forma, verifica-se que, em um modelo de dois passos à frente, uma constante jamais poderá ser considerada um termo primário, mas somente secundário (Exemplo 4.3.2) ou misto, ou seja, que não é puramente primário nem puramente secundário (Exemplo 4.3.1). Vale a pena mencionar que, como o modelo (4.14) não apresenta termo constante, também não há tal termo em seu respectivo modelo de dois passos à frente (equação (4.16)).

Tabela 4.2: Termos primários e termos secundários dos modelos de dois passos à frente apresentados nos Exemplos 4.3.1 e 4.3.2.

	Termos primários	Termos secundários
Exemplo 4.3.1	$y(k-1)$ $y(k-2)^2$ $u(k-1)u(k-2)$ $u(k-1)$ constante ($\theta_1\theta_5$)	$y(k-1)^2$ $u(k)u(k-1)$ $u(k)$ constante (θ_5)
Exemplo 4.3.2	$y(k-1)u(k-2)^2u(k-1)^2$ $u(k-1)u(k-3)u(k-1)^2$ $u(k-2)u(k-1)^2$ $y(k-2)u(k-1)u(k-1)^2$ $u(k-1)^2$	$u(k)u(k-2)$ $u(k-1)$ $y(k-1)u(k)$ constante (θ_5)

A fim de obter um modelo de um passo à frente a partir de seu respectivo modelo de dois passos, foi criada uma tabela (Tabela 4.3) que relaciona cada um dos termos primários de um modelo de dois passos à frente (\mathcal{M}_{2PF}), com grau de não-linearidade até cinco, com os possíveis termos de seu respectivo modelo de um passo à frente (\mathcal{M}_{1PF}), com grau de não-linearidade até três. Nessa tabela, a , b , c , d e e indicam os atrasos presentes nos regressores, e os termos em x representam tanto o sinal de entrada u quanto o sinal de saída y .

Tabela 4.3: Relação entre os termos primários de um modelo de dois passos à frente e os termos de seu respectivo modelo de um passo à frente. As letras a , b , c , d e e representam os atrasos presentes nos regressores, e os termos em x se referem tanto ao sinal de entrada u quanto ao sinal de saída y .

Situação	Termo de \mathcal{M}_{2PF}	Opção	Posíveis termos de \mathcal{M}_{1PF}
1	$x(k-a)$	1.1	$y(k-1), x(k-a)$ (para $a \neq 0$)
		1.2	$y(k-1)x(k-a-1)$, constante
2	$x(k-a)x(k-b)$	2.1	$y(k-1), x(k-a)x(k-b)$ (para $a, b \neq 0$)
		2.2	$y(k-1)x(k-a-1), x(k-b)$ (para $b \neq 0$) (podendo apresentar até 2 formas distintas de representação)
		2.3	$y(k-1)x(k-a-1)x(k-b-1)$, constante
3	$x(k-a)x(k-b)x(k-c)$	3.1	$y(k-1), x(k-a)x(k-b)x(k-c)$ (para $a, b, c \neq 0$)
		3.2	$y(k-1)x(k-a-1), x(k-b)x(k-c)$ (para $b, c \neq 0$) (podendo apresentar até 3 formas distintas de representação)
		3.3	$y(k-1)x(k-a-1)x(k-b-1), x(k-c)$ (para $c \neq 0$) (podendo apresentar até 3 formas distintas de representação)
4	$x(k-a)x(k-b)x(k-c)x(k-d)$	4.1	$y(k-1)x(k-a-1), x(k-b)x(k-c)x(k-d)$ (para $b, c, d \neq 0$) (podendo apresentar até 4 formas distintas de representação)
		4.2	$y(k-1)x(k-a-1)x(k-b-1), x(k-c)x(k-d)$ (para $c, d \neq 0$) (podendo apresentar até 6 formas distintas de representação)
5	$x(k-a)x(k-b)x(k-c)x(k-d)x(k-e)$	5.1	$y(k-1)x(k-a-1)x(k-b-1), x(k-c)x(k-d)x(k-e)$ (para $c, d, e \neq 0$) (podendo apresentar até 10 formas distintas de representação)

A Tabela 4.3 apresenta cinco situações distintas, que podem ser empregadas de acordo com o grau de não-linearidade dos termos do modelo de dois passos à frente. Para cada uma dessas situações, há um determinado número de opções de estrutura que podem ser utilizadas na formação dos possíveis termos do modelo de um passo à frente. Deve-se ressaltar que, dependendo da escolha dos atrasos presentes nos regressores (a , b , c , d , e), as estruturas podem apresentar mais de uma forma de representação. Considerando, por exemplo, o termo $y(k-2)u(k-1)u(k-3)$ de um modelo de dois passos e empregando a opção 3.3, cuja estrutura pode apresentar até três formas distintas de representação, chega-se aos seguintes termos de um passo à frente: (i) $y(k-1)y(k-3)u(k-2)$ e $u(k-3)$ (para $a = 2$, $b = 1$, $c = 3$), (ii) $y(k-1)y(k-3)u(k-4)$ e $u(k-1)$ (para $a = 2$, $b = 3$, $c = 1$) e (iii) $y(k-1)u(k-2)u(k-4)$ e $y(k-2)$ (para $a = 1$, $b = 3$, $c = 2$).

Para que as relações presentes na Tabela 4.3 possam ser empregadas na obtenção de um modelo de um passo à frente a partir de seu respectivo modelo de dois passos, será adotado o seguinte procedimento:

1. Aplicar as relações apresentadas na Tabela 4.3 a todos os termos de \mathcal{M}_{2PF} .

Como, na prática, não é possível saber quais termos de um modelo de dois passos à frente são, de fato, primários, é necessário que as relações apresentadas na Tabela 4.3 sejam aplicadas a todos os termos de \mathcal{M}_{2PF} (com exceção do termo constante, que não pode ser considerado um termo primário). Além disso, recomenda-se utilizar, na formação dos termos de um passo à frente, todas as opções possíveis de estrutura.

2. Representar, na forma de predição de dois passos à frente, todos os termos de um passo obtidos pelas relações presentes na Tabela 4.3 a fim de verificar quais novos termos de dois passos são gerados.

Para que os termos de um passo à frente possam ser considerados, a princípio, pertencentes a \mathcal{M}_{1PF} , é necessário que todos os novos termos de dois passos, gerados por tais termos de um passo à frente, estejam presentes em \mathcal{M}_{2PF} . Se pelo menos um desses

termos não pertencer ao modelo de dois passos, os termos de um passo à frente em questão deverão, então, ser descartados. Considerando, por exemplo, o termo $y(k-3)u(k-2)$ de um modelo de dois passos e empregando a opção 2.1, chega-se aos termos $y(k-1)$ e $y(k-3)u(k-2)$ de um passo à frente, que, por sua vez, implicam nos termos $y(k-1)$, $y(k-3)u(k-2)$ e $y(k-2)u(k-1)$ de dois passos. Assim, além do termo $y(k-3)u(k-2)$, outros dois termos são gerados por $y(k-1)$ e $y(k-3)u(k-2)$ na predição de dois passos à frente. Portanto, nesse caso, é necessário que os termos $y(k-1)$ e $y(k-2)u(k-1)$ estejam presentes em \mathcal{M}_{2PF} para que $y(k-1)$ e $y(k-3)u(k-2)$ sejam considerados, a princípio, pertencentes a \mathcal{M}_{1PF} . Deve-se ressaltar que, como este trabalho se limitou a lidar apenas com modelos de um passo à frente nos quais o regressor $y(k-1)$ apresenta grau de multiplicidade igual a um, os termos candidatos a \mathcal{M}_{1PF} que apresentarem tal regressor com grau de multiplicidade maior que um também devem ser eliminados.

3. Verificar quais dos termos considerados pertencentes a \mathcal{M}_{1PF} geram, na predição de dois passos à frente, os termos de \mathcal{M}_{2PF} nos quais a aplicação das relações apresentadas na Tabela 4.3 leva à obtenção de termos de um passo à frente incapazes de serem considerados pertencentes ao modelo de um passo.

Ao aplicar as relações apresentadas na Tabela 4.3, que considera apenas os termos primários de um modelo de dois passos à frente, a um termo secundário, verifica-se que nenhum dos termos candidatos a \mathcal{M}_{1PF} obtidos nesse caso pode ser considerado pertencente ao modelo de um passo devido ao fato de tais termos implicarem a geração de termos não pertencentes ao modelo de dois passos à frente. Portanto, quando isso acontecer a um determinado termo de \mathcal{M}_{2PF} , basta verificar quais dos termos de um passo à frente que, obtidos por algum outro termo de \mathcal{M}_{2PF} a partir das relações da Tabela 4.3 e considerados pertencentes a \mathcal{M}_{1PF} , implicam na sua geração.

Com base nesse procedimento, é possível obter um modelo de um passo à frente a partir de seu respectivo modelo de dois passos utilizando as relações presentes na Tabela 4.3. Isso será demonstrado nos exemplos

apresentados na próxima seção.

4.4 Obtenção de $\mathcal{M}_{1\text{PF}}$ a partir de $\mathcal{M}_{2\text{PF}}$

Os Exemplos 4.4.1, 4.4.2 e 4.4.3, apresentados a seguir, mostram como um modelo de um passo à frente pode ser obtido a partir de seu respectivo modelo de dois passos utilizando, para isso, as relações apresentadas na Tabela 4.3. A fim de auxiliar na obtenção da representação de um modelo na predição de dois passos à frente, foi implementado, na função `predicao`, que se encontra no Apêndice A.7 deste trabalho, o procedimento necessário para tal fim.

Exemplo 4.4.1

Considere o seguinte modelo de um passo à frente:

$$y(k) = \theta_1 y(k-1)u(k-1)^2 + \theta_2 u(k-2) + \theta_3 y(k-2)u(k-2) + \theta_4 u(k-2)^3 + \theta_5, \quad (4.17)$$

que, na predição de dois passos à frente, pode ser representado da seguinte forma:

$$\begin{aligned} y(k+1) = & \theta_1^2 u(k)^2 y(k-1)u(k-1)^2 + \theta_1 \theta_2 u(k)^2 u(k-2) + \\ & \theta_1 \theta_3 u(k)^2 y(k-2)u(k-2) + \theta_1 \theta_4 u(k)^2 u(k-2)^3 + \\ & \theta_1 \theta_5 u(k)^2 + \theta_2 u(k-1) + \theta_3 y(k-1)u(k-1) + \\ & \theta_4 u(k-1)^3 + \theta_5. \end{aligned} \quad (4.18)$$

A fim de se obter o modelo de um passo à frente (equação (4.17)) a partir do modelo (4.18), utilizou-se a Tabela 4.3, que relaciona os termos primários de um modelo de dois passos ($\mathcal{M}_{2\text{PF}}$) aos possíveis termos de seu respectivo modelo de um passo à frente ($\mathcal{M}_{1\text{PF}}$). A Tabela 4.4 apresenta a relação entre os termos do modelo (4.18) e os possíveis termos de seu respectivo modelo de um passo à frente.

De acordo com a Tabela 4.4, o único termo de um passo à frente obtido pelo termo $u(k)^2 y(k-1)u(k-1)^2$ do modelo (4.18) não gera nenhum outro termo no modelo de dois passos à frente a não ser o próprio $u(k)^2 y(k-1)u(k-1)^2$. Assim, pode-se afirmar que o termo $y(k-1)u(k-1)^2$ pode ser considerado, a princípio, pertencente a $\mathcal{M}_{1\text{PF}}$.

Tabela 4.4: Relação entre os termos do modelo (4.18) e os possíveis termos de seu respectivo modelo de um passo à frente.

Termo de \mathcal{M}_{2PF}	Opção	Possíveis termos de \mathcal{M}_{1PF}	Demais termos gerados em \mathcal{M}_{2PF}	Situação dos termos gerados em \mathcal{M}_{2PF}
$u(k)^2 y(k-1)u(k-1)^2$	5.1	$y(k-1)u(k-1)^2$		
$u(k)^2 u(k-2)$	3.3	$y(k-1)u(k-1)^2$	$u(k)^2 y(k-1)u(k-1)^2$	Termo pertencente ao modelo (4.18)
		$u(k-2)$	$u(k-1)$	Termo pertencente ao modelo (4.18)
$u(k)^2 y(k-2)u(k-2)$	4.2	$y(k-1)u(k-1)^2$	$u(k)^2 y(k-1)u(k-1)^2$	Termo pertencente ao modelo (4.18)
		$y(k-2)u(k-2)$	$y(k-1)u(k-1)$	Termo pertencente ao modelo (4.18)
$u(k)^2 u(k-2)^3$	5.1	$y(k-1)u(k-1)^2$	$u(k)^2 y(k-1)u(k-1)^2$	Termo pertencente ao modelo (4.18)
		$u(k-2)^3$	$u(k-1)^3$	Termo pertencente ao modelo (4.18)
$u(k)^2$	2.3	$y(k-1)u(k-1)^2$	$u(k)^2 y(k-1)u(k-1)^2$	Termo pertencente ao modelo (4.18)
		constante	constante	Termo pertencente ao modelo (4.18)
$u(k-1)$	1.1	$y(k-1)$	$y(k-1)$	Termo não pertencente ao modelo (4.18)
	1.2	$u(k-1)$	$u(k)$	Termo não pertencente ao modelo (4.18)
$y(k-1)u(k-1)$	2.1	$y(k-1)u(k-2)$	$y(k-1)u(k-1)u(k-2)$	Termo não pertencente ao modelo (4.18)
		constante	constante	Termo não pertencente ao modelo (4.18)
$y(k-1)u(k-1)$	2.2	$y(k-1)$	$y(k-1)$	Termo não pertencente ao modelo (4.18)
		$y(k-1)u(k-1)$	$u(k)$	Termo não pertencente ao modelo (4.18)
$y(k-1)u(k-1)$	2.2	$y(k-1)u(k-2)$	$y(k-1)u(k-1)u(k-2)$	Termo não pertencente ao modelo (4.18)
		$y(k-1)$	$y(k-1)$	Termo não pertencente ao modelo (4.18)
$y(k-1)u(k-1)$	2.3	$y(k-1)u(k-2)u(k-2)$	$y(k-1)^2 y(k-2)u(k-1)u(k-2)$	Termo não pertencente ao modelo (4.18)
		constante	constante	Termo pertencente ao modelo (4.18)
$u(k-1)^3$	3.1	$y(k-1)$	$y(k-1)$	Termo não pertencente ao modelo (4.18)
	3.2	$u(k-1)^3$	$u(k)^3$	Termo não pertencente ao modelo (4.18)
$u(k-1)^3$	3.3	$y(k-1)u(k-2)$	$y(k-1)u(k-1)u(k-2)$	Termo não pertencente ao modelo (4.18)
		$u(k-1)^2$	$u(k)^2$	Termo pertencente ao modelo (4.18)
$u(k-1)^3$	3.3	$y(k-1)u(k-2)$	$y(k-1)u(k-1)^2 u(k-2)^2$	Termo não pertencente ao modelo (4.18)
		$u(k-1)$	$u(k)$	Termo não pertencente ao modelo (4.18)

Além disso, verifica-se que os termos de um passo à frente obtidos pelos termos $u(k)^2u(k-2)$, $u(k)^2y(k-2)u(k-2)$, $u(k)^2u(k-2)^3$ e $u(k)^2$ do modelo (4.18) implicam a geração de termos pertencentes a $\mathcal{M}_{2\text{PF}}$. Portanto, além do termo $y(k-1)u(k-1)^2$, os termos $u(k-2)$, $y(k-2)u(k-2)$, $u(k-2)^3$ e constante também podem ser considerados, a princípio, pertencentes a $\mathcal{M}_{1\text{PF}}$.

Entretanto, observa-se que os termos de um passo à frente obtidos pelos termos $u(k-1)$, $y(k-1)u(k-1)$ e $u(k-1)^3$ do modelo (4.18) implicam a geração de termos não pertencentes a $\mathcal{M}_{2\text{PF}}$. Apesar disso, verifica-se, na Tabela 4.4, que $u(k-1)$, $y(k-1)u(k-1)$ e $u(k-1)^3$ são gerados, respectivamente, pelos termos $y(k-1)u(k-1)^2$ e $u(k-2)$, $y(k-1)u(k-1)^2$ e $y(k-2)u(k-2)$, $y(k-1)u(k-1)^2$ e $u(k-2)^3$ na predição de dois passos à frente.

Assim, de acordo com a análise da Tabela 4.4, pode-se afirmar que apenas os termos $y(k-1)u(k-1)^2$, $u(k-2)$, $y(k-2)u(k-2)$, $u(k-2)^3$ e constante podem ser considerados pertencentes a $\mathcal{M}_{1\text{PF}}$. Portanto, verifica-se que os termos de um passo à frente sugeridos pela análise da Tabela 4.4 correspondem exatamente aos termos do modelo (4.17). \square

Exemplo 4.4.2

Considere o seguinte modelo de um passo à frente:

$$y(k) = \theta_1 y(k-1) + \theta_2 u(k-2) + \theta_3 y(k-2)^2 + \theta_4 u(k-1)^2 u(k-2), \quad (4.19)$$

que, na predição de dois passos à frente, pode ser representado da seguinte forma:

$$\begin{aligned} y(k+1) = & \theta_1^2 y(k-1) + \theta_1 \theta_2 u(k-2) + \theta_1 \theta_3 y(k-2)^2 + \\ & \theta_1 \theta_4 u(k-1)^2 u(k-2) + \theta_2 u(k-1) + \theta_3 y(k-1)^2 + \\ & \theta_4 u(k)^2 u(k-1). \end{aligned} \quad (4.20)$$

A Tabela 4.3 foi novamente utilizada a fim de se obter o modelo de um passo à frente (equação (4.19)) a partir do modelo (4.20). Como ausência de termo constante em modelos de dois passos à frente implica ausência de termo constante em modelos de um passo e como o modelo (4.20) não apresenta esse tipo de termo, as opções 1.2 e 2.3 não foram empregadas. A Tabela 4.5 apresenta a relação entre os termos do modelo (4.20) e os possíveis termos de seu respectivo modelo de um passo à frente.

Tabela 4.5: Relação entre os termos do modelo (4.20) e os possíveis termos de seu respectivo modelo de um passo à frente.

Termo de \mathcal{M}_{2PF}	Opção	Possíveis termos de \mathcal{M}_{1PF}	Demais termos gerados em \mathcal{M}_{2PF}	Situação dos termos gerados em \mathcal{M}_{2PF}
$y(k-1)$	1.1	$y(k-1)$		
$u(k-2)$	1.1	$y(k-1)$	$y(k-1)$	Termo pertencente ao modelo (4.20)
		$u(k-2)$	$u(k-1)$	Termo pertencente ao modelo (4.20)
$y(k-2)^2$	2.1	$y(k-1)$	$y(k-1)$	Termo pertencente ao modelo (4.20)
		$y(k-2)^2$	$y(k-1)^2$	Termo pertencente ao modelo (4.20)
	2.2	$y(k-1)y(k-3)$	$y(k-1)y(k-2)y(k-3)$	Termo não pertencente ao modelo (4.20)
		$y(k-2)$	$y(k-1)$	Termo pertencente ao modelo (4.20)
$u(k-1)^2u(k-2)$	3.1	$y(k-1)$	$y(k-1)$	Termo pertencente ao modelo (4.20)
		$u(k-1)^2u(k-2)$	$u(k)^2u(k-1)$	Termo pertencente ao modelo (4.20)
	3.2	$y(k-1)u(k-2)$	$y(k-1)u(k-1)u(k-2)$	Termo não pertencente ao modelo (4.20)
		$u(k-1)u(k-2)$	$u(k)u(k-1)$	Termo não pertencente ao modelo (4.20)
$u(k-1)^2u(k-2)$	3.2	$y(k-1)u(k-3)$	$y(k-1)u(k-2)u(k-3)$	Termo não pertencente ao modelo (4.20)
		$u(k-1)^2$	$u(k)^2$	Termo não pertencente ao modelo (4.20)
	3.3	$y(k-1)u(k-2)^2$	$y(k-1)u(k-1)^2u(k-2)^2$	Termo não pertencente ao modelo (4.20)
		$u(k-2)$	$u(k-1)$	Termo pertencente ao modelo (4.20)
$u(k-1)$	3.3	$y(k-1)u(k-2)u(k-3)$	$y(k-1)u(k-1)u(k-2)^2u(k-3)$	Termo não pertencente ao modelo (4.20)
		$u(k-1)$	$u(k)$	Termo não pertencente ao modelo (4.20)
	1.1	$y(k-1)$	$y(k-1)$	Termo pertencente ao modelo (4.20)
		$u(k-1)$	$u(k)$	Termo não pertencente ao modelo (4.20)
$y(k-1)^2$	2.1	$y(k-1)$		$y(k-1)^n$ com $n > 1$ em \mathcal{M}_{1PF}
		$y(k-1)^2$		
	2.2	$y(k-1)y(k-2)$	$y(k-1)^2y(k-2)$	Termo não pertencente ao modelo (4.20)
		$y(k-1)$	$y(k-1)y(k-2)$	Termo não pertencente ao modelo (4.20)
$u(k)^2u(k-1)$	3.3	$y(k-1)u(k-1)^2$	$y(k-1)$	Termo pertencente ao modelo (4.20)
		$u(k-1)$	$u(k)^2y(k-1)u(k-1)^2$	Termo não pertencente ao modelo (4.20)
	3.3	$y(k-1)u(k-1)$	$u(k)$	Termo não pertencente ao modelo (4.20)
		$u(k-1)$	$u(k)$	Termo não pertencente ao modelo (4.20)

Em relação aos termos $y(k-1)$, $u(k-2)$, $y(k-2)^2$ e $u(k-1)^2u(k-2)$ do modelo (4.20), a Tabela 4.5 mostra que apenas os termos $y(k-1)$, $u(k-2)$, $y(k-2)^2$ e $u(k-1)^2u(k-2)$ de um passo à frente podem ser considerados, a princípio, pertencentes a \mathcal{M}_{1PF} .

Por outro lado, verifica-se que os termos de um passo à frente obtidos pelos demais termos do modelo (4.20) não podem ser considerados, a princípio, pertencentes a \mathcal{M}_{1PF} . Entretanto, como pode ser visto na Tabela 4.5, os termos $u(k-1)$, $y(k-1)^2$ e $u(k)^2u(k-1)$ são gerados, respectivamente, pelos termos $y(k-1)$ e $u(k-2)$, $y(k-1)$ e $y(k-2)^2$, $y(k-1)$ e $u(k-1)^2u(k-2)$ na predição de dois passos à frente.

Portanto, de acordo com a análise da Tabela 4.5, é possível afirmar que apenas os termos $y(k-1)$, $u(k-2)$, $y(k-2)^2$ e $u(k-1)^2u(k-2)$ podem ser considerados pertencentes a \mathcal{M}_{1PF} . Dessa forma, verifica-se que os termos de um passo à frente sugeridos nesse caso correspondem exatamente aos termos do modelo (4.19). \square

Exemplo 4.4.3

Considere o seguinte modelo de um passo à frente:

$$y(k) = \theta_1 y(k-1)y(k-2) + \theta_2 u(k-1)u(k-2) + \theta_3 y(k-2) + \theta_4 u(k-2)^3 + \theta_5, \quad (4.21)$$

que, na predição de dois passos à frente, pode ser representado da seguinte forma:

$$\begin{aligned} y(k+1) = & \theta_1^2 y(k-1)^2 y(k-2) + \theta_1 \theta_2 y(k-1) u(k-1) u(k-2) + \\ & \theta_1 \theta_3 y(k-1) y(k-2) + \theta_1 \theta_4 y(k-1) u(k-2)^3 + \\ & \theta_2 u(k) u(k-1) + (\theta_1 \theta_5 + \theta_3) y(k-1) + \theta_4 u(k-1)^3 + \\ & \theta_5. \end{aligned} \quad (4.22)$$

Assim como nos Exemplos 4.4.1 e 4.4.2, a Tabela 4.3 foi utilizada com o intuito de obter o modelo de um passo à frente (equação (4.21)) a partir do modelo (4.22). Diferentemente do Exemplo 4.4.2, as opções 1.2 e 2.3 foram empregadas neste caso, pois o modelo (4.22) apresenta termo constante. A relação entre os termos do modelo (4.22) e os possíveis termos de seu respectivo modelo de um passo à frente é apresentada nas Tabelas 4.6, 4.7 e 4.8.

Tabela 4.6: Relação entre os termos do modelo (4.22) e os possíveis termos de seu respectivo modelo de um passo à frente (parte 1).

Termo de \mathcal{M}_{2PF}	Opção	Possíveis termos de \mathcal{M}_{1PF}	Demais termos gerados em \mathcal{M}_{2PF}	Situação dos termos gerados em \mathcal{M}_{2PF}	
$y(k-1)^2y(k-2)$	3.1	$y(k-1)$		$y(k-1)^n$ com $n > 1$ em \mathcal{M}_{1PF}	
		$y(k-1)^2y(k-2)$			
	3.2	$y(k-1)y(k-3)$			$y(k-1)^n$ com $n > 1$ em \mathcal{M}_{1PF}
		$y(k-1)^2$			
	3.3	$y(k-1)y(k-2)^2$	$y(k-1)^3y(k-2)^2$	Termo não pertencente ao modelo (4.22)	
		$y(k-2)$	$y(k-1)$	Termo pertencente ao modelo (4.22)	
$y(k-1)y(k-2)y(k-3)$		$y(k-1)^2y(k-2)^2y(k-3)$	Termo não pertencente ao modelo (4.22)		
$y(k-1)$		$y(k-1)y(k-2)y(k-3)$	Termo não pertencente ao modelo (4.22)		
$y(k-1)u(k-1)u(k-2)$	3.1	$y(k-1)$	$y(k-1)$	Termo pertencente ao modelo (4.22)	
		$y(k-1)u(k-1)u(k-2)$	$u(k)y(k-1)u(k-1)$	Termo não pertencente ao modelo (4.22)	
			$u(k)y(k-1)u(k-1)^2u(k-2)$	Termo não pertencente ao modelo (4.22)	
			$y(k-1)^2y(k-2)$	Termo pertencente ao modelo (4.22)	
	3.2	$y(k-1)y(k-2)$	$u(k)u(k-1)$	Termo pertencente ao modelo (4.22)	
		$u(k-1)u(k-2)$			
		$y(k-1)u(k-2)$			
			$y(k-1)u(k-2)u(k-3)$	Termo não pertencente ao modelo (4.22)	
	3.3	$y(k-1)u(k-3)$	$u(k)y(k-1)u(k-3)$	Termo não pertencente ao modelo (4.22)	
		$y(k-1)u(k-1)$	$u(k)y(k-1)u(k-1)$	Termo não pertencente ao modelo (4.22)	
		$y(k-1)y(k-2)u(k-2)$	$y(k-1)^2y(k-2)u(k-1)u(k-2)$	Termo não pertencente ao modelo (4.22)	
		$u(k-2)$	$u(k-1)$	Termo não pertencente ao modelo (4.22)	
3.3	$y(k-1)y(k-2)u(k-3)$	$y(k-1)^2y(k-2)u(k-2)u(k-3)$	Termo não pertencente ao modelo (4.22)		
	$u(k-1)$	$u(k)$	Termo não pertencente ao modelo (4.22)		
	$y(k-1)u(k-2)u(k-3)$	$y(k-1)u(k-1)u(k-2)^2u(k-3)$	Termo não pertencente ao modelo (4.22)		
	$y(k-1)$	$y(k-1)u(k-2)u(k-3)$	Termo não pertencente ao modelo (4.22)		
		$y(k-1)u(k-2)u(k-3)$	Termo pertencente ao modelo (4.22)		

Tabela 4.7: Relação entre os termos do modelo (4.22) e os possíveis termos de seu respectivo modelo de um passo à frente (parte 2).

Termo de \mathcal{M}_{2PF}	Opção	Possíveis termos de \mathcal{M}_{1PF}	Demais termos gerados em \mathcal{M}_{2PF}	Situação dos termos gerados em \mathcal{M}_{2PF}
$y^{(k-1)}y^{(k-2)}$	2.1	$y^{(k-1)}$	$y^{(k-1)}$	Termo pertencente ao modelo (4.22)
		$y^{(k-1)}y^{(k-2)}$	$y^{(k-1)^2}$	Termo não pertencente ao modelo (4.22)
	2.2	$y^{(k-1)}y^{(k-2)}$	$y^{(k-1)^2}y^{(k-2)}$	Termo pertencente ao modelo (4.22)
		$y^{(k-1)}y^{(k-3)}$	$y^{(k-1)}$	Termo pertencente ao modelo (4.22)
		$y^{(k-1)}$	$y^{(k-1)}y^{(k-3)}$	Termo não pertencente ao modelo (4.22)
		constante	$y^{(k-1)^2}y^{(k-2)}y^{(k-3)}$	Termo não pertencente ao modelo (4.22)
2.3	$y^{(k-1)}y^{(k-2)}y^{(k-3)}$	constante	Termo pertencente ao modelo (4.22)	
	constante	$y^{(k-1)^2}y^{(k-2)}$	Termo pertencente ao modelo (4.22)	
$y^{(k-1)u^{(k-2)^3}}$	4.1	$y^{(k-1)}y^{(k-2)}$	$u^{(k-1)^3}$	Termo pertencente ao modelo (4.22)
		$u^{(k-2)^3}$	$y^{(k-1)u^{(k-2)}}$	Termo não pertencente ao modelo (4.22)
		$y^{(k-1)u^{(k-3)}}$	$y^{(k-1)u^{(k-1)^2}u^{(k-3)}}$	Termo não pertencente ao modelo (4.22)
	4.2	$y^{(k-1)u^{(k-2)^2}}$	$y^{(k-1)u^{(k-1)^2}}$	Termo não pertencente ao modelo (4.22)
		$y^{(k-1)u^{(k-3)^2}}$	$y^{(k-1)u^{(k-1)u^{(k-3)^2}}$	Termo não pertencente ao modelo (4.22)
		$y^{(k-1)u^{(k-2)}}$	$y^{(k-1)u^{(k-1)u^{(k-2)}}$	Termo pertencente ao modelo (4.22)

Tabela 4.8: Relação entre os termos do modelo (4.22) e os possíveis termos de seu respectivo modelo de um passo à frente (parte 3).

Termo de \mathcal{M}_{2PF}	Opção	Possíveis termos de \mathcal{M}_{1PF}	Demais termos gerados em \mathcal{M}_{2PF}	Situação dos termos gerados em \mathcal{M}_{2PF}
$u(k)u(k-1)$	2.2	$y(k-1)u(k-1)$ $u(k-1)$	$u(k)y(k-1)u(k-1)$ $u(k)$	Termo não pertencente ao modelo (4.22) Termo não pertencente ao modelo (4.22)
	2.3	$y(k-1)u(k-1)u(k-2)$ constante	$u(k)y(k-1)u(k-1)^2u(k-2)$ constante	Termo não pertencente ao modelo (4.22) Termo pertencente ao modelo (4.22)
$y(k-1)$	1.1	$y(k-1)$		
	1.2	$y(k-1)y(k-2)$ constante	$y(k-1)^2y(k-2)$ constante	Termo pertencente ao modelo (4.22) Termo pertencente ao modelo (4.22)
$u(k-1)^3$	3.1	$y(k-1)$ $u(k-1)^3$	$y(k-1)$ $u(k)^3$	Termo pertencente ao modelo (4.22) Termo não pertencente ao modelo (4.22)
	3.2	$y(k-1)u(k-2)$ $u(k-1)^2$	$y(k-1)u(k-1)u(k-2)$ $u(k)^2$	Termo pertencente ao modelo (4.22) Termo não pertencente ao modelo (4.22)
	3.3	$y(k-1)u(k-2)^2$ $u(k-1)$	$y(k-1)u(k-1)^2u(k-2)^2$ $u(k)$	Termo não pertencente ao modelo (4.22) Termo não pertencente ao modelo (4.22)

Em relação aos termos $y(k-1)^2y(k-2)$, $y(k-1)y(k-2)$ e $y(k-1)u(k-2)^3$ do modelo (4.22), verifica-se que apenas os termos $y(k-1)y(k-2)$, $y(k-2)$ e $u(k-2)^3$ de um passo à frente podem ser considerados, a princípio, pertencentes a \mathcal{M}_{1PF} .

Apesar de os termos de um passo à frente obtidos pelo termo $u(k-1)^3$ do modelo (4.22) não poderem ser considerados, a princípio, pertencentes a \mathcal{M}_{1PF} , a Tabela 4.7 mostra que $u(k-1)^3$ é gerado pelos termos $y(k-1)y(k-2)$ e $u(k-2)^3$ na predição de dois passos à frente.

Além disso, verifica-se que os termos de um passo à frente obtidos pelo termo $u(k)u(k-1)$ do modelo (4.22) também não podem ser considerados, a princípio, pertencentes a \mathcal{M}_{1PF} . Entretanto, de acordo com a Tabela 4.6, é possível perceber que $u(k)u(k-1)$ é gerado pelos termos $y(k-1)y(k-2)$ e $u(k-1)u(k-2)$ na predição de dois passos à frente. Diante disso, é importante ressaltar que, no caso do termo $y(k-1)u(k-1)u(k-2)$ do modelo (4.22), tanto os termos $y(k-1)y(k-2)$ e $u(k-1)u(k-2)$ quanto o termo $y(k-1)u(k-2)$, ambos obtidos a partir da opção 3.2, podem ser considerados, a princípio, pertencentes a \mathcal{M}_{1PF} . Porém, deve ser observado que, se o termo $y(k-1)u(k-2)$ for escolhido ao invés dos termos $y(k-1)y(k-2)$ e $u(k-1)u(k-2)$ para compor o modelo de um passo, o termo $u(k)u(k-1)$ não estará presente no modelo de dois passos à frente. Em virtude disso, o termo $y(k-1)u(k-2)$ deve ser descartado do modelo de um passo. Portanto, além dos termos $y(k-1)y(k-2)$, $y(k-2)$ e $u(k-2)^3$, o termo $u(k-1)u(k-2)$ também pode ser considerado, a princípio, pertencente a \mathcal{M}_{1PF} .

Em relação ao termo $y(k-1)$ do modelo (4.22), verifica-se que tanto o termo $y(k-1)$ quanto os termos $y(k-1)y(k-2)$ e constante de um passo à frente, obtidos, respectivamente, a partir das opções 1.1 e 1.2, podem ser considerados, a princípio, pertencentes a \mathcal{M}_{1PF} . Entretanto, deve ser ressaltado que o fato de o termo $y(k-1)$ ser escolhido ao invés dos termos $y(k-1)y(k-2)$ e constante para compor o modelo de um passo implica que todos os demais termos considerados pertencentes a \mathcal{M}_{1PF} de acordo com a análise das Tabelas 4.6, 4.7 e 4.8 ($y(k-1)y(k-2)$, $y(k-2)$, $u(k-2)^3$ e $u(k-1)u(k-2)$) estejam presentes no modelo (4.22) em virtude do termo $y(k-1)$, que se transforma em $y(k)$ na predição de dois passos à frente, o que não acontece. Em virtude disso, o termo $y(k-1)$ deve ser descartado do modelo de um passo. Portanto, os termos $y(k-1)y(k-2)$ e constante podem ser considerados, a princípio, pertencentes a \mathcal{M}_{1PF} . Além disso,

pode-se observar, de acordo com a Tabela 4.7, que o termo $y(k-1)$ é gerado pelos termos $y(k-1)y(k-2)$ e $y(k-2)$ na predição de dois passos à frente. Portanto, verifica-se que o termo $y(k-1)$ do modelo (4.22) corresponde à soma de um termo primário (devido aos termos $y(k-1)y(k-2)$ e constante) com um termo secundário (devido ao termo $y(k-2)$).

Assim, após a análise das Tabelas 4.6, 4.7 e 4.8, pode-se afirmar que apenas os termos $y(k-1)y(k-2)$, $y(k-2)$, $u(k-2)^3$, $u(k-1)u(k-2)$ e constante podem ser considerados pertencentes a \mathcal{M}_{1PF} . Dessa forma, verifica-se que os termos de um passo à frente sugeridos nesse caso correspondem exatamente aos termos do modelo (4.21). \square

4.5 Metodologia Preliminar

Nesta seção, é apresentada a metodologia adotada neste trabalho. Na Figura 4.1, encontra-se um fluxograma com os passos que devem ser seguidos ao empregar tal metodologia. A descrição detalhada de cada um desses passos encontra-se nos itens apresentados a seguir.

1. Dividir os dados de identificação em janelas.

O uso dessa técnica, conhecida como janelas deslizantes, é comum em diversos problemas de controle (Mitra, 1992), processamento de sinais (Pierce e Plemmons, 1992) e seleção de estrutura de modelos a partir de estimação *on-line* (Luo e Billings, 1995; Ruano et al., 2006).

2. Definir o número de termos, os máximos atrasos nos sinais de entrada n_u e saída n_y e o grau de não-linearidade ℓ a serem utilizados durante o processo de identificação. Em seguida, utilizando o critério ERR_1 , deve-se estimar um modelo para cada uma das janelas de dados. Assim, o número de modelos obtidos será igual ao número de janelas formadas durante a divisão dos dados de identificação.

Sabe-se que, para diferentes realizações de dados correspondentes a um mesmo sistema, nem sempre obtém-se um mesmo modelo em todas as realizações, ou seja, os termos escolhidos por um determinado

critério de seleção de estrutura para compor um modelo nem sempre são os mesmos em todos os modelos, mesmo quando os máximos atrasos nos sinais de entrada e saída e o grau de não-linearidade são mantidos constantes. Assim, para evitar que uma única estimação fosse priorizada, optou-se por trabalhar com um modelo moda obtido a partir dos modelos estimados em cada uma das janelas de dados.

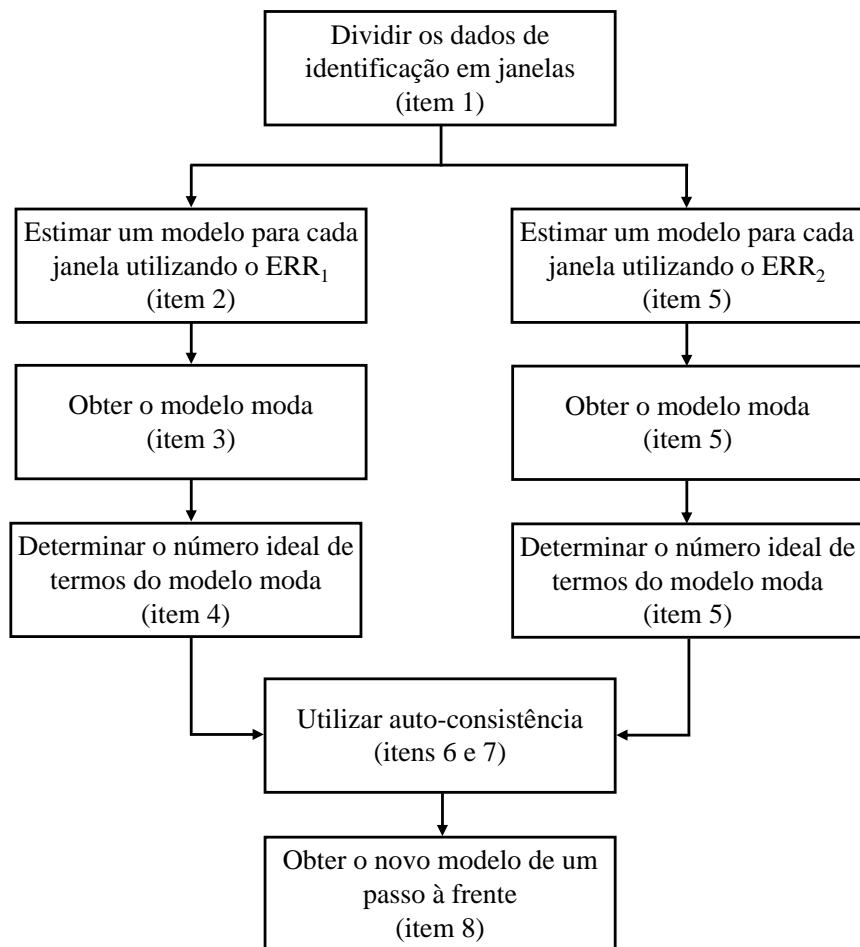


Figura 4.1: Fluxograma que representa os passos utilizados na metodologia adotada neste trabalho.

3. Obter o modelo moda.

Primeiramente, é preciso mencionar que, neste trabalho, um modelo moda consiste num modelo formado pelos termos escolhidos com maior frequência pelo critério de seleção. Para isso, tal modelo é obtido utilizando-se o seguinte procedimento: primeiramente, verifica-se o primeiro termo de cada um dos modelos estimados nas janelas de dados e, em seguida, selecionam-se apenas os modelos que apresentam o termo encontrado no maior número de modelos. Considere, por exemplo, os seguintes modelos:

$$\begin{aligned}
 M_1 &= [y(k-1), y(k-2), u(k-1), y(k-3)], \\
 M_2 &= [y(k-1), y(k-2), u(k-1), y(k-3)], \\
 M_3 &= [y(k-1), y(k-2), u(k-1), y(k-3)], \\
 M_4 &= [y(k-1), y(k-2), u(k-1), u(k-2)], \\
 M_5 &= [y(k-1), y(k-2), u(k-1), u(k-2)], \\
 M_6 &= [y(k-1), u(k-1), y(k-2), y(k-4)], \\
 M_7 &= [y(k-2), y(k-1), u(k-1), u(k-3)],
 \end{aligned}$$

Analisando-se o primeiro termo de cada um desses modelos, verifica-se que o termo $y(k-1)$ se encontra em seis dos sete modelos. Portanto, o modelo M_7 , cujo primeiro termo é $y(k-2)$, deve, então, ser descartado. Em seguida, deve-se verificar o segundo termo de cada um dos modelos pré-selecionados (no caso do exemplo, referem-se aos modelos M_1 , M_2 , M_3 , M_4 , M_5 e M_6) para que sejam selecionados apenas os modelos que apresentam o termo mais freqüente. Como cinco dos seis modelos apresentam o termo $y(k-2)$, o modelo M_6 , cujo segundo termo é $u(k-1)$, também deve ser eliminado. Esse processo continua até que o último termo seja analisado. Os modelos que chegam até o fim desse processo apresentam exatamente a mesma estrutura, que é, então, usada para compor o modelo moda. Assim, o modelo moda obtido neste exemplo deve ser formado pelos termos $y(k-1)$, $y(k-2)$, $u(k-1)$ e $y(k-3)$, que correspondem à estrutura dos modelos M_1 , M_2 e M_3 . O procedimento necessário para se chegar ao modelo moda foi implementado na função `modelo_moda`, que se encontra no Apêndice A.3 deste trabalho. Deve-se observar

que, se, por acaso, o último termo do modelo M_3 fosse, por exemplo, $u(k-4)$ em vez de $y(k-3)$, haveria, neste caso, dois modelos moda, sendo que um deles apresentaria a estrutura dos modelos M_1 e M_2 e o outro, a estrutura dos modelos M_4 e M_5 . Portanto, apesar de todos os modelos moda obtidos neste trabalho apresentarem uma única estrutura, a função `modelo_moda` pode, eventualmente, retornar dois ou mais modelos moda.

4. Determinar o número ideal de termos do modelo moda.

Para isso, serão empregados o critério de informação de Akaike (AIC) e a estatística APRESS, apresentados na Seção 3.2. Deve-se ressaltar que, ao aplicar esses métodos a um modelo moda, serão utilizadas somente as amostras responsáveis pela geração de tal modelo, ou seja, apenas aquelas usadas na obtenção dos modelos que apresentam a mesma estrutura do modelo moda.

5. Retornar ao item 2 utilizando o critério ERR_2 ao invés do ERR_1 e, em seguida, prosseguir com os itens 3 e 4, chegando-se, assim, a dois modelos moda: um obtido pelo critério ERR_1 e outro obtido pelo ERR_2 .

É importante ressaltar que o modelo moda obtido pelo ERR_2 corresponde a um modelo de dois passos à frente.

6. Determinar, a partir das relações apresentadas na Tabela 4.3, o modelo de um passo à frente correspondente ao modelo moda obtido pelo critério ERR_2 a fim de que possa ser comparado com o modelo moda obtido pelo critério ERR_1 .

Dessa forma, a auto-consistência entre os critérios ERR_1 e ERR_2 é utilizada a partir da comparação entre modelos de um passo à frente.

7. Identificar os termos comuns ao modelo moda obtido pelo critério ERR_1 e ao modelo de um passo à frente correspondente ao modelo moda obtido pelo ERR_2 .

8. Utilizar, na formação do novo modelo de um passo à frente, apenas os termos comuns ao modelo moda obtido pelo critério ERR_1 e ao modelo de um passo à frente correspondente ao modelo moda obtido pelo ERR_2 .

4.6 Considerações Finais

Neste capítulo, foi apresentado um método baseado na predição de dois passos à frente que tem por finalidade auxiliar na etapa de seleção de estrutura de modelos NARX polinomiais. O desempenho e a robustez desse método serão verificados no próximo capítulo.

Estudo de Casos Simulados

“O pessimista se queixa do vento, o otimista espera que ele mude e o realista ajusta as velas.”

William George Ward

5.1 Introdução

Este capítulo apresenta um estudo de casos em que se pretende verificar o desempenho e a robustez da técnica de predição de dois passos à frente na seleção de estrutura de modelos NARX polinomiais no caso em que o sinal de entrada é relativamente lento utilizando-se, para isso, dados simulados. No Capítulo 4, foi visto que a forma como um modelo se expande na predição de dois passos à frente depende de como o regressor $y(k-1)$ aparece no modelo de um passo. Em virtude disso, optou-se por trabalhar com sistemas nos quais tal regressor se encontra de três formas distintas: em um termo linear, multiplicado por um regressor de entrada e multiplicado por um outro regressor de saída.

5.2 Estudo de Casos

Três casos são analisados, neste capítulo, com o propósito de avaliar o desempenho e a utilidade do critério ERR_2 como uma ferramenta auxiliar na seleção de estrutura de modelos NARX polinomiais. O sistema empregado no primeiro caso (Seção 5.2.1) foi proposto em Piroddi e Spinelli (2003a), onde foi utilizada, no processo de identificação, uma massa de

dados contendo 500 amostras. Conforme discutido na Seção 4.5, os dados de identificação empregados neste trabalho devem ser divididos em janelas a fim de se trabalhar com um modelo moda obtido a partir dos modelos estimados em cada uma dessas janelas. Dessa forma, optou-se por utilizar 500 janelas contendo 500 amostras cada, resultando, assim, numa massa de dados composta por 999 amostras (Figura 5.1). Portanto, são estimados, nesse caso, 500 modelos que, posteriormente, são empregados na determinação do modelo moda.

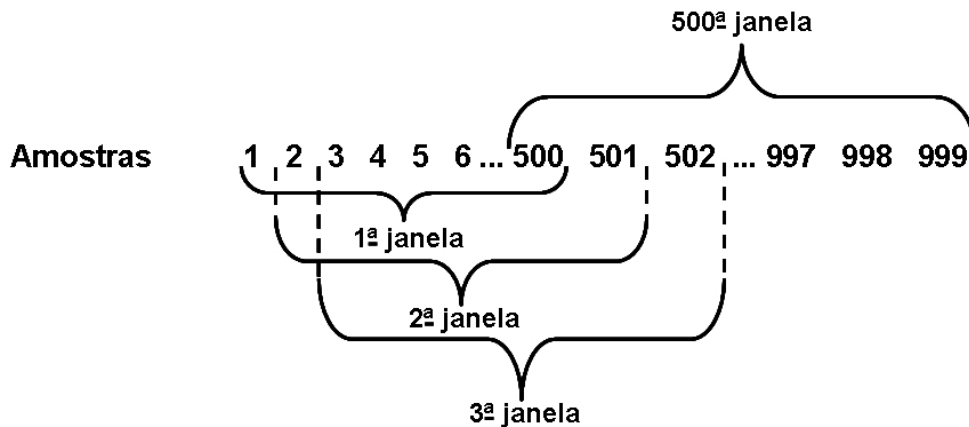


Figura 5.1: Representação de uma massa de dados dividida em 500 janelas contendo 500 amostras cada.

O procedimento descrito acima também foi adotado na identificação dos sistemas empregados nos outros dois casos (Seções 5.2.2 e 5.2.3), que, ao contrário do sistema considerado no primeiro caso, foram criados de forma arbitrária. Como a estrutura dos sistemas utilizados nesses três casos é conhecida, o desempenho dos critérios ERR_1 e ERR_2 pode ser facilmente avaliado comparando-se a estrutura dos modelos utilizados na geração dos dados com a dos modelos obtidos por tais critérios.

5.2.1 Primeiro Caso

Considere o seguinte sistema \mathcal{S}_1 representado pela equação (5.1) (Piroddi e Spinelli, 2003a):

$$\begin{aligned} w(k) &= 0,5w(k-1) + 0,8u(k-2) + u(k-1)^2 - 0,05w(k-2)^2 + 0,5 \\ y(k) &= w(k) + \frac{1}{1-0,5q^{-1}}\nu(k). \end{aligned} \quad (5.1)$$

Seguindo o experimento realizado em Piroddi e Spinelli (2003a), o sinal de entrada $u(k)$, cuja função de autocorrelação¹ (FAC) se encontra na Figura 5.3, foi gerado a partir de um sinal tipo ruído branco com distribuição Gaussiana, média nula e variância unitária. O ruído $\nu(k)$ foi gerado da mesma forma, porém, com variância de 0,05. A Figura 5.2 apresenta os sinais de entrada e saída do sistema \mathcal{S}_1 , cuja relação sinal/ruído é de 38,95 dB.

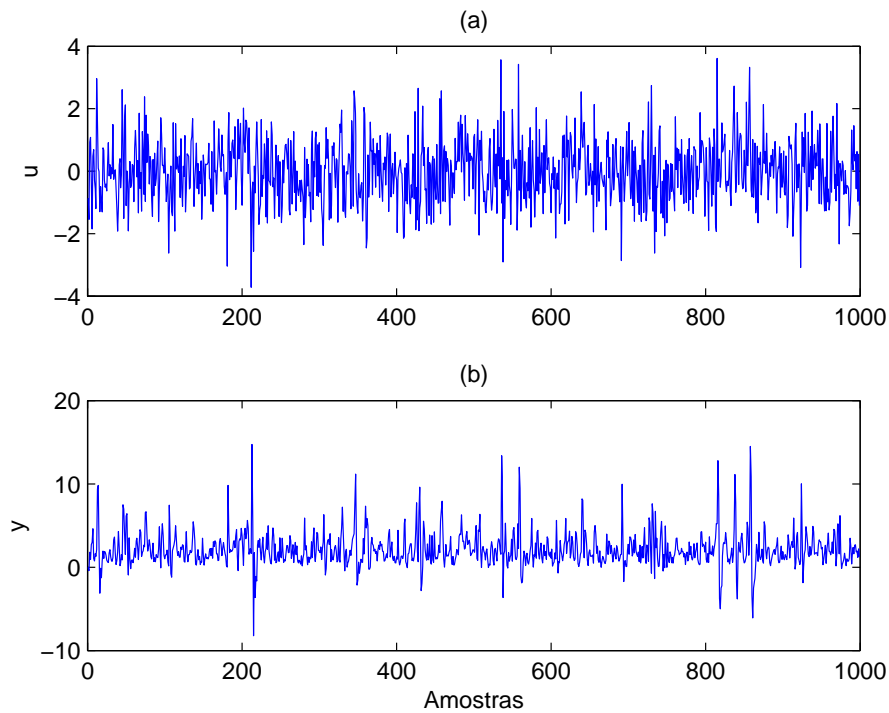


Figura 5.2: Dados do sistema \mathcal{S}_1 para uma entrada aleatória.

Sinais de (a) entrada e (b) saída (esses dados encontram-se disponíveis no arquivo `dados_myex1_ent_aleat_cr`).

¹As funções de correlação apresentadas neste trabalho foram calculadas com um nível de confiança de 95%. Como tais funções encontram-se normalizadas (entre -1 e 1), seus limites de confiança variam de $1,96/\sqrt{N}$ a $-1,96/\sqrt{N}$, sendo que N representa o número de amostras do sinal.

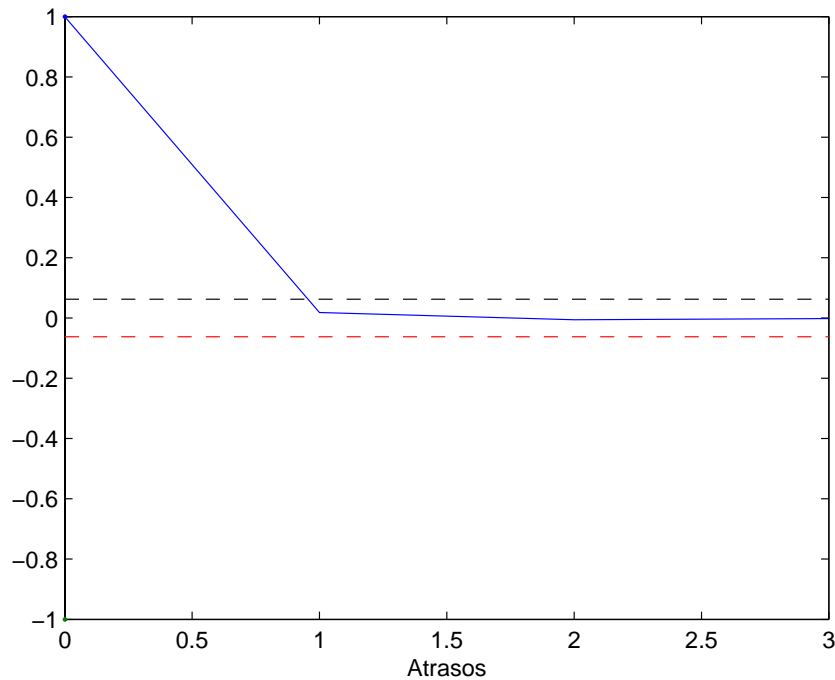


Figura 5.3: Resultado gráfico da aplicação da função FAC ao sinal de entrada apresentado na Figura 5.2.

O sinal de entrada apresenta FAC semelhante à de um ruído branco, que é impulsiva.

Assim como em Piroddi e Spinelli (2003a), na identificação do sistema \mathcal{S}_1 a partir da técnica de predição de um passo à frente, foram utilizados modelos NARX com grau de não-linearidade e máximos atrasos nos sinais de entrada e saída iguais a 2, totalizando 15 termos candidatos.

A aplicação do critério de informação de Akaike e da estatística APRESS ao modelo moda obtido pelo critério ERR_1 pode ser visualizada, respectivamente, nas Figuras 5.4 e 5.5.

A Figura 5.4 mostra claramente que, segundo o critério de informação de Akaike, o modelo deve ser formado por cinco termos. Além disso, de acordo com a Figura 5.5, a estatística APRESS também sugere que o número ideal de termos desse modelo seja cinco. Como o sistema \mathcal{S}_1 é composto por cinco termos, pode-se afirmar que, neste caso, os dois métodos estimaram corretamente o número de termos do modelo.

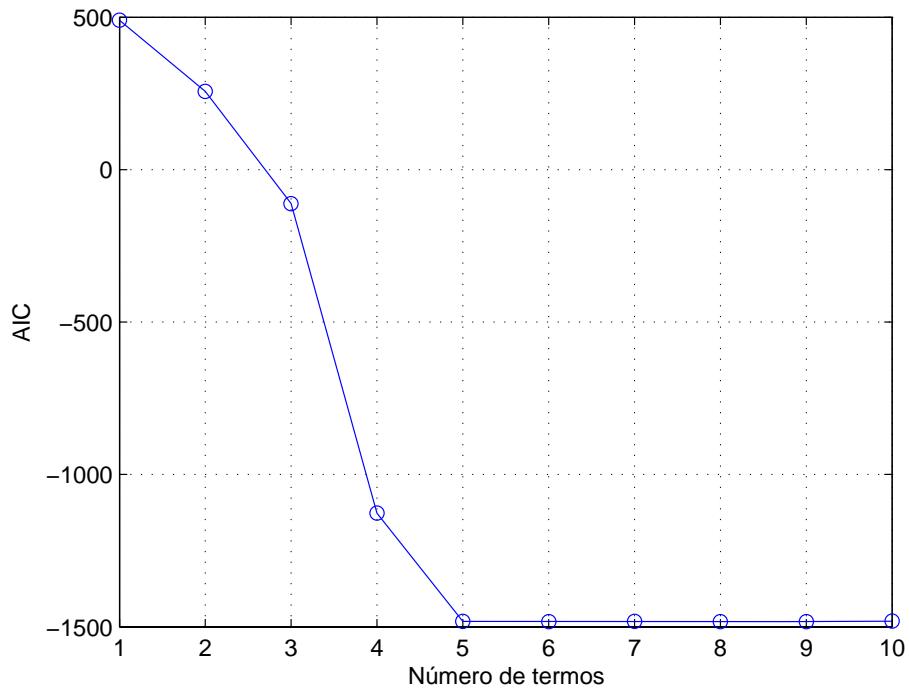


Figura 5.4: Resultado gráfico da aplicação do critério de informação de Akaike ao modelo moda obtido pelo critério ERR_1 a partir dos dados apresentados na Figura 5.2.

A Tabela 5.1 apresenta o modelo moda formado apenas pelo número de termos estimado pelo AIC e pela estatística APRESS. Ressalta-se que o número de ocorrências de um determinado termo refere-se ao número de modelos em que tal termo foi selecionado. É importante mencionar que, neste trabalho, os termos corretos são apresentados em **negrito**, e os espúrios, em *itálico*. Como pode ser observado na Tabela 5.1, os termos que compõem o modelo moda correspondem exatamente aos cinco termos presentes no sistema \mathcal{S}_1 e, por isso, encontram-se em **negrito**.

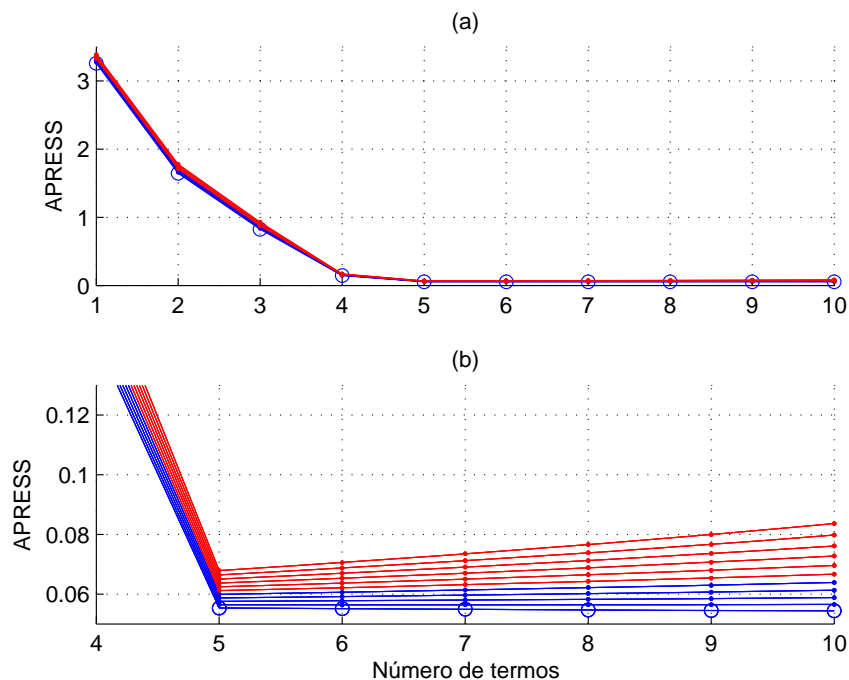


Figura 5.5: Resultado gráfico da aplicação da estatística APRESS ao modelo moda obtido pelo critério ERR_1 a partir dos dados apresentados na Figura 5.2.

Gráfico (a) em seu tamanho real e (b) em escala ampliada. As linhas, de baixo para cima, correspondem a $\alpha = 0, 1, \dots, 10$. A linha inferior (com círculos), correspondente a $\alpha = 0$, indica o MSE.

Tabela 5.1: Modelo moda obtido pelo critério ERR_1 a partir dos dados apresentados na Figura 5.2.

Termo	Parâmetro real	Parâmetro estimado	ERR_1	Número de ocorrências ²
$u(k-1)^2$	1	1,0073	0,6417	500
$y(k-1)$	0,5	0,4985	0,1774	500
$y(k-2)^2$	-0,05	-0,0484	0,0902	355
$u(k-2)$	0,8	0,8121	0,0747	355
constante	0,5	0,4829	0,0099	355

²Ressalta-se que, em todos os casos apresentados neste capítulo, o máximo valor do número de ocorrências corresponde a 500.

Na identificação do sistema \mathcal{S}_1 a partir da técnica de predição de dois passos à frente, foram utilizados modelos NARX com grau de não-linearidade e máximos atrasos nos sinais de entrada e saída iguais a 2, assim como na predição de um passo à frente. Deve-se observar que o sistema \mathcal{S}_1 apresenta o termo $u(k)^2$ quando representado na forma de predição de dois passos à frente (equação (5.2)). Conforme discutido na Seção 4.2, para que tal termo pudesse ser selecionado, todos os termos de u com atraso nulo foram incluídos no conjunto de termos candidatos. Em virtude disso, neste caso, foram utilizados 6 termos a mais ($u(k)$, $u(k)y(k-1)$, $u(k)y(k-2)$, $u(k)^2$, $u(k)u(k-1)$, $u(k)u(k-2)$) que na predição de um passo à frente, totalizando, assim, 21 termos candidatos em vez de 15.

$$y(k+1) = 0,25y(k-1) + 0,4u(k-2) + 0,5u(k-1)^2 - 0,025y(k-2)^2 + 0,75 + 0,8u(k-1) + u(k)^2 - 0,05y(k-1)^2. \quad (5.2)$$

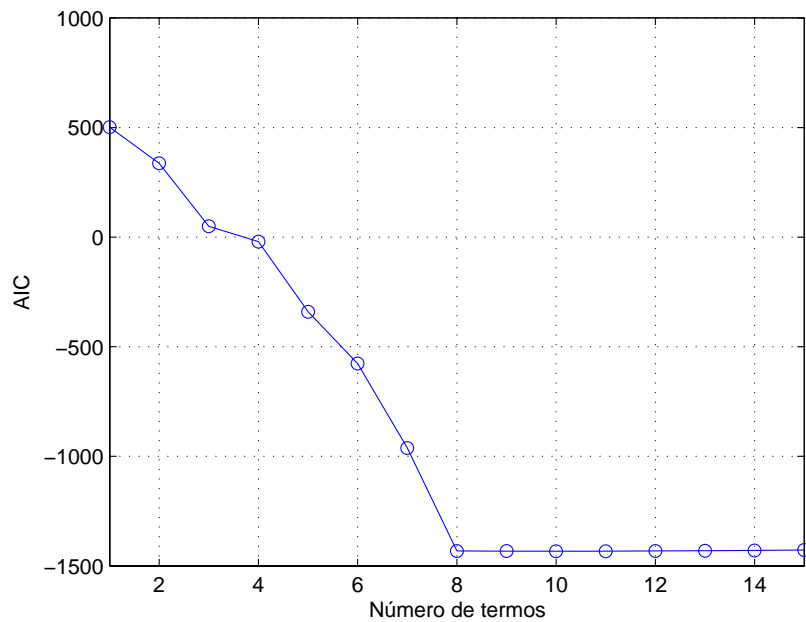


Figura 5.6: Resultado gráfico da aplicação do critério de informação de Akaike ao modelo moda obtido pelo critério ERR_2 a partir dos dados apresentados na Figura 5.2.

As Figuras 5.6 e 5.7 mostram, respectivamente, o resultado gráfico da aplicação do critério de informação de Akaike e da estatística APRESS ao modelo moda obtido pelo critério ERR_2 . Analisando-se essas figuras, percebe-se claramente que tanto o AIC quanto a estatística APRESS sugerem que o modelo seja formado por oito termos.

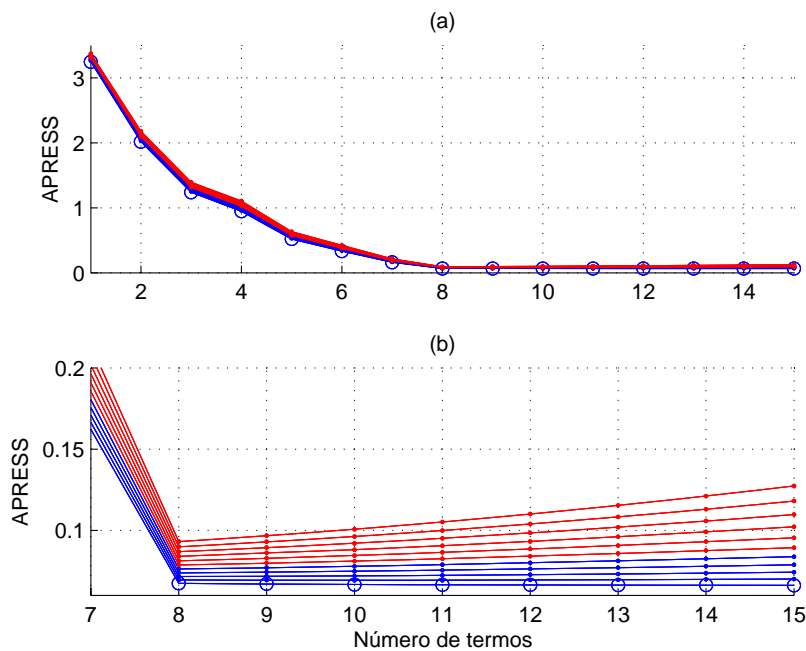


Figura 5.7: Resultado gráfico da aplicação da estatística APRESS ao modelo moda obtido pelo critério ERR_2 a partir dos dados apresentados na Figura 5.2.

Gráfico (a) em seu tamanho real e (b) em escala ampliada. As linhas, de baixo para cima, correspondem a $\alpha = 0, 1, \dots, 10$. A linha inferior (com círculos), correspondente a $\alpha = 0$, indica o MSE.

Como a equação (5.2) é composta por oito termos, pode-se afirmar que, novamente, o número de termos do modelo foi corretamente estimado pelos dois métodos. A Tabela 5.2 apresenta o modelo moda formado apenas pelo número de termos estimado pelo AIC e pela estatística APRESS. Como os termos que compõem o modelo moda correspondem exatamente aos oito termos presentes na equação (5.2), é possível afirmar que, neste caso, tanto o critério ERR_1 quanto o ERR_2 mostraram-se eficazes ao identificar o sistema \mathcal{S}_1 excitado por uma entrada aleatória.

Tabela 5.2: Modelo moda obtido pelo critério ERR_2 a partir dos dados apresentados na Figura 5.2.

Termo	Parâmetro real	Parâmetro estimado	ERR_2	Número de ocorrências
$u(k)^2$	1	1,0032	0,6434	500
$u(k-1)^2$	0,5	0,5015	0,1349	460
$u(k-1)$	0,8	0,8090	0,0854	460
constante	0,75	0,7265	0,0323	316
$y(k-1)^2$	-0,05	-0,0487	0,0468	171
$u(k-2)$	0,4	0,4162	0,0205	171
$y(k-2)^2$	-0,025	-0,0236	0,0188	171
$y(k-1)$	0,25	0,2503	0,0104	171

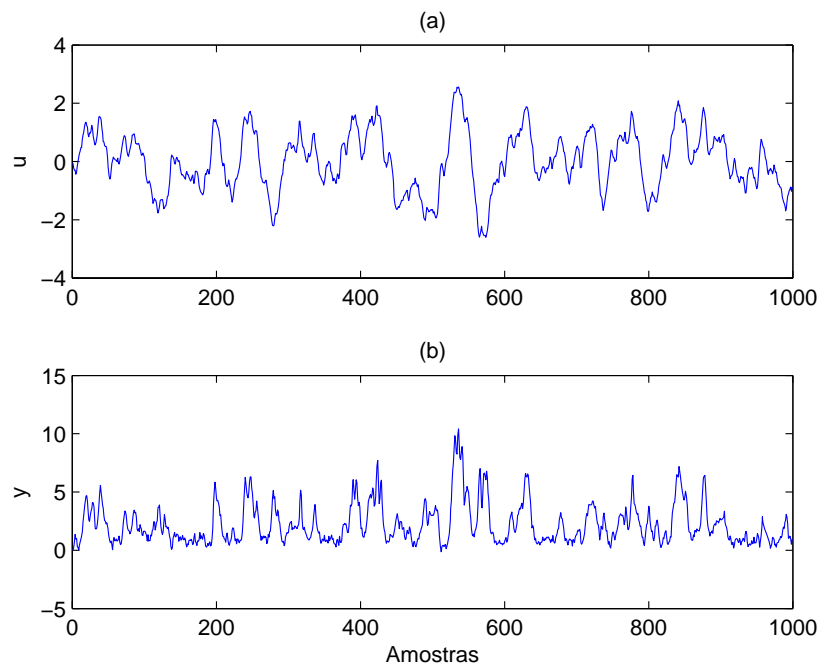


Figura 5.8: Dados do sistema \mathcal{S}_1 para uma entrada AR.

Sinais de (a) entrada e (b) saída (esses dados encontram-se disponíveis no arquivo `dados_myex1_ent_AR_cr`).

Entretanto, se este experimento for repetido utilizando uma entrada de baixa frequência em vez de uma entrada aleatória, os resultados mudam drasticamente. Para isso, assim como em Piroddi e Spinelli (2003a), foi gerado um sinal de entrada auto-regressivo de ordem dois AR(2) com média nula, variância unitária e dois pólos reais na faixa entre 0,75 e 0,90. O resultado gráfico da aplicação da função de autocorrelação (FAC) a esse novo sinal de entrada se encontra na Figura 5.9. O ruído usado neste experimento foi gerado conforme descrito no experimento anterior. A Figura 5.8 apresenta os novos sinais de entrada e saída do sistema \mathcal{S}_1 , cuja relação sinal/ruído é de 35,66 dB.

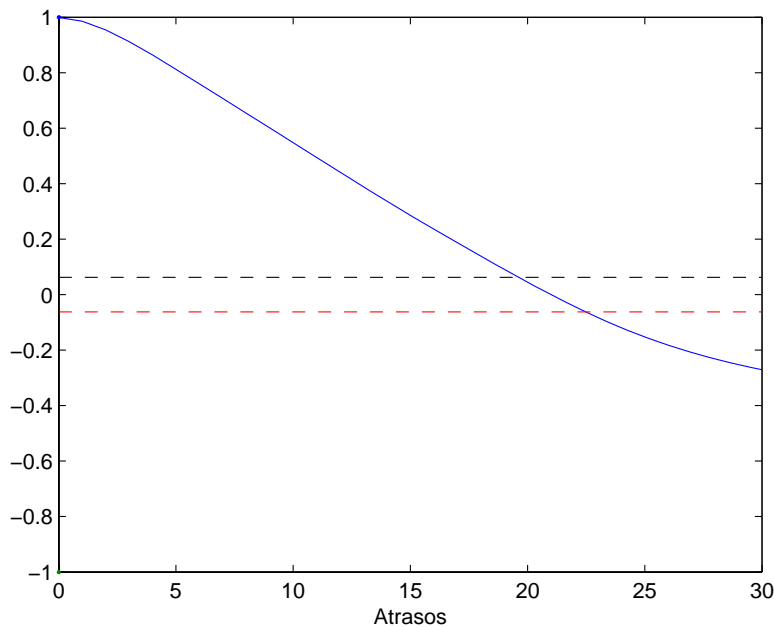


Figura 5.9: Resultado gráfico da aplicação da função FAC ao sinal de entrada apresentado na Figura 5.8.

O sinal de entrada não apresenta FAC semelhante à de um ruído branco, que é impulsiva.

Assim como no primeiro experimento, na identificação do sistema \mathcal{S}_1 a partir da técnica de predição de um passo à frente, foram utilizados modelos NARX com grau de não-linearidade e máximos atrasos nos sinais de entrada e saída iguais a 2, totalizando 15 termos candidatos.

Analisando-se as Figuras 5.10 e 5.11, que apresentam, respectivamente, o resultado gráfico da aplicação do critério de informação de Akaike e da estatística APRESS ao modelo moda obtido pelo critério ERR_1 , verifica-se que ambos os métodos sugerem que o modelo seja formado por oito termos. Entretanto, como o sistema \mathcal{S}_1 é composto por cinco termos, uma avaliação precipitada levaria a julgar erroneamente o desempenho de tais métodos.

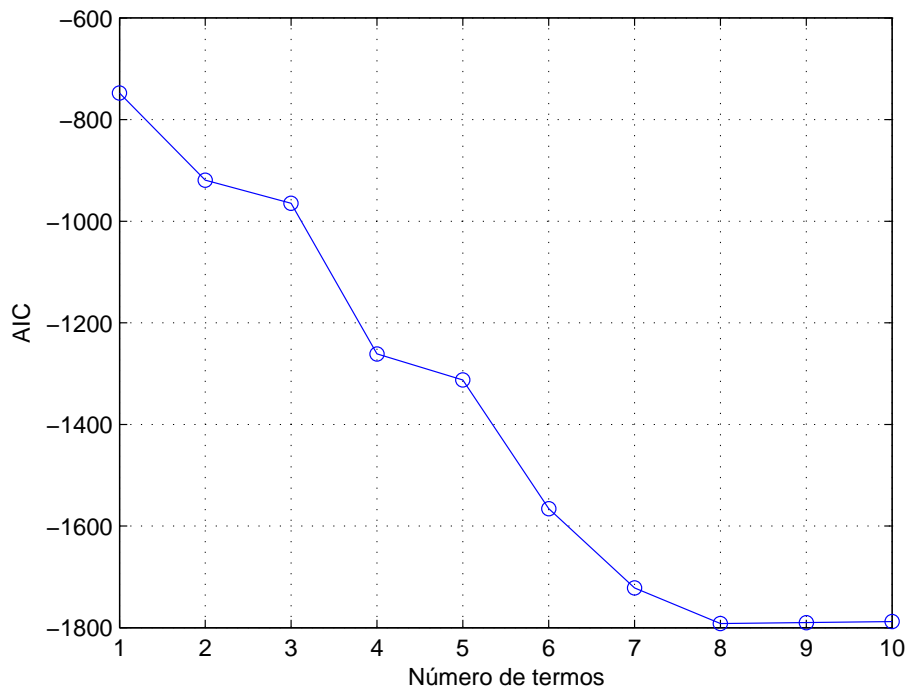


Figura 5.10: Resultado gráfico da aplicação do critério de informação de Akaike ao modelo moda obtido pelo critério ERR_1 a partir dos dados apresentados na Figura 5.8.

A Tabela 5.3 apresenta o modelo moda formado apenas pelo número de termos estimado pelo AIC e pela estatística APRESS. Como pode ser observado, todos os termos do sistema \mathcal{S}_1 encontram-se presentes no modelo moda. Além disso, o modelo também apresenta três termos espúrios. É interessante observar que os termos espúrios pertencem a agrupamentos não espúrios. Entretanto, nota-se que foram necessários exatamente oito termos para que o critério ERR_1 selecionasse todos os termos corretos,

apresentando, assim, 62,50% de acerto na escolha dos regressores. Portanto, pode-se afirmar que, neste caso, os métodos usados na estimação do número de termos do modelo mostraram ser úteis no processo de identificação diante do erro cometido pelo ERR_1 ao selecionar termos espúrios na tentativa de identificar o sistema \mathcal{S}_1 excitado por uma entrada AR.

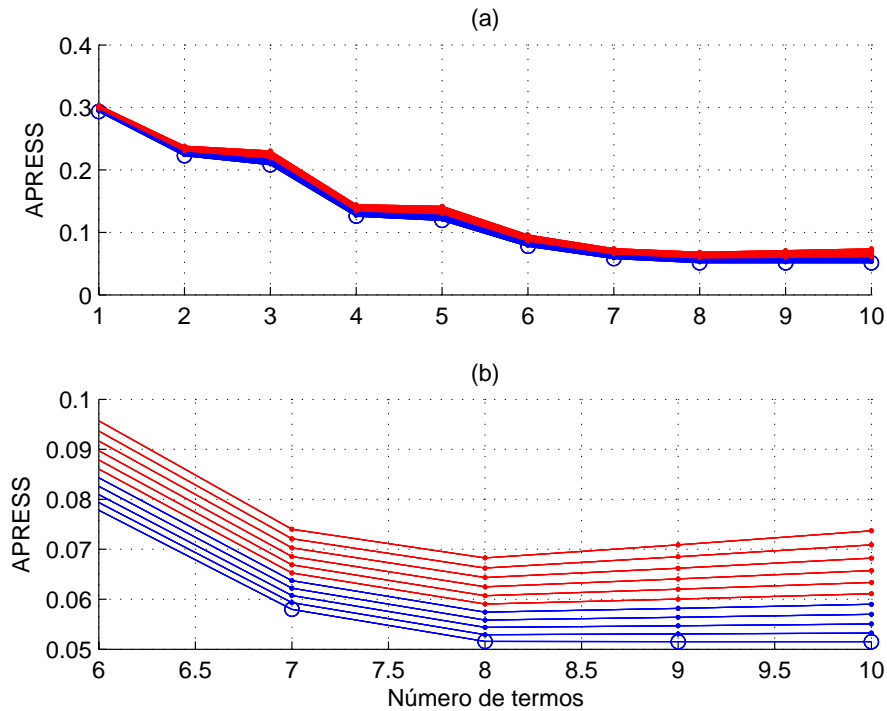


Figura 5.11: Resultado gráfico da aplicação da estatística APRESS ao modelo moda obtido pelo critério ERR_1 a partir dos dados apresentados na Figura 5.8.

Gráfico (a) em seu tamanho real e (b) em escala ampliada. As linhas, de baixo para cima, correspondem a $\alpha = 0, 1, \dots, 10$. A linha inferior (com círculos), correspondente a $\alpha = 0$, indica o MSE.

Na identificação do sistema \mathcal{S}_1 a partir da técnica de predição de dois passos à frente, foram novamente utilizados modelos NARX empregando-se, para isso, os mesmos 21 termos candidatos usados na predição de dois passos realizada no primeiro experimento.

Tabela 5.3: Modelo moda obtido pelo critério ERR_1 a partir dos dados apresentados na Figura 5.8.

Termo	Parâmetro real	Parâmetro estimado	ERR_1	Número de ocorrências
$\mathbf{y}(\mathbf{k} - 1)$	0,5	0,5523	0,9690	500
$y(k - 2)$	0	-0,0313	0,0075	472
$\mathbf{u}(\mathbf{k} - 1)^2$	1	1,0180	0,0015	472
$u(k - 2)^2$	0	-0,1132	0,0087	305
$u(k - 1)$	0	0,1336	0,0007	299
$\mathbf{y}(\mathbf{k} - 2)^2$	-0,05	-0,0431	0,0044	299
constante	0,5	0,4827	0,0021	299
$\mathbf{u}(\mathbf{k} - 2)$	0,8	0,6031	0,0007	299

As Figuras 5.12 e 5.13 mostram, respectivamente, o resultado gráfico da aplicação do critério de informação de Akaike e da estatística APRESS ao modelo moda obtido pelo critério ERR_2 . Apesar de o AIC sugerir um número de termos entre oito e dez, a Figura 5.12b mostra que o menor valor de AIC corresponde a dez termos. Entretanto, a estatística APRESS sugere que o número ideal de termos do modelo seja nove.

A Tabela 5.4 apresenta o modelo moda formado apenas pelo número de termos correspondente ao menor valor de AIC. Como pode ser novamente observado, todos os termos da equação (5.2) encontram-se presentes no modelo moda. Além disso, o modelo também apresenta dois termos espúrios. É interessante observar que, assim como na predição de um passo à frente, os termos espúrios pertencem a agrupamentos não espúrios. Entretanto, nota-se que foram necessários exatamente dez termos para que o critério ERR_2 selecionasse todos os termos corretos, apresentando, assim, 80% de acerto na escolha dos regressores. Portanto, verifica-se, neste caso, um aumento da robustez do critério ERR_2 em relação ao ERR_1 .

Entretanto, se fossem selecionados apenas nove termos do modelo moda conforme sugerido pela estatística APRESS, o termo $u(k - 2)$, que corresponde a um termo correto, seria eliminado. Portanto, neste caso, o critério de informação de Akaike se mostrou mais eficiente na estimação do número ideal de termos do modelo moda. Isso sugere que até mesmo as ferramentas mais recentes nem sempre conseguem superar as “clássicas”.

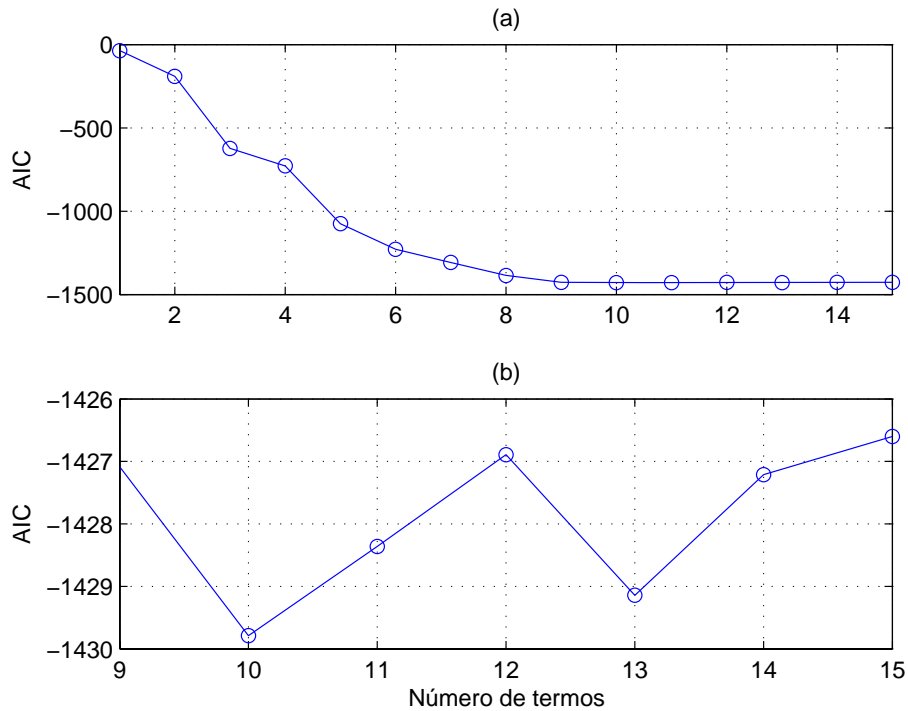


Figura 5.12: Resultado gráfico da aplicação do critério de informação de Akaike ao modelo moda obtido pelo critério ERR_2 a partir dos dados apresentados na Figura 5.8.

Gráfico (a) em seu tamanho real e (b) em escala ampliada.

Como o critério ERR_2 se mostrou mais robusto que o ERR_1 , o modelo apresentado na Tabela 5.4 foi usado na tentativa de se chegar a um modelo de um passo à frente livre de termos espúrios. Para isso, utilizou-se a Tabela 4.3 a fim de se obter tal modelo. A relação entre os termos do modelo de dois passos (Tabela 5.4) e os possíveis termos de seu respectivo modelo de um passo à frente é apresentada nas Tabelas 5.5 e 5.6, a partir das quais chega-se aos seguintes termos de um passo à frente: $y(k-1)$, $u(k-2)^2$, $u(k-1)$, $y(k-2)^2$, $u(k-1)^2$ e $u(k-2)$. Como termo constante em modelos de dois passos à frente implica termo constante em modelos de um passo e como o modelo apresentado na Tabela 5.4 apresenta esse tipo de termo, a constante deve, então, ser acrescentada aos termos de um passo obtidos pela análise das Tabelas 5.5 e 5.6. Assim, chega-se à seguinte estrutura de um passo à frente:

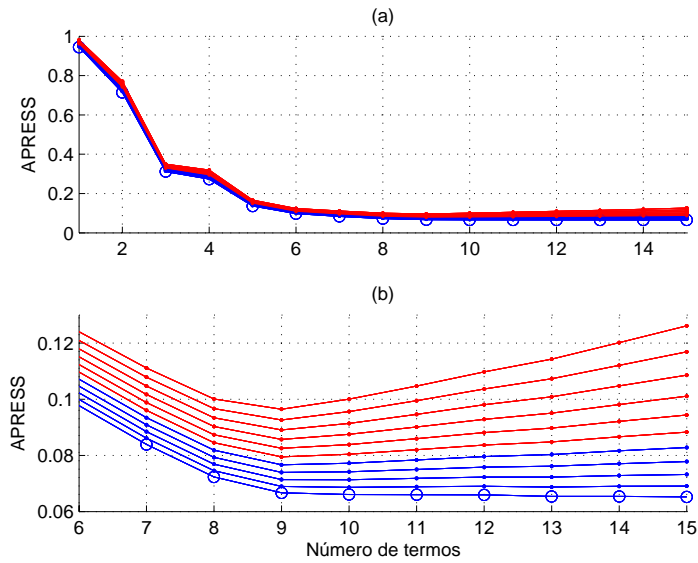


Figura 5.13: Resultado gráfico da aplicação da estatística APRESS ao modelo moda obtido pelo critério ERR_2 a partir dos dados apresentados na Figura 5.8.

Gráfico (a) em seu tamanho real e (b) em escala ampliada. As linhas, de baixo para cima, correspondem a $\alpha = 0, 1, \dots, 10$. A linha inferior (com círculos), correspondente a $\alpha = 0$, indica o MSE.

Tabela 5.4: Modelo moda obtido pelo critério ERR_2 a partir dos dados apresentados na Figura 5.8.

Termo	Parâmetro real	Parâmetro estimado	ERR_2	Número de ocorrências
$y(k-1)$	0,25	0,3132	0,8968	500
$u(k)^2$	1	1,0092	0,0252	500
$u(k-2)^2$	0	-0,1751	0,0440	500
$u(k)$	0	0,1022	0,0041	500
$y(k-1)^2$	-0,05	-0,0460	0,0149	466
constante	0,75	0,7228	0,0043	466
$u(k-1)$	0,8	0,8030	0,0015	466
$y(k-2)^2$	-0,025	-0,0215	0,0013	162
$u(k-1)^2$	0,5	0,5167	0,0006	125
$u(k-2)$	0,4	0,2296	0,0001	122

Tabela 5.5: Relação entre os termos do modelo da Tabela 5.4 e os possíveis termos de seu respectivo modelo de um passo à frente (parte 1).

Termo de \mathcal{M}_{2PF}	Opção	Possíveis termos de \mathcal{M}_{1PF}	Demais termos gerados em \mathcal{M}_{2PF}	Situação dos termos gerados em \mathcal{M}_{2PF}
$y(k-1)$	1.1	$y(k-1)$		
	1.2	$y(k-1)y(k-2)$ constante	$y(k-1)^2y(k-2)$ constante	Termo não pertencente ao modelo da Tabela 5.4 Termo pertencente ao modelo da Tabela 5.4
$u(k)^2$	2.3	$y(k-1)u(k-1)^2$ constante	$u(k)^2y(k-1)u(k-1)^2$	Termo não pertencente ao modelo da Tabela 5.4
			constante	Termo pertencente ao modelo da Tabela 5.4
	2.1	$y(k-1)$ $u(k-2)^2$	$y(k-1)$ $u(k-1)^2$	Termo pertencente ao modelo da Tabela 5.4 Termo pertencente ao modelo da Tabela 5.4
$u(k-2)^2$	2.2	$y(k-1)u(k-3)$ $u(k-2)$	$y(k-1)u(k-2)u(k-3)$ $u(k-1)$	Termo não pertencente ao modelo da Tabela 5.4 Termo pertencente ao modelo da Tabela 5.4
	2.3	$y(k-1)u(k-3)^2$ constante	$y(k-1)u(k-2)^2u(k-3)^2$ constante	Termo não pertencente ao modelo da Tabela 5.4 Termo pertencente ao modelo da Tabela 5.4
$u(k)$	1.2	$y(k-1)u(k-1)$ constante	$u(k)y(k-1)u(k-1)$ constante	Termo não pertencente ao modelo da Tabela 5.4 Termo pertencente ao modelo da Tabela 5.4
	2.1	$y(k-1)$ $y(k-1)^2$		$y(k-1)^n$ com $n > 1$ em \mathcal{M}_{1PF}
$y(k-1)^2$	2.2	$y(k-1)y(k-2)$ $y(k-1)$	$y(k-1)^2y(k-2)$	Termo não pertencente ao modelo da Tabela 5.4
			$y(k-1)y(k-2)$	Termo não pertencente ao modelo da Tabela 5.4
	2.3	$y(k-1)y(k-2)^2$ constante	$y(k-1)^3y(k-2)^2$ constante	Termo não pertencente ao modelo da Tabela 5.4 Termo pertencente ao modelo da Tabela 5.4

Tabela 5.6: Relação entre os termos do modelo da Tabela 5.4 e os possíveis termos de seu respectivo modelo de um passo à frente (parte 2).

Termo de \mathcal{M}_{2PF}	Opção	Possíveis termos de \mathcal{M}_{1PF}	Demais termos gerados em \mathcal{M}_{2PF}	Situação dos termos gerados em \mathcal{M}_{2PF}	
$u(k-1)$	1.1	$y(k-1)$ $u(k-1)$	$y(k-1)$	Termo pertencente ao modelo da Tabela 5.4	
	1.2	$y(k-1)u(k-2)$ constante	$u(k)$	Termo pertencente ao modelo da Tabela 5.4	
			$y(k-1)u(k-1)u(k-2)$ constante	Termo não pertencente ao modelo da Tabela 5.4 Termo pertencente ao modelo da Tabela 5.4	
	$y(k-2)^2$	2.1	$y(k-1)$ $y(k-2)^2$	$y(k-1)$	Termo pertencente ao modelo da Tabela 5.4
2.2		$y(k-1)y(k-3)$ $y(k-2)$	$y(k-1)y(k-2)y(k-3)$	Termo pertencente ao modelo da Tabela 5.4	
			$y(k-1)$	Termo não pertencente ao modelo da Tabela 5.4	
2.3		$y(k-1)y(k-3)^2$ constante	$y(k-1)y(k-2)^2y(k-3)^2$ constante	Termo pertencente ao modelo da Tabela 5.4 Termo não pertencente ao modelo da Tabela 5.4	
$u(k-1)^2$		2.1	$y(k-1)$ $u(k-1)^2$	$y(k-1)$ $u(k)^2$	Termo pertencente ao modelo da Tabela 5.4 Termo pertencente ao modelo da Tabela 5.4
		2.2	$y(k-1)u(k-2)$ $u(k-1)$	$y(k-1)u(k-1)u(k-2)$ $u(k)$	Termo não pertencente ao modelo da Tabela 5.4 Termo pertencente ao modelo da Tabela 5.4
	2.3	$y(k-1)u(k-2)^2$ constante	$y(k-1)u(k-1)^2u(k-2)^2$ constante	Termo não pertencente ao modelo da Tabela 5.4 Termo pertencente ao modelo da Tabela 5.4	
$u(k-2)$	1.1	$y(k-1)$ $u(k-2)$	$y(k-1)$ $u(k-1)$	Termo pertencente ao modelo da Tabela 5.4 Termo pertencente ao modelo da Tabela 5.4	
	1.2	$y(k-1)u(k-3)$ constante	$y(k-1)u(k-2)u(k-3)$ constante	Termo não pertencente ao modelo da Tabela 5.4 Termo pertencente ao modelo da Tabela 5.4	

$$y(k) = \theta_1 y(k-1) + \theta_2 u(k-2)^2 + \theta_3 u(k-1) + \theta_4 y(k-2)^2 + \theta_5 u(k-1)^2 + \theta_6 u(k-2) + \theta_7, \quad (5.3)$$

que corresponde à seguinte estrutura de dois passos à frente:

$$y(k+1) = \theta_1^2 y(k-1) + \theta_1 \theta_2 u(k-2)^2 + (\theta_1 \theta_3 + \theta_6) u(k-1) + \theta_1 \theta_4 y(k-2)^2 + (\theta_1 \theta_5 + \theta_2) u(k-1)^2 + \theta_1 \theta_6 u(k-2) + \theta_3 u(k) + \theta_4 y(k-1)^2 + \theta_5 u(k)^2 + \theta_7 (\theta_1 + 1). \quad (5.4)$$

Conforme discutido anteriormente, se fossem selecionados apenas nove termos do modelo moda de acordo com a estatística APRESS, a ausência do termo $u(k-2)$ levaria a um modelo de um passo à frente idêntico ao da equação (5.3) com exceção do termo $\theta_6 u(k-2)$. Entretanto, se fossem selecionados apenas oito termos do modelo moda, a ausência do termo $u(k-1)^2$ impediria a obtenção de um modelo de um passo à frente por dois motivos. Apesar de os termos de um passo à frente obtidos pelo termo $u(k)^2$ não poderem ser considerados pertencentes a \mathcal{M}_{1PF} (Tabela 5.5), verifica-se que a aplicação da opção 2.1 ao termo $u(k-1)^2$ de \mathcal{M}_{2PF} leva à geração do termo $u(k)^2$ (Tabela 5.6). Além disso, de acordo com a Tabela 5.5, se $u(k-1)^2$ fosse eliminado do modelo moda, todas as relações obtidas entre o termo $u(k-2)^2$ e os possíveis termos de \mathcal{M}_{1PF} implicariam a geração de termos não pertencentes ao modelo da Tabela 5.4. Portanto, se $u(k-1)^2$ fosse eliminado do modelo moda, não seria possível, a partir das relações presentes na Tabela 4.3, obter um modelo de um passo à frente capaz de gerar os termos $u(k)^2$ e $u(k-2)^2$ em seu respectivo modelo de dois passos. Diante disso, verifica-se que, dependendo do número de termos selecionado para compor um modelo de dois passos, a obtenção de um modelo de um passo à frente pode ser uma tarefa impossível. Isso mostra que a técnica de predição de dois passos à frente é sensível à escolha do número de termos utilizado no modelo.

Assim, quando os gráficos obtidos pelo critério de informação de Akaike e/ou pela estatística APRESS não sugerirem de forma clara o número de termos a ser utilizado em um modelo de dois passos à frente, ou seja, quando esses critérios apresentarem valores muito baixos (com pequenas variações) para mais de um número de termos, será considerado, neste trabalho, o número de termos correspondente ao menor valor obtido para o critério de informação. Portanto, no caso da Figura 5.12, em que o AIC

sugere um número de termos entre oito e dez, o modelo de dois passos à frente deve ser formado por dez termos. Quando não for possível determinar o modelo de um passo utilizando-se o número de termos proposto, deve-se, então, reduzir o número de termos gradativamente até que se consiga obter o modelo de um passo à frente³.

Analisando-se o modelo da equação (5.4), verifica-se que seus termos correspondem exatamente aos dez termos presentes no modelo da Tabela 5.4. Comparando-se, então, o modelo da equação (5.3) com o modelo da Tabela 5.3, observa-se que o critério ERR_2 eliminou o termo espúrio $y(k-2)$ do modelo de um passo à frente. Assim, um novo modelo foi estimado a partir do critério ERR_1 utilizando os termos presentes na Tabela 5.3 (com exceção do termo $y(k-2)$) como candidatos. A Tabela 5.7 apresenta esse novo modelo.

Tabela 5.7: Modelo obtido pelo critério ERR_1 a partir dos dados apresentados na Figura 5.8 após a eliminação do termo $y(k-2)$.

Termo	Parâmetro real	Parâmetro estimado	ERR_1
$y(k-1)$	0,5	0,5294	0,9690
$y(k-2)^2$	-0,05	-0,0447	0,0041
$u(k-1)^2$	1	1,0339	0,0038
$u(k-2)$	0,8	0,5995	0,0133
constante	0,5	0,4711	0,0043
$u(k-2)^2$	0	-0,1238	0,0000
$u(k-1)$	0	0,1413	0,0001

Aplicando-se o critério de informação de Akaike e a estatística APRESS ao novo modelo, verifica-se que ambos os métodos sugerem que o modelo seja formado por cinco termos (Figuras 5.14 e 5.15).

O novo modelo formado apenas pelo número de termos estimado pelo AIC e pela estatística APRESS encontra-se na Tabela 5.8. Verifica-se que os termos apresentados nessa tabela correspondem exatamente aos cinco termos presentes no sistema \mathcal{S}_1 .

³Ressalta-se que, no Capítulo 6, não será necessário esse nível de detalhamento.

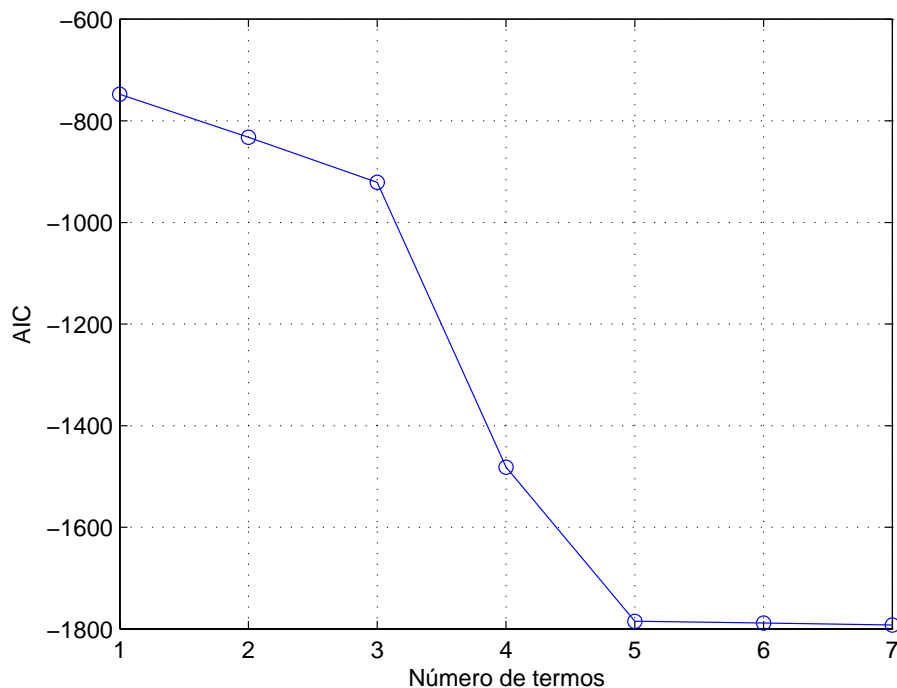


Figura 5.14: Resultado gráfico da aplicação do critério de informação de Akaike ao modelo obtido pelo critério ERR_1 a partir dos dados apresentados na Figura 5.8 após a eliminação do termo $y(k-2)$.

Portanto, para o caso analisado nesta seção, verificou-se que a obtenção da estrutura correta do sistema \mathcal{S}_1 a partir do critério ERR_1 , quando excitado por uma entrada AR, só foi possível após a aplicação da técnica de predição de dois passos à frente.

Para resolver esse mesmo problema, Piroddi e Spinelli (2003a) empregaram os algoritmos FRP (algoritmo baseado na redução do erro de predição + *pruning*), SEM (algoritmo baseado na redução do erro de simulação) e SEMP (algoritmo baseado na redução do erro de simulação + *pruning*)⁴. A Tabela 5.9 apresenta os modelos obtidos a partir de tais algoritmos. Deve-se ressaltar que, diferentemente deste trabalho, em Piroddi e Spinelli (2003a), foi identificado um único modelo utilizando uma massa de dados contendo 500 amostras.

⁴Tais algoritmos são apresentados na Seção 3.5.

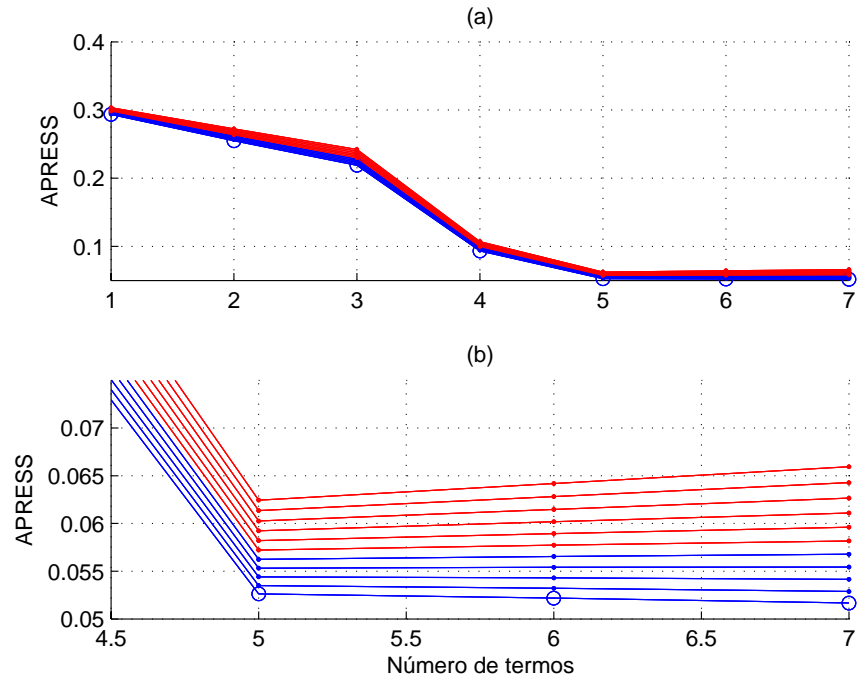


Figura 5.15: Resultado gráfico da aplicação da estatística APRESS ao modelo obtido pelo critério ERR_1 a partir dos dados apresentados na Figura 5.8 após a eliminação do termo $y(k-2)$.

Gráfico (a) em seu tamanho real e (b) em escala ampliada. As linhas, de baixo para cima, correspondem a $\alpha = 0, 1, \dots, 10$. A linha inferior (com círculos), correspondente a $\alpha = 0$, indica o MSE.

Tabela 5.8: Modelo final obtido pelo critério ERR_1 a partir dos dados apresentados na Figura 5.8.

Termo	Parâmetro real	Parâmetro estimado	ERR_1
$y(k-1)$	0,5	0,4912	0,9690
$y(k-2)^2$	-0,05	-0,0482	0,0041
$u(k-1)^2$	1	0,9850	0,0038
$u(k-2)$	0,8	0,7978	0,0133
constante	0,5	0,5070	0,0043

Como pode ser visto na Tabela 5.9, foram necessários sete termos para que o algoritmo SEM selecionasse todos os termos corretos, gerando, assim, dois termos espúrios. É interessante observar que um dos termos espúrios pertence a um agrupamento também espúrio. Por outro lado, tanto com o FRP quanto com o SEMP, a estrutura do modelo foi corretamente obtida. Além disso, os parâmetros estimados por esses dois algoritmos são idênticos.

Tabela 5.9: Modelos obtidos em Piroddi e Spinelli (2003a) para o sistema \mathcal{S}_1 a partir de uma entrada AR.

Algoritmo SEM		Algoritmos FRP e SEMP	
Termo	Parâmetro estimado	Termo	Parâmetro estimado
$\mathbf{u}(\mathbf{k} - 1)^2$	1,16900	$\mathbf{y}(\mathbf{k} - 1)$	0,49620
constante	0,25790	$\mathbf{u}(\mathbf{k} - 1)^2$	1,05100
$\mathbf{u}(\mathbf{k} - 2)$	0,72740	$\mathbf{y}(\mathbf{k} - 2)^2$	-0,05299
$y(k - 2)u(k - 1)$	0,03818	constante	0,48380
$y(k - 2)$	0,10040	$\mathbf{u}(\mathbf{k} - 2)$	0,78050
$\mathbf{y}(\mathbf{k} - 2)^2$	-0,08915		
$\mathbf{y}(\mathbf{k} - 1)$	0,70710		

Recentemente, esse mesmo problema foi abordado em Wei e Billings (2008). Para resolvê-lo, os autores empregaram um algoritmo denominado IFOS (*integrated forward orthogonal search*). Dependendo da técnica associada a ele, tal algoritmo pode apresentar as seguintes versões: IFOS-SC (quando associado à correlação quadrada) ou IFOS-MI (quando associado à informação mútua). A Tabela 5.10 apresenta os modelos obtidos a partir dos algoritmos IFOS-SC e IFOS-MI. É importante ressaltar que, em Wei e Billings (2008), diferentemente de Piroddi e Spinelli (2003a), foi utilizado um ruído $\nu(k)$ com variância de $0,05^2$ em vez de 0,05.

Como pode ser visto na Tabela 5.10, os cinco primeiros termos selecionados pelo algoritmo IFOS-SC correspondem exatamente aos cinco termos presentes no sistema \mathcal{S}_1 . Por outro lado, foram necessários sete termos para que o algoritmo IFOS-MI selecionasse todos os termos corretos,

gerando, assim, dois termos espúrios. Apesar de o método proposto em Wei e Billings (2008) depender um tempo computacional muito pequeno, tal método incorpora diversas técnicas, tais como correlação quadrada, informação mútua, testes de hipóteses e validação cruzada, aumentando, assim, a complexidade dos algoritmos.

Tabela 5.10: Modelos obtidos em Wei e Billings (2008) para o sistema \mathcal{S}_1 a partir de uma entrada AR.

Algoritmo IFOS-SC		Algoritmo IFOS-MI	
Termo	Parâmetro estimado	Termo	Parâmetro estimado
$\mathbf{y}(\mathbf{k} - 1)$	0,500106	$u(k - 1)$	0,006148
$\mathbf{y}(\mathbf{k} - 2)^2$	-0,049757	$\mathbf{u}(\mathbf{k} - 1)^2$	0,994118
$\mathbf{u}(\mathbf{k} - 1)^2$	1,000401	$\mathbf{y}(\mathbf{k} - 1)$	0,496906
$\mathbf{u}(\mathbf{k} - 2)$	0,806721	$\mathbf{y}(\mathbf{k} - 2)^2$	-0,049833
constante	0,493459	$u(k - 1)u(k - 2)$	0,011942
$y(k - 1)^2$	-0,000419	constante	0,499216
$u(k - 2)^2$	0,006367	$\mathbf{u}(\mathbf{k} - 2)$	0,800587
		$y(k - 1)u(k - 1)$	0,000024

Diante da complexidade dos algoritmos IFOS e do grande esforço computacional requerido pelos algoritmos usados em Piroddi e Spinelli (2003a), a técnica de predição de dois passos à frente se revelou, para o caso analisado nesta seção, como uma alternativa na seleção de estrutura de modelos polinomiais em condições não ideais de identificação.

5.2.2 Segundo Caso

Considere o seguinte sistema \mathcal{S}_2 representado pela equação (5.5):

$$y(k) = 0,5y(k-1) - 0,75u(k-1) + 0,55u(k-2) + 0,55y(k-1)u(k-1) + \nu(k), \quad (5.5)$$

que, na predição de dois passos à frente, pode ser representado da seguinte

forma:

$$\begin{aligned}
 y(k+1) = & 0,25y(k-1) + 0,175u(k-1) + 0,275u(k-2) + \\
 & 0,275y(k-1)u(k-1) - 0,75u(k) + 0,275u(k)y(k-1) - \\
 & 0,4125u(k)u(k-1) + 0,3025u(k)u(k-2) + \\
 & 0,3025u(k)y(k-1)u(k-1).
 \end{aligned} \tag{5.6}$$

Neste caso, foi utilizado um sinal de entrada $u(k)$ auto-regressivo de ordem dois AR(2) com média nula, variância unitária e dois pólos reais em 0,10 e 0,20, fazendo com que tal sinal, cuja função de autocorrelação (FAC) se encontra na Figura 5.17, fosse relativamente lento. O ruído $\nu(k)$ usado neste experimento foi gerado conforme descrito no primeiro caso. A Figura 5.16 apresenta os sinais de entrada e saída do sistema \mathcal{S}_2 , sendo que a relação sinal/ruído é de 29,06 dB.

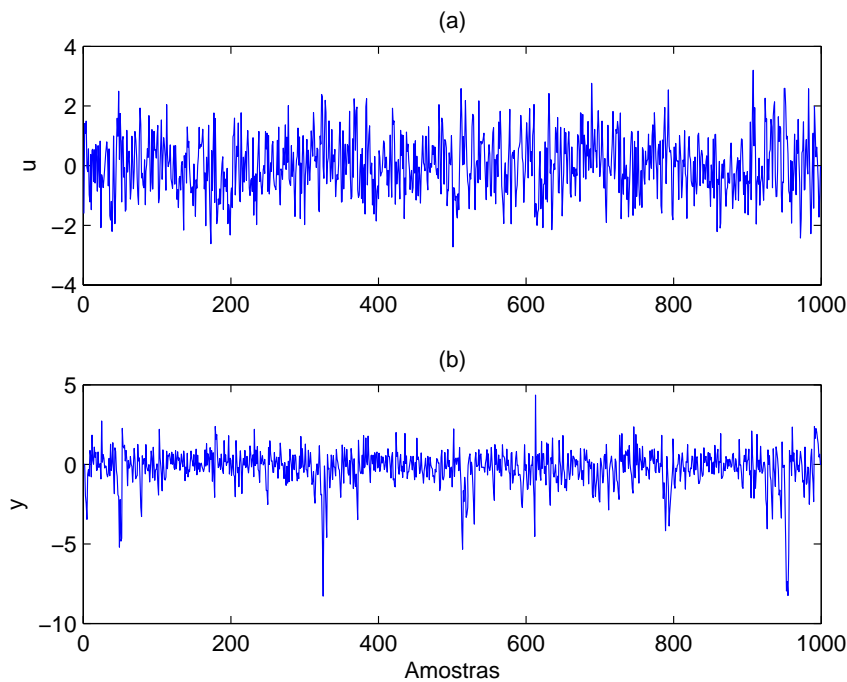


Figura 5.16: Dados do sistema \mathcal{S}_2 para uma entrada AR.

Sinais de (a) entrada e (b) saída (esses dados encontram-se disponíveis no arquivo `dados_myex2_ent_AR_cr`).

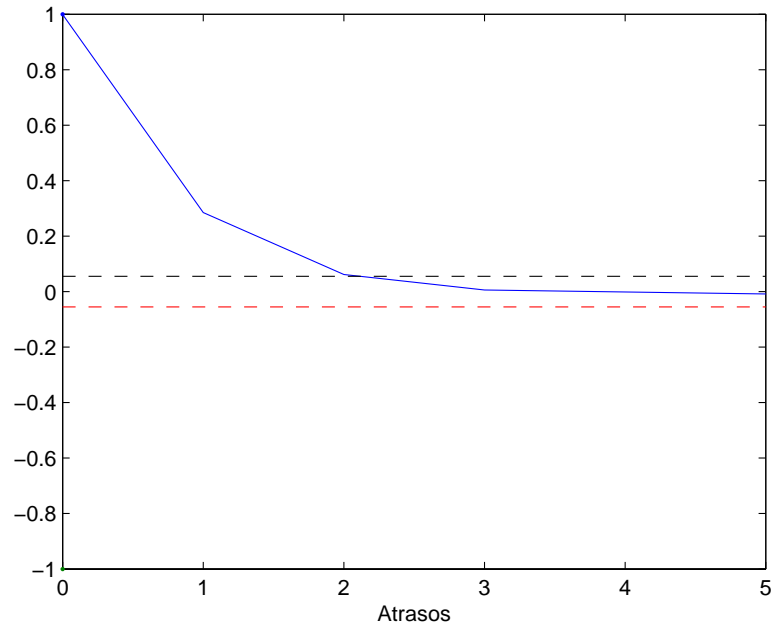


Figura 5.17: Resultado gráfico da aplicação da função FAC ao sinal de entrada apresentado na Figura 5.16.

O sinal de entrada não apresenta FAC semelhante à de um ruído branco, que é impulsiva.

Na identificação do sistema \mathcal{S}_2 , tanto na predição de um passo quanto na de dois passos à frente, foram utilizados, a princípio, modelos NARX com grau de não-linearidade igual a 3, máximo atraso no sinal de entrada igual a 2 e máximo atraso no sinal de saída igual a 1, totalizando, assim, 20 termos candidatos na predição de um passo e 35 na de dois passos à frente. É importante ressaltar que, nesse caso, os máximos atrasos utilizados durante o processo de identificação correspondem aos valores reais de n_u e n_y (ver equações (5.5) e (5.6)).

Analisando-se as Figuras 5.18 e 5.19, que apresentam, respectivamente, o resultado gráfico da aplicação do critério de informação de Akaike e da estatística APRESS ao modelo moda obtido pelo critério ERR_1 , verifica-se que ambos os métodos sugerem que o número ideal de termos do modelo seja cinco, sendo que o sistema \mathcal{S}_2 é composto por quatro termos.

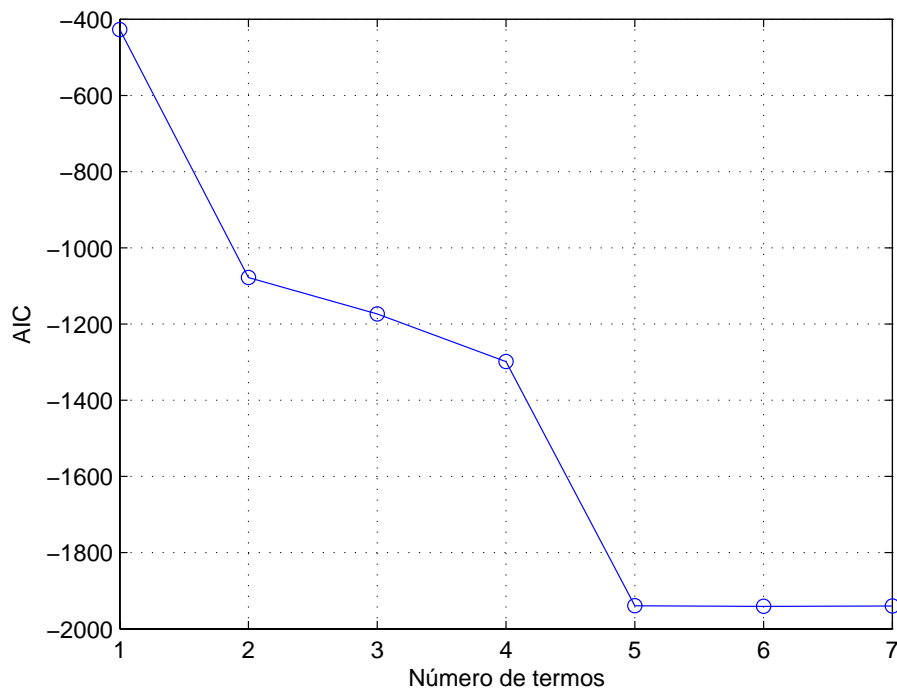


Figura 5.18: Resultado gráfico da aplicação do critério de informação de Akaike ao modelo moda obtido pelo critério ERR_1 a partir dos dados apresentados na Figura 5.16.

A Tabela 5.11 apresenta o modelo moda formado apenas pelo número de termos estimado pelo AIC e pela estatística APRESS. Apesar de todos os termos do sistema \mathcal{S}_2 se encontrarem presentes no modelo moda, tal modelo apresenta um termo espúrio pertencente a um agrupamento também espúrio. Deve-se observar, entretanto, que foram necessários exatamente cinco termos para que o critério ERR_1 selecionasse todos os termos corretos, apresentando, assim, 80% de acerto na escolha dos regressores. Dessa forma, verifica-se, neste caso, que os métodos empregados na estimação do número de termos do modelo novamente se mostraram úteis diante do erro cometido pelo ERR_1 ao selecionar um termo espúrio na tentativa de identificar o sistema \mathcal{S}_2 excitado por uma entrada AR.

Em relação à técnica de predição de dois passos à frente, verifica-se que tanto o critério de informação de Akaike (Figura 5.20) quanto a estatística APRESS (Figura 5.21) sugerem que o modelo moda obtido pelo critério ERR_2 seja formado por nove termos.

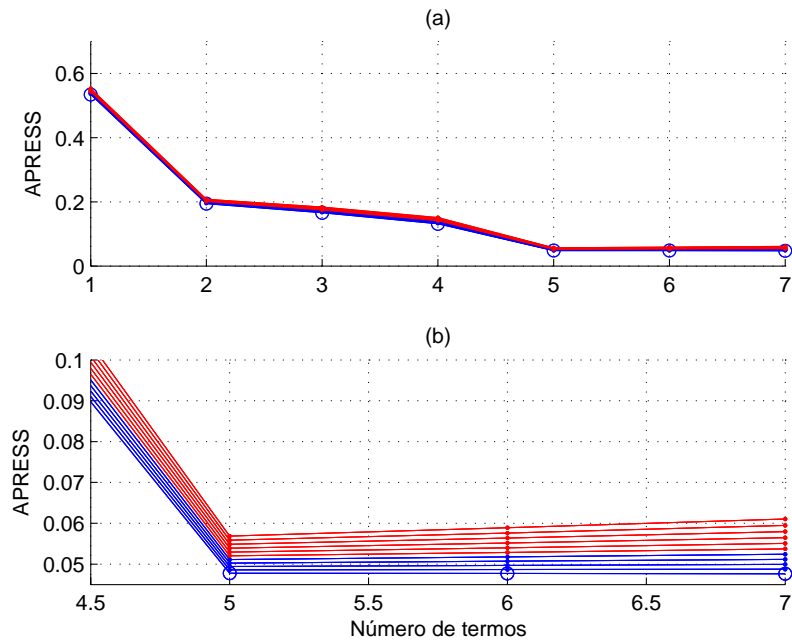


Figura 5.19: Resultado gráfico da aplicação da estatística APRESS ao modelo moda obtido pelo critério ERR_1 a partir dos dados apresentados na Figura 5.16.

Gráfico (a) em seu tamanho real e (b) em escala ampliada. As linhas, de baixo para cima, correspondem a $\alpha = 0, 1, \dots, 10$. A linha inferior (com círculos), correspondente a $\alpha = 0$, indica o MSE.

Tabela 5.11: Modelo moda obtido pelo critério ERR_1 a partir dos dados apresentados na Figura 5.16.

Termo	Parâmetro real	Parâmetro estimado	ERR_1	Número de ocorrências
$\mathbf{u}(\mathbf{k} - 1)$	-0,75	-0,7522	0,5537	453
$\mathbf{y}(\mathbf{k} - 1)\mathbf{u}(\mathbf{k} - 1)$	0,55	0,5453	0,2898	453
$y(k - 1)^2$	0	-0,0021	0,0264	257
$\mathbf{u}(\mathbf{k} - 2)$	0,55	0,5426	0,0237	257
$\mathbf{y}(\mathbf{k} - 1)$	0,5	0,4813	0,0676	257

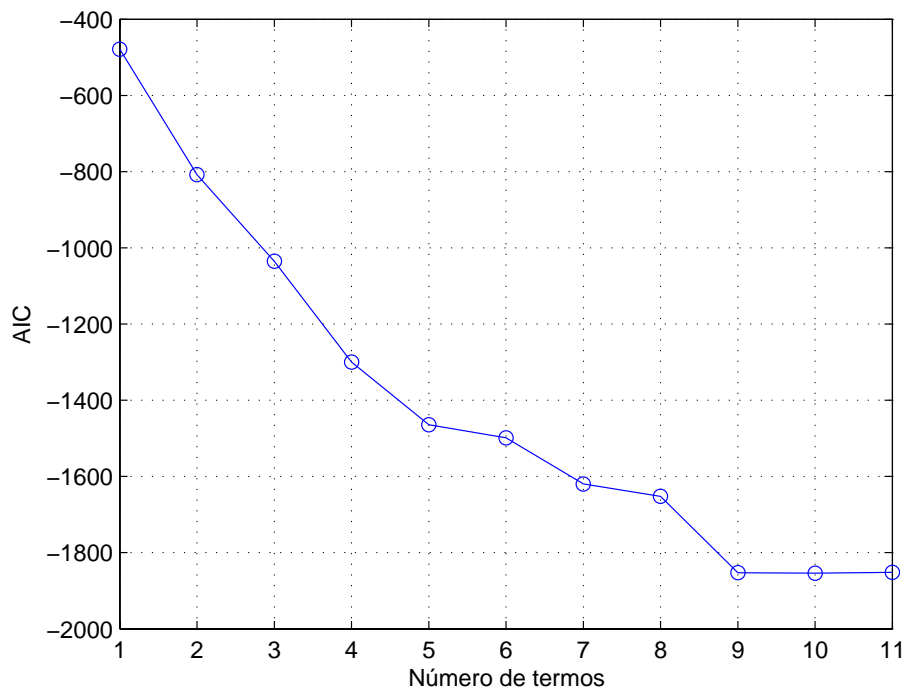


Figura 5.20: Resultado gráfico da aplicação do critério de informação de Akaike ao modelo moda obtido pelo critério ERR_2 a partir dos dados apresentados na Figura 5.16.

Como a equação (5.6) é composta por nove termos, pode-se afirmar que o número de termos do modelo foi corretamente determinado pelos dois métodos. A Tabela 5.12 apresenta o modelo moda formado apenas pelo número de termos estimado pelo AIC e pela estatística APRESS. Verifica-se que os termos apresentados nessa tabela correspondem exatamente aos nove termos presentes na equação (5.6). É importante observar que, apesar de o conjunto de termos candidatos usado na predição de dois passos à frente ser muito maior que o utilizado na de um passo, o critério ERR_2 não selecionou nenhum termo espúrio, apresentando, assim, 100% de acerto na escolha dos regressores.

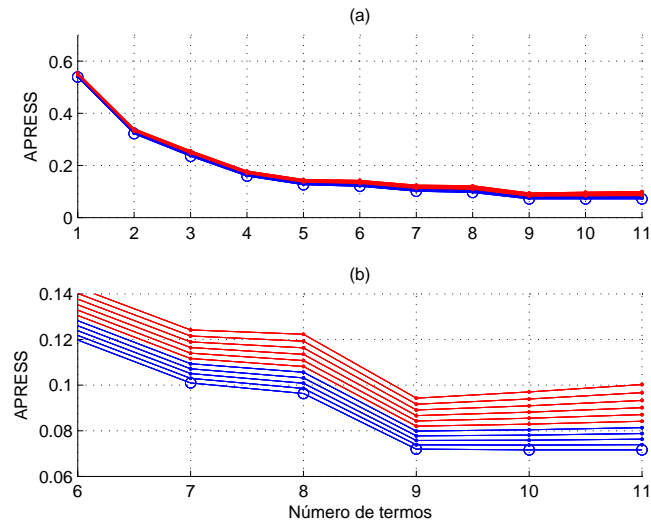


Figura 5.21: Resultado gráfico da aplicação da estatística APRESS ao modelo moda obtido pelo critério ERR_2 a partir dos dados apresentados na Figura 5.16.

Gráfico (a) em seu tamanho real e (b) em escala ampliada. As linhas, de baixo para cima, correspondem a $\alpha = 0, 1, \dots, 10$. A linha inferior (com círculos), correspondente a $\alpha = 0$, indica o MSE.

Tabela 5.12: Modelo moda obtido pelo critério ERR_2 a partir dos dados apresentados na Figura 5.16.

Termo	Parâmetro real	Parâmetro estimado	ERR_2	Número de ocorrências
$u(k)$	-0,75	-0,7553	0,5372	500
$u(k)u(k-1)$	-0,4125	-0,4171	0,1865	455
$u(k)y(k-1)u(k-1)$	0,3025	0,3034	0,0741	454
$y(k-1)u(k-1)$	0,275	0,2776	0,0653	454
$u(k-1)$	0,175	0,1801	0,0286	454
$u(k)u(k-2)$	0,3025	0,2977	0,0054	337
$u(k)y(k-1)$	0,275	0,2508	0,0161	337
$u(k-2)$	0,275	0,2537	0,0040	135
$y(k-1)$	0,25	0,2189	0,0209	135

A Tabela 5.14 apresenta a relação entre os termos do modelo de dois passos (Tabela 5.12) e os possíveis termos de seu respectivo modelo de um passo à frente. Ressalta-se que, como ausência de termo constante em modelos de dois passos à frente implica ausência de termo constante em modelos de um passo e como o modelo apresentado na Tabela 5.12 não apresenta esse tipo de termo, as opções 1.2 e 2.3 da Tabela 4.3 não foram empregadas. De acordo com a análise da Tabela 5.14, pôde-se chegar à seguinte estrutura de um passo à frente:

$$y(k) = \theta_1 y(k-1) + \theta_2 u(k-1) + \theta_3 u(k-2) + \theta_4 y(k-1)u(k-1). \quad (5.7)$$

Como era esperado, os termos do modelo (5.7) são idênticos aos da equação (5.5). Comparando-se, então, o modelo da equação (5.7) com o modelo da Tabela 5.11, verifica-se que o critério ERR_2 eliminou o único termo espúrio do modelo de um passo à frente ($y(k-1)^2$). Assim, um novo modelo foi estimado a partir do critério ERR_1 utilizando os termos presentes na Tabela 5.11 (com exceção do termo $y(k-1)^2$) como candidatos. A Tabela 5.13 apresenta esse novo modelo.

Tabela 5.13: Modelo obtido pelo critério ERR_1 a partir dos dados apresentados na Figura 5.16 após a eliminação do termo $y(k-1)^2$.

Termo	Parâmetro real	Parâmetro estimado	ERR_1
$u(k-1)$	-0,75	-0,7517	0,5537
$y(k-1)u(k-1)$	0,55	0,5464	0,2898
$y(k-1)$	0,5	0,4876	0,0217
$u(k-2)$	0,55	0,5456	0,0959

Como o único termo espúrio foi eliminado do conjunto de termos candidatos, os termos do modelo apresentado na Tabela 5.13 correspondem exatamente aos da equação (5.5). Assim, apesar de o sinal de entrada utilizado neste caso ser tão próximo de um sinal aleatório, o critério ERR_1 não foi capaz de recuperar a estrutura correta do sistema \mathcal{S}_2 sem o auxílio da técnica de predição de dois passos à frente. Portanto, como o modelo obtido pelo ERR_2 não apresentou nenhum termo espúrio, pode-se afirmar que, neste caso, tal critério se mostrou mais robusto que o ERR_1 .

Tabela 5.14: Relação entre os termos do modelo da Tabela 5.12 e os possíveis termos de seu respectivo modelo de um passo à frente.

Termo de \mathcal{M}_{2PF}	Opção	Possíveis termos de \mathcal{M}_{1PF}	Demais termos gerados em \mathcal{M}_{2PF}	Situação dos termos gerados em \mathcal{M}_{2PF}
$u(k)$				
$u(k)u(k-1)$	2.2	$y(k-1)u(k-1)$	$u(k)y(k-1)u(k-1)$	Termo pertencente ao modelo da Tabela 5.12
		$u(k-1)$	$u(k)$	Termo pertencente ao modelo da Tabela 5.12
$u(k)y(k-1)u(k-1)$	3.2	$y(k-1)u(k-1)$		
		$y(k-1)u(k-1)y(k-2)$	$u(k)y(k-1)^2y(k-2)u(k-1)$	Termo não pertencente ao modelo da Tabela 5.12
	3.3	$u(k-1)$	$u(k)$	Termo pertencente ao modelo da Tabela 5.12
		$y(k-1)u(k-1)u(k-2)$	$u(k)y(k-1)u(k-1)^2u(k-2)$	Termo não pertencente ao modelo da Tabela 5.12
		$y(k-1)$	$y(k-1)u(k-1)u(k-2)$	Termo não pertencente ao modelo da Tabela 5.12
		$y(k-1)$	$y(k-1)$	Termo pertencente ao modelo da Tabela 5.12
		$y(k-1)u(k-1)$	$y(k-1)$	Termo pertencente ao modelo da Tabela 5.12
		$y(k-1)u(k-1)u(k-1)$	$u(k)y(k-1)$	Termo pertencente ao modelo da Tabela 5.12
$y(k-1)u(k-1)$	2.1	$y(k-1)$	$y(k-1)$	Termo pertencente ao modelo da Tabela 5.12
		$y(k-1)u(k-1)$	$u(k)y(k-1)$	Termo pertencente ao modelo da Tabela 5.12
	2.2	$y(k-1)y(k-2)$	$u(k)y(k-1)u(k-1)$	Termo pertencente ao modelo da Tabela 5.12
		$u(k-1)$	$y(k-1)^2y(k-2)$	Termo não pertencente ao modelo da Tabela 5.12
		$y(k-1)u(k-2)$	$u(k)$	Termo pertencente ao modelo da Tabela 5.12
		$y(k-1)u(k-2)$	$y(k-1)u(k-1)u(k-2)$	Termo não pertencente ao modelo da Tabela 5.12
		$y(k-1)$	$y(k-1)u(k-2)$	Termo não pertencente ao modelo da Tabela 5.12
		$y(k-1)u(k-1)$	$y(k-1)$	Termo pertencente ao modelo da Tabela 5.12
$u(k)u(k-2)$	2.2	$y(k-1)u(k-1)$	$u(k)y(k-1)u(k-1)$	Termo pertencente ao modelo da Tabela 5.12
		$u(k-2)$	$u(k-1)$	Termo pertencente ao modelo da Tabela 5.12
$u(k)y(k-1)$	2.2	$y(k-1)u(k-1)$	$u(k)y(k-1)u(k-1)$	Termo pertencente ao modelo da Tabela 5.12
		$y(k-1)$	$y(k-1)u(k-1)$	Termo pertencente ao modelo da Tabela 5.12
$u(k-2)$	1.1	$y(k-1)$	$y(k-1)$	Termo pertencente ao modelo da Tabela 5.12
		$u(k-2)$	$u(k-1)$	Termo pertencente ao modelo da Tabela 5.12
$y(k-1)$	1.1	$y(k-1)$	$y(k-1)$	Termo pertencente ao modelo da Tabela 5.12
		$y(k-1)$	$u(k-1)$	Termo pertencente ao modelo da Tabela 5.12

Entretanto, ao se aumentar os valores dos máximos atrasos utilizados durante o processo de identificação, os resultados, que serão apresentados a seguir, mostram que o critério ERR_2 é sensível à escolha de n_u e n_y . Para isso, tanto na predição de um passo quanto na de dois passos à frente, foram utilizados modelos NARX com grau de não-linearidade e máximos atrasos, tanto no sinal de entrada quanto no sinal de saída, iguais a 3, totalizando, assim, 84 termos candidatos na predição de um passo e 120 na de dois passos à frente.

A partir das Figuras 5.22 e 5.23, verifica-se que, assim como no experimento anterior, tanto o critério de informação de Akaike quanto a estatística APRESS sugerem que o novo modelo moda obtido pelo critério ERR_1 seja formado por cinco termos.

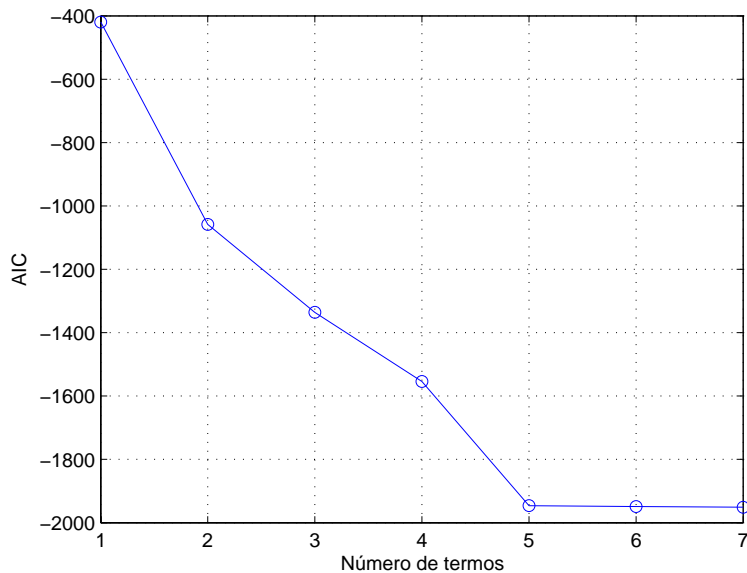


Figura 5.22: Resultado gráfico da aplicação do critério de informação de Akaike ao novo modelo moda obtido pelo critério ERR_1 a partir dos dados apresentados na Figura 5.16.

A Tabela 5.15, que apresenta o modelo moda formado apenas pelo número de termos estimado pelo AIC e pela estatística APRESS, mostra que, apesar de tal modelo ser formado por todos os termos do sistema \mathcal{S}_2 , ele também apresenta um termo espúrio, que, neste caso, pertence a um agrupamento não espúrio. Entretanto, assim como no experimento

anterior, foram necessários exatamente cinco termos para que o critério ERR_1 selecionasse todos os termos corretos. Assim, verifica-se que tal critério apresentou novamente 80% de acerto na escolha dos regressores.

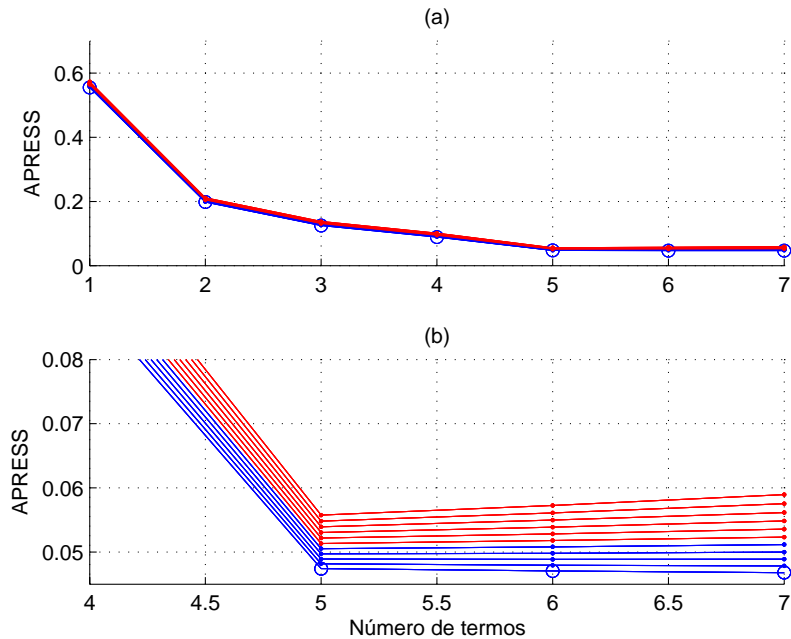


Figura 5.23: Resultado gráfico da aplicação da estatística APRESS ao novo modelo moda obtido pelo critério ERR_1 a partir dos dados apresentados na Figura 5.16.

Gráfico (a) em seu tamanho real e (b) em escala ampliada. As linhas, de baixo para cima, correspondem a $\alpha = 0,1, \dots, 10$. A linha inferior (com círculos), correspondente a $\alpha = 0$, indica o MSE.

Tabela 5.15: Novo modelo moda obtido pelo critério ERR_1 a partir dos dados apresentados na Figura 5.16.

Termo	Parâmetro real	Parâmetro estimado	ERR_1	Número de ocorrências
$u(k-1)$	-0,75	-0,7532	0,5518	453
$y(k-1)u(k-1)$	0,55	0,5469	0,2881	453
$y(k-2)u(k-2)$	0	0,0277	0,0592	453
$u(k-2)$	0,55	0,5291	0,0291	453
$y(k-1)$	0,5	0,4580	0,0335	453

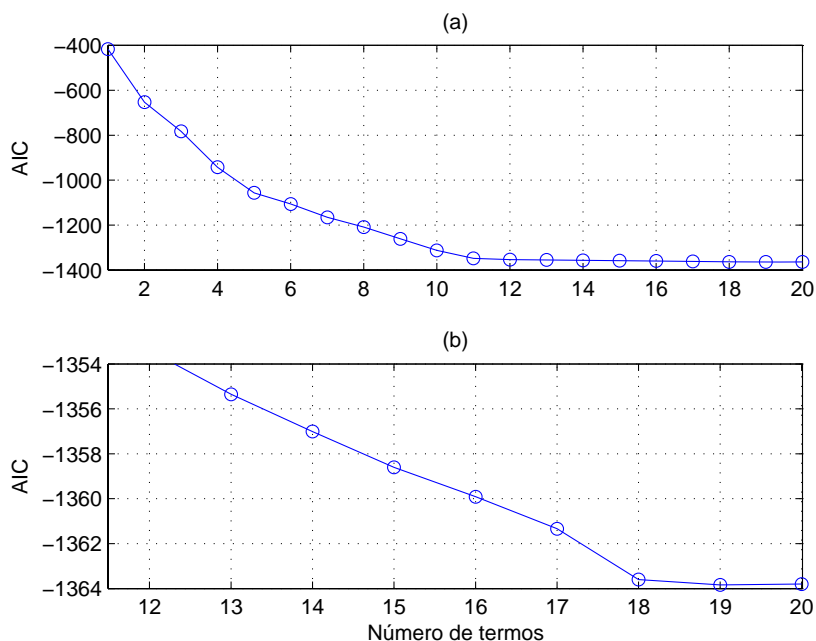


Figura 5.24: Resultado gráfico da aplicação do critério de informação de Akaike ao novo modelo moda obtido pelo critério ERR_2 a partir dos dados apresentados na Figura 5.16.

Gráfico (a) em seu tamanho real e (b) em escala ampliada.

Em relação à técnica de predição de dois passos à frente, verifica-se que, apesar de o critério de informação de Akaike sugerir um número de termos entre dez e doze para o novo modelo moda obtido pelo critério ERR_2 , a Figura 5.24b mostra que o menor valor de AIC corresponde a dezenove termos. Conforme discutido na seção anterior, quando o gráfico obtido por algum critério de informação não sugere de forma clara o número de termos a ser utilizado em um modelo de dois passos à frente, deve-se considerar, então, o número de termos correspondente ao menor valor obtido para o critério de informação, que, neste caso, equivale a dezenove. Entretanto, verificou-se que não foi possível obter um modelo de um passo à frente a partir dos dezenove termos considerados no modelo de dois passos, pois muitos desses termos implicam a geração de termos não selecionados pelo ERR_2 . Assim, ainda de acordo com o que foi discutido na Seção 5.2.1, reduziu-se o número de termos até que fosse possível determinar o modelo de um passo à frente. Procedendo dessa forma, verificou-se que o número

de termos a ser utilizado deveria ser igual a onze, que corresponde exatamente ao valor sugerido pela estatística APRESS (Figura 5.25). Portanto, neste caso, pode-se afirmar que tal critério, ao sugerir de forma clara um número de termos igual a onze, se mostrou mais eficiente que o AIC.

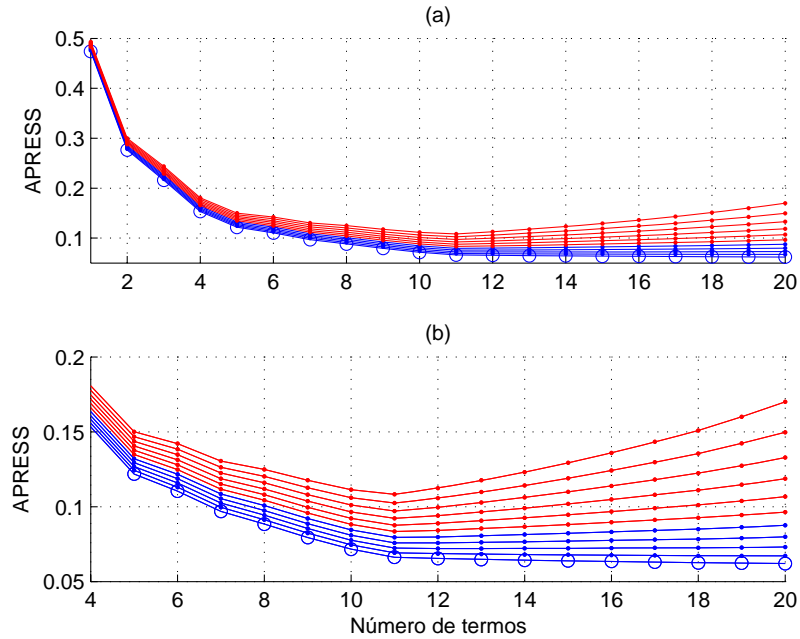


Figura 5.25: Resultado gráfico da aplicação da estatística APRESS ao novo modelo moda obtido pelo critério ERR_2 a partir dos dados apresentados na Figura 5.16.

Gráfico (a) em seu tamanho real e (b) em escala ampliada. As linhas, de baixo para cima, correspondem a $\alpha = 0, 1, \dots, 10$. A linha inferior (com círculos), correspondente a $\alpha = 0$, indica o MSE.

A Tabela 5.16 apresenta o modelo moda formado por apenas onze termos. Como pode ser observado, todos os termos da equação (5.6) encontram-se presentes no modelo moda. Entretanto, tal modelo apresenta dois termos espúrios pertencentes a agrupamentos não espúrios. Apesar disso, verifica-se que foram necessários exatamente onze termos para que o critério ERR_2 selecionasse todos os termos corretos, apresentando, assim, 81,82% de acerto na escolha dos regressores. Portanto, neste caso, verifica-se uma redução da robustez do ERR_2 em comparação com o experimento anterior.

Tabela 5.16: Novo modelo moda obtido pelo critério ERR_2 a partir dos dados apresentados na Figura 5.16.

Termo	Parâmetro real	Parâmetro estimado	ERR_2	Número de ocorrências
$\mathbf{u}(\mathbf{k})$	-0,75	-0,7549	0,5500	500
$\mathbf{u}(\mathbf{k})\mathbf{u}(\mathbf{k} - 1)$	-0,4125	-0,4284	0,1873	455
$\mathbf{u}(\mathbf{k})\mathbf{y}(\mathbf{k} - 1)\mathbf{u}(\mathbf{k} - 1)$	0,3025	0,3210	0,0578	453
$\mathbf{y}(\mathbf{k} - 1)\mathbf{u}(\mathbf{k} - 1)$	0,275	0,2770	0,0593	453
$\mathbf{u}(\mathbf{k} - 1)$	0,175	0,1798	0,0299	453
$u(k)y(k - 2)u(k - 2)$	0	0,0221	0,0109	234
$y(k - 2)u(k - 2)$	0	0,0243	0,0129	172
$\mathbf{u}(\mathbf{k})\mathbf{u}(\mathbf{k} - 2)$	0,3025	0,2983	0,0079	133
$\mathbf{u}(\mathbf{k} - 2)$	0,275	0,2353	0,0086	67
$\mathbf{u}(\mathbf{k})\mathbf{y}(\mathbf{k} - 1)$	0,275	0,2419	0,0074	67
$\mathbf{y}(\mathbf{k} - 1)$	0,25	0,1823	0,0051	67

As Tabelas 5.17 e 5.18 apresentam a relação entre os termos do modelo da Tabela 5.16 e os possíveis termos de seu respectivo modelo de um passo à frente. Assim como no experimento anterior, em virtude da ausência de termo constante no modelo de dois passos à frente (Tabela 5.16), as opções 1.2 e 2.3 da Tabela 4.3 não foram empregadas neste caso. Dessa forma, foi possível se chegar à seguinte estrutura de um passo à frente:

$$y(k) = \theta_1 u(k - 1) + \theta_2 y(k - 1)u(k - 1) + \theta_3 y(k - 2)u(k - 2) + \theta_4 u(k - 2) + \theta_5 y(k - 1). \quad (5.8)$$

Como pode ser observado, os termos do modelo (5.8) são idênticos aos do modelo moda obtido pelo critério ERR_1 (Tabela 5.15). Dessa forma, verifica-se que o ERR_2 não foi capaz de eliminar o único termo espúrio do modelo de um passo à frente ($y(k - 2)u(k - 2)$). Assim, pode-se afirmar que, neste caso, a técnica de predição de dois passos à frente não pôde ser empregada como uma ferramenta auxiliar na identificação do sistema \mathcal{S}_2 . Isso mostra que o critério ERR_2 é sensível à escolha dos máximos atrasos.

Tabela 5.17: Relação entre os termos do modelo da Tabela 5.16 e os possíveis termos de seu respectivo modelo de um passo à frente (parte 1).

Termo de \mathcal{M}_{2PF}	Opção	Possíveis termos de \mathcal{M}_{1PF}	Demais termos gerados em \mathcal{M}_{2PF}	Situação dos termos gerados em \mathcal{M}_{2PF}
$u(k)$				
$u(k)u(k-1)$	2.2	$y(k-1)u(k-1)$	$u(k)y(k-1)u(k-1)$	Termo pertencente ao modelo da Tabela 5.16
		$u(k-1)$	$u(k)$	Termo pertencente ao modelo da Tabela 5.16
$u(k)y(k-1)u(k-1)$	3.2	$y(k-1)u(k-1)$		
		$y(k-1)u(k-1)y(k-2)$	$u(k)y(k-1)^2y(k-2)u(k-1)$	Termo não pertencente ao modelo da Tabela 5.16
	3.3	$u(k-1)$	$u(k)$	Termo pertencente ao modelo da Tabela 5.16
		$y(k-1)u(k-1)u(k-2)$	$u(k)y(k-1)u(k-1)^2u(k-2)$	Termo não pertencente ao modelo da Tabela 5.16
		$y(k-1)$	$y(k-1)u(k-1)u(k-2)$	Termo não pertencente ao modelo da Tabela 5.16
		$y(k-1)$	$y(k-1)$	Termo pertencente ao modelo da Tabela 5.16
	2.1	$y(k-1)$	$y(k-1)$	Termo pertencente ao modelo da Tabela 5.16
		$y(k-1)u(k-1)$	$u(k)y(k-1)$	Termo pertencente ao modelo da Tabela 5.16
$y(k-1)u(k-1)$	2.2	$y(k-1)y(k-2)$	$u(k)y(k-1)u(k-1)$	Termo pertencente ao modelo da Tabela 5.16
		$u(k-1)$	$y(k-1)^2y(k-2)$	Termo não pertencente ao modelo da Tabela 5.16
		$u(k)$	$u(k)$	Termo pertencente ao modelo da Tabela 5.16
			$y(k-1)u(k-1)u(k-2)$	$y(k-1)u(k-1)u(k-2)$
		$y(k-1)$	$y(k-1)u(k-2)$	Termo não pertencente ao modelo da Tabela 5.16
		$y(k-1)$	$y(k-1)$	Termo pertencente ao modelo da Tabela 5.16
$u(k-1)$	1.1	$y(k-1)$	$y(k-1)$	Termo pertencente ao modelo da Tabela 5.16
		$u(k-1)$	$u(k)$	Termo pertencente ao modelo da Tabela 5.16

Tabela 5.18: Relação entre os termos do modelo da Tabela 5.16 e os possíveis termos de seu respectivo modelo de um passo à frente (parte 2).

Termo de \mathcal{M}_{2PF}	Opção	Possíveis termos de \mathcal{M}_{1PF}	Demais termos gerados em \mathcal{M}_{2PF}	Situação dos termos gerados em \mathcal{M}_{2PF}
$u^{(k)}y^{(k-2)}u^{(k-2)}$	3.3	$y^{(k-1)}u^{(k-1)}$ $y^{(k-2)}u^{(k-2)}$	$u^{(k)}y^{(k-1)}u^{(k-1)}$ $y^{(k-1)}u^{(k-1)}$	Termo pertencente ao modelo da Tabela 5.16 Termo pertencente ao modelo da Tabela 5.16
		$y^{(k-1)}u^{(k-1)}y^{(k-3)}$ $u^{(k-2)}$ $y^{(k-1)}u^{(k-1)}u^{(k-3)}$ $y^{(k-2)}$	$u^{(k)}y^{(k-1)}y^{(k-2)}y^{(k-3)}u^{(k-1)}$ $u^{(k-1)}$ $u^{(k)}y^{(k-1)}u^{(k-1)}u^{(k-2)}u^{(k-3)}$ $y^{(k-1)}$	Termo não pertencente ao modelo da Tabela 5.16 Termo pertencente ao modelo da Tabela 5.16 Termo não pertencente ao modelo da Tabela 5.16 Termo pertencente ao modelo da Tabela 5.16
$y^{(k-2)}u^{(k-2)}$	2.1	$y^{(k-1)}$ $y^{(k-2)}u^{(k-2)}$	$y^{(k-1)}$ $y^{(k-1)}u^{(k-1)}$	Termo pertencente ao modelo da Tabela 5.16 Termo pertencente ao modelo da Tabela 5.16
		$y^{(k-1)}y^{(k-3)}$ $u^{(k-2)}$ $y^{(k-1)}u^{(k-3)}$ $y^{(k-2)}$	$y^{(k-1)}y^{(k-2)}y^{(k-3)}$ $u^{(k-1)}$ $y^{(k-1)}u^{(k-2)}u^{(k-3)}$ $y^{(k-1)}$	Termo não pertencente ao modelo da Tabela 5.16 Termo pertencente ao modelo da Tabela 5.16 Termo não pertencente ao modelo da Tabela 5.16 Termo pertencente ao modelo da Tabela 5.16
$u^{(k)}u^{(k-2)}$	2.2	$y^{(k-1)}u^{(k-1)}$ $u^{(k-2)}$	$u^{(k)}y^{(k-1)}u^{(k-1)}$ $u^{(k-1)}$	Termo pertencente ao modelo da Tabela 5.16 Termo pertencente ao modelo da Tabela 5.16
$u^{(k-2)}$	1.1	$y^{(k-1)}$ $u^{(k-2)}$	$y^{(k-1)}$ $u^{(k-1)}$	Termo pertencente ao modelo da Tabela 5.16 Termo pertencente ao modelo da Tabela 5.16
$u^{(k)}y^{(k-1)}$	2.2	$y^{(k-1)}u^{(k-1)}$ $y^{(k-1)}$	$u^{(k)}y^{(k-1)}u^{(k-1)}$ $y^{(k-1)}u^{(k-1)}$	Termo pertencente ao modelo da Tabela 5.16 Termo pertencente ao modelo da Tabela 5.16
$y^{(k-1)}$	1.1	$y^{(k-1)}$	$y^{(k-1)}$	Termo pertencente ao modelo da Tabela 5.16

5.2.3 Terceiro Caso

Considere o seguinte sistema \mathcal{S}_3 representado pela equação (5.9):

$$y(k) = 0,7y(k-1) - 0,25y(k-2) - 0,5u(k-2)^2 + 0,3y(k-1)y(k-2) + \nu(k), \quad (5.9)$$

que, na predição de dois passos à frente, pode ser representado da seguinte forma:

$$\begin{aligned} y(k+1) = & 0,24y(k-1) - 0,175y(k-2) - 0,35u(k-2)^2 + \\ & 0,135y(k-1)y(k-2) - 0,5u(k-1)^2 + 0,21y(k-1)^2 - \\ & 0,15y(k-1)u(k-2)^2 + 0,09y(k-1)^2y(k-2). \end{aligned} \quad (5.10)$$

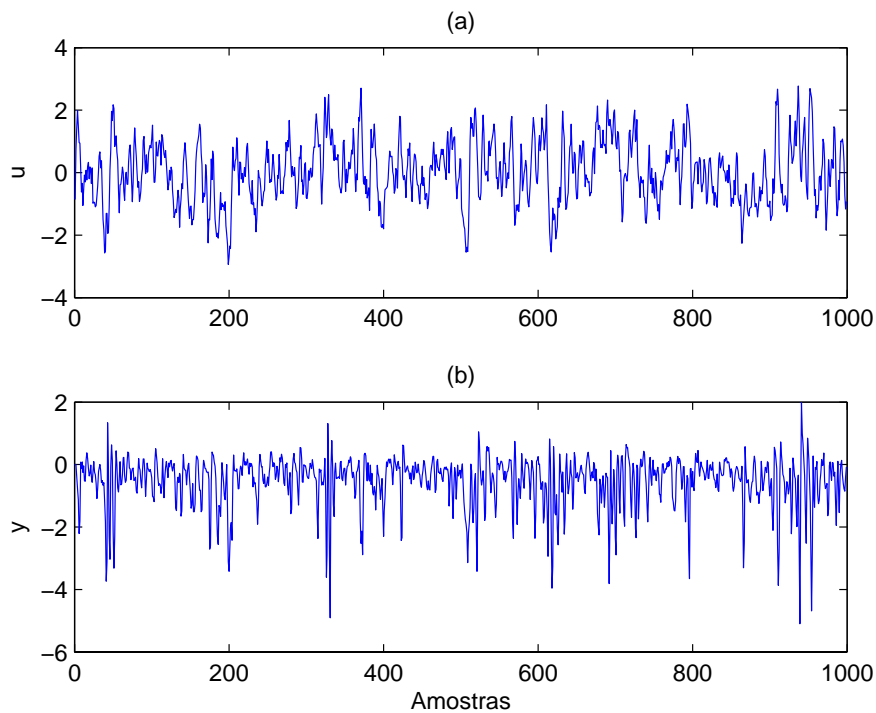


Figura 5.26: Dados do sistema \mathcal{S}_3 para uma entrada AR.

Sinais de (a) entrada e (b) saída (esses dados encontram-se disponíveis no arquivo `dados_myex3_ent_AR_cr`).

Assim como na Seção 5.2.2, o sistema considerado neste caso foi excitado com um sinal de entrada $u(k)$ auto-regressivo de ordem dois AR(2) com média nula e variância unitária. Para que tal sinal fosse relativamente lento, foram utilizados dois pólos reais em 0,50 e 0,60. A Figura 5.27 apresenta o resultado gráfico da aplicação da função de autocorrelação (FAC) ao sinal $u(k)$. O ruído $\nu(k)$ usado neste experimento foi gerado conforme descrito no primeiro caso, porém, com variância de 0,03. A Figura 5.26 mostra os sinais de entrada e saída do sistema \mathcal{S}_3 , cuja relação sinal/ruído é de 26,83 dB.

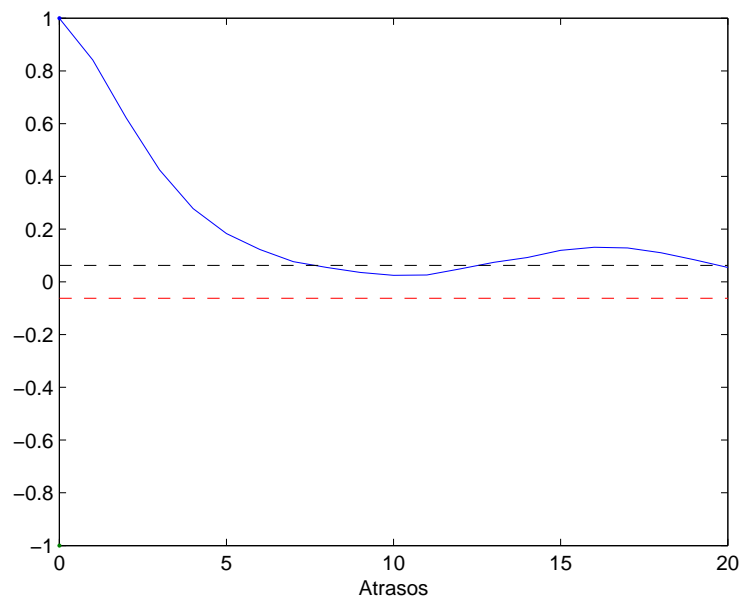


Figura 5.27: Resultado gráfico da aplicação da função FAC ao sinal de entrada apresentado na Figura 5.26.

O sinal de entrada não apresenta FAC semelhante à de um ruído branco, que é impulsiva.

A fim de identificar o sistema \mathcal{S}_3 , foram utilizados, tanto na predição de um passo quanto na de dois passos à frente, modelos NARX com grau de não-linearidade e máximos atrasos iguais a 3, sendo que, neste caso, os valores reais de n_u e n_y correspondem a 2 (ver equações (5.9) e (5.10)). Dessa forma, foram gerados 84 termos candidatos na predição de um passo e 120 na de dois passos à frente.

As Figuras 5.28 e 5.29 mostram que tanto o critério de informação de Akaike quanto a estatística APRESS sugerem que o número ideal de termos do modelo moda obtido pelo critério ERR_1 seja cinco.

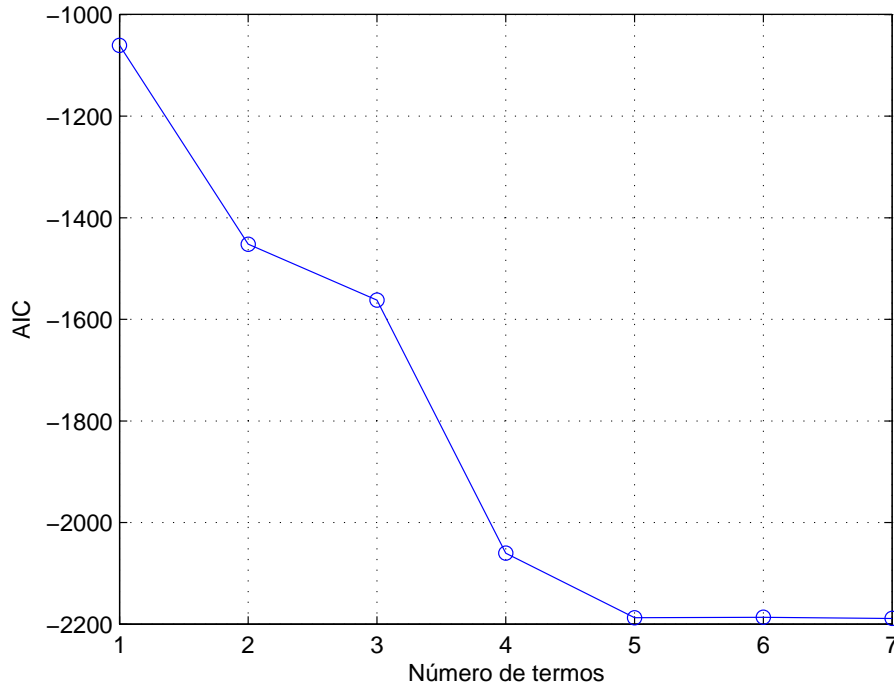


Figura 5.28: Resultado gráfico da aplicação do critério de informação de Akaike ao modelo moda obtido pelo critério ERR_1 a partir dos dados apresentados na Figura 5.26.

O modelo moda, composto apenas pelo número de termos estimado pelo AIC e pela estatística APRESS, é apresentado na Tabela 5.19. Como pode ser observado, todos os termos do sistema \mathcal{S}_3 encontram-se presentes no modelo moda. Além disso, tal modelo também apresenta um termo espúrio pertencente a um agrupamento não espúrio. Apesar disso, verifica-se que foram necessários exatamente cinco termos para que o ERR_1 conseguisse selecionar todos os termos corretos, correspondendo, assim, a 80% de acerto na escolha dos regressores.

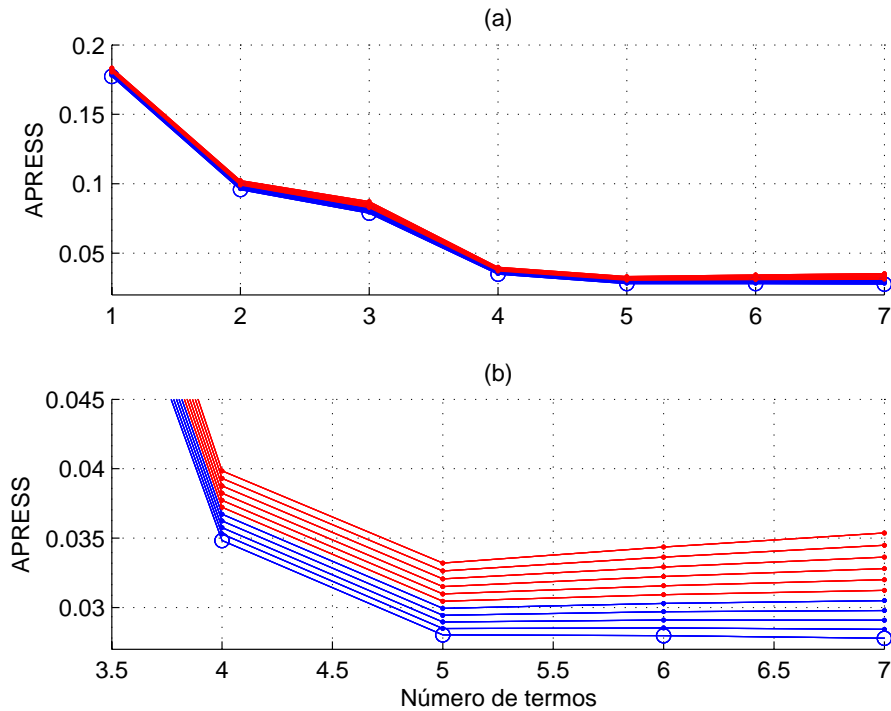


Figura 5.29: Resultado gráfico da aplicação da estatística APRESS ao modelo moda obtido pelo critério ERR_1 a partir dos dados apresentados na Figura 5.26.

Gráfico (a) em seu tamanho real e (b) em escala ampliada. As linhas, de baixo para cima, correspondem a $\alpha = 0, 1, \dots, 10$. A linha inferior (com círculos), correspondente a $\alpha = 0$, indica o MSE.

Tabela 5.19: Modelo moda obtido pelo critério ERR_1 a partir dos dados apresentados na Figura 5.26.

Termo	Parâmetro real	Parâmetro estimado	ERR_1	Número de ocorrências
$u(k-2)^2$	-0,5	-0,5100	0,8171	500
$y(k-2)^2$	0	0,0080	0,0842	436
$y(k-1)$	0,7	0,6437	0,0174	411
$y(k-1)y(k-2)$	0,3	0,2828	0,0454	411
$y(k-2)$	-0,25	-0,2193	0,0070	411

Na predição de dois passos à frente, verifica-se que, de acordo com o critério de informação de Akaike (Figura 5.30) e a estatística APRESS (Figura 5.31), o modelo moda obtido pelo critério ERR_2 deve ser formado por oito termos.

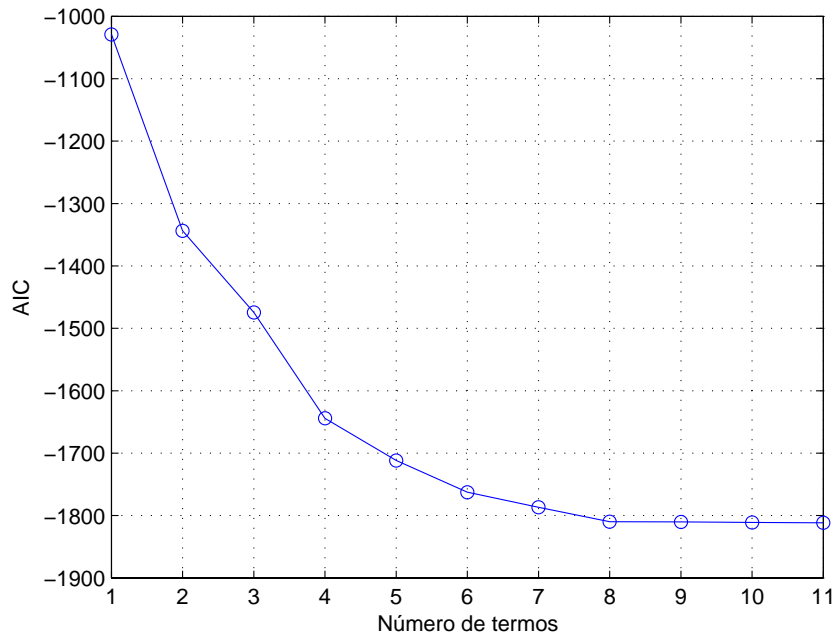


Figura 5.30: Resultado gráfico da aplicação do critério de informação de Akaike ao modelo moda obtido pelo critério ERR_2 a partir dos dados apresentados na Figura 5.26.

A partir da Tabela 5.20, que apresenta o modelo moda formado por apenas oito termos, é possível perceber que, neste caso, o critério ERR_2 não selecionou nenhum termo espúrio. Dessa forma, tem-se, então, 100% de acerto na escolha dos regressores.

A relação entre os termos do modelo de dois passos (Tabela 5.20) e os possíveis termos de seu respectivo modelo de um passo à frente se encontra nas Tabelas 5.21 e 5.22. Assim como na seção anterior, devido à ausência de termo constante no modelo de dois passos, as opções 1.2 e 2.3 da Tabela 4.3 não foram empregadas neste caso. Assim, obteve-se a seguinte estrutura de um passo à frente:

$$y(k) = \theta_1 y(k-1) + \theta_2 y(k-2) + \theta_3 u(k-2)^2 + \theta_4 y(k-1)y(k-2). \quad (5.11)$$

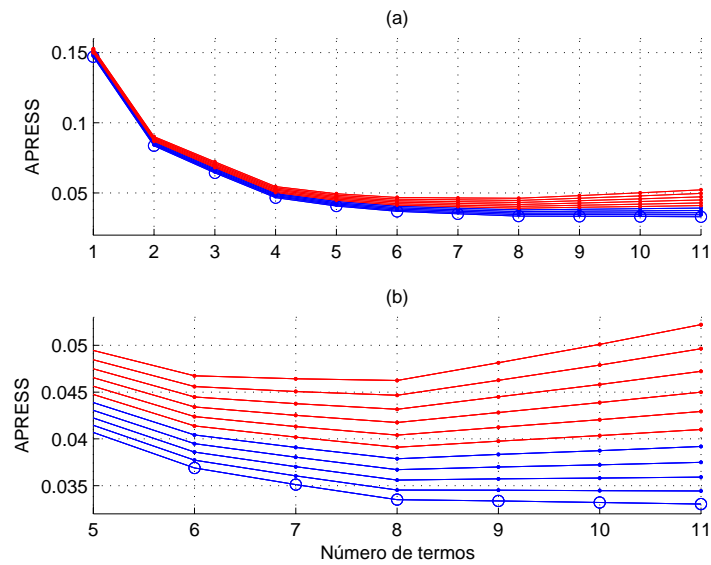


Figura 5.31: Resultado gráfico da aplicação da estatística APRESS ao modelo moda obtido pelo critério ERR_2 a partir dos dados apresentados na Figura 5.26.

Gráfico (a) em seu tamanho real e (b) em escala ampliada. As linhas, de baixo para cima, correspondem a $\alpha = 0, 1, \dots, 10$. A linha inferior (com círculos), correspondente a $\alpha = 0$, indica o MSE.

Tabela 5.20: Modelo moda obtido pelo critério ERR_2 a partir dos dados apresentados na Figura 5.26.

Termo	Parâmetro real	Parâmetro estimado	ERR_2	Número de ocorrências
$u(k-1)^2$	-0,5	-0,5049	0,8236	500
$y(k-1)^2$	0,21	0,2102	0,0761	296
$u(k-2)^2$	-0,35	-0,3346	0,0231	294
$y(k-1)u(k-2)^2$	-0,15	-0,1472	0,0212	294
$y(k-2)$	-0,175	-0,1473	0,0072	294
$y(k-1)^2y(k-2)$	0,09	0,0886	0,0045	294
$y(k-1)$	0,24	0,2163	0,0021	294
$y(k-1)y(k-2)$	0,135	0,1202	0,0019	294

Tabela 5.21: Relação entre os termos do modelo da Tabela 5.20 e os possíveis termos de seu respectivo modelo de um passo à frente (parte 1).

Termo de \mathcal{M}_{2PF}	Opção	Possíveis termos de \mathcal{M}_{1PF}	Demais termos gerados em \mathcal{M}_{2PF}	Situação dos termos gerados em \mathcal{M}_{2PF}
$u(k-1)^2$	2.1	$y(k-1)$	$y(k-1)$	Termo pertencente ao modelo da Tabela 5.20
		$u(k-1)^2$	$u(k)^2$	Termo não pertencente ao modelo da Tabela 5.20
	2.2	$y(k-1)u(k-2)$	$y(k-1)u(k-1)u(k-2)$	Termo não pertencente ao modelo da Tabela 5.20
		$u(k-1)$	$u(k)$	Termo não pertencente ao modelo da Tabela 5.20
$y(k-1)^2$	2.1	$y(k-1)$		$y(k-1)^n$ com $n > 1$ em \mathcal{M}_{1PF}
		$y(k-1)^2$		
	2.2	$y(k-1)y(k-2)$	$y(k-1)^2y(k-2)$	Termo pertencente ao modelo da Tabela 5.20
		$y(k-1)$	$y(k-1)y(k-2)$	Termo pertencente ao modelo da Tabela 5.20
			$y(k-1)$	Termo pertencente ao modelo da Tabela 5.20
				Termo pertencente ao modelo da Tabela 5.20
$u(k-2)^2$	2.1	$y(k-1)$	$y(k-1)$	Termo pertencente ao modelo da Tabela 5.20
		$u(k-2)^2$	$u(k-1)^2$	Termo pertencente ao modelo da Tabela 5.20
	2.2	$y(k-1)u(k-3)$	$y(k-1)u(k-2)u(k-3)$	Termo não pertencente ao modelo da Tabela 5.20
		$u(k-2)$	$u(k-1)$	Termo não pertencente ao modelo da Tabela 5.20
			$y(k-1)$	Termo pertencente ao modelo da Tabela 5.20
				Termo pertencente ao modelo da Tabela 5.20
$y(k-1)u(k-2)^2$	3.1	$y(k-1)$	$y(k-1)u(k-1)^2$	Termo não pertencente ao modelo da Tabela 5.20
		$y(k-1)u(k-2)^2$	$y(k-1)u(k-1)^2u(k-2)^2$	Termo não pertencente ao modelo da Tabela 5.20
	3.2	$y(k-1)y(k-2)$	$y(k-1)^2y(k-2)$	Termo pertencente ao modelo da Tabela 5.20
		$u(k-2)^2$	$u(k-1)^2$	Termo pertencente ao modelo da Tabela 5.20
		$y(k-1)u(k-3)$	$y(k-1)u(k-2)u(k-3)$	Termo não pertencente ao modelo da Tabela 5.20
		$y(k-1)u(k-2)$	$y(k-1)u(k-1)u(k-3)$	Termo não pertencente ao modelo da Tabela 5.20
3.3	$y(k-1)y(k-2)u(k-3)$	$y(k-1)u(k-2)u(k-3)$	$y(k-1)^2y(k-2)u(k-2)u(k-3)$	Termo não pertencente ao modelo da Tabela 5.20
		$u(k-2)$	$u(k-1)$	Termo não pertencente ao modelo da Tabela 5.20
	$y(k-1)u(k-3)^2$	$y(k-1)u(k-2)^2u(k-3)^2$	$y(k-1)u(k-2)^2u(k-3)^2$	Termo não pertencente ao modelo da Tabela 5.20
		$y(k-1)$	$y(k-1)u(k-3)^2$	Termo não pertencente ao modelo da Tabela 5.20

Tabela 5.22: Relação entre os termos do modelo da Tabela 5.20 e os possíveis termos de seu respectivo modelo de um passo à frente (parte 2).

Termo de \mathcal{M}_{2PF}	Opção	Possíveis termos de \mathcal{M}_{1PF}	Demais termos gerados em \mathcal{M}_{2PF}	Situação dos termos gerados em \mathcal{M}_{2PF}
$y^{(k-2)}$	1.1	$y^{(k-1)}$ $y^{(k-2)}$	$y^{(k-1)}$	Termo pertencente ao modelo da Tabela 5.20
	3.1	$y^{(k-1)}$ $y^{(k-1)^2}y^{(k-2)}$ $y^{(k-1)}y^{(k-2)}$		$y^{(k-1)^n}$ com $n > 1$ em \mathcal{M}_{1PF}
$y^{(k-1)^2}y^{(k-2)}$	3.2	$y^{(k-1)}y^{(k-3)}$ $y^{(k-1)^2}$		$y^{(k-1)^n}$ com $n > 1$ em \mathcal{M}_{1PF}
			$y^{(k-1)}y^{(k-2)^2}$ $y^{(k-2)}$	Termo não pertencente ao modelo da Tabela 5.20 Termo pertencente ao modelo da Tabela 5.20
	3.3	$y^{(k-1)}y^{(k-2)}y^{(k-3)}$ $y^{(k-1)}$	$y^{(k-1)^2}y^{(k-2)^2}y^{(k-3)}$ $y^{(k-1)}y^{(k-2)}y^{(k-3)}$ $y^{(k-1)}$	Termo não pertencente ao modelo da Tabela 5.20 Termo não pertencente ao modelo da Tabela 5.20 Termo pertencente ao modelo da Tabela 5.20
$y^{(k-1)}$	1.1	$y^{(k-1)}$		
			$y^{(k-1)}$	Termo pertencente ao modelo da Tabela 5.20
			$y^{(k-1)^2}$	Termo pertencente ao modelo da Tabela 5.20
			$y^{(k-1)^2}y^{(k-2)}$ $y^{(k-1)}$	Termo pertencente ao modelo da Tabela 5.20 Termo pertencente ao modelo da Tabela 5.20
$y^{(k-1)}y^{(k-2)}$	2.1	$y^{(k-1)}$ $y^{(k-1)}y^{(k-2)}$		
			$y^{(k-1)}y^{(k-2)}$ $y^{(k-2)}$	Termo pertencente ao modelo da Tabela 5.20 Termo pertencente ao modelo da Tabela 5.20
	2.2	$y^{(k-1)}y^{(k-3)}$ $y^{(k-1)}$	$y^{(k-1)}y^{(k-2)}$ $y^{(k-1)}$	Termo não pertencente ao modelo da Tabela 5.20 Termo não pertencente ao modelo da Tabela 5.20
			$y^{(k-1)}y^{(k-3)}$ $y^{(k-1)}$	Termo não pertencente ao modelo da Tabela 5.20 Termo pertencente ao modelo da Tabela 5.20

Conforme esperado, os termos do modelo de um passo à frente obtido a partir do modelo de dois passos (equação (5.11)) correspondem aos mesmos termos da equação (5.9). Dessa forma, ao se comparar o modelo da equação (5.11) com o modelo da Tabela 5.19, observa-se que o critério ERR_2 foi capaz de eliminar o único termo espúrio do modelo de um passo à frente ($y(k-2)^2$). Utilizando-se os termos presentes na Tabela 5.19 (com exceção do termo $y(k-2)^2$) como candidatos, um novo modelo, cujos termos evidentemente são idênticos aos da equação (5.9), foi estimado a partir do critério ERR_1 (Tabela 5.23).

Tabela 5.23: Modelo obtido pelo critério ERR_1 a partir dos dados apresentados na Figura 5.26 após a eliminação do termo $y(k-2)^2$.

Termo	Parâmetro real	Parâmetro estimado	ERR_1
$\mathbf{u(k-2)^2}$	-0,5	-0,5087	0,8171
$\mathbf{y(k-2)}$	-0,25	-0,2362	0,0529
$\mathbf{y(k-1)}$	0,7	0,6538	0,0291
$\mathbf{y(k-1)y(k-2)}$	0,3	0,2878	0,0720

Como a obtenção da estrutura correta do sistema \mathcal{S}_3 a partir do critério ERR_1 só foi possível após a aplicação da técnica de predição de dois passos à frente, pode-se afirmar que o ERR_2 , neste caso, se mostrou útil ao ser empregado como uma ferramenta auxiliar. Além disso, vale a pena observar que, diferentemente do que ocorreu na seção anterior, o ERR_2 se mostrou robusto mesmo utilizando-se máximos atrasos com valores maiores que os reais.

Também é interessante destacar que os termos espúrios presentes nos modelos de um passo à frente obtidos neste capítulo não foram os últimos a serem escolhidos pelo ERR_1 . Mesmo assim, verificou-se que, na maioria dos casos, ainda foi possível identificar tais termos por meio do ERR_2 . Isso mostra que tal critério pode realmente auxiliar na seleção de estrutura de modelos polinomiais no caso em que o sinal de entrada é relativamente lento.

5.3 Metodologia Revisada

O estudo de casos discutido neste capítulo mostrou que o critério ERR_2 pode ser útil ao ser empregado como uma ferramenta auxiliar na seleção de estrutura de modelos NARX polinomiais no caso em que o sinal de entrada é relativamente lento. Entretanto, verificou-se que tal critério é sensível à escolha dos máximos atrasos e também do número de termos. É importante ressaltar que tais sensibilidades também podem ser atribuídas ao critério ERR_1 .

Uma outra desvantagem do ERR_2 que também deve ser ressaltada é o exaustivo esforço empregado na obtenção de um modelo de um passo a partir de seu respectivo modelo de dois passos à frente⁵, que é necessário em virtude de a auto-consistência entre os critérios ERR_1 e ERR_2 ser utilizada a partir da comparação entre modelos de um passo à frente. Por esse motivo, decidiu-se, então, alterar a maneira como empregar a auto-consistência. Assim, em vez de comparar modelos de um passo à frente, serão, agora, comparados modelos de dois passos, que podem ser facilmente obtidos a partir de seus respectivos modelos de um passo à frente. Dessa forma, é necessário fazer algumas adaptações na metodologia inicialmente proposta na Seção 4.5, chegando-se, assim, a uma nova metodologia, cujos passos encontram-se listados a seguir. É importante ressaltar que as adaptações realizadas na metodologia não implicam alterações no fluxograma (Figura 4.1), o que o torna útil tanto para a metodologia apresentada na Seção 4.5 quanto para a nova metodologia aqui proposta.

1. Dividir os dados de identificação em janelas.

O uso dessa técnica, conhecida como janelas deslizantes, é comum em diversos problemas de controle (Mitra, 1992), processamento de sinais (Pierce e Plemmons, 1992) e seleção de estrutura de modelos a partir de estimação *on-line* (Luo e Billings, 1995; Ruano et al., 2006).

2. Definir o número de termos, os máximos atrasos nos sinais de entrada n_u e saída n_y e o grau de não-linearidade ℓ a serem utilizados durante o processo de identificação. Em seguida, utilizando o critério ERR_1 ,

⁵Utilizando-se, para isso, as relações apresentadas na Tabela 4.3.

deve-se estimar um modelo para cada uma das janelas de dados. Assim, o número de modelos obtidos será igual ao número de janelas formadas durante a divisão dos dados de identificação.

Sabe-se que, para diferentes realizações de dados correspondentes a um mesmo sistema, nem sempre obtém-se um mesmo modelo em todas as realizações, ou seja, os termos escolhidos por um determinado critério de seleção de estrutura para compor um modelo nem sempre são os mesmos em todos os modelos, mesmo quando os máximos atrasos nos sinais de entrada e saída e o grau de não-linearidade são mantidos constantes. Assim, para evitar que uma única estimação fosse priorizada, optou-se por trabalhar com um modelo moda obtido a partir dos modelos estimados em cada uma das janelas de dados.

3. Obter o modelo moda.

Primeiramente, é preciso mencionar que, neste trabalho, um modelo moda consiste num modelo formado pelos termos escolhidos com maior frequência pelo critério de seleção. Para isso, tal modelo é obtido utilizando-se o seguinte procedimento: primeiramente, verifica-se o primeiro termo de cada um dos modelos estimados nas janelas de dados e, em seguida, selecionam-se apenas os modelos que apresentam o termo encontrado no maior número de modelos. Considere, por exemplo, os seguintes modelos:

$$\begin{aligned}
 M_1 &= [y(k-1), y(k-2), u(k-1), y(k-3)], \\
 M_2 &= [y(k-1), y(k-2), u(k-1), y(k-3)], \\
 M_3 &= [y(k-1), y(k-2), u(k-1), y(k-3)], \\
 M_4 &= [y(k-1), y(k-2), u(k-1), u(k-2)], \\
 M_5 &= [y(k-1), y(k-2), u(k-1), u(k-2)], \\
 M_6 &= [y(k-1), u(k-1), y(k-2), y(k-4)], \\
 M_7 &= [y(k-2), y(k-1), u(k-1), u(k-3)],
 \end{aligned}$$

Analisando-se o primeiro termo de cada um desses modelos, verifica-se que o termo $y(k-1)$ se encontra em seis dos sete modelos. Por-

tanto, o modelo M_7 , cujo primeiro termo é $y(k-2)$, deve, então, ser descartado. Em seguida, deve-se verificar o segundo termo de cada um dos modelos pré-selecionados (no caso do exemplo, referem-se aos modelos M_1 , M_2 , M_3 , M_4 , M_5 e M_6) para que sejam selecionados apenas os modelos que apresentam o termo mais freqüente. Como cinco dos seis modelos apresentam o termo $y(k-2)$, o modelo M_6 , cujo segundo termo é $u(k-1)$, também deve ser eliminado. Esse processo continua até que o último termo seja analisado. Os modelos que chegam até o fim desse processo apresentam exatamente a mesma estrutura, que é, então, usada para compor o modelo moda. Assim, o modelo moda obtido neste exemplo deve ser formado pelos termos $y(k-1)$, $y(k-2)$, $u(k-1)$ e $y(k-3)$, que correspondem à estrutura dos modelos M_1 , M_2 e M_3 . O procedimento necessário para se chegar ao modelo moda foi implementado na função `modelo_moda`, que se encontra no Apêndice A.3 deste trabalho. Deve-se observar que, se, por acaso, o último termo do modelo M_3 fosse, por exemplo, $u(k-4)$ em vez de $y(k-3)$, haveria, neste caso, dois modelos moda, sendo que um deles apresentaria a estrutura dos modelos M_1 e M_2 e o outro, a estrutura dos modelos M_4 e M_5 . Portanto, apesar de todos os modelos moda obtidos neste trabalho apresentarem uma única estrutura, a função `modelo_moda` pode, eventualmente, retornar dois ou mais modelos moda.

4. Determinar o número ideal de termos do modelo moda.

Para isso, serão empregados o critério de informação de Akaike (AIC) e a estatística APRESS, apresentados na Seção 3.2. Deve-se ressaltar que, ao aplicar esses métodos a um modelo moda, serão utilizadas somente as amostras responsáveis pela geração de tal modelo, ou seja, apenas aquelas usadas na obtenção dos modelos que apresentam a mesma estrutura do modelo moda.

5. Retornar ao item 2 utilizando o critério ERR_2 ao invés do ERR_1 e, em seguida, prosseguir com os itens 3 e 4, chegando-se, assim, a dois modelos moda: um obtido pelo critério ERR_1 e outro obtido pelo ERR_2 .

É importante ressaltar que o modelo moda obtido pelo ERR_2 corresponde a um modelo de dois passos à frente.

6. Determinar, por substituição analítica⁶, o modelo de dois passos à frente correspondente ao modelo moda obtido pelo critério ERR_1 a fim de que possa ser comparado com o modelo moda obtido pelo critério ERR_2 .

Dessa forma, a auto-consistência entre os critérios ERR_1 e ERR_2 é utilizada a partir da comparação entre modelos de dois passos à frente.

7. Identificar os termos comuns ao modelo moda obtido pelo critério ERR_2 e ao modelo de dois passos à frente correspondente ao modelo moda obtido pelo ERR_1 .
8. Utilizar, na formação do novo modelo de um passo à frente, apenas os termos do modelo moda obtido pelo critério ERR_1 responsáveis pela geração dos termos comuns ao modelo moda obtido pelo ERR_2 e ao modelo de dois passos à frente correspondente ao modelo moda obtido pelo critério ERR_1 na predição de dois passos à frente.

Como pode ser observado, na nova metodologia apresentada nesta seção, a auto-consistência entre os critérios ERR_1 e ERR_2 é utilizada a partir da comparação entre modelos de dois passos à frente. Assim, a obtenção de um modelo de um passo a partir de seu respectivo modelo de dois passos à frente, utilizando-se, para isso, as relações apresentadas na Tabela 4.3, não é mais necessária, evitando, assim, o exaustivo esforço empregado na realização de tal tarefa. Dessa forma, a restrição aos modelos de um passo à frente nos quais o regressor $y(k-1)$ apresenta grau de multiplicidade apenas igual a um pode ser, então, eliminada.

⁶Conforme mostrado na Seção 4.2.

5.4 Considerações Finais

Neste capítulo, verificou-se que o critério ERR_2 pode ser útil ao ser empregado como uma ferramenta auxiliar na seleção de estrutura de modelos NARX polinomiais no caso em que o sinal de entrada é relativamente lento. Além disso, com o auxílio de tal critério, foi possível identificar termos espúrios presentes em modelos de um passo à frente mesmo não sendo esses os últimos termos a serem escolhidos pelo ERR_1 .

Entretanto, foram observadas algumas desvantagens do critério ERR_2 , tais como: (i) sensibilidade à escolha dos máximos atrasos, (ii) sensibilidade à escolha do número de termos utilizado para compor o modelo de dois passos e (iii) exaustivo esforço empregado na obtenção de um modelo de um passo a partir de seu respectivo modelo de dois passos à frente utilizando-se as relações apresentadas na Tabela 4.3. É importante ressaltar que as desvantagens listadas nos itens (i) e (ii) também podem ser atribuídas ao critério ERR_1 .

Ao realizar algumas adaptações na metodologia inicialmente proposta na Seção 4.5, chegou-se, neste capítulo, a uma nova metodologia na qual a auto-consistência entre os critérios ERR_1 e ERR_2 é utilizada a partir da comparação entre modelos de dois passos à frente em vez de modelos de um passo. O uso dessa metodologia será apresentado no Capítulo 6, onde são utilizados dados experimentais de sistemas reais.

Estudo de Casos Experimentais

“A insatisfação é a principal motivadora do progresso.”

Thomas Alva Edison

6.1 Introdução

No Capítulo 5, foi apresentada uma nova metodologia na qual a auto-consistência entre os critérios ERR_1 e ERR_2 , como ferramenta auxiliar, é utilizada a partir da comparação entre modelos de dois passos à frente, e não mais de modelos de um passo, como no caso da metodologia proposta na Seção 4.5. Este capítulo mostra o uso dessa nova metodologia na seleção de estrutura de modelos NARX polinomiais no caso em que o sinal de entrada é relativamente lento utilizando-se, para isso, dados experimentais de sistemas reais.

6.2 Estudo de Casos

Neste capítulo, são utilizados dados coletados de dois sistemas distintos. O primeiro sistema, considerado na Seção 6.2.1, consiste num forno a gás (Box e Jenkins, 1976). O sistema apresentado na Seção 6.2.2, por sua vez, consiste num pequeno aquecedor elétrico (Cassini, 1999).

Como, no caso de sistemas reais, a estrutura não é conhecida, o desempenho dos critérios ERR_1 e ERR_2 , nesse caso, não pode ser avaliado como no Capítulo 5, ou seja, comparando-se a estrutura do sistema com a dos modelos obtidos por tais critérios. Em virtude disso, foi utilizado, neste capítulo, o método de validação dinâmica a partir de predição livre com o intuito de verificar o desempenho dos modelos na recuperação do comportamento dinâmico do sistema. A fim de quantificar a eficiência desses modelos em relação ao seu desempenho dinâmico, foi empregado o índice RMSE, apresentado na Seção 2.2.5. Para isso, foi necessário que as massas de dados utilizadas neste capítulo fossem divididas em duas partes: uma destinada ao processo de identificação e outra reservada para a etapa de validação.

Dessa forma, a massa de dados empregada na Seção 6.2.1, que contém 296 amostras, foi dividida de forma que suas 200 primeiras amostras fossem utilizadas no processo de identificação e as 96 restantes, na etapa de validação. Em seguida, para que se pudesse trabalhar com um modelo moda, as amostras referentes aos dados de identificação foram divididas em 51 janelas contendo 150 amostras cada (Figura 6.1).

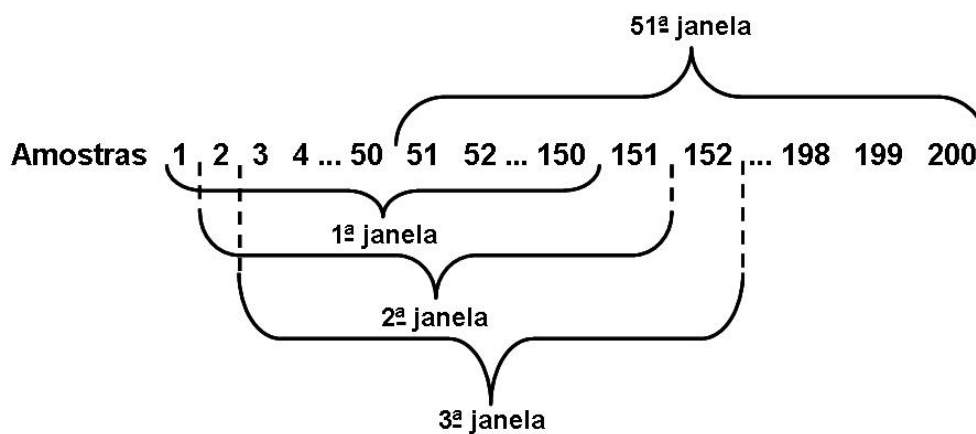


Figura 6.1: Representação de uma massa de dados dividida em 51 janelas contendo 150 amostras cada.

Das 1255 amostras presentes na massa de dados utilizada na Seção 6.2.2, as 999 primeiras foram empregadas no processo de identificação e as 256 restantes, na etapa de validação. Assim como no Capítulo 5, a fim de se trabalhar com um modelo moda, as amostras referentes aos dados de

identificação foram divididas em 500 janelas contendo 500 amostras cada (ver Figura 5.1).

6.2.1 Primeiro Caso

Os dados empregados nesta seção foram coletados de um sistema que consiste num forno a gás, cuja entrada é a vazão de gás na entrada do forno (ft^3/min) e a saída é a concentração de gás carbônico (%) (Box e Jenkins, 1976). A Figura 6.2 apresenta os dados normalizados utilizados no processo de identificação desse sistema. Ao observá-los, é possível perceber que um aumento na entrada provoca uma redução no sinal de saída. Na Figura 6.3, pode ser visualizado o resultado gráfico da aplicação da função de autocorrelação (FAC) ao sinal de entrada.

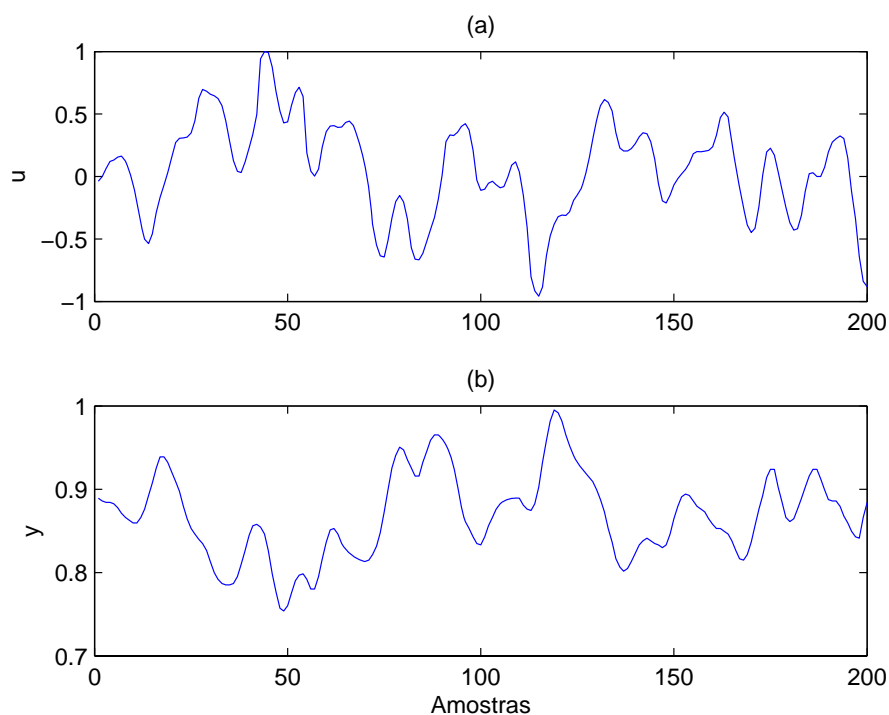


Figura 6.2: Dados utilizados na identificação de um forno a gás.

Sinais de (a) entrada (vazão de gás (ft^3/min)) e (b) saída (concentração de CO_2 (%)). (Fonte: Box e Jenkins (1976) - arquivo `boxjenk.dat`).

Ao se tentar identificar o sistema de onde foram coletados os dados mostrados na Figura 6.2, verificou-se que o uso de grau de não-linearidade e máximos atrasos maiores que três era desnecessário. Sendo assim, deu-se início ao processo de identificação a partir da técnica de predição de um passo à frente utilizando-se, para isso, modelos NARMAX com vinte termos lineares de ruído e grau de não-linearidade e máximos atrasos nos sinais de entrada e saída iguais a três a partir dos quais obteve-se um modelo moda.

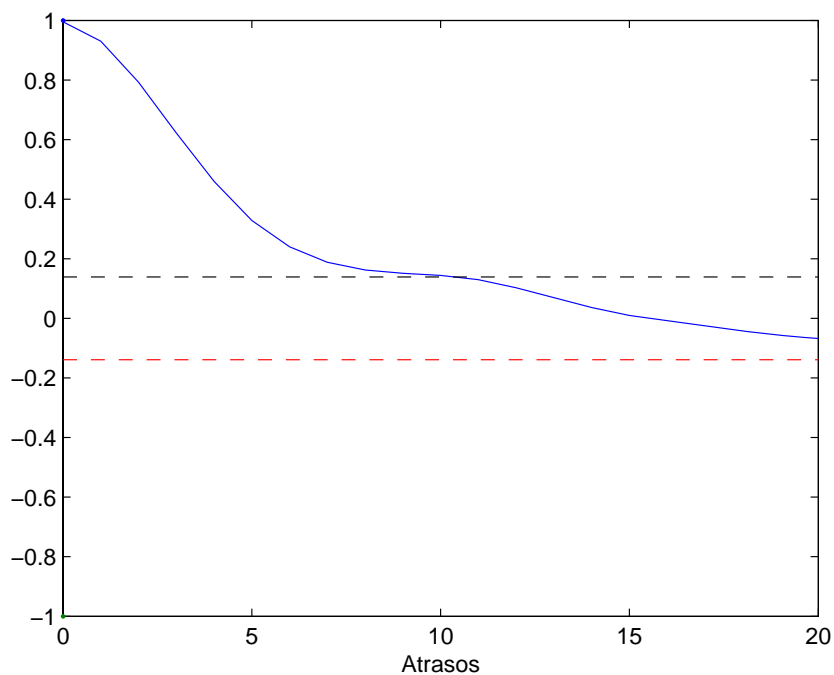


Figura 6.3: Resultado gráfico da aplicação da função FAC ao sinal de entrada apresentado na Figura 6.2.

O sinal de entrada não apresenta FAC semelhante à de um ruído branco, que é impulsiva.

Apesar de o critério de informação de Akaike sugerir um número de termos entre seis e nove (Figura 6.4), a Figura 6.5 mostra claramente que, de acordo com a estatística APRESS, o modelo moda obtido pelo critério ERR_1 deve ser formado por seis termos. Considerando, então, que

o número ideal de termos seja seis, tal modelo pode ser representado por:

$$y(k) = 1,0789y(k-1) - 0,3300y(k-2) + 0,0259y(k-3) + 0,0021y(k-3)u(k-1) + 0,1979 - 0,0354u(k-3) + \sum_{i=1}^{20} \hat{\theta}_i \xi(k-i) + \xi(k), \quad (6.1)$$

sendo que o somatório, que corresponde à parte média móvel (MA) do modelo, é utilizado, neste trabalho, apenas para reduzir o efeito de polarização durante o procedimento de estimação de parâmetros.

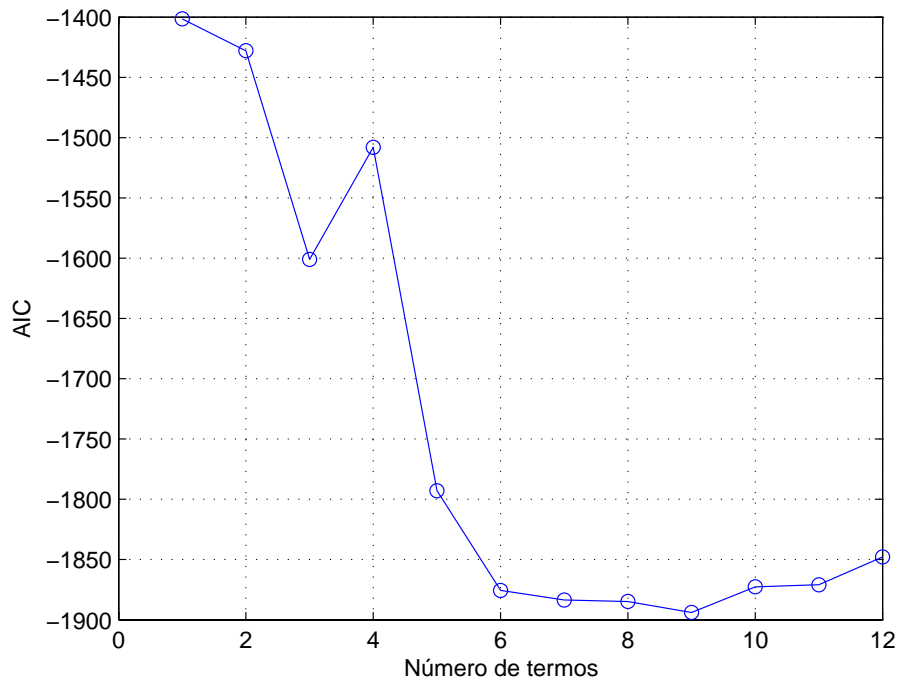


Figura 6.4: Resultado gráfico da aplicação do critério de informação de Akaike ao modelo moda obtido pelo critério ERR_1 a partir dos dados apresentados na Figura 6.2.

A Figura 6.6a mostra a predição livre do modelo (6.1) utilizando-se tanto os dados de identificação (200 primeiras amostras) quanto os de validação (96 últimas amostras). O RMSE obtido a partir da saída predita nesse caso foi de 0,2665. A Figura 6.6b, por sua vez, é uma ampliação

da Figura 6.6a delimitada apenas pela região correspondente aos dados de validação. Assim, o RMSE obtido considerando-se apenas as 96 últimas amostras da saída predita foi de 0,5646. Como pode ser observado, o modelo (6.1) apresentou dificuldade em acompanhar os dados nas regiões compreendidas pelo antepenúltimo e último picos. Uma possível explicação para isso poderia ser o suposto fato de não estar contido, nos dados de identificação, um comportamento semelhante ao apresentado nas regiões citadas ou, ainda, uma provável quebra de invariância temporal. Uma outra justificativa poderia estar relacionada com o pequeno número de amostras presentes na massa de dados empregada neste caso. É importante ressaltar que, neste trabalho, a parte MA dos modelos não é utilizada no cálculo das predições.

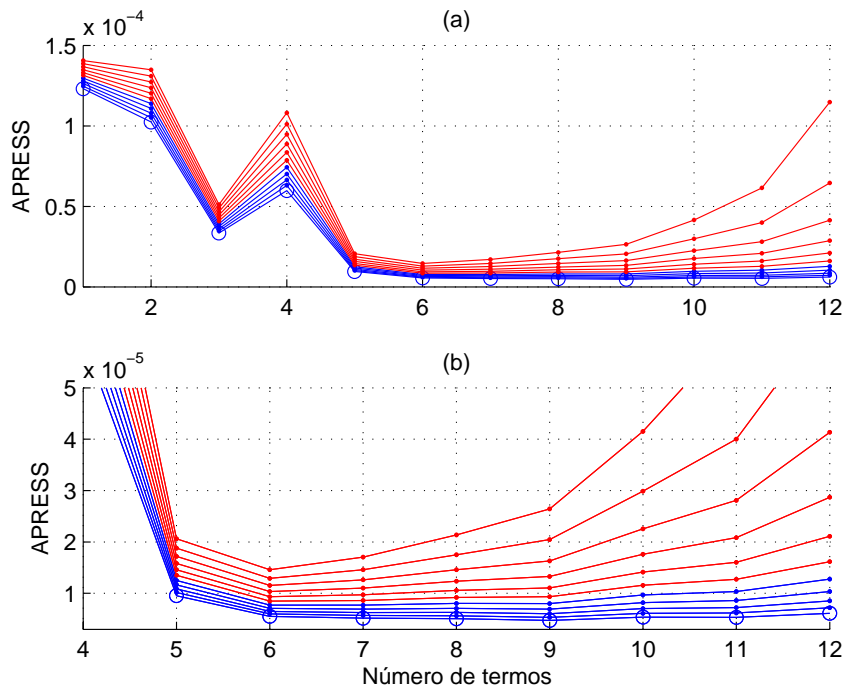


Figura 6.5: Resultado gráfico da aplicação da estatística APRESS ao modelo moda obtido pelo critério ERR_1 a partir dos dados apresentados na Figura 6.2.

Gráfico (a) em seu tamanho real e (b) em escala ampliada. As linhas, de baixo para cima, correspondem a $\alpha = 0, 1, \dots, 10$. A linha inferior (com círculos), correspondente a $\alpha = 0$, indica o MSE.

Como pode ser observado no modelo (6.1), o coeficiente do termo $y(k-3)u(k-1)$ é muito menor que o dos demais termos do modelo. Apesar disso, tal termo não deve ser eliminado baseando-se apenas nesta informação, pois, conforme mencionado no Capítulo 1, a técnica de se eliminar os termos cujos coeficientes são muito menores que os demais é, em geral, ineficiente no caso de sistemas não-lineares, especialmente quando há ruído presente nos dados.

Assim, com o intuito de eliminar possíveis termos espúrios do modelo (6.1), utilizou-se a técnica de predição de dois passos à frente a fim de que se pudesse empregar auto-consistência entre os critérios ERR_1 e ERR_2 . Para isso, assim como na predição de um passo, foram utilizados, neste caso, modelos NARMAX com vinte termos lineares de ruído e grau de não-linearidade e máximos atrasos nos sinais de entrada e saída iguais a três a partir dos quais obteve-se um modelo moda.

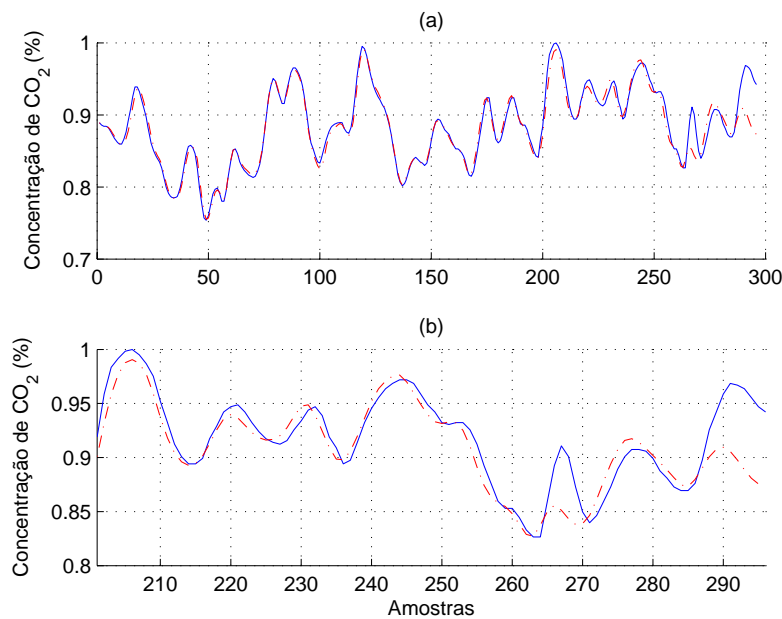


Figura 6.6: Validação dinâmica do modelo (6.1).

(a) Predição livre utilizando-se tanto os dados de identificação quanto os de validação ($RMSE = 0,2665$). (b) Ampliação da Figura 6.6a delimitada apenas pela região correspondente aos dados de validação ($RMSE = 0,5646$). O traço contínuo representa a saída medida, e o tracejado, a saída predita.

As Figuras 6.7 e 6.8 mostram, respectivamente, que, assim como na predição de um passo à frente, apesar de o critério de informação de Akaike sugerir um número de termos entre seis e nove, a estatística APRESS propõe que o número ideal de termos do modelo moda obtido pelo critério ERR_2 seja seis. Como os modelos de dois passos à frente apresentam número de termos maior que os de um passo e como o modelo (6.1) é formado por seis termos, optou-se, então, por considerar, na predição de dois passos à frente, um número de termos igual a nove, chegando-se, assim, ao seguinte modelo:

$$\begin{aligned}
 y(k+1) = & 0,6369y(k-1) - 0,0812y(k-2) - 0,0625y(k-3) + \\
 & 0,0078u(k)y(k-3)^2 + 0,4451 - 0,0402u(k-2) - \\
 & 0,0112u(k)^3 - 0,0397u(k-3) + 0,0059u(k)u(k-3)^2 + \\
 & \sum_{i=1}^{20} \hat{\theta}_i \xi(k-i) + \xi(k). \quad (6.2)
 \end{aligned}$$

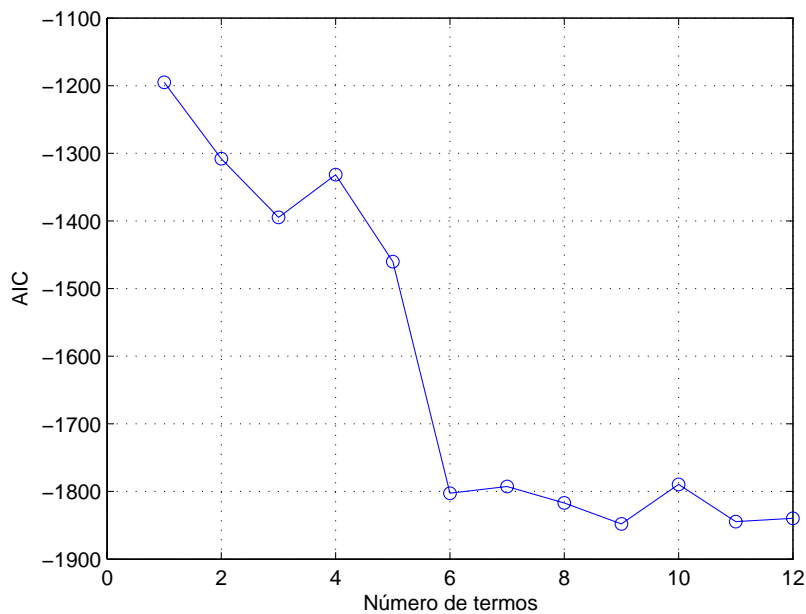


Figura 6.7: Resultado gráfico da aplicação do critério de informação de Akaike ao modelo moda obtido pelo critério ERR_2 a partir dos dados apresentados na Figura 6.2.

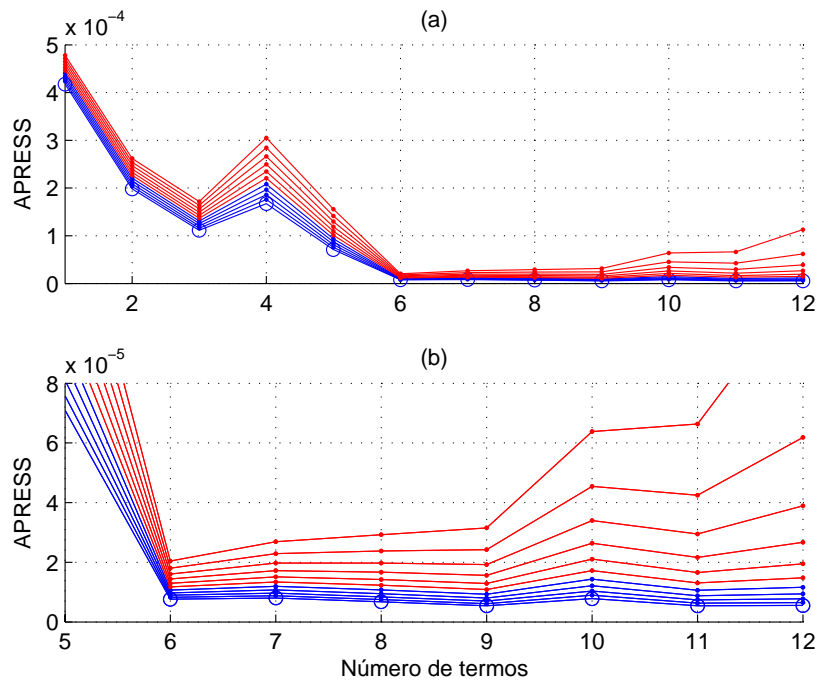


Figura 6.8: Resultado gráfico da aplicação da estatística APRESS ao modelo moda obtido pelo critério ERR_2 a partir dos dados apresentados na Figura 6.2.

Gráfico (a) em seu tamanho real e (b) em escala ampliada. As linhas, de baixo para cima, correspondem a $\alpha = 0, 1, \dots, 10$. A linha inferior (com círculos), correspondente a $\alpha = 0$, indica o MSE.

Seguindo a nova metodologia proposta na Seção 5.3, é necessário, então, de acordo com o item 6, determinar o modelo de dois passos à frente correspondente ao modelo (6.1) a fim de que possa ser comparado com o modelo (6.2). Assim, representando-se o modelo (6.1) na forma de predição de dois passos à frente, tem-se:

$$\begin{aligned}
 y(k+1) = & 0,8340y(k-1) - 0,3301y(k-2) + 0,0279y(k-3) + \\
 & 0,0023y(k-3)u(k-1) + 0,4114 - 0,0382u(k-3) + \\
 & 0,0021u(k)y(k-2) - 0,0354u(k-2). \quad (6.3)
 \end{aligned}$$

Comparando-se o modelo acima com o da equação (6.2), verifica-se que os termos comuns a esses modelos são: $y(k-1)$, $y(k-2)$, $y(k-3)$,

constante, $u(k-3)$ e $u(k-2)$. Utilizando-se, então, apenas os termos do modelo (6.1) responsáveis pela geração de tais termos comuns na predição de dois passos à frente, um novo modelo, apresentado na equação (6.4), foi estimado a partir do critério ERR_1 .

$$y(k) = 1,1940y(k-1) - 0,4727y(k-2) + 0,0733y(k-3) + 0,1805 - 0,0310u(k-3) + \sum_{i=1}^{20} \hat{\theta}_i \xi(k-i) + \xi(k). \quad (6.4)$$

Ao se comparar o modelo (6.4) com o modelo da equação (6.1), é possível perceber que a utilização de auto-consistência entre os critérios ERR_1 e ERR_2 como método de seleção de estrutura sugeriu a eliminação do termo $y(k-3)u(k-1)$.

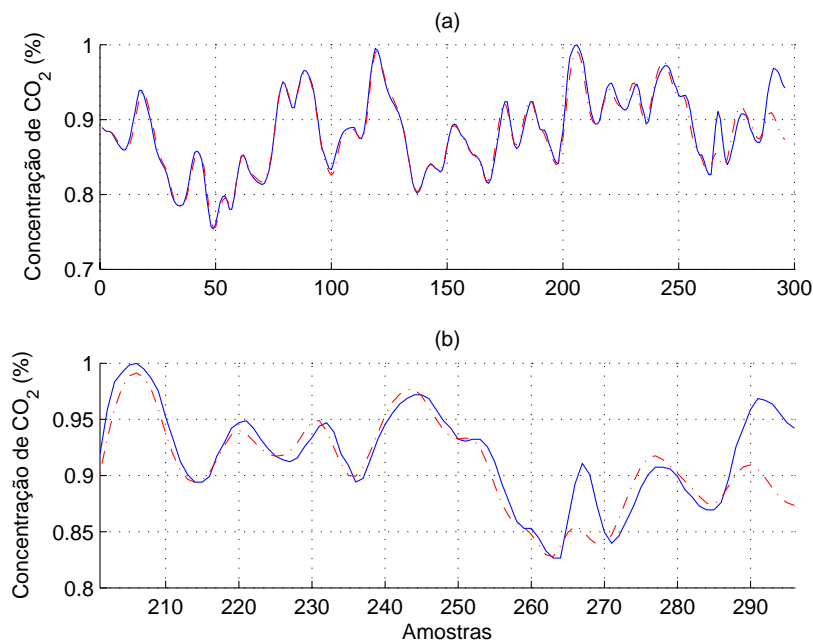


Figura 6.9: Validação dinâmica do modelo (6.4).

(a) Predição livre utilizando-se tanto os dados de identificação quanto os de validação ($RMSE = 0,2650$). (b) Ampliação da Figura 6.9a delimitada apenas pela região correspondente aos dados de validação ($RMSE = 0,5630$). O traço contínuo representa a saída medida, e o tracejado, a saída predita.

Na predição livre do modelo (6.4) a partir das 296 amostras presentes na massa de dados utilizada neste caso (Figura 6.9a), o RMSE foi de 0,2650. Por outro lado, ao se considerar apenas as 96 últimas amostras da saída predita (Figura 6.9b), referentes aos dados de validação, obteve-se $\text{RMSE} = 0,5630$. Como os valores de RMSE obtidos pelos modelos (6.1) e (6.4) são praticamente idênticos, pode-se afirmar que o desempenho dinâmico de tais modelos é equivalente. Isso mostra que, neste caso, a utilização de um modelo com grau de não-linearidade igual ou maior que dois é desnecessária. Portanto, verifica-se que o termo $y(k-3)u(k-1)$ pode, de fato, ser eliminado do modelo de um passo à frente.

A fim de verificar se mais algum termo, além do $y(k-3)u(k-1)$, poderia ser descartado, foram estimados outros cinco modelos, sendo cada um deles formado por cinco dos seis termos presentes no modelo (6.1). Após obter a predição livre de tais modelos, calculou-se o índice RMSE correspondente a cada um deles considerando-se, para isso, apenas as amostras referentes aos dados de validação. A comparação dos valores de RMSE obtidos por tais modelos é mostrada na Figura 6.10. É importante mencionar que a linha contínua se refere ao valor de RMSE correspondente ao modelo (6.1) ($\text{RMSE} = 0,5646$). Além disso, ressalta-se que o modelo m1, formado por todos os termos do modelo (6.1) com exceção do $y(k-1)$, se mostrou instável na predição livre. De acordo com a Figura 6.10, é possível perceber que os modelos m5 e m6, compostos, respectivamente, pelos termos $y(k-1)$, $y(k-2)$, $y(k-3)$, $y(k-3)u(k-1)$, $u(k-3)$ e $y(k-1)$, $y(k-2)$, $y(k-3)$, $y(k-3)u(k-1)$, constante, apresentam valores de RMSE muito maiores que 0,5646. Assim, pode-se afirmar que a constante e o $u(k-3)$, ausentes, respectivamente, em m5 e m6, não devem ser eliminados do modelo (6.1), já que mostraram ser importantes para o seu desempenho dinâmico. O modelo m2, que é formado por todos os termos do modelo (6.1) com exceção do $y(k-2)$, apresenta valor de RMSE pouco maior que 0,5646. O único modelo cujo RMSE encontra-se abaixo da linha contínua é o m4, que corresponde ao modelo da equação (6.4). Dessa forma, fica claro que o termo $y(k-3)u(k-1)$ pode ser, de fato, descartado do modelo (6.1). Entretanto, como o RMSE obtido pelo modelo m3 é muito próximo de 0,5646 e como tal modelo é constituído por todos os termos do modelo (6.1) com exceção do $y(k-3)$, pode-se afirmar que tal termo também pode ser excluído.

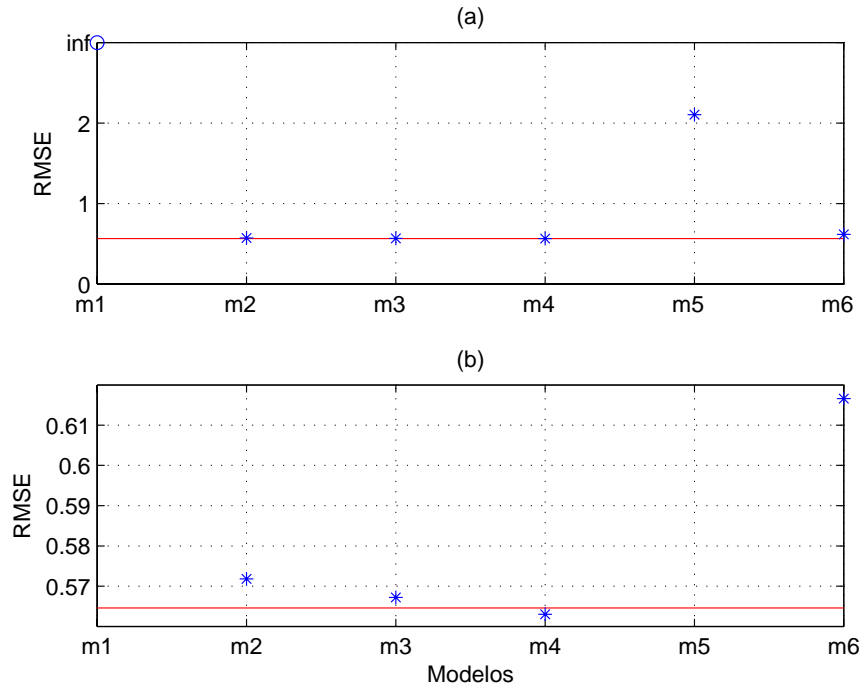


Figura 6.10: Comparação dos valores de RMSE correspondentes a modelos formados por termos presentes no modelo (6.1).

Gráfico (a) em seu tamanho real e (b) em escala ampliada. A linha contínua corresponde ao valor de RMSE obtido pelo modelo (6.1) ($\text{RMSE} = 0,5646$). Os asteriscos representam os valores de RMSE correspondentes aos modelos m2 ($\text{RMSE} = 0,5718$), m3 ($\text{RMSE} = 0,5672$), m4 (modelo (6.4)) ($\text{RMSE} = 0,5630$), m5 ($\text{RMSE} = 2,1033$) e m6 ($\text{RMSE} = 0,6166$). O círculo indica que o modelo m1 é instável.

Assim, descartando-se os termos $y(k-3)u(k-1)$ e $y(k-3)$, um outro modelo foi, então, estimado considerando-se apenas os termos $y(k-1)$, $y(k-2)$, constante e $u(k-3)$. O índice RMSE obtido por tal modelo foi de 0,5669, que é praticamente idêntico ao valor correspondente ao modelo (6.1). Portanto, pode-se concluir que a ausência dos termos $y(k-3)u(k-1)$ e $y(k-3)$ não prejudicou o desempenho dinâmico do modelo (6.1), podendo, então, ser eliminados de tal modelo. Apesar de o uso de auto-consistência entre os critérios ERR_1 e ERR_2 não ter sugerido a exclusão do termo $y(k-3)$, ainda é possível afirmar que a metodologia proposta neste trabalho contribuiu para a obtenção de um modelo mais compacto (modelo (6.4)).

6.2.2 Segundo Caso

Os dados utilizados nesta seção foram coletados de uma planta piloto que consiste de um pequeno aquecedor elétrico, cuja entrada é a tensão elétrica aplicada a um ferro de solda (p.u.) e a saída é a tensão nos terminais de um termopar conectado a um amplificador de instrumentação (p.u.) (Cassini, 1999). Esse sistema também é constituído de um ventilador que, quando acionado, altera a dinâmica do mesmo. Na Figura 6.11, são apresentados os dados normalizados, coletados com o ventilador ligado, utilizados no processo de identificação desse sistema. A Figura 6.12 mostra o resultado gráfico da aplicação da função de autocorrelação (FAC) ao sinal de entrada.

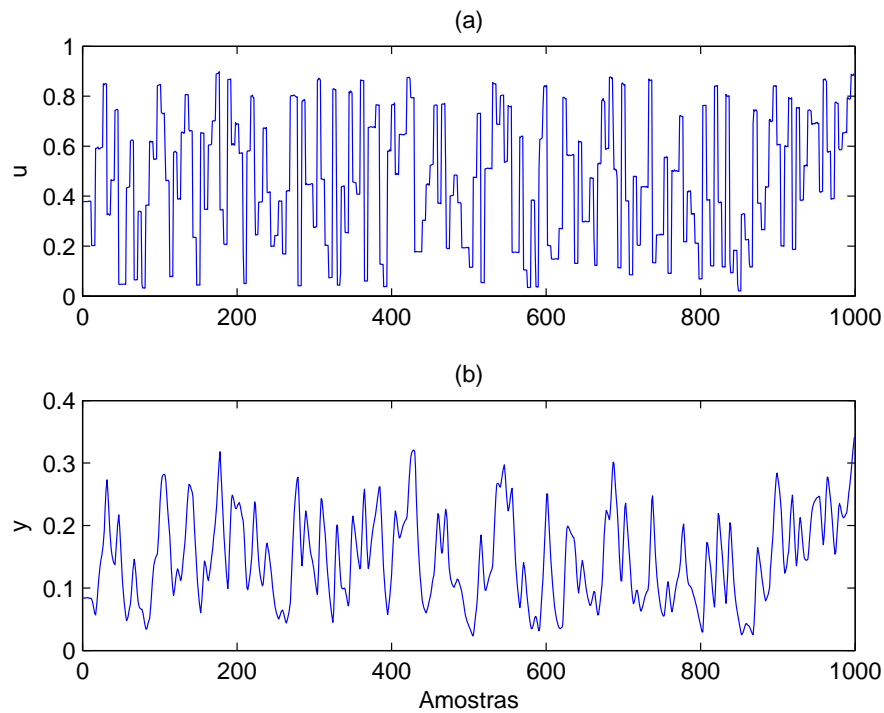


Figura 6.11: Dados utilizados na identificação de um pequeno aquecedor elétrico.

Sinais de (a) entrada (tensão elétrica aplicada a um ferro de solda (p.u.)) e (b) saída (tensão nos terminais de um termopar (p.u.)). (Fonte: Cassini (1999) - arquivo `din3.dat`).

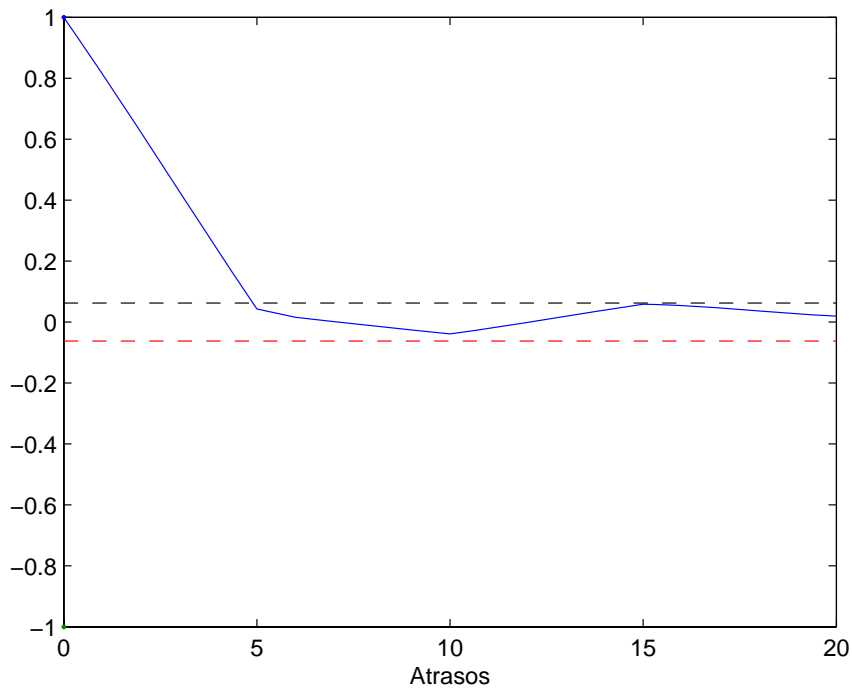


Figura 6.12: Resultado gráfico da aplicação da função FAC ao sinal de entrada apresentado na Figura 6.11.

O sinal de entrada não apresenta FAC semelhante à de um ruído branco, que é impulsiva.

De acordo com Cassini (1999), a utilização de grau de não-linearidade maior que dois na identificação do sistema aqui considerado é desnecessária. Como foi verificado que o uso de máximos atrasos maiores que três também não era necessário, a técnica de predição de um passo à frente foi, então, empregada utilizando-se, para isso, modelos NARMAX com cinco termos lineares de ruído, grau de não-linearidade igual a dois e máximos atrasos nos sinais de entrada e saída iguais a três a partir dos quais obteve-se um modelo moda.

Como pode ser visto na Figura 6.13, o critério de informação de Akaike propõe um número de termos entre seis e oito. A estatística APRESS, por sua vez, também parece indicar que o modelo moda obtido pelo critério ERR_1 possa ser formado por seis ou oito termos (Figura 6.14). Ao considerar o número de termos sendo igual a oito, chegou-se ao seguinte modelo:

$$\begin{aligned}
 y(k) = & 1,4879y(k-1) + 0,0413u(k-1)^2 - 0,5637y(k-2) - \\
 & 0,5207y(k-1)^2 + 0,4603y(k-1)y(k-3) + \\
 & 0,0244y(k-1)u(k-1) + 0,0012 - 0,0032u(k-2) + \\
 & \sum_{i=1}^5 \hat{\theta}_i \xi(k-i) + \xi(k).
 \end{aligned} \tag{6.5}$$

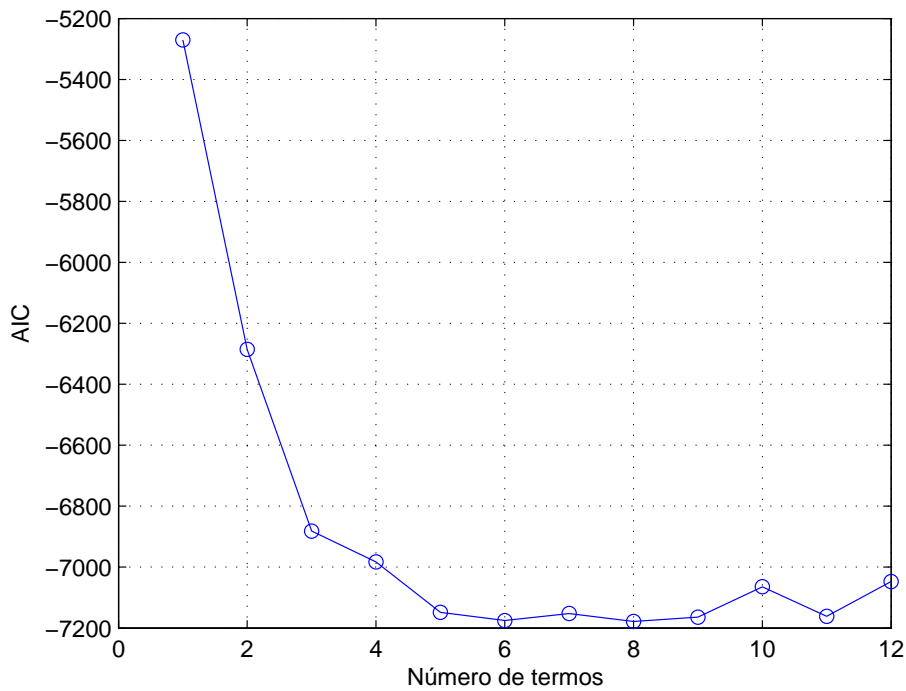


Figura 6.13: Resultado gráfico da aplicação do critério de informação de Akaike ao modelo moda obtido pelo critério ERR_1 a partir dos dados apresentados na Figura 6.11.

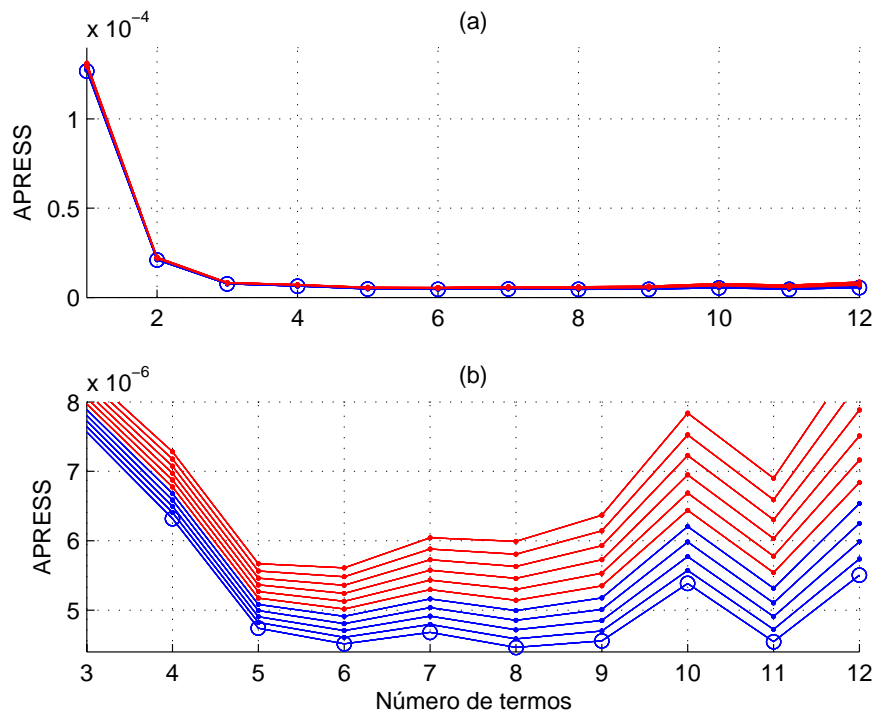


Figura 6.14: Resultado gráfico da aplicação da estatística APRESS ao modelo moda obtido pelo critério ERR_1 a partir dos dados apresentados na Figura 6.11.

Gráfico (a) em seu tamanho real e (b) em escala ampliada. As linhas, de baixo para cima, correspondem a $\alpha = 0, 1, \dots, 10$. A linha inferior (com círculos), correspondente a $\alpha = 0$, indica o MSE.

A Figura 6.15a mostra a predição livre do modelo (6.5) utilizando-se as 1255 amostras presentes na massa de dados empregada neste caso ($RMSE = 0,0809$). Ao se considerar apenas as amostras referentes aos dados de validação, obteve-se $RMSE = 0,0900$ (Figura 6.15b). Como pode ser visualizado, o modelo (6.5) foi capaz de acompanhar os dados de forma satisfatória, ou seja, seu desempenho dinâmico se mostrou muito semelhante ao do sistema. Entretanto, deve-se observar que os termos $y(k-1)^2$ e $y(k-1)y(k-3)$, presentes no modelo (6.5), são espúrios, pois, num processo térmico como o considerado nesta seção, não pode haver dois pontos fixos para uma mesma entrada.

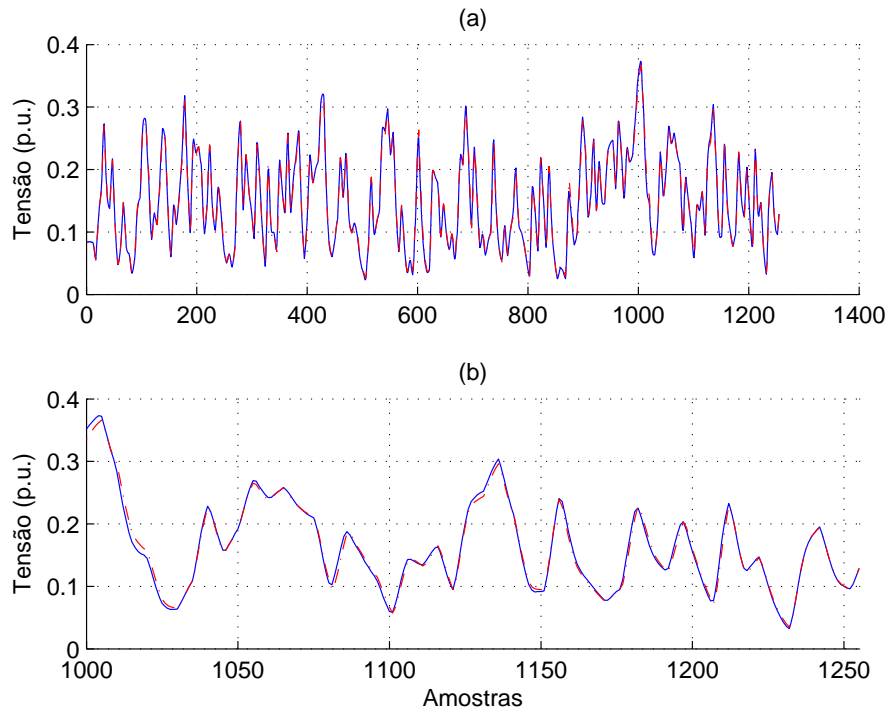


Figura 6.15: Validação dinâmica do modelo (6.5).

(a) Predição livre utilizando-se tanto os dados de identificação quanto os de validação ($RMSE = 0,0809$). (b) Ampliação da Figura 6.15a delimitada apenas pela região correspondente aos dados de validação ($RMSE = 0,0900$). O traço contínuo representa a saída medida, e o tracejado, a saída predita.

Desta forma, assim como na seção anterior, utilizou-se a técnica de predição de dois passos à frente na tentativa de eliminar termos espúrios do modelo (6.5). Como, de acordo com a nova metodologia proposta na Seção 5.3, o modelo de dois passos à frente correspondente ao da equação (6.5) deve ser comparado com o modelo moda a ser obtido pelo critério ERR_2 , é necessário ressaltar que a presença do termo $y(k-1)^2$ no modelo (6.5) é responsável por gerar, em seu respectivo modelo de dois passos (equação (6.6)), termos com grau de não-linearidade igual a quatro. Em virtude disso, para que o modelo moda obtido pela técnica de predição de dois passos à frente possa vir a apresentar grau de não-linearidade igual a quatro, foram utilizados, neste caso, modelos NARMAX com cinco termos

lineares de ruído, máximos atrasos nos sinais de entrada e saída iguais a três (como na predição de um passo à frente) e grau de não-linearidade igual a quatro a partir dos quais obteve-se um modelo moda.

$$\begin{aligned}
y(k+1) = & 0,0112u(k)y(k-1)y(k-3) + 0,6843y(k-1)y(k-3) + \\
& 0,8131 \times 10^{-4}y(k-1)u(k-1)u(k-2) + 0,0413u(k)^2 + \\
& 0,0256y(k-1)y(k-2)u(k-1) + 0,0363y(k-1)u(k-1) - \\
& 0,0117y(k-1)^2y(k-3)u(k-1) + 1,6483y(k-1) + \\
& 0,4821y(k-1)y(k-2)y(k-3) - 0,0032u(k-1) - \\
& 0,0198y(k-1)y(k-3)u(k-1)^2 - 0,8375y(k-2) + \\
& 0,0015y(k-1)y(k-3)u(k-2) - 0,0048u(k-2) + \\
& 0,5954 \times 10^{-3}u(k)y(k-1)u(k-1) + 0,0614u(k-1)^2 - \\
& 1,9268y(k-1)^2 + 0,2928 \times 10^{-4}u(k) + 0,8068y(k-1)^3 - \\
& 0,8882 \times 10^{-3}u(k-1)^4 - 0,0640y(k-1)u(k-1)^2 + \\
& 1,5583y(k-1)y(k-2) - 0,7132y(k-1)^2y(k-3) - \\
& 0,0378y(k-1)^2u(k-1) + 0,0050y(k-1)u(k-2) + \\
& 0,0433u(k-1)^2y(k-2) + 0,0221u(k-1)^2y(k-1)^2 - \\
& 0,0010y(k-1)u(k-1)^3 - 0,5453y(k-1)^2y(k-2) + \\
& 0,1376 \times 10^{-3}u(k-1)^2u(k-2) - 0,4249y(k-2)^2 - \\
& 0,0034y(k-2)u(k-2) + 0,2496y(k-1)^3y(k-3) + \\
& 0,0132y(k-1)^3u(k-1) - 0,0017y(k-1)^2u(k-2) - \\
& 0,1103y(k-1)^2y(k-3)^2 - 0,5332 \times 10^{-5}u(k-2)^2 + \\
& 0,0363u(k)y(k-1) + 0,0010u(k)u(k-1)^2 - \\
& 0,0138u(k)y(k-2) - 0,0127u(k)y(k-1)^2 - \\
& 0,7808 \times 10^{-4}u(k)u(k-2) - 0,1412y(k-1)^4 + \\
& 0,0030.
\end{aligned} \tag{6.6}$$

Analisando-se as Figuras 6.16 e 6.17, verifica-se que tanto o critério de informação de Akaike quanto a estatística APRESS parecem sugerir um número de termos entre sete e onze. Considerando-se, então, que o número ideal de termos do modelo moda obtido pelo critério ERR_2 seja onze, tem-se:

$$\begin{aligned}
y(k+1) = & 1,0532y(k-1) + 0,0118u(k)u(k-1) - 0,2824y(k-2) + \\
& 0,0472u(k)^2 + 0,0496u(k-1)^2 - \\
& 0,1913y(k-1)u(k-2)u(k-3)^2 + 0,0070u(k-2) + \\
& 0,3122u(k)y(k-3)u(k-3)^2 - \\
& 0,0872u(k)y(k-2)u(k-3) + 0,0360u(k-2)^2u(k-3) - \\
& 0,1230u(k)^2y(k-1)u(k-3) + \sum_{i=1}^5 \hat{\theta}_i \xi(k-i) + \xi(k). \quad (6.7)
\end{aligned}$$

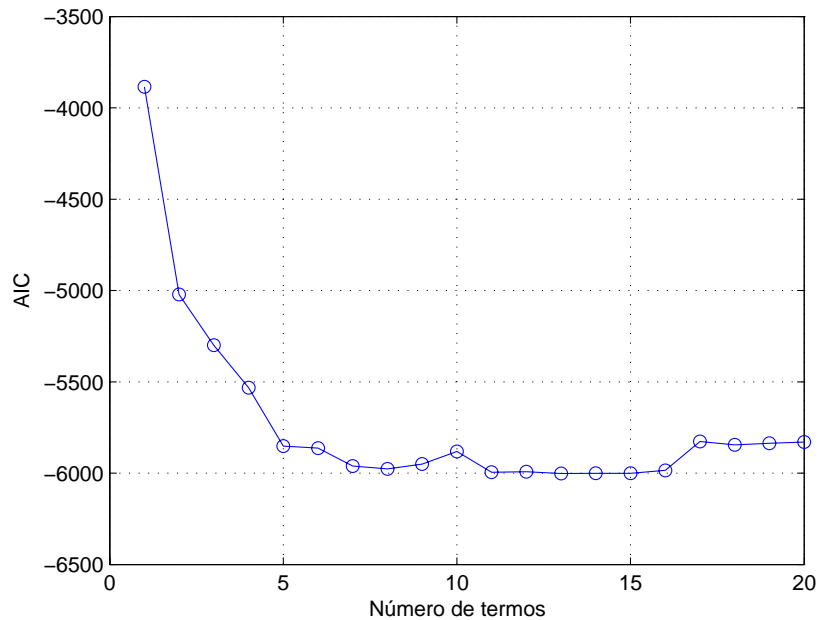


Figura 6.16: Resultado gráfico da aplicação do critério de informação de Akaike ao modelo moda obtido pelo critério ERR_2 a partir dos dados apresentados na Figura 6.11.

Comparando-se, então, o modelo (6.6) com o da equação (6.7), verifica-se que os termos comuns a esses modelos são: $y(k-1)$, $y(k-2)$, $u(k)^2$, $u(k-1)^2$ e $u(k-2)$. Como o próximo passo, de acordo com a nova metodologia proposta na Seção 5.3, consiste em identificar os termos do modelo (6.5) responsáveis pela geração de tais termos comuns na predição de dois passos à frente para que possam ser utilizados na obtenção do

novo modelo de um passo à frente, deve-se ressaltar que o termo $u(k-2)$ (modelo (6.5)), que, juntamente com o termo $y(k-1)$, é responsável por gerar o termo $u(k-2)$ na predição de dois passos, também implica a geração do termo $u(k-1)$, que, por sua vez, não pertence ao conjunto de termos comuns. Portanto, o termo $u(k-2)$ não pode ser empregado na estimação do novo modelo de um passo à frente, restando, assim, apenas os termos $y(k-1)$, $y(k-2)$ e $u(k-1)^2$ do modelo (6.5), responsáveis pela geração dos termos $y(k-1)$, $y(k-2)$, $u(k)^2$ e $u(k-1)^2$ na predição de dois passos à frente. Dessa forma, obteve-se o seguinte modelo:

$$y(k) = 1,2563y(k-1) + 0,0475u(k-1)^2 - 0,3466y(k-2) + \sum_{i=1}^5 \hat{\theta}_i \xi(k-i) + \xi(k). \quad (6.8)$$

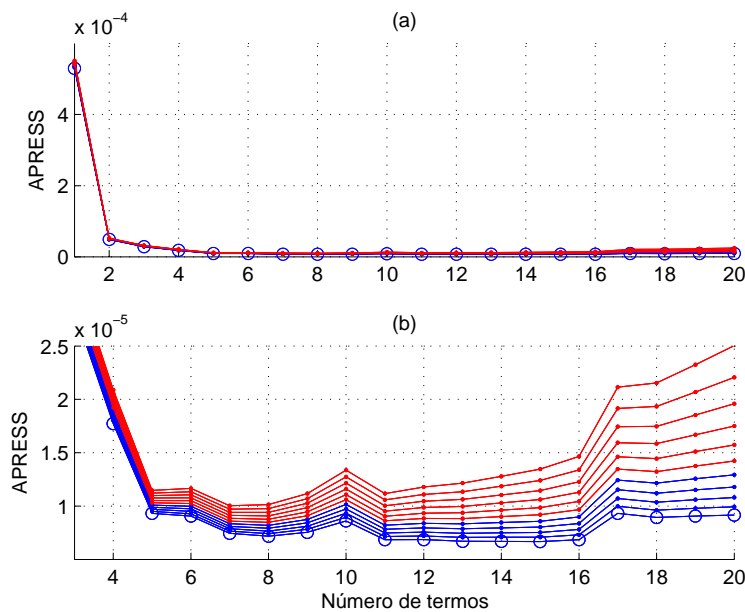


Figura 6.17: Resultado gráfico da aplicação da estatística APRESS ao modelo moda obtido pelo critério ERR_2 a partir dos dados apresentados na Figura 6.11.

Gráfico (a) em seu tamanho real e (b) em escala ampliada. As linhas, de baixo para cima, correspondem a $\alpha = 0, 1, \dots, 10$. A linha inferior (com círculos), correspondente a $\alpha = 0$, indica o MSE.

O modelo (6.8) mostra que a utilização de auto-consistência entre os critérios ERR_1 e ERR_2 como método de seleção de estrutura sugeriu corretamente a eliminação dos termos pertencentes ao agrupamento Ω_{y^2} .

Na predição livre do modelo (6.8) a partir das 1255 amostras presentes na massa de dados utilizada neste caso (Figura 6.18a), obteve-se $RMSE = 0,0948$. Considerando-se apenas as 256 últimas amostras da saída predita (Figura 6.18b), referentes aos dados de validação, o $RMSE$ foi de 0,0992. É importante observar que, apesar de o modelo (6.8) ser formado por apenas três termos, seu desempenho dinâmico se mostrou muito semelhante ao do sistema.

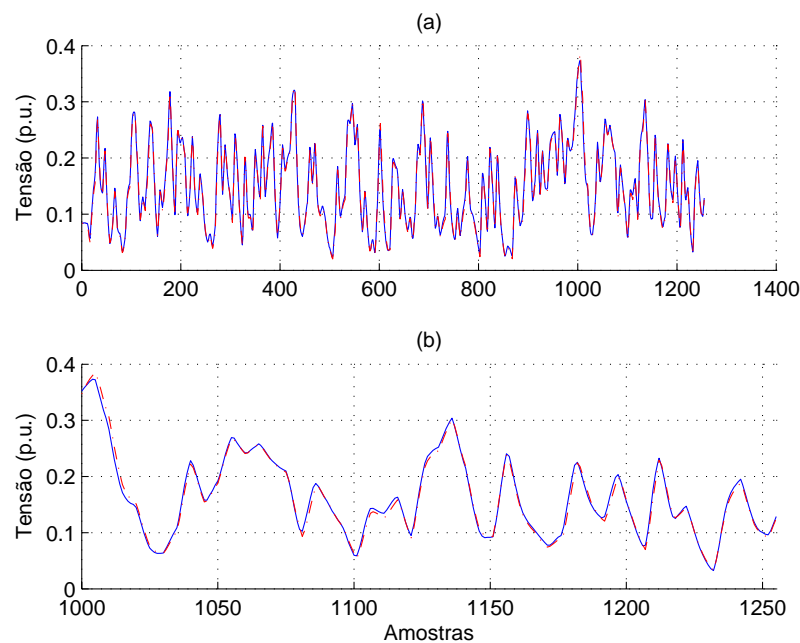


Figura 6.18: Validação dinâmica do modelo (6.8).

(a) Predição livre utilizando-se tanto os dados de identificação quanto os de validação ($RMSE = 0,0948$). (b) Ampliação da Figura 6.18a delimitada apenas pela região correspondente aos dados de validação ($RMSE = 0,0992$). O traço contínuo representa a saída medida, e o tracejado, a saída predita.

Apesar de o termo $u(k-1)$ não ser comum aos modelos (6.6) e (6.7), é provável que o fato de o termo $u(k-2)$ pertencer a ambos os modelos indique que este termo não deva ser eliminado do modelo (6.5). Em virtude

disso, um outro modelo, apresentado na equação (6.9), foi, então, estimado considerando-se os termos $y(k-1)$, $y(k-2)$, $u(k-1)^2$ e $u(k-2)$. O índice RMSE obtido nesse caso foi de 0,0753 (Figura 6.19), ou seja, menor que o RMSE correspondente ao modelo (6.5) (0,0900). Isso mostra que a eliminação dos termos $y(k-1)^2$, $y(k-1)y(k-3)$, constante e $y(k-1)u(k-1)$ do modelo (6.5) e a permanência dos termos $y(k-1)$, $y(k-2)$, $u(k-1)^2$ e $u(k-2)$ em tal modelo, resultantes da utilização de auto-consistência entre os critérios ERR_1 e ERR_2 , levaram a um modelo mais compacto e com desempenho dinâmico mais próximo do sistema (menor valor de RMSE).

$$y(k) = 1,1936y(k-1) + 0,0457u(k-1)^2 - 0,2988y(k-2) + 0,0062u(k-2) + \sum_{i=1}^5 \hat{\theta}_i \xi(k-i) + \xi(k). \quad (6.9)$$

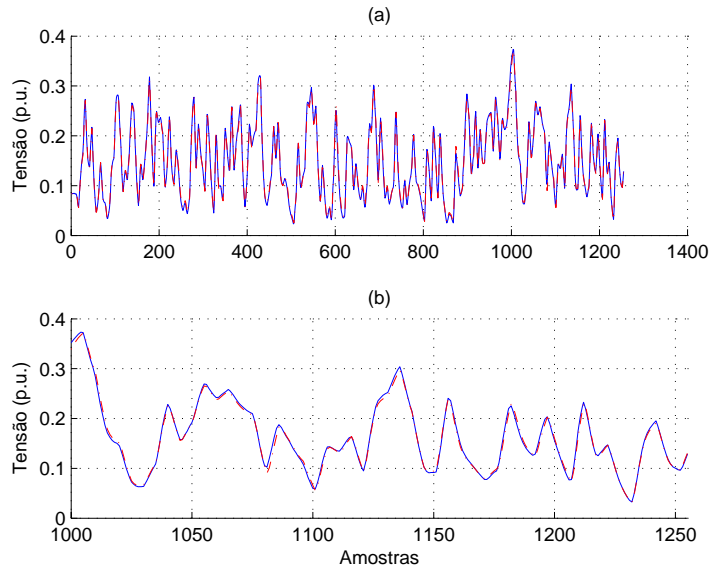


Figura 6.19: Validação dinâmica do modelo (6.9).

(a) Predição livre utilizando-se tanto os dados de identificação quanto os de validação (RMSE = 0,0755). (b) Ampliação da Figura 6.19a delimitada apenas pela região correspondente aos dados de validação (RMSE = 0,0753). O traço contínuo representa a saída medida, e o tracejado, a saída predita.

6.3 Considerações Finais

A aplicação da nova metodologia, proposta na Seção 5.3, na identificação dos sistemas considerados neste capítulo revelou que, ao utilizar auto-consistência entre os critérios ERR_1 e ERR_2 a partir da comparação entre modelos de dois passos à frente, é possível identificar e, conseqüentemente, eliminar termos espúrios, gerando, assim, modelos mais compactos. Isso sugere que o método proposto neste trabalho pode ser empregado como uma ferramenta auxiliar na seleção de estrutura de modelos NARX polinomiais.

Conclusões e Propostas Futuras

“Se eu vi mais longe, foi por estar de pé sobre ombros de gigantes.”

Isaac Newton

Diversos métodos têm sido propostos a fim de se tentar solucionar o problema da seleção de estrutura, que representa uma etapa extremamente importante no processo de identificação de sistemas. Entretanto, ressalta-se que cada um deles apresenta alguma desvantagem ou limitação.

Um critério amplamente utilizado na seleção de estrutura é o ERR (denominado, neste trabalho, de ERR_1), que se baseia na redução do erro de predição (de um passo à frente). Apesar de sua ampla aplicabilidade, tal critério pode escolher termos incorretos ou redundantes em condições não ideais de identificação, ou seja, quando os dados disponíveis não são adequados (superamostrados ou ruidosos) ou quando o sinal de entrada é relativamente lento.

A fim de investigar esse problema, foi proposto, neste trabalho, um critério baseado na redução do erro de predição de dois passos à frente (ERR_2) a fim de ser utilizado como uma ferramenta auxiliar na seleção de estrutura de modelos NARX polinomiais em um caso específico de condição não ideal de identificação, caso no qual o sinal de entrada é relativamente lento. A fim de identificar e eliminar possíveis termos espúrios selecionados pelo ERR_1 , foi empregado, neste trabalho, auto-consistência entre os critérios ERR_1 e ERR_2 , utilizando, a princípio, a comparação entre modelos de um passo à frente. Como a estrutura obtida pelo ERR_2

corresponde a um modelo de dois passos, foi preciso, então, que os modelos obtidos por tal critério fossem representados na forma de modelos de um passo à frente. Para isso, realizou-se um estudo a partir do qual foi possível determinar as relações entre os termos dos modelos de um e dois passos à frente para que pudessem ser empregadas na obtenção de um modelo de um passo a partir de seu respectivo modelo de dois passos à frente. Para isso, foram considerados apenas os modelos de um passo à frente nos quais o regressor $y(k-1)$ apresenta grau de multiplicidade igual a um.

O estudo de casos discutido no Capítulo 5 mostrou que o critério ERR_2 pode ser útil ao ser empregado como uma ferramenta auxiliar na seleção de estrutura de modelos NARX polinomiais no caso em que o sinal de entrada é relativamente lento. Além disso, verificou-se também que, com o auxílio de tal critério, foi possível identificar termos espúrios presentes em modelos de um passo à frente mesmo não sendo esses os últimos termos a serem escolhidos pelo ERR_1 .

Entretanto, observou-se que o critério ERR_2 é sensível à escolha do número de termos, pois, dependendo do número de termos selecionado, a obtenção do modelo de um passo à frente pode ser uma tarefa impossível. Além disso, constatou-se que tal critério também é sensível à escolha dos máximos atrasos. Deve-se ressaltar que tais sensibilidades também podem ser atribuídas ao critério ERR_1 . A fim de auxiliar na determinação dos valores de n_u e n_y a serem considerados durante o processo de identificação, chegou a ser empregado, neste trabalho, o método de falsos vizinhos (FNN). Entretanto, devido à sua subjetividade em virtude da sensibilidade à escolha de R , ϵ^y e τ , tal método não se mostrou uma ferramenta útil na estimação dos máximos atrasos e, por esse motivo, não foi considerado.

Uma outra desvantagem apresentada pelo ERR_2 foi o exaustivo esforço empregado na obtenção de um modelo de um passo a partir de seu respectivo modelo de dois passos à frente, que é necessário em virtude de a auto-consistência entre os critérios ERR_1 e ERR_2 ser utilizada a partir da comparação entre modelos de um passo à frente. Por isso, decidiu-se, então, alterar a maneira como empregar a auto-consistência. Assim, em vez de comparar modelos de um passo, passaram a ser comparados modelos de dois passos à frente, que podem ser facilmente obtidos a partir de seus respectivos modelos de um passo. Dessa forma, a restrição aos modelos de um passo à frente nos quais o regressor $y(k-1)$ apresenta grau de

multiplicidade apenas igual a um foi, então, eliminada.

Os resultados apresentados no Capítulo 6 mostraram que a utilização de auto-consistência entre os critérios ERR_1 e ERR_2 a partir da comparação entre modelos de dois passos à frente pode auxiliar na seleção de estrutura de modelos NARX polinomiais, gerando modelos mais compactos.

Desta forma, a fim de dar prosseguimento ao estudo iniciado neste trabalho, propõe-se, como proposta futura, investigar a sensibilidade do critério ERR_2 em relação à escolha dos máximos atrasos e do número de termos. Além disso, também seria interessante, como proposta de continuidade deste trabalho, verificar a aplicação do método de seleção de estrutura aqui proposto considerando-se outros casos de condições não ideais de identificação.

Referências Bibliográficas

- Abarbanel, H. D. I., Brown, R., Sidorowich, J. J., e Tsimring, L. S. (1993). The analysis of observed chaotic data in physical systems. *Reviews of Modern Physics*, 65(4):1331–1392.
- Aguirre, L. A. (1994a). Some remarks on structure selection for nonlinear models. *International Journal of 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 reconstructions. *Physics Letters A*, 203(2,3):88–94.
- Aguirre, L. A. (1997). On the structure of nonlinear polynomial models: higher order correlation functions, spectra and term clusters. *Circuits and Systems I – Fundamental Theory and Applications*, 44:450–453.
- Aguirre, L. A. (2004). *Introdução à identificação de sistemas – Técnicas lineares e não-lineares aplicadas a sistemas reais*. Editora UFMG, 2. ed. Belo Horizonte.
- Aguirre, L. A. e Billings, S. A. (1994). Validating identified nonlinear models with chaotic dynamics. *International Journal of Bifurcation and Chaos*, 4(1):109–125.
- Aguirre, L. A. e Billings, S. A. (1995a). Dynamical effects of overparametrization in nonlinear models. *Physica D*, 80(1,2):26–40.

- Aguirre, L. A. e Billings, S. A. (1995b). Improved structure selection for nonlinear models based on term clustering. *International Journal of Control*, 62(3):569–587.
- Aguirre, L. A., Corrêa, M. V., e Cassini, C. C. S. (2002). Nonlinearities in NARX polynomial models: representation and estimation. *IEE Proceedings - Control Theory and Applications*, 149(4):343–348.
- Aguirre, L. A., Donoso-Garcia, P. F., e Santos-Filho, R. (2000). Use of a priori information in the identification of global nonlinear models – A case study using a buck converter. *IEEE Transactions on Circuits and Systems – I: Fundamental Theory and Applications*, 47(7):1081–1085.
- Aguirre, L. A., Furtado, E. C., e Tôrres, L. A. B. (2006). Evaluation of dynamical models: dissipative synchronization and other techniques. *Physical Review E*, 74(6).
- Aguirre, L. A. e Jácome, C. R. F. (1998). Cluster analysis of NARMAX models for signal-dependent systems. In *IEE Proceedings – Control Theory Applications*, v. 145, p. 409–414.
- Aguirre, L. A. e Mendes, E. M. A. M. (1996). Global nonlinear polynomial models: structure, term clusters and fixed points. *International Journal of Bifurcation and Chaos*, 6(2):279–294.
- Aguirre, L. A., Rodrigues, G. G., e Jácome, C. R. F. (1998). Identificação de sistemas não-lineares utilizando modelos NARMAX polinomiais – uma revisão e novos resultados. *Controle & Automação*, 9(2):90–106.
- Akaike, H. (1970). Statistical predictor identification. *Annals of the Institute of Statistical Mathematics*, 22(2):203–217.
- Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19(6):716–723.
- Allen, D. M. (1974). The relationship between variable selection and data augmentation and a method for prediction. *Technometrics*, 16:125–127.
- Alves, G. B. (2004). *Propriedades das funções de base radiais aplicadas à identificação de sistemas dinâmicos não-lineares*. Belo Horizonte: Universidade Federal de Minas Gerais. (Dissertação de Mestrado, PPGEE).

- Amaral, G. F. V. (2001). *Uso de redes neurais e conhecimento a priori na identificação de sistemas dinâmicos não-lineares*. Belo Horizonte: Universidade Federal de Minas Gerais. (Dissertação de Mestrado, PPGEE).
- Barroso, M. F. S., Takahashi, R. H. C., e Aguirre, L. A. (2007). Multi-objective parameter estimation via minimal correlation criterion. *Journal of Process Control*, 17(4):321–332.
- Beligiannis, G. N., Skarlas, L. V., Likothanassis, S. D., e Perdikouri, K. G. (2005). Nonlinear model structure identification of complex biomedical data using a genetic-programming-based technique. *IEEE Transactions on Instrumentation and Measurement*, 54(6):2184–2190.
- Billings, S. A. (1980). Identification of nonlinear systems – A survey. *IEE Proceedings, Part D*, 127(6):272–285.
- Billings, S. A. e Aguirre, L. A. (1995). Effects of the sampling time and identification of nonlinear models. *International Journal of Bifurcation and Chaos*, 5(6):1541–1556.
- Billings, S. A. e Chen, S. (1989). Identification of nonlinear rational systems using a prediction-error estimation algorithm. *International Journal of Control*, 20(3):467–494.
- Billings, S. A., Chen, S., e Korenberg, M. J. (1989). Identification of MIMO nonlinear systems using a forward-regression orthogonal estimator. *International Journal of Control*, 49:2157–2189.
- Billings, S. A. e Coca, D. (1999). Discrete wavelet models for identification and qualitative analysis of chaotic systems. *International Journal of Bifurcation and Chaos*, 9(7):1263–1284.
- Billings, S. A. e Tao, Q. H. (1991). Model validity tests for nonlinear signal processing applications. *International Journal of Control*, 54(1):157–194.
- Billings, S. A. e Voon, W. S. F. (1983). Structure detection and model validity tests in the identification of nonlinear system. In *Proceedings of the Institution of Electronic Engineers*, v. 130, p. 193–199.

- Billings, S. A. e Voon, W. S. F. (1986). Correlation based model validity tests for nonlinear models. *International Journal of Control*, 44(1):235–244.
- Billings, S. A. e Wei, H. L. (2008). An adaptive orthogonal search algorithm for model subset selection and non-linear system identification. *International Journal of Control*, 81(5):714–724.
- Billings, S. A. e Zhu, Q. M. (1991). Rational model identification using an extended least-squares algorithm. *International Journal of Control*, 54(3):529–546.
- Billings, S. A. e Zhu, Q. M. (1994). Nonlinear model validation using correlation tests. *International Journal of Control*, 60(6):1107–1120.
- Bland, R. G., Goldfarb, D., e Todd, M. J. (1981). The ellipsoid method: a survey. *Operations Research*, 29:1039–1091.
- Bomberger, J. D. e Seborg, D. E. (1998). Determination of model order for NARX models directly from input-output data. *Journal of Process Control*, 8(5,6):459–468.
- Box, G. E. P. e Jenkins, G. M. (1976). *Time series analysis: forecasting and control*. Holden-Day, São Francisco.
- Braga, A. P., Carvalho, A. P. L. F., e Ludemir, T. B. (2000). *Redes neurais artificiais: teoria e aplicações*. Editora LTC.
- Campos, R. C. C. (2007). *Projeto e construção de planta piloto de neutralização de pH e proposta de metodologia para incorporação de informações auxiliares na identificação NARX racional*. Coronel Fabriciano: Centro Universitário do Leste de Minas Gerais – Unileste-MG. (Dissertação de Mestrado, PPGE).
- Casdagli, M. (1989). Nonlinear prediction of chaotic time series. *Physica D*, 35:335–356.
- Cassini, C. (1999). *Estimação recursiva de não-linearidades estáticas usando modelos NARMAX polinomiais*. Belo Horizonte: Universidade Federal de Minas Gerais. (Dissertação de Mestrado, PPGE).

- Chen, S. e Billings, S. A. (1989). Representation of non-linear systems: the NARMAX model. *International Journal of Control*, 49(3):1013–1032.
- Chen, S., Billings, S. A., e Luo, W. (1989). Orthogonal least squares methods and their application to nonlinear system identification. *International Journal of Control*, 50(5):1873–1896.
- Coelho, M. C. S. (2002). *Modelos de Hammerstein e de Wiener: conexões com modelos NARX e sua aplicação em identificação de sistemas não-lineares*. Belo Horizonte: Universidade Federal de Minas Gerais. (Dissertação de Mestrado, PPGEE).
- Corrêa, M. V. (1997). *Identificação de sistemas dinâmicos não-lineares utilizando modelos NARMAX racionais – Aplicação a sistemas reais*. Belo Horizonte: Universidade Federal de Minas Gerais. (Dissertação de Mestrado, PPGEE).
- Corrêa, M. V. (2001). *Identificação caixa-cinza de sistemas não-lineares utilizando representações NARMAX racionais e polinomiais*. Belo Horizonte: Universidade Federal de Minas Gerais. (Tese de Doutorado, PPGEE).
- Corrêa, M. V., Aguirre, L. A., e Saldanha, R. R. (2002). Using steady-state prior knowledge to constrain parameter estimates in nonlinear system identification. *IEEE Transactions on Circuits and Systems – I: Fundamental Theory and Applications*, 49(9):1376–1381.
- Elsner, J. B. (1992). Predicting time series using a neural network as a method of distinguishing chaos from noise. *Journal of Physics A*, 25:843–850.
- Fair, R. C. (1986). *Evaluating the predictive accuracy of models*. Handbook of econometrics, volume 3, capítulo 33. Elsevier Science, Amsterdã.
- Feil, B., Abonyi, J., e Szeifert, F. (2004). Model order selection of nonlinear input-output models – A clustering based approach. *Journal of Process Control*, 14(6):593–602.
- Fiedler-Ferrara, N. e Prado, C. P. C. (1994). *Caos: uma introdução*. Editora Edgar Blücher Ltda, São Paulo.

- Foss, B. A. e Johansen, T. A. (1992). A NARMAX model representation for adaptive control based on local models. *Modeling, Identification and Control*, 13(1):25–39.
- Gooijer, J. G., Abraham, B., Gould, A., e Robinson, L. (1985). Methods for determining the order of an autoregressive moving average process: a survey. *International Statistical Review*, 53(3):301–329.
- Graps, A. (1995). An introduction to wavelets. *IEEE Computational Science and Engineering*, 2(2).
- Gray, G. J., Murray-Smith, D. J., Li, Y., Sharman, K. C., e Weinbrenner, T. (1998). Nonlinear model structure identification using genetic programming. *Control Engineering Practice*, 6:1341–1352.
- Haber, R. e Unbehauen, H. (1990). Structure identification of nonlinear dynamic systems – A survey on input/output approaches. *Automatica*, 26(4):651–677.
- Haykin, S. (2001). *Redes neurais: princípios e prática*. Editora Bookman, 2. ed. Porto Alegre.
- He, X. e Asada, H. (1993). A new method for identifying orders of input-output models for nonlinear dynamic systems. In *Proceedings of the ACC*, São Francisco – Califórnia. p. 2520–2523.
- Henrique, H. M., Lima, E. L., e Pinto, J. C. (1998). A bifurcation study on neural network models for nonlinear dynamics systems. *Latin American Applied Research*, 28:187–200.
- Hernández, E. e Arkun, Y. (1993). Control of nonlinear systems using polynomial arma models. *AIChE Journal*, 39(3):446–460.
- Iglesias, R., Kyriacou, T., Nehmzow, U., e Billings, S. A. (2007). Task identification and characterisation in mobile robotics through nonlinear modelling. *Robotics and Autonomous Systems*, 55:267–275.
- Jazayeri, P., Rosehart, W., e Westwick, D. T. (2007). A multistage algorithm for identification of nonlinear aggregate power system loads. *IEEE Transactions on Power Systems*, 22(3):1072–1079.

- Jácome, C. R. F. (1996). *Uso de conhecimento prévio na identificação de modelos polinomiais NARMAX*. Belo Horizonte: Universidade Federal de Minas Gerais. (Dissertação de Mestrado, PPGEE).
- Judd, K. e Nakamura, T. (2006). Degeneracy of time series models: the best model is not always the correct model. *Chaos*, 16(3):033105–033105–3.
- Kadtke, J. B., Brush, J., e Holzfuss, J. (1993). Global dynamical equations and Lyapunov exponents from noisy chaotic time series. *International Journal of Bifurcation and Chaos*, 3(3):607–616.
- Kennel, M. B., Brown, R., e Abarbanel, H. D. I. (1992). Determining embedding dimension for phase-space reconstruction using a geometrical construction. *Physical Review A*, 45(6):3403–3411.
- Korenberg, M. J., Billings, S. A., Liu, Y. P., e Mcilroy, P. J. (1987). Orthogonal parameter estimation algorithm for nonlinear stochastic systems. *International Journal of Control*, 48:193–210.
- Koza, J. R. (1992). *Genetic programming: on the programming of computers by means of natural selection*. MIT Press, Cambridge, MA.
- Leontaritis, I. J. e Billings, S. A. (1985a). Input-output parametric models for nonlinear systems – Part 1: deterministic nonlinear systems. *International Journal of Control*, 41:303–328.
- Leontaritis, I. J. e Billings, S. A. (1985b). Input-output parametric models for nonlinear systems – Part 2: stochastic nonlinear systems. *International Journal of Control*, 41:329–344.
- Leontaritis, I. J. e Billings, S. A. (1987). Model selection and validation methods for nonlinear systems. *International Journal of Control*, 45:311–341.
- Ljung, L. (1999). *System identification – Theory for the user*. Prentice-Hall, 2. ed. Nova Jersey.
- Luna, I., Ballini, R., e Soares, S. (2006). Técnica de identificação de modelos lineares e não-lineares de séries temporais. *Controle & Automação*, 17(3):245–256.

- Luo, W. e Billings, S. A. (1995). Adaptive model selection and estimation for nonlinear systems using a sliding data window. *Signal Processing*, 46(2):179–202.
- Maddess, T., Nagai, Y., James, A. C., e Ankiewicz, A. (2004). Binary and ternary textures containing higher-order spatial correlations. *Vision Research*, 44(11):1093–1113.
- Madár, J., Abonyi, J., e Szeifert, F. (2005). Genetic programming for the identification of nonlinear input-output models. *Industrial & Engineering Chemistry Research*, 44(9):3178–3186.
- Mao, K. Z. e Billings, S. A. (1997). Algorithms for minimal model structure detection in nonlinear dynamic system identification. *International Journal of Control*, 68:311–330.
- Mao, K. Z. e Billings, S. A. (2000). Multi-directional model validity tests for non-linear system identification. *International Journal of Control*, 73(2):132–143.
- Matko, D., Zupančič, B., e Karba, R. (1992). *Simulation and modeling of continuous systems – A case study approach*. Prentice Hall International.
- Mendes, E. M. A. M. e Billings, S. A. (1998). On overparametrization of nonlinear discrete systems. *International Journal of Bifurcation and Chaos*, 8(3):535–556.
- Mendes, E. M. A. M. e Billings, S. A. (2001). An alternative solution to the model structure selection problem. *IEEE Transactions on Systems, Man and Cybernetics – Part A: Systems and Humans*, 31(6):597–608.
- Metenidis, M. F., Witczak, M., e Korbicz, J. (2004). A novel genetic programming approach to nonlinear system modelling: application to the DAMADICS benchmark problem. *Engineering Applications of Artificial Intelligence*, 17:363–370.
- Mitra, D. (1992). Asymptotically optimal design of congestion control for high speed data networks. *IEEE Transactions on Communications*, 40(2):301–311.

- Monteiro, L. H. A. (2002). *Sistemas dinâmicos*. Livraria da Física, São Paulo.
- Nakamura, T., Judd, K., Mees, A. I., e Small, M. (2006). A comparative study of information criteria for model selection. *International Journal of Bifurcation and Chaos*, 16(8):2153–2175.
- Nakamura, T. e Small, M. (2006). Modeling nonlinear time series using improved least squares method. *International Journal of Bifurcation and Chaos*, 16(2):445–464.
- Ogata, K. (2003). *Engenharia de controle moderno*. Editora Prentice Hall, 4. ed. São Paulo.
- Palumbo, P. e Piroddi, L. (2001). Seismic behaviour of buttress dams: nonlinear modelling of a damaged buttress based on ARX/NARX models. *Journal of Sound and Vibration*, 239(3):405–422.
- Patwardhan, R. S., Lakshminarayanan, S., e L., S. S. (1998). Constrained nonlinear MPC using Hammerstein and Wiener models: PLS framework. *AIChE Journal*, 44(7):1611–1622.
- Pearson, R. K. e Pottmann, M. (2000). Gray-box identification of block-oriented nonlinear models. *Journal of Process Control*, 10(4):301–315.
- Peyton-Jones, J. E. e Billings, S. A. (1989). Recursive algorithm for computing the frequency response of a class of nonlinear difference equations models. *International Journal of Control*, 50(5):1925–1940.
- Pierce, D. J. e Plemmons, R. J. (1992). Tracking the condition number for RLS in signal processing. *Mathematics of Control, Signals and Systems*, 5(1):23–39.
- Piroddi, L. (2008). Simulation error minimisation methods for NARX model identification. *International Journal of Modelling, Identification and Control*, 3(4):392–403.
- Piroddi, L. e Spinelli, W. (2003a). An identification algorithm for polynomial NARX models based on simulation error minimization. *International Journal of Control*, 76(17):1767–1781.

- Piroddi, L. e Spinelli, W. (2003b). Long-range nonlinear prediction: a case study. In *Proceedings of the 42^o IEEE CDC*, p. 3984–3989.
- Piroddi, L. e Spinelli, W. (2003c). A pruning method for the identification of polynomial NARMAX models. In *Proceedings of the 13^o IFAC SYSID*, p. 371–376.
- Piroddi, L. e Spinelli, W. (2003d). Structure selection for polynomial NARX models based on simulation error minimization. In *Proceedings of the 13^o IFAC SYSID*, p. 1108–1113.
- Porcaro, C., Zappasodi, F., Rossini, P. M., e Tecchio, F. (2009). Choice of multivariate autoregressive model order affecting real network functional connectivity estimate. *Clinical Neurophysiology*, 120(2):436–448.
- Previdi, F. (2002). Identification of black-box nonlinear models for lower limb movement control using functional electrical stimulation. *Control Engineering Practice*, 10:91–99.
- Rhodes, C. e Morari, M. (1995). Determining the model order of nonlinear input/output systems directly from data. In *Proceedings of the American Control Conference*, Seattle – Washington. v. 3, p. 2190–2194.
- Rhodes, C. e Morari, M. (1996). The FNN algorithm and noise corrupted time series. *IfA – ETH Technical Report AUT 96-03*.
- Rhodes, C. e Morari, M. (1997a). The false nearest neighbors algorithm: an overview. *Computers & Chemical Engineering*, 21:1149–1154.
- Rhodes, C. e Morari, M. (1997b). False-nearest-neighbors algorithm and noise-corrupted time series. *Physical Review E*, 55(5):6162–6170.
- Rissanen, J. (1978). Modeling by shortest data description. *Automatica*, 14(5):465–471.
- Rodríguez-Vázquez, K., Fonseca, C. M., e Fleming, P. J. (2004). Identifying the structure of nonlinear dynamic systems using multiobjective genetic programming. *IEEE Transactions on Systems, Man and Cybernetics – Part A: Systems and Humans*, 34(4):531–545.

- Rodrigues, G. G. (1996). *Identificação de sistemas dinâmicos não-lineares utilizando modelos NARMAX polinomiais – Aplicação a sistemas reais*. Belo Horizonte: Universidade Federal de Minas Gerais. (Dissertação de Mestrado, PPGEE).
- Ruano, A. E., Crispim, E. M., Conceição, E. Z. E., e Lúcio, M. M. J. R. (2006). Prediction of building's temperature using neural networks models. *Energy and Buildings*, 38(6):682–694.
- Schwarz, G. (1978). Estimating dimension of a model. *Annals of Statistics*, 6(2):461–464.
- Spinelli, W., Piroddi, L., e Lovera, M. (2006). A two-stage algorithm for structure identification of polynomial NARX models. In *Proceedings of the 2006 American Control Conference*, Minneapolis – Minnesota. p. 2387–2392.
- Stone, M. (1974). Cross-validatory choice and assessment of statistical predictions. *Journal of the Royal Statistical Society*, 36(2):111–147.
- van den Boom, A. J. W. e van den Enden, A. W. M. (1974). The determination of the orders of process and noise dynamics. *Automatica*, 10:245–256.
- Volterra, V. (1930). *Theory of functions*. Blackie and Sons.
- Wei, H. L. e Billings, S. A. (2008). Model structure selection using an integrated forward orthogonal search algorithm assisted by squared correlation and mutual information. *International Journal of Modeling, Identification and Control*, 3(4):341–356.
- Wiener, N. (1958). *Nonlinear problems in random theory*. In Wiley, J. and Sons, New York.
- Wigren, T. (1993). Recursive prediction error identification using nonlinear Wiener model. *Automatica*, 29(4):1011–1025.
- Zhang, L. F., Zhu, Q. M., e Longden, A. (2007). A set of novel correlation tests for nonlinear system variables. *International Journal of Systems Science*, 38(1):47–60.

- Zhu, Q. M. e Billings, S. A. (1991). Recursive parameter estimation for nonlinear rational models. *Journal of Systems Engineering*, 1:63–76.
- Zhu, Q. M. e Billings, S. A. (1993). Parameter estimation for stochastic nonlinear rational models. *International Journal of Control*, 57(2):309–333.
- Zhu, Q. M. e Billings, S. A. (1997). Properties of higher order correlation function tests for nonlinear model validation. In *Proceedings of the 23^o International Conference on Industrial Electronics, Control and Instrumentation*, New Orleans. v. 1, p. 306–310.
- Zhu, Q. M., Zhang, L. F., e Longden, A. (2007). Development of omnidirectional correlation functions for nonlinear model validation. *Automatica*, 43:1519–1531.

Rotinas Computacionais

Ao longo deste trabalho, foi necessário desenvolver algumas rotinas computacionais. Tais rotinas, implementadas no MATLABTM, são apresentadas, de forma sucinta, nas seções a seguir.

A.1 `apress`

Sintaxe

```
[npr_apress modelos_apress] = apress(modelo,u,y,nno,nitn,
    kp,alpha_max,num_fig,apress_tot)
```

Descrição

Plota o gráfico da estatística APRESS em função do número de termos de processo (termos de entrada + termos de saída) a fim de determinar o número ideal de termos de processo de um modelo*.

Dados de entrada

```
modelo -> modelo contendo os termos candidatos;
u -> sinal de entrada;
y -> sinal de saída;
nno -> número de termos de ruído;
nitn -> número de iterações de ruído;
kp -> kp passos à frente;
alpha_max -> máximo valor de alpha;
num_fig -> número da figura;
apress_tot -> se apress_tot = 1, serão apresentados
```

todos os valores de APRESS;
se $\text{apress_tot} \approx 1$, não serão
apresentados os valores de APRESS.

Dados de saída

`npr_apress` -> número de termos de processo com o
menor valor de APRESS;
`modelos_apress` -> célula contendo todos os modelos
analisados. Cada um desses modelos
possui um determinado número de termos
de processo.

* Método proposto em Billings, S. A. e Wei, H. L. (2008).
An adaptive orthogonal search algorithm **for** model subset
selection and non-linear system identification.
International Journal of Control, 81(5):714–724.

Marcela Andrade Alves
Universidade Federal de Minas Gerais – UFMG
Programa de Pós Graduação em Engenharia Elétrica – PPGEE
Belo Horizonte, 13/08/2008

A.2 `fnn_modificado`

Sintaxe

```
[nu ny] = fnn_modificado(u,y,tau,m_max,R,epsilon_u,  
epsilon_y,fnn_total)
```

Descrição

Calcula os máximos atrasos nos sinais de entrada e saída
de um modelo utilizando o método de falsos vizinhos (FNN)
modificado (considerando o ruído)*.

Dados de entrada

u → sinal de entrada;
y → sinal de saída;
tau → atraso de tempo;
m_max → máximo atraso considerado (tanto em relação à u quanto em relação à y);
R → valor limiar (normalmente, o valor de R encontra-se entre 10 e 50);
epsilon_u → máximo valor absoluto do ruído presente no sinal de entrada;
epsilon_y → máximo valor absoluto do ruído presente no sinal de saída;
fnn_total → se fnn_total = 1, todos os valores de FNN serão mostrados;
se fnn_total \neq 1, os valores de FNN não serão mostrados.

Dados de saída

nu → menor atraso do modelo (em relação ao sinal de entrada) quando fnn = 0%;
ny → menor atraso do modelo (em relação ao sinal de saída) quando fnn = 0%.

* Método proposto em Rhodes, C. e Morari, M. (1997).
The false nearest neighbors algorithm: an overview.
Computers and Chemical Engineering, 21:1149–1154.

Marcela Andrade Alves
Universidade Federal de Minas Gerais – UFMG
Programa de Pós Graduação em Engenharia Elétrica – PPGEE
Belo Horizonte, 01/08/2008

A.3 modelo_moda

Sintaxe

```
[mod_moda mmm amostras] = modelo_moda(m_total, Cand, Na)
```

Descrição

Analisa todos os modelos e retorna aquele cujos termos aparecem com maior frequência.

Dados de entrada

m_total -> célula contendo todos os modelos a serem analisados;
Cand -> matriz contendo todos os termos candidatos;
Na -> número de amostras usadas em cada modelo.

Dados de saída

mod_moda -> célula contendo o modelo cujos termos aparecem com maior frequência (modelo moda);
mmm -> número de ocorrências do modelo moda;
amostras -> amostras responsáveis pela geração do modelo moda.

Marcela Andrade Alves

Universidade Federal de Minas Gerais – UFMG

Programa de Pós Graduação em Engenharia Elétrica – PPGEE

Belo Horizonte, 26/08/2008

A.4 my_aic

Sintaxe

```
[npr_aic modelos_aic] = my_aic(modelo,u,y,nno,nitn,kp,  
    num_fig,aic_tot)
```

Descrição

Plota o gráfico do critério de informação de Akaike (AIC) em função do número de termos de processo (termos de entrada + termos de saída) a fim de determinar o número ideal de termos de processo de um modelo.

Dados de entrada

```
modelo -> modelo contendo os termos candidatos;  
    u -> sinal de entrada;  
    y -> sinal de saída;  
    nno -> número de termos de ruído;  
    nitn -> número de iterações de ruído;  
    kp -> kp passos à frente;  
num_fig -> número da figura;  
aic_tot -> se aic_tot = 1, serão apresentados  
    todos os valores de AIC;  
    se aic_tot ~= 1, não serão apresentados  
    os valores de AIC.
```

Dados de saída

```
npr_aic -> número de termos de processo com o  
    menor valor de AIC;  
modelos_aic -> célula contendo todos os modelos  
    analisados. Cada um desses modelos  
    possui um determinado número de termos  
    de processo.
```

Marcela Andrade Alves
Universidade Federal de Minas Gerais – UFMG
Programa de Pós Graduação em Engenharia Elétrica – PPGEE
Belo Horizonte, 01/08/2008

A.5 odccf

Sintaxe

```
rho = odccf(m, lag, flag, tfig, s)
```

Descrição

Calcula a função de correlação ODCCF* entre os sinais da matriz m.

Dados de entrada

m → matriz (Nx2) que contém os sinais a serem analisados;
lag → máximo atraso de correlação entre os sinais;
flag → se **flag** = 1, a função de correlação será calculada de $-\text{lag}/2$ a $\text{lag}/2$;
se **flag** = 0, a função de correlação será calculada de 0 a lag;
tfig → número da figura onde o gráfico será plotado;
s → título do gráfico.

Dado de saída

rho → função de correlação ODCCF.

* Função proposta em Zhang, L. F., Zhu, Q. M. e Longden, A. (2007). A set of novel correlation tests for nonlinear system variables. International Journal of Systems Science, 38(1):47–60.

Marcela Andrade Alves
 Universidade Federal de Minas Gerais – UFMG
 Programa de Pós Graduação em Engenharia Elétrica – PPGEE
 Belo Horizonte, 16/06/2007

A.6 orthregkp

Sintaxe

```
[m,x,e,va] = orthregkp(model,u,y,values,N,kp)
```

Descrição

Retorna o modelo obtido na predição de kp passos à frente.

Dados de entrada

```
model -> modelo contendo os termos candidatos
        (atenção: lagu = lagu+kp-1);
u -> sinal de entrada;
y -> sinal de saída;
values -> [(número de termos de processo) (número de
           termos de ruído)];
N -> número de iterações de ruído;
kp -> kp passos à frente.
```

Dados de saída

```
m -> modelo identificado
    (se 1 passo à frente: u(k-X) -> u(k-X));
    (se 2 passos à frente: u(k-X) -> u(k-X+1));
    (se 3 passos à frente: u(k-X) -> u(k-X+2));
    (           .           );
    (           .           );
    (           .           );
    (se kp passos à frente: u(k-X) -> u(k-X+kp-1));
x -> [(coeficientes) (ERR) (desvio padrão)];
e -> resíduo;
va -> variância.
```

Eduardo Mazoni Andrade Marçal Mendes
ACSE – Sheffield , 11/08/1994
Modificado por Marcela Andrade Alves e
Marcelo Vieira Corrêa
Universidade Federal de Minas Gerais – UFMG
Programa de Pós Graduação em Engenharia Elétrica – PPGEE
Belo Horizonte , 26/09/2008

A.7 predicao

Sintaxe

```
y_pred = predicao(y_1pf,k)
```

Descrição

Retorna o modelo de k passos a partir de seu respectivo modelo de 1 passo à frente.

Dados de entrada

```
y_1pf -> modelo de 1 passo à frente;  
k -> número de passos utilizados na predição.
```

Dado de saída

```
y_pred -> modelo de k passos à frente.
```

Observações

* O modelo de 1 passo à frente (y_1pf) deve estar na forma simbólica.

* Os termos com grau de não-linearidade maior que 1, por exemplo, $u(k-2)^2$, devem estar na forma expandida, ou seja, $u(k-2)*u(k-2)$.

* Os sinais de entrada e saída do modelo de 1 passo à

frente (y_{1pf}) devem ser representados por u e y respectivamente.

Exemplo:

$$y_{1pf} = '0.5*y(k-1) + 0.4*u(k-2)*u(k-2)'$$

Marcela Andrade Alves

Universidade Federal de Minas Gerais – UFMG

Programa de Pós Graduação em Engenharia Elétrica – PPGEE

Belo Horizonte, 18/02/2008

Funções de Correlação

A função de correlação é uma ferramenta que apresenta várias aplicações na área de identificação de sistemas, podendo ser utilizada, por exemplo, na determinação do tempo de amostragem, na estimação do tempo morto¹ de um processo (a partir do máximo valor da função de correlação cruzada entre seus sinais de entrada e saída), na verificação de redundância de informações entre as variáveis de um sistema, na validação estatística e na determinação dos atrasos presentes nos regressores de um modelo².

As funções de correlação linear³, tais como a função de autocorrelação (FAC) e a função de correlação cruzada (FCC), são normalmente utilizadas quando se deseja analisar associações lineares entre os sinais. Por esse motivo, elas são amplamente usadas na identificação de sistemas lineares. Entretanto, essas funções podem fornecer resultados incorretos se efeitos não-lineares estiverem presentes nos dados (Zhang et al., 2007). Além disso, as funções de correlação linear não são capazes de detectar associações não-lineares entre sinais.

A FCC normalizada entre $u_i(k)$ e $y_j(k)$ pode ser expressa como:

$$r_{u_i y_j}(\tau) = \frac{\sum_{k=\tau+1}^N (y_j'(k))(u_i'(k - \tau))}{\left[\left(\sum_{k=1}^N (y_j'(k))^2 \right) \left(\sum_{k=1}^N (u_i'(k))^2 \right) \right]^{1/2}}, \quad (\text{B.1})$$

sendo que N corresponde ao número de amostras e o sinal ' indica que a média foi removida. Quando $y_j(k) = u_i(k)$, a equação (B.1) é chamada de função de autocorrelação normalizada.

¹Também conhecido como atraso puro de tempo ou atraso de transporte.

²Conforme será apresentado neste apêndice.

³Apesar do pleonasmo existente no termo *correlação linear*, pois toda correlação é, de fato, linear, essa denominação será adotada neste trabalho para se referir à dependência linear entre variáveis.

Vários métodos, tais como testes de correlação multi-dimensional (Billings e Zhu, 1994; Aguirre, 1997; Maddess et al., 2004), FCC de ordem elevada (Aguirre, 1995; Zhu e Billings, 1997) e testes de correlação multi-direcional (Mao e Billings, 2000), foram desenvolvidos com o intuito de detectar associações não-lineares. Entretanto, a maioria dos testes de correlação não-linear⁴ não é tão simples quanto os testes de correlação linear (Billings e Voon, 1983).

A FCC de ordem elevada normalizada pode ser representada por (Zhu e Billings, 1997):

$$r_{(u_i^2)'(y_j^2)'}(\tau) = \frac{\sum_{k=\tau+1}^N (((y_j^2(k))')((u_i^2(k-\tau))'))}{\left[\left(\sum_{k=1}^N ((y_j^2(k))')^2 \right) \left(\sum_{k=1}^N ((u_i^2(k))')^2 \right) \right]^{1/2}}. \quad (\text{B.2})$$

É importante ressaltar que a FCC de ordem elevada não pode ser usada para lidar com todos os modelos não-lineares e pode falhar na detecção de associações não-lineares sob algumas condições especiais.

A fim de detectar correlações lineares e não-lineares, uma série de FCCs de primeira ordem, denominada ODCCFs (*omni-directional cross-correlation functions*) normalizadas, foi proposta em Zhang et al. (2007). Essas funções são usadas na detecção de quatro tipos de associações que incluem todos os possíveis efeitos não-lineares. Tais associações são representadas da seguinte forma (Zhang et al., 2007):

⁴Neste trabalho, será empregado o termo *correlação não-linear* ainda que, a rigor, a denominação *dependência não-linear* seja mais adequada.

$$\begin{aligned}
r_{\alpha\beta}(\tau) &= \frac{\sum_{k=\tau+1}^N (\beta(k) - \bar{\beta})(\alpha(k - \tau) - \bar{\alpha})}{\left[\left(\sum_{k=1}^N (\beta(k) - \bar{\beta})^2 \right) \left(\sum_{k=1}^N (\alpha(k) - \bar{\alpha})^2 \right) \right]^{1/2}}, \\
r_{\alpha y'_j}(\tau) &= \frac{\sum_{k=\tau+1}^N y'_j(k) (\alpha(k - \tau) - \bar{\alpha})}{\left[\left(\sum_{k=1}^N (y'_j(k))^2 \right) \left(\sum_{k=1}^N (\alpha(k) - \bar{\alpha})^2 \right) \right]^{1/2}}, \\
r_{u'_i y'_j}(\tau) &= \frac{\sum_{k=\tau+1}^N y'_j(k) u'_i(k - \tau)}{\left[\left(\sum_{k=1}^N (y'_j(k))^2 \right) \left(\sum_{k=1}^N (u'_i(k))^2 \right) \right]^{1/2}}, \\
r_{u'_i \beta}(\tau) &= \frac{\sum_{k=\tau+1}^N (\beta(k) - \bar{\beta}) u'_i(k - \tau)}{\left[\left(\sum_{k=1}^N (\beta(k) - \bar{\beta})^2 \right) \left(\sum_{k=1}^N (u'_i(k))^2 \right) \right]^{1/2}}, \tag{B.3}
\end{aligned}$$

sendo que

$$\begin{aligned}
\alpha(k) &= |u'_i(k)|, \\
\beta(k) &= |y'_j(k)|. \tag{B.4}
\end{aligned}$$

Quando $y_j(k) = u_i(k)$, o conjunto de equações (B.3) é denominado ODACFs (*omni-directional auto-correlation functions*) normalizadas.

Os resultados obtidos a partir das funções ODCCFs podem ser combinados da seguinte forma (Zhang et al., 2007):

- Se $|\max(r_{\alpha\beta}(\tau), r_{\alpha y'_j}(\tau), r_{u'_i y'_j}(\tau), r_{u'_i \beta}(\tau))| > |\min(r_{\alpha\beta}(\tau), r_{\alpha y'_j}(\tau), r_{u'_i y'_j}(\tau), r_{u'_i \beta}(\tau))|$, então:

$$\rho_{uy}(\tau) = \max(r_{\alpha\beta}(\tau), r_{\alpha y'_j}(\tau), r_{u'_i y'_j}(\tau), r_{u'_i \beta}(\tau)). \tag{B.5}$$

- Se $|\max(r_{\alpha\beta}(\tau), r_{\alpha y'_j}(\tau), r_{u'_i y'_j}(\tau), r_{u'_i \beta}(\tau))| \leq |\min(r_{\alpha\beta}(\tau), r_{\alpha y'_j}(\tau), r_{u'_i y'_j}(\tau), r_{u'_i \beta}(\tau))|$, então:

$$\rho_{uy}(\tau) = \min(r_{\alpha\beta}(\tau), r_{\alpha y'_j}(\tau), r_{u'_i y'_j}(\tau), r_{u'_i \beta}(\tau)). \tag{B.6}$$

Os Exemplos B.0.1 e B.0.2, apresentados a seguir, mostram a eficiência da função ODCCF em relação às funções FCC e FCC de ordem elevada na detecção de associações não-lineares. Tais exemplos foram apresentados em Zhang et al. (2007).

Exemplo B.0.1 ⁵

O primeiro sistema a ser analisado é representado por:

$$y(k) = u_1(k-1)^2 - 2u_3(k-5)^2 + 2u_2(k-3)u_4(k-7) - u_2(k)^3 + e(k), \quad (\text{B.7})$$

sendo que as variáveis independentes $u_1(k)$, $u_2(k)$, $u_3(k)$ e $u_4(k)$, com dimensão (1000×1) , são formadas por seqüências aleatórias com distribuição uniforme, média nula e amplitude entre -1 e 1. O ruído $e(k)$ é formado por uma seqüência de dados aleatórios com distribuição gaussiana, média nula e variância de 0,01.

Para efeito de comparação, utilizaram-se as funções FCC (Figura B.1), FCC de ordem elevada (Figura B.2) e ODCCF (Figura B.3) a fim de se verificar a eficiência de cada função na detecção das associações não-lineares presentes no sistema.

A Figura B.1 mostra que somente $r_{u_2y}(0)$ localiza-se fora da região de confiança. Portanto, a função FCC não foi capaz de detectar as outras associações presentes no sistema.

Analisando-se a Figura B.2, é possível perceber que $r_{(u_2^2)'(y^2)'(\tau)}$ encontra-se fora da região de confiança em $\tau = 0$ e $\tau = 3$. Além disso, $r_{(u_3^2)'(y^2)'(5)}$ e $r_{(u_4^2)'(y^2)'(7)}$ também localizam-se fora da região de confiança. Apesar disso, a função FCC de ordem elevada apresentou as seguintes falhas: (i) $r_{(u_1^2)'(y^2)'(1)}$ não se encontra fora da região de confiança e (ii) $r_{(u_2^2)'(y^2)'(0)}$ e $r_{(u_3^2)'(y^2)'(5)}$ deveriam ser negativas. Portanto, as funções FCC e FCC de ordem elevada mostraram-se ineficientes na tentativa de detectar as associações não-lineares presentes no sistema.

⁵Zhang et al., 2007.

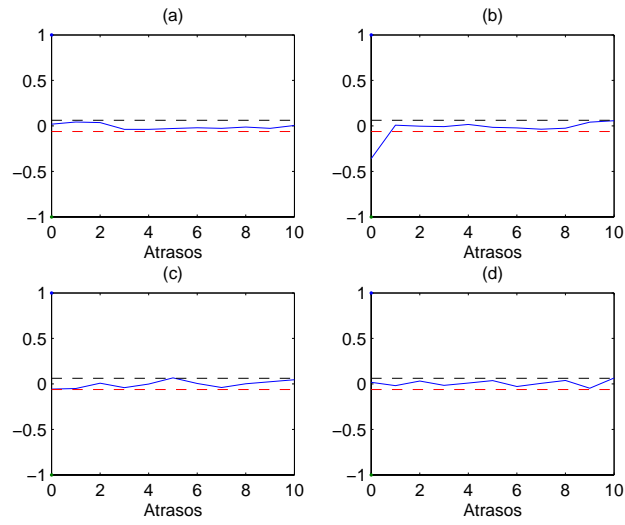


Figura B.1: Resultado gráfico da aplicação da função FCC aos dados do sistema (B.7).

(a) $r_{u_1y}(\tau)$, (b) $r_{u_2y}(\tau)$, (c) $r_{u_3y}(\tau)$ e (d) $r_{u_4y}(\tau)$.

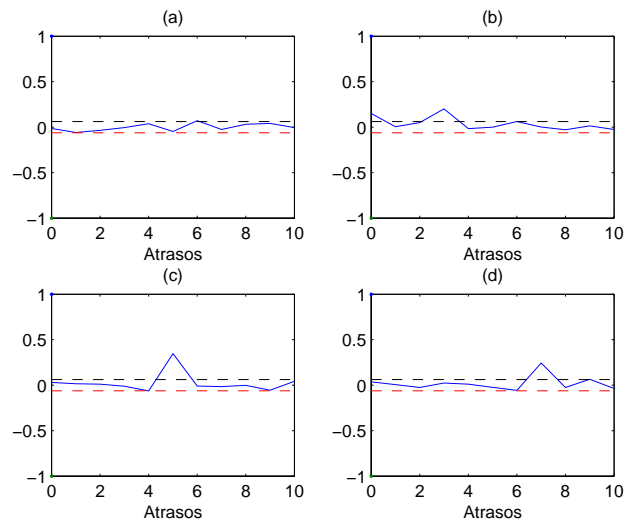


Figura B.2: Resultado gráfico da aplicação da função FCC de ordem elevada aos dados do sistema (B.7).

(a) $r_{(u_1^2)'(y^2)'(\tau)}$, (b) $r_{(u_2^2)'(y^2)'(\tau)}$, (c) $r_{(u_3^2)'(y^2)'(\tau)}$ e (d) $r_{(u_4^2)'(y^2)'(\tau)}$.

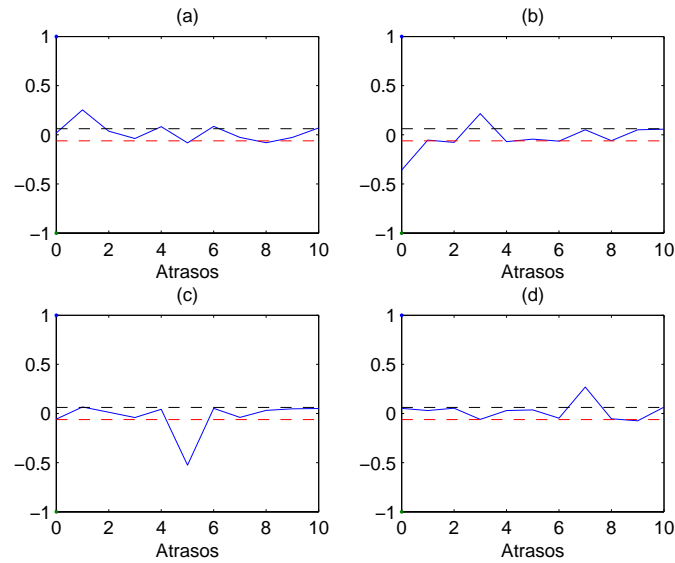


Figura B.3: Resultado gráfico da aplicação da função ODCCF aos dados do sistema (B.7).

(a) $\rho_{u_1y}(\tau)$, (b) $\rho_{u_2y}(\tau)$, (c) $\rho_{u_3y}(\tau)$ e (d) $\rho_{u_4y}(\tau)$.

A Figura B.3, por sua vez, revela que a função ODCCF foi capaz de detectar corretamente todas as associações não-lineares presentes no sistema. Portanto, tal função mostrou ser mais eficiente em relação às funções FCC e FCC de ordem elevada. \square

Exemplo B.0.2 ⁶

O segundo sistema a ser analisado é um sistema complexo que envolve termos do tipo senoidal e exponencial. Tal sistema é representado por:

$$y(k) = \frac{\text{sen}(2u_2(k-3)u_3(k-5) + u_1(k-1)^2)}{\exp(u_4(k-7))} + e(k), \quad (\text{B.8})$$

sendo que as variáveis independentes $u_1(k)$, $u_2(k)$, $u_3(k)$ e $u_4(k)$ e o ruído $e(k)$ foram gerados conforme descrito no Exemplo B.0.1.

⁶Zhang et al., 2007.

Para efeito de comparação, utilizaram-se, novamente, as funções FCC (Figura B.4), FCC de ordem elevada (Figura B.5) e ODCCF (Figura B.6) a fim de se verificar a eficiência de cada função na detecção de associações não-lineares presentes no sistema.

Analisando-se a Figura B.4, é possível perceber que somente $r_{u_4y}(7)$ localiza-se fora da região de confiança. Portanto, a função FCC mostrou-se ineficiente na tentativa de detectar as associações não-lineares presentes no sistema.

Em relação à Figura B.5, verifica-se que, apesar de a função FCC de ordem elevada ter sido capaz de detectar todas as associações não-lineares do sistema, $r_{(u_4^2)'(y^2)'(7)}$ deveria ser negativa. Além disso, $r_{(u_1^2)'(y^2)'(1)}$ não se encontra tão claramente fora da região de confiança.

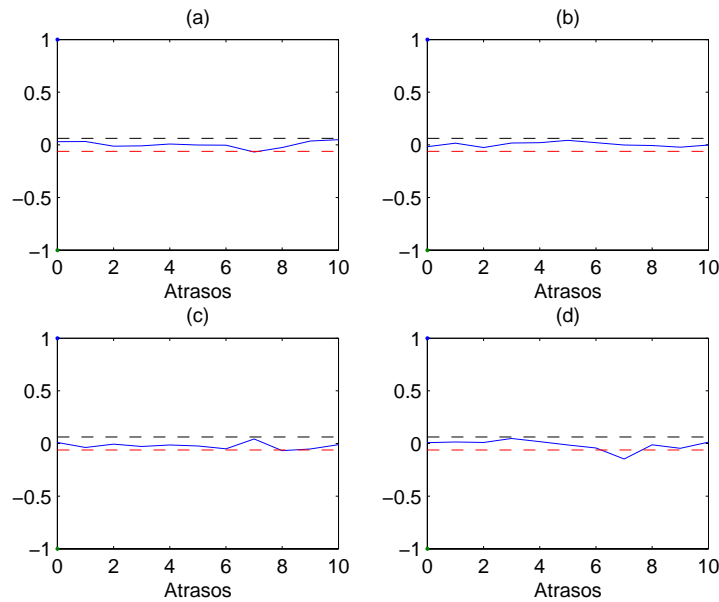


Figura B.4: Resultado gráfico da aplicação da função FCC aos dados do sistema (B.8).

(a) $r_{u_1y}(\tau)$, (b) $r_{u_2y}(\tau)$, (c) $r_{u_3y}(\tau)$ e (d) $r_{u_4y}(\tau)$.

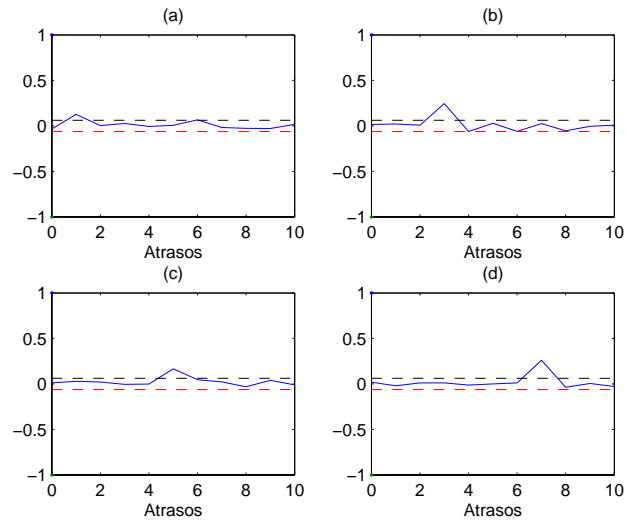


Figura B.5: Resultado gráfico da aplicação da função FCC de ordem elevada aos dados do sistema (B.8).

(a) $r_{(u_1^2)'(y^2)'(\tau)}$, (b) $r_{(u_2^2)'(y^2)'(\tau)}$, (c) $r_{(u_3^2)'(y^2)'(\tau)}$ e (d) $r_{(u_4^2)'(y^2)'(\tau)}$.

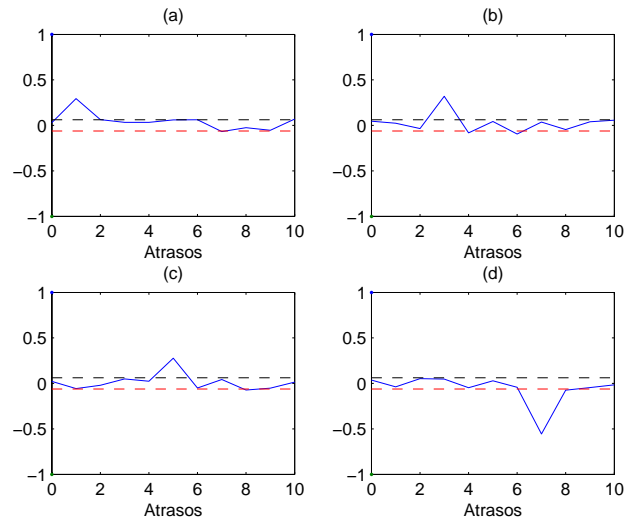


Figura B.6: Resultado gráfico da aplicação da função ODCCF aos dados do sistema (B.8).

(a) $\rho_{u_1y}(\tau)$, (b) $\rho_{u_2y}(\tau)$, (c) $\rho_{u_3y}(\tau)$ e (d) $\rho_{u_4y}(\tau)$.

A Figura B.6, por sua vez, revela que a função ODCCF foi capaz de detectar corretamente todas as associações não-lineares presentes no sistema. Além disso, $\rho_{u_1y}(1)$ localiza-se significativamente fora da região de confiança. Portanto, a função ODCCF mostrou, novamente, ser mais eficiente em relação às funções FCC e FCC de ordem elevada. \square

Nos Exemplos B.0.1 e B.0.2, foi possível constatar a eficiência da função ODCCF em relação às funções FCC e FCC de ordem elevada tanto no caso de um sistema simples (equação (B.7)) quanto no caso de um sistema complexo (equação (B.8)). Entretanto, é importante observar que os sistemas considerados nesses exemplos são de resposta finita ao impulso (FIR). Em virtude disso, um novo teste foi realizado acrescentando-se um termo de saída ao sistema (B.7), tornando-o, assim, auto-regressivo. Os resultados obtidos a partir da aplicação das funções FCC, FCC de ordem elevada e ODCCF nesse novo sistema encontram-se no Exemplo B.0.3.

Exemplo B.0.3

Um terceiro sistema foi criado acrescentando-se um termo de saída ao sistema (B.7) a fim de se verificar a eficiência da função ODCCF no caso de sistemas auto-regressivos. Para isso, as variáveis independentes $u_1(k)$, $u_2(k)$, $u_3(k)$ e $u_4(k)$ e o ruído $e(k)$ foram novamente gerados conforme descrito no Exemplo B.0.1. Tal sistema é representado por:

$$y(k) = u_1(k-1)^2 - 2u_3(k-5)^2 + 2u_2(k-3)u_4(k-7) - u_2(k)^3 - 0,7y(k-1) + e(k). \quad (\text{B.9})$$

Assim como nos Exemplos B.0.1 e B.0.2, as funções FCC (Figura B.7), FCC de ordem elevada (Figura B.8) e ODCCF (Figura B.9) foram utilizadas com o intuito de se verificar a eficiência de cada função na detecção de associações não-lineares presentes no sistema.

A Figura B.7 mostra que a função FCC só foi capaz de detectar associações entre o sinal de saída e a entrada u_2 (Figura B.7b). Além disso, $r_{u_2y}(\tau)$ encontra-se erroneamente fora da região de confiança em $\tau = 1$ e $\tau = 2$. Portanto, a função FCC mostrou-se ineficiente ao tentar detectar as associações não-lineares presentes no sistema.

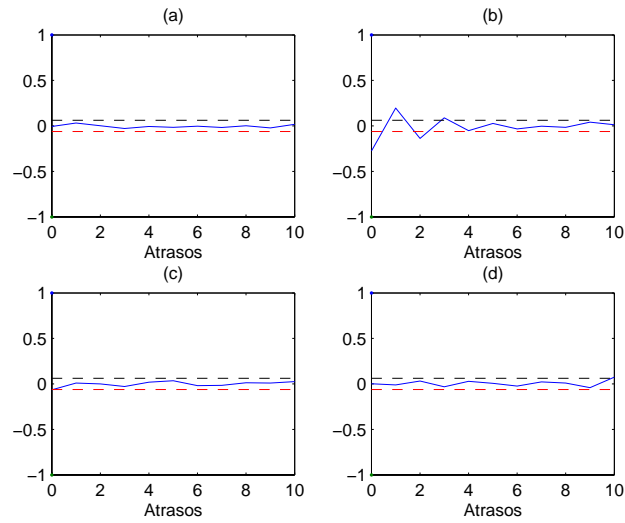


Figura B.7: Resultado gráfico da aplicação da função FCC aos dados do sistema (B.9).

(a) $r_{u_1y}(\tau)$, (b) $r_{u_2y}(\tau)$, (c) $r_{u_3y}(\tau)$ e (d) $r_{u_4y}(\tau)$.

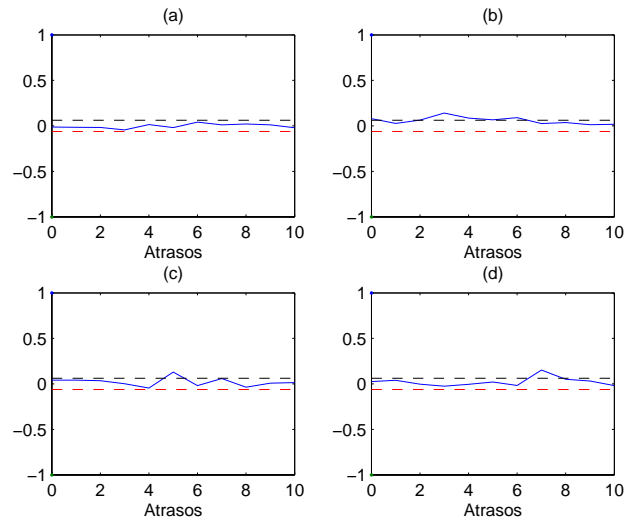


Figura B.8: Resultado gráfico da aplicação da função FCC de ordem elevada aos dados do sistema (B.9).

(a) $r_{(u_1^2)'(y^2)'(\tau)}$, (b) $r_{(u_2^2)'(y^2)'(\tau)}$, (c) $r_{(u_3^2)'(y^2)'(\tau)}$ e (d) $r_{(u_4^2)'(y^2)'(\tau)}$.

Na Figura B.8, verifica-se que $r_{(u_2)'(y^2)'(3)}$, $r_{(u_3)'(y^2)'(5)}$ e $r_{(u_4)'(y^2)'(7)}$ localizam-se ligeiramente fora da região de confiança. Nota-se também que a função FCC de ordem elevada apresentou as seguintes falhas: (i) $r_{(u_1)'(y^2)'(1)}$ e $r_{(u_2)'(y^2)'(0)}$ não se encontram fora da região de confiança e (ii) $r_{(u_3)'(y^2)'(5)}$ deveria ser negativa. Portanto, tanto a FCC quanto a FCC de ordem elevada mostraram-se ineficientes na detecção das associações não-lineares presentes no sistema.

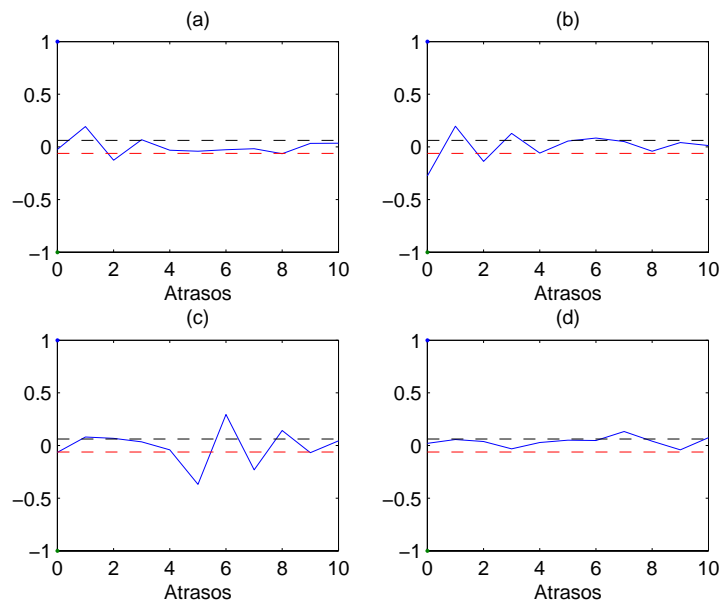


Figura B.9: Resultado gráfico da aplicação da função ODCCF aos dados do sistema (B.9).

(a) $\rho_{u_1y}(\tau)$, (b) $\rho_{u_2y}(\tau)$, (c) $\rho_{u_3y}(\tau)$ e (d) $\rho_{u_4y}(\tau)$.

A Figura B.9 mostra que $\rho_{u_1y}(1)$, $\rho_{u_3y}(5)$, $\rho_{u_2y}(0)$, $\rho_{u_2y}(3)$ e $\rho_{u_4y}(7)$ localizam-se, de fato, fora da região de confiança, mesmo que ligeiramente nos dois últimos casos. Apesar de a função ODCCF ter sido capaz de detectar corretamente todas as associações não-lineares presentes no sistema, verifica-se que $\rho_{u_1y}(\tau)$, $\rho_{u_2y}(\tau)$ e $\rho_{u_3y}(\tau)$ extrapolam os limites da região de confiança em vários outros valores de τ . Portanto, neste caso, a função ODCCF também se mostrou ineficiente na tentativa de detectar as associações não-lineares presentes no sistema. \square

Os Exemplos B.0.1 e B.0.2 mostraram a eficiência da função ODCCF na seleção de estrutura de modelos do tipo FIR. Entretanto, no Exemplo B.0.3, verificou-se que a inclusão de um único termo de saída no sistema (B.7) prejudicou significativamente o desempenho de tal função. É provável que isso tenha acontecido em virtude de o sistema (B.9) ser autorregressivo. Portanto, suspeita-se que a função ODCCF seja indicada somente no caso de modelos do tipo FIR, tais como os sistemas (B.7) e (B.8). Em Zhu et al. (2007), a função ODCCF foi empregada na validação estatística de modelos com o propósito de verificar a aleatoriedade dos resíduos. A função ODCCF foi implementada na função `odccf`, que se encontra no Apêndice A.5 deste trabalho.