GCoM
Grupo de Controle e Modelagem
Universidade Federal de São João del-Rei - UFSJ
Departamento de Engenharia Elétrica - DEPEL

UFSJ
UNIVERSIDADE FEDERAL
DE SÃO JOÃO DEL-REI

Samir Angelo Milani Martins

Aplicação de Otimização Multi-objetivo na Detecção de Estruturas de Modelos NARMAX Polinomiais

Dissertação apresentada à banca examinadora designada pelo Colegiado do Programa de Pós-Graduação em Engenharia Elétrica, associação ampla entre a Universidade Federal de São João del-Rei e o Centro Federal de Educação Tecnológica de Minas Gerais, como parte dos requisitos necessários à obtenção do grau de Mestre em Engenharia Elétrica.

Orientador: Prof. Dr. Erivelton Geraldo Nepomuceno

Co-orientador: Prof. Dr. Márcio Falcão Santos Barroso

São João del-Rei
2012

Dedicatória

A DEUS, que sempre me dá mais que mereço.

Aos meus pais, José Antônio e Edy, que me ensinaram valores e princípios que a vida moderna não ensina, que nenhum grau em nenhuma universidade poderia ensinar.

Agradecimentos

A DEUS!

À MARIA, MÃE de DEUS, pelas incessantes intercessões. Obrigado Padre Pio, por ser meu guia espiritual!

Aos meus pais, meus grandes mestres e amigos! Hoje sou o que vocês me inspiraram a ser!

Aos meus irmãos Gustavo e Ramon, pela amizade, pela experiência adquirida nas trocas de conversas e pelo apoio quando necessário.

Ao meu grande amor, Érica! Obrigado por me fazer sorrir todos os dias! Obrigado por me conquistar sempre, cada vez de um modo mais intenso! Te amo!

Ao grande amigo Erivelton, pelo grande apoio dado desde 2006, por ter me iniciado na ciência e pelos crescimentos decorrentes de longas conversas.

Ao co-orientador Márcio Barroso, pelo apoio prestado durante o desenvolvimento deste e de outros trabalhos.

Ao movimento católico Comunhão e Libertação. Obrigado por serem a face de Cristo para mim.

Ao meu amigo João Paulo, pelo apoio e crescimento durante a graduação e mestrado. À República Caverna, pela hospitalidade.

Aos amigos da República Calabouço e aos de Bom Sucesso, que sempre se fazem presente em minhas conquistas pessoais!

Aos amigos que fiz durante a graduação e mestrado! Vocês farão sempre parte de minha vida!

Ao amigo Alípio, pelas conversas científicas, políticas, pessoais e pela amizade ao longo dos anos!

A todos das famílias Milani Martins!

Ao GCoM, onde dei meus primeiros passos na ciência.

À CAPES, FAPEMIG e CNPq, pelo apoio financeiro ao longo de anos.

Àqueles que, por limitação de espaço, não pude citar aqui. Estão sempre

comigo!

"Veni Sancte Spiritus, Veni Per Mariam"

Epígrafe

"(· · ·) Como não ter Deus?! Com Deus existindo, tudo dá esperança: sempre um milagre é possível, o mundo se resolve. Mas, se não tem Deus, há-de a gente perdidos no vai-vem, e a vida é burra."

Grande Sertão Veredas, de Guimarães Rosa

"O primeiro gole do copo das ciências o torna ateu; mas no fundo do copo Deus aguarda."

Werner Heisenberg

"A fé é o ápice da razão"

Dom Luigi Giussani - Movimento Católico Comunhão e Libertação

Sumário

Resumo

A identificação de sistemas é a área da ciência que busca encontrar padrões e propor modelos matemáticos a sistemas reais. Pode-se resumir o processo de identificação de sistemas em cinco principais etapas: 1) testes e coleta de dados, 2) escolha da representação matemática, 3) detecção da estrutura do modelo, 4) estimação dos parâmetros e, 5) validação dos modelos. Dessas etapas, tem-se a determinação da estrutura do modelo como sendo uma das mais complexas e importantes. Tal problema foi estudado neste trabalho.

Dentre as principais técnicas de detecção de estruturas, tem-se a taxa de redução de erro (ERR). Essa técnica quantifica a contribuição de cada regressor ao explicar a variância dos dados dinâmicos de um sistema. Contudo, apesar de ser uma técnica consagrada, tem como grande limitação o fato de quantificar os regressores utilizando somente dados dinâmicos do sistema. Neste sentido, foi desenvolvida a taxa de redução de erro multi-objetivo (MERR, do inglês *Multi-objective Error Reduction Ratio*), uma generalização multi-objetivo da ERR em que outras informações acerca do sistema também podem ser consideradas.

A metodologia proposta foi aplicada na identificação de dois sistemas simulados, amplamente estudados, e na identificação de dois sistemas reais (um conversor CC-CC buck e uma planta de neutralização de pH). Os resultados simulados mostraram que a técnica foi capaz de discriminar os regressores genuínos dos espúrios, ao passo que os resultados reais mostraram grande eficiência na identificação do sistema. Pôde-se obter modelos com dinâmica e estática representativos e adequados, na identificação do conversor CC-CC buck (sistema com dinâmica não-linear e estática afim). Considerando a planta de neutralização de pH (sistema com dinâmica e estática não-linear), obteve-se um modelo estável e representativo com uma pequena massa de dados, o que não foi possível por meio da técnica ERR.

Palavras-chave: Identificação de sistemas, MERR, detecção de estruturas, modelos NARMAX polinomiais.

Abstract

The system identification is an area of the science that studies several ways to find patterns, obtain mathematical models to represent systems, using observations and theoretical data. This process can be summarized in five main steps, namely: 1) Tests and data acquisition, 2) choice of the mathematical representation, 3) the structure detection, 4) parameter estimation and, 5) model validation. The structure detection is one of the most complex and significant system identification's steps. This issue was studied on this work.

Among structure detection techniques, the Error Reduction Ratio (ERR) can be cited. This approach aims to quantify the contribution of each regressor, when inserted into the model, aiming reduce the dynamic residual. However, this technique has as one of its limitations the use of only dynamic data to obtain the model's regressors. In this sense, a multi-objective error reduction ratio (MERR) was developed on this work, a generalization of the ERR, where other information about the system can be used.

The proposed metodology was applied in the identification of two simulated systems broadly studied in the science and to identify two real systems (a DC-DC buck converter and a pH neutralization process). The real structure could be obtained for the simulated systems using the MERR approach, whereas the real results showed a good efficiency in the identification process. Representative and well fitted models could be obtained in the identification of the buck converter (system with non-linear dynamic and affine static behavior). Considering the pH neutralization process, a stable and representative model was obtained using a small amount of dynamical and static data, which was not possible by mean of ERR.

Keywords: System identification, MERR, structure detection, NARMAX polynomial models.

Lista de Tabelas

Lista de Figuras

Lista de Símbolos

b	Número de bits;
C	Termo constante;
D	Razão cíclica;
$e(k)$	Regressor de ruído;
f_s	Frequência de amostragem;
$\hat{g}_j$	Parâmetros associados aos regressores multi-objetivo ortogonais;
J_{AIC}	Função custo do critério de informação de Akaike;
J_d	Função custo quadrático dinâmico;
J_m	Função custo quadrático multi-objetivo;
J_s	Função custo quadrático estático;
$\hat{K}$	Ganho estático;
n_e	Máximo atraso do regressor de ruído;
n_u	Máximo atraso do regressor de entrada;
n_y	Máximo atraso do regressor de saída;
n_θ	Número de parâmetros de um modelo;
Q	Matriz que mapeia os pontos fixos;
R	Matriz que mapeia os parâmetros;
T_B	Tempo de chaveamento;
T_s	Período de amostragem;

T_{on}	Tempo em que o conversor se encontra ligado;
T	Tempo de operação do conversor;
$u(k)$	Regressor de entrada;
V_d	Tensão CC de alimentação do conversor;
V_o	Tensão de saída do conversor;
y_d	Valor da saída dinâmico;
y_m	Valor da saída composto;
y_s	Valor da saída estático;
$\hat{y}_d$	Valor da saída dinâmico estimado;
$\hat{y}_i$	Valor de dada informação i estimado;
$\hat{y}_s$	Valor estático estimado;
w_i	Regressores do modelo ortogonal;
$w(k)$	Ruído gaussiano branco;
$\Xi(k)$	Erro de predição;
ξ_s	Resíduo estático;
ξ_d	Resíduo dinâmico;
ξ_m	Resíduo composto;
$\sum_{n_1,n_m}^{n_y,n_u} c_{p,m-p}(n_1,...,n_m)$	Coeficientes de agrupamentos;
Σ_0	termo constante;
$\Sigma_y \bar{y}$	termos lineares em y;
$\Sigma_u \bar{u}$	termos lineares em u;
$\sum_{m=1}^{\ell-1} \sum_{p=1}^{\ell-m} \Sigma_{y^p u^m} \bar{y}^p \bar{u}^m$	termos cruzados;
$\sum_{i=1}^{\ell} \Sigma_{y^i} \bar{y}^i$	termos não-lineares em y;

$\sum_{i=1}^{l} \Sigma_{u^i} \bar{u}^i$	termos não-lineares em u;
θ	Vetor de parâmetros reais;
$\hat{\theta}$	Vetor de parâmetros estimados;
ψ_d	Matriz de regressores dinâmicos;
ψ_s	Matriz de regressores estáticos;
Ψ_m	Matriz de regressores composta;
ω_d	Peso dado à informação dinâmica;
ω_s	Peso dado à informação estática;
Ω_m^i	Regressores ortogonais compostos;
Λ	Modelos NARMAX obtidos;
Δ	Fator de decimação;

Lista de Abreviações

AC	Corrente alternada;
AIC	Critério de informação de Akaike;
CC	Corrente contínua;
CE	Expansão completa;
CSI	Identificaçao de estruturas baseado em agrupamentos;
ELS	Mínimos quadrados estendidos;
ERR	Taxa de redução de erro;
ERR_2	Taxa de redução de erro dois passos a frente;
GE	Expansão gradual;
IE	Expansão iterativa;
LS	Mínimos quadrados;
MACSIN	Modelagem, análise e controle de sistemas não-lineares;
MERR	Taxa de redução de erro multi-objetivo;
MOSFET	Transistor de efeito de campo de metal oxido semicondutor;
NARMAX	Modelo não-linear, auto-regressivo com média móvel e entrada exógena;
pH	Potencial hidrogeniônico;
PPGEE	Programa de Pós-graduação em Engenharia Elétrica da UFMG;
PRBS	Sinal binário pseudo-aleatório;
PWM	Modulação por largura de pulso;
RMSE	Raiz quadrada do erro médio quadrático;
SEMP	Minimização do erro de simulação com punição;
SISO	Sistema com uma entrada e uma saída;
SRR	Taxa de redução de erro de simulação;
UFMG	Universidade Federal de Minas Gerais;

Capítulo 1

Introdução

"No princípio era o Verbo, e o Verbo estava junto de Deus e o Verbo era Deus. Ele estava no princípio junto de Deus. $(\cdots)$. E o Verbo se fez carne e habitou entre nós, e vimos sua glória, a glória que o Filho único recebe do seu Pai, cheio de graça e de verdade."

Evangelho de São João, $1, 1-2, 14$

A Identificação de Sistemas é uma das atividades mais antigas e relevantes da ciência. É a área do conhecimento que estuda maneiras de modelar e analisar sistemas na tentativa de encontrar algum padrão em observações. Define-se modelo como o conjunto de hipóteses sobre a estrutura ou comportamento de um sistema físico. Do lado matemático, um modelo nada mais é do que uma abstração de um sistema real expresso por meio de equações (Aguirre, 2007).

As técnicas de Identificação de Sistemas podem ser empregadas na proposição de modelos para sistemas de diferentes áreas. Na engenharia elétrica, pode-se identificar aquecedores elétricos (Martins et al., 2009; Furtado et al., 2002) e conversores eletrônicos de potência, como conversores CC-CC buck (Algreer et al., 2011; Martins et al., 2011; Yang et al., 2009). Como sistemas químicos, pode-se, por exemplo, modelar plantas de neutralização de pH (Margoti, 2011; Lima, 2009). Já na área biológica, modela-se o comportamento dinâmico de epidemias, propondo ações de controle baseada em modelos (Fellmann et al., 2011; Wickwire, 1977). A identificação também

é utilizada no contexto de predição de séries temporais em sistemas econômicos, que são sistemas altamente sazonais (Valackiene, 2011; Neves et al., 2007).

Existem algumas etapas típicas no processo de construção de modelos matemáticos a partir de dados (Ljung, 1987): i) execução de testes e coletas de dados, ii) escolha da classe de modelos a ser utilizada, iii) escolha dos regressores do modelo, iv) estimação de parâmetros associados a cada regressor e v) validação do modelo.

Neste âmbito, a escolha da representação matemática a ser utilizada é de grande importância na qualidade do modelo obtido. Dentre as principais abordagens matemáticas utilizadas na literatura, pode-se citar a representação NARMAX (Nonlinear AutoRegressive Moving Average model with eXogenous input) polinomial (Billings, 1980) como sendo uma grande ferramenta na identificação de sistemas dinâmicos não-lineares. Ressalta-se que a nomenclatura NARMAX será utilizada para denominar modelos NARX (Nonlinear AutoRegressive model with eXogenous input) ou NARMAX, uma vez que a técnica aqui desenvolvida é aplicável à ambas classes de modelos.

Modelos NARMAX permitem a representação de dinâmicas altamente não-lineares por meio de modelos de diferença relativamente simples. Ademais, tem-se como ponto positivo a grande capacidade de incorporação de outros tipos de informação do sistema no modelo a um baixo esforço matemático.

Atualmente, a etapa de determinação dos parâmetros encontra-se bastante consolidada. Várias técnicas de caráteres mono e multi-objetivo foram testadas com êxito (Ding et al., 2011; Wang e Ding, 2011; Martins et al., 2009; Nepomuceno et al., 2007; Barroso et al., 2007; Koop e Orford, 1963; Lichtenberger, 1961), permitindo quantificar os regressores de modo adequado, quando da representação de um dado sistema.

Diferentemente, encontra-se ainda hoje um grande desafio na determinação de estruturas de modelos matemáticos. Dentre as técnicas presentes na literatura, pode-se citar a taxa de redução de erro (ERR - do inglês *Error Reduction Ratio*), que pode ser aplicada em conjunto com o critério de informação de Akaike (AIC - do inglês *Akaike Information Criteria*). Outras técnicas podem ser vistas em (Barbosa, 2010; Bonin et al., 2010; Alves, 2009; Hong et al., 2008; Lind e Ljung, 2008; Wei e Billings, 2008; Pulecchi e Piroddi, 2007; Lind e Ljung, 2005; Mendes e Billings, 2001). Contudo, ainda existem limitações nestas técnicas, pois a grande maioria utiliza somente de

informações contidas nos dados dinâmicos para obtenção dos regressores que irão compor o modelo, considerando predição um passo à frente.

A partir de trabalhos de Johansen (1996) e colaboradores, a área de identificação de sistemas começou a se preocupar com a possibilidade de utilizar outros tipos de informações sobre o sistema no modelo, e não apenas dados dinâmicos coletados. Dessa preocupação, surge uma nova técnica, chamada de Identificação Multi-objetivo de Sistemas (IMS) (Barroso et al., 2007; Johansen e Babuska, 2003; Nepomuceno, 2002; Johansen, 2000). Porém, tradicionalmente, o uso de informações auxiliares se dá exclusivamente na estimação dos parâmetros do modelo. A grande maioria das técnicas de detecção de estruturas são de caráter mono-objetivo.

Alguns trabalho, como (Martins et al., 2011; Barbosa, 2010) utilizam a abordagem multi-objetivo de determinação de parâmetros como ferramenta auxiliar na detecção de estruturas. Contudo, até onde foi verificado (Martins et al., 2011; Bonin et al., 2010; Barbosa, 2010; Alves, 2009; Wei e Billings, 2008; Lind e Ljung, 2008; Pulecchi e Piroddi, 2007; Lind e Ljung, 2005; Piroddi e Spinelli, 2003; Mendes e Billings, 2001; Aguirre e Billings, 1995a), não se encontrou técnicas de determinação de estruturas que fossem puramente multi-objetivo, ou seja, em que a abordagem multi-objetivo fosse aplicada diretamente para quantificar a contribuição de cada regressor.

Surge a partir daí a motivação para este trabalho, onde é proposta uma técnica multi-objetivo para detecção de estruturas de modelos NARMAX polinomiais. O desenvolvimento deste tipo de técnicas aplicada à detecção de estruturas é importante por muitos aspectos. Conjectura-se que dados dinâmicos apenas podem não ser suficientes para uma correta detecção de estruturas, dadas limitações em coleta de dados (limitações em tempo de amostragem, bem como número de dados disponíveis para identificação), ruídos dos processos e outros problemas. Espera-se que tais efeitos negativos possam ser minimizados com a incorporação de outras informações que se tem do sistema na estrutura do modelo.

1.1 Objetivo

O objetivo geral deste trabalho é desenvolver uma técnica multi-objetivo a ser aplicada e desenvolvida diretamente na etapa de detecção de estruturas de modelos NARMAX polinomiais.

Diferentemente das abordagens mono-objetivo, tais abordagens permitem

encontrar soluções intermediárias, levando em conta no modelo as diferentes características do sistema. O conjunto das soluções ótimas é denominado de conjunto Pareto-ótimo e apresenta soluções intermediárias entre os extremos do problema. Especificamente, será apresentada a técnica de taxa de redução de erro multi-objetivo (MERR, do inglês *Multi-objective Error Reduction Ratio*), aplicada a modelos NARMAX polinomiais. Mais uma vez, ressalta-se que nomenclatura NARMAX será utilizada para denominar modelos NARX ou NARMAX, uma vez que a técnica aqui desenvolvida é aplicável à ambas as classes.

1.2 Estrutura da Dissertação

Este trabalho esta dividido em seis capítulos. No Capítulo 1, são introduzidos conceitos de identificação de sistemas, as suas principais etapas e alguns conceitos teóricos básicos de identificação de sistemas via abordagem mono e multi-objetivo. Também foi apresentada a justificativa do tema investigado, compondo o objetivo do trabalho. Por fim, a estrutura da dissertação é apresentada.

É feita uma revisão teórica e bibliográfica no Capítulo 2 a respeito de identificação de sistemas, em particular em técnicas de detecção de estruturas de modelos NARMAX polinomiais. São estudados e apresentados alguns trabalhos recentes desenvolvidos na área e publicados por meio de periódicos e dissertações de mestrado.

Já no Capítulo 3 são apresentados conceitos básicos, necessários para o bom entendimento deste trabalho. Análise de propriedades de modelos NARMAX polinomiais, agrupamentos de termos, obtenção da curva estática, dentre outros conceitos são expostos de modo claro e objetivo. Técnicas de detecção de estrutura (a taxa de redução de erro - ERR - e o critério de informação de Akaike - AIC) também são apresentadas neste momento.

O Capítulo 4 consiste na apresentação da metodologia empregada para a realização deste trabalho. É feita uma introdução à técnica proposta, seguida da definição do problema multi-objetivo e da taxa de redução de erro multi-objetivo (MERR).

No Capítulo 5 são apresentados os principais resultados obtidos por meio da aplicação desta técnica em dois sistemas simulados, em um conversor CC-CC do tipo buck e em uma planta de neutralização de pH. Também é mostrado como a MERR pode ser aplicada quando se tem como objetivo a

incorporação de pontos fixos ou curva estática. Os resultados são comparados com técnicas presentes na literatura e é feita uma discussão crítica dos resultados. Por fim, é feita uma análise crítica da técnica apresentada.

Por fim, no Capítulo 6 são apresentadas as conclusões do trabalho e as considerações finais. Pesquisas para trabalhos futuros também são apresentadas, além de ser feito um resumo das principais contribuições deste trabalho.

Capítulo 2

Revisão de Literatura

"If you can't explain it simply, you don't understand it well enough[a]*."*

Albert Einstein

[a]Se você não pode explicar um problema de um modo simples, você ainda não o entendeu suficientemente bem.

Este capítulo apresenta alguns dos principais resultados obtidos nos últimos anos, referente às técnicas de detecção de estruturas de modelos NARMAX polinomiais. São apresentadas duas dissertações de mestrado desenvolvidas na área (Barbosa, 2010; Alves, 2009), bem como dois artigos de pesquisadores que atualmente têm pesquisado sobre o tema em questão (Pulecchi e Piroddi, 2007; Bonin et al., 2010).

Destaca-se que a maioria das técnicas apresentadas são de caráter mono-objetivo, sendo que somente Barbosa (2010) utilizou de técnicas multi-objetivo na determinação de parâmetros como ferramenta auxiliar na detecção de estruturas de modelos matemáticos. É digno de ressalva que até onde se pesquisou, não foram constatadas técnicas puramente multi-objetivos para detecção de estruturas capazes de quantificar a contribuição individual de cada regressor.

Deve ser ressaltado que foi feita uma extensa revisão bibliográfica (Martins et al., 2011; Bonin et al., 2010; Barbosa, 2010; Alves, 2009; Wei e Billings, 2008; Lind e Ljung, 2008; Pulecchi e Piroddi, 2007; Lind e Ljung, 2005; Piroddi e Spinelli, 2003; Mendes e Billings, 2001; Aguirre e Billings, 1995a) e que somente alguns trabalhos, considerados relevantes e de algum modo

associados com o desenvolvimento desta dissertação, são apresentados neste Capítulo.

2.1 (Pulecchi e Piroddi, 2007)

A princípio, os autores destacam a grande importância de se estudar modelos NARX polinomiais, bem como sua crescente quantidade de publicações ao decorrer dos anos. Ressaltam que, o processo de identificação utilizando este tipo de modelo, pode ser resumido em duas etapas principais: a escolha dos regressores que irão compor o modelo e a obtenção dos parâmetros, utilizados para quantificar a contribuição de cada regressor.

Em um segundo momento, é dito que o número de termos candidatos é grande em demasia, aumentando este exponencialmente, proporcional à não-linearidade do modelo. Deste modo, é necessário o desenvolvimento de novas técnicas para detecção de quais regressores devem ser inclusos no modelo do sistema. O autor enfoca que as técnicas presentes, ao contrário da apresentada neste trabalho, geralmente envolvem um custo computacional grande, pois necessitam de muitos cálculos, realizados de modo iterativo.

É então apresentada a técnica SEMP (*Simulation Error Minimization with Pruning*, proposta em (Piroddi e Spinelli, 2003), em que o modelo é obtido baseado na minimização do erro quadrático de simulação. Apesar de eficiente, apresenta como limitação o alto custo computacional, que é tratado em (Pulecchi e Piroddi, 2007), utilizando conceitos de agrupamento de termos.

A técnica apresentada é denominada pelo autor de CSI (*Cluster-based Structure Identification*) e, como o nome diz, baseia-se nos conceitos de agrupamento de termos para a obtenção de modelos NARX polinomiais. Possui como principais vantagens o baixo custo computacional e seleção automática dos regressores que descrevem as não-linearidades do modelo.

A técnica CSI apresenta dois principais estágios. No primeiro estágio, apenas um termo de cada agrupamento de termos é escolhido para compor o modelo. Agrupamentos espúrios são identificados e eliminados. Em um segundo momento, o modelo é refinado, com a inclusão de mais regressores, referentes aos agrupamentos já presentes no modelo.

A primeira etapa é feita baseada na técnica apresentada em (Aguirre e Billings, 1995a), em que agrupamentos com coeficientes de agrupamentos pequenos em relação à outros são agrupamentos espúrios e podem ser excluídos

do conjunto de termos candidatos. A partir daí, Pulecchi e Piroddi (2007) propõem métodos de expansão do modelo, de forma a refiná-lo e compor o modelo final do sistema, a saber:

1. Na Expansão Completa (CE - *Complet Expansion*), o estágio de refinamento do modelo é conduzido estritamente de modo sequencial. A cada inserção de termo, o modelo é reavaliado de acordo com o erro de simulação e decide-se de modo automático a permanência ou não deste termo. Todos os agrupamentos de uma ordem específica são considerados em um primeiro estágio. A seguir, todos os agrupamentos são expandidos para todos os atrasos de tempo disponíveis.

2. A Expansão Gradual (GE - *Gradual Expansion*) expande gradualmente, a ordem de não-linearidade considerada, a começar por 1, até o valor de ℓ, sendo ℓ o grau de não-linearidade máxima do modelo. Na i-ésima iteração, um termo por cada agrupamento de i-ésima ordem é adicionado ao conjunto de termos selecionados, obtidos por meio do modelo $(i-1)$. Avalia-se o modelo e decide-se pela inclusão ou não do regressor no modelo, por meio do erro de simulação.

3. A Expansão Iterativa (IE - *Iterative Expansion*) é similar à expansão completa. A princípio, todos os agrupamentos são especificados ao máximo grau do polinômio. Então, os agrupamentos selecionados são expandidos à todos os atrasos de tempo possíveis. Nesta abordagem, um procedimento iterativo é adotado, em que um agrupamento é expandido por vez e o modelo é refinado em relação aos termos somente deste agrupamento. O procedimento termina quando os agrupamentos não mais modificam o modelo.

A estratégia CE apresenta uma redução significante no conjunto de regressores candidatos utilizados pelo algoritmo de identificação. As abordagens GE e IE dividem o processo de identificação em vários passos, sendo que cada qual executa uma seleção de regressores em um conjunto de regressores candidatos extremamente compacto, beneficiando regressores de alta ordem, quando os mesmos devem ser inclusos no modelo. Contudo, as duas últimas estratégias consomem mais recursos computacionais. O autor mostra que as três abordagens são computacionalmente superiores à técnica SEMP, ou seja, demandam um menor custo computacional.

A técnica CSI é aplicada na identificação de um problema de análise de eventos sísmicos, simulado por meio de uma mesa que vibra de acordo com a intensidade do sinal de entrada. Segundo Pulecchi e Piroddi (2007), este tipo de processo requer uma técnica de modelagem consistente, de modo a descrever as dinâmicas do sistema quando o mesmo é excitado por valores de entradas de alto, média e baixa intensidades. As vibrações foram coletadas por meio de sensores acelerômetros, acoplados à bancada, quando excitados pelos diferentes valores de entrada citados anteriormente.

O autor iniciou os algoritmos (CSI e SEMP) com um modelo de estrutura linear, incluindo somente termos auto-regressivos. O máximo grau de não-linearidade foi definido como sendo 4 e o máximo atraso dos regressores de entrada e saída definidos como 5.

As duas abordagens apresentaram modelos e resultados similares, assim como satisfatórios tanto em termos de estruturas, quanto em termos de precisão, sendo que a técnica apresentada obteve uma estrutura mais compacta. O modelo obtido pelo algoritmo CSI apresenta uma pequena perda de qualidade dinâmica quando comparado à técnica SEMP. Porém, apresenta um grande ganho no custo computacional, em que o tempo de identificação é reduzido por um fator de 60% a 70%, quando comparado ao SEMP. Por fim, ressalta-se que a técnica CSI, quando empregada utilizando as expansões CE, GE e IE, gerou o mesmo modelo, sendo a estratégia de expansão CE a mais eficiente computacionalmente (reduziu o custo computacional de 10% a 30% relativo às expansões GE e IE).

Apesar da técnica ser eficiente no quesito custo computacional, possui como grande deficiência não considerar outros tipos de erros, senão o dinâmico. Isso pode prejudicar consideravelmente a representatividade do modelo, quando visto de um modo mais global, analisados em outras faixas de operação, em condições em que ruídos estejam presentes nos dados dinâmicos ou ainda quando apenas uma pequena massa de dados estiver disponível.

2.2 (Alves, 2009)

Consiste de uma dissertação de mestrado, desenvolvida no PPGEE (Programa de Pós-graduação em Engenharia Elétrica) da UFMG (Universidade Federal de Minas Gerais). Vale destacar que o PPGEE, em particular o grupo de pesquisa MACSIN (Modelagem, Análise e Controle de SIstemas Não-lineares), é um dos maiores centros de pesquisa brasileiros na área de

Identificação de Sistemas.

Neste trabalho, a autora apresenta a taxa de redução de erro dois passos a frente (ERR_2) aplicada a identificação de modelos NARX (Nonlinear Auto-Regressive model with eXogenous input) polinomiais, considerando um meio termo entre as técnicas ERR (*Error Reduction Ratio)* e SRR (*Simulation Error Ratio*). Tais abordagens se baseiam em erros de predição um passo à frente e livre, respectivamente.

No Capítulo 1, a autora introduz conceitos acerca de Identificação de sistemas, bem como dá a motivação e justificativa da pesquisa realizada. Já no Capítulo 2, conceitos básicos e bem conhecidos da área são apresentados e abordados de forma clara e objetiva. Cada uma das principais etapas da identificação de sistemas é discutida.

O Capítulo 3 torna-se de grande importância, pois apresenta uma vasta revisão bibliográfica acerca de técnicas de detecção de estruturas. A princípio, são apresentados dois critérios de informação frequentemente utilizados para determinação do número de regressores que serão incluídos no modelo (*Akaike Information Criteria* e a estatística *APRESS*). Ademais, ainda neste capítulo, são apresentadas a ERR e a SRR, bem como conceitos de agrupamentos de termos, frequentemente utilizados na detecção de estruturas. São destacados pontos positivos e negativos de todas as técnicas supracitadas.

Já o Capítulo 4 pode ser visto como um dos mais importantes do trabalho, pois apresenta a descrição da técnica utilizada. A princípio, o modelo de dois passos à frente é definido matematicamente, e explicitado sua obtenção por meio do modelo um passo a frente. A equivalência entre os modelos de um e dois passos à frente é provada matematicamente.

O resíduo gerado na predição de dois passos à frente, que consiste obviamente no erro de predição de dois passos à frente, dá origem ao novo critério de seleção de estrutura, denominado ERR_2 (taxa de redução de erro utilizando o resíduo de dois passos à frente).

A metodologia é apresentada em 8 principais etapas, a seguir, descritas:

1. Os dados de identificação são divididos em janelas, para identificação por meio de janelas deslizantes.

2. Os máximos atrasos e graus de não-linearidades são definidos. A técnica ERR é aplicada aos termos candidatos, estimando um modelo para cada janela de dados.

3. Obtém-se o modelo moda, que consiste no modelo formado pelos termos escolhidos com maior frequência pelo critério de seleção ERR.

4. O número *ideal* de regressores que devem compor o modelo é obtido por meio do critério de informação de Akaike ou da estatística APRESS.

5. Retornar ao item 2 utilizando a ERR_2, ao invés da ERR. Prosseguir com as etapas 3 e 4, obtendo um outro modelo moda (modelo de dois passos à frente).

6. Obter o modelo um passo à frente associado ao modelo moda obtido pela aplicação da ERR_2. Compara-se os modelos obtidos pelas taxas ERR e ERR_2 por meio de conceitos de auto-consistência.

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

8. Utilizar apenas os termos comuns ao modelos moda obtido pelo critério ERR e ao modelo de um passo à frente correspondente ao modelo moda obtido pelo ERR_2.

Nos Capítulos restantes, a técnica é empregada na identificação de três sistemas simulados e dois sistemas reais. Para os sistemas simulados, foram utilizadas entradas relativamente lentas, visando dificultar o processo de identificação. Como sistemas reais, foram utilizados um forno a gás e um pequeno aquecedor elétrico.

Verificou-se maior robustez na técnica ERR_2 quando comparada à ERR, uma vez que os termos corretos algumas vezes não foram selecionados por meio da técnica ERR. Contudo, foi verificado que a técnica apresentada é sensível às escolhas dos máximos atrasos, e pode não ser capaz de escolher corretamente os devidos regressores.

Ademais, destaca-se como grande ponto negativo a alta complexidade da técnica e seu grande custo computacional, uma vez que apresenta uma série de processos iterativos e vários cálculos matriciais. Outrossim, tal técnica mostra-se limitada por utilizar apenas um tipo de informação disponível dos sistemas.

2.3 (Barbosa, 2010)

Consiste de uma outra dissertação de mestrado desenvolvida dentro do grupo de pesquisa MACSIN, no PPGEE da UFMG. Desenvolveu-se um critério baseado em técnicas multi-objetivo para auxiliar na detecção de estrutura de modelos NARX polinomiais. O objetivo principal foi o de determinar a estrutura de um modelo NARX polinomial por meio da análise de curvas de Pareto, geradas pela variação combinacional de possíveis estruturas do modelo. A metodologia foi aplicada na identificação de três exemplos simulados e dois exemplos reais (um conversor CC-CC buck e uma válvula pneumática).

Os primeiros dois Capítulos apresentam uma introdução ao problema de identificação de sistemas e a fundamentação teórica necessária para a compreensão do conteúdo restante do trabalho. Conceitos básicos de modelos polinomiais NARMAX são apresentados, bem como conceitos relativos a agrupamento de termos, frequentemente utilizado como ferramenta auxiliar na detecção de estruturas. Conceitos sobre a representação da curva estática, bem como sua incorporação na etapa de determinação dos parâmetros também são apresentados, uma vez que foram utilizados nos capítulos subsequentes.

O Capítulo 3 apresenta a metodologia proposta para detecção de estruturas de modelos NARX polinomiais, via abordagem multi-objetivo. A técnica é apresentada em duas rotinas (**Rotina I** e **Rotina II**), a seguir apresentadas. Ressalta-se que ferramentas multi-objetivo são utilizadas somente na determinação de parâmetros, sendo esta utilizada como ferramenta auxiliar para determinação de estruturas.

ROTINA I

1. Considere uma condição inicial de n estruturas candidatas.

2. Para $i = 1$ ate n, faça:

   Determine o conjunto Pareto-ótimo dos parâmetros da i-ésima estrutura, utilizando a abordagem multi-objetivo para determinação dos parâmetros.

3. Identifique, por meio dos conjuntos Pareto, os agrupamentos potencialmente espúrios.

4. Eliminar os termos pertencentes aos agrupamentos potencialmente espúrios.

ROTINA II

1. Seja $M_1, \cdots, M_m$ um conjunto previamente selecionado de estruturas candidatas.

2. Para $i = 1$ ate n, faça:

   Determine o conjunto Pareto-ótimo dos parâmetros da i-ésima estrutura;

   Determine o modelo de mínima autocorrelação, dentre os modelos obtidos pela etapa anterior.

3. Escolher, o modelo que apresentar maior sincronismo.

O algoritmo apresentado pela Rotina I tem o objetivo de discriminar os agrupamentos espúrios dos genuínos. A Rotina II é uma continuação da Rotina I e visa *filtrar* o modelo até então obtido. Ou seja, busca-se eliminar todos os termos espúrios e fazer com que o modelo apresente somente termos genuínos. A partir daí, obtém-se a estrutura do modelo por abordagem multi-objetivo, via análise de curvas de Pareto.

O primeiro exemplo simulado teve como objetivo mostrar um possível comportamento das curvas Pareto para modelos de estruturas diferentes. Já o segundo busca averiguar a aplicação da técnica, bem como caracterização de incertezas e particularidades. Por fim, o terceiro exemplo simulado sugere a avaliação da técnica quando o sinal de entrada é persistentemente excitante, porém de ordem menor. Essa é uma situação menos favorável, importante para avaliar a robustez do método proposto.

No primeiro exemplo simulado, ficou evidente o comportamento das curvas de Pareto a partir dos funcionais estático e dinâmico. Conjecturou-se que curvas de Pareto construídas a partir de outros funcionais, conflitantes entre si, resultariam em resultados similares.

O segundo exemplo provou que é possível discriminar estruturas subparametrizadas por meio do critério bi-objetivo e que, dado um modelo sobreparametrizado que contenha a estrutura nominal, é possível, por meio da técnica apresentada, discriminar termos espúrios e genuínos. Também foi

mostrada uma região de incerteza, na qual estruturas sobre-parametrizadas, a princípio, seriam equivalentes entre si.

No terceiro caso, o procedimento de identificação proposto mostrou-se eficiente na escolha dos termos pertinentes ao modelo, mesmo em condições não ideais de identificação.

O primeiro sistema real identificado foi um conversor CC-CC buck, quando excitado por um sinal PRBS (Sinal Binário Pseudo-Aleatório, do inglês *Pseudo Random Binary Signal*). Acredita-se que este sinal de excitação é capaz de excitar o sistema nas mais diferentes componentes de frequências. A abordagem empregada neste sistema propiciou resultados dinâmicos similares aos encontrados na literatura. Contudo, apresenta como grande vantagem a melhora na representação estática.

O segundo sistema real foi uma válvula pneumática, consistente em um sistema de bombeamento de água e controle de nível. A planta é composta por um reservatório que permite o envio de água para um tanque cilíndrico de diâmetro uniforme. A vazão é manipulada por uma válvula de atuação pneumática. Para este caso, os resultados foram satisfatórios, quando avaliados mediante diferentes métricas. Contudo, não pôde ser comparado com outros resultados, uma vez que a planta ainda não tinha sido identificada.

Deve-se ressaltar o alto custo computacional da técnica empregada. A mesma é composta por diversos processos iterativos e vários cálculos matriciais, incluindo inversões de matrizes, gerando um alto gasto computacional. Vale lembrar que o custo computacional cresce exponencialmente com o aumento dos termos candidatos.

Por fim, observa-se como caráter inovador a abordagem multi-objetivo para determinação dos parâmetros como ferramenta auxiliar na detecção de estruturas de modelos NARX polinomiais. Todavia, ainda nesta técnica não se obtém os regressores diretamente de uma abordagem multi-objetivo, pois usa-se de determinação dos parâmetros via técnicas multi-objetivos para escolher os regressores.

Necessita-se de uma abordagem que considere outras informações diretamente na determinação de estruturas, visando a obtenção de um modelo mais global de um dado sistema. Visando preencher tal lacuna, desenvolveu-se o trabalho apresentado nesta dissertação, que expande a ERR de modo que a mesma permita a incorporação de outras características do sistema no modelo.

2.4 (Bonin et al., 2010)

O autor apresenta técnicas mono-objetivo de detecção de estruturas para modelos NARX polinomiais. Como visto em um grande número de trabalhos, vários sistemas podem ser modelados precisamente por meio de modelos NARX polinomiais com baixa complexidade. Contudo, o autor afirma que se a estrutura do modelo não é conhecida *a priori*, o número de termos candidatos se torna extremamente grande. Portanto, cada vez mais são necessárias aprimoramentos em técnicas de detecção de estruturas de modelos matemáticos.

A grande maioria das técnicas de detecção de estruturas de modelos NARX polinomiais são da classe PEM (*Preditcion Error Minimization*). Recentemente, a técnica SEM (*Simulation Error Minimization*) foi proposta, em que utilizava-se o erro de simulação (predição livre) para obter os regressores que comporiam o modelo. Deste modo, pode-se aumentar a robustez do processo de seleção e do modelo, quando em condições de identificação não ideias.

Contudo, infelizmente um alto custo computacional é requerido para a aplicação da técnica SEM. Visando reduzir este esforço, foi proposta a técnica SEMP (*Simulation Error Minimization with Pruning*) em seus trabalhos anteriores.

Alguns outros trabalhos investigam o uso de informações contidas nos agrupamento de termos, de forma a reduzir o tamanho do conjunto de termos candidatos, como o caso da técnica LASSO (*Least Absolute Shrinkage and Selection Operator*). Deste modo o presente trabalho estuda a combinação da técnica LASSO com a abordagem SEM, visando obter uma melhor combinação de regressores, a um menor custo computacional.

Os resultados mostram que enquanto a técnica LASSO eficientemente elimina vários regressores redundantes, sua performance de seleção é altamente dependente do peso do termo de regularização da função custo. Isso sugere que a técnica LASSO não possa oferecer como resultados regressores suficientemente eficientes para identificar o sistema. Desse modo, pode-se combinar de um modo eficiente a técnica LASSO com a técnica SEM, formando a técnica SEM-LASSO.

O algoritmo SEM-LASSO a princípio utiliza das características do método LASSO como uma ferramenta de pré-escaneamento, para a eliminação de termos redundantes do conjunto de termos candidatos. Deste modo, o custo

computacional (elevado na técnica SEM) é consideravelmente reduzido. A partir daí, aplica-se a técnica baseada em erro de simulação (SEM) de modo a obter o modelo final.

A técnica foi aplicada em dois exemplos simulados e um sistema real. O primeiro sistema consiste de um modelo NARX gerador de dados, utilizando como entrada um ruído branco com característica gaussiana de média 0 e variância 1 e como ruído um ruído branco com média zero e variância 0,33. Sistemas simulados são ferramentas eficientes para detectar a qualidade da técnica, uma vez que, ao contrário de sistemas reais, a estrutura do modelo é totalmente conhecida.

O método empregado foi capaz de identificar o modelo do sistema a um custo computacional bastante reduzido, quando comparado à técnica *SEMP* (também capaz de identificar o modelo do sistema).

O segundo exemplo explora o fato da técnica não somente ser benéfica no aspecto computacional, mas também na melhora de precisão do modelo, ao identificar um sistema. Dessa vez, o mesmo sistema do é identificado, modificando a entrada para um filtro passa-baixas composto pela combinação de um sinal auto-regressivo com um ruído gaussiano branco com média 0 e desvio padrão unitário.

Segundo o autor, este tipo de excitação pode confundir o algoritmo seletor reduzindo a diferença de termos regressores como $y(t)$ e $y(t-1)$. Neste caso, o algoritmo *SEMP* não foi capaz de recuperar a estrutura correta do modelo. Já a técnica *SEM-LASSO* reconstrói o modelo, com um acréscimo de 2 termos redundantes. Todavia, os parâmetros associados aos termos redundantes são irrelevantes, não contribuindo para a dinâmica da evolução do modelo. Por fim, um terceiro sistema eletrônico não-linear, SISO, foi identificado. Por meio de métricas que levam em conta o erro de simulação, o algoritmo foi validado e o modelo considerado representativo do sistema.

A técnica apresentada mostrou-se bastante eficiente no aspecto dinâmico, na identificação de sistemas reais e simulados. Em particular nos sistemas simulados, a mesma conseguiu resgatar as estruturas originais do modelo. Contudo, ainda possui o inconveniente de somente considerar dados dinâmicos durante a obtenção da estrutura do modelo, sendo que características importantes do sistema podem não ser levadas em conta.

2.5 Conclusões do Capítulo

O presente Capítulo apresentou detalhadamente alguns trabalhos na área de identificação de sistemas, em particular na etapa de detecção de estruturas de modelos. Outros trabalhos também foram estudados, mas não expostos de forma detalhada nesta dissertação.

Pulecchi e Piroddi (2007) apresentaram nova abordagem baseada em agrupamento de termos e erros de simulação para detecção de estruturas de modelos NARX polinomiais. Já Alves (2009) elabora a ERR_2 (taxa de redução de erro dois passos a frente), considerando um meio termo entre a ERR e a SRR, que consideram erros de predição um passo à frente e erros de simulação, respectivamente.

Estas técnicas possuem como pontos negativos comum o fato de considerarem uma pequena quantia de informações acerca do sistema para a obtenção da estrutura. Ademais, a técnica apresentada em (Alves, 2009) é cara, do ponto de vista matemático e computacional. A técnica apresentada em (Bonin et al., 2010) também apresenta limitações semelhantes às anteriores, por ser de caráter puramente mono-objetivo e consiste de uma combinação de duas técnicas já existentes (SEM-LASSO).

Por fim, a abordagem de Barbosa (2010), apesar de eficiente à luz das métricas consideradas naquele trabalho, engloba altos custo e complexidade computacional para determinar a estrutura. Utiliza a determinação multi-objetivo de parâmetros como ferramenta auxiliar para obter a estrutura do modelo, mas ainda assim não consiste de uma técnica puramente multi-objetivo para obtenção da estrutura do modelo.

Deste modo, encontra-se uma grande lacuna nas técnicas de detecção de estruturas de modelos NARMAX polinomiais, que é a quantificação de regressores considerando diversas informações que se tem acerca do sistema. Esta dissertação explora justamente este aspecto, e propõe uma nova taxa de redução de erro multi-objetivo, capaz de quantificar regressores utilizando diferentes tipos de características dos sistemas.

Capítulo 3

Conceitos Preliminares

"(· · ·) o bravo é quem suporta mais pancada e continua em pé, sem perder a própria fé!"

Terminal Guadalupe - Chico Balboa

Este capítulo apresenta os conceitos preliminares necessários para melhor entendimento dos capítulos posteriores. Serão apresentados conceitos de modelos NARMAX polinomiais e suas propriedades, bem como técnicas de detecção de estruturas que foram utilizadas como técnicas motivadoras para o desenvolvimento da técnica apresentada nesta dissertação.

3.1 Identificação de Sistemas

Para se identificar um sistema qualquer, geralmente utiliza-se de 5 (cinco) principais etapas, descritas em (Ljung, 1987), a saber:

1. Teste Dinâmico e Coleta de Dados

   Para se obter modelos matemáticos, são necessários dados. Para se coletar tais dados, necessita-se de um estudo acerca das frequências componentes do sinal a ser coletado, de maneira a determinar taxas de amostragens, dentre outros parâmetros necessários.

   Dessa forma, o que é feito, em geral, é a execução de testes de modo a obter dados para identificação e validação do modelo, separados em

duas massas distintas de dados. Nesta etapa, deve-se procurar minimizar os efeitos do ruído nos sinais coletados.

Para tal, aplica-se um sinal de excitação que tenham espectros de frequência largos o bastante para excitar persistentemente a dinâmica de interesse do sistema. Para o caso de sistemas não-lineares, tais efeitos devem ser excitados por estes sinais, de modo a estarem contidos nos dados utilizados para identificação e validação.

Um sinal geralmente utilizado para excitação é o sinal binário pseudo-aleatório (PRBS - do inglês *Pseudo Random Binary Signal*). Uma descrição detalhada deste sinal pode ser encontrada em (Aguirre, 2007). Este sinal foi utilizado neste trabalho, para a excitação dos sistemas reais identificados.

Uma vez excitado, faz-se necessária a coleta de dados do sistema. Sendo assim, é de extrema importância a escolha correta do tempo de amostragem (T_s) do sinal. A maioria dos processos reais são contínuos, contudo, há a necessidade de se fazer esta amostragem, visando tratamento e análise mais aprofundados do sinal.

O período entre amostras é conhecido como tempo de amostragem (T_s), e o mesmo deve ser suficientemente curto, de modo que o sinal possa ser reconstruído *a posteriori*. De acordo com o teorema de Shannon (Shannon, 1949), para que o sinal possa ser reconstruído, a frequência de amostragem ($f_s = T_s^{-1}$) deve ser ao menos duas vezes a maior componente de frequência do sinal original. Contudo, em situações práticas, essa frequência é de 5 a 10 vezes a maior componente de frequência do sinal, sendo que o sinal é posteriormente decimado, gerando o sinal de trabalho. Técnicas baseadas em correlação linear e não linear para a decimação do sinal também podem ser encontradas em (Aguirre, 2007).

2. Escolha da Representação Matemática a ser Utilizada

   Existem na literatura, várias representações matemáticas de sistemas reais. Dentre elas, pode-se citar as representações de Hammerstein e de Wiener, representações NARMAX (Leontaritis e Billings, 1985a,b) polinomiais, neurais, racionais, dentre outras. Em particular, modelos NARMAX polinomiais são bastante utilizados, pois permitem com relativa facilidade a incorporação de informações *a priori* de um sistema no modelo, bem como a extração de informações analíticas do

modelo. Neste sentido, tal representação será utilizada neste trabalho e apresentada com mais detalhes na próxima seção.

3. Determinação da Estrutura do Modelo

   É a etapa mais difícil da Identificação de Sistemas, por conter um universo de busca bastante extenso. Fenômenos como sub ou sobre-parametrização [1] podem ocorrer se a estrutura não for escolhida de forma adequada, piorando a representatividade do modelo. Na literatura, tem-se, em sua maioria, técnicas mono-objetivo para a determinação de estruturas. Estas são complexas e são em menor quantidade. Ademais, levam em conta para determinação de estruturas somente questões relativas ao erro de predição um passo a frente, deixando de lado outras características do sistema.

4. Estimação dos Parâmetros

   É nessa etapa que se quantifica cada regressor da estrutura obtida na etapa antecedente. Podem ser feitas abordagens mono ou multi-objetivo na determinação dos parâmetros de um modelo. Dentro do âmbito da otimização em projeto assistido por computador a otimização escalar emprega, classicamente, a abordagem mono-objetivo. Isso significa que a figura de mérito que o mecanismo de otimização deve minimizar é um funcional (uma função cuja imagem é escalar). A otimização vetorial ou otimização multi-objetivo parte da constatação de que, havendo diversos objetivos em questão num problema de projeto, haverá soluções que, sob todos os objetivos simultaneamente considerados, serão suplantadas por outras soluções (Takahashi, 1998). Haverá ainda soluções que, comparadas com outras soluções, serão melhores em algum ou alguns objetivos, mas piores em outros . As soluções desse último grupo são denominadas soluções eficientes ou soluções Pareto-ótimas. Os parâmetros de modelos matemáticos podem ser obtidos de modo eficiente utilizando qualquer uma das abordagens citadas neste tópico.

[1] Sobre-parametrização é o fenômeno que ocorre quando se tem em um modelo de um sistema mais regressores do que o necessário. Isso faz com que mais de um regressor seja responsável por descrever a mesma característica dinâmica do sistema, sendo que os mesmos em conjunto, não representem adequadamente o sistema. Já a sub-parametrização é o fenômeno causado por um número pequeno de regressores no modelo, sendo que os mesmos não são capazes de descreverem as características do sistema.

5. Validação dos Modelos

   Quando se obtém modelos de sistemas reais, é necessário que o mesmo seja validado. Em geral, pode-se calcular alguns índices como o *RMSE* estático e dinâmico (raiz quadrada da média dos erros quadráticos, do inglês *Root Mean Squared Error*). Outros métodos de validação serão discutidos ainda neste capítulo, tal como a fidelidade do comportamento estático do modelo do sistema, bem como as funções custo estático e dinâmico, compostas pelo somatório dos erros quadráticos de predição.

3.2 Abordagens Mono e Multi-objetivo Aplicadas à Identificação de Sistemas

A Identificação de Sistemas pode ser vista sob a ótica de otimização, onde deseja-se minimizar funcionais associados aos erros cometidos pelo modelo relativo a um sistema, em suas diversas peculiaridades. A otimização mono-objetivo é dada nos casos em que o funcional que deve ser otimizado é um escalar, ou seja, apenas uma função objetivo J. Já na abordagem multi-objetivo, o funcional que será minimizado é composto por n funcionais J, que devem ser simultaneamente otimizados. A seguir, tais abordagens serão descritas de modo mais apropriado. Podem ser utilizadas tanto na determinação de estruturas, quanto na determinação de parâmetros do modelo.

3.2.1 Otimização Mono-objetivo

Neste tipo de abordagem, geralmente o funcional a ser minimizado é composto pelo erro quadrático dinâmico do modelo. Um clássico exemplo da aplicação da otimização mono-objetivo se encontra no método dos mínimos quadrados apresentado neste capítulo, onde a função custo a ser minimizada, apresentada na equação (3.21), é composta pelo erro quadrático dinâmico de predição, dado um passo a frente.

Neste processo, obtém-se um conjunto de parâmetros, os quais minimizam o erro quadrático de predição um passo à frente referente à massa de dados de identificação. A técnica é dada como mono-objetivo por considerar somente um objetivo (erro de predição) para a estimação dos parâmetros do modelo.

Outro exemplo de otimização mono-objetivo, dessa vez aplicada a obtenção de estruturas de modelos é a abordagem ERR. Nela, obtém-se os regressores de modo a minimizar a variância normalizada dos dados dinâmicos do sistema.

3.2.2 Otimização Multi-objetivo

Analisada à luz da Identificação de Sistemas, a otimização multi-objetivo considera um vetor de funções custo, composto como segue:

$$\mathbf{J}(\theta) = \begin{bmatrix} J_1(\theta) \\ \vdots \\ J_n(\theta) \end{bmatrix}. \tag{3.1}$$

Cada funcional $J_1(\theta) \dots J_n(\theta)$ representa uma função a ser otimizada, sendo que o problema se define de tal modo que todas as funções devam ser otimizadas, simultaneamente.

Os funcionais apresentados na equação (3.1) podem ser compostos por associações a diversas características do sistema envolvido, como ganho estático, curva estática, informações acerca do ponto fixo, dentre outras características importantes do sistema. Uma abordagem bastante utilizada, é a composição de uma função objetivo pela combinação linear convexa das funções objetivos individuais:

$$J_c = \omega_1 J_1(\theta) + \omega_2 J_2(\theta) + \cdots + \omega_n J_n(\theta). \tag{3.2}$$

em que ω_i, $i = 1, \cdots, n$ são ponderações entre 0 e 1, tal que $\omega_1 + \cdots + \omega_n = 1$. Deste modo, o problema se assemelha a um problema de otimização mono-objetivo, em que as variáveis $\omega_1, \omega_2, \cdots, \omega_n$ são variadas de forma a gerar o conjunto de soluções eficientes (conjunto Pareto-ótimo). A existência de um conjunto de soluções eficientes é dada uma vez que, em geral, os objetivos são conflitantes e não podem ser otimizados simultaneamente. O conjunto de soluções eficientes apresenta as *melhores* soluções possíveis, no âmbito do processo de otimização.

3.3 Modelos NARMAX

Modelos NARMAX descrevem sistemas não-lineares em equações de diferença, relacionando a entrada atual em combinação das saídas e entradas passadas. Este tipo de abordagem pode ser utilizada para problemas de controle onde o principal objetivo é encontrar uma descrição simples para o sistema. Em particular, o modelo NARMAX polinomial pode ser representado como:

$$\begin{aligned} y(k) &= F^{\ell}[y(k-1), \cdots, y(k-n_y), \\ &\quad u(k-1), \cdots u(k-n_u), \\ &\quad e(k-1), \cdots, e(k-n_e)] + \Xi(k), \end{aligned} \tag{3.3}$$

em que $y(k)$ é a saída, $u(k)$ é a entrada exógena, e $e(k)$ é o sinal de ruído. $\Xi(k)$ representa o erro de predição. n_y , n_u , e n_e são as ordens da saída, da entrada exógena e do ruído modelado por processo de média móvel, respectivamente. A função F^{ℓ} pode representar uma grande variedade de funções, incluindo funções lineares. Neste trabalho, F^{ℓ} é restrita a funções polinomiais não-lineares. $k = 1, \cdots, N$, sendo $N \in \mathbb{Z}^+$ o número de passos discretos.

Podem ser destacadas grandes características deste tipo de representação matemática, tal como a baixa complexidade. Outra característica a ser explorada neste trabalho é a relativa facilidade de se incorporar informações auxiliares no modelo, bem como a facilidade na obtenção de informações analíticas do sistema por meio do modelo.

3.4 Propriedades dos Modelos NARMAX Polinomiais

Como dito anteriormente, uma das justificativas de se utilizar modelos NARMAX polinomiais na identificação de sistemas é a sua alta capacidade de incorporação e extração de informações auxiliares no modelo. As subseções seguintes descreverão algumas de suas propriedades.

3.4.1 Agrupamento de Termos

Em regime permanente, as entradas e saídas de um modelo NARMAX polinomial podem ser expressas como:

$$\begin{aligned} y(k-1) &= y(k-2) = \cdots = y(k-n_y), \\ u(k-1) &= u(k-2) = \cdots = u(k-n_u). \end{aligned} \tag{3.4}$$

Substituindo a equação (3.4) na equação (3.3), vem que (Aguirre e Billings, 1995b):

$$y(k) = \sum_{n_1,n_m}^{n_y,n_u} c_{p,m-p}(n_1,...,n_m) \sum_{m=0}^{\ell} \sum_{p=0}^{m} y(k-1)^p u(k-1)^{m-p}, \tag{3.5}$$

sendo o conjunto de termos $y(k-1)^p u(k-1)^{m-p}$ denominado de agrupamento de termos. As constantes $\sum_{n_1,n_m}^{n_y,n_u} c_{p,m-p}(n_1,...,n_m)$ são definidas como coeficientes de agrupamentos. De acordo com (Aguirre e Billings, 1995b), todos os termos pertinentes a um dado agrupamento explicam o mesmo tipo de não-linearidade do modelo. Neste sentido, Aguirre (1997) propôs uma técnica que pode ser utilizada como ferramenta auxiliar na detecção de estruturas de modelos NARMAX polinomiais. Ademais, Pulecchi e Piroddi (2007) propuseram uma técnica baseada em agrupamento de termos para detecção de estruturas deste mesmo tipo de modelo.

Detalhes mais aprofundados a respeito da importância de agrupamento de termos como ferramenta auxiliar de detecção de estruturas podem ser encontrados em (Aguirre, 2007).

3.5 Determinação da Estrutura

O grau de não-linearidade de modelos está diretamente relacionado à dificuldade de determinação de estruturas. Isso pode ser observado, já que a expansão do espaço de busca é dado de modo exponencial, de acordo com o aumento do grau de não-linearidade e ordem máxima dos regressores.

Uma vez que a grande maioria dos sistemas reais apresenta alto grau de não-linearidade em seus comportamentos, tem-se um universo de busca bastante extenso, quando da determinação de estruturas de modelos. Neste trabalho serão somente estudadas e apresentadas técnicas para detecção de estruturas de modelos NARMAX polinomiais.

Alguns dos fatores que mais influenciam no tamanho do espaço de busca

é o grau de não-linearidade e os máximos atrasos dos regressores de entrada e saída (ℓ, n_y e n_u, respectivamente). O aumento de qualquer uma destas grandezas provoca um aumento exponencial na quantidade de termos candidatos. Estes, por sua vez, podem ser determinados pela seguinte expressão (Korenberg et al., 1988):

$$n_\theta = M + 1, \tag{3.6}$$

em que n_θ é o número de termos (de processo e de ruído) no modelo e

$$M = \sum_{i=1}^{\ell} n_i, \tag{3.7}$$

$$n_i = n_{i-1} \frac{(n_y + n_u + n_e + i - 1)}{i}, n_0 = 1. \tag{3.8}$$

Devido ao aumento exponencial do número de termos candidatos, tornam-se necessárias técnicas para determinar quais regressores, se inclusos no modelo, melhoram sua representatividade. Considere um modelo linear ($\ell = 1$), com máximo atraso na saída, entrada e ruído de valor 2 ($n_y = n_u = n_e = 2$). Neste caso, o número de termos candidatos seria de $n_\theta = 7$. Contudo, ao aumentar somente uma unidade no grau de não-linearidade ($\ell = 2$) e nos máximos atrasos ($n_y = n_u = n_e = 3$), o tamanho do espaço de busca cresce para $n_\theta = 55$. Ao considerar como 6 o grau de não-linearidade e os máximos atrasos ($\ell = n_y = n_u = n_e = 6$) tem-se o número de termos candidatos da ordem de 10^5 (exatos 134.596 termos candidatos!). Tal aumento na quantidade de termos candidato justifica a pesquisa na área de detecção de estruturas.

Se escolhidos de forma equivocada, os regressores podem causar fenômenos como sub ou sobre-parametrização do modelo, sendo sua representatividade deteriorada devido a tais fenômenos. Ademais, por apresentar um grande número de termos candidatos, a busca por força bruta mostra-se por várias vezes inadequada e de alto custo computacional.

O restante desta seção descreverá duas das principais técnicas de detecção de estruturas de modelos NARMAX polinomiais, que geralmente são empregadas em conjunto. Tais técnicas serão utilizadas como critério comparativo com a técnica apresentada neste trabalho.

3.5.1 Determinação da Estrutura via ERR

A taxa de redução de erro - ERR (do inglês *Error Reduction Ratio*) é uma das técnicas utilizadas para detecção de estruturas de modelos NARMAX polinomiais. Tal técnica baseia-se no erro dinâmico de predição, quando visto um passo à frente, associando cada termo candidato à um índice correspondente à contribuição na explicação da variância dos dados de saída do sistema. A variância dos dados de saída pode ser definida como:

$$\mathrm{var}[\xi(k)] = \lim_{N\to\infty} \frac{1}{N}\left[y^T y - \sum_{i=1}^{N} g_i^2 w_i^T w_i\right], \tag{3.9}$$

em que g_i indica cada elemento do vetor de parâmetros g, w_i representa os regressores do modelo ortogonal associado e y representa a série temporal dos dados de saída e N o tamanho da série. O modelo ortogonal associado pode ser obtido por meio de transformações de Householder do modelo original (maiores detalhes a respeito dessa transformação em (Aguirre, 2007)).

Caso nenhum termo fosse acrescentado ao modelo, a variância $\xi(k)$ seria exatamente igual ao erro quadrático de saída ($N = 0$). A cada termo acrescentado ao modelo, a variância decresce um valor de $\frac{1}{N}(g_i^2 w_i^T w_i)$, sendo w_i o regressor incluído e g_i o respectivo parâmetro ortogonal associado.

O fator de redução da variância devido a inclusão de um dado termo pode ser normalizada em relação ao erro quadrático médio do sinal de saída. Deste modo, tem-se a ERR como sendo:

$$[ERR] = \frac{(g_i^2 w_i^T w_i)}{y^T y}. \tag{3.10}$$

A taxa de redução de erro pode ser utilizada para quantificar cada regressor, ou seja, qual a parcela de cada regressor na contribuição da redução da variância dos dados de saída. Contudo, acredita-se que o número de termos que devem ser inseridos no modelo não possa ser determinado por esta técnica, sendo necessário a utilização de uma técnica auxiliar, como o critério de informação de Akaike (AIC), por exemplo.

Por fim, ressalta-se como grande fator limitante da técnica a utilização exclusiva de dados dinâmicos e das informações ali contidas, para a quantificação da contribuição de cada regressor, para a composição do modelo final.

3.5.2 O Critério de Informação de Akaike - AIC

Alguns dos procedimentos que podem ser utilizados para a determinação do número de termos que devem ser inclusos no modelo são os critérios de informação. Em particular, o critério de informação de Akaike fornece o número de regressores que devem ser inclusos no modelo, de forma que o mesmo represente de forma adequada o sistema. Quando aplicada em conjunto com a ERR (que classifica os termos de modo hierárquico), pode fornecer um modelo, baseado em dados dinâmicos, para representar um sistema.

De acordo com esta técnica, o número de termos de um modelo deve minimizar a função custo J_{AIC}:

$$J_{AIC} = N\log(\text{var}[\xi(k)]) + 2n_\theta, \tag{3.11}$$

em que N é o tamanho da série temporal dos dados de identificação, n_θ o número de regressores candidatos ao modelo e $\text{var}[\xi(k)]$ é a variância do erro de modelagem (resíduos ou erro de predição um passo a frente). O resultado obtido pelo AIC pode ser visto como um número *ótimo* de termos que devem ser inseridos no modelo do sistema (Barbosa, 2010; Nepomuceno, 2002).

3.6 Os Estimadores LS e ELS

Uma vez determinada a estrutura do modelo, faz-se necessária a determinação dos parâmetros associados a cada regressor, de modo a quantificá-los. Várias técnicas podem ser empregadas, sendo a técnica dos mínimos quadrados uma das mais conhecidas e estudadas. Uma extensão desta técnica, os mínimos quadrados estendidos, também podem ser utilizados de modo a evitar a polarização dos parâmetros.

3.6.1 O Método dos Mínimos Quadrados (LS)

Para obter os parâmetros de um modelo polinomial do tipo apresentado na equação (3.3), é necessário que a mesma seja expressa na forma de regressão linear, ou seja:

$$y(k) = \psi^T(k-1)\hat{\theta} + \xi(k), \tag{3.12}$$

sendo $\psi^T(k-1)$ todos os regressores do modelo, quando tomados até o ins-

tante $(k-1)$, $\hat{\theta}$ o vetor de parâmetros a ser estimado e $\xi(k)$ são os resíduos ou erro de predição um passo a frente.

Quando o modelo apresentado na equação (3.12) é obtido sobre uma massa de dados finita, o mesmo pode ser representado por uma equação matricial, de acordo com a equação (3.13):

$$y = \Psi\hat{\theta} + \xi, \tag{3.13}$$

sendo Ψ conhecida como matriz de regressores. Pode-se provar matematicamente (Aguirre, 2007) que o vetor de parâmetros que minimiza a função custo composta pelo erro quadrático de predição $\xi^T\xi$ é dado por:

$$\hat{\theta}_{MQ} = [\Psi^T\Psi]^{-1}\Psi^T y, \tag{3.14}$$

em que y é o vetor de dados, $\Psi^T\Psi$ é a matriz de informação e $[\Psi^T\Psi]^{-1}\Psi^T$ é a matriz pseudo-inversa.

Caso a estimativa dos parâmetros não seja adequada, ou seja, se os parâmetros obtidos não forem os parâmetros *reais* do sistema, a estimativa é dita polarizada e os resíduos apresentam alguma dinâmica que não foi devidamente explicada pelo modelo. Matematicamente, tem-se:

$$E[\hat{\theta}_{MQ}] \neq \theta, \tag{3.15}$$

sendo E o operador esperança. Existem na literatura estimadores que permitem a estimação não-polarizada dos parâmetros. Um estimador pertencente à esta classe é o estimador de mínimos quadrados estendidos (ELS), a ser descrito na próxima subseção.

3.6.2 O Método dos Mínimos Quadrados Estendidos (ELS)

A polarização na estimação de parâmetros acontece quando o resíduo apresentado na equação 3.13 for autocorrelacionado e o modelo incluir regressores de saída, simultaneamente. A partir daí, é importante ressaltar que se os dados puderem ser modelados utilizando somente regressores de entrada, o estimador MQ é não-polarizado, independentemente do tipo de ruido presente nos dados.

Para evitar este tipo de efeito na estimação dos parâmetros, pode-se utilizar o método dos mínimos quadrados estendidos. Neste método, o resíduos

são modelados por um processo média móvel, do tipo:

$$e(k) = c_i\nu(k-i) + \nu(k), \tag{3.16}$$

sendo $\nu(k)$ ruído branco. Os termos de $\nu(k-i)$ podem então serem incorporados à matriz de regressores e os parâmetros associados à estes termos, no vetor de parâmetros, formando o modelo estendido.

$$y^* = \Psi^*\theta^* + e^*, \tag{3.17}$$

em que $y^* = y$, $e^* = [\nu(k) \cdots \nu(k+N-1)]^T$,

$$\Psi^* = \begin{bmatrix} & \vdots & \nu(k-1) \\ & \vdots & \nu(k) \\ \Psi & \vdots & \nu(k+1) \\ & \vdots & \vdots \\ & \vdots & \nu(k+N-2) \end{bmatrix} \tag{3.18}$$

e $\theta^* = [\theta \vdots c_i]^T$. Os regressores referentes à c_i foram inclusos na matriz de regressores estendida Ψ^*. Deste modo, o ruído e^* apresentado em (3.17) é do tipo branco, e não correlacionado com a matriz de regressores, que contém regressores de saída. Sendo assim, a estimativa por MQ da equação (3.17) não apresenta polarização, e pode ser dada por:

$$\hat{\theta}^* = [\Psi^{*T}\Psi^*]^{-1}\Psi^{*T}y. \tag{3.19}$$

Deste modo, tem-se que os parâmetros estimados são aproximados dos parâmetros *reais* do sistema. Deve-se ressaltar que geralmente o ruído é utilizado apenas para evitar a polarização na estimação de parâmetros de modelos NARMAX polinomiais, sendo que o modelo final contém apenas termos de entrada e saída do processo.

Obviamente, os parâmetros não podem ser estimados diretamente como da forma apresentada em 3.19. Isto pois o ruído $\nu(k-1)$ que compõe a matriz estendida de regressores Ψ^* não é medido. Sendo assim, (Aguirre, 2007) apresenta um método sistemático, de forma a se estimar o ruído por meio de resíduos. O procedimento será descrito a seguir:

1. A partir da equação de regressão $y(k) = \Psi^T(k-1)\theta + e(k)$ e dos dados

disponíveis, monte a Equação matricial $\mathbf{y} = \Psi\theta + \mathbf{e}$, como no método de mínimos quadrados, e determine $\hat{\theta}_{\text{MQ}} = [\Psi^T\Psi]^{-1}\Psi^T\mathbf{y}$;

2. Calcule o vetor de resíduos $\xi_i = \mathbf{y} - \Psi\hat{\theta}_{\text{MQ}}$;

3. Faça $i = 2$ (i indica o número de iterações);

4. Com ξ_{i-1}, monte a matriz estendida de regressores, Ψ_i^*, e estime $\hat{\theta}^*_{\text{MQE}} = [\Psi_i^{*T}\Psi_i^*]^{-1}\Psi_i^{*T}\mathbf{y}$;

5. Determine o vetor de resíduos $\xi_i = \mathbf{y} - \Psi_i^*\hat{\theta}^*_{\text{MQE}_i}$;

6. Faça $i = i + 1$ e volte ao passo 4. Repita até convergir.

A convergência pode ser verificada por meio de diversos métodos. Pode-se verificar a variância dos resíduos ou dos parâmetros estimados, por exemplo. Normalmente, são necessárias aproximadamente 10 iterações ou pouco mais para a convergência do método (Aguirre, 2007). Para o presente trabalho utilizou-se um regressor para estimar o ruído com um processo com 30 iterações.

3.7 Validação do Modelo

Uma vez obtido um modelo, é necessário verificar se o mesmo apresenta as principais características do sistema. Sendo assim, diferentes técnicas de validação de modelos foram desenvolvidas.

As técnicas de validação podem depender da aplicação do modelo e da quantidade de informações disponíveis sobre o sistema real.

Para validação, utilizou-se os seguintes aspectos neste trabalho:

1. Predição livre do modelo.

2. Representatividade da característica estática do sistema.

3. Erro quadrático médio estático e dinâmico (RMSE - Root Mean Squared Error).

4. Funções custos quadráticos estáticos e dinâmicos normalizados (J_d e J_s).

O índice RMSE, quando normalizado, pode ser expresso por:

$$\text{RMSE} = \frac{\sqrt{\sum_{k=1}^{N}[y(k) - \hat{y}(k)]^2}}{\sqrt{\sum_{k=1}^{N}[y(k) - \bar{y}]^2}}, \tag{3.20}$$

sendo $\hat{y}(k)$ a *simulação livre* do sinal para o caso dinâmico e o valor estático de saída para o caso estático e $\bar{y}$ é o valor médio do sinal medido $y(k)$. Este índice mede o erro em uma unidade de medida coerente com os dados reais (Nepomuceno, 2002). Geralmente, consideram-se bons modelos aqueles que apresentam índice RMSE normalizado menor que a unidade. Isto significa que, em média, o erro apresentado pelo modelo é menor que o erro dado pela média da série temporal.

O somatório dos erros quadráticos de predição dinâmico e estático é uma outra métrica que pode ser utilizada como critério para validação e comparação dos modelos obtidos. A curva estática pode ser estimada pela equação (3.33), dada por $QS\hat{\theta}$, a seguir apresentada. Portanto, o somatório dos erros quadráticos estático de predição pode ser definido como:

$$J_s = (y_s - QS\theta)^T(y_s - QS\theta). \tag{3.21}$$

Uma vez que o modelo dinâmico pode ser escrito por meio da equação (3.13) e que a estimativa do valor real, um passo à frente, pode ser dada por $\Psi\hat{\theta}$, o somatório dos erros quadráticos dinâmicos pode ser definido como:

$$J_d = (y - \Psi\hat{\theta})^T(y - \Psi\hat{\theta}). \tag{3.22}$$

As funções custo, definidas por meio das equações (3.22) e (3.21) podem ser, por conveniência, normalizadas. Isso faz com que a análise e entendimento dos resultados sejam facilitados, além de contribuir no processo de decisão dos modelos. Ademais, tal procedimento deve ser realizado, uma vez que nem sempre as grandezas envolvidas nas funções custo J_s e J_d são equivalentes.

Sendo assim, define a normalização (Nepomuceno, 2002) das funções custo como sendo:

$$\bar{J}_i(\hat{\theta}) = \frac{J_i(\hat{\theta}) - \min(J_i(\hat{\theta}))}{\max(J_i(\hat{\theta}) - \min(J_i(\hat{\theta})))}, \tag{3.23}$$

em que $\min(\cdot)$ e $\max(\cdot)$ são o valor mínimo e valor máximo do funcional

$J_i(\hat{\theta})$, que normalizado pode ser escrito como $\bar{J}_i$. $\hat{\theta}$ é o vetor de parâmetros estimados e J_i, no caso estudado neste trabalho é equivalente aos custos estáticos e dinâmicos (J_s e J_d).

Dessa forma, obtém-se um outro meio de se quantificar o modelo do sistema obtido. As técnicas de validação aqui apresentadas serão utilizadas ao longo deste trabalho. Ressalta-se que, caso a validação dos modelos não seja executada de modo satisfatório, reformulações nas etapas anteriores devem ser realizadas.

3.8 Informações Auxiliares

Nesta seção será apresentado o conceito básico de informações auxiliares utilizado neste trabalho. Tais informações quando expressas na forma de erro de estimação, podem ser utilizadas pela técnica multi-objetivo de detecção de estruturas, que será apresentada no capítulo subsequente.

Entende-se como informação auxiliar alguma característica peculiar do sistema. Pode-se considerar informações auxiliares características a respeito dos pontos fixos (localização e estabilidade), curva estática e ganho estático, por exemplo. É importante o desenvolvimento de técnicas que utilizam este tipo de informação, pois na prática nem sempre os dados dinâmicos estão disponíveis. Ademais, há casos em que os mesmos se apresentam em pequenas quantidades, ou até mesmo contaminados por ruído. Tais informações podem ser estimadas por um modelo NARMAX polinomial na forma do modelo de erro de estimação (equação 3.24) :

$$y_{aux} = \hat{y}_{aux} + \xi_{aux}, \tag{3.24}$$

em que y_{aux} é o valor real da informação auxiliar, $\hat{y}_{aux}$ seu valor estimado pelo modelo NARMAX e ξ_{aux} o erro de estimação, dado pela diferença entre o valor real e estimado. Em geral, tal característica é estimada em função de uma matriz contendo tais informações que multiplica a matriz mapa, composta por zeros e uns (responsável por mapear os parâmetros) e os parâmetros. Dessa forma, a equação (3.24) pode ser reescrita como:

$$y_{aux} = \psi_{aux}\hat{\theta} + \xi_{aux}, \tag{3.25}$$

sendo ψ_{aux} a matriz de regressores da informação e $\hat{\theta}$ os parâmetros estimados. A metodologia apresentada neste trabalho permite a incorporação na etapa

de detecção de estruturas de qualquer informação acerca do sistema que possa ser estimada por meio da equação (3.25). Tais tipos de informações já foram utilizadas para obtenção de parâmetros (Martins et al., 2009; Barroso et al., 2007; Nepomuceno, 2002), porém não se identificou trabalhos em que as mesmas fossem utilizadas para quantificar a contribuição de cada regressor.

A subseção seguinte apresentará como a curva estática pode ser representadas na forma da equação de erro (equação 3.25).

3.8.1 Pontos Fixos em Modelos NARMAX Polinomiais

Pontos fixos de um sistema discreto, são os pontos de operação que apresentam a seguinte característica:

$$y(k) = y(k+i), \forall i \ \in \mathbb{Z}^+. \tag{3.26}$$

Os pontos fixos (também conhecidos por pontos de equilíbrio) podem ser obtidos a partir de modelos NARX polinomiais por meio de conceitos de agrupamentos de termos e coeficientes de agrupamentos:

$$\Sigma_{y^\ell}\ y_f^\ell + \cdots + \Sigma_{y^2}\ y_f^2 + (\Sigma_y - 1)\ y_f + \Sigma_0 = 0, \tag{3.27}$$

sendo que Σ_0 é o termo constante do modelo. O modelo poderá apresentar até ℓ pontos de operação na saída se o termo $\Sigma_{y^\ell} \neq 0$. Os pontos fixos do modelo são as soluções de (3.27).

Foi mostrado em (Nepomuceno, 2002) que os pontos fixos estimados por um modelo NARMAX polinomial podem ser obtidos por:

$$\hat{y}_f = S\hat{\theta}, \tag{3.28}$$

em que S é uma matriz mapa composta por zeros e uns, responsável por mapear os regressores. Sendo assim, o ponto fixo do sistema pode ser representado como:

$$\begin{aligned} y_f &= \hat{y}_f + \xi_f, \\ &= \psi_f\hat{\theta} + \xi_f \\ &= S\hat{\theta} + \xi_f. \end{aligned} \tag{3.29}$$

Deste modo, pode-se caracterizar o ponto fixo como um tipo de informação a ser expressa na forma da equação 3.24, ou seja, é um tipo de informação auxiliar que pode ser utilizada na metodologia apresentada neste trabalho.

3.8.2 Curva estática de modelos NARMAX polinomiais

A partir da definição de agrupamento de termos, pode-se derivar uma expressão para a curva estática de modelos NARMAX polinomiais. Todo o procedimento seguinte foi originalmente apresentado em (Nepomuceno, 2002). Em regime estacionário, o modelo NARMAX polinomial pode ser escrito da seguinte forma:

$$\hat{\bar{y}} = \Sigma_0 + \Sigma_y \bar{y} + \Sigma_u \bar{u} + \sum_{m=1}^{l-1} \sum_{p=1}^{\ell-m} \Sigma_{y^p u^m} \bar{y}^p \bar{u}^m + \sum_{p=2}^{\ell} \Sigma_{y^p} \bar{y}^p + \sum_{m=2}^{\ell} \Sigma_{u^m} \bar{u}^m, \quad (3.30)$$

sendo que os termos de processo e seus respectivos parâmetros foram agrupados da seguinte forma:

- termo constante : Σ_0
- termos lineares em y : $\Sigma_y \bar{y}$
- termos lineares em u : $\Sigma_u \bar{u}$
- termos cruzados : $\sum_{m=1}^{\ell-1} \sum_{p=1}^{\ell-m} \Sigma_{y^p u^m} \bar{y}^p \bar{u}^m$
- termos não-lineares em y : $\sum_{i=1}^{\ell} \Sigma_{y^i} \bar{y}^i$
- termos não-lineares em u : $\sum_{i=1}^{\ell} \Sigma_{u^i} \bar{u}^i$.

Deste modo, o ganho estático $\hat{K}$ pode ser estimado por:

$$\hat{K}(\bar{y},\bar{u}) = \frac{\bar{y}}{\bar{u}} = \frac{\dfrac{\Sigma_0}{\bar{u}} + \Sigma_u + \displaystyle\sum_{m=2}^{\ell} \Sigma_{u^m} \bar{u}^{m-1}}{1 - \Sigma_y - \displaystyle\sum_{m=1}^{\ell-1} \sum_{p=1}^{\ell-m} \Sigma_{y^p u^m} \bar{y}^{(p-1)} \bar{u}^m - \sum_{p=2}^{\ell} \Sigma_{y^p} \bar{y}^{(p-1)}}. \quad (3.31)$$

Para obter a curva estática, basta multiplicar ambos os lados da igualdade apresentada na equação (3.31) por $\bar{u}$, resultando em:

$$\hat{y_s} = \frac{\Sigma_0 + \Sigma_u \bar{u} + \sum_{m=2}^{\ell} \Sigma_{u^m} \bar{u}^m}{1 - \Sigma_y - \sum_{m=1}^{\ell-1} \sum_{p=1}^{\ell-m} \Sigma_{y^p u^m} \bar{y}^{(p-1)} \bar{u}^m - \sum_{p=2}^{\ell} \Sigma_{y^p} \bar{y}^{(p-1)}}. \tag{3.32}$$

Pode ser mostrada, como em (Nepomuceno et al., 2007), que a curva estática de um modelo NARMAX polinomial pode ser estimada por:

$$\hat{y_s} = QS\hat{\theta}, \tag{3.33}$$

ou ainda

$$y_s = QS\hat{\theta} + \xi_s, \tag{3.34}$$

que é a informação auxiliar expressão na forma de erro de estimação. Neste sentido $\hat{y_s}$ é o valor estático estimado por $QS\theta$. y_s é a curva estática do sistema e ξ_s o erro de estimação. Para modelos NARMAX polinomiais, é possível escrever:

$$\hat{y_{si}} = f(\bar{y}_i,\bar{u}_i) = q_i^T S\theta,\ i = 1,2,\ldots,n_{\text{sf}}, \tag{3.35}$$

sendo n_{sf} o número de pontos em estado estacionário $(\bar{u}_i,\bar{y}_i)$ considerado e

$$\mathbf{q}_i^T = \left[1\ \bar{y}_i \ldots \bar{y}_i^{\ell}\ \bar{u}_i \ldots \bar{u}_i^{\ell}\ F_{yu}\right],\ i = 1,2,\ldots,n_{\text{sf}}, \tag{3.36}$$

sendo que F_{yu} representa todos os monômios não-lineares no modelo que envolvam $y(k)$ e $u(k)$, e

$$R = S\theta = \left[\Sigma_0\ \Sigma_y \ldots \Sigma_{y^\ell}\ \Sigma_u \ldots \Sigma_{u^\ell}\ F_\Sigma\right]^T, \tag{3.37}$$

em que F_Σ indica os coeficientes de agrupamentos de todos os agrupamentos de termos em F_{yu} e $S\theta$ é um vetor de constantes. Portanto, a estimativa estática é valida, e ainda:

$$Q = \begin{bmatrix} \mathbf{q}_1^T \\ \mathbf{q}_2^T \\ \vdots \\ \mathbf{q}_{n_{\text{sf}}} \end{bmatrix}. \tag{3.38}$$

Uma vez que a curva estática pode ser representada de acordo com a equação de erro apresentada na seção 3.8, a mesma pode ser classificada como uma característica do sistema a ser incorporada na estrutura do modelo, por meio da metodologia posteriormente apresentada.

3.9 Conclusões do Capítulo

O presente capítulo teve como principal objetivo contextualizar o leitor com o principal tema abordado nesta dissertação. Foram abordadas as cinco principais etapas empregadas na identificação de sistemas, sendo dado um enfoque principal na modelagem por meio de modelos NARMAX polinomiais.

Em particular, na etapa de detecção de estruturas, abordou-se as técnicas de taxa de redução de erro (ERR) e o critério de informação de Akaike. Isto se fez necessário, uma vez que estas técnicas foram base para o desenvolvimento deste trabalho.

Ademais, foi elucidado o conceito de informação auxiliar, exemplificado por meio de conceitos de pontos fixos e curva estática, além de explicitar como informações auxiliares podem ser utilizadas na metodologia proposta neste trabalho.

No próximo capítulo será apresentada a metodologia multi-objetivo para detecção de estruturas de modelos NARMAX polinomiais, desenvolvida e empregada nesta pesquisa.

Capítulo 4

Metodologia

"O grande prazer da vida é fazer as coisas que as pessoas dizem que você não é capaz!"

Walter Bagehot

4.1 Introdução

Nos capítulos anteriores foram descritos os conceitos preliminares, bem como revisões de literatura acerca de Identificação de Sistemas. Dentre as cinco principais etapas da Identificação de Sistemas, enfocou-se no processo de detecção de estruturas de modelos NARMAX polinomiais.

Neste capítulo será apresentada a metodologia multi-objetivo utilizada para detecção de estruturas de modelos NARMAX polinomiais, aplicada no âmbito deste trabalho em sistemas reais e simulados.

4.2 Definição do Problema

A Figura 4.1 apresenta um problema de otimização multi-objetivo, em que os dois objetivos apresentados são conflitantes, ou seja, não podem ser minimizados simultaneamente. Ao minimizar um dos objetivos, o outro necessariamente maximiza-se gradativamente.

As soluções eficientes do problema de otimização vetorial (ou multi-objetivo) são conhecidas como o conjunto Pareto-ótimo. São as melhores soluções pos-

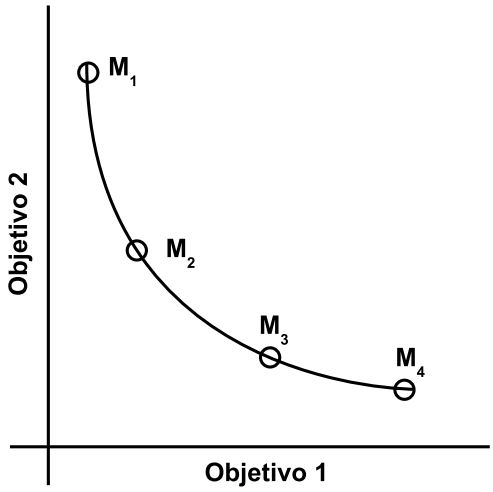

Figura 4.1: Problema de otimização bi-objetivo com objetivos conflitantes. Os modelos M_1 e M_4 apresentam as soluções individuais dos objetivos 1 e 2, ao passo que os modelos M_2 e M_3 são modelos intermediários pertinentes à curva de soluções eficientes.

síveis de serem obtidas. Não há como definir *a priori*, a partir das avaliações dos funcionais, se existe uma solução melhor que outra.

Na Identificação de Sistemas, deseja-se obter um modelo que tenha bom comportamento perante diversas características do sistema. Em geral, deseja-se minimizar simultaneamente erros estáticos e dinâmicos. É difícil obter um modelo capaz de minimizar simultaneamente erros estáticos e dinâmicos de predição, pois os mesmos são, usualmente, conflitantes. Sendo assim, o que se tem é um problema de otimização multi-objetivo que consequentemente apresente um conjunto de soluções eficientes, compondo o conjunto Pareto-ótimo.

A abordagem multi-objetivo é desenvolvida classicamente durante a etapa de determinação dos parâmetros do modelo (Martins et al., 2009; Nepomuceno et al., 2007; Barroso et al., 2007). Recentemente, trabalhos como (Martins et al., 2011; Barbosa, 2010) utilizaram de abordagens multi-objetivo para determinação de estruturas de modelos NARX polinomiais. Contudo, tais trabalhos utilizaram o processo multi-objetivo para determinação de parâmetros como ferramenta auxiliar na detecção de estruturas. Até onde se viu, não se tem notícia de um processo de detecção de estruturas puramente

multi-objetivo, sem o auxílio da etapa de obtenção multi-objetivo dos parâmetros.

A ERR e o AIC são inerentemente técnicas mono-objetivo, pois visam minimizar unicamente o erro dinâmico de predição. Deste modo, o problema consiste em estabelecer uma estrutura ou um conjunto de estruturas de modelos que possua compromisso mútuo entre as mais diversas características do sistema. Neste trabalho considerou-se como características a serem utilizadas os comportamentos dinâmicos e estáticos do sistema. Sendo assim, a taxa de redução de erro multi-objetivo (MERR, do inglês *Multi-objective Error Reduction Ratio*) toma seu lugar, e é aqui definida, proposta e apresentada.

4.3 A Taxa de Redução de Erro Multi-objetivo (MERR)

Modelos NARMAX polinomiais possuem como característica a se destacar a fácil extração de informações acerca do sistema, estimadas pelo modelo. Informações como curva estática, pontos fixos, ganhos estáticos e estabilidade podem ser extraídas do modelo, de forma estimada. Sendo assim, tal informação pode ser expressa como:

$$\begin{aligned} y_q &= \hat{y}_q + \xi_q, \\ y_q &= \psi_q \hat{\theta} + \xi_q, \end{aligned} \tag{4.1}$$

sendo y_q a informação do sistema (dados dinâmicos de saída, ponto fixo, curva estática, ganho estático), $\hat{y}_q$ o valor da informação estimada pelo modelo, ψ_q a matriz de informações e ξ_q o erro de estimação da informação.

Utilizando algoritmos de ortogonalização (Gram-Schmidt normal ou modificado ou transformações de HouseHolder, por exemplo), pode-se obter um modelo estimador auxiliar em que os regressores pertinentes a cada informação ψ_q são ortogonais sobre as informações y_q, formando uma base ortogonal às informações do sistema.

$$y_q = \sum_{j=1}^{n_\theta} \hat{g}_j \Omega_q^j + \xi_q, \tag{4.2}$$

em que Ω_q^j é o j-ésimo regressor da informação ortogonal y_q, $\hat{g}_j$ o parâmetro

ortogonal a ele associado e n_θ o número de regressores candidatos. A partir daí, pode-se definir a variância afim (ou variância auxiliar) composta pelo erro de estimação, como sendo:

$$var[\xi_q] = \lim_{N\to\infty} \frac{1}{N}\left[y_q^T y_q - \sum_{j=1}^{n_\theta} \hat{g}_j{}^2 \Omega_q^{jT}\Omega_q^j\right], \tag{4.3}$$

sendo $var[\xi_q]$ a variância correspondente à q-ésima informação afim do sistema e o operador T correspondente à transposição do vetor.

Considerando que nenhum termo fosse acrescentado ao modelo, a variância afim da i-ésima informação auxiliar seria exatamente igual ao erro quadrático da informação do sistema. A cada termo acrescido ao modelo, a variância afim decresce um valor de $\frac{1}{N}\hat{g}_j{}^2\Omega_q^{jT}\Omega_q^j$, sendo Ω_q^j o regressor ortogonal inserido e $\hat{g}_j$ seu respectivo parâmetro. Ressalta-se que pode-se tomar a equação (4.3) como um problema de otimização mono-objetivo em que o objetivo (função custo a ser otimizada) a ser minimizado é a variância afim.

Objetivando o desenvolvimento de uma técnica multi-objetivo, pode-se trabalhar com a combinação ponderada convexa de diferentes informações auxiliares do sistema utilizando pesos diferentes ($\omega_1, \cdots, \omega_z$), tal que $\sum_{q=1}^{z} \omega_z = 1$, sendo z o número de informações a serem incorporadas). Nesse sentido, tem-se:

$$\begin{aligned} \omega_1 var[\xi_1] + \cdots + \omega_z var[\xi_z] &= \omega_1 \lim_{N\to\infty} \frac{1}{N}\left[y_1^T y_1 - \sum_{j=1}^{n_\theta} \hat{g}_j{}^2 \Omega_1^{jT}\Omega_1^j\right] \\ &+ \cdots + \omega_z \lim_{N\to\infty} \frac{1}{N}\left[y_z^T y_z - \sum_{j=1}^{n_\theta} \hat{g}_j{}^2 \Omega_z^{jT}\Omega_z^j\right]. \end{aligned} \tag{4.4}$$

Rearranjando a equação (4.4), tem-se;

$$\begin{aligned} \omega_1 var[\xi_1] + \cdots + \omega_z var[\xi_z] &= \lim_{N\to\infty} \frac{1}{N}\left[\omega_1\left(y_1^T y_1 - \sum_{j=1}^{n_\theta} \hat{g}_j{}^2 \Omega_1^{jT}\Omega_1^j\right)\right] \\ &+ \cdots + \lim_{N\to\infty} \frac{1}{N}\left[\omega_z\left(y_z^T y_z - \sum_{j=1}^{n_\theta} \hat{g}_j{}^2 \Omega_z^{jT}\Omega_z^j\right)\right]. \end{aligned} \tag{4.5}$$

Uma vez que pode-se distribuir a variável ω e que o limite da soma é igual

à soma dos limites, a parcela:

$$\lim_{N\to\infty}\frac{1}{N}\left[\omega_1\left(y_1^T y_1-\sum_{j=1}^{n_\theta}\hat{g}_j{}^2\Omega_1^{jT}\Omega_1^j\right)\right]+\cdots \tag{4.6}$$
$$+\lim_{N\to\infty}\frac{1}{N}\left[\omega_z\left(y_z^T y_z-\sum_{j=1}^{n_\theta}\hat{g}_j{}^2\Omega_z^{jT}\Omega_z^j\right)\right],$$

pode ser reescrita como:

$$\lim_{N\to\infty}\frac{1}{N}\left[\sum_{q=1}^{z}\omega_q y_q^T y_q-\sum_{q=1}^{z}\sum_{j=1}^{n_\theta}\omega_q\hat{g}_j{}^2\Omega_q^{jT}\Omega_q^j\right]. \tag{4.7}$$

A variância multi-objetivo composta pelas variâncias das informações do modelo, apresentada na equação (4.4), pode ser reescrita como:

$$\sum_{q=1}^{z}\omega_q var[\xi_q]=\lim_{N\to\infty}\frac{1}{N}\left[\sum_{q=1}^{z}\omega_q y_q^T y_q-\sum_{q=1}^{z}\sum_{j=1}^{n_\theta}\omega_q\hat{g}_j{}^2\Omega_q^{jT}\Omega_q^j\right], \tag{4.8}$$

ou ainda:

$$\sum_{q=1}^{z}\omega_q var[\xi_q]=\lim_{N\to\infty}\frac{1}{N}\sum_{q=1}^{z}\omega_q\left[y_q^T y_q-\sum_{j=1}^{n_\theta}\hat{g}_j{}^2\Omega_q^{jT}\Omega_q^j\right], \tag{4.9}$$

em que $z\in\mathbb{N}$ é o número de informações auxiliares do sistema.

A redução no valor da variância multi-objetivo pode ser normalizada com relação ao erro quadrático médio do sinal multi-objetivo composto pela combinação convexa das informações a serem incorporadas. Sendo assim, define-se a taxa de redução de erro multi-objetivo (MERR - Multi-objective Error Reduction Ratio), dada pela inclusão do j-ésimo regressor como sendo:

$$MERR^i=\frac{\hat{g}_j{}^2\langle\sum_{q=1}^{z}\omega_q\Omega_q^j,\sum_{q=1}^{z}\omega_q\Omega_q^j\rangle}{\langle\sum_{q=1}^{z}\omega_q y_q,\sum_{q=1}^{z}\omega_q y_q\rangle}. \tag{4.10}$$

O operador $\langle\cdot,\cdot\rangle$ é o operador produto interno. Pode-se quantificar em

porcentagem a contribuição de cada regressor ao representar dada informação do sistema em questão. Para se fazer estimativas dos parâmetros representados em base ortogonal $\hat{g}_j$, pode-se proceder da seguinte forma:

$$\hat{g}_i = \frac{\langle \sum_{q=1}^{z} \omega_q \Omega_q^j, \sum_{q=1}^{z} \omega_q y_q \rangle}{\langle \sum_{q=1}^{z} \omega_q \Omega_q^j, \sum_{q=1}^{z} \omega_q \Omega_q^j \rangle}, i = 1, \cdots, n_\theta, \tag{4.11}$$

que nada mais é que o método LS aplicado à combinação linear convexa dos regressores ortogonais às informações. Destaca-se que os modelos pertinentes à curva Pareto-ótima são obtidos por meio da variação de ω_q.

Um ponto importante a ser ressaltado é que considerando apenas uma informação a ser incorporada na equação (4.10) e se esta informação for a informação referente aos dados dinâmicos, a técnica retorna exatamente à ERR (taxa de redução de erro). Neste caso haveria apenas um peso ($\omega_1 = 1$) e somente dados dinâmicos estariam envolvidos no processo.

O uso da equação 4.10 é viável desde que se tenha o modelo representado em uma base ortogonal, e é necessário a aplicação de um método para determinar tal base. Em (Aguirre, 2007) podem ser encontradas uma lista de técnicas com este objetivo. Em particular, cita-se o método das transformações sucessivas de HouseHolder que fornece uma representação ortogonal às informações, e que foi utilizado neste trabalho.

O método de Golub-HouseHolder utiliza de transformações de HouseHolder para quantificar de modo iterativo a contribuição de cada regressor ao explicar o quadrado da variância dos dados dinâmicos do sistema. Tal método, amplamente utilizado no emprego da ERR, foi expandido para a MERR e será apresentado na subseção seguinte.

4.3.1 O Algoritmo de Golub-HouseHolder para a MERR

O algoritmo de Golub-HouseHolder, presente em (Aguirre, 2007) apresenta uma forma iterativa e sequencial de se compor o modelo, considerando os regressores com maiores ERR. Nesta subseção será apresentado a adaptação deste algoritmo à MERR.

O ponto de partida é montar a seguinte matriz estendida, antes mesmo da primeira iteração.

$$\Psi_m^{(0)} = [\Psi_m \; y_m], \tag{4.12}$$

sendo $\Psi_m^{(0)}$ a matriz de regressores multi-objetivo e y_m os dados compostos pelas informações do sistema a serem incorporadas a seguir definidos:

$$\Psi_m = \sum_{q=1}^{z} \omega_q \psi_q, \tag{4.13}$$

$$y_m = \sum_{q=1}^{z} \omega_q y_q, \tag{4.14}$$

em que $z \in \mathbb{N}$ é o número de informações auxiliares do sistema. A matriz $\Psi_m^{(k-1)}$ é obtida após a aplicação de $(k-1)$ transformações de HouseHolder em $\Psi_m^{(0)}$. Um modelo que contenha os primeiros $(k-1)$ parâmetros e regressores terá a seguinte soma dos resíduos multi-objetivo, ao quadrado:

$$J_m^{(k-1)} = \sum_{T=k}^{N} \left(y_{m,T}^{(k-1)} \right)^2, \tag{4.15}$$

sendo esta parcela devido a não inserção dos regressores $k, \cdots, N$. Quando a $k-$ésima transformação é aplicada a $\Psi_m^{(k-1)}$, obtém-se $\Psi_m^{(k)}$ e, se o novo regressor for adicionado ao modelo, a soma dos resíduos multi-objetivo, ao quadrado (equação 4.15) é reduzida a:

$$\begin{aligned} J_m^{(k-1)} &= \sum_{T=k}^{N} \left(y_{m,T}^{(k-1)} \right)^2 - \left(y_{m,T}^{(k)} \right) \\ &= \sum_{T=k+1}^{N} \left(y_{m,T}^{(k)} \right)^2. \end{aligned} \tag{4.16}$$

Na $k-$ésima iteração, deseja-se escolher o $k-$ésimo regressor que, quando incluído no modelo, resultará na maior diminuição da soma dos quadrados do resíduo multi-objetivo. Tal benefício devido a inclusão pode ser quantificado por meio da taxa de redução de erro multi-objetivo (MERR) do regressor.

Finalmente, o procedimento para determinação de estrutura na $k-$ésima iteração consiste em calcular a MERR para todos os regressores candidatos que ainda não tenham sido incluídos no modelo, e escolher o que apresentar maior valor para ser incluso no mesmo. Vale lembrar que tanto a ERR, como a MERR devem ser empregadas em conjunto com algum critério de informação, de modo a precisar o número de termos que deve ser incluído no

modelo.

4.4 Aspectos Gerais e Recursos Computacionais

Esta seção tem como propósito apresentar aspectos gerais, utilizados para o desenvolvimento deste trabalho. Tem como objetivo facilitar a reprodução dos resultados aqui apresentados.

Todos os processos foram simulados via *software MATLAB*®, utilizando um *laptop* da marca *ASUS*® com 4*Gb* de memória RAM, processador Intel core *i*3 à 2,27 GHz, executando sistema operacional Windows 7 *Home Premium*®.

Utilizou-se 2 como máximo atraso para regressores de entrada e saída e máximo grau de não-linearidade. A estimação dos parâmetros foi realizada via ELS, utilizando um regressor para modelar o ruído ($e(k-1)$) e um processo com 30 iterações.

Os sistemas foram cuidadosamente escolhidos, de modo a tentar abranger uma maior variedade possível de classes, necessárias para validar a técnica desenvolvida. Como sistemas simulados, utilizou-se modelos NARX polinomiais com entradas distintas, composta em um primeiro momento por uma variável aleatória e em um segundo momento por um filtro passa-baixas, construído pela combinação linear de uma entrada auto-regressiva e de uma variável aleatória. Como sistemas reais utilizou-se um conversor CC-CC buck, que apresenta dinâmica não linear e comportamento estático afim. Por outro lado, a planta de neutralização de pH (segundo sistema real utilizado) apresenta uma não-linearidade estática severa, bem como dinâmica.

Quanto da obtenção dos pesos ω_d e ω_s para utilização no procedimento multi-objetivo, sabe-se que não é uma tarefa trivial. Pardalos et al. (2011) apresentam dois métodos de variação dos pesos, de forma a obter um Pareto uniformemente povoado. Maiores informações acerca de técnicas para preenchimento de curvas de Pareto podem ser encontradas na literatura citada e nas outras ali referenciadas. Como não se trata do escopo principal deste trabalho, considerou-se uma variação uniforme nos pesos, de forma a obter o conjunto Pareto ($\omega_d = [0{,}1; 0{,}35; 0{,}7; 0{,}9]$ e consequentemente $\omega_s = [0{,}9; 0{,}65; 0{,}3; 0{,}1]$) para pesos dinâmicos e estáticos, respectivamente.

4.5 Conclusões do Capítulo

O presente capítulo apresentou uma técnica de detecção de estruturas de modelos NARMAX polinomiais via abordagem multi-objetivo. Para tal, foi necessária a definição da taxa de redução de erro multi-objetivo, responsável por quantificar a contribuição de cada regressor ao explicar a variância composta pela combinação convexa das informações do sistema.

A metodologia apresentada permite a obtenção do conjunto Pareto-ótimo de estruturas de modo simples, direto e sistemático. Sendo assim, pode-se escolher o modelo a ser utilizado para representar o sistema, que está contido no conjunto de soluções eficientes (Pareto-ótimo).

Mostrou-se que, no caso particular em que o peso dinâmico tem o valor unitário e os demais pesos assumem valor nulo, a técnica volta a ser exatamente a taxa de redução de erro (ERR). Por fim, destacou-se que a técnica, diferentemente das técnicas encontradas pelo autor na literatura, apresenta caráter puramente multi-objetivo na detecção de estruturas e quantificação de regressores, sendo este seu principal diferencial inovador. Apesar de sua aplicação particular em modelos NARMAX (NARX ou NARMAX), a metodologia apresentou-se eficiente, como pode ser relatado nos próximos capítulos deste trabalho.

Capítulo 5

Resultados

"Toda virtude é um justo meio entre dois extremos: o excesso e a carência."

Aristóteles

5.1 Introdução

Este capítulo tem por principal objetivo mostrar os resultados obtidos, quando a metodologia desenvolvida no Capítulo 4 foi aplicada a dois sistemas reais e dois sistemas simulados. Como sistemas reais, utilizou-se uma planta de neutralização de pH e um conversor abaixador CC-CC buck. Como sistemas simulados, utilizou-se os sistemas apresentados em (Piroddi e Spinelli, 2003) e em (Bonin et al., 2010), que consistem de modelos NARX polinomiais. Ademais, serão explicitadas formas de incorporação de pontos fixos e curva estática por meio da MERR.

5.2 Aplicação da MERR para Pontos Fixos e Curva Estática

5.2.1 Pontos Fixos

Como visto em conceitos preliminares, os pontos fixos de um modelo NARMAX polinomial podem ser estimados por:

$$\begin{aligned} y_1 &= \hat{y_1} + \xi_1, \\ y_1 &= S\hat{\theta} + \xi_1, \\ y_1 &= \psi_1\hat{\theta} + \xi_1, \end{aligned} \tag{5.1}$$

sendo y_1 o valor dos pontos fixos, S a matriz que contém informações acerca do mapeamento dos pontos fixos, $\hat{\theta}$ os parâmetros estimados e ξ_1 o erro de estimação do ponto fixo. $\hat{y_1}$ são os pontos fixos do modelo, estimados por $S\hat{\theta}$ e ψ_1 é a matriz de regressores referente aos pontos fixos.

Por outro lado, tem-se o modelo dinâmico geral NARMAX polinomial, que pode ser escrito como:

$$y_2 = \psi_2\hat{\theta} + \xi_2, \tag{5.2}$$

sendo y_2 o valor dinâmico do modelo, ψ_2 a matriz de regressores dinâmicos e ξ_2 o erro de predição um passo a frente. Neste caso, $z = 2$ é o número de informações a serem incorporadas no modelo.

Utilizando transformações de HouseHolder, pode-se obter modelos auxiliares ortogonais às informações, da forma apresentada nas equações 5.3 e 5.4. Deste modo, obtém-se:

$$y_1 = \sum_{j=1}^{n_\theta} \hat{g}_j \Omega_1^j + \xi_1, \tag{5.3}$$

$$y_2 = \sum_{j=1}^{n_\theta} \hat{g}_j \Omega_2^j + \xi_2, \tag{5.4}$$

em que Ω_q^j, $q = 1{,}2$ é o j-ésimo regressor ortogonal à informação(pontos fixos e dados dinâmicos, respectivamente), n_θ o número de regressores candidatos e $\hat{g}_j$ o parâmetro ortogonal associado ao regressor ortogonal. A partir daí, pode-se definir a variância afim, referente aos pontos fixos e aos dados dinâmicos como sendo:

$$var[\xi_1] = \lim_{N\to\infty} \frac{1}{N} \left[y_1^T y_1 - \sum_{j=1}^{n_\theta} \hat{g}_j{}^2 \Omega_1^{jT} \Omega_1^j \right], \tag{5.5}$$

$$var[\xi_2] = \lim_{N\to\infty} \frac{1}{N} \left[y_2^T y_2 - \sum_{j=1}^{n_\theta} \hat{g}_j{}^2 \Omega_2^{jT} \Omega_2^j \right]. \tag{5.6}$$

As variâncias acima apresentadas se referem às informações dos pontos fixos e dos dados dinâmicos. Por meio de uma combinação ponderada convexa dos objetivos a serem minimizados (variância dinâmica e de pontos fixos), baseando-se na equação 4.9 tem-se:

$$\sum_{q=1}^{2} \omega_q var[\xi_q] = \lim_{N\to\infty} \frac{1}{N} \sum_{q=1}^{2} \omega_q \left[y_q^T y_q - \sum_{j=1}^{n_\theta} \hat{g}_j{}^2 \Omega_q^{jT} \Omega_q^j \right], \tag{5.7}$$

em que $z = 2$ é correspondente ao número de informações do modelo a serem utilizadas.

A redução no valor da variância multi-objetivo composta por informações dinâmicas e de pontos fixos podem ser normalizada com relação ao erro quadrático médio da informação composta pela combinação ponderada convexa das informações utilizadas. Deste modo, a taxa de redução de erro multi-objetivo (MERR - Multi-objective Error Reduction Ratio), dada pela inclusão do j-ésimo regressor considerando informações dinâmicas e referente aos pontos fixos é dada como:

$$MERR^i = \frac{\hat{g}_j{}^2 \langle \sum_{q=1}^{2} \omega_q \Omega_q^j, \sum_{q=1}^{2} \omega_q \Omega_q^j \rangle}{\langle \sum_{q=1}^{2} \omega_q y_q, \sum_{q=1}^{2} \omega_q y_q \rangle}. \tag{5.8}$$

Pode-se estimar os parâmetros $\hat{g}_i$ exatamente como já apresentado, reproduzido aqui por conveniência ao leitor:

$$\hat{g}_i = \frac{\langle \sum_{q=1}^{2} \omega_q \Omega_q^j, \sum_{q=1}^{2} \omega_q y_q \rangle}{\langle \sum_{q=1}^{2} \omega_q \Omega_q^j, \sum_{q=1}^{2} \omega_q \Omega_q^j \rangle}, i = 1, \cdots, n_\theta, \tag{5.9}$$

que nada mais é que o método LS aplicado à combinação linear convexa dos regressores ortogonais às informações dinâmicas e referente aos pontos fixos. Destaca-se que os modelos pertinentes à curva Pareto-ótima são obtidos por meio da variação de ω_q, $q = 1{,}2$ e ainda que para $\omega_1 = 0$ e $\omega_2 = 1$ a técnica retorna exatamente à ERR.

Foi mostrado como se pode incorporar informações acerca dos pontos fixos, em conjunto com informações dinâmicas, na quantificação da contri-

buição de cada regressor ao representar a combinação ponderada convexa, composta pelas variâncias referente aos dados dinâmicos e pontos fixos. A seguir será elucidado como se pode incorporar a curva estática do sistema na etapa de detecção de estruturas.

5.2.2 A Curva Estática

Como visto em conceitos preliminares, a curva estática de um modelo NARMAX polinomial pode ser estimada por:

$$\begin{aligned} y_1 &= \hat{y_1} + \xi_1, \\ y_1 &= QR\hat{\theta} + \xi_1, \\ y_1 &= \psi_1\hat{\theta} + \xi_1, \end{aligned} \tag{5.10}$$

sendo y_1 a curva estática do sistema, Q a matriz que contém informações acerca dos pontos fixos, R a matriz que mapeia os parâmetros de acordo com os regressores do modelo, $\hat{\theta}$ os parâmetros estimados e ξ_1 o erro de estimação da curva estática. $\hat{y_1}$ é a curva estática estimada por $QR\hat{\theta}$ e ψ_1 é a matriz de regressores estáticos.

Por outro lado, tem-se o modelo dinâmico geral NARMAX polinomial, que pode ser escrito como:

$$y_2 = \psi_2\hat{\theta} + \xi_2, \tag{5.11}$$

sendo y_2 o valor dinâmico do modelo, ψ_2 a matriz de regressores dinâmicos e ξ_2 o resíduo dinâmico do modelo. Também neste caso, $z = 2$ é o número de informações a serem incorporadas no modelo.

A partir daqui, o procedimento é idêntico ao apresentado para incorporação dos pontos fixos. Utilizando transformações de HouseHolder, pode-se obter modelos auxiliares ortogonais às informações, da forma apresentada na equação 5.3. Deste modo, obtém-se:

$$y_1 = \sum_{j=1}^{n_\theta} \hat{g}_j \Omega_1^j + \xi_1, \tag{5.12}$$

$$y_2 = \sum_{j=1}^{n_\theta} \hat{g}_j \Omega_2^j + \xi_2, \tag{5.13}$$

em que Ω_q^j, $q = 1{,}2$ é o j-ésimo regressor ortogonal à cada informação (dados estáticos e dinâmicos, respectivamente), n_θ o número de regressores candidatos e $\hat{g}_j$ o parâmetro ortogonal associado ao regressor ortogonal. Define-se a variância afim estática e dinâmica como sendo:

$$var[\xi_1] = \lim_{N\to\infty} \frac{1}{N}\left[y_1^T y_1 - \sum_{j=1}^{n_\theta} \hat{g}_j{}^2 \Omega_1^{jT}\Omega_1^j\right], \tag{5.14}$$

$$var[\xi_2] = \lim_{N\to\infty} \frac{1}{N}\left[y_2^T y_2 - \sum_{j=1}^{n_\theta} \hat{g}_j{}^2 \Omega_2^{jT}\Omega_2^j\right]. \tag{5.15}$$

Por meio de uma combinação ponderada convexa dos objetivos a serem minimizados (variâncias estática e dinâmica), baseando-se na equação 4.9 tem-se:

$$\sum_{q=1}^{2} \omega_q var[\xi_q] = \lim_{N\to\infty} \frac{1}{N}\sum_{q=1}^{2} \omega_q \left[y_q^T y_q - \sum_{j=1}^{n_\theta} \hat{g}_j{}^2 \Omega_q^{jT}\Omega_q^j\right], \tag{5.16}$$

A redução no valor da variância multi-objetivo composta por informações dinâmicas e estáticas pode ser normalizada com relação ao erro quadrático médio da informação composta pela combinação ponderada convexa das informações utilizadas. Deste modo, a taxa de redução de erro multi-objetivo (MERR - Multi-objective Error Reduction Ratio), dada pela inclusão do j-ésimo regressor considerando informações estáticas e dinâmicas é dada como:

$$MERR^i = \frac{\hat{g}_j{}^2 \langle \sum_{q=1}^{2} \omega_q \Omega_q^j, \sum_{q=1}^{2} \omega_q \Omega_q^j \rangle}{\langle \sum_{q=1}^{2} \omega_q y_q, \sum_{q=1}^{2} \omega_q y_q \rangle}. \tag{5.17}$$

Os parâmetros $\hat{g}_i$ são estimados como já apresentado, reproduzido mais uma vez por conveniência ao leitor:

$$\hat{g}_i = \frac{\langle \sum_{q=1}^{2} \omega_q \Omega_q^j, \sum_{q=1}^{2} \omega_q y_q \rangle}{\langle \sum_{q=1}^{2} \omega_q \Omega_q^j, \sum_{q=1}^{2} \omega_q \Omega_q^j \rangle}, i = 1, \cdots, n_\theta, \tag{5.18}$$

que nada mais é que o método LS aplicado à combinação linear convexa dos regressores ortogonais às informações dinâmicas e estáticas. Destaca-se que os modelos pertinentes à curva Pareto-ótima são obtidos por meio da variação de ω_q, $q = 1{,}2$ e ainda que para $\omega_1 = 0$ e $\omega_2 = 1$ tem-se exatamente a ERR,

conforme provado no Capítulo 4.

Uma vez obtido o conjunto Pareto-ótimo, tem-se o problema do critério de decisão. Nepomuceno (2002) apresenta técnicas como equilíbrio entre polarização e variância e mínimo das normas de objetivo como critérios decisores. Já Barroso et al. (2007) desenvolve a técnica baseada no critério de correlação como seletor automático de modelos. Neste trabalho serão estudados todos os modelos obtidos, pertinentes ou não à curva Pareto-ótimo, suas características e peculiaridades. Assim, não será utilizado nenhum critério decisor, sendo que todos os modelos serão analisados e avaliados.

Deve ser ressaltado que não se tem por objetivo esgotar o número de informações auxiliares que podem ser incorporadas ao modelo no processo, e sim mostrar que há mais de um tipo de informação que pode ser utilizada por meio da metodologia apresentada. Conjectura-se que o comportamento da técnica seja semelhante perante a utilização de pontos fixos ou curva estática, uma vez que a diferença conceitual entre estes dois tipos de informações é bastante sutil. Por fim, destaca-se que mais de dois tipos de informação auxiliar podem ser incorporados simultaneamente, bastando que os mesmos possam ser escritos na forma de erro de estimação apresentada.

5.3 Resultados Simulados

5.3.1 Exemplo Simulado 1

Considere o seguinte sistema simulado, apresentado originalmente em (Piroddi e Spinelli, 2003):

$$y(k) = 0{,}5y(k-1)+0{,}8u(k-2)+u(k-1)^2-0{,}05y(k-2)^2+0{,}5+e(k), \quad (5.19)$$

em que a entrada u foi gerada a partir de um sinal do tipo ruído branco, com característica gaussiana, média nula e variância unitário. O ruído $e(k)$ foi gerado de forma semelhante, a exceção da variância, que foi de 0,05. Este exemplo também foi estudado em (Bonin et al., 2010).

Os dados estáticos foram gerados de acordo com a teoria de coeficientes de agrupamentos e curva estática, apresentados no Capítulo 3 e são apresentados na Figura (5.3). Ressalta-se que o modelo do sistema é composto pelos agrupamentos constante, termos lineares em y e u e termos quadráticos não-

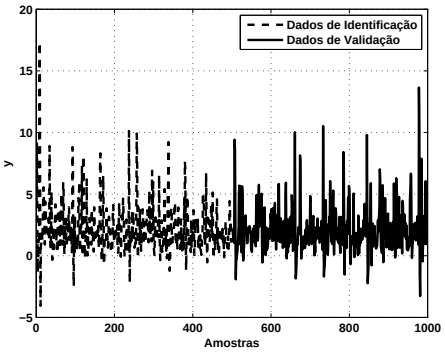

Figura 5.1: Dados de saída do sistema simulado 1. Em pontilhado encontram-se os dados utilizados para identificação e em linha contínua os dados que foram utilizados para validação.

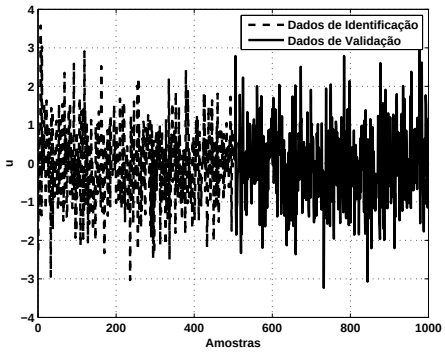

Figura 5.2: Dados de entrada do sistema simulado 1. Em pontilhado encontram-se os dados utilizados para identificação e em linha contínua os dados que foram utilizados para validação.

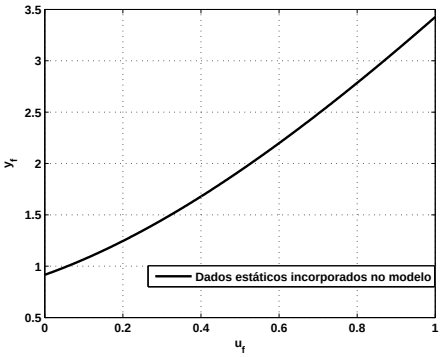

Figura 5.3: Dados estáticos do sistema simulado 1 incorporados ao modelo do sistema.

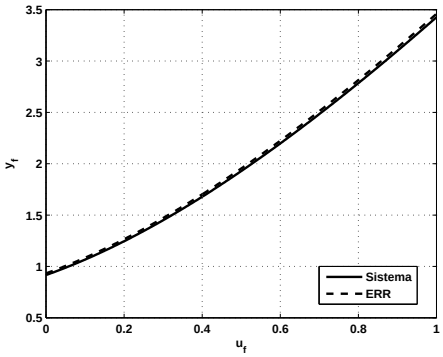

Figura 5.4: Curva estática do modelo Λ_{ERR}. Sistema 1

lineares em y e u. Os dados de saída e entrada podem ser vistos nas Figuras (5.1) e (5.2), respectivamente.

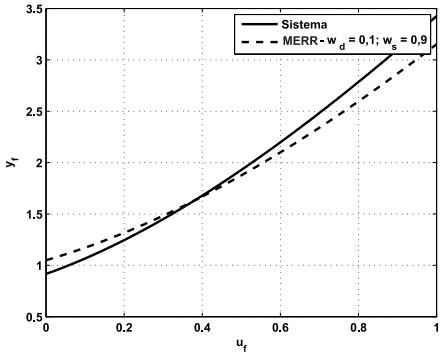

Figura 5.5: Curva estática do modelo Λ_1. Sistema 1

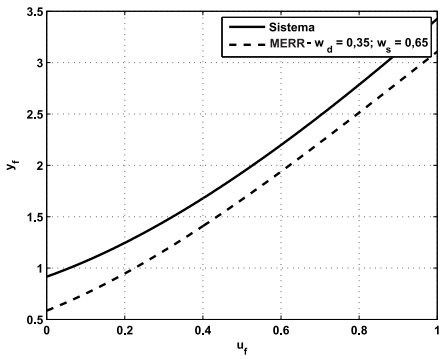

Figura 5.6: Curva estática do modelo Λ_2. Sistema 1

5.3.2 Resultados e Discussão dos Resultados - Exemplo Simulado 1

A metodologia da taxa de redução de erro multi-objetivo (MERR) foi aplicada a identificação deste sistema, utilizando para tal $\omega_d = [0{,}1; 0{,}35; 0{,}7; 0{,}9]$ e consequentemente $\omega_s = [0{,}9; 0{,}65; 0{,}3; 0{,}1]$ (informações dinâmicas e estáticas do sistema). Buscou-se uma variação uniforme dos pesos (ω_d e ω_s) de modo a melhor estudar os resultados obtidos pela aplicação da técnica.

A fim de facilitar a exposição dos resultados, os modelos serão denominados de Λ_1, Λ_2, Λ_3 e Λ_4. Também foi obtido um modelo por meio da técnica ERR, que será denominado de Λ_{ERR}. O critério de informação de Akaike foi utilizado para determinar o número de termos do modelo, resultando em 6 termos. Porém, uma vez que a estrutura é conhecida, sabe-se que existem somente 5 regressores genuínos.

Após a aplicação da técnica, obteve-se os seguintes modelos para o sistema simulado 1:

$$\begin{aligned} y_{\Lambda_{ERR}}(k) &= 1{,}0091u(k-1)^2 + 0{,}4948y(k-1) \\ &\quad -0{,}0492y(k-2)^2 + 0{,}8278u(k-2) \\ &\quad +0.5203 - 0{,}0047u(k-2)y(k-1), \end{aligned} \tag{5.20}$$

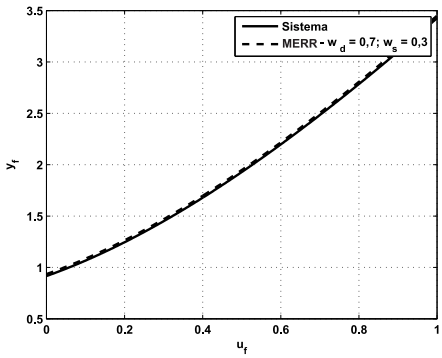

Figura 5.7: Curva estática do modelo Λ_3. Sistema 1

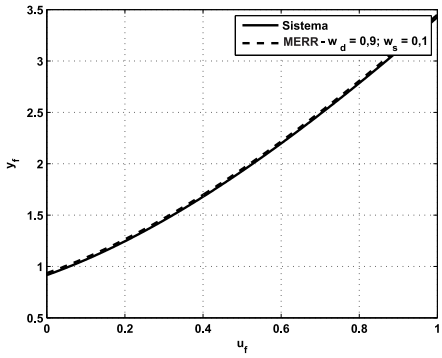

Figura 5.8: Curva estática do modelo Λ_4. Sistema 1

$$
\begin{aligned}
y_{\Lambda_1}(k) &= 0{,}5821y(k-1) + 0{,}8184u(k-2) + 0{,}8348 \\
&\quad +0{,}9459u(k-1)^2 - 0{,}3209y(k-2) \\
&\quad -0{,}0292y(k-2)y(k-1),
\end{aligned} \tag{5.21}
$$

$$
\begin{aligned}
y_{\Lambda_2}(k) &= 0{,}7019y(k-1) + 0{,}8018u(k-2) - \\
&\quad 0{,}0695y(k-2)y(k-1) + 0{,}9786u(k-1)^2 \\
&\quad +0{,}0042u(k-1)y(k-1) - 0{,}0812u(k-1)y(k-2),
\end{aligned} \tag{5.22}
$$

$$
\begin{aligned}
y_{\Lambda_3}(k) &= 0{,}5245 + 1{,}0083u(k-1)^2 + 0{,}8074u(k-2) \\
&\quad +0{,}4933y(k-1) - 0{,}0493y(k-2)^2 - 0{,}0015u(k-1),
\end{aligned} \tag{5.23}
$$

$$
\begin{aligned}
y_{\Lambda_4}(k) &= 1{,}0083u(k-1)^2 + 0{,}4933y(k-1) - 0{,}0493y(k-2)^2 \\
&\quad +0{,}8074u(k-2) + 0{,}5245 - 0{,}0015u(k-1).
\end{aligned} \tag{5.24}
$$

As Figuras (5.4), (5.5), (5.6), (5.7) e (5.8) apresentam a curva estática obtida para os modelos Λ_{ERR}, Λ_1, Λ_2, Λ_3 e Λ_4, respectivamente. Os modelos

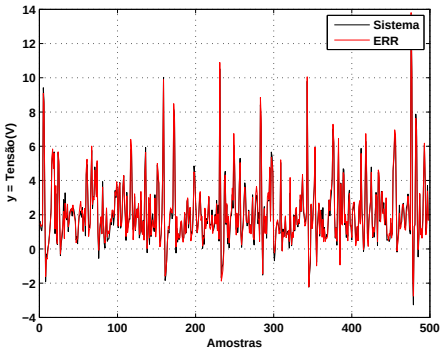

Figura 5.9: Simulação livre do modelo Λ_{ERR}.

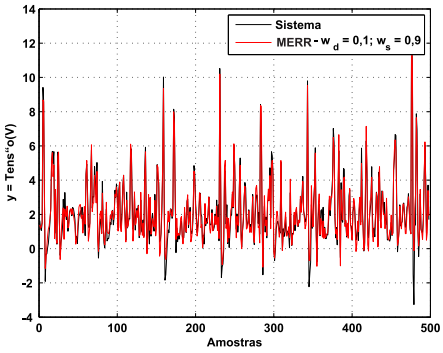

Figura 5.10: Simulação livre do modelo Λ_1.

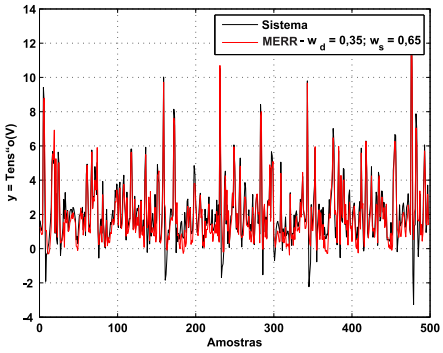

Figura 5.11: Simulação livre do modelo Λ_2.

obtidos foram simulados por meio de simulação livre, sendo os desempenhos apresentados nas Figuras (5.9), (5.10), (5.11), (5.12) e (5.13).

A Tabela (5.1) mostra o valor dos índices utilizados para validação dos modelos. Os custos quadráticos dinâmicos e estáticos foram normalizados, visando facilitar a interpretação dos resultados. Esses mesmos índices de validação podem ser vistos graficamente nas Figuras (5.14) e (5.15).

Pode ser percebido que a técnica MERR foi capaz de reconstruir o modelo por meio dos modelos Λ_3 e Λ_4, ou seja, a técnica MERR quantificou os regressores genuínos como sendo os mais importante. O modelo Λ_1 errou somente um regressor, dos 5 regressores genuínos do sistema. A técnica Λ_{ERR} também teve a capacidade de reconstruir o sistema por meio do modelo obtido.

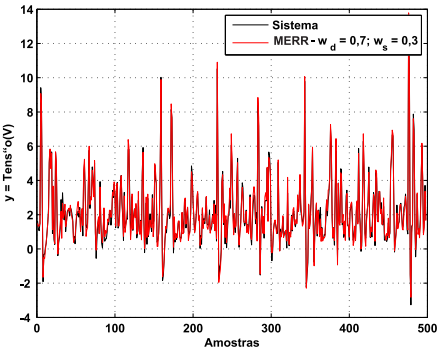

Figura 5.12: Simulação livre do modelo Λ_3.

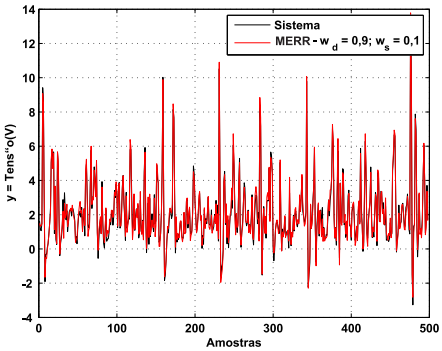

Figura 5.13: Simulação livre do modelo Λ_4.

Tabela 5.1: Índices calculados para validação dos modelos - sistema simulado 1. $RMSE_S$ se refere ao índice RMSE calculado por meio dos dados estáticos ao passo que $RMSE_D$ refere-se ao índice calculado utilizando dados dinâmicos.

Modelo	$RMSE_S$	$RMSE_D$	J_s	J_d
Λ_{ERR}	0,0342	0,1418	0,0028	0,0023
Λ_1	0,1837	0,2467	0,2234	0,3599
Λ_2	0,3843	0,3731	1	1
Λ_3	0,0275	0,1405	0	0
Λ_4	0,0275	0,1405	0	0

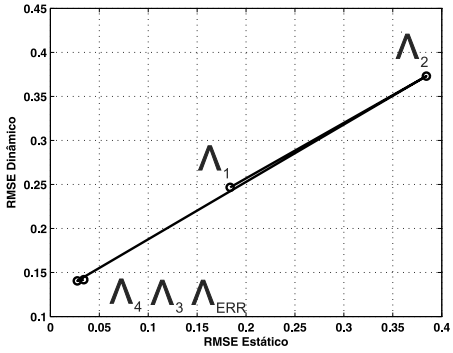

Figura 5.14: Espaço dos RMSE - Sistema Simulado 1.

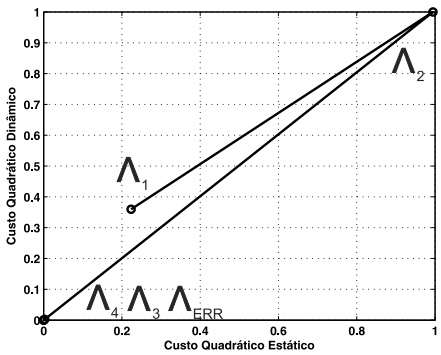

Figura 5.15: Espaço dos Erros Qudráticos - Sistema Simulado 1.

Tabela 5.2: MERR dos regressores e parâmetros- Sistema simulado 1.

Λ_{ERR}	Regressor	ERR (%)	Parâmetro do Sistema	Parâmetros do modelo
	$u(k-1)^2$	65,351	1,0000	1,0091
	y(k-1)	15,766	0,5000	0,4948
	$y(k-2)^2$	9,455	-0,0500	-0,0492
	u(k-2)	7,590	0,8000	0,8278
	C	1,224	0,5000	0,5203
	u(k-2)y(k-1)	0,003	-	0,0047
Λ_1	Regressor	MERR (%)	Parâmetro do Sistema	Parâmetros do modelo
	y(k-1)	98,898	0,5000	0,5821
	u(k-2)	0,172	0,8000	0,8184
	C	0,240	0,5000	0,8348
	$u(k-1)^2$	0,034	1,0000	0,9459
	y(k-2)	0,044	-	-0,3209
	y(k-2)y(k-1)	0,008	-	-0,0292
Λ_2	Regressor	MERR (%)	Parâmetro do Sistema	Parâmetros do modelo
	y(k-1)	87,768	0,5000	0,7019
	u(k-2)	2,707	0,8000	0,8018
	y(k-2)y(k-1)	2,301	-	-0,0695
	$u(k-1)^2$	2,080	1,0000	0,9786
	u(k-1)y(k-1)	2,004	-	0,0042
	u(k-1)y(k-2)	0,138	-	0,0812
Λ_3	Regressor	MERR (%)	Parâmetro do Sistema	Parâmetros do modelo
	C	69,301	0,5000	0,5245
	$u(k-1)^2$	14,083	1,0000	1,0083
	u(k-2)	4,823	0,8000	0,8074
	y(k-1)	3,399	0,5000	0,4933
	$y(k-2)^2$	6,324	-0,0500	-0,0493
	u(k-1)	1,068	-	-0,0015
Λ_4	Regressor	MERR (%)	Parâmetro do Sistema	Parâmetros do modelo
	$u(k-1)^2$	64,458	1,0000	1,0083
	y(k-1)	17,799	0,5000	0,4933
	$y(k-2)^2$	8,486	-0,0500	-0,0493
	u(k-2)	7,099	0,8000	0,8074
	C	1,479	0,5000	0,5245
	u(k-1)	0,106	-	-0,0015

O critério de informação de Akaike sugeriu um modelo com 6 termos ao invés de 5. Isso fez com que todos os modelos apresentassem um termo espúrio. Contudo, nos modelos Λ_3, Λ_4 e Λ_{ERR}, a quantificação de tal termo pela taxa de redução de erro multi(mono)-objetivo apresentou valor desprezível. Vale destacar que o valor do parâmetro associado ao termo espúrio é reduzido para os modelos obtidos via MERR, sugerindo a contribuição da incorporação de outras informações também na etapa de determinação dos parâmetros.

A técnica pode gerar alguns modelos dominados por outros, que apresentam termos espúrios, devido a combinação convexa dos pesos estático e dinâmico. Este foi o caso dos modelos Λ_1 e Λ_2, que devem ser descartados e não utilizados para identificação do sistema, embora tenham tido um bom desempenho dinâmico. Isso se dá uma vez que um modelo mais global, representativo do sistema, é requerido.

A Tabela (5.2) apresenta a taxa de redução de erro multi-objetivo dos modelos obtidos para a representação do sistema simulado 1, bem como os respectivos parâmetros associados. Observa-se que o valor da taxa de redução de erro do último regressor (não-genuíno) nos modelos Λ_3 e Λ_4 é expressivamente inferior aos demais, sugerindo a sua não-incorporação ao modelo do sistema.

Sendo assim, de acordo com os resultados obtidos para este sistema simulado, a técnica MERR, uma expansão da ERR, foi capaz de obter modelos genuínos, pertinentes ao conjunto pareto-ótimo, e separá-los dos espúrios, quando no processo de validação dos mesmos.

5.3.3 Exemplo Simulado 2

O exemplo simulado 2 é uma extensão do exemplo 1, trabalhado em (Bonin et al., 2010), quando do desenvolvimento da técnica *SEM-LASSO*. O sistema é praticamente o mesmo, com exceção da entrada. Esta é dada por uma combinação linear de uma componente auto-regressiva e um ruído branco gaussiano, com média 0 e variância 1, formando um filtro passa-baixas, a saber:

$$u(k) = 0{,}5u(k-1) + w(k), \tag{5.25}$$

sendo $w(\cdot)$ o ruído gaussiano branco.

O uso deste tipo de excitação pode confundir o algoritmo durante o pro-

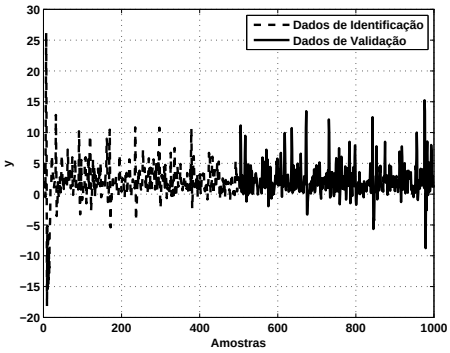

Figura 5.16: Dados de saída do sistema simulado 2. Em pontilhado encontram-se os dados utilizados para identificação e em linha contínua os dados que foram utilizados para validação.

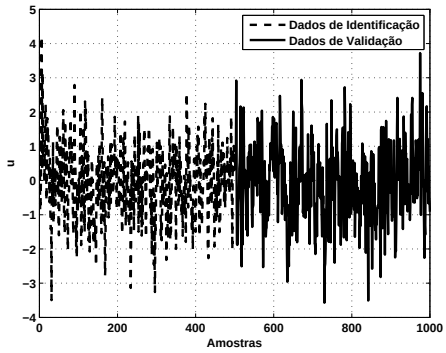

Figura 5.17: Dados de entrada do sistema simulado 2. Em pontilhado encontram-se os dados utilizados para identificação e em linha contínua os dados que foram utilizados para validação.

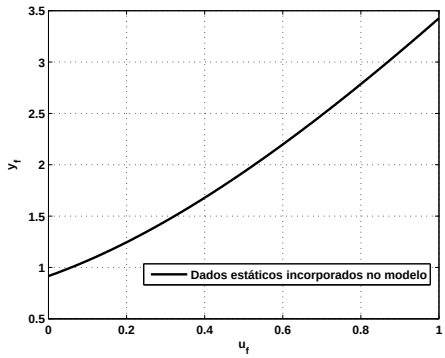

Figura 5.18: Dados estáticos do sistema simulado 2 incorporados ao modelo do sistema.

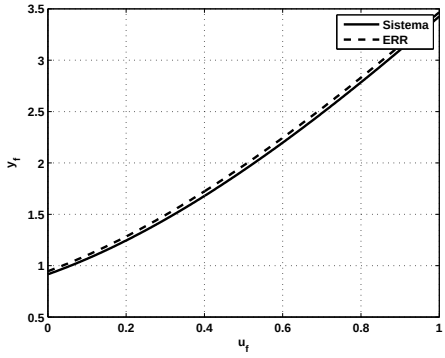

Figura 5.19: Curva estática do modelo Λ_{ERR}. Sistema 2

cesso de seleção, pois reduz a diferença de termos como $y(t)$ e $y(t-1)$ (Bonin et al., 2010).

Os dados estáticos foram gerados de acordo com a teoria de coeficientes de agrupamentos e curva estática, apresentados no Capítulo 3 e são apresentados na Figura (5.18). Os dados de saída e entrada podem ser vistos nas Figuras (5.16) e (5.17), respectivamente.

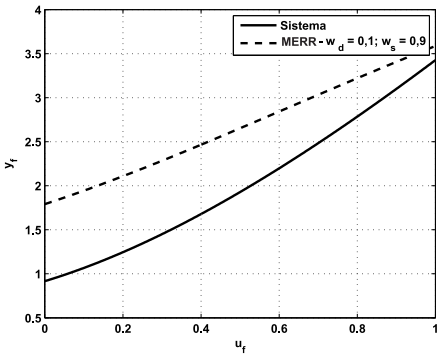

Figura 5.20: Curva estática do modelo Λ_1. Sistema 2

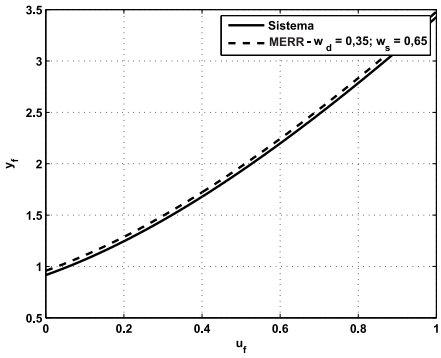

Figura 5.21: Curva estática do modelo Λ_2. Sistema 2

5.3.4 Resultados e Discussão dos Resultados - Exemplo Simulado 2

A metodologia da taxa de redução de erro multi-objetivo (MERR) foi empregada na identificação do exemplo simulado 2, utilizando para tal $\omega_d = [0{,}1; 0{,}35; 0{,}7; 0{,}9]$ e consequentemente $\omega_s = [0{,}9; 0{,}65; 0{,}3; 0{,}1]$ (informações dinâmicas e estáticas do sistema). Visando analisar a técnica ao longo da curva de Pareto, buscou-se a variação uniforme dos pesos.

A fim de facilitar a exposição dos resultados, os modelos serão denominados de Λ_1, Λ_2, Λ_3 e Λ_4. Também foi obtido um modelo por meio da técnica ERR, que será denominado de Λ_{ERR}. O critério de informação de Akaike foi utilizado para determinar o número de termos do modelo, resultando em 7 termos. De modo semelhante ao sistema 1, sabe-se que há somente 5 regressores genuínos. Outra vez, o AIC não foi capaz de obter o número correto de regressores genuínos.

Após a aplicação da técnica, obteve-se os seguintes modelos para o sistema simulado 2:

$$\begin{aligned} y_{\Lambda_{ERR}}(k) &= 1{,}0185u(k-1)^2 - 0{,}0490y(k-2)^2 + 0{,}5098y(k-1) \\ &\quad +0{,}8436u(k-2) + 0{,}5199 - 0{,}0573u(k-2)u(k-1) \\ &\quad -0{,}0084u(k-2)y(k-1), \end{aligned} \tag{5.26}$$

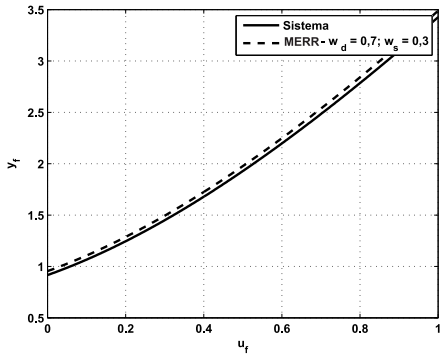

Figura 5.22: Curva estática do modelo Λ_3. Sistema 2

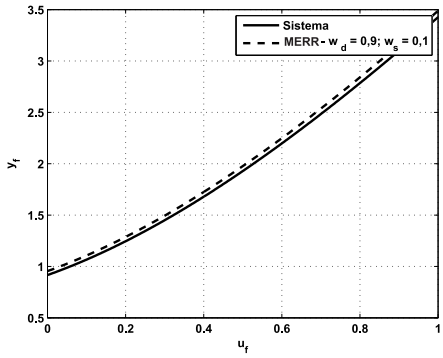

Figura 5.23: Curva estática do modelo Λ_4. Sistema 2

$$
\begin{aligned}
y_{\Lambda_1}(k) &= 0{,}4790y(k-1) + 0{,}8936u(k-2) - 0{,}0444y(k-2)^2 \quad (5.27)\\
&\quad +0{,}0427y(k-2) - 0{,}0395u(k-2)y(k-2)\\
&\quad +1{,}3504 + 0{,}2121u(k-2)^2,
\end{aligned}
$$

$$
\begin{aligned}
y_{\Lambda_2}(k) &= 0{,}4962y(k-1) + 0{,}9962u(k-1)^2 - 0{,}0489y(k-2)^2 \quad (5.28)\\
&\quad +0{,}5446 + 0{,}0037u(k-2)y(k-2)\\
&\quad +0{,}0099u(k-1) + 0{,}7904u(k-2),
\end{aligned}
$$

$$
\begin{aligned}
y_{\Lambda_3}(k) &= 0{,}4932y(k-1) + 0{,}9960u(k-1)^2 - 0{,}0491y(k-2)^2 \quad (5.29)\\
&\quad +0{,}8007u(k-2) + 0{,}5355\\
&\quad +0{,}0118u(k-1) + 0{,}0089y(k-2),
\end{aligned}
$$

$$
\begin{aligned}
y_{\Lambda_4}(k) &= 0{,}4932y(k-1) + 0{,}9960u(k-1)^2 - 0{,}0491y(k-2)^2 \quad (5.30)\\
&\quad +0{,}8007u(k-2) + 0{,}5355\\
&\quad +0{,}0118u(k-1) + 0{,}0089y(k-2).
\end{aligned}
$$

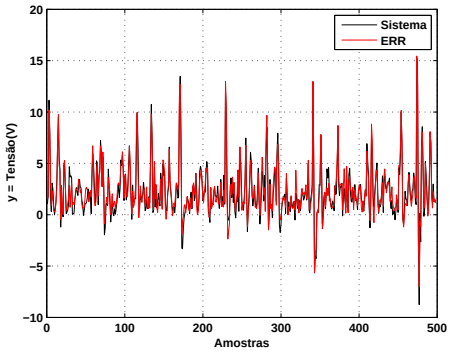

Figura 5.24: Simulação livre do modelo Λ_{ERR}.

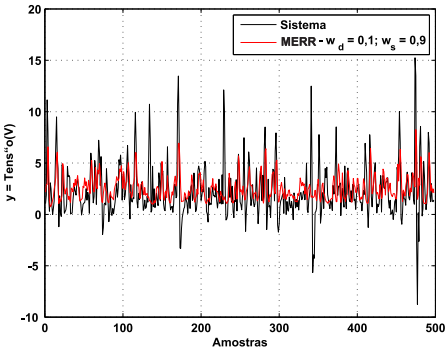

Figura 5.25: Simulação livre do modelo Λ_1.

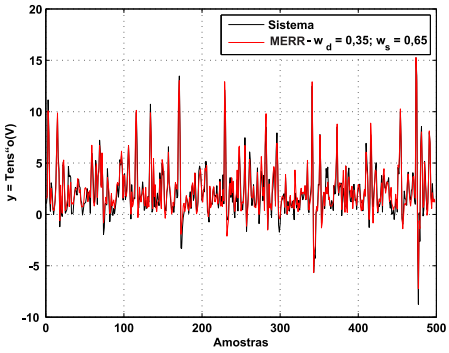

Figura 5.26: Simulação livre do modelo Λ_2.

As Figuras (5.19), (5.20), (5.21), (5.22) e (5.23) apresentam a curva estática obtida para os modelos Λ_{ERR}, Λ_1, Λ_2, Λ_3 e Λ_4, respectivamente. Os modelos obtidos foram simulados por meio de simulação livre, sendo os desempenhos apresentados nas Figuras (5.24), (5.25), (5.26), (5.27) e (5.28).

A Tabela (5.3) mostra o valor dos índices utilizados para validação dos modelos. Os custos quadráticos dinâmicos e estáticos foram normalizados, visando facilitar a interpretação dos resultados. Esses mesmos índices de validação podem ser vistos graficamente nas Figuras (5.29) e (5.30).

A Tabela (5.4) apresenta a taxa de redução de erro multi-objetivo dos modelos obtidos para a representação do sistema simulado 2 e os respectivos parâmetros associados. Ainda neste sistema, as técnicas ERR e MERR foram capazes de reconstruir o sistema original (modelos Λ_3 e Λ_4), utilizando dados

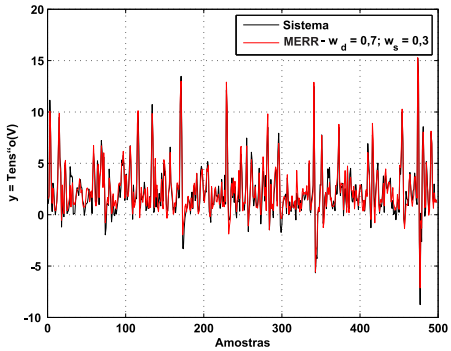

Figura 5.27: Simulação livre do modelo Λ_3.

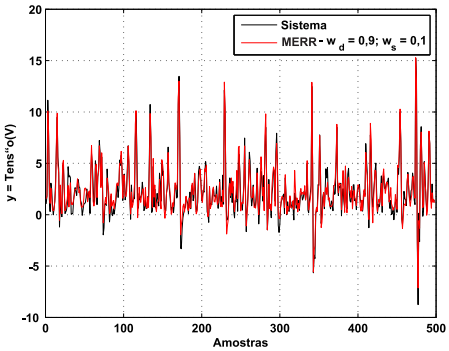

Figura 5.28: Simulação livre do modelo Λ_4.

Tabela 5.3: Índices calculados para validação dos modelos - sistema simulado 2. $RMSE_S$ se refere ao índice RMSE calculado por meio dos dados estáticos ao passo que $RMSE_D$ refere-se ao índice calculado utilizando dados dinâmicos.

Modelo	$RMSE_S$	$RMSE_D$	J_s	J_d
Λ_{ERR}	0,0581	0,2834	0	0,0034
Λ_1	0,9324	0,8937	1	1
Λ_2	0,0615	0,2775	0,0005	0
Λ_3	0,0667	0,2779	0,0012	0,0006
Λ_4	0,0667	0,2779	0,0012	0,0006

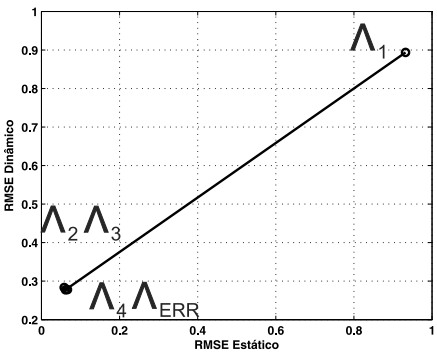

Figura 5.29: Espaço dos RMSE - Sistema Simulado 2.

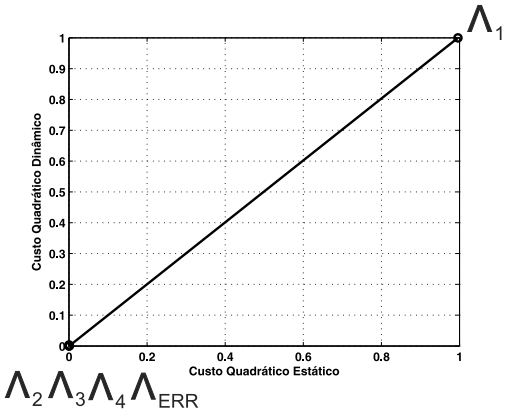

Figura 5.30: Espaço dos Erros Qudráticos - Sistema Simulado 2.

Tabela 5.4: MERR dos regressores e parâmetros- Sistema simulado 2.

Λ_{ERR}	Regressor	ERR (%)	Parâmetro do Sistema	Parâmetros do modelo
	$u(k-1)^2$	44,850	1,0000	1,0185
	$y(k-2)^2$	20,650	-0,0500	-0,0490
	y(k-1)	25,720	0,5000	0,5098
	u(k-2)	5,234	0,8000	0,8436
	C	1,128	0,5000	0,5199
	u(k-2)u(k-1)	0,022	-	-0,0573
	u(k-2)y(k-1)	0,017	-	-0,0084
Λ_1	Regressor	MERR (%)	Parâmetro do Sistema	Parâmetros do modelo
	y(k-1)	97,980	0,5000	0,4790
	u(k-2)	0,265	0,8000	0,8936
	$y(k-2)^2$	0,300	-0,0500	-0,0444
	y(k-2)	0,124	-	0,0427
	u(k-2)y(k-2)	0,165	-	-0,0395
	C	0,053	0,5000	1,3504
	$u(k-2)^2$	0,085	-	0,2121
Λ_2	Regressor	MERR (%)	Parâmetro do Sistema	Parâmetros do modelo
	y(k-1)	80,298	0,5000	0,4962
	$u(k-1)^2$	4,749	1,0000	0,9962
	$y(k-2)^2$	4,945	-0,0500	-0,0489
	C	2,918	0,5000	0,5446
	u(k-2)y(k-2)	0,879	-	0,0037
	u(k-1)	1,591	1,0000	0,0099
	u(k-2)	0,346	0,8000	0,7904
Λ_3	Regressor	MERR (%)	Parâmetro do Sistema	Parâmetros do modelo
	y(k-1)	55,218	0,5000	0,4932
	$u(k-1)^2$	17,630	1,0000	0,9960
	$y(k-2)^2$	18,906	-0,0500	-0,0491
	u(k-2)	3,053	0,8000	0,8007
	C	2,050	0,5000	0,5355
	u(k-1)	0,812	-	0,0118
	y(k-2)	0,155	-	0,0089
Λ_4	Regressor	MERR (%)	Parâmetro do Sistema	Parâmetros do modelo
	y(k-1)	46,779	0,5000	0,4932
	$u(k-1)^2$	20,764	1,0000	0,9960
	$y(k-2)^2$	24,223	-0,0500	-0,0491
	u(k-2)	4,510	0,8000	0,8007
	C	1,386	0,5000	0,5355
	u(k-1)	0,066	-	0,0118
	y(k-2)	0,028	-	0,0089

dinâmicos (ERR) ou dinâmicos e estáticos (MERR). Mais uma vez, a contribuição dos parâmetros associados aos regressores não-genuínos (inseridos no modelo devido à falha do AIC) é menor quando da utilização da MERR, sugerindo novamente que a inserção de outras informações também contribuem para a estimação dos parâmetros do modelo. Os modelos dados por Λ_1 e Λ_2 se encontram suplantados pelos demais, e não devem ser utilizados para representar o sistema por não pertencerem ao conjunto Pareto-ótimo.

A variação da entrada como uma combinação linear de uma componente auto-regressivo com um ruído gaussiano branco fez com que a modelagem fosse dificultada, quando baseada somente nos dados dinâmicos. Prova disso é que o AIC sugeriu uma quantidade de 7 termos ao invés de 5 para identificar o sistema. Contudo, a técnica MERR foi robusta o suficiente para classificar os regressores genuínos com maior importância.

Verifica-se por meio dos exemplos simulados a qualidade da taxa de redução de erro multi-objetivo, uma vez que tal técnica foi capaz de classificar os regressores de modo correto. Os modelos pertinentes ao conjunto Pareto-ótimo apresentaram parâmetros bem próximos aos do sistema modelado. Em ambos os casos, existem modelos pertinentes ao conjunto Pareto-ótimo que apresentam todos os regressores genuínos com maior índice MERR.

5.4 O Conversor CC-CC Buck

5.4.1 Descrição do Sistema

Um dos sistemas reais utilizados para teste da técnica desenvolvida foi um conversor CC-CC do tipo buck. Esse sistema eletrônico de potência, como o próprio nome sugere, é um conversor abaixador, e foi utilizado quando operando em modo contínuo, ou seja, a corrente através do indutor não se anulava. Tal sistema foi utilizado por apresentar uma dinâmica altamente não-linear e um comportamento estático afim. Esta característica faz com que a modelagem, visto de uma forma global, seja dificultada.

Durante os testes, a fonte de alimentação CC foi mantida constante, em $24V$. A razão cíclica, definida como o tempo em que a chave se encontra ligada em relação ao tempo total de operação ($D = T_{on}/T$), foi variada utilizando técnicas de modulação por largura de pulso (PWM) a uma taxa de $33kHz$, por meio de um circuito integrado LM3524 [1].

[1] Os dados foram originalmente coletados em (Aguirre et al., 2000)

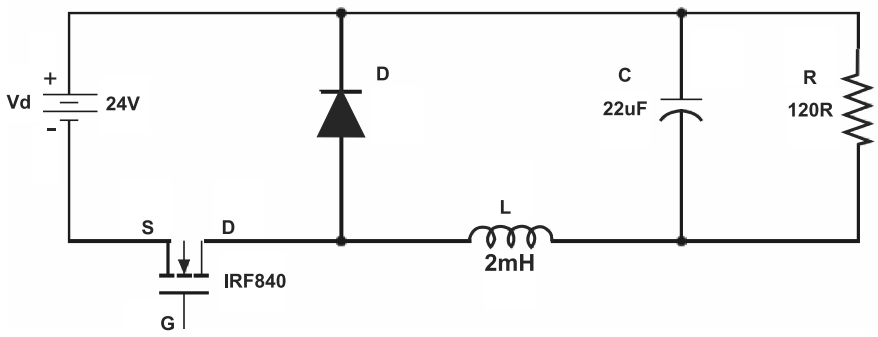

Figura 5.31: Conversor CC-CC buck.

Quando a razão cíclica está perto da unidade, a corrente através do indutor e a tensão de saída na carga aumentam, pois a fonte CC energiza a malha formada por ela, pelo capacitor e pelo indutor. Quando a razão tende a zero, a tensão de carga diminui, com um regime dinâmico distinto, caracterizando a não-linearidade do sistema (Aguirre et al., 2000).

5.4.2 Testes Dinâmicos

Visando estimular toda a dinâmica do sistema, foi projetado um sinal PRBS (sinal binário pseudo-aleatório, do inglês *Pseudo Random Binary Signal*) para ser aplicado no *gate* do MOSFET IRF840.

Para o projeto deste tipo de sinal, é necessário que a constante de tempo predominante do sistema seja estimada. De acordo com o teste de resposta ao degrau realizado, o valor dessa constante foi de aproximadamente $2ms$, com um tempo de amostragem de $T_s = 10\mu s$. Esses parâmetros permitiram a construção do sinal PRBS.

Em sinais PRBS, dois parâmetros são necessários para sua construção. O número de *bits*, b, e o tempo de chaveamento, T_B (Aguirre, 2007). O número de bits deve ser escolhido de forma a garantir que o período do sinal seja maior que o tempo de acomodação do sistema (cinco vezes a constante de tempo predominante do sistema).

Os sinais coletados foram superamostrados. Deste modo, foi necessário decimar o sinal por um fator $\Delta = 12$, gerando os sinais de trabalho, apresentados nas Figuras (5.33) e (5.32). Ressalta-se que os dados dinâmicos excursiona somente a faixa de $2{,}2V < u < 2{,}5V$.

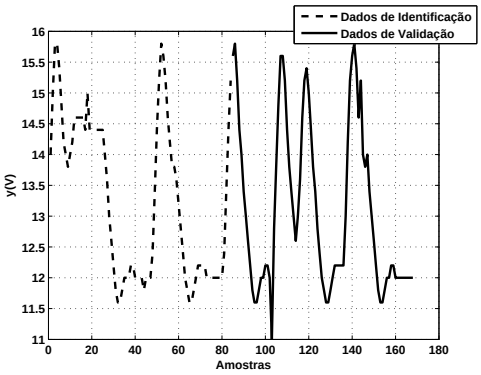

Figura 5.32: Dados de saída do conversor CC-CC buck. Em pontilhado encontram-se os dados utilizados para identificação e em linha contínua os dados que foram utilizados para validação.

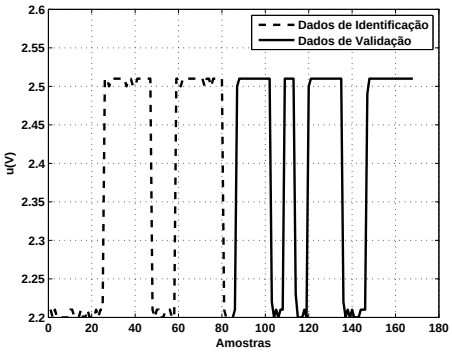

Figura 5.33: Dados de entrada do conversor CC-CC buck. Em pontilhado encontram-se os dados utilizados para identificação e em linha contínua os dados que foram utilizados para validação.

5.4.3 Curva Estática do Conversor CC-CC Buck

Além das características dinâmicas do sistema, é altamente desejável que o modelo apresente também suas características estáticas. Pela teoria de eletrônica de potência, a curva estática de um conversor CC-CC buck pode ser expressa pela equação (5.31).

$$\begin{aligned} V_{\mathrm{o}} &= (1-D)V_{\mathrm{d}} \\ &= \left(1-\frac{\bar{u}-1}{3}\right)V_{\mathrm{d}} \\ &= \frac{4V_{\mathrm{d}}}{3}-\frac{V_{\mathrm{d}}}{3}\bar{u}, \end{aligned} \tag{5.31}$$

sendo $y = V_{\mathrm{o}}$ a tensão de saída, D a razão ciclica, V_{d} o valor da tensão CC de alimentação ($24V$ neste trabalho) e $\bar{u}$ o valor de entrada em estado estacionário. Variando-se a entrada em estado estacionário, excursiona-se a curva estática na região de interesse. Neste trabalho, foram incorporados pontos estáticos na região de $0V < u < 10V$, como mostra a Figura (5.34)

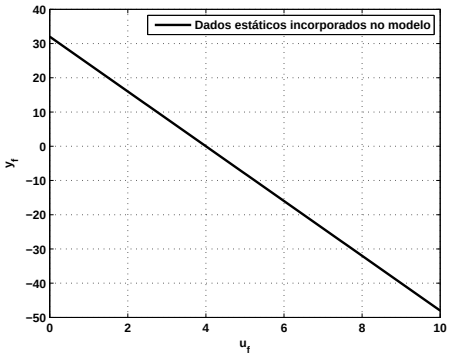

Figura 5.34: Dados estáticos do conversor CC-CC buck que foram incorporados ao modelo do sistema.

5.4.4 Resultados e Discussão dos Resultados - Conversor CC-CC Buck

A metodologia da taxa de redução de erro multi-objetivo (MERR) foi empregada na identificação de um conversor CC-CC buck, quando excitado por um sinal *PRBS*. Para tal, utilizou-se $\omega_d = [0{,}1; 0{,}35; 0{,}7; 0{,}9]$ e consequentemente $\omega_s = [0{,}9; 0{,}65; 0{,}3; 0{,}1]$ (foram utilizadas informações dinâmicas e estáticas do sistema). Mais uma vez, buscou-se a variação uniforme dos pesos ω_d e ω_s.

A fim de facilitar a exposição dos resultados, os modelos serão denominados de Λ_1 (modelo 1), Λ_2 (modelo 2), Λ_3 (modelo 3) e Λ_4 (modelo 4). Também foi obtido um modelo por meio da técnica ERR, que será denominado de Λ_{ERR}. O critério de informação de Akaike foi utilizado para determinar o número de termos do modelo, resultando em 9 termos.

Os modelos obtidos para Λ_1 e Λ_2 foram exatamente os mesmos. Por esta razão, o modelo será denominado de $\Lambda_{1,2}$ durante a exposição e análise dos resultados. Tais modelos serão expostos a seguir:

$$\begin{aligned} y_{\Lambda_{ERR}}(k) &= -0{,}6867y(k-1) - 0{,}3253y(k-2) \\ &\quad -0{,}0081y(k-1)^2 - 2{,}8685u(k-1)^2 - 0{,}0404y(k-2)^2 \\ &\quad +0{,}0593y(k-2)y(k-1) + 14{,}0007 \\ &\quad +0{,}6153u(k-1)y(k-1) + 3{,}2458u(k-1), \end{aligned} \tag{5.32}$$

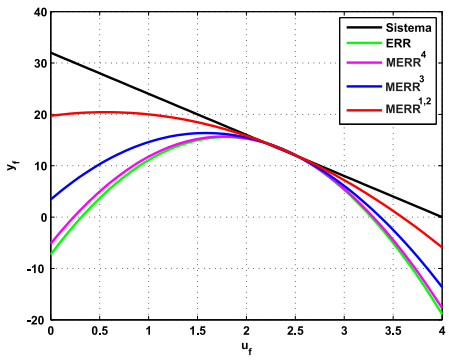

Figura 5.35: Curva estática do conversor CC-CC buck e dos modelos obtidos.

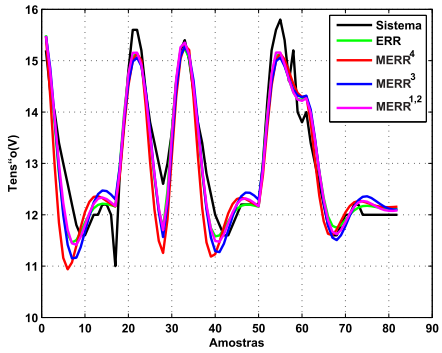

Figura 5.36: Comportamento dinâmico (simulação livre) dos modelos obtidos.

$$
\begin{aligned}
y_{\Lambda_{1,2}}(k) &= 1{,}6168y(k-1) - 1{,}1554y(k-2) - 2{,}1340u(k-1) \\
&\quad +13{,}9972 - 2{,}8396u(k-2) - 0{,}0744y(k-2)^2 \\
&\quad +0{,}1519y(k-2)y(k-1) - 0{,}0863y(k-1)^2 \\
&\quad +0{,}2078u(k-2)y(k-2),
\end{aligned} \tag{5.33}
$$

$$
\begin{aligned}
y_{\Lambda_3}(k) &= -0{,}4849y(k-1) + 0{,}1546y(k-2) - 3{,}0944u(k-1)^2 \\
&\quad +0{,}3150u(k-2)^2 + 1{,}2441u(k-2)u(k-1) + 14{,}0006 \\
&\quad -0{,}7677u(k-1) + 0{,}7952u(k-1)y(k-1) \\
&\quad -0{,}3438u(k-2)y(k-2),
\end{aligned} \tag{5.34}
$$

$$
\begin{aligned}
y_{\Lambda_4}(k) &= -0{,}0396y(k-1) - 1{,}0472y(k-2) + 14{,}0005 \\
&\quad -5{,}6780u(k-1) + 9{,}3834u(k-2) - 2{,}0313u(k-2)^2 \\
&\quad +0{,}5945u(k-1)y(k-1) + 0{,}0152y(k-2)^2 \\
&\quad -0{,}8788u(k-1)^2.
\end{aligned} \tag{5.35}
$$

A Figura (5.35) apresenta o comportamento estático dos modelos obtidos,

Tabela 5.5: Índices calculados para validação dos modelos do conversor CC-CC buck.

Modelo	$RMSE_S$	$RMSE_D$	J_s	J_d
Λ_{ERR}	1,5824	0,3004	1	0
Λ_1	0,4930	0,4802	0	1
Λ_2	0,4930	0,4802	0	1
Λ_3	1,1493	0,3545	0,4767	0,2737
Λ_4	1,4959	0,2935	0,8822	0,2510

bem como do sistema. Observa-se claramente um grande aumento na faixa de precisão estática do modelo, de acordo com o aumento de ω_s, ou seja, o erro estático é drasticamente diminuído, de acordo com o incremento da contribuição estática durante o processo de detecção de estruturas. Neste caso dados estáticos mostraram-se importante para a obtenção de um modelo mais global, valido em uma maior faixa de operação.

A fim de validar o comportamento dinâmico dos modelos, realizou-se a predição livre dos mesmos (Figura (5.36)). Qualitativamente, observa-se que todos os modelos conseguem seguir a curva de tendência, oscilações e não-linearidades do sistema.

Visando quantificar os resultados estáticos e dinâmicos obtidos pela técnica por meio dos diversos modelos, calculou-se o índice *RMSE* estático e dinâmico, bem como o custo quadrático, estático e dinâmico (J_s e J_d). Tais índices são representados graficamente nas Figuras (5.37) e (5.38), e por meio da Tabela (5.5), no intuito de facilitar a análise dos resultados.

Um ponto interessante e que deve ser ressaltado é que, embora por um pequeno valor, a luz dos índices RMSE, a ERR encontra-se suplantada por outros modelos. Dessa forma, conjectura-se que existem informações presentes nos dados estáticos necessários para a modelagem do sistema como um todo. Ademais, a incorporação de dados estáticos foi capaz de gerar modelos mais robustos e globais, válidos em uma maior faixa de operação estática.

As Tabelas (5.6) e (5.7) apresentam a taxa de redução de erro multi-objetivo dos modelos obtidos para a representação do conversor CC-CC buck estudado nessa seção. Interessante ressaltar que ainda que os modelos Λ_1 e Λ_2 sejam os mesmos, o indice MERR referente a cada regressor não necessariamente é o mesmo. Isso pode ser explicado, pois estes variam de acordo com o peso dado a cada objetivo a ser minimizado.

Tabela 5.6: Taxa de Redução de Erro Multi-objetivo - Conversor CC-CC buck.

Λ_{ERR}	Regressor	ERR (%)
	y(k-1)	99,918601
	y(k-2)	0,040465
	$y(k-1)^2$	0,005679
	$u(k-1)^2$	0,014101
	$y(k-2)^2$	0,000065
	y(k-2)y(k-1)	0,001073
	C	0,000218
	u(k-1)y(k-1)	0,000605
	u(k-1)	0,000596
Λ_1	Regressor	MERR (%)
	y(k-1)	99,850869
	y(k-2)	0,148949
	u(k-1)	0,000036
	C	0,000063
	u(k-2)	0,000001
	$y(k-2)^2$	0,000001
	y(k-2)y(k-1)	0,000001
	$y(k-1)^2$	0,000003
	u(k-2)y(k-2)	0,000002
Λ_2	Regressor	MERR (%)
	y(k-1)	99,820737
	y(k-2)	0,174741
	u(k-1)	0,000856
	C	0,001604
	u(k-2)	0,000010
	$y(k-2)^2$	0,000008
	y(k-2)y(k-1)	0,000037
	$y(k-1)^2$	0,000076
	u(k-2)y(k-2)	0,000045

Tabela 5.7: Taxa de Redução de Erro Multi-objetivo - Conversor CC-CC buck (continuação).

Λ_3	Regressor	MERR (%)
	y(k-1)	99,899770
	y(k-2)	0,064284
	$u(k-1)^2$	0,001337
	$u(k-2)^2$	0,005358
	u(k-2)u(k-1)	0,001061
	C	0,000807
	u(k-1)	0,009267
	u(k-1)y(k-1)	0,000115
	u(k-2)y(k-2)	0,000323
Λ_4	Regressor	MERR (%)
	y(k-1)	99,913069
	y(k-2)	0,042932
	C	0,002313
	u(k-1)	0,017904
	u(k-2)	0,000103
	$u(k-2)^2$	0,000057
	u(k-1)y(k-1)	0,001063
	$y(k-2)^2$	0,000315
	$u(k-1)^2$	0,000291

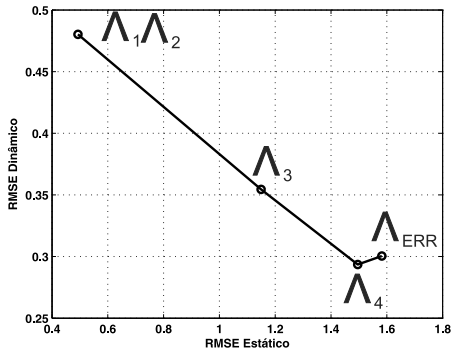

Figura 5.37: Espaço dos RMSE - conversor CC-CC buck.

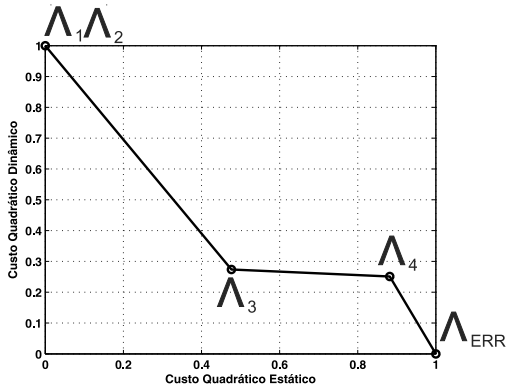

Figura 5.38: Espaço dos Erros Qudráticos - conversor CC-CC buck.

Uma vez que não se tem o modelo original do sistema, não se pode ter certeza que os regressores obtidos são genuínos. Contudo, pode-se afirmar que o acréscimo da informação da curva estática fez com que a representação estática do modelo fosse aprimorada, a um baixo custo no seu desempenho dinâmico (à luz de J_s e J_d), ou até mesmo melhorando seu desempenho dinâmico (considerando os índices RMSE estático e dinâmico). Isso pode ser explicado, uma vez que os objetivos aparentam ser conflitantes e que há informações contidas nos dados estáticos, relevantes durante o processo de identificação. Outro ponto extremamente positivo da técnica é o fato de disponibilizar uma família de soluções eficientes, sendo a escolha de qual modelo objetivada pela aplicação em questão.

No que se refere ao comportamento dinâmico, todos os modelos são satisfatórios. De acordo com a teoria, modelos que apresentam RMSE dinâmico menor que a unidade podem ser utilizados, uma vez que apresente o erro quadrático, em média, menor que o erro quadrático cometido pela média da série temporal. Já o comportamento estático de alguns modelos (Λ_3, Λ_4 e Λ_{ERR}, por exemplo), não são satisfatórios a luz do índice RMSE por apresentarem um erro maior que a unidade.

Por fim, vale lembrar que durante a identificação deste conversor não houveram modelos suplantados por outros (analisando os erros quadráticos), sendo todos os modelos eficientes do ponto de vista multi-objetivo. Deve ser lembrado que todo o processo de validação foi realizado sob a massa de dados

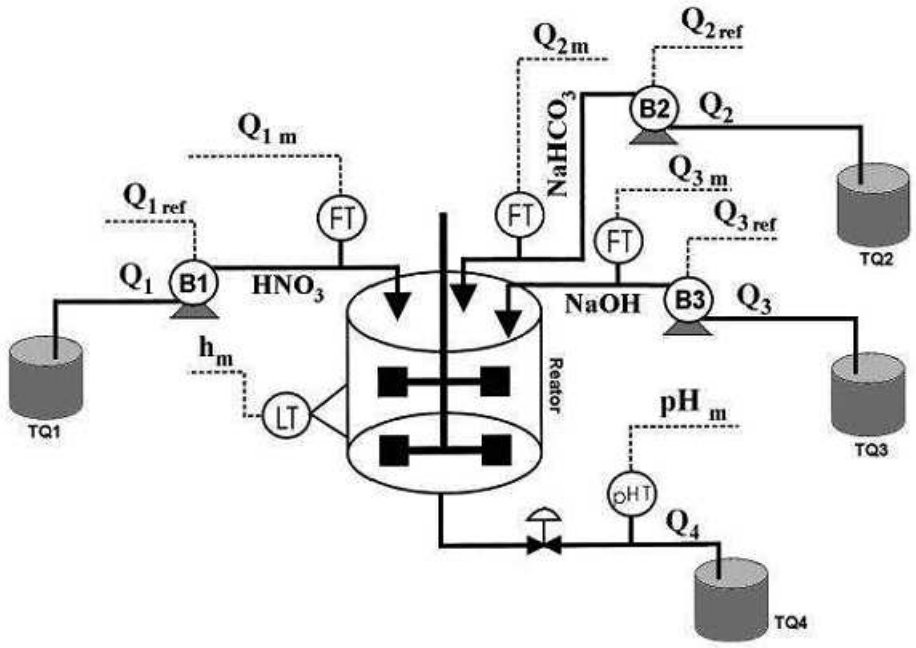

Figura 5.39: Esquema do processo de neutralização de pH.

de validação.

5.5 A Planta de Neutralização de PH

Processos de neutralização de pH têm sido por muito tempo uma referência para problemas não-lineares no controle de processos químicos devido à grande dificuldade apresentada em sua identificação e controle (Lima, 2009). Em particular, destaca-se a alta não-linearidade em suas características estáticas.

O processo de neutralização de pH apresentado neste trabalho [2] se resume na adição de três soluções, ácido, base e tampão, a um reator (Margoti, 2011). A vazão de base é definida como variável de entrada e o nível de acidez do pH é a variável de saída. A Figura 5.39 apresenta o diagrama esquemático da planta de neutralização de pH (Campos, 2007).

[2] A planta encontra-se atualmente no Laboratório de Identificação, Análise e Controle de Sistemas Não-lineares do Centro Universitário do Leste de Minas Gerais - UNILESTE.

Na Figura 5.39, $TQ1$ é o tanque de ácido ($HNO3$), $TQ2$ é o tanque de solução tampão ($NaHCO3$), $TQ3$ o tanque de base ($NaOH$) e $TQ4$ o coletor da mistura. $B1$, $B2$ e $B3$ são as bombas dosadoras de cada solução. A mistura das soluções ocorre no reator. As vazões de tampão e ácido são mantidas constantes, assim como o volume no reator.

5.5.1 Testes Dinâmicos

Os dados dinâmicos foram coletados na planta. Visando excitar toda a dinâmica não-linear da planta, um sinal PRBS foi projetado e aplicado à entrada da mesma.

Mais uma vez, os dados foram divididos em duas massas distintas, apresentadas nas Figuras (5.41) e (5.40), respectivamente. Os dados pontilhados foram utilizados para identificação, ao passo que os dados representados por meio de uma linha contínua foram utilizados para validação do modelo.

Com o intuito de verificar a eficiência da técnica apresentada, utilizou-se uma massa de dados dinâmicos bastante reduzida para identificação do sistema (utilizou-se poucas informações de frequência e amplitude do sistema). Isso permite ratificar a importância da incorporação da informação estática na detecção de estruturas de modelos NARMAX polinomiais. Outrossim, é possível verificar se a adição deste tipo de informação é realmente relevante, a ponto de influenciar na estabilidade e robustez do modelo, tornando-o mais global.

5.5.2 Curva Estática

A característica estática do processo de neutralização de pH pode ser obtida por meio do modelo fenomenológico que descreve a curva de titulação de pH, apresentada em (Campos, 2007). A vazão $Q3$ é variada, ponto a ponto, de modo a obter a resposta y (pH), em regime permanente. A Figura (5.42) apresenta os dados estáticos que foram incorporados ao modelo.

Deve ser ressaltado a alta não linearidade da característica estática da planta de neutralização de pH. Isso faz com que haja uma grande dificuldade na representação dessa característica do sistema pelo modelo, principalmente quando se trabalha com modelos de ordem reduzida.

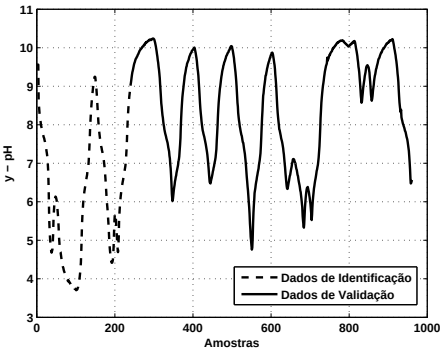

Figura 5.40: Dados de saída do processo de neutralização de pH. Em pontilhado encontram-se os dados utilizados para identificação e em linha contínua os dados que foram utilizados para validação.

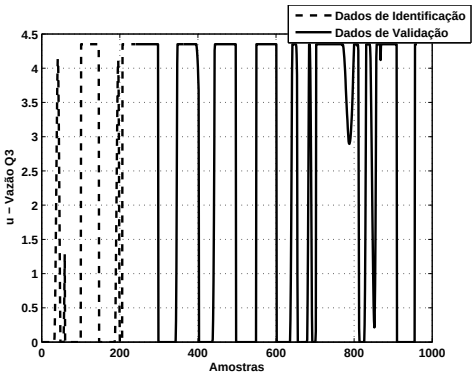

Figura 5.41: Dados de entrada do processo de neutralização de pH. Em pontilhado encontram-se os dados utilizados para identificação e em linha contínua os dados que foram utilizados para validação.

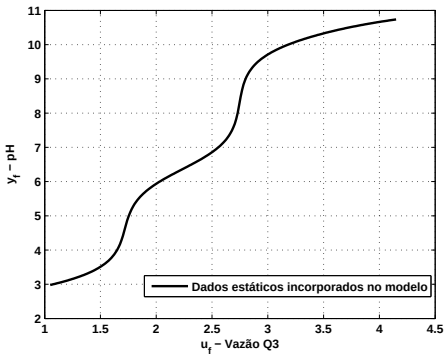

Figura 5.42: Dados estáticos do processo de neutralização de pH que foram incorporados ao modelo do sistema.

5.5.3 Resultados e Discussão dos Resultados - Planta de Neutralização de pH

A metodologia da taxa de redução de erro multi-objetivo (MERR) foi empregada na identificação de uma planta de neutralização de pH (Campos, 2007), quando excitado por um sinal PRBS. Para tal, utilizou-se $\omega_d = [0{,}1; 0{,}35; 0{,}7; 0{,}9]$ e consequentemente $\omega_s = [0{,}9; 0{,}65; 0{,}3; 0{,}1]$ (foram utilizadas informações dinâmicas e estáticas do sistema).

A fim de facilitar a exposição dos resultados, os modelos serão denominados de Λ_1 (modelo 1), Λ_2 (modelo 2), Λ_3 (modelo 3) e Λ_4 (modelo 4). Também foi obtido um modelo por meio da técnica ERR, que será denominado de Λ_{ERR}. O critério de informação de Akaike foi utilizado para determinar o número de termos do modelo, resultando em um modelo com 10 termos. Isso ressalta a alta não-linearidade presente também nos dados dinâmicos.

No intuito de verificar a aplicabilidade da técnica apresentada, foi utilizada uma massa de dados bastante reduzida para identificação. Sendo assim, desejou-se simular um caso em que apenas uma pequena quantidade de dados dinâmicos estivesse disponível. Os seguintes modelos foram obtidos:

$$\begin{aligned} y_{\Lambda_{ERR}}(k) &= 0{,}4923y(k-1) - 1{,}0972y(k-2) + 0{,}0687u(k-2) \\ &\quad -0{,}0355u(k-2)^2 + 0{,}3573u(k-1)y(k-1) \\ &\quad -0{,}3546u(k-1)y(k-2) + 4{,}8465 + 0{,}0636y(k-2)^2 \\ &\quad +0{,}0571y(k-2)y(k-1) + 0{,}0161u(k-2)y(k-2), \end{aligned} \tag{5.36}$$

$$\begin{aligned} y_{\Lambda_1}(k) &= 0{,}4810y(k-1) - 1{,}0604y(k-2) - 0{,}0806u(k-1) \\ &\quad +0{,}0340u(k-2)u(k-1) + 4{,}7234 + 0{,}1900u(k-2) \\ &\quad -0{,}0461u(k-2)^2 + 0{,}0027y(k-1)^2 \\ &\quad +0{,}1162y(k-2)y(k-1) + 0{,}0021y(k-2)^2, \end{aligned} \tag{5.37}$$

$$
\begin{aligned}
y_{\Lambda_2}(k) &= 2{,}6294y(k-1) - 1{,}6238y(k-2) - 0{,}0005u(k-2) \\
&\quad -0{,}0099u(k-2)^2 + 0{,}0045u(k-2)u(k-1) \\
&\quad +0{,}0082y(k-2)^2 - 0{,}1704y(k-1)^2 + 0{,}1601y(k-2)y(k-1) \\
&\quad -0{,}0266u(k-2)y(k-2) + 0{,}0346u(k-2)y(k-1),
\end{aligned} \tag{5.38}
$$

$$
\begin{aligned}
y_{\Lambda_3}(k) &= 0{,}1784y(k-1) - 0{,}7791y(k-2) - 0{,}0398u(k-2) \\
&\quad +4{,}8431 + 0{,}0547y(k-2)^2 + 0{,}0327y(k-2)y(k-1) \\
&\quad +0{,}0329y(k-1)^2 - 0{,}3422u(k-2)y(k-2) \\
&\quad -0{,}0228u(k-1) + 0{,}3576u(k-2)y(k-1),
\end{aligned} \tag{5.39}
$$

$$
\begin{aligned}
y_{\Lambda_4}(k) &= 0{,}9173y(k-1) - 0{,}0629y(k-2) + 0{,}0983u(k-2) \\
&\quad 0{,}0029u(k-2)^2 + 0{,}2155u(k-1)y(k-1) \\
&\quad -0{,}1848u(k-1)y(k-2) - 0{,}0242u(k-2)u(k-1) \\
&\quad +0{,}0422u(k-1)^2 - 0{,}3638u(k-1) + 0{,}8646.
\end{aligned} \tag{5.40}
$$

A Figura (5.44) apresenta a dinâmica de simulação livre dos modelos Λ_2 e Λ_4. Estes modelos foram os únicos que apresentaram estabilidade. A dinâmica do modelo dominante Λ_2 é mostrada na Figura (5.45). O modelo obtido pela ERR também apresentou instabilidade dinâmica. Isso ratifica a importância da expansão da ERR para a MERR, uma vez que podem haver informações importantes no processo de modelagem não presentes nos dados dinâmicos.

O termo constante provavelmente consiste de um termo espúrio, pois não se encontra no modelo dominante e se encontra presente em todos os modelos dominados, estáveis ou não. O único modelo estável obtido que apresenta o termo constante tem o parâmetro a ele associado bastante reduzido, quando comparado aos demais. Isso faz com que a sua contribuição seja pequena na composição final do modelo. Ademais, ressalta-se que, ainda neste modelo, o termo constante apresentou parâmetro de menor importância, com valor consideravelmente reduzido.

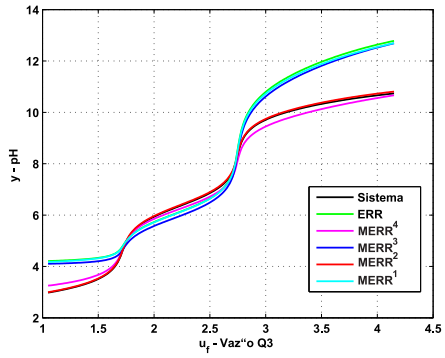

Figura 5.43: Curva estática da planta de neutralização de pH e dos modelos obtidos. A representação estática do modelo é aprimorada por meio da incorporação de informação estática.

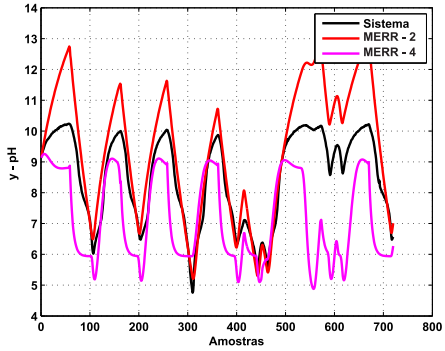

Figura 5.44: Simulação livre dos modelos. Qualitativamente, observa-se um bom desempenho particular do modelo Λ_2.

A Figura (5.43) mostra o comportamento estático dos modelos obtidos, bem como do sistema. Observa-se claramente um grande aumento na faixa de precisão do modelo, de acordo com o incremento da contribuição estática durante o processo de detecção de estruturas. Os modelos Λ_2 e Λ_4 praticamente sobrepõem a curva estática apresentada pelo sistema.

Com o propósito de validar o comportamento dinâmico dos modelos, realizou-se a predição livre dos mesmos (Figura (5.44)). Qualitativamente, observa-se que o modelo Λ_2 consegue seguir tendências não-lineares e oscilações presentes no sistema. Seu comportamento dinâmico é comparado ao sistema na Figura (5.45). Observa-se, também qualitativamente por meio da Figura (5.43), que este foi o modelo mais fiel ao comportamento estático do sistema.

Visando quantificar os resultados estáticos e dinâmicos obtidos pela técnica por meio dos diversos modelos, calculou-se o índice RMSE estático e dinâmico, bem como o erro quadrático de predição, estático e dinâmico. Tais índices são representados graficamente nas Figuras (5.46) e (5.47), e por meio da Tabela (5.8), no intuito de facilitar a análise dos resultados.

O modelo Λ_2 é notavelmente superior aos demais, tanto no comportamento estático, onde apresenta erro desprezível, quando no comportamento

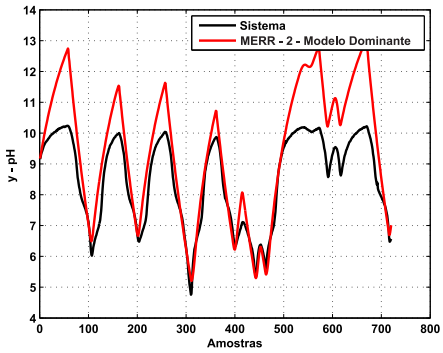

Figura 5.45: Simulação livre do modelo dominante (Λ_2).

Tabela 5.8: Índices calculados para validação dos modelos da planta de neutralização de pH.

Modelo	$RMSE_S$	$RMSE_D$	J_s	J_d
Λ_{ERR}	0,4167	∞	–	–
Λ_1	0,3972	∞	–	–
Λ_2	0,0147	0,8332	0	0
Λ_3	0,3827	∞	–	–
Λ_4	0,0648	1,2955	1	1

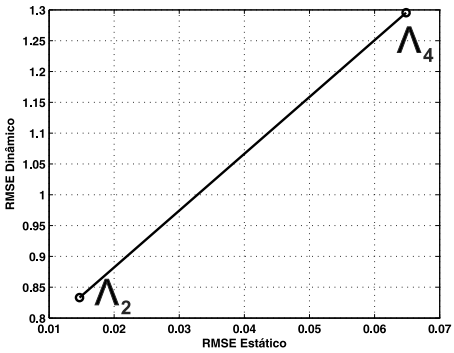

Figura 5.46: Espaço dos RMSE - planta de neutralização de pH

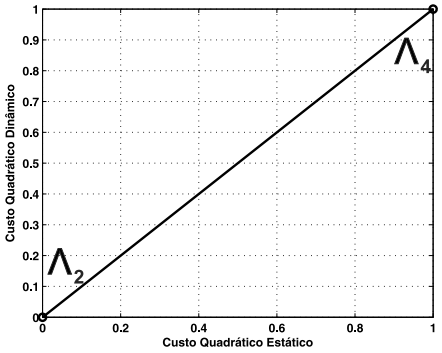

Figura 5.47: Espaço dos Erros Qudráticos - planta de neutralização de pH.

dinâmico, apresentando erro quadrático, em média, inferior ao erro da média da série temporal. O modelo Λ_4 também apresenta uma excelente qualidade estática. Contudo, seu comportamento dinâmico é prejudicado, provavelmente pelo termo constante, não-genuíno.

Outro fator relevante é que em todos os modelos obtidos, a taxa de redução de erro (mono ou multi-objetivo) do regressor $y(k-1)$ foi de mais de 99%. O parâmetro associado à este regressor no modelo Λ_2 é maior que os demais, aumentando assim a sua contribuição nos regimes dinâmicos e estáticos do modelo.

Ressalta-se que a ERR foi incapaz de obter um modelo dinamicamente estável do sistema a ser modelado, além de apresentar um enorme erro estático. Fica evidente, mais uma vez, a importância da expansão da ERR feita neste trabalho (MERR), em que são consideradas outras informações do sistema, além da importância de incorporação de outros tipos de informação na detecção de estruturas de modelos.

As Tabelas (5.9) e (5.10) apresentam a taxa de redução de erro multi-objetivo dos modelos obtidos para a representação da planta de neutralização de pH.

Como grande resultado desta seção, observa-se a grande capacidade de se obter um modelo do sistema, quando somente uma pequena parte dos dados dinâmicos se encontram disponíveis. Destaca-se que isso não foi possível utilizando somente dados dinâmicos. Ressalta-se também o grande aprimoramento estático do modelo do sistema, oriundo da incorporação deste tipo de informação.

Mais uma vez, não se pode ter certeza que os regressores obtidos são exatamente os mesmo do sistema, pois o modelo *real* do sistema encontra-se indisponível. Contudo, pode-se afirmar que o acréscimo da informação da curva estática fez com que as representações estática e dinâmica do modelo fossem aprimoradas.

O modelo Λ_2 suplantou todos os outros modelos obtidos, apresentando comportamentos estático e dinâmico satisfatórios, ou seja, apresentando erro quadrático, em média, menor que o erro cometido pela média das séries temporais. Ressalta-se que a alta não-linearidade estática presente no sistema foi devidamente modelada pela técnica MERR.

Tabela 5.9: Taxa de Redução de Erro Multi-objetivo - Planta de Neutralização de pH.

Λ_{ERR}	Regressor	ERR (%)
	y(k-1)	99,9429277
	y(k-2)	0,0333347
	u(k-2)	0,0049005
	$u(k-2)^2$	0,0000604
	u(k-1)y(k-1)	0,0000319
	u(k-1)y(k-2)	0,0009509
	C	0,0000280
	$y(k-2)^2$	0,0002091
	y(k-2)y(k-1)	0,0016860
	u(k-2)y(k-2)	0,0003360
Λ_1	Regressor	MERR (%)
	y(k-1)	99,9973033
	y(k-2)	0,0024497
	u(k-1)	0,0000105
	u(k-2)u(k-1)	0,0000071
	C	0,0000020
	u(k-2)	0,0000003
	$u(k-2)^2$	0,0000004
	$y(k-1)^2$	0,0000003
	y(k-2)y(k-1)	0,0000020
	$y(k-2)^2$	0,0000185
Λ_2	Regressor	MERR (%)
	y(k-1)	99,9931493
	y(k-2)	0,0045640
	u(k-2)	0,0003047
	$u(k-2)^2$	0,0000441
	u(k-2)u(k-1)	0,0000012
	$y(k-2)^2$	0,0000007
	$y(k-1)^2$	0,0000868
	y(k-2)y(k-1)	0,0002956
	u(k-2)y(k-2)	0,0000069
	u(k-2)y(k-1)	0,0000172

Tabela 5.10: Taxa de Redução de Erro Multi-objetivo - planta de neutralização de pH (continuação).

Λ_3	Regressor	MERR (%)
	y(k-1)	99,9742454
	y(k-2)	0,0153336
	u(k-2)	0,0021457
	C	0,0000193
	$y(k-2)^2$	0,0000412
	y(k-2)y(k-1)	0,0002623
	$y(k-1)^2$	0,0009446
	u(k-2)y(k-2)	0,0001375
	u(k-1)	0,0000545
	u(k-2)y(k-1)	0,0000433
Λ_4	Regressor	MERR (%)
	y(k-1)	99,9549976
	y(k-2)	0,0264050
	u(k-2)	0,0038710
	$u(k-2)^2$	0,0000361
	u(k-1)y(k-1)	0,0000325
	u(k-1)y(k-2)	0,0007904
	u(k-2)u(k-1)	0,0000145
	$u(k-1)^2$	0,0000405
	u(k-1)	0,0000152
	C	0,0000164

5.6 Análise da Técnica apresentada - MERR

Após ser apresentada a técnica de taxa de redução de erro multi-objetivo, várias análises podem ser feitas. A começar, uma vez que se trata de um problema multi-objetivo onde os objetivos foram combinados de forma ponderada, pode-se retornar exatamente à taxa de redução de erro. Para tal, basta que o peso referente aos dados dinâmicos seja de valor unitário, e os demais pesos nulos. A ERR é amplamente conhecida, estudada e validada na área de Identificação de Sistemas.

Por outro lado, variando os pesos associados às informações, podem ser obtidos modelos pertinentes ao conjunto de soluções eficientes, que incorporam diversas características do sistema. Este conjunto de soluções eficientes também é conhecido de Pareto-ótimo, onde estão contidas todas as soluções não suplantadas por outras do problema de otimização multi-objetivo. Ressalta-se também que a diferente combinação dos pesos pode gerar soluções não eficientes, suplantadas por outras. Uma vez que estas não representam fielmente o sistema, devem ser descartadas e não fazem parte do conjunto Pareto-ótimo.

A multi-objetividade da técnica apresentada é importante, pois em certos casos somente a utilização de dados dinâmicos não é suficiente para a obtenção da estrutura adequada de um modelo. Ademais, alguns sistemas apresentam limitações na coleta de dados dinâmicos ou ruído nos mesmos, problemas que podem ser minimizados com o desenvolvimento desta técnica que torna o modelo mais global.

A partir daí surge um novo problema, conhecido como problema da decisão, que consiste em escolher um modelo, dentre os apresentados pela taxa de redução de erro multi-objetivo para representar o sistema em questão. Na literatura existem diversas técnicas para o critério de decisão. Nepomuceno (2002) apresenta uma técnica de decisão baseada na mínima norma dos objetivos normalizados. Dessa forma, garante-se um compromisso mútuo entre os objetivos considerados no problema. Já Barroso (2006) apresenta o decisor de correlação, que correlaciona o modelo obtido com os dados de validação. Vale observar que apesar de várias técnicas, a escolha do modelo a ser utilizado pode ser objetivada pela aplicação em questão.

Do ponto de vista computacional, a técnica apresenta um baixo custo. Poucos recursos computacionais são necessários para sua implementação. Ressalta-se o caráter multi-objetivo inovador da técnica apresentada, uma vez

que as principais técnicas de detecção de estruturas de modelos NARMAX polinomiais presentes na literatura, são de caráter mono-objetivo. As poucas técnicas multi-objetivos não se mostram capazes de quantificar a contribuição de cada regressor ao explicar diferentes características do sistema.

A MERR mostrou-se eficaz na identificação de todos os estudos de caso em questão. Nos dois casos simulados, a técnica foi capaz de quantificar regressores genuínos como aqueles que apresentavam maior contribuição na representação do sistema. Ressalta-se que, para os sistemas simulados, a contribuição dos parâmetros associados aos termos espúrios (incluído nos devido à sugestão do AIC e portanto, espúrios) é menor nos modelos MERR, quando comparados à ERR.

Outro ponto a se destacar é que no caso simulado 1, o modelo Λ_4 ($\omega_d = 0{,}9$ e $\omega_s = 0{,}1$) classificou os regressores genuínos da mesma forma que a ERR, confirmando o que foi provado matematicamente quando apresentada a metodologia, ou seja, a técnica tende à ERR quando o peso dinâmico tende à unidade.

No caso do conversor buck, pôde-se obter modelos consideravelmente mais globais do ponto de vista estático, perante uma pequena perda de qualidade dinâmica. O caráter multi-objetivo permite a obtenção de várias soluções eficientes, sendo este um outro aspecto positivo da MERR.

O exemplo utilizando a planta de neutralização de pH simulou um processo em que uma pequena massa de dados estivesse disponível, situação em que dados teóricos estáticos poderiam ser utilizados. Neste problema de identificação a ERR não pode identificar corretamente o sistema, pois apresentou um modelo dinamicamente instável e com um grande erro estático. Já a MERR foi capaz de obter um modelo representativo do sistema dos pontos de vista estático e dinâmico. Deste modo, observa-se que, havendo informações estáticas disponíveis, as mesmas devem ser utilizadas para obter um bom modelo do sistema.

5.7 Conclusões do Capítulo

O presente Capítulo mostrou os resultados oriundos da aplicação da metodologia aqui desenvolvida em dois sistemas reais e dois sistemas simulados. A metodologia consiste na taxa de redução de erro multi-objetivo (MERR), em que informações acerca do sistema são utilizadas na etapa de detecção de estruturas de modelos, quantificando a contribuição de cada regressor.

Como estudo de casos, buscou-se uma maior variedade de classes de sistemas possível. Os sistemas simulados foram estudados nos trabalhos de (Piroddi e Spinelli, 2003) e (Bonin et al., 2010), e tratam-se de modelos NARX polinomiais acrescidos de ruídos, quando excitados por uma fonte de ruído branco gaussiano (sistema simulado 1 - (Piroddi e Spinelli, 2003)) e por uma combinação linear de uma entrada auto-regressiva e de uma fonte de ruído branco gaussiano (sistema simulado 2 - (Bonin et al., 2010)).

Como sistemas reais, utilizou-se um conversor CC-CC buck (Aguirre et al., 2000) e uma planta de neutralização de pH, inicialmente estudada em (Campos, 2007). O conversor CC-CC buck possui como principais características seu comportamento dinâmico não-linear e seu comportamento estático afim, o que dificulta a obtenção de um bom modelo, do ponto de vista global. Já a planta de neutralização de pH apresenta alto grau de não-linearidade estática e dinâmica. Foram utilizados poucos dados dinâmicos para a identificação da planta de neutralização de pH, visando simular um processo em que poucos dados dinâmicos estivessem disponíveis.

Nos exemplos simulados, a abordagem via MERR mostrou-se eficiente, uma vez que foi capaz de quantificar os regressores genuínos como mais importantes em modelos pertinentes ao Pareto de estruturas. Ressalta-se que nos casos simulados o critério de informação de Akaike (AIC) não foi capaz de precisar o número real de regressores do sistema.

Os exemplos reais ratificaram a qualidade e aplicabilidade da técnica MERR e a necessidade de se expandir a ERR, de modo a permitir a incorporação de outras informações no modelo. No caso da identificação do conversor CC-CC buck, todos os modelos obtidos, com exceção do modelo obtido pela ERR, pertencem ao conjunto Pareto-ótimo. Ainda neste sistema, notou-se uma grande melhora na representação estática à um baixo custo dinâmico, fornecendo um modelo mais robusto e global.

Por fim, testou-se a robustez da técnica por meio da identificação da planta de neutralização de pH. Neste processo, utilizou-se uma massa de dados dinâmicos consideravelmente reduzida, em adição de informação estática. Pôde ser provada a obtenção de um modelo mais robusto e fiel ao sistema, considerando aspectos estáticos e dinâmicos.

Ressalta-se o caráter inovador da técnica, uma vez que, até onde se viu, não existem técnica puramente multi-objetivo para detecção de estruturas de modelos NARMAX polinomiais capazes de quantificar a contribuição de cada regressor na representação de diferentes informações do sistema.

Capítulo 6

Conclusões e Pesquisas Futuras

"A prova mais imediata da compatibilidade entre religião e ciência natural, mesmo sob análise detalhada e crítica, é o fato histórico de que justamente os maiores cientistas de todos os tempos, homens como Kepler, Newton, Leibniz, estavam imbuídos de profunda religiosidade."

Max Planck

Este Capítulo apresenta as conclusões obtidas por meio da aplicação da técnica desenvolvida neste trabalho, elaborando uma análise crítica dos resultados. As principais contribuições e propostas de trabalhos futuros também são apresentadas, geradas por meio de questões investigativas que surgiram durante o desenvolvimento deste trabalho.

6.1 Discussões Finais

O presente trabalho apresentou uma metodologia para a criação da taxa de redução de erro multi-objetivo (MERR), em que foi possibilitada a incorporação de informações estáticas do sistema na etapa de detecção de estrutura do modelo.

A princípio, foi feita uma revisão teórica das técnicas presentes na literatura. Pôde-se concluir que a grande maioria das técnicas de detecção de estruturas de modelos NARMAX polinomiais são de caráter mono-objetivo,

com exceção das que utilizam a detecção multi-objetivo de parâmetros como ferramenta auxiliar na obtenção da estrutura do modelo, sendo estas incapazes de quantificar a contribuição de cada regressor ao representar diferentes características do sistema. Conceitos básicos também foram apresentados, buscando contextualizar o leitor com o conteúdo apresentado neste trabalho.

A metodologia desenvolvida foi apresentada e aplicada em 4 sistemas, sendo 2 sistemas reais e 2 sistemas simulados. No caso dos sistemas simulados (onde a estrutura do modelo do sistema é totalmente conhecida), a técnica foi capaz de obter exatamente os regressores genuínos, que compõem o sistema provando a eficiência da técnica apresentada. Nestes casos, a ERR também foi capaz de quantificar com maior importância os regressores genuínos.

O primeiro sistema simulado consiste de um modelo NARX polinomial estudado em diversos trabalhos da área, quando no estudo de técnicas de detecção de estruturas. Neste primeiro momento, o sistema foi excitado por meio de um ruído gaussiano branco com média zero e variância unitária. O outro sistema simulado consiste do mesmo modelo NARX do primeiro sistema simulado. Contudo, possui como principal peculiaridade o fato da entrada ser composta por uma combinação linear de uma entrada auto-regressiva com um ruído branco gaussiano com média zero e variância unitária, formando um filtro passa-baixas. Destaca-se que o uso deste tipo de excitação pode confundir o algoritmo de seleção de estruturas, pois reduz a diferença entre os termos $y(t)$ e $y(t-1)$ (Bonin et al., 2010).

Como sistemas reais, utilizou-se um conversor CC-CC buck e uma planta de neutralização de pH. No caso do conversor CC-CC buck, obteve-se um Pareto-ótimo por meio da variação dos pesos ω_d e ω_S, sendo que a medida que se dava uma maior importância ao peso estático, o modelo se tornava mais global, à custo de uma pequena perda de sua qualidade dinâmica. O conversor utilizado, apesar de apresentar alta não-linearidade dinâmica, possui comportamento estático afim, o que dificulta sua modelagem precisa de modo mais global, ou seja, envolvendo as diversas características do sistema.

Já na planta de neutralização de pH, utilizou-se uma pequena massa de dados dinâmicos, acrescidos de informações contidas em sua curva estática. Deste modo pôde-se verificar que há informações presentes na estática do sistema que influenciam seu comportamento dinâmico. Isso pode ser justificado, uma vez que o modelo obtido por meio da ERR não apresentou estabilidade dinâmica. Outro ponto que deve ser destacado foi a alta capacidade do modelo em representar a característica estática do sistema, que é

altamente não-linear.

Por fim, ressalta-se que buscou-se a utilização de sistemas gerais, que apresentassem diversas peculiaridades inerentes de sistemas reais e simulados, de modo a testar de forma precisa e eficiente a nova abordagem desenvolvida e apresentada neste trabalho. Trabalhos similares na literatura, como (Margoti, 2011; Bonin et al., 2010; Barbosa, 2010; Barroso et al., 2007) e outros, utilizaram destes mesmos sistemas para validar as técnicas ali desenvolvidas.

6.2 Contribuições deste trabalho

Abaixo serão enumeradas as principais contribuições deste trabalho:

1. Elaboração de uma técnica puramente multi-objetivo para determinação de estruturas de modelos NARMAX polinomiais, capaz de quantificar a contribuição de cada regressor ao identificar diferentes peculiaridades de um sistema.

   Foi desenvolvida a MERR (taxa de redução de erro multi-objetivo, do inglês *Multi-objective Error Reduction Ratio*), permitindo a incorporação de diferentes características que se têm do sistema no modelo. Tal técnica é uma extensão da conhecida ERR (taxa de redução de erro, do inglês *Error Reduction Ratio*), que considera somente informações acerca da dinâmica do sistema para a obtenção dos regressores do modelo. Ressalta-se o caráter inovador da técnica, uma vez que até onde se pesquisou não se encontrou técnicas puramente multi-objetivo aplicadas à determinação de estruturas de modelos NARMAX capazes de quantificar a contribuição dos regressores.

2. Descrição de como incorporar informações estáticas no modelo do sistema.

   Foi descrita de forma sistemática, matemática e computacional como incorporar pontos fixos e curva estática de um modelo, durante a etapa de obtenção da estrutura do modelo.

3. Elucidação de conceitos básicos de Identificação de Sistemas.

   Foram apresentados de forma clara conceitos acerca de duas das principais etapas da Identificação de Sistemas (determinação da estrutura

e dos parâmetros). Diferentes técnicas foram apresentadas de forma clara, permitindo um maior aprendizado do leitor.

4. Revisão de algumas das principais técnicas de detecção de estruturas de modelos NARMAX polinomiais.

   Foram apresentadas algumas das técnicas de detecção de estruturas desenvolvidas nos últimos anos, por grandes pesquisadores da área. Ademais, uma análise crítica acerca dessas técnicas também foi feita.

5. Expansão do Algoritmo de Golub-HouseHolder

   Foi apresentada uma expansão do algoritmo de Golub-HouseHolder visando sua aplicação na MERR.

6.3 Pesquisas Futuras

Durante o desenvolvimento do trabalho, surgiram questões ainda não respondidas, tampouco investigadas. Dentre elas, serão destacadas algumas:

1. Considerando a identificação dos sistemas simulados 1 e 2, o critério de informação de Akaike (AIC) não foi capaz de estimar o número real de regressores genuínos. O AIC é uma técnica inerentemente mono-objetivo, baseada exclusivamente em dados dinâmicos. Sugere-se uma readaptação neste índice, a fim de incorporar outras características do sistema em questão.

2. Foi apresentado um meio de incorporação da característica estática do sistema na estrutura do modelo. Como trabalhos futuros, pretende-se estudar situações em que um número maior de informações esteja disponível para a identificação do sistema em questão e a utilização destas informações na detecção de estruturas de modelos NARMAX.

3. A técnica apresentada gera um conjunto de soluções eficientes. Desse modo, pretende-se verificar a aplicabilidade dos critérios decisores já presentes na literatura, como o decisor de mínima correlação (Barroso, 2006) e a norma mínima dos objetivos (Nepomuceno et al., 2007). Ademais, pretende-se criar outros mecanismos decisores.

4. Neste trabalho, foi utilizado a combinação convexa dos objetivos em questão. Como utilizar outros mecanismos de otimização multi-objetivo, especificamente no caso da MERR?

5. Uma vez realizada a etapa anterior, quais as principais vantagens e desvantagens de cada abordagem? Quando devem ser utilizadas?

6. Qual a melhor forma de variar os valores dos pesos, de modo a obter uma curva de Pareto uniformemente povoada?

7. Como a técnica se comporta perante a incorporação de outras características do sistema?

8. Como incorporar características de estabilidade durante a obtenção de estruturas do modelo?

Referências Bibliográficas

Aguirre, L. e Billings, S. (1995a). Improved structure selection for nonlinear models based on term clustering. *International Journal of Control*, 62(3):569–587.

Aguirre, L. A. (1997). On the structure of nonlinear polynomial models: higher order correlation functions, spectra, and term clusters. *IEEE Transactions on Circuits and Systems Part I: Fundamental Theory and Applications*, 44(5):450–453.

Aguirre, L. A. (2007). *Introdução à Identificação de Sistemas: Técnicas lineares e não-lineares aplicadas a sistemas reais*. Editora da UFMG. 3ª edição.

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., Donoso-Garcia, P. F., e Santos, 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. e Jácome, C. R. F. (1998). Cluster analysis of narmax models for signal-dependent systems. *IEE Proceedings-D Control Theory and Applications*, 145(4):409–414.

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. *SBA Controle & Automação*, 9(2):90–106.

Algreer, M., Armstrong, M., e Giaouris, D. (2011). Adaptive PD plus I control of a switch-mode DC-DC power converter using a recursive FIR predictor. *IEEE Transactions on Iindustry Applications*, 47(5):2135–2144.

Alves, M. A. (2009). *Utilização de auto-consistência como ferramenta auxiliar na seleção de estrutura de modelos NARX polinomiais*. Dissertação de Mestrado, Universidade Federal de Minas Gerais.

Barbosa, A. M. (2010). *Técnicas de otimização bi-objetivo para a determinação da estrutura de modelos NARX*. Dissertação de Mestrado, Universidade Federal de Minas Gerais.

Barroso, M. F. S. (2006). *Otimização bi-objetivo aplicada à estimação de parâmetros de modelos não-Lineares: caracterização e tomada de decisão*. Tese de Doutorado, Universidade Federal de Minas Gerais.

Barroso, M. F. S., Takahashi, R., e Aguirre, L. (2007). Multi-objective parameter estimation via minimal correlation criterion. *Journal of Process Control*, 17(4):321–332.

Billings, S. A. (1980). Identification of nonlinear systems - a survey. *IEE Proceedings-D Control Theory and Applications*, 127(6):272–285.

Bograd, S., Reuss, P., Schmidt, A., Gaul, L., e Mayer, M. (2011). Modeling the dynamics of mechanical joints. *Mechanical Systems and Signal Processing*, 25(8):2801–2826.

Bonin, M., Seghezza, V., e Piroddi, L. (2010). NARX model selection based on simulation error minimisation and LASSO. *IET Control Theory and Applications*, 4(7):1157–1168.

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*. Dissertação de Mestrado, UNILESTE.

Chen, Q., Worden, K., Peng, P., e Leung, A. (2007). Genetic algorithm with an improved fitness function for (n)arx modelling. *Mechanical Systems and Signal Processing*, 21(2):994 – 1007.

Ding, F., Liu, X. P., e Liu, G. (2011). Identification methods for hammerstein nonlinear systems. *Digital Signal Processing*, 21(2):215–238.

Ding, N. e Taheri, S. (2010). Application of Recursive least square algorithm on estimation of vehicle sideslip angle and road friction. *Mathematical Problems in Engineering*, páginas 1–18.

Fellmann, C., Zuber, J., McJunkin, K., Chang, K., Malone, C. D., Dickins, R. A., Xu, Q., Hengartner, M. O., Elledge, S. J., Hannon, G. J., e Lowe, S. W. (2011). Functional identification of optimized RNAi triggers using a massively parallel sensor assay. *Molecular Cell*, 41(6):733–746.

Fung, R.-F., Hsu, Y.-L., e Huang, M.-S. (2009). System identification of a dual-stage xy precision positioning table. *Precision Engineering*, 33(1):71 – 80.

Furtado, E. C., Mendes, E. M. A. M., Nepomuceno, E. G., e Silva, V. V. R. E. (2002). Identificação de sistemas dinâmicos não-lineares contínuos utilizando modelos NARMAX: estudo de caso de um forno a arco elétrico. In *XIV Congresso Brasileiro de Automática*, páginas 2150–2155, Natal, Brasil.

Gomes, T. V., Tavares, S. E., Martins, S. A. M., Kurchbart, S. M., e Nepomuceno, E. G. (2008). Simulation of complex systems by means of scicos. *II Workshop on Simulation and Analysis of Complex Systems*.

Gunnar, J., Wernholt, E., Hovland, G., e Brogardh, T. (2006). Nonlinear grey-box identification of linear actuators containing hysteresis. In *IEEE International Conference on Robotics and Automation. ICRA 2006*, páginas 1818 –1823.

Hong, X., Mitchell, R. J., Chen, S., Harris, C. J., Li, K., e Irwin, G. W. (2008). Model selection approaches for non-linear system identification: a review. *International Journal of Systems Science*, 39(10):925–946.

Hu, H., Dong, J., Zhang, J., Cheng, Y., e Xu, T. (2011). Identification of gas/solid two-phase flow regimes using electrostatic sensors and neural-network techniques. *Flow Measurement and Instrumentation*, (0):–.

Johansen, T. A. (1996). Identification of non-linear systems using empirical data and prior knowledge - an optimization approach. *Automatica*, 32(3):337–356.

Johansen, T. A. (2000). Multi-objective identification of FIR models. In *Proceedings of 12th IFAC Symposium on System Identification 2000*, Santa Barbara, USA.

Johansen, T. A. e Babuska, R. (2003). Multi-objective identification of Takagi-Sugeno fuzzy models. *IEEE Transactions on Fuzzy Systems*, 11(6):847–860.

Kobayashi, Y. e Okita, T. (2002). Identification of nonlinear systems with hysteresis characteristics. In *SICE 2002. Proceedings of the 41st SICE Annual Conference*, volume 3, páginas 1577 – 1581 vol.3.

Koop, R. E. e Orford, R. J. (1963). Linear regression applied to system identification for adaptive control systems. *AIAA Journal*, 1(10):2300–2306.

Korenberg, M., Billings, S., Liu, Y., e Mcilroy, P. (1988). Orthogonal parameter estimation algorithm for non-linear stochastic systems. *International Journal of Control*, 48(1):193–210.

Lee, D. J. e Park, Y. S. (2011). Sliding-mode-based parameter identification with application to tire pressure and tire-road friction. *International Journal of Automotive Technology*, 12(4):571–577.

Leontaritis, I. J. e Billings, S. A. (1985a). Input-output parametric models for non–linear systems - part i: deterministic non–linear systems. *International Journal of Control*, 41(2):303–328.

Leontaritis, I. J. e Billings, S. A. (1985b). Input-output parametric models for non-linear systems - part ii: sthocastic non-linear systems. *International Journal of Control*, 41(2):329–344.

Lichtenberger, W. (1961). Technique of linear system identification using correlating filters. *Ire Transactions On Automatic Control*, AC 6(2):183–&.

Lima, C. B. S. (2009). *Uso de informação a priori para identificação de um processo de neutralização de pH em malha fechada*. Dissertação de Mestrado, Universidade Federal de Minas Gerais.

Lind, I. e Ljung, L. (2005). Regressor selection with the analysis of variance method. *Automatica*, 41(4):693 – 700.

Lind, I. e Ljung, L. (2008). Regressor and structure selection in narx models using a structured anova approach. *Automatica*, 44(2):383 – 395.

Ljung, L. (1987). *System Identification: Theory for the User*. Prentice-Hall, London.

Margoti, L. M. (2011). *Determinação de estruturas de modelos neurais*. Dissertação de Mestrado, Universidade Federal de São João del-Rei.

Martins, S. A. M., Braga, D. C. S., Nepomuceno, E. G., Gomes, T. V., e Reis, M. L. F. (2009). Investigation of the static curve information for multiobjective system identification. *Journal of Computational Interdisciplinary Sciences*, 1(2):149–157.

Martins, S. A. M., Nepomuceno, E. G., e Figueiredo, J. P. (2011). Detecção de estruturas de modelos narmax polinomiais: uma abordagem inteligente multi-objetivo. In *Anais do X Simpósio Brasileiro de Automação Inteligente*, páginas 320–325.

Martins, S. A. M., Reis, M. L. F., Braga, D. C. S., Gomes, T. V., e Nepomuceno, E. G. (2008). Multiobjective modeling of a buck converter. *II Workshop on Simulation and Analysis of Complex Systems*.

Marton, L. e Lantos, B. (2011). Control of robotic systems with unknown friction and payload. *IEEE Transactions on Control Systems Technology*, 19(6):1534–1539.

Mendes, E. e Billings, S. (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.

Nakamura, T., Judd, K., Mees, A. I., e Small, M. (2006). A comparative study of information criteria for model selection. *International Journal of Bifurcations and Chaos*, 16(8):2153–2175.

Nepomuceno, E. G. (2002). *Identificação multiobjetivo de sistemas não-lineares*. Dissertação de Mestrado, PPGEE, Universidade Federal de Minas Gerais, Belo Horizonte, Brasil.

Nepomuceno, E. G., Takahashi, R. H. C., e Aguirre, L. A. (2007). Multiobjective parameter estimation for non-linear systems: affine information and least-squares formulation. *International Journal of Control*, 80(6):863–871.

Neves, L. D., Agostinho, N. U., e Caetano, S. M. (2007). Inflation forecast by means of narmax models. In *Anais do VII Encontro Brasileiro de Finanças*, páginas 1–9, São Paulo, SP.

Oliveira, E., Martins, S. A. M., Barbosa, A., e Nepomuceno, E. G. (2009). Combate de epidemias utilizando controle em malha fechada por meio do método p-narmax. In *Anais do IX Simpósio Brasileiro de Automação Inteligente*.

Pardalos, P. M., Steponavice, I., e Zilinskas, A. (2011). Pareto set approximation by the method of adjustable weights and successive lexicographic goal programming. *Optimization Letters*, páginas 1–14.

Piroddi, L. e Spinelli, W. (2003). An identification algorithm for polynomial NARX models based on simulation error minimization. *International Journal of Control*, 76(17):1767–1781.

Pulecchi, T. e Piroddi, L. (2007). A cluster selection approach to polynomial narx identification. In *American Control Conference, 2007. ACC '07*, páginas 852 –857.

Rubio, J. d. J., Torres, C., Rivera, R., e Hernandez, C. A. (2011). Comparison of four mathematical models for braking of a motorcycle. *IEEE Latin America Transactions*, 9(5):630–637.

Shannon, C. (1949). Communication in the presence of noise. *Porceedings of the Institute of Radio Engineers*, 37(1):10–21.

Takahashi, R. H. C. (1998). *Controle singular de sistemas incertos*. Tese de Doutorado, Faculdade de Engenharia Elétrica e Computação, Universidade de Campinas.

Tan, X. e Baras, J. S. (2004). Modeling and control of hysteresis in magnetostrictive actuators. *Automatica*, 40(9):1469 – 1480.

Valackiene, A. (2011). Theoretical substation of the model for crisis management in organization. *Engineering Economics*, 22(1):78–90.

Vance, R. R., Steele, M. A., e Forrester, G. E. (2010). Using an individual-based model to quantify scale transition in demographic rate functions: Deaths in a coral reef fish. *Ecological Modelling*, 221:1907 – 1921.

Wang, D. e Ding, F. (2011). Least squares based and gradient based iterative identification for wiener nonlinear systems. *Signal Processing*, 91(5):1182–1189.

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 Modelling Identification and Control*, 3(4):341–356.

Wickwire, K. (1977). Mathematical-models for control of pest and infectious-diseases – survey. *Theoretical Population Biology*, 11(2):182–238.

Worden, K., Wong, C., Parlitz, U., Hornstein, A., Engster, D., Tjahjowidodo, T., Al-Bender, F., Rizos, D., e Fassois, S. (2007). Identification of pre-sliding and sliding friction dynamics: Grey-box and black-box models. *Mechanical Systems and Signal Processing*, 21(1):514 – 534.

Yang, X., Zhang, H., e Ma, X. (2009). Modeling and stability analysis of cascade buck converters with n power stages. *Mathematics and Computers in Simulation*, 80(3):533 – 546.

Zhang, Y., Chen, Y., e Chang, Y.-J. (2011). Estimating biological reference points using individual-based per-recruit models for the gulf of maine american lobster, homarus americanus, fishery. *Fisheries Research*, 108(2-3):385 – 392.