

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL  
ESCOLA DE ENGENHARIA  
DEPARTAMENTO DE ENGENHARIA QUÍMICA  
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA QUÍMICA

**Aplicação do CEKF na Estimação de  
Parâmetros em Modelo Dinâmico**

DISSERTAÇÃO DE MESTRADO

Cristiane Martins Schweinberger

**Porto Alegre**

**2009**

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL  
ESCOLA DE ENGENHARIA  
DEPARTAMENTO DE ENGENHARIA QUÍMICA  
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA QUÍMICA

## **Aplicação do CEKF na Estimação de Parâmetros em Modelo Dinâmico**

Cristiane Martins Schweinberger

Dissertação de mestrado apresentada como requisito parcial para obtenção do título de Mestre em Engenharia.

Área de concentração: Controle e Otimização de Processos.

**Orientador:**  
**Prof. Dr. Jorge Otávio Trierweiler**

**Porto Alegre**

**2009**

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL  
ESCOLA DE ENGENHARIA  
DEPARTAMENTO DE ENGENHARIA QUÍMICA  
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA QUÍMICA

A Comissão Examinadora, abaixo assinada, aprova a Dissertação *Aplicação do CEKF na estimação de Parâmetros em modelo dinâmico*, elaborada por Cristiane Martins Schweinberger, como requisito parcial para obtenção do Grau de Mestre em Engenharia.

Comissão Examinadora:

---

Prof. Dr. Marcelo Farenzena

---

Prof. Dr. Nilo Sérgio Medeiros Cardozo

---

Prof. Dr. Rafael Pelegrini Soares

---

Ms. Nina Paula Gonçalves Salau

# Agradecimentos

Agradeço a Deus, que me deu:

- A vida;
- A saúde;
- O intelecto;
- Minha família, meus Amigos, meus Mestres...
- Todas as pessoas com quem muito aprendi e ainda aprendo. Sejam aquelas que passaram brevemente em minha vida ou as que andam comigo lado a lado nesta peregrinação...
- A consciência, as oportunidades de arrependimento, as chances de mudar de direção...
- A paciência, o amor incondicional

Amamos a ciência e a sabedoria porque são ferramentas que utilizamos para

Agradeço a orientação do Jorge, que foi um orientador muito mais que técnico. Também agradeço muito ao Tiago e a Fabi, também ao Numa e a Simone, sendo apenas o Tiago meu parceiro na engenharia, mas todos são cientistas da vida e muito me contribuíram ao longo do meu mestrado. Meu esposo amado, meu pai, minha mãe e minhas irmãs, que amo muuuito!!!

*“E, como Moisés levantou a serpente no deserto, assim importa que o Filho do homem seja levantado”. Jo 3.14.*

## Resumo

Dado a larga aplicação dos modelos matemáticos, a procura pelo aprimoramento dos métodos de identificação de modelos sempre é bem vinda. Foi em cima disto que se desenvolveu o presente trabalho. Considerando a problemática de que os distúrbios desconhecidos nas entradas comprometem a qualidade da estimação dos parâmetros, se estudou a estimação combinada com o estimador de estados CEKF (Constrained Extended Kalman Filter). Isto porque se supôs que com o emprego do CEKF o efeito do distúrbio desconhecido seria reduzido. Para tanto, se tomou um caso de estudo de seis tanques esféricos interligados e foram simulados dados experimentais computacionalmente. Foi escolhido um ponto de operação para a estimação com base na estimabilidade dos parâmetros. Antes da estimação os dados experimentais foram filtrados. Para a estimação dos parâmetros foi elaborada uma metodologia iterativa, caracterizada pelos três procedimentos: (1) estimação dos parâmetros, (2) ajuste de função objetivo: possibilidade de incorporar as derivadas medidas, bem como os estados e derivadas estimados pelo CEKF e (3) melhoria dos dados com a passagem do CEKF. Em outras palavras, a metodologia desenvolvida consiste em estimações de parâmetros seqüenciais acompanhadas com a melhoria progressiva dos dados devido à aplicação do CEKF. Para o estudo de caso tomado foi perceptível a melhoria na estimação dos parâmetros, tornando interessante a continuidade do estudo que foi iniciado no presente trabalho.

## Abstract

Given the wide application of mathematical models, the demand for improvement of methods of identification of models is always welcome. It was upon this that has developed in this work. Considering the problem of the unknown disturbances in the inputs affect the quality of the estimation of parameters, we studied the estimation combined with the estimator states CEKF (Extended Constrained Filter Kalmar). This is because it was assumed that with the use of CEKF the effect of unknown disturbance would be reduced. As such, it took a case study of six spherical tanks interconnected to form simulated experimental data computationally. Was chosen an operating point for the estimation based on the estimated parameters. Before estimating the experimental data were filtered. For the estimation has been elaborated iterative methodology, characterized by three procedures: (1) parameter estimation, (2) adjustment of the objective function: the possibility of incorporating the derived measures, and states and derive the estimated CEKF and (3 ) improvement of data with the passage of CEFK. In other words, the methodology consists of sequential estimation of parameters together with the gradual improvement of data due to application of CEKF. For the case study was taken a noticeable improvement in the estimation of parameters, makes it interesting to continue the study that was initiated in this work.

# Sumário

<b>Introdução .....</b>	<b>1</b>
<b>Revisão: identificação de processos .....</b>	<b>4</b>
2.1 Modelos e definições .....	4
2.2 Identificação de processos.....	5
2.2.1 Estimabilidade dos parâmetros .....	9
2.2.2 Estimação de parâmetros.....	13
2.2.3 Validação do modelo .....	18
<b>Processos com distúrbios e proposta de aplicação do CEKF .....</b>	<b>22</b>
3.1 Problema da identificação de processos com distúrbios .....	22
3.1.1 Processo estocástico .....	24
3.1.2 Processo com distúrbio determinístico e estocástico .....	26
3.1.3 Motivação para o emprego do estimador de estados .....	27
3.2 Estimadores de estados .....	29
3.3 Proposta de aplicação do CEKF e estimação dos parâmetros.....	33
<b>Metodologia.....</b>	<b>37</b>
4.1 Estimabilidade dos parâmetros .....	38
4.2 Tratamento prévio dos dados experimentais.....	40
4.3 Implementação do método de Gauss.....	40
<b>Resultados.....</b>	<b>45</b>
5.1 Descrição do estudo de caso .....	45
5.2 Estimabilidade dos parâmetros .....	46
5.3 Gerando dados experimentais para estimação .....	51
5.4 Tratamento prévio dos dados experimentais.....	52
5.5 Estimação dos parâmetros.....	56
<b>Conclusão.....</b>	<b>73</b>
<b>Referências Bibliográficas .....</b>	<b>76</b>

# Lista de Figuras

<b>Figura 1.1:</b> Forma geral de se identificar processos vinculada com o objetivo do presente trabalho. ....	2
<b>Figura 1.2:</b> Forma resumida da metodologia que se elaborou no que diz respeito a estimação dos parâmetros e aplicação do CEKF. ....	3
<b>Figura 2.1:</b> Fluxograma referente a metodologia de identificação de processos.....	8
<b>Figura 2.2:</b> Resultado para as regiões de confiança construídas com 80 e 95% de confiança. O lado esquerdo se refere as regiões propriamente ditas com 80% e o lado direito se refere a região condicional com 95%. No lado esquerdo, a linha sólida adicionada se refere a projeção do autovetor correspondente ao menor autovalor (BRUN et al., 2001). ....	11
<b>Figura 2.3:</b> Característica da estimação dos parâmetros: diversos valores possíveis para os parâmetros.....	13
<b>Figura 2.4:</b> Representação esquemática da planta estudada pelos autores: processo de aquecimento com dois tanques (MACHADO e TRIERWEILER, 2004). ....	16
<b>Figura 2.5:</b> Entradas experimentais: variável manipulada $Q$ e distúrbio $T_0$ , considerado como não medido (MACHADO e TRIERWEILER, 2004). ....	16
<b>Figura 2.6:</b> Resultados obtidos para as três formas de identificação em comparação com os dados experimentais da planta (MACHADO, TRIERWEILER, 2004)...	17
<b>Figura 3.1:</b> Esquematização da influência dos distúrbios nas entradas sobre as saídas. ....	22
<b>Figura 3.2:</b> Processo de dois tanques tomado como exemplo. ....	23
<b>Figura 3.3:</b> Entrada estocástica de um determinado processo realizado três vezes (adaptado de FIKAR; MIKLES, 2007). ....	25
<b>Figura 3.4:</b> Exemplo do processo sujeito ao distúrbio provocado pelo ambiente. ....	26
<b>Figura 3.5:</b> Visualização da temperatura ambiente formada por duas parcelas: determinística e estocástica. ....	27
<b>Figura 3.7:</b> Esquema de funcionamento de um estimador de estados. ....	30
<b>Figura 3.8:</b> Loop realizado pelo estimador de estados.....	31
<b>Figura 3.10:</b> Fluxograma elaborado para o ajuste da sintonia do CEKF.....	35
<b>Tabela 3.11:</b> Fluxograma para o ajuste da função objetivo atualizado com sua nova formulação. ....	36
<b>Figura 4.1:</b> Seqüência de passos realizados no trabalho. ....	37
<b>Figura 4.2:</b> Cálculo de $J$ , $q$ , e $N$ para o sistema dinâmico.....	43
<b>Figura 4.3:</b> Resumo de cada iteração. ....	43
<b>Figura 4.4:</b> Fluxograma da iteração. ....	44
<b>Figura 5.1:</b> Desenho esquemático da planta estudada. ....	45
<b>Figura 5.3:</b> Comparação ente as dinâmicas dos Pontos 3 e 9 para os estados $h1$ e $h2$ . ...	48
<b>Figura 5.4:</b> Comparação ente as dinâmicas dos Pontos 3 e 9 para os estados $h3$ e $h4$ . ...	48
<b>Figura 5.7:</b> Sensibilidades dinâmicas para o Ponto 8. ....	50
<b>Figura 5.9:</b> Frações $x1$ e $x2$ com distúrbios determinísticos. ....	51
<b>Figura 5.10:</b> Comparação dos estados com os distúrbios determinísticos ( $DistDest$ ) em relação ao modelo fundamental que é sem distúrbio ( $ModFund$ ). ....	51
<b>Figura 5.11:</b> Saídas $h1$ e $h2$ com todos os distúrbios atuantes: $h1_{Exp}$ e $h2_{Exp}$ . $Exp$ : dados experimentais. ....	52

<b>Figura 5.13:</b> Filtro de ordem 2 e frequência de corte 0,2: (a) estados medidos e (b) derivadas. ....	54
<b>Tabela 5.14:</b> Filtro de ordem 5 e frequência de corte 0,035: (a) estados medidos e (b) derivadas. ....	54
<b>Figura 5.15:</b> Derivadas experimentais filtradas ( <i>ExpFilt</i> ) com filtro de ordem 5 e frequência de corte 0,035. Com defasagem (a) e com defasagem corrigida (b). ....	55
<b>Figura 5.16:</b> Efeito da ordem do filtro: (a) ordem 2 e frequência de corte 0,035 e, (b) ordem 5 e frequência de corte 0,035. ....	55
<b>Figura 5.17:</b> Comparação dos resultados para diferentes frequências de corte com ordem 5: (a) 0,015, (b) 0,035, (c) 0,055 e (d) referência. ....	56
<b>Figura 5.18:</b> Comparação dos resultados apresentados na Tabela 5.6: Erro (primeiro ajuste), Derivada do erro (terceiro ajuste). <i>ModFund</i> : Referência, modelo ideal. ....	59
<b>Figura 5.19:</b> Comparação dos resultados apresentados na Tabela 5.6: Erro (primeiro ajuste), Derivada do erro (terceiro ajuste). <i>ModFund</i> : Referência, modelo ideal. ....	59
<b>Figura 5.20:</b> Comparação dos resultados apresentados na Tabela 5.6: Erro (primeiro ajuste), Derivada do erro (terceiro ajuste). <i>ModFund</i> : Referência, modelo ideal. ....	60
<b>Figura 5.21:</b> <i>ModFund</i> : Se refere aos valores verdadeiros, referência para comparação. <i>ExpFilt</i> : derivadas experimentais filtradas com a função <i>IdFilt</i> do MATLAB <sup>®</sup> , utilizadas na primeira etapa para obtenção de $\hat{\theta}_1$ . <i>CEKF<sub>2</sub></i> : derivadas melhoradas com a passagem do CEKF nesta segunda etapa. ....	62
<b>Figura 5.22:</b> Comparação entre dados utilizados para estimação dos parâmetros na segunda ( <i>CEKF<sub>2</sub></i> ) e terceira etapa ( <i>CEKF<sub>3</sub></i> ). ....	65
A primeira vista os gráficos da Figura 5.22 são intrigantes, levando-nos a questionar: “como a estimação com <i>CEKF<sub>3</sub></i> pôde ser melhor que com <i>CEKF<sub>2</sub></i> , sendo que existe um considerável erro ou discrepância inicial em <i>CEKF<sub>3</sub></i> ?”. Na verdade, tal erro é responsável pela melhoria na estimação. Constatamos que este erro conduz a um melhor balanço numérico de forma que a estimação dos parâmetros é melhorada. E, salientamos: esse balanço numérico somente foi possível com a passagem do CEKF, reforçando ainda mais a vantagem do seu uso. ....	66
<b>Figura 5.23:</b> Visualização gráfica do segundo caso da Tabela 5.15. ....	67
<b>Figura 5.24:</b> Gráficos com o erro inicial em <i>CEKF<sub>3</sub></i> removido. ....	67
<b>Figura 5.25:</b> <i>h1</i> e <i>h2</i> : Comparação do modelo obtido ( <i>Etapa 4</i> ) com o modelo ideal ( <i>ModFund</i> ) e o processo com os distúrbios determinísticos ( <i>DistDet</i> ). ....	70
<b>Figura 5.26:</b> <i>h3</i> e <i>h4</i> : Comparação do modelo obtido ( <i>Etapa 4</i> ) com o modelo ideal ( <i>ModFund</i> ) e o processo com os distúrbios determinísticos ( <i>DistDet</i> ). ....	70
<b>Figura 5.27:</b> <i>h5</i> e <i>h6</i> : Comparação do modelo obtido ( <i>Etapa 4</i> ) com o modelo ideal ( <i>ModFund</i> ) e o processo com os distúrbios determinísticos ( <i>DistDet</i> ). ....	71
<b>Figura 5.28:</b> Comparação entre os resultados obtidos nas quatro etapas. ....	71
<b>Figura 5.29:</b> Comparação entre os resultados obtidos nas quatro etapas. ....	72
<b>Tabela 5.30:</b> Comparação entre os resultados obtidos nas quatro etapas. ....	72

## Lista de Tabelas

<b>Tabela 2.1:</b> Matriz de correlação (BRUN et al., 2001).....	11
<b>Tabela 2.2:</b> Ranking da importância dos Parâmetros (BRUN et al., 2001).....	12
<b>Tabela 2.3:</b> Índice de Colinearidade dos Grupos de Parâmetros (BRUN et al., 2001)....	12
<b>Tabela 3.1:</b> Duas situações para modelar o processo de exemplo. ....	23
<b>Tabela 3.2:</b> Modelos fundamental e específico para o exemplo do trocador com distúrbio. ....	26
<b>Tabela 3.3:</b> Modelos para o processo e o sensor.....	29
<b>Tabela 5.1:</b> Valores dos Coeficientes de Descarga $CD (\theta^*)$ . ....	46
<b>Tabela 5.3:</b> Índice de importância dos parâmetros $\delta$ . ....	47
<b>Tabela 5.5:</b> Chutes iniciais. ....	57
<b>Tabela 5.6:</b> Resultados para o ajuste da função objetivo*. ....	58
<b>Tabela 5.7:</b> Comparação da estimação com outras sintonias (escolhida destacada). ....	61
<b>Tabela 5.8:</b> Resultado obtido na etapa anterior em comparação com a sintonia selecionada. ....	61
<b>Tabela 5.9:</b> Resultado da função objetivo formada apenas com a parcela relativa a $\lambda_A$ .....	62
<b>Tabela 5.10:</b> Avaliação da influência da parcela relativa a $\zeta_A$ , tendo $\lambda_A = 0$ e $\zeta_B = 0$ .....	63
<b>Tabela 5.11:</b> Avaliação da influência da parcela relativa a $\zeta_B$ , tendo $\lambda_A = 0$ e $\zeta_A = 0$ .....	63
<b>Tabela 5.12:</b> Comparação de resultados para o ajuste da matriz $Q (P_{ii} = 10^{-3}$ e $R_{ii} = 10^{-3})$ .....	64
<b>Tabela 5.13:</b> Comparação de resultados para o ajuste da matriz $R (Q_{ii} = 10^{-10}$ e $P_{ii} = 10^{-3})$ .....	64
<b>Tabela 5.14:</b> Comparação de resultados para o ajuste da matriz $P (Q_{ii} = 10^{-10}$ e $R_{ii} = 10^{-6})$ . ....	65
<b>Tabela 5.15:</b> Resultado obtido na etapa anterior em comparação com a sintonia selecionada. ....	65
<b>Tabela 5.16:</b> Dois casos tomados para comparação.....	66
<b>Tabela 5.17:</b> Estimação com a remoção do erro inicial da Figura 5.22.....	67
<b>Tabela 5.18:</b> Resultado da função objetivo formada apenas com a parcela relativa a $\lambda_A$ .....	68
<b>Tabela 5.19:</b> Avaliação da influência da parcela relativa a $\zeta_A$ , tendo $\lambda_A = 0$ e $\zeta_B = 0$ .....	68
<b>Tabela 5.20:</b> Resultado da função objetivo formada apenas com a parcela relativa a $\zeta_B$ . ....	69
<b>Tabela 5.21:</b> Comparação entre os resultados obtidos nas quatro etapas. ....	69
<b>Tabela 6.1:</b> Comparação dos índices de desempenho nas quatro etapas. ....	74

## Notação e Simbologia

$e$	Diferença entre medição com estimativa do modelo
$f$	Função da taxa dos estados
$F$	Matriz jacobiana de $f$ em relação à $x$
$g$	Função das saídas
$H$	Matriz Hessiana de $J$ em relação à $\theta$ / Matriz jacobiana de $g$ em relação à $x$
$J$	Função objetivo para a estimação dos parâmetros
$L$	Função de verossimilhança
$ne$	Número de pontos experimentais
$n\theta$	Número de parâmetros
$N$	Aproximação da hessiana pelo método de Gauss
$p$	Função de probabilidade
$P$	Matriz de covariância do estimador de estados
$q$	Gradiente
$Q$	Matriz de covariância de $w$
$R$	Matriz de covariância de $v$
$S$	Matriz de sensibilidade local das saídas
$u$	Vetor das entradas
$v$	Distúrbios de medição
$V$	Matriz de covariância
$w$	Distúrbios de processo
$x$	Vetor dos estados
$y$	Vetor das saídas
$\gamma$	Índice de colinearidade do grupo de parâmetros
$\delta$	Índice de importância dos parâmetros / Direção de busca
$\zeta_A$	Peso para a parcela referente aos estados estimados na função objetivo
$\zeta_B$	Peso para a parcela referente às derivadas dos estados estimados na função objetivo
$\theta$	Vetor dos parâmetros
$\hat{\theta}$	Estimativa dos parâmetros
$\theta_0$	Chute inicial dos parâmetros
$\lambda_A$	Peso para a parcela referente às saídas medidas da função objetivo
$\lambda_B$	Peso para a parcela referente às derivadas das saídas medidas da função objetivo
$\rho$	Tamanho do passo
$\sigma$	Desvio padrão

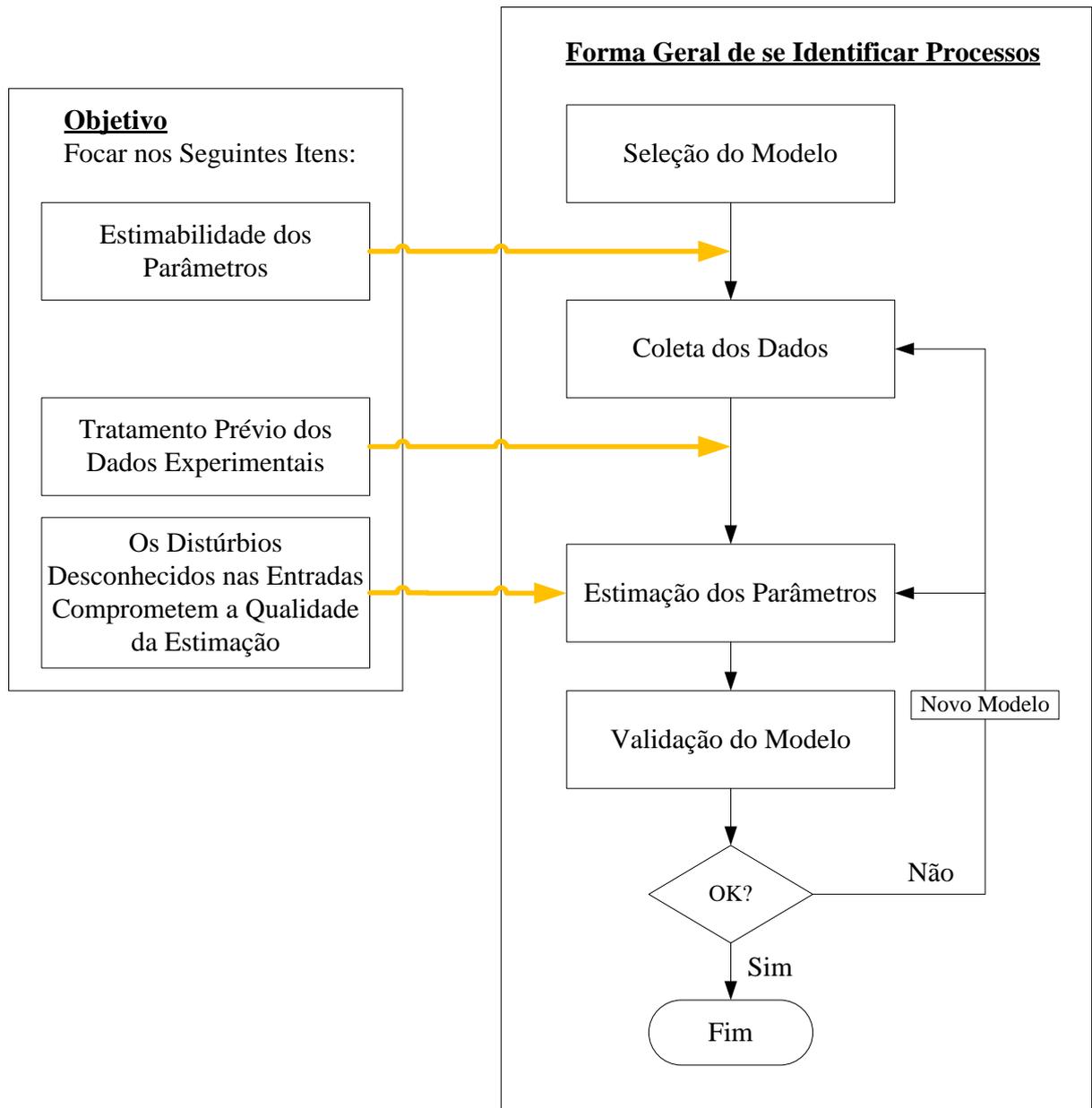
# Capítulo 1

## Introdução

No campo das engenharias é bem sabido que os modelos matemáticos constituem-se uma ferramenta largamente utilizada para diversas finalidades. Por meio dos modelos matemáticos que se realizam previsões, otimizações, projetos, etc. Sendo assim, seguramente se pode dizer que muitos avanços que demarcaram a história no sentido tecnológico e a quebra de paradigmas não teriam sido conquistados sem que antes se dispusesse de um modelo matemático, ficando claro a indiscutível importância dos modelos matemáticos. Desta forma, obter bons modelos é um alvo que sempre deve ser almejado, o que está diretamente ligado com a técnica de identificação (PINTO e SCHWAAB, 2007). Dentro deste contexto foi desenvolvido o presente trabalho, que possui como destaque a implementação da análise da estimabilidade dos parâmetros, o tratamento prévio dos dados experimentais através da filtragem e o maior foco na implementação do CEKF (Constrained Extended Kalman Filter) para a redução do efeito dos distúrbios desconhecidos nas entradas no resultado da estimação dos parâmetros. Na Figura 1.1 se apresenta a forma geral de identificar processos vinculada com o objetivo do presente trabalho.

Em especial, os distúrbios determinísticos comprometem a qualidade da estimação. Ocorrendo determinismo nos distúrbios, o ideal é que se estabeleçam estruturas para eles e sejam identificados simultaneamente com o modelo fundamental. Entretanto, na prática, por diversas razões tais como: o não acesso a medição de um distúrbio, o seu desconhecimento, sua complexidade, etc., um distúrbio determinístico pode ser parcialmente ou nem mesmo caracterizado. Para casos como este, julgou-se o emprego de um estimador de estados, como o CEKF, durante a estimação dos parâmetros ser uma interessante alternativa. Pois, o estimador de estados ao considerar a possibilidade de distúrbios no processo resulta em uma melhor estimativa para os estados.

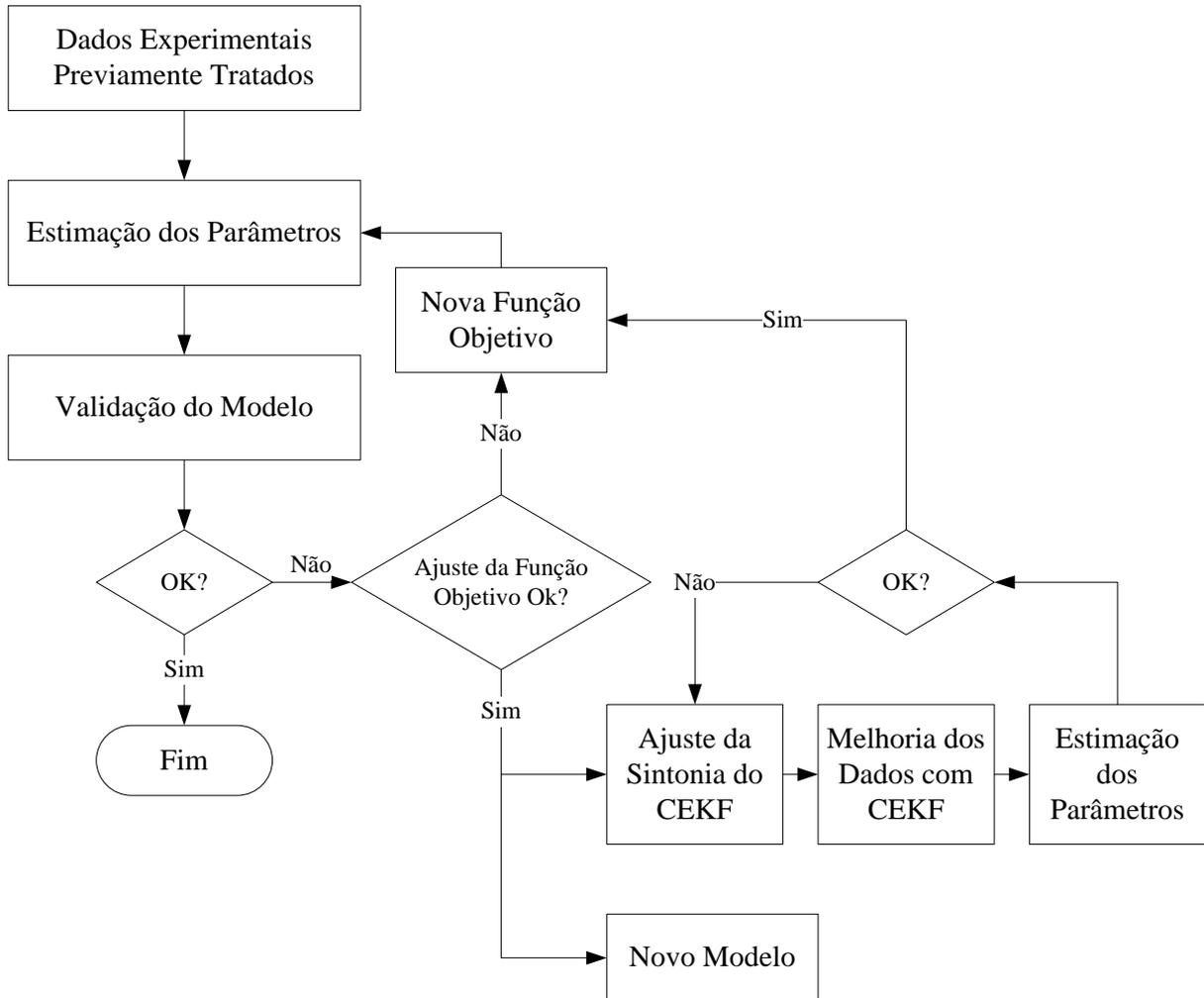
Em Bard (1974), Arnold e Beck (1977), que são livros clássicos sobre estimação de parâmetros não se encontram métodos similares ao proposto.



**Figura 1.1:** Forma geral de se identificar processos vinculada com o objetivo do presente trabalho.

A fim de aplicar a metodologia proposta foi tomado um estudo de caso que consiste em uma planta de seis tanques esféricos com vazões interligadas. De forma geral, na Figura 1.2 é apresentada a metodologia que se elaborou no que diz respeito a estimação dos parâmetros e aplicação do CEKF.

No presente trabalho, como técnica de estimação de parâmetros se utilizou mínimos quadrados. Na Figura 1.2 quando se fala em alteração da função objetivo se refere em considerar também as derivadas medidas, bem como os estados e derivadas estimados pelo CEKF. Isso tudo com atribuição de pesos a estas parcelas.



**Figura 1.2:** Forma resumida da metodologia que se elaborou no que diz respeito a estimação dos parâmetros e aplicação do CEKF.

No estudo de caso utilizado os valores dos parâmetros eram conhecidos previamente, pois os valores experimentais foram simulados computacionalmente. Conhecer de antemão os valores dos parâmetros ( $\theta$ ) foi uma estratégia tomada a fim de analisar a hipótese levantada sobre o uso do estimador de estados, isso por meio da comparação direta  $\theta_{Estimados} \times \theta_{Originais}$ .

No Capítulo 2, onde é feita uma revisão sobre o tema identificação de processos, é bem enfatizada a importância da estimabilidade dos parâmetros. O Capítulo 3 esclarecerá a ideia que levou a investir no estudo do estimador de estados como ferramenta para a estimação de parâmetros. No capítulo 4 é apresentada a metodologia. No Capítulo 5, onde são apresentados os resultados, é observada de forma prática a importância do tratamento prévio dos dados experimentais, no caso estudado se limitou a filtragem dos dados.

## Capítulo 2

### Revisão: identificação de processos

#### 2.1 Modelos e definições

Pode-se definir o modelo como uma aproximação da realidade que procura representar fenômenos observados, permitindo estimar fatos ou eventos sob diferentes circunstâncias e de uma forma ou de outra permitir um melhor entendimento da realidade. Assim, o modelo é construído a partir de um conjunto de observações experimentais nas quais a ocorrência e evolução são relacionadas a uma ou mais variáveis: as variáveis explicam a observação.

É importante lembrar que o modelo deve ser jamais confundido com a realidade. O modelo é apenas uma tentativa de explicar a realidade, baseado nas observações disponíveis e em um conjunto de considerações admitidas e em muitos casos baseadas em leis físicas, químicas e biológicas incluindo normalmente equações constitutivas. Nenhum modelo é capaz de descrever a realidade completamente por diversos motivos, a saber: (i) porque todo dado experimental é corrompido por erros de medida, (ii) porque nem todas as variáveis podem ser controladas e/ou medidas precisamente durante os testes experimentais e existe uma série de distúrbios não medidos e (iii) o modelo tem capacidade de representar parcialmente o sistema. Por isso, o bom pesquisador não tem apego a qualquer modelo e está pronto a modificá-lo sempre que começam a surgir observações experimentais confiáveis e que já não podem mais ser explicadas pelo seu modelo original.

O modelo pode ser apresentado de várias formas, vejamos algumas (PINTO e SCHWAAB, 2007):

- **Modelo Matemático.** O modelo matemático é aquele que propõe que as relações entre as diversas variáveis de um fenômeno podem ser descritas de forma matemática precisa, a grande vantagem é a previsão quantitativa.
- **Modelo Conceitual.** O modelo conceitual é aquele que estabelece vínculos qualitativos entre as variáveis de um problema, sem que se estabeleçam

necessariamente vínculos matemáticos. Esses modelos são muito utilizados em campos da ciência onde a mensuração das variáveis é complexa ou impossível, como na área de ciências humanas e, como protótipo de modelos matemáticos a serem desenvolvidos a *posteriori*. Um exemplo clássico de modelo conceitual é o conhecido complexo de Édipo, usado para explicar as relações que se estabelecem entre pais e filhos na psicologia. Modelos conceituais extremamente importantes na engenharia são as considerações de que a massa, a energia e a quantidade de movimento do universo se conservam. O primeiro modelo é o de princípio de Lavoisier, o segundo é a primeira lei da termodinâmica e o terceiro é constituído pelas leis fundamentais de Newton para o movimento. Tais pressupostos consistem no ponto de partida para a quase totalidade dos modelos matemáticos que se estabelecem nos vários ramos da engenharia. Ou seja, na engenharia muitos modelos matemáticos são precedidos por modelos conceituais. Salienta-se que já se sabe há muitos anos que o princípio de Lavoisier não é válido quando os objetos se movem a alta velocidade. Nesse caso, é necessário utilizar a Teoria da Relatividade para se acompanhar com mais precisão as variações de massa do universo. O princípio de Lavoisier, no entanto, descreve muito bem o comportamento que fazem parte do nosso cotidiano, constituindo por isso muito da base da engenharia química. Isso mostra que um modelo não precisa ser completo nem descrever todos os detalhes da realidade para que seja útil na prática. Na realidade, um modelo é útil quando fornece as respostas desejadas, com exatidão compatível com a qualidade dos dados experimentais, com o maior grau de simplicidade possível. É importante se frisar que a utilidade e complexidade não são sinônimos.

- **Modelo físico – a planta piloto.** As plantas piloto são outro tipo de modelo: um modelo físico de uma planta real, porém em escala menor. A planta piloto se enquadra na definição de modelo pois ela é tomada como uma representação da planta real, comportamentos esperados para a planta real são concluídos à partir de simulações na planta piloto.

Assim, podemos dizer que um modelo resulta sempre de um trabalho de investigação em qualquer área de conhecimento, já que o objetivo central da ciência é correlacionar dados e fatos. Sob o ponto de vista da engenharia, pela própria natureza prática e exata, o grande objetivo na construção de modelos é se chegar em relações matemáticas precisas entre as várias variáveis de um problema. Para os fins dessa dissertação, quando se fala em desenvolvimento de modelos passamos a nos referir especificamente aos modelos matemáticos.

## 2.2 Identificação de processos

No conceito geral de gerência de operações, o processo é definido como a atividade principal de uma unidade de produção, onde, com o auxílio de atividades secundárias, através do processo a matéria-prima é convertida em um produto. Nesse contexto, o processo pode ser dito como uma seqüência de passos, tarefas e atividades que resultam em uma transformação.

No âmbito da engenharia química, que consiste no foco do presente trabalho, o processo está fundamentado em fenômenos de transferência de calor, massa, movimento e reações químicas. Ou seja, em decorrência de entradas ( $u$ ) ocorrem tais fenômenos resultando em um ou mais estados. Processos deste tipo possuem saídas ( $y$ ) que são diretamente ou indiretamente relacionadas com os estados resultantes ( $x$ ) e é através das medições das saídas que se consegue fazer inferências à respeito dos estados.

Quando se tem como problema a construção do modelo de um processo se deseja obter tanto as relações para os estados  $x$  como para as saídas  $y$ .

$$\begin{aligned}x &= f(x, u; \theta) \\ y &= g(x; \theta)\end{aligned}\tag{2.1}$$

Onde  $x$  representa o vetor dos estados,  $u$  o vetor das entradas,  $y$  o vetor das saídas e  $\theta$  o vetor dos parâmetros.

Todo o processo de construção de um modelo (experimentações e hipóteses) visa à definição de duas entidades básicas distintas: a estrutura do modelo e os parâmetros. A estrutura do modelo é a forma funcional através da qual as diversas variáveis do problema estão relacionadas. Os parâmetros do modelo são os números que tornam possível a previsão quantitativa das relações existentes entre as várias variáveis do problema. Por exemplo, quando se diz que duas variáveis  $x$  e  $y$  estão relacionadas linearmente, apenas se estabelece o tipo de relação funcional que existe entre as duas variáveis analisadas, como:

$$y = \theta_1 x + \theta_2\tag{2.2}$$

Estabelece-se a estrutura do modelo. No entanto, para que o modelo possa ser utilizado para previsões ou simulações, é necessário definir adicionalmente quem são os parâmetros  $\theta_1$  e  $\theta_2$ .

Quando o desafio consiste na determinação da estrutura do modelo bem como a estimação dos parâmetros, se está diante de um problema de identificação do modelo. Em um contexto mais completo, a identificação consiste em um procedimento iterativo incluindo as seguintes etapas (ASTRÖM, WITTENMARK, 1997):

- seleção da estrutura do modelo;
- planejamento de experimentos;
- estimação de parâmetros;
- validação do modelo.

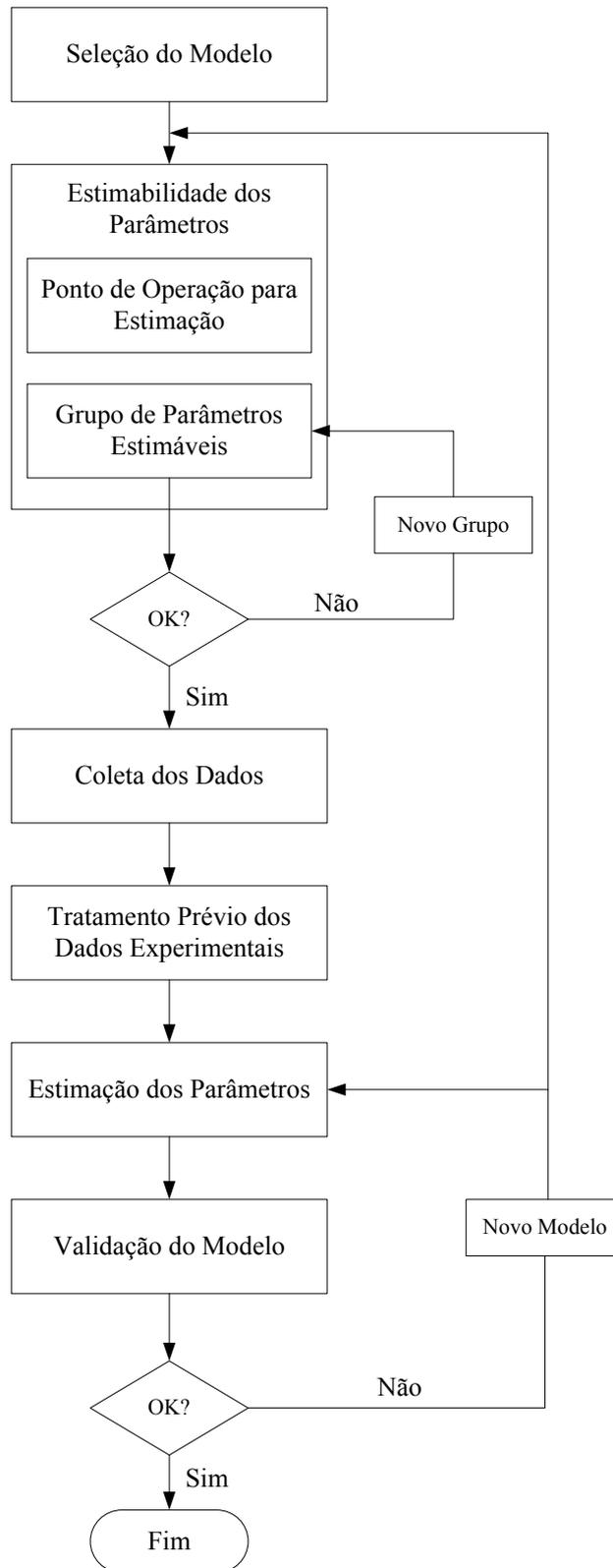
Na literatura é comum encontrar as duas técnicas clássicas de identificação: caixa branca e caixa preta. Quando se fala sobre modelos caixa branca se está falando sobre

modelos teóricos e, quando se fala sobre modelos caixa preta se está falando sobre modelos empíricos.

A caixa branca procura suprir a busca científica pela compreensão e descrição da natureza e mecanismos que regem os fenômenos físicos e/ou químicos. As relações entre as diversas variáveis do problema são construídas a partir de pressupostos teóricos fundamentais, como as leis de conservação de massa, energia e quantidade de movimento. Pode-se dizer que os modelos caixa branca derivam de modelos conceituais, que são aqueles que procuram interpretar os fenômenos. Com isso, muito frequentemente os modelos caixa branca podem ser propostos *a priori*, antes mesmo da observação experimental. Em um modelo caixa preta as relações entre as diversas variáveis do problema são estabelecidas não havendo qualquer pressuposto teórico que as justifique. O desenvolvimento de modelos caixa preta depende completamente da obtenção de dados experimentais confiáveis e da criatividade do analista. Como exemplo para este tipo de identificação citamos a ferramenta *System Identification ToolBox* implementada no MATLAB<sup>®</sup>, tal método requer que o pesquisador aceite uma das estruturas genéricas de modelos, onde são especificados ordens, fatores de peso etc.. A grande vantagem da caixa preta em relação a caixa branca é a possibilidade de se identificar até uma certa extensão um modelo para representar o padrão do distúrbios, como acontece através dos modelos ARMAX e BJ (BOHLIN, 2007).

Definidos os modelos caixa branca e preta podemos então definir um modelo caixa cinza, que nada mais é do que a combinação de ambos os métodos.

Dentro do enfoque da identificação de processos já introduzido no Capítulo 1 (Figura 1.1), agora na Figura 2.1 é apresentado um fluxograma mais detalhado. Observa-se em tal fluxograma que a coleta dos dados experimentais é precedida pela etapa de estimabilidade dos parâmetros, que por sua vez está sub-divida em: ponto de operação para estimação e grupos de parâmetros estimáveis. Tendo em mãos uma estrutura de modelo para o processo, mediante simulações de um determinado número de pontos de operação previamente estabelecidos é possível escolher o ponto mais interessante para a coleta dos dados experimentais, no presente trabalho sugerimos o critério baseado na sensibilidade dos parâmetros quanto a estimação. Na verdade, o processo descrito nada mais é do que um método de planejamento de experimentos. No que diz respeito ao grupo de parâmetros estimáveis, aqui, o objetivo é obter uma idéia prévia sobre quais parâmetros são realmente possíveis de serem estimados, se precavendo de problemas futuros na etapa da estimação. Na presente revisão, bem como no que foi desenvolvido nessa dissertação, se baseou em Brun et al. (2001) para a etapa da estimabilidade dos parâmetros. A próxima seção abordará em maiores detalhes este assunto.



**Figura 2.1:** Fluxograma referente a metodologia de identificação de processos.

### 2.2.1 Estimabilidade dos parâmetros

Quando se trabalha com modelos mais complexos, que são essencialmente os modelos caixa branca, se torna difícil e normalmente inviável se estimar todos os parâmetros. A fraca estimabilidade dos parâmetros ocorre quando existe forte dependência entre eles, isso quer dizer que mesmo que ocorram diferentes modificações nos valores dos parâmetros os modelos resultantes irão sempre convergir para a mesma resposta. Conforme foi introduzido anteriormente, uma solução consiste em identificar quais são os parâmetros possíveis de serem estimados e estimá-los mantendo os demais fixos em valores coerentes conhecidos ou especificados previamente (BATES e WATTS, 1988).

As técnicas mais tradicionais de avaliar a dependência entre os parâmetros são a matriz de covariância e a construção da região de confiança, representada graficamente por uma elipsóide  $n\theta$ -dimensional – maiores detalhes veja Bard (1974, p.176 e p.187) e Arnold e Beck (1977, p.378 e p.380). A matriz de covariância possui a desvantagem de permitir apenas a avaliação da dependência dos parâmetros aos pares. Brun et al. (2001) sugerem uma técnica com melhor acuracidade e facilidade, baseada na sensibilidade local das saídas:

$$S = \left. \frac{\partial g(x; \theta)}{\partial \theta} \right|_{\theta = \theta_0} \quad (2.3)$$

Caso os componentes da matriz de sensibilidade  $S$  não estejam na mesma escala recomenda-se que sejam escalonados. Com isso, é possível obter o índice de importância dos parâmetros  $\delta$  definido na Equação 2.4, onde  $ne$  e  $n\theta$  representam os números de dados experimentais e de parâmetros, respectivamente.

$$\delta_j = \sqrt{\frac{1}{ne} \sum_{i=1}^{ne} s_{ij}^2} \quad j = 1, 2, \dots, n\theta. \quad (2.4)$$

Localizados os parâmetros mais importantes, constroem-se de maneira criteriosa diferentes grupos de parâmetros identificados pela letra “K”. O objetivo é avaliar o índice de colinearidade  $\gamma_K$  de cada grupo. O índice de colinearidade, ou dependência linear, indica o quanto modificações nos parâmetros do grupo podem se auto-cancelar, ou seja, manter a resposta pouco sensível as alterações. Isso quer dizer que poderão existir casos de grupos compostos por parâmetros vistos como importantes individualmente, mas, quando agrupados passam a ser fracamente estimados. Antes de entrarmos em maiores detalhes sobre a determinação do  $\gamma_K$ , salientamos nesse ponto que além do índice de importância  $\delta$  ser útil no processo de construção dos grupos, ele também serve de critério para determinar o ponto de operação para estimação.

Ao determinar o índice de colinearidade  $\gamma_K$ , as colunas da matriz  $S$  devem ser normalizadas obtendo-se  $\tilde{S}$ .

$$\tilde{s}_j = \frac{s_j}{\|s_j\|} \quad j = 1, 2, \dots, n\theta. \quad (2.5)$$

$$\gamma_K = \frac{1}{\min_{\|\beta\|=1} \|\tilde{S}_K \beta\|} = \frac{1}{\sqrt{\lambda_K}} \quad (2.6)$$

Onde  $\beta$  corresponde a um determinado vetor com norma igual a 1,  $\tilde{S}_K$  é uma submatriz de  $\tilde{S}$  com dimensão  $ne \times k$ , sendo  $1 < k \leq n\theta$  e,  $\lambda_K$  é o menor autovalor de  $\tilde{S}_K^T \tilde{S}_K$ , ou seja é o menor valor singular da submatriz, i.e.,  $\sigma(\tilde{S}_K)$ . Note que  $\gamma_K$  é fortemente dependente do escalonamento empregado. Portanto, se o sistema não estiver escalonado apropriadamente, as conclusões advindas deste tipo de análise podem ser equivocadas.

Quanto maior o índice de colinearidade menor é a estimabilidade dos parâmetros do grupo, índices superiores a 5 já são críticos. Por exemplo, um índice de colinearidade de 20 quer dizer que a modificação na resposta causada pela alteração de um determinado parâmetro  $\theta_j$  pertencente ao grupo  $K$  pode ser compensada com apropriadas modificações nos outros parâmetros do grupo na ordem de 5% ( $1/\gamma_K \times 100$ ) (BRUN et al., 2001).

Na seqüência é apresentado resumidamente um estudo de caso abordado em Brun et al. (2001). No entanto, antes disto, salienta-se que as conclusões dos autores foram baseadas no estudo da estimabilidade em  $\hat{\theta}$ , ou seja, verificando a qualidade dos parâmetros pela dependência entre eles após a estimação. Sendo que nesta seção inicialmente se propõe em  $\theta_0$ . Na verdade, a análise da estimabilidade dos parâmetros é um processo iterativo, devendo ser aplicada antes e após a estimação, quando aplicada após passa a ser uma das ferramentas da validação do modelo.

O referido estudo de caso consiste em um modelo matemático que descreve a concentração de oxigênio dissolvido no rio Glatt localizado no nordeste da Suíça, o modelo possui quatro parâmetros:

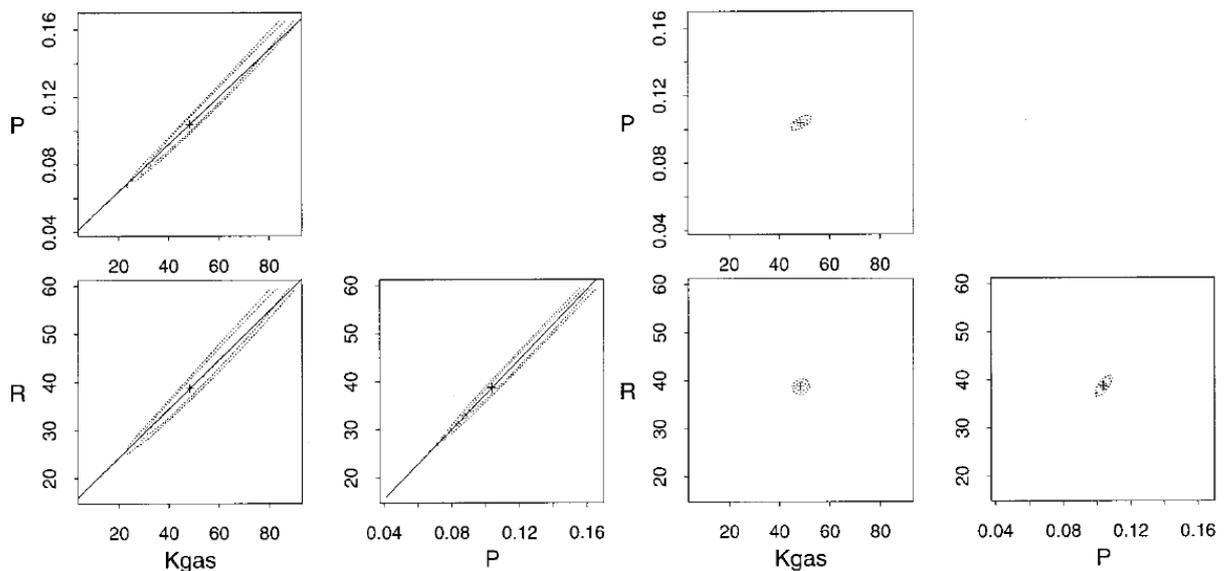
- $P$ : referente à produção de oxigênio;
- $R$ : referente à respiração do ecossistema;
- $K_{gas}$ : referente ao fenômeno de transferência de massa, que é a troca do gás com o ambiente e;
- $E_{gas}$ : referente à eficiência de reaeração.

Os autores relatam que inicialmente tentaram estimar os quatro parâmetros, entretanto, não foi possível devido a falta de convergência. Em seguida, estimaram o grupo de parâmetros  $\{P, R, K_{gas}\}$ , mantendo  $E_{gas}$  fixo. A partir do resultado da estimação ( $\hat{\theta}$ ), foram calculados a matriz de correlação, as regiões de confiança e os índices de importância e colinearidade. O objetivo no caso de estudo foi comparar as técnicas clássicas com o índice de colinearidade.

Na tabela 2.1 vemos os resultados obtidos por Brun et al. (2001) para a matriz de correlação e, na Figura 2.2 os resultados para as regiões de confiança e regiões de confiança condicional. A construção da região de confiança tem o lado negativo de que para cada construção do gráfico bidimensional  $\theta_j \times \theta_l$   $n\theta - 1$  parâmetros devem ser estimados a cada ponto, isso implica em uma demanda computacional exaustiva. A vantagem da região de confiança condicional é que ela é feita fixando os demais parâmetros no resultado da estimação:  $\hat{\theta}$ , no entanto, o resultado pode ser enganoso, como podemos ver na Figura 2.2.

**Tabela 2.1:** Matriz de correlação (BRUN et al., 2001).

Parâmetro	Matrix de Correlação		
	$K_{gas}$	$P$	$R$
$K_{gas}$	1,00	0,995	0,993
$P$	0,995	1,00	0,997
$R$	0,993	0,997	1,00



**Figura 2.2:** Resultado para as regiões de confiança construídas com 80 e 95% de confiança. O lado esquerdo se refere as regiões propriamente ditas com 80% e o lado direito se refere a região condicional com 95%. No lado esquerdo, a linha sólida adicionada se refere a projeção do autovetor correspondente ao menor autovalor (BRUN et al., 2001).

A análise prévia da matriz de correlação e da região de confiança indicaram uma fraca estimabilidade. Entretanto, na seqüência foram determinados os índices de importância de cada parâmetro e os índices de colinearidade para todos os possíveis grupos com os quatro parâmetros a fim de uma análise mais profunda, os resultados são apresentados nas Tabelas 2.2 e 2.3.

**Tabela 2.2:** Ranking da importância dos Parâmetros (BRUN et al., 2001).

Parâmetro	$\delta_j$
$P$	1,96
$R$	1,63
$K_{gas}$	0,84
$E_{gas}$	0,22

**Tabela 2.3:** Índice de Colinearidade dos Grupos de Parâmetros (BRUN et al., 2001).

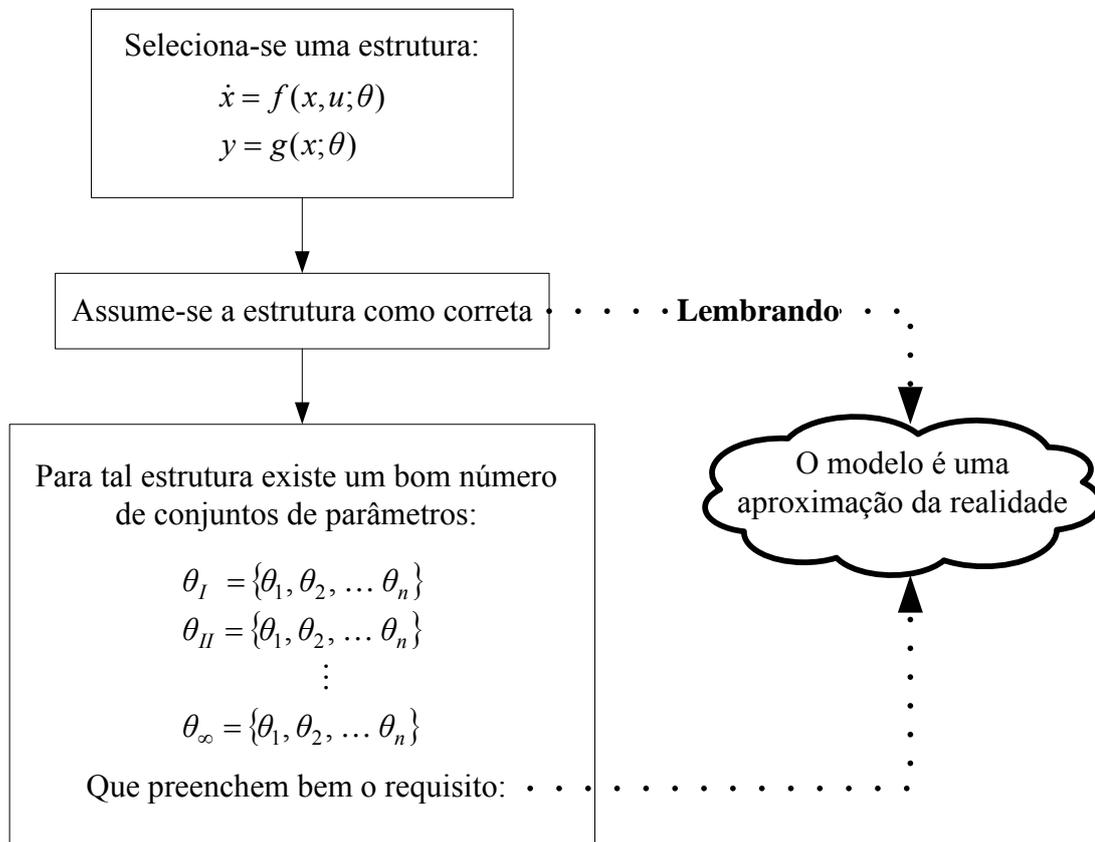
Grupo	$\gamma_K$
$E_{gas}, K_{gas}$	9,07
$P, R$	2,02
$K_{gas}, P$	1,57
$E_{gas}, P$	1,55
$K_{gas}, R$	1,04
$E_{gas}, R$	1,04
$K_{gas}, P, R$	23,81
$E_{gas}, P, R$	9,13
$E_{gas} K_{gas}, P$	9,10
$E_{gas} K_{gas}, R$	9,09
$E_{gas} K_{gas}, P, R$	59,49

Ao avaliar os índices de colinearidade vemos que o grupo composto por todos os parâmetros possui um índice extremamente alto (59,49), isso justifica o problema de convergência verificado inicialmente. Sobre o grupo  $\{P, R, K_{gas}\}$ , também se verifica um alto índice de colinearidade (23,81) e, é isto que faz com que os elementos da matriz de correlação

sejam altos e não a correlação que existe ao pares, como pode ser interpretado de forma precipitada. Este fato pode ser verificado observando o índice de colinearidade de todos os pares formados pelo grupo  $\{P, R, K_{gas}\}$ , que são baixos, o maior é de  $\{P, R\}$  sendo 2,02. A região de confiança retrata bem a colinearidade existente entre os três parâmetros. A região condicional, por sua vez, por seus contornos aproximadamente circulares apenas indica que o problema de estimabilidade não é causado aos pares. Ou seja, Brun et al. (2001) afirmam que conclusões corretas à respeito da estimabilidade dos parâmetros são feitas à partir da região de confiança ou do índice de colinearidade e, devido a construção da região poder ser exaustiva numericamente o recomendado é adotar o índice de colinearidade por sua praticidade e de fácil interpretação.

### 2.2.2 Estimação de parâmetros

Uma característica da estimação dos parâmetros é descrita a seguir:



**Figura 2.3:** Característica da estimação dos parâmetros: diversos valores possíveis para os parâmetros.

Entretanto, para a estrutura adotada como correta, assume-se a existência de  $\theta^*$ , que representa o vetor dos parâmetros que melhor representa a realidade. Logo, o desafio na

estimação dos parâmetros é obter um  $\theta$  o mais próximo de  $\theta^*$ , o resultado é a estimativa  $\hat{\theta}$ , sendo  $E\{\hat{\theta}\} = \theta^*$ .

Na literatura se encontram diversas técnicas para se chegar a  $\hat{\theta}$ , todas são baseadas na função erro:  $e(\theta)$ , seja de forma direta ou indireta (como a derivada). Como o presente trabalho enfoca os processos controlados e amostrados computacionalmente, o erro é definido como:

$$e = \underbrace{y^{sensor}}_{\text{considerado}=y} - \underbrace{y^{processo}}_{\text{considerado}=g} \quad (2.7)$$

Logo,

$$e = y - g(x; \theta) \quad (2.8)$$

Na seqüência é feita uma breve revisão sobre alguns métodos para estimação de parâmetros.

### Mínimos Quadrados

Na clássica técnica dos mínimos quadrados os parâmetros são estimados ao minimizar erro:

$$\begin{aligned} \min_{\theta} J(\theta) &= \min_{\theta} \sum_{i=1}^{ne} \sum_{j=1}^{ny} e_{ij}^2 \\ \min_{\theta} J(\theta) &= \min_{\theta} \sum_{i=1}^{ne} \sum_{j=1}^{ny} [y_{ij} - g(x_i; \theta)]^2 \end{aligned} \quad (2.9)$$

Apesar de ser bastante útil e permitir a solução de uma série de problemas práticos, a técnica de mínimos quadrados é bastante limitada porque admite implicitamente que todas as variáveis medidas pertencem a um mesmo conjunto amostral e que são medidas com a mesma precisão experimental, bem como assume que não existem desvios nas variáveis independentes.

### Máxima Verossimilhança

A técnica da máxima verossimilhança é baseada na função densidade probabilidade da variável medida:  $p(y|x; \theta)$ . Nesta seção trabalharemos com o caso mais comum, que é quando  $p(y|x; \theta)$  é uma distribuição normal, ou gaussiana. A distribuição gaussiana multivariável é igual a:

$$p(y|x; \theta) = \frac{1}{(2\pi)^{ny/2} \det^{1/2}(V_y)} \exp \left[ -\frac{1}{2} \times (y - g(x; \theta))^T V_y^{-1} (y - g(x; \theta)) \right] \quad (2.10)$$

Sendo  $V_y$  a matriz de covariância experimental e  $g(x; \theta)$  a média de  $y$ . Se as realizações experimentais são independentes, temos

$$p(y|x; \theta) = \prod_{j=1}^{ny} \frac{1}{\sqrt{2\pi\sigma_{y_j}}} \exp\left[-\frac{(y_j - g(x; \theta))^2}{2\sigma_{y_j}}\right] \quad (2.11)$$

Assim, a função de verossimilhança  $L$  fica sendo

$$L(y|x; \theta) = \prod_{i=1}^{ne} \prod_{j=1}^{ny} \frac{1}{\sqrt{2\pi\sigma_{y_{ij}}}} \exp\left[-\frac{1}{2} \times \frac{(y_{ij} - g(x_i; \theta))^2}{\sigma_{y_{ij}}}\right] \quad (2.12)$$

Os parâmetros podem ser estimados de forma a maximizar a função de verossimilhança. Isto é, maximizar a probabilidade de encontrar os dados experimentais obtidos. Pois, se o experimento foi bem feito, não há razão para acreditar que os resultados experimentais sejam pouco prováveis e não representem satisfatoriamente a realidade. Reescrevendo a função de verossimilhança na forma logarítmica obtém-se (HIMMELBLAU, 1970):

$$\ln L(y|x; \theta) = \sum_{i=1}^{ne} \sum_{j=1}^{ny} \ln \frac{1}{\sqrt{2\pi\sigma_{y_{ij}}}} - \sum_{i=1}^{ne} \sum_{j=1}^{ny} \frac{1}{2\sigma_{y_{ij}}^2} [y_{ij} - g(x_i; \theta)]^2 \quad (2.13)$$

Observa-se que maximizar  $\ln L(y|x; \theta)$  em relação aos parâmetros é equivalente a

$$\min_{\theta} J(\theta) = \min_{\theta} \sum_{i=1}^{ne} \sum_{j=1}^{ny} \frac{1}{\sigma_{y_{ij}}^2} [y_{ij} - g(x_i; \theta)]^2 \quad (2.14)$$

A Equação 2.14 também pode ser chamada como um problema de mínimos quadrados ponderados. Ou seja, quando a estimação por máxima verossimilhança está baseada na distribuição Gaussiana e as variáveis independentes não possuem desvios o problema se reduz a mínimos quadrados ponderados.

Uma das grandes vantagens da função objetivo da Equação 2.14 é a possibilidade de misturar diferentes conjuntos de dados. Por exemplo, durante a análise de uma reação, podem ser medidas temperaturas e pressões como variáveis dependentes, que são variáveis de conjuntos de dados diferentes. Além disso, os erros de medição de cada uma das variáveis podem mudar de ponto para ponto. Por exemplo, a incerteza de medição de temperatura costuma aumentar com o aumento da temperatura.

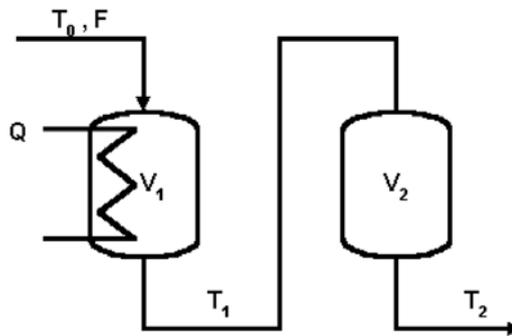
O grande mérito da máxima verossimilhança é permitir a estimação considerando os desvios nas variáveis independentes  $x$ . Admitindo-se que ambas as variáveis:  $y$  e  $x$  apresentam

distribuição Gaussiana em torno de suas médias,  $g(\bar{x})$  e  $\bar{x}$ , e as realizações experimentais são independentes, a função objetivo fica sendo:

$$\min_{\theta, \bar{x}} J(\theta, \bar{x}) = \min_{\theta, \bar{x}} \left\{ \sum_{i=1}^{ne} \sum_{j=1}^{ny} \frac{1}{\sigma_{y_{ij}}^2} [y_{ij} - g(\bar{x}_i; \theta)]^2 + \sum_{i=1}^{ne} \sum_{k=1}^{nx} \frac{1}{\sigma_{x_{ik}}^2} [x_{ik} - \bar{x}_{ik}]^2 \right\} \quad (2.15)$$

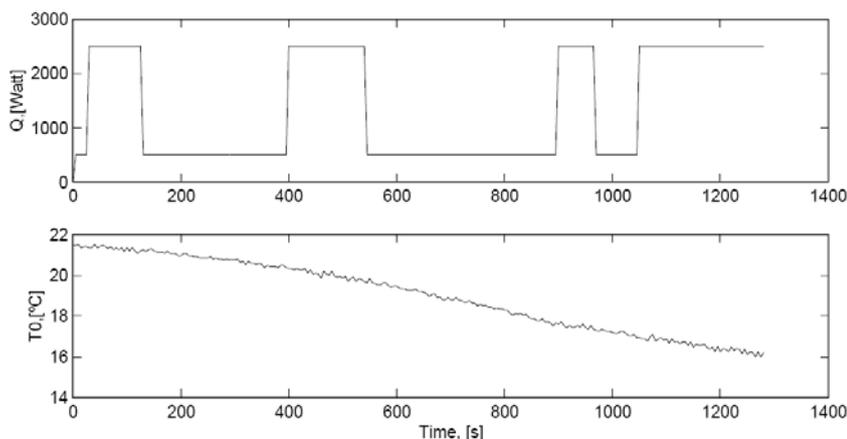
### Minimização da Derivada do Erro

Machado e Trierweiler (2004) propuseram e compararam duas funções objetivo:  $J = \min \sum e^2$  e  $J = \min \sum (de/dt)^2$ , quando um processo é submetido a distúrbios não medidos, consideraram como caso de estudo a planta descrita na Figura 2.4.



**Figura 2.4:** Representação esquemática da planta estudada pelos autores: processo de aquecimento com dois tanques (MACHADO e TRIERWEILER, 2004).

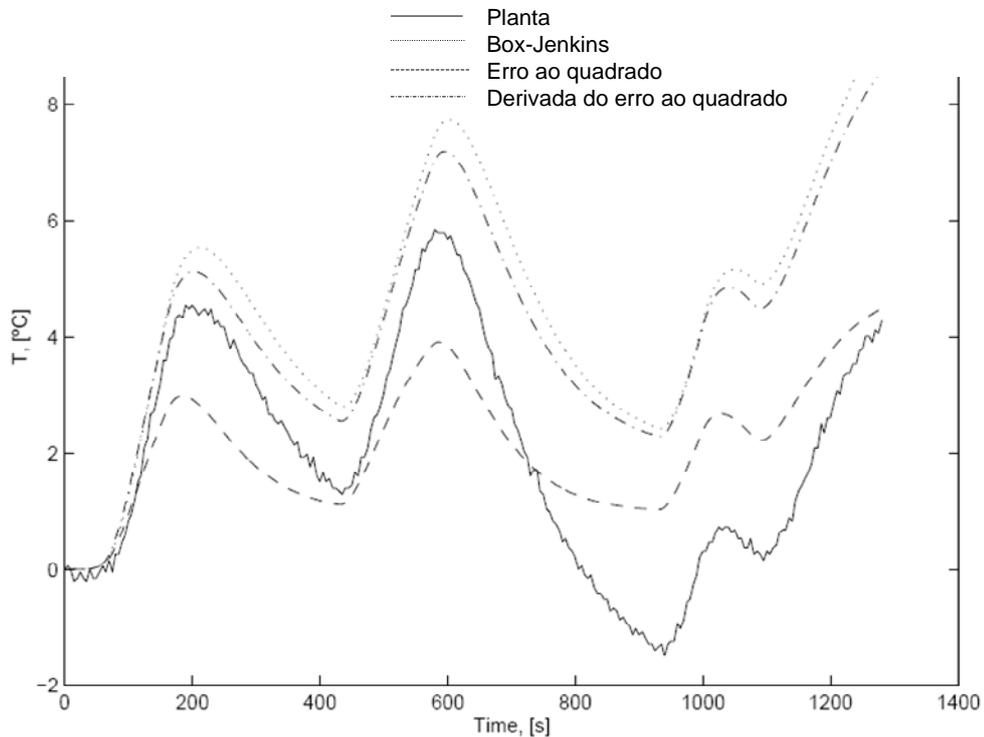
O processo possui duas entradas: a variável manipulada  $Q$  e o distúrbio  $T_0$ . O objetivo foi obter um modelo para a variável controlada  $T_2$ . Apesar de no processo de identificação Machado e Trierweiler (2004) considerarem  $T_0$  como não medida, eles dispunham de tal informação. Na Figura 2.5 são apresentadas as entradas.



**Figura 2.5:** Entradas experimentais: variável manipulada  $Q$  e distúrbio  $T_0$ , considerado como não medido (MACHADO e TRIERWEILER, 2004).

Os autores estabeleceram um modelo caixa preta para o processo e utilizando ambos critérios:  $J = \min \sum e^2$  e  $J = \min \sum (de/dt)^2$  realizaram a identificação. Também,

utilizando a estrutura Box-Jenkins do MATLAB<sup>®</sup> *System Identification Toolbox* realizaram mais uma identificação, o resultados obtidos foram:



**Figura 2.6:** Resultados obtidos para as três formas de identificação em comparação com os dados experimentais da planta (MACHADO, TRIERWEILER, 2004).

Analisando os resultados, os autores concluíram que o método  $J = \min \sum (de/dt)^2$  produz melhores resultados do que  $J = \min \sum e^2$  quando uma planta sofre com distúrbios não medidos. Pois, a capacidade de seguir a tendência do processo é melhor do que tentar passar por todos os pontos experimentais. Em outras palavras, como podemos observar na Figura 2.6, com  $\min \sum (de/dt)^2$  se obtêm um melhor acompanhamento do perfil do comportamento da planta.

Desde que distúrbios não medidos estão sempre presentes nos processos reais, formular uma função objetivo considerando o termo  $de/dt$  consiste em uma interessante alternativa. Esta alternativa irá dar bons resultados somente para os casos em que a variação provocada pelo distúrbio não medido seja mais lenta que as variações provocadas pelas variáveis manipuladas. Isto é fácil de entender, pois através da derivada uma variação que pode ser aproximada por uma rampa (como a que representa a temperatura  $T_0$  no estudo de Machado e Trierweiler (2004)), se obterá uma constante. Os distúrbios que variam lentamente são chamados de *drift*. Aqui deve se considerar a derivada como sendo apenas mais uma alternativa que em casos de distúrbios tipo *drift* trarão uma significativa melhora no processo de identificação de modelos dinâmicos.

## Outros Métodos de Estimação

Além dos métodos já abordados citam-se outros de forma breve (ARNOLD e BECK, 1977; BARD, 1970):

- **Estimação de Bayesian.** O teorema de Bayesian define uma função densidade probabilidade posterior:  $p''(y|x;\theta) = cL(y|x;\theta)p'(y|x;\theta)$ , considerando que se parte de um conhecimento prévio:  $p'(y|x;\theta)$ . Com base no teorema de Bayesian, pode-se estimar os parâmetros pela busca da moda de  $p''$ , este método também é referido como MPD (Máxima da Distribuição Posterior). Outra aplicação do teorema de Bayesian é a Estimação com o Mínimo Risco (MRE). Embora usualmente se procure estimar os parâmetros o mais próximo dos reais, em certos casos a estimação tem um objetivo mais específico. Ou seja, deseja-se encontrar os valores de  $\theta$  que são ótimos sim, mas para um propósito particular. Em certos casos o ótimo é determinado por considerações econômicas. A uma função “Risco” é baseada a um custo atribuído a cada erro de estimação.
- **Desvio Minimax.** Os parâmetros são determinados quando se minimiza o máximo desvio existente entre o modelo e os dados experimentais.
- **Pseudomáxima Verossimilhança.** É quando se aplica o método de Máxima Verossimilhança assumindo que a distribuição é Gaussiana sem se ter de fato essa convicção. Podemos considerar mínimos quadrados ponderados e não-linear como aplicações deste método.
- **Mínimo Qui-Quadrado.** Dado um determinado intervalo de valores possíveis para  $y$   $[y_a, y_b]$ , e um número de observações experimentais  $N$ , por meio da distribuição de probabilidade é possível se determinar o número esperado de observações a cair no intervalo pré-definido ( $N_{eo}$ ). Desta forma, a função objetivo é formulada a minimizar a diferença entre o número esperado de observações com o número de fato ocorrido ( $N_{oo} - N_{eo}$ ) para diferentes intervalos  $[y_a, y_b]$  e, a função objetivo também é ponderada por  $N_{oo}$  ou  $N_{eo}$ .

### 2.2.3 Validação do modelo

Apesar não terem sido aplicados na presente dissertação, nesta seção 2.2.3 é feita uma breve revisão sobre os métodos tradicionais de validação do modelo com objetivo informativo. Como já introduzido no Capítulo 1, no estudo de caso da presente dissertação os valores dos parâmetros  $\theta^*$  são conhecidos previamente, pois os valores experimentais foram simulados computacionalmente. Desta forma, a avaliação da adequação do modelo pôde ser simplificada com a comparação direta  $\theta_{Estimados} \times \theta_{Originais}$ , o que será visto de forma clara como foi realizada tal comparação no Capítulo 5 dos resultados, mas não indo além do que o cálculo convencional de médias e desvios padrão referente às diferenças:  $\theta_{Estimado} - \theta_{Original}$ .

Na prática, a validação do modelo consiste em tomar o modelo obtido com a estimação e aplicar testes estatísticos a fim de se encontrar respostas para as seguintes questões:

- se o modelo é adequado;
- se os parâmetros são de boa qualidade;
- o intervalo de predição do modelo.

### Adequação do modelo

A adequação do modelo pode ser verificada mediante a utilização das seguintes ferramentas (PINTO e SCHWAAB, 2007):

- Construção de um gráfico  $y \times g$  e verificar o coeficiente de correlação, que indica o quão próximos estão os dados do modelo com os dados experimentais. Usualmente, se o coeficiente de correlação tem módulo superior a 0,9, o modelo é considerado satisfatório. Valores inferiores a 0,9 podem indicar tanto desajuste do modelo quanto erros experimentais excessivos:
- Testes de hipóteses.

No caso da estimação por mínimos quadrados, ao se considerar a hipótese do modelo perfeito e do experimento bem feito se obtêm a estimativa da variância dos erros experimentais  $\hat{\sigma}_y^2$ :

$$\hat{\sigma}_y = \frac{J(\hat{\theta})}{ne - n\theta} \quad (2.16)$$

$$\hat{\sigma}_y = \frac{1}{ne - n\theta} \min_{\theta} \sum_{i=1}^{ne} \sum_{j=1}^{ny} \frac{1}{\sigma_{y_{ij}}^2} [y_{ij} - g(x_i; \theta)]^2$$

Sendo  $ne - n\theta$  o número de graus de liberdade.

Se a variância dos erros experimentais  $\sigma_y^2$  também for conhecida é possível verificar a adequação da hipótese do modelo perfeito pela aplicação do teste  $F$ . Para isso,  $F$  deve atender o critério:

$$\frac{1 - \alpha}{2} < p(F) < \frac{1 + \alpha}{2} \quad (2.17)$$

Sendo  $\alpha$  o grau de certeza e  $F$ :

$$F = \frac{\hat{\sigma}_y^2}{\sigma_y^2} \quad (2.18)$$

Se as duas variâncias são semelhantes estatisticamente, o modelo deve ser considerado satisfatório. Caso contrário, é necessário continuar investindo na modelagem do sistema, como a busca por uma nova estrutura para o modelo.

No caso da função de máxima verossimilhança com erros experimentais de distribuição Gaussiana, a função objetivo resultante tem interpretação de  $\chi^2$ . Nesse caso, o modelo é considerado adequado se

$$\chi_{\min}^2 < J(\hat{\theta}) < \chi_{\max}^2 \quad (2.19)$$

### Qualidade dos parâmetros obtidos

Sobre a avaliação da qualidade dos parâmetros muito já foi visto na seção 2.2.1, em especial sobre a vantagem do uso do coeficiente de colinearidade  $\gamma$ . Contudo, a determinação da matriz de covariância  $V_{\theta}$  ainda pode ser interessante, aqui é considerado o caso simples em que os erros nas variáveis independentes são negligenciáveis, para casos mais complexos é sugerido Himmelblau (1970), Arnold e Beck (1977) e Bard (1974). Assim, se pode obter  $V_{\theta}$  com (HIMMELBLAU, 1970):

$$V_{\theta} = H^{-1} G V_y G^T (H^{-1})^T \quad (2.20)$$

Onde,

$$\begin{aligned} H &= \frac{\partial}{\partial \theta} \nabla_{\theta} J(x, y; \hat{\theta}) \\ G &= \frac{\partial}{\partial y} \nabla_{\theta} J(x, y; \hat{\theta}) \end{aligned} \quad (2.21)$$

Com isso, também se pode obter o grau de significância do parâmetro, que é normalmente calculado em relação à referência zero, no qual o efeito paramétrico desaparece. De forma simples, o grau de significância pode ser obtido variando-se o grau de confiança até que o intervalo de confiança do parâmetro inclua o zero. Se o grau de confiança necessário para incluir o zero é inferior ao grau de confiança estabelecido pelo investigador, diz-se que o parâmetro é não significativo. Nesse caso, há argumentos estatísticos suficientes para remover o parâmetro (e o respectivo efeito) do modelo matemático (PINTO e SCHWAAB, 2007).

### Intervalo de predição do modelo

Após o modelo passar pelo teste de adequação é necessário determinar em qual intervalo das variáveis independentes o modelo pode ser aplicado. Dessa forma, se verifica como os erros paramétricos se propagam através do modelo e viram erros de predição. Para isso o erro de predição é tido como (HIMMELBLAU, 1970):

$$\Delta \tilde{y} = y - g(x; \hat{\theta}) \quad (2.22)$$

Considerando um problema de estimação onde as variáveis independentes não estão sujeitas a erros de medição obtemos  $y$  através de uma aproximação por da série de Taylor:

$$y = g(x; \hat{\theta}) + \frac{\partial g}{\partial \theta} \Delta\theta + \tilde{e} \quad (2.23)$$

Onde,  $\tilde{e}$  se refere aos erros de predição e  $\Delta\theta = \hat{\theta} - \theta^*$ . Sendo assim, podemos quantificar  $\Delta\tilde{y}$ , o que nos permite calcular a matriz de covariância da predição  $V_{\tilde{y}}$ .

$$\begin{aligned} \Delta\tilde{y} &= \frac{\partial g}{\partial \theta} \Delta\theta + \tilde{e} \\ V_{\tilde{y}} &= E\{\Delta\tilde{y}\Delta\tilde{y}^T\} = BV_{\theta}B^T + V_{\tilde{e}} \end{aligned} \quad (2.24)$$

Sendo,

$$B = \left. \frac{\partial g}{\partial \theta} \right|_{\theta=\hat{\theta}} \quad (2.25)$$

Na prática, costuma-se assumir que  $\sigma_{\tilde{e}}$  é equivalente a  $\sigma_y$ , dessa forma:

$$V_{\tilde{y}} = BV_{\theta}B^T + V_y \quad (2.26)$$

Com isso, se observa que as predições são afetadas:

- pelos erros passados, através dos erros nos parâmetros:  $V_{\theta}$ ;
- pelos erros futuros:  $V_y$ .

Munidos da Equação 2.26, é possível conhecer a região onde a predição pode ser feita de forma adequada, isto é, o intervalo de valores da variável independente  $x$  que fazem com que as variâncias dos erros de predição  $\sigma_{\tilde{y}}$  e de experimentação  $\sigma_y$  passem no teste  $F$ .

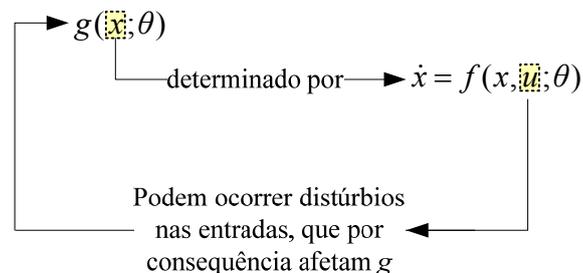
## Capítulo 3

# Processos com distúrbios e proposta de aplicação do CEKF

Frente ao comprometimento da estimação dos parâmetros devido a distúrbios nas entradas e ao se saber que estimadores de estados são ferramentas indicadas para se estimar estados quando os processos são sujeitos a distúrbios, surgiu a proposta do presente trabalho: empregar o CEKF conjuntamente com a estimação dos parâmetros e verificar se de fato há uma melhoria. Para tanto, nesta seção será abordada a problemática da identificação de processos com distúrbios nas entradas. Em especial, quando existe uma componente determinística no distúrbio o comprometimento na estimação é maior. Também serão abordados exemplos de processos com distúrbios, a motivação que levou a estudar o emprego do CEKF, bem como uma breve abordagem sobre o funcionamento de um estimador de estados. E por fim, a proposta do presente trabalho propriamente dita.

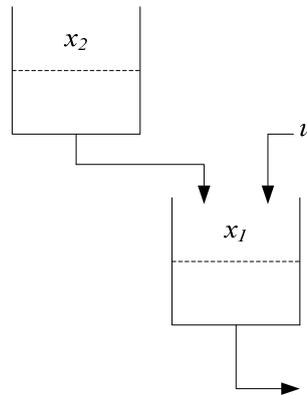
### 3.1 Problema da identificação de processos com distúrbios

Sabe-se que ao identificar um processo é estabelecido um modelo  $g(x; \theta)$  para prever a saída e a seguir se calcular o erro, que mede a discrepância entre o valor predito e o medido, sendo que a Figura 3.1 ilustra o procedimento de cálculo.



**Figura 3.1:** Esquematização da influência dos distúrbios nas entradas sobre as saídas.

Vejamos tal caso considerando o exemplo clássico de dois tanques interligados (Figura 3.2). Aqui é focado o modelo do tanque inferior.



**Figura 3.2:** Processo de dois tanques tomado como exemplo.

O modelo adotado para descrever a taxa do estado  $x_1$  no tempo é

$$A \frac{dx_1}{dt} = \theta_2 \sqrt{x_2} - \theta_1 \sqrt{x_1} + u \quad (3.1)$$

Com respeito à vazão de entrada  $u$ , considera-se que se deseja que  $u = cte$  ( $cte = constante$ ). No entanto, é sabido que nos processos reais isso não acontece. Em cada momento que o processo é colocado em operação  $u$  pode sofrer a influência de um determinado distúrbio que não significa, necessariamente, que tal distúrbio será reproduzido em uma segunda operação. Os maiores adversários da identificação de processos são os distúrbios determinísticos desconhecidos. Seguindo o exemplo da Figura 3.2, é considerado o caso em que um pesquisador coleta dados na planta em que ocorreu um distúrbio determinístico em  $u$ , sendo melhor modelada como  $u = cte - \theta_3 t$ . Logo, o distúrbio determinístico no processo é caracterizado pelo termo  $\theta_3 t$ . O pesquisador pode estipular um dos dois modelos da Tabela 3.1, segundo o seu conhecimento

**Tabela 3.1:** Duas situações para modelar o processo de exemplo.

	Modelo	Pesquisador
1	$A\dot{x}_1 = \theta_2 \sqrt{x_2} - \theta_1 \sqrt{x_1} + cte$	Desconhece a existência do distúrbio e acredita, portanto, que $u = cte$ . Ou, sabe da existência do distúrbio, mas, o negligencia no modelo.
2	$A\dot{x}_1 = \theta_2 \sqrt{x_2} - \theta_1 \sqrt{x_1} + cte - \theta_3 t$	Sabe da existência do distúrbio e o considera no modelo. Nesse caso, têm-se mais um parâmetro a ser identificado: $\theta_3$ .

No primeiro caso da Tabela 3.1 os valores de  $\theta_1$  e  $\theta_2$  acabam sendo corrompidos pelo distúrbio. Já no segundo caso, o distúrbio é filtrado pelo termo  $-\theta_3 t$  de forma que  $\theta_1$  e  $\theta_2$  já não são mais corrompidos como no caso anterior. Percebe-se aqui a importância da filtragem dos distúrbios.

Com base no que foi apresentado, podemos dizer que a melhor forma para se estimar parâmetros quando ocorrem distúrbios segue da seguinte maneira: modelo fundamental + distúrbios = modelo específico. Ou seja, partimos de uma estrutura fundamental, como na Equação 3.1 e, à medida que vamos agregando termos referentes aos distúrbios que ocorreram durante a coleta dos dados o modelo vai ganhando especificidade.

Na linguagem caixa branca e preta, dizemos que o modelo fundamental é basicamente caixa branca. Já os modelos para os distúrbios podem ter natureza caixa branca ou caixa preta. Na prática, para distúrbios complexos e difíceis de serem explicados o melhor recurso que se tem é a caixa preta.

No que diz respeito aos distúrbios relacionados às entradas, eles podem ter natureza estocástica e determinística. Para melhor tratamento matemático um distúrbio pode ser decomposto em ambas as naturezas:

- parcela determinística: média;
- parcela estocástica: desvio da média.

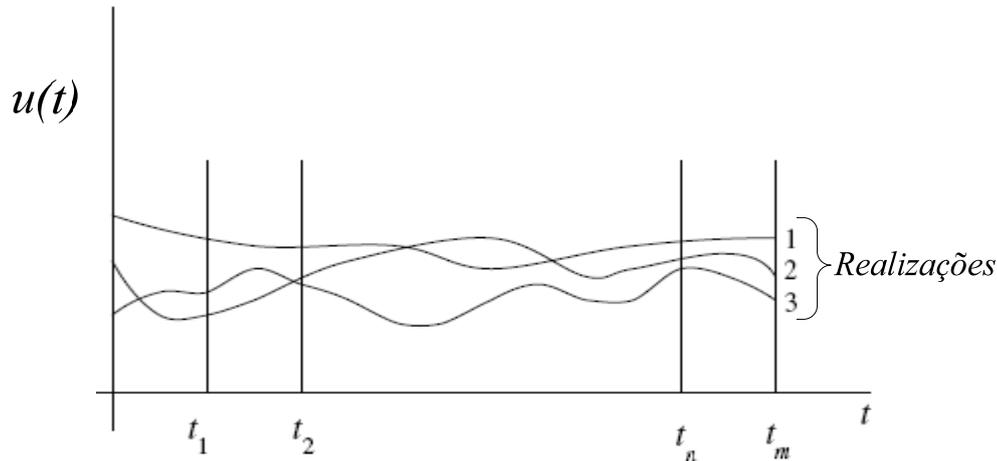
Com isso, ao identificar um processo o pesquisador deve atentar nos distúrbios atuantes, pois a qualidade dos parâmetros será comprometida caso eles sejam desprezados. A caracterização da parcela estocástica permite delimitar uma região de possibilidades para o processo segundo ferramentas e critérios probabilísticos e, a importância da caracterização da parcela determinística já se pôde ver no exemplo dos dois tanques. Para trazer um exemplo à respeito deste assunto: parcela determinística/estocástica, antes faremos uma breve abordagem sobre o processo estocástico e, na seqüência, o simples exemplo de um processo que tem como distúrbio o calor proveniente do ambiente.

### **3.1.1 Processo estocástico**

Sistemas dinâmicos são muitas vezes sujeitos as variáveis de entrada que não são funções exatamente especificadas no tempo, em oposição às tradicionais funções degrau, impulso, harmônica ou outras funções padrões. Já sabemos que quando as variáveis de entrada podem assumir valores aleatórios ao longo do tempo o recurso que se tem é a determinação da probabilidade de sua influência em um determinado tempo. Isso não quer dizer que sua influência não possa ser prevista, apenas a acuracidade é reduzida em relação a uma entrada determinística (ASTRÖM e WITTENMARK, 1997).

Uma variável estocástica é uma simples generalização de variável aleatória: a variável estocástica é uma variável aleatória que muda em função de outra variável, que no nosso caso é o tempo. Suponhamos um processo que foi realizado três vezes com uma entrada estocástica

$u(t)$  (Figura 3.3). Observa-se que os valores de  $u(t)$  nas três realizações são diferentes nos mesmos instantes de tempo. Assim, um processo estocástico é aquele que está sujeito a variáveis estocásticas. Em outras palavras, em qualquer instante de tempo o valor do processo estocástico é uma variável aleatória em decorrência de suas entradas aleatórias (FIKAR e MIKLES, 2007; SIMON, 2006).



**Figura 3.3:** Entrada estocástica de um determinado processo realizado três vezes (adaptado de FIKAR; MIKLES, 2007).

Sabendo que  $u(t)$  é uma variável aleatória, podemos dizer

$$u(t) = u_{m\acute{e}dia} + desvio \quad (3.2)$$

Quando  $u_{m\acute{e}dia}$  é uma entrada desejável para o processo, o desvio nada mais é do que um distúrbio, ou ruído de processo, usualmente representado por “ $w$ ”. Possuindo  $u_{m\acute{e}dia}$  um valor bem determinado, a estocasticidade da entrada se deve ao distúrbio “ $w$ ”. Assim, o modelo do processo estocástico pode ser representado

$$\dot{x} = f(x, u; \theta) = f(x, \underbrace{u_{m\acute{e}dia}}_u, w; \theta) \quad (3.3)$$

Sendo

$$f(x, \underbrace{u_{m\acute{e}dia}}_{omitido}, w; \theta) = f(x, u, w; \theta) \quad (3.4)$$

Considerando que  $w$  entra na forma linear e, assim será considerado daqui por diante neste trabalho, obtemos

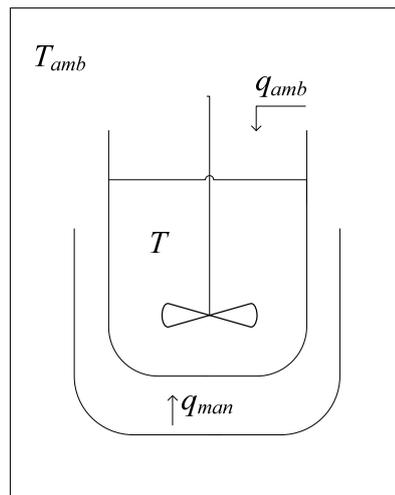
$$\dot{x} = f(x, u, w; \theta) = f(x, u; \theta) + w \quad (3.5)$$

### 3.1.2 Processo com distúrbio determinístico e estocástico

Nesta seção é apresentado um simples exemplo de um distúrbio causado pelo ambiente, para isto é adotado o processo apresentado na Figura 3.4, que consiste em um trocador de calor encamisado. Suponhamos que quando o pesquisador foi realizar o experimento o sistema de isolamento não estava bem acondicionado, logo, além da entrada desejada, que é a manipulada  $q_{man}$ , o processo também recebeu o calor proveniente do ambiente  $q_{amb}$  como distúrbio. Neste caso, seguindo a nomenclatura de modelo fundamental e específico tem-se a Tabela 3.2.

**Tabela 3.2:** Modelos fundamental e específico para o exemplo do trocador com distúrbio.

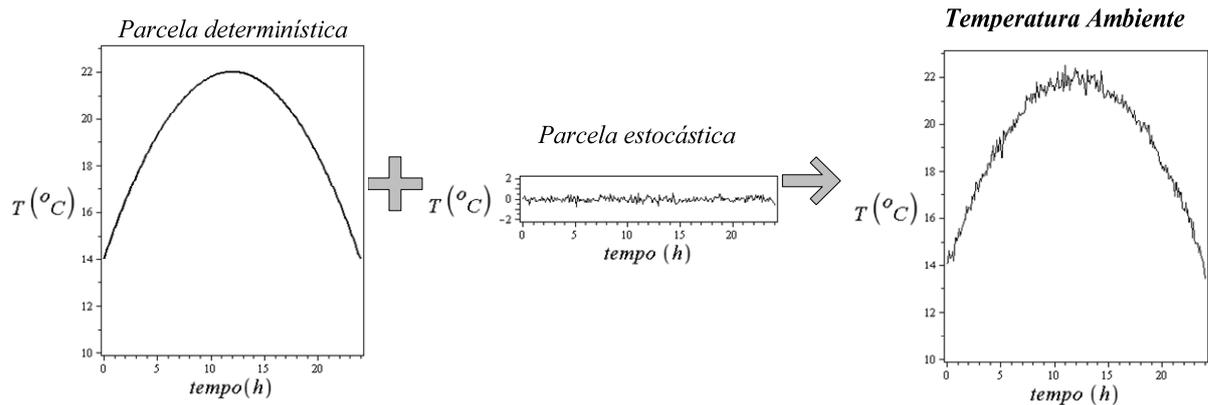
Modelo Fundamental	Modelo Específico
$\rho V c_p \frac{dT}{dt} = q_{man}$	$\rho V c_p \frac{dT}{dt} = q_{man} + q_{amb}$



**Figura 3.4:** Exemplo do processo sujeito ao distúrbio provocado pelo ambiente.

Neste exemplo nota-se a importância do conhecimento da temperatura ambiente  $T_{amb}$  para uma identificação de qualidade, pois sabemos:  $q_{amb} = hA(T - T_{amb})$ .

A temperatura ambiente é uma entrada estocástica, no entanto, pode ser dividida nas duas parcelas de modo a melhorar a qualidade da estimação dos parâmetros, como pode ser visto na Figura 3.5.



**Figura 3.5:** Visualização da temperatura ambiente formada por duas parcelas: determinística e estocástica.

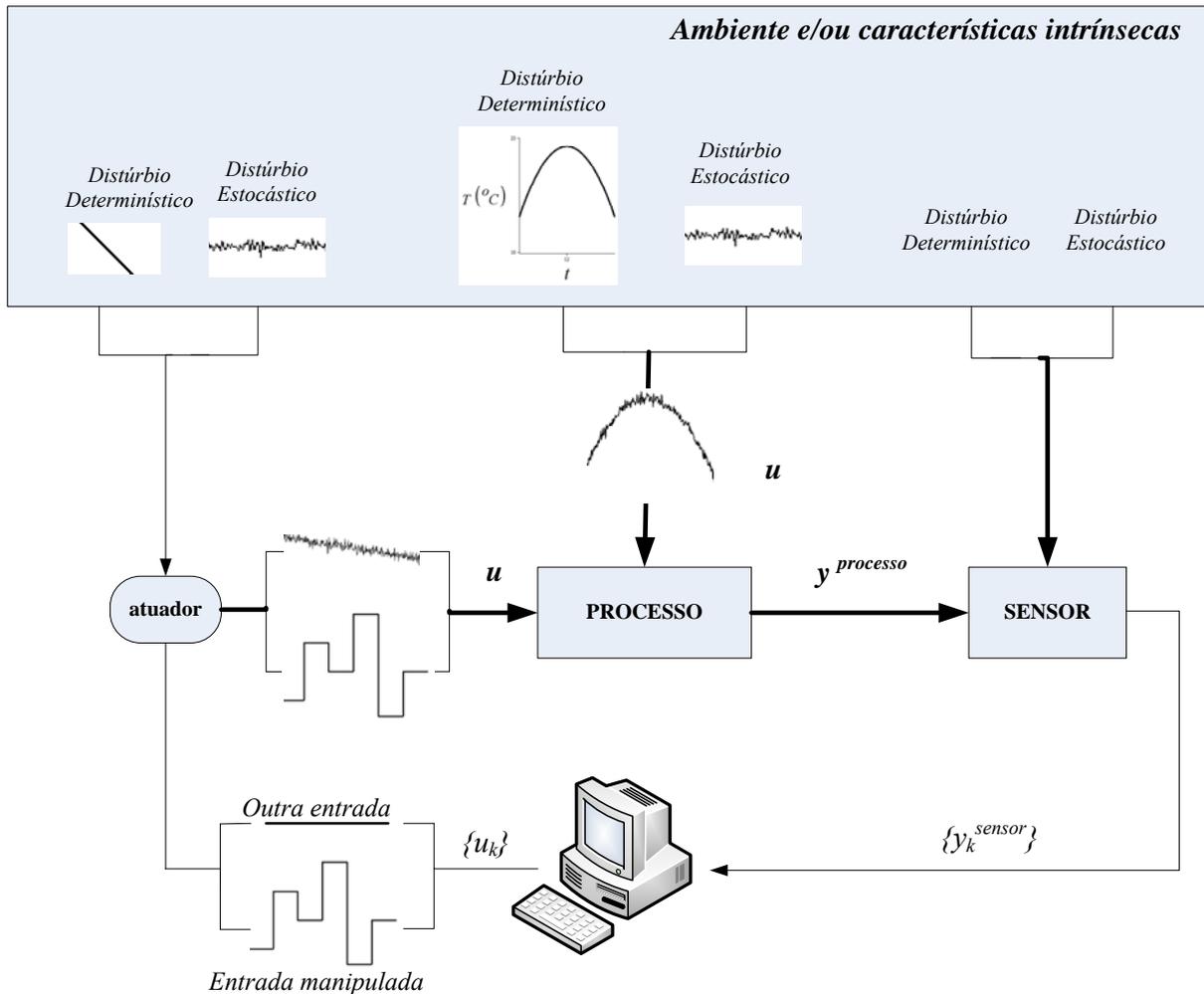
### 3.1.3 Motivação para o emprego do estimador de estados

Baseado no que foi exposto até aqui, podemos concluir que ocorrendo determinismo nos distúrbios o ideal é que se estabeleçam estruturas para eles e sejam identificados simultaneamente com o modelo fundamental. Entretanto, na prática, por diversas razões tais como: o não acesso a medição de um distúrbio, o seu desconhecimento, sua complexidade, etc., um distúrbio determinístico pode ser parcialmente ou nem mesmo caracterizado. Para casos como este, o emprego de um estimador de estados durante a estimação dos parâmetros pode ser uma interessante alternativa. Pois, o estimador de estados ao considerar a possibilidade de distúrbios no processo resulta em uma melhor estimativa para os estados.

Ao introduzirmos o assunto a respeito do uso de um estimador de estados, nesta seção 3.1.3 abordaremos as equações básicas constituintes do estimador. Na seção 3.2 abordamos o funcionamento do estimador propriamente dito, com enfoque no filtro de Kalman estendido com restrições (CEKF), que é a ferramenta usada no presente trabalho. E por fim, na seção 3.3 detalhamos a proposta do presente trabalho, explicando como o estimador de estados participa na identificação, sendo a metodologia desenvolvida nesta dissertação.

A já vista Equação 3.5 faz parte da formulação do estimador de estados. Como vimos na seção 3.1.1 sobre processos estocásticos, a rigor a variável  $w$  representa os distúrbios estocásticos sem qualquer determinismo. Entretanto, na prática, o distúrbio  $w$  pode incorporar mais informações, como: erros de modelagem na estrutura fundamental e/ou nos parâmetros ou algum distúrbio determinístico não modelado.

Até o momento enfocamos os distúrbios no processo, mas, o estimador de estados tem a vantagem de considerar também os distúrbios no sensor. Pois, a identificação de processos controlados e amostrados por computador (que constitui o estudo do presente trabalho) deve considerar que o sistema é formado basicamente por quatro itens: processo, computador, atuador e sensor, como se pode ver na Figura 3.6.



**Figura 3.6:** Esquematização de processo controlado e amostrado pelo computador.

Assim, como base do estimador de estados, os modelos para o processo e o sensor podem ser vistos na Tabela 3.3. Os distúrbios de medição  $v$  no sensor são equivalentes aos distúrbios do processo, a única diferença é que ao invés de atuarem no processo atuam nos sensores. Pois, nos sensores também observamos fenômenos físicos e/ou químicos. Por exemplo: um termopar, com o termopar sempre trabalhamos com uma temperatura de referência para poder determinar a outra, perturbações na temperatura de referência causadas pelo ambiente consiste em um tipo de distúrbio de medição, outro exemplo é a corrosão ocorrida ao longo do tempo nos metais constituintes do termopar. Claro que estes problemas relacionados a termopares são significativamente reduzidos com o simples uso de conversores e junta fria direto na cabeça do sensor e com posterior transmissão através de um sinal de 4-20 mA. Mas independentemente do aprimoramento da qualidade do medidor, qualquer sensor apresenta um certo grau de ruído de medição que pode ser originário da conversão ou do processo de transmissão do sinal. Este ruído de medição será representado aqui pela letra  $v$ .

**Tabela 3.3:** Modelos para o processo e o sensor

Processo	Sensor
$\dot{x} = f(x, u; \theta) + w$ $y^{processo} = g(x; \theta)$	$y^{sensor} = g(x; \theta) + v$
Como iremos identificar os parâmetros, esta relação será explicitada por:	
$g(x; \theta) = g[x(\dots, w); \theta]$	

### 3.2 Estimadores de estados

Para obtermos os estados considerando os distúrbios calculamos:

$$x = x_0 + \int_{t_0}^t [f(x, u; \theta) + w] dt \quad (3.6)$$

A variável  $w$ , que representa o distúrbio na taxa de  $x$ , se for gaussiano com média nula poderá ser colocada para fora da integral passando a representar o distúrbio no processo resultante:

$$x = x_0 + \int_{t_0}^t f(x, u; \theta) dt + w \quad (3.7)$$

Nos casos em que o distúrbio  $w$  não seja gaussiano, pode-se simplesmente considerar que o  $w$  da Equação 3.7 é resultado da integração do  $w$  da equação anterior.

Os dois primeiros termos da Equação 3.8 são uma simulação do processo como se não houvesse distúrbios.

$$x = x_0 + \underbrace{\int_{t_0}^t f(x, u; \theta) dt}_{x^{sim}} + w \quad (3.8)$$

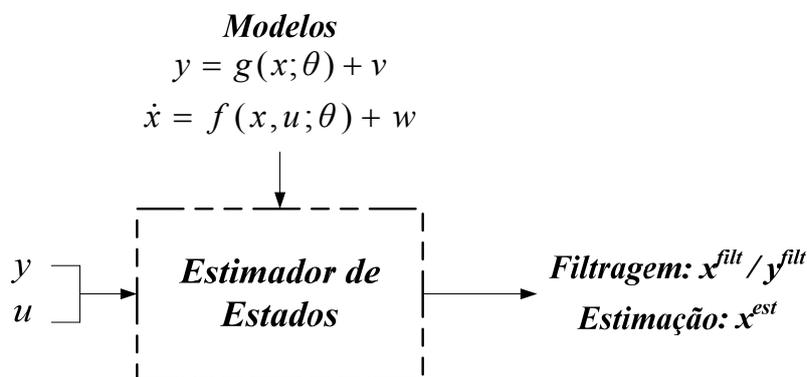
$$x = x^{sim} + w$$

Assim, para aquele estado que é calculado considerando os distúrbios adotamos a nomenclatura de estado corrigido:  $x^{cor} = x^{sim} + w$ .

Além de poder levar melhor em conta o efeito dos distúrbios não medidos, os estimadores de estados nos permitem estimar os estados que não são medidos diretamente,

desde que eles sejam observáveis. Desta forma, o resultado de um estimador de estados ( $x^{cor}$ ) pode ser uma filtragem ( $x^{filt}$ ) ou uma estimação ( $x^{est}$ ), no primeiro caso o estado é medido diretamente e no segundo não.

Lembrando que para um estimador de estados os distúrbios  $w$  e  $v$  podem incorporar mais informações além dos distúrbios nas entradas, como: erros de modelagem na estrutura fundamental e/ou nos parâmetros, o papel de um estimador de estados nada mais é do que uma reconciliação de dados entre os modelos e os dados disponíveis. Desta forma, um estimador de estados é alimentado pela as variáveis medidas, as entradas conhecidas e os modelos matemáticos, na Figura 3.7 é apresentado o esquema de funcionamento de um estimador de estados.



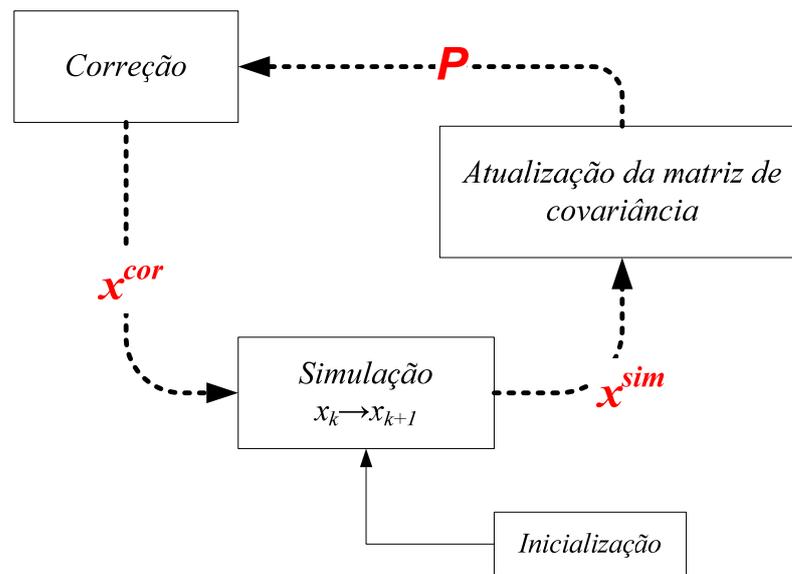
**Figura 3.7:** Esquema de funcionamento de um estimador de estados.

Quando estamos controlando um processo on-line o objetivo da reconciliação de dados é capturar e quantificar distúrbios no processo a fim de se fazer as devidas correções. No entanto, no presente trabalho o objetivo é contrário, de forma que as confianças nos modelos e nos dados são ajustadas a se obter dados corrigidos ( $x^{cor}$ ) removendo o efeito dos distúrbios nas entradas. Assim, com os dados melhorados, podemos partir para estimação dos parâmetros, onde, o corrompimento dos mesmos será diminuído devido a prévia reconciliação. Embora o ideal seja que se disponha de um modelo específico, com este método trabalhando unicamente com o modelo fundamental ainda é possível obter resultados satisfatórios.

A operação de um estimador de estados pode ser descrita basicamente em três etapas fundamentais:

- **Predição:** simulação do processo através do modelo dinâmico como se não houvessem os distúrbios  $w$  e  $v$ ;
- **Atualização de P:** Atualização da matriz de covariância  $P$  do estimador;
- **Correção:** realimentação da discrepância entre a predição e a medição. Nesta etapa os distúrbios  $w$  e  $v$  são distribuídos entre os diversos estados. Isso é feito mediante o resultado de um problema de otimização ou pelo ganho de Kalman;

Na Figura 3.8 apresenta-se o loop realizado pelo estimador de estados:



**Figura 3.8:** Loop realizado pelo estimador de estados.

Os métodos discutidos na presente seção (EKF, CEKF e MHE) são baseados em um problema de otimização de mínimos quadrados resultante de uma formulação probabilística. Veja (SALAU, 2009; TONEL, 2008) para uma descrição mais detalhada da origem do problema de mínimos quadrados. O cálculo da estimação de estados pode ser visto como uma busca por um ótimo da função densidade probabilística com respeito a uma norma adequada (JAZWINSKI, 1970).

Para sistemas lineares é possível formular um algoritmo de solução recursiva para o método dos mínimos quadrados que permita um cálculo rápido do próximo estado estimado em função do último calculado. Esta formulação recursiva foi primeiro proposta por Kalman (1960). O método de espaço-estados de Kalman promoveu um intenso interesse na teoria do filtro, por superar as suposições estacionárias restritivas da teoria de Wiener-Kolmogorov do filtro linear, abrindo novas perspectivas. Na seqüência, surgiu o filtro de Kalman estendido (EKF), que é a técnica mais utilizada para problemas de estimação não-lineares. As formulações usam linearização local para estender o espaço do filtro de Kalman. As vantagens das formulações do EKF são: simplicidade, recursividade dos algoritmos e baixa carga computacional. Estas formulações são adequadas para aplicação em tempo real em escala industrial (BECERRA *et al.*, 2001). Porém formulações recursivas não possibilitam a incorporação de restrições adicionais, as quais permitiriam descrever um conhecimento prévio do sistema, como a restrição do estado a uma certa região ou as limitações físicas influenciadas pelos distúrbios.

Já a estimação baseada no horizonte móvel (MHE), possui a vantagem de lidar com restrições. O horizonte móvel tenta preservar informações passadas pelo uso de uma janela de informações que desliza sobre as medidas. O estado é estimado através de um horizonte de  $N$  medidas de saída mais recentes. A desvantagem deste método consiste na alta carga computacional para grandes horizontes de estimação.

A formulação do CEKF é basicamente a formulação do MHE com horizonte de estimação igual a zero ( $N=0$ ). Isso é similar ao filtro de Kalman Estendido, que também possui horizonte de estimação igual a zero. Ou seja, o CEKF combina a vantagem de se trabalhar com restrições do MHE e a vantagem da redução de cálculos computacionais devido ao horizonte igual a zero como no EKF.

Marcon et al. (2002) testaram duas versões do Filtro de Kalman, o EKF e o CEKF. O objetivo foi averiguar qual delas apresentava maior robustez na predição. Para isso foi utilizado como estudo de caso um modelo de quatro tanques, que apresenta característica de inversão de comportamento, migrando da fase mínima para fase não mínima. As formulações do filtro foram testadas em distintas condições: para um modelo do processo com elevada incerteza e com condições iniciais distintas da verdadeira. Os resultados mostraram que o CEKF é mais robusto que o EKF. Tanto que só foi possível obter estimações do EKF com interferências no algoritmo, forçando as matrizes de covariância a ficarem em valores aceitáveis. Isso foi devido principalmente ao fato das restrições não poderem ser consideradas diretamente durante a resolução do EKF.

Como se pôde ver, foi feita uma abordagem geral sobre os métodos de estimação, para maiores detalhes sugere-se Tonel (2008), Salau (2007) e Simon (2006), na seqüência focaremos em uma das formulações do CEKF (as variações ocorrem no cálculo da matriz de covariância):

1º) Predição

$$x_k^{sim} = x_{k-1}^{cor} + \int_{t_{k-1}}^{t_k} f(x, u, \theta) dt \quad (3.9)$$

2º) Atualização da matriz de covariância

Como já mencionado, existem diferentes formulações para a atualização da matriz de covariância, ainda divididas em duas classes: contínuas e discretas. Salau et al. (2007) chegaram a conclusão de que uma das melhores formulações para atualizar a matriz de covariância na sua versão discreta é dada por:

$$P_k = Q_k + \varphi_k P_{k-1} \varphi_k^T - (\varphi_k P_{k-1} H_k^T)(H_k P_{k-1} H_k^T + R)^{-1} (H_k P_{k-1} \varphi_k^T)$$

Onde,

$$H_k = \left. \frac{\partial g(x)}{\partial x} \right|_{x=x^{sim}}$$

$$F_k = \left. \frac{\partial f(x, u)}{\partial x} \right|_{x=x^{sim}}$$

$$\varphi_k = e^{F_k(t_k - t_{k-1})}$$

(3.10)

As matrizes  $Q$  e  $R$  são parâmetros de sintonia para reconciliar o modelo com as medidas de processo. A matriz  $Q$  é a matriz de covariância dos distúrbios  $w$  e, a  $R$  dos distúrbios  $v$ . Usualmente, a matriz  $R$  é igualada a matriz  $V_v$  já vista na seção 2.2.

### 3º) Correção

Nesta última etapa os distúrbios estimados  $\hat{w}_k$  são acrescentados aos estados preditos anteriormente.

$$x_k^{cor} = x_k^{sim} + \hat{w}_k \quad (3.11)$$

Sendo  $\hat{w}_k$  resultado do seguinte problema de otimização.

$$\min_{w,v} \Psi = w_k^T P_k^{-1} w_k + v_k^T R^{-1} v_k \quad (3.12)$$

Sujeito as seguintes restrições:

$$\begin{aligned} x_{\min} - x_k^{sim} &\leq w_k \leq x_{\max} - x_k^{sim} \\ y_k &= g(x_k^{sim}) + Hw_k + v_k \\ y_{\min} - g(x_k^{sim}) &\leq Hw_k + v_k \leq y_{\max} - g(x_k^{sim}) \end{aligned} \quad (3.13)$$

## 3.3 Proposta de aplicação do CEKF e estimação dos parâmetros

A proposta de aplicação do CEKF com a estimação de parâmetros do presente trabalho consiste em um processo iterativo que se pode separar em etapas repetitivas até se chegar ao resultado desejável. Cada etapa é separada em duas sub-etapas: ajuste da sintonia do estimador e ajuste da função objetivo.

A primeira etapa não possui o ajuste da sintonia do CEKF, as demais etapas são equivalentes a segunda. Portanto, a seguir descrevemos mais detalhadamente as 1ª e 2ª etapas.

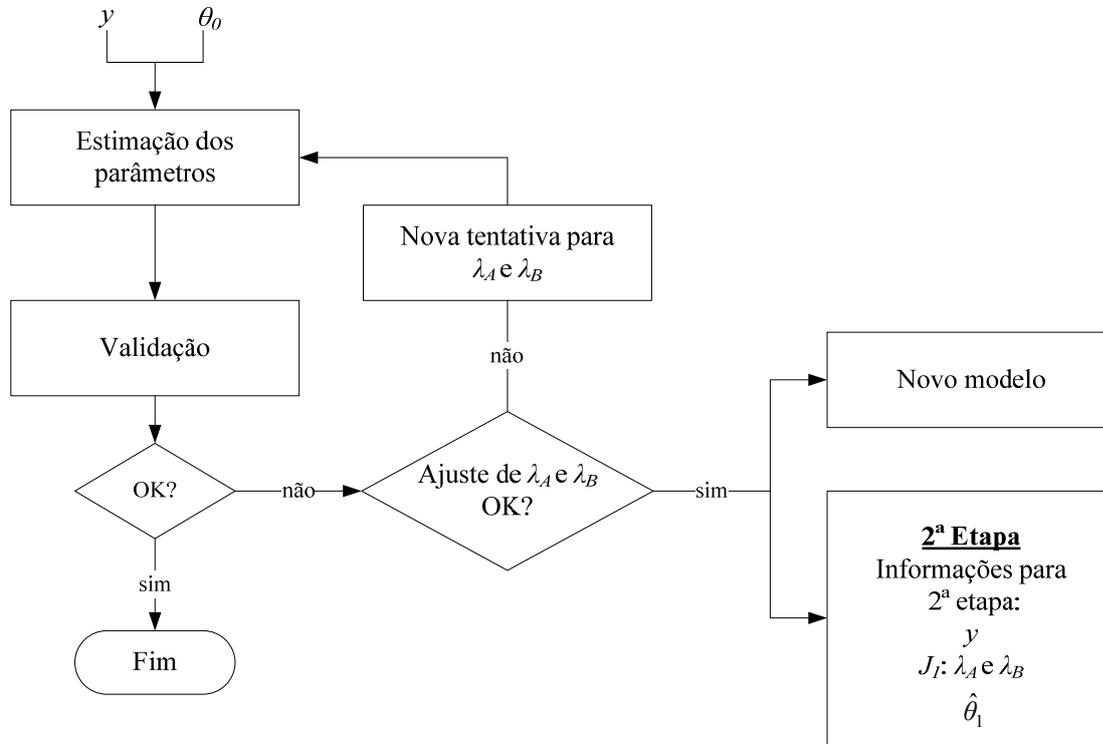
### 1ª Etapa

→ *Ajuste da função objetivo*

A função objetivo genérica que se propõe é:

$$J = \lambda_A \left( \sum_{i=1}^{ne} \sum_{j=1}^{ny} [y_{ij} - g(x_i^{sim}; \theta)]^2 \right) + \lambda_B \left( \sum_{i=1}^{ne} \sum_{j=1}^{ny} \left[ \frac{dy_{ij}}{dt} - \frac{dg(x_i^{sim}; \theta)}{dt} \right]^2 \right) \quad (3.14)$$

Desta forma, aqui o objetivo é obter um bom balanço entre as duas parcelas da função objetivo através dos pesos  $\lambda_A$  e  $\lambda_B$ . Na Figura 3.9 é apresentado o fluxograma para o ajuste da função objetivo:



**Figura 3.9:** Fluxograma para o ajuste da função objetivo.

Como se pode observar na Figura 3.9, os pesos  $\lambda_A$  e  $\lambda_B$  não são estimados junto com os parâmetros, mas sim estipulados previamente de forma iterativa. Realizada uma estimação o resultado desta e das anteriores servem de informações para determinar uma nova tentativa para  $\lambda_A$  e  $\lambda_B$ . São feitas tentativas para  $\lambda_A$  e  $\lambda_B$  até que o pesquisador julgue ser suficiente segundo a qualidade dos resultados obtidos. Portanto, o ajuste consiste em construir uma boa quantidade de possibilidades de funções objetivo e ao final selecionar qual resultou em um melhor resultado. É interessante a construção de uma tabela para se poder visualizar os resultados parciais e assim selecionar o melhor.

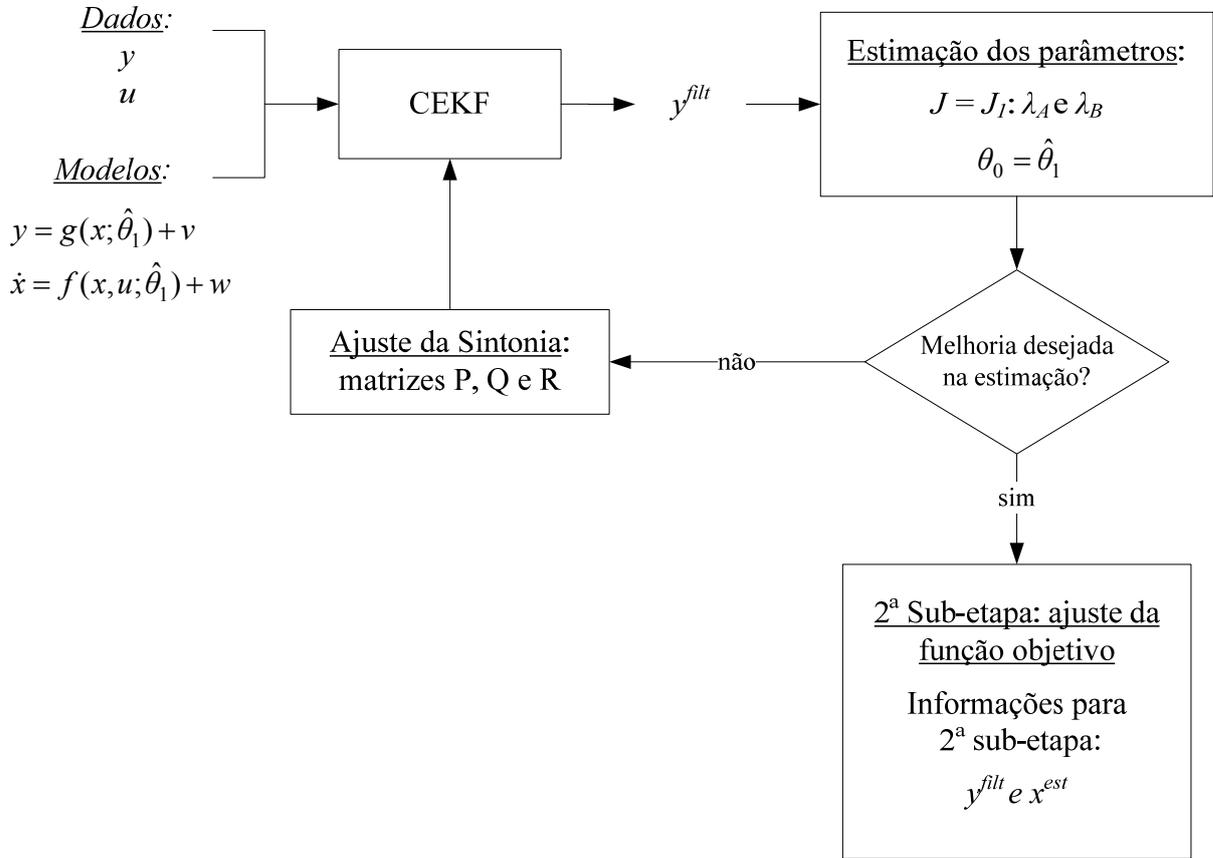
## 2ª Etapa

→ *Ajuste da sintonia do CEKF*

Na Figura 3.10 é apresentado o fluxograma elaborado para o ajuste da sintonia, observe nesta figura que:

- O modelo que alimenta o filtro foi obtido na etapa anterior ( $\hat{\theta}_1$ );
- Na etapa da estimação os dados são melhorados, pois no lugar de  $y$  usamos  $y^{filt}$ ;

- Foi fixada a função objetivo tida como melhor na etapa anterior:  $J = J_I: \lambda_A$  e  $\lambda_B$ .



**Figura 3.10:** Fluxograma elaborado para o ajuste da sintonia do CEKF.

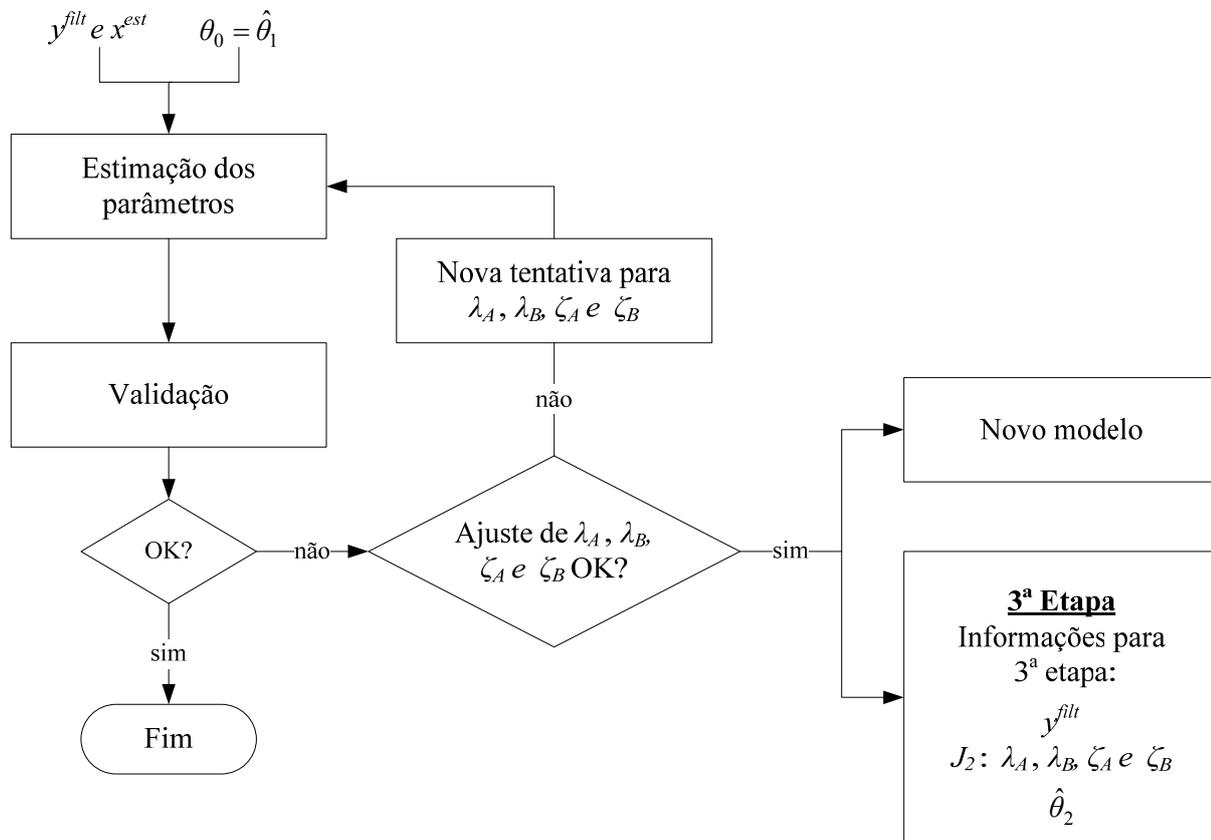
O ajuste da sintonia do CEKF  similar ao j abordado para a funo objetivo. So testadas de forma iterativa um boa quantidade de sintonias e ao final  selecionada a melhor.

→ *Ajuste da funo objetivo*

Com o resultado da aplicaco do CEKF no momento anterior, se pode incrementar a funo objetivo com os estados estimados ( $x^{est}$ ) alm dos filtrados ( $y^{filt}$ ), sendo assim a funo objetivo genrica passa a ser:

$$\begin{aligned}
 J = & \lambda_A \left( \sum_{i=1}^{ne} \sum_{j=1}^{ny} [y_{ij}^{filt} - g(x_i^{sim}; \theta)]^2 \right) + \lambda_B \left( \sum_{i=1}^{ne} \sum_{j=1}^{ny} \left[ \frac{dy_{ij}^{filt}}{dt} - \frac{dg(x_i^{sim}; \theta)}{dt} \right]^2 \right) + \\
 & \zeta_A \left( \sum_{i=1}^{ne} \sum_{j=1}^{nx^{est}} [x_{ij}^{est} - x_{ij}^{sim}]^2 \right) + \zeta_B \left( \sum_{i=1}^{ne} \sum_{j=1}^{nx^{est}} \left[ \frac{dx_{ij}^{est}}{dt} - \frac{dx_{ij}^{sim}}{dt} \right]^2 \right)
 \end{aligned} \tag{3.15}$$

O fluxograma para o ajuste da função objetivo é essencialmente igual ao já apresentado na primeira etapa (Figura 3.9). No entanto, na Figura 3.11 é apresentado novamente com as pequenas diferenças devido à mudança na função objetivo.



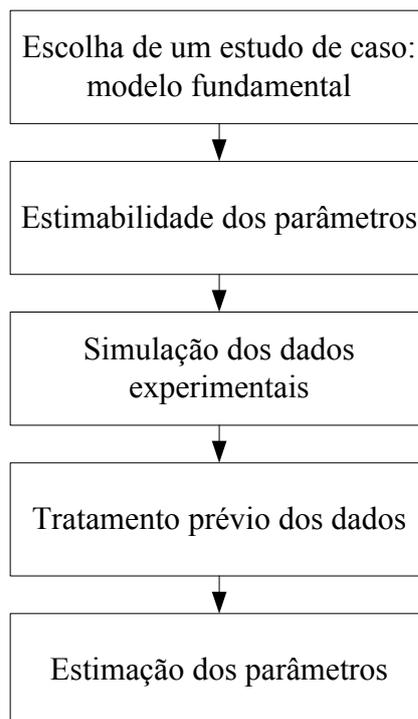
**Tabela 3.11:** Fluxograma para o ajuste da função objetivo atualizado com sua nova formulação.

Salienta-se que à medida que vai se avançando nas etapas a sub-etapa de ajuste da função objetivo vai se tornando cada vez mais rápida, até o momento em que se chega a um ajuste constante. Isto porque a cada etapa que se passa mais possibilidades para  $\lambda_A$ ,  $\lambda_B$ ,  $\zeta_A$  e  $\zeta_B$  vão sendo eliminadas.

# Capítulo 4

## Metodologia

Na Figura 4.1 apresentamos a seqüência de passos realizados no presente trabalho:



**Figura 4.1:** Seqüência de passos realizados no trabalho.

A metodologia descrita no presente trabalho foi aplicada tendo-se como instrumento de simulação e resolução de problemas matemáticos o software Maple<sup>TM</sup>, com exceção da etapa do tratamento prévio dos dados onde se utilizou o MATLAB<sup>®</sup>.

Como já mencionado no Capítulo 1, tomou-se como estratégia desenvolver o trabalho partindo de um conhecimento prévio dos valores ideais dos parâmetros ( $\theta^*$ ), o que facilita a análise dos resultados. Assim, a partir de  $\theta^*$  e do modelo fundamental foram simulados os dados experimentais, gerados com o acréscimo de distúrbios determinísticos e estocásticos.

A metodologia empregada para a estimação dos parâmetros em conjunto com o CEKF já foi abordada na seção 3.3. Na presente dissertação trabalhamos com a estrutura do modelo correta desde o início, desta forma, a vertente “novo modelo” da metodologia vista na seção 3.3 não foi necessária na presente identificação.

O CEKF implementado corresponde ao descrito na seção 3.2.

Na primeira etapa do processo de estimação dos parâmetros testamos três diferentes chutes iniciais  $\theta_0$  a fim de verificar a convergência. Na seção 4.3 abordaremos em maior profundidade o método de busca pelo ótimo implementado: de Gauss (ou Gauss Newton).

E por fim, as derivadas dos dados experimentais, que são necessárias na função objetivo (ver seção 3.3), foram calculadas utilizando o método de diferenças finitas centrais, dada por:

$$\frac{dy}{dt} \approx \frac{y_{t+h} - y_{t-h}}{2h} \quad (4.1)$$

## 4.1 Estimabilidade dos parâmetros

A análise da estimabilidade dos parâmetros é muito interessante porque nos permite determinar um ponto de operação no qual os parâmetros são mais sensíveis a estimação e, também, identificar quais são os parâmetros possíveis de estimação.

Tendo em mãos o modelo fundamental do estudo de caso, aqui implementamos o que já vimos na seção 2.2 sobre estimabilidade dos parâmetros.

Lembramos que para a análise da estimabilidade precisamos da matriz de sensibilidade  $S$ , que se refere às saídas  $y = g(x; \theta)$  em relação aos parâmetros. Como no presente trabalho já conhecemos  $\theta^*$ , acaba sendo mais vantajoso calcular  $S$  em  $\theta^*$ :

$$S = \left. \frac{\partial g(x; \theta)}{\partial \theta} \right|_{\theta=\theta^*} \quad (4.2)$$

Lembrando que em um caso real se usaria a estimativa inicial disponível  $\theta_0$  na Equação 4.2. Vemos também, que a Equação 4.2 é totalmente dependente da sensibilidade dos estados em relação aos parâmetros,  $S_x$ :

$$S_x = \left. \frac{\partial x}{\partial \theta} \right|_{\theta=\theta^*} \quad (4.3)$$

que no caso de processos dinâmicos pode ser obtida através da integração da seguinte equação diferencial obtida pela aplicação da regra da cadeia:

$$\frac{d}{dt} \left( \frac{\partial x}{\partial \theta} \right) = \frac{\partial f}{\partial x} \cdot \frac{\partial x}{\partial \theta} + \frac{\partial f}{\partial \theta} \quad (4.4)$$

Assim, a matriz  $S$  genérica, considerando todos instantes de tempo, necessária para o cálculo dos índices de importância  $\delta$  e colinearidade  $\gamma$  é apresentada a seguir:

$$S = \begin{bmatrix} S_{1,1}|_{t_0} & \cdots & S_{1,n\theta}|_{t_0} \\ \vdots & & \vdots \\ S_{ny,1}|_{t_0} & \cdots & S_{ny,n\theta}|_{t_0} \\ \vdots & & \vdots \\ S_{1,1}|_{t_n} & \cdots & S_{1,n\theta}|_{t_n} \\ \vdots & & \vdots \\ S_{ny,1}|_{t_n} & \cdots & S_{ny,n\theta}|_{t_n} \end{bmatrix} \quad (4.5)$$

Esta foi a forma que se implementou a determinação da matriz  $S$  para todos os pontos no tempo. A integração da Equação 4.4 é feita simultaneamente com as equações dos estados:  $\dot{x} = f(x, u; \theta)$ . Usualmente, as sensibilidades iniciais são  $S(0) = 0$ , isto porque os estados iniciais não são dependentes dos parâmetros, pois se pressupõe que os valores de  $x(0)$  são provenientes de informações “externas”, ou seja, de medição, de conhecimento histórico, etc. Por exemplo, consideramos um produto que possui a concentração  $C$  de seu princípio ativo e, com o passar do tempo essa concentração vai diminuindo, o fabricante alega que ao comprarmos o produto a concentração é  $C_0$ . Nesse caso  $C_0$  é um estado inicial histórico, se quisermos estimar  $C$  em qualquer outro instante de tempo a não ser  $t_0$  podemos nos valer de um modelo matemático, aí sim os demais estados serão dependentes e, portanto, sensíveis aos parâmetros.

A montagem do sistema da Equação 4.4 com as equações dos estados  $\dot{x} = f(x, u; \theta)$  foi implementada no Maple<sup>TM</sup>. Todo o sistema foi resolvido com um método numérico já implementado no Maple<sup>TM</sup>: *lsode* (*Livermore Stiff ODE solver*). As referências do método *lsode* estão em Hindmarsh et al. (1983). Um método alternativo ao que foi elaborado é o *CVODES* solver, desenvolvido para solução de equações diferenciais simultaneamente com a sensibilidade dos estados, uma vantagem é que precisamos apenas informar o sistema  $\dot{x} = f(x, u; \theta)$ , o resto fica por conta do solver, outra vantagem é a redução no tempo computacional inerente ao solver. O *CVODES* é uma parte da família de software Sundials (Suite of Nonlinear and Differential/Algebraic Equation Solvers), ao utilizar este software é necessário que se crie uma interface com a ferramenta matemática em uso, que no caso do

presente trabalho é o Maple<sup>TM</sup>. Mencionamos aqui a respeito do *CVODES* com o propósito para trabalhos futuros, ao aperfeiçoar o que foi desenvolvido no presente trabalho.

## 4.2 Tratamento prévio dos dados experimentais

Esta etapa significa filtrar os dados a fim de reduzir de forma significativa o efeito dos distúrbios estocásticos. Ressaltamos que o cálculo das derivadas experimentais é altamente sensível a distúrbios estocásticos quando não tratados previamente.

O clássico filtro *low pass* geralmente retorna bons resultados, sendo assim, utilizamos a função *Idfilt* já implementada no MATLAB<sup>®</sup> (Ljung, 1999).

Outras técnicas podem e devem ser testadas e utilizadas, mas como o foco deste trabalho não foi este, limitou-se apenas a aplicação desta técnica.

## 4.3 Implementação do método de Gauss

Na seção 2.2.2 vimos que a estimação dos parâmetros é baseada na função erro:  $e = y - g(x; \theta)$ , seja de forma direta ou indireta. Logo, de forma mais abrangente podemos dizer que a estimação dos parâmetros é baseada na diferença:  $\varepsilon = y_\varepsilon - g_\varepsilon(x; \theta)$ . Tomamos como exemplo a Equação 4.6 para ilustrar esta abrangência:

$$\min_{\theta} \sum_{i=1}^{ne} \sum_{j=1}^{ny} [y_{ij} - g(x_i^{sim}; \theta)]^2 + \sum_{i=1}^{ne} \sum_{j=1}^{ny} \left[ \frac{dy_{ij}}{dt} - \frac{dg(x_i^{sim}; \theta)}{dt} \right]^2 \quad (4.6)$$

Onde,

$$y_\varepsilon = \begin{bmatrix} y \\ \frac{dy}{dt} \end{bmatrix} \quad g_\varepsilon = \begin{bmatrix} g \\ \frac{dg}{dt} \end{bmatrix} \quad (4.7)$$

A adoção dessa nomenclatura foi tomada para que as formulações matemáticas que seguem sejam aplicáveis para todos os casos possíveis de funções objetivo, como vimos na seção 3.3. Na Equação 4.6 tomada como exemplo temos  $\lambda_A = 1$  e  $\lambda_B = 1$ , escolhemos esse caso por ser didaticamente mais interessante ao deduzir o gradiente ( $q$ ) e a aproximação da hessiana ( $N$ ).

Sabemos que os métodos de busca analíticos requerem o gradiente ( $q$ ) e a matriz hessiana ( $H$ ). Considerando a função objetivo com apenas uma saída para simplificar:

$$J(\theta) = \sum_{i=1}^{ne} \varepsilon_i^2 \quad (4.8)$$

O gradiente fica:

$$q_j = \frac{\partial J}{\partial \theta_j} = 2 \sum_{i=1}^{ne} \varepsilon \frac{\partial \varepsilon}{\partial \theta_j} = -2 \sum_{i=1}^{ne} \varepsilon \frac{\partial g_\varepsilon}{\partial \theta_j} \quad j = 1, 2, \dots, n\theta \quad (4.9)$$

A matriz Hessiana:

$$H_{jl} = \frac{\partial^2 J}{\partial \theta_j \partial \theta_l} = -2 \sum_{i=1}^{ne} \varepsilon \frac{\partial^2 g_\varepsilon}{\partial \theta_j \partial \theta_l} + 2 \sum_{i=1}^{ne} \left( \frac{\partial g_\varepsilon}{\partial \theta_j} \right) \left( \frac{\partial g_\varepsilon}{\partial \theta_l} \right) \quad j = 1, 2, \dots, n\theta \quad l = 1, 2, \dots, n\theta \quad (4.10)$$

Utilizando o método de Gauss, negligenciamos o primeiro termo e temos  $N$  como uma aproximação de  $H$ :

$$N_{jl} = 2 \sum_{i=1}^{ne} \left( \frac{\partial g_\varepsilon}{\partial \theta_j} \right) \left( \frac{\partial g_\varepsilon}{\partial \theta_l} \right) \quad j = 1, 2, \dots, n\theta \quad l = 1, 2, \dots, n\theta \quad (4.11)$$

Aqui, definimos a matriz de sensibilidade  $S_\varepsilon$ :

$$S_\varepsilon = \frac{\partial g_\varepsilon}{\partial \theta} \quad (4.12)$$

Que é dependente de  $S_x$ , determinada pela integração da mesma equação diferencial apresentada na Equação 4.4.

Passando para notação matricial temos:

$$N = 2 \sum_{i=1}^{ne} S_\varepsilon^T S_\varepsilon \quad (4.13)$$

$$q = -2 \sum_{i=1}^{ne} S_\varepsilon^T \varepsilon$$

No presente trabalho as funções objetivo formuladas são dependentes do resultado de um sistema dinâmico, portanto,  $N$ ,  $q$  e  $J$  foram computados simultaneamente com a integração do sistema, como pode ser visto na Figura 4.2.

O método iterativo para a busca de  $\hat{\theta}$  segue a seguinte formulação:

$$\theta_{i+1} = \theta_i + \rho_i \delta_i \quad (4.14)$$

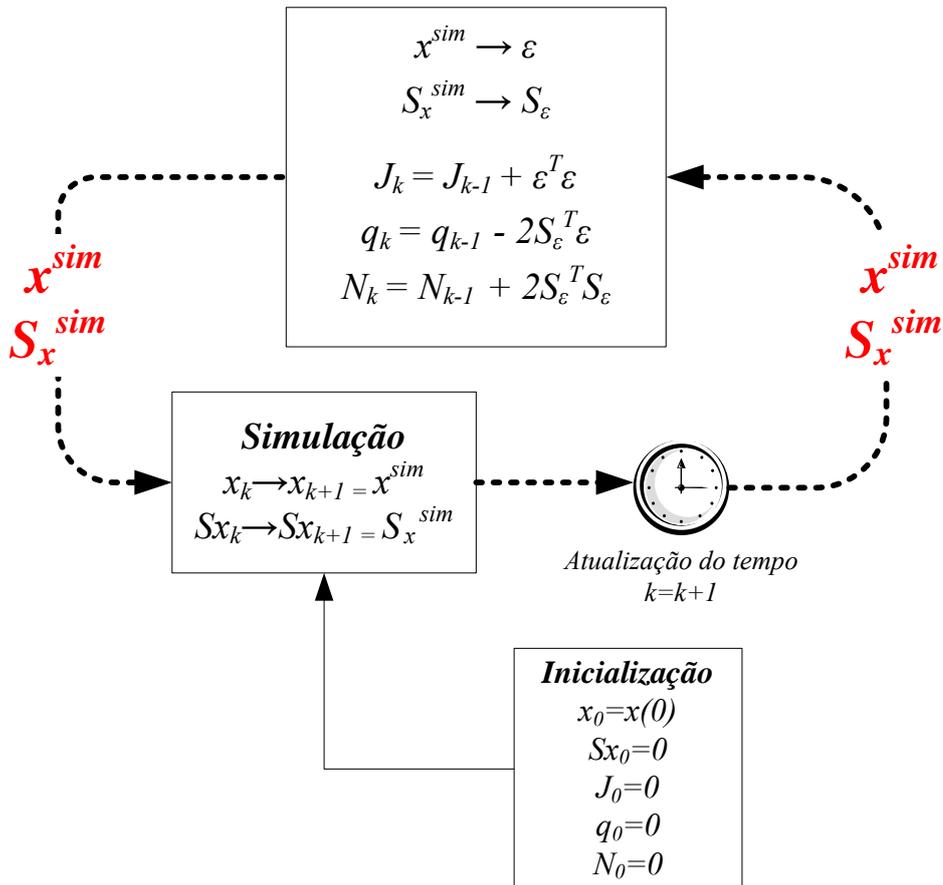
Onde,  $\delta$  é a direção de busca, definida como  $\delta = -N^{-1}q$ . E,  $\rho_i$  é o tamanho do passo, que assim como a direção de busca deve ser definido a cada iteração. Os critérios para

a determinação do  $\rho_i$  estão inseridos no algoritmo da busca pelo ótimo. Contudo, antes de abordarmos este algoritmo veremos o critério para a determinação do ótimo:

$$|\theta_{i+1} - \hat{\theta}_i| \leq E \quad (4.15)$$

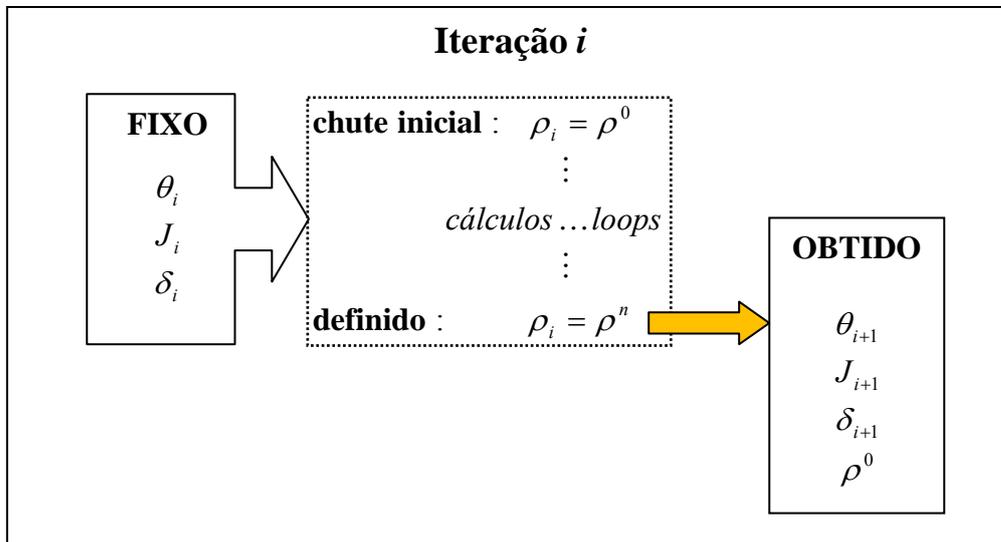
A adoção da Equação 4.15 é interessante em relação ao critério  $q \approx 0$  porque é eficiente nos casos em que se está muito perto do mínimo e o gradiente ainda é significativo. Portanto, economizam-se cálculos. Como  $E$  tomou-se a recomendação de Marquardt (1963) apud Bard (1974):  $E = 10^{-4}(\theta_i + 10^{-3})$ . Assim, podemos estabelecer  $\rho_{min}$ :

$$\rho_{min} = \min \left( \frac{E_j}{\delta_j} \right) \quad j = 1, 2, \dots, n\theta \quad (4.16)$$



**Figura 4.2:** Cálculo de  $J$ ,  $q$ , e  $N$  para o sistema dinâmico.

De forma resumida, cada iteração é processada como na Figura 4.3.



**Figura 4.3:** Resumo de cada iteração.

Como pudemos ver na Figura 4.3, a cada iteração o objetivo é determinar o tamanho do passo  $\rho_i$  a fim de que  $J_{i+1} < J_i$ . Salientamos também que o valor na função objetivo  $J_{i+1}$  passa a ser unicamente dependente de  $\rho_i$ , com isso definimos uma função  $\Psi(\rho) = J(\theta(\rho))$ , a qual terá aplicação no procedimento de busca pelo ótimo.

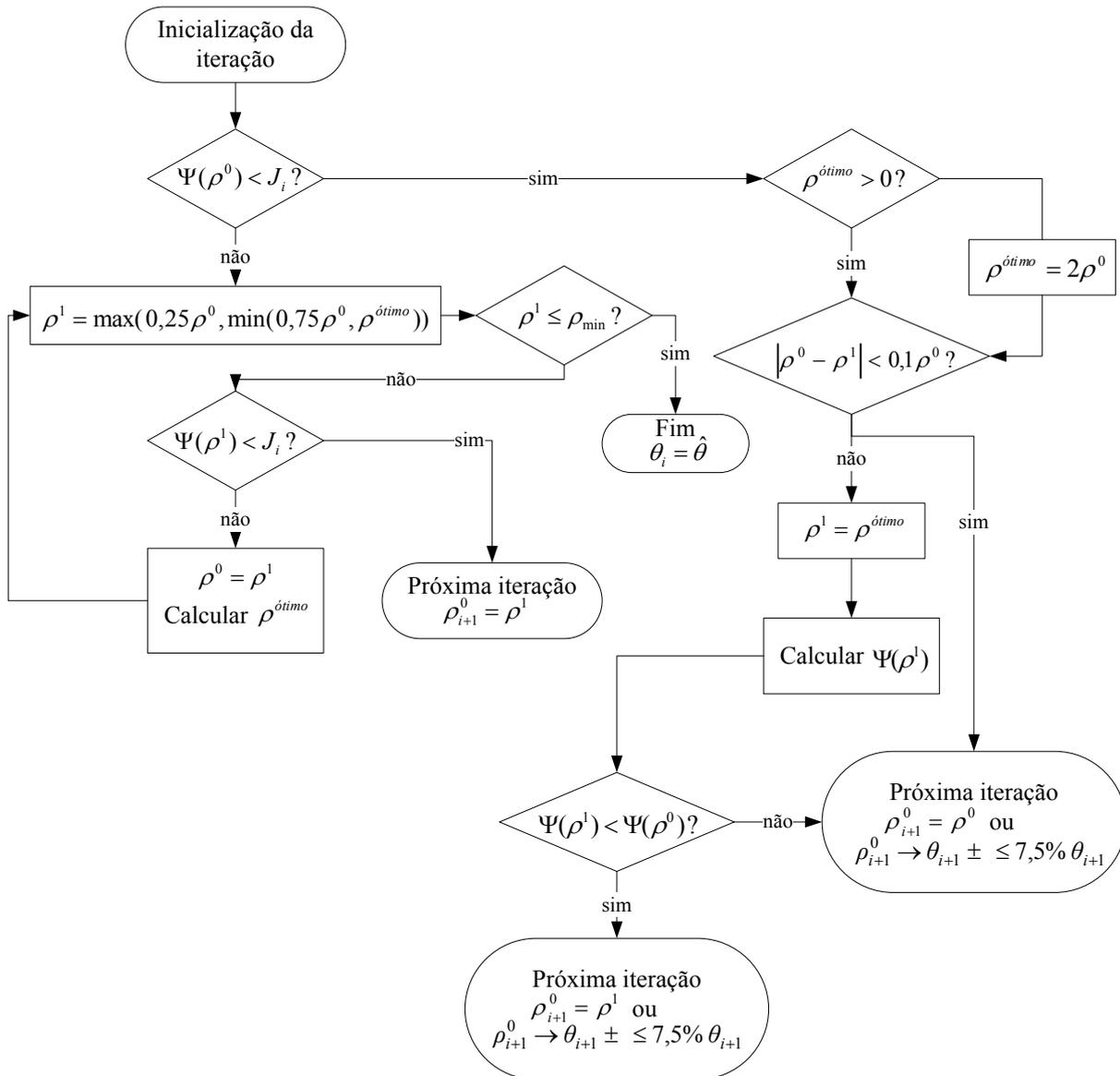
Para definir  $\rho_i$ , a interpolação – extrapolação é uma interessante técnica. Na seqüência apresentaremos a técnica vista em Bard (1974). Supondo que inicialmente  $\Psi(\rho^0) > J_i$ , o próximo passo é determinar um  $\rho_i$  menor e continuar esse processo até que chegue a um ponto aceitável:  $\Psi(\rho^n) = J_{i+1} < J_i$ . Já no caso inicial em que  $\Psi(\rho^0) < J_i$ , ainda tenta-se outro valor para  $\rho_i$  objetivando-se reduzir ainda mais a função  $\Psi$ .

Nos processos de interpolação/extrapolação o algoritmo de otimização utiliza uma função quadrática para determinar  $\rho^{\text{ótimo}}$ :

$$a + b\rho^0 + c(\rho^0)^2 = \beta$$

$$\rho^{\text{ótimo}} = \frac{b(\rho^0)^2}{2(b\rho^0 + a - \beta)} \quad (4.17)$$

Sendo,  $a = J_i$ ,  $b = q_i^T \delta_i$ ,  $\beta = \Psi(\rho^0)$ , e  $c$  é determinado por isolamento, embora ele não seja necessário em  $\rho^{\text{ótimo}}$ .



**Figura 4.4:** Fluxograma da iteração.

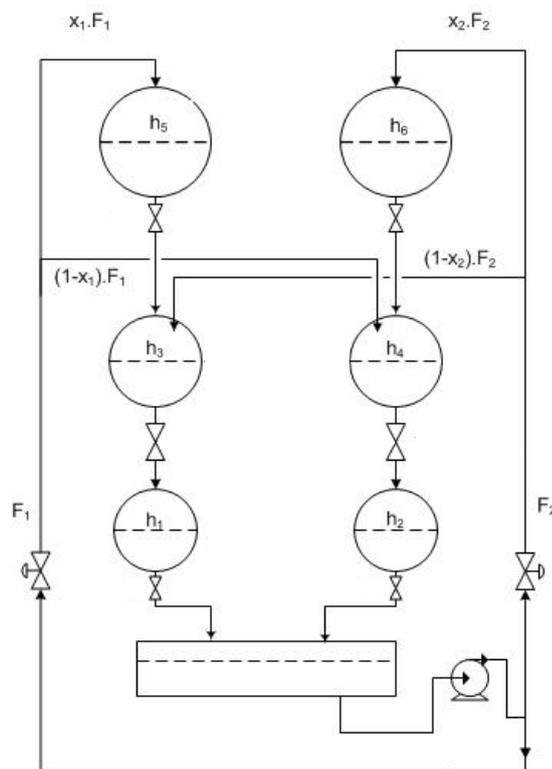
Observe no lado direito da Figura 4.4, que se refere a parte da extrapolação, onde temos nas partes “próxima iteração” a opção:  $\rho_{i+1}^0 \rightarrow \theta_{i+1} \pm \leq 7,5\% \theta_{i+1}$ . Na verdade, o acréscimo de  $\pm \leq 7,5\%$  no valor dos parâmetros é feito somente até o momento que ocorre a primeira interpolação. À partir do momento em que é necessário fazer uma interpolação significa que já estamos consideravelmente próximos do mínimo e  $\pm \leq 7,5\%$  se torna muito alto para o próximo chute para os parâmetros. Após a primeira interpolação a idéia é assumir como  $\rho^0$  para a próxima iteração o  $\rho$  atualmente determinado.

# Capítulo 5

## Resultados

### 5.1 Descrição do estudo de caso

A planta estudada consiste em uma planta real de seis tanques esféricos interligados, situada no Departamento de Engenharia Química da UFRGS. O desenho esquemático da planta é ilustrado na Figura 5.1:



**Figura 5.1:** Desenho esquemático da planta estudada.

O modelo fundamental adotado para a planta pode ser visto na Equação 5.1, onde os parâmetros são os coeficientes de descarga ( $CD$ ).

*Estados :*

$$\begin{aligned}
 A_6(h_6) \frac{dh_6}{dt} &= x_2 \cdot F_2 - CD_6 \sqrt{h_6} \\
 A_5(h_5) \frac{dh_5}{dt} &= x_1 \cdot F_1 - CD_5 \sqrt{h_5} \\
 A_4(h_4) \frac{dh_4}{dt} &= (1 - x_1) \cdot F_1 + CD_6 \sqrt{h_6} - CD_4 \sqrt{h_4} \\
 A_3(h_3) \frac{dh_3}{dt} &= (1 - x_2) \cdot F_2 + CD_5 \sqrt{h_5} - CD_3 \sqrt{h_3} \\
 A_2(h_2) \frac{dh_2}{dt} &= CD_4 \sqrt{h_4} - CD_2 \sqrt{h_2} \\
 A_1(h_1) \frac{dh_1}{dt} &= CD_3 \sqrt{h_3} - CD_1 \sqrt{h_1}
 \end{aligned} \tag{5.1}$$

*Saídas*

$$y_1 = h_1$$

$$y_2 = h_2$$

Os valores de todos os diâmetros foram adotados como  $D = 22,5$  cm, estando em concordância com a planta real, da mesma forma, na Tabela 5.1 apresentam-se os valores dos coeficientes de descarga  $CD$  ( $\theta^*$ ).

**Tabela 5.1:** Valores dos Coeficientes de Descarga  $CD$  ( $\theta^*$ ).

$CD_1$	17,800 cm <sup>2,5</sup> /s	$CD_4$	15,990 cm <sup>2,5</sup> /s
$CD_2$	19,100 cm <sup>2,5</sup> /s	$CD_5$	13,650 cm <sup>2,5</sup> /s
$CD_3$	15,955 cm <sup>2,5</sup> /s	$CD_6$	13,650 cm <sup>2,5</sup> /s

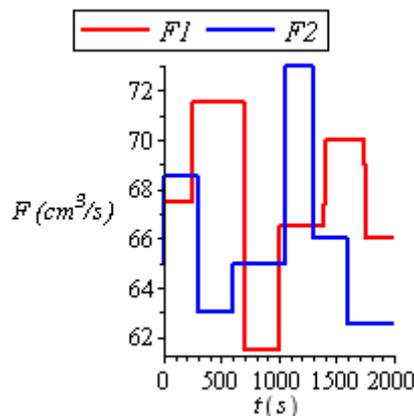
## 5.2 Estimabilidade dos parâmetros

Para se determinar as condições operacionais foram estudadas nove situações conforme as frações  $x_1$  e  $x_2$  e as vazões iniciais  $F1_0$  e  $F2_0$ . Todas as simulações foram feitas partindo-se do estado estacionário. As frações  $x_1$  e  $x_2$  foram assumidas constantes e degraus foram aplicados nas vazões  $F1$  e  $F2$ . Na tabela 5.2 são caracterizados os pontos de operação analisados.

**Tabela 5.2:** Características de cada Condição de Operação.

Ponto	$x$	$F_0$ (cm <sup>3</sup> /s)
1	$x1 = x2 = 0,25$	$F1_0 = F2_0 = 25$
2	$x1 = x2 = 0,48$	
3	$x1 = x2 = 0,75$	
4	$x1 = x2 = 0,25$	$F1_0 = F2_0 = 45$
5	$x1 = x2 = 0,48$	
6	$x1 = x2 = 0,75$	
7	$x1 = x2 = 0,25$	$F1_0 = F2_0 = 65$
8	$x1 = x2 = 0,48$	
9	$x1 = x2 = 0,75$	

Os degraus aplicados em  $F1$  e  $F2$  em todas as simulações tiveram o mesmo perfil (amplitudes e instantes de tempos ocorridos), mudando apenas as vazões iniciais  $F_0$ . Na Figura 5.2 temos a representação gráfica dos degraus tendo-se como vazão inicial  $F_0 = 65$  cm<sup>3</sup>/s.

**Figura 5.2:** Representação gráfica dos degraus aplicados ( $F_0 = 65$  cm<sup>3</sup>/s).

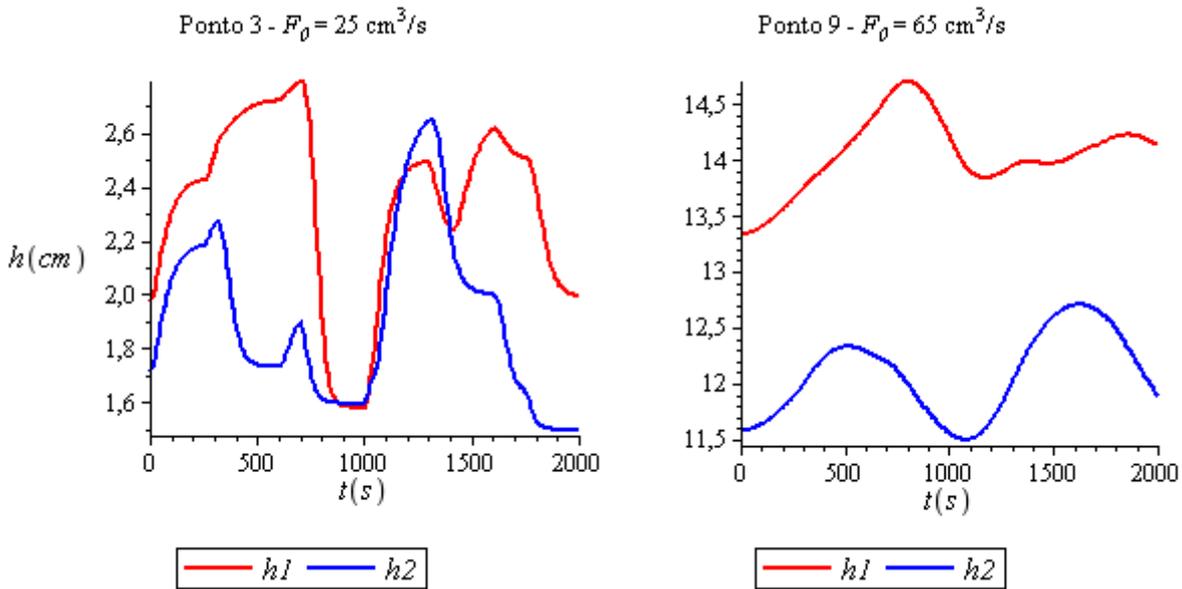
Para os nove pontos de operação calculamos o índice de importância, isso com o objetivo de selecionar o ponto mais estratégico. Os resultados são apresentados na Tabela 5.3.

**Tabela 5.3:** Índice de importância dos parâmetros  $\delta$ .

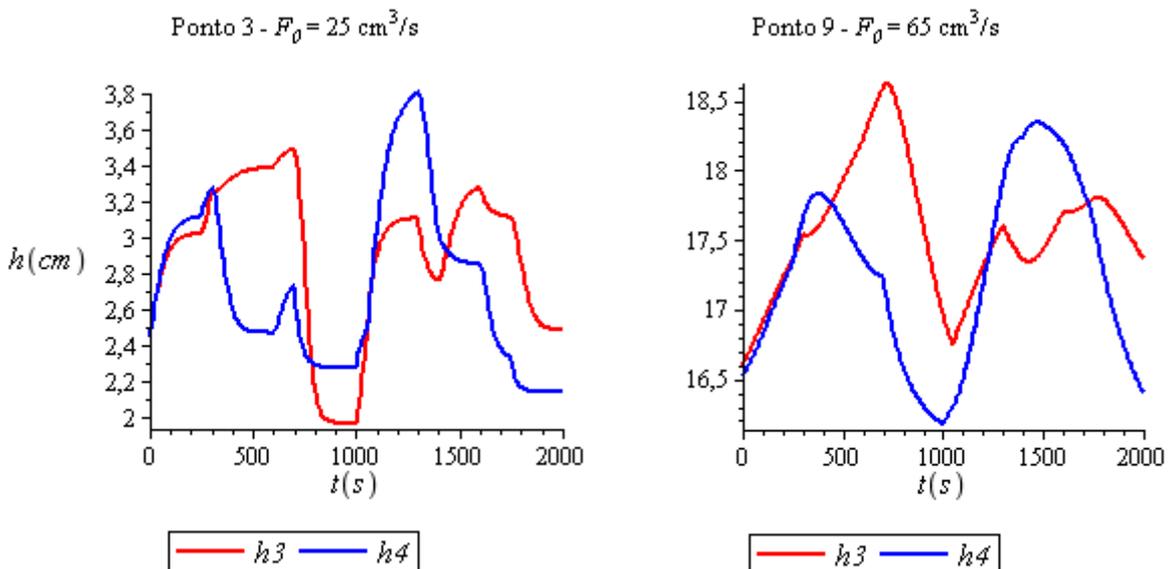
Ponto	Parâmetros					
	$CD1$	$CD2$	$CD3$	$CD4$	$CD5$	$CD6$
1	0,176	0,151	0,024	0,021	0,000	0,000
2	0,180	0,147	0,022	0,019	0,004	0,004
3	0,186	0,143	0,022	0,019	0,020	0,017
4	0,524	0,437	0,091	0,084	0,002	0,001
5	0,528	0,433	0,089	0,082	0,020	0,016
6	0,531	0,428	0,088	0,080	0,070	0,057
7	1,041	0,861	0,167	0,142	<b>0,007</b>	<b>0,006</b>
8	1,047	0,855	0,162	0,143	<b>0,068</b>	<b>0,055</b>
9	1,052	0,846	0,161	0,145	<b>0,170</b>	<b>0,145</b>

Ao analisarmos a tabela 5.3 concluímos que os parâmetros são mais sensíveis quando se opera em altas vazões (Pontos 7 ao 9). Em especial no Ponto 9, os parâmetros  $CD5$  e  $CD6$  são significativamente mais sensíveis em relação aos Pontos 7 e 8, lembrando que no Ponto 9 a alimentação é preferencialmente por cima ( $x = 0,75$ ).

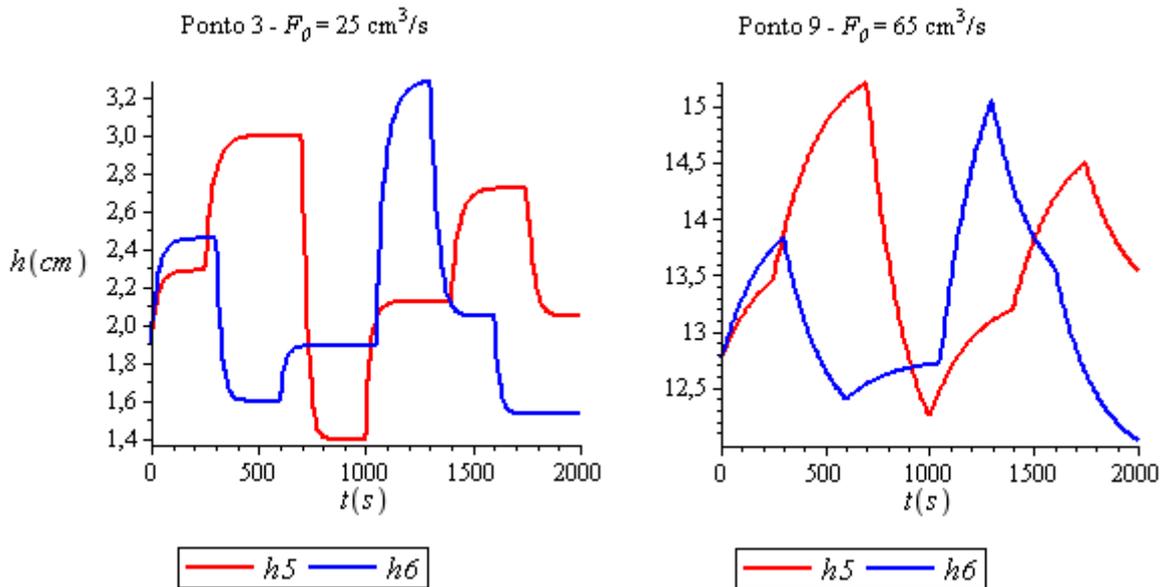
A fim observar graficamente a dinâmica dos estados e correlacionar com a diferença nas sensibilidades (vazão baixa x vazão alta), compararemos os resultados dos Pontos 3 e 9, os quais possuem as mesmas frações ( $x_1 = x_2 = 0,75$ ) e  $F_0$  diferentes: 25 e 65  $\text{cm}^3/\text{s}$ , respectivamente.



**Figura 5.3:** Comparação entre as dinâmicas dos Pontos 3 e 9 para os estados  $h1$  e  $h2$ .

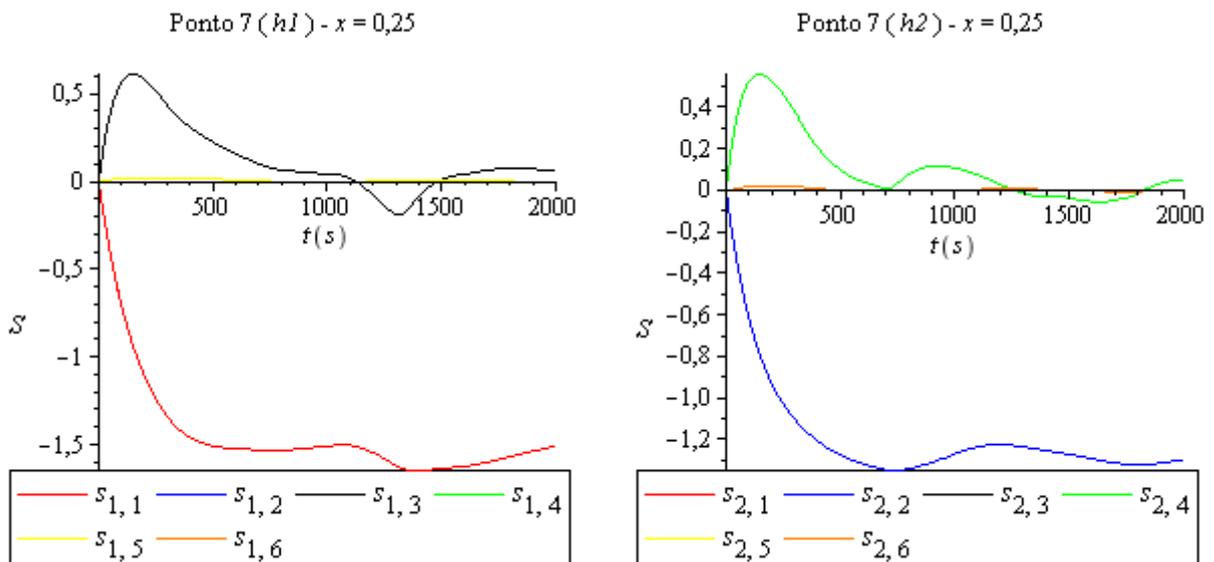


**Figura 5.4:** Comparação entre as dinâmicas dos Pontos 3 e 9 para os estados  $h3$  e  $h4$ .

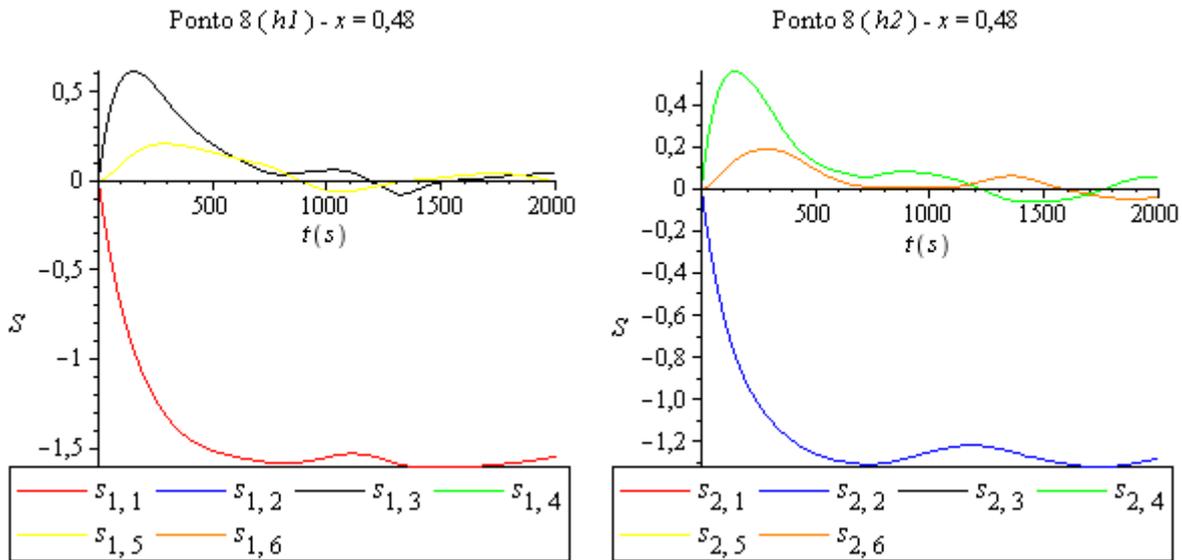


**Figura 5.5:** Comparação entre as dinâmicas dos Pontos 3 e 9 para os estados  $h5$  e  $h6$ .

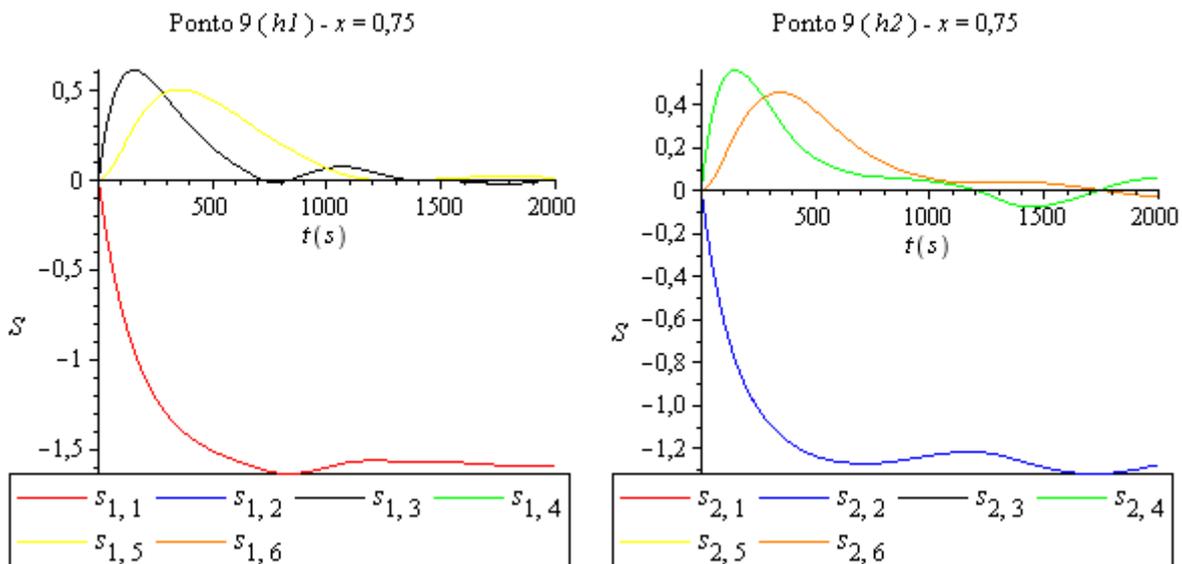
Concluimos aqui que nas vazões mais altas a dinâmica é mais lenta, isso se deve a configuração esférica dos tanques, que apresentam dinâmicas mais rápidas quando o nível dos tanques é baixo (i.e.,  $h < 4$ ). Nas dinâmicas mais lentas os parâmetros são mais sensíveis a estimação. Agora, nas Figuras 5.6 a 5.8 apresentam-se os resultados das sensibilidades dinâmicas para os Pontos 7 ao 9. Analisando esses três pontos perceberemos o efeito da mudança nas frações  $x1$  e  $x2$ , sendo que tais pontos possuem a mesma vazão  $F_0 = 65 \text{ cm}^3/\text{s}$ . Iremos acompanhar como as sensibilidades de  $CD5$  e  $CD6$  são mais proeminentes à medida que se aumenta as frações  $x1$  e  $x2$ , estando em conformidade com os resultados verificados na Tabela 5.3. Como suporte para análise das figuras, lembramos que como saídas  $y$  foram assumidos os estados  $h1$  e  $h2$ , logo, como resultados para  $h1$  obtemos as sensibilidades:  $s_{1,1} \dots s_{1,6}$ , sendo o primeiro índice referente à saída e o segundo ao parâmetro ( $CD1 \dots CD6$ ). De forma análoga, para  $h2$  obtemos as sensibilidades:  $s_{2,1} \dots s_{2,6}$ .



**Figura 5.6:** Sensibilidades dinâmicas para o Ponto 7.



**Figura 5.7:** Sensibilidades dinâmicas para o Ponto 8.

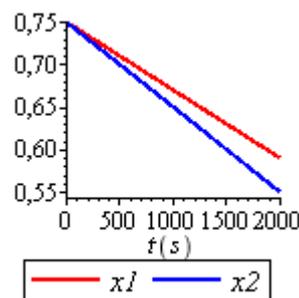


**Figura 5.8:** Sensibilidades dinâmicas para o Ponto 9.

Com base nos resultados apresentados, selecionou-se o Ponto 9 para seguir a análise. Primeiramente, formou-se um grupo composto por todos os parâmetros para determinar o índice de colinearidade. Como resultado obteve-se um índice de  $\gamma = 2,43$ , isso quer dizer que a modificação na resposta causada pela alteração de um determinado parâmetro  $\theta_j$  pertencente este grupo pode ser compensada com apropriadas modificações nos outros parâmetros na ordem de 41 % ( $1/\gamma_K \times 100$ ). Se tomarmos como base a recomendação de Brun et al. (2001), que apresenta valores de colinearidades superiores a 5 como sendo críticos. Tomando como base este valor, podemos concluir nessa etapa que, pelo fato de  $\gamma = 2,43$ , todos os parâmetros poderão ser estimados sem maiores problemas de colinearidade.

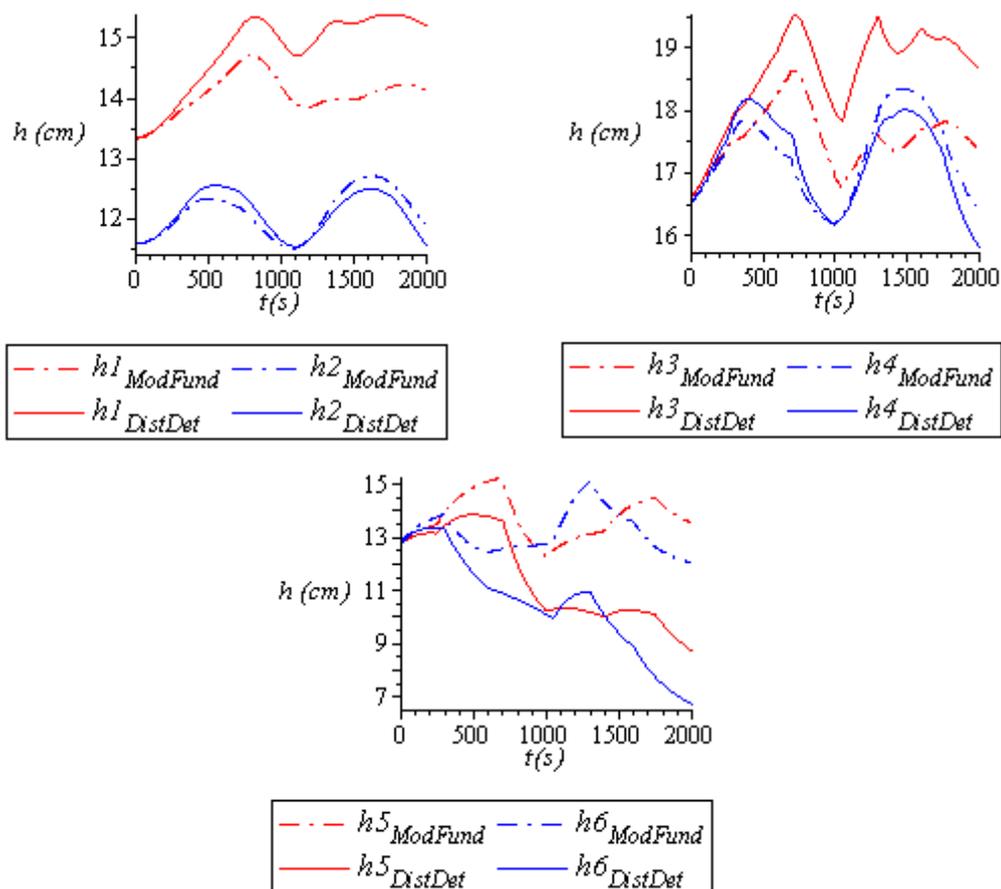
### 5.3 Gerando dados experimentais para estimação

Depois de selecionado o Ponto 9, foram gerados computacionalmente dados experimentais agregando distúrbios determinísticos e estocásticos. Na etapa da estimabilidade dos parâmetros as frações  $x_1$  e  $x_2$  foram simuladas constantes, entretanto, aqui criamos um caso hipotético que embora se deseje que estas frações sejam constantes, na planta real existiu um distúrbio determinístico desconhecido de maneira que elas caíram linearmente ao longo do tempo. Assim, as frações  $x_1$  e  $x_2$  foram geradas pela Equação 5.2 e podem ser visualizadas graficamente na Figura 5.9. Já os estados com a ocorrência do distúrbio determinístico são apresentados na Figura 5.10 em comparação aos estados sem os distúrbios.



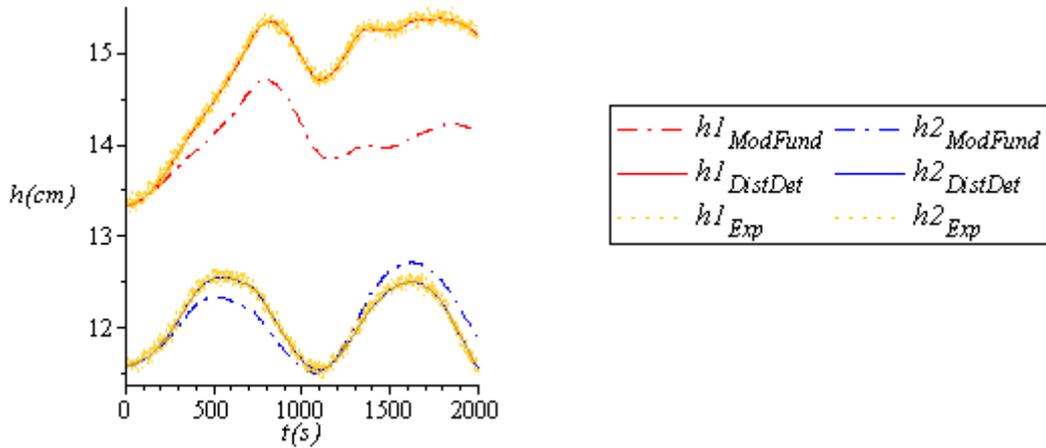
$$\begin{aligned} x_1 &= -0,00008t \\ x_2 &= -0,0001t \end{aligned} \quad (5.2)$$

**Figura 5.9:** Frações  $x_1$  e  $x_2$  com distúrbios determinísticos.



**Figura 5.10:** Comparação dos estados com os distúrbios determinísticos (*DistDest*) em relação ao modelo fundamental que é sem distúrbio (*ModFund*).

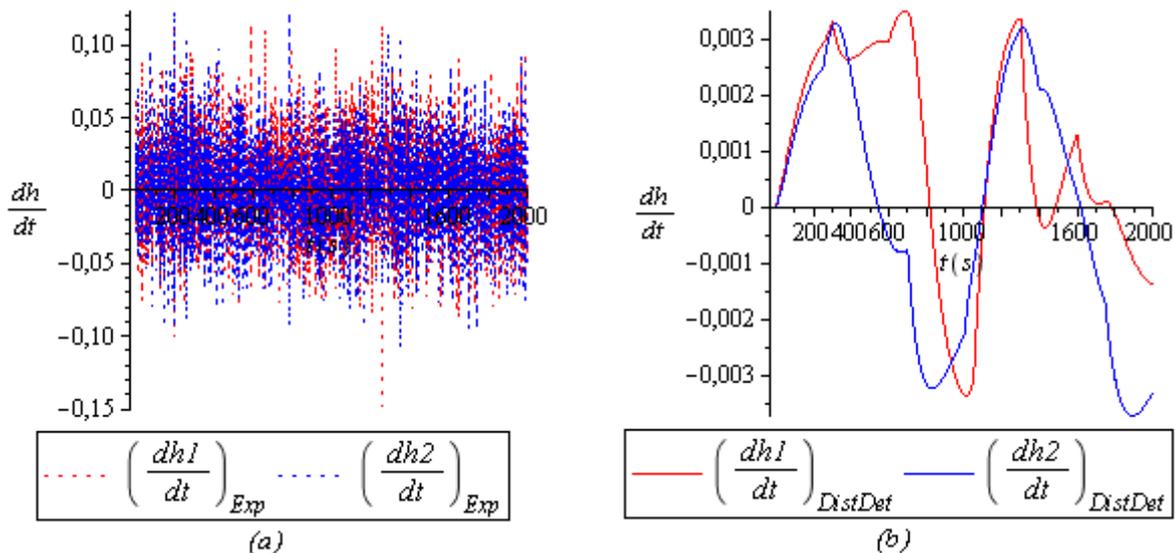
Por fim, após o acréscimo do distúrbio determinístico foi acrescentado o distúrbio estocástico, representando a combinação das incertezas no processo e nas medições. Este distúrbio é caracterizado por distribuição normal, média zero e desvio padrão 0,05. Na Figura 5.11 visualizamos as saídas  $h1$  e  $h2$  com todos os distúrbios atuantes.



**Figura 5.11:** Saídas  $h1$  e  $h2$  com todos os distúrbios atuantes:  $h1_{Exp}$  e  $h2_{Exp}$ . *Exp*: dados experimentais.

## 5.4 Tratamento prévio dos dados experimentais

A filtragem dos dados antes da estimação consiste em uma ferramenta extremamente útil para melhorar a qualidade dos resultados, principalmente no caso do presente trabalho, onde é necessário calcular as derivadas experimentais. Sem o prévio tratamento dos dados os distúrbios estocásticos corrompem de forma notável o resultado das derivadas, como podemos ver Na Figura 5.12.



**Figura 5.12:** (a) Efeito dos distúrbios estocásticos no cálculo das derivadas quando não aplicamos a filtragem. (b) Para fins de comparação: as derivadas sem os distúrbios estocásticos, ou seja, com apenas os distúrbios determinísticos atuantes.

Foram testadas duas formas para a filtragem das derivadas:

- 1) Filtragem dos dados experimentais e cálculo das derivadas;
- 2) Cálculo das derivadas e posterior filtragem.

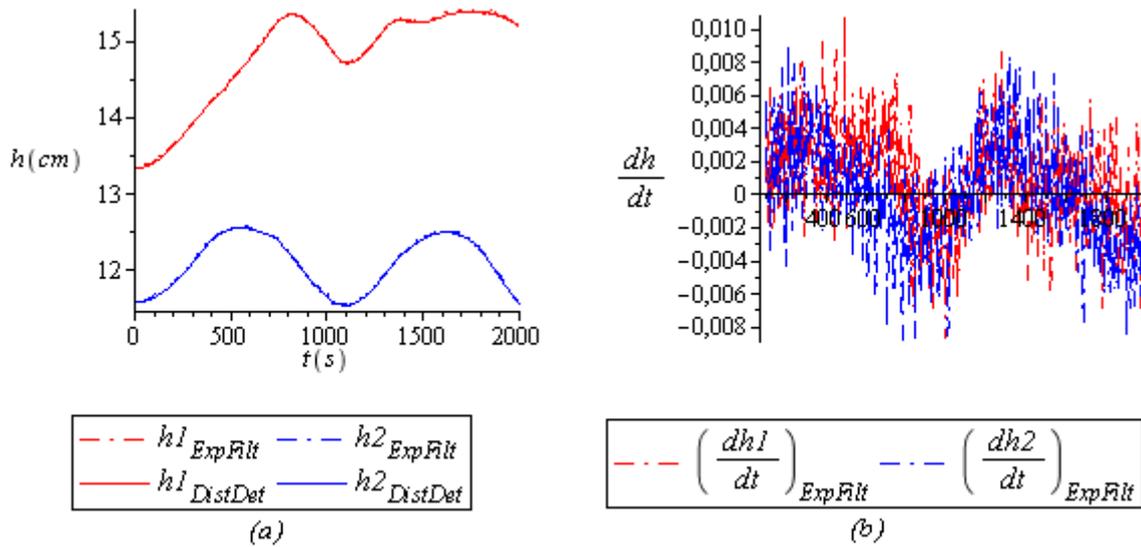
Como foi abordado na metodologia do presente trabalho, tanto os dados experimentais quanto suas derivadas foram testados em distintas formulações de função objetivo, sendo, portanto, necessário o tratamento de ambos. Aqui, se poderia obter apenas um filtro, caso o efeito do filtro nos dados experimentais repercutisse satisfatoriamente no cálculo das derivadas posteriormente, este é o primeiro método mencionado. Mas, o melhor método para as derivadas foi o segundo: derivar os dados experimentais brutos e em seguida filtrar. Assim, foram utilizados dois filtros: um para os dados experimentais e, outro para as derivadas dos dados experimentais brutos, como se pode ver na Tabela 5.4:

**Tabela 5.4:** Filtros selecionados.

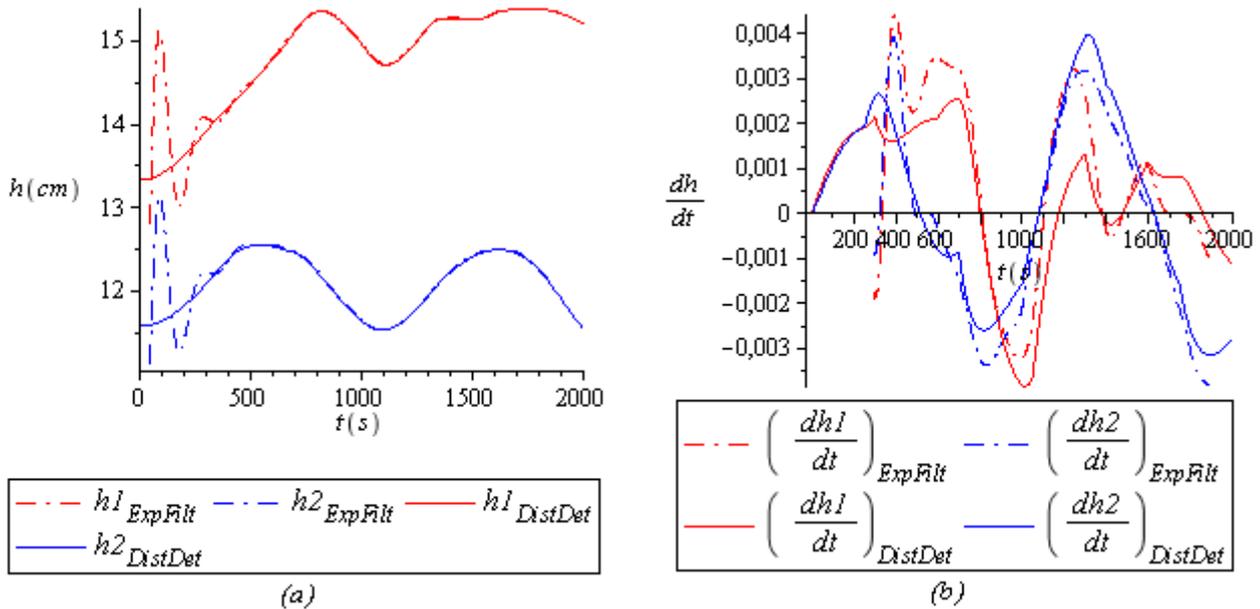
	<b>Filtro</b>	
	Ordem	Frequência de corte
(1) Dados experimentais: estados medidos	2	0,2
(2) Derivada dos dados experimentais	5	0,035

Sobre a busca do filtro para os estados medidos, foi testado um número razoável de possíveis ordens e frequência de corte. Observou-se que à medida que se aumentava a ordem, bem como se reduzia a frequência de corte, a defasagem era aumentada e a oscilação inicial também. Para se ter uma idéia, com o filtro selecionado para os estados medidos a defasagem foi corrigida com o deslocamento de 5 pontos e a oscilação inicial atingiu os 50 primeiros pontos. Agora, os dados medidos quando tratados com o filtro selecionado para as derivadas, a defasagem e a oscilação inicial já foram consideravelmente maiores: a defasagem foi corrigida com o deslocamento de 100 pontos e a oscilação inicial avançou dos 500 primeiros pontos. Na Figura 5.13 (a) apresentamos graficamente os dados experimentais tratados com o filtro correto: com a defasagem corrigida e a remoção dos 50 primeiros pontos instáveis. Já na parte (b) apresentamos as derivadas calculadas com os dados da parte (a), ou seja, o método: filtrar em seguida derivar, podemos ver que assim as derivadas não foram satisfatórias.

Na Figura 5.14 Apresentamos os dados experimentais tratados com o filtro selecionado para as derivadas. Em 5.14 (a) a defasagem está corrigida e, os 50 primeiros pontos foram removidos assim como na Figura 5.13(a) para fins comparativos, podemos ver a notável oscilação inicial em 5.14 (a). Agora, em 5.14 (b) calculamos as derivadas da mesma forma que anteriormente, ainda removemos os 300 primeiros pontos para melhor visualização em decorrência do efeito da oscilação inicial vista em (a). Aqui vimos que definitivamente o método filtrar e em seguida derivar não retornou resultados satisfatórios.

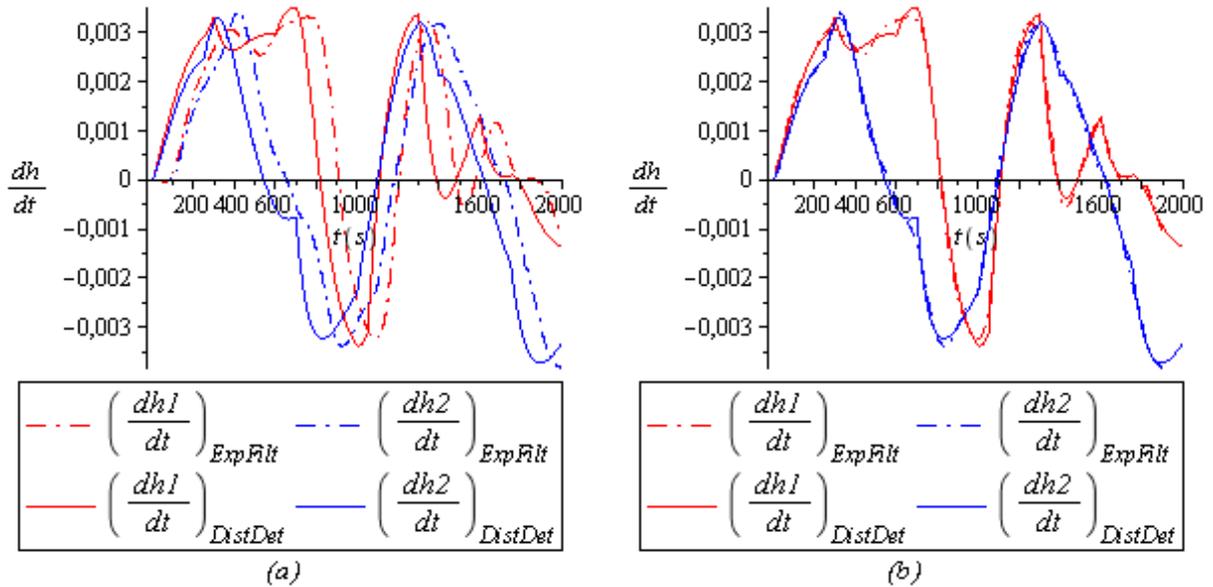


**Figura 5.13:** Filtro de ordem 2 e frequência de corte 0,2: (a) estados medidos e (b) derivadas.



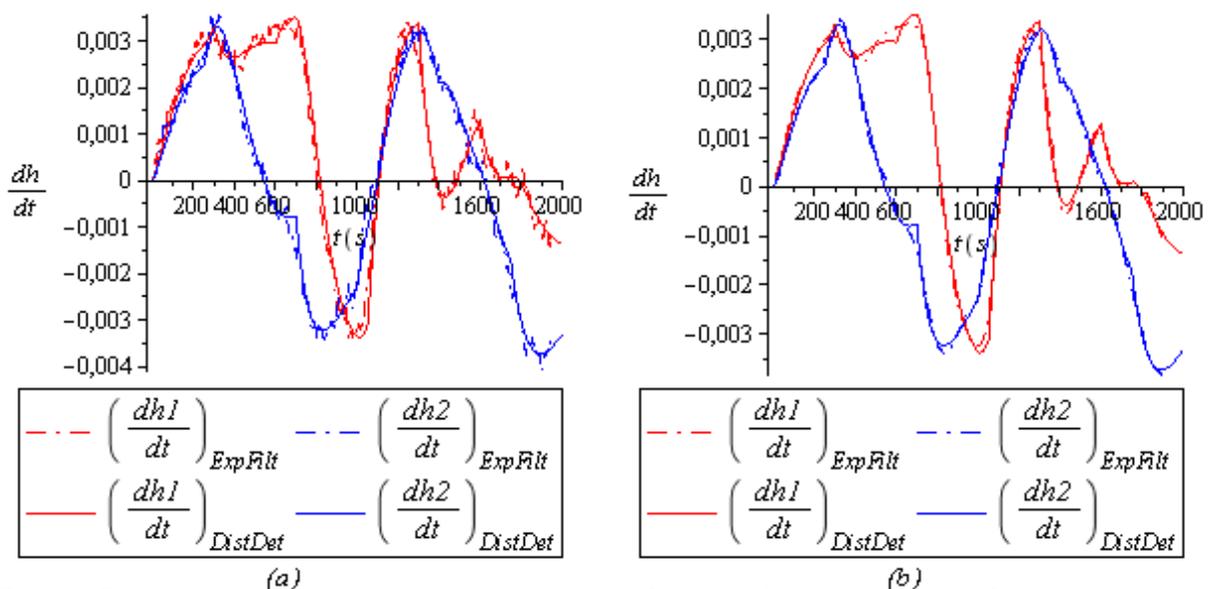
**Tabela 5.14:** Filtro de ordem 5 e frequência de corte 0,035: (a) estados medidos e (b) derivadas.

Partindo para o segundo método de filtragem das derivadas: derivar e filtrar, na Figura 5.15 podemos ver os resultados com o filtro selecionado. Neste caso também se observou uma defasagem (Figura 5.15 (a)) que foi corrigida com o deslocamento de 103 pontos (Figura 5.15 (b)). Ao contrário do caso anterior dos estados medidos, aqui não ocorreu oscilação inicial.



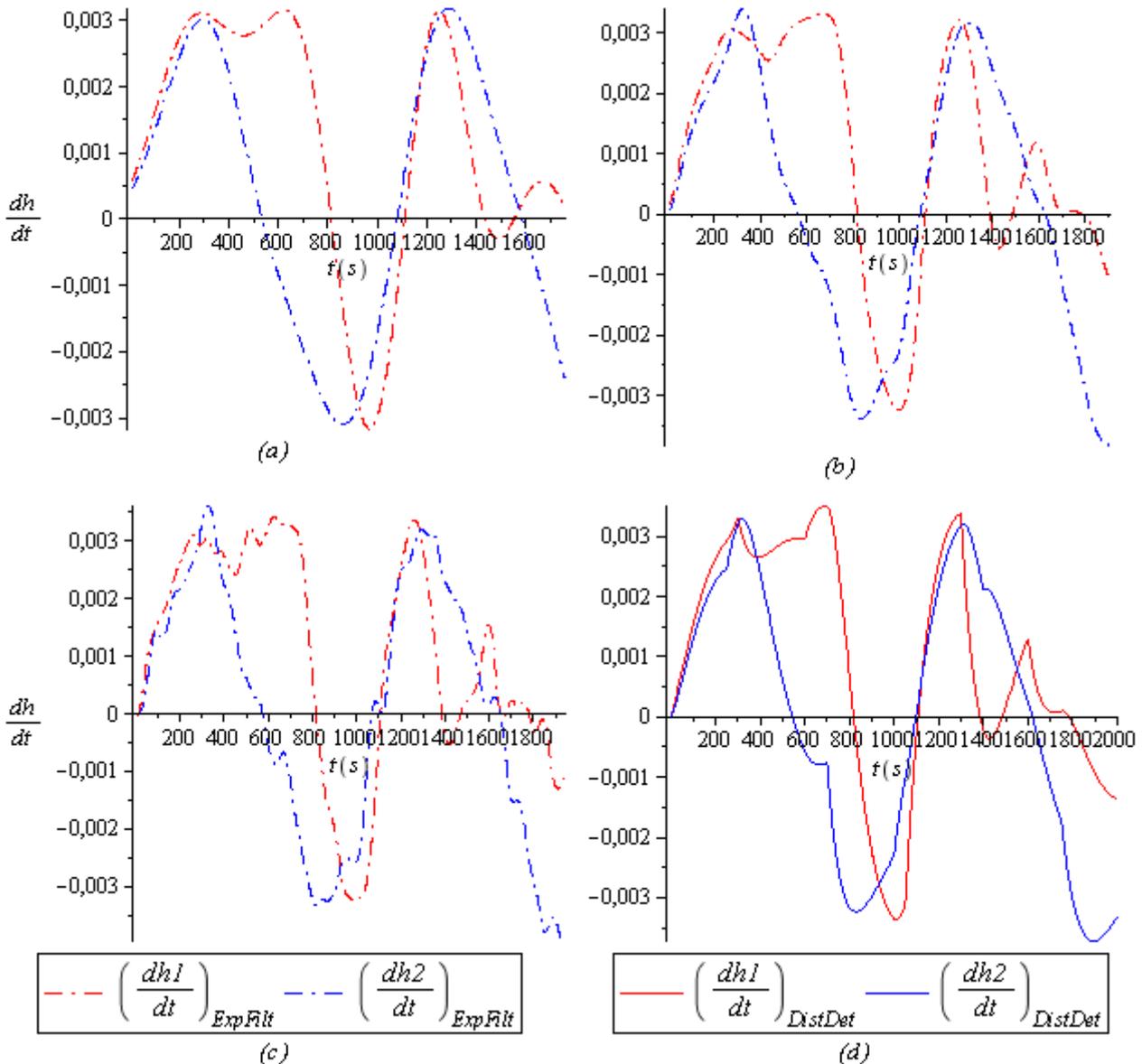
**Figura 5.15:** Derivadas experimentais filtradas (*ExpFilt*) com filtro de ordem 5 e frequência de corte 0,035. Com defasagem (a) e com defasagem corrigida (b).

Intuitivamente, nos parece que o primeiro método para filtragem das derivadas iria repercutir em um melhor resultado. No entanto, constatamos que o segundo foi melhor, talvez por que no processo de derivação, por diferenças centrais, já se está filtrando um pouco os dados, nesta linha significa que se usar uma fórmula de maior ordem para calcular a derivada será ainda melhor, pois usaremos mais pontos para o cálculo da derivada, o que conferimos foi uma suavização ao sinal. Pensando desta forma, realmente é melhor se derivar e depois filtrar. A seguir, discutiremos um pouco mais sobre o processo que nos levou a chegar ao filtro selecionado para as derivadas. No que diz respeito à ordem do filtro, observou-se que em ordens menores o efeito dos distúrbios estocásticos se tornou mais proeminente, como podemos ver na Figura 5.16 onde comparamos a ordem 2 com a 5. Para ordens maiores que 5 não se observou melhora significativa.



**Figura 5.16:** Efeito da ordem do filtro: (a) ordem 2 e frequência de corte 0,035 e, (b) ordem 5 e frequência de corte 0,035.

Sobre a frequência de corte, com 0,035 se obteve um resultado satisfatório por acompanhar bem o formato das curvas, como podemos ver na Figura 5.17: comparamos as frequências de corte 0,015, 0,035 e 0,055 mantendo a ordem fixa em 5.



**Figura 5.17:** Comparação dos resultados para diferentes frequências de corte com ordem 5: (a) 0,015, (b) 0,035, (c) 0,055 e (d) referência.

## 5.5 Estimação dos parâmetros

Com os dados experimentais filtrados na etapa anterior:  $h1_{ExpFilt}$ ,  $h2_{ExpFilt}$ ,  $(dh1/dt)_{ExpFilt}$  e  $(dh2/dt)_{ExpFilt}$ , pudemos partir para a estimação dos parâmetros. Aqui aplicamos a metodologia apresentada na seção 3.3: até se obter uma estimativa definitiva  $\hat{\theta}$ , vamos realizando diversas estimações de parâmetros de forma iterativa. Como vimos na seção 3.3, tal procedimento iterativo pode ser separado por etapas e será assim que abordaremos os resultados daqui em diante.

Lembrando que a metodologia elaborada e neste momento aplicada foi desenvolvida com a principal finalidade de suavizar o efeito dos distúrbios determinísticos nos processos, por essa razão em nosso estudo de caso aplicamos distúrbios determinísticos nas frações  $x_1$  e  $x_2$  (rever seção 5.3 e Figuras 5.9 e 5.10). Portanto, aqui o alvo é obter um modelo o mais próximo possível do fundamental (Figuras 5.10), que é o processo ideal: sem distúrbios. Por fim, lembramos também que partimos da estrutura correta do modelo (Equação 5.1).

### **1ª Etapa**

→ *Ajuste da função objetivo*

Nesta primeira etapa todas as formulações de função objetivo foram testadas com três diferentes chutes iniciais a fim de verificar a convergência. O chute inicial não podia ser qualquer vetor arbitrado, pois dependendo da combinação podiam ocorrer problemas numéricos pelo transbordamento dos tanques. Na Tabela 5.5 apresentamos os três chutes iniciais.

**Tabela 5.5:** Chutes iniciais.

	<i>CD1</i>	<i>CD2</i>	<i>CD3</i>	<i>CD4</i>	<i>CD5</i>	<i>CD6</i>
$\theta^*$	17,8	19,1	15,955	15,99	13,65	13,65
$\theta_0^1$	20	21	18	18	15,5	15,5
$\theta_0^2$	20	17,5	15	17,5	11,75	16
$\theta_0^3$	16	17	20,5	19,5	11,75	11,75

Testamos oito formas diferentes de funções objetivo, os resultados podemos observar na Tabela 5.6, lembrando que  $\zeta_A = 0$  e  $\zeta_B = 0$  porque nesta etapa não utilizamos o CEKF. No que diz respeito ao chute inicial, nas três primeiras funções objetivo os resultados convergiram para a mesma resposta, já nas demais houve uma ligeira diferença em função do chute, como podemos verificar na Tabela 5.6.

De todas as formulações da Tabela 5.6 selecionamos a terceira e o seu resultado para prosseguir para a segunda etapa, sendo assim:

$$\begin{aligned}
 J_1 : \lambda_A = 0 \quad \lambda_B = 1 \quad \zeta_A = 0 \quad \zeta_B = 0 \\
 \hat{\theta}_1 = [17,314 \quad 18,977 \quad 15,564 \quad 15,799 \quad 13,051 \quad 13,695]
 \end{aligned}
 \tag{5.3}$$

**Tabela 5.6:** Resultados para o ajuste da função objetivo\*.

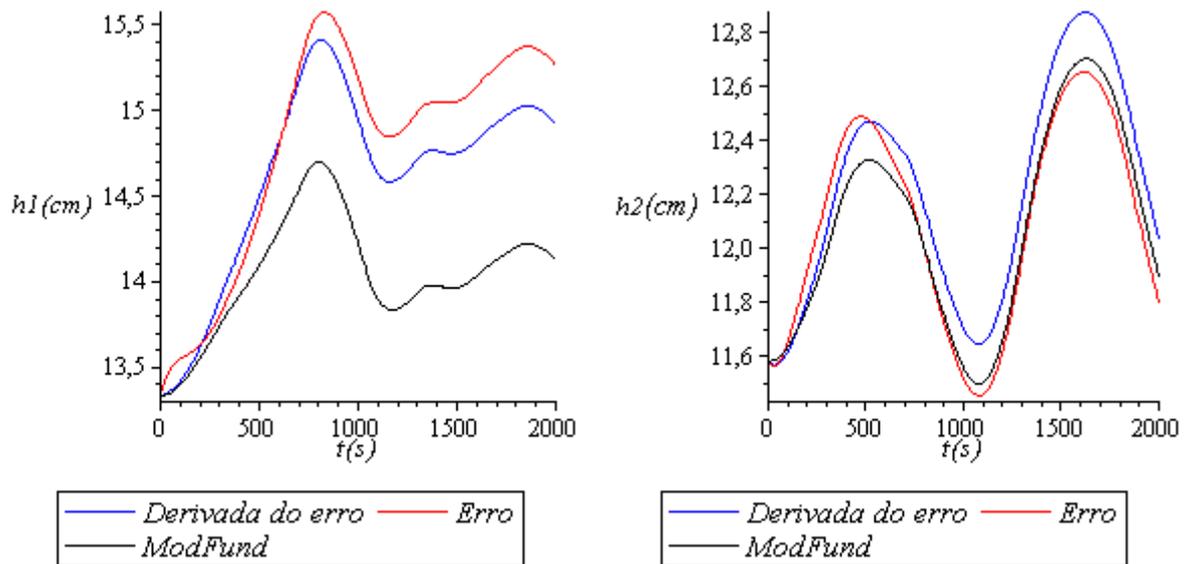
		<i>CD1</i>	<i>CD2</i>	<i>CD3</i>	<i>CD4</i>	<i>CD5</i>	<i>CD6</i>	$ \hat{\theta} - \theta^* _{\text{média}}$	$\sigma_{\hat{\theta} - \theta^*}$
$\theta^*$		17,8	19,1	15,955	15,99	13,65	13,65	-	-
$\lambda_A = 1 \quad \lambda_B = 0$	$\hat{\theta}$	17,126	19,153	15,817	15,883	12,166	14,335	0,523	0,502
$\lambda_A = 1 \quad \lambda_B = 1$	$\hat{\theta}$	17,126	19,153	15,817	15,883	12,166	14,335	0,523	0,502
$\lambda_A = 0 \quad \lambda_B = 1$	$\hat{\theta}$	17,314	18,977	15,564	15,799	13,051	13,695	0,306	0,200
$\lambda_A = 1 \quad \lambda_B = 10^2$	$\theta_0^1$	17,137	19,151	15,737	15,877	12,299	14,276	0,503	
	$\theta_0^2$	17,135	19,145	15,731	15,906	12,310	14,238	0,491	
	$\theta_0^3$	17,134	19,144	15,735	15,906	12,308	14,237	0,490	
	$\hat{\theta}_{\text{média}}$	17,135	19,147	15,734	15,896	12,306	14,250	0,495	0,447
$\lambda_A = 1 \quad \lambda_B = 10^3$	$\theta_0^1$	17,204	19,045	15,551	15,792	12,897	13,793	0,358	
	$\theta_0^2$	17,193	19,087	15,810	15,815	12,512	13,980	0,387	
	$\theta_0^3$	17,231	19,087	15,559	15,762	12,934	13,886	0,359	
	$\hat{\theta}_{\text{média}}$	17,209	19,073	15,640	15,790	12,781	13,886	0,373	0,278
$\lambda_A = 1 \quad \lambda_B = 10^4$	$\theta_0^1$	17,262	18,943	15,489	15,751	13,031	13,683	0,342	
	$\theta_0^2$	17,286	19,030	15,589	15,764	13,065	13,692	0,300	
	$\theta_0^3$	17,248	19,047	15,472	15,721	13,002	13,868	0,370	
	$\hat{\theta}_{\text{média}}$	17,265	19,007	15,517	15,745	13,033	13,748	0,338	0,205
$\lambda_A = 1 \quad \lambda_B = 10^5$	$\theta_0^1$	17,266	18,942	15,490	15,752	13,035	13,682	0,340	
	$\theta_0^2$	17,315	18,987	15,563	15,795	13,044	13,727	0,311	
	$\theta_0^3$	17,248	19,045	15,473	15,721	13,001	13,866	0,370	
	$\hat{\theta}_{\text{média}}$	17,276	18,991	15,509	15,756	13,027	13,758	0,341	0,202
$\lambda_A = 1 \quad \lambda_B = 10^6$	$\theta_0^1$	17,308	18,972	15,545	15,793	13,055	13,693	0,311	
	$\theta_0^2$	17,297	19,003	15,623	15,782	13,021	13,678	0,300	
	$\theta_0^3$	17,271	19,019	15,499	15,750	13,021	13,801	0,348	
	$\hat{\theta}_{\text{média}}$	17,292	18,998	15,556	15,775	13,032	13,724	0,319	0,204

\* Lembrando que nas três primeiras funções objetivo os resultados convergiram para a mesma resposta nos três chutes iniciais, por isso apresentamos apenas um resultado. Já nas demais houve uma ligeira diferença em função do chute, por isso apresentamos os três resultados.

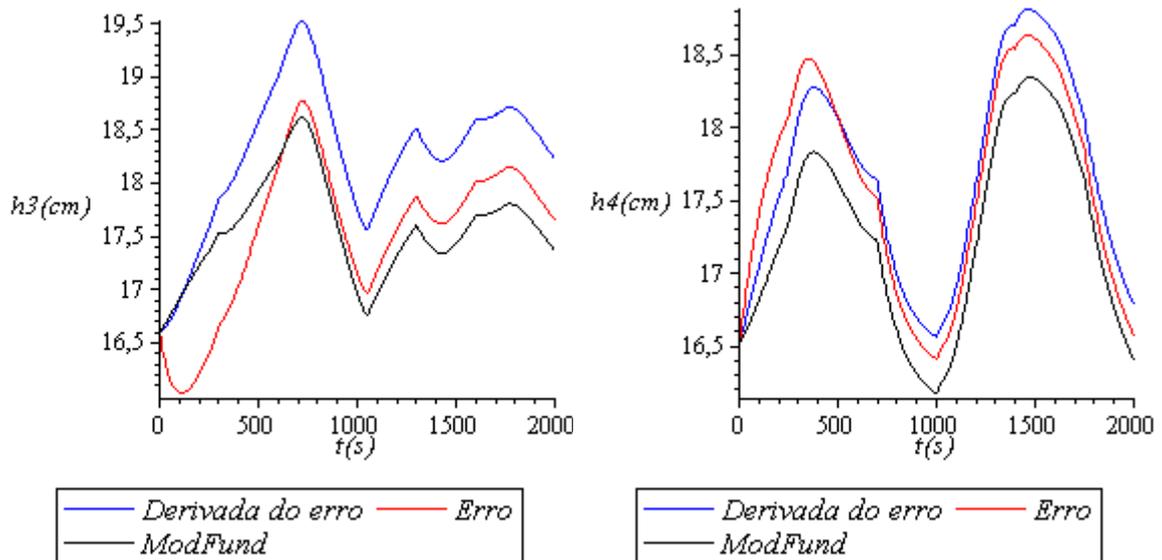
Escolhemos o terceiro ajuste para dar continuidade na segunda etapa, pois na análise geral apresentou o melhor resultado. Aqui, como conhecemos os valores reais dos parâmetros  $\theta^*$  podemos observar que outros ajustes apresentam resultados melhores para alguns

parâmetros. No entanto, na prática, o pesquisador não conhecendo  $\theta^*$  estará condicionado a escolher o resultado de apenas um ajuste. Como cada ajuste leva a uma distinta curva, a decisão mais certa é escolher o resultado daquela que menos agrega o efeito dos distúrbios. Neste caso, sugerimos escolher o resultado com menores  $\hat{\sigma}_\theta^2$  e menores diferenças entre elas.

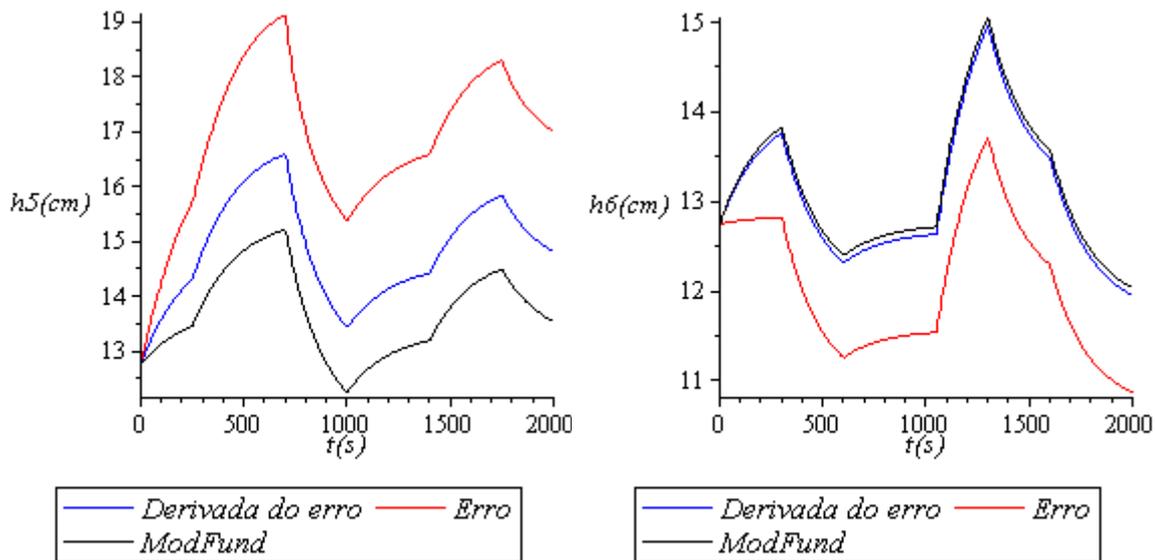
Ressaltamos ainda, que o ajuste escolhido nesta etapa se refere a uma função objetivo formulada apenas com as derivadas. Podemos ver numericamente na Tabela 5.6 que tal função objetivo foi significativamente melhor que a tradicional, que é sem as derivadas (primeiro resultado da Tabela 5.6). Podemos visualizar graficamente a comparação entre elas nas Figuras 5.18 a 5.20.



**Figura 5.18:** Comparação dos resultados apresentados na Tabela 5.6: Erro (primeiro ajuste), Derivada do erro (terceiro ajuste). *ModFund*: Referência, modelo ideal.



**Figura 5.19:** Comparação dos resultados apresentados na Tabela 5.6: Erro (primeiro ajuste), Derivada do erro (terceiro ajuste). *ModFund*: Referência, modelo ideal.



**Figura 5.20:** Comparação dos resultados apresentados na Tabela 5.6: Erro (primeiro ajuste), Derivada do erro (terceiro ajuste). *ModFund*: Referência, modelo ideal.

Ao observar as Figuras 5.18 a 5.20 verificamos que com a derivada do erro se tem um melhor acompanhamento do perfil da curva ideal, assim como no estudo de caso de Machado e Trierweiler (2004) discutido na seção 2.2.2.

## 2ª Etapa

→ *Ajuste da sintonia do CEKF*

Considerando  $J = J_1$  e  $\theta_0 = \hat{\theta}_1$ , bem como o CEKF com os modelos:  $\dot{x} = f(x, u; \hat{\theta}_1) + w$  e  $y = g(x; \hat{\theta}_1) + v$ , foram testadas diversas formas de sintonia, dentre elas selecionou-se como melhor as matrizes:  $P_{ii} = 10^{-3}$ ,  $Q_{ii} = 10^{-6}$  e  $R_{ii} = 10^{-3}$ , sendo todas diagonais.

Para fins de comparação, na Tabela 5.7 apresentamos outras formas de sintonia e seus respectivos resultados. Na Tabela 5.7 todas as matrizes especificadas são diagonais e, quando avaliamos o efeito de uma matriz as demais foram mantidas constantes nos valores escolhidos.

Como restrições para os estados estipulou-se:  $h1=11..17$ ,  $h2=9..15$ ,  $h3=13..22$ ,  $h4=13..22$ ,  $h5=5..17$ ,  $h6=5..17$ .

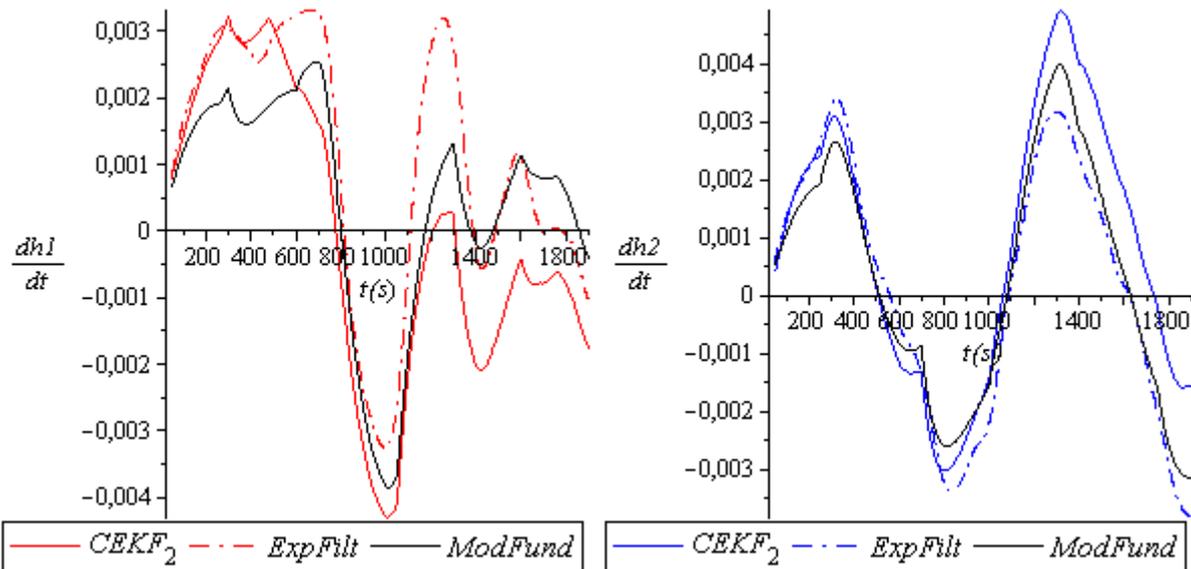
**Tabela 5.7:** Comparação da estimação com outras sintonias (escolhida destacada).

		<i>CD1</i>	<i>CD2</i>	<i>CD3</i>	<i>CD4</i>	<i>CD5</i>	<i>CD6</i>	$ \hat{\theta} - \theta^* _{m\u00e9dia}$	$\sigma_{\hat{\theta} - \theta^*}$
$\theta^*$		17,8	19,1	15,955	15,99	13,65	13,65	-	-
$P_{ii}$	$10^{-1}$	17,681	18,917	15,595	16,034	13,844	13,532	0,170	0,098
	$10^{-3}$	17,669	18,968	15,602	15,948	13,811	13,627	0,140	0,107
	$10^{-5}$	17,666	18,973	15,601	15,953	13,801	13,634	0,137	0,110
$Q_{ii}$	$10^{-5}$	17,698	18,982	15,603	15,959	13,878	13,637	0,140	0,117
	$10^{-6}$	17,669	18,968	15,602	15,948	13,811	13,627	0,140	0,107
	$10^{-7}$	17,573	18,967	15,619	15,910	13,584	13,625	0,144	0,107
$R_{ii}$	$10^{-1}$	17,409	18,980	15,656	15,838	13,185	13,679	0,242	0,155
	$10^{-3}$	17,669	18,968	15,602	15,948	13,811	13,627	0,140	0,107
	$10^{-5}$	17,714	18,950	15,603	15,980	13,914	13,630	0,147	0,125

Podemos ver na Tabela 5.7 que o emprego do CEKF melhorou de forma significativa a estimação dos parâmetros, observe a comparação direta na Tabela 5.8. Ainda podemos ver a melhoria na Figura 5.21, onde fazemos a comparação entre os dados utilizados para a estimação, observe que com a passagem do CEKF as curvas acompanharam melhor o formato referente ao modelo fundamental.

**Tabela 5.8:** Resultado obtido na etapa anterior em comparação com a sintonia selecionada.

	<i>CD1</i>	<i>CD2</i>	<i>CD3</i>	<i>CD4</i>	<i>CD5</i>	<i>CD6</i>	$ \hat{\theta} - \theta^* _{m\u00e9dia}$	$\sigma_{\hat{\theta} - \theta^*}$
$\theta^*$	17,8	19,1	15,955	15,99	13,65	13,65	-	-
1ª Etapa: $\hat{\theta}_1$	17,314	18,977	15,564	15,799	13,051	13,695	0,306	0,200
2ª Etapa: Sintonia Escolhida	17,669	18,968	15,602	15,948	13,811	13,627	0,140	0,107



**Figura 5.21:** *ModFund*: Se refere aos valores verdadeiros, referência para comparação. *ExpFilt*: derivadas experimentais filtradas com a função *IdFilt* do MATLAB<sup>®</sup>, utilizadas na primeira etapa para obtenção de  $\hat{\theta}_1$ . *CEKF<sub>2</sub>*: derivadas melhoradas com a passagem do CEKF nesta segunda etapa.

Ao obtermos  $y^{filt}$  e  $x^{est}$  com o emprego do CEKF, pudemos passar para a segunda parte desta segunda etapa: ajuste da função objetivo.

→ *Ajuste da função objetivo*

No momento anterior (ajuste da sintonia do CEKF) obtivemos resultados com a parcela relativa à  $\lambda_B$ , pois consideramos  $J = J_1$ . Agora iremos nos aprofundar um pouco mais na composição da função objetivo. Na primeira etapa vimos que agregar a parcela:

$$\sum_{i=1}^{ne} \sum_{j=1}^{ny} [y_{ij} - g(x_i^{sim}; \theta)]^2 \quad (5.4)$$

não foi vantajoso, pois obtivemos  $\lambda_A = 0$ . Mas, com os dados melhorados com o CEKF não descartamos essa possibilidade, testando diretamente a influência de tal parcela com  $\lambda_A = 1$  e demais zeradas. O resultado obtido é apresentado na Tabela 5.9.

**Tabela 5.9:** Resultado da função objetivo formada apenas com a parcela relativa a  $\lambda_A$ .

	<i>CD1</i>	<i>CD2</i>	<i>CD3</i>	<i>CD4</i>	<i>CD5</i>	<i>CD6</i>	$ \hat{\theta} - \theta^* _{m\u00e9dia}$	$\sigma_{\hat{\theta} - \theta^*}$
$\theta^*$	17,8	19,1	15,955	15,99	13,65	13,65	-	-
$\lambda_A = 1 \quad \lambda_B = 0$ $\zeta_A = 0 \quad \zeta_B = 0$	17,158	19,121	15,857	15,867	12,209	14,221	0,483	0,490

Como não observamos melhora significativa em relação a primeira etapa, aqui já descartamos outras possibilidades de combinações com  $\lambda_A \neq 0$ .

Na seqüência, avaliamos a influência da parcela:

$$\sum_{i=1}^{ne} \sum_{j=1}^{nx^{est}} [x_{ij}^{est} - x_{ij}^{sim}]^2 \quad (5.5)$$

relativa a  $\zeta_A$ , o resultado observamos na Tabela 5.10.

**Tabela 5.10:** Avaliação da influência da parcela relativa a  $\zeta_A$ , tendo  $\lambda_A = 0$  e  $\zeta_B = 0$ .

	CD1	CD2	CD3	CD4	CD5	CD6	$ \hat{\theta} - \theta^* _{m\acute{e}dia}$	$\sigma_{\hat{\theta} - \theta^*}$
$\theta^*$	17,8	19,1	15,955	15,99	13,65	13,65	-	-
$\lambda_B = 0 \quad \zeta_A = 1$	17,378	18,988	15,581	15,824	13,078	13,710	0,284	0,184
$\lambda_B = 10^3 \quad \zeta_A = 1$	17,370	18,985	15,589	15,823	13,106	13,703	0,279	0,178
$\lambda_B = 10^6 \quad \zeta_A = 1$	17,369	18,976	15,563	15,820	13,173	13,686	0,272	0,168

Como nenhum resultado da Tabela 5.10 se apresentou melhor que quando  $\lambda_B = 1$  e demais parcelas zeradas (Tabela 5.8), também descartamos a parcela relativa a  $\zeta_A$  na função objetivo.

Por fim, ao verificar a influência da parcela:

$$\sum_{i=1}^{ne} \sum_{j=1}^{nx^{est}} \left[ \frac{dx_{ij}^{est}}{dt} - \frac{dx_{ij}^{sim}}{dt} \right]^2 \quad (5.6)$$

relativa a  $\zeta_B$  (Tabela 5.11), concluímos que a melhor formulação para a função objetivo nesta segunda etapa foi quando:  $\lambda_B = 1$  e demais iguais a zero. Desta forma, prosseguimos para a terceira etapa como vemos na Equação 5.7.

**Tabela 5.11:** Avaliação da influência da parcela relativa a  $\zeta_B$ , tendo  $\lambda_A = 0$  e  $\zeta_A = 0$ .

	CD1	CD2	CD3	CD4	CD5	CD6	$ \hat{\theta} - \theta^* _{m\acute{e}dia}$	$\sigma_{\hat{\theta} - \theta^*}$
$\theta^*$	17,8	19,1	15,955	15,99	13,65	13,65	-	-
$\lambda_B = 0 \quad \zeta_B = 1$	17,394	18,987	15,629	15,826	13,112	13,708	0,267	0,170
$\lambda_B = 1 \quad \zeta_B = 1$	17,413	18,910	15,647	15,838	13,153	13,709	0,266	0,148
$\lambda_B = 1 \quad \zeta_B = 0,5$	17,446	18,991	15,667	15,858	13,240	13,707	0,225	0,132

$$\begin{aligned}
 J_2 : \lambda_A = 0 \quad \lambda_B = 1 \quad \zeta_A = 0 \quad \zeta_B = 0 \\
 \hat{\theta}_2 = [17,669 \quad 18,968 \quad 15,602 \quad 15,948 \quad 13,811 \quad 13,627]
 \end{aligned}
 \tag{5.7}$$

### **3ª Etapa**

→ Ajuste da sintonia do CEKF

Com a atualização:

- $J = J_2$ ,
- $\theta_0 = \hat{\theta}_2$ ,
- $\dot{x} = f(x, u; \hat{\theta}_2) + w$  e
- $y = g(x; \hat{\theta}_2) + v$ .

Inicialmente ajustamos a matriz  $Q$ , dado que houve uma melhoria na confiança do modelo em função de  $\hat{\theta}_2$ . Para o ajuste da matriz  $Q$  mantivemos as demais matrizes fixas nos valores ajustados na etapa anterior:  $P_{ii} = 10^{-3}$  e  $R_{ii} = 10^{-3}$ . Após diversos testes selecionamos  $Q_{ii} = 10^{-10}$ , lembrando que na etapa anterior  $Q_{ii} = 10^{-6}$ . Para fins de comparação apresentamos alguns outros testes na Tabela 5.12.

**Tabela 5.12:** Comparação de resultados para o ajuste da matriz  $Q$  ( $P_{ii} = 10^{-3}$  e  $R_{ii} = 10^{-3}$ ).

	CD1	CD2	CD3	CD4	CD5	CD6	$ \hat{\theta} - \theta^* _{\text{média}}$	$\sigma_{\hat{\theta} - \theta^*}$
$\theta^*$	17,8	19,1	15,955	15,99	13,65	13,65	-	-
$Q_{ii} = 10^{-6}$	17,990	18,888	15,701	16,046	14,350	13,568	0,249	0,213
$Q_{ii} = 10^{-8}$	17,832	18,913	15,607	15,976	14,031	13,599	0,169	0,150
$Q_{ii} = 10^{-10}$	17,819	18,918	15,586	15,955	14,013	13,617	0,167	0,151

Da mesma forma, fizemos o ajuste das matrizes  $R$  e  $P$ , como podemos ver nas Tabelas 5.13 e 5.14.

**Tabela 5.13:** Comparação de resultados para o ajuste da matriz  $R$  ( $Q_{ii} = 10^{-10}$  e  $P_{ii} = 10^{-3}$ ).

	CD1	CD2	CD3	CD4	CD5	CD6	$ \hat{\theta} - \theta^* _{\text{média}}$	$\sigma_{\hat{\theta} - \theta^*}$
$\theta^*$	17,8	19,1	15,955	15,99	13,65	13,65	-	-
$R_{ii} = 10^{-5}$	17,768	18,949	15,703	15,947	13,983	13,594	0,145	0,114
$R_{ii} = 10^{-6}$	17,759	18,964	15,863	15,940	14,011	13,566	0,127	0,109
$R_{ii} = 10^{-7}$	17,760	19,040	16,058	15,876	14,086	13,642	0,127	0,143

**Tabela 5.14:** Comparação de resultados para o ajuste da matriz  $P$  ( $Q_{ii} = 10^{-10}$  e  $R_{ii} = 10^{-6}$ ).

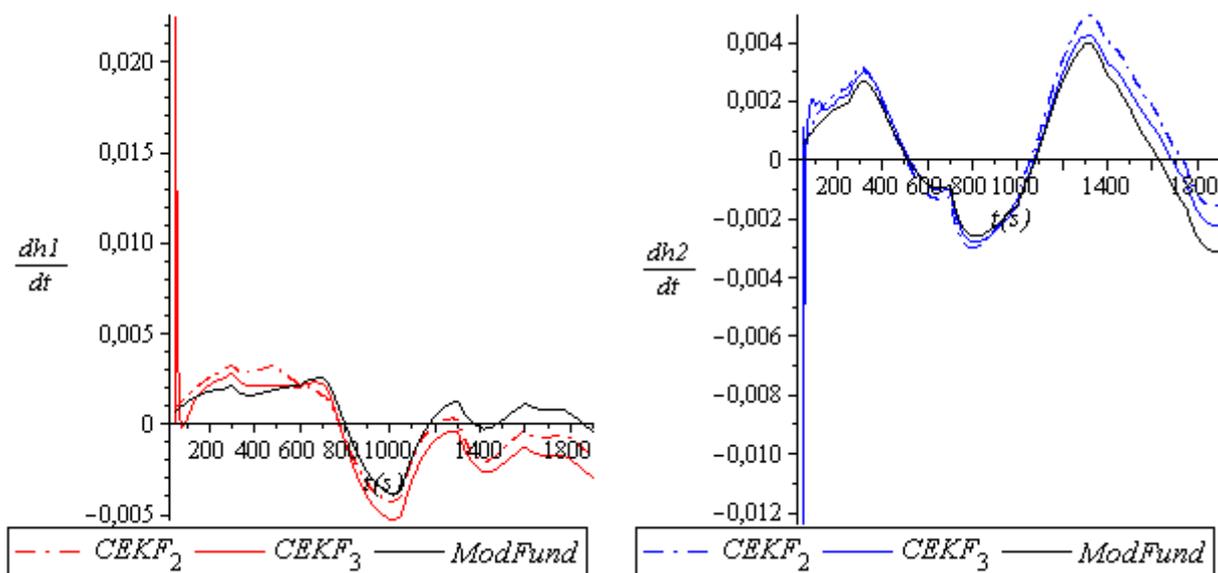
	CD1	CD2	CD3	CD4	CD5	CD6	$ \hat{\theta} - \theta^* _{m\u00e9dia}$	$\sigma_{\hat{\theta} - \theta^*}$
$\theta^*$	17,8	19,1	15,955	15,99	13,65	13,65	-	-
$P_{ii} = 10$	17,662	19,066	16,141	15,819	13,767	13,689	0,114	0,059
$P_{ii} = 1$	17,670	19,066	16,107	15,818	13,794	13,689	0,112	0,055
$P_{ii} = 0,1$	17,663	19,072	16,110	15,812	13,794	13,695	0,115	0,057

Com base nos resultados das Tabelas 5.12 a 5.14 selecionamos a sintonia  $P_{ii} = 1$ ,  $Q_{ii} = 10^{-10}$  e  $R_{ii} = 10^{-6}$ . Neste estágio já verificamos uma melhoria em relação à etapa anterior, como podemos ver na Tabela 5.15.

**Tabela 5.15:** Resultado obtido na etapa anterior em comparação com a sintonia selecionada.

	CD1	CD2	CD3	CD4	CD5	CD6	$ \hat{\theta} - \theta^* _{m\u00e9dia}$	$\sigma_{\hat{\theta} - \theta^*}$
$\theta^*$	17,8	19,1	15,955	15,99	13,65	13,65	-	-
2ª Etapa: $\hat{\theta}_2$	17,669	18,968	15,602	15,948	13,811	13,627	0,140	0,107
3ª Etapa: Sintonia Escolhida	17,670	19,066	16,107	15,818	13,794	13,689	0,112	0,055

Na Figura 5.22 fazemos a comparação entre os dados utilizados para a estimação na etapa anterior ( $CEKF_2$ ) com a presente etapa ( $CEKF_3$ ).

**Figura 5.22:** Comparação entre dados utilizados para estimação dos parâmetros na segunda ( $CEKF_2$ ) e terceira etapa ( $CEKF_3$ ).

A primeira vista os gráficos da Figura 5.22 são intrigantes, levando-nos a questionar: “como a estimação com  $CEKF_3$  pôde ser melhor que com  $CEKF_2$ , sendo que existe um considerável erro ou discrepância inicial em  $CEKF_3$ ?”. Na verdade, tal erro é responsável pela melhoria na estimação. Constatamos que este erro conduz a um melhor balanço numérico de forma que a estimação dos parâmetros é melhorada. E, salientamos: esse balanço numérico somente foi possível com a passagem do CEKF, reforçando ainda mais a vantagem do seu uso.

Sobre essa afirmação: “melhoria no balanço numérico” ainda iremos discutir um pouco mais. Começamos comparando o resultado de  $CEKF_3$  com o terceiro da Tabela 5.12 ( $P_{ii} = 10^{-3}$ ,  $Q_{ii} = 10^{-10}$  e  $R_{ii} = 10^{-3}$ ), para facilitar a comparação na Tabela 5.16 analisamos diretamente os resultados dos dois casos tomados neste momento.

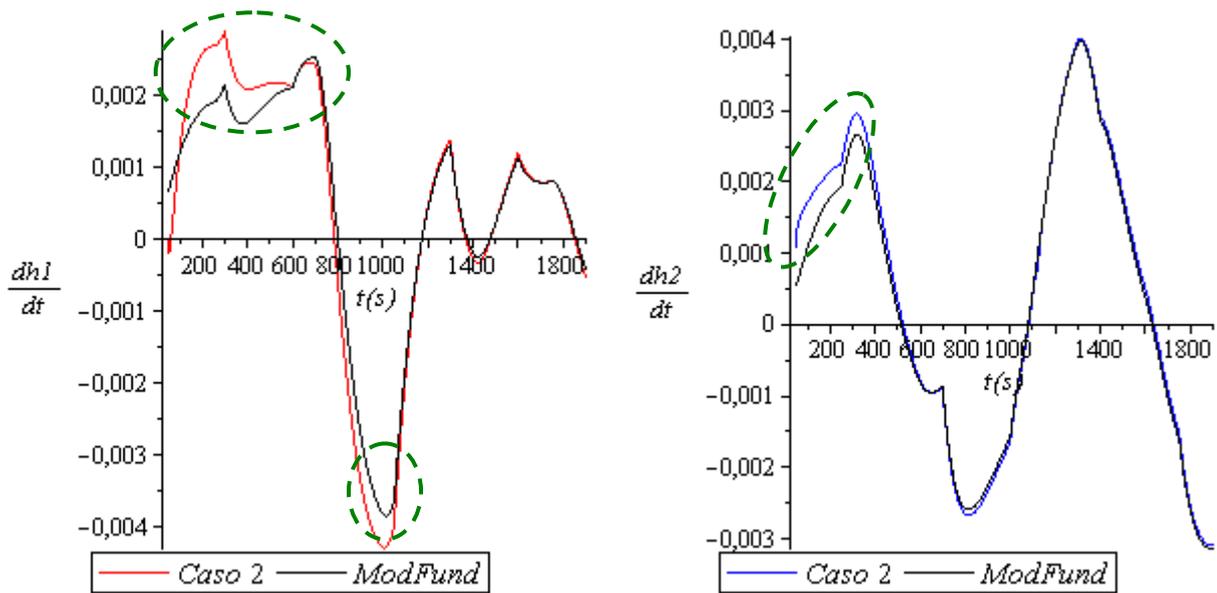
**Tabela 5.16:** Dois casos tomados para comparação.

	CD1	CD2	CD3	CD4	CD5	CD6	$ \hat{\theta} - \theta^* _{\text{média}}$	$\sigma_{\hat{\theta} - \theta^*}$
$\theta^*$	17,8	19,1	15,955	15,99	13,65	13,65	-	-
<b>Caso 1:</b> “ $CEKF_3$ ”, $P_{ii} = 1$ , $Q_{ii} = 10^{-10}$ e $R_{ii} = 10^{-6}$	17,670	19,066	16,107	15,818	13,794	13,689	0,112	0,055
<b>Caso 2:</b> $P_{ii} = 10^{-3}$ , $Q_{ii} = 10^{-10}$ e $R_{ii} = 10^{-3}$	17,819	18,918	15,586	15,955	14,013	13,617	0,167	0,151

Na Tabela 5.16 pode-se ver que o resultado do primeiro caso é melhor que do segundo, embora as curvas  $(dh/dt)_{CEKF}$  do mesmo sejam mais próximas das verdadeiras que se pode ver na Figura 5.23. Como no segundo caso grande parte dos valores  $(dh/dt)_{CEKF}$  estão muito próximos dos verdadeiros, o maior trabalho do problema de otimização acaba se concentrando em minimizar o erro das regiões destacadas na Figura 5.23. Em outras palavras, os demais pontos não exercem influência significativa na estimação dos parâmetros, sendo as regiões destacadas determinantes. Assim, se caiu em uma situação em que o erro está longe de ser considerado uniformemente distribuído.

Tomamos o segundo caso como exemplo por ter a maior parte dos pontos muito próximos dos verdadeiros, isso para demonstrar que tal situação não implica necessariamente na melhoria da estimação.

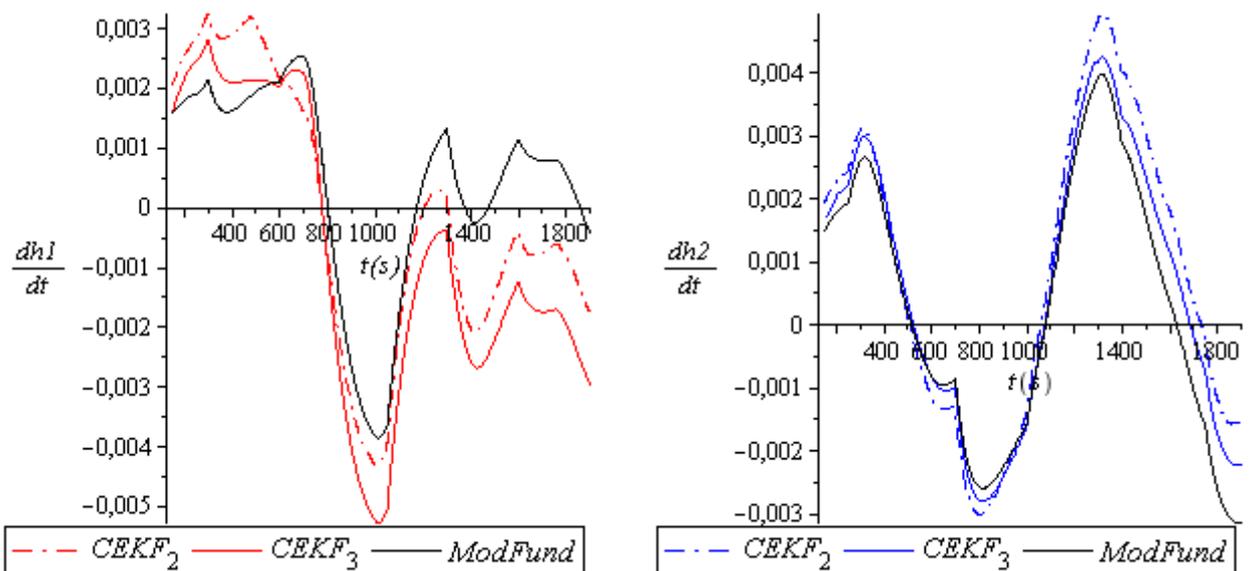
Ainda para sustentar o que foi argumentado sobre a importância do erro inicial para contrabalançar a não uniformidade dos erros, foi feita uma nova estimação removendo a discrepância inicial verificada na Figura 5.22. Na Tabela 5.17 apresenta-se o resultado obtido e se pode constatar que não foi melhorado. Também, na Figura 5.24 traz-se os gráficos com o erro inicial removido.



**Figura 5.23:** Visualização gráfica do segundo caso da Tabela 5.15.

**Tabela 5.17:** Estimação com a remoção do erro inicial da Figura 5.22.

	CD1	CD2	CD3	CD4	CD5	CD6	$ \hat{\theta} - \theta^* _{m\u00e9dia}$	$\sigma_{\hat{\theta} - \theta^*}$
$\theta^*$	17,8	19,1	15,955	15,99	13,65	13,65	-	-
$\hat{\theta}$	17,954	18,546	15,640	16,375	14,110	13,319	0,366	0,125



**Figura 5.24:** Gráficos com o erro inicial em  $CEKF_3$  removido.

Dados os fatos, vimos na Tabela 5.17 que ao remover a discrepância inicial o resultado ficou pior, o que nos confirmou que tais pontos são fundamentais para a melhoria na estimação, embora aparentemente nos leve a crer que não. Sendo assim, encerramos esta parte passando para a segunda: ajuste da função objetivo.

→ *Ajuste da função objetivo*

Já tendo uma direção das etapas anteriores, analisamos individualmente a influência de cada parcela relativa a  $\lambda_A$ ,  $\zeta_A$  e  $\zeta_B$ , caso verificássemos alguma vantagem aprofundaríamos a investigação, caso contrário não. Com respeito à  $\lambda_A$  apresentamos o resultado na Tabela 5.18.

**Tabela 5.18:** Resultado da função objetivo formada apenas com a parcela relativa a  $\lambda_A$ .

	CD1	CD2	CD3	CD4	CD5	CD6	$ \hat{\theta} - \theta^* _{\text{média}}$	$\sigma_{\hat{\theta} - \theta^*}$
$\theta^*$	17,8	19,1	15,955	15,99	13,65	13,65	-	-
$\lambda_A = 1 \quad \lambda_B = 0$ $\zeta_A = 0 \quad \zeta_B = 0$	17,364	19,049	15,784	15,847	12,842	13,968	0,321	0,251

Como não observamos vantagem significativa, aqui já descartamos outras possibilidades de combinações com  $\lambda_A \neq 0$ .

Sobre a parcela relativa à  $\zeta_A$ , ainda aprofundamos um pouco mais a investigação realizando três testes, como podemos ver na Tabela 5.19. No entanto, no fim acabamos descartando a possibilidade de incorporar tal parcela na função objetivo.

**Tabela 5.19:** Avaliação da influência da parcela relativa a  $\zeta_A$ , tendo  $\lambda_A = 0$  e  $\zeta_B = 0$ .

	CD1	CD2	CD3	CD4	CD5	CD6	$ \hat{\theta} - \theta^* _{\text{média}}$	$\sigma_{\hat{\theta} - \theta^*}$
$\theta^*$	17,8	19,1	15,955	15,99	13,65	13,65	-	-
$\lambda_B = 0 \quad \zeta_A = 1$	17,607	19,061	15,530	15,967	13,795	13,630	0,141	0,143
$\lambda_B = 10^3 \quad \zeta_A = 1$	17,669	18,968	15,602	15,948	13,811	13,627	0,140	0,107
$\lambda_B = 10^6 \quad \zeta_A = 1$	17,668	18,973	15,622	15,941	13,811	13,630	0,137	0,101

Por fim, ao avaliar a influência de  $\zeta_B$  constatamos também não ser vantajoso sua utilização na função objetivo, como podemos ver na Tabela 5.20.

**Tabela 5.20:** Resultado da função objetivo formada apenas com a parcela relativa a  $\zeta_B$ .

	CD1	CD2	CD3	CD4	CD5	CD6	$ \hat{\theta} - \theta^* _{\text{média}}$	$\sigma_{\hat{\theta} - \theta^*}$
$\theta^*$	17,8	19,1	15,955	15,99	13,65	13,65	-	-
$\lambda_A = 0 \quad \lambda_B = 0$ $\zeta_A = 0 \quad \zeta_B = 1$	17,823	19,256	19,171	15,685	12,652	14,135	0,864	1,096

Com base nos resultados obtidos, a partir desta etapa adotamos como fixa para as demais a seguinte função objetivo:  $J: \lambda_A = 0 \quad \lambda_B = 1 \quad \zeta_A = 0 \quad \zeta_B = 0$ . Desta forma partimos para a quarta etapa com:

$$\hat{\theta}_3 = [17,670 \quad 19,066 \quad 16,107 \quad 15,818 \quad 13,794 \quad 13,689] \quad (5.8)$$

#### **4ª Etapa**

→ *Ajuste da sintonia do CEKF*

Foram realizados diversos testes buscando uma melhoria significativa em relação à etapa anterior, no entanto, este alvo não foi alcançado. A qualidade do melhor resultado obtido foi equivalente a etapa anterior, o que significa que aqui encerramos o processo iterativo da busca por  $\hat{\theta}$ .

Nesta etapa o melhor resultado que se obteve foi com:  $P_{ii} = 10^{-8}$ ,  $Q_{ii} = 10^{-12}$  e  $R_{ii} = 10^{-6}$ , podemos visualizá-lo ( $\hat{\theta}_4$ ) em comparação com os resultados das demais etapas na Tabela 5.21.

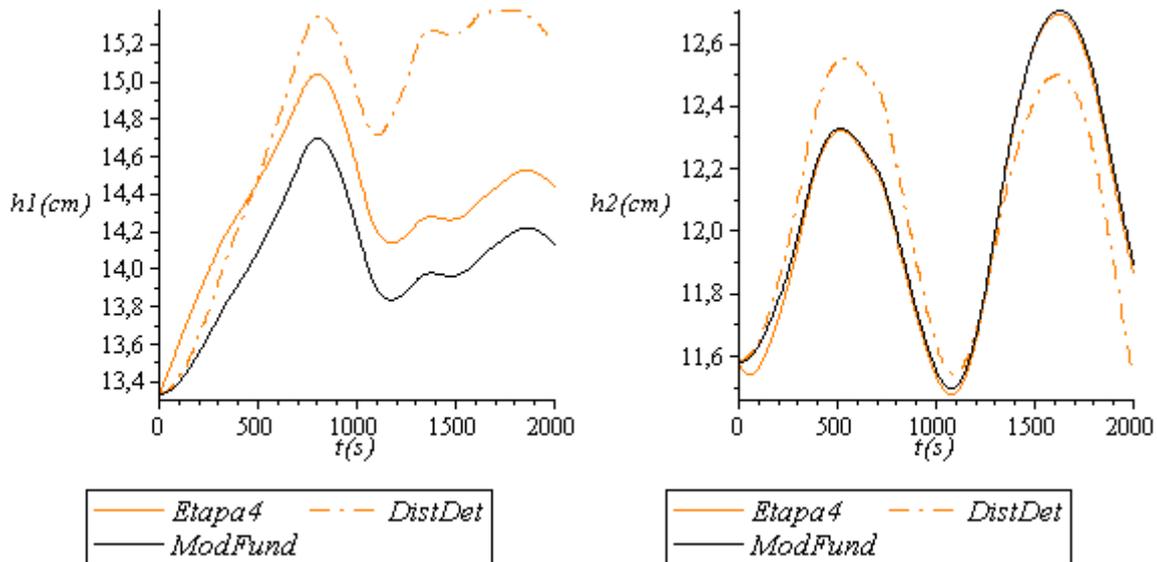
**Tabela 5.21:** Comparação entre os resultados obtidos nas quatro etapas.

	CD1	CD2	CD3	CD4	CD5	CD6	$ \hat{\theta} - \theta^* _{\text{média}}$	$\sigma_{\hat{\theta} - \theta^*}$
$\theta^*$	17,8	19,1	15,955	15,99	13,65	13,65	-	-
$\hat{\theta}_1$	17,314	18,977	15,564	15,799	13,051	13,695	0,306	0,200
$\hat{\theta}_2$	17,669	18,968	15,602	15,948	13,811	13,627	0,140	0,107
$\hat{\theta}_3$	17,670	19,066	16,107	15,818	13,794	13,689	0,112	0,055
$\hat{\theta}_4$	17,612	19,114	16,045	15,845	13,794	13,734	0,111	0,056

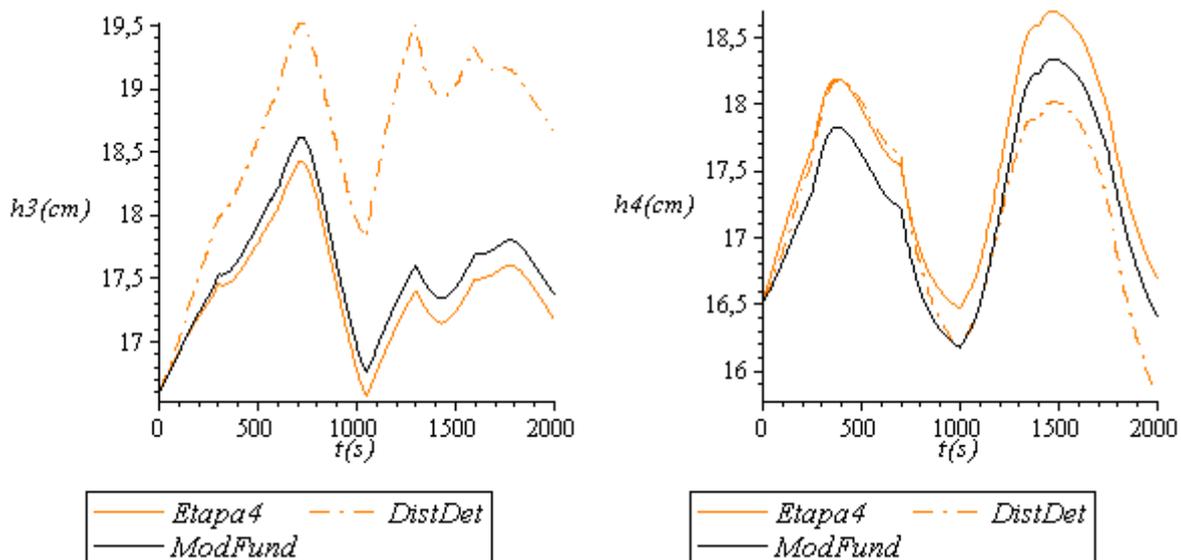
Como podemos constatar na Tabela 5.21, a diferença dos resultados para  $\hat{\theta}_3$  e  $\hat{\theta}_4$  é sutil. Sendo assim, escolhemos como definitivo  $\hat{\theta} = \hat{\theta}_4$ .

Nas Figuras 5.25 a 5.27 podemos observar graficamente o modelo obtido com  $\hat{\theta}$  (*Etapa 4*) em comparação com o modelo ideal (*ModFund*) e o processo com os distúrbios determinísticos (*DistDet*) adicionados, como vimos na seção 5.3.

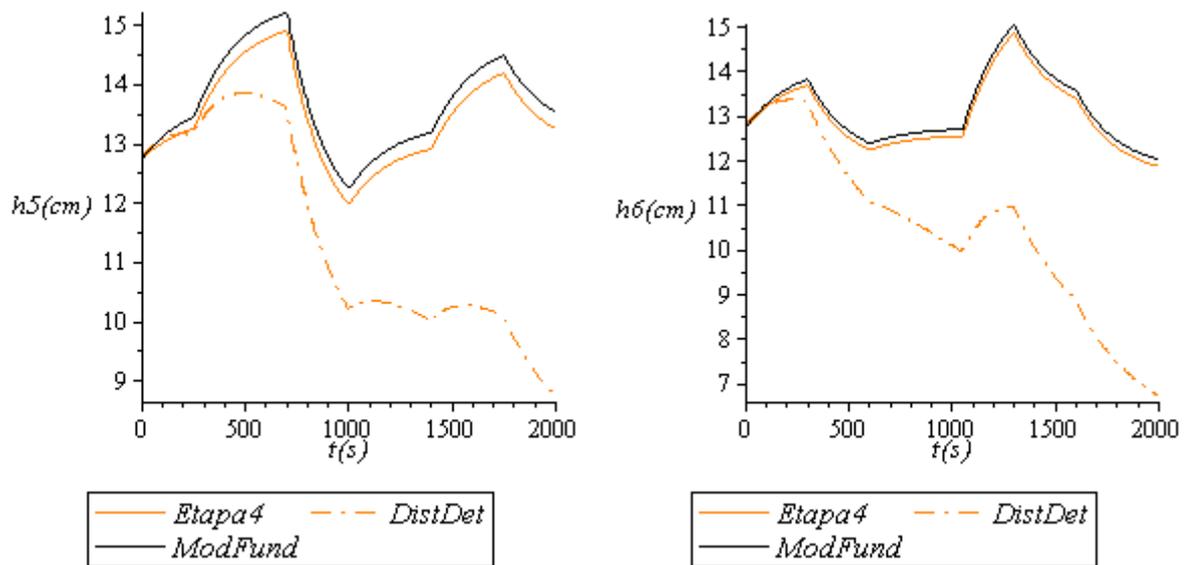
Por fim, nas Figuras 5.28 a 5.30 apresentamos graficamente os resultados obtidos nas quatro etapas.



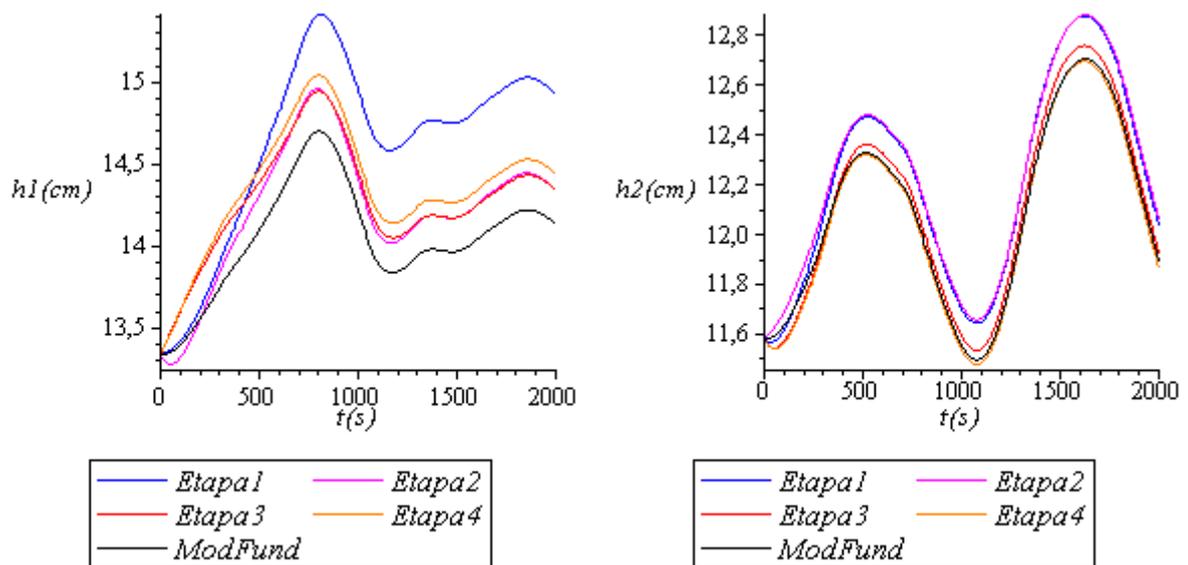
**Figura 5.25:**  $h1$  e  $h2$ : Comparação do modelo obtido (*Etapa 4*) com o modelo ideal (*ModFund*) e o processo com os distúrbios determinísticos (*DistDet*).



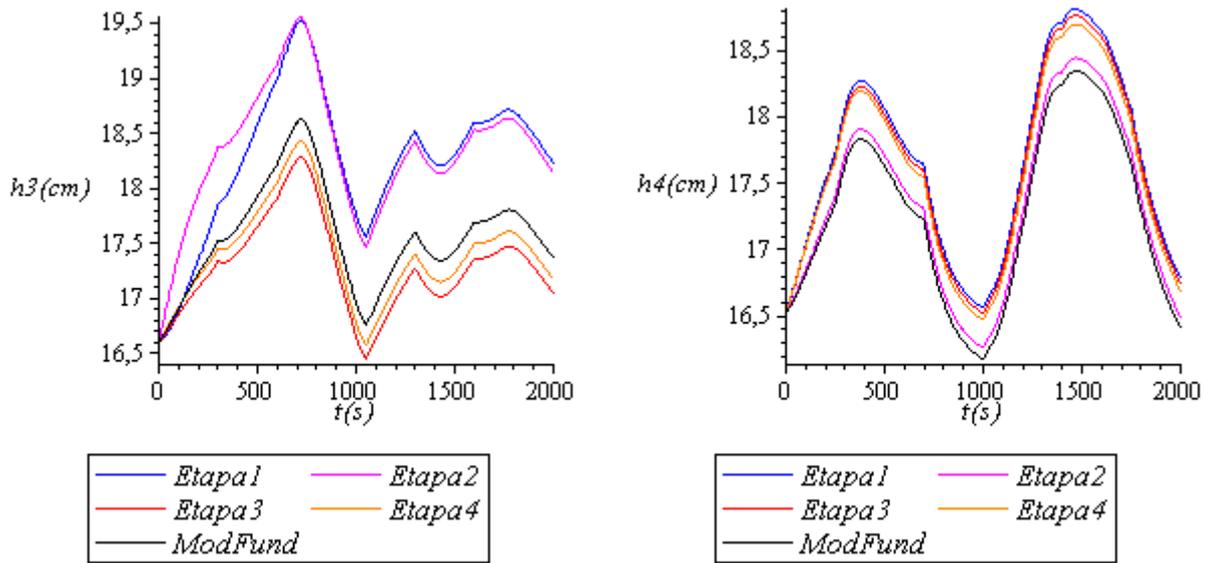
**Figura 5.26:**  $h3$  e  $h4$ : Comparação do modelo obtido (*Etapa 4*) com o modelo ideal (*ModFund*) e o processo com os distúrbios determinísticos (*DistDet*).



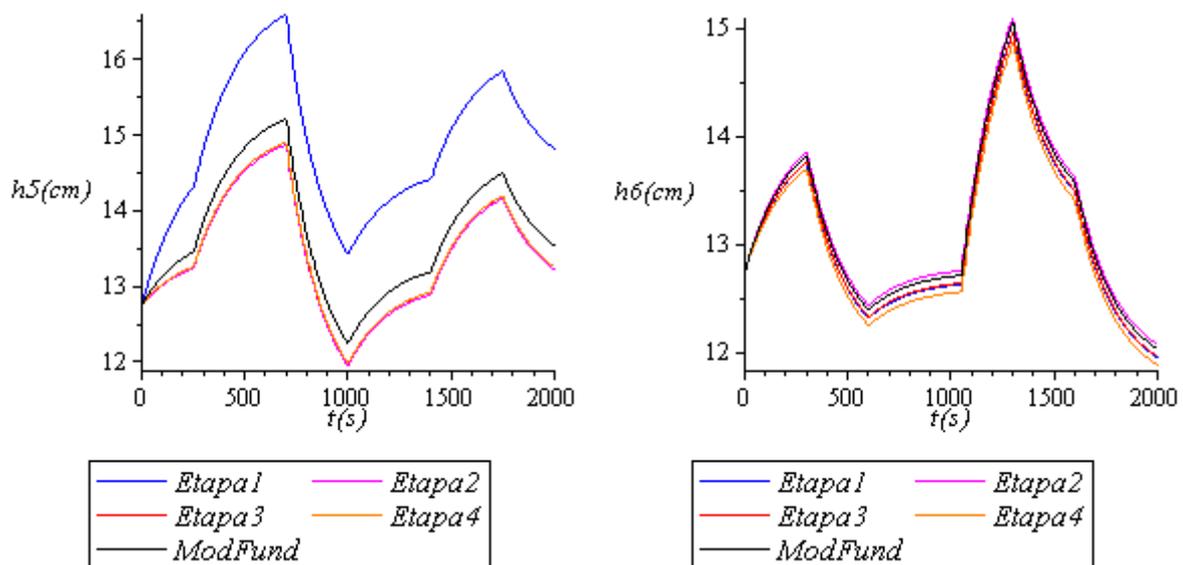
**Figura 5.27:**  $h_5$  e  $h_6$ : Comparação do modelo obtido (*Etapa 4*) com o modelo ideal (*ModFund*) e o processo com os distúrbios determinísticos (*DistDet*).



**Figura 5.28:** Comparação entre os resultados obtidos nas quatro etapas.



**Figura 5.29:** Comparação entre os resultados obtidos nas quatro etapas.



**Tabela 5.30:** Comparação entre os resultados obtidos nas quatro etapas.

# Capítulo 6

## Conclusão

Podemos dizer que o presente trabalho é o início de uma metodologia de identificação de processos inovadora que ainda deve ser continuado. Contudo, o primeiro passo já foi dado. Trazendo a tona novas direções a serem tomadas em vista do aperfeiçoamento da técnica.

Ao aumentar o horizonte de visão, podemos aspirar com a elaboração de um sistema iterativo com o usuário ao construir uma interface gráfica. Como pudemos ver no presente trabalho, a etapa da estimabilidade dos parâmetros já foi desenvolvida.

Visto que o uso das derivadas na função objetivo resultou no melhor resultado, aperfeiçoar a técnica do tratamento prévio dos dados experimentais é uma vertente de pesquisa interessante. No presente trabalho nos valem de uma ferramenta já oferecida no pacote do Matlab<sup>®</sup>, e o que estamos sugerindo para trabalhos futuros é a elaboração de uma nova, eliminando o problema da defasagem e ainda melhorando a qualidade da filtragem. O cálculo da derivada simultaneamente à filtragem consiste em uma idéia que pode retornar bons resultados.

Ainda, no presente trabalho se utilizou o método de diferenças finitas centrais para o cálculo das derivadas experimentais, sugerimos também o estudo do desempenho de outros métodos numéricos mais elaborados.

O objetivo central do presente trabalho foi: verificar se emprego do CEKF melhoraria qualidade da estimação dos parâmetros em um caso com distúrbios determinísticos desconhecidos nas entradas. Com base nos resultados obtidos e apresentados concluímos que sim, o emprego do CEKF ocasionou em uma melhoria perceptível. Argumentando mais tal afirmação de forma prática, comparamos os resultados das estimações nas quatro etapas. Na Tabela 6.1 apresentamos os índices de desempenho adotados.

**Tabela 6.1:** Comparação dos índices de desempenho nas quatro etapas.

	$ \hat{\theta} - \theta^* _{\text{média}}$	$\sigma_{\hat{\theta} - \theta^*}$
1ª Etapa	0,306	0,200
2ª Etapa	0,140	0,107
3ª Etapa	0,112	0,055
4ª Etapa	0,111	0,056

De acordo com a Tabela 6.1, observamos uma melhoria aproximada de 50% da 1ª Etapa para a 2ª, justo quando se passou a empregar o CEKF. Tal constatação por si só nos foi suficiente para afirmar que o emprego do CEKF ocasionou em uma melhoria significativa no estudo de caso. Já da 3ª Etapa para a 2ª ainda consideramos numericamente significativa a melhoria, já entre a 3ª e a 4ª não consideramos, motivo por encerrarmos o processo iterativo.

Ressaltamos também, o interessante comportamento da estimação com o CEKF constatado no caso em que os erros estavam longe de serem considerados uniformemente distribuídos (terceira etapa da seção 5.5). Tal situação que é desfavorável para a estimação dos parâmetros foi contrabalançada com o emprego do CEKF, sendo mais um bom argumento para o seu uso.

No estudo de caso adotado no presente vimos que a melhor função objetivo na primeira etapa foi:

$$J = \sum_{i=1}^{ne} \sum_{j=1}^{ny} \left[ \frac{dy_{ij}}{dt} - \frac{dg(x_i^{sim}; \theta)}{dt} \right]^2 \quad (6.1)$$

nas etapas seguintes onde houve um aprimoramento dos dados ainda testamos outras formulações, considerando a possibilidade da mudança na melhor formulação de função objetivo. Mas, isto não ocorreu, sendo de fato a melhor função objetivo em todos os casos a Equação 6.1. Isso reforça a importância das derivadas na melhoria da estimação. Para um novo caso de estudo, devido a particularidade dos processos, não podemos afirmar que o emprego das derivadas será sempre o melhor. No entanto, sugerimos o teste, pois existe considerável chance de ser o melhor, ainda com o respaldo do trabalho de Machado e Trierweiler (2004).

Um próximo passo a ser tomado para continuidade deste estudo é a aplicação em um caso real.

Por fim, o último comentário que abordamos é que no presente trabalho fizemos o ajuste da matriz Q de forma simples: tentativa e erro, devido à praticidade e bons resultados. No entanto, surge a questão: “E se a matriz Q fosse formada com um método mais avançado, considerando  $\hat{\sigma}_\theta$ , de maneira que os elementos da matriz Q já não seriam mais constantes e

sim funções. Como ficariam os resultados?”. A questão levantada é uma idéia interessante para um trabalho futuro.

## Referências Bibliográficas

ARNOLD, Kenneth J; BECK, James V. **Parameter Estimation: in engineering and science**. New York: John Wiley & Sons, 1977.

ASTRÖM, K. J.; WITTENMARK, B. **Computer Controlled Systems: Theory and Design**. 3. ed. Upper Saddle River, NJ, USA: Prentice Hall, 1997.

BARD, Yonathan. **Nonlinear Parameter Estimation**. New York and London: Academic Press, 1974.

BATES, D. M.; WATTS, D. G. **Nonlinear Regression Analysis and its Applications**. New York: John Wiley, 1988.

BAGAJEWLCS, Miguel. **Process Plant Instrumentation: Design and Upgrade**. Lancaster, Pennsylvania, USA: Technomic, 2001.

BOHLIN, Torsten. **Practical Grey-Box Proces Identification: Theory and applications**. Springer-Verlag London Limited, 2006.

BRUN, Roland; REICHERT, Peter; KÜNSCH, Hans R. Practical Identifiability analysis of large environmental simulation models, **Water Resources Research**, v. 37, n. 4, p. 1015-1030, abr. 2001.

CATEN, Carla ten; RIBEIRO, José Luis Duarte. **Estatística industrial**. Porto Alegre, RS: Programa de pós graduação em engenharia de produção, 2000. Escola de engenharia da UFRGS.

BECERRA, M.; GRIFFITHS, W; ROBERTS, D. Applying the extended Kalman filter to systems described by nonlinear differential algebraic equations. **Control Engineering Practice**, v. 9, n. 3, p. 267- 281, mar. 2001.

---

CHACHUAT, B. C. **Nonlinear and dynamic optimization: from theory to practice**. [S.l.]: Laboratoire d'Automatique, École Polytechnique Fédérale de Lausanne, winter semester 2006/2007.

EDGAR, T. F.; HIMMELBLAU, D. M.; LASDON, L. S. **Optimization of chemical process**. 2. ed. New York: MacGraw-Hill, 2001.

FRANCESCHINI, Gaia. **New Formulations for Model-Based Experiment Design and Application to a Biodiesel Production Process**. Thesis (Doctor of Chemical Engineering) - Imperial College London, London, may 2007.

FIKAR, Miroslav; MIKLES, Ján. **Process, Modelling, Identification and Control**. Berlin: Springer-Verlag, 2007.

HIMMELBLAU, David M. **Process Analysis by Statistical Methods**. New York: John Wiley & Sons, 1970.

HINDMARSH, Alan C.; STEPLEMAN, R. S.; et al. **Odepack, a Systemized Collection of ODE Solvers**. Amsterdam: North-Holland, 1983.

HOFFMANN, Christian. **Numerical methods for the discrimination of DAE models with applications in enzyme kinetics**. Thesis (Doctor of Physics) – University Heidelberg, Germany, dec. 15, 2005.

JAZWINSKI, A. H. **Stochastic Processes and Filtering Theory**. New York: Academic Press, 1970. v. 3.

LJUNG, L. **System Identification: Theory for the User**. 2. ed. Upper Saddle River, NJ, USA: Prentice Hall, 1999.

KALMAN, R. E. A New Approach to Linear Filtering and Prediction Problems, **Transactions of the ASME - Journal of Basic Engineering**, v. 82, p. 35-45, 1960.

MACHADO, Vinícius C .; TRIERWEILER, J. O. Which is the best criterion for identification of dynamic models? In: SYMPOSIUM ON DYNAMICS AND CONTROL OF PROCESS SYSTEMS, 7., 2004, Boston. **Anais...** Boston: International federation of automatic control (IFAC), 2004. v. 1. p. 178-188, CD-ROM.

MARCON, Simone Maria; TRIERWEILER, J. O.; SECCHI, A. R. EKF e CEKF: Comparação entre duas formulações do filtro de Kalman Estendido. In: CONGRESSO BRASILEIRO DE AUTOMAÇÃO, 14., 2002, Natal, RN. **Anais...** Natal, RN: SBA, 2002. v. 1, p. 739-744, CD-ROM.

MORETTIN, Luiza Gonzaga. **Estatística básica**. São Paulo: Makron Books, 2000. v. 2.

MUSKE, K. R., RAWLINGS, J.B. Nonlinear Moving Horizon State Estimation. **NATO ASI Series (Kluwer Academic)**, v. 293, p. 349-365, 1994.

---

PINTO, José Carlos; SCHWAAB, Márcio. **Análise de Dados Experimentais I: Fundamentos de Estatística e Estimação de Parâmetros**. 1. ed. Rio de Janeiro: E-pappers, 2007.

QUARTERONI, Alfio; SACCO, Riccardo; SALERI, Fausto. **Numerical Mathematics**. New York: Springer-Verlag, 2000.

ROBERTSON, D.G. et al. A moving horizon-based approach for least squares estimation. **AIChE Journal**, v. 42, n. 8, p. 2209-2224, 1995.

SALAU, Nina Paula Gonçalves; SECCHI, Argimiro Resende; TRIERWEILER, J. O. Five Formulations of Extended Kalman Filter: Which is the best for D-RTO?. In: European Symposium on Computer Aided Process Engineering (ESCAPE), 17., 2007, Bucharest. **Anais...** Bucharest: European Federation of Chemical Engineering, 2007. v. 1, p. T3-438, CD-ROM.

SIMON, Dan. **Optimal State Estimation**. Hoboken, New Jersey, USA: John Wiley & Sons, 2006.

SOROUSH, M. State and parameter estimations and their applications in process control. **Computers and chemical engineering**, v. 23, p. 229-245, 1998.

TONEL, G. **Aplicação dos Estimadores EKF, CEKF e MHE para Aplicações em Linha em Processos**. Dissertação (Mestrado em engenharia química) - Universidade Federal do Rio Grande do Sul, Brasil, 2007.