

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL  
INSTITUTO DE MATEMÁTICA  
PROGRAMA DE PÓS-GRADUAÇÃO EM MATEMÁTICA APLICADA

# **Problemas Inversos em Identificação de Biorreatores e Transporte de Radiação**

por

Michel Michelon

Profa. Dra. Liliane Basso Barichello  
Orientadora

Prof. Dr. Diego Eckhard  
Coorientador

Porto Alegre, Dezembro de 2016.

## CIP - CATALOGAÇÃO NA PUBLICAÇÃO

Michelon, Michel

Problemas Inversos em Identificação de Biorreatores e Transporte de Radiação / Michel Michelon.—Porto Alegre: PPGMAp da UFRGS, 2016.

104 p.: il.

Dissertação (mestrado) —Universidade Federal do Rio Grande do Sul, Programa de Pós-Graduação em Matemática Aplicada, Porto Alegre, 2016.

Orientadora: Barichello, Liliane Basso; Coorientador: Eckhard, Diego

Dissertação: Teoria de Transporte de Partículas Matemática Aplicada, Problemas Inversos, Teoria de Transporte, Sistemas Dinâmicos, Biorreatores Anaeróbicos

# Problemas Inversos em Identificação de Biorreatores e Transporte de Radiação

por

Michel Michelon

## Mestre em Matemática Aplicada

Linha de Pesquisa: Teoria de Transporte de Partículas

Orientadora: Profa. Dra. Liliane Basso Barichello

Coorientador: Prof. Dr. Diego Eckhard

Banca examinadora:

Prof. Dr. Helcio Rangel Barreto Orlande  
POLI - UFRJ

Prof. Dr. Leonardo Dagnino Chiwiacowsky  
CCET - UCS

Prof. Dr. João Batista da Paz Carvalho  
PPGMAp - UFRGS

Dissertação apresentada em  
2 de dezembro de 2016.

Prof. Dr. Carlos Hoppen  
Coordenador

*"Learning is the only thing  
the mind never exhausts,  
never fears,  
and never regrets."*

*(Leonardo Da Vinci)*

## AGRADECIMENTOS

Agradeço a todos que contribuíram, direta ou indiretamente, para a realização deste trabalho. Devo, contudo, dedicar especial atenção a algumas pessoas, sem as quais não venceria essa etapa.

Aos meus orientadores Profa. Dra. Liliane Basso Barichello e Prof. Dr. Diego Eckhard, pela orientação, pelas aulas, pelo incentivo e pela confiança em mim e no trabalho que desenvolvia.

Aos meus pais, Denise e Valmor, tanto pelo modo como sempre primaram minha educação, quanto pela formação ética e moral que me foi dada.

A meu irmão, Diego, por sempre me receber de portas abertas em sua casa, permitindo que pudesse melhor aproveitar meus estudos.

Ao grande amigo César, por todo o indispensável auxílio prestado durante a graduação e pós-graduação.

A Letícia, por não permitir que meu medo levasse à estagnação, por compreender minha ausência e por sempre incentivar meus estudos.

A Francine, pela capacidade de transformar em luz todos aqueles momentos de trevas no meio do caminho, permitindo que eu tivesse forças para continuar.

A Ezequiel, por todo o incentivo e por estar sempre ao meu lado, me fazendo acreditar que seria capaz de chegar ao final desta difícil, porém gratificante etapa.

Aos colegas e amigos da pós-graduação, em especial a Cássio, Daniella, Lucas e Yuri, por todo o auxílio e discussões a respeito dos estudos.

A todos os professores da graduação e pós-graduação, pela contribuição e ensinamentos.

Ao PPGMAp e à UFRGS, pela oportunidade.

À CAPES, pelo auxílio financeiro.

## Sumário

<b>LISTA DE FIGURAS</b> . . . . .	<b>x</b>
<b>LISTA DE SIGLAS</b> . . . . .	<b>xiii</b>
<b>LISTA DE SÍMBOLOS</b> . . . . .	<b>xiv</b>
<b>LISTA DE TABELAS</b> . . . . .	<b>xvi</b>
<b>RESUMO</b> . . . . .	<b>xviii</b>
<b>ABSTRACT</b> . . . . .	<b>xix</b>
<b>1 INTRODUÇÃO</b> . . . . .	<b>1</b>
<b>2 PROBLEMAS INVERSOS: ASPECTOS FUNDAMENTAIS</b> . . . . .	<b>5</b>
2.1 Classificação dos problemas inversos . . . . .	6
2.2 Dificuldades na solução de problemas inversos . . . . .	8
<b>3 APROXIMAÇÃO POR MÍNIMOS QUADRADOS</b> . . . . .	<b>12</b>
3.1 O caso linear e as equações normais . . . . .	14
3.2 Caso não-linear e os métodos iterativos determinísticos . . . . .	16
3.2.1 Método do Gradiente . . . . .	17
3.2.2 Métodos de Newton . . . . .	18
3.2.2.1 Método de Newton Clássico . . . . .	20
3.2.2.2 Método de Gauss-Newton . . . . .	21

3.2.2.3	Método de Levenberg-Marquart . . . . .	22
3.2.3	Métodos <i>Trust-Region</i> . . . . .	25
<b>3.3</b>	<b>Análise estocástica do método para o caso linear</b> . . . . .	<b>26</b>
3.3.1	Qualidade da solução aproximada por mínimos quadrados . . . . .	28
3.3.2	Elipses de erro para as soluções aproximadas . . . . .	31
<b>4</b>	<b>BIORREATORES ANAERÓBICOS</b> . . . . .	<b>33</b>
4.1	Descrição do modelo matemático considerado . . . . .	37
4.2	Solução aproximada do problema direto pelo método de Runge-Kutta . . . . .	40
4.3	Resultados Numéricos . . . . .	42
4.3.1	Implementação do problema direto . . . . .	43
4.3.2	Análise das estimativas de $\theta$ pela solução de problemas inversos . . . . .	44
<b>5</b>	<b>TRANSPORTE DE PARTÍCULAS</b> . . . . .	<b>54</b>
5.1	Descrição do modelo considerado . . . . .	58
5.2	Solução em forma fechada para problema direto usando o método <i>ADO</i> . . . . .	62
5.3	Resultados Numéricos . . . . .	69
5.3.1	Geração de dados sintéticos a partir implementação do problema direto . . . . .	70
5.3.2	Análise do comportamento da função objetivo em relação ao albedo . . . . .	73



5.3.3	Resultados numéricos para a identificação do albedo . . . . .	75
5.3.3.1	Análise da média e variância quando da alteração do número de estimações consideradas . . . . .	79
5.3.4	Resultados numéricos para a identificação do albedo e grau de anisotropia	82
5.3.5	Resultados numéricos adicionais . . . . .	88
<b>6</b>	<b>CONSIDERAÇÕES FINAIS . . . . .</b>	<b>94</b>
	<b>REFERÊNCIAS BIBLIOGRÁFICAS . . . . .</b>	<b>96</b>

## LISTA DE FIGURAS

Figura 2.1	Caracterização de problemas diretos e inversos . . . . .	6
Figura 4.1	Diferentes tipos de biorreatores, de acordo com seu modo de operação . . . . .	35
Figura 4.2	Diagrama do processo de decomposição de matéria orgânica em um biorreator descontínuo . . . . .	36
Figura 4.3	Gráfico da simulação da saída de metano no biorreator anaeróbico	44
Figura 4.4	Matriz de covariância das estimativas para chute inicial exato e perturbação gaussiana com desvio padrão de 1% do valor máximo da saída do modelo usando LM . . . . .	47
Figura 4.5	Gráficos do modelo usando parâmetros corretos, aquele que gera o menor e aquele que gera o maior resíduos, na reconstrução com perturbação gaussiana com desvio padrão de 1% do valor máximo da saída do modelo direto usando LM . . . . .	48
Figura 4.6	Elipses de erro das estimativas para chute inicial exato e perturbação gaussiana com desvio padrão de 1% do valor máximo da saída do modelo usando LM . . . . .	49
Figura 4.7	Matriz de covariância das estimativas para chute inicial exato e perturbação gaussiana com desvio padrão de 10% do valor máximo da saída do modelo usando LM . . . . .	50

Figura 4.8	Gráficos do modelo usando parâmetros corretos, aquele que gera o menor e aquele que gera o maior resíduos, na reconstrução com perturbação gaussiana com desvio padrão de 10% do valor máximo da saída do modelo direto usando LSQNONLIN . . . . .	51
Figura 4.9	Elipses de erro das estimativas para chute inicial exato e perturbação gaussiana com desvio padrão de 10% do valor máximo da saída do modelo usando LSQNONLIN . . . . .	52
Figura 5.1	Gráfico da densidade radiativa $\rho(x)$ sobre a barra . . . . .	72
Figura 5.2	Gráfico da função objetivo $\mathcal{S}(\mathbf{x}, \mathbf{P})$ para $L$ conhecido, dados não perturbados e $c$ contínuo . . . . .	73
Figura 5.3	Gráficos da função objetivo $\mathcal{S}(x, \mathbf{P})$ para grau de anisotropia conhecido e ruídos da ordem de 0% (azul), 1% (verde) e 10% (vermelho) na saída do modelo . . . . .	75
Figura 5.4	Histograma dos valores obtidos para o parâmetro de albedo a partir de dados perturbados em 1% . . . . .	76
Figura 5.5	Intervalos de confiança para as estimativas na determinação do parâmetro de albedo a partir de dados perturbados em 1% . . . . .	77
Figura 5.6	Histograma dos valores obtidos para o parâmetro de albedo a partir de dados perturbados em 10% . . . . .	77
Figura 5.7	Intervalos de confiança para as estimativas na determinação do parâmetro de albedo a partir de dados perturbados em 10% . . . . .	78
Figura 5.8	Intervalos de Confiança para perturbações da ordem de 1% e 10, 20, 50 e 100 estimativas . . . . .	80

Figura 5.9	Intervalos de Confiança para perturbações da ordem de 10% e 10, 20, 50 e 100 estimativas . . . . .	81
Figura 5.10	Histograma de estimativas para o grau de anisotropia na identificação conjunta com dados perturbados em 1% . . . . .	83
Figura 5.11	Histograma de estimativas para a constante de albedo na identificação conjunta com dados perturbados em 1% . . . . .	84
Figura 5.12	Comparação entre a saída do problema direto com os parâmetros corretos e com um dos conjuntos de parâmetros estimados no caso de uma perturbação de 1% . . . . .	85
Figura 5.13	Histograma de estimativas para o grau de anisotropia na identificação conjunta com dados perturbados em 5% . . . . .	86
Figura 5.14	Histograma de estimativas para a constante de albedo na identificação conjunta com dados perturbados em 5% . . . . .	87
Figura 5.15	Comparação entre a saída do problema direto com os parâmetros corretos e com um dos conjuntos de parâmetros estimados no caso de uma perturbação de 5% . . . . .	88

## LISTA DE SIGLAS

ADO	<i>Analytical Discrete Ordinates Method</i>
COD	Demanda química de oxigênio
LM	Método de <i>Levenberg-Marquardt</i>
LSQ	Método <i>Trust-Region</i> LSQNONLIN
PD	Problema Direto
PI	Problema Inverso

## LISTA DE SÍMBOLOS

$\mathcal{A}(\cdot)$	Modelo não necessariamente linear
$Y$	Saída de um modelo
$P$	Conjunto de parâmetros
$P_0$	Conjunto de parâmetros que descreve corretamente a saída de um modelo
$Y_{observado}$	Conjunto de dados experimentais
$P_{observado}$	Solução observada de um problema inverso
$Y_{aproximado}$	Modelo aplicado à solução aproximada do PI
$\epsilon, \delta$	Perturbações
$R(\cdot)$	Vetor de resíduos
$\mathcal{S}(\cdot)$	Função objetivo dos mínimos quadrados
$\nabla$	Gradiente
$\mathbf{A}$	Modelo linear
$P_k$	k-ésima iteração na determinação dos parâmetros
$\mathbf{J}$	Matriz Jacobiana
$\mathbf{H}$	Matriz Hessiana
$E[.]$	Valor esperado para uma variável aleatória
$\sigma_0^2(\cdot), var(\cdot)$	Variância de uma variável aleatória
$\sigma_0(\cdot, \cdot) = cov(\cdot, \cdot)$	Covariância entre duas variáveis aleatórias
$\bar{\mathcal{X}}$	Média de um conjunto de amostras de uma variável aleatória
$\sigma^2(\cdot)$	Aproximação da variância
$\sigma(\cdot, \cdot)$	Aproximação da covariância
$\mathbf{W}$	Vetor de perturbações gaussianas com média zero
$\hat{P}$	Solução aproximada de um problema inverso
$\mathcal{N}(m, n)$	Distribuição normal com média $m$ e variância $n$
$\chi^2(n)$	Distribuição qui-quadrada com $n$ graus de liberdade
$\mathbf{C}$	Matriz de covariância
$x_1$	Concentração de bactérias acidogênicas
$x_2$	Concentração de bactérias metanogênicas
$S_1$	Concentração de demanda química de oxigênio
$S_2$	Concentração de ácidos graxos voláteis

$S^{in}$	Concentração de influentes
$D$	Coefficiente de diluição dos influentes
$k_1$	Coefficiente de degradação de COD
$k_2$	Coefficiente de produção de ácidos graxos voláteis
$k_3$	Coefficiente de consumo de ácidos graxos voláteis
$K_{S_i}$	Coefficiente de saturação do $S_i$ correspondente
$v_i(S_i(\cdot))$	Taxas de crescimento específico das bactérias
$q_M(t)$	Taxa de fluxo de metano
$k_6$	Taxa de saída de gás
$\theta$	Vetor contendo os parâmetros do biorreator
$\mathbf{X}$	Vetor das componentes espaciais de uma partícula
$\mathbf{v}$	Vetor contendo as componentes de velocidade de uma partícula
$\Omega$	Vetor unitário de direção
$v$	Módulo da velocidade
$f(\mathbf{X}, E, \Omega, t)$	Função de distribuição
$N(\mathbf{X}, E, \Omega, t)$	Densidade angular de nêutrons
$\sigma(\mathbf{X}, E)$	Seção de choque macroscópica
$\sigma(\mathbf{X}, E' \rightarrow E, \Omega' \rightarrow \Omega)$	Seção de choque diferencial
$f(\mathbf{X}, E' \rightarrow E, \Omega' \rightarrow \Omega)$	Probabilidade de transferência
$\Psi(\mathbf{X}, \Omega, E, t)$	Fluxo angular de nêutrons
$v\Omega \cdot \nabla N(\mathbf{X}, \Omega, E, t)$	Termo de fuga
$c$	Parâmetro de albedo
$P_l(\cdot)$	Polinômio de Legendre de ordem $l$
$\beta_l$	Coefficiente da lei de espalhamento na forma binomial
$L$	Grau de anisotropia
$\mathbf{F}_i(\cdot)$	Condições de contorno do problema de transporte
$S(\cdot, \cdot), Q(\cdot, \cdot)$	Fonte externa de radiação
$\nu$	Constante de separação da solução hipotética no método ADO
$\Phi(\nu, \pm\mu_i)$	Função considerada na solução do método ADO
$\rho(x)$	Densidade radiativa
$\mathbf{I}$	Matriz Identidade
$t$	Tempo

## LISTA DE TABELAS

Tabela 4.1	Distribuição estimada dos produtos finais em um biorreator descontínuo . . . . .	37
Tabela 4.2	Média, variância e erro relativo das estimativas para chute inicial exato e perturbação gaussiana com desvio padrão de 1% do valor máximo da saída do modelo . . . . .	46
Tabela 4.3	Média, variância e erro relativo das estimativas para chute inicial exato e perturbação gaussiana com desvio padrão de 10% do valor máximo da saída do modelo . . . . .	50
Tabela 5.1	Comparação dos resultados numéricos com os apresentados na literatura . . . . .	71
Tabela 5.2	Valores de $c$ que minimizam $S(x, \mathbf{P})$ quando o grau de anisotropia é conhecido . . . . .	74
Tabela 5.3	Resultados médios para a identificação da constante de albedo com dados perturbados em 1% . . . . .	76
Tabela 5.4	Resultados médios para a identificação da constante de albedo com dados perturbados em 10% . . . . .	78
Tabela 5.5	Resultados médios com diferentes números de estimativas na determinação do albedo com perturbações da ordem de 1% . . . .	80
Tabela 5.6	Resultados médios com diferentes números de estimativas na determinação do albedo com perturbações da ordem de 10% . . . .	81
Tabela 5.7	Resultados médios para a identificação da constante de albedo e grau de anisotropia com dados perturbados em 1% . . . . .	84



Tabela 5.8	Resultados médios para a identificação da constante de albedo e grau de anisotropia com dados perturbados em 5% . . . . .	87
Tabela 5.9	Resultados médios para a identificação da constante de albedo e grau de anisotropia no primeiro problema, considerando apenas três pontos e perturbação da ordem de 1% . . . . .	89
Tabela 5.10	Resultados médios para a identificação da constante de albedo e grau de anisotropia no primeiro problema, considerando apenas três pontos e perturbação da ordem de 5% . . . . .	89
Tabela 5.11	Resultados médios para a identificação da constante de albedo e grau de anisotropia no primeiro problema, considerando nove pontos e perturbação da ordem de 1% . . . . .	90
Tabela 5.12	Resultados médios para a identificação da constante de albedo e grau de anisotropia no primeiro problema, considerando nove pontos e perturbação da ordem de 5% . . . . .	91
Tabela 5.13	Resultados médios para a identificação da constante de albedo no segundo problema considerado . . . . .	92
Tabela 5.14	Resultados médios para a identificação da constante de albedo no terceiro problema considerado . . . . .	92

## RESUMO

Neste trabalho, apresenta-se um estudo preliminar sobre problemas inversos com aplicação na identificação de parâmetros em biorreatores e transferência radiativa. A identificação se baseia na busca de soluções aproximadas a partir do critério dos mínimos quadrados. Para tal, métodos determinísticos implícitos são analisados e posteriormente utilizados, destacando-se o método de Levenberg-Marquardt. No caso do biorreator anaeróbico, discute-se brevemente a teoria acerca dos processos internos e, a partir da implementação do modelo matemático de Antonelli, considerada a solução exata do problema direto, procede-se com a adição de diferentes níveis de ruídos gaussianos e análise de seus efeitos na reconstrução dos parâmetros do modelo. Iniciando a estimação sempre pelos parâmetros corretos, conhecidos *a priori*, pode-se perceber que é possível proceder com a estimação destes quando a magnitude do erro é da ordem de 1%. No caso de perturbações maiores nos dados de saída do modelo, contudo, embora não seja possível estimar os parâmetros, a simulação continua confiáveis. No que tange à transferência radiativa, usa-se a solução em forma fechada em relação à variável espacial da equação íntegro-diferencial linear de Boltzmann pelo método das Ordenadas Discretas Analíticas como solução exata do problema direto. Investiga-se a eficiência do método de Levenberg-Marquardt na estimação dos parâmetros de albedo e grau de anisotropia em problemas de transporte de radiação. É possível concluir que bons resultados são obtidos na identificação do parâmetro de albedo, mas que o grau de anisotropia é mais suscetível às perturbações.

## ABSTRACT

In this report, a preliminary study on the application of inverse problems in parameter estimation problems in transport of radiation and anaerobic bioreactors is presented. This estimation is based on the search of approximate solution to the inverse problem by the least squares criterion. Implicit deterministic methods are analyzed and subsequently used, mainly the Levenberg-Marquardt method. The anaerobic bioreactor theory is discussed and the mathematical model of Antonelli is considered the exact solution of the direct problem. After, the effects of the addition of different levels of Gaussian noise on the reconstruction of the model parameters is analysed. Starting the estimation always by the correct parameters, known *a priori*, it is possible to verify that the correct estimation can be achieved when the magnitude of the error is 1%. In the case of larger errors in the model output data, however, although it is not possible to correctly estimate the parameters, the simulation remains reliable. In the case of transport of neutral particles, the closed solution of the linear integral-differential Boltzmann equation obtained by the Analytic Discrete Ordinates method is considered as the exact solution of the direct problem. The efficiency of Levenberg-Marquardt method for estimating the albedo and degree of anisotropy is then investigated. It is possible to conclude that good results are obtained for the reconstruction of albedo, even with larger orders of error, but the degree of anisotropy is more susceptible to the perturbations.

# 1 INTRODUÇÃO

Uma questão de importância sempre atual e de interesse mundial é a prospecção de petróleo. Para a otimização de processos e obtenção de progressos nas pesquisas, é um procedimento natural que, por exemplo, a partir da análise de dados obtidos por medições de sondas nucleares, antes da perfuração de poços, ou mesmo, após perfurados, para monitoramento de suas condições, se busque estimar as características das rochas [28]. Por outro lado, o crescimento industrial e agrícola das últimas décadas trouxe consigo o problema do aumento da poluição e geração de lixo, com graves consequências ao meio ambiente, o que conduziu à criação, pelos órgãos ambientais, de normas cada vez mais rígidas em relação à quantidade de matéria orgânica despejada no meio ambiente pelos efluentes industriais e municipais [11]. Busca-se hoje desenvolver uma noção de economia verde, que alie o desenvolvimento tecnológico e econômico ao desenvolvimento sustentável, a fim de reduzir os impactos ambientais decorrentes das atividades produtivas. Nesse sentido, o tratamento de resíduos orgânicos por biorreatores passa a desempenhar papel importante, já que alia o tratamento de efluentes à produção de combustível renovável na forma de biogás. Estudar, portanto, meios de identificar os parâmetros internos de biorreatores, de modo a otimizar a produção de gás metano e, portanto, energia gerada, se torna fundamental [11, 12, 22].

Os dois processos descritos envolvem aspectos que caracterizam uma classe de problemas chamados problemas inversos. De fato, esta classe de problemas está relacionada a diversas áreas da ciência, dentre as quais pode-se citar, além dos problemas apresentados, a reconstrução de imagens médicas, como tomografias [10, 38], eletrocardiogramas e ultrassom [24, 32]; explorações sísmicas [39, 58], calibração de modelos econômicos [23], determinação da intensidade de fontes de radiação, em problemas de segurança, propriedades radiativas do meio ou os co-

eficientes de absorção e de espalhamento em tomografia [54, 55], determinação de parâmetros atmosféricos, a partir da intensidade de radiação medida [16], entre outras. Além disso, há problemas em engenharia nos quais, devido à dificuldade de medir determinados parâmetros, é usual proceder com a estimação destes [43]. Desse modo, são problemas de ordem multidisciplinar que têm atraído interesse de pesquisadores por suas implicações em fatores sociais e econômicos e também por propiciarem desenvolvimento de novas tecnologias.

Neste trabalho, tem-se como objetivo realizar um estudo preliminar sobre problemas inversos com aplicação na identificação de parâmetros em biorreatores e transferência radiativa. Para tal, apresenta-se a formulação geral de problemas inversos e procede com a estimação de parâmetros pelo critério dos mínimos quadrados. Para tal, métodos determinísticos implícitos são analisados e posteriormente utilizados, destacando-se o método de Levenberg-Marquardt e métodos do tipo *Trust-Region*.

Dessa forma, o texto está estruturado em seis capítulos. Neste primeiro capítulo, apresenta-se uma visão geral do trabalho que será desenvolvido, bem como as motivações e objetivos do estudo conduzido. No segundo capítulo, apresenta-se um conjunto de definições e caracterizações importantes a respeito dos problemas inversos, bem como as dificuldades associadas à solução deste tipo de problema. No terceiro capítulo, define-se a abordagem utilizada para a busca de uma solução aproximada para os problemas inversos, o critério dos mínimos quadrados, que se baseia na minimização dos resíduos quadráticos entre os dados observados e os modelos considerados para os problemas diretos associados. Os casos linear e não-linear são abordados, apresentando-se métodos determinísticos implícitos para a minimização no caso não-linear. Além disso, a análise da qualidade das estimativas a partir dessa metodologia nos casos lineares é observada do ponto de vista estatístico. No quarto capítulo, após breve caracterização e descrição dos processos bioquímicos internos

em um biorreator, define-se o modelo considerado como problema direto, onde se escolhe utilizar o modelo de Antonelli, um modelo de média complexidade obtido a partir de um balanço de massa. Discute-se, então, a implementação do modelo em tempo discreto a partir do método de Runge-Kutta de quarta ordem. Sobre a saída do modelo discreto, são adicionadas perturbações gaussianas de ordem conhecida e, após a repetição do experimento diversas vezes, a fim de coletar informações estatisticamente relevantes, procede-se com a análise dos resultados da estimação dos parâmetros do biorreator a partir dos dados perturbados. Para a estimação, usa-se como vetor inicializador sempre o conjunto de parâmetros que, conhecida-mente, leva à saída não-perturbada do modelo e, como resultado, observa-se que, no caso de pequenas perturbações, é possível estimar os parâmetros do modelo pois as estimativas se aproximam, em média, dos valores corretos. No caso de perturbações de ordem maior, as estimativas não são confiáveis, mas, mesmo assim, a simulação condiz com o resultado esperado. No quinto capítulo, introduz-se a teoria de transporte de partículas neutras de forma breve. Após, apresenta-se o modelo considerado como problema direto, a equação íntegro-diferencial de transporte linear, que é a versão linear da equação de Ludwig Boltzmann. A equação de Boltzmann foi derivada, originalmente, para aplicação em teoria cinética dos gases, a fim de tratar fenômenos em dinâmica de gases rarefeitos [25], sendo posteriormente aplicada para o tratamento de problemas de transferência radiativa e transporte de partículas neutras, sendo hoje usada como modelo matemático em diversas aplicações, desde tomografias [3, 5] até caracterização de vegetações [14]. Uma solução por ordenadas discretas para a equação de transporte é determinada, usando o método *ADO*, e passa a ser considerada a solução fechada exata do problema. De modo análogo ao capítulo anterior, sobre a saída do modelo são adicionadas perturbações gaussianas de ordem conhecida e deseja-se, então, investigar a qualidade das estimativas para o albedo e grau de anisotropia para diferentes níveis de ruído. A determinação desses parâmetros é importante em problemas inversos onde a caracterização do meio é

importante, como em problemas de tomografia ou sensoriamento. Conclui-se que o parâmetro de albedo pode ser estimado com boa precisão mesmo nos casos em que há considerável quantidade de ruído nos dados mas que o mesmo não pode ser afirmado na estimação do grau de anisotropia. O último capítulo trata das considerações finais a respeito dos resultados encontrados e apresenta perspectivas de estudo futuras dentro dos tópicos apresentados.

## 2 PROBLEMAS INVERSOS: ASPECTOS FUNDAMENTAIS

Neste capítulo, trata-se de tópicos fundamentais relativos aos problemas inversos, tal como a conceituação e classificações, ressaltando as dificuldades de solução. Antes, contudo, faz-se necessário tratar e caracterizar aquele que se denomina problema direto.

Conforme aponta Tarantola [56], “cientistas e engenheiros desejam, frequentemente, relacionar um determinado conjunto de parâmetros físicos,  $P$ , que caracterizam um modelo, a observações coletadas, que constituem um conjunto de dados,  $Y$ ”. Assumindo que se compreenda adequadamente a física aplicada ao sistema em estudo, é possível considerar uma determinada relação *mathcal{A}* entre o conjunto de parâmetros e os dados coletados. Ou seja,

$$\mathcal{A}(P) = Y, \quad (2.1)$$

onde  $Y$  pode depender continuamente do tempo/espço ou se referir a uma coleção discreta de observações.

Quando  $P$  e  $Y$  são funções, é usual referir-se a  $\mathcal{A}$  como um operador. O operador  $\mathcal{A}$ , por sua vez, pode se apresentar de diversas maneiras, como, por exemplo, uma equação diferencial ordinária ou parcial ou, ainda, uma equação integral. Quando esse operador é capaz de representar com suficiente exatidão o comportamento do sistema em estudo, é comum chamá-lo de modelo matemático do sistema em estudo.

Define-se, então, **problema direto (PD)** como aquele em que se deseja encontrar  $Y$  a partir de  $P$  dado. Em outras palavras, se deseja encontrar a saída ou efeitos de um sistema, conhecidos os parâmetros de entrada e considerado um determinado modelo matemático que descreva o comportamento interno do mesmo. Um



**problema inverso (PI)**, por sua vez, é aquele em que, conhecido um determinado conjunto de dados  $\mathbf{Y}$ , deseja-se determinar o conjunto de parâmetros  $\mathbf{P}$ , ou parte deste, que, aplicado ao modelo, conduz aos dados observados. Pode-se entender ainda como aquele problema em que, conhecido o efeito de saída de um determinado sistema, deseja-se encontrar os parâmetros ou, ainda, o comportamento que conduziu a essa resposta [15]. A Figura 2.1 procura caracterizar, de forma sucinta, as definições de problemas direto e inverso.



Figura 2.1: Caracterização de problemas diretos e inversos

## 2.1 Classificação dos problemas inversos

Deve-se considerar que, por haver mais de uma maneira de formular e resolver problemas inversos, haja vista que a própria “causa” a determinar pode ser usada para sua distinção [13], estes podem ser, também, classificados de diversas formas [15, 54–56]. Campos Velho [13], por exemplo, indica que algumas das categorias de classificação possíveis são em relação à natureza da solução, em relação à

natureza matemática ou estatística do método empregado para a solução ou mesmo em relação à natureza da propriedade estimada.

Beck e colaboradores [8] classificam os problemas inversos em relação à natureza de sua solução. Essa caracterização, separa os problemas inversos basicamente em dois grupos: discretos ou contínuos. O problema inverso discreto, ou **problema de estimação de parâmetros**, é aquele em que se faz necessário determinar apenas um número finito  $n$  de parâmetros que definem o modelo. Estes parâmetros podem estar relacionados diretamente a alguma entidade física relevante para o problema físico estudado ou então serem coeficientes ou outras constantes em uma relação funcional que descreva o processo em estudo [56]. Nesse caso, o conjunto de parâmetros  $\mathbf{P}$  pode ser expresso como um vetor de dimensão  $n$  e, de modo análogo, se houver um número finito  $m$  de dados experimentais à disposição,  $\mathbf{Y}$  pode ser representado como um vetor de dimensão  $m$ . O problema inverso contínuo, ou **problema de estimação de funções**, por sua vez, se dá quando temos parâmetros que representam funções dependentes de tempo e espaço, ou seja, que dependem de dados contínuos. Esse tipo de problema inverso costuma ser mais complicado que o de estimação de parâmetros e, por esse motivo, tende a ser aproximado por uma discretização do problema original [56].

Levando em consideração a natureza dos métodos matemáticos usados na solução dos problemas inversos, Colaço et al. [18] indicam que os mesmos podem ser classificados segundo duas abordagens: métodos inversos puros e métodos baseados em otimização. Com base nessa classificação, Silva Neto e Campos Velho [54] se referem às formulações implícitas e explícitas para os problemas inversos.

Nesse contexto, a **formulação explícita** de um problema inverso refere-se à manipulação direta das equações usadas na modelagem do problema direto, de modo a obter-se uma expressão analítica na qual as incógnitas a determinar apareçam de forma explícita [54]. A **formulação implícita**, por sua vez, consiste na

“minimização de uma função objetivo, também denominada função custo (...). O ponto focal da busca da solução do problema inverso passa então a ser o emprego de métodos determinísticos, métodos estocásticos e/ou métodos híbridos para a minimização da função objetivo” [54].

## 2.2 Dificuldades na solução de problemas inversos

Deve-se considerar que, em geral, o conjunto completo de dados reais,  $\mathbf{Y}$ , é desconhecido, tendo-se acesso apenas a um conjunto de observações experimentais,  $\mathbf{Y}_{\text{observado}}$ , suscetível à presença de ruídos, que podem estar relacionados a fatores externos como erro na leitura dos dados ou erros de arredondamento numérico [56]. As medidas experimentais, portanto, consistem geralmente de um “conjunto de observações de um experimento ‘perfeito’ acrescidos de um componente de erro” [56]. Ou seja, considerando-se

$$\mathbf{Y} = \mathcal{A}(\mathbf{P}_0), \quad (2.2)$$

em que  $\mathbf{Y}$  satisfaz de modo exato ao modelo proposto para um conjunto de parâmetros,  $\mathbf{P}_0$ , considerada a **solução exata** do problema inverso, implicando que os dados experimentais sejam dados por

$$\mathbf{Y}_{\text{observado}} = \mathbf{Y} + \boldsymbol{\epsilon}, \quad (2.3)$$

onde  $\boldsymbol{\epsilon}$  representa uma perturbação nos dados reais. Um outro conjunto de parâmetros,  $\mathbf{P}_{\text{observado}}$ , passa então a satisfazer os dados experimentais com o modelo proposto, ou seja,

$$\mathbf{Y}_{\text{observado}} = \mathcal{A}(\mathbf{P}_{\text{observado}}), \quad (2.4)$$

de modo que  $\mathbf{P}_{\text{observado}}$  representa a **solução observada** do problema inverso. É importante ressaltar que nem sempre existe, contudo, um conjunto  $\mathbf{P}_{\text{observado}}$  que atenda 2.4.

Uma série de dificuldades no tratamento dos problemas inversos advém, contudo, justamente dos ruídos presentes nas observações. Uma característica importante que diferencia os problemas diretos dos inversos é que o segundo, em geral, é mal-posto no sentido de Hadamard. Citando Cézaro [15], um problema de qualquer natureza “é dito bem-posto no sentido de Hadamard se satisfaz as condições de **existência, unicidade e dependência contínua dos dados iniciais.**” Quando qualquer uma dessas condições não é satisfeita, o problema é então dito mal-posto nesse sentido. Ao proceder com uma análise dos problemas inversos em relação a estas condições, percebe-se que:

1. **Existência:** um problema inverso nem sempre admite uma solução observada, ou seja, nem sempre será possível determinar um conjunto de parâmetros  $\mathbf{P}$  que, aplicado ao modelo considerado, conduza às observações experimentais, o que pode decorrer das perturbações presentes nos mesmos ou da imprecisão do modelo utilizado para explicar o fenômeno, entre outros. Nesse caso, faz-se necessário flexibilizar a ideia de solução, procurando-se uma aproximação da solução observada. Define-se **solução aproximada**,  $\hat{\mathbf{P}}$ , como aquela que, quando aplicada no modelo considerado, retorna um conjunto  $\mathbf{Y}_{aproximado}$ , ou seja,

$$\mathbf{Y}_{aproximado} = \mathcal{A}(\hat{\mathbf{P}}), \quad (2.5)$$

que é o mais próximo possível de  $\mathbf{Y}_{observado}$ , de modo que

$$\mathbf{Y}_{observado} = \mathbf{Y}_{aproximado} + \boldsymbol{\delta}, \quad (2.6)$$

onde  $\boldsymbol{\delta}$  representa um resíduo.

2. **Unicidade:** mesmo no caso em que o problema inverso apresente solução, contudo, esta geralmente não é única [55, 56]. Isso significa que pode existir uma infinidade de conjuntos de parâmetros diferentes de

$\mathbf{P}_0$  que, quando aplicados ao modelo, estimem perfeitamente as observações experimentais  $\mathbf{Y}_{observado}$  [56]. Nesse caso, Silva Neto [55] aponta que é necessário buscar restrições adicionais que obriguem a unicidade da solução, já que a falta de unicidade, em geral, está relacionada à falta de informação.

3. **Dependência contínua dos dados iniciais:** a última condição está relacionada à dependência da solução do problema inverso aos dados. Isso significa que, em muitos problemas inversos, o nível de ruídos  $\epsilon$  presente nas observações, mesmo quando de pequena magnitude, pode conduzir a uma solução observada do problema inverso,  $\mathbf{P}_{observado}$ , que apresenta pouca ou nenhuma correspondência com a solução exata,  $\mathbf{P}_0$ . No caso de sistemas lineares, essa condição é traduzida pelo condicionamento do operador  $\mathcal{A}$ . Um exemplo simples desse tipo de problema é dado pelo sistema linear

$$\begin{bmatrix} 1 & 1 \\ 1 & 1,001 \end{bmatrix} \mathbf{P} = \begin{bmatrix} 2 \\ 2 \end{bmatrix},$$

mal-condicionado. A solução exata do problema inverso é dada por  $\mathbf{P}_0 = (2, 0)^T$ . Supondo, contudo, que os dados tenham sido obtidos experimentalmente, sendo expressos por  $\mathbf{Y}_{observado} = (2; 2, 001)^T$ . Ao resolver o problema inverso e estimar os parâmetros, temos, então uma solução observada  $\mathbf{P}_{observado} = (1, 1)^T$ , pouco relacionada à solução exata.

Percebe-se, então, que tanto os ruídos nos dados quanto a própria natureza dos problemas inversos levam a dificuldades matemáticas e numéricas para sua solução.

Ao longo desse trabalho, consideradas as caracterizações acima apresentadas, serão tratados problemas inversos discretos, formulados implicitamente como

problemas de minimização, cuja função objetivo a minimizar é dada pela soma dos resíduos quadráticos entre os dados observados e previsões teóricas dadas pelo problema direto. Esta estratégia, capaz de fornecer uma solução aproximada para o problema inverso, é conhecida como aproximação por Mínimos Quadrados, e será abordada no próximo capítulo.

### 3 APROXIMAÇÃO POR MÍNIMOS QUADRADOS

Considere um modelo no qual determinado efeito de saída  $Y$  é dado em função de um conjunto de parâmetros  $P_0$ , ou seja,

$$\mathbf{Y} = \mathcal{A}(\mathbf{P}_0), \quad (3.1)$$

e que estejam disponíveis  $m$  dados experimentais na forma

$$\mathbf{Y} = \begin{bmatrix} y_1 & y_2 & y_3 & \dots & y_m \end{bmatrix}^T, \quad (3.2)$$

deseja-se estimar o vetor de parâmetros  $P \in \mathbb{R}^n$ , dado por

$$\mathbf{P} = \begin{bmatrix} p_1 & p_2 & p_3 & \dots & p_n \end{bmatrix}^T, \quad (3.3)$$

com  $n \leq m$ , tal que (3.1) seja satisfeito. Em outras palavras, deseja-se encontrar a solução de um problema inverso discreto.

Caso o problema inverso apresente solução exata, tem-se que

$$\mathbf{P} = \mathbf{P}_0. \quad (3.4)$$

Sabe-se, contudo, que este problema inverso pode não apresentar uma solução exata. Supondo então que se deseje encontrar a melhor solução aproximada possível. Neste capítulo, apresenta-se uma metodologia, referenciada na literatura como aproximação pelo **critério dos mínimos quadrados** [56], para a determinação desta solução aproximada.

Seja o vetor de resíduos

$$\mathbf{R}(\mathbf{P}) = \begin{bmatrix} r_1(\mathbf{P}) & r_2(\mathbf{P}) & r_3(\mathbf{P}) & \dots & r_m(\mathbf{P}) \end{bmatrix}^T, \quad (3.5)$$

em que

$$\mathbf{R}(\mathbf{P}) = Y - \mathcal{A}(\mathbf{P}), \quad (3.6)$$

onde  $\mathbf{Y}$  representa o vetor contendo os dados disponíveis e  $\mathcal{A}(\mathbf{P})$  a solução do problema direto para um determinado vetor de parâmetros  $\mathbf{P}$ . Define-se a **função objetivo dos mínimos quadrados**, ou, ainda, funcional dos mínimos quadrados como

$$\mathcal{S}(\mathbf{P}) = \frac{1}{2} \|\mathbf{R}(\mathbf{P})\|_2^2 = \frac{1}{2} \sum_{i=1}^m r_i^2(\mathbf{P}). \quad (3.7)$$

A **solução aproximada por mínimos quadrados** para o problema inverso proposto, é, então, aquela que gera um vetor de parâmetros  $\mathbf{P}$  capaz de minimizar (3.7), ou seja

$$\mathbf{P} = \arg \min_{\mathbf{P}} [\mathcal{S}(\mathbf{P})]. \quad (3.8)$$

Há diversas maneiras de encontrar esse argumento mínimo, que dependem exclusivamente da natureza do problema estudado. Conforme citado anteriormente, pode-se utilizar métodos determinísticos ou, então, métodos de cunho estocástico, como os métodos de Monte Carlo ou Algoritmos Genéticos. Este trabalho, contudo, se atém a estudar apenas métodos determinísticos. Nesse caso, sabe-se que o mínimo ocorre quando o gradiente de  $\mathcal{S}(\mathbf{P})$ , definido como

$$\nabla \mathcal{S}(\mathbf{P}) = \left[ \frac{\partial \mathcal{S}(\mathbf{P})}{\partial p_1} \quad \frac{\partial \mathcal{S}(\mathbf{P})}{\partial p_2} \quad \frac{\partial \mathcal{S}(\mathbf{P})}{\partial p_3} \quad \dots \quad \frac{\partial \mathcal{S}(\mathbf{P})}{\partial p_n} \right]^T, \quad (3.9)$$

é nulo, ou seja, quando

$$\frac{\partial \mathcal{S}(\mathbf{P})}{\partial p_j} = 0, \quad j = 1, \dots, n. \quad (3.10)$$

Embora este trabalho se preocupe, sobretudo, com a aplicação do método em operadores não-lineares, devido à importância e quantidade de resultados interessantes encontrados na literatura para os problemas em que  $\mathcal{A}(\mathbf{P})$  é linear, ambos os casos serão tratados nas próximas seções.



### 3.1 O caso linear e as equações normais

Uma classe de modelos matemáticos para a qual uma série de resultados úteis estão disponíveis são os sistemas lineares em relação aos parâmetros [56]. Esses sistemas são aqueles que obedecem as propriedades

$$\mathcal{A}(\mathbf{P}_1 + \mathbf{P}_2) = \mathcal{A}(\mathbf{P}_1) + \mathcal{A}(\mathbf{P}_2), \quad (3.11)$$

e

$$\mathcal{A}(\alpha\mathbf{P}) = \alpha\mathcal{A}(\mathbf{P}), \quad (3.12)$$

onde  $\alpha$  é escalar. No caso de um problema inverso linear discreto, é sempre possível escrever (3.1) como um sistema linear de equações algébricas [56], de modo que se pode escrever

$$\mathcal{A}(\mathbf{P}) = \mathbf{A}\mathbf{P} = \mathbf{Y}. \quad (3.13)$$

O caso linear dos mínimos quadrados é de interesse devido à grande quantidade de problemas na literatura que são lineares ou podem, então, ser linearizados.

Além disso, uma propriedade interessante que se observa no caso linear é que, a aplicação da condição dada por (3.10) conduz sempre a um sistema de  $n$  equações algébricas que dependem de  $P$ , de modo que é possível estabelecer uma solução aproximada determinística.

Considere o modelo (3.1) linear em relação aos parâmetros, ou seja, aquele que pode ser escrito como (3.13) e, substituindo essa definição em (3.7), tem-se

$$\mathcal{S}(\mathbf{P}) = [\mathbf{Y} - \mathcal{A}(\mathbf{P})]^T[\mathbf{Y} - \mathcal{A}(\mathbf{P})] = [\mathbf{Y} - \mathbf{A}\mathbf{P}]^T[\mathbf{Y} - \mathbf{A}\mathbf{P}] \quad (3.14)$$

ou, ainda,

$$\mathcal{S}(\mathbf{P}) = \mathbf{P}^T \mathbf{A}^T \mathbf{A} \mathbf{P} - (\mathbf{P}^T \mathbf{A}^T \mathbf{Y} + \mathbf{Y} \mathbf{A} \mathbf{P}) + \mathbf{Y}^T \mathbf{Y} \quad (3.15)$$

onde  $\mathcal{S}(\mathbf{P})$  é uma função quadrática. Tem-se que  $\mathbf{A}^T \mathbf{A}$  é simétrica e, quando for, também, positiva definida, então a função objetivo admite um mínimo global único. Deseja-se, então, determinar qual o vetor de parâmetros capaz de minimizá-lo ou, ainda, encontrar  $\mathbf{P}$  tal que

$$\nabla S(\mathbf{P}) = 0. \quad (3.16)$$

Aplicando-se essa condição, pode-se facilmente encontrar que

$$0 = 2\mathbf{A}^T \mathbf{A} \mathbf{P} - (\mathbf{A}^T \mathbf{Y} + \mathbf{A}^T \mathbf{Y}) + 0, \quad (3.17)$$

e, por fim,

$$\mathbf{A}^T \mathbf{A} \mathbf{P} - \mathbf{A}^T \mathbf{Y} = 0, \quad (3.18)$$

que pode ser escrito como o sistema de equações lineares dado por

$$\mathbf{A}^T \mathbf{A} \mathbf{P} = \mathbf{A}^T \mathbf{Y}. \quad (3.19)$$

As equações dadas por (3.19) são chamadas de **equações normais** do sistema. Esse sistema sempre admite soluções e, quando  $\mathbf{A}^T \mathbf{A}$  for inversível, é possível determinar uma solução aproximada única via mínimos quadrados na forma

$$\mathbf{P} = (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T \mathbf{Y}. \quad (3.20)$$

Com isso demonstra-se que, quando o sistema é linear em relação aos parâmetros, é possível determinar uma solução analítica em relação aos mesmos. A solução por Equações Normais, contudo, não é recomendada em aplicações computacionais, visto que o condicionamento de  $\mathbf{A}^T \mathbf{A}$  será ampliado quando o condicionamento do operador  $\mathbf{A}$  for alto. Dá-se preferência, nesse caso, a soluções como a obtida via Fatoração QR ou então por Decomposição SVD. Mais informações a respeito dessas metodologias de resolução podem ser encontradas na literatura [19].

Ressalta-se ainda que, no caso de a matriz  $\mathbf{A}$  do sistema ser quadrada e inversível, a solução do sistema de equações normais naturalmente equivale à solução

do problema  $\mathbf{A}\mathbf{P} = \mathbf{Y}$ , já que é possível estabelecer a relação

$$\mathbf{A}^{-1} = (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T, \quad (3.21)$$

onde a expressão do lado direito da igualdade recebe o nome de **pseudo-inversa** da matriz ou, ainda, inversa generalizada de Moore-Penrose [19] e existe mesmo no caso de  $\mathbf{A}$  não ser quadrada.

### 3.2 Caso não-linear e os métodos iterativos determinísticos

No caso de sistemas não-lineares, é comum que não seja possível encontrar uma solução diretamente do gradiente da função custo. Isso se deve ao fato de as derivadas

$$\frac{\partial r_i}{\partial p_j}, \quad (3.22)$$

que surgem em (3.10) serem dadas por funções nas quais não é possível isolar  $P$ .

Nesses casos, portanto, a solução aproximada por mínimos quadrados ou, ainda, o conjunto de parâmetros que satisfaz (3.8), passa a ser encontrado de forma iterativa, sendo refinado a cada passo. De modo genérico, temos que, dado um vetor inicializador, proceder-se-á com o processo iterativo

$$\mathbf{P} \approx \mathbf{P}_{k+1} = \mathbf{P}_k + (\Delta \mathbf{P})_k, \quad (3.23)$$

em que  $k$  indica o número da iteração e o **vetor de incrementos**  $\Delta \mathbf{P}$ , chamado na literatura de *shift vector* [56], deve ser escolhido de modo que, a cada passo,  $\mathbf{P}_k$  se aproxime mais daquele que minimiza a função objetivo.

Há mais de uma maneira de determinar o vetor de incrementos a cada iteração. Três formas muito utilizadas são os métodos conhecidos como o método do Gradiente, os métodos de Newton e os métodos de *Trust-Region*. Nas próximas seções descreve-se as ideias que fundamentam a escolha do *shift vector* em cada um desses métodos, apontando as vantagens e desvantagens de cada um deles.

### 3.2.1 Método do Gradiente

Um dos modos mais simples de determinar a atualização do vetor de parâmetros, de forma que este se aproxime cada vez mais daquele que minimiza (3.7) é o método do gradiente ou da descida máxima. A ideia é, a partir de uma escolha inicial livre, criar um processo iterativo no qual cada nova estimativa para o mínimo da função seja dada por um passo na direção de maior decréscimo da função na estimativa atual [40]. Nesse sentido, como o gradiente de  $\mathcal{S}$ , ou seja,  $\nabla\mathcal{S}(\mathbf{P})$ , representa a máxima variação no ponto aplicado, considera-se que um passo na direção oposta será, também, uma “descida máxima” em direção ao mínimo da função [27].

Um passo de magnitude do gradiente, contudo, pode ser exagerado e levar o método não convergir. Define-se, portanto, uma variável de *damping*, ou **fator de amortecimento**,  $\lambda > 0$ , cujo papel é regular o tamanho do passo dado na direção desse gradiente. Desse modo, a atualização do parâmetro passa a ser dada pelo processo iterativo

$$\mathbf{P}_{k+1} = \mathbf{P}_k - \lambda_k \nabla\mathcal{S}(\mathbf{P}_k), \quad (3.24)$$

que poderá, eventualmente, convergir para  $P$  que minimiza a função objetivo.

O comportamento do método do gradiente fica, portanto, atrelado à correta escolha do fator de amortecimento, já que é este que regula o tamanho dos passos no algoritmo, de forma que sejam pequenos o suficiente para garantir sua estabilidade. A escolha dessa variável pode ser fixa ou então variar dinamicamente a cada iteração, sendo que várias regras de atualização estão disponíveis na literatura [27, 40, 41].

Em linhas gerais, tem-se que o método do gradiente é simples, cada iteração é computada de forma menos complexa, em relação a outros métodos apresentados na literatura, e que, do ponto de vista teórico, existe uma sequência de passos que converge, caso exista, ao menos para um mínimo local - desde que  $\lambda$  seja suficientemente pequeno - já que cada atualização deverá caminhar em direção ao mínimo [19].

O maior problema relacionado ao método do gradiente está no número de iterações necessárias para que se atinja a convergência. Considerando um  $\lambda$  fixo pequeno o suficiente para garantir a convergência, é possível que um número considerável de iterações sejam necessárias até que o método convirja para o mínimo local. Desse modo, o método, quando usado isoladamente, é considerado “uma má escolha para problemas de otimização” [27] mas, quando usado em conjunto com outras ferramentas, aumenta as chances de se encontrar a melhor solução aproximada para o problema. Embora o método do gradiente não seja, neste trabalho, aplicado diretamente para a solução de problemas inversos, ele ajuda a compreender o comportamento do método de Levenberg-Marquardt, tratado na próxima seção.

### 3.2.2 Métodos de Newton

No caso dos métodos de Newton, a função objetivo é expandida em série truncada de Taylor e a direção de atualização do vetor passa a representar, então, a direção ótima em relação ao mínimo dessa aproximação em cada passo. Antes de proceder com uma descrição mais formal do método de Newton, é importante que algumas informações sejam reforçadas e novas definições sejam estabelecidas.

Considerando novamente a função objetivo dos mínimos quadrados, dada por

$$\mathcal{S}(\mathbf{P}) = \frac{1}{2} \|\mathbf{Y} - \mathcal{A}(\mathbf{P})\|_2^2 = \frac{1}{2} \|\mathbf{R}(\mathbf{P})\|_2^2, \quad (3.25)$$

tem-se que o gradiente de  $\mathcal{S}$  é dado por

$$\nabla \mathcal{S}_j(\mathbf{P}) = - \sum_{i=1}^m r_i(\mathbf{P}) \frac{\partial \mathcal{A}(\mathbf{P})}{\partial p_j}; j = 1, \dots, n. \quad (3.26)$$

Seja

$$\mathcal{A}_i(\mathbf{P}) = y_i, \quad (3.27)$$

define-se a **matriz jacobiana**,  $\mathbf{J}$ , como aquela cujos elementos são tais que

$$J_{ij} = \frac{\partial \mathcal{A}_i(\mathbf{P})}{\partial p_j}, \quad (3.28)$$

com  $i = 1, \dots, m$  e  $j = 1, \dots, n$ , ou seja,

$$\mathbf{J} = \begin{bmatrix} \frac{\partial \mathcal{A}_1(\mathbf{P})}{\partial p_1} & \frac{\partial \mathcal{A}_1(\mathbf{P})}{\partial p_2} & \dots & \frac{\partial \mathcal{A}_1(\mathbf{P})}{\partial p_n} \\ \frac{\partial \mathcal{A}_2(\mathbf{P})}{\partial p_1} & \frac{\partial \mathcal{A}_2(\mathbf{P})}{\partial p_2} & \dots & \frac{\partial \mathcal{A}_2(\mathbf{P})}{\partial p_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial \mathcal{A}_m(\mathbf{P})}{\partial p_1} & \frac{\partial \mathcal{A}_m(\mathbf{P})}{\partial p_2} & \dots & \frac{\partial \mathcal{A}_m(\mathbf{P})}{\partial p_n} \end{bmatrix}. \quad (3.29)$$

Então é possível escrever o gradiente da função objetivo na forma vetorial como

$$\nabla \mathcal{S}(\mathbf{P}) = -\mathbf{J}^T \mathbf{R}(\mathbf{P}). \quad (3.30)$$

Além disso, considere que

$$\nabla^2 \mathcal{S}(\mathbf{P}) = \frac{\partial \nabla \mathcal{S}(\mathbf{P})}{\partial \mathbf{P}} \quad (3.31)$$

representa a **matriz Hessiana**,  $\mathbf{H}$ , de  $\mathcal{S}(\mathbf{P})$ , cujos elementos são dados por

$$H_{i,j} = \frac{\partial^2 \mathcal{S}(\mathbf{P})}{\partial p_i \partial p_j}.$$

As definições acima apresentadas serão utilizadas no desenvolvimento dos métodos de Newton nas próximas subseções.

### 3.2.2.1 Método de Newton Clássico

Considere que a função objetivo  $\mathcal{S}(\mathbf{P})$  seja aproximada, a cada iteração, por uma série de Taylor truncada nos termos de segunda ordem em torno de  $\mathbf{P}_k$ . Tem-se

$$\mathcal{S}(\mathbf{P}_{k+1}) \approx \mathcal{S}(\mathbf{P}_k) + \nabla \mathcal{S}(\mathbf{P}_k)(\mathbf{P}_{k+1} - \mathbf{P}_k) + \frac{1}{2}(\mathbf{P}_{k+1} - \mathbf{P}_k)^T \nabla^2 \mathcal{S}(\mathbf{P}_k)(\mathbf{P}_{k+1} - \mathbf{P}_k) \quad (3.32)$$

e derivando essa expressão em relação a  $\mathbf{P}_{k+1}$ , além de presumir que  $\mathcal{S}(\mathbf{P}_{k+1})$  é o mínimo, tem-se

$$0 = 0 + \nabla \mathcal{S}(\mathbf{P}_k) + \nabla^2 \mathcal{S}(\mathbf{P}_k)(\mathbf{P}_{k+1} - \mathbf{P}_k). \quad (3.33)$$

Manipulando a expressão acima, tem-se que

$$\nabla^2 \mathcal{S}(\mathbf{P}_k) \mathbf{P}_{k+1} = \nabla^2 \mathcal{S}(\mathbf{P}_k) \mathbf{P}_k - \nabla \mathcal{S}(\mathbf{P}_k). \quad (3.34)$$

Substituindo-se, finalmente, as definições vetoriais do gradiente de  $\mathcal{S}(\mathbf{P})$  e da Hessiana, tem-se que o passo ótimo para  $\mathbf{P}$  é aquele dado pela solução do sistema linear

$$\mathbf{H}_k \mathbf{P}_{k+1} = \mathbf{H}_k \mathbf{P}_k - \mathbf{J}_k^T \mathbf{R}(\mathbf{P}_k) \quad (3.35)$$

e, portanto, o passo  $\Delta \mathbf{P}$ , na  $k$ -ésima iteração, passa a ser dado por

$$\Delta \mathbf{P}_k = (\mathbf{H}_k)^{-1} \mathbf{J}_k^T \mathbf{R}(\mathbf{P}_k), \quad (3.36)$$

sempre que  $\mathbf{H}_k$  for inversível.

O Método de Newton, embora bem fundamentado e, geralmente, eficiente quando aplicado a uma condição inicial de qualidade, traz consigo dificuldades computacionais. A mais imediata dessas dificuldades reside na necessidade de, além de computar a Jacobiana, também a Hessiana associadas ao problema direto  $\mathcal{A}(\mathbf{P})$ , bem como sua inversa, a cada passo, o que acaba tornando o método computacionalmente oneroso. Na sequência, serão apresentadas sugestões feitas na literatura

que amenizam o custo computacional do método de Newton por evitar computação direta da Hessiana, dando preferência a aproximações da mesma. A estas dá-se o nome de Métodos Quasi-Newton.

Outro problema enfrentado pelos métodos do tipo Newton se refere à escolha do vetor de chute inicial  $\mathbf{P}_0$ . Numa grande gama de aplicações, uma escolha inapropriada do vetor inicial pode conduzir a soluções que não se referem àquelas de “norma mínima”. Ou seja, o método eventualmente convergirá para mínimos locais e não para o mínimo global do funcional de resíduos quadráticos.

### 3.2.2.2 Método de Gauss-Newton

O método de Gauss-Newton se trata de uma adaptação do método de Newton, que considera uma aproximação de primeira ordem para a computação da Hessiana.

Retornando à definição da Hessiana, tem-se que cada um de seus elementos é dado por

$$H_{i,j} = \frac{\partial^2[\mathcal{S}(\mathbf{P})]}{\partial p_i \partial p_j} = \frac{\partial[\nabla \mathcal{S}(\mathbf{P})]}{\partial p_j} = \sum_{j=1}^n \left( \frac{\partial[r_i(\mathbf{P})]}{\partial p_i} \frac{\partial[r_j(\mathbf{P})]}{\partial p_j} + r_j(\mathbf{P}) \frac{\partial^2[r_j(\mathbf{P})]}{\partial p_i \partial p_j} \right). \quad (3.37)$$

Levando-se em consideração, ainda, que

$$\frac{\partial r_i}{\partial p_j} = \frac{\partial[y_i - \mathcal{A}(\mathbf{P})]}{\partial p_j} = -\frac{\partial \mathcal{A}(\mathbf{P})}{\partial p_j}, \quad (3.38)$$

é possível reescrever a Hessiana, na forma matricial, como

$$\mathbf{H} = \mathbf{J}^T \mathbf{J} + \sum_{j=1}^n \left[ r_j(\mathbf{P}) \frac{\partial^2[r_j(\mathbf{P})]}{\partial p_i \partial p_j} \right]. \quad (3.39)$$



Desconsiderando-se, finalmente, os termos de segunda ordem, tem-se uma aproximação da Hessiana dada por

$$\mathbf{H} \approx \mathbf{J}^T \mathbf{J}, \quad (3.40)$$

usualmente chamada de **matriz de informações de Fischer** [2].

O método de Gauss-Newton se baseia na substituição, em (3.35), da Hessiana por essa aproximação [4], de modo a obter o sistema

$$\mathbf{J}_k^T \mathbf{J}_k \Delta \mathbf{P}_k = -\mathbf{J}_k^T \mathbf{R}(\mathbf{P}_k). \quad (3.41)$$

no qual a matriz de informações de Fischer é simétrica e positiva semidefinida [4]. Desse modo, o processo de atualização de  $\mathbf{P}$  passa a ser dado por

$$\mathbf{P}_{k+1} = \mathbf{P}_k - [\mathbf{J}_k^T \mathbf{J}_k]^{-1} \mathbf{J}_k^T \mathbf{R}(\mathbf{P}_k). \quad (3.42)$$

Embora a convergência de Gauss-Newton exija um número, em geral, maior de iterações do que o método original, ela costuma ser mais rápida, já que o custo de cada iteração é minimizado pelo fato de não mais ser necessário computar a Hessiana. Os problemas de convergência do método anterior são, contudo, replicados aqui. O fato de a matriz de Fischer ser positiva semidefinida, contudo, impede que a atualização se dê em sentido "oposto" ao do mínimo, resultando em mais estabilidade na convergência, dada uma estimativa inicial de qualidade.

### 3.2.2.3 Método de Levenberg-Marquart

Visando permitir a convergência do método de Gauss-Newton para uma gama mais ampla de estimativas iniciais, Levenberg propôs a introdução de um fator de amortecimento, de modo semelhante ao método do Gradiente, na diagonal

da matriz de informações de Fischer, em (3.41). Na prática, isso garante que o gradiente tenha maior peso na atualização de  $\mathbf{P}_k$ . Nessa perspectiva, o passo  $\Delta\mathbf{P}_k$  passa, então, a ser dado como solução do sistema [47]

$$[\mathbf{J}_k^T \mathbf{J}_k + \lambda_k \mathbf{I}] \Delta\mathbf{P}_k = -\mathbf{J}_k^T \mathbf{R}(\mathbf{P}_k) \quad (3.43)$$

É fácil de perceber que, quando esse fator de amortecimento é muito pequeno, a atualização obtida é muito próxima da que se obtém por Gauss-Newton, enquanto que, quando é muito grande, tem-se

$$\mathbf{J}_k^T \mathbf{J}_k + \lambda_k \mathbf{I} \approx \lambda_k \mathbf{I}, \quad (3.44)$$

e o passo de atualização é dada por

$$\Delta\mathbf{P}_k \approx -\frac{1}{\lambda_k} \mathbf{J}_k^T \mathbf{R}(\mathbf{P}_k) = \frac{1}{\lambda_k} \nabla \mathcal{S}(\mathbf{P}_k), \quad (3.45)$$

uma aproximação do tipo Gradiente.

O maior desafio associado ao método de Levenberg envolve a determinação de um valor ótimo para  $\lambda$  em cada iteração. Na literatura, é possível encontrar uma série de sugestões a respeito de como essa atualização pode ser controlada. Uma delas é controlar esse valor de  $\lambda_k$  de acordo com o comportamento da última iteração. Nesse caso, atribui-se, inicialmente, um valor relativamente grande a  $\lambda_{inicial}$ , de modo que a diagonal tenha um peso maior na primeira iteração e se obtenha um passo do tipo Gradiente, cuja convergência tende a ser estável para uma gama maior de valores inicializadores. Ao longo das iterações, quando um passo se mostrar eficiente, indicando que o método está *funcionando bem*, ou seja, que o módulo da função objetivo está reduzindo, procede-se com uma atualização que reduza também o valor de  $\lambda$ . De modo análogo, quando o método estiver *funcionando mal*, ou seja, quando o módulo da função objetivo aumentar, descarta-se a atualização do *shift vector* e se procede com a ampliação do fator  $\lambda$ .

Desse modo, quando o método estiver funcionando bem, tenta-se impor uma atualização mais próxima da de Gauss-Newton, cuja convergência é rápida e estável próximo do mínimo e, quando estiver funcionando mal, força-se uma atualização do tipo Gradiente, que é lenta, mas mais estável. Com isso, espera-se que o método conduza  $\mathbf{P}_k$  para a vizinhança de  $\mathbf{P}$ , dentro do raio de convergência dos métodos de Newton e que, conseqüentemente,  $\lambda$  seja reduzido de tal modo que seja quase nulo quando houver a convergência [26, 33].

Uma desvantagem do algoritmo proposto por Levenberg é que, nos casos em que  $\lambda$  é muito grande, como na primeira iteração, a matriz de informações de Fischer, calculada, não apresenta relevância alguma para a solução do sistema, e pode-se acabar direcionando a próxima estimativa a alguma área de baixa convergência. Uma solução para o problema foi introduzida por Marquardt, que sugeriu a substituição da matriz identidade, no passo de atualização, pela matriz com os elementos da diagonal da matriz de informações de Fischer. Assim, o passo  $\Delta\mathbf{P}_k$  fica dado pela solução do seguinte sistema de equações [47],

$$[\mathbf{J}_k^T \mathbf{J}_k + \lambda_k \mathbf{D}_k] \Delta\mathbf{P}_k = -\mathbf{J}_k^T \mathbf{R}(\mathbf{P}_k) \quad (3.46)$$

onde

$$\mathbf{D}_k = \text{diag}(\mathbf{J}_k^T \mathbf{J}_k), \quad (3.47)$$

de modo que cada iteração produz uma nova aproximação na forma,

$$\mathbf{P}_{k+1} = \mathbf{P}_k + [\mathbf{J}_k^T \mathbf{J}_k + \lambda_k \mathbf{D}_k]^{-1} \mathbf{J}_k^T \mathbf{R}(\mathbf{P}_k). \quad (3.48)$$

O método descrito acima é conhecido como método de Levenberg-Marquardt e é um dos mais importantes na solução de problemas de otimização, já que, em comparação com o de Newton, aumenta a gama de chutes iniciais que levam à convergência, dispensando a necessidade de calcular a Hessiana.

### 3.2.3 Métodos *Trust-Region*

A classe de métodos do tipo *Trust-Region* é relativamente nova e se baseia em limitar a uma região de confiança, as atualizações de passo dadas pelo vetor de incrementos  $\Delta \mathbf{P}$ . Essa ideia é razoável, considerando que os modelos não-lineares, quando aproximados linear ou quadraticamente, como nos métodos anteriores, só são capazes de coincidir com o modelo real localmente [60].

Supondo que uma determinada iteração  $k$  gere a estimativa  $\mathbf{P}_k$  para o vetor de parâmetros que minimiza a função objetivo, a região de confiança (ou *trust-region*) é definida, nesse passo, como um círculo em torno da estimativa atual de raio  $l_k$ .

A ideia é iniciar com um raio pequeno e, quando a próxima aproximação é calculada, verifica-se, então, se há indicativos de que essa região do modelo aproximado representa bem o modelo real, caso no qual, a região de confiança pode ser ampliada, aumentando-se o raio  $l_k$ . No caso, contudo, em que uma nova iteração indique que a região não mais representa de forma adequada o modelo real, o raio  $l_k$  é diminuído [60].

De modo geral, define-se então um método do tipo *trust-region* para mínimos quadrados como aquele que visa encontrar

$$\mathbf{P} = \arg \min_{\mathbf{P}} [\mathcal{S}(\mathbf{P})]$$

a partir de um processo iterativo

$$\mathbf{P}_{k+1} = \mathbf{P}_k + \Delta \mathbf{P}_k \tag{3.49}$$

e

$$\|\Delta \mathbf{P}_k\|_2 \leq l_k. \tag{3.50}$$

Há, na literatura, uma série de sugestões a respeito de como a região de confiança pode ser testada e atualizada a cada passo, a fim de verificar se a nova iteração deve ser aceita ou rejeitada. Uma discussão mais aprofundada sobre o tema pode ser encontrada no artigo de Ya-xiang Yuan [60] e foge do escopo desse trabalho.

Por conta de a região de confiança ser limitada, esse tipo de método se comporta bem mesmo quando a função objetivo considerada não é convexa ou quando aplicados a problemas altamente mal-condicionados, conservando fortes propriedades de convergência [60]. Neste trabalho, será utilizada a implementação do *Matlab* de um método do tipo *trust-region*, chamado pelo comando LSQNONLIN.

### 3.3 Análise estocástica do método para o caso linear

Nesta seção, procede-se com a introdução de conceitos fundamentais em probabilidade e estatística, posteriormente utilizados para a execução de análise mais aprofundada da qualidade esperada para a solução aproximada por mínimos quadrados de problemas inversos discretos lineares.

Define-se **erro** como toda discrepância entre os valores reais e observados para uma determinada grandeza. Em geral, não é possível conhecer o **erro real** nos dados, mas sim um **erro aparente**, que é a diferença entre os valores observados e o valor mais provável. Em relação à sua natureza, pode-se classificar os erros dentro de três categorias:

1. **Grosseiro:** em geral de ordem humana, como, por exemplo, o erro na leitura de uma medida, ou então por defeitos advindos do equipamento de medição. Esse tipo de erro, comumente chamado de **outlier**, pode ser detectado e corrigido antes mesmo da aplicação do método dos mínimos quadrados.

2. **Sistemático:** proveniente de uma causa permanente, conhecida ou não. Ocorre sempre no mesmo sentido, ou seja, é polarizado, podendo ter causa física. Ocorrem, geralmente, devido à falta de calibração do equipamento de medição ou acabam sendo introduzidos pelo observador. Esse tipo de erro costuma se acumular, **polarizando**, ou seja, desviando em um determinado sentido, os parâmetros aproximados por mínimos quadrados.
3. **Estocástico:** erros de ordem aleatória. São, em geral, pequenos e inevitáveis e causam discrepâncias que, a princípio não contêm conformidade matemática. Sua influência sobre as observações é aleatória e, como tendem a se apresentar em todas as direções, acabam por se anular quando uma grande quantidade de dados está disponível.

Antes mesmo de aplicar o método dos mínimos quadrados, faz-se essencial proceder com uma análise inicial das observações, de modo a eliminar os erros grosseiros e corrigir os erros sistemáticos.

Considere-se  $\mathcal{X}$  uma variável aleatória contínua, considerando  $f(x)$  a função de densidade de probabilidade, define-se **valor esperado** de  $\mathcal{X}$  como

$$E[\mathcal{X}] = \int_{-\infty}^{\infty} xf(x)dx. \quad (3.51)$$

Considerando-se ainda um vetor  $T$  de dimensão  $m$ , onde cada uma de suas componentes sejam variáveis aleatórias,  $\mathcal{T}_i$ , tem-se

$$E[T] = \left[ E[\mathcal{T}_1] \quad E[\mathcal{T}_2] \quad E[\mathcal{T}_3] \quad \cdots \quad E[\mathcal{T}_m] \right]^T. \quad (3.52)$$

A precisão ou acurácia aparente entre os dados aponta para o grau de refinamento da medição, ou seja, o quanto os valores medidos estão próximos uns dos outros. Esta é, em geral, representada pela **variância**, que, para uma variável

aleatória  $\mathcal{X}$  é definida como

$$\sigma_0^2(\mathcal{X}) = \text{var}(\mathcal{X}) = E[(\mathcal{X} - E[\mathcal{X}])^2]. \quad (3.53)$$

Pode-se, ainda, definir a **covariância** entre duas variáveis aleatórias  $\mathcal{X}$  e  $\mathcal{Z}$ . Essa relação indica a tendência e força da relação linear entre essas duas variáveis e é dada por

$$\sigma_0(\mathcal{X}, \mathcal{Z}) = \text{cov}(\mathcal{X}, \mathcal{Z}) = E[(\mathcal{X} - E[\mathcal{X}])(\mathcal{Z} - E[\mathcal{Z}])]. \quad (3.54)$$

Quando apenas um conjunto de  $n$  observações de uma variável aleatória  $\mathcal{X}$  estiver disponível, contudo, não é possível encontrar o valor esperado exato. Nesse caso, quando  $n$  for suficientemente grande, considera-se que a **média** das observações representa uma boa **aproximação do valor esperado**. Define-se média das  $n$  observações de  $\mathcal{X}$  na forma  $\{x_i\}_{i=1}^n$  como

$$\bar{\mathcal{X}} = \frac{1}{n} \sum_{i=1}^n x_i. \quad (3.55)$$

Na mesma situação, define-se, a **aproximação da variância** de determinada variável aleatória como

$$\sigma^2(\mathcal{X}) = V(\mathcal{X}) = \frac{1}{n-g} \sum_{i=1}^n (x_i - \bar{\mathcal{X}})^2, \quad (3.56)$$

onde  $g$  representa o número de **graus de liberdade** na aproximação. De modo análogo, a **aproximação da covariância** entre duas variáveis aleatórias é dada por

$$\sigma(\mathcal{X}, \mathcal{Z}) = \frac{1}{n-g} \sum_{i=1}^n (x_i - \bar{\mathcal{X}})(z_i - \bar{\mathcal{Z}}). \quad (3.57)$$

### 3.3.1 Qualidade da solução aproximada por mínimos quadrados

Suponha-se que as observações em  $\mathbf{Y}$  possam ser descritas por

$$\mathbf{Y} = \mathcal{A}(\mathbf{P}) + \mathbf{W}. \quad (3.58)$$

onde  $\mathcal{A}(\mathbf{P})$  é determinística e  $\mathbf{W}$  representa um vetor de perturbações de variáveis aleatórias com média zero e variância individual  $\sigma_0^2$ , representado por

$$\mathbf{W} = \begin{bmatrix} \mathcal{W}_1 & \mathcal{W}_2 & \mathcal{W}_3 & \cdots & \mathcal{W}_m \end{bmatrix}^T. \quad (3.59)$$

Pelas características de  $\mathbf{W}$ , sabe-se que

$$E[\mathbf{W}] = 0 \quad (3.60)$$

e, portanto,

$$\text{cov}(\mathbf{W}, \mathbf{W}) = E[(\mathbf{W} - E(\mathbf{W}))(\mathbf{W} - E(\mathbf{W}))^T] = E[\mathbf{W}\mathbf{W}^T] = \mathbf{L}. \quad (3.61)$$

Se o problema em questão for linear em relação aos parâmetros, sabe-se ainda que sua solução aproximada por mínimos quadrados é aquela que atende às equações normais, ou seja, com  $\mathbf{P}$  tal que

$$\hat{\mathbf{P}} = [\mathbf{A}^T \mathbf{A}]^{-1} \mathbf{A}^T \mathbf{Y} \quad (3.62)$$

conforme anteriormente demonstrado. Considerando-se  $\mathbf{P}_0$  a solução exata do problema inverso, de modo que

$$\mathbf{Y} = \mathbf{A}\mathbf{P}_0 + \mathbf{W} \quad (3.63)$$

e substituindo essa formulação na equação(3.62), tem-se

$$\begin{aligned} \hat{\mathbf{P}} &= [\mathbf{A}^T \mathbf{A}]^{-1} \mathbf{A}^T (\mathbf{A}\mathbf{P}_0 + \mathbf{W}) = [\mathbf{A}^T \mathbf{A}]^{-1} [\mathbf{A}^T \mathbf{A}] \mathbf{P}_0 + [\mathbf{A}^T \mathbf{A}]^{-1} \mathbf{A}^T \mathbf{W} = \\ &= \mathbf{P}_0 + [\mathbf{A}^T \mathbf{A}]^{-1} \mathbf{A}^T \mathbf{W} \end{aligned} \quad (3.64)$$

e, por (3.60), segue que

$$E[\hat{\mathbf{P}}] = E[\mathbf{P}_0 + [\mathbf{A}^T \mathbf{A}]^{-1} \mathbf{A}^T \mathbf{W}] = E[\mathbf{P}_0] + [\mathbf{A}^T \mathbf{A}]^{-1} \mathbf{A}^T E[\mathbf{W}] = E[\mathbf{P}_0], \quad (3.65)$$

ou seja, que

$$E[\hat{\mathbf{P}}] = \mathbf{P}_0, \quad (3.66)$$



donde pode-se concluir que  $\hat{\mathbf{P}}$  é não polarizada no caso em que uma descrição verdadeira do sistema esteja disponível.

Considerando-se então as equações (3.58), (3.65) e (3.62), é possível concluir que

$$\text{cov}(\mathbf{P}, \mathbf{P}) = E[(\mathbf{P} - E[\mathbf{P}])(\mathbf{P} - E[\mathbf{P}])^T] = E[(\mathbf{P} - \mathbf{P}_0)(\mathbf{P} - \mathbf{P}_0)^T] \quad (3.67)$$

e, por (3.64), tem-se que

$$\hat{\mathbf{P}} - \mathbf{P}_0 = [\mathbf{A}^T \mathbf{A}]^{-1} \mathbf{A}^T \mathbf{W}, \quad (3.68)$$

de modo que, levando-se em consideração (3.61), conclui-se que

$$\text{cov}(\hat{\mathbf{P}}, \hat{\mathbf{P}}) = E(([\mathbf{A}^T \mathbf{A}]^{-1} \mathbf{A}^T \mathbf{W})(\mathbf{W}^T \mathbf{A} [\mathbf{A}^T \mathbf{A}]^{-1})) = ([\mathbf{A}^T \mathbf{A}]^{-1} \mathbf{A}^T) \mathbf{L} (\mathbf{A} [\mathbf{A}^T \mathbf{A}]^{-1}) \quad (3.69)$$

que independe da forma de  $\mathcal{A}(\mathbf{P})$ , desde que  $[\mathbf{A}^T \mathbf{A}]^{-1}$  exista.

Caso a variância do ruído seja igual para todas as amostras, e dada por  $\sigma_0^2$ , e estas sejam independentes entre si tem-se, ainda, que  $\mathbf{L} = \sigma_0^2 \mathbf{I}$  [31], e, nesse caso, a matriz de covariância das soluções aproximadas é dada por

$$\mathbf{C} = \sigma_0^2 [\mathbf{A}^T \mathbf{A}]^{-1}. \quad (3.70)$$

Em relação à sua distribuição, as estimativas  $\hat{\mathbf{P}}$  e  $\mathbf{P}_0$  são variáveis aleatórias, visto que foram construídas a partir de variáveis aleatórias  $\mathbf{Y}$ . Assumindo que o vetor  $\mathbf{W}$  das perturbações tenha distribuição gaussiana, pode-se concluir que  $\mathbf{Y}$  também tenha distribuição gaussiana, com média  $\mathcal{A}(\mathbf{P})$  e variância  $\mathbf{L}$ , ou seja

$$\mathbf{Y} \sim \mathcal{N}(\mathcal{A}(\mathbf{P}), \mathbf{L}). \quad (3.71)$$

Como  $\mathbf{P}$  é dada por uma combinação linear de  $\mathbf{Y}$ , será, também, gaussiana

$$\hat{\mathbf{P}} \sim \mathcal{N}(\mathbf{P}_0, \mathbf{C}), \quad (3.72)$$

onde  $\mathbf{C}$  é a matriz de covariância dos dados.

Mesmo no caso em que as observações não estejam distribuídas normalmente, “é comum que a distribuição de  $\mathbf{P}$  se aproxime de uma distribuição normal conforme  $n \rightarrow \infty$ , o que segue da aplicação do teorema central do limite para a soma das variáveis aleatórias” que constituem as estimativas [31]. É possível verificar na literatura [31], também, que a variância real,  $\sigma_0^2$ , é uma variável aleatória e segue uma distribuição qui-quadrada com  $n$  graus de liberdade, ou seja

$$\frac{n}{\sigma_0^2} V(\mathbf{P}_0) \sim \chi^2(n). \quad (3.73)$$

### 3.3.2 Elipses de erro para as soluções aproximadas

Considerando-se, novamente, o caso linear dado por

$$\mathcal{A}(\mathbf{P}_0) = \mathbf{A}\mathbf{P}_0, \quad (3.74)$$

onde o parâmetro  $\mathbf{P}_0$  correto existe, os resultados anteriores permitem inferir que a diferença entre  $\mathbf{P}$  e  $\mathbf{P}_0$  está distribuída normalmente, de modo que

$$\hat{\mathbf{P}} - \mathbf{P}_0 \sim \mathcal{N}(0, \mathbf{C}). \quad (3.75)$$

Como  $\mathbf{C}$  é matriz de covariância da distribuição do vetor  $\mathbf{P}$ , pode-se obter informações úteis a respeito da covariância e correlação entre as diferentes componentes de  $\mathbf{P}$ . A partir da equação (3.75), é possível determinar que [31]:

$$(\hat{\mathbf{P}} - \mathbf{P}_0)^T \mathbf{C}^{-1} (\hat{\mathbf{P}} - \mathbf{P}_0) \sim \chi^2(d) \quad (3.76)$$

e, por aplicação direta da definição da distribuição qui-quadrada, segue que a probabilidade de

$$(\hat{\mathbf{P}} - \mathbf{P}_0)^T \mathbf{C}^{-1} (\hat{\mathbf{P}} - \mathbf{P}_0) \geq \eta \quad (3.77)$$

é aquela de nível  $\eta$  em uma distribuição qui-quadrada com  $d$  graus de liberdade. Essa expressão nos permite determinar elipsóides em  $\mathbb{R}^d$ , cuja forma é determinada por  $\mathbf{C}$ .

Quando  $\sigma_0^2$  for conhecida, verificou-se que é possível determinar essa matriz pela relação

$$\mathbf{C} = \sigma_0^2 [\mathbf{A}^T \mathbf{A}]^{-1}. \quad (3.78)$$

Nos casos em que  $\sigma_0^2$  for desconhecida, é importante que a variância seja estimada a partir dos dados.

Neste capítulo, o critério dos mínimos quadrados foi apresentado como uma forma eficiente para a determinação de uma solução aproximada para problemas inversos. Demonstrou-se que, ao menos no caso linear, a solução aproximada  $\hat{\mathbf{P}}$  tende a representar uma boa estimativa do conjunto de parâmetros  $\mathbf{P}_0$  quando apenas ruídos gaussianos de média zero estiverem presentes nos dados experimentais. Além disso, ferramentas para a análise da qualidade das estimativas foram apresentadas e serão posteriormente utilizadas para a verificação dos resultados numéricos deste trabalho.

## 4 BIORREATORES ANAERÓBICOS

Um **biorreator**, ou reator biológico, trata-se de um equipamento em que se dá o crescimento de biomassa de um ou mais microorganismos, resultando na criação de algum produto. O biorreator contém um inóculo, ou seja, um ambiente com alta concentração de microorganismos e substrato, que é um meio fermentativo com os ingredientes necessários para o crescimento desses microorganismos.

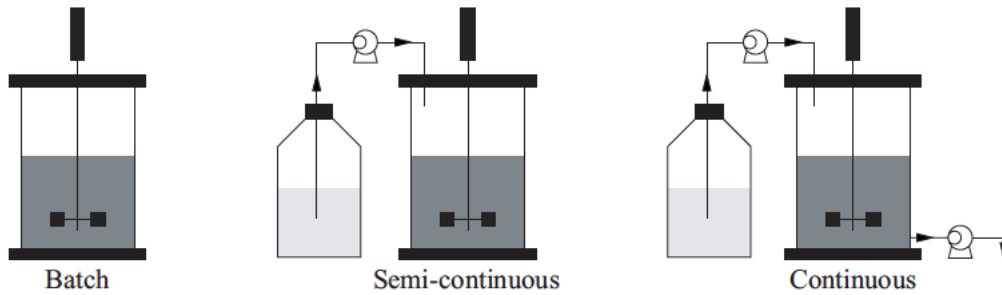
Entende-se como **biomassa** "qualquer matéria orgânica que possa ser transformada em energia mecânica, térmica ou elétrica" [45]. Esta é, em geral, classificada de acordo com sua origem, podendo ser florestal, agrícola ou de rejeitos urbanos e industriais. Há diversas técnicas utilizadas para transformar matéria-prima em energético, sendo que cada uma dá origem a determinado derivado. A **fermentação microbiana**, no interior do biorreator, é um processo no qual uma população de micro-organismos, como bactérias, são cultivadas usando elementos nutritivos específicos em condições favoráveis - como temperatura, pH, agitação, aeração, entre outros. Esquemáticamente corresponde "à transformação de substâncias (geralmente substratos carbonáceos) em produtos, resultantes da atividade metabólica das células" [20]. Os principais componentes desse tipo de reação são os substratos, "necessários para o crescimento de micro-organismos, ou mesmo precursores de um composto a ser produzido", a biomassa microbiana e os produtos finais [20].

Em relação ao tipo de digestão bacteriana, os biorreatores podem ser classificados em **aeróbicos** ou **anaeróbicos**, ou seja, com ou sem a presença de ar [12]. A digestão anaeróbica, contudo, apresenta uma série de vantagens em relação à aeróbica. Além de possuir "alta capacidade de degradação de substratos concentrados e complexos (resíduos de plantas, dejetos animais, resíduos da indústria de alimentos, entre outros), produz poucos resíduos, requer pouca energia e, em

alguns casos, ainda é possível recuperar energia usando a combustão de metano” [9]. Nesse sentido, se justificam os estudos feitos na área nos últimos anos, no sentido de desenvolver e analisar modelos para o processo de digestão anaeróbica considerando a concepção, operação e otimização de biorreatores anaeróbicos [11].

Já em relação ao modo de operação, os biorreatores são em geral caracterizados pelo método de alimentação de substratos. Nesse sentido, é possível caracterizá-los dentro de três categorias: biorreatores descontínuos (ou batelada), semi-contínuos ou contínuos.

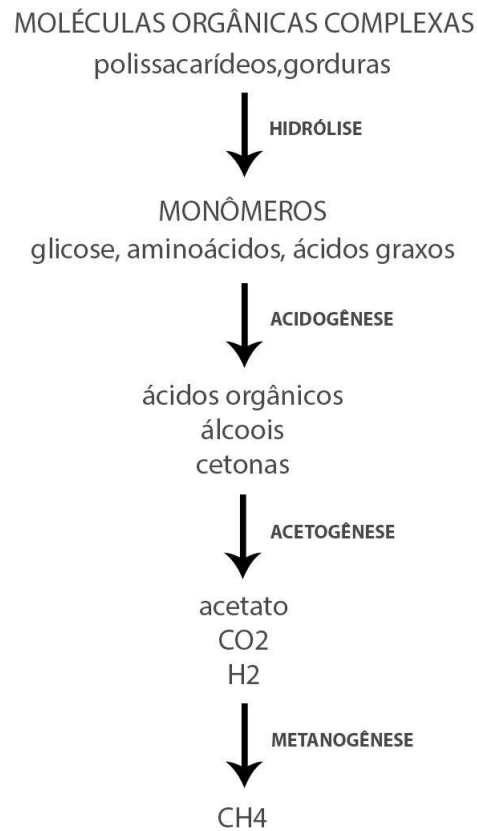
No caso dos  **biorreatores batelada**, todos os elementos nutritivos necessários para o crescimento biológico são introduzidos no início da reação, não havendo alimentação de influentes ou remoção de resíduos após o início do processo, de modo que a reação ocorre com volume constante. Este tipo de biorreator é de fácil implementação e há a garantia da pureza das culturas inseridas, já que há pouco risco de contaminação. Isso acarreta, contudo, na dificuldade de operar o meio fermentador a fim de otimizar o uso dos micro-organismos, além de reduzir a duração do processo, já que a introdução de uma grande quantidade de substrato inicial, em geral, acaba por inibir o crescimento dos micro-organismos que o consomem [20]. Já no caso dos  **biorreatores semi-contínuos**, é possível alimentar o sistema, após o início da reação, com os elementos nutritivos necessários no meio fermentativo. Com isso, é possível eliminar os problemas de inibição presentes nos modelos batelada, podendo otimizar as taxas de crescimento biológico. Os  **biorreatores contínuos**, por sua vez, são caracterizados por um circuito aberto de alimentação, com volume constante de reação, onde a taxa de saída do meio é igual à de entrada de nutrientes. Isso permite produções significativas mesmo em reatores de dimensões pequenas e, por isso mesmo, são largamente utilizados. É possível encontrar, no livro de Dochain [20], uma ilustração que evidencia a diferença entre os três tipos de biorreatores, que é reproduzida na Figura 4.1.



Fonte: Dochain [20] (2008)

Figura 4.1: Diferentes tipos de biorreatores, de acordo com seu modo de operação

Por fim, em relação ao processo interno do reator, após a introdução de moléculas orgânicas complexas, como polissacarídeos e gorduras, há quatro estágios de decomposição que conduzem aos produtos finais: hidrólise, acidogênese, acetogênese e metanogênese. Após a alimentação, procede-se com a adição de água. Passa então a ocorrer um processo de hidrólise, no qual ocorre uma primeira quebra de ligação química das moléculas. A molécula de água é, também, quebrada em íons de hidrogênio e hidroxila, que se ligam às moléculas resultantes da quebra das moléculas poliméricas. Esse processo se dá devido à alteração do pH do meio. Após essa primeira reação, estarão presentes no interior do biorreator monômeros, como a glicose, aminoácidos e ácidos graxos. Nesse momento, há a ação das bactérias acidogênicas, que passam a decompor esses materiais em ácidos orgânicos, álcoois e cetonas, em um processo que recebe o nome de acidogênese. Em paralelo, as bactérias acetogênicas decompõem essas matérias em acetato, dióxido de carbono ( $CO_2$ ) e gás hidrogênio ( $H_2$ ). Após a acetogênese, as bactérias metanogênicas decompõem os acetatos, produzindo gás metano ( $CH_4$ ). Todo o processo pode ser visualizado, de forma resumida em diagrama encontrado em material de Pedroso [46], reproduzido na Figura 4.2.



Fonte: Pedroso [46] (2012)

Figura 4.2: Diagrama do processo de decomposição de matéria orgânica em um biorreator descontínuo

Ao final do processo, a matéria originalmente introduzida é, então, decomposta em produtos como gás metano, dióxido de carbono, hidrogênio, nitrogênio, oxigênio e outros. Um quadro do percentual estimado de saída de cada um destes produtos [46] pode ser visualizado na Tabela 4.1.

Tabela 4.1: Distribuição estimada dos produtos finais em um biorreator descontínuo

<b>Produto</b>	<b>Porcentagem</b>
Metano	50 a 75
Dióxido de Carbono	25 a 40
Hidrogênio	1 a 3
Nitrogênio	0,5 a 2,5
Oxigênio	0,1 a 1,0
Ácido Sulfídrico	0,1 a 0,5
Amônia	0,1 a 0,5
Água	variável

O objetivo deste trabalho é avaliar a qualidade das estimativas para os estágios de decomposição em um reator, sendo conhecida apenas da saída de gás metano do mesmo. Para tal, se faz necessário, contudo, definir primeiro qual será o modelo matemático que se considera suficientemente preciso para explicar os processos internos do mesmo, ou seja, o problema direto.

#### 4.1 Descrição do modelo matemático considerado

É possível encontrar na literatura modelos capazes de descrever o funcionamento de um biorreator anaeróbico em diferentes graus de complexidade. Um modelo simples, por exemplo, é o proposto por ANDREWS (1968), que descreve os processos internos do reator com apenas uma população de bactérias. Um modelo mais complexo, capaz de representar tanto as reações bioquímicas quanto as reações físico-químicas envolvidas na digestão anaeróbica foi proposto pelo grupo IWA Anaerobic Digestion Modelling Task Group e é chamado de ADM1 [12]. Esse modelo



é composto por 32 variáveis de estado e, embora teoricamente represente o sistema de forma mais acurada, a identificação de um grande número de parâmetros com desejada precisão se torna uma tarefa difícil [12], além de computacionalmente onerosa [11]. Considerando-se que, para a otimização e, sobretudo, controle desejados do biorreator, além de não ser possível ou necessário conhecer todos esses estados, o tempo necessário para a identificação passa a desempenhar papel importante. Nesse sentido, vê-se larga aceitação nos artigos mais recentes [1,22] de um modelo de média complexidade, com quatro a seis variáveis de estado, representando as reações de acidogênese, metanogênese e balanço de íons do biorreator, ou seja, desconsiderando o processo de acetogênese.

Esses modelos são capazes de descrever a dinâmica do processo com precisão suficiente [11] em menos tempo e com menor complexidade computacional. Neste trabalho, considera-se o modelo proposto por Antonelli [1], baseado em balanço de massa e composto por quatro estados e uma saída, descrito por

$$x_1'(t) = [v_1(S_1(t)) - \alpha D]x_1(t) \quad (4.1)$$

$$x_2'(t) = [v_2(S_2(t)) - \alpha D]x_2(t) \quad (4.2)$$

$$S_1'(t) = D[S_1^{in}(t) - S_1(t)] - k_1 v_1(S_1(t))x_1(t) \quad (4.3)$$

$$S_2'(t) = D[S_2^{in}(t) - S_2(t)] + k_2 v_1(S_1(t))x_1(t) - k_3 v_2(S_2(t))x_2(t), \quad (4.4)$$

onde  $x_1(t)$  representa a concentração de bactérias acidogênicas (mg/L),  $x_2(t)$  a concentração de bactérias metanogênicas (mg/L),  $S_1(t)$  é a concentração de demanda química de oxigênio, ou COD (mg/L),  $S_2(t)$  é a concentração de ácidos graxos voláteis (mmol/L),  $S_1^{in}(t)$  e  $S_2^{in}(t)$  representam, respectivamente, as concentrações de influentes, ou seja, de entrada de  $S_1(t)$  e  $S_2(t)$ ,  $D$ , por sua vez, é a taxa de diluição dos influentes ( $dia^{-1}$ ) enquanto que  $\alpha \in (0, 1]$  é um parâmetro de proporcionalidade determinado experimentalmente. Tem-se, ainda  $k_1$ , o coeficiente de degradação de COD (mgCOD/mg de  $x_1$ ),  $k_2$  o coeficiente de produção de ácidos graxos volá-

teis (mmolVFA/mg de  $x_2$ ) e  $k_3$  o coeficiente de consumo de ácidos graxos voláteis (mmolVFA/mg de  $x_2$ ).

Por fim,  $v_1(S_1)$  e  $v_2(S_2)$  representam as taxas de crescimento específico das duas bactérias, e são responsáveis pela não-linearidade do processo. Há mais de uma forma de representar esse crescimento específico, dentre as quais as mais utilizadas são a lei de Monod e a lei de Haldane [12]. Neste trabalho, considera-se que sejam expressas pela lei de Monod, de tal modo que

$$v_i(S_i(t)) = \mu_{mi} \frac{S_i(t)}{K_{S_i} + S_i(t)}, \quad i = 1, 2 \quad (4.5)$$

onde  $\mu_{m1}$  e  $\mu_{m2}$  ( $dia^{-1}$ ) representam, respectivamente, as taxas de crescimento máximas de biomassa acidogênica e metanogênica e  $K_{S_i}$  é o parâmetro de saturação associado à  $S_i(t)$  correspondente.

Esse modelo pode representar biorreatores contínuos, semi-contínuos ou batelada [22]. No caso de biorreatores do tipo batelada, que será alvo deste trabalho, visto que a alimentação se dá apenas no início do processo, é possível proceder com uma série de simplificações. Primeiro, considerando a ausência de influentes dos substratos, tem-se que, para todo  $t$ ,

$$S_1^{in}(t) = S_2^{in}(t) = 0.$$

Além disso, se não há influentes, tampouco é necessário considerar a diluição dos mesmos no meio, de tal modo que  $D = 0$ . Considerando essas simplificações, o modelo, no caso batelada, se reduz a

$$x_1'(t) = v_1(S_1(t))x_1(t) \quad (4.6)$$

$$x_2'(t) = v_2(S_2(t))x_2(t) \quad (4.7)$$

$$S_1'(t) = -k_1 v_1(S_1(t))x_1(t) \quad (4.8)$$

$$S_2'(t) = k_2 v_1(S_1(t))x_1(t) - k_3 v_2(S_2(t))x_2(t), \quad (4.9)$$

cujo tratamento é mais simples. A medida experimental disponível é a saída da taxa de fluxo de gás, onde

$$Q_{gas}(t) = q_C(t) + q_M(t) \quad (4.10)$$

dada pela soma da taxa de fluxo de dióxido de carbono  $q_C(t)$  e pela taxa de fluxo de metano  $q_M(t)$  produzidos pelo reator. Em particular, conhece-se a taxa de fluxo de gás metano que sai do sistema, expressa por

$$q_M(t) = k_6 v_2 (S_2(t)) x_2(t), \quad (4.11)$$

onde  $k_6$  representa a taxa de saída de gás.

## 4.2 Solução aproximada do problema direto pelo método de Runge-Kutta

Uma maneira de implementar a solução do problema direto, a partir do modelo definido anteriormente, é através dos métodos de Runge-Kutta. Estes podem ser vistos como um desenvolvimento do método de Euler [21] e se baseia na solução de um problema de valor inicial a partir da integração da equação diferencial envolvida e de um procedimento iterativo.

Considere um problema de valor inicial dado na forma

$$y' = f(u, y) \quad (4.12)$$

$$y(u_0) = y_0 \quad (4.13)$$

e considere que  $y(u)$  seja contínua e diferenciável no intervalo  $(a, b)$ , então (4.12) pode ser integrada no intervalo  $[u_i, u_{i+1}] \in (a, b)$  como

$$\int_{u_i}^{u_{i+1}} y' du = \int_{u_i}^{u_{i+1}} f(u, y) du \quad (4.14)$$

dando origem ao procedimento iterativo

$$y_{i+1} = y_i + Y_i, \quad (4.15)$$

no qual

$$Y_i = \int_{u_i}^{u_{i+1}} f(u, y) du. \quad (4.16)$$

O método de Euler se baseia na aproximação da integral dada por  $Y_i$  a partir da regra do retângulo, tomando um passo  $h$  e considerando

$$Y_i \approx hf(u_i, y_i), \quad (4.17)$$

onde  $u_i$  é o nodo do início do intervalo  $[u_i, u_{i+1}]$ .

Um modo de melhorar essa estimativa é usar, para a aproximação da integral em questão, é usar um polinômio interpolador de primeiro grau entre o nodo do início e fim do intervalo, obtendo-se

$$b_1 = f(u_i, y_i) \quad (4.18)$$

$$b_2 = f(u_i + h, y_i + b_1 h) \quad (4.19)$$

$$y_{i+1} = y_i + \frac{1}{2}(b_1 + b_2)h. \quad (4.20)$$

Esse método é conhecido na literatura como método de Runge-Kutta de ordem 2. Nesse caso,  $b_1$  representa o valor exato de  $f(u, y)$  em  $u_i$  enquanto que  $b_2$  dá uma estimativa para  $f$  em  $u_{i+1}$ .

Uma estimativa ainda melhor é encontrada quando  $Y_i$  é aproximada usando-se

$$b_1 = f(u_i, y_i) \quad (4.21)$$

$$b_2 = f\left(u_i + \frac{h}{2}, y_i + b_1 \frac{h}{2}\right) \quad (4.22)$$

$$b_3 = f\left(u_i + \frac{h}{2}, y_i + b_2 \frac{h}{2}\right) \quad (4.23)$$

$$b_4 = f(u_i + h, y_i + b_3 h) \quad (4.24)$$

$$y_{i+1} = y_i + \frac{1}{6}(b_1 + 2b_2 + 2b_3 + b_4)h, \quad (4.25)$$

Esse modo de aproximar os resultados do problema de valor inicial é, em geral, conhecido como método de Runge-Kutta de ordem 4. Nesse caso,  $b_1$  representa

o valor exato de  $f(u, y)$  em  $u_i$ ,  $b_2$  e  $b_3$  dão uma estimativa para  $f$  na metade do intervalo e  $b_4$  dá uma estimativa para  $f$  no fim do intervalo.

Assim sendo, considerando o modelo de biorreator anaeróbico dado acima na forma vetorial, tem-se

$$y' = \begin{bmatrix} x'_1(t) \\ x'_2(t) \\ S'_1(t) \\ S'_2(t) \end{bmatrix} = \begin{bmatrix} v_1(S_1(t))x_1(t) \\ v_2(S_2(t))x_2(t) \\ -k_1v_1(S_1(t))x_1(t) \\ k_2v_1(S_1(t))x_1(t) - k_3v_2(S_2(t))x_2(t) \end{bmatrix} = f(t, y), \quad (4.26)$$

onde

$$y = \begin{bmatrix} x_1(t) \\ x_2(t) \\ S_1(t) \\ S_2(t) \end{bmatrix} \quad (4.27)$$

e, por fim, encontra-se a saída de metano a partir da expressão

$$q_M(t) = k_6v_2(S_2(t))x_2(t). \quad (4.28)$$

### 4.3 Resultados Numéricos

Pretende-se, nesta seção, investigar a eficiência dos métodos de Levenberg-Marquardt e *Trust-Region* na identificação de parâmetros de um biorreator do tipo batelada, a partir do Modelo de Antonelli, quando conhecida apenas a saída de gás metano. Se considerará a implementação discreta do modelo com Runge-Kutta de quarta ordem, conforme definida na seção anterior, como a solução real do problema direto associado. Sobre essa solução, serão inseridas perturbações gaussianas de ordem controlada e, então, proceder-se-á com a aplicação de problemas inversos para a estimativa dos parâmetros reais, com posterior análise das soluções aproximadas via mínimos quadrados em relação à sua qualidade.

### 4.3.1 Implementação do problema direto

O modelo apresentado neste capítulo foi implementado, com período de amostragem fixo, usando o método de Runge-Kutta de quarta ordem, que, conforme se observa na literatura [12,22] é suficientemente preciso para a solução do problema direto. Descreve-se, nos próximos passos, brevemente o modo como o problema foi implementado, bem como verifica-se a qualidade da implementação.

Considere-se que o modelo dado por (4.6)-(4.9), cujos estados são representados por (4.27) esteja parametrizado pelo vetor

$$\boldsymbol{\theta} = \left[ \mu_{m1} \quad K_{S1} \quad \mu_{m2} \quad K_{S2} \quad k_1 \quad k_2 \quad k_3 \quad k_6 \right]^T. \quad (4.29)$$

A saída do modelo, dada pela quantidade de gás metano, depende portanto dos parâmetros em  $\boldsymbol{\theta}$ . Para melhor representar essa situação, denomine-se por  $q_M(\mathbf{T})$  os dados experimentais coletados e por  $\hat{q}_M(\mathbf{T}, \boldsymbol{\theta})$  os valores de saída do modelo discretizado, onde  $\mathbf{T}$  representa o conjunto de valores discretos de tempo nos quais o modelo foi aplicado.

Neste trabalho, considerar-se-á o problema exemplo encontrado na literatura [11, 12, 22] em que os parâmetros reais do sistema são dados por

$$\boldsymbol{\theta}_0 = \left[ 0,42912 \quad 13,065 \quad 2,6493 \quad 571,27 \quad 0,31204 \quad 0,062776 \quad 3,1473 \quad 278,62 \right]^T \quad (4.30)$$

e no qual as condições iniciais são dadas por

$$\mathbf{y}(0) = \begin{bmatrix} 0,2 \\ 0,8 \\ 74 \\ 93 \end{bmatrix}. \quad (4.31)$$

A fim de verificar a eficácia da estratégia, compara-se a saída do modelo com as amostras experimentais disponíveis. Como estão disponíveis 120 amostras

$q_M(\mathbf{T})$ , referentes a coletas feitas três vezes ao dia pelo período de 40 dias, usa-se

$$\mathbf{T} = \left[ 0 \quad \frac{1}{3} \quad \frac{2}{3} \quad \cdots \quad 40 \right]^T$$

para a simulação numérica. Considerando esse conjunto de dados inicializadores, obtém-se, a partir do método de Runge-Kutta de 4ª ordem o gráfico apresentado na Figura 4.3, que, percebe-se, está de acordo com os resultados encontrados na literatura [11, 12, 22].

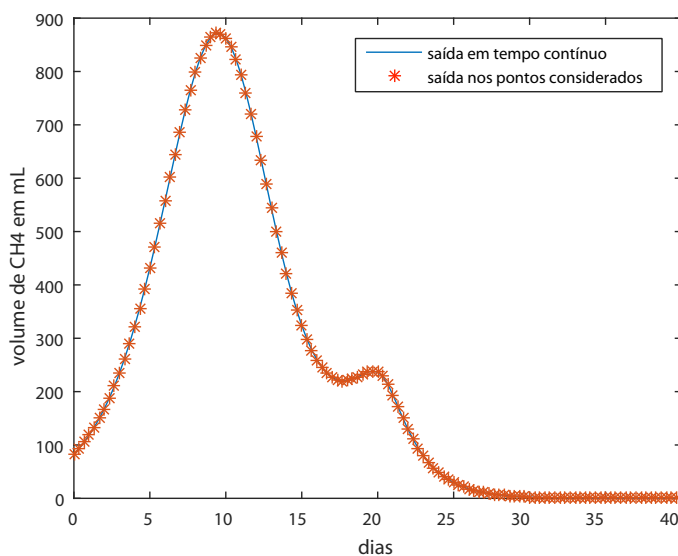


Figura 4.3: Gráfico da simulação da saída de metano no biorreator anaeróbico

### 4.3.2 Análise das estimativas de $\theta$ pela solução de problemas inversos

Considere-se o vetor solução da simulação numérica realizada anteriormente como a solução real do problema direto associado e  $\hat{q}_M$  o conjunto de dados experimentais coletados. Deseja-se, então, determinar os parâmetros do sistema que minimizam o funcional dos mínimos quadrados dado por

$$\mathcal{S}(\theta) = \sum_{i=1}^N [q_M(t_i) - \hat{q}_M(t_i, \theta)]^2. \quad (4.32)$$

Sabe-se que, devido à alta não-linearidade do sistema, esta função objetivo é não-convexa e, portanto, pequenas alterações nas condições iniciais podem conduzir a estimação de mínimos locais, resultando em parâmetros que não explicam de forma adequada a dinâmica interna do reator. Por esse motivo, neste trabalho, a estimativa inicial dada será sempre o vetor  $\theta_0$ , ou seja, pela solução exata do problema inverso, desejando-se observar qual o reflexo da adição de perturbações apenas na saída do modelo para a reconstrução deste conjunto.

Sobre a saída do sistema, adiciona-se ruído gaussiano com média zero e desvio padrão dado por uma porcentagem ( $i \in (0, 1)$ ) do maior valor dentre os  $q_M(T)$ . Após a execução de cem estimações, a partir da solução exata acrescida de erros distintos mas de mesma ordem, a análise estocástica da qualidade dos parâmetros estimados se baseará na análise da média e variância dos resultados encontrados, bem como pela construção de elipses de erros entre pares de parâmetros. Deseja-se verificar se, em média, os resultados são satisfatórios ou se as soluções se desviam do valor original do problema, mesmo quando a média dos erros agregados é nula. Este modo de operação para a análise caracteriza o que se convencionou chamar de métodos de Monte Carlo. Para tal, serão utilizados dois dos métodos apresentados anteriormente, o método de Levenberg-Marquardt, implementado pelo autor, e o Método do tipo *Trust-Region* pré-implementado na biblioteca do Matlab 2015 e chamado pelo comando LSQNONLIN. Os critérios utilizados nos dois métodos são um um resíduo da ordem de  $10^{-12}$  ou então um máximo de dez mil iterações.

Para compreensão das tabelas abaixo, considerar-se-á que  $\bar{\theta}_{LM}$  e  $\bar{\theta}_{LSQ}$  representam as médias das iterações utilizando-se o método de Levenberg-Marquardt e LSQNONLIN, respectivamente. De modo análogo,  $\sigma_{LM}^2$  e  $\sigma_{LSQ}^2$  representam as variâncias de cada um dos parâmetros a partir de cada método e  $errel_{LM}$  e  $errel_{LSQ}$  os erros relativos em relação ao valor real, dado por  $\theta_0$ . Para tal, define-se erro



relativo como

$$errel(\bar{x}, x_0) = \left| \frac{x_0 - \bar{x}}{x_0} \right|, \quad (4.33)$$

onde  $x_0$  refere-se ao valor real do parâmetro e  $\bar{x}$  representa o valor médio encontrado. Em todos os testes, utilizar-se-á estimativa inicial coincidente com o vetor de parâmetros real,  $\theta_0$ .

Considerando-se, inicialmente, a adição de perturbação gaussiana da ordem de 1% ( $i = 0.01$ ) - ou seja, com média zero e desvio padrão de 1% do maior dentre os valores de saída de gás metano do problema direto associado - nos dados de saída do modelo discreto e após a repetição de cem estimações, observa-se o resultado apresentado na Tabela 4.2.

Tabela 4.2: Média, variância e erro relativo das estimativas para chute inicial exato e perturbação gaussiana com desvio padrão de 1% do valor máximo da saída do modelo

	$\theta_0$	Levenberg-Marquardt			LSQNONLIN		
		$\bar{\theta}_{LM}$	$\sigma_{LM}^2$	$errel_{LM}$	$\bar{\theta}_{LSQ}$	$\sigma_{LSQ}^2$	$errel_{LSQ}$
$\mu_{m1}$	0,42912	0,42719	0,00070	0,451%	0,42718	0,00070	0,451%
$K_{S1}$	13,065	12,99404	2,42595	0,543%	12,99398	2,42588	0,544%
$\mu_{m2}$	2,6493	2,65385	0,03976	0,172%	2,65385	0,03976	0,172%
$K_{S2}$	571,27	572,44107	2253,66896	0,205%	572,44005	2253,67204	0,205%
$k_1$	0,31204	0,33678	0,01108	7,927%	0,33678	0,01108	7,929%
$k_2$	0,062776	0,06839	0,00053	8,949%	0,06839	0,00053	8,951%
$k_3$	3,1473	3,15167	0,00141	0,139%	3,15167	0,00141	0,139%
$k_6$	278,62	278,90019	8,46762	0,101%	278,90021	8,46768	0,101%

É possível perceber, nesse caso, que alguns parâmetros são estimados, em média, com maior precisão que outros. Na Figura 4.4, é possível verificar a matriz de covariância, representada por  $\mathbf{C}_{LM}$ , dos resultados obtidos nas cem aproximações usando o método de Levenberg-Marquardt.

$$C_{LM} = \begin{bmatrix} 7,02 \cdot 10^{-4} & 3,79 \cdot 10^{-2} & 2,52 \cdot 10^{-3} & 6,14 \cdot 10^{-1} & -2,65 \cdot 10^{-3} & -5,72 \cdot 10^{-4} & -3,91 \cdot 10^{-4} & -1,44 \cdot 10^{-2} \\ 3,79 \cdot 10^{-2} & 2,426 & 7,51 \cdot 10^{-2} & 1,83 \cdot 10^1 & -1,29 \cdot 10^{-1} & -2,77 \cdot 10^{-2} & -1,23 \cdot 10^{-2} & -2,83 \cdot 10^{-1} \\ 2,52 \cdot 10^{-3} & 7,51 \cdot 10^{-2} & 3,97 \cdot 10^{-2} & 9,463 & -1,13 \cdot 10^{-2} & -2,53 \cdot 10^{-3} & -6,85 \cdot 10^{-3} & -4,67 \cdot 10^{-1} \\ 6,14 \cdot 10^{-1} & 1,83 \cdot 10^1 & 9,463 & 2,25 \cdot 10^3 & -2,756 & -6,16 \cdot 10^{-1} & -1,617 & -1,09 \cdot 10^2 \\ -2,65 \cdot 10^{-3} & -1,29 \cdot 10^{-1} & -1,13 \cdot 10^{-2} & -2,756 & 1,10 \cdot 10^{-2} & 2,41 \cdot 10^{-3} & 1,71 \cdot 10^{-3} & 6,42 \cdot 10^{-2} \\ -5,72 \cdot 10^{-4} & -2,77 \cdot 10^{-2} & -2,53 \cdot 10^{-3} & -6,16 \cdot 10^{-1} & 2,41 \cdot 10^{-3} & 5,27 \cdot 10^{-4} & 3,84 \cdot 10^{-4} & 1,46 \cdot 10^{-2} \\ -3,91 \cdot 10^{-4} & -1,23 \cdot 10^{-2} & -6,85 \cdot 10^{-3} & -1,617 & 1,71 \cdot 10^{-3} & 3,84 \cdot 10^{-4} & 1,41 \cdot 10^{-3} & 1,05 \cdot 10^{-1} \\ -1,44 \cdot 10^{-2} & -2,83 \cdot 10^{-1} & -4,67 \cdot 10^{-1} & -1,09 \cdot 10^2 & 6,42 \cdot 10^{-2} & 1,46 \cdot 10^{-2} & 1,05 \cdot 10^{-1} & 8,468 \end{bmatrix}$$

Figura 4.4: Matriz de covariância das estimativas para chute inicial exato e perturbação gaussiana com desvio padrão de 1% do valor máximo da saída do modelo usando LM

É possível verificar que, embora a variância que  $K_{S2}$  seja grande, em média o valor estimado não se distancia muito do valor real. Na Figura 4.5, apresenta-se o gráfico da saída do problema direto (em azul), da comparação entre esta saída e uma das perturbações agregadas (em vermelho), da comparação entre esta saída e a simulação quando se utiliza o conjunto de parâmetros estimado cujo resíduo com a solução exata do problema direto é o menor entre as execuções (laranja) e, por fim, da comparação no caso em que se utiliza o conjunto de parâmetros estimado cujo resíduo com a solução exata do problema direto é o maior entre as execuções (roxo), onde se percebe que, mesmo no último caso, as saídas praticamente coincidem.

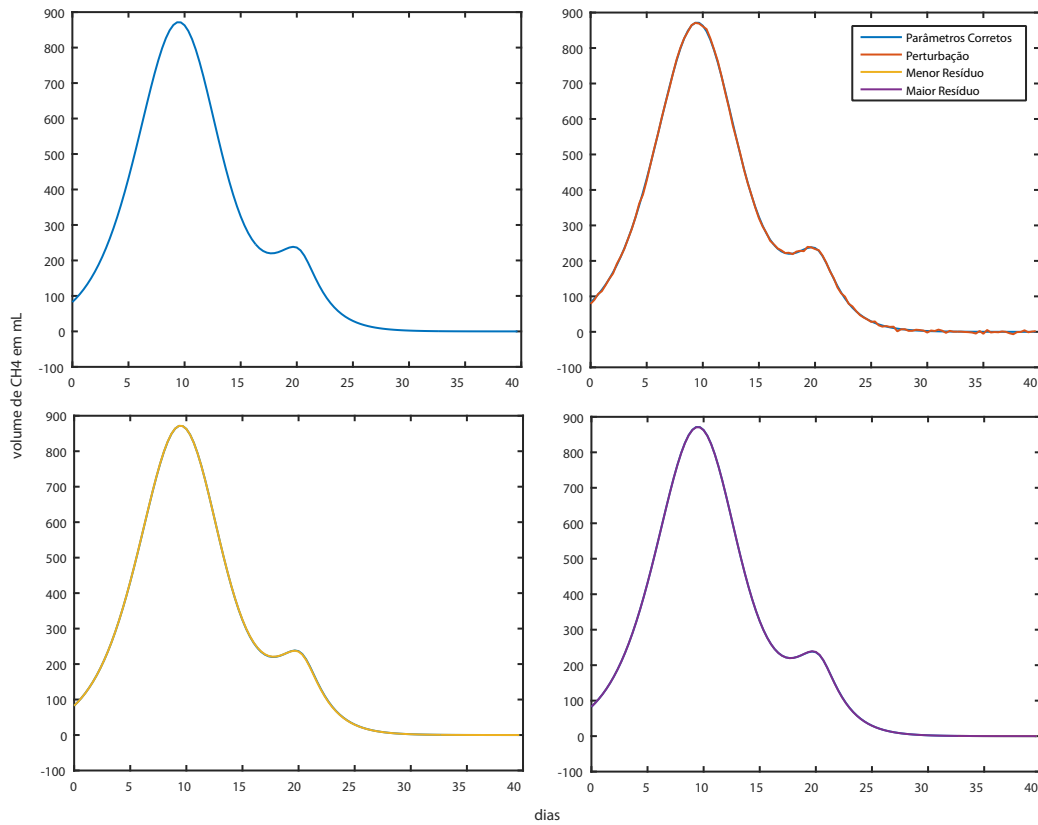


Figura 4.5: Gráficos do modelo usando parâmetros corretos, aquele que gera o menor e aquele que gera o maior resíduos, na reconstrução com perturbação gaussiana com desvio padrão de 1% do valor máximo da saída do modelo direto usando LM

Com o auxílio da matriz de covariância, apresentada anteriormente, foi possível proceder com a construção de elipses de erros para os pares de parâmetros  $\mu_{m1}$  e  $\mu_{m2}$ ,  $K_{S1}$  e  $K_{S2}$ ,  $k_1$  e  $k_2$  e  $k_3$  e  $k_6$ . Conforme verificado anteriormente, mesmo nos casos em que a variância entre essas estimativas foi alta, os resultados tenderam a se concentrar no entorno da média (ponto preto), levando ao resultado anteriormente observado, conforme se observa na Figura 4.6.

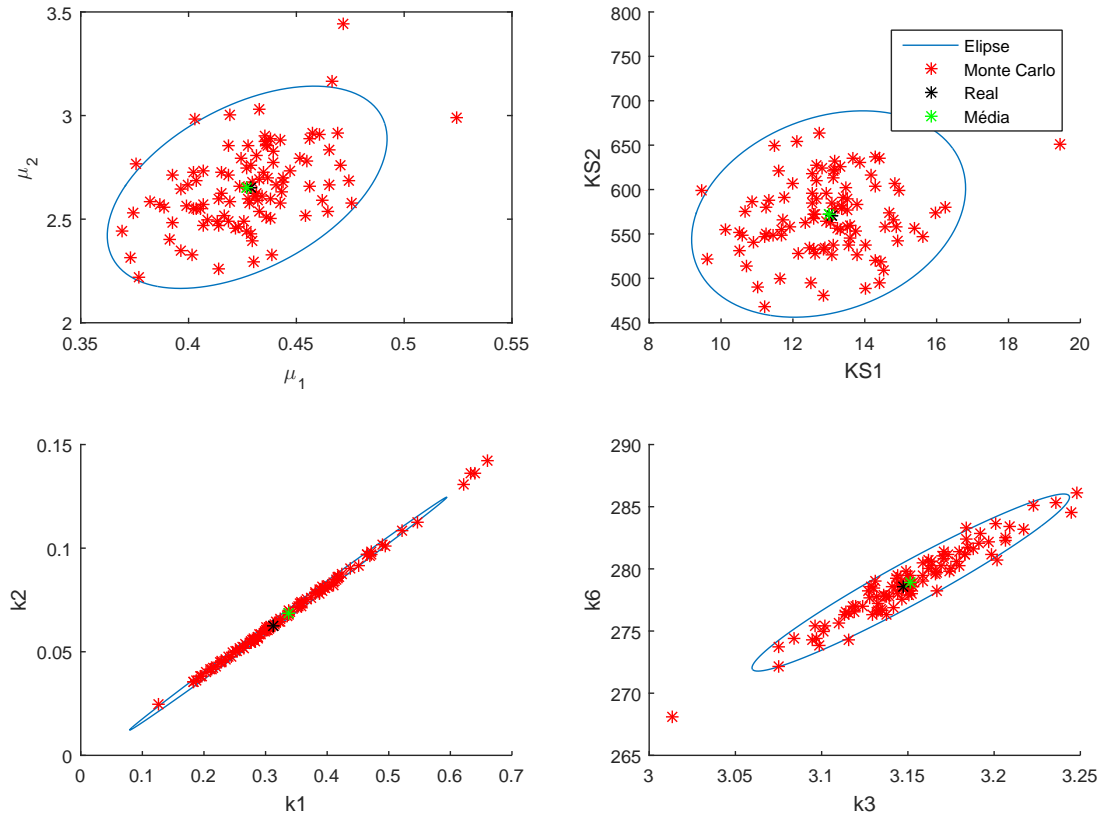


Figura 4.6: Elipses de erro das estimativas para chute inicial exato e perturbação gaussiana com desvio padrão de 1% do valor máximo da saída do modelo usando LM

Conclui-se, portanto, que para perturbações gaussianas da ordem 1% do valor máximo do modelo, os dois métodos se mostram eficientes tanto para a estimação dos parâmetros do sistema, quanto para a simulação numérica.

Considera-se, então, a adição de perturbações gaussianas da ordem de 10% ( $i = 0.1$ ) nos dados de saída do modelo discreto e após a repetição de cem estimações, é possível observar o resultado apresentado na Tabela 4.3.

Tabela 4.3: Média, variância e erro relativo das estimativas para chute inicial exato e perturbação gaussiana com desvio padrão de 10% do valor máximo da saída do modelo

	$\theta_0$	Levenberg-Marquardt			LSQNONLIN		
		$\bar{\theta}_{LM}$	$\sigma_{LM}^2$	$errel_{LM}$	$\bar{\theta}_{LSQ}$	$\sigma_{LSQ}^2$	$errel_{LSQ}$
$\mu_{m1}$	0,42912	4,530	1612,74	956%	0,939	20,789	119%
$K_{S1}$	13,065	312,887	8527921	2295%	51,818	109275,9	297%
$\mu_{m2}$	2,6493	1598,646	129470232	60242%	211,413	1111577	7880%
$K_{S2}$	571,27	381213,190	7,36E+12	66631%	50408,446	63154112057	8724%
$k_1$	0,31204	8,217	800,398	2533%	10,885	1317,97	3388%
$k_2$	0,062776	3,939	290,391	6175%	5,537	483,049	8721%
$k_3$	3,1473	3,329	0,1843	6%	3,337	0,2357	6%
$k_6$	278,62	286,513	766,7048	3%	286,110	926,821	3%

É possível perceber, nesse caso, que as estimativas, além de apresentarem grande variância, ampliam o ruído nos dados, conduzindo a resultados desviados em relação ao valor esperado para os parâmetros. Ou seja, já não há confiança suficiente nos resultados obtidos a partir de problemas inversos quando o erro na saída é de maior magnitude. Na Figura 4.7, por sua vez, é possível verificar a matriz de covariância, representada por  $\mathbf{C}_{LM}$ , dos resultados obtidos nas cem aproximações usando o método de Levenberg-Marquardt.

$$\mathbf{C}_{LM} = \begin{bmatrix} 1,61 \cdot 10^3 & 1,17 \cdot 10^5 & -5,59 \cdot 10^3 & -1,33 \cdot 10^6 & -3,66 \cdot 10^1 & -1,77 \cdot 10^1 & -1,93 \cdot 10^0 & 8,46 \cdot 10^1 \\ 1,17 \cdot 10^5 & 8,53 \cdot 10^6 & -4,20 \cdot 10^5 & -1,00 \cdot 10^8 & -2,57 \cdot 10^3 & -1,24 \cdot 10^3 & -1,38 \cdot 10^2 & -6,11 \cdot 10^3 \\ -5,59 \cdot 10^3 & -4,20 \cdot 10^5 & 1,29 \cdot 10^8 & 3,09 \cdot 10^{10} & -1,32 \cdot 10^4 & -6,34 \cdot 10^3 & -9,65 \cdot 10^2 & -6,37 \cdot 10^4 \\ -1,33 \cdot 10^6 & -1,00 \cdot 10^8 & 3,09 \cdot 10^{10} & 7,37 \cdot 10^{12} & -3,14 \cdot 10^6 & -1,51 \cdot 10^6 & -2,30 \cdot 10^5 & -1,52 \cdot 10^7 \\ -3,66 \cdot 10^1 & -2,57 \cdot 10^3 & -1,32 \cdot 10^4 & -3,14 \cdot 10^6 & 8,00 \cdot 10^2 & 4,79 \cdot 10^2 & 8,02 \cdot 10^0 & 2,52 \cdot 10^2 \\ -1,77 \cdot 10^1 & -1,24 \cdot 10^3 & -6,34 \cdot 10^3 & -1,51 \cdot 10^6 & 4,79 \cdot 10^2 & 2,90 \cdot 10^2 & 3,78 \cdot 10^0 & 1,60 \cdot 10^2 \\ -1,93 \cdot 10^0 & -1,38 \cdot 10^2 & -9,65 \cdot 10^2 & -2,30 \cdot 10^5 & 8,02 \cdot 10^0 & 3,78 \cdot 10^0 & 1,84 \cdot 10^{-1} & 1,05 \cdot 10^1 \\ -8,46 \cdot 10^1 & -6,11 \cdot 10^3 & -6,37 \cdot 10^4 & -1,52 \cdot 10^7 & 2,52 \cdot 10^2 & 1,60 \cdot 10^2 & 1,05 \cdot 10^1 & 7,67 \cdot 10^2 \end{bmatrix}$$

Figura 4.7: Matriz de covariância das estimativas para chute inicial exato e perturbação gaussiana com desvio padrão de 10% do valor máximo da saída do modelo usando LM

Na Figura 4.8, apresenta-se, de modo análogo ao feito na Figura 4.5, a simulação usando os parâmetros corretos, uma das perturbações agregadas sobre este e a comparação entre as simulações nos casos em que se utiliza os parâmetros estimados que geram, com a solução exata do problema direto, o menor e maior resíduo. É interessante verificar que, novamente, mesmo no caso em que se obtém o maior resíduo quadrático, com um conjunto estimado de parâmetros pouco relacionados com  $\theta_0$ , a simulação numérica obtida é consideravelmente boa. Pode-se reparar, na Figura 4.8, que, mesmo nesse caso, os gráficos da simulação real e aproximada quase coincidem.

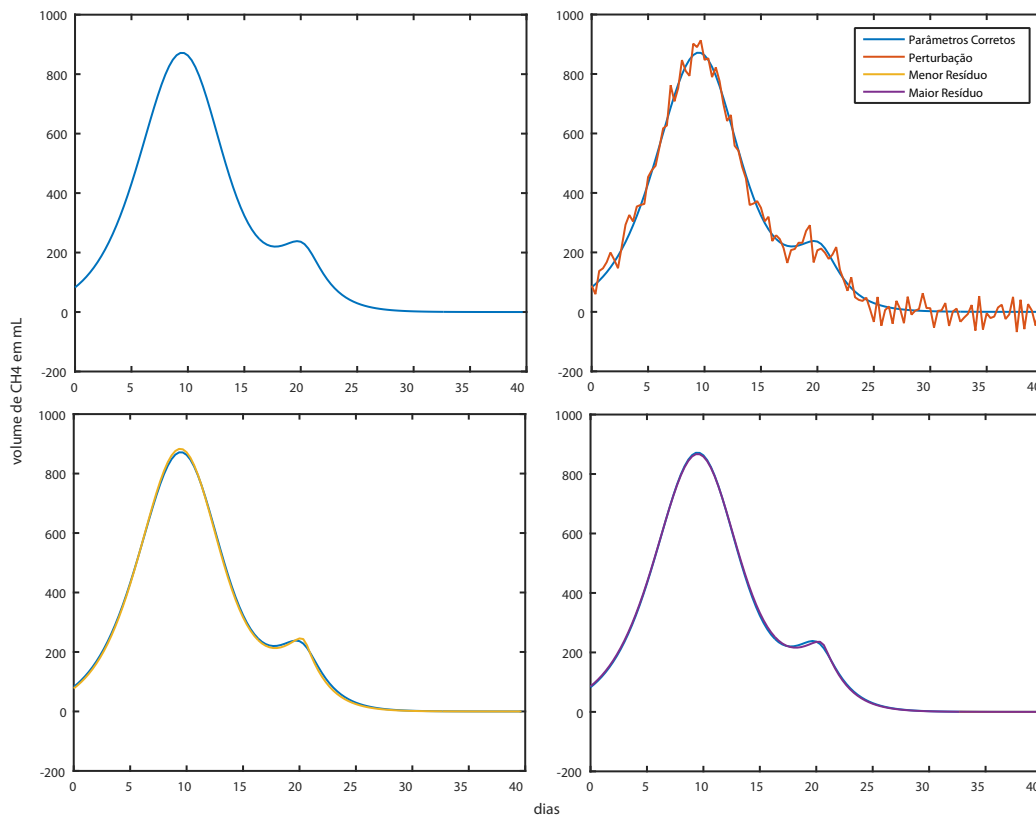


Figura 4.8: Gráficos do modelo usando parâmetros corretos, aquele que gera o menor e aquele que gera o maior resíduo, na reconstrução com perturbação gaussiana com desvio padrão de 10% do valor máximo da saída do modelo direto usando LSQNONLIN

Com o auxílio da matriz de covariância, uma investigação mais apurada pode ser feita com a construção de elipses de erros para os pares de parâmetros  $\mu_{m1}$  e  $\mu_{m2}$ ,  $K_{S1}$  e  $K_{S2}$ ,  $k_1$  e  $k_2$ ,  $k_3$  e, por fim,  $k_6$ .

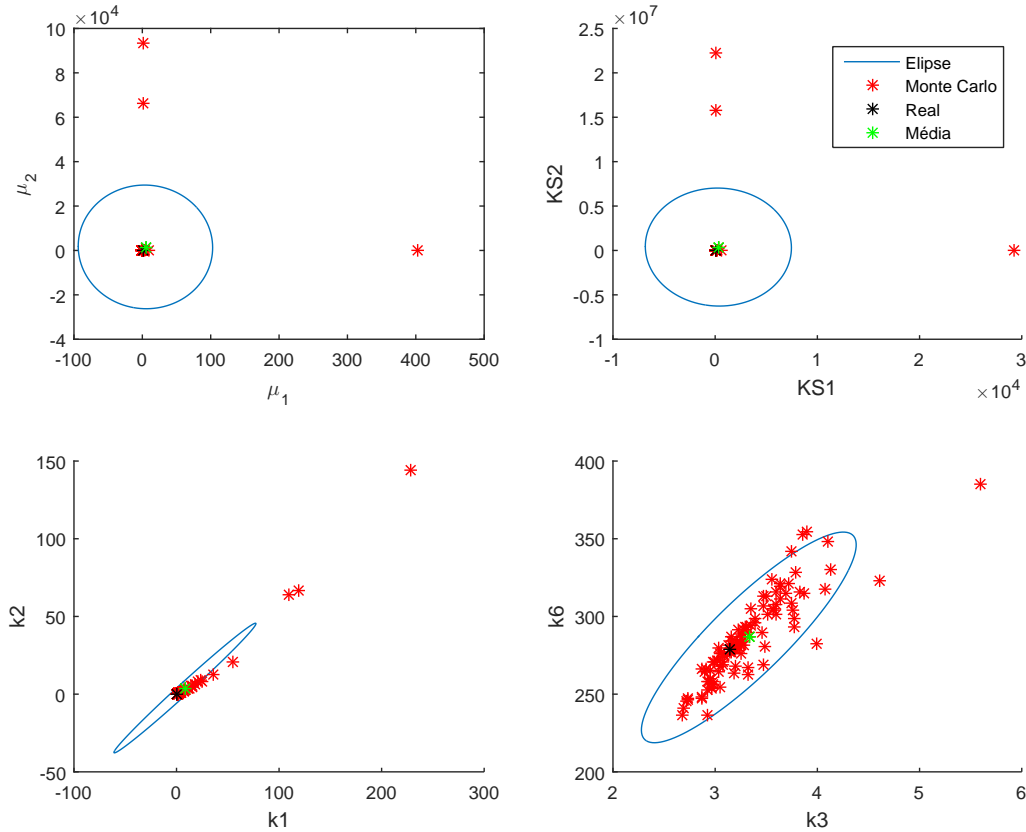


Figura 4.9: Elipses de erro das estimativas para chute inicial exato e perturbação gaussiana com desvio padrão de 10% do valor máximo da saída do modelo usando LSQNONLIN

A partir das elipses de erro, é possível verificar que a maioria dos valores concentra-se, também, em torno da média, havendo, contudo, polarização das estimativas. É possível verificar, ainda, que algumas das execuções conduziram a

mínimos locais distantes dos valores esperados, o que pode explicar os resultados obtidos na Tabela 4.3.

Deve-se observar que, mesmo usando-se a solução exata do problema inverso, quando a saída não é perturbada,  $\theta_0$ , como chute inicial para a reconstrução de parâmetros, mínimos locais foram encontrados, nos dois níveis de erros, o que mostra a alta não-linearidade do modelo ou, então, a característica de não-unicidade da solução quando não se têm acesso aos dados não perturbados de saída do problema direto. A partir das observações pode-se concluir, contudo, que quando a perturbação sobre os valores observados é suficientemente pequena, é possível estimar com boa precisão os parâmetros do sistema, já nos casos em que a perturbação nos dados é mais significativa, os resultados das estimativas não são suficientemente confiáveis mas, ainda assim, podem ser utilizados para o estudo do comportamento do reator via simulações numéricas.



## 5 TRANSPORTE DE PARTÍCULAS

O transporte de partículas e radiação é, em geral, modelado pela versão linear da Equação de Boltzmann, esta que foi derivada, inicialmente, para o tratamento de fenômenos em dinâmica de gases rarefeitos [14]. Um dado de interesse nas aplicações é a estimativa do fluxo de radiação ou intensidade radiativa. A **intensidade radiativa** é a quantidade de radiação recebida por um objeto desde uma determinada fonte, podendo se apresentar, por exemplo, na forma de partículas.

A composição dos materiais nos quais ocorre o transporte ou, então, a estimação das fontes de radiação representam problemas de interesse com aplicações tanto na solução do problema direto, quanto de problemas inversos em transporte. No que se refere aos problemas inversos em radiação, muito já foi feito desde a década de oitenta. Cita-se, como exemplos, a determinação de radiação incidente nos contornos de uma barra [30], em radiação de calor, a estimação de propriedades radiativas e termofísicas em materiais a altas temperaturas [34], a identificação de perfis de temperatura em um meio [43], estimação da lei de espalhamento em fissão em termos do fluxo radiativo [35,36,52,53] e do albedo em radiação atmosférica [51], a reconstrução de fontes externas de radiação [44] e a determinação do meio em tomografias computadorizadas [3,5].

Neste trabalho, considera-se a possibilidade de identificar sobretudo dois parâmetros a partir do fluxo de radiação conhecido: o albedo e o grau de anisotropia, em transferência radiativa. Estes coeficientes, relacionados ao modelo considerado, descrevem o meio material em relação a quanto absorvedor este é ou, então, como se dá o espalhamento de radiação, sendo de vital importância no estudo de tomografias [3].

Ainda em 1917, Johann Radon [49] atentou para a possibilidade de reconstruir imagens a partir de diversas projeções. Sua abordagem levou, mais tarde, ao desenvolvimento de procedimentos médicos para diagnóstico, como é o caso da tomografia computadorizada. Esta foi desenvolvida durante as décadas de sessenta e setenta, sobretudo por Cormack e Hounsfield, em um estudo que lhes rendeu o Nobel em 1979. Neste procedimento, “a absorção é o mecanismo dominante e medidas experimentais da radiação atenuada pelo meio sob análise permitem identificar estruturas do mesmo [57]. Em outro tipo de tomografia, mais recente, do tipo *Near Infrared Optical Tomography*, o espalhamento também passa a ser levado em consideração, “estando então no mesmo contexto de transferência de calor por radiação térmica em meios participantes e transporte de nêutrons em reatores nucleares” [57].

Para a descrição do transporte tridimensional de partículas em um meio material, em particular o nêutron, considera-se o mesmo como uma **partícula pontual**. Isso significa que, nesse contexto, uma partícula pode ser caracterizada por sua posição e velocidade. Dessa forma, além da variável temporal, o chamado **espaço de fase**, no qual se considera o balanço de partículas, inclui ainda três variáveis espaciais e três referentes ao vetor de velocidade, totalizando sete variáveis. Sejam, então

$$\mathbf{X} = \begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix}^T, \quad (5.1)$$

$$\mathbf{v} = \begin{bmatrix} v_1 & v_2 & v_3 \end{bmatrix}^T \quad (5.2)$$

e o instante de tempo  $t$  considerado e caracterize-se a velocidade por

$$\mathbf{v} = v\boldsymbol{\Omega},$$

onde  $v = |\mathbf{v}|$  e  $\boldsymbol{\Omega}$  é um vetor unitário de direção. Levando-se em conta, ainda, a energia cinética descrita por

$$E = \frac{1}{2}mv^2, \quad (5.3)$$

onde  $m$  é a massa da partícula, conclui-se que é possível, também, substituir a magnitude da velocidade das partículas por sua energia.

O fenômeno físico considerado é a migração dessas partículas pontuais no interior do meio material, com a probabilidade de interação com os núcleos dos átomos do meio. Uma das formas de interação é a **colisão**, definida como o choque dos nêutrons com núcleos, seguida pela absorção ou emissão de energia ou outras partículas [25]. Algumas das reações esperadas são a captura radiativa, ou absorção, o espalhamento e a fissão. A **absorção** é uma reação na qual um nêutron incidente é absorvido pelo núcleo-alvo, deixando-o em um estado excitado. O **espalhamento**, por sua vez, que pode ser elástico ou inelástico, é aquele onde o nêutron atinge o núcleo e é desviado. No espalhamento elástico, o núcleo não tem seu estado alterado, já no espalhamento inelástico, o núcleo é excitado, decaindo posteriormente [44]. Nessas reações, uma partícula costuma ser considerada um projétil que atinge outra partícula-alvo.

Ao tratar de partículas neutras, é necessário compreender sua natureza randômica, o que implica na impossibilidade de determinar um número exato de partículas em uma certa região, mas sim trabalhar com uma distribuição de probabilidades e prever um número esperado de partículas nessa. Define-se a função de distribuição

$$f(\mathbf{X}, E, \boldsymbol{\Omega}, t) \quad (5.4)$$

como aquela que descreve a distribuição de partículas em um meio material em relação a seu espaço, direção e energia. A **densidade angular de nêutrons**

$$N(\mathbf{X}, E, \boldsymbol{\Omega}, t) \quad (5.5)$$

representa, por sua vez, o número provável de nêutrons presente numa esfera de volume  $dV$  ao redor de  $\mathbf{X}$  que viajam em direções  $d\boldsymbol{\Omega}$  ao redor de  $\boldsymbol{\Omega}$  com energias no intervalo  $dE$  ao redor de  $E$  em um determinado instante  $t$ . Assim, considerando-se a caracterização pontual, a taxa de variação do número esperado de partículas em

torno de  $\mathbf{X}$  e, portanto, na região delimitada pela bola de volume  $dV$ , com energia no intervalo  $[E, E+dE]$  no ângulo sólido  $d\Omega$  e no instante  $t$  pode ser representada [25] por

$$\left[ \frac{d}{dt} N(\mathbf{X}, E, \Omega, t) \right] dV dE d\Omega. \quad (5.6)$$

A probabilidade de interação de um nêutron por unidade de comprimento de sua trajetória é chamada de **seção de choque macroscópica** e depende da posição e da energia da partícula. Esta costuma ser representada por uma letra sigma maiúscula ( $\Sigma$ ), mas é comum, nos trabalhos sobre teoria de transporte, ver a representação por uma letra sigma minúscula ( $\sigma$ ) [25], notação que será adotada doravante. A seção de choque macroscópica total é definida [6, 25] como

$$\sigma(\mathbf{X}, E) = \sigma_a(\mathbf{X}, E) + \sigma_s(\mathbf{X}, E) + \sigma_{op}(\mathbf{X}, E), \quad (5.7)$$

onde,  $\mathbf{X}$  representa o vetor de posição,  $E$  a energia,  $\sigma_a(\mathbf{X}, E)$  a seção de choque macroscópica de absorção,  $\sigma_s(\mathbf{X}, E)$  a seção de choque macroscópica de espalhamento e  $\sigma_{op}(\mathbf{X}, E)$  está associada à seção de choque que leva em conta outros processos de interação do nêutron com o núcleo, menos importantes que a absorção e espalhamento [25]. Estas seções de choque podem, ainda, ser decompostas como

$$\sigma_a(\mathbf{X}, E) = \sigma_{cr}(\mathbf{X}, E) + \sigma_f(\mathbf{X}, E) \quad (5.8)$$

$$\sigma_s(\mathbf{X}, E) = \sigma_n(\mathbf{X}, E) + \sigma_{n'}(\mathbf{X}, E), \quad (5.9)$$

onde os subíndices de  $\sigma$  representam, respectivamente, as seções de choque de captura radiativa ( $cr$ ), de fissão ( $f$ ), de espalhamento elástico ( $n$ ) e de espalhamento inelástico ( $n'$ ).

Há ainda a chamada **seção de choque diferencial**, dada por

$$\sigma(\mathbf{X}, E' \rightarrow E, \Omega' \rightarrow \Omega) = \sigma(\mathbf{X}, E') f(\mathbf{X}, E' \rightarrow E, \Omega' \rightarrow \Omega), \quad (5.10)$$

que depende da seção de choque macroscópica e da função **probabilidade de transferência**, dada por

$$f(\mathbf{X}, E' \rightarrow E, \Omega' \rightarrow \Omega) \quad (5.11)$$

que representa a probabilidade de um nêutron com direção  $\Omega'$  e energia  $E'$ , ao sofrer reação com um núcleo, causar a emissão de nêutrons no intervalo com direções  $d\Omega$  em torno de  $\Omega$  e energias no intervalo  $dE$  em torno de  $E$  [50]. Essa reação é normalizada sobre todas as energias finais e a probabilidade total passa a ser definida sobre todas as reações [25] que produzem nêutrons na forma

$$\sigma(\mathbf{X}, E')f(\mathbf{X}, E' \rightarrow E, \Omega' \rightarrow \Omega) = \sum_{\mathbf{X}} \sigma(\mathbf{X}, E')f(\mathbf{X}, E' \rightarrow E, \Omega' \rightarrow \Omega). \quad (5.12)$$

O **fluxo angular de nêutrons**,  $\Psi(\mathbf{X}, \Omega, E, t)$ , por sua vez, é dado por

$$\Psi(\mathbf{X}, \Omega, E, t) = vN(\mathbf{X}, \Omega, E, t), \quad (5.13)$$

ou seja, sua unidade é dada em quantidade de nêutrons por unidade de área, por unidade de tempo, por unidade de ângulo sólido e por unidade de energia.

## 5.1 Descrição do modelo considerado

A dedução da equação de transporte pode ser encontrada em diversas referências na literatura [6, 17, 25, 37, 42] e não será detalhada aqui. No entanto, considerando-se as definições anteriores, pode-se introduzir uma ideia geral do balanço de partículas. Considere-se  $v'$  o módulo da velocidade do nêutron com energia  $E'$ , a taxa de produção de nêutrons é dada [25] por

$$\left[ \int_{E'} \int_{\Omega'} \sigma(\mathbf{X}, E')f(\mathbf{X}, \Omega' \rightarrow \Omega, E' \rightarrow E)v'N(\mathbf{X}, \Omega', E', t)d\Omega'dE' + S(\mathbf{X}, \Omega, E) \right] dV d\Omega dE, \quad (5.14)$$

onde o primeiro termo representa os nêutrons que são transferidos a outras direções e energias após colidirem em  $dV$ , cuja integral se dá em todas as direções e energias,

e o segundo termo representa nêutrons produzidos por **fontes externas** localizadas no interior do meio material, enquanto que a taxa de perda de nêutrons é dada por

$$\left[ v\boldsymbol{\Omega} \cdot \nabla N(\mathbf{X}, \boldsymbol{\Omega}, E, t) + \sigma(\mathbf{X}, E)vN(\mathbf{X}, \boldsymbol{\Omega}, E, t) \right] dV d\boldsymbol{\Omega} dE \quad (5.15)$$

onde o primeiro termo, chamado de **termo de fuga**, representa a fuga líquida de nêutrons do volume angular considerado, enquanto que o segundo termo corresponde aos nêutrons que deixam os intervalos angulares e energéticos considerados devido a interações no volume espacial considerado [25, 50]. Assim, considerando que a taxa de variação do número mais provável de nêutrons em  $dV$ , dada por (5.6) e procedendo com a substituição pelo fluxo angular de nêutrons, dado por (5.13), é possível encontrar a forma mais geral da equação de transporte, dada por

$$\begin{aligned} & \frac{1}{v} \frac{\partial}{\partial t} \Psi(\mathbf{X}, \boldsymbol{\Omega}, E, t) + \boldsymbol{\Omega} \cdot \nabla \Psi(\mathbf{X}, \boldsymbol{\Omega}, E, t) + \sigma(\mathbf{X}, E) \Psi(\mathbf{X}, \boldsymbol{\Omega}, E, t) = \\ & = \int_{E'} \int_{\boldsymbol{\Omega}'} \sigma(\mathbf{X}, E') f(\mathbf{X}, \boldsymbol{\Omega}' \rightarrow \boldsymbol{\Omega}, E' \rightarrow E) \Psi(\mathbf{X}, \boldsymbol{\Omega}', E', t) d\boldsymbol{\Omega}' dE' + S(\mathbf{X}, \boldsymbol{\Omega}, E). \end{aligned} \quad (5.16)$$

Pode-se observar que, nesta forma de escrita, todos os termos que se encontram no lado esquerdo da igualdade representam perdas enquanto que, os termos do lado direito da igualdade, representam produção de novos nêutrons na região considerada.

Deve-se ressaltar que, nesta equação, considera-se que as partículas descrevam trajetória retilínea entre colisões e se desconsidera qualquer espécie de interação entre as mesmas. Considera-se, ainda, que as propriedades do material independam do tempo e que as colisões entre as mesmas sejam instantâneas [44]. No caso da equação linearizada, negligencia-se ainda a interação entre partículas e se considera que as colisões são instantâneas, ou seja, que as partículas são emitidas exatamente após as colisões [25].

Neste trabalho, considera-se apenas problemas em estado estacionário, ou seja, aqueles nos quais os nêutrons atingiram o equilíbrio em relação à variável de tempo, de modo que a equação geral pode ser reduzida a

$$\begin{aligned}
& \boldsymbol{\Omega} \cdot \nabla \Psi(\mathbf{X}, \boldsymbol{\Omega}, E) + \sigma(\mathbf{X}, E) \Psi(\mathbf{X}, \boldsymbol{\Omega}, E) = \\
& = \int_{E'} \int_{\boldsymbol{\Omega}'} \sigma(\mathbf{X}, E') f(\mathbf{X}, \boldsymbol{\Omega}' \rightarrow \boldsymbol{\Omega}, E' \rightarrow E) \Psi(\mathbf{X}, \boldsymbol{\Omega}', E') d\boldsymbol{\Omega}' dE' + S(\mathbf{X}, \boldsymbol{\Omega}, E).
\end{aligned} \tag{5.17}$$

Além disso, sabe-se que o termo de fuga é o único na equação que depende explicitamente da geometria do problema [25]. Tomando o problema em geometria plana com simetria axial, tem-se

$$\boldsymbol{\Omega} \cdot \nabla = \mu \frac{\partial}{\partial z}, \tag{5.18}$$

com

$$\mu = \boldsymbol{\Omega} \cdot \mathbf{k} = \cos\theta, \tag{5.19}$$

onde  $\mathbf{k}$  representa o versor na direção do eixo  $z$ . Considerando-se ainda simetria azimutal, nos casos em que as seções de choque forem independentes de energia ou, de modo equivalente, em que todas as partículas neutras tenham a mesma energia, pode-se escrever a equação como

$$\mu \frac{\partial}{\partial z} \Psi(z, \mu) + \sigma \Psi(z, \mu) = \sigma c \int_{-1}^1 f(\mu' \rightarrow \mu) \Psi(z, \mu') d\mu' + S(z, \mu), \tag{5.20}$$

onde  $c$  representa o número **médio de nêutrons produzido pela interação nêutron-núcleo** e está associado ao número médio de nêutrons no espalhamento.

Definindo-se, por fim, a variável ótica  $x = \sigma z$ , adimensional [25], é possível reescrever a equação como

$$\mu \frac{\partial}{\partial x} \Psi(x, \mu) + \Psi(x, \mu) = c \int_{-1}^1 f(\mu' \rightarrow \mu) \Psi(x, \mu') d\mu' + Q(x, \mu), \tag{5.21}$$

onde

$$Q(x, \mu) = \frac{S(x, \mu)}{\sigma}. \tag{5.22}$$

Por fim, é comum que se expresse a probabilidade de transferência  $f(\mu' \rightarrow \mu)$  através de uma expansão truncada de polinômios de Legendre [6, 25],

na forma

$$f(\mu' \rightarrow \mu) = \frac{1}{2} \sum_{l=0}^L \beta_l P_l(\mu') P_l(\mu), \quad (5.23)$$

onde  $L$  representa o grau de anisotropia do sistema, os coeficientes  $\{\beta_l\}$  são tais que  $\beta_0 = 1$  e  $|\beta_l| < 1$  para  $1 \leq l \leq L$  e  $P_l(\mu)$  representa o polinômio de Legendre de ordem  $l$  computado em  $\mu$ . Há ao menos duas formas de descrever a lei de espalhamento da função de transferência, uma delas é a forma binomial, na qual os coeficientes  $\beta_l$  são calculados a partir de uma fórmula de recursão dada [6] por

$$\beta_l = \left( \frac{2l+1}{2l-1} \right) \left( \frac{L+1-l}{L+1+l} \right) \beta_{l-1}, \quad (5.24)$$

para  $l = 1, 2, \dots, L$ , com  $\beta_0 = 1$ .

A equação de transporte é, então, finalmente escrita como

$$\begin{aligned} \mu \frac{\partial}{\partial z} \Psi(x, \mu) + \Psi(x, \mu) = \\ = \frac{c}{2} \sum_{l=0}^L \beta_l P_l(\mu) \int_{-1}^1 P(\mu') \Psi(x, \mu') d\mu' + Q(x, \mu), \end{aligned} \quad (5.25)$$

sujeita às condições de contorno

$$\Psi(0, \mu) = F_1(\mu) \quad (5.26)$$

e

$$\Psi(a, -\mu) = F_2(\mu), \quad (5.27)$$

onde  $x \in (0, a)$ ,  $\mu \in [-1, 1]$  e  $F_1(\mu)$  e  $F_2(\mu)$  são funções conhecidas que descrevem a incidência externa de partículas nos contornos da placa. Além disso, (5.25) é chamada de homogênea quando  $Q(x, \mu) = 0$ . Esta equação representa o problema de transporte unidimensional com simetria azimutal, que será objeto de estudo deste trabalho.

Deve-se observar que esta mesma equação modela o problema em que se considera a propagação de um feixe de radiação com um determinado intervalo



de comprimento de onda transportado em um elemento de área  $dA$  em todas as direções de um elemento de ângulo sólido  $d\Omega$ . A dedução da equação, nesse caso, pode ser encontrada na literatura [6]. Destaca-se que o coeficiente  $c$ , o número médio de nêutrons na interação, representa então o **albedo** do espalhamento simples do problema de radiação, dado pela razão entre o coeficiente de espalhamento e o coeficiente de extinção [6], enquanto que  $x \in (0, a)$  representa a variável ótica, com  $a$  representando a espessura do meio plano-paralelo.

## 5.2 Solução em forma fechada para problema direto usando o método *ADO*

Considere a equação de transporte dada pela equação (5.25) para  $x \in (0, a)$ ,  $\mu \in [-1, 1]$ , com as condições de contorno (5.26) e (5.27) e  $c < 1$ , ou seja, um meio não-multiplicador e não-conservativo.

Diversas técnicas analíticas, estocásticas e numéricas foram propostas para a solução da equação de transporte. Cita-se, como exemplos, o método de Case, a expansão em harmônicos esféricos, também conhecido como método  $P_N$ , o método das ordenadas discretas, ou método  $S_N$ , métodos estocásticos como os de Monte Carlo [6, 25, 48].

O método de ordenadas discretas, originalmente desenvolvido por Wick [59] e Chandrasekhar [17], foi proposto pelo segundo em seu trabalho sobre transferência radiativa na atmosfera e se baseia na discretização da variável angular da intensidade radiativa ( $\mu$ ). A solução por ordenadas discretas para o problema de transporte é encontrada, então, em direções discretas ao longo da esfera unitária. O método de ordenadas discretas considera portanto, a aproximação da integral angular no termo de espalhamento por uma fórmula de quadraturas [25].

O método *ADO*, uma modificação do método das ordenadas discretas proposto por Chandrasekhar, e desenvolvido por Barichello e Siewert [6, 7] apresenta uma série de contribuições positivas em relação à proposta original. Além de manter as soluções analíticas em relação à variável espacial, os problemas de autovalores associados à determinação das constantes de separação na solução têm ordem menor, em geral da metade do tamanho do problema original. Além disso, por trabalhar com o mapeamento das direções no intervalo  $\mu \in [0, 1]$ , é possível utilizar esquemas arbitrários de quadratura, o que torna o método viável para a aplicação em diversas áreas. Recentemente, problemas inversos de reconstrução de fontes [44] foram apresentados usando-se como solução do problema direto o método *ADO*.

O primeiro passo para encontrar a solução por este método se baseia em reescrever o termo integral da equação (5.25) em um semi-intervalo, obtendo-se

$$\begin{aligned} \mu \frac{\partial}{\partial z} \Psi(x, \mu) + \Psi(x, \mu) &= \\ &= \frac{c}{2} \sum_{l=0}^L \beta_l P_l(\mu) \int_0^1 P(\mu') [\Psi(x, \mu') + (-1)^l \Psi(x, -\mu')] d\mu' + Q(x, \mu). \end{aligned} \quad (5.28)$$

Após, procede-se com a definição da quadratura a ser utilizada na aproximação da integral. Sabe-se que a aproximação da integral de uma função por quadraturas numéricas, usando-se a quadratura de Gauss-Legendre, segue a relação [6]

$$\int_{-1}^1 f(x) dx = \sum_{i=1}^m w_i f(x_i) \quad (5.29)$$

onde os chamados nós  $\{x_i\}$  são os zeros do polinômio de Legendre de ordem  $m$  e  $\{w_i\}$  são pesos associados, com a aproximação se tornando exata quando  $f(x)$  for um polinômio de grau

$$\text{grau}(f) \leq 2m - 1. \quad (5.30)$$

Quando o intervalo de integração é um segmento finito arbitrário  $[a, b]$ , é possível estabelecer um mapeamento linear que relacione a integral original com

uma integral definida sobre o intervalo  $[-1, 1]$ , como na relação anterior [25]. Conforme citado anteriormente, uma das contribuições do método *ADO* é o mapeamento do intervalo  $[-1, 1]$  no intervalo  $[0, 1]$ , o que permite a utilização não apenas da quadratura de Gauss-Legendre [6], mas também outros esquemas arbitrários.

Embora existam tabelas, na literatura [29], apontando os pontos e pesos dessa quadratura para diversos valores de  $m$ , é possível derivar um método eficiente para computar esses nós [6], que se baseia na fórmula de recorrência dos polinômios de Legendre. Do mesmo modo, os pesos podem ser computados a partir de relações já encontradas na literatura [29]. O mapeamento dos nós e pesos para o intervalo de interesse é dado, então, por [6, 25]

$$\mu_k = \frac{y_k + 1}{2} \text{ e } w_k = \frac{1}{2v_k}. \quad (5.31)$$

Após definida a quadratura  $\{\mu_i, w_i\}_{i=1}^m$ , pode-se derivar as equações das ordenadas discretas da equação de transporte na forma

$$\mu_i \frac{\partial}{\partial x} \Psi(x, \mu_i) + \Psi(x, \mu_i) = \frac{c}{2} \sum_{l=0}^L \beta_l P_l(\mu_i) \sum_{k=1}^m w_k P_l(\mu_k) [\Psi(x, \mu_k) + (-1)^l \Psi(x, -\mu_k)] \quad (5.32)$$

e

$$-\mu_i \frac{\partial}{\partial x} \Psi(x, -\mu_i) + \Psi(x, -\mu_i) = \frac{c}{2} \sum_{l=0}^L \beta_l P_l(\mu_i) \sum_{k=1}^m w_k P_l(\mu_k) [\Psi(x, -\mu_k) + (-1)^l \Psi(x, \mu_k)], \quad (5.33)$$

que compõem um sistema de  $2m$  equações diferenciais ordinárias [6, 25]. Esse conjunto de equações certamente admite soluções exponenciais [25]. Propõe-se, então, pelo método *ADO*, a hipótese de solução [6]

$$\Psi(x, \pm\mu_i) = \Phi(\nu, \pm\mu_i) e^{-x/\nu}, \quad (5.34)$$

onde  $\nu$  representa uma constante de separação e  $\Phi(\nu, \pm\mu_i)$  é uma função cujo comportamento é, por hora, desconhecido.

Após essa substituição, obtém-se [6]

$$\begin{aligned} & \pm \mu_i \frac{\partial}{\partial x} [\Phi(\nu, \pm \mu_i) e^{-x/\nu}] + \Phi(\nu, \pm \mu_i) e^{-x/\nu} = \\ & = \frac{c}{2} \sum_{l=0}^L \beta_l P_l(\mu_i) \sum_{k=1}^m w_k P_l(\mu_k) [\Phi(\nu, \pm \mu_k) e^{-x/\nu} + (-1)^l \Phi(\nu, \mp \mu_k) e^{-x/\nu}]. \end{aligned} \quad (5.35)$$

Operando essa expressão, é possível obter

$$\begin{aligned} & [\pm \mu_i \frac{\partial}{\partial x} [e^{-x/\nu}] + e^{-x/\nu}] \Phi(\nu, \pm \mu_i) = \\ & = e^{-x/\nu} \frac{c}{2} \sum_{l=0}^L \beta_l P_l(\mu_i) \sum_{k=1}^m w_k P_l(\mu_k) [\Phi(\nu, \pm \mu_k) + (-1)^l \Phi(\nu, \mp \mu_k)] \end{aligned} \quad (5.36)$$

que pode ser simplificada, após aplicação da derivada, para

$$[1 \mp \frac{\mu_i}{\nu}] \Phi(\nu, \pm \mu_i) = \frac{c}{2} \sum_{l=0}^L \beta_l P_l(\mu_i) \sum_{k=1}^m w_k P_l(\mu_k) [\Phi(\nu, \pm \mu_k) + (-1)^l \Phi(\nu, \mp \mu_k)]. \quad (5.37)$$

Introduz-se, então, notação matricial. Sejam os vetores  $m \times 1$

$$\Phi_{\pm}(\nu) = [\Phi(\nu, \pm \mu_1), \Phi(\nu, \pm \mu_2), \dots, \Phi(\nu, \pm \mu_N)]^T \quad (5.38)$$

e

$$\mathbf{\Pi}(l) = [P_l(\mu_1), P_l(\mu_2), \dots, P_l(\mu_N)]^T \quad (5.39)$$

e as matrizes  $N \times N$

$$\mathbf{M} = \text{diag}\{\mu_1, \mu_2, \dots, \mu_N\} \quad (5.40)$$

e

$$\mathbf{W} = \text{diag}(w_1, w_2, \dots, w_N), \quad (5.41)$$

é possível reescrever as equações dadas por (5.37) na forma de duas equações matriciais, que podem ser expressas por [6]

$$(\mathbf{I}_m - \frac{1}{\nu} \mathbf{M}) \Phi_+(\nu) = \frac{c}{2} \sum_{l=0}^L \beta_l \mathbf{\Pi}(l) \mathbf{\Pi}^T(l) \mathbf{W} [\Phi_+(\nu) + (-1)^l \Phi_-(\nu)] \quad (5.42)$$

e

$$(\mathbf{I}_m + \frac{1}{\nu}\mathbf{M})\Phi_-(\nu) = \frac{c}{2} \sum_{l=0}^L \beta_l \mathbf{\Pi}(l) \mathbf{\Pi}^T(l) \mathbf{W} [\Phi_-(\nu) + (-1)^l \Phi_+(\nu)], \quad (5.43)$$

onde  $\mathbf{I}_m$  representa a matriz identidade de ordem  $m$ .

Por questão de simplicidade de notação, defina-se, também a matriz  $m \times m$  dada por

$$\Upsilon(l) = \beta_l \mathbf{\Pi}(l) \mathbf{\Pi}^T(l) \mathbf{W} \quad (5.44)$$

de modo que as equações (5.42) e (5.43) podem ser reescritas como

$$(\mathbf{I}_m \mp \frac{1}{\nu}\mathbf{M})\Phi_{\pm}(\nu) = \frac{c}{2} \sum_{l=0}^L \Upsilon(l) [\Phi_{\pm}(\nu) + (-1)^l \Phi_{\mp}(\nu)]. \quad (5.45)$$

Somando as duas equações anteriormente apresentadas, tem-se

$$\begin{aligned} (\mathbf{I}_m - \frac{1}{\nu}\mathbf{M})\Phi_+(\nu) + (\mathbf{I}_m + \frac{1}{\nu}\mathbf{M})\Phi_-(\nu) &= \\ &= \frac{c}{2} \sum_{l=0}^L \Upsilon(l) [\Phi_+(\nu) + (-1)^l \Phi_-(\nu) + \Phi_-(\nu) + (-1)^l \Phi_+(\nu)] \end{aligned} \quad (5.46)$$

ou, ainda

$$[\Phi_+(\nu) + \Phi_-(\nu)] - \frac{1}{\nu}\mathbf{M}[\Phi_+ - \Phi_-(\nu)] = \frac{c}{2} \sum_{l=0}^L [1 + (-1)^l] \Upsilon(l) [\Phi_+(\nu) + \Phi_-(\nu)]. \quad (5.47)$$

Definindo, então, os vetores

$$\mathbf{U} = \Phi_+(\nu) + \Phi_-(\nu) \quad (5.48)$$

e

$$\mathbf{V} = \Phi_+(\nu) - \Phi_-(\nu), \quad (5.49)$$

a expressão pode ser ainda reescrita como

$$\mathbf{U} - \frac{1}{\nu}\mathbf{M}\mathbf{V} = \frac{c}{2} \sum_{l=0}^L [1 + (-1)^l] \Upsilon(l) \mathbf{U} \quad (5.50)$$

donde se tem

$$[\mathbf{I}_m - \frac{c}{2} \sum_{l=0}^L [1 + (-1)^l] \Upsilon(l)] \mathbf{M}^{-1} \mathbf{M} \mathbf{U} = \frac{1}{\nu} \mathbf{M} \mathbf{V}. \quad (5.51)$$

De modo análogo, subtraindo as duas equações dadas por (5.42) e (5.43), obtém-se a expressão [6]

$$[\mathbf{I}_m - \frac{c}{2} \sum_{l=0}^L [1 - (-1)^l] \Upsilon(l)] \mathbf{M}^{-1} \mathbf{M} \mathbf{V} = \frac{1}{\nu} \mathbf{M} \mathbf{U}. \quad (5.52)$$

Defina-se, então, as matrizes

$$\mathbf{E} = \mathbf{I}_m - \frac{c}{2} \sum_{l=0}^L [1 + (-1)^l] \Upsilon(l) \mathbf{M}^{-1} \quad (5.53)$$

$$\mathbf{B} = \mathbf{I}_m - \frac{c}{2} \sum_{l=0}^L [1 - (-1)^l] \Upsilon(l) \mathbf{M}^{-1} \quad (5.54)$$

e os vetores

$$\mathbf{X} = \mathbf{M} \mathbf{U} \quad (5.55)$$

$$\mathbf{Y} = \mathbf{M} \mathbf{V}, \quad (5.56)$$

é possível reescrever (5.51) e (5.52) como [6]

$$\mathbf{E} \mathbf{X} = \frac{1}{\nu} \mathbf{Y} \quad (5.57)$$

e

$$\mathbf{B} \mathbf{Y} = \frac{1}{\nu} \mathbf{X} \quad (5.58)$$

que podem ser combinadas de forma a obter o problema de autovalores [6, 7]

$$(\mathbf{B} \mathbf{E}) \mathbf{X} = \frac{1}{\nu^2} \mathbf{X} \quad (5.59)$$

de onde se pode obter as constantes de separação  $\nu_j$  usadas na suposição de solução dada pela equação(5.34), que sempre aparecem em pares  $\pm \nu_i$  [6].

Escrevendo então

$$\nu \mathbf{E} \mathbf{X} = \mathbf{Y}$$

e somando  $\mathbf{X}$  nos dois termos dessa expressão, tem-se

$$(\mathbf{I}_m + \nu \mathbf{E})\mathbf{X} = \mathbf{X} + \mathbf{Y} = \mathbf{M}(\mathbf{U} + \mathbf{V}) \quad (5.60)$$

ou, ainda,

$$(\mathbf{I}_m + \nu_j \mathbf{E})\mathbf{X}(\nu_j) = 2\mathbf{M}\Phi_+(\nu_j), \quad (5.61)$$

de onde se obtém [6]

$$\Phi_+(\nu_j) = \frac{1}{2}\mathbf{M}^{-1}(\mathbf{I}_m + \nu_j \mathbf{E})\mathbf{X}(\nu_j). \quad (5.62)$$

De modo análogo, subtraindo-se  $X$  dos dois lados da expressão, é possível obter [6]

$$\Phi_-(\nu_j) = \frac{1}{2}\mathbf{M}^{-1}(\mathbf{I}_m - \nu_j \mathbf{E})\mathbf{X}(\nu_j). \quad (5.63)$$

Assim, com as constantes de separação dadas pelo problema de autovalores (5.59) e as soluções elementares de dadas pelas equações (5.62) e (5.63), pode-se escrever as soluções por ordenadas discretas do problema dado, na forma vetorial [6], como

$$\Psi_{\pm}(x) = [\Psi(x, \pm\mu_1), \Psi(x, \pm\mu_2), \dots, \Psi(x, \pm\mu_N)]^T \quad (5.64)$$

na forma

$$\Psi_{\pm}(x) = \sum_{j=1}^N [a_j \Phi_{\pm}(\nu_j) e^{-x/\nu_j} + b_j \Phi_{\mp}(\nu_j) e^{-(x_0-x)/\nu_j}], \quad (5.65)$$

onde  $\{a_j\}$  e  $\{b_j\}$  são coeficientes a determinar e a segunda exponencial é escrita dessa maneira a fim de evitar problemas de *overflow* durante o cálculo computacional [6,7], o que representa outra vantagem do método *ADO*.

A fim de determinar os coeficientes  $\{a_j\}$  e  $\{b_j\}$ , substitui-se (5.65) nas condições iniciais  $\Psi(0, \mu_i) = \mathbf{F}_1(\mu_i)$  e  $\Psi(x_0, -\mu_i) = \mathbf{F}_2(\mu_i)$ , com  $i = 1, 2, \dots, m$ , obtendo o sistema linear de tamanho  $2m$  dado pelas equações

$$\sum_{j=1}^N [a_j \Phi_+(\nu_j) + b_j \Phi_-(\nu_j) e^{-x_0/\nu_j}] = \mathbf{F}_1(\mu_i) \quad (5.66)$$

$$\sum_{j=1}^N [a_j \Phi_-(\nu_j) e^{-x_0/\nu_j} + b_j \Phi_+(\nu_j)] = \mathbf{F}_2(\mu_i). \quad (5.67)$$

A partir dos passos anteriores, é possível determinar a solução da equação de transporte pelo método das ordenadas discretas, mas mantendo essa solução analítica em termos da variável espacial. Este tipo de solução costuma ser chamada de solução por ordenadas discretas em forma fechada para a variável espacial da equação de transporte considerada.

Considere-se, então, que a intensidade em todas as direções não está disponível, mas que se tem a informação relativa à densidade radiativa em alguns pontos no intervalo  $x \in [0, x_0]$ . A **densidade radiativa** é definida em termos das intensidades de radiação em todas as direções. Considerando o esquema de quadratura escolhido [6], pode-se escrever

$$\rho(x) = \int_0^1 [\Psi(x, +\mu) + \Psi(x, -\mu)] d\mu, \quad (5.68)$$

que, em termos da solução em ordenadas discretas, é dada [6, 25] por

$$\rho(x) = \sum_{j=1}^N [a_j e^{-x/\nu_j} + b_j e^{-(a-x)/\nu_j}] \Phi_0(\nu_j) \quad (5.69)$$

onde

$$\Phi_0(\nu_j) = \sum_{i=1}^N w_i [\Phi(\nu_j, \mu_i) + \Phi(\nu_j, -\mu_i)]. \quad (5.70)$$

### 5.3 Resultados Numéricos

Discute-se, nessa seção, a eficiência de métodos numéricos na identificação do parâmetro de albedo e detecção do grau de anisotropia em um problema inverso de transporte de radiação. Esses parâmetros identificam o meio e o problema inverso associado pode ser relacionado com reconstruções, por exemplo, em problemas de tomografia ótica [3, 5].



A metodologia consiste na solução de problema direto usando o método *ADO* e posterior identificação dos parâmetros fixados com técnicas inversas. Usa-se como conjunto de dados experimentais a informação relativa à intensidade de radiação ( $\rho(x)$ ) em diferentes pontos de uma barra de comprimento  $x_0$ . Pretende-se concluir a respeito da qualidade das estimativas para o parâmetro de albedo ( $c$ ) e grau de anisotropia ( $L$ ) após a introdução de perturbação gaussiana aos dados experimentais, estes que serão gerados de forma sintética.

### 5.3.1 Geração de dados sintéticos a partir implementação do problema direto

A solução do problema apresentado neste capítulo foi derivada segundo o método das Ordenadas Discretas Analíticas, apresentado anteriormente. A fim de verificar a correta implementação do método, considera-se inicialmente o problema teste de transporte unidimensional homogêneo com simetria azimutal e sem dependência energética, dado pela equação

$$\mu \frac{\partial}{\partial x} \Psi(x, \mu) + \Psi(x, \mu) = \frac{c}{2} \sum_{l=0}^L \beta_l P_l(\mu) \int_{-1}^1 P_l(\mu') \Psi(x, \mu') d\mu' \quad (5.71)$$

com grau de anisotropia  $L = 6$ , constante de albedo  $c = 0.99$ , e quadratura de ordem  $m = 10$ . Além disso, considera-se que as condições de contorno sejam conhecidas e dadas por

$$\Psi(0, \mu) = 1 \quad (5.72)$$

e

$$\Psi(x_0, -\mu) = 0 \quad (5.73)$$

e que o comprimento da barra seja dado por  $x_0 = 1$ . Compara-se os resultados da densidade de radiação com aqueles apresentados na literatura [6].

Tabela 5.1: Comparação dos resultados numéricos com os apresentados na literatura

$x$	literatura	implementação
<b>0.0</b>	1,29320	1,29320
<b>0.1</b>	1,20038	1,20037
<b>0.2</b>	1,13796	1,13796
<b>0.3</b>	1,08352	1,08352
<b>0.4</b>	1,03279	1,03279
<b>0.5</b>	0,983728	0,983728
<b>0.6</b>	0,934947	0,934947
<b>0.7</b>	0,885085	0,885085
<b>0.8</b>	0,832215	0,832215
<b>0.9</b>	0,772365	0,772365
<b>1.0</b>	0,684710	0,684710

Percebe-se, na Tabela 5.1, que os resultados encontrados via implementação estão de acordo com aqueles disponíveis na literatura, o que permite concluir que a mesma pode ser utilizada como solução do problema direto considerado neste trabalho.

Nesse sentido, passa-se a considerar um problema alternativo, que será utilizado como problema teste neste trabalho. Neste, o grau de anisotropia é dado por  $L = 8$  e o albedo por  $c = 0.8$ . Mantém-se as condições de contorno idênticas ao problema anterior e comprimento da barra  $x_0 = 2$ . A densidade de nêutrons ( $\rho$ ) é tomada para os pontos

$$\mathbf{x} = [0, 1 \quad 0, 5 \quad 0, 9 \quad 1, 3 \quad 1, 7]^T. \quad (5.74)$$

Além disso, a variável angular é discretizada com  $m = 10$ . A saída do modelo é apresentada na Figura 5.1.

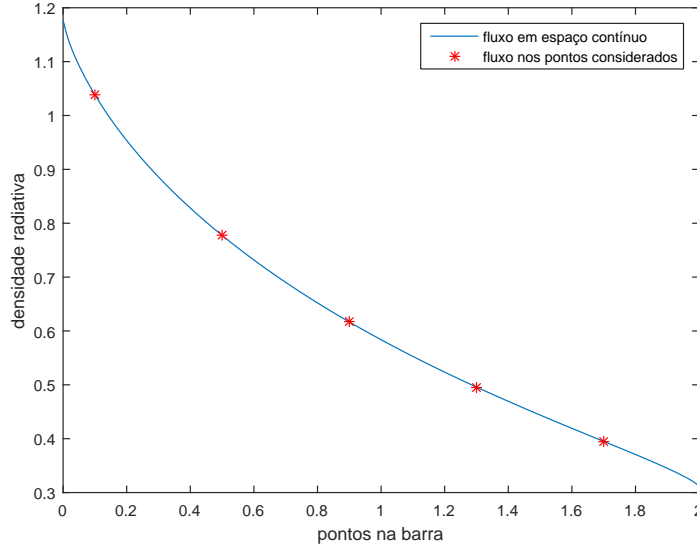


Figura 5.1: Gráfico da densidade radiativa  $\rho(x)$  sobre a barra

Este problema será utilizado na geração dos primeiros resultados numéricos a respeito da identificação dos parâmetros de albedo e grau de anisotropia. Mais tarde, outros problemas serão considerados a fim de verificar os resultados obtidos.

Considere a saída da simulação numérica realizada, denotada por  $\rho_0(x)$ , como a solução exata do problema direto associado. Sejam os parâmetros a identificar dados por

$$\mathbf{P} = \begin{bmatrix} c & L \end{bmatrix}^T. \quad (5.75)$$

com dados experimentais sintéticos gerados a partir de uma perturbação gaussiana sobre a saída do modelo com os parâmetros definidos nos problemas, com média zero e desvio padrão dado por uma porcentagem ( $i$ ) do maior valor do fluxo de radiação em cada caso. Deseja-se, determinar o conjunto  $\mathbf{P}$  de parâmetros que minimize a função objetivo dada por

$$\mathcal{S}(\mathbf{P}) = [\hat{\rho}(\mathbf{x}) - \rho(\mathbf{x}, \mathbf{P})]^T [\hat{\rho}(\mathbf{x}) - \rho(\mathbf{x}, \mathbf{P})], \quad (5.76)$$

onde  $\rho(\mathbf{x}, \mathbf{P})$  representa a densidade de nêutrons calculada para o vetor de parâmetros  $\mathbf{P}$  e nos pontos  $\mathbf{x}$  dados anteriormente.

### 5.3.2 Análise do comportamento da função objetivo em relação ao albedo

Antes de proceder com a identificação dos parâmetros a partir de métodos numéricos, uma breve análise da confiabilidade esperada para os resultados, a partir da análise dos gráficos da função objetivo, como problema direto, para diferentes valores de albedo, foi estabelecida.

Considerando, inicialmente, que o grau de anisotropia do problema ( $L = 8$ ) seja conhecido, procede-se com a análise do caso em que não há perturbação nos dados. É natural esperar que o mínimo exista e seja nulo exatamente no valor do albedo, dado no problema teste. Na Figura 5.2, que representa o gráfico da função objetivo para  $c$  contínuo, pode-se verificar que o comportamento está de acordo com o esperado, ou seja, que o mínimo ocorre em  $c = 0,8$  e que, nesse ponto, o resíduo é nulo.

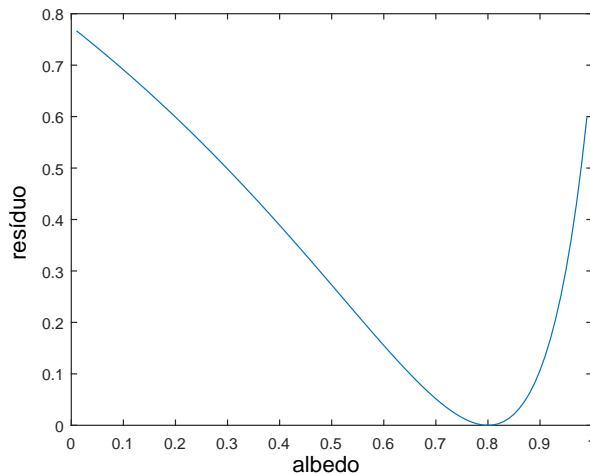


Figura 5.2: Gráfico da função objetivo  $\mathcal{S}(\mathbf{x}, \mathbf{P})$  para  $L$  conhecido, dados não perturbados e  $c$  contínuo

Pode-se notar, além disso, a suavidade da curva que caracteriza a função objetivo, fator positivo no processo de identificação de parâmetros quando se utiliza processos iterativos de refinamento, como neste trabalho.

Em sequência, consideram-se perturbações da ordem de 1% e 10% com distribuição gaussiana aplicada sobre a saída do problema direto. Sabe-se que o mínimo obtido não será necessariamente nulo, mas deseja-se observar se, nesses casos, o parâmetro associado ao menor resíduo se mantém ou se há a tendência de algum desvio do valor correto.

Com base na Tabela 5.2, conclui-se que, no caso de de perturbações de baixa magnitude, é possível identificar o albedo obtendo resíduo na ordem de  $10^{-5}$ . No caso de perturbações de maior magnitude, fica clara a tendência de um desvio em relação ao valor correto. Em ambos os casos, contudo, o erro relativo, conforme definido em (4.33), é menor que o nível de perturbação agregado aos dados.

Tabela 5.2: Valores de  $c$  que minimizam  $S(x, \mathbf{P})$  quando o grau de anisotropia é conhecido

<b>grau da perturbação</b>	<b>c</b>	<b>erro relativo</b>	<b>resíduo</b>
1%	0,80167	0,21%	$2,025 \cdot 10^{-5}$
10%	0,73685	7,8932%	$2,307 \cdot 10^{-2}$

A Figura 5.3 apresenta o gráfico que relaciona os diferentes valores de  $c$  ao resíduo encontrado na identificação do albedo, que permite verificar o comportamento observado na Tabela 5.2.

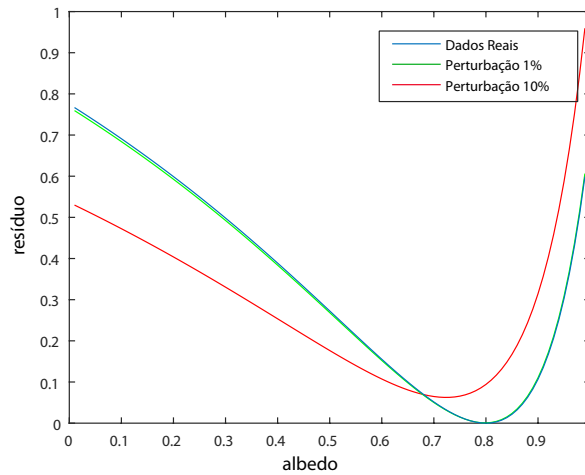


Figura 5.3: Gráficos da função objetivo  $\mathcal{S}(x, \mathbf{P})$  para grau de anisotropia conhecido e ruídos da ordem de 0% (azul), 1% (verde) e 10% (vermelho) na saída do modelo

### 5.3.3 Resultados numéricos para a identificação do albedo

A identificação do parâmetro é efetuada, neste trabalho, a partir do método de Levenberg-Marquardt, fixando-se como critérios de parada, o máximo de 15 iterações ou a obtenção de resíduo de magnitude inferior a  $10^{-15}$ . A estimativa inicial é dada por valores aleatórios dentro do intervalo  $c \in (0, 1)$ . Ao efetuar os testes iniciais, sem perturbação nos dados, foi possível obter o valor exato para o parâmetro de albedo,  $c = 0,8$ , a partir de estimativas iniciais em, na média, cinco iterações.

Ao introduzir perturbações da ordem de 1% na saída do modelo e, assim como realizado no caso do biorreator, estimar o parâmetro de albedo um total de cem vezes, foi possível encontrar uma distribuição aproximadamente normal dos resultados, com média muito próxima do valor real e baixa variância. Um histograma desses resultados pode ser observado na Figura 5.4.

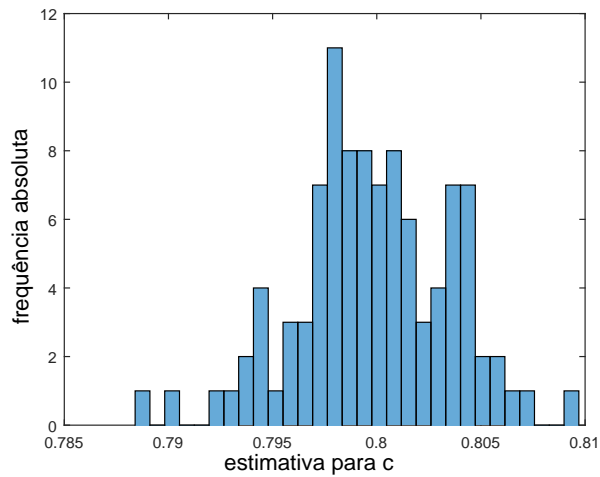


Figura 5.4: Histograma dos valores obtidos para o parâmetro de albedo a partir de dados perturbados em 1%

Levando em consideração que o ruído agregado é da ordem de 1% e que o valor encontrado apresenta erro relativo, em média, inferior a 0,1% ao parâmetro real, é possível concluir que esse tipo de abordagem se mostra eficiente para a identificação do parâmetro de albedo quando se puder garantir que os erros nos dados sejam suficientemente pequenos, conforme se observa na Tabela 5.3.

Tabela 5.3: Resultados médios para a identificação da constante de albedo com dados perturbados em 1%

	<b>média</b>	<b>variância</b>	<b>erro relativo</b>
<b>c</b>	0,79983	$1,35047 \cdot 10^{-5}$	0,021446%

Além disso, ao tomar intervalos de confiança de 95% em relação à média dos resultados é possível, ainda, verificar que as estimativas encontradas se distribuem de acordo com o esperado na teoria. Pode-se observar, na Figura 5.5, o conjunto de estimativas (em vermelho), bem como o valor real do parâmetro (em verde) e a média e limites do intervalo de confiança (em azul).

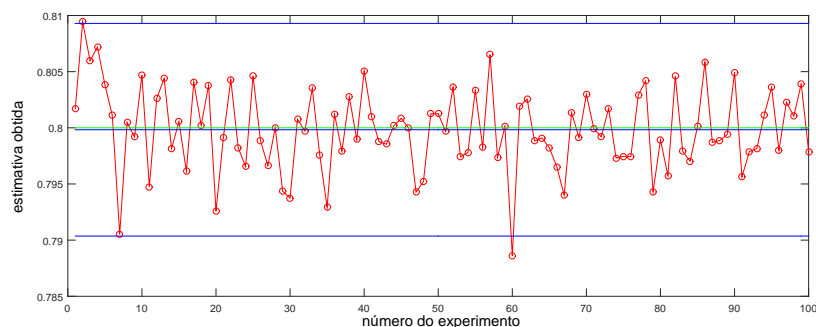


Figura 5.5: Intervalos de confiança para as estimativas na determinação do parâmetro de albedo a partir de dados perturbados em 1%

Repetindo o experimento para perturbações maiores, da ordem de 10%, é possível verificar, no histograma apresentado na Figura 5.6, que a variância dos resultados é ampliada. Além disso, diversas estimativas se distanciam do valor real do parâmetro para o problema considerado.

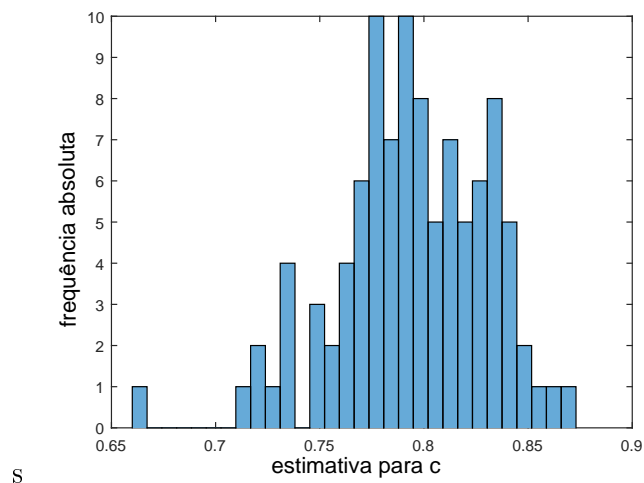


Figura 5.6: Histograma dos valores obtidos para o parâmetro de albedo a partir de dados perturbados em 10%



Em média, contudo, o método consegue estimar o parâmetro de albedo de forma adequada, com erro relativo médio inferior a 1%, como se verifica na Tabela 5.4.

Tabela 5.4: Resultados médios para a identificação da constante de albedo com dados perturbados em 10%

	<b>média</b>	<b>variância</b>	<b>erro relativo</b>
<b>c</b>	0,79417	0,00128	0,7281%

Além disso, é possível verificar, comparando-se com o resultado anterior, que a variância está relacionada ao quadrado do nível de perturbação agregada às saídas, conforme esperado. Ao construir o intervalo de confiança de 95% em relação ao valor esperado, verifica-se, na Figura 5.7, novamente, que as estimativas estão de acordo com a teoria.

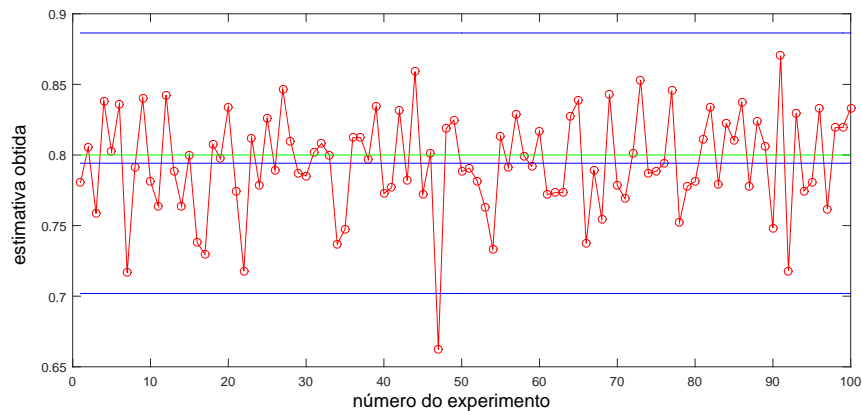


Figura 5.7: Intervalos de confiança para as estimativas na determinação do parâmetro de albedo a partir de dados perturbados em 10%

### *5.3.3.1 Análise da média e variância quando da alteração do número de estimações consideradas*

A fim de complementar o estudo conduzido nessa seção, foram construídos intervalos de confiança para os resultados obtidos parcialmente, ou seja, para parte dos experimentos realizados. A ideia é verificar se o número de estimativas realizadas exerce grande influência sobre a qualidade da estimativa média obtida.

Ao proceder com a análise dos resultados parciais na inserção de perturbações da ordem de 1% nos dados de saída do modelo, pode-se verificar que, mesmo com um número inferior de estimativas, a teoria dos intervalos de confiança é satisfeita. Além disso, é possível perceber pela Figura 5.8 e Tabela 5.5, que a média dos resultados tende a se aproximar do valor real quanto mais estimativas são executadas, bem como a variância tem uma tendência de redução conforme se aumenta o número de testes, conforme esperado.

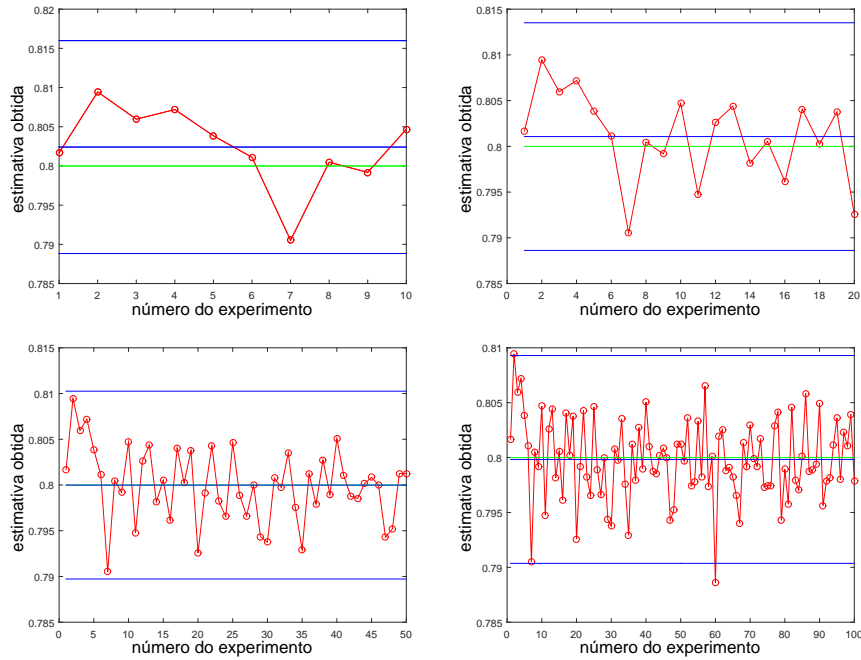


Figura 5.8: Intervalos de Confiança para perturbações da ordem de 1% e 10, 20, 50 e 100 estimativas

Tabela 5.5: Resultados médios com diferentes números de estimativas na determinação do albedo com perturbações da ordem de 1%

estimativas	média	variância	erro relativo
10	0,80241	$2,77 \cdot 10^{-5}$	0,30%
20	0,80107	$2,33 \cdot 10^{-5}$	0,134%
50	0,79999	$1,58 \cdot 10^{-5}$	0,001%
100	0,79983	$1,35 \cdot 10^{-5}$	0,02%

De modo análogo, no caso em que a perturbação agregada é da ordem de 10%, a teoria dos intervalos de confiança é satisfeita, mas a variância se mostra menos sensível ao aumento número de estimativas tomadas, conforme se observa na Figura 5.9 e Tabela 5.6.

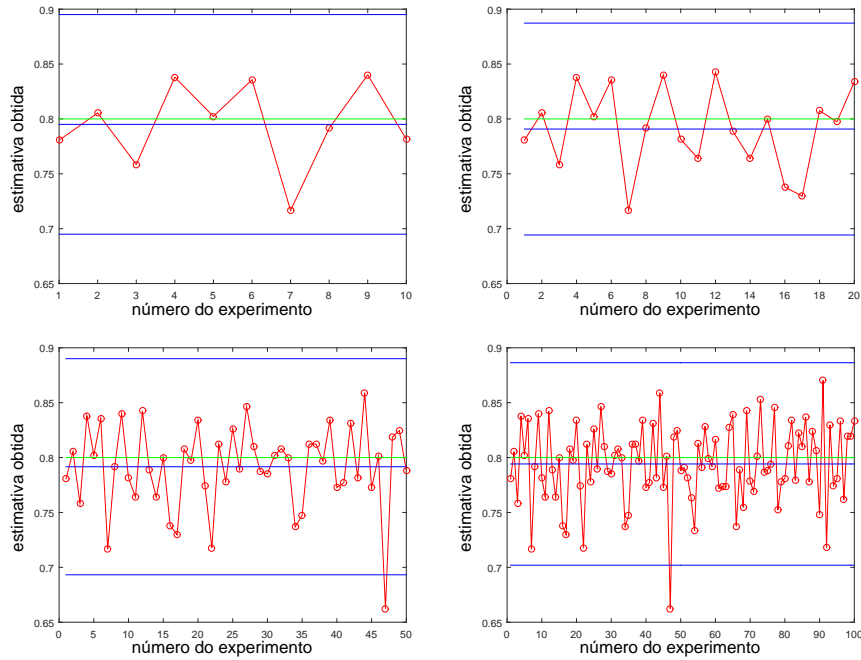


Figura 5.9: Intervalos de Confiança para perturbações da ordem de 10% e 10, 20, 50 e 100 estimativas

Tabela 5.6: Resultados médios com diferentes números de estimativas na determinação do albedo com perturbações da ordem de 10%

<b>estimativas</b>	<b>média</b>	<b>variância</b>	<b>erro relativo</b>
10	0,79502	$1,51 \cdot 10^{-3}$	0,620%
20	0,79077	$1,4 \cdot 10^{-3}$	1,154%
50	0,79163	$1,46 \cdot 10^{-3}$	1,046%
100	0,79417	$1,28 \cdot 10^{-3}$	0,729%

É possível concluir, portanto, que o número de experimentos não é determinante na identificação do parâmetro de albedo. Ou seja, boas estimativas podem ser obtidas mesmo sem a realização elevado número de experimentos.

### 5.3.4 Resultados numéricos para a identificação do albedo e grau de anisotropia

Nesta seção, investiga-se a possibilidade estimar, em conjunto, o parâmetro de albedo e o grau de anisotropia, a partir dos dados sintéticos gerados para o problema teste, usando-se o método de Levenberg-Marquardt. O grau de anisotropia, relacionado à discretização da função de probabilidade de distribuição de partículas, dá uma ideia do termo a partir do qual a série de polinômios de Legendre, dada por

$$f(\mu' \rightarrow \mu) = \frac{1}{2} \sum_{l=0}^L \beta_l P_l(\mu') P_l(\mu), \quad (5.77)$$

foi truncada, e explica como é dado o espalhamento da radiação no meio material.

Cabe ressaltar que o método de Levenberg-Marquardt foi concebido para a estimação de parâmetros contínuos, e não discretos. Além disso, as constantes  $\beta_l$ , na equação anterior, dependem (5.24) diretamente do valor inteiro desse parâmetro. Desse modo, a cada passo iterativo do método, toma-se os valores de  $L$  arredondados para o número inteiro mais próximo, de acordo com a convenção, como nova estimativa.

Após a introdução de perturbações da ordem de 1% na saída do modelo, e da realização de cem estimações, usando o método de Levenberg-Marquardt com os mesmos critérios de parada definidos anteriormente e chute inicial aleatório no intervalo  $c \in (0, 1)$ , é possível observar, na Figura 5.10, a seguinte distribuição de resultados.

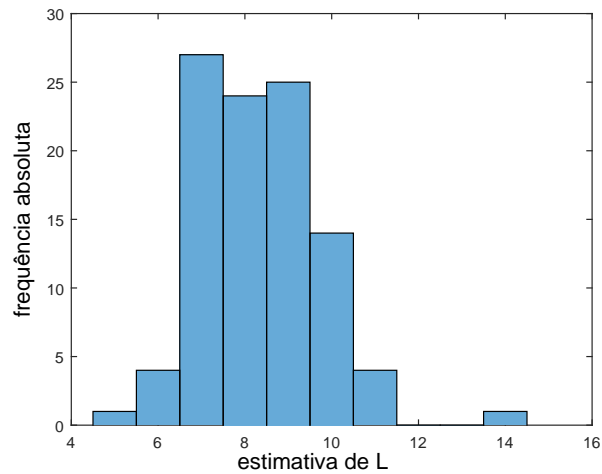


Figura 5.10: Histograma de estimativas para o grau de anisotropia na identificação conjunta com dados perturbados em 1%

Destaca-se que, embora a reconstrução nem sempre tenha conduzido ao valor real do grau de anisotropia ( $L=8$ ), a maior parte dos resultados se concentraram em uma vizinhança aceitável, já que não se faz necessário estimar exatamente  $L$ , mas sim determinar um valor para o qual o modelo discreto represente bem o modelo contínuo e os dados experimentais. O parâmetro de albedo, por sua vez, apresenta distribuição semelhante à encontrada quando estimado isoladamente. Os resultados se distribuíram conforme histograma apresentado na Figura 5.11.

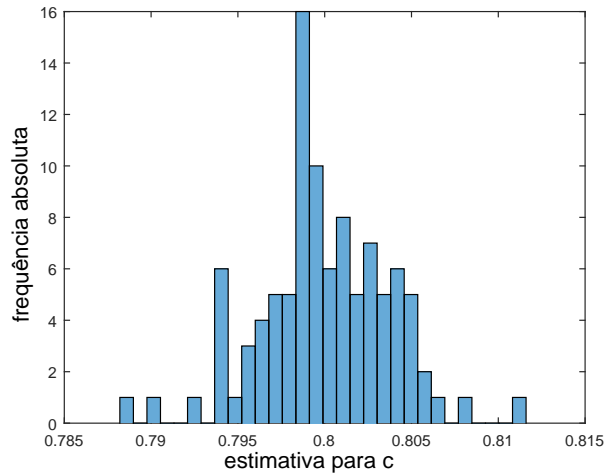


Figura 5.11: Histograma de estimativas para a constante de albedo na identificação conjunta com dados perturbados em 1%

Ao proceder com uma análise mais aprofundada, é possível perceber, na Tabela 5.7, que, em média, as estimativas de ambos os parâmetros são precisas, apresentando erros relativos baixos em relação aos valores reais dos parâmetros do problema.

Tabela 5.7: Resultados médios para a identificação da constante de albedo e grau de anisotropia com dados perturbados em 1%

	<b>média</b>	<b>variância</b>	<b>erro relativo</b>
<b>c</b>	0,79996	$1,41246 \cdot 10^{-5}$	0,04125%
<b>L</b>	$8,33 \approx 8$	$1,961 \approx 2$	4,12%

Na Figura 5.12, apresentada abaixo, é possível observar a comparação entre as saídas do problema direto computado com o conjunto correto de parâmetros do problema e com uma das estimativas encontradas com o método de Levenberg-Marquardt com dados experimentais perturbados em 1%, onde se verifica uma boa concordância entre as simulações.

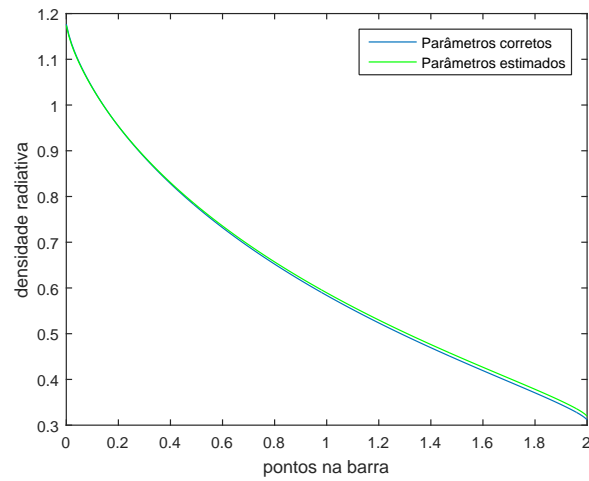


Figura 5.12: Comparação entre a saída do problema direto com os parâmetros corretos e com um dos conjuntos de parâmetros estimados no caso de uma perturbação de 1%

Ao proceder com a análise das estimativas no caso de uma saída perturbada em 5%, contudo, é possível perceber uma tendência de que o grau de anisotropia passe a não ser estimado de forma correta, como pode se observar no histograma presente na Figura 5.13.



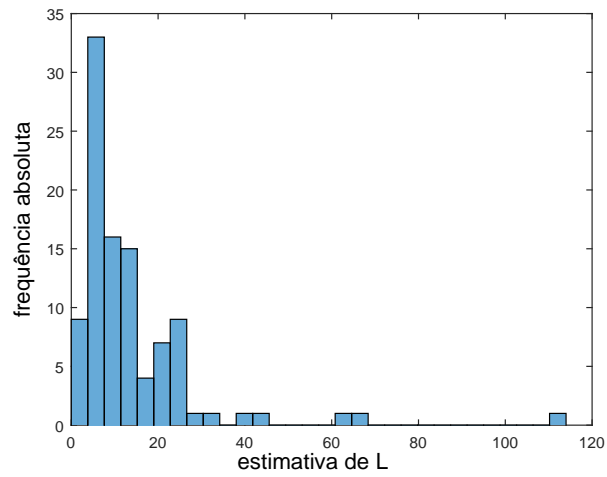


Figura 5.13: Histograma de estimativas para o grau de anisotropia na identificação conjunta com dados perturbados em 5%

A constante de albedo, contudo, apresenta uma boa distribuição de estimativas  $e$ , embora se verifique uma dispersão maior de resultados, verifica-se que, em média, tende a ser estimada de forma suficientemente precisa, com erro relativo reduzido em relação aos ruídos agregados, conforme se verifica na Figura 5.14.

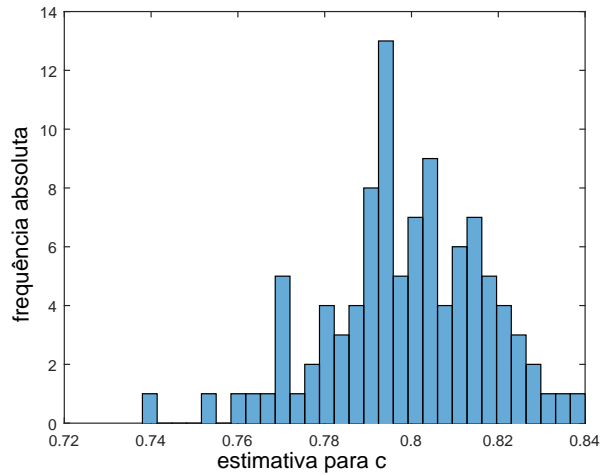


Figura 5.14: Histograma de estimativas para a constante de albedo na identificação conjunta com dados perturbados em 5%

Estes resultados podem ser confirmados na Tabela 5.8, apresentada abaixo.

Tabela 5.8: Resultados médios para a identificação da constante de albedo e grau de anisotropia com dados perturbados em 5%

	<b>média</b>	<b>variância</b>	<b>erro relativo</b>
<b>c</b>	0,79879	$3,317199 \cdot 10^{-4}$	0,15057%
<b>L</b>	$13,6 \approx 14$	$227,37 \approx 227$	70%

De modo análogo ao anterior, é possível observar, na Figura 5.15, a comparação entre as saídas do problema direto computado com o conjunto correto de parâmetros do problema e com uma das estimativas encontradas com o método de Levenberg-Marquardt com dados experimentais perturbados em 5%. Nesse caso, embora o resultado continue satisfatório, com uma boa concordância entre as simulações, é possível notar uma maior discrepância que no caso anterior, devido à estimação incorreta do grau de anisotropia.

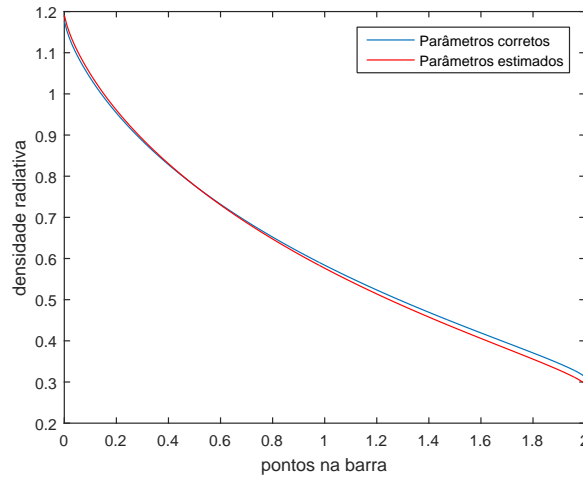


Figura 5.15: Comparação entre a saída do problema direto com os parâmetros corretos e com um dos conjuntos de parâmetros estimados no caso de uma perturbação de 5%

### 5.3.5 Resultados numéricos adicionais

Nesta subseção, resultados para casos teste adicionais são apresentados. Em um primeiro momento, considera-se o mesmo problema, investigando, no entanto, os efeitos da disponibilidade de diferentes quantidades de dados experimentais para a estimação do albedo e grau de anisotropia. Posteriormente, verifica-se se os resultados obtidos na identificação do albedo em casos limite na solução do problema direto, conforme implementada nesse trabalho.

Considere-se o problema inicial, com os fluxos conhecidos apenas sobre os pontos dados por

$$\mathbf{x} = [0,4 \quad 0,8 \quad 1,4]^T. \quad (5.78)$$

É possível obter, após a repetição de cinquenta experimentos e identificação dos parâmetros de albedo e grau de anisotropia, com o método de Levenberg-Marquardt, com os mesmos critérios definidos na seção anterior, os resultados médios apresentados na Tabela 5.9.

Tabela 5.9: Resultados médios para a identificação da constante de albedo e grau de anisotropia no primeiro problema, considerando apenas três pontos e perturbação da ordem de 1%

<b>perturbação</b>	<b>média</b>	<b>variância</b>	<b>erro relativo</b>
<b>c</b>	0,8005	$1,9 \cdot 10^{-5}$	0,06%
<b>L</b>	8,5951	4,7882	7,44%

Pode-se perceber que, no caso de perturbações da ordem de 1%, foi possível identificar com suficiente precisão a constante de albedo. O grau de anisotropia, em média, também pôde ser encontrado, mas com maior variância nos resultados obtidos.

Tabela 5.10: Resultados médios para a identificação da constante de albedo e grau de anisotropia no primeiro problema, considerando apenas três pontos e perturbação da ordem de 5%

<b>perturbação</b>	<b>média</b>	<b>variância</b>	<b>erro relativo</b>
<b>c</b>	0,7807	$2,1 \cdot 10^{-2}$	2,41%
<b>L</b>	132,36	$9,96 \cdot 10^4$	> 1000%

Quando perturbações maiores, da ordem de 5%, são adicionadas à saída, verifica-se, na Tabela 5.10, a impossibilidade de identificar, corretamente, ambos os parâmetros. Em testes anteriores, nos quais mais informação estava disponível, o

grau de anisotropia tampouco havia sido estimado corretamente, mas a constante de albedo havia sido identificada.

Pode-se concluir, portanto, que quando níveis maiores de ruído estão presentes nos dados experimentais e há menos informação disponível, a estimação dos parâmetros considerados nesta tarefa é comprometida.

Em seguida, deseja-se investigar se uma maior quantidade de informação a respeito da saída do sistema tende a melhorar as estimativas encontradas. Considerando o problema inicial, com os fluxos conhecidos sobre os pontos dados por

$$\mathbf{x} = [0,1 \quad 0,3 \quad 0,5 \quad 0,7 \quad 0,9 \quad 1,1 \quad 1,3 \quad 1,5 \quad 1,7 \quad 1,9]^T, \quad (5.79)$$

e procedendo com cinquenta estimações pelo método de Levenberg-Marquardt, é possível encontrar os resultados médios apresentados na Tabela 5.11.

Tabela 5.11: Resultados médios para a identificação da constante de albedo e grau de anisotropia no primeiro problema, considerando nove pontos e perturbação da ordem de 1%

<b>perturbação</b>	<b>média</b>	<b>variância</b>	<b>erro relativo</b>
<b>c</b>	0,7995	$5,82 \cdot 10^{-6}$	0,06%
<b>L</b>	8,2379	$6,94 \cdot 10^{-1}$	2,97%

Percebe-se que, nesse caso, ambos os parâmetros podem ser identificados, em média, de forma correta, ou seja, com baixo erro relativo. Vale notar, também, que a variância entre as diferentes estimações é baixa.

Comparando-se com os resultados anteriormente obtidos, contudo, não é possível afirmar que o maior número de informações disponíveis *a priori* resulte em melhor estimação dos parâmetros considerados.

Tabela 5.12: Resultados médios para a identificação da constante de albedo e grau de anisotropia no primeiro problema, considerando nove pontos e perturbação da ordem de 5%

<b>perturbação</b>	<b>média</b>	<b>variância</b>	<b>erro relativo</b>
<b>c</b>	0,7936	$1,1 \cdot 10^{-3}$	0,8%
<b>L</b>	16,0059	819,15	100%

No caso em que perturbações da magnitude de 5% são adicionadas aos dados experimentais, é possível verificar, na Tabela 5.12, que, mesmo com mais informações disponíveis, o grau de anisotropia continua não sendo identificado de forma correta, apresentando erro relativo de 100% em relação ao valor no problema original ( $L = 8$ ) e grande variância nos resultados obtidos. Conclui-se, novamente, que o método utilizado não é eficaz na determinação do grau de anisotropia quando perturbações dessa magnitude estão presentes nos dados experimentais.

Por fim, levando em consideração que neste trabalho estamos considerando apenas os casos em que  $c < 1$ , verifica-se, a seguir, a possibilidade de identificar o albedo em caso limite, ou seja, quando  $c \approx 1$ .

Para tal, considera-se, como segundo problema, aquele em que o albedo é dado por  $c = 0,99$ . Todos os demais parâmetros são mantidos em relação ao problema anterior e as medidas experimentais continuam sendo tomadas sobre os pontos

$$\mathbf{x} = \left[ 0,1 \quad 0,3 \quad 0,5 \quad 0,7 \quad 0,9 \quad 1,1 \quad 1,3 \quad 1,5 \quad 1,7 \quad 1,9 \right]^T. \quad (5.80)$$

Após a adição de perturbações da ordem de 1% e 10% nos dados e posterior execução do algoritmo num total de cinquenta vezes, pode-se perceber que o parâmetro de albedo pôde ser estimado no caso em que a perturbação nos dados era da ordem de 1%. No segundo caso, embora o parâmetro estimado não seja exato, o erro em

relação ao valor real é da ordem de 0,25%, ou seja, inferior à perturbação agregada nos dados experimentais. Estes resultados podem ser observados na Tabela 5.13, apresentada abaixo.

Tabela 5.13: Resultados médios para a identificação da constante de albedo no segundo problema considerado

<b>perturbação</b>	<b>média</b>	<b>variância</b>	<b>erro relativo</b>
1%	0,9906	$2,24 \cdot 10^{-6}$	0,06%
10%	0,9925	$1,27 \cdot 10^{-4}$	0,25%

Considera-se, ainda, um terceiro problema teste, onde o albedo tem valor dado por  $c = 0,05$ . Verifica-se, a possibilidade de identificação do albedo no caso em que  $c \approx 0$ . Após a adição de perturbações da ordem de 1% e 10% nos dados e da repetição de cinquenta estimações, os resultados presentes na Tabela 5.14 foram obtidos, na identificação.

Tabela 5.14: Resultados médios para a identificação da constante de albedo no terceiro problema considerado

<b>perturbação</b>	<b>média</b>	<b>variância</b>	<b>erro relativo</b>
1%	0,0474	$9,66 \cdot 10^{-5}$	5,19%
10%	0,0470	$1,45 \cdot 10^{-2}$	5,93%

No caso do terceiro problema considerado, é possível perceber que o albedo não pôde ser estimado com boa precisão mesmo no caso em que perturbações da ordem de 1% foram adicionadas sobre os dados experimentais. Em casos limites nos quais  $c \approx 0$ , portanto, pode-se concluir que o método de Levenberg-Marquardt não produz resultados tão eficientes na identificação do albedo.

Nesta última seção, os resultados encontrados para o problema teste considerado anteriormente foram confirmados nos casos em que menos e mais dados experimentais estavam disponíveis. Além disso, verificou-se a possibilidade de estimar o albedo em problemas limite.

Em suma, foi possível concluir que o albedo é um parâmetro identificável por Levenberg-Marquardt mesmo quando ruídos de maior magnitude estão presentes nos dados, exceto quando  $c \approx 0$ . O grau de anisotropia, por sua vez, se mostra mais sensível às perturbações nos dados, deixando de ser identificável quando ruídos de maior magnitude estão presentes.



## 6 CONSIDERAÇÕES FINAIS

Neste trabalho, problemas inversos foram estudados através da identificação de parâmetros em biorreatores e problemas unidimensionais de transporte de partículas. No caso dos biorreatores anaeróbicos, os problemas diretos foram tratados com a utilização de esquemas numéricos. No caso de transporte de partículas, uma solução em forma fechada foi determinada, usando o método *ADO*. Os problemas inversos de identificação se basearam no critério dos mínimos quadrados, a partir da implementação de métodos iterativos determinísticos, dos quais se destaca o método de Levenberg-Marquardt. Foi possível concluir que os parâmetros do biorreator são identificáveis apenas quando ruídos de baixa magnitude estão presentes nos dados. Observou-se, contudo, que mesmo os parâmetros incorretos, encontrados nos casos em que maiores perturbações estão presentes nos dados, podem ser utilizados para a simulação numérica da saída de gás metano do biorreator. No caso analisado em transferência radiativa, o parâmetro de albedo pôde ser identificado, usando Levenberg-Marquardt, mesmo quando ruídos de maior magnitude foram introduzidos nos dados e foi possível perceber que a qualidade da estimativa encontrada não depende diretamente do número de experimentos realizados. O grau de anisotropia, por sua vez, se mostrou mais suscetível a erros experimentais e, embora identificável quando os ruídos são de baixa magnitude, não pôde ser estimado de forma correta quando perturbações maiores foram adicionadas nos dados de saída do modelo considerado.

Este estudo apresentou, contudo, apenas resultados preliminares, que devem ser investigados com mais atenção na continuidade dessa pesquisa. Nesse sentido, destacam-se alguns tópicos que fazem parte das perspectivas de estudo:

1. A identificação dos parâmetros do biorreator é obtida, na literatura, sempre a partir dos dados de saída de gás metano do sistema. Outros

dados experimentais, contudo, como a saída de gás carbônico, estão disponíveis. Levando em consideração que mais informação possibilita a conduzir a melhores estimativas, pretende-se revisitar a identificação dos parâmetros com métodos que permitam a estimação dos mesmos simultaneamente para ambos os conjuntos de dados.

2. No caso do transporte de partículas, mais resultados devem ser analisados a fim de confirmar as conclusões obtidas neste trabalho. Além disso, deseja-se verificar a aplicabilidade dos métodos citados na identificação das condições de contorno e fontes externas de radiação.

## Referências Bibliográficas

- [1] ANTONELLI, R.; HARMAND, J.; STEYER, J. P.; ASTOLFI, A. *Set-point regulation of an anaerobic digestion process with bounded output feedback*. Transactions on Control Systems Technology, v. 11, n. 4, 2003. pp. 495-504.
- [2] ARIAS, A. C.; BANKS, H. T.; CEPALDI, A.; LLOYD, A. L. *A Sensitivity Matrix Based Methodology for Inverse Problem Formulation*. North Carolina State University, Raleigh, 2009. Disponível em: <<http://www.ncsu.edu/crsc/reports/ftp/pdf/crsc-tr09-09.pdf>>. Acesso em: 20 nov. 2015.
- [3] ARRIDGE, S.R.; SCHOTLAND, J.C. *Optical tomography: forward and inverse problems*. arxiv, v. 1, University of Pennsylvania, Philadelphia, 2009. Disponível em: <<https://arxiv.org/pdf/0907.2586.pdf>>. Acesso em 27 out. 2016.
- [4] ASTER, R.; BORCHERS, B.; THURBER, C. *Parameter Estimation and Inverse Problems*. Elsevier Academic Press, Amsterdam, 2 ed., 2013. 376 p.
- [5] BAL, G.. *Introduction to Inverse Problems*. Columbia University, New York, 2012. 199 p. Disponível em: <<http://www.columbia.edu/~gb2030/PAPERS/IntroductionInverseProblems.pdf>>. Acesso em 27 out 2016.
- [6] BARICHELLO, L. B. *Thermal Measurements and Inverse Techniques. Capítulo 15: Explicit formulations for radiative transfer problems*. CRC Press, 1 ed., 2011. pp. 541-562.

- [7] BARICHELLO, L. B.; SIEWERT, C. *A new version of the discrete-ordinates method*. Proceedings of the 2nd International Conference on Computational Heat and Mass Transfer, COPPE, Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brasil, 2001.
- [8] BECK, J.; BLACKWELL, B.; ST. CLAIR JR, C. *Inverse Heat Conduction: Ill-posed Problems*. John Wiley & Sons Inc, New York, 1 ed., 1985. 308 p.
- [9] BERNARD, O.; HADJ-SADOK, Z.; DOCHAIN, D.; GENOVESI, A.; STEYER, J. *Dynamical Model Development and Parameter Identification for an Anaerobic Wastewater Treatment Process*. Biotechnology and Bioengineering, v. 75, n. 4, 2001. pp. 424–438.
- [10] BERTERO, M.; BOCCACCI, P. *Introduction to inverse problems in imaging*. Institute of Physics Publishing Ltd, Editora: CRC Press, Bristol, USA, 1 ed., 1998. 352 p.
- [11] CAMPESTRINI, L.; ECKHARD, D.; RUI, R.; BAZANELLA, A. S. *Identifiability Analysis and Prediction Error Identification of Anaerobic Batch Bioreactors*. Journal of Control, Automation and Electrical Systems, v. 25, n. 4, 2014. pp 438–447.
- [12] CAMPESTRINI, L.; ECKHARD, D.; KONRAD, O.; BAZANELLA, A. S. *Identificação não-linear de um biorreator através da minimização do erro de predição*. Congresso Brasileiro de Automática, XIX Congresso Brasileiro de Automática. Campina Grande, SBA, 2012. pp. 3066–3072.
- [13] CAMPOS VELHO, H. F. *Problemas Inversos: Conceitos Básicos e Aplicações*. INPE, Laboratório de Computação e Matemática Aplicada. IV Encontro de Modelagem Computacional, Nova Friburgo, RJ, 2008. 17 p.

- [14] CERCIGNANI, C. *The Boltzmann Equation and It's Application*. Springer Science & Business Media, v. 67, New York, 1988. 455 p.
- [15] CÉZARO, A. *Problemas Inversos: Uma Introdução*. SBM, Santa Maria, 2010. 165 p. Disponível em: <<http://www.sbm.org.br/docs/coloquios/SU-1.06.pdf>>. Acesso em: 12 set. 2015.
- [16] CHAHINE, M. T. *Inverse Problems in Radiative Transfer: Determination of Atmospheric Parameters*. Journal of the Atmospheric Sciences, v. 27, Jet Propulsion Laboratory, California Institute of Technology, Pasadena, 1970. pp. 960-967. Disponível em: <<http://web.gps.caltech.edu/~drf/misc/airs/i1520-0469-027-06-0960.pdf>>. Acesso em: 20 out 2016.
- [17] CHANDRASEKHAR, S. *Radiative Transfer*. Dover Books, Oxford University Press, London, 1 ed., 1950. 393 p.
- [18] COLAÇO, M.J.; ORLANDE, H.R.B.; DULIKRAVICH, G.S. *Inverse Problems, Design and Optimization*. v. 2, Editora: e-papers, Rio de Janeiro, 1 ed., 2005. 354 p.
- [19] DATTA, B.N. *Numerical Linear Algebra and Applications*. Northern Illinois University, SIAM, Brooks/Cole Publishing, 2 ed., 1995. 680p.
- [20] DOCHAIN, D. *Automatic Control of Bioprocesses Control Systems*. ISTE Ltd, Wiley, 1 ed., [S.l.], 2008. 243 p.
- [21] DORNELLES FILHO, A. A. *Cálculo Numérico*. Notas de Aula. Universidade de Caxias do Sul, Caxias do Sul, 11 ed., 2015.
- [22] ECKHARD, D.; CAMPESTRINI L. *Análise do uso de modelos discretizados para identificação de modelos de biorreatores anaeróbicos*. Proceeding

Series of the Brazilian Society of Computational and Applied Math, v. 4, n. 1, 2016. pp. 010059-1-010059-7.

- [23] EKELAND, I.; DJITTÈ, N. *An inverse problem in the economic theory of demand*. Elsevier SAS, Canadá, 2005. pp. 269-281. 13p. Disponível em: <<https://www.ceremade.dauphine.fr/~ekeland/Articles/Djitte.pdf>>. Acesso em: 29 set. 2015.
- [24] EPSTEIN, C. L.. *Introduction to the Mathematics of Medical Imaging*. SIAM, Pearson Prentice Hall, Upper Saddle River, NJ, 2 ed., 2008. 764 p.
- [25] GARCIA, R.D.M.. *Métodos para a solução da equação de transporte de partículas Íntegro-diferencial*. Escola de Verão em Teoria de Transporte de Partículas Neutras, PUC - Porto Alegre, RS, 2002.
- [26] GAVIN, H. P. *The Levenberg-Marquardt method for nonlinear least squares curve-fitting problems*. Department of Civil and Environmental Engineering, Duke University, 2016. 18 p. Disponível em: <<http://people.duke.edu/~hpgavin/ce281/lm.pdf>>. Acesso em: 10 jun. 2016.
- [27] HJORTELAND, T. *The Action Variational Principle in Cosmology*. Tese para o grau de Candidatus Scientiarum - Institute of Theoretical Astrophysics University of Oslo, University of Oslo, Oslo, 1999. Disponível em: <<http://trond.hjorteland.com/thesis/node26.html>>. Acesso em: 30 dez 2015.
- [28] ISAKOV, V. *On inverse problems in secondary oil recovery*. Department of Mathematics and Statistics, Wichita State University, Wichita, USA, 2007. 17 p. Disponível em: <<http://www.math.wichita.edu/~isakov/oil2007.pdf>>. Acesso em: 29 set. 2015.

- [29] KAMERMANS, M. P. *Gaussian Quadrature Weights and Abscissae*. Website. Github, 2011. Disponível em: <<https://pomax.github.io/bezierinfo/legendre-gauss.html>>. Acesso em: 8 maio 2016.
- [30] LINHUA, L.; HEPING, T.; QIZHENG, Y. *Inverse radiation problem of boundary incident radiation heat flux in semitransparent planar slab with semitransparent boundaries*. Journal of Thermal Science, v. 7, n. 2, 1998. pp.131-138.
- [31] LJUNG, L. *System Identification: Theory for the User*. Wiley Encyclopedia of Electrical and Electronics Engineering, Prentice-Hall, Upper Saddle River, 2ed., 1999.
- [32] LOUIS, A. K. *Inverse problems in medicine*. Proceedings of the Third German-Italian Symposium Applications of Mathematics in Industry and Technology, Teubner Stuttgart, 1 ed., 1989. pp. 277-287. 344 p.
- [33] MAFFRA, F. A.; GATTASS, M. *Uma implementação do método de Levenberg-Marquardt em C*. Website, 2015. Disponível em: <[http://webserver2.tecgraf.puc-rio.br/~mgattass/LM\\_Fabiola/](http://webserver2.tecgraf.puc-rio.br/~mgattass/LM_Fabiola/)>. Acesso em: 18 nov. 2015.
- [34] MATHEWS, L. K.; VISKANTA, R.; INCROPERA, F. P. *Development of Inverse Methods for Determining Thermophysical and Radiative Properties of High Temperature Fibrous Materials*. International Journal of Heat and Mass Transfer, v. 27, n. 4, Heat Transfer Laboratory, School of Mechanical Engineering, Purdue University, West Lafayette, 1984. pp. 487-495.
- [35] MCCORMICK, N. J. *Recent Developments in Inverse Scattering Transport Method*. Transport Theory and Statistical Physics, v. 13, n. 1-2, Taylor & Francis, USA, 1984. pp. 15-28.

- [36] MCCORMICK, N. J. *Inverse Radiative Transfer Problems: A Review*. Nuclear Science and Engineering, v. 112, 1992. pp. 185-198.
- [37] MODEST, M. F. *Radiative Heat Transfer*. McGraw-Hill, New York, 1 ed., 1993.
- [38] NATTERER, F.; WÜBBELING, F. *Mathematical methods in image reconstruction*. SIAM Monographs on Mathematical Modeling and Computation, Society for Industrial and Applied Mathematics, Philadelphia, PA, 1 ed., 2001. 216 p.
- [39] OLIVEIRA, J. P. L. *Caracterização da pluma de contaminação numa antiga lixeira com o método de resistividade elétrica*. Dissertação de Mestrado - Universidade Nova de Lisboa, Faculdade de Ciências e Tecnologia, Departamento de Ciências da Terra, Lisboa, 2009. 73 p. Disponível em: <[http://run.unl.pt/bitstream/10362/2364/1/Oliveira\\_2009.pdf](http://run.unl.pt/bitstream/10362/2364/1/Oliveira_2009.pdf)>. Acesso em: 29 set. 2015.
- [40] On My PHD. *Gradient Descent*. Website. Disponível em: <<http://www.onmyphd.com/?p=gradient.descent>>. Acesso em: 18 nov 2015.
- [41] On My PHD. *Gradient Descent Lemma*. Website. Disponível em: <[http://www.onmyphd.com/?p=math.statements#descent\\_lemma](http://www.onmyphd.com/?p=math.statements#descent_lemma)>.
- [42] ÖZISIK, M. N. *Radiative Transfer and interactions with conduction and convection*. John Wiley & Sons, New York, USA, 1 ed., 1973. 575 p.
- [43] ÖZISIK, M. N.; ORLANDE, H. R. B. *Inverse Heat Transfer: Fundamentals and Applications*. Editora: Taylor & Francis, New York, 1ed., 2000. 352 p.



- [44] PAZINATTO, C. B. *Formulação ADO para o problema Adjunto de Transporte Unidimensional e Aplicação em um Problema Inverso de Reconstrução de Fonte*. Dissertação de Mestrado. PPGMAp, UFRGS, Porto Alegre, 2015.
- [45] Plano Nacional de Energia 2030. *Parte II, Fontes Renováveis: Biomassa*. Aneel, Brasil, s.d. 12 p. Disponível em: <[http://www2.aneel.gov.br/arquivos/pdf/atlas\\_par2\\_cap4.pdf](http://www2.aneel.gov.br/arquivos/pdf/atlas_par2_cap4.pdf)>. Acesso em: 21 set. 2016.
- [46] PEDROSO, A.G.; KOCH, F. F. *Uso de Resíduos na Compostagem e Energia. Projeto de Biodigestão Anaeróbia Ecocitrus Narutovos*. Painel. Congresso Mercosul de Biomassa de Bioenergia, Teatro da UCS, Universidade de Caxias do Sul, Caxias do Sul, 2012.
- [47] RANGANATHAN, A. *The Levenberg-Marquardt Algorithm*. 2014. 5 p. Disponível em: <<http://www.ananth.in/docs/lmtut.pdf>>. Acesso em: 21 nov 2015.
- [48] RUPP, K. *Numerical Solution of the Boltzmann Transport Equation using Spherical Harmonics Expansions*. Dissertação. Institut für Analysis und Scientific Computing, Technischen Universität Wien, Wien, 2003.
- [49] SCHMETTERER, L. *Reminiscences to Johann Radon*. Lecture Notes in Mathematics and Physics, v. 4, 75 years of Radon transform, Cambridge University, Cambridge, 1 ed., 1994.
- [50] SCHULZ, D. M. Schulz. *Métodos Analíticos e Computacionais em Geofísica Nuclear*. Dissertação de Mestrado. PPGMAp, UFRGS, Porto Alegre, 2014.

- [51] SIEWERT, C. E.; MAIORINO, J. R. *The Inverse Problem for a Finite Rayleigh-Scattering Atmosphere*. Journal of Applied Mathematics and Physics, v. 31, USA, 1980. pp. 767-770.
- [52] SIEWERT, C. E.; ÖZISIK, M. N.; YENER, Y. *On the Inverse Problem for Multigroup Neutron Transport Theory*. Nuclear Science and Engineering, v. 63, n. 1, American Nuclear Society, USA, 1977. pp. 95-96.
- [53] SIEWERT, C. E. *Inverse solutions to radiative-transfer problems based on the binomial or the Henyey-Greenstein scattering law*. Journal of Quantitative Spectroscopy and Radiative Transfer v. 72, n. 6, Elsevier, North Carolina State University, Raleigh, 2002. pp. 827-835.
- [54] SILVA NETO, A. J.; CAMPOS VELHO, H. F. *Inteligência Computacional Aplicada a Problemas Inversos. Capítulo 3: Problemas Inversos em Transferência Radiativa, uma formulação implícita*. Editora: EdUERJ, Rio de Janeiro, 2011. p. 278.
- [55] SILVA NETO, A. J., MOURA NETO, F. D. *Problemas Inversos: Conceitos Fundamentais e Aplicações*. Editora: EdUERJ, Rio de Janeiro, 1 ed., 2005. 168 p.
- [56] TARANTOLA, A. *Inverse Problem Theory and Methods for Model Parameter Estimation*. SIAM, Paris, 2005.
- [57] TITO, M. J. B. *Problemas Inversos em Transferência Radiativa em Meios Heterogêneos Bidimensionais*. Dissertação de Mestrado. COPPE, UFRJ, Rio de Janeiro, 2001.
- [58] TYGEL, M. *Imageamento de estruturas geológicas na subsuperfície com vistas a prospecção sísmica de hidrocarbonetos*. UNICAMP, Instituto de Matemática, Estatística e Ciência da Computação (IMECC), Campinas, Editora: EdUNICAMP, 1993.

- [59] WICK, G. C. *Über Ebene Diffusions probleme*. Springer, Z. Physik, v. 121, n. 11-12, 1943. pp. 702-718
- [60] YUAN, Y. *A review of trust region algorithms for optimization*. State Key Laboratory of Scientific and Engineering Computing, Institute of Computational Mathematics and Scientific/Engineering Computing, Chinese Academy of Sciences, Beijing, China, s.d. Disponível em: <<ftp://lsec.cc.ac.cn/pub/yyx/papers/p995.pdf>>. Acesso em: 20 set 2016.