

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

**NOVA METODOLOGIA PARA ESTIMAÇÃO DE
PARÂMETROS DE MODELOS COM
BIFURCAÇÃO HOPF**

DISSERTAÇÃO DE MESTRADO

Ricardo França Rodrigues

**Porto Alegre
2018**

NOVA METODOLOGIA PARA ESTIMAÇÃO DE PARÂMETROS DE MODELOS COM BIFURCAÇÃO HOPF

Ricardo França Rodrigues

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

Área de concentração: Pesquisa e Desenvolvimento de Processos.

Linha de Pesquisa: Projeto, Simulação, Modelagem, Controle e Otimização de Processo Químicos e Bioprocessos.

Orientadores:

Prof. Dr. Marcelo Farenzena

Prof. Dr. Jorge Otávio Trierweiller

**Porto Alegre
2018**

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

A Comissão Examinadora, abaixo assinada, aprova a Nova Metodologia para Estimação de Parâmetros de Modelos com Bifurcação Hopf, elaborada por Ricardo França Rodrigues, como requisito parcial para obtenção do Grau de Mestre em Engenharia.

Comissão Examinadora:



Profa. Dra. Luciane Ferreira Trierweiler – DEQUI/UFRGS



Prof. Dr. Márcio Schwaab – DEQUI/UFRGS



Dr. André Domingues Quelhas – PETROBRAS

Agradecimentos

Aos meus pais, Bete e Sérgio, pelo apoio, amor e carinho, hoje e sempre.

Aos irmãos que encontrei e que me acompanharam nessa jornada, no melhor e no pior: Lopes, Richard, Renan, Vítor, Bruno, Pedro.

À família que a UFRGS me proporcionou na forma de uma turma com os melhores colegas e amigos que eu poderia ter pedido.

Ao Farenzena e ao Jorge pelo apoio, compreensão e orientação que levaram a esse trabalho do qual me orgulho.

Ao GIMSCOP pelo espaço e acolhimento.

À Petrobras pelo apoio financeiro.

*"I learned that courage was not the absence of fear, but the triumph over it.
The brave man is not he who does not feel afraid, but he who conquers that fear."
(Nelson Mandela)*

Resumo

A crescente importância da extração de petróleo e gás natural na economia brasileira tem resultado em relevantes investimentos no setor, resultando nos últimos anos recordes históricos de produção. Em contrapartida, à medida que um reservatório de petróleo amadurece, começam a ser observadas instabilidades na produção. A principal fonte desse fenômeno é o escoamento em golfadas (*slugging*). Seu surgimento é indesejado por causar dificuldades de operação e perdas na produção.

Estudos sobre supressão das golfadas são realizados desde 1930. Dentre as medidas possíveis, o uso de controle automático no sistema de elevação tem recebido especial atenção em anos recentes. A metodologia mais comum consiste em controlar uma medida de pressão na tubulação através da manipulação da válvula *choke* de produção. A incorporação de um modelo simplificado dá ao controlador recursos de antecipação para as medidas de controle.

Contudo, a qualidade de predição do modelo é limitada pelos parâmetros de ajuste utilizados. Os métodos de obtenção dos parâmetros são complexos e geralmente requerem tentativa e erro. Poucos autores propuseram metodologias de estimação desses parâmetros com base em algoritmos de otimização. O grande obstáculo para essa linha é inerente aos dados de referência utilizados por apresentarem ciclo limite estável.

Assim, o presente trabalho propõe uma nova metodologia para estimação de parâmetros de modelos com ciclo limite. Seu uso está condicionado a uma metodologia de tratamento dos dados que visa concentrar as informações de múltiplos ciclos em um único ciclo representativo através de superposição de dados. Além disso, critérios de estabilidade foram estabelecidos devido à natureza dos modelos. O método foi aplicado a dois modelos da literatura, cujos resultados publicados foram comparados estatisticamente aos aqui obtidos.

Os parâmetros obtidos com a nova metodologia apresentaram ajustes melhores nos dois casos estudados. Dado que a capacidade preditiva de uma estrutura de controle baseada em modelos depende justamente da representatividade do modelo utilizado, a obtenção de bons ajustes se traduz em resultados melhores para o controlador. Assim, a metodologia proposta neste trabalho se comprova como um importante acessório na melhoria da eficiência de recuperação dos sistemas de elevação.

Abstract

The rising importance of the oil and gas industry for the Brazilian economy has brought significant investments to this sector. Recent years have registered record levels of production. On the other hand, as a reservoir reaches advanced stages of its life cycle, production instabilities begin to arise due to the slugging phenomenon. This causes operation problems as well as production losses that justify researching methods of suppression.

Slugging suppression has been investigated as early as the 1930s. Among the anti-slugging measures is the usage of automated control in oil elevation systems. The most common methodology revolves around controlling a pressure measure in the production line by manipulating the production choke valve. The addition of simplified slugging models incorporates a predictive response to the control system.

However, the quality of model predictions is limited by the parameters used. Methods of obtaining these parameters are complex and often require trial and error. Few authors have proposed parameter estimation methods based on optimization algorithms. The main issue for this approach is the limit cycle behavior of the reference data required to adjust these models.

Therefore, the present work proposes a new objective function for parameter estimation of limit cycle models. Its usage is conditioned to a data treatment methodology that aims at representing multiple cycles in a reference data set into a single mean cycle. Furthermore, stability criteria were established due to the dual nature of these models, which may present steady state or limit cycle responses. The proposed method was applied to two published models, whose results were statistically compared to those obtained hereafter.

The parameters estimated using the new method presented better model fits for both case studies. Given that the predictive capacity of a control structure based on models heavily depends on the quality of the employed models, obtaining better fits will present better results for the controller. Therefore, the proposed method is proven to be an important accessory in the road to improving the recovery efficiency of elevation systems.

LISTA DE FIGURAS

Figura 1.1: Participações estimadas do setor de petróleo e da Petrobras no PIB a preços de mercado (PIBpm) brasileiro de 1955 a 2003. (ARAGÃO, 2004).....	1
Figura 1.2: Ilustração da influência das instabilidades oriundas do regime de golfadas na produção de petróleo em malha aberta. Adaptado de Hu (2004) e Di Meglio <i>et al.</i> (2012).....	2
Figura 2.1: Representação de sistema de elevação offshore. Adaptado de Di Meglio <i>et al.</i> (2012).....	6
Figura 2.2: Representação de um sistema <i>gas-lift</i> . Adaptado de Aamo <i>et al.</i> (2005).....	6
Figura 2.3: Tipos de escoamento horizontal. Adaptado de Jahanshahi (2013).....	8
Figura 2.4: Tipos de escoamento vertical. Adaptado de Jahanshahi (2013).....	8
Figura 2.5: Mapa de regimes para escoamento vertical. Zonas tracejadas indicam regiões de transição. Adaptado de Weisman e Kang (1981).	9
Figura 2.6: Exemplo de ciclo das golfadas por acidente em sistemas <i>pipeline-riser</i> sendo a fase líquida representada em azul. Adaptado de Storkaas (2005).	10
Figura 2.7: Bifurcação Hopf gerada a partir do exemplo. (SEYDEL, 2010).....	11
Figura 2.8: Representação do modelo de três estados. (DI MEGLIO <i>et al.</i> , 2012)	13
Figura 2.9: Representação do modelo de seis estados. (DIEHL <i>et al.</i> , 2017).....	17
Figura 2.10: Ciclo completo de uma golfada obtido com o simulador OLGA. Adaptado de Diehl <i>et al.</i> (2017).....	21
Figura 4.1: (a) Superfície de otimização e (b) curvas de níveis, da função objetivo de mínimos quadrados para o exemplo.	32
Figura 4.2: Funções senoidais utilizadas como exemplo.	32
Figura 4.3: Ciclos de referência para comparação ponto a ponto.	33
Figura 4.4: (a) Superfície de otimização e (b) curvas de níveis. da função objetivo proposta por Nabavi e Williams para o exemplo.	33
Figura 4.5: Influência da taxa de amostragem na resposta da função objetivo proposta por Nabavi e Williams.....	34
Figura 4.6: Medições reais de pressão de um poço de petróleo.....	34
Figura 4.7: Ilustração passo a passo do método proposto: (a) identificação de picos e vales; (b) identificação de todos os ciclos; (c) separação dos ciclos; (d,e) avaliação da diferença entre os ciclos; (f,g) interpolação através de <i>splines</i> ; (h,i) verificação da amostragem; (j) sobreposição dos ciclos; (h) ciclo médio com variância.....	36
Figura 4.8: Resultados da otimização do exemplo (azul) usando: WLS (vermelho); LS (verde).....	37
Figura 4.9: (a) Superfície de otimização do exemplo para Mínimos Quadrados, e (b) curvas de nível da função objetivo.....	38
Figura 4.10: (a) Superfície de otimização do exemplo para Mínimos Quadrados Ponderados, e (b) curvas de nível da função objetivo.....	38
Figura 4.11: Resposta da pressão na base do <i>riser</i> , P_r, b , do modelo de Meglio, Kaasa e Petit (2009) para diferentes aberturas de válvula <i>choke</i> , z . Os pontos representam a média (quando o valor de <i>choke</i> é inferior ao valor crítico) ou os picos e vales (quando esse valor é superior ao crítico). A linha tracejada cinza representa o diagrama de bifurcação do sistema.....	39
Figura 4.12: Representação de um oscilador harmônico com amortecimento.....	40
Figura 4.13: Resposta da posição de um oscilador harmônico com amortecimento para diferentes valores de ζ	41

Figura 4.14: (a) Exemplo de ajuste da parcela exponencial de um oscilador harmônico subamortecido para valores de abertura de <i>choke</i> inferiores ao crítico o modelo de Meglio, Kaasa e Petit. (b) Quando os parâmetros ζ e A são ajustados para picos ou vales (linhas cheias), o lado recíproco apresenta um ajuste ruim (linhas tracejadas).	42
Figura 5.1: Dados utilizados para estimação de parâmetros: (a) modelo Meglio (azul) com erro normalmente distribuído, todos os picos (vermelho) e vales (verde) detectados, e variação da pressão do separador P_s (roxo); (b) resultado da reconstrução (vermelho) dos dados (azul) baseada na diferença entre a média dos vales dos ciclos sem distúrbio (verde tracejado) e os valores da <i>spline</i> dos vales dos ciclos com distúrbio (verde cheio).	45
Figura 5.2: Comparação entre ciclo representativo dos dados reconstruídos (vermelho) e dos dados sem reconstrução (azul).	46
Figura 5.3: Resultados das estimações (a) sem reconstrução e (b) com reconstrução.	46
Figura 5.4: Comparação entre ciclo dos dados reconstruídos (pontos vermelhos), ciclo do modelo estimado com dados reconstruídos (linha vermelha) e ciclo do modelo estimado sem reconstrução (linha azul).	47
Figura 5.5: (a) Superfície de otimização e (b) curvas de níveis de Mínimos Quadrados para os parâmetros m_l , $still$ e $Cout$	49
Figura 5.6: (a) Superfície de otimização e (b) curvas de níveis da função objetivo proposta por Diehl <i>et al.</i> (2017) para os parâmetros m_l , $still$ e $Cout$	49
Figura 5.7: (a) Superfície de otimização e (b) curvas de níveis da função objetivo proposta neste trabalho para os parâmetros m_l , $still$ e $Cout$	49
Figura 5.8: Regiões de confiança para um grau de confiança $\alpha = 99\%$ dos parâmetros obtidos com reconstrução dos dados: região elíptica (vermelho) e região de máxima verossimilhança (azul).	50
Figura 5.9: Dados obtidos do artigo de Di Meglio <i>et al.</i> (2012).	51
Figura 5.10: Ciclos representativos: dados de referência (azul); modelo Meglio conforme publicado (vermelho).	52
Figura 5.11: Ciclos representativos: dados originais (azul); modelo Meglio original (vermelho); modelo Meglio estimado com base na pressão Pr, b (verde).	53
Figura 5.12: Evolução da melhor posição da função objetivo para a estimação com Pr, b	53
Figura 5.13: Valores preditos e valores observados de Pr, b e Pr, t do modelo Meglio estimado com Pr, b	54
Figura 5.14: Regiões de confiança de máxima verossimilhança para o modelo de Di Meglio <i>et al.</i> (2009) estimado com Pr, b . Valores ótimos são representados em vermelho.	55
Figura 5.15: Ciclos representativos: dados originais (azul); modelo Meglio original (vermelho); modelo Meglio estimado com base na pressão Pr, t (verde).	55
Figura 5.16: Evolução da melhor posição da função objetivo para a estimação com Pr, t	56
Figura 5.17: Valores preditos e valores observados de Pr, b e Pr, t do modelo Meglio estimado com Pr, t	57
Figura 5.18: Regiões de confiança de máxima verossimilhança para o modelo Meglio estimado com Pr, t . Valores ótimos são representados em vermelho.	57

Figura 5.19: Ciclos representativos: dados originais (azul); modelo Meglio original (vermelho); modelo Meglio estimado com base nas pressões Pr, b e Pr, t (verde).	59
Figura 5.20: Evolução da melhor posição da função objetivo para a estimação com Pr, b e Pr, t .	59
Figura 5.21: Valores preditos e valores observados de Pr, b e Pr, t do modelo Meglio estimado com Pr, b e Pr, t .	60
Figura 5.22: Regiões de confiança de máxima verossimilhança para o modelo Meglio estimado com Pr, b e Pr, t . Valores ótimos são representados em vermelho.	61
Figura 5.23: Dados de referência para o <i>Real Well A</i> .	62
Figura 5.24: Comparação entre ciclos obtidos para o modelo FOWM.	64
Figura 5.25: Valores preditos e observados da pressão <i>PPDG</i> : (a) dados originais, (b) cenário 1, (c) cenário 2 e (d) cenário 3.	67
Figura 5.26: Regiões de máxima verossimilhança para o cenário 1.	68
Figura 5.27: Regiões de máxima verossimilhança para o cenário 2.	69
Figura 5.28: Regiões de máxima verossimilhança para o cenário 3.	70
Figura 5.29: Ciclos médios das pressões (a) <i>PPDG</i> , (b) <i>PTPT</i> e (c) Pr, t , considerando a referência e os resultados de diferentes valores de ωu . Os resultados foram obtidos utilizando somente <i>PPDG</i> como referência.	71
Figura 5.30: Evolução da função objetivo para a estimação do modelo FOWM com base em três pressões.	72
Figura 5.31: Comparação entre as pressões (a) <i>PPDG</i> , (b) <i>PTPT</i> e (c) Pr, t para a referência e os diferentes conjuntos de parâmetros obtidos para o modelo FOWM.	73
Figura 5.32: Comparação entre as pressões (a) <i>PPDG</i> , (b) <i>PTPT</i> e (c) Pr, t para a referência e o conjunto de parâmetros estimado para o modelo FOWM.	74
Figura 5.33: Valores preditos e observados da (a) <i>PPDG</i> , (b) <i>PTPT</i> e (c) Pr, t do modelo FOWM estimado com três pressões.	77
Figura 5.34: Regiões de máxima verossimilhança para o modelo FOWM estimado com três pressões.	78

LISTA DE TABELAS

Tabela 1: Comparação entre os resultados das estimações de parâmetros do modelo de Meglio et al. (2010).	47
Tabela 2: Teste Chi Quadrado entre os resultados com e sem reconstrução dos dados.....	47
Tabela 3: Teste F entre os resultados com e sem reconstrução dos dados.	48
Tabela 4: Matriz de correlações dos resultados com reconstrução dos dados. ...	48
Tabela 5: Comparação entre os parâmetros estimados para o modelo Meglio....	51
Tabela 6: Limites de otimização para os parâmetros do modelo Meglio.	51
Tabela 7: Teste Chi quadrado para o modelo Meglio original.	52
Tabela 8: Teste Chi Quadrado para o modelo Meglio estimado com a pressão Pr, b	53
Tabela 9: Matriz de correlações do modelo Meglio estimado com a pressão Pr, b	54
Tabela 10: Teste F do modelo Meglio estimado com a pressão Pr, b	54
Tabela 11: Teste Chi Quadrado do modelo Meglio estimado com a pressão Pr, t	56
Tabela 12: Matriz de correlações do modelo Meglio estimado com a pressão Pr, t	56
Tabela 13: Teste F do modelo Meglio estimado com a pressão Pr, t	57
Tabela 14: Teste Chi Quadrado para o modelo Meglio estimado com as pressões Pr, b e Pr, t	59
Tabela 15: Teste F do modelo Meglio estimado com as pressões Pr, b e Pr, t	60
Tabela 16: Matriz de correlações para as medidas de Pr, b do modelo Meglio estimado com as pressões Pr, b e Pr, t	60
Tabela 17: Matriz de correlações para as medidas de Pr, t do modelo Meglio estimado com as pressões Pr, b e Pr, t	60
Tabela 18: Comparação entre resultados originais e resultados obtidos com a nova metodologia.	63
Tabela 19: Teste Chi Quadrado para o modelo FOWM baseado em dados reais.	64
Tabela 20: Coeficientes de correlação obtidos do resultado original e dos cenários para o modelo FOWM.	64
Tabela 21: Teste F para o modelo FOWM entre os cenários.....	65
Tabela 22: Teste F para o modelo FOWM entre os cenários e o resultado original.	65
Tabela 23: Matriz de correlações do modelo FOWM para cenário 1.....	66
Tabela 24: Matriz de correlações do modelo FOWM para cenário 2.....	66
Tabela 25: Matriz de correlações do modelo FOWM para cenário 3.....	66
Tabela 26: Comparação entre parâmetros estimados do modelo FOWM para o <i>Real Well A</i>	72
Tabela 27: Coeficientes de correlação para os conjuntos de parâmetros estudados para o modelo FOWM.	73
Tabela 28: Teste Chi Quadrado para o modelo FOWM estimado com três pressões.	74

Tabela 29: Teste F para o modelo FOWM entre os diferentes conjuntos de parâmetros obtidos para o <i>Real Well A</i>	75
Tabela 30: Matriz de correlações da <i>PPDG</i> do modelo FOWM estimado com três pressões.....	75
Tabela 31: Matriz de correlações da <i>PTPT</i> do modelo FOWM estimado com três pressões.....	76
Tabela 32: Matriz de correlações da <i>Pr, t</i> do modelo FOWM estimado com três pressões.....	76

NOTAÇÃO E SIMBOLOGIA

u	Abertura de válvula
z	Abertura de válvula
g	Aceleração da gravidade
A_{SS}	Área da seção transversal da tubulação <i>subsea</i>
ρ^m	Coefficiente de correlação do modelo
C_{out}	Coefficiente de fluxo da válvula <i>choke</i>
C_g	Coefficiente de fluxo da válvula virtual
K_a	Coefficiente de fluxo do casco anular para o <i>tubing</i>
K_{wh}	Coefficiente de fluxo do <i>manifold</i>
K_r	Coefficiente de fluxo do reservatório
L_a	Comprimento do casco anular
L_{fl}	Comprimento do <i>flowline</i>
L_t	Comprimento do <i>riser</i>
π	Constante pi
R	Constante universal dos gases
D_{SS}	Diâmetro da tubulação <i>subsea</i>
D_a	Diâmetro do casco anular
D_t	Diâmetro do <i>tubing</i>
H_{vgl}	Distância vertical do <i>manifold</i> até a válvula de <i>gas-lift</i>
H_t	Distância vertical do <i>manifold</i> até o reservatório
H_{pdg}	Distância vertical do <i>manifold</i> até o sensor PDG
$\chi_{\alpha, \nu}^2$	Distribuição Chi Quadrado
F_{ν_2, ν_1}^α	Distribuição F de Fisher
t_ν^α	Distribuição t de Student
EDO	Equações Diferenciais Ordinárias
EDP	Equações Diferenciais Parciais
φ	Fase
FOWM	<i>Fast Offshore Wells Model</i>
ω_u	Fator de correção de volume
w_{gout}	Fluxo mássico de gás da válvula <i>choke</i>
w_{iv}	Fluxo mássico de gás da válvula de <i>gas-lift</i>

w_g	Fluxo mássico de gás da válvula virtual
w_{gc}	Fluxo mássico de gás entrante no casco anular
w_{whg}	Fluxo mássico de gás no <i>manifold</i>
w_{lout}	Fluxo mássico de líquido da válvula <i>choke</i>
w_{whl}	Fluxo mássico de líquido no <i>manifold</i>
w_r	Fluxo mássico do reservatório
ϵ	Fração de gás que entra diretamente no <i>riser</i>
α_{gw}	Fração mássica de gás do reservatório
α_{gt}	Fração mássica de gás do <i>tubing</i>
α_g	Fração mássica de gás na tubulação <i>subsea</i>
α_l	Fração mássica de líquido na tubulação <i>subsea</i>
F_{obj}	Função objetivo
α	Grau de confiança
θ	Inclinação média do <i>riser</i>
m_{gb}	Massa de gás na bolha alongada
m_{ga}	Massa de gás no casco anular
m_{gr}	Massa de gás no <i>riser</i>
m_{gt}	Massa de gás no <i>tubing</i>
m_{lr}	Massa de líquido no <i>riser</i>
m_{lt}	Massa de líquido no <i>tubing</i>
$m_{l,still}$	Massa de líquido sempre presente no <i>riser</i>
ρ_{mres}	Massa específica da mistura no reservatório
ρ_{mt}	Massa específica da mistura no <i>tubing</i>
ρ_{ai}	Massa específica de gás no casco anular
ρ_{gt}	Massa específica de gás no <i>tubing</i>
ρ_l	Massa específica de líquido
M	Massa molar de gás
R_θ	Matriz de correlações dos parâmetros
V_θ	Matriz de covariâncias dos parâmetros
$\hat{\sigma}_y^2$	Matriz de variâncias de predição
V_y	Matriz dos erros experimentais
LS	Mínimos Quadrados
WLS	Mínimos Quadrados Ponderados
NE	Número de Experimentos
ν	Número de graus de liberdade

NP	Número de parâmetros
P_a	Pressão Anular
P_{eb}	Pressão Bolha Alongada
P_{bh}	Pressão <i>Bottomhole</i>
P_r	Pressão do reservatório
P_s	Pressão do Separador
P_{PDG}	Pressão <i>Permanent Downhole Gauge</i>
$P_{r,b}$	Pressão <i>Riser Base</i>
$P_{r,t}$	Pressão <i>Riser Top</i>
P_{TPT}	Pressão <i>Temperature and Pressure Transmitter</i>
$P_{t,b}$	Pressão <i>Tubing Base</i>
$P_{t,t}$	Pressão <i>Tubing Top</i>
PIBpm	Produto Interno Bruto a preço de mercado
T	Temperatura
t	Tempo
σ^2	Variância
\dot{Q}	Vazão volumétrica de líquido
u_l	Velocidade superficial de líquido
u_g	Velocidade superficial de gás
θ	Vetor de parâmetros
V_{SS}	Voluma da tubulação <i>subsea</i>
V_{eb}	Volume da bolha alongada
V_{gt}	Volume de gás no <i>tubing</i>
V_a	Volume do casco anular
V_t	Volume total do <i>tubing</i>

SUMÁRIO

Capítulo 1 – Introdução.....	1
1.1 Motivação.....	2
1.2 Objetivos	3
1.3 Estrutura da dissertação.....	3
Capítulo 2 – Revisão bibliográfica	5
2.1 Sistemas de elevação de petróleo.....	5
2.2 Escoamento multifásico.....	7
2.3 Regime de golfadas	9
2.3.1 Golfadas por acidente	9
2.3.2 Golfadas por casing-heading	10
2.3.3 Golfadas por ondas de densidade	10
2.4 Ciclo limite e bifurcação Hopf	11
2.5 Modelagem aplicada a automação e controle	12
2.6 Modelo de Di Meglio, Kaasa e Petit (2009).....	12
2.6.1 Descrição do modelo.....	13
2.6.2 Ajuste de parâmetros	15
2.7 Modelo de Diehl <i>et al.</i> (2017).....	16
2.7.1 Descrição do modelo.....	16
2.7.2 Ajuste de parâmetros	20
2.8 Estimação de parâmetros para modelos com bifurcação Hopf	22
Capítulo 3 – Fundamentos de avaliação estatística para modelos.....	24
3.1 Qualidade dos modelos.....	24
3.2 Qualidade dos parâmetros.....	25
3.2.1 Intervalos de confiança	27
3.2.2 Regiões de confiança.....	27
3.2.3 Comparação entre conjuntos de parâmetros	29
3.3 Qualidade de predição do modelo.....	29
Capítulo 4 – Metodologia	30
4.1 Funções objetivo.....	30
4.1.1 Modelos com ciclo limite	30
4.1.2 Modelos com ciclo limite e alta variabilidade.....	34
4.1.3 Critérios de convergência para modelos com bifurcação Hopf	38
4.2 Recursos computacionais	43
Capítulo 5 – Estudos de caso.....	44
5.1 Modelo Meglio com distúrbio	44
5.2 Modelo Meglio aplicado a dados reais.....	50
5.2.1 Resultados originais	52
5.2.2 Estimação com base em Pr, b	52
5.2.3 Estimação com base em Pr, t	55
5.2.4 Estimação com base em Pr, b e Pr, t	58
5.3 Modelo FOWM com dados reais.....	61
5.3.1 Influência do parâmetro ωu	63
5.3.2 Estimação de parâmetros com outras referências	71
Capítulo 6 – Conclusões e Trabalhos futuros.....	79
6.1 Conclusões quanto à metodologia	79
6.2 Estratégia de otimização do modelo FOWM.....	79
6.3 Sugestões de trabalhos futuros	80

Referências	81
Apêndice A – Códigos implementados em Python.....	87

Capítulo 1 – Introdução

A extração de petróleo e gás natural representam uma significativa parcela do PIB nacional. Após o decreto da chamada Lei do Petróleo, em 1998, foi criada a Agência Nacional do Petróleo, Gás Natural e Biocombustíveis (ANP) e instituiu-se o fim do monopólio de extração da Petrobras. Isso causou o início da crescente participação do setor de petróleo no PIB nacional, como pode ser observado na Figura 1.1:

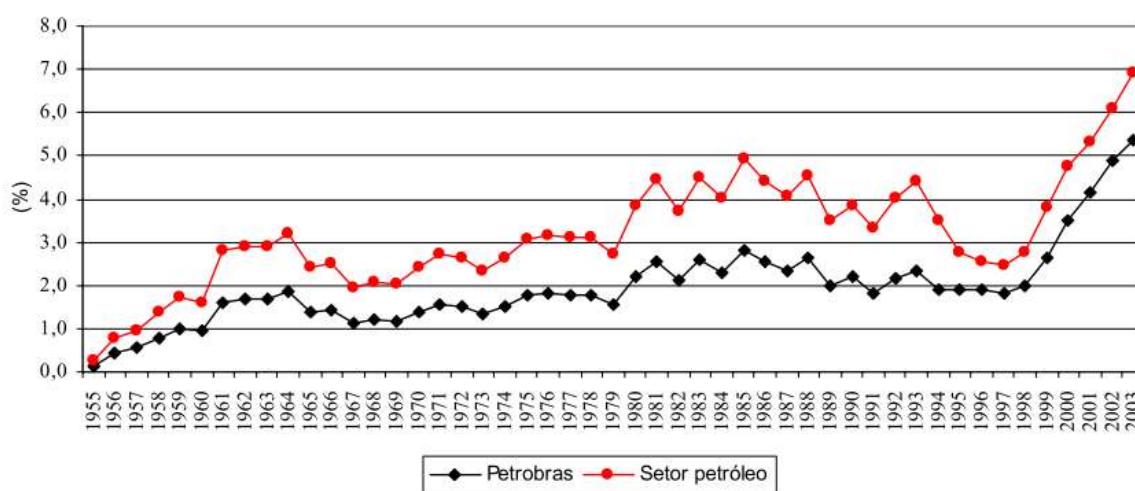


Figura 1.1: Participações estimadas do setor de petróleo e da Petrobras no PIB a preços de mercado (PIBpm) brasileiro de 1955 a 2003. (ARAGÃO, 2004)

Somente no período entre 1998 e 2003, essa participação cresceu de 2,5% para 6,9% (ARAGÃO, 2004). Estima-se que em 2017 a participação tenha sido de 13% do PIB nacional, representando 50% da oferta interna de energia do país. A produção de barris por dia de petróleo cresceu 268% só entre o período de 1998 a 2017, aumentando de 970 mil para 2,6 milhões de barris por dia. Já a produção de gás natural cresceu 370%, de 30 milhões para 111 milhões de metros cúbicos por dia (ANP, 2018). Um fator que contribuiu para o aumento da participação do setor foi a descoberta da camada pré-sal em 2007, que em julho de 2017 representou 76% da produção total de barris por dia (PETROBRAS, 2017).

1.1 Motivação

À medida que um poço alcança um estágio maduro de sua vida útil, torna-se necessário o emprego de técnicas de elevação artificial que aumentem a recuperação de produto do poço. A técnica mais comum para essa etapa é conhecida como *gas-lift* e representava em 2009 uma parcela de 72% das técnicas de elevação artificial alocadas à produção da Petrobras. (RIZZO FILHO, 2011)

A popularidade do *gas-lift* se dá pela versatilidade: pode ser aplicado em poços com até 2.600 metros de profundidade em fluidos que apresentem alto teor de areia e elevada razão gás-líquido. Consiste basicamente na aeração do sistema de elevação pela injeção de gás natural. A contrapartida do método de *gas-lift* consiste no surgimento do regime de escoamento chamado de golfadas ou *slug*. (THOMAS, 2001)

Sob esse regime, o sistema de elevação de petróleo apresenta períodos intermitentes de produção de óleo e gás, configurando um ciclo limite. O excesso de produção de óleo pode fazer com que o separador transborde, ocasionando *shut down* do mesmo. Já um período de alta produção de gás causa dificuldades operacionais durante o *flaring*. Além disso, as altas oscilações de pressão no sistema diminuem a produção total obtida. (JANSEN, F.E.; SHOHAM; TAITEL, 1996)

As origens dessas flutuações podem ser relacionadas tanto a aberturas maiores da válvula *choke* de produção (Figura 1.2a) quanto a vazões menores de *gas-lift* (Figura 1.2b), configurando um sistema com bifurcação Hopf. É visível que existe uma margem considerável para aumento da eficiência de recuperação de petróleo se as instabilidades puderem ser atenuadas. Com isso em mente, uma das soluções possíveis consiste no emprego de métodos de automação e controle. (FOSS, 2012)

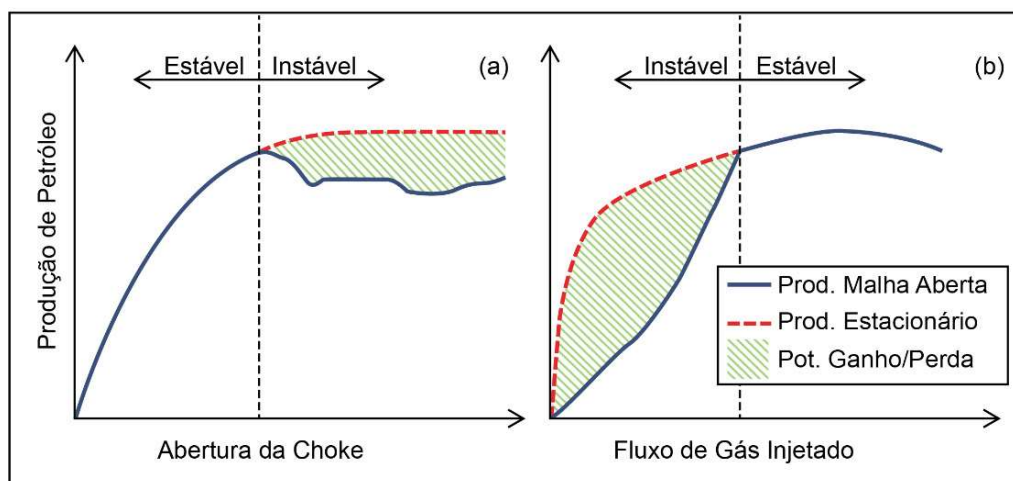


Figura 1.2: Ilustração da influência das instabilidades oriundas do regime de golfadas na produção de petróleo em malha aberta. Adaptado de Hu (2004) e Di Meglio *et al.* (2012).

Alguns autores apresentaram resultados promissores utilizando um controlador PI com uma medida de pressão do fundo da tubulação (DALSMO; HALVORSEN; SLUPPHAUG, 2002; GODHAVN; FARD; FUCHS, 2005; HAVRE; STORNES; STRAY, 2000). Contudo, medidas desse tipo nem sempre estão disponíveis. Além disso, um controlador tão simples pode apresentar dificuldades de aplicação para faixas maiores de pontos operacionais porque não incorpora termos de antecipação (DI MEGLIO *et al.*, 2012).

1.2 Objetivos

Esses obstáculos impulsionaram o desenvolvimento de modelos simplificados que pudessem descrever sistemas em regime de golfadas para melhorar a eficiência dos métodos de controle. Modelos baseados em Equações Diferenciais Ordinárias são suficientes para capturar as principais características das golfadas (DI MEGLIO; KAASA; PETIT, 2009; DIEHL *et al.*, 2017; JANSEN, BÅRD *et al.*, 1999; STORKAAS, E, 2005). Contudo, a obtenção dos parâmetros de ajuste desses modelos é complexa devido ao surgimento de ciclo limite. A literatura sobre a estimação dos parâmetros utilizando algoritmos de otimização também é escassa (APIO; BOTELHO; TRIERWEILER, 2018; DIEHL *et al.*, 2017).

Dessa forma, o objetivo primário do presente trabalho é melhorar a performance de algoritmos de otimização para estimação de parâmetros de modelos com bifurcação Hopf através de uma nova metodologia baseada em Mínimos Quadrados Ponderados. Foram estudados dois modelos simplificados para sistemas de elevação de petróleo. A nova metodologia foi aplicada em três casos de estudo: um caso teórico com distúrbio e discrepância de modelo; e dois casos com dados reais de operação. O objetivo secundário do trabalho é comparar estatisticamente os resultados obtidos com as metodologias disponíveis na literatura.

1.3 Estrutura da dissertação

O Capítulo 2 apresenta os conhecimentos necessários para o melhor entendimento do trabalho. São apresentados os sistemas de elevação de petróleo, bem como as misturas multifásicas envolvidas e as origens principais do regime de golfadas. Também são discutidos os conceitos de ciclo limite e bifurcação Hopf, assim como as metodologias disponíveis na literatura para estimação de parâmetros de modelos com esse tipo de bifurcação. Os principais métodos de controle automático baseado em modelo são comentados. Por fim, são apresentados os modelos estudados e os métodos de ajuste dos parâmetros sugeridos pelos respectivos autores.

O Capítulo 3 mostra a influência do ciclo limite na estimação de parâmetros com Mínimos Quadrados. A nova função objetivo é descrita em detalhes e suas vantagens são ilustradas com uma função senoidal. Para sua aplicação em modelos de golfadas foi necessária a determinação de critérios de estabilidade dos conjuntos de dados estudados.

O Capítulo 4 apresenta os recursos estatísticos empregados na avaliação dos resultados. A fundamentação teórica é baseada na literatura e visa explicar as métricas utilizadas tanto para os resultados obtidos quanto para os resultados de outros trabalhos para comparar metodologias.

No Capítulo 5 são apresentados os resultados para três estudos de caso com base em dois modelos. Para o primeiro modelo, são estudados três cenários em que diferentes referências são utilizadas na estimação; para o segundo, estudaram-se três cenários relacionados a um dos parâmetros do modelo e também um cenário em que três referências são usadas em uma mesma estimação.

O Capítulo 6 apresenta as conclusões e as sugestões de trabalhos futuros. Comenta-se a qualidade dos resultados obtidos e também possíveis estratégias para a estimação de parâmetros de um dos modelos, baseadas nos resultados apresentados. Outras aplicações da metodologia são sugeridas com base na versatilidade da obtenção de um ciclo representativo de um conjunto de dados em ciclo limite.

Por fim, o Apêndice A – apresenta os principais códigos implementados para a obtenção dos resultados. São apresentados tanto códigos autorais quanto códigos disponíveis de forma gratuita.

Capítulo 2 – Revisão bibliográfica

Neste capítulo serão descritos os sistemas de elevação de petróleo que utilizam *gas-lift* como fonte de elevação artificial. São descritos os escoamentos multifásicos apresentados por esses sistemas, bem como as condições para o surgimento do regime de golfadas. Também são definidos os conceitos de ciclo limite e bifurcação Hopf. Comenta-se a metodologia de automação e controle baseada em modelos para atenuação de golfadas. Por fim, são apresentados os modelos utilizados para a obtenção dos resultados deste trabalho.

2.1 Sistemas de elevação de petróleo

Na produção de petróleo, os sistemas de elevação de poços *offshore* são um conjunto de válvulas e longos tubos que transportam uma mistura complexa de água, areia, gás e óleo proveniente de um reservatório para suas instalações de processamento de superfície, conforme ilustrado na Figura 2.1. As seções do sistema podem ser divididas em reservatório, poço, *flowline*, *riser* e instalações de superfície. Os poços transportam a mistura do reservatório até o fundo do mar. Múltiplos poços podem ser ligados a uma única linha de produção por um conjunto de válvulas chamado de *manifold* ou "árvore de Natal". Do *manifold*, a mistura segue pelo *flowline* conforme a disposição da tubulação em relação ao terreno. Finalmente, a tubulação inclina até a plataforma de produção, na seção do *riser*. (DI MEGLIO *et al.*, 2012)

Contudo, certas condições de temperatura e pressão podem favorecer o surgimento de um sistema bifásico, um sistema trifásico ou até um sistema com emulsões. Ainda, muitos poços apresentam o sistema *gas-lift* como fonte de elevação natural. Como ilustrado na Figura 2.2, esse sistema consiste em uma casca externa ao poço que injeta o gás produzido na superfície de volta à linha de produção para diminuir a pressão hidrodinâmica e facilitar o fluxo da mistura. O uso de *gas-lift* é comum porque aumenta a recuperação de produtos conforme o reservatório se aproxima do fim de sua vida útil, o que diminui sua pressão (RIZZO FILHO, 2011).

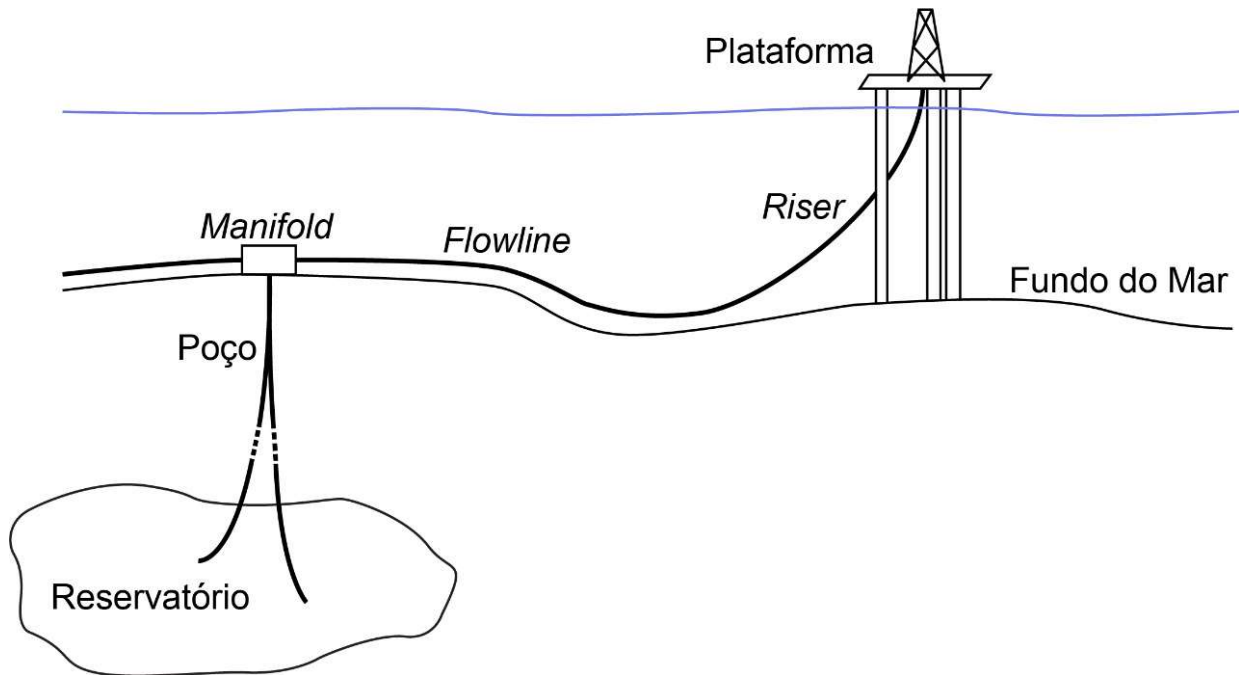


Figura 2.1: Representação de sistema de elevação offshore. Adaptado de Di Meglio et al. (2012).

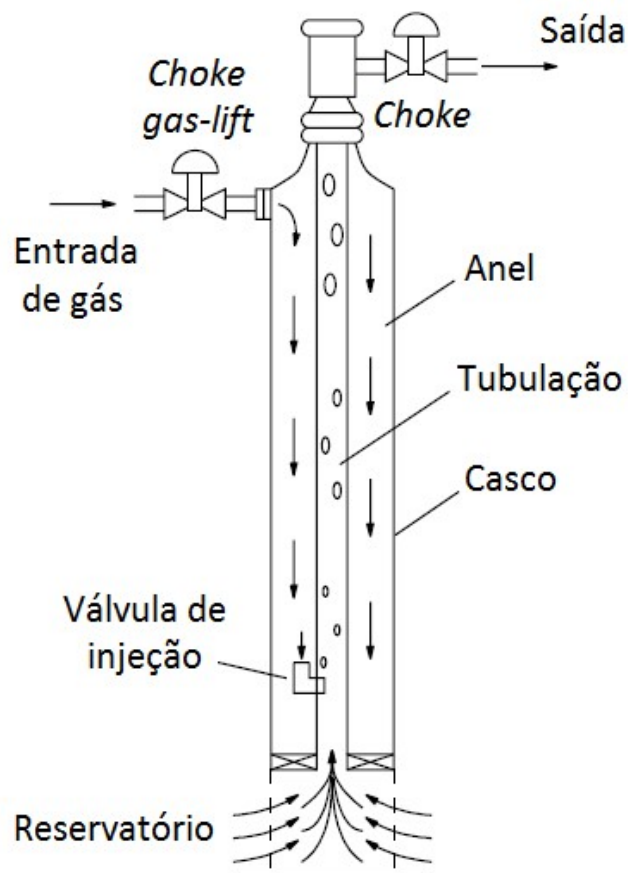


Figura 2.2: Representação de um sistema *gas-lift*. Adaptado de Aamo et al. (2005).

2.2 Escoamento multifásico

A natureza da mistura obtida dos reservatórios faz com que o processo de deslocar os produtos até plataformas de processamento envolva complexos escoamentos multifásicos. A disposição espacial das fases nos dutos pode caracterizar regimes ou padrões de escoamento com propriedades bem diferentes em relação à clássica abordagem monofásica para escoamentos (BAKER, 1954). A classificação e as características de escoamentos líquido-gás são bem conhecidas e documentadas em revisões sobre escoamento multifásico (WALLIS, 1969; WEISMAN, 1983). A partir da configuração da tubulação (vertical ou horizontal), a maior influência no regime se dá pela velocidade superficial de cada fase:

$$u_l = \frac{\dot{Q}_l}{A} \quad (2.1)$$

$$u_g = \frac{\dot{Q}_g}{A} \quad (2.2)$$

onde u_i é a velocidade superficial, \dot{Q}_i é a taxa de fluxo volumétrico, A é a área da seção transversal e os subíndices l e g representam líquido e gás, respectivamente. Regimes comuns são ilustrados na Figura 2.3 e na Figura 2.4. As classificações gerais de regime são (JAHANSHAH, 2013):

- *Escoamento estratificado*: somente possível para geometria horizontal. Apresenta duas fases distintas de gás e líquido em que o gás permanece acima do líquido por diferença de massa específica.
- *Escoamento em bolha alongada*: apresenta bolhas alongadas com três regiões distintas: nariz, corpo e cauda.
- *Escoamento em golfadas*: caracterizado por bolhas intermitentes cuja seção transversal tende a ocupar toda a área disponível para escoamento da tubulação.
- *Escoamento em bolhas*: apresenta bolhas de gás dispersas no meio líquido.
- *Escoamento anular*: a fase gás é predominante enquanto um filme de líquido escoava nas paredes do tubo.

Estudos foram feitos para relacionar as velocidades superficiais de gás e líquido com os regimes de escoamento observados (BAKER, 1954; JANSEN, F.E.; SHOHAM; TAITEL, 1996; SCHMIDT; BRILL; BEGGS, 1980, 1979; WEISMAN; KANG, 1981). Os resultados foram usados para gerar mapas de escoamento que representassem essas relações graficamente, como ilustrado na Figura 2.5. Além disso, esses estudos incluem esforços feitos na tentativa de matematicamente prever as regiões de transição. Apesar disso, os experimentos realizados restringem a validade dos modelos preditivos por serem feitos utilizando tubulações de diâmetro específico e misturas bifásicas com características físico-químicas distintas das misturas reais (BRENNEN, 2005).

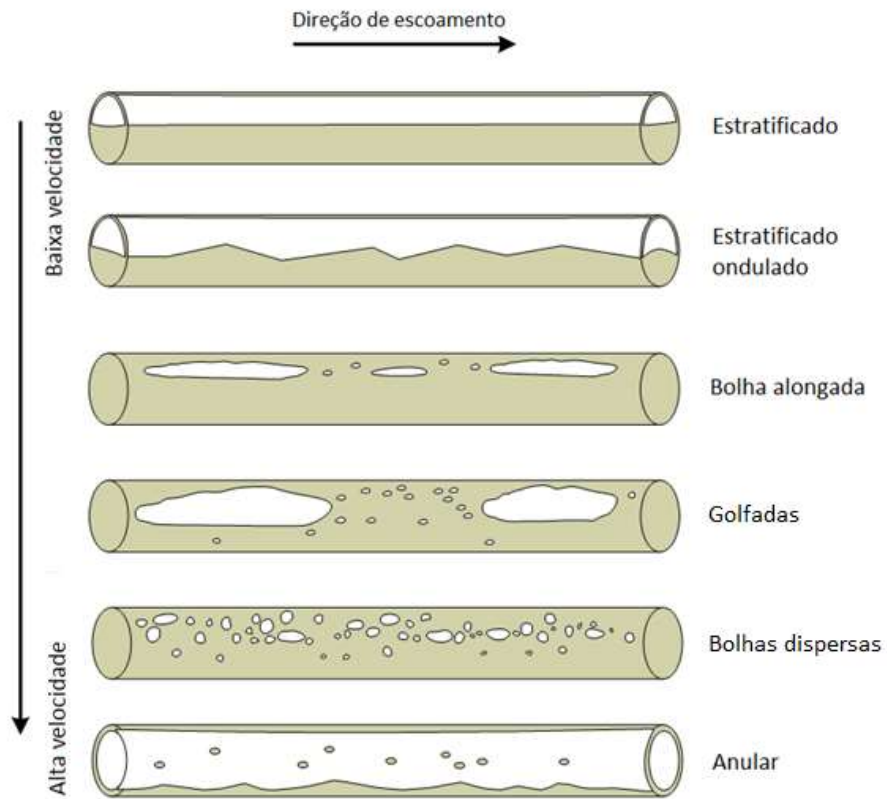


Figura 2.3: Tipos de escoamento horizontal. Adaptado de Jahanshahi (2013).

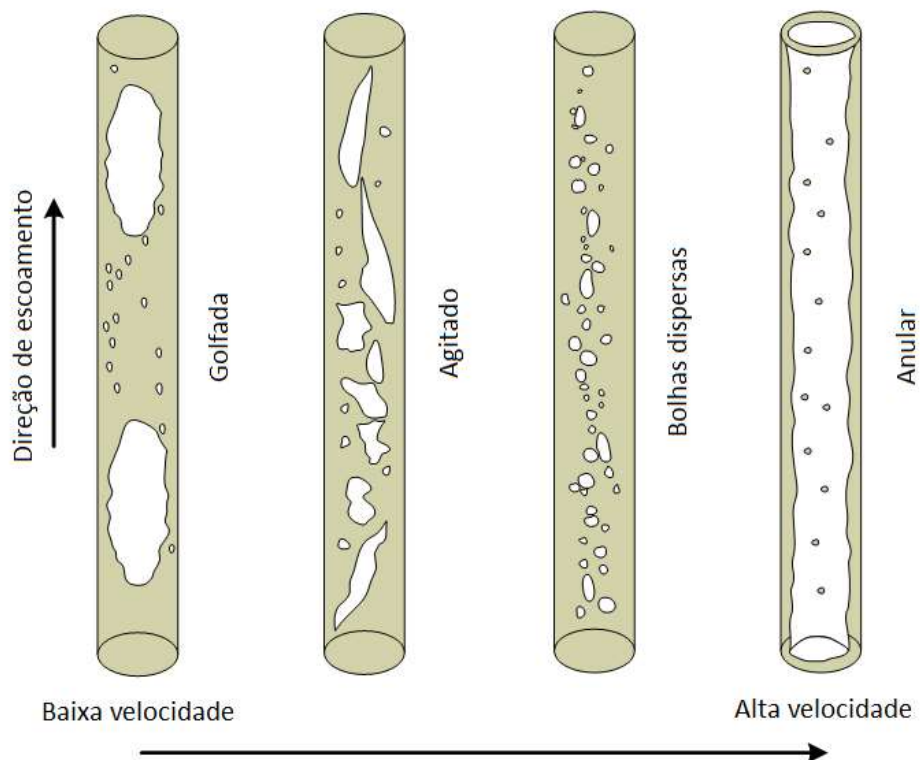


Figura 2.4: Tipos de escoamento vertical. Adaptado de Jahanshahi (2013).

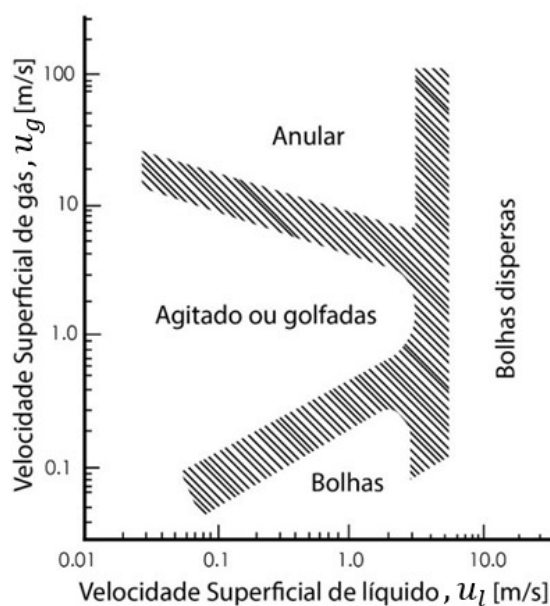


Figura 2.5: Mapa de regimes para escoamento vertical. Zonas tracejadas indicam regiões de transição. Adaptado de Weisman e Kang (1981).

2.3 Regime de golfadas

Dentre os padrões de escoamento multifásicos observados, o regime de golfadas é um grande obstáculo para a produção de petróleo. Seu surgimento decorre de uma distribuição não-homogênea de gás e líquido em tubulações de transporte de petróleo, podendo causar significativas perdas na produção e, conseqüentemente, no lucro obtido. (DI MEGLIO *et al.*, 2012)

Sob regime de golfadas, o sistema alterna entre períodos de alta produção de gás e baixa produção de líquido e períodos com baixa produção de gás e alta produção de líquido. Existem seis tipos distintos de mecanismos para golfadas em sistemas de elevação (JAHANSHAHI, 2013), três dos quais têm recebido maior interesse na área de modelagem: golfadas por acidente, golfadas por *casing-heading* e golfadas por ondas de densidade.

2.3.1 Golfadas por acidente

Esse tipo de mecanismo ocorre quando acidentes na disposição espacial da tubulação favorecem o acúmulo de líquido em um ponto mínimo ou acidente. Seu surgimento é mais frequente no limite entre *pipeline* e *riser*. O mecanismo pode ser descrito por um ciclo em quatro etapas, ilustrados pela Figura 2.6 (STORKAAS, 2005):

1. Líquido acumula em um ponto de mínimo, bloqueando a passagem de gás;
2. Gás acumula à montante do acidente em uma bolha;
3. A pressão da bolha vence a pressão à jusante, gás começa a fluir e líquido é abruptamente expulso da tubulação;
4. A pressão de gás no acidente é insuficiente para manter o fluxo através do ponto de mínimo e líquido volta a acumular.

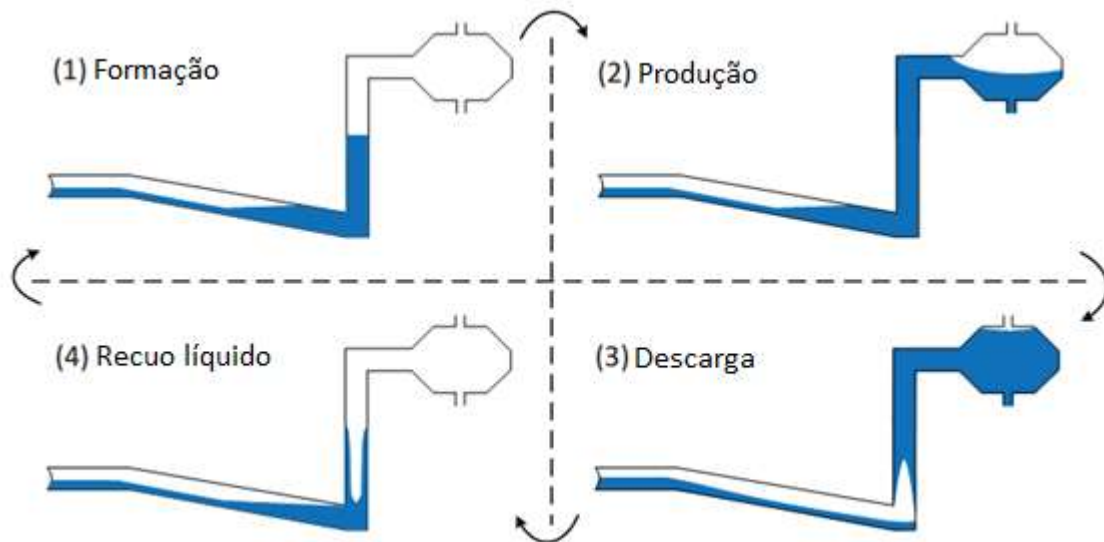


Figura 2.6: Exemplo de ciclo das golfadas por acidente em sistemas *pipeline-riser* sendo a fase líquida representada em azul. Adaptado de Storkaas (2005).

2.3.2 Golfadas por casing-heading

Ocorre em sistemas *gas-lift*, quando a válvula de injeção de gás opera fora da zona supersônica. Nessas condições, variações na diferença de pressão entre casco e *tubing* têm impacto direto sobre a vazão de *gas-lift* para o *tubing*. O surgimento das golfadas se dá quando a vazão de gás da superfície é insuficiente para manter a pressão do casco constante frente às flutuações de vazão do gás injetado por causa das variações de pressão no *tubing*. O processo pode ser descrito como um ciclo com as seguintes etapas (AAMO *et al.*, 2005):

1. Conforme o gás do casco anular flui para o *tubing*, a pressão do *tubing* diminui, acelerando a entrada de gás;
2. Gás empurra o líquido para fora do *tubing*, enquanto a pressão do casco diminui drasticamente;
3. A pressão do casco é insuficiente para permitir o fluxo de gás. Líquido começa a acumular no *tubing*, e gás, no casco;
4. Eventualmente, o acúmulo de gás no casco permite a reentrada de gás no *tubing*, e o ciclo recomeça.

2.3.3 Golfadas por ondas de densidade

Esse mecanismo ocorre em sistemas *gas-lift* quando a válvula de injeção opera em condições supersônicas. Nessa situação, a taxa de injeção de *gas-lift* depende apenas da pressão do casco, e não da diferença de pressão entre casco e *tubing*. Pode-se descrever o processo cíclico com as seguintes etapas (SINEGRE; PETIT; MENEGATTI, 2006):

1. A produção de óleo na superfície é nula, mas a pressão de fundo do *tubing* é menor do que a pressão do reservatório. Assim, há fluxo de óleo do reservatório para o sistema e a pressão de fundo do *tubing* aumenta;
2. A pressão de fundo do *tubing* se torna constante e a entrada de óleo cessa. Contudo, a vazão de *gas-lift* é constante, fazendo com que o óleo produzido na etapa 1 alcance a superfície.
3. A produção de óleo na etapa 2 causa uma queda de pressão no poço. A pressão de fundo do *tubing* diminui, permitindo novamente a entrada de óleo no sistema.

2.4 Ciclo limite e bifurcação Hopf

As perturbações causadas pelas golfadas se apresentam nos sistemas como um ciclo limite estável. Ciclo limite é uma solução periódica de um sistema dinâmico não linear quando para uma variável $x(t)$ existe um tempo \mathcal{T} tal que $x(t + \mathcal{T}) = x(t)$ para todo $t \in \mathfrak{R}$. Para modelos, se a variação de um parâmetro λ faz com que o sistema migre de um estado de equilíbrio para um ciclo limite a partir de um valor λ^* , então o sistema apresenta bifurcação Hopf, onde λ^* é o ponto de bifurcação. (SEYDEL, 2010)

Um exemplo de sistema com bifurcação Hopf é dado pelas Equações (2.3) e (2.4). Quando $\lambda < 0$, o sistema apresenta equilíbrio. A partir de $\lambda \geq 0$, o sistema sempre apresenta ciclo limite. As duas condições independem dos valores iniciais de y_1 e y_2 , caracterizando a bifurcação Hopf (Figura 2.7).

$$\dot{y}_1 = -y_2 + y_1(\lambda - y_1^2 - y_2^2) \quad (2.3)$$

$$\dot{y}_2 = y_1 + y_2(\lambda - y_1^2 - y_2^2) \quad (2.4)$$

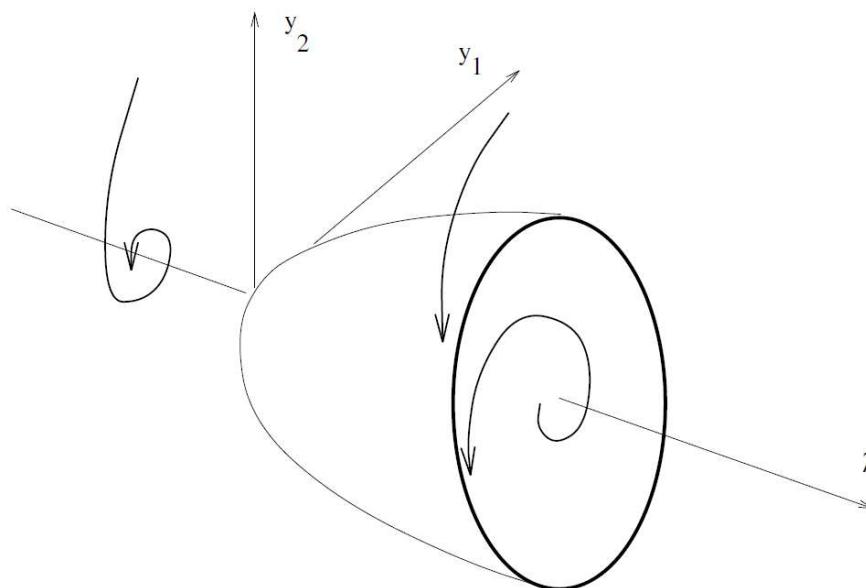


Figura 2.7: Bifurcação Hopf gerada a partir do exemplo. (SEYDEL, 2010)

2.5 Modelagem aplicada a automação e controle

Dadas as instabilidades descritas anteriormente, torna-se interessante o uso de métodos que atenuem as oscilações geradas pelo regime de golfadas. A solução mais simples consiste na diminuição da abertura da válvula de produção (SCHMIDT; BRILL; BEGGS, 1980, 1979). Isso elimina as oscilações, caracterizando um sistema com bifurcação Hopf: valores menores do que o valor crítico de abertura apresentam resposta amortecida, enquanto valores maiores apresentam oscilações em ciclo limite estável. Contudo, a produção total depende diretamente da abertura da válvula. Assim, o aumento de produção sem influência de golfadas se torna dependente do desenvolvimento de outros métodos.

Considerando-se que os sistemas de elevação estão submersos, a manutenção de qualquer equipamento representa um custo elevado de dinheiro e tempo. Além disso, o número de instrumentos e de medidas disponíveis é geralmente mínimo devido a restrições impostas ao projeto do sistema (PEDERSEN; DURDEVIC; YANG, 2017). Com isso em mente, muitos autores têm proposto soluções pouco custosas para aumento de produção nessas condições (BAILEY *et al.*, 2014; DI MEGLIO *et al.*, 2012; GODHAVN; FARD; FUCHS, 2005; HAVRE; DALSMO, 2002; JANSEN, F.E.; SHOHAM; TAITEL, 1996; SARICA; TENGESDAL, 2000; SKOGESTAD; JAHANSHAH; HANSEN, 2012; STORNES; HAVRE; STRAY, 2000).

Dentre os métodos estudados está o controle preditivo baseado em modelos. Esse tipo de medida anti-golfada utiliza medidas de pressão do sistema de elevação para manipular a válvula de produção. Contudo, a aplicação desse método é condicionada a modelos de baixa complexidade, que apresentem pouco custo computacional. Modelos como o OLGA (BENDIKSEN *et al.*, 1991), baseados em Equações Diferenciais Parciais (EDP), apresentam alta precisão ao custo de um grande tempo computacional que impossibilita sua aplicação *online*.

Já modelos baseados em Equações Diferenciais Ordinárias apresentam baixo custo computacional, gerando resultados numa pequena fração do tempo de simulação de modelos com EDP. Em contrapartida, esses modelos são baseados em simplificações do sistema que requerem parâmetros de ajuste de complexa determinação analítica e de difícil estimação numérica (AAMO *et al.*, 2005; DI MEGLIO; KAASA; PETIT, 2009; DIEHL *et al.*, 2017; EIKREM; AAMO; FOSS, 2008; JAHANSHAH, ESMAEIL; SKOGESTAD, 2011; SKOGESTAD; JAHANSHAH; HANSEN, 2012; STORKAAS, ESPEN, 2005).

2.6 Modelo de Di Meglio, Kaasa e Petit (2009)

Em seu trabalho, Di Meglio, Kaasa e Petit (2009), tendo em vista os modelos desenvolvidos até a data, propuseram um modelo simples que pudesse reproduzir as golfadas severas observadas na seção do *riser*. A grande contribuição dos autores está em utilizar fenômenos físicos em conjunto com um mecanismo artificial para a simulação das golfadas.

2.6.1 Descrição do modelo

Considere a seção ilustrada na Figura 2.8. É considerado que a geometria da tubulação favorece a obstrução e o acúmulo de gás. O fluxo de gás através dessa obstrução é modelado por uma válvula virtual. É suposto que à montante da válvula existe um volume fixo de gás, denominado de bolha alongada. Assumindo que a válvula não permite fluxo inverso, enquanto a pressão da bolha não excede a pressão hidrodinâmica do *riser*, não existe fluxo de gás através da válvula, gerando o acúmulo de gás que causa as instabilidades observadas.

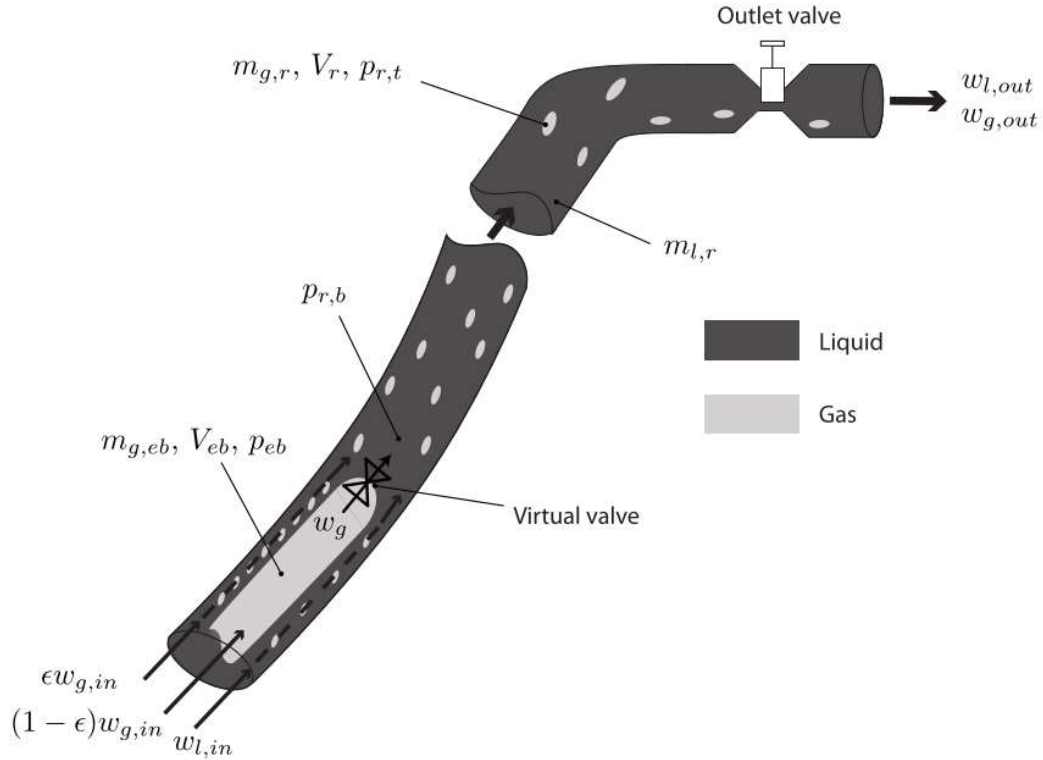


Figura 2.8: Representação do modelo de três estados. (DI MEGLIO *et al.*, 2012)

O modelo pressupõe então a existência de três seções de acúmulo distintas: dois volumes de gás separados pela válvula virtual e um volume de líquido. Assim, as variáveis de estado são a massa de gás na bolha alongada, $m_{g,eb}$, a massa de gás no *riser*, $m_{g,r}$, e a massa de líquido no *riser*, $m_{l,r}$. Do balanço de massa, obtém-se:

$$\dot{m}_{g,eb}(t) = (1 - \epsilon)w_{g,in} - w_g(t) \quad (2.5)$$

$$\dot{m}_{g,r}(t) = \epsilon w_{g,in} + w_g(t) - w_{g,out}(t) \quad (2.6)$$

$$\dot{m}_{l,r}(t) = w_{l,in} - w_{l,out}(t) \quad (2.7)$$

onde $w_{i,in}$ e $w_{i,out}$ são as taxas de vazão mássica entrando (*in*) e saindo (*out*) do *riser* para o líquido (*l*) e para o gás (*g*), w_g é a taxa de vazão mássica do gás através da válvula virtual e $\epsilon \in (0,1)$ corresponde à fração do gás que entra diretamente no *riser*, sendo o restante acumulado na bolha alongada de volume fixo V_{eb} . Dada a vazão mássica total w_{out} , as seguintes suposições são feitas:

$$w_{g,out} = \frac{m_{g,r}}{m_{g,r} + m_{l,r}} w_{out} \approx \frac{m_{g,r}}{m_{l,r}} w_{out} \quad (2.8)$$

$$w_{l,out} = \frac{m_{l,r}}{m_{g,r} + m_{l,r}} w_{out} \approx w_{out} \quad (2.9)$$

ou seja, considera-se que $m_{l,r} \gg m_{g,r}$. Originalmente, os autores propuseram que ambas as válvulas virtual e de produção fossem modeladas como válvulas lineares em relação às diferenças de pressão, sem permitir fluxo inverso:

$$w_g = C_g \max(p_{eb} - p_{r,b}, 0) \quad (2.10)$$

$$w_{out} = u C_{out} \max(p_{r,t} - p_s, 0) \quad (2.11)$$

onde p_{eb} é a pressão da bolha alongada, $p_{r,b}$ é a pressão na saída da válvula virtual, $u \in (0,1)$ é a abertura da válvula de produção, C_g é a constante da válvula virtual, C_{out} é a constante da válvula de produção e p_s é a pressão do separador na saída da válvula de produção. Contudo, em trabalhos recentes a válvula de produção foi modificada de modo a melhor representar o fluxo bifásico (DI MEGLIO *et al.*, 2010, 2012):

$$w_{out} = u C_{out} \sqrt{\max(p_{r,t} - p_{r,s}, 0)} \quad (2.12)$$

O volume de gás no *riser* $V_{g,r}$ é considerado como todo o volume que não é ocupado pelo líquido no *riser*:

$$V_{g,r} = V_r - \frac{m_{l,r} + m_{l,still}}{\rho_l} \quad (2.13)$$

$$V_r = LA \quad (2.14)$$

onde V_r é o volume do *riser*, L é o comprimento da tubulação, A é a área da seção transversal, ρ_l é a massa específica da fase líquida e $m_{l,still}$ representa a massa de líquido sempre presente no *riser* que não é afetada pela dinâmica do sistema. Para o cálculo das pressões, o sistema é considerado como dominado pelo efeito da gravidade, sendo a perda de carga desconsiderada. Além disso, assume-se que a Lei dos Gases Ideais é capaz de corretamente descrever o comportamento das fases gasosas:

$$p_{eb} = \frac{RT}{MV_{eb}} m_{g,eb} \quad (2.15)$$

$$p_{r,b} = p_{r,t} + \frac{g \sin \theta}{A} \quad (2.16)$$

$$p_{r,t} = \frac{RT}{MV_{g,r}} m_{g,r} \quad (2.17)$$

onde $p_{r,t}$ é a pressão no *riser*, R é a constante universal dos gases, T é a temperatura, M é a massa molar da fase gasosa, g é a aceleração da gravidade e θ é a inclinação do *riser*.

Assim, o modelo apresenta onze parâmetros fixos relacionados às características físico-químicas e geométricas do sistema: $L, \theta, A, g, \rho_l, R, T, M, p_s, w_{g,in}, w_{l,in}$; e cinco parâmetros de ajuste - $C_{out}, C_g, m_{l,still}, V_{eb}$ e ϵ . Procedimentos de obtenção dos parâmetros foram propostos (DI MEGLIO *et al.*, 2010; DI MEGLIO; KAASA; PETIT, 2009). Contudo, trabalhos que incluíram o modelo para fins de

comparação tentaram utilizar esses procedimentos e não obtiveram resultados satisfatórios, sendo que os parâmetros foram então ajustados por tentativa e erro (JAHANSHAH, E. *et al.*, 2014; JAHANSHAH, 2013).

2.6.2 Ajuste de parâmetros

Em seu trabalho, Di Meglio *et al.* (2010) descrevem métodos de obtenção dos parâmetros. Contudo, diversas considerações são feitas para a determinação analítica ou numérica dos parâmetros. Primeiramente, o modelo assume que as vazões de entrada são constantes. Isso facilita os cálculos das equações diferenciais, mas os autores notam que esses valores não são constantes nem são medidos nos sistemas reais. Assim, são utilizadas as próprias vazões médias de saída, obtidas com medidores multifásicos de vazão ou com simuladores como o OLGA.

Segundo os autores, é matematicamente difícil de se estudar os impactos da constante da válvula virtual C_g sobre a dinâmica do sistema, de modo que é sugerido um valor de 10^{-4} (kg.s.m^{-1}). Já a constante da válvula de produção C_{out} é geralmente fornecida pelo fabricante. Porém, a Equação (2.12) assume que a massa específica da mistura é constante, o que significa que a pressão de topo na condição de equilíbrio não representa o sistema para todos os valores de abertura.

A partir das vazões de entrada, os autores propõem que a constante de válvula da saída, C_{out} , seja calculada em uma condição de equilíbrio utilizando-se uma pressão de topo medida $\bar{P}_{r,t}$ para uma única abertura \bar{u} tal que $\bar{u} < u^*$, onde u^* é o ponto de bifurcação Hopf do sistema:

$$C_{out} = \frac{w_{l,in}}{\bar{u} \sqrt{\bar{P}_{r,t}(\bar{u}) - P_s}} \quad (2.18)$$

O parâmetro $m_{l,still}$ é obtido para uma medida real da pressão de fundo $\bar{P}_{r,b}$ para a mesma abertura \bar{u} :

$$\bar{P}_{r,b}(\bar{u}) = \bar{P}_{r,t}(\bar{u}) + (\bar{m}_{l,r}(\bar{u}) + m_{l,still}) \frac{g \sin \theta}{A} \quad (2.19)$$

$$m_{l,still} = \frac{\bar{P}_{r,b} - \bar{P}_{r,t}}{g/A \sin \theta} \left(1 + \frac{\bar{P}_{r,t}}{b\bar{x}} \right) - \frac{\bar{P}_{r,t}}{b\bar{x}} V_r \rho_l \quad (2.20)$$

$$\bar{x} = \frac{w_{g,in}}{w_{l,in}} \quad (2.21)$$

$$b = \frac{\rho_l RT}{M} \quad (2.22)$$

O volume da bolha alongada V_{eb} é determinado a partir da matriz Jacobiana do sistema na condição crítica, $J(u^*)$. Com o auxílio de um algoritmo de otimização, o valor de V_{eb} deve satisfazer a seguinte equação:

$$\Re\{\lambda(u^*, V_{eb})\} = 0 \quad (2.23)$$

onde λ é um dos dois autovalores complexos conjugados de $J(u^*)$. Por fim, a obtenção do parâmetro ϵ depende da disponibilidade de medidas da pressão de fundo do riser. Quando essa medida existe, é possível calcular o valor do parâmetro com o pico de pressão, $P_{r,b}^{max}$:

$$\epsilon = \frac{\bar{P}_{r,t} \rho_l V_r - (P_{r,b}^{max} - \bar{P}_{r,t}) \frac{A}{g \sin \theta}}{b \bar{x} (P_{r,b}^{max} - \bar{P}_{r,t}) \frac{A}{g \sin \theta} - m_{l,still}} \quad (2.24)$$

Sem essa medida, o valor deve ser obtido por tentativa e erro de forma a ajustar a magnitude das oscilações da pressão de topo.

2.7 Modelo de Diehl *et al.* (2017)

De forma a modelar o sistema de elevação completo, ilustrado na Figura 2.9, Diehl *et al.* (2017) uniram e aprimoraram os trabalhos de Jahanshahi *et al.* (2012) e Di Meglio *et al.* (2012). O *Fast Offshore Wells Model* (FOWM) inclui as dinâmicas de reservatório, *gas-lift*, poço e *riser*.

2.7.1 Descrição do modelo

Primeiramente, considera-se que a principal força-motriz do processo é a pressão do reservatório, P_r , assumida constante. A taxa mássica de produção do reservatório, w_r , é dada pelo modelo não-linear proposto por Vogel (1968):

$$w_r = K_r \cdot \left(1 - \left(0.2 \cdot \frac{P_{bh}}{P_r} \right) - \left(0.8 \cdot \frac{P_{bh}}{P_r} \right)^2 \right) \quad (2.25)$$

onde K_r é o coeficiente de fluxo do reservatório e P_{bh} é a pressão no fundo do poço (*bottomhole*). Do reservatório em diante, as dinâmicas do sistema são modeladas por balanço de massa. Considera-se que a mistura produzida pelo reservatório é composta de uma fase gás e uma fase líquida cuja fração gasosa é dada por $\alpha_{g,w}$. As variações de massa das fases no poço são então equacionadas por:

$$\dot{m}_{g,t} = w_r \cdot \alpha_{g,w} + w_{iv} - w_{g,wh} \quad (2.26)$$

$$\dot{m}_{l,t} = w_r (1 - \alpha_{g,w}) - w_{l,wh} \quad (2.27)$$

onde $\dot{m}_{g,t}$ e $\dot{m}_{l,t}$ são as taxas de variação mássica no *tubing* do gás e do líquido, w_{iv} é a taxa de vazão mássica de *gas-lift* do casco anular para o *tubing* e $w_{l,wh}$ é a taxa de vazão mássica do topo do poço (*wellhead*) para o *flowline*. É suposto que o casco anular recebe uma vazão mássica de *gas-lift* constante, w_{gc} . A variação mássica no casco é equacionada por:

$$\dot{m}_{g,a} = w_{gc} - w_{iv} \quad (2.28)$$

Por fim, a seção *flowline-riser* apresenta os três balanços de massa de Di Meglio *et al.*:

$$\dot{m}_{g,eb} = (1 - \epsilon) w_{g,wh} - w_g \quad (2.29)$$

$$\dot{m}_{g,r} = \epsilon w_{g,wh} + w_g - w_{g,out} \quad (2.30)$$

$$\dot{m}_{l,r} = w_{l,wh} - w_{l,out} \quad (2.31)$$

As taxas de vazão mássica remanescentes são dadas por:

$$w_{iv} = K_a \sqrt{\rho_{ai} \max(P_{ai} - P_{t,b}, 0)} \quad (2.32)$$

$$w_{g,wh} = \alpha_{g,t} K_w \sqrt{\rho_l \max(P_{t,t} - P_{r,b}, 0)} \quad (2.33)$$

$$w_{l,wh} = (1 - \alpha_{g,t}) K_w \sqrt{\rho_l \max(P_{t,t} - P_{r,b}, 0)} \quad (2.34)$$

$$w_g = C_g \max(P_{eb} - P_{r,b}, 0) \quad (2.35)$$

$$w_{g,out} = \alpha_{g,r} C_{out} z \sqrt{\rho_l \max(P_{r,t} - P_s, 0)} \quad (2.36)$$

$$w_{l,out} = (1 - \alpha_{g,r}) C_{out} z \sqrt{\rho_l \max(P_{r,t} - P_s, 0)} \quad (2.37)$$

onde w_g e $w_{l,out}$ são as taxas mássicas da válvula virtual e da válvula de produção, respectivamente, ρ_{ai} é a massa específica no casco anular, ρ_l é a massa específica média do líquido, considerada constante, $\alpha_{g,t}$ é a fração de gás no *tubing* e $\alpha_{g,r}$ é a fração de gás no *riser*.

Considera-se que a pressão de saída do *manifold* é a própria pressão na base do *riser*, ou seja, as dinâmicas do *flowline* são desconsideradas. K_a , K_w , C_g e C_{out} são constantes das equações de válvula do sistema, enquanto z é a fração de abertura da válvula de produção. Assume-se que não ocorre fluxo inverso em nenhum momento, o que é equacionado pela função *max*. A partir da Lei dos Gases ideais, são obtidas as equações:

$$\rho_{ai} = \frac{MP_{ai}}{RT} \quad (2.38)$$

$$\alpha_{g,t} = \frac{m_{g,t}}{m_{g,t} + m_{l,t}} \quad (2.39)$$

$$\alpha_{g,r} = \frac{m_{g,r}}{m_{g,r} + m_{l,r}} \quad (2.40)$$

$$P_{ai} = \left(\frac{RT}{V_a M} + \frac{gL_a}{V_a} \right) m_{g,a} \quad (2.41)$$

$$P_{t,b} = P_{t,t} + \rho_{m,t} g H_{vgl} \quad (2.42)$$

$$P_{bh} = P_{pdg} + \rho_{m,res} g (H_t - H_{pdg}) \quad (2.43)$$

$$P_{pdg} = P_{t,b} + \rho_{m,res} g (H_{pdg} - H_{vgl}) \quad (2.44)$$

$$P_{t,t} = \frac{\rho_{g,t} RT}{M} \quad (2.45)$$

$$P_{r,b} = P_{r,t} + \frac{(m_{l,r} + m_{l,still}) g \sin \theta}{A_{ss}} \quad (2.46)$$

$$P_{eb} = \frac{m_{g,b}RT}{MV_{eb}} \quad (2.47)$$

$$P_{r,t} = \frac{m_{g,r}RT}{M\left(\omega_u V_{ss} - \frac{m_{l,r} + m_{l,still}}{\rho_l}\right)} \quad (2.48)$$

onde M é a massa molar do gás, R é a constante universal dos gases, T é a temperatura, g é a aceleração da gravidade, θ é a inclinação média do *riser*, V_a é o volume do casco, L_a é o comprimento do casco, P_{ai} é a pressão no casco, $P_{t,t}$ é a pressão de topo no *tubing*, $P_{t,b}$ é a pressão de fundo no *tubing*, P_{pdg} é a pressão na posição PDG (*permanent downhole gauge*), $P_{r,b}$ é a pressão de fundo no *riser* e $P_{r,t}$ é a pressão de topo no *riser*. A massa específica do reservatório $\rho_{m,res}$ é considerada constante.

As distâncias verticais a partir do *manifold* até a válvula de *gas-lift*, o transmissor da PDG e a pressão de fundo no poço são representadas respectivamente por H_{vgl} , H_{pdg} e H_t . Os parâmetros geométricos do sistema são dados por:

$$A_{ss} = \frac{\pi D_{ss}^2}{4} \quad (2.49)$$

$$V_{ss} = \frac{\pi D_{ss}^2 L_r}{4} + \frac{\pi D_{ss}^2 L_{fl}}{4} \quad (2.50)$$

$$V_a = \frac{\pi D_a^2 L_a}{4} \quad (2.51)$$

$$V_t = \frac{\pi D_t^2 L_t}{4} \quad (2.52)$$

onde D_{ss} , D_a e D_t são respectivamente os diâmetros da tubulação *subsea*, do casco e do *tubing*, A_{ss} e V_{ss} são a área da seção transversal e o volume da tubulação *subsea* e V_t é o volume do *tubing*.

As massas específicas de gás e da mistura no *tubing*, $\rho_{g,t}$ e $\rho_{m,t}$, assim como o volume de gás do *tubing* $V_{g,t}$ são dados por:

$$\rho_{g,t} = \frac{m_{g,t}}{V_{g,t}} \quad (2.53)$$

$$\rho_{m,t} = \frac{m_{g,t} + m_{l,t}}{V_t} \quad (2.54)$$

$$V_{g,t} = V_t - \frac{m_{l,t}}{\rho_l} \quad (2.55)$$

Assim, o modelo FOWM utiliza nove parâmetros de ajuste: os cinco parâmetros do modelo de Di Meglio, Kaasa e Petit ($m_{l,still}$, C_g , V_{eb} , C_{out} e ϵ) e quatro parâmetros adicionais (K_r , K_a , K_w e ω_u). O parâmetro $\omega_u \in [1, \infty)$ é utilizado para ajuste do volume da tubulação *subsea*. Como o sistema é baseado em parâmetros concentrados, esse parâmetro compensaria essa simplificação de modo a melhor ajustar os resultados do modelo. Contudo, os autores recomendam o ajuste desse parâmetro como último recurso, quando a otimização não apresenta resultados satisfatórios.

2.7.2 Ajuste de parâmetros

Para que o modelo consiga descrever o comportamento de diferentes sistemas, os algoritmos de otimização global *Direct* (FINKEL, 2005) e *Particle Swarm Optimization* (POLI; KENNEDY; BLACKWELL, 2007) são empregados. Como os sistemas reais apresentam bifurcação Hopf, uma função objetivo baseada em mínimos quadrados foi proposta de forma a penalizar condições de equilíbrio em favor de situações de ciclo limite. O conjunto de dados é dividido em M janelas com N pontos:

$$J_s = \sum_{j=1}^M \frac{1}{\omega_j} \left[\frac{1}{N} \sum_{i=1}^N \omega_i (y_{j,i} - x_{j,i})^2 \right] \quad (2.56)$$

onde J_s é a função objetivo, y é o valor medido de pressão e x é o valor simulado de pressão. Dado que a otimização do problema costuma atingir um mínimo local em que o sistema está em equilíbrio, os parâmetros ω_i e ω_j foram adicionados como um peso para cada ponto amostral i e para cada janela j :

$$\omega_i = \left(\frac{x_{j,i} - \bar{x}_j}{\sigma_{j,x}} \right)^2 \quad (2.57)$$

$$\omega_j = \frac{\rho_{y,x} + 1}{2} \quad (2.58)$$

onde \bar{x}_j é a média dos dados simulados na janela j , $\sigma_{j,x}$ é o desvio padrão dos dados simulados na janela j e $\rho_{y,x}$ é o coeficiente de correlação de Pearson entre dados reais e simulados na janela j . Assim, ω_j é um coeficiente normalizado, de modo que $\omega_j = 0$, $\omega_j = 0,5$ e $\omega_j = 1$ representam perfeita correlação negativa, nenhuma correlação e perfeita correlação positiva, respectivamente (DIEHL *et al.*, 2017).

Contudo, a hipersuperfície de otimização é complexa. Os autores notam que mesmo um algoritmo global pode ser incapaz de encontrar uma solução adequada sem um ajuste cuidadoso dos limites de estimação dos parâmetros. Por isso, também é sugerida uma metodologia de ajuste dos parâmetros com base em dados de projeto, dados de operação ou dados de simuladores rigorosos. Os parâmetros K_r , K_a e K_w podem ser diretamente calculados com as Equações (2.59) a (2.61), sabendo-se as vazões mássicas w_r , w_{iv} e w_{wh} :

$$K_r = \frac{w_r}{1 - \left(0,2 \frac{P_{bh}}{P_r}\right) - \left(0,8 \frac{P_{bh}}{P_r}\right)^2} \quad (2.59)$$

$$K_a = \frac{w_{iv}}{\sqrt{\rho_{ai}(P_{ai} - P_{t,b})}} \quad (2.60)$$

$$K_w = \frac{w_{wh}}{\sqrt{\rho_l(P_{t,t} - P_{r,b})}} \quad (2.61)$$

A massa de líquido $m_{l,still}$ pode ser determinada pelo método proposto por Di Meglio *et al.* (2010). O parâmetro C_g é ajustado baseado em experiência com o modelo e requer um processo de tentativa e erro, embora $10^{-4} \text{ kg s m}^{-1}$ seja considerado um valor inicial adequado para qualquer problema. A constante de produção C_{out} pode ser obtida através das informações fornecidas pelo fabricante ou através da Equação (2.62):

$$C_{out} = \frac{1}{z \sqrt{\rho_l (P_{r,t} - P_s)}} \quad (2.62)$$

Já os demais parâmetros requerem dados de um simulador rigoroso. As informações de interesse nesse caso são as vazões mássicas de gás no *manifold* e na saída da válvula de produção. A Figura 2.10 ilustra o comportamento desses dados, obtidos com o simulador OLGA. Os tempos t_1 e t_2 caracterizam os inícios das fases de formação e produção da golfada de gás, respectivamente (FABRE *et al.*, 1990). A fração de gás que entra pelo meio líquido pode ser então calculada por:

$$\epsilon = \frac{\int_{t_1}^{t_2} w_{g,out} dt}{\int_{t_1}^{t_2} w_{g,wh} dt + \int_{t_1}^{t_2} w_{g,out} dt} \quad (2.63)$$

O gás acumulado é considerado como a massa de gás da bolha alongada. Assumindo-se uma massa específica média da bolha, $\bar{\rho}_g$, o volume da bolha é dado por:

$$V_{eb} = \frac{\int_{t_1}^{t_2} w_{g,wh} dt - \int_{t_1}^{t_2} w_{g,out} dt}{\bar{\rho}_g} \quad (2.64)$$

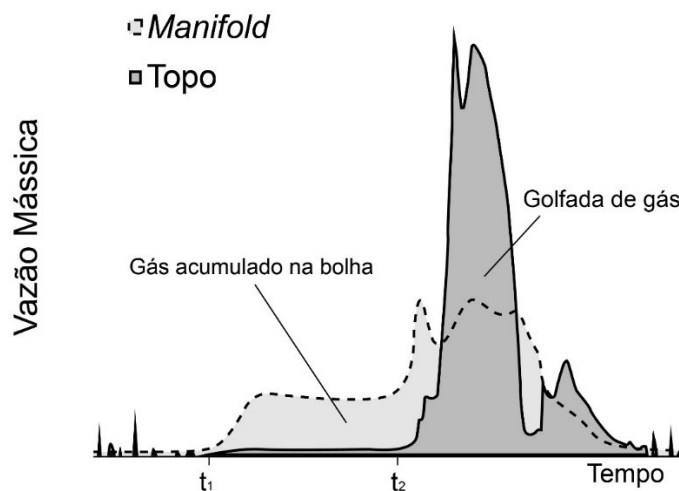


Figura 2.10: Ciclo completo de uma golfada obtido com o simulador OLGA. Adaptado de Diehl *et al.* (2017).

2.8 Estimação de parâmetros para modelos com bifurcação Hopf

Para melhor compreender as metodologias disponíveis para estimação de parâmetros de modelos com ciclo limite, é interessante destacar a importância do Método de Mínimos Quadrados. Esse método baseia a função objetivo do problema de otimização em duas hipóteses. A primeira é a Hipótese do Modelo Perfeito. Considere um modelo y^m dado pela função f com x estados e θ parâmetros:

$$y^m = f(x; \theta) \quad (2.65)$$

Assumindo que o modelo descreve exatamente as relações entre as variáveis do problema, o modelo é dito perfeito. Isso significa que as medidas experimentais y^e apresentam desvios em relação ao modelo somente devido aos erros experimentais ε :

$$y^e = f(x^e + \varepsilon_x; \theta + \varepsilon_\theta) + \varepsilon_y. \quad (2.66)$$

A segunda hipótese é a Hipótese do Experimento Bem Feito. Se os erros de medição durante os experimentos equivalem ao mínimo erro admissível no processo de medição, os experimentos são ditos bem feitos. Assim, os erros experimentais não devem apresentar tendência ou polarização (média zero). Em decorrência disso, é possível avaliar a variância dos erros experimentais, σ_y^2 , na forma:

$$\sigma_y^2 = \frac{\sum_{i=1}^{NE} (y_i^e - y_i^m)^2}{\nu} \quad (2.67)$$

onde ν é o número de graus de liberdade e NE é o número de experimentos. Dado que a variância σ_y^2 é uma métrica dos erros experimentais, torna-se interessante minimizá-la. Como ν é constante, é possível definir a função objetivo F_{obj} como:

$$F_{obj} = \sum_{i=1}^{NE} (y_i^e - y_i^m)^2 \quad (2.68)$$

Minimizar a Equação (2.68) corresponde ao Método de Mínimos Quadrados. Sua importância se dá pela interpretação estatística associada à função objetivo relativa à variância dos erros experimentais. Contudo, sua aplicação em modelos com bifurcação Hopf não é capaz de produzir resultados que consigam representar a periodicidade dos sistemas (APIO, 2017).

Na literatura, metodologias têm sido desenvolvidas para estimação de parâmetros de modelos com ciclo limite baseada no ponto de bifurcação (APIO; BOTELHO; TRIERWEILER, 2018; CEDERSUND; KNUDSEN, 2005; ELMAN *et al.*, 2012; GUCKENHEIMER; MYERS; STURMFELS, 1997; HYNNE; DANØ; SØRENSEN, 2001; HYNNE; GRAAE SØRENSEN; MØLLER, 1993; MEERBERGEN; SPENCE, 2010). Dentre essas, Apio (2017) propôs uma função objetivo baseada em mínimos quadrados com a soma de uma parcela relacionada ao ponto de bifurcação:

$$F_{obj} = \sum_{i=1}^{NE} (y_i^e - y_i^m(\boldsymbol{\theta}))^2 + \varphi \cdot \zeta(\boldsymbol{\theta}) \quad (2.69)$$

onde φ é um peso cujo valor decai com o número de iterações da otimização, $\zeta(\boldsymbol{\theta})$ é o fator de amortecimento associado ao modelo linearizado no seu estado estacionário e $\boldsymbol{\theta}$ é o conjunto de parâmetros estimado.

Contudo, métodos baseados no ponto de bifurcação requerem a determinação exata desse ponto. Apesar de apresentarem resultados teóricos eficientes, esses métodos também apresentam dificuldades de aplicação prática. A localização exata do ponto de bifurcação nem sempre é conhecida e pode mudar conforme as condições do sistema.

Com isso em mente, alguns autores têm proposto modificações à função objetivo como alternativa para melhorar a performance de algoritmos de otimização em sistemas com ciclo limite (DIEHL *et al.*, 2017; GEREVINI, 2017; HÜFFNER, 2018). Diehl *et al.* (2017) propuseram uma função objetivo baseada em mínimos quadrados que divide o conjunto de dados em M janelas e utiliza pesos ω_i e ω_j para melhor ajustar dados em ciclo limite:

$$J_s = \sum_{j=1}^M \frac{1}{\omega_j} \left[\frac{1}{N} \sum_{i=1}^{NE} \omega_i (y_{j,i}^e - y_{j,i}^m)^2 \right] \quad (2.70)$$

$$\omega_i = \left(\frac{y_{j,i}^m - \bar{y}_j^m}{\sigma_{j,x}} \right)^2 \quad (2.71)$$

$$\omega_j = \frac{\rho_{y,x} + 1}{2} \quad (2.72)$$

onde \bar{y}_j^m e $\sigma_{j,x}$ são respectivamente a média e o desvio padrão das medidas simuladas da janela j , e $\rho_{y,x}$ é o coeficiente de correlação de Pearson entre as medidas e o modelo. Gerevini (2017) propôs uma função objetivo baseada em mínimos quadrados com a soma de uma parcela relacionada ao período e à derivada em cada ponto experimental:

$$F_{obj} = \sum_{i=1}^{NE} (y_i^e - y_i^m(\boldsymbol{\theta}))^2 + \left[\frac{1}{NE} \sum_{i=1}^{NE} \left(\frac{dy_i^e}{dt} - \frac{dy_i^m(\boldsymbol{\theta})}{dt} \right)^2 \right] (\mathcal{T}^e - \mathcal{T}^m(\boldsymbol{\theta}))^2 \quad (2.73)$$

onde \mathcal{T} é o período. As derivadas são calculadas numericamente através de diferenças finitas:

$$\frac{dy_i}{dt} = \frac{y_{i+1} - y_i}{\Delta t} \quad (2.74)$$

Hüffner (2018) propôs uma função objetivo baseada na soma de diferenças quadráticas relacionadas ao período, ao número de oscilações e à média:

$$F_{obj} = [(\mathcal{T}^e - \mathcal{T}^m)^2 + (\aleph^e - \aleph^m)^2 + (\bar{y}^e - \bar{y}^m)^2] \quad (2.75)$$

onde \mathcal{T} é o período, \aleph é o número de oscilações e \bar{y} é a média das medidas de pressão.

Capítulo 3 – Fundamentos de avaliação estatística para modelos

Para se quantificar a qualidade dos resultados obtidos, é necessário o uso de recursos estatísticos. As seções a seguir descrevem a metodologia adotada para fazer essa avaliação dos resultados. Todos os métodos foram baseados no livro de Schwaab e Pinto (2007).

3.1 Qualidade dos modelos

Para comprovar a qualidade representativa dos modelos em relação aos dados de referência utilizados, dois critérios estatísticos são empregados. O primeiro critério é um teste Chi Quadrado com o valor ótimo da função objetivo:

$$\chi_{min}^2 < F_{obj}(\hat{\boldsymbol{\theta}}_{opt}) < \chi_{max}^2 \quad (3.1)$$

$$\chi_{min}^2 = \chi^2\left(\frac{1-\alpha}{2}, \nu\right) \quad (3.2)$$

$$\chi_{max}^2 = \chi^2\left(\frac{1+\alpha}{2}, \nu\right) \quad (3.3)$$

onde F_{obj} é a função objetivo de Mínimos Quadrados Ponderados, $\hat{\boldsymbol{\theta}}_{opt}$ é o conjunto de parâmetros obtidos na estimação, χ^2 é a função Chi Quadrado, α é o grau de confiança e ν é o número de graus de liberdade, definido pela diferença entre o número de experimentos e o número de parâmetros. Caso a desigualdade seja satisfeita, o modelo representa suficientemente bem os dados experimentais. Caso contrário, duas situações podem ocorrer:

- $F_{obj}(\hat{\boldsymbol{\theta}}_{opt}) < \chi_{min}^2$: o modelo reproduz os dados experimentais melhor do que o esperado;
- $\chi_{max}^2 < F_{obj}(\hat{\boldsymbol{\theta}}_{opt})$: o modelo não explica os erros experimentais de forma satisfatória.

A primeira situação indica que há algo errado, porque um modelo não pode levar a previsões melhores do que os dados usados para gerá-lo. Portanto, o modelo está provavelmente superparametrizado e pode ser simplificado. Também é possível que os erros experimentais estejam superestimados, abrindo margem para se reavaliar a precisão das informações de referência. Já a segunda situação indica que os erros de predição são significativamente maiores do que os erros experimentais, indicando que o modelo pode ser aperfeiçoado. Analogamente à primeira situação, é possível que os erros experimentais estejam subestimados, também sugerindo a reavaliação dos dados de referência.

O segundo critério é definido a partir de um coeficiente de correlação entre os valores experimentais e os valores calculados, $\rho^m \in [0,1]$, definido como:

$$\rho^m = \frac{\sum_{i=1}^{NE} (z_i^e - \bar{z}_i^e) (z_i^m - \bar{z}_i^m)}{\sqrt{\left[\sum_{i=1}^{NE} (z_i^e - \bar{z}_i^e)^2 \right] \left[\sum_{i=1}^{NE} (z_i^m - \bar{z}_i^m)^2 \right]}} \quad (3.4)$$

onde NE é o número de experimentos, z^e são as amostras do ciclo obtido para os dados experimentais, z^m são as amostras do ciclo obtido para os dados simulados e \bar{z} é o valor médio. Esse critério quantifica o quanto os dados simulados acompanham a tendência dos dados experimentais. Se o valor obtido é superior a 0,9, o modelo é satisfatório e os valores preditos variam linear e proporcionalmente com os valores experimentais. Assim, um modelo representativo e um conjunto de dados de boa qualidade satisfazem os dois critérios.

3.2 Qualidade dos parâmetros

Para se determinar a qualidade dos parâmetros, é necessário que se calcule a matriz de covariância dos parâmetros, V_θ . Contudo, soluções analíticas para esse cálculo só existem para modelos lineares. Dada a não-linearidade dos modelos estudados, é conveniente que se obtenha a matriz V_θ de forma numérica. Considerando a ausência de erros de medida nas variáveis independentes, é possível definir a matriz V_θ a partir da matriz Hessiana H_θ para NP parâmetros:

$$V_\theta = 2H_\theta^{-1} = \begin{bmatrix} \sigma_{\theta_1}^2 & \sigma_{\theta_1, \theta_2} & \cdots & \sigma_{\theta_1, \theta_{NP}} \\ \sigma_{\theta_2, \theta_1} & \sigma_{\theta_2}^2 & \cdots & \sigma_{\theta_2, \theta_{NP}} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{\theta_{NP}, \theta_1} & \sigma_{\theta_{NP}, \theta_2} & \cdots & \sigma_{\theta_{NP}}^2 \end{bmatrix} \quad (3.5)$$

$$H_\theta = \begin{bmatrix} \frac{\partial^2 F_{obj}}{\partial \theta_1 \partial \theta_1} & \frac{\partial^2 F_{obj}}{\partial \theta_1 \partial \theta_2} & \cdots & \frac{\partial^2 F_{obj}}{\partial \theta_1 \partial \theta_{NP}} \\ \frac{\partial^2 F_{obj}}{\partial \theta_2 \partial \theta_1} & \frac{\partial^2 F_{obj}}{\partial \theta_2 \partial \theta_2} & \cdots & \frac{\partial^2 F_{obj}}{\partial \theta_2 \partial \theta_{NP}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 F_{obj}}{\partial \theta_{NP} \partial \theta_1} & \frac{\partial^2 F_{obj}}{\partial \theta_{NP} \partial \theta_2} & \cdots & \frac{\partial^2 F_{obj}}{\partial \theta_{NP} \partial \theta_{NP}} \end{bmatrix} \quad (3.6)$$

Definindo-se a função objetivo como:

$$F_{obj} = (\mathbf{z}^e - \mathbf{z}^m)^T V_z^{-1} (\mathbf{z}^e - \mathbf{z}^m) \quad (3.7)$$

$$\mathbf{V}_z = \begin{bmatrix} \sigma_{z_1}^2 & 0 & \cdots & 0 \\ 0 & \sigma_{z_2}^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_{z_{NZ}}^2 \end{bmatrix} \quad (3.8)$$

É possível reescrever os elementos $[h_{r,s}]$ da matriz Hessiana \mathbf{H}_θ como:

$$h_{r,s} = \frac{\partial^2 F_{obj}}{\partial \theta_r \partial \theta_s} = 2 \left(\frac{\partial \mathbf{z}^m}{\partial \theta_r} \right)^T \mathbf{V}_z^{-1} \left(\frac{\partial \mathbf{z}^m}{\partial \theta_r} \right) - 2 \left(\frac{\partial^2 \mathbf{z}^m}{\partial \theta_r \partial \theta_s} \right)^T \mathbf{V}_z^{-1} (\mathbf{z}^e - \mathbf{z}^m) \quad (3.9)$$

Supondo que as diferenças entres os valores preditos e os valores observados são pequenas e que flutuam ao redor do zero de forma aleatória (hipótese do modelo perfeito e hipótese do experimento bem feito), o termo $(\mathbf{z}^e - \mathbf{z}^m)$ pode ser desconsiderado. Assim, a matriz Hessiana pode ser calculada através da matriz de sensibilidade \mathbf{B} :

$$\mathbf{H}_\theta \cong 2\mathbf{B}^T \mathbf{V}_z^{-1} \mathbf{B} \quad (3.10)$$

$$\mathbf{B} = \begin{bmatrix} \frac{\partial z_1^m}{\partial \theta_1} & \frac{\partial z_1^m}{\partial \theta_2} & \cdots & \frac{\partial z_1^m}{\partial \theta_{NP}} \\ \frac{\partial z_2^m}{\partial \theta_1} & \frac{\partial z_2^m}{\partial \theta_2} & \cdots & \frac{\partial z_2^m}{\partial \theta_{NP}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial z_{NZ}^m}{\partial \theta_1} & \frac{\partial z_{NZ}^m}{\partial \theta_2} & \cdots & \frac{\partial z_{NZ}^m}{\partial \theta_{NP}} \end{bmatrix} \quad (3.11)$$

Esse resultado é chamado de Aproximação de Gauss e sua importância surge pelo fato de que a matriz \mathbf{V}_θ pode então ser calculada utilizando-se apenas derivadas primeiras dos valores preditos em relação aos parâmetros:

$$\mathbf{V}_\theta \cong [\mathbf{B}^T \mathbf{V}_z^{-1} \mathbf{B}]^{-1} \quad (3.12)$$

As derivadas podem ser obtidas numericamente através de diferenças finitas centrais para um passo h com um erro de ordem $\varphi(h^2)$:

$$\frac{df(x)}{dx} = \frac{f(x+h) - f(x-h)}{h^2} + \varphi(h^2) \quad (3.13)$$

Assim, obtém-se uma matriz que contém as variâncias de cada parâmetro bem como a covariância entre cada par de parâmetros. Para facilitar a interpretação das covariâncias, é interessante definir o coeficiente de correlação paramétrica entre um par de parâmetros i e j :

$$\rho_{ij} = \frac{\sigma_{ij}}{\sqrt{\sigma_i^2 \sigma_j^2}} \quad (3.14)$$

Essa definição leva a uma matriz de correlações \mathbf{R}_θ da forma:

$$\mathbf{R}_\theta = \begin{bmatrix} 1 & \rho_{1,2} & \cdots & \rho_{1,NP} \\ \rho_{2,1} & 1 & \cdots & \rho_{2,NP} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{NP,1} & \rho_{NP,2} & \cdots & 1 \end{bmatrix} \quad (3.15)$$

Essa matriz contém todas as correlações entre todos os pares de parâmetros e sua interpretação é facilitada, visto que os coeficientes de correlação sempre apresentam valores entre -1 e 1 . Valores positivos ou negativos indicam correlações positivas ou negativas, ou seja, o aumento de um parâmetro causa um aumento ou uma diminuição em seu par em questão, respectivamente.

Valores próximos de zero aumentam a eficiência dos procedimentos de estimação. Por outro lado, valores absolutos de correlação acima de 0.9 indicam correlação alta. O resultado disso é que os parâmetros não podem ser obtidos independentemente e que, ao invés de um intervalo de confiança, torna-se necessária uma região de confiança para avaliar a faixa de validade dos parâmetros.

3.2.1 Intervalos de confiança

Define-se como intervalo de confiança o conjunto de resultados que concentra $\alpha\%$ dos resultados possíveis, dada uma distribuição de probabilidades. Em outras palavras, intervalos de confiança são faixas de valores dentro das quais qualquer valor do parâmetro em questão mantém a qualidade do ajuste do modelo aos dados experimentais com grau de confiança α . Se os erros são normalmente distribuídos, o intervalo de confiança toma a forma:

$$\hat{\theta} + \sigma_\theta u \left| \frac{1-\alpha}{2} \right. < \theta < \hat{\theta} + \sigma_\theta u \left| \frac{1+\alpha}{2} \right. \quad (3.16)$$

onde $\hat{\theta}$ é o parâmetro ótimo, σ_θ é o desvio padrão do parâmetro, u é o valor da distribuição normal para uma probabilidade $(1 \pm \alpha)/2$ e α é o grau de confiança utilizado. Qualquer parâmetro θ que satisfaça a desigualdade é capaz de manter a qualidade do modelo.

Porém, a Equação (3.16) assume que o erro dos parâmetros segue a distribuição normal, ou seja, muitas medidas experimentais estão disponíveis. Para um número pequeno de graus de liberdade, torna-se conveniente utilizar a distribuição t de Student. Isso causa o aumento do intervalo de confiança frente à limitação dos graus de liberdade, modificando a desigualdade:

$$\hat{\theta} + s_\theta t \left| \frac{1-\alpha}{2}, \nu \right. < \theta < \hat{\theta} + s_\theta t \left| \frac{1+\alpha}{2}, \nu \right. \quad (3.17)$$

Onde s_θ é a estimativa do desvio padrão do parâmetro e t é o valor da distribuição t para uma probabilidade $(1 \pm \alpha)/2$ com ν graus de liberdade.

3.2.2 Regiões de confiança

Quando existem altas correlações entre parâmetros, os intervalos de confiança não representam adequadamente a faixa de validade dos parâmetros. Ao invés de um intervalo, utiliza-se uma região dentro da qual os valores dos parâmetros mantêm a qualidade do modelo. Assumindo que a variância experimental é conhecida com precisão e que os parâmetros seguem a distribuição normal, é possível aproximar a função objetivo por uma série de Taylor truncada no termo quadrático em torno do ponto ótimo, obtendo-se duas relações em termos da distribuição Chi Quadrado:

$$(\boldsymbol{\theta} - \hat{\boldsymbol{\theta}})^T \mathbf{V}_{\boldsymbol{\theta}}^{-1} (\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}) \leq \chi_{NP, \alpha}^2 \quad (3.18)$$

$$F_{obj}(\boldsymbol{\theta}) - F_{obj}(\hat{\boldsymbol{\theta}}) \leq \chi_{NP, \alpha}^2 \quad (3.19)$$

onde NP é o número de parâmetros e $\chi_{NP, \alpha}^2$ é o valor da função Chi Quadrado com NP graus de liberdade e grau de confiança α .

Quando a variância experimental não é precisamente conhecida, é possível reescrever as relações anteriores em termos da distribuição F de Fisher:

$$(\boldsymbol{\theta} - \hat{\boldsymbol{\theta}})^T \mathbf{V}_{\boldsymbol{\theta}}^{-1} (\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}) \leq F_{obj}(\hat{\boldsymbol{\theta}}) \frac{NP}{N - NP} F_{NP, N - NP}^{\alpha} \quad (3.20)$$

$$F_{obj}(\boldsymbol{\theta}) \leq F_{obj}(\hat{\boldsymbol{\theta}}) \left(1 + \frac{NP}{N - NP} F_{NP, N - NP}^{\alpha} \right) \quad (3.21)$$

onde N é o número de amostras e $F_{NP, N - NP}^{\alpha}$ é o valor da função F com NP e $N - NP$ graus de liberdade para um grau de confiança α . Esse formato leva a regiões de confiança maiores, mas isso representa o custo por não se conhecer a variância experimental. As Equações (3.18) e (3.20) levam a regiões elípticas. Já as Equações (3.19) e (3.21) não forçam o formato elíptico, o que para modelos não-lineares pode levar a resultados distintos e mais corretos, mas sua obtenção representa um alto custo computacional porque requer muitas avaliações da função objetivo.

A região de confiança torna-se mais complexa quanto maior o número de parâmetros. Quando esse número é maior do que três, a visualização gráfica só é possível através da projeção da sombra da hiper-região para cada par de parâmetros. Para o formato elíptico, é conveniente decompor a matriz $\mathbf{V}_{\boldsymbol{\theta}}$ em autovetores e autovalores:

$$\mathbf{V}_{\boldsymbol{\theta}} = \mathbf{D} \boldsymbol{\Lambda} \mathbf{D}^T = [\mathbf{d}_1 \quad \mathbf{d}_2 \quad \dots \quad \mathbf{d}_{NP}] \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_{NP} \end{bmatrix} \begin{bmatrix} \mathbf{d}_1 \\ \mathbf{d}_2 \\ \vdots \\ \mathbf{d}_{NP} \end{bmatrix} \quad (3.22)$$

onde \mathbf{d}_i são os autovetores e λ_i são os autovalores. Além disso, definindo-se:

$$\mathbf{z} = \mathbf{D}^T (\boldsymbol{\theta} - \hat{\boldsymbol{\theta}}) \quad (3.23)$$

Da Equação (3.18), obtém-se:

$$\mathbf{z}^T \boldsymbol{\Lambda}^{-1} \mathbf{z} \leq \chi_{NP, \alpha}^2 \quad (3.24)$$

Que tem a forma explícita:

$$\sum_{i=1}^{NP} \frac{z_i^2}{\lambda_i} = \chi_{NP, \alpha}^2 \quad (3.25)$$

A partir da Equação (3.25), torna-se trivial o uso de métodos implícitos de plotagem de elipses. Considerando-se dois parâmetros i e j , qualquer par de parcelas da forma $\frac{z_i^2}{\lambda_i} + \frac{z_j^2}{\lambda_j} = \chi_{NP, \alpha}^2$ representa uma projeção da sombra da hiper-elipse. Considerando que $\lambda_i > \lambda_j$, a visualização da sombra requer a obtenção dos

comprimentos do semi-eixo principal, a , e do semi-eixo secundário, b , assim como do ângulo de inclinação, φ :

$$a = \sqrt{\lambda_i \chi_{NP,\alpha}^2} \quad (3.26)$$

$$b = \sqrt{\lambda_j \chi_{NP,\alpha}^2} \quad (3.27)$$

$$\varphi = \arctan \frac{d_{i,i}}{d_{i,j}} \quad (3.28)$$

3.2.3 Comparação entre conjuntos de parâmetros

Para avaliar os parâmetros obtidos neste trabalho frente a outros parâmetros, um teste F bilateral é empregado. Primeiro, obtém-se um ciclo característico do modelo para cada conjunto de parâmetros. A partir do valor da Função Objetivo F_{obj} , dividindo-se pelo número de graus de liberdade ν , obtém-se a variância dos modelos:

$$\hat{\sigma}_i^2 = \frac{F_{obj,i}}{\nu} \quad (3.29)$$

A hipótese nula H_0 é de que não há diferença entre as duas variâncias para um nível de confiança α :

$$H_0: F_{\nu,\nu}^{(1-\alpha)/2} < \frac{\hat{\sigma}_1^2}{\hat{\sigma}_2^2} < F_{\nu,\nu}^{(1+\alpha)/2} \quad (3.30)$$

onde F é o valor da distribuição de Fisher. Se a hipótese nula é aceita, não é possível comprovar estatisticamente com grau de confiança α que existe diferença na qualidade de ajuste entre os conjuntos de parâmetros avaliados.

3.3 Qualidade de predição do modelo

Na Seção 3.1, comentou-se sobre a avaliação da qualidade do modelo. Os critérios dessa avaliação são relacionados às características gerais do modelo e quantificam a qualidade do ajuste aos dados. Por outro lado, torna-se interessante também quantificar os erros de predição do modelo, que mostram explicitamente a influência da propagação dos erros paramétricos sobre a capacidade preditiva. Através da Aproximação de Gauss, é possível calcular as variâncias de predição:

$$\hat{\sigma}_y^2 \cong \mathbf{B} \mathbf{V}_\theta \mathbf{B}^T \quad (3.31)$$

onde $\hat{\sigma}_y^2$ é a matriz das variâncias de predição. O eixo diagonal de $\hat{\sigma}_y^2$ é composto dos valores das variâncias de predição $\hat{\sigma}_{y_i}^2$ para cada ponto experimental y_i . A partir disso é possível plotar os valores preditos com os valores observados e suas respectivas variâncias, o que mostra graficamente a influência dos desvios experimentais e dos desvios de predição sobre o resultado do modelo.

Modelos bons com dados de referência adequados apresentam diferenças entre valores preditos e valores observados condizentes com os desvios envolvidos. Quando isso não ocorre, é possível que o modelo seja ruim ou que os desvios experimentais estão subestimados. Além disso, se os valores apresentam uma tendência fora da linha central dos valores observados, é possível que o modelo não apresente desempenho uniforme na região dos dados de referência.

Capítulo 4 –Metodologia

Com base nos procedimentos de ajuste de parâmetros dos modelos de interesse, fica evidente a possibilidade de melhorias frente aos métodos propostos. Neste capítulo serão abordadas as dificuldades de estimação de parâmetros para modelos com ciclo limite, em particular os que apresentam bifurcação Hopf. Uma nova função objetivo é proposta de forma a melhorar o desempenho de algoritmos de otimização utilizando tratamento e sobreposição de dados.

4.1 Funções objetivo

O alicerce de um problema de otimização é a boa formulação da função objetivo a ser minimizada. A função mais simples consiste na soma dos mínimos quadrados. As seções a seguir visam ilustrar a resposta desta e outras funções objetivo para modelos com ciclo limite.

4.1.1 Modelos com ciclo limite

Em seu trabalho, Nabavi e Williams (2012) propuseram uma nova função objetivo para sistemas biológicos oscilatórios. O cerne da metodologia está em utilizar um conjunto de dados dinâmicos em ciclo limite para se obter um único ciclo representativo a ser usado na estimação de parâmetros. A motivação desse método é o comportamento da soma de mínimos quadrados para modelos oscilatórios frente a parâmetros que tenham impacto direto na frequência do sistema. Para ilustrar o problema em questão, mínimos quadrados foram aplicados entre uma função seno de referência e uma função seno estimada em relação ao período, da forma:

$$y(t) = \sin\left(\frac{2\pi}{10}t\right) \quad (4.1)$$

$$\hat{y}(t, \theta) = \sin\left(\frac{2\pi}{\theta_1}t + \theta_2\right) \quad (4.2)$$

$$F_{obj} = \sum_i^N (y(t_i) - \hat{y}(t_i, \theta))^2 \quad (4.3)$$

onde $y(t)$ é a função de referência, $\hat{y}(t, \theta)$ é a função estimada, θ é o conjunto de parâmetros de ajuste da função estimada (representando período e fase) e F_{obj} é a função objetivo de mínimos quadrados. Como mostra a Figura 4.1, a variação do período faz com que a função objetivo oscile ao redor do ótimo, criando mínimos locais. Além disso, a condição inicial do sistema, dada pela fase, também causa oscilações da função objetivo. Isso inviabiliza a utilização de algoritmos locais de otimização, enquanto algoritmos globais têm sua performance prejudicada porque a região convexa ao redor do ótimo é extremamente restrita.

Com isso em mente, Nabavi e Williams utilizaram tratamento de dados para tornar a função objetivo mais convexa. Para ilustrar o método, considere três funções senoidais idênticas à equação (4.2) de períodos $\theta_{1,1} = 70s$, $\theta_{ref,1} = 100s$ e $\theta_{2,1} = 130s$ (Figura 4.2), com fase $\theta_{1,2} = \theta_{ref,2} = \theta_{2,2} = 0$. Assumindo que um ciclo é constituído por todos os pontos entre dois vales consecutivos, obtém-se um ciclo característico para cada conjunto de dados. Para a soma de mínimos quadrados utiliza-se o ciclo de referência, z_{ref} , e o ciclo estimado, z_i (Figura 4.3).

Os autores propõem, dada uma diferença de períodos entre ciclos, que a falta de amostras entre conjuntos de dados seja suprida preenchendo-se o menor ciclo com zeros. Assim, as equações dos ciclos são dadas por:

$$z(t_i) = \begin{cases} y(t_i) & 0 \leq t_i < \mathcal{T} \\ 0 & \mathcal{T} < t_i \leq \max(\mathcal{T}, \hat{\mathcal{T}}) \end{cases} \quad (4.4)$$

$$\hat{z}(t_i, \hat{\theta}) = \begin{cases} \hat{y}(t_i, \hat{\theta}) & 0 \leq t_i < \hat{\mathcal{T}} \\ 0 & \hat{\mathcal{T}} < t_i \leq \max(\mathcal{T}, \hat{\mathcal{T}}) \end{cases} \quad (4.5)$$

onde $z(t)$ é o ciclo de referência, $\hat{z}(t, \theta)$ é o ciclo estimado, $\hat{\theta}$ é o vetor dos parâmetros estimados, \mathcal{T} é o período de referência e $\hat{\mathcal{T}}$ é o período estimado. Quando condições periódicas não são identificadas, nenhuma alteração aos dados é feita:

$$z(t) = y(t) \quad (4.6)$$

$$\hat{z}(t) = \hat{y}(t) \quad (4.7)$$

Essas modificações resultam em uma função objetivo convexa, como ilustra a Figura 4.4. Contudo, a convexidade obtida é altamente dependente da taxa de amostragem f_s utilizada. Para taxas suficientemente pequenas, as equações senoidais apresentam o resultado anterior, mostrando comportamento de funções contínuas. Conforme a taxa de amostragem aumenta, a função objetivo começa a apresentar uma tendência convexa com comportamento tipo serrote ao redor do ótimo, como mostrado na Figura 4.5. Além disso, o método requer o uso de um ciclo característico para cada conjunto de dados (referência e estimativa).

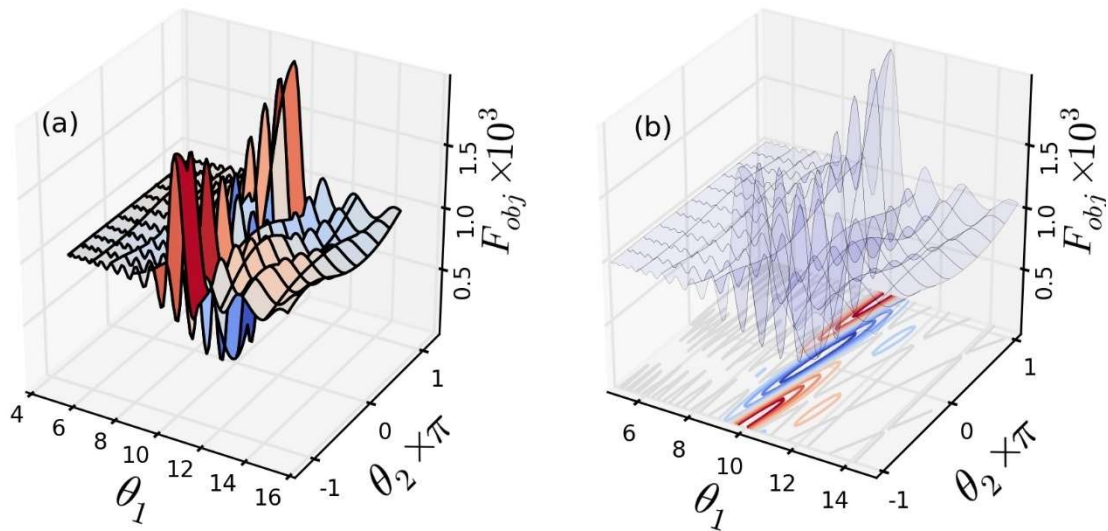


Figura 4.1: (a) Superfície de otimização e (b) curvas de níveis, da função objetivo de mínimos quadrados para o exemplo.

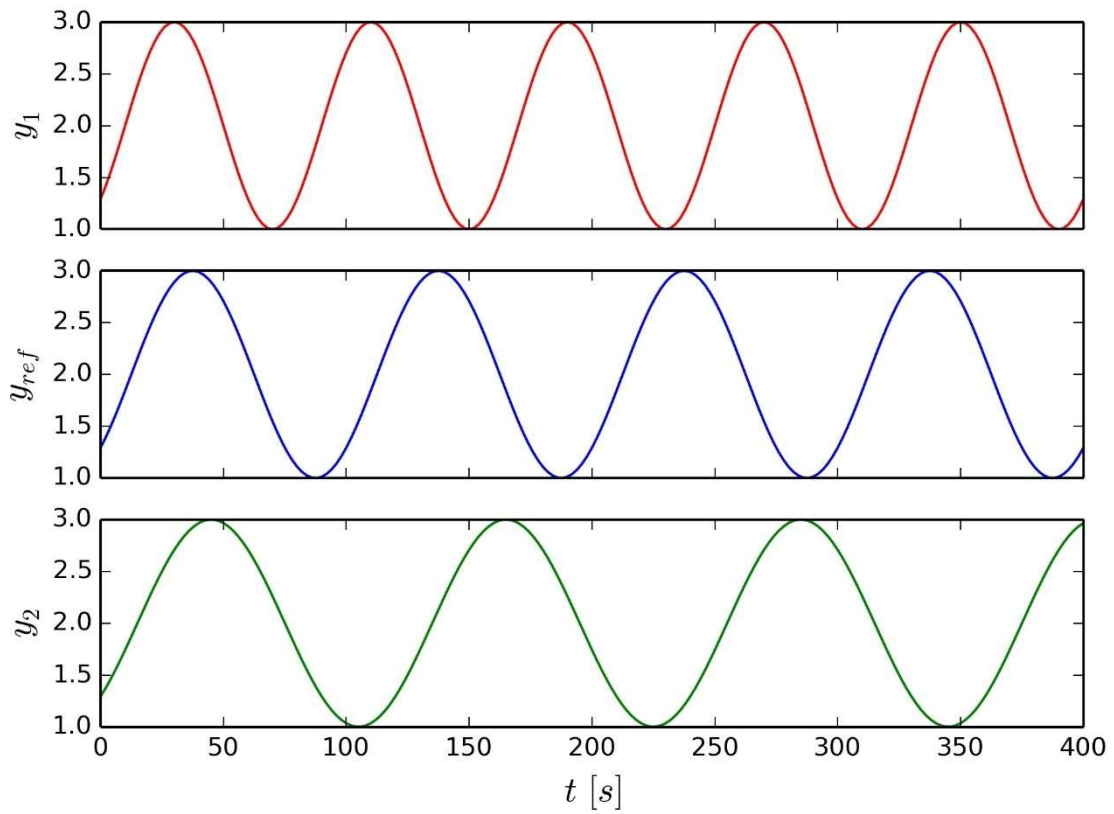


Figura 4.2: Funções senoidais utilizadas como exemplo.

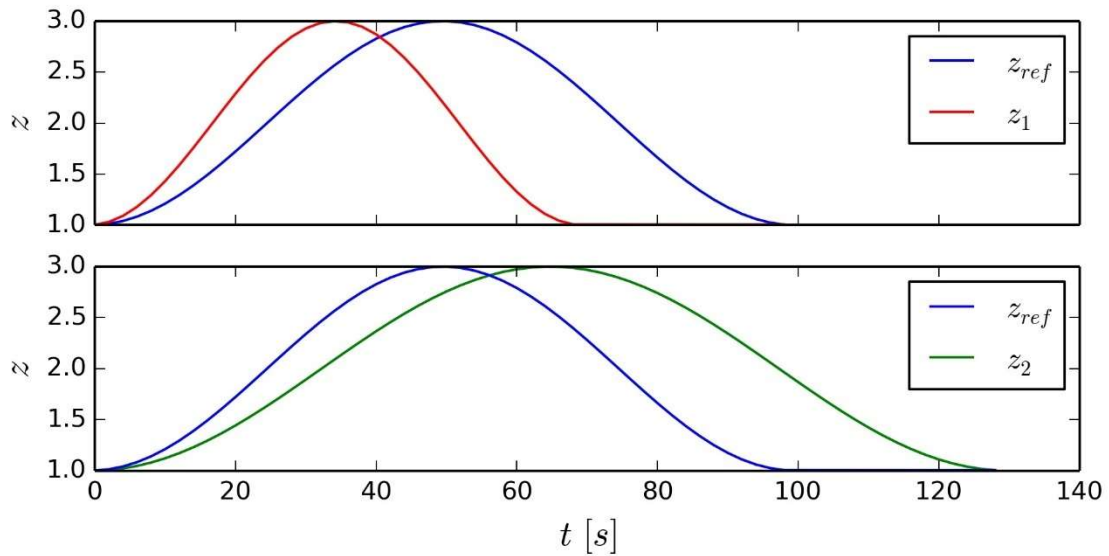


Figura 4.3: Ciclos de referência para comparação ponto a ponto.

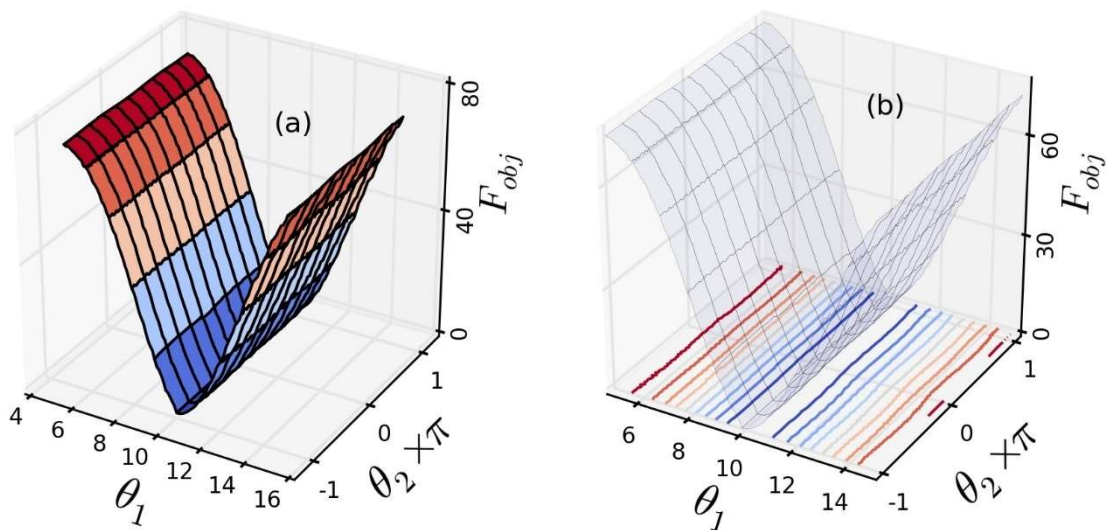


Figura 4.4: (a) Superfície de otimização e (b) curvas de níveis. da função objetivo proposta por Nabavi e Williams para o exemplo.

Para sistemas perfeitamente periódicos como a função seno, a obtenção de um ciclo representativo é direta e não requer nenhuma modificação adicional. Na prática, sistemas sob regime de golfadas não são perfeitamente periódicos e ainda podem apresentar flutuações relevantes nas medições. Assim, o objetivo deste trabalho foi estender a aplicação da função objetivo de Nabavi e Williams a sistemas quase periódicos com variações nas medições. A Figura 4.6 apresenta dados de processo da pressão de um poço de petróleo e ilustra esse tipo de comportamento.

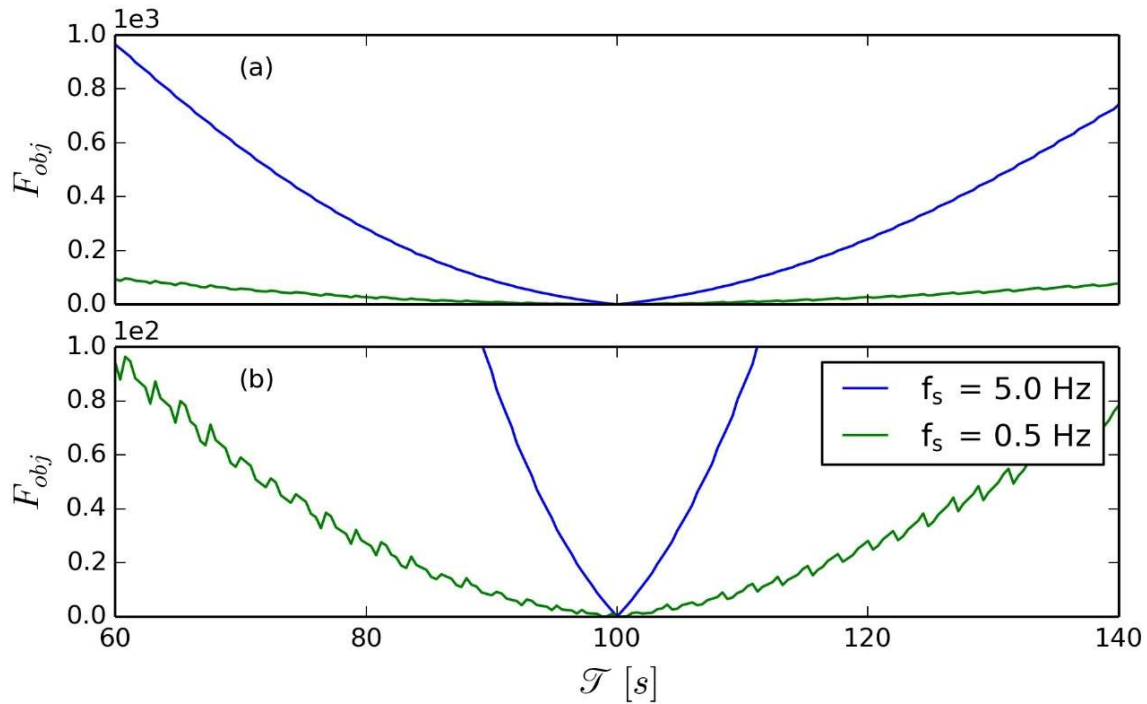


Figura 4.5: Influência da taxa de amostragem na resposta da função objetivo proposta por Nabavi e Williams.

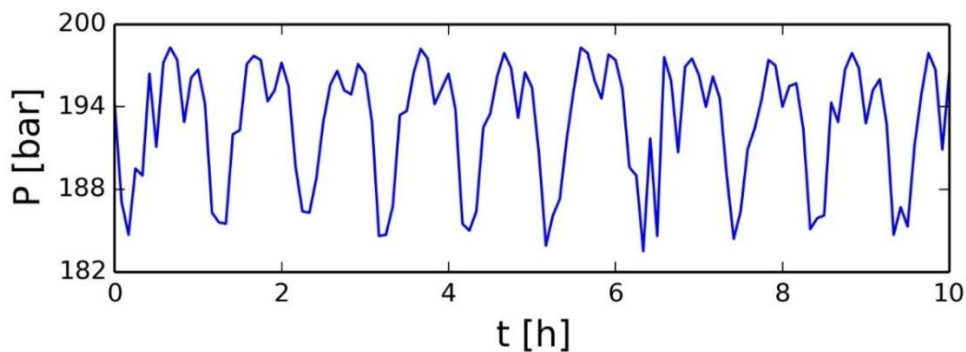


Figura 4.6: Medições reais de pressão de um poço de petróleo.

4.1.2 Modelos com ciclo limite e alta variabilidade

Para ilustrar a aplicação do método, erro com distribuição normal foi adicionado a uma função seno de período variável, dado por:

$$y(x) = \sin(2\pi\sqrt{x + 6\pi}) + e \quad (4.8)$$

$$e \sim \mathcal{N}(0, 0.2) \quad (4.9)$$

onde y é o vetor da variável dependente, x é o vetor da variável independente e e é o vetor do erro com distribuição normal. Primeiramente, é necessária a identificação dos picos e vales nos dados de referência (Figura 4.7a). Isso é feito através de um algoritmo de detecção de picos (BERGMAN, 2016) e permite a separação de todos os ciclos identificados: todos os pontos entre e incluindo dois vales consecutivos configuram um ciclo (Figura 4.7b e Figura 4.7c).

Os ciclos identificados são então deslocados a um novo espaço amostral a partir do tempo $t_0 = 0$ (Figura 4.7d e Figura 4.7e). Contudo, fica claro que os ciclos obtidos não apresentam o mesmo número de pontos por ciclo. Assim, para cada ciclo os dados são regularmente interpolados através de *splines* sem suavização (Figura 4.7f e Figura 4.7g). Um novo tempo amostral é criado de modo a permitir a comparação ponto a ponto:

$$t^C = t_1^C, t_2^C, \dots, t_{\bar{N}-1}^C, t_{\bar{N}}^C \quad (4.10)$$

$$t' = 0, t'_2, \dots, t'_{\bar{N}-1}, \bar{T} \quad (4.11)$$

onde t^C é um tempo amostral com \bar{N} amostras de igual frequência de amostragem do ciclo C , t' é o novo tempo amostral médio, \bar{N} é a média de amostras por ciclo e \bar{T} é o período médio dos dados de referência, obtido por autocorrelação (GERHARD, 2003).

Aplicando o espaço amostral t^C para cada *spline*, são obtidos \bar{N} pontos para cada ciclo, mas é considerado que o espaço amostral real é dado por t' . Isso faz com que os ciclos sejam expandidos ou reduzidos a um mesmo tempo amostral, como se todas as medidas fossem tomadas à mesma taxa amostral (Figura 4.7h e Figura 4.7i). Assim, torna-se possível a comparação ponto a ponto (Figura 4.7j) que resulta em um ciclo médio com uma variância experimental correspondente para cada amostra (Figura 4.7k). Esse é o ciclo experimental de referência utilizado então para a estimação de parâmetros.

Para interpretar os passos descritos, é conveniente considerar como se cada ciclo fosse um único experimento. Como será mostrado posteriormente nos resultados, a obtenção de um ciclo médio é particularmente útil quando os dados de referência apresentam uma tendência clara com muita variabilidade das amostras. Para a otimização, cada iteração aplica o método proposto aos dados simulados de modo a atenuar erros numéricos relacionados à tolerância do *solver*. Isso diminui consideravelmente o tempo de otimização por possibilitar o uso de tolerâncias mais brandas. A função *spline* obtida gera então o ciclo médio baseado no tempo amostral da referência.

Caso o período do modelo seja menor do que o período da referência, a abordagem de Nabavi e Williams de preenchimento dos dados é aplicada utilizando o valor médio do ciclo ao invés de zeros. As pressões de fundo das tubulações são da ordem de 200 bar, de modo que o preenchimento com zeros gera um erro alto para o algoritmo de otimização. Finalmente, Mínimos Quadrados Ponderados (WLS) são aplicados entre o ciclo de referência e o ciclo do modelo, utilizando a variância experimental $\sigma_{z_i}^2$ obtida pelo método:

$$F_{obj} = \sum_{i=1}^N \frac{(z(t'_i) - \hat{z}(t'_i, \theta))^2}{\sigma_{z_i}^2} \quad (4.12)$$

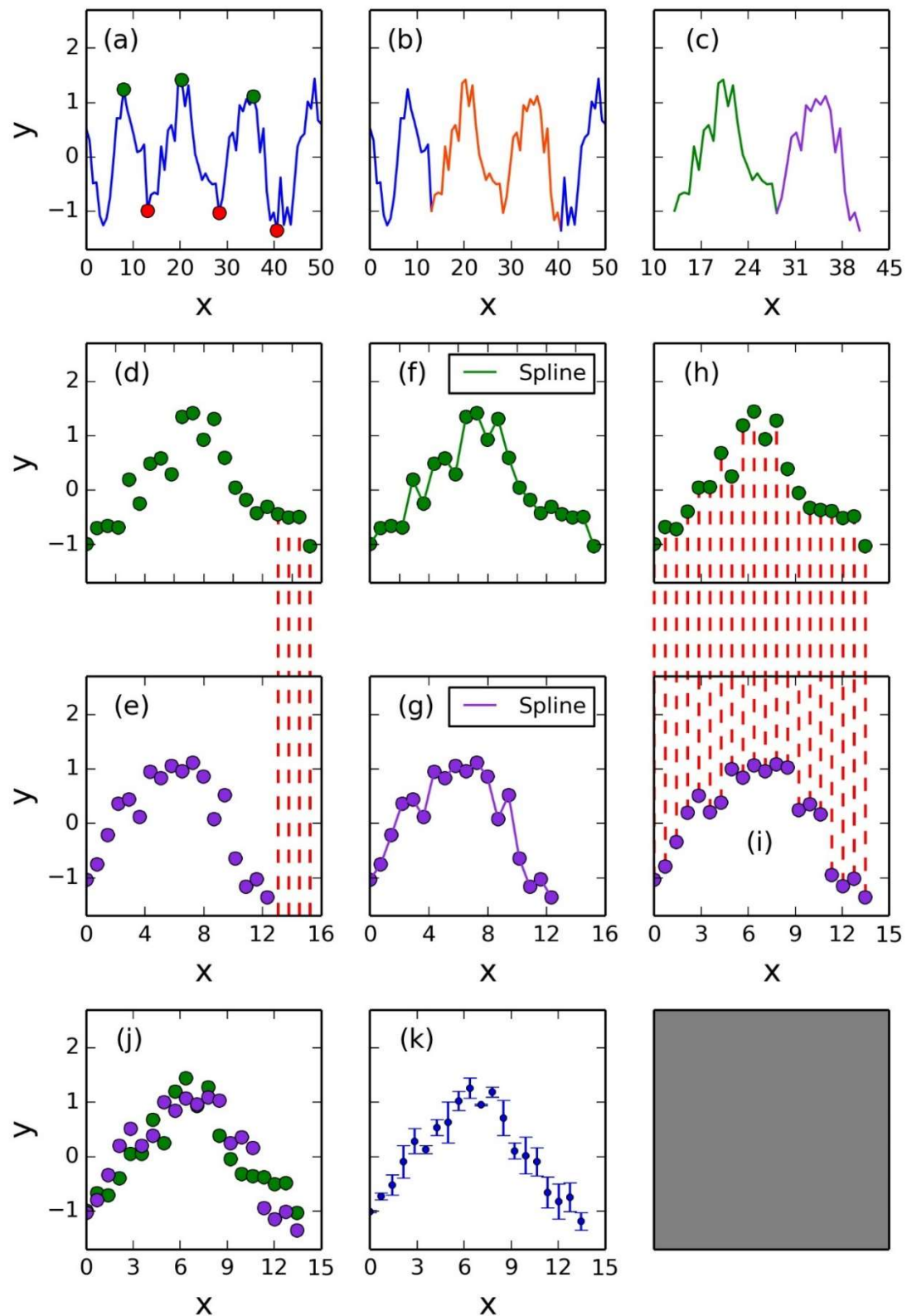


Figura 4.7: Ilustração passo a passo do método proposto: (a) identificação de picos e vales; (b) identificação de todos os ciclos; (c) separação dos ciclos; (d,e) avaliação da diferença entre os ciclos; (f,g) interpolação através de *splines*; (h,i) verificação da amostragem; (j) sobreposição dos ciclos; (k) ciclo médio com variância.

Para comparar a função objetivo de Mínimos Quadrados (LS) com a função objetivo proposta por este trabalho, os parâmetros de uma função seno foram estimados:

$$\hat{y}(x, \theta) = \theta_1 \sin\left(\frac{2\pi x}{\theta_2} + \varphi\right) \quad (4.13)$$

onde \hat{y} é o seno estimado e θ é o vetor dos parâmetros. Para ajuste gráfico, é conveniente definir um valor de fase φ para melhor observar os resultados. A Equação (4.14) define esse valor de modo que o modelo apresente um vale no valor de x em que é identificado o início do primeiro ciclo, x_{v_0} :

$$\varphi = \frac{3\pi}{2} - x_{v_0} \frac{2\pi}{\theta_2} \quad (4.14)$$

A escolha de um seno com período constante está baseada no fato de que os modelos simplificados de golfadas apresentam ciclo limite de período aproximadamente constante para valores fixos de abertura de *choke*. O resultado para as duas estimações está representado na Figura 4.8. É possível observar que o modelo obtido com WLS representa as principais características do ciclo limite melhor do que o modelo obtido com LS. Além disso, a Figura 4.9 e a Figura 4.10 mostram as superfícies de otimização para os dois métodos. Fica evidente como a superfície se torna convexa com o método proposto mesmo para dados ruidosos, embora apresente o comportamento tipo serrote mencionado anteriormente.

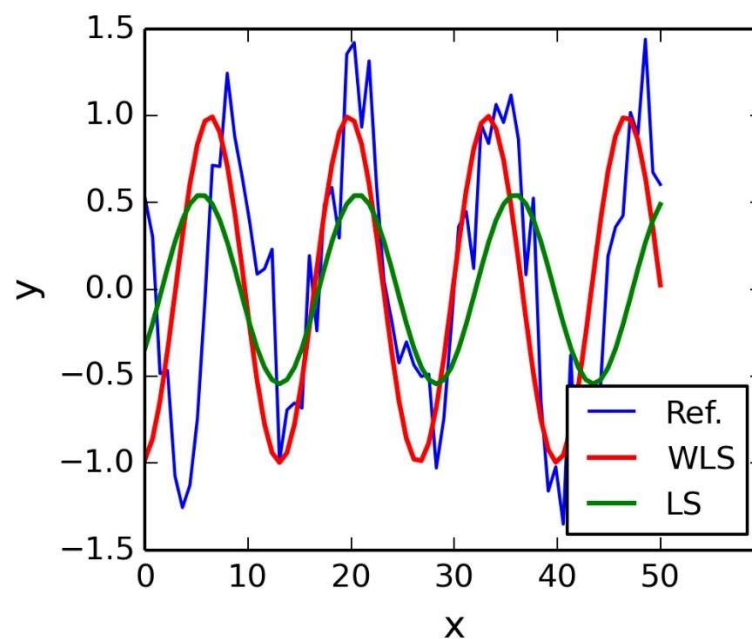


Figura 4.8: Resultados da otimização do exemplo (azul) usando: WLS (vermelho); LS (verde).

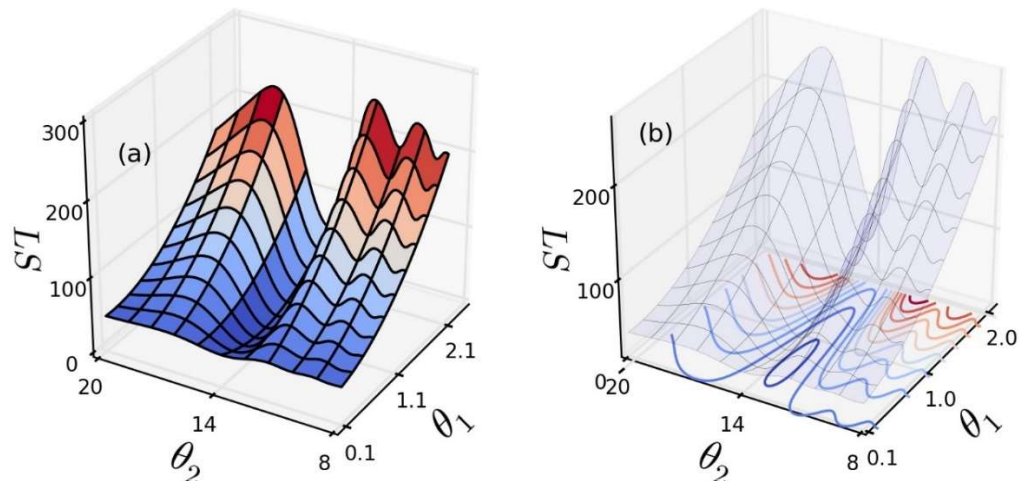


Figura 4.9: (a) Superfície de otimização do exemplo para Mínimos Quadrados, e (b) curvas de nível da função objetivo.

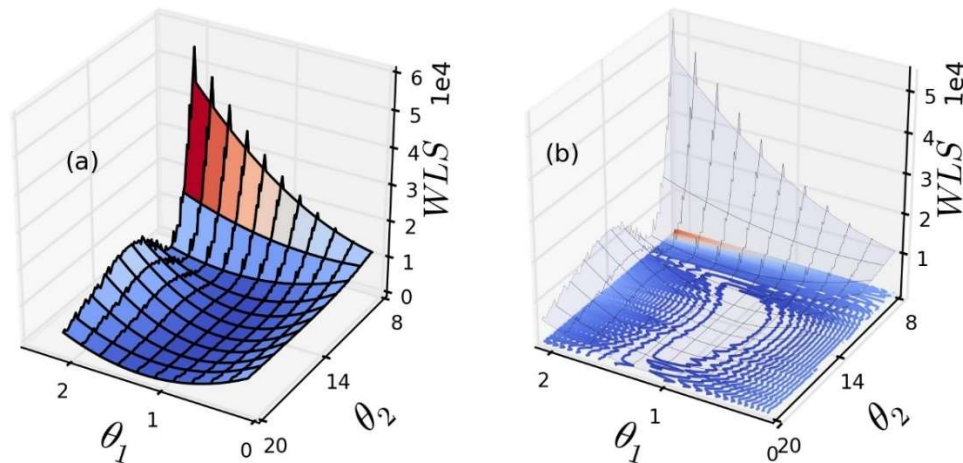


Figura 4.10: (a) Superfície de otimização do exemplo para Mínimos Quadrados Ponderados, e (b) curvas de nível da função objetivo.

4.1.3 Critérios de convergência para modelos com bifurcação Hopf

O método desenvolvido para dados com ciclo limite é restrito a essas condições periódicas ou quase periódicas. Caso essas condições não sejam identificadas, a função objetivo de Mínimos Quadrados clássica é aplicada. Um peso arbitrário ω é adicionado para penalizar dados convergentes:

$$F_{obj} = \omega \sum_{i=1}^N (y_i - \hat{y}_i)^2 \quad (4.15)$$

Os parâmetros estimados podem fazer com que o sistema lentamente entre em equilíbrio, como ilustrado pela Figura 4.11. Essa condição em particular apresenta um comportamento similar a um oscilador harmônico subamortecido. Para se obter a distinção entre as condições de equilíbrio, estado transiente e ciclo limite, alguns critérios são utilizados para se determinar a convergência dos dados:

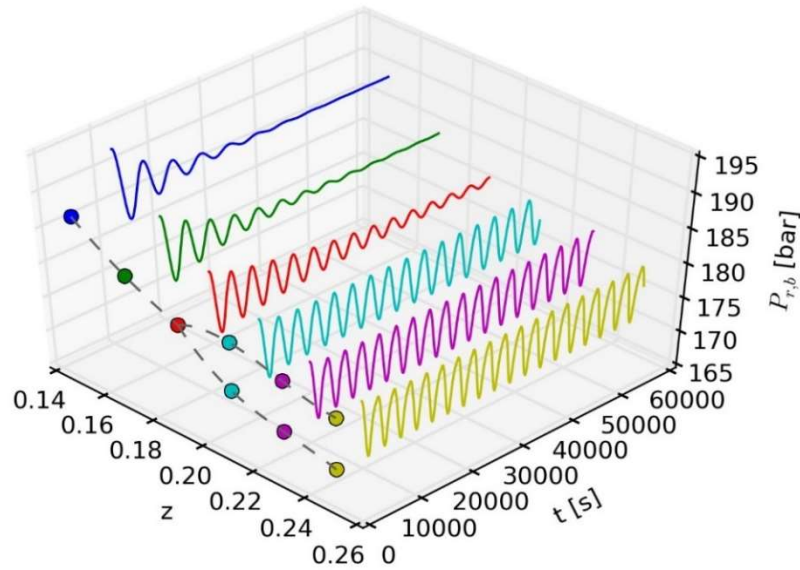


Figura 4.11: Resposta da pressão na base do *riser*, $P_{r,b}$, do modelo de Meglio, Kaasa e Petit (2009) para diferentes aberturas de válvula *choke*, z . Os pontos representam a média (quando o valor de *choke* é inferior ao valor crítico) ou os picos e vales (quando esse valor é superior ao crítico). A linha tracejada cinza representa o diagrama de bifurcação do sistema.

1. O conjunto de dados deve apresentar pelo menos três ciclos de forma que:

$$t_f > 3T \quad (4.16)$$

2. As últimas amostras de um conjunto devem apresentar uma variância maior do que uma dada tolerância ε :

$$\lim_{t \rightarrow t_f} \sigma_y^2 > \varepsilon \quad (4.17)$$

3. A equação limitante do modelo de um oscilador harmônico subamortecido não pode representar adequadamente os picos ou vales do conjunto de dados.

A primeira condição identifica o ciclo limite e visa produzir um ciclo médio de modo a atenuar os erros numéricos relacionados ao *solver*, anteriormente mencionados. A segunda condição é uma medida simples para determinar se um conjunto de dados varia significativamente com o tempo, o que indica a presença ou não de equilíbrio. A última condição é mais complexa e é implementada de modo a identificar condições que caracterizem o estado transiente.

Como dito anteriormente, nessas condições o sistema se comporta analogamente a um oscilador harmônico subamortecido. Esses modelos são representados por um sistema de massa-mola-amortecedor, como ilustrado no exemplo da Figura 4.12 onde o objeto de massa m encontra-se suspenso por uma mola de constante elástica k e também por um amortecedor de coeficiente de amortecimento viscoso c . O equacionamento da posição do objeto em relação ao tempo $x(t)$ pode ser obtido através da Segunda Lei de Newton e, dada a ausência de forças externas, é descrito por uma Equação Diferencial:

$$m\ddot{x} + c\dot{x} + kx = 0 \quad (4.18)$$

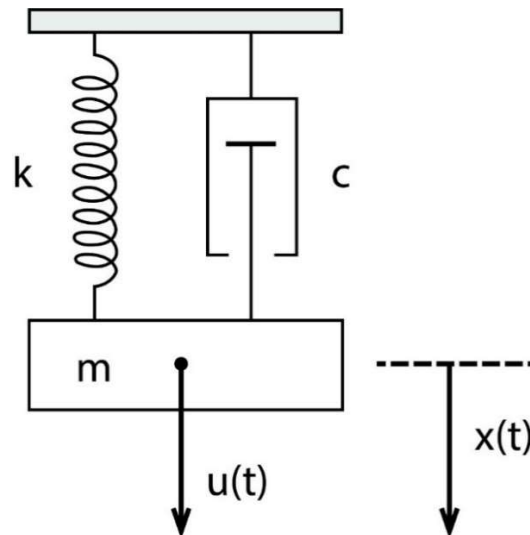


Figura 4.12: Representação de um oscilador harmônico com amortecimento. que pode ser reescrita da forma:

$$\ddot{x} + 2\zeta\omega_0\dot{x} + \omega_0^2x = 0 \quad (4.19)$$

onde ζ é a razão de amortecimento e ω_0 é a frequência angular não-amortecida, respectivamente dados por:

$$\zeta = \frac{c}{2\sqrt{mk}} \quad (4.20)$$

$$\omega_0 = \sqrt{k/m} \quad (4.21)$$

O valor de ζ indica o tipo de sistema em questão (Figura 4.13):

- $\zeta > 1$: sistema superamortecido
- $\zeta = 1$: sistema criticamente amortecido
- $\zeta < 1$: sistema subamortecido

Para sistemas subamortecidos, a solução da Equação (4.19) é uma função seno multiplicada por um fator exponencial responsável pela dissipação da energia do movimento, da forma:

$$y(t) = Ae^{-\zeta\omega_0 t} \sin(\omega_1 t + \phi) \quad (4.22)$$

$$\omega_1 = \omega_0\sqrt{1 - \zeta^2} \quad (4.23)$$

onde A é a amplitude, ϕ é a fase, ω_0 é a frequência angular não-amortecida e ω_1 é a frequência angular amortecida. Dada a similaridade da resposta com modelos de golfadas em estado transiente, é possível que a parcela exponencial do sistema subamortecido consiga adequadamente prever os picos e vales dos modelos. Torna-se necessária a obtenção dos parâmetros da parcela exponencial. A determinação da frequência não-amortecida ω_0 é feita encontrando-se a frequência fundamental f_0 do sistema através da Transformada Discreta de Fourier e utilizando a Equação (4.24):

$$\omega_0 = 2\pi f_0 \quad (4.24)$$

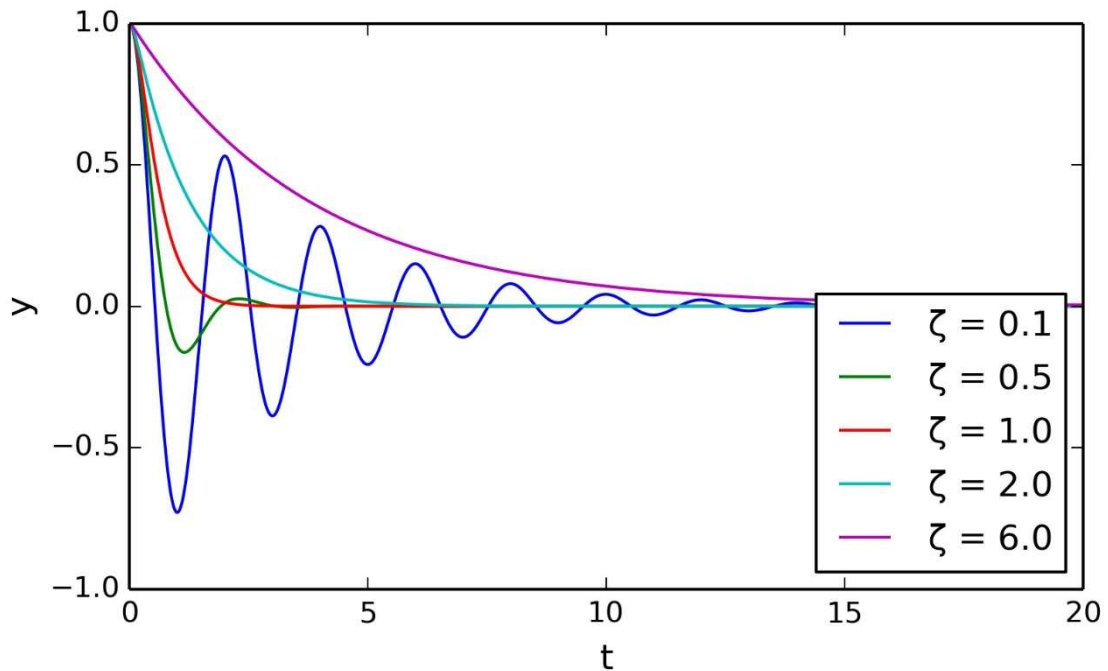


Figura 4.13: Resposta da posição de um oscilador harmônico com amortecimento para diferentes valores de ζ .

Os picos e vales do sistema ocorrem quando a parcela senoidal da equação é máxima ou mínima, respectivamente. Dessa forma, é possível obter os parâmetros ζ e A através de manipulações da Equação (4.22) utilizando-se a média dos valores dos parâmetros para todos os pares de N picos ou vales:

$$\zeta = \frac{1}{N-1} \sum_{i=1}^{N-1} \frac{\ln(y_{i+1}/y_i)}{\omega_0(t_i - t_{i+1})} \quad (4.25)$$

$$A = \frac{1}{N-1} \sum_{i=1}^{N-1} \frac{y_{i+1} - y_i}{\exp(-\zeta\omega_0 t_{i+1}) - \exp(-\zeta\omega_0 t_i)} \quad (4.26)$$

Como ilustra a Figura 4.14, a referência de picos ou vales tem grande influência na equação de contorno obtida quando os dados de referência provêm de um modelo de golfadas. Isso se dá porque o sistema não é perfeitamente descrito como um oscilador harmônico subamortecido, mas a equação de contorno é capaz de corretamente prever os valores dos picos ou dos vales.

Para modelos de golfadas, quando o sistema se encontra em equilíbrio, não existem picos a serem identificados; no caso do ciclo limite, a relação entre picos é uma reta de valor constante porque o modelo é perfeitamente periódico. Para comparação, neste caso, utiliza-se como referência uma reta constante composta do valor médio dos picos. Assim, caso a equação de contorno exponencial consiga representar os picos melhor do que uma equação de reta, o sistema é dito em estado transiente.

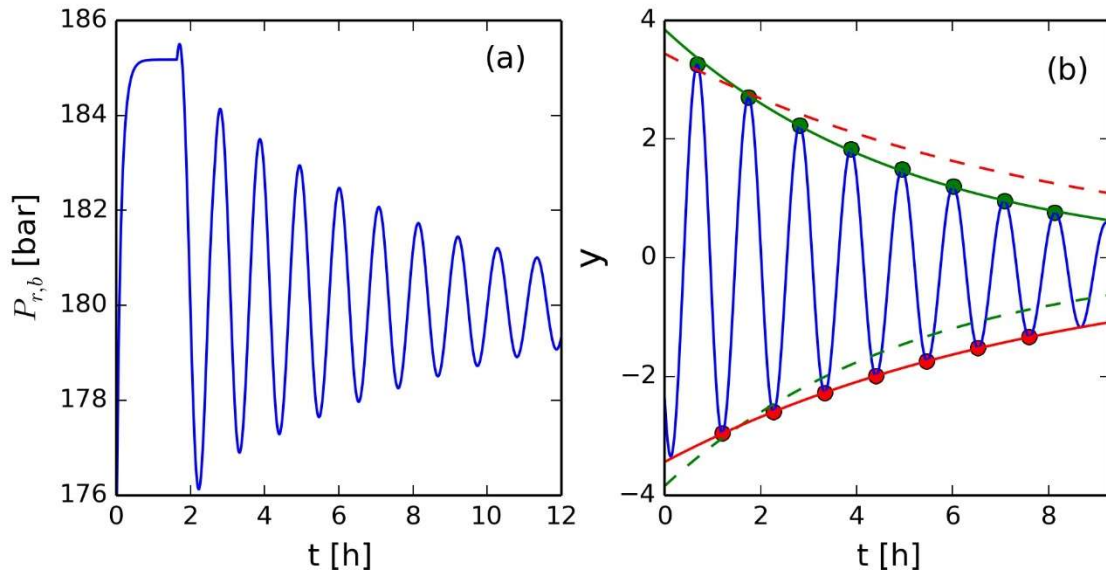


Figura 4.14: (a) Exemplo de ajuste da parcela exponencial de um oscilador harmônico subamortecido para valores de abertura de *choke* inferiores ao crítico o modelo de Meglio, Kaasa e Petit. (b) Quando os parâmetros ζ e A são ajustados para picos ou vales (linhas cheias), o lado recíproco apresenta um ajuste ruim (linhas tracejadas).

Essa diferenciação é feita através de um teste F, com a hipótese nula H_0 de que a variância do modelo exponencial é menor do que a variância do modelo linear. As Equações (4.27) a (4.36) mostram os cálculos necessário para o teste F:

$$\hat{y}_1(t) = Ae^{-\zeta\omega t} \quad (4.27)$$

$$F_{obj,1} = \sum_{i=1}^N [y(t_i) - \hat{y}_1(t_i)]^2 \quad (4.28)$$

$$v_1 = N - 2 \quad (4.29)$$

$$\hat{\sigma}_1^2 = \frac{F_{obj,1}}{v_1} \quad (4.30)$$

$$\hat{y}_2(t) = \bar{y}(t) \quad (4.31)$$

$$F_{obj,2} = \sum_{i=1}^N [y(t_i) - \hat{y}_2(t_i)]^2 \quad (4.32)$$

$$v_2 = N - 1 \quad (4.33)$$

$$\hat{\sigma}_2^2 = \frac{F_{obj,2}}{v_2} \quad (4.34)$$

$$H_0: \frac{\hat{\sigma}_1^2}{\hat{\sigma}_2^2} < F \quad (4.35)$$

$$F = F_{v_2, v_1}^\alpha \quad (4.36)$$

onde $\hat{y}_1(t)$ é o modelo exponencial, $\hat{y}_2(t)$ é o modelo linear, $\hat{\sigma}_i^2$ é a variância estimada, $F_{obj,i}$ é a soma dos mínimos quadrados, v_i é o número de graus de liberdade, N é o número de picos/vales de referência, α é o nível de confiança e F_{v_2, v_1}^α é o valor da distribuição de Fisher para um nível de confiança α com v_2 e v_1 graus de liberdade. Se a hipótese H_0 for aceita, o sistema não apresenta ciclo limite.

4.2 Recursos computacionais

Todos os resultados foram obtidos no ambiente Python 2.7.1. Utilizou-se o algoritmo de otimização global *Particle Swarm Optimization* (POLI; KENNEDY; BLACKWELL, 2007) a partir da implementação Pyswarm, disponível sob licença BSD (LEE, 2014). Os parâmetros de ajuste escolhidos para o algoritmo foram de $\omega = 0,7298$ e $\phi_p = \phi_g = 1,49618$ (CLERC; KENNEDY, 2002). As rotinas de otimização foram realizadas com 50 partículas e um máximo de 2000 iterações, o que significa que a função objetivo é avaliada até 100.050 vezes.

Os modelos foram simulados com a plataforma *open source* JModelica. Suas implementações estão disponíveis nas seções A.1 e A.2 do Apêndice A. Para o modelo de Di Meglio et al. (2009), utilizou-se o *solver* CVode com um passo mínimo de $h = 10^{-4}$; para o de Diehl et al. (2017), o *solver* Radau5ODE.

O algoritmo de detecção de picos utilizado foi o Peakdetect, disponível sob licença WTFPL (BERGMAN, 2016). O código é mostrado na seção A.3 do Apêndice A. Essa implementação é derivada do algoritmo de domínio público Peakdet, proposto em MATLAB (BILLAUER, 2012).

Um exemplo de implementação da função objetivo é mostrado na seção A.9 do Apêndice A.

Valores extraídos de artigos foram obtidos diretamente dos gráficos onde os resultados são apresentados através do *software* Pega Ponto 1.0 (OLIVEIRA; GAMBETTA; PINTO, 2006).

Capítulo 5 – Estudos de caso

Este capítulo apresenta os resultados obtidos para a estimação de parâmetros dos modelos estudados utilizando a nova metodologia proposta para três cenários distintos: um cenário com distúrbio não medido e discrepância de modelo e dois cenários com dados de poços reais. Os modelos e seus parâmetros são então analisados estatisticamente. Também é apresentada uma comparação entre a qualidade de ajuste entre os parâmetros estimados pelo novo método e os parâmetros originalmente ajustados pelos autores.

5.1 Modelo Meglio com distúrbio

O modelo de Di Meglio, Kaasa e Petit (2009), a partir daqui denominado de modelo Meglio, foi simulado para uma abertura fixa de $u = 0,30$ com os parâmetros propostos no artigo de Di Meglio *et al.* (2010). A pressão do separador P_s , inicialmente 6,6 bar, foi lentamente reduzida até 5 bar. Passado um período de tempo, a pressão foi novamente elevada até 6,6 bar. Aos resultados da pressão de fundo $P_{r,b}$ foram adicionados valores de erro com distribuição normal, média $\mu = 0$ e desvio padrão $\sigma = 0,5$ (Figura 5.1a).

O objetivo para este conjunto de dados é a estimação dos parâmetros do modelo Meglio baseado em uma pressão de separação fixa $P_s = 6,6$ bar, considerando que a variação de pressão que gerou os dados seja um distúrbio não medido. Contudo, a aplicação direta do método desenvolvido cria um ciclo médio intermediário entre os ciclos de $P_s = 6,6$ bar e os de $P_s = 5$ bar.

Para melhorar o resultado obtido, considerou-se que os seis primeiros ciclos podem ser usados como referência para incluir os outros ciclos presentes na rotina de otimização (Figura 5.1b). A partir do valor médio dos sete primeiros vales, \bar{P}_v , todas as amostras a partir do último vale foram ajustadas através da Equação (5.1):

$$P'_{r,b,i} = P_{r,b,i} + \Delta P_i \quad (5.1)$$

$$\Delta P_i = \bar{P}_v - sf(t_i) \quad (5.2)$$

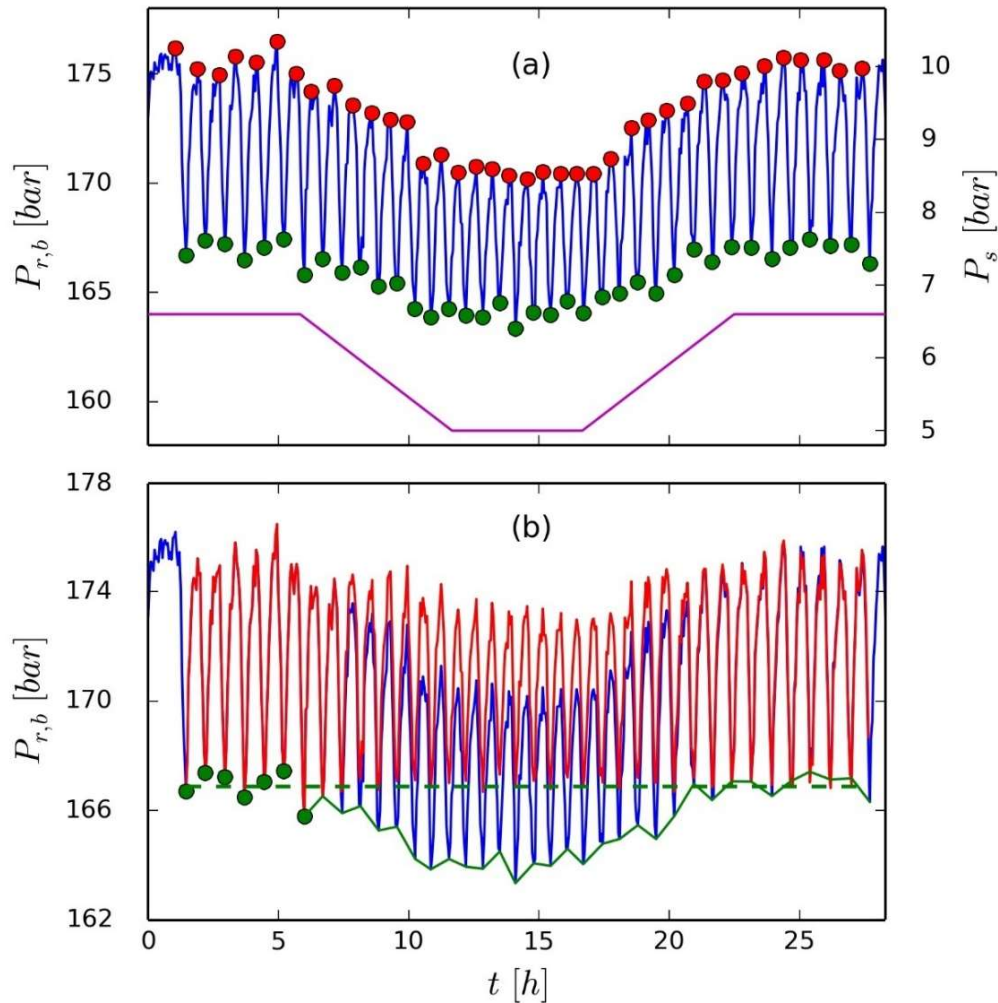


Figura 5.1: Dados utilizados para estimação de parâmetros: (a) modelo Meglio (azul) com erro normalmente distribuído, todos os picos (vermelho) e vales (verde) detectados, e variação da pressão do separador P_s (roxo); (b) resultado da reconstrução (vermelho) dos dados (azul) baseada na diferença entre a média dos vales dos ciclos sem distúrbio (verde tracejado) e os valores da *spline* dos vales dos ciclos com distúrbio (verde cheio).

onde $sf(t_i)$ é o valor da função *spline* de todos os vales a partir do último considerado estável aplicada no ponto t_i , ΔP_i é a diferença de pressão entre referência e o valor da função *spline*, $P_{r,b,i}$ é o valor original de pressão e $P'_{r,b,i}$ é o valor reconstruído de pressão.

Esse processo de reconstrução apresenta um ciclo mais representativo com menor variância (Figura 5.2). Os parâmetros obtidos são mostrados na Tabela 1. Fica evidente como os resultados da estimação sem reconstrução (Figura 5.3a) apresentam menor aderência aos dados do que os resultados obtidos a partir da reconstrução (Figura 5.3b). A Figura 5.4 mostra a comparação entre os ciclos.

A aderência pode ser comprovada pelo teste Chi Quadrado entre os dois resultados, apresentado na Tabela 2. Os parâmetros obtidos com reconstrução apresentam resultados muito bons, o que é esperado se considerando que os dados foram gerados a partir do modelo. Já os parâmetros obtidos sem reconstrução não conseguem justificar os erros experimentais introduzidos.

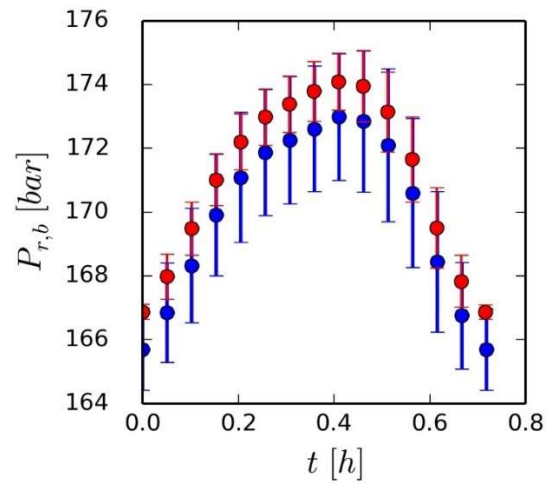


Figura 5.2: Comparação entre ciclo representativo dos dados reconstruídos (vermelho) e dos dados sem reconstrução (azul).

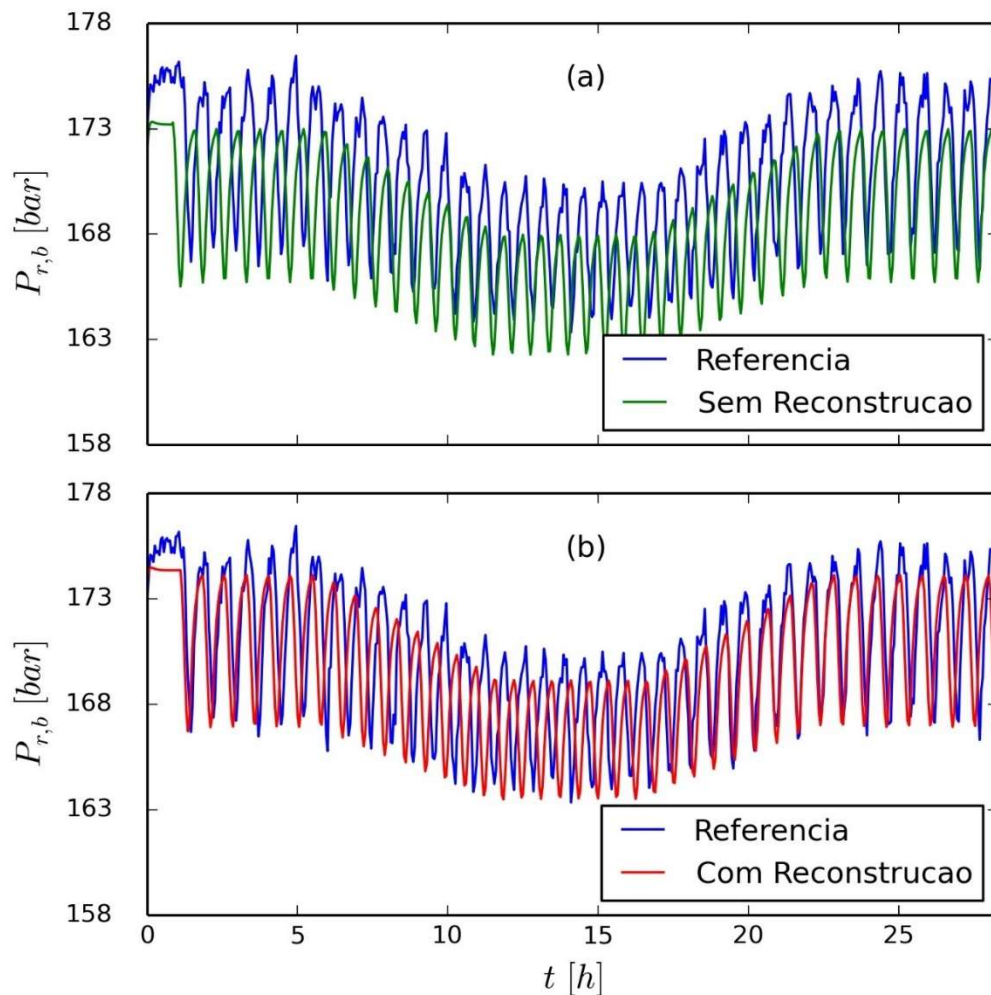


Figura 5.3: Resultados das estimações (a) sem reconstrução e (b) com reconstrução.

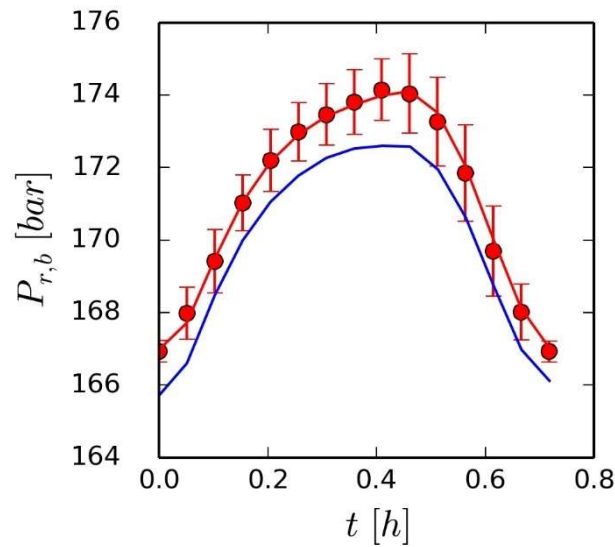


Figura 5.4: Comparação entre ciclo dos dados reconstruídos (pontos vermelhos), ciclo do modelo estimado com dados reconstruídos (linha vermelha) e ciclo do modelo estimado sem reconstrução (linha azul).

Tabela 1: Comparação entre os resultados das estimações de parâmetros do modelo de Meglio et al. (2010).

Parâmetro	Di Meglio et al. (2010)	Dados sem reconstrução	Dados com reconstrução
$m_{l,still}$ [kg]	37.300,00	36.860,74	37.301,89
C_g [m · s]	$1,00 \cdot 10^{-4}$	$9,98 \cdot 10^{-5}$	$1,00 \cdot 10^{-5}$
V_{eb} [m ³]	48,00	46,72	47,63
C_{out} [m ²]	$2,80 \cdot 10^{-3}$	$2,88 \cdot 10^{-3}$	$3,00 \cdot 10^{-3}$
ϵ [-]	0,78	0,80	0,79

Tabela 2: Teste Chi Quadrado entre os resultados com e sem reconstrução dos dados.

$\alpha = 99\%, \nu = 10$	$\chi^2_{\frac{1-\alpha}{2}, \nu}$	$F_{obj} = WLS$	$\chi^2_{\frac{1+\alpha}{2}, \nu}$
Dados originais	2,156	57,718	25,188
Dados reconstruídos	2,156	0,487	25,188

A qualidade do ajuste também pode ser quantificada pelo teste F (Tabela 3):

Tabela 3: Teste F entre os resultados com e sem reconstrução dos dados.

$\alpha = 99\%, \nu = 10$	$F_{\frac{1-\alpha}{2}, \nu, \nu}$	$\frac{\hat{\sigma}_1^2}{\hat{\sigma}_2^2}$	$F_{\frac{1+\alpha}{2}, \nu, \nu}$
Sem (1) e com reconstrução (2)	0,171	118,555	5,847
Di Meglio (1) e reconstrução (2)		54,305	

A Tabela 3 mostra que o ajuste obtido com reconstrução apresenta variância estimada de predição menor do que o ajuste sem reconstrução. Os parâmetros originais apresentam resultado pior do que o esperado porque o ciclo de referência utilizado leva em consideração os distúrbios não medidos. Também é possível obter uma matriz de correlações (Tabela 4). Fica evidente a alta correlação entre parâmetros ($m_{l,still}$ e C_{out} ; C_g e ϵ) que é uma das fontes de dificuldade para estimação desses parâmetros.

A Figura 5.5, a Figura 5.6 e a Figura 5.7 mostram as superfícies de otimização entre os parâmetros $m_{l,still}$ e C_{out} considerando Mínimos Quadrados, a função objetivo de Diehl *et al.* (2017) e a função objetivo proposta neste trabalho, respectivamente. Destaca-se a diminuição da não-convexidade, atingindo a quase total convexidade para a função proposta. O degrau apresentado é referente ao comportamento tipo serrote comentado na Seção 4.1.1 (Figura 4.5). Já a Figura 5.8 mostra como as regiões de confiança dos parâmetros são complexas. As regiões elípticas são pouco representativas frente às regiões de máxima verossimilhança, obtidas a partir das avaliações da função objetivo com o algoritmo PSO.

Tabela 4: Matriz de correlações dos resultados com reconstrução dos dados.

	$m_{l,still} [kg]$	$C_g [m \cdot s]$	$V_{eb} [m^3]$	$C_{out} [m^2]$	$\epsilon [-]$
$m_{l,still} [kg]$	1	0.482672	-0.25262	0.895107	0.487792
$C_g [m \cdot s]$	0.482672	1	-0.50732	0.10248	0.919645
$V_{eb} [m^3]$	-0.25262	-0.50732	1	-0.11667	-0.33283
$C_{out} [m^2]$	0.895107	0.10248	-0.11667	1	0.180514
$\epsilon [-]$	0.487792	0.919645	-0.33283	0.180514	1

Assim, é possível concluir que a função objetivo proposta foi capaz de estimar parâmetros adequados. O tratamento dos dados auxilia na obtenção de um ciclo de referência mais representativo, o que leva a um ajuste melhor do modelo. Também fica explícita a alta correlação entre parâmetros do modelo e a melhor representatividade da região de confiança de máxima verossimilhança.

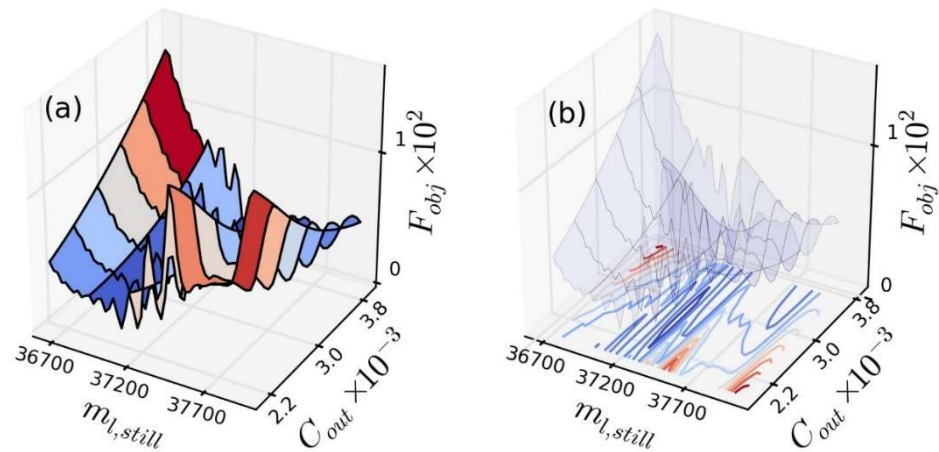


Figura 5.5: (a) Superfície de otimização e (b) curvas de níveis de Mínimos Quadrados para os parâmetros $m_{l,still}$ e C_{out} .

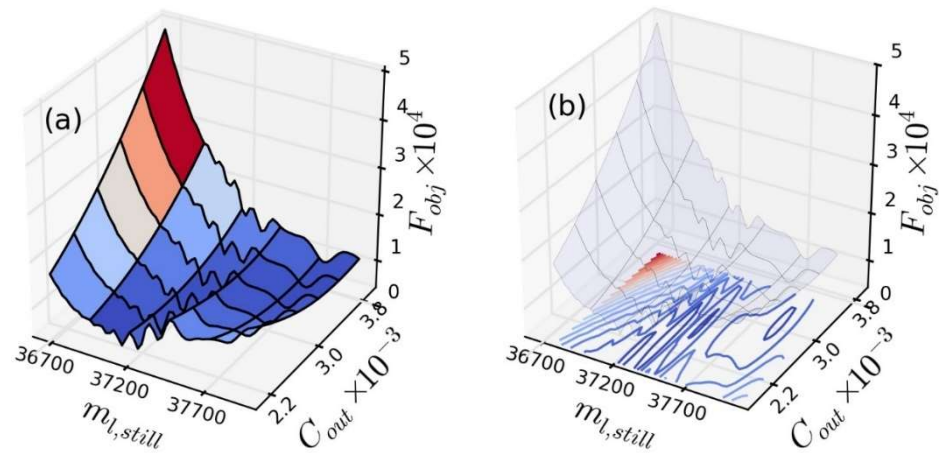


Figura 5.6: (a) Superfície de otimização e (b) curvas de níveis da função objetivo proposta por Diehl *et al.* (2017) para os parâmetros $m_{l,still}$ e C_{out} .

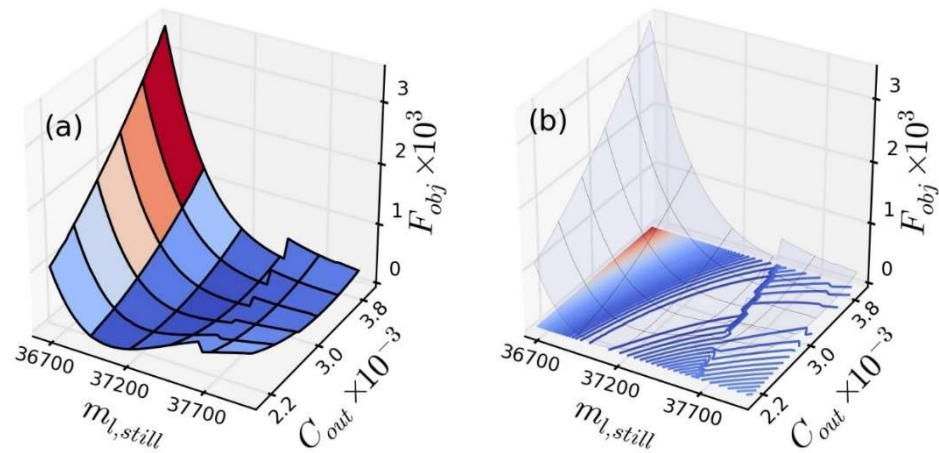


Figura 5.7: (a) Superfície de otimização e (b) curvas de níveis da função objetivo proposta neste trabalho para os parâmetros $m_{l,still}$ e C_{out} .

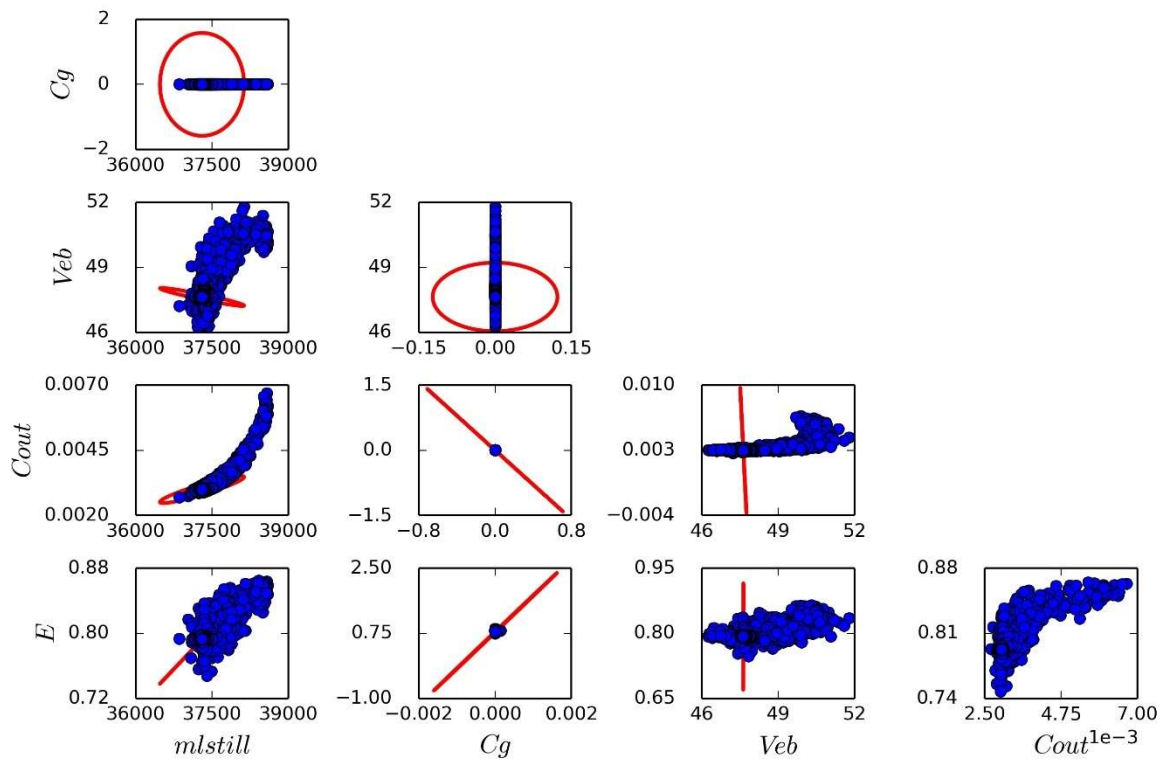


Figura 5.8: Regiões de confiança para um grau de confiança $\alpha = 99\%$ dos parâmetros obtidos com reconstrução dos dados: região elíptica (vermelho) e região de máxima verossimilhança (azul).

5.2 Modelo Meglio aplicado a dados reais

Os dados reais de pressão de topo $P_{r,t}$ e de pressão de fundo $P_{r,b}$ foram obtidos da Figura 7 do artigo de Di Meglio *et al.* (2012). O *software* Pega Ponto gerou os dados representados na Figura 5.9. As leituras de $P_{r,b}$ apresentam uma baixa frequência de amostragem que foi reproduzida manualmente para as leituras de $P_{r,t}$. Por outro lado, os dados de $P_{r,t}$ obtidos apresentam pouca variabilidade; assim, suas variâncias foram consideradas constantes e removidas do cálculo da função objetivo quando a $P_{r,t}$ é usada como referência. Isso aumenta significativamente a performance da otimização.

Os parâmetros originais são idênticos aos apresentados anteriormente. Sabe-se que a pressão de separação é de $P_s = 6,6$ bar, mas é assumido que a abertura de válvula é de $u = 30\%$. Discrepâncias nesse valor devem ser compensadas pelo ajuste do parâmetro C_{out} .

Primeiramente, o modelo foi simulado nessas condições com os parâmetros publicados para se determinar a qualidade original com base nas métricas anteriormente discutidas. Os parâmetros foram então estimados em três cenários: utilizando como referência somente $P_{r,b}$; somente $P_{r,t}$; e a combinação de $P_{r,b}$ e $P_{r,t}$. A Tabela 5 mostra a comparação entre os parâmetros obtidos nesses cenários.

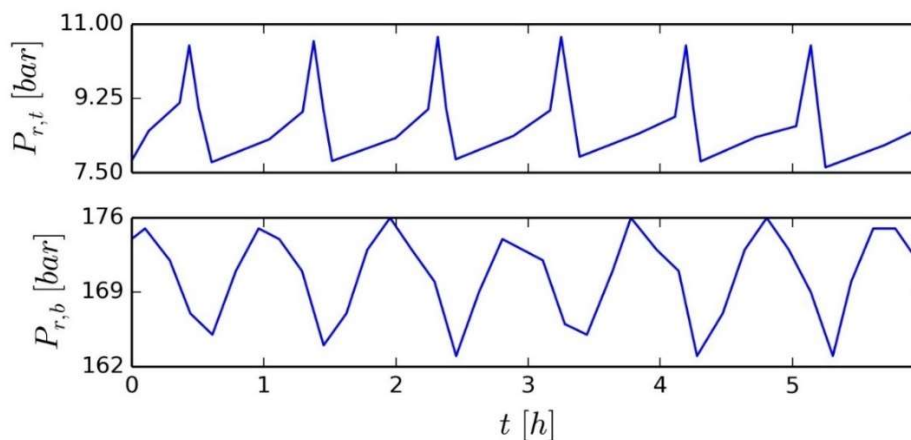


Figura 5.9: Dados obtidos do artigo de Di Meglio et al. (2012).

Tabela 5: Comparação entre os parâmetros estimados para o modelo Meglio.

Parâmetro	Di Meglio <i>et al.</i>	Somente $P_{r,b}$	Somente $P_{r,t}$	$P_{r,b}$ e $P_{r,t}$
$m_{l,still}$ [kg]	37.300,00	37.643,05	43.883,73	36698,30
C_g [$m \cdot s$]	$1,00 \cdot 10^{-4}$	$1,00 \cdot 10^{-5}$	$7,74 \cdot 10^{-4}$	$3,27 \cdot 10^{-4}$
V_{eb} [m^3]	48,00	43,05	45,75	40,31
C_{out} [m^2]	$2,80 \cdot 10^{-3}$	$4,29 \cdot 10^{-3}$	$2,72 \cdot 10^{-3}$	$2,79 \cdot 10^{-3}$
ϵ [-]	0,78	0,63	0,70	0,65

Os limites superior e inferior da otimização são mostrados na Tabela 6. Os limites dos parâmetros $m_{l,still}$, V_{eb} e ϵ foram estabelecidos por experiência com a resposta do modelo: limites mais largos prejudicam muito a otimização. Já os parâmetros C_g e C_{out} foram limitados com uma diferença de uma ordem de grandeza em relação à grandeza original.

A média de tempo por avaliação da função objetivo foi de 576 ms, obtida com auxílio da função “%%timeit” do Python. Contudo, esse tempo é avaliado com os parâmetros ótimos, que apresentam ciclo limite. Como os valores iniciais dos estados nunca são atualizados, um conjunto de parâmetros que cause dificuldades iniciais de convergência pode aumentar o tempo de avaliação para mais de 1 s.

Tabela 6: Limites de otimização para os parâmetros do modelo Meglio.

Parâmetro	Limite inferior	Di Meglio <i>et al.</i>	Limite superior
$m_{l,still}$ [kg]	0,00	37.300,00	50.000,00
C_g [$m \cdot s$]	$1,00 \cdot 10^{-5}$	$1,00 \cdot 10^{-4}$	$1,00 \cdot 10^{-3}$
V_{eb} [m^3]	10,00	48,00	60,00
C_{out} [m^2]	$1,00 \cdot 10^{-4}$	$2,80 \cdot 10^{-3}$	$1,00 \cdot 10^{-2}$
ϵ [-]	0,30	0,78	0,90

5.2.1 Resultados originais

A Figura 5.10 mostra a comparação entre os ciclos representativos dos dados e do modelo original. Os coeficientes de correlação obtidos para as pressões $P_{r,b}$ e $P_{r,t}$ foram $\rho_{r,b}^m = 0,452$ e $\rho_{r,t}^m = 0,114$, respectivamente. A Tabela 7 mostra o teste Chi Quadrado para esses dados. Fica evidente que o modelo por si só não é capaz de prever satisfatoriamente os dados experimentais.

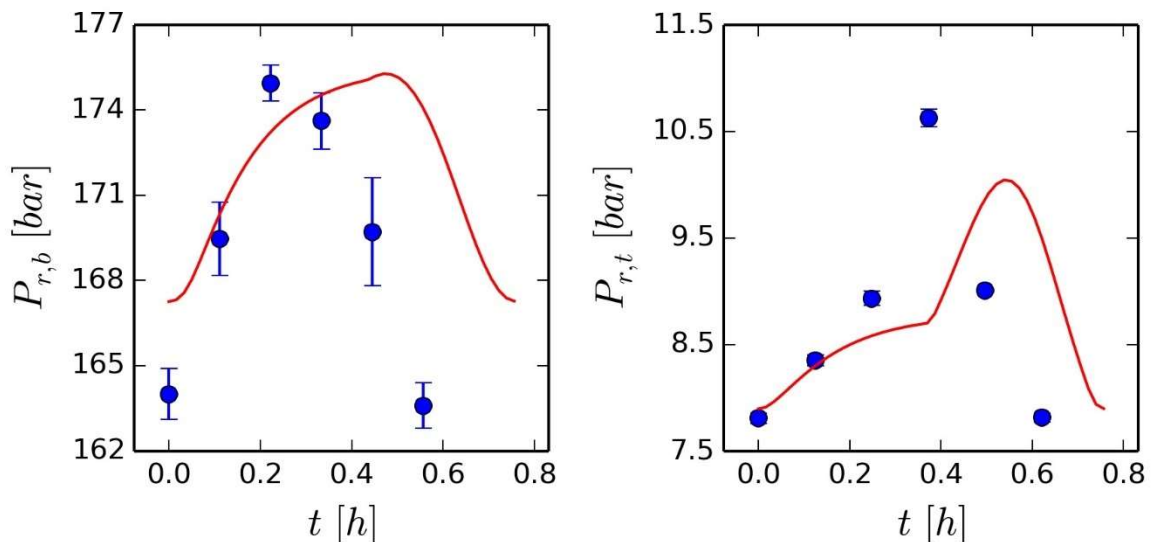


Figura 5.10: Ciclos representativos: dados de referência (azul); modelo Meglio conforme publicado (vermelho).

Tabela 7: Teste Chi quadrado para o modelo Meglio original.

$\alpha = 95\%, \nu = 1$	$\chi_{\frac{1-\alpha}{2}, \nu}^2$	$F_{obj} = WLS$	$\chi_{\frac{1+\alpha}{2}, \nu}^2$
$P_{r,b}$	0,001	201,273	5,024
$P_{r,t}$		7,417	

5.2.2 Estimação com base em $P_{r,b}$

A Figura 5.11 mostra os resultados obtidos para a estimação baseada em $P_{r,b}$, enquanto a Figura 5.12 mostra a evolução da melhor posição global das partículas com as iterações. O número total de iterações foi de 1373. Os coeficientes de correlação obtidos foram de $\rho_{r,b}^m = 0,978$ e $\rho_{r,t}^m = 0,869$. A Tabela 8 mostra os resultados do teste Chi Quadrado. O modelo consegue acompanhar as tendências da $P_{r,b}$, mas não consegue justificar os erros experimentais nem para $P_{r,b}$ nem para $P_{r,t}$ embora apresente claras melhorias na representatividade dos dados de $P_{r,b}$.

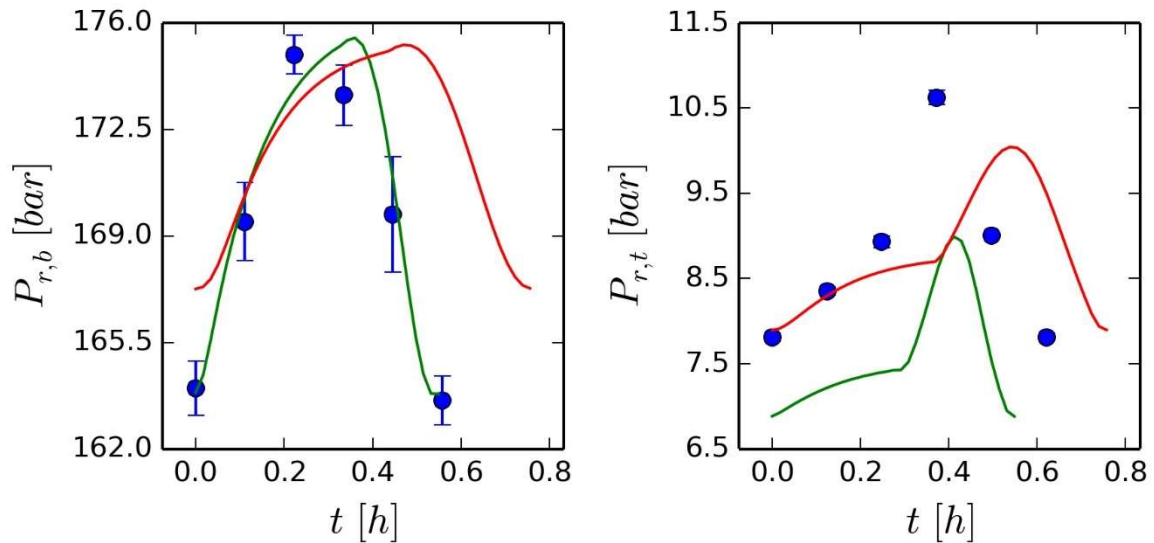


Figura 5.11: Ciclos representativos: dados originais (azul); modelo Meglio original (vermelho); modelo Meglio estimado com base na pressão $P_{r,b}$ (verde).

Tabela 8: Teste Chi Quadrado para o modelo Meglio estimado com a pressão $P_{r,b}$.

$\alpha = 95\%, \nu = 1$	$\chi^2_{\frac{1-\alpha}{2}, \nu}$	$F_{obj} = WLS$	$\chi^2_{\frac{1+\alpha}{2}, \nu}$
$P_{r,b}$	0,001	7,434	5,024
$P_{r,t}$		10,922	

A matriz de correlações obtida é mostrada na Tabela 9. Fica evidente a alta correlação entre os parâmetros. Como pode ser observado na Figura 5.13, o modelo resultante tem dificuldade de prever os picos de pressão da $P_{r,b}$. Já a $P_{r,t}$ está totalmente desajustada em relação aos dados.

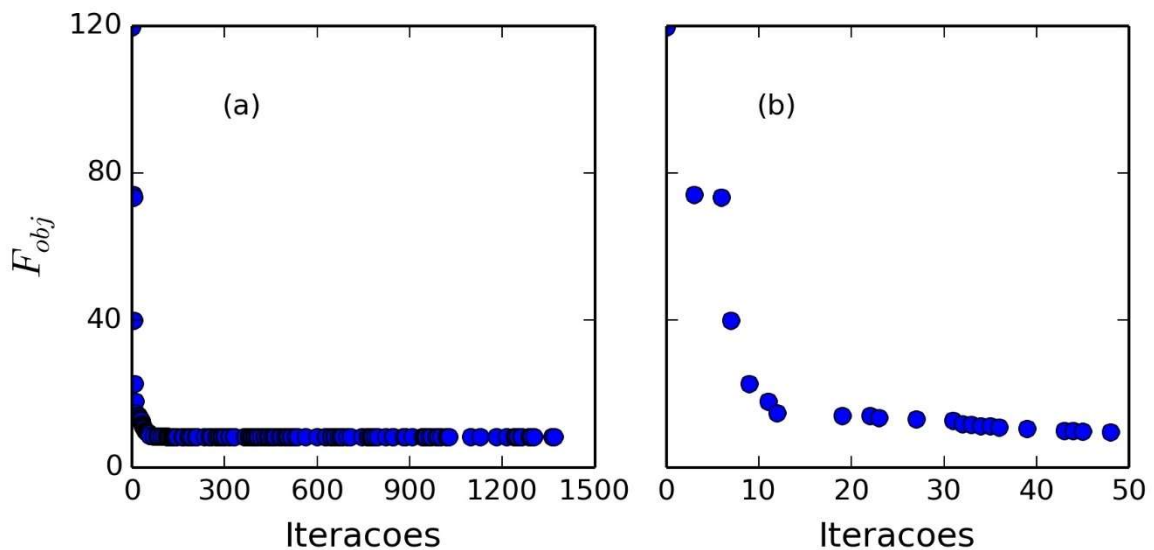


Figura 5.12: Evolução da melhor posição da função objetivo para a estimação com $P_{r,b}$.

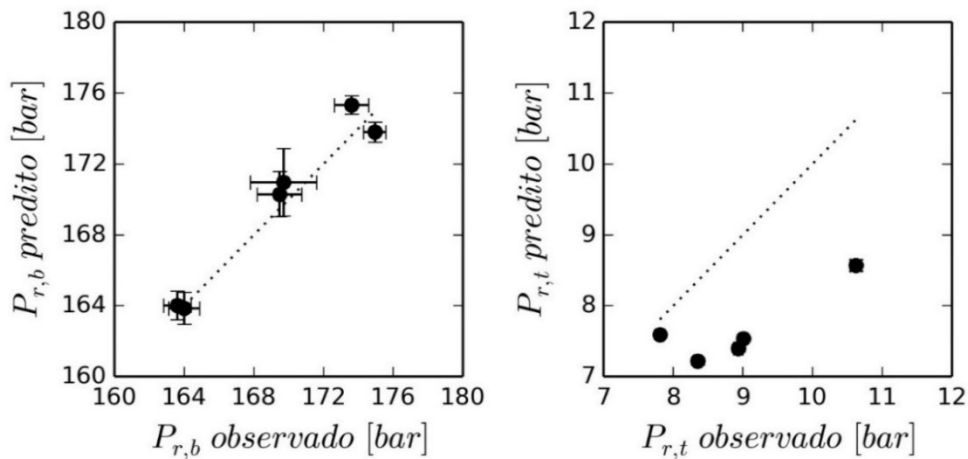
Tabela 9: Matriz de correlações do modelo Meglio estimado com a pressão $P_{r,b}$.

	$m_{l,still}$ [kg]	C_g [m · s]	V_{eb} [m ³]	C_{out} [m ²]	ϵ [–]
$m_{l,still}$ [kg]	1	-0.90911	-0.91422	0.968896	-0.96352
C_g [m · s]	-0.90911	1	0.74841	-0.93961	0.878606
V_{eb} [m ³]	-0.91422	0.74841	1	-0.86222	0.906134
C_{out} [m ²]	0.968896	-0.93961	-0.86222	1	-0.9875
ϵ [–]	-0.96352	0.878606	0.906134	-0.9875	1

A Figura 5.14 mostra as regiões de máxima verossimilhança resultantes. Como foi observado para o modelo teórico com distúrbio, as regiões elípticas são pouco representativas e dificultam a representação das regiões de máxima verossimilhança; portanto, não foram plotadas. Com essas regiões é possível observar graficamente a alta correlação entre os parâmetros. A Tabela 10 mostra o teste F entre o modelo original e o modelo estimado. Os parâmetros são melhores para representar $P_{r,b}$, mas não apresentam diferença quanto à estimação da $P_{r,t}$.

Tabela 10: Teste F do modelo Meglio estimado com a pressão $P_{r,b}$.

$\alpha = 95\%, \nu = 1$	$F_{\frac{1-\alpha}{2}, \nu, \nu}$	$\frac{\hat{\sigma}_1^2}{\hat{\sigma}_2^2}$	$F_{\frac{1+\alpha}{2}, \nu, \nu}$
$P_{r,b}$ estimada (1)		0,037	
$P_{r,b}$ original (2)	1,001		6,024
$P_{r,t}$ estimada (1)		1,473	
$P_{r,t}$ original (2)			

Figura 5.13: Valores preditos e valores observados de $P_{r,b}$ e $P_{r,t}$ do modelo Meglio estimado com $P_{r,b}$.

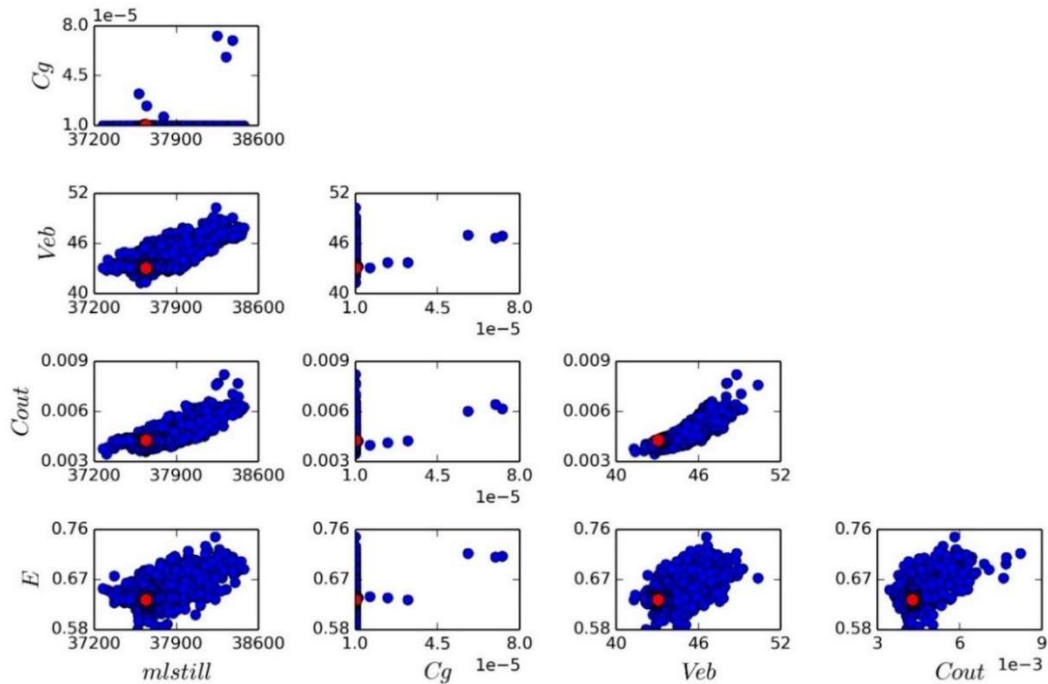


Figura 5.14: Regiões de confiança de máxima verossimilhança para o modelo de Di Meglio et al. (2009) estimado com $P_{r,b}$. Valores ótimos são representados em vermelho.

5.2.3 Estimação com base em $P_{r,t}$

Os resultados da estimação baseada em $P_{r,t}$ são mostrados na Figura 5.15. Foram avaliadas 155 iterações. A evolução da função objetivo é apresentada na Figura 5.16. Os coeficientes de correlação obtidos foram de $\rho_{r,b}^m = 0,852$ e $\rho_{r,t}^m = 0,872$. A Tabela 11 mostra os resultados do teste Chi Quadrado. O modelo apresenta dificuldades de acompanhar as tendências de $P_{r,b}$ e $P_{r,t}$, embora o teste Chi Quadrado recomende o uso do modelo para a estimação da $P_{r,t}$. A Tabela 12 mostra a matriz de correlações obtida.

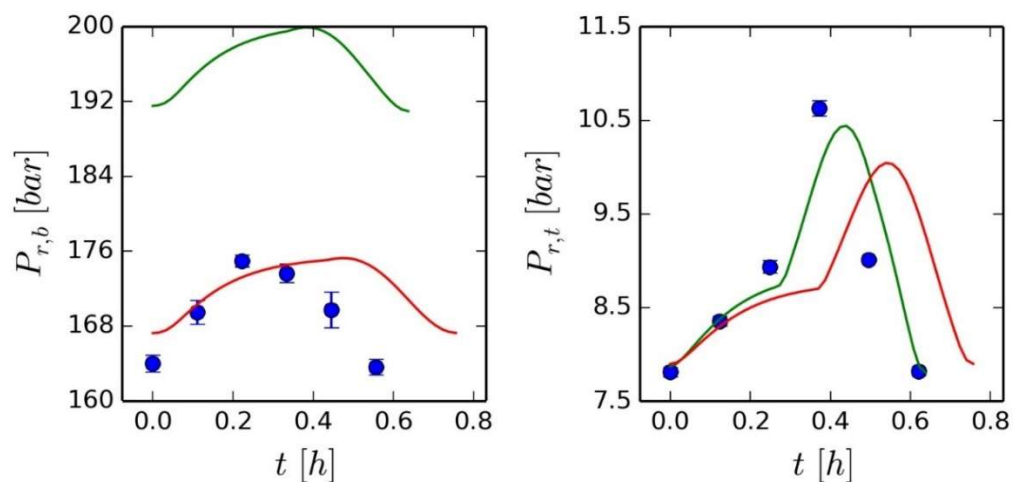


Figura 5.15: Ciclos representativos: dados originais (azul); modelo Meglio original (vermelho); modelo Meglio estimado com base na pressão $P_{r,t}$ (verde).

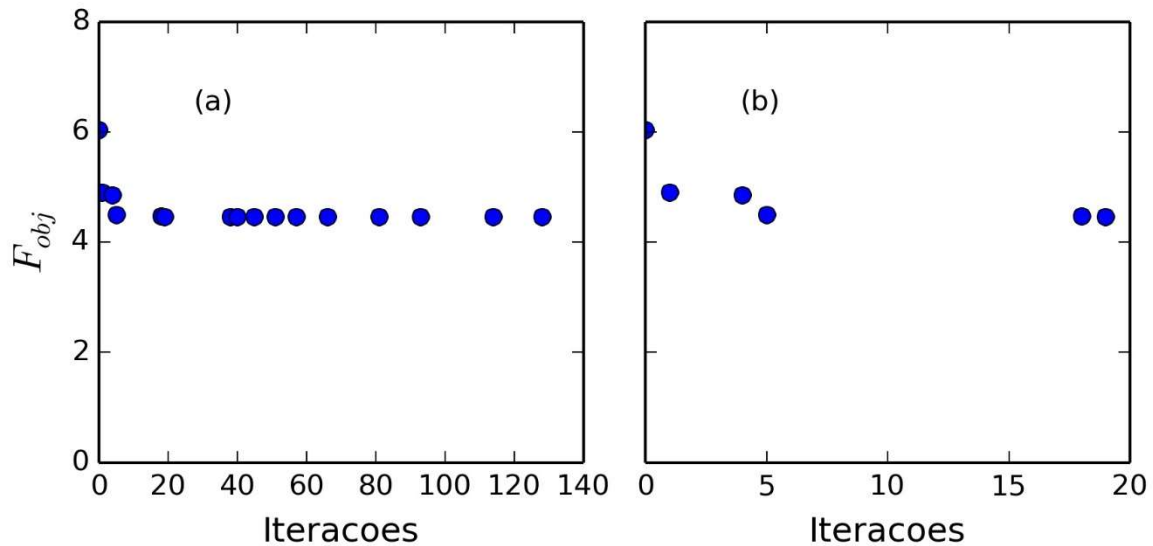


Figura 5.16: Evolução da melhor posição da função objetivo para a estimação com $P_{r,t}$.

O modelo claramente não consegue ajustar os valores de $P_{r,b}$ e a $P_{r,t}$ estimada não é adequada para todos os pontos experimentais, como ilustrado na Figura 5.17. Apesar disso, os resultados obtidos apresentam correlações menores entre os parâmetros, o que é representado pelas regiões de confiança menos correlacionadas e mais abrangentes do que os resultados anteriores (Figura 5.18). A Tabela 13 mostra o teste F entre o modelo original e o modelo estimado. Os parâmetros originais representam melhor a $P_{r,b}$, enquanto os parâmetros estimados representam melhor a $P_{r,t}$.

Tabela 11: Teste Chi Quadrado do modelo Meglio estimado com a pressão $P_{r,t}$.

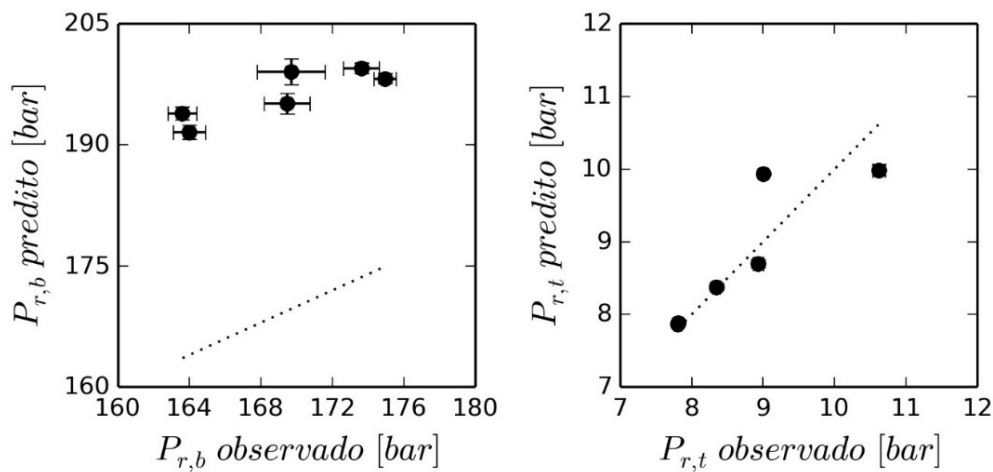
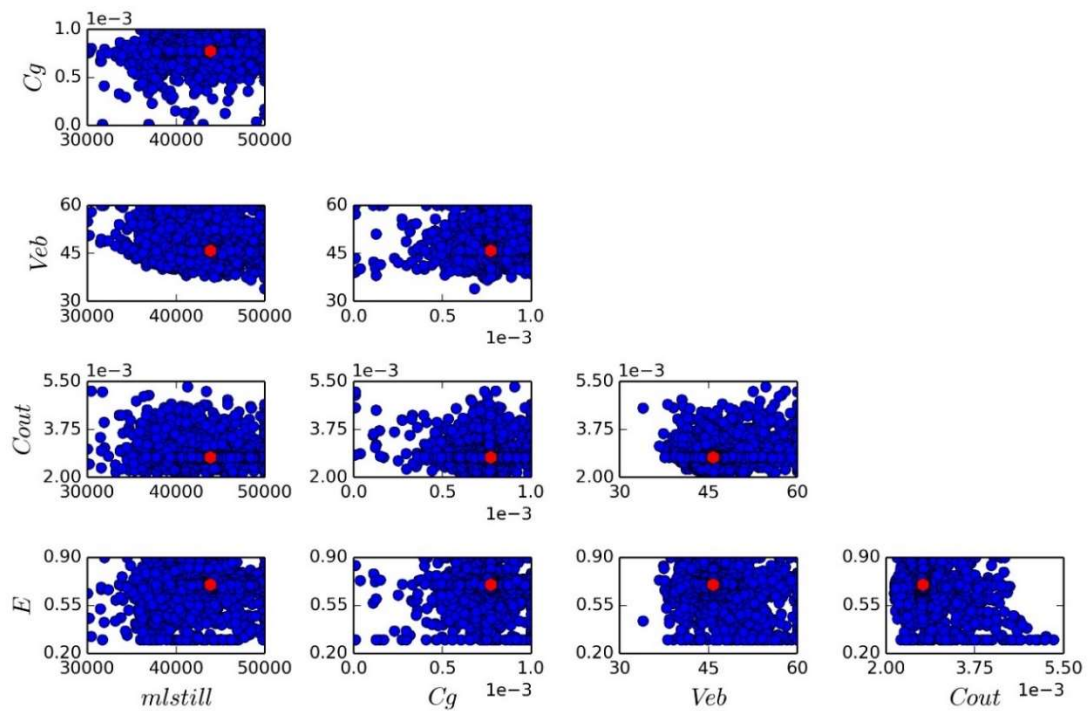
$\alpha = 95\%, \nu = 1$	$\chi_{\frac{1-\alpha}{2}, \nu}^2$	$F_{obj} = WLS$	$\chi_{\frac{1+\alpha}{2}, \nu}^2$
$P_{r,b}$	0,001	5.037,403	5,024
$P_{r,t}$		1,340	

Tabela 12: Matriz de correlações do modelo Meglio estimado com a pressão $P_{r,t}$.

	$m_{l,still} [kg]$	$C_g [m \cdot s]$	$V_{eb} [m^3]$	$C_{out} [m^2]$	$\epsilon [-]$
$m_{l,still} [kg]$	1	0.867379	-0.71949	-0.91368	0.358474
$C_g [m \cdot s]$	0.867379	1	-0.8786	-0.9661	0.757863
$V_{eb} [m^3]$	-0.71949	-0.8786	1	0.857801	-0.69855
$C_{out} [m^2]$	-0.91368	-0.9661	0.857801	1	-0.58781
$\epsilon [-]$	0.358474	0.757863	-0.69855	-0.58781	1

Tabela 13: Teste F do modelo Meglio estimado com a pressão $P_{r,t}$.

$\alpha = 95\%, \nu = 1$	$F_{\frac{1-\alpha}{2}, \nu, \nu}$	$\frac{\hat{\sigma}_1^2}{\hat{\sigma}_2^2}$	$F_{\frac{1+\alpha}{2}, \nu, \nu}$
$P_{r,b}$ estimada (1)	1,001	25,028	6,024
$P_{r,b}$ original (2)			
$P_{r,t}$ estimada (1)	0,181		
$P_{r,t}$ original (2)			

Figura 5.17: Valores preditos e valores observados de $P_{r,b}$ e $P_{r,t}$ do modelo Meglio estimado com $P_{r,t}$.Figura 5.18: Regiões de confiança de máxima verossimilhança para o modelo Meglio estimado com $P_{r,t}$. Valores ótimos são representados em vermelho.

5.2.4 Estimação com base em $P_{r,b}$ e $P_{r,t}$

De modo a aumentar a performance dos parâmetros estimados, incluiu-se como referência tanto $P_{r,b}$ quanto $P_{r,t}$ na mesma estimação. O cálculo da função objetivo foi modificado da forma:

$$F_{obj} = WLS_{r,b} + WLS_{r,t} \quad (5.3)$$

onde F_{obj} é a função objetivo e $WLS_{r,b}$ é a função de Mínimos Quadrados Ponderados aplicada aos dados de $P_{r,b}$. $WLS_{r,t}$ é a função de Mínimos Quadrados Ponderados aplicada aos dados de $P_{r,t}$, considerando uma variância unitária para os experimentos porque as variâncias da $P_{r,t}$ são muito menores do que da $P_{r,b}$.

Essa modificação pode ser interpretada como a inclusão de mais pontos experimentais ao conjunto de dados. Contudo, considere um caso experimental hipotético em que são medidas as frações molares x_1 e x_2 de uma mistura binária qualquer. Como x_1 e x_2 estão matematicamente correlacionados, tem-se de fato duas medidas experimentais de uma única variável independente, e não duas medidas de duas variáveis como pode parecer à primeira vista.

Analogamente, se os parâmetros do modelo Meglio fossem capazes de representar adequadamente os valores de $P_{r,b}$ e $P_{r,t}$ simultaneamente para qualquer resultado da otimização, a modificação da função objetivo seria redundante. Porém, como observado nos resultados anteriores, a referência utilizada tem um impacto profundo nas capacidades preditivas do modelo.

Isso levou à investigação da estimação com mais de um conjunto de dados. A Figura 5.19 ilustra os resultados obtidos. A Figura 5.20 mostra a evolução da função objetivo. Os coeficientes de correlação obtidos foram de $\rho_{r,b}^m = 0,878$ e $\rho_{r,t}^m = 0,987$. As Tabela 14 e Tabela 15 mostram os resultados do teste Chi Quadrado e do teste F, respectivamente. O teste Chi Quadrado foi feito considerando a função objetivo modificada, enquanto o teste F levou em conta os resultados de $P_{r,b}$ e $P_{r,t}$ separadamente.

O teste Chi Quadrado recomenda o uso do modelo e o teste F mostra que os parâmetros obtidos são mais adequados do que os parâmetros originais para $P_{r,b}$ e para $P_{r,t}$. Fica evidente que a inclusão dos dois conjuntos de dados auxiliou na obtenção de parâmetros mais representativos para as duas pressões. Além disso, as Tabela 16 e Tabela 17 mostram uma significativa redução na correlação entre os parâmetros. A Figura 5.21 ilustra como os resultados dos valores preditos estão mais próximos do eixo dos valores observados para $P_{r,b}$ e $P_{r,t}$. Em contrapartida, a Figura 5.22 mostra como também diminuíram as regiões de confiança para os parâmetros.

Esse resultado é de suma importância. Em seu artigo, Di Meglio *et al.* (2012) estudaram quatro possíveis estruturas de controle para um sistema *pipeline-riser*. A solução mais estável é baseada na diferença de pressão entre $P_{r,b}$ e $P_{r,t}$, aplicando uma lei de controle que utiliza a estimação das pressões dada pelo modelo. Portanto, fica claro que um bom ajuste das pressões apresentará melhores resultados para a estrutura de controle.

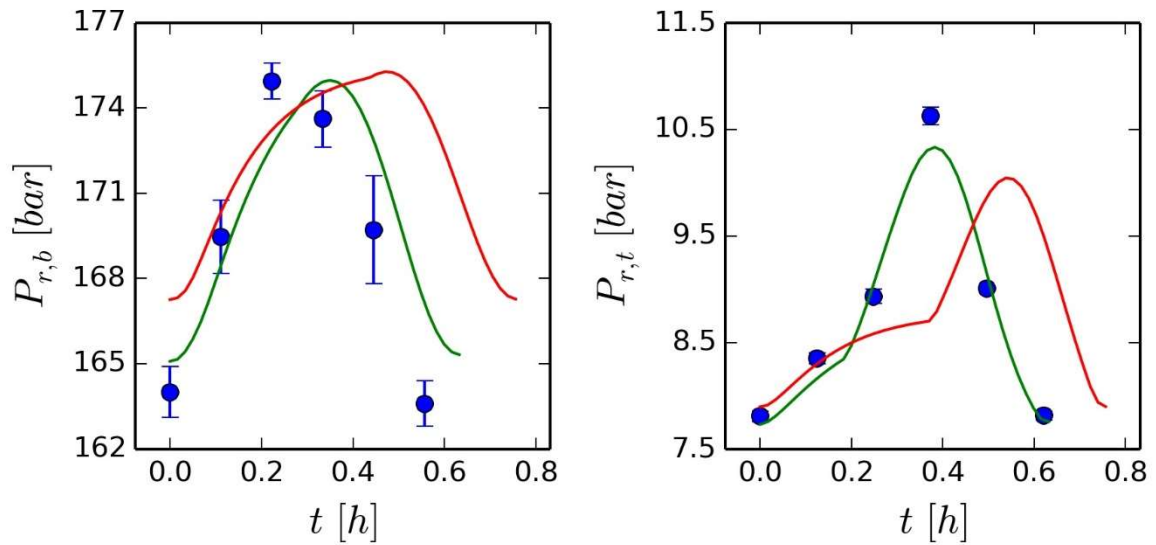


Figura 5.19: Ciclos representativos: dados originais (azul); modelo Meglio original (vermelho); modelo Meglio estimado com base nas pressões $P_{r,b}$ e $P_{r,t}$ (verde).

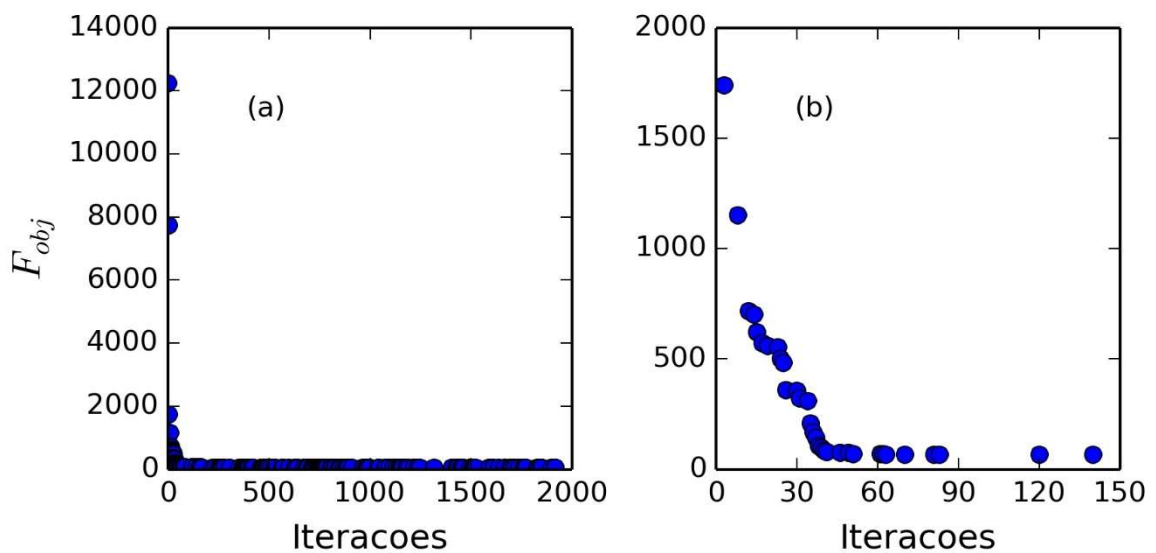


Figura 5.20: Evolução da melhor posição da função objetivo para a estimação com $P_{r,b}$ e $P_{r,t}$.

Tabela 14: Teste Chi Quadrado para o modelo Meglio estimado com as pressões $P_{r,b}$ e $P_{r,t}$.

$\alpha = 95\%, \nu = 7$	$\chi^2_{\frac{1-\alpha}{2}, \nu}$	F_{obj}	$\chi^2_{\frac{1+\alpha}{2}, \nu}$
$P_{r,b}$ e $P_{r,t}$	24,433	39,217	59,342

Tabela 15: Teste F do modelo Meglio estimado com as pressões $P_{r,b}$ e $P_{r,t}$.

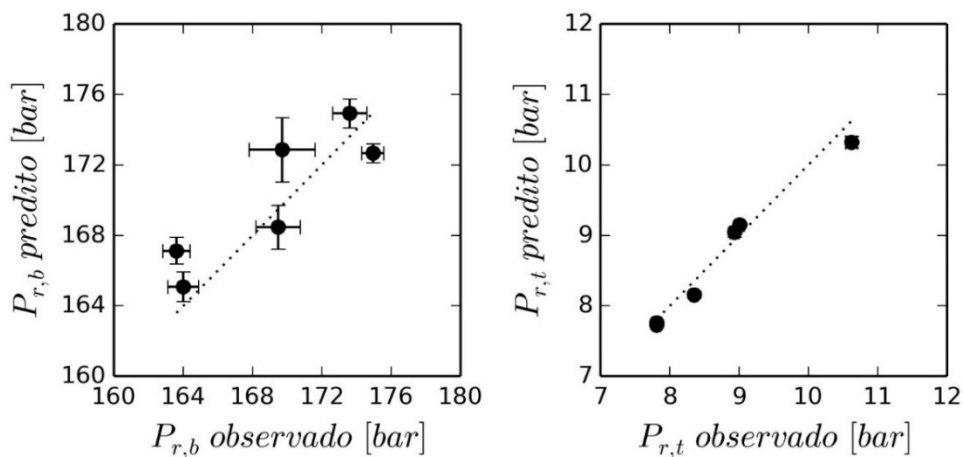
$\alpha = 95\%, \nu = 7$	$F_{\frac{1-\alpha}{2}, \nu, \nu}$	$\frac{\hat{\sigma}_1^2}{\hat{\sigma}_2^2}$	$F_{\frac{1+\alpha}{2}, \nu, \nu}$
$P_{r,b}$ estimada (1)	0,200	0,194	4,995
$P_{r,b}$ original (2)			
$P_{r,t}$ estimada (1)		0,023	
$P_{r,t}$ original (2)			

Tabela 16: Matriz de correlações para as medidas de $P_{r,b}$ do modelo Meglio estimado com as pressões $P_{r,b}$ e $P_{r,t}$.

	$m_{l,still} [kg]$	$C_g [m \cdot s]$	$V_{eb} [m^3]$	$C_{out} [m^2]$	$\epsilon [-]$
$m_{l,still} [kg]$	1	-0.99311	-0.28624	0.99908	0.603872
$C_g [m \cdot s]$	-0.99311	1	0.391658	-0.99289	-0.51642
$V_{eb} [m^3]$	-0.28624	0.391658	1	-0.27962	0.504617
$C_{out} [m^2]$	0.99908	-0.99289	-0.27962	1	0.601391
$\epsilon [-]$	0.603872	-0.51642	0.504617	0.601391	1

Tabela 17: Matriz de correlações para as medidas de $P_{r,t}$ do modelo Meglio estimado com as pressões $P_{r,b}$ e $P_{r,t}$.

	$m_{l,still} [kg]$	$C_g [m \cdot s]$	$V_{eb} [m^3]$	$C_{out} [m^2]$	$\epsilon [-]$
$m_{l,still} [kg]$	1	-0.99181	-0.06739	0.110723	-0.81511
$C_g [m \cdot s]$	-0.99181	1	0.172231	-0.22856	0.834784
$V_{eb} [m^3]$	-0.06739	0.172231	1	-0.63925	0.439121
$C_{out} [m^2]$	0.110723	-0.22856	-0.63925	1	-0.28884
$\epsilon [-]$	-0.81511	0.834784	0.439121	-0.28884	1

Figura 5.21: Valores preditos e valores observados de $P_{r,b}$ e $P_{r,t}$ do modelo Meglio estimado com $P_{r,b}$ e $P_{r,t}$.

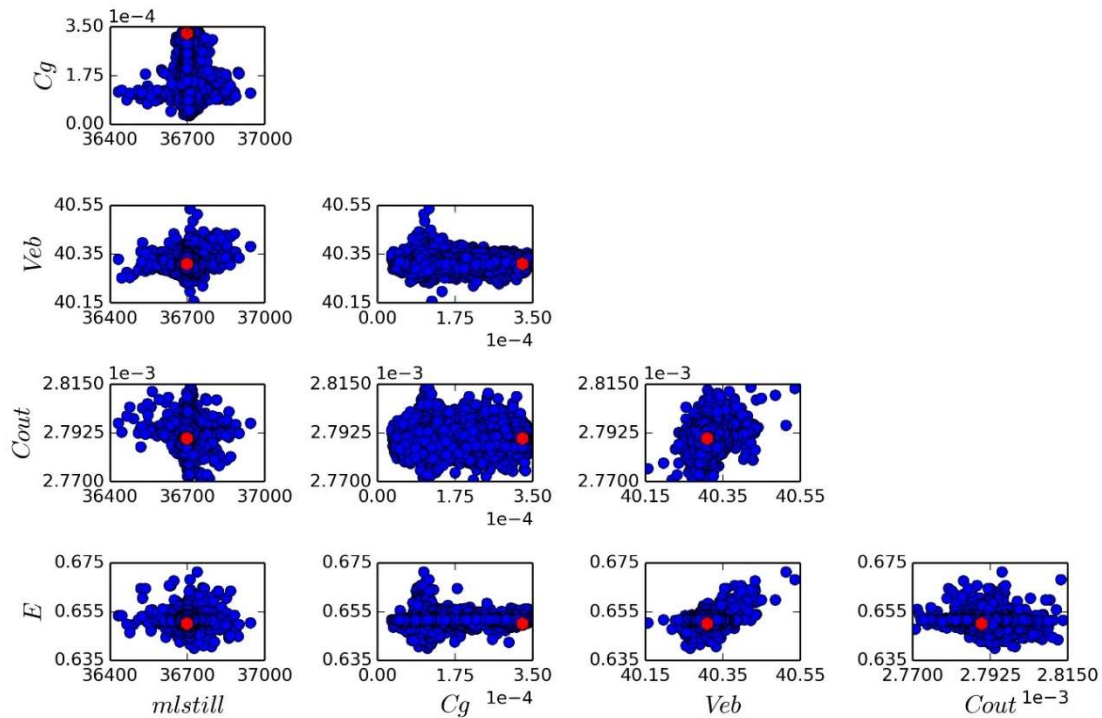


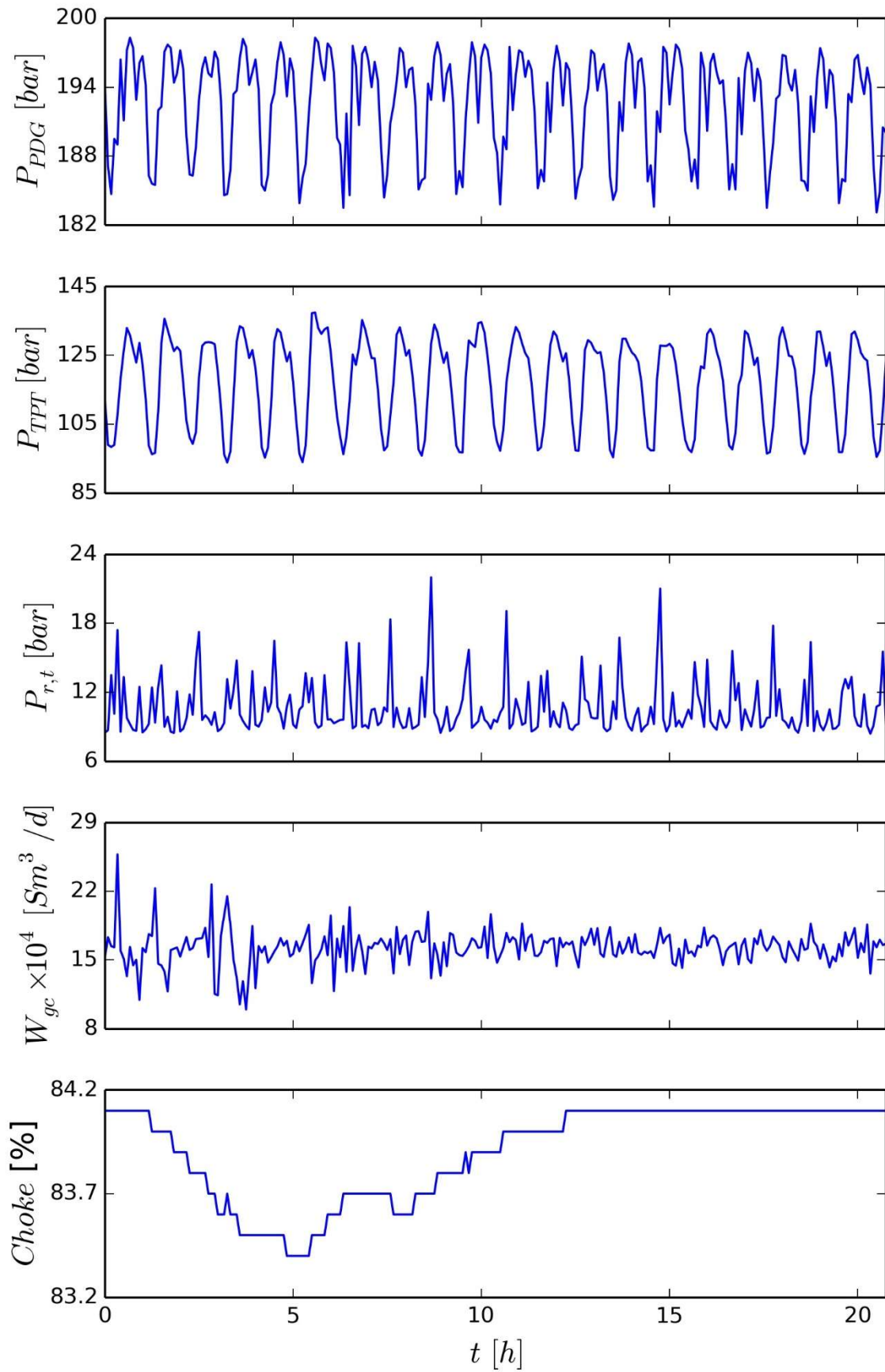
Figura 5.22: Regiões de confiança de máxima verossimilhança para o modelo Meglio estimado com $P_{r,b}$ e $P_{r,t}$. Valores ótimos são representados em vermelho.

5.3 Modelo FOWM com dados reais

Os parâmetros do modelo FOWM foram estimados com base nos dados do poço *Real Well A*, apresentados no artigo de Diehl *et al.* (2017) e ilustrados na Figura 5.23. Considera-se uma abertura média de válvula de $z = 84\%$ e uma vazão média de entrada de *gas-lift* de $w_{gc} = 1,406$ kg/s.

Alcançar 2000 iterações da função objetivo leva em torno de 24 h. O tempo computacional com o *solver* Radau5ODE para um conjunto de parâmetros com fácil convergência do modelo é de aproximadamente 870 ms. O tempo para o *solver* CVode, com passo mínimo de $h = 10^{-4}$, nas mesmas condições, é de 1,5 s. Ambos os tempos foram medidos com a função `%%timeit` do Python.

Primeiramente, foram estimados três conjuntos de parâmetros baseados na P_{PDG} considerando três cenários para o parâmetro ω_u , que modifica multiplicativamente o volume da tubulação *subsea*, V_{SS} . O estudo de diferentes cenários possibilita a melhor compreensão dos efeitos do parâmetro sobre a estimação. Em seguida, foram estimados parâmetros do modelo com base nas pressões P_{PDG} , P_{TPT} (que no modelo é dada por $P_{r,b}$) e $P_{r,t}$. Esses parâmetros foram comparados com os obtidos por Diehl *et al.* (2017), Apio (2017) e Huffner (2018), que também estimaram parâmetros para o modelo FOWM utilizando as mesmas três pressões.

Figura 5.23: Dados de referência para o *Real Well A*.

5.3.1 Influência do parâmetro ω_u

Como o parâmetro ω_u representa a modificação de um parâmetro geométrico do sistema, é interessante estudar sua influência nos outros parâmetros. No primeiro cenário, ω_u é fixo e não é incluso na estimação de parâmetros. Isso representa um custo computacional menor e, possivelmente, a diminuição de correlações entre os parâmetros. No segundo cenário, os limites de busca do algoritmo de otimização para ω_u são definidos entre 1 e 1,8 para avaliar o aumento do volume real de tubulação na resposta do modelo. O terceiro cenário estende o limite superior até 2,2 somente para demonstrar a influência do limite desse parâmetro.

A Tabela 18 mostra a comparação entre os parâmetros originais e os parâmetros estimados com a nova função objetivo. A Figura 5.24 ilustra os ciclos médios dos dados, dos parâmetros originais e dos novos parâmetros. É preciso notar que não foi possível reproduzir o modelo como apresentado no artigo provavelmente devido à diferença de integrador utilizado. Os dados de referência apresentam um comportamento similar, mas uma média maior do que o visto no artigo; os dados originais não estavam disponíveis. Trata-se do mesmo poço, mas provavelmente em condições diferentes das utilizadas na publicação. O resultado apresentado em vermelho na Figura 5.24 foi extraído com o *software* Pega Ponto.

Tabela 18: Comparação entre resultados originais e resultados obtidos com a nova metodologia.

Parâmetro	Real Well A			
	Diehl <i>et al.</i>	Cenário 1 $\omega_u = 1$	Cenário 2 $1 \leq \omega_u \leq 1,8$	Cenário 3 $1 \leq \omega_u \leq 2,2$
$m_{l,still} [kg]$	$6.222 \cdot 10^1$	$2,014 \cdot 10^4$	$2,096 \cdot 10^4$	$1,794 \cdot 10^4$
$C_g [m \cdot s]$	$1.137 \cdot 10^{-3}$	$1,000 \cdot 10^{-3}$	$1,000 \cdot 10^{-3}$	$5,000 \cdot 10^{-3}$
$C_{out} [m^2]$	$2.039 \cdot 10^{-3}$	$9,066 \cdot 10^{-4}$	$1,300 \cdot 10^{-3}$	$1,584 \cdot 10^{-3}$
$V_{eb} [m^3]$	$6.098 \cdot 10^1$	$1,500 \cdot 10^2$	$1,145 \cdot 10^2$	$9,500 \cdot 10^1$
$\epsilon [-]$	$1.545 \cdot 10^{-1}$	$4,691 \cdot 10^{-1}$	$7,000 \cdot 10^{-1}$	$7,430 \cdot 10^{-1}$
$K_w [m^2]$	$6.876 \cdot 10^{-4}$	$1,000 \cdot 10^{-3}$	$1,120 \cdot 10^{-3}$	$1,010 \cdot 10^{-3}$
$K_a [m^2]$	$2.293 \cdot 10^{-5}$	$4,940 \cdot 10^{-2}$	$9,883 \cdot 10^{-3}$	$5,000 \cdot 10^{-4}$
$K_r [kg/s]$	$1.269 \cdot 10^2$	$1,200 \cdot 10^2$	$1,000 \cdot 10^2$	$1,079 \cdot 10^2$
$\omega_u [-]$	$2.780 \cdot 10^0$	$1,000 \cdot 10^0$	$1,800 \cdot 10^0$	$2,200 \cdot 10^0$

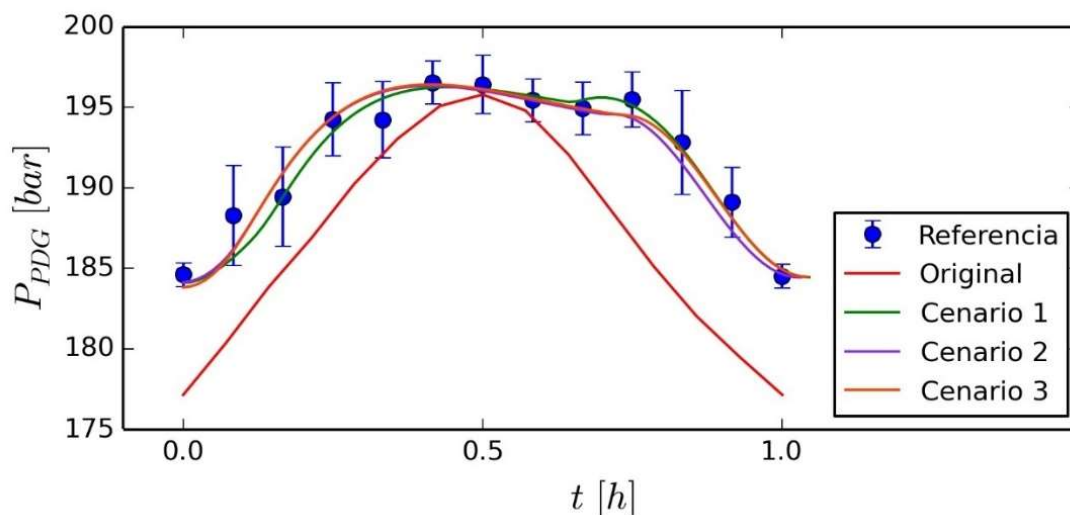


Figura 5.24: Comparação entre ciclos obtidos entre os cenários para o modelo FOWM.

Os testes Chi Quadrado para os resultados são mostrados na Tabela 19. A Tabela 20 mostra os coeficientes de correlação obtidos. Como pode ser observado, os parâmetros originais apresentam tendência similar à referência, mas o uso do modelo não é recomendado pelo teste. Já os parâmetros estimados se adequam bem à tendência e o modelo satisfaz o teste.

O teste F para os cenários mostra que não existe diferença significativa entre os ajustes obtidos (Tabela 21). Assim, os resultados obtidos em cada cenário são igualmente capazes de representar os dados de referência. Já o teste F entre os cenários e os resultados originais mostra que os parâmetros estimados apresentam melhores resultados do que os originais (Tabela 22), embora isso fosse previsto dado que a referência não é a mesma para os dois modelos.

Tabela 19: Teste Chi Quadrado para o modelo FOWM baseado em dados reais.

$\alpha = 95\%, \nu = 4$	$\chi^2_{\frac{1-\alpha}{2}, \nu}$	$F_{obj} = WLS$	$\chi^2_{\frac{1+\alpha}{2}, \nu}$
Original		275,969	
Cenário 1	0,484	2,026	11,143
Cenário 2		2,260	
Cenário 3		2,148	

Tabela 20: Coeficientes de correlação obtidos do resultado original e dos cenários para o modelo FOWM.

Conjunto	ρ^m
Diehl et al.	0,917
Cenário 1	0,981
Cenário 2	0,970
Cenário 3	0,974

Tabela 21: Teste F para o modelo FOWM entre os cenários.

$\alpha = 95\%, \nu = 4$	$F_{\frac{1-\alpha}{2}, \nu, \nu}$	$\frac{\hat{\sigma}_1^2}{\hat{\sigma}_2^2}$	$F_{\frac{1+\alpha}{2}, \nu, \nu}$
Cenário 1 (1) e Cenário 2 (2)	0,104	0,896	9,604
Cenário 1 (1) e Cenário 3 (2)		0,943	

Tabela 22: Teste F para o modelo FOWM entre os cenários e o resultado original.

$\alpha = 95\%, \nu = 4$	$F_{\frac{1-\alpha}{2}, \nu, \nu}$	$\frac{\hat{\sigma}_1^2}{\hat{\sigma}_2^2}$	$F_{\frac{1+\alpha}{2}, \nu, \nu}$
Cenário 1 (1) e parâmetros originais (2)	0,104	0,007	9,604
Cenário 2 (1) e parâmetros originais (2)		0,008	
Cenário 3 (1) e parâmetros originais (2)		0,008	

As Tabela 23, Tabela 24 e Tabela 25 mostram as matrizes de correlações dos cenários 1, 2 e 3, respectivamente. Como não foi possível reproduzir o modelo original, também não é possível obter a matriz de correlações original. Fica evidente que o parâmetro ω_u tem uma alta influência nas correlações entre os parâmetros. O cenário 3, que apresenta o ω_u mais elevado, é também o que apresenta o maior número de correlações altas. Dado que não há diferença estatística entre os ajustes obtidos, a necessidade de inclusão de ω_u na estimação de parâmetros se torna questionável.

A Figura 5.25 mostra a comparação entre os valores preditos e os valores observados. Considera-se a variância de predição do modelo original como igual à variância experimental obtida, dado que representa a variância mínima associada à predição. Os erros observados para os resultados são justificados pela variância experimental e pela de predição. Já os erros relacionados ao modelo original simplesmente reforçam a ideia de que esse conjunto de parâmetros é inadequado.

A Figura 5.26, a Figura 5.27 e a Figura 5.28 mostram as regiões de confiança de máxima verossimilhança para os cenários 1, 2 e 3, respectivamente. O algoritmo de otimização tende a priorizar valores-limite dos parâmetros, como mostrado na Tabela 18. Isso gera regiões menores a menos que se faça uma varredura manual com alto custo computacional alto e foi desconsiderado. Assim, foram obtidos bons ajustes com a metodologia proposta.

A distinção entre os resultados dos cenários requer novos pontos experimentais. Para tanto, pode-se incluir dados de outras aberturas de válvula *choke*, ou dados de outras pressões de forma análoga à realizada para o modelo Meglio. Sem outros dados para comparação, os modelos obtidos possuem igual capacidade de predição da pressão P_{PDG} nas condições operacionais simuladas.

Tabela 23: Matriz de correlações do modelo FOWM para cenário 1.

	$m_{l,still}$	C_g	C_{out}	V_{eb}	ϵ	K_w	K_a	K_r
$m_{l,still}$	1	0.859505	-0.98947	0.757071	-0.76258	-0.46734	0.113646	-0.96956
C_g	0.859505	1	-0.85035	0.320865	-0.97459	-0.85275	0.187543	-0.95357
C_{out}	-0.98947	-0.85035	1	-0.75932	0.774462	0.462402	-0.08661	0.968955
V_{eb}	0.757071	0.320865	-0.75932	1	-0.17918	0.219017	-0.06229	-0.58616
ϵ	-0.76258	-0.97459	0.774462	-0.17918	1	0.912277	-0.17493	0.896534
K_w	-0.46734	-0.85275	0.462402	0.219017	0.912277	1	-0.21449	0.661371
K_a	0.113646	0.187543	-0.08661	-0.06229	-0.17493	-0.21449	1	-0.14174
K_r	-0.96956	-0.95357	0.968955	-0.58616	0.896534	0.661371	-0.14174	1

Tabela 24: Matriz de correlações do modelo FOWM para cenário 2.

	$m_{l,still}$	C_g	C_{out}	V_{eb}	ϵ	K_w	K_a	K_r	ω_u
$m_{l,still}$	1	0.191502	0.553975	-0.78354	0.668743	-0.34176	0.555433	-0.97621	0.750607
C_g	0.191502	1	-0.26591	0.157149	-0.56543	-0.96791	0.56933	0.001676	-0.48368
C_{out}	0.553975	-0.26591	1	-0.93183	0.801253	0.015565	0.474695	-0.58317	0.725134
V_{eb}	-0.78354	0.157149	-0.93183	1	-0.86759	0.080935	-0.51713	0.81789	-0.85582
ϵ	0.668743	-0.56543	0.801253	-0.86759	1	0.377638	0.132467	-0.78564	0.982712
K_w	-0.34176	-0.96791	0.015565	0.080935	0.377638	1	-0.70937	0.150029	0.31211
K_a	0.555433	0.56933	0.474695	-0.51713	0.132467	-0.70937	1	-0.43091	0.156794
K_r	-0.97621	0.001676	-0.58317	0.81789	-0.78564	0.150029	-0.43091	1	-0.86473
ω_u	0.750607	-0.48368	0.725134	-0.85582	0.982712	0.31211	0.156794	-0.86473	1

Tabela 25: Matriz de correlações do modelo FOWM para cenário 3.

	$m_{l,still}$	C_g	C_{out}	V_{eb}	ϵ	K_w	K_a	K_r	ω_u
$m_{l,still}$	1	-0.23863	-0.82479	0.803583	-0.85388	-0.95853	-0.02192	-0.47475	-0.87555
C_g	-0.23863	1	-0.29716	0.234233	-0.18802	0.185886	0.770638	0.448932	-0.09593
C_{out}	-0.82479	-0.29716	1	-0.97559	0.988923	0.806064	-0.26618	0.053929	0.965054
V_{eb}	0.803583	0.234233	-0.97559	1	-0.98666	-0.73596	0.09277	0.081546	-0.9843
ϵ	-0.85388	-0.18802	0.988923	-0.98666	1	0.823787	-0.13959	0.027372	0.991751
K_w	-0.95853	0.185886	0.806064	-0.73596	0.823787	1	-0.13669	0.540471	0.823545
K_a	-0.02192	0.770638	-0.26618	0.09277	-0.13959	-0.13669	1	-0.1576	-0.03563
K_r	-0.47475	0.448932	0.053929	0.081546	0.027372	0.540471	-0.1576	1	0.017865
ω_u	-0.87555	-0.09593	0.965054	-0.9843	0.991751	0.823545	-0.03563	0.017865	1

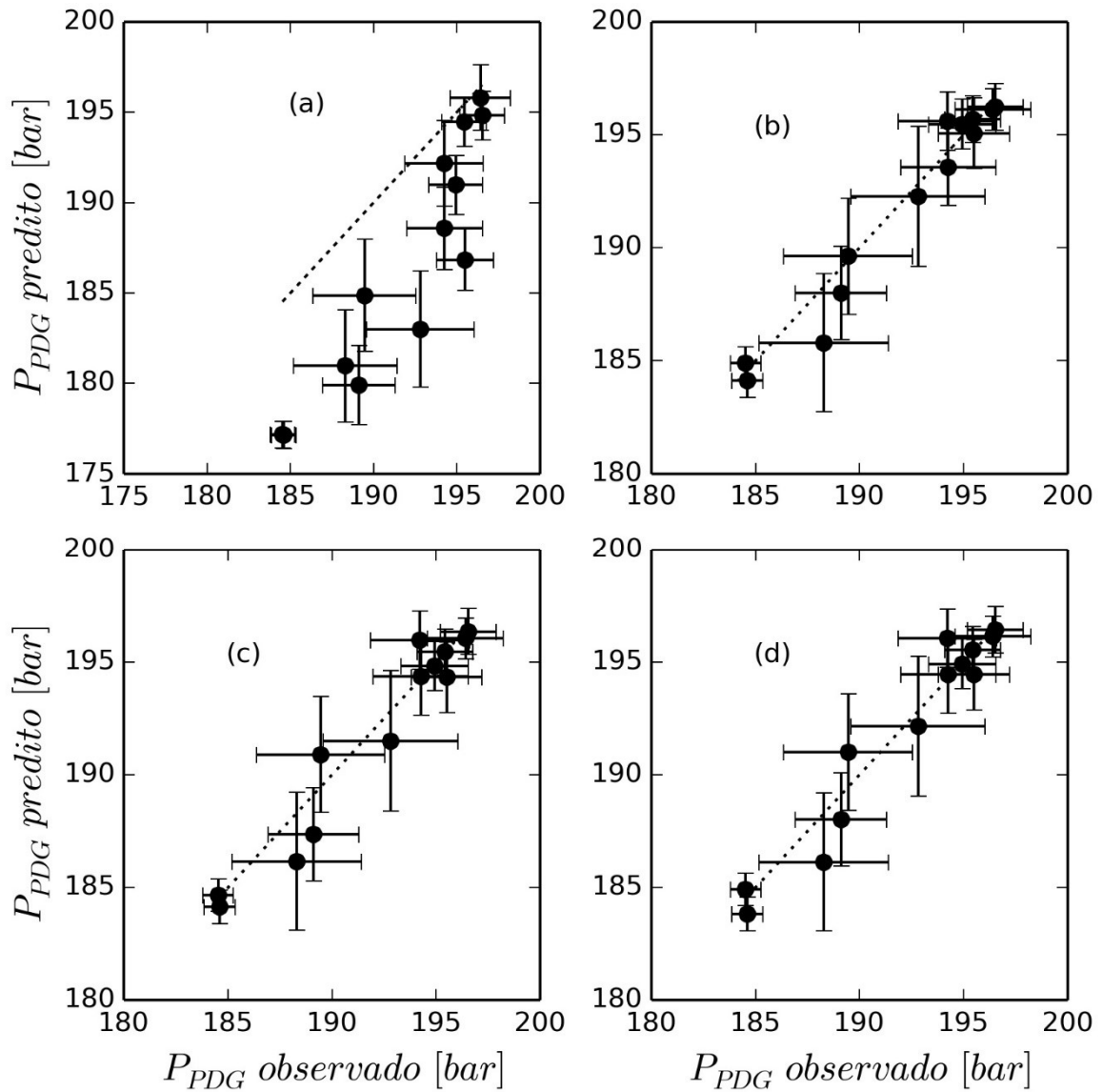


Figura 5.25: Valores preditos e observados da pressão P_{PDG} : (a) dados originais, (b) cenário 1, (c) cenário 2 e (d) cenário 3.

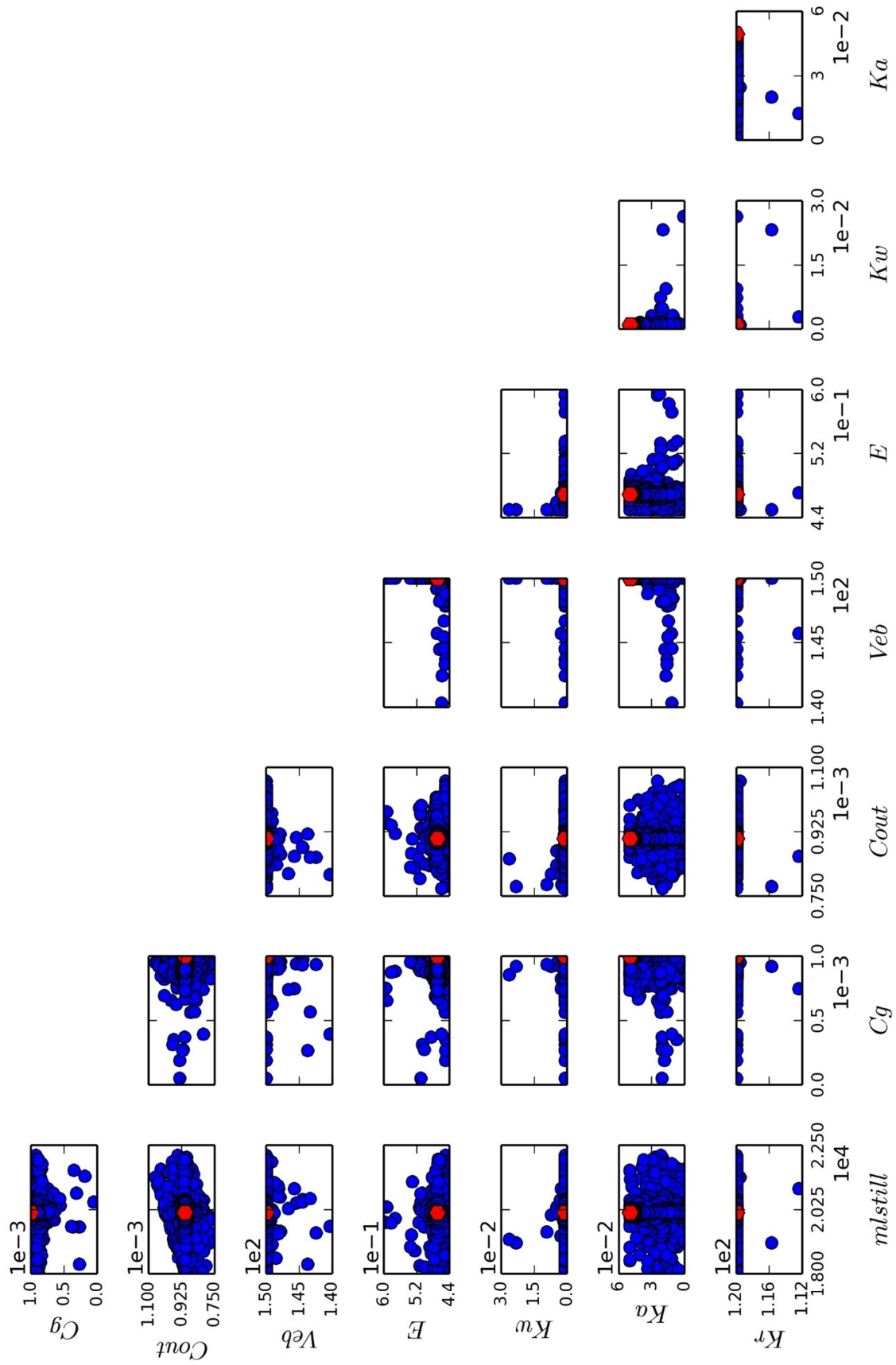


Figura 5.26: Regiões de máxima verossimilhança para o cenário 1.

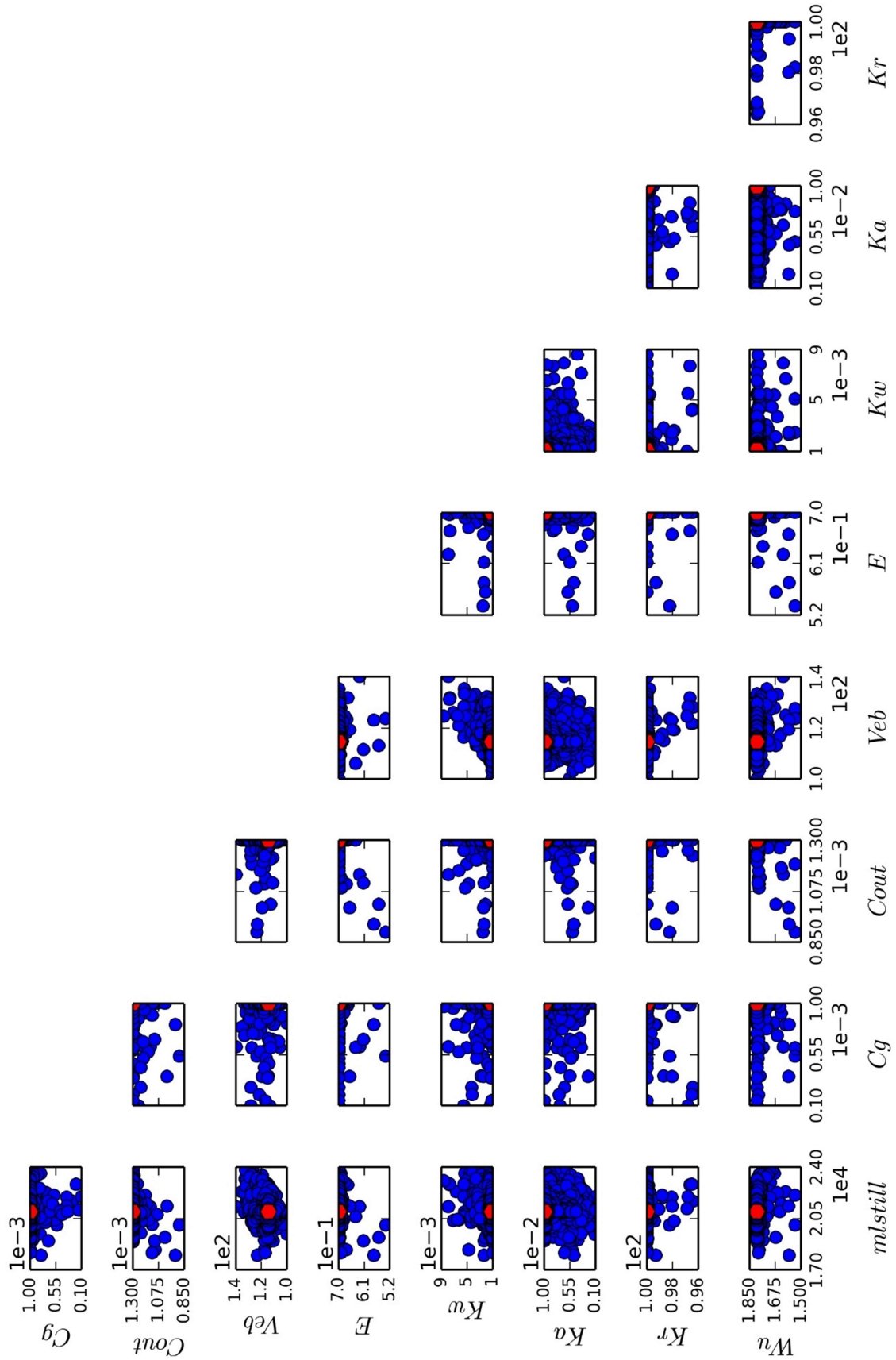


Figura 5.27: Regiões de máxima verossimilhança para o cenário 2.

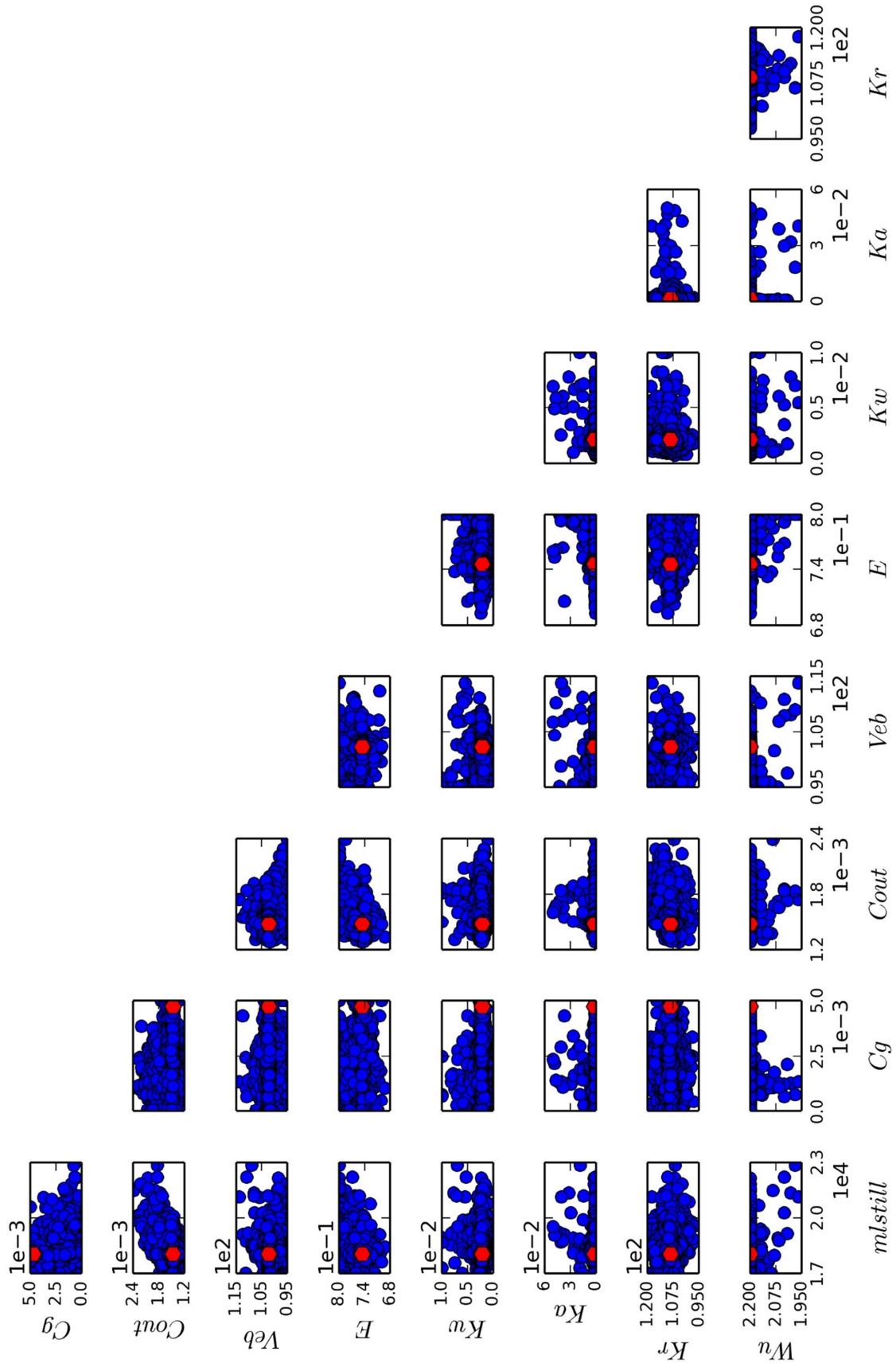


Figura 5.28: Regiões de máxima verossimilhança para o cenário 3.

5.3.2 Estimação de parâmetros com outras referências

Analisando a Figura 5.29, que mostra o comportamento das pressões P_{PDG} , P_{TPT} e $P_{r,t}$ dos modelos estimados com a P_{PDG} , é possível notar que o uso da P_{PDG} como única referência não garante a aderência do modelo às outras pressões. Assim, analogamente ao que foi feito com o modelo de Di Meglio et al. (2009), foram estimados os parâmetros do modelo FOWM utilizando três conjuntos de dados de referência: P_{PDG} , $P_{r,b}$ e $P_{r,t}$. Foi necessário considerar o início e o fim do ciclo da $P_{r,t}$ com base nos picos ao invés dos vales porque são mais facilmente identificados. Os resultados foram então comparados com os obtidos por Diehl et al. (2017), Apio (2017) e Hüffner (2018). A Figura 5.30 mostra a evolução da função objetivo, enquanto a Tabela 26 mostra a comparação entre os parâmetros obtidos.

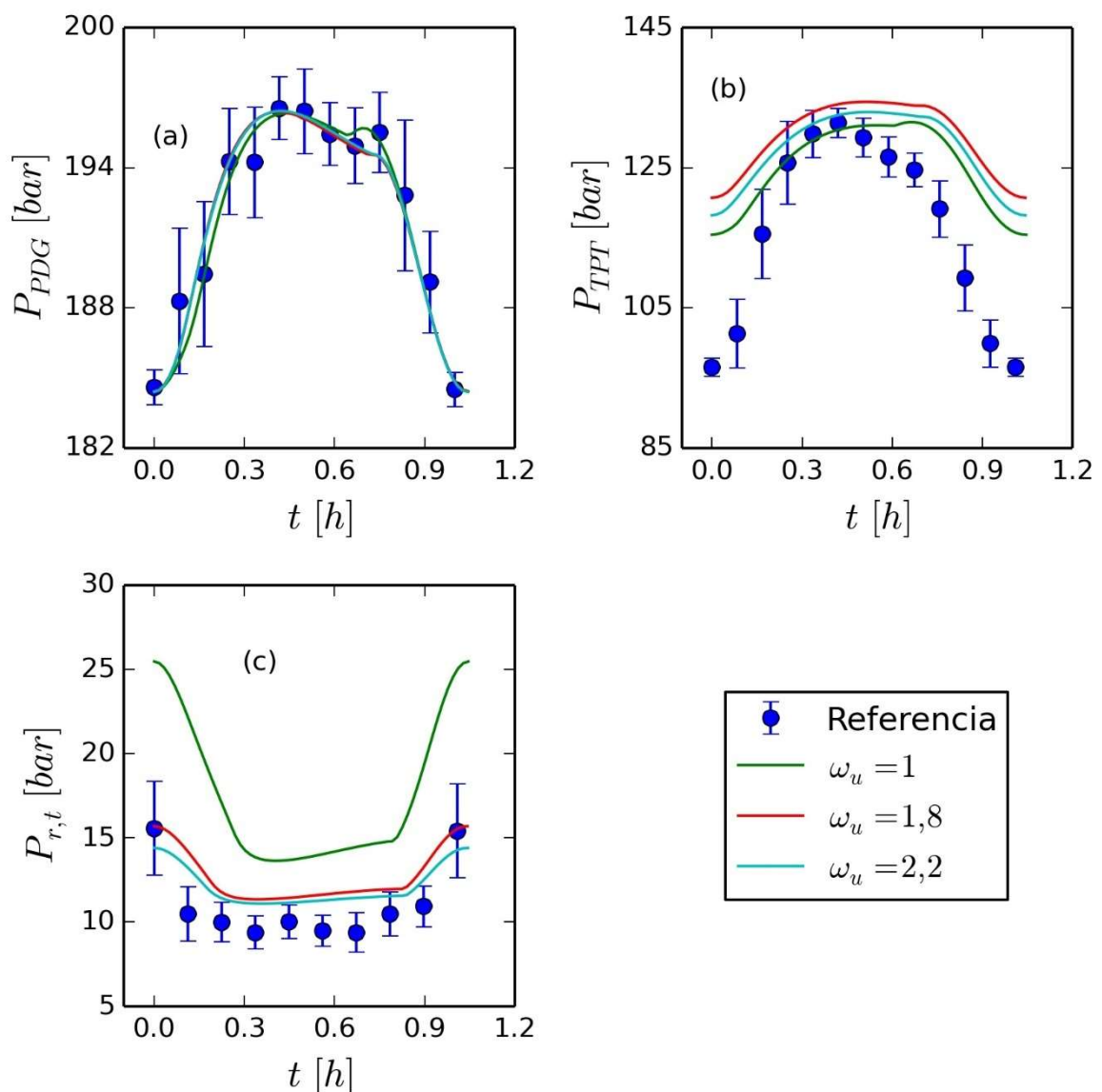


Figura 5.29: Ciclos médios das pressões (a) P_{PDG} , (b) P_{TPT} e (c) $P_{r,t}$, considerando a referência e os resultados de diferentes valores de ω_u . Os resultados foram obtidos utilizando somente P_{PDG} como referência.

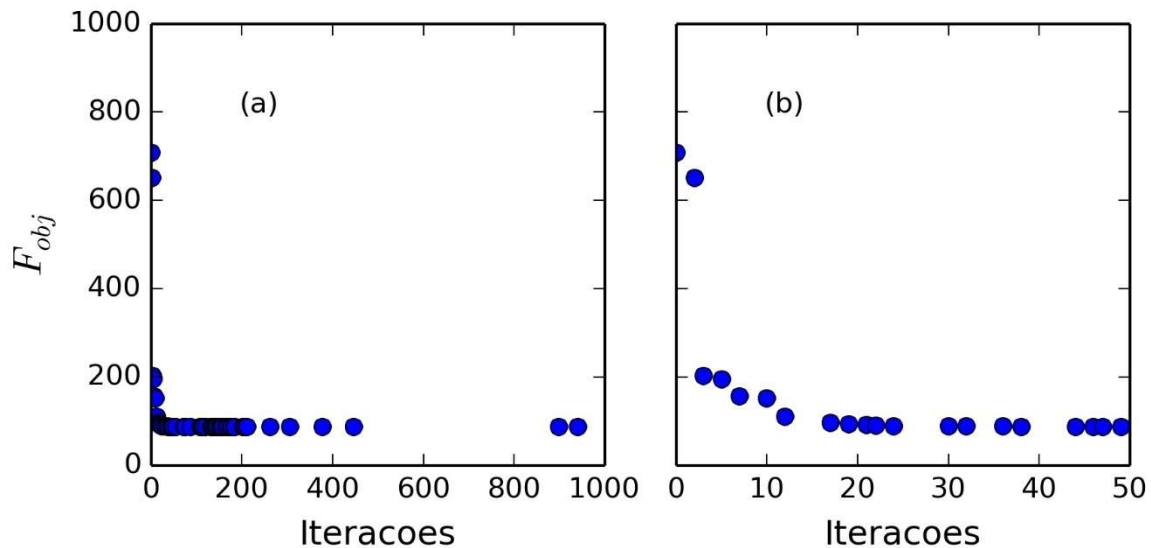


Figura 5.30: Evolução da função objetivo para a estimação do modelo FOWM com base em três pressões.

Tabela 26: Comparação entre parâmetros estimados do modelo FOWM para o *Real Well A*.

Parâmetro	Diehl et al. (2017)	Apio (2017)	Hüffner (2018)	Rodrigues
$m_{l,still}$ [kg]	$6,222 \cdot 10^1$	$7,183 \cdot 10^1$	$1,210 \cdot 10^3$	$2,020 \cdot 10^4$
C_g [m · s]	$1,137 \cdot 10^{-3}$	$1,329 \cdot 10^{-3}$	$5,080 \cdot 10^{-4}$	$1,000 \cdot 10^{-3}$
C_{out} [m ²]	$2,039 \cdot 10^{-3}$	$3,975 \cdot 10^{-3}$	$7,360 \cdot 10^{-3}$	$1,626 \cdot 10^{-3}$
V_{eb} [m ³]	$6,098 \cdot 10^1$	$7,047 \cdot 10^1$	$7,180 \cdot 10^1$	$1,500 \cdot 10^2$
ϵ [–]	$1,545 \cdot 10^{-1}$	$2,095 \cdot 10^{-1}$	$1,000 \cdot 10^{-1}$	$4,500 \cdot 10^{-1}$
K_w [m ²]	$6,876 \cdot 10^{-4}$	$5,936 \cdot 10^{-4}$	$2,740 \cdot 10^{-4}$	$4,958 \cdot 10^{-4}$
K_a [m ²]	$2,293 \cdot 10^{-5}$	$4,236 \cdot 10^{-5}$	$2,750 \cdot 10^{-4}$	$1,000 \cdot 10^{-3}$
K_r [kg/s]	$1,269 \cdot 10^2$	$2,470 \cdot 10^2$	$1,320 \cdot 10^2$	$1,200 \cdot 10^2$
ω_u [–]	$2,780 \cdot 10^0$	$1,877 \cdot 10^0$	$1,020 \cdot 10^0$	$1,000 \cdot 10^0$

Como o trabalho de Diehl et al. (2017) não apresenta os resultados do modelo para a $P_{r,t}$, foram utilizados somente dados simulados para a avaliação dos resultados. A Figura 5.31 mostra os ciclos médios das três pressões para a referência e os quatro conjuntos de parâmetros do modelo FOWM. Os parâmetros de Diehl et al., Apio e Hüffner apresentam baixa aderência (Tabela 27) e não são recomendados pelo teste Chi Quadrado (Tabela 28). Os parâmetros obtidos com a nova função objetivo, por outro lado, apresentam maior aderência, mas também não são recomendados pelo teste. Contudo, a representatividade dos dados é maior com os parâmetros estimados (Figura 5.32). Isso indica que há margem para melhorias no equacionamento do modelo.

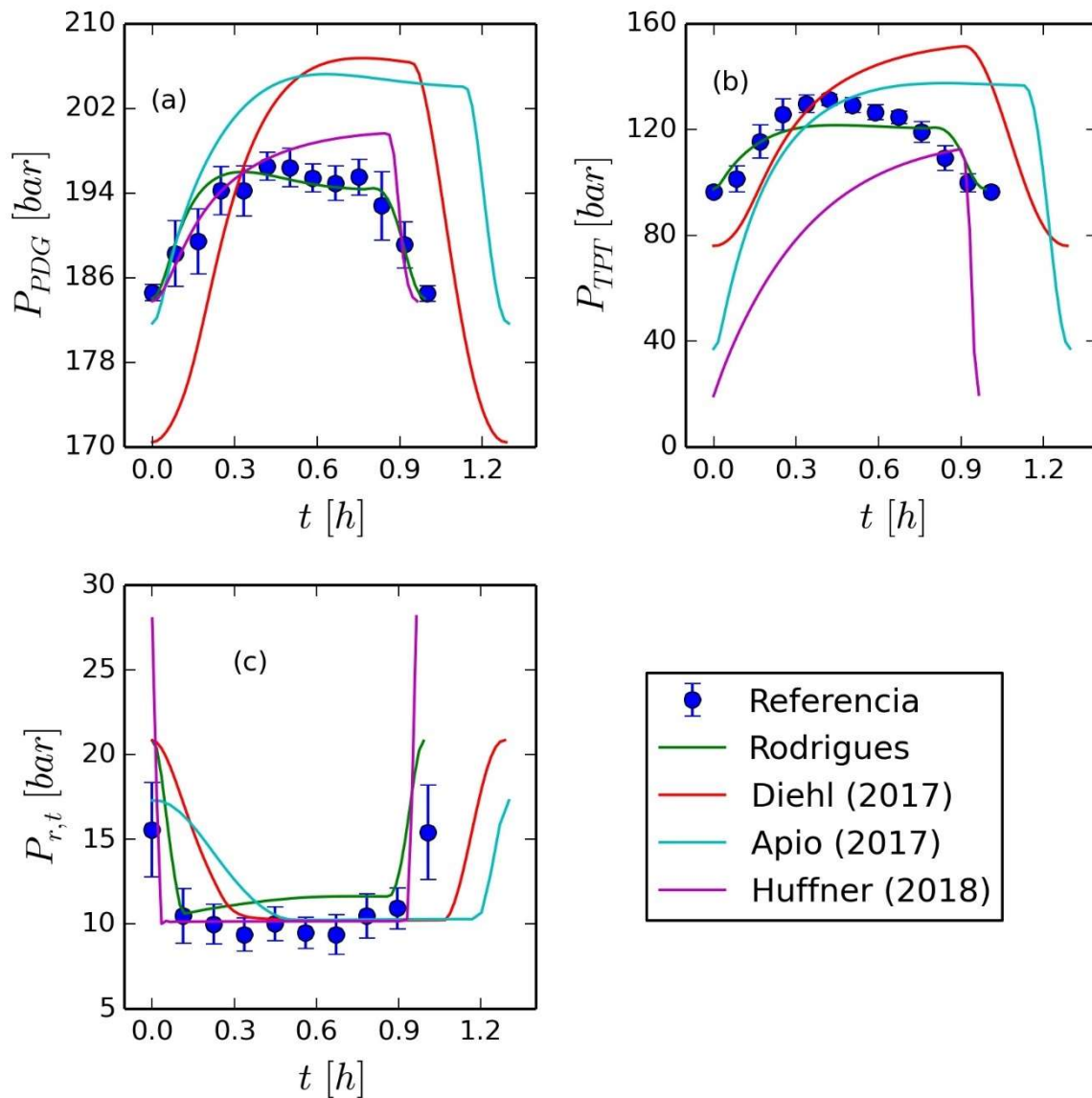


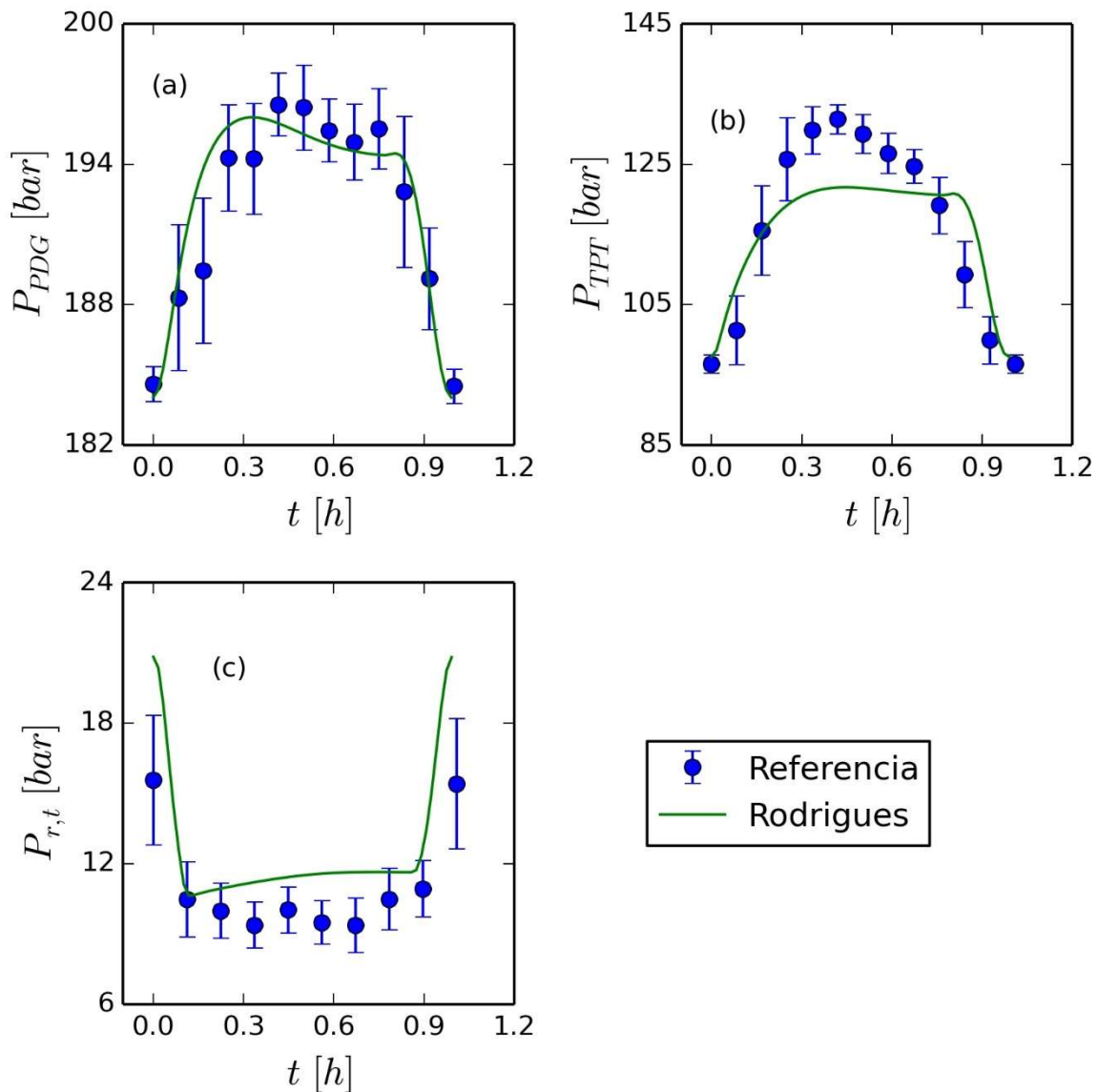
Figura 5.31: Comparação entre as pressões (a) P_{PDG} , (b) P_{TPT} e (c) $P_{r,t}$ para a referência e os diferentes conjuntos de parâmetros obtidos para o modelo FOWM.

Tabela 27: Coeficientes de correlação para os conjuntos de parâmetros estudados para o modelo FOWM.

ρ^m	P_{PDG}	P_{TPT}	$P_{r,t}$
Diehl et al. (2017)	0,573	0,358	0,498
Apio (2017)	0,625	0,433	0,337
Hüffner (2018)	0,774	0,480	0,696
Rodrigues	0,933	0,804	0,977

Tabela 28: Teste Chi Quadrado para o modelo FOWM estimado com três pressões.

$\alpha = 95\%, \nu = 27$	$\chi_{\frac{1-\alpha}{2}, \nu}^2$	$F_{obj} = WLS$	$\chi_{\frac{1+\alpha}{2}, \nu}^2$
Diehl et al. (2017)	14,573	3076,065	43,194
Apio (2017)		4497,095	
Hüffner (2018)		5170,191	
Rodrigues		86,873	

Figura 5.32: Comparação entre as pressões (a) P_{PDG} , (b) P_{TPT} e (c) $P_{r,t}$ para a referência e o conjunto de parâmetros estimado para o modelo FOWM.

O teste F é mostrado na Tabela 29. As Tabela 30, Tabela 31 e Tabela 32 mostram as matrizes de correlação das pressões P_{PDG} , P_{TPT} e $P_{r,t}$, respectivamente. A Figura 5.33 mostra a comparação entre valores preditos e observados. Por fim, a Figura 5.34 mostra as regiões de confiança de máxima verossimilhança para os parâmetros considerando a estimação com a nova função objetivo.

Para cada referência foram observados conjuntos de até três parâmetros que não apresentam correlações altas ($|\rho_{ij}| < 0,90$). No caso da P_{PDG} , são os parâmetros C_g , C_{out} e K_a ; para a P_{TPT} , os parâmetros $m_{l,still}$, V_{eb} e K_w ; para a $P_{r,t}$, os parâmetros C_{out} e K_r . Isso significa, por exemplo, que utilizar a P_{PDG} como referência facilita a estimação de C_g , C_{out} e K_a .

Além disso, é possível afirmar com 95% de confiança que os parâmetros estimados ajustam melhor o modelo aos dados do que os trabalhos anteriores. O modelo também nem sempre consegue justificar os erros experimentais observados, embora os valores preditos sejam próximos dos valores observados na maior parte dos pontos experimentais.

Tabela 29: Teste F para o modelo FOWM entre os diferentes conjuntos de parâmetros obtidos para o *Real Well A*.

$\alpha = 95\%, \nu = 27$	$F_{\frac{1-\alpha}{2}, \nu, \nu}$	$\frac{\hat{\sigma}_1^2}{\hat{\sigma}_2^2}$	$F_{\frac{1+\alpha}{2}, \nu, \nu}$
Rodrigues (1) e Diehl et al. (2017) (2)		0,017	
Rodrigues (1) e Apio (2017) (2)	0,463	0,019	2,161
Rodrigues (1) e Hüffner (2018) (2)		0,028	

Tabela 30: Matriz de correlações da P_{PDG} do modelo FOWM estimado com três pressões.

	$m_{l,still}$	C_g	C_{out}	V_{eb}	ϵ	K_w	K_a	K_r
$m_{l,still}$	1	-0.07251	-0.12838	0.980913	-0.90167	-0.92722	-0.1755	-0.93949
C_g	-0.07251	1	0.494797	-0.06895	0.104728	-0.00352	0.025953	0.087619
C_{out}	-0.12838	0.494797	1	-0.22102	0.418463	0.383015	-0.21042	0.397949
V_{eb}	0.980913	-0.06895	-0.22102	1	-0.96353	-0.96805	-0.28179	-0.97877
ϵ	-0.90167	0.104728	0.418463	-0.96353	1	0.979945	0.326742	0.988839
K_w	-0.92722	-0.00352	0.383015	-0.96805	0.979945	1	0.176479	0.994814
K_a	-0.1755	0.025953	-0.21042	-0.28179	0.326742	0.176479	1	0.21138
K_r	-0.93949	0.087619	0.397949	-0.97877	0.988839	0.994814	0.21138	1

Tabela 31: Matriz de correlações da P_{TPP} do modelo FOWM estimado com três pressões.

	$m_{l,still}$	C_g	C_{out}	V_{eb}	ϵ	K_w	K_a	K_r
$m_{l,still}$	1	0.371235	-0.12657	0.602054	-0.36309	-0.54725	0.37121	-0.25651
C_g	0.371235	1	-0.95457	0.892786	-0.98247	-0.13372	0.99999	-0.98316
C_{out}	-0.12657	-0.95457	1	-0.76921	0.963847	0.124756	-0.95458	0.956612
V_{eb}	0.602054	0.892786	-0.76921	1	-0.88489	-0.12256	0.892902	-0.88597
ϵ	-0.36309	-0.98247	0.963847	-0.88489	1	0.224104	-0.98249	0.966246
K_w	-0.54725	-0.13372	0.124756	-0.12256	0.224104	1	-0.1336	-0.02034
K_a	0.37121	0.99999	-0.95458	0.892902	-0.98249	-0.1336	1	-0.98321
K_r	-0.25651	-0.98316	0.956612	-0.88597	0.966246	-0.02034	-0.98321	1

Tabela 32: Matriz de correlações da $P_{r,t}$ do modelo FOWM estimado com três pressões.

	$m_{l,still}$	C_g	C_{out}	V_{eb}	ϵ	K_w	K_a	K_r
$m_{l,still}$	1	0.992312	-0.62444	-0.99946	-0.99161	0.821848	0.967209	0.67277
C_g	0.992312	1	-0.53149	-0.98777	-0.99419	0.880795	0.988988	0.599102
C_{out}	-0.62444	-0.53149	1	0.64849	0.559376	-0.08241	-0.4474	-0.83721
V_{eb}	-0.99946	-0.98777	0.64849	1	0.988138	-0.80397	-0.95946	-0.68927
ϵ	-0.99161	-0.99419	0.559376	0.988138	1	-0.85487	-0.97526	-0.63666
K_w	0.821848	0.880795	-0.08241	-0.80397	-0.85487	1	0.923214	0.186155
K_a	0.967209	0.988988	-0.4474	-0.95946	-0.97526	0.923214	1	0.483164
K_r	0.67277	0.599102	-0.83721	-0.68927	-0.63666	0.186155	0.483164	1

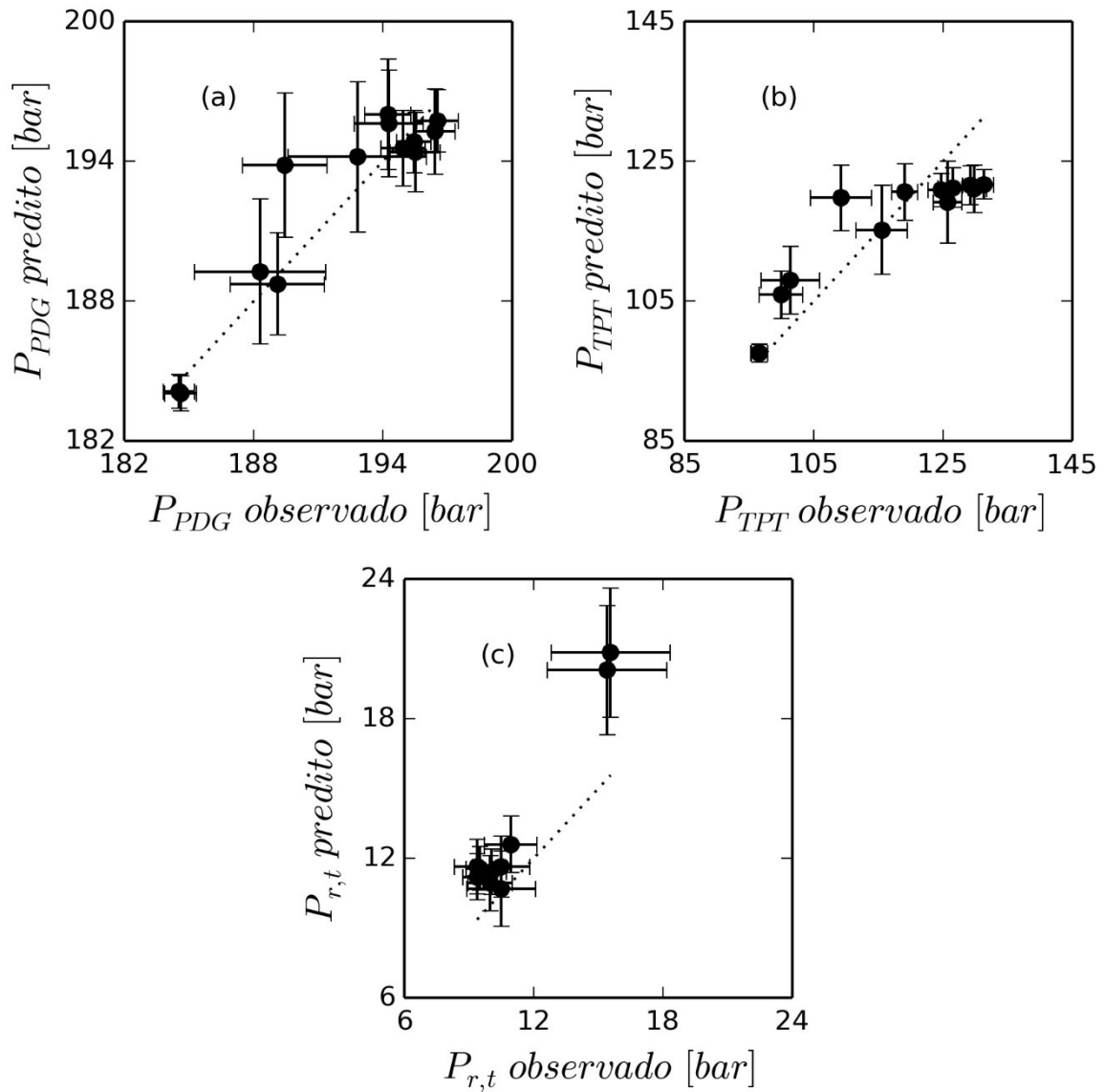


Figura 5.33: Valores preditos e observados da (a) P_{PDG} , (b) P_{TPT} e (c) $P_{r,t}$ do modelo FOWM estimado com três pressões.

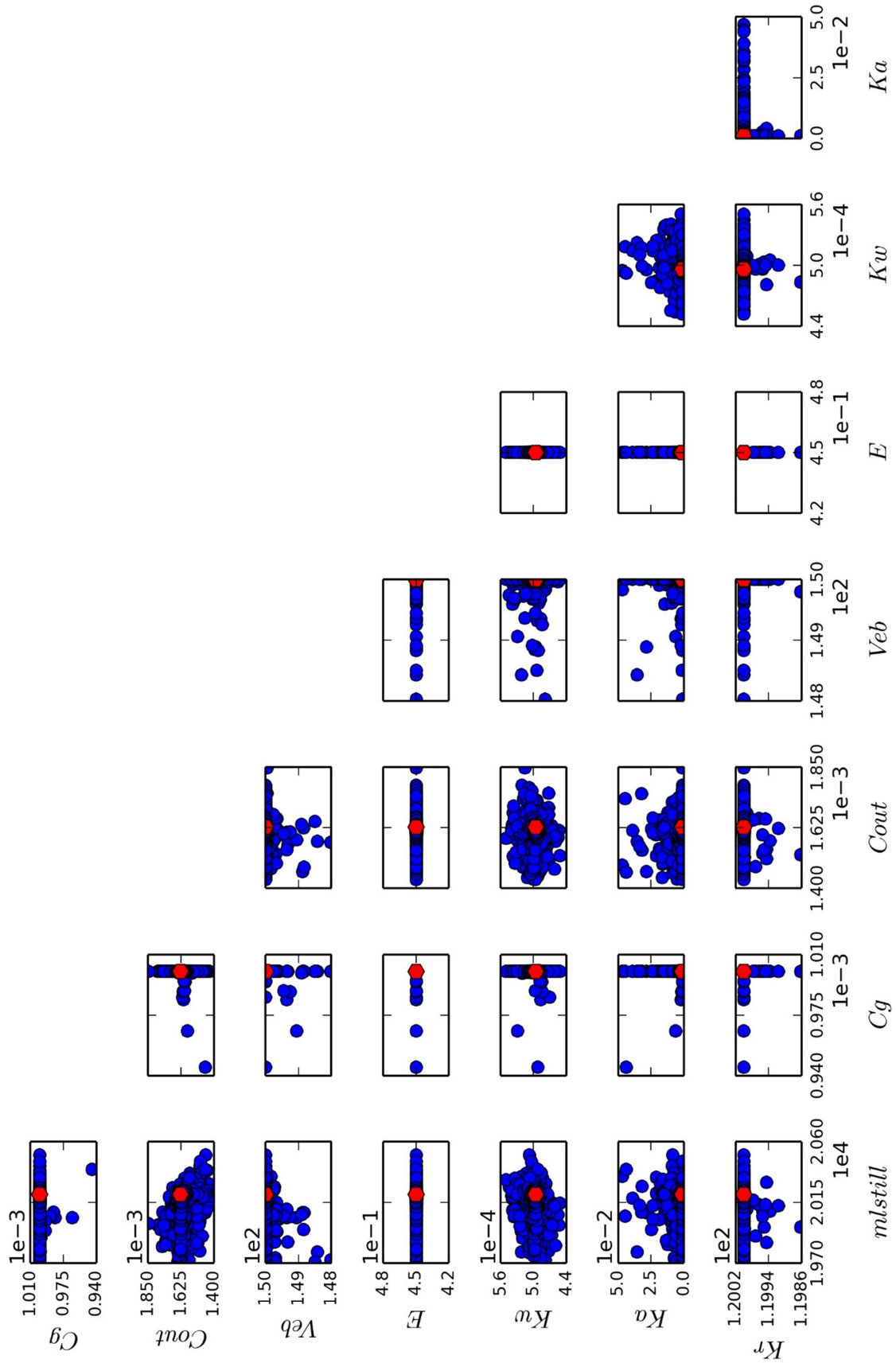


Figura 5.34: Regiões de máxima verossimilhança para o modelo FOWM estimado com três pressões.

Capítulo 6 – **Conclusões e Trabalhos futuros**

Neste capítulo são apresentadas as conclusões a respeito dos resultados. Também são feitas sugestões de trabalhos futuros tendo em vista outras aplicações do método proposto.

6.1 Conclusões quanto à metodologia

A nova metodologia desenvolvida neste trabalho propõe que se obtenha um ciclo representativo para todo o conjunto de dados em ciclo limite que incorpore a variabilidade dos dados. Isso facilita as avaliações feitas a respeito dos dados, considerando que um grande número de amostras pode ser condensado em um conjunto menor sem perdas de informação. A interpretação estatística também se torna mais direta porque se aproveita da natureza periódica dos dados como se fossem realizados vários experimentos seguidos.

Como foi mostrado no Capítulo 5 –, a nova abordagem foi capaz de apresentar resultados mais significativos para os dois modelos estudados. As variâncias experimentais foram essenciais na estimação de parâmetros quando a referência apresenta alta variabilidade. Isso foi comprovado estatisticamente através dos testes F e Chi Quadrado.

6.2 Estratégia de otimização do modelo FOWM

Considerando a alta correlação entre parâmetros apresentada nos resultados para o modelo FOWM, sugere-se a avaliação da necessidade da inclusão de todos os parâmetros na rotina de otimização. A determinação da sensibilidade dos parâmetros pode auxiliar na decisão de se obter um valor fixo ou um valor estimado. Isso teria uma influência direta na correlação entre os parâmetros, que precisaria ser novamente estudada. As estimações mostram que o parâmetro ω_u muda dramaticamente as regiões de confiança dos parâmetros. Contudo, não existe melhora significativa quanto ao ajuste do modelo. Assim, esse parâmetro poderia ser fixado em $\omega_u = 1$ até que não seja possível obter um ajuste adequado com a estimação dos demais parâmetros.

Além disso, seria interessante a estimação simultânea com diferentes aberturas de válvulas. Isso melhoraria a capacidade preditiva do modelo e a representatividade de múltiplos conjuntos de dados. Analogamente como foi feito ao modelo Meglio, outras pressões, como a P_{TPT} e a $P_{r,t}$, também poderiam ser inclusas na otimização.

6.3 Sugestões de trabalhos futuros

A comparação entre diferentes conjuntos de parâmetros é feita através do teste F. Contudo, o teste assume que os erros seguem a distribuição normal. Determinar se essa suposição é válida ou não através de um teste Kolmogorov–Smirnov (KS) permitiria decidir se há necessidade de outro teste para comparação entre conjuntos de parâmetros.

A obtenção de um ciclo médio para um espaço amostral gera uma base de comparação para dados similares. Da mesma forma que o modelo é comparado com a referência na função objetivo, referências diferentes podem ser avaliadas a fim de determinar a presença ou não de diferenças significativas.

No segmento de elevação de petróleo, pode-se aplicar o método de modo a auditar dados de pressão e produção em diferentes condições operacionais. A criação de uma base de dados para cada condição auxiliaria na identificação de distúrbios do sistema.

No segmento de automação, válvulas de controle podem apresentar agarramento. Nesses casos, surge uma condição de ciclo limite devido à resistência ao movimento da haste da válvula (DAMBROS, 2016). Como as oscilações prejudicam a performance da malha de controle, vários autores estudam métodos de detecção de agarramento. Um ciclo representativo das oscilações poderia auxiliar essa linha de pesquisa.

Referências

AAMO, O.M. *et al.* Observer design for multiphase flow in vertical pipes with gas-lift—theory and experiments. *Journal of Process Control*, v. 15, n. 3, p. 247–257, Apr. 2005. Disponível em:

<<http://linkinghub.elsevier.com/retrieve/pii/S0959152404000848>>. Acesso em: 21 fev. 2017.

ANP. *Especial ANP 20 anos*. Disponível em:

<<http://www.anp.gov.br/wwwanp/institucional/especial-anp-20-anos>>.

APIO, Andressa. Estimaco de parâmetros em modelos com ciclo limite. 2017.

Disponível em: <<http://www.lume.ufrgs.br/handle/10183/163746>>. Acesso em: 18 fev. 2018.

APIO, Andressa; BOTELHO, Viviane R.; TRIERWEILER, Jorge O. Parameter estimation of models with limit cycle based on the reformulation of the objective function. *Computers & Chemical Engineering*, v. 109, p. 236–248, 4 Jan. 2018.

Disponível em: <<http://linkinghub.elsevier.com/retrieve/pii/S0098135417304106>>. Acesso em: 9 jan. 2018.

ARAGO, Amanda Pereira. Estimativa Da Contribuio Do Setor Petrleo Ao Produto Interno Bruto Brasileiro: 1955/2004. p. 165, 2004.

BAILEY, R *et al.* *From Market Basket Analysis to Wellhead Monitoring: Using Events to Increase Oil Recovery*. *SPE Intelligent Energy Conference & Exhibition*. [S.l.]: Society of Petroleum Engineers. Disponível em:

<<http://www.onepetro.org/doi/10.2118/167836-MS>>. , 2014

BAKER, D. Simultaneous flow of oil and gas. *Oil and Gas J.*, v. 53, p. 183–195, 1954.

BENDIKSEN, Kjell H *et al.* The Dynamic Two-Fluid Model OLGA: Theory and Application. *Society of Petroleum Engineers*, 1991.

BERGMAN, Sixten. Peakdetect. 2016. Disponível em:

<<https://gist.github.com/sixtenbe/1178136>>.

BILLAUER, Eli. Peakdet. 2012. Disponível em: <<http://billauer.co.il/peakdet.html>>.

BRENNEN, Christopher E. *Fundamentals of Multiphase Flows*. Pasadena, California: Cambridge University Press, 2005.

CEDERSUND, G; KNUDSEN, C. Improved parameter estimation for systems with an experimentally located Hopf bifurcation. *IEE Proceedings - Systems Biology*, v. 152, n. 3, p. 161–168, Sep. 2005. Disponível em: <<http://link.aip.org/link/?BDJ/152/161/1>>.

CLERC, M; KENNEDY, J. The particle swarm - explosion, stability, and convergence in a multidimensional complex space. *IEEE Transactions on Evolutionary Computation*, v. 6, n. 1, p. 58–73, 2002. Disponível em: <<http://ieeexplore.ieee.org/document/985692/>>.

DALSMO, M; HALVORSEN, E; SLUPPHAUG, O. *Active Feedback Control of Unstable Wells at the Brage Field*. . [S.l.]: Society of Petroleum Engineers. , 2002

DAMBROS, J. W. V. *Detecção do agarramento em válvulas de controle para sinais com referência variável*. 2016. Universidade Federal do Rio Grande do Sul, 2016. Disponível em: <<http://hdl.handle.net/10183/149795>>.

DI MEGLIO, Florent *et al*. Reproducing slugging oscillations of a real oil well. Dec. 2010, [S.l.]: IEEE, Dec. 2010. p. 4473–4479. Disponível em: <<http://ieeexplore.ieee.org/document/5717367/>>.

DI MEGLIO, Florent *et al*. Stabilization of slugging in oil production facilities with or without upstream pressure sensors. *Journal of Process Control*, v. 22, n. 4, p. 809–822, Apr. 2012. Disponível em: <<http://www.scopus.com/inward/record.url?eid=2-s2.0-84859550558&partnerID=40&md5=c9c552e1838d482207b2ce03faf7e388>>.

DI MEGLIO, Florent; KAASA, Glenn-Ole Ole; PETIT, Nicolas. A first principle model for multiphase slugging flow in vertical risers. Dec. 2009, [S.l.]: IEEE, Dec. 2009. p. 8244–8251. Disponível em: <<http://ieeexplore.ieee.org/document/5400680/>>.

DIEHL, Fabio C. *et al*. Fast Offshore Wells Model (FOWM): A practical dynamic model for multiphase oil production systems in deepwater and ultra-deepwater scenarios. *Computers & Chemical Engineering*, v. 99, p. 304–313, Apr. 2017. Disponível em: <<http://linkinghub.elsevier.com/retrieve/pii/S0098135417300443>>. Acesso em: 21 fev. 2017.

EIKREM, Gisle Otto; AAMO, Ole Morten; FOSS, Bjarne. On Instability in Gas Lift Wells and Schemes for Stabilization by Automatic Control. *SPE Production & Operations*, v. 23, n. 2, p. 268–279, 1 May 2008. Disponível em: <<http://www.onepetro.org/mslib/servlet/onepetropreview?id=SPE-101502-PA%5Cnhttps://www.onepetro.org/journal-paper/SPE-101502-PA>>.

ELMAN, Howard C *et al.* Lyapunov Inverse Iteration for Identifying Hopf Bifurcations in Models of Incompressible Flow. *SIAM Journal on Scientific Computing*, doi: 10.1137/110827600, v. 34, n. 3, p. A1584–A1606, 1 Jan. 2012. Disponível em: <<http://epubs.siam.org/doi/10.1137/110827600>>.

FABRE, Jean *et al.* *Severe Slugging in Pipeline/Riser Systems*. [S.l.: s.n.], 1990. v. 5. Disponível em: <<http://www.onepetro.org/doi/10.2118/16846-PA>>.

FINKEL, Daniel E. *Global optimization with the direct algorithm*. 2005. 2005.

FOSS, Bjarne. Process control in conventional oil and gas fields—Challenges and opportunities. *Control Engineering Practice*, v. 20, n. 10, p. 1058–1064, 2012. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0967066112000202>>.

GEREVINI, Giovanni. Atenuação de golfadas em sistemas de elevação de petróleo em ambiente offshore. p. 100, 2017.

GEREVINI, Giovanni Gonçalves. Atenuação de golfadas em sistemas de elevação de petróleo em ambiente offshore. 2017. Disponível em: <<http://www.lume.ufrgs.br/handle/10183/172287>>. Acesso em: 18 feb. 2018.

GERHARD, D. *Pitch extraction and fundamental frequency: history and current techniques*. . Department of Computer Science, University of Regina, Regina, Canada: [s.n.]. , 2003

GODHAVN, John Morten; FARD, Mehrdad P.; FUCHS, Per H. New slug control strategies, tuning rules and experimental results. *Journal of Process Control*, v. 15, n. 5, p. 547–557, 1 Aug. 2005. Disponível em: <<http://www.sciencedirect.com/science/article/pii/S0959152404001131>>. Acesso em: 25 sep. 2017.

GUCKENHEIMER, John; MYERS, Mark; STURMFELS, Bernd. Computing Hopf Bifurcations I. *SIAM Journal on Numerical Analysis*, doi: 10.1137/S0036142993253461, v. 34, n. 1, p. 1–21, 1 Feb. 1997. Disponível em: <<http://epubs.siam.org/doi/10.1137/S0036142993253461>>.

HAVRE, Kjetil; DALSMO, Morten. Active Feedback Control as a Solution to Severe Slugging. *SPE Production & Facilities*, v. 17, n. 3, p. 138–148, 2002. Disponível em: <<http://www.onepetro.org/doi/10.2118/79252-PA>>.

HÜFFNER, Leonardo Nardi. *Formulação de Função Objetivo para Estimação de Parâmetros de Modelos com Ciclo Limite*. 2018.

HYNNE, F.; DANØ, S.; SØRENSEN, P. G. Full-scale model of glycolysis in *Saccharomyces cerevisiae*. *Biophysical Chemistry*, v. 94, n. 1–2, p. 121–163, Dec. 2001.

HYNNE, F.; GRAAE SØRENSEN, P.; MØLLER, T. Complete optimization of models of the Belousov-Zhabotinsky reaction at a Hopf bifurcation. *The Journal of Chemical Physics*, v. 98, n. 1, p. 219–230, Jan. 1993. Disponível em: <<http://aip.scitation.org/doi/10.1063/1.464667>>.

JAHANSHAHI, E. *et al.* A comparison between Internal Model Control, optimal PIDF and robust controllers for unstable flow in risers. *IFAC Proceedings Volumes (IFAC-PapersOnline)*, v. 19, n. March 2016, p. 5752–5759, 2014.

JAHANSHAHI, Esmaeil. *Control Solutions for Multiphase Flow*. 2013. Norwegian University of Science and Technology, 2013.

JAHANSHAHI, Esmaeil; SKOGESTAD, Sigurd. Simplified Dynamical Models for Control of Severe Slugging in Multiphase Risers. *IFAC Proceedings Volumes*, v. 44, n. 1, p. 1634–1639, Jan. 2011. Disponível em: <<http://linkinghub.elsevier.com/retrieve/pii/S1474667016438449>>. Acesso em: 21 fev. 2017.

JANSEN, Bård *et al.* Automatic Control of Unstable Gas Lifted Wells. 4 Apr. 1999, [S.l.]: Society of Petroleum Engineers, 4 Apr. 1999. p. 1–9. Disponível em: <<http://www.onepetro.org/doi/10.2118/56832-MS>>.

JANSEN, F.E.; SHOHAM, O.; TAITEL, Y. The elimination of severe slugging—experiments and modeling. *International Journal of Multiphase Flow*, v. 22, n. 6, p. 1055–1072, Nov. 1996. Disponível em: <<http://linkinghub.elsevier.com/retrieve/pii/0301932296000274>>. Acesso em: 21 fev. 2017.

LEE, Abraham. *Pyswarm v0.6*. . [S.l.: s.n.]. Disponível em: <<https://pypi.python.org/pypi/pyswarm/#downloads>>. , 2014

MEERBERGEN, Karl; SPENCE, Alastair. Inverse Iteration for Purely Imaginary Eigenvalues with Application to the Detection of Hopf Bifurcations in Large-Scale Problems. *SIAM Journal on Matrix Analysis and Applications*, doi: 10.1137/080742890, v. 31, n. 4, p. 1982–1999, 1 Jan. 2010. Disponível em: <<http://epubs.siam.org/doi/abs/10.1137/080742890>>.

NABAVI, Seyedbehzad; WILLIAMS, Cranos M. A novel cost function to estimate parameters of oscillatory biochemical systems. *EURASIP Journal on Bioinformatics and Systems Biology*, v. 2012, n. 1, p. 3, 2012. Disponível em: <<https://bsb-eurasipjournals.springeropen.com/articles/10.1186/1687-4153-2012-3>>.

OLIVEIRA, J.A.; GAMBETTA, R.; PINTO, J.C. *Pega Ponto v1.0*. . [S.l.: s.n.]. Disponível em: <<http://pega-ponto.software.informer.com/1.0/>>. , 2006

PEDERSEN, Simon; DURDEVIC, Petar; YANG, Zhenyu. Challenges in slug modeling and control for offshore oil and gas productions: A review study. *International Journal of Multiphase Flow*, v. 88, p. 270–284, 2017.

PETROBRAS. *Confira nossa produção de petróleo e gás natural em julho*. Disponível em: <<http://www.petrobras.com.br/fatos-e-dados/confira-nossa-producao-de-petroleo-e-gas-natural-em-julho.htm>>.

POLI, Riccardo; KENNEDY, James; BLACKWELL, Tim. Particle swarm optimization. *Swarm Intelligence*, v. 1, n. 1, p. 33–57, 17 Oct. 2007. Disponível

em: <<http://dx.doi.org/10.1007/s11721-007-0002-0>>.

RIZZO FILHO, Haroldo dos Santos. *A Otimização de Gás Lift na Produção de Petróleo: Avaliação da Curva de Performance do Poço*. 2011. 92 f. UFRJ, 2011.

SARICA, Cem; TENGESDAL, Jarl Å. *A New Technique to Eliminate Severe Slugging in Pipeline/Riser Systems*. *SPE Annual Technical Conference and Exhibition*. [S.l.]: Society of Petroleum Engineers. Disponível em: <<http://www.onepetro.org/mslib/app/Preview.do?paperNumber=00063185&societyCode=SPE>>. , 2000

SCHMIDT, Z; BRILL, J.P.; BEGGS, H.D. Experimental Study of Severe Slugging in a Two-Phase-Flow Pipeline - Riser Pipe System. *Society of Petroleum Engineers Journal*, v. 20, n. 5, p. 407–414, 1 Oct. 1980. Disponível em: <<http://www.onepetro.org/doi/10.2118/8306-PA>>.

SCHMIDT, Z; BRILL, J P; BEGGS, H D. Choking can eliminate pipeline slugging. *Oil Gas J.*, v. 12, p. 230–238, 1979.

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

SEYDEL, Rüdiger. *Practical Bifurcation and Stability Analysis*. New York, NY: Springer New York, 2010. v. 5. Disponível em: <<http://link.springer.com/10.1007/978-1-4419-1740-9>>. (Interdisciplinary Applied Mathematics).

SINEGRE, L.; PETIT, Nicolas; MENEGATTI, P. Predicting instabilities in gas-lifted wells simulation. *2006 American Control Conference*, p. 8 pp., 2006. Disponível em: <<http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=1657604>>.

SKOGESTAD, Sigurd; JAHANSHAHI, Esmail; HANSEN, Henrik. Control structure design for stabilizing unstable gas-lift oil wells. *IFAC Proceedings Volumes (IFAC-PapersOnline)*, v. 8, n. PART 1, p. 93–100, 2012. Disponível em: <<http://linkinghub.elsevier.com/retrieve/pii/S1474667016304256>>. Acesso em: 21 feb. 2017.

STORKAAS, E. *Stabilizing control and controllability*. 2005. 156 f. Norwegian University of Science and Technology, 2005. Disponível em: <http://www.nt.ntnu.no/users/skoge/diplom/prosjekt08/tuvnes/Articles/STORKAAS_PHD_THESIS_STABILIZING_CONTROL_AND_Controllerability_Control_Solutions_to_Slug_FLOW.pdf>.

STORKAAS, Espen. *Stabilizing control and controllability: control solutions to avoid slug flow in pipeline-riser systems*. 2005. 2005.

STORNES, Karl Ole; HAVRE, Kjetil; STRAY, Henrik. Taming slug flow in pipelines. *ABB Review*, v. 46, n. 4, p. 55–63, 2000.

THOMAS, José Eduardo. *Fundamentos de Engenharia de Petróleo*. Rio de Janeiro: Interciência, 2001.

VOGEL, J V. Inflow Performance Relationships for Solution-Gas Drive Wells. 1968.

WALLIS, Graham B. *One-dimensional two-phase flow*. New York: McGraw-Hill, 1969.

WEISMAN, J. Two-phase flow patterns. Cited By :1Export Date: 3 November 2017, 1983. Disponível em: <<https://www.scopus.com/inward/record.uri?eid=2-s2.0-0020556019&partnerID=40&md5=08df04132425fe33ae561c40f8326541>>.

WEISMAN, J.; KANG, S.Y. Flow pattern transitions in vertical and upwardly inclined lines. *International Journal of Multiphase Flow*, v. 7, n. 3, p. 271–291, Jun. 1981. Disponível em: <<http://linkinghub.elsevier.com/retrieve/pii/0301932281900227>>. Acesso em: 21 fev. 2017.

Apêndice A – Códigos implementados em Python

Esta seção é dedicada a apresentar os principais códigos utilizados neste trabalho. Para os arquivos .py, o título em itálico representa a função descrita. O fluxograma da Figura A.1 ilustra as tomadas de decisão para uma iteração.

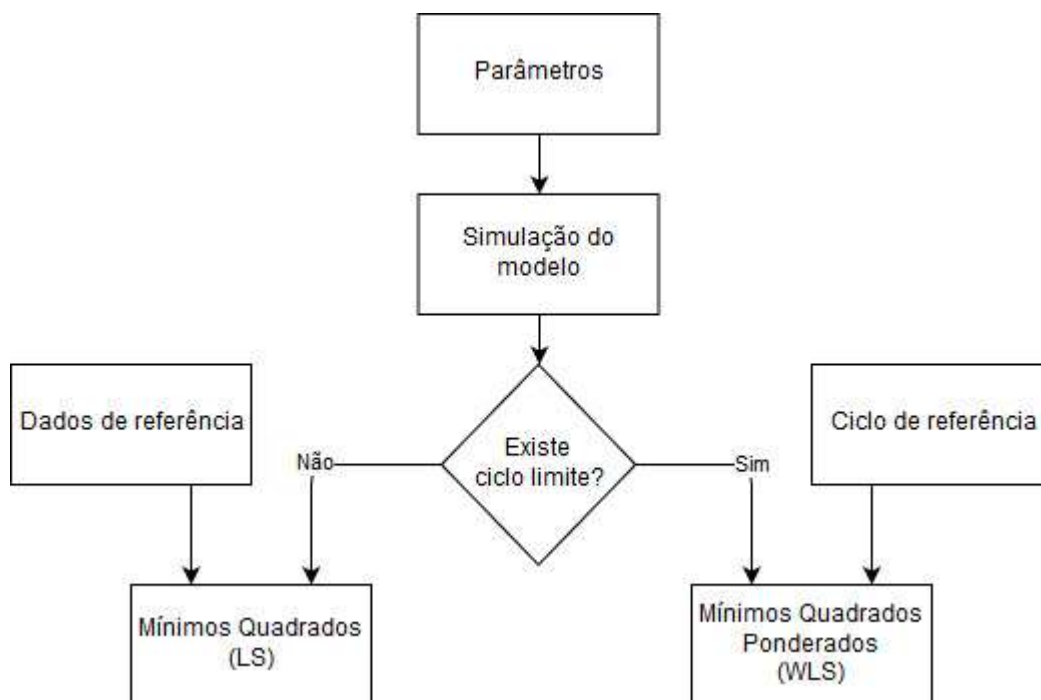


Figura A.1: Fluxograma para obtenção do valor da função objetivo para uma iteração.

1.1 meglio.mo

Modelo Meglio implementado em JModelica.

```

1. model meglio
2.
3.     //PARAMETERS
4.     parameter Real rol=900;
5.     parameter Real Ps=6.6e5;
6.     parameter Real R=8.314;
7.     parameter Real T=363.0;
8.     parameter Real M=2.2e-2;
9.     parameter Real g=9.81;
10.    parameter Real teta=0.7854;
11.    parameter Real A=1.77e-2;
12.    parameter Real mlstill=3.73e4;
13.    parameter Real D=1.5e-1;
14.    parameter Real Wgin=8.2e-1;
15.    parameter Real Wlin=11.75;
16.    parameter Real L=5200;
17.
18.    parameter Real Cg=1e-4;
19.    parameter Real E=0.78;
20.    parameter Real Veb=48.0;
21.    parameter Real Cout=2.8e-3;
22.
23.    parameter Real x1_0=5.4132e3;
24.    parameter Real x2_0=0.2704e3;
25.    parameter Real x3_0=4.9669e3;
26.
27.    //parameter Real u=0.5;
28.    parameter Real pi=3.14;
29.    parameter Real epsi = 1e-10;
30.
31.    //Initial Conditions
32.    Real x1(fixed=true,start=x1_0);
33.    Real x2(fixed=true,start=x2_0);
34.    Real x3(fixed=true,start=x3_0);
35.
36.    //Outputs
37.    Real Vr;
38.    Real Peb;
39.    Real Prt;
40.    Real Prb;
41.    Real Wout;
42.    Real Wlout;
43.    Real Wgout;
44.    Real Wg;
45.
46.    input Real u;
47.
48. equation
49.
50.    Vr = L*pi*D*D/4;
51.    Peb = x1*R*T/(M*Veb);
52.    Prt = x2*R*T/(M*(Vr-(x3+mlstill)/rol));
53.
54.    Prb = Prt + (x3+ mlstill)*g*sin(teta)/A;
55.    Wout = Cout*u*sqrt(rol)*sqrt(((Prt-Ps)+sqrt((Prt-Ps)^2+epsi))/2);
56.    Wgout = x2*Wout/x3;
57.    Wlout = Wout;
58.    Wg = Cg*((Peb - Prb)+sqrt((Peb - Prb)^2+epsi))*(0.5);
59.
60.    der(x1) = (1-E)*Wgin-Wg;
61.    der(x2) = E*Wgin+Wg-Wgout;

```

```

62.     der(x3) = Wlin - Wlout;
63.
64. end meglio;

```

1.2 FOWM.mo

Modelo FOWM implementado em JModelica.

```

1. model ABL56
2.
3.     //PARAMETERS
4.     parameter Real Rol    = 900;
5.     parameter Real R      = 8314;
6.     parameter Real T      = 298;
7.     parameter Real M      = 18;
8.     parameter Real g      = 9.81;
9.     parameter Real teta   = 3.14/4;
10.    parameter Real Ps     = 1013250;
11.    parameter Real Pr     = 2.25e7;
12.    parameter Real ALFagw = 0.0188;
13.    parameter Real Romres = 891.9523;
14.    parameter Real L      = 4497;
15.    parameter Real Lt     = 1639;
16.    parameter Real La     = 1118;
17.    parameter Real D      = 0.152;
18.    parameter Real Dt     = 0.150;
19.    parameter Real Da     = 0.140;
20.    parameter Real Hvgl   = 916;
21.    parameter Real Hpdg   = 1117;
22.    parameter Real Ht     = 1279;
23.    parameter Real epsi   = 1e-10;
24.
25.    //ADJUST PARAMETERS
26.    parameter Real mlstill = 6.222e1;
27.    parameter Real Cg      = 1.137e-3;
28.    parameter Real Cout    = 2.039e-3;
29.    parameter Real Veb     = 6.098e1;
30.    parameter Real E       = 1.545e-1;
31.    parameter Real Kw      = 6.876e-4;
32.    parameter Real Ka      = 2.293e-5;
33.    parameter Real Vr      = 226.74;
34.    parameter Real Kr      = 1.269e2;
35.
36.
37.    //PARAMETERS OF DECISION
38.    parameter Real z       = 0.05;
39.    parameter Real Wgc     = 0.693246557269;
40.
41.
42.    parameter Real A       = D*D*3.14/4;
43.    parameter Real Vt     = Lt*Dt*Dt*3.14/4;
44.    parameter Real Va     = La*Da*Da*3.14/4;
45.
46.    //Initial Conditions
47.    parameter Real x1_0    = 9347.5467727;
48.    parameter Real x2_0    = 5347.53586993;
49.    parameter Real x3_0    = 35014.1218045;
50.    parameter Real x4_0    = 2395.94295262;
51.    parameter Real x5_0    = 1788.01439135;
52.    parameter Real x6_0    = 11725.3198923;
53.
54.    //Outputs
55.    Real x1(fixed=true,start=x1_0);
56.    Real x2(fixed=true,start=x2_0);
57.    Real x3(fixed=true,start=x3_0);
58.    Real x4(fixed=true,start=x4_0);

```

```

59.      Real x5(fixed=true,start=x5_0);
60.      Real x6(fixed=true,start=x6_0);
61.
62.      Real Peb;
63.      Real Prt;
64.      Real Prb;
65.      Real ALFAg;
66.      Real ALFA1;
67.      Real Wout;
68.      Real Wlout;
69.      Real Wgout;
70.      Real Wg;
71.
72.
73.      Real Vgt;
74.      Real ROgt;
75.      Real ROmt;
76.      Real Ptt;
77.      Real Ptb;
78.      Real Ppdg;
79.      Real Pbh;
80.      Real ALFAgt;
81.      Real Wwh;
82.      Real Wwhg;
83.      Real Wwh1;
84.      Real Wr;
85.
86.      Real Pai;
87.      Real ROai;
88.      Real Wiv;
89.
90. equation
91.
92.      Peb = x1*R*T/(M*Veb);
93.      Prt = x2*R*T/(M*(Vr-(x3+mlstill)/Rol));
94.      Prb = Prt + (x3 + mlstill)*g*sin(teta)/A;
95.      ALFAg = x2/(x2 + x3);
96.      ALFA1 = 1 - ALFAg;
97.      Wout = Cout*z*sqrt(Rol*((Prt - Ps)+sqrt((Prt - Ps)^2+epsi))*(1/2));
98.      Wlout = ALFA1*Wout;
99.      Wgout = ALFAg*Wout;
100.     Wg = Cg*((Peb - Prb)+sqrt((Peb - Prb)^2+epsi))*(1/2);
101.
102.
103.     Vgt = Vt - x6/Rol;
104.     ROgt = x5/Vgt;
105.     ROmt = (x5 + x6)/Vt;
106.     Ptt = ROgt*R*T/M;
107.     Ptb = Ptt + ROmt*g*Hvgl;
108.     Ppdg = Ptb + Romres*g*(Hpdg-Hvgl);
109.     Pbh = Ppdg + Romres*g*(Ht-Hpdg);
110.     ALFAgt = x5/(x6 + x5);
111.     Wwh = Kw*sqrt(Rol*((Ptt - Prb)+sqrt((Ptt - Prb)^2+epsi))*(1/2));
112.     Wwhg = Wwh*ALFAgt;
113.     Wwh1 = Wwh*(1 - ALFAgt);
114.     Wr = Kr*(1-0.2*Pbh./Pr-0.8*(Pbh./Pr).^2);
115.
116.     Pai = ((R*T/(Va*M)) + (g*La/Va))*x4;
117.     ROai = M*Pai/(R*T);
118.     Wiv = Ka*sqrt(ROai*((Pai - Ptb)+sqrt((Pai - Ptb)^2+epsi))*(1/2));
119.
120.     der(x1) = (1 - E)*(Wwhg) - Wg;
121.     der(x2) = E*(Wwhg) + Wg - Wgout;
122.     der(x3) = Wwh1 - Wlout;
123.     der(x4) = Wgc - Wiv;

```

```

124.         der(x5) = Wr * ALFAgw + Wiv - Wwhg;
125.         der(x6) = Wr * ( 1 - ALFAgw ) - Wwhl;
126.
127.end ABL56;

```

1.3 peakdetect.py

peakdetect - Algoritmo de detecção de picos.

```

1. import numpy as np
2.
3. def _datacheck_peakdetect(x_axis, y_axis):
4.     if x_axis is None:
5.         x_axis = range(len(y_axis))
6.
7.     if len(y_axis) != len(x_axis):
8.         raise ValueError(
9.             "Input vectors y_axis and x_axis must have same length")
10.
11.     #needs to be a numpy array
12.     y_axis = np.array(y_axis)
13.     x_axis = np.array(x_axis)
14.     return x_axis, y_axis
15.
16. def peakdetect(y_axis, x_axis = None, lookahead = 200, delta=0):
17.     """
18.     Converted from/based on a MATLAB script at:
19.     http://billauer.co.il/peakdet.html
20.
21.     function for detecting local maxima and minima in a signal.
22.     Discovers peaks by searching for values which are surrounded by lower
23.     or larger values for maxima and minima respectively
24.
25.     keyword arguments:
26.     y_axis -- A list containing the signal over which to find peaks
27.
28.     x_axis -- A x-axis whose values correspond to the y_axis list and is used
29.               in the return to specify the position of the peaks. If omitted an
30.               index of the y_axis is used.
31.               (default: None)
32.
33.     lookahead -- distance to look ahead from a peak candidate to determine if
34.                  it is the actual peak
35.                  (default: 200)
36.                  '(samples / period) / f' where '4 >= f >= 1.25' might be a good value
37.
38.     delta -- this specifies a minimum difference between a peak and
39.              the following points, before a peak may be considered a peak. Useful
40.              to hinder the function from picking up false peaks towards to end of
41.              the signal. To work well delta should be set to delta >= RMSnoise * 5.
42.              (default: 0)
43.              When omitted delta function causes a 20% decrease in speed.
44.              When used Correctly it can double the speed of the function
45.
46.
47.     return: two lists [max_peaks, min_peaks] containing the positive and
48.            negative peaks respectively. Each cell of the lists contains a tuple
49.            of: (position, peak_value)
50.            to get the average peak value do: np.mean(max_peaks, 0)[1] on the
51.            results to unpack one of the lists into x, y coordinates do:
52.            x, y = zip(*max_peaks)
53.            """
54.     max_peaks = []
55.     min_peaks = []
56.     dump = [] #Used to pop the first hit which almost always is false
57.

```

```

58. # check input data
59. x_axis, y_axis = _datacheck_peakdetect(x_axis, y_axis)
60. # store data length for later use
61. length = len(y_axis)
62.
63.
64. #perform some checks
65. if lookahead < 1:
66.     raise ValueError("Lookahead must be '1' or above in value")
67. if not (np.isscalar(delta) and delta >= 0):
68.     raise ValueError("delta must be a positive number")
69.
70. #maxima and minima candidates are temporarily stored in
71. #mx and mn respectively
72. mn, mx = np.Inf, -np.Inf
73.
74. #Only detect peak if there is 'lookahead' amount of points after it
75. for index, (x, y) in enumerate(zip(x_axis[:-lookahead],
76.                                   y_axis[:-lookahead])):
77.     if y > mx:
78.         mx = y
79.         mxpos = x
80.     if y < mn:
81.         mn = y
82.         mnpos = x
83.
84.     #####look for max####
85.     if y < mx-delta and mx != np.Inf:
86.         #Maxima peak candidate found
87.         #look ahead in signal to ensure that this is a peak and not jitter
88.         if y_axis[index:index+lookahead].max() < mx:
89.             max_peaks.append([mxpos, mx])
90.             dump.append(True)
91.             #set algorithm to only find minima now
92.             mx = np.Inf
93.             mn = np.Inf
94.             if index+lookahead >= length:
95.                 #end is within lookahead no more peaks can be found
96.                 break
97.             continue
98.     #else: #slows shit down this does
99.     #     mx = ahead
100.    #     mxpos = x_axis[np.where(y_axis[index:index+lookahead]==mx)]
101.
102.    #####look for min####
103.    if y > mn+delta and mn != -np.Inf:
104.        #Minima peak candidate found
105.        #look ahead in signal to ensure that this is a peak and not jitter
106.        if y_axis[index:index+lookahead].min() > mn:
107.            min_peaks.append([mnpos, mn])
108.            dump.append(False)
109.            #set algorithm to only find maxima now
110.            mn = -np.Inf
111.            mx = -np.Inf
112.            if index+lookahead >= length:
113.                #end is within lookahead no more peaks can be found
114.                break
115.            #else: #slows shit down this does
116.            #     mn = ahead
117.            #     mnpos = x_axis[np.where(y_axis[index:index+lookahead]==mn)]
118.
119.
120. #Remove the false hit on the first value of the y_axis
121. try:
122.     if dump[0]:

```

```

123.         max_peaks.pop(0)
124.     else:
125.         min_peaks.pop(0)
126.     del dump
127. except IndexError:
128.     #no peaks were found, should the function return empty lists?
129.     pass
130.
131.     return [max_peaks, min_peaks]

```

1.4 freq.py

freq - Determinação de frequência baseada em autocorrelação.

```

1. from numpy.fft import rfft
2. from numpy import argmax, mean, diff, log, copy, arange
3. from scipy.signal import fftconvolve
4.
5.
6. def parabolic(f, x):
7.
8.     """Quadratic interpolation for estimating the true position of an
9.     inter-sample maximum when nearby samples are known.
10.
11.     f is a vector and x is an index for that vector.
12.
13.     Returns (vx, vy), the coordinates of the vertex of a parabola that goes
14.     through point x and its two neighbors.
15.
16.     Example:
17.     Defining a vector f with a local maximum at index 3 (= 6), find local
18.     maximum if points 2, 3, and 4 actually defined a parabola.
19.
20.     In [3]: f = [2, 3, 1, 6, 4, 2, 3, 1]
21.
22.     In [4]: parabolic(f, argmax(f))
23.     Out[4]: (3.2142857142857144, 6.1607142857142856)
24.
25.     """
26.     xv = 1/2. * (f[x-1] - f[x+1]) / (f[x-1] - 2 * f[x] + f[x+1]) + x
27.     yv = f[x] - 1/4. * (f[x-1] - f[x+1]) * (xv - x)
28.     return (xv, yv)
29.
30. def freq(signal, fs):
31.
32.     """Estimate frequency using autocorrelation
33.
34.     Pros: Best method for finding the true fundamental of any repeating wave,
35.     even with strong harmonics or completely missing fundamental
36.
37.     Cons: Not as accurate, doesn't work for inharmonic things like musical
38.     instruments, this implementation has trouble with finding the true peak
39.
40.     """
41.     # Calculate autocorrelation (same thing as convolution, but with one input
42.     # reversed in time), and throw away the negative lags
43.     signal = signal - mean(signal) # Remove DC offset
44.     corr = fftconvolve(signal, signal[::-1], mode='full')
45.     corr = corr[len(corr)/2:]
46.
47.     # Find the first low point
48.     d = diff(corr)
49.     start = next(i for i in range(len(d)) if d[i] > 0)
50.
51.     # Find the next peak after the low point (other than 0 lag). This bit is
52.     # not reliable for long signals, due to the desired peak occurring between

```

```

53.     # samples, and other peaks appearing higher.
54.     i_peak = argmax(corr[start:]) + start
55.     i_interp = parabolic(corr, i_peak)[0]
56.
57.     return fs / i_interp

```

1.5 cycle.py

Cycle - Obtenção do ciclo representativo.

```

1. import scipy.interpolate as inter
2. from numpy import *
3. from freq import freq
4. from peakdetect import peakdetect
5.
6. def Cycle(traw,yraw):
7.
8.     """
9.     Input 1 - time data
10.    Input 2 - signal data
11.
12.    Output 1 - cycle time data
13.    Output 2 - cycle signal data
14.    Output 3 - cycle variance data
15.
16.    """
17.
18.    if len(traw) != len(yraw): raise ValueError('Input arrays must have same length')
19.
20.    Ts = traw[1]-traw[0]
21.    T = 1./freq(yraw,1./Ts)
22.    L = len(traw)
23.    LA = T/Ts/1.5
24.    if LA < 1.: LA = 1
25.    peaks = peakdetect(yraw, lookahead=LA)
26.
27.    tshift = []
28.    yshift = []
29.    for i in range(len(peaks[1])-1):
30.        tshift.append(array(traw[peaks[1][i][0]:peaks[1][i+1][0]+1])-
31.            traw[peaks[1][i][0]])
32.        yshift.append(yraw[peaks[1][i][0]:peaks[1][i+1][0]+1])
33.        #plot(tshift[i],yshift[i])
34.
35.    NP = int(mean([ len(i) for i in tshift ]))
36.
37.    tspline = []
38.    yspline = []
39.    yshift = []
40.    spliner = linspace(0,1./freq(yraw,1./(float(traw[1]-traw[0]))),NP)
41.    for i in range(len(peaks[1])-1):
42.        yshift.append(yraw[peaks[1][i][0]:peaks[1][i+1][0]+1])
43.        ranger = linspace(0,1./freq(yraw,1./(float(traw[1]-
44.            traw[0]))),len(yshift[i]))
45.        tspline.append(ranger)
46.        yfspline = inter.InterpolatedUnivariateSpline(tspline[i],yshift[i])
47.        yspline.append(yfspline(spliner))
48.
49.    ymean = [ mean([ yspline[k][i] for k in range(len(yspline))]) for i in range(
50.        len(spliner)) ]

```



```

49.     yvar = [ var([ yspline[k][i] for k in range(len(yspline))]) for i in range(le
50.               n(spliner)) ]
51.     return spliner,ymean,yvar

```

1.6 stability.py

StabilityCheck - Determina a condição do sistema (estado estacionário, transiente ou ciclo limite);

```

1. import scipy
2. import numpy as np
3. from numpy import mean, array as A, log, exp
4. from peakdetect import peakdetect
5.
6.
7. def LS(ye,ym,g1): return sum( ( A(ye) - A(ym) ) **2 ) / g1
8.
9. def f(t,y):
10.
11.     DFT = np.fft.fft(y)/len(y)
12.     fs = 1./(t[1]-t[0])
13.     xDFT = np.fft.fftfreq(len(DFT),d=1./fs)
14.     M = np.argmax(abs(DFT))
15.     return abs(xDFT[M])
16.
17. def b(t,y,la):
18.
19.     pmax,pmin = peakdetect(y,lookahead=la)
20.     tpeak,ypeak = [],[]
21.     for i in range(len(pmax)):
22.         tpeak.append(t[pmax[i][0]])
23.         ypeak.append(pmax[i][1])
24.
25.     bmean = mean([log(ypeak[i]/ypeak[i+1])/(tpeak[i+1]-
26.               tpeak[i]) for i in range(len(tpeak)-1)])
27.     return bmean
28.
29. def C(t,y,la):
30.
31.     pmax,pmin = peakdetect(y,lookahead=la)
32.     tpeak,ypeak = [],[]
33.     for i in range(len(pmax)):
34.         tpeak.append(t[pmax[i][0]])
35.         ypeak.append(pmax[i][1])
36.
37.     Cmean = 0.5*mean([(ypeak[i]/exp(-b(t,y,la)*tpeak[i])+ypeak[i+1]/exp(-
38.               b(t,y,la)*tpeak[i+1])) for i in range(len(tpeak)-1)])
39.     return Cmean
40.
41. def Exp(b,c,t): return c*exp(-b*A(t))
42.
43. def StabilityCheck(t,y,debug=False):
44.
45.     """
46.     Input 1 - time data
47.     Input 2 - signal data
48.
49.     Output - True = Yes, the system is stabilizing / False = No, the system prese
50.               nts limit cycle
51.     """
52.
53.     ## 1 - Tem periodo e pelo menos 3 ciclos?

```

```

54.     try:
55.         T = 1./f(t,A(y)-mean(y))
56.         ts = t[1] - t[0]
57.         N = int(t[-1]/T)
58.         n = int(3.*T/ts) + 1
59.         if not N >= 6:
60.             if debug: print('N < 6, T = '+str(T)+' , N = '+str(N)+' , t[-
1] = '+str(t[-1]))
61.             return True # com menos de 6 ciclos, o metodo nao funciona (3 ciclos
de corte, 3 ciclos de ajuste)
62.         except:
63.             return True
64.             if debug:
65.                 print 'erro na avaliacao de periodos'
66.                 return True
67.
68.     ## 2 - A variancia dos tres ultimos periodos tem significancia? / Todos os pi
cos sao decrescentes?
69.     if np.var(y[-n:]) < 1e-5:
70.         if debug: print 'variancia insignificante'
71.         return True
72.     t,y = A(t[n:])-t[n],y[n:]
73.     la = int(len(t)/(t[-1]/T)/1.5)
74.     if la < 1: la = 1
75.     pmax,pmin = peakdetect(y,lookahead=la)
76.     thpeak,yhpeak,tlpeak,ylpeak = [],[],[],[]
77.     for i in range(len(pmax)):
78.         thpeak.append(t[pmax[i][0]])
79.         yhpeak.append(pmax[i][1])
80.     for i in range(len(pmin)):
81.         tlpeak.append(t[pmin[i][0]])
82.         ylpeak.append(pmin[i][1])
83.
84.     if np.var(yhpeak) < 1e-3:
85.         if debug: print 'picos nao variam significativamente'
86.         return False
87.     elif np.var(ylpeak) < 1e-3:
88.         if debug: print 'picos nao variam significativamente'
89.         return False
90.
91.     dec = all([yhpeak[i] > yhpeak[i+1] for i in range(len(yhpeak)-1)])
92.     if not dec:
93.         if debug: print 'picos nao sao decrescentes'
94.         return False
95.
96.     ## 3 - Um oscilador harmonico subamortecido ajusta os picos melhor do que a m
edia dos picos?
97.     med = mean(yhpeak+ylpeak)
98.     yh = Exp(b(t,A(y)-med,la),C(t,A(y)-med,la),thpeak)
99.     gl1 = len(thpeak)-2
100.    ls1 = LS(A(yhpeak)-med,yh,gl1)
101.    gl2 = len(tlpeak)-1
102.    ls2 = LS(yhpeak,mean(yhpeak),gl2)
103.    rls = ls1/ls2
104.    F1 = 1./scipy.stats.f.ppf(0.90,gl1,gl2)
105.    F2 = scipy.stats.f.ppf(0.90,gl2,gl1)
106.
107.    if F1 < rls and rls < F2:
108.        if debug: print 'passou no teste de Fisher'
109.        return False
110.    else:
111.        if debug: print 'estavel'
112.        return True

```

1.7 Settings.py

Settings – Compila o modelo.

```

1. def Settings(SimName,N,minh = 0.001,solver='Radau5ODE',directory=''):
2.
3.     """
4.     Inputs:
5.     SimName - Nome do modelo
6.     N - Numero de pontos
7.     minh - Valor do passo minimo (somente valido para BDF)
8.     solver - Nome do solver a ser utilizado. Opcoes recomendadas: Radau5ODE ou CV
           ode.
9.     directory - Nome do diretório do arquivo. Exemplos: FOWM/, Meglio/...
10.
11.    Outputs:
12.    slugging - Modelo compilado
13.    opts - Opções de simulação
14.    """
15.    # Compila e carrega a simulacao
16.    fmu_name = compile_fmud(SimName, directory+SimName+'.mo')
17.    slugging = load_fmud(fmu_name, log_level=0)
18.
19.    opts = slugging.simulate_options()
20.    opts['ncp'] = N - 1 #numero de pontos no intervalo
21.    opts['solver'] = solver
22.    opts[solver+'_options']['verbosity'] = 50 # omite dados de simulacao
23.    opts[solver+'_options']['atol'] = 1e-8
24.    opts[solver+'_options']['rtol'] = 1e-8
25.    opts['result_handling'] = 'memory'
26.    if solver == 'CVode':
27.        opts['CVode_options']['discr'] = 'BDF' # Backward differentiation formula
           - metodo de passo multiplo
28.        opts['CVode_options']['minh'] = minh
29.
30.    return slugging,opts

```

Exemplo de utilização:

```

1. sim,opts = Settings('Real Well A',700,solver='CVode',minh=1e-4,directory='Real
           Well A/')

```

1.8 ZeroPad.py

ZeroPad – Preenchimento do ciclo do modelo com a própria média.

```

1. import scipy.interpolate as inter
2. from numpy import array as A
3.
4. def ZeroPad(ym,tm,te):
5.
6.     ym,tm,te = A(ym),A(tm),A(te)
7.     if debug: print(tm,te)
8.     Tsm = tm[1]-tm[0]
9.     if tm[-1] < te[-1]:
10.        Sm = inter.InterpolatedUnivariateSpline(tm,ym)
11.        if tm[-1] + Tsm < te[-1]:
12.            i2 = next(list(te).index(i) for i in te if i > tm[-1])
13.            Ym = list(Sm(te[:i2])) + [mean(ym)]*(len(te)-i2)
14.        else:
15.            Ym = Sm(te)
16.
17.    elif te[-1] < tm[-1]:
18.        Sm = inter.InterpolatedUnivariateSpline(tm,ym)

```

```

19.         Ym = Sm(te)
20.     else:
21.         Ym = ym
22.
23.     return Ym

```

1.9 FO.py

Abaixo segue um exemplo de função objetivo *Fobj*. Supõe-se que:

- O modelo FOWM está compilado com o exemplo da função *Settings* (*sim* e *opts*);
- O número de pontos compilado com a função *Settings* é maior do que o número de pontos dos dados de referência;
- Os dados de referência têm a mesma amostragem do modelo (somente relevante se o modelo não está apresentando ciclo limite);
- A estimação é para o Cenário 1 dos resultados, onde 8 parâmetros são estimados enquanto o volume do riser $V_r = \omega_u V_{SS} = V_{SS}$ é fixo;
- Os valores de tempo da referência estão em horas;
- Os valores do sinal da referência estão em bar.

```

1. from stability import f, StabilityCheck
2. from ZeroPad import ZeroPad
3. from Cycle import Cycle
4. import numpy as *
5.
6. def Fobj(P,yraw,te,ye,vare,NP,sim=(),opts=(),IN=()):
7.
8.     """
9.     Inputs:
10.    P - Lista de parametros
11.    yraw - dados de referencia
12.    te - tempo do ciclo de referencia
13.    ye - sinal do ciclo de referencia
14.    vare - variancia do ciclo de referencia
15.    NP - numero de pontos a serem simulados.
16.    //CUIDADO// ==> !!! NP deve ser igual ao numero de pontos da compilacao do mo
delo !!!
17.    sim,opts - Resultado da compilacao obtida com a funcao Settings
18.    IN - informacao adicional a ser adicionada na simulacao
19.
20.    Output:
21.    Valor da funcao objetivo
22.    """
23.
24.    # Simulacao
25.    sim.reset()
26.    sim.set(['mlstill','Cg','Cout','Veb','E','Kw','Ka','Kr'],P)
27.    sim.set(['Wgc','Vr'],[1.406,4497*0.152**2*3.14/4])
28.    try:
29.        Sim = sim.simulate(start_time=0,final_time=(NP-
1)*60.,options=opts,input=IN)
30.        t,ymod = Sim['time']/3600.,array(Sim['Ppdg'])/1e5
31.    except: return 1e12

```

```
32.
33.     # Tenta descartar dados da convergencia inicial (3 ciclos)
34.     T = 1./f(t,ymod-mean(ymod))
35.     ts = t[1] - t[0]
36.     n = int(3*T/ts)
37.     ym = []
38.     try: t,ym = A(t[n:])-t[n],ymod[n:]
39.     except: t,ym = t,ymod
40.
41.     FO = 0.
42.
43.     # O sistema esta em estado estacionario, estado transiente?
44.     try: steadystate = Stable(t,ym)
45.     except: steadystate = True
46.
47.     # Se apresenta ciclo limite, aplica WLS;
48.     # Se nao apresenta ciclo limite, aplica LS.
49.     if not steadystate:
50.         try:
51.             ctm,cym,cvarm = Cycle(t,ym)
52.             Ym = ZeroPad(cym,ctm,te)
53.             FO += sum( ( array(Ym) - array(ye) ) ** 2 / array(vare))
54.         except:
55.             FO += sum( ( array(ymod[:len(yraw)]) - array(yraw) ) ** 2 )
56.         else: FO += sum( ( array(ymod[:len(yraw)]) - array(yraw) ) ** 2 )
57.
58.     return FO
```