



APLICAÇÃO DO FILTRO DE KALMAN POR CONJUNTO DE AMOSTRAS À
CALIBRAÇÃO E À VALIDAÇÃO DE MODELOS PREVISIONAIS DE
RESERVATÓRIOS DE PETRÓLEO

Edison Fabián Caballero Pérez

Tese de Doutorado apresentada ao Programa de Pós-graduação em Engenharia Mecânica, COPPE, da Universidade Federal do Rio de Janeiro, como parte dos requisitos necessários à obtenção do título de Doutor em Engenharia Mecânica.

Orientadores: Fernando Alves Rochinha
Márcio Rentes Borges

Rio de Janeiro
Abril de 2018

APLICAÇÃO DO FILTRO DE KALMAN POR CONJUNTO DE AMOSTRAS À
CALIBRAÇÃO E À VALIDAÇÃO DE MODELOS PREVISIONAIS DE
RESERVATÓRIOS DE PETRÓLEO

Edison Fabián Caballero Pérez

TESE SUBMETIDA AO CORPO DOCENTE DO INSTITUTO ALBERTO LUIZ
COIMBRA DE PÓS-GRADUAÇÃO E PESQUISA DE ENGENHARIA (COPPE)
DA UNIVERSIDADE FEDERAL DO RIO DE JANEIRO COMO PARTE DOS
REQUISITOS NECESSÁRIOS PARA A OBTENÇÃO DO GRAU DE DOUTOR
EM CIÊNCIAS EM ENGENHARIA MECÂNICA.

Examinada por:



Prof. Fernando Alves Rochinha, D.Sc.



Prof. Márcio Rentes Borges, D.Sc.



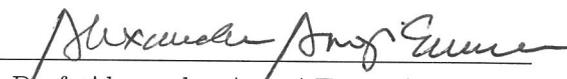
Prof. Marcio Arab Murad, D.Sc.



Prof. Fernando Pereira Duda, D.Sc.



Prof. Alvaro Luiz Coutinho, D.Sc.



Prof. Alexandre Anozé Emerick, Ph.D.

RIO DE JANEIRO, RJ – BRASIL

ABRIL DE 2018

Caballero Pérez, Edison Fabián

Aplicação do Filtro de Kalman por Conjunto de Amostras à Calibração e à validação de Modelos Previsionais de Reservatórios de Petróleo/Edison Fabián Caballero Pérez. – Rio de Janeiro: UFRJ/COPPE, 2018.

XVIII, 77 p.: il.; 29, 7cm.

Orientadores: Fernando Alves Rochinha

Márcio Rentes Borges

Tese (doutorado) – UFRJ/COPPE/Programa de Engenharia Mecânica, 2018.

Referências Bibliográficas: p. 69 – 77.

1. Assimilação de dados. 2. EnKF. 3. Modelagem de Reservatórios. I. Rochinha, Fernando Alves *et al.* II. Universidade Federal do Rio de Janeiro, COPPE, Programa de Engenharia Mecânica. III. Título.

a meus pais que tanto me apoiaram em este processo, a meus amigos que sempre estiveram para me dar uma força nos momentos certos e a todas as pessoas que de uma u outra forma fizeram parte do caminho percorrido no meu doutorado até hoje

Agradecimentos

Agradeço ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) e à Rede Temática de Simulação e Gerenciamento de Reservatórios da Petrobras (SIGER) pelo suporte financeiro.

Resumo da Tese apresentada à COPPE/UFRJ como parte dos requisitos necessários para a obtenção do grau de Doutor em Ciências (D.Sc.)

APLICAÇÃO DO FILTRO DE KALMAN POR CONJUNTO DE AMOSTRAS À
CALIBRAÇÃO E À VALIDAÇÃO DE MODELOS PREVISIONAIS DE
RESERVATÓRIOS DE PETRÓLEO

Edison Fabián Caballero Pérez

Abril/2018

Orientadores: Fernando Alves Rochinha
Márcio Rentes Borges

Programa: Engenharia Mecânica

Neste trabalho, foi construída uma nova metodologia para melhorar a precisão das previsões geradas com métodos sequenciais para o acoplamento da hidrodinâmica e da geomecânica de reservatórios de petróleo, preservando o baixo custo computacional. A nova abordagem computacional é desenvolvida dentro da estrutura do procedimento de algoritmo de divisão de tensão fixa em conjunto com a assimilação de dados baseada no filtro de Kalman por Conjunto de amostras (EnKF). Neste contexto, identificamos o modelo de alta fidelidade como a formulação de duas vias onde o termo de fonte adicional aparece na equação de fluxo contendo a derivada temporal da tensão média total. O esquema iterativo é então entrelaçado com as etapas de assimilação de dados que também incorporam o erro de modelagem inerente à estrutura do EnKF. Tal procedimento dá origem a uma formulação de uma via aprimorada exibindo melhora substancial na precisão em comparação com o método de uma via clássico. As equações governantes são discretizadas por elementos finitos mistos, e simulações numéricas de um problema slab 2D entre os poços de injeção e de produção ilustra o alcance do método aqui apresentado.

Abstract of Thesis presented to COPPE/UFRJ as a partial fulfillment of the requirements for the degree of Doctor of Science (D.Sc.)

ENKF APPLICATION TO CALIBRATION AND DATA ASSIMILATION ON
PETROLEUM RESERVOIRS FORECASTING MODELS

Edison Fabián Caballero Pérez

April/2018

Advisors: Fernando Alves Rochinha
Márcio Rentes Borges

Department: Mechanical Engineering

In this work we construct a new methodology for enhancing the predictive accuracy of sequential methods for coupling flow and geomechanics whilst preserving low computational cost. The new computational approach is developed within the framework of the fixed stress split algorithm procedure in conjunction with data assimilation based on the Ensemble Kalman Filter (EnKF). In this context we identify the high fidelity model with the two-way formulation where additional source term appears in the flow equation containing the time derivative of total mean stress. The iterative scheme is then interlaced with data assimilation steps which also incorporate the modeling error inherent to the EnKF framework. Such a procedure give rise to an "Enhanced one-way formulation" exhibiting substantial improvement in accuracy compared to the classical one-way method. The governing equations are discretized by mixed finite elements, and numerical simulation of a 2D slab problem between injection and production wells illustrate the tremendous achievement of the method proposed herein.

Sumário

Lista de Figuras	x
Lista de Tabelas	xiii
Lista de Símbolos	xiv
Lista de Abreviaturas	xviii
1 Introdução	1
1.1 Motivação e Contextualização	2
1.2 O <i>EnKF</i> na assimilação de Dados para Modelos de Reservatórios de Petróleo	5
1.3 O <i>EnKF</i> e o <i>ES</i>	7
1.4 O erro de modelo (ME) na assimilação de Dados	9
1.5 Organização do texto	10
2 Modelagem de Reservatórios de Petróleo	11
2.1 Propriedades dos Reservatórios de Petróleo	11
2.1.1 Porosidade	11
2.1.2 Permeabilidade	12
2.2 Lei de Darcy	13
2.3 Acoplamento Sequencial Entre o Escoamento e a Geomecânica	14
2.3.1 Formulação do Problema	15
3 Implementação Numérica e Algoritmo	19
3.1 Algoritmo Iterativo	19
3.2 Discretização do Subproblema de Escoamento	20
3.3 Discretização das Equações Poromecânicas	21
4 Assimilação de Dados Usando o <i>EnKF</i>	23
4.1 Assimilação Sequencial de Dados	24
4.2 O <i>Ensemble Kalman Filter (EnKF)</i>	25
4.3 Inversão Usando <i>TSVD</i>	27

4.4	Consistência das Variáveis de Estado	28
4.5	Aproximação do Erro de Modelagem	31
4.6	Proposta para Aprimorar o <i>EnKF</i>	33
4.7	Validação do Algoritmo do <i>EnKF</i> empregando o <i>MCMC</i>	36
5	Resultados e Discussões	46
5.1	Problema Protótipo	46
5.2	Cenário A	51
5.3	Cenário B	57
6	Conclusões	67
	Referências Bibliográficas	69

Lista de Figuras

4.1	Representação esquemática do <i>EnKF</i> iterativo [91].	30
4.2	Campo log-permeabilidade usado como referência no problema de estimação do tempo de <i>breakthrough</i> . O símbolos “o” indicam as posições nas quais a pressão é monitorada	38
4.3	Comparação entre a pressão de referência (linha contínua) e a pressão média (círculos) nos pontos de monitoração, obtidas com os métodos <i>EnKF</i> , <i>McMC</i> e <i>MC</i> . Barras indicam o desvio padrão. Linha segmentada vertical indica o tempo utilizado para a caracterização . . .	39
4.4	Comparação entre as previsões para a pressão de referência (linha contínua) obtidas com os métodos <i>EnKF</i> e <i>McMC</i> . Os círculos representam as pressões médias nos pontos de monitoração obtidas por cada método e as barras indicam o desvio padrão	40
4.5	Comparação entre as previsões para a pressão do campo de referência (linha contínua) e a pressão média (círculos) em três pontos do reservatório, obtidas com os métodos <i>EnKF</i> e <i>MC</i> . Barras indicam o desvio padrão. Linha segmentada vertical indica o tempo utilizado para a caracterização	41
4.6	Campo de referência e campos log-permeabilidade médios obtidos pelos métodos <i>EnKF</i> e <i>McMC</i>	42
4.7	Variância dos campos log-permeabilidade obtidos pelos métodos <i>EnKF</i> e <i>McMC</i>	42
4.8	Função de densidade de probabilidade do tempo de <i>breakthrough</i> previsto. A linha a segmentada vertical representa o tempo de <i>breakthrough</i> do modelo de referência.	43
4.9	Gráficos de caixa dos tempos de <i>breakthrough</i> estimados. A linha vermelha representa o tempo de <i>breakthrough</i> do modelo de referência	44
4.10	Função de densidade de probabilidade do tempo de <i>breakthrough</i> previsto. A linha a segmentada vertical representa o tempo de <i>breakthrough</i> do modelo de referência	44

4.11	Função de densidade de probabilidade do tempo de <i>breakthrough</i> previsto. A linha a segmentada vertical representa o tempo de <i>breakthrough</i> do modelo de referência	45
5.1	Domínio computacional: Decomposição em diferentes regiões geológicas; posição dos poços e estações de monitoração da subsidência	47
5.2	Campos espacialmente heterogêneos do modelo de referência: (a) permeabilidade e (b) módulo de Young	48
5.3	Termo fonte da equação de balanço	51
5.4	Comparação entre o desvio dos campos de permeabilidade assimilados e o campo de referência em $t = 72$ dias usando os modelos a)duas vias, b) uma via e c) uma via aprimorado	52
5.5	Comparação entre o desvio dos campos de módulo de Young assimilados e o campo de referência em $t = 72$ dias usando os modelos a)duas vias, b) uma via e c) uma via aprimorado	53
5.6	Análise do desempenho dos métodos de assimilação de dados	54
5.7	Previsão de produção empregando os modelos a) a priori, b) de duas vias, c) de uma via e d) de uma via aprimorado	55
5.8	Previsão de produção média acumulada	56
5.9	Média e de desvio da previsão de pressão de injeção empregando os modelos duas vias, umas via e uma aprimorada	56
5.10	Previsão de subsidência de superfície média estimada pelos modelos duas vias (azul), uma via (cinza) e uma via aprimorada (vermelho) .	57
5.11	Média e desvio padrão da previsão de subsidência da superfície nos sensores de deslocamento obtida com os modelos: de referência (preto), duas vias (verde), uma via (vermelho), uma via aprimorado (azul) e a priori (magenta). Barras de erro indicam desvio padrão de \pm nas posições monitoradas	58
5.12	Termo fonte da equação de balanço	59
5.13	Comparação entre o desvio dos campos de permeabilidade assimilados e o campo de referência em $t = 72$ dias usando os modelos a)duas vias, b) uma via e c) uma via aprimorado	60
5.14	Comparação entre o desvio dos campos de módulo de Young assimilados e o campo de referência em $t = 72$ dias usando os modelos a)duas vias, b) uma via e c) uma via aprimorado	62
5.15	Análise do desempenho dos métodos de assimilação de dados	63
5.16	Previsão de produção empregando os modelos a) a priori, b) de duas vias, c) de uma via e d) de uma via aprimorado	64
5.17	Previsão de produção média acumulada	64

5.18	Média e de desvio da previsão de pressão de injeção empregando os modelos duas vias, umas via e uma aprimorada	65
5.19	Previsão de subsidência de superfície média estimada pelos modelos duas vias (azul), uma via (cinza) e uma via aprimorada (vermelho) .	65
5.20	Média e desvio padrão da previsão de subsidência da superfície nos sensores de deslocamento obtida com os modelos: de referência (preto), duas vias (verde), uma via (vermelho), uma via aprimorado (azul) e a priori (magenta). Barras de erro indicam desvio padrão de \pm nas posições monitoradas	66

Lista de Tabelas

5.1	Parâmetros e dados da simulação	50
-----	---	----

Lista de Símbolos

A	Seção transversal à direção do fluxo, p. 16
E	Módulo de <i>Young</i> , p. 19
$E(\varepsilon(i + 1))$	Média de ε no instante de tempo i , p. 34
F	Simulador do problema direto, p. 27
Fr	Função fracionária do fluxo, p. 21
J	Jacobiano de formações, p. 14
K	Permeabilidade, p. 15
$K(i)$	Matriz de ganhos do filtro de <i>Kalman</i> , p. 26
K_H	Permeabilidade horizontal, p. 15
K_V	Permeabilidade vertical, p. 15
M	Módulo de <i>Biot</i> , p. 18
M	Quantidade de modelos empregados na construção da aproximação do erro de modelagem, p. 34
N_d	X, p. 29
N_r	X, p. 29
P	Pressão no meio, p. 16
$R(i)$	Matriz de covariância do erro de medição, p. 27
S	Matriz diagonal que contém valores singulares, p. 28
S_S	Superfície específica, p. 15
U	Matriz ortogonal que contém autovetores, p. 28
V	Matriz ortogonal que contém autovetores, p. 28

Δt_{micro}	Passo de tempo problema de transporte, p. 23
α	Coefficiente de <i>Biot-Willis</i> , p. 18
\bar{F}	Modelo computacional de maior precisão, p. 32
β	Compressibilidade da matriz porosa, p. 20
δ_{ij}	Delta de <i>Kronecker</i> , p. 19
λ_s, μ_s	Constantes de <i>Lame</i> , p. 19
$\langle S_k^f(i) \rangle$	Media dos estados estimados, p. 26
μ	Viscosidade do fluido, p. 16
ϕ^0	Porosidade à Pressão de Referência, p. 14
ϕ	Porosidade Lagrangiana, p. 14
ρ	Densidade do fluido, p. 16, 17
ρ_s	densidade da fase sólida, p. 18
\emptyset	Porcentagem da informação contida em valores singulares a ser recuperada, p. 29
C	Módulo de elasticidade da fase sólida, p. 19
σ_T	Tensão total para a mistura sólido-fluido, p. 19
σ	Tensor de tensões efetivas da matriz porosa, p. 19
n	Vetor unitário normal à superfície, p. 17
u	Deslocamento da fase sólida, p. 16
\mathbf{v}_D	velocidade de <i>Darcy</i> , p. 16
\mathbf{v}_f	Velocidade do fluido, p. 16
c	Constante de <i>Kozeny</i> , p. 15
$cov(\varepsilon(i+1), \varepsilon(i+1))$	Covariância de ε no instante de tempo i , p. 34
g	Aceleração da gravidade, p. 16
k_{ro}^*	Permeabilidade limite do ógua, p. 22
k_{rw}^*	Permeabilidade limite da água, p. 22

m	Massa, p. 17
n	Porosidade Euleriana, p. 14
p^0	Pressão de Referência, p. 14
q	Fluxo no meio poroso, p. 16
t	Instante de tempo, p. 14
v	Coefficiente de <i>Poisson</i> , p. 19
z	Profundidade, p. 16
V_0	Volume da Rocha no Tempo Inicial, p. 14
V_P	Volume Poroso, p. 14
V_t	Volume da Rocha, p. 14
\mathbb{k}_R	Módulo de compressibilidade da matriz porosa, p. 18
\mathbb{k}_S	Módulo de compressibilidade do grão sólido, p. 18
\mathbb{k}_f	Módulo de compressibilidade do fluido, p. 18
ρ_o	Densidade de referência do fluido, p. 18
N_e	Quantidade de realizações do conjunto, p. 26
$P^f(i)$	Matriz de covariância dos estados estimados, p. 26
$S_k^a(i)$	Vetor de estados assimilado, p. 26
$S_k^f(i)$	Vetor de estados e parâmetros estimados, p. 26
$d_k(i)$	Vetor de medições no instante de tempo i , p. 26
$e_1k(i)$	Erro de modelagem, p. 27
ε_k	Aproximação do erro de modelagem, p. 33
m_f^*	Conteúdo de massa fluido, p. 18
S_{orw}	Saturação crítica do óleo, p. 22
S_{ro}	Saturação residual do óleo, p. 22
S_{rw}	Saturação residual da água, p. 22

S_{wc}	Saturação crítica da água, p. 22
S_w	Saturação da água, p. 21
λ_o	Mobilidade da fase óleo, p. 21
λ_t	Mobilidade da mistura água-óleo, p. 21
λ_w	Mobilidade da fase água, p. 21
μ_o	Viscosidade do óleo, p. 21
μ_w	Viscosidade da água, p. 21
$\tilde{\epsilon}_{1k}$	Erro de modelagem do modelo computacionalmente eficiente, p. 33
f_w	Função fraçãoária da fase água, p. 21
k_{ro}	Permeabilidade relativa do óleo, p. 21
k_{rw}	Permeabilidade relativa da água, p. 21

Lista de Abreviaturas

<i>CFL</i>	Condição de estabilidade Courant Friedrichs Lewy, p. 22
<i>EnKF</i>	Ensemble Kalman Filter, p. 1
<i>GPS</i>	Sistema de Posicionamento Global, p. 5
<i>MAP</i>	Maximum a Posteriori, p. 7
<i>MCMC</i>	Markov Chain Monte Carlo, p. 7
<i>MC</i>	Método de <i>Monte Carlo</i> , p. 34
<i>RML</i>	Radomized Maximum Likelihood, p. 7
<i>TSVD</i>	Truncated Singular Value Decomposition, p. 12

Capítulo 1

Introdução

Previsões e diagnósticos construídos a partir de simulações computacionais podem servir como base para tomada de importantes decisões, mas isso fica condicionado à capacidade em estimar as incertezas neles contidos. Estas, inevitáveis e de presença generalizada, resultam da propagação das incertezas contidas nos dados, na geologia do reservatório e nas hipóteses que conformam o modelo.

Parâmetros de modelos podem ser incertos dado a limitação das técnicas e equipamentos de medição ou da escassez de informações (caso típico em observações de campo). Outra fonte de incertezas reside na utilização de modelos fenomenológicos simplificados, que descrevem, limitadamente, a física dos processos envolvidos. Modelos simplificados tem seu uso justificado por representarem um balanço atraente entre o custo computacional e os aspectos físicos capturados.

A compreensão e mitigação das incertezas nas previsões de modelos computacionais podem ser perseguidas através da combinação de dados monitorados do sistema real com métodos de estimação de parâmetros (problemas inversos e assimilação de dados) e propagação de incertezas. Tendo esse quadro como motivação, neste trabalho são adaptados e implementados métodos baseados em filtros estocásticos para a assimilação de dados em modelos de reservatórios [16]. O objetivo é robustecer modelos previsionais que consideram o efeito acoplado entre o escoamento no meio poroso e a deformação induzida na rocha. Esta abordagem é pouco tratada na literatura atual, entre os trabalhos publicados com este propósito destacam-se [12, 50].

Trabalhos como [13, 49, 74] mostraram a capacidade de assimilação de dados do filtro de *Kalman* por conjunto de amostras (tradução livre para o português de *Ensemble Kalman Filter, EnKF*). Este tipo de filtro tem sido amplamente empregado na assimilação de dados em modelos de escoamento em reservatórios de petróleo por seu balanço atraente entre custo computacional e qualidade das previsões geradas pelos modelos assimilados. Nesta tese procura-se estender estas características do *EnKF* à assimilação de modelos de reservatório de petróleo que consideram o

acoplamento entre os efeitos hidrodinâmicos e geomecânicos, a metodologia de aproximação do erro de modelagem proposta em [86] é introduzida na formulação com intento de aprimorar o desempenho deste filtro em este tipo de problema.

1.1 Motivação e Contextualização

Para caracterizar um reservatório de petróleo diversos tipos de medições e análises são feitos. Amostras da rocha do reservatório podem ser obtidas através da perfuração de poços. Testes nessas amostras permitem quantificar localmente propriedades como porosidade, permeabilidade, saturações residuais, módulo de elasticidade entre outros. É possível também medir propriedades dos fluidos nas vizinhanças dos poços. Além da perfuração de poços, estudos sísmicos e geológicos dão suporte ao trabalho de delimitação do contorno, identificação de falhas e uma primeira caracterização das heterogeneidades da geologia do reservatório.

Estas informações comumente são usadas na construção de modelos que visam simular a dinâmica do reservatório. A importância deste tipo de ferramenta está em seu uso como suporte à gestão de reservatórios. A qualidade dos resultados obtidos dependem, em grande medida, da combinação adequada entre a complexidade dos modelos empregados na formulação fenomenológica, o conhecimento geológico do reservatório e a capacidade de quantificação das incertezas envolvidas. Geocientistas e engenheiros não podem caracterizar o reservatório e suas propriedades com precisão, por isto normalmente recorrem ao conhecimento a priori do reservatório e a uma descrição probabilística [78].

Na Indústria do Petróleo, a simulação de modelos de reservatórios adquire um papel relevante no cálculo do tamanho das reservas de petróleo, de fatores de recuperação em função dos cenários de produção e da previsão das curvas de produção de um determinado reservatório de petróleo. O conhecimento do comportamento dos fluidos no reservatório, assim como a estrutura geofísica deste, são básicos na tentativa de otimização do processo de recuperação de petróleo [54].

As incertezas podem ser distinguidas entre as originadas por processos aleatórios e as que tem sua origem no conhecimento ou compreensão limitada de um processo [11]. Tipicamente as incertezas contidas no processo de modelagem, simulação e assimilação de dados em reservatórios de petróleo tem sua origem no conhecimento limitado dos processos que intervêm.

Estas incertezas, classificadas como de conhecimento limitado, podem-se encaixar nas seguintes fontes de incertezas. Incertezas originadas por erro de medição; incertezas originadas na pouca informação disponível (informação esparsa); incertezas originadas pelo erro de modelagem; incertezas originadas pela heterogeneidade característica das rochas [11].

Pyrcz e White [78] definem uma outra fonte de incertezas que está relacionada com a validade da decisão de assumir as propriedades dos reservatório como estacionárias do ponto de vista probabilístico e considerando a variação espacial. Esta fonte de incertezas poderia ser classificada dentro do erro de modelagem, mas, [78] dão um destaque especial ao fato de que propriedades como a porosidade podem sofrer grandes alterações durante o processo de extração de petróleo. A porosidade e a permeabilidade estão, por definição, correlacionadas, entretanto, são poucas as formulações que recuperam esta relação, a mais conhecida é a relação empírica de Kozeny-Carman [93].

Dados sísmicos oferecem uma boa cobertura do reservatório, mas com baixa precisão e resolução no que diz respeito a caracterização das heterogeneidades das rochas. Conceitos estratigráficos podem guiar a interpolação entre os poços e prover um quadro plausível para uma interpretação quantitativa da sísmica, porém ainda longe de uma compreensão exaustiva do reservatório. Por isto, as incertezas devem ser integradas aos modelos do reservatório [78]. Ainda, se o modelo do processo geológico conseguisse modelar perfeitamente a interação das variáveis que controlam o reservatório, sua distribuição espacial assim como as condições iniciais e de contorno continuam sendo incertas [78].

A Geoestatística oferece uma abordagem conveniente para combinar várias fontes de informação e seus modelos de incertezas em um conjunto de realizações consistentes da heterogeneidade do reservatório, e que, juntas conseguem representar o modelo geológico [24, 38]. Não é possível, em geral, restringir a distribuição das propriedades da rocha para honrar as restrições físicas que governam a erosão, transporte, deposição e diagênese de materiais sedimentários e as observações das propriedades nas posições dos poços. Em muitos casos, contudo, a covariância espacial pode ser usada para quantificar a plausibilidade geológica de uma distribuição das propriedades da rocha [75].

A heterogeneidade de um meio poroso, além de ser uma fonte de incertezas, também desempenha um papel relevante na hidrodinâmica do escoamento dos fluídos no interior do reservatório. Um exemplo disto é o fenômeno conhecido como formação de dedos ou *viscous fingering*, nele a interface de dois fluídos com viscosidades diferentes ao escoar em um meio poroso heterogêneo gera uma região de mistura macroscópica com um perfil de deslocamento irregular na frente de avanço do fluido injetado. Dependendo do tamanho da região de mistura algumas secções do reservatório podem ser isoladas pelas irregularidades [36].

Processos como a recuperação secundária de petróleo, onde um fluido, geralmente água, é injetado no reservatório com o objetivo de deslocar o óleo na direção do poço produtor podem se tornar ineficientes como resultado da formação de dedos. A formação de frentes irregulares (*viscous fingering*) em um reservatório heterogêneo

pode levar a que a água injetada isole volumes significativos de petróleo recuperável, ou até ter uma chegada prematura da água aos poços produtores. A heterogeneidade do meio poroso afeta diretamente o crescimento da região de mistura macroscópica [9]. A chegada da água ao poço produtor é conhecida como *breakthrough* e pode inviabilizar o processo de produção, levando em alguns casos à desativação do poço.

A caracterização das heterogeneidades geológicas dos meios porosos, bem como as propriedades dos fluídos são fatores importantes na identificação do potencial de produção das rochas portadoras de petróleo, ajudando a prever o comportamento futuro desses reservatórios [76].

Na tentativa de construir modelos de reservatórios confiáveis empregando todas as fontes disponíveis de informação surgiu o processo chamado de ajuste de histórico ou assimilação de dados. O ajuste histórico envolve a determinação de um conjunto de parâmetros do reservatório que vai fazer a previsão do modelo o mais próximo possível para os dados do histórico observado. Há duas áreas de interesse no ajuste histórico. Em primeiro lugar, as diferentes abordagens para a construção de modelos de reservatórios ajustados e, em segundo lugar, os diversos métodos para gerar um algoritmo apropriado que calcule e corrija a diferença entre os dados do modelo e os dados históricos [72]. No decorrer do processo de produção do reservatório, novas medições são feitas e geram a oportunidade de corrigir constantemente os resultados obtidos pela simulação do reservatório para dar confiabilidade às predições feitas com o modelo. Um modelo, calibrado adequadamente permite que diferentes alternativas de produção possam ser avaliadas.

Com o aumento de capacidade computacional, tem-se tornado possível simular modelos computacionais que acoplam a hidrodinâmica como os efeitos geomecânicos. Diversos trabalhos têm mostrado a importância de considerar os efeitos geomecânicos induzidos no reservatório pelo escoamento dos fluidos. O estudo do acoplamento entre a geomecânica e o fluxo multifásico está se tornando cada vez mais importante na engenharia de reservatórios. Este acoplamento desempenha um papel de destaque em formações descobertas e exploradas em grandes profundidades. Durante a recuperação secundária do hidrocarbonetos devido as forças de imbibição da água mudanças na pressão dos poros desencadeia perturbações no equilíbrio mecânico do meio poroso conduzindo a variações nas tensões que alteram propriedades da rocha como a permeabilidade e a porosidade [70].

Recentemente, tem sido reconhecido que um acoplamento eficiente entre os efeitos geomecânicos e hidrodinâmicos do escoamento em reservatórios é relevante para gerar previsões precisas de fluxo [50]. Enquanto a modelagem acoplada dos efeitos Geomecânicos e Hidrodinâmicos tem evoluído significativamente nos últimos anos, a análise e implementação do correspondente problema inverso continua em estágios iniciais. Somente umas poucas tentativas tem sido feitas para inverter con-

juntamente dados do poço e informações de deformações da superfície da terra em modelos que acoplam a geomecânica e a Hidrodinâmica [50].

As propriedades dos reservatórios não podem ser medidas extensivamente e de forma precisa. Em contraste, as respostas dos modelos podem ser medidas com relativa precisão, pelo exemplo, taxas de produção de água e óleo são medidas regularmente, bem como a pressão no fundo do poço. O ajuste histórico de dados comumente emprega estes dados de produção na estimação das propriedades de escoamento do reservatório. Este tipo de dados provê informações médias entre os poços ou na vizinhança da localização do poço enquanto observações dos processos Geomecânicos, tais como subsidência da superfície, podem prover informação extensiva adicional. O sistema de posicionamento global (*GPS*), pelo exemplo, podem monitorar o processo de deformação da superfície continuamente [12].

1.2 O *EnKF* na assimilação de Dados para Modelos de Reservatórios de Petróleo

O *EnKF* é baseado no filtro de *Kalman* introduzido por [52]. Este é um filtro recursivo empregado na estimação de estados de sistemas dinâmicos lineares a partir de medições perturbadas com ruído. No filtro de *Kalman* o estado atual do sistema está associado a uma incerteza que é expressada através da matriz de covariância e a uma equação de observação. A equação de observação relaciona uma combinação linear dos estados às medições.

Para dar conta de não linearidades o filtro de *Kalman* foi estendido (*Extended Kalman Filter*)[39]. Esta extensão emprega linearizações do modelo e das equações de observação ao redor da média estimada. O filtro de *Kalman* estendido apresenta bom desempenho em várias aplicações, como por exemplo, estimação de trajetórias e estimação de parâmetros em modelos hidrológicos [27, 29, 44, 61].

Ao trabalhar com modelos altamente não lineares, o filtro de *Kalman* estendido mostrou-se deficiente, além disso, o custo computacional de empregar esta técnica em sistemas de grande escala torna-se inviável, uma vez que é preciso armazenar e atualizar matrizes de covariância de dimensão elevada. Para trabalhar com problemas não lineares de grande escala foi proposto então o filtro de *Kalman* por conjunto de amostras (tradução livre do *Ensemble Kalman Filter*) *EnKF* [10].

O *EnKF*, originalmente introduzido por [30], é uma variante do filtro de *Kalman* que combina o Método de Monte Carlo com as características recursivas do filtro de *Kalman*. No *EnKF* a matriz de covariância do erro é aproximada à matriz de covariância de um conjunto finito de amostras que são recursivamente calibradas com medições. Em [30], o *EnKF* foi empregado na previsão de variáveis dinâmicas

em sistemas meteorológicos oceânicos. O primeiro trabalho empregando o *EnKF* na indústria do petróleo foi [66], posteriormente outros trabalhos nesta indústria começaram a surgir [34, 40, 65, 67, 71].

O *EnKF* apresenta desempenho ótimo quando aplicado a problemas lineares com distribuição *Gaussiana* e empregando grandes quantidades de amostras. O uso de conjuntos de amostras pequenos induz, nos resultados obtidos com o *EnKF*, crescimento espúrio das correlações ao longo do cômputo da matriz de covariância. Para reduzir os efeitos do crescimento espúrio da correlação com conjuntos de amostras pequenos é introduzido na formulação do *EnKF* a função de localização (*Localization*), aplicações em oceanografia e meteorologia frequentemente empregam esta função devido à quantidade de observações empregadas [43, 46, 47].

O erro na estimação da covariância depende do tamanho do conjunto de amostras e da correlação real. Este erro aumenta com a diminuição da correlação e do tamanho do conjunto de amostras [46]. Em aplicações na indústria do petróleo o uso da função de localização tem se tornado relevante em casos nos quais o tamanho do conjunto de amostras é pequeno [13, 25, 83]. Ao empregar observações do processo geomecânico, tais como subsidência, a quantidade de informações aumenta consideravelmente, o que também tem influência na importância do uso da função de localização no *EnKF*.

A estratégia de solução do *EnKF* se baseia na execução de dois passos. O primeiro é o passo de estimação, no qual o problema direto é resolvido empregando o modelo do problema estudado, tendo como resultado o cálculo das variáveis de estado do problema. O segundo passo é o de assimilação, no qual as medições são empregadas no ajuste dos parâmetros e das variáveis do problema. No passo de assimilação, o *EnKF* atualiza simultaneamente, a partir dos dados observados, parâmetros do modelo e variáveis de estado do modelo. Isto induz inconsistências físicas entre as variáveis de estado [41, 85]. Para superar este problema uma formulação iterativa do *EnKF* é empregada. Nesta formulação, um ciclo de consistência é executado para garantir a satisfação das equações de balanço [90].

Em [95] é apresentado um estudo do desempenho do *EnKF* em diversos problemas (*toy problems*) em que o *EnKF* apresenta dificuldades. Isto é feito com o propósito de avaliar a qualidade da assimilação efetuada empregando o *EnKF*. [95] chega à conclusão que ainda quando o *EnKF* não fornece uma caracterização perfeita da variabilidade das variáveis do reservatório, suas características de custo computacional e desempenho fazem ele atraente para seu uso em problemas de ajuste histórico de dados em reservatórios de petróleo. Igualmente, para um caso linear, [95] indica que executar novamente o passo de tempo com os parâmetros assimilados, para recalcular as variáveis dependentes do tempo, procurando evitar valores não-físicos (em esta tese denominado ciclo de consistência) não tem base teórica.

Ainda assim, nos testes efetuados no desenvolvimento desta tese a utilização da consistência se mostrou absolutamente necessária para estabilizar a assimilação.

Entre as características mais relevantes do *EnKF* encontram-se [91]:

- O *EnKF* incrementalmente atualiza os modelos do reservatório, assimilando os dados de produção sequencialmente no tempo assim que estes são disponibilizados; isto é ideal para aplicações em tempo real.
- Um conjunto de modelos de reservatório que reflete os dados de produção mais atuais é sempre mantido. Assim, as previsões e incertezas do desempenho do reservatório estão sempre disponíveis no conjunto de modelos e podem ser empregados para estudos de otimização.
- O *EnKF* é computacionalmente eficiente devido a sua formulação e potencial uso de computação paralela.
- O *EnKF* pode ser empregado junto com qualquer simulador de reservatórios.
- O *EnKF* não precisa de cálculos de otimização e de cálculos de coeficientes de sensibilidade.

Iglesias [49] apresenta uma comparação entre métodos de assimilação de dados que se baseiam em aproximações *Gaussianas*. Estes métodos são o *EnKF*, *MAP* (*Maximum a Posteriori*) e *RML* (*Radomized Maximum Likelihood*). É avaliada a capacidade desses métodos de reproduzir a média e a variância da distribuição posterior caracterizada empregando o *MCMC* (Markov Chain Monte Carlo) .

O *EnKF* tem a desvantagem de assumir aproximações Gaussianas no passo de assimilação [64], para alguns problemas esta aproximação pode resultar muito restritiva podendo levar a soluções inconsistentes ou até induzir instabilidades numéricas. Esta restrição é particularmente relevante em problemas fortemente não Gaussianos. Igualmente, uma vez a assimilação de dados empregando o *EnKF* é um processo sequencial no tempo, isto implica a reinicialização das realizações após cada passo de assimilação, o que desde uma perspectiva computacional pode levar a gasto de tempo de computação acessando disco ou memória e adiciona complexidade à implementação computacional do processo de assimilação de dados.

1.3 O *EnKF* e o *ES*

O *Ensemble Smoother* (*ES*) é um método de assimilação de dados alternativo ao *EnKF*, ele surgiu na tentativa de evitar a reinicialização das realizações de forma sequencial quando a assimilação é executada. O *ES* difere do *EnKF* em que em este

método é computada uma única atualização global em um único passo no domínio do espaço-tempo evitando as atualizações recursivas no tempo que são feitas ao empregar o *EnKF* [82].

Por outro lado, tanto o *EnKF* quanto o *ES* empregam um conjunto (*ensemble*) de realizações para representar as incertezas e a atualização das variáveis assimiladas é feita minimizando a variância. O *ES* foi introduzido por [88], nesse trabalho foi encontrado que aplicado a um modelo de circulação do oceano o *EnKF* oferece melhores resultados do que o *ES*. Em [35] são comparados o *EnKF* e o *ES* aplicados às equações de *Lorenz*, em [32] *ES* foi testado aplicado a modelos lineares.

O *EnKF* e o *ES* resolvem a mesma formulação Bayesiana, o *EnKF* tem uma abordagem recursiva no tempo, em ele é assumido que as medições são independentes no tempo, igualmente é assumido que o reservatório pode ser modelado como uma cadeia de *Markov*. Para problemas dinâmicos lineares o *EnKF* e o *ES* chegam a soluções idênticas [32, 82]. Ainda assim, para problemas dinâmicos não lineares e particularmente para modelos com dinâmica caótica, o *EnKF* tem apresentado um desempenho superior ao desempenho do *ES* [35, 88] uma vez as assimilações recursivas efetuadas pelo *EnKF* mantém o desempenho do modelo com um comportamento próximo à solução real que é representada pelas medições [82].

Em [82] também concluem que os modelos de simulação de reservatórios corresponde a problemas matemáticos mais bem condicionados quando comparados com aqueles dos sistemas dinâmicos caóticos que foram empregados em [35, 88] para testar o *ES*. Se é possível assumir que o modelo do reservatório é estável quando são aplicadas pequenas perturbações nas condições iniciais e nos parâmetros assimilados, então o *ES* tem a capacidade de fornecer resultados similares aos obtidos com o *EnKF* mas com a vantagem de que o *ES* é um método mais eficiente computacionalmente e mais fácil de implementar [82].

Em [14] é demonstrado que ainda quando o *ES* consegue acabar com o custo de reiniciar o simulador de reservatório e com os problemas de convergência que induze a formulação original do *EnKF*, é difícil obter bons resultados empregando o *ES* sem empregar iterações. Igualmente, é preciso empregar a formulação iterativa do *EnKF* como as apresentadas em [42, 62, 90] para evitar problemas de convergência e inconsistências físicas entre as variáveis de estado [41, 85].

Versões iterativas do *ES* como o *Multi Data Assimilation - Ensemble Smoother* (*ES-MDA*) proposto em [28] conseguem efetuar uma atualização global dos parâmetros assimilados eliminando a necessidade de executar o processo de assimilação de forma sequencial o que reduz consideravelmente o custo computacional e aumenta a eficiência desta técnica preservando bons resultados quando comparados a abordagens iterativas do *EnKF*.

1.4 O erro de modelo (ME) na assimilação de Dados

O foco desta tese é pesquisar o uso de aproximações do erro de modelagem na assimilação de dados procurando possibilitar o uso de formulações de modelos de reservatórios acoplados à geomecânica simplificados e que ainda permitam gerar boas previsões, mas aliviando o custo computacional de executar modelos com formulações complexas. Desta forma, ainda quando o *ES-MDA* apresenta um custo computacional menor com um bom desempenho à hora de assimilar dados, o uso em esta pesquisa do *EnKF* iterativo baseia-se na possibilidade de para um determinado instante de tempo, durante o processo de assimilação, poder quantificar de uma forma aproximada quanto da discrepância das variáveis estimadas com respeito às medições é induzida pela discrepância entre formulações do mesmo problema e quanto pelo processo de calibração paramétrica executado no ajuste histórico de dados.

Em esta tese é proposto o uso de um *EnKF* alternativo capaz de levar em consideração, junto com o processo de calibração, erros de modelo [53] induzidos pela prática de usar modelos com física limitada, tipicamente motivado por menores custos computacionais. Os efeitos prejudiciais de não levar em conta os erros do modelo na calibração de modelos computacionais complexos atraíram atenção substancial nos últimos vinte anos, conforme relatado na revisão recente [77] e suas referências, e, particularmente na modelagem de subsuperfície [60, 73, 94]. Estes estudos abrangem aspectos formais e teóricos [51] e aplicações em diferentes áreas [1, 48, 84, 86]. É de interesse em esta tese conjugar filtros de Kalman e o erro de modelo, de forma sequencial e recursiva, para melhorar o desempenho do *EnKF* em um cenário de escoamento acoplado à geomecânica em reservatórios de petróleo.

Uma aplicação em hidrologia envolvendo erro de modelo dentro da assimilação de dados usando *EnKF* foi introduzida e detalhada em [86]. Os erros de modelo provêm de diferentes fontes relacionadas à física simplificada, condições de contorno e geometria incompletas, e métodos de redução, entre muitos outros. Tal diversidade dá origem a diferentes abordagens para tratar as limitações do modelo. Este trabalho parte das ideias propostas em [53] e [51] com base na inserção de um erro de modelo aditivo na equação de observação. No contexto desta tese, associado à etapa de assimilação da filtragem. Dentro do contexto do método de divisão de tensão fixa, exploramos o erro de modelagem visando construir uma "formulação aprimorada de uma via" exibindo maior precisão e desempenho computacional quando comparado ao *EnKF* empregando a formulação de uma via clássica.

1.5 Organização do texto

Este texto está organizado da seguinte forma. No capítulo 2 é apresentada a formulação empregada na modelagem de reservatórios. Primeiro são definidas as principais propriedades petrofísicas da rocha reservatório, a seguir é apresentada a lei de *Darcy* e o modelo de escoamento monofásico em meio poroso poroelástico com acoplamento geomecânico.

No capítulo 3 é apresentada a implementação numérica dos modelos de duas vias e de uma via, baseados na estratégia sequencial de acoplamento, divisão da tensão média fixa. Na primeira parte deste capítulo é apresentado o algoritmo que ilustra a estratégia de solução iterativa entre as equações de escoamento em meio poroso e a geomecânica do reservatório, posteriormente é apresentada a formulação discreta do problema de escoamento em meio poroso e finalmente é apresentada a formulação discreta do problema de deformação elástica tanto do reservatório como das rochas adjacentes ao reservatório.

No capítulo 4 é introduzida a assimilação de dados. No começo é definida a assimilação sequencial de dados, depois é apresentado o *Ensemble Kalman Filter (EnKF)*, após isso é documentada a inversão de dados empregando a metodologia *Truncated Singular Value Decomposition (TSVD)*, posteriormente é explicado o ciclo de consistência das variáveis de estado e a aproximação do erro de modelagem. Na parte final do capítulo é apresentado o algoritmo de assimilação de dados que entrelaça as formulações de duas vias e de uma via ao *EnKF* e ao erro de modelo construído a partir das discrepâncias entre as formulações.

No capítulo 5 são apresentados os resultados numéricos e a análise desses resultados. Neste capítulo são apresentadas comparações entre o desempenho no processo de calibração e identificação paramétrica assim como a qualidade das previsões geradas pelos modelos de duas vias, de uma via e de uma via aprimorado. Comparações com as respostas dos modelos empregados a priori assim como com o modelo de referência empregado na calibração também foram feitas. Por último, no capítulo 6 são apresentadas as conclusões.

Capítulo 2

Modelagem de Reservatórios de Petróleo

Um reservatório de petróleo é um meio poroso que contém hidrocarbonetos. A função do reservatório é conter e permitir, no seu interior, o escoamento dos hidrocarbonetos que originalmente foram produzidos por uma rocha geradora. Assim, um reservatório de petróleo é conformado por dois tipos de rocha, rochas porosas onde os hidrocarbonetos são armazenados e rochas selantes que são as encarregadas de inibir a saída dos hidrocarbonetos do reservatório.

Neste capítulo, primeiro são documentadas algumas das principais propriedades dos reservatório de petróleo; em seguida, são apresentadas a lei de *Darcy* e as equações de balanço aplicadas ao reservatório e é formulado o problema de deformação elástica das rochas adjacentes ao reservatório; finalmente é apresentado o modelo de escoamento em meio poroso acoplado à geomecânica do reservatório.

2.1 Propriedades dos Reservatórios de Petróleo

2.1.1 Porosidade

Os poros são pequenos espaços que existem dentro de uma rocha, tipicamente de tamanho entre 1 a 200 μm . Estes espaços podem estar ou não interconectados. A porosidade é definida como a fração da rocha que é espaço poroso, ou seja, o volume total dos poros dividido pelo volume total da rocha. Assim, a porosidade pode ser classificada em dois tipos: porosidade efetiva que é aquela que somente leva em conta o volume dos poros interconectados e porosidade total que considera o volume total dos poros. Somente poros interconectados podem transmitir fluidos, por isto em simulação de reservatórios ao falar da porosidade é considerada unicamente a porosidade efetiva [16].

A porosidade de um reservatório de petróleo pode variar dependendo da posição

no reservatório na qual seja tirada a amostra para a medição, se este é o caso, o campo de porosidade é considerado heterogêneo. Além da variabilidade espacial, o campo de porosidades também pode apresentar variabilidade temporal, em função da compressibilidade da rocha. O gradiente de pressão pode induzir tensões e deformações consideráveis na rocha o que pode levar à mudanças na distribuição e volume do espaço poroso e até ao colapso de alguns dos poros [70].

A porosidade em um meio poroelástico pode ser descrita como: porosidade Euleriana e porosidade Lagrangiana. A porosidade Euleriana n é definida em função do volume total da rocha no instante de tempo atual, já a porosidade Lagrangiana ϕ é definida a partir do volume total em um determinado tempo de referência [20].

$$n = \frac{dV_P}{dV_t} \quad \phi = \frac{dV_P}{dV_0} \quad (2.1)$$

Na equação 2.1, dV_P denota o volume poroso infinitesimal, dV_t denota o volume da rocha infinitesimal no tempo t e dV_0 é o volume da rocha infinitesimal no tempo inicial. A equação 2.2 apresenta a relação entre a porosidade Euleriana e a porosidade Lagrangiana.

$$\phi = Jn, \quad (2.2)$$

onde J é o Jacobiano de deformações. Assumindo a compressibilidade da rocha como constante, esta pode ser definida através da seguinte relação:

$$c_R = \frac{1}{\phi} \frac{d\phi}{dp}. \quad (2.3)$$

Integrando a equação 2.3 podemos obter uma expressão da variação da porosidade em função do gradiente de pressão:

$$\phi = \phi^0 e^{c_R(p-p^0)} \quad (2.4)$$

onde ϕ^0 é a porosidade à pressão de referência p^0 .

2.1.2 Permeabilidade

A permeabilidade é a capacidade da rocha de conduzir fluidos através de seus poros interconectados [16]; a permeabilidade também pode ser interpretada como a medida da condutividade de fluidos de um material [79]. Como sua definição evidencia, existe uma relação muito estreita entre a permeabilidade e a porosidade, e é de se esperar que a maior maior porosidade maior seja a permeabilidade da rocha, ainda assim, a permeabilidade depende de outros fatores próprios do reservatório como sua geometria, por isto, modelos que relacionam a permeabilidade e a poro-

sidade costumam ser empíricos e depender de parâmetros que são ajustados para cada reservatório. Um dos modelos empíricos mais conhecidos que relacionam a permeabilidade e a porosidade é o de Kozeny-Carman [93]:

$$K = \frac{\phi^3}{c(1 - \phi)^2 S_S^2}; \quad (2.5)$$

onde K é a permeabilidade, c é a constante de Kozeny e S_S a superfície específica. A permeabilidade tem dimensões de área e frequentemente é expressa através de unidades de Darcy (D) ou mili-Darcy (mD). Assim como porosidade, a permeabilidade pode variar espacialmente. Além disso, a permeabilidade também pode ter valores diferentes para uma mesma localização em função da direção do fluxo. Por isto em situações praticas é recomendavel assumir a permeabilidade K como um tensor diagonal [16]:

$$\mathbf{K} = \begin{bmatrix} K_{11} & & \\ & K_{22} & \\ & & K_{33} \end{bmatrix} \quad (2.6)$$

Considerando que K_{11} e K_{22} representam o plano horizontal, comumente é possível assumir que $K_{11} = K_{22}$, neste caso poderíamos falar de permeabilidade horizontal $K_H = K_{11} = K_{22}$. A permeabilidade vertical K_V , representada com K_{33} , usualmente é menor que K_H . A permeabilidade é reduzida pela pressão que as rochas capeadoras induzem no reservatório. Este fator deve ser considerado na estimação da permeabilidade da rocha reservatório em poços profundos [3]. Uma característica importante do reservatório, definida pela permeabilidade, é que se $K_{11} = K_{22} = K_{33}$ seu campo de permeabilidades é isotrópico, caso contrario seu campo de permeabilidades é anisotrópico, ou seja, para efeitos práticos, um reservatório com permeabilidade vertical K_V diferente da permeabilidade horizontal K_H é um reservatório anisotrópico.

2.2 Lei de Darcy

O escoamento de um fluido em um reservatório de petróleo é governado pelas leis de conservação da massa, do momento e da energia. Na simulação do fluxo no reservatório, a equação do momento é dada na forma da lei de *Darcy* [23]. Derivada empiricamente, esta lei indica uma relação linear entre a velocidade do fluido (relativa ao sólido) e o gradiente de pressão [17]. Além do gradiente de pressão, a lei de *Darcy* também envolve a taxa de fluxo através do meio poroso à viscosidade do fluido e à permeabilidade do meio.

$$q = -\frac{KA}{\mu}\nabla p. \quad (2.7)$$

Na equação 2.7, q é o fluxo no meio poroso, A é a seção transversal à direção do fluxo, p é a pressão no meio, μ é a viscosidade do fluido e K é o tensor de permeabilidades. A equação 2.7 tipicamente é apresentada em termos da velocidade de *Darcy* \mathbf{v}_D , dita velocidade é definida como o fluxo por unidade de área q/A no meio poroso:

$$\mathbf{v}_D = -\frac{K}{\mu}\nabla p. \quad (2.8)$$

Nota-se, considerando a definição da porosidade, que a velocidade do fluido nos poros \mathbf{v}_f é igual a velocidade de *Darcy* dividida pela porosidade Euleriana $\mathbf{v}_f = \mathbf{v}_D/n$ [16]. As heterogeneidades do meio, assim como a localização dos poços injetores e produtores no reservatório, fazem que a variação espacial da permeabilidade e da pressão seja refletida em variação espacial da velocidade de *Darcy* e das direções de fluxo no domínio do reservatório.

Na equação 2.9 foram introduzidos na lei de *Darcy* os efeitos gravitacionais, nela ρ é a densidade do fluido, g é a aceleração da gravidade e z é a profundidade. Caso o meio seja considerado poroelástico, é preciso considerar a taxa de deformação do sólido na equação do momento, isto leva à seguinte expressão generalizada da lei de *Darcy* [8, 18]:

$$\mathbf{v}_D = -\frac{K}{\mu}(\nabla p - \rho g \nabla z) \quad (2.9)$$

$$\mathbf{v}_D = n(\mathbf{v}_f - \dot{\mathbf{u}}) = -\frac{K}{\mu}(\nabla p - \rho g \nabla z). \quad (2.10)$$

na equação 2.10 \mathbf{u} é o deslocamento da fase sólida e \mathbf{v}_f é a velocidade do fluido.

2.3 Acoplamento Sequencial Entre o Escoamento e a Geomecânica

Para estabelecer a formulação iterativa, que resolve o acoplamento entre o fluxo e a geomecânica do reservatório, foi empregado o algoritmo apresentado em [70] e que se baseia no conceito de divisão de tensões fixas. Este algoritmo foi empregado considerando um meio poroelástico saturado por um líquido “levemente” compressivo. A formulação totalmente acoplada tem como base a teoria de Biot-Coussy [20] para uma formação geológica fortemente heterogênea. Mais especificamente, a variabilidade espacial e a incerteza são incorporadas na permeabilidade e no módulo de Young, e por simplicidade, o coeficiente de Poisson juntamente com o módulo de

bulk local de grãos são assumidos constantes. Nesta formulação também é incorporada a iteração mecânica entre o reservatório e o as rochas adjacentes modeladas como rochas elásticas.

2.3.1 Formulação do Problema

Definindo $\Omega \in \mathbb{R}^d$ ($1 \leq d \leq 3$) como o domínio espacial com contorno suave Γ , ocupado por um meio poroelástico isotrópico, altamente heterogêneo, saturado por um fluido compressível. Considere o intervalo de tempo $I = [0, t_{final}]$. As equações da formulação totalmente acoplada para cada $\mathbf{x} \in \Omega$ e $t \in I$ estão dadas por:

$$\nabla \cdot \sigma_T + \rho_T \mathbf{g} = 0 \quad (2.11)$$

onde σ_T é o tensor de tensão total da mistura sólido-fluido, ρ_T é a densidade total e \mathbf{g} a gravidade. Continuando com a formulação de Coussy [20], é considerada a formulação baseada na porosidade Lagrangiana ϕ . Assim, denotando ρ_s and ρ_f as densidades das fases sólida e fluida, respectivamente, temos:

$$\rho_T = (1 - \phi)\rho_s + \phi\rho_f, \quad (2.12)$$

e o princípio clássico de tensões efetivas dado por:

$$\sigma_T = \sigma_S - \alpha p \mathbf{I}. \quad (2.13)$$

O parâmetro α representa o coeficiente de Biot-Willis dado por:

$$\alpha = 1 - \frac{\mathbb{K}_r}{\mathbb{K}_s} \quad (2.14)$$

onde \mathbb{K}_r e \mathbb{K}_s são o módulo do Bulk do esqueleto da rocha e dos constituintes sólidos respectivamente. O tensor de tensões efetivas σ_S está ligado à deformação do meio pela lei elástica linear:

$$\sigma_S = \lambda_S \nabla \cdot \mathbf{u} \mathbf{I} + 2\mu_S \varepsilon(\mathbf{u}) \quad (2.15)$$

onde \mathbf{u} é o vetor de deslocamentos, ε é o tensor de deformações dado pela parte simétrica de $\nabla \mathbf{u}$ e λ_S e μ_S são o par de constantes de Lamé. As constantes λ_S e μ_S e o módulo de Young \mathbb{E} e o coeficiente de Poisson ν estão relacionados da seguinte forma: $\mu_S = \mathbb{E}/[2(1 + \nu)]$ e $\lambda_S = \mathbb{E}\nu/[(1 + \nu)(1 - 2\nu)]$. Neste trabalho especificamente, ν é considerado homogêneo, assim a variabilidade inerente a λ_S , μ_S , \mathbb{K}_s e α decorre unicamente da heterogeneidade em \mathbb{E} [70]. O balanço de massa da fase fluida é dado por

$$\frac{\partial(\rho_f \phi)}{\partial t} + \nabla \cdot (\rho_f \mathbf{v}_D) = 0, \quad (2.16)$$

onde \mathbf{v}_D é a velocidade do fluido satisfazendo a lei de Darcy:

$$\mathbf{v}_D = -K(\nabla p - \rho_f \mathbf{g} \nabla z) \quad (2.17)$$

com a permeabilidade absoluta K tratada como uma função aleatória no espaço. A lei constitutiva para o conteúdo de massa fluida é dada por [20]:

$$m_f^* = m_{f,o}^* + \rho_f \alpha \nabla \cdot (\mathbf{u} - \mathbf{u}_o) + \frac{\rho_f}{M} (p - p_o), \quad (2.18)$$

o que implica

$$\frac{\partial m_f^*}{\partial t} = \rho_f \alpha \nabla \cdot \dot{\mathbf{u}} + \frac{\rho_f}{M} \frac{\partial p}{\partial t}. \quad (2.19)$$

As equações (2.18) e (2.19) relacionam o conteúdo de massa fluida com a deformação sólida e pressão nos poros. Nas relações acima $m_{f,o}^*$ é o conteúdo de massa de fluida inicial do reservatório, ela é calculada empregando a densidade inicial do fluido $\rho_{f,o}$, ϕ_o a porosidade inicial, $\mathbf{u}_o = \mathbf{0}$ os deslocamentos iniciais, p_o a pressão inicial do reservatório e M o módulo de Biot[8]:

$$\frac{1}{M} = \frac{\phi}{\mathbb{K}_f} + \frac{\alpha - \phi}{\mathbb{K}_s}, \quad (2.20)$$

onde \mathbb{K}_f é o módulo de Bulk da fase fluida. Tomando o traço da decomposição das tensões efetivas (2.13) e combinando com a lei elástica linear (2.15) é obtida a expressão

$$tr \sigma_T = (3\lambda + 2\mu) \nabla \cdot \mathbf{u} - 3\alpha p. \quad (2.21)$$

Inserindo os resultados anteriores em (2.21) e definindo a compressibilidade da matriz porosa β como

$$\beta = \frac{1}{\mathbb{K}_r} = \frac{3}{3\lambda + 2\mu}, \quad (2.22)$$

é obtida a seguinte equação que relaciona as derivadas temporais de deslocamento, pressão e tensão total

$$\nabla \cdot \dot{\mathbf{u}} = \beta \alpha \frac{\partial p}{\partial t} + \frac{\beta}{3} \frac{\partial tr \sigma_T}{\partial t} \quad (2.23)$$

onde $\dot{\mathbf{u}}$ denota a velocidade da fase sólida. A equação acima descreve um mecanismo de “feedback” no qual as mudanças na tensão impactam diretamente a pressão dos poros e representam um aspecto importante nos desenvolvimentos do método se-

quencial a ser apresentado posteriormente. Substituindo (2.19) em (2.16) é obtido

$$\frac{1}{M} \frac{\partial p}{\partial t} + \alpha \nabla \cdot \dot{\mathbf{u}} + \nabla \cdot \mathbf{v}_D = 0. \quad (2.24)$$

Adicionalmente, combinando (2.23) e (2.24), com a compressibilidade total β^* definida por

$$\beta^* = \frac{1}{M} + \beta \alpha^2, \quad (2.25)$$

é obtida a equação de balanço global, em termos do tensor médio de tensões totais σ_T , a pressão dos poros p e a velocidade de Darcy do fluido \mathbf{v}_D

$$\beta^* \frac{\partial p}{\partial t} + \nabla \cdot \mathbf{v}_D = -\beta \alpha \frac{\partial \tilde{\sigma}_T}{\partial t} \quad (2.26)$$

com a tensão total média definida como $\tilde{\sigma}_T = \frac{1}{3} \text{tr}(\sigma_T)$. Após as manipulações anteriores, chega-se ao seguinte conjunto de equações cujas soluções serão aproximadas de forma sequencial, com base no algoritmo divisão de tensões fixas:

$$\nabla \cdot (\lambda_S \nabla \cdot \mathbf{u} \mathbf{I} + 2\mu_S \varepsilon(\mathbf{u})) - \alpha \nabla p - \rho_T \mathbf{g} = 0,$$

$$\beta^* \frac{\partial p}{\partial t} - \nabla \cdot K \nabla p = -\beta \alpha \frac{\partial \tilde{\sigma}_T}{\partial t}, \quad (2.27)$$

$$\mathbf{v}_D = -K(\nabla p - \rho_f \mathbf{g}).$$

Para o domínio adjacente ao reservatório Ω_c ocupado por formações rochosas impermeáveis, é formulado o problema de elasticidade clássico:

$$\nabla \cdot \sigma_S = -\rho_T \mathbf{g}$$

$$\sigma_S = \lambda_{S,l} \nabla \cdot \mathbf{u} \mathbf{I} + 2\mu_{S,l} \varepsilon(\mathbf{u}) \quad (2.28)$$

O índice $l = 1, 2, \dots, 5$ em (2.28) denota o par de constantes de Lamé para cada subdomínio $\Omega_{c,l}$ de Ω_c . O sistema anterior é complementado pelas seguintes condições iniciais e de contorno:

$$v_D \cdot n = Q \text{ on } \Gamma_N^p, \quad p = 0 \text{ on } \Gamma_D^p, \quad p(\mathbf{x}, 0) = p_0 \text{ in } \Omega. \quad (2.29)$$

para o problema de escoamento e para a geomecânica

$$\sigma_T \cdot \mathbf{n} = \mathbf{t}; \text{ on } \Gamma_D^u; \text{ and } \mathbf{u} = 0, \text{ on } \Gamma_D^u, \quad (2.30)$$

onde \mathbf{n} denota o vetor normal à superfície do contorno do domínio e \mathbf{t} é uma força de tração prescrita.

Este problema é usualmente resolvido iterativamente na presença de incertezas, através de um espaço de parâmetro de entrada de alta dimensão, levando a um grande número de graus de liberdade. A fim de reduzir a carga computacional, muitas vezes são adotados modelos acoplados de forma fraca, considerando explicitamente apenas o impacto do escoamento no subsistema da geomecânica, comumente referido como modelos de 1-via. Esta abordagem simplificada, amplamente adotada na geomecânica de reservatórios, consiste em postular como estacionária a tensão média total. Tal suposição é exata apenas no cenário unidimensional, como no problema de consolidação de Terzaghi [5].

A construção de uma formulação de 1-via aprimorada, desconsiderando a variação temporal da tensão média total, termo fonte $\frac{\partial \sigma_T}{\partial t}$ no lado direito de (2.27) procura evitar a solução simultânea dos subsistemas hidrodinâmico e geomecânico. Isto requer conhecimento no desenvolvimento de métodos alternativos de assimilação de dados, incorporando o erro de modelo [57], para construir uma formulação de 1-via aprimorada capaz de reduzir a perda de precisão quando comparada com a formulação de 2-vias mas procurando preservar um baixo custo computacional. Intrinsecamente procura-se mitigar a perda de precisão por meio de abordagens que permitem estimar as possíveis imprecisões envolvidas.

Capítulo 3

Implementação Numérica e Algoritmo

A seguir, apresentamos o método de discretização juntamente com o algoritmo iterativo para resolver o problema direto que acopla o modelo de escoamentos à geomecânica do reservatório e das rochas adjacentes. São definidas também as bases para incorporar o erro de modelagem dentro da abordagem de assimilação de dados empregando o *EnKF*.

3.1 Algoritmo Iterativo

A formulação de divisão de tensão fixa baseia-se em um procedimento iterativo em que o subsistema de escoamento é resolvido antes da geomecânica. A fim de apresentar os passos do algoritmo em evolução, começamos por discretizar as equações governantes usando o método de Euler. Consideramos um passo de tempo constante Δt e um tempo fixo denotado por $t^n = n\Delta t$, onde n denota o número de passos de tempo. Para cada passo de tempo, o sistema é resolvido iterativamente com o subsistema geomecânico. O índice k introduzido indica o número de iteração.

A solução do problema de escoamento a t^n na k -ésima iteração é denotada por $p^{n,k}$, está é computada substituindo a derivada temporal $\frac{\partial p}{\partial t}$ pelo quociente discreto $\frac{(p^{n,k} - p^{n-1})}{\Delta t}$. O termo fonte $\frac{\partial \tilde{\sigma}_T}{\partial t}$ é aproximado usando a metodologia de tensão fixa o que tem como resultado o problema:

$$\beta^* \frac{(p^{n,k} - p^{n-1})}{\Delta t} - \nabla \cdot K \nabla p^{n,k} = -\beta \alpha \frac{(\tilde{\sigma}_T^{n,k-1} - \tilde{\sigma}_T^{n-1})}{\Delta t} \quad (3.1)$$

$$\mathbf{v}_D^{n,k} = -K(\nabla p^{n,k} - \rho_f \mathbf{g}) \quad \text{with} \quad n = 1, \dots, N_T \quad \text{and} \quad k = 1, \dots, N_k$$

Na sequencia, o subproblema de geomecânica na $k - ith$ iteração é presentation como: dado $p^{n,k}$ solução de (3.1) achar $\mathbf{u}^{n,k}$ satisfazendo

$$\nabla \cdot (\lambda_S \nabla \cdot \mathbf{u}^{n,k} \mathbf{I} + 2\mu_S \varepsilon(\mathbf{u}^{n,k})) - \alpha \nabla p^{n,k} - \rho_T \mathbf{g} = 0 \quad (3.2)$$

após resolver para $\mathbf{u}^{n,k}$ o par $(\tilde{\sigma}_T^{n,k}, \tilde{\sigma}_s^{n,k})$ pode ser computado dentro de uma abordagem de pós-processamento considerando $\sigma_T^{n,k} = \sigma_s^{n,k} - \alpha p^{n,k} \mathbf{I}$. Para resumir o procedimento iterativo, apresentamos o algoritmo 1:

Algoritmo 1: ALGORITMO DE MODELO DE DUAS VIAS

- 1 **Passo 0** Conjunto de dados iniciais $p^0 = p_o(\mathbf{x})$, $\phi^0 = \phi^0(\mathbf{x}, 0)$ e $\tilde{\sigma}_T^0 = -\alpha_s p^0(\mathbf{x})$ com $n = 0$ e $k = 0$.
 - 2 **Passo 1** Dadas as condições iniciais ajustar o índice temporal $n = 1$ e o índice de número de iteração $k = 1$ e atribuir $p^n = p^0$, $\phi^n = \phi^0$ e $\tilde{\sigma}_T^n = \tilde{\sigma}_T^0$.
 - 3 **Passo 2** Dada p^n , v_D^n , \mathbf{u}^n , $\tilde{\sigma}_T^n$, ϕ^n ajustar o índice temporal $n = n + 1$ e o índice de iteração $k = 1$ e atribuir $p^{n-1} = p^n$, $v_D^{n-1} = v_D^n$, $\mathbf{u}^{n-1} = \mathbf{u}^n$, $\tilde{\sigma}_T^{n-1} = \tilde{\sigma}_T^n$, $\phi^{n-1} = \phi^n$.
 - 4 **Passo 3** Dada p^{n-1} , ϕ^{n-1} , $\tilde{\sigma}_T^{n-1}$, \mathbf{u}^{n-1} e $\tilde{\sigma}_T^{n-1}$ computar $(p^{n,k}, v_D^{n,k})$.
 - 5 **Passo 4** Dada $(p^{n,k}, v_D^{n,k})$, obtida no passo 3, computar $(\mathbf{u}^{n,k}, \tilde{\sigma}_T^{n,k})$.
 - 6 **Passo 5** Avaliar a convergência do algoritmo de divisão de tensão fixa na iteração k . Se for atingida estabelecer $p^n = p^{n,k}$, $v_D^n = v_D^{n,k}$, $\mathbf{u}^n = \mathbf{u}^{n,k}$, $\tilde{\sigma}_T^n = \tilde{\sigma}_T^{n,k}$, $\phi^n = \phi^{n,k}$ and continue to **Passo 6**; de outra forma estabelecer $k = k + 1$ e retornar para **Passo 2**.
 - 7 **Passo 6** Dada p^n , v_D^n , \mathbf{u}^n , $\tilde{\sigma}_T^n$ computar ϕ^n e retornar a **Passo 2**
-

Vale a pena notar que, para calcular o erro do modelo e obter um modelo direto mais barato no *EnKF*, os resultados numéricos associados à formulação de uma via podem ser facilmente computados (e armazenados) na primeira iteração do loop escoamento - geomecânica ($k = 1$).

3.2 Discretização do Subproblema de Escoamento

Para cada iteração do algoritmo sequencial, foi feita a discretização do problema de escoamento pelo método de elementos finitos mistos localmente conservativo. Assim, foi introduzida a formulação variacional juntamente com as funções de espaço apropriados. Seja $L^2(\Omega)$ o espaço de funções escalares integráveis quadradas com produto interno e norma dada por $(f, g) = \int_{\Omega} fg \, d\Omega$, $\|f\| = \sqrt{(f, f)}$ e seja $H^1(\Omega)$ o espaço de funções com derivadas quadradas integráveis com produto interno natural $(f, g)_1 = (f, g) + (\nabla f, \nabla g)$, $\|f\|_1 = \sqrt{(f, f)_1}$.

A aproximação da hidrodinâmica é baseada no espaço misto de Raviart-Thomas de menor ordem $RT_h^0 \times \mathbb{P}^0$, com \mathbb{P}^0 denotando o espaço as funções constantes por

partes. Introduzindo os conjuntos dimensionais finitos: $V_h = \{\mathbf{v}_h \in RT_h^0, \mathbf{v}_h \cdot \mathbf{n} = Q \text{ on } \Gamma_N^p\}$ e $X_h = \{\mathbf{v}_h \in RT_h^0, \mathbf{v}_h \cdot \mathbf{n} = 0 \text{ on } \Gamma_N^p\}$, o problema matemático resultante pode ser representado como: para cada passo de tempo t^n dado $\{K^n, \beta^*, \beta, \alpha, \tilde{\sigma}_{T_h}^{n,k-1}, \tilde{\sigma}_{T_h}^{n-1}\}$ achar $(\mathbf{v}_{Dh}^{n,k}, p_h^{n,k}) \in V_h \times \mathbb{P}^0$ de tal modo que:

$$((K^n)^{-1} \mathbf{v}_{Dh}^{n,k}, \mathbf{v}_h) - (p_h^{n,k}, \nabla \cdot \mathbf{v}_h) = 0 \quad \forall (\mathbf{v}_h, q_h) \in X_h \times \mathbb{P}^0 \quad (3.3)$$

$$(\beta^* p_h^{n,k}, q_h) + \Delta t (\nabla \cdot \mathbf{v}_{Dh}^{n,k}, q_h) = (\beta^* p_h^{n-1}, q_h) + \beta \alpha (\tilde{\sigma}_{T_h}^{n,k} - \tilde{\sigma}_{T_h}^{n-1}, q_h) \quad (3.4)$$

Desta forma a pressão dos poros é aproximada por interpolações constantes por partes descontínuas, com isto um procedimento de condensação estática é adotado para eliminar os graus de liberdade de pressão e formular o problema somente em termos de velocidade de Darcy. Escolhendo q_h em (3.4) uma função de base constante por partes, é possível eliminar a pressão de poros discreta no nível do elemento. Denotando a restrição de cada variável ao nível do elemento pelo índice T tem-se

$$\beta_T^* p_h^{n,k}|_T = -\Delta t (\nabla \cdot \mathbf{v}_{Dh}^{n,k})|_T + \beta_T^* p_h^{n-1}|_T + \beta \alpha \tilde{\sigma}_{T_h}^{n,k}|_T - \beta \alpha \tilde{\sigma}_{T_h}^{n-1}|_T \quad (3.5)$$

Inserindo (3.5) em (3.3) para todos os elementos, temos o problema colocado apenas em termos de velocidade de Darcy

$$\begin{aligned} (K^n \Lambda_T^n)^{-1} \mathbf{v}_{Dh}^{n,k}, \mathbf{v}_h) + \Delta t \beta^{*-1} (\nabla \cdot \mathbf{v}_{Dh}^{n,k}, \nabla \cdot \mathbf{v}_h) &= (p_h^{n-1}, \nabla \cdot \mathbf{v}_h) \\ &+ \beta^{*-1} \beta \alpha (\tilde{\sigma}_{T_h}^{n,k} - \tilde{\sigma}_{T_h}^{n-1}, \nabla \cdot \mathbf{v}_h) \quad \forall \mathbf{v}_h \in V_h \end{aligned} \quad (3.6)$$

Após resolver a velocidade do Darcy, a pressão descontínua pode ser pós-processada no nível do elemento usando (3.5).

3.3 Discretização das Equações Poromecânicas

Nesta seção é apresentada a aproximação por elementos finitos do subsistema geomecânico. Seja W_h^u e W_h polinômios bilineares contínuos clássicos para deslocamento e descontínuos constantes de partes para a tensão efetivas e as deformações ε . Dada a solução do problema de escoamento $p_h^{n,k}$, ache $(\mathbf{u}_h, \sigma_h, \varepsilon_h) \in W_h^u \times W_h \times W_h$ satisfazendo o seguinte problema de elasticidade variacional discreta

$$(\nabla^s \mathbf{w}_h, \sigma_h^{n,k}) = (\mathbf{w}_h, \mathbf{t})_{\Gamma_N^q} + (p_h^{n,k}, \Delta \cdot \mathbf{w}_h) \quad \forall \mathbf{w}_h \in W_h^u \quad (3.7)$$

$$(\xi_h, \sigma_h^{n,k} - \mathbf{C} \varepsilon_h^{n,k}) = 0 \quad \forall \xi_h \in W_h \quad (3.8)$$

$$(\tau_h, \varepsilon_h^{n,k} - \nabla^s \mathbf{u}_h^{n,k}) = 0 \quad \forall \tau_h \in W_h, \quad (3.9)$$

Explorando a flexibilidade fornecida pelo emprego de polinômios descontínuos, são escolhidas $\{\xi_h, \tau_h\}$ funções de base constante por partes, isto leva às seguintes expressões locais para a deformação e a tensão efetiva no nível do elemento:

$$\varepsilon_h^{n,k} |_{T=} \overline{\nabla^s \mathbf{u}_h^{n,k}} = \frac{1}{|T|} \int_T \nabla^s \mathbf{u}_h^n dT, \quad \sigma_h^{n,k} |_{T=} \frac{1}{|T|} \int_T \mathbf{C} \varepsilon_h^{n,k} dT, \quad (3.10)$$

onde $|T|$ designa a área do elemento T , a linha acima das variáveis denota a média sobre cada elemento, ∇^s é a parte simétrica do gradiente e N_e denota o número de elementos. Substituindo (3.10) em (3.7), e adicionando sobre os elementos N_e é obtido o seguinte problema para os deslocamentos do sólido

$$\sum_{T=1}^{N_e} \int_T \overline{\nabla^s \mathbf{w}_h} \mathbf{C} \overline{\nabla^s \mathbf{u}_h^{n,k}} dT = \sum_{T=1}^{N_e} \int_T p_h^{n,k} |_{T} \nabla \cdot \mathbf{w}_h dT + \int_{\Gamma_N^u} \mathbf{w}_h \cdot \mathbf{g}_h d\Gamma_N^u. \quad (3.11)$$

Depois de resolver o problema anterior com a aproximação bilinear de menor ordem para $\mathbf{u}_h^{n,k}$ as deformações constantes por partes e as tensões podem ser pós-processadas empregando (3.10).

Capítulo 4

Assimilação de Dados Usando o *EnKF*

A assimilação de dados também conhecida na Engenharia de Reservatórios como ajuste de histórico de dados é um tipo de problema inverso, no qual são usadas medições diretas ou indiretas de um fenômeno, sistema ou evento na estimação e calibração de modelos previsionais para simular esse fenômeno, sistema ou evento.

Especificamente, na calibração de modelos previsionais de reservatórios de petróleo procura-se empregar medições como vazão de produção, pressão em poços, leituras de subsidência, medições de propriedades, sísmica, entre outros. Estas medições são registradas durante a operação do reservatório estudado. Dependendo da capacidade do modelo de representar características e aspectos físicos envolvidos na dinâmica do reservatório, podem ser empregados diversos tipos de fontes de informação em sua calibração, pelo exemplo, modelos que consideram o reservatório rígido não podem usar informação de subsidência durante sua calibração.

O objetivo central da assimilação de dados é que os modelos matemáticos calibrados consigam reproduzir o comportamento do reservatório para prever o comportamento futuro do reservatório com um aumento da confiança, e permitir combinar simulações computacionais com métodos de gestão do reservatório. Com modelos confiáveis é possível, por exemplo, explorar diferentes localizações para os poços de injeção, ou otimizar a gestão do reservatório [74].

Como é um tipo de problema inverso, o ajuste de histórico tem a característica de ser mal-posto. Especificamente, a falta de unicidade na solução do problema dá origem a um grande número de combinações possíveis dos parâmetros do reservatório que levam a um bom ajuste dos dados observados. Isto faz com que um único modelo ajustado seja insuficiente para seu uso na gestão do reservatório. Uma solução completa de um problema de ajuste de histórico deve levar, idealmente, a estimativas das incertezas nas propriedades e previsões do reservatório [74].

4.1 Assimilação Sequencial de Dados

A dinâmica e a geologia do reservatório mudam constantemente no decorrer do tempo de sua etapa produtiva. Isto leva a que novas medições e informações do reservatório sejam necessárias. Nas estratégias convencionais de assimilação de dados, é empregado todo ou parte da história de medições na procura de um cálculo aproximado do erro para gerar, assim, a nova calibração do modelo de estados. Isto é refletido em um elevado custo computacional na hora de introduzir novas informações para calibrar o modelo.

Esse quadro deu origem à aplicação de técnicas baseadas em filtros ao ajuste de histórico. Este tipo de técnica permite a, chamada, assimilação sequencial de dados, na qual, o processo de calibração dos modelos previsionais é feito sequencialmente, uma vez que novas medições estão disponíveis. Este tipo de técnicas somente emprega a informação nova, evitando desta forma o re-ajuste dos modelos com informação que já foi inserida. Isto, obriga a que uma vez vão ser usadas as novas medições, seja suspenso o processo de previsão e simultaneamente sejam ajustadas todas as variáveis e parâmetros de todos os modelos, depois de efetuada a calibração, volta a ser executada a etapa de previsão, mas desta vez com os modelos atualizados e para o tempo subsequente, por isto, a assimilação sequencial prevê simultaneamente múltiplos modelos calibrados a diferença de outras abordagens de assimilação de dados.

Se o modelo de previsão e a covariância do erro de previsão são conhecidos para um instante de tempo t_i , onde temos medições disponíveis, com uma determinada covariância do erro de medição, então, é possível calcular um modelo de estados assimilado, com uma covariância de erros assimilados. A grande questão é, então, como estimar ou prever a covariância do erro de previsão do modelo no instante t_i [33].

As matrizes de covariância do erro são definidas no filtro de *Kalman* [52] em termos do estado real. Contudo, o estado real não é conhecido, portanto, são definidas as matrizes de covariância em torno da média do conjunto de modelos (conhecido como *ensemble* na literatura). Desta forma é usada uma interpretação na qual a média do conjunto de modelos é a melhor estimativa e o espalhamento em torno à média é uma definição natural do erro no conjunto médio. Assim, em vez de armazenar uma matriz de covariância completa, é possível representar a mesma estatística do erro empregando um conjunto de modelos de estados apropriado. Dada uma matriz de covariância do erro, um conjunto de modelos de tamanho finito é capaz de prover uma aproximação da matriz de covariância do erro [31].

4.2 O Ensemble Kalman Filter (EnKF)

O *Ensemble Kalman Filter (EnKF)* introduzido por Evensen [30] é uma evolução da aplicação do filtro de Kalman [52], ela foi originada porque no filtro de Kalman a matriz de covariância do vetor de estados é muito grande o que implica em elevado custo computacional em problemas de grande escala. Além disso, o filtro de Kalman, em sua formulação original, é um estimador de estados linear o que limita suas possibilidades de aplicação.

O *EnKF* emprega uma abordagem do Método de Monte Carlo para representar e evoluir a função de densidade de probabilidade conjunta dos estados e parâmetros do modelo. Ele computa passos de atualização recursivos introduzindo uma aproximação onde somente os momentos de primeiro e segundo ordem da função de densidade de probabilidade (média e covariância) prevista são empregados para computar o ajuste dos modelos. Cada realização é atualizada individualmente somando os incrementos de atualização como é feito no filtro de *Kalman* tradicional [10, 80]. A seguir são apresentadas as equações para o cálculo da média e da covariância do conjunto de modelos de estados e parâmetros.

$$\langle S_r^f(i) \rangle = \frac{1}{N_e} \sum_{r=1}^{N_e} S_k^f(i) \quad (4.1)$$

$$P^f(i) = \frac{1}{N_e - 1} \sum_{r=1}^{N_e} \left\{ [S_r^f(i) - \langle S_r^f(i) \rangle] [S_r^f(i) - \langle S_r^f(i) \rangle]^T \right\} \quad (4.2)$$

Nas equações anteriores o índice f faz referência ao processo de estimação das variáveis, N_e é a quantidade de realizações com que é gerado o conjunto, $S_r^f(i)$ é o vetor de estados e parâmetros estimados da realização r , $\langle S_r^f(i) \rangle$ é a média dos estados estimados e $P^f(i)$ sua matriz de covariância. Como foi exposto anteriormente, assim que o sistema evolui no tempo, tanto a média como a matriz de covariância do conjunto podem ser atualizadas e, uma vez novas medições tornam-se disponíveis, a assimilação dos novos dados é executada:

$$S_r^a(i) = S_r^f(i) + G(i) [d_r(i) - HS_r^f(i)], \quad (4.3)$$

onde $S_r^a(i)$ é o vetor de estados assimilado da realização r no instante de tempo em que os dados $d_r(i)$ foram medidos, $G(i)$ representa os ganhos do filtro e são definidos através de:

$$G(i) = P^f(i)H^T [HP^f(i)H^T + R(i)]^{-1}. \quad (4.4)$$

Os dados $d_r(i)$ com que é assimilada a realização r são perturbados por um ruído gaussiano de média zero permitindo considerar as incertezas de medição, assim, a

matriz $R(i)$ é a matriz de covariância dos ruídos do conjunto em um determinado instante de tempo. Os ruídos das diversas realizações são assumidos independentes, por isto a matriz de $R(i)$ é uma matriz diagonal que armazena a variância com que foram geradas as perturbações de cada realização.

A matriz $HP^f(i)H^T + R(i)$ é definida e assim lembrando que tanto $HP^f(i)H^T$ como $R(i)$ são matrizes simétricas a existência da inversa da operação esta garantida, mas não entanto, nem sempre o resultado é uma matriz bem condicionada.

As incertezas no modelo continuam sendo formalmente estimadas empregando o conjunto de vetores de estados. Os modelos ou realizações iniciais usados no *EnKF* são amostras aleatórias da distribuição inicial, que, até certo ponto, permitem superar as limitações da consideração de distribuições Gaussianas. Contudo, passo de assimilação continua baseado nos momentos de primeira e segunda ordem da distribuição, o qual limita o desempenho do método se as distribuições estão muito afastadas da distribuição Gaussiana [2].

Assim que a assimilação dos dados é terminada em um instante de tempo fixo, o vetor de estados de previsão é atualizado com os valores do vetor de estados assimilados e pode-se proceder a um novo passo de integração temporal,

$$S_r^f(i+1) = \mathcal{F} [S_r^a(i)] + e_{1,r}(i) \quad (4.5)$$

aqui, \mathcal{F} representa um operador que pode ser não linear, especificamente nesta tese, \mathcal{F} representará o modelo matemático do reservatório. Portanto \mathcal{F} representa um simulador de fluxo monofásico em meio poroelástico heterogêneo. $e_{1,r}(i)$ é um ruído gaussiano de média zero que é usado para levar em consideração o erro de modelagem. Geralmente, na literatura atual de assimilação de dados em modelos de reservatórios, $e_{1,r}(i)$ é considerado como um erro Gaussiano com média zero, em aplicações práticas, assumir média zero no erro de modelagem dificilmente corresponde à realidade. Para abordar esta problemática na seção 4.5, será introduzido um tratamento especial para este termo.

O uso do *EnKF* na assimilação sequencial de dados é dividida em duas etapas, uma etapa de estimação ou previsão que é executada pelas equações de governo fenomenológicas contidas no simulador (equação 4.5) e uma segunda etapa de assimilação ou análise, a qual é executada sobre os parâmetros e variáveis de estado estimadas, enquanto o passo de tempo se encontra fixo (equação 4.3).

Das equações 4.3 e 4.5 pode-se concluir que mesmo quando \mathcal{F} possa representar operadores não lineares, como é o caso dos modelos de reservatórios, a assimilação empregando o *EnKF* é uma operação linear, na qual, o ajuste dos modelos estimados S_r^f é feito somando o produto do ganho do filtro de Kalman pela diferença entre os dados estimados e de referência, como será tratado na seção 4.4. Isto faz com que os

resultados obtidos na assimilação das variáveis dinâmicas possam apresentar valores sem sentido físico quando o modelo dinâmico é não linear.

4.3 Inversão Usando *TSVD*

Uma das características das matrizes de covariância é que são positivas semidefinidas. O fato de que a matriz de covariância possa conter valores singulares iguais a zero pode resultar em mau condicionamento da matriz do filtro, o que acaba representando uma potencial instabilidade numérica.

Para superar este obstáculo, o ruído de medição ao longo do tempo é considerado como independente ou não correlacionado com o ruído dos outros modelos faz com que a matriz de covariância do ruído de medição $R(i)$ seja uma matriz diagonal. Assim, a variância com que foi gerado o ruído de cada realização constitui os valores da diagonal da matriz $R(i)$ garantindo com isso que $R(i)$ seja uma matriz positiva definida e que o resultado de $HP^f(i)H^T + R(i)$ também seja uma matriz positiva definida.

Essa hipótese sobre o ruído tem implicações significativas. Por um lado, se a variância do ruído é muito pequena, a regularização que $R(i)$ gera vai ser mínima e a estimação da inversa na equação (4.4) pode ser afetada por problemas numéricos, na medida em que as variáveis estimadas ficam mais próximas das variáveis de referência. Por outro lado, se a variância do ruído é muito grande, vai afetar os resultados da estimação reduzindo a precisão. Para tentar diminuir os efeitos negativos do mau condicionamento, implementou-se uma estimativa para a matriz (pseudo) inversa presente na equação (4.4). A metodologia empregada foi *TSVD* (*Truncated Singular Value Decomposition*) [4, 33]. Desta forma a equação (4.6) após sua decomposição em valores singulares *SVD*. pode ser escrita como:

$$HP^f(i)H^T + R(i) = U S_{SVD} V^T, \quad (4.6)$$

onde U e V são matrizes ortogonais que contém os autovetores da soma $HP^f(i)H^T + R(i)$ e S_{SVD} é uma matriz diagonal que contém seus valores singulares. Como as matrizes U e V são ortogonais então:

$$U U^T = I \quad V V^T = I \quad (4.7)$$

o que leva a formulação da pseudo-inversa empregando a decomposição em valores singulares:

$$(HP^f(i)H^T + R(i))^{-1} = V S_{SVD}^{-1} U^T \quad (4.8)$$

Onde S_{SVD}^{-1} é a matriz diagonal com seus componentes iguais ao inverso dos valores singulares contidos em S_{SVD} . Em caso da matriz ter valores singulares iguais a zero, esses valores são truncados imediatamente e não são levados em consideração no cálculo da pseudo inversa. Entretanto, mesmo valores singulares pequenos podem gerar grandes erros no momento do cálculo de S_{SVD}^{-1} . Para superar isto foi implementado o truncamento dos valores singulares, desta forma é fixada uma porcentagem de toda a informação contida nos valores singulares \emptyset a ser recuperada. Arrumando os valores singulares do maior até o menor tem-se:

$$\emptyset \leq \frac{\sum_{i=1}^{N_d} \lambda_i}{\sum_{i=1}^{N_r} \lambda_i} \quad (4.9)$$

Onde N_d é a quantidade de valores singulares a ser mantida na pseudo inversa e N_r é a quantidade total de valores singulares de $HP^f(i)H^T + R(i)$. Assim a pseudo inversa é calculada apenas com os N_d valores singulares que cumprem com o critério. O *TSVD* estabiliza a solução, mas, esta solução não pode ajustar os dados tão bem quanto a solução construída com todos os vetores base do espaço do modelo. Estabilidade é obtida a custa de resolução [7]. Previamente ao processo de aplicação de *SVD* e de inversão, a matriz $HP^f(i)H^T + R(i)$ é reescalada para evitar o truncamento de informações relevantes uma vez as variáveis e parâmetros que compõem o vetor de estados estendido contem grandezas de diferentes magnitudes.

4.4 Consistência das Variáveis de Estado

Como foi comentado anteriormente, o *EnKF* emprega uma estratégia de assimilação sequencial. A utilização do filtro introduz um vetor de estados composto pelas variáveis que efetivamente descrevem o escoamento e por parâmetros que se quer identificar. No presente contexto, tal vetor toma a forma abaixo:

$$S_r^f(i) = \begin{bmatrix} p_r^f \\ \mathbf{v}_{x,r}^f \\ \mathbf{v}_{y,r}^f \\ \mathbf{u}_r^f \\ \phi_r^f \\ \ln(K_r^f) \\ \mathbb{E}_r^f \end{bmatrix} \quad (4.10)$$

A dimensão do vetor de estados $S_r^f(i)$ depende da quantidade de variáveis e parâmetros a serem assimilados, mas também depende da discretização espacial empregada na simulação computacional do modelo. O filtro, necessariamente, identifica em cada instante os parâmetros alvos e, também, as variáveis de estado (como

pressão, saturação e velocidade). Isto promove implicações importantes na satisfação das equações básicas do modelo. Uma limitação referente ao algoritmo *EnKF* convencional é que inconsistência significativa pode ocorrer dentro das variáveis de estado após a etapa de atualização uma vez que os parâmetros dinâmicos e os parâmetros estáticos são atualizados simultaneamente [12, 63].

Note-se na equação 4.10, que a despeito dos parâmetros serem constantes ao longo do tempo, aparecem indexados com o índice i , que está associado ao instante de tempo. Isso é necessário, uma vez que os parâmetros, alvo da estratégia de calibração do modelo, assumem novos valores na medida em que a assimilação é levada a cabo com a incorporação de dados em instantes t_i .

A equação 4.3 constitui o passo de assimilação do algoritmo. Os super índices a e f nos vetores S indicam que são o vetor de estados assimilado e estimado, o vetor d_r representa os dados medidos que são usados no ajuste dos modelos e $G(i)$ representa a matriz de ganhos do filtro de Kalman. O vetor de estados estimados $S_r^f(i)$ é calculado resolvendo o problema de escoamento monofásico em meio poroso poroelástico formado pelas em equações 2.27 para o reservatório e 2.28 para as rochas adjacentes ao reservatório.

O processo de assimilação descrito na equação 4.3 não garante que, tanto as variáveis, como os parâmetros do vetor de estados assimilado $S_k^a(i)$ respeitem restrições ou satisfaçam as leis de balanço. Para corrigir isto, no que diz respeito a não satisfação das equações de balanço por parte das variáveis de estado assimiladas, foi implementada uma estratégia adicional, denominada de consistência, em razão dela objetivar a satisfação que potencialmente foi quebrada [12]. Essa estratégia foi adota pela sua efetividade, uma vez que a satisfação das equações de balanço é recuperada. Desta forma, o *EnKF* implementado pasa a ter a forma denominada na literatura como *EnKF* iterativo [90],[91].

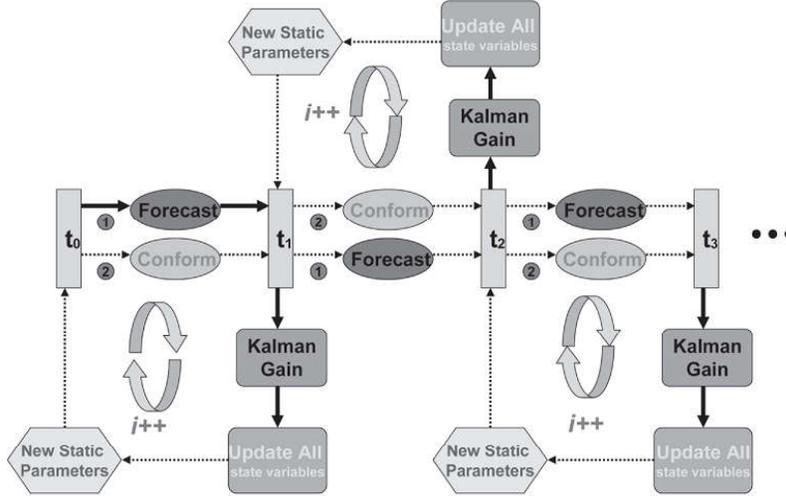


Figura 4.1: Representação esquemática do *EnKF* iterativo [91].

A figura 4.1 apresenta um esquema que permite visualizar o passo de consistência no algoritmo do *EnKF* iterativo. A consistência utiliza os parâmetros assimilados no final do intervalo $[t_i, t_{i+1}]$ para resolver, de novo, as equações do modelo, tendo como condições iniciais as variáveis de estado obtidas no passo anterior, correspondentes ao instante de tempo t_i . Portanto, nesse intervalo o sistema de equações do modelo é resolvido duas vezes, implicando, assim, em um aumento do custo computacional quando comparado com o do filtro *EnKF* convencional. Por outro lado, alguma economia computacional é conseguida, já que na equação de assimilação 4.3 apenas as operações algébricas referentes aos parâmetros são necessárias.

É relevante impor restrições aos parâmetros na execução da assimilação dos dados. A porosidade e as saturações devem ter seus valores limitados entre 0 e 1. Não respeitar estas restrições pode levar a que a solução das equações 2.27 e 2.28 não tenha sentido físico. O uso do logaritmo da permeabilidade $\ln(K_k^f)$ no vetor de estados (equação 4.10) preserva a restrição da permeabilidade ser sempre positiva. O passo de consistência implica o recálculo das variáveis do problema a partir dos campos paramétricos assimilados em t_i , partindo do estado calculado para as variáveis em t_{i-1} , com o objetivo de otimizar o custo computacional do filtro, na fase de assimilação de dados, unicamente são ajustados os parâmetros de cada um dos modelos, isto porque as grandezas das variáveis de estados ajustadas é desconsiderada no início do passo de consistência. Os valores das variáveis de estado obtidos no passo de consistência serão considerados como assimilados preservando a física do problema.

4.5 Aproximação do Erro de Modelagem

Na equação 4.5 o termo $e_{1,r}$ representa o erro de modelagem. Tipicamente, este erro é desconhecido e é influenciado por erros na parametrização dos problemas, incertezas nas condições de contorno e iniciais que definem o problema, simplificações na formulação e, no caso de reservatórios, o surgimento de processos físicos que não foram considerados na construção do modelo. Comumente a literatura de assimilação de dados em reservatórios de petróleo desconsidera esta fonte de erro devido a dificuldade em sua estimação e a mistura de seus efeitos com os efeitos originados por outras fontes de incertezas.

Os trabalhos [48, 84, 86] apresentam uma metodologia empregada na aproximação do erro de modelagem em modelos de hidrologia, tomografia e identificação de parâmetros térmicos. Esta metodologia estabelece que ainda quando o erro de modelagem é desconhecido, é possível tentar estimar sua estatística de segunda ordem e emprega-la na construção de modelos de ruído. Isto assume que o erro de modelagem pode ser descrito empregando distribuições *Gaussianas*. Estas propriedades estatísticas podem ser obtidas através de simulações de *Monte Carlo*.

Nesta abordagem a informação a priori é empregada na construção de modelos que contêm as incertezas primárias. A estatística do erro de modelagem são aproximadas sobre da distribuição a priori das incertezas primarias [55]. O cálculo da diferença entre as respostas obtidas com modelos mais complexos e modelos menos complexos produzem um conjunto de amostras de erro de modelagem. A distribuição deste erro depende explicitamente da capacidade das formulações comparadas, permitindo que as incertezas induzidas pelo uso de modelos computacionalmente menos custosos possam ser aproximadas e avaliadas no erro de modelagem:

$$S_r(i+1) = \mathcal{F}^2[S_r(i), \mu_r^*] + \tilde{e}_{1,r}(i+1), \quad (4.11)$$

$$S_r(i+1) = \mathcal{F}^1[S_r(i), \mu_r^*] + e_{1,r}(i+1). \quad (4.12)$$

Nas equações 4.11 e 4.12, \mathcal{F}^2 e \mathcal{F}^1 representam o modelo computacional de duas vias e o modelo computacional de uma via, respectivamente, $S_r(i)$ é o vetor de estados da realização r no passo de tempo i , μ_r^* são os parâmetros incertos gerados da informação a priori empregada na construção do conjunto de modelos, $\tilde{e}_{1,r}$ é o ruído que considera todas as incertezas no modelo computacionalmente eficiente e $e_{1,r}$ é o ruído que representa erro de modelagem que possa ser induzido pela formulação mais precisa.

$$\varepsilon_r(i+1) = \mathcal{F}^2[S_r(i), \mu_r^*] - \mathcal{F}^1[S_r(i), \mu_r^*], \quad (4.13)$$

$$S_r(i+1) = \mathcal{F}^1[S_r(i), \mu_r^*] + \varepsilon_r(i+1) + \tilde{e}_{1,r}(i+1). \quad (4.14)$$

Nas equações 4.13 e 4.14 ε_r é a amostrar r do erro de modelagem aproximado. Estas amostras são usadas para construir a distribuição do erro de modelo ε_m . O erro de modelo é induzido pelo uso de formulações simplificadas ou mais eficientes (do ponto de vista computacional) na modelagem do problema. O erro de modelagem aproximado descreve a discrepância entre o modelo preciso e o modelo reduzido, onde, o modelo reduzido pode ser uma aproximação do modelo físico preciso, ou, um modelo com uma discretização mais grosseira, ou com um domínio computacional menor [84]. Desta forma, a estimação de ε_m vai permitir considerar os efeitos na resposta dos modelos do uso de uma abordagem computacionalmente menos custosa. Na equação 4.14 fica evidente que é preciso ter a distribuição explícita de μ_r e $\tilde{e}_{1,r}$ para conseguir calcular ε_r e S_r . A seguir é apresentado o algoritmo proposto por [86] para a estimação do erro de modelagem aproximado.

Algoritmo 2: APROXIMAÇÃO DO ERRO DE MODELAGEM

Data: $S_j(0), \mu_j^*, N_M, N_e$

Result: ε_m

```

1 begin
2   for  $i = 1 \rightarrow T$  do
3     for  $j = 1 \rightarrow N_M$  do
4       Computar  $\mathcal{F}^2[S_j(i), \mu_j^*]$  e  $\mathcal{F}^1[S_j(i), \mu_j^*]$ 
5       Calcular  $\varepsilon_j(i+1) = \mathcal{F}^2[S_j(i), \mu_j^*] - \mathcal{F}^1[S_j(i), \mu_j^*]$ 
6     end
7     Calcular  $E(\varepsilon_j(i+1))$  e  $cov(\varepsilon_j(i+1), \varepsilon_j(i+1))$ 
8   end
9   for  $i = 1 \rightarrow T$  do
10    for  $r = 1 \rightarrow N_e$  do
11      Gerar a amostra  $e_{1,r}(i)$  da variável aleatória  $\varepsilon_m(i)$ 
12    end
13  end
14 end
15 return  $\varepsilon_m \in \mathbb{R}^{N_e \times T}$ 

```

No algoritmo 2 N_M é a quantidade de modelos empregados na construção das distribuições do erro de modelagem, N_e é a quantidade de modelos que conformam o conjunto de modelos empregados na assimilação de dados com o *EnKF*, T é o número de passos que serão executados na assimilação de dados e $E(\varepsilon_j(i+1))$ e

$cov(\varepsilon_j(i+1), \varepsilon_j(i+1))$ são a média e a covariância da distribuição ε_m para o passo de tempo i .

O método de *Monte Carlo (MC)* é empregado no computo das respostas dos modelos \mathcal{F}^2 e \mathcal{F}^1 , e assim, partindo das amostras $\varepsilon_j(i+1)$ são estimadas a média e a covariância do erro de modelagem aproximado ε_m para cada passo de tempo. Desta forma realizações $e_{1,r}(i)$ da distribuição $\varepsilon_m(i)$ serão geradas para seu computo na assimilação de dados empregando o *EnKF*. A metodologia de aproximação do erro de modelagem, aplicada à assimilação de dados, empregando o *EnKF* em modelos hidrológicos é apresentada em [86]. Destaca-se que as amostras de $e_{1,r}$ não dependem das medições, assim, todos os cálculos relacionados à estatística do erro de modelagem podem ser realizados previamente ao processo de assimilação de dados.

4.6 Proposta para Aprimorar o *EnKF*

O uso de modelos computacionais de alta fidelidade para a previsão é considerado importante para garantir estimativas precisas a partir dos modelos assimilados empregando o *EnKF*. No entanto, esse procedimento pode ser computacionalmente caro, pois é preciso resolver os problemas diretos de um grande número de amostras, necessárias para alcançar a convergência estatística na assimilação. Modelos de baixa fidelidade são frequentemente empregados, visando alcançar um equilíbrio justo entre custo computacional e a precisão. Para melhorar a capacidade do *EnKF* em produzir resultados confiáveis é preciso ter uma abordagem sistemática para lidar com a discrepância entre os resultados de alta e baixa fidelidade.

Com este objetivo é construída uma formulação de uma via aprimorada, explorando o erro de modelo calculado a partir das discrepâncias entre os dois modelos. Para realizar essa tarefa é aproveitada a natureza estocástica do *EnKF* combinada com a incerteza engendrada através do erro do modelo ε_m , considerada como uma variável randômica. Para isto, são geradas amostras das distribuições a priori da permeabilidade e do módulo de Young, usadas no cálculo da resposta tanto do modelo de 2-vias quanto do modelo de 1-via o que permite estimar ε_m através do método de Monte Carlo.

Desta forma, para produzir estimativas do vetor de variáveis a cada passo de tempo são empregadas N_m amostras a priori do vetor de parâmetros $\tau_j = \begin{bmatrix} K_j & \mathbb{E}_j \end{bmatrix}^T$ ($j = 1, \dots, N_m$) empregando os dois modelos:

$$Y_j^2(i+1) = \mathcal{F}^2(Y_j^2(i); \tau_j), \quad (4.15)$$

$$Y_j^1(i+1) = \mathcal{F}^1(Y_j^2(i); \tau_j). \quad (4.16)$$

nas equações anteriores \mathcal{F}^2 e \mathcal{F}^1 representam as formulações de duas vias e uma via, respectivamente. Y_j^2 e Y_j^1 são os vetores de estados correspondentes a cada uma das formulações. Destaca-se que as j amostras do vetor de estados $Y_j^2(i+1)$ e $Y_j^1(i+1)$ são ambas computadas partindo do último estado calculado pelo modelo de 2-vias $Y_j^2(i)$ no passo de tempo i . O propósito disto é eliminar a acumulação de diferença entre os dois modelos gerada em passos de tempo prévios. Assim amostras do erro de modelo são obtidas

$$\varepsilon_j(i) = Y_j^2(i+1) - Y_j^1(i+1) \text{ with } j = 1, \dots, N_m. \quad (4.17)$$

A discrepância entre os dois modelos foi calculada empregando a mesma discretização espaço-temporal para os dois modelos. Igualmente, procurando simplificar os cálculos, foi assumido que ε_m é um processo Gaussiano que pode ser caracterizado pela média e a covariância. Para cada passo de tempo $i+1$ são computadas as estatísticas de primeira ordem:

$$\mu_m(i+1) = \frac{1}{N_m} \sum_{j=1}^{N_m} \varepsilon_j(i+1) \quad (4.18)$$

$$C_m(i+1) = \frac{1}{N_m - 1} \sum_{i=1}^{N_m} [\varepsilon_j(i+1) - \mu_m(i+1)] [\varepsilon_j(i+1) - \mu_m(i+1)]^T \quad (4.19)$$

O erro de modelo ε_m estimado em (4.17) representa uma marginalização dos parâmetros incertos, não consistindo em uma abordagem totalmente bayesiana. Na literatura, abordagens mais sofisticadas para a estimação da distribuição do erro de modelo são propostos [57, 89] levando a distribuições condicionais para a média e a covariância obtidas com maiores custos computacionais.

Dentro do contexto atual, amostras da distribuição ε_m são computadas na assimilação de dados empregando o *EnKF*. Observe-se que as amostras de ε_m são independentes das medições, e portanto, os cálculos relacionados às estatísticas de erro de modelagem podem ser realizados antes do processo de assimilação de dados. Tal procedimento emprega todo o vetor de estados que não necessariamente é monitorado, o que representa um ganho do ponto de vista da flexibilidade de implementação e eficiência computacional. O processo completo de assimilação é apresentado esquematicamente no algoritmo 3.

A fim de estabelecer um equilíbrio entre a consistência física e os custos computacionais do processo de assimilação, no algoritmo 3 é implementada uma variante do processo de assimilação iterativo. Nesta abordagem, depois de atualizar os parâmetros no tempo $i+1$, é executado novamente o modelo do problema direto

Algoritmo 3: ALGORITMO DO PROCESSO DE ASSIMILAÇÃO DE DADOS

Data: $n = 0$, $p_o(\mathbf{x})$, $\phi_o(\mathbf{x}, 0)$, N_e amostras de K e \mathbb{E}

```
1 begin
2   for  $i = 1 \rightarrow T$  do
3     Loop de evolução temporal
4     Atualizar  $n = n + 1$ 
5     for  $c = 1 \rightarrow 2$  do
6       Loop de consistência
7       for  $j = 1 \rightarrow N_e$  do
8         Loop de construção do Ensemble
9         Inicializar  $k = 0$ , valores iniciais  $\tilde{\sigma}_{T,j}^{k=0}(\mathbf{x}, t^n) = \tilde{\sigma}_{T,j}(\mathbf{x}, t^{n-1})$ 
10        for  $k$  do
11          Atualizar  $k = k + 1$ 
12          if  $k = 1$  then
13             $\tilde{\sigma}_{T,j}^{k-1}(\mathbf{x}, t^n) = \tilde{\sigma}_{T,j}(\mathbf{x}, t^{n-1})$ 
14          end
15          Dado  $\tilde{\sigma}_{Tn,j}^k$  e  $\tilde{\sigma}_{Tn,j}^{k-1}$ , computar  $\mathbf{v}_{D,j}^k(\mathbf{x}, t^n)$  e  $p_j^k(\mathbf{x}, t^n)$ 
16          Dado  $p_j^k(\mathbf{x}, t^n)$  calcular  $\mathbf{u}_j^k(\mathbf{x}, t^n)$  e  $\sigma_{s,j}(\mathbf{x}, t^n)$ 
17          Dado  $p_j^k(\mathbf{x}, t^n)$  e  $\sigma_{s,j}(\mathbf{x}, t^n)$  calcular  $\tilde{\sigma}_{T,j}^k(\mathbf{x}, t^n)$ 
18          until  $\|\sigma_{Tn,j}^k(x) - \sigma_{Tn,j}^o(x)\| \leq \text{tolerância}$ 
19        end
20        com  $\phi_j(\mathbf{x}, t^{n-1})$ ,  $p_j(\mathbf{x}, t^n)$  e  $\mathbf{u}_j(\mathbf{x}, t^n)$  calcular  $\phi_j(\mathbf{x}, t^n)$ 
21        Construir o vetor de estados estimados  $S_j^f$  contendo  $p_j(\mathbf{x}, t^n)$ ,
           $\mathbf{v}_{D,j}(\mathbf{x}, t^n)$ ,  $\mathbf{u}_j(\mathbf{x}, t^n)$ ,  $K_j(\mathbf{x})$  e  $\mathbb{E}_j(\mathbf{x})$ 
22      end
23      if  $c = 1$  then
24        Calcular a matriz de covariância  $P^f(t_n)$ 
25        if Aproximação do erro de modelo está ativo then
26          incluir  $\epsilon_j$  amostras do erro de modelo no cálculo de  $G(t_n)$ 
27        end
28        Calcular os ganhos do filtro de Kalman  $G(t_n)$ 
29        Calcular os parâmetros assimilados  $K_j^a(\mathbf{x})$  e  $\mathbb{E}_j^a(\mathbf{x})$ 
30        atualizar  $K_j(\mathbf{x}) = K_j^a(\mathbf{x})$  e  $\mathbb{E}_j(\mathbf{x}) = \mathbb{E}_j^a(\mathbf{x})$ 
31      end
32      if  $c = 2$  then
33        atualizar  $p_j^a(\mathbf{x}, t^n) = p_j(\mathbf{x}, t^n)$ ,  $\mathbf{v}_{D,j}^a(\mathbf{x}, t^n) = \mathbf{v}_{D,j}(\mathbf{x}, t^n)$ ,
           $\mathbf{u}_j^a(\mathbf{x}, t^n) = \mathbf{u}_j(\mathbf{x}, t^n)$ ,  $\phi_j^a(\mathbf{x}, t^n) = \phi_j(\mathbf{x}, t^n)$ 
34      end
35    end
36  until  $n = N$ 
37 end
38 end
```

com a condição inicial dada pelo estado no tempo i obtendo um novo estado em $i + 1$ com a diferença que nesta segunda rodada são empregados os campos assimilados de K e \mathbb{E} . Para simplificar a terminologia, tal procedimento em duas etapas é chamado de etapa de consistência (seção 4.4) e está representado no algoritmo 3 no loop c entre as linhas 5 e 35.

Além disso, para evitar instabilidades numéricas no cálculo da inversão $[\mathbf{HP}^f(i)\mathbf{H}^T + \mathbf{R}(i)]^{-1}$ em (4.4), é empregado o truncamento de valores singulares (TSVD) apresentado na seção 4.3 durante o cálculo dos ganhos do filtro de Kalman (linha 28 do algoritmo 3).

4.7 Validação do Algoritmo do *EnKF* empregando o *McMC*

Nesta seção é apresentada a formulação e resultados obtidos assimilando dados em um problema de escoamento bifásico incompressível em um meio poroso rígido. Estes testes foram empregados para validar a implementação do *EnKF* uma vez os resultados obtidos com ele são comparados com os resultados obtidos empregando o *McMC* documentado em []. Estimativas do tempo de *breakthrough* desse problema são geradas com os modelos assimilados empregando tanto o *EnKF* quanto o *McMC*.

O escoamento bifásico em um meio poroso rígido é governado pelas seguintes equações de balanço:

$$\mathbf{v}_{Dt} = -K\lambda_t(S_w)\nabla p; \quad (4.20)$$

$$\nabla \cdot \mathbf{v}_{Dt} = 0; \quad (4.21)$$

$$\phi \frac{\partial S_w}{\partial t} + \nabla \cdot F(S_w) = 0. \quad (4.22)$$

que formam um sistema não linear de equações diferenciais acopladas, onde \mathbf{v}_{Dt} é a velocidade de Darcy para mistura água-óleo, K é a permeabilidade da rocha, p é a pressão nos poros, ϕ é a porosidade, S_w a saturação da água, F é a função fracionária de fluxo e λ_t a mobilidade da mistura água-óleo. A mobilidade e a função fracionária de fluxo dependem da permeabilidade relativa de cada uma das fases no correspondente meio poroso:

$$\lambda_t(S_w) = \lambda_w + \lambda_o = \frac{k_{rw}(S_w)}{\mu_w} + \frac{k_{ro}(S_w)}{\mu_o} \quad (4.23)$$

$$F = f_w \mathbf{v}_{Dt} = \frac{\lambda_w}{\lambda_t} \mathbf{v}_{Dt} = \frac{k_{rw}}{\mu_w \lambda_t} \mathbf{v}_{Dt} \quad (4.24)$$

Nas equações (4.23) e (4.24) k_{rw} e k_{ro} são as permeabilidades relativas da água e do óleo, respectivamente, μ_w e μ_o são as viscosidades da água e o óleo, λ_w e λ_o a mobilidade da fase água e da fase óleo e f_w a função fracionária de fluxo da fase água. As permeabilidades relativas quantificam a mobilidade de uma determinada fase na mistura em função da saturação da fase molhante. Nesta formulação emprega-se o modelo de Corey para descrever as relações entre as permeabilidades relativas e a saturação da fase água. A seguir é apresentada a versão do modelo de Corey empregado em estes testes:

$$k_{rw}(S_w) = \frac{(S_w - S_{rw})^2}{(1 - S_{rw})^2} \quad (4.25)$$

$$k_{ro}(S_o) = \left[1 - \frac{S_w}{1 - S_{ro}} \right]^2 \quad (4.26)$$

Note-se que nas equações (4.25) e (4.26) estão formuladas utilizando-se dois parâmetros: as saturações residuais da água e do óleo S_{rw} e S_{ro} . Esses parâmetros representam as saturações mínimas que ficarão de cada fase no reservatório. Esse tipo de equação de fechamento que dita o comportamento de um processo físico, como já mencionado anteriormente, pode introduzir tanto incertezas paramétricas, como os valores das saturações mínimas, como erros de modelos (conhecidos também como incertezas de estruturas).

Estimação do Tempo de Breakthrough Usando *McMC* e *EnKF*

Na execução dos testes, o campo de permeabilidades heterogêneo é modelado como um campo log-normal (equação (4.27)) onde Y é um campo Gaussiano com média zero e função de covariância exponencial (equação (4.28)),

$$K(\mathbf{x}) = e^{Y(\mathbf{x})}, \quad (4.27)$$

$$C_Y(\mathbf{x}, \mathbf{y}) = \sigma^2 e^{\left(-\frac{|x_1 - y_1|}{2L_1} - \frac{|x_2 - y_2|}{2L_2} \right)}. \quad (4.28)$$

A Figura 4.2 mostra o campo log-permeabilidade, gerado a partir do modelo apresentado em (4.27) e (4.28), usado como campo de referência para gerar os dados sintéticos empregados nos processos de assimilação. Os parâmetros apresentados na tabela 4.7 são usados em todas as simulações. Cada um dos métodos usou um conjunto de 200 amostras na obtenção dos resultados apresentados.

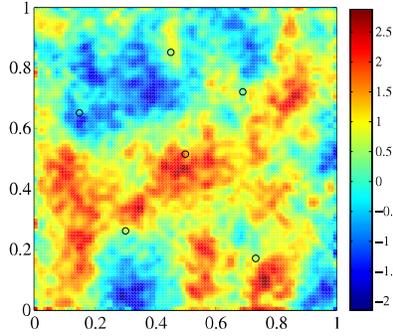


Figura 4.2: Campo log-permeabilidade usado como referência no problema de estimação do tempo de *breakthrough*. O símbolos “o” indicam as posições nas quais a pressão é monitorada

PARÂMETRO	VALOR
Taxa de injeção	0,1 m^3/s
Viscosidade da água	$1 \times 10^{-3} Ns/m^2$
Viscosidade do óleo	$20 \times 10^{-3} Ns/m^2$
Tempo de assimilação	0,15 <i>dia</i>
Malha computacional	100 x 100 elementos
Malha geológica	100 x 100 elementos
Comprimentos de correlação (L_1 e L_2)	0,2 <i>m</i>
Quantidade de realizações	200

Nos experimentos, dados de pressão obtidos em alguns pontos de monitoração foram utilizados nos métodos *EnKF* e *McMC* visando melhorar as previsões para o tempo de *breakthrough*. Duas configurações de medições são usadas: na primeira é usada uma única medição da pressão no poço injetor do reservatório, na outra configuração são usadas 6 estações de monitoração da pressão no reservatório (Figura 4.2) como fonte de dados na assimilação. Os objetivos principais foram: (i) evidenciar o impacto da quantidade de informação na qualidade das previsões geradas com o *EnKF* e; (ii) comparar a qualidade das previsões obtidas pelos métodos de Monte Carlo clássico (*MC*), *McMC* e *EnKF*.

Na Figura 4.3 são apresentadas a evolução temporal média da pressão nos pontos de monitoração obtidas para os conjuntos de campos assimilados no método *EnKF* (com 1 e 6 medições, Figuras 4.3(a) e 4.3(b), respectivamente), campos selecionados no método *McMC* (com 6 monitores, Figura 4.3(c)) e campos do método de *MC* clássico (Figura 4.3(d)). Nestas figuras, a linha contínua representa o valor da pressão obtida para o campo de referência enquanto os círculos e as barras representam a média e o desvio padrão, respectivamente. A linha segmentada vertical

em $t = 0.15$ indica o tempo até o qual foi executado o processo de assimilação. Após este ponto são apresentadas as previsões para a pressão nos pontos de monitoração obtidas para cada experimento. Observa-se claramente que os métodos *EnKF* e *McMC* foram capazes de reduzir drasticamente as incertezas associadas ao problema e produziram previsões para a pressão muito melhores quando comparados ao método de *MC*.

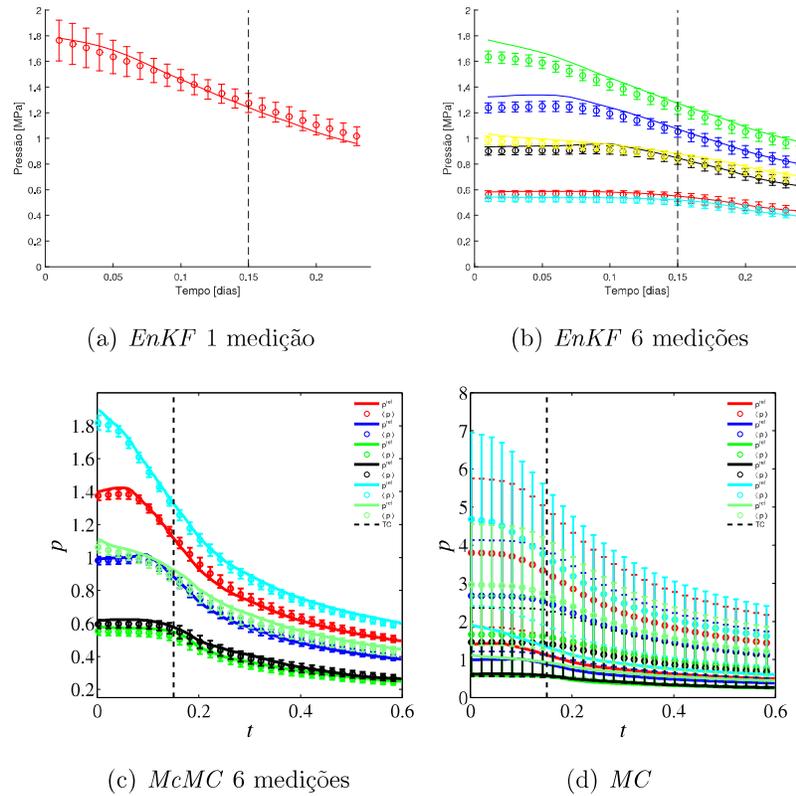


Figura 4.3: Comparação entre a pressão de referência (linha contínua) e a pressão média (círculos) nos pontos de monitoração, obtidas com os métodos *EnKF*, *McMC* e *MC*. Barras indicam o desvio padrão. Linha segmentada vertical indica o tempo utilizado para a caracterização

A Figura 4.4 mostra, com maior detalhe, o resultado da previsão nos pontos de monitoração. Destaca-se que ambos os métodos produziram boas previsões para a pressão nesses pontos em todos os casos.

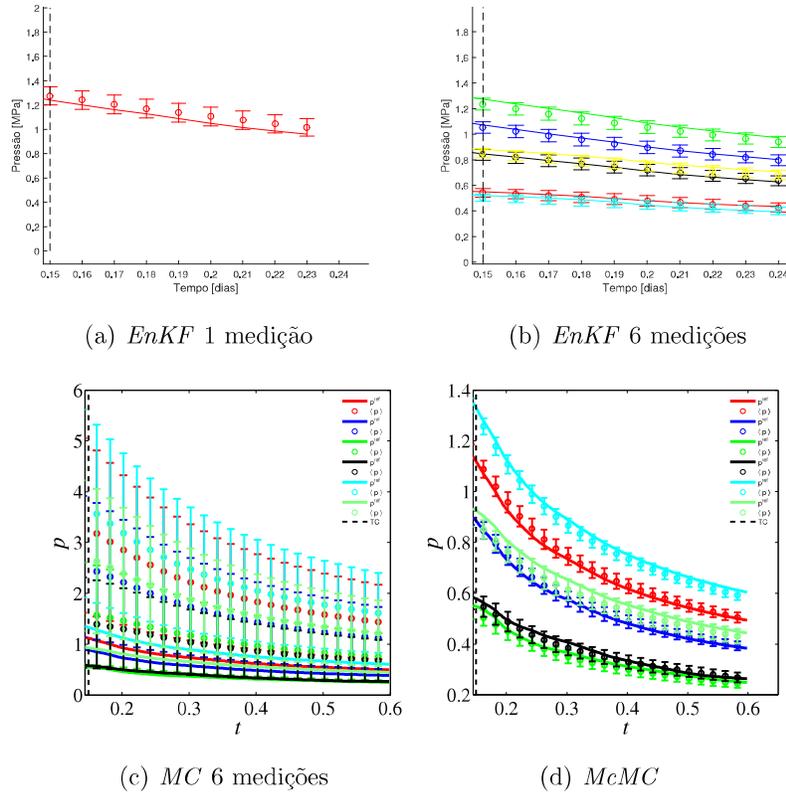
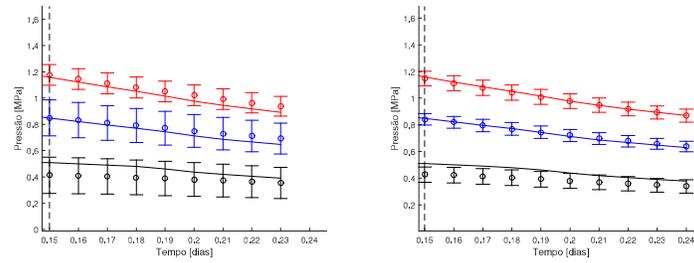


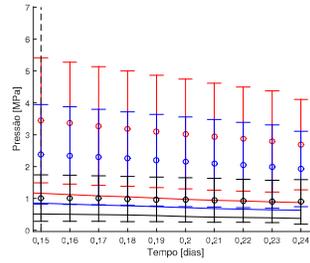
Figura 4.4: Comparação entre as previsões para a pressão de referência (linha contínua) obtidas com os métodos *EnKF* e *McMC*. Os círculos representam as pressões médias nos pontos de monitoração obtidas por cada método e as barras indicam o desvio padrão

Por se tratar de um caso sintético, previsões para a pressão foram realizadas em pontos do reservatório que não foram monitorados, no caso do *EnKF*, permitindo verificar a capacidade da técnica de extrapolar a informação obtida em pontos específicos ao domínio todo do reservatório. Portanto, foram selecionados três pontos, um perto do lado do reservatório onde é injetada a água, outro na metade do domínio do reservatório e um último ponto próximo ao poço produtor. Os resultados, apresentados na Figura 4.5, mostram que o método *EnKF* melhorou significativamente as previsões para a pressão, quando comparado ao método de *MC*, mesmo em locais nos quais os dados não foram utilizados na etapa de assimilação. Isto deixa em evidência o impacto que tem a assimilação de dados no processo de previsão. Um aumento no número de medições de 1 (Figura 4.5(a)) para 6 (Figura 4.5(b)) proporcionou uma maior redução no grau de incerteza representada pelo desvio padrão.



(a) *EnKF* 1 medição

(b) *EnKF* 6 medições



(c) MC

Figura 4.5: Comparação entre as previsões para a pressão do campo de referência (linha contínua) e a pressão média (círculos) em três pontos do reservatório, obtidas com os métodos *EnKF* e *MC*. Barras indicam o desvio padrão. Linha segmentada vertical indica o tempo utilizado para a caracterização

Os campos log-permeabilidades médio obtidos em cada um dos experimentos e suas variâncias são apresentados nas Figuras 4.6 e 4.7, respectivamente. Verifica-se que os campos médios obtidos em todas as situações apresentam uma região de maior permeabilidade (representada pela cor vermelha) que guarda semelhança visual com o campo de referência (Figura 4.6(a)). Além disso, verifica-se uma grande diminuição na variância dos campos obtidos pelas duas técnicas (Figura 4.7), lembrando que a distribuição *a priori* foi gerada com variância igual a 1.

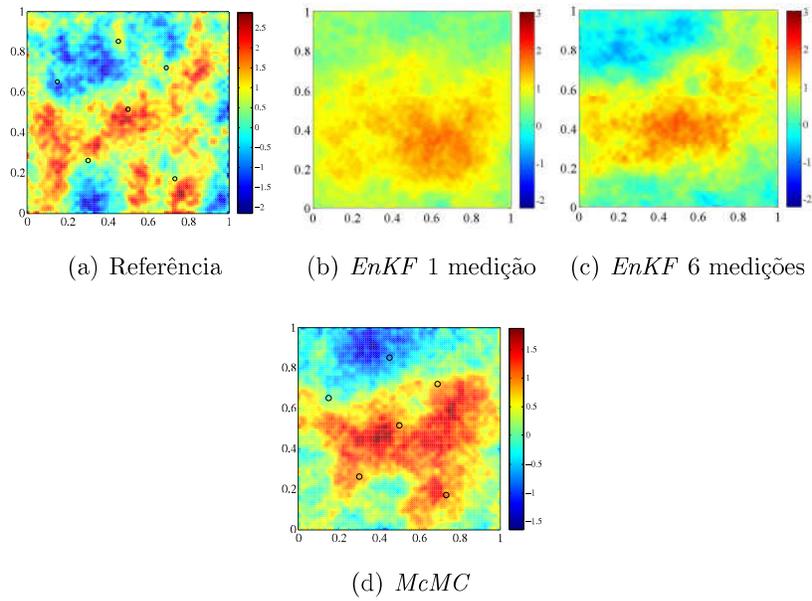


Figura 4.6: Campo de referência e campos log-permeabilidade médios obtidos pelos métodos *EnKF* e *MCMC*

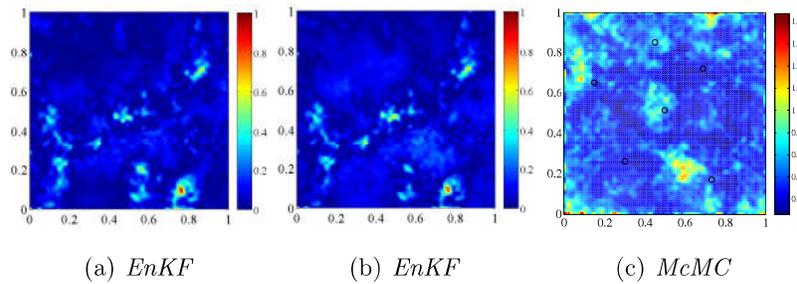


Figura 4.7: Variância dos campos log-permeabilidade obtidos pelos métodos *EnKF* e *MCMC*

Na etapa de caracterização foram utilizados somente dados de pressão nos pontos monitores, entretanto, previsões para esta variável não tem interesse prático, ao menos neste caso em que é considerado o meio poroso rígido. Maior interesse reside na obtenção de boas previsões para o tempo de *breakthrough*, por exemplo.

Com o objetivo de garantir a convergência do método, além de estabelecer um balanço entre custo computacional e qualidade de resultados foram avaliadas diversos tamanhos do conjunto de amostras empregadas na assimilação de dados com o *EnKF*, na Figura 4.8 são apresentadas as funções de densidade de probabilidade da estimação do tempo de *breakthrough* obtidas empregando conjuntos de 200, 400 e 600 amostras.

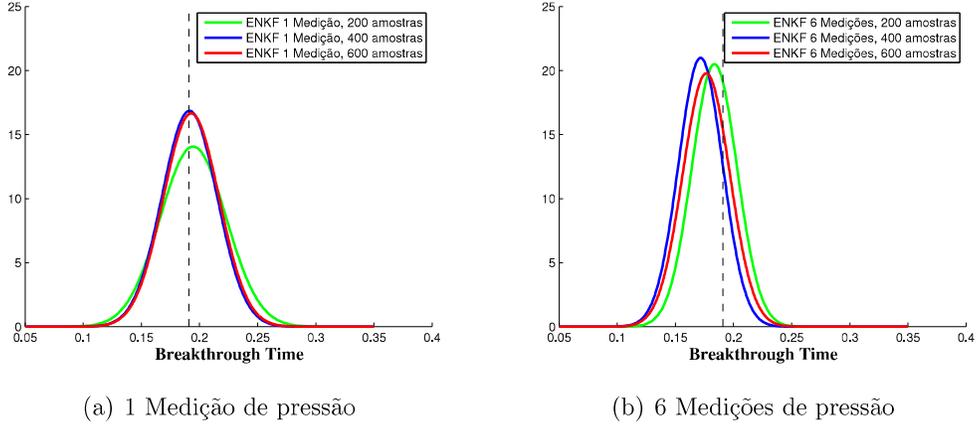


Figura 4.8: Função de densidade de probabilidade do tempo de *breakthrough* previsto. A linha a segmentada vertical representa o tempo de *breakthrough* do modelo de referência.

Na figura 4.8 a linha verde apresenta os resultados obtidos com um conjunto de 200 amostras, a linha azul com um conjunto de 400 amostras e a linha verde com um conjunto de 600 amostras. O aumento de amostras para os casos em que somente foi empregada uma medição de pressão na assimilação, figura da esquerda, mostrou-se eficaz na redução da variância das previsões, conservando uma média cercana ao tempo de *breakthrough* do modelo de referência.

Para o caso em que foram empregadas 6 estações de monitoração de pressão na assimilação dos dados, figura da direita, a variância foi um pouco reduzida, mas, as pdfs se deslocaram levemente à esquerda mostrando uma tendencia a ter um valor médio menor do que o tempo de *breakthrough* do modelo de referência. Dos resultados anteriores se pode apreciar claramente como a quantidade de amostras impacta diretamente no ajuste do intervalo de incertezas das previsões; mas não necessariamente, o valor médio ou a mediana coincidem com o valor de referência. Em problemas reais, dito valor é desconhecido e torna-se relevante ter uma estimação do intervalo de incertezas confiável.

A Figura 4.9 mostra as previsões para o tempo de *breakthrough* obtidas pelos métodos *EnKF* (com 1 e 6 pontos de monitoração), *McMC* (com 6 pontos de monitoração) e *MC* na forma de gráficos de caixa. A linha vermelha horizontal indica o tempo de *breakthrough* do campo de referência. Nestes gráficos, verifica-se que ambos os métodos utilizados para a redução de incertezas produziram previsões superiores ao método de *MC* tanto na proximidade das medianas com o valor referência, bem como na menor variabilidade dos dados. Comparando-se os experimentos com *EnKF* com 1 e 6 medições percebe-se ainda que uma maior quantidade de informação produziu uma menor variabilidade na previsão, embora a mediana tenha se afastado do valor de referência quando comparado ao caso com apenas 1 monitor (Figura 4.9(a)).

Nos gráficos da Figura 4.9 os conjuntos de campos dos experimentos do método *MC* são diferentes, entretanto, ambos produziram resultados bastante próximos, como esperado.

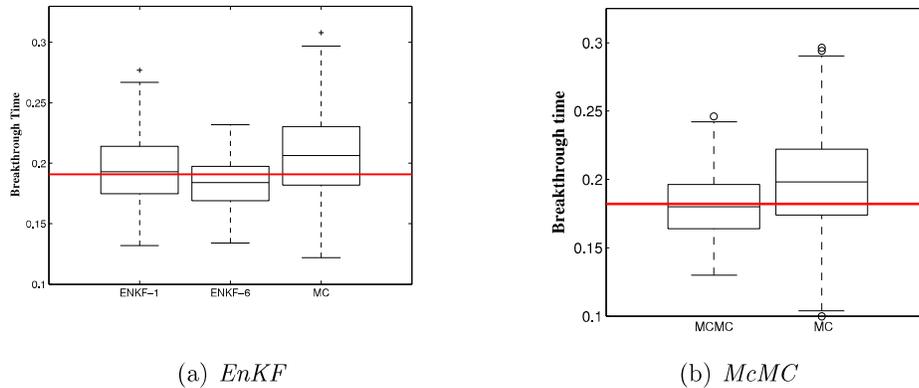


Figura 4.9: Gráficos de caixa dos tempos de *breakthrough* estimados. A linha vermelha representa o tempo de *breakthrough* do modelo de referência

Na Figura 4.10 são apresentadas as funções de densidade de probabilidade (pdf) calculadas para os tempos de *breakthrough* obtidos os métodos *EnKF* e *MC*. Verifica-se que os dados obtidos para os tempos de *breakthrough* tiveram distribuição aproximadamente normal. A linha verde representa a pdf obtida assimilando os dados com o *EnKF* e 1 medição de pressão, a linha azul é a pdf obtida com o *EnKF* e 6 medições de pressão e a linha vermelha é a pdf obtida empregando o *MC*. O tempo de *breakthrough* do modelo de referência é representado nessa figura com a linha segmentada vertical.

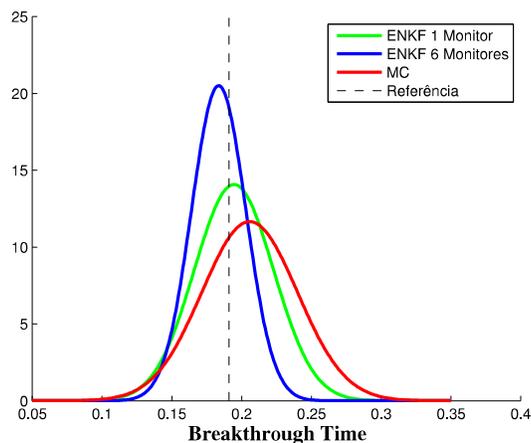


Figura 4.10: Função de densidade de probabilidade do tempo de *breakthrough* previsto. A linha a segmentada vertical representa o tempo de *breakthrough* do modelo de referência

Na imagem da esquerda da figura 4.11 é apresentada a função de densidade de probabilidade estimada para o tempo de *breakthrough* empregando o *McMC*. Comparando estes resultados com os obtidos pelo *EnKF* iterativo (apresentados na figura 4.10) observa-se que os dois métodos estimaram distribuições semelhantes mantendo uma variância e media próximas para o tempo de *breakthrough*. Este resultado permitiu validar a implementação do *Ensemble Kalman Filter (EnKF)* com ciclo de consistência empregado em esta pesquisa uma vez é aceito que o *McMC* é a técnica mais robusta para conseguir recuperar a distribuição das incertezas.

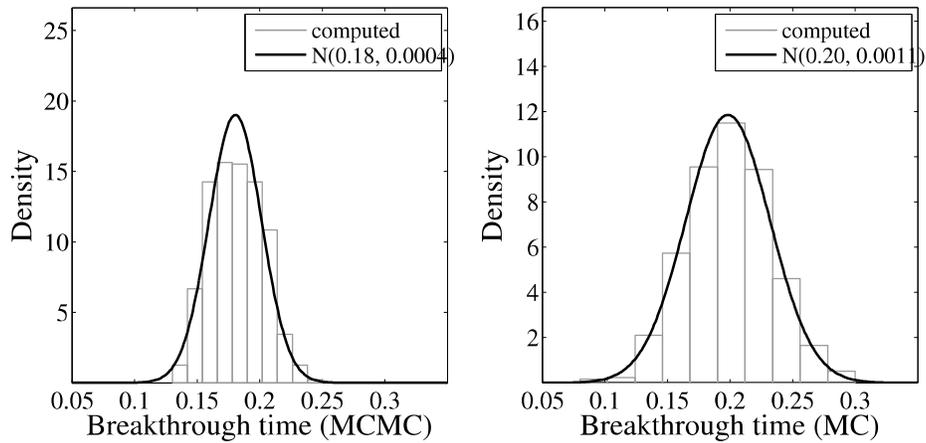


Figura 4.11: Função de densidade de probabilidade do tempo de *breakthrough* previsto. A linha a segmentada vertical representa o tempo de *breakthrough* do modelo de referência

Capítulo 5

Resultados e Discussões

Neste capítulo, são apresentados testes numéricos que permitem comparar o desempenho da assimilação de dados empregando a formulação de uma via aprimorada com o desempenho da assimilação de dados empregando as formulações de duas vias e uma via tradicional. Concentramo-nos na eficácia da introdução de erros de modelo para avaliar sua real capacidade de diminuir os impactos negativos de erros introduzidos pelo uso de formulações simplificadas em problemas de escoamento em reservatórios de petróleo considerando os efeitos geomecânicos. Os resultados obtidos com o modelo de duas vias assim como o modelo de referência, empregado na geração sintética dos dados, são considerados como referência de desempenho, desta forma a qualidade das calibrações e previsões dos modelos de uma via tradicional e aprimorado são medidas em função do modelo de duas vias e o modelo de referência.

5.1 Problema Protótipo

Na Figura 5.1, apresentamos uma visão esquemática da formação geológica considerada nos testes numéricos. O domínio computacional é decomposto em sub-domínios, cada um, com diferentes configurações geológicas e propriedades mecânicas. O foco das análises reside no reservatório formado por uma rocha carbonática que ocupa o domínio Ω_r , cercado por rochas impermeáveis. Os sub-domínios são representados na figura como Ω_2 , Ω_3 , Ω_4 e Ω_5 . Neste cenário idealizado, a operação do reservatório é feito por dois poços verticais de injeção e produção (Figura 5.1). As formações rochosas adjacentes são consideradas homogêneas e elásticas, e são as responsáveis pela imposição e transmissão de cargas mecânicas sobre o reservatório. A superfície superior Ω_1 é constituída de rocha pós-sal, e a subjacente, Ω_3 , é uma camada rígida que atua limitando o deslocamento do reservatório de formação.

O sub-problema de escoamento só é resolvido em Ω_r , enquanto o da geomecânica no domínio todo ($\Omega_r \cup \Omega_1 \cup \Omega_2 \cup \Omega_3 \cup \Omega_4 \cup \Omega_5$) com o termo envolvendo o gradiente de

pressão nos poros aparecendo apenas em Ω_r . O contorno selado de Ω_r é imposto por condições de contorno de Neumann homogêneas para o fluxo. Quando o processo de injeção começa, o respectivo poço é aberto com uma velocidade de Darcy prescrita constante $\mathbf{v}_D \cdot \mathbf{n} = Q$. O contorno direito de Ω_r corresponde a um poço de produção onde as condições de contorno de Dirichlet são aplicadas para a pressão de poros $p = T_{des} * t$ onde T_{des} é uma taxa de despressurização linear e t é o tempo decorrido. Para o problema de elasticidade são postuladas as condições de contorno (equação 2.30) nos limites externos do domínio geral $\Omega = \Omega_r \cup \Omega_1 \cup \Omega_2 \cup \Omega_3 \cup \Omega_4 \cup \Omega_5$. Nas laterais de Ω_r condições de contorno mistas impõem a ausência de deslocamento horizontal e tensão de cisalhamento. O lado inferior do domínio é considerado rígido, impondo deslocamento homogêneo e nulo.

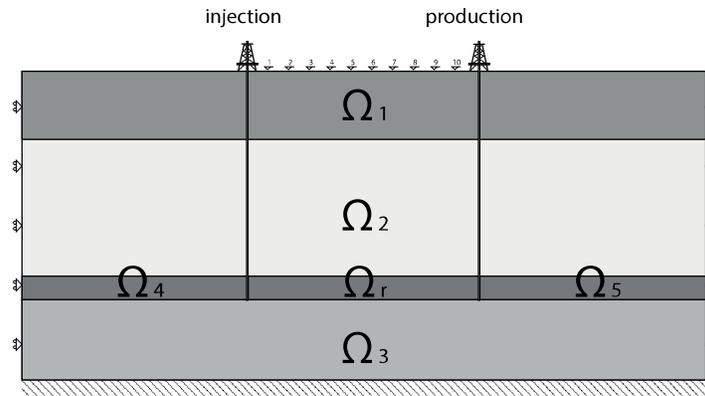
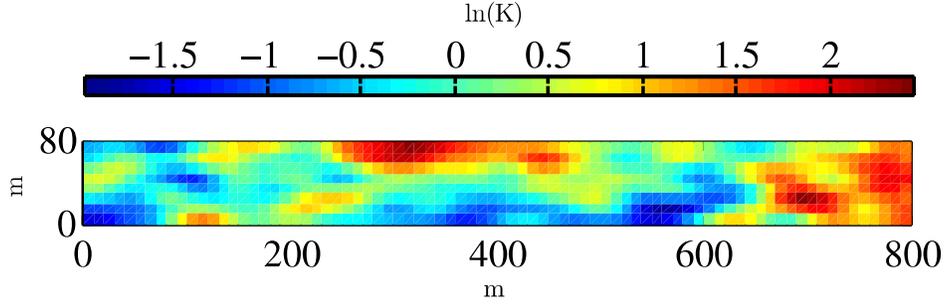


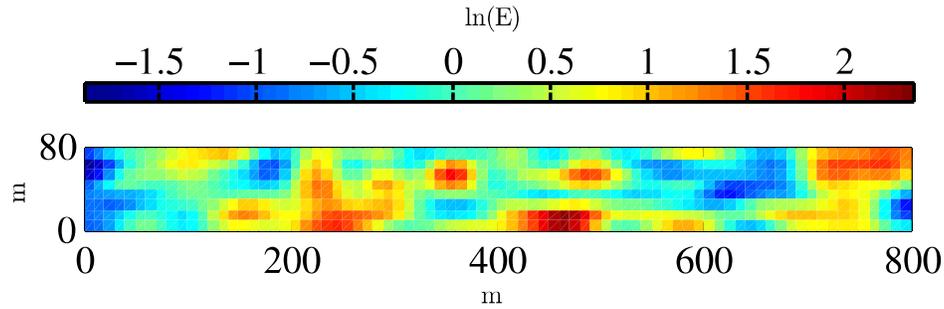
Figura 5.1: Domínio computacional: Decomposição em diferentes regiões geológicas; posição dos poços e estações de monitoração da subsidência

As dimensões do reservatório são $80m$ de altura (direção vertical) por $800m$ de comprimento (direção horizontal), enquanto as dimensões do domínio total $1080m$ de altura por $2400m$ de comprimento. A malha de elementos finitos usada na discretização de Ω_r tem 8×80 elementos quadrilaterais de $10m \times 10m$ cada. O restante do domínio é discretizado com uma malha mais grossa, usando principalmente elementos quadriláteros de $40m \times 40m$. Nas proximidades da interface entre as rochas capeadoras e o reservatório, a discretização é gradualmente refinada até elementos de $10m \times 10m$.

Na Figura 5.1, as dez posições indicadas na superfície superior correspondem à localização das estações de medição da subsidência. Estas estão posicionadas a $720m$ acima da superfície superior de Ω_r . No processo de assimilação, além das medições de subsidência, a vazão de produção de óleo e a pressão de injeção no fundo do poço são explorados como fontes de informação.



(a) Permeabilidade de referência



(b) Módulo de Young de referência

Figura 5.2: Campos espacialmente heterogêneos do modelo de referência: (a) permeabilidade e (b) módulo de Young

A heterogeneidade espacial do reservatório é caracterizada por campos estocásticos que contemplam as incertezas. Como citado anteriormente, os campos de permeabilidade K e módulo de Young \mathbb{E} são selecionados como alvos dos esquemas de assimilação de dados permitindo sua identificação e calibração a partir de medições indiretas.

Seguindo as abordagens geoestatísticas usuais, os campos estocásticos são considerados estacionários exibindo um comportamento Gaussiano. Para reforçar essas características, a permeabilidade e o módulo de Young são modelados como campos espaciais log-normalmente distribuídos. Tomando \mathbf{Z} um campo aleatório Gaussiano, estacionário, normalizado e isotrópico caracterizado pela função de covariância exponencial de dois pontos C [22, 37]:

$$K(\mathbf{x}) = \alpha_K e^{\delta_K \mathbf{Z}(\mathbf{x})} \quad (5.1)$$

$$E(\mathbf{x}) = \alpha_E e^{\delta_E \mathbf{Z}(\mathbf{x})} \quad (5.2)$$

$$C(\mathbf{x}, \mathbf{y}) = \sigma^2 e^{-\frac{|\mathbf{x}-\mathbf{y}|}{l}} \quad (5.3)$$

Nas equações (5.1) e (5.2), os hiperparâmetros α_K , δ_K , α_E e δ_E são a intensidade de heterogeneidade dos parâmetros enunciados. Além disso, σ^2 e l em (5.3) são a variância e comprimento de correlação.

Com a finalidade de avaliar se a metodologia de assimilação proposta aumenta a capacidade dos modelos computacionais do reservatório de fornecer previsões confiáveis, empregamos dados sintéticos, juntamente com a assimilação, escolhendo o modelo de duas vias como o modelo de maior fidelidade. As simulações do problema direto produzem estados transitórios que são perturbados por ruído para evitar os chamados crimes em problemas inversos [51]. Posteriormente, estes estados são usados como dados para a assimilação. A figura 5.2 apresenta uma amostra para cada campo estocástico adotado na produção dos dados sintéticos, estes campos serão referidos no decorrer do texto como campos de referência. Os valores de α_K , δ_K , α_E , δ_E , l e os valores dos principais parâmetros empregados nos testes são apresentados na Tabela 5.1. Vale ressaltar que essas amostras particulares não pertencem ao conjunto utilizado para a assimilação.

Como mencionado anteriormente, as saídas da simulação direta para obter os dados de referência são perturbadas com ruído Gaussiano com média igual a 0 e desvios padrão de $1 \times 10^{-3}m$ para subsidência, $0,1m^3/s$ para vazão de produção e $5 \times 10^5 Pa$ para pressão de injeção.

Em todas as simulações, o processo de recuperação de óleo segue duas etapas sequenciais, reproduzindo as estratégias operacionais padrão de reservatórios. A produção de petróleo no primeiro estágio é impulsionada pela diferença de pressão entre o reservatório e as condições externas. Posteriormente, no estágio de recuperação secundária inicia-se a injeção de fluido no poço localizado no limite esquerdo de Ω_r . Nos cenários aqui apresentados, o primeiro estágio começa em $t = 0$ dias e termina em $t = 96$ dias. Nesta janela de tempo, o poço de produção é despressurizado a uma taxa linear de $1,47 \times 10^5 Pa/dia$. O estágio de recuperação secundário começa em $t = 102$ dias, e a injeção de fluido é realizada com um fluxo constante de $1,5m^3/s$ terminando em $t = 300$ dias.

A calibração do modelo começa em $t = 0$ dias e vai até $t = 72$ dias. As assimilações ocorreram a cada 3 passos de tempo do problema direto, ou seja, a atualização do vetor estados estendido ocorreu a cada 18 dias. Os parâmetros identificados são então empregados no modelo computacional de duas vias para fazer previsões da taxa de produção, da pressão de injeção e da subsidência da superfície até $t = 300$ dias.

Antes de executar a análise completa, foram realizados testes preliminares para determinar o número de amostras a serem empregadas na construção da aproximação do erro do modelo e o número de membros do conjunto necessário para realizar o processo de assimilação de dados com o *EnKF*. O uso de dados sintéticos permite

controlar o erro através do erro quadrático médio residual (RMSE) apresentado na equação 5.4. Este indicador calcula, a qualquer instante de tempo t para qualquer componente do vetor de estado estendido, um desvio normalizado do conjunto a partir do valor de referência.

$$RMSE = \frac{1}{\max(S^t) - \min(S^t)} \sqrt{\frac{1}{N_r} \sum_{i=1}^{N_r} (S^t - S_i^a)^2} \quad (5.4)$$

onde N_r é o número de amostras, e S^t e S_j^a são respectivamente, o vetor de estado de referência e o vetor de estados assimilado correspondente ao membro j do conjunto de modelos assimilados. Em todos os cenários aqui ilustrados, as aproximações do erro do modelo foram construídas empregando 300 amostras e o *EnKF* empregou 400 amostras no conjunto de modelos (*Ensemble*). Além disso, o número máximo de iterações em cada passo de tempo para o modelo de duas vias foi fixado em 10 iterações.

Tabela 5.1: Parâmetros e dados da simulação

Propiedade	Valor
Viscosidade do fluido	$\mu w = 6.33 \times 10^{-9}$ Pa·s
Injeção	$q = 1.5$ m ³ /day
Pressão inicial	$p(\mathbf{x}, 0) = 7.16$ MPa
Pressão de referência	$p^{ST} = 0.1$ MPa
Pressão no fundo do poço produtor	$p = 7.16$ MPa
Amostras a priori da permeabilidade	$\delta_K = 4.87 \times 10^{-15}$ m ² $\alpha_K = 0.8$
Amostras a priori do módulo de Young	$\delta_E = 1.68 \times 10^9$ Pa $\alpha_E = 0.6$
Porosidade média do reservatório	$\bar{\phi} = 0.25$
Cumprimento de correlação	$l = 20$ m
Módulo de Young de Ω_1	$\mathbb{E} = 5.5 \times 10^{10}$ Pa
Módulo de Young de Ω_2	$\mathbb{E} = 5.0 \times 10^{10}$ Pa
Módulo de Young de Ω_3	$\mathbb{E} = 8.0 \times 10^{10}$ Pa
Módulo de Young de Ω_4	$\mathbb{E} = 2.0 \times 10^9$ Pa
Módulo de Young de Ω_5	$\mathbb{E} = 2.0 \times 10^9$ Pa
Passo de tempo na simulação	$\Delta t = 6$ dias

Foram analisados dois cenários distintos, A e B, que serão descritos em seguida, acompanhado dos respectivos resultados.

5.2 Cenário A

Na figura 5.3 é apresentada uma comparação entre a evolução temporal do termo fonte, $\frac{\beta^* \alpha \partial \bar{\sigma}_T}{\partial t}$, integrado em Ω_r da formulação de duas vias com a equação de balanço 2.26 aplicada na formulação de uma via (desconsiderando o termo fonte do lado direito). Nos dois casos foram empregadas as amostras de referência de K e \mathbb{E} apresentadas na figura 5.2. Esta figura ilustra a potencial diferença existente entre as duas formulações, e como a formulação de uma via tende a sobre estimar a produção. Lembrando que as respostas aqui apresentadas com o modelo de duas vias empregando os campos de referência de K e \mathbb{E} constituem o modelo de referência empregado na geração dos dados sintéticos que são usados na assimilação, tanto empregando a formulação de duas vias, como as formulações de uma via tradicional e aprimorada.

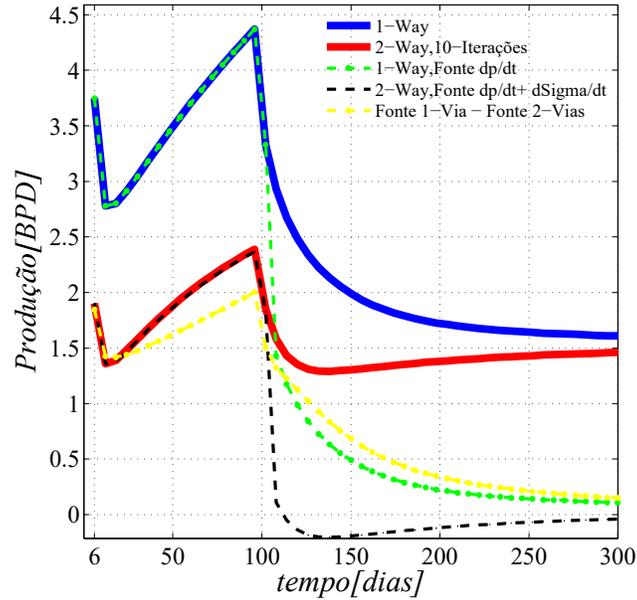
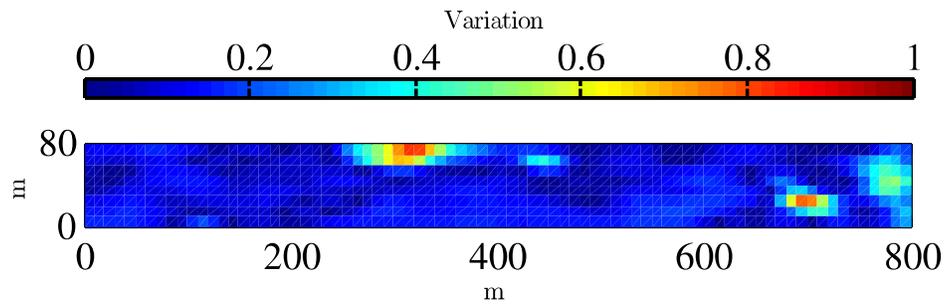


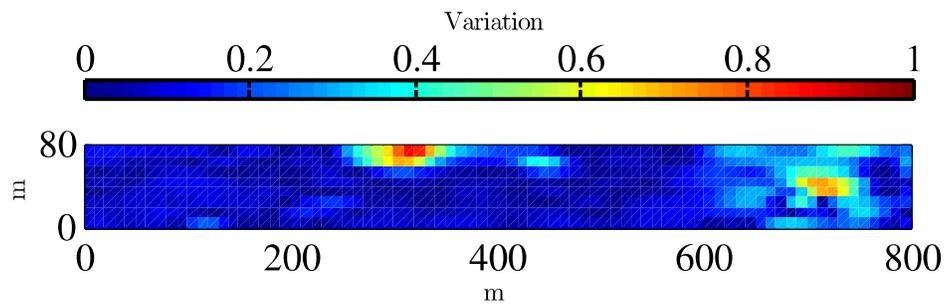
Figura 5.3: Termo fonte da equação de balanço

A seguir são, apresentados os valores identificados para parâmetros de permeabilidade e elasticidade após a assimilação de 72 dias. A distribuição espacial dos desvios dos valores de referência é medida pela média, em cada elemento da malha, sobre o conjunto de campos de propriedades identificados e normalizado pelos valores reais. A figura 5.4 mostra os campos espaciais resultantes para a permeabilidade em relação a cada um dos três esquemas de assimilação. Os desvios são maiores nas regiões de maior permeabilidade (ver a figura 5.2). Observamos que o esquema no qual o modelo aprimorado de uma via é usado obteve resultados mais precisos em comparação com o mesmo esquema sem incorporar o erro de modelo

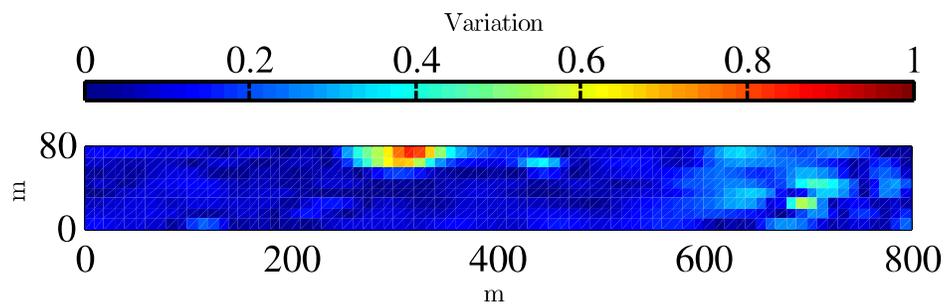
Uma análise semelhante é apresentada na Figura 5.5 para o módulo de Young.



(a) Modelo de duas vias

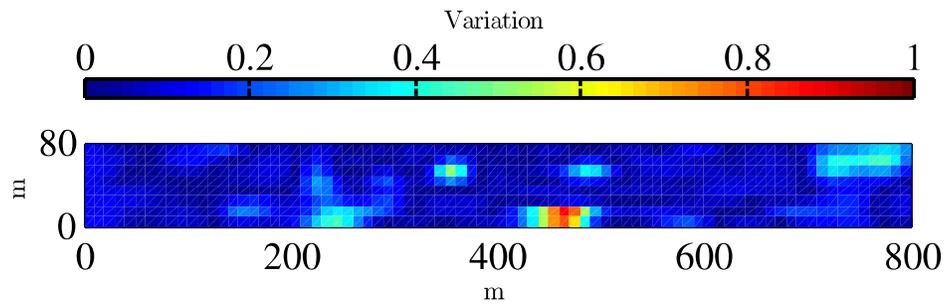


(b) Modelo de uma via

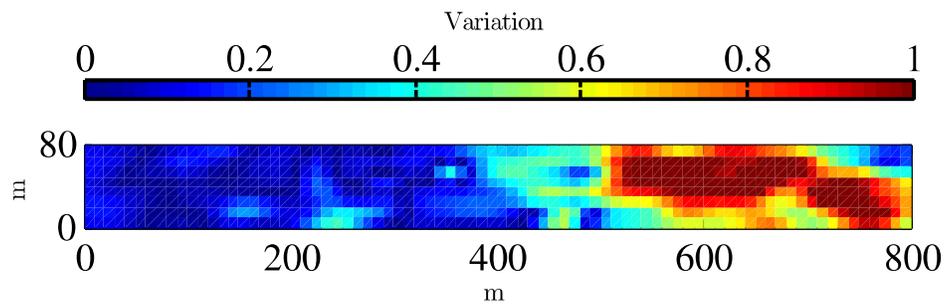


(c) Modelo de uma via aprimorado

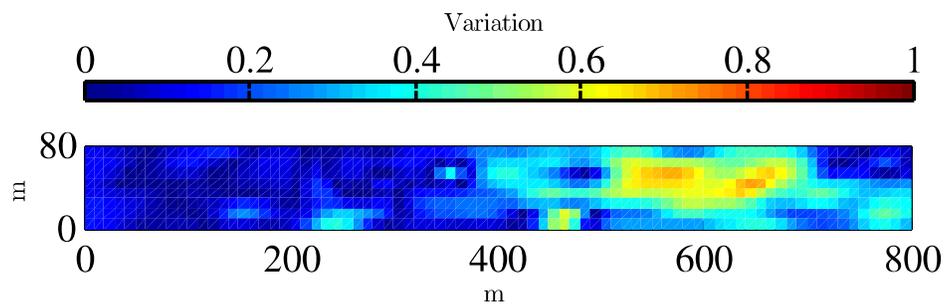
Figura 5.4: Comparação entre o desvio dos campos de permeabilidade assimilados e o campo de referência em $t = 72$ dias usando os modelos a) duas vias, b) uma via e c) uma via aprimorado



(a) Modelo de duas vias



(b) Modelo de uma via



(c) Modelo de uma via aprimorado

Figura 5.5: Comparação entre o desvio dos campos de módulo de Young assimilados e o campo de referência em $t = 72$ dias usando os modelos a)duas vias, b) uma via e c) uma via aprimorado

A região vermelha na imagem (b) mostra que a calibração com o modelo de uma via superestima o módulo de Young sobre uma região significativa do domínio do reservatório. Novamente, destacamos o papel desempenhado pela inclusão do erro de modelo no procedimento de assimilação permitindo que a formulação de uma via consiga ter uma calibração mais adequada dos parâmetros identificados.

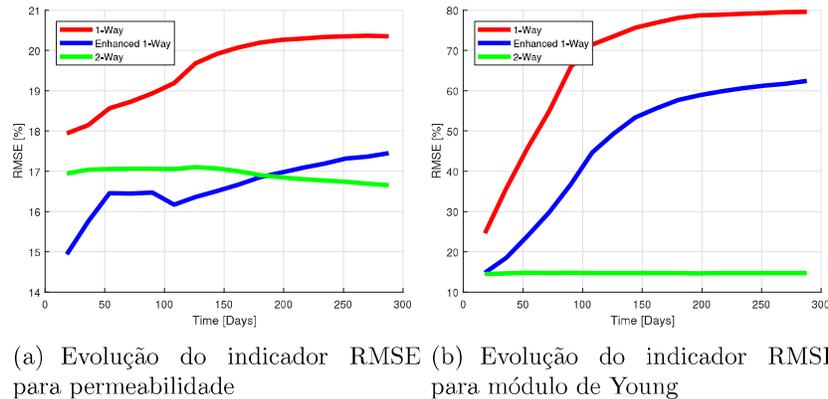


Figura 5.6: Análise do desempenho dos métodos de assimilação de dados

A assimilação é estendida além de $t = 72$ dias para investigar a capacidade dos três métodos de melhorar a calibração dos modelos com dados em um intervalo maior. Isto é ilustrado na figura 5.6 que mostra a evolução temporal do indicador RMSE computado para ambos os campos. Em todas as situações, as tendências observadas são semelhantes e os dados após certo número de iterações não afetam mais a identificação dos parâmetros. De fato, observamos fraca capacidade de obter resultados precisos para o módulo de elasticidade.

A capacidade preditiva dos modelos de reservatório após a calibração realizada com o método *EnKF* é avaliada e usada como referência da qualidade do processo de assimilação de dados executado com cada um dos três modelos. Apesar da precisão dos parâmetros identificados, os modelos calibrados produzem resultados mais precisos e exibem menos incertezas nos resultados quando comparados com seus correspondentes antes da assimilação. Tal melhoria é ilustrada nas Figuras 5.7 e 5.8 para a produção de óleo. Todas as simulações começam em $t = 0$ e empregam os campos assimilados, exceto o que emprega o modelo a priori. Cada uma das quatro imagens contidas na figura 5.7 mostra a curva de produção verdadeira (em vermelho) comparada com a produção média (em azul) juntamente com uma estimativa das incertezas nessa saída fornecida pelos resultados obtidos com cada membro da produção (linhas verdes). Além disso, na figura 5.7 se observa o cenário no caso de ausência de assimilação. Uma análise similar é resumida na figura 5.8 onde é exibida a produção média acumulada para cada modelo

Destaca-se novamente o desempenho da assimilação baseada no modelo de uma

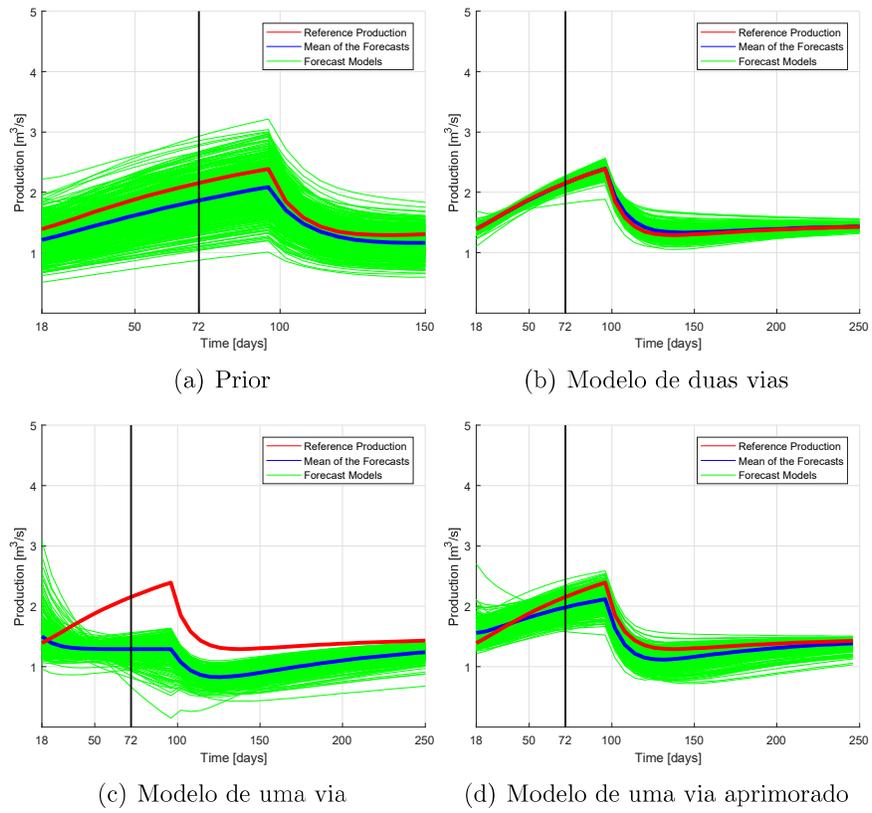


Figura 5.7: Previsão de produção empregando os modelos a) a priori, b) de duas vias, c) de uma via e d) de uma via aprimorado

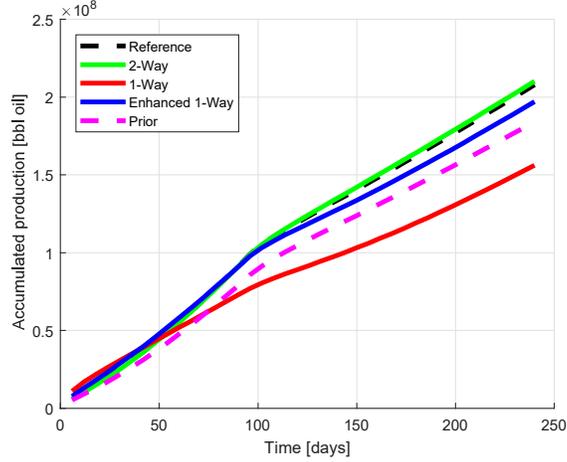
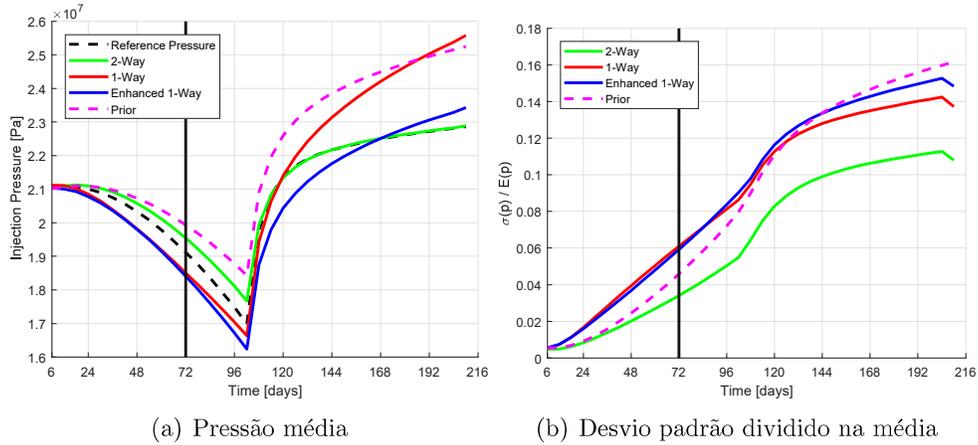


Figura 5.8: Previsão de produção média acumulada



(a) Pressão média

(b) Desvio padrão dividido na média

Figura 5.9: Média e de desvio da previsão de pressão de injeção empregando os modelos duas vias, umas via e uma aprimorada

via aprimorado. Esta última abordagem supera o modelo de uma via e fornece resultados mais precisos e próximos do obtido uma vez com o método duas vias que requer processos de computação mais intensos. Vale a pena mencionar que o modelo produzido pela calibração usando o modelo de uma via leva a uma previsão da produção global que é menos precisa na média quando comparada com a ausência de calibração. Por outro lado, as incertezas correspondentes diminuem com a assimilação, conforme ilustrado na Figura 5.7 comparando as imagens (a) e (c).

Também é comparado o desempenho dos diferentes esquemas do *EnKF* avaliando o comportamento da pressão no fundo do poço de injeção e a subsidência na superfície superior da formação geológica. A figura 5.9 apresenta a evolução temporal da pressão e uma estimativa das incertezas proporcionadas pela razão entre o desvio padrão e a média estatística. Novamente, observamos os efeitos positivos do uso do erro do modelo na assimilação, particularmente no período de injeção.

As figuras 5.10 e 5.11 mostram a previsão de subsidência da superfície. Podemos observar diferenças significativas entre as previsões obtidas com os três diferentes modelos calibrados. A figura 5.11 exibe a evolução temporal da subsidência da superfície superior entre os poços. Cada imagem representa o deslocamento vertical em momentos diferentes. A assimilação empregando o método de uma via aprimorado fornece melhores resultados em comparação com o método de uma via tradicional. Por outro lado, para esta saída particular, a média obtida com o modelo a priori produz resultados próximos ao modelo de referência e a assimilação com o modelo de duas vias. Por outro lado, suas incertezas associadas, representadas por barras de erro (desvio padrão \pm), são maiores do que aquelas calculadas com os modelos calibrados.

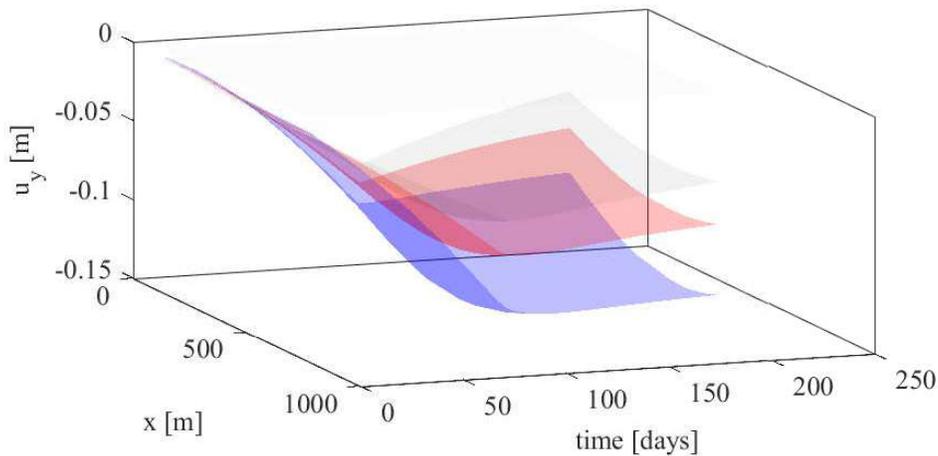
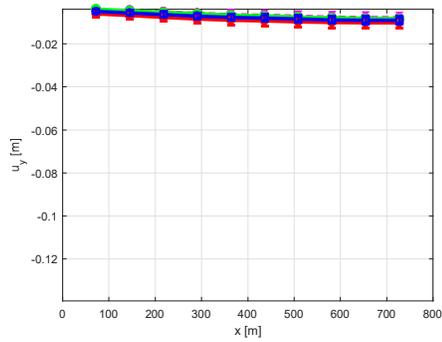


Figura 5.10: Previsão de subsidência de superfície média estimada pelos modelos duas vias (azul), uma via (cinza) e uma via aprimorada (vermelho)

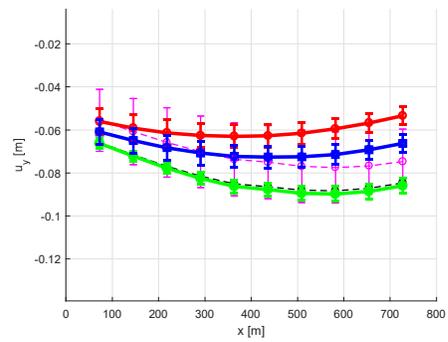
5.3 Cenário B

Este segundo cenário é construído com uma permeabilidade média maior ($\delta_K = 7.3 \times 10^{-14} m^2$ mantendo $\alpha_K = 0,8$), o que impacta diretamente o escoamento e, indiretamente, o acoplamento com a geomecânica. Essa nova configuração leva a simulação a atingir um comportamento pseudo-rígido em alguns intervalos de tempo.

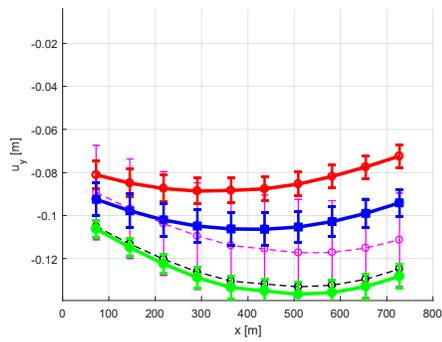
Na figura 5.12 fica novamente evidente a enorme diferença de comportamento



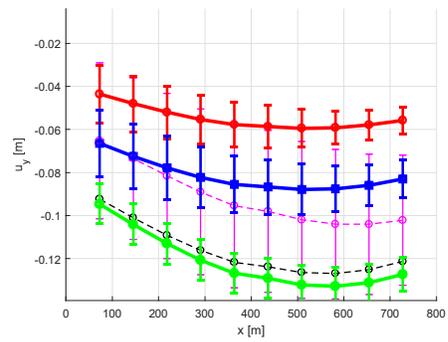
(a) $t = 6$ days



(b) $t = 72$ days



(c) $t = 102$ days



(d) $t = 204$ days

Figura 5.11: Média e desvio padrão da previsão de subsidência da superfície nos sensores de deslocamento obtida com os modelos: de referência (preto), duas vias (verde), uma via (vermelho), uma via aprimorado (azul) e a priori (magenta). Barras de erro indicam desvio padrão de \pm nas posições monitoradas

entre os modelos de referência quando calculados com a formulação de duas vias (linha vermelha) e de uma via (linha azul), neste cenário foi mantida a estratégia de empregar a formulação de duas vias como uma formulação para a geração de dados sintéticos para o processo de assimilação. Desta forma se manteve o critério de avaliar as formulações de uma via tradicional e aprimorada desde a perspectiva de sua capacidade para ter uma performance próxima à exibida pela formulação de duas vias nos processo de calibração e capacidade de previsão. Observe-se também na figura 5.12 o comportamento pseudo-rígido do reservatório neste cenário específico, ele fica evidente em que para grandes intervalos de tempo o termo fonte e a vazão de produção ficam constantes no tempo, isso implica que os efeitos geomecânicos deixaram de afetar o equilíbrio mecânico da rocha.

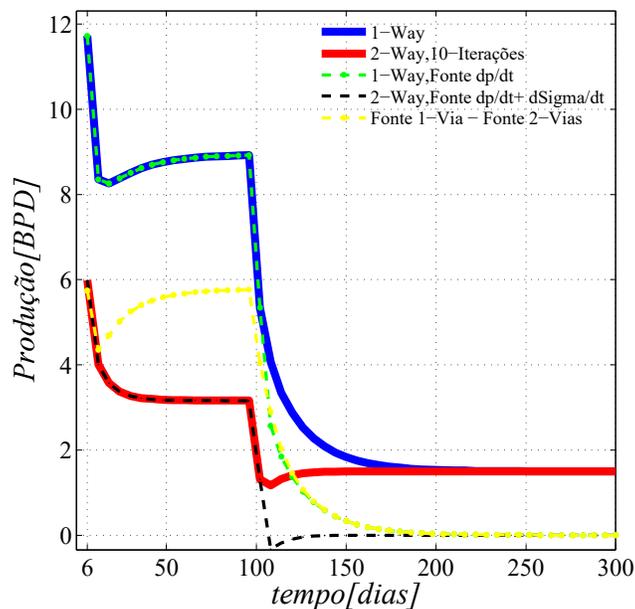
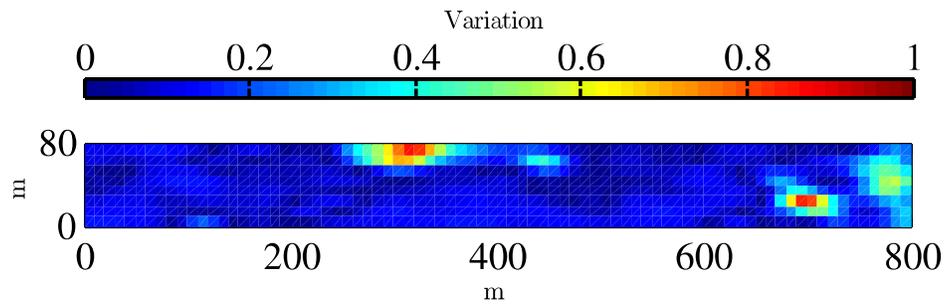


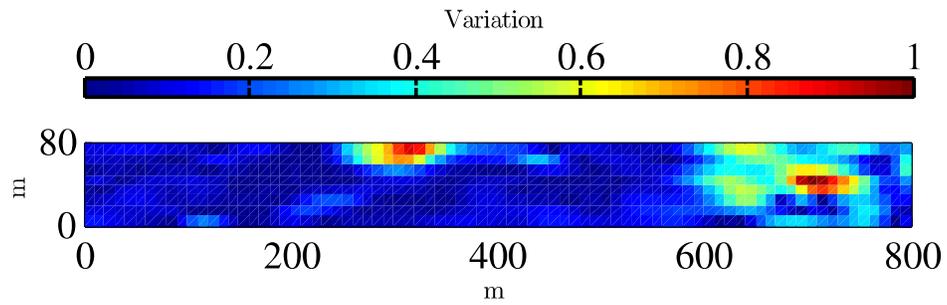
Figura 5.12: Termo fonte da equação de balanço

Na figura 5.13 a diferença espacial do campo médio assimilado com respeito ao campo de referência concentrou a discrepância nas mesmas regiões documentadas na figura 5.4 para todos os modelos. Na figura 5.2 se observa que para ditas regiões com maior discrepância na calibração de K também são regiões de grande variabilidade da permeabilidade do modelo de referência.

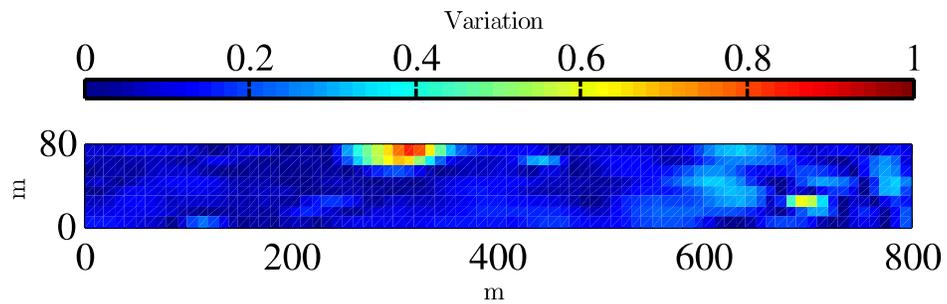
A figura 5.14, contém a distribuição espacial do desvio da média dos campos de módulo de Young assimilados com respeito ao módulo de Young de referência. A imagem (b) da figura 5.14 evidencia que o uso do modelo uma via não foi bem sucedido na tarefa de identificar o módulo de Young. Isto mostra o grande impacto que tem o uso da formulação simplificada na hora de calibrar os parâmetros diretamente relacionados com a geomecânica quando o transiente entre estados de



(a) Modelo de duas vias



(b) Modelo de uma via



(c) Modelo de uma via aprimorado

Figura 5.13: Comparação entre o desvio dos campos de permeabilidade assimilados e o campo de referência em $t = 72$ dias usando os modelos a)duas vias, b) uma via e c) uma via aprimorado

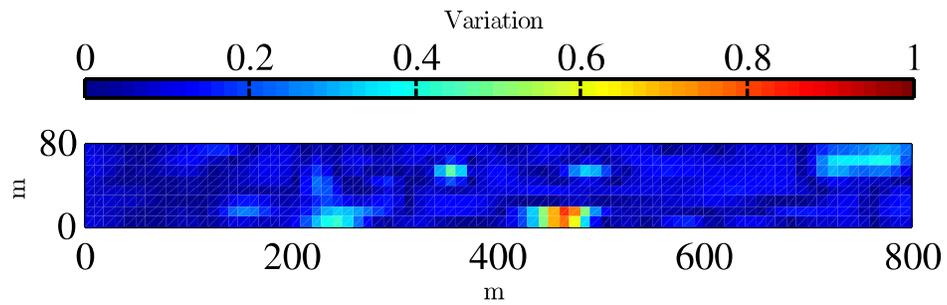
equilíbrio estrutural da rocha é curto. A imagem (b) e imagem (c) mostram que, apesar da incidência do comportamento pseudo-rígido do meio, o efeito de correção induzido pelo erro do modelo permite que o modelo de uma via obtenha um melhor desempenho identificando o campo de referência do módulo de Young.

Na assimilação simultânea de K e \mathbb{E} as informações de produção obtêm uma visão geral do comportamento do escoamento no domínio do reservatório, pressão de injeção e taxa de produção estão diretamente relacionadas com o sistema hidrodinâmico e portanto com o campo de permeabilidade mas, com um comportamento pseudo-rígido do meio, o tempo transiente até atingir o equilíbrio mecânico diminui, isso reduz o desvio da informação de subsidência levando a um processo de calibração de módulo de Young pobre. Na figura 5.15 é apresentada a evolução temporal do erro médio quadrático RMSE (equação 5.4) para o processo de calibração dos campos de permeabilidade e módulo de Young.

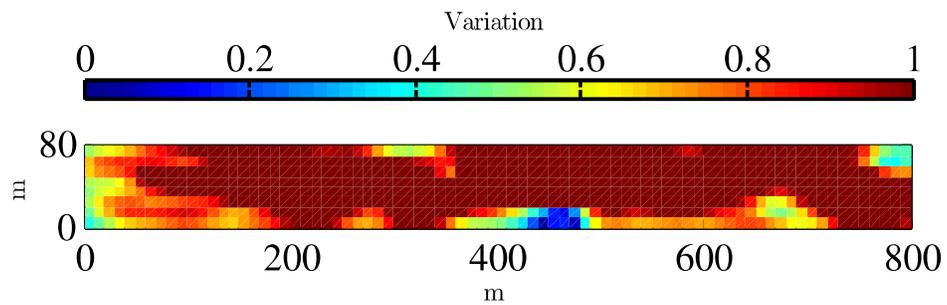
Como no teste da seção 5.2, os campos assimilados em $t = 72$ dias foram selecionados para a geração de previsões. Os valores de erro RMSE da calibração de permeabilidade para os campos selecionados foram de aproximadamente 17,54%, 24,7% e 16,5% para o modelo duas vias, uma via e uma via aprimorado respectivamente, imagem a) figura 5.15. Na calibração do módulo Young, os valores do erro RMSE foram de 13,6%, 128,5% e 35,4% para o modelo de duas vias, uma via e uma via aprimorada respectivamente, imagem b) da figura 5.15.

Na figura 5.16 são apresentadas as curvas de previsão de produção para os 400 modelos assimilados. A linha vermelha corresponde à produção calculada com o modelo de referência, a linha azul é a média da produção dos 400 modelos, as linhas verdes são a produção de cada modelo assimilado. Lembrando que os campos empregados na geração de previsões foram assimilados até $t = 72$ dias.

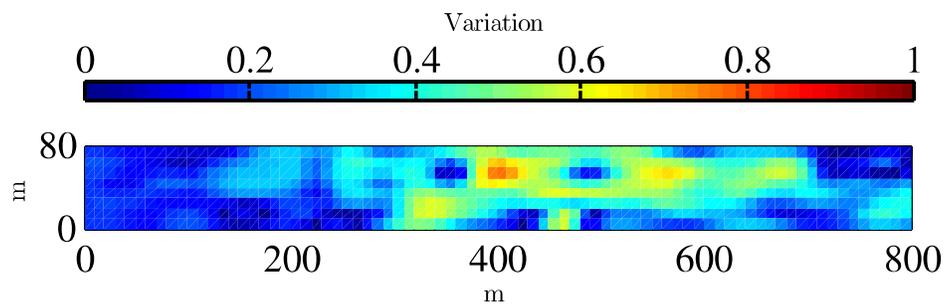
Apesar do processo de calibração de campo de permeabilidade obter valores de erro RMSE semelhantes para os resultados com o modelo de uma via e de duas vias, neste teste, a dificuldade apresentada pelo modelo de uma via tradicional na calibração do módulo de Young é refletida no desempenho da previsão de produção. O termo de aproximação do erro do modelo consegue melhorar o desempenho das previsões geradas com a formulação de uma via aprimorado, imagem (d) figura 5.16, deslocando as curvas de produção e aproximando-as da produção de referência, porém, nesta configuração com uma aproximação rápida da simulação para um estado de equilíbrio mecânico, a eficácia do método para estimar simultaneamente a permeabilidade e os campos de módulo do Young foi reduzida. Depois de $t = 102$ dias fica claro que o fluxo de injeção domina a dinâmica do reservatório e o modelo unidirecional chega à produção de referência com um pequeno atraso. Este atraso é curto porque os efeitos geomecânicos para esta configuração não apresentam grandes perturbações após esse tempo. As curvas de produção geradas com a distribuição



(a) Modelo de duas vias

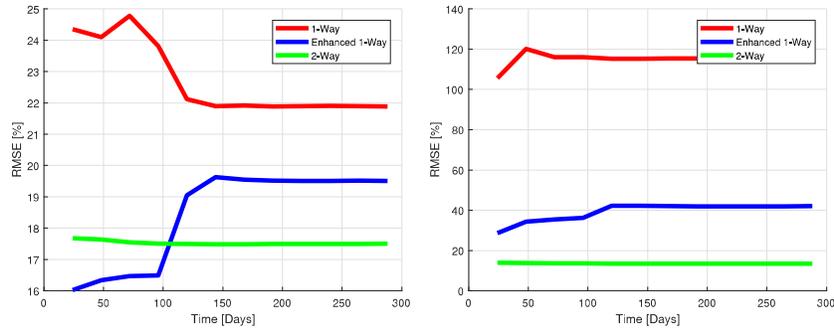


(b) Modelo de uma via



(c) Modelo de uma via aprimorado

Figura 5.14: Comparação entre o desvio dos campos de módulo de Young assimilados e o campo de referência em $t = 72$ dias usando os modelos a)duas vias, b) uma via e c) uma via aprimorado



(a) Evolução do indicador RMSE para permeabilidade (b) Evolução do indicador RMSE para módulo de Young

Figura 5.15: Análise do desempenho dos métodos de assimilação de dados

a priori, imagem (a) da figura 5.16 apresentaram uma média próxima á media á produção de referência mas com uma variabilidade entre os modelos muito grande, o que faz com que as incertezas associadas a o modelo a priori sejam muito altas.

As curvas de produção acumulada, figura 5.17, mostram como realmente o modelo de uma via aprimorado consegue ter previsões de produção mais acuradas do que o modelo de uma via, igualmente continua em evidencia que a média das produções estimadas com o modelo a priori apresenta um comportamento próximo da produção de referência e como de se esperar, a assimilação empregando o modelo de duas vias consegue capturar a produção do modelo de referência.

As previsões de pressão no poço injetor, figura 5.18 continuam mantendo a tendência de corroborar que o uso do erro de modelo aproximado a partir da discrepância entre as formulações de maior e menor fidelidade permite corrigir e ajustar as previsões geradas com o modelo de uma via. Na previsão da média da produção no poço injetor o modelo de uma via se aproxima ao resultado do modelo de duas vias, quando avaliada a variabilidade dos resultados obtido com cada um dos métodos, fica claro que o erro de modelo permitiu que o modelo de uma via aprimorado gere previsões com um desvio padrão próximo do que o desvio dos modelos assimilados empregando a formulação de duas vias.

Na figura 5.19 observa-se a grande diferença entre as previsões de subsidência da superfície geradas com as formulações de duas vias (azul), uma via aprimorada (vermelha), e uma via (cinza). Esta tendência foi a mesma evidenciada na figura 5.10 para os teste com permeabilidade media mais baixa. Isto mostra que realmente o erro de modelo permite induzir a discrepância entre as formulações corrigendo um pouco o efeito do uso de formulações computacionalmente mais eficientes, mas também mostra que é difícil para a formulação de uma via capturar plenamente os efeitos geomecânicos na assimilação de dados.

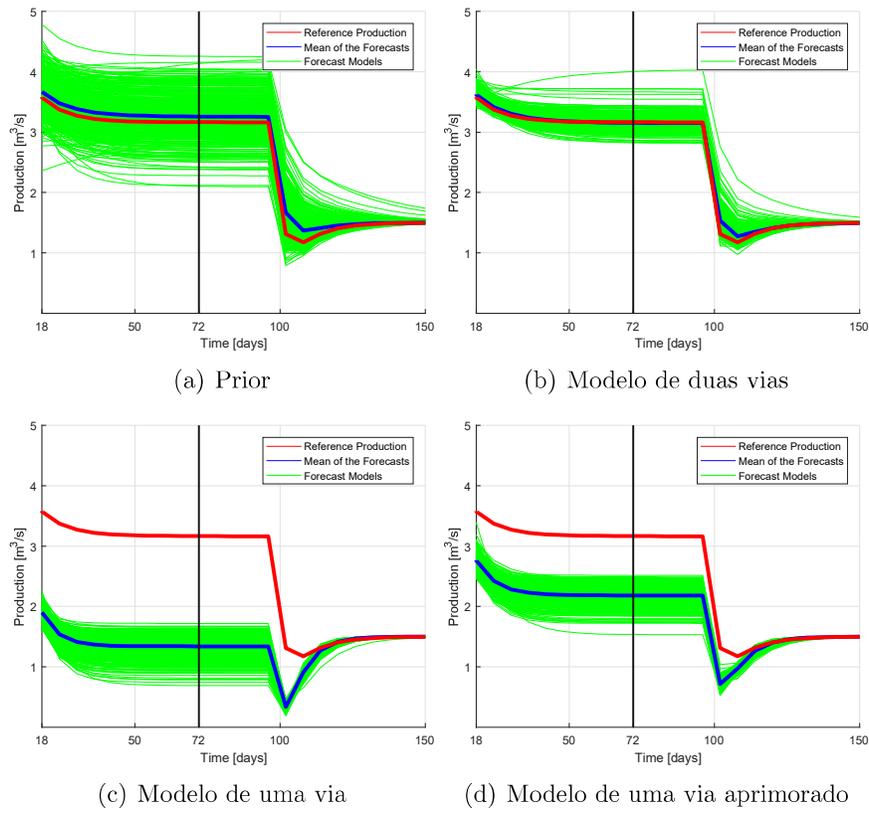


Figura 5.16: Previsão de produção empregando os modelos a) a priori, b) de duas vias, c) de uma via e d) de uma via aprimorado

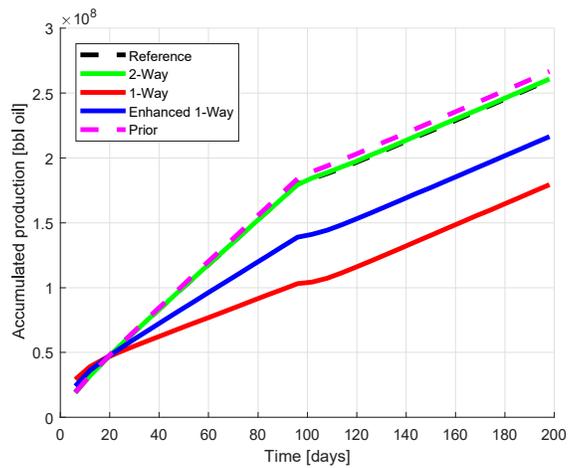


Figura 5.17: Previsão de produção média acumulada

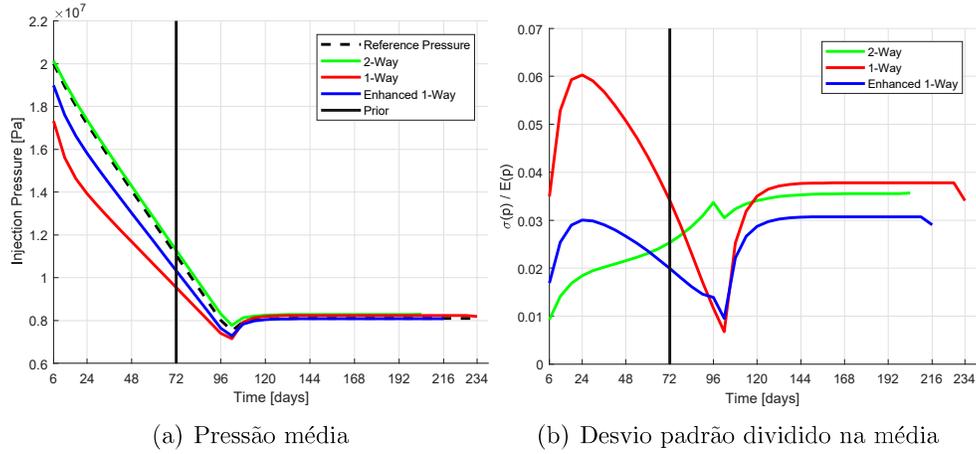


Figura 5.18: Média e de desvio da previsão de pressão de injeção empregando os modelos duas vias, uma via e uma aprimorada

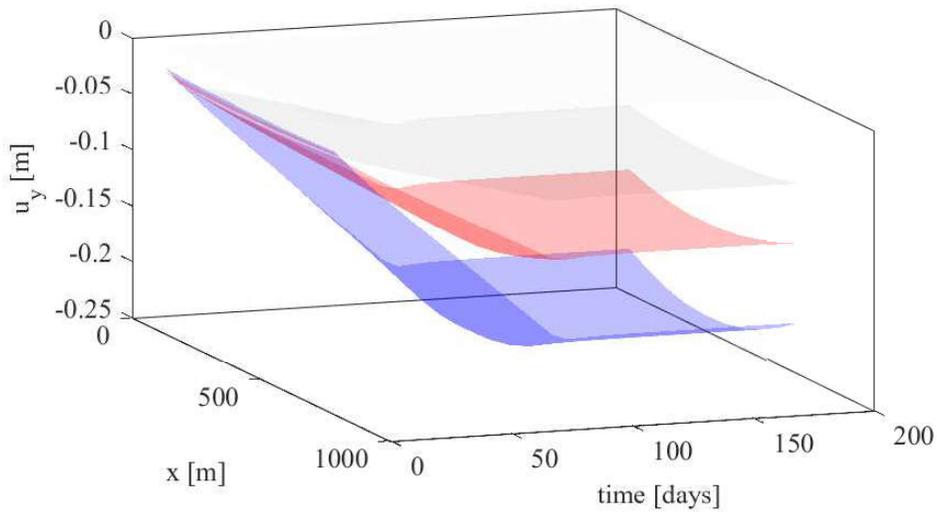
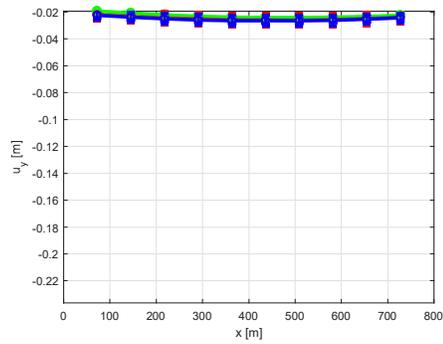
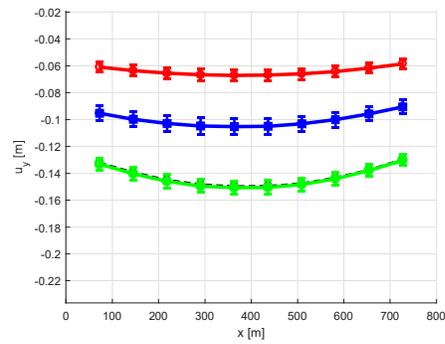


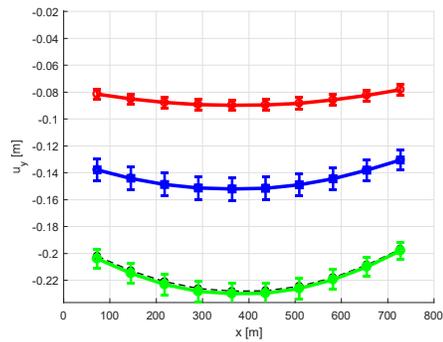
Figura 5.19: Previsão de subsidência de superfície média estimada pelos modelos duas vias (azul), uma via (cinza) e uma via aprimorada (vermelho)



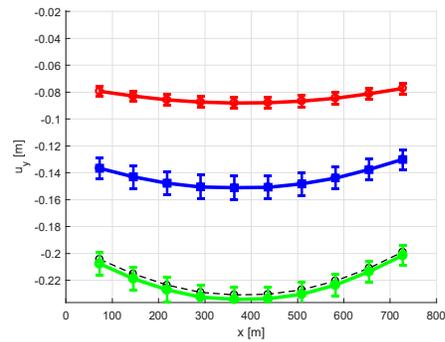
(a) $t = 6$ dias



(b) $t = 60$ dias



(c) $t = 96$ dias



(d) $t = 198$ dias

Figura 5.20: Média e desvio padrão da previsão de subsidência da superfície nos sensores de deslocamento obtida com os modelos: de referência (preto), duas vias (verde), uma via (vermelho), uma via aprimorado (azul) e a priori (magenta). Barras de erro indicam desvio padrão de \pm nas posições monitoradas

Capítulo 6

Conclusões

O *EnKF* tem-se mostrado como uma ferramenta útil e eficaz na calibração de modelos e geração de estimativas do intervalo de incertezas, mas, o fato da metodologia desde sua concepção correlacionar todas as variáveis e todos os parâmetros do problema, além, de em sua formulação mais essencial não respeitar restrições físicas nos valores assimilados faz com que ela se torne ineficiente em problemas complexos em que seja preciso empregar formulações computacionalmente complexas e possa chegar a calibrações que não representem o problema estudado.

A aproximação do erro de modelagem surge como uma alternativa para assistir a quantificação de incertezas em modelos de reservatórios de petróleo com formulações complexas. Este tipo de modelos tem grande dificuldade para se integrar às técnicas de assimilação de dados atuais. A capacidade de estimar a discrepância entre modelos de reservatórios, além da possibilidade de introduzi-la na assimilação de dados, permite um aumento na capacidade de previsão dos modelos e também gera um ponto de referência enquanto à necessidade de investir esforços na construção de modelos mais precisos.

Neste trabalho foi proposto um novo esquema de assimilação de dados baseado na estrutura do *EnKF* para calibrar um modelo sequencial no contexto do algoritmo de divisão de tensão fixa para o escoamento em reservatórios acoplado e a geomecânica. O impacto das estratégias de acoplamento na assimilação de dados foi avaliado através da comparação de testes numéricos que ilustram o desempenho da assimilação de dados com modelos de duas vias e uma via. A última abordagem consiste em um modelo simplificado que pode ser explorado para reduzir os custos computacionais envolvidos. Nossas simulações numéricas mostraram que, em alguns casos, a estratégia de uma via não produz resultados confiáveis no cenário de ajuste histórico de dados.

Para superar essa desvantagem, introduzimos uma abordagem de aproximação de erro de modelo para compensar a perda de precisão inerente ao modelo unidirecional. Nesse contexto, buscamos uma solução equilibrada em termos de precisão e custos

computacionais. A principal conquista é a construção de um método alternativo ao modelo de duas vias de alta fidelidade, que leva a custos mais altos no processo de assimilação.

Simulações numéricas ilustraram a viabilidade de incorporar este termo de compensação na estrutura do *EnKF* com custos computacionais mais baixos quando comparados ao uso do modelo de alta fidelidade. Estes resultados numéricos mostraram a melhoria na qualidade da calibração de parâmetros e a capacidade de previsão dos modelos assimilados. Por fim, observa-se que esta estratégia aqui proposta utiliza apenas dados produzidos pelo modelo de duas vias e não depende de medições de campo. Assumindo certas suposições, o erro de modelo pode ser facilmente computado antes do início do processo de assimilação.

A experiência adquirida na assimilação de dados em modelos de reservatórios de petróleo acoplados à geomecânica e no uso de aproximações de erro de modelo, para criar uma estimativa do impacto da formulação na qualidade dos campos paramétricos recuperados, leva a propor os seguintes passos futuros que permitem continuar o desenvolvimento de esta pesquisa:

- Explorar a construção de erros de modelo empregando algoritmos de *machine learning* e avaliar a precisão que pode ser atingida na assimilação de dados;
- paralelizar o algoritmo de assimilação de dados empregando o erro de modelo para testar a formulação de uma via aprimorada em geometrias de reservatórios 3D com escoamentos multifásicos;
- estudar o desempenho da assimilação de dados empregando *Iterative Ensemble Smoother* na formulação de duas vias aplicada a modelos de reservatórios acoplados à geomecânica e comparar os resultados com aspectos críticos da assimilação de dados empregando o *Iterative Ensemble Kalman Filter*;
- propor e avaliar o desempenho de uma formulação da aproximação de erro de modelagem que possa ser aliada com outros algoritmos de assimilação de dados como o *Ensemble Randomized Maximum Likelihood (EnRML)* e o *iterative Ensemble Smoother*;
- avaliar de forma sistemática a sensibilidade do processo de assimilação e da aproximação do erro de modelo em função do tipo, quantidade e localização das estações de medição empregando técnicas de projeto ótimo de experimentos;
- propor um termo, a ser construído empregando métodos estocásticos e/ou algoritmos de *machine learning*, que possa ser inserido na equação de balanço da formulação de uma via para compensar a discrepância entre os modelos diretamente em suas equações de governo.

Referências Bibliográficas

- [1] Particle filter and approximation error model for state estimation in hyperthermia. *Journal of Heat Transfer*, 139(1), 2016.
- [2] Sigurd I. Aanonsen, Geir Nævdal, Dean S. Oliver, Albert C. Reynolds, and Brice Vallès. The Ensemble Kalman Filter in Reservoir Engineering—a Review. *SPE Journal*, 14(03):393–412, September 2009.
- [3] Tarek Ahmed. *Reservoir Engineering Handbook*. ELSEVIER, fourth edition, 2010.
- [4] Richard C. Aster, Brian Borchers, and Clifford H. Thurber. Rank Deficiency and Ill-Conditioning. In *Parameter Estimation and Inverse Problems*, pages 55–91. Elsevier, 2013.
- [5] Juarez S. Azevedo, Márcio A. Murad, Marcio R. Borges, and Saulo P. Oliveira. A space-time multiscale method for computing statistical moments in strongly heterogeneous poroelastic media of evolving scales. *International Journal for Numerical Methods in Engineering*, 90(6):671–706, may 2012.
- [6] Jacob Bear. *Dynamics of Fluids in Porous Media*. Dover Publications, 1972.
- [7] Brede Rem Bergo. *Quality of the Analysis Step in EnKF*. PhD thesis, University of Bergen, 2011.
- [8] M. A. Biot. Theory of Elasticity and Consolidation for a Porous Anisotropic Solid. *Journal of Applied Physics*, 26(2):182, 1955.
- [9] M. R. Borges, F. Furtado, F. Pereira, and H. P. Amaral Souto. Scaling Analysis for the Tracer Flow Problem in Self-Similar Permeability Fields. *Multiscale Modeling & Simulation*, 7(3):1130–1147, January 2009.
- [10] Gerrit Burgers, Peter Jan van Leeuwen, and Geir Evensen. Analysis Scheme in the Ensemble Kalman Filter. *Monthly Weather Review*, 126(6):1719–1724, June 1998.

- [11] Jef Caers. *Modeling Uncertainty in the Earth Sciences*. John Wiley & Sons, Ltd, Chichester, UK, June 2011.
- [12] Haibin Chang, Yan Chen, and Dongxiao Zhang. Data Assimilation of Coupled Fluid Flow and Geomechanics Using the Ensemble Kalman Filter. *SPE Journal*, 15(02):382–394, June 2010.
- [13] Yan Chen and Dean S. Oliver. Ensemble-Based Closed-Loop Optimization Applied to Brugge Field. In *SPE Reservoir Simulation Symposium*. Society of Petroleum Engineers, April 2009.
- [14] Yan Chen and Dean S. Oliver. Ensemble Randomized Maximum Likelihood Method as an Iterative Ensemble Smoother. *Mathematical Geosciences*, 44(1):1–26, jan 2012.
- [15] Yan Chen and Dongxiao Zhang. Data assimilation for transient flow in geologic formations via ensemble Kalman filter. *Advances in Water Resources*, 29(8):1107–1122, August 2006.
- [16] Zhangxin Chen. *Reservoir Simulation*. Society for Industrial and Applied Mathematics, January 2007.
- [17] Zhangxin Chen, Guanren Huan, and Yuanle Ma. *Computational Methods for Multiphase Flows in Porous Media*. Society for Industrial and Applied Mathematics, January 2006.
- [18] Zhangxin Chen, Stephen L Lyons, and Guan Qin. The mechanical behavior of a poroelastic medium saturated with a newtonian fluid. *International Journal of Numerical Analysis and Modeling*, 1(1):75–97, 2004.
- [19] M.R. Correa and M.R. Borges. A semi-discrete central scheme for scalar hyperbolic conservation laws with heterogeneous storage coefficient and its application to porous media flow. *International Journal for Numerical Methods in Fluids*, 73(3):205–224, September 2013.
- [20] Olivier Coussy. *Poromechanics*. John Wiley & Sons, Ltd, Chichester, UK, December 2003.
- [21] Gedeon Dagan. *Flow and Transport in Porous Formations*. Springer Berlin Heidelberg, Berlin, Heidelberg, 1989.
- [22] Gedeon Dagan and Sholmo Neuman. Subsurface Flow and Transport: A Stochastic Approach. In Gedeon Dagan and Sholmo Neuman, editors, *International Hydrology Series*. Cambridge University Press, 1997.

- [23] Henry Darcy. *Les Fontaines Publiques de la Ville de Dijon*. Victor Dalmont, 1856.
- [24] Clayton V. Deutsch and André G. Journel. *GSLIB: Geostatistical Software Library and User's Guide*. Oxford University Press, second edition, 1998.
- [25] Deepak Devegowda, Elkin Arroyo, Akhil Datta-Gupta, and Sippe G. Douma. Efficient and Robust Reservoir Model Updating Using Ensemble Kalman Filter With Sensitivity-Based Covariance Localization. In *SPE Reservoir Simulation Symposium*. Society of Petroleum Engineers, April 2007.
- [26] Y. Efendiev, T. Hou, and W. Luo. Preconditioning Markov Chain Monte Carlo Simulations Using Coarse-Scale Models. *SIAM Journal on Scientific Computing*, 28(2):776–803, January 2006.
- [27] U. Eigbe, M. B. Beck, H. S. Wheeler, and F. Hirano. Kalman filtering in groundwater flow modelling: problems and prospects. *Stochastic Hydrology and Hydraulics*, 12(1):15–32, April 1998.
- [28] Alexandre A. Emerick and Albert C. Reynolds. Ensemble smoother with multiple data assimilation. *Computers & Geosciences*, 55:3–15, jun 2013.
- [29] Margaret J. Eppstein and David E. Dougherty. Simultaneous Estimation of Transmissivity Values and Zonation. *Water Resources Research*, 32(11):3321–3336, November 1996.
- [30] Geir Evensen. Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *Journal of Geophysical Research*, 99(C5):10143, 1994.
- [31] Geir Evensen. The Ensemble Kalman Filter: theoretical formulation and practical implementation. *Ocean Dynamics*, 53(4):343–367, November 2003.
- [32] Geir Evensen. Sampling strategies and square root analysis schemes for the EnKF. *Ocean Dynamics*, 54(6):539–560, December 2004.
- [33] Geir Evensen. *Data Assimilation*. Springer Berlin Heidelberg, Berlin, Heidelberg, 2009.
- [34] Geir Evensen, Joakim Hove, Hilde Meisingset, Edel Reiso, Knut Sponheim Seim, and Ø. Espelid. Using the EnKF for Assisted History Matching of a North Sea Reservoir Model. In *SPE Reservoir Simulation Symposium*. Society of Petroleum Engineers, April 2007.

- [35] Geir Evensen and Peter Jan van Leeuwen. An Ensemble Kalman Smoother for Nonlinear Dynamics. *Monthly Weather Review*, 128(6):1852–1867, jun 2000.
- [36] Frederico Furtado and Felipe Pereira. Crossover from nonlinearity controlled to heterogeneity controlled mixing in two-phase porous media flows. *Computational Geosciences*, pages 115–135, 2003.
- [37] L W Gelhar. *Stochastic Subsurface Hydrology*. Prentice-Hall, 1993.
- [38] Pierre Goovaerts. *Geostatistics for Natural Resources Evaluation*. Oxford University Press, 1997.
- [39] Wendy Graham and Dennis McLaughlin. Stochastic analysis of nonstationary subsurface solute transport: 1. Unconditional moments. *Water Resources Research*, 25(2):215–232, February 1989.
- [40] Yaqing Gu and Dean S. Oliver. History Matching of the PUNQ-S3 Reservoir Model Using the Ensemble Kalman Filter. *SPE Journal*, 10(02):217–224, June 2005.
- [41] Yaqing Gu and Dean S. Oliver. The Ensemble Kalman Filter for Continuous Updating of Reservoir Simulation Models. *Journal of Energy Resources Technology*, 128(1):79, 2006.
- [42] Yaqing Gu and Dean S. Oliver. An Iterative Ensemble Kalman Filter for Multiphase Fluid Flow Data Assimilation. *SPE Journal*, 12(04):438–446, November 2007.
- [43] Thomas M. Hamill, Jeffrey S. Whitaker, and Chris Snyder. Distance-Dependent Filtering of Background Error Covariance Estimates in an Ensemble Kalman Filter. *Monthly Weather Review*, 129(11):2776–2790, November 2001.
- [44] Mohamed M. Hantush and Miguel A. Mariño. Estimation of Spatially Variable Aquifer Hydraulic Properties Using Kalman Filtering. *Journal of Hydraulic Engineering*, 123(11):1027–1035, November 1997.
- [45] P. L. Houtekamer and Herschel L. Mitchell. Data Assimilation Using an Ensemble Kalman Filter Technique. *Monthly Weather Review*, 126(3):796–811, March 1998.
- [46] P. L. Houtekamer and Herschel L. Mitchell. A Sequential Ensemble Kalman Filter for Atmospheric Data Assimilation. *Monthly Weather Review*, 129(1):123–137, January 2001.

- [47] P. L. Houtekamer, Herschel L. Mitchell, Gérard Pellerin, Mark Buehner, Martin Charron, Lubos Spacek, and Bjarne Hansen. Atmospheric Data Assimilation with an Ensemble Kalman Filter: Results with Real Observations. *Monthly Weather Review*, 133(3):604–620, March 2005.
- [48] Janne M.J. Huttunen and Jari P. Kaipio. Model reduction in state identification problems with an application to determination of thermal parameters. *Applied Numerical Mathematics*, 59(5):877–890, May 2009.
- [49] Marco A. Iglesias, Kody J. H. Law, and Andrew M. Stuart. Evaluation of Gaussian approximations for data assimilation in reservoir models. *Computational Geosciences*, 17(5):851–885, October 2013.
- [50] Marco A Iglesias and Dennis McLaughlin. Data inversion in coupled subsurface flow and geomechanics models. *Inverse Problems*, 28(11):115009, November 2012.
- [51] Jari Kaipio and Erkki Somersalo. *Statistical and Computational Inverse Problems*. Springer, 2004.
- [52] R. E. Kalman. A New Approach to Linear Filtering and Prediction Problems. *Journal of Basic Engineering*, 82(1):35, 1960.
- [53] Marc C. Kennedy and Anthony O’Hagan. Bayesian calibration of computer models. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 63(3):425–464, aug 2001.
- [54] khalid Aziz and Antonín Settari. *Petroleum Reservoir Simulation*. Applied Science Publishers LTD, 1979.
- [55] Ville Kolehmainen, Tanja Tarvainen, Simon R. Arridge, and Jari P. Kaipio. MARGINALIZATION OF UNINTERESTING DISTRIBUTED PARAMETERS IN INVERSE PROBLEMS - APPLICATION TO DIFFUSE OPTICAL TOMOGRAPHY. *International Journal for Uncertainty Quantification*, 1(1):1–17, 2011.
- [56] Janne Koponen, Tomi Huttunen, Tanja Tarvainen, and Jari Kaipio. Bayesian approximation error approach in full-wave ultrasound tomography. *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, 61(10):1627–1637, October 2014.
- [57] Janne Koponen, Tomi Huttunen, Tanja Tarvainen, and Jari P. Kaipio. Bayesian approximation error approach in full-wave ultrasound tomography.

IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control, 61(10):1627–1637, 2014.

- [58] Alexander Kurganov and Eitan Tadmor. New High Resolution Central Schemes for Nonlinear Conservation Laws and Convection Diffusion Equations. *Journal of Computational Physics*, 160(1):241–282, May 2000.
- [59] Timo Lähivaara, Nicholas F Dudley Ward, Tomi Huttunen, Janne Koponen, and Jari P Kaipio. Estimation of aquifer dimensions from passive seismic signals with approximate wave propagation models. *Inverse Problems*, 30(1):015003, January 2014.
- [60] Ahmed H. Elsheikh Laureline Josset, Vasily Demyanov and Ivan Lunati. Accelerating monte carlo markov chains with proxy and error models. *Computer and Geosciences*, 85:38–48, 2015.
- [61] C. H. Leng and H. D. Yeh. Aquifer parameter identification using the extended Kalman filter. *Water Resources Research*, 39(3):n/a–n/a, March 2003.
- [62] Gaoming Li and Albert Coburn Reynolds. An Iterative Ensemble Kalman Filter for Data Assimilation. In *SPE Annual Technical Conference and Exhibition*. Society of Petroleum Engineers, April 2007.
- [63] Heng Li, Shengnan Chen, Daoyong Yang, and Paitoon Tontiwachwuthikul. Ensemble-Based Relative Permeability Estimation Using B-Spline Model. *Transport in Porous Media*, 85(3):703–721, December 2010.
- [64] N. Liu and D.S. Oliver. Critical Evaluation of the Ensemble Kalman Filter on History Matching of Geologic Facies. In *SPE Reservoir Simulation Symposium*. Society of Petroleum Engineers, apr 2005.
- [65] Ning Liu and Dean S. Oliver. Ensemble Kalman filter for automatic history matching of geologic facies. *Journal of Petroleum Science and Engineering*, 47(3-4):147–161, June 2005.
- [66] Rolf J. Lorentzen, Kjell Kåre Fjelde, Jonny Frøyen, Antonio C.V.M. Lage, Geir Nævdal, and Erlend H. Vefring. Underbalanced and Low-head Drilling Operations: Real Time Interpretation of Measured Data and Operational Support. In *SPE Annual Technical Conference and Exhibition*. Society of Petroleum Engineers, April 2001.
- [67] Rolf Johan Lorentzen, Geir Nævdal, Brice Valles, Aina Berg, and Alv-Arne Grimstad. Analysis of the Ensemble Kalman Filter for Estimation of

- Permeability and Porosity in Reservoir Models. In *SPE Annual Technical Conference and Exhibition*. Society of Petroleum Engineers, April 2005.
- [68] Marcos Alcoforado Mendes, Marcio A. Murad, and Felipe Pereira. A new computational strategy for solving two-phase flow in strongly heterogeneous poroelastic media of evolving scales. *International Journal for Numerical and Analytical Methods in Geomechanics*, 36(15):1683–1716, October 2012.
- [69] Meghdoot Mozumder, Tanja Tarvainen, Jari P. Kaipio, Simon R. Arridge, and Ville Kolehmainen. Compensation of modeling errors due to unknown domain boundary in diffuse optical tomography. *Journal of the Optical Society of America A*, 31(8):1847, August 2014.
- [70] Marcio a. Murad, Marcio Borges, Jesus a. Obregón, and Maicon Correa. A new locally conservative numerical method for two-phase flow in heterogeneous poroelastic media. *Computers and Geotechnics*, 48:192–207, 2013.
- [71] Geir Naevdal, Liv Merete Johnsen, Sigurd I. Aanonsen, and Erlend H. Vefring. Reservoir Monitoring and Continuous Model Updating Using Ensemble Kalman Filter. *SPE Journal*, 10(01):66–74, March 2005.
- [72] Andrew Oghena. *Quantification of Uncertainties Associated With Reservoir Performance Simulation*. PhD thesis, Texas Tech University, 2007.
- [73] Dean S. Oliver and Miguel Alfonzo. Calibration of imperfect models to biased observations. *Computational Geosciences*, 2017.
- [74] Dean S. Oliver and Yan Chen. Recent progress on reservoir history matching: a review. *Computational Geosciences*, 15(1):185–221, January 2011.
- [75] Dean S. Oliver, Albert C. Reynolds, and Ning Liu. *Inverse Theory for Petroleum Reservoir Characterization and History Matching*. Cambridge University Press, Cambridge, 2008.
- [76] Donald Peaceman. *Fundamentals of Numerical Reservoir Simulation*. ELSEVIER, 1977.
- [77] Pascal Pernot and Fabien Cailliez. A critical review of statistical calibration/prediction models handling data inconsistency and model inadequacy. *American Institute of Chemical Engineers*, 63(10):4642–4665, 2017.
- [78] Michael J. Pyrcz and Christopher D. White. Uncertainty in reservoir modeling. *Interpretation*, 3(2):SQ7–SQ19, May 2015.

- [79] Adalberto José Rosa, Renato Carvalho, and José Xavier. *Engenharia de Reservatórios de Petróleo*. Interciência, 2006.
- [80] A. Seiler, G. Evensen, J.-A. Skjervheim, J. Hove, and J. G. Vabø. Using the Ensemble Kalman Filter for History Matching and Uncertainty Quantification of Complex Reservoir Models. In *Large-Scale Inverse Problems and Quantification of Uncertainty*, pages 247–271. John Wiley & Sons, Ltd, Chichester, UK, October 2010.
- [81] A. Settari, D.A. Walters, and G.A. Behie. Use of Coupled Reservoir and Geomechanical Modelling for Integrated Reservoir Analysis and Management. *Journal of Canadian Petroleum Technology*, 40(12), December 2001.
- [82] Jan-arild Skjervheim and Geir Evensen. An Ensemble Smoother for Assisted History Matching. In *SPE Reservoir Simulation Symposium*. Society of Petroleum Engineers, apr 2011.
- [83] Jan-Arild Skjervheim, Geir Evensen, Sigurd Ivar Aanonsen, Bent Ole Ruud, and Tor-Arne Johansen. Incorporating 4D Seismic Data in Reservoir Simulation Models Using Ensemble Kalman Filter. *SPE Journal*, 12(03):282–292, September 2007.
- [84] T Tarvainen, V Kolehmainen, A Pulkkinen, M Vauhkonen, M Schweiger, S R Arridge, and J P Kaipio. An approximation error approach for compensating for modelling errors between the radiative transfer equation and the diffusion approximation in diffuse optical tomography. *Inverse Problems*, 26(1):015005, January 2010.
- [85] Kristian Thulin, Gaoming Li, Sigurd Aanonsen, and Albert Coburn Reynolds. Estimation of Initial Fluid Contacts by Assimilation of Production Data With EnKF. In *SPE Annual Technical Conference and Exhibition*. Society of Petroleum Engineers, April 2007.
- [86] Olli-Pekka Tossavainen, Julie Percelay, Mark Stacey, Jari P. Kaipio, and Alexandre Bayen. State estimation and modeling error approach for 2-D shallow water equations and Lagrangian measurements. *Water Resources Research*, 47(10):n/a–n/a, October 2011.
- [87] David Tran, Antonin Settari, and Long Nghiem. New Iterative Coupling Between a Reservoir Simulator and a Geomechanics Module. In *SPE/ISRM Rock Mechanics Conference*. Society of Petroleum Engineers, April 2002.

- [88] Peter Jan van Leeuwen and Geir Evensen. Data Assimilation and Inverse Methods in Terms of a Probabilistic Formulation. *Monthly Weather Review*, 124(12):2898–2913, dec 1996.
- [89] Jian-xun Wang, J. Roy Christopher, and Xiao Heng. A Multi-Model Approach for Uncertainty Propagation and Model Calibration in CFD Applications. 2015.
- [90] Xian-Huan Wen and Wen H. Chen. Real-Time Reservoir Model Updating Using Ensemble Kalman Filter With Confirming Option. *SPE Journal*, 11(04):431–442, December 2006.
- [91] Xian-Huan Wen and Wen H. Chen. Some Practical Issues on Real-Time Reservoir Model Updating Using Ensemble Kalman Filter. *SPE Journal*, 12(02):156–166, June 2007.
- [92] Guocan Wu, Xiaogu Zheng, Liqun Wang, Shupeng Zhang, Xiao Liang, and Yong Li. A new structure for error covariance matrices and their adaptive estimation in EnKF assimilation. *Quarterly Journal of the Royal Meteorological Society*, 139(672):795–804, 2013.
- [93] Peng Xu and Boming Yu. Developing a new form of permeability and Kozeny-Carman constant for homogeneous porous media by means of fractal geometry. *Advances in Water Resources*, 31:74–81, 2008.
- [94] Tianfang Xu and Albert J. Valocchi. A bayesian approach to improved calibration and prediction of groundwater models with structural error. *Water Resources Research*, 51:9290–9311, 2015.
- [95] Mohammad Zafari and Albert Reynolds. Assessing the Uncertainty in Reservoir Description and Performance Predictions With the Ensemble Kalman Filter. *SPE Journal*, 12(03):382–391, September 2007.
- [96] Shike Zhang and Shunde Yin. Reservoir geomechanical parameters identification based on ground surface movements. *Acta Geotechnica*, 8(3):279–292, June 2013.
- [97] Mark D. Zoback. *Reservoir Geomechanics*. Cambridge University Press, Cambridge, 2007.