



Universidade Federal de Pernambuco
Centro de Tecnologia e Geociências
Pós-Graduação em Engenharia Civil

ROBERTO NAVARRO PAREJA

Ajuste ao Histórico em Reservatórios de Petróleo
Usando o Método do Filtro de Kalman
con Ensembles (EnKF)

Recife, 2014

ROBERTO NAVARRO PAREJA

Ajuste ao Histórico em Reservatórios de Petróleo
Usando o Método do Filtro de Kalman
con Ensembles (EnKF)

Dissertação de Mestrado submetida ao Corpo de Docente do Curso de Pós-Graduação em Engenharia Civil da Universidade Federal de Pernambuco, como parte dos requisitos necessários à obtenção do grau de Mestre em Ciências em Engenharia Civil.

Ramiro Brito Willmersdorf, Ph. D.
Orientador

Área de Concentração: Simulação e Gerenciamento de Reservatórios

Recife, 2014.

Catálogo na fonte
Bibliotecária Margareth Malta, CRB-4 / 1198

P227a Pareja, Roberto Navarro.
Ajuste ao histórico em reservatórios de petróleo usando o Método do Filtro de Kalman con Ensembles (EnKF) / Roberto Navarro Pareja. - Recife: O Autor, 2014.
108 folhas, il., gráfs., tabs.

Orientador: Prof. Dr. Ramiro Brito Willmersdorf.
Dissertação (Mestrado) – Universidade Federal de Pernambuco. CTG. Programa de Pós-Graduação em Engenharia Civil, 2014.
Inclui Referências e Anexo.

1. Engenharia Civil. 2. Ajuste ao Histórico. 3. EnKF. 4. Assimilação de Dados. 5. MRST. I. Willmersdorf, Ramiro Brito. (Orientador). II. Título.

UFPE

624 CDD (22. ed.)

BCTG/2015-223



UNIVERSIDADE FEDERAL DE PERNAMBUCO
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA CIVIL
A comissão examinadora da Defesa de Dissertação de Mestrado

AJUSTE AO HISTÓRICO EM RESERVATÓRIOS DE PETRÓLEO USANDO
O FILTRO DE KALMAN CON ENSEMBLES

defendida por

Roberto Navarro Pareja

Considera o candidato APROVADO

Recife, 26 de Agosto de 2014

Banca Examinadora:

Prof. Dr. Ramiro Brito Willmersdorf - UFPE
(Orientador)

Prof. Dr. Raydonal Ospina Martinez - UFPE
(examinador externo)

Prof. Dr. Ézio da Rocha Araujo - UFPE
(examinador interno)

Dedicado

Dedico este trabalho a minha família, fonte de energia necessária para acordar, trabalhar, lutar e sonhar.

Aos meus pais Vicente e Marcelina, meu eterno agradecimento. Meu pai exemplo de superação e esforço constante na vida, minha mãe exemplo de amor e fé.

A minha noiva Paola, agradeço o amor, carinho e compreensão que teve durante o desenvolvimento de meu trabalho.

Aos meus irmãos, sogros e cunhados, meu agradecimento pelo apoio brindado em todos os momentos.

Agradecimentos

Ao longo deste trabalho, tenho aprendido muito e gostaria de aproveitar nesta oportunidade para agradecer a todas as pessoas que ajudaram e me orientaram que, de alguma forma, contribuíram para a sua realização.

A Universidade Federal de Pernambuco (UFPE), ao programa de Pós-graduação em Engenharia Civil, a CAPES pela concessão da bolsa e Rede SIGER.

Ao professor Ramiro Brito Willmersdorf, pela sua valiosa orientação para a realização deste trabalho e pelo tempo investido durante seus ensinamentos.

Aos professores Antônio Barboza e Darlan Carvalho por as aulas motivadoras, por suas recomendações e amizade.

Aos professores Leonardo Guimarães, Ivaldo Pontes, Lícia Costa, Silvana Afonso e Bernardo Horowitz, que contribuíram na minha formação profissional.

A meus amigos e companheiros do LABCOM, Marcio, Eduardo, Felipe, Gustavo, Lorena, Marcelo, Túlio, Luciolo, Tammhy, Nilson, Tássio, pelo convívio e amizade ao longo destes dois anos compartilhando momentos difíceis do Mestrado.

Para Manuel, Fernando e Liliane pelo convívio, solidariedade, amizade e conhecimentos brindados durante o desenvolvimento do trabalho.

A meus amigos Ricardo, Fernando, Omar, Ernesto e Alberto, que com paciência estavam ao meu lado me incentivando sempre a continuar.

Aos colegas da Pós-graduação, Rodrigo, Jefferson pela troca de ideais.

E, principalmente, a Deus, pela força que me dá para que eu sempre continue buscando meu aperfeiçoamento.

Resumo

A simulação de reservatórios é uma importante ferramenta usada pela indústria do petróleo para o gerenciamento de reservatórios. A fim de obter previsões da produção de óleo confiáveis, diferentes propriedades petrofísicas do reservatório, como porosidade e permeabilidade são usadas nos modelos de reservatórios. Porém, medições diretas dessas propriedades são possíveis apenas em alguns poucos poços. Uma forma de melhorar o conhecimento sobre essas propriedades é através do processo de ajuste ao histórico. O ajuste ao histórico consiste em melhorar estimativas de propriedades do reservatório usadas na construção de um modelo de reservatório de forma que as previsões do modelo se aproximem dos dados medidos em campo. Nesta dissertação apresentamos um estudo para o ajuste ao histórico automático baseado em um modelo areal, isto é, que considera o reservatório plano e horizontal, descrito por apenas duas dimensões geométricas, de um reservatório bifásico (óleo/água), onde desejamos estimar a distribuição de permeabilidades do reservatório.

Devido à sua simplicidade e eficiência, o método do Filtro de Kalman com Ensembles (EnKF), é usado para assimilar as medições estáticas e dinâmicas, atualizando continuamente as propriedades do reservatório. O EnKF nos últimos anos tem ganhado muita popularidade, é um método de assimilação de dados para modelos dinâmicos não lineares de alta dimensão e portanto adequado para ser usado no ajuste ao histórico de modelos de simulação de reservatórios.

O EnKF foi implementado em Matlab e acoplado ao Matlab Reservoir Simulation Toolbox (MRST), que foi desenvolvido pelo SINTEF para simulação de reservatórios, e foi aplicado a dois casos sintéticos simples. Os resultados mostraram que a rotina EnKF funcionou corretamente, mostrando-se que, para muitos dos parâmetros com incerteza inicial, esta foi reduzida a um nível aceitável, para a produção de petróleo e água.

Palavras chaves: Ajuste ao Histórico. EnKF. Assimilação de Dados. MRST.

Abstract

Reservoir simulation is an important tool used by the oil industry for reservoir management. In order to obtain reliable predictions of oil production, different petrophysical properties such as porosity and permeability are used to build the reservoir models. However, direct measurements of these properties are only possible in a few wells. One way to improve the knowledge of these properties is through the history matching process. History matching improves the estimates of reservoir properties used in the construction of the reservoir model so that the model predictions are closer to the measured production of the field. In this paper we present a study for an automatic history matching based upon and two-dimensional model of two-phase (oil/water) reservoir, where we wish to improve the estimate of the distribution of the reservoir permeabilities.

Due to its simplicity and efficiency, the method of the Ensemble Kalman Filter (EnKF) is used to assimilate the static and dynamic measurements, continuously updating the properties of the reservoir. The EnKF, in recent years has gained much popularity, as it is a method for dynamic data assimilation for nonlinear models of high dimension and therefore suitable for use in history matching models of reservoir simulations.

The EnKF was implemented in Matlab and coupled to Matlab Reservoir Simulation Toolbox (MRST), which was developed by SINTEF for reservoir simulation, and was applied to two simple synthetic cases. The results showed that the EnKF routine works properly, showing that, for many of the parameters and initial uncertainty has been reduced to an acceptable level for the production of oil and water.

Keywords: History Matching. EnKF. Data Assimilation. MRST.

Lista de Figuras

Figura 2.1: Seção transversal de uma amostra de rocha de reservatório.....	22
Figura 2.2: Rocha-reservatório contendo três fluidos: água, óleo e gás.....	24
Figura 4.1: Esquema de como assimilação de dados (DA) funciona e agrega valor à observação e informações do modelo.....	24
Figura 4.2: O filtro de Kalman discreto.....	27
Figura 4.3: Assimilação de dados com EnKF para o membro j do ensemble.....	32
Figura 4.4: Algoritmo Básico do EnKF.....	36
Figura 5.1: Simulação da Produção em Reservatórios.....	38
Figura 5.2: Exemplos de Módulos MRST.....	39
Figura 5.3: Amostragem gerada pelo método de Monte Carlo.....	41
Figura 5.4: Amostragem gerada pelo método de Quasi-Monte Carlo.....	41
Figura 5.5: Conjunto inicial de pontos geradores.....	42
Figura 5.6: O método de tesselação de Voronoi para os pontos da Figura 3.3.....	42
Figura 5.7: O método de tesselação de Voronoi (CVT) após atingir a condição final de uniformidade.....	42
Figura 5.8: Amostragem gerada pelo método LHS num espaço bidimensional (Giunta et al, 2003).....	43
Figura 5.9: Ilustração da EnKF a partir do ponto de vista da concepção bayesiana.....	44
Figura 5.10: Fluxo de trabalho de ajuste ao histórico baseado em EnKF para os dados de produção.....	45
Figura 5.11: Ajuste ao histórico tradicional e o baseado em EnKF.....	47
Figura 6.1: Campo de permeabilidade inicial.....	49

Figura 6.2: Campo Log-permeabilidade para 1825 dias e 3650 dias.....	51
Figura 6.3: Campo Log-permeabilidade para 5475 dias e 7300 dias.....	51
Figura 6.4: Campo Log-permeabilidade e Porosidade do ensemble médio.....	51
Figura 6.5: Campo de Pressão e Saturação do ensemble médio.....	52
Figura 6.6: Produção de óleo acumulada. A linha vermelha corresponde ao modelo verdadeiro, a linha azul corresponde ao ensemble médio e as linhas cinzas às realizações.....	52
Figura 6.7: Produção de óleo acumulada. A linha vermelha corresponde ao modelo verdadeiro, a linha azul corresponde ao ensemble médio e as linhas cinzas às realizações.....	53
Figura 6.8: Taxa de Óleo. A linha vermelha corresponde ao modelo verdadeiro, a linha azul corresponde ao ensemble médio e as linhas cinzas às realizações.....	53
Figura 6.9: Taxa de Óleo. A linha vermelha corresponde ao modelo verdadeiro, a linha azul corresponde ao ensemble médio e as linhas cinzas às realizações	54
Figura 6.10: Corte de Água. A linha vermelha corresponde ao modelo verdadeiro, a linha azul corresponde ao ensemble médio e as linhas cinzas às realizações	54
Figura 6.11: Corte de Água. A linha vermelha corresponde ao modelo verdadeiro, a linha azul corresponde ao ensemble médio e as linhas cinzas às realizações.....	55
Figura 6.12: Campo de permeabilidade inicial.....	57
Figura 6.13: Campo Log-permeabilidade para 1825 dias e 3650 dias.....	58
Figura 6.14: Campo Log-permeabilidade para 5475 dias e 7300 dias.....	58
Figura 6.15: Campo de Log-permeabilidade e porosidade do ensemble médio.....	59
Figura 6.16: Campo de Pressão e Saturação do ensemble médio.....	59
Figura 6.17: Produção de óleo acumulada. A linha vermelha corresponde ao modelo verdadeiro, a linha azul corresponde ao ensemble médio e as linhas cinzas às realizações.....	60
Figura 6.18: Produção de óleo acumulada. A linha vermelha corresponde ao modelo verdadeiro, a linha azul corresponde ao ensemble médio e as linhas cinzas às realizações.....	60
Figura 6.19: Produção de óleo acumulada. A linha vermelha corresponde ao modelo verdadeiro, a linha azul corresponde ao ensemble médio e as linhas cinzas às realizações.....	61

Figura 6.20: Produção de óleo acumulada. A linha vermelha corresponde ao modelo verdadeiro, a linha azul corresponde ao ensemble médio e as linhas cinzas às realizações.....	61
Figura 6.21: Taxa de Óleo. A linha vermelha corresponde ao modelo verdadeiro, a linha azul corresponde ao ensemble médio e as linhas cinzas às realizações.....	62
Figura 6.22: Taxa de Óleo. A linha vermelha corresponde ao modelo verdadeiro, a linha azul corresponde ao ensemble médio e as linhas cinzas às realizações.....	62
Figura 6.23: Taxa de Óleo. A linha vermelha corresponde ao modelo verdadeiro, a linha azul corresponde ao ensemble médio e as linhas cinzas às realizações.....	63
Figura 6.24: Taxa de Óleo. A linha vermelha corresponde ao modelo verdadeiro, a linha azul corresponde ao ensemble médio e as linhas cinzas às realizações.....	63
Figura 6.25: Corte de Água. A linha vermelha corresponde ao modelo verdadeiro, a linha azul corresponde ao ensemble médio e as linhas cinzas às realizações.....	64
Figura 6.26: Corte de Água. A linha vermelha corresponde ao modelo verdadeiro, a linha azul corresponde ao ensemble médio e as linhas cinzas às realizações.....	64
Figura 6.27: Corte de Água. A linha vermelha corresponde ao modelo verdadeiro, a linha azul corresponde ao ensemble médio e as linhas cinzas às realizações.....	65
Figura 6.28: Corte de Água. A linha vermelha corresponde ao modelo verdadeiro, a linha azul corresponde ao ensemble médio e as linhas cinzas às realizações.....	65

ANEXO A

Figura A.1: Mapeamento de resultados de experimentos a números em eixo real..	85
Figura A.2: FDA e pdf para a distribuição de Gauss.....	92

Lista de Tabelas

Tabela 6.1: Resumo das características do modelo do Caso 1.....	49
Tabela 6.2: Resultados obtidos após das simulações.....	55
Tabela 6.3: Resumo das características do modelo do Caso 2.....	56
Tabela 6.4: Resultados obtidos após das simulações.....	66
ANEXO A	
Tabela A.1: Operações entre eventos.....	78

SUMÁRIO

1.	Introdução.....	16
1.1	Introdução.....	16
1.2	Visão Geral da Dissertação.....	17
1.3	Objetivo.....	17
1.4	Trabalhos Prévios.....	18
2.	Reservatórios Petróleo.....	22
2.1	Introdução.....	22
2.2	Reservatórios de Petróleo.....	23
2.3	Principais Propriedades do Reservatório.....	23
2.3.1	Propriedades da Rocha.....	24
2.3.1.1	Porosidade (ϕ).....	24
2.3.1.2	Compressibilidade efetiva (c_f).....	25
2.3.1.3	Saturação (S).....	25
2.3.1.4	Permeabilidade (k).....	26
2.3.2	Propriedades dos Fluidos.....	27
2.3.2.1	Viscosidade (μ).....	27
2.3.2.2	Compressibilidade (c).....	28
2.3.2.3	Fator Volume Formação (B).....	28
2.4	Simulação de Reservatórios.....	29
2.4.1	Formulação Matemática.....	29
2.4.2	Equação de Fluxo para o Escoamento Multifásico.....	29
3.	Ajuste de Histórico	31
3.1	Introdução.....	31
3.2	Tipos de Ajuste ao Históricos.....	32
3.2.1	Ajuste ao Histórico Manual.....	32
3.2.2	Ajuste ao Histórico Automatizado.....	32
3.2.2.1	Recozimento Simulado (Annealing).....	33
3.2.2.2	Algoritmos Genéticos.....	34
3.2.2.3	Busca Scatter e Tabu.....	34
3.2.2.4	Algoritmo de Vizinhança.....	35
3.2.2.5	Filtro de Kalman.....	35
3.3	Diferenças entre o Ajuste ao Histórico Tradicional e EnKF.....	36
3.4	Estimativa de Parâmetros.....	37
3.4.1	Construir um Modelo Matemático.....	37

3.4.2	Definição de uma Função Objetivo.....	38
3.4.3	Limitações do Ajuste ao Histórico.....	39
4.	Assimilação de Dados com EnKF.....	41
4.1	Introdução.....	41
4.2	Revisão do Filtro de Kalman.....	42
4.3	Filtro de Kalman.....	43
4.4	Filtro de Kalman Baseado em Ensemble.....	43
4.4.1	Representação das Estatísticas de Erro.....	44
4.5	Análise do Esquema.....	45
4.6	Implementação Prática.....	48
4.6.1	Representação do Ensemble da Covariância.....	48
4.6.2	Perturbações das Medições.....	49
4.6.3	Equação de Análise.....	49
5.	Software e Metodologia.....	54
5.1	Introdução.....	54
5.2	Matlab Reservoir Simulation Toolbox (MRST).....	56
5.3	Métodos de Amostragem.....	57
5.3.1	Método de Quasi-Monte Carlo (QMC).....	58
5.3.2	Método da Tesselação de Voronoi (CVT).....	59
5.3.3	Hipercubo Latino (LHS).....	59
5.4	Ajuste de Histórico Baseado em EnKF.....	60
5.5	Fluxo de Trabalho de Ajuste ao Histórico Baseado em EnKF.....	61
5.6	Metodologia EnKF.....	62
5.7	Ajuste ao Histórico Tradicional Versus Baseado em EnKF.....	63
5.8	Recursos Computacionais.....	64
6.	Estudos de Caso.....	66
6.1	Introdução.....	66
6.2	Caso 1.....	67
6.2.1	Resultados do Caso 1.....	68
6.3	Caso 2.....	73
6.3.1	Resultados do Caso 2.....	75
7.	Conclusões e Trabalhos Futuros.....	86
7.1	Conclusão.....	86
7.2	Trabalhos Futuros.....	87
	Referências	88

Anexo A

A.	Fundamentos da Estatística e Probabilidades.....	95
A.1	Introdução.....	95
A.2	Espaço Amostral.....	96
A.2.1	Classificação de um Espaço Amostral.....	96
A.3	Evento.....	96
A.3.1	Combinação de Elementos.....	96
A.4	Álgebra de Conjuntos.....	96
A.5	Conceptos de Probabilidade.....	96
A.5.1	Definição clássica de probabilidade.....	96
A.5.2	Definição de Probabilidade como Frequência Relativa.....	97
A.5.2.1	Propriedades da Frequência Relativa.....	97
A.5.3	Definição Axiomática de Probabilidade.....	98
A.6	Eventos Independentes.....	99
A.7	Probabilidade Condicional.....	99
A.8	Probabilidade Condicional para Eventos Independentes.....	99
A.9	Probabilidade Total e Teorema de Bayes.....	100
A.9.1	Probabilidade Total.....	100
A.9.2	Teorema de Bayes.....	101
A.10	Espaço de Probabilidade.....	102
A.11	Variáveis Aleatórias.....	102
A.11.1	Tipos de Variáveis Aleatórias.....	102
A.11.1.1	Variáveis Aleatórias Discretas.....	102
A.11.1.2	Variáveis Aleatórias Contínuas.....	103
A.12	Distribuições Bivariadas.....	104
A.12.1	Distribuições Conjuntas Discretas.....	105
A.12.2	Distribuições Conjuntas Contínuas.....	105
A.13	Distribuições Marginais.....	105
A.13.1	Caso Discreto.....	106
A.13.2	Caso Contínuo.....	106
A.14	Esperança de uma Variável Aleatória.....	106
A.15	Variância de uma Variável Aleatória.....	107
A.16	Covariância.....	107
A.17	Distribuições de Probabilidade.....	108
A.18	Trabalhando com Amostras de uma Distribuição.....	109

CAPÍTULO 1

1. Introdução

1.1 Introdução

Os modelos para prever a produção de um reservatório de petróleo são baseados em dados sísmicos, testes de poços e outras informações onde há grandes incertezas associadas a esses dados. Em geral, o poder preditivo de tais modelos é limitado e tende a deteriorar-se ao longo do tempo. Para evitar isto os modelos têm que ser atualizados a intervalos regulares. A atualização usa novos dados do sistema de produção para atualizar o modelo. Este procedimento é conhecido como Ajuste ao Histórico e é um campo bem estabelecido em engenharia de reservatório. Este processo pode ser feito manualmente, mas demora muito tempo e demanda muita experiência dos engenheiros de reservatórios, além de não haver garantia que os reservatórios ajustados manualmente forneçam previsões satisfatórias da produção. Uma alternativa é realizar o ajuste de forma assistida por computador. O ajuste ao histórico é uma tarefa difícil, com modelos que contêm muitos parâmetros, sujeitos a elevada incerteza e com poucas medidas disponíveis. Normalmente, os modelos de reservatórios atuais contêm centenas de milhares de células¹, descritos por uma série de parâmetros dinâmicos e estáticos. Portanto, existe uma demanda por métodos mais automatizados que sejam computacionalmente eficientes. Um método popular para ajuste ao histórico é o Filtro de Kalman com Ensembles² (EnKF). Este método foi desenvolvido por G. Evensen em 1994 para uso em modelos oceânicos, enquanto que a primeira aplicação de ajuste ao histórico foi dada por G. Nævdal et al. em 2002. Este método tem mostrado resultados promissores em vários casos de ajuste ao histórico e é computacionalmente eficiente. Esta última propriedade possibilita um processo de atualização contínua para o modelo, o que significa que os dados de produção, assim que se tornam disponíveis, podem ser usados para atualizar o mo-

¹ **Células:** Células são subdomínios geométricos simples, hexaédricos ou quadriláteros, resultantes do procedimento de discretização pelo método das volumes finitos usado para a solução das equações diferenciais parciais que descrevem o escoamento multifásico nos reservatórios.

² **Ensemble:** Conjunto de realizações, onde cada realização representa um estado de um reservatório de petróleo, com suas propriedades petrofísicas estáticas (permeabilidade, porosidade) e dinâmicas (saturação, pressão) em cada célula. Na página 28 é dada uma definição mais formal de ensemble.

delo e tentar melhorar a sua capacidade preditiva. Todas essas qualidades fazem o EnKF adequado para a integração com o gerenciamento do reservatório.

1.2 Visão geral da dissertação

Os demais capítulos desse trabalho estão organizados da seguinte maneira.

No capítulo 2 são apresentados os conceitos fundamentais da simulação de reservatórios, mostrando a importância que tem no estudo de ajuste ao histórico. No capítulo 3, é apresentado o ajuste ao histórico, tipos de ajustes e métodos determinísticos e estocásticos que são aplicados no ajuste ao histórico. No capítulo 4, são descritos conceitos fundamentais da Estatística. Isso funciona como uma base para a compreensão da teoria do Filtro de Kalman apresentado no capítulo 5. No capítulo 5 é apresentado o algoritmo EnKF com base no Filtro de Kalman linear (KF). No capítulo 6, dá-se a conhecer a metodologia aplicada neste trabalho e além do software usado durante a simulação. No capítulo 7, são apresentados os casos sintéticos que foram estudados neste trabalho, um campo com 3 poços (1 injetor e 2 produtores) e outro campo com 5 poços (1 injetor e 4 produtores). Além disso, mostra-se os resultados obtidos com o método EnKF. Finalmente, as conclusões da dissertação e trabalhos futuros são resumidas no capítulo 8.

1.3 Objetivo

O objetivo deste trabalho é estudar uma técnica de assimilação de dados, o Filtro de Kalman com Ensembles, para realizar ajuste ao histórico automatizado de reservatórios de petróleo convencionais. Além disso, EnKF também fornece informações valiosas sobre o estado do reservatório após o processo de assimilação de dados. O foco principal foi estudar a estimação das propriedades petrofísicas do reservatório (permeabilidade e porosidade) através do tempo de produção e observar as variáveis de produção de óleo e corte de água (WCT) após realizar os ajustes, para realizar uma previsão adequada. O EnKF foi implementado com o Matlab, e as simulações de reservatório foram realizadas com o Matlab Reservoir Simulation Toolbox (MRST), que é uma caixa de ferramentas open-source de rotinas de Matlab para a realização de simulação e visualização do reservatório. Este simulador acadêmico é

desenvolvido no SINTEF³, um centro de pesquisas da Noruega, e seu código fonte pode ser acessado livremente do site <http://www.sintef.no/Projectweb/MRST/>. Os exemplos simulados mostraram bons resultados em termos de assimilação de dados e atualização de parâmetros do modelo após a aplicação de EnKF para o ajuste ao histórico.

1.4 Trabalhos prévios

Vários artigos sobre a aplicação com êxito do Filtro de Kalman com Ensembles na indústria do petróleo têm sido publicados. A pesquisa sobre Filtro de Kalman com Ensembles (EnKF) começou com Evensen (1994) e mais tarde por Houtekamer e Mitchell (1998), onde este método é classificado como de observações perturbadas EnKF e são essencialmente ensembles de sistemas de assimilação de dados. Outro tipo de EnKF é uma classe de filtros de raiz quadrada (Anderson, 2001, Whitaker e Hamill, 2002, Bishop, et al. 2001, e Tippett, et al., 2003). Whitaker e Hamill, (2002) mostraram que os filtros de raiz quadrada são mais precisos que os filtros de observação perturbadas por causa de erros de amostragem introduzidos pelas observações perturbadas com erros aleatórios. Os filtros de raiz quadrada discutidos por Tippett, et al. (2003) assimilam observações sequencialmente (Houtekamer e Mitchell, 1998) o que aumenta a eficiência ao evitar a inversão de matrizes grandes. Nævdal, et al. (2002) utilizaram o EnKF para a estimativa da permeabilidade num reservatório simples. Em seu trabalho o EnKF foi usado para atualizar parâmetros estáticos em modelos de reservatórios ajustando o campo permeabilidade. Nævdal, et al. (2005) utilizaram o EnKF para estimar a permeabilidade de todo o reservatório utilizando um modelo de reservatório 2D⁴ simplificado de um campo do Mar do Norte⁵. Naquele trabalho, o modelo de reservatório trifásico bidimensional⁶ foi atualizado continuamente, ajustando o campo permeabilidade estático e campos dinâmicos de saturação e pressão em cada passo de assimilação. Gu e Oliver (2005) analisaram o uso do EnKF para o estado e estimativa de parâmetros no caso de teste do reserva-

³ **SINTEF**: É a maior organização de pesquisa independente na Escandinávia que atua na geração de conhecimento, pesquisa e inovação, e desenvolver soluções tecnológicas que são trazidas para o uso prático. Realiza investigação multidisciplinar em tecnologia, medicina e ciências sociais.

⁴ **Modelo de reservatório 2D**: É um modelo simplificado de um reservatório, onde se consideram apenas duas dimensões geométricas e propriedades constantes ao longo de sua profundidade.

⁵ **Campo do Mar do Norte**: O campo de Buzzard, no Mar do Norte, é o maior e mais importante dos campos petrolíferos do Mar do Norte.

⁶ **Modelo de reservatório trifásico bidimensional**: É um modelo que considera o escoamento de três fases, água, óleo e gás.

tório PUNQ-S3⁷, um modelo de reservatório padrão usado como benchmark para técnicas de ajuste ao histórico pela indústria do petróleo. O EnKF foi usado para atualizar os campos de porosidade e permeabilidade, assim como os campos de saturação e pressão, e em seguida, foi aplicado o ajuste aos dados de produção das três fases. Brouwer, et al. (2004) investigaram o EnKF em uma malha fechado de ajuste com controle ótimo em um modelo de reservatório bidimensional simples. Nævdal, et al. (2006) usou a combinação de EnKF para a atualização do modelo contínua com otimização baseada em adjuntos de injeção de água para otimizar a estratégia de injeção de água. Gao, Zafari e Reynolds (2005), comparam o EnKF com outro método chamado máxima verossimilhança randomizado (RML). Uma diferença importante entre o EnKF e o RML é que o RML é baseado em adjuntos, o que requer que o simulador de reservatórios forneça derivadas, enquanto o EnKF não exige muito esforço no acoplamento com o modelo de reservatório. Zafari e Reynolds (2005) testaram o método em alguns problemas não lineares para validar o EnKF. Foi mostrado que o EnKF tem dificuldades em conseguir bons ajustes quando as distribuições da produção são multi-modais e que a hipótese Gaussiana das variáveis de produção embutida no EnKF é muito crítica, no sentido de que o afastamento da normalidade prejudica fortemente o ajuste. Wen e Chen (2005) apresentaram uma versão modificada do EnKF onde eles adicionaram um passo de confirmação, "confirming", para executar a simulação do reservatório usando os mais recentes parâmetros estáticos atualizados do modelo, para que os parâmetros estáticos e dinâmicos atualizados sejam sempre consistentes. No entanto, Zafari e Reynolds (2005) utilizaram um processo linear, para mostrar que o algoritmo com este passo de confirmação é inconsistente. Liu e Oliver (2005a e 2005b) usaram o EnKF tanto para o ajuste ao histórico quanto para a estimativa das fácies⁸ de um modelo de simulação de reservatórios. Eles descobriram que o método EnKF superou o método de minimização da norma da diferenças entre as curvas de produção observadas e previstas, baseado em gradientes, em termos de eficiência de computação e aplicabilidade. Lorentzen, et al. (2005) estudaram a robustez do EnKF executando dez casos do ensemble de diferentes condições iniciais onde encontraram que o EnKF é bem adequado para a previsão de incerteza. Skjervheim, et al. (2005) sugeriram um

⁷ **PUNQ-S3**: <http://www3.imperial.ac.uk/earthscienceandengineering/research/perm/punq-s3model>

⁸ **Fácies**: Denomina-se **fácies** o conjunto de rochas com determinadas características distintas dentro de um reservatório.

método baseado no EnKF para incorporar dados sísmicos 4D⁹ no modelo de atualização contínua. Eles mostraram que o EnKF poderia lidar com grande quantidade de dados sísmicos e teve um impacto positivo no ajuste do campo permeabilidade a partir de dados de medição muito ruidosos. Dong, et al. (2006) também identificaram uma conclusão semelhante à Skjervheim, et al. (2005), utilizando o EnKF para o ajuste ao histórico na produção e dados de sísmica 4D. Reinlie (2006) usou o EnKF tradicional e condicionando as informações de permeabilidade local ao redor do poço. Haugen, et al. (2006) usaram o EnKF para ajuste ao histórico em um caso de campo do Mar do Norte, usando dados de produção reais. Eles mostraram como EnKF pode melhorar os parâmetros do modelo para o ajuste ao histórico e também discutiram a atualização dos estados do reservatório. Eles mostraram resultados promissores, mas concluíram que mais trabalhos devem ser feitos para ser capaz de estimar outros parâmetros de reservatório utilizando o EnKF. Park e Choe (2006) estudaram o baixo valor da estimativa da covariância de erro depois de alguns períodos de ajuste ao histórico e do número versus qualidade dos dados de medição. Eles encontraram que as medidas de saturação de água perto da saturação de água irreduzível ou a saturação de óleo residual não são sensíveis aos parâmetros estáticos, como a permeabilidade e porosidade, do reservatório e podem ser ignoradas. Evensen, et al. (2007) também mostraram como o EnKF pode ser utilizado para o ajuste ao histórico de um modelo para um reservatório do Mar do Norte. Eles investigaram mais parâmetros, como a WOC e GOC¹⁰, multiplicadores de transmissividade verticais e multiplicadores de transmissividade de falhas. O interesse e frequência de uso de EnKF como uma técnica de ajuste ao histórico aumentou significativamente e muitas aplicações recentes de EnKF para o ajuste ao histórico foram dadas. Saad e Ghanem (2009) desenvolveram uma abordagem para a assimilação de dados sequencial para representar variáveis aleatórias com a expansão em Caos Polinomial (Wiener, 1938) e adotando o esquema de análise da EnKF. Zeng et al. (2010) propuseram colocação probabilística baseada em filtro de Kalman (PCKF). Emerick e Reynolds (2011) usam o EnKF com Localização de Covariância para o ajuste ao histórico, e posteriormente (Emerick et al., 2012) combinaram o EnKF com

⁹ **Sísmica 4D:** Usa-se este termo quando testes sísmicos são realizados periodicamente ao longo da vida do reservatório e seus resultados usados para inferir a distribuição dos fluidos dentro dele.

¹⁰ **WOC e GOC:** Contato Água – Óleo e Contato Gás – Óleo, termos usuais da indústria do petróleo que indicam qual é a profundidade na qual o aquífero se encontra e qual a profundidade na qual a capa de gás do reservatório começa.

MCMC para realizar o ajuste ao histórico de produção. Tavakoli et al. (2013) utilizam o EnKF para atualizar modelos estocásticos. Trani e Leeuwenburgh (2013), usam o EnKF no ajuste de histórico sísmico dos fluidos frontais.

CAPÍTULO 2

2. Reservatórios de Petróleo

2.1 Introdução

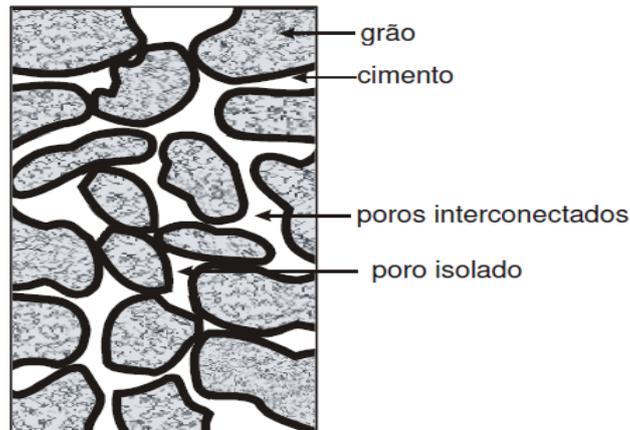
Realizar ajuste ao histórico em um reservatório de petróleo significa calibrar o modelo usado na simulação variando as propriedades de rocha e fluido com o objetivo de aproximar as curvas de produção previstas e observadas. Portanto, para obter sucesso na calibração é necessário entender como funciona um reservatório, quais são as suas propriedades, como elas interferem no escoamento dos fluidos e quais são as equações que as governam.

Neste capítulo, descrevemos alguns conceitos fundamentais de engenharia de reservatórios, que são essenciais para a discussão de qualquer aplicação da simulação de reservatórios, e introduzimos algumas terminologias comuns usadas em engenharia de reservatórios. Começamos com algumas definições formais do reservatório de petróleo, em seguida, descrevemos brevemente as equações governantes.

2.2 Reservatórios de Petróleo

Um reservatório de petróleo é um aglomerado de rochas sedimentares, geralmente arenitos ou carbonatos, portador de hidrocarbonetos (óleo e gás), água e outros fluidos. A Figura 2.1 ilustra uma seção transversal de uma rocha-reservatório com os seus principais componentes.

Figura 2.1: Seção transversal de uma amostra de rocha de reservatório.



(Rosa *et al.*, 2006)

Os fluidos contidos na rocha reservatório (água, óleo e gás) ficam armazenados sob pressão nos poros da rocha. Quando um poço é aberto, devido à diferença de pressões, esses fluidos escoam da rocha-reservatório para o poço e, em seguida, para a unidade de produção na superfície onde são tratados. Este processo é conhecido como recuperação primária. Depois de um certo tempo de produção, normalmente é iniciado um procedimento de recuperação secundária, no qual é injetada água em poços injetores¹¹, para manter a pressão no reservatório e deslocar o óleo na direção dos poços produtores.

2.3 Principais Propriedades do Reservatório

2.3.1 Propriedades da Rocha

As propriedades da rocha variam bastante em função ao histórico de deposição e compactação dos seus sedimentos, como o tamanho e seleção dos grãos, grau de cimentação, histórico de pressão e temperatura, etc. Raramente, uma rocha reservatório é homogênea ao longo de toda sua extensão, sendo necessário realizar estudos sobre a variação de suas propriedades. As propriedades petrofísicas tem uma variabilidade inerente aos processos de formação da subsuperfície, e está sujeita a grandes incertezas. Apenas em poucos pontos do reservatório as propriedades são conhecidas com alguma confiança, em função de dados oriundos de poços exploratórios e de correlações com outros tipos de informações, como dados sísmicos e

¹¹ **Poços injetores:** É um poço usado para injetar fluidos no reservatório visando melhorar a recuperação de petróleo, de gás natural.

afloramentos. A escolha natural para a descrição destas propriedades é através de campos estocásticos e as ferramentas usadas são baseadas em geoestatística.

2.3.1.1 Porosidade (ϕ)

A porosidade é uma medida dos espaços vazios na rocha, e mede a capacidade de armazenar fluidos. Ela é representada pela letra grega ϕ e definida como:

$$\phi = \frac{V_p}{V_T} \quad (2.1)$$

onde V_T é um volume representativo de rocha¹² e V_p é o volume dos poros (vazios) interconectados ou não. A porosidade é expressa, geralmente, em porcentagem.

A porosidade efetiva (ϕ_{ef}) é a razão entre o volume de poros conectados e o volume representativo. Essa porosidade é a mais importante para o estudo de reservatórios, pois o escoamento dos fluidos do reservatório se dá através dos poros conectados. (Rosa *et al.*, 2006).

2.3.1.2 Compressibilidade Efetiva (c_f)

A compressibilidade efetiva expressa a variação do volume de poros em função da variação de pressão ao qual a rocha está submetida:

$$c_f = \frac{1}{V_p} \frac{\partial V_p}{\partial p} \quad (2.2)$$

onde p é a pressão. A compressibilidade é medida, geralmente, em $(kg / cm^2)^{-1}$.

¹² **Volume representativo de rocha:** O volume representativo é um conceito associado à escala de variação das propriedades da rocha. Ele é pequeno o suficiente para que as propriedades de interesse possam ser consideradas constantes, mas não tão pequeno que as propriedades, calculadas com alguma média espacial, variem excessiva ou mesmo descontinuamente em torno da sua vizinhança. Por exemplo, se um volume considerado estivesse contido apenas em um poro ou em um grão maciço do material, a porosidade como definida acima não faria sentido, pois variaria entre 1 ou 0, sendo que nenhum destes dois valores representa a capacidade de um pequeno volume desta rocha armazenar fluidos. Conforme o volume considerado aumenta, o valor da porosidade calculado como em (2.1) estabiliza-se em torno de um valor fixo. O menor volume para o qual isto ocorre é o volume representativo.

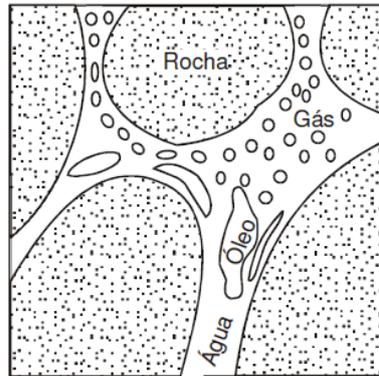
Considerando o volume total constante e substituindo a Equação (2.1) na Equação (2.2), e obtemos a compressibilidade da rocha:

$$c_R = \frac{1}{\phi} \frac{\partial \phi}{\partial p} \quad (2.3)$$

2.3.1.3 Saturação (S)

A saturação é a fração ou porcentagem do volume de poro ocupado por um fluido em particular (óleo, gás, água). A Figura 2 mostra um corte transversal de uma rocha-reservatório com presença de fluidos, água, óleo e gás nos poros.

Figura 2.2: Rocha-reservatório contendo três fluidos: água, óleo e gás.



(Rosa *et al.*, 2006)

$$S_f = \frac{V_f}{V_p} \quad (2.4)$$

Onde S_f é a saturação do fluido (água, óleo e gás), V_f o volume que ocupa o fluido e V_p o volume poroso. A saturação é expressa, geralmente, em porcentagem.

2.3.1.4 Permeabilidade (k)

A permeabilidade absoluta é uma medida que indica o quão fácil é a passagem de fluidos pelo meio poroso. Quanto maior for o seu valor, maior será a vazão dos fluidos quando o meio é submetido a uma mesma diferença de pressões. O escoamento em meios rochosos acontece através de uma infinidade de microcanais formados pelos poros, distribuído em direções e com conexões completamente aleatórias.

Considerações experimentais e teóricas permitem que este escoamento possa ser descrito por velocidades volumétricas ou velocidades médias, sem que seja necessária a descrição detalhada do escoamento nos poros. No entanto, o que se verifica tanto experimental quanto teoricamente é que uma diferença de pressões em uma direção pode causar escoamento não só nesta direção, mas também em direções perpendiculares a ela. Para quantificar este fenômeno, a permeabilidade pode ser representada por uma matriz quadrada, onde os coeficientes da diagonal relacionam a velocidade média com o gradiente de pressão na mesma direção e os termos fora da diagonal relacionam o gradiente de pressão com velocidades de escoamento perpendiculares a ele. Outras considerações teóricas mostram que a permeabilidade atende aos critérios para ser considerada um tensor de segunda ordem. Em muitas aplicações de simulações de reservatórios usa-se uma simplificação considerando a permeabilidade como um escalar ou tensor diagonal, isto é, que os gradientes de pressão só produzem escoamento na sua própria direção. Neste caso, como há três direções espaciais, a permeabilidade pode ser representada por um vetor ou uma matriz diagonal. A simplificação final ocorre quando se considera que a permeabilidade é a mesma nas três direções espaciais, e portanto pode ser representada por um único escalar. No restante deste capítulo a permeabilidade será considerada um escalar. Geralmente, é expressa em miliDarcy (mD) (Rosa et al., 2006).

Em escoamentos multifásicos, isto é, quando consideramos três fases diferentes coexistindo no reservatório, a fase aquosa (água e outros componentes dissolvidos), oleica (hidrocarbonetos na fase líquida) e gasosa (hidrocarbonetos na fase gasosa), a presença de cada fase influencia o escoamento das outras, e portanto a permeabilidade deve ser determinada para cada fase, sendo chamada neste caso de permeabilidade efetiva (k_o, k_g, k_w). A razão entre a permeabilidade efetiva de cada fluido e a permeabilidade absoluta do reservatório é chamada de permeabilidade relativa (k_{ro}, k_{rg}, k_{rw}). A permeabilidade deve ser medida experimentalmente, porém, há fórmulas analíticas, com parâmetros ajustáveis em função de propriedades das rochas e das condições de escoamento, que podem ser usadas em simuladores de reservatórios e fornecem resultados adequados em algumas situações.

2.3.2 Propriedades dos Fluidos

As propriedades dos fluidos variam conforme a sua composição, pressão e temperatura ao qual estão submetidos e à composição da própria rocha, devido às interações rocha-fluido. Elas são medidas através de amostras coletadas em testes específicos (viscosidade, PVT, etc.).

2.3.2.1 Viscosidade (μ)

A viscosidade de um fluido é a propriedade que mede a resistência ao cisalhamento. É definida como a razão entre a tensão cisalhante e a gradiente de velocidades. Normalmente é medida em centipoise (cp) na engenharia de reservatórios

2.3.2.2 Compressibilidade (c)

Como na rocha reservatório, é importante conhecer a variação de volume de cada fluido em função da variação de pressão. A determinação da compressibilidade o fluido é análoga à da rocha e usa a mesma unidade.

2.3.2.3 Fator Volume de Formação (B)

O fator volume de formação é a razão entre a massa específica de um fluido nas condições de superfície (1.033 kg/cm^2 e 20°C) e nas condições de pressão e temperatura do reservatório.

$$B = \frac{\rho_{std}}{\rho_{res}} \quad (2.5)$$

Onde ρ_{std} é a massa específica nas condições de superfície e ρ_{res} é a massa específica nas condições de reservatório.

2.4 Simulação de Reservatórios

Nos itens anteriores, foram apresentadas algumas das propriedades do sistema rocha-fluido. Nesta seção, será apresentada uma formulação matemática que embasa o modelo numérico utilizado na simulação de reservatórios com a finalidade de mostrar a mecânica que envolve o escoamento. Dentre os modelos aplicados na indús-

tria, o modelo Black Oil (Rosa et al., 2006) é adotado neste trabalho devido a sua simplicidade no tratamento do problema e boa aplicabilidade a problemas realistas.

Neste modelo, é considerado um escoamento trifásico e com três componentes (água, óleo e gás) onde:

- ❖ A água e o óleo são imiscíveis;
- ❖ O gás pode existir na forma livre ou dissolvida no óleo. A solubilidade do gás na água é desprezada;
- ❖ A temperatura do reservatório é assumida constante ao longo do tempo;
- ❖ Os fluidos estão em equilíbrio termodinâmico dentro do reservatório (Ertekin et al., 2001).

2.4.1 Formulação Matemática

Para o desenvolvimento de um simulador numérico, o primeiro passo é o estudo da física do problema e desenvolvimento das equações que a governam o comportamento do escoamento no meio poroso. A engenharia de reservatório, no que concerne a técnicas de simulação computacional, é uma disciplina conservadora. Nos simuladores comerciais, em geral, são usadas técnicas numéricas baseadas em coordenadas cartesianas e malhas estruturadas¹³, devido principalmente à familiaridade da indústria com as mesmas e a confiabilidade percebida dos atuais simuladores que foram assim construídos. As equações que regem o escoamento multifásico em meios porosos podem, no entanto, ser escritas também em outros sistemas de coordenadas e para outros tipos de malhas.

2.4.2 Equação de Fluxo para o Escoamento Trifásico

A formulação da equação de fluxo adotada no simulador de escoamento empregado neste trabalho é baseada na conservação de massa para cada fase (líquida ou gasosa) e o modelo Black Oil o gás se encontra em duas formas: uma livre, formando a fase gasosa, e a outra solubilizada no óleo, compondo a fase de hidrocarbonetos líquidos. A modelagem matemática do problema parte do princípio da conservação de massa, que fornece uma equação de conservação de massa para cada fase, e

¹³ **Malhas estruturadas:** São malhas localmente ortogonais que obedece a um sistema de coordenadas usualmente formadas por quadriláteros (2D) e prismas retangulares geradas facilmente (3D).

da lei de Darcy (Ertekin et al., 2001), uma lei fenomenológica que relaciona a velocidade média de cada fase com os gradientes de pressão. A conservação de massa e a Lei de Darcy são expressas em termos de equações diferenciais parciais, que podem ser resolvidas com uma série de técnicas numéricas, a técnica mais convencional é das diferenças finitas baseadas em volumes de controle, que quando aplicadas à conservação de massa e Lei de Darcy, generalizada para escoamentos multifásicos, leva às Equações para as fases oleica, aquosa e gasosa de 2.6 a 2.8.

$$\begin{aligned} & \frac{\partial}{\partial x} \left(A_x k_x \frac{k_{ro}}{B_o \mu_o} \left(\frac{\partial p_o}{\partial x} - \gamma_o \frac{\partial Z}{\partial x} \right) \right) \Delta x + \frac{\partial}{\partial y} \left(A_y k_y \frac{k_{ro}}{B_o \mu_o} \left(\frac{\partial p_o}{\partial y} - \gamma_o \frac{\partial Z}{\partial y} \right) \right) \Delta y \\ & + \frac{\partial}{\partial z} \left(A_z k_z \frac{k_{ro}}{B_o \mu_o} \left(\frac{\partial p_o}{\partial z} - \gamma_o \frac{\partial Z}{\partial z} \right) \right) \Delta z + q_{o,std} = V_b \frac{\partial}{\partial t} \left(\frac{\phi S_o}{B_o} \right) \end{aligned} \quad (2.6)$$

$$\begin{aligned} & \frac{\partial}{\partial x} \left(A_x k_x \frac{k_{rw}}{B_w \mu_w} \left(\frac{\partial p_w}{\partial x} - \gamma_w \frac{\partial Z}{\partial x} \right) \right) \Delta x + \frac{\partial}{\partial y} \left(A_y k_y \frac{k_{rw}}{B_w \mu_w} \left(\frac{\partial p_w}{\partial y} - \gamma_w \frac{\partial Z}{\partial y} \right) \right) \Delta y \\ & + \frac{\partial}{\partial z} \left(A_z k_z \frac{k_{rw}}{B_w \mu_w} \left(\frac{\partial p_w}{\partial z} - \gamma_w \frac{\partial Z}{\partial z} \right) \right) \Delta z + q_{w,std} = V_b \frac{\partial}{\partial t} \left(\frac{\phi S_w}{B_w} \right) \end{aligned} \quad (2.7)$$

$$\begin{aligned} & \frac{\partial}{\partial x} \left(A_x k_x \frac{k_{rg}}{B_g \mu_g} \left(\frac{\partial p_g}{\partial x} - \gamma_g \frac{\partial Z}{\partial x} \right) + A_x k_x \frac{k_{ro} R_s}{B_o \mu_o} \left(\frac{\partial p_o}{\partial x} - \gamma_o \frac{\partial Z}{\partial x} \right) \right) \Delta x \\ & + \frac{\partial}{\partial y} \left(A_y k_y \frac{k_{rg}}{B_g \mu_g} \left(\frac{\partial p_g}{\partial y} - \gamma_g \frac{\partial Z}{\partial y} \right) + A_y k_y \frac{k_{ro} R_s}{B_o \mu_o} \left(\frac{\partial p_o}{\partial y} - \gamma_o \frac{\partial Z}{\partial y} \right) \right) \Delta y \\ & + \frac{\partial}{\partial z} \left(A_z k_z \frac{k_{rg}}{B_g \mu_g} \left(\frac{\partial p_g}{\partial z} - \gamma_g \frac{\partial Z}{\partial z} \right) + A_z k_z \frac{k_{ro} R_s}{B_o \mu_o} \left(\frac{\partial p_o}{\partial z} - \gamma_o \frac{\partial Z}{\partial z} \right) \right) \Delta z \\ & + q_{g,std} = \frac{\partial}{\partial t} \left(\frac{\phi S_g}{B_g} + \frac{\phi R_s S_o}{B_o} \right) \end{aligned} \quad (2.8)$$

Nestas equações, A é a área de uma face de um volume de controle, também chamado de célula computacional ou apenas célula, k é a permeabilidade, k_r é a permeabilidade relativa, p é a pressão, B é o fator volume de formação, μ é a viscosidade, γ é o peso específico, Z é a altura do volume de controle em relação a uma altura de referência, x, y, z são as coordenadas cartesianas, ϕ é a porosidade, S é a saturação, q é a vazão volumétrica e R_s é a razão de solubilidade do gás no óleo.

Essas são as equações que governam o fluxo para todos os fluidos que compõem o reservatório. Elas foram obtidas através da combinação das equações de conservação de massa, de estado e da lei de Darcy, porém não são suficientes para completar a descrição do fluido em um meio poroso. Para isso são necessárias relações adicionais de restrição de saturação de cada fase e expressões para as pressões capilares como funções da saturação de fases (Ertekin et al., 2001). As equações que representam estas restrições estão escritas abaixo:

$$S_o + S_g + S_w = 1 \quad (2.9)$$

$$p_w = p_o - p_w = f(S_w) \quad (2.10)$$

$$p_{cgo} = p_g - p_o = f(S_g) \quad (2.11)$$

Nestas equações, os subscritos o , w e g representam as fases oleica, aquosa e gasosa, respectivamente, e f é uma função dada que relaciona pressões capilares e saturações, e deve ser determinada experimentalmente. As equações apresentadas acima junto com as condições de contorno do reservatório descrevem a formulação *Black Oil*. Observe que o número de incógnitas para uma única célula $(p_o, p_w, p_g, S_o, S_w, S_g)$ coincide com o número de equações. Todas as outras variáveis são medidas ou estimadas, ou seja, são parâmetros de entrada na simulação de reservatórios e compõem o domínio do espaço de busca no ajuste ao histórico.

CAPÍTULO 3

3. Ajuste ao Histórico

3.1 Introdução

O processo de modificação de parâmetros do modelo de reservatório usado em uma simulação para tentar fazer com que as previsões de taxas de injeção e produção, pressões, saturações e outras variáveis se aproximem dos valores observados quando o campo entra em produção é chamado de ajuste ao histórico.

O procedimento consiste no ajuste de um ou mais modelos matemáticos aos dados históricos, até que se obtenha uma discrepância aceitável entre os dados observados e os resultados previstos pelo modelo. Durante o processo, os parâmetros do reservatório são alterados para que se obtenha, por exemplo, a menor diferença possível entre uma curva de produção de óleo prevista e observada, em alguma norma conveniente. O ajuste ao histórico é, portanto, um processo de otimização, onde se procura minimizar as discrepâncias entre os resultados do modelo de reservatório e as observações feitas no campo. (Rosa et al., 2006).

O histórico de dados de produção contém informações do reservatório que podem ser usadas para reduzir a incerteza na previsão do desempenho futuro. Avaliação da incerteza, que pode ser feita quando o ajuste ao histórico é feito de forma estocástica, é importante por permitir uma avaliação dos riscos associados à produção daquele reservatório. Espera-se também que a assimilação ao longo do tempo de dados observados de produção permita que haja uma redução da incerteza associada ao modelo do reservatório. Neste trabalho, o Filtro de Kalman com Ensembles, que é um método estocástico para ajuste ao histórico automático tem com resultado útil uma avaliação da incerteza associada ao modelo do reservatório. Neste caso, o modelo do reservatório deve ser entendido como todo o conjunto de realizações petrofísicas que são atualizadas ao longo do tempo durante o processo de ajuste. A previsão da produção é feita para cada membro do conjunto e estatísticas de interesse podem ser calculadas sobre o ensemble para qualquer instante no futuro, e podem ser usadas para quantificar a incerteza no modelo.

3.2 Tipos de Ajuste ao Histórico

Na engenharia de petróleo classifica-se o ajuste ao histórico em manual, onde as modificações no modelo são feitas apenas de acordo com a experiência dos profissionais envolvidos e automatizado ou assistido por computador, onde programas de computador que implementam algum algoritmo matemático ditam as modificações no modelo de forma a aproximar a previsão do modelo aos dados observados.

3.2.1 Ajuste ao Histórico Manual

O ajuste ao histórico manual se caracteriza por ser um processo de tentativa e erro, que depende essencialmente da experiência e intuição dos engenheiros de reservatórios e geólogos. Normalmente, é mais trabalhoso do ponto de vista de esforço do profissional envolvido no processo, pois todas as tarefas, tais como alterações do modelo de simulação, execução das simulações, análise dos resultados, dentre outras, são realizadas de forma manual e exigem mais interação com o processo, tomando mais tempo de trabalho. Além disto, as decisões sobre quanto modificar os parâmetros de projeto são tomadas sem um grande embasamento matemático. Muitas vezes, o conhecimento sobre a geologia do campo é ignorado na busca de um melhor ajuste entre curvas previstas e observadas. Este procedimento é oneroso, demorado e propenso a erros, e os ajustes manuais tendem a não serem confiáveis para previsões de longo prazo.

3.2.2 Ajuste ao Histórico Automatizado

O ajuste ao histórico automatizado tende a ser muito mais rápido do que o procedimento manual. Ele pode tratar um grande número de diferentes parâmetros de ajuste, em contraste ao procedimento manual, e também pode tratar com muito mais facilidade um grande número de poços, campos de grande dimensão e complexidade geológica. O ajuste ao histórico automático também pode fornecer resultados mais confiáveis nos casos onde as litologias são mais complexas, com grande heterogeneidade. A ideia geral do ajuste ao histórico automático é começar a partir de uma estimativa inicial de um parâmetro e melhorá-la por meio da assimilação de dados de campo em um laço automático, segundo algum mecanismo de otimização. As alterações dos parâmetros do modelo do reservatório são feitas diretamente por programas de computadores, para minimizar uma função objetivo, por exemplo, a

integral dos quadrados das diferenças entre os dados simulados e observados para algumas variáveis de interesse. Usualmente, o procedimento não é completamente automatizado, a experiência dos profissionais de engenharia de reservatórios é aproveitada impondo restrições aos modelos admissíveis no processo de otimização, como a existência de características estruturais geológicas importantes. O ajuste ao histórico automatizado pode ser feito por etapas, e pode incluir etapas de interferência manual direta no processo.

Várias investigações foram feitas sobre muitos métodos de ajuste ao histórico automático nas últimas décadas no setor de petróleo. Estes métodos podem ser classificados como:

a) Algoritmos Determinísticos

Algoritmos determinísticos usam abordagens tradicionais de otimização e usam um modelo de reservatório ideal como um modelo que torna uma determinada função objetivo estacionária, sujeita a um conjunto de restrições. Nas implementações usuais, o gradiente da função objetivo é calculado, alguma direção de busca de otimização é então determinada e um algoritmo do tipo “line search” é aplicado ao longo desta direção (Liang, 2007). Os algoritmos determinísticos normalmente minimizam a diferença entre as medidas observadas no campo e os resultados do simulador de reservatórios, dentro de um processo iterativo que inclui entre suas principais etapas:

- ❖ Execução do simulador de reservatório para o período completo de ajuste ao histórico;
- ❖ Avaliação da função objetivo;
- ❖ Atualização das propriedades petrofísicas do modelo.

Alguns algoritmos que são comumente usados para o ajuste ao histórico baseado em algoritmos de otimização determinística são listados a seguir. (Landa, 1979 e Liang, 2007):

❖ Algoritmos Baseados em Gradiente:

- Método de Gauss - Newton (GN);
- Algoritmo de Levenberg – Marquardt;

- Método de Gradientes Conjugados;
- Método Quase –Newton;
- Método de Broyden, Fletcher, Goldfarb e Shannon (BFGS);

b) Algoritmos Estocásticos

Os algoritmos estocásticos podem consumir uma quantidade considerável de tempo computacional, em comparação com um algoritmo determinístico, mas devido ao rápido desenvolvimento da capacidade de memória do computador e velocidade de processamento, algoritmos estocásticos estão recebendo cada vez mais atenção.

Algoritmos estocásticos têm três principais vantagens diretas:

- ❖ A abordagem estocástica gera um conjunto de modelos de reservatório e, portanto, é mais apropriada para os problemas de ajuste ao histórico que consideram incertezas;
- ❖ A quantificação de incertezas da previsão de desempenho é relativamente simples e direta.
- ❖ Algoritmos estocásticos podem, teoricamente, atingir o ótimo global.

Alguns importantes algoritmos estocásticos para otimização são. (Liang, 2007):

- ❖ **Algoritmos Baseados em Metaheurísticas:**
 - Recozimento simulado (Simulated Annealing);
 - Algoritmos Genéticos;
 - Buscas Scatter e Tabu;
 - Algoritmo de Busca na Vizinhança (Neighborhood Search);
 - Algoritmos baseados em Ensembles (Filtro de Kalman);

3.2.2.1 Recozimento Simulado

Recozimento simulado (simulated annealing) é um algoritmo probabilístico para problemas de otimização global, que normalmente fornece uma boa aproximação para o ótimo global de uma função dada em um grande espaço de busca. Para certos problemas o recozimento simulado (annealing) encontra uma solução aceitável em um período de tempo razoável. O nome de recozimento (annealing) vem da metalur-

gia, de uma técnica que envolve aquecimento e resfriamento controlado de material para aumentar o tamanho dos seus cristais e reduzir os seus defeitos. Recozimento simulado é uma generalização de um método de Monte Carlo para analisar as equações de estado e estados congelados de sistemas n-corpos (Metropolis, et al. 1953). Panda e Lake (1993), e Portellaand e Prais (1999) aplicaram a técnica de recozimento simulado para o ajuste ao histórico em reservatórios.

3.2.2.2 Algoritmos Genéticos

Algoritmos Genéticos (GA) são algoritmos de busca heurística adaptativa que tem como premissa as ideias evolucionárias de seleção natural e genética. O método envolve uma população de cromossomos, onde cada cromossomo é normalmente codificado como uma sequência de bits e processado por "seleção natural" de uma geração para a próxima geração, associada com a herança, mutação, seleção e cruzamento. As soluções viáveis são chamadas indivíduos. O conceito básico dos GA é o de simular processos do ambiente natural necessários para a evolução, especificamente aqueles que seguem os princípios estabelecidos por Charles Darwin da sobrevivência do mais apto. Como tal, eles representam uma exploração inteligente dentro de um espaço de busca definido para resolver um problema. Algoritmos genéticos têm sido amplamente estudados, experimentados e aplicados em muitos campos do mundo da engenharia (Holland, 1975). Sua aplicação direta a problemas realistas de grande porte é difícil, devido ao alto custo computacional inerente à sua convergência lenta. Os algoritmos genéticos podem resolver bem problemas de grande escala com muitos mínimos locais.

3.2.2.3 Busca Scatter e Tabu

Busca Scatter e Tabu são métodos heurísticos, estabelecidos por Fred Glover na Universidade do Colorado (Glover, 1977). A Busca Scatter (Glover, 1994) opera em um conjunto de pontos chamados de pontos de referência que constituem boas soluções obtidas a partir de soluções anteriores. A abordagem sistematicamente gera combinações lineares de pontos de referência para criar novos pontos, cada um dos quais está mapeado e associado a pontos possíveis. A busca Tabu (Glover e Laguna, 1997) é uma orientação inteligente para o processo de busca. Na indústria do petróleo, para a exploração e produção de petróleo, um otimizador contendo busca

de scatter, busca tabu e redes neurais foi usado em alguns exemplos simples (April, et al. 2003a) e também utilizado para gerenciamento de portfólio (April, et al. 2003b). Sousa e colaboradores (2006) utilizaram a busca scatter para casos simples de ajuste ao histórico e Cullick, et al. (2006) combinaram a busca scatter com um modelo substituto baseado em rede neural não-linear para problemas de ajuste ao histórico com um pequeno número de parâmetros desconhecidos.

3.2.2.4 Algoritmo de Vizinhança

O algoritmo de vizinhança é um algoritmo de otimização estocástico inicialmente empregado para problemas de inversão sísmica (Sambridge, 1999a, 1999b), que agora é usado para ajuste ao histórico (Christie, et al 2002 e 2006;. Suzuki e Caers, 2006). É uma técnica popular de inversão por busca direta, semelhante ao recozimento simulado (annealing) e algoritmos genéticos.

3.2.2.5 Filtro de Kalman

O Filtro de Kalman (KF) foi desenvolvido por Kalman (1960) e é o esquema de assimilação de dados sequencial mais conhecido. O Filtro de Kalman é um procedimento computacional que proporciona uma boa eficiência para estimar o estado de um processo, de uma maneira que minimiza a média do erro quadrático. Ele estima o estado de um sistema dinâmico linear a partir de uma série de medições ruidosas. Em teoria, o Filtro de Kalman Estendido (EKF) é a versão não-linear do Filtro de Kalman que lineariza em torno da média e covariância atuais. A desvantagem do EKF é que ele não é eficiente no caso de sistemas altamente não lineares (Emerick, 2012).

Muitas variações de Filtro de Kalman têm sido propostos para lidar com estas dificuldades. Entre eles o Filtro de Kalman com Ensembles (EnKF) é promissor. É um filtro recursivo, o que significa que quando novos dados chegam, ele pode apenas calcular o próximo passo e não há necessidade de executar uma nova otimização completa e é adequado para problemas com um grande número de variáveis, tais como discretizações de equações diferenciais parciais em modelos geofísicos. O EnKF originou-se como uma versão do filtro de Kalman para grandes problemas em que a matriz de covariância é calculada aproximadamente a partir das realizações que compõem o ensemble, e agora é um importante componente de assimilação de

dados. Devido à formulação simples e fácil implementação, o EnKF ganhou popularidade para previsão do tempo, oceanografia, hidrologia e engenharia de petróleo. Nesta dissertação, o EnKF é usado para ajuste ao histórico automático e assimilação seqüencial de dados.

3.3 Diferenças entre o Ajuste ao Histórico Tradicional e EnKF

O ajuste ao histórico tradicional atualiza os parâmetros estáticos do reservatório como porosidade e permeabilidade até que um ajuste aceitável seja alcançado executando o modelo de forma iterativa (Liang, 2007), e suas principais características são:

- ❖ Repetidas simulações de reservatório para todo o histórico da produção;
- ❖ Requer o cálculo de coeficientes de sensibilidade, que são as derivadas da função objetivo em relação às variáveis de otimização;
- ❖ Normalmente não é totalmente automatizado;
- ❖ O ajuste é repetido com todo o histórico de produção quando novos dados estiverem disponíveis;
- ❖ Não é adequado para a atualização do modelo de reservatório em tempo real;
- ❖ Não incorpora facilmente a avaliação da incerteza.

O EnKF atualiza o modelo de reservatório sequencialmente para ambos os parâmetros estáticos como porosidade e permeabilidade e parâmetros dinâmicos como pressão e saturação, e suas principais características são:

- ❖ Adequado para a atualização do modelo não-linear de simulação de reservatório de grande porte, com centenas de milhares ou milhões de células;
- ❖ Uma simulação independente é realizada para cada membro do ensemble o que permite uma implementação em computação distribuída trivial;
- ❖ Não há necessidade cálculos de derivadas, pois não são usados algoritmos de otimização sequencial;
- ❖ Pode ser totalmente automatizado;
- ❖ Os dados de produção são assimilados sequencialmente no tempo;
- ❖ Os membros do ensemble são atualizados sequencialmente no tempo e refletem os últimos dados assimilados;

- ❖ A previsão de incerteza é sempre atualizada e obtida diretamente dos membros do ensemble.

3.4 Estimativa de Parâmetros

Estimação é uma aproximação calculada para um resultado que é utilizável se os dados de entrada são incompletos ou incertos. Em matemática, a aproximação ou estimação normalmente significa encontrar limites superiores ou inferiores de uma quantidade que não pode ser facilmente calculada com precisão. Em um enfoque onde a quantificação de incertezas é importante, no entanto, estamos mais interessados em determinar estatísticas, como médias, variâncias e percentis, das variáveis estimadas.

A estimação de parâmetros é uma aplicação útil do ajuste ao histórico para a determinação das propriedades do reservatório. Todos os métodos de estimação de parâmetros têm três etapas principais:

- ❖ Construção de um modelo matemático.
- ❖ Definição de uma função-objetivo.
- ❖ Escolha de um algoritmo de minimização (otimização).

3.4.1 Construção de um Modelo Matemático

Um modelo matemático descreve o comportamento do sistema físico sob diferentes condições, com uma precisão adequada. O problema de calcular a resposta do modelo matemático a um conjunto de condições que representam as características de operação do sistema é referido como o problema para frente.

As seguintes leis fundamentais são relevantes para a dinâmica do reservatório, e são usadas para a construção do modelo matemático:

- ❖ Lei de conservação de massa.
- ❖ Lei de Darcy.
- ❖ Equações de Estado.

3.4.2 Definição de uma Função-Objetivo

A função-objetivo mede a discrepância entre os dados observados e a resposta do simulador para um determinado conjunto de parâmetros. Existem três formas básicas para definir a função-objetivo.

a) Formulação por Mínimos Quadrados

$$F = (d^{obs} - d^{sim})^T (d^{obs} - d^{sim}) \quad (3.12)$$

onde,

- F é a função-objetivo.
- d^{obs} é um vetor formado pelos dados de produção histórica.
- d^{sim} é um vetor formado pelos dados de produção simulada.

b) Formulação por Mínimos Quadrados Ponderada

$$F = (d^{obs} - d^{sim})^T W (d^{obs} - d^{sim}) \quad (3.13)$$

Onde,

- W é uma matriz diagonal que atribui pesos individuais para cada uma das medições.

c) Formulação Generalizada dos Mínimos Quadrados

$$F = F_{obs} + \mu^* F_{previo} \quad (3.14)$$

Onde,

- F_{obs} é a soma ponderada dos quadrados das diferenças entre os valores de produção simulados d^{sim} e o histórico de dados de produção d^{obs} , conforme mostrado na Equação 3.4.
- μ^* é o fator de ponderação que expressa a importância relativa no modelo inicial.
- F_{previo} é a função objetivo prévio definida na Equação 3.5.

$$F_{obs} = \frac{1}{2} \sum_{i=1}^{n_{obs}} w_i \frac{(d_i^{sim} - d_i^{obs})}{\sigma_i^2} \quad (3.15)$$

onde, na Equação 3.4, temos que,

- n_{obs} é o número de observações (dados)
- w_i é um fator de ponderação, que pode ser usado equilibrar a magnitude dos componentes da função objetivo, que poder, por exemplo, ter unidades diferentes.
- σ_i é o desvio padrão do erro da medição de i^{th} observação.
- d_i^{sim} são os dados simulados pelo modelo matemático do reservatório.
- d_i^{obs} são os dados observados.

O fator

$$F_{previo} = (\alpha - \alpha_{previo})^T C_{\alpha}^{-1} (\alpha - \alpha_{previo}) \quad (3.16)$$

Onde,

- α é um vetor com os parâmetros do reservatório após a atualização.
- α_{previo} é um vetor com os parâmetros do reservatório antes da atualização
- C_{α} é a matriz de covariância dos estimadores dos parâmetros do modelo matemático.

3.4.3 Limitações do Ajuste ao Histórico

O melhor ajuste para os dados de produção, em alguns casos, não fornece uma boa previsão para a produção futura, que é o objetivo principal do ajuste ao histórico. As vezes, um modelo com bem ajustado ao histórico de produção produz previsões ruins, e vice-versa, um modelo pouco ajustado ao histórico pode levar a uma boa previsão (Tavassoli, et al. 2004). A experiência da comunidade de engenheiros de reservatório registra que qualidade dos dados de entrada é essencial para o ajuste ao histórico. Não existe ainda na engenharia de petróleo uma teoria que indique para um campo de petróleo específico e um conjunto de dados de produção se as previsões feitas após um procedimento de ajuste serão aceitáveis ou não. Esta é uma área muito interessante para pesquisa.

CAPÍTULO 4

4. Assimilação de Dados com EnKF

4.1 Introdução

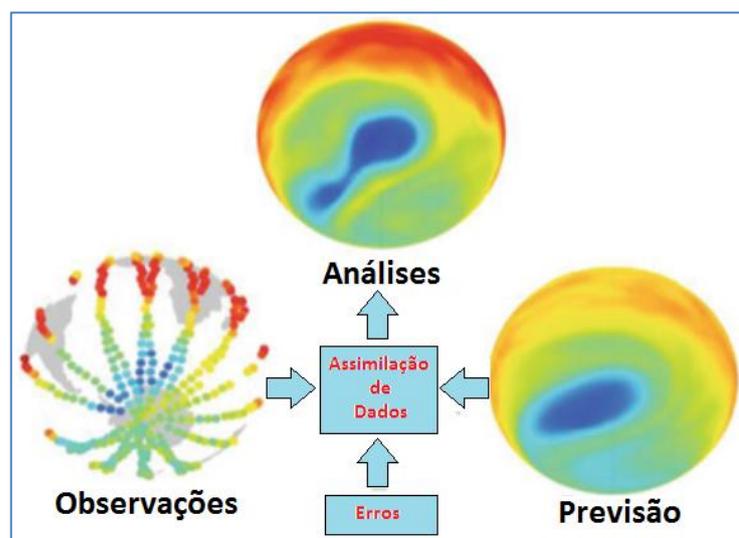
Assimilação de dados é um processo no qual os parâmetros do modelo e variáveis dinâmicas são atualizadas, mantendo consistência com as observações e as equações governantes. Métodos eficazes são necessários para atualizar continuamente os modelos dinâmicos através da assimilação de dados dinâmicos. O EnKF e suas variantes fornecem uma alternativa eficiente, que é explorada neste estudo para assimilação de dados dinâmicos.

Para qualquer sistema dinâmico, temos duas fontes de informação:

1. Medições do sistema ou observações.
2. Compreensão da evolução temporal e espacial do sistema ou modelos.

A Figura 4.1 mostra a relação entre os participantes de um processo de assimilação de dados.

Figura 4.1: Esquema de como assimilação de dados (DA) funciona e agrega valor à observação e informações do modelo.



(Lahoz, 2010)

Começaremos com uma revisão do filtro de Kalman. Em seguida, apresentaremos o Filtro de Kalman baseado em Ensembles (EnKF) e terminaremos este capítulo com os detalhes práticos da implementação do EnKF.

4.2 Revisão do Filtro de Kalman

O Filtro de Kalman (KF) foi desenvolvido em 1960 por Rudolf Emil Kalman, que publicou o seu famoso artigo (Kalman, 1960) descrevendo uma solução recursiva para o problema da filtragem linear de dados discretos. Desde então, devido aos grandes avanços da computação digital, o KF tem sido uma ferramenta muito importante nas áreas da navegação, climatologia, oceanografia, astronomia, economia, etc.

Na realidade o KF é um algoritmo que atualiza a estimativa do estado sempre que as medições estão disponíveis. Vamos supor que as medições e que a previsão do modelo estejam sujeitas a erros, isto é, que sejam ruidosas. Vamos supor também que o ruído seja Gaussiano, isto é, que seja uma variável aleatória com função densidade de probabilidade normal, e que nenhuma correlação esteja presente entre estes ruídos. Considere um função $\psi(x,t)$, definida em algum domínio espacial de interesse D com coordenadas espaciais x . Quando esta função representa a resposta de um modelo computacional de um fenômeno físico que foi criado com um método numérico como o método dos elementos finitos ou das diferenças finitas, esta função é aproximada por um conjunto finito de valores associados a pontos no domínio espacial. O conjunto destes pontos associados à informações de vizinhança forma uma malha computacional para o método numérico. A função $\psi(x,t)$ pode então ser representada por um vetor de estado ψ_k em cada instante de tempo t_k . Consideremos também que novas medições estão disponíveis no instante t_k . Definimos a nova estimativa para o vetor de estado como

$$\psi_k^a = \psi_k^f + K_k (d_k - M_k \psi_k^f), \quad (4.1)$$

(Evensen, 2007), em que ψ_k^a e ψ_k^f são as estimativas analisadas e previsão respectivamente, d_k é o vetor de medições, M_k é a matriz que fornece a relação entre as medições d_k e os estados do modelo ψ_k e K_k é o fator de ponderação ou o ganho de Kalman.

A covariância do estado analisado é dada por

$$\left(C_{\psi\psi}\right)_k^a = (1 - K_k M_k)^2 \left(C_{\psi\psi}\right)_k^f \quad (4.2)$$

Onde $\left(C_{\psi\psi}\right)_k^f$ é a covariância de erro do estado do modelo previsto, e $\left(C_{\epsilon\epsilon}\right)_k$ é a matriz de covariância de erro de medição. Um valor ideal para K_k pode ser encontrado escolhendo K_k de tal forma que a covariância de ψ_k^a seja minimizada. A solução deste problema de minimização é dada por: (Evensen, 2007).

$$K_k = \left(C_{\psi\psi}\right)_k^f M_k^T \left(M_k \left(C_{\psi\psi}\right)_k^f M_k^T + \left(C_{\epsilon\epsilon}\right)_k \right)^{-1} \quad (4.3)$$

e a covariância correspondente é dada por: (Evensen, 2007).

$$\left(C_{\psi\psi}\right)_k^a = (1 - K_k M_k) \left(C_{\psi\psi}\right)_k^f \quad (4.4)$$

4.3 Filtro de Kalman

Supomos que o verdadeiro estado ψ^t evolui no tempo de acordo com o modelo dinâmico linear

$$\psi_k^t = F \psi_{k-1}^t + q_{k-1} \quad (4.5)$$

em que F é um operador linear e q_{k-1} é o erro de modelo desconhecido sobre um passo de tempo de $k-1$ para k e o sobrescrito t é a sigla para verdade. Neste caso, o modelo evolui de acordo com,

$$\psi_k^f = F \psi_{k-1}^a \quad (4.6)$$

onde os sobrescritos a e f denotam análise e previsão respectivamente. Ou seja, dada a melhor estimativa possível (análise) para ψ no tempo t_{k-1} , a previsão é calculada para o tempo t_k , usando a equação aproximada (4.6).

A equação da covariância do erro é então derivada das equações (4.5) e (4.6). (Evensen, 2007).

$$(C_{\psi\psi})_k^f = F(C_{\psi\psi})_{k-1}^a F^T + (C_{qq})_{k-1} \quad (4.7)$$

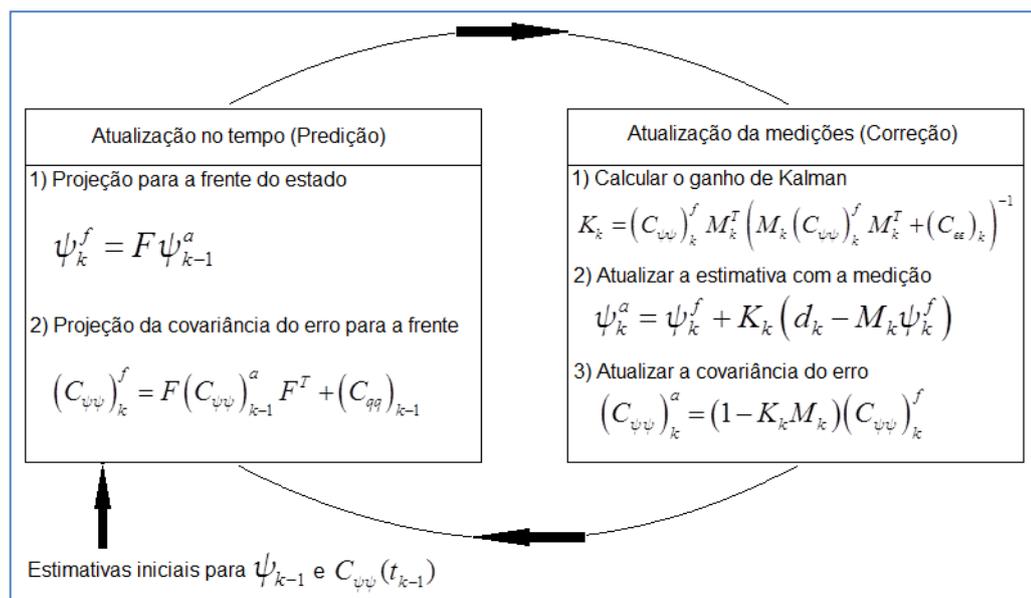
onde as matrizes de covariância de erro para as estimativas previstas e analisados são definidos como

$$C_{\psi\psi}^f = \overline{(\psi^f - \psi^t)(\psi^f - \psi^t)^T} \quad (4.8)$$

$$C_{\psi\psi}^a = \overline{(\psi^a - \psi^t)(\psi^a - \psi^t)^T} \quad (4.9)$$

A linha continua sobre uma variável está sendo empregada como o operador de expectativa. A Figura 4.2 descreve o princípio de funcionamento de assimilação de dados pelo filtro de Kalman.

Figura 4.2: O filtro de Kalman discreto.



(Fonte: Internet¹⁴)

4.4 Filtro Kalman Baseado em Ensemble

Enquanto o filtro de Kalman clássico fornece uma solução completa para a estimação de estado de sistemas lineares com ruído Gaussiano, a estimação de sistemas não lineares continua sendo um problema. Resultados rigorosos para problemas não-lineares têm pouca aplicabilidade ou são computacionalmente caros (S. Gillijns,

¹⁴ <http://www.cimat.mx/~alram/VC/MsAA.htm>

2006). O Filtro de Kalman baseado em Ensembles é uma abordagem de Monte Carlo (Weisstein, 2012), para superar as limitações filtro de Kalman mencionadas acima.

4.4.1 Representação das estatísticas de erro

As matrizes de covariância de erros $C_{\psi\psi}^f$ e $C_{\psi\psi}^a$ para as estimativas prevista e analisada no filtro de Kalman são definidos em termos do estado verdadeiro em (4.8) e (4.9). No entanto, já que o estado verdadeiro nunca é conhecido, no EnKF utilizamos um conjunto de modelos distintos para o mesmo sistema físico, o ensemble, que são ajustados e avançados no tempo simultaneamente. Aproximamos as matrizes de covariância do modelo em termos da média do ensemble $\bar{\psi}$ ao invés de usarmos o estado verdadeiro, de acordo com,

$$\left(C_{\psi\psi}^e\right)^f = \overline{\left(\psi^f - \bar{\psi}^f\right)\left(\psi^f - \bar{\psi}^f\right)^T} \quad (4.10)$$

$$\left(C_{\psi\psi}^e\right)^a = \overline{\left(\psi^a - \bar{\psi}^a\right)\left(\psi^a - \bar{\psi}^a\right)^T} \quad (4.11)$$

onde a sobrelinha denota o operador de expectativa. Assim, podemos usar uma interpretação onde o ensemble médio é a melhor estimativa e a dispersão do ensemble em torno da média é uma definição natural do erro associado ao ensemble médio.

4.5 Análise do Esquema

O EnKF depende do ensemble inicial do modelo de reservatório condicionado a todos os dados estáticos disponíveis até o momento da assimilação. Os modelos de reservatórios que formam os ensembles são geralmente criados usando técnicas de interpolação geoestatística como Simulação Sequencial Gaussiana, (Deutsche e Journal, 1998). O ensemble pode ser descrito como um conjunto de amostras aleatórias de uma função de distribuição de probabilidade Gaussiana previa multidimensional, com uma média e variância especificadas, que refletem a incerteza nessas estimativas iniciais. O processo de assimilação de dados tenta melhorar as estimativas iniciais destas variáveis do modelo, que podem incluir a distribuição espacial da permeabilidade, porosidade, saturação de fluidos, etc. Os modelos atualizados com-

preendem amostras de uma pdf posterior que são consistentes com todos os dados estáticos previamente adquiridos e observações.

O EnKF para o modelo de atualização contínua propaga um ensemble de realizações iniciais do reservatório ao longo do tempo para assimilar dados. O ensemble de vetores é denotado por:

$$A = [\psi_1, \psi_2, \dots, \psi_{N_e}] \quad (4.12)$$

onde N_e é o número total de membros do ensemble e $\psi_j; j=1 \dots N_e$ são vetores de estado. Cada vetor de estado consiste em parâmetros do modelo (m), tais como porosidade e permeabilidade, variáveis de estado (u), como pressão e saturação, que são variáveis dependentes do tempo e observações (d), como a taxa de produção de fluido do reservatório e o corte de água (WCT). Em um certo passo de tempo $t_k, k=1 \dots N_t$; o vetor de estado para o modelo de simulação do reservatório é representado como,

$$\psi_k = \begin{bmatrix} m_k \\ u_k \\ d_k \end{bmatrix} \quad (4.13)$$

onde k indica o passo de tempo em que os dados são assimilados.

A ideia principal do EnKF para assimilação de dados contém duas etapas sequenciais. Uma delas é a etapa de previsão, em que o modelo de previsão é aplicado a cada membro do ensemble separadamente usando um simulador de reservatório com base na solução das equações dinâmicas de fluxo e transporte de fluidos no meio poroso,

$$\psi_{k,j}^f = F(\psi_{k-1,j}^a) + q_{k-1} \quad (4.14)$$

onde, $F(\bullet)$ é o operador previsão, que é representado pelo simulador de reservatório e q_{k-1} representa o erro de modelo. Os sobrescritos f e a denotam os estados previstos e analisados, respectivamente. Nesta etapa do modelo as propriedades da rocha são inalteradas, mas é substituída a pressão, saturação e dados simulados no vetor de estado previsto. O passo de previsão é muitas vezes o passo mais exigente

computacionalmente, para grandes modelos de reservatórios e grandes tamanhos do ensemble.

A outra é a etapa de análise, no qual a atualização do modelo é feita corrigindo as variáveis que representam o estado do sistema para honrar as observações. O esquema de atualização no KF usando as covariâncias do ensemble é definido em (4.10) e (4.11). É essencial que as observações possam ser tratadas como variáveis aleatórias tendo uma distribuição com média igual ao valor observado e covariância igual a C_{ee} (Burgers, et al. 1998). Começamos por definir um ensemble de observações

$$d_j = d_{obs,j} + e_j, j = 1 \dots N \quad (4.15)$$

onde N é o número de membros do ensemble. Em seguida, a matriz de covariância do ensemble dos erros de medição é definida como,

$$C_{ee}^e = \overline{ee^T} \quad (4.16)$$

Esta matriz converge para o erro atual da matriz covariância C_{ee} usado no filtro de Kalman, quando temos um número infinito de realizações e o sobescrito e indica que a matriz corresponde aos erros calculados para o ensemble.

No passo de análise do EnKF, atualizações são realizadas em cada um dos membros do ensemble, dadas por

$$\psi_j^a = \psi_j^f + (C_{\psi\psi}^e)^f M^T \left(M (C_{\psi\psi}^e)^f M^T + C_{ee}^e \right)^{-1} (d_j - M \psi_j^f) \quad (4.17)$$

Onde $M = [0|I]$ é o operador de observação, que extrai os dados simulados a partir do vetor de estado ψ_j^f . Além disto, 0 é uma matriz $N_d \times (N_m + N_u)$, com todos os elementos nulos e I é a matriz identidade de $N_d \times N_d$.

Devemos notar que com um tamanho finito do ensemble, o uso das covariâncias do ensemble é uma aproximação das verdadeiras covariâncias. Além disso, as matrizes $M (C_{\psi\psi}^e)^f M^T$ e C_{ee}^e são singulares quando o número de medições é maior do que o

número de membros do ensemble (Evensen, 2007), e devemos usar pseudo-inversão em vez de inversão para avaliar (4.17).

A equação (4.17), implica que,

$$\bar{\psi}^a = \bar{\psi}^f + (C_{\psi\psi}^e)^f M^T \left(M (C_{\psi\psi}^e)^f M^T + C_{\epsilon\epsilon}^e \right)^{-1} (\bar{d} - M \bar{\psi}^f) \quad (4.18)$$

em que $\bar{d} = d$, uma vez que temos um ensemble de média zero das perturbações de medição. Assim, temos a mesma relação entre o ensemble médio analisado e previsto, como temos entre o estado analisado e previsto no filtro de Kalman padrão, além de utilizar $(C_{\psi\psi}^e)^{f,a}$ e $C_{\epsilon\epsilon}^e$ em vez de $C_{\psi\psi}^{f,a}$ e $C_{\epsilon\epsilon}$ respectivamente. Devemos também observar que a introdução de um ensemble de observações não afeta a atualização do ensemble médio.

Vamos agora mostrar que, atualizando cada um dos membros do ensemble usando as observações perturbadas, podemos criar um novo ensemble com as estatísticas de erro corretas. Obtivemos a estimativa da covariância de erro analisada como resultado do esquema de análise acima, apesar de manter a forma padrão do filtro de Kalman para as equações de estado. Primeiro, obtemos usando (4.17) e (4.18).

$$\psi_j^a - \bar{\psi}^a = (I - K_e M) (\psi_j^f - \bar{\psi}^f) + K_e (d_j - \bar{d}), \quad (4.19)$$

onde o ganho de Kalman é, (Evensen, 2007)

$$K_e = (C_{\psi\psi}^e)^f M^T \left(M (C_{\psi\psi}^e)^f M^T + C_{\epsilon\epsilon}^e \right)^{-1} \quad (4.20)$$

Usando a equação (4.11), obtemos a atualização do erro da covariância como,

$$(C_{\psi\psi}^e)^a = (I - K_e M) (C_{\psi\psi}^e)^f \quad (4.21)$$

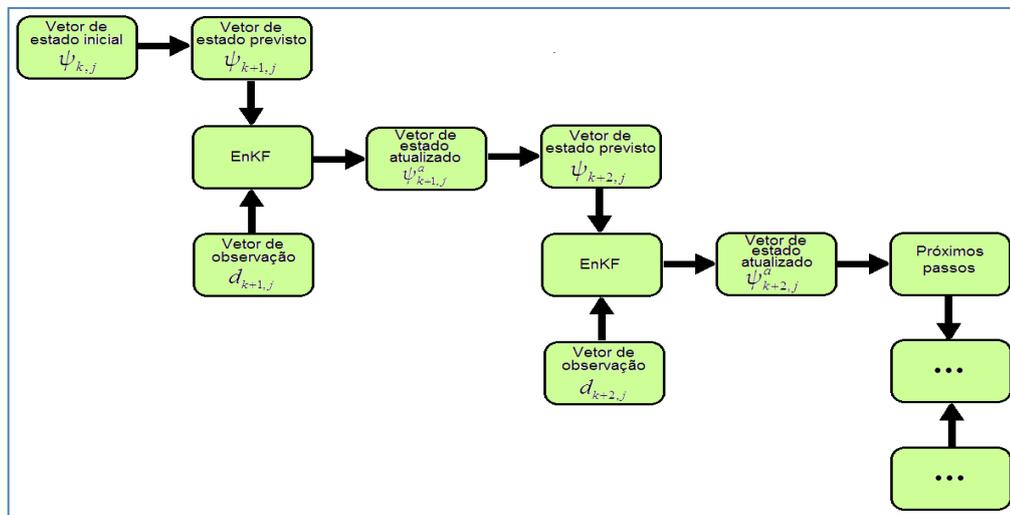
Agora, consideramos que as incertezas do modelo e as incertezas das medições são independentes. No entanto, usando um ensemble de tamanho finito e negligenciando o termo cruzado introduzem-se erros de amostragem.

Quando a matriz de ganho de Kalman é singular, a inversão em (4.20) pode ser substituída pela pseudo inversão, e pode-se escrever o ganho de Kalman como:

$$K_e = (C_{\psi\psi}^e)^f M^T (M (C_{\psi\psi}^e)^f M^T + C_{\epsilon\epsilon}^e)^+ \quad (4.22)$$

Quando a matriz na inversão tem posto completo, (4.22) torna-se idêntico a (4.20) (Evensen, 2007). A Figura 4.3 mostra o princípio de funcionamento do EnKF para um membro do ensemble.

Figura 4.3: Assimilação de dados com EnKF para o membro j do ensemble.



(Khairullah, 2012)

4.6 Implementação Prática

4.6.1 Representação do Ensemble da Covariância

Definimos a matriz A cujas colunas são os membros do ensemble $\psi_i \in \mathbb{R}^n$ (Evensen, 2007) por,

$$A = (\psi_1, \psi_2, \dots, \psi_N) \in \mathbb{R}^{n \times N} \quad (4.23)$$

onde N é o número de realizações e n é o tamanho do vetor de estado do modelo. O ensemble médio é armazenado em cada coluna de \bar{A} , que é definida como:

$$\bar{A} = A 1_N, \quad (4.24)$$

em que $1_N \in \mathbb{R}^{N \times N}$ é uma matriz cujas entradas são todas iguais $1/N$. A matriz de perturbação do ensemble é definida como

$$A' = A - \bar{A} = A(I - 1_N) \quad (4.25)$$

A matriz de covariância do ensemble $C_{\psi\psi}^e \in \mathbb{R}^{n \times n}$ pode ser definida como

$$C_{\psi\psi}^e = \frac{1}{N-1} A' (A')^T \quad (4.26)$$

4.6.2 Perturbações das Medições

Dado um vetor de medições de $d \in \mathbb{R}^m$, onde m é o número de medições, que definem os N vetores perturbados das observações (Evensen, 2007) como;

$$d_j = d + \epsilon_j, \quad j = 1 \dots N \quad (4.27)$$

que estão armazenados nas colunas da matriz

$$D = (d_1, d_2, \dots, d_N) \in \mathbb{R}^{m \times N} \quad (4.28)$$

enquanto que o ensemble médio das perturbações são armazenados na matriz

$$E = (\epsilon_1, \epsilon_2, \dots, \epsilon_N) \in \mathbb{R}^{m \times N} \quad (4.29)$$

a partir do qual construímos a representação da matriz de covariância de erro de medição do ensemble

$$C_{\epsilon\epsilon}^e = \frac{1}{N-1} EE^T \quad (4.30)$$

4.6.3 Equação de Análise

A equação de análise (4.17), expressada em termos de matrizes do ensemble (Evensen, 2007), é,

$$A^a = A + C_{\psi\psi}^e M^T (M C_{\psi\psi}^e M^T + C_{\epsilon\epsilon}^e)^{-1} (D - MA) \quad (4.31)$$

em que a matriz de medição é $M \in \mathbb{N}^{m \times n}$.

Definimos o conjunto de vetores de inovação,

$$D' = D - MA \quad (4.32)$$

em que $D' \in \mathbb{R}^{m \times N}$. Usando D' , juntamente com as definições das matrizes de covariância de erro do ensemble em (4.26) e (4.30), a matriz de perturbação pode ser expressa como,

$$A^a = A + A'A'^T M^T (MA'A'^T M^T + EE^T)^{-1} D' \quad (4.33)$$

Vamos agora definir a matriz $S \in \mathbb{R}^{m \times N}$, mantendo as medições das perturbações do conjunto, por,

$$S = MA' \quad (4.34)$$

E a matriz $C \in \mathbb{R}^{m \times m}$,

$$C = SS^T + (N-1)C_{\epsilon\epsilon}. \quad (4.35)$$

Aqui podemos usar a matriz de covariância do erro das medições exata $C_{\epsilon\epsilon}$ ou a representação de posto reduzido $C_{\epsilon\epsilon}^e$ definida em (4.30).

A equação de análise (4.33) pode então ser escrita como,

$$\begin{aligned} A^a &= A + A'S^T C^{-1} D' = A + A(I - 1_N)S^T C^{-1} D' \\ &= A(I + (I - 1_N)S^T C^{-1} D') = AX \end{aligned} \quad (4.36)$$

em que usamos a identidade em (4.25). A matriz $X \in \mathbb{R}^{N \times N}$ é definida como,

$$X = I + (I - 1_N)S^T C^{-1} D' \quad (4.37)$$

Com base no pressuposto de que $1_N S^T \equiv 0$, (4.36) pode ser escrita como,

$$A^a = A(I + S^T C^{-1} D') = AX \quad (4.38)$$

Em seguida, (4.37) se torna

$$X = I + S^T C^{-1} D' \quad (4.39)$$

Quando o número de medições é maior do que o número de membros das realizações, a matriz C é singular e em vez de usar inversa devemos usar a pseudo-

inversa de C nas expressões acima. Uma forma popular para calcular a pseudo-inversa de uma matriz C é usando a decomposição em valores singulares (SVD) usando a seguinte formulação (Evensen, 2007).

$$C^+ = VS^+U^T \quad (4.40)$$

em que U e V são as matrizes singulares à esquerda e à direita de C , respectivamente, S é uma matriz diagonal contendo os valores singulares de C , em ordem decrescente, e S^+ é definido como

$$S^+ = \begin{cases} \frac{1}{S_{ij}}, & \text{se } i = j \text{ e } S_{ij} \neq 0 \\ 0, & \text{caso contrário} \end{cases} \quad (4.41)$$

A outra maneira de escrever (4.35) é

$$C = SS^T + EE^T \quad (4.42)$$

$$C = (S + E)(S + E)^T \quad (4.43)$$

Assumindo que as perturbações do conjunto e as perturbações de medição são correlacionadas, que é equivalente a

$$SE^T = MA'E^T = 0 \quad (4.44)$$

Calcula-se o SVD

$$S + E = U\Sigma V^T, \quad (4.45)$$

em seguida (4.33) se torna

$$A^a = A + A'S^T U \Lambda^{-1} U^T D', \quad (4.46)$$

em que

$$\Lambda = \Sigma \Sigma^T \quad (4.47)$$

Esta é outra abordagem de cálculo da matriz de perturbação, evitando a inversão de uma matriz grande quando $m \gg N$, baseada nas suposições acima mencionadas.

Existem duas principais fontes de erros de amostragem no EnKF: o uso de um ensemble finito de realizações de modelos estocásticos e a introdução de perturbações na (Evensen, 2004, 2007).

Na Figura 4.4 se mostra um algoritmo base para o Filtro de Kalman baseado em ensemble.

Figura 4.4: Algoritmo Básico do EnKf.

```

Entrada :  $Q, C_{ee}^e, d_{obs}$ 
 $A = [\psi_1, \psi_2, \dots, \psi_{N_e}] \rightarrow$  Gerar Ensemble inicial dos parâmetros do modelo
 $\psi_{0,j} \rightarrow$  O número de estados iniciais do sistema, de acordo com o número de realizações
for  $k = 1 : t_k$ 

    Passo em Frente
     $q_{k-1} \rightarrow$  Calcule modelo de ruído com  $N(0, Q)$ 
     $\psi_{k,j}^f = F(\psi_{k-1,j}^a) + q_{k-1}$  % Passo em Frente
     $(C_{\psi\psi}^e)^f = \frac{1}{N_e - 1} \sum_{j=1}^{N_e} \left\{ (\psi_j^f - \overline{\psi^f})(\psi_j^f - \overline{\psi^f})^T \right\}$  % Cálculo da covariância

    Passo de atualização
     $\epsilon_{k,j}^d \rightarrow$  Calcule ruído de medição com  $N(0, C_{ee}^e)$ 
     $d_{k,j} = d_{obs,j} + \epsilon_{k,j}^d$  % Medições perturbadas
     $K_e = (C_{\psi\psi}^e)^f M^T (M (C_{\psi\psi}^e)^f M^T + C_{ee}^e)^{-1}$  % Cálculo de ganho de Kalman
     $\psi_{k,j}^a = \psi_{k,j}^f + K_e (d_j - M \psi_{k,j}^f)$  % Parâmetros do modelo de atualização
     $(C_{\psi\psi}^e)^a = (I - K_e M) (C_{\psi\psi}^e)^f$  % Atualização da covariância do erro
     $k \rightarrow k + 1$  % Próximo passo
end

```

(Fonte: Autor)

CAPÍTULO 5

5. Software e Metodologia

5.1 Introdução

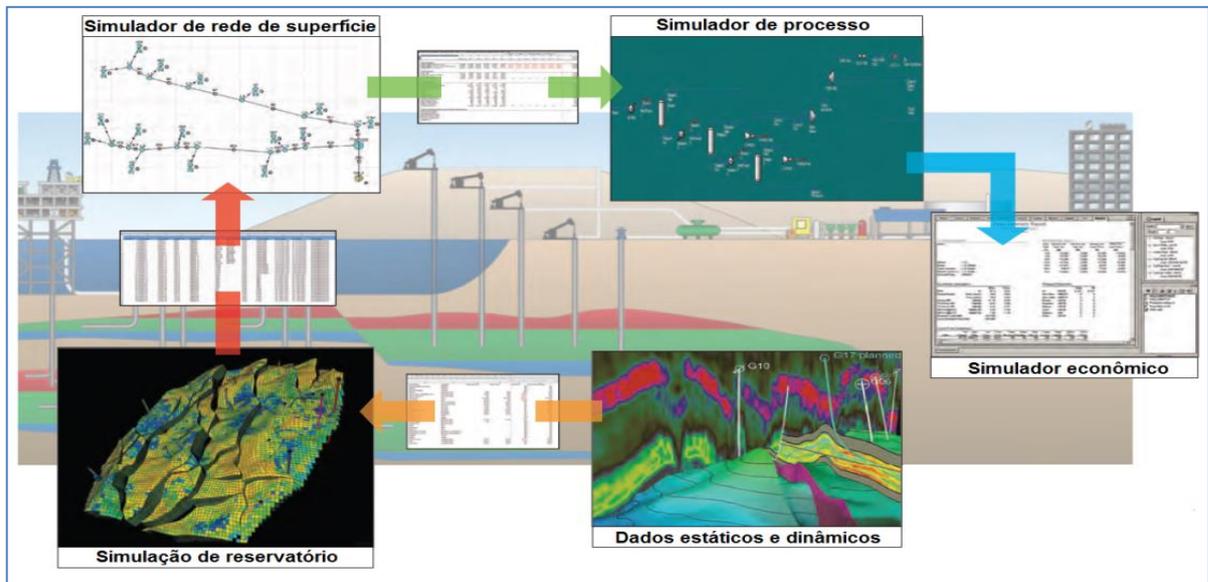
Atualmente, vários programas de computador são usados dentro dos processos de gerenciamento e engenharia de reservatórios. Simuladores de reservatórios, programas para interpretação sísmica e criação de modelos, assim como software de visualização 3D são exemplos típicos. Estes programas são usados para modelar, compreender melhor e fazer previsões sobre o comportamento dos reservatórios. Uma melhor compreensão do reservatório possibilita previsões mais precisas, maiores oportunidades para aumentar a recuperação de petróleo, maior retorno de investimento e maior tempo de vida de um campo e um maior lucro.

No mercado atual encontramos muitos simuladores de reservatórios que modelam escoamentos em meios porosos considerando alta complexidade geológica, trajetórias complexas dos poços, métodos avançados de recuperação, recuperação de óleo pesado e reservatórios não convencionais. Há simuladores comerciais (Eclipse, Imex, Petrel, etc) e simuladores com distribuição livre, livres (BOAST, WinB4D, MRST, etc), que tem sido utilizados para simular estratégias complexas de produção. Um número grande de fatores leva a simulação da produção a obter resultados precisos no menor tempo possível.

Nesta dissertação foi usado como plataforma de desenvolvimento o Matlab Reservoir Simulation Toolbox (MRST), que é um simulador com distribuição livre. O simulador é desenvolvido pela SINTEF e pode ser obtido a partir de <http://www.sintef.no/Projectweb/MRST/>. Neste capítulo será mostrado o ajuste ao histórico usando o EnKF acoplado ao simulador de reservatórios MRST.

A Figura 5.1 mostra um processo completo e integrado para o gerenciamento de um campo de petróleo. Esta dissertação se concentra na etapa da simulação de reservatórios, tentando melhorar seus modelos com a assimilação de dados de produção.

Figura 5.1: Simulação da Produção em Reservatórios.



(Adaptado de Schlumberger e Chevron, 2012)

Um engenheiro de reservatório usa dados estáticos e dinâmicos (canto inferior direito, Fig. 5.1) e desenvolve um modelo para um simulador de reservatório (canto inferior esquerda, Fig. 5.1). O simulador de reservatórios, cuja principal tarefa é o de modelar o escoamento multifásico em meios porosos, produz perfis de produção calculados em função do tempo de operação do campo para os poços do reservatório. Estes dados são transferidos ao engenheiro de produção para desenvolver modelos de elevação, que tratam o escoamento multifásico dentro das tubulações que levam o petróleo do reservatório até a superfície, e para alimentarem os simuladores de rede de superfície (canto superior esquerdo, Fig. 5.1). Um engenheiro de instalações de superfície utiliza os dados de produção e composição do óleo produzido para construir um modelo de planta de processo com a ajuda de um simulador de processos (canto superior direito, Fig. 5.1). Finalmente, os dados de todos os simuladores, de reservatórios, elevação e processos, são usados para alimentar um simulador econômico (à direita, Fig. 5.1). Este procedimento leva ao planejamento da produção para obter os melhores resultados econômicos dentro das limitações operacionais e econômicas do campo (Schlumberger e Chevron, 2012).

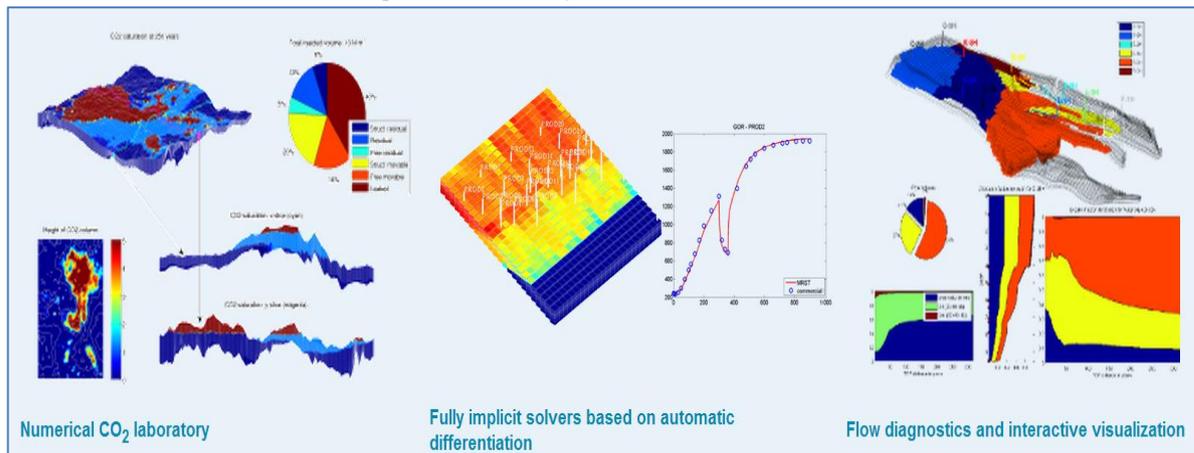
5.2 Matlab Reservoir Simulation Toolbox (MRST)

O MRST é um programa de código aberto desenvolvido em conjunto com a iniciativa Open Porous Media (OPM) do SINTEF, um Instituto de Matemática Aplicada localizado em Oslo. O objetivo da MRST é de funcionar como uma plataforma de testes eficiente para novas implementações de técnicas de discretização, formulações numéricas e métodos de solução de sistemas de equações lineares e não lineares em aplicações de simulação de reservatório. Ele contém módulos para geração de malhas, um conjunto completo de funções e rotinas para manipulação de parâmetros de fluidos, poços e condições de contorno, ferramentas de pós-processamento e visualização em 2D e 3D, além de uma série de módulos para os métodos de solução numérica e otimização.

O MRST foi lançado sob a licença GNU General Public License (GPL). As equações governantes do escoamento em meios porosos são avaliadas utilizando o método dos volumes finitos e o sistema de equações algébricas, resultante da discretização das equações diferenciais parciais que modelam o escoamento é resolvido de forma totalmente implícita usando o método de Newton-Raphson a cada passo de tempo de integração. Em uma técnica implícita, os valores das variáveis em todas as células do domínio computacional, em um instante de tempo, são funções dos valores destas variáveis em outras células do domínio no mesmo instante de tempo, o que requer a solução de um sistema de equações algébricas para a determinação de seus valores. Em contraste, em uma técnica explícita, os valores nas células dependem apenas de valores calculados em instantes de tempo anteriores, de forma que cada célula pode ser atualizada independentemente, sem a necessidade da solução de um sistema de equações algébricas.

Na Figura 5.2, mostramos os exemplos de módulos que são fornecidos pelo MRST.

Figura 5.2: Exemplos de Módulos MRST.



(Adaptado: <http://www.sintef.no/Projectweb/MRST/>)

Em razão de o MRST ser um programa cujo código fonte está pública e livremente disponível, nesta dissertação o autor usou uma rotina existente do EnKF para MRST 2013a, que foi fornecida por Olwijn Leeuwenburgh¹⁵ (TNO), além disso, o autor modificou vários arquivos do código fonte fornecido, para gerar os cálculos de produção do óleo, geração de malhas e gráficos como os de produção acumulada e instantânea, permeabilidades, porosidade, saturação e pressão, que não eram fornecidos no código fonte original, após realizar ajustes ao histórico de produção.

5.3 Métodos de Amostragem

O planejamento de experimentos é uma técnica importante da estatística que é usada nas ciências experimentais para que experimentos com múltiplos fatores possam ser realizados com um número razoável de amostras, o que seria inviável com, por exemplo, um planejamento fatorial completo (Montgomery e Douglas, 2012). Há uma diferença fundamental entre o planejamento para experimentos físicos e experimentos computacionais, visto que os primeiros são sujeitos a erros de medição enquanto que os experimentos computacionais, salvo interesse particular, são determinísticos. Desta forma, normalmente não se amostra várias vezes o mesmo ponto no planejamento de experimentos computacionais. Várias metodologias estão disponíveis para geração dos valores dos fatores que compõem um ponto amostral, tais como Quasi-Monte Carlo (QMC), Hipercubo latino (LHS - *Latin Hypercube Sampling*) e método

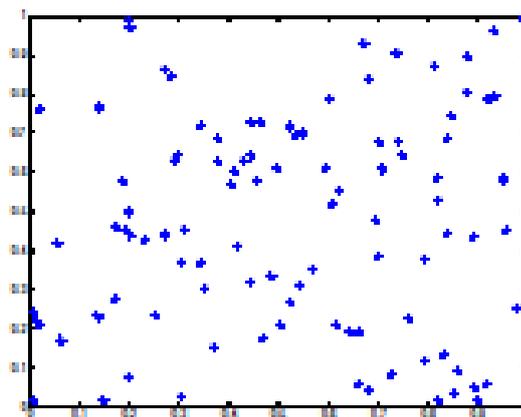
¹⁵ Engenheiro de Reservatório - Pesquisador da TNO.

da tesselação de Voronoi (CVT - *Centroidal Voronoi Tessellation*), (Giunta, 2002; Giunta et al, 2003). Todas as técnicas tentam gerar boa cobertura do espaço amostral e uma distribuição espacial que se aproxime da uniforme, com um número razoável de amostras, particularmente quando o número de fatores é muito grande. Neste trabalho as técnicas para geração dos pontos amostrais serão chamadas de DOE, do inglês Design of Experiments. Neste trabalho foi utilizada como técnica de DOE o Hipercubo Latino (LHS), mas faremos uma breve definição de algumas alternativas.

A característica principal do DOE aplicado em experimentos computacionais determinísticos é que não existe erro aleatório, ou seja, se uma simulação computacional é executada duas vezes teremos a geração do mesmo conjunto de amostras.

Consideraremos como exemplo um experimento computacional com apenas dois fatores, cujos valores variam entre 0 e 1, por simplicidade. Uma técnica clássica para geração de amostras é o método de Monte Carlo (Halton, 1970; Oliver et al, 2008). Para um pequeno número de amostras, no entanto, tem-se possivelmente a geração de amostras mal distribuídas no domínio, com a geração de vários vazios e regiões com uma concentração excessiva de pontos amostrais, conforme mostrado na Figura 5.3, onde temos a geração 100 pontos amostrais.

Figura 5.3: Amostragem gerada pelo método de Monte Carlo.



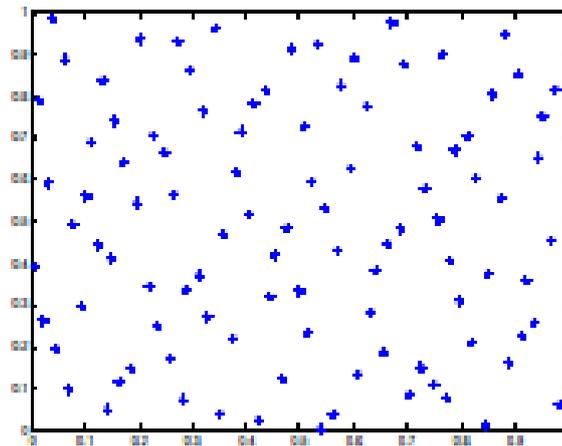
(Giunta et al, 2003)

5.3.1 Método de Quasi-Monte Carlo (QMC)

O método de Quasi-Monte Carlo (Giunta et al, 2003; Kocis e Whiten, 1997) é puramente determinístico e gera amostragens com uma melhor distribuição do que seria obtido com um método de geração aleatória, como Monte Carlo. Este método é caracterizado por usar sequências de baixa discrepância, onde discrepância é a medida quantitativa de quanto a distribuição de amostras desvia de uma distribuição uniforme ideal.

A Figura 5.4 mostra um planejamento usando o método Quasi-Monte Carlo, onde temos a geração de 100 pontos amostrais para o mesmo exemplo da Figura 5.3.

Figura 5.4: Amostragem gerada pelo método de Quasi-Monte Carlo.



(Giunta et al, 2003)

5.3.2 Método da Tesselação de Voronoi (CVT)

Considerando um conjunto de pontos gerados aleatoriamente, o método da tesselação de Voronoi (CVT – *Centroidal Voronoi Tessellation*) é baseado na construção de mosaicos de Voronoi (Du et al, 1999). Nas Figuras 5.5 a 5.7 temos exemplo de um conjunto de amostras gerados pelo método CVT. Este método pode ser usado na forma LCVT (*Latinized Centroidal Voronoi Tessellations*). Neste caso, para todas as projeções unidimensionais das amostras, somente deve existir uma única projeção no intervalo particionado de cada variável. Está ideia será mostrada na próxima seção referente a método LHS.

Figura 5.5: Conjunto inicial de pontos geradores.

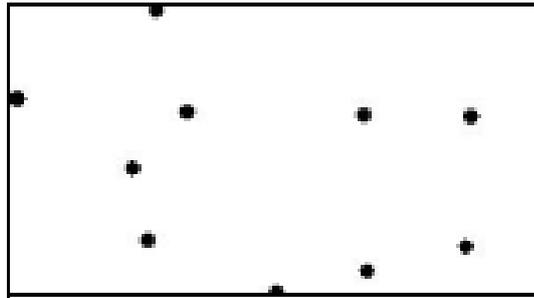


Figura 5.6: O método de tesselação de Voronoi para os pontos da Figura 3.3.

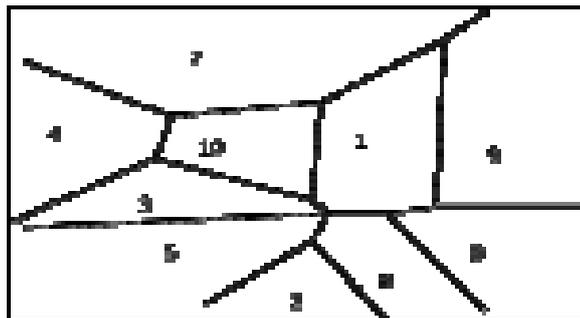
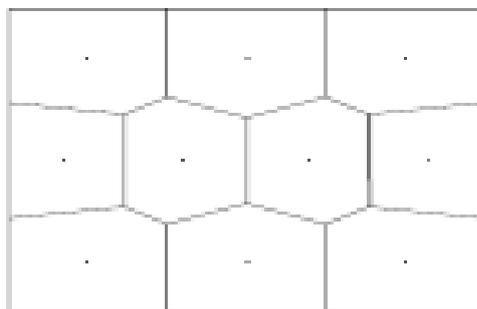


Figura 5.7: O método de tesselação de Voronoi (CVT) após atingir a condição final de uniformidade.



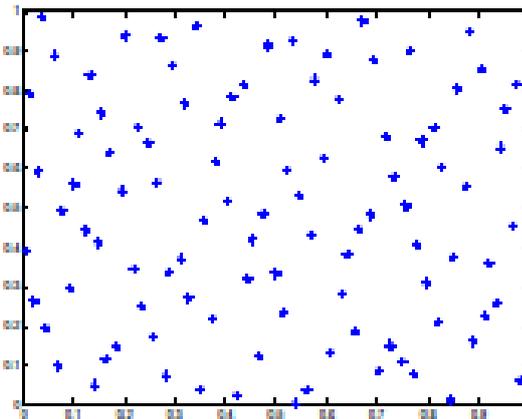
5.3.3 Hipercubo Latino (LHS)

A ideia principal do Hipercubo Latino (LHS - *Latin Hypercube Sampling*) consiste em dividir a faixa de valores de cada variável (x_j) em m intervalos com mesma probabi-

lidade de incidência, sendo os pontos da amostragem gerados aleatoriamente em cada intervalo (Giunta et al, 2003).

Considerando n amostras com dimensão d , (x_1, x_2, \dots, x_d) , e as faixas de valores de cada variável, teremos um total de n^d intervalos. Um exemplo de amostras escolhidas aleatoriamente num espaço de dimensão dois é mostrado na Figura 5.8.

Figura 5.8: Amostragem gerada pelo método LHS num espaço bidimensional.



(Giunta et al, 2003)

As condições para geração das amostras são:

1. Somente uma amostra é escolhida dentro de cada área retangular (*Bin*).
2. Para todas as projeções unidimensionais das amostras, somente deve existir uma única projeção no intervalo particionado de cada variável.

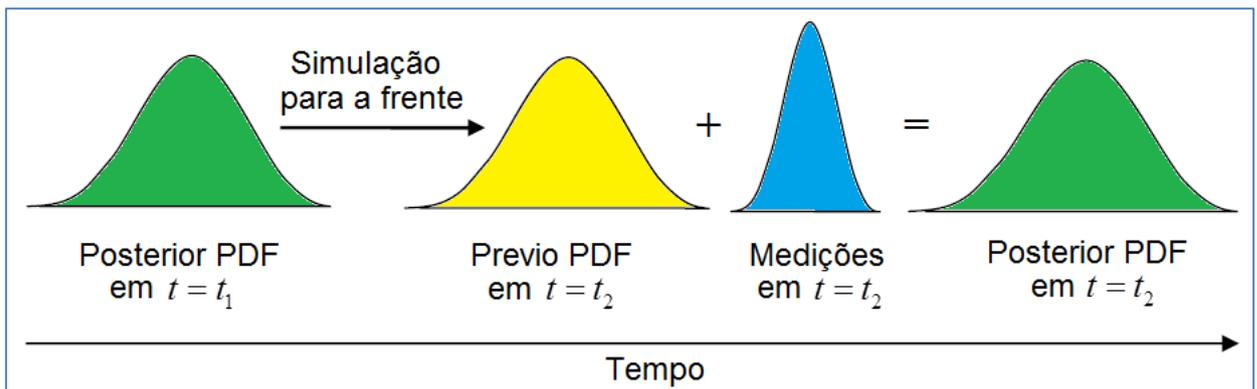
5.4 Ajuste ao Histórico com EnKF

O EnKF é uma abordagem de Monte Carlo que atualiza um conjunto de modelos sequencialmente, usando assimilação de dados. As realizações iniciais geralmente são geradas a partir do conhecimento prévio geológico (Figura 5.5). As etapas seguintes do ajuste ao histórico são um laço iterativo de um circuito fechado que consiste na previsão e assimilação de dados. O EnKF oferece um cenário ideal para a atualização em tempo real e previsão em modelos de simulação de reservatórios. Periodicamente, ao longo da vida operacional do reservatório, novos valores de produção de óleo, gás e água estão disponíveis e podem ser assimiladas para melhorar os parâmetros do modelo (permeabilidade e porosidade) e as variáveis associadas

(saturação e pressão). Portanto, o ensemble de modelos ao final do processo de ajuste proporciona modelos que são condicionados aos dados de produção existentes, e podem ser utilizados para previsão da futura estratégia de planejamento da produção. Alguns detalhes sobre a codificação de cada uma destas etapas serão apresentados a seguir. As abordagens para verificar os resultados correspondentes e as previsões também são discutidas.

Este capítulo é uma base para o próximo capítulo, onde será mostrado o impacto do esquema de assimilação, como melhorar a estratégia inicial de amostragem aleatória, e uma modificação do método EnKF usando uma média ponderada na equação assimilação. O EnKF poder ser visto sob um ponto de vista Bayesiano, ilustrado na Figura 5.9.

Figura 5.9: Ilustração da EnKF a partir do ponto de vista da concepção bayesiana.



(Liang, 2007)

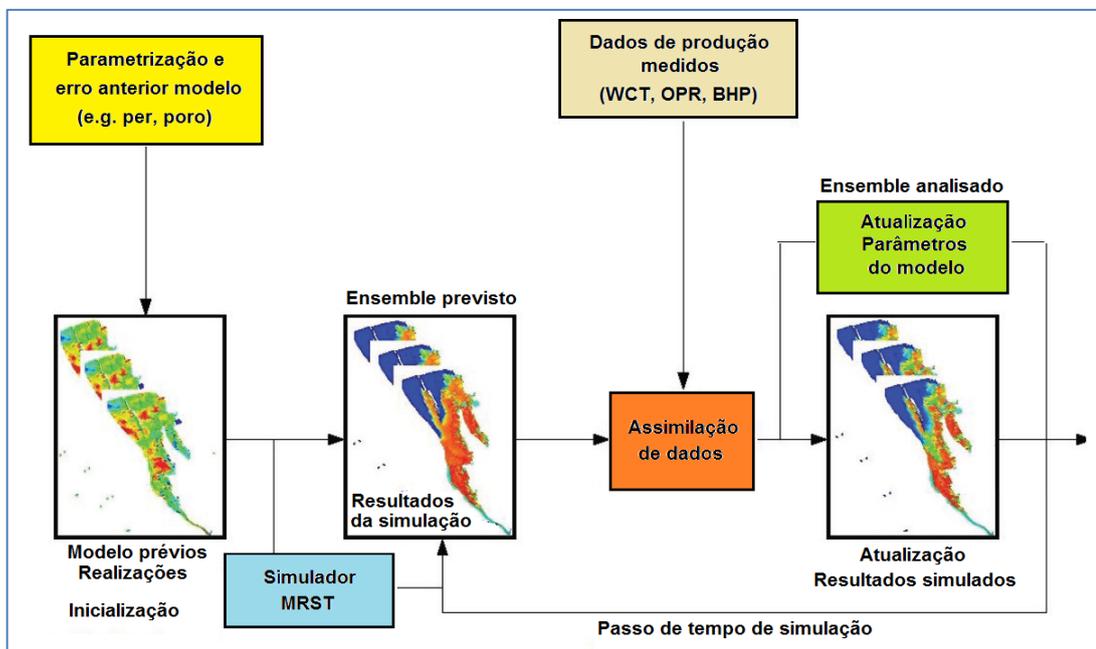
A aproximação do ensemble posterior é dependente apenas da média e da covariância calculadas a partir do ensemble anterior e das novas medições. O ensemble posterior no tempo t_i é projetado para frente no tempo e utilizado como conjunto prévio no tempo t_{i+1} , e o processo repete-se sequencialmente ao longo do tempo.

5.5 Fluxo de Trabalho de Ajuste ao Histórico Baseado em EnKF

Quando um ensemble inicial de modelos de reservatórios é gerado, o EnKF é usado para atualizar sequencialmente o ensemble no tempo para honrar as novas observações de produção no momento em que eles chegam. O EnKF consiste em uma integração para a frente gerando a previsão seguido pela atualização da variável de estado para gerar a análise, como mostrado na Figura 5.10. A cada passo de assimi-

lação, o modelo é atualizado de acordo com a Equação 4.14 e não há necessidade de executar o simulador a partir do tempo inicial novamente. As observações assimiladas são consideradas como variáveis aleatórias com uma distribuição normal com média igual ao valor observado e uma covariância de erro que reflete a precisão da medição. O conjunto atualizado é então integrado no tempo até o próximo tempo de atualização. O resultado é um conjunto atualizado de realizações, condicionado a todos os dados anteriores de produção e, portanto, fornece o ponto de partida ideal para as previsões de produção futura (Seiler et. al. 2009).

Figura 5.10: Fluxo de trabalho de ajuste ao histórico baseado em EnKF para os dados de produção



(Fahimuddin, 2010)

5.6 Metodologia EnKF

O EnKF é um procedimento para ajuste ao histórico automático que requer colaboração entre diversas áreas da engenharia e geologia do petróleo e tem diversas etapas:

- ❖ Criação de um modelo inicial do reservatório com base na interpretação geológica;
- ❖ Parametrização do problema de estimação (variáveis estáticas e dinâmicas, medições, etc);
- ❖ Definição do problema em termos matemáticos;

- ❖ Criação do ensemble inicial;
- ❖ Execução do laço de assimilação de dados e atualização dos modelos;
- ❖ Interpretação de resultados.

Os resultados insatisfatórios devem ser analisados novamente desde o início para encontrar o erro e refazer o procedimento a partir do ponto onde foi detectado o problema na assimilação.

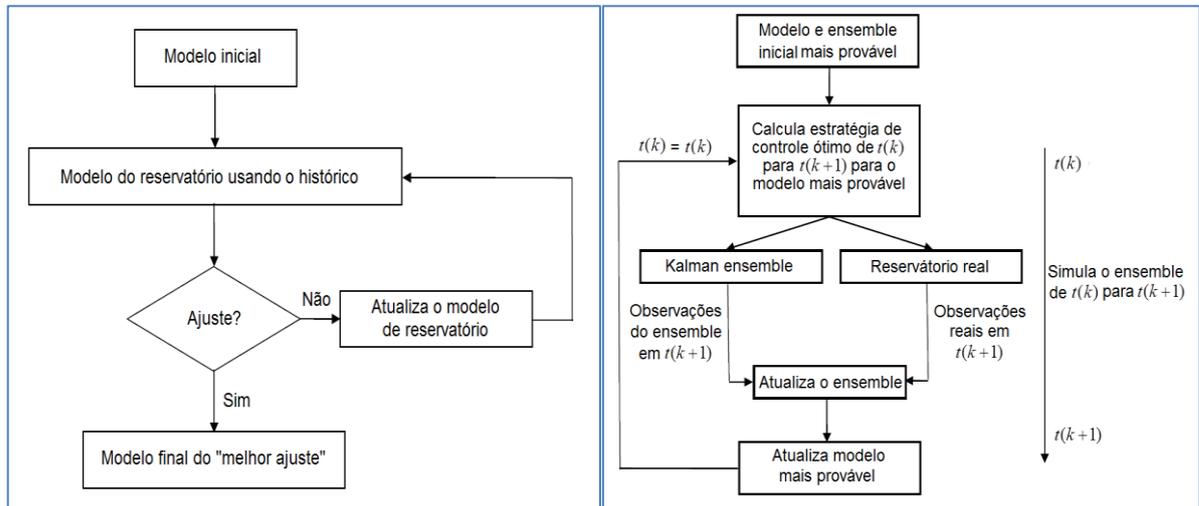
5.7 Ajuste de Histórico Tradicional Versus Baseado em EnKF

A primeira desvantagem do método tradicional é o fato de que para os métodos tradicionais de ajuste ao histórico, os dados de produção para todo o histórico são ajustados ao mesmo tempo, e, portanto, é necessário repetir simulações de fluxos de todo o histórico. Isso faz com que o ajuste seja extremamente demorado. Na Figura 5.11, mostramos uma análise comparativa entre o ajuste ao histórico baseada em EnKF e um método de ajuste ao histórico tradicional. Observe o recurso de atualização sequencial do EnKF comparado com o método tradicional.

Em segundo lugar, métodos de otimização baseados em gradientes para o ajuste ao histórico exigem cálculos de coeficiente de sensibilidade e de minimização que são complicados, caros e muitas vezes, ficam presos em mínimos locais. No entanto, há também técnicas estocásticas de otimização global que não precisam de gradiente e não ficam presos em mínimos locais, porém estas técnicas são muito mais custosas em termos de tempo de execução do que a aplicação do EnKF (Fahimuddin, 2010).

É difícil avaliar a incerteza com os métodos tradicionais, e este processo pode incluir a repetição do processo de ajuste ao histórico com diferentes modelos iniciais. Com o EnKF a avaliação da incerteza é imediata e intrínseca ao método, com a avaliação direta de estatísticas sobre o ensemble. Uma limitação de EnKF é o fato de que seus cálculos baseiam-se apenas em momentos de primeira e segunda ordem das variáveis de produção, e há problemas que são difíceis de tratar quando as distribuições de probabilidade destas variáveis são multi-modais (Haugen et. al., 2008). Apesar de questões teóricas e práticas sobre a aplicabilidade do EnKF para problemas reservatório altamente não-lineares e não-gaussianos, tem sido encontrados resultados promissores do uso do EnKF como uma ferramenta de ajuste ao histórico (Aanonsen et al., 2009).

Figura 5.11: Ajuste ao histórico tradicional e o baseado em EnKF



(Liang, 2007)

5.8 Recursos Computacionais

Como o tempo de processamento e a demanda por memória são as principais dificuldades em simulações computacionais de problemas de grande porte um sistema computacional para ajuste ao histórico deve utilizar técnicas para a redução do esforço computacional requerido pelos processos de ajuste e de propagação de incerteza. Da mesma forma, uma vez que decisões críticas serão tomadas baseadas nas informações geradas pelas simulações numéricas, é necessário que o sistema forneça dados confiáveis. Nesta dissertação foram tratados apenas problemas acadêmicos de pequeno porte, portanto foi usado um computador pessoal convencional, com um processador Intel(R) core (TM) i7-3930k CPU 3.80GHz com sistema operacional Windows 7 e memória RAM de 8,00 GB, para realizar as simulações. Atualmente, problemas com até alguns milhares de células computacionais são considerados de pequeno porte.

CAPÍTULO 6

6. Estudos de Caso

6.1 Introdução

A metodologia do EnKF foi aplicada a dois casos com características distintas. Cada um dos casos possui peculiaridades no que se refere à complexidade de sua geometria, heterogeneidade e esquema de produção. A permeabilidade e a porosidade são os principais parâmetros petrofísicos para o modelo de simulação de reservatório, e têm um impacto significativo sobre as reservas, previsões de produção e variáveis econômicas dos reservatórios. Os principais parâmetros de saída são saturações e pressões para cada célula nos intervalos de tempo desejados e as variáveis em observação são a produção de óleo e corte de água (WCT).

O EnKF é usado juntamente com o simulador de reservatório MRST nestes dois casos para a estimativa dos parâmetros petrofísicos do reservatório e criar um conjunto de modelos ajustados que forneça uma previsão adequada da produção futura do reservatório, com uma quantificação da sua incerteza.

O objetivo principal desses casos é estudar e avaliar o EnKF dentro de condições simplificadas e de fácil controle.

6.2 Caso 1

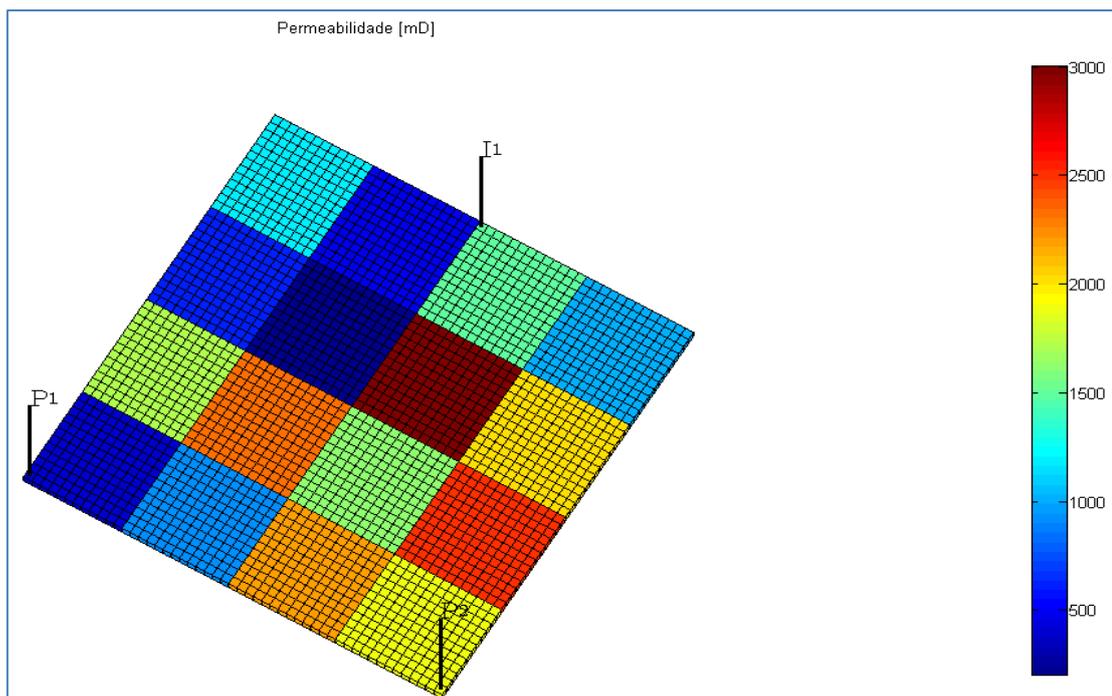
O caso 1 é um problema modelo artificial, mas cuja simplicidade permite a criação de situações, sua finalidade principal é verificar a correta implementação do método e das ferramentas desenvolvidas nesta dissertação.

O modelo de simulação do Caso 1 tem 51x51x1 células e apresenta as propriedades resumidas na Tabela 6.1.

Tabela 6.1: Resumo das características do modelo do Caso 1.

PROPRIEDADE	DADOS	
Malha de Simulação	51(510m)x51(510m)x1(4m)	
Porosidade	30%	
Permeabilidade horizontal (kh)	1000 mD	
Permeabilidade vertical (kv)	10% de kh	
Compressibilidade da Rocha @ 200 kgf/cm ²	5.10 ⁻⁵ (kgf/cm ²) ⁻¹	
Viscosidade	Água	Óleo
	0.97 cP	0.97 cP
Densidade	Água	Óleo
	965.284 kg/m ³	903.249 kg/m ³
Contatos entre Fluidos	Sem Contatos entre Fluidos	
Pressão de Saturação (Psat)	273 kgf/cm ²	
Razão Gás-Óleo de Formação (RGOF)	115.5 m ³ /m ³ std	

Esse modelo resulta num sistema com 7.083 graus de liberdade. O modelo usado em nosso estudo foi inspirado pelo usado na dissertação de Oliveira (2006), que é um reservatório com 1 poço injetor e 2 poços produtores como mostrado na Figura 6.1, mas que no modelo de Oliveira possui apenas três regiões de permeabilidade distintas. O campo de permeabilidades suposto como real é mostrado na Figura 6.1, e tem 16 regiões de permeabilidade, com valores variando entre 200 mD e 3000 mD.

Figura 6.1: Campo de permeabilidade inicial (Adaptado de Oliveira, 2006).

Foi considerado que a permeabilidade de cada bloco é uma realização de uma variável aleatória padrão independente, e foram geradas 10 realizações aleatórias da permeabilidade a partir do modelo inicial, usando um gerador de amostras Latin Hypercube Sampling (LHS). Cada variável aleatória tem densidade de probabilidade normal, sendo a média igual a 1200 mD, e desvio padrão igual a 400 mD. O reservatório foi considerado bifásico, isto é, com apenas óleo e água, e a produção de óleo ao longo do tempo foi calculada com o método implícito adaptativo disponível no simulador de reservatórios MRST.

Os poços produtores são restringidos por uma vazão máxima de líquido de $40\text{m}^3/\text{dia}$, sem ter nenhum controle secundário complementar, e o poço injetor foi controlado apenas pela vazão máxima de água injetada, definida em $44\text{m}^3/\text{dia}$. Esse valor foi adotado buscando-se a manutenção da pressão média do reservatório.

Para analisar o modelo do caso 1 foram consideradas três situações com diferentes números de passos de assimilação de dados (iterações). Em todas as situações a permeabilidade foi considerada ortotrópica, a permeabilidade horizontal igual nas duas direções e a permeabilidade vertical igual a dez por cento da horizontal. As variáveis observadas são a produção de óleo e corte de água (WCT). Nestas situações foi utilizado um período de 30 anos de produção, com 20 anos de histórico e 10 anos de previsão.

6.2.1 Resultados do caso 1.

Neste caso foram gerados os campos de permeabilidade, porosidade e pressão mostrados nas Figuras 6.2 a 6.5, que são gerados a partir da média dos ensembles, além, que os campos de permeabilidade são mostrados em momentos diferentes e são feitas com 10 realizações e 20 passos de tempo (iterações) de ajuste de histórico. As Figuras 6.6 a 6.11 mostram a produção de óleo acumulado, taxa de produção de óleo e corte de água para os poços produtores para todo o período de concessão.

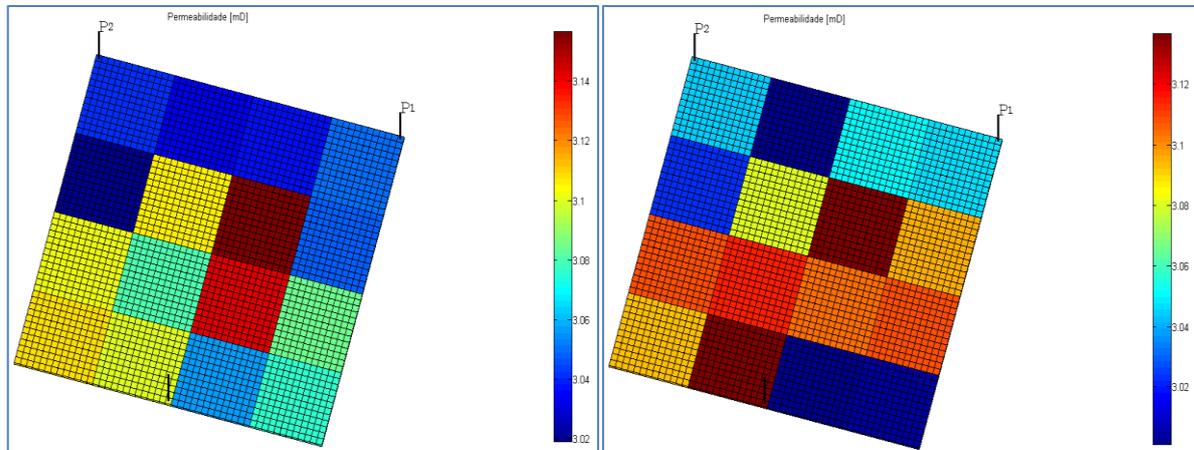
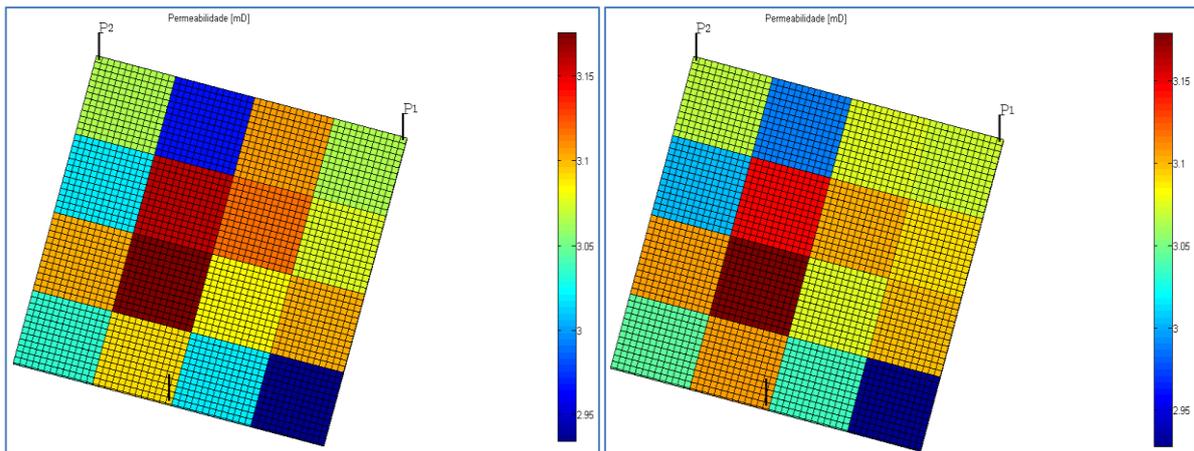
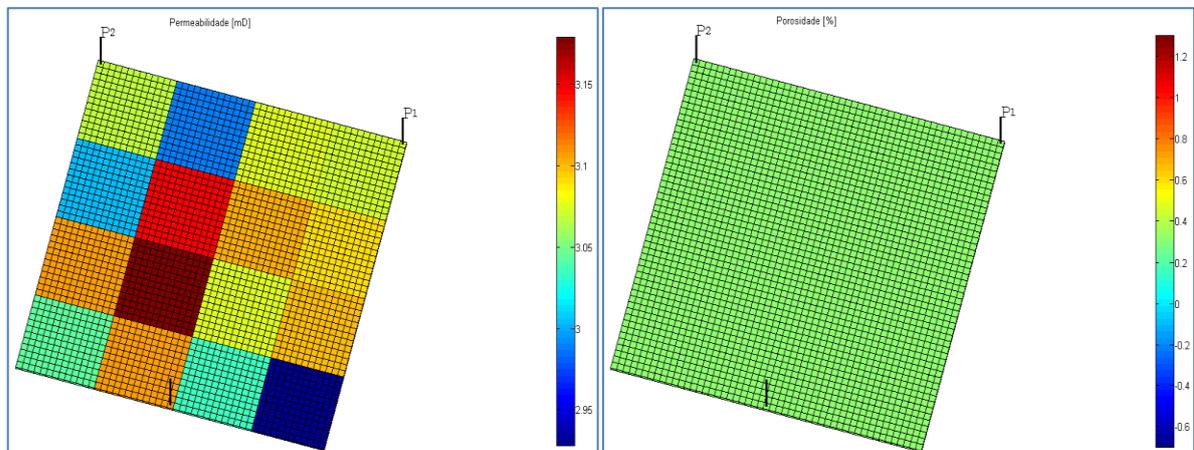
Figura 6.2: Campo Log-permeabilidade para 1825 dias e 3650 dias.**Figura 6.3:** Campo Log-permeabilidade para 5475 dias e 7300 dias.**Figura 6.4:** Campo Log-permeabilidade e Porosidade do ensemble médio.

Figura 6.5: Campo de Pressão e Saturação do ensemble médio.

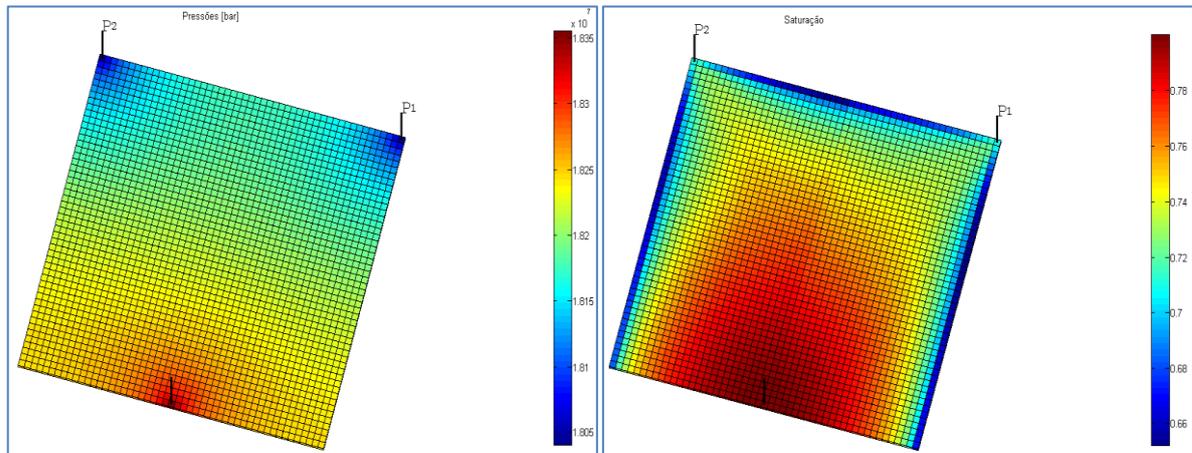


Figura 6.6: Produção de óleo acumulada. A linha vermelha corresponde ao modelo verdadeiro, a linha azul corresponde ao ensemble médio e as linhas cinzas às realizações.

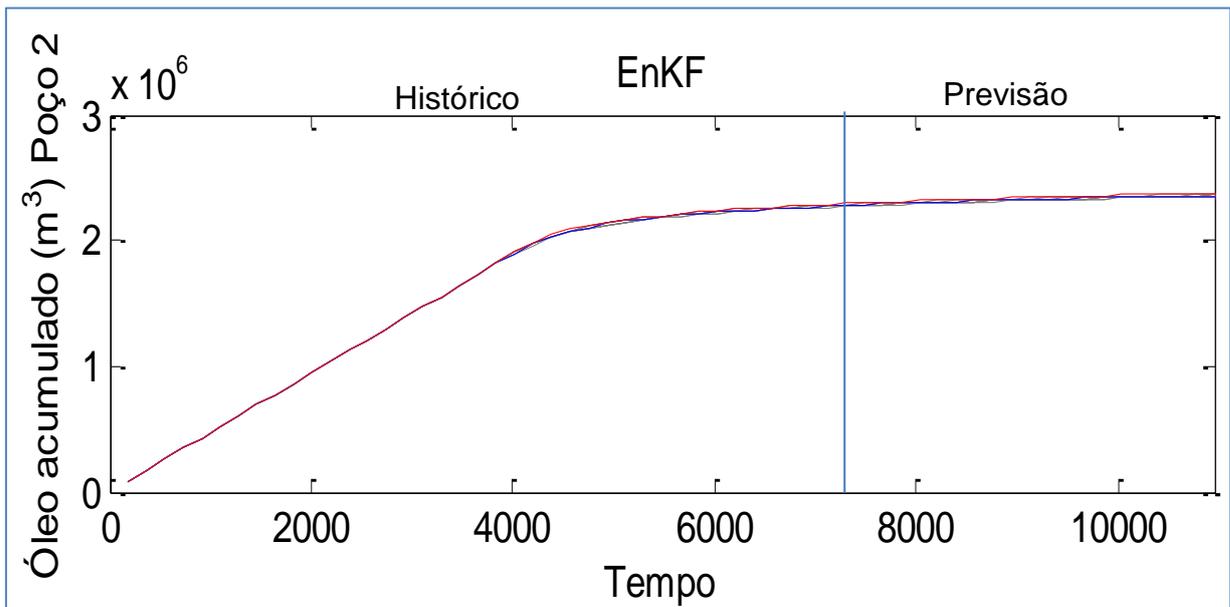


Figura 6.7: Produção de óleo acumulada. A linha vermelha corresponde ao modelo verdadeiro, a linha azul corresponde ao ensemble médio e as linhas cinzas às realizações.

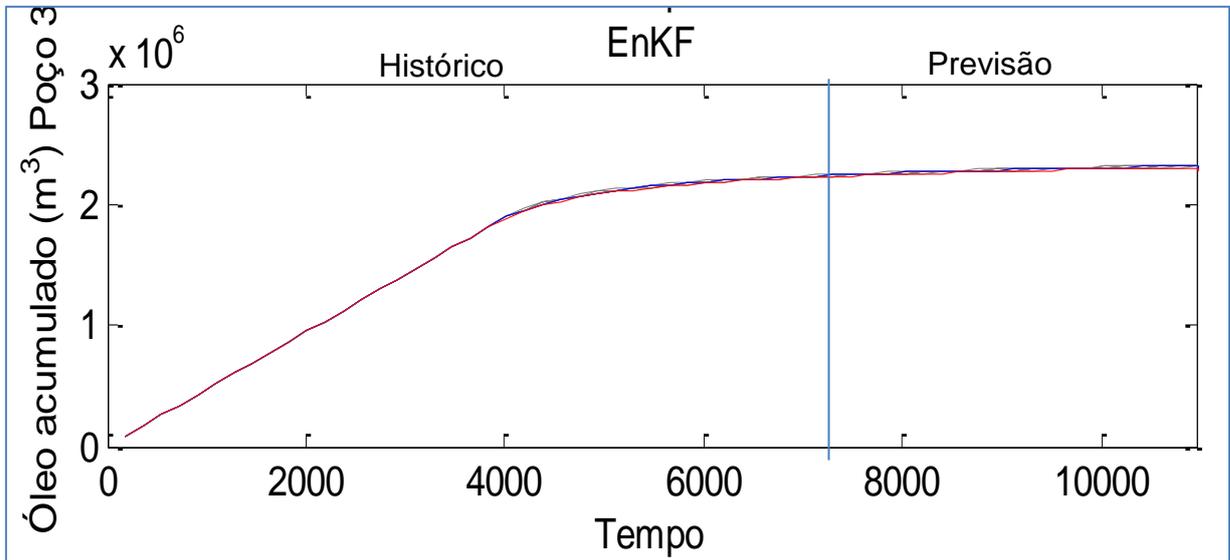


Figura 6.8: Taxa de Óleo. A linha vermelha corresponde ao modelo verdadeiro, a linha azul corresponde ao ensemble médio e as linhas cinzas às realizações.

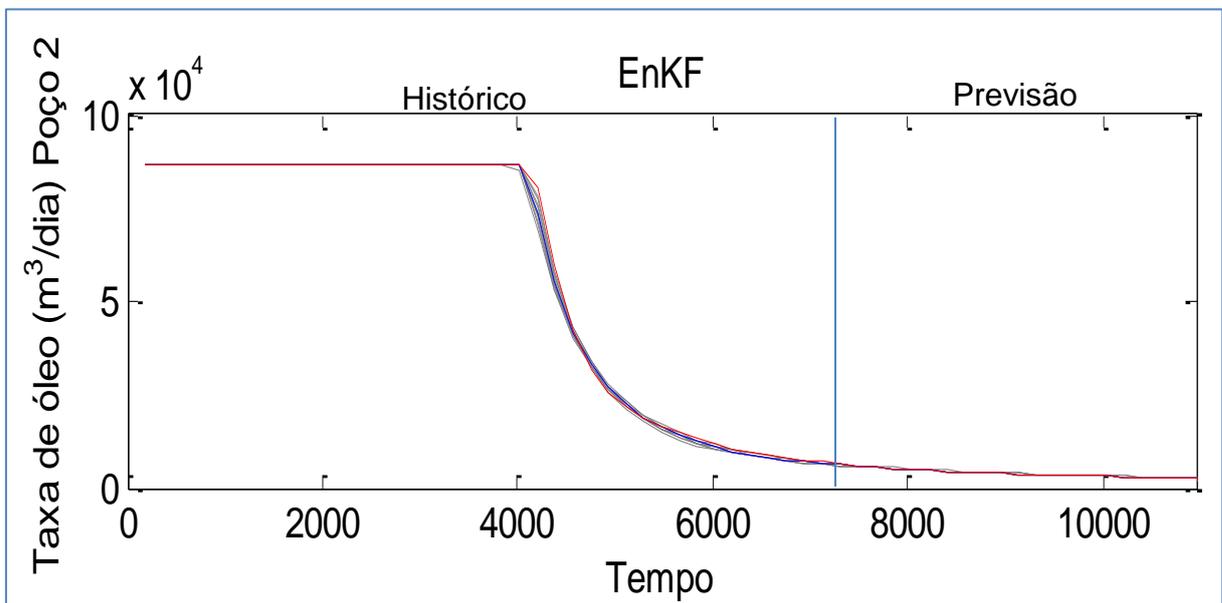


Figura 6.9: Taxa de Óleo. A linha vermelha corresponde ao modelo verdadeiro, a linha azul corresponde ao ensemble médio e as linhas cinzas às realizações.

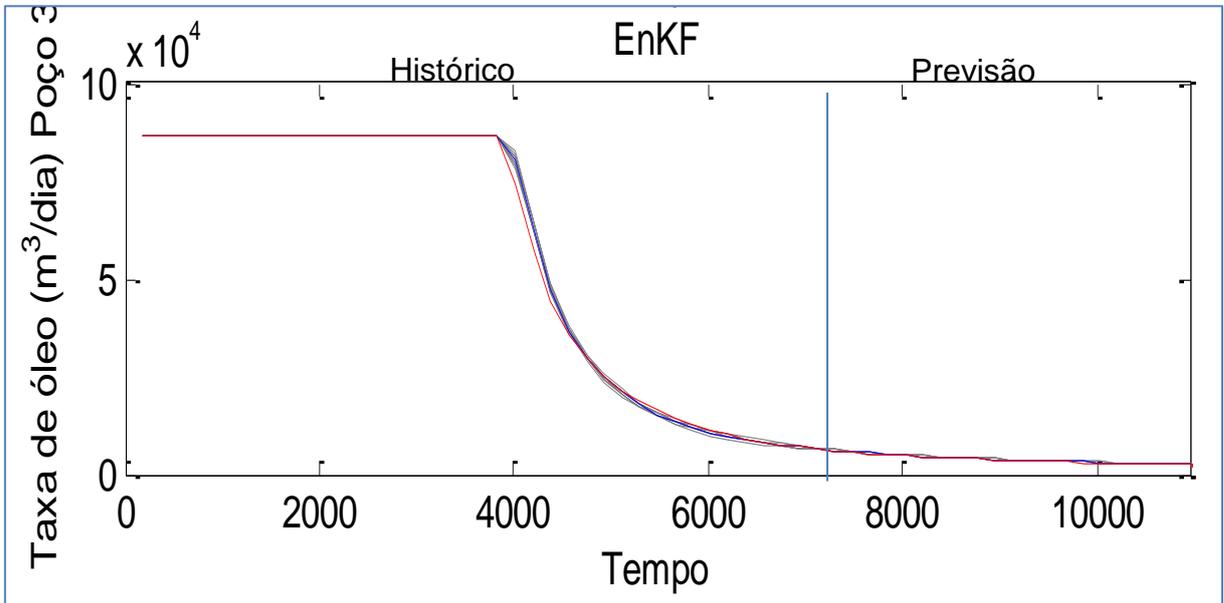


Figura 6.10: Corte de Água. A linha vermelha corresponde ao modelo verdadeiro, a linha azul corresponde ao ensemble médio e as linhas cinzas às realizações.

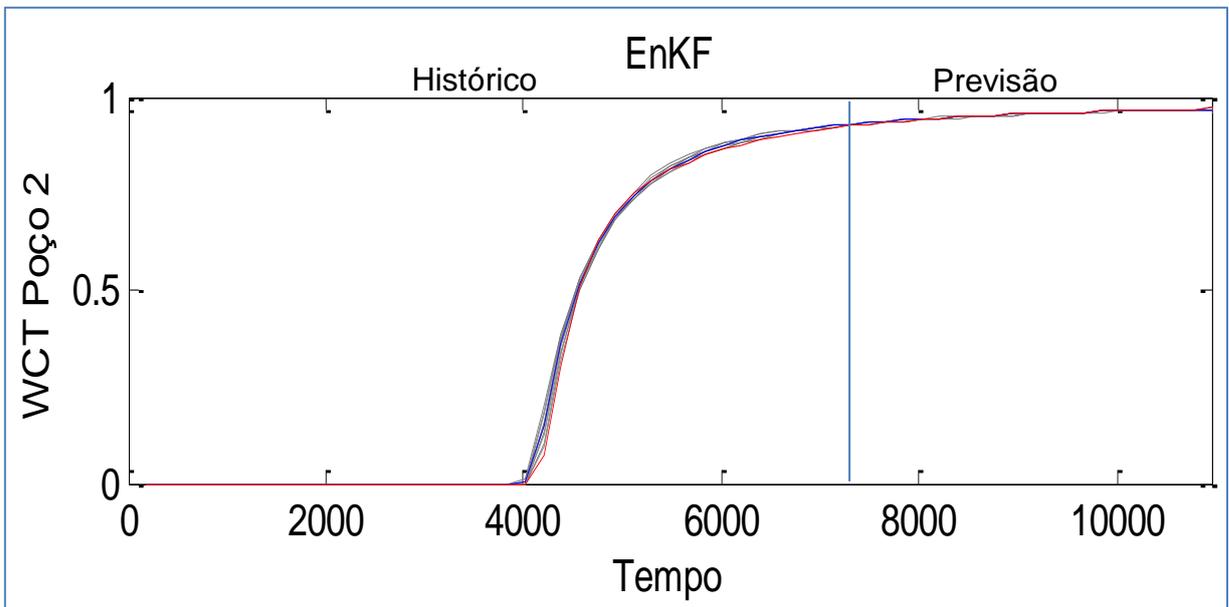
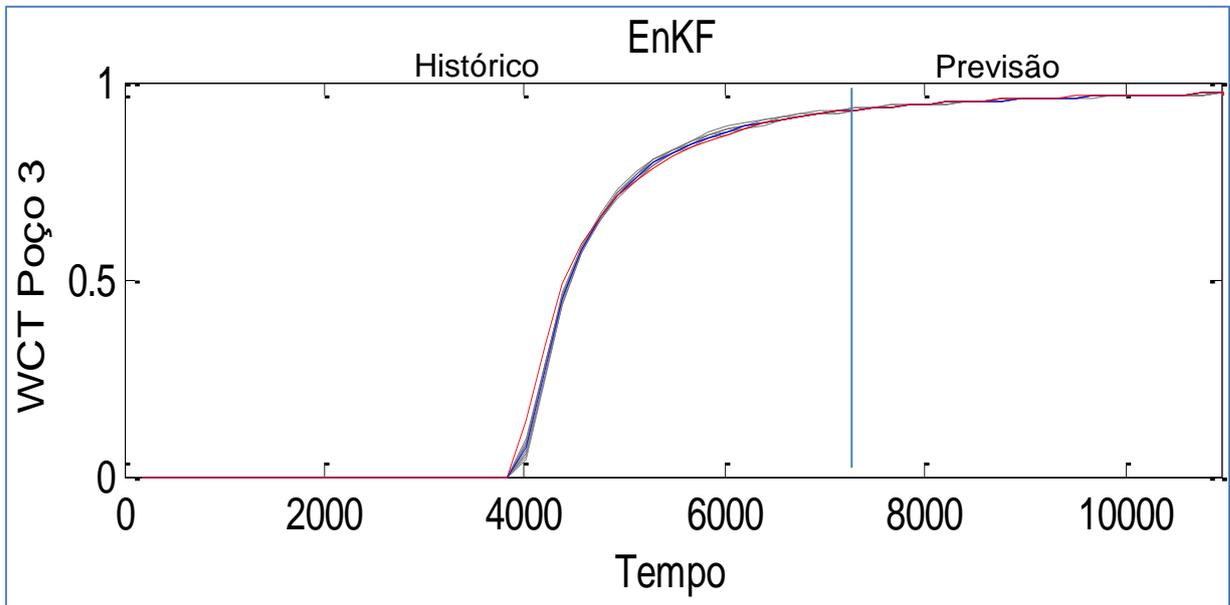


Figura 6.11: Corte de Água. A linha vermelha corresponde ao modelo verdadeiro, a linha azul corresponde ao ensemble médio e as linhas cinzas às realizações.



Após de realizar as simulações aplicando o método EnKF e realizar o ajuste no caso 1, obtendo-se as curvas de Óleo Acumulado, Fator de Recuperação e Corte de Água para os poços de produção, observamos que o método EnKF fornece modelos muito ajustados com relação ao modelo verdadeiro. Infelizmente este resultado não deve ser generalizado, pois as curvas de produção deste modelo são muito pouco sensíveis às variações de permeabilidade. A convergência dos históricos de produção no final do período de concessão deve-se mais ao esgotamento do reservatório do que ao procedimento de ajuste.

A Tabela 6.2 mostra o número de iterações e tempo de simulação para o EnKF, utilizado para realizar o ajuste de histórico durante o tempo de concessão.

Tabela 6.2: Resultados obtidos após das simulações.

Caso 1			
Histórico de produção: 20 anos		Tempo de produção: 30 anos	
Método	Realizações	Iterações	Tempo de simulação
EnKF	10	20	3019.68 seg.

6.3 Caso 2

O caso 2 é um caso sintético com propriedades mais realistas, que se compõe de um modelo geológico um pouco mais complexo, mais próximo de um modelo de reservatório real. O modelo tem um campo de permeabilidade aleatório Gaussiano, gerado no simulador de reservatórios MRST e os parâmetros petrofísicos foram obtidos pela relação de Carman – Kozeny, conforme pode ser visto no MRST (K,-A. Lie, 2012).

O modelo usa uma malha de 21x21x1 células, onde foram locados os 5 poços que compõem o plano de desenvolvimento proposto para esse caso. Tratam-se de 4 poços produtores e 1 poço injetor, estão situados nos quatros cantos e no meio do reservatório, respectivamente, posicionados segundo as características permo-porosas e de continuidade do reservatório. É um sistema de duas fases (água e óleo). Com fluidos incompressíveis e não há forças de capilaridade presentes.

Suas principais características estão resumidas na Tabela 6.3.

Tabela 6.3: Resumo das características do modelo do Caso 2.

PROPRIEDADE	DADOS	
Malha de Simulação	21(1000m)x21(1000m)x1(10m)	
Porosidade	Entre 10% e 26%	
Permeabilidade horizontal (kh)	Ver Mapas	
Permeabilidade vertical (kv)	10% de kh	
Compressibilidade da Rocha @ 200 kgf/cm ²	$5 \cdot 10^{-5} \text{ (kgf/cm}^2\text{)}^{-1}$	
Viscosidade	Água	Óleo
	0.40 cP	0.90 cP
Densidade	Água	Óleo
	1014 kg/m ³	859 kg/m ³
Contatos entre Fluidos	Sem Contatos entre Fluidos	
Pressão de Saturação (P _{sat})	350 kgf/cm ²	
Razão Gás-Óleo de Formação (RGOF)	115.5 m ³ /m ³ std	

Como no caso anterior, a permeabilidade foi considerada ortotrópica, a permeabilidade horizontal igual nas duas direções e a permeabilidade vertical igual a dez por

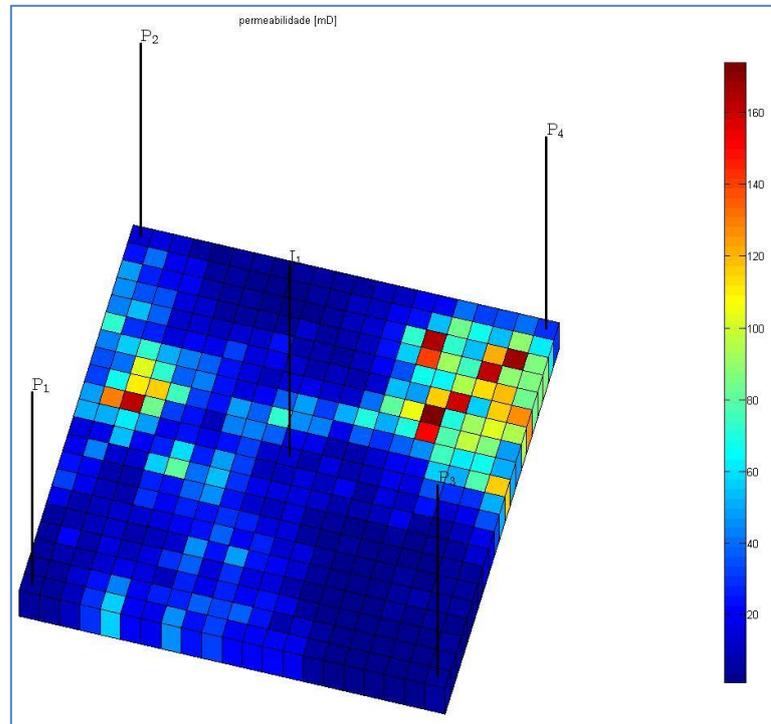
cento da horizontal. As variáveis a ser observadas são a produção de óleo e corte de água (WCT), que são medidos a cada 30 dias. As medições são criadas adicionando algum ruído Gaussiano, com média zero e desvio padrão igual a 10% do valor produzido no campo real no instante de tempo correspondente, para simular erros de medição.

O tempo total de produção é de 30 anos, onde foi utilizado um histórico de produção de um período de 20 anos e 10 anos de previsão. As vazões de produção e injeção foram distribuídas equitativamente entre os poços. Assim, todos os poços produtores estarão produzindo a uma mesma vazão de $150\text{m}^3/\text{dia}$, além disso, a pressão do fundo de poço (BHP) é de 300 bar para os poços produtores e de 350 bar para o poço injetor, desprezando a gravidade.

Foi gerado um ensemble inicial com 50 realizações, considerando-se uma permeabilidade média de 90 mD e desvio padrão de 47 mD, com densidade de probabilidade normal. Este caso também é resolvido com o método adaptativo, que determina durante a execução, em cada célula, se a mesma será tratada de forma implícita ou explícita, de acordo com as características físicas da célula naquele instante de tempo, disponível no simulador de reservatórios MRST.

Este modelo será considerado aquele que descreve o Caso 2. O mapa de permeabilidade para o Caso 2 está apresentado na Figura 6.12, onde as permeabilidades estão entre 10 mD a 170 mD.

Figura 6.12: Campo de permeabilidade inicial.



6.3.1 Resultados do caso 2.

Similarmente ao caso anterior, neste caso foram gerados os campos de permeabilidade, porosidade e pressão mostrados nas Figuras 6.13 a 6.16, que são gerados a partir da média dos ensembles, além, que os campos de permeabilidade são mostrados em momentos diferentes e são feitas com 50 realizações e 50 passos de tempo (iterações) de ajuste de histórico. As Figuras 6.17 a 6.28 mostram a produção de óleo acumulado, taxa de produção de óleo e corte de água para os poços produtores para todo o período de concessão. Obtendo-se os seguintes resultados.

Figura 6.13: Campo Log-permeabilidade para 1825 dias e 3650 dias.

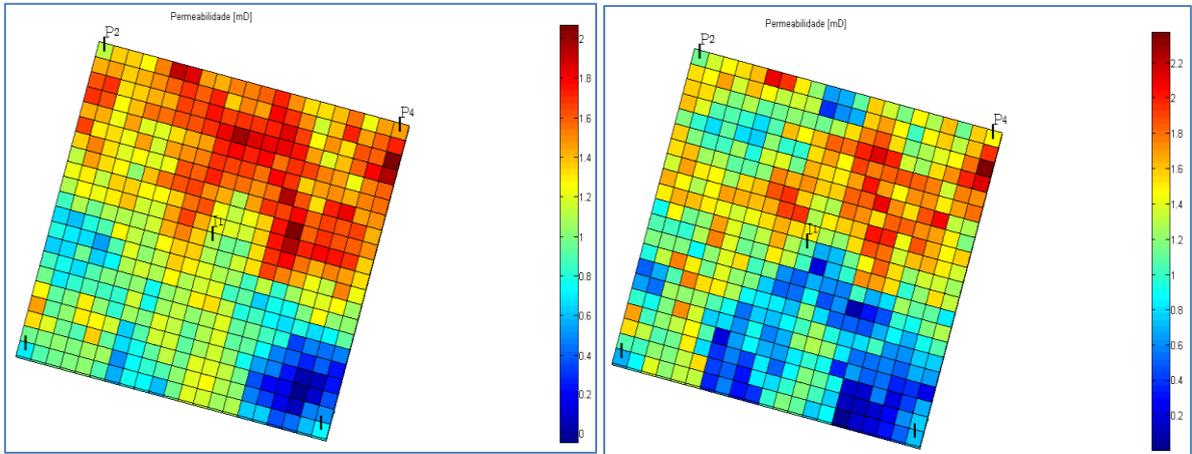


Figura 6.14: Campo Log-permeabilidade para 5475 dias e 7300 dias.

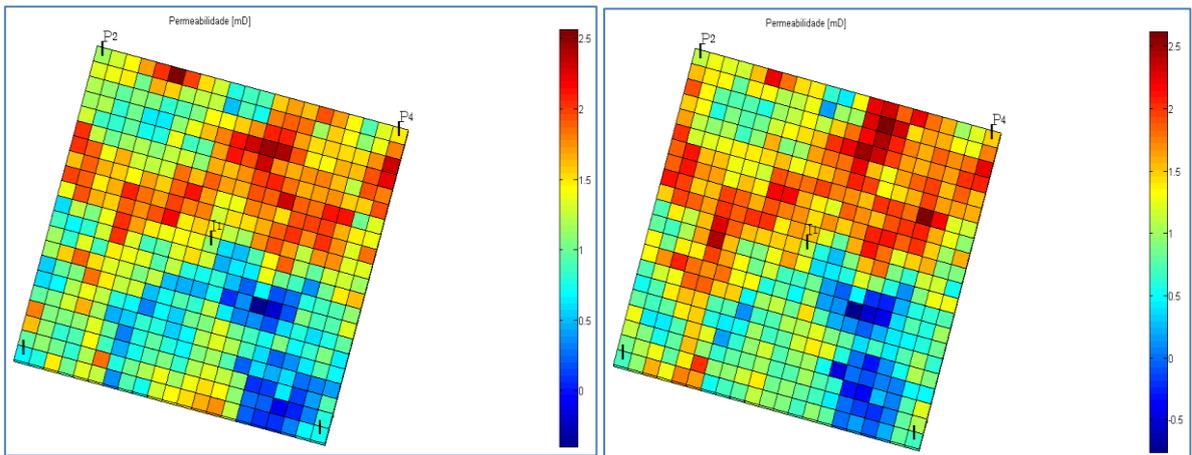


Figura 6.15: Campo de Log-permeabilidade e porosidade do ensemble médio.

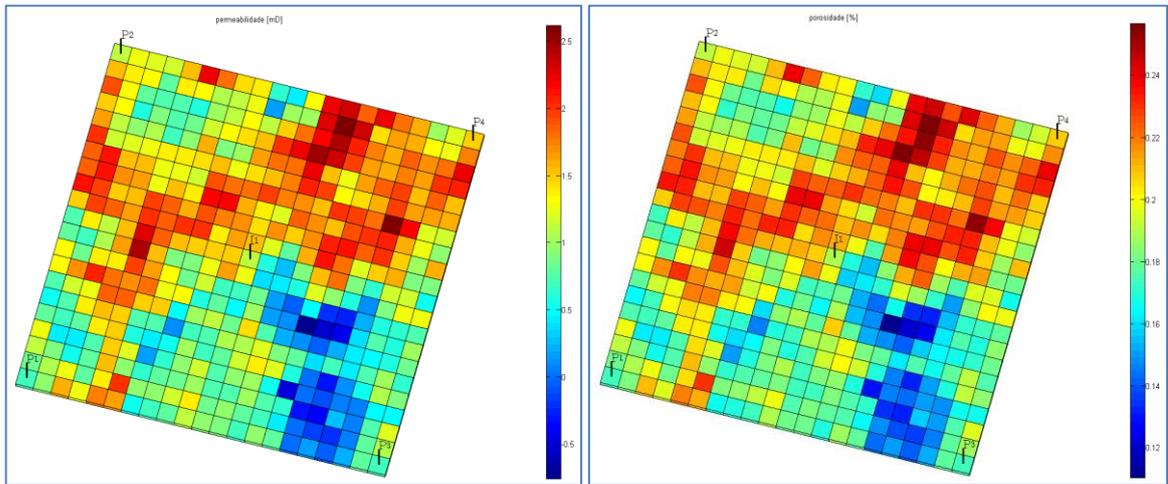


Figura 6.16: Campo de Pressão e Saturação do ensemble médio.

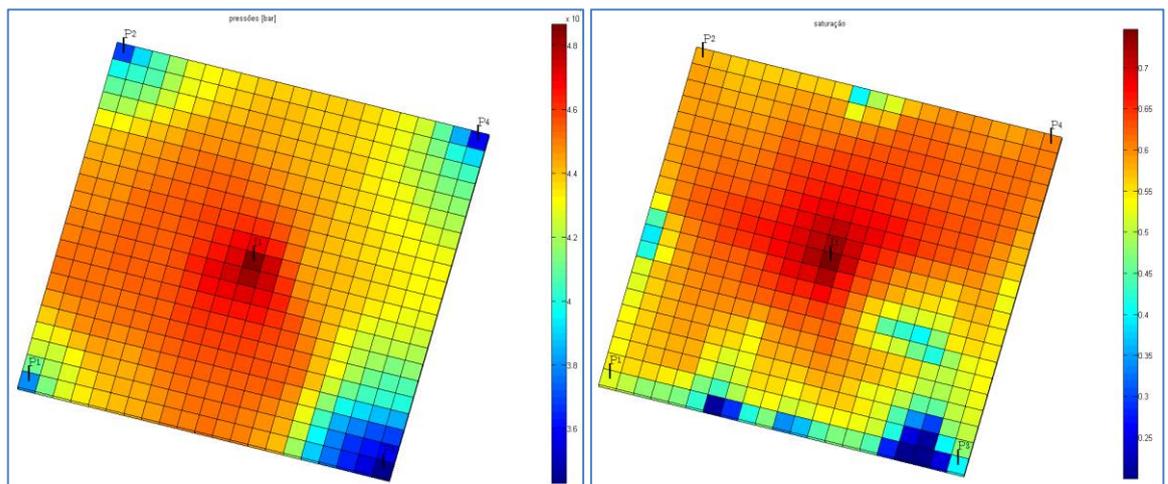


Figura 6.17: Produção de óleo acumulada. A linha vermelha corresponde ao modelo verdadeiro, a linha azul corresponde ao ensemble médio e as linhas cinzas às realizações.

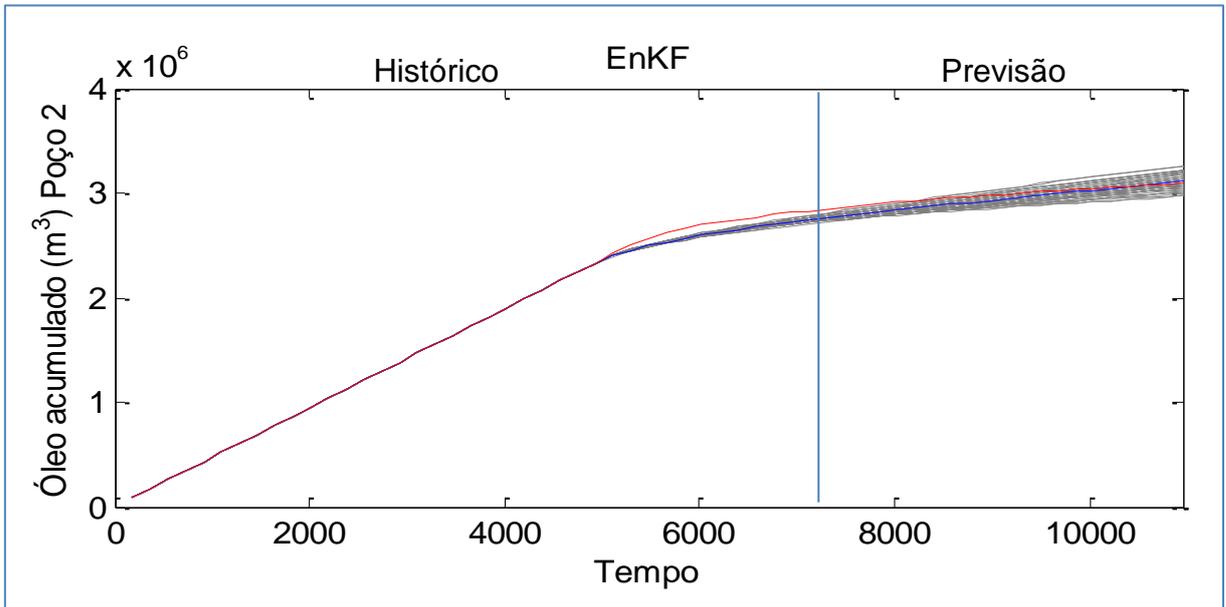


Figura 6.18: Produção de óleo acumulada. A linha vermelha corresponde ao modelo verdadeiro, a linha azul corresponde ao ensemble médio e as linhas cinzas às realizações.

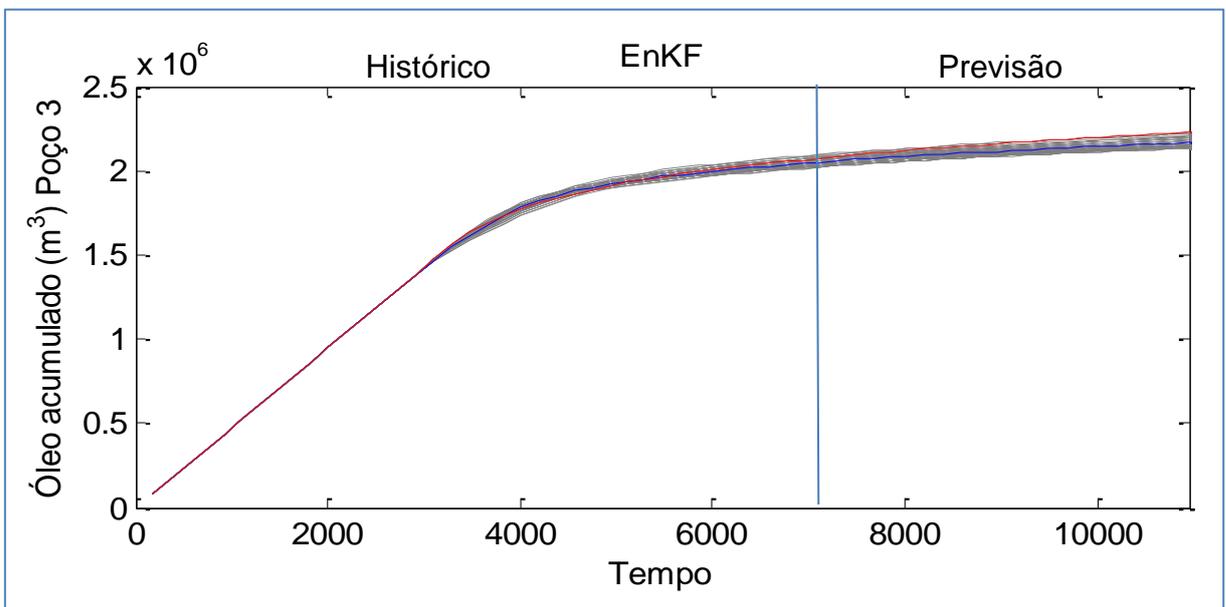


Figura 6.19: Produção de óleo acumulada. A linha vermelha corresponde ao modelo verdadeiro, a linha azul corresponde ao ensemble médio e as linhas cinzas às realizações.

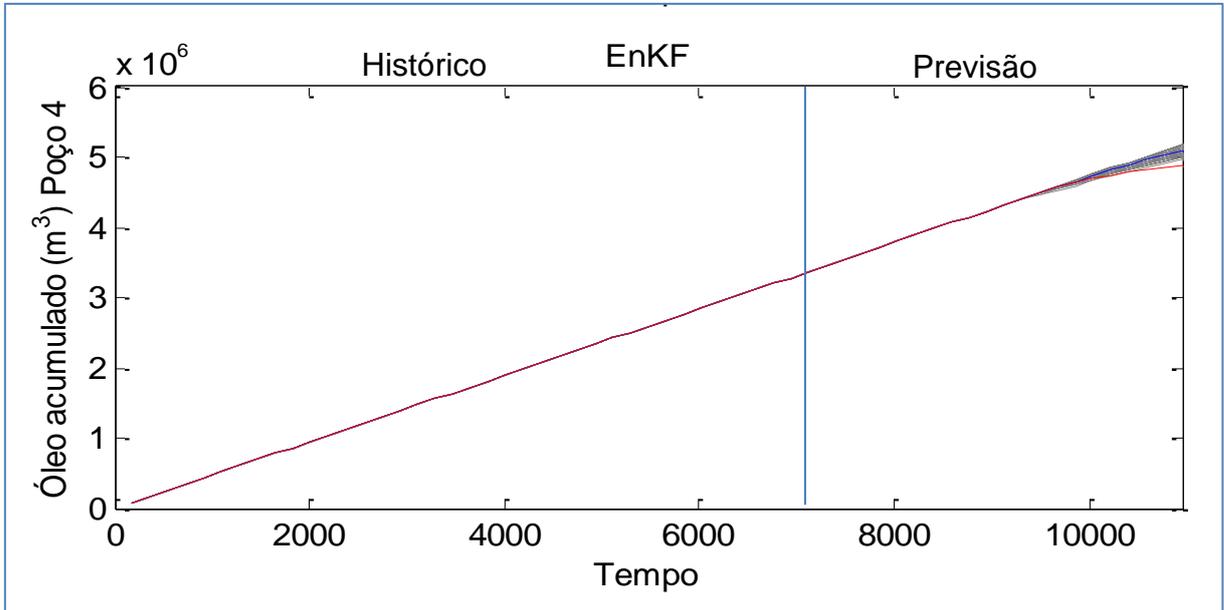


Figura 6.20: Produção de óleo acumulada. A linha vermelha corresponde ao modelo verdadeiro, a linha azul corresponde ao ensemble médio e as linhas cinzas às realizações.

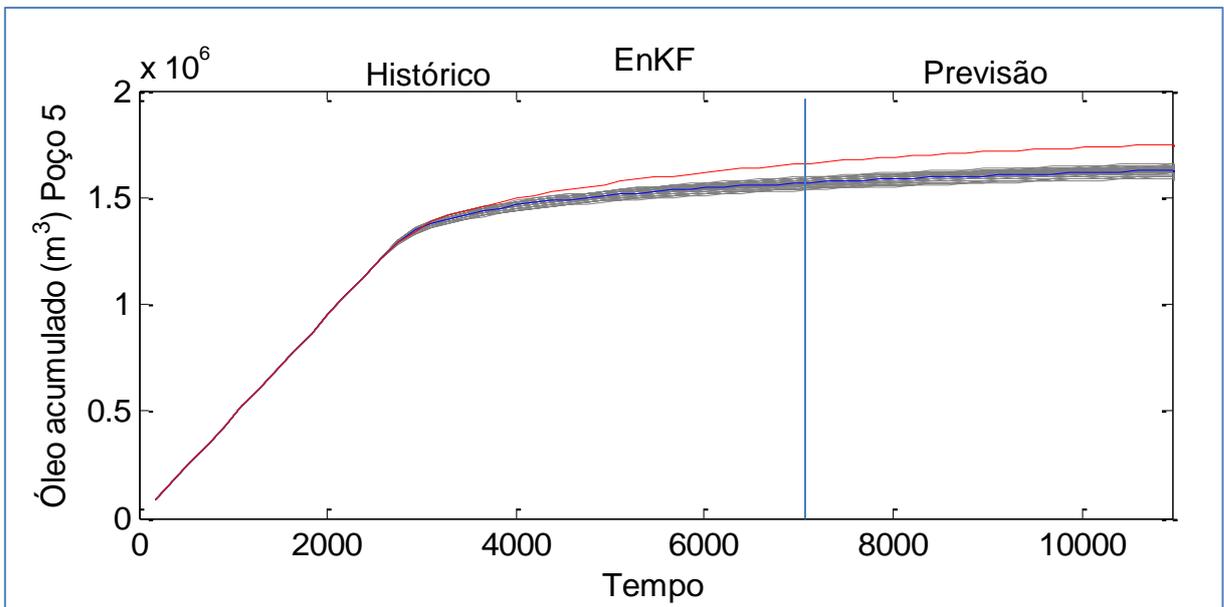


Figura 6.21: Taxa de Óleo. A linha vermelha corresponde ao modelo verdadeiro, a linha azul corresponde ao ensemble médio e as linhas cinzas às realizações.

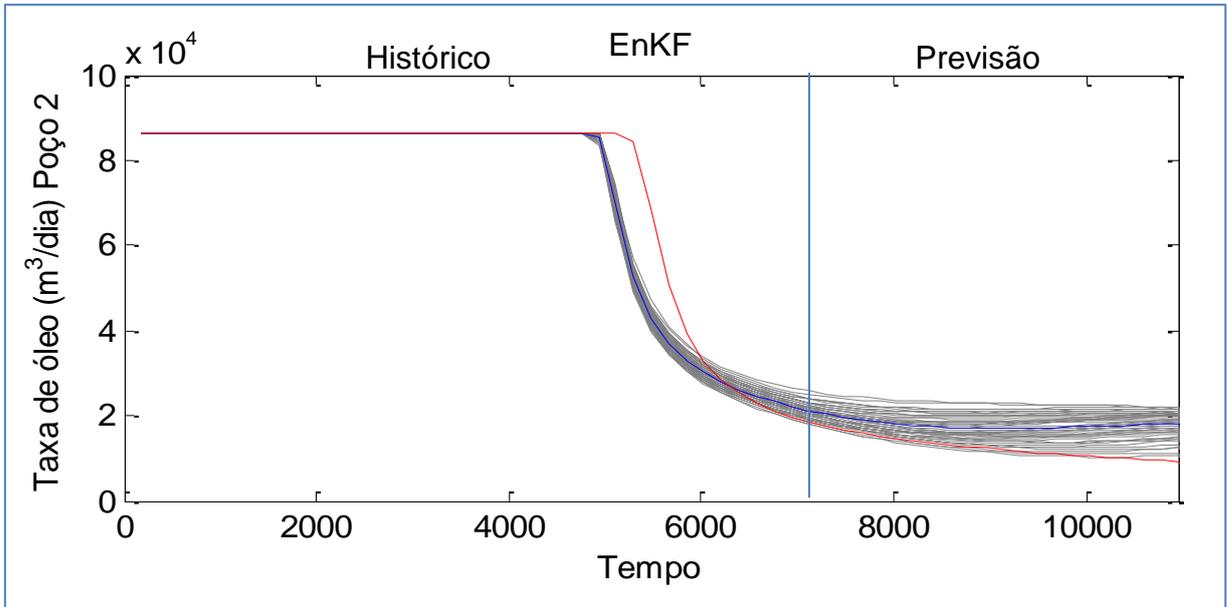


Figura 6.22: Taxa de Óleo. A linha vermelha corresponde ao modelo verdadeiro, a linha azul corresponde ao ensemble médio e as linhas cinzas às realizações.

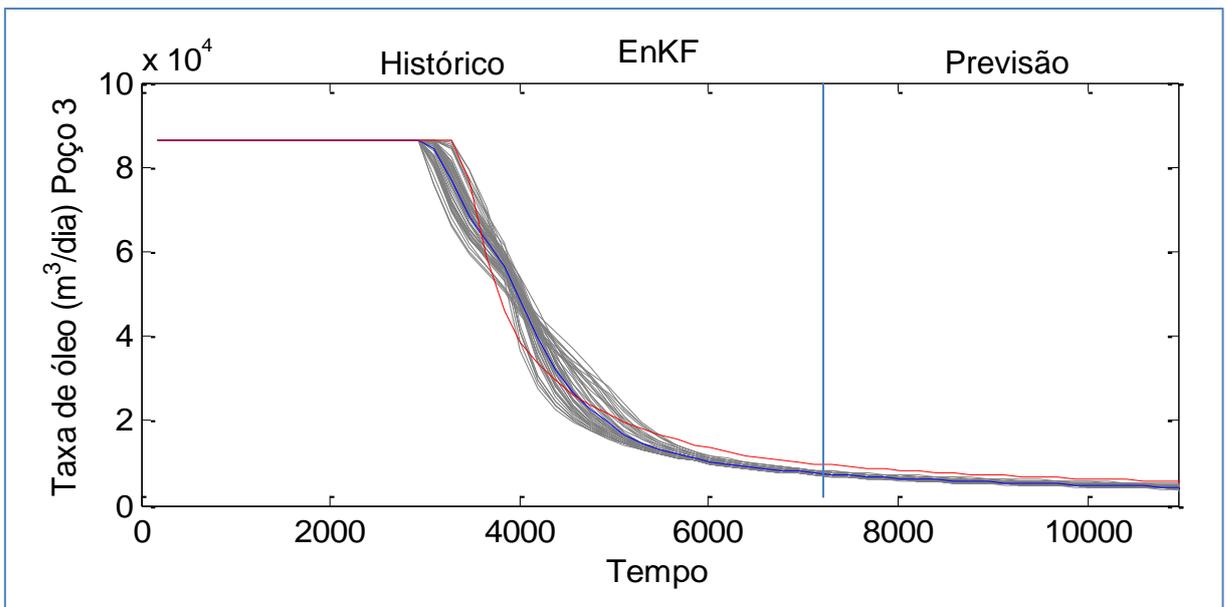


Figura 6.23: Taxa de Óleo. A linha vermelha corresponde ao modelo verdadeiro, a linha azul corresponde ao ensemble médio e as linhas cinzas às realizações.

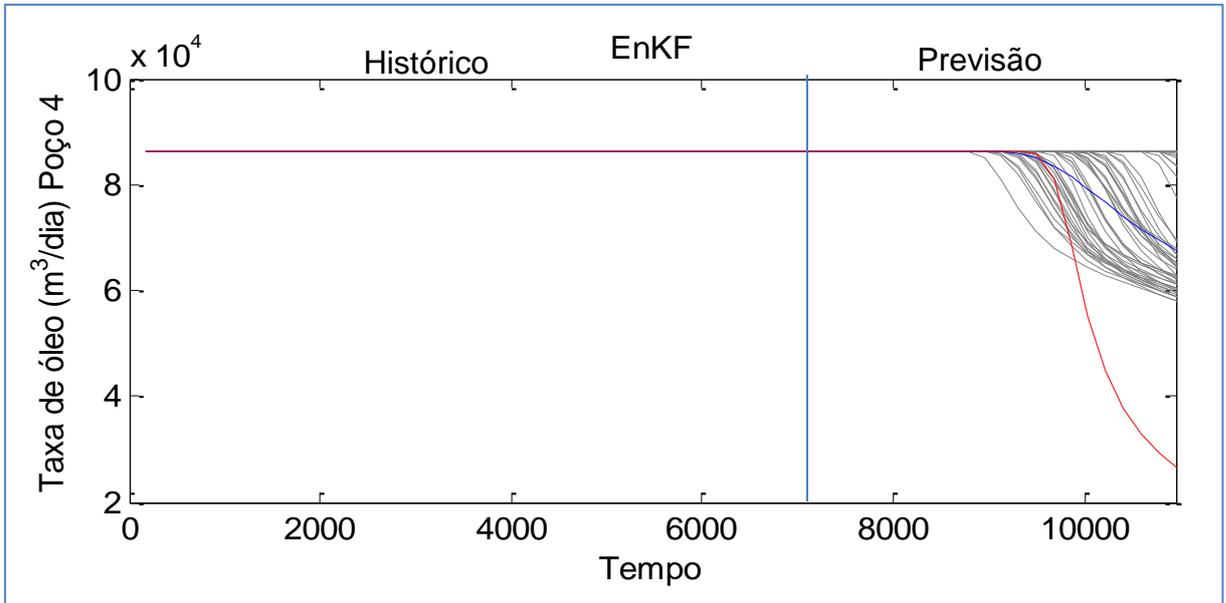


Figura 6.24: Taxa de Óleo. A linha vermelha corresponde ao modelo verdadeiro, a linha azul corresponde ao ensemble médio e as linhas cinzas às realizações.

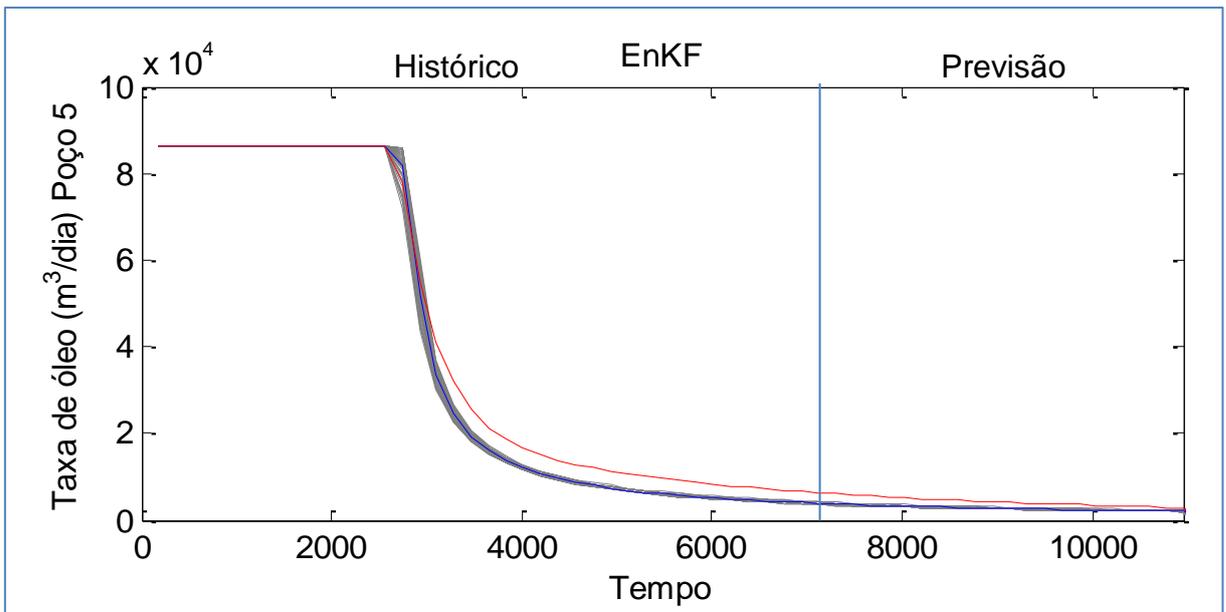


Figura 6.25: Corte de Água. A linha vermelha corresponde ao modelo verdadeiro, a linha azul corresponde ao ensemble médio e as linhas cinzas às realizações.

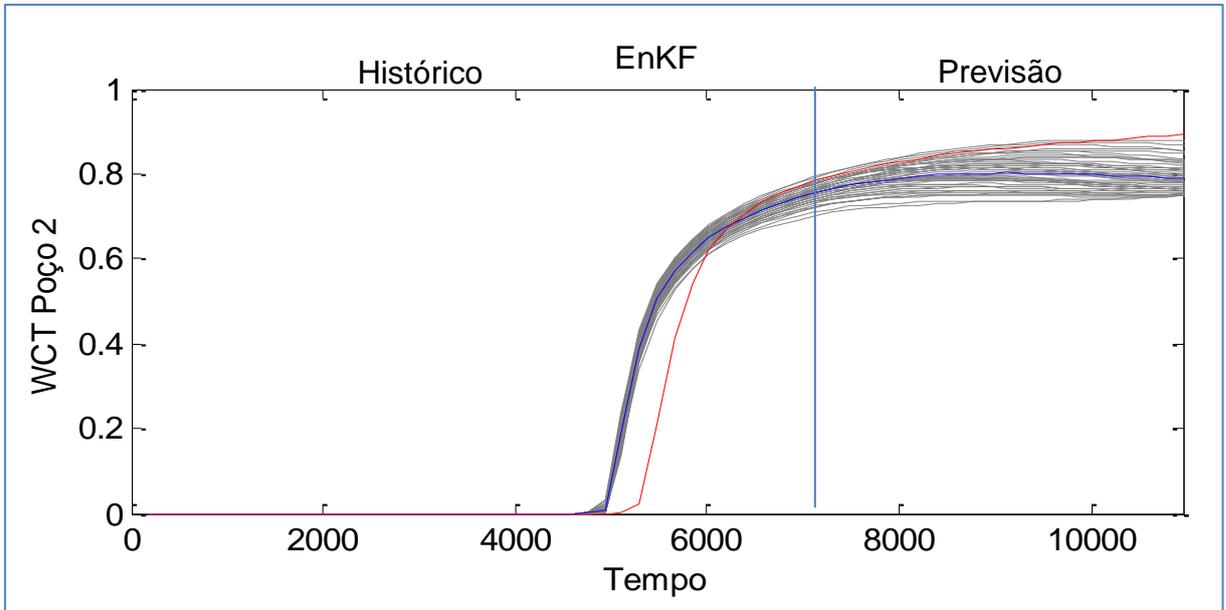


Figura 6.26: Corte de Água. A linha vermelha corresponde ao modelo verdadeiro, a linha azul corresponde ao ensemble médio e as linhas cinzas às realizações.

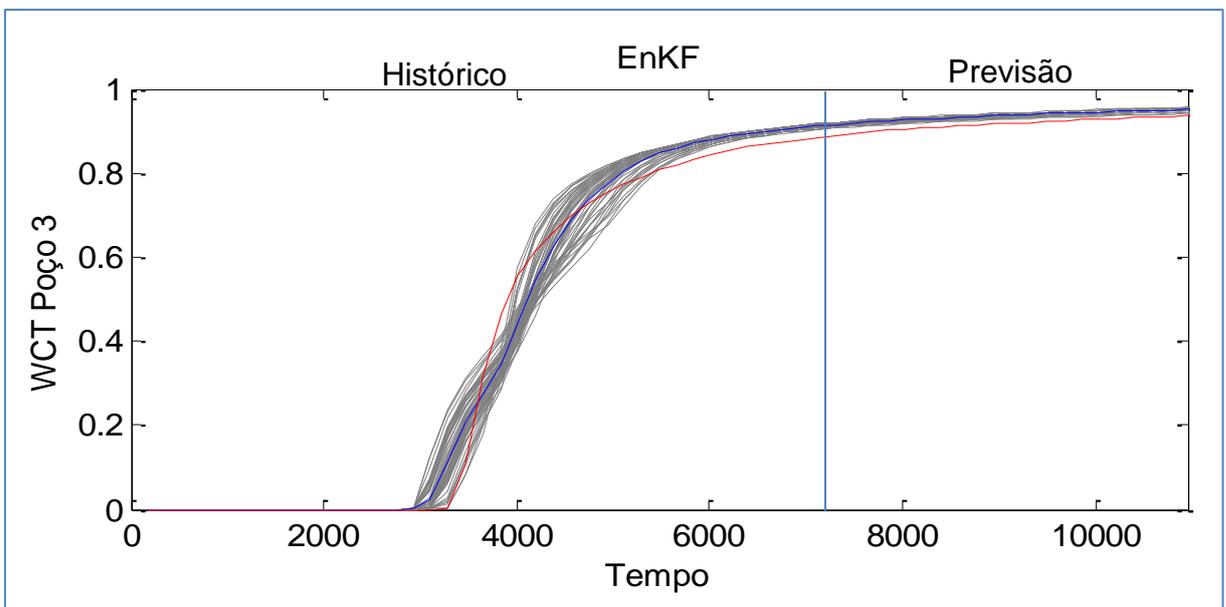


Figura 6.27: Corte de Água. A linha vermelha corresponde ao modelo verdadeiro, a linha azul corresponde ao ensemble médio e as linhas cinzas às realizações.

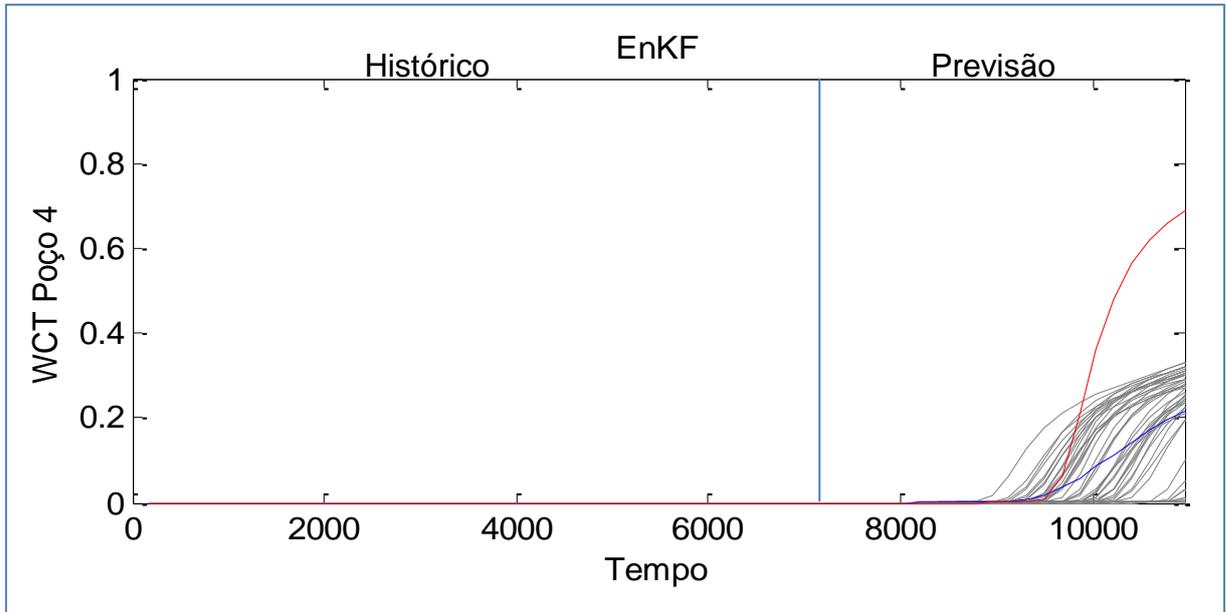
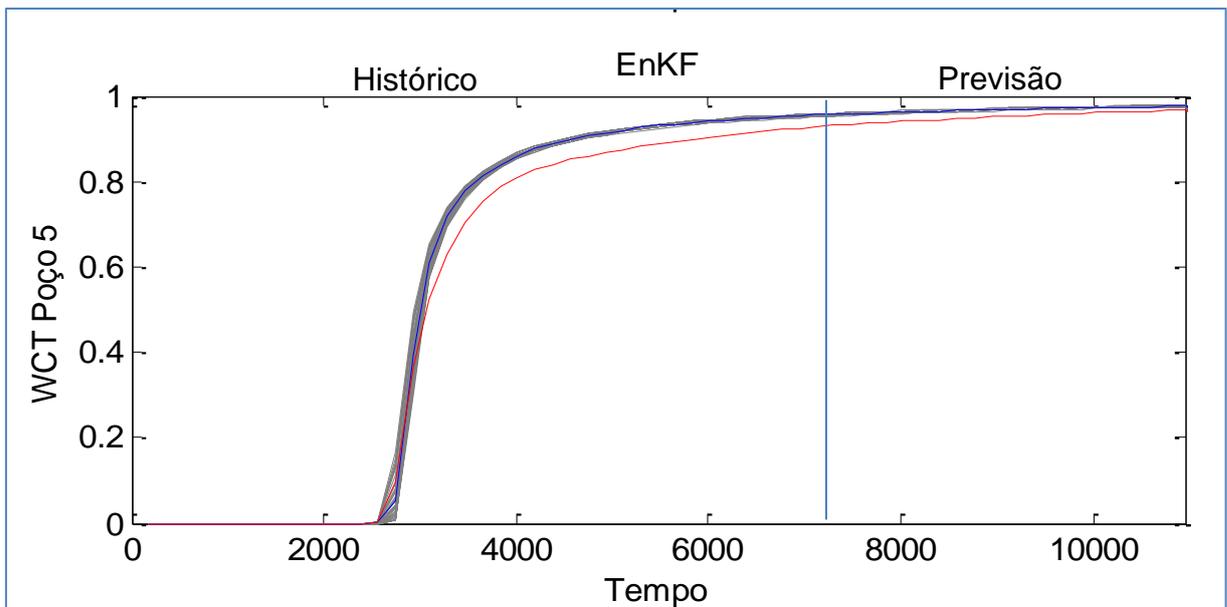


Figura 6.28: Corte de Água. A linha vermelha corresponde ao modelo verdadeiro, a linha azul corresponde ao ensemble médio e as linhas cinzas às realizações.



Observamos que, o quarto poço produtor apresenta resultados não aceitáveis. Isso ocorre porque ainda não haver irrupção de água no período de ajuste. Além disso,

observamos que o método EnKF gera modelos mais aceitáveis nos outros poços produtores ao longo da vida do campo, as outras figuras mostram os campos estimados para a permeabilidade, porosidade, pressão e saturação após a execução do EnKF. Na qual mostram um aumento de porosidade ao final da atualização com o EnKF, na qual foram utilizadas diferentes tempos de assimilação de dados.

A Tabela 6.4, mostra o número de iterações e tempo de simulação do método EnKF, utilizados para realizar o ajuste ao histórico no caso 2.

Tabela 6.4: Resultados obtidos após das simulações.

Caso 2			
Histórico de produção: 20 anos		Tempo de produção: 30 anos	
Método	Realizações	Iterações	Tempo de simulação
EnKF	50	50	7821.84 seg.

CAPÍTULO 7

7. Conclusão e Trabalhos Futuros

7.1 Conclusão

Atualmente as companhias petroleiras estão se esforçando para atender à demanda crescente de energia. Portanto é importante desenvolver novas tecnologias e estratégias para otimizar a recuperação de óleo em campos existentes e nos novos campos que entram em produção. Qualquer que seja a tecnologia ou estratégia, esta requer um melhor conhecimento e caracterização do reservatório. Isto é conseguido através do ajuste de parâmetros incertos do modelo que descreve o reservatório, tais como a permeabilidade, usando o ajuste ao histórico ou assimilação de dados. O EnKF mostrou-se uma ferramenta útil para esta assimilação. Reduzir o tempo de assimilação de dados em todo o ciclo de gestão do reservatório é uma questão importante.

Nesta dissertação foi descrito o processo do ajuste ao histórico baseado no método EnKF, aplicado a reservatórios de petróleo. Abaixo resumimos outras questões importantes relevantes para a implementação de EnKF em reservatórios de petróleo.

1. O EnKF é fácil de implementar e seu desempenho é muito satisfatório em comparação com outros métodos que são usados para o ajuste ao histórico em reservatórios de petróleo.
2. Foram usadas de 10 até 50 realizações, para avaliar cada um dos casos estudados neste trabalho e a eficiência do EnKF para o ajuste foi aceitável.
3. O EnKF tradicional pode fazer a assimilação de dados, minimização de variância e fornece previsões adequadas que utilizam resultados de previsões anteriores e as medições passadas e presentes.
4. Aumentando o número de realizações proporcionalmente aumenta o tamanho do problema para a etapa de previsão e o tempo de execução é proporcionalmente aumentado, no entanto, o tamanho do problema para a etapa de análise permanece praticamente o mesmo e o tempo de execução não se altera significativamente.
5. O simulador de reservatórios com código livre MRST pode ser modificado com facilidade e adaptado para outros métodos e técnicas.

6. Partir dos resultados apresentados, o trabalho de implementação do EnKF usando o simulador MRST, foi bem sucedido.

7.2 Trabalhos futuros

Durante este trabalho algumas ideias para desenvolvimentos futuros foram levantadas, como por exemplo.

1. Acoplar o método EnKF com outros métodos estocásticos, como o método de Colocação Estocástica (Dongxiao et. al. 2011).
2. Estender o código do MRST para explorar as possibilidades de uso iterativo EnKF em caso de utilização de dados sísmicos.
3. A construção de uma interface de usuário e simplificar o processo de aplicação do filtro.
4. O sistema desenvolvido é acadêmico e seria interessante leva-lo para um estágio em que pudesse ser incorporado aos procedimentos mais avançados de gestão do reservatório.

Referências

- Aanonsen, S.; Nævdal, G.; Oliver, D.; Reynolds, A.; and Vallès, B.; *“The Ensemble Kalman Filter in Reservoir Engineering - a Review”*, SPE Journal, SPE 117274, 2009.
- Anderson, J.L., *“An Ensemble Adjustment Kalman Filter for Data Assimilation”*. Monthly Weather Review, 129: 2884-2903. 2001.
- Fahimuddin Abul; *“4D Seismic History Matching Using the Ensemble Kalman Filter (EnKF): Possibilities and Challenges”*, PhD Thesis Department of Mathematics University of Bergen. March, 2010.
- April, J.; Glover, F.; Kelly, J.; Laguna, M.; Erdogen, M.; Mudford, B. and Stegemeier, D.; *“Advanced Optimization Methodology in the Oil and Gas Industry: The Theory of Scatter Search Techniques with Simple Examples”*. Paper presented at the SPE Hydrocarbon Economics and Evaluation Symposium, Dallas, TX. SPE 82009. 5 - 8 April, 2003a.
- April, J.; Glover, F.; Kelly, J. and Laguna, M.; *“A New Optimization Methodology for Portfolio Management”*. Paper presented at the SPE Annual Technical Conference and Exhibition, Denver, CO. SPE 84332. 5 - 8 October 2003b.
- Bishop, C.H., Etherton, B.J., and Majumdar, S.J., 2001. *“Adaptive Sampling with Ensemble Transform Kalman Filter”*. Part I: Theoretical Aspects. Monthly Weather Review, 129: 420-436.
- Brouwer, D.R., Nævdal, G., Jansen, J.D., Vefring, E.H., and Van Kruijsdijk, C.P.J.W., 2004. *“Improved Reservoir Management through Optimal Control and Continuous Model Updating”*. Paper presented at the SPE Annual Technical Conference and Exhibition, Houston, TX, 26-29 September, SPE 90149.
- Brown, R. G. and Hwang, P. Y. C.; *“Introduction to Random Signals and Applied Kalman Filtering: with MATLAB Exercises and Solutions”*. Wiley New York, ISBN 0-471-12839-2. 1997.
- Burgers, G.; Van Leeuwen, P. J. and Evensen, G.; *“Analysis Scheme in the Ensemble Kalman Filter”*. Mon. Weather Rev., 126:1719–1724, 1998.
- Christie, M.; Macbeth, C. and Subbey, S.; *“Multiple History-Matched Models for Teal South”*. The Leading Edge, 21(3): 286-289. 2002.
- Christie, M.; Demyanov, V. and Erbas, D.; *“Uncertainty Quantification of Porous Media Flows”*. Journal of Computational Physics, 217: 143-158. 2006.
- Coxeter, H.S.M., *“Regular Polytopes”*. 3rd ed. New York: Dover, p. 45. 1973.
- Cullick, A.S.; Johnson, D. and Shi, G.; *“Improved and More Rapid History Matching with a Nonlinear Proxy and Global Optimization”*. Paper presented at the SPE Annual

Technical Conference and Exhibition, San Antonio, TX, 24-27 September, SPE 101933. 2006.

Da Costa Freitas C.; “*Estatística – Curso I*”, Ministério da Ciência e Tecnologia Instituto Nacional de Pesquisas Espaciais, São José dos Campos 2003.

De Lira Junior J. D.; “*Otimização com Modelos Substitutos Considerando Incertezas em Reservatórios de Petróleo*”. Tese de Doutorado, Engenharia Civil, Universidade Federal de Pernambuco. Recife – Brasil. 2012.

Deutsch, C.V. and Journel, A.G.; “*GSLIB: Geostatistical Software Library and User's Guide*”; Oxford University Press, New York. 1998.

Dong, Y., Gu, Y., Oliver, D.S., 2006. “*Sequential Assimilation of 4D Seismic Data for Reservoir Description Using the Ensemble Kalman Filter*”. Journal of Petroleum Science and Engineering, 53 (1-2): 83-99.

Dongxiao Zhang, Heng Li, and Haibin Chang. “*History Matching for Non-Gaussian Random Fields Using the Probabilistic Collocation Based Kalman Filter*”. SPE, Peking University, 2011.

Du, Q., Vance F., Gunzburger; M.; “*Centroidal Voronoi Tessellations: Applications and Algorithms*”, Society for Industrial and Applied Mathematics, 1999.

Emerick, A. A. and Reynolds, A. C.; “*History Matching a Field Case Using the Ensemble Kalman Filter With Covariance Localization*”, SPE Reservoir Evaluation & Engineering, 14(4), 423-432, 2011.

Emerick, A. A. and Reynolds, A. C.; “*Combining the Ensemble Kalman Filter With Markov Chain Monte Carlo for Improved History Matching and Uncertainty Characterization*”, SPE Journal, in press, 2012.

Emerick, A. A., “*History Matching and Uncertainty Characterization Using Ensemble Based Methods*”, Ph. D. Thesis, Petroleum Engineering, University of Tulsa, 2012.

Ertekin, T.; Abou-Kassem, J. H. and King, G. R.; “*Basic Applied Reservoir Simulation*”. 1. ed. Rirchadson, Texas: Society of Petroleum Engineers, (Textbook, 7). 2001.

Evensen, G., 1994. “*Sequential Data Assimilation with a Nonlinear Quasi-Geostrophic Model using Monte Carlo Methods to Forecast Error Statistics*”. Journal of Geophys. Res. 99 (C5): 10143-10162.

Evensen, G.; “*Sampling Strategies and Square Root Analysis Schemes for the EnKF*”. Ocean Dyn., 54:539–560, 2004.

Evensen, G.; “*Data Assimilation: The Ensemble Kalman Filter*”. Springer: New York, 2007.

Evensen, G., Hove, J., Meisingset, E., Reiso, K.S., Seim and Espelid, Ø., 2007. “Using the EnKF for Assisted History Matching of a North Sea Reservoir Model”. Paper presented in the SPE Reservoir Simulation Symposium, Houston, TX, 26-28 February, SPE 106184.

Fahimuddin, A., “4D Seismic History Matching Using the Ensemble Kalman Filter (EnKF): Possibilities and Challenges”, Ph. D. Thesis, Department of Mathematics, University of Bergen, 2010.

Gao, G., Zafari, M., and Reynolds, A.C., 2005. “Quantifying Uncertainty for the PUNQ-S3 Problem in a Bayesian Setting with RML and EnKF”. Paper presented at SPE Reservoir Simulation Symposium, Houston, TX, 31 January-02 February, SPE 99324.

Gillijns, S.; Barrero M. O.; Chandrasekar, J. B; de Moor, L. R.; Bernstein, D. S. and Ridley, A.; “What is the Ensemble Kalman Filter and How Well Does it Work?”. In Proceedings of the 2006 American Control Conference, pages 4448–4453, 2006.

Giunta, A. A.; “Use of Data Sampling, Surrogate Models, and Numerical Optimization in Engineering Design”, Proceedings of the 40th AIAA Aerospace Sciences Meeting and Exhibit, Reno, NV, 2002.

Giunta, A. A., Wojtkiewicz, S. F., Eldred, M. S.; “Overview of Modern Design of Experiments Methods for Computational Simulations”, Proceedings of the 41st AIAA Aerospace Sciences Meeting and Exhibit, Reno, NV, 2003.

Glover, F.; “Heuristics for Integer Programming Using Surrogate Constraints”. Decision Sciences, 8: 156-166. 1977.

Glover, F.; “Genetic Algorithms and Scatter Search: Unsuspected Potentials”. Statistics and Computing, 4: 131-140. 1994.

Glover, F. and Laguna, M.; “Tabu Search”. Kluwer Academic Publishers, Boston. 1997.

Gu, Y., and Oliver, D.S., 2005. “History Matching of the PUNQ-S3 Reservoir Model Using the Ensemble Kalman Filter”. Society of Petroleum Engineering Journal 10 (2) 217- 224.

Halton, J. H.; “A Retrospective and Prospective Survey of the Monte Carlo Method”. SIAM, 1970.

Haugen, V., Natvik, L.; Evensen, G.; Berg, A.; Flornes, K. and Nævdal, G., “History Matching Using the Ensemble Kalman Filter on a North Sea Field Case”. SPE Journal, v. 13, no. 4, p. 382–391. 2006.

Haugen, V. and Nævdal, G.; “History Matching Using the Ensemble Kalman Filter on a North Sea Field Case”, SPE Journal, SPE 102430, 2008.

Holland, J. H.; *“Adaptation in Natural and Artificial Systems”*. University of Michigan Press, Ann Arbor. 1975.

Houtekamer, P.L., and Mitchell, H.L., 1998. *“Data assimilation using an ensemble Kalman filter technique”*. Monthly Weather Review, 126: 796-811.

Jazwinski, A. H.; *“Stochastic Processes and Filtering Theory”*. Academic Press New York, ISBN 3-540-38300-X. 1970.

Jensen, J. P.; *“Ensemble Kalman Filter for State and Parameter Estimation on a Reservoir Model”*. M.Sc. Thesis paper submitted in the Department of Science in Engineering Cybernetics. 2007.

Kalman, R. E.; *“A New Approach to Linear Filtering and Predictions Problems”*. Transaction of the ASME-Journal of Basic Engineering, pp. 35-45, March 1960.

K.-A. Lie; S. Krogstad; I. S. Ligaarden; J. R. Natvig; H. M. Nilsen and B. Skaflestad.; *“Open Source MATLAB Implementation of Consistent Discretisations on Complex Grids”*. Comput. Geosci., Vol. 16, No. 2, pp. 297-322, 2012. DOI: 10.1007/s10596-011-9244-4.

Khairullah, Md., *“Parallelization of Ensemble Kalman Filter (EnKF) For Oil Reservoirs”*. Master’s Thesis in Computer Simulations for Science and Engineering (COSSE). Delft University of Technology. The Netherlands. 2012.

Kocis, L, Whiten, J. W., *“Computational Investigations of Low-discrepancy Sequences”*, ACM Transactions on Mathematical Software, 1997.

Landa, J. L.; *“Reservoir Parameter Estimation Constrained to Pressure Transients, Performance History and Distributed Saturation Data”*. PhD thesis, Stanford University. 1979.

Lahoz, W.; Khattatov, B. and Menard, R., *“Data Assimilation: Making Sense of Observations”*. editors. Springer, 2010.

Liang, B.; *“An Ensemble Kalman Filter Module for Automatic History Matching”*. Dissertation presented to the Faculty of Graduate School, University of Texas, Austin. 2007.

Liu, N., and Oliver, D.S.; *“Critical Evaluation of the Ensemble Kalman Filter on History Matching of Geological Facies”*. Paper presented at the SPE Reservoir Simulation Symposium, Houston, TX, 31 January-2 February, SPE 92867. 2005a.

Liu, N., and Oliver, D.S.; *“Ensemble Kalman Filter for Automatic History Matching of Geologic Facies”*. Journal of Petroleum Science and Engineering 47: 147–161. 2005b.

Lorentzen, R.J., Nævdal, G., Vallès, B., SPE, and Grimstad, A.-A., RF-Rogaland Research.; *“Analysis of the Ensemble Kalman Filter for Estimation of Permeability and*

Porosity in Reservoir Models". Paper presented at the SPE Annual Technical Conference and Exhibition, Dallas, Texas, USA, SPE 96375. 2005.

Machale, D.; *"George Boole: His Life and Work"*. Dublin, Ireland: Boole. 1985.

Maschio, C.; Campana, V. A. and Schiozer, D. J.; *"Integração do Processo de Ajuste de Histórico com a Modelagem Geoestatística em Reservatórios de Petróleo"*, Revista Brasileira de Geociências, 38(1 - Suplemento): 75-81, março 2008.

Mei Han; Gaoming Li and Jingyi Chen.; *"Assimilating Microseismic and Welltest Data Using EnKF for Accurate Reservoir Characterisation"*, International Petroleum Technology Conference (IPTC), University of Tulsa, 26-28, March 2013.

Metropolis, N.; Rosenbluth, A.; Rosenbluth, M.; Teller, A. and Teller, E.; *"Equation of State Calculations by Fast Computing Machines"*. Journal of Chem. Phys., 21, 6: 1087-1092. 1953.

Nævdal, G., Mannseth, T., Vefring, E.H.; *"Near-Well Reservoir Monitoring through Ensemble Kalman Filter"*. Presented at SPE/DOE Improved Oil Recovery Symposium, Tulsa, Oklahoma, 13-17 April, USA, SPE 75235. 2002.

Montgomery, D. C., Runger, G. C., *"Estatística Aplicada e Probabilidade para Engenheiros"*, LTC, 2012.

Nævdal, G., Johnsen, L.M., Aanonsen, S.I., Vefring, E.H.; *"Reservoir Monitoring and Continuous Model Updating using Ensemble Kalman Filter"*. Society of Petroleum Engineering Journal 10 (1): 66-74. 2005.

Nævdal, G., Brouwer, D.R., and Jansen, J.D.; *"Waterflooding Using Closed-Loop Control"*. Computational Geosciences 10 (1): 37-60. 2006.

Oliveira, Diego F. B.; *"Técnicas de Otimização da produção para Reservatórios de Petróleo"*, 137 f. Dissertação (Mestrado em Engenharia Civil), Departamento de Engenharia Civil, UFPE, Recife. 2006.

Oliver, D. S., Reynolds, A. C, Liu, N.; *"Inverse Theory for Petroleum Reservoir Characterization and History Matching"*, Cambridge University Press, 2008.

Panda, M. N. and Lake, L. W.; *"Parallel Simulated Annealing for Stochastic Reservoir Modeling"*. Paper SPE 26418 presented at the SPE Annual Technical Conference and Exhibition, Houston, TX, 3-6 October. 1993.

Park, K., and Choe, J.; *"Use of Ensemble Kalman Filter with 3-Dimensional Reservoir Characterization"*. Paper presented at the SPE Europec/EDGE Annual Conference and Exhibition, Vienna, Austria, 12-15 June, SPE 100178. 2006.

Portellaand, R. C. M. and Prais, F.; *"Use of Automatic History Matching and Geostatistical Simulation to Improve Production Forecast"*. Paper SPE 53976 presented at

the Latin American and Caribbean Petroleum Engineering Conference, Caracas, Venezuela, 21-23 April. 1999.

Press, W. H; Teukolsky, S. A.; Vetterling W. T. and Flannery, B. P.; *“Numerical Recipes in Fortran”*. Cambridge University Press, p. 963. 1989.

Reinlie, S.T.; *“Analysis of Continuous Monitoring Data and Rapid, Stochastic Updating of Reservoir Models”*. PhD Dissertation, University of Texas, Austin. 2006

Reza, T.; Sanjay, S. and Mary, F. W.; *“Rapid Updating of Stochastic Models Using an Ensemble Filter Approach”*, SPE, U. of Texas 2013, Society of Petroleum Engineers, Texas USA, 18–20 February 2013.

Rosa, A. J.; Carvalho, R. D. S.; Xavier, J. A. D., *“Engenharia de Reservatórios de Petróleo”*. 1. ed. Rio de Janeiro: Editora Interciência, 2006.

Saad G, Ghanen, R, *“Characterization of Reservoir Simulation using a polynomial chaos based ensemble Kalman Filter”*, Water Resources Research, vol. 45, n. 4, W04417, 2009.

Sambridge, M.; *“Geophysical Inversion with a Neighborhood Algorithm-I: Searching a Parameter Space”*. Geophys. J. Int., 138: 479-494. 1999a.

Sambridge, M.; *“Geophysical Inversion with a Neighborhood Algorithm-I: Appraising the Ensemble”*. Geophys. J. Int., 138: 727-746. 1999b.

Seiler, A.; Evensen, G.; Skjervheim, J. A.; Hove, J.; *“Advanced Reservoir Management Work Flow Using an EnKF Based Assisted History Matching Method”*. SPE Reservoir Simulation Symposium, SPE 118906, 2009.

Skjervheim, J.-A., Evensen, G., Aanonsen, S.I., Ruud, B.O., and Johnansen, T.A.; *“Incorporating 4D Seismic Data in Reservoir Simulation Models Using Ensemble Kalman Filter”*. Paper presented at the SPE Annual Technical Conference and Exhibition, Dallas, Texas, USA, 9-12 October, SPE 95789. 2005.

SINTEF. Homepage of the MRST project. url: <http://www.sintef.no/Projectweb/MRST> (visited on 13/03/2014).

Sousa, S. H. G.; Maschio, C. and Schiozer, D. J.; *“Scatter Search Metaheuristic Applied to the History-Matching Problem”*. Paper SPE 102975 presented at the SPE Annual Technical Conference and Exhibition, San Antonio, TX, 24-27 September. 2006.

Suzuki, S. and Caers, J.; *“History Matching with an Uncertain Geological Scenario”*. Paper SPE 102154 presented at the SPE Annual Technical Conference and Exhibition, San Antonio, TX, 24-27 September. 2006.

Tavakoli, R., Pencheva, G., Wheeler, M. F., Ganis, B., “*A parallel ensemble-based framework for reservoir history matching and characterization*”, Computational Geosciences, n. 17, pp. 83-97, 2013.

Tavassoli, Z.; Carter, J. N.; King, P. R.; “*Errors in History Matching*”. SPE 86883. 2004.

Tippett, M.K., Anderson, J.L., Bishop, C.H., Hamill, T.M., and Whitaker, J.S.; “*Ensemble Square Root Filters*”. Monthly Weather Review, 131: 1485-1490. 2003.

Trani, M.; Arts, R. and Leeuwenburgh, O.; “*Seismic History Matching of Fluid Fronts Using the Ensemble Kalman Filter*”, Delft University of Technology, TNO, 2013.

Weisstein, Eric.; Monte Carlo Method. <http://mathworld.wolfram.com/MonteCarloMethod.html>. 2012.

Wen, X.-H., and Chen, W.H.; “*Real-Time Reservoir Model Updating Using Ensemble Kalman Filter*”. Paper presented at the SPE Annual Reservoir Simulation Symposium, Houston, TX, 31 January-2 February, SPE 92991. 2005.

Whitaker, J.S., and Hamill, T.M.; “*Ensemble Data Assimilation without Perturbed Observations*”. Monthly Weather Review, 130: 1913-1924. 2002.

Wiener, N., “*The Homogeneous Chaos*”, American Journal of Mathematics, vol. 60, n. 4, pp 897-936, 1938.

Zafari, M., and Reynolds, A.C.; “*Assessing the Uncertainty in the Reservoir Description and Performance Predictions with the Ensemble Kalman Filter*”. Paper presented at the SPE Annual Technical Conference and Exhibition, Dallas, TX, 9- 12 October, SPE 95750. 2005.

ANEXO A

Fundamentos da Estatística e Probabilidade

A.1 Introdução

O material neste capítulo foi baseado do material de aula do curso de Estatística (Da Costa, 2003), e tem como objetivo proporcionar ao leitor uma base para a compreensão da teoria do filtro. Alguns termos básicos de estatística são apresentados primeiro, seguido de uma introdução para a distribuição de probabilidade de Gauss.

No nosso cotidiano, lidamos sempre com situações em que está presente a incerteza do resultado, embora, muitas vezes, os resultados possíveis sejam conhecidos. É conveniente, então, dispormos de uma medida que exprima a incerteza presente em cada um destes acontecimentos. Tal medida é a probabilidade.

A Teoria da Probabilidade que permite estudar, por exemplo, de forma abstrata um fenômeno físico ao qual está associado uma incerteza. Toda a vez que se emprega a matemática com a finalidade de estudar algum fenômeno deve-se começar por construir um modelo. Este modelo pode ser construído de fenômenos **determinísticos** ou **probabilísticos**.

❖ Fenômenos Determinísticos

São aqueles em que os resultados são sempre os mesmos, para as mesmas condições da experiência, qualquer que seja o número de ocorrências verificadas.

❖ Fenômenos Probabilísticos

São aqueles que, mesmo repetidos várias vezes sob condições idênticas, apresentam resultados diferentes e imprevisíveis.

A.2. Espaço Amostral

Define-se espaço amostral (Ω) o conjunto de todos os resultados possíveis de um experimento ou fenômeno aleatório.

Um resultado do experimento (ω) é chamado ponto amostral e o espaço amostral é o conjunto de todos os pontos amostrais.

A.2.1 Classificação de um Espaço Amostral

Um espaço amostral pode ser classificado em:

- ❖ Finito, quando sua cardinalidade é finita. Por exemplo: $\Omega = \{1, 2, 3, 4, 5, 6\}$
- ❖ Infinitos, caso contrário. Neste caso, podem ser:
 - Enumeráveis (ou contáveis), quando podemos associar a cada elemento do espaço biunivocamente um número natural. Por exemplo: $\Omega = \{1, 2, 3, 4, 5, \dots\}$
 - Não-enumeráveis, caso isto não seja possível (ou não contáveis), Por exemplo: $\Omega = \{t \in \mathbb{R} / t \geq 0\}$

A.3 Evento

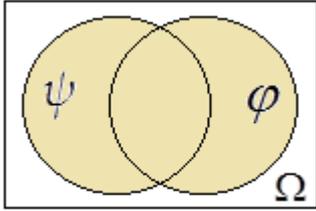
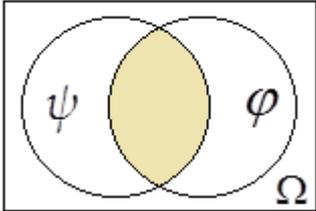
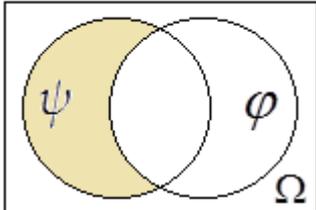
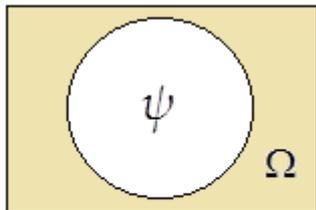
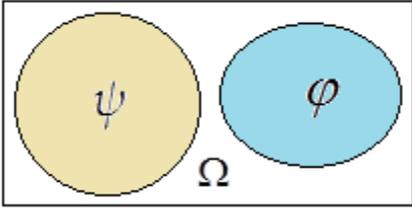
Chamamos de evento (ψ) qualquer subconjunto do espaço amostral (Ω) de um experimento aleatório. Assim, um evento aleatório pode ser um único ponto amostral ou uma reunião deles.

- ❖ Se $\psi = \Omega$, ψ é chamado evento certo.
- ❖ Se $\psi \subset \Omega$ e ψ é um conjunto unitário, ψ é chamado evento elementar.
- ❖ Se $\psi = \emptyset$, ψ é chamado evento impossível.

A.3.1 Combinação de Eventos

Podem-se realizar operações entre eventos da mesma forma que elas são realizadas entre conjuntos. Sejam ψ e φ dois eventos de um mesmo espaço amostra Ω . A Tabela mostra as operações entre eventos,

Tabela A.1: Operações entre eventos, a região sombreada é o resultado da operação.

<p>ψ união φ ou ψ soma φ, anotado por $\psi \cup \varphi$, se e somente se ψ ocorre ou φ ocorre.</p>  <p style="text-align: center;">$\psi \cup \varphi$</p>	<p>ψ produto φ ou ψ interseção φ, anotado por $\psi \cap \varphi$, se e somente se ψ ocorre e φ ocorre.</p>  <p style="text-align: center;">$\psi \cap \varphi$</p>
<p>ψ menos φ ou ψ diferença φ, anota-se $\psi - \varphi$, se e somente se ψ ocorre e φ não ocorre.</p>  <p style="text-align: center;">$\psi - \varphi$</p>	<p>O complementar de ψ, anotado por ψ^c ou ψ', se e somente se ψ não ocorre.</p>  <p style="text-align: center;">ψ^c</p>
<p>Dois eventos ψ e φ, são denominados mutuamente exclusivos ou excludentes, se eles não puderem ocorrer juntos, isto é, se $\psi \cap \varphi = \emptyset$.</p> 	

A.4 Álgebra de Conjuntos

Estamos interessados em trabalhar com subconjuntos do espaço amostral, e estes subconjuntos formam uma classe. Para consistência da Teoria da Probabilidade, esta classe deve ter algumas propriedades que definem uma álgebra.

É definido como uma álgebra de subconjuntos de Ω , um conjunto F qualquer, que seja subconjunto não-vazio do espaço amostral Ω e satisfaça as seguintes propriedades:

- ❖ Se $\psi \in F$, então $\psi^c \in F$.
- ❖ Se $\psi \in F$ e $\varphi \in F$, então $\psi \cup \varphi \in F$.

A partir dessas propriedades podemos observar que uma álgebra de conjuntos também seguirá as propriedades a seguir:

- ❖ $\Omega \in F$.
- ❖ $\emptyset \in F$.
- ❖ Se $\psi_i \in F, i \geq 1$, então $\bigcup_{i=1}^n \psi_i \in F$.

Para que F seja chamado uma σ -álgebra, as mesmas propriedades anteriores são validas, mas com uma modificação na última: a união deve ser verificada “exaustivamente”, ou seja, de forma a que todas as possíveis uniões venham a pertencer ao conjunto,

- ❖ Se $\psi_i \in F, i \geq 1$, então $\bigcup_{i=1}^{\infty} \psi_i \in F$.

A.5 Conceitos de Probabilidade

Existem três formas de se definir probabilidade. A definição clássica, a definição frequentista e a definição axiomática.

A.5.1 Definição Clássica de Probabilidade

Seja ω um experimento aleatório e Ω um espaço amostral. Seja $\psi \subseteq \Omega$ um evento, a probabilidade de ψ , é definida:

$$P(\psi) = \frac{\text{Possíveis resultados favorecendo o evento } \psi}{\text{Número total de resultados possíveis}}$$

Crítica à Definição Clássica

- ❖ A definição clássica é dúbia, já que a idéia de “igualmente provável” é a mesma de “com probabilidade igual”, isto é, a definição é circular, porque está definindo essencialmente a probabilidade com seus próprios termos.
- ❖ A definição não pode ser aplicada quando o espaço amostral é infinito.

A.5.2 Definição de Probabilidade como Frequência Relativa

Uma das definições de probabilidade utiliza a frequência relativa, já que as frequências relativas são estimativas de probabilidades. Podemos então definir a probabilidade como a proporção (ou frequência relativa) em uma seqüência muito grande de experimentos.

$$P(\psi) = \lim_{N \rightarrow \infty} \frac{n(\psi)}{N} = \frac{\text{número de vezes que } \psi \text{ ocorre}}{\text{número de vezes que } \omega \text{ é repetido}}$$

A.5.2.1 Propriedades da Frequência Relativa

Seja ω um experimento e ψ e φ dois eventos de um espaço amostral Ω . Sejam $P(\psi)$ e $P(\varphi)$ as frequências relativas de ψ e φ respectivamente. Então,

- ❖ $0 \leq P(\psi) \leq 1$, isto é, a frequência relativa do evento ψ é um número que varia entre 0 e 1.
- ❖ $P(\psi) = 1$, se e somente se, ψ ocorre em todas as repetições de ω .
- ❖ $P(\psi) = 0$, se e somente se, ψ nunca ocorre nas repetições de ω .
- ❖ $P(\psi \cup \varphi) = P(\psi) + P(\varphi)$, se ψ e φ forem eventos mutuamente excludentes.

Crítica à Definição Frequentista

Esta definição, embora útil na prática, apresenta dificuldades matemáticas, pois o limite pode não existir. Em virtude dos problemas apresentados pela definição clássica

sica e pela definição frequentista, foi desenvolvida uma teoria, na qual a probabilidade é um conceito indefinido, como o ponto e a reta o são na geometria.

A.5.3 Definição Axiomática de Probabilidade

Para um dado experimento, é necessário atribuir para cada evento ψ no espaço amostral Ω um número $P(\psi)$ que indica a probabilidade de ψ ocorrer. Para satisfazer a definição matemática de probabilidade, este número $P(\psi)$ deve satisfazer três axiomas específicos:

❖ **Axioma 1:** Para qualquer evento ψ , $P(\psi) \geq 0$.

❖ **Axioma 2:** $P(\Omega) = 1$.

❖ **Axioma 3:** Para qualquer seqüência infinita de eventos disjuntos $\psi_i \cap \psi_j = \emptyset$; $i, j = 1, 2, \dots, (i \neq j)$, então,

$$P\left(\bigcup_{i=1}^{\infty} \psi_i\right) = \sum_{i=1}^{\infty} P(\psi_i) \quad \text{A.1}$$

A definição matemática de probabilidade pode agora ser dada como segue:

A distribuição de probabilidade, ou simplesmente a probabilidade, no espaço amostral Ω é uma especificação de números $p(\psi)$ que satisfazem os axiomas 1, 2, e 3. Destes axiomas decorrem um número de teoremas úteis cuja demonstração é trivial.

Teorema 1: $P(\emptyset) = 0$

Teorema 2: Para qualquer seqüência finita de eventos disjuntos $\psi_i \cap \psi_j = \emptyset$; $i, j = 1, 2, \dots, n (i \neq j)$

$$P\left(\bigcup_{i=1}^n \psi_i\right) = \sum_{i=1}^n P(\psi_i) \quad \text{A.2}$$

Teorema 3: Para qualquer evento ψ ,

$$P(\psi^c) = 1 - P(\psi) \quad \text{A.3}$$

Teorema 4: Para qualquer evento ψ ,

$$0 \leq P(\psi) \leq 1 \quad \text{A.4}$$

Teorema 5: Se $\psi \subset \varphi$, então

$$P(\psi) \leq P(\varphi) \quad \text{A.5}$$

Teorema 6: Para qualquer dois eventos ψ e φ ,

$$P(\psi \cup \varphi) = P(\psi) + P(\varphi) - P(\psi \cap \varphi) \quad \text{A.6}$$

A.6 Eventos Independentes

Suponha que dois eventos ψ e φ ocorram independentes um do outro no sentido que a ocorrência ou não de um deles tenha nenhuma relação, e nenhuma influência na ocorrência ou na não ocorrência do outro. Nessas condições

$$P(\psi \cap \varphi) = P(\psi) \cdot P(\varphi) \quad \text{A.7}$$

A.7 Probabilidade Condicional

Se ψ e φ são dois eventos, a probabilidade de ψ ocorrer, depois de φ ter acontecido, é representada por $P(\psi \setminus \varphi)$ (probabilidade de ψ dado φ) e é definida como:

$$P(\psi \setminus \varphi) = \frac{P(\psi \cap \varphi)}{P(\varphi)}; \quad P(\varphi) > 0 \quad \text{A.8}$$

Se $P(\varphi) = 0$ a $P(\psi \setminus \varphi)$ não é definida.

A.8 Probabilidade Condicional para Eventos Independentes

Se ψ e φ são independentes, então $P(\psi \cap \varphi) = P(\psi) \cdot P(\varphi)$. Logo,

$$P(\psi \setminus \varphi) = \frac{P(\psi) \cdot P(\varphi)}{P(\varphi)} = P(\psi) \quad \text{A.9}$$

Da mesma forma,

$$P(\varphi \setminus \psi) = \frac{P(\varphi) \cdot P(\psi)}{P(\psi)} = P(\varphi) \quad \text{A.10}$$

Teorema:

Suponha que $\psi_1, \psi_2, \dots, \psi_n$ sejam quaisquer eventos tais que:

$P(\psi_1) > 0, P(\psi_1 \cap \psi_2) > 0, \dots, P(\psi_1 \cap \psi_2 \cap \dots \cap \psi_{n-1}) > 0$, então,

$$P(\psi_1 \cap \psi_2 \cap \dots \cap \psi_n) = P(\psi_1) \cdot P(\psi_2 \setminus \psi_1) \cdot P(\psi_3 \setminus \psi_1 \cap \psi_2) \cdot \dots \cdot P(\psi_n \setminus \psi_1 \cap \psi_2 \cap \dots \cap \psi_{n-1}) \quad \text{A.11}$$

A.9 Probabilidade Total e Teorema de Bayes

Diz-se que os conjuntos $\psi_1, \psi_2, \dots, \psi_n$ eventos de um mesmo espaço amostral Ω , formam uma partição deste espaço se:

- ❖ $\psi_i \cap \psi_j = \emptyset$, para todo $i \neq j$.
- ❖ $\psi_1 \cup \psi_2 \cup \dots \cup \psi_n = \Omega$.
- ❖ $P(\psi_i) > 0$, para todo i .

A.9.1 Probabilidade total

Considere-se um espaço amostra Ω e $\psi_1, \psi_2, \dots, \psi_n$ uma partição deste espaço amostral. Seja φ um evento de Ω . Então φ , pode ser escrito como:

$$\varphi = (\varphi \cap \psi_1) \cup (\varphi \cap \psi_2) \cup \dots \cup (\varphi \cap \psi_n) \quad \text{A.12}$$

É claro que, alguns destes conjuntos $\varphi \cap \psi_j$, poderão ser vazios, mas isto não representa nenhum problema na decomposição de φ . O importante é que todos os conjuntos $\varphi \cap \psi_1, \varphi \cap \psi_2, \dots, \varphi \cap \psi_n$ são dois a dois mutuamente excludentes. E por isto, pode-se aplicar a propriedade da adição de eventos mutuamente excludentes e escrever.

$$\begin{aligned} P(\varphi) &= P((\varphi \cap \psi_1) \cup (\varphi \cap \psi_2) \cup \dots \cup (\varphi \cap \psi_n)) \\ &= P(\varphi \cap \psi_1) + P(\varphi \cap \psi_2) + \dots + P(\varphi \cap \psi_n) \end{aligned} \quad \text{A.13}$$

Mas cada um dos termos $P(\varphi \cap \psi_j)$ pode ser escrito na forma:

$$P(\varphi \cap \psi_j) = P(\psi_j) \cdot P(\varphi \setminus \psi_j) \quad \text{A.14}$$

Pela definição de probabilidade condicional, obtém-se então o denominado teorema da probabilidade total:

$$P(\varphi) = P(\psi_1) \cdot P(\varphi \setminus \psi_1) + P(\psi_2) \cdot P(\varphi \setminus \psi_2) + \dots + P(\psi_n) \cdot P(\varphi \setminus \psi_n) \quad \text{A.15}$$

Ou

$$P(\varphi) = \sum_{j=1}^n P(\psi_j) \cdot P(\varphi \setminus \psi_j) \quad \text{A.16}$$

A.9.2 Teorema de Bayes

Pela notação já vista acima, e generalizando-o que se está interessado em obter é a probabilidade de ocorrência de um dos ψ_j dado que φ ocorreu, isto é, o que se quer é saber o valor de $P(\psi_j \setminus \varphi)$, em que os eventos $\psi_1, \psi_2, \dots, \psi_n$ formam uma partição de Ω e φ é um evento qualquer de Ω .

Aplicando a definição de probabilidade condicionada segue que:

$$P(\psi_j \setminus \varphi) = \frac{P(\psi_j \cap \varphi)}{P(\varphi)} = \frac{P(\psi_j) \cdot P(\varphi \setminus \psi_j)}{P(\varphi)} \quad \text{A.17}$$

onde $P(\varphi)$ é avaliado pelo teorema da probabilidade total. Este resultado é conhecido como teorema de Bayes. Assim:

$$P(\psi_j \setminus \varphi) = \frac{P(\psi_j) \cdot P(\varphi \setminus \psi_j)}{P(\psi_1) \cdot P(\varphi \setminus \psi_1) + P(\psi_2) \cdot P(\varphi \setminus \psi_2) + \dots + P(\psi_n) \cdot P(\varphi \setminus \psi_n)} \quad \text{A.18}$$

Ou

$$P(\psi_j \setminus \varphi) = \frac{P(\psi_j) \cdot P(\varphi \setminus \psi_j)}{\sum_{j=1}^n P(\psi_j) \cdot P(\varphi \setminus \psi_j)} \quad \text{A.19}$$

A.10 Sistema de Probabilidade

O trio formado por:

- ❖ Um espaço amostral Ω .
- ❖ Uma σ -álgebra F .
- ❖ Uma media de probabilidade P .

Configura um espaço de probabilidade, denotado por:

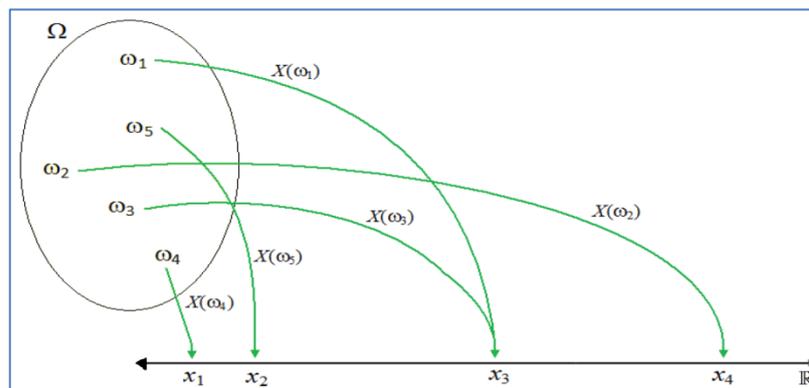
$$S = (\Omega, F, P)$$

A.20

A.11 Variáveis Aleatórias (VA)

Uma função X que associe a cada elemento de Ω ($\omega \in \Omega$) um número real $x = X(\omega)$ é denominada **variável real aleatória (VA)**, mostrada na Figura 4.1,

Figura A.1: Mapeamento de resultados de experimentos a números em eixo real.



A.11.1 Tipos de Variáveis Aleatórias

Os tipos de variáveis são discretas e contínuas ou mistas, a continuação falaremos de cada uma delas, mas para nosso caso faremos uso das variáveis aleatórias contínuas.

A.11.1.1 Variáveis Aleatórias Discretas (VAD)

Uma VA X é dita discreta se o seu conjunto de valores $X(\Omega)$ é finito ou então infinito contável ou enumerável.

a) Função de Probabilidade

Se a cada resultado $x_i \in \mathbb{R}$ de $X(\Omega)$ associa-se um número $f(x_i) = P(X = x_i)$, é denominada função de probabilidade de x_i e tal que satisfaça as seguintes propriedades:

$$\diamond f(x_i) \geq 0, \text{ para todo "i".}$$

$$\diamond \sum_{i=1}^{\infty} f(x_i) = 1$$

A coleção dos pares $(x_i, f(x_i))$, para $i = 1, 2, 3, \dots$ é denominada de distribuição de probabilidade da VAD X .

A.11.1.2 Variável Aleatória Contínua (VAC)

Seja X uma variável aleatória contínua. A função densidade de probabilidade (pdf) é uma função f que satisfaz às seguintes condições:

$$\diamond f(x) \geq 0, \quad x \in \mathbb{R}_x$$

$$\diamond \int_{\mathbb{R}_x} f(x) dx = 1$$

Além disso, definimos, para qualquer $c < d$ (em \mathbb{R}_x ¹⁶)

$$P(c < x < d) = \int_c^d f(x) dx \tag{A.21}$$

Obs:

$\diamond P(c < x < d)$ representa a área sob a curva, como exemplificado no gráfico da Figura 4.2, da fdp. f , entre $x = c$ e $x = d$.

\diamond Como consequência da descrição probabilística de X que, para qualquer valor especificado de X , digamos x_0 , teremos $P(X = x_0) = 0$, porque

$$P(X = x_0) = \int_{x_0}^{x_0} f(x) dx = 0.$$

¹⁶ \mathbb{R}_x : x é a variável aleatória nos reais.

a) Função de Distribuição Acumulada (FDA)

A função de distribuição acumulada de uma VA Real X , representada por F_X ou simplesmente por F , é definida por:

$$F_X(x) = P(X \leq x) \quad \text{A.22}$$

Observação:

❖ A função de distribuição de X é também frequentemente chamada de função de distribuição acumulada de X .

❖ A função $F_X(x)$ é não-decrescente quando x aumenta, isto é, se $x_1 < x_2$, então,

$$F_X(x_1) \leq F_X(x_2) \quad \text{A.23}$$

❖ Para qualquer valor de x ,

$$P(X > x) = 1 - F_X(x) \quad \text{A.24}$$

❖ Para quaisquer valores x_1 e x_2 , tais que $x_1 < x_2$,

$$P(x_1 < X \leq x_2) = F_X(x_2) - F_X(x_1) \quad \text{A.25}$$

Teoremas:

❖ Se X for uma VAD,

$$F(x) = P(X \leq x) = \sum_{x_i \leq x} x_i \quad \text{A.26}$$

❖ Se X for uma VAC com pdf f ,

$$F_X(x) = \int_{-\infty}^x f(s) ds \quad \text{A.27}$$

A.12 Distribuições Bivariadas

A distribuição de probabilidade conjunta de duas VA é denominada uma distribuição bivariada.

A.12.1 Distribuições Conjuntas Discretas

A função de probabilidade conjunta de X e Y é definida pela função f tal que qualquer ponto (x, y) no plano xy ,

$$f(x, y) = P(X = x, Y = y) \quad \text{A.28}$$

Observação:

$$\diamond \sum_i f(x_i, y_i) = 1.$$

\diamond Se X e Y forem independentes, $f(x_i, y_i) = P(X = x_i) \cdot P(Y = y_i)$, para todo i .

A.12.2 Distribuições Conjuntas Contínuas

Duas VA, X e Y possuem uma distribuição conjunta contínua se existe uma função f não negativa, definida sobre o plano xy , tal que para qualquer subconjunto S do plano,

$$P[(x, y) \in S] = \iint_S f(x, y) dx dy \quad \text{A.29}$$

A função f é denominada pdf conjunta de X e Y . Esta função deve satisfazer

$$f(x, y) > 0 \text{ para } -\infty < x < \infty, -\infty < y < \infty$$

e

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) dx dy = 1 \quad \text{A.30}$$

A.13 Distribuições Marginais

Denomina-se função densidade de probabilidade marginal de x à pdf de x quando ela é obtida através da pdf conjunta de X e Y .

A.13.1 Caso Discreto

Se X e Y são VA discretas com função de probabilidade conjunta $f(x, y)$, então a função de probabilidade marginal de X é obtida por:

$$P_X(x) = P(X = x) = \sum_y f(x, y) \quad \text{A.31}$$

Similarmente, a função de probabilidade marginal de Y é:

$$P_Y(y) = P(Y = y) = \sum_x f(x, y) \quad \text{A.32}$$

A.13.2 Caso Contínuo

Se X e Y possuem uma distribuição conjunta com pdf conjunta $f(x, y)$, então a pdf marginal $f_X(x)$ de X é obtida por:

$$f_X(x) = \int_{-\infty}^{\infty} f(x, y) dy \quad \text{A.33}$$

Similarmente, a pdf marginal de Y é obtida por:

$$f_Y(y) = \int_{-\infty}^{\infty} f(x, y) dx \quad \text{A.34}$$

Observação

Se X e Y forem independentes

- ❖ $P(x, y) = P_X(x) \cdot P_Y(y)$ (Caso discreto)
- ❖ $f(x, y) = f_X(x) \cdot f_Y(y)$ (Caso contínuo)

A.14 Esperança de uma Variável Aleatória

Seja X uma VA real, a esperança de X , denotada por $E(X)$, é um número definido por:

$$\diamond \mu = E(X) = \sum_{i=1}^n x_i f(x_i) \quad (\text{Caso discreto})$$

$$\diamond \mu = E(X) = \int_{-\infty}^{\infty} x f(x) dx \quad (\text{Caso contínuo})$$

Dada a média da amostra¹⁷, onde o valor esperado que é o resultado esperado se infinitos dados estão presentes pode ser aproximada por,

$$\mu = E(X) = \bar{x} = \frac{1}{N} \sum_{i=1}^N x_i \quad \text{A.35}$$

onde,

- ❖ N é o número de realizações independentes de x .
- ❖ x tem dimensão m e a distribuição $f(x)$.

A.15 Variância de uma Variável Aleatória

Seja x uma VA real, a variância de x , denotada por $Var(X)$, é um número definido por:

$$\diamond \sigma^2 = Var(X) = \sum_{i=1}^N (x_i - \mu)^2 f(x_i) \quad (\text{Caso discreto})$$

$$\diamond \sigma^2 = Var(X) = \int_{-\infty}^{\infty} (x - \mu)^2 f(x) dx \quad (\text{Caso contínuo})$$

A variância amostral pode ser aproximada como mostrado na Equação (A.36).

$$\sigma^2 = Var(X) = \frac{1}{N-1} \sum_{i=1}^N (x_i - \bar{x})^2 \quad \text{A.36}$$

onde, $N-1$ é utilizado em vez de N para proporcionar um estimador para a variância.

A.16 Covariância

¹⁷ São os objetos selecionados da população. Se esses objetos são selecionados de tal maneira que cada objeto tem a mesma chance de ser selecionado do que o outro, temos uma amostra aleatória.

Dadas duas VA X e Y , e suas respectivas pdfs $f(x)$ e $f(y)$, sua covariância é definida como,

$$COV[X, Y] = E[(X - E[X])(Y - E[Y])] \quad A.37$$

Se X e Y são independentes $f(X, Y) = f(X) \cdot f(Y)$ e assim $COV[X, Y] = 0$. Qualitativamente a covariância descreve a dependência entre duas variáveis aleatórias X e Y .

A covariância da amostra é derivada utilizando as mesmas propriedades da variância da amostra como acima e é dada pela

$$COV[X, Y] = E[(X - E[X])(Y - E[Y])] = \frac{1}{N-1} \sum_{i=1}^N (x_i - \bar{x})(y_i - \bar{y}) \quad A.38$$

A.17 Distribuições de Probabilidade

Existem muitas distribuições de variáveis aleatórias, nesta dissertação faremos uso um dos principais modelos de distribuição contínua, que é a distribuição normal ou de Gauss. Sua importância para a Estatística (prática) reside no fato que muitas variáveis encontradas na natureza se distribuem de acordo com o modelo normal. Este modelo também tem uma importância teórica devido ao fato de ser uma distribuição limite. A distribuição de Gauss é dada pela pdf,

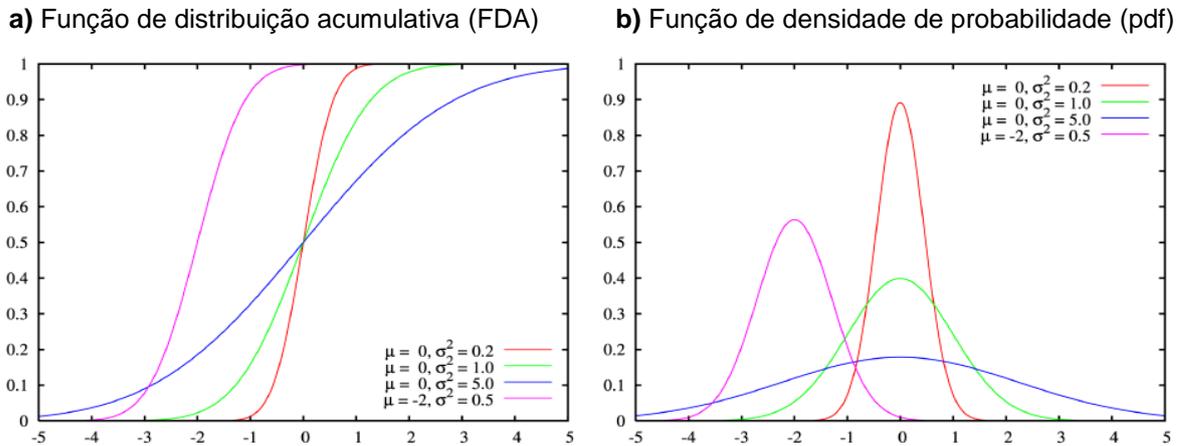
$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2\sigma^2}(x-E[X])^2}, \quad -\infty < x < \infty \quad A.39$$

A distribuição normal tem uma função de distribuição e pdf exemplificadas na Figura 4.3. Aqui nesta figura a FDA e pdf para uma distribuição gaussiana são representadas para diferentes desvios padrão σ e media μ .

$X \sim N(E[X], \sigma^2)$ é uma notação compacta para dizer que uma variável aleatória X é normalmente distribuída com um valor esperado $E[X]$ e desvio padrão σ . Se X é um vetor de variáveis aleatórias com distribuição gaussiana com média $E[X]$ e uma matriz de covariância Q , então isso pode ser denotado como $X \sim N(E[X], Q)$. Os elementos da diagonal de Q denotam a variância para cada uma das variáveis alea-

tórias em x e os elementos fora da diagonal representam a covariância entre as variáveis.

Figura A.2: FDA e pdf para a distribuição de Gauss



(Jensen, 2007)

A.18 Trabalhando com Amostras de uma Distribuição

Quando a dimensão de uma função de probabilidade aumenta torna-se impraticável, se não impossível, avaliar os integrais por integração numérica em uma malha regular. Suponha que a dimensão é 10 e precisamos de 10 pontos de grade em cada direção para ter uma representação adequada da densidade. Uma grade com 10^{10} nós teria que ser armazenado o que exigiria 40 Gigabytes de armazenamento e 10^{10} adições seriam necessários para calcular a integral.

Felizmente, há técnicas para a integração numérica direta, que muitas vezes funcionam muito bem para sistemas de alta dimensão. Uma é a abordagem chamada o método de Markov Chain Monte Carlo (MCMC), (Robert e Casella, 2004), e assume que nós temos disponível um grande número N , de realizações da distribuição $f(x)$