



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



Dissertação de Mestrado

Propagação de Incertezas na Exploração e Produção de Petróleo:

Abordagens com o uso do Método do Caos Polinomial
e da Colocação Estocástica

Jane Cassimiro Carneiro da Cunha

Orientador:
Ramiro Brito Willmersdorf, D. Sc.
(UFPE/DEMEC)

**Recife – PE
Agosto de 2010**

Propagação de Incertezas na Exploração e Produção de Petróleo:

*Abordagens Com o Uso do Método do Caos Polinomial e
da Colocação Estocástica*

Jane Cassimiro Carneiro da Cunha

Orientador: **Professor Ramiro Brito Willmersdorf, D.Sc..**

Esta dissertação foi julgada e aprovada para a obtenção do grau de Mestre em Ciências em Engenharia Mecânica do Departamento de Engenharia Mecânica da Universidade Federal de Pernambuco.

Recife, Agosto de 2010.

Catálogo na fonte
Bibliotecário Marcos Aurélio Soares da Silva, CRB-4 / 1175

C972p

Cunha, Jane Cassimiro Carneiro da.

Propagação de incertezas na exploração e produção de Petróleo: abordagens com o uso do método do Caos Polinomial e da Colocação Estocástica / Jane Cassimiro Carneiro da Cunha. - Recife: O Autor, 2010.

x, 108 folhas, il., gráfs., tabs.

Orientador: Prof^o. Dr. Ramiro Brito Willmersdorf.

Dissertação (Mestrado) – Universidade Federal de Pernambuco. CTG. Programa de Pós-Graduação em Engenharia Mecânica, 2010.

Inclui Referências.

1. Engenharia Mecânica. 2. Propagação de Incertezas. 3. Reservatório de Petróleo. 4. Riscos. I. Willmersdorf, Ramiro Brito (orientador). II. Título.

621 CDD (22. ed.)

UFPE
BCTG/2012-006

"PROPAGAÇÃO DE INCERTEZA NA EXPLORAÇÃO E PRODUÇÃO DE
PETRÓLEO: ABORDAGENS COM O USO DOS MÉTODOS DE CAOS
POLINOMIAL E COLOCAÇÃO ESTOCÁSTICA".

JANE CASSIMIRO CARNEIRO DA CUNHA

ESTA DISSERTAÇÃO FOI JULGADA ADEQUADA PARA OBTENÇÃO DO
TÍTULO DE MESTRE EM ENGENHARIA MECÂNICA

ÁREA DE CONCENTRAÇÃO: MECÂNICA COMPUTACIONAL
APROVADA EM SUA FORMA FINAL PELO
CURSO DE PÓS-GRADUAÇÃO EM ENGENHARIA MECÂNICA/CTG/EEP/UFPE

Prof. Dr. RAMIRO BRITO WILLMERSDORF
ORIENTADOR/PRESIDENTE

Prof. Dr. SEVERINO LÉOPOLDINO URTIGA FILHO
COORDENADOR DO CURSO

BANCA EXAMINADORA:

Prof. Dr. RAMIRO BRITO WILLMERSDORF (UFPE)

Prof. Dr. DARLAN KARLO ELISIÁRIO DE CARVALHO (UFPE)

Profa. Dra. SILVANA MARIA BASTOS AFONSO (UFPE)

Agradecimentos

Com a finalização deste trabalho, aproveito a oportunidade para agradecer a todas aquelas pessoas que, de alguma forma, contribuíram para a sua realização.

Ao professor Ramiro Brito Willmersdorf pelos ensinamentos transmitidos e pela sua valiosa orientação para a realização desta pesquisa.

Aos professores Silvana Afonso e Bernardo Horowitz, do departamento de Engenharia Civil da UFPE que, juntamente com o professor Ramiro, conduziram as pesquisas em determinada fase do trabalho criando as alternativas para a execução dos experimentos necessários a dissertação.

A todos os meus amigos por sempre acreditarem nas minhas empreitadas.

À Andrea Leite, em especial, pela sua amizade e companheirismo ao longo dos anos, também compartilhando em sala de aula os momentos difíceis do Mestrado.

Ao meu companheiro, Cássio, que com paciência esteve ao meu lado me incentivando sempre a continuar.

Em especial, aos meus filhos Daniel e Leonardo, por tentarem entender que os momentos de ausência são necessários para o meu crescimento profissional e pessoal.

Aos meus pais, Rivaldo e Janete, que com muito amor e sacrifício me prepararam para a vida, me ensinando sempre o caminho do bem e da verdade, e mostrando, sobretudo que com esforço e empenho a realização dos meus sonhos seria possível.

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

Resumo

Os grandes investimentos alocados na área de exploração e produção de petróleo, em geral, são realizados sob grandes riscos, devido às incertezas envolvidas no processo. Em virtude da grande quantidade de variáveis no processo de exploração e produção de petróleo, a análise de riscos – uma prática comum nessa área – vem se tornando cada vez mais complexa, exigindo metodologias mais eficientes. Dessa forma, as empresas petrolíferas vêm investindo pesadamente em pesquisas, visando à obtenção de inovações tecnológicas relativas a metodologias para análise de riscos, de forma que suas tomadas de decisões ocorram sob menores riscos. Como parte do processo de busca por essas inovações tecnológicas, este trabalho tem por objetivo o desenvolvimento de um sistema computacional para o estudo da propagação das incertezas envolvidas no processo de exploração e produção de petróleo.

Este trabalho faz abordagens sobre aspectos contextuais da exploração e produção de petróleo e sobre os recursos utilizados pelo sistema computacional desenvolvido, como: o DAKOTA, ambiente para otimização e propagação de incertezas; o IMEX, simulador numérico para o estudo do comportamento dos reservatórios de petróleo, e os métodos aplicáveis à propagação de incertezas. O sistema em si será composto pelos programas DAKOTA, IMEX e por programas adicionais desenvolvidos especificamente para possibilitar a troca de dados entre os dois primeiros programas.

Este trabalho consiste na realização de experimentos computacionais utilizando-se como métodos de propagação de incertezas o Caos Polinomial e a Colocação Estocástica. Esses métodos são empregados em problemas-modelo de simulação de reservatórios na determinação do efeito das incertezas das variáveis petrolíferas, como a porosidade e a permeabilidade, em uma variável de interesse – no caso, o Valor Presente Líquido (VPL) – ou a produção acumulada de óleo.

Palavras-chaves: Propagação de incertezas, Reservatórios de petróleo, Riscos.

Abstract

The major investments in the area of Oil's Exploration and Production are usually allocated under great risks due to the involved uncertainties in the process. Due to the big amount of involved variables in the process of oil's exploration and production, the analysis of risks - common practice in this area -, has become more complex also demanding more complex methodologies. Thus, more frequently, petroliferous companies have been investing in research for technological innovations attainment, in order to guarantee that decisions will be taken under lesser risks. As part of this process of searching for technological innovations, relative to risks analysis methodologies, this work has the objective of developing a computational system for the study of the propagation of involved uncertainties in the process of oil's exploration and production.

The work approach's includes the aspects of the exploration and production of oil and the resources used on the developed system as: DAKOTA, an environment for optimization and propagation of uncertainty; the numerical simulator IMEX, for the study of behavior of oil reservoirs, and the applicable methods for propagation of uncertainty. The system will be constructed by using DAKOTA, IMEX and other programs specifically designed to make possible data exchange between IMEX and DAKOTA.

The work consists of computational experiments using Polynomial Chaos and Stochastic Collocation as techniques of propagation of uncertainty, applied to model-problems of reservoir simulation to determine the effect of the uncertainties in petrophysical variables, such as porosity and permeability, on output variable - in this case the Net Present Value (NPV) - or the accumulated oil production.

Key-words: Propagation of uncertainties, Oil reservoirs, Risks

Lista de Figuras

<i>Figura 1. Representação Esquemática do Problema de Propagação de Incertezas.</i>	14
<i>Figura 2. Polinômios Ortogonais do Esquema de Askey.</i>	36
<i>Figura 3. Esquema representativo do sistema proposto.</i>	59
<i>Figura 4. Fluxo de informações no IMEX.</i>	61
<i>Figura 5. Regiões de permeabilidade horizontal para o Caso 1.</i>	74
<i>Figura 6. Estudo paramétrico do VPL com uma variável incerta.</i>	78
<i>Figura 7. Histogramas das variáveis de entrada e saída.</i>	80
<i>Figura 8. Realizações normais – primeiro conjunto de amostras.</i>	89
<i>Figura 9. Realizações normais – segundo conjunto de amostras.</i>	89
<i>Figura 10. Realizações normais – terceiro conjunto de amostras.</i>	90
<i>Figura 11. Realizações normais – quarto conjunto de amostras.</i>	90
<i>Figura 12. Magnitude dos autovalores.</i>	91
<i>Figura 13. Soma acumulada normalizada dos autovalores.</i>	92
<i>Figura 14. Realizações com o log natural da permeabilidade – primeiro conjunto de amostras.</i>	93
<i>Figura 15. Realizações com o log natural da permeabilidade – segundo conjunto de amostras.</i>	94
<i>Figura 16. Realizações com o log natural da permeabilidade – terceiro conjunto de amostras.</i>	94
<i>Figura 17. Realizações com o log natural da permeabilidade – quarto conjunto de amostras.</i>	95
<i>Figura 18. Saturação de água ao final da análise, correspondente às realizações da Figura 14.</i>	96
<i>Figura 19. Saturação de água ao final da análise, correspondente às realizações da Figura 15.</i>	96
<i>Figura 20. Saturação de água ao final da análise, correspondente às realizações da Figura 16.</i>	97
<i>Figura 21. Saturação de água ao final da análise, correspondente às realizações da Figura 17.</i>	97
<i>Figura 22. Produções acumuladas de óleo e água.</i>	98

Lista de Tabelas

<i>Tabela 1. Ligação entre as formas padrão das distribuições de probabilidades.</i>	36
<i>Tabela 2. Número de coeficientes na Expansão em Caos Polinomial.</i>	39
<i>Tabela 3. Tipos de distribuição suportados pela Expansão em Caos Polinomial.</i>	70
<i>Tabela 4. Scripts do pré-processamento e suas finalidades.</i>	72
<i>Tabela 5. Scripts do pós-processamento e suas finalidades.</i>	72
<i>Tabela 6. Características do modelo do Caso 1.</i>	73
<i>Tabela 7. Resultados para o método de Monte Carlo.</i>	79
<i>Tabela 8. Expansão em Caos Polinomial linear ($p = 1$; $OQ=p+1$).</i>	81
<i>Tabela 9. Expansão em Caos Polinomial de ordem 2 ($p=2$; $OQ=p+1$).</i>	81
<i>Tabela 10. Expansão em Caos Polinomial de ordem 3 ($p=3$; $OQ=p+1$).</i>	81
<i>Tabela 11. Expansão em Caos Polinomial de ordem 4 ($p=4$; $OQ=p+1$).</i>	82
<i>Tabela 12. Expansão polinomial de ordem 5 ($p=5$; $OQ=p+1$).</i>	82
<i>Tabela 13. Caos polinomial, uma variável incerta, com $OQ = p + 1$.</i>	82
<i>Tabela 14. Diferença entre os valores do Caos Polinomial e Monte Carlo para 1000 amostras (Valores da ordem de 10^6 US\$ para $E(VPL)$ e 10^5 US\$ para $S(VPL)$).</i>	83
<i>Tabela 15. Colocação Estocástica.</i>	83
<i>Tabela 16. Diferença entre os valores da Colocação Estocástica e Caos Polinomial.</i>	84
<i>Tabela 17. Caos Polinomial, para duas variáveis incertas.</i>	85
<i>Tabela 18. Colocação Estocástica, para duas variáveis incertas.</i>	85
<i>Tabela 19. Diferença entre os valores do Caos Polinomial e Colocação Estocástica para duas variáveis incertas.</i>	85
<i>Tabela 20. Método de Monte Carlo, para três variáveis incertas.</i>	86
<i>Tabela 21. Caos Polinomial, para três variáveis incertas, por quadratura numérica.</i>	86
<i>Tabela 22. Caos Polinomial de ordem 2, para 3 variáveis incertas, por regressão.</i>	87
<i>Tabela 23. Caos Polinomial de ordem 3, para 3 variáveis incertas, por regressão.</i>	87
<i>Tabela 24. Expansão em Caos Polinomial.</i>	98

Sumário

1 INTRODUÇÃO	11
1.1 SITUAÇÃO DO PROBLEMA	11
1.1.1 <i>Objetivo Geral</i>	13
1.1.2 <i>Objetivos Específicos</i>	13
1.2 DECLARAÇÃO DO PROBLEMA	14
1.3 ESCOPO DA DISSERTAÇÃO	15
2 EXPLORAÇÃO E PRODUÇÃO DE PETRÓLEO	18
2.1 ESTIMATIVAS DE VOLUMES	19
2.2 PROCESSOS DE EXTRAÇÃO DE PETRÓLEO	22
2.3 SIMULADORES DE RESERVATÓRIOS	24
3 PROPAGAÇÃO DE INCERTEZAS	27
3.1 VARIÁVEIS INCERTAS	28
3.2 MÉTODOS PARA PROPAGAÇÃO DE INCERTEZAS	29
3.2.1 <i>Métodos que Usam Amostragem Direta</i>	30
3.2.2 <i>Métodos que Não Usam Amostragem Direta</i>	32
4 EXPANSÃO EM CAOS POLINOMIAL	34
4.1 EXPANSÃO EM CAOS POLINOMIAL DE WIENER-ASKEY	35
4.2 PROJEÇÃO DE GALERKIN	41
4.2.1 <i>Aproximação por Amostragem</i>	42
4.2.2 <i>Aproximação por Quadratura/Cubatura</i>	42
4.3 REGRESSÃO LINEAR	43
4.4 MOMENTOS ANALÍTICOS	44
4.5 ANÁLISE DE SENSIBILIDADE DA VARIÁVEL ALEATÓRIA	45
5 COLOCAÇÃO ESTOCÁSTICA	46
5.1 MOMENTOS ANALÍTICOS	47
5.2 ANÁLISE DE SENSIBILIDADE DA VARIÁVEL ALEATÓRIA	47
6 REDUÇÃO DO ESPAÇO PROBABILÍSTICO	49
6.1 EXPANSÃO DE KARHUNEN-LOÈVE	49
6.1.1 <i>Detalhamento da Expansão de Karhunen-Loève (KLE)</i>	51
6.2 EXPANSÃO <i>KERNEL</i> DE KARHUNEN-LOÈVE (KKLE)	53
6.2.1 <i>Detalhamento da Expansão Kernel de Karhunen-Loève (KKLE)</i>	54
7 CARACTERIZAÇÃO DO SISTEMA.....	58
7.1 AMBIENTE COMPUTACIONAL	58
7.2 PROGRAMAS	59
7.2.1 <i>IMEX</i>	60
7.2.2 <i>DAKOTA</i>	62
7.2.3 <i>Programas Adicionais</i>	71
8 APLICAÇÕES.....	73
8.1 CASO EM ESTUDO	73
8.2 SITUAÇÕES PARA ANÁLISE	77
8.2.1 <i>Situação-1</i>	77
8.2.2 <i>Situação-2</i>	84
8.2.3 <i>Situação-3</i>	86
8.3 PROBLEMA DE ALTA DIMENSÃO	87
9 CONSIDERAÇÕES FINAIS.....	99
9.1 CONCLUSÕES	99
9.2 SUGESTÕES PARA TRABALHOS FUTUROS.....	102

REFERÊNCIAS 104

1 Introdução

1.1 Situação do Problema

Em geral, a alocação de grandes investimentos implica também no envolvimento de grandes riscos os quais, se não considerados, levam a decisões equivocadas que podem prejudicar a rentabilidade do negócio e até mesmo comprometer a sobrevivência da própria empresa. Atualmente, a avaliação dos riscos envolvidos nesses processos já é uma prática comum, como meio de garantir que as decisões sejam tomadas mais acertadamente, baseadas em estimativas mais confiáveis. Numa avaliação de riscos, todas as vulnerabilidades que podem levar o processo ao insucesso são mapeadas visando o estabelecimento de estratégias para reduzi-las ou contorná-las. Essas vulnerabilidades são caracterizadas pelas probabilidades de ocorrência de eventos adversos sobre determinadas variáveis do sistema, eventos estes que possam promover um grande distanciamento entre o resultado prático da implantação e operação de um projeto, e as projeções e premissas adotadas inicialmente. Dessa forma, quando existem variáveis sujeitas a vulnerabilidades, o fator risco é inserido no processo, uma vez que não há certeza sobre o comportamento exato delas.

O processo de exploração e produção de petróleo é uma das áreas na qual são alocados grandes investimentos e, ao mesmo tempo, onde há grandes riscos envolvidos. Para o gerenciamento de reservatórios de petróleo, as empresas vêm utilizando a simulação de reservatórios como peça-chave na propagação de incertezas e análise de riscos, uma vez que a mesma permite integrar informações relacionadas aos parâmetros petrofísicos com os relativos ao escoamento de fluidos e à estratégia de produção. Apesar de ser útil em todas as fases da exploração e produção do petróleo, a simulação de reservatórios torna-se mais empregada na avaliação, no desenvolvimento e no gerenciamento dos campos, em razão do grande esforço computacional envolvido. A simulação de reservatórios influencia muito o processo de tomada de decisão relacionado à exploração e produção de petróleo, pois é através dela que são testadas as várias alternativas para esses

processos como: a escolha do método de recuperação, do número e posição dos poços produtores e injetores, entre outras.

A precisão das estimativas fornecidas pelos simuladores de reservatórios depende da qualidade das informações que os alimentam e dos modelos utilizados. Nos anos 80, era usual a previsão de produção baseada em um único modelo determinístico de reservatório. Essa previsão possuía também um caráter determinístico e, na maioria dos casos, até otimista [29]. Nessa época, muitas das propriedades do reservatório eram definidas por um único valor, embora apresentassem variações de acordo com a sua posição, tornando as previsões imprecisas. Ainda assim, os altos preços praticados no mercado, na época, garantiam a viabilidade econômica dos projetos de campos de óleo e de gás [29]. Com a mudança de cenário, devido à crescente globalização dos negócios, ao elevado investimento em campos de petróleo e a outros diversos fatores envolvidos no processo [53], uma avaliação inadequada de riscos e estimativas imprecisas e/ou subestimadas podem provocar consequências desastrosas.

No cenário atual, para a geração de estimativas mais precisas, as propriedades do reservatório cujos comportamentos não são totalmente conhecidos, denominadas por variáveis incertas, devem ser representadas por valores não determinísticos e devem ser tratadas por metodologias específicas, denominadas aqui como métodos para a propagação de incertezas. Em geral, dependendo da dimensionalidade do problema, os métodos de propagação de incertezas demandam excessivo tempo e enorme custo computacional para o tratamento das variáveis incertas envolvidas, além do já exigido pela simulação de reservatórios. Nesse contexto, a identificação das maiores incertezas, a quantificação do impacto dessas incertezas nas decisões a serem tomadas e a avaliação dos riscos envolvidos nos projetos, tem sido alguns dos problemas desafiadores para as empresas de Óleo e Gás, motivando-as a investir em pesquisas na busca por metodologias mais eficientes em termos de maior precisão, menor esforço computacional e tempo de processamento aceitável, para a geração das informações que servirão como base para seus processos decisórios.

1.1.1 Objetivo Geral

De forma a contribuir com as empresas de Óleo e Gás no processo de busca por metodologias mais eficientes, que poderão ser utilizadas em suas avaliações de riscos, este trabalho tem como objetivo geral o desenvolvimento de um sistema computacional para o estudo do problema de propagação de incertezas dentro da abrangência de reservatórios de petróleo. O sistema tem como proposta o emprego, simultâneo, de técnicas de redução tanto para os modelos computacionais como para a dimensão do espaço probabilístico, de forma a possibilitar a minimização do esforço computacional e do tempo de processamento exigido, mantendo uma precisão razoável em relação àquela encontrada sem o emprego das técnicas de redução. Neste trabalho a precisão aceitável será em relação aos valores da média e desvio padrão do VPL gerados pelo método de Monte Carlo, sendo esses considerados os valores reais.

1.1.2 Objetivos Específicos

Como objetivos específicos, esse trabalho apresenta:

- a) Estudar as metodologias que podem ser empregadas para a propagação de incertezas em reservatórios de petróleo;
- b) Identificar dentre as metodologias de propagação de incertezas estudadas as que preenchem os requisitos exigidos pelo sistema proposto como: menor tempo de processamento, menor esforço computacional e obtenção de precisão igual ou superior em análise comparativa de outras metodologias já empregadas;
- c) Aplicar técnica para a redução do espaço probabilístico em problemas com entradas estocásticas de ordens mais altas, de forma a verificar a viabilidade do emprego das metodologias estudadas na propagação de incertezas dentro da abrangência de reservatórios de petróleo.

1.2 Declaração do Problema

No presente estudo, em razão do simulador de reservatório do sistema proposto ser um programa cujo código fonte não está disponível para o projeto, apenas as técnicas não intrusivas para a propagação de incertezas foram utilizadas. As técnicas não intrusivas são aquelas que consideram o simulador como uma “caixa preta” e, sendo assim, representam-no como uma função não linear, que mapeia as variáveis de entrada incertas (aleatórias) e determinísticas, em variáveis de saída incertas, conforme representação esquemática do problema de propagação de incertezas, apresentado na Figura 1.



Figura 1. Representação Esquemática do Problema de Propagação de Incertezas.

O Caos Polinomial ([59], [62]) e a Colocação Estocástica ([25], [43]), detalhados respectivamente nos capítulos 4 e 5, foram os métodos de propagação de incertezas empregados no sistema proposto, devido a proporcionarem altas taxas de convergência e precisão aceitável para uma quantidade de avaliações de função bem inferior aos outros métodos de propagação de incertezas; bem como estarem implementados no DAKOTA [43] como técnicas não intrusivas [12], satisfazendo assim as condições exigidas para sua aplicação. Essas técnicas têm em comum a ideia de aproximar a função não linear que representa o simulador com alguma função conhecida, como um polinômio global nas variáveis incertas. Os valores dessa função não linear, desconhecida, são obtidos mediante a realização de avaliações de função, em pontos específicos. Cada avaliação de função requer uma análise feita pelo simulador para o problema de interesse, e esta análise para problemas realistas de reservatórios, em geral, demanda muitas horas de processamento. Sendo assim, qualquer técnica de propagação de incertezas só terá

aplicabilidade prática se o número de pontos, exigidos para a avaliação de função, for relativamente pequeno. Para isso foram empregadas técnicas para a redução do espaço probabilístico.

O problema da propagação de incertezas consiste no cálculo de estatísticas (momentos) de variáveis de interesse, em geral calculadas em uma etapa de pós-processamento e a partir dos resultados de um simulador numérico de reservatórios, considerando que algumas das entradas deste simulador sejam incertas. Uma vez que a aproximação da função é obtida, estatísticas de interesse das variáveis de saída podem ser geradas com facilidade, através de fórmulas analíticas ou métodos baseados em amostragem da função aproximada. Nas técnicas aplicadas neste projeto, existe o requisito aparentemente forte de que é necessário o conhecimento da função de densidade de probabilidade (PDF – *Probability Density Function*) das variáveis de entrada, e da função de distribuição cumulativa (CDF - *Cumulative Distribution Function*), no caso de múltiplas variáveis incertas. Para o caso de algumas poucas variáveis incertas, esta densidade é fornecida diretamente, e em geral é escolhida entre algumas famílias típicas (uniforme, triangular, normal, lognormal, etc.), de acordo com as particularidades de cada problema e com a experiência do engenheiro.

É importante ressaltar ainda que este projeto concentra-se apenas nas incertezas referentes às propriedades petrofísicas – como a permeabilidade e a porosidade – quando consideradas como campos estocásticos, e não nas incertezas provenientes de estruturas mais complexas, como a presença ou não de falhas, localização de fácies, etc. Este segundo tipo de problemas será considerado em trabalhos futuros.

1.3 Escopo da Dissertação

Os demais capítulos deste trabalho estão organizados da seguinte forma:

No capítulo 2, para uma boa compreensão do assunto, é feito um breve estudo sobre os aspectos contextuais da exploração e produção de petróleo, e sobre os simuladores de reservatórios.

No capítulo 3 é apresentado o conceito de incerteza no contexto da propagação de incertezas e a sua importância nos processos decisórios das empresas.

São apresentadas, também nesse capítulo, as formas de representação das variáveis incertas e os métodos utilizados para a propagação de incertezas, classificados de acordo com a abordagem que utilizam.

Nos capítulos 4 e 5 são apresentados os conceitos básicos referentes às metodologias empregadas neste trabalho para o estudo da propagação de incertezas em reservatórios de petróleo: o Caos Polinomial e a Colocação Estocástica, respectivamente. Nesses capítulos são apresentados também os requisitos principais para a escolha dessas metodologias.

No capítulo 6 são apresentados os conceitos básicos referentes às técnicas empregadas para a redução do espaço probabilístico.

No capítulo 7 são discutidos os requisitos essenciais à construção do sistema proposto, assim como os conceitos básicos referentes ao IMEX [7], ao DAKOTA [43] e aos programas adicionais desenvolvidos com a finalidade de prover uma interface entre o DAKOTA e o IMEX. Esses programas adicionais realizam o tratamento prévio das informações geradas pelo DAKOTA, extraindo e formatando adequadamente apenas os dados que serão utilizados pelo IMEX, e vice-versa. Nesse mesmo capítulo é apresentado o programa desenvolvido para a geração das informações sobre o Valor Presente Líquido (VPL), que neste trabalho é a variável de interesse.

No capítulo 8 são divulgados os resultados obtidos com a aplicação do sistema desenvolvido, utilizando-se os seguintes métodos de propagação de incertezas, aplicados a reservatórios de petróleo: método de Monte Carlo ([15], [33], [41], [43]), método do Caos Polinomial ([59], [62]) e método da Colocação Estocástica ([25], [43]). Os resultados apresentados referem-se aos experimentos realizados com uma, duas e três variáveis incertas. Para cada uma das quantidades de variáveis definidas foram também criadas situações variando-se a ordem de quadratura/cubatura e a ordem da expansão, para o método do Caos Polinomial, e variando-se a ordem de interpolação, para o método da Colocação Estocástica. Como forma de avaliar o desempenho de cada uma das metodologias estudadas, é feita uma comparação entre os resultados obtidos nos experimentos para cada situação de estudo criada. Ainda neste capítulo é feita a aplicação da Expansão de Karhunen-Loève ([44], [45]) em um problema de alta dimensão, visando à redução do espaço probabilístico.

No capítulo 9 é feito um resumo das principais conclusões sobre as metodologias empregadas no sistema construído para a propagação de incertezas em reservatórios de petróleo. São apresentadas também nesse capítulo, questões importantes sobre os assuntos que ainda precisam ser tratados em trabalhos futuros.

2 Exploração e Produção de Petróleo

O petróleo é hoje uma importante fonte não renovável da matriz energética mundial para a qual a localização de jazidas com indícios suficientemente atrativos – ou seja, que justifiquem a continuidade das atividades de exploração – vem se tornando cada vez mais difícil. No passado, quando existia maior possibilidade de descoberta de boas jazidas, a maioria dos processos decisórios na área da exploração e produção de petróleo ainda podia ser realizada de forma simples e intuitiva, baseada apenas na observação das variáveis geológicas mais relevantes (volumes, dinâmica das acumulações, etc.) e nas tendências gerais do mercado (preço do óleo, demanda de derivados, etc.) [53]. No decorrer dos anos, com a redução das possibilidades de localização de boas jazidas, juntamente com outros acontecimentos – como a crescente globalização dos negócios e o envolvimento de diversos novos agentes no cenário atual – o processo decisório das empresas petrolíferas tornou-se bastante complexo e a análise de riscos, além de extremamente necessária, passou a utilizar metodologias complexas e mais abrangentes, para o tratamento da grande quantidade de variáveis envolvidas, sobre grande parte das quais se tem conhecimento limitado do comportamento, o que gera incertezas e riscos.

No processo de exploração e produção de petróleo vários tipos de riscos devem ser analisados, de acordo com Suslick [53]:

- a) Risco de um poço exploratório ou de desenvolvimento ser seco;
- b) Risco de uma descoberta não possuir um volume de óleo suficiente para os custos envolvidos no seu aproveitamento econômico;
- c) Risco relacionado com o preço futuro de óleo e gás natural;
- d) Risco econômico-financeiro;
- e) Risco ambiental;
- f) Risco político vinculado às incertezas jurídico-institucionais de um país detentor dos recursos petrolíferos.

No entanto, para esta dissertação, apenas os riscos relativos ao tipo “Risco de uma descoberta não possuir um volume de óleo suficiente para os custos

envolvidos no seu aproveitamento econômico”, categorizado como riscos geológicos, estão sendo quantificados [53]. Nesse contexto, os riscos geológicos estão diretamente vinculados aos modelos geológicos de acumulação, a definição dos limites e dos volumes envolvidos e aos riscos inerentes às variáveis críticas da dinâmica das acumulações petrolíferas (geração, reservatórios, suprimento e “*timing*”) [53].

No cenário atual não é mais aceitável que decisões sejam tomadas baseadas apenas em observações e intuições, como no passado, pois estimativas com grandes margens de erro podem levar a grandes perdas financeiras, comprometendo não só a viabilidade do projeto, mas também a sobrevivência da empresa. É possível obter uma noção da magnitude desses valores observando-se que – de acordo com Suslick – o custo estimado para perfuração de um poço na bacia de Campos, em lâmina d'água acima de 2.500 metros, não é menos de US\$ 15.000.000 (quinze milhões de dólares); e estima-se que a cada dez poços perfurados, de um a três resultam em acumulações comerciais. Ainda de acordo com Suslick, as taxas médias de sucesso de poços pioneiros perfurados nas bacias petrolíferas localizadas em zonas de fronteira (com escasso conhecimento geológico) situam-se numa faixa entre 20-30%, dependendo da complexidade da bacia.

Dessa forma, conclui-se que quanto mais precisas forem as estimativas dos volumes recuperáveis, menores serão as perdas financeiras e, sendo assim, as empresas de petróleo vêm investindo cada vez mais em pesquisas, com a intenção de obter inovações tecnológicas que reduzam não somente os riscos do processo exploratório e produtivo de petróleo, mas principalmente promovam a redução de custos exploratórios possibilitando o acesso a prospectos em zonas ainda mais remotas.

2.1 Estimativas de Volumes

Os investimentos necessários para a implantação de um projeto de exploração e produção de petróleo, assim como os custos para mantê-lo em operação, devem ser pagos com a receita obtida com a comercialização dos fluidos a serem produzidos a partir da exploração do reservatório. Dessa forma, o conhecimento da

quantidade de fluido existente em uma jazida de petróleo ou, mais especificamente, da quantidade de fluido que dela pode ser extraído, desempenha um papel fundamental na decisão de se implantar ou não um projeto exploratório [57]. Sendo assim, encontrar métodos mais eficientes e confiáveis para obter estimativas dos volumes a serem extraídos das jazidas de petróleo tem sido uma das grandes preocupações das empresas petrolíferas. De acordo José Eduardo Thomas, não existe uma maneira única de se estimar os volumes originais de hidrocarbonetos e as reservas de uma jazida de petróleo. Esses volumes podem ser calculados em geral pelos seguintes métodos: volumétrico, analogia, análise de riscos e análise de desempenho do reservatório [57].

O cálculo do volume original de um reservatório de óleo ou de gás, ou seja, da quantidade de fluido existente no reservatório na época de sua descoberta, é feito pelo método volumétrico. Nesse processo são necessárias as seguintes informações sobre o reservatório: volume total da rocha portadora de hidrocarbonetos, que é obtido por meio da sísmica de reflexão; porosidade média da rocha e saturações dos fluidos, que podem ser obtidos tanto por meio de interpretação de perfis como em ensaios de laboratório; e o fator volume de formação do fluido, que se obtém por meio de uma análise feita em laboratório [57].

O cálculo do volume recuperável, ou seja, da quantidade de óleo ou gás que se espera produzir de uma acumulação de petróleo, pode ser feito tanto pelo método da analogia como pelo método da análise de riscos, ambos em uma época que se precede a perfuração do primeiro poço a penetrar na jazida, ou seja, do poço descobridor. Nessa época tem-se uma série de evidências, mas não se tem a comprovação da existência de uma acumulação de petróleo na região que está sendo pesquisada; assim, as estimativas são feitas a partir de dados sísmicos e resultados de reservatórios localizados nas proximidades, os quais, se acredita, tenham características semelhantes as do reservatório que está sendo estudado. A diferença entre esses dois métodos consiste no fato de que o método da análise de riscos exige certa sofisticação no tratamento estatístico dos dados e os resultados são apresentados não como um valor único (diferente do método de analogia), mas como uma faixa de resultados possíveis [57].

Para o cálculo do volume de reservas, ou seja, da quantidade de fluido que ainda pode ser obtida de um reservatório numa época qualquer de sua vida produtiva, é feito por métodos baseados na análise do desempenho do reservatório.

Esses métodos fazem uma previsão do comportamento futuro do reservatório a partir do seu comportamento passado, utilizando registros do histórico de produção sobre as pressões, as vazões, as relações RGO (Razão Gás-Óleo), RAO (Razão Água-Óleo) e BSW (*Basic Sediments and Water* – Sedimentos Básicos e Água), e as produções acumuladas, ou seja, os volumes totais que foram produzidos de cada fluido. Em alguns casos são necessárias também informações sobre o mecanismo de produção do reservatório [57]. Dependendo de fatores tais como a quantidade e o tipo de dados disponíveis sobre a rocha e fluido e dos recursos de informática existentes (*software* e *hardware*), a análise de desempenho dos reservatórios ou a previsão de comportamento futuro pode ser realizada pelos seguintes métodos:

- a) **Análise de declínio de produção:** é um processo bastante simplificado, uma vez que não se utilizam informações sobre as propriedades da rocha-reservatório, sobre o comportamento dos fluidos ou sobre as relações rocha-fluido. Este método se baseia apenas na observação do comportamento das vazões de produção ao longo do tempo e, a partir da análise desse histórico, pode-se caracterizar a tendência de declínio da vazão. A partir dessa tendência passada, estima-se o comportamento futuro da produção [57].

- b) **Equação de balanço de materiais:** é uma relação que associa o balanço de massa dos fluidos do reservatório com as reduções de pressão no seu interior. A equação de balanço de materiais representa matematicamente a seguinte expressão: “em uma determinada época da vida de um reservatório, a massa dos fluidos existentes no reservatório somada à massa dos fluidos já produzidos é igual à massa de fluidos originalmente existente nesse meio poroso”. A equação de balanço de materiais é escrita em função das propriedades da rocha e do comportamento do fluido em função da pressão, das propriedades rocha-fluido, do histórico de produção, e é particularizada para cada caso, dependendo dos mecanismos de produção atuantes no reservatório. A equação de balanço de materiais trata o reservatório como se fosse um único bloco com propriedades uniformes e fornece apenas relações de produção acumulada de fluidos versus queda de

pressão. Para o estudo do comportamento futuro do reservatório são necessárias outras equações que relacionem as produções acumuladas com vazões de produção e tempos [57].

- c) **Simulação numérica de reservatório:** este método utiliza os mesmos procedimentos da equação de balanço de materiais, mas se diferencia, pois fornece os seus resultados em função do tempo, permite a subdivisão do reservatório em células com propriedades diferentes e envolve a solução simultânea de um grande número de equações que representam o fluxo no meio poroso [57].

A precisão das estimativas de volumes irá depender basicamente da quantidade e qualidade das informações usadas e do próprio método de cálculo. Para a estimativa do volume de reservas ou para o estudo do comportamento de reservatórios, os simuladores numéricos de reservatórios tem sido os mais utilizados, pois além de apresentar mais sofisticação que os outros métodos, permitem também o emprego de técnicas para a quantificação de incertezas. Pela importância dos simuladores de reservatórios no processo exploratório e produtivo de petróleo, é feita uma apresentação mais detalhada sobre eles na seção 2.3.

2.2 Processos de Extração de Petróleo

Ao longo da vida útil de um reservatório de petróleo, muitas decisões importantes são tomadas e grandes investimentos são realizados para que a produção do óleo ocorra conforme premissas adotadas para a implantação do projeto.

O processo de extração ou retirada do petróleo dos espaços vazios da rocha, ao contrário do que se pensa, não é uma tarefa tão fácil, pode ser subdividido em três categorias principais: primário, secundário e terciário [57].

O processo de extração classificado como primário é resultante da atuação da energia natural do reservatório, ou seja, quando a pressão do reservatório é ainda suficiente para provocar a liberação do petróleo nos poços perfurados. Já o processo secundário tem início quando a pressão do reservatório cai, bem como, conseqüentemente, o nível de produção. Para o aumento da pressão e manutenção

da produção são empregados mecanismos de injeção de fluidos nos poços injetores, de forma a bombear o petróleo para os poços produtores.

Estimativas feitas em diversos locais têm conduzido a um fator de recuperação médio de 30% considerando-se apenas os processos convencionais, ou seja, o primário e o secundário [57]. Como nem sempre o aspecto mais crítico do fluxo dos fluidos nos meios porosos é a baixa pressão, as baixas recuperações resultantes do processo secundário podem ser creditadas basicamente a dois aspectos principais: alta viscosidade do óleo e altas tensões interfaciais entre o fluido injetado e o óleo. Dessa forma, após avaliação prévia de sua viabilidade, o processo terciário é iniciado, visando à retirada de parte do óleo que não pode ser recuperado pelo processo secundário. O processo terciário emprega mecanismos para aumentar a mobilidade do óleo no meio poroso e facilitar a sua retirada. Dependendo das características e propriedades do reservatório, o processo terciário ou a recuperação aprimorada de petróleo pode empregar diferentes métodos para a retirada do óleo dos poros das rochas, os quais podem ser classificados em três categorias principais [57]:

- a) **Métodos Térmicos:** muito utilizados em reservatórios cujo óleo é altamente viscoso. Nesse método o aumento da mobilidade do óleo no meio poroso ocorre pelo aumento da temperatura do reservatório com a injeção de fluidos aquecidos (água na forma de vapor ou em estado líquido com temperatura elevada) ou pela combustão de parte do óleo existente no interior do próprio reservatório [57].
- b) **Métodos Miscíveis:** muito utilizados quando são identificadas baixas eficiências de deslocamento, ou seja, onde o fluido injetado não consegue retirar o óleo para fora dos poros da rocha devido a altas tensões interfaciais. Para a eliminação dessas tensões interfaciais, esses métodos promovem a injeção de fluidos como dióxido de carbono, gás natural ou nitrogênio, os quais se tornam miscíveis com o óleo do reservatório deslocando-o para fora da área que for contatada pelo fluido injetado [57].

c) **Métodos Químicos:** eles pressupõem certa interação química entre o fluido injetado e o do reservatório e, por esta razão, em geral, se apresentam mais eficientes do que os métodos miscíveis em relação ao processo de varrida do óleo. Dessa forma, dependendo do que se quer atacar, se a viscosidade ou as tensões interfaciais, os fluidos injetados podem conter soluções de polímeros, soluções de tensoativos, soluções alcalinas e micro emulsões que irão interagir com o fluido do reservatório e agir cada um deles de forma diferenciada [57].

Alguns outros métodos vêm sendo pesquisados para a extração aprimorada do óleo e não se enquadram em nenhuma dessas categorias, como é o caso da recuperação microbiológica e da recuperação utilizando-se ondas eletromagnéticas [57]. A recuperação microbiológica é obtida a partir da utilização de diferentes microorganismos que, quando adequadamente escolhidos, produzem através dos seus processos biológicos, uma série de substâncias que causam diversos efeitos no interior do reservatório, aumentando a recuperação do óleo [57]. Já a recuperação através de ondas eletromagnéticas ocorre pelo aquecimento do reservatório por meio de ondas eletromagnéticas geradas pela aplicação de uma diferença de potencial entre os poços do campo [57].

2.3 Simuladores de Reservatórios

A simulação numérica é hoje uma das mais poderosas ferramentas para guiar decisões de gerenciamento em todos os estágios da vida de um reservatório de petróleo. Nos últimos 10 anos, com o desenvolvimento do modelo computacional geostatístico e geológico, os simuladores numéricos têm ajudado inclusive a testar a validade de seus próprios modelos. Um simulador numérico contendo informações corretas pode emular o comportamento de um reservatório, predizendo a sua produção sob condições de operação corrente, ou sua reação mediante alteração dessas condições [1].

As simulações numéricas de reservatórios em geral visam responder, entre outras, as seguintes questões abaixo, sob a forma de recomendações para melhoria do processo produtivo, as quais estão diretamente associadas à questão econômica

[1]: i) Onde deveriam estar localizados os poços de forma a maximizar o retorno do investimento?; ii) Quantos poços são necessários para produzir óleo em quantidade suficiente para garantir a entrega contratual?; iii) O óleo deveria ser recuperado por esgotamento natural ou por injeção de água?; iv) Qual o tamanho ótimo de um poço horizontal?; v) O dióxido de carbono (CO₂) é uma injeção viável?; vi) O reservatório deveria ser mantido ativo?.

Uma das funções da simulação do reservatório é a de auxiliar no entendimento do comportamento da pressão e produção de um reservatório e, conseqüentemente, na predição das taxas de produção como uma função do tempo. A previsão da produção futura, quando expressa em termos de rendimentos e comparada com custos e investimentos, determina a viabilidade da operação [1].

Uma vez que o objetivo da simulação tenha sido determinado, o passo seguinte é a descrição do reservatório em termos de volume de óleo ou gás localizado, a quantia a ser recuperada e a taxa na qual ocorrerá essa recuperação. Para a estimativa das reservas recuperáveis, deve ser construído um modelo de reservatório, incluindo as falhas, camadas e as propriedades associadas a elas. Este modelo, chamado de estático, é criado através da combinação dos esforços de geólogos, geofísicos, petrofísicos e engenheiros de reservatórios, e nele são gastos milhões de dólares em serviços de campo de óleo, para a obtenção de informações confiáveis que alimentem os simuladores de reservatório [1].

Em geral, os simuladores de reservatórios se encaixam em duas categorias principais: na primeira categoria estão os simuladores “*black-oil*”, para reservatórios contendo água, gás e óleo. O gás pode se mover dentro ou fora da solução com o óleo. Na segunda categoria estão os simuladores térmicos e composicionais, para reservatórios que requerem descrição mais detalhada da composição do fluido. Uma descrição composicional poderia abranger as quantidades e propriedades do hexano, pentano, butano, benzeno, asfaltenos e outros hidrocarbonetos, e pode ser usada quando a composição dos fluidos muda durante a simulação. Um simulador térmico considera a mudança de temperatura para modificar a composição do fluido. Uma situação que exemplifica esse caso seria a injeção de vapor ou injeção de água dentro de um reservatório profundo e quente [1].

Em todos os tipos de simuladores, uma das principais dificuldades é a aquisição das informações utilizadas para a sua alimentação, uma vez que não se tem o total conhecimento do comportamento de todas as propriedades petrofísicas de um

reservatório de petróleo. Essas informações incompletas geram incertezas no processo e podem provocar grandes distorções nas estimativas produzidas pelos simuladores de reservatórios se não tratadas adequadamente. Para a correção ou minimização dessas distorções, as informações com essas características são descritas pelos métodos de propagação de incertezas, sob forma probabilística, e são denominadas como variáveis incertas. No entanto, os simuladores de reservatórios, que por natureza são determinísticos, devem ser alimentados com informações também determinísticas. Sendo assim, os métodos de propagação de incertezas, de acordo com a sua abordagem, calculam valores determinísticos para as variáveis incertas a partir de suas distribuições de probabilidades e de seus parâmetros estatísticos, para a alimentação do simulador de reservatório. Dessa forma, havendo entradas de dados com essas características, ou seja, variáveis incertas, os simuladores de reservatórios irão propagar essas incertezas produzindo também saídas incertas.

3 Propagação de Incertezas

Na linguagem comum, a palavra incerteza está associada ao conceito geral de dúvida. No entanto, no contexto da propagação de incertezas, a palavra refere-se à incerteza de um resultado em virtude do conhecimento limitado a respeito de um dado valor que interfere neste resultado. Dessa forma, o conhecimento da “incerteza” e da sua propagação implica no acréscimo de confiança e na validade do resultado de uma avaliação. A não caracterização da incerteza, nas estimativas usadas como base para a tomada de decisões importantes, pode causar grandes distorções em relação aos valores reais e, por essa razão, há uma grande preocupação atual em encontrar métodos que possam caracterizar as incertezas de forma adequada para a redução ou eliminação desses tipos de erros.

A prática comum na engenharia é analisar sistemas baseados em modelos matemáticos determinísticos e com dados de entrada definidos precisamente. No mundo real, raramente essas condições ideais são encontradas, como é o caso de reservatórios de petróleo, onde algumas propriedades não podem ser simplesmente representadas por um único valor. Essas propriedades são consideradas não determinísticas, ou seja, um único valor não irá representar o seu comportamento em todo o campo. O conhecimento sobre elas é extremamente limitado, pois só podem ser medidas - ou até mesmo estimadas através de correlações - em apenas alguns poucos poços exploratórios. No restante do reservatório estas propriedades são estimadas através de técnicas geostatísticas, que são, por definição, estatísticas, de forma que, em um ponto arbitrário do reservatório, estas propriedades não são conhecidas exatamente, mas sim através de momentos como a média e a variância. A porosidade, definida como a característica de uma rocha armazenar fluidos em seus espaços interiores, pode ser citada como exemplo de uma dessas propriedades. Devido às variabilidades geológicas existentes na região que está sendo explorada, a porosidade obtida através de uma amostra extraída apenas de alguns poços exploratórios não pode definir exatamente o comportamento dessa variável para todo o campo. Fica evidente então que, se a “porosidade”, variável de entrada na simulação de reservatório, for representada como uma variável determinística, ou seja, apenas com um valor conhecido e não com todos os valores possíveis, os resultados da simulação serão inconsistentes. A inclusão da incerteza

na representação de variáveis cujo comportamento não é conhecido por completo em todas as situações é essencial para representar a incerteza nas saídas resultantes dos sistemas. Dessa forma, essas variáveis, denominadas como variáveis incertas, devem ser descritas sob forma probabilística, pois assumem valores aleatórios. Para o tratamento das variáveis incertas, modelos matemáticos específicos são utilizados, e devido a isso a propagação de incertezas passa a ser naturalmente conceituada como uma modelagem estocástica. A seção 3.2 apresenta os métodos mais utilizados para a propagação de incertezas.

3.1 Variáveis Incertas

As variáveis incertas são definidas como aleatórias, podendo ser contínuas ou discretas. Se em um problema existirem variáveis aleatórias discretas, então estas variáveis podem ser modeladas usando-se a representação de histograma baseada em pontos. Em geral, as variáveis incertas são modeladas como variáveis aleatórias contínuas ou, no caso de histograma, com uma representação de um histograma empírico. A descrição das variáveis contínuas normalmente é feita por uma distribuição de probabilidade, com parâmetros estatísticos calculados com base em dados experimentais. Os tipos de distribuição de probabilidade mais utilizados para a representação desses tipos de variáveis são [43]:

- Normal
- Lognormal
- Uniforme
- Loguniforme
- Triangular
- Exponential
- Beta
- Gamma
- Gumbel
- Frechet
- Weibull
- Histograma
- Intervalo

A permeabilidade de um meio poroso é a medida de sua capacidade de se deixar atravessar por fluidos. No caso de reservatórios de petróleo, a permeabilidade é a propriedade que indica a medida da condutividade dos fluidos (óleo, gás e água) na rocha, ou seja, no meio poroso. Tal qual a porosidade, o valor da permeabilidade é determinado a partir de amostras extraídas em alguns poços exploratórios; portanto, devido às variabilidades geológicas, não é possível afirmar que seja a mesma para todos os pontos do reservatório. Sendo assim, a permeabilidade é uma variável incerta, pois o seu comportamento não é totalmente conhecido, e para pontos arbitrários do reservatório o seu valor será determinado através de correlações que utilizam a posição do ponto, a média e a variância.

Por ser uma propriedade que interfere no cálculo das vazões do reservatório, a permeabilidade, selecionada neste trabalho como sendo a variável incerta de interesse, foi definida nos casos de estudo apresentados, como aleatória contínua com distribuições do tipo normal e lognormal.

3.2 Métodos para Propagação de Incertezas

Os métodos de propagação de incertezas, também conhecidos como métodos de análise não determinística, envolvem o cálculo de informação probabilística sobre as funções respostas, baseados num conjunto de simulações que tiveram como entrada uma distribuição de probabilidade especificada, referente às variáveis incertas. Esses métodos efetuam a propagação das incertezas, nos quais a informação de probabilidade dos parâmetros de entrada é mapeada para uma informação de probabilidade nas funções respostas de saída [43].

Muitos métodos têm sido desenvolvidos para solucionar sistemas estocásticos, visando sempre maior rapidez na convergência. Existe atualmente um conjunto de métodos que são mais populares nas aplicações de engenharia, os quais podem ser classificados genericamente em duas categorias principais: métodos que usam amostragem direta e métodos que não usam amostragem direta.

3.2.1 Métodos que Usam Amostragem Direta

A abordagem por amostragem direta inclui: Simulação de Monte Carlo (SMC) ([15], [33], [41]), *Latin Hypercube Sampling* (LHS) ([30], [51], [54]), entre outros. Esses métodos envolvem amostragem e estimativa e, na maioria dos casos, possuem implementações relativamente simples. A Simulação de Monte Carlo, técnica de amostragem mais antiga e bastante utilizada até os dias atuais, embora apresente excelente precisão torna-se inviável para certas aplicações, em virtude da sua baixa taxa de convergência. Para aceleração da convergência, outras técnicas de amostragem foram sendo desenvolvidas, como o *Latin Hypercube Sampling*, o método de Quasi-Monte Carlo (QMC) ([16], [33]) o método de Cadeia de Monte Carlo Markov (CMCM) ([17], [21]), o método de Aproximação de Superfície de Resposta para Análise de Confiabilidade ([14], [39]), etc. Esses métodos são mais eficientes do que o método de Monte Carlo, mas algumas restrições adicionais, baseadas em seus próprios projetos específicos, lhes são impostas, tornando sua aplicabilidade limitada.

Nos casos de simulações de reservatórios, onde as incertezas estão geralmente associadas a campos de permeabilidade e porosidade, certas dificuldades podem ser previstas no que diz respeito à propagação dessas incertezas com o uso de técnicas clássicas de amostragem, como por exemplo, Monte Carlo ou Quasi-Monte Carlo. Nesses casos é necessário um número de amostras muito grande (simulações determinísticas), acarretando num tempo de processamento muito elevado e não aceitável para a solução do problema. Sendo assim, mesmo com o auxílio da computação paralela com *clusters*, ou mesmo com processamento distribuído em várias máquinas, o estudo pode tornar-se inviável devido ao tempo de processamento da quantidade de amostras exigidas.

Como a Simulação de Monte Carlo é utilizada neste trabalho como parâmetro de comparação com os métodos que não usam amostragem direta, é feita uma breve apresentação desse método na próxima seção.

3.2.1.1 Simulação de Monte Carlo (SMC)

De acordo com Hammersley [23], o nome "Monte Carlo" dado ao método pelo matemático austríaco Stanislaw Ulam, em virtude dos jogos de roletas nos cassinos de Mônaco, surgiu durante a Segunda Guerra Mundial no projeto Manhattan, para a construção da bomba atômica [60]. Nesse projeto, Stanislaw Ulam sugeriu o uso do método de Monte Carlo para avaliar integrais matemáticas complicadas que surgiam na teoria das reações nucleares em cadeia. Embora a atenção desses cientistas tenha sido despertada em 1948, a lógica do método já era conhecida há bastante tempo, fato comprovado pela existência de um registro de um artigo escrito por Lord Kelvin, dezenas de anos antes, que já utilizava técnicas de Monte Carlo em uma discussão das equações de Boltzmann [60]. No entanto, a sugestão de Stanislaw Ulam para a utilização do método de Monte Carlo no projeto de construção da bomba atômica, levou Von Neumann, Nicholas Metropolis e outros a um desenvolvimento mais sistemático do método de Monte Carlo [60].

Como mencionado anteriormente, o método de Monte Carlo tem sido utilizado há bastante tempo como forma de se obter aproximações numéricas de funções complexas e a sua maior aplicação ocorre na computação numérica para avaliação de integrais. A ideia do método é escrever a integral que se deseja calcular como um valor esperado, o que torna possível a simulação de qualquer processo influenciado por fatores aleatórios. Desta forma, este método permite a resolução aproximada de problemas físicos e matemáticos através da simulação de quantidades aleatórias de valores associados a um determinado domínio.

Sua construção é relativamente simples, sendo necessário apenas definir seus processos de transformação, os valores de parâmetros e ter um eficiente gerador de números aleatórios à disposição. No entanto, apresenta baixa taxa de convergência, sendo às vezes necessárias milhares de simulações para se alcançar um valor com alto grau de precisão para a variável estimada.

Visto que o objetivo do método Simulação de Monte Carlo é desenvolver diversos eventos com a mesma possibilidade de ocorrência, este gera continua e aleatoriamente números não tendenciosos que representam uma distribuição de probabilidade de um evento ou cenário provável de acontecer. Através de ferramen-

tas estatísticas, como a média e o desvio padrão, é possível avaliar a probabilidade destes eventos em uma distribuição.

Para este método podem ser abordados três tipos de variáveis: as aleatórias que serão geradas aleatoriamente, as independentes e as dependentes. De acordo com J. F. Correa [8], as primeiras têm seus possíveis valores dentro de uma faixa identificada por um valor mínimo e um máximo. As variáveis independentes são constantes (desvio padrão igual à zero) ou variam de outra forma, independentemente das variáveis inicialmente identificadas. Outras variáveis podem ser diagnosticadas como dependentes das variáveis a serem geradas aleatoriamente; identificadas essas variáveis, uma relação de proporcionalidade deve ser determinada entre elas. Dessa forma, em cada geração aleatória de valores, as variáveis dependentes terão automaticamente seus valores calculados, através das relações percentuais firmadas entre elas [8].

A precisão do método de Monte Carlo não é dependente da dimensionalidade do sistema, mas sim do número de realizações usadas para caracterizá-lo. Nesse tipo de abordagem, a grande quantidade de avaliações de função, necessárias para gerar estatísticas convergidas, pode resultar numa análise computacionalmente muito cara, e em alguns casos impraticáveis, em aplicações de engenharia do mundo real.

3.2.2 Métodos que Não Usam Amostragem Direta

Muitos esforços em pesquisas vêm sendo feitos para o desenvolvimento de métodos que não usam amostragem direta. Um dos métodos mais populares é o método de perturbação ([26], [32]), onde um campo aleatório é expandido em torno de sua média, via séries de Taylor, e truncado numa certa ordem. Em geral, apenas expansões de segunda ordem são empregadas, visto que o sistema de equações se torna extremamente complicado para expansões além dessa ordem. Essa abordagem, também conhecida como “Análise de Segundo Momento”, tem sido muito usada em vários campos. Uma limitação inerente dos métodos de perturbação é que as incertezas não podem ser muito grandes, isto é, os desvios de campos aleatórios não podem ser muito maiores que os valores de suas médias, ou seja,

devem ser menores que 10%. Além disso, estatísticas de mais alta ordem não estão prontamente disponíveis para métodos de Segundo Momento [62].

Outra abordagem para métodos que não usam amostragem direta é baseada na manipulação de operadores estocásticos. Métodos ao longo dessa linha de abordagem incluem a Expansão de Neumann, que é baseada no inverso do operador estocástico numa série de Neumann [63], e o método da integral ponderada ([10], [55]). Esses métodos têm limitações no tipo das equações modelo que eles podem tratar, e também estão restritos a pequenas incertezas [62].

De acordo com a Xiu e Karniadakis [62], outra metodologia de abordagem não amostral é a discretização direta de um campo aleatório no espaço aleatório, na qual Ghanem e Spanos [20] se basearam para desenvolver pioneiramente um método de expansão denominado “Caos Polinomial”, o qual foi aplicado com sucesso em vários problemas da mecânica [20] e, sendo objeto de estudo desse trabalho, será apresentada em mais detalhes no capítulo 4.

4 Expansão em Caos Polinomial

Dentre as metodologias de abordagem não amostral para a propagação de incertezas destaca-se a Expansão em Caos Polinomial [43], a qual é baseada na teoria do caos homogêneo de Wiener [59]. A Expansão em Caos Polinomial foi proposta seguindo o trabalho mais fundamental da teoria estocástica [49] e dos polinômios ortogonais [5]. Seu destaque se dá principalmente por sua rápida convergência e por expressar a solução final como um processo aleatório, e não meramente como um conjunto de estatísticas, como é o caso de muitas metodologias não determinísticas. Isto faz a técnica particularmente atrativa para uso em aplicações multifísicas que ligam pacotes de análise diferentes. O Caos Polinomial, embora estime probabilidades pontuais, forma uma aproximação para a relação funcional entre as funções respostas e as suas entradas aleatórias, provendo uma representação mais completa da incerteza, para ser usada em simulações computacionais [43].

O desenvolvimento do método do Caos Polinomial espelha a análise de elementos finitos determinísticos, utilizando noções de projeção, ortogonalidade e convergência fraca ([19], [20]), e tem como objetivo a caracterização das respostas dos sistemas cujas equações governantes envolvem coeficientes estocásticos [43]. Este método se enquadra na linha de abordagem dos métodos de expansão estocástica, e teve o desenvolvimento de sua expansão original baseada apenas nos polinômios de Hermite, razão esta de ser denominada aqui por “Expansão de Wiener-Hermite”. Essa expansão tem sido efetiva na resolução de equações diferenciais estocásticas com entradas Gaussianas e certos tipos de entradas não Gaussianas, a exemplo da distribuição lognormal, o que pode ser justificado pelo teorema Cameron-Martin [4]. Entretanto, para entradas aleatórias não Gaussianas, a convergência pode ser lenta e em alguns casos inaceitável. De forma a permitir que o método do Caos Polinomial também proporcionasse rápida convergência para entradas aleatórias não Gaussianas, a Expansão em Caos Polinomial foi generalizada, mediante a ampliação da base de campos aleatórios, com a utilização dos polinômios ortogonais do Esquema de Askey ([28], [49]). Por esta razão essa Expansão em Caos Polinomial generalizada é denominada aqui como “Expansão de Wiener-Askey”. Devido ao fato de que a diferença entre as duas expansões diz

respeito apenas aos tipos de polinômios ortogonais que cada uma utiliza, e por ser a Expansão em Caos Polinomial de Wiener-Askey a mais abrangente, na seção 4.1 será feito o detalhamento apenas desta, em termos de bases para representação de campos aleatórios.

Para possibilitar maior rapidez na convergência, no cálculo dos coeficientes da Expansão em Caos Polinomial são empregadas as seguintes técnicas: a projeção de Galerkin, detalhada na seção 4.2, e a regressão linear, detalhada na seção 4.3. Os momentos estatísticos da saída podem ser facilmente calculados utilizando-se diretamente os coeficientes da Expansão em Caos Polinomial, conforme mostrado na seção 4.4. Da mesma forma, a análise de sensibilidade da variável aleatória gerada por essa metodologia será apresentada na seção 4.5.

4.1 Expansão em Caos Polinomial de Wiener-Askey

A Expansão de Wiener-Askey emprega, além dos polinômios de Hermite, outros polinômios ortogonais pertencentes ao Esquema de Askey, que podem ser classificados e relacionados entre si numa estrutura em árvore, conforme Figura 2. Nesta expansão, as variáveis aleatórias de base não estão restritas apenas as variáveis aleatórias Gaussianas, ou seja, os polinômios do Esquema de Askey, Hermite, Legendre, Laguerre, Jacobi e Legendre generalizado são usados para a modelagem do efeito das variáveis aleatórias contínuas descritas por distribuições de probabilidade uniforme, exponencial, beta e gama, respectivamente. Essas seleções de polinômios são ideais para esses tipos de distribuição uma vez que o produto interno das funções peso (w) e suas faixas de suporte associadas correspondem às funções de densidade para esses tipos de distribuições, como pode ser visto na Tabela 1. Na teoria, taxas de convergência exponenciais em momentos respostas podem ser obtidas com o uso destas bases. Além disso, diversos tipos de expansões discretas são igualmente incorporados, aumentando ainda mais a flexibilidade da Expansão em Caos Polinomial generalizada [43].

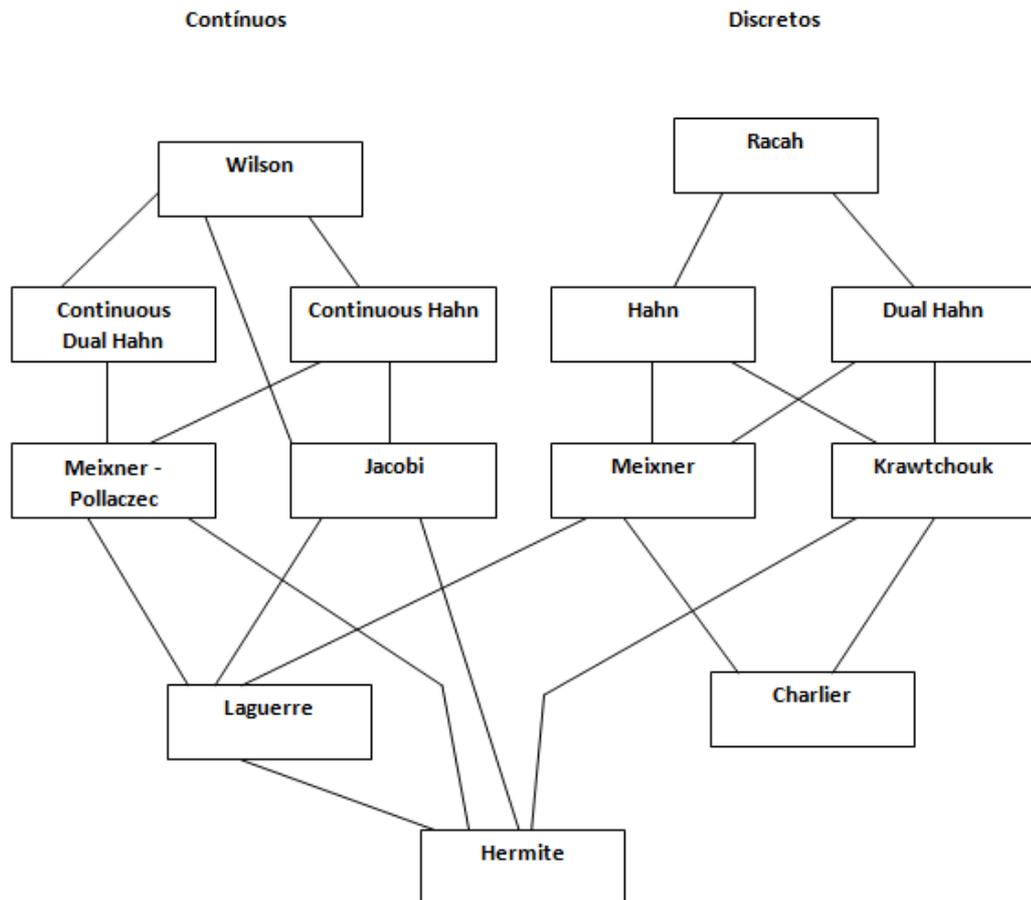


Figura 2. Polinômios Ortogonais do Esquema de Askey.

Tabela 1. Ligação entre as formas padrão das distribuições de probabilidades.

Distribuição	Função Densidade de Probabilidade	Família de Polinômios	Função Peso $w(x)$	Faixa Suporte
Gaussiana (Normal)	$\frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$	Hermite	$e^{-\frac{x^2}{2}}$	$[-\infty, \infty]$
Uniforme	$\frac{1}{2}$	Legendre	1	$[-1, 1]$
Beta	$\frac{(1-x)^\alpha (1+x)^\beta}{2^{\alpha+\beta+1} B(\alpha+1, \beta+1)}$	Jacobi	$(1-x)^\alpha (1+x)^\beta$	$[-1, 1]$
Exponencial	e^{-x}	Laguerre	e^{-x}	$[0, \infty]$
Gama	$\frac{x^\alpha e^{-x}}{\Gamma(\alpha+1)}$	Laguerre Generalizado	$x^\alpha e^{-x}$	$[0, \infty]$

A maioria dos processos físicos é composta de processos aleatórios de segunda ordem com variância finita. Sendo assim, um processo aleatório de segunda ordem $R(w)$, (o evento aleatório) visto como uma função de w (função peso especificada na Tabela 1), pode ser representado na Expansão em Caos Polinomial, na forma da Equação (1):

$$\begin{aligned}
 R(w) &= c_0 I_0 \\
 &+ \sum_{i_1=1}^{\infty} c_{i_1} I_1(\zeta_{i_1}(w)) \\
 &+ \sum_{i_1=1}^{\infty} \sum_{i_2=1}^{i_1} c_{i_1 i_2} I_2(\zeta_{i_1}(w), \zeta_{i_2}(w)) \\
 &+ \sum_{i_1=1}^{\infty} \sum_{i_2=1}^{i_1} \sum_{i_3=1}^{i_2} c_{i_1 i_2 i_3} I_3(\zeta_{i_1}(w), \zeta_{i_2}(w), \zeta_{i_3}(w)) + \dots
 \end{aligned} \tag{1}$$

onde $I_n(\zeta_{i_1}, \dots, \zeta_{i_n})$ representam os polinômios do Caos Polinomial generalizado de ordem n , em termos de variáveis aleatórias multidimensionais $\zeta = (\zeta_{i_1}, \dots, \zeta_{i_n})$ e os coeficientes da expansão c_{i_1, \dots, i_n} .

A Equação (1) é uma versão da Expansão em Caos Polinomial de Wiener para variáveis aleatórias discretas, na qual para o caso de variáveis aleatórias contínuas os somatórios são substituídos por integrais.

Na Expansão em Caos Polinomial de Wiener-Askey, os polinômios I_n não estão restritos aos polinômios de Hermite, ou seja, podem ser qualquer um dos membros do Esquema de Askey, como mostrado na Figura 2. Na Equação (1) cada conjunto adicional do somatório aninhado indica uma ordem adicional dos polinômios na expansão. A Equação (1) pode ser simplificada pela reformulação da expansão de uma indexação baseada em ordem para uma indexação baseada em termo, conforme a Equação (2):

$$R(w) = \sum_{j=0}^{\infty} \hat{c}_j \Phi_j(\zeta) \tag{2}$$

onde há uma correspondência de um para um entre as funções $I_n(\zeta_{i_1}, \dots, \zeta_{i_n})$ e $\Phi_j(\zeta)$ e também entre os coeficientes \hat{c}_j e c_{i_1, \dots, i_r} . O Caos Polinomial de Wiener-Askey forma uma base. As funções base multidimensionais, ou seja, os polinômios

multivariados $\Phi_j(\zeta)$ são produtos tensoriais de bases polinomiais de uma dimensão. Por exemplo, o polinômio de Hermite multivariado $H(\zeta)$ de ordem n é definido pela expressão geral, conforme Equação (3) [43]:

$$H_n(\zeta_{i_1}, \dots, \zeta_{i_n}) = e^{\frac{1}{2}\zeta^T \zeta} (-1)^n \frac{\partial^n}{\partial \zeta_{i_1} \dots \partial \zeta_{i_n}} e^{-\frac{1}{2}\zeta^T \zeta} \quad (3)$$

que pode ser mostrado como sendo um produto de polinômios de uma dimensão envolvendo um índice múltiplo conforme Equação (4):

$$H_n(\zeta_{i_1}, \zeta_{i_2}, \dots, \zeta_{i_n}) = \Phi_j(\zeta) = \prod_{i=1}^n \Phi_i^{l_i}(\zeta_i), \quad j = 1, 2, \dots, S; \quad l_i = 1, 2, \dots, p \quad (4)$$

onde $S = \frac{(n+p)!}{n!p!}$, sendo n o número de variáveis aleatórias ζ e p a maior ordem da base polinomial.

Os primeiros polinômios de Hermite para uma dimensão são:

$$\Phi_0 = 1, \quad \Phi_1 = \zeta, \quad \Phi_2 = \zeta^2 - 1, \quad \Phi_3 = \zeta^3 - 3\zeta \quad (5)$$

Os primeiros polinômios de Hermite multidimensionais para um caso de duas dimensões (cobrindo os termos de zero, primeira e segunda ordem) são dados pelas equações a seguir:

$$\Phi_0(\zeta) = \Phi_0(\zeta_1)\Phi_0(\zeta_2) = 1 \quad (6)$$

$$\Phi_1(\zeta) = \Phi_1(\zeta_1)\Phi_0(\zeta_2) = \zeta_1 \quad (7)$$

$$\Phi_2(\zeta) = \Phi_0(\zeta_1)\Phi_1(\zeta_2) = \zeta_2 \quad (8)$$

$$\Phi_3(\zeta) = \Phi_2(\zeta_1)\Phi_0(\zeta_2) = \zeta_1^2 - 1 \quad (9)$$

$$\Phi_4(\zeta) = \Phi_1(\zeta_1)\Phi_1(\zeta_2) = \zeta_1\zeta_2 \quad (10)$$

$$\Phi_5(\zeta) = \Phi_0(\zeta_1)\Phi_2(\zeta_2) = \zeta_2^2 - 1 \quad (11)$$

Na prática, a expansão infinita é truncada em um número finito de variáveis aleatórias e em uma ordem finita de expansão, podendo a Equação (2) ser reescrita conforme a Equação (12):

$$R(w) = \sum_{j=0}^P \hat{c}_j \Phi_j(\zeta) \quad (12)$$

De acordo com a referência [43], o número total de termos numa Expansão em Caos Polinomial para uma função resposta envolvendo n variáveis incertas de entrada é dado pela Equação (13):

$$N_t = 1 + P = 1 + \sum_{s=1}^p \frac{1}{s!} \prod_{r=0}^{s-1} (n+r) = \frac{(n+p)!}{n!p!} \quad (13)$$

onde p é a maior ordem dos polinômios que representam as variáveis incertas.

A Tabela 2 mostra a quantidade necessária de coeficientes da expansão, em função do número de variáveis incertas e do grau do polinômio, para alguns valores típicos de variáveis incertas. Observa-se que a quantidade de coeficientes cresce rapidamente, à medida que se aumenta a quantidade de variáveis incertas e a ordem dos polinômios que as representam.

Tabela 2. Número de coeficientes na Expansão em Caos Polinomial.

Grau do Polinômio - p	Número de variáveis incertas - n			
	5	10	15	20
2	21	66	136	231
3	56	286	816	1771
4	126	1001	3876	10626
5	252	3003	15504	53130

Para maior clareza, uma Expansão em Caos Polinomial na forma expandida completa (ver Equação (1)) para duas dimensões é dada pela Equação (14):

$$R(w) = c_0 I_0 + c_{11} I_1(\zeta_1) + c_{21} I_1(\zeta_2) + c_{111} I_2(\zeta_1, \zeta_1) + c_{211} I_2(\zeta_2, \zeta_1) + c_{221} I_2(\zeta_2, \zeta_2) + \dots \quad (14)$$

E na forma simplificada (ver Equação (2)), a Expansão em Caos Polinomial é dada pela Equação (15):

$$R(w) = \hat{c}_0 \Phi_0 + \hat{c}_1 \Phi_1 + \hat{c}_2 \Phi_2 + \hat{c}_3 \Phi_3 + \hat{c}_4 \Phi_4 + \hat{c}_5 \Phi_5 + \dots \quad (15)$$

Substituindo-se, na Equação (15), os valores de Φ_0 , Φ_1 , Φ_2 , Φ_3 , Φ_4 , e Φ_5 dados pelas Equações (6), (7), (8), (9), (10) e (11), respectivamente, tem-se que:

$$R(w) = \hat{c}_0 + \hat{c}_1 \zeta_1 + \hat{c}_2 \zeta_2 + \hat{c}_3 (\zeta_1^2 - 1) + \hat{c}_4 (\zeta_1 \zeta_2) + \hat{c}_5 (\zeta_1^2 - 1) + \dots \quad (16)$$

Desde que cada tipo de polinômio do Esquema de Askey forma uma base completa no espaço de Hilbert, determinado por seus vetores aleatórios correspondentes ζ , espera-se que cada tipo de Expansão em Caos Polinomial generalizado convirja para uma funcional L^2 no espaço funcional de Hilbert. Sendo assim, a relação de ortogonalidade do método do Caos Polinomial generalizado toma a forma da Equação (17):

$$\langle \Phi_i, \Phi_j \rangle = \langle \Phi_i^2 \rangle \delta_{ij} \quad (17)$$

onde δ_{ij} é o delta de Kronecker e $\langle \dots \rangle$ representa o valor esperado da função, que é o produto interno no espaço de Hilbert das variáveis aleatórias ζ , dado pela Equação (18).

$$\langle f(\zeta)g(\zeta) \rangle = \int f(\zeta)g(\zeta)W(\zeta)d\zeta \quad (18)$$

ou no caso de variáveis discretas, conforme Equação (19):

$$\langle f(\zeta)g(\zeta) \rangle = \sum_{\zeta} f(\zeta)g(\zeta)W(\zeta) \quad (19)$$

onde $W(\zeta)$ é a função peso correspondente a base do Caos Polinomial generalizado $\{\Phi_i\}$.

4.2 Projeção de Galerkin

A projeção de Galerkin é utilizada no cálculo dos coeficientes da Expansão em Caos Polinomial. A abordagem de Galerkin consiste na projeção da função resposta contra cada função de uma base usando produtos internos, e no emprego das propriedades da ortogonalidade polinomial para extrair cada coeficiente. Sendo R a função resposta e Φ_j uma função de uma base tem-se então a Equação (20) para o cálculo de cada coeficiente:

$$\hat{c}_j = \frac{\langle R, \Phi_j \rangle}{\langle \Phi_j^2 \rangle} \quad (20)$$

A Equação (20) pode ser escrita de forma equivalente conforme a Equação (21):

$$\hat{c}_j = \frac{E[R(h(\zeta))\Phi_j(\zeta)]}{E[\Phi_j(\zeta)^2]} \quad (21)$$

onde $E[\bullet]$ denota o operador da expectativa matemática e $h(\zeta)$ denota a dependência explícita da resposta em um conjunto de variáveis aleatórias $\{\zeta\}$. A equação (21) é equivalente a Equação (22) com integrais multidimensionais:

$$\hat{c}_j = \frac{\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} R(h(\zeta))\Phi_j(\zeta)f_{\{\zeta\}}(\zeta)d\zeta}{\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} \Phi_j(\zeta)^2 f_{\{\zeta\}}(\zeta)d\zeta} = \frac{\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} R(h(\zeta))\Phi_j(\zeta)f_{\{\zeta\}}(\zeta)d\zeta}{\langle \Phi_j^2 \rangle} \quad (22)$$

Nesta equação $f_{\{\zeta\}}(\zeta)$ denota a densidade de probabilidade conjunta de dimensão m da coleção de variáveis aleatórias normais padrão.

Conforme a referência [43], como a dependência explícita de R em $\{\zeta\}$ não é geralmente conhecida, o numerador da Equação (22) deve ser calculado numericamente usando abordagens de amostragem ou de quadratura/cubatura, apresentadas nas seções 4.2.1 e 4.2.2, respectivamente.

Já o termo do denominador da Equação (22) pode ser resolvido analiticamente, usando-se o produto do quadrado das normas uni-variadas, como mostra a Equação (23):

$$\langle \Phi_j^2 \rangle = \prod_{i=1}^n \langle \Phi_{m_i^j}^2 \rangle \quad (23)$$

onde os produtos internos uni-variados têm expressões simples de forma fechada para cada polinômio pertencente ao esquema de Askey [43]. Assim, o esforço computacional principal para o cálculo dos coeficientes da Expansão em Caos Polinomial reside na avaliação do numerador.

4.2.1 Aproximação por Amostragem

Na aproximação por amostragem, a avaliação da integral do numerador da Equação (22) é equivalente a calcular a média do produto da função resposta por cada termo da expansão, quando amostrado dentro da densidade da função peso. Na prática computacional, nas estimativas de coeficientes baseadas em amostragem, a média da função resposta, que é o primeiro coeficiente da Expansão em Caos Polinomial, é estimada primeiramente e esse valor é então removido das avaliações do valor esperado para todos os coeficientes subsequentes. Um dos métodos bastante utilizado nesse tipo de abordagem é a Simulação de Monte Carlo.

O custo computacional para a estimativa do termo do numerador da Equação (22) via médias amostradas é dependente do problema, já que irá requerer N_s cálculos, sendo N_s a quantidade de amostras. Dessa forma, para uma boa precisão é necessário um grande número de amostras, o que acarreta num custo excessivamente alto, tornando o método impraticável.

4.2.2 Aproximação por Quadratura/Cubatura

Na aproximação baseada em quadratura, a mais simples abordagem para a integração multidimensional, um produto tensorial das regras de quadratura Gaussiana de uma dimensão é empregado para cada tipo de polinômio ortogonal

(quadratura de uma dimensão de Gauss-Hermite, Gauss-Legendre, Gauss-Laguerre, Gauss-Laguerre generalizado, ou quadratura de uma dimensão Gauss-Jacobi).

Esta abordagem é excelente do ponto de vista da precisão, por integrar exatamente polinômios até a ordem combinada $2m - 1$, mostrado na Equação (22), para ordem de quadratura m . No entanto, o custo computacional dessa abordagem também é dependente do problema, ou seja, ela sofre com o problema da “maldição da dimensionalidade”, na qual o custo aumenta exponencialmente com o número de variáveis aleatórias (m^n avaliações de função, sendo n o número de dimensões aleatórias e m a mesma ordem para todas as dimensões). Por essa razão, métodos de quadratura de produtos tensoriais anisotrópicos (ou de ordem mista) e de cubatura de Smolyak de grade esparsa têm sido desenvolvidos exibindo grande melhora com o tamanho do problema [34]. Em particular, os métodos de Smolyak demandam $m^{\log m}$ avaliações de função quando os pontos de quadratura para diferentes ordens podem ser aninhados (exemplo, Clenshaw-Curtis ao invés de quadraturas Gaussianas) [43]. Além disso, quadraturas aninhadas permitem a implementação eficiente de esquemas adaptativos: baseados nos erros estimam ou diminuem retornos. Para problemas de baixa dimensão, entretanto, as quadraturas Gaussianas são as mais utilizadas [43].

Sendo assim, nas avaliações dos coeficientes de mais alta ordem (Equação (22)), dada uma ordem p da expansão, presume-se que serão envolvidos integrandos de ordem polinomial de pelo menos $2p$, tal que uma quadratura Gaussiana mínima de ordem $p + 1$ é requerida para obtenção de uma boa precisão nesses coeficientes. Em outras palavras isto quer dizer que, uma vez que a ordem de R é geralmente desconhecida, uma seleção de uma expansão de ordem p deveria ser acompanhada por uma seleção de ordem de quadratura razoável para a captura precisa de um R de ordem p .

4.3 Regressão Linear

A aproximação por regressão linear, também conhecida como colocação de pontos ou superfície de resposta estocástica ([24], [58]), usa uma única solução de

mínimos quadrados linear da forma da Equação (24) para resolver o conjunto completo de coeficientes da Expansão em Caos Polinomial $\hat{\mathbf{c}}$ que melhor se ajuste aos valores resposta \mathbf{R} :

$$\mathbf{\Phi}\hat{\mathbf{c}} = \mathbf{R} \quad (24)$$

O conjunto de valores respostas é tipicamente obtido pela execução de um projeto de experimentos de computador dentro da função densidade de ζ , onde cada linha da matriz $\mathbf{\Phi}$ contém os N_t termos polinomiais multivariados Φ_j avaliados em uma amostra ζ , em particular. Uma amostragem maior é aconselhável (Hosder, Walters e Balch [24] recomendam $2N_t$ amostras), resultando numa solução de mínimos quadrados para o sistema sobre determinado.

4.4 Momentos Analíticos

De acordo com a referência [43], os primeiro e segundo momentos da Expansão em Caos Polinomial estão disponíveis na forma fechada simples, indicadas respectivamente nas Equações (25) e (26), a seguir:

$$\mu_R = \langle R \rangle = \sum_{j=0}^P \hat{c}_j \langle \Phi_j(\zeta) \rangle = \hat{c}_0 \quad (25)$$

$$\sigma_R^2 = \langle (R - \mu_R)^2 \rangle = \left\langle \left(\sum_{j=1}^P \hat{c}_j \Phi_j(\zeta) \right)^2 \right\rangle = \left\langle \sum_{j=1}^P \sum_{k=1}^P \hat{c}_j \hat{c}_k \Phi_j(\zeta) \Phi_k(\zeta) \right\rangle = \sum_{j=1}^P \hat{c}_j^2 \langle \Phi_j^2 \rangle \quad (26)$$

onde o quadrado da norma de cada polinômio multivariado é calculado da Equação (23). Esses são momentos exatos da expansão, que convergem para momentos da função resposta verdadeiros.

4.5 Análise de Sensibilidade da Variável Aleatória

As expansões em Caos Polinomial são facilmente diferenciáveis em relação às variáveis aleatórias [43]. Sendo assim, usando-se primeiramente a Equação (12) tem-se que:

$$\frac{dR}{d\zeta} = \sum_{j=0}^p \hat{c}_j \frac{d\Phi_j(\zeta)}{d\zeta_i} \quad (27)$$

e então usando a Equação (23), chega-se a Equação (28):

$$\frac{d\Phi_j(\zeta)}{d\zeta_i} = \frac{d\Phi_i}{d\zeta_i} \prod_{\substack{k=1 \\ k \neq i}}^n \Phi_{m_k}(\zeta_k) \quad (28)$$

onde as derivadas polinomiais univariadas $\frac{d\Phi_i}{d\zeta_i}$ têm expressões simples de forma fechada para cada polinômio no Esquema de Askey. Finalmente, usa-se o Jacobiano da transformação de variável de Nataf [27], conforme a Equação (29):

$$\frac{dR}{dx_i} = \frac{dR}{d\zeta} \frac{d\zeta}{dx_i} \quad (29)$$

que é simplificada para a Equação (30), no caso de x_i não correlacionada:

$$\frac{dR}{d\zeta_i} \frac{d\zeta_i}{dx_i} \quad (30)$$

5 Colocação Estocástica

O método da Colocação Estocástica é outra técnica de expansão estocástica para a quantificação de incerteza que é conceitualmente bem mais simples do que o método do Caos Polinomial, pois consiste simplesmente em adotar um polinômio interpolador Lagrangeano multidimensional, que passa em uma série de pontos de colocação pré-definidos, N_p , em uma malha regular, como mostrado na Equação (31):

$$R(\xi) \approx \sum_{j=1}^{N_p} r_j(\xi) L_j(\xi) \quad (31)$$

onde $r_j(\xi)$ são os valores respostas em cada ponto de interpolação, e $L_j(\xi)$ os polinômios interpoladores Lagrangeanos.

Como os polinômios de Lagrange têm o valor 1 (um) em um ponto de colocação e o valor 0 (zero) nos demais, os coeficientes da Equação (31) são diretamente os valores da função resposta nos pontos de colocação. A formação de polinômios interpoladores multidimensionais com esta propriedade requer o uso de um conjunto de pontos de colocação estruturado, derivado de produto tensorial ou de malhas esparsas. A chave para a aproximação da função é executar a colocação usando os pontos de Gauss e os pesos dos mesmos polinômios usados no Caos Polinomial generalizado, o que vai resultar na mesma taxa de convergência exponencial. Esta equação é então usada diretamente como aproximação da função desconhecida. Em comum com a quadratura Gaussiana para o cálculo dos coeficientes da Expansão em Caos Polinomial, a Colocação Estocástica em uma malha isotrópica regular pode demandar um número excessivo de pontos de colocação, já que o número de pontos também cresce exponencialmente com o número de variáveis incertas. Para minimização desse problema são utilizadas técnicas de malhas esparsas [43].

A diferença prática entre os métodos do Caos Polinomial e da Colocação Estocástica, é que no primeiro os coeficientes da expansão são desconhecidos e devem ser estimados para funções base conhecidas, enquanto que no segundo os

polinômios interpoladores Lagrangeanos multidimensionais devem ser formados para coeficientes já conhecidos [43].

5.1 Momentos Analíticos

De acordo com a referência [43], o primeiro e segundo momentos do método da Colocação Estocástica estão disponíveis na forma fechada simples, indicadas respectivamente nas Equações (32) e (33), a seguir:

$$\mu_R = \langle R \rangle = \sum_{j=1}^{N_p} r_j \langle \mathbf{L}_j(\xi) \rangle = \sum_{j=1}^{N_p} r_j w_j \quad (32)$$

$$\sigma_R^2 = \langle R^2 \rangle - \mu_R^2 \cong \sum_{j=1}^{N_p} \sum_{k=1}^{N_p} r_j r_k \langle \mathbf{L}_j(\xi) \mathbf{L}_k(\xi) \rangle - \mu_R^2 = \sum_{j=1}^{N_p} r_j^2 w_j - \mu_R^2 \quad (33)$$

onde a expectativa de um polinômio Lagrangeano em particular construído em pontos de Gauss e integrado nesses mesmos pontos deixa apenas o peso correspondente ao ponto para o qual o valor de interpolação é 1.

5.2 Análise de Sensibilidade da Variável Aleatória

As expansões do método da Colocação Estocástica são facilmente diferenciáveis em relação às variáveis aleatórias [43].

Primeiramente, usando-se a Equação (31) tem-se a Equação (34):

$$\frac{dR}{d\xi_i} = \sum_{j=1}^{N_p} r_j \frac{d\mathbf{L}_j(\xi)}{d\xi_i} \quad (34)$$

onde o polinômio interpolador multidimensional L_j é formado sobre pontos de quadratura do produto tensorial ou sobre uma malha esparsa.

Para o primeiro caso, a derivada do polinômio interpolador multidimensional L_j envolve uma regra de produto de polinômios interpoladores de uma dimensão L_k , conforme a Equação (35) [43]:

$$\frac{dL_j(\xi)}{d\xi_i} = \frac{dL_i}{d\xi_i} \prod_{\substack{k=1 \\ k \neq i}}^n L_k(\xi_k) \quad (35)$$

Para o segundo caso, a derivada do polinômio interpolador multidimensional L_j envolve uma combinação linear dessas regras de produtos.

Finalmente, de forma similar ao apresentado na seção 4.5 para o método do Caos Polinomial, usa-se o Jacobiano da transformação de variável de Nataf [27], conforme a Equação (29) onde, para o caso de x_i não correlacionada, essa equação é simplificada para a forma da Equação (30).

6 Redução do Espaço Probabilístico

Como alternativa para a melhoria do desempenho do método do Caos Polinomial e o da Colocação Estocástica pode ser acoplada a eles a decomposição de Karhunen-Loève ([44], [45]). A decomposição de Karhunen-Loève é particularmente útil para os métodos estocásticos baseados em expansões, pois tem por objetivo a redução do espaço probabilístico das entradas estocásticas de ordens mais altas. Com um número reduzido de variáveis estocásticas descrevendo o campo de entrada - produzidas previamente pelo processo de decomposição de Karhunen-Loève - torna-se viável a aplicação de métodos avançados de propagação de incertezas, como é o caso dos métodos do Caos Polinomial (PCE – *Polynomial Chaos Expansion*) e da Colocação Estocástica (SC – *Stochastic Collocation*).

No entanto, técnicas convencionais para o cálculo da Expansão de Karhunen-Loève (KLE) podem ser computacionalmente caras quando aplicadas a modelos de reservatórios de larga escala, uma vez que seria requerida, no mínimo, a decomposição do valor singular de uma matriz muito grande. Para a minimização desse problema foi desenvolvida a Expansão *Kernel* de Karhunen-Loève (KKLE), na qual a Expansão de Karhunen-Loève (KLE) é determinada, sem o cálculo explícito da matriz de covariância, ambas apresentadas nas seções seguintes.

6.1 Expansão de Karhunen-Loève

A Expansão de Karhunen-Loève (KLE) é um método de representação de um processo aleatório baseado na expansão espectral da função de correlação do processo e é particularmente útil para a Expansão em Caos Polinomial generalizada e para a Colocação Estocástica, uma vez que provê meios de reduzir a dimensionalidade no espaço aleatório, permitindo o uso dessas expansões como método para a propagação de incertezas em problemas de ordens mais altas.

A Expansão de Karhunen-Loève representa o processo aleatório na forma da Equação (36):

$$R(\mathbf{x}, w) = \bar{R}(\mathbf{x}) + \sum_{i=1}^{\infty} \xi_i(w) \sqrt{\lambda_i} \mathbf{v}_i(\mathbf{x}) \quad (36)$$

onde $R(\mathbf{x}, w)$ é um campo estocástico estacionário, $\bar{R}(\mathbf{x})$ é o valor médio do campo, $\xi_i(w)$ forma um conjunto de variáveis aleatórias não correlacionadas, com média 0 e variância 1, e $\mathbf{v}_i(x)$ e $\lambda_i(x)$ são as autofunções e autovalores da função de correlação, respectivamente, conforme descreve a Equação (37):

$$\int C_{RR}(x, y) \mathbf{v}_i(y) dy = \lambda_i \mathbf{v}_i(x) \quad (37)$$

onde x e y são as coordenadas temporais ou espaciais.

A série mostrada na Equação (36) converge para campos multigaussianos, preservando a covariância da matriz de covariância original. A expansão mostrada na Equação (36) é infinita, mas pode ser usada na prática visto que, em muitos casos, esta série pode ser truncada em um número pequeno de termos n sem efeito deletério. O número de termos n é determinado pela deterioração dos autovalores da Equação (37), para garantir que o erro de truncamento seja suficientemente pequeno para ser aceitável. Isto é possível nos casos em que a magnitude dos autovalores da matriz de covariância cai muito rapidamente, situação característica da existência de grandes comprimentos de correlação no campo estocástico expandido. Quando o campo tende a ter comprimentos de correlação muito curtos e, conseqüentemente, tende a ser um campo de ruído branco espacial, as magnitudes dos autovalores decaem muito lentamente e a Expansão de Karhunen-Loève se torna ineficiente por necessitar de um número muito grande de termos. Vale ressaltar que aleatoriedade do processo está concentrada nas variáveis $\xi_i(w)$, pois tanto $\mathbf{v}_i(x)$ quanto $\lambda_i(x)$ são determinísticos. Dessa forma, a Expansão de Karhunen-Loève expressa o campo estocástico como uma série de autovalores e autovetores da matriz de covariância, multiplicado pelas variáveis aleatórias normais padrão (SNRV – *Standard Normal Random Variables*), fornecendo uma forma

efetiva para representar os processos aleatórios de entrada com função de correlação conhecida.

A Expansão de Karhunen-Loève pode ser considerada uma excelente expansão para campos aleatórios multigaussianos e, mesmo quando o campo não é multigaussiano, minimiza o erro do quadrado da média da representação do termo finito.

Em aplicações de reservatórios de petróleo, uma vez que a matriz de covariância é constante e determinística, como são seus autovalores ($\lambda_i(x)$) e autovetores ($v_i(x)$), a aleatoriedade do campo de entrada para a simulação do reservatório é dada apenas pelo conjunto de SNRV's denominado na Equação (36) por $\xi_i(w)$. A Expansão de Karhunen-Loève preserva estatísticas de dois pontos do campo de entrada, que servem para testar se as funções de distribuições cumulativas para dois, três ou mais pontos seguem uma distribuição Gaussiana também, de forma a ratificar a adoção do modelo multigaussiano.

A Expansão de Karhunen-Loève pode ser vista em termos abrangentes, como uma analogia de uma decomposição modal de um problema de vibração estrutural, onde os autovetores da matriz de covariância correspondem aos modos de vibração, e seus autovalores às frequências naturais de vibração. Um resultado importante na pesquisa da Expansão de Karhunen-Loève é que as características estocásticas mais importantes do campo são bem representadas ainda que a série seja truncada num pequeno número dos primeiros termos (número pequeno dependendo do problema). Isto significa que, um grande campo estocástico - a exemplo do campo de permeabilidade de um reservatório, que possui dezenas ou centenas de milhões de células - pode ser aproximado com precisão aceitável por dezenas ou centenas de variáveis apenas, nos casos em que as magnitudes dos autovalores da matriz de covariância decaem muito rapidamente.

6.1.1 Detalhamento da Expansão de Karhunen-Loève (KLE)

Supondo que há um conjunto de realizações de um campo aleatório com média zero, dado por R_j , $j = 1, \dots, N_R$, onde N_R é o número de realizações e cada vetor tem N_C elementos correspondendo, no caso de simulações de reservatório,

ao valor da propriedade de interesse na célula do reservatório. A matriz de covariância do campo de entrada pode ser calculada conforme Equação (38):

$$\mathbf{C} = \frac{1}{N_R} \sum_{j=1}^{N_R} \mathbf{R}_j \mathbf{R}_j^t \quad (38)$$

onde o número de realizações deve ser suficientemente grande para que a matriz de covariância seja bem aproximada.

Nos problemas de simulações de reservatório, as realizações serão normalmente criadas usando-se um sistema de software que implemente algoritmos geostatísticos. A Expansão de Karhunen-Loève calcula novas realizações do campo \mathbf{R} em termos de autovalores e autovetores da matriz de covariância, de acordo com a Equação (39):

$$\mathbf{R} = \mathbf{E}\mathbf{\Lambda}^{1/2}\boldsymbol{\xi} \quad (39)$$

onde \mathbf{E} é a matriz de autovetores da matriz de covariância, $\mathbf{\Lambda}$ é a matriz diagonal dos autovalores de \mathbf{C} e $\boldsymbol{\xi}$ é um vetor de variáveis normais padrão não correlacionadas, que são independentes se o campo original é multigaussiano.

A realização \mathbf{R} obtida com a Equação (39) será multigaussiana e, principalmente, honrará as restrições geológicas do campo de entrada (até o segundo momento), pois terá a mesma matriz de covariância \mathbf{C} . É importante, também, notar que as matrizes \mathbf{C} , \mathbf{E} e $\mathbf{\Lambda}$ tem dimensão $N_C \times N_C$ e o vetor $\boldsymbol{\xi}$, dimensão $N_C \times 1$. Como essas variáveis aleatórias não são correlacionadas, elas podem ser variadas de qualquer maneira, em particular em algoritmos determinísticos e sob incerteza, para propagação de incertezas e otimização. Tipicamente, para a propagação de incertezas, essas variáveis poderiam ser amostradas tanto numa técnica como Monte Carlo, o que poderá ser extremamente caro se as simulações levarem muito tempo, ou usadas como variáveis de base para uma Expansão em Caos Polinomial, o que poderia ser muito mais econômico.

Na prática atual, o número de células num modelo de reservatório é muito grande e conseqüentemente o tamanho das matrizes mencionadas também, tornando o cálculo dos seus autovalores e autovetores impraticável. Uma primeira

alternativa, para a solução desse problema, é simplesmente truncar as séries e não calcular todos os autopares da matriz de covariância. Essa abordagem pode proporcionar resultados aceitáveis, devido à maioria da energia do campo estocástico estar associada aos maiores autovalores. Entretanto, essa operação ainda pode ser muito cara para um campo muito grande, podendo ser necessárias para isso técnicas mais avançadas.

O custo computacional excessivo e o fato de ele expressar apenas estatísticas de dois pontos dos campos de entrada, fazendo com que estruturas geológicas complexas, tais como canais, não sejam reproduzidas pela expansão, tornam a Expansão de Karhunen-Loève não vantajosa para o uso em todas as situações. Para a solução desses problemas deverá ser utilizada a versão *kernel* (núcleo) dessa expansão, detalhada na seção 6.2.

6.2 Expansão *Kernel* de Karhunen-Loève (KKLE)

A Expansão *Kernel* de Karhunen-Loève (KKLE) é um algoritmo que calcula a KLE de forma mais eficiente. Além disso, possui outras propriedades atrativas como a capacidade de reproduzir momentos estatísticos de ordens mais altas, com a escolha das funções *kernel* apropriadas.

A KKLE é baseada no cálculo dos autovalores e autovetores do núcleo (*kernel*) da matriz de covariância, que é uma matriz quadrada de ordem igual ao número de realizações, que é normalmente muito menor do que o número de células. Cada elemento do núcleo da matriz de covariância é uma função do produto interno de realizações do campo de entrada aleatório, da função do *kernel*. A KKLE é feita num espaço característico e , quando estatísticas multipontos devem ser preservadas na expansão, a transformação deve ser não linear. O cálculo dos autopares da matriz de covariância original se torna um tanto mais elaborado, requerendo a transformação das realizações do espaço característico transformado no espaço original das realizações. Isto pode ser aproximadamente resolvido como um problema de minimização sem qualquer dificuldade mais séria. A KKLE é, conseqüentemente, a ferramenta para calcular uma expansão diferenciável do campo de entrada, de uma maneira mais eficiente computacionalmente, preservando momentos estatísticos do campo de entrada como desejado.

6.2.1 Detalhamento da Expansão *Kernel* de Karhunen-Loève (KKLE)

Cada autovetor e autovalor da matriz de covariância, descrita pela Equação (38), é dado pela solução de $\lambda \mathbf{v} = \mathbf{Cv}$, onde λ é um autovalor e \mathbf{v} é um autovetor da matriz de covariância. Multiplicando-se ambos os lados da igualdade na Equação (38) pelo autovetor \mathbf{v} resulta na Equação (40):

$$\mathbf{Cv} = \frac{1}{N_R} \sum_{j=1}^{N_R} (\mathbf{R}_j \cdot \mathbf{v}) \mathbf{R}_j \quad (40)$$

o que mostra que todos os autovetores devem estar no subespaço definido pelas realizações \mathbf{R}_j e, conseqüentemente, apenas pode haver N_R autovetores independentes (e autovetores diferentes de zero). Além disso, é possível reescrever a equação de autopares na forma da Equação (41):

$$\lambda (\mathbf{R}_j \cdot \mathbf{v}) = (\mathbf{R}_j \cdot \mathbf{Cv}) \quad j = 1, \dots, N_R \quad (41)$$

Para implementar a Expansão *Kernel* de Karhunen-Loève, uma transformação não linear de \mathfrak{R}^{N_C} para um espaço característico arbitrário F , com membros \mathbf{Y} , é definida por $\Phi: \mathfrak{R}^{N_C} \rightarrow F$, onde N_C é o número de células no modelo. O espaço característico pode ser arbitrariamente grande, e possivelmente ter dimensão infinita, podendo-se supor, a partir de agora, que os vetores $\mathbf{Y} = \Phi(\mathbf{R})$ são centralizados (isto é, possuem média zero). A covariância pode ser calculada no espaço transformado conforme a Equação (42):

$$\bar{\mathbf{C}} = \frac{1}{N_R} \sum_{j=1}^{N_R} \Phi(\mathbf{R}_j) \Phi(\mathbf{R}_j)^t \quad (42)$$

É possível calcular os autovalores λ e os autovetores \mathbf{V} dessa matriz resolvendo o problema, de acordo com a Equação (43):

$$\lambda \mathbf{V} = \bar{\mathbf{C}} \mathbf{V} \quad (43)$$

As dimensões da matriz $\bar{\mathbf{C}}$ podem ter ordens de magnitude maiores que aquelas da matriz de covariância original, devido à transformação não linear, logo a Equação (43) não pode ser resolvida com um algoritmo convencional. É possível concluir, pelos mesmos argumentos acima, que as soluções da Equação (43) caem em um subespaço definido por $\Phi(\mathbf{R}_j), \dots, \Phi(\mathbf{R}_{N_R})$; desta forma, uma equação análoga a Equação (41) pode ser escrita sob a forma da Equação (44):

$$\lambda(\Phi(\mathbf{R}_j) \cdot \mathbf{V}) = (\Phi(\mathbf{R}_j) \cdot \bar{\mathbf{C}}\mathbf{V}) \quad j = 1, \dots, N_R \quad (44)$$

Como todos os autovetores no espaço característico podem ser escritos como combinações lineares das realizações transformadas, há coeficientes reais \mathbf{a}_j que podem ser descritos na Equação (45):

$$\mathbf{V} = \sum_{j=1}^{N_R} \mathbf{a}_j \Phi(\mathbf{R}_j) \quad (45)$$

Substituindo-se os valores de $\bar{\mathbf{C}}$ da Equação (42) e de \mathbf{V} da Equação (45) na Equação (44) obtém-se a Equação (46):

$$\lambda \left(\Phi(\mathbf{R}_j) \cdot \sum_{j=1}^{N_R} \mathbf{a}_j \Phi(\mathbf{R}_j) \right) = \left(\Phi(\mathbf{R}_j) \cdot \left(\frac{1}{N_R} \sum_{j=1}^{N_R} \Phi(\mathbf{R}_j) \Phi(\mathbf{R}_j)^t \right) \left(\sum_{j=1}^{N_R} \mathbf{a}_j \Phi(\mathbf{R}_j) \right) \right) \quad (46)$$

A Equação (46) pode então ser escrita conforme Equação (47):

$$\lambda \sum_{j=1}^{N_R} \mathbf{a}_j (\Phi(\mathbf{R}_j) \cdot \Phi(\mathbf{R}_j)) = \frac{1}{N_R} \sum_{j=1}^{N_R} \mathbf{a}_j \left(\Phi(\mathbf{R}_j) \cdot \sum_{j=1}^{N_R} \Phi(\mathbf{R}_j) \right) (\Phi(\mathbf{R}_j) \cdot \Phi(\mathbf{R}_j)) \quad (47)$$

definindo-se a matriz *kernel* \mathbf{K} , de dimensão $N_R \times N_R$, tal que cada elemento da matriz seja calculado pela Equação (48):

$$K(\Phi(\mathbf{R}_i), \Phi(\mathbf{R}_j))_{ij} = (\Phi(\mathbf{R}_i) \cdot \Phi(\mathbf{R}_j)) \quad (48)$$

A Equação (47) pode ser reescrita conforme a Equação (49):

$$N_R \lambda \mathbf{K} \mathbf{a} = \mathbf{K}^2 \mathbf{a} \quad (49)$$

onde \mathbf{a} é um vetor coluna com entradas a_j . As soluções da Equação (48) são dadas pelas soluções do problema do autovalor da matriz do *kernel* dada pela Equação (50):

$$N_R \lambda \mathbf{a} = \mathbf{K} \mathbf{a} \quad (50)$$

O problema dado pela Equação (50) pode ser resolvido muito mais rapidamente que o problema original do autovalor, uma vez que as dimensões da matriz do *kernel* são $N_R \times N_R$ ao invés de $N_C \times N_C$, e para problemas reais $N_R \ll N_C$. Por outro lado, o cálculo de realizações transformadas $\mathbf{Y} = \Phi(\mathbf{R})$ pode ser impossível se suas dimensões fossem muito grandes ou ainda infinitas, dependendo da transformação escolhida. Pode ser mostrado, entretanto, que não é necessário calcular essa transformação explicitamente, ou seja, é possível definir uma função *kernel* $k(\mathbf{R}_i, \mathbf{R}_j)$ que implementa o produto interno das realizações diretamente no espaço característico das realizações em \mathfrak{R}^{N_c} , conforme mostrado na Equação (51):

$$\langle \Phi(\mathbf{R}_i) \cdot \Phi(\mathbf{R}_j) \rangle = k(\mathbf{R}_i, \mathbf{R}_j) \quad (51)$$

sendo \mathbf{R}_i e \mathbf{R}_j o conjunto de realizações, cada um com N_c elementos.

Muitas funções *kernel* estão em uso em aplicações diferentes, mas nas simulações de reservatório bons resultados têm sido obtidos com as funções *kernel* polinomiais da forma:

$$k(\mathbf{R}_i, \mathbf{R}_j) = (\mathbf{R}_i \cdot \mathbf{R}_j)^d \quad (52)$$

onde d é um número inteiro pequeno. Quando $d = 1$, a função *kernel* é linear e uma simplificação importante da KKLE é obtida, mas apenas estatísticas de dois pontos do campo de entrada são preservados na expansão. Neste caso, a função *kernel* é simplesmente o produto interno de duas realizações originais, a transformação Φ se

torna a transformação identidade, o espaço característico é o próprio \mathfrak{R}^{N_c} , a matriz de covariância no espaço característico é a mesma matriz de covariância original e naturalmente, assim são seus autovetores. Consequentemente, os autovetores da matriz de covariância podem ser facilmente calculados com a Equação (45), e seus autovalores são simplesmente os autovalores da matriz *kernel* divididos por N_R (das Equações (43) e (50)). Neste trabalho foi concluído apenas este caso.

Quando $d > 1$, a função *kernel* é não linear e as estatísticas de mais alta ordem são preservadas, sob um custo adicional da complexidade da formulação. Torna-se necessário resolver um *problema pré-imagem*, isto é, já que a expansão será feita num espaço característico (que não é o \mathfrak{R}^{N_c} ,) será necessário então encontrar a realização que faz o seu mapeamento, resolvendo a equação $y - \Phi^{-1}(\mathbf{Y})$. Esta equação pode não ter apenas uma solução em \mathfrak{R}^{N_c} , assim como a função Φ pode não ter inversa e, na maioria dos casos, Φ não ser explicitamente conhecida. Este problema é resolvido de uma maneira aproximada com técnicas de minimização [47]. Além disso, para garantir que os vetores \mathbf{Y} sejam centralizados no espaço característico demanda-se uma manipulação algébrica extra. Os detalhes dessas duas técnicas não serão explorados neste trabalho.

A Expansão *Kernel* de Karhunen-Loève é de fácil implementação, pode ser calculada de forma eficiente e, como mencionado anteriormente, provê uma parametrização diferenciável do campo da entrada estocástica, em termos de poucas variáveis aleatórias normais padrão. Esta parametrização preserva estatísticas de dois pontos do campo de entrada e pode ser usada em algoritmos de propagação de incertezas e para a otimização sob incerteza, sendo uma das técnicas matemáticas mais robustas para a descrição das características estocásticas das propriedades petrofísicas de um reservatório.

Como forma de exemplificação, a Expansão *Kernel* de Karhunen-Loève será aplicada em um problema-modelo na seção 8.3.

7 Caracterização do Sistema

O sistema computacional desenvolvido para o estudo da propagação de incertezas na área de exploração e produção de petróleo faz uso de recursos tecnológicos, classificados em dois grupos distintos de acordo com a sua função, os quais são apresentados a seguir, nas seções 7.1 e 7.2.

7.1 Ambiente Computacional

Neste grupo encontram-se os elementos necessários ao desenvolvimento e execução do sistema computacional, os quais se encontram relacionados abaixo:

- a) Cluster tipo Beowulf para processamento paralelo;
- b) Sistema Operacional Linux, com bibliotecas para *Message Passing Interface* (MPI);
- c) Octave GNU [13].

Mesmo com a utilização de técnicas de redução, tanto para os modelos computacionais como para a dimensão do espaço probabilístico, as simulações, em geral, esbarram nos limites impostos pelo equipamento computacional disponível e, dessa forma, é recomendado o uso da computação paralela com clusters. Isto posto, para atender à demanda de memória e velocidade de processamento na manipulação dos dados durante o processo de construção dos sistemas de equações e na execução das iterações necessárias às simulações, foi usada, em várias etapas deste trabalho, a computação paralela com clusters.

O Octave GNU é uma linguagem de alto nível, que foi utilizada para o desenvolvimento de programas com a finalidade de permitir a troca de dados entre o programa simulador de reservatórios e o programa contendo os métodos de propagação de incertezas. O Octave GNU foi utilizado por ser compatível com o MATLAB e ser um software livre [56].

7.2 Programas

Neste grupo encontram-se os programas que foram usados para o estudo da propagação de incertezas, os quais se encontram relacionados abaixo:

- a) IMEX - Programa computacional para a simulação de reservatórios de petróleo;
- b) DAKOTA - Programa contendo os métodos de análise para a propagação de incertezas;
- c) Programas adicionais – *scripts* desenvolvidos no Octave GNU, para a filtragem dos dados relevantes utilizados pelo IMEX, quando esses dados são resultantes do DAKOTA, e vice-versa.

A estrutura do sistema computacional desenvolvido neste trabalho pode ser visualizada na Figura 3.

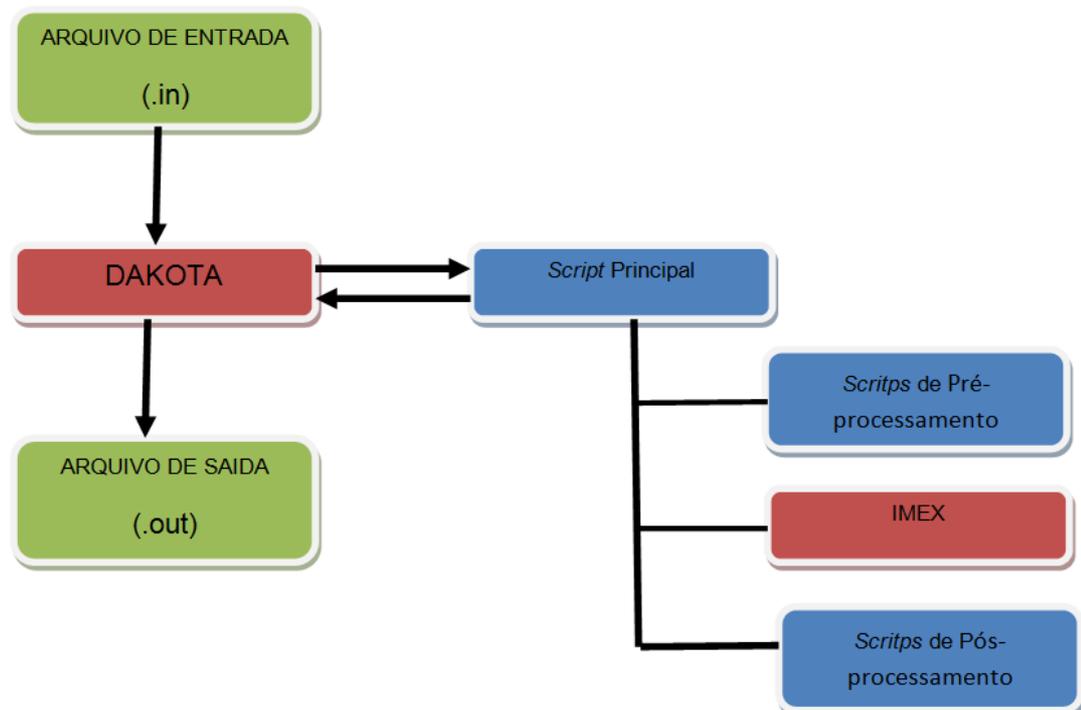


Figura 3. Esquema representativo do sistema proposto.

7.2.1 IMEX

De forma a manter a compatibilidade com o programa em uso na Petrobrás, empresa estatal responsável pela exploração de petróleo no Brasil, o IMEX foi o programa adotado neste projeto como o simulador numérico, com o papel de um protótipo virtual de um reservatório de petróleo.

O IMEX é um simulador numérico de escoamento em meios porosos, desenvolvido pela corporação CMG (*Computing Modelling Group*) como parte de um pacote de aplicações para simulação de reservatórios. Ele é um simulador “*black-oil*”, ou seja, para reservatórios contendo água, gás e óleo, que leva em consideração a gravidade e efeitos capilares, e suporta configurações de malha cartesiana, cilíndrica e com profundidade e espessura variável, possibilitando assim o mapeamento das configurações de campos em duas e três dimensões [7].

A execução do simulador IMEX pode ocorrer em três modos: explícito, implícito ou adaptativo. Estes modos referem-se à forma como são resolvidas as equações não lineares dos modelos de fluxo. No caso do modo explícito, o passo de tempo utilizado deve ser sempre menor que o máximo valor do passo que mantenha a estabilidade do sistema. No modo implícito, é necessário resolver um sistema de equações não lineares a cada passo, podendo ser escolhidos passos de tempo maiores. O modo explícito é vantajoso quando as variações dos estados são lentas, enquanto que o modo implícito é necessário quando as variações nos estados são muito rápidas. O modo adaptativo permite escolher entre a solução implícita ou explícita segundo a característica das variações dos estados do sistema [7].

Geralmente, só um pequeno número de blocos necessita ser tratado de forma totalmente implícita, pois, na maioria deles, as propriedades variam lentamente. Este procedimento é bastante proveitoso para problemas onde só existem grandes fluxos nas proximidades dos poços (efeito cone) ou em reservatórios estratificados com camadas bem finas. Utilizando o modo adaptativo, o tratamento de forma totalmente implícita é feito apenas onde for necessário, obtendo-se com isto uma grande economia no esforço computacional requerido [7].

O IMEX suporta também as opções de dupla porosidade e dupla permeabilidade, as quais permitem o tratamento de reservatórios com fraturas, onde estas são modeladas como duas partes: com porosidade primária e secundária. Já pela dupla

permeabilidade é possível modelar a comunicação de fluidos entre as duas partes definidas na fratura [7].

O simulador IMEX usa o GMRES (*Generalized Minimum Residual Method*) [42] como solucionador de sistemas de equações lineares pré-condicionadas com fatoração de Gauss incompleta. O GMRES é um método iterativo que permite minimizar a norma do resíduo baseado na geração dos espaços de Krylov. O pré-condicionamento dos dados juntamente com o método GMRES permitem diminuir o custo computacional e os requerimentos de memória quando o número de equações a resolver é muito grande [7].

Neste simulador a comunicação com o usuário pode ser efetuada através de arquivos de entrada e saída. O usuário deve criar um arquivo texto de entrada (formato .DAT) que contém: os parâmetros da malha, informação geológica, tipo de fluxo de óleo, condições iniciais, métodos numéricos, lista de poços, entre outras informações. O simulador, após a execução, fornece os valores de saída em um grupo de arquivos. Esse fluxo de informações é apresentado na Figura 4 [7].

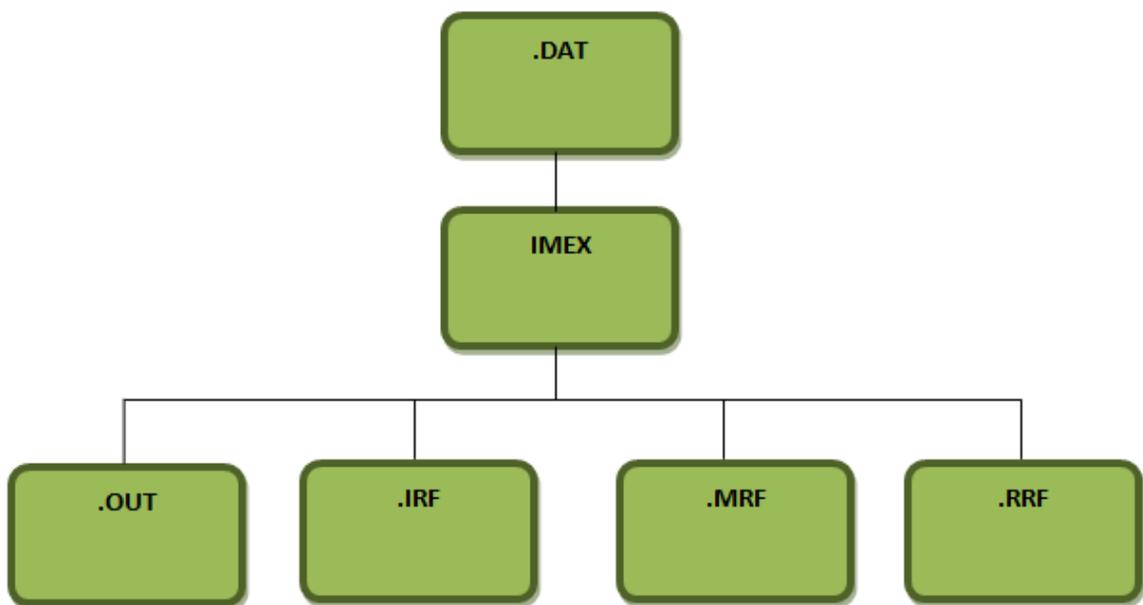


Figura 4. Fluxo de informações no IMEX.

O arquivo de extensão .DAT é imprescindível, uma vez que este arquivo contém todas as informações que definem a configuração completa do reservatório a ser simulado, conforme os tipos de dados apresentados a seguir:

- a) Dados de controle de Entrada e Saída
- b) Dados de descrição do reservatório
- c) Propriedades dos componentes
- d) Dados de Rochas e Fluidos
- e) Condições Iniciais
- f) Informação sobre os métodos numéricos
- g) Informação dos poços e dados recorrentes

Os arquivos de extensão .OUT e .IRF contêm os resultados globais da simulação, isto é, valores de saturações de óleo, gás, água, etc., em cada célula. O arquivo .IRF contém o índice dos resultados e o arquivo .OUT armazena os resultados numéricos em si [7].

Os arquivos de extensão .MRF contêm os resultados principais da simulação e os .RRF, que são opcionais, os resultados recorrentes [7].

Todos os resultados da simulação são armazenados em arquivos de extensões .OUT e .IRF, mas para a obtenção da curva de produção do campo a maioria destes resultados é desnecessária. Desta forma o aplicativo “*results report*” (CMG 2000:2) permite a escolha apenas dos dados necessários, fazendo o papel de um gerador de relatórios [7]. O “*results report*” (CMG 2000:2) não foi utilizado neste trabalho e, para a extração dos dados necessários a determinação do Valor Presente líquido, foram desenvolvidos programas adicionais (*scripts*) com esse fim.

7.2.2 DAKOTA

No sistema desenvolvido, os métodos destinados a realizar a propagação de incertezas, aplicada as simulações de reservatórios de petróleo, foram os implementados no DAKOTA. O DAKOTA, desenvolvido pela *Sandia National Laboratories* [43], foi escolhido pelo fato de ser um software livre e pela facilidade de implementar interfaces com métodos de análise iterativos diversos, e principalmente, métodos não determinísticos de propagação de incertezas. O DAKOTA é um “*framework*” orientado a objeto, multi-paralelo, que tem como funcionalidades principais a otimização de projetos, estimativa de parâmetros, análise de sensibilidade e propagação de incertezas.

Os objetivos de desempenho de um sistema podem ser formulados, por exemplo, para minimizar o peso, o custo, ou os defeitos; para limitar uma temperatura, tensão ou resposta crítica a vibração; ou para maximizar o desempenho, a confiabilidade, a saída, a agilidade ou a robustez do projeto. Sendo assim, o DAKOTA foi desenvolvido para fornecer meios sistemáticos e rápidos para a obtenção de projetos melhorados ou ótimos, com o uso de seus modelos baseados em simuladores [43].

As ferramentas do DAKOTA fornecem uma interface flexível e extensa entre um código de simulação e uma variedade de estratégias e métodos iterativos. O DAKOTA foi concebido originalmente como uma interface de uso fácil entre códigos de simulação e algoritmos de otimização. No entanto, suas versões mais recentes foram expandidas para implementar interfaces com outros tipos de métodos de análise iterativos, tais como propagação de incertezas com métodos não determinísticos de propagação, estimativa de parâmetros com o método dos mínimos quadrados, e análise de sensibilidade/variância, para projeto de experimentos e estudos paramétricos. Estas potencialidades podem ser usadas em seus próprios blocos ou em blocos construídos com estratégias mais sofisticadas, tais como otimização híbrida, otimização baseada em modelos substitutos, programação linear e não linear, ou otimização sob incerteza. Assim, uma das vantagens que o DAKOTA oferece é o acesso a uma escala muito abrangente de potencialidades iterativas, que podem ser obtidas através de uma única interface relativamente simples entre o DAKOTA e o simulador em uso. Para a utilização de um tipo diferente de método ou de estratégia iterativa com o simulador em uso, faz-se necessário apenas a mudança de alguns comandos de entrada no DAKOTA e dar início a uma nova análise [43].

No DAKOTA, as Expansões em Caos Polinomial proveem estatísticas (momentos) de resposta analítica, mas as probabilidades CDF e CCDF são avaliadas na expansão por amostragem. Após o cálculo dos coeficientes da expansão, algumas estatísticas já ficam disponíveis analiticamente e outras devem ser avaliadas numericamente. Para a avaliação numérica, deve ser fornecida uma quantidade de amostras que serão avaliadas na expansão, para o cálculo das probabilidades CDF. As estatísticas analíticas para a média, desvio padrão e covariância são calculadas conforme descrito na seção 4.4 e da mesma forma, as sensibilidades das variáveis

aleatórias analíticas são calculadas conforme descrito na seção 4.5, avaliadas nos valores da média [43].

Sendo a propagação de incertezas a potencialidade do DAKOTA de interesse neste trabalho, serão apresentados nas seções 7.2.2.1 e 7.2.2.2, respectivamente, os tipos de variáveis contínuas e os métodos para a propagação de incertezas atualmente disponíveis no DAKOTA. Da mesma forma, será feita, na seção 7.2.2.3, uma breve apresentação do processo implementado no DAKOTA, para transformação das variáveis aleatórias originais para variáveis aleatórias independentes padrão, forma exigida pelos métodos de propagação de incertezas utilizados.

7.2.2.1 Representação das Variáveis Aleatórias no DAKOTA

As variáveis incertas contínuas são representadas por uma distribuição de probabilidade, com seus parâmetros estatísticos. As distribuições de probabilidade que estão disponíveis no DAKOTA para a representação das variáveis incertas são apresentadas a seguir, com suas respectivas características [43]:

- **Normal** - Também conhecida como Gaussiana. Caracterizada por uma média e desvio padrão.
- **Normal Limitada** - Similar à normal, mas com especificação adicional para o limite superior e inferior.
- **Lognormal** - Caracterizada por uma média, um desvio padrão ou um fator de erro. O logaritmo natural de uma variável lognormal tem uma distribuição normal.
- **Lognormal limitada** - Similar à lognormal, mas com uma especificação adicional para o limite superior e inferior.
- **Uniforme** - Caracterizada pelos limites superior e inferior e probabilidade constante entre os limites.

- **Loguniforme** - Caracterizada pelos limites superior e inferior. O logaritmo natural de uma variável loguniforme tem uma distribuição uniforme.
- **Triangular** - Caracterizada por uma moda e limites superior e inferior.
- **Beta** – Flexível. Caracterizada por limites inferior e superior e parâmetros alfa e beta.
- **Gamma** – Flexível. Caracterizada por limites inferior e superior e parâmetros alfa e beta. A distribuição exponencial é um caso especial.
- **Gumbel** - Distribuição de probabilidade de valor extremo maior do tipo I, ou seja, a função de distribuição cumulativa é dada por $F(x; \mu, \sigma, 0) = e^{-e^{-(x-\mu)/\sigma}}$ para $x \in \mathfrak{R}$. É caracterizada pelos parâmetros alfa e beta.
- **Frechet** - Distribuição de probabilidade de valor extremo maior do tipo II, ou seja, se $\zeta = \alpha^{-1}$ com $\alpha > 0$, a função de distribuição cumulativa é dada por

$$F(x; \mu, \sigma, \zeta) = \begin{cases} 0 & x \leq \mu \\ e^{-\left(\frac{x-\mu}{\sigma}\right)^{-\alpha}} & x > \mu \end{cases} . \text{ É caracterizada pelos parâmetros alfa e beta.}$$
- **Weibull** - Distribuição de probabilidade do menor extremo valor do tipo III, ou seja, se $\zeta = -\alpha^{-1}$ com $\alpha > 0$, a função de distribuição cumulativa é dada por

$$F(x; \mu, \sigma, \zeta) = \begin{cases} e^{-\left(\frac{-(x-\mu)}{\sigma}\right)^{\alpha}} & x < \mu \\ 1 & x \geq \mu \end{cases} . \text{ É caracterizada pelos parâmetros alfa e beta.}$$
- **Histograma** - Uma distribuição de probabilidade determinada empiricamente. Caracterizada por um conjunto de pares (x, y).

7.2.2.2 Métodos de Propagação de Incertezas no DAKOTA

Dentro da abordagem que utiliza amostragem direta, o DAKOTA dispõe dos métodos de Monte Carlo (MC) e *Latin Hypercube Sampling* (LHS) para a propaga-

ção de incertezas [43]. Na abordagem que não utiliza amostragem direta, o DAKOTA dispõe dos métodos baseados em confiabilidade, não determinísticos epistêmicos e de expansão estocástica. Os métodos utilizados neste trabalho para o estudo da propagação de incertezas - o método do Caos Polinomial e o da Colocação Estocástica - classificam-se como métodos de expansão estocástica [43].

- **Métodos Baseados em Confiabilidade**

Os métodos baseados em confiabilidade oferecem uma abordagem alternativa para a propagação de incertezas, uma vez que podem demandar menos esforço computacional do que as técnicas baseadas em amostragem [43]. Eles são baseados em abordagens probabilísticas, que calculam as estatísticas da distribuição da função resposta baseadas nas distribuições das variáveis incertas especificadas. Estas estatísticas incluem como saída: a média, o desvio padrão e as funções de distribuição cumulativas (CDF – *Cumulative Distribution Function*) ou funções de distribuição cumulativas complementares (CCDF – *Complementary Cumulative Distribution Function*) [43].

- **Métodos Não Determinísticos Epistêmicos**

Os métodos não determinísticos epistêmicos são utilizados quando não se pode especificar uma distribuição de probabilidade para uma variável, por não se ter o conhecimento suficiente sobre ela ou pela falta de informação sobre um aspecto particular do modelo de simulação, incluindo o sistema e o ambiente que está sendo modelado. Um exemplo de incerteza epistêmica é o caso de um parâmetro físico existente, mas desconhecido, onde se tem apenas pouco ou nenhum dado experimental sobre ele. O conhecimento limitado sobre o fenômeno físico complexo e a incerteza sobre a forma do modelo correto a ser usado, são outros exemplos de incerteza epistêmica. Muitas abordagens têm sido desenvolvidas para o modelo de incerteza epistêmica, dentre elas: a Teoria do Conjunto Vazio, a Teoria da Possibilidade e a Teoria da Evidência [43].

A Teoria da Evidência é a abordagem mais recente, sendo também conhecida como Teoria de Dempster-Shafer ou teoria dos conjuntos aleatórios ([35], [50]), e tem como objetivo a modelagem dos efeitos das incertezas epistêmicas. Nesta teoria

há duas medidas complementares de incerteza: Crença e Plausibilidade. Crença é o limite mais baixo de uma estimativa de probabilidade e deve ser consistente com a evidência. Da mesma forma a Plausibilidade é o limite mais alto numa estimativa de probabilidade e também deve ser consistente com a evidência.

Na Teoria da Evidência não é possível especificar um valor de probabilidade, e sim uma faixa de valores que deve ser consistente com a evidência. Essa faixa de valores é definida pela Crença e Plausibilidade. Nessa teoria, as variáveis de entrada incertas são modeladas como um conjunto de intervalos, os quais podem ser contínuos ou ter lacunas. Para cada intervalo é assinalada uma Sentença de Probabilidade Básica (BPA – *Basic Probability Assignment*), indicando como ela cai dentro do intervalo. O intervalo e suas BPA's associadas são então propagados através da simulação, para obtenção de funções de distribuição cumulativas, de Crença e Plausibilidade [43].

- **Método de Expansão Estocástica**

Dentro dessa abordagem se enquadram o método do Caos Polinomial e o da Colocação Estocástica. Nessa classe de métodos, a aproximação da relação desconhecida entre as entradas e saídas incertas é realizada por meio de uma expansão polinomial, que é uma série infinita em polinômios ortogonais, multidimensionais, de variáveis estocásticas. Um dos aspectos de grande importância desses métodos, e que os tornam ainda mais atrativos em relação aos outros, é o alcance da convergência com a realização de pouquíssimas avaliações de funções. Por esse e outros aspectos não menos importantes, os métodos do Caos Polinomial e da Colocação Estocástica se apresentam como excelentes alternativas para a propagação de incertezas, e foram selecionados para uso no sistema desenvolvido para o estudo da propagação de incertezas no processo de exploração e produção de petróleo. O DAKOTA implementa, por padrão, a abordagem do método do Caos Polinomial generalizado usando o esquema de Wiener-Askey. Esse padrão pode ser alterado para a Expansão em Caos Polinomial de Wiener-Hermite, apenas se for necessário suportar correlações prescritas [43].

No entanto, para problemas com dimensões de ordens mais altas, ou seja, com grande número de variáveis aleatórias, esses métodos só se tornam viáveis com a utilização, em conjunto, de técnicas para redução de espaço probabilístico.

7.2.2.3 Transformação de Variáveis

O Caos Polinomial e a Colocação Estocástica são expandidos usando polinômios que são funções de variáveis aleatórias independentes padrão ζ . O uso de variáveis aleatórias independentes padrão é o componente crítico que possibilita o desacoplamento das integrais multidimensionais numa expansão de base mista. Com o uso de variáveis independentes, as integrais multidimensionais, que estão envolvidas nos produtos internos dos polinômios multivariados, são simplificadas para um produto de integrais de uma dimensão, envolvendo apenas a base polinomial em particular e a função peso correspondente, selecionada para cada dimensão aleatória.

Os produtos internos multidimensionais são diferentes de zero apenas se cada um dos produtos internos de uma dimensão for diferente de zero, o que preserva as desejadas propriedades da ortogonalidade multivariadas para o caso de uma base mista [43]. Quando as variáveis aleatórias envolvidas não se encontram nesta forma, é necessário executar uma transformação das variáveis aleatórias originais x para variáveis aleatórias independentes padrão ζ e então aplicar a expansão estocástica no espaço transformado. Para simplificação, presume-se que ζ possui a mesma dimensão de x .

Para problemas envolvendo variáveis aleatórias de entrada com distribuição normal, uniforme, exponencial, beta e gama, independentes, a conversão para a forma padrão envolve uma transformação linear simples (para a forma das funções de densidade descritas na Tabela 1). Para outras distribuições independentes há duas alternativas diferentes para resolver esse problema. A primeira alternativa consiste na geração numérica de uma base polinomial para cada distribuição, juntamente com seus pesos e pontos de Gauss [43]. Essa alternativa preserva as taxas exponenciais de convergência para aplicações de quantificação de incerteza com entradas probabilísticas gerais, mas a execução deste processo para funções de densidade conjunta geral com correlação é um tópico de pesquisas em andamento. A segunda alternativa consiste na execução de uma transformação de variável não linear de uma dada distribuição de entrada - para uma base mais similar a uma do Esquema de Askey - e no emprego dos polinômios ortogonais de Askey e dos pontos/pesos de Gauss, associados. Como exemplo dessa alternativa, tem-se a

distribuição lognormal, para a qual se pode empregar uma base de Hermite num espaço normal padrão transformado. Essa alternativa pode ser efetiva [12], mas as taxas de convergência são tipicamente degradadas e os coeficientes de correlação podem ser distorcidos pela transformação [43].

Para distribuições não normais correlacionadas, a única opção aceitável, embora outras opções estejam numa área de pesquisa ativa, é executar uma transformação de variável não linear de todas as distribuições de entradas dadas, para distribuições normais padrão não correlacionadas e empregar estritamente bases polinomiais ortogonais de Hermite e pesos e pontos de Gauss, associados. No DAKOTA, a transformação não linear das variáveis aleatórias potencialmente correlacionadas para variáveis aleatórias não correlacionadas padrão envolve uma transformação de Nataf [27], a qual ocorre nos dois passos a seguir, conforme [43]:

Primeiro Passo: para transformar variáveis correlacionadas originais (x – espaço) para normais correlacionadas padrão (z – espaço), é usada a condição baseada na Função de Distribuição Cumulativa, apresentada na Equação (53) [43]:

$$\Phi(z_i) = F(x_i) \quad (53)$$

onde $F(\bullet)$ é a função de distribuição cumulativa da distribuição de probabilidade da variável correlacionada original x e $\Phi(\bullet)$ é a função de distribuição cumulativa da distribuição de probabilidade da variável correlacionada padrão z .

Segundo Passo: para transformar variáveis correlacionadas padrão \mathbf{z} (z – espaço) para não correlacionadas \mathbf{u} (u – espaço). é usado o fator de Cholesky \mathbf{L} de uma matriz de correlação modificada, isto é, a matriz de correlação original para não normais no espaço x é modificada para representar a correlação correspondente no espaço z . Essa transformação é dada pela Equação (54) [43]:

$$\mathbf{z} = \mathbf{L}\mathbf{u} \quad (54)$$

Pressupõe-se nesta abordagem que as variáveis aleatórias padrão não correlacionadas resultantes dessas transformações também sejam independentes [43].

A Tabela 3 apresenta os tipos de distribuições e correlações suportadas para as duas variantes da Expansão em Caos Polinomial implementadas no DAKOTA, a de Wiener-Hermite e a de Wiener-Askey [43].

Tabela 3. Tipos de distribuição suportados pela Expansão em Caos Polinomial.

Tipo de Distribuição	Expansão em Caos Polinomial de Wiener-Hermite	Expansão em Caos Polinomial de Wiener-Askey
NORMAL	C	C1, C2
LOGNORMAL	C	C2
UNIFORME	C	C1
LOGUNIFORME	U	U
TRIANGULAR	U	U
EXPONENCIAL	C	C1
BETA	U	C1
GAMMA	C	C1
GUMBEL	C	C2
FRECHET	C	C2
WEIBUL	C	C2
HISTOGRAMA	-	-
INTERVALO	-	-

As sinalizações apresentadas na Tabela 3 significam:

- “-“ - a distribuição de probabilidade não é suportada pelo método.
- U - as variáveis de entrada incertas para este tipo não devem ser correlacionadas
- C - as correlações são suportadas envolvendo variáveis de entrada incertas para este tipo.
- C# - as correlações são suportadas apenas entre variáveis deste tipo e outras variáveis do mesmo conjunto numerado.

A Tabela 3 indica, para o caso da Expansão em Caos Polinomial de Wiener-Hermite, que as variáveis com os tipos de distribuição sinalizados por “C” se correlacionam apenas com outras variáveis com o seu próprio tipo de distribuição. Já para o caso da Expansão em Caos Polinomial de Wiener-Askey, as variáveis com os

tipos de distribuição sinalizados por “C1” ou “C2”, ou por ambos, se correlacionam com outras variáveis com o seu próprio tipo de distribuição e com os tipos de distribuição sinalizados por “C1” ou “C2”. Isto quer dizer, por exemplo, que no caso de variáveis com o tipo de distribuição Normal, sinalizado na Tabela 3 por “C1” e “C2”, elas irão se correlacionar com outras variáveis com tipo de distribuição: Normal, Lognormal (“C2”), Uniforme (“C1”), Exponencial (“C1”), Beta (“C1”), Gamma (“C1”), Gumbel (“C2”), Frechet (“C2”), e Weibul (“C2”). Já os demais tipos de distribuição, sinalizados por “C1” ou por “C2” apenas irão se correlacionar com variáveis com o seu próprio tipo de distribuição e com variáveis com os tipos de distribuição que apresentam a mesma sinalização. Este caso pode ser exemplificado por variáveis com distribuição Lognormal (“C2”), as quais poderão se correlacionar com outras variáveis com os seguintes tipos de distribuição: Lognormal, Normal (“C1” e “C2”), Gumbel (“C2”), Frechet (“C2”), e Weibul (“C2”) [43].

7.2.3 Programas Adicionais

Para o tratamento das informações resultantes do IMEX e DAKOTA foram elaborados programas adicionais, aqui denominados *scripts*, que foram desenvolvidos utilizando-se o Octave. Esses *scripts* têm a finalidade principal de prover uma interface entre o DAKOTA e o IMEX, realizando um tratamento prévio das informações fornecidas pelo DAKOTA para manipulação pelo IMEX, e vice-versa. Além disso, foram desenvolvidos também programas com finalidades específicas, a exemplo de geração de informações sobre o Valor Presente Líquido (VPL), geração de relatórios estatísticos e outros.

Esses *scripts* são integrados ao sistema compondo a arquitetura visualizada anteriormente na Figura 3 da seção 7.2.

Os *scripts* foram agrupados da seguinte forma:

- *Scripts* de pré-processamento
- *Scripts* de pós-processamento

Na Tabela 4 podem ser visualizados os *scripts* de pré-processamento, juntamente com suas respectivas finalidades.

Tabela 4. *Scripts* do pré-processamento e suas finalidades.

Script	Finalidade
analysisDriver.m	Programa principal que chama todos os outros <i>scripts</i> , assim também como executa o programa IMEX.
readDakotaParam.m	Cria uma tabela de pares (nome da variável, valor) a partir de um arquivo de parâmetros do DAKOTA.
makeFileName.m	Gera nome dos arquivos que serão manipulados pelo IMEX em cada avaliação de função.
processFilesN.m	Copia o conteúdo do arquivo modelo (<i>template</i>) para um novo arquivo de nome gerado pelo <i>script</i> makeFilename.m, substituindo as variáveis indicadas com marcações (por exemplo \$) por seus valores correspondentes

Tabela 5 Tabela 5 apresenta a relação dos *scripts* de pós-processamento, todos com suas respectivas finalidades.

Tabela 5. *Scripts* do pós-processamento e suas finalidades.

Script	Finalidade
posProcessadorImex.m	Tem a finalidade de, a partir do arquivo de saída do IMEX, extrair informações sobre a produção acumulada de óleo, água e gás por tempo e gerar um arquivo para ser usado por outros <i>scripts</i> , a exemplo do VPL.m.
VPL.m	Tem a finalidade de, a partir das informações geradas pelo IMEX sobre a produção acumulada de óleo, água e gás, calcular o Valor Presente Líquido (VPL).
cleanup.m	Tem a finalidade de remover todos os arquivos temporários criados em cada avaliação de função.

8 Aplicações

8.1 Caso em Estudo

Esse estudo foi aplicado a um caso com peculiaridades bem definidas no que se refere à complexidade de sua geometria, heterogeneidade e esquema de produção. Esse caso apresenta características bem simples e teve como propósito principal estudar e avaliar as técnicas de propagação de incertezas dentro de condições simplificadas e de fácil controle. A simplicidade desse caso permite a criação de situações, sem alteração do contexto, de forma que possa ser feita uma avaliação das peculiaridades do comportamento e identificar dificuldades dos métodos em estudo. É possível que, em modelos com características complexas, essas peculiaridades de comportamento não se mostrem tão evidentes em função de outras variáveis envolvidas, e assim a avaliação poderia não ser satisfatória.

Este trabalho utiliza o mesmo modelo de simulação do Caso 1 da referência [37], o qual tem 51x51x1 células e apresenta as propriedades resumidas na Tabela 6.

Tabela 6. Características do modelo do Caso 1.

PROPRIEDADE	DADO
Malha de Simulação	51(510m) x 51(510m) x 1(4m)
Porosidade	30%
Permeabilidade horizontal média (Kh)	1000 mD
Permeabilidade vertical (Kv)	10% de Kh
Compressibilidade da Rocha @ 200 kgf/cm ²	$510^{-5} \text{ (kgf/cm}^2\text{)}^{-1}$
Contatos entre Fluidos	Sem contatos WOC (<i>Water Oil Contact</i>) e GOC (<i>Gas Oil Contact</i>)
Pressão de Saturação (P _{sat})	273 kgf/cm ²
Viscosidade @ Tres, P _{sat}	0.97 cP
Razão Gás-Óleo de Formação (RGOF)	115.5 m ³ /m ³ std

Esse modelo é resolvido com o método implícito adaptativo disponível no simulador de reservatórios utilizado. Essa opção controla a troca entre a formulação

IMPES e a “Fully Implicit” (Totalmente Implícita) nos blocos da malha. Esse caso em uma máquina SUN com um processador AMD de 2,6 GHz e 4,0 GB de RAM é simulado em cerca de 30 segundos.

O modelo possui um poço injetor de água (I-1), além de dois poços produtores de óleo (P-1 e P-2). São poços verticais completados na única camada do modelo. A permeabilidade horizontal, apresentada na Tabela 6, é um valor base a partir do qual se define valor de permeabilidade de três regiões do modelo, mostradas na Figura 5.

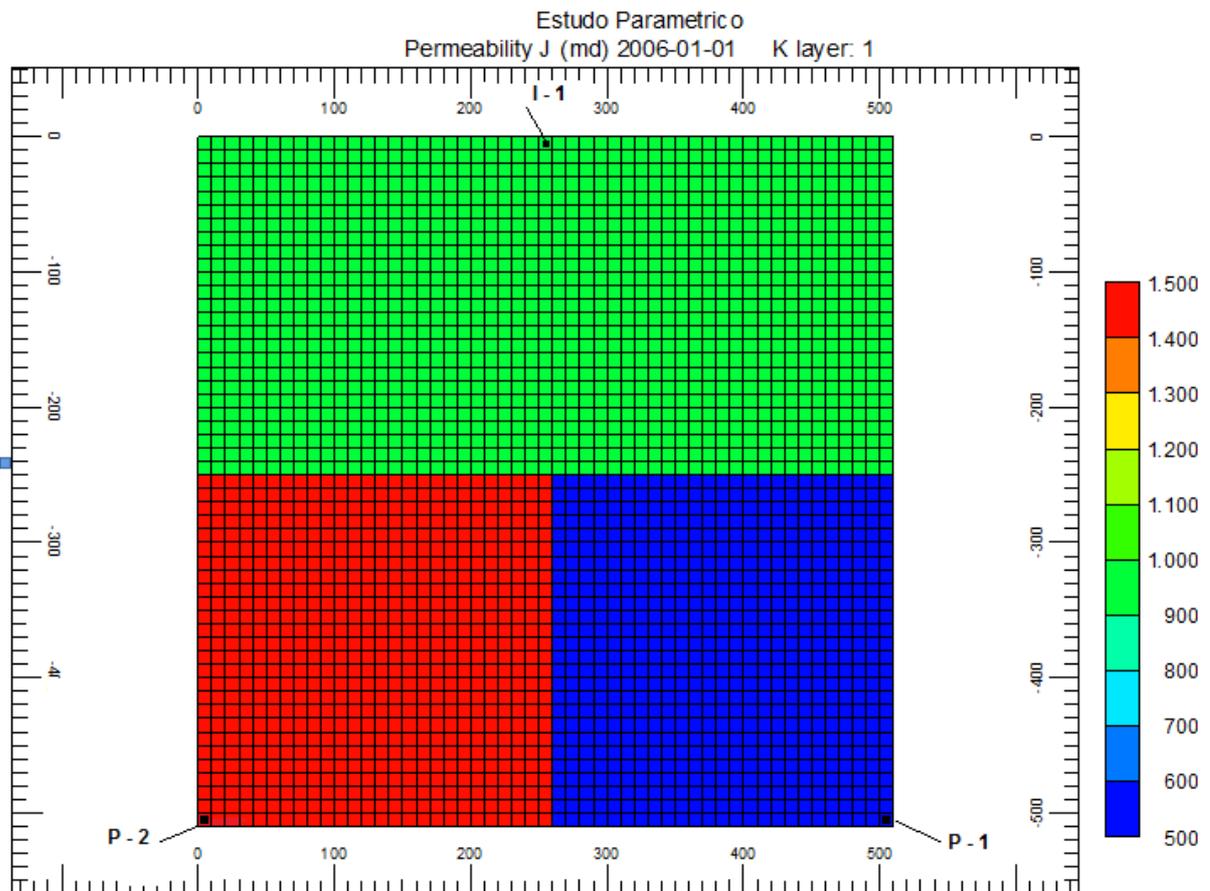


Figura 5. Regiões de permeabilidade horizontal para o Caso 1.

A primeira região, próxima ao poço injetor (I-1), tem valores para a permeabilidade iguais ao valor de base. A segunda região, próxima ao poço produtor P-1, possui valores inferiores ao valor de base; enquanto a terceira região, próxima ao poço produtor P-2, tem valores superiores à permeabilidade base. A razão entre a permeabilidade da região do poço P-2 e a do poço P-1 será doravante designada por “contraste de permeabilidade” ou, simplesmente, “contraste”. Trabalhou-se com

um contraste de permeabilidade de 3:1, a partir do uso dos valores de permeabilidade de 500mD, próximo ao poço P-1, e 1500mD, na região do poço P-2.

O modelo contém ainda a definição de um grupo, representante de uma unidade de produção, onde estão ligados os poços produtores. Esse grupo apresenta um limite à vazão de líquido fixada em 40m³/dia, refletindo a máxima vazão diária de produção capaz de ser tratada pela unidade.

No modelo de simulação, foi utilizada uma opção para definir o rateio com o qual os poços irão produzir, considerando a máxima capacidade especificada para o grupo. Esses parâmetros de rateio foram usados como variáveis de controle do problema, de modo a garantir que em todos os momentos o grupo estará produzindo na sua máxima capacidade.

Os poços produtores são restringidos por uma vazão máxima de líquido de 30m³/dia, sem ter nenhum controle secundário complementar. A ausência de algum controle adicional na operação dos poços produtores se apresenta como uma condição em desacordo com a prática comum nos estudos de reservatório. Usualmente, se especifica algum controle na pressão de fundo de poço ou na pressão da cabeça do poço, complementando a restrição de vazão máxima. Entretanto, essa condição foi mantida visando prover o máximo de simplicidade ao modelo como a proposta original.

Mais adiante a discussão sobre o efeito dos controles será aprofundada, identificando seus inconvenientes. A partir da limitação do grupo e dos controles adotados para os poços, observa-se que os dois produtores não podem operar nas suas máximas capacidades, pois assim ultrapassariam o limite estabelecido para a vazão máxima de líquido de 30m³/dia. Por outro lado, individualmente nenhum dos produtores é capaz de atingir a capacidade do grupo, exigindo a operação simultânea de ambos na maximização da produção.

Seguindo o mesmo princípio de controle dos produtores, o poço injetor foi controlado apenas pela vazão máxima de água injetada, definida em 44m³/dia. Esse valor foi adotado buscando-se a manutenção da pressão média do reservatório.

Como o indicador de risco na maioria das aplicações se refere a algum parâmetro econômico ou financeiro, foi adotado neste trabalho como a variável de interesse para o estudo da quantificação de incerteza o Valor Presente Líquido (VPL) do fluxo de caixa da operação do campo. Sendo assim, seguindo a mesma

proposta de Oliveira [37] a função adotada para representar Valor Presente Líquido (VPL) do fluxo de caixa da operação do campo é dada pela Equação (55):

$$f(\mathbf{x}_{p,t}, \mathbf{u}) = \sum_{\tau=0}^T \left[\frac{1}{(1+d)^\tau} \cdot F_\tau(\mathbf{x}_{p,t}, \mathbf{u}) \right] \quad (55)$$

sendo:

d a taxa de desconto aplicada ao capital;

T o tempo final (horizonte de tempo do projeto ou tempo de concessão do projeto);

F_τ 2ºo fluxo de caixa no tempo τ , tal que:

Sendo F_τ dado pela Equação (56):

$$F_\tau = R_\tau - (Capex_\tau + Opex_\tau + GT_\tau) \quad (56)$$

onde cada termo da Equação (56) é calculado conforme as equações a seguir :

$$R_\tau = \sum_{l=o,g} P_\tau^l \cdot q_\tau^l \cdot \Delta t; \quad (57)$$

$$Capex_\tau = N_w \cdot C_w + C_{f,UEP} + C_{v,UEP} \cdot Q_{\max}, \quad \forall \tau \leq \tau_i; \quad (58)$$

$$Opex_\tau = \sum_{l=o,g,w,wi} C_\tau^l \cdot q_\tau^l \cdot \Delta t; \quad (59)$$

GT_τ (*Government Take*) são as tributações aplicadas ao projeto no tempo τ .

sendo:

P^l o preço de venda do óleo (o) ou do gás (g);

q^l a vazão de produção de óleo (o) ou gás (g), em termos das receitas, e a vazão de óleo (o), gás (g), água produzida (w) e água injetada (wi), em termos dos custos;

N_w o número de poços perfurados;

C_w o custo médio de perfuração de um poço. Pode assumir valores diferentes para tipos distintos de poços;

$C_{f,UEP}$ o custo fixo da Unidade Estacionária de Produção;

$C_{v,UEP}$ o custo variável da Unidade Estacionária de Produção;

- Q_{\max} a vazão máxima compatível com a Unidade Estacionária de Produção;
- C^l o custo de manuseio do óleo (o), gás (g), água produzida (w) e água injetada (wi);
- Δt o intervalo de tempo considerado.

Efetivamente, foi utilizada uma versão simplificada do VPL, levando a efeito apenas as receitas provenientes do óleo, bem como os custos de produção e injeção. A premissa do plano de desenvolvimento dos reservatórios em ser considerado fixo e o fato da unidade de produção também ser constante implica em todos os parâmetros do investimento inicial serem os mesmos para todos os casos estudados dentro de um mesmo modelo. Dessa forma, tratando-se de uma parcela constante, o valor do *Capex* foi adotado igual à zero nesse VPL simplificado, sem qualquer perda de generalidade ou de qualidade desse estudo [37].

8.2 Situações para Análise

Neste trabalho foram consideradas três situações para análise, em termos de quantidades de variáveis incertas, definidas como:

- Situação-1 – Uma variável incerta.
- Situação- 2 – Duas variáveis incertas.
- Situação- 3 – Três variáveis incertas.

Problemas de reservatório como este não possuem solução analítica e, para obtenção de um parâmetro de comparação para as soluções com os métodos do Caos Polinomial e da Colocação Estocástica, foi feita, sempre que necessária, uma análise empregando-se o método de Monte Carlo.

8.2.1 Situação-1

Na Situação-1, o problema é definido com apenas uma variável incerta, que é a permeabilidade na região do poço injetor, chamada de permeabilidade base (k_b). Neste caso, as permeabilidades nas regiões dos poços P-1 e P-2 são definidas em

função da permeabilidade base, e são respectivamente 0,50 e 1,50 vezes a permeabilidade base. Inicialmente, o valor médio da permeabilidade base foi definido como 1000 mD, com uma distribuição normal com desvio padrão igual a 100 mD.

Infelizmente, este esquema não se mostrou apropriado para demonstração de técnicas de propagação de incertezas, pois para esta faixa de valores, de acordo o estudo paramétrico realizado inicialmente, cujo gráfico é apresentado na Figura 6, o VPL é praticamente independente da permeabilidade base.

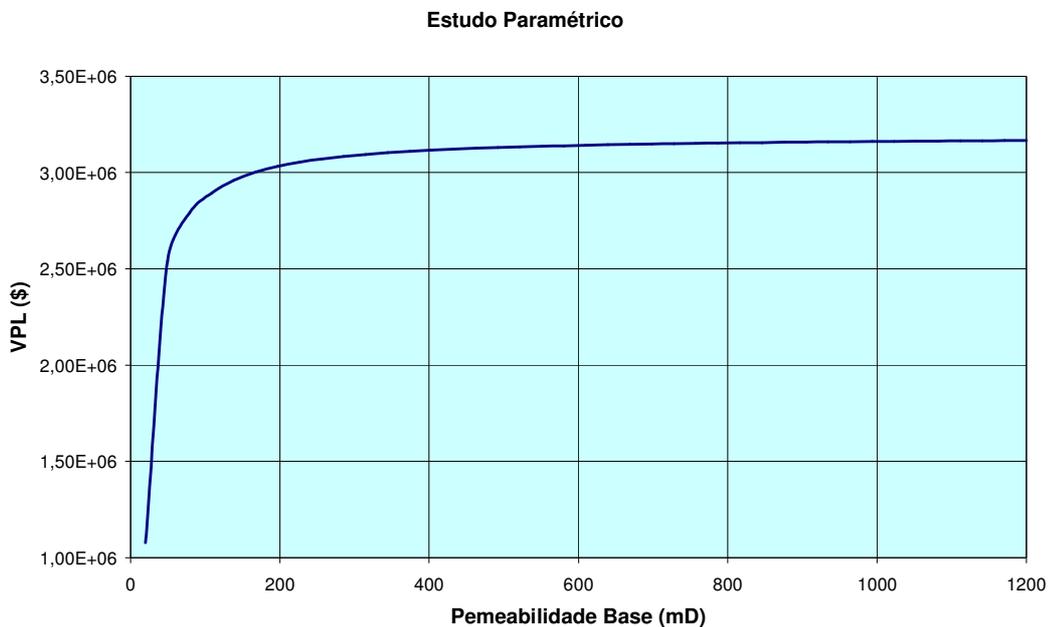


Figura 6. Estudo paramétrico do VPL com uma variável incerta.

Como o VPL é praticamente invariante com a permeabilidade nesta faixa de valores, não há, na prática, incerteza na reposta, sendo o coeficiente de variação calculado da ordem de 10^{-4} . Isto ocorre devido ao cálculo empregado para o VPL ser basicamente uma relação entre o volume de óleo produzido e o volume de água injetado. Com as condições de vazão prescritas como descrito acima, estes volumes não dependem da permeabilidade do campo (muitas outras variáveis, como a pressão média no reservatório, claramente dependem) assim o VPL é praticamente constante. Como este exemplo de nenhuma forma visa representar uma situação real, mas apenas demonstrar as técnicas de propagação de incertezas, o exemplo foi modificado para que a permeabilidade se situasse na faixa na qual sua variação tem efetivamente influência sobre o VPL. Foi tomada então a permeabilidade base

igual a 70 mD, com densidade de probabilidade lognormal e desvio padrão igual a 21 mD.

8.2.1.1 Método de Monte Carlo

Os resultados obtidos com o uso do método de Monte Carlo para a Situação-1, tomando-se a permeabilidade base igual a 70 mD, a densidade de probabilidade como lognormal e o desvio padrão igual a 21 mD, são apresentados na Tabela 7, onde de acordo com o número de amostras, podem ser visualizados: o valor esperado do VPL (representado por $E(VPL)$), o desvio padrão do VPL (representado por $S(VPL)$) e o coeficiente de variação do VPL (representado por COV).

Tabela 7. Resultados para o método de Monte Carlo.

Número de Amostras	$E(VPL)$ (10^6 US\$)	$S(VPL)$ (10^5 US\$)	COV
100	2,6398	2,8711	0,1087
200	2,6844	1,8665	0,0695
300	2,6194	2,7899	0,1060
400	2,4984	2,4984	0,0936
500	2,6883	2,0686	0,0769
1000	2,6727	2,3784	0,0889
2000	2,6713	2,1864	0,0819
4000	2,6706	2,1933	0,0821
6000	2,6666	2,2921	0,0896
8000	2,6673	2,2873	0,0858

Como previsto, as características usuais do método de Monte Carlo são aparentes, a convergência é lenta e não monotônica, e mesmo com 8.000 amostras, em um problema relativamente simples, não está realmente garantida a convergência para três algarismos significativos, quando se considera o valor esperado. Para a variância, a situação é ainda pior.

A Figura 7 mostra histogramas para os valores da variável de entrada, no caso, a permeabilidade base e para os valores da variável de saída, o VPL. Percebe-se claramente que a relação entre estas variáveis é não linear (que pode ser inferido, dada a forma da curva da Figura 6 na região em torno de 70 mD).

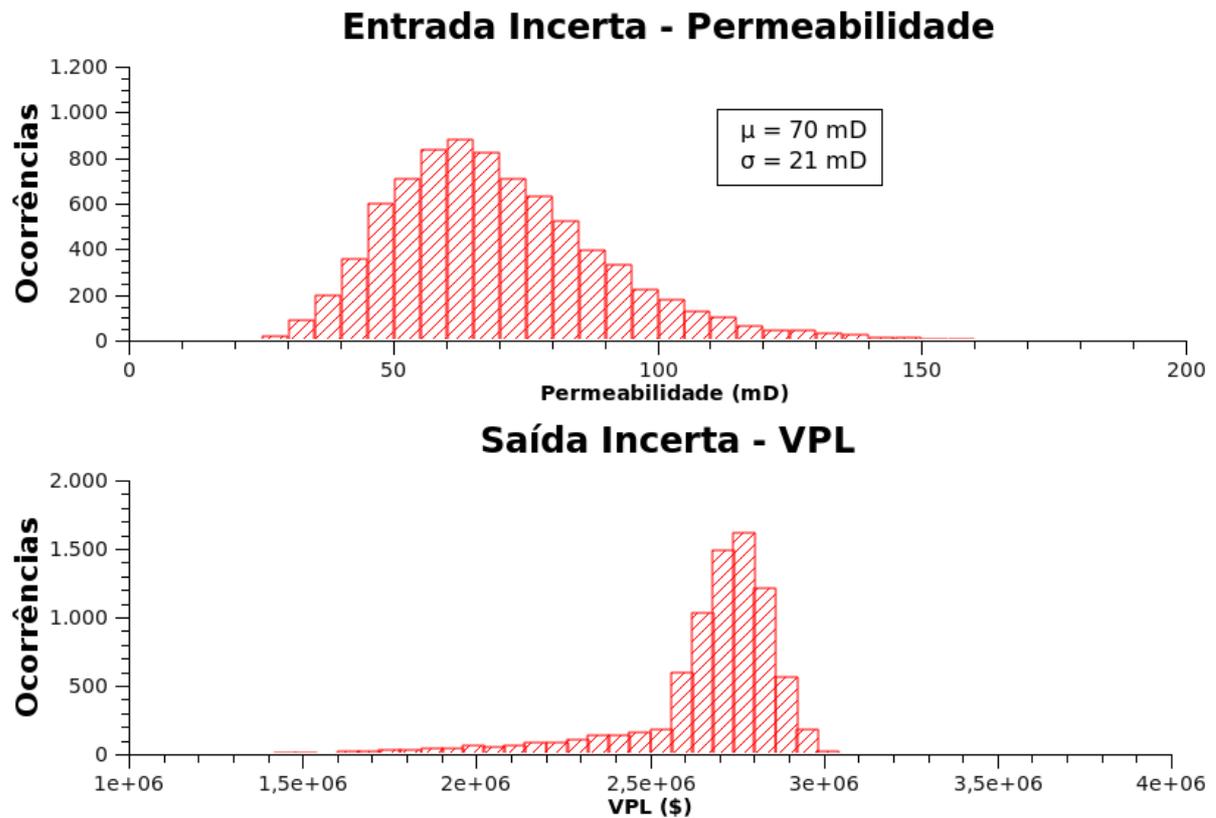


Figura 7. Histogramas das variáveis de entrada e saída.

8.2.1.2 Método do Caos Polinomial

Há duas escolhas importantes na Expansão em Caos Polinomial: a ordem da expansão, isto é, o grau máximo do polinômio usado na expansão, e a ordem de quadratura, isto é, o número de pontos usado na integração numérica. A quadratura Gaussiana empregada integra exatamente um polinômio de grau $2m-1$ com m pontos de integração e, como na integral dos coeficientes aparece um polinômio de grau pelo menos $2p$ (um polinômio de grau p , multiplicado pela função R , demonstrada na Equação (1), aproximada a mais ou menos esta ordem), é previsto que seja necessária a ordem de quadratura $p+1$ para o cálculo dos coeficientes com precisão adequada. Nesta seção foram feitas experiências numéricas variando-se a ordem de expansão e a ordem de quadratura, inclusive com a utilização de ordens de quadratura nitidamente abaixo do necessário, apenas para levantar o comportamento do método nesta situação extrema.

Na Tabela 8 são mostrados o valor esperado $E(VPL)$, o desvio padrão $S(VPL)$ e o coeficiente de variação $COV = S(VPL)/E(VPL)$ computados com uma Expansão em Caos Polinomial de primeira ordem, isto é, empregando-se polinômios lineares apenas, para várias ordens de quadratura OQ . Não poderiam ser esperados bons resultados com uma aproximação utilizando-se ordem da expansão e quadraturas tão baixas, mas é interessante observar que os resultados para o valor esperado são muito bons, quando comparados àqueles obtidos pelo método de Monte Carlo; lembrando que neste caso unidimensional, o número de pontos de avaliação da função é igual à ordem de quadratura $p + 1$.

Tabela 8. Expansão em Caos Polinomial linear ($p = 1$; $OQ=p+1$).

OQ	E(VPL) (10^6 US\$)	S(VPL) (10^5 US\$)	COV
2	2,6961	1,3828	0,05138
3	2,6499	2,3357	0,08814
4	2,6798	1,9169	0,07153
5	2,6580	2,1052	0,07921

Nas Tabelas 9, 10, 11 e 12, são mostrados os resultados correspondentes para expansões de grau 2 a 5, respectivamente, onde percebemos que, para este problema, as ordens de quadratura acima de $p + 1$ fornecem bons resultados, mas não necessariamente os melhores. É interessante observar também que o uso de pontos de integração insuficientes fornece resultados ainda razoáveis para o valor esperado, mas pode fornecer resultados completamente absurdos para a variância.

Tabela 9. Expansão em Caos Polinomial de ordem 2 ($p=2$; $OQ=p+1$).

OQ	E(VPL) (10^6 US\$)	S(VPL) (10^5 US\$)	COV
3	2,6499	2,5750	0,0972
4	2,6798	2,1726	0,0811
5	2,6758	2,2820	0,0859
6	2,6690	2,3102	0,0866

Tabela 10. Expansão em Caos Polinomial de ordem 3 ($p=3$; $OQ=p+1$).

OQ	E(VPL) (10^6 US\$)	S(VPL) (10^5 US\$)	COV
3	2,6495	2,5757	0,0972
4	2,6798	2,2583	0,0843
5	2,6580	2,2920	0,0862
6	2,6685	2,3440	0,0878

Tabela 11. Expansão em Caos Polinomial de ordem 4 ($p=4$; $OQ=p+1$).

OQ	E(VPL) (10^6 US\$)	S(VPL) (10^5 US\$)	COV
3	2,6499	2,7420	0,1035
4	2,6798	2,2583	0,0843
5	2,6580	2,2943	0,0832
6	2,6685	2,3526	0,0882

Tabela 12. Expansão polinomial de ordem 5 ($p=5$; $OQ=p+1$).

OQ	E(VPL) (10^6 US\$)	S(VPL) (10^5 US\$)	COV
4	2,6798	2,3246	0,0867
5	2,6580	2,2943	0,0863
6	2,6685	2,3788	0,0891
7	2,6676	2,2253	0,0834

A Tabela 13 mostra os resultados para o valor esperado e a variância em função da ordem de expansão, obtidos com a ordem de quadratura $p+1$, onde é possível se observar a convergência do método.

Tabela 13. Caos polinomial, uma variável incerta, com $OQ = p + 1$.

OQ	E(VPL) (10^6 US\$)	S(VPL) (10^5 US\$)	COV
2	2,6961	1,3828	0,05138
3	2,6499	2,5750	0,0972
4	2,6798	2,2583	0,0843
5	2,6580	2,2943	0,0832
6	2,6685	2,3788	0,0891

A Tabela 14 apresenta a diferença entre os resultados dos experimentos obtidos com o Caos Polinomial, para as ordens de quadratura de 2 a 6 e ordens de expansão de 1 a 5, com os valores do método de Monte Carlo, para 1.000 amostras. Sabendo-se de antemão que o método de Monte Carlo exige uma grande quantidade de amostras para uma boa aproximação, mesmo assim já é possível uma conclusão a partir da Tabela 14, observando-se que bons resultados podem ser obtidos com expansões de baixa ordem, extremamente econômicas.

No entanto, não deve ser esperado muito mais do que dois ou três algarismos significativos de precisão no resultado. A baixa precisão observada pode ser devida ao ruído numérico da simulação, pois na utilização do método para aproximar funções não lineares analíticas são obtidos resultados mais precisos, condição que precisa ser investigada em trabalhos futuros.

Tabela 14. Diferença entre os valores do Caos Polinomial e Monte Carlo para 1000 amostras (Valores da ordem de 10^6 US\$ para E(VPL) e 10^5 US\$ para S(VPL)).

OQ	Ordem de expansão p=1		Ordem de expansão p=2		Ordem de expansão p=3		Ordem de expansão p=4		Ordem de expansão p=5	
	E(VPL)	S(VPL)								
2	0,0234	-0,9956								
3	-0,0228	-0,0427	-0,0228	-0,0682	-0,0232	0,1973	-0,0228	0,3636		
4	0,0071	-0,4615	0,0071	-0,2058	0,0071	-0,1201	0,0071	-0,1201	0,0071	-0,0538
5	-0,0147	-0,2732	0,0031	-0,0964	-0,0147	-0,0864	-0,0147	-0,0841	-0,0147	-0,0841
6			-0,0037	-0,0682	-0,0042	-0,0344	-0,0042	-0,0258	-0,0042	-0,0004
7									-0,0051	-0,1531

8.2.1.3 Método da Colocação Estocástica

Na Tabela 15 são apresentados os resultados obtidos com o método da Colocação Estocástica, para diversas ordens de interpolação. A ordem de interpolação, que é o grau do polinômio interpolador, será denominada aqui por *OE*. Neste caso, o número de avaliações de função é um a mais do que a ordem de interpolação, ou seja, $OE+1$.

Tabela 15. Colocação Estocástica.

OE	E(VPL) (10^6 US\$)	S(VPL) (10^5 US\$)	COV
1	2,7267	0	0
2	2,6962	1,3820	0,05128
3	2,6500	2,5757	0,09720
4	2,6800	2,2583	0,08427
5	2,6580	2,2943	0,08632
6	2,6685	2,3788	0,08915

Foi observado nos resultados apresentados nas tabelas 13 e 15 que os valores do *E(VPL)* e *S(VPL)* obtidos para a Colocação Estocástica e para o Caos Polinomial convergiam quando a ordem de interpolação *OE* do primeiro método era igual à ordem de quadratura, $OQ+1$, do segundo método, ou seja, quando $OE = OQ + 1$. Dessa forma, a Tabela 16 apresenta a diferença entre os valores *E(VPL)* e *S(VPL)* das tabelas 13 e 15, para as ordens de interpolação *OE*, de 2 a 6, correspondente ao método da Colocação Estocástica, e ordens de quadratura, *OQ*, de 1 a 5, do Caos Polinomial.

Tabela 16. Diferença entre os valores da Colocação Estocástica e Caos Polinomial (Valores da ordem de 10^6 US\$).

OE=OQ+1	Diferença entre E(VPL)	Diferença entre S(VPL)
2	-0,0001	0,0008
3	-0,0001	-0,0007
4	-0,0002	0,0000
5	0,0000	0,0000
6	-0,0000	0,0000

Observa-se, comparando-se os valores obtidos pelos métodos da Colocação Estocástica e de Monte Carlo, que a convergência para o valor esperado com poucos números significativos é muito rápida, entretanto, novamente, se resultados com alta precisão são necessários, mais investigação é necessária. Por outro lado, observa-se pela Tabela 16, que os resultados obtidos com os métodos do Caos Polinomial e da Colocação Estocástica são equivalentes, principalmente quando $OE = OQ + 1$. No entanto não é possível afirmar que este fato ocorra em qualquer situação. A Colocação Estocástica é conceitualmente bem mais simples do que o Caos Polinomial e é uma técnica muito interessante para implementações independentes.

8.2.2 Situação-2

Para testar as técnicas com duas variáveis incertas, o problema descrito na seção 8.2.1 foi modificado para que a razão entre as permeabilidades nas regiões dos poços produtores fosse também uma variável incerta. Considerando que um valor para esta razão seja r e que a permeabilidade base seja k_b , a maior permeabilidade, na região do poço P-1 será dada por $2k_b r / (r+1)$, enquanto a menor, na região do poço P-2, será dada por $2k_b / (r+1)$. Para este exemplo, esta variável foi tomada com uma distribuição normal, com média 3,0 e desvio padrão 0,66. A permeabilidade base foi tomada como no exemplo anterior. Um sumário dos resultados obtidos com o método do Caos Polinomial é mostrado na Tabela 17, onde são apresentados apenas os melhores resultados obtidos em cada ordem de expansão. É importante notar que, com duas variáveis, o número de pontos de integração necessário é igual a OQ^2 .

Tabela 17. Caos Polinomial, para duas variáveis incertas.

OE	OQ	E(VPL) (10 ⁶ US\$)	S(VPL) (10 ⁵ US\$)	COV
1	2	2,6788	1,9184	0,07162
2	3	2,6788	2,1744	0,08117
3	4	2,6570	2,2917	0,08251

Na Tabela 18 são mostrados os resultados para o método da Colocação Estocástica, sendo o número de pontos de integração $(OE+1)^2$.

Tabela 18. Colocação Estocástica, para duas variáveis incertas.

OE	E(VPL) (10 ⁶ US\$)	S(VPL) (10 ⁵ US\$)	COV
1	2,7284	0	0
2	2,6951	1,3891	0,05134
3	2,6488	2,5742	0,09718
4	2,6788	2,2600	0,08437
5	2,6569	2,2940	0,08624

Foi observado também, para o caso de duas variáveis incertas, que os valores do $E(VPL)$ e $S(VPL)$, obtidos para a Colocação Estocástica e para o Caos Polinomial, convergiam quando a ordem de interpolação OE do primeiro método era igual à ordem de quadratura, $OQ+1$, do segundo método, ou seja, quando $OE = OQ + 1$. Dessa forma, a Tabela 19 apresenta a diferença entre os valores $E(VPL)$ e $S(VPL)$ das tabelas 17 e 18, para as ordens de interpolação OE , de 3 a 5, correspondente ao método da Colocação Estocástica, e ordens de quadratura, OQ , de 2 a 4, do Caos Polinomial.

Tabela 19. Diferença entre os valores do Caos Polinomial e Colocação Estocástica para duas variáveis incertas.

OE=OQ+1	Diferença entre E(VPL) (10 ⁶ US\$)	Diferença entre S(VPL) (10 ⁵ US\$)
3	0,0300	0,5293
4	0,0000	-0,0856
5	0,0001	-0,0023

Observa-se pela Tabela 19 que o comportamento obtido com uma variável aleatória repete-se para o caso de duas variáveis, ou seja, os resultados obtidos com os métodos do Caos Polinomial e da Colocação Estocástica são equivalentes, principalmente quando $OE = OQ + 1$. No entanto não é possível afirmar também que este fato ocorra em qualquer situação.

8.2.3 Situação-3

Nesta Situação, as permeabilidades das três camadas do reservatório foram tomadas como variáveis incertas independentes, com distribuição lognormal, médias iguais a 70, 50 e 35 mD, e desvios padrão iguais a 21, 15 e 11 mD, tomados a partir da camada superior do reservatório. Um resumo dos resultados obtidos com o método de Monte Carlo é mostrado na Tabela 20.

Tabela 20. Método de Monte Carlo, para três variáveis incertas.

Amostras	E(VPL) (10 ⁶ US\$)	S(VPL) (10 ⁵ US\$)	COV
2000	2,3698	2,7747	0,1108
4000	2,3595	2,8483	0,1207
6000	2,3647	2,8137	0,1190
8000	2,3674	2,8016	0,1183

A Tabela 21 mostra o resumo dos melhores resultados obtidos com o método do Caos Polinomial, para ordens de expansão iguais a 2 e 3.

Tabela 21. Caos Polinomial, para três variáveis incertas, por quadratura numérica.

OE	OQ	E(VPL) (10 ⁶ US\$)	S(VPL) (10 ⁵ US\$)	COV
2	4	2,3709	2,7655	0,11664
3	4	2,3709	2,7870	0,11756
3	5	2,3682	2,7727	0,11710

Para os resultados da Tabela 21, os coeficientes foram calculados por quadratura numérica. Como se trata de um problema tridimensional, o número de pontos de avaliação de função neste caso é OQ^3 , o que já é bastante significativo. Por exemplo, para a expansão de grau 3 com ordem de quadratura 5, são necessários 125 pontos de avaliação de função.

Devido a este alto número de análises requeridas, avaliou-se também neste problema a técnica de regressão linear para o cálculo aproximado dos coeficientes. Nas Tabelas 22 e 23 são mostrados os resultados para expansões de ordem 2 e 3 respectivamente, empregando-se a técnica de regressão linear.

Tabela 22. Caos Polinomial de ordem 2, para 3 variáveis incertas, por regressão.

OE	NP	E(VPL) (10 ⁶ US\$)	S(VPL) (10 ⁵ US\$)	COV
2	10	2,3569	3,4227	0,14522
2	15	2,3659	3,0825	0,13029
2	20	2,3518	2,8902	0,12289
2	25	2,3724	2,8935	0,12200
2	30	2,3593	3,0784	0,13040

Tabela 23. Caos Polinomial de ordem 3, para 3 variáveis incertas, por regressão.

OE	NP	E(VPL) (10 ⁶ US\$)	S(VPL) (10 ⁵ US\$)	COV
3	20	2,3339	4,2340	0,18141
3	30	2,3863	3,1046	0,13018
3	40	2,3607	2,7823	0,11786
3	50	2,3737	2,9119	0,12267
3	60	2,3603	2,6973	0,11428
3	80	2,3727	2,8634	0,12681

Nas Tabelas 22 e 23, o número de pontos de interpolação empregado é mostrado na coluna *NP*. Por resultados experimentais, sabe-se que um bom número de pontos de interpolação é da ordem de 2 vezes o número de coeficientes necessários na expansão, e estas tabelas confirmam este resultado. Usando-se a Equação (49), o número de coeficientes necessários para expansão de ordem 2 com 3 variáveis é 10, e para expansão de ordem 3 com 3 variáveis é 20.

Neste problema foram obtidos resultados de qualidade equivalente, de maneira mais econômica, mas não é possível afirmar que isto se aplique a todos os casos.

8.3 Problema de Alta Dimensão

Em problemas mais realistas de simulação de reservatórios não é possível a aplicação direta das técnicas de propagação de incertezas. Isto ocorre por não ser possível nem correta a associação de uma variável incerta a cada célula do reservatório. Sendo assim, para criar uma descrição dos campos estocásticos de permeabilidade do reservatório em termos de um número finito e pequeno de variáveis aleatórias é necessária que seja aplicada, previamente, uma técnica de redução dimensional. Nesse estudo foi empregada a Expansão de Karhunen-Loève, descrita na seção 6.1.

Para demonstrar essa técnica foi criado um problema modelo, com a mesma geometria e condições operacionais do problema descrito na seção 8.1, porém com o campo de permeabilidades gerado aleatoriamente, a partir de uma dada função de correlação espacial. Esta função de correlação é usada para gerar realizações equiprováveis do campo, que são os dados de entrada para a decomposição de Karhunen-Loève via núcleo (KKLE). Em um problema real, a função de covariância não é sequer conhecida, e estas realizações seriam determinadas por procedimentos estatísticos mais elaborados, e não apenas com a aplicação direta de um gerador de números aleatórios correlacionados. É importante ressaltar que, na técnica aqui empregada, os autovetores e autovalores da matriz de covariância são calculados sem que jamais a própria matriz seja montada, sendo esta a essência da técnica via núcleo. A função de correlação usada para gerar as realizações é dada pela Equação (60):

$$\text{cov}(\mathbf{x}, \mathbf{y}) = \sigma e^{-\frac{|x_1 - y_1|}{l_1}} e^{-\frac{|x_2 - y_2|}{l_2}} \quad (60)$$

onde \mathbf{x} e \mathbf{y} são coordenadas dos centros das células, σ é a variância e l_1 e l_2 são comprimentos de correlação. Esta função foi empregada para gerar campos normais com média zero e a variância empregada foi 1.61. Este valor foi adotado, pois os campos normais gerados foram transformados em campos lognormais (este valor corresponde a um desvio padrão de 2.000 mD na distribuição lognormal final), e os comprimentos de correlação empregados foram de 25 e 35 comprimentos de célula.

A distribuição de probabilidades mais natural para campos de permeabilidades gerados aleatoriamente é a lognormal, porém os campos recuperados com a Equação (36) são Gaussianos. A expansão foi aplicada, portanto, no logaritmo natural dos campos de permeabilidade desejada. Neste exemplo, desejava-se campos lognormais com média 1000 mD e desvio padrão igual a 2.000 mD, e os parâmetros para a distribuição normal associada são média igual a 6,103 e desvio padrão igual a 1,27 (variância igual a 1,61). Algumas realizações normais geradas com a Equação (36), aplicando-se a função de correlação dada pela Equação (60), são mostradas nas Figuras 8 a 11. Cada uma dessas figuras apresenta realizações geradas a partir de conjuntos de amostras diferentes. Foram usadas 1000 realizações, como as mostradas nessas figuras, para computar os autovalores e autovetores da matriz de covariância.

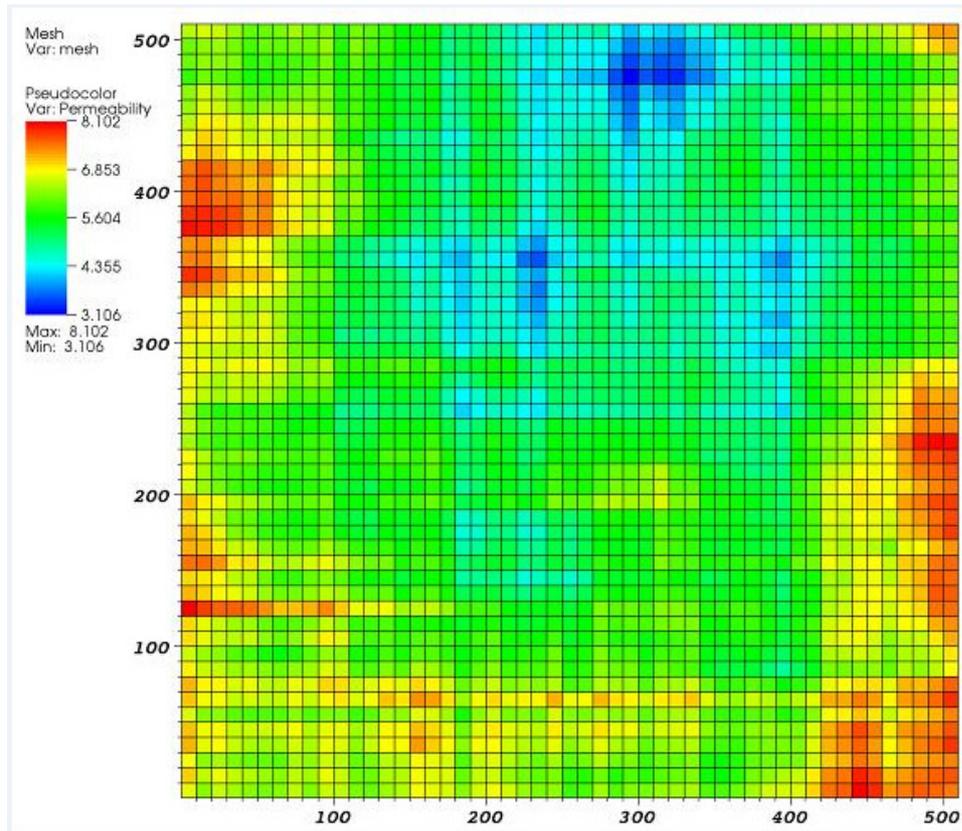


Figura 8. Realizações normais – primeiro conjunto de amostras.

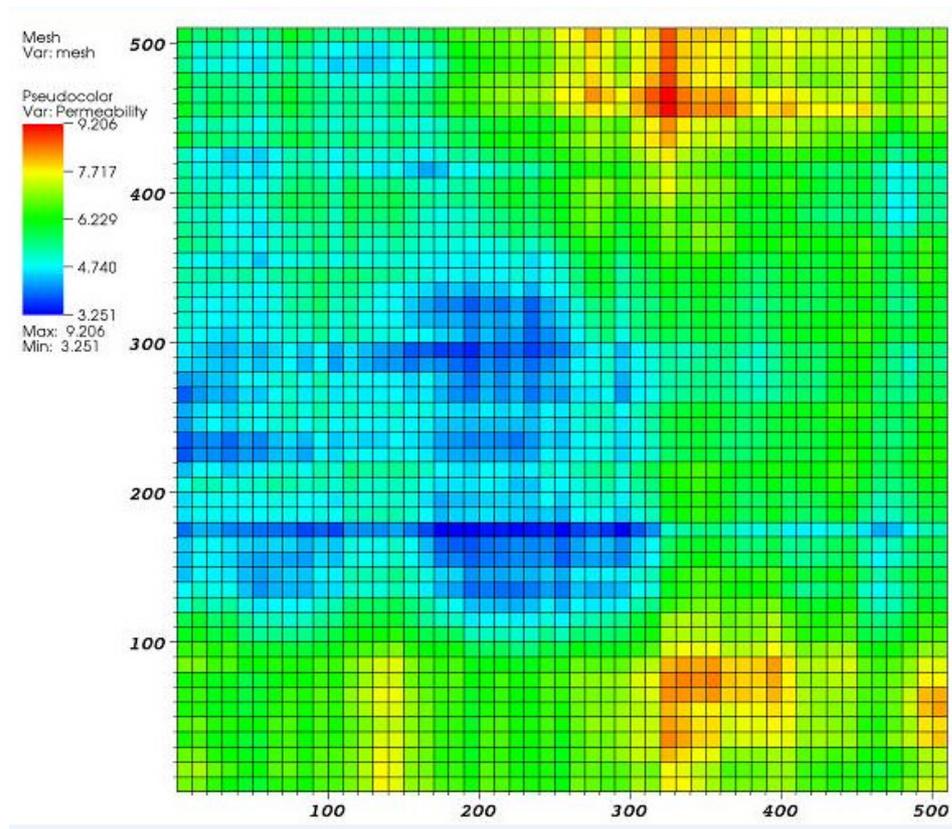


Figura 9. Realizações normais – segundo conjunto de amostras.

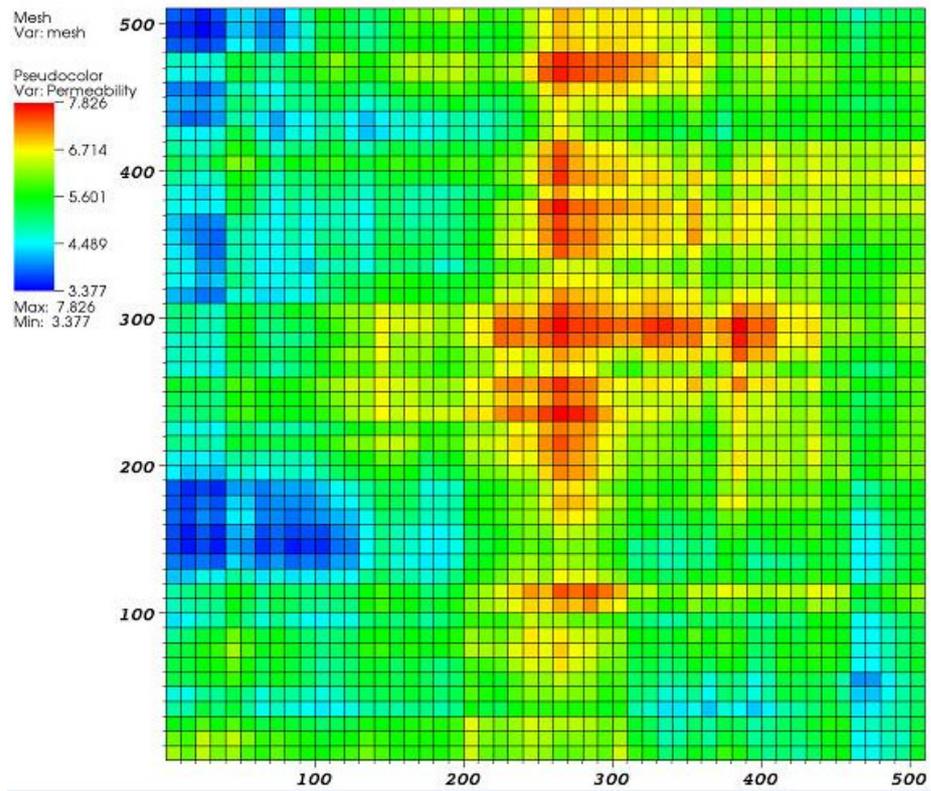


Figura 10. Realizações normais – terceiro conjunto de amostras.

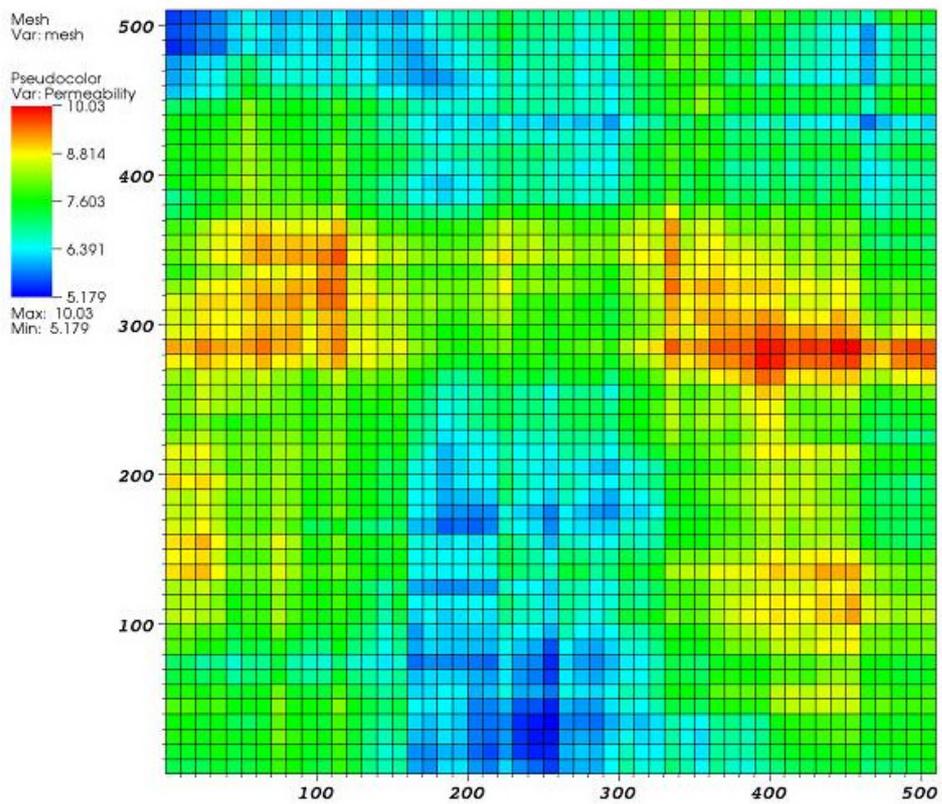


Figura 11. Realizações normais – quarto conjunto de amostras.

A Figura 12 mostra um gráfico da magnitude destes autovalores, ordenados em ordem decrescente.

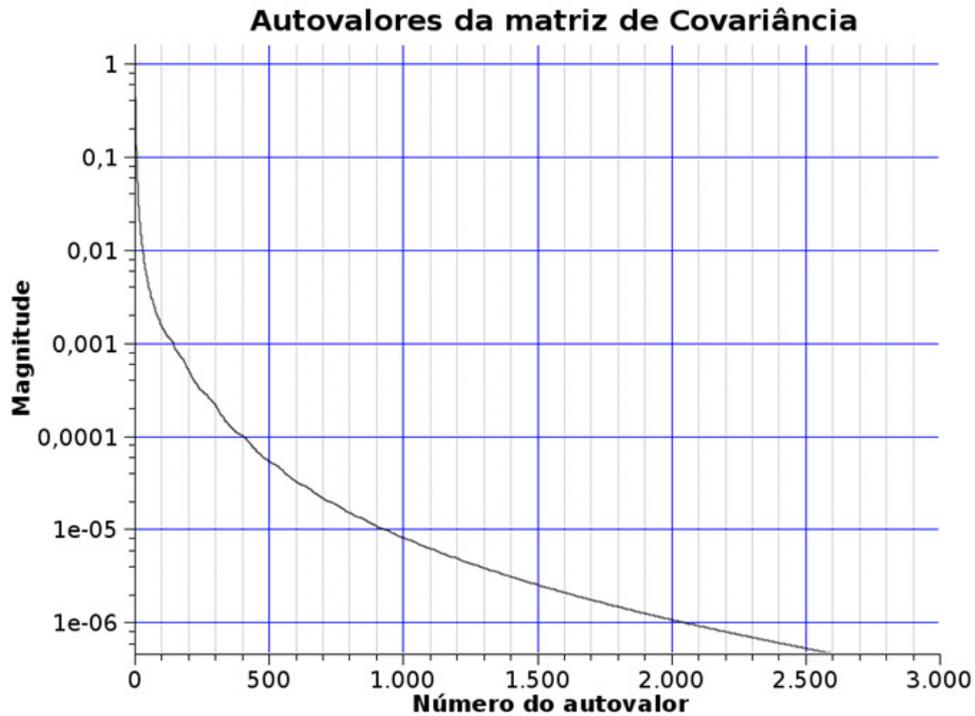


Figura 12. Magnitude dos autovalores.

Através da Figura 12 podemos perceber que, para os parâmetros empregados, os autovalores decaem realmente muito rapidamente.

A Figura 13 mostra a soma acumulada normalizada das magnitudes dos 200 maiores autovalores e pode ser usada na escolha do número de autovalores e autovetores usados na expansão. Normalmente, uma quantidade de autovalores cuja soma acumulada normalizada seja da ordem de 0,80 é uma escolha razoável. Neste caso, obtêm-se este valor com aproximadamente 10 autovalores.

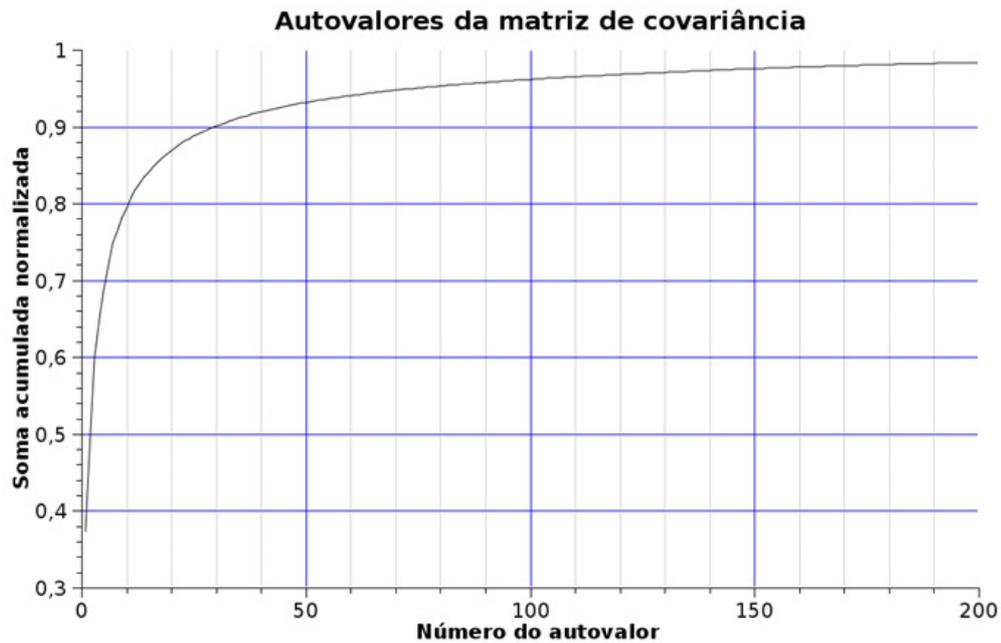


Figura 13. Soma acumulada normalizada dos autovalores.

A propagação de incertezas é feita então considerando as variáveis ξ_i como as variáveis incertas. Para a aplicação do método de Monte Carlo, por exemplo, são tomadas dez amostras independentes e identicamente distribuídas da distribuição normal padrão. Cada uma destas amostras é associada a uma das variáveis ξ_i e o somatório mostrado na Equação (36) é calculado. Isto gera uma nova realização, que é usada para uma análise do simulador, as variáveis de interesse da saída são armazenadas e estatísticas destas variáveis são calculadas. Este procedimento é repetido até que a convergência seja considerada atingida.

Para aplicação dos métodos do Caos Polinomial ou da Colocação Estocástica, as variáveis ξ_i são usadas diretamente como as variáveis incertas na expansão ou interpolação, sendo seus valores não mais gerados aleatoriamente, mas vindos dos pontos de quadratura numérica no caso do método do Caos Polinomial, e dos pontos de interpolação, no caso do método da Colocação Estocástica. Novamente o somatório da Equação (36) é feito, e a realização assim gerada é usada em uma simulação. Os valores de interesse na saída são usados como apropriado na técnica correspondente: como integrandos no método do Caos Polinomial, e como valores, nos pontos de interpolação para o método da Colocação Estocástica.

Um problema óbvio surge aqui, pois, mesmo com apenas 10 variáveis incertas, com uma expansão polinomial de ordem 3, por exemplo, provavelmente será

necessária uma ordem de quadratura igual a 4, o que implica em $4^{10}=1.048.576$ avaliações de função, o que é, claramente, um número impraticável. Conforme mostrado na Tabela 2, o número de coeficientes a determinar para ordem 3 com 10 variáveis é igual a 286, de forma que o cálculo aproximado dos coeficientes por regressão linear requer apenas 572 avaliações, um número viável quando se considera o uso da computação paralela. Esta foi a técnica usada neste exemplo.

É importante ressaltar, no entanto, que a Expansão de Karhunen-Loève não é a única expansão possível para campos estocásticos, e uma de suas desvantagens é considerar apenas a magnitude dos autovalores, e não o seu efeito na solução final, para a escolha dos pares que formarão a expansão. Dessa forma, devem ser investigadas expansões mais aprimoradas que levem isto em consideração, podendo assim diminuir o número de variáveis incertas na entrada do problema.

Algumas realizações criadas com a Expansão de Karhunen-Loève usando-se o logaritmo natural da permeabilidade estão mostradas nas Figuras 14 a 17.

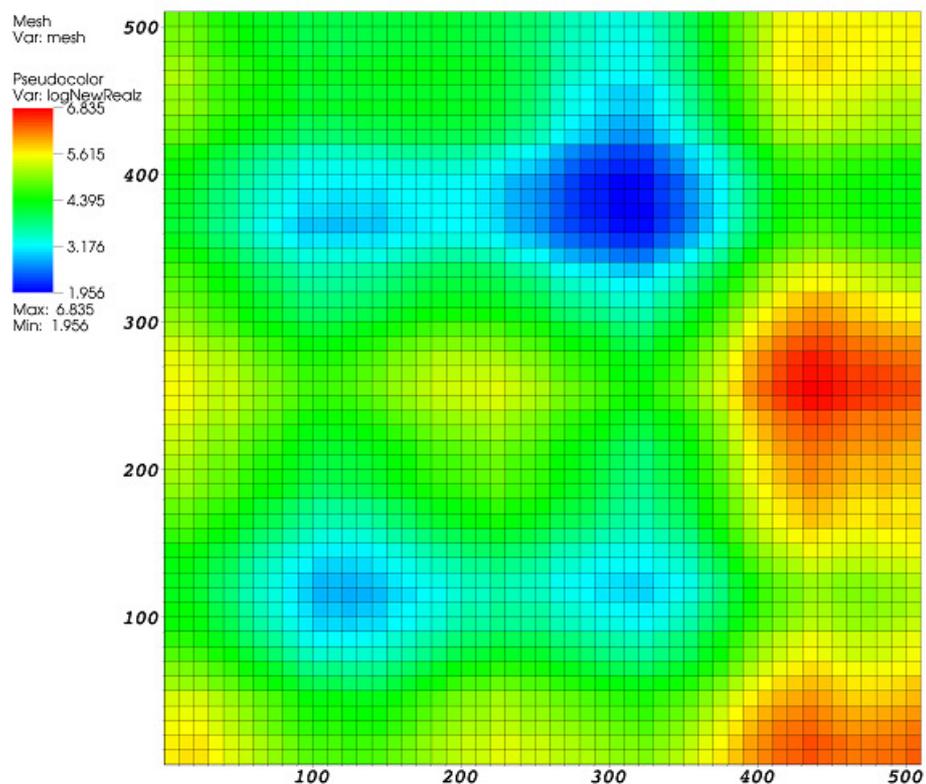


Figura 14. Realizações com o log natural da permeabilidade – primeiro conjunto de amostras.

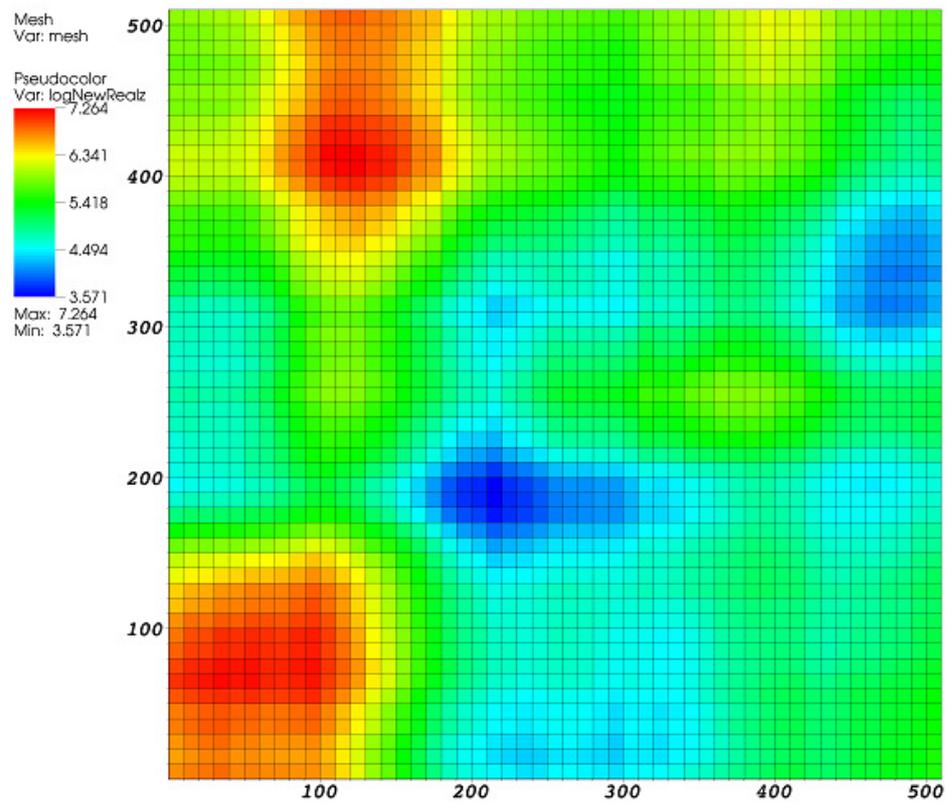


Figura 15. Realizações com o log natural da permeabilidade – segundo conjunto de amostras.

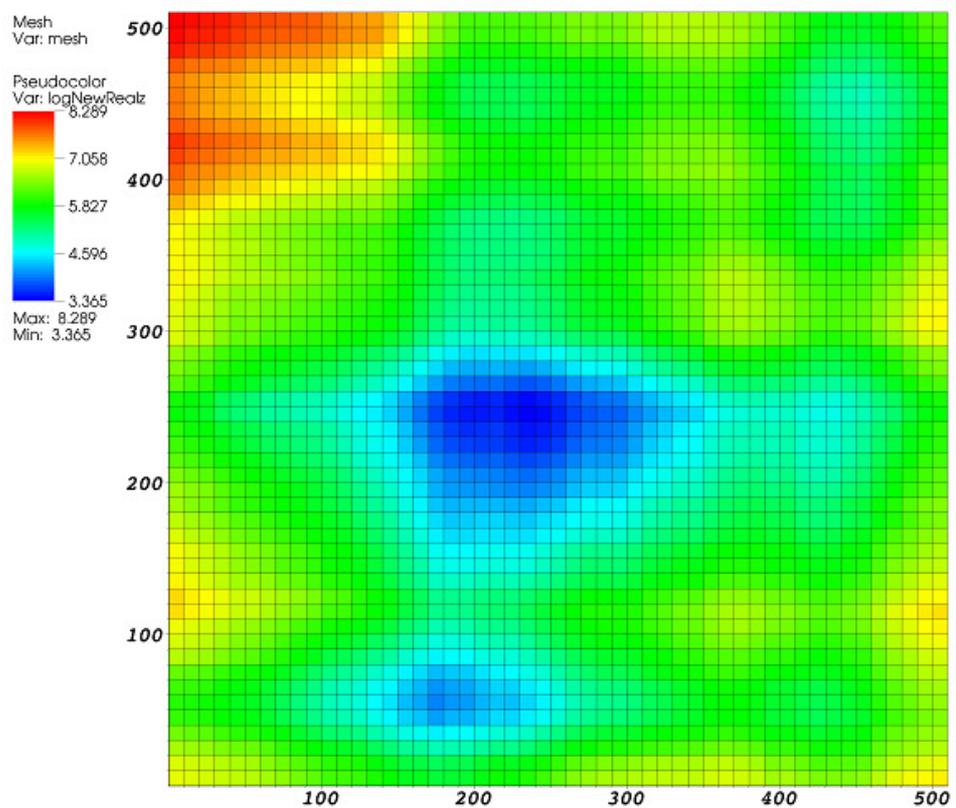


Figura 16. Realizações com o log natural da permeabilidade – terceiro conjunto de amostras.

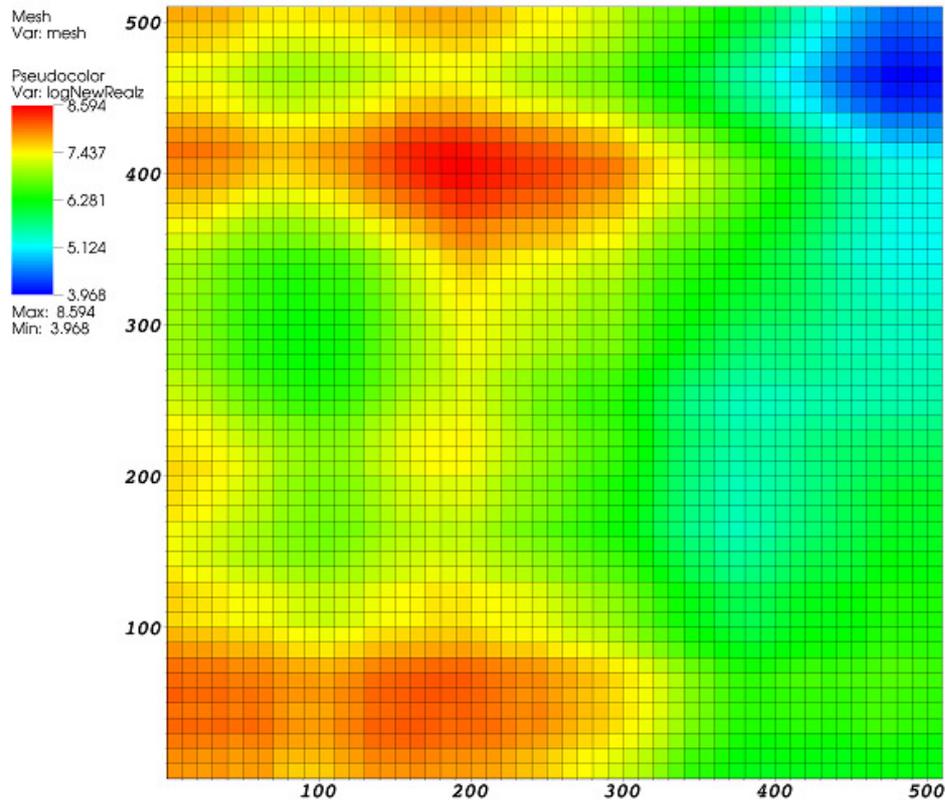


Figura 17. Realizações com o log natural da permeabilidade – quarto conjunto de amostras.

É interessante notar que, como esperado, que a variabilidade de alta frequência espacial é filtrada pela Expansão de Karhunen-Loève. Como o problema com a configuração inicial de injeção e produção é muito pouco sensível à permeabilidade, o problema foi novamente modificado. Nestas análises, não foi imposta restrição de vazão nos poços produtores, mas a vazão total da plataforma foi restrita a 250 m³/dia, e os poços produtores eram fechados quando o corte de água atingia 92%. Além disto, o campo era abandonado (isto é, a análise era interrompida) quando os dois poços produtores fossem fechados. O poço injetor foi operado com a vazão prescrita de 200 m³/dia, e as pressões mínimas de fundo de poço adotadas foram 180atm para os poços produtores e 220atm para o poço injetor. As análises foram executadas por 30 anos, ou até o abandono do campo.

Para ilustração, as Figuras 18 a 21 mostram os campos de saturação de água ao final de cada análise, correspondendo às realizações mostradas nas Figuras 14 a 17.

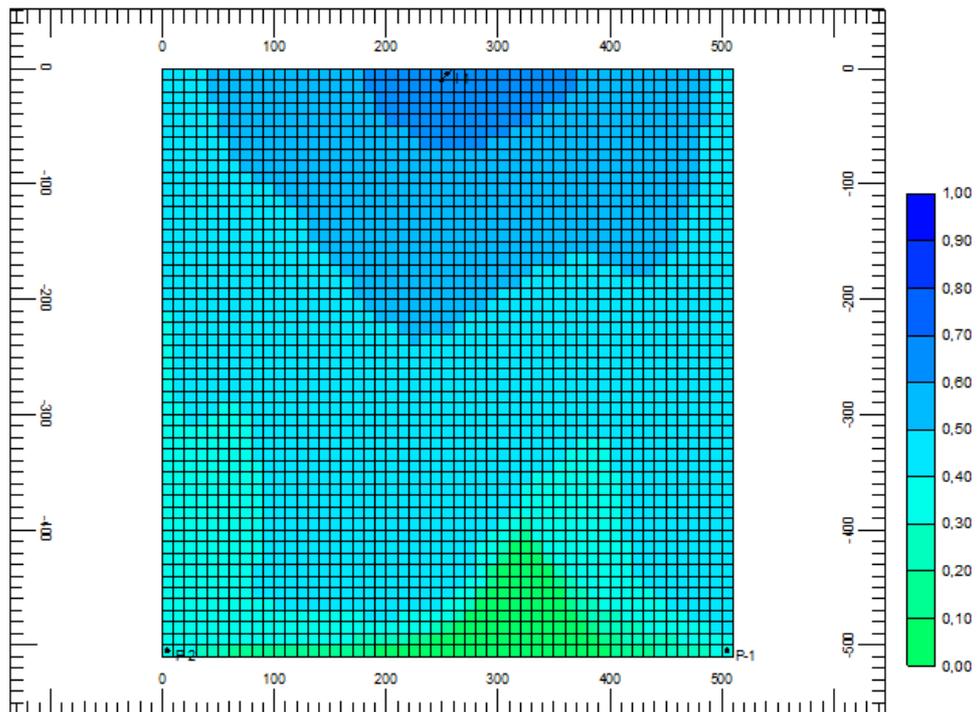


Figura 18. Saturação de água ao final da análise, correspondente às realizações da Figura 14.

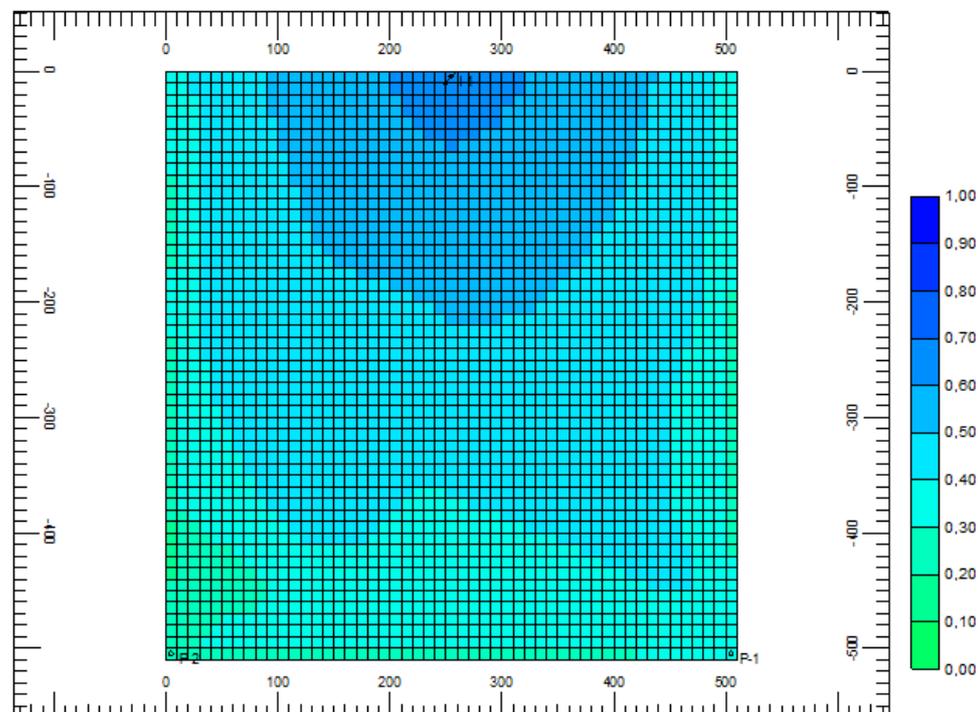


Figura 19. Saturação de água ao final da análise, correspondente às realizações da Figura 15.

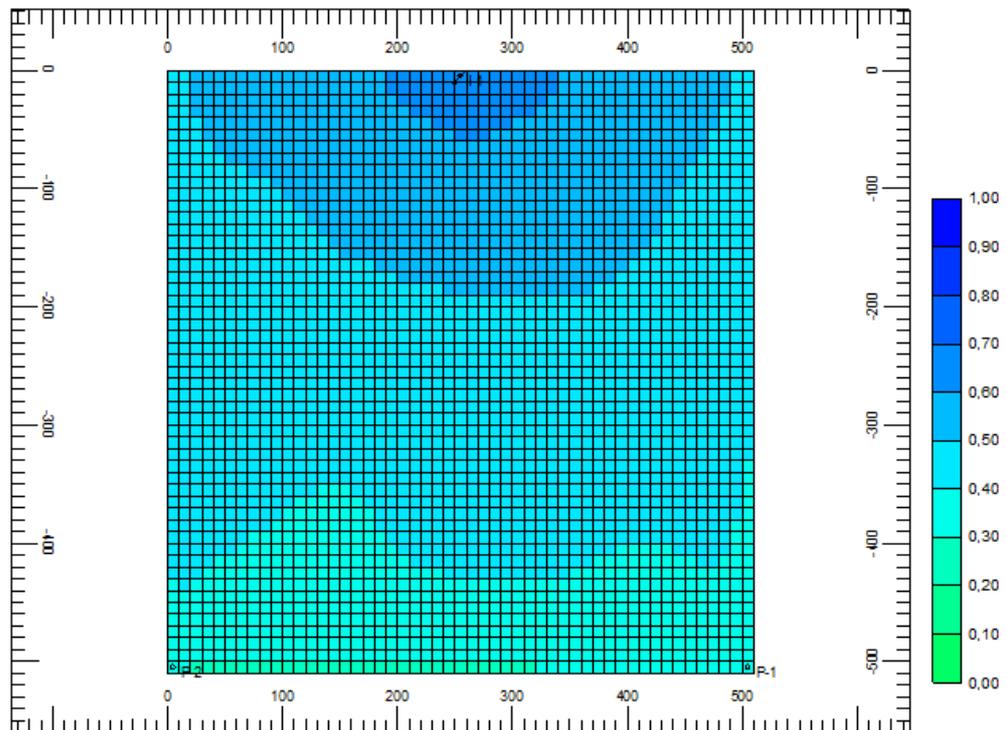


Figura 20. Saturação de água ao final da análise, correspondente às realizações da Figura 16.

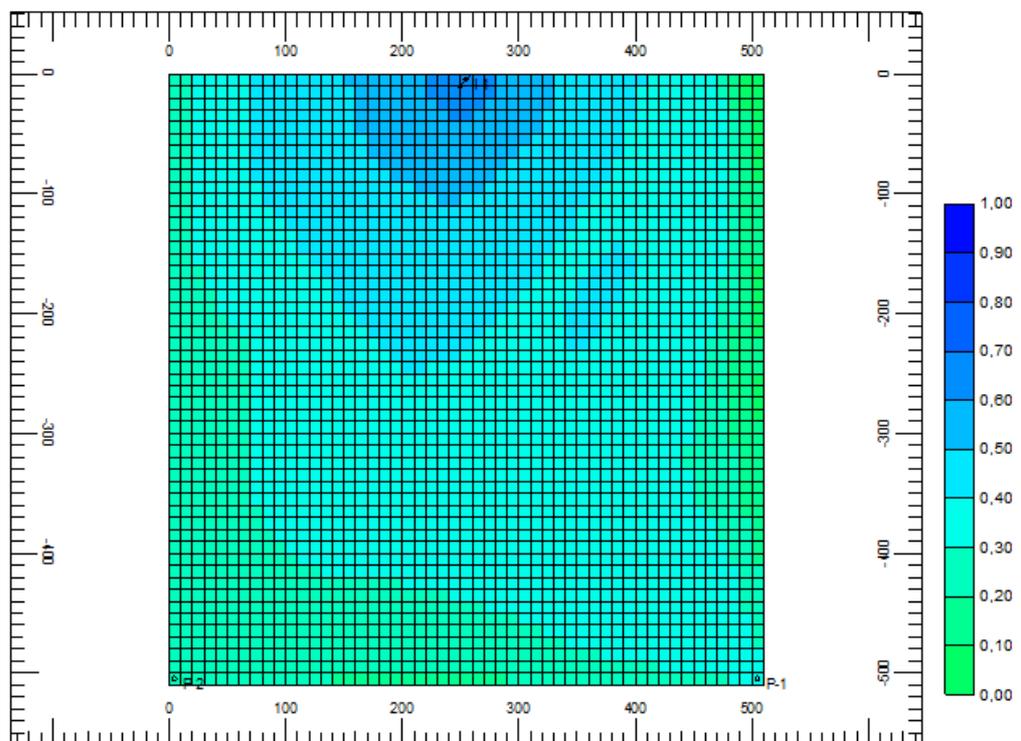


Figura 21. Saturação de água ao final da análise, correspondente às realizações da Figura 17.

A Figura 22 mostra as produções acumuladas de óleo e água para cada uma das realizações acima.

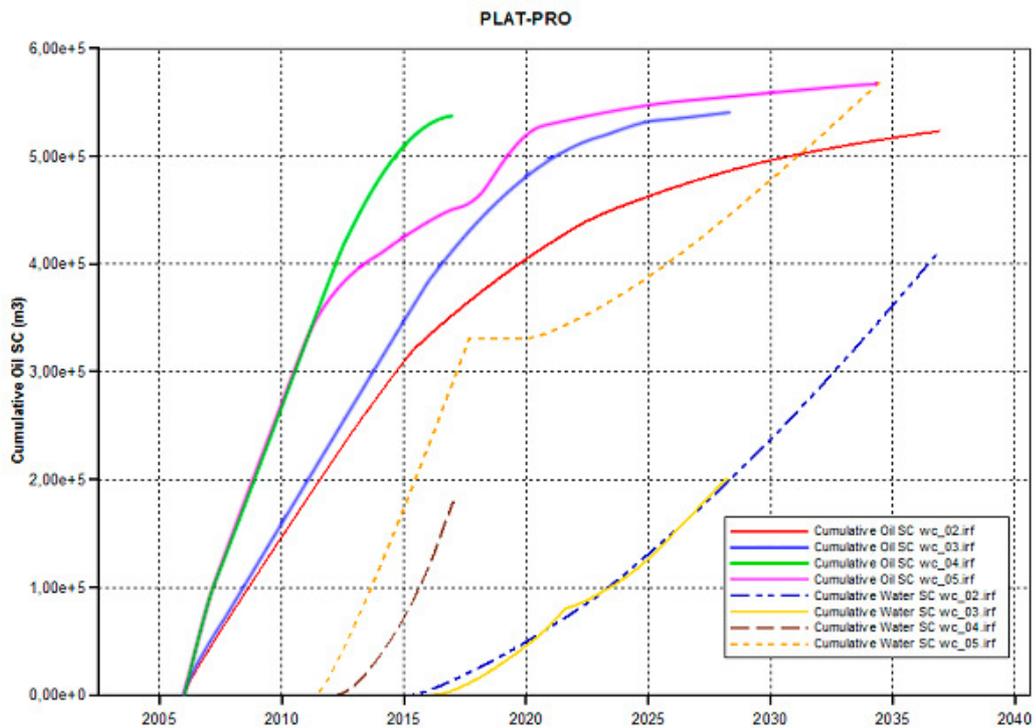


Figura 22. Produções acumuladas de óleo e água.

Na Figura 22, as legendas wc_02, wc_03, wc_04 e wc_05 correspondem às realizações mostradas nas Figuras 14 a 17, respectivamente. É muito importante ressaltar que as realizações mostradas nas Figuras 14 a 17, bem como os resultados mostrados nas Figuras 18 a 22, são resultados aleatórios, pois derivam da Expansão de Karhunen-Loève, que é feita em termos de variáveis aleatórias. Desta forma, para cada técnica de propagação de incertezas e até mesmo em diferentes execuções da mesma técnica, realizações e resultados diferentes serão gerados.

Os resultados obtidos com a Expansão em Caos Polinomial, para o caso das 10 variáveis, estão mostrados na Tabela 24. Percebe-se que os resultados obtidos utilizando-se a regressão linear com a expansão quadrática e apenas 132 avaliações de função foram bastante razoáveis em comparação com aqueles obtidos com a expansão cúbica e 572 avaliações de função.

Tabela 24. Expansão em Caos Polinomial.

OE	NP	E(VPL) (10 ⁶ US\$)	S(VPL) (10 ⁵ US\$)	COV
2	132	7,2752	8,6033	0,1183
3	572	7,2964	8,6190	0,1182

9 Considerações Finais

9.1 Conclusões

No cenário atual, um dos grandes desafios das empresas da área de exploração e produção do petróleo vem sendo a busca por inovações tecnológicas que lhes proporcionem maior rentabilidade, mediante o aumento de produção e redução das perdas financeiras. De forma a contribuir para o atendimento das necessidades das empresas de Óleo e Gás, com o fornecimento de ferramentas mais eficientes para a análise dos riscos envolvidos no processo, este trabalho desenvolveu e testou um sistema computacional que utiliza os métodos do Caos Polinomial e da Colocação Estocástica como técnicas para a propagação de incertezas, aplicado a reservatórios de petróleo. Na aplicação do sistema desenvolvido nas situações de estudo propostas, a análise dos resultados obtidos, com o emprego dos métodos supracitados para a propagação de incertezas, nos conduz às seguintes conclusões, quando analisadas do ponto de vista do desempenho:

- Foi observado nos resultados obtidos para o caso de apenas uma variável incerta que a diferença entre os valores esperados $E(VPL)$ apresentados pelo método de Monte Carlo e o método do Caos Polinomial é praticamente a mesma dentro de uma mesma ordem de quadratura, independente da ordem de expansão avaliada. Supõe-se dessa forma que para o caso de uma variável incerta apenas, se for considerada uma ordem de expansão $p = 1$, qualquer ordem de quadratura superior a $OQ = p + 1$, com OQ pontos de avaliação de função, já apresentarão resultados satisfatórios.
- Considerando-se ainda apenas uma variável incerta, foi observado que os valores esperados $E(VPL)$ e os desvios padrão $S(VPL)$ apresentados pelo método da Colocação Estocástica são equivalentes aos apresentados pelo método do Caos Polinomial, quando a ordem de interpo-

lação OE , do primeiro método, é igual a $OQ + 1$, do segundo. O método do Caos Polinomial demanda: OQ^n pontos de avaliação de função, por quadratura e $2N_t$ por regressão, sendo n o número de variáveis incertas e N_t o número de termos da expansão. Já o método da Colocação Estocástica demanda $(OE + 1)^n$. Considerando-se aqui que, para a mesma precisão apresentada pelo método do Caos Polinomial, a ordem de interpolação da Colocação Estocástica deve ser $OE = OQ + 1$, a quantidade de avaliações de função demandada pelo método da Colocação Estocástica deverá ser $(OQ + 2)^1$, para o caso de uma variável incerta. Sendo assim, o método do Caos Polinomial se mostra mais econômico, pois demanda apenas OQ^1 . No entanto, não é certo afirmar que isto se aplique a todos os problemas.

- Para o caso de duas variáveis incertas os valores esperados $E(VPL)$ e os desvios padrão $S(VPL)$ obtidos pelo método do Caos Polinomial para ordem de quadratura $OQ \geq 2$ se equiparam aos valores esperados $E(VPL)$ e aos desvios padrão $S(VPL)$ obtidos pelo método da Colocação Estocástica, a partir da ordem de interpolação $OQ + 1$. Para o caso de duas variáveis incertas, o método do Caos Polinomial demandou OQ^2 pontos de avaliação de função, enquanto o método da Colocação Estocástica $(OQ + 2)^2$. Percebe-se novamente que, para duas variáveis incertas, o método do Caos Polinomial ainda é mais econômico que o método da Colocação Estocástica uma vez que irá demandar menos pontos de avaliação de função, para a mesma precisão. Da mesma forma, não é certo afirmar que isto se aplique a todos os problemas.
- Para o caso de três variáveis incertas, a quantidade de pontos de avaliação de função demandada pelo método do Caos Polinomial para o cálculo dos coeficientes da expansão por quadratura numérica é igual a OQ^3 . Esta quantidade de pontos de avaliação de função é ainda inferior à quantidade demandada pelo método da Colocação Estocástica, ou seja, $(OQ + 2)^3$. Dessa forma, na medida em que a ordem de quadratura

umenta, para a obtenção de maior precisão, a quantidade de pontos de avaliação se tornará impraticável. Sendo assim é recomendado que a partir de três variáveis incertas seja utilizado o método do Caos Polinomial com o emprego da técnica da regressão linear para o cálculo aproximado dos coeficientes da equação, o que irá demandar uma quantidade de pontos de avaliação de função igual a duas vezes a quantidade de termos da expansão. Em outras palavras isto quer dizer que, com o emprego da regressão linear no caso de três variáveis incertas e ordens de expansão $p = 2$ ou $p = 3$, recomenda-se apenas 20 ou 40 pontos de avaliações de função, respectivamente, para obtenção de resultados satisfatórios. De acordo com os experimentos realizados observou-se que foram obtidos resultados de qualidade equivalente, de maneira mais econômica, ou seja, com quantidade de pontos de avaliação inferior ao recomendado. No entanto não é possível afirmar que isto se aplique a todos os casos.

- Para os dois métodos estudados, os valores obtidos para a média do VPL foram equivalentes em todas as situações criadas em termos de quantidade de variáveis incertas e de variações feitas em relação à ordem de expansão dos polinômios, a forma utilizada para o cálculo dos coeficientes da expansão (por quadratura/cubatura ou regressão linear) e a ordem de quadratura e ordem de interpolação. No entanto, a rápida convergência para a média e desvio padrão do valor do VPL só ocorreu para poucos dígitos significativos, ainda com maiores dificuldades na determinação do desvio padrão do VPL, fato que pode ser devido ao ruído numérico na simulação. Dessa forma, não havendo necessidade de maior precisão, as Expansões do método do Caos Polinomial e as da Colocação Estocástica poderão ser utilizadas como aproximação da função não linear, no caso o simulador de reservatórios. A depender da necessidade de maior precisão, mais investigações devem ser feitas com o incremento das ordens de quadratura/cubatura e/ou da ordem da expansão dos polinômios, visando o alcance da convergência para maior quantidade de dígitos significativos.

- Deve ser ressaltado também que, para que as técnicas de propagação de incertezas pudessem ser avaliadas dentro de condições simplificadas e de fácil controle neste trabalho, foi estudado um caso de reservatório com características bem simples significando com isso que, em condições mais complexas, o desempenho desses métodos pode não ser o mesmo.
- Em relação ao custo computacional e ao tempo de processamento, para as condições de simplicidade estabelecidas no estudo, ou seja, com apenas poucas variáveis incertas, os dois métodos de propagação de incertezas apresentaram desempenho bastante satisfatório. No entanto, para uma maior quantidade de variáveis incertas, como é o caso de aplicações mais realistas, confirma-se a necessidade do emprego de técnicas para a redução do espaço probabilístico, a exemplo da Expansão *Kernel* de Karhunen-Loève (KLE).
- Para driblar os limites impostos pelo equipamento computacional disponível, principalmente em problemas realistas, onde a quantidade de variáveis incertas é muito grande, confirma-se a necessidade de uso de *clusters* de servidores para a computação paralela, visando não apenas a redução do esforço computacional como o tempo de processamento.

9.2 Sugestões para Trabalhos Futuros

Sabendo-se que as técnicas aqui demonstradas ainda são relativamente novas, há ainda muito a investigar sobre elas, sendo necessário ainda muito esforço em pesquisas para a obtenção de todas as suas vantagens e conhecimento de suas limitações. Dessa forma, sugere-se como perspectiva de estudos futuros para consolidação das técnicas empregadas aqui:

- a) Aplicação das técnicas demonstradas em problemas mais complexos e de maior porte.

- b) Experimentos com o método do Caos Polinomiais e o da Colocação Estocástica utilizando-se funções não lineares analíticas.
- c) Investigação de técnicas para a geração de realizações equiprováveis a partir de uma realização dada e não partindo de funções de correlação pré-definidas, como foi feito até aqui.
- d) Verificação do número de autovalores a ser mantido nas decomposições de Karhunen-Loève aplicadas a problemas reais, no intuito de verificar se as técnicas aqui previstas podem ser empregadas.
- e) Implementação da decomposição de Karhunen-Loève não linear, de forma a permitir estatísticas de ordem superior nas expansões, o que irá permitir o tratamento de características geológicas como fácies, falhas e canais de alta permeabilidade.

REFERÊNCIAS

- [1] ADAMSON, GORDON ; CRICK, MARTIN ; GANE, BRIAN ; GURPINAR, OMER ; HARDIMAN, JIM; PONTING, DAVE. *Simulation Throughout the life of a reservoir*. 1996. Disponível em: <<http://www.slb.com/media/services/resources/oilfieldreview/ors96/sum96/06961627.pdf>>. Acesso em 12 mar. 2010.
- [2] ASKEY, R.; WILSON, J. A. *Some basic Hypergeometric polynomials that generalize Jacobi polynomials*. *Memoirs Amer. Math. Soc.*, AMS Providence RI, 319, 1985.
- [3] AYLOR, W.K. *Measuring the impact of 3D seismic on business performance*. SPE, HEEC Hydrocarbon Economics and Conference, Dallas, 21-23 March, 1999, 12p.
- [4] CAMERON, R.H.; MARTIN, W.T. *The orthogonal development of nonlinear functionals in series of Fourier-Hermite functionals*. *Ann. Math.*, 48:385 {392, 1947.
- [5] CHIHARA, T. S. *An Introduction to Orthogonal Polynomials*. Gordon and Breach, Science Publishers, Inc., 1978.
- [6] CHORIN, A. J. *Hermite expansion in Monte-Carlo simulations*. *J. Comput. Phys.*, Science Publishers, Inc., 1978.
- [7] COMPUTER MODELLING GROUP Ltd. *IMEX Advanced Oil/Gas Reservoir Simulator Version 2000 User's Guide*. Calgary, Alberta, Canada, 746pp. 2000.
- [8] CORREIA NETO, J. F.; MOURA, H. J.; FORTE, S. H.C. A. *Modelo Prático de Previsão de Fluxo de Caixa Operacional para Empresas Comerciais Considerando os Efeitos do Risco, Através do Método de Monte Carlo*. *Revista Eletrônica de Administração*. Edição 27, n. 3, v. 8, 2002.
- [9] CULLICK, A.S.; HEATH, D.; NARAYANAN, K.; KELLY, J; LANDMARK GRAPHICS CORPORATION; JAY APRIL; OPTTEK SYSTEMS SOURCE. *Optimizing Multiple-Field Scheduling and Production Strategy with Reduced Risk*. SPE Annual Technical Conference and Exhibition, Denver, Colorado, 5-8 October, 2003, 12p.
- [10] DEODATIS, G. *Weighted integral method. I: stochastic stiffness matrix*. *J. Eng. Mech.*, 117(8):1851-1864, 1991.
- [11] DOUGLAS Jr, J.; ARBOGAST, T. *Dual-porosity models for flow in naturally fractured reservoirs "Dynamics of fluids in hierarchical porous formations"*, J. Cushman, ed., Academic Press, London, 1990.

- [12] ELDRED, M. S.; WEBSTER, C. G.; CONSTATINE, P. G. *Evaluation of Non-Intrusive Approaches for Wiener-Askey Generalized Polynomial Chaos*, AIAA Paper 2008-1892.
- [13] EATON, J.W. *Octave*. Versão 3.2.2. 2009. Disponível em: < www.Octave.org>. Acesso em 30 jul. 2010.
- [14] FAVARELLI, L. *Response surface approach for reability analysis*. J. eng. Mech, 115:2763-2781, 1989.
- [15] FISHMAN, G.S. *Monte Carlo: Concepts, Algorithms, and Applications*. Springer Verlag New York, Inc., 1996.
- [16] FOX, B.L. *Strategies for Quasi-Monte Carlo*. Kluwer Academic Pub., 1999.
- [17] GAMERMAN, D. *Markov Chain Monte Carlo: stochastic simulation for Bayesian inference*. Chapman & Hall, 1997.
- [18] GHANEM, ROGER. *Probabilistic Characterization of transport in Porous Media*, Comput. Methods. Appl. Mech. Engrg., 158:199-220, 1998.
- [19] GHANEM, ROGER; RED-HORSE, J. R. *Propagation of probabilistic uncertainty in complex physical systems using a stochastic finite element technique*. Physica D, 133:137–144, 1999. 55, 81, 100.
- [20] GHANEM, R.G.; SPANOS P.D. *Stochastic Finite Elements: a Spectral Approach*. Springer-Verlag, 1991.
- [21] GILKS, W.R.; RICHARDSON S.; SPIEGELHALTER D.J. *Markov chain Monte Carlo in practice*. Chapman & Hall, 1995.
- [22] HALDAR, A.; MAHADEVAN, S. *Probability, Reliability, and Statistical Methods in Engineering Design*. Wiley, New York, 2000. 17, 55, 89, 90, 167, 285.
- [23] HAMMERSELEY, D.C; HANDSCOMB, J. M. *Monte Carlo Methods*. Methuen, 1964.
- [24] HOSDER, S.; WALTERS, R.W.; BALCH, M. *Efficient sampling for non-intrusive polynomial chaos applications with multiple uncertain input variables*. In Proceedings of the 48th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference, number AIAA-2007-1939, Honolulu, HI, April 23–26, 2007. 103.
- [25] HUANG, SHUPING; MAHADEVAN, SANKARAN; REBBA, RAMESH. *Collocation-based Stochastic Finit Element Analysis for Randon Field Problems*. Probabilistic Engineering Mechanics 22:194-205, 2007.
- [26] KEVORKIAN, J.; COLE, J.D. *Perturbation Methods in Applied Mathematics*. Springer Verlag (1981).

- [27] KIUREGHIAN, A. Der; LIU, P. L.OEK, R.; SWARTTOUW, R.F. *Structural reliability under incomplete information*. J. Eng. Mech., ASCE, Vol. 112 , No. 1, pp: 85-104, 1986.
- [28] KOEKOEK, R.; SWARTTOUW, R.F. *The Askey-scheme of hypergeometric orthogonal polynomials and its q-analogue*. Technical Report 98-17, Department of Technical Mathematics and Informatics, Delft University of Technology, 1998.
- [29] LIGERO, ELIANA L. *Quantificação de Incertezas e Análise de Risco em Projetos de Exploração e Produção de Petróleo*. Disponível em: <[http://www.dep.fem.unicamp.br/boletim/BE29/artigo_elianna.htm](http://www.dep.fem.unicamp.br/boletim/BE29/artigo_eliانا.htm)>. Acesso em 24 jul. 2010.
- [30] LOH, W.L. *On Latin hypercube sampling*. *Ann Stat.* 24(5):2058-2080, 1996.
- [31] MATHELIN L.; HUSSAINI, M. Y.; ZANG, T.A. *Stochastic approaches to uncertainty quantification in CFD simulations*. Numer. Algo., submitted, 2003.
- [32] NAYFEH, A.H. *Perturbation Methods*. John Willey .1973.
- [33] NIEDERREITER H.; HELLEKALEK, P.; LARCHER, G.; ZIENTERHOF, P. *Monte Carlo and Quasi-Monte Carlo methods*. 1996. Springer-Verlag, 1998.
- [34] NOBILE, F.; TEMPONE, R.; WEBSTER, C. G. *A sparse grid stochastic collocation method for partial differential equations with random input data*. Technical Report Technical report TRITA-NA 2007:7, Royal Institute of Technology, Stockholm, Sweden, 2007.
- [35] OBERKAMPF , W. .L.; HELTON, J. C. *Evidence theory for engineering applications*. Technical Report SAND2003-3559P, Sandia National Laboratories, Albuquerque, NM, 2003.
- [36] OGURA, H. *Orthogonal functionals of the Poisson process*. IEEE Trans. Info. Theory, IT-18:473{481, 1972.
- [37] OLIVEIRA, DIEGO. F. B. *Técnicas de Otimização da produção para Reservatórios de Petróleo, 2006, 137 f.* Dissertação (Mestrado em Engenharia Civil), Departamento de Engenharia Civil, UFPE, Recife.
- [38] PEIXOTO, LOURENÇO L. *Quadratura de Gauss iterativa com base nos polinômios ortogonais clássicos*. Dissertação de mestrado, Centro Federal de Educação tecnológica de Minas, 2008, Belo Horizonte, MG, 142p.
- [39] RAJASCHEKHAR, M.R.; ELLINGWOOD, B.R. *A new look at the response surface approach for reability analysis*. *Struc. Safety*, 123:205-220, 1993.
- [40] REAGAN, M. T. REAGAN; NAJM, H. N.; PEBAY, P. P.; KNIO, O. M.; GHANEM R. G. *Quantifying uncertainty in chemical systems modeling*. *Int. J. Chem. Kinet.*, 37:368–382, 2005.

- [41] RUBINSTEIN, R.Y. *Simulation and the Monte Carlo method*. John Wiley & Sons, New York, 1981.
- [42] SAAD Y.; SCHULTZ M. H. *GMRES: Generalized Minimum Residual Algorithm for Solving Nonsymmetric Linear Systems*, SIAM Journal of Scientific and Statistical Computing, 7, pp. 856-869, 1986.
- [43] SANDIA NACIONAL LABORATORIES. *DAKOTA, A Multilevel Parallel Object-Oriented Framework for Design Optimization, Parameter Estimation, Uncertainty Quantification, and Sensitivity Analysis: Version 4.2 User's Manual*. California, 2008. 330 p. Disponível em: <<http://www.cs.sandia.gov/DAKOTA/software.html>>. Acesso em 12 mar 2010.
- [44] SARMA,PALLAV; DURLOFSKY, LOUIS J.; AZIZ, KHALID. *Kernel Principal Component Analysis for Efficient, Differentiable Parameterization of Multipoint Geostatistics*, Math. Geosci, 40:3-32, 2008.
- [45] SARMA,PALLAV; DURLOFSKY, LOUIS J.; AZIZ, KHALID; CHEN, WEN H. *A New Approach to Automatic History Matching Using Kernel PCA*, SPE Reservoir Simulation Symposium, 26-28 February 2007, Houston, Texas, U.S.A. Paper 106176, 2007.
- [46] SARMA,PALLAV; DURLOFSKY, LOUIS J.; AZIZ, KHALID; CHEN, WEN H. *Efficient Real-Time Reservoir Management Using Adjoint-based Optimal Control and Model Updating*, Computacional Geosciences, 10:3-36,2006.
- [47] SCHOKKOPF, BERNHARD; SMOLA, ALEXANDER; MULLER, KLAUS-ROBER. *Nonlinear Componente Analysis as a Kernel Eigenvalue Problem*, Technical Report No. 44, Max-Planck Institut fur biologische Kybernetik, Arbeitsgruppe Bulthoff, December 1996.
- [48] SCHOUTENS, W. *Stochastic Processes and ortogonal polynomials*. Springer-Verlag, New York, 2000.
- [49] SCHOUTENS, W. *Stochastic processes in the Askey scheme*. PhD thesis, K.U. Leuven, 1999.
- [50] SHAFER, GLENN. *A Mathematical Theory of Evidence*. Princeton University Press, 1976, ISBN 0-608-02508-9
- [51] STEIN, M. *Large sample properties of simulations using Latin Hypercube Sampling*. Technometrics, 29(2):143-151, 1987.
- [52] SUSLICK,S. B; FURTADO, R. *Quantifying the Value of Technological, Environmental, and Financial Gain in Decision Models for Offshore Oil Exploration*, Journ.Petrol Sci. Eng., 2001.
- [53] SUSLICK, SAUL B. *Conhecer as incertezas: o desafio da indústria do petróleo*. Com Ciência, São Paulo, 2002. Disponível em:

<<http://www.comciencia.br/reportagens/petroleo/pet20.shtml>>. Acesso em 12 mar. 2010.

- [54] SWILER, L. P.; WYSS, G. D. *A user's guide to Sandia's latin hypercube sampling software: LHS UNIX library and standalone version*. Technical Report SAND04-2439, Sandia National Laboratories, Albuquerque, NM, July 2004. 71, 82, 167.
- [55] TAKADA, T. *Weighted integral method in stochastic finite element analysis*. Prob. Eng. Mech., 5(3):146-156, 1990.
- [56] THE MATHWORKS, INC. *MATLAB: Version 7.1.0*. Disponível em: <<http://www.mathworks.com/products/matlab/>>. Acesso em 30 jul. 2010.
- [57] THOMAS, JOSÉ EDUARDO. *Fundamentos de Engenharia de Petróleo*. 2. ed. São Paulo, Interciência, 2004.
- [58] WALTERS, R. W. *Towards stochastic fluid mechanics via polynomial chaos*. In Proceedings of the 41st AIAA Aerospace Sciences Meeting and Exhibit, number AIAA-2003-0413, Reno, NV, January 6–9, 2003. 103.
- [59] WIENER, N. *The homogeneous chaos*. Amer. J. Math., 60:897{936, 1938.
- [60] WIKIPEDIA, a enciclopédia livre. *Método de Monte Carlo*. [S.l.]: [s.n.], 2010. Disponível em: <http://pt.wikipedia.org/wiki/M%C3%A9todo_de_Monte_Carlo>. Acesso em: 15 jun. 2010.
- [61] WILLMERSDORF, RAMIRO. B. *Relatório do Subprojeto SP-2, Novembro de 2008*.
- [62] XIU, D.; KARNIADAKIS, G. E. *The Wiener-Askey polynomial chaos for stochastic differential equations*. SIAM J. Sci. Comput., 24(2):619{644, 2002.
- [63] YAMAZAKI, F.; SHINOZUKA, M.; DASGUPTA, G. *Neumann expansion for stochastic finite element analysis*. J. Eng. Mech., 114(8):1335-1354, 1988.